Konvolüsyonel Yapay Sinir Ağları ile Diş Tespiti

Tooth Detection with Convolutional Neural Networks

Ayse Betul Oktay
Department of Computer Engineering,
Faculty of Engineering and Natural Sciences,
Istanbul Medeniyet University,
Istanbul, Turkey
abetul.oktay@medeniyet.edu.tr

***Özetçe*—Dişlerin panoramik X-ray görüntülerinden otomatik tespit edilmesi bilgisayar destekli diş uygulamaları için çok önemli bir adımdır. Bu çalışmada, Konvolüsyonel Sinir Ağları ile panoramik diş görüntülerindeki dişleri tespit etmeye yönelik bir yöntem sunulmuştur. İlk olarak ağız boşluğunu bulduktan sonra, üç diş tipi için (kesici dişler, küçük azılar ve büyük azılar) uygun pozisyonlar belirlenir. Dişler, çoklu sınıflandırmanın yapıldığı AlexNet mimarisinin değiştirilmiş bir versiyonu ile tespit edilir. Yapılan testler, yöntemin doğruluğunun ümit verici olduğunu ve sunulan yöntemin bilgisayar destekli uygulamaların ilk adımı olarak kullanılabileceğini göstermiştir.**

Anahtar Kelimeler*—konvolüsyonel sinir ağları; derin öğrenme; diş görüntüsü.

***Abstract*—Detection of teeth from dental panoramic X-ray images is the a crucial step for the computerized tooth applications. In this a paper, we present a method for detecting teeth in dental panoramic X-ray images with Convolutional Neural Networks (CNN). After finding the mouth gap, possible positions of three tooth type (incisors, premolars, and molars) are determined. Teeth are detected with a modified version of AlexNet architecture where multi-class classification is performed. The accuracy of the method is over 90% and the presented method can be used as the first step of the computerized dental applications.**

Keywords*—convolutional neural networks; deep learning; dental image.

I. INTRODUCTION

Dental biometric information is used for human identification like fingerprint and DNA. Teeth are not generally damaged after disasters because of their hard and durable structures. In addition, human identification from teeth is more rapid and effective than fingerprints and DNA. During the tsunami at Thailand in 2004, nearly 46.2% of human identification was performed with dental images in a short time period [1]. In addition, for computer aided detection of tooth diseases and dental operations each tooth should first be localized in dental images. Thus, detecting teeth in dental images is a crucial task.

There are two types of dental X-ray images which are intra-oral and extra-oral. The intra-oral images are taken by placing the image receptor inside the patient's mouth and they contain a few teeth. The extra-oral images are taken while the X-ray source rotates around the head and they show the entire mouth, including all the teeth in the upper and lower arch, in one image. Panoramic X-rays are used to plan treatment for dental implants, detect impacted wisdom teeth and jaw problems, and diagnose tumors and cysts. Fig. 1 shows a panoramic dental X-ray image.

In recent years, deep learning has been widely used at many applications for computer vision and speech recognition and it has improved the state-of-the-art in many other domains [11]. Convolutional Neural Networks (CNN) with deep architectures produce state-of-the-art results especially for detection and classification tasks for computer vision.

In this paper, we present a tooth detection method with CNN for panoramic dental images (Figure 1). First, the mouth gap is determined with the pre-processing step. Determination of the mouth gap provides the possible placement of the teeth. A modified version of AlexNet architecture is used for detecting each tooth at the detected area. The method is tested on a dataset containing 100 dental panoramic images and the results are promising.

978-1-5386-0633-9/17/$31.00 ©2017 IEEE

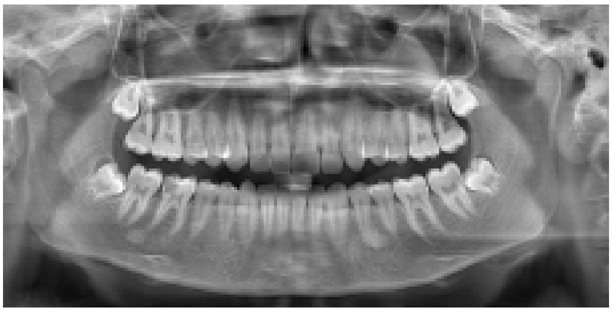


Fig. 1.A dental panoramic image

II. LITERATURE REVIEW

In the literature, many methods are presented for isolating and segmenting the teeth. Many of them are proposed as the first step of human identification methods and they use low-level image processing techniques. Automated Dental Identification System [2] is a project that includes many studies for human identification with dental intra-oral images. Jain ve Chen [3] proposed a semi-automatic method with two steps which are feature extraction and matching. They first isolated the teeth and found tooth contours with an intensity based probabilistic method and compared the contours for matching. Zhou ve Abdel-Mottaleb [4] presented a method for classifying the dental images and they segmented the tooth contours with snake method. Nomir and Abdel-Mottaleb [5] applied iterative and adaptive thresholding for enhancing the tooth contours. Then the final contours were extracted with integral projection.

Pattanachai et. al [6] marked the teeth manually and found the tooth borders with Otsu thresholding method. Then they used Hu's moment invariants for matching the tooth contours. In [7], a method with histogram projection and adaptive thresholding is used for classification and numbering of teeth for dental bitewing radiographs. In [8], a tooth isolation method for dental X-ray images, which contains upper-lower jaw separation, single tooth isolation, over-segmentation verification, and under-segmentation detection is presented. Guzel et al. [9] used Haar filters for filter extraction and classified the teeth with SVM and labeled with an atlas based model.

Miki et al. [12] employed CNN for classifying teeth into seven classes in cone-beam CT images and the performance of the system is 88.8%. Raith et al. [13] used neural networks for classifying the teeth at 3D range images and their method has 93% accuracy.

III. PRE-PROCESSING: MOUTH GAP DETECTION

Panoramic dental images contain entire mouth with jaws and sinuses besides teeth. Determining the mouth gap and potential placement of teeth brings the advantage of decreasing the search space. Thus, we first determine the mouth gap.

Let I be an image that has width of w and height of h. The center of the mouth gap exists between the 40%-60% of the image height. So, the integral projection of the pixels between $w/2-100$ and $w/2+100$ at width and pixels from h*0.4 to h*0.6 at height is calculated. The minimum value of the integral projection is determined as the center of the mouth gap. In order to determine the whole mouth gap, a method similar to [3],[10] is used. Let v_i be a local minimum of the intensity projection of a patch around the mouth gap center, y_i is location of vi and D_i is intensity value of the projection at y_i. $p_{vi}(D_i)$ is the likelihood of local minima having the pixel intensity D_i and $p_{vi}(y_i)$ is a Gaussian with expected value equal to the position of the last detected line at the former iteration. The whole mouth gap is found iteratively as follows:

$$P(v_i,D_i)=(1-P_{vi}(D_i))P_{vi}(y_i) \qquad (1)$$

At each iteration a patch 20 pixels from left and right are chosen and the whole mouth gap line is determined for each patch.

The mouth gap gives the potential positions of the three tooth types which are molars, premolars, and canine&incisors. The sizes of the areas for these three tooth types are determined by taking the maximum tooth areas at the training images. Fig. 2 shows the mouth gap and tooth areas for a panoramic image.

Another important process is the affine transformation of the potential teeth areas. Each tooth type is found at four quarter of the mouth which are upper left, upper right, lower left, and lower right. The quarters except the lower right are rotated and mirrored in order to get a tooth appearance like lower right. This process makes all quarters as they are all lower right quarter of mouth and increases the number of examples in the training set. Since CNN need a lot of samples for training, the transformation process provides four times more samples for training.

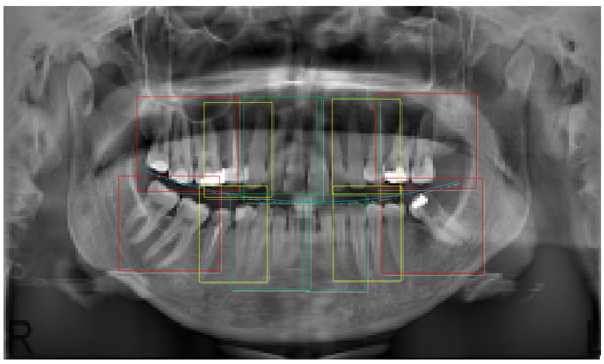


Fig. 2. The detected mouth gap and the potantial tooth areas.

IV. TOOTH DETECTION

Deep learning enables learning on large data by creating very large neural networks with low cost and high computing power. The representation of the data is learnt hierarchically with many layers and there is no need of extracting features with hand-crafted methods. Convolutional neural network is one of the deep learning methods that can calculate millions of parameters and learn the image representations effectively. The accuracy of the CNN is much higher than hand crafted methods like local binary patterns or histogram of oriented gradients especially in image classification and object detection. Therefore, we use one of the popular CNN architectures –AlexNet [11]- for tooth detection.

After finding the possible tooth locations and affine transformation, each mouth quarter is fed into the CNN. The original AlexNet takes images that have size 227 by 227 pixels as input and they use patches of 11x11 pixels. However, the size of the image of the mouth quarter is much bigger than 227x227 pixels. The average size of the incisor tooth in the images of the training set is 180x80 pixels and the average of the molar tooth images is 212x200 pixels. Using patches at these sizes has a very high computational cost. So, we modified the AlexNet architecture by using the sliding window technique. The image quarters are resized to half and patches are sliding windows with stride of 4 pixels. The original AlexNet architecture has output of 1000 classes however our method has 4 outputs/classes which are incisor, premolar, molar, and background. The size of the sliding window changes between the minimum and maximum tooth size.

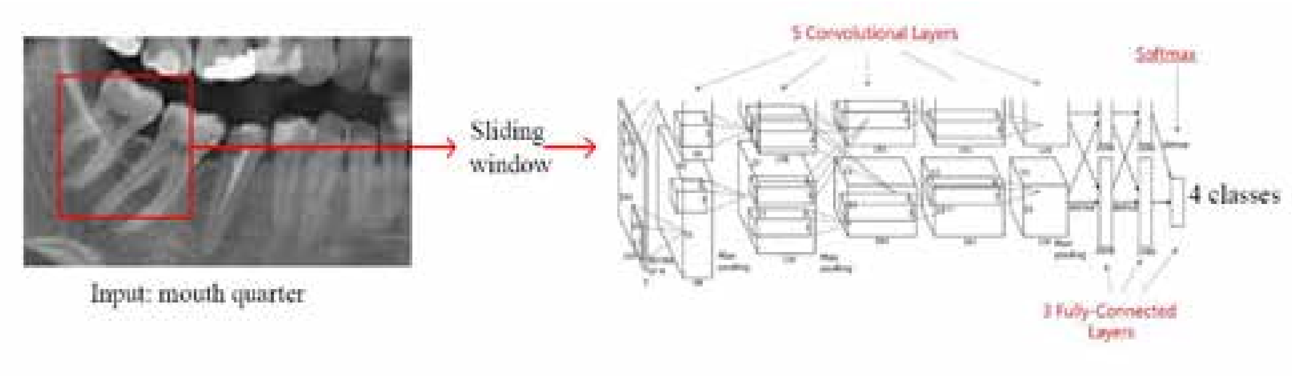


Fig. 3. AlexNet framework [11]

V. EXPERIMENTS

The system is tested and validated on a dataset containing dental panoramic images of 100 different people. The images are taken with 3 different X-ray machines and the images have sizes of 2871x1577, 1435x791 or 2612x1244 pixels. Each tooth in the images are delineated by a dentist and these delineations are used for training and testing the system.

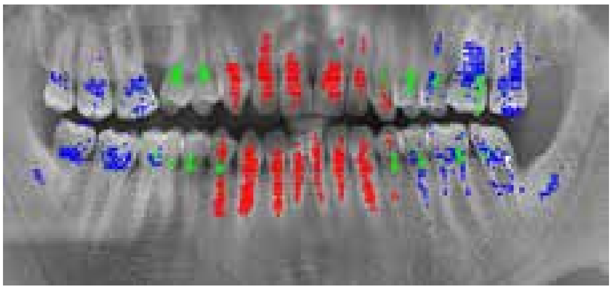
At first the mouth gap is determined automatically and the mouth quarters are transformed. For evaluation of the system 10-fold cross validation is used. For the background samples, the image patches that do not contain any tooth is selected randomly.

The accuracy of the system is calculated according to Dice coefficient and a detected tooth is considered as correct if the intersection of the detected tooth window and the ground truth window over their union is greater than 0.5. The accuracy of the molar teeth is 94.32%, premolar is 91.74% and canine and incisors is 92.47%. The accuracy of the premolars are lower than other tooth types because neighboring canine and teeth premolar teeth are very similar to premolars and they are misclassified as premolars.

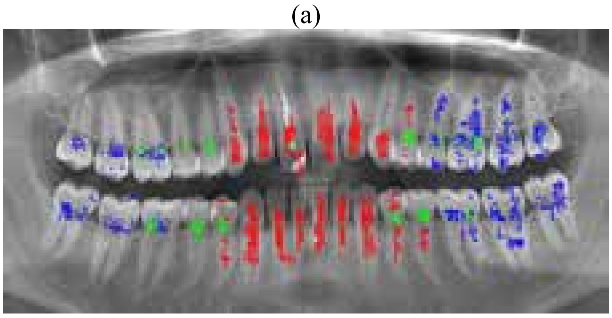
TABLE 1. THE CLASSIFICATION ACCURACY OF THE PROPOSED SYSTEM

Tooth Class	Accuracy
Molar	94.32%
Premolar	91.74%
Canine & incisor	92.47%

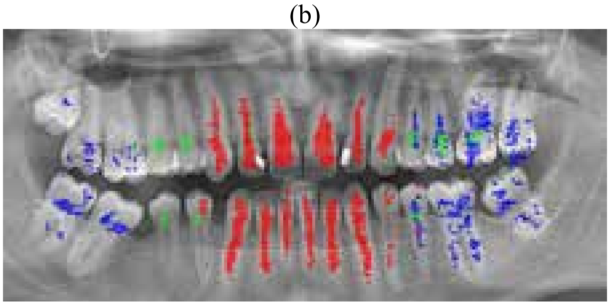
The centers of the detected windows are shown in Fig. 4. The incisors and canines are shown with red, premolars are shown with green and molars are shown with blue. The detected tooth centers are seem centered at the tooth root. Some mis-detected tooth positions are present especially at neighboring teeth and tooth gums for the third molar tooth.



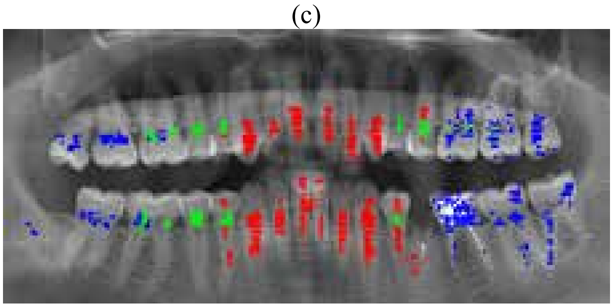
(a)



(b)



(c)



(d)

Fig. 4. The centers of the detected tooth positions. Incisors and canine teeth are marked with red, promolar aremarked with green and molars are marked with blue.

VI. CONCLUSIONS

In this study a new method is designed for detecting teeth in panoramic dental X-ray images with AlexNet architecture. The preprocessing step detected the possible areas at image for each tooth type after detecting the mouth gap. The AlexNet architecture is modified for taking the inputs as sliding windows and the output of the CNN is 4 classes. The system is tested on a dataset containing 100 panoramic dental images and the accuracy of the system is over 90%. This study shows that CNN architecture can be effectively used for detection of the teeth.

REFERENCES

[1] International Federation of Red Cross and Red Cresdent Societies. 2005."world disaster report 2005", London: Eurospan.http://www.ifrc.org/Global/Publications/disasters/WDR/69001-WDR2005-english-LR.pdf.
Access Date:21.01.2017

[2] G. Fahmy et al., "Towards an automated dental identification system (ADIS)," Biometric Authentication'04, pp. 789–796, 2004

[3] A. K. Jain., H. Chen, "Matching of dental X-ray images for human identification", Pattern Recognition, 2004, 37, (7), pp. 1519 – 1532

[4] J. Zhou, M. Abdel-Mottaleb, "A content-based system for human identification based on bitewing dental x-ray images", Pattern Recognition, 2005, 38, (11), pp. 2132 – 2142

[5] O. Nomir., M. Abdel.Mottaleb, "A system for human identification from x-ray dental radiographs", Pattern Recognition, 2005, 38, (8), pp. 1295–1305

[6] N. Pattanachai, N. Covavisaruch, C. Sinthanayothin, "Tooth recognition in dental radiographs via Hu's moment invariants", International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology. 2012. pp. 1–4

[7] P.L. Lin, Y. H. Lai, P. W. Huang, "An effective classification and numbering system for dental bitewing radiographs using teeth region and contour information", Pattern Recognition, 2010, 43, (4), pp. 1380 – 1392

[8] P.L. Lin, P.W. Huang, Y.S. Cho, C.H. Kuo, "An automatic and effective tooth isolation method for dental radiographs", Opto-Electronics Review, 2013, 21, pp. 126–136

[9] S. Guzel., A.B. Oktay, K. Tufan, "Automatic tooth identification in dental panoramic images with atlas-based models". Int. Conf. on Pattern Recognition Applications and Methods, Lizbon, Portugal, January 2015, pp.136-141

[10] D. Frejlichowski, P. Czapiewski P., "An Application of the Curvature Scale Space Shape Descriptor for Forensic Human Identification Based on Orthopantomograms", in K. Saeed et al. (eds.): 'Lecture Notes in Computer Science' (Springer, Berlin, Heidelberg, 2011), 8104, pp. 67-76

[11] A. Krizhevsky, I. Sutskever, G.E. Hinton, "ımagenet Classification with Deep Convolutional Neural Networks", NIPS 2012: Neural Information Processing Systems, Lake Tahoe, Nevada

[12] Y. Miki, et al., "Classification of teeth in cone-beam CT using deep convolutional neural network", Computers in Biology and Medicine, Volume 80, 2017, Pages 24-29.

[13] S. Raith et al., "Artificial Neural Networks as a powerful numerical tool to classify specific features of a tooth based on 3D scan data", Computers in Biology and Medicine, Volume 80, 2017, Pages 65-76.