RUSSIAN ELECTRONIC JOURNAL OF RADIOLOGY

ОРИГИНАЛЬНАЯ СТАТЬЯ

DETECTION AND NUMBERING OF TEETH IN PANORAMIC RADIOGRAPHIC IMAGES USING DEEP NEURAL NETWORKS

Neda Alizad [1], Masoumeh Johari [1], Hadi Abbasi [2]

1 - Department of Oral and Maxillofacial Radiology, Faculty of Dentistry, Tabriz University of Medical Sciences.
2 - Faculty of Electrical and Computer Engineering, Tabriz University. Tabriz, Iran.

Today, with the advancement of the artificial intelligence methods, it is possible to automatically evaluate these images in order to save the clinician's time.

Purpose. To employ Convolutional Neural Networks (CNNs) for tooth segmentation and numbering in panoramic radiographic images. The study utilized a dataset with ample volume and diversity, and employed cutting-edge deep learning algorithms for the task of tooth segmentation and numbering. Implementing and utilizing this method can enhance the efficiency of clinical diagnosis and treatment procedures.

Materials and Methods. The data set includes 527 panoramic images that were selected from the archives of the Radiology Department of the Faculty of Dentistry of Tabriz. After that, the images were labeled by an oral and maxillofacial radiologist, according to the FDI numbering system. The segmentation was done by using the U-Net architecture and its output entered the VGG-16 network for numbering. Eighty percent of the data was used for network training and 10% for validation and another 10% for network testing.

Results. The results obtained from the U-Net network for tooth segmentation, based on the original data; sensitivity, specificity, and Dice, are 98.9%, 98.4%, and 95.4%, respectively. Also, for teeth numbering by using the VGG-16 Network Architecture, we obtained sensitivity, specificity and accuracy equal to 98.58%, 99.93% and 96.8%, respectively. In the examination and diagnosis of the implant, the retained roots, and the extracted teeth, the accuracy of 98.45%, 97.1%, and 98.2% was obtained, respectively.

Conclusion. The obtained results are favorable compared to similar studies, and in the future, with the development of these methods, it can be a useful help in the automatic analysis of panoramic images and other dental images.

Keywords: Panoramic Radiography; Teeth Detection; Teeth Numbering; Convolutional Neural Networks.

Corresponding author: Masoumeh Johari, e-mail: Khatoonm87@yahoo.com

For citation: Neda Alizad, Masoumeh Johari, Hadi Abbasi. Detection and numbering of teeth in panoramic radiographic images using deep neural networks. REJR 2024; 14(1):42-54. DOI: 10.21569/2222-7415-2024-14-1-42-54.

Received: **07.12.23** **Accepted:** **17.12.23**

RUSSIAN ELECTRONIC JOURNAL OF RADIOLOGY

АНАЛИЗ И НУМЕРАЦИЯ ЗУБОВ НА ОРТОПАНТОМОГРАММАХ С ИСПОЛЬЗОВАНИЕМ ГЛУБОКИХ НЕЙРОННЫХ СЕТЕЙ

Неда Ализад[1], Масуме Джохари[1], Хади Аббаси[2]

1 - Отделение лучевой диагностики челюстно-лицевой области, стоматологический факультет Тебризского университета медицинских наук.

2 - Факультет электротехники и вычислительной техники, Тебризский университет. Тебриз, Иран.

Сегодня, с развитием методов искусственного интеллекта, появилась возможность автоматически оценивать изображения с целью экономии времени.

Цель исследования. Использовать сверточные нейронные сети для сегментации и нумерации зубов на ортопантомограммах. В исследовании использовался набор данных достаточного объема и разнообразия, а также передовые алгоритмы глубокого обучения для сегментации и маркировки зубов. Внедрение и использование этого метода может повысить эффективность диагностики и лечебных процедур.

Материалы и методы. Исследование включало 527 ортопантомограмм, которые были отобраны из архивов отделения лучевой диагностики стоматологического факультета Тебриза. После этого рентгенолог, работающий в челюстно-лицевой области, маркировал снимки в соответствии с системой нумерации FDI. Сегментация была выполнена с использованием архитектуры U-Net и ее выходные данные поступали в сеть VGG-16 для нумерации. Восемьдесят процентов данных было использовано для обучения сети, 10% – для валидации и еще 10% – для тестирования сети.

Результаты. Результаты, полученные с помощью сети U-Net для сегментации зубов, основаны на исходных данных; чувствительность, специфичность и Dice составила 98,9%, 98,4% и 95,4% соответственно. Кроме того, для маркировки зубов с использованием сетевой архитектуры VGG-16 была получена чувствительность, специфичность и точность, равные 98,58%, 99,93% и 96,8% соответственно. При обследовании и диагностике имплантатов, сохраненных корней и удаленных зубов была получена точность 98,45%, 97,1% и 98,2% соответственно.

Вывод. Было проведено сравнение полученных результатов с аналогичными исследованиями, в будущем, с развитием этих методов, это может стать полезным подспорьем при автоматическом анализе ортопантомограмм и других изображений зубо-челюстной системы.

Ключевые слова: ортопантомограмма, сегментация зубов, маркировка зубов, сверточные нейронные сети.

Контактный автор: Масуме Джохари, e-mail: Khatoonm87@yahoo.com

Для цитирования: Неда Ализад, Масуме Джохари, Хади Аббаси. Анализ и нумерация зубов на ортопантомограммах с использованием глубоких нейронных сетей. REJR 2024; 14(1):42-54. DOI: 10.21569/2222-7415-2024-14-1-42-54.

Статья получена: **07.12.23** **Статья принята:** **17.12.23**

In recent decades, medical imaging methods such as Computed tomography have been used to discover, diagnose and treat various diseases. The new and emerging field in dentistry is informatics dentistry, which in addition to improving the treatment and diagnosis plan, also saves time and reduces stress and fatigue caused by daily practices. These applications in medicine, and especially in dentistry, generate large amounts of data from sources such as high-resolution medical images, medical sensors, and electronic medical records. The use of computer programs can assist the dental professionals in decisions related to prevention, diagnosis or treatment planning [1].

Artificial intelligence is defined as the ability of a machine to imitate intelligent behavior to perform complex tasks such as decision-making, recognition of words and objects, and problem solving. Currently, one of the methods of artificial intelligence used in the therapeutic fields is called Deep Learning. Deep learning methods have been developed to improve the performance of traditional Artificial Neural Networks in which the network architecture becomes more complex. The success of the deep learning is mainly due to the advancement of the capacity of computers, the huge amount of the available data and the development of the algorithms. This method has been used effectively in image-based diagnoses in different fields and its application has been confirmed. Deep learning methods are characterized by different levels of representations and raw data are processed to perform the classification or diagnostic tasks. In deep learning, several layers of the algorithms are hierarchically placed on top of each other and produce meaningful data. These layers receive the input data and the resulting output changes gradually by learning new features based on the provided data. Artificial neural networks consist of thousands to millions of nodes or connected units. The communication between the nodes is activated like the human nervous system and this activity spreads from one unit to another, by which a numerical value that is the weight of the connection, the strength of the desired connection is determined. A type of deep learning network called CNNs is usually used in applications that rely on the deep learning, which has been greatly developed during the last decade and is mainly used for the analysis of the medical images, which is more accurate and modern than the other methods [2, 3].

Dental Panoramic Radiography was developed by Pattero in 1945 and has been widely used for the dental diagnostic purposes since then [4]. Panoramic Radiography is a radiological technique that provides an overview of the jaws and surrounding structures. This method is often used when specialists want to evaluate some structures such as unerupted third molar teeth, orthodontic treatment, tooth growth, developmental abnormalities, trauma, large lesions, etc. [5-7]. Panoramic Radiography allows the dentist to see a large area of the upper and lower jaw on a film.6 Panoramic Radiography is often used as the initial diagnostic image, and based on it, the specialist confirms the need for other more detailed and accurate examinations [5].

Tooth detection and implant and tooth numbering are the basic concepts that have been used in various studies by applying the artificial intelligence to panoramic radiographic images. Tooth numbering refers to the numbering of the prostheses and teeth and achieving the communication with the surrounding teeth. In the study of Sing et al., classification was done by using a deep neural network with 6 layers, including 3 convolutional layers and 3 fully connected layers, and the accuracy of 92% was achieved for the original data set [8]. In a study of Muramatsu et al., one hundred panoramic images were used for diagnosis, and in this method, the diagnosed teeth are divided into four types and three positions [9]. The sensitivity of the tooth detection in this method is 96.4%.

In a study by Zhao et al., a method has been presented for classification of teeth in panoramic radiographic images [10]. In designing the model, it is adapted from the Mask R-CNN method and includes branches for classification and segmentation. In this study, 400 radiographic images have been used, and the segmentation accuracy for U-Net method has been reported as 92% and for Mask-R-CNN as 98%. In terms of classification, the accuracy rate for the proposed network is 95%. In the study of Jader et al., a method for segmenting a tooth sample is presented by using the deep learning methods [11]. For this purpose, the CNN based on the desired area is used. In this study, 1500 images have been used, and the accuracy obtained was reported as 98%. Precision and specificity values are 94% and 99%, respectively. In 2019, Koch et al. investigated the segmentation of the panoramic dental radiographs in a study by using a fully connected network based on U-Net architecture [12]. The proposed network in Koch et al.'s study was tested on a dataset with 1500 images and the Dice value was reported as 93.4% [12]. In Min Yung et al.'s study, a method has been presented for detecting tooth in panoramic radiographic images [13]. In this way, not only the existing teeth but also the missing teeth are identified. The accuracy of the identification has reached 99.7%. In Nader et al.'s study, tooth segmentation was investigated on the panoramic radiographic images by using the U-Net network [14]. The results of that study have been reported to be about 90% according to Dice.

The purpose of the present study is to use the CNNs to segment the teeth and number them in the panoramic radiographic images. In the current study, a data set with appropriate volume and diversity has been used, and tooth segmentation and numbering were done by using the latest deep learning algorithms. The results show that the presented method was effective. Through the present study, the usual convolutional networks with special settings have been used, and the aim is to provide an end-to-end network for tooth segmentation and numbering, as the existing studies often perform this operation separately. Also, the used data set was collected by the researchers of the present study, which has high generalizability, and it has been tried to include different types of teeth. Also, the images are available in different qualities, which increases the generalizability and difficulty of the training. Also, in this study, diagnosing the missing teeth, dental implant and retaining root have been investigated, which according to our knowledge, has not been investigated before.

Methods and Materials.

The aim of the present study was to segment and number the teeth in Panoramic Radiographic images. For this purpose, a convolutional deep learning network was used. The network has been used for segmenting the U-Net network and numbering the VGG-16 network.

Data Collection

The data set used in the present study included 527 Panoramic Radiographic images which were collected from the patients referring to the Radiology Department of the Faculty of Dentistry of Tabriz. The collection of images was selective and it was tried to include healthy teeth, dental prosthesis, extracted or exfoliated teeth and decayed teeth in the dataset. Panoramic Images were selected from the visually acceptable images (appropriate contrast, resolution and brightness) and without the technical errors and distortions, and included patients over 12 years old and with permanent dental system, people with all 32 teeth (12 Incisors, 8 Premolars and 12 Muller) and patients with partial edentulousness. The presence of caries, dental restorations, fixed prostheses and endodontic teeth are not among the inclusion and exclusion criteria. The images were free of any lesions (except for a small periapical lesion and mild periodontitis), fractures, congenital anomalies, and dental anomalies (except for missing cases). Patients with complete edentulousness of one or two jaws were excluded from the study. Images of patients with implants and retaining roots were also selected to identify and number these cases.

After collecting the data set, the images were labeled by an oral and maxillofacial radiologist. This labeling is based on the teeth and a bounding box and a number were assigned to each tooth. These steps were done by using the tools available on the CVAT website. Then, the masks of the images as labels and the images themselves as the desired data set were entered into the pre-processing stage. In this study, the FDI numbering system was used, which is one of the most common tooth numbering methods [15].

Preprocessing.

In the pre-processing stage, we first changed the size of all images to 512x512. This resizing was to improve learning. Then we normalized the images between 0 and 1. This normalization was also done to improve learning.

Due to the small amount of data, data augmentation has also been used in network training. A variety of data augmentation methods such as Vertical and Horizontal Reflection, Gaussian Random Noise, Brightness Change and Salt and Pepper Noise have been added to the data. The amount of added noise in Gaussian and Salt-Pepper Noise is chosen equal to 10%.

Also, 80% of images were separated as training images, 10% of images as validation data, and 10% of images as test data. During the training of the network, the test data was not used in any way so that the final results could be relied upon. Data augmentation has not been applied to the test data.

Convolutional Neural Networks(CNNs)

CNNs, which are called ConvNet, are a type of artificial neural network that have a pre-deep architecture and compared to other networks, along with the fully connected layer, it has an amazing generalization capability that can learn very abstract features. This function is especially effective in the spatial features and can identify them well. The basic concepts of these networks have been examined in the previous sections and details are provided in this section.

A deep CNN model consists of a limited set of processing layers that can learn different

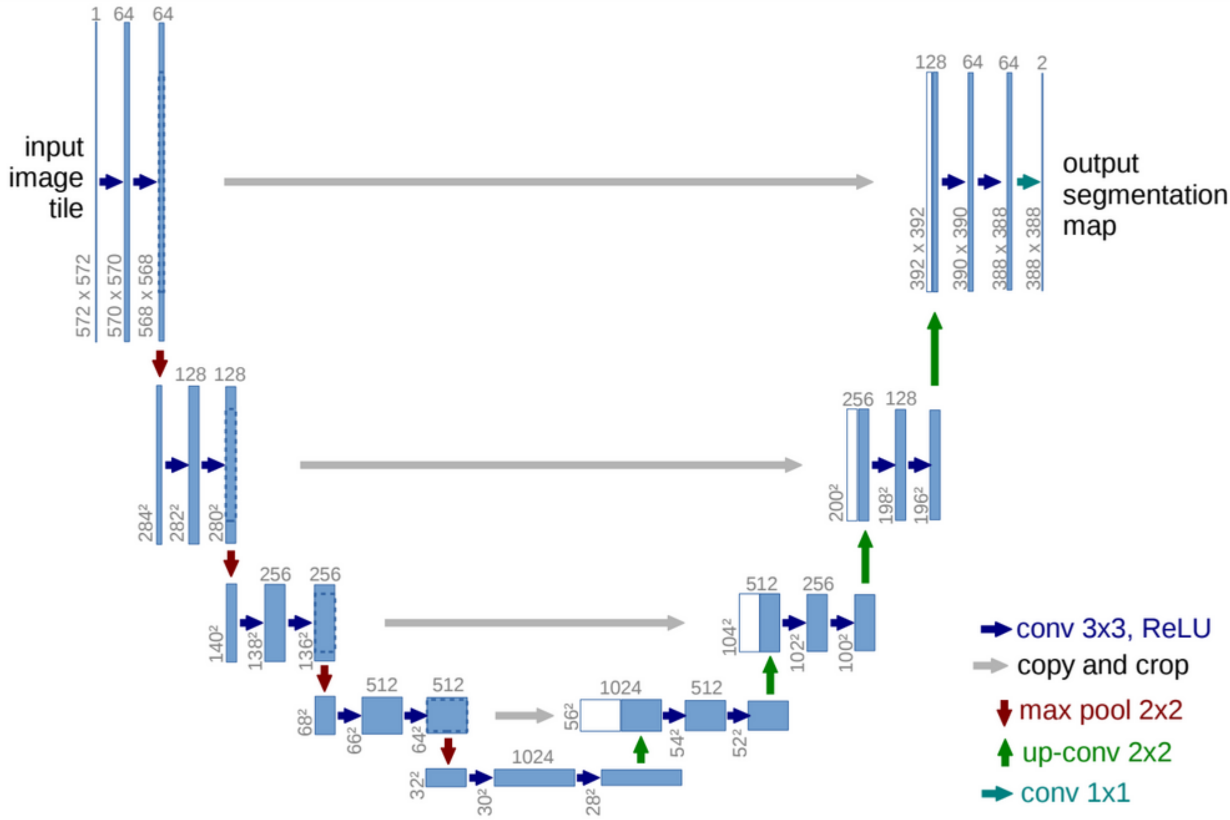

Fig. 1 (Рис. 1)

Fig. 1. **Diagram.**

U-Net Network Architecture for Image Segmentation.

Рис. 1. **Диаграмма.**

Архитектура сети U-Net для сегментации изображений.

features of the input data (such as an image) with multiple levels of abstraction. The primary layer's extract and learn the high-level features (with lower abstraction) and the deeper layer's extract and learn the low-level features (with higher abstraction). For example, in an image, the overall shape is the high-level feature and the edges and corners are counted as the low-level feature.

Compared to other networks, CNNs have a very good performance in the field of image processing, which has several reasons. One of the features of CNN is its weight sharing, which reduces the number of parameters that can be trained in the network, prevents overfitting of the model, and increases its generalization ability. Also, in CNN, the classification layer and the feature extraction layer are trained together, which makes the output of the model organized and the dependence of the output on the extracted features increases. Finally, it is much more difficult to implement a large network by using the other neural networks instead of the CNNs.

After examining different networks for segmentation and numbering teeth, two networks were used for this purpose, the first network was responsible for segmenting and the second network was used for numbering.

A typical use of the convolutional networks is classification tasks where the output of an image is a class label. However, in many visual tasks and especially the medical image processing, the desired output must have localization, that is, a class label must be assigned to each pixel. In the study, a network is introduced in which the class label of each pixel is taught by a local area around that pixel as input. This network is called U-Net [16].

U-Net architecture, which is a CNN and is popular in the medical image segmentation, has been used in the present study. The network architecture is shown in Figure 1. U-Net has a symmetric architecture that consists of an encoder network followed by a decoder network. The task of the encoder network is to map the images to a low-dimensional hidden representation. The function of the decoder

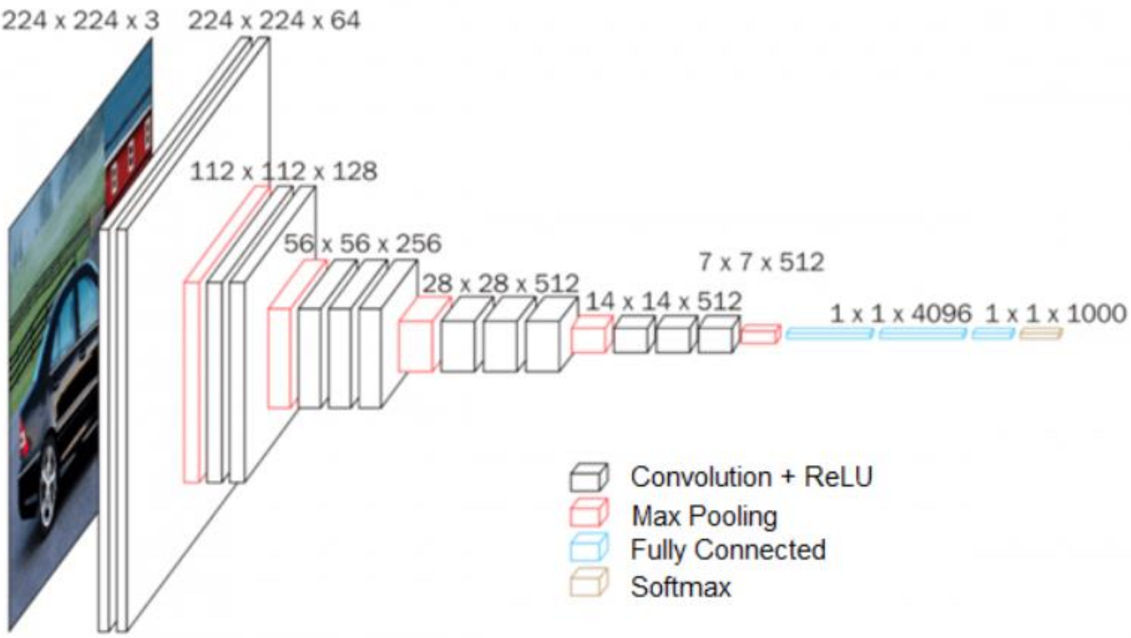

Fig. 2 (Рис. 2)

Fig. 2. Diagram.

VGG-16 Network Architecture.

Рис. 2. Диаграмма.

Архитектура сети VGG-16.

network is to reconstruct the output by increasing the rate of hidden vector samples to the size of the input.

In each level of the contraction path, two convolutional layers with 3x3 kernel and ReLU activity function is used, followed by a batch normalization layer. The samples of the feature plan are reduced by a factor of 2, and at the same time, the number of features is doubled, which is done by the 2x2 maximum integration operation in each step. In the expanding path, the increase of the samples is done with transposed 4x4 convolution. The random removal operation rate in each level, from the input to the output level, is equal to 0.15, 0.2, 0.3, 0.4, 0.5, 0.4, 0.3, 0.2, and 0.1, respectively. By the same connections, the features from each level of the contraction path are transferred to the same level of the expansion path. After each scaling operation, two convolutional layers with 3x3 kernel and ReLU activity function are applied, followed by batch normalization. In the final step, the output of the network is obtained by applying a 1x1 convolution and a Sigmoid activity function. We have used binary mutual entropy cost function in network training. The weights of the network were obtained by using the Adam Optimizer in 250 iterations with a group size of 4.

In order to apply the outputs of the U-Net network for numbering the teeth, each of the bounding boxes was cut and then entered into the VGG-16 network. Then the VGG CNN learned each image to predict a number from 1 to 32. The output of the classifier is a set of confidence scores in all classes for each bounding box that indicates how likely each tooth belongs to each of the 32 classes.

Then these data were post-processed to improve the prediction results. This post-processing was based on the assumption that each tooth can be created only once in the image and in a certain order. This algorithm works in such a way that first the bounding boxes of the predicted teeth are arranged based on the coordinates in each jaw. Then the number of missing teeth is counted based on the number of teeth. All possible combinations of teeth are examined and the total confidence score is calculated. The combination that has the highest confidence score is selected.

The output of the U-Net network enters a dental numbering network. This section was created based on the VGG-16 convolution architecture. To classify the teeth based on their number, this module uses the output of the tooth detection module. The teeth are cut based on the predicted bounding boxes. Then these cropped images are classified by the VGG-16 to predict a dental number.

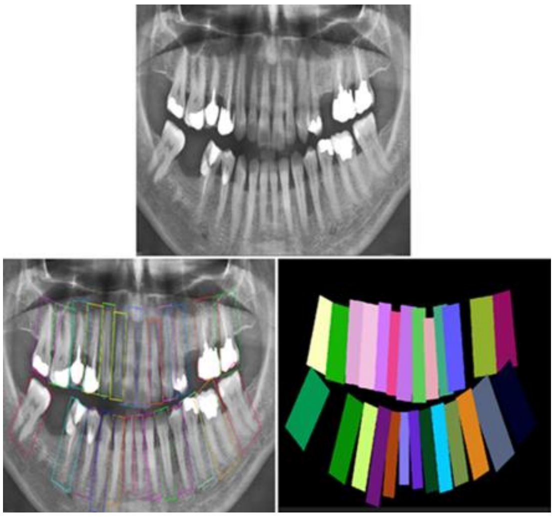

Fig. 3 a (Рис. 3 a)

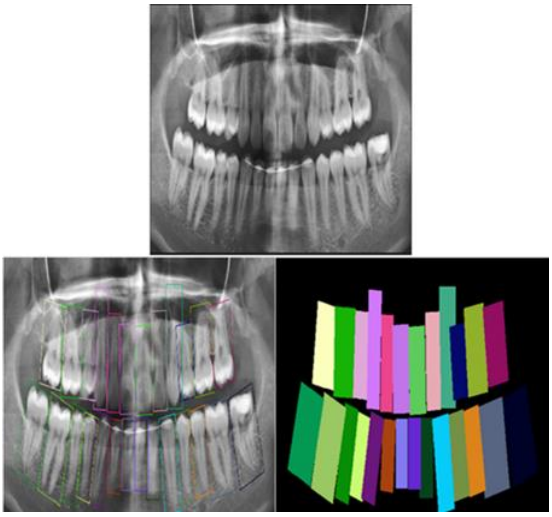

Fig. 3 b (Рис. 3 б)

Fig. 3. Panoramic Images and Related Labels of (a) person A, (b) person B.

Upper figures – the dental panoramic radiographic images for person A and B. Left lower figures – the corresponding mask made by expert placed on panoramic radiographic images. Right lower figures – the corresponding mask made by expert for each person.

Рис. 3. Панорамные изображения и соответствующая маркировка (а) пациента А, (б) пациента В.

Верхний ряд – ортопантомограммы пациентов А и В. Нижний ряд слева – соответствующая маркировка экспертом на ортопантомограмме. Нижний ряд справа – соответствующая маркировка экспертом для каждого пациента.

The VGG network has different settings based on the depth of the network. Usually, the VGG network has several stacks of the convolution layers with filter size 3x3, step 1 and integration 1, after the convolution layer, a maximum integration layer with size 2x2 is placed. Among all available settings for VGG, VGG-16 has the best performance on the available data sets, whose general structure is shown in Figure 2.

As shown in Figure 3, first the input is entered into two convolutions with a filter size of 3x3, each of which is followed by a ReLU activity function. Each of these two layers contains 64 filters. The convolution step is equal to 1 pixel and the padding value is equal to 1. Then the feature plan is entered into a maximum integration layer whose size is equal to 2x2 and its step is equal to 2. In the next step, the number of filters is doubled and becomes equal to 128 and continued until the last step. The final three fully-connected layers have the number of neurons 4096, 4096 and 32. The number of final layers has been selected for the number of desired classes. In Figure 3, the panoramic dental images of two people are shown in figures (a) and (b) of the upper figure. It is also clear that the labeling done by the expert is in the lower figures, that the left figure in figure (a) corresponds to the labels made by the expert for each tooth, and the right figure shows these labels outside the dental image. Each color represents a tooth number. These modes are also true for figure (b). Also, the teeth that are not present in the figure also do not have their corresponding color in the figure, and in this way the network notices the absence of the corresponding tooth.

The success criterion indicates the result of the network training, which provides an overall evaluation of the network. To evaluate the segmentation results, we used the criteria of sensitivity, specificity, Positive Predictive Value (PPV), Negative Predictive Value (NPV), and Dice similarity coefficients, which are specified in relation (1) of the relationships of these criteria. Also, the accuracy criterion was used to report the numbering of teeth. In these relationships, TP indicates a True Positive, TN indicates a True Negative, FN indicates a False Negative, and FP indicates a False Positive. These criteria are obtained pixel by pixel in the segmentation mode.

Sensitivity, Specificity, PPV, NPV and Dice Formula.

$$\text{Sensitivity} = \frac{TP}{TP + FN} \qquad \text{Specificity} = \frac{TN}{TN + FP} \qquad \text{PPV} = \frac{TP}{TP + FP}$$

$$\text{NPV} = \frac{TN}{TN + FN} \qquad \text{Dice} = \frac{2TP}{2TP + FP + FN}$$

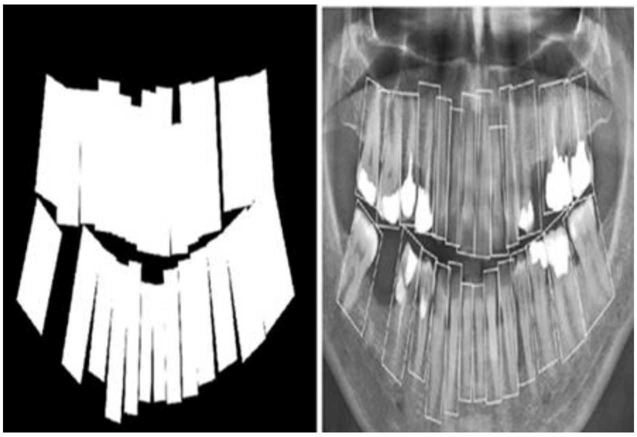

Fig. 4 a (Рис. 4 а)

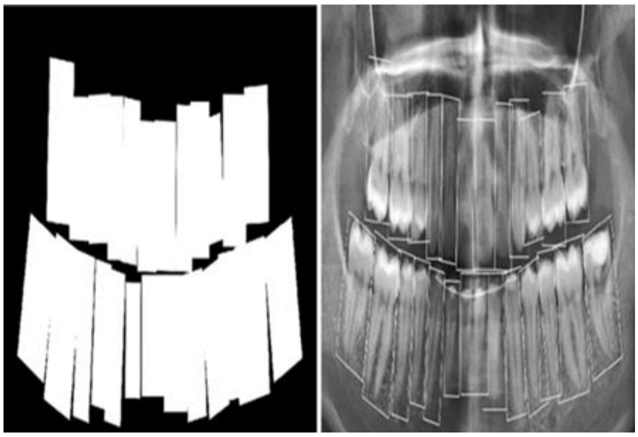

Fig. 4 b (Рис. 4 б)

Fig. 4. U-Net Output Masks of (a) person A, (b) person B.

Left figures – output mask of the person A and B.

Right figures – Placement of output mask on panoramic radiographic image

Рис. 4. Результаты маркировки U-Net для (a) пациента A, (b) пациента B.

Рисунки слева – результаты маркировки для пациентов A и B.

Рисунки справа – расположение маркировки на ортопантомограмме.

In the present study, we also have used the Dice criterion to evaluate the segmentation. This criterion is used to calculate the similarity of two images, which indicates the difference between the resulting mask and the original image mask. The value of Dice is equal to twice the commonality of two images divided by the union of the two images.

Results.

The aim of this study is to segment and number the teeth in a convolution network in an End-to-End manner. For this purpose, we have used a U-Net network for segmenting and a VGG-16 network for teeth numbering, and the results of these two networks are analyzed in this section.

Radiographic panoramic images were resized to 512x512 during pre-processing. The existing masks and labels have been obtained manually by an expert. The labeling was done by an oral and maxillofacial radiologist and was approved by several experts. An example of the corresponding image and label is shown in Figure 3.

The U-Net network was responsible for segmenting the images. After entering the training and validation images into U-Net network, network training was done. The output of the network is blocks of segmentation, where a bounding box is defined for each block. \ In Figure 4, the output mask of the network is shown for person A and B. These people are the same people in figure 3, where figures (a) and (b) represent the respective person respectively. Also, the figure on the left shows the output mask for each person, and in the figure on the right, this mask is placed on the person's teeth to show the segmentation more accurately. Also, according to this figure, the accuracy of the network can be seen visually.

As it can be seen in Figure 4, a good output has been obtained and in some values the difference with the expert labels is clear. The image of person A has a lot of errors in recognition and that is why it has been shown in the sample images.

The final results of the U-Net network were measured based on the various criteria. Various experiments were also performed by using the different data augmentation, which are specified in Table №1. Each of the data augmentation used in each experiment is marked with * symbol in the table.

The results of this network in terms of sensitivity, specificity, PPV, NPV and Dice are shown in Table №2. The reported values are the average result of several repetitions of the experiment. Also, the results were obtained on test data and the network had no observations on them before that.

After this step, the results obtained from the U-Net network were entered into the VGG network for teeth numbering. The VGG CNN learned each image to predict a number from 1 to 32. The output of the classifier is a set of confidence scores in all classes for each bounding box that indicates how likely each tooth belongs to each of the 32 classes. Then these data

Table №1. Different experiment designs of segmentation task.

Experiment	Salt and Pepper	Vertical Reflection	Horizontal Reflection	Gaussian Random Noise	Brightness Change	Without Data Augmentation
1	*	*	*	*	*	-
2	*	*	*	-	-	-
3	-	*	*	*	-	-
4	-	*	*	-	*	-
5	-	-	-	-	-	*

Table №2. Results of U-Net network for segmentation of teeth.

Experiment	Sensitivity	Specificity	PPV	NPV	Dice
1	96.3 %	98.9 %	95.1 %	99.4 %	95.3 %
2	96.4 %	99 %	94.4 %	99.3 %	94.5 %
3	95.9 %	99.1 %	95.2 %	99.1 %	95.2 %
4	96.1 %	98.9 %	94.8 %	99.3 %	95.1 %
5	98.9 %	98.4 %	92.3 %	99.8 %	95.4 %

Fig. 5 a (Рис. 5 а)

Fig. 5 b (Рис. 5 б)

Fig. 5. Examples of Detected Teeth Bounding Boxes of (a) person A, (b) person B.

Рис. 5. Примеры обнаруженных ограничивающих рамок для зубов (а) пациента А, (б) пациента В.

were post-processed to improve the prediction results. An example of cut images for the person A and B is shown in figure 5.

Finally, the results of the numbering of person A and B are shown in figure 6. In this figure, the final results for two persons A and B are shown along with the numbering using the VGG-16 network.

The results of tooth numbering have been shown in Table №3. As it is clear in figures 6, the numbering has an acceptable accuracy and the missing teeth are also marked with a bounding box. The results of the VGG-16 network for teeth numbering have been shown in Table 3. The results reported in this table are the result of averaging several tests. As it is clear in the table, the numbering accuracy for teeth is equal to 96.8%, which is considered a suitable value for 32 teeth.

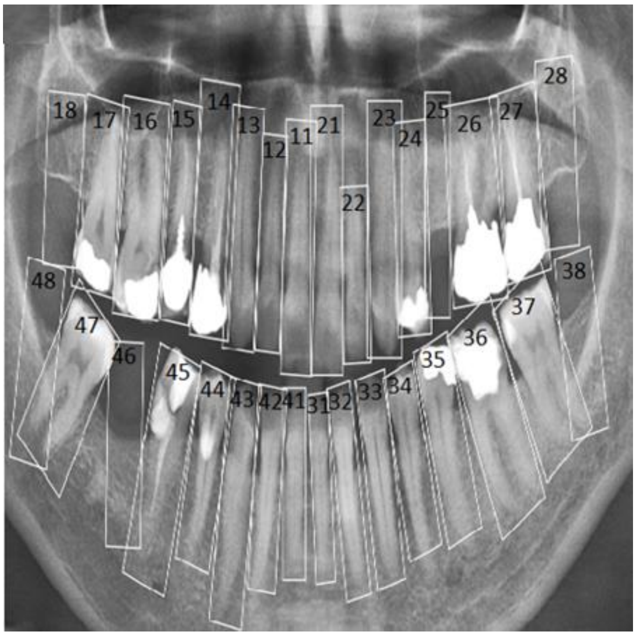

Fig. 6 a (Рис. 6 а)

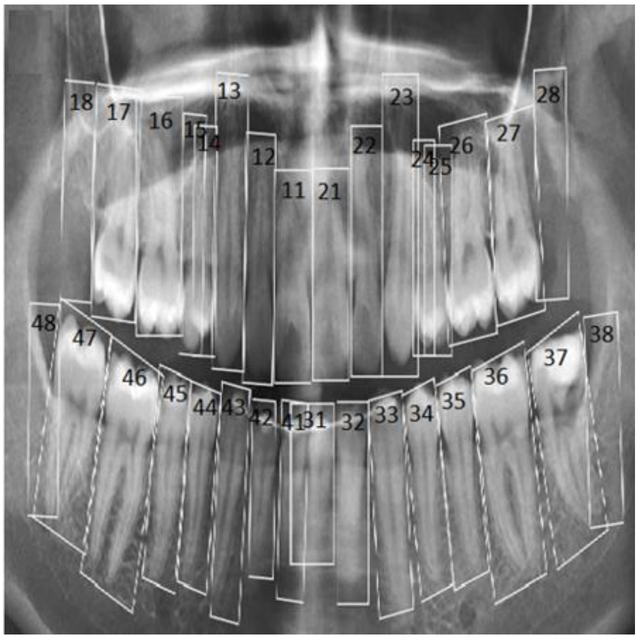

Fig. 6 b (Рис. 6 б)

Fig. 6. Numbering output of VGG-16 network (a) person A, (b) person B.

Рис. 6. Результаты нумерации сети VGG-16 (а) пациента А, (б) пациента В.

Table №3. Results of teeth numbering using VGG-16.

Variable	Value
Sensitivity	98.58 %
Specificity	99.93 %
Accuracy	96.8 %

Table №4. Detection result of extracted teeth region, retaining roots, and dental implants.

Variable	Accuracy
Implant	98.45 %
Retained root	97.1 %
Extracted tooth region	98.2 %

Also, the VGG-16 network has been used to detect extracted teeth region, retaining roots, and dental implants. In this case, the number of classes is chosen equal to 3 classes and the segmenting is done like the dental numbering. The different accuracies of this network for these three types of classes have been shown in Table №4. The values reported in this table are the average of several experiments. As it is clear in this table, the accuracy obtained for the dental implant is 98.45%, for the retaining root it is 97.1% and for the extracted tooth it is 98.2%. For example, in Figure 7, all the cases have been correctly diagnosed, except for one of the teeth, which was diagnosed as a retained root instead of an extracted tooth.

Discussion.

Automatic segmentation of teeth in the Panoramic Radiographic Images is one of the most important processes in the field of dentistry. For this purpose, in the present study, our goal was segmenting and numbering of teeth in the Panoramic Radiographic Images.

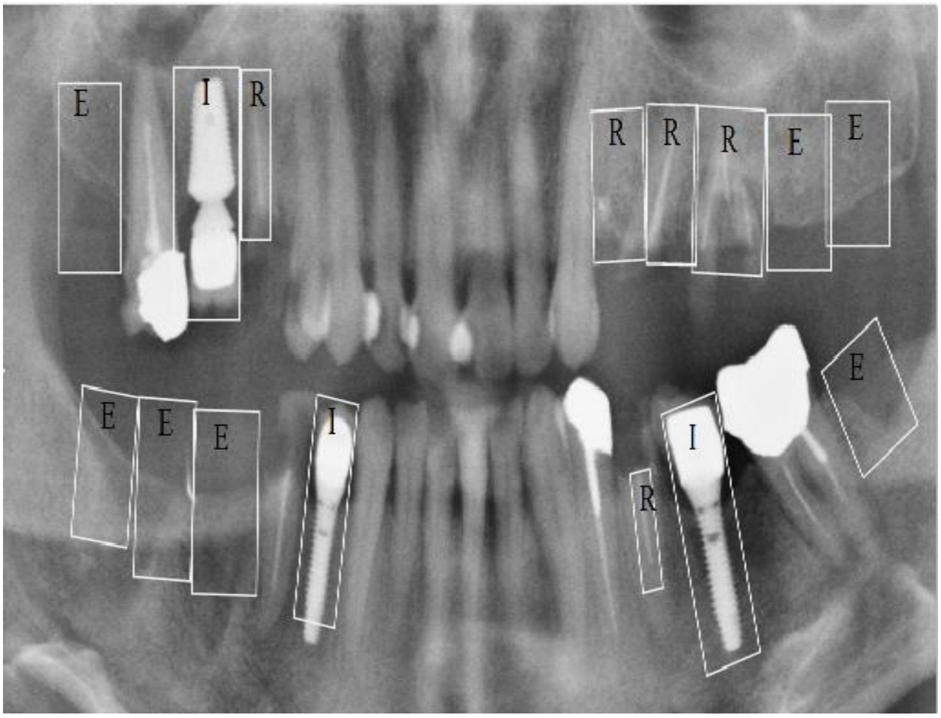

Fig. 7 (Рис. 7)

Fig. 7. An example of implant (I), retained root (R) and extracted tooth area (E).

Рис. 7. Пример имплантата (I), остаточного корня (R) и области удаленного зуба (E).

Table №5. Segmentation result for comparison of related works and proposed method.

Study	Year	Method	Number of Data	Sensitivity	Precision	Accuracy	Specificity	Dice
Tuzoff et al. [17]	2019	Fast R-CNN	1572	99.41%	99.45%	-	-	-
Silva et al. [18]	2020	PANet	778	-	94.4%	96.7%	98.7%	-
Singh et al. [8]	2020	Phase C-Mean	400	-	-	-	-	-
Zhao et al. [10]	2020	Mask R-CNN	400	-	98%	-	-	-
Jader et al. [11]	2019	Region-based CNN	1500	-	94%	98%	99%	-
Koch et al. [12]	2019	U-Net	1500	94.37%	93.31%	95.21%	96.14%	93.6%
Nader et al. [14]	2022	U-Net	543	-	-	-	-	90%
Our study	2022	U-Net	527	96.3%	-	-	98.9%	95.3%

Table №6. Numbering result for comparison of related works and proposed method.

Study	Year	Method	Number of Data	Sensitivity	Specificity	Precision	Accuracy
Tuzoff et al. [17]	2019	VGG-16	1572	98%	99.95%	-	-
Silva et al. [18]	2020	PANet	778	-	-	74 % mAP	-
Singh et al. [8]	2020	6-Layer CNN	400	98%	-	-	-
Zhao et al. [10]	2020	Mask R-CNN	400	98%	92%	-	95%
Our study	2022	VGG-16	527	98.58%	99.93%	-	96.8%

The proposed network for tooth segmentation consists of a U-Net convolutional network. In the current study, the images obtained from the radiology center of the Faculty of Dentistry of Tabriz were labeled by an expert. These images were entered into the U-Net network after pre-processing, and the label prepared by the expert was also entered into the network separately. After training the network, the sensitivity value was 98.9%, the specificity was 98.4% and the Dice value was 95.3%. The results of the proposed method have acceptable performance compared to the current methods and are suitable for the next step. Also, the comparison of the proposed network in terms of segmentation with existing and updated methods is provided in Table №5. In the existing methods for tooth segmentation, out of the seven studied investigations, three studies have 1500 samples and the rest of the studies have between 400 and 700 samples. Among these studies, two studies have used different states of U-Net for segmentation. In the study of Tuzoff and et al., the Fast R-CNN method was used, the number of images was 1572 and the accuracy value was 99.45% [17]. This value is the highest accuracy reported in studies. Among the studies that used U-Net, study of Nader et al. with 543 samples has a Dice value of 90%, and in the study of Koch et al. with 1500 samples, the Dice value is reported as 93.6% [12, 14]. According to the network settings and its hyperparameters, in our proposed method, even though the number of samples is 527, the Dice value is 95.3%.

Then, for the purpose of numbering, the cropped images are entered into the network and a confidence score is calculated for each image. This confidence score shows how likely each image belongs to each of the classes. Finally, by using the hypotheses included in the algorithm, such as the order of the teeth and the presence of one tooth in each image, the teeth are numbered. Also, the teeth that are not in the image have a bounding box in their place, which indicates the missing teeth. Our proposed method has a sensitivity value of 98.58%, a specificity of 99.93% and an accuracy of 96.8%, which has good performance compared to the existing methods and promising results have been obtained. A comparison of the proposed method with other tooth numbering methods has been shown in Table 6. As you can see in Table №6, the study method of Tuzoff and et al. has a sensitivity of 98 and a specificity of 95.99%, which has similar values with the results of the present study [17]. The numbering network of the proposed method is also similar to the network of the study of Tuzoff and et al., but the main difference is the number of images in the data set, which is almost three times of the present study [17].

In comparison with studies, in present study implant, retained root and extracted tooth region were considered in numbering and the present neural network showed high accuracy.

Conclusion.

Our study showed good results in both segmentation and numbering methods even with limitations in data set amount. The use of deep learning networks, despite creating very good results, has other limitations. Due to the fact that feature learning in deep learning is done by the network, training images should have a good variety than any other types of data. Also, the amount of data required for training the network is very important, and the more the amount of data, the better the training will be.

As suggestions for the future studies in this field, the following can be mentioned:

• Using a larger and more comprehensive data set for segmenting and numbering purposes.

• Using deeper networks with more learning capabilities.

• Labeling the teeth accurately according to the shape of the teeth and segmenting and numbering them.

• Using the pre-processing methods to increase image quality.

References:

1. *Kim C, Kim D, Jeong H, Yoon S-J, Youm S. Automatic tooth detection and numbering using a combination of a CNN and heuristic algorithm. Applied Sciences. 2020; 10 (16): 5624.*
2. *Bilgir E, Bayrakdar İŞ, Çelik Ö, et al. An artificial ıntelligence approach to automatic tooth detection and numbering in panoramic radiographs. BMC Medical Imaging. 2021; 21 (1): 1-9.*
3. *Chen H, Zhang K, Lyu P, et al. A deep learning approach to automatic teeth detection and numbering based on object detection in dental periapical films. Scientific reports. 2019; 9 (1): 1-11.*
4. *Hallikainen D. History of panoramic radiography. Acta radiologica. 1996; 37 (3): 441-445.*
5. *White SC, Pharoah MJ. White and Pharoah's Oral Radiology: Principles and Interpretation. Elsevier Health Sciences, 2018.*
6. *Haring JI, Jansen L. Dental radiography: principles and techniques. WB Saunders, 2000.*
7. *Moghaddam S.A., Yadekar M., Vahdat A.S., Esmaeili F. Automatic detection of mandibular fractures on panoramic radiographs using the convolutional neural network. Russ Electron J Radiol. 2023; 13 (3): 5-13.*
8. *Singh P, Sehgal P. Numbering and classification of panoramic dental images using 6-layer convolutional neural network. Pattern Recognition and Image Analysis. 2020; 30 (1):125-133.*
9. *Muramatsu C, Morishita T, Takahashi R, et al. Tooth detection and classification on panoramic radiographs for automatic dental chart filing: improved classification by multi-sized input data. Oral Radiology. 2021; 37 (1): 13-19.*
10. *Zhao Y, Li P, Gao C, et al. TSASNet: Tooth segmentation on dental panoramic X-ray images by Two-Stage Attention Segmentation Network. Knowledge-Based Systems. 2020; 206: 106338.*
11. *Jader G, Fontineli J, Ruiz M, Abdalla K, Pithon M, Oliveira L. Deep instance segmentation of teeth in panoramic X-ray images. Paper presented at: 2018 31st SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI), 2018.*
12. *Koch TL, Perslev M, Igel C, Brandt SS. Accurate segmentation of dental panoramic radiographs with U-NETS. Paper presented at: 2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019), 2019.*
13. *Chung M, Lee J, Park S, et al. Individual tooth detection and identification from dental panoramic x-ray images via point-wise localization and distance regularization. Artificial Intelligence in Medicine. 2021; 111: 101996.*
14. *Nader R, Smorodin A, De La Fourniere N, Amouriq Y, Autrusseau F. Automatic teeth segmentation on panoramic X-rays using deep neural networks. Paper presented at: International Conference on Pattern Recognition, 2022.*
15. *Harris EF. Tooth-coding systems in the clinical dental setting. Dental Anthropology Journal. 2005; 18 (2): 43-49.*
16. *Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation. Paper presented at: International Conference on Medical image computing and computer-assisted intervention, 2015.*
17. *Tuzoff DV, Tuzova LN, Bornstein MM, et al. Tooth detection and numbering in panoramic radiographs using convolutional neural networks. Dentomaxillofacial Radiology. 2019; 48 (4): 20180051.*
18. *Silva B, Pinheiro L, Oliveira L, Pithon M. A study on tooth segmentation and numbering using end-to-end deep neural networks. Paper presented at: 2020 33rd SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI), 2020.*