UNet Architecture Based Dental Panoramic Image Segmentation

S. Sivagami[1], P. Chitra[2], G. Sri Ram Kailash[3], and S.R. Muralidharan[4]

[1,2,3,4]Department of CSE, Thiagarajar College of Engineering, Madurai
[1]ssivagami186@gmail.com, [2]pccse@gmail.com, [3]sriramkailash88@gmail.com

***Abstract*—This paper proposes an UNet architecture that uses convolutional neural networks to achieve accurate segmentation of Dental panoramic x-ray images. In dentistry, Radiographic images help medical experts to identify and diagnose the disease in an accurate manner X-rays, Computed Tomography (CT), Magnetic resonance imaging (MRI) are some of the radiographic images. Generally, X-ray images are complex in nature. The presence of noise results in lack of reliable separation between the various parts of teeth. This makes the segmentation process very difficult. Dental image segmentation helps the dentist to detect the impacted teeth, find the accurate position for the placement of dental implants and determine the orientation of teeth structure. UNet architecture model is a recent approach used for medical image segmentation. In this paper we took an advantage of the UNet architecture for dental x-ray image segmentation and achieved an accuracy of 97 % and Dice score of 94 %. Also, the performance of UNet architecture for dental x-ray image segmentation is compared with other image segmentation algorithms.**

***Index Terms*—Segmentation, UNet architecture, F1 score, dental implants.**

I. Introduction

In image processing and computer vision techniques, image segmentation is one of the major hotspots. The main aim of the image segmentation is to divide an image into several segments having similar features or attributes. Image segmentation acts as an important basis for image recognition and image analysis. Medical imaging, object detection and recognition tasks, content-based image retrieval, automatic traffic control systems and video surveillance are the basic applications of image segmentation.

There are several types of image segmentation techniques namely region-based, cluster based, boundary based and watershed-based image segmentation etc. Medical image Segmentation is the basic requirements in medical applications for diagnosing and identifying the conditions of the diseases earlier.

Convolutional neural network (CNN) is the most common techniques used in computer vision for image segmentation. CNN have a great attention towards medical image analysis. CNN uses a machine learning classifier [1] for segmenting the normal gingiva with inflamed gingiva from intra-oral images. Through this segmentation, dental professionals get the accurate information for early diagnosis of periodontal diseases. [2] focus on dental image segmentation and labelling the dental images using two level hierarchical CNN approach. First approach is teeth-gingiva labelling and other is the inter teeth labelling. This helps to segment the teeth and label all the teeth in an accurate manner. Semi-automatic segmentation method in CNN is used to separate the teeth from socket in CT images using threshold values. This threshold values will be calculated based on the reference images [3].

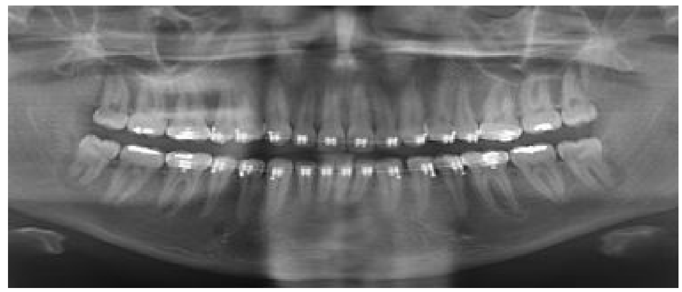

Fig. 1. Panoramic x-ray image.

The paper is organized as follows. Section II presents input modalities used in this work. Section III explains the related work. Section IV discusses the proposed work. Section V address about Results and Discussions.

II. Input Modalities

In dentistry, there are different modalities for acquiring the dental images. They are x-rays, Computed tomography (CT), cone-based chromatography (CBCT). Radiographic x-rays are commonly divided into intra oral x-rays, and extra oral x-rays. Intra oral x-ray image is obtained inside the patients mouth and extra oral x-ray image is obtained outside the patients mouth [4].

Among the other types of radiographic x-ray images, Dental Panoramic x-ray image give the detailed anatomical structure of teeth. Fig. 1 shows two-dimensional panoramic radiographic x-ray image. This type of radiographic examinations allows the visualization of dental irregularities such as cysts, tumors, oral cancers.

III. Related Works

A. Global Thresholding

The main idea behind this thresholding segmentation is to select the initial threshold to separate the image into two regions 1. Background regions 2. Foreground regions using threshold value. Image objects are set as background in gray

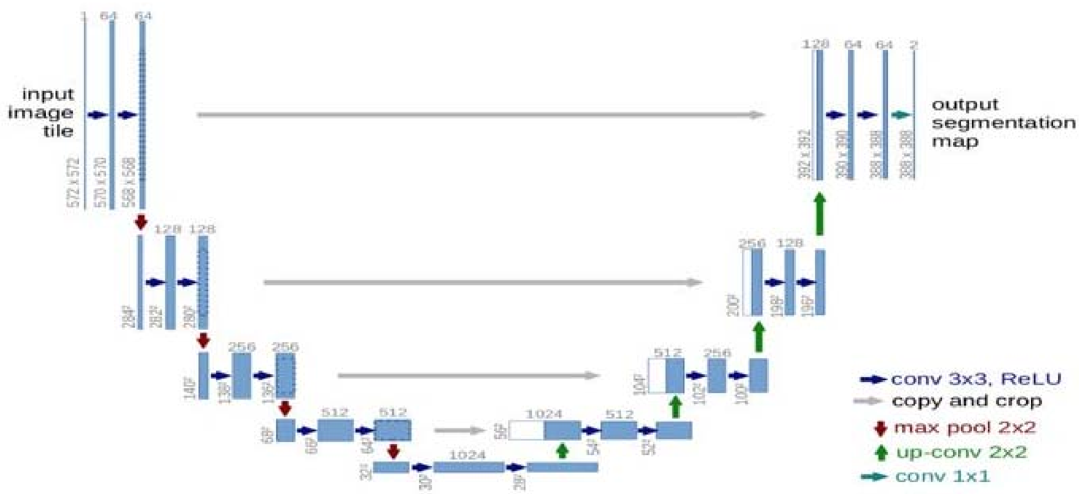

Fig. 2. UNet architecture [11].

colors are usually distributed with different means of values The threshold value separate the objects from background easily [5]. There are several types of thresholding methods in image segmentation i.e. multilevel Otsu thresholding method and sauvola thresholding. These are the methods which are most commonly used for thresholding based dental image segmentation [6].

B. Fuzzy C-Means

Fuzzy C-Means (FCM) is one of the iterative clustering-based segmentation approaches and it also named as soft clustering algorithm. FCM is the generalized form of classical K-Means algorithms. FCM is the popular method in image segmentation could retain many information than hard segmentation methods. Hybrid of fuzzy c-means and neutrophonic approach is used to segment the jaw and the lesions present in the jaw using panoramic radiographic images [7]. Semi supervised fuzzy clustering approach with spatial constraints (SFCM-SC) approach is used in segmenting the teeth. In this SFCM-SC approach, number of clusters is very small compared with other approaches and segmentation could be done in very accurate manner [8].

C. Watershed Algorithms

This watershed segmentation algorithm which is a mathematical morphology method for image segmentation based on image processing. It is a transformation on grayscale images. The steps followed in watershed algorithms are a) finding gradients on image ii) Applying watershed on gradient image iii) Applying morphology to get the image segmentation. Marker controlled watershed algorithms is used to segment the tumor cells from the jaw in radiographic images [9].

D. Canny Edge Detection

Canny edge detection process is a trade-off between noise reduction and edge localization in an image. There are four steps followed in canny edge segmentation algorithm 1. Apply gaussian filter to smooth the image 2. Find the intensity gradient 3. Apply non maximum suppression 4. Track the edge by Hysteresis. [10] In radiographic imaging, edge detection plays on important role in image segmentation for detecting and diagnosing the diseases.

IV. Proposed Work

The network architecture used in this image segmentation model is UNet architecture. This architecture is mainly designed for segmentation of medical images. UNet architecture is an end-to-end fully convolutional neural network. This architecture consists of two paths namely contraction path and expansion path. Looks like an 'U' shaped structure. In Contraction path, convolution operations have been carried out and followed by maximum pooling operations with a stride size. Transposed convolution operation is to be done in expansion path.

UNet architecture comprises of two 3×3 convolutions, followed by Rectified Linear Unit (ReLu) and 2×2 maximum pooling operations with the stride of 2 for down sampling path. In Up sampling path, 2×2 transposed convolution operation taken place for reducing the feature channels [11]. Convolution path Skip connections also introduced in the UNet architecture. This connection is used to skip the features from the contracting path to the expanding path in order to recover

TABLE I
Parameter Description for Proposed Model

Model used	Sequential parameters
Activation function (Input)	ReLU
Activation function (Output)	SoftMax
Optimizer	Adam
Loss function	Binary Cross entropy
Number of epochs	20
Batch size	1
Validation split	0.15

TABLE II
RESULT SUMMARY

Segmentation Method	Accuracy	Specificity	Precision	Recall	F1-Score	Dice Score
Global thresholding	0.79	0.81	0.52	0.69	0.56	0.74
Fuzzy C-means	0.82	0.91	0.61	0.45	0.49	0.81
Watershed	0.77	0.75	0.48	0.82	0.58	0.75
Canny Edge detection	0.79	0.94	0.45	0.11	0.17	0.54
UNet	0.97	0.95	0.93	0.94	0.93	0.94

the spatial feature lost during down sampling operations [12]. So, the segmentation is very fast and accurate when compared with other segmentation methods.

UNet which outperformed the other deep architectures based on the following attractive characteristics:

1. This architecture has the combinations of convolutional, pooling and Up-Sampling layers.
2. Instead of Tanh, logistic, arctan or Sigmoid as activation function it uses ReLU function which reduce likelihood of vanishing gradient problem.
3. It trains faster than other deeper architectures.

Fig. 2 shows the UNet architecture. Each box corresponds to a multi-channel feature map. The number of channels is mentioned on top of the box. The size of x, y size is provided at the lower left edge of the box. White boxes represent copied feature maps. The blue arrows denote the convolution operations followed by pooling operations that should be represent as in red arrow, green arrow tells the up-sampling operations.

V. Results and Discussions

A. Dataset

Data set was obtained from Ivisionlab [13]. The data set consists of 1171 dental panoramic X-ray images in total with the classes of 15, including 942 images for training, 166 images for validation and 63 images for testing. This dataset consists of both panoramic x-ray image and ground truth image. UNet architecture were implemented using the keras framework in Python and performed our experiment on NVIDIA GEFORCE GPU.

B. Evaluation Metrics

i) Accuracy—Accuracy is the most instinctive performance measure and it is a ratio of correctly predicted observation to the total observations.

$$\text{Accuracy} = (\text{TP+TN})/(\text{TP+FP+FN+TN}) \quad (1)$$

ii) Precision—Precision is the ratio of correctly predicted positive observations to the total predicted positive observations.

$$\text{Precision} = \text{TP/TP+FP} \quad (2)$$

iii) Recall (Sensitivity)—Recall is the ratio of correctly predicted positive observations to the all observations in actual class.

$$\text{Recall} = \text{TP/TP+FN} \quad (3)$$

iv) F1 score—It is the weighted average of Precision and Recall. Therefore, this score takes both false negatives and false positives into considerations.

$$\text{F1 Score} = 2*(\text{Precision} * \text{Recall}) / (\text{Precision} + \text{Recall}) \quad (4)$$

v) Specificity—Specificity (SP) is calculated as the number of correct negative predictions divided by the total number of negatives.

$$\text{Specificity} = \text{TN/TN+FP} \quad (5)$$

vi) Dice Score—Dice score is calculated as the twice the number of true positives divided by the sum of twice the number of true positives and number of false positive and false negatives.

$$\text{Dice score} = \text{2TP/2TP+FP+FN (6)} \quad (6)$$

Table I shows that the parameters description used in the proposed model. ReLU activation function is a piecewise linear function and ReLU is responsible for weighted input from the node into the activation of the output node. ReLU is easier to train and attains better performance than other activation function, so ReLU is one of the added advantages of UNet architecture. SoftMax function gives the probability values of the output. Adam optimizer is used to achieve the results very fast.

Tables II, III shows the summary of results, UNet architecture have the highest accuracy and F2-score is also high when compared with state-of-art segmentation approach.

VI. Conclusion

In general, Medical image segmentation plays an essential role on Computer diagnosis system. Through medical image segmentation, Doctors could able to diagnose the disease and make accurate decisions. Particularly, Dental image segmentation in panoramic x-ray images has been search for many years in supervised methods and unsupervised methods in machine learning approaches. In this paper, UNet architecture is the proposed model for accurate segmentation of panoramic radiographic dental images. UNet is one the popular convolutional neural network (CNN) architecture for medical image segmentation. Our proposed method works well without the characteristics of dense connection, residual connections and inception module. UNet architecture works on both supervised and unsupervised methods. Images of different size is given as input in UNet architecture which generates the high-resolution image from blurry images in a fast manner. UNet is an end-to-end encoder-decoder network architecture. The encoder part

TABLE III
EXPERIMENTAL ANALYSIS OF VARIOUS SEGMENTATION METHODS

Name of the Segmentation	X- ray Image	Ground Truth	Segmented Image
Global Thresholding			
Watershed Segmentation			
Canny edge detection			
Fuzzy C- means Clustering			
U-Net Architecture			

of UNet architecture learns low level features and decoder part learns high level features from the encoder features. Skip connections also used to concatenate the features of both encoder and decoder path. This concatenation operations allows deep supervision of the networks. So, the segmentation of medical images with U-Net architecture is very accurate with the accuracy of 97% and Dice score of 94 % compared with other start-of-the-art methods. As our work is done with the panoramic X-ray images, we can use this architecture for other types of dental X-ray images like buccal images, bitewing images etc. In future, our proposed architecture could be used for segmentation of three-dimensional medical images.

REFERENCES

[1] A. Rana, G. Yauney, L.C. Wong, O. Gupta, A. Muftu, P. Shah, "Automated segmentation of gingival diseases from oral simages," In *2017 IEEE Healthcare Innovations and Point of Care Technologies (HI-POCT)*, vol. 6, pp. 144–147, IEEE, Nov 2017.

[2] X. Xu, C. Liu, and Y. Zheng, "3D tooth segmentation and labelling using deep convolutional neural networks," *IEEE Trans. on Visualization and Computer Graphics*, vol. 25, no. 7, pp. 2336–2348, 2018.

[3] H.C. Kang, C. Choi, J. Shin, J. Lee, and Y.G. Shin, "Fast and accurate semiautomatic segmentation of individual teeth from dental CT images," *Computational and Mathematical Methods in Medicine*, 2015.

[4] G. Silva, L. Oliveira, and M. Pithon, "Automatic segmenting teeth in X-ray images: Trends, a novel data set, benchmarking and future perspectives," *Expert Systems with Applications*, vol. 107, pp. 15–31, 2018.

[5] C.H. Huang and C.Y. Hsu, " Computer-assisted orientation of dental periapical radiographs to the occlusal plane," *Oral Surgery, Oral Medicine, Oral Pathology, Oral Radiology, and Endodontology*, vol. 105, no. 5, pp. 649–653, 2008.

[6] R. Indraswari, A.Z. Arifin, D.A. Navastara, and N. Jawas, "Teeth segmentation on dental panoramic Radiographs using decimation-free directional filter bank thresholding and multistage adaptive thresholding," In *2015 International Conference on Information & Communication. Technology and Systems (ICTS)*, pp. 49–54, IEEE, September 2015.

[7] M.K. Alsmadi, "A hybrid Fuzzy C-Means and Neutrosophic for jaw lesions segmentation," *Ain Shams Engineering Journal*, vol. 9, no. 4, pp. 697–706, 2018.

[8] T.M. Tuan, "Dental segmentation from X-ray images using semi-supervised fuzzy clustering with spatial constraints," *Engineering Applications of Artificial Intelligence*, vol. 59, pp. 186–195, 2017.

[9] K. Anuradha and K. Sankaranarayanan, "Detection of oral tumor based on marker-controlled watershed algorithm," *International Journal of Computer Applications*, vol. 52, no. 2, 2012.

[10] A. Agrawal and R.K. Bhogal, "A ReviewEdge Detection Techniques in Dental Images," In *International Conference on ISMAC in Computational Vision and Bio-Engineering*, pp. 1359–1378, Springer, Cham., May 2018.

[11] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," October 2015.

[12] Drozdzal, Michal, Vorontsov, Eugene, Chartrand, Gabriel, Kadoury, Samuel, & Pal, Chris, "The importance of skip connections in biomedical image segmentation," In *Deep learning and data labeling for medical applications*, pp. 179–187, Springer, 2016.

[13] G. Jader, J. Fontineli, M. Ruiz, K. Abdalla, M. Pithon, and L. Oliveira, Deep instance segmentation of teeth in panoramic X-ray images. In *2018 31st SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI)*, pp. 400–407, IEEE, October 2018.