Assessment of Segmentation Models on Panoramic Radiographic Dental Images

Apporv Upadhye[1], Peixi Liao[2], Reem Alasleh[2], Vissuta Khampatee[2], and Farshid Alizadeh-Shabdiz[1](✉)

[1] Department of Computer Science, Boston University, Boston, MA 02410, USA
alizadeh@gwu.edu

[2] Henry M. Goldman School of Dental Medicine, Boston University, Boston MA. 02410, USA

Abstract. Computer-aided diagnostics and treatment is one of the fastest-growing areas in the dental field . In this effort, dental X-ray image segmentation plays a crucial role in enabling many dental analyses and interpretations and also enables accurate analysis. Recent advancements in image segmentation have been instrumental in this effort. This paper assesses different segmentation models accuracy on dental X-ray panoramic images. Among these models, Residual UNET with binary cross-entropy achieved the best results. Despite obtaining favorable accuracy, other UNET models exhibited lower Intersection of Union (IOU) values in segmentation masks.

Keywords: Segmentation · Dental images · deep learning model · UNET

1 Introduction

Artificial Intelligence (AI) has significantly advanced dentistry. AI technology can revolutionize the dental industry by improving diagnostics, treatment planning, patient care, and practice management. While clinical dentistry is built on visional detection and judgment of normal or abnormal structures, visional detection plays a critical role in dental practice. Image segmentation is the fundamental step for developing image recognition projects. Machine learning as a general area is a branch of AI that creates models that enable computers to learn from data and make predictions. Dentistry uses it to segment dental images by analyzing large clinical datasets. It can match or even exceed human clinician performance, thanks to advanced techniques like deep learning, which play a vital role in dentistry's advancements [7].

Dental X-ray images like periapical, bitewing, and panoramic are the most common images analyzed with machine learning algorithms. Abdi AH et al. [6] automatically developed an algorithm to segment the mandible in panoramic X-rays, which achieves over 93% average performance in terms of Dice similarity, specificity, and sensitivity when evaluated on a dataset of 95 panoramic X-rays. Zhang et al. and Chen et al. conducted studies on automated detection and labeling of 2D teeth using a convolution neural network to detect teeth in periapical radiographs [7]. The experiments of

T. Zlateva and G. Tuparov (Eds.): CSECS 2024, LNICST 609, pp. 169–182, 2025.
https://doi.org/10.1007/978-3-031-84312-9_12

their work shows precision rates of 95% and 90%, respectively, which is high and helpful in diagnosis by clinicians in cases like detecting missing teeth regions for implant placement.

X-rays and image source data from dental clinics have different settings depending on the machine brand, technician preference, and patient tolerance. Therefore, preprocessing also plays a crucial role in image segmentation and helps improve the accuracy and effectiveness of the segmentation algorithm by increasing the diversity of images [8].

Augmentation increases the accuracy of image segmentation by increasing the robustness and generalization ability of the segmentation algorithm. The algorithm can learn to recognize and segment objects under various conditions and variations by generating new training samples with different variations, making it more accurate and reliable when applied to real-world images. By generating new training samples with different variations, the algorithm can better generalize to new, unseen images, reducing the risk of overfitting and improving accuracy [11].

This is the pioneer study to seek the solution of finding the best dental segmentation models and report the performance of the state-of-the-art segmentation models for clinical dental applications.

2 Training Data Description

One hundred twenty anonymous panoramic dental images extracted from CBCT are collected from the patient database of Boston University Henry M. Goldman School of Dental Medicine (IRB approval number: H-42139). A team of dentists labeled the images into six different items: teeth, root canal, restoration, maxillary sinus, bone, and inferior alveolar nerve, using Label Studio.

3 Neural Network Models

Popular deep-learning segmentation models have been selected to assess preprocessing methods. The U-NET model is one of the most popular segmentation deep learning models, which has been studied in [1, 2]. In this study, the original U-NET model (Base U-NET) and a couple of its extensions, which are UNET++, residual UNET, and inception UNET, have been considered.

The neural network model settings are as follows. The deep learning models have been implemented in TensorFlow. As the cost function, binary cross entropy has been used. Adam optimizer has been selected after testing different optimizers.

In this study, the results of the segmentation models have been given to an edge detection algorithm. By incorporating edge information into the loss function, the model is trained to produce segmentation maps that are more accurate and visually appealing. These maps have sharper and more well-defined object boundaries. In the end, the loss function that is used is as follows:

Edge-aware Focal Loss(α set to be 0.5) [10]

$$\text{loss}=(1-\alpha)\cdot\text{binary\_cross\_entropy}(\text{ytrue},\text{ypred}) \\ +\alpha\cdot\text{mean\_squared\_error}(\text{SobelEdges}(\text{ytrue}),\text{SobelEdges}(\text{ypred}))$$

Input to the neural network models is 512 × 512 gray-scaled images.

3.1 Augmentation

We have applied no image augmentation in the assessment of the models.

Applying different data augmentation techniques can improve overall performance of the model. For example, applying contrast-limited adaptive histogram equalization to the image improves the contrast and visibility of details in low-contrast images [9].

Since these effects are going to be the same across different models, the models have been assessed in absence of augmentation.

3.2 Assessment and Metrics

The following metrics have been used to assess the final results.

1. Intersection over Union (IoU) or Jaccard Index: This metric measures the overlap between the predicted segments and ground truth. It is calculated as the area of overlap between the two segmentations divided by the area of their union.
2. Pixel Accuracy: This metric calculates the percentage of pixels in the predicted segmentation correctly classified as belonging to the foreground or background. It is calculated as the ratio of correctly classified pixels to the total number of pixels in the image [3].

It is important to note that different metrics give different insights into the model's performance. Therefore, evaluating the model using multiple metrics is recommended to understand its performance comprehensively.

4 Training Images

Description and examples of images used in this study are shown in Figs. 1, 2, 3, 4, 5, 6 and 7.

Panoramic View: A panoramic X-ray view, known as a panoramic radiograph, is a type of dental X-ray that captures a broad view of the entire mouth, including the teeth, upper and lower jaws, temporomandibular joints (TMJ), and surrounding structures.

Teeth: Hard, calcified structures found in the mouths of humans and many other animals. They are primarily used for biting, chewing, and grinding food but also play an essential role in speech and the face's overall appearance. Teeth are classified into four main types: incisors, canines, premolars, and molars. Teeth comprise several layers, including enamel, dentin, and pulp.

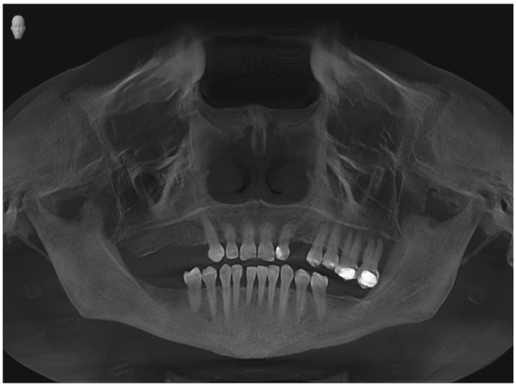

Fig. 1. Panoramic Image

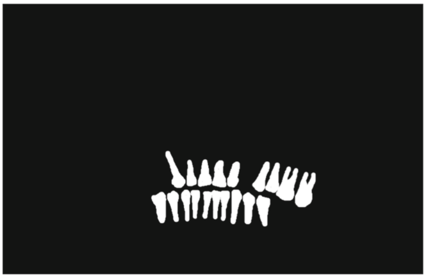

Fig. 2. Teeth Segment

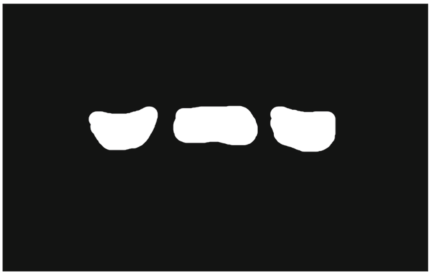

Fig. 3. Sinus Segment

Sinus: In dental terms, the sinus is a hollow cavity in the facial bones surrounding the nose, known as the maxillary sinus. The maxillary sinus is the most prominent paranasal sinus and is located above the upper teeth in the back of the mouth.

Restoration: Any dental procedure that repairs or replaces missing or damaged teeth or oral structures. Restorations can be used to restore the function, strength, and appearance of teeth damaged due to decay, trauma, or other factors. By segmenting the restoration process, dental professionals can carefully plan and prioritize each restoration procedure

Fig. 4. Restoration Segment

to ensure the patient receives the most effective and efficient treatment. This can minimize the number of dental appointments required, reduce overall treatment time, and maximize the success and longevity of each restoration.

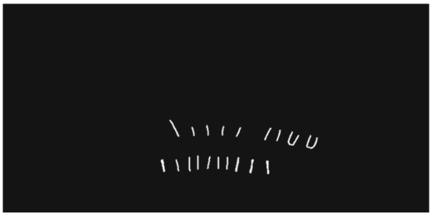

Fig. 5. Root Segment

Root: In dental terms, a root in a dental X-ray refers to the part of a tooth located below the gum line and anchors the tooth in the jawbone. The root of a tooth can vary in shape and size depending on the type of tooth and the individual patient but typically consists of one or more narrow, cone-shaped channels called root canals [4].

Fig. 6. Major Nerve Segment

Major Nerve: In dental terms, the central nerve in an X-ray refers to the inferior alveolar nerve, which is a branch of the mandibular nerve and provides sensation to the lower teeth, gums, and lips. The inferior alveolar nerve runs through a canal in the jawbone called the mandibular canal, which is visible on dental X-rays as a thin, dark line or shadow [5].

Fig. 7. Bone Segment

Bone: In dental terms, bone in an X-ray refers to the alveolar bone, the specialized bone surrounding and supporting the teeth. The alveolar bone forms the sockets, or alveoli, in which the tooth roots are embedded and help to anchor the teeth firmly in place.

5 Test Results

The results of the neural network models have been presented in this section.

As the baseline, the accuracy of different neural network models with no augmentation has been measured and captured in the tables in this section (Table 1, Fig. 8, Table 2, Fig. 9).

Table 1. Training Data accuracy of different neural network models

Section	Teeth	Sinus	Root	Restoration	Nerve	Bone
Base UNET Accuracy	94.76654	90.91201	97.1778	97.97836	97.07918	91.20427
Inception Unet Accuracy	95.45506	92.60849	97.45686	98.92324	97.21782	91.55782
UNET++ Accuracy	93.76233	90.7336	96.1383	98.41879	96.9121	84.0106
Residual Unet Accuracy	96.78777	97.16851	98.6734	99.08474	97.23479	95.04914
Base UNET IOU	78.37333	49.17806	67.64625	80.3425	45.57964	76.18763
Inception Unet IOU	80.84756	54.21964	71.17388	84.83342	49.34683	76.05482
UNET++ IOU	71.1927	42.0675	58.1953	77.54502	44.3472	57.6547
Residual Unet IOU	87.10189	78.85537	83.7476	89.08599	49.70025	85.2447

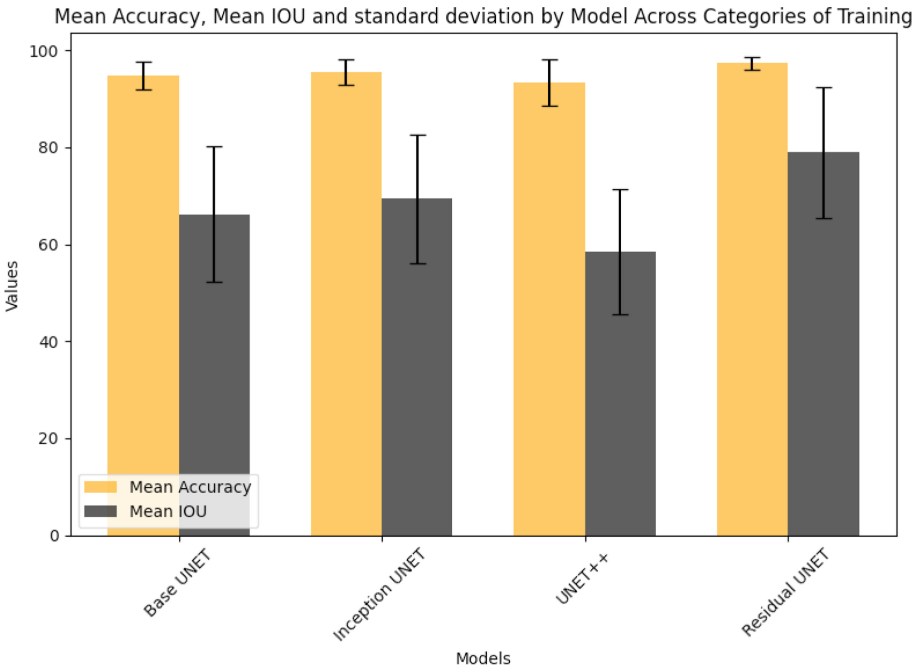

Fig. 8. Training of Mean Accuracy, Mean IOU with standard deviation grouped by model across all the categories

Table 2. Validation of Data accuracy of different neural network models.

Section	Teeth	Sinus	Root	Restoration	Nerve	Bone
Base UNET Accuracy	95.68249	91.20748	97.29434	98.15587	97.11299	90.69254
Inception Unet Accuracy	95.39721	92.53312	97.30266	98.95716	97.27156	89.61229
UNET++ Accuracy	89.216	91.07114	96.17388	98.08631	96.95522	71.96896
Residual Unet Accuracy	96.24427	94.67407	97.58293	99.0468	97.23754	91.10404
Base UNET IOU	83.19162	49.87472	68.77294	82.03536	44.6103	75.01012
Inception Unet IOU	80.13799	53.05991	70.00132	85.78766	49.59429	72.87867
UNET++ IOU	63.38982	49.36427	59.60315	81.65454	44.39679	24.94519
Residual Unet IOU	84.60757	64.87876	73.1733	88.79906	48.80972	75.70124

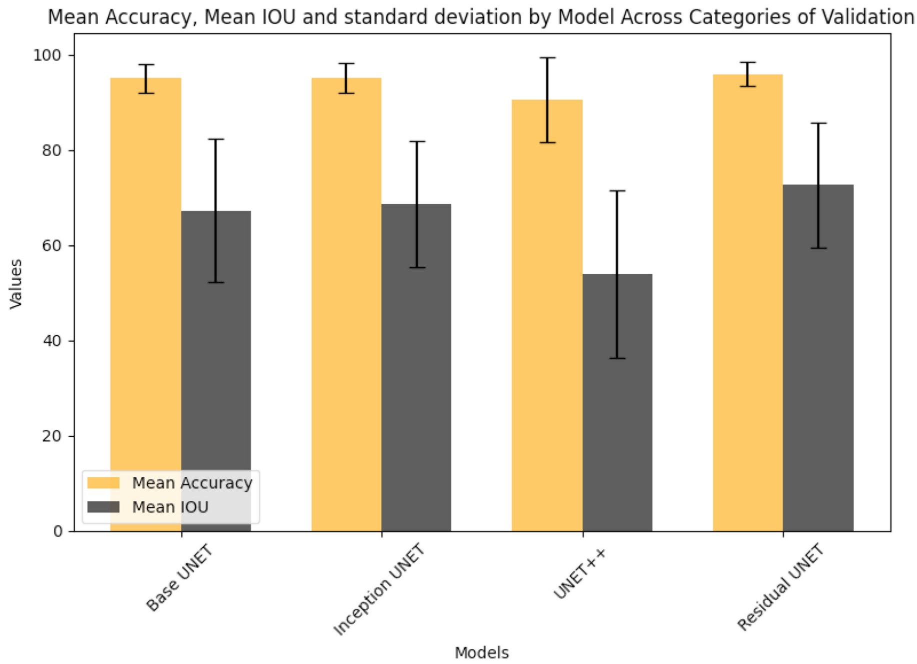

Fig. 9. Validation of Mean Accuracy, Mean IOU with standard deviation grouped by model across all the categories

Samples of segmentation results:
The provided tables offer an insight into the performance of various network architectures across distinct dental segmentation sections, explicitly focusing on accuracy and Intersection over Union (IOU) scores. A comprehensive analysis reveals the following key findings:

Accuracy Scores
Residual UNET consistently emerged as the architecture with the highest accuracy scores across the different dental sections. Notably, for sections like "Root" and "Restoration," Residual UNET achieved remarkable accuracy levels of 97.58% and 99.05%, respectively. This robust accuracy performance underscores Residual UNET's suitability for intricate segmentation tasks.

IOU Scores
Regarding IOU scores, Residual UNET again demonstrated noteworthy performance, particularly excelling in the "Restoration" section with an impressive IOU score of 88.80%. This emphasizes the model's ability to capture boundaries and contours precisely, resulting in highly accurate segmentation.

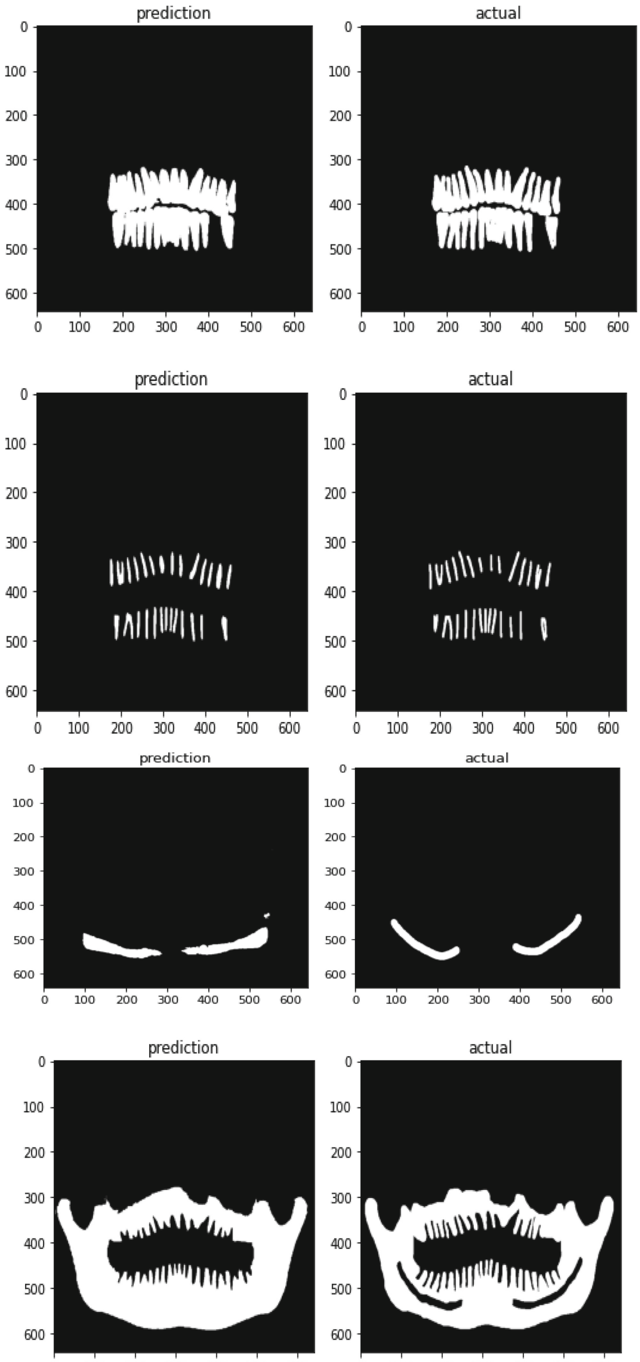

Fig. 10. Masks vs predicted Masks

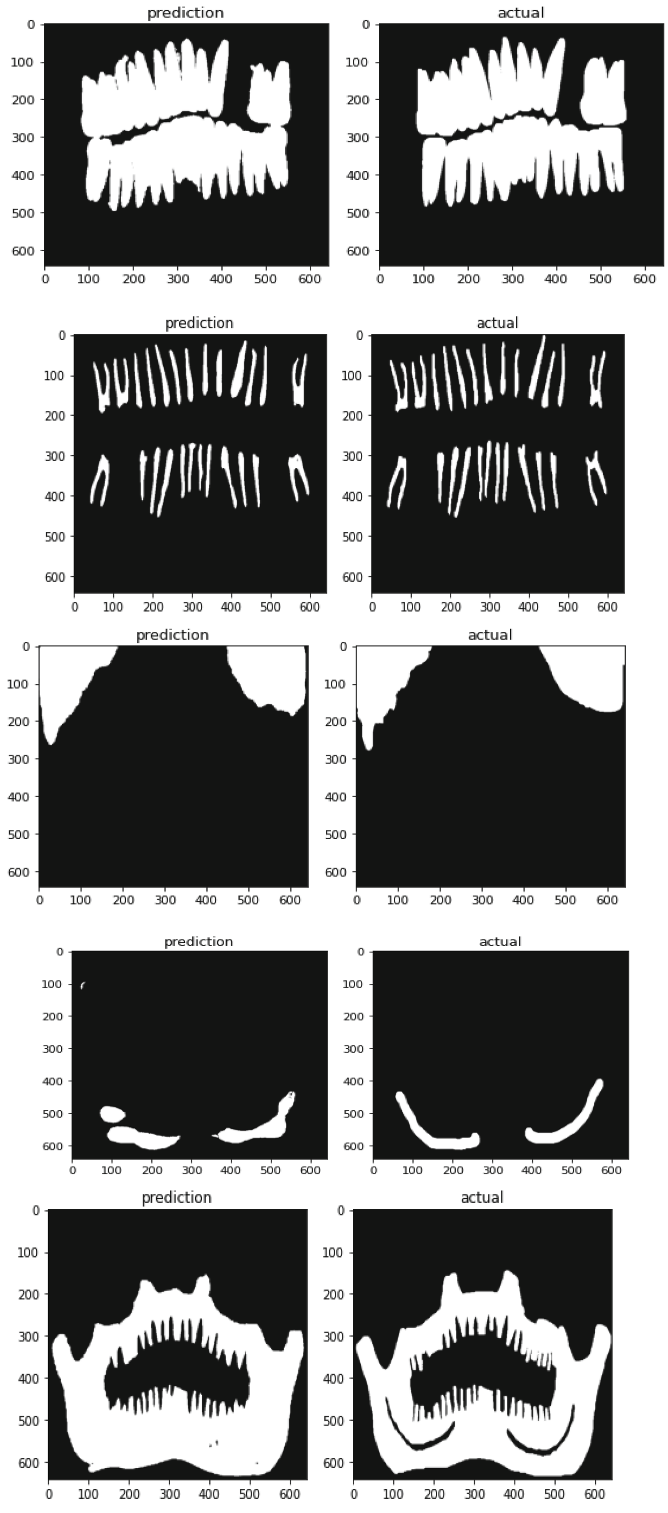

Fig. 10. (*continued*)

In conclusion, Residual UNET consistently showcased superior accuracy and IOU scores across various dental sections, highlighting its proficiency in intricate dental segmentation tasks. These findings emphasize the efficacy of Residual UNET architecture in producing accurate and precise segmentations across diverse dental structures, solidifying its position as the preferred choice for achieving high-quality dental image analysis outcomes (Fig. 10).

6 Review of Previous Work

Panoramic image has been widely used in previous studies, and an accuracy of about 90% precision has been reported. Park et al. [1] used automated machine learning methods for the entire process from annotation to segmentation, and finally detection of missing tooth regions. They trained and tested the model with 455 panoramic radiographic images. Their model identifies the position and number of missing teeth in the panoramic radiographic image with 92.14% mean Average Precision (mAP) for tooth segmentation and 59.09% mAP for missing tooth detection [1]. Zhang et al. [3] and Chen et al. [4] study on automated detection and labeling of 2D teeth using a CNN – (convolution neural network) to detect teeth in radiographs of periapical. The results of two studies show precision rate of 95% and 90%, respectively. The achieved accuracy is high enough to assist clinicians for implant placement [1].

Multiple machine-learning models have been tested for dental segmentation in the panoramic image. Particularly, Mask R–CNN-based models are tried by many research teams for segmenting and identifying teeth (Koch et al. [16], Oktay et al. [17]). The [18] built an end-to-end deep learning architecture for teeth segmentation and numbering based on Mask R–CNN and PointRend. Zhao et al. [19] took different approach and built a two-staged attention segmentation network (TSASNet) to localize and classify teeth in dental x-rays. The tooth borders were recognized using a fully convolutional network with an accuracy of 96.94%. Moreover, Jader et al. [20] examined Mask R–CNN to segment teeth in cases of difficult panoramic radiographs.

Wirtz et al. [15] proposed a coupled-shaped model for tooth segmentation in low-quality panoramic radiographs. Lee et al. [21] tried an Mask-R-CNN model to automate teeth segmentation on dental panoramic images. Silva et al. [22] evaluated four neural network models, Mask R-CNN, HTC, PANet, and ResNet for tooth numbering and segmentation on dental radiographs. According to Silva et al. [22] PANet demonstrated superior accuracy in this paper.

One of the most popular segmenation models is U-Net, which was introduced in the paper "U-Net: Convolutional Networks for Biomedical Image Segmentation" by Olaf Ronneberger, et al. in 2015 [14]. U-Net achieves state-of-the-art results on various biomedical image segmentation tasks.

7 Conclusion

7.1 Best Models

In this research, Dental segmentation is carried out using various image segmentation models.

Residual- UNET provides the best results using binary cross-entropy. After experimenting with various UNET models, we received good accuracy but a low Intersection of Union values for all our segmentation masks.

7.2 Best Loss Function

In dental segmentation, our findings have revealed that employing a binary cross-entropy loss function consistently yields superior results compared to utilizing edge-aware loss functions. This observation can be attributed to several factors intrinsic to the nature of dental imaging and the characteristics of the segmentation task.

Firstly, dental images often present well-defined and distinct structures, such as teeth, against a relatively uniform background. The binary cross-entropy loss function, tailored for binary segmentation tasks, distinguishes between foreground structures and the background due to its intrinsic ability to accurately measure pixel-wise differences. In scenarios where the primary objective is to identify and isolate specific dental structures accurately, the inherent simplicity and effectiveness of the binary cross-entropy loss function prove advantageous.

Furthermore, the binary cross-entropy loss function's computational efficiency and straightforward nature enable efficient convergence during training. Given that dental segmentation tasks can be sensitive to dataset size and quality, the binary cross-entropy loss function's capacity to offer stable and consistent outcomes, even in scenarios with limited data or noise, becomes particularly advantageous. This robustness is essential for generating reliable segmentation results, especially when dealing with dental images that might exhibit variations in lighting conditions, imaging quality, and anatomical structures.

Additionally, the dental segmentation task may rely on something other than preserving intricate edge details. Dental structures, while integral to the analysis, may possess distinct boundaries that the binary cross-entropy loss function effectively captures. In such cases, the demand for sophisticated edge-aware loss functions designed to enhance boundary detection might need to be more pronounced.

Ultimately, choosing an appropriate loss function depends on the specific characteristics of the segmentation task and the nature of the analyzed images. **The favorable performance of the binary cross-entropy loss function** in dental segmentation can be attributed to its alignment with the task's simplicity, the well-defined nature of dental structures, and the ability to provide stable outcomes under varying data conditions. These findings underscore the significance of tailoring loss function selection to the unique attributes of the segmentation problem, leading to enhanced accuracy and efficiency in dental image analysis.

References

1. Park, J.M., Lee, J., Moon, S.Y., Lee, K.: Deep learning based detection of missing tooth regions for dental implant planning in panoramic radiographic images. Appl. Sci. (2022). https://doi.org/10.3390/app12031595
2. Types of Contrast Enhancement Algorithms and Implementation in Python. https://iq.opengenus.org/contrast-enhancement-algorithms/

3. Instance segmentation quality metrics - MATLAB - MathWorks Australia. https://au.mathworks.com/help/vision/ref/instancesegmentationmetrics.html
4. dental anatomy – Dental illustrations and videos for your clinic. https://alexmit.com/tag/dental-anatomy/
5. Rice, M.: Regional nerve blocks for equine dentistry. J. Vet. Dent. **2017** (2017). https://doi.org/10.1177/0898756417713730
6. Abdi, A.H., Kasaei, S., Mehdizadeh, M.: Automatic segmentation of mandible in panoramic x-ray. J. Med. Imaging (Bellingham, Wash.) **2**(4), 044003. https://doi.org/10.1117/1.JMI.2.4.044003 (2015)
7. Chen, Y., Wang, Z., Shen, W., Liu, Y., Zhang, Y.: A deep learning approach to automatic teeth detection and numbering based on object detection in dental periapical films. Dent. Mater. **36**(1), 65–74 (2020)
8. Jeyavathana, R.B., Balasubramanian, R., Pandian, A.A.: A survey: analysis on pre-processing and segmentation techniques for medical images. Int. J. Res. Sci. Eng. **5**(2), 113–120 (2019)
9. Yadav, G., Maheshwari, S., Agarwal, A.: Contrast limited adaptive histogram equalization based enhancement for real-time video system. In: 2014 International Conference on Advances in Computing, Communications and Informatics (ICACCI), pp. 2392–2397. Delhi, India (2014). https://doi.org/10.1109/ICACCI.2014.6968381
10. Xie, S., Tu, Z.: Holistically-nested edge detection. In: 2015 IEEE International Conference on Computer Vision (ICCV), pp. 1395–1403. Santiago, Chile (2015). https://doi.org/10.1109/ICCV.2015.164
11. Shorten, C., Khoshgoftaar, T.M.: A survey on image data augmentation for deep learning. J. Big Data **6**(1), 1–48 (2019). https://doi.org/10.1186/s40537-019-0197-0
12. Buslaev, A., Iglovikov, V.I., Khvedchenya, E., Parinov, A., Druzhinin, M., Kalinin, A.A.: Albumentations: fast and flexible image augmentations. Information **11**(2), 125 (2020). https://doi.org/10.3390/info11020125
13. Chen, L.-C., Papandreou, G., Kokkinos, I., Murphy, K., Yuille, A.L.: Classification with an edge: Improving semantic image segmentation with boundary detection. arXiv preprint arXiv:1612.01337 (2016)
14. Ronneberger, O., Fischer, P., Brox, T.: U-Net: convolutional networks for biomedical image segmentation. In: Navab, N., Hornegger, J., Wells, W., Frangi, A. (eds.) Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015. MICCAI 2015. Lecture Notes in Computer Science(), vol. 9351, pp. 234–241. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-24574-4_28
15. Wirtz, A., Mirashi, S.G., Wesarg, S.: Automatic teeth segmentation in panoramic x-ray images using a coupled shape model in combination with a neural network. In: Frangi, A., Schnabel, J., Davatzikos, C., Alberola-López, C., Fichtinger, G. (eds.) Medical Image Computing and Computer Assisted Intervention – MICCAI 2018. MICCAI 2018. Lecture Notes in Computer Science(), vol. 11073, pp. 712–719. Springer, Cham (2018). https://doi.org/10.1007/978-3-030-00937-3_81
16. Koch, G., Staab, T., Spitzer, M.: Accurate tooth segmentation in panoramic radiographs using fully convolutional neural networks. In: International Conference on Medical Image Computing and Computer-Assisted Intervention, pp. 703–711. Springer, Cham (2018)
17. Oktay, O., et al.: Concurrent detection, segmentation, and counting of teeth in panoramic X-ray images. Med. Image Anal. **54**, 31–44 (2019)
18. Pinheiro, J.P., Vieira, A.X., Silva, A.C., de Mendonça, F.M., Campilho, A.: Deciduous teeth segmentation and numbering using Mask R-CNN and PointRend. arXiv preprint arXiv:2110.08959 (2021)
19. Zhao, H., Gu, Y., Li, X., Fu, Y.: TSASNet: Two-stage attention segmentation network for tooth segmentation on dental panoramic X-ray images. arXiv preprint arXiv:2103.16012 (2021)

20. Jader, B.F., de Mendonça, F.M., Campilho, A., Silva, A.C.: Deep instance segmentation of teeth in panoramic X-ray images. Med. Image Anal. **54**, 1–12 (2018)
21. Lee, S.H., Kim, H.E., Jung, K.H., Lee, S.K.: Automated teeth segmentation on dental panoramic images using Mask R-CNN. J. Digit. Imaging **33**(5), 1460–1468 (2020)
22. Silva, B.F., Pinheiro, J.P., Oliveira, L.F., Campilho, A.: A study on tooth segmentation and numbering using end-to-end deep neural networks. In: 2020 IEEE 33rd SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI), pp. 1–8 (2020)