



Research Article

Optimized IANSegNet: Deep Segmentation for the Detection of Inferior Alveolar Nerve Canal

V. Gokula Krishnan ¹, M. Navaneethakrishnan,² Sangeetha Ganesan,³
M. V. Vijaya Saradhi,⁴ K. Hemapriya,⁵ D. Selvaraj,⁶ J. Deepa,⁷ K. Sreerama Murthy,⁸
and Srinath Doss ⁹

¹Department of CSE, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Thandalam, Chennai, Tamil Nadu, India

²CSE Department, St. Joseph College of Engineering, Sriperumbudur, Chennai, Tamil Nadu, India

³AIDS Department, RMK College of Engineering and Technology, Kavaraipettai, Tamil Nadu, India

⁴CSE Department, ACE Engineering College, Ghatkesar, Hyderabad, Telangana, India

⁵CSE Department, Panimalar Engineering College, Chennai, Tamil Nadu, India

⁶Department of ECE, Panimalar Engineering College, Chennai, Tamil Nadu, India

⁷CSE Department, Easwari Engineering College, Ramapuram, Chennai, Tamil Nadu, India

⁸IT Department, Sreenidhi Institute of Science and Technology, Ghatkesar, Hyderabad, Telangana, India

⁹Lesotho Campus, Botho University, Lesotho

Correspondence should be addressed to V. Gokula Krishnan; gokul_kris143@yahoo.com

Received 30 September 2022; Revised 10 November 2022; Accepted 3 April 2023; Published 28 April 2023

Academic Editor: Abid Yahya

Copyright © 2023 V. Gokula Krishnan et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Imaging studies in dentistry and maxillofacial pathology have recently concentrated on detecting the inferior alveolar nerve (IAN) canal. In spite of the minor dimensions of 3D maxillofacial datasets, deep learning-based algorithms have shown encouraging consequences in this study area. This study describes a mandibular cone-beam CT (CBCT) dataset with 2D and 3D hand comments. It is huge and freely available. It was possible to utilise this dataset by applying the residual neural network (IANSegNet), which consumed less GPU memory and computational complexity. As an encoder, IANSegNet uses the computationally efficient 3D ShuffleNetV2 network to reduce graphics processing unit (GPU) memory usage and improve efficiency. After that, a decoder with leftover blocks is added to keep the quality high. To address network convergence and data inequity, Dice's loss and cross-entropy loss were created. Optimized postprocessing techniques are also recommended for fine-tuning the coarse segmentation findings that are generated by IANSegNet. The results of the validation show that IANSegNet outperformed other deep learning models in a variety of criteria.

1. Introduction

Deep learning has recently been used to automate several sectors [1], and the field is fast expanding [2]. When it comes to medical applications, deep learning [3] has proven to be extremely effective [4]. Magnetic resonance imaging (MRI) and computerized tomography (CT) images, as well as electroencephalogram (EEGs) [5], can be utilised to forecast and diagnose diseases using deep learning [6]. In dentistry, it

can be used to automatically diagnose a variety of disorders [7]. Cone-beam computed tomography (CBCT) images have been used to classify cystic lesions, and teeth have been used to estimate age [8]. Third molars are common in the mouths of most people; however, they can be removed for a variety of reasons. In oral and maxillofacial operation, the removal of third molars is a shared process [9].

After the third molar is extracted, symptoms can appear in anywhere from 30 to 68% of cases, depending on the

impaction type of the tooth [10]. Impaction patterns of mandibular third molars vary because they grow in a variety of locations and directions [11]. Consequently, it is essential to regulate the mandibular third molar impaction pattern prior to the procedure to select the most appropriate surgical technique [12]. Third molar impaction can be defined using a variety of criteria [13]. After extraction, a variety of issues can arise in certain impaction patterns [14].

1.1. Problem Statement. Complications may emerge following the extraction depending on the impaction pattern [15]. Damage to the mediocre alveolar nerve [16] is the most shared complication following the removal of the mandibular third tooth (IAN) [17]. It is crucial to know what to expect following an IAN injury because of the varied results reported [18] on the surgical therapy options [19]. Because of the recognised relationship between the molar and the IAN, the risk of IAN injury increases [20].

1.2. Contribution. In this research, IANSegNet is presented to improve segmentation accuracy and reduce computing complexity. IANSegNet's efficiency is initially improved by using 3D ShuffleNetV2, a new efficient CNN, as an encoder. We also use the Res-decoder to keep the neural network from degrading when it is very deep. For the third time, we have devised a fusion loss function that combines entropy with Dice to ease meeting concerns and keep too many negative samples from swamping the good ones. In addition, we use 3D patches to apply the augmentation during testing and mirror them. Additional postprocessing techniques are employed to enhance performance, where threshold values are selected by proposed optimization model (spotted hyena optimizer algorithm (SHA)). By delivering both the uppermost segmentation accuracy and the lowest computational difficulty, the suggested method exceeds the state-of-the-art styles.

2. Related Works

The curved MPR image set was recently developed by Wei and Wang [21]. One-pixel sampling is used for the MPR and total regular intensity forecast panoramic picture, and K -means are used to cluster texture characteristics from the grey level-gradient cooccurrence matrix of the region of interest to recover image difference in the IAN canal. The fourth-order polynomial is used to fit the results of the final segmentation of the canal margins in 2D line tracking. Jaskari et al. [22] trained a complete network using a dataset of 3D scans with coarse annotations to segment the mandibular canal for the first time using deep learning. Using an average diameter of 3.0 millimetres, a volume of each canal is expected to be created by interpolating 10 control points in a spline. However, the lack of hand-annotated voxels and the poor quality of segmentation limit the approach's effectiveness in comparison to previous approaches based on SSM. Cipriano et al. [23] provided publicly available CBCT (3-dimensional dataset with expertly generated 3D annotations for download). Then, using a freshly trained architecture, it improved the state-of-the-art mandibular canal segmenta-

tion accuracy. On the basis of the U-Net, a 3D convolutional neural network (CNN) is used to segment images in 3D feature maps that are compressed in the network's contractive path using stride 2 convolutions, whereas they are expanded in the expanding path using stride 2 transposes. Cipriano et al. [23] describe a unique label propagation method, based on deep learning, which can transform sparse 2D labels into 3D voxel-level annotations. This method can cover the gap between the most contemporary and advanced methods for 3D segmentation and the absence of viable annotated data in the maxillofacial field. Moreover, with the purpose of pushing the state of the art in 3D IAN segmentation, a novel 3D segmentation CNN is built that uses positional information to generate the final 3D prediction.

Qi et al. [24] raised the possibility of IAN injury during the removal of impacted lower third molars. A total of 200 wedged lower third molars were found around the IAN. Data from CBCT was utilised to divide the four categories into AR, LE, and AE: apical region of the root, lateral region of tapering root, and area adjacent to root. Surgeons used a tooth sectioning technique to extract all of the patient's teeth, relocating them down the root's long axis or arc. IAN's postoperative neurosensory impairment was the most significant outcome variable. X2 testing was performed to examine the variations in postoperative IAN damage between the two categories. Using a retrospective evaluation of panoramic and cone-beam computed tomography images of two hundred mandibular third molars, Tassoker [25] conducted a study to examine the reliability of a panoramic view to detect the diversion of the inferior alveolar canal. Data was analysed to see whether or not there was a correlation between the panoramic view and cone-beam computed tomography findings based on the mandibular canal wall interruption, darkening of roots, canal diversion, and narrowing of the mandibular canal in the panoramic pictures. On cone-beam computed tomographic images, there was no canal cortication in 136 cases, which was most strongly associated with mandibular canal diversion (96 percent) and least strongly associated with mandibular canal wall interruption (65 percent). Reexposing twenty-five patients to cone-beam computed tomography allowed Pandey et al. [26] to evaluate the accuracy of radiographic signals in a panoramic perspective. 63.8% were discovered to be in direct touch with the mandibular canal's superior border, resulting in damage to the cortical layer. For teeth with two or more radiological signals on panoramic view, the canal was either primarily buccal (61.7 percent) or followed by inferior (23.4 percent) in relation to lower 3rd molar root in cone-beam computed tomography (CBCT). Cone-beam computed tomography is always recommended when there are two or more indicators of mandibular canal injury, they concluded. The affected lower third molar root end and the inferior alveolar canal were assessed and compared using panoramic radiography and cone-beam computed tomography in a 40-sample study by Nayak et al. [27]. Cone-beam computed tomography revealed a real association between 23 of the roots in the panoramic view and darkening and constriction of the canal. Cone-beam computed

tomography, on the other hand, may often predict a real association when any of the radiographic signals are present, although panoramic radiography cannot. Twenty-three individuals with totally impacted teeth were studied for their ability to predict inferior alveolar nerve injury following lower 3rd molar extraction using panoramic radiography and computed tomography. There has been a perfect correlation found between the development of paraesthesia after mandibular nerve damage and a panoramic view based on prevision. After impacted lower third molar surgery, the study panoramic is a first-level diagnostic exam that gives sufficient information to forecast inferior alveolar canal damage lesions [28].

3. Proposed System

Patches are initially preprocessed and augmented with data during training in this study. Second, for IAN segmentation, we introduce IANSegNet, a powerful residual neural network. Finally, in order to maintain training stability and minimise the influence of unbalanced data, we suggest a loss function that combines a Dice with a cross loss. IANSegNet's segmentation results can be improved by a simple and effective postprocessing procedure.

3.1. Dataset. The Affidea facility in Modena, Italy, has provided us with the 3D CBCT capacities that make up our dataset. Rehabilitation, cancer detection and treatment, and advanced diagnostics are only few of Affidea's specialties. It has 312 facilities in 15 countries, with a workforce of approximately 11,000 people.

Cone-beam CT dental scans are included in the collection. Here's a download link for the dataset if you would want to take a look at it. Interslice and interpixel distances are always the same 0.3 millimetre. Between 1 000 and 5 264, data capacities have already been converted to the Hounsfield unit (HU). We also made sure that the window width and centre were properly processed in accordance with the DICOM protocol during the conversion to HU (148, 265 312 to 178, 423 463) for the X-, Y-, and Z-axes of volume. We only had access to a few personal details about each patient because they had been anonymised, such as gender, age, and the year of the scan. All scans were performed between 2019 and 2020; 59% of the patients were female.

The dataset can be used for panoramic views in both 2D and 3D. It is, however, a 2D model that is used in the suggested study. Figure 1(a) is an axial slice derived from the CT capacity, and Figure 1(b) is a 2D panoramic image. The jawbone can be recognised by the red contour known as the panoramic base curve. Using the CT-panoramic volume's perspective, Figure 1(b) provides an orthogonal image of the curved formed by curve. Figure 1(c) shows an experienced technician's hand annotation of the IAN, which is identical to the view in Figure 1(b).

3.2. Data Augmentation and Preprocessing. For CNN training, normalisation is essential since it speeds up training and prevents overfitting. For CBCTs, the significance of

CBCT intensity varies among images because of the variety of modalities, patients, and devices. Since the IAN region is first cropped, we then compute mean and SD for each modality just in the IAN region in every patient to normalise images to a consistent scale. There are many ways that IAN-SegNet can learn and generalise, such as gamma correction and random.

3.3. Proposed IANSegNet

3.3.1. Encoder Based on 3D ShuffleNetV2. Semantic segmentation, picture rebuilding, and image classification are all possible with encoder, a feature extractor. VGG16 [29] not only had significant achievement classification but also produced remarkable presentation in medical picture segmentation, which is an effective practice in deep learning deep encoders like deeplabV3 [30], deeplabV3+ [31], ResNets [28], or Xception [32] and has recently been shown to advance the results of semantic segmentation. A deep encoder is used to segment 3D IAN in this way. Our encoder uses 3D ShuffleNetV2 [33], an extremely deep but efficient network, because 3D operations typically take a long time and utilise a lot of GPU memory. As compared to other deep networks, the encoder has a smaller number of floating point operations per second (FLOPS) and parameters.

For a 3D convolutional layer with no calculation for nonlinear functions, the FLOPS are calculated as follows:

$$\text{FLOPS} = 2 \times x \times y \times z \times (a_{\text{in}} \times B^3 + 1) a_{\text{out}}, \quad (1)$$

where the output feature maps, depth, height, and width are denoted by the letters x , y , and z ; a_{in} and a_{out} denote the sum of input and output channels, correspondingly; and the kernel size is B .

A 3D convolutional layer has a fixed number of parameters.

$$\text{Num}_{\text{Parameters}} = (a_{\text{in}} \times B^3 + 1) a_{\text{out}}. \quad (2)$$

More advanced and efficient than ShuffleNetV1, ShuffleNetV2 is a descendant of ShuffleNetV1. As a result of the work presented in this paper, 3D IANSegNet can now use it as an encoder. Figure 2 shows the encoder's shuffle components (see Figure 3), identifying the most important structures for minimising the amount of encoder parameters. Because of the depth-wise separable convolutions performed on the shuffle unit, ShuffleNetV2 and other networks have higher FLOPS and parameter gaps. Operations on each of the input channels are utilised to perform depth-wise separable convolutions in order to fuse the maps. If the given input size a_{in} , a , y , z , and output size are the same, a typical convolution requires the parameters R G FPLOPS and S M.

$$R = 2 \times x \times y \times z \times (a_{\text{in}} \times B^3 +) \times a_{\text{out}} \times 10^{-9} G, \quad (3)$$

$$S = (a_{\text{in}} \times B^3 + 1) \times a_{\text{out}} \times 10^{-6}. \quad (4)$$

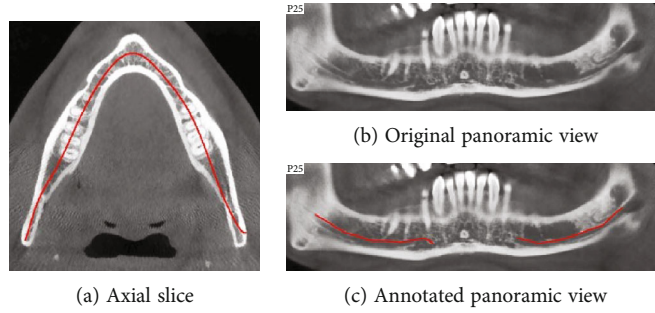


FIGURE 1: CT annotation based on 2D panoramic views.

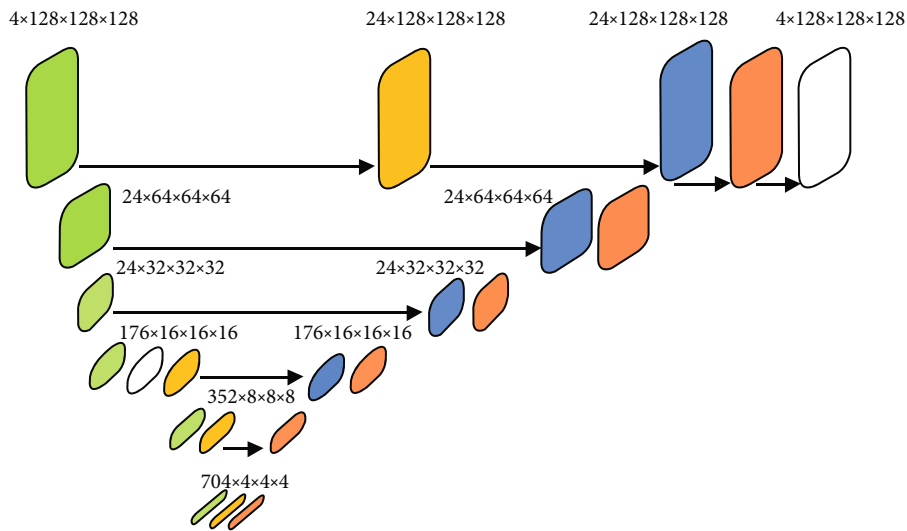


FIGURE 2: The proposed IANSegNet architecture.

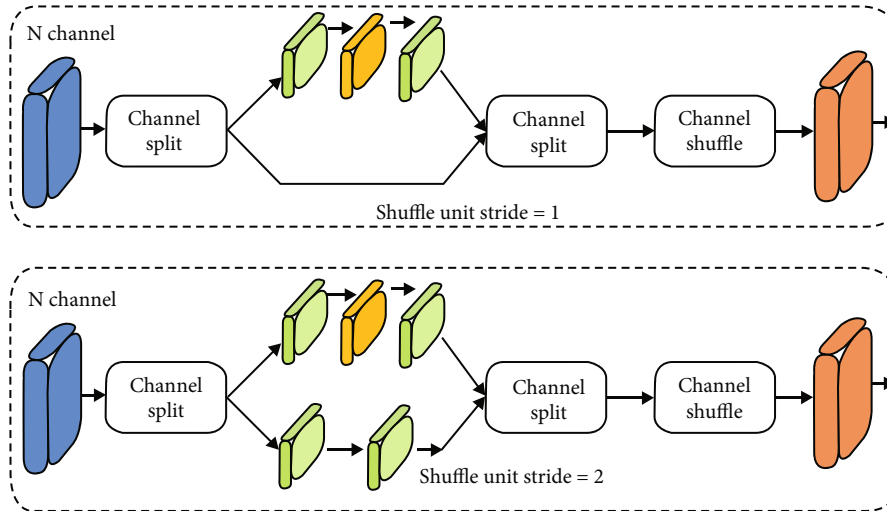


FIGURE 3: Shuffle unit.

Only R' G FLOPS and S' M parameters are needed for depth-separable convolutions, though.

$$R' = 2 \times x \times y \times z \times (a_{in} a_{out} + a_{in} \times B^3 + a_{out} + 1) \times 10^{-9}, \quad (5)$$

$$S' = (a_{in} \times B^3 + 1) + (a_{in} + 1) \times a_{out} \times 10^{-6}. \quad (6)$$

It has a significant impact on efficiency while also conserving GPU memory.

In addition, element-wise actions such as division sum, depth-wise convolutions, and batch normalisation take a

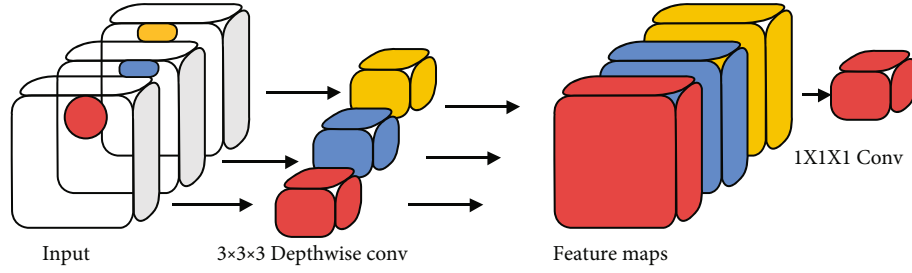


FIGURE 4: Depth-wise separable convolutions.

long time. Shuffling unit introduces “channel split” operation, which divides in order to decrease element-wise operations. Finally, by utilising “concatenate” instead of summation, it substantially minimises the element-wise processes. “Channel shuffle” is used after “concatenate” in order to improve data transmission between the various channels. Also depicted in Figure 3 is “channel shuffle,” which occurs when a convolution layer’s feature maps comprise n channels, and $c = d$. Transpose and flatten operations are used to transform the channel dimension of the feature maps into cd . IANSegNet’s encoder is completed with the last convolutional layer and the shuffle unit. In order to get a large enough receptive field for segmentation, the encoder’s output size is $1/32$ of the unique image’s size. Figure 4 depicts the convolutions that can be separated by depth.

Relative architecture-based shallow decoder has been shown to be sufficient to provide outstanding presentation in a deep network; hence, this paper uses a shallow decoder for IAN segmentation. Transposed convolution, as shown in Figure 2, is used to increase the number of feature maps by stride 2 and to reduce the sum of feature maps by half. Concatenating high-resolution feature maps from encoders in each spatial level is an effective way to compensate for the loss of information after down sampling (see Figure 2). A further issue with the deep network is that its accuracy is prone to saturation. The problem persists since the encoder-whole decoder’s design is based on a deep neural network. Thus, the use of residual blocks in decoders helps to relieve the difficulties of convergence (see Figure 2). A residual block and skip connection in the last spatial level is followed by a residual block again. Convolution with four channels and a stride of one completes the decoder’s output layer.

It is possible to think about the remaining block as

$$\text{Outputs} = f(\text{Input}, t) + \text{Input}. \quad (7)$$

The outstanding value to be erudite by the network is denoted by $f()$, and t indicates the weights of a convolution layer. It is a lot like the identification map’s formulation.

3.4. Fusion Loss Function Based on Cross-Entropy and Dice. The fundamental problem with the CBCT dataset is that only 2% of the IAN is accounted for in most situations. In addition, the loss functions employed to handle the imbalanced situation have convergence concerns. Thus, we suggest a loss function that combines Dice and cross-entropy

to produce an effective fusion loss function. As an example of a data loss function, Dice can be used to lessen the impact of data imbalances, such as in 3D segmentation. IAN segmentation relies heavily on Dice as a measure. IANSegNet has introduced a multiclass Dice as a result. IANSegNet output is normalised before Dice loss using the softmax method. The remaining value to be erudite by the neural network is denoted by $f()$ in the definition of multiclass Dice. It is a lot like with the identification map’s formulation

$$L_{\text{Dice}} = -\frac{2}{|C|} \sum_{j \in C} \frac{\sum_i s_i^j s_i^j}{\sum_i s_i^j + \sum_i g_i^j}. \quad (8)$$

IANSegNet’s likelihood of acquiring softmax is s . With a one-hot encoding, g is the actual truth. The training patch’s voxel count is indicated by the letter i . In the CBCT data, class j is one of the C classes.

It is easy to run into convergence problems when your Dice loss data is so unbalancing. To get Dice’s differentiable form, $2sg/(s+g)$, we first need to simplify its formula to $2g^2/(s+g)^2$. When s and g are small, a greater gradient will be obtained, which will lead to unstable training. As a result, we implement cross-entropy loss in order to counteract this effect. IANSegNet’s total fusion loss can be calculated by taking the sum of Dice and cross-entropy into account.

$$\text{Fusion loss} = -\frac{1}{n} \sum_{i \in n} \sum_{j \in C} g_i^j \log \left(s_i^j \right) \frac{2}{|C|} \sum_{j \in C} \frac{\sum_i s_i^j s_i^j}{\sum_i s_i^j + \sum_i g_i^j}. \quad (9)$$

The patch’s size, in this case, is n .

3.5. Postprocessing. Postprocessing methods have been shown to enhance the presentation of neural networks in numerous studies. There are a number of examples of automated methods that use K -means clustering [34]. Consequently, to fine-tune the segmentation outcome, a post-processing technique is required. Edema can be mistaken for necrotic and nonenhancing tumour cores, and there are some false positives for tumour enhancement. Therefore, the author of [35] devised a postprocessing technique based on the neural network’s properties and the distribution of tumours. Increasing the number of tumour core positives and eliminating false positives is the goal. A two-step segmentation model is developed in this study, and postprocessing is used to improve segmentation accuracy.

Input: Predictions of neural networks (**Pre**), probability maps of neural networks (**ProMap**), and the threshold for comparison (**T**)

Output: Final predictions (**FPre**)

1. $Voxels_{mandibular\ third\ molars} \leftarrow extract_voxels_in_mandibular\ third\ molars(Pre)$
2. $TC_{pro} \leftarrow get_probability_in_TC(Voxels_{mandibular\ third\ molars}, ProMap)$
3. for $Voxels_{mandibular\ third\ molars} \in Voxels_{mandibular\ third\ molars}$ **do**
4. if TC_{pro} of $Voxels_{mandibular\ third\ molars} > T$ using SHA optimization **then**
5. $replace_voxel_with_TC(Voxels_{mandibular\ third\ molars}, Pre)$
6. **end if**
7. if TC_{pro} of $Voxels_{mandibular\ third\ molars} < T$ using SHA optimization **then**
8. $remain_voxel_unchanged(Voxels_{mandibular\ third\ molars}, Pre)$
9. **end if**
10. **end for**
11. $Voxels_{enhancing} \leftarrow extract_voxels_{in_enhancing}(Pre)$
12. $NumVoxel_{enhancing} \leftarrow compute_num(Voxels_{enhancing})$
13. if $NumVoxel_{enhancing} < 150$ **then**
14. for $Voxel_{enhancing} \in Voxel_{enhancing}$ **do**
15. $replace_voxel_with_TC(Voxel_{enhancing}, Pre)$
16. **end for**
17. **end if**
18. $FPre \leftarrow Pre$
19. **Return FPre**

ALGORITHM 1: Postprocessing.

(i) A voxel is more likely to be classified as belonging to the proper class if the neural network is generally confident. As a result, the first step differs from Chen et al. [31] because it is based on the softmax probability of the neural network. All the voxels that can be separated into mandibular third molars are extracted first because IANSegNet makes it straightforward to partition the core. For these voxels, we get the likelihood of TCpro (the core region's probability). When TCpro is bigger than a threshold T , which is obtained by using the proposed optimization algorithm called SHA, where the existing technique [36] choose these threshold value between 0.1 to 0.3 and based on this value only, TCpro leads to high computation time for training the algorithm. In order to minimise the training time for large data, the threshold value must be identified by optimization (SHA) technique, which is explained as follows

3.5.1. Spotted Hyenas' Algorithm. The simulation of spotted hyenas' social behaviors is the main concept behind Algorithm 1. Only four steps are followed in this algorithm that are taken from the spotted hyenas' behaviors including encircling, hunting the prey, attacking, and searching behavior. Towards the best search agent for threshold, group of trusted hyenas is guided in the hunting process, and finally, the best optimal solution for threshold is saved.

(1) *Encircling Prey.* Here, the best solution is considered as either objective or target prey (i.e., threshold value), where the position and location of other search agents are updated according to the obtained best search solution. Equations

(10) and (11) provide the mathematical model for this search agent's behavior:

$$\vec{L}_h = \left| \vec{M} \times \vec{N}_p(x) - \vec{N}(x) \right|, \quad (10)$$

$$N(b+1) = \vec{N}_p(x) - \vec{O} \times \vec{L}_h. \quad (11)$$

At prey point, the distance between hyenas is described as \vec{L}_h , common vectors are depicted as \vec{M} and \vec{O} , current frequency is represented as b , location vector of prey is described as \vec{N}_p , and, finally, location vector of hyena is used as \vec{N} . The next two equations ((12) and (13)) are used to calculate the common vectors, where Equation (14) is used to find the parameter iterations that are used in Equation (13).

$$\vec{M} = 2 \times r \vec{d}_1, \quad (12)$$

$$\vec{O} = 2 \vec{s} \times r \vec{d}_2 - \vec{h}, \quad (13)$$

$$\vec{s} = 5 - \left(\text{Iteration} \times \frac{5}{\text{Max}_{\text{Iteration}}} \right), \quad (14)$$

where $\text{Iteration} = 0, 1, 2, \dots, \text{Max}_{\text{Iteration}}$. Random vectors between the range of 0 and 1 are denoted as $r \vec{d}_1$ and $r \vec{d}_2$, where the parameter \vec{s} is linearly minimised from 5 to 0.

(2) *Hunting.* Equations (15)–(17) are used to describe the strategy of spotted hyenas' algorithm (SHA) for hunting process.

$$\vec{L}_h = \left| \vec{M} \times \vec{N}_h - \vec{N}_k \right|, \quad (15)$$

$$\vec{N}_k = \vec{N}_h - \vec{O} \times \vec{L}_h, \quad (16)$$

$$\vec{O}_h = \vec{N}_k + \vec{N}_{k+1} + \dots + \vec{N}_{k+K}, \quad (17)$$

where initial best spotted hyena's position is denoted as \vec{N}_h and other spotted hyena's position is represented as \vec{N}_k . But, Equation (18) is used to calculate the overall number of spotted hyenas, which is denoted by the variable K .

$$K = \text{Count}_{\text{nos}} \left(\vec{N}_h, \vec{N}_{h+1}, \vec{N}_{h+2}, \dots, \vec{N}_h + \vec{R} \right). \quad (18)$$

In the range of 0.5 to 1, the value of random vector lies, which is denoted as \vec{R} , and it is used to describe the total number of solutions and counts for all candidate. For the number K , the solution for optimal set is represented as \vec{O}_h .

(3) *Attacking the Target.* Equation (19) is used to attack the target/prey of spotted hyena.

$$\vec{N}(b+1) = \frac{\vec{O}_h}{K}. \quad (19)$$

According to the best search agents' state, the best solution is protected by using $\vec{N}(b+1)$ and also updates the state of other spotted hyenas' search agent. The threshold value can be determined for parameter values from $\vec{N}(b+1)$.

(4) *Search for Prey.* \vec{O} is responsible if it is greater than or equal to 1 using Equation (13) to find a suitable solution. In addition, SHA algorithm uses the other important component, i.e., \vec{M} for research process. As provided in Equation (12), random weight of the prey is given by the random value that is presented in this another component. To show the SHA algorithm's inconsistent behavior and the effect of divergence, the vector $\vec{M} > 1$ is preferred over $\vec{M} < 1$. The local optimal problem is eliminated in SHA algorithm by solving the high dimensional issues with less computational efforts.

- (i) As long as there are fewer than 150 projected enhancing mandibular third molar voxels, we replace all of them with the core. The CBCT dataset's performance dictated the selection of the aforementioned cut-off points

4. Results and Discussion

An n1-highmem-8 system with an NVIDIA Tesla V100 GPU and PyTorch software [33, 34] were utilised. Fast.ai is

a deep learning framework that makes it possible to use CovidXrayNet thanks to its unique ability to connect several transformers into a pipeline that manages the least amount of computations and lossy operations possible.

4.1. *Evaluation Metrics.* The proposed technique is evaluated using Dice, sensitivity, and Hausdorff's distance. Dice and Hausdorff are two of the measures employed on the CBCT dataset's online review platform.

Segmentation results and the ground truth are compared using Dice to determine how similar they are to one other.

$$\text{Dice} = \frac{2\text{TP}}{\text{FP} + 2\text{TP} + \text{FN}}. \quad (20)$$

Sensitivity, which is utilised to determine the sum of TP, is provided by

$$\text{Sensitivity} = \frac{\text{TP}}{\text{TP} + \text{FN}}. \quad (21)$$

There were a total of FN, TP, and FP voxels that were both false negative and false positive in this study.

Hausdorff's distance is used to quantify the separation between the segmentation areas' and the ground truth's surfaces. It is referred to as the

$$\text{Haus}(G, P) = \max \left\{ \text{su}_{e \in E} \text{in}_{f \in Y} d(e, f), \text{su}_{f \in F} \text{in}_{e \in E} d(e, f) \right\}. \quad (22)$$

The supremum and infimum of the IAN area are denoted by su and in, respectively. The points on the surface represented by e are considered to be part of the ground truth (G). f represents the segmentation points on the Y -axis. To calculate the distance between e and f , use P and $d()$.

4.2. *Proposed Performance Evaluation.* Here, the suggested that the model's segmentation results are contrasted with those of already established methods like ResNet50, VGG16, and UNet (see Figure 5).

Ground truth prevarious processing's shapes, such as dense point sets (DP), alpha-shape polygonal meshes (PM), and voxelized raster volumes (VV), are taken into account for validation in this case. Here, the model from [32] is also considered and tested with dataset which is used in this work.

Table 1 shows that ERV-Net outperforms the most recent state-of-the-art techniques in terms of Dice. It is difficult to learn local information in vast patches, which results in lower Hausdorff distance findings. Our solution, on the other hand, surpasses ResNet50 in terms of Hausdorff's distance, with improvements in the VV and PM regions of 1.12 mm and 0.53 mm, respectively. The PM and VV regions of Dice have also been upgraded by a combined 53% and 60%. VGG16 came in second place when compared to the proposed model, which was first. When it came to performance, they proved that a well-trained network was enough. Table 1 shows that our approach outperforms the VGG16

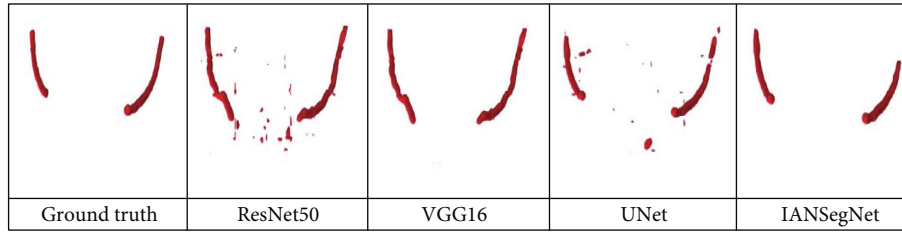


FIGURE 5: Postprocessing procedure when practical to proposed and existing forecasts.

TABLE 1: Presentation analysis with various existing pretrained models.

Method	Dice (%)			Hausdorff (mm)		
	DP	PM	VV	DP	PM	VV
ResNet50	81.48	91.09	85.49	2.63	4.20	6.69
VGG16	80.13	90.60	86.02	3.66	4.66	6.68
UNet	79.60	90.52	86.05	2.55	4.59	6.67
ERV-Net	82.81	91.21	86.62	2.70	3.88	6.81
Optimized IANSegNet	83.59	92.18	88.70	2.51	3.71	6.60

TABLE 2: Performance comparison of IANSegNet with diverse loss functions.

Loss	Postprocessing	Dice (%)			Sensitivity (%)			Hausdorff (mm)		
		DP	VV	PM	DP	PM	VV	DP	PM	VV
Cross-entropy	Yes	79.88	84.60	91.12	82.32	90.66	84.75	2.87	4.46	7.27
Dice	Yes	81.21	86.11	90.71	82.25	91.33	84.66	2.88	3.92	6.75
Fusion	Yes	83.59	92.18	88.70	85.20	91.39	85.59	2.51	3.71	6.60
Cross-entropy	No	75.49	84.18	91.12	81.57	90.66	82.61	3.69	4.46	7.27
Dice	No	79.05	85.80	90.71	82.33	91.33	83.33	2.87	3.92	6.80
Fusion	No	77.58	85.85	91.21	84.89	91.39	82.77	3.80	3.88	6.90

model. Noteworthy is that the IANSegNet technique delivers 1.6 percent and 2.03 percent Dice improvement, respectively, as well as a 1.95 mm Hausdorff distance improvement in the PM region. In the DP and VV regions, the suggested technique outperforms U-Net by 1.54 percent and 1.38 percent, respectively. With the exception of Hausdorff's distance in the VV region, IANSegNet exceeds theirs in every other parameter.

4.3. Performance Evaluation on Loss Function. For classification, cross-entropy is the most usually employed loss function. However, it frequently fails when applied to datasets that are unbalanced. The goal of the Dice loss function is to maximise the resemblance among neural network segmentation results and the actual data. The lack of cross-entropy is compensated for by applying the Dice loss to the imbalanced data. However, when it comes to small structures, it is susceptible to convergence difficulties. A loss function that combines the effects of both Dice and cross-entropy is used to solve this problem. These loss functions are discussed in this section primarily, and Table 2 offers validation results.

Table 2 demonstrates that when the cross-entropy loss function is used, none of the measures are at the top. Unbalanced data is an issue for cross-entropy loss function, which has worse DP metrics than other loss functions. Even while

TABLE 3: Proportional analysis of processing time for evaluated segmentation methods.

Pretrained model	Training time (hours)	Inference time -1 slice (seconds)
ResNet50	±132	±25
VGG16	±118	±17
UNet	±120	±12
EVR-Net	±115	±10
IANSegNet	±108	±8

the Dice loss function has a better Dice and Hausdorff distance, it has a lower sensitivity. Given that sensitivity measures the number of true positives, IANSegNet finds a significant number of them. Loss function fusion is more affected by IAN than are the loss function dice, as can be observed from the data. To make matters worse, the loss functions of fusions outperform Dice and cross-entropy. When Dice and cross-entropy are postprocessed to remove IAN false positives, the fusion loss function outperforms both by a considerable margin. The fusion loss function, which combines the benefits of cross-entropy and Dice, can be used to address the issue of data imbalance. The proposed fusion loss function beats other loss functions, such as

Dice or cross-entropy, in terms of performance. Additionally, the processing times for each strategy are exposed in Table 3.

The training time is less for proposed model, when compared with existing techniques, i.e., IANSegNet consumed nearly 108 hours, where other models consumed nearly 120 hrs to 130 hrs for training the same samples.

5. Conclusion

For the segmentation of IAN, we provide IANSegNet, a 3D residual neural network. When developing IANSegNet, we decided to first use an extremely deep neural network with a lightweight encoder, which allowed us to reduce the sum of parameters while still maintaining good performance. Using three pretrained models, we compared the suggested method's performance and efficiency to our own. To address the issue of degradation caused by the excessive number of convolutional layers in the encoder-decoder construction, we proposed a shallow decoder with residual blocks. Experiments have shown that decoders with residual blocks outperform those without them in terms of performance. Using entropy and Dice's advantages, a fusion loss function was created. Because of data imbalances or worries about convergence, using entropy or Dice was ineffective. A new optimized postprocessing technique took advantage of neural network probabilities and IAN distribution properties, greatly improving segmentation accuracy. Data from 3D CBCT scans was compared to the most up-to-date practices. IANSegNet was determined to be superior to its competitors in terms of both performance and computational complexity. We can further reduce the number of network parameters by employing a lightweight decoding algorithm. We intend to increase the IANSegNet's receptive field by adding more layers to improve the network's performance.

Data Availability

The data will be provided upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

References

- [1] S. Shin, Y. Lee, S. Kim, S. Choi, J. G. Kim, and K. Lee, "Rapid and non-destructive spectroscopic method for classifying beef freshness using a deep spectral network fused with myoglobin information," *Food Chemistry*, vol. 352, article 129329, 2021.
- [2] A. I. Maqueda, A. Loquercio, G. Gallego, N. García, and D. Scaramuzza, "Event-based vision meets deep learning on steering prediction for self-driving cars," in *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 5419–5427, Salt Lake City, UT, USA, 2018.
- [3] J. Mahler, M. Matl, X. Liu, A. Li, D. Gealy, and K. Goldberg, "Dex-net 3.0: computing robust vacuum suction grasp targets in point clouds using a new analytic model and deep learning," in *2018 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 5620–5627, Brisbane, QLD, Australia, 2018.
- [4] S. Back, S. Lee, S. Shin et al., "Robust skin disease classification by distilling deep neural network ensemble for the mobile diagnosis of herpes zoster," *IEEE Access*, vol. 9, pp. 20156–20169, 2021.
- [5] R. Poplin, A. V. Varadarajan, K. Blumer et al., "Prediction of cardiovascular risk factors from retinal fundus photographs via deep learning," *Nature Biomedical Engineering*, vol. 2, no. 3, pp. 158–164, 2018.
- [6] F. Kanavati, G. Toyokawa, S. Momosaki et al., "Weakly-supervised learning for lung carcinoma classification using deep learning," *Scientific Reports*, vol. 10, no. 1, 2020.
- [7] J. Lee, J. Park, S. Y. Moon, and K. Lee, "Automated prediction of extraction difficulty and inferior alveolar nerve injury for mandibular third molar using a deep neural network," *Applied Sciences*, vol. 12, no. 1, p. 475, 2022.
- [8] H. Tang, X. Chen, Y. Liu et al., "Clinically applicable deep learning framework for organs at risk delineation in CT images," *Nature Machine Intelligence*, vol. 1, no. 10, pp. 480–491, 2019.
- [9] H. Seo, S. Back, S. Lee, D. Park, T. Kim, and K. Lee, "Intra- and inter-epoch temporal context network (IITNet) using sub-epoch features for automatic sleep scoring on raw single-channel EEG," *Biomedical Signal Processing and Control*, vol. 61, article 102037, 2020.
- [10] J. Kühnisch, O. Meyer, M. Hesenius, R. Hickel, and V. Gruhn, "Caries detection on intraoral images using artificial intelligence," *Journal of Dental Research*, vol. 101, no. 2, pp. 158–165, 2022.
- [11] H. Wang, J. Minnema, K. J. Batenburg, T. Forouzanfar, F. J. Hu, and G. Wu, "Multiclass CBCT image segmentation for orthodontics with deep learning," *Journal of Dental Research*, vol. 100, no. 9, pp. 943–949, 2021.
- [12] F. Casalegno, T. Newton, R. Daher et al., "Caries detection with near-infrared transillumination using deep learning," *Journal of Dental Research*, vol. 98, no. 11, pp. 1227–1233, 2019.
- [13] V. Majanga and S. Viriri, "Automatic blob detection for dental caries," *Applied Sciences*, vol. 11, no. 19, p. 9232, 2021.
- [14] J. H. Lee, D. H. Kim, and S. N. Jeong, "Diagnosis of cystic lesions using panoramic and cone beam computed tomographic images based on deep learning neural network," *Oral Diseases*, vol. 26, no. 1, pp. 152–158, 2020.
- [15] S. Kim, Y. H. Lee, Y. K. Noh, F. C. Park, and Q. Auh, "Age-group determination of living individuals using first molar images based on artificial intelligence," *Scientific Reports*, vol. 11, no. 1, pp. 1073–1111, 2021.
- [16] L. Alfadil and E. Almajed, "Prevalence of impacted third molars and the reason for extraction in Saudi Arabia," *The Saudi Dental Journal*, vol. 32, no. 5, pp. 262–268, 2020.
- [17] S. Yilmaz, M. Z. Adisen, M. Misirlioglu, and S. Yorubulut, "Assessment of third molar impaction pattern and associated clinical symptoms in a central anatolian turkish population," *Medical Principles and Practice*, vol. 25, no. 2, pp. 169–175, 2016.
- [18] V. R. Kumar, P. Yadav, E. Kahsu, F. Girkar, and R. Chakraborty, "Prevalence and pattern of mandibular third molar impaction in Eritrean population: a retrospective study," *The Journal of Contemporary Dental Practice*, vol. 18, no. 2, pp. 100–106, 2017.

- [19] M. Eshghpour, A. Nezadi, A. Moradi, R. M. Shamsabadi, N. M. Rezaer, and A. Nejat, "Pattern of mandibular third molar impaction: a cross-sectional study in northeast of Iran," *Nigerian Journal of Clinical Practice*, vol. 17, no. 6, pp. 673–677, 2014.
- [20] A. Jaroń and G. Trybek, "The pattern of mandibular third molar impaction and assessment of surgery difficulty: a retrospective study of radiographs in East Baltic population," *International Journal of Environmental Research and Public Health*, vol. 18, no. 11, p. 6016, 2021.
- [21] X. Wei and Y. Wang, "Inferior alveolar canal segmentation based on cone-beam computed tomography," *Medical Physics*, vol. 48, no. 11, pp. 7074–7088, 2021.
- [22] J. Jaskari, J. Sahlsten, J. Järnstedt et al., "Deep learning method for mandibular canal segmentation in dental cone beam computed tomography volumes," *Scientific Reports*, vol. 10, no. 1, pp. 1–8, 2020.
- [23] M. Cipriano, S. Allegretti, F. Bolelli, F. Pollastri, and C. Grana, "Improving segmentation of the inferior alveolar nerve through deep label propagation," in *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 21137–21146, New Orleans, LA, USA, 2022.
- [24] W. Qi, J. Lei, Y. N. Liu, J. N. Li, J. Pan, and G. Y. Yu, "Evaluating the risk of post-extraction inferior alveolar nerve injury through the relative position of the lower third molar root and inferior alveolar canal," *International Journal of Oral and Maxillofacial Surgery*, vol. 48, no. 12, pp. 1577–1583, 2019.
- [25] M. Tassoker, "Diversion of the mandibular canal: is it the best predictor of inferior alveolar nerve damage during mandibular third molar surgery on panoramic radiographs?," *Imaging science in dentistry*, vol. 49, no. 3, pp. 213–218, 2019.
- [26] R. Pandey, C. Ravindran, D. Pandiyan, A. Gupta, A. Aggarwal, and S. Aryasri, "Assessment of roods and Shehab criteria if one or more radiological signs are present in orthopantomogram and position of the mandibular canal in relation to the third molar apices using cone beam computed tomography: a radiographic study," *Tanta Dental Journal*, vol. 15, no. 1, p. 33, 2018.
- [27] D. S. Nayak, S. A. Raghavan, P. Birur, S. Gurudath, and G. Keerthi, "Determination of proximity of mandibular third molar to mandibular canal using panoramic radiography and cone-beam computed tomography," *Journal of Indian Academy of Oral Medicine and Radiology*, vol. 29, no. 4, p. 273, 2017.
- [28] F. Fusaro, F. Costantinides, M. Maglione, D. Dalessandri, E. Vettori, and R. di Lenarda, "Comparison of orthopantomography and CT accuracy in predicting inferior alveolar nerve injury following lower third molar surgical extraction," *Minerva Stomatologica*, vol. 66, no. 5, pp. 193–200, 2017.
- [29] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2014, <https://arxiv.org/abs/1409.1556>.
- [30] L. C. Chen, G. Papandreou, F. Schroff, and H. Adam, "Rethinking atrous convolution for semantic image segmentation," 2017, <https://arxiv.org/abs/1706.05587>.
- [31] L. C. Chen, Y. Zhu, G. Papandreou, F. Schroff, and H. Adam, "Encoder-decoder with atrous separable convolution for semantic image segmentation," in *Proceedings of the European Conference on Computer Vision (ECCV)*, pp. 801–818, Munich, Germany, 2018.
- [32] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 770–778, Las Vegas, NV, USA, 2016.
- [33] F. Chollet, "Xception: deep learning with depthwise separable convolutions," in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1251–1258, Honolulu, HI, USA, 2017.
- [34] N. Ma, X. Zhang, H. T. Zheng, and J. Sun, "Shufflenet v2: practical guidelines for efficient cnn architecture design," in *Proceedings of the European Conference on Computer Vision (ECCV)*, pp. 116–131, Munich, Germany, 2018.
- [35] C. Zhou, C. Ding, X. Wang, Z. Lu, and D. Tao, "One-pass multi-task networks with cross-task guided attention for brain tumor segmentation," *IEEE Transactions on Image Processing*, vol. 29, pp. 4516–4529, 2020.
- [36] X. Zhou, X. Li, K. Hu, Y. Zhang, Z. Chen, and X. Gao, "ERV-Net: an efficient 3D residual neural network for brain tumor segmentation," *Expert Systems with Applications*, vol. 170, article 114566, 2021.