Computers in Biology and Medicine 168 (2024) 107821

Contents lists available at ScienceDirect

Computers in Biology and Medicine

journal homepage: www.elsevier.com/locate/compbiomed

A fine-grained orthodontics segmentation model for 3D intraoral scan data

Juncheng Li [a,*], Bodong Cheng [b], Najun Niu [c], Guangwei Gao [d], Shihui Ying [e], Jun Shi [a], Tieyong Zeng [f]

[a] School of Communication Information Engineering, Shanghai University, Shanghai, China
[b] School of Computer Science and Technology, East China Normal University, Shanghai, China
[c] School of Stomatology, Nanjing Medical University, Nanjing, China
[d] Institute of Advanced Technology, Nanjing University of Posts and Telecommunications, Nanjing, China
[e] Department of Mathematics, School of Science, Shanghai University, Shanghai, China
[f] Department of Mathematics, The Chinese University of Hong Kong, New Territories, Hong Kong

ARTICLE INFO

Keywords:
Tooth segmentation
Intraoral scan
GNN
3D mesh

ABSTRACT

With the widespread application of digital orthodontics in the diagnosis and treatment of oral diseases, more and more researchers focus on the accurate segmentation of teeth from intraoral scan data. The accuracy of the segmentation results will directly affect the follow-up diagnosis of dentists. Although the current research on tooth segmentation has achieved promising results, the 3D intraoral scan datasets they use are almost all indirect scans of plaster models, and only contain limited samples of abnormal teeth, so it is difficult to apply them to clinical scenarios under orthodontic treatment. The current issue is the lack of a unified and standardized dataset for analyzing and validating the effectiveness of tooth segmentation. In this work, we focus on deformed teeth segmentation and provide a fine-grained tooth segmentation dataset (3D-IOSSeg). The dataset consists of 3D intraoral scan data from more than 200 patients, with each sample labeled with a fine-grained mesh unit. Meanwhile, 3D-IOSSeg meticulously classified every tooth in the upper and lower jaws. In addition, we propose a fast graph convolutional network for 3D tooth segmentation named Fast-TGCN. In the model, the relationship between adjacent mesh cells is directly established by the naive adjacency matrix to better extract the local geometric features of the tooth. Extensive experiments show that Fast-TGCN can quickly and accurately segment teeth from the mouth with complex structures and outperforms other methods in various evaluation metrics. Moreover, we present the results of multiple classical tooth segmentation methods on this dataset, providing a comprehensive analysis of the field. All code and data will be available at https://github.com/MIVRC/Fast-TGCN.

1. Introduction

With the development of society and the improvement of people's livelihood, more and more people have begun to pay attention to dental problems and hope to ensure the normal function of teeth and facial shape through dental treatment, orthodontics, restoration, and dental implants. Traditional diagnosis and treatment require dentists to make diagnoses and treatment plans based on their own experience, which is extremely difficult. More importantly, the outcome of the treatment depends directly on the doctor's judgment. And due to the current shortage of medical resources, dentists need to deal with many patients, which makes doctors face a huge challenge. In recent years, thanks to the progress of computer technology, the digital dentistry has made remarkable achievements in orthodontics, prosthesis, and implant. By leveraging AI, dental care is becoming more convenient and scientific, directly benefiting patients and doctors. In dental diagnosis and treatment, digital medical images are an important basis for doctors to diagnose. Depending on the mode of acquisition, there are typically three modality data based on 2D X-ray, 3D cone beam computed tomography (CBCT) and 3D intraoral scan. Different from the other two modes, 3D intraoral scans can obtain high-precision tooth surface information, which can help doctors design orthodontic schemes [1], deduce the orthodontic process [2] and produce braces. However, manual labeling is tedious, so designing an automatic segmentation and labeling method for tooth data is necessary (see Fig. 1).

In the past decade, deep neural networks have shined in various fields and have also greatly promoted the development of tooth segmentation, such as segmentation methods for 2D X-ray [3–8], 3D cone beam computed tomography (CBCT) [9–14] and 3D intraoral scans

* Corresponding author.
E-mail addresses: junchengli@shu.edu.cn (J. Li), bdcheng@stu.ecnu.edu.cn (B. Cheng), najunniu@njmu.edu.cn (N. Niu), csgwgao@njupt.edu.cn (G. Gao), shying@shu.edu.cn (S. Ying), junshi@shu.edu.cn (J. Shi), zeng@math.cuhk.edu.hk (T. Zeng).

https://doi.org/10.1016/j.compbiomed.2023.107821
Received 27 April 2023; Received in revised form 1 November 2023; Accepted 4 December 2023
Available online 6 December 2023

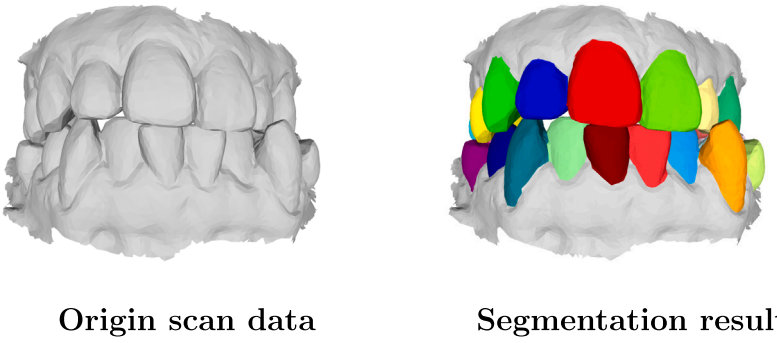

Fig. 1. Example of 3D tooth segmentation. Different colors represent different types of teeth.

[15–18]. Most of these methods are data-driven, so how to obtain high-precision patient data is crucial. However, most methods do not disclose the used datasets or use different datasets, thus lacking a common benchmark to fairly evaluate the performance of these methods. Moreover, these methods are often only applicable to tooth data with normal morphology and cannot effectively segment the malformed teeth in clinical applications.

To solve the aforementioned problems, in this work, we construct the first fine-grained orthodontic tooth segmentation dataset (3D-IOSSeg) based on 3D intraoral scan data and propose a new tooth segmentation network. In the dataset, we collected oral scans from 90 patients with all common dental abnormalities. Subsequently, we added an additional 130 data samples to encompass a broader range of patient cases. Thus, the entire dataset comprises data from 220 patients, encompassing cases of dental anomalies such as missing teeth, overlapping, misalignment, and malocclusion. Specifically, each patient sample consisted of two rows of upper/lower tooth data. Meanwhile, we used the most common form of triangular mesh to describe the internal surface of the mouth to improve accuracy. It is worth noting that each row of teeth (including the gums) consists of 100,000 to 30,000 points and 100,000 to 450,000 mesh cells, and there are more than 20 million cells in the entire dataset. Based on this dataset, we also propose a graph convolutional network for fast tooth segmentation, called Fast-TGCN. In summary, this paper introduces the more extreme dental conditions in practical applications into the dataset, aiming to help researchers in the field address potential challenges that may arise and to analyze them to inform future work. The main contributions of this paper are as follows:

(1) We construct a standardized clinical orthodontic tooth segmentation dataset (3D-IOSSeg). 3D-IOSSeg is based on 3D intraoral scan data, which contain a rich array of deformed teeth, and all of them are fine-grained labeled. Moreover, this will be a publicly available dataset.

(2) We propose a novel network for fast tooth segmentation called Fast-TGCN, which uses a more naive adjacency matrix to construct the relationship of 3D surfaces, and combines the information flow characteristics of graph convolutional network (GCN), setting a new benchmark for orthodontics segmentation.

(3) We present the results of multiple classical tooth segmentation methods on the propose dataset, providing a comprehensive comparison and analysis of the field.

2. Related works

Common digital dental imaging includes 2D X-rays, 3D cone beam computed tomography (CBCT), and 3D intraoral scans. Due to its excellent surface modeling capabilities, 3D intraoral scans have been widely used in the fields of prosthodontics and orthodontics. In order to assist doctors to achieve accurate diagnosis and treatment, more and more work has begun to focus on using intelligent algorithms to segment different types of teeth from the 3D intraoral scan data.

3D intraoral scan data usually have extremely strong geometric properties, so traditional segmentation methods include curvature-based methods [19–23], contour-based methods [24,25]and harmonic field-based methods [26]. There are also methods [27,28] to project a 3D data into a 2D plane segment and then project it back into 3D space. However, these methods lack data-driven, resulting in poor performance, and cannot be applied clinically.

With the great success of deep learning in 3D point cloud segmentation, some neural network-based 3D tooth segmentation methods [29–34] have also been proposed.

At first, some methods [29,30,34] were inspired by voxelization, which converts the 3D intraoral scans into a regular structure, which is then directly processed by CNN, for example. However, this indirect feature conversion method, as a compromise solution, will lead to the loss of the original geometric information of the data and affect the quality of tooth segmentation. At the same time, some tooth segmentation methods that combine geometry and CNN have emerged, such as Xu et al. [35] used a two-level layered method to extract the features of the intraoral scan data, and relied on CNN to extract different geometric features, and finally obtained a refined tooth segmentation result. Tian et al. [36] also adopted a two-level hierarchical network structure, and this method also introduced a sparse voxel o-ctree to perform pre-feature extraction on the original data. Although these methods using a hierarchical network structure have better performance than traditional geometric segmentation methods, due to the limitations of the hierarchical structure, errors at the upper level cannot be corrected at the next level, which affects the segmentation performance.

Later, with the rise and maturity of 3D point cloud processing, a faster and more effective processing strategy [37] was provided for the study of intraoral scan data. [38] presents an end-to-end deep learning framework for semantic segmentation of individual teeth as well as gingiva from point clouds representing IOS. By introducing non-uniform resampling techniques, models can be trained and deployed at higher spatial resolutions.

Since PointNet does not take into account the local features of the 3D data, Lian et al. [39] used KNN to build an end-to-end 3D intraoral scans segmentation network MeshSegNet for automatic tooth marking on the original tooth surface. This method integrates a series of graph constraint learning modules to extract multi-scale local layers hierarchically. Contextual features, achieving decent performance. Based on MeshSegNet, the surface normal vector information of 3D teeth is further considered, and a dual-branch tooth segmentation network TSGCNet [40] is proposed that combines coordinate and normal vector information. However, the KNN graph used in the above methods usually contains cross-surface information, which will confuse the information on different surfaces and is not conducive to fine-grained tooth segmentation. To extract the geometric features of the 3D tooth data more effectively, we will construct an adjacency graph from the spatial position adjacency relationship between mesh cells and use the adjacency matrix to guide the network to focus on the local geometric features of the teeth.

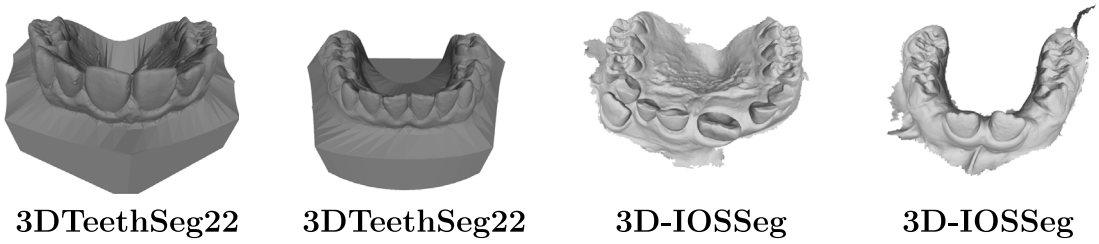

Fig. 2. Visual comparison of 3DTeethSeg22 and 3D-IOSSeg datasets. Obviously, although 3DTeethSeg22 obtained a relatively regular 3D intraoral scans data through indirect scanning, its sample will be accompanied by a plaster base, and the gingival part is not very obvious. In contrast, our 3D-IOSSeg can reflect the scanning results in the real diagnostic environment through direct scanning modeling, and more importantly, it contains more and complex abnormal tooth samples.

3. Proposed 3D-IOSSeg dataset

To promote the development of digital orthodontics, we propose a standardized clinical orthodontic tooth segmentation dataset, named 3D-IOSSeg. In this section, we will describe the acquisition of intraoral scan data, organizational form, as well as the classification and tooth labeling process. Finally, we will conduct a detailed analysis of the sample distribution of the dataset.

3.1. Dataset acquisition and annotation

The acquisition of 3D intraoral scans data in clinical practice usually uses a dedicated oral scanner and adopts digital intraoral impression technology. The principle is to obtain digital impressions of teeth, gingiva, and mucosal tissues in the oral cavity through optical compression technology. This work adopts the 3Shape Trios oral digital. It is worth noting that digital intraoral scans data impression technology is divided into direct and indirect methods. The indirect method is to obtain a 3D intraoral scans data from a plaster cast or fine model of the patient's mouth. This is also a commonly used data set construction method in previous works. The direct method is a technique of obtaining a digital model directly through intraoral scans, which reduces errors in the mold-making process. To better reflect the characteristics of the oral cavity of patients, we adopted the direct method for data acquisition.

The 3D-IOSSeg dataset serves as a fully public 3D intraoral scan dataset geared towards more specific orthodontic treatments. It is worth noting that the 3DTeethSeg22,[1] challenge in MICCAI 2022 has released a segmentation dataset of 3D intraoral scan data. As shown in Fig. 2 compared with 3DTeethSeg22, it can be clearly found that the samples in 3D-IOSSeg have abundant abnormal teeth. Meanwhile, both 3DTeethSeg22 and existing open-source datasets were obtained by indirect scanning, accompanied by a plaster base. However, the sample data in our 3D-IOSSeg is directly scanned, which contain a large amount of oral tissue, which is closer to the human oral environment. The above has a more positive effect on the diagnosis and treatment of orthodontics. More importantly, both original and annotated data of the proposed 3D-IOSSeg will be publicly available, and researchers can freely manipulate the dataset through 3D view software such as Meshlab to adapt to different research needs. In summary, 3D-IOSSeg has the following properties:

(1) **Broad coverage of the population.** 3D-IOSSeg used in previous works did not consider the age stage of the sample, which could not guarantee the coverage of various populations. Through the statistics of the age of patients (Fig. 3), we can clearly observe that our dataset includes not only the oral data of adults but also a large number of samples of adolescents under the age of 20. Teenagers have slightly different tooth shapes than adults, and the shape and size of their jaws are also different from adults. As one of the groups with a high incidence of dental problems, teenagers should also be given extensive attention. Therefore, the research on this kind of data is urgent and challenging.

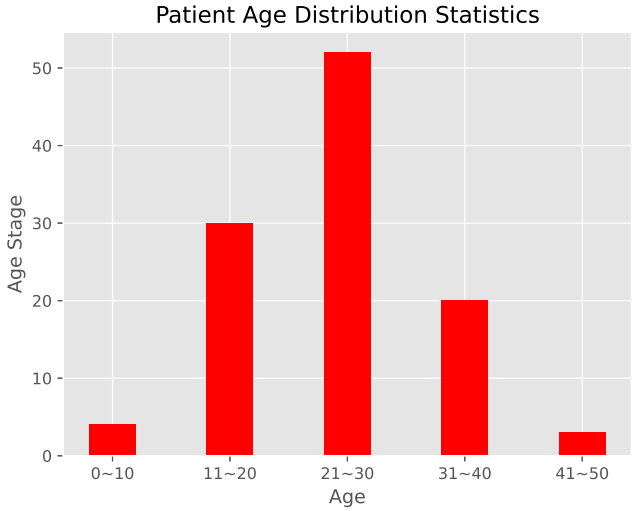

Fig. 3. Patient age distribution of 3D-IOSSeg dataset.

(2) **Diverse and unusual teeth.** Segmentation of abnormal teeth is always a challenging task by investigating related tooth segmentation methods. Because these teeth are often crowded, misplaced, or even largely displaced, it is very difficult to extract them accurately from other teeth. Moreover, this abnormality will lead to large changes in the topology of the data, which will bring great challenges to the learning of the segmentation network. Considering this situation, 3D-IOSSeg introduces a higher proportion of abnormal tooth samples for training and testing.

(3) **Fine-grained division.** Different from the method of symmetrical division of tooth categories adopted by other methods, we treat each tooth as an independent class division, so that the number of target categories in this dataset reaches 33 categories. Such fine-grained division can better assist dentists in their diagnosis in practice. Meanwhile, it also is more difficult to accurately distinguish individual teeth, as many teeth have similar shapes, making the dataset more challenging.

After obtaining patient oral data, we started to label each group of samples. This work adopts a combination of expert guidance and cross-labeling. 3D-IOSSeg uses Meshlab as a marking tool, which can easily color the target area of the 3D data directly. We use different color marks to distinguish different types of teeth. Each sample in 3D-IOSSeg contains more than 100,000 mesh cells, and more than 20 million mesh cells in the entire dataset need to be classified and labeled. Finally, we export the annotated tooth data as a PLY file. Specifically, we invited 10 people to carry out data labeling work, and each person was assigned 22 groups of samples. It is worth noting that each group of samples contained two 3D intraoral scans of the upper and lower jaws. Meanwhile, the whole labeling process is divided into three rounds.

[1] https://3dteethseg.grand-challenge.org/.

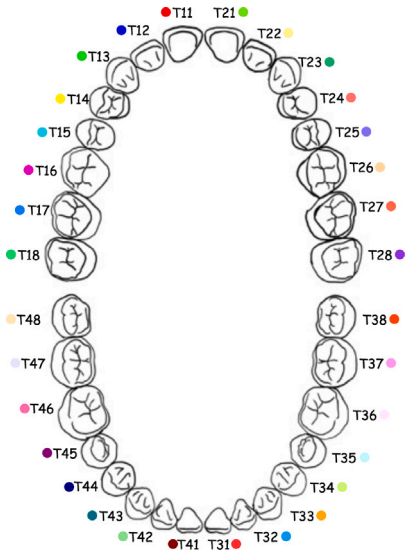

Fig. 4. The corresponding color coding for FDI tooth notation.

The first round of marking is carried out under the guidance of dental experts. After each marking is completed, the experts will check and give feedback on the marking results. The second round adopts the crossover method between different groups, and the annotators will be reassigned data for modification. The third round of annotation will fine-tune the edges of the different teeth in the data and check again for errors. Therefore, the data obtained after the above three rounds of labeling are used as the final result of the dataset. The advantages of the method include: 1. Avoid human factors that cause data to be mislabeled, such as wrong coloring of tooth categories. 2. Reduce the impact of the subjective bias of the annotators on the accuracy of the data. With the help of this method, we are able to build a tooth segmentation dataset that is as accurate as possible.

3.2. Organizational form of 3D-IOSSeg dataset

3D-IOSSeg takes 3D mesh data as the processing object. 3D mesh data is a kind of 3D data that is easier to operate and observe based on 3D point cloud data. As shown in Fig. 5, mesh is a refined extension of point clouds. Therefore, dentists can quickly make a diagnosis through mesh data.

In particular, each tooth sample of the 3D-IOSSeg is stored in a PLY file format that consist of a header, a list of vertex, and a list of face. Among them, the header defines the internal organization of the file, such as which face are used by the 3D sample to describe the mesh, and the data type and attribute arrangement order of the spatial coordinates in the vertex and face lists. The vertex list and face list are used to store vertex data (spatial coordinates and normal vectors) and mesh data (vertex numbers and color labels).

Adults usually have 32 permanent teeth, which are divided into incisors, canines, premolars, and molars according to their function and morphology, and previous work usually makes a coarse-grained classification according to the above categories. However, in the actual diagnosis process, it is often necessary to precisely locate the position of a certain tooth, which makes the previous methods unable to accurately segment specific teeth. Therefore, fine-grained labeled data samples are needed to drive current segmentation models. As shown in Fig. 4, in this work, we analyze the relevant data for a fine-grained division of 32 human teeth. According to the different positions of the upper and lower jaws and the left and right distribution, we divide the teeth into frontal incisors, lateral incisors, canines, first premolars, second premolars, first molars, second molars, and third molars, a total of 32 tooth categories. More detail definitions can be found in the Tables 1 and 2 for details.

Table 1
Statistical details of upper jaw teeth.

Class	Name	Mesh cells	Proportion	Definition
0	T11	770 113	3.17%	Upper jaw front incisor (left)
1	T12	764 112	3.12%	Upper jaw side incisor (left)
2	T13	630 428	2.57%	Upper jaw canine (left)
3	T14	775 465	3.20%	Upper jaw first premolar (left)
4	T15	606 966	2.47%	Upper jaw second premolar (left)
5	T16	1 186 691	4.87%	Upper jaw first molar (left)
6	T17	959 921	3.88%	Upper jaw second molar (left)
7	T18	85 656	0.32%	Upper jaw third molar (left)
8	T21	788 404	3.24%	Upper jaw front incisor (right)
9	T22	595 660	2.43%	Upper jaw side incisor (right)
10	T23	628 243	2.57%	Upper jaw canine (right)
11	T24	788 811	3.22%	Upper jaw first premolar (right)
12	T25	752 012	3.09%	Upper jaw second premolar (right)
13	T26	1 202 707	4.94%	Upper jaw first molar (right)
14	T27	935 754	3.80%	Upper jaw second molar (right)
15	T28	111 378	0.42%	Upper jaw third molar (right)
32	Gingiva	12 931 958	52.69%	Upper jaw gingiva

Table 2
Statistical details of lower jaw teeth.

Class	Name	Mesh cells	Proportion	Definition
16	T31	464 287	2.58%	Lower jaw front incisor (left)
17	T32	494 379	2.71%	Lower jaw lateral incisor (left)
18	T33	569 232	3.09%	Lower jaw canine (left)
19	T34	654 778	3.57%	Lower jaw first premolar (left)
20	T35	690 873	3.78%	Lower jaw second premolar (left)
21	T36	1 152 077	6.28%	Lower jaw first molar (left)
22	T37	889 824	4.81%	Lower jaw second molar (left)
23	T38	77 463	0.40%	Lower jaw third molar (left)
24	T41	446 978	2.45%	Lower jaw front incisor (right)
25	T42	487 693	2.67%	Lower jaw lateral incisor (right)
26	T43	580 307	3.16%	Lower jaw canine (right)
27	T44	660 538	3.59%	Lower jaw first premolar (right)
28	T45	726 970	4.01%	Lower jaw second premolar (right)
29	T46	1 121 742	6.10%	Lower jaw first molar (right)
30	T47	894 965	4.83%	Lower jaw second molar (right)
31	T48	102 208	0.53%	Lower jaw third molar (right)
32	Gingiva	8 347 289	45.47%	Lower jaw gingiva

3.3. Classification of teeth in intraoral scan data

Unlike coarse-grained segmentation, fine-grained tooth segmentation is more difficult. The table shows the proportion of surface area occupied by different teeth in the upper and lower jaws in the dataset. It can be observed that the proportion of each tooth is very small, and it is challenging to segment them accurately. According to the table, we found large differences in distribution between different teeth. Among them, the proportion of molars is the largest, and the proportion of wisdom teeth is the least. This reflects the larger volume and more complex crowns of the molars. For wisdom teeth, which are only present in a small number of samples, segmenting them is difficult. In addition, we also found some samples of missing teeth and broken teeth. This means that the dataset contains a rich variety of samples, which further increases the challenge of the task.

3.4. Overview of abnormal teeth

Abnormal teeth are very common in the clinical diagnosis of orthodontics, and their irregular distribution in position and geometry leads to the complexity of spatial distribution, which brings greater challenges to 3D tooth segmentation. In this work, we introduced many scan samples with abnormal teeth in the 3D-IOSSeg dataset. Fig. 6 shows some representative abnormal tooth morphology in 3D-IOSSeg. In addition to the common misalignment of the teeth, we also collected some samples of patients with labially inclined anterior teeth, such as sample 26, the front teeth of these patients are towards the lip The directional tilt causes the overall arch to protrude outward. We also

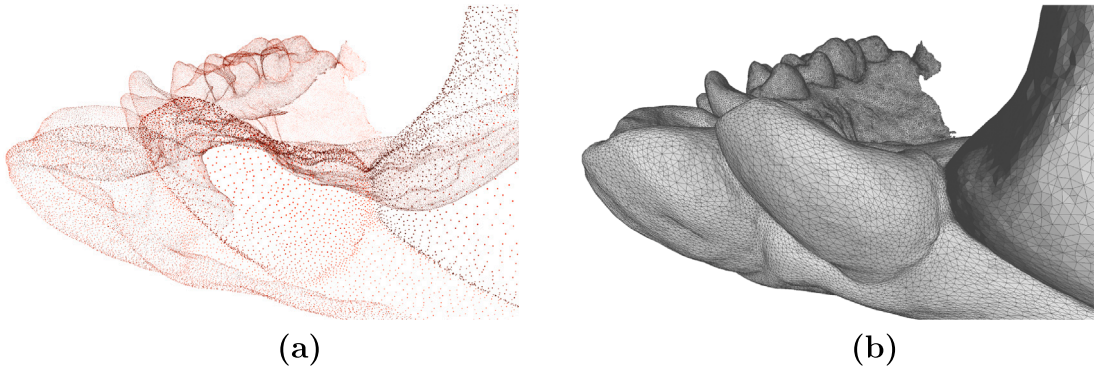

Fig. 5. (a) represents point cloud data, which is composed of a large number of discrete points. Although point cloud can be processed well by computer, it is difficult to present intuitively; (b) represents mesh data. Obviously, it has more advantages than point cloud. Good visual expression.

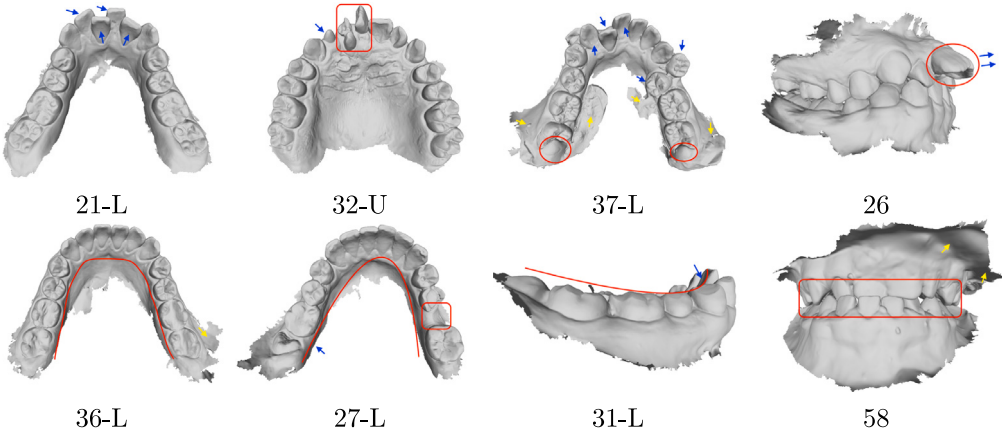

Fig. 6. Some samples of abnormal teeth in 3D-IOSSeg. 3D-IOSSeg covers a wealth of abnormal dental conditions, including common irregular incisors (indicated by the blue arrows in 21-L, 37-L, and the red boxes in 32-U), as well as the less common atrophied teeth (indicated by blue arrows in 32-U); Sample 26 represents a data of a patient with anteverted lips (the incisors in the area circled by a red frame slope in the direction indicated by the blue arrow); Samples 36-L and 27 -L shows two kinds of mandibular arch shapes (the red curve shows the outline of the dental arch), and 27-L also has abnormal distortion of the mandibular arch; sample 31-L shows the situation that the spee curve is too large, which is also exclusive to 3D-IOSSeg; Sample 58 is a dental data of a child whose teeth are shaped differently than those of an adult. In addition, the yellow arrows in the figure indicate other complex situations in the oral cavity, which pose a great challenge for tooth segmentation.

considered an abnormal Spee Curve, which is one of the important factors for the success of orthodontics. Sample 31-L reflects the excessive curvature of the Spee Curve, which affects the overall tooth alignment. Not only that, but we further considered some children's tooth data (58), these samples will contain a large number of deciduous teeth, which are different from the permanent teeth in size and arrangement, so they also need to be widely concerned.

3.5. Dataset split

In this work, we divide the annotated dataset regarding the recommendations of professional dentists and the results of data analysis, respectively. First, dentists divided the samples into three categories of low, medium, and high degrees of deformity based on their own experience. On this basis, we use the principal component analysis (PCA) method to reduce the dimensionality of the samples and map the samples into a two-dimensional space. Furthermore, we searched for patient samples corresponding to different locations and found that the degree of deformity of the samples increases with the distance from the origin. Through the data distribution information and the doctor's division results, it is found that the sample points distributed in the marginal area correspond to samples of highly abnormal categories. According to the above results, we selected some abnormal tooth samples as the test data set to ensure that the distribution of test samples and training samples is different, which increases the challenge of segmentation and also helps other researchers to verify the effectiveness of the proposed method.

4. Proposed method

4.1. Challenge

It is well known that 3D data have richer geometric information than 2D images. In 3D mesh data, these features are usually composed of several adjacent mesh cells. However, depending on the save format and software, the order of mesh cells in the generated 3D data file is random. For the intraoral scan data, the spatial geometric information contained in it is more diverse. Especially for data with missing, misplaced, and deformed teeth, they often lead to higher uncertainty in the tooth arrangement of the entire data and the topology of mesh cells. Therefore, effectively guiding the network to extract stable feature representation from the unstable distribution of the oral data is a challenging task.

4.2. Method

4.2.1. Adjacency matrix graph

Previous works often use multi-layer perceptrons or one-dimensional convolution to extract stable features and combine global pooling to obtain global features to solve the problem of disorder of data input. Further, to solve the problem that the heterogeneity between samples is too large and difficult to learn due to various deformities in the oral cavity, many studies have introduced the idea of constructing KNN graphs using spatial Euclidean distance. Undoubtedly, this method

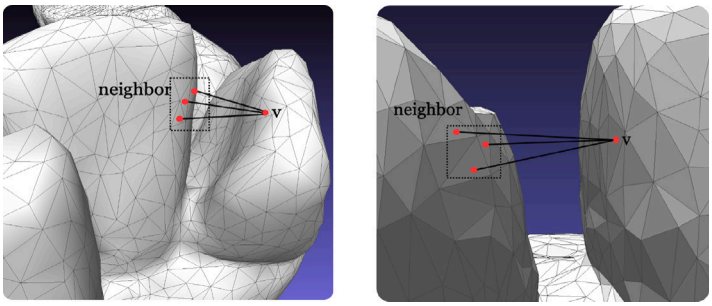

Fig. 7. Schematic diagram of the discontinuity phenomenon in the construction of KNN graph.

effectively enhances the expression of local information and can significantly improve the segmentation performance of the 3D intraoral scan data by matching the normal vector of the space. However, we found that the KNN diagram could not reflect the local characteristics of the data well under certain circumstances. When the distance between two surfaces is very close, the neighbor cell of the mesh cell in the center of one surface is likely to contain the mesh cell of the other surface, which will result in a discontinuous area. According to Fig. 7, it can be observed that the neighbor cell of one tooth contains the cell of the other tooth, while the middle part of the two teeth is not divided. Since the 3D data is formed by the interrelation of mesh cells in different positions, the curvature of the surface can usually be described by the normal vector of the mesh cells of the surface. In addition, the 3D tooth data is composed of continuous and uninterrupted clusters of curved surfaces. If discontinuity occurs in the process of constructing the KNN diagram, a lot of information will be lost, and the normal vector information cannot be effectively transmitted between adjacent cells.

To solve the above problems, we replace the traditional KNN graph with the adjacency matrix graph in graph theory. First of all, the natural topology of the 3D mesh data can be directly described by the adjacency matrix, and the adjacency matrix diagram has very good stability. It is only generated according to the 3D mesh data file and will not change with the characteristics, which is impossible with KNN diagrams. In Fig. 8, we compare the difference between the adjacency matrix graph and the KNN graph. It can be seen that the adjacency matrix strictly follows the dependencies between adjacent points, and there will be no jumping relationship across the surface of the KNN graph. Meanwhile, unlike the implicit feature representation of the KNN graph, the adjacency matrix graph is an explicit feature representation that can be directly reflected by the surface of the tooth, which significantly improves the interpretability of the segmentation network. Moreover, the construction of the adjacency matrix also avoids the unnecessary overhead of frequent calculation of euclidean distance, thereby improving the efficiency of the network segmentation.

Specifically, we use each mesh cell in the data as a center point to construct an adjacency matrix. Since the triangular mesh data is used in this paper, each mesh cell is composed of three vertexes. First, we define a set family V as

$$V = \left\{ v_k \mid k \in N^* \wedge k \leq n \right\}, \tag{1}$$

where v_k represents the kth mesh cell in the sample, and $v_k = \{a_k, b_k, c_k\}$ represents that v_k is composed of three vertexes a_k, b_k, and c_k. N^* is a set of positive integers, and n is the number of mesh cells included in the sample. Then, we define that the cells sharing the same vertexes are adjacent to each other, and initialize an N-order matrix A, where the elements a_{ij} are

$$a_{ij} = \begin{cases} 0 & \text{if } v_i \cap v_i = \emptyset \\ 1 & \text{if } v_i \cap v_i \neq \emptyset, \end{cases} \tag{2}$$

According to the above definition, the adjacency matrix of the intraoral scan data can be quickly obtained without additional calculation. After constructing the topology of the data, we introduced the classic graph convolutional network (GCN) [41] in graph data mining to extract the geometry of the tooth data. Since GCN can spread the information contained in the graph (Graph) along the edges of the topological structure to aggregate the information around the nodes:

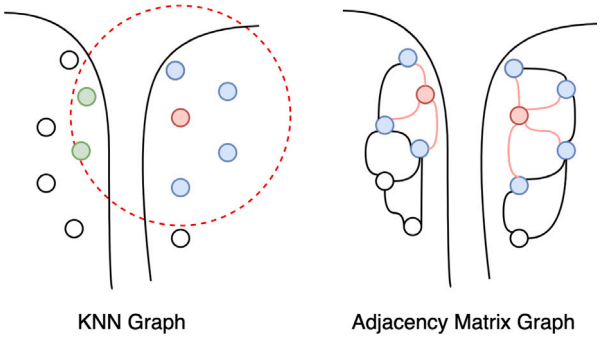

Fig. 8. Schematic diagram of KNN graph and adjacency matrix graph.

$$GCN(F) = act(A \cdot F \cdot W), \tag{3}$$

where $F \in R^{N \times D}$ represents the characteristics of the mesh cell, $act()$ represents the activation function, and W represents the learnable weight.

Therefore, GCN can extract local features of the intraoral under the guidance of the above mesh adjacency matrix. According to the fact that the normal vector can describe the geometric information of the 3D data, we feed the vertex normal vector and surface normal vector of the mesh cells into the GCN, thus simulating the generation process of the 3D surface through the message-passing mechanism.

4.2.2. Fast-TGCN

With the above analytical modeling, we propose a Fast Teeth Segmentation Graph Convolutional Network (Fast-TGCN). As shown in Fig. 9, Fast-TGCN uses the same 24-dimensional feature input as TSGCNet (the first 12 dimensions $\mathcal{F}_c$ represent the center coordinates of three vertexes and grid cells, and the last 12 dimensions $\mathcal{F}_n$ represent the normal vectors of vertexes and grid cells). And it adopts parallel stream processing with two branches, which has been proved to be a good representation of 3D oral data in TSGCNet. In addition, in order to construct the constraints of the adjacency graph, we use the method introduced in the previous section to regard the cells that have common vertexes with each cell as adjacent to each other, so as to construct an N-order binary adjacency matrix A. After that, the graph convolution blocks (GCBs) are used to compose the graph convolution network to transfer the normal vector feature along the path specified by the adjacency matrix A, so that all cells can obtain the direction information of the surrounding cells, and get the feature $\mathcal{F}_n'$:

$$GCB(\mathcal{F}_n) = \sigma\left(A \cdot Conv1D\left(\mathcal{F}_n\right) \cdot W\right), \tag{4}$$

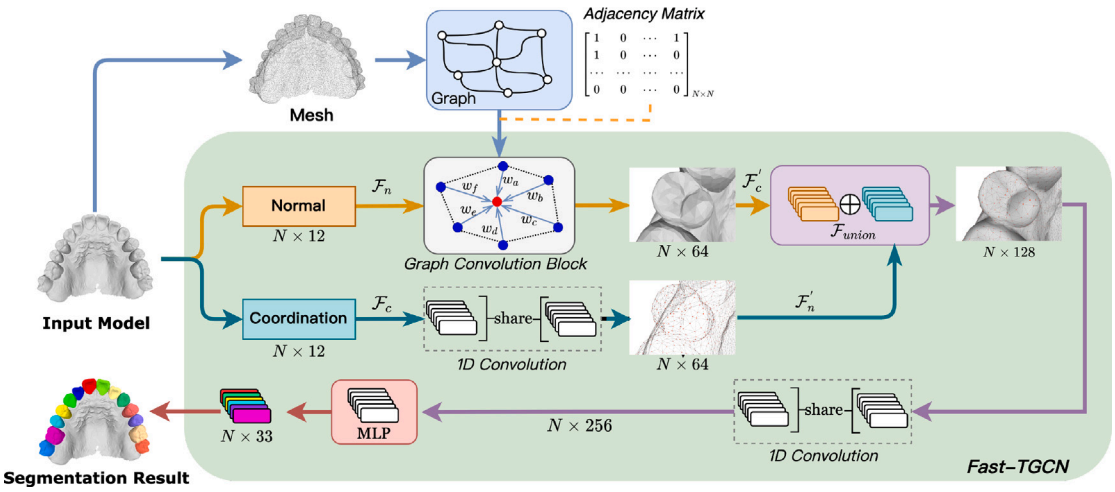

Fig. 9. Schematic diagram of the proposed Fast Tooth Graph Convolutional Network (Fast-TGCN).

$$\mathcal{F}_n' = GCB(\mathcal{F}_n). \tag{5}$$

The reason for performing graph convolution on the normal vector instead of the coordinates is that the coordinates of each cell are determined relative to other cells, which already contain all the information of its neighbors, so it does not make sense to perform graph convolution on coordinates. As a representation of the global features of the 3D data, Fast-TGCN only obtains the feature $\mathcal{F}_c$ through 1D convolution processing in parallel with GCB:

$$\mathcal{F}_c' = Conv1D\left(\mathcal{F}_c\right). \tag{6}$$

Meanwhile, we use the local surface feature $\mathcal{F}_n'$ obtained by the normal vector branch and the global position feature $\mathcal{F}_c'$ of the same dimension obtained by the coordinate branch splicing, get $\mathcal{F}_{union} = \mathcal{F}_n' \oplus \mathcal{F}_c'$, $\oplus$ represents tensor concatenation. The advantage of this is to maintain the independence of the two features, which is convenient for the network to learn features from different aspects. After fully connected prediction, we can obtain the final segmentation result.

5. Experiments

5.1. Experimental setup

In this work, we used 380 intraoral scan samples of the upper and lower jaws from 190 patients for training and the other 60 intraoral scan samples for testing. Like other 3D mesh-based tooth segmentation methods, we sample all samples in 3D-IOSSeg to 16,000 cells, which maximizes the retention of tooth details in the oral cavity and facilitates network learning. Meanwhile, we also added random translation and random rotation operations during training. Specifically, we consider the distribution of all training samples on the X, Y, and Z axes for random translation, and the moving intervals are $[-6, 6]$, $[-8, 8]$ and $[-5, 5]$, and randomly rotate along the Z axis between $\left[\frac{-\pi}{10}, \frac{\pi}{10}\right]$. In this way, the position and shape distribution of abnormal teeth can be effectively learned, which is conducive to improving the learning efficiency of the network and the accuracy of tooth segmentation.

This work is implemented under Python, and the neural network model is constructed through the Pytorch deep learning framework. Since the samples and labels in 3D-IOSSeg are stored in the same PLY file, they can be read directly through the PlyData dependency package in Python. Then use the pandas data processing code package to separate the coordinates, normal vectors and category labels of the samples. Finally, 200 Epochs are iterated for training with the support of the NVIDIA RTX 3090 graphics processing unit.

5.2. Evaluation metrics

We use Overall Accuracy (OA), mean class Accuracy (mAcc), mean Intersection-over-Union (mIoU), and mean Dice coefficient (mDice), four common performance measures on point cloud tasks, as the main evaluation metrics for dataset benchmarks. The calculation formula for each indicator is as follows:

$$OA = \frac{TP}{TP + FP + TN + FN}. \tag{7}$$

$$mAcc = \frac{1}{C}\sum_{c}^{C}\frac{TP_c}{TP_c + FN_c}. \tag{8}$$

$$mIoU = \frac{1}{C}\sum_{c}^{C}\frac{TP_c}{TP_c + FN_c + FP_c}. \tag{9}$$

$$mDice = \frac{1}{C}\sum_{c}^{C}\frac{2 \times TP_c}{FN_c + 2 \times TP_c + FP_c} \tag{10}$$

Among them, TP, TN, FP, and FN represent true positive, true negative, false positive, and false negative, respectively. Meanwhile, C is the total number of semantic categories, which equals 33 in this work.

5.3. Classic baseline methods

In order to comprehensively evaluate the performance of existing methods on abnormal teeth data, we choose 3 classic point cloud segmentation pipelines and 2 representative tooth segmentation methods as reliable baselines in this work.

PointNet [37]: This is an end-to-end network that can directly process point clouds. It encodes the features of the input points through MLP and then uses global pooling to extract the overall features of the point cloud. PointNet addresses the challenges posed by the unstructured characteristics of point clouds and is a pioneering work in 3D point clouds.

PointNet++ [42]: This network is an improvement based on PointNet, which adds multiple layers to focus on local and global features, and introduces farthest-point sampling to reduce computational cost.

Point Transformer (PT) [43]: This network introduces the Transformer into the 3D point cloud task. It uses the global extraction ability of the Transformer to obtain the geometric information of the 3D objects well.

MeshSegNet [39]: This is a tooth segmentation network for 3D intraoral scan data, which solves the defect that the traditional point cloud segmentation network cannot pay attention to the geometric details of the 3D data.

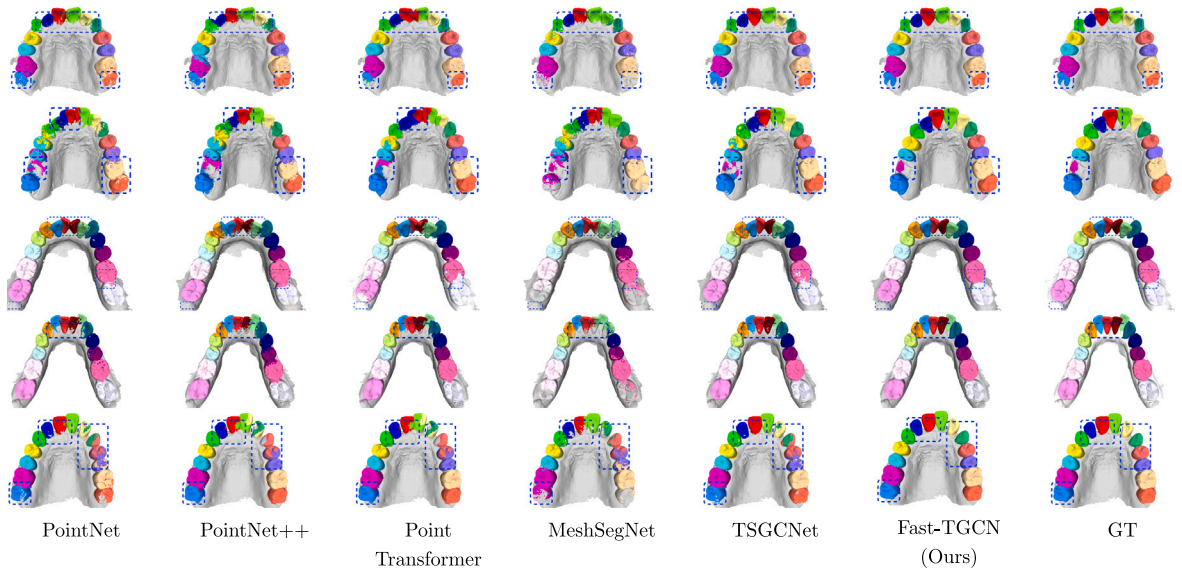

Fig. 10. Visualization of representative segmentation results produced by five competing methods and Fast-TGCN. The results in the first and second rows show that Fast-TGCN can effectively distinguish teeth with abnormalities; the results of the second to fifth rows show that Fast-TGCN does not confuse adjacent teeth, reflecting the effectiveness of the adjacency graph.

Table 3
Overall Accuracy (OA), mean class Accuracy (mAcc), mean Intersection-over-Union (mIoU), mean Dice coefficient (mDice), and score statistics of each tooth class IoU on the 3D-IOSSeg dataset for classic benchmark methods (under 5-fold cross-validation condition). Among them, the best results are highlighted.

Class name	Pointnet [37]	Pointnet++ [42]	Point Transformer [43]	MeshSegNet [39]	TSGCNet [40]	Fast-TGCN (Ours)
T11/T31	78.42%/81.01%	66.37%/53.64%	60.93%/55.81%	70.01%/72.07%	**88.00%**/90.06%	87.20%/**92.25%**
T12/T32	70.82%/77.71%	68.77%/61.43%	77.70%/62.06%	79.91%/80.43%	90.72%/84.87%	**91.04%/85.52%**
T13/T33	82.26%/75.60%	58.04%/58.18%	72.30%/74.12%	72.00%/78.79%	**94.32%/90.88%**	94.27%/90.60%
T14/T34	74.82%/78.82%	52.32%/63.83%	78.96%/74.60%	67.82%/74.06%	90.82%/**81.24%**	**91.40%**/79.54%
T15/T35	74.92%/83.42%	56.67%/59.80%	64.16%/67.55%	69.01%/74.90%	87.92%/**90.87%**	**90.49%**/89.28%
T16/T36	72.24%/81.11%	58.42%/66.79%	74.75%/79.83%	73.49%/78.56%	88.34%/82.46%	**90.92%/82.79%**
T17/T37	87.59%/81.44%	67.52%/64.37%	71.75%/74.26%	77.21%/78.58%	89.48%/**87.83%**	**93.64%**/86.53%
T18/T38	68.35%/71.08%	66.72%/60.11%	69.50%/71.17%	65.31%/70.10%	76.46%/79.92%	**80.64%/83.32%**
T21/T41	83.52%/79.48%	61.93%/57.88%	65.91%/58.77%	76.45%/71.13%	**85.48%**/89.66%	85.28%/**91.99%**
T22/T42	75.10%/79.63%	53.92%/50.54%	58.97%/52.44%	68.46%/64.57%	**84.84%**/85.51%	84.49%/**90.11%**
T23/T43	71.82%/77.19%	82.48%/64.92%	66.07%/73.38%	71.28%/76.86%	**91.03%**/87.37%	90.16%/**92.86%**
T24/T44	76.38%/80.14%	60.26%/67.42%	70.36%/74.85%	67.23%/78.44%	**92.29%**/86.37%	90.32%/**92.84%**
T25/T45	64.53%/73.92%	65.47%/58.45%	67.22%/68.77%	66.47%/65.03%	89.70%/82.98%	**89.83%/88.52%**
T26/T46	66.54%/79.04%	54.88%/64.59%	72.46%/75.85%	56.70%/56.20%	89.43%/83.58%	**93.76%/89.82%**
T27/T47	68.30%/79.37%	55.88%/61.31%	70.95%/66.12%	68.41%/58.31%	90.70%/86.92%	**92.52%/88.82%**
T28/T48	72.49%/71.61%	64.30%/60.58%	67.14%/69.05%	73.44%/72.69%	80.78%/78.21%	**81.60%/84.42%**
mIoU	76.42 ± 0.45%	64.83 ± 0.32%	71.42 ± 0.20%	73.32 ± 0.39%	87.73 ± 0.22%	**90.05 ± 0.40%**
mAcc	73.49 ± 0.26%	60.94 ± 0.31%	70.18 ± 0.12%	64.71 ± 0.23%	78.76 ± 0.38%	**80.71 ± 0.27%**
OA	90.07 ± 0.13%	81.66 ± 0.36%	88.53 ± 0.41%	86.47 ± 0.29%	94.81 ± 0.25%	**96.57 ± 0.41%**
Dice	84.32 ± 0.74%	76.91 ± 0.25%	81.74 ± 0.13%	80.06 ± 0.16%	92.10 ± 0.33%	**93.23 ± 0.39%**

TSGCNet [40]: This is an end-to-end network for 3D intraoral scan data segmentation. It ingeniously processes the coordinate and normal vector information of the 3D mesh data in parallel and constructs the attention convolution of the KNN graph to learn the two types of information.

5.4. Comparison and analysis

According to the above evaluation metrics, we evaluate the performance of these classic benchmarks on 3D-IOSSeg via 5-fold cross-validation. It is worth noting that since none of the datasets of existing work is publicly available, testing was not possible. In order to guarantee the accuracy of the evaluation, the above benchmarks are trained with the same method to ensure data fairness and result reliability. The scores of each method on mIoU, mAcc, OA, and Dice are reported in Table 3. Moreover, we provide the result of the mIoU score for each tooth. Compared to the results on other tooth segmentation datasets, the performance of all these classic models degraded significantly on the new dataset. Classical algorithms for 3D point cloud segmentation cannot recognize complex geometric features in dental data, limiting their performance on abnormal dental data. And because 3D-IOSSeg has a greater degree of abnormality and fine-grained division, these methods have obtained worse scores than other dental datasets. Meanwhile, we found that methods specifically designed for tooth segmentation also do not perform satisfactorily on the dataset. In the segmentation of regular 3D intraoral scanning data, TSGCNet demonstrates outstanding performance with an impressive OA (Overall Accuracy) score of 96.69% and mIoU (mean Intersection over Union) score of 91.19%. However, results on 3D-IOSSe show a decrease in performance, particularly with respect to the mIoU metric, where TSGCNet achieves only 87.73 ± 0.72%. Furthermore, MeshSegNet, as an advanced tooth segmentation method, also cannot effectively deal with large-scale abnormal teeth on 3D-IOSSeg. Different from the above methods, our Fast-TGCN is the first tooth segmentation method specially designed for abnormal teeth. According to the table, it can be observed that our method outperforms previous methods in terms of overall performance. In particular, the score on the overall mIoU performance index has a very obvious

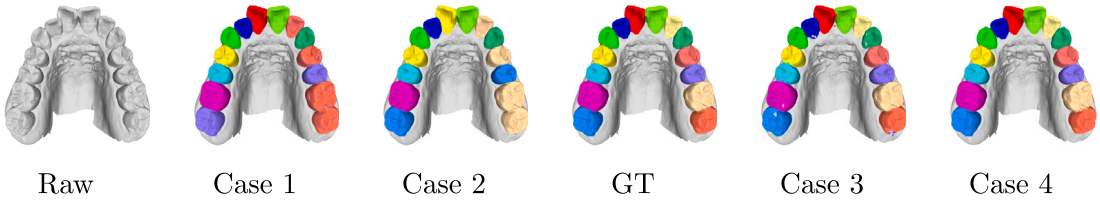

Fig. 11. Results of Fast-TGCN with different strategies.

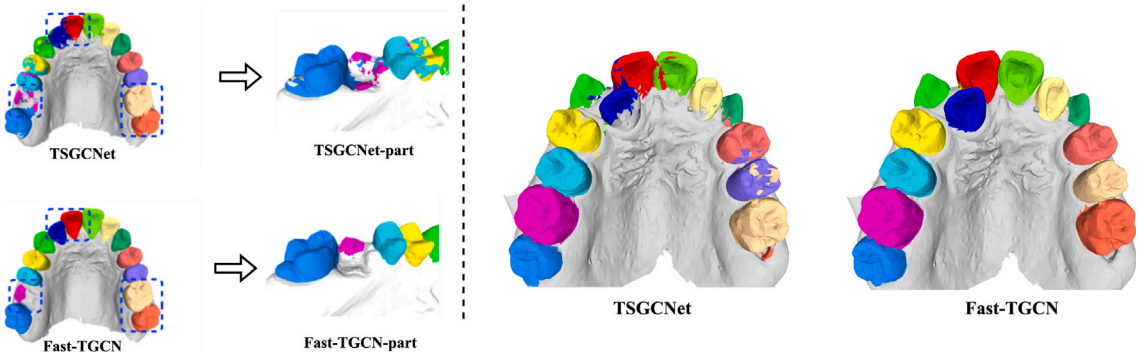

Fig. 12. Segmentation comparison between TSGCN and our Fast-TGCN on abnormal tooth samples. Obviously, our Fast-TGCN achieves better segmentation results. (The left part shows the segmentation results when the tooth is broken and the boundary between the tooth and gingiva is not clear, and the right part shows the segmentation result when the tooth is misaligned/overlapping.)

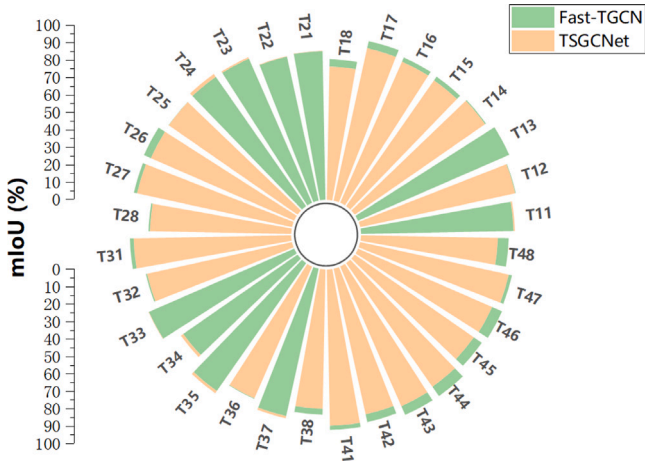

Fig. 13. Radial diagram of Fast-TGCN with TSGCNet. Longer bars represent better results.

Table 4
Evaluation results of all competing methods and Fast-TGCN on OA, mAcc and mIoU.

Matrix/Method	mAcc	mIoU	Para	Time	GPU
PointNet	73.49 ± 0.26%	76.42 ± 0.45%	1.82M	0.32 s	2457MiB
PointNet++	60.94 ± 0.31%	64.83 ± 0.32%	0.97M	0.64 s	2035MiB
PointTransformer	70.18 ± 0.12%	71.42 ± 0.20%	19.39M	1.18 s	3465MiB
MeshSegNet	64.71 ± 0.23%	73.32 ± 0.39%	1.79M	11.85 s	14655MiB
TSGCNet	78.76 ± 0.38%	87.73 ± 0.22%	4.13M	0.54 s	19967MiB
Fast-TGCN (Ours)	80.71 ± 0.27%	90.05 ± 0.40%	1.83M	0.35 s	12629MiB

Table 5
Performance comparison on normal samples.

Methods	mACC	mIoU	OA	Dice
TSGCNet	79.46 ± 0.24%	88.75 ± 0.10%	95.84 ± 0.56%	95.27 ± 0.28%
Fast-TGCN (Ours)	84.13 ± 0.21%	91.62 ± 0.33%	97.60 ± 0.23%	96.39 ± 0.19%

Table 6
Performance comparison on abnormal samples.

Methods	mACC	mIoU	OA	Dice
TSGCNet	74.70 ± 0.14%	80.36 ± 0.17%	87.21 ± 0.34%	89.31 ± 0.11%
Fast-TGCN (Ours)	78.13 ± 0.25%	83.09 ± 0.19%	89.60 ± 0.28%	90.14 ± 0.20%

advantage, which benefits from the natural advantages of the adjacency matrix in representing the complex geometric information of the tooth. It is worth noting that mACC represents the ratio of the number of correctly segmented cells to the number of all cells in the entire dataset. In 3D-IOSSeg, 60 dental samples are used for testing, and each sample contains 16000 cells, so the entire data set has 960,000 cells. Therefore, although our Fast-TGCN only improves mACC by 1.95% compared to TSGCNet, it means that our method has correctly segmented 18720 cells more than TSGCNet. This further validates the effectiveness of our proposed Fast-TGCN. In the Table 4, we also compare the inference cost of these segmentation methods. Obviously, Fast-TGCN achieves an excellent balance in execution time (time from loading data to completion of segmentation), GPU memory, model size and performance, which is of great significance for improving the efficiency of clinical diagnosis.

In Fig. 10, we provide the visual segmentation effect of all these methods in the 3D intraoral scan samples. From the results, we can see that our Fast-TGCN outperforms all other methods in segmenting abnormal teeth. Specifically: 1. Three classic point cloud segmentation methods, PointNet, PointNet++, and Point Transformer cannot accurately segment the tooth target. 2. The visualization results of MeshSegNet and TSGCNet in the second and third lines prove that the discontinuous area of the KNN graph caused the wrong segmentation. Both methods use the KNN graph constructed by the spatial Euclidean distance to guide the network segmentation. Although this method can build a dynamic graph to learn the 3D data, it will confuse the adjacent teeth when facing the fine-grained 3D-IOSSeg dataset full of abnormal teeth. In contrast, our Fast-TGCN can accurately distinguish the boundaries of adjacent teeth when dealing with such complex oral data, which proves the effectiveness and feasibility of our adjacency graph for modeling tooth surfaces.

Furthermore, we divided the dataset into two parts, one part only contains pure abnormal samples, and the other part contains pure

Fig. 14. Challenging cases of our proposed Fast-TGCN.

Table 7
Ablation study of graph convolutional operation and adjacency matrix graph strategy.

Case	Matrix/Case	OA	mAcc	mIoU
1	Coordinates	85.12%	71.88%	78.94%
2	Normal vector + Coordinates	86.69%	70.21%	76.52%
3	KNN graph	92.73%	75.68%	82.39%
4	Fast-TGCN (Ours)	**96.57%**	**80.71%**	**90.05%**

normal samples. Then, we use the pre-trained TSGCNet and Fast-TGCN to test these two dataset, and the results are indicated in the Tables 5 and 6, respectively. According to the result, we can clearly observe that our proposed Fast-TGCN achieves better results on normal and abnormal samples. This verifies the effectiveness and generalizability of the proposed Fast-TGCN. Meanwhile, we provide a visual comparison between TSGCN and our Fast-TGCN on abnormal tooth samples in Fig. 12. Among them, the upper part shows the segmentation results when the tooth is broken and the boundary between the tooth and gingiva is not clear, and the lower part shows the segmentation result when the tooth is misaligned/overlapping. Obviously, our Fast-TGCN achieves better segmentation results when confronted with samples exhibiting tooth damage, displacement, and atypical tooth morphology. This further verifies the effectiveness and generalizability of the proposed Fast-TGCN.

5.5. Ablation study

The key idea of the proposed Fast-TGCN is the introduced adjacency matrix graph strategy. To verify the effectiveness of this strategy, we designed a series of ablation experiments and provided the results in Table 7. Among them, case 1 represents the network that uses graph convolution only in the coordinates branch, and case 2 represents the use of graph convolution in both coordinates and normal vector branches. According to the result, we found that convolving coordinates alone cannot achieve good segmentation results, and convoluting two kinds of information simultaneously will also affect the segmentation accuracy. This is because the coordinates contain obvious global features, and their role is usually to place spatial constraints on the segmentation target, such as distinguishing between incisors and molars. However, for two types of teeth whose spatial distribution is significantly different, no such information is needed. For the normal vector, it only contains limited local features. If the network wants to use these features to segment the 3D data more carefully, it needs to obtain the overall surface information of a certain area. At this point, the graph convolution guided by the adjacency matrix graph can effectively connect all mesh cell normal vectors so that the network can perceive various curved features of the tooth surface. Meanwhile, in case 3, we also replace the GCB in our network with the KNN graph attention method in TSGCNet [40] for training and testing. According to the table, we can observe that the performance of the modified network drops significantly. In contrast, our Fast-TGCN achieves great results. This is because Fast-TGCN adopts the expression method of adjacency matrix graph, which greatly avoids the discontinuity problem of the KNN graph and transmits the surface information contained in the normal vector of adjacent mesh cells to the surrounding area. In Fig. 11, we also provide their segmentation results. This further verifies the effectiveness of our proposed method.

6. Discussion

The challenge of fine-grained segmentation. Although our Fast-TGCN achieves satisfactory results in 3D-IOSSeg, we found that the segmentation results of some teeth was not accurate due to the difference in the jaw. In addition, through Fig. 13, we found that the results of our Fast-TGCN were not the best on some teeth. This is mainly due to the lack of a constraint between teeth for the transmission of information in the adjacency graph, resulting in some information being diffused to other tooth regions. Fig. 14 displays some less satisfactory segmentation results. We observed that individual meshes are erroneously segmented in certain areas. Additionally, when faced with horizontally misaligned teeth, the current segmentation method requires further improvement. Since we employ a direct scanning approach, the oral environment in our samples is more intricate, including irregular gums, further emphasizing the greater complexity the 3D-IOSSeg dataset presents as a challenge for future oral segmentation. In future research, we intend to study tooth segmentation with a more refined network.

Simplified mesh data. Original 3D intraoral scan data are usually composed of more than 100,000 cells. It is impossible to perform directly on the original data with existing methods. The current solution is usually to simplify the original tooth data into a unified and smaller mesh data, but this also leads to the loss of some geometric details. In future work, we will explore a simplified algorithm for the mesh data, which enables they to perform directly on raw data.

7. Conclusion

In this work, we proposed a fine-grained tooth segmentation dataset 3D-IOSSeg dedicated to orthodontics, which not only has rich abnormal tooth distribution and finer-grained segmentation markers, but also uses real intraoral scan data. These characteristics make 3D-IOSSeg better reflect the diversity of clinical patients, and lay the foundation for the further application of tooth segmentation. Meanwhile, 3D-IOSSeg is a open and editable orthodontic dataset, which means that researchers can flexibly adjust each sample through 3D view software to suit specific works. In addition, we proposed a fast graph convolutional network (Fast-TGCN) for tooth segmentation. Different from previous tooth segmentation methods, Fast-TGCN replaces the KNN graph with a more naive adjacency matrix graph and use the node message passing mechanism of GCN to gradually transmit the generation process of

the simulated 3D surfaces to the entire space. This method has a significant impact on the geometric information modeling of the 3D intraoral scans, greatly improving the efficiency of tooth segmentation. Extensive results show that Fast-TGCN has strong advantages in fine-grained abnormal tooth segmentation, laying the foundation for digital dentistry.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work is supported in part by the National Key R&D Program of China under Grant 2021YFA1003004, in part by the National Natural Science Foundation of China under Grant 62301306, in part by the Natural Science Foundation of Shanghai, China under Grand 23ZR1422200, in part by the Shanghai Sailing Program under Grant 23YF1412800, and in part by the NSFC/RGC under Grant N_CUHK 415/19. All authors approved the version of the manuscript to be published.

References

[1] Y. Wang, L. Wu, H. Guo, T. Qiu, Y. Huang, B. Lin, L. Wang, Computation of tooth axes of existent and missing teeth from 3D CT images, Biomed. Eng./Biomed. Tech. 60 (6) (2015) 623–632.
[2] G. Wei, Z. Cui, Y. Liu, N. Chen, R. Chen, G. Li, W. Wang, TANet: towards fully automatic tooth arrangement, in: Proceedings of the European Conference on Computer Vision, 2020.
[3] B. Silva, L. Pinheiro, L. Oliveira, M. Pithon, A study on tooth segmentation and numbering using end-to-end deep neural networks, in: Conference on Graphics, Patterns and Images, 2020.
[4] Y. Zhao, P. Li, C. Gao, Y. Liu, Q. Chen, F. Yang, D. Meng, TSASNet: Tooth segmentation on dental panoramic X-ray images by Two-Stage Attention Segmentation Network, Knowl.-Based Syst. 206 (2020) 106338.
[5] M.P. Muresan, A.R. Barbura, S. Nedevschi, Teeth detection and dental problem classification in panoramic X-ray images using deep learning and image processing techniques, in: International Conference on Intelligent Computer Communication and Processing, IEEE, 2020, pp. 457–463.
[6] K. Panetta, R. Rajendran, A. Ramesh, S.P. Rao, S. Agaian, Tufts dental database: a multimodal panoramic x-ray dataset for benchmarking diagnostic systems, IEEE J. Biomed. Health Inf. 26 (4) (2021) 1650–1659.
[7] G. Chandrashekar, S. AlQarni, E.E. Bumann, Y. Lee, Collaborative deep learning model for tooth segmentation and identification using panoramic radiographs, Comput. Biol. Med. 148 (2022) 105829.
[8] S. Lin, X. Hao, Y. Liu, D. Yan, J. Liu, M. Zhong, Lightweight deep learning methods for panoramic dental X-ray image segmentation, Neural Comput. Appl. 35 (11) (2023) 8295–8306.
[9] Y. Gan, Z. Xia, J. Xiong, G. Li, Q. Zhao, Tooth and alveolar bone segmentation from dental computed tomography images, IEEE J. Biomed. Health Inf. 22 (1) (2017) 196–204.
[10] Z. Cui, C. Li, W. Wang, ToothNet: automatic tooth instance segmentation and identification from cone beam CT images, in: Proceedings of the IEEE Conference on Computer Vision and Rattern Recognition, 2019.
[11] M. Chung, M. Lee, J. Hong, S. Park, J. Lee, J. Lee, I.-H. Yang, J. Lee, Y.-G. Shin, Pose-aware instance segmentation framework from cone beam CT images for tooth segmentation, Comput. Biol. Med. 120 (2020) 103720.
[12] Z. Cui, Y. Fang, L. Mei, B. Zhang, B. Yu, J. Liu, C. Jiang, Y. Sun, L. Ma, J. Huang, et al., A fully automatic AI system for tooth and alveolar bone segmentation from cone-beam CT images, Nature Commun. 13 (1) (2022) 1–11.
[13] P. Li, Y. Liu, Z. Cui, F. Yang, Y. Zhao, C. Lian, C. Gao, Semantic graph attention with explicit anatomical association modeling for tooth segmentation from CBCT images, IEEE Trans. Med. Imaging (2022).
[14] S. Gao, X. Li, X. Li, Z. Li, Y. Deng, Transformer based tooth classification from cone-beam computed tomography for dental charting, Comput. Biol. Med. 148 (2022) 105880.
[15] D. Sun, Y. Pei, G. Song, Y. Guo, G. Ma, T. Xu, H. Zha, Tooth segmentation and labeling from digital dental casts, in: International Symposium on Biomedical Imaging, 2020.
[16] L. Qiu, C. Ye, P. Chen, Y. Liu, X. Han, S. Cui, DArch: Dental arch prior-assisted 3D tooth instance segmentation with weak annotations, in: Proceedings of the IEEE Conference on Computer Vision and Rattern Recognition, 2022.
[17] Z. Liu, X. He, H. Wang, H. Xiong, Y. Zhang, G. Wang, J. Hao, Y. Feng, F. Zhu, H. Hu, Hierarchical self-supervised learning for 3D tooth segmentation in intra-oral mesh scans, IEEE Trans. Med. Imaging (2022).
[18] T.-H. Wu, C. Lian, S. Lee, M. Pastewait, C. Piers, J. Liu, F. Wang, L. Wang, C.-Y. Chiu, W. Wang, et al., Two-stage mesh deep learning for automated tooth segmentation and landmark localization on 3D intraoral scans, IEEE Trans. Med. Imaging (2022).
[19] Y. Kumar, R. Janardan, B. Larson, J. Moon, Improved segmentation of teeth in dental models, Comput.-Aided Des. Appl. 8 (2) (2011) 211–224.
[20] T. Yuan, W. Liao, N. Dai, X. Cheng, Q. Yu, Single-tooth modeling for 3D dental model, Int. J. Biomed. Imaging (2010).
[21] M. Zhao, L. Ma, W. Tan, D. Nie, Interactive tooth segmentation of dental models, in: IEEE Engineering in Medicine and Biology Conference, 2006.
[22] T. Kronfeld, D. Brunner, G. Brunnett, Snake-based segmentation of teeth from virtual dental casts, Comput.-Aided Des. Appl. 7 (2) (2010) 221–233.
[23] K. Wu, L. Chen, J. Li, Y. Zhou, Tooth segmentation on dental meshes using morphologic skeleton, Comput. Graph. 38 (2014) 199–211.
[24] C. Sinthanayothin, W. Tharanont, Orthodontics treatment simulation by teeth segmentation and setup, in: International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology, 2008.
[25] M. Yaqi, L. Zhongke, Computer aided orthodontics treatment by virtual segmentation and adjustment, in: International Conference on Image Analysis and Signal Processing, 2010.
[26] B.-j. Zou, S.-j. Liu, S.-h. Liao, X. Ding, Y. Liang, Interactive tooth partition of dental mesh base on tooth-target harmonic field, Comput. Biol. Med. 56 (2015) 132–144.
[27] T. Kondo, S.H. Ong, K.W. Foong, Tooth segmentation of dental study models using range images, IEEE Trans. Med. Imaging 23 (3) (2004) 350–362.
[28] N. Wongwaen, C. Sinthanayothin, Computerized algorithm for 3D teeth segmentation, in: International Conference on Electronics and Information Engineering, 2010.
[29] D. Maturana, S. Scherer, Voxnet: A 3d convolutional neural network for real-time object recognition, in: International Conference on Intelligent Robots and Systems, 2015.
[30] C.R. Qi, H. Su, M. Nießner, A. Dai, M. Yan, L.J. Guibas, Volumetric and multi-view cnns for object classification on 3d data, in: Proceedings of the IEEE Conference on Computer Vision and Rattern Recognition, 2016.
[31] P.-S. Wang, Y. Liu, Y.-X. Guo, C.-Y. Sun, X. Tong, O-cnn: Octree-based convolutional neural networks for 3d shape analysis, ACM Trans. Graph. 36 (4) (2017) 1–11.
[32] G. Riegler, A. Osman Ulusoy, A. Geiger, Octnet: Learning deep 3d representations at high resolutions, in: Proceedings of the IEEE Conference on Computer Vision and Rattern Recognition, 2017.
[33] B. Graham, M. Engelcke, L. Van Der Maaten, 3D semantic segmentation with submanifold sparse convolutional networks, in: Proceedings of the IEEE Conference on Computer Vision and Rattern Recognition, 2018.
[34] Z. Wang, F. Lu, VoxSegNet: Volumetric CNNs for semantic part segmentation of 3D shapes, IEEE Trans. Vis. Comput. Graphics 26 (9) (2019) 2919–2930.
[35] X. Xu, C. Liu, Y. Zheng, 3D tooth segmentation and labeling using deep convolutional neural networks, IEEE Trans. Vis. Vomputer Graph. 25 (7) (2018) 2336–2348.
[36] S. Tian, N. Dai, B. Zhang, F. Yuan, Q. Yu, X. Cheng, Automatic classification and segmentation of teeth on 3D dental model using hierarchical deep learning networks, IEEE Access 7 (2019) 84817–84828.
[37] C.R. Qi, H. Su, K. Mo, L.J. Guibas, Pointnet: Deep learning on point sets for 3d classification and segmentation, in: Proceedings of the IEEE Conference on Computer Vision and Rattern Recognition, 2017.
[38] F.G. Zanjani, D.A. Moin, B. Verheij, F. Claessen, T. Cherici, T. Tan, et al., Deep learning approach to semantic segmentation in 3D point cloud intra-oral scans of teeth, in: International Conference on Medical Imaging with Deep Learning, 2019, pp. 557–571.
[39] C. Lian, L. Wang, T.-H. Wu, F. Wang, P.-T. Yap, C.-C. Ko, D. Shen, Deep multi-scale mesh feature learning for automated labeling of raw dental surfaces from 3D intraoral scanners, IEEE Trans. Med. Imaging 39 (7) (2020) 2440–2450.
[40] Y. Zhao, L. Zhang, Y. Liu, D. Meng, Z. Cui, C. Gao, X. Gao, C. Lian, D. Shen, Two-stream graph convolutional network for intra-oral scanner image segmentation, IEEE Trans. Med. Imaging 41 (4) (2021) 826–835.
[41] M. Welling, T.N. Kipf, Semi-supervised classification with graph convolutional networks, in: International Conference on Learning Representations, 2016.
[42] C.R. Qi, L. Yi, H. Su, L.J. Guibas, Pointnet++: Deep hierarchical feature learning on point sets in a metric space, in: Conference on Neural Information Processing Systems, 2017.
[43] H. Zhao, L. Jiang, J. Jia, P.H. Torr, V. Koltun, Point transformer, in: Proceedings of the IEEE Conference on International Conference on Computer Vision, 2021.