

# Dense Representative Tooth Landmark/axis Detection

## Network on 3D Model 复现报告

Guangshun Wei, Zhiming Cui, Jie Zhu, Lei Yang, Yuanfeng Zhou, Pradeep Singh, Min Gu, Wenping Wang

Fellow, IEEE

### 摘要

人工智能技术越来越多地用于数字化正畸，但其中一个挑战是自动准确地检测牙齿的关键点和牙轴。一部分是由于牙齿复杂的几何定义，还有部分由于单个牙齿和不同类型的牙齿之间存在很大的差异。该论文提出一种深度学习方法，以三维牙齿模型作为输入，并预测各种类型的牙关键点和牙轴。具体来说，即将关键点和牙轴编码为定义在牙齿模型表面的稠密场，从中提取稀疏关键点和牙轴。实验结果表明，该论文方法能够生成具有较高精度的牙齿关键点和牙轴。

**关键词：**稀疏特征；牙齿关键点；牙轴；牙齿治疗

## 1 引言

错位咬合是指颌骨关闭时牙齿的不对齐。严重者可导致不适感，甚至出现言语或呼吸问题。为了治疗这种疾病，牙科治疗的主要目标是根据某些经验规则和条件，如 Andrews Six Keys<sup>[1]</sup>，使牙齿对齐。这些规则和条件被认为是实现正常咬合从而成功地进行正畸治疗的核心。

简而言之，这样的充分条件可以有效地表示为在正常咬合 (见图 1) 中应该满足的预定义的牙齿关键点和牙轴之间的一组期望的关系。例如，每个牙齿在正常咬合时的接触点可以用于相邻牙齿与自身在当前构型 (如图 1 (d)) 上的对齐，而尖点与咬合点对及其对侧颌骨上的对应点对之间的关系则表示上下颌骨的牙齿尖点-牙槽、牙齿尖点-牙槽骨或牙齿尖点-咬合关系。

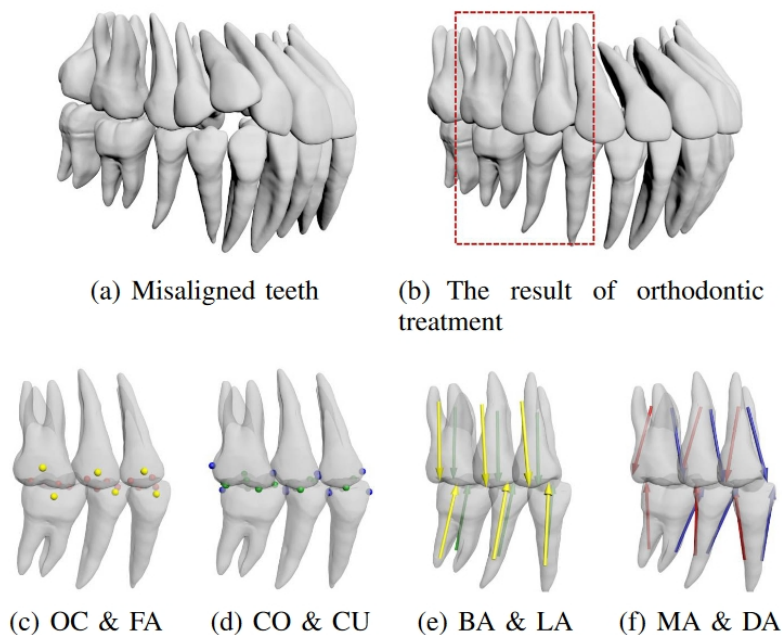


图 1: 使用 Andrews ' Six Keys 中定义的牙齿标志点和坐标轴来表征当前牙齿构型偏离正常咬合的程度

随着自动化技术的普及，尤其是数字化诊断辅助设备和定制设备的使用无托槽隐形矫治器，对人工智能技术优化当代正畸治疗计划的需求越来越高。现有的许多方法被提出，以提高正畸治疗的自动化水平，节省临床医生的时间。然而，对于定义正常咬合至关重要的前述牙齿关键点和牙轴的检测在很大程度上留给经验丰富的牙科医生，并由人工进行。仅有少数研究探讨了自动检测牙体标志点进行牙齿排列的可能性。为了提高牙齿排列的自动化水平，该论文旨在提供一种自动的、数据驱动的方法，从不同类型的牙齿中提取 Andrews Six Keys 中使用的关键点和牙轴，如图 1 所示，红色、黄色、蓝色和绿色点分别代表咬合点 (OC)、面轴点 (FA)、接触点 (CO) 和尖点 (CU)。绿色、黄色、蓝色和红色箭头分别代表舌轴 (LA)、颊面轴 (BA)、近中轴 (MA) 和远中轴 (DA)。(a) 表示错位齿的状态。(b) 表示基于牙齿标志点和轴的正畸治疗后所需的排列方式。(c) (f) 显示了牙齿标志点和轴在正畸中的作用和意义。并列于表 I。

接触点 (CO)	为相邻两牙近中面的接触点。
尖点 (CU)	牙齿牙冠上圆锥形突起的端部。
面中轴点 (FA)	中心点/牙齿面中点 (近嘴唇或脸颊)。
咬合点 (OC)	上下颌骨闭合性咬合时，上下颌牙之间正常接触的点。
颊面中轴 (BA)	沿着与唇部或颊部相邻的牙齿表面的斜率画出的一条切线。
舌轴 (LA)	沿牙面面向舌面，平行于牙体长轴画出一条切线。
近中轴 (MA)	沿朝向 cen 的近端表面画一条切线。

表 1: 牙齿标志点和轴线的描述

一个主要的挑战是，在给定的牙齿模型上，大多数牙齿关键点和牙轴不对应任何尖锐的几何形状。例如，接触点位于齿面 (见图 1) 的光滑区域，所有牙轴都定义为齿的全局和局部几何形状。现有研究方法局限于检测具有独特几何形状的关键点或被设计用于一般几何处理任务的识别或分类，不足以定义正常的咬合关系，且给定网格模型的稀疏信号时，它们的精度较低，因此对于正畸治疗来说是不切实际的。

为了解决上述问题，该论文提出了一种新的神经网络，用于在 3D 牙齿模型上可靠地检测牙齿关键点和牙轴。因此，给定三维牙齿表面上的一个牙齿标志点，该论文首先将其转换为定义在牙齿模型表面上的测地距离场。因此，牙齿关键点的定位可以首先转化为在给定的牙齿表面预测一个距离场，然后检测距离场的峰值。类似地，对于给定的牙轴，用一束投影向量来表示它，其中每个投影向量将曲面上的关联点投影到牙轴上。这些稠密的表示在论文中统称为点域编码。考虑到牙齿的关键点和牙轴同时依赖于牙齿的局部和全局几何结构，因此该论文利用多尺度机制，通过在输入点云上设置不同的尺度半径并聚合它们来学习潜在特征表示，以提取分层的潜在特征。此外，该论文使用了一个增强模块，将空间坐标、输入点的法线和提取的潜在特征串联起来，以进一步提高预测结果。

该论文的贡献总结如下：

- 提出了一种新颖的框架，可以从三维牙齿模型中稳健地检测牙齿关键点和牙轴。
- 提出了一种新的机制，将稀疏的几何实体 (即点和轴) 编码成稠密的表示，并进一步采用多

尺度和增强模块，在此稠密表示的基础上学习可靠的潜在特征。

## 2 相关工作

近年来，人们提出了许多数字化方法和工具来促进数字化牙科，包括牙齿分割<sup>[2]</sup>、牙科诊断<sup>[3]</sup>、修复<sup>[4]</sup>、正畸治疗过程<sup>[5]</sup>等。

### 2.1 点云数据的深度学习

大多数现有的工作在特征学习之前将点云转换为图像或体积形式。然而，这些基于 CNN 的方法仍然需要对具有体素分辨率的点云进行量化，对硬件设备要求较高。最近，PointNet<sup>[6]</sup>提出将原始点云作为输入的 DNN 用于分类和分割，为处理点云提供了有趣的理论见解。随后，PointNet++<sup>[7]</sup>作为点网络的改进版本被提出，使用了分层特征学习和多尺度点云卷积操作。随后，基于点云的深度神经网络的其他模型也相继被提出。此外，随着计算机图形学在深度学习中的发展，许多方法试图将点集连接成图并提取潜在特征。此外，基于点云学习的许多领域都受到了广泛的关注，如法向量估计、点云配准、点云超分辨率以及点云去噪等。

### 2.2 关键点检测

关键点检测的方法有很多，包括传统方法和较新的数据驱动方法。传统的方法大多利用曲率或表面法线等局部几何信息进行关键点检测。这条工作路线假设关键点与几何变化有关，甚至是由几何变化定义的，因此可以很好地检测机械零件中的尖锐特征。然而，这些方法不能鲁棒地处理具有较大形状变化的复杂三维形状 (例如, 牙齿)。一种替代方法是数据驱动的方法。对于 2D 图像，关键点检测和手部通常将稀疏路标转化为热力图并将其设置为监督学习的目标。对于 3D 模型，有监督算法预测定义在曲面上的概率分布。然而，这些方法通常会受到许多伪影的影响，这归因于对局部显著特征和局部几何变化的依赖，特别是当牙齿特征不突出时。

### 2.3 牙齿标志点和轴线检测

近年来，数字化正畸已成为研究热点。然而，由于复杂的排齐规则和患者个性化的功能美学的约束，错牙合畸<sup>[8]</sup>形给正畸医生带来了挑战。牙齿排列不齐的及时有效治疗依赖于标志点和坐标轴检测的质量。Andrews' Six Keys<sup>[1]</sup>的构想是最早提出的牙齿相关特征。然后，随着数字化技术和计算机辅助牙齿移动的出现，临床医生已经能够进行全面和仔细的评估，从而提高了美学和改善了治疗效果。然而，这些方法都是手动或半自动的。由于数据采集的困难和齿面特征不明显，现有的方法都不能完全检测出所有的特征点和坐标轴。在本文中，提出了一种基于深度学习的全自动、高效的关键点和坐标轴检测方法。

## 3 本文方法

### 3.1 本文方法概述

网络由三个主要部分组成，用于预测牙齿标志点和轴的密集表示，如图 2 所示，包括一个多尺度潜在特征提取模块和两个子网络，分别用于牙轴和关键点检测。以基于数据处理的三维牙齿模型的点云作为输入，通过多尺度特征提取初步提取特征向量。然后将提取的特征图同时传递给两个分支子网络。然后通过两个子网络对距离场和投影向量进行预测。最后，通过特征增强模块和后处理步骤分别得到牙齿关键点和牙轴。

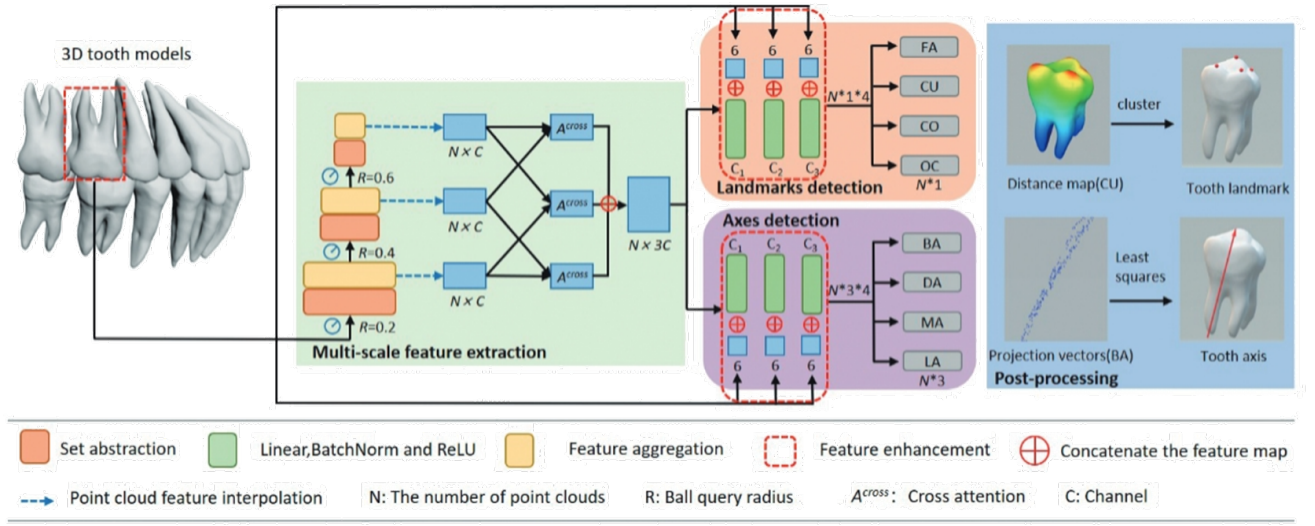


图 2: 该论文方法

### 3.2 多尺度潜在特征提取模块

虽然该论文方法是针对两个不同的任务 (即牙齿关键点和牙轴线检测) 设计的, 但从几何角度来看, 它们是内在关联的。因此, 给定一个三维牙齿模型的点云, 如图 3 所述。首先通过多尺度潜在特征提取模块提取公共点特征。由于牙齿标志点和轴线的定位依赖于牙冠几何形状和整体牙齿形状, 同时需要局部信息及整体信息。例如, 尖点主要取决于局部齿冠几何形状, 而牙轴主要取决于整体齿形信息。

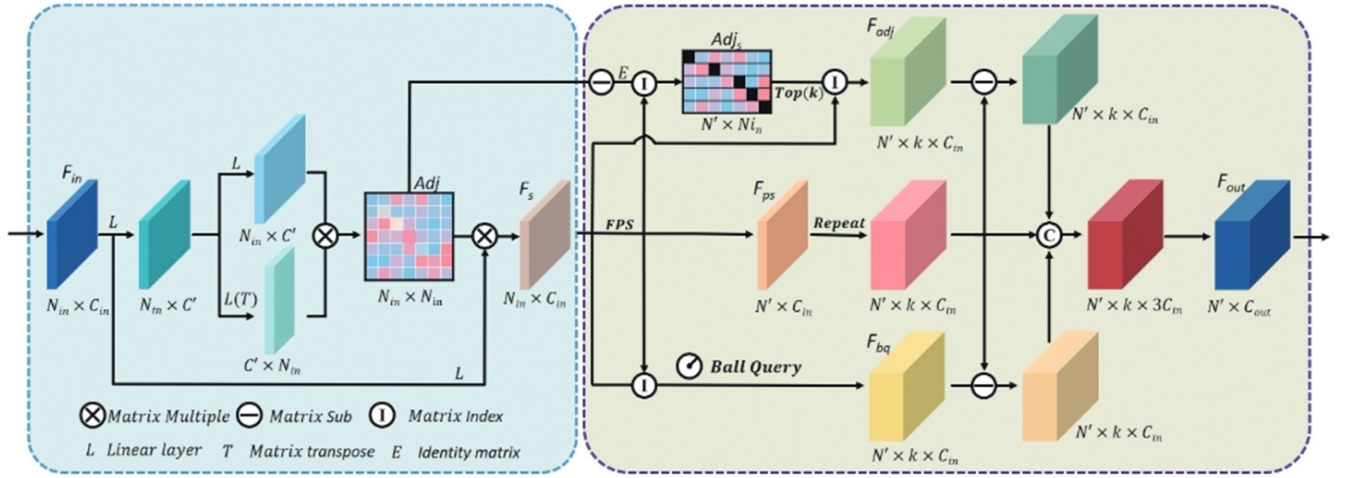


图 3: 点云特征提取与特征聚合模块。FSP 是最远点采样策略。

如图 2 所示, 首先采用多尺度特征提取方案来捕获不同半径的局部信息。在此过程中, 论文提出了一种新颖的特征聚合模块, 该模块同时聚合了点云在局部空间和非局部空间的特征。如图 3 所示, 给定输入特征图  $F_{in}$ , 经过特征提取层  $L$ , 由  $\text{Softmax}(L(F_{in}) \times L(F_{in})^T)$  得到一个邻接矩阵  $Adj$ 。  $Adj$  中的每个元素表示  $F_{in}$  中不同点特征的相似度。然后通过  $Adj \times L(F_{in})$  计算捕获非局部形状信息的新特征映射  $F_s$ 。随后, 通过最远点采样 (FPS) 策略得到具有逐点特征的子采样点云  $p_s$ , 相应的邻接矩阵由  $Adj_s = \text{Index}(Adj - E)$  计算得到, 其中  $E$  为单位矩阵。对于  $p_s$  上的每个点, 该论文根据邻接矩阵  $adj_s$  的每一行选择  $\text{Top}(k)$  点, 以获得聚合特征  $F_{adj}$ 。此外, 为了捕获局部空间中的特征, 对于  $p_s$  上的每个点, 采用 Ball Query 在半径  $r$  内搜索最近的  $k$  个点, 以获得聚合特征  $F_{bp}$ 。最后, 将子采样点

云 $p_s$ 的原始特征 $F_{ps}$ 、非局部特征 $F_{adj}$ 和局部特征 $F_{bp}$ 组合在一起进行最终预测。

$$\begin{aligned} F_{cat} &= \text{concat}(F_{ps}, F_{adj} - F_{ps}, F_{bp} - F_{ps}), \\ F_{out} &= MP(MLP(F_{cat})), \end{aligned} \quad (1)$$

式中： $MP$ 为最大池化算子， $LP$ 为多层感知器。

在每个特征聚合模块之后，通过插值得到与输入点云数量相同的分层特征。为了进一步捕获点云的形状和上下文信息，在不同尺度下应用交叉注意力层来提取多尺度特征，定义如下：

$$A^{\text{cross}}(x, y, y) = \alpha \left( \frac{QK^T}{\sqrt{d_k}} \right) V, \quad \begin{cases} Q = \text{Conv1D}(x) \\ K = \text{Conv1D}(y) \\ V = \text{Conv1D}(y) \end{cases} \quad (2)$$

式中： $x, y$ 为分层特征， $dk$ 为增加稳定性的维度 $k$ 。最后，将提取的所有特征拼接在一起，然后将解码后的逐点特征输入到以下两个子网络中，分别预测距离场和投影向量场。

### 3.3 特征增强模块

在共同的点特征基础上，采用两个任务特有的模块来预测牙齿的轴线和标志点。如图2所示，由于牙齿标志点和轴的检测对空间坐标和法线敏感，论文提出了一个特征增强模块，将输入点坐标和法线与提取的潜在特征进行级联。具体来说，绿色框表示多尺度特征提取模块提取的潜在特征，蓝色框表示输入点的坐标信息和法向信息。然后将这两个信息作为LBR模块(Linear、BatchNorm和ReLU)的输入，最终预测关键点的距离场和轴的投影向量场。

### 3.4 损失函数

#### 3.4.1 距离场预测：

经过特征增强模块后，输出齿面上不同的距离场。该论文利用均方误差作为损失函数来监督训练过程，具体如下：

$$\mathcal{L}_p = \frac{\sum_{i=1}^N (D_i - D'_i)^2}{N} \quad (3)$$

其中 $D'_i$ 为点 $p_i$ 的预测距离值， $D_i$ 为真值， $N$ 为点数。

#### 3.4.2 投影向量场预测

牙轴线的逐点场编码体现在齿面上的点到其轴线的位移。该论文的牙轴检测网络架构如图3所示。以点特征为输入，预测齿面与牙轴的映射位移关系。最后得到不同的投影向量。该论文计算真值对应点与预测投影向量之间的L2距离作为损失函数：

$$\mathcal{L}_{axis} = \frac{1}{N} \sum_{i=1}^N \|v(p_i) - v(p'_i)\| \quad (4)$$

式中 $v(p_i)$ 为 $p_i$ 处的真值投影向量， $v(p'_i)$ 为其估计结果。

## 4 复现细节及结果

### 4.1 与已有开源代码对比

本次复现的论文在特征提取与增强部分使用PintNet++的开源代码来实现，其余数据预处理、关键点检测、牙轴检测及结果可视化部分代码未开源。



## 4.2 数据预处理

为了获得适合特定任务的牙齿三维模型，同时利用口内扫描数据和 CBCT 数据重建单个牙齿的三维模型。CBCT 数据提供了牙齿 (包括牙冠和牙根) 的全面三维信息，适用于三维牙齿模型的获取。然而，基于 CBCT 数据重建的牙齿模型丢失了牙冠的精细几何细节。另一方面，口内扫描数据可以捕获牙齿牙冠的高精度几何细节。对于牙齿标志点检测，大部分牙齿标志点位于牙冠区域。对于牙轴检测，它更注重齿的整体形状。因此，结合两种数据来获得最终的高精度三维牙齿模型。

首先通过 RANSAC 进行全局配准，然后使用 ICP 对全局配准结果进行细化。配准后的点云在 blender 中渲染如所示，其中黄色点云为 CBCT 数据，蓝色点云为口扫数据。

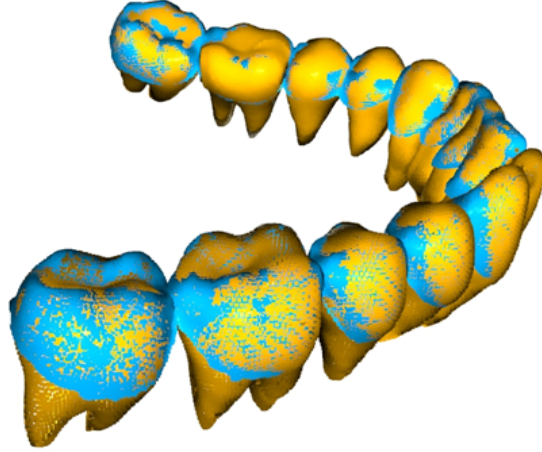


图 4: CBCT、口扫数据预处理结果

## 4.3 牙齿关键点检测

与其他几何关键点检测问题不同的是，牙齿定义的关键点大多没有显著的局部几何特征，如图 1 所示。例如，面轴点位于牙齿颊侧的平滑区域，该区域不存在尖锐特征。在这种情况下，带有目标坐标的直接回归方法提供了不充分的监督，通常会导致模棱两可的预测。受图像域人脸特征点检测的启发，研究人员使用热图回归策略来减少这种歧义。该论文将这一思想扩展到 3D 领域，通过对牙齿模型的每个点进行关键点编码来实现更好的性能。对于关键点检测，该论文通过计算每个点到地面的测

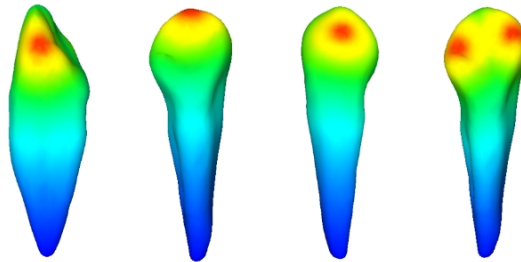


图 5: 距离场对应不同的标志点，其中依次代表接触点、尖点、面轴点和咬合点。

地距离<sup>[9]</sup>，定义一个距离场  $D$  作为真值  $P_i$  到 3D 牙齿网格上的关键点  $f_{pj}$  如下：

$$D_i = \exp \left( -\frac{(G(f_{pj}, p_i))^2}{2\sigma^2} \right) \quad (5)$$

其中  $G(\cdot)$  函数旨在计算牙齿表面两点的测地距离，该论文的方法中  $\sigma$  取为 0.3。该论文通过公式得到最终的距离场  $D$ 。如果牙齿模型上有多个标志点，则对计算得到的多个距离场进行组合，为每个点选择最大的距离值。如图 5 所示，点的测地距离越小，di (红色) 越大，反之亦然。

给定每个牙齿的预测距离场，首先获得一些距离值相对较大的部分点。换句话说，距离场可以看作是一个置信场，其中一个点的值越大，该点上的关键点可信度越大。然后设定一个阈值过滤掉这些置信度低的点。最后，根据牙齿模型上的关键点数量，将聚类中心作为关键点。

本次复现任务选取前磨牙分别检测接触点 (OC) 和面中轴点 (FA)，结果如下图所示。牙齿面片中间的点为面中轴点，两个牙齿之间的点位接触点。

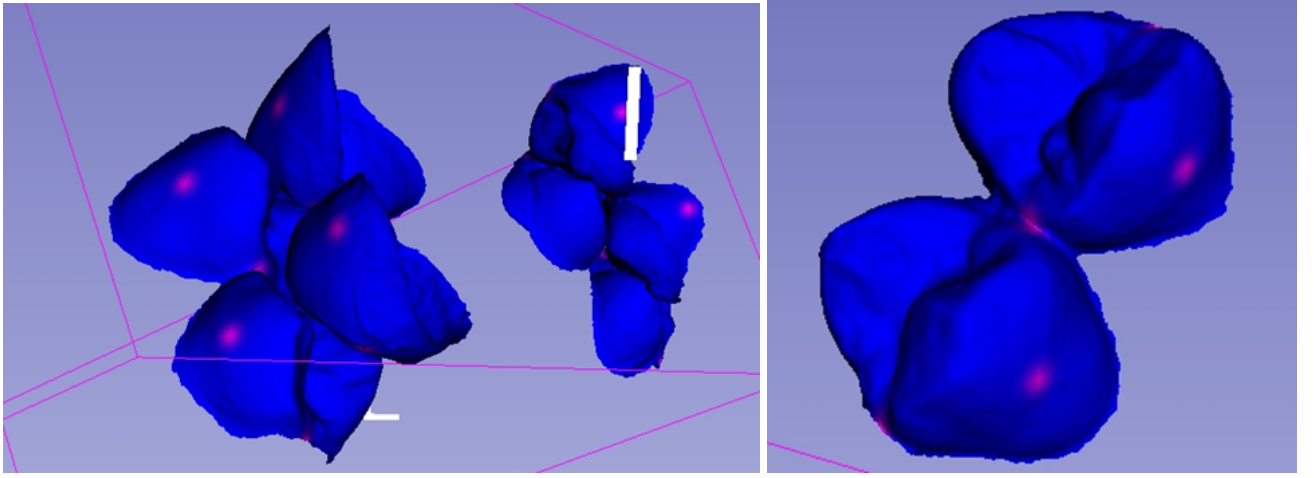


图 6: OC 和 FA 检测结果

#### 4.4 牙轴检测

齿轴检测与齿标检测类似，齿轴的适当稠密表示也可以在训练过程中提供更丰富的监督信息，对异常值具有鲁棒性，相比于直接进行轴回归。

对于齿轴检测任务，我们首先根据给定的线性方程定义齿轴的稠密投影向量场：

$$p_{\text{center}} = \frac{\sum_{i=1}^N p_i}{N}, \quad (6)$$

$$\mathbf{l} = p_{\text{center}} + \mathbf{n} \cdot t$$

其中  $p_{\text{center}}$  为齿的中心点， $\mathbf{n}$  为齿轴线的单位矢量， $\mathbf{l}$  为齿轴线与模数  $t$  的直线。然后将齿面上所有点  $p_i$  垂直投影到该齿轴线  $\mathbf{l}$  上，得到相应的投影向量作为监督信号。换言之，如图 7所示，齿轴检测网络的输出即为点云的投影坐标。

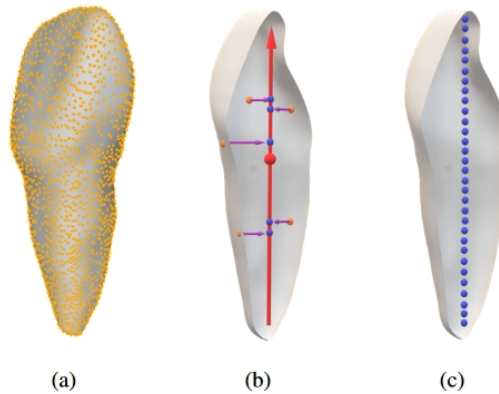


图 7: 齿轴检测的逐点场编码，其中 (a) 为三维牙齿模型，(b) 表示获取投影向量的过程，(c) 为每个点对应的最终编码特征。橙色圆点代表牙齿表面的点云。红色圆点表示齿的中心点。红色箭头表示穿过的牙轴

通过学习齿面到齿轴的位移关系，得到如图 2中齿轴检测分支所示的投影向量。然后，我们将输入点云与它们的预测投影向量进行投影以逼近目标齿轴。改论文采用最小二乘法<sup>[10]</sup>拟合直线，并将其

设为目标齿轴线的最终结果。将点云与预测的投影向量进行投影，以近似目标牙轴。采用最小二乘法拟合直线，设为目标牙轴的最终结果。

由于缺少专业牙医标注的数据集，本次复现选取牙长轴进行预测，复现结果如下所示。

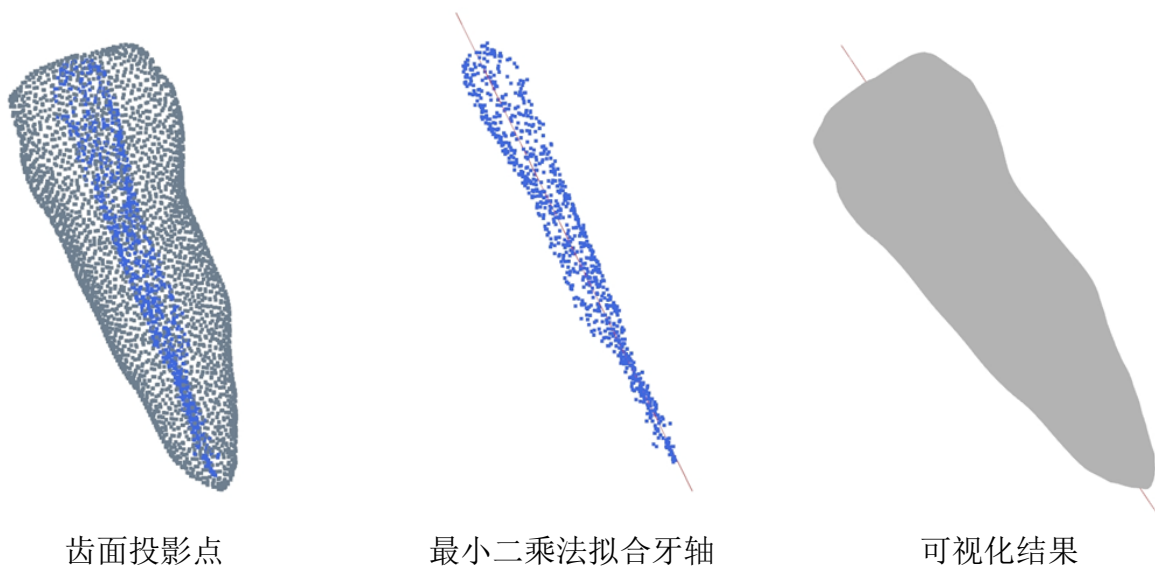


图 8: 牙轴生成

## 5 总结与展望

本次复现论文三维模型上牙齿关键点和牙轴的密集表示检测网络，通过将牙齿点云模型转化为距离场和投影向量场，实现对没有明显几何特征的关键点进行检测。本次复现实验实现内容包括数据集标记、CBCT 及口扫数据处理、牙齿关键点检测、牙轴检测、结果可视化等内容。由于缺少专业牙医标注的完整数据集用于训练等原因，复现过程中仅完成部分关键点及牙轴检测，在未来工作将在 CBCT 和口扫数据配准、距离场的计算精度、检测完整性等方面进一步完善。

## 参考文献

- [1] ANDREWS L F. The six keys to normal occlusion[J]. American journal of orthodontics, 1972, 62(3): 296-309.
- [2] CUI Z, LI C, WANG W. ToothNet: Automatic Tooth Instance Segmentation and Identification From Cone Beam CT Images[C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 2019: 6361-6370. DOI: 10.1109/CVPR.2019.00653.
- [3] RATH S, VOGEL E P, ANEES N, et al. Artificial Neural Networks as a powerful numerical tool to classify specific features of a tooth based on 3D scan data[J]. Computers in Biology & Medicine, 2017, 80: 65-76.
- [4] CHEN R, MA Y, CHEN N, et al. Cephalometric Landmark Detection by Attentive Feature Pyramid Fusion and Regression-Voting[C/OL]//Medical Image Computing and Computer Assisted Intervention - MICCAI 2019: 22nd International Conference, Shenzhen, China, October 13-17, 2019, Proceedings, Part III. Shenzhen, China: Springer-Verlag, 2019: 873-881. [https://doi.org/10.1007/978-3-030-32248-9\\_97](https://doi.org/10.1007/978-3-030-32248-9_97). DOI: 10.1007/978-3-030-32248-9\_97.



- [5] WEI G, CUI Z, LIU Y, et al. TANet: Towards Fully Automatic Tooth Arrangement[C]//VEDALDI A, BISCHOF H, BROX T, et al. Computer Vision – ECCV 2020. Cham: Springer International Publishing, 2020: 481-497.
- [6] GUO Y, WANG H, HU Q, et al. Deep Learning for 3D Point Clouds: A Survey[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2021, 43(12): 4338-4364. DOI: 10.1109/TPAMI.2020.3005434.
- [7] QI C R, YI L, SU H, et al. PointNet++: Deep Hierarchical Feature Learning on Point Sets in a Metric Space[J]. 2017.
- [8] JAFRI Z, AHMAD N, SAWAI M, et al. Digital Smile Design-An innovative tool in aesthetic dentistry [J]. Journal of Oral Biology and Craniofacial Research, 2020, 10(2): 194-198. DOI: <https://doi.org/10.1016/j.jobcr.2020.04.010>.
- [9] XIN S Q, WANG G J. Improving Chen and Han's Algorithm on the Discrete Geodesic Problem[J/OL]. ACM Trans. Graph., 2009, 28(4). <https://doi.org/10.1145/1559755.1559761>. DOI: 10.1145/1559755.1559761.
- [10] SANFORD R J. Application of the least-squares method to photoelastic analysis[J]. Experimental Mechanics, 1980, 20(6): 192-197.