

DOI: 10.13382/j.jemi.B2508107

错颌畸形的自动化诊断方法*

蒋晓彤¹ 刘小峰²

(1. 河海大学信息科学与工程学院 常州 213200; 2. 河海大学人工智能与自动化学院 常州 213200)

摘要:针对目前传统的错颌畸形诊断的主观性强、效率低等问题,设计了一种基于图神经网络(GNN)的错颌畸形自动化诊断方法 GraphTeeth。GraphTeeth 结合了图神经网络的架构优势,能够有效捕捉牙齿及其周围结构的拓扑信息。通过将牙齿的位置、形态以及相互关系建模为图结构,利用节点间的消息传递机制学习到更加精细的局部特征和全局特征。在实验阶段,使用了包含各类错颌畸形病例的大型数据集来训练和测试 GraphTeeth。实验结果显示,GraphTeeth 在关键性能指标上显著优于现有的目标检测方法。在平均精确率均值(mAP)指标上,GraphTeeth 达到了 43.45%,相较于传统的目标检测算法如 Mask R-CNN 的 32.26%、EfficientDet 的 38.73% 和 DETR 的 25.05% 有显著提高。此外,对于特定类型的错颌畸形——如固定矫治器的配戴——GraphTeeth 的准确率高达 91.28%,而对于健康牙齿的识别率则达到了 83.91%。结果表明,GraphTeeth 能够提供更快、更准确、更客观的诊断,为正畸治疗提供可靠支持。

关键词: 错颌畸形; 自动化检测; 图神经网络; 图卷积网络; 深度学习

中图分类号: TH3914.41; TN91 **文献标识码:** A **国家标准学科分类代码:** 520.20

Automated diagnostics of malocclusion

Jiang Xiaotong¹ Liu Xiaofeng²

(1. School of Information Science and Engineering, Hohai University, Changzhou 213200, China;

2. School of Artificial Intelligence and Automation, Hohai University, Changzhou 213200, China)

Abstract: Aiming at the problems of high subjectivity and low efficiency of the current traditional diagnosis of malocclusion, the article designs GraphTeeth, an automated diagnosis method of malocclusion based on graph neural network. GraphTeeth combines the architectural advantages of graph neural network, which is able to efficiently capture the topological information of the teeth and their surrounding structures. By modeling the position, morphology, and interrelationships of teeth as graph structures, finer local and global features are learned using a message passing mechanism between nodes. In the experimental phase, a large dataset containing various types of malocclusion cases was used to train and test GraphTeeth. The experimental results show that GraphTeeth significantly outperforms existing target detection methods in the key performance metrics. On the mAP metric, GraphTeeth achieves 43.45%, which is a significant improvement over traditional target detection algorithms such as Mask R-CNN at 32.26%, EfficientDet at 38.73%, and DETR at 25.05%. In addition, for specific types of malocclusions—such as fixed orthodontic appliance fitting—GraphTeeth achieves an accuracy of 91.28%, while the recognition rate for healthy teeth reaches 83.91%. The results suggest that GraphTeeth is able to provide faster, more accurate and objective diagnosis, providing reliable support for orthodontic treatment.

Keywords: malocclusion; automated detection; graph neural networks; graph convolutional networks; deep learning

0 引言

错颌畸形^[1]是一种牙齿在咬合过程中不能正确排列的异常牙齿状况,导致牙齿、颌面部软组织之间的功能和美学问题。传统的错颌畸形诊断主要依靠牙医的经验和视觉观察,往往受到各种主观因素^[2]的影响(如诊断医师的专业知识水平、诊断疲劳程度、光照条件和观察角度等),导致诊断结果不一致,有误诊^[3]的可能。与传统方法相比,基于深度学习的方法不仅能提高诊断的准确性和效率,还能简化诊断过程,减少患者的等待时间。具体来说,基于深度学习的方法^[4]通过训练学习大量标注的牙科图像数据,使训练有素的模型能够自动从目标口腔图像中提取和识别牙科特征。通过这种方法使得经过训练的模型提供客观一致的诊断结果,这样不仅可以消除人为因素的影响,还能很好地处理复杂病例。

目前,基于深度学习的牙科疾病自动诊断方法^[5-7]已被广泛开发,包括卷积神经网络(convolutional neural networks, CNN)和 Transformers。这些方法在提升诊断客观性和一致性方面取得了显著进展,但仍存在局限性。基于 CNN 的方法主要利用卷积层来提取图像的局部特征,通过层层卷积和池化操作来逐步提取更高级的特征。CNN 的优势在于其强大的局部特征提取能力和计算效率。然而,由于 CNN 的感受野有限,难以捕捉全局上下文信息。此外,卷积网络的参数量大,对计算资源要求较高,在处理 3D 牙科影像时计算成本更高^[8]。而基于 Transformer 的方法^[9-10]通过自注意力机制捕捉全局和远程语义信息的交互。此外,Transformer 结构复杂,计算量较大,在实际应用中可能受到硬件限制^[11]。相比之下,图神经网络(graph neural network, GNN)方法^[12]具有独特的优势,能够通过图结构建模牙齿及其相互关系,适用于具有结构化信息的数据分析。此外,GNN 方法能够通过图卷积操作(graph convolutional network, GCN)提取牙齿节点特征,并通过边连接建模牙齿间的空间关系。相比 CNN 和 Transformer, GNN 更适合处理牙齿的拓扑结构信息,特别是在错颌畸形分析任务中,能够弥补传统方法在空间信息建模上的不足^[13]。

本文提出了基于 GNN 的错颌畸形自动诊断模型 GraphTeeth。基于 GNN 的 GraphTeeth 方法采用了一种改进的图卷积网络架构,通过图卷积运算提取和更新节点特征,其独特优势在于能够捕获每个牙齿对象内的局部和线索,同时促进所有对象之间通过边缘进行全局信息建模的交互。通过实验发现 GraphTeeth 方法在错颌畸形自动化诊断方面展现了显著的优势。本文创新性地提出了一种基于 GNN 的错颌畸形自动诊断方法——GraphTeeth。通过将牙齿及其相互关系建模为图结构,

GraphTeeth 能够捕捉牙齿之间的空间关系及错位特征,克服了传统方法依赖经验和主观判断的局限性,为错颌畸形的诊断提供了一种更加客观、自动化的解决方案;在 GraphTeeth 模型中采用了改进的 GCN 架构,通过图卷积操作提取节点特征,并在全局范围内建模牙齿对象之间的关系。该架构增强了模型处理复杂结构数据的能力,显著提升了在错颌畸形诊断中的精度,相比于传统的 CNN 方法,能够更有效地捕捉牙齿结构的多维度特征;通过大量实验对 GraphTeeth 模型进行了验证,评估了其在错颌畸形自动诊断任务中的表现。实验结果表明,GraphTeeth 模型的诊断准确率优于传统方法,该方法的优越性能证明了其在临床应用中的潜力。

1 相关工作

近年来,越来越多的研究人员致力于将基于深度学习的方法应用于牙科领域的自动诊断。目前,在该领域的研究中主要使用了 3 种基于深度学习的方法分别为基于 CNN、基于 Transformer 和基于 GNN。在本节中将对这 3 种基于深度学习的自动诊断方法在牙科领域中的应用进行分类和总结。

1.1 基于 CNN 的方法

目前已经提出的各种基于 CNN 的牙齿图像分析方法,大致可以分为基于传统 CNN 的方法、基于 Faster R-CNN 的方法和基于 Mask R-CNN 的方法。Prajapat 等^[14]应用迁移学习对预训练的 VGG-16 模型^[15]进行再训练,然后将其与另一个包含 5 个卷积层的 CNN 相结合,对龋齿、根尖周感染和牙周炎 3 种主要的牙科疾病进行分类。Lee 等^[16]使用预训练好的 GoogLeNet Inception v3 网络^[17]进行数据预处理,并利用迁移学习将预训练好的 GoogLeNet Inception v3 网络适应新的数据集,从而可以更准确地检测和诊断龋病。

近十年来,区域卷积神经网络(region-based convolutional neural network, R-CNN)在牙齿图像分析任务中取得了重大进展,大致可分为基于 Faster R-CNN^[18]和基于 Mask R-CNN^[19]的方法。Alalharith 等^[20]提出基于改进的 Faster R-CNN-ResNet50 特征提取器自动检测牙齿和牙龈炎症区域。Tuzoff 等^[21]将 Faster R-CNN 与使用 ImageNet 预训练的 VGG-16 相结合,其中使用 Faster R-CNN 模块定义每个牙齿的边界,而 VGG-16 则对检测到的牙齿进行分类。Chen 等^[22]提出了一种基于 Faster R-CNN 和 Inception ResNet v2 作为神经网络模型的牙齿根尖切片检测和计数方法。在这项工作中,使用 Faster R-CNN 进行牙齿检测,然后进行对象定位和分类,其中使用 Inception ResNet v2 来提高检测效率并降低计算成本。Kim 等^[23]结合 Faster R-CNN、单会话多帧检测

(single-session multi-frame detection, SSD) 和启发式算法实现了全景放射图像中牙齿、种植体和冠的检测和编号。结合 Faster R-CNN 和 SSD 的检测结果,启发式算法利用位置和形状信息进一步优化牙齿检测和编号过程,提高整体精度和效率。

另外, Anantharaman 等^[24]提出了一种基于 Mask R-CNN 的口腔唇疱疹和口腔溃疡的检测和分割方法。Jader 等^[25]首次将 Mask R-CNN 应用于全景射线图像上进行自动检测和牙齿实例分割,对 193 张带注释的口腔图像进行训练,并从 MSCOCO 数据集迁移学习预训练权重,初始化 ResNet101 的参数,以弥补数据不足的挑战。Fatima 等^[26]使用改进的轻量级 Mask R-CNN 模型自动检测、分类和定位牙周病变。为了减少参数数量,该研究使用轻量级的 MobileNet-v2 作为主干网络进行特征提取,并在分类时增加了全局平均池化层和 dropout 层,防止模型过拟合。尽管基于 CNN 的方法在自动检测和分类牙齿疾病方面取得了重大进展,但仍然有一个主要缺点,在处理复杂变形和多尺度特征时的能力有限。

1.2 基于 Transformer 的方法

近年来,Transformer 也被广泛应用于牙科领域自动化诊断,以弥补 CNN 方法的不足。Jiang 等^[27]和 Hossain 等^[28]提出了基于 Transformer 的龋齿检测方法。具体而言,文献^[27]提出了一种快速龋病检测方法 RDFNet,其特征提取模块由 3 个堆叠的 Transformer 编码器组成,并去除了原始 Transformer 编码器中的归一化层以简化模型。此外,RDFNet 引入了 FReLU 激活功能,激活视觉空间信息,提高龋齿检测速度。文献^[28]扩展了 Vision Transformer 的应用来处理智能手机记录的牙齿图像,旨在将图像分类为晚期龋齿,早期龋齿或无龋齿。

此外,Almalki 等^[30]应用了 Swin-Transformer^[31]对牙齿全景 X-ray 图像进行分析。虽然 Swin-Transformer 需要大量的数据进行训练,而可用的牙科 X-ray 图像非常有限,但可利用自监督学习技术来缓解全景 X-ray 图像中数据稀缺的问题。Kadi 等^[32]训练 Detection Transformer^[33]对 52 个不同牙齿类别进行检测、分割,并提出了一种稳定的基于弥散的数据增强技术。Yang 等^[34]和 Li 等^[35]都对 CBCT 图像进行了研究,其中文献^[34]提出了一种基于 Transformer 的种植体位置回归网络,利用口腔 CBCT 数据自动预测种植体位置。使用牙冠区域的二维轴向视图来预测种植体的位置,并通过拟合种植体的中心线来确定其在根部的实际位置。同时,文献^[35]提出了一种用于牙齿分割、识别和牙髓钙化识别的新型 Transformer。使用了一种从粗到细的方法来分割低分辨率 CBCT 图像中的牙齿实例。在获得初始分割结果后,将这些结果映射到高分辨率 CBCT 图像中,提取出精确的牙齿区域进行进一步处理。在这些高分辨

率区域,采用牙齿实例相关和三重损失来提高钙化牙齿的识别性能。然而,基于 Transformer 的方法经常受到高训练数据需求和高计算复杂度的困扰。因此,限制了基于 Transformer 方法在牙科领域的应用。

1.3 基于 GNN 的方法

与以往基于 CNN 和 Transformer 的方法相比,基于 GNN 的方法可以灵活处理复杂的牙齿几何形状,减少信息损失,计算效率高^[36],适用于牙齿重建和自动化诊断等应用。

在计算机辅助正畸治疗中,关键任务之一是准确标记牙齿的三维表面。现有的分割方法不仅使用耗时的人工操作,而且通常依赖简单的几何属性作为分割标准,存在一定的局限性,难以适应不同病人牙齿外观的显著差异。为此,Lian 等^[37]提出了一种深度多尺度网格特征学习方法 MeshSNet,用于端到端三维牙齿表面标注。该方法采用了一种新颖的图约束学习模块,分层提取多尺度上下文特征,然后将局部和全局几何特征密集融合,从而全面表征网格元素,用于分割任务。进一步将这种方法扩展为 MeshSegNet^[38],该方法应用多层感知器(multilayer perceptron, MLP)提取网格单元的高阶几何特征。通过特征转换模块(feature conversion module, FTM)对齐初始特征,并通过多尺度图约束学习模块逐步增加接收域,以分层方式捕获多尺度局部几何特征。Wu 等^[39]提出了一个两阶段的框架。TS-MDL,用于分割牙齿和识别原始口腔内扫描的地标。TS-MDL 的创新之处在于引入了一种改进的基于 GNN 的 iMeshSegNet,用于高效准确的牙齿分割。用 EdgeConv 代替对称平均池和邻接矩阵,在保留局部几何信息的同时,大大降低了计算复杂度。Zheng 等^[40]提出了一种用于三维牙齿分割的 TeethGNN 算法。利用 GNN 直接处理 3D 网格数据,而不是将几何特征转换为规则网格输入,这可以更好地捕捉网格的非均匀和不规则结构。

2 基于 GNN 的错颌畸形自动化诊断方法

本文提出了一种基于 GNN 的方法,称为 GraphTeeth (图 1),其灵感来自通用 GRARIS^[41-43]。GraphTeeth 从目标口腔的表示中学习一个独特的牙齿之间的关系图,该图基于关系建模模块显式捕获各个牙齿之间的相关关系线索。GraphTeeth 首先从输入的牙齿图像中提取每颗牙齿的目标边界框。然后,从原始牙齿图像中裁剪出提取的边界框所对应的单个牙齿区域。基于这些裁剪的图像,牙齿节点特征生成(tooth node feature generator, TFG)模块构建节点特征图来学习每个牙齿节点的特征表示。与此同时,将整幅图像送入边缘特征提取(edge feature extractor, EFE)模块,得到牙齿关系图来描述每对节点之

间的特定任务关系。学习到的牙齿关系图将第 i 个牙齿表示为 graph 中的节点 $v_i \in \mathbf{V}$, 其中包含 1 个描述第 i 个牙齿的状态及其与口腔中其他牙齿的关系的向量。此

外, v_i 和 v_j 节点之间的特定任务关系线索也由两个有向边 $e_{i,j}, e_{j,i} \in \mathbf{E}$ 表示。

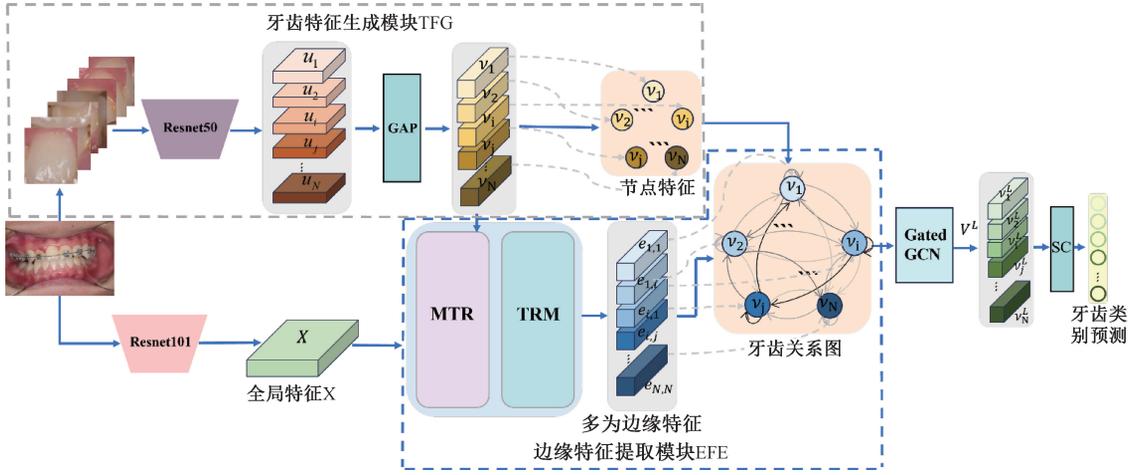


图 1 GraphTeeth 的流程基线

Fig. 1 GraphTeeth baseline workflow diagram

2.1 牙齿间关系节点特征学习

TFG 模块的主体是 N 个特征提取器, 每个特征提取器包一个全局平均池化层(global average pooling, GAP)。在 TFG 模块中, 采用整个口腔图像作为输入, 接着利用 N 个特征提取器对图像中的每个牙齿提取特征得到 U_i , 然后将其馈送到 GAP, 生成第 i 颗牙齿的特征向量。于是, 该模块学习到的牙齿间的关系图将第 i 个牙齿表示为图 1 中的节点 $v_i \in \mathbf{V}$, 其中包含 1 个描述第 i 个牙齿的特征及其与该口腔内其他牙齿的关联的向量。因此, 从整个口腔中可以分别学习到 N 个牙齿的特征表示。

2.2 多维边缘特征学习

除了节点特征中编码的关系线索外, 还提出了 1 个 EFE 模块来多维边缘的特征, 目的是明确表示不同节点特征之间的多个特定任务的关系线索。即使 1 对节点具有较低的相似性, 它们之间的关系仍然可能包含影响错颌畸形的关键线索, 这些线索在节点特征学习过程中被忽略。因此, EFE 模块中主要包括两个主要模块, 颌面牙齿表征建模(maxillofacial teeth representation modelling, MTRM)和牙齿关系建模(teeth relationship modelling, TRM)。如图 2 所示。将全局特征 $X \in \mathbf{R}^{D \times C}$ 作为输入, 其中 D 为维度, C 为通道数, 可以由目前任何标准的深度学习主干产生。

1) MTRM

MTRM 块以 1 对牙齿节点特征 v_i 和 v_j 的具体特征, 以及原始牙齿图像的全局特征 $X \in \mathbf{R}^{D \times C}$ 作为输入。首先分别在 v_i 和 X 以及 v_j 和 X 之间进行交叉注意, 其中仅将牙齿特定的特征映射 v_i 和 v_j 作为 query, 将整个口腔中

所有牙齿的表示 X 作为 key 和 value。首先, 交叉注意力机制定义为:

$$MTRM(\mathbf{Y}, \mathbf{Z}) = \text{softmax}\left(\frac{\mathbf{Y}\mathbf{W}_q(\mathbf{Z}\mathbf{W}_k)^T}{\sqrt{d_k}}\right)\mathbf{Z}\mathbf{W}_v \quad (1)$$

因此, 整个过程表示为:

$$\mathbf{F}_{i,x}^{TS}, \mathbf{F}_{j,x}^{TS} = MTRM(\mathbf{V}_i, \mathbf{X}), MTRM(\mathbf{V}_j, \mathbf{X}) \quad (2)$$

式中: $\mathbf{W}_q, \mathbf{W}_k, \mathbf{W}_v$ 是可学习的权重; d_k 是相当于 key 通道的缩放因子。

2) TRM

在通过 MTRM 块后, 每颗牙齿都被独立地编码了特定的线索。因此, 需要 TRM 进一步提取牙间的相关线索。在 $\mathbf{F}_{i,x}^{TS}$ 和 $\mathbf{F}_{j,x}^{TS}$ 之间再次使用交叉注意力机制, 生成 $\mathbf{F}_{i,j,x}^{TR}$ 和 $\mathbf{F}_{j,i,x}^{TR}$ 。 $\mathbf{F}_{i,j,x}^{TR}$ 是用 $\mathbf{F}_{j,x}^{TS}$ 作为 query, $\mathbf{F}_{i,x}^{TS}$ 作为 key 和 value 生成的, 而 $\mathbf{F}_{j,i,x}^{TR}$ 是用相反的方式生成的。因此, $\mathbf{F}_{i,j,x}^{TR}$ 总结了 $\mathbf{F}_{i,x}^{TS}$ 中牙齿的相关线索, $\mathbf{F}_{j,i,x}^{TR}$ 总结了 $\mathbf{F}_{j,x}^{TS}$ 中牙齿的相关线索。最后, 将得到的 $\mathbf{F}_{i,j,x}^{TR}, \mathbf{F}_{j,i,x}^{TR}$ 输入到 GAP 层中, 得到多维边缘特征向量 $e_{i,j}$ 和 $e_{j,i}$ 。这样, 边缘特征向量 $e_{i,j}$ 和 $e_{j,i}$ 总结了整个颌面图像中的多个牙齿线索。以上过程可以概括为:

$$e_{i,j}, e_{j,i} = \text{GAP}\left(\text{TRM}(\mathbf{F}_{j,x}^{TS}, \mathbf{F}_{i,x}^{TS}), \text{TRM}(\mathbf{F}_{i,x}^{TS}, \mathbf{F}_{j,x}^{TS})\right) \quad (3)$$

经过以上两个块, 学习到牙齿关系图 $\mathbf{G}^0 = (\mathbf{V}^0, \mathbf{E}^0)$, 其中 \mathbf{G}^0 包含 N 个节点特征和 $N \times N$ 个多维有向边缘特征, 全面描述了每对牙齿之间的特定任务关系。

2.3 损失函数

在本文模型中, 使用加权负对数似然损失 $NLLLoss$

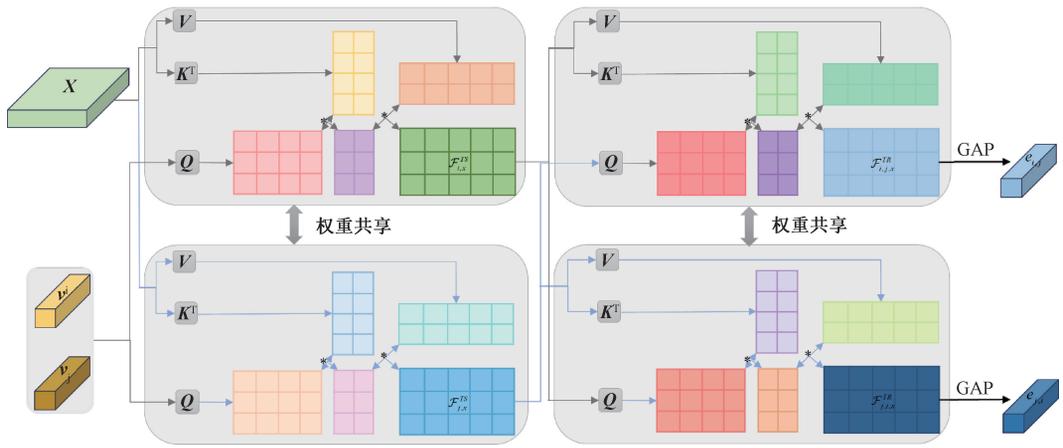


图2 EFE 模块示意图

Fig. 2 Diagram of the EFE module

来计算分类输出与真实值之间的损失值 L_{WA} 。从分类器中获得类的对数概率向量后,应用 $\log_softmax$ 函数获得对数概率,如式(4)和(5)所示。

$$\sigma(z)_j = \frac{e^{z_j}}{\sum_{k=1}^n e^{z_k}} \quad (4)$$

$$\log_softmax = \ln \sigma(z)_j \quad (5)$$

对于给定的对数概率分布, $NLLLoss$ 计算负对数似然函数。因此, $NLLLoss$ 的公式为:

$$NLLLoss = -\frac{1}{N} \sum_{k=1}^N y_k (\log_softmax) \quad (6)$$

式中: $\sigma(z)_j$ 为第 j 类的概率; z_j 是输入向量 Z 的第 j 个元素; n 是类别的个数。 $\log_softmax(z)_j$ 表示第 j 类的对数概率; y_k 是第 k 个样本的真实标签。在实践中,真实标签 y_k 通常用指数表示,即第 k 个样本的真实类别指数。因此,损失函数可以简化为:

$$NLLLoss = -\frac{1}{N} \sum_{k=1}^N \log(P_{i,y_i}) \quad (7)$$

式中: P_{i,y_i} 表示第 i 个样本中真实标签; y_i 对应的对数概率。使用加权负对数似然损失确保数值稳定性,因为在计算概率和对数概率时直接使用对数函数可能导致数值不稳定问题。先计算 $softmax$ 再取对数,小概率值将接近于零,取对数时将得到负无穷 ($-inf$) 或 NaN 。但是, $\log_softmax$ 是一个单独的操作,可以更稳定地计算日志概率,避免上述问题。另外, $NLLLoss$ 可以更灵活地应用于对数概率损失计算,这在一些自定义损失函数或特定概率模型中非常有用。

3 实验

3.1 实验细节

使用一个由医院口腔科提供的未公开口腔数据集来

评估本文提出的方法。该数据集中共包含 1 500 张口腔图像,将这些所有数据一共划分为 10 个类别。并将其划分为训练集(60%)、验证集(20%)和测试集(20%)。在实验中,采用特定的配置来评估 GraphTeeth 以确保性能,在训练过程中采用 AdamW 优化器和 Cosine Annealing Schedule,对 GraphTeeth 进行了 50 个 epoch 训练,学习率降低至 2×10^{-5} 。所有实验均在 NVIDIA GeForce RTX 3090 上通过 PyTorch 库进行。

3.2 实验结果

本文进行了全面的实验,以评估比较 GraphTeeth 方法和其他几种最先进的方法。图 3 所示为 GraphTeeth 的诊断效果。表 1 为各模型的整体性能。

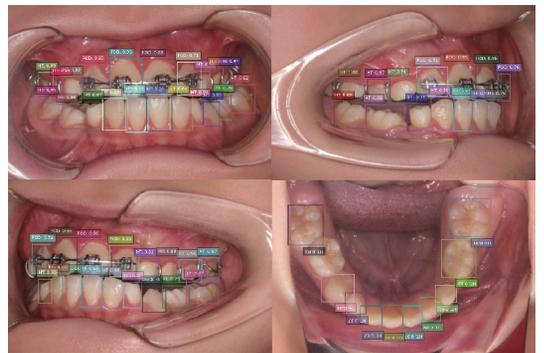


图3 GraphTeeth 的诊断应用效果

Fig. 3 Diagnostic application results of GraphTeeth

1) 整体性能分析

从表 1 可以看出, GraphTeeth 模型仅经过 50 个 epoch 训练平均精确率均值 (mean average precision, mAP) 就达到了 43.45%。相比之下 Mask R-CNN 和 EfficientDet 需要 100 个 epoch 训练才能达到最优 mAP,而 DETR 需要更长的 epoch 训练才能达到收敛。这说明

GraphTeeth 在大幅度减少训练时间的情况下仍然能够获得相对优越的 mAP。这是由于 GraphTeeth 采用轻量级 GatedGCN 结构,相比其他方法减少了计算资源的消耗,

适用于大规模牙科影像诊断任务。同时,相比其他深度学习方法,GraphTeeth 能够利用 GNN 建模牙齿的拓扑关系,从而提升诊断精度。

表 1 各方法整体性能比较

Table 1 Overall performance comparison of different methods

指标	Mask R-CNN	EfficientDet	DETR	GraphTeeth(本文)
Epoch	100	150	300	50
Backbone	ResNet50	EfficientNet-D3	ResNet50	ResNet50+ResNet101
mAP/%	32.26	<u>38.73</u>	25.05	43.45

2) 分类精确度分析

表 2 和图 4 所示为 10 个类别的 mAP 的详细比较。表 2 中加粗数据为各类别在不同模型中的最好结果,加下划线的数据为第 2 优的结果。从表中数据来看,GraphTeeth 模型在各个类别上表现出色,与其他先进方法比,在各类别的 mAP 指标上都有不同程度的提高。首先,其整体 mAP 为 41.45%,表明该方法相较于其他经典方法在整体上具有较高的准确诊断能力。GraphTeeth 在特定类别上的表现尤为突出,为错颌畸形的诊断提供了强有力的支持。例如,在“活动矫治器”类别中,GraphTeeth 的 mAP 达到了 63.17%,相比于 EfficientDet 的 53.73%和 DETR 的 22.92%,显示了更优的识别能力。同样,在“反颌”这一关键分类中,GraphTeeth 的 mAP 为

42.19%,虽然与 EfficientDet 的 41.89%接近,但明显优于 DETR 的 27.55%,证明了其在此类任务上的优势。最引人注目的是在“固定矫治器”这一类别,GraphTeeth 的表现尤为出色,mAP 达到 91.28%,远超 Mask R-CNN 的 59.79%和 DETR 的 44.23%,这一结果不仅强调了 GraphTeeth 在精确度上的卓越表现,也表明了其在处理具有挑战性的医学图像方面的出色能力。这种高精度的诊断能力对于确保患者接受到准确及时的治疗至关重要。综合来看,GraphTeeth 模型不仅在通用的 mAP 指标上展示了优异的表现,更在关键的诊断类别中展示了其精确度和鲁棒性,使其成为错颌畸形自动化诊断领域的优选模型。

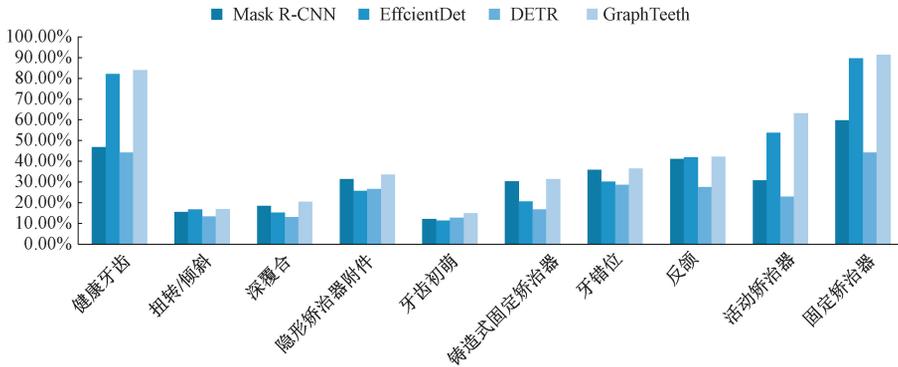


图 4 各类别的识别准确率

Fig. 4 Classification accuracy for each category

表 2 每个模型上各类别的识别 mAP 比较

Table 2 Each class recognition mAP comparison for each model

(%)

方法	Mask R-CNN	EfficientDet	DETR	GraphTeeth(本文)
健康牙齿	46.78	<u>82.11</u>	44.16	83.91
扭转/倾斜	15.62	16.74	13.42	16.89
深覆合	<u>18.51</u>	15.28	13.10	20.56
隐形矫治器附件	<u>31.46</u>	25.71	26.73	33.64
牙齿初萌	12.17	11.36	<u>12.83</u>	14.89
铸造式固定矫治器	30.40	20.63	16.86	31.44
牙错位	<u>35.88</u>	30.23	28.70	36.53
反颌	41.20	<u>41.89</u>	27.55	42.19
活动矫治器	30.79	<u>53.73</u>	22.92	63.17
固定矫治器	59.79	<u>89.62</u>	44.23	91.28

3.3 消融实验

表3为边缘特征在增强模型性能方面的作用。在加入边缘特征时,GraphTeeth的结果表明,加入边缘特征可以使模型更好地捕获牙齿之间的关系信息,从而提高定位和分类精度。在边缘特征下,GraphTeeth的mAP为43.45%。在错颌畸形的诊断中,牙齿的位置关系和结构特征对准确分类至关重要。边缘特征有效地补充了节点特征无法完全捕获的信息,使模型能够更好地理解牙齿之间的复杂关系。相比之下,在去除边缘特征后,GraphTeeth的性能有明显的变化,mAP都有明显的下降。

表3 边缘特征在 GraphTeeth 中的作用

Table 3 The analysis of the role of edge features in GraphTeeth (%)

方法	mAP
GraphTeeth(本文)	43.45
GraphTeeth without edge features	41.27

4 结论

本文提出了一种基于GNN的错颌畸形自动化诊断方法 GraphTeeth,该方法通过构建牙齿及其相互关系的拓扑结构,并结合局部特征提取和全局信息建模,提高错颌畸形的自动化诊断能力。实验结果表明,GraphTeeth在关键性能指标上优于现有方法,mAP达到43.45%,相较于Mask R-CNN(32.26%)、EfficientDet(38.73%)和DETR(25.05%)有显著提升,特别是在固定矫治器的识别任务上取得了91.28%的高准确率,健康牙齿识别率达到83.91%。与传统基于CNN和Transformer的方法相比,GraphTeeth在训练效率、特征提取能力和泛化能力上均具有优势。通过构建牙齿关系图,该方法能够更好地捕捉牙齿间的错颌特征,提高诊断精度,同时采用GatedGCN结构优化计算效率,在50个epoch训练内即达到最佳性能,显著降低计算开销。然而,尽管GraphTeeth在实验中展现出良好的性能,该方法仍存在一定局限性,如对高质量标注数据的依赖以及在处理大规模数据时计算资源消耗较大。未来研究可以进一步优化GraphTeeth的模型结构,结合多模态数据(如X-ray、CBCT、口腔扫描)提升诊断精度,并探索自监督学习方法减少对人工标注的依赖。此外,通过改进GNN结构,如引入轻量化GNN设计,可以提升模型的计算效率,使其更适用于临床应用。同时,GraphTeeth还可进一步拓展至个性化正畸方案推荐系统,结合强化学习和生成模型,为患者提供更精准的矫治方案。综合来看,本文为错颌

畸形的智能诊断提供了一种新的思路,验证了GNN在牙科智能诊断中的可行性和应用潜力,未来有望进一步推动智能正畸辅助系统的发展,提高诊断效率和治疗效果。

参考文献

- [1] 姜金刚,郭晓伟,张永德,等.基于有限点展成法的正畸弓丝成形控制点规划[J].仪器仪表学报,2017,38(3):612-619.
JIANG J G, GUO X W, ZHANG Y D, et al. Formed control point planning of orthodontic archwire based on finite point extension method[J]. Chinese Journal of Scientific Instrument, 2017, 38(3):612-619.
- [2] HASSAN R, RAHIMAH A K. Occlusion, malocclusion and method of measurements-an overview[J]. Archives of Orofacial Sciences, 2007, 2: 3-9.
- [3] ALBALAWI F, ALAMOUD K A. Trends and application of artificial intelligence technology in orthodontic diagnosis and treatment planning: A review[J]. Applied Sciences, 2022, 12(22): 11864.
- [4] 赵阳,李俊诚,成博栋,等.深度学习在口腔医学影像中的应用与挑战[J].中国图象图形学报,2024,29(3):586-607.
ZHAO Y, LI J CH, CHENG B D, et al. Applications and challenges of deep learning in dental imaging[J]. Journal of Image and Graphics, 2024, 29(3): 586-607.
- [5] 余杨杨,陈洁玉,高峰,等.深度学习在口腔颌面部影像诊断领域的研究进展[J].中华口腔医学研究杂志(电子版),2021,15(3):185.
SHE Y Y, CHEN J Y, GAO F, et al. Deep learning: Advance in the field of oral and maxillofacial diagnostic imaging[J]. Chinese Journal of Stomatological Research (Electronic Edition), 2021, 15(3): 185.
- [6] SONG S, JAISWAL S, SHEN L, et al. Spectral representation of behaviour primitives for depression analysis[J]. IEEE Transactions on Affective Computing, 2020, 13(2): 829-844.
- [7] SHAMSHIRBAND S, FATHI M, DEHZANGI A, et al. A review on deep learning approaches in healthcare systems: Taxonomies, challenges, and open issues[J]. Journal of Biomedical Informatics, 2021, 1(113): 103627.
- [8] 林景栋,吴欣怡,柴毅,等.卷积神经网络结构优化综述[J].自动化学报,2020,46(1):24-37.
LIN J D, WU X Y, CHAI Y, et al. Structure optimization of convolutional neural networks: A Survey[J]. Acta Automatica Sinica, 2020, 46(1): 24-37.
- [9] SHENG C, WANG L, HUANG Z, et al. Transformer-based deep learning network for tooth segmentation on panoramic radiographs[J]. Journal of Systems Science and Complexity, 2023, 36(1): 257-272.

- [10] MEHREEN K, MUTTI M R U, UMAR M F, et al. Mask-transformer-based networks for teeth segmentation in panoramic radiographs [J]. *Bioengineering* (Basel, Switzerland), 2023, 10(7).
- [11] 田永林, 王雨桐, 王建功, 等. 视觉 Transformer 研究的关键问题: 现状及展望 [J]. *自动化学报*, 2022, 48(4): 957-979.
TIAN Y L, WANG Y T, WANG J G, et al. Key problems and progress of vision transformers: The state of the art and prospects [J]. *Acta Automatica Sinica*, 2022, 48(4): 957-979.
- [12] LIU M, LI X, LIU J, et al. TUCNet: A channel and spatial attention-based graph convolutional network for teeth upsampling and completion [J]. *Computers in Biology and Medicine*, 2023, 15(166): 107519.
- [13] WU Z, PAN S, CHEN F, et al. A comprehensive survey on graph neural networks [J]. *IEEE Transactions on Neural Networks and Learning Systems*, 2020, 32(1): 4-24.
- [14] PRAJAPATI S A, NAGARAJ R, MITRA S. Classification of dental diseases using CNN and transfer learning [C]. 2017 5th International Symposium on Computational and Business Intelligence. IEEE, 2017: 70-74.
- [15] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition [J]. *ArXiv preprint arXiv:1409.1556*, 2014.
- [16] LEE J H, KIM D H, JEONG S N, et al. Detection and diagnosis of dental caries using a deep learning-based convolutional neural network algorithm [J]. *Journal of Dentistry*, 2018, 10(77): 106-111.
- [17] SZEGEDY C, VANHOUCKE V, IOFFE S, et al. Rethinking the inception architecture for computer vision [C]. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2016: 2818-2826.
- [18] REN S, HE K, GIRSHICK R, et al. Faster R-CNN: Towards real-time object detection with region proposal networks [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2016, 39(6): 1137-1149.
- [19] HE K, GKIOXARI G, DOLLÁR P, et al. Mask R-CNN [C]. *Proceedings of the IEEE International Conference on Computer Vision*. 2017: 2961-2969.
- [20] ALALHARITH D M, ALHARTHI H M, ALGHAMDI W M, et al. A deep learning-based approach for the detection of early signs of gingivitis in orthodontic patients using faster region-based convolutional neural networks [J]. *International Journal of Environmental Research and Public Health*, 2020, 17(22): 8447.
- [21] TUZOFF D V, TUZOVA L N, BORNSTEIN M M, et al. Tooth detection and numbering in panoramic radiographs using convolutional neural networks [J]. *Dentomaxillofacial Radiology*, 2019, 48(4): 20180051.
- [22] CHEN H, ZHANG K, LYU P, et al. A deep learning approach to automatic teeth detection and numbering based on object detection in dental periapical films [J]. *Scientific Reports*, 2019, 9(1): 3840.
- [23] KIM C, KIM D, JEONG H G, et al. Automatic tooth detection and numbering using a combination of a CNN and heuristic algorithm [J]. *Applied Sciences*, 2020, 10(16): 5624.
- [24] ANANTHARAMAN R, VELAZQUEZ M, LEE Y. Utilizing mask R-CNN for detection and segmentation of oral diseases [C]. 2018 IEEE International Conference on Bioinformatics and Biomedicine. IEEE, 2018: 2197-2204.
- [25] JADER G, FONTINELI J, RUIZ M, et al. Deep instance segmentation of teeth in panoramic X-ray images [C]. 2018 31st SIBGRAPI Conference on Graphics, Patterns and Images. IEEE, 2018: 400-407.
- [26] FATIMA A, SHAFI I, AFZAL H, et al. Deep learning-based multiclass instance segmentation for dental lesion detection [C]. *Healthcare*. MDPI, 2023, 11(3): 347.
- [27] JIANG H, ZHANG P, CHE C, et al. Rdfnet: A fast caries detection method incorporating transformer mechanism [J]. *Computational and Mathematical Methods in Medicine*, 2021, 2021(1): 9773917.
- [28] HOSSAIN M S, RAHMAN M M, SYEED M M M, et al. Cavit: Early stage dental caries detection from smartphone-image using vision transformer [C]. 2023 4th International Conference on Artificial Intelligence, Robotics and Control. IEEE, 2023: 9-14.
- [29] DOSOVITSKIY A. An image is worth 16×16 words: Transformers for image recognition at scale [J]. *ArXiv preprint arXiv:2010.11929*, 2020.
- [30] ALMALKI A, LATECKI L J. Self-supervised learning with masked image modeling for teeth numbering, detection of dental restorations, and instance segmentation in dental panoramic radiographs [C]. *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, 2023: 5594-5603.
- [31] LIU Z, LIN Y, CAO Y, et al. Swin Transformer: Hierarchical vision transformer using shifted windows [J]. 2021, DOI:10.48550/arXiv.2103.14030.
- [32] KADI H, SOURGET T, KAWCZYNSKI M, et al. Detection transformer for teeth detection, segmentation, and numbering in oral rare diseases: Focus on data

- augmentation and inpainting techniques [J]. ArXiv preprint arXiv:2402.04408, 2024.
- [33] CARION N, MASSA F, SYNNAEVE G, et al. End-to-end object detection with transformers [C]. European Conference on Computer Vision. Cham: Springer International Publishing, 2020: 213-229.
- [34] YANG X, LI X, LI X, et al. Implantformer: Vision transformer-based implant position regression using dental cbct data [J]. Neural Computing and Applications, 2024, 36(12): 6643-6658.
- [35] LI S, LI C, DU Y, et al. Transformer-based tooth segmentation, identification and pulp calcification recognition in CBCT [C]. International Conference on Medical Image Computing and Computer-Assisted Intervention. Cham: Springer Nature Switzerland, 2023: 706-714.
- [36] HAN K, WANG Y, GUO J, et al. Vision GNN: An image is worth graph of nodes [J]. Advances in Neural Information Processing Systems, 2022, 35: 8291-8303.
- [37] LIAN C, WANG L, WU T H, et al. Meshsnet: Deep multi-scale mesh feature learning for end-to-end tooth labeling on 3D dental surfaces [C]. Medical Image Computing and Computer Assisted Intervention-MICCAI 2019: 22nd International Conference, Shenzhen, China, October 13-17, 2019, Proceedings, Part VI 22. Springer International Publishing, 2019: 837-845.
- [38] LIAN C, WANG L, WU T H, et al. Deep multi-scale mesh feature learning for automated labeling of raw dental surfaces from 3D intraoral scanners [J]. IEEE Transactions on Medical Imaging, 2020, 39(7): 2440-2450.
- [39] WU T H, LIAN C, LEE S, et al. Two-stage mesh deep learning for automated tooth segmentation and landmark localization on 3D intraoral scans [J]. IEEE Transactions on Medical Imaging, 2022, 41(11): 3158-3166.
- [40] ZHENG Y, CHEN B, SHEN Y, et al. TeethGNN: Semantic 3D teeth segmentation with graph neural networks [J]. IEEE Transactions on Visualization and Computer Graphics, 2022, 29(7): 3158-3168.
- [41] SONG S, SONG Y, LUO C, et al. Gratis: Deep learning graph representation with task-specific topology and multi-dimensional edge features [J]. ArXiv preprint arXiv: 2211.12482, 2022.
- [42] LUO C, SONG S, XIE W, et al. Learning multi-dimensional edge feature-based au relation graph for facial action unit recognition [J]. ArXiv preprint arXiv: 2205.01782, 2022.
- [43] SONG Y, LUO C, JACKSON A, et al. MERG: Multi-dimensional edge representation generation layer for graph neural networks [C]. ICASSP 2024-2024 IEEE International Conference on Acoustics, Speech and Signal Processing. IEEE, 2024: 5475-5479.

作者简介



蒋晓彤, 2022 年于南京信息工程大学获得学士学位, 现为河海大学硕士研究生, 主要研究方向为深度学习、人工智能、错颌畸形研究和计算机视觉等。

E-mail: 1747521405@qq.com

Jiang Xiaotong received her B. Sc. degree from Nanjing University of Information Engineering in 2022. Now she is a M. Sc. candidate at Hohai University. Her main research interests in deep learning, artificial intelligence orthodontic malocclusion studies and computer vision, etc.



刘小峰(通信作者), 1996 年于太原理工大学获得学士学位, 1999 年于太原理工大学获得硕士学位, 2006 年于西安交通大学获得博士学位, 现为河海大学人工智能与自动化学院教授兼院长, 主要研究方向为自然启发式导航及人机交互。

E-mail: xfliu@hhu.edu.cn

Liu Xiaofeng (Corresponding author) received his B. Sc. degree from Taiyuan University of Technology in 1996, M. Sc. degree from Taiyuan University of Technology in 1999 and Ph. D. degree from Xi'an Jiaotong University in 2006, respectively. Now he is a professor and president of the School of Artificial Intelligence and Automation at Hohai University. His main research interests include nature inspired navigation and human-robot interaction.