**摘 要**

随着社会发展，人群对生活品质的要求渐渐提高，越来越多的成年人也在寻求牙齿正畸治疗，以改善他们的笑容，改正牙齿咬合状况或纠正因受伤、疾病或长久忽视口腔护理所造成的其它问题。

正畸即矫正牙齿、解除错牙和畸形。正畸主要研究错颌畸形的病因机制，诊断分析及其预防和治疗，重新排列病人牙弓上牙齿的位置来达到审美上美观和功能上最佳的结果。在实际诊疗中，医生需要花费大量时间和精力在识别病人牙齿的特征点，从而产生了自动识别牙齿特征点算法的需求。特征点是一组由医生确定的一组定位标记点，使得每个牙弓上取得想要的牙齿排列方式和上下牙弓相互对应的匹配情况。

本文对华西口腔提供的牙齿数据进行过滤筛选，并提出一个牙列上的54类特征点类型满足牙医的临床需求，使用软件标注大量牙齿数据，对文件和IGES文件做预处理转换，改进深度学习网络使其适用于牙齿网格数据，使用高斯函数计算出高斯热图矩阵，使用改进作网络预测每个牙齿的高斯热图，对高斯热图进行后处理找出特征点位置，使用改进后的模型在大量数据集上进行实验，分析实验结果，改善其性能。

本文主要分为六个部分，第一部分介绍机器学习在计算机视觉领域的的研究背景和研究现状，特别是三维口腔领域；第二部分阐述本文提出的54类牙齿特征类型的定义和作用；第三部分介绍本文算法设计用到的关键技术；第四部分介绍本文设计的如何将从点云分割和识别转换到为牙齿网格的特征点自动动识别，主要由网格文件和特征点文件的数据预处理、高斯热图的使用和深度学习网络的改进和数据后处理确定特征点位置三个部分构成；第五部分对展示实验结果并进行分析和比较；第六部分对于本文提出的算法的总结与展望。

**关键词**：正畸；牙齿特征；机器学习；深度学习;计算机视觉；

**Abstract**

With the development of society, the requirements for the quality of life of the population are gradually increasing, and more and more adults are also seeking orthodontic treatment to improve their smiles, correct the occlusion of teeth, or correct oral care due to injury, disease or long-term neglect. other problems caused.

Orthodontics is the correction of teeth, the removal of misaligned teeth, and deformities. Orthodontics mainly studies the etiological mechanism, diagnosis, analysis, prevention, and treatment of malocclusion, and rearranges the position of the teeth on the patient's dental arch to achieve the best aesthetic and functional results. In actual diagnosis and treatment, doctors need to spend a lot of time and energy in identifying the feature points of a patient's teeth, thus creating a need for an algorithm to automatically identify tooth feature points. The feature points are a set of positioning mark points determined by the doctor so that the desired arrangement of teeth and the matching situation of the upper and lower dental arches can be obtained on each dental arch.

In this paper, the dental data provided by Dental are filtered and screened, and 54 characteristic point types on a dentition are proposed to meet the clinical needs of dentists. Geomagic Studio software is used to mark a large amount of dental data, and STL files and IGES files are preprocessed and converted. Improve the deep learning network to make it suitable for tooth mesh data, use the Gaussian function to calculate the Gaussian heatmap matrix, use the improved network to predict the Gaussian heatmap of each tooth, and post-process the Gaussian heatmap to find the location of feature points, using the improved model to conduct experiments on a large number of datasets, analyze the experimental results, and improve its performance.

This paper is mainly divided into six parts. The first part introduces the research background and research status of machine learning in the field of computer vision, especially in the field of the 3D oral cavity; the second part describes the definition and function of the tooth feature types proposed in this paper; the third part introduces The key technologies used in the algorithm design of this paper; the fourth part introduces how to convert from point cloud segmentation and recognition to automatic recognition of feature points of tooth meshes designed in this paper, mainly composed of grid file STL and feature point file IGES. Data preprocessing, the use of Gaussian heat map, the improvement of deep learning network, and the post-processing of data to determine the location of feature points are composed of three parts; the fifth part shows the experimental results and analyzes and compares; the sixth part is about the algorithm proposed in this paper summary and outlook.

**Key Words**: orthodontics; machine learning; deep learning; computer vision;

目录

[摘 要 I](#_Toc103895239)

[Abstract III](#_Toc103895240)

[1 绪论 1](#_Toc103895241)

[**1.1 课题背景** 1](#_Toc103895242)

[**1.2 国内外研究概况** 2](#_Toc103895243)

[**1.3 本文主要工作** 3](#_Toc103895244)

[2 牙齿特征 5](#_Toc103895245)

[**2.1 牙齿解剖学概念** 5](#_Toc103895246)

[**2.2 牙齿数据来源** 6](#_Toc103895247)

[**2.3牙齿特征点** 7](#_Toc103895248)

[**2.4小节** 8](#_Toc103895249)

[3 关键技术 9](#_Toc103895250)

[**3.1神经元模型** 9](#_Toc103895251)

[**3.2激活函数** 10](#_Toc103895252)

[**3.3多层感知器** 11](#_Toc103895253)

[**3.4小节** 12](#_Toc103895254)

[4 算法设计 14](#_Toc103895255)

[**4.1 数据的获取与处理** 14](#_Toc103895256)

[**4.2 高斯热图** 21](#_Toc103895257)

[**4.3 神经网络模型** 22](#_Toc103895258)

[**4.5 后处理** 24](#_Toc103895259)

[**4.6 小节** 24](#_Toc103895260)

[5 实验与分析 25](#_Toc103895261)

[**5.1 数据集** 25](#_Toc103895262)

[**5.2 实验过程** 25](#_Toc103895263)

[**5.3 实验分析** 28](#_Toc103895264)

[**5.4 实验环境** 30](#_Toc103895265)

[**5.4 小节** 31](#_Toc103895266)

[6 总结与展望 32](#_Toc103895267)

[**6.1 本文总结** 32](#_Toc103895268)

[**6.2 展望** 32](#_Toc103895269)

[参考文献 34](#_Toc103895270)

**1 绪论**

机器学习是人工智能研究中最早的课题之一，也是人工智能中最具有智能特

征和最前沿的研究课题之一。机器学习在自然语言处理、模式识别、计算机视等领域得到了普遍的应用。本文搭建机器学习中的神经网络架构，设计自动识别牙齿特征点的算法。深度学习以人工神经网络为架构，对资料进行表征学习，可以分为非监督式学习和监督式学习来提取目标特征信息。一个监督式学习者的任务在观察完一些事先标记过的训练范例（输入和预期输出）后，去预测这个函数对任何可能出现的输入的输出。要达到此目的，学习者必须以"合理"的方式从现有的资料中一般化到非观察到的情况。本文使用的基于改进的网络模型是监督式学习，预测每一个输入牙齿网格对应的高斯热图矩阵，之后对高斯热图矩阵做后处理，取高斯热图值中，有局部最大值的网格中心点作为特征点的位置。

**1.1 课题背景**

本文旨在解决临床正畸正畸的特征点识别问题。正畸就是矫正牙齿、解除错牙和畸形。正畸主要研究错颌畸形的病因机制，诊断分析及其预防和治疗，重新排列病人牙弓上牙齿的位置来达到审美上美观和功能上最佳的结果。

随着人们生活水平的提高和社会技术的进步，正畸治疗的费用已逐渐降低，使得普通民众也能够选择正畸治疗来改正牙齿咬合状况或纠正因受伤、疾病或长久忽视口腔护理所造成的其它问题，或者是保证牙齿的美观性，拥有一个美丽的笑容。王兴详细介绍了目前国内中国口腔医疗行业的发展形式[1]，反应了正畸治疗的必要性。

在正畸治疗中，相关牙齿特征集的识别对于临床应用至关重要。数字 3D 牙科模型因其效率和安全性而被广泛用于正畸。创建一个患者的正畸治疗计划（例如，用于制作清晰的对准器），正畸医生需要分割牙齿并注释3D 牙科模型上的相应解剖标志分析和重新排列牙齿位置。即使在大多数商业软件（集成半自动算法）提供的帮助下完成牙齿的解剖标志分析和重新排列这两项任务也是耗时、乏味且依赖操作人的专业知识。真正的临床需求是开发全自动的方法而不是手动操作，但实际上仅是手动操作就相当具有挑战性，特别是对于牙齿标志定位，主要是由于以下几点原因：

（1）不同牙齿的形状差异变化大；

（2）一些患者的牙齿排列紊乱、因先天或者后天因素造成牙齿缺失或磨损（3)口内扫描仪捕获的牙齿三维口腔模型不完整，扫描结果并非是牙齿的全貌（牙齿的侧面难以扫描到且牙齿的排列不整齐导致相互遮挡）。

三维牙齿分割和识别一直是计算机视觉医疗领域的一个小众分支，近年来随着机器学习在计算机视觉的崛起，

因此，本文设计一个自动识别牙齿特征点的算法来减少正畸治疗过程中临床医生的任务。科技的进步使得正畸治疗计划几乎可以用计算机进行模拟。首先通过构建患者特定牙印的各个牙弓的石膏模型，然后通过激光扫描该牙弓的石膏模型来生成虚拟网格对象，然后将包含整个口腔元素的网格模型分割为单独的牙齿对象。随着机器学习技术的发展，机器学习的基本理论和综合系统的研究不断得到加强和发展，人们也开始使用机器学习的方法解决三维口腔领域的问题，在计算机医疗视觉领域上取得了突破，特别是三维口腔网格牙齿的划分和识别方面。

**1.2 国内外研究概况**

随着三维扫描技术和计算机视觉的广泛应用，数字化的口腔模型得到越来越多人的关注和研究[1-4]。如何识别牙齿特征来帮助临床医生诊疗逐渐成为计算机图形学领域的研究热点问题。随着机器学习的兴起，牙齿特征点识别方法也从传统方法逐渐转向为机器学习的方法。

Mangan利用分水岭算法对牙齿曲面进行切割，得到拥有不同曲率的区域，进而得到牙尖点的位置[5]。之后，Kumar[6]等人在Mangan的分水岭算法上进一步改进，提出新的高度函数，并取每个区域的局部最大值作为牙尖点。但是该算法会导致划分结果过多，需要区域之间进行合并操作。于是，仲哲等人[7]提出一种基于Z轴夹角判断的牙尖点识别算法。该方法对每颗牙齿的局部坐标进行标记后计算各个顶点的法向量，通过筛选法向量和Z轴的夹角来得到牙尖点位置。

传统算法对牙齿类型具有较高的要求，难以适应实际临床需求，病人的牙齿形状多变难以有统一的标准。于是越来越多人采用机器学习的方法来解决三维口腔问题。

深度学习是机器学习中的一个分支，以人工神经网络为架构，对数据进行表征学习的算法。它能够使用半监督式或者非监督式的特征学习和分层特征提取高效算法来替代手工获取特征。至今已有数种深度学习框架，如深度神经网络、卷积神经网络和深度置信网络和循环神经网络已被应用在实例分割[8-13]、物体检测[14-16]、分类任务[18]与医疗处理[19-20]并取得了良好的效果。

目前已有许多使用深度学习网络的方法实现3D扫描口腔内牙齿的自动分割和识别[21-28]。Charles R. Qi和Hao Su等人提出一种简单快速的基于点云分割和分类的深度学习网络[11]，直接处理点云数据可以解决3D识别的物体分类、语义分割等多项任务。网络被广泛应用于多个计算机视觉领域：Charles R. Qi和Li Yi等人又对PointNet网络进行改进提出[29]，使得该网络能更好地考虑局部信息来识别物体; Xinxin Chen和 Kang Jiang等人对改进应用于无人机载激光雷达扫描点云数据学习识别分类树木[30]； Chunfeng Lian和Li Wang等人对进行拓展提出一种端到端的深度神经网络应用于牙齿分割，称为网络[18]；赵庶旭、罗庆和王小龙等人对区域卷积神经网络[31]和[32]网络的改进作用于牙齿的分割与分类[33];Chunfeng Lian和Tai-Hsien Wu等人提出了深度多尺度网格特征学习的用于自动牙齿分割的网格深度学习[23]；学习 变换以重组相邻顶点/单元以进行一般卷积操作，已成功应用于自动标记牙齿表面[34]。

**1.3 本文主要工作**

本文使用基于机器学习的方法实现牙齿特征点的自动识别，根据临床需求设计54个牙齿特征点。基于传统方法自动识别牙齿特征点的算法具有局限性，对牙印模型的三维图像的特征进行识别的方法噪声大，基于区域分割的牙齿特征点识别算法识别出过多的特征点，基于Z轴夹角的特征点识别方法准确率有待提升，且以上传统算法只能识别单类特征点，比如牙尖，这远远不能够满足临床需求：重新排列牙齿并确定排列结果。目前已有许多研究使用机器学习的方法应用于牙齿三维切割和识别，但是基于机器学习的牙齿特征点识别算法还缺乏探索和创新。本文实现使用机器学习中的深度学习方法，搭建神经网络架构，训练模型解决牙齿特征点自动识别，对数据集进行筛选和过滤，提出54个牙齿特征点帮助定位前牙切缘和后牙中央凹槽，标记大量牙齿数据集，并在此数据集上展开大量实验，提升精度，分析结果。

本文的工作有：

1.对计算机医疗领域中的牙齿分割和识别任务进行调研，总结各类特征点类型，并根据临床需求提出54类特征点类型。

2.对牙齿数据集进行过滤筛选，使用软件对3D牙齿网格模型进行特征点标记，标记有30个病人的牙列，共420颗牙齿。

3.对特征点IGES文件和单个牙齿STL文件进行数据预处理，对比分析点云下采样方法和网格简化方法，设计算法提取牙齿的指定网格数目，生成所需要的网络输入矩阵和目标高斯热图矩阵。

4.改进基于点云分割的深度学习网络模型，使其适用于牙齿网格模型，使用高斯热图进行回归。

5.在标记的牙齿网格数据集上展开大量实验，提升性能，从各个方面分析实验结果，提出自己的改进意见。

**2 牙齿特征**

**2.1 牙齿解剖学概念**

每个牙列共有14个牙齿左侧牙齿的序号为11~17，右侧牙齿的序号为21~27，牙齿序号的尾数决定牙齿的类型。牙齿分为第一门牙、第二门牙、犬牙、第一前磨牙、第二前磨牙、第一磨牙和第二磨牙。每个牙弓，即一排牙齿，可以分为左右两侧。每侧有两个门牙（1号和2号），一个犬齿（3号），两个前磨牙（4号和5号）和两个磨牙（6号和7号）。门牙和犬齿统称为前牙，用于切割动作。前磨牙和磨牙称为后牙，参与咀嚼动作。在正畸过程中，医生通过前牙的切缘和后牙的凹槽来定位排列牙齿。

**切缘**：这些是沿着近远中线延伸的前牙（门牙和犬齿）上的尖锐脊。图 3 显示了一个示例。

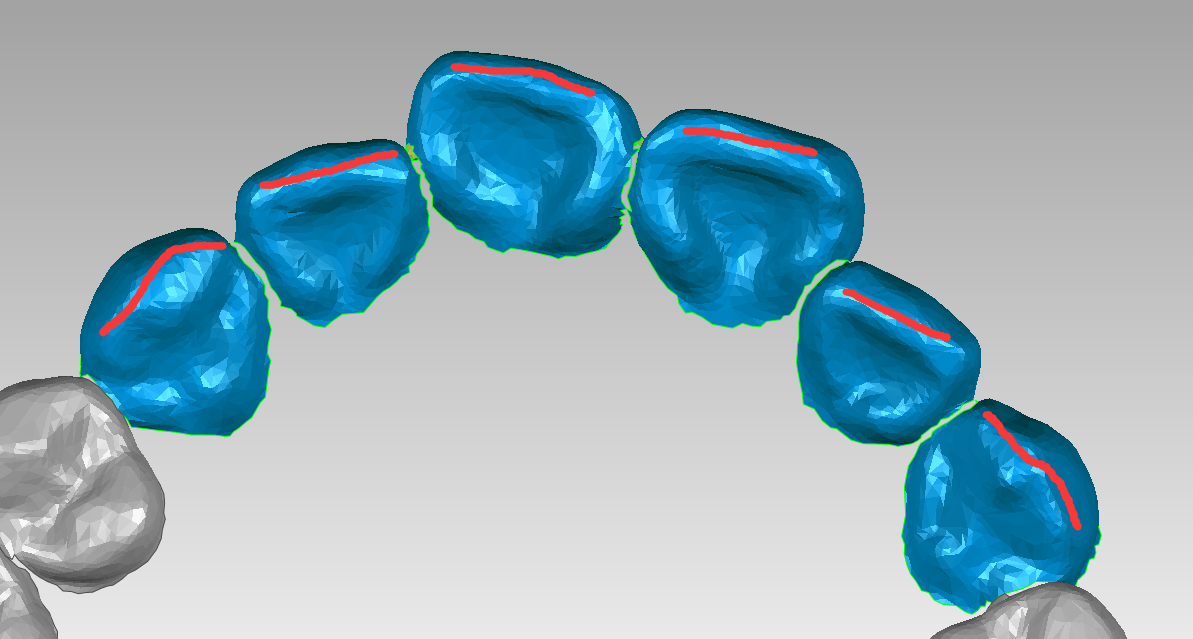


图 1 前牙切缘，如红线所示

**牙尖**：这些是表面上的山峰状结构，位于牙齿的角落。每个尖峰都有从其尖端辐射的尖峰脊，类似于将山峰连接到地形上的山谷的脊，本文利用前磨牙和磨牙的牙尖来定位识别牙齿的凹槽位置。犬齿有一个牙尖，它在确定对齐的整体质量方面起着重要作用。前磨牙和磨牙的牙尖数目不固定，取决于牙弓和个体。

**凹槽**：这些是后牙咬合面上的凹陷和裂缝，类似于地形上的河床和山谷。凹槽有多种类型和相应的分类命名约定。在我们的研究中，我们对沿着牙齿近远中线延伸的长凹槽感兴趣，称为中央发育凹槽，或只是中央凹槽。

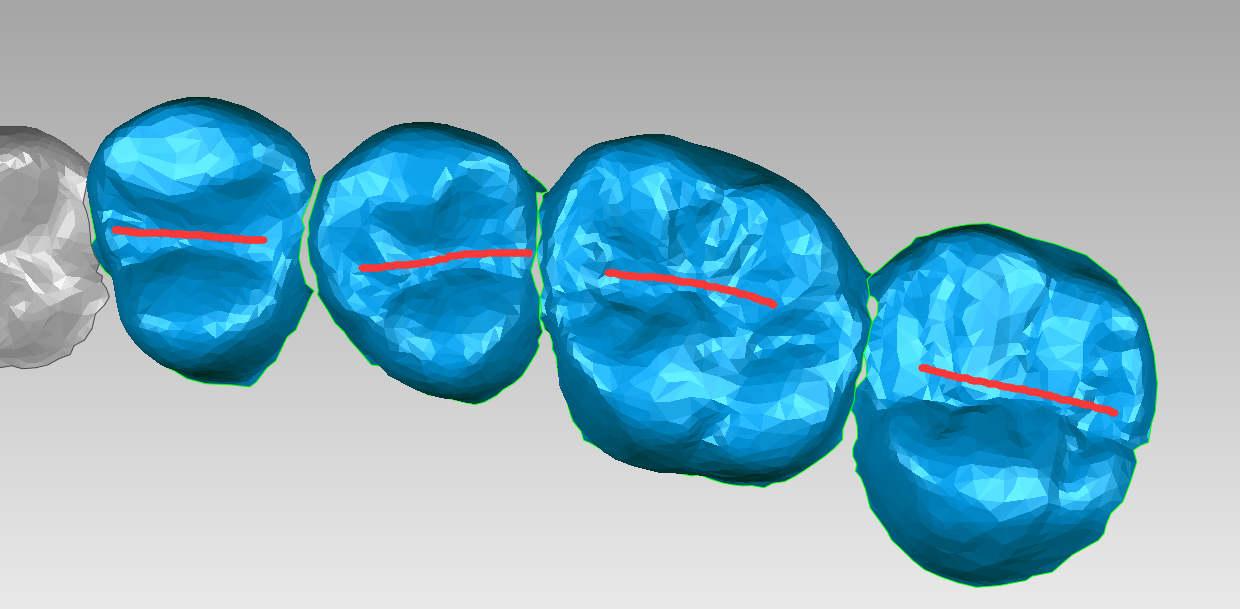


图 2 后牙的中央凹槽，如红线所示

为了方便定义特征点的前后关系，本文使用到近中、远中、近颊、近舌的牙齿朝向的解剖概念，近中远中和近颊近舌定义了一个牙齿的两个垂直方向。

**近中**：靠近中间门牙的一侧。

**远中**：远离中间门牙的一侧。

**近颊**：靠近脸颊，远离舌头的一侧。

**近舌**：靠近舌头，远离脸颊的一侧。

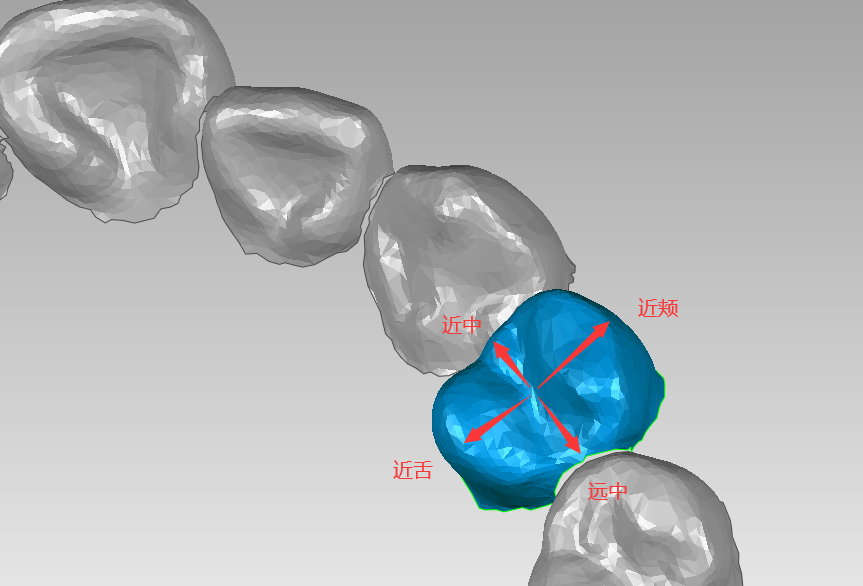


图 3 第一前磨牙的近中侧、远中侧、近舌侧和近颊侧

**2.2 牙齿数据来源**

本文使用到的牙齿数据是由华西口腔提供，先是使用激光扫描由病人特定的牙印建立的牙弓石膏模型获得带有噪音的牙齿三维表面后，对三维表面进行分割得到单个的牙齿对象。我从中挑选满足要求（一个牙列有14颗完整的牙齿）的上牙列进行标记，得到30个左右的上牙列，即420个牙齿。本文关于特征点的定位是我与华西口腔的医生在沟通中确定的特征点类型，通过这些特征点进一步识别前牙的切缘与后牙的中央凹槽。这些特征对进行调整和评估此次调整的质量都起到重要作用。

**2.3牙齿特征点**

牙齿特征是医生在正畸治疗中用于重新排列病人牙齿的重要指标。这些特征提供了一组地标，可用于定量定义对齐要求。此外，这些特征本身不会随着时间而改变，因此它们可用于监控对齐的进度。此外，由于某些特征需要在整个治疗过程中保持不变，内在特征决定了可以实现的最佳对齐方式，并提供了一种评估计算对齐方式的方法。

目前关于牙齿特征点识别的论文使用的特征点类型不固定，在[]文章中一个牙列共有66个特征点，是由专业的医生标记。但是由于其数据的保密性，因此本文无法使用该数据集。本文到使用的数据集是由华西口腔提供，并由学生标记的一组数据集合。

本文定义了一个牙列上54个特征点，按照近中到远中，近颊到近舌的顺序有序标记。以左侧牙齿（11~17号）为例，11号和12号门牙各有两个特征点位于切缘两端；13号尖牙有三个特征点，分别位于切缘两端和中间的尖点端；14号和15号前磨牙各有6个特征点，分别位于中央凹槽的两侧边缘线的两端和边缘线上的尖点；16号和17号磨牙各标注有4个特征点，分别位于四个尖点上。右侧牙齿同理，一个牙列就有54个特征点。

有些牙齿因为后天磨损或者先天原因，其特征点并不明显，并非一定要有尖点，须根据实际情况而定。本文标记的特征点的目的是确定前牙的切缘和后牙的中央凹槽而非确定尖点位置。我一共标记有30组左右的上牙列，每个牙列有14个牙齿。

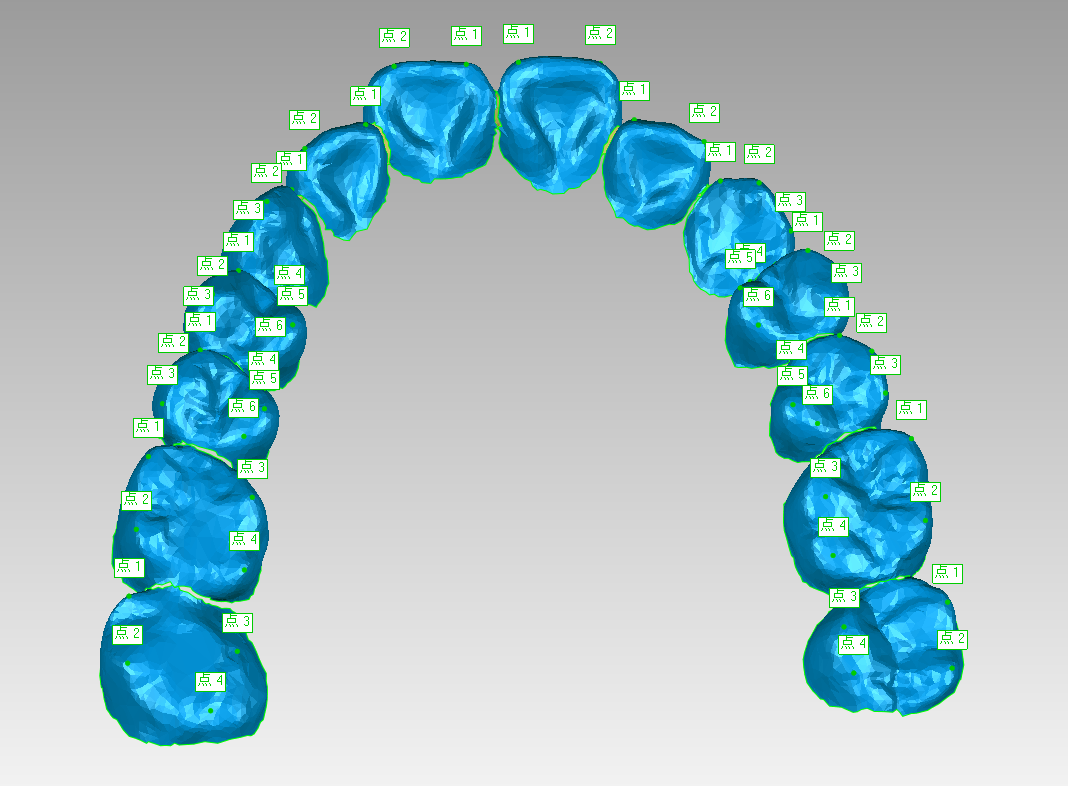


图 4 一个牙列的54个特征点

我使用处理牙齿和标记牙齿的特征点。 是公司的一款逆向软件产品，可根据任何实物零部件通过扫描点云自动生成准确的数字模型，是一款自动化逆向工程软件，广泛应用在在众多工业领域，比如汽车、航空、医疗设备和消费产品领域。该软件结合了实时三维扫描、三维点云和三角网格编辑功能以及全面CAD造型设计、装配建模。它可以输出行业标准格式，包括STL、IGES、STEP和CAD等众多文件格式。本文主要用到STL和IGES两种文件格式。

**2.4小节**

本节介绍了牙齿的解剖学概念和已有的特征点类型，再根据医生的临床需求而制定出一个牙列的特征点类型、标记顺序。阐述了在特征点标记过程中使用到的软件和导出格式等，为之后的算法介绍做铺垫。

**3 关键技术**

机器学习是用数据或以往的经验，以此优化计算机程序的性能标准。给定世界的有限或者无限的观测对象的集合，通过该世界的一个有限的子集合，即样本集，来推算整个世界的模型，使它对这个世界为真。

人的神经系统可以大致分为三阶段：感受器、神经网络和效应器。携带信息的刺激信号从左向右通过系统向前传输，从右到左的箭头表示神经系统中的反馈。感受器把来自外界的刺激信号转换为点冲击，向神经网络传输信息。神经网络的效应器将电冲击转换为可识别的响应输出。

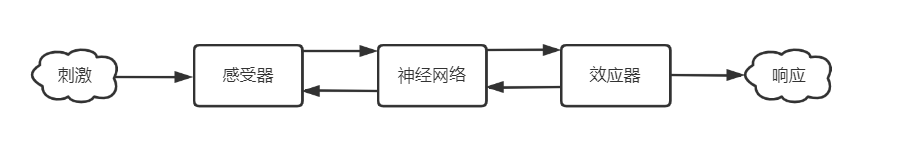


图 5 神经系统的三阶段

类神经网络仿真人脑的神经突触传递结构的工作原理，透过脉冲传递信息。神经元之间的每个连接可将信号传递给另一个神经元，处理信号的方式通过树突向连接到它的下游神经元发出信号，突触具有随着学习的进行而变化的权重，从而增加或减小它发送到下游的信号的强度。

**3.1神经元模型**

神经元是神经网络操作的基本信息处理单位，是神经网络模型的基础。神经元模型由三种基本元素：突触、加法器和激活函数。

**突触**：每一个突触都有一个权值作为特征。在连接到神经元的突触上的输入信号乘以突触的权值。

**加法器**：用于求输入信号被神经元的相应突触加权的和，构成一个线性组合器。

**激活函数**：用来限制神经元输出的振幅，将输出信号限制到允许范围内的一定值，通常取或者。

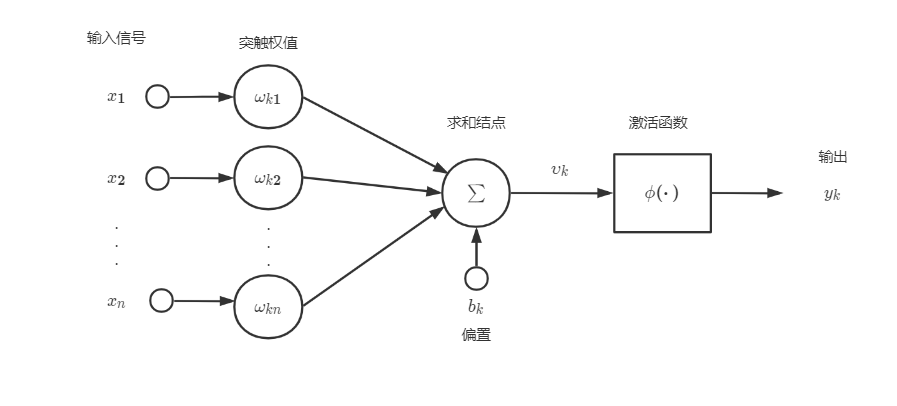


图 6 神经元模型，第k个神经元

图中是输入信号，是神经元k的突触权值，通过线性组合器得到输入信号根据其对应权值相加得到输出，再根据偏置值对线性组合器的输出作仿射变换，即:

。

经过激活函数将输出范围限制在指定区间内得到输出，有:

**3.2激活函数**

激活函数的类型：

1. 阈值函数：若输入大于等于0取1，小于等0取0，即：
2. Sigmoid激活函数：又称Logistic函数，在神经网络中最常用的激活函数，用于隐藏层神经元的输出。该函数图像呈S型，将输出范围限制在区间，易于求导，严格递增，在线性和非线性的行为之间展现除了较好的平衡性。

sigmoid函数的定义如下：

其中是sigmoid函数的倾斜参数，可以决定sigmoid函数的倾斜程度。

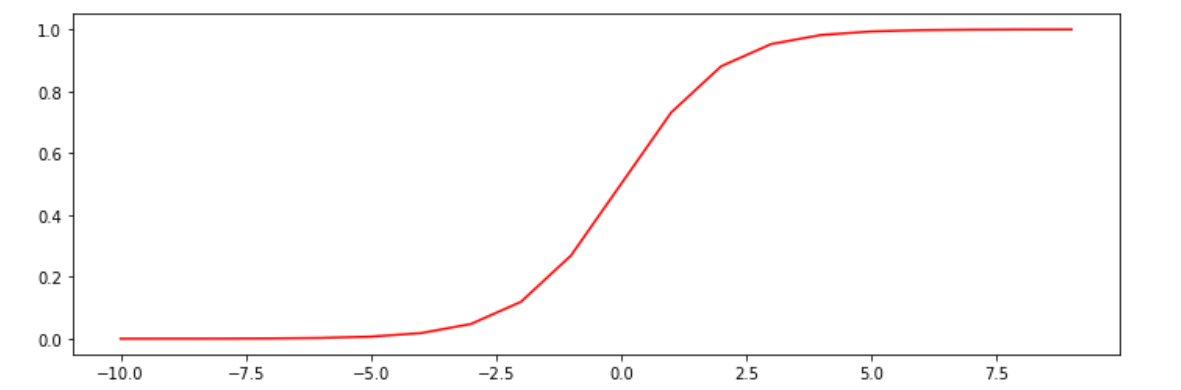


图 7 取1的sigmoid函数图形

1. softmax激活函数：又称归一化指数函数，是有限项离散概率分布的梯度对数归一化，计算每个可能类别的置信度。主要应用在多项逻辑回归，多项线性判别分析，朴素贝叶斯分类器和人工神经网络等的多种基于机率的多分类问题方法。样本向量属于第个分类的概率为：

5.ReLU函数：又称线性整流函数，线性整流被认为有一定的生物学原理，并且由于在实践中通常有着比其他常用激励函数（譬如Logistic函数）更好的效果，而被如今的深度神经网络广泛使用于诸如图像识别等计算机视觉人工智能领域。

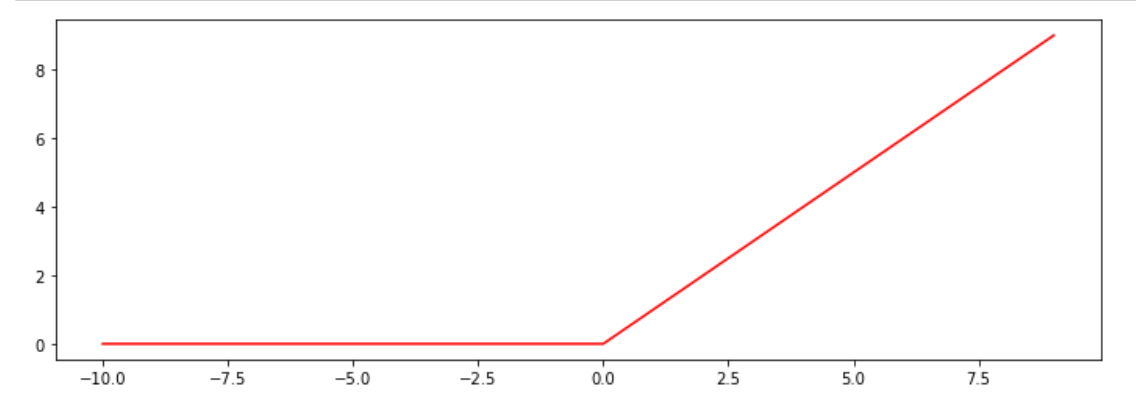


图 8 线性整流函数图形

**3.3多层感知器**

Rosenblatt感知器本质上是一个单层的神经网络，只能解决线性可分模式的分类问题。Minsky等人于1961年提出对Rosenblatt感知器的批评，他们之处Rosenblatt感知器本质无法推广到二进制数的奇偶校验对的情况，更不能做到一般的抽奖，后来Minsky等人也用数学分析证明Rosenblatt感知器无法进行全局的泛化。

多层感知器（,缩写）是Rosenblatt感知器的推广，克服了感知器的局限性。它是一种前向结构的人工神经网络，映射一组输入向量到一组输出向量，由多个节点层组成，每一层全连接到下一层，除了输入节点，每个节点都是一个带有非线性激活函数的神经元（或称处理单元）。常被MLP用来进行学习的反向传播算法，在模式识别的领域中算是标准监督学习算法，并在计算神经学及并行分布式处理领域中，持续成为被研究的课题。MLP已被证明是一种通用的函数近似方法，可以被用来拟合复杂的函数，或解决分类问题。

MLP在80年代的时候曾是相当流行的机器学习方法，拥有广泛的应用场景，譬如语音识别、图像识别、机器翻译等等，但自90年代以来，MLP遇到来自更为简单的支持向量机的强劲竞争。近来，由于深度学习的成功，MLP又重新得到了关注。

多层感知器的特点如下：

1. 网络中每个神经元模型包含一个可谓的非线性激活函数
2. 网络中包括一个或者多个隐藏在输入和输出神经节点之间的层
3. 网络的突触权值决定了网络高度的连接性。

多层感知机首先学习，然后使用权重存储数据，并使用算法来调整权重并减少训练过程中的偏差，即实际值和预测值之间的误差。主要优势在于其快速解决复杂问题的能力。

以只有一个简单全连接的多层感知器MLP（m,M）为例，如图9所示。N表示多层感知器的输入节点数目，隐藏层有m个神经元，最后网络的输出层神经元的数目为M，相当于将N维的输入空间映射到M为的输出空间中。每个连接对应一个权值，神经网络在学习过程中不断调整的大小。

**3.4小节**

本节着重介绍了机器学习中的神经网络的重要基础知识。分别介绍了网络中常用的激活函数：阈值函数、Sigmoid函数、Softmax函数和线性整流函数，以及这些激活函数的特性和使用范围。接着，本节介绍了神经网络应用到的相关技术和算法，从Rosenblatt感知器再到多层感知器的发展，以及反向传播算法是如何训练多层感知器，为下节介绍本文使用的基于机器学习中的神经网络算法奠定基础。

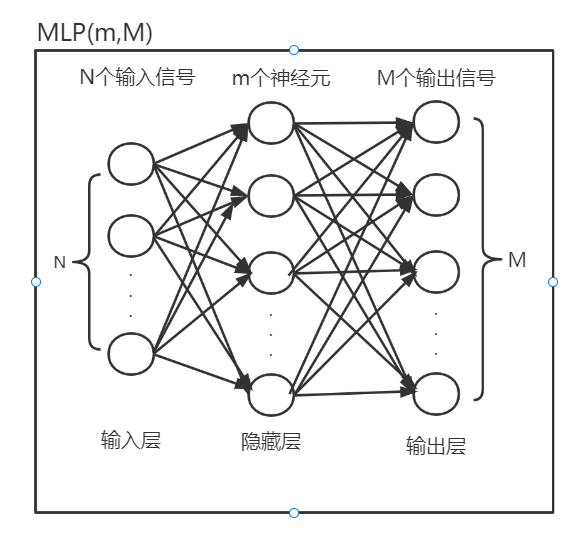


图 9 多层感知器MLP（m, M）

**4 算法设计**

本文对深度学习网络[]进行改进，结合高斯热图回归应用于牙齿特征点识别任务中。该算法的主要步骤有：对于网格STL文件和目标IGES文件的数据预处理，对，每一个牙齿数据进行转换得到的矩阵作为改进后的网络的输入，经过多次特征变换和多层感知机运算，输出预测的牙齿高斯热图矩阵，本文对热图矩阵进行后处理将对应网格转换为预测的特征点位置。

**4.1 数据的获取与处理**

网格是一种三维建模的重要数据结构体，也是本文的主要数据来源。计算机生成的三维模型和实际获取的数据表示模式是不同的。计算机生成的模型可能是平滑曲线曲面，而实际获取的数据，如激光扫描得到的，一般都是以点云的形式存在。图形学中需要一个统一的表示方式，同时要求视觉精度和处理速度都在可以接受的范围内。于是就选择了网格，用多边形来近似曲面，三角网格最为简单高效，再加上图形硬件的快速发展，三角网格和光栅化已经可以嵌入到硬件中去渲染。每一个网格有三个顶点信息和法向量信息，本文使用的网格数据是表面模型，只考虑物体的表面细节并直接进行处理，易于显示。三维牙齿网格是对病人口腔模型进行激光扫描后得到的整体口腔三维网格模型，然后进一步分割为单个牙齿模型从而得到单个牙齿的网格文件。我们选取的数据必须是具有完整14颗牙齿的牙列，单个牙齿以STL(Stereo lithographic)格式保存。

STL文件格式是美国3D SYSTEMS公司提出的三维实体造型系统的一个接口标准，是采用一系列三角形面片离散地、近似地表示三维实体模型。目前被工业界认为是3D打印领域的标准描述文件格式，在逆向工程、有限元分析、医学成像、增材制造等方面有广泛的应用。

STL文件是由一系列的三角形面片无序排列组合在一起的，没有反映三角形面片之间的拓扑关系。

目前的STL文件格式包括二进制格式(BINARY)和文本格式(ASCII)。一个完整的STL文件记载了组成实体模型的所有三角形面片的法向量数据（3个浮点数分量，用于确定三角面片的正反方向）和顶点坐标数据信息（3\*3个浮点数分量）。在该项目中是以二进制的格式保存。

二进制文件格式——文件起始的80个字节是文件头，用于存贮文件名；紧接着用 4 个字节的整数来描述模型的三角面片个数， 依次给出每个三角面片的几何信息。每个三角面片信息占用固定的50个字节，依次是: 3个4字节浮点数(法向量数据) 、3个4字节浮点数(顶点1笛卡尔坐标)、3个4字节浮点数(顶点2笛卡尔坐标) 、3个4字节浮点数(顶点3笛卡尔坐标)，最后2个字节用来描述三角面片的属性信息，一般只保留位置，补齐50字节。

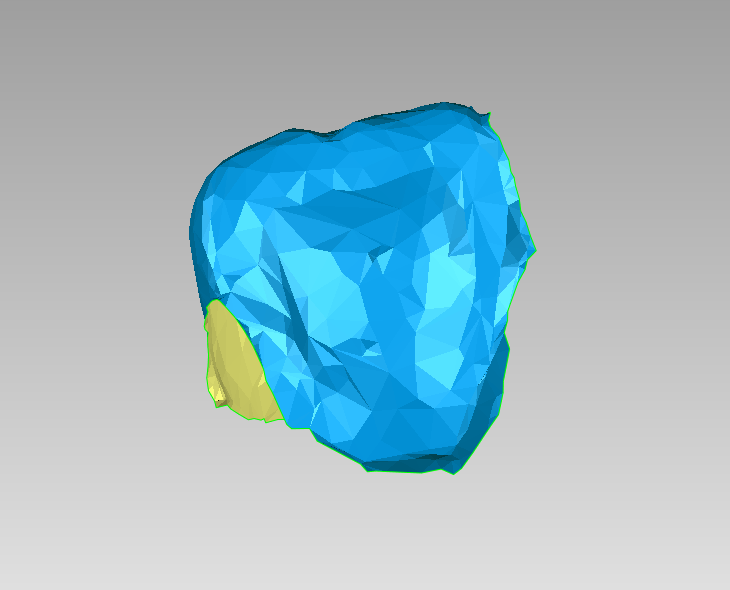


图 10 STL格式的牙齿网格文件

读取牙齿网格STL文件后，每一个牙齿的网格大小不一，从2000个网格到500个网格不等。需要使用采样方式来统一牙齿网格的数量，输入到改进的 网络中。

普遍的下采样方法是面向点云数据，可以将网格数据转为点云数据执行下采样方法，和都有对点云数据下采样函数的封装处理，点云数据的下采样方法如下：

1. 体素算法：确立一个三维体素栅格的大小，将三维点云转为一个个体素栅格，用每个体素的重心或者中心来近似显示体素中其他点，这样将体素内所有点转换为该体素的中心点坐标，即用体素的重心点坐标表示体素包含的全部点。若使用体素的质心点去近似体素内的所有点，虽然该方法更慢但是质心能够更准确表示底层表面信息。
2. 均匀采样：指定网格中采样的点云数目，对三角形网格的三维表面均匀采样。

将牙齿从网格形式转为点云形式如图11所示，通过点云下采样后得到指定点云个数的点云文件，再从点云形式转为网格形式需要点云表面法线估计来计算表面的法向量生成网格文件。可以使用最小二乘法平面拟合估计或者使用PCL法线估计方法。PCL法线估计是通过寻找待估计点与其邻域点共同构成的局部平面的法向量实现的。通过主成分分析，最小特征向量就是该点法向量，通常法向量具有二义性，所以需要指明法向量的朝向。

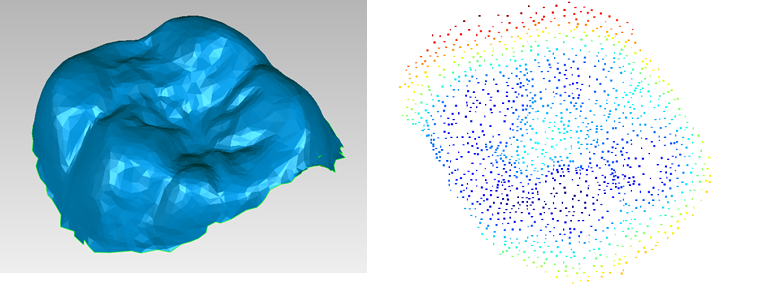


图 11 牙齿从网格数据转为点云数据

但是该转换过程会丢失原网格数据的法向量信息。点云数据独立看待各个顶点，没有考虑各顶点之间的关系。而网格数据保留了点与点之间的关系，能够更好地表示物体表面信息。使用点云的下采样方法得到的点云文件做法向量估计得到的网格文件是不准确的。此外，对点云采样只能确定点云数目转换为网格形式后网格的数目不单一。

对网格直接进行操作使用较少的三角形来表示一个高分辨率的网格并最大程度保证信息不丢失，这类操作称为网格简化。网格简化的操作有如下方法：

1. 顶点聚类：确立一个三维体素栅格的大小，将所有落入给定大小的顶点聚集到单个顶点，该顶点可以是体素内各顶点的平均值或者到相邻平面最小距离的顶点。
2. 连通分量: 对连接的三角形（即通过边连接的三角形）进行聚类的函数被分配相同的聚类索引，获得包含每个三角形的簇索引，每个簇的三角形数量，和每个簇的表面积。
3. 网格抽取：选择一个使误差度量最小化的三角形并将其删除。重复此过程直到满足指定的三角形数量时停止。方法有Garland和Heckbert的二次误差度量抽取提出的最小化误差平方（到相邻平面的距离）[36],和William等人提出的三角形网格抽取方法[37]。

使用顶点聚类方法只能根据输入网格空间和指定的体素大小得网格数目，但是每个牙齿的网格空间是不固定的，因此无法得到指定数目的网格。比如如图13所示的原牙齿，有5567个顶点和1889三角形的切牙网格数据，经过体素大小为0.3的顶点聚类后得到885个顶点和1684个三角形的网格数据；而一个8157个顶点和2719个三角形的磨牙网格数据通过0.3大小的体素的顶点聚类得到1218个顶点和2331个三角形。所以顶点聚类得到三角形的顶点个数和三角形的网格个数是不可控的，无法获得指定数目的网格数量。

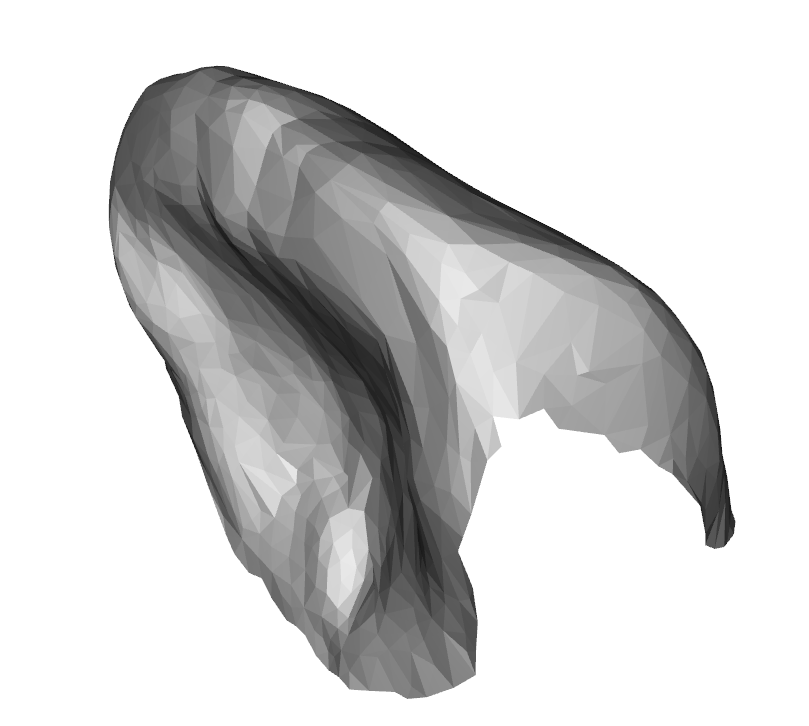


图 13 原牙齿

连通分量的网格简化方法只适用于有多个连通分量的输入，多用于取出网格数据中的噪声。本文的数据集合处理对象是连通的单个牙齿，所以无法使用连通分量操作简化网格。

使用最小误差平方的网格抽取方法虽然可以指定三角形网格的数量，获得指定三角形数目的牙齿网格数据，但是大大损坏原有牙齿的拓扑结构，如图13所示的未处理的切牙经过网格抽取至1000个三角形，得到如图14所示的结果，可以看到造成大量的空洞。

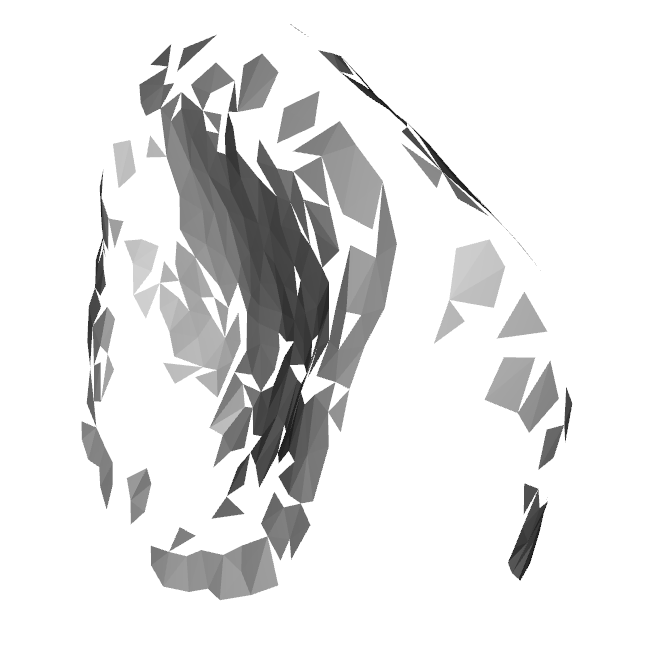


图 14 最小误差平方的网格抽取方法[10]后的牙齿

将图11所示的牙齿经过William[36]等人提出的对三角形网格抽取方法抽取，设置要削减的三角形数目的百分比，

如图15所示的牙齿网格有1889个三角形网格，目标三角形网格数目为1000，所以。抽取后得到有2997个顶点和999个三角形网格的数据。该抽取方法最大程度地保留了牙齿的拓扑结构，没有造成空洞，最能符合本文需求。然而，该方法只能够按照比例抽取，得到的结果不一定是目标三角形数目，比如处理后的图15的牙齿是999个网格而不是目标数1000个网格，所以William的网格抽取结果会在目标网格数目上下浮动。

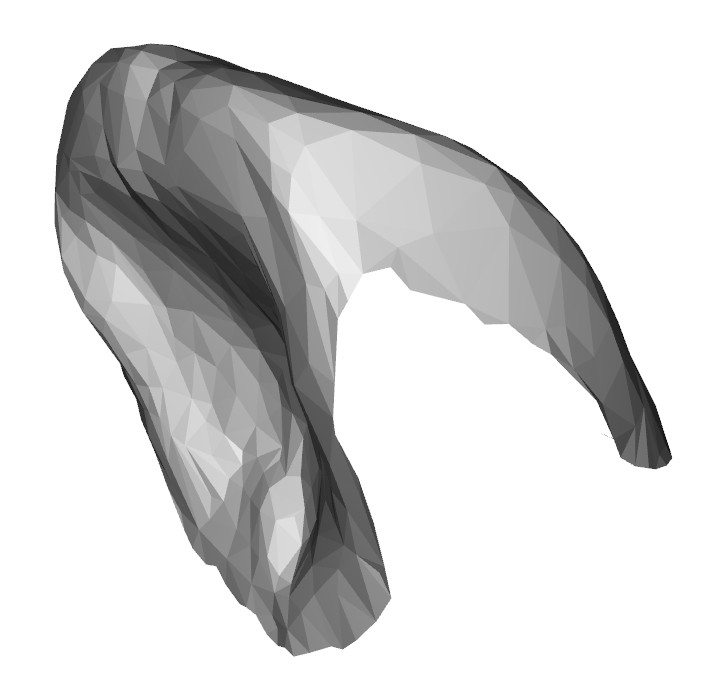


图 15 使用William提出的网格抽取方法后的牙齿

为了抽取后得到固定的网格数，本文制定的网格下采样流程图如图16所示。首先设置下采样的目标网格数目，初始化参数为0，根据式（）计算出削减三角形数目的比例：

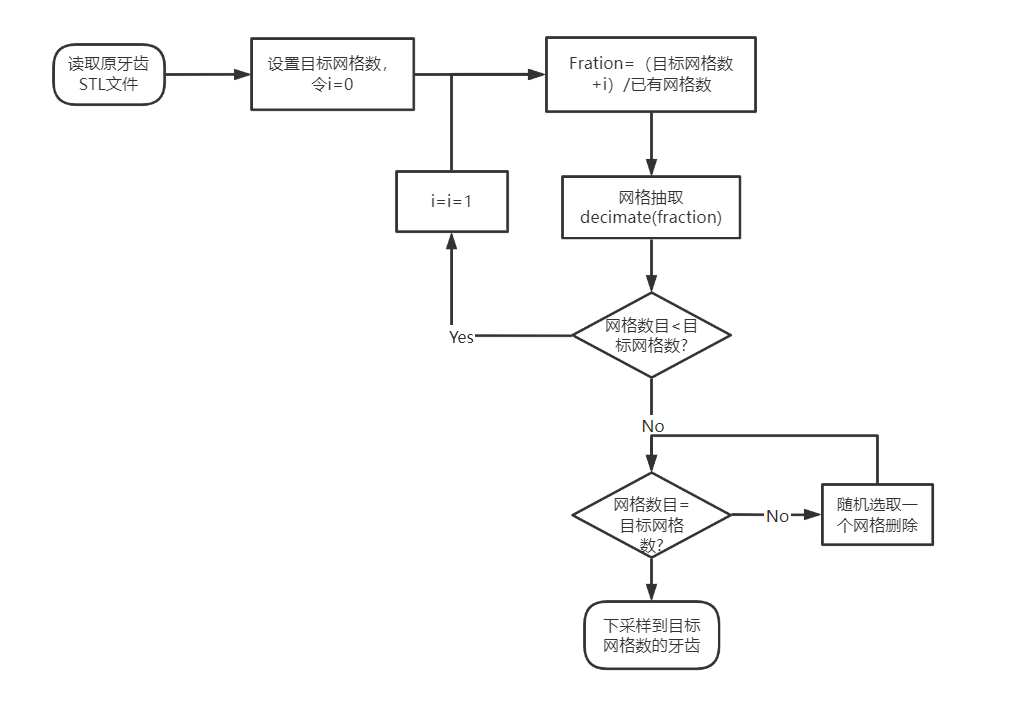
执行网格抽取，若得到的牙齿网格数目小于目标网格数目则令，重复上述操作，直到网格抽取后的牙齿网格数目大于目标网格数目。对于多出的网格数目采用随机抽取法删除直到得到指定目标数目的牙齿网格文件。

图 16 网格下采样流程图

对于每一个牙齿执行网格抽取后得到指定数目的统一网格后，对每一个网格进行计算，得到的输入矩阵。矩阵中的第1列到第9列表示三角形三个顶点的坐标值，第10列到第12列表示网格的法向量，第12列到第15列表示网格的中心点，以上数据都是经过正则化后的结果。

每个牙齿的特征点坐标以文件格式进行保存，需要根据文件的格式形式读取文件中的特征点坐标进行数据处理。

文件由五或六段组成：

（1）标志（FLAG)段；

（2）开始（START）段；

（3）全局（GLOBAL）段；

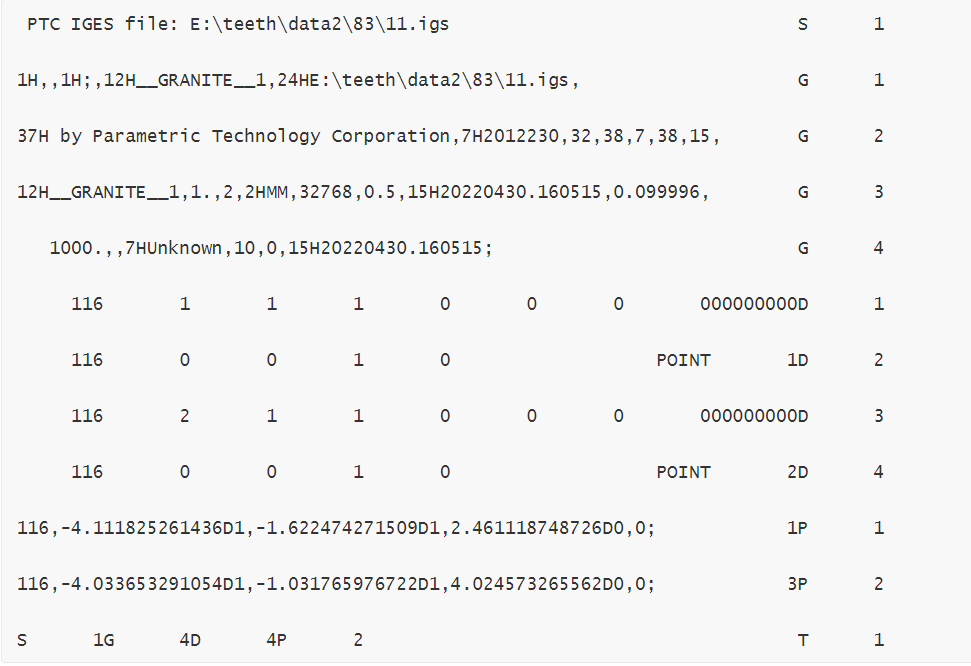
（4）元素索引（DIRECTORY ENTRY）段；

（5）参数数据（PARAMTER DATA）段；

（6）结束（TERMINATE）段。

在数据交换文件中表示信息的基本单位就是元素，每种元素都有唯一的元素类型号与之对应。元素类型号0000到0599和0700到5000由IGES标准本身使用；元素类型号0600到0699和10000到99999作为宏元素。特征点以116类型号进行保存。

文件每行80个字符。每段若干行，每行的第1~ 72个字符为该段的内容；第73个字符为该段的段码；第74~80个字符为该段每行的序号。段码的规定如下：字符“B”或“C”表示标志段；“S”表示开始段；“G”表示全局段；“D”表示元素索引段；“P”表示参数数据段；“T”表示结束段。图17展示了单颗牙齿特征点真实值文件。



**图 17 特征点的IGES文件内容**

读出IGES文件中的参数数据段，找出116类型号开头的行，读出对应点的三个坐标数据并转换为浮点数。如图17所示，IGES文件中点数据找是以科学计数法的形式存储，其中D表示10的幂。读取出IGES文件中的特征点坐标后，需要对数据进行转换处理得到对应浮点数值，转换流程如图18所示：

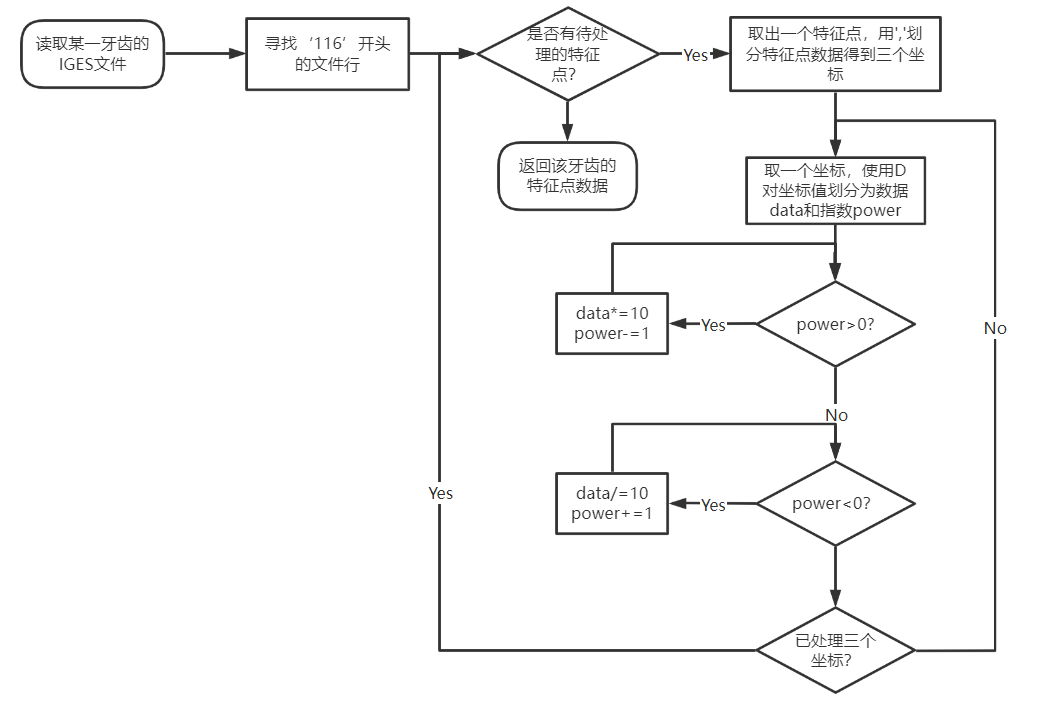


图 18 IGES文件特征点处理流程图

一个IGES文件保存一个牙齿的所有特征点坐标，首先读取IGES文件后找到以‘116’开头的行，每一行对应一个特征点的坐标。对每一个特征点坐标以‘,’分割，得到三个坐标值，对每一个坐标值用‘D’划分分别得到数据部分和指数部分。若指数部分为正数，则将数据部分每次乘10，指数减1，直到指数为0；若指数为负数，则将数据部分每次除以10，指数加1，直到指数部分为0。重复上述步骤直到该IGES文件的所有特征点数据处理完毕。

**4.2 高斯热图**

牙齿特征点的检测本质上是对于关键点的检测。在传统的回归方式是直接用目标点坐标进行监督，这样网络是一个坐标点回归网络，最后用一个全连接层输出目标点的坐标集合。

使用高斯热图矩阵去监督，网络输出的是特征矩阵，在检测点的位置为1，靠近检测点附近的值也较大，一直向周围递减到0。多个点的检测就输出多个通道的特征图，特征图矩阵的每一列就对应一类牙齿特征点。Nibali和Aiden等人对比了热图回归和坐标点回归进行了深入分析[38]。

在牙齿特征点识别中，实际上特征点很难准确地被某一个点的位置定义，很难被准确地标注。如果将特征点周围的点直接设置为负样本，会给网络带来干扰，所以我使用高斯函数对其进行“软标注”，使得热图矩阵的值以特征点为中心向外递减到0。通过高斯热图，网络不仅能够更快地收敛，也给网络的训练增加了一个方向性的引导，距离特征点越近，高斯热图的值就越接近1，网络就能有方向地快速达到目标点。

基于热图回归的特征定位已应用于医学成像领域。Payer和Stern等人将医学解剖标志的空间信息合并到基于热图的卷积网络回归框架中，使得在有限的训练数据集上也能够得到较低的误差[37]。张军和刘明霞的等人使用的用于骨骼分割和特征数字化定位的联合学习的卷积神经网络架构，也使用热图回归来数字化骨骼特征[39]。

我使用高斯函数计算出高斯热图矩阵作为网络的真实值，其计算方式如下：

其中表示第i个牙齿网格的中心坐标，表示第k类特征点的坐标，表示高斯高度，是高斯高度的最大值，即1，是高斯宽度的均方根，设为5.0mm。

第一步是对牙齿进行预处理。读取对应的牙齿网格，网格进行下采样，将每一个牙齿的网格数目下采样到500个网格。对每一个网格计算出对应的15维数据：网格的三个顶点的坐标值（9个），正则化的法向量（3个）和相对于整个表面的网格中心点（3个）。每一个牙齿得到了500×15的矩阵，并对整个矩阵进行z-score归一化。

**4.3 神经网络模型**

PointNet是斯坦福大学研究人员提出的一个基于点云的深度学习网络模型，该网络结构简单能够直接输入点云数据进行处理，在物体分割和识别取得相当好的效果。提出了一种新型的处理点云数据的神经网络模型——PointNet，并验证了它能够用于点云数据的多种认知任务，如分类、语义分割和目标识别。本文基于PointNet网络模型进行改进，使其适用于网格牙齿的自动特征点识别问题。Chunfeng Lian和Li Wang等人指出了PointNet网络的缺陷：独立处理不同的顶点而忽略顶点之间的空间依赖性[1]，本文算法对网络进行改进作用于牙齿网格，解决模型的不足之处，使其能够结合顶点之间的相互作用信息，而非单一考虑顶点位置信息。

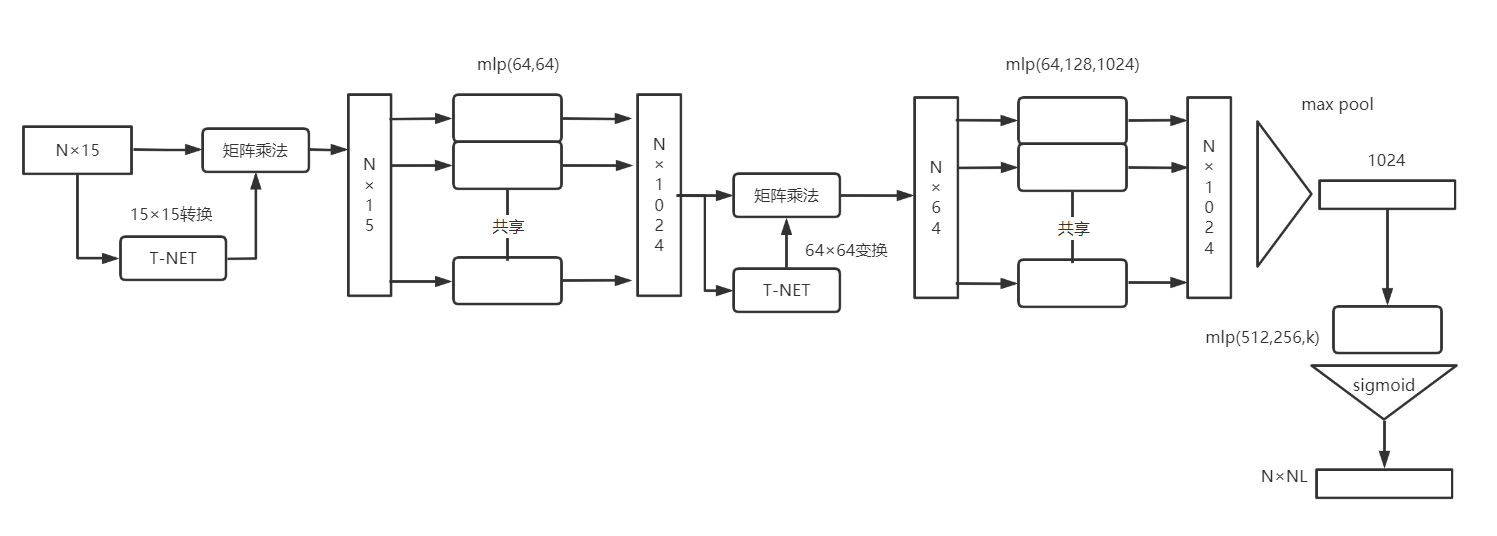


图 19 PointNet网络结构

1、网络的输入不再是一帧的全部点云数据集合（一个nx3的二维张量，其中n代表点云数量，3对应三维x,y,z坐标），而是网格牙齿经过数据预处理后得到的N×15的二维张量，其中N表示牙齿下采样的网格数目。

2、输入数据先通过和一个T-Net学习到的转换矩阵相乘来对齐，保证了模型对特定空间转换的不变性。

3、之后，通过多层感知器MLP（64,64），括号中记录的是感知机的层数。通过多层感知机对各点云数据进行特征提取后，用一个T-Net对特征进行对齐，得到一个N×64的张量，再通过多层感知机MLP（64,128,1024），得到一个N×1024的张量，使用最大池化操作得到全局特征张量1×1024。全局特征层感知机（512,256，k），其中k在原网络中是分类数目，在改进网络中代表牙齿的特征点数目，最后在一维卷积的全连接层中使用的softmax函数替换为sigmoid函数，输出N×NL的高斯热图，NL是特征点数目。

4、本文的网络中，除了最后一层，所有层都使用线性整流函数和批量标准化处理一维卷积后的结果。在网络的最后一层全连接层dropout设置为0.7防止网络训练结果过拟合。

5、第一个T-NET网络由多层感知器MLP(64,128,1024)组成，对每一个网格做最大池化操作后，输入到两个输出大小分别为512和256的全连接层，最后输出为k×k的仿射变换矩阵，输出矩阵初始化为单位矩阵。第二个T-NET网络和第一个T-NET网络有相同的架构，但是输出的是64×64的仿射变换矩阵，

**4.5 后处理**

本文使用改进后的网络去预测输入牙齿对应的高斯热图矩阵，是一个牙列的特征点个数，为54,是牙齿网格数目，为下采样的网格数500。高斯热图矩阵的第行第列的值位于0到1之间，表示第个网格是第类特征的概率。在后处理中，我们寻找拥有高斯热图矩阵的最大值，根据其行数找到对应的网格中心点为预测的特征点位置，根据牙齿的类型取出对应个数的最大值点。

为了解决预测特征点位置集中在单类特征点周围的问题，本文加入阈值判断来避免该问题。阈值判断流程：若识别出的最大值网格中心点的位置与已存在的预测特征点的欧式距离，若该距离小于指定距离阈值，则判定该特征点与已预测的特征点类型发生重复，忽略此最大值，将寻找下一个最大值点位置，直到找到满足阈值条件的位置。

**4.6 小节**

本节详细介绍了算法的设计流程：数据预处理、基于PointNet改进的神经网络和数据后处理。对算法中使用到高斯热图、网格文件处理方式、改进后的PointNet网络框架和阈值处理等算法中应用的技术和知识进行了详细介绍。

**5 实验与分析**

**5.1 数据集**

本文使用的数据集合由华西口腔提供的牙齿口腔数据集上，使用[20]网络对整个口腔的牙齿进行分割，得到每一个牙齿的网格数据。对每一个患者的牙齿进行筛选过滤，只保留有14颗牙齿的上牙列并对其进行标记。每个牙列标记有54个特征点，得到约30名病人的上牙列，共约420个牙齿。本文在实验中分别使用网络一次性预测一个牙列所有54个特征点类型和使用网络只预测前牙（切牙和尖牙）的14类特征点，比对实验效果，分析结果。

**5.2 实验过程**

有效数据集合共有29个上牙列数据，每个牙列有7颗牙齿。24个牙列用于训练，5个牙列进行验证。本节将展示本论文设计的算法作用于该数据集合上的结果并进行分析。本论文采用自适应学习率的Adam优化方法中的AMS-Grad，采用20个牙列（每个牙列14颗牙齿）进行训练。每一颗牙齿下采样到500网格数目，使用MSE损失进行向后传播，Epoch大小为100，每次训练14个数据。在训练中使用均方差损失函数对模型进行收敛，如图13所示。

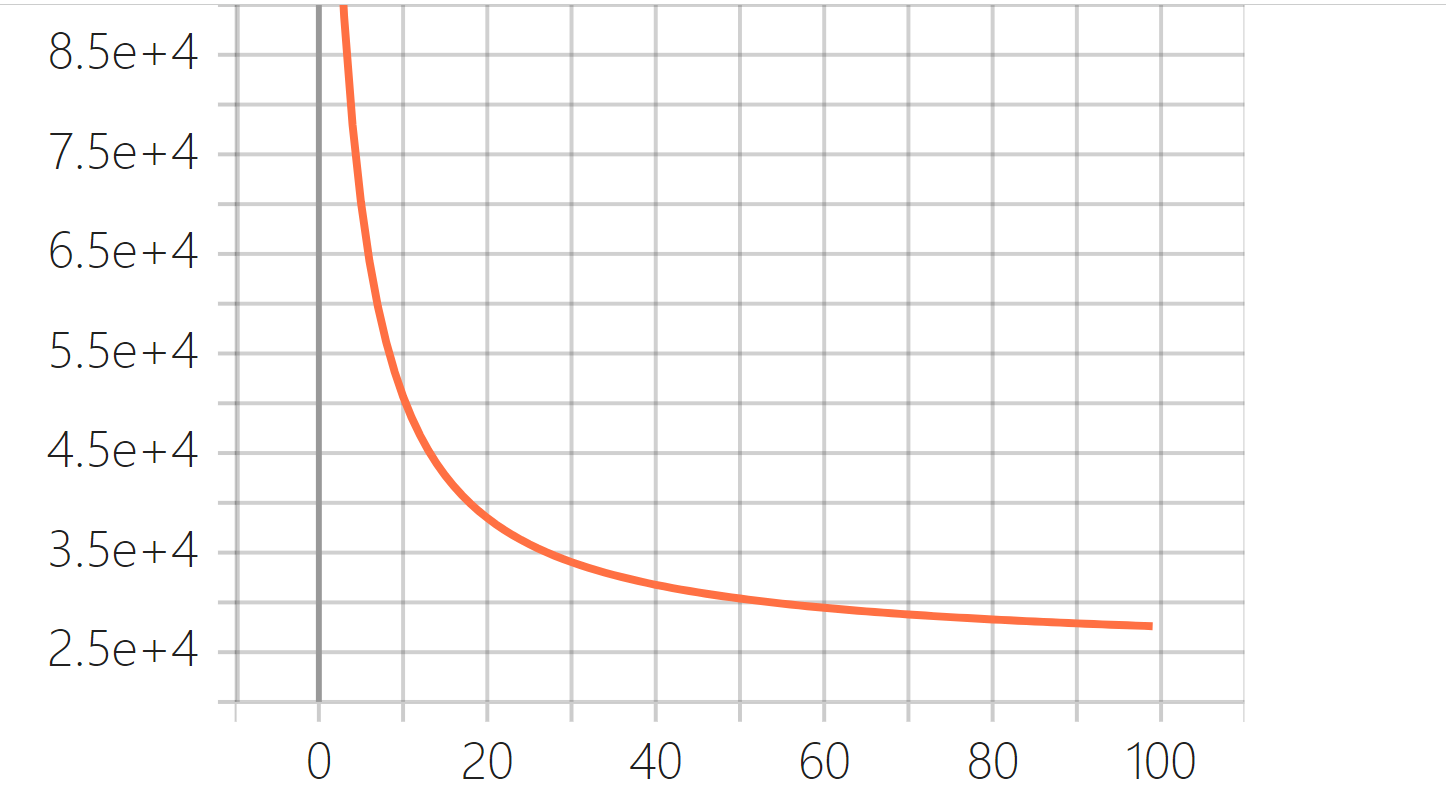


图 20 训练Loss曲线图

在测试集中，预测的特征点集中在某一类特征点周围，如图11的左图所示。比如切牙的特征点应位于切缘的左右两侧，但是预测出的最大值点都集中在了左侧特征点的周围而忽略了右侧特征点。为了解决该情况，在寻找高斯热图的最大值中加入距离阈值thresh的判定，确保预测的各个特征点位置之间大于阈值距离。

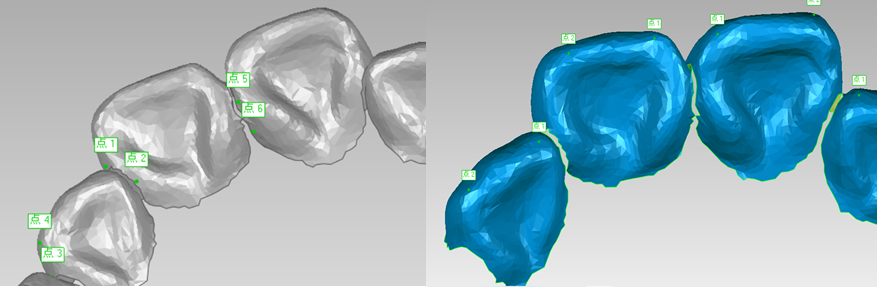


图 21 左图：特征点的预测值（集中在左侧特征点周围）

右图：特征点的真实值

使用6组病人的上牙列进行测试，下面分别列出切牙（2类特征点）、尖牙（3类特征点）、前磨牙（6类特征点）和磨牙（4类特征点）的预测特征点可视化结果。

切牙的预测特征点位置如图22所示，可以看到预测结果都在牙齿的切缘上，符合预期要求。

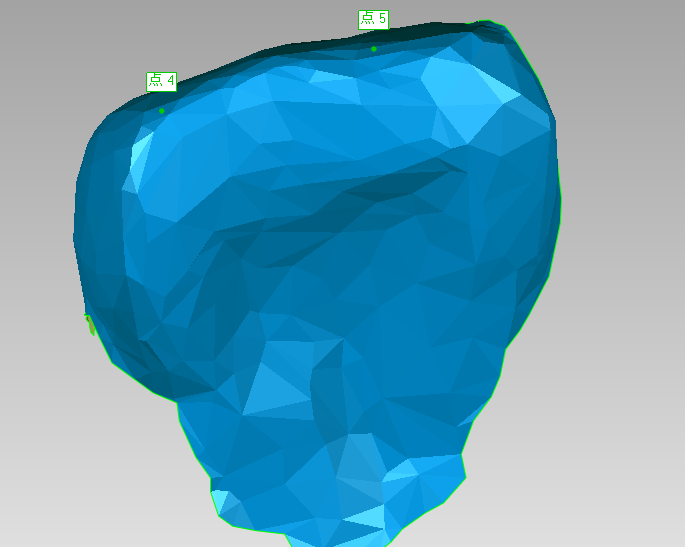


图 22 预测切牙的2个特征点位置

尖牙预测的三个特征点如图23所示，可以看到左侧两个点位于尖牙的切缘上，但是右侧的点预测位置偏离了切缘线。

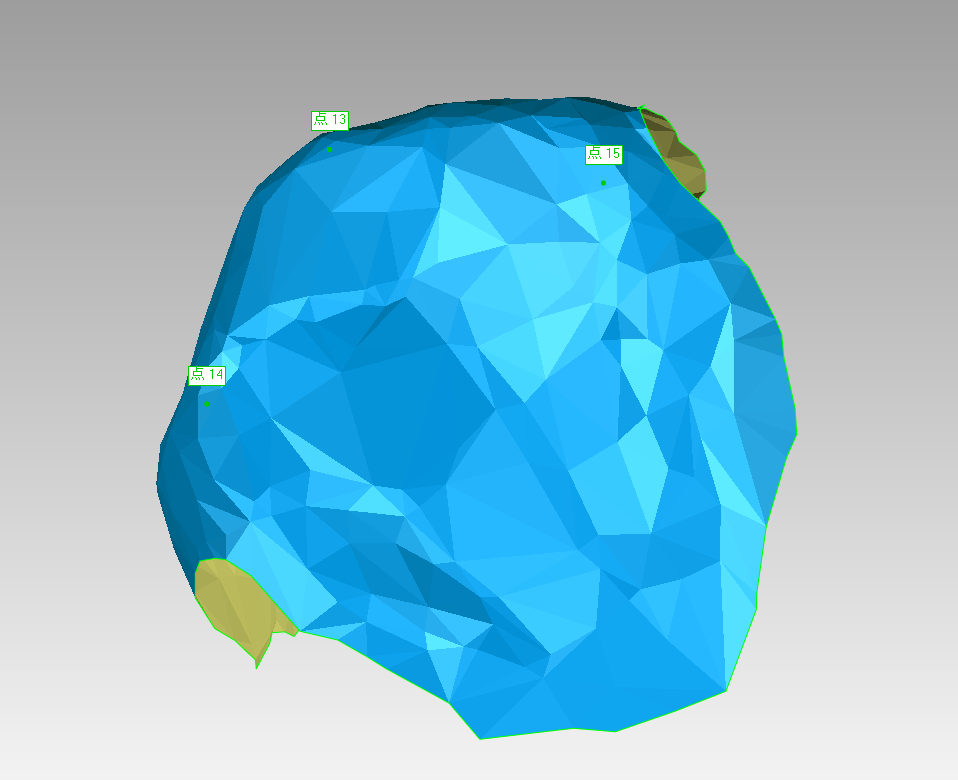


图 23 预测尖牙的3个特征点位置

前磨牙的预测特征点如图24所示，可以看到除了个别点（点16和点18）在预期位置上，其他预测特征点位置大大偏离预期位置。

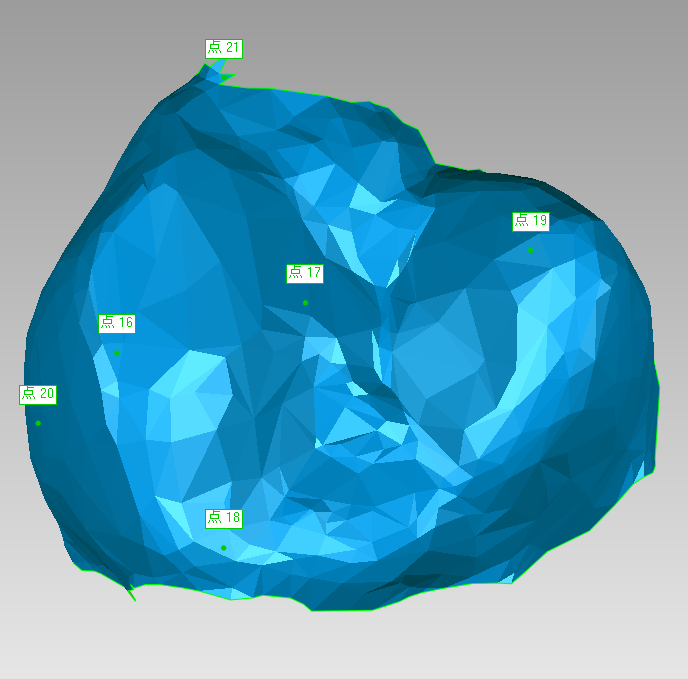


图 24 预测前磨牙的6个特征点位置

磨牙的预测特征点如图25所示，可以看到除了点29和点28在预期尖点，其余点偏离预期位置。

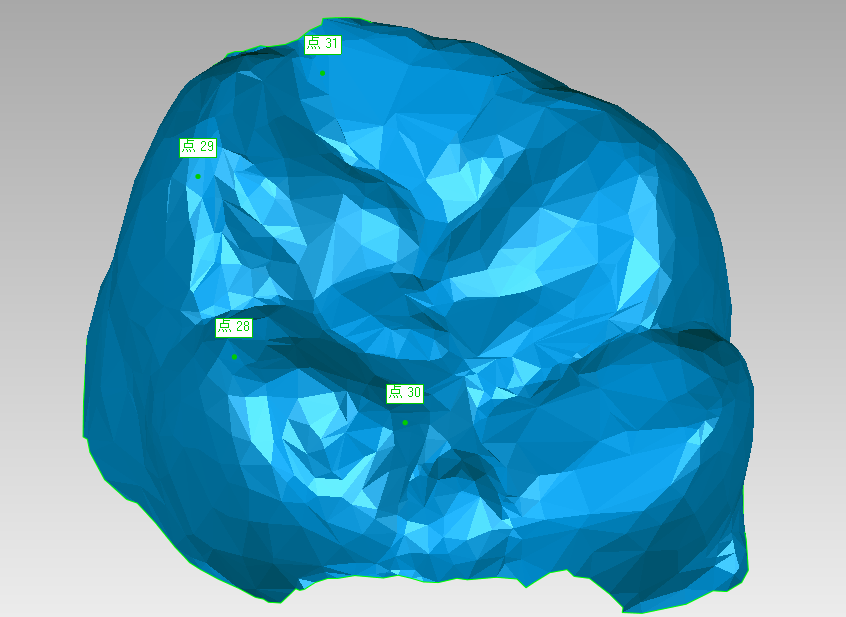


图 25 预测磨牙的4个特征点位置

后牙的识别精度低于前牙，可能是因为患者磨牙形状差异极大，网络难以学习，于是本文尝试只针对前牙（切牙和尖牙）的14类特征点单独训练基于改进的网络。

训练的曲线如图26所示，可以看到训练集收敛值比网络一次性训练54类特征点的值更小。

得到预测特征点结果可视化如下：

两个切牙的预测特征点如图27所示，可以看到预测特征点虽然基本位于牙齿切缘上，但是集中在右边一侧的特征点附近。

尖牙的预测特征点位置如图28所示，可以看到尖牙的预测效果远不如一次性预测54类特征点的效果好，三类特征点集中在了牙齿的一侧。

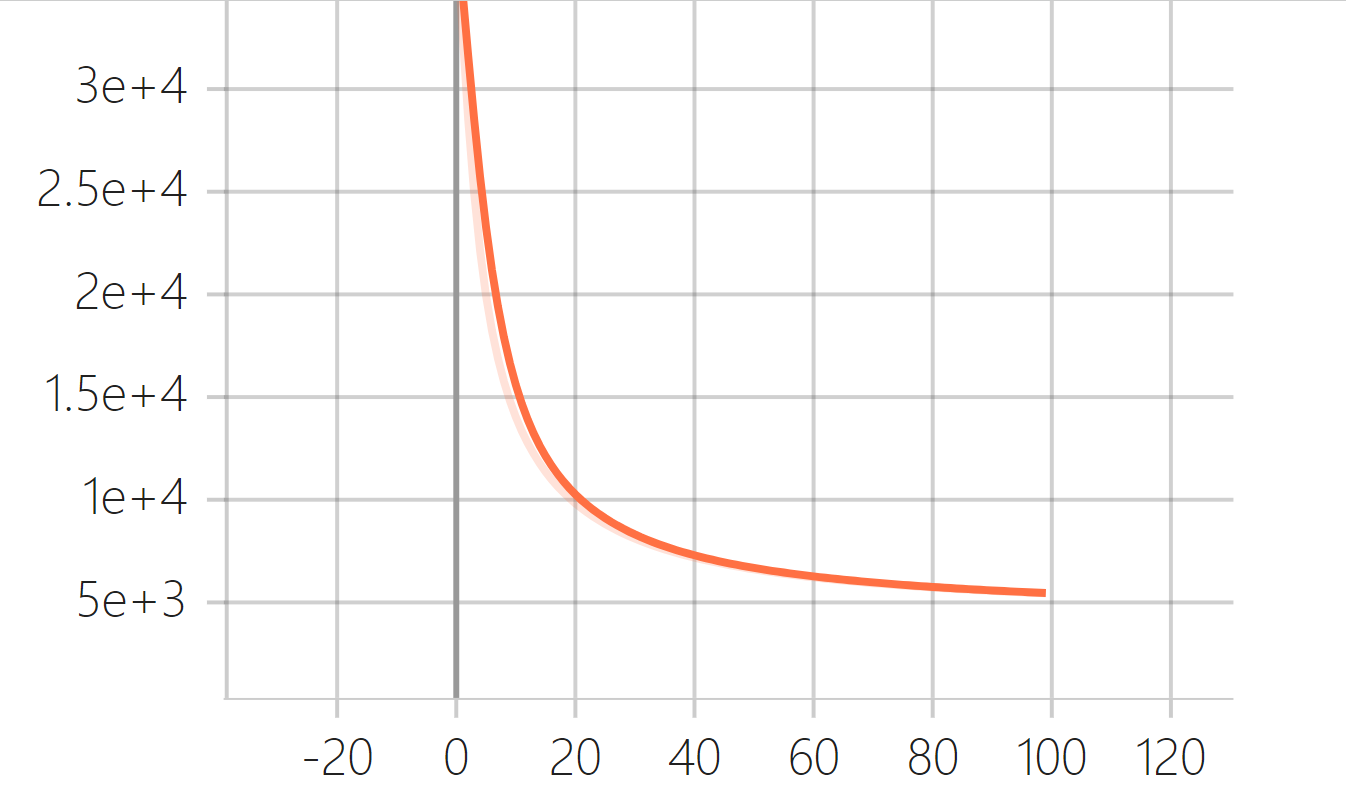


图 26 针对前牙训练的Loss曲线

只对前牙部分进行训练效果虽然基于高斯热图的网络回归Loss值更小，但是预测的特征点位置没有更准确，预测的特征点集中在了单侧位置。所以实际上使用一个完整牙列中54类特征点对于网络进行训练的效果更好。

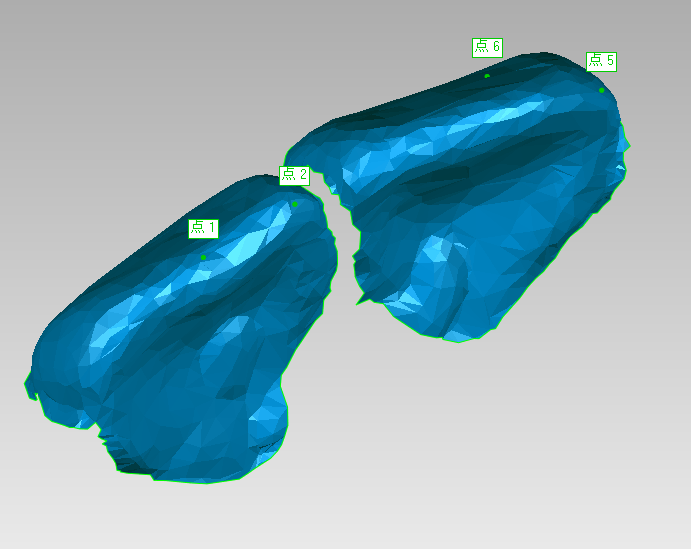


图 27 针对前牙训练的网络预测的切牙特征点位置

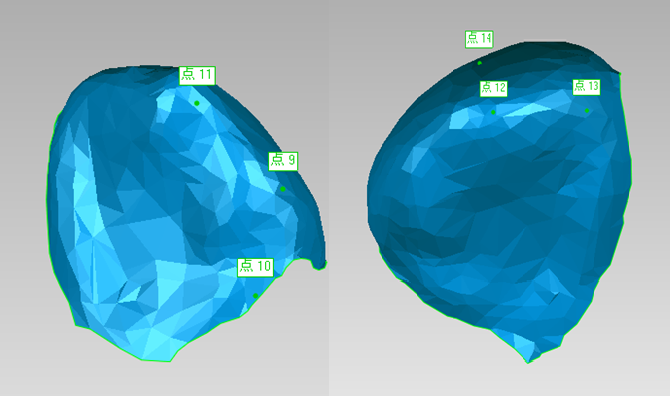


图 28 针对前牙训练的网络预测的尖牙特征点位置

**5.3 实验分析**

表1： 54类特征点的平均误差值

|  |  |  |  |
| --- | --- | --- | --- |
| 特征点序号 | 误差距离（mm） | 特征点序号 | 误差距离（mm） |
| 1 | 3.5462511638398095 | 28 | 4.133966063731832 |
| 2 | 5.280026409634219 | 29 | 4.565465709183772 |
| 3 | 4.672337113815887 | 30 | 3.662070551687025 |
| 4 | 3.025593627046208 | 31 | 3.403062825245622 |
| 5 | 2.7016634573041816 | 32 | 6.154360172684698 |
| 6 | 2.8279728764890657 | 33 | 3.8846865018737127 |
| 7 | 2.9816397872574156 | 34 | 1.8517270590793815 |
| 8 | 4.215131241461075 | 35 | 2.4363976915875902 |
| 9 | 4.428782414446258 | 36 | 2.9396118214768023 |
| 10 | 4.454852951160448 | 37 | 3.94556148341068 |
| 11 | 4.146720975515195 | 38 | 3.0873270231702645 |
| 12 | 2.135614019076339 | 39 | 3.1142640345077996 |
| 13 | 3.300956400870485 | 40 | 3.649554296985547 |
| 14 | 3.86617230605077 | 41 | 2.4023382540980402 |
| 15 | 5.060138441729216 | 42 | 2.3865896228943213 |
| 16 | 5.030954780005199 | 43 | 2.4723495215073075 |
| 17 | 3.8559765931572705 | 44 | 1.4301385078540265 |
| 18 | 3.5977376258204483 | 45 | 3.2784857475634435 |
| 19 | 2.804389657157919 | 46 | 4.36027871794118 |
| 20 | 4.072732303868509 | 47 | 2.7128009076287376 |
| 21 | 3.548392989255443 | 48 | 3.6563509974009274 |
| 22 | 5.473015377205685 | 49 | 5.391522923356908 |
| 23 | 7.919726054513945 | 50 | 7.451155203916879 |
| 24 | 3.241421846000343 | 51 | 2.8614897561103056 |
| 25 | 6.095514215208925 | 52 | 4.967370935106093 |
| 26 | 5.122925534197631 | 54 | 5.445482332166468 |
| 27 | 6.701170237093347 | 54 | 5.065971498086337 |
| 总平均误差（mm） | | 3.9781148 | |

上表给出了测试集中6个牙列54类特征点的平均误差（毫米为单位），即预测的特征点位置与真实特征点位置的平均欧式距离。从表1可以看出不同特征类别的精度不同，从最小的44号特征点1mm精度到最大的23号特征点6mm的精度，误差大小不一。表2根据牙齿类型对预测的特征点误差值进行计算。

表2：牙齿类型的平均误差

|  |  |
| --- | --- |
| 牙齿类别 | 平均误差（mm） |
| 切牙 | 4.0360967 |
| 尖牙 | 3.4003416 |
| 前磨牙 | 3.433347 |
| 磨牙 | 4.98294019 |
| 前牙 | 3.763163 |
| 后牙 | 4.208144 |

表2分别给出了切牙、尖牙、前磨牙和磨牙的平均误差，最后分析前牙和后牙的平均误差。可以看出前牙（切牙和尖牙）的预测效果比后牙（前磨牙和磨牙）的精确度高了1mm左右。前牙精度更高的原因有：

1. 后牙的特征点比前牙的特征点多。一个牙列中，前牙共有14类特征点，后牙共有40个特征点（具体可见2.3节牙齿特征点）。
2. 后牙的特征点真实值标记存在困难。由于数据集中病人的前磨牙和磨牙的牙齿差异很大，难以以一个统一的标准进行标注，标记出的同一类特征点之间仍然存在较大差异，从而难以得到理想的训练结果。

**5.4 实验环境**

表3：实验环境配置

|  |  |
| --- | --- |
| 环境 | 版本 |
| NVIDIA驱动 | 470.82.01 |
| CUDA | 1.4 |
| Python | 3.9.12 |
| Pytorch | 1.10.1 |

Epoch大小为100，网络在GPU上每次训练的时间大致在五分钟左右。可以看出该算法不需要耗费大量的GPU运算资源，网络模型能够快速收敛。

**5.4 小节**

本小节阐述了本文算法的实验过程，如实验相关参数的设置，可视化实验结果。接着对本算法在6个测试集病人上牙列共84颗牙齿的实验结果进行分析，从特征点类型和牙齿类型等多个角度对实验结果进行比较，分析实验结果背后的原因。

**6 总结与展望**

**6.1 本文总结**

随着机器学习的热烈蓬勃发展，其广泛应用在各领域，如搜索引擎、数据挖掘、自然语言处理，特别是计算机视觉领域。

本文旨在用机器学习的方法解决自动识别牙齿特征点问题，减少牙医在临床中需要花费大量精力标记病人牙齿特征点的问题。传统的数学方法识别牙齿特征点只局限在牙尖，难以应用于临床上。本文在三维牙齿模型视觉问题上已有前人使用机器学习的方法来分割标记牙齿，但在特征点识别问题上由于特征点种类多样不固定等因素，还未提出一种完善的牙齿特征点类型和牙齿特征点的自动识别算法。本文使用神经网络改进来识别不同的牙齿特征点。我对数据集合进行分类筛选，制定了一个牙列的54类特征点，帮助牙医能够快速确定牙前牙切缘和后牙的中央凹槽。根据这54个特征点类型，对30个牙列共420个牙齿进行标记。

本文提出的算法主要分为三个步骤：数据预处理、网络设计和数据后处理。本文指定网格数目，对牙齿网格STL文件做下采样操作，计算出其15维度的特征：网格三个顶点坐标、正则化法向量和网格中心坐标。特征点的真实值以IGES文件保存，本文根据IGES文件的格式读取出对应点的坐标和数目。

本文使用基于高斯热图回归的神经网络方法来实现牙齿特征点自动识别。在基于点云分类的PointNet神经网络架构进行改进，得到一个基于牙齿网格模型的网络。该网络使用高斯热图回归，而非传统点回归方法，预测输入的每一个牙齿的高斯热图。在算法后处理过程中，对高斯热图进行取最大值的网格中心位置作为每个牙齿对应特征点的位置，并加入了距离阈值的限制来避免预测特征点过于集中的实验现象。

**6.2 展望**

计算机医疗图像领域是一个较为小众的方向，可使用的数据集合较少，一般由于病人隐私问题已有研究使用的数据集都是保密，所以本文使用的数据集都是需要自己从头开始进行筛选和标记特征点，需要花费大量时间处理数据。本文共标记了30个上牙列，只有24个牙列作用于网络的训练，训练的数据较少从而导致网络学习结果不佳。病人牙齿存在异常，比如牙齿被磨平或者先天性的缺陷，牙齿难以以一个统一的标准度量其确切特征点，导致标记的特征点真实值不准确统一，增加了网络训练的难度。

实验结果中54类特征点的识别误差大小范围差异大，总体上看前牙的识别精度高于后牙精度。如何平均算法识别的不同牙齿特征点的误差也是一个问题。对于算法预测的牙齿特征点位置目前也没有做到快速批量的可视化方法，需要手动将预测的坐标点位输入到Geomagic Studio软件中，耗时耗力。若能够将实现结果批量导入Geomagic Studio软件中，或者使用其他方法可视化预测特征点，就可以更加快速方便地查看实验效果。

本文需要解决的问题是一次性预测一个牙列中54个特征点类型，这是一个相当困难的任务，加上数据集的缺乏，本文的实验精度还有很大提升空间。之后，可以尝试对不同类型的牙齿分类进行网络训练，从而减少实验误差距离。

总之，本文仍然存有许多地方可以继续改进提升，包括扩大数据集，提升实验精度和批量可视化预测特征点等。

**参考文献**

1. 马银玲．口腔固定义齿数字化设计技术研究：[硕士学位论文]。哈尔滨：哈尔滨工业大学图书馆，2014．
2. 马逸丹．数字图像处理技术在三维重建中的应用研究．数字技术与应用， 2016(10)：66．
3. 董园园，余小清，万旺根．基于虚拟现实技术的医学实验仿真教学系统．工 业控制计算机，2018，31卷(11期)：111-112，115．
4. 王兴. 我国口腔医疗行业的现状、机遇与挑战. 科技与金融, 2018: 10-13
5. Mangan A P， Whitaker R T．Partitioning 3D surface meshes using watershed segmentation．IEEE Transactions on Visualization and Computer Graphics，1999， 5(4)：308—321．
6. Kumar Y， Janardan R， Larson B．Automatic feature identification in dental meshes．Computer-Aided Design and Applications，2012，9(6)：747-769．
7. 仲哲．自动化排牙系统的研究与实现．硕士学位论文，杭州：浙江大学图书馆，2012．
8. Zhenqian Chen, Yongheng Shang, Andre Python, et al. DB-BlendMask: Decomposed Attention and Balanced BlendMask for Instance Segmentation of High-Resolution Remote Sensing Images. in IEEE Transactions on Geoscience and Remote Sensing, vol. 60, pp. 1-15, 2022
9. Q. Huang, W. Wang, and U. Neumann, “Recurrent slice networks for 3D segmentation of point clouds,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2018, pp. 2626–2635.
10. C. R. Qi, H. Su, M. Niebner, A. Dai, M. Yan, and L. J. Guibas, “Volumetric and multi-view CNNs for object classification on 3D data,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2016, pp. 5648–5656.
11. Charles R. Qi, Hao Su, Kaichun Mo, Leonidas J. Guibas. PointNet: Deep Learning on Point Sets for 3D Classification and Segmentation in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017: 652-660
12. C. Lian, S. Ruan, T. Denoeux, H. Li, and P. Vera, “Joint tumor segmentation in PET-CT images using co-clustering and fusion based on belief functions” IEEE Trans. Image Process., vol. 28, no. 2, pp. 755–766, Feb. 2019.
13. Kaiming He, Georgia Gkioxari, Piotr Dollar, Ross Girshick. Mask R-CNN in Proceedings of the IEEE International Conference on Computer Vision (ICCV), 2017, pp. 2961-2969
14. Georgia Gkioxari, Jitendra Malik, Justin Johnson. Mesh R-CNN in Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), 2019, pp. 9785-9795
15. Z. Wu et al .3D ShapeNets: A deep representation for volumetric shapes in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2015: 1912–1920.
16. D. Maturana and S. Scherer, VoxNet: A 3D convolutional neural network for real-time object recognition, in Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS), Sep. 2015: 922–928.
17. H. Su, S. Maji, E. Kalogerakis, and E. Learned-Miller, “Multi-view convolutional neural networks for 3D shape recognition,” in Proc. IEEE Int. Conf. Comput. Vis. (ICCV), Dec. 2015: 945–953.
18. C. R. Qi, H. Su, M. Niebner, A. Dai, M. Yan, and L. J. Guibas.Volumetric and multi-view CNNs for object classification on 3D data in Proc. IEEEInt. Conf. Comput. Vis. (ICCV), Dec. 2015: 945–953.
19. CHEN H, ZHANG K L, LYU P J, et al. A deep learning approach to automatic teeth detection and numbering based on object detection in dental periapical films［J］.Sci Rep, 2019, 9: 1-11.
20. C. H. Sudre, W. Li, T. Vercauteren, S. Ourselin, and M. Jorge Cardoso,“Generalised Dice Overlap as a Deep Learning Loss Function for Highly Unbalanced Segmentations BT - Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support,” M. J. Cardoso, T. Arbel, G. Carneiro, T. Syeda-Mahmood, J. M. R. S. Tavares, M. Moradi, A. Bradley, H. Greenspan, J. P. Papa, A. Madabhushi, J. C. Nascimento, J. S. Cardoso, V. Belagiannis, and Z. Lu, Eds. Cham: Springer International Publishing, 2017, pp. 240–248.
21. C. Lian et al., “MeshSNet: Deep multi-scale mesh feature learning for end-to-end tooth labeling on 3D dental surfaces,” in Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent. (MICCAI) (Lecture Notes in Computer Science), vol. 11769. Cham, Switzerland: Springer, 2019, pp. 837–845.
22. S. Tian, N. Dai, B. Zhang, F. Yuan, Q. Yu, and X. Cheng, “Automatic classification and segmentation of teeth on 3D dental model using hierarchical deep learning networks,” IEEE Access, vol. 7, pp. 84817–84828, 2019.
23. Chunfeng Lian, Li Wang, Tai-Hsien Wu,et al. Deep Multi-Scale Mesh Feature Learning for Automated Labeling of Raw Dental Surfaces From 3D Intraoral Scanners in IEEE TRANSACTIONS ON MEDICAL IMAGING, VOL. 39, NO. 7, JULY 2020.
24. X. Xu, C. Liu, and Y. Zheng, “3D Tooth Segmentation and Labeling Using Deep Convolutional Neural Networks,” IEEE Transactions on Visualization and Computer Graphics, vol. 25, no. 7, pp. 2336–2348, 2019.
25. S. Tian, N. Dai, B. Zhang, F. Yuan, Q. Yu, and X. Cheng, “Automatic Classification and Segmentation of Teeth on 3D Dental Model Using Hierarchical Deep Learning Networks,” IEEE Access, vol. 7, pp. 84 817–84 828, 2019.
26. F. Ghazvinian Zanjani, D. Anssari Moin, B. Verheij, F. Claessen, T. Cherici, T. Tan, and P. H. N. de With, “Deep Learning Approach to Semantic Segmentation in 3D Point Cloud Intra-oral Scans of Teeth,” in Proceedings of The 2nd International Conference on Medical Imaging with Deep Learning, ser. Proceedings of Machine Learning Research, M. J. Cardoso, A. Feragen, B. Glocker, E. Konukoglu, I. Oguz, G. Unal, and T. Vercauteren, Eds., vol. 102. London, United Kingdom: PMLR, 2019, pp. 557–571.
27. D. Sun, Y. Pei, G. Song, Y. Guo, G. Ma, T. Xu, and H. Zha, “Tooth Segmentation and Labeling from Digital Dental Casts,” in 2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI), 2020, pp. 669–673.
28. J. Zhang, C. Li, Q. Song, L. Gao, and Y.-K. Lai, “Automatic 3D tooth segmentation using convolutional neural networks in harmonic parameter space,” Graphical Models, vol. 109, p. 101071, 2020.
29. Charles Ruizhongtai Qi, Li Yi, Hao Su, Leonidas J. Guibas. PointNet++: Deep Hierarchical Feature Learning on Point Sets in a Metric Space, in Advances in Neural Information Processing Systems 30 (NIPS 2017)
30. Xinxin Chen, Kang Jiang, Yushi Zhu, et al. Individual Tree Crown Segmentation Directly from UAV-Borne LiDAR Data Using the PointNet of Deep Learning in Forests 2021, 12, 131-143
31. HE K, GEORGIA G, PIOTR D, et al. Mask R-CNN［J］.IEEE T Pattern Anal, 2020, 42(2): 386-397
32. REN S, HE K, GIRSHICK R, et al. Faster R-CNN: towards real-time object detection with region proposal networks［J］. IEEE Trans Pattern Anal, 2017, 39(6): 1137-1149.
33. 赵庶旭，罗庆，王小龙 .中国医学物理学杂志. 第38卷 第10期：1229-1236，2021年 10月
34. Y. Li, R. Bu, M. Sun, W. Wu, X. Di, and B. Chen, “PointCNN: Convolution On X-Transformed Points,”in Advances in Neural Information Processing Systems 31 (NeurIPS 2018), jan 2018.
35. Michael Garland, aul S. Heckbert .Surface Simplification Using Quadric Error Metrics in Proceedings of the 24th annual conference on Computer graphics and interactive techniques August 1997: 209–216
36. William J. SchroederJonathan A. ZargeWilliam E. Lorensen. Decimation of Triangle Meshes in ACM SIGGRAPH Computer Graphics, June 1997, 26(2):65-70
37. Nibali, Aiden, et al. "Numerical coordinate regression with convolutional neural networks." arXiv preprint arXiv:1801.07372 (2018).
38. C. Payer, D. Stern, H. Bischof, and M. Urschler, “Integrating spatial configuration into heatmap regression based CNNs for landmark localization,” Medical Image Analysis, vol. 54, pp. 207–219, 2019.
39. J. Zhang, M. Liu, L. Wang, S. Chen, P. Yuan, J. Li, S. G.-F. Shen, Z. Tang, K.-C. Chen, J. J. Xia, and D. Shen, “Context-guided fully convolutional networks for joint craniomaxillofacial bone segmentation and landmark digitization,” Medical Image Analysis, vol. 60, p. 101621, 2020.