

PECS-Net：一种融合多尺度局部 – 全局特征与双重注意力机制的断裂面识别算法

钟 玉^{1,2} 谢晓尧^{1,2} 刘 嵩^{1,2*} 刘建成³ 王 冲³ 刘旭斌^{1,2}
 ZHONG Yu XIE Xiaoyao LIU Song LIU Jiancheng WANG Chong LIU Xubin

摘 要

在人类历史的长河中，积累了许多珍贵的文物。随着时间的推移和各种因素的影响，这些文物不可避免地遭受了损坏。修复受损的文物不仅有助于保护文化遗产，还对弘扬和传承历史文化也起着重要作用。对此，文章提出了一种基于深度学习的模型 PECS-Net，用于识别文物残片的断裂面和原始表面。该模型以改进的 PointNet++ 网络为基础，修改原有的特征提取方法，提取多尺度的局部 - 全局特征，同时结合双重注意力机制增强特征表示。测试结果表明，PECS-Net 模型在 FragTag3D 数据集和 3D Puzzles 数据集上分别实现了 94.241% 和 76.859% 的断裂面识别准确率，平均交并比(MIoU)得分分别为 0.874 和 0.623。与 Baseline 相比，PECS-Net 在上述数据集上将断裂面识别准确率提高了 19.529% 和 15.022%，MIoU 得分分别提高了 0.277 和 0.201。

关键词

PointNet++；多尺度局部特征；全局特征；注意力机制；断裂面识别；碎片拼接

doi: 10.3969/j.issn.1672-9528.2025.09.029

0 引言

文物的保护与修复具有极为重要的意义。传统上，受损文物的修复依赖专业人员手工拼接碎片。这种方法耗时费力，严重依赖专家的经验，同时在修复过程中可能会对文物造成二次损害。

近年来，随着计算机技术的飞速发展，计算机视觉领域涌现出大量人才和创新技术，使得具备强大计算能力的计算机能够在碎片匹配与拼接任务中替代人工^[1-3]。在利用计算机修复破损文物的过程中，根据碎片厚度的特征，可将其分为非薄壁碎片与薄壁碎片两类^[4]。对于非薄壁碎片，可以从断裂表面提取特征，并利用这些特征进行点对点匹配，以完成碎片的拼接。因此，准确高效地区分断裂表面与原始表面，成为实现非薄壁碎片自动拼接的关键步骤之一。

为了克服这一挑战，本文提出了一种新的网络——PECS-Net，该网络能够识别非薄壁文物碎片的断裂表面，为文物的

- 贵州师范大学贵州省先进计算全省重点实验室
贵州贵阳 550001
- 贵州师范大学网络空间安全学院 贵州贵阳 550001
- 四川省文物考古研究院 四川成都 610093

- [18] OORD A V D, KALCHBRENNER N, KAVUKCUOGLU K. Pixel recurrent neural networks[C]//ICML'16: Proceedings of the 33rd International Conference on International Conference on Machine Learning. New York: ACM, 2016: 1747 - 1756.
- [19] OORD A V D, KALCHBRENNER N, VINYALS O, et al. Conditional image generation with PixelCNN decoders[EB/OL]. (2016-06-16)[2025-02-22]. <https://www.semanticscholar.org/paper/Conditional-Image-Generation-with-PixelCNN-Decoders-Oord-Kalchbrenner/0936352b78a52bc5d2b5e3f042>. DOI:10.48550/arXiv.1606.05328.
- [20] SALIMANS T, KARPATY A, CHEN X, et al. PixelCNN++: improving the PixelCNN with discretized logistic mixture

likelihood and other modifications[EB/OL]. (2017-01-19)[2025-03-22]. <https://doi.org/10.48550/arXiv.1701.05517>.

【作者简介】

潘俊安(1997—)，男，江西南昌人，硕士，教师、工程师，研究方向：图像处理、深度学习。

段润涛(1998—)，男，河南平顶山人，硕士，教师，研究方向：图像处理。

丁要男(1996—)，男，江西上饶人，硕士，教师，研究方向：图像处理、深度学习。

张煜(1996—)，女，江西南昌人，硕士，教师，研究方向：计算机视觉、深度学习。

(收稿日期：2025-03-24 修回日期：2025-09-03)

自动拼接与修复奠定基础。PECS-Net 基于经典的 PointNet++ 网络^[5], 并针对断裂表面与原始表面的区分任务进行了优化调整。在此基础上, 改进了原有的特征提取方法, 采用多尺度局部-全局特征提取策略, 并结合双重注意力机制^[6]以增强特征处理能力。此外, 优化了损失函数, 以进一步提升模型的表达能力。该网络在公共 FragTag3D 数据集^[7]和维也纳工业大学的 3D Puzzles 数据集^[8]上, 分别实现了 94.241% 和 76.859% 的断裂表面识别准确率, MIoU (平均交并比) 得分分别为 0.874 和 0.623。

1 国内外研究现状

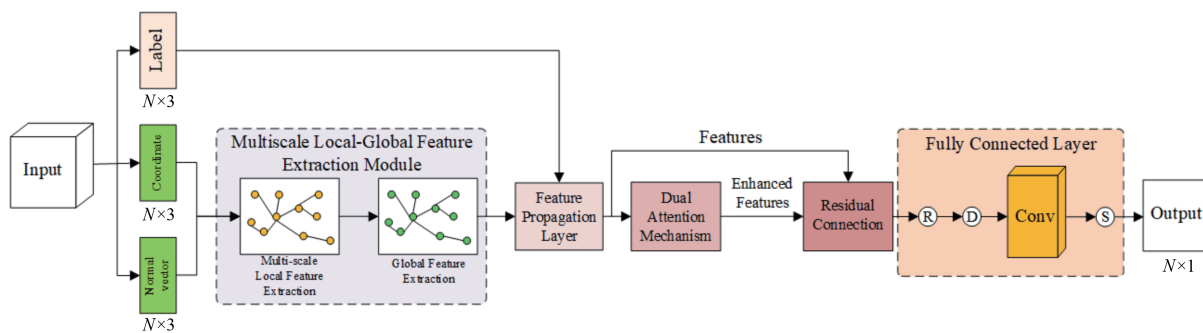
在传统的文物碎片匹配流程中, 通常遵循以下步骤: 首先, 应用断裂表面识别算法区分断裂表面与原始表面; 随后, 利用特征提取算法, 从断裂表面中提取关键特征; 最后, 基于提取的特征信息执行粗匹配与精细匹配。Papaioannou 等人^[9]采用区域生长算法对碎片表面进行分割, 进一步分离出断裂表面, 并提取每个碎片的封闭边界曲线。随后, 利用曲线拟合技术结合表面匹配算法, 实现 3D 碎片的拼接。Huang 等人^[8]提出了一种能够处理粗糙断裂表面的方法。使用多尺度边缘提取算法生成初始的表面分割结果, 并采用图割 (graph-cut) 算法将这些表面划分为断裂表面和原始表面。最终, 他们在断裂表面上计算积分不变量, 以此作为碎片匹配和拼接的特征描述符。Yi 等人^[10]提出了一种自动碎片匹配算法, 该算法依赖于从碎片中提取几何特征和外观特征进行匹配。这一方法的前提是必须准确识别断裂表面。Chen 等人^[11]通过分析表面粗糙度来识别断裂表面, 并对断裂表面上的坑洼进行聚类, 以形成凹凸区域。首先基于碎片的边界曲线进行粗匹配, 随后利用改进的迭代最近点 (ICP) 算法, 在凹凸表面基础上执行精细匹配。Palmas 等人^[12]通过检测封闭边缘轮廓提取断裂曲线, 并根据这些曲线分割的区域进行碎片匹配。尽管该方法未明确提及断裂表面, 但匹配过程中使用的断裂曲线分割区域, 实际上对应于断裂表面和原始表面。

随着硬件的快速发展, 尤其是 GPU 内存容量和计算速度的飞跃式提升, 深度学习方法在计算机视觉领域得到了广泛应用, 并在实际应用中展现了其高效性和优越性^[13-14]。Yao 等人^[15]采用 Chen 等人^[11]提出的断裂表面分割方法, 从断裂表面提取特征并生成高维特征向量, 然后将这些向量输入 SiamesePointNet 进行相似度计算, 并基于计算结果进行碎片拼接。Lu 等人^[16]以 PointNet++ 作为骨干网络, 并引入自注意力和交叉注意力机制, 从断裂表面提取特征后输入匹配网络。尽管该算法未明确区分断裂表面与原始表面, 但在提取特征描述符的过程中完成了断裂表面识别。刘旭斌等人^[17]同样以 PointNet++ 作为骨干网络, 并基于该网络提出了一种断裂表面识别算法。这些基于深度学习的拼接算法, 无论是显式还是隐式地进行断裂表面识别, 均需要从断裂表面提取特征描述符, 并将其作为输入数据, 以完成碎片拼接。

2 理论与方法

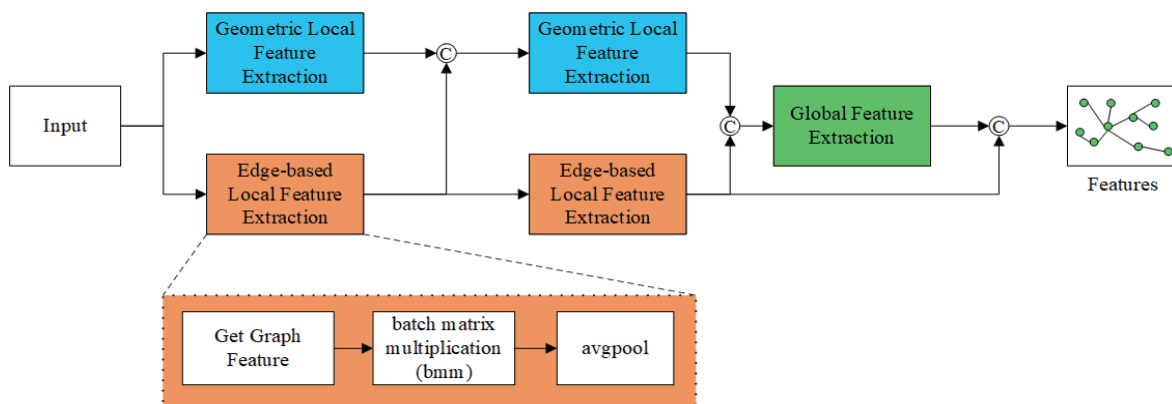
受刘旭斌等人的启发, 本文将复杂的断裂表面识别任务视为一个语义分割问题, 并提出了 PECS-Net (如图 1 所示)。该模型以改进版的 PointNet++ 作为骨干框架, 对原有的特征提取方法进行了优化。该模型在公开的 FragTag3D 数据集上进行训练、验证和测试, 同时在维也纳工业大学提供的 3D Puzzles 数据集上进行泛化测试。

在 PECS-Net 中, 输入的点云数据包含 $N \times 3$ 的坐标信息和 $N \times 3$ 的法向量信息, 其中 N 为输入点云中的点数。这些原始特征经过几何特征提取和边缘特征提取, 生成多尺度局部特征 F_{mul} , 进一步提取全局特征 F_{global} ; F_{mul} 与 F_{global} 结合, 最终形成新的特征信息 F , 即多尺度局部-全局特征。然后, F 进入双重注意力模块进行进一步处理, 生成增强特征 E_F 。接着, F 和 E_F 通过残差连接 (residual connection)^[18] 相连。最后, 通过全连接层输出一个 $N \times 1$ 矩阵, 该矩阵用于判断输入点云数据中的每个点是否属于断裂表面。



注: R 表示 ReLU 激活函数; D 表示 Dropout 操作, 用于防止模型过拟合, 从而提高模型的泛化能力; S 表示 Sigmoid 激活函数。

图 1 PECS-Net 结构图



注：C 表示矩阵拼接（matrix concatenation）。

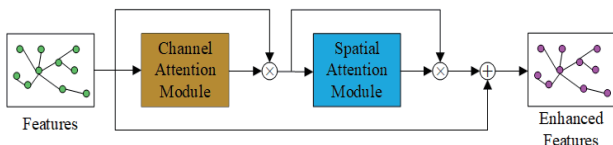
图2 多尺度局部-全局特征提取模块结构图

2.1 多尺度局部-全局特征（MS-LG）

由于 PointNet++ 采用 PointNet^[19] 作为其基本特征提取单元，而 PointNet 主要侧重于几何信息，因此在某些碎片的识别性能上可能存在不足。受 Wang 等人^[20] 的启发，将原始几何特征与边缘特征相结合，形成多尺度局部特征，并将其与全局特征融合，以获取多尺度局部-全局特征。在多尺度局部-全局特征提取模块中，分别从输入点云数据中提取边缘局部特征和几何局部特征，得到边缘局部特征 F_E 和几何局部特征 F_G 。随后，将 F_E 和 F_G 结合，形成多尺度局部特征 F_{mul} 。接着，对 F_{mul} 进行全局池化（global pooling），以获得全局特征 F_{global} 。最后，融合 F_E 和 F_{global} ，得到多尺度全局特征 F 。多尺度局部-全局特征提取模块结构如图2所示。

2.2 双重注意力机制（DA）

Woo 等人^[21] 提出了一种用于二维图像的（convolutional block attention module, CBAM），受此启发，为 PECS-Net 引入双重注意力机制。双重注意力机制模块如图3所示。



注：× 表示矩阵乘法；+ 表示矩阵元素按位相加。

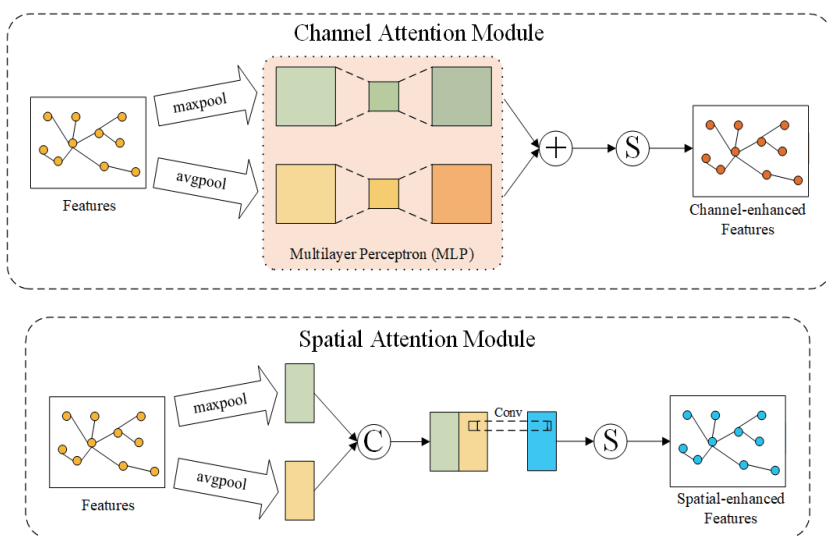
图3 双重注意力机制模块结构图

假设输入数据为 F ，则输出数据 E_F 用公式表示为：

$$E_F = S(C(F) \times F) \times (C(F) \times F) \quad (1)$$

式中： $C(F)$ 表示对 F 进行通道增强； $S(F)$ 表示对 F 进行空间增强。

通道注意力模块与空间注意力模块如图4所示。



注：+ 表示矩阵元素按位相加；C 表示矩阵拼接；S 表示使用 Sigmoid 函数进行激活。

图4 通道注意力模块与空间注意力模块结构图

2.3 特征变换正则化损失函数（FTR）

向 PECS-Net 中引入特征变换正则化器确保变换矩阵尽可能接近正交矩阵，从而保存数据的几何结构特征。同时，正则化项通过增加一个小的惩罚，使得模型在训练过程中不会过于依赖某些特定的特征变换。本文引入了特征变换正则化器，用公式表示为：

$$L_{trans} = \frac{1}{B} \sum_{i=1}^B \|I - T_i^T T_i\|_F^2 \quad (2)$$

式中： B 表示批量大小； T_i 表示第 i 个样本的特征变换矩阵； I 表示单位矩阵； $\|\cdots\|_F$ 表示弗罗贝尼乌斯范数，用公式表示为：

$$\|A\|_F = \sqrt{\sum_{i=1}^m \sum_{j=1}^n |a_{ij}|^2} \quad (3)$$

式中： A 表示一个 $m \times n$ 的矩阵； a_{ij} 表示矩阵 A 的第 i 行第 j 列的元素； $\|A\|_F$ 表示矩阵元素的平方和的平方根。

3 实验

本文的实验部署在一台搭载 AMD2700X 处理器、64 GB 内存、8 GB NVIDIA GTX 1080 显卡的电脑上，其软件方面的环境为：Ubuntu 20.04、python 3.7.16、torch 1.13.1、CUDA 11.1、open3d 0.17.0、NumPy 1.21.6 等。

3.1 数据集

本文采用公开数据集 FragTag3D 进行训练，该数据集中的数据基于 Thingy10k^[22]。该数据集以 ASCII 码形式的 PLY 文件格式存储，其中包含了每个点的坐标信息、法向量信息以及颜色信息。PLY 文件格式其具备良好的可读性。然而，该结构不适合直接进行数值计算和矩阵操作。在深度学习中，通常将数据存储为 NPY 文件格式。NPY 文件是 NumPy 的专有格式，旨在存储多维数组。

3.2 实验过程

本研究按照 7:1.5:1.5 的比例将数据划分为训练集、验证集和测试集，用来进行网络训练、调整网络参数。Batch size 设置为 8，数据下采样的数目 npoints 为 2 048，使用 Adam 优化器来加速和稳定神经网络的训练过程。训练的 epoch 设置为 200，初始学习率为 0.001，并在训练过程中对学习率进行调整。

在验证阶段中，本研究统计验证过程中预测值与真值相同的点的数量，记预测值与真值相等的点为 r_i ，其总数量为 sum_r ，参与训练的点为 p_j ，总数量为 sum_p ， $\text{accuracy} = \text{sum}_r / \text{sum}_p$ 。最后根据 accuracy 对网络训练效果进行数值化的评估。同时，本文通过语义分割的方式进行断裂面与原始面的识别工作，计算出模型的均交并比 (MIoU)，用公式表示为：

$$\text{MIoU} = \frac{1}{N} \sum_{i=1}^N \frac{A_i \cap B_i}{A_i \cup B_i} \quad (4)$$

式中： A_i 表示模型预测出的第 i 类的区域； B_i 表示实际标注的第 i 类的区域。

PECS-Net 在数据集 FragTag3D 上的实验过程中各指标如图 5 所示，其中平滑度为 0.6。图 (a) 展示模型在验证集上的准确率随着验证步数的增加而逐渐提升。可以看到随着训练的进行，验证准确率不断提升，最终达到最高值 96.026%；图 (b) 展示了模型在训练过程中 Best MIoU 的变化情况。随着训练 epoch 的增加，Best MIoU 也逐步提升，最高达到 0.838；图 (c) 显式模型在训练集上的准确率随着

训练步数的增加而上升。可以观察到，训练集准确率在不断提高，达到的最高值为 94.149%；图 (d) 显式模型在训练过程中的损失函数值随训练步数的增加而逐渐减少，并最终收敛于 0.13。

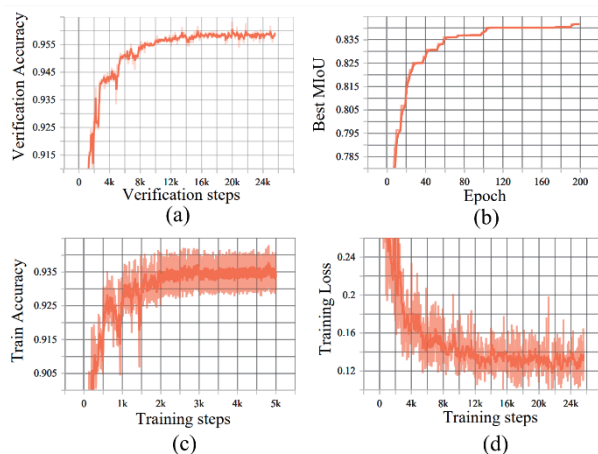


图 5 训练过程指标

3.3 实验结果

在测试阶段使用的 batch size 值为 1，数据下采样的目标值 npoints 为 15 000。本文提出的模型在 FragTag3D 数据集上选取 252 个未参与训练和验证的数据进行测试，得到的断裂面识别准确率为 94.241%，MIoU 为 0.874。识别效果如图 6 所示。本文从 47 个碎片集中选取了 3 个具有代表性的碎片集进行展示，可以看出本研究提出的网络模型成功地识别出点云数据中的断裂面，展现出良好的性能。

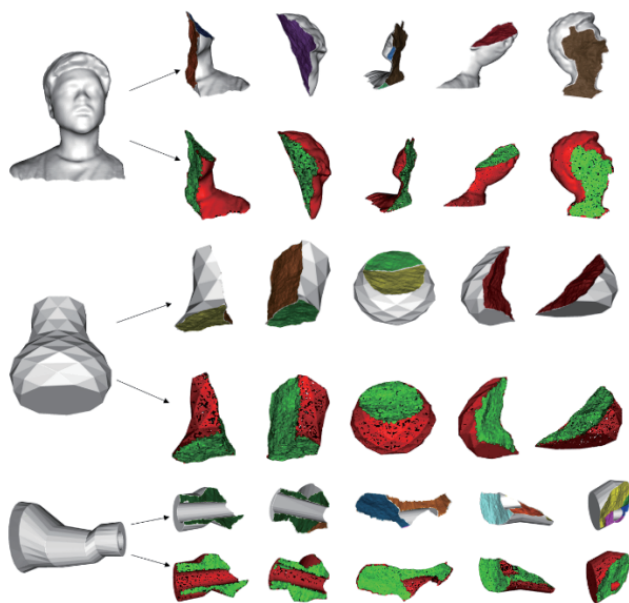


图 6 PECS-Net 识别效果图

3.4 泛化实验

本研究在维也纳工业大学公开数据集 3DPuzzle 上对本

文提出的模型进行泛化性测试。本文提出的模型在 3DPuzzle 碎片集上进行断裂面识别的实验，并进行效果展示，如图 7 所示。

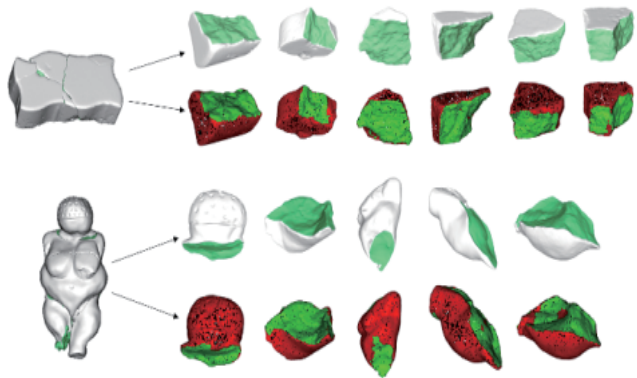


图 7 泛化效果图

图 7 展示了 PECS-Net 对断裂面和原始面识别效果与人工标注断裂面的对比图。本文对维也纳工业大学公开数据集进行识别准确率计算，其结果如表 1 所示。

表 1 各碎片集泛化测试准确率

碎片集	Accuracy/%	MIoU
Brick	80.540	0.671
Cake	81.374	0.681
Gargoyle	78.358	0.642
Sculpture	71.700	0.562
Venus	75.586	0.602
Overall	76.859	0.623

3.5 消融实验

本文进行了消融实验，以评估对 PointNet++ 网络进行修改的有效性。本研究从 FragTag3D 中选取 252 个未参与训练和验证的数据，从 3DPuzzle 中选取 69 个数据，分别对上述模块进行测试。其结果如表 2~3 所示。

表 2 FragTag3D 上的结果

Baseline	MS-LG	DA	FTR	Accuracy/%	MIoU
√				74.712	0.597
√	√			84.747	0.725
√		√		86.229	0.736
√			√	79.458	0.652
√	√	√		94.521	0.879
√	√		√	91.603	0.827
√		√	√	88.190	0.769
√	√	√	√	94.241	0.874

表 3 3DPuzzle 上的结果

Baseline	MS-LG	DA	FTR	Accuracy/%	MIoU
√				61.837	0.422
√	√			69.709	0.543
√		√		71.335	0.546
√			√	67.152	0.485
√	√	√		74.306	0.590
√	√		√	71.926	0.548
√		√	√	72.474	0.555
√	√	√	√	76.859	0.623

4 结论

在本研究中提出了一种改进的网络模型 PECS-Net 以提升识别文物碎片断裂面和原始面的准确率，该网络在数据集 FragTag3D 上对断裂面的识别准确率为 94.241%，MIoU 为 0.874，在维也纳工业大学公开数据集 3DPuzzle 上对断裂面的识别准确率为 76.859%，MIoU 为 0.623。PECS-Net 基于 PointNet++ 的网络架构，同时提取出几何局部特征和边缘局部特征，结合成多尺度局部特征，并与全局特征融合得到多尺度局部 - 全局特征，再使用双重注意力机制对特征进行动态调整，优化了原本的损失函数。通过这些改进方法，本文得到一个能够更好地提取碎片特征的新模型，从而更准确地识别 3D 点云数据中的断裂面和原始面。在数据集 FragTag3D 上的测试结果表明，本文提出的方法在准确识别断裂面和原始面方面表现出色，同时在维也纳工业大学的数据集 3DPuzzle 上的测试结果也证明本研究提出的方法在不同类别的对象上同样有效。

致谢：

感谢维也纳科技大学为本研究提供 3D 拼图数据集集中使用的 3D 模型。

参考文献：

[1] CHETVERIKOV D, SVIRKO D, STEPANOV D, et al. The trimmed iterative closest point algorithm[C]// 2002 International Conference on Pattern Recognition.Piscataway: IEEE, 2002: 545-548.

[2] ZINBER T, SCHMIDT J, NIEMANN H, et al. Point set registration with integrated scale estimation[C]//International conference on pattern recognition and image processing. Piscataway: IEEE, 2005:116-119.

- [3] WILLIS A R, COOPER D B. Computational reconstruction of ancient artifacts[J]. IEEE signal processing magazine, 2008, 25(4): 65-83.
- [4] YANG W, ZHOU M Q, ZHANG P F, et al. Matching method of cultural relic fragments constrained by thickness and contour feature[J]. IEEE access, 2020, 8: 25892-25904.
- [5] QI C R, YI L, SU H, et al. PointNet++: deep hierarchical feature learning on point sets in a metric space[C]//NIPS'17: Proceedings of the 31st International Conference on Neural Information Processing Systems. New York: ACM, 2017: 5105-5114.
- [6] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need[J]. NIPS'17: Proceedings of the 31st International Conference on Neural Information Processing Systems. New York: ACM, 2017: 6000-6010.
- [7] SU H T. FragTag3D: a 3D object fragment dataset with annotations for overlapping regions [EB/OL]. [2025-05-16]. <https://bailanciaga.github.io/FragTag3D/>.
- [8] HUANG Q X, FLÖRY S, GELFAND N, et al. Reassembling fractured objects by geometric matching[J]. ACM transactions on graphics (TOG), 2006, 25(3): 569-578.
- [9] PAPAIOANNOU G, KARABASSI E A. On the automatic assemblage of arbitrary broken solid artefacts[J]. Image and vision computing, 2003, 21(5): 401-412.
- [10] YI R B, LI J L, LUO L, et al. DOPNet: achieving accurate and efficient point cloud registration based on deep learning and multi-level features[J]. Sensors, 2022, 22(21): 8217.
- [11] CHEN Y C, LI H D, TURPIN D, et al. Neural shape matting: self-supervised object assembly with adversarial shape priors[EB/OL]. (2022-05-30)[2025-05-25]. <https://doi.org/10.48550/arXiv.2205.14886>.
- [12] PALMAS G, PIETRONI N, CIGNONI P, et al. A computer-assisted constraint-based system for assembling fragmented objects[C]//2013 Digital Heritage International Congress (DigitalHeritage). Piscataway: IEEE, 2013: 529-536.
- [13] LI Q H, GENG G H, ZHOU M Q. Pairwise matching for 3D fragment reassembly based on boundary curves and concave-convex patches[J]. IEEE access, 2019, 8: 6153-6161.
- [14] ALAGRAMI A, PALMIERI L, ASLAN S, et al. Reassembling broken objects using breaking curves[C]//Pattern recognition. Berlin: Springer, 2024: 197-208.
- [15] YAO W M, CHU T, TANG W L, et al. SPPD: a novel reassembly method for 3D terracotta warrior fragments based on fracture surface information[J]. ISPRS international journal of geo-Information, 2021, 10(8): 525.
- [16] LU J X, SUN Y F, HUANG Q X. Jigsaw: learning to assemble multiple fractured objects[EB/OL]. (2023-10-27)[2025-02-22]. <https://doi.org/10.48550/arXiv.2305.17975>.
- [17] 刘旭斌, 刘嵩, 刘建成, 等. 基于 Point Net++ 的碎片断裂面提取算法 [J]. 自动化应用, 2024, 65(6): 129-132.
- [18] HE K M, ZHANG X Y, REN S Q, et al. Deep residual learning for image recognition[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE, 2016: 770-778.
- [19] QI C R, SU H, MO K C, et al. PointNet: deep learning on point sets for 3D classification and segmentation[EB/OL]. (2017-04-10)[2025-05-13]. <https://doi.org/10.48550/arXiv.1612.00593>.
- [20] WANG Y, SUN Y B, LIU Z W, et al. Dynamic graph CNN for learning on point clouds[J]. ACM transactions on graphics (ToG), 2019, 38(5): 1-12.
- [21] WOO S, PARK J, LEE J Y, et al. CBAM: convolutional block attention module[C]//Computer Vision-ECCV 2018. Berlin: Springer, 2018: 3-19.
- [22] SELLÁN S, CHEN Y C, WU Z Y, et al. Breaking bad: a dataset for geometric fracture and reassembly[EB/OL]. (2022-10-20)[2025-03-16]. <https://doi.org/10.48550/arXiv.2210.11463>.

【作者简介】

钟玉（1999—），男，贵州安顺人，硕士研究生，研究方向：计算机视觉、碎片拼接，email: 232100210687@gznu.edu.cn。

谢晓尧（1952—），男，贵州贵阳人，博士，教授，研究方向：网络通信、信息安全与人工智能，email: xyx@gznu.edu.cn。

刘嵩（1983—），通信作者（email: songliu@gznu.edu.cn），男，贵州贵阳人，博士，副教授，研究方向：计算机视觉、三维重建。

刘建成（1990—），男，四川成都人，硕士，馆员，研究方向：文物保护，email: liujc90@163.com。

王冲（1981—），男，四川广安人，硕士，副研究馆员，研究方向：文物保护，email: 67692034@qq.com。

刘旭斌（1997—），男，广东韶关人，硕士研究生，研究方向：计算机视觉、三维重建，email: 21010210660@gznu.edu.cn。

（收稿日期：2025-03-26 修回日期：2025-09-04）