

doi: 10.3969/j.issn.1000-8349.2024.02.08

深度学习语义分割方法用于双致密天体 并合引力波数据处理研究

马存良¹, 钟国健¹, 闵源¹, 嘉明珍¹, 贺观圣^{2,3}

(1. 江西理工大学, 赣州 341000; 2. 南华大学, 衡阳 421001; 3. 中国科学院紫金山天文台, 南京 210023)

摘要: 研究表明, 基于深度学习的引力波搜寻, 有望解决匹配滤波方法的低效率问题。与传统的匹配滤波引力波信号搜寻、基于贝叶斯后验或深度学习的波源参数估计方法相比, 基于深度学习的端对端的引力波搜寻输出缺乏波源的时频信息, 如双星并合信号到达时间和信号在探测站灵敏度区间的持续时间等。首次实现了深度学习用于应变信号时频点的细粒度分类, 研究了基于图像语义分割的引力波搜寻方法。通过真实噪声和仿真物理信号合成数据, 使用 Q 变换将其转换为时频图, 构建像素级标注数据集。构建并训练时频图像语义分割模型, 研究了其应用于引力波信号数据分析的可行性。结果表明语义分割方法未来有望作为基于深度学习的引力波信号搜寻流水线的一个组件, 还可用于进一步提取信号的时频信息。

关键词: 引力波数据分析; 深度学习; 语义分割

中图分类号: P129 **文献标识码:** A

1 引言

对引力波的直接观测分析, 有益于物理理论的验证和完善, 及多信使天文学和宇宙学的发展。先进激光干涉引力波天文台 (advanced Laser Interferometer Gravitational wave Observatory, aLIGO) 所代表的第二代地基引力波探测器在前 3 次观测运行中获得 90 多个可信引力波源^[1, 2]。这些数据已应用于广义相对论预言的验证^[3]、引力波速度的测定以及宇宙学常数的测定^[4]等; 并且初步应用于宇宙弦^[5]、连续引力波^[6]和随机引力波背景^[7]等的证据搜寻。

收稿日期: 2023-10-30; 修回日期: 2024-04-02

资助项目: 国家自然科学基金 (12205139); 江西省自然科学基金 (20224BAB1012); 湖南省自然科学基金 (2022JJ40347)

通讯作者: 嘉明珍, mingzhenj@jxust.edu.cn

未来第三代地基引力波探测器（如爱因斯坦望远镜^[8]）和空间引力波探测器（如激光干涉空间天线^[9]、天琴^[10]和太极^[11]等）的并网探测将使得引力波可观测天体源在数量和种类上极大提高。探测器数量和灵敏度的提高使得引力波搜寻数据处理方法面临新的挑战^[12]。传统的基于模板库的方法效率低，深度学习能够解决引力波信号搜寻的效率问题。2018年Gebhard等人^[13]与George和Huerta^[14, 15]几乎同时得到基于深度学习的引力波搜寻，在精度上与匹配滤波相当。研究表明，卷积神经网络^[16-19]、空洞卷积神经网络^[20, 21]、多深度网络的集成^[22, 23]、循环神经网络^[24]等多种深度学习网络结构可以应用于引力波信号搜寻。许多用于引力波搜寻工作的深度学习对数据预处理方面进行了研究，如加入匹配感知层^[25]，在小波变换后去除高频分量^[26]等。近年来，除了地基引力波搜寻之外，深度学习已经对引力波数据处理的多个层面进行了研究，包括引力波的去噪^[27, 29]、波源参数估计^[30-32]、空基引力波信号处理^[33-35]、脉冲星计时阵列引力波信号处理^[36]、分赫兹频段引力波信号处理^[37]等。

与白化的一维信号相比，二维时频信号更易于表示非平稳信号的特征。可以采用小波变换^[18]或Q变换^[38]将一维信号变换为二维时频图像。其与计算机视觉的数据表示一致，因此许多工作将计算机视觉相关算法迁移到引力波信号处理中。2021年，Menendez-Vazquez等人^[38]将计算机视觉中的图像分类算法应用于引力波搜寻。2022年，Aveiro等人^[39]首次将计算机视觉中的目标检测算法应用于双中子星信号的搜寻。与图像分类相比，目标检测不但预测了引力波的存在，而且预测了引力波存在的时间范围。2023年，Jin等人^[40]通过对时频图进行分析，开发了一种方法来确定双黑洞并合是否会释放引力波信号。

受文献[39]的启发，我们首次提出基于图像语义分割的引力波搜寻方法。优于文献[39]的是，语义分割能够实现像素级分类，本文所提出的方法的输出包含丰富的频率和时间信息，不但能够预测引力波的并合时间，而且能够预测信号在探测器灵敏频率区间的瞬时频率范围。本方法有望融入深度学习引力波信号分析流水线中，图像分割模型的输出具有对原信号进行自适应滤波的潜在应用；可以根据图像分割的输出，构建时间自适应带通滤波器，使得不同的时间片通过不同的带通滤波器滤波；有望构建算法将分割结果作为滤波掩码，滤除噪声频率成分，保留信号频率成分。

图像的语义分割已成为自动驾驶^[41]、医学成像诊断^[42]和遥感图像智能分析^[43]等应用的基本感知组件。与目标检测任务不同，语义分割不仅需要标识图像中存在的对象，还需要为每个像素赋予一个语义标签。其可以帮助计算机实现对图像中对象的更精细和准确的理解。

阈值分割法、边缘检测法和模糊聚类方法等传统分割方法存在分割不连续或欠分割的问题。近年来，深度学习的兴起，使得基于数据驱动的图像分割方法得到快速发展。卷积神经网络^[44]具有提取图像特征的能力，已被广泛用于不同的语义分割任务。卷积层的有限感受野，使其性能受到限制。金字塔池化模块(pyramid pooling module, PPM^[45])、可变形卷积^[46]、扩张卷积^[47-49]等方法可以用来改善语义分割中的局部感受野问题。最近的发展使得基于深度学习的图像分割可以捕获更精确的全局^[50-52]和更好的边界信息^[53, 54]。

本文对深度学习用于引力波信号时频图的语义分割进行了研究。第2章对数据分析的框架进行了介绍；第3章详细论述了系统的结构、数据的生成方法、语义分割模型和模型

的训练方法; 第 4 章介绍了 5 种语义分割模型在测试集和真实事件的测试结果; 第 5 章为论文的总结。

2 数据分析框架

本文首次探索使用语义分割的方法提取引力波的时频特征, 所提出方法的框架如图 1 所示。首先通过 SEOBNRv4 方法仿真信号的波形, 随机截取真实背景噪声并对其通过 Welch 方法进行功率谱 (power spectral density, PSD) 计算。然后通过 PSD 计算噪声幅度谱 (amplitude spectral density, ASD), 使用 ASD 白化“信号”和“噪声+信号”。将白化的“信号”和“噪声+信号”分别转换为时频图, 并利用“信号”的时频图构建标签。接下来, 建立语义分割模型, 使用“噪声+信号”的时频图作为输入, 训练该模型, 最后通过测试集与真实事件数据集来评估模型性能。

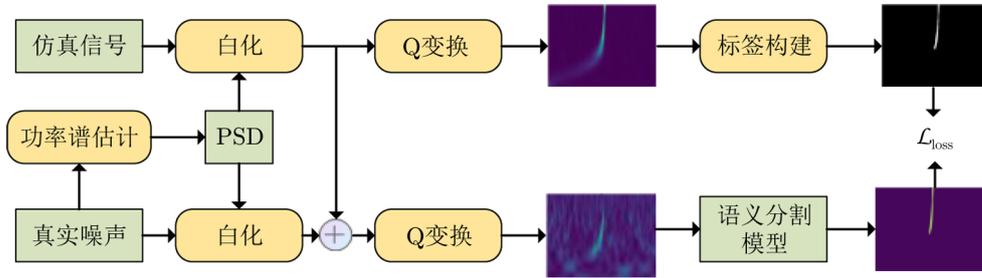


图 1 基于深度学习的引力波信号语义分割数据分析框架

3 数据准备与模型介绍

3.1 数据集

本文的训练数据由激光干涉引力波天文台获取的“噪声”数据和通过 SEOBNRv4 仿真得到的“信号”数据合成。通过 SEOBNRv4 仿真得到引力波的两个偏振模式 h_+ 和 h_\times , 获得汉福德 (Hanford) 引力波探测器探测到的引力波信号:

$$h'(t) = F^+ h_+(t) + F^\times h_\times(t) \quad , \quad (1)$$

其中, F^+ 和 F^\times 为探测器的响应函数。读取从引力波开放科学中心 (Gravitational Wave Open Science Center, GWOSC) 获得的真实噪声数据 $n(t)$, 试探性地将 $h'(t)$ 注入 $n(t)$ 。将 $n(t) + h'(t)$ 与 $h'(t)$ 进行匹配滤波计算, 得到初始信噪比值, 然后通过目标信噪比对 $h'(t)$ 进行幅度调整得到 $h(t)$, 最终合成数据为:

$$s(t) = h(t) + n(t) \quad , \quad (2)$$

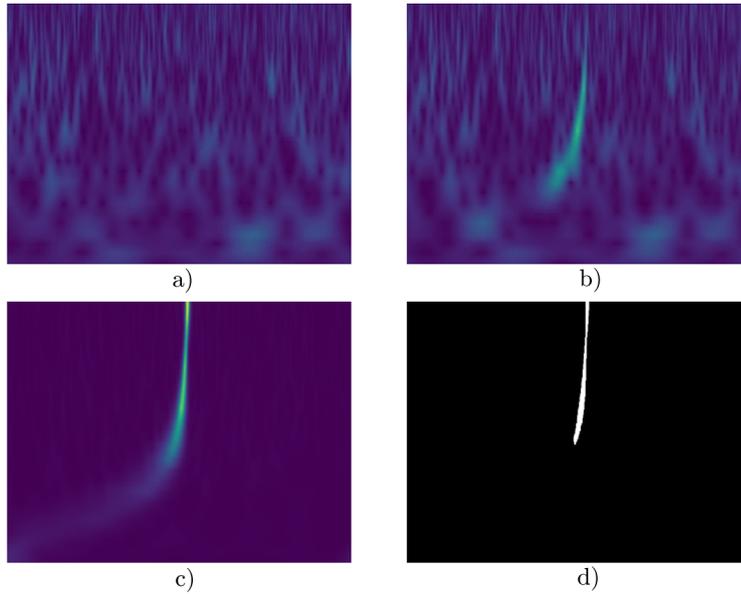
其中, $n(t)$ 为从汉福德探测器获得的 O3 数据中采样的噪声。本文采用的数据集合成的程序源代码已在 github^①公开。数据的采样频率为 4096 Hz, 双黑洞并合质量 $m_{1,2} \in (10M_{\odot}, 80M_{\odot})$, 双黑洞自旋分量分别为 $s_{1z} \in (0, 0.998)$ 和 $s_{2z} \in (0, 0.998)$, 赤经和赤纬分别在空间方位均匀采样, 极化角在 $0 \sim 2\pi$ 之间均匀采样, 合成数据的信噪比 $SNR \in (6, 20)$ 。为了过滤背景噪声的光谱分量和消除牛顿噪声的影响, 数据集中的所有数据被白化。

通过 Q 变换, 将 1 s 的时序信号转换成时频图。离散时域信号 s 通过 Q 变换转换为 X^Q 的公式:

$$X^Q(k, n) = \sum_{j=n-\lfloor N_k/2 \rfloor}^{n+\lfloor N_k/2 \rfloor} s(j)a_k^*(j-n+N_k/2) \quad , \quad (3)$$

$$a_k(n) = \frac{1}{N_k} W\left(\frac{n}{N_k}\right) \exp\left(-i2\pi n \frac{f_k}{f_s}\right) \quad , \quad (4)$$

其中, $k = 1, 2, 3, \dots, K$ 为频率点的索引, N_k 为时间窗口的长度, $\lfloor \cdot \rfloor$ 表示向下取整, $a_k^*(n)$ 表示 $a_k(n)$ 的复共轭, $W(n)$ 为高斯窗函数, f_k 为第 k 个点的中心频率。图 2 展示了一个时频图的例子, 包括“噪声”(图 2a)、“噪声+信号”(图 2b)以及“信号”(图 2c)的时频图。由“信号”的时频图生成的图像分割标签, 如图 2d) 所示。



注: a) “噪声”时频图; b) “信号+噪声”时频图; c) “信号”时频图; d) “信号”对应标签。

图 2 时频图像示意图

3.2 语义标签构建

“信号”的时频图不能直接用作训练数据的标签。图像语义分割任务中, 每一个像素的

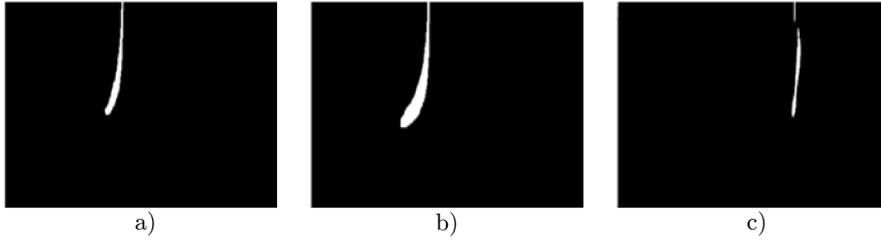
^①<https://github.com/WangSen-1006/Generate-data.git>

标签都拥有特定的物理含义。激光干涉引力波天文台获取的应变数据中, 时频图的每一个像素都可以分割为引力波、背景噪声或毛刺信号 (glitch)。本文仅研究引力波和背景噪声两种物理信息的分割, 研究方案易于迁移到更多类别的分割中。运用图像分割技术对致密双星并合释放引力波的数据进行处理, 但在实际操作中, 我们遇到了数据标注的问题。参考计算机视觉领域, 图像分割的公开数据集包括 COCO 和 PASCAL VOC 等, 这些数据集通过人工标注制作, 包括类别标注、实例定位、像素级的语义标注以及标注质量把控等多个步骤, 需要大量的人力投入, 耗费数万工时, 成本极为高昂。2022 年, Aveiro 等人^[39]采用自动标注的方法对引力波目标检测数据集进行标注。该论文将时频引力波目标存在的时间范围通过注入信号的时间范围进行标注, 将时频图中的全频率范围作为频率区间的标注, 该方法造成目标检测结果频率信息缺失的问题。与 Aveiro 等人不同的是, 本文通过对注入引力波的时频信息分析, 对引力波的时频像素进行更细致的标注。假设应变数据中埋藏的信号 $h(t)$ 的时频表示为 $H^Q(k, n)$, 对于每一个像素, 设定阈值 g , 遍历所有像素 (k, n) , 假设 $H^Q(k, n) > g$, 则信号和噪声合成应变数据的时频图 X^Q 的像素 (k, n) 标注为正类, 反之该像素标注为负类。 g 可以采用与注入信号无关的固定阈值以及与注入信号 $H^Q(k, n)$ 有关的自适应阈值这两套方案。我们发现, 固定阈值方法的标注结果依赖于应变数据的信噪比, 对于低信噪比的数据甚至所有像素都被标注为噪声类别。在本工作中, 通过自适应方法对阈值 g 进行选择, g 的取值为:

$$g = (v_{\max} - v_b)\delta + v_b \quad , \quad (5)$$

其中, $v_{\max} = \max_{k,n}(|H^Q(k, n)|)$ 表示像素最大值, $v_b = \text{avg}_{k,n}(|H^Q(k, n)|)$ 表示像素平均值。 $\delta \in [0, 1)$ 为比例因子, 当 $\delta = 0$ 时以平均像素值作为标注的阈值; 当 $\delta \approx 1$ 时, 几乎所有像素都被标注为背景噪声类别。 δ 的选择是一个困难问题, 比较部分数据在不同 δ 下的标签结果 (图 3a) 中 $\delta = 0.4$, 图 3b) 中 $\delta = 0.3$), 发现当 δ 取值小时标签覆盖范围广, 因此容易将噪声标注为引力波。通过自适应阈值对像素进行标注将导致少部分应变数据的标注结果不连通的问题 (如图 3c) 所示), 在物理上容易把单天体源判断为多天体源。不连通比例会随着 δ 的增大而增大。 $\delta = 0.4$ 时仅少部分不连通, $\delta = 0.5$ 时数据不连通问题变得普遍, 因此, 本文取 $\delta = 0.4$ 。为了解决少部分不连通问题, 借鉴计算机视觉领域人工质量把控过程, 通过人工筛选把少部分不连通的标注样本删除。构建的时频图数据集共 7982 张, 按训练集与验证集 9 : 1 的比例划分为两部分。

在未来, 本论文提出的方法还需要进一步改进, 比例因子可以作为一项优化参数。部分标注结果不连通的问题有望通过形态图像处理方法解决。未来, 还可以在分割一切大模型, 如 SAM (segment anything model)^[55] 的辅助下进行数据标注。从图 3 可以看出, 本方法存在标注不全面的问题。本方法仅对合并周期后期的分割结果进行了标注, 缺乏对环绕周期的标注。数据驱动的方法得到的模型往往能够模拟训练数据的分布。本文中生成的训练数据将环绕周期标注为背景类别, 这是本文标注偏置的体现。另外, 通过对白化信号的时频图进行分析后发现, 并合附近信号最强, 更易于分割。在未来, 改进数据的标注方法, 在数据标注中将环绕周期标注为信号类别, 有望解决标注不全面的问题。



注: a) $\delta = 0.4$ 标注结果; b) $\delta = 0.3$ 标注结果; c) 不连通标注结果。

图3 标注结果示意图

3.3 语义分割模型

语义分割模型可以看作参数化的系统。本工作中,模型的输入 $M^i \in \mathbb{R}^{3 \times 512 \times 512}$, 其中通道数为3, 图像的长和宽均为512, 模型的输出可表示为:

$$\hat{M}^o = \text{SegNet}(M^i | W_{\text{seg}}) \quad , \quad (6)$$

其中, SegNet 表示参数化的语义分割模型, W_{seg} 表示模型的可训练参数。模型的输出 $\hat{M}^o \in \mathbb{R}^{512 \times 512}$, 其中每一像素 $\hat{M}_{i,j}^o \in [0, 1]$, 能够反映每一时频点包含引力波信号的可信程度。模型参数 W_{seg} 通过交叉熵误差损失函数 (cross entropy loss, CEL) 优化, 其定义式为:

$$\zeta_{\text{loss}} = \frac{1}{NHW} \sum_{k=1}^N \sum_{i=1}^H \sum_{j=1}^W \text{CrossEntropy}(M_{i,j}^{o,k}, \hat{M}_{i,j}^{o,k}) \quad , \quad (7)$$

其中, N 表示样本数量, $M_{i,j}^{o,k}$ 和 $\hat{M}_{i,j}^{o,k}$ 分别为第 k 张时频图的真实标签和网络预测结果。

图像语义分割模型的结构能够影响分割的效果。我们分别研究了近年来应用广泛的5种图像分割模型, 它们是 U-Net^[56]、PSPNet^[45]、DeepLabv3+^[48]、HRNetV2^[58] 和 SegFormer^[59]。本文将这5种模型的构建、训练、测试以及数据集在 github[®]上公开。下面对这5种语义分割模型进行简要介绍。

2015年, Ronneberger 等人^[56]首次提出 U-Net 图像分割模型, 并将其应用于医疗图像处理。该模型主要由编码器和解码器两部分组成, 编码器采用堆叠卷积层逐步下采样提取特征, 解码器则逐步上采样还原图像的细节信息。编码-解码器间使用跳跃连接, 将浅层特征和深层特征相融合, 融合后的特征能够更准确地捕捉图像细节, 从而提高分割的准确性。

2017年, Zhao 等人^[45]提出了 PSPNet 图像分割模型, 该模型融合了深度残差网络 (ResNet) 和金字塔池化模块的优势, 能够更好地理解不同尺度下的语境关系。PSPNet 在复杂场景的图像分割任务中表现出色, 有着广泛的应用前景。

DeepLabv3+ 于2018年由谷歌的研究人员^[48]提出, 它融合了金字塔池化模块和编码器-解码器结构的优点。金字塔池化模块帮助模型获取丰富的全局信息, 而编码器-解码器结

[®]https://github.com/xianpiguozhi/GW_SegFormer.git

构能够捕获更清晰的对象边界。DeepLabv3+ 代表了当时自然图像语义分割的最高水平, 至今仍然是被广泛使用的语义分割模型之一。本文中模型的 backbone 为 MobileNetV2。

中国科学院和亚洲微软研究院^[57]于 2019 年共同提出的 HRNet, 最初用于 2D 人体姿态估计任务。HRNet 采用了一种由高到低的分辨率特征结合方式, 保留高分辨率特征的同时实现不同分辨率特征之间的信息交换。在 COCO 数据集的关键点检测、姿态估计、多人姿态估计任务中, 这一创新使 HRNet 在性能上超越了以往的工作。HRNetV2^[58]进一步将这一由高到低的分辨率特征结合方式应用于语义分割任务。本文中的模型为 HRNetV2-W48。

2021 年, Xie 等人^[59]提出 SegFormer 图像分割模型, 该模型采用分层 TransFormer 编码器, 同时还采用了轻量级的多层感知机 (multilayer perceptron, MLP) 解码器。其优势是能够融合多尺度特征, 在提升分割精度的同时也实现了更高的实时性, 使其成为一款性能出色且适用于实时应用的图像分割模型。本文中的模型为 SegFormer-B3。

4 实验结果

4.1 模型评价指标

为了评估模型的有效性, 本文采用平均交并比 ($mIoU$) 作为衡量模型性能的评价指标, $mIoU$ 定义为:

$$mIoU = \frac{1}{N} \sum_{i=1}^N IoU_i \quad , \quad (8)$$

其中, N 为像素级分类的类别数, IoU_i 为第 i 个类别的交并比, 其定义式为:

$$IoU_i = \frac{|S_i \cap \hat{S}_i|}{|S_i \cup \hat{S}_i|} \quad , \quad (9)$$

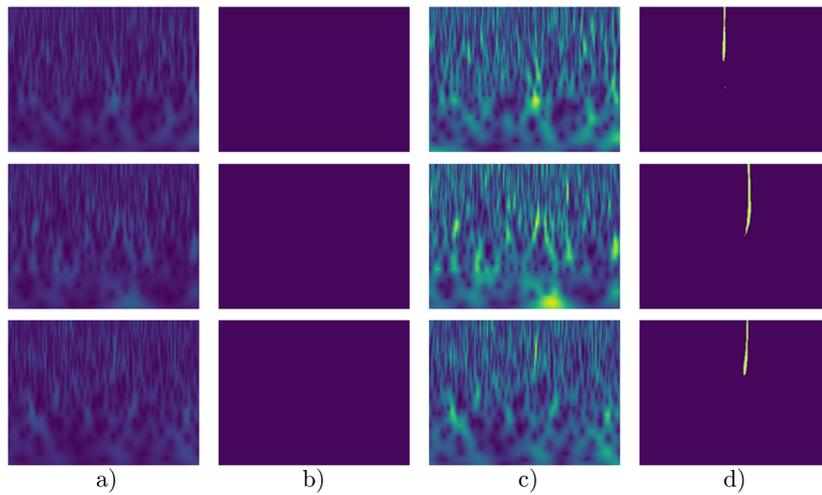
其中, S_i 为标签中属于第 i 个类别像素的集合, \hat{S}_i 为模型预测的属于第 i 个类别像素的集合, $|A|$ 表示集合 A 中元素的个数。

4.2 时频成像方法与实验结果分析

本文研究了归一化预处理对图像分割性能的影响。我们分别使用归一化和未归一化处理的数据集对模型进行了训练, 并测试其性能。图 4 展示了 SegFormer 模型对部分背景噪声测试数据的分割结果。由图 4 可知, SegFormer 模型在未归一化的噪声数据上更容易将其错误地辨识为引力波信号。因此, 本文只展示归一化预处理数据训练得到的模型的实验结果。

在训练中, 每 5 个 epoch 对验证集进行一次 $mIoU$ 计算。图 5 展示了各模型的 $mIoU$ 随训练过程的变化: SegFormer 模型的 $mIoU$ 最高, 其在大约 10 个 epoch 后趋于稳定, 达到 80% 以上; PSPNet 模型的 $mIoU$ 最低, 最终低于 70%。

目前, 引力波开放科学中心 (GWOSC) 已经公开第一次观测 (O1)、第二次观测 (O2) 和第三次观测 (O3) 得到的应变数据。探测器硬件等的升级, 使得 O3 和 O2 的背景噪声功率



注：a) 归一化的背景噪声时频图，归一化处理将着色数值归一化至 (0 ~ 11)；b) 归一化数据集训练后的 SegFormer 模型分割结果；c) 未归一化背景噪声时频图；d) 未归一化数据集训练后的 SegFormer 模型分割结果。

图 4 部分背景噪声时频图及其分割结果

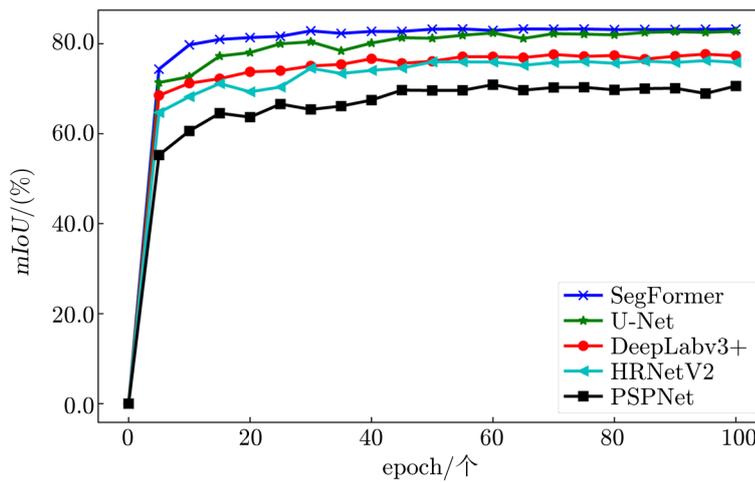


图 5 各模型平均交并比随训练的变化

谱有异。我们定量测试了各分割模型对不同背景噪声分布 (O3 和 O2) 和不同信噪比范围的分割结果。低信噪比 ($SNR \in (6, 13)$) 和高信噪比 ($SNR \in (13, 20)$) 两个信噪比范围内以 O3 和 O2 的噪声为背景分别合成 1 000 数据作为测试集 (共 4 000 数据)。图 6 和图 7 展示了 5 种分割模型分别在 O3 和 O2 背景噪声下不同信噪比范围的 $mIoU$ 。

从图 6 和图 7 可以看出, PSPNet 模型分割性能最低, 这说明 PSPNet 难以应用于引力波时频图的分割。除 PSPNet 外, 其他模型在高信噪比测试数据集上的分割结果集中分布

于 0.6 ~ 1。低信噪比的数据分割性能低于高信噪比, 这表明信噪比能够影响模型的分割性能。在 SegFormer 模型和 U-Net 模型的低信噪比分割结果中, 0.6 ~ 1 间的 IoU 占比都在 80% 左右; 在 0.8 ~ 1 间的占比中, Segformer 模型 (约 30%) 高于 U-Net 模型 (约 20%)。这些结果表明 SegFormer 模型分割性能更好, 与图 4 中验证集测试结果一致。通过图 6 和图 7 的对比, 各模型对 O2 噪声与 O3 噪声分布的分割统计特性一致, 这说明训练得到的模型对背景噪声分布具有一定的泛化能力。

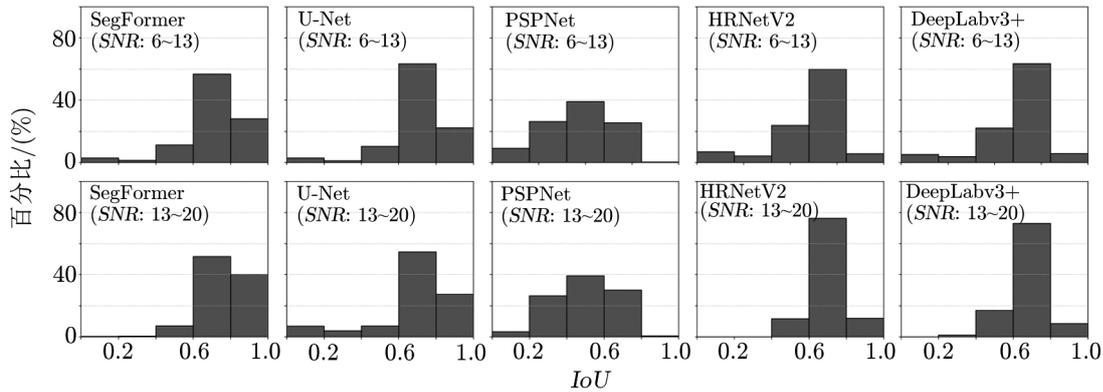


图 6 模型分割 O3 背景噪声测试集交并比统计结果

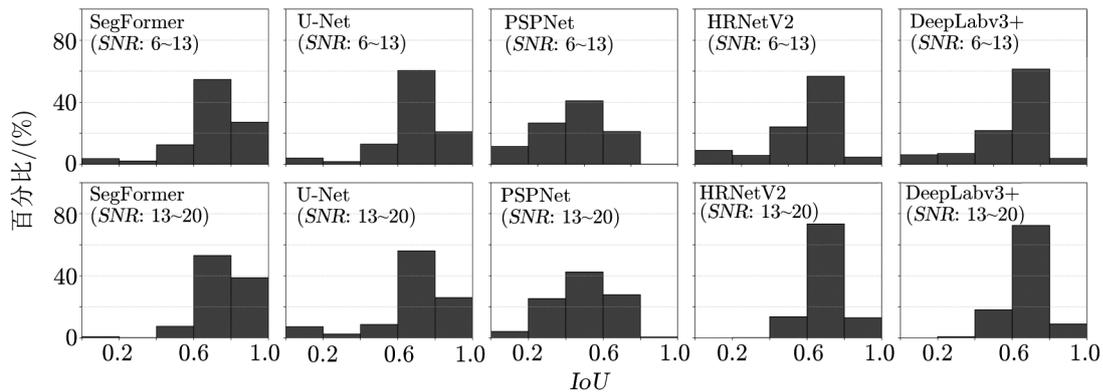
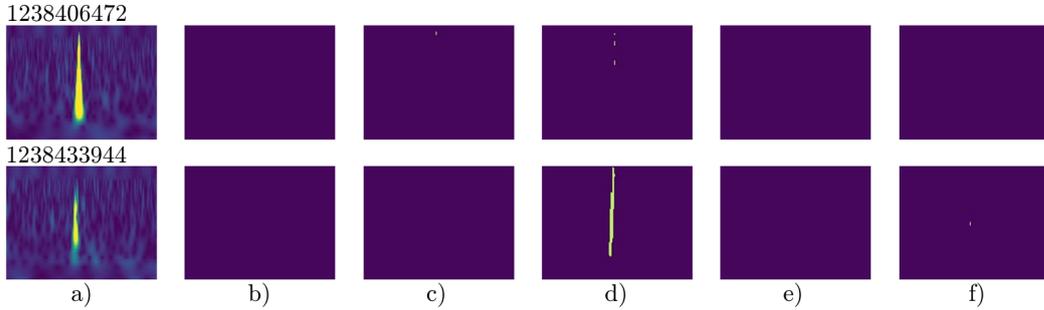


图 7 模型分割 O2 背景噪声测试集交并比统计结果

引力波探测器探测得到的背景噪声包含非平稳非高斯的毛刺成分, 这会引入模型的错误判别, 其中 Blip 毛刺在时频图上与引力波形态类似, 更易于引起错误分割。我们研究了 5 种分割模型对 Blip 毛刺的分割结果: 分别在 GPS 时间 1238406472 和 1238433944 附近获取包含 Blip 毛刺的噪声数据, 将其时频图分别输入 5 个分割模型, 获取其输出结果。由图 8 可知, HRNetV2 模型将 GPS 时间为 1238433944 的毛刺识别为引力波, DeepLabv3+ 模型将 GPS 时间为 1238433944 的毛刺少量时频点分类错误, U-Net 模型将 GPS 时间为 1238406472 的毛刺部分识别为引力波, SegFormer 和 PSPNet 两个模型都成功将 Blip 毛刺

分割为背景噪声。Transformer 结构使得 SegFormer 更易于获取时频图的全局特征，能够避免局部亮斑引起的错误分割。在各个层面的分割结果表明，SegFormer 更适合用于引力波信号时频图的分割。



注: a) 毛刺事件时频图; b) SegFormer 模型分割结果; c) U-Net 模型分割结果; d) HRNetV2 模型分割结果; e) PSPNet 模型分割结果; f) DeepLabv3+ 模型分割结果。

图 8 毛刺事件分割结果

联合分析分割模型对正类（噪声+信号）和负类（噪声）的响应。假设第 k 张时频图的分割结果为 $M^{o,k}$ ，分割得到的正类和负类的像素集合分别为 S_+^k 和 S_-^k 。设定整数阈值 th ，当 $|S_+^k| > th$ 时将整张时频图分类为正类，反之将该时频图分类为负类。遍历每一 th 值，获得测试集的正确预警概率 (TAP) 和虚警概率 (FAP)，绘制受试者工作特征曲线 (receiver operating characteristic, ROC) 如图 9 所示。由图可知，SegFormer 在正类中表现最佳，其虚警率也最高，难以直接应用于引力波的探测任务。未来，在训练集中加入负类数据可能会提升模型对正负类的区分度。目前，该方法在引力波探测方面没有优于基于二分类的深度学习模型。

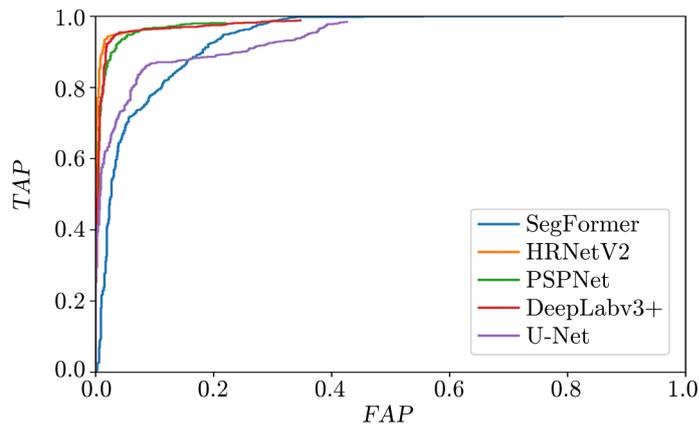


图 9 5 种分割模型应用于引力波搜寻的受试者工作特征曲线 (ROC)

5 真实事件结果

在 O1、O2 和 O3 这 3 次观测中, LVK 组织共公布 90 多个置信双星并合事件, 本章展示图像分割模型对 Hanford 和 Livingston 干涉仪获得的共同置信引力波事件的分析结果。

5.1 语义分割模型在 O1 和 O2 真实事件中的分割结果

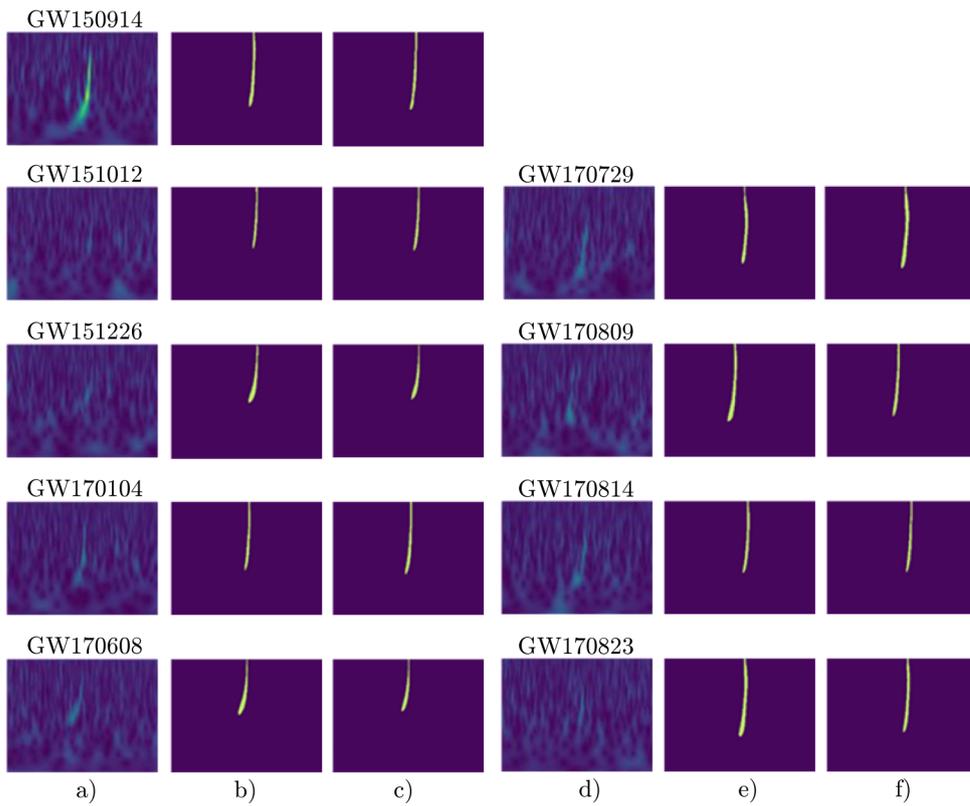
将模型用于 GWTC1 公布的引力波事件数据分析, 每个引力波事件的应变数据持续时间为 16 s, 并且波源并合信号位于数据的中心附近。时频图的时长设置为 1 s, 通过滑动窗口法对 16 s 的数据进行扫描, 滑动步长设置为 0.5 s。我们对数据集 (训练集和测试集) 中标记为引力波的像素数量进行了详细计算, 数据集中标记为引力波的像素个数大于 850 的图像占比达 95.3%, 因此, 对于分割后的时频图像, 当分割为引力波的像素个数大于 850 时, 将其视为含有引力波。表 1 展示了各模型对真实事件附近数据的分割结果, \checkmark 表示成功对引力波信号进行分割。

表 1 Hanford 干涉仪 O1、O2 真实事件的分割结果

	PSPNet	HRNet	DeepLabv3+	U-Net	SegFormer
GW150914	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark
GW151012			\checkmark	\checkmark	\checkmark
GW151226			\checkmark	\checkmark	\checkmark
GW170104	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark
GW170608		\checkmark		\checkmark	\checkmark
GW170729	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark
GW170809		\checkmark	\checkmark	\checkmark	\checkmark
GW170814	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark
GW170817					
GW170818					
GW170823	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark

SegFormer 和 U-Net 能够成功分割 9 个真实事件。5 个分割模型都未能成功分割 GW170817 和 GW170818, 原因如下: GW170817 属于双中子星并合事件, 本文训练集仅包含双黑洞参数空间, 其在本文训练集参数之外; GW170818 在 Hanford 干涉仪的匹配信噪比仅为 4.1, 即使理论最优的匹配滤波流水线如 PyCBC 也不能将其成功探测。SegFormer 和 U-Net 对除 GW170817 和 GW170818 以外的 O1 和 O2 置信事件的分割结果如图 10 所示。从局部像素来看, GW151226 噪声更强, 从全局来看其具有致密双星并合释放引力波的啁啾特征。SegFormer 中的自注意特征提取方法更能够提取全局特征, 所以该方法能够分割信噪比较低的 GW151226 的时频图。虽然卷积操作中感受野有限, 但 U-Net 通过堆叠卷积层的方式, 逐渐扩大感受野, 使其具备了一定的全局识别能力, 能够辨识出 GW151226 中的啁啾特征。从图中可以看出, SegFormer 和 U-Net 的分割结果一致。

需要注意的是, 模型在滑动窗口分析过程中会产生错误分割, 通过 Hanford 和 Liv-



注: a)、d) Hanford 干涉仪时频图; b)、e) SegFormer 模型分割结果; c)、f) U-Net 模型分割结果。

图 10 SegFormer 与 U-Net 在 Hanford 干涉仪的 O1 和 O2 时期真实事件的分割结果

Livingston 干涉仪分割结果的交互验证, 能够降低错误分割的发生。与端对端分类相比, 图像分割包含更加丰富的信息, 可以交互验证 Hanford 和 Livingston 干涉仪时频图上相近时间和频率位置的分割。图 11 展示了交互验证的基本原理。

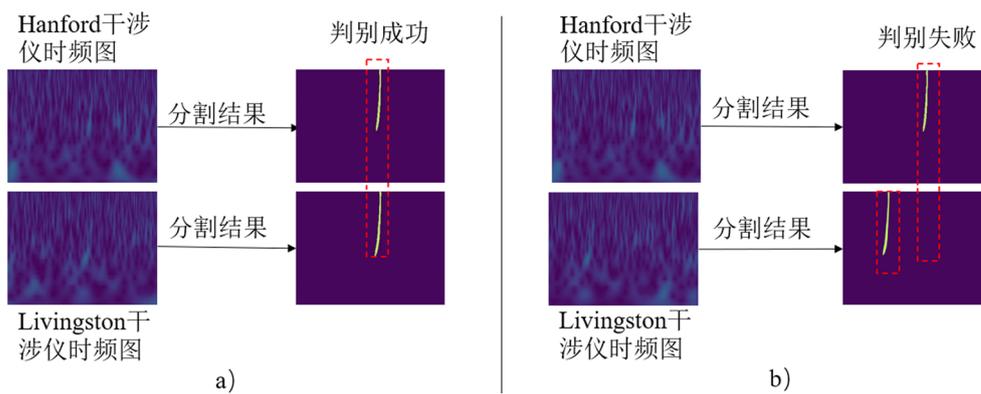
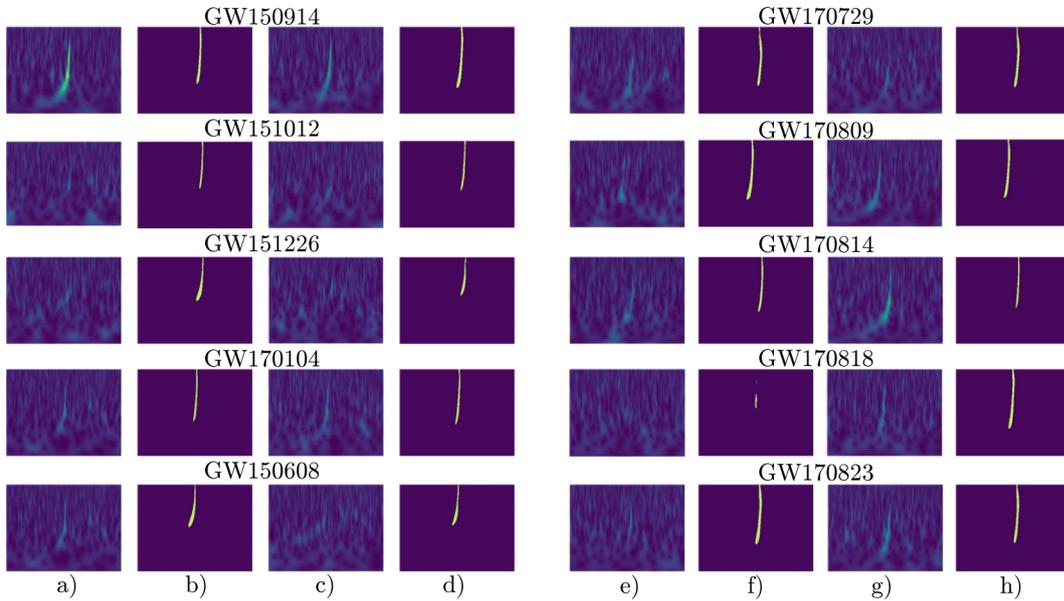


图 11 交互验证判别示意图

使用交互验证, 11 个事件的 341 张时频图, 未出现错误判别结果。图 12 展示了 SegFormer 模型在 Hanford 和 Livingston 干涉仪获得的 10 个双黑洞合并事件中的分割结果。除了 GW17081, 其他 9 个双黑洞合并事件都已成功探测。



注: a)、e) Hanford 干涉仪时频图; b)、f) SegFormer 模型分割结果; c)、g) Livingston 干涉仪时频图; d)、h) Livingston 干涉仪时频图分割结果。

图 12 Hanford 和 Livingston 干涉仪的 O1 与 O2 时期真实事件的分割结果

5.2 语义分割模型在 O3 真实事件分割结果

本节研究了 SegFormer 模型对 O3 数据的搜寻性能。在 O3a 和 O3b 的置信事件中, Hanford 和 Livingston 干涉仪共同探测到的事件数量为 40 和 32。在 Hanford 和 Livingston 干涉仪上交互验证, SegFormer 模型正确分割 50 个真实事件, 同时产生一个错误分割。错误分割结果如图 13 所示, 该错误发生在 GW191113 事件后的第 5 秒。



注: a) Hanford 干涉仪时频图; b) Hanford 干涉仪时频图分割结果; c) Livingston 干涉仪时频图; d) Livingston 干涉仪时频图分割结果。

图 13 O3 时期 Hanford 和 Livingston 干涉仪上的错误判别

6 总结与展望

本文采用时频图作为输入, 构建并训练了不同的图像语义分割模型, 以研究其在引力波时频数据分析中的可行性。通过定量分析模型性能, 结果显示 SegFormer 模型更适合用于引力波时频图的分割。此外, 还初步探索了 Blip 毛刺对模型分割的影响。最后, 将模型应用于 Hanford 和 Livingston 干涉仪共同探测的置信引力波事件。对于 O1 和 O2 探测的 10 个双黑洞合并事件, SegFormer 模型成功交互验证了 9 个事件; 在 Hanford 端, GW170818 未能成功分割。对于 O3 事件的应用, SegFormer 模型在 Hanford 和 Livingston 干涉仪数据上交互验证了 72 个真实事件中的 50 个事件, 同时产生一个错误判断。这些结果表明, SegFormer 模型对引力波时频数据分割可行, 并有望成为未来基于深度学习的引力波信号搜寻流水线中的一个关键组件, 用于进一步提取信号的时频信息。未来, 可以将这一工作扩展到引力波、背景噪声、毛刺信号的多类别分割, 有望实现像素级多类别分析。

参考文献:

- [1] Abbott B P, Abbott R, Abbott T D, et al. Phys Rev Lett, 2016, 116: 061102
- [2] Abbott B P, Abbott R, Abbott T D, et al. Phys Rev Lett, 2017, 119: 161101
- [3] Abbott R, Abbott T D, Acernese F, et al. Phys Rev X, 2023, 13: 011048
- [4] Abbott B P, Abbott R, Abbott T D, et al. Nature, 2017, 551: 85
- [5] Abbott R, Abbott T D, Abraham S, et al. Phys Rev Lett, 2021, 126: 241102
- [6] Abbott R, Abe H, Acernese F, et al. Phys Rev D, 2022, 106: 102008
- [7] Romero-Rodríguez A, Martínez M, Pujolós O, et al. Phys Rev Lett, 2022, 128: 051301
- [8] Li Y F, Heng I S, Chan M L, et al. Phys Rev D, 2022, 105: 043010
- [9] Baghi Q, Korsakova N, Slutsky J, et al. Phys Rev D, 2022, 105: 042002
- [10] Fan H M, Hu Y M, Barausse E, et al. Phys Rev D, 2020, 102: 063016
- [11] Luo Z, Wang Y, Wu Y L, et al. Progress of Theoretical and Experimental Physics, 2021, 2021(5): A108
- [12] Cuoco E, Patricelli B, Iess A, et al. Nature Computational Science, 2022, 2(8): 479
- [13] Gebhard T D, Kilbertus N, Harry I, et al. Phys Rev D, 2019, 100: 063015
- [14] George D, Huerta E A. Phys Rev D, 2018, 97: 044039
- [15] George D, Huerta E A. Physics Letters B, 2018, 778: 64
- [16] Fan X L, Li J, Li X, et al. Science China Physics, Mechanics & Astronomy, 2019, 62: 1
- [17] Luo H M, Lin W, Chen Z C, et al. Frontiers of Physics, 2020, 15: 1
- [18] Li X R, Yu W L, Fan X L, et al. Frontiers of physics, 2020, 15: 7
- [19] Schöfer M B, Zelenka O, Nitz A H, et al. Phys Rev D, 2023, 107: 023021
- [20] Gebhard T D, Kilbertus N, Harry I, et al. Phys Rev D, 2019, 100: 063015
- [21] Wei W, Khan A, Huerta E A, et al. Phys Lett B, 2021, 812: 136029
- [22] Huerta E A, Khan A, Huang X, et al. Nature Astronomy, 2021, 5(10): 1062
- [23] Ma C L, Wang W, Wang H, et al. Phys Rev D, 2022, 105: 083013
- [24] Zhang Y, Xu H, Liu M, et al. Phys Rev D, 2022, 106: 122002
- [25] Wang H, Wu S, Cao Z, et al. Phys Rev D, 2020, 101: 104003
- [26] Ma C, Wang W, Wang H, et al. Phys Rev D, 2023, 107: 063029
- [27] Wei W, Huerta E A. Phys Lett B, 2020, 800: 135081

- [28] Chatterjee C, Wen L, Diakogiannis F, et al. *Phys Rev D*, 2021, 104: 064046
- [29] Ren Z, Wang H, Zhou Y, et al. *arXiv:2212.14283*, <https://arxiv.org/abs/2212.14283>, 2022
- [30] Green S R, Simpson C, Gair J. *Phys Rev D*, 2020, 102: 104057
- [31] Kuo H S, Lin F L. *Phys Rev D*, 2022, 105: 044016
- [32] Dax M, Green S R, Gair J, et al. *Phys Rev Lett*, 2021, 127: 241103
- [33] Chua A J K, Katz M L, Warburton N, et al. *Phys Rev Lett*, 2021, 126: 051102
- [34] Zhao T, Lyu R, Wang H, et al. *Communications Physics*, 2023, 6(1): 212
- [35] Zhang X T, Messenger C, Korsakova N, et al. *Phys Rev D*, 2022, 105: 123027
- [36] Chen M N, Zhong Y H, Feng Y, et al. *Science China Physics, Mechanics & Astronomy*, 2020, 63: 129511
- [37] Sun M F, Li J, Cao S, et al. *arXiv:2307.16437*, <https://arxiv.org/abs/2307.16437>, 2023
- [38] Menendez-Vazquez A, Kolstein M, Martinez M, et al. *Phys Rev D*, 2021, 103: 062004
- [39] Aveiro J, Freitas F F, Ferreira M, et al. *Phys Rev D*, 2022, 106: 084059
- [40] Jin S J, Wang Y X, Sun T Y, et al. *arXiv:2305.19003*, <https://arxiv.org/abs/2305.19003>, 2023
- [41] Chen L C, Papandreou G, Kokkinos I, et al. *arXiv:1412.7062*, <https://arxiv.org/abs/1412.7062>, 2014
- [42] Taghanaki S A, Abhishek K, Cohen J P, et al. *Artificial Intelligence Review*, 2021, 54(1): 137
- [43] Yuan X, Shi J, Gu L. *Expert Systems with Applications*, 2021, 169: 114417
- [44] Long J, Shelhamer E, Darrell T. *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015: 3431
- [45] Zhao H, Shi J, Qi X, et al. *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017: 2881
- [46] Dai J, Qi H, Xiong Y, et al. *Proceedings of the IEEE international conference on computer vision*, 2017: 764
- [47] Chen L C, Papandreou G, Schroff F, et al. *arXiv:1706.05587*, <https://arxiv.org/abs/1706.05587>, 2017
- [48] Chen L C, Zhu Y, Papandreou G, et al. *Proceedings of the European conference on computer vision (ECCV)*, 2018: 801
- [49] Chen L C, Papandreou G, Kokkinos I, et al. *IEEE transactions on pattern analysis and machine intelligence*, 2017, 40(4): 834
- [50] Yuan Y, Chen X, Wang J. *Computer Vision - ECCV 2020: 16th European Conference, Glasgow: Springer International Publishing*, 2020: 173
- [51] Jin Z, Liu B, Chu Q, et al. *Proceedings of the IEEE/CVF International Conference on Computer Vision*. 2021: 7189
- [52] Liu M, Schonfeld D, Tang W. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2021: 9726
- [53] Yuan Y, Xie J, Chen X, et al. *Computer Vision - ECCV 2020: 16th European Conference, Glasgow: Springer International Publishing*, 2020: 489
- [54] Borse S, Wang Y, Zhang Y, et al. *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 2021: 5901
- [55] Kirillov A, Mintun E, Ravi N, et al. *arXiv:2304.02643*, <https://arxiv.org/abs/2304.02643>, 2023
- [56] Ronneberger O, Fischer P, Brox T. *Medical Image Computing and Computer Assisted Intervention MICCAI 2015: 18th International Conference, Munich: Springer International Publishing*, 2015: 234
- [57] Sun K, Xiao B, Liu D, et al. *Proceeding of the IEEE/CVF conference on computer vision and pattern recognition*. 2019: 5696
- [58] Sun K, Zhao Y, Jiang B, et al. *arXiv:1904.04514*, <https://arxiv.org/abs/1904.04514>, 2019
- [59] Xie E, Wang W, Yu Z, et al. *Advances in Neural Information Processing Systems*, 2021, 34: 12077

Research on Gravitational Wave Data Processing for Binary Compact Object Mergers Using Deep Learning Semantic Segmentation Methods

MA Cun-liang¹, ZHONG Guo-jian¹, MIN Yuan¹, JIA Ming-zhen¹, HE Guan-sheng^{2,3}

(1. Jiangxi University of Science and Technology, Ganzhou 341000, China; 2. University of South China, Hengyang 421001, China; 3. Purple Mountain Observatory, Chinese Academy of Sciences, Nanjing 210023, China)

Abstract: Recent studies have shown that deep learning (DL) based gravitational wave (GW) search holds the promise of addressing the inefficiencies of matched filtering method. Compared to matched filtering for GW search and Bayesian posterior-based or deep learning-based methods for GW source parameter estimation, the DL based end-to-end GW search lacks the time-frequency information of the source, such as the arrival time and the duration of the signals within the sensitivity range of the detectors. This study represents the first implementation of fine-grained classification of time-frequency points in strain signals and explores a gravitational wave search approach based on image semantic segmentation. The data is synthesized by combining real noise detected by the Hanford interferometer and simulated physical signals. It is then transformed into time-frequency images using the Q-transform, and a pixel-level annotated dataset is constructed. The time-frequency image semantic segmentation model is built and trained, and the results demonstrate the feasibility of image semantic segmentation methods in the analysis of gravitational wave signal data. The image semantic segmentation methods have the potential to become a component of the gravitational wave signal search pipeline for further extraction of time-frequency information of signals in the future.

Key words: gravitational wave data analysis; deep learning; semantic segmentation