arXiv:2303.00693v1 [hep-ph] 1 Mar 2023

PE-GAN：Belle II 的 PXD 图像先验嵌入生成对抗网络

Baran Hashemi慕尼黑路德维希-马克西米利安大学物理学院，德国慕尼黑
gh.hashemi@physik.uni-muenchen.de

Nikolai Hartmann慕尼黑路德维希-马克西米利安大学物理学院，德国慕尼黑
ni
kolai.hartmann@physik.uni-muenchen.de

Thomas Kuhr慕尼黑路德维希-马克西米利安大学物理学院，德国慕尼黑
thomas.Kuhr@lmu.de

Martin Ritter慕尼黑路德维希-马克西米利安大学物理学院，德国慕尼黑
martin.ritter@lmu.de

Matej Srebre慕尼黑路德维希-马克西米利安大学物理学院，德国慕尼黑
m
atej.srebre@physik.uni-muenchen.de

摘要

像素顶点探测器（PXD）是 Belle II 探测器的重要组成部分，用于记录粒子位置。PXD 和其他传感器的数据使我们能够重建粒子径迹和衰变顶点。我们通过将测得或模拟的背景击中模式添加到模拟信号粒子产生的击中中，来模拟本底击中对径迹重建的影响。为避免重建径迹产生系统偏差，该模型需要大量统计独立的 PXD 背景噪声样本。然而，精细像素化 PXD 的数据需要占用大量存储空间。为高效地产生背景噪声，我们探索了使用条件生成对抗网络（GAN）构建按需 PXD 背景生成器的思路，并根据 PXD 传感器的数量对其进行调整，以提高图像保真度并生成依传感器而异的 PXD 击中图。

关键词 生成对抗网络 · 高颗粒度探测器模拟 · PXD Belle II · 对比学习

1 引言

Belle II [1]是位于筑波、日本 SuperKEKB [2]加速器上的 B 工厂实验，其峰值瞬时亮度为 $2 \times 10^{34}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$。该加速器以 $\sqrt{s} \approx 10.6$GeV 的质心能量碰撞电子和正电子，从而产生 $\Upsilon(4S)$ 共振态，该态几乎只衰变为带电或中性的纠缠 B 介子对。这些过程称为信号衰变，属于各向同性事件。此外还有其他过程，即喷注状连续谱事件（$e^+e^- \longrightarrow q\bar{q},\ \ q = u,\ \ d,\ \ s,\ \ c$）。精确测量 CP-破坏，以及搜寻 B 介子和 τ 轻子的稀有或“禁戒”衰变以探测新物理，很大程度上依赖于大量记录事件，以及重建 B 介子和轻子衰变顶点的精度。为此，Belle II 计划收集比其前身 Belle 多 50 倍的数据集。加速器升级使粒子碰撞率如期提高。因此，我们预计在设计亮度下，所有子探测器的本底都会增加。为了充分利用大幅增加的记录数据集，必须严格控制模拟数据集的系统不确定度。这要求在探测器模拟中准确纳入束流本底效应。

各种本底过程会在探测器中产生信号，而这些信号并非源自我们感兴趣的物理过程。这些过程可分为两类：束流诱发过程和依赖瞬时亮度的过程。束流诱发过程源于同步辐射，以及束流粒子与束流管、弯转磁铁中的残余气体或束团内粒子的碰撞；依赖瞬时亮度的过程则包括正负电子碰撞所产生的“不感兴趣”的物理过程，例如巴巴散射或双光子过程。

有两种方法可以纳入并模拟这些束流本底效应：本底混合方法和本底叠加方法。在本底混合方法中，预先模拟的 Geant4 [3] 本底击中会作为信号事件的模拟击中，与模拟能量沉积相加。随后，本底和信号的贡献会合并，并同时进行数字化。这样做的优点是，后续的数字化模拟能够正确考虑本底粒子和信号粒子的能量沉积对同一通道产生贡献的情况。另一方面，将采用这种所谓混合技术的本底模拟与测量结果进行比较时，会发现存在明显差异。本底叠加方法利用随机触发事件和真实数据生成本底事件样本。这种方法在数字信号级进行，因此需要针对探测器对本底和信号共同影响同一通道的情况进行专门处理。尽管如此，它预期能更真实地模拟束流本底效应。

2 像素顶点探测器

像素顶点探测器（PXD） [4] 被设计为 Belle II 最内层的半导体子探测器。PXD 测量穿过的粒子的位置，这些粒子源自粒子碰撞，以便精确重建衰变顶点。因此，它被放置在非常靠近相互作用点的位置，也因此暴露于大量本底辐射之中。PXD 的设计配置由 2 个探测器层中的 40 个模块组成。内层由 16 个模块构成，组装在 8 个平面梯形组件上。外层由 12 个梯形组件或 24 个模块组成。每个模块由一个 250×768 DEPFET [5] 像素传感器的像素矩阵组成，最终生成一幅 250×768 命中图。目前仅安装了 20 个模块。PXD 的读出时间为 $20\ \mu s$，在此期间会累积约 5000 次束流交叉的数据。因此，绝大多数电荷高于阈值的像素信号都来自束流本底。只有当像素信号的值超过阈值时才会被读出。在我们的模拟数据中，阈值为 7 ADU。我们将在第 4.2 节讨论这一问题。

2.1 问题

本底叠加方法的一个问题是，读出、存储和分发随机触发事件需要大量资源。用于本底叠加的格式中，单个事件的大小约为 200 kB，约为分析所用格式中事件大小的 20 倍。此外，束流本底文件必须分发到所有生成依赖运行条件的蒙特卡罗样本的站点。存储需求会因此随站点数量成倍增加。尤其是较小的站点，没有足够的存储资源和带宽来存放全束流本底数据集。即使将本底文件从主服务器复制出来，并通过网络访问，每个作业也必须访问多个文件。例如，要生成 1 Tb 的蒙特卡罗文件，就必须复制 20 Tb 的 PXD 本底文件，这会大幅增加数据传输量和网络流量。

我们针对这一问题的解决方案是将整个流程端到端化。首先，使用条件生成智能体，按需为每个传感器即时生成这些命中图，然后从触发像素中提取 PXD命中数据。唯一的存储成本是生成模型的权重。

3 PXD背景生成

本研究属于粒子物理学“快速模拟”项目的研究方向。该项目在 Goodfellow 等人将生成对抗网络（GAN）引入图像数据后启动。 [6] GAN 是一种灵活的无监督深度学习架构，可通过训练学习一个函数，使其匹配给定的数据分布。生成对抗网络已广泛应用于粒子物理学，以实现大型强子对撞机 [7] [8], 的更快速探测器模拟，但尚未用于像素探测器。生成对抗网络包含两个网络：生成器和判别器，它们参与一场双人零和博弈，分别负责将随机噪声映射为样本，以及区分真实样本和生成样本。最终目标是找到双人极大极小问题的纳什均衡。在我们的案例中，我们需要构建一个模型，以尽可能快的速度、高保真度地生成依赖传感器的 PXD 本底命中图（图像）。目前采用的方法暂不考虑两个层之间的相关性，因为大多数本底来自非同时时间事件。由于通常无法获得与同一批粒子造成的两个相关 PXD 本底击中相对应的 SVD 命中，因此不会重建出径迹。因此，忽略这些相关性对径迹重建性能的预期影响较小，但仍需进行详细研究。

每个事件中的数据由 40 幅依赖于传感器的灰度 250×768 图像组成。在训练中，我们使用了 40000 个蒙特卡罗模拟事例；在评估中，我们使用了 10000 个蒙特卡罗模拟事例。这些事件是在设计亮度下预期本底水平两倍的条件下模拟得到的。未来我们还计划使用真实数据（随机触发）。

在这里，我们面临着几个挑战：

1. 获得高分辨率图像。一般来说，对于生成对抗网络，无法保证生成器和判别器的训练保持平衡。因此，其中一个网络可能会比另一个更强大，而在大多数情况下，判别器会更强。当判别器变得比生成器强大得多时，伪造图像就很容易与真实图像区分开来，从而达到判别器的梯度消失的状态，无法为生成器的进一步训练提供指导。由于生成有意义的高频细节较为困难，在处理高分辨率图像时，这种情况发生得更加频繁。另一方面，判别器能力不足会导致模式坍塌，因为生成器会贪婪地优化其损失函数以“欺骗”判别器，而每次生成器更新都会部分地向模式坍塌发展 [9]。在消除梯度消失和模式坍塌之间保持平衡的关键，是使用具有较强泛化能力的判别器。因此，最终采用了更深层网络和非局部块，并将 PXD梯形编号（传感器 ID）作为条件类别，以尝试捕获全局和局部信息及依赖关系，而这些信息和依赖关系通常会消耗大量内存。

2. 图像稀疏。作为判别器一部分的卷积神经网络（CNN）难以处理稀疏数据 [10]。它们需要稠密数据才能充分学习。直接将 CNN 应用于稀疏数据只能取得部分效果，因为它们对数据缺失很敏感。因此，我们需要采用更复杂的方法，并将传感器 ID 等先验信息注入学习过程。

3. 训练期间缺乏评估。处理 ImageNet [11], 这类通用图像数据集时，可以使用 Inception 分数（IS） [13] 或 Fréchet Inception 分数（ FID ） [14] 来评估和验证训练效果，从而选出最佳模型，因为生成对抗网络的训练极不稳定。但在我们的案例中，由于数据集具有特殊性，常用于自然图像的评估指标并不适用。

3.1 模型

在综合考虑多种模型和结构，并根据第 5 节所述的评估方法，通过反复试验和错误最小化调整了许多超参数后，我们选择了采用对比条件化 [16] 和一致性正则化 [18] 的 BigGAN-deep [12] 修改版。

我们首先在原有模型 WGAN-gp [20], 中增加了一个嵌入层，将 40 个传感器编号作为类别标签输入模型，以此作为研究基线。由于卷积核尺寸较大，该模型的基于事件的图像生成速度相对于整体模拟时间而言较慢，而且未能完整捕捉像素强度分布。随后，我们尝试了结构更复杂但速度更快的模型。这些模型的卷积核尺寸较小，并采用了专为高分辨率图像生成设计的特殊残差块，例如 BigGAN [12]。在

训练过程中，我们的模型出现了训练崩溃，因此及早停止训练至关重要。正如 BigGAN 论文 [12],中所建议的，我们监测了生成器和判别器各自权重矩阵的奇异值，并在图 2 中展示了监测结果。随着训练进行，两者的首个奇异值都会增大；但每当出现模式坍塌或判别器对训练集过拟合时，生成器的首个奇异值就会急剧增大。我们尝试了多种处理方法，例如降低判别器的学习率，以及采用不同的正则化方法，如正交正则化、小批量判别 [13], Dropout 和 L2正则化；潜变量优化 [21], 和特征量化 [22]也同样无效。因此，我们认为必须改用更简单的模型。

随后，结合图像数据集的特点、内存限制，以及训练过程中模型的反馈和表现不佳之处，我们决定采用一个更简单的模型。因此，我们使用了 SAGAN [15]。自注意力生成对抗网络是一种架构，能够让生成器和判别器对图像流形上的长程依赖进行建模。其核心思想是让生成器能够“看到”卷积核范围内的内容，以及全局细节。通常，卷积作为局部模块，具有参数共享和平移不变性等优良特性，并以分层的方式进行学习。然而，对于复杂结构和精细的几何轮廓，则需要长程细节。这正是注意力发挥作用的地方。通过选择这种架构，我们深入了解了模型之间的差异会如何影响结果；最终评估结果显示，与 BigGAN 相比，一个更深、结构更简洁且层数多于 SAGAN 的模型表现更好。

在我们的模型中，我们增加了生成器和判别器中每个块的深度，并将每层的乘法通道数设为 32。经过一段时间的训练以及针对 BigGAN-deep 的消融研究后，我们决定移除最初在 SNGAN 中引入的投影判别 [23]。投影判别的目标是：当图像为真实图像时，最大化其嵌入在相应目标嵌入上的投影；而当图像为伪造图像时，最小化内积值。然而，由于它只利用数据到类别关系，因此会导致判别器过拟合和训练崩溃问题 [12]。因此，我们将其替换为对比条件判别机制 [16]。该机制基于一种源自度量学习的条件对比损失，该损失考虑每个批次中的数据到数据关系。当类别标签相同时，它会使多幅图像的嵌入在嵌入空间中彼此更加接近。换言之，判别器不仅能够捕捉样本之间的数据到类别关系，还能够捕捉数据到数据关系。

在度量学习损失的背景下，由于我们处理的是近条件图像生成任务，因此还将先验信息的嵌入与类别标签一起添加到判别器和生成器中。判别器中的关系嵌入模块由语义信息的线性投影，以及对标签嵌入和语义信息投影拼接后结果进行的自注意力组成。设 $\mathbf{X} = \{x_1, ..., x_m\}$，其中 $x_i \in \mathbb{R}^{W \times H}$ 为采样得到的训练图像，$\mathbf{y} = \{y_1, ..., y_m\}$，其中 $y_i \in \mathbf{Z}$ 为对应标签，取值范围为 1 到 40；$\mathbf{s} = \{s_1, ..., s_m\}$，其中 $s_i \in \mathbf{R}$ 为各类别传感器的平均占用率。此外，我们定义线性投影函数 $l^{(d)}: \mathbb{R}^k \to \mathbb{S}^d$ 以及嵌入函数 $\mathbf{e}: y_i \to \mathbb{R}^{d/4}$。对于关系嵌入，我们将拼接结果输出的 40 个词元作为查询、键和值，每个词元的维度为 d；点积自注意力 [17] 根据表示对之间的点积相似度分数确定权重，并据此计算特征表示的加权平均，

$$Q = l^{(d)}\ [l^{(d/4)}(\mathbf{s}) \mid\mid \mathbf{e}]$$

$$A = softmax\ (\frac{QK^T}{\sqrt{d}})\ V.$$

然后，使用线性层 判别器输出的投影，可以定义条件对比损失为 s,

$$L(x_i, y_i, s_i; t) = -\log\left(\frac{\exp\left(h_i^T a_i/t\right) + \sum_{k=1}^{m} \mathbf{1}_{k=i}.\exp\left(h_i^T h_k/t\right)}{\exp\left(h_i^T a_i/t\right) + \sum_{k=1}^{m} \mathbf{1}_{k\neq i}.\exp\left(h_i^T h_k/t\right)}\right)$$

$$H = l^{(d)}[D(x_i)]$$

其中 $a_i \in A$ 和 $h_i \in H$。最小化上述对比损失会缩小具有相同标签的图像嵌入之间的距离，同时增大其他图像嵌入之间的距离。它考虑了数据到数据关系 $h_i^T h_k$ 和数据到类别关系 $h_i^T a_i$，且无需进行数据增强。因此，我们不仅帮助判别器依据图像进行引导，还通过归纳偏置注入，使其利用细粒度先验信息。这些先验语义信息基于模拟图像中不同的平均占用率。注入这些额外信息，不仅提升了平均占用率分布方面的表现，还使训练稳定性提高，并额外稳定了近 20000 次迭代。

我们还在训练过程中使用了一致性正则化 [18]，显著延长了训练稳定性，并延缓了坍塌。一致性正则化的核心思想是惩罚分类器对某些语义保持噪声的敏感性。在我们的案例中，噪声扰动包括图像翻转和随机平移，因为生成器的输出中存在带有这些增强伪影的图像。

最终，如表1所示，我们的模型结合上述技术，并在生成器和判别器中采用正交初始化的 [24]残差块 [12] 以及铰链损失 [15], , 成功实现了依赖传感器的PXD命中图生成。我们使用 PyTorch 框架 [19]，在 Tesla V100 GPU 上训练了该模型。该模型在一台 CPU 上约 6 秒即可为一个事件生成 40 张命中图。

$z \in \mathbb{R}^{128} \sim \mathcal{N}(0, I)$
嵌入(y) $\in \mathbb{R}^{128}$
$ch = 32$
线性层(128+128) $\rightarrow 4 \times 12 \times 16ch$
残差块 $16ch \rightarrow 16ch$
残差块 上采样 $16ch \rightarrow 16ch$
残差块 $16ch \rightarrow 16ch$
残差块 上采样 $16ch \rightarrow 8ch$
残差块 $8ch \rightarrow 8ch$
残差块 上采样 $8ch \rightarrow 8ch$
残差块 $8ch \rightarrow 8ch$
上采样残差块 $8ch \rightarrow 4ch$
非局部注意力块
残差块 $4ch \rightarrow 4ch$
上采样残差块 $4ch \rightarrow 2ch$
残差块 $2ch \rightarrow 2ch$
上采样残差块 $2ch \rightarrow ch$
批归一化、ReLU、 3×3 卷积 $ch \rightarrow 1$
Tanh

零填充灰度图像
$x \in \mathbb{R}^{256 \times 768 \times 1}$
卷积块 $1 \rightarrow ch$
残差块下采样 $ch \rightarrow 2ch$
残差块 $2ch \rightarrow 2ch$
下采样残差块 $2ch \rightarrow 4ch$
残差块 $4ch \rightarrow 4ch$
下采样残差块 $4ch \rightarrow 8ch$
残差块 $8ch \rightarrow 8ch$
下采样残差块 $8ch \rightarrow 8ch$
残差块 $8ch \rightarrow 8ch$
下采样残差块 $8ch \rightarrow 16ch$
残差块 $16ch \rightarrow 16ch$
下采样残差块 $16ch \rightarrow 16ch$
残差块 $16ch \rightarrow 16ch$
ReLU，全局求和池化
自注意力 [投影(嵌入(y) ⊮线性层(s))], 归一化(线性层512$\rightarrow$ 1024)
条件对比损失

(b) 判别器 表1：生成器和判别器的架构。“ch”表示乘法通道数，“残差块*”表示 BigGAN-deep 中引入的残差块，“批归一化”表示批归一化操作，“ReLU”表示修正线性单元，“归一化”表示归一化算子，“投影”和“线性层”表示线性投影或多层感知机。

4 模型验证

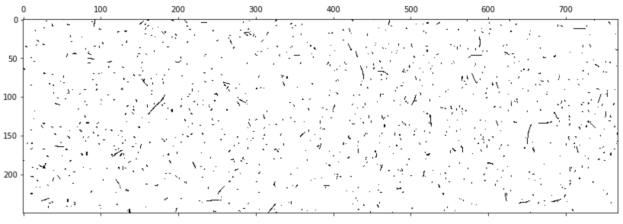
模型评估方法分为两类：一类是在训练过程中使用的方法，另一类是在训练后用于验证的方法。

1. 训练期间：
   - 图像可视化（图灵测试）
   - 监测权重矩阵的奇异值
2. 训练后：
   - 图像像素强度分析
   - 传感器数量占用分布
   - 对撞击参数分辨率进行物理分析

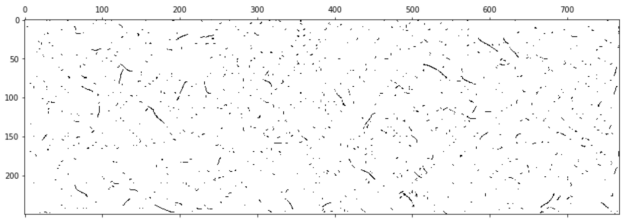
4.1 训练评估方法

图1展示了内层某个传感器的生成命中图与实数命中图的对比。可以为任意层和梯级编号生成依赖于传感器的图像。模型通过首次定性评估测试非常重要。模糊度、图像保真度、总体聚类细节和清晰度都是可以检查的视觉因素。毋庸置疑，我们也在开发这种评估的自动化定量版本，并将在展望部分中讨论。

关于不同层的权重矩阵的第一个奇异值，我们的观察与 [12] 一致：训练期间，当判别器的损失持续接近零时，这是由坍塌造成的。然而，我们还观察到，判别器的谱非常嘈杂；发生模式坍塌或判别器



（a）实数



（b）生成图像

图 1：颜色反转的 PXD 图像（上）与生成图像（下）的比较

记忆化时，奇异值会骤然下降。最重要的是观察其绝对变化。图 2 展示了模型各层权重矩阵的第一个奇异值。可以使用 Arnoldi 迭代法 [25] 高效地计算这些值。第一次大幅下降及随之而来的坍塌发生在约第 25000 次迭代时，图 2a 中绿色和橙色的层可以看出这一点。因此，图像内容逐渐变得完全荒谬。为了解决这个问题并延缓坍塌，我们尝试并评估了多种方法，例如对生成器和判别器使用 Dropout、正交正则化、谱正则化 [26] 以及梯度惩罚 [27]。然而，这些方法在稳定性和持久性方面带来的性能提升并不理想。之后，我们对 BigGAN-deep 模型进行了修改，在判别器中加入对比条件化，最终实现了良好的稳定性。

4.2 验证方法

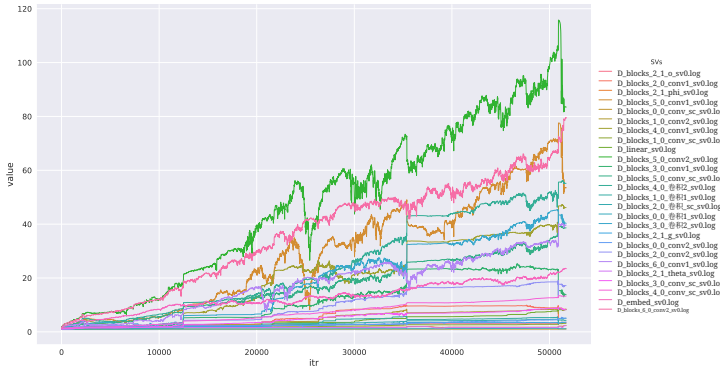
在训练后验证过程中，我们使用了 10000 个新的模拟事件。首先，我们观察了像素强度值。像素值 0 表示完全黑暗。根据蒙特卡罗模拟，如图 3 所示，0 到 7 ADU 之间存在一个间隙。对于低于像素强度 7 的值，我们使用一个函数来截除残余像素值。这样，我们便能恢复图中两部分之间的分隔。不过，我们会损失许多像素值，而这些值原则上对应于图像中的某个特定背景过程。此外，我们尚未完全学会强度曲线的非线性特征。我们将在讨论部分处理这一问题。

我们用于验证的另一项指标是图 4 中蒙特卡罗模拟测试集与生成图像之间的平均占有率频率比较。这是考虑类别条件生成对抗网络模型中的传感器 ID 所带来的结果。通过考虑标签，模型部分学会了各个传感器上的 PXD 击中频率，这一点非常重要。随后，这表明模型生成的命中图更加接近真实情况，且具有更好的多样性和保真度。通过进一步调整模型，可以缩小真实命中图与生成命中图之间的平均占有率差异。我们在第 5 节提出了更多解决方案。

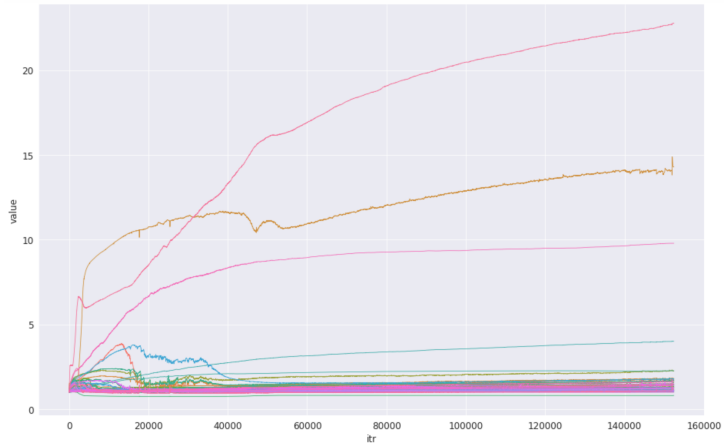
最后，我们采用的最后一项指标是通过研究冲击参数分辨率，对生成的命中图与模拟的命中图进行定量比较。为此，我们模拟了 $\Upsilon(4S)$ 个信号事件，并考虑了三种背景叠加情景进行比较：

- PXD 中无本底
- Geant4 模拟的 PXD 本底
- 由改进的 BigGAN-deep 生成的 PXD 本底

PXD 中本底的主要影响是将错误击中附加到径迹上。因此，本底的影响在分发的尾部清晰可见。Geant4 模拟结果与条件生成对抗网络之间的差异很小，而我们的模型成功再现了 Geant4 模拟的本底影响。为了量化分辨率之间的差异，我们引入二点Wasserstein距离（Fréchet距离） [28]，它是一种



（a）坍塌的模型



（b）我们的模型

图 2：判别器各层的首批奇异值：(a) 发生坍塌的模型；(b) 我们的模型。这里最值得注意的是，在一个绝对值较大的范围内，值骤然下降，这表明判别器出现了过拟合。

用于高斯分布的相似性度量。对于均值为 μ_x 和 μ_y、协方差矩阵为 Σ_x 和 Σ_y的两个多元高斯分布，Fréchet距离定义为

$$W^2(x,y) = |\mu_x - \mu_y|^2 + \mathrm{Tr}(\Sigma_x + \Sigma_y - 2\sqrt{(\Sigma_x\Sigma_y)}). \tag{1}$$

观察分发之间的两点瓦瑟斯坦距离函数可知，Fréchet距离（FD）越低，从物理学角度来看图像质量就越好。由于对传感器 ID 进行了条件化，与先前的模型相比，Fréchet距离得分进一步降低。

模型	Geant4 与无。	生成对抗网络与无。	生成对抗网络与 Geant4。
改进的 BigGAN-deep	5.46e−3	**6.70e−3**	**2.51e−3**
WGAN-gp	5.46e−3	6。98e−3	2.82e−3

表 2：三种背景场景下模型之间的 FD 分数比较。

5 总结与展望

我们开发了一种基于 PXD 图像及其传感器 ID 的类别条件生成对抗网络。因此，我们实现了与真实数据更合理的一致性，并且现在能够生成依赖传感器的 PXD 背景命中图。使用生成对抗网络生成背景数据后，我们将以增加 CPU 资源为代价，减少存储消耗

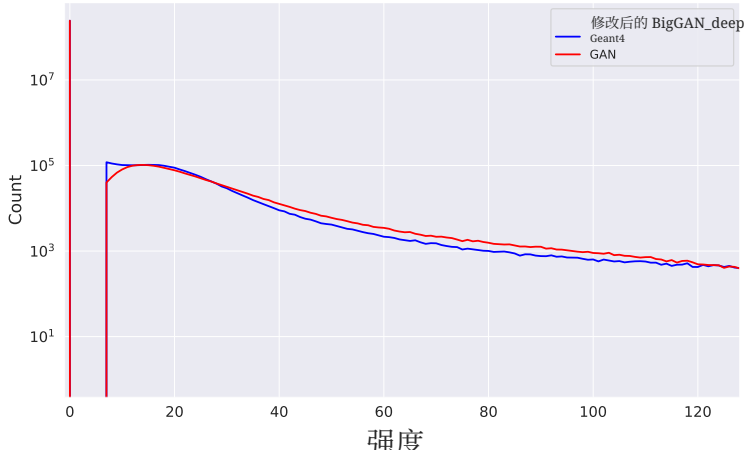


图 3：像素强度值对比。显然，图像中有很大一部分完全是黑的。Geant4 模拟生成的像素值与生成对抗网络生成的像素值之间存在非常微小的差异。

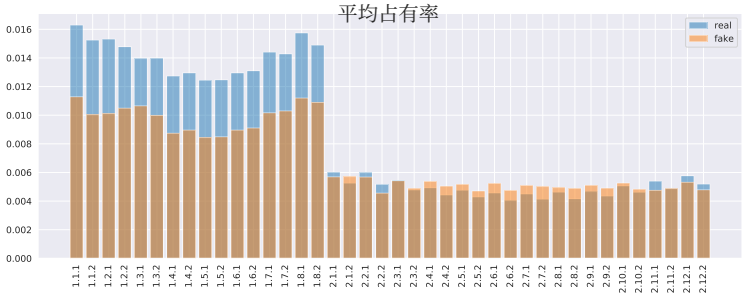


图 4：传感器　Geant4 模拟图像与生成对抗网络生成图像的相关平均占有率对比　s

。在这一过程中，我们面临着许多挑战，例如严重的模式崩溃和记忆化问题；由于缺乏具体的训练评估方法，这些问题在我们的案例中更加难以解决。我们还针对条件图像生成任务中的许多最先进的方法进行了广泛的消融研究，并开发了一种稳定的改进的生成对抗网络模型。

我们仍在尝试开发方法来解决遗留问题。例如，一种想法是将像素强度值的先验信息加入生成器的损失函数中，类似于 L2 正则化方案，以惩罚每个批次中的错误强度分布。换句话说，通过最小化损失，该模型也会最小化生成器常规输出与经过像素强度截断后的生成器输出之间的 L2 距离。理论上，这会比硬截断方法更有效。我们认为，如果仅在后处理步骤中进行截断，只将像素值置黑，就会丢失图像中的基于位置的数据。因为，尽管低于 7 ADU 的像素值不正确，但它们对应着可能存在的物理径迹；原则上，即使强度较低，我们也需要保留这些径迹。

另一个值得尝试的方向是改进生成对抗网络的设置，以捕捉 PXD 探测器两层之间的绝对关系，类似于使用基于序列的生成对抗网络模型生成动态图像。此外，鉴于我们迫切需要可靠的训练指标，可以构建一种自定义的 Inception Score（IS） [13], ：使用基于模拟事件和图像训练的图像分类模型来预测 PXD 传感器 ID。原始 Inception Score 基于预训练的 Inception v3 模型 [29] ，并利用该模型对大量生成的伪造图像进行分类，从而预测图像属于各个类别的概率。随后将这些预测汇总为 Inception 分数，以实现完全自动化的训练评估指标。此外，要全面验证生成的命中图，下一步应估算使用生成对抗网络对径迹重建效率、假信号率和分辨率造成的系统不确定度。如果该不确定度与其他系统不确定度相比足够小，我们就可以使用该模型。最后，随着 PXD 本底生成被集成到 Belle II 软件 [30], 中，我们将获得许多新经验，这些经验也可能对其他实验非常有用。

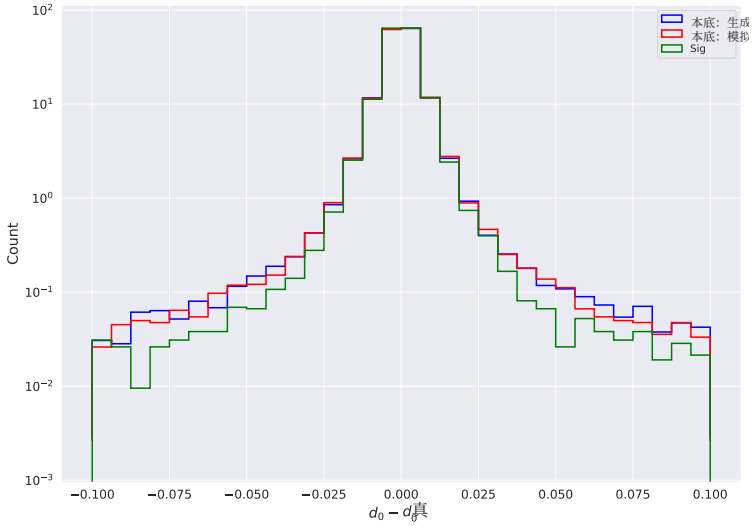


图 5：比较从枢轴点到螺旋线的横向平面有符号距离的分辨率，包括模拟本底、生成背景和无本底三种情况

参考文献

[1] T. Abe（Belle II合作组），1011.0352，（2010）[2] K. Akai、K. Furukawa、H. Koiso（SuperKEKB），Nucl. Instrum. Meth.**A907**，188（2018）[3] S. Agostinelli 等（GEANT4），Nucl.Instrum.Meth.**A506**，250（2003）[4] F. Mueller（DEPFET），JINST**9**，C10007（2014）[5] J. Kemmer、G. Lutz，Nucl. Instrum. Meth.**A253**，365（1987）[6] Goodfellow, Ian J.、Jean Pouget-Abadie、Mehdi Mirza、Bing Xu、David Warde-Farley、Sherjil Ozair、Aaron Courville 和 Yoshua Bengio。1406.2661（2014）[7] de Oliveira, Luke、Michela Paganini 和 Benjamin Nachman。《大科学计算与软件》，1–24（2017）[8] Di Sipio, Riccardo 等。《高能物理学杂志》**8**，1–17（2019）。[9] Luke Metz、Ben Poole、David Pfau 和 Jascha Sohl-Dickstein，《国际学习表征会议》，第 1–25 页（2017）[10] Jaritz, Maximilian、RaoulDe Charette、Emilie Wirbel、XavierPerrotton 和 Fawzi Nashashibi。2018年国际三维视觉会议（3DV），第 52–60 页（2018）[11] Deng, Jia、Wei Dong、Richard Socher、Li-Jia Li、Kai Li 和 Li Fei-Fei。第 248–255 页（2009）。[12] Brock, Andrew、Jeff Donahue 和 Karen Simonyan，1809.11096（2018）[13] Salimans, Tim 等，1606.03498，（2016）[14] Heusel, Martin、Hubert Ramsauer、Thomas Unterthiner、Bernhard Nessler 和 Sepp Hochreiter。arXiv 预印本 arXiv: 1706.08500（2017）[15] Zhang, Han、Ian Goodfellow、Dimitris Metaxas 和 Augustus Odena。《国际机器学习会议》，第 7354–7363 页（2019）[16] Kang, Minguk 和 Jaesik Park。2006.12681（2020）[17] Vaswani, Ashish 等。arXiv 预印本 arXiv:1706.03762（2017）[18] Zhang, Han、Zizhao Zhang、Augustus Odena 和 Honglak Lee。1910.12027（2019）[19] Paszke, Adam 等。arXiv 预印本 arXiv:1912.01703（2019）[20] Srebre, Matej、Pascal Schmolz、Baran Hashemi、Martin Ritter 和 Thomas Kuhr。《EPJ会议网络》**245**，第 02010 页（2020）

[21] Wu、Yan、Jeff Donahue、David Balduzzi、Karen Simonyan 和 Timothy Lillicrap，1912.00953（2019）[22] Zhao、Yang、Chunyuan Li、Ping Yu、Jianfeng Gao 和 Changyou Chen，2004.02088（2020）[23] Miyato、Takeru、Toshiki Kataoka、Masanori Koyama 和 Yuichi Yoshida，1802.05957（2018）[24] Saxe、Andrew M.、James L. McClelland 和 Surya Ganguli。1312.6120（2013）[25] Gene Golub 和 Henk Van der Vorst。《计算与应用数学杂志》**123**，35–65（2000）[26] Liu、Kanglin、Wenming Tang、Fei Zhou 和 Guoping Qiu。IEEE/CVF国际计算机视觉会议论文集，第6382-6390页（2019）[27] Mescheder、Lars、Andreas Geiger 和 Sebastian Nowozin。国际机器学习会议，第3481-3490页（2018）[28] Agueh、Martial 和 Guillaume Carlier。《SIAM数学分析杂志》**43**，第2期，第904-924页（2011）[29] Szegedy、Christian、Vincent Vanhoucke、Sergey Ioffe、Jon Shlens 和 Zbigniew Wojna。IEEE计算机视觉与模式识别会议论文集，第2818-2826页（2016）[30] T. Kuhr、C. Pulvermacher、M. Ritter、T. Hauth、N. Braun（Belle-II框架软件组），《计算软件与大科学》**3**，1（2019）