# CSI-MAE：一种基于掩码自编码器的信道基础模型

Jun Jiang*‡、Xiaolong Ruan†‡ 和 Shugong Xu†§
*上海大学通信与信息工程学院，中国上海
†西交利物浦大学，中国江苏 电子邮箱：jun jiang@shu.edu.cn、
Xiaolong.Ruan25@student.sjtlu.edu.cn、shugong.xu@xjtlu.edu.cn
‡ 共同第一作者 § 通讯作者

arXiv:2601.03789v1 [eess.SP] 7 Jan 2026

摘要——**自监督学习（SSL）已成为机器学习中的一项关键技术，能够应对标注数据有限、标注成本高以及无线信道条件多变等挑战。它对于开发信道基础模型（CFM）至关重要；这类模型可从信道状态信息（CSI）中提取潜在特征，并适应不同的无线环境。然而，现有信道基础模型存在明显不足：过度依赖特定场景的数据，限制了泛化能力；关注的任务类型单一或仅涉及双任务；且缺乏零样本学习能力。本文提出 CSI-MAE，这是一种利用掩码自编码器实现跨场景泛化的通用信道基础模型。该模型基于 3GPP 信道模型数据集进行训练，通过 CSI 感知与生成整合感知和通信，并已证实在多种任务中有效。轻量级解码器微调策略降低了训练成本，同时保持了有竞争力的性能。在这一策略下，CSI-MAE 的表现可与监督模型媲美或更胜一筹。通过全参数微调，它实现了最先进的性能。其卓越的零样本迁移能力在跨场景应用中也可与监督技术相媲美，从而推动无线通信创新。**

关键词—**信道基础模型、自监督学习、定位、信道外推、信道反馈、感知与通信一体化（ISAC）**

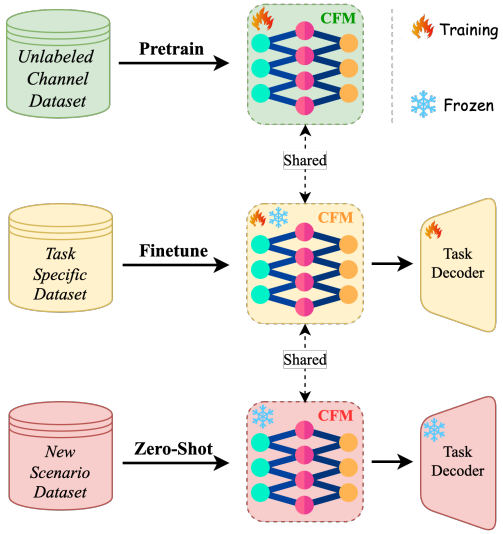



图 1. 信道基础模型的整体框架。

## I. 引言

自监督学习（SSL）已成为机器学习中的关键范式，并在人工智能（AI）领域中确立了基石地位 [1]。在快速发展的无线通信领域，传统监督学习方法正面临严峻挑战。这些挑战主要源于标注数据匮乏、标注成本高昂以及场景泛化能力有限，所有这些因素都阻碍了无线通信系统中稳健且适应性强的模型的发展。自监督学习提供了一种很有前景的解决方案，使模型能够从未标注数据中提取知识，并有效利用信道状态信息（CSI）中蕴含的丰富信息。

这些自监督学习范式在跨不同领域学习通用表征方面的有效性，促使研究人员探索其在专业领域中的适用性。具体而言，无线通信中独特的挑战和数据特性，例如多样的场景和任务，为运用自监督学习提供了理想条件。这使得人们对信道基础模型的兴趣日益浓厚（CFM） [2], [3],旨在发现通用的信道表征。如图 1 所示，信道基础模型通过预训练学习具有迁移性和泛化能力的信道状态信息表征。这种方法显著提升了跨任务和跨场景的泛化能力以及数据效率。

自监督学习旨在从大规模未标注数据中提取具有泛化性和可迁移性的表征，其成功可归因于两种主要范式：掩码X建模（MXM）和对比学习（CL）。MXM 技术依赖重构推理，通过训练模型从不完整数据中预测缺失部分 [4]–[6]。这种设计能够有效捕捉语义和结构先验，有助于生成具备上下文感知能力的表征。例如，在自然语言处理领域，BERT [4] 采用双向自注意力机制来预测掩码标记；而在计算机视觉领域，MAE [5]则在预训练阶段利用非对称编码器-解码器架构重构掩码图像块。相较之下，对比学习（CL）方法（如 MoCo [7] 和 SimCLR [8], ）侧重于判别式特征学习。这些方法会最大化同一数据实例的增强视图之间的相似度

同时最小化其与其他实例之间的相似度，从而生成鲁棒、域不变的表征，提升泛化能力，且无需明确的监督。值得注意的是，近期研究 [9], [10] 将 MXM 与 CL 结合，以发挥二者互补的优势。这些混合方法将 MXM 的上下文重构能力与 CL 的判别能力相结合，既通过重构实现细粒度语义捕捉，又增强实例判别能力。由此产生的表征具有丰富的上下文信息和很强的判别性，能够提升各种下游任务的表现，并增强自监督模型的泛化能力。

多项开创性研究 [11]–[15] 采用掩码信道建模（MCM），为信道基础模型（CFM）的发展奠定了基础。LWM [11]引入了首个信道感知基础模型，专注于信道识别和 6 GHz 以下至毫米波的波束预测。BERT4MIMO [12]将 BERT架构应用于大规模多输入多输出系统，以实现高维信道状态信息重建。WiMAE [13] 融入了 MAE 的非对称编码器-解码器设计，以促进高效且稳健的信道表示学习。WiFo [14] 提出了统一的时空频信道基础模型，用于跨配置信道预测，无需针对特定任务进行微调。

尽管取得了这些成果，现有信道基础模型仍面临三项重大局限，这些局限与无线通信领域更广泛的挑战相一致：

1）依赖特定场景数据集。大多数模型使用从特定环境中采集的射线追踪数据（例如 DeepMIMO [16], SionnaRT [17]）。与 3GPP TR 38.901 [18], 规定的标准化统计信道模型相比，这导致模型的泛化能力有限；后者包括农村宏蜂窝（RMa）、城市宏蜂窝（UMa）和城市微蜂窝（UMi）模型，涵盖了广泛的现实世界传播条件。

2）任务通用性不足。当前信道基础模型主要针对生成类任务（如信道重建/预测）或感知类任务（如用户定位）进行设计。现有方法缺乏统一框架，无法同时预测信道状态信息并将其用于环境感知，因而限制了无线系统的整体优化。3）资源需求高且迁移能力有限。训练和微调这些模型需要大量计算资源。此外，缺乏零样本学习能力，使其难以在没有额外数据的情况下适应新场景。

为解决上述局限性，我们提出了一种新型信道基础模型 CSI-MAE。我们的方法首先模拟一个大规模数据集，该数据集涵盖多种配置，并纳入各种传播模型用于预训练。通过在预训练过程中采用掩码信道建模策略，CSI-MAE 学会了无线信道的稳健表征。此外，通过冻结预训练模型并仅微调一个轻量级任务解码器，CSI-MAE 实现了具有竞争力的性能，同时显著减少了计算资源消耗。值得注意的是，该模型在不同频率下展现出出色的零样本性能。

本文的贡献可总结如下：

1）**通用信道基础模型**：我们提出了 CSI-MAE，这是一种基于标准化 3GPP 统计信道模型生成的数据集进行训练的信道基础模型。该方法摆脱了对特定场景数据的依赖，实现了可复用性和跨场景泛化。

2）**面向感知与通信一体化的通用表征**：CSI-MAE 支持同步进行 CSI感知与生成，并提取具有高度泛化能力的特征。通过在信道外推、反馈和用户定位任务上的严格验证，证明了其在无线信道处理中的有效性。

3)**参数高效且可迁移的适配**：通过仅微调轻量级解码器，CSI-MAE 在显著降低训练成本的同时，取得了与监督模型相当或更优的性能。在全参数微调下，我们的模型超越了以往方法，达到最先进水平，并显著提升了准确率。此外，CSI-MAE 展现出卓越的零样本迁移能力，在跨场景任务中表现出与监督方法相当的能力。

## II. 提出的框架

### *A.* 模型架构

CSI-MAE 遵循 MAE [5] 框架，并针对 CSI 数据的结构特征进行了调整。如图2所示，它采用非对称编码器–解码器架构，其中 CSI 输入被视为由实部和虚部构成的双通道图像。基于视觉Transformer (ViT) 的 [19] 编码器仅处理部分可见图像块，而轻量级解码器则根据潜在特征和掩码标记重建完整信道。这种非对称设计确保了自监督预训练期间的高计算效率和可扩展性。

**掩码策略**。 输入的复值CSI矩阵$H$定义在天线域和子载波域上，被划分为互不重叠的图像块$\mathbf{h}_i$，并序列化为一维标记序列。通过均匀随机采样施加较高的掩码率（75%），仅让25%的可见图像块进入编码器。这种设计促使模型推断信道的全局结构，而不是依赖局部插值。较低的掩码率可能会让模型找到捷径，使其只需通过插值即可恢复矩阵，而无需学习通用特征。

**位置嵌入**。位置嵌入 $P_{\text{emb}}$保留了信道矩阵的二维空间相关结构。具体而言，对于沿第 $t$ 维的每个空间索引 $(i, j)$，嵌入分量定义如下：

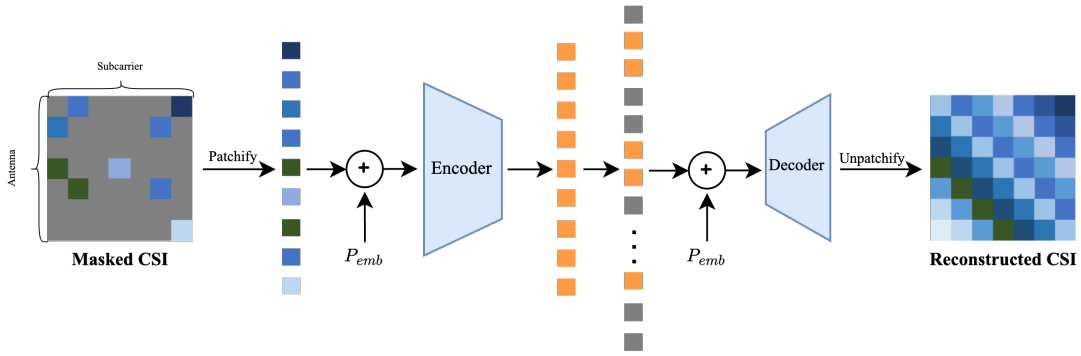



图2. 我们提出的 CSI-MAE 架构。

$$P^t_{\text{emb}}[i, 2j] = \sin\left(\frac{\text{pos}^t_i}{10000^{\frac{2j}{D^t_{\text{emb}}}}}\right), \tag{1}$$

$$P^t_{\text{emb}}[i, 2j+1] = \cos\left(\frac{\text{pos}^t_i}{10000^{\frac{2j}{D^t_{\text{emb}}}}}\right), \tag{2}$$

其中，$\text{pos}^t_i$ 表示离散空间索引，$D^t_{\text{emb}}$ 表示轴 $t$上的嵌入维度。在经过 Transformer 编码器和解码器层处理之前，这些固定嵌入会逐元素添加到图块序列中。

**CSI-MAE 编码器**。编码器采用标准视觉 Transformer架构处理可见图像块序列。为保留天线和子载波之间的空间关系，我们将固定的二维正弦–余弦位置嵌入融入图块嵌入，使编码器能够捕获 $H$ 中的空频依赖关系。此外，在图块嵌入序列前添加一个可学习的类别标记 [CLS] 。这一特殊标记作为整个输入序列的全局表示，可用于后续的分类等任务，也可作为编码信息的摘要。编码后的潜在表示可表示为：

$$H_{\text{emb}} = f_{\text{enc}}\left([\text{CLS}; H_{\text{vis}}] + P_{\text{emb}}\right), \tag{3}$$

其中，$H_{\text{vis}}$ 表示可见图像块的嵌入，$f_{\text{enc}}(\cdot)$表示编码器，$[\text{CLS}; H_{\text{vis}}]$ 表示将可学习的类别标记与可见图像块嵌入拼接。

**CSI-MAE 解码器**。 解码器利用潜在表征以及放置在缺失位置的共享可学习掩码标记，重建被掩码的信道图块。位置嵌入在这里也发挥着至关重要的作用。它们帮助解码器理解序列中各元素的相对位置，这对于准确重建被掩码图块至关重要。没有位置嵌入，解码器就缺少恢复原始信道值所需的空间上下文。与编码器类似，类别标记也包含在解码器的输入中。这使解码器能够全面掌握输入序列，并在重建过程中利用类别标记承载的全局信息。

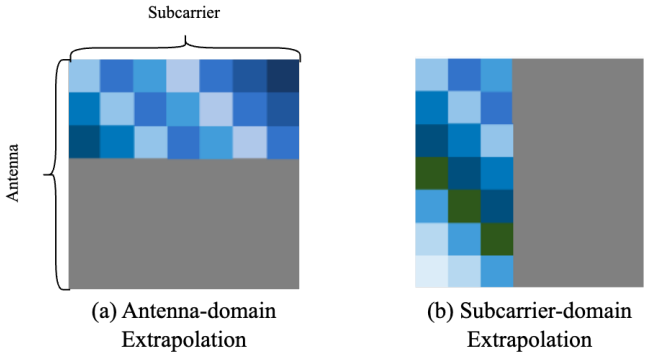



图3. 天线域和子载波域外推的输入。

重新加入位置嵌入后，完整序列会由轻量级 Transformer 解码器处理，随后通过线性投影头预测原始信道值。解码器比编码器更浅、更窄，侧重于局部重建，而非全局语义。

**训练目标**。模型通过最小化仅针对掩码图块计算的均方误差（MSE）进行优化，从而确保未掩码区域不参与损失函数计算。MSE 定义如下：

$$\mathcal{L}_{\text{MSE}} = \frac{1}{N}\sum_{i=1}^{N} \|\mathbf{h}_i - \hat{\mathbf{h}}_i\|^2, \tag{4}$$

其中 $N$ 表示掩码图块的数量，$\mathbf{h}_i$ ˆ和$\mathbf{h}_i$ 分别表示输入和重建的信道图块。该掩码重建目标促使模型学习有意义的潜在表示，以捕捉信道状态信息固有的空频相关性。

### *B.* 下游任务

在下游任务的微调阶段，我们利用迁移学习来挖掘预训练模型中蕴含的丰富表征，从而提升任务性能。具体而言，预训练的 CSI-MAE编码器已经具备从 CSI数据中提取通用特征的能力。这些特征在各种基于 CSI 的应用中都展现出广泛的适用性。通过附加一个

轻量级任务解码器，并随后使用标注数据进行微调，CSI-MAE 可以针对特定任务进行优化。

**信道外推与反馈**。信道外推旨在利用部分信道状态信息预测未知的信道状态信息，从而解决实际通信场景中信道状态信息获取不完整的问题，并提升系统性能。具体而言，我们考虑天线域和子载波域中的外推。如图 3 所示，我们采用 50% 的掩码率来模拟待预测的未知信道状态信息。另一方面，信道反馈会先压缩输入的信道状态信息，再对其进行重建。

信道外推和反馈都被建模为信道重建任务，目标是根据部分或压缩后的观测恢复完整的信道矩阵。在外推中，只有部分 $H$ 可见；而在反馈中，完整的 $H$ 会先被编码为紧凑的潜在表示，以便高效传输。在这两种情况下，重建都遵循统一的前向过程：

$$\hat{H} = f_{\text{dec}}(f_{\text{enc}}(H_{\text{in}})), \tag{5}$$

其中，$H_{\text{in}}$ 表示输入，$f_{\text{dec}}(\cdot)$是任务解码器。学习目标是尽可能减小预测结果与真实值之间的重建差异。

**用户定位**。用户定位被定义为一项回归任务，用于估计用户设备的空间坐标。CSI-MAE 从输入的信道状态信息中提取潜在表示，并通过一个轻量级线性回归头，根据作为全局表示的类别标记 [CLS], 预测二维坐标：

$$\hat{\mathbf{p}} = W_{\text{pos}} \cdot f_{\text{enc}}(H_{\text{in}})_{[\text{CLS}]} + b_{\text{pos}}, \tag{6}$$

其中，$W_{\text{pos}}$ 和 $b_{\text{pos}}$ 是可学习参数。模型通过均方误差损失进行优化。

## III. 实验

### A. 数据集

该数据集基于 3GPP TR 38.901 [18] 配置，使用 Sionna [17] 生成，涵盖三种具有代表性的传播环境，如表I所示。仿真涵盖五种载波频率和三种子载波间隔。生成的数据集总计约含 145 万个信道状态信息样本，可为评估提供不同环境、频率和移动性配置下的多样化信道条件。所有数据均用于预训练。选取五种环境与载波频率的组合，即 RMa-0.7、RMa-2.4、RMa-3.5、UMa-4.9 和 UMi-5，作为测试场景。

### B. 结果分析

在预训练阶段，我们采用针对双通道输入进行调整的 ViT-Base [19] 架构作为 CSI-MAE 编码器。归一化均方误差（NMSE）用作主要评估指标。为提供全面的比较框架，实验设置中纳入了一个监督式基线模型（监督式）。该基线

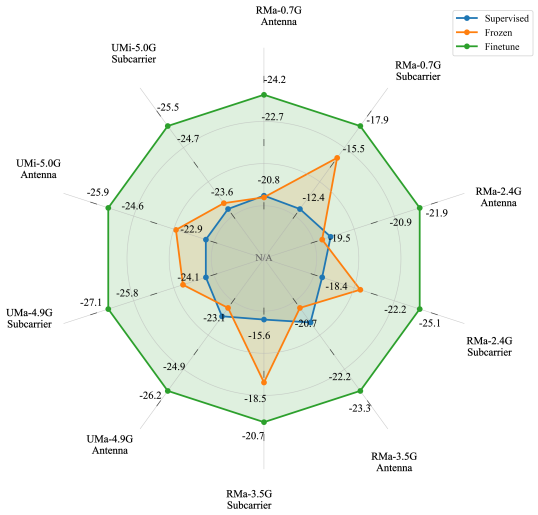



图4. 该图比较了不同场景以及天线/子载波域外推任务中的性能。

模型采用相同的网络架构骨干，但训练时不利用任何预训练。我们探索了两种迁移学习方法。在冻结 方法中，我们冻结 CSI-MAE 编码器，仅训练一个轻量级任务解码器。相反，微调方法则针对目标数据集对整个模型进行端到端调整。

**信道外推**。 图4显示，在所有五种场景的信道外推任务中，我们的 CSI-MAE 显著优于监督基线。尽管两种方法采用完全相同的编码器架构，其核心差异在于训练范式。监督式模型仅使用各场景的本地数据进行训练，因此学到的表征具有针对性，但也较为有限。相比之下，CSI-MAE 编码器在海量数据上进行预训练，获得了通用的 CSI表征，从而在多种场景下具备更强的泛化能力。

这一点也得到了冻结 实验的进一步证明。即使编码器被冻结、仅训练一个轻量级变换器解码器，该模型仍能达到与监督基线相当甚至更优的性能。这一结果有力地验证了所学习通用特征的强大能力，并凸显了 CSI-MAE 的实际价值，尤其适用于数据受限场景中的应用。

表I 仿真参数设置

| 参数 | 数值 | | |
|---|---|---|---|
| 信道模型 | UMi | UMa | RMa |
| 载波频率（GHz） | 0.7、2.4、3.5、4.9、5 | | |
| 子载波间隔（KHz） | 15, 30, 60 | | |
| 子载波数量 | 256 | | |
| 基站天线 | 8 × 8 | | |
| 基站高度（m） | 10 | 25 | 35 |
| 用户设备天线 | 2 × 2 | | |
| 用户设备高度（m） | 1.5 | | |
| 用户设备速度（m/s） | [0, 27.78] | | |

表 II CSI-MAE 信道反馈性能比较。

| 方法 | 归一化均方误差 (dB) | | | | |
|---|---|---|---|---|---|
| | RMa-0.7 | RMa-2.4 | RMa-3.5 | UMa-4.9 | UMi-5 |
| 监督式 | -35.14 | -43.55 | -38.73 | -44.08 | -48.86 |
| 冻结 | -36.18 | -44.07 | -40.83 | -44.88 | -49.01 |
| 微调 | **-37.11** | **-48.25** | **-43.31** | **-46.52** | **-50.76** |

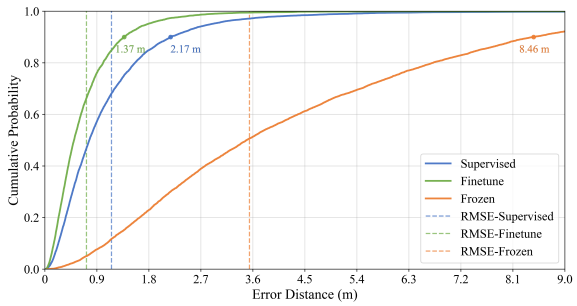



图 5. CSI-MAE 在累积分布函数中的定位性能比较。

**信道反馈**。 如表 II 所示，我们的 CSI-MAE 在信道反馈任务上同样展现出卓越性能。与信道外推类似，该任务本质上是信道重建。因此，在预训练阶段学习到的通用 CSI 特征非常有效。这一点从 冻结实验中可以得到证明：尽管编码器被冻结，该模型仍在所有场景下持续优于监督基线。通过微调整个模型，微调 进一步提升了性能，取得了最佳结果。这再次验证了 CSI-MAE 为下游任务提供稳健起点的能力，证明了其在不同任务上的强大泛化能力。

**用户定位**。 如图 5 所示，定位结果揭示了有关 CSI-MAE 的关键见解。微调配置表现出色，90%的误差在 1.37 m 以内，均方根误差为 0.718 m，优于监督基线（2.17 m，1.151 m）。相比之下，冻结 配置表现不佳，90%的误差低于 8.46 m，均方根误差为 3.539 m。这种差异源于两个因素：首先，用于用户定位任务的解码器仅为单个线性层，学习能力有限。更重要的是，与掩码信道建模相比，定位属于不同的回归问题。它要求模型将高维 CSI 数据映射到低维空间坐标，因此需要更高层次的语义信息，而这些信息不同于模型在预训练期间为结构重建所学习的信息。

此外，我们在表III中将 MDA [20], LWM [11], CSI-CLIP [2],和 CSI-MAE 与 LWM 提出的基于 DeepMIMO的数据集 [16] 进行了比较。鉴于不同数据集在发射和接收天线配置、子载波等方面存在差异，我们采用双线性插值来对齐预训练的输入尺寸。结果表明，即使仅冻结，CSI-MAE 的表现也优于现有的 LWM。它

表III 用户定位任务中最先进方法的性能比较（平均距离误差（m））。

| 方法 | 场景 | | | |
|---|---|---|---|---|
| | 芝加哥 | 圣迭戈 | 夏洛特 | 室外 |
| MDA [20] | 48.75 | 285.90 | 295.45 | 3.75 |
| LWM [11] | 69.45 | 381.58 | 788.11 | 16.84 |
| CSI-CLIP [2] | **3.95** | **6.79** | 7.75 | **0.45** |
| CSI-MAE（冻结） | 14.29 | 15.35 | 20.69 | 0.93 |
| CSI-MAE（微调） | 5.11 | 9.59 | **7.18** | **0.45** |

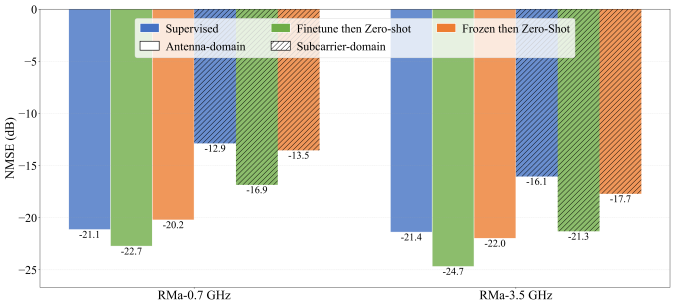



图6. 信道外推的零样本性能对比。模型在 RMa-2.4 上进行微调，并在 RMa-0.7 和 RMa-3.5 上进行测试。

甚至超越了专为用户定位设计的 MDA。通过微调，性能进一步提升，平均误差降低了 10 倍以上。然而，在三个场景中，基于 DeepMIMO 预训练的 CSI-CLIP 表现最佳。

### *C.* 零样本性能

为进一步验证 CSI-MAE 的泛化能力，我们进行了零样本实验。我们将仅在 RMa-2.4 场景上微调的模型直接应用于 RMa-0.7 和 RMa-3.5 场景，未进行任何额外训练。如图6所示，我们的 CSI-MAE 显著优于在目标数据上专门训练的监督基线。这一出色的零样本性能凸显了模型学习通用、可迁移 CSI 特征的能力，使其能够有效适应全新的环境和频段。

### *D.* 缩放定律

缩放定律对于提升模型性能和泛化能力至关重要。在我们的研究中，理解其对 CSI-MAE框架的影响是实现优化的关键。为全面验证 CSI-MAE 的可扩展性，我们开展了数据和模型缩放实验。

**数据缩放**。我们在一个部分数据集上对模型进行预训练，该数据集包含 0.48 M 个样本，且仅限于单一子载波间隔，并用于信道外推任务。如表IV所示，在更多样化的数据上预训练的 CSI-MAE 性能更优，尤其是在天线域外推任务中采用冻结设置时，性能提升最高达 62.7%。这表明，大规模预训练数据使模型能够学习到泛化能力更强的信道表征，展现了该框架出色的数据可扩展性。

表IV CSI-MAE 在部分数据集和完整数据集上预训练后的信道外推性能比较

| 领域 | 方法 | 归一化均方误差（dB） | | |
|---|---|---|---|---|
| | | 部分数据集-0.48M | 完整数据集-1.45M | Δ |
| 天线 | 冻结 | -17.43 | -21.83 | +62.7% |
| | 微调 | -23.69 | -24.31 | +12.2% |
| 子载波 | 冻结 | -18.85 | -21.04 | +31.2% |
| | 微调 | -21.41 | -23.28 | +27.3% |

[a]T改进幅度 Δ 是根据线性归一化均方误差值计算的 es.

表 V CSI-MAE 编码器用于信道外推时 ViT-Base 与 ViT-Large 的性能比较

| 领域 | 方法 | 归一化均方误差 (dB) | | |
|---|---|---|---|---|
| | | ViT-Base | ViT-Large | Δ |
| 天线 | 冻结 | -21.83 | -18.60 | -109.1% |
| | 微调 | -24.31 | -23.04 | -34.0% |
| 子载波 | 冻结 | -21.04 | -21.52 | +10.1% |
| | 微调 | -23.28 | -26.75 | +54.8% |

[a]改进幅度 Δ 是根据线性归一化均方误差值计算的。

**模型缩放**。我们比较了 CSI-MAE 编码器采用 ViT-Base 和 ViT-Large 时，在完整数据集上进行预训练后执行信道外推任务的性能。ViT-Large 的架构比 ViT-Base 更深、更宽，我们对其进行评估以考察模型容量的影响。如表 V 所示，模型缩放的效果高度依赖于数据维度。在子载波维度上，ViT-Large 表现出显著优势，其微调后的性能提升了 54.8%，这表明更大的模型能够更有效地捕捉频域的复杂性。然而，在结构性更强的天线域中，ViT-Large 的性能有所下降，降幅最高达 109.1%。这说明，对于空间结构更强的天线维度，更大的模型容量并不能带来性能优势，甚至可能因过参数化而导致性能退化。这表明当前的模型训练策略存在局限性，需要针对物理天线阵列几何结构进行专门设计。

## IV. 结论

总之，我们提出了 CSI-MAE，这是一种基于掩码自编码器、使用 3GPP 定义的统计信道模型进行训练的信道基础模型。CSI-MAE 摆脱了对特定场景数据集的依赖，实现了标准化、可复用和跨场景泛化，并能够提取泛化能力极强的特征，在多种无线信道处理任务中展现出有效性。此外，只需微调一个轻量级任务解码器，CSI-MAE 就能显著降低训练成本，同时保持与全监督模型相当或更优的性能，并展现出极强的零样本迁移能力。值得注意的是，我们从数据规模和模型规模两个维度严格验证了缩放定律。随着数据规模扩大、模型参数量增加，CSI-MAE 的性能也逐步增强。我们的工作是推动无线通信技术发展的重要一步，为利用信道基础模型提供了一种更高效、灵活且适应性更强的方法。

## 参考文献


[1] J. Gui、T. Chen、J. Zhang、Q. Cao、Z. Sun、H. Luo 和 D. Tao，“自监督学习综述：算法、应用与未来趋势”，IEEE模式分析与机器智能汇刊，第46卷，第12期，第9052–9071页，2024年。[2] J. Jiang、W. Yu、Y. Li、Y. Gao 和 S. Xu，“通过CIR-CSI一致性构建大规模多输入多输出无线信道基础模型”，载于2025年IEEE通信与网络机器学习国际会议（ICMLCN），2025年，第1–6页。[3] J. Jiang、Y. Gao、X. Wu 和 S. Xu，“迈向信道基础模型（CFM）：动机、方法与机遇”，arXiv预印本 arXiv:2507.13637，2025年。[4] J. Devlin，“BERT：深度双向变换器的预训练与语言理解”，arXiv预印本 arXiv:1810.04805，2018年。[5] K. He、X. Chen、S. Xie、Y. Li、P. Doll´ar 和 R. Girshick，“掩码自编码器是可扩展的视觉学习器”，载于IEEE/CVF计算机视觉与模式识别会议论文集，2022年，第16 000–16 009页。[6] Y. Gao、Y. Liu、R. Yu、S. Liu、Y. Jin、S. Zhang、S. Xu 和 X. Chu，“SSNet：基于自监督学习框架实现的灵活、稳健的流体天线系统信道外推”，IEEE通信领域精选期刊，2025年。[7] K. He、H. Fan、Y. Wu、S. Xie 和 R. Girshick，“用于无监督视觉表征学习的动量对比”，载于IEEE/CVF计算机视觉与模式识别会议论文集，2020年，第9729–9738页。[8] T. Chen、S. Kornblith、M. Norouzi 和 G. Hinton，“一种用于视觉表征对比学习的简单框架”，载于国际机器学习会议。PMLR，2020年，第1597–1607页。[9] M. Oquab、T. Darcet、T. Moutakanni、H. Vo、M. Szafraniec、V. Khalidov、P. Fernandez、D. Haziza、F. Massa、A. El-Nouby 等，“DINOv2：无需监督地学习稳健的视觉特征”，arXiv预印本 arXiv:2304.07193，2023年。[10] J. Zhou、C. Wei、H. Wang、W. Shen、C. Xie、A. Yuille 和 T. Kong，“iBOT：采用在线分词器的图像BERT预训练”，arXiv预印本 arXiv:2111.07832，2021年。[11] S. Alikhani、G. Charan 和 A. Alkhateeb，“大型无线模型（LWM）：无线信道基础模型”，arXiv预印本 arXiv:2411.08872，2024年。[12] F. O. Catak、M. Kuzlu 和 U. Cali，“BERT4MIMO：一种采用BERT架构进行大规模多输入多输出信道状态信息预测的基础模型”，arXiv预印本 arXiv:2501.01802，2025年。[13] B. Guler、G. Geraci 和 H. Jafarkhani，“利用对比学习和掩码自编码器学习构建无线信道表征多任务基础模型”，arXiv预印本 arXiv:2505.09160，2025年。[14] B. Liu、S. Gao、X. Liu、X. Cheng 和 L. Yang，“WIFO：用于信道预测的无线基础模型”，中国科学：信息科学，第68卷，第6期，文章编号162302，2025年。[15] T. Yang、P. Zhang、M. Zheng、Y. Shi、L. Jing、J. Huang 和 N. Li，“WirelessGPT：一种面向无线通信的生成式预训练多任务学习框架”，IEEE网络，2025年。[16] A. Alkhateeb，“DeepMIMO：面向毫米波和大规模多输入多输出应用的通用深度学习数据集”，载于信息论与应用研讨会（ITA）论文集，美国加利福尼亚州圣迭戈，2019年2月，第1–8页。[17] J. Hoydis、S. Cammerer、F. Ait Aoudia、A. Vem、N. Binder、G. Marcus 和 A. Keller，“Sionna：用于下一代物理层研究的开源库”，arXiv预印本，2022年3月。[18] 3GPP，“3GPP无线接入网工作组：0.5至100 GHz频段的信道模型研究（第18版）”，3GPP，技术报告 TR 38.901 V18.0.0（2024-03），2024年。[19] A. Dosovitskiy，“一幅图像胜过16×16个词：用于大规模图像识别的变换器”，arXiv预印本 arXiv:2010.11929，2020年。[20] L. Chu、A. Alghafis 和 A. F. Molisch，“利用深度同方差域适应，在高度动态无线网络中挖掘语义定位信息”，IEEE通信汇刊，第73卷，第3期，第2032–2046页，2025年。