

20

实验在两个域内进行（在同一数据集上进行预训练和微调）和跨域表中的TF-C和TS-TCC的下标“I”或“R”分别表示编码器变体InceptionTime或ResNet1d_wang。表中的SimMTM的下标“D”或“R”分别表示其默认编码器或ResNet1d_wang。

最佳结果以粗体突出显示，次佳结果以下划线突出显示。所有度量均以百分比（%）表示。

域内评估

在本节中，所有实验都在相同的数据集上进行，以评估模型的域内泛化能力。首先，模型在不使用标签的情况下进行自监督预训练。随后，在微调阶段引入标签来进一步优化模型。按照这种两阶段训练协议，最终在测试集上评估模型的性能。表1展示了我们的方法与基线方法的分类结果。总体而言，我们的BMIRC在宁波和PTB-XL数据集上表现优于所有基线方法，但在Chapman数据集上的表现略微逊色。随着数据集规模的扩大，我们的BMIRC相较于基线方法的性能提升也逐渐增加。例如，在PTB-XL数据集上，BMIRC相比表现次优的模型CPC提升了1.4%。同样，在宁波数据集上，与表现次优的模型CPC相比，AUPRC显著提升了9.1%。虽然BMIRC在Chapman数据集上的整体表现不及TS-TCCR，我们认为通过迁移学习可以改变这一情况（详细信息在下一节提供）。与未使用预训练权重的随机初始化（Random Init）相比，BMIRC表现出显著优势，在所有数据集的大多数指标上提升超过3%。值得注意的是，虽然基于对比学习的基线方法总体上优于基于掩码的基线方法，但它们仍未超越我们的BMIRC。我们将此现象归因于在我们的方法中引入了频率模态并对传统MAE架构进行了增强。

21

跨域评估

为了评估模型在跨域设置下的性能，在对较小数据集进行监督学习之前，使用在较大数据集上进行预训练得到的权重来初始化模型。这个过程代表了迁移学习中遇到的典型场景，用于验证预训练权重的可迁移性。我们在三种场景下进行实验：Ningbo→PTB-XL、Ningbo→Chapman和PTB-XL→Chapman，确保从较大数据集获得的预训练权重应用于较小数据集。表2显示了跨域设置的评估结果。在大多数场景中，我们的BMIRC超过基线。对于Chapman数据集，采用来自较大数据集的预训练权重显著增强了大多数模型的性能，如表1中的Chapman结果与表2中的Ningbo→Chapman、PTB-XL→Chapman结果之间的比较所证明的。此外，观察到在我们的方法中，更大的预训练数据集对应着更显著的改进。例如，我们的BMIRC的AUPRC在PTB-XL→Chapman场景中增加了4%，在Ningbo→Chapman场景中增加了12.7%，这一发现表明，我们的BMIRC可以通过在大型数据集上进行广泛的预训练提高其泛化能力。

22

使用更多下游数据集进行微调

为了进一步验证我们的模型在更多下游任务上的泛化能力（跨域微调），我们的BMIRC在格鲁吉亚和合肥数据集上与Random Init和TS-TCCR进行了比较。除了基于Ningbo数据集的预训练权重外，还使用了Ningbo、PTB-XL和Chapman数据集，以获得更强大的预训练权重。如表3所示，我们的BMIRC在这些下游任务上的表现优于基线，更大的预训练数据集与更高的性能相关。

23、24

不同比例训练集的微调

预训练模型已被证明在标签稀疏的情况下表现出增强的泛化能力【36】。在本节中，我们评估了在微调阶段使用不同比例的训练集（25%、50%、75%、100%）对模型性能的影响。如表4所示，我们展示了两个不同场景下的分类结果：宁波（域内）和宁波→Chapman（跨域）。总体而言，我们的BMIRC在大多数场景中都优于所有基线，仅在使用宁波数据集的25%时表现不及TS-TCCR。这一观察结果与Chapman数据集上的评估结果一致。值得注意的是，随着有标签数据量的增加或在跨域设置下，BMIRC持续优于TS-TCCR。我们注意到，随着训练集比例的增加，基线模型的性能逐渐提高，但性能提升的幅度逐渐放缓。我们的BMIRC在域内微调时表现出类似的趋势。然而，在跨域设置下，仅使用75%的训练集便可达到近似最优性能。此外，在这两个不同场景下，使用仅50%训练集训练的BMIRC，其性能已超过在整个数据集上训练的随机初始化模型（Random Init）。这一特性使得我们的BMIRC在下游任务标签稀疏时也能取得令人满意的结果。

25

双峰掩蔽率分析

时间和频率模态的独特特征，以及它们在双峰联合表示中占据的信息比例的差异，决定了两种模态的不同掩蔽率组合会影响模型的表示学习。我们研究了不同掩蔽率组合对模型性能的影响，并报告了ACC和AUPRC。如图5所示，我们注意到一个一致的趋势，频率模态掩蔽率的增加对应于模型性能的逐渐改善。这一观察结果由所有子图的每列内从底部到顶部的进展证明。保持频率模态中的掩蔽率恒定，经常观察到模型性能在时间模态中的50%的掩蔽率处达到其峰值。最终，最佳配置被确定为时间模态掩蔽率为50%，频率模态掩蔽率为75%。我们将这一观察结果归因于时间模态固有的复杂性，因此，过高的时间模态掩蔽率对模型有效地从时间模态中提取潜在模式提出了挑战。

26

双模态重构损失权分析

在双峰联合建模过程中，重建损失的权重可能会影响模型的性能。使用不同的权重组合进行预训练，最终的分类结果如表6所示。当 α 和 β 都设置为1.0时，模型的性能达到最佳，表明两种模态的重建损失权重同等重要。

27

心电图采样率的分析

为了验证不同输入长度对模型性能的影响，我们使用采样率为 50 Hz、100 Hz 和 200 Hz 的心电图（ECG）进行实验，分别对应输入长度为 500、1000 和 2000。正如表 7 所示，随着输入长度从 500 增加到 1000，模型性能有所提升，但更大的输入长度（2000）并未带来更好的性能。这表明，过多的数据冗余对模型学习是有害的。

额外频率模态的影响

表 8 展示了在缺少主要组件时所有场景下的实验结果。可以看到，当去除频率模态时，所有场景的性能都出现了显著下降。我们将这种现象归因于频率模态能够为时间模态提供有价值的补充信息。去除频率模态会削弱模型提取判别性表示的能力，最终导致性能下降。

双模联合编码器的效果

如表 9 所示，我们的 Specific-Shared 架构在大多数场景中优于 All Shared 架构，仅在 Ningbo→Chapman 场景中略微表现不佳。我们呈现了两种架构在所有场景下的平均排名，以全面评估它们的整体性能。Specific-Shared 架构在 ACC、F1 和 AUPRC 指标上显示出优于 All Shared 架构的表现，而在 AUROC 指标上则保持了相当的表现。与将两种模态视为一个整体的 All Shared 架构相比，Specific-Shared 架构旨在显式捕捉模态内和模态间的交互关系。实验结果强调了该分层结构在我们框架中的重要性。

IRC的影响

我们首先研究去除 IRC 对模型性能的影响。比较表 8（28页）中 w/o IRC 与 Full Model 的分类结果，可以明显看出引入 IRC 全面提升了模型的性能。

这一观察表明，为解码器提供不同层次的信息以进行重建，能够增强编码器学习高级区分性表示的能力。为了进一步探究 IRC 的影响，我们进行了专门分析，重点考察 IRC 在单一模态中的应用以及 IRC 层数（定义为连接编码器和解码器的层数）的变化影响。为便于解释，后续分析中我们采用对类别不平衡下模型性能敏感度更高的 AUPRC 指标。图 6 (a) 展示了 AUPRC 在四种情况下的结果：1) w/o IRC：无 IRC，2) IRC(Time)：仅在时间模态上使用 IRC，3) IRC(Frequency)：仅在频率模态上使用 IRC，4) IRC(All)：在时间和频率模态上都使用 IRC。

我们的研究表明，在大多数场景中，IRC 的整合都带来了性能提升，突显了其有效性。在域内设置下，IRC(All) 一贯表现优于其他选项。尽管在跨域设置下 IRC(All) 未达到最佳结果，但其性能接近最优。因此，由于其更均衡的表现，我们选择 IRC(All) 作为默认配置。在解码器深度为 4 的情况下，我们从解码器的第二层开始逐步增加 IRC 层数，用 F1、F2 和 F3 表示，其中 F3 表示完整模型。图 6 (b) 展示了这三种情况下的 AUPRC 结果。可以看出，F3 在大多数场景中取得了最佳性能，仅在 PTB-XL→Chapman 场景中表现较差。F2 的整体性能最差，我们认为这是因为第二层融合涉及从双模态共享表示转换到单模态独有表示的过程。更多单模态表示的融合（如 F3）帮助模型适应这一过程，从而提升了整体性能。

GRM的影响

表10显示了图4所示的三种融合方法在所有场景中的分类结果。在域内设置下，与竞争对手相比，我们的 GRM表现出上级的整体性能。然而，在跨域设置下，似乎没有明显的赢家。三种融合方法在所有场景中的平均排名见表10，这表明GRM独特的自适应选通融合策略可以增强模型的泛化能力。

原始灰度图像：显示原始心电图。

边缘检测图像：应用 Canny 边缘检测以突出心电图曲线的边缘。

形态学处理图像：通过形态学操作去除噪声，进一步清晰化曲线轮廓。