

从路径到全盘：通过多模态知识扩展推进降水检索

Zheng Wang

Kai Ying

zhengwang@zjut.edu.cn

221123120165@zjut.edu.cn

College of Computer Science,
Zhejiang University of Technology
Hangzhou, Zhejiang, China

Bin Xu

Chunjiao Wang

xubin@cma.gov.cn

wangcj@cma.gov.cn

National Meteorological Information
Center
Beijing, China

Cong Bai[†]

congbai@zjut.edu.cn

College of Computer Science,
Zhejiang University of Technology &
Zhejiang Key Laboratory of Visual
Information Intelligent Processing
Hangzhou, Zhejiang, China

Abstract

基于卫星的技术提高了精确的近实时降水检索。然而，由于与地面降水的关系较弱，红外算法的精度较低，而被动微波和雷达方法则更准确但范围有限。这一挑战激发了降水检索扩展（PRE）任务，其目标是在扫描观测区域之外实现精确的、基于红外的完整圆盘降水检索。我们介绍了多模态知识扩展，这是一个包含拟议的 PRE-Net 模型的两阶段流程。在扫描蒸馏阶段，PRE-Net 通过协调掩蔽和小波增强（CoMWE），将多模态数据集成模型的知识转移到扫描观测区域内的红外模型。在全圆盘适应阶段，自我掩蔽调优（Self-MaskTune）通过平衡多模态和全圆盘红外知识来细化全圆盘的预测。在引入的 PRE 基准上的实验表明，PRE-Net 显著提高了降水检索性能，超过了 PERSIANN-CCS、PDIR 和 IMERG 等领先产品。代码将在 <https://github.com/Zjut-MultimediaPlus/PRE-Net> 处提供。

CCS Concepts

• Applied computing → Earth and atmospheric sciences; • Information systems → Multimedia and multimodal retrieval.

Keywords

Precipitation retrieval expansion, Multimodal satellite data, Incomplete multimodal learning, Multimodal knowledge distillation

ACM Reference Format:

Zheng Wang, Kai Ying, Bin Xu, Chunjiao Wang, and Cong Bai[†]. 2025. 从路径到全盘：通过多模态知识扩展推进降水检索. In *Proceedings of the 31st ACM SIGKDD Conference on Knowledge Discovery and Data Mining V.2 (KDD '25)*, August 3–7, 2025, Toronto, ON, Canada. ACM, New York, NY, USA, 11 pages. <https://doi.org/10.1145/3711896.3736971>

1 引言

准确的降水测量对于水文气象学研究、水资源管理和灾害预防至关重要 [27, 34]。这些通常分为通过地面雨量计的直接观测和通过卫星的间接检索。虽然雨量计可靠，但其分布稀疏且无法捕捉海洋降水，限制了其有效性 [31]。相比之下，基于卫

[†]Corresponding author.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.
KDD '25, Toronto, ON, Canada

© 2025 Copyright held by the owner/author(s). Publication rights licensed to ACM.
ACM ISBN 979-8-4007-1454-2/2025/08
<https://doi.org/10.1145/3711896.3736971>

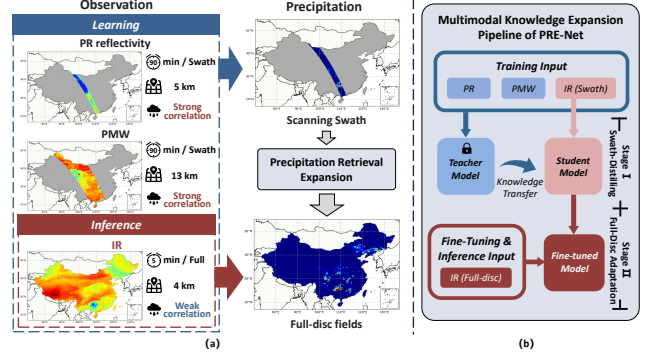


Figure 1: (a) 降水反演扩展（PRE）任务利用了强相关的微波辐射计（PMW）和降水雷达（PR）数据，以及弱相关的红外（IR）数据进行学习，而只有 IR 数据提供了完整的区域覆盖。PRE 任务利用这些异构数据实现高质量的全区域降水反演。**(b)** PRE-Net 的多模态知识扩展流程。多模态知识通过在条带蒸馏阶段的知识蒸馏进行转移，并在全区域适应阶段从扫描条带进一步扩展到全区域降水反演。

星的检索提供了更广泛的覆盖，使其成为降水估计的有前途的替代方案。

基于卫星的降水提取主要使用三种类型的传感器数据：红外（IR）、被动微波（PMW）和空间降水雷达（PR）。由地球静止轨道（GEO）卫星收集的 IR 数据提供了高时空分辨率（≤ 4 公里，5-30 分钟间隔）和接近全球的覆盖范围 [22]，使得其在操作性提取中被广泛使用 [16, 25, 26]。然而，IR 数据通常与降水的相关性较弱，因为它主要捕捉云顶特征而非水成物剖面 [1, 5]，导致降水提取质量较低。相比之下，从低地球轨道（LEO）卫星获取的 PMW 和 PR 反射率可以穿透云层以检测降水粒子的三维微观结构 [6]。高级提取方法 [7, 20, 23] 利用这些直接的降水-云关系，实现了卓越的性能。然而，与 IR 数据（全圆盘、5 分钟、4 公里）相比，PMW 和 PR 的空间覆盖（扫描带宽狭窄，2,873 公里/245 公里宽）和时空分辨率（90 分钟，13 公里 & 90 分钟，5 公里）都有限。数据的异质性问题阻碍了它们提供连续的全球覆盖的能力，这对于有效的降水监测至关重要。

为了克服这些限制，提供高质量的全圆盘降水检索，我们提出了一项新颖的降水检索扩展（PRE）任务，如图 1 (a) 所示。该任务利用多模态数据（IR、PMW 和 PR）的互补优势来解决窄覆盖范围的限制并提高检索精度。虽然多模态数据在单个扫描轨迹内是可用的，但 IR 是唯一具有全圆盘覆盖的模式，这使其在超出扫描轨迹范围的降水检索扩展中成为必不可少的。本论文将 PRE 任务构建为一个不完全的多模态学习

问题 [8, 37]。为了解决这一挑战，一种可能的方法是生成丢失的多模态数据 [24, 28, 33]。然而，由于上述不同数据捕获的物理特性差异，缺失的模式很难恢复。或者，模式不可知学习方法 [9, 13, 43] 旨在学习对特定模式无差别的表示。然而，当每种模式与降水的相关性存在严重不平衡时，学习共享特征可能会毫无用处。相对而言，我们发现知识蒸馏 (KD) [38–40] 提供了一个更可行的解决方案。它通过教师-学生模型将与降水相关的知识从高质量的多模态数据转移到低质量的 IR 数据，从而在缺失模式的情况下促进更有效的学习。尽管如此，转移知识仍然具有挑战性，特别是因为 IR 数据的更大覆盖范围并不一定转化为更多有用的降水知识。

为了应对上述挑战，我们提出了一个由两个阶段组成的多模态知识扩展管道，如图 1 所示。在第一阶段，Swath-Distilling，我们使用基于特征的知识蒸馏来将降水相关的知识从一个在多模态数据 (PMW, PR, IR) 上训练过的教师转移到仅操作 IR 的学生。我们设计了一个只有 48M 参数的轻量级 PRE-Net 模型来实现这一目标，并引入了协同掩蔽和小波增强 (CoMWE) 模块。它由两个关键组件组成：重新掩蔽知识蒸馏 (RMKD)，通过应用策略性重构的高掩蔽率掩码来编码更丰富的内容知识；细节感知小波增强 (DAWE)，通过利用高频小波成分来保留细粒度细节。在第二阶段，全圆盘适应中，我们提出了一个简单但有效的微调方法 Self-MaskTune，它将降水检索从扫描条带推广到全圆盘。Self-MaskTune 通过指导的参数更新来平衡多模态知识和全圆盘 IR 知识，实现这一点。该管道使得 PRE-Net 能够有效地从多模态数据中学习，同时解决窄覆盖和缺失模式的限制，从而最终增强全圆盘降水检索。

我们的贡献如下：(i) 我们引入了降水检索扩展任务，旨在实现高质量的基于红外的全圆盘降水检索，并提供了一个基准。(ii) 我们提出了 PRE-Net，具有两个阶段的多模态知识扩展管道：Swath-Distilling 和 Full-Disc Adaptation，使得有效的多模态知识转移成为可能，以克服质量和覆盖差异的限制。(iii) 我们引入了 CoMWE 模块来有效编码更丰富的知识和高频成分，以及 Self-MaskTune 来自适应平衡从多模态数据中提取的知识与从全圆盘红外上学习的新知识。(iv) 实验结果表明，PRE-Net 超过了几个基于红外的降水产品（例如，PERSIANN-CCS, PDIR），并达到了与多模态算法 IMERG 相当的性能。

许多基于卫星观测的降水反演算法已经被提出。早期的方法依赖于物理模型和经验关系。例如，基于红外的算法，如 GPI [1] 和 GMSRA [3]，建模了亮温 (BTs) 与降水之间的关系。同样地，GPROF [20] 使用贝叶斯方法来匹配微波 (PMW) 与降水剖面。随着机器学习的出现，数据驱动的方法 [7, 23, 36] 在降水反演中获得了显著的关注。其中，PERSIANN [25] 及其衍生产品（例如，PERSIANN-CCS [16] 和 PDIR [26]）是基于红外数据的卫星降水产品 (SPPs)，被科学界广泛使用。

1.0.1 多模态降水检索 将多个传感器的数据整合在一起的努力产生了像 Deep-STEP [10] 这样的算法和其他一些算法 [2, 41]。然而，由于卫星观测的稀疏性和异步性，假设完全的多模态数据通常是不现实的。一些基于物理的方法通过计算云运动矢量来扩展扫描观测，以应对模式稀疏性 [17, 19, 29]。综合多卫星检索 GPM (IMERG) [17] 是降水监测和研究中最权威的 SPP。IMERG 将多种模式集成用于偏差校正，提供近全球范围内时空分辨率高的降水估算 ($0.1^\circ \times 0.1^\circ$, 30 分钟间隔)。然而，这些物理方法需要大量的计算资源，并在复杂的地形和季节性中表现出系统偏差 [12, 30]。此外，IMERG 受到实时可用性有限的影响。即使是近实时使用设计的产品也有 4 到 14 小

时的延迟。此外，利用统计方法 [35] 进行扫描降水扩展的尝试已发表，但这些方法通常随着扩展范围的增加而导致误差的增加。因此，我们提出了降水检索扩展 (PRE) 任务，以实现高质量、高分辨率、低成本的降水检索，覆盖全球。

大多数多模态学习任务假定可以获得完整的数据，但由于设备限制、隐私问题等原因，这种情况并不现实。为了解决不完整的多模态学习，最近的努力集中在两种主要方法上：基于生成和基于学习的方法。基于生成的方法试图重建缺失的模式，将任务转换为完整的多模态任务。这些方法在有足够的配对多模态数据用于训练时是有效的。基于学习的方法侧重于从完整的多模态中学习表示。这些方法可以进一步分为模式无关学习和知识蒸馏 (KD)。模式无关学习方法捕捉对任何缺失模式场景都适用的模式不变特征。这些方法适用于缺失模式多样但模式与目标之间的相关性强度平衡的情况。知识蒸馏从教师模型（在完整模式上训练）向学生模型（处理不完整输入）转移知识。KD 可以分为两种类型：基于响应的 KD，利用软标签进行蒸馏，以及基于特征的 KD，利用中间特征捕捉更丰富的知识。在这项工作中，我们采用基于特征的 KD 来处理 PRE，因为它将强相关的 PMW/PR 中的降水相关知识转移到 IR 上。

2 方法论

2.1 任务定义

正式地，我们将 $X = \{X_{IR}, X_{PMW}, X_{PR}\}$ 定义为由红外、被动微波和星载雷达反射率组成的多模态数据集，其中 X_{IR} 在扫描和全圆盘覆盖中均可用，而 X_{PMW} 和 X_{PR} 仅限于扫描覆盖。相应的降水标签用 $Y = \{Y^s, Y^f\}$ 表示，代表扫描和全圆盘区域的降水地面实况。在训练过程中，集成的多模态数据 X 用于预测降水 Y 。在推理时，仅给定一个红外观测 $x_{IR} \in X_{IR}$ ，模型预测全圆盘降水 $\hat{y}^f$ ，从而实现超出扫描覆盖的高质量降水检索。PRE 任务包括两个目标：降水像素的分类和降水强度的回归。

2.2 PRE-Net 和多模态知识扩展概述

多模态知识扩展 (MKE) 流程如图 2 所示。PRE-Net 是一种 U 形模型，结合了编码器端的专用协调掩蔽和小波增强 (CoMWE) 模块，使 PRE-Net 能够编码更丰富和更详细的特征。整体而言，MKE 流程包括条带提取和全圆盘适应两个阶段。在阶段 I 中，我们预训练一个整合了各种条带数据的多模态 UNet 模型，包括 X_{IR}, X_{PMW}, X_{PR} ，以学习用于降水反演的联合特征表示。这个预训练的模型然后作为多模态教师，其知识传递到单模态 PRE-Net，该网络仅用条带红外数据进行降水反演。在阶段 II 中，单模态 PRE-Net 进行全圆盘适应，我们使用 LoRA [18] 和自掩式调谐来优化其全圆盘降水反演能力，同时保持从教师模型转移的知识，以应对红外数据提供的更广泛空间覆盖。这个两阶段流程使得 PRE-Net 能够弥合多模态数据整合与有效的单模态泛化之间的差距，提高降水反演的准确性和鲁棒性。结果，PRE-Net 能够仅依赖红外数据实现高质量的全圆盘反演。

2.3 条带蒸馏：协调屏蔽与小波增强

我们首先整合多模态数据 (IR, PMW 和 PR)，并训练了一个教师 UNet 模型用于扫描区域降水检索。这种多模态融合显著优于单模态基线，并且比领先的卫星降水产品 (SPP) GPM 2B-CMB [11] 更出色（有关架构和消融研究，请参见补充材料 A）。为了将这种能力从扫描区域推广到全圆盘降水检索，我

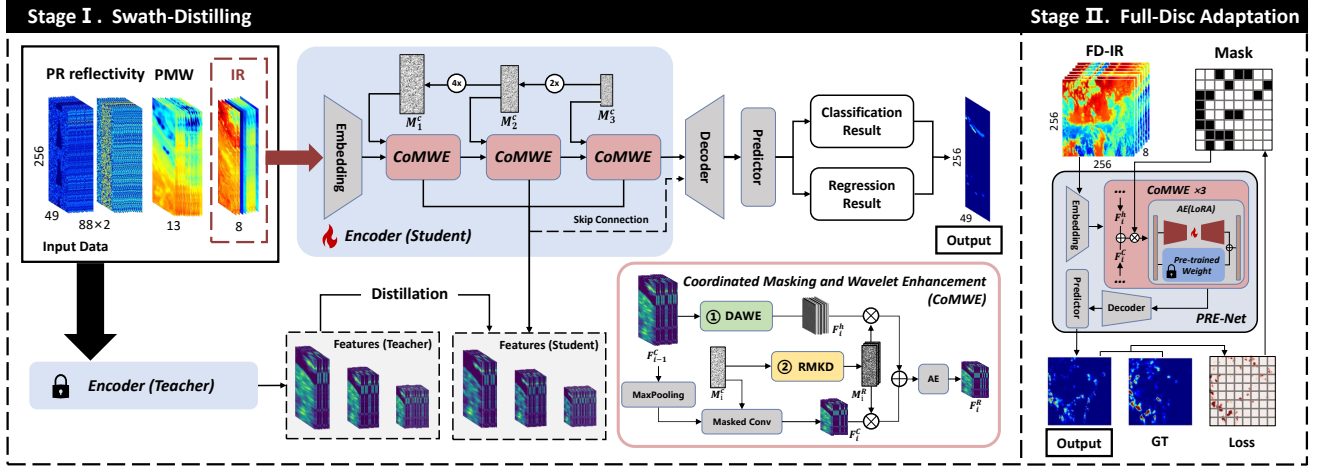


Figure 2: PRE-Net 的多模态知识扩展流程。在阶段 I 中详细介绍了 PRE-Net 的框架。训练两个相同的 PRE-Net 模型来获得分类和回归结果。

们采用基于特征的知识蒸馏方法，其中一个限于仅输入 IR 的学生 PRE-Net 模型，与教师的多模态表示对齐其特征图。

然而，教师的多模态输入和学生的仅红外 (IR) 输入之间形成了显著的特征表示差距。传统基于特征的知识蒸馏 (KD) 难以弥合这种不对称性，因为学生无法访问教师的补充 PMW/PR 数据。为了在知识蒸馏过程中增强特征学习，我们利用了掩码图像建模 (MIM)。MIM 会掩盖随机输入块并训练模型从可见上下文中重建缺失特征。MIM 的一个著名进展是掩码自动编码器 (MAE) [14]。虽然 MAE 在高掩码比例（例如，75% 掩码）下表现出色，但其在掩码知识蒸馏 (MKD) [21] 中的整合揭示了一个关键的不匹配：MKD 在低掩码比例（约 20% 掩码）下实现了最佳蒸馏效果，未能利用 MAE 的高掩码重建优势。MAE 的高掩码设计和 MKD 的低掩码操作模式之间的这种不匹配造成了低效，因为 MKD 未能在其预期的高掩码范式下充分利用 MAE 的重建能力。这种低效限制了学生从部分多模态观测（即，仅红外）中恢复高质量降水模式的能力。为了解决这一矛盾，我们提出了协调掩码和小波增强 (CoMWE) 模块，该模块自适应地平衡高比例掩码和细节保留。如图 2 所示，CoMWE 结合了两个组件：RMKD 和 DAWE。CoMWE 的过程可以表示为如下：

$$F_i^c = \text{MaskedConv}(\phi(F_{i-1}^c) \times M_i^c; \theta_{MC}), \quad (1)$$

$$F_i^h = \text{DAWE}(\text{embed}(X_{IR})), \quad (2)$$

$$F_i^R = \text{AE}((F_i^c + F_i^h) \times M_i^R; \theta_{AE}), \quad (3)$$

，这里的 MaskedConv 和 AE 代表多尺度掩码卷积和自动编码器。 θ_{MC} 和 θ_{AE} 是对应的权重。 ϕ 是一个具有 1×1 核大小的最大池化和卷积层，用于对齐降采样通道的维度。 F_i^c 、 F_i^h 和 F_i^R 分别表示多尺度卷积特征、高频提示特征和重建特征。 M_i^c 、 M_i^h 是用于多尺度掩码卷积和自动编码器的掩码。公式 2 将在接下来的部分中描述。

我们开发了 RMKD，其中结合了针对 F_i^c 定制的不同掩码来重建 F_i^R 。如图 3 所示，使用预定义的掩码比例 α 应用于多尺度卷积层。这个初始掩码 M_i^c 隐藏了少量特征，使得模型能够从可见的块中学习到健壮的特征。为了在高掩码比例下优化特征重建，对 F_i^c 剩余的可见块进行进一步打乱，并均等分为 N 部

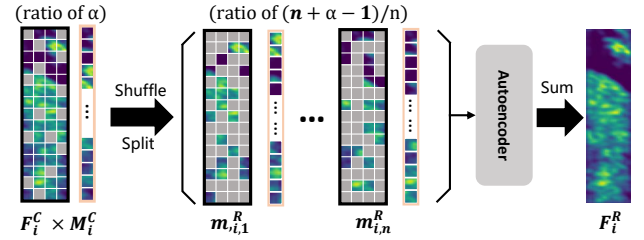


Figure 3: RMKD 的说明。输入是特征图 F_i^c 和掩码 M_i^c 。

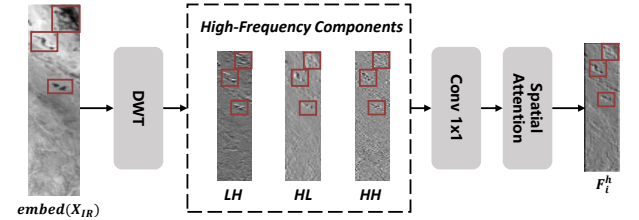


Figure 4: DAWE 的示意图。它将输入图像或特征图分解为小波系数组件。红色高亮区域是与降水相关特征的高频细节。DAWE 精确地保留了这些细粒度特征，并以清晰的边界对它们进行重建。

分。然后，将这些 N 部分与掩码集 $M_i^R = \{m_{i,1}^R, m_{i,2}^R, \dots, m_{i,n}^R\}$ 分别结合，生成 N 掩码块，每个都有高掩码比例 $\frac{n+\alpha-1}{n}$ 。然后，这些掩码独立应用于 F_i^c 和 F_i^h ，并输入到自动编码器中以进行特征重建。最后， F_i^R 将重构后的特征与可学习权重结合起来。

高频细节，如降水边界和局部纹理，对于在蒸馏过程中的精确知识转移至关重要。然而，增加掩码率可能放大错误的高频组件（例如，虚假边缘或噪声），而传统掩码结合池化层进一步削弱了高频组件。为了解决这个问题，我们提出了细节感

知小波增强 (DAWE) 模块 (图), 该模块利用离散小波变换 (DWT) 来保留和增强高频细节, 确保更准确的特征恢复。

DWT 将输入图像或特征图分解为四个子带: 低频 (LL): 捕捉全局语义模式 (例如, 大规模降水系统)。高频 (HL, LH, HH): 编码方向细节——水平 (HL)、垂直 (LH) 和对角线 (HH) 成分是关键 (例如, 细粒度降水边界)。尽管高频子带天生包含噪声, 但它们保留了基本的结构细节。DAWE 首先提取并连接 HL、LH 和 HH 成分。然后, 通过一个 1×1 卷积层对它们的通道进行对齐。为了在抑制噪声的同时增强判别细节, 应用了一种空间注意力 (SA) 机制, 该机制根据高频特征与降水模式的空间相关性动态地重新校准其权重。DAWE 的过程可以被表述为如下:

$$LL_i, (HL_i, LH_i, HH_i) = DWT(embed(X_{IR}); f_{haar}), \quad (4)$$

$$F_i^h = SA(\phi(cat(HL_i, LH_i, HH_i))), \quad (5)$$

其中, ϕ 表示用于通道对齐的 1×1 卷积层, SA 是空间注意力模块。方程 4 表示对 $embed(X_{IR})$ 应用第 i 次 DWT, 使用 Haar 小波滤波器。

2.3.1 优化目标. PRE-Net 最后由一个轻量级的预测模块收尾, 该模块将解码器的输出特征转换为降水类别或估计值。该预测器采用一个 1×1 卷积层将特征映射到单通道表示中, 随后使用 sigmoid 函数进行降水分类。PRE-Net 的训练目标由三个损失组成: (i) 任务损失 $\mathcal{L}_{task}$, 例如, 用于分类的交叉熵 (CE) 损失和用于回归的均方误差 (MSE) 损失; (ii) 基于特征的知识蒸馏损失 $\mathcal{L}_{feat}$, 采用 KL 散度来对齐学生和教师的潜在空间, 以及 (iii) 特征重构损失 $\mathcal{L}_{rec}$, 在学生从掩码输入中重构的特征与教师的原始特征之间强制像素级一致性。在片段蒸馏阶段的总训练损失如下:

$$\mathcal{L}_{SD} = \mathcal{L}_{task} + \lambda \mathcal{L}_{feat} + \gamma \mathcal{L}_{rec}, \quad (6)$$

, 其中 λ 和 γ 是用于平衡蒸馏损失的参数。

在全盘适应阶段, 模型从扫描条带过渡到全盘降水检索。一个核心挑战是保持在阶段 I 中提炼的多模态知识, 同时适应全盘红外数据的分布特征。一种简单的策略, 即冻结 CoMWE 模块并仅微调解码器, 往往会产生次优结果, 因为它无法完全考虑条带和全盘降水模式之间的差异。相反, 允许 CoMWE 完全可学习则有遗忘知识的风险, 因为虽然仅有红外数据很丰富, 但它缺乏多模态输入对于精确检索的可靠性。为了解决这个问题, 我们采用了一种混合策略: LoRA 使得高效适应基于新红外的数据成为可能, 而 Self-MaskTune 通过选择性更新模型参数以应对高知识差异, 从而保留先前获得的多模态知识。

为了避免 CoMWE 模块在微调期间严重的知识遗忘, 我们选择性地仅更新自编码器 (AE) 参数。卷积层将单模输入转化为多模兼容特征, 而 AE 负责优化这些特征以进行降水检索。冻结卷积层可以保留其预训练的多模映射能力, 而自适应 AE 则确保任务特定特征的针对性优化。具体来说, 我们在 AE 中集成了低秩适配 (LoRA) [18], 通过旁路学习约束参数更新: $W_{updated} = W_{pretrained} + \Delta W$ 。为了指导微调过程, 我们提出了 Self-MaskTune (图 2), 一种新颖的动态掩蔽策略, 优先考虑预测不确定性较高的区域, 其中基于条带的多模态知识与全光盘红外模式发生冲突。通过动态优先考虑这些区域, 模型能够超越空间限制的降水系统 (例如, 对流核心) 进行泛化, 减少对特定条带偏差的过拟合。具体来说, 在前 K 个周期中, 不施加任何掩蔽, 使模型适应全光盘红外特征。经过 K 个周期后, 引入掩膜 $M^{AE}(y, t-1)$ 以突出 F^R 中预测错误较高的

Data Type	Time Frequency	Spatial Resolution	Spatial Coverage	Data Usage
IR (FY4A/AGRI)	0.07 h	4 km	Full-disc	Train / Infer
PMW (GPM/GMI)	1.5 h	13 km	Swath	Train
PR (GPM/DPR)	1.5 h	5 km	Swath	Train
CLDAS-V2.0	1 h	~ 7 km	Full-disc	Ground Truth
PERSIANN-CCS	1 h	~ 4.5 km	Full-disc	
PDIR	1 h	~ 4.5 km	Full-disc	Competitive
GPM 2B-CMB	1 h	5 km	Swath	Products
IMERG	0.5 h	~ 11 km	Full-disc	

Table 1: 本研究中使用的数据和产品。

RMKD	DAWE	RMSE ↓	CC ↑	POD ↑	FAR ↓	CSI ↑
×	×	0.9782	0.3408	0.2159	0.4536	0.1831
✓	×	0.9484	0.4051	0.4288	0.4702	0.3106
×	✓	0.9689	0.3859	0.2456	0.3772	0.2138
✓	✓	0.9457	0.4106	0.4921	0.4944	0.3322

Table 2: CoMWE 在 Swath-Distilling 阶段的消融实验, 最佳结果用粗体表示。

区域, 这需要重新学习以提高泛化能力:

$$F^R = \begin{cases} AE((F^c + F^h); \theta_{LoRA}), & t \leq K, \\ AE((F^c + F^h) \times M^{AE}(y^f, t-1); \theta_{LoRA}), & t > K. \end{cases} \quad (7)$$

$M^{AE}(y^f, t-1)$ 是一个标识样本 y 在之前 $t-1$ 时期中预测误差较高的区域的掩码, 并应用于 AE 内的特征图。为了获得这个掩码, 我们引入 $\mathcal{P}_{max}$ 作为阈值, 用于确定某个特征区域是否需要进一步学习:

$$\mathcal{P}_{max} = \rho \cdot \max(\mathcal{L}_{task}^{t-1}), \quad (8)$$

$$M^{AE}(y^f, t-1) = \begin{cases} 1, & \text{where } \mathcal{L}_{task}^{t-1}(y^f, \hat{y}^f) \geq \mathcal{P}_{max}, \\ 0, & \text{otherwise.} \end{cases} \quad (9)$$

其中 y^f 和 $\hat{y}^f$ 分别表示全盘降水的真实值和预测值。 ρ 是用于调整阈值的参数。全盘自适应阶段的训练损失表述为:

$$\mathcal{L}_{FF} = \mathcal{L}_{task}(y^f, \hat{y}^f), \text{ subject to } M^{AE}(y^f, t-1). \quad (10)$$

通过该损失, 自掩码调参首先巩固已有的知识并吸收新的知识, 然后在不确定性的指导下逐渐整合新的知识, 实现两种知识的平衡。

本研究重点关注覆盖东亚及其附近地区的 $0-65^\circ\text{N}$ 和 $60-160^\circ\text{E}$ 区域的降水检索。

如表 1 所示, 本文使用的数据包括具有不同时间和空间分辨率的多模态观测。来自 FY-4A/AGRI 的红外数据是训练和推理的主要输入, 具有 0.08 小时的时间分辨率和 4 公里的空间分辨率, 覆盖整个圆盘。来自 GPM/GMI 的被动微波 (PMW) 数据和来自 GPM/DPR 的降水雷达 (PR) 数据提供云的微物理特性, 提供扫描带覆盖, 时间分辨率为 1.5 小时, 空间分辨率分别为 13 公里和 5 公里。

由于

3.2.2 数据集构建 . CLDAS-V2.0 [32] 在研究区域内通过地面站观测验证时表现优于其他产品，它被用作评估的真实值。我们使用竞争产品 PERSIANN-CCS [16]、PDIR [26]、GPM 2B-CMB [11] 和 IMERG [17] 作为数值天气预报 (NWP) 方法 [4] 进行比较。我们的数据集分为两个部分：用于 Swath-Distilling 阶段的 Swath-MPR 和用于 Full-Disc Adaptation 阶段以及 PRE 任务的 Full-IPR。我们评估模型在子午和全盘场景下的整体性能。
Swath-MPR：我们收集了从 2021 年到 2022 年的多模态数据作为 Swath 多模态降水检索数据集 (Swath-MPR) 用于训练，该数据集包括红外 (IR)、微波 (PMW)、PR 反射率、地理信息和降水产品：CLDAS-V2.0 和 GPM 2B-CMB。PMW 和 PR 观测数据来源于 GPM 核心观测站 (GPM-CO)，提供每个扫描进入研究区域的精确时间戳。对于每个 PMW/PR 扫描，我们在时间上识别最接近的 FY-4A IR 数据和 CLDAS-V2.0，确保最大时间偏差为 10 分钟，以减少不匹配。在空间上，所有数据在 5 公里的分辨率下进行了空间匹配和裁剪处理 256×49 。我们排除了降水极少或没有降水的扫描，因为它们对评估指标的方差贡献很小 [44]。Swath-MPR 由 3597 个扫描组成，其中 2281 个扫描和 279 个扫描来自 2021 年用于训练和验证，1037 个扫描来自 2022 年用于测试。
Full-IPR：我们还收集了 2022 年和 2023 年的 CLDAS-V2.0 红外和降水数据，以创建完整的圆盘红外降水检索数据集 (Full-IPR)，用于微调和评估 PRE-Net。我们将这些数据对齐到 4 公里的空间分辨率，并执行了 256×256 裁剪的分块处理。降水很少或没有降水的分块被排除。2022 年的数据被分成训练集 (前 10 个月) 和验证集 (最后两个月)，而 2023 年的所有数据用于测试。总共使用了 6,220 个分块进行训练，754 个分块用于验证，12,819 个分块用于测试。

我们采用一组统计指标来量化各种方法的性能。这些指标的选择基于其在评估降水检索方面的已建立有效性，如先前研究所示。
对于降水分类，我们使用探测概率 (POD)、误报率 (FAR) 和关键成功指数 (CSI)，其中 CSI 被认为是最平衡的指标。应用 0.1 毫米/小时的降水阈值来区分下雨和不下雨的像素，这符合标准实践。此外，为了解决比较产品 (PERSIANN-CCS 和 PDIR 为 0.04° ，IMERG 为 0.1°) 与 PRE-Net 结果 (4 公里) 之间的分辨率差异，我们引入了 CSI-邻近 (CSI-4 和 CSI-8) [44]。这些指标通过使用核大小为 4 和 8 的最大池化计算，以评估预测和观测到的降水模式之间的局部空间一致性。
对于降水强度的估计，我们使用均方根误差 (RMSE) 和皮尔逊相关系数 (CC)。RMSE 量化了检索到的降水与 GT 之间误差的大小，而 CC 则衡量了预测的强度与 GT 之间的线性关系。
我们使用 PyTorch 实现了 PRE-Net。训练和推理均在 NVIDIA RTX A6000 GPU 上进行。我们使用 RMSprop 来优化 PRE-Net，初始学习率为 $1e-6$ 。我们每 30 个 epoch 将学习率降低 0.5 倍。训练批量大小设为 32，epoch 分别在两个阶段设置为 200 和 100。初始掩码比例 α 设置为 0.25，RMKD 中生成的掩码数量 n 设置为 3。 λ 设置为 0.2， γ 设置为 50。我们选择在验证集上表现最好的模型，并报告其在测试集上的性能。

Method	RMSE ↓	CC ↑	POD ↑	FAR ↓	CSI ↑
UNet	0.9782	0.3408	0.2159	0.4536	0.1831
KD [15]	0.9807	0.3567	0.2927	0.4323	0.2393
MGD [42]	0.9779	0.3429	0.3036	0.4720	0.2388
MKD [21]	0.9541	0.3940	0.3654	0.4430	0.2831
RMKD	0.9484	0.4051	0.4288	0.4702	0.3106

Table 3: 对于 Swath-Distillation 阶段的知识蒸馏方法进行消融实验，最佳结果用粗体显示。

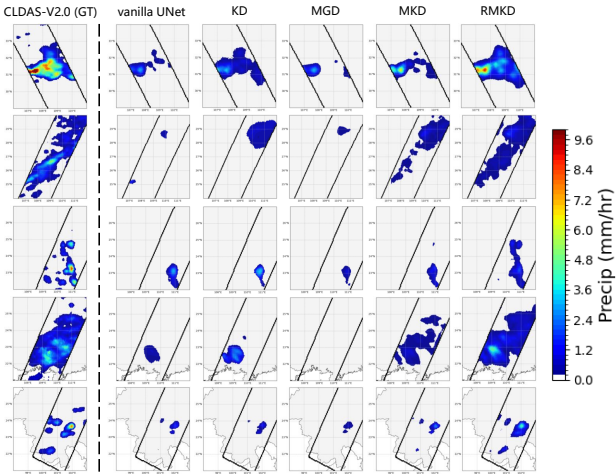


Figure 5: 在不同扫描带宽上的视觉比较。

3 实验结果

3.1 消融研究

为了验证我们 MKE 流水线和模块设计的有效性，我们在三个关键维度上进行了消融研究：(i) CoMWE 模块组件，(ii) Swath-Distilling 阶段的知识蒸馏方法，以及 (iii) Full-Disc Adaptation 阶段的微调策略。
CoMWE 组件。表 2 中提供的 CoMWE 模块的消融结果验证了其设计。结果是在 Swath-MPR 测试集中报告的。首先，只有 RMKD 在基线之上实现了显著改进，特别是在 CC (+5.57 %)、POD (+21.29 %) 和 CSI (+12.75 %) 上。这表明 RMKD 通过策略性地遮盖不太相关的区域并专注于知识丰富的内容，有效增强了特征表示。其次，仅 DAWE 导致表现不佳。虽然它降低了误报率为 0.3772 的水平，但在其他指标上，它比仅 RMKD 差，因为它在增强详细知识的同时也放大了噪音。最后，结合 RMKD 和 DAWE 获得了最佳结果 (CC: 0.4106, CSI: 0.3322)。这些结果验证了 CoMWE 模块的设计，表明 RMKD 和 DAWE 是互补的，它们的整合对于实现最佳降水检索性能是必不可少的。
蒸馏策略。表 3 比较了在 Swath-Distilling 阶段应用的各种知识蒸馏策略。结果是在 Swath-MPR 测试集上报告的。仅在 IR 数据上训练的 UNet 基线性能有限 (CC: 0.3408, CSI: 0.1831)，突出了仅用 IR 数据进行降水检索的挑战。包含基础知识蒸馏带来了适度的改进 (CC 提高了 1.59 %，CSI 提高了 5.62 %)，证实了将多模态教师知识蒸馏到基于 IR 的学生中的价值。MGD 仅通过引入掩码策略而不通过 AE 提炼多模态特征，其表现不

Method			RMSE ↓	CC ↑	POD ↑	FAR ↓	CSI ↑
Train from scratch	-		0.9082	0.4589	0.3309	0.3169	0.2869
Train with MKE	KD	Non-MaskTune	0.8976	0.4613	0.3649	0.3357	0.3081
	CoMWE	/	1.0291	0.1063	0.1173	0.4606	0.1066
		Non-MaskTune	0.8990	0.4616	0.3171	0.3006	0.2791
		Rand-MaskTune	0.8853	0.4850	0.4136	0.3887	0.3274
		Self-MaskTune	0.8714	0.5031	0.5598	0.4349	0.3912

Table 4: 全盘降水检索的策略组合消融研究，最佳结果以粗体显示。

如基础 KD。MKD 通过结合掩码特征重建和蒸馏进一步提升性能 (CC 提高了 5.32%，CSI 提高了 10.0%)。我们的 RMKD 方法达到最佳结果 (CC: 0.4051, CSI: 0.3106)，比基线提高 12.75% 的 CSI，展示了其在表示学习中平衡掩码比率和特征重构的能力。此外，RMKD 显著提高了 POD (+21.29%)，并将 RMSE 降至 0.9484 (-3.08%)，凸显其将多模态知识传输至仅 IR 学生的能力。

图 5 中的可视化比较了不同的 Swath-Distilling 知识蒸馏方法。UNet 基线、KD 和 MGD 都低估了降水强度和覆盖范围。MKD 稍微增强了降水细节，但在一些区域仍然存在不足。RMKD 恢复了最准确和详细的降水模式，在强度和空间分布上都与真实值非常接近。

微调策略。我们在表 4 中比较了用于全盘降水检索的 PRE-Net。一个在全盘红外数据上从头训练的 UNet 被用作基线 (第 1 行)。与其他使用 MKE 训练的模型相比，它可以看作是一个指标，表明多模式知识是否隐含在微调的模型中。在第 3 行中，我们显示出在条带检索中表现优异的 PRE-Net 模型在全盘检索中未能泛化，尽管两者都是 IR 数据。这强调了全盘适应阶段对于全盘检索的必要性。第 4-6 行比第 3 行取得了更好的检索结果，进一步证明了这一点。这证实了这些微调策略有效地平衡了先前的多模式知识与全盘 IR 知识，从而实现了更优越的降水检索。为了进一步验证 Self-MaskTune 的有效性及其在减轻知识遗忘中的作用，我们将 Self-MaskTune 方法与 Non-MaskTune 和 Rand-MaskTune 进行了比较。Non-MaskTune 在没有任何屏蔽的情况下微调所有参数，仅依赖于 LoRA。Rand-MaskTune 使用 LoRA 进行微调，同时使用随机生成的掩码而不依赖损失指导。Self-MaskTune 采用了一个动态掩码策略，其中掩码由预测误差引导。将 KD 与 Non-MaskTune (第 2 行) 结合起来比基准有了提升，这验证了多模态知识扩展流程的有效性。与 Non-MaskTune 一起使用的 PRE-Net (第 4 行) 表现逊于与 Non-MaskTune 一起使用的 KD (第 2 行)，甚至不如基准 (第 1 行)，这表明仅靠 LoRA 不足以实现有效的适应。在没有引导更新的情况下，LoRA 可能会保留与完整光盘 IR 领域不一致的冲突的多模态特征，从而降低模型学习新模式的能力。这突出了有针对适应的必要性，这正是 Self-MaskTune 提供的。对比第 6 行和第 1 行，我们的方法 (PRE-Net 与 Self-MaskTune) 展示了优异的检索性能 (RMSE 提升 2.76%，CC 提升 2.65%，POD 提升 24.27%，CSI 提升 11.21%)。

请注意，在降水检索中，FAR 指标略有增加是可以接受的。一方面，解决漏报比错误报告更为关键，因为它可以防止对潜在灾害的及时响应。另一方面，实现更高的 POD 通常以更高的 FAR 为代价。

定量比较。表 5 量化了 PRE-Net 相较于某些降水产品的优越性。与基于 NWP 的 IR 方法相比，PRE-Net 显著优于 PERSIANN-CCS 和 PDIR。在分类指标中，PRE-Net 取得了显著的改进：与 PERSIANN-CCS 相比，POD 提高了 27.46%，FAR 降低了 8.81%，CSI 提高了 17.40%；与 PDIR 相比，POD 提高了 10.10%，FAR 降低了 15.12%，CSI 提高了 11.31%。对于回归指标，PRE-Net 实现了较低的 RMSE (0.8714)，分别比 PERSIANN-CCS 和 PDIR 减少了 0.7344 和 3.3747，并且 CC 更高 (0.5028)。这些结果展示了 PRE-Net 在解决精细尺度的降水强度变化时，能够最大限度地减少误报的能力。

值得注意的是，尽管 IMERG 依赖于多模态数据和更强大的计算资源，PRE-Net 的表现却与之相当。PRE-Net 以显著较低的计算成本实现了基于 IR 的检索，具备更快的时间分辨率 (5 分钟) 和更高的空间分辨率 (4 公里)。在分类指标中，PRE-Net 降低了误警率 (减少 4.28%) 并提高了临界成功指数 (增加 4.30%)，同时保持了较高的探测率 (增加 4.88%)。在回归指标中，它实现了更低的均方根误差 (减少 0.2360) 和更高的相关系数 (增加 5.44%)。这些结果使 PRE-Net 成为 IMERG 的一种高度可行且具成本效益的替代方案，具有在资源需求减少的情况下仍然具备竞争力的表现。虽然 IMERG 在 CSI-4 (0.4176) 和 CSI-8 (0.4891) 得分上表现更高，但这一优势来自于在插值到 4 公里分辨率时的空间平滑。此内在的空间平滑削弱了其 CSI。PRE-Net 在原生 4 公里分辨率下运行，在不降级分辨率的情况下在 CSI-4 (0.4405) 和 CSI-8 (0.4815) 得分上实现了竞争力，证实了其在不同空间分辨率下保持结构保真度的能力。

定性比较。我们在图 6 中可视化并比较了不同降水事件的检索结果。PRE-Net 始终重建降水结构，准确捕捉空间分布和强度。(i) 锋面雨。在中部地区这种带状降水中，PERSIANN-CCS、PDIR 和 UNet 无法准确捕捉该降水带的强度。IMERG 提供了更精细的降水强度，但未能形成锋面雨的形状。PRE-Net 在分布和强度上都与 GT 非常接近。(ii) 对流雨。在大面积分布的强降水中，PERSIANN-CCS 只显示稀疏降水，PDIR 显示弱降水强度，UNet 未能检测到大部分降水。IMERG 在分布上表现较好，但其强度被高估。PRE-Net 准确定位和量化了强降水。(iii) 地形雨。喜马拉雅地区密集的降水。PERSIANN-CCS、PDIR 和 UNet 完全未能捕捉到降水。IMERG 略微低估了强度。PRE-Net 成功识别出这种地形雨，紧密符合 GT。

区域泛化。尽管所选区域已经涵盖了从热带到温带的多种降水模式，我们进一步评估了 PRE-Net 的泛化能力。具体来说，我们在 2022 年 12 月 1 日至 2023 年 2 月 28 日期间的澳大利亚 (南纬 10-40°，东经 113-153°) 湿季进行了实验。所有评估都使用了由国家海洋和大气管理局 (NOAA) 维护的综合地表数据集 (ISD) 提供的站点级地面真实数据进行。与 IMERG 相比，

Method	RMSE ↓	CC ↑	POD ↑	FAR ↓	CSI ↑	CSI-4 ↑	CSI-8 ↑
PERSIANN-CCS	1.6058	0.2147	0.2851	0.5230	0.2172	0.2928	0.3672
PDIR	4.2461	0.0833	0.4587	0.5861	0.2781	0.3573	0.4354
IMERG	1.1074	0.4484	0.5109	0.4777	0.3482	0.4176	0.4891
PRE-Net	0.8714	0.5028	0.5597	0.4349	0.3912	0.4405	0.4815

Table 5: 与降水检索产品的比较。PERSIANN-CCS、PDIR 和 IMERG 是 NWP 方法，IMERG 是基于多模式的检索方法，而其他的是基于 IR 的方法。

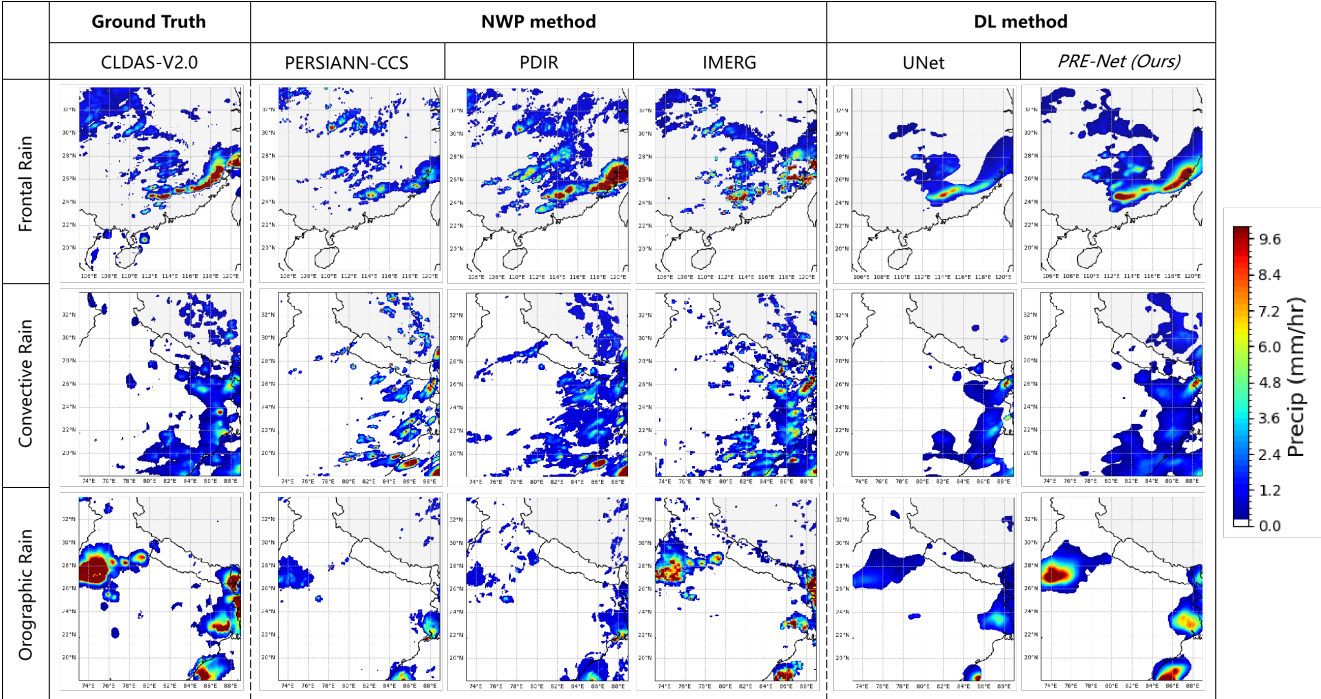


Figure 6: 关于不同全盘降水事件的视觉比较示例。

Method	RMSE ↓	CC ↑	POD ↑	FAR ↓	CSI ↑
PERSIANN-CCS	11.1723	0.1204	0.3083	0.0113	0.3073
PDIR	11.1150	0.1280	0.4113	0.0090	0.4098
IMERG	11.0625	0.1738	0.3650	0.0118	0.3634
PRE-Net	11.0303	0.1989	0.3906	0.0068	0.3896

Table 6: 澳大利亚区域概化结果（2022-2023 年湿季）。

Noise injection	RMSE ↓	CC ↑	POD ↑	FAR ↓	CSI ↑
-	0.8714	0.5028	0.5597	0.4349	0.3912
Additive Noise	0.8751	0.4999	0.5198	0.4256	0.3753
Multiplicative Noise	0.8731	0.5023	0.5530	0.4315	0.3895

Table 7: PRE-Net 在高斯噪声扰动下的鲁棒性。

PRE-Net 在所有关键指标上都优于 IMERG，如 6 所示。这些结果显示了 PRE-Net 在全球部署方面的潜力。

噪声敏感性。在实际检索系统中，卫星数据常常受到各种噪声的影响，例如大气干扰、传感器校准漂移。这些扰动可能会降低检索性能，并在实际操作中影响可靠性。因此，我们通过在输入数据中注入合成噪声（加性高斯噪声和乘性高斯噪声）来评估 PRE-Net 的噪声敏感性。如 7 所示，尽管存在噪声，PRE-Net 依然表现出显著的稳定性能，表明 PRE-Net 对噪声具有很强的抵抗力。

操作效率。通过结合 4 公里分辨率、5 分钟更新和仅红外输入，PRE-Net 提供了一种计算效率高的替代方案，相较于 SPPs 的资源密集型多模态流程。它在准确性和操作可行性之间的平衡使其成为实时高分辨率降水监测的一个有前景的工具。

我们在本文中提出了降水检索扩展 (PRE) 任务，其目标是实现超出扫描带范围的基于红外技术的高质量全盘降水检索。为了处理 PRE 任务，我们开发了一个两阶段的多模态知识扩展管道和 PRE-Net 模型，该模型包括从多模态到红外的知识转移的扫描提炼阶段和从扫描带到全盘降水检索的泛化的全盘适应阶段。在 PRE-Net 模型中，我们引入了协调遮掩和小波增强模块来有效编码更丰富的知识。对于全盘适应阶段，我们开发了自我遮掩调整，以平滑地关注提炼的多模态知识与新的仅红外知识之间的分歧，减轻知识差异的影响，促进领域转

移以实现可扩展的预测。大量的实验和分析表明，PRE-Net 是 PRE 任务的有效解决方案。

4

致谢 本工作得到了中国浙江省自然科学基金（项目编号：LRG25F020002，LR21F020002）、中国气象服务协会项目（编号：CMSA2023MD001）、中国自然科学基金（编号：62302453）、浙江省自然科学基金（编号：LMS25F020003）及中国气象局重点创新团队（CMA2023ZD01）的部分支持。

References

- [1] Phillip A Arkin and Bernard N Meisner. 1987. The relationship between large-scale convective rainfall and cold cloud over the western hemisphere during 1982–84. *Monthly Weather Review* 115, 1 (1987), 51–74.
- [2] Malarvizhi Arulraj and Ana P Barros. 2021. Automatic detection and classification of low-level orographic precipitation processes from space-borne radars using machine learning. *Remote Sensing of Environment* 257 (2021), 112355.
- [3] Mamoudou B Ba and Arnold Gruber. 2001. GOES multispectral rainfall algorithm (GMSRA). *Journal of Applied Meteorology and Climatology* 40, 8 (2001), 1500–1514.
- [4] Peter Bauer, Alan Thorpe, and Gilbert Brunet. 2015. The quiet revolution of numerical weather prediction. *Nature* 525, 7567 (2015), 47–55.
- [5] Ali Behrang, Kuo-lin Hsu, Bisher Imam, Soroosh Sorooshian, and Robert J Kuligowski. 2009. Evaluating the utility of multispectral information in delineating the areal extent of precipitation. *Journal of Hydrometeorology* 10, 3 (2009), 684–700.
- [6] Haonan Chen, V Chandrasekar, Haiming Tan, and Robert Cifelli. 2019. Rainfall estimation from ground radar and TRMM precipitation radar using hybrid deep neural networks. *Geophysical Research Letters* 46, 17–18 (2019), 10669–10678.
- [7] Haonan Chen, V Chandrasekar, Haiming Tan, and Robert Cifelli. 2019. Rainfall estimation from ground radar and TRMM precipitation radar using hybrid deep neural networks. *Geophysical Research Letters* 46, 17–18 (2019), 10669–10678.
- [8] Jiayi Chen and Aidong Zhang. 2020. HGMF: Heterogeneous Graph-based Fusion for Multimodal Data with Incompleteness. In *KDD '20: The 26th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, Virtual Event, CA, USA, August 23–27, 2020*, Rajesh Gupta, Yan Liu, Jiliang Tang, and B. Aditya Prakash (Eds.). ACM, 1295–1305.
- [9] Reuben Dorent, Samuel Joutard, Marc Modat, Sébastien Ourselin, and Tom Vercauteren. 2019. Hetero-Modal Variational Encoder-Decoder for Joint Modality Completion and Segmentation. In *Medical Image Computing and Computer Assisted Intervention - MICCAI 2019 - 22nd International Conference, Shenzhen, China, October 13–17, 2019, Proceedings, Part II (Lecture Notes in Computer Science, Vol. 11765)*, Dinggang Shen, Tianming Liu, Terry M. Peters, Lawrence H. Staib, Caroline Essert, Sean Zhou, Pew-Thian Yap, and Ali R. Khan (Eds.). Springer, 74–82.
- [10] Vesta Afzali Gorooh, Ata Akbari Asanjan, Phu Nguyen, Kuolin Hsu, and Soroosh Sorooshian. 2022. Deep neural network high SpatioTemporal resolution precipitation estimation (Deep-STEP) using passive microwave and infrared data. *Journal of Hydrometeorology* 23, 4 (2022), 597–617.
- [11] Mircea Grecu, William S. Olson, Stephen Joseph Munchak, Sarah Ringerud, Liang Liao, Ziad Haddad, Bartie L. Kelley, and Steven F. McLaughlin. 2016. The GPM Combined Algorithm. *Journal of Atmospheric and Oceanic Technology* 33, 10 (2016), 2225–2245.
- [12] Webster Gumindoga, Tom HM Rientjes, Alemseged Tamiru Haile, Hodson Makurira, and Paolo Reggiani. 2019. Performance of bias-correction schemes for CMORPH rainfall estimates in the Zambezi River basin. *Hydrology and Earth System Sciences* 23, 7 (2019), 2915–2938.
- [13] Mohammad Havaei, Nicolas Guizard, Nicolas Chapados, and Yoshua Bengio. 2016. HeMIS: Hetero-Modal Image Segmentation. In *Medical Image Computing and Computer-Assisted Intervention - MICCAI 2016 - 19th International Conference, Athens, Greece, October 17–21, 2016, Proceedings, Part II (Lecture Notes in Computer Science, Vol. 9901)*, Sébastien Ourselin, Leo Joskowicz, Mert R. Sabuncu, Gözde B. Ünal, and William M. Wells III (Eds.). 469–477.
- [14] Kaiming He, Xinlei Chen, Saining Xie, Yanghao Li, Piotr Dollár, and Ross B. Girshick. 2022. Masked Autoencoders Are Scalable Vision Learners. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition, CVPR 2022, New Orleans, LA, USA, June 18–24, 2022*. IEEE, 15979–15988.
- [15] Geoffrey E. Hinton, Oriol Vinyals, and Jeffrey Dean. 2015. Distilling the Knowledge in a Neural Network. *CoRR abs/1503.02531* (2015).
- [16] Yang Hong, Kuo-Lin Hsu, Soroosh Sorooshian, and Xiaogang Gao. 2004. Precipitation estimation from remotely sensed imagery using an artificial neural network cloud classification system. *Journal of Applied Meteorology* 43, 12 (2004), 1834–1853.
- [17] Arthur Y Hou, Ramesh K Kakar, Steven Neeck, Ardesir A Azarbarzin, Christian D Kummerow, Masahiro Kojima, Riko Oki, Kenji Nakamura, and Toshio Iguchi. 2014. The global precipitation measurement mission. *Bulletin of the American meteorological Society* 95, 5 (2014), 701–722.
- [18] Edward J. Hu, Yelong Shen, Phillip Wallis, Zeyuan Allen-Zhu, Yuanzhi Li, Shean Wang, Lu Wang, and Weizhu Chen. 2022. LoRA: Low-Rank Adaptation of Large Language Models. In *The Tenth International Conference on Learning Representations, ICLR 2022, Virtual Event, April 25–29, 2022*. OpenReview.net.
- [19] Robert J Joyce, John E Janowiak, Phillip A Arkin, and Pingping Xie. 2004. CMORPH: A method that produces global precipitation estimates from passive microwave and infrared data at high spatial and temporal resolution. *Journal of hydrometeorology* 5, 3 (2004), 487–503.
- [20] Christian Kummerow, Y Hong, WS Olson, S Yang, RF Adler, J McCollum, R Ferraro, G Petty, Dong-Bin Shin, and TT Wilheit. 2001. The evolution of the Goddard Profiling Algorithm (GPROF) for rainfall estimation from passive microwave sensors. *Journal of Applied Meteorology and Climatology* 40, 11 (2001), 1801–1820.
- [21] Shanshan Lao, Guanglu Song, Boxiao Liu, Yu Liu, and Yujiu Yang. 2023. Masked Autoencoders Are Stronger Knowledge Distillers. In *IEEE/CVF International Conference on Computer Vision, ICCV 2023, Paris, France, October 1–6, 2023*. IEEE, 6361–6370.
- [22] Vadim Lebedev, Vladimir Ivashkin, Irina Rudenko, Alexander Ganshin, Alexander Molchanov, Sergey Ovcharenko, Ruslan Grokhovetskiy, Ivan Bushmarinov, and Dmitry Solomentsev. 2019. Precipitation nowcasting with satellite imagery. In *Proceedings of the 25th ACM SIGKDD international conference on knowledge discovery & data mining*. 2680–2688.
- [23] Kangwen Liu, Jieying He, and Haonan Chen. 2022. Precipitation retrieval from Fengyun-3D MWHTS and MWRI data using deep learning. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 15 (2022), 7619–7630.
- [24] Mengmeng Ma, Jian Ren, Long Zhao, Sergey Tulyakov, Cathy Wu, and Xi Peng. 2021. SMIL: Multimodal Learning with Severely Missing Modality. In *Thirty-Fifth AAAI Conference on Artificial Intelligence, AAAI 2021, Thirty-Third Conference on Innovative Applications of Artificial Intelligence, IAAI 2021, The Eleventh Symposium on Educational Advances in Artificial Intelligence, EAAI 2021, Virtual Event, February 2–9, 2021*. AAAI Press, 2302–2310.
- [25] Phu Nguyen, Mohammed Ombadi, Soroosh Sorooshian, Kuolin Hsu, Amir AghaKouchak, Dan Braithwaite, Hamed Ashouri, and Andrea Rose Thorstensen. 2018. The PERSIANN family of global satellite precipitation data: A review and evaluation of products. *Hydrology and Earth System Sciences* 22, 11 (2018), 5801–5816.
- [26] Phu Nguyen, Eric J Shearer, Mohammed Ombadi, Vesta Afzali Gorooh, Kuolin Hsu, Soroosh Sorooshian, William S Logan, and Marty Ralph. 2020. PERSIANN Dynamic Infrared–Rain rate model (PDIR) for high-resolution, real-time satellite precipitation estimation. *Bulletin of the American Meteorological Society* 101, 3 (2020), E286–E302.
- [27] Taikan Oki and Shinjiro Kanae. 2006. Global hydrological cycles and world water resources. *science* 313, 5790 (2006), 1068–1072.
- [28] Yongsheng Pan, Ming-Xia Liu, Yong Xia, and Dinggang Shen. 2022. Disease-Image-Specific Learning for Diagnosis-Oriented Neuroimage Synthesis With Incomplete Multi-Modality Data. *IEEE Trans. Pattern Anal. Mach. Intell.* 44, 10 (2022), 6839–6853.
- [29] Giulia Panegrossi, Stefano Dietrich, Frank S Marzano, Alberto Mugnai, Eric A Smith, Xuwu Xiang, Gregory J Tripoli, Pao K Wang, and JPV Poires Baptista. 1998. Use of cloud model microphysics for passive microwave-based precipitation retrieval: Significance of consistency between model and measurement manifolds. *Journal of the atmospheric sciences* 55, 9 (1998), 1644–1673.
- [30] Rajani K Pradhan, Yannis Markonis, Mijael Rodrigo Vargas Godoy, Anahi Villalba-Pradas, Konstantinos M Andreadis, Efthymios I Nikolopoulos, Simon Michael Papalexiou, Akif Rahim, Francisco J Tapiador, and Martin Hanel. 2022. Review of GPM IMERG performance: A global perspective. *Remote Sensing of Environment* 268 (2022), 112754.
- [31] Catherine Prigent. 2010. Precipitation retrieval from space: An overview. *Comptes Rendus Geoscience* 342, 4–5 (2010), 380–389.
- [32] Chunxiang Shi, Lipeng Jiang, Tao Zhang, Bin Xu, and Shuai Han. 2014. Status and plans of CMA land data assimilation system (CLDAS) project. In *EGU General Assembly Conference Abstracts*. 5671.
- [33] Wangbin Sun, Fei Ma, Yang Li, Shao-Lun Huang, Shiguang Ni, and Lin Zhang. 2021. Semi-Supervised Multimodal Image Translation for Missing Modality Imputation. In *IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP 2021, Toronto, ON, Canada, June 6–11, 2021*. IEEE, 4320–4324.
- [34] Kevin E Trenberth, Aiguo Dai, Roy M Rasmussen, and David B Parsons. 2003. The changing character of precipitation. *Bulletin of the American Meteorological Society* 84, 9 (2003), 1205–1218.
- [35] Cunguang Wang, Guoqiang Tang, and Pierre Gentile. 2021. PrecipGAN: Merging microwave and infrared data for satellite precipitation estimation using generative adversarial network. *Geophysical Research Letters* 48, 5 (2021), e2020GL092032.
- [36] Cunguang Wang, Jing Xu, Guoqiang Tang, Yi Yang, and Yang Hong. 2020. Infrared Precipitation Estimation Using Convolutional Neural Network. *IEEE Trans. Geosci. Remote. Sens.* 58, 12 (2020), 8612–8625.
- [37] Qi Wang, Liang Zhan, Paul M. Thompson, and Jiayu Zhou. 2020. Multimodal Learning with Incomplete Modalities by Knowledge Distillation. In *KDD '20: The 26th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, Virtual Event, CA, USA, August 23–27, 2020*, Rajesh Gupta, Yan Liu, Jiliang Tang, and B. Aditya Prakash (Eds.). ACM, 1828–1838.
- [38] Qi Wang, Liang Zhan, Paul M. Thompson, and Jiayu Zhou. 2020. Multimodal Learning with Incomplete Modalities by Knowledge Distillation. In *KDD '20: The 26th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, Virtual Event, CA, USA, August 23–27, 2020*, Rajesh Gupta, Yan Liu, Jiliang Tang, and B. Aditya Prakash (Eds.). ACM, 1828–1838.

- [39] Shuai Wang, Zipei Yan, Daoan Zhang, Haining Wei, Zhongsen Li, and Rui Li. 2023. Prototype Knowledge Distillation for Medical Segmentation with Missing Modality. In *IEEE International Conference on Acoustics, Speech and Signal Processing ICASSP 2023, Rhodes Island, Greece, June 4-10, 2023*. IEEE, 1–5.
- [40] Shicai Wei, Chunbo Luo, and Yang Luo. 2023. MMANet: Margin-Aware Distillation and Modality-Aware Regularization for Incomplete Multimodal Learning. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition, CVPR 2023, Vancouver, BC, Canada, June 17-24, 2023*. IEEE, 20039–20049.
- [41] Songkun Yan, Ziqiang Ma, Xiaoqing Li, Hao Hu, Jintao Xu, Qingwen Ji, and Fuzhong Weng. 2023. PCSSR-DNNWA: A physical constraints based surface snowfall rate retrieval algorithm using deep neural networks with attention module. *Geophysical Research Letters* 50, 13 (2023), e2023GL103923.
- [42] Zhendong Yang, Zhe Li, Mingqi Shao, Dachuan Shi, Zehuan Yuan, and Chun Yuan. 2022. Masked Generative Distillation. In *Computer Vision - ECCV 2022 - 17th European Conference, Tel Aviv, Israel, October 23-27, 2022, Proceedings, Part XI (Lecture Notes in Computer Science, Vol. 13671)*, Shai Avidan, Gabriel J. Brostow, Moustapha Cissé, Giovanni Maria Farinella, and Tal Hassner (Eds.). Springer, 53–69.
- [43] Bo Zhang, Zuheng Ming, Wei Feng, Yaqian Liu, Liang He, and Kaixing Zhao. 2023. MMFormer: Multimodal Transformer Using Multiscale Self-Attention for Remote Sensing Image Classification. *CoRR* abs/2303.13101 (2023). doi:10.48550/ARXIV.2303.13101 arXiv:2303.13101
- [44] Yuchen Zhang, Mingsheng Long, Kaiyuan Chen, Lanxiang Xing, Ronghua Jin, Michael I Jordan, and Jianmin Wang. 2023. Skilful nowcasting of extreme precipitation with NowcastNet. *Nature* 619, 7970 (2023), 526–532.

A 教师网络的详细信息

这里我们详细介绍了用于 Swath-Distilling 阶段的教师网络的信息。教师网络通过多模态数据进行训练，用于扫描条带降水检索，基于基础的 U-Net 框架构建。网络处理综合的多模态数据（即 IR、PMW 和 PR 反射率）作为输入。编码器中的下采样由双卷积块组成，后接 2×2 最大池化。负责上采样的解码器结构与其相似。特征层的初始数量从 64 开始，每次下采样时翻倍，每次上采样时减半。此外，我们在网络后期加入地理数据（海拔、纬度和经度），以解决地形对降水的影响。

我们在 NVIDIA RTX A6000 GPU 上进行训练和推理来实现教师模型。该模型使用 RMSprop 优化器进行优化，初始学习率为 1e-6，每 30 个 epoch 减半。训练在 200 个 epoch 上进行，批量大小为 32。由于在中国的优越质量，CLDAS-V2.0 被用作地面实况。教师模型每 20 个 epoch 进行一次验证，并选择验证集上表现最好的模型用于测试集评估。

表 8 和表 9 中性能指标的比较显示了整合多种数据源（IR、PMW 和 PR）进行逐小时降水反演的有效性。在表 8 中，三种数据类型的组合在所有指标中始终表现最佳，具有最低的 RMSE (0.7290 mm/h) 和最高的 CSI (0.4524)，表明降水反演更准确可靠。值得注意的是，PR 数据的纳入显著提高了性能，特别是在减少 FAR 和提高 CC 方面，表明 PR 数据对降水检测的准确性有重要贡献。同时，结果突出了仅使用 IR 数据进行降水反演的局限性，因为它产生了最高的 RMSE (0.9781 mm/h) 和最低的 CC (0.3408)、POD (0.2159) 以及 CSI (0.1831) 在所有配置中。这表明 IR 数据自身缺乏准确捕捉降水模式的能力，因此与其他数据类型（PMW 和 PR）的结合对于提高反演精度和减少错误是至关重要的。

表 9 比较了教师模型（使用所有三种数据类型）与 GPM 2B-CMB 产品的性能。教师模型在所有指标上均优于 GPM 2B-CMB，在各个方面有显著提升（RMSE 提高 41.27 %，CC 提高 25.30 %，POD 提高 8.96 %，FAR 提高 8.84 %，CSI 提高 9.32 %），显示出其在精确检测降水事件方面的优越能力。这表明，利用多模态数据的教师模型在捕捉降水模式的复杂性方面比 GPM 2B-CMB 产品更有效。

IR	PMW	PR	RMSE(mm/h) ↓	CC ↑	POD ↑	FAR ↓	CSI ↑
✓			0.9781	0.3408	0.2159	0.4536	0.1831
	✓		0.7545	0.6824	0.5279	0.2678	0.4425
		✓	0.7383	0.7040	0.5409	0.2605	0.4544
✓	✓		0.7684	0.6705	0.5295	0.2917	0.4347
✓		✓	0.7351	0.7048	0.4853	0.2273	0.4274
	✓	✓	0.7290	0.7076	0.5245	0.2330	0.4524
✓	✓	✓	0.7436	0.7213	0.5511	0.2358	0.4710

Table 8: 来自不同模式的每小时降水检索性能消融。

	RMSE(mm/h) ↓	CC ↑	POD ↑	FAR ↓	CSI ↑
Teacher	0.7436	0.7213	0.5511	0.2358	0.4710
GPM 2B-CMB	1.1563	0.4683	0.4615	0.3242	0.3778

Table 9: 与 GPM 2B-CMB 在 Swath-MPR 上的性能比较。