arXiv:2304.02767v1 [eess.IV] 5 Apr 2023

MethaneMapper: Spectral Absorption aware Hyperspectral Transformer for Methane Detection

Satish Kumar satishkumar@ucsb.edu
Ivan Arevalo ifa@ucsb.edu
ASM Iftekhar iftekhar@ucsb.edu
B S Manjunath manj@ucsb.edu

Department of Electrical and Computer Engineering
University of California Santa Barbara

Abstract

甲烷（*CH_4*）是全球气候变化的主要推手。近期，航空可见光-红外成像光谱仪新一代（*AVIRIS-NG*）在甲烷排放的定量制图方面发挥了重要作用。现有的数据分析方法对局部地形条件敏感，通常需要领域专家进行人工检查，容易产生重大误差，因此难以扩展。为应对这些挑战，我们提出了一种新颖的端到端光谱吸收波长感知变换器网络 *MethaneMapper*，用于检测和量化甲烷排放。*MethaneMapper* 引入了两个新模块，帮助在光谱域中定位最相关的甲烷羽流区域，并利用这些区域实现精准定位。全面评估表明，*MethaneMapper* 在检测中实现了 0.63 平均精度均值，并且与当前最先进方法相比，模型规模缩小了 5×。此外，我们还推出了一个大规模甲烷羽流分割掩码数据集，涵盖 *2015*年至*2022*年间 *1200* 多条 *AVIRIS-NG* 飞行航线。该数据集包含 *4000* 多个甲烷羽流地点。我们的数据集将为研究人员提供机会，开发和改进应对这一具有重大社会影响的棘手温室气体检测问题的新方法。数据集和源代码链接[1]。

1. Introduction

我们研究如何利用高光谱成像数据检测并定位甲烷（CH_4）羽流。检测并定位潜在的 CH_4 热点，是应对温室气体排放导致的全球变暖的必要第一步。据估计，甲烷气体导致的全球变暖约占温室气体引发的全球变暖的 20%，其全球变暖潜能值（GWP）在 20 年内是二氧化碳（CO_2）的 86 倍 [36]。为了直观地说，二氧化碳在 100 年内造成的环境损害，甲烷在 1.2 年内就能造成。因此，监测并遏制甲烷排放至关重要。甲烷排放有多种来源，其中尤为值得关注的是石油和天然气行业的排放。根据美国环境保护署的报告，这些行业每年排放 8400 万吨甲烷 [19]。这些甲烷排放来自特定地点，主要源于管道泄漏、储罐泄漏或石油开采点泄漏。

当前检测这些源的工作主要依赖航空影像。喷气推进实验室（JPL）在过去十年中开展了数千次航空勘测，使用机载传感器 AVIRIS-NG 采集数据 [22]。已有多种方法被提出，用于从此类影像中检测潜在的排放点，例如参见 [8, 9, 39, 43, 45, 46]。然而，这些方法通常对背景环境和土地覆盖类型非常敏感，导致产生大量假阳性，而修正这些检测结果往往需要领域专家投入大量时间。主要原因在于，这些基于像素的方法仅依赖光谱相关性进行检测。空间信息可以有效减少这些假阳性，因为 CH_4 羽流呈现出类似羽流的形态结构。近期已有研究尝试利用深度学习方法来运用空间相关性 [23, 31], , 然而这些工作并未利用光谱特性来过滤混淆因素。例如，甲烷的光谱特性与白色涂漆商业屋顶或机场沥青路面等铺装路面相似 [1]。本文提出了一种新颖的基于深度网络的解决方案，以尽量减少此类混淆因素对准确定位甲烷羽流造成的影响。

我们提出的方法称为 MethaneMap- per (MM)，它改造了 DETR [4], a Transformer模型，将成像数据中的光谱和空间相关性相结合，以生成潜在甲烷（CH_4）羽流候选项的地图。这些候选项缩小了高光谱解码器检测 CH_4 羽流的搜索空间，并重新-

[1] *https://github.com/UCSB-VRL/MethaneMapper-Spectral-Absorption-aware-Hyperspectral-Transformer-for-Methane-Detection*

移动潜在混淆因素。MM 是一种轻量级端到端单阶段 CH_4 检测器，并引入了两个新模块：光谱特征生成器和 查询细化器。前者通过线性滤波器生成光谱特征，该滤波器能在存在加性背景噪声的情况下最大化 CH_4 与噪声之比；后者则整合这些特征以进行解码。

CH检测方法开发的一个主要瓶颈是$_4$ 公开训练数据的有限可用性。为解决这一问题，本研究的另一项重要贡献是引入了一个新的甲烷热点（MHS）数据集，这是计算机视觉研究人员可用的同类中最大的。MHS通过系统地从不同公开可用数据集（机载传感器 [6], 非营利组织 [3, 34] 和卫星 [38]）收集信息并生成如第4.1.1节所述的注释来策划。该策划数据集包含超过1200条AVIRIS-NG飞行航线从2015年到2022年的甲烷分割掩码。每条飞行航线包含3-4个CH_4羽流站点，MHS 数据集中总共有4000个。

Our contributions can be summarized as follows:

1.我们提出了一种使用高光谱变换器进行甲烷羽流检测的新型单阶段端到端方法。两个模块，即 光谱特征生成器 和查询细化器，协同工作，通过使用光谱感知线性滤波器改进传统的Transformer 设计，并实现对高光谱图像中潜在甲烷热点的定位，同时细化查询表示，以实现更好的解码。 2. 新的光谱线性滤波器（*SLF*）通过在光谱域中策略性地选择相关像素，从而更好地对背景分布进行白化并增强甲烷信号，改进了传统的线性滤波器。

3. 一个新的基准数据集MHS提供了最大的（∼ 35×）公开可用的标注AVIRIS-NG飞行航线数据集，来自2015年至2022年。

2. 相关工作

我们的工作处于用于CH_4 检测的高光谱数据、光谱特征的确定性线性滤波方法以及基于编码器-解码器的变换器的交叉点。下面给出相关工作的综述。

近期有几篇论文研究了利用机载成像光谱仪 AVIRIS- NG 检测甲烷羽流 [22]。其中包括迭代最大后验差分光学吸收光谱算法 (IMAP-DOAS) [7, 8] 和匹配滤波器 [9, 39, 43,45, 46]。IMAP-DOAS需要来自 2 个高光谱传感器的数据，一个位于机载平台，另一个位于地面，因此对于大多数应用场景而言并不实用。基于匹配滤波器的方法使用背景统计量进行归一化

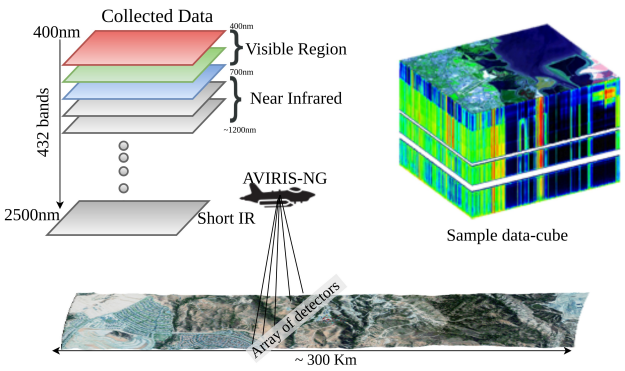

图 1. 数据采集过程的描述。每条飞行航线长∼ 300 公里。*598*个传感器阵列以*1.5*米/像素的空间分辨率记录数据。所有飞行航线均经过正射校正。每个数据立方体的维度为 $\sim 25000\times \sim 1500 \times 432$。

光谱信号，并在每个空间位置（逐像素）与 CH_4 光谱特征进行匹配。然而，这一过程对地表反照率以及具有与 CH_4, 相似光谱吸收特性的土地覆盖十分敏感，从而导致虚假检测。领域专家随后必须手动检查每条飞行航线，以识别并勾勒真实的 CH_4 羽流 [44]。为抑制地面元素变化所导致的假阳性影响，*Christopher et.al.* [10] 和 *Thorbe et.al.* [46]引入了聚类调谐匹配滤波器。该方法使用k-means聚类，将具有相似光谱特性的像素聚类在一起。由于信息是逐像素处理的，IMAP-DOAS 和所有版本的匹配滤波器都极易产生假阳性。

机器学习方法已被用于高光谱图像中的目标检测，包括CH_4 识别 [12,18,30,40,41]。与匹配滤波类似，这些方法未考虑混淆物对CH_4 光谱特征的影响，并存在类似的假阳性问题。最近引入的基于深度学习的H-mrcnn模型 [31] 专注于捕获空间相关性。H-mrcnn是处理高光谱数据块的mask-rcnncite网络集成。这种光谱域中的块处理效率低下，通常导致整体性能较差。Methanet [23] 是一项更近期的工作，专注于从匹配滤波数据中估计甲烷浓度。在这方面，我们提出的MethaneMapper同时使用光谱和空间相关性来准确描绘CH_4 羽流。

Datasets: The only dataset publicly available with annotation for CH_4 plume detection is JPL-CH_4-detection2017-V1.0 dataset [44]. It contains only 46 AVIRIS-NG [22] flight lines in the US Four-Corners region. Deep learning architectures require a large number of annotated samples, and for this reason we introduce the new MHS dataset with over 1200 annotated flightlines and $\sim$ 4000 plume sites.

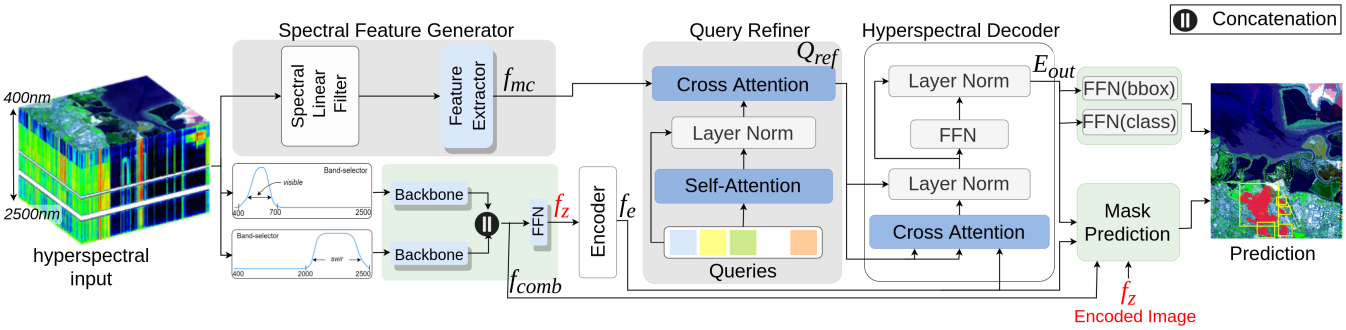

图 2. *MethaneMapper*（*MM*）架构概览。给定一幅高光谱图像，我们的 *RGB*（$400nm - 700nm$）和 *SWIR*（$2000nm - 2500nm$）带通滤波器会筛选出所需波长范围内的部分通道，并将其输入 *CNN* 骨干网络（*ResNet*）以提取特征。这些特征经过拼接后输入 *Transformer* 编码器。与此同时，我们的光谱特征生成器（*SFG*）模块接收输入图像的所有通道，并生成甲烷候选特征。随后，这些候选特征被送入查询细化器（*QR*）以细化查询。接着，使用 *Transformer* 编码器编码后的特征对这些查询进行解码。最后，每个解码后的查询都用于通过掩膜预测来预测羽流掩膜，并通过 *FFN*（前馈网络）预测边界框和类别。

3. MethaneMapper (MM) Architecture

3.1. Data Overview

AVIRIS-NG 高光谱成像传感器采集 N 个通道的光谱辐亮度值$_0$（$N_0 = 432$），对应的波长范围为 $400nm - 2500nm$，如图 1所示。完整的高光谱图像表示为 $\mathbf{x} \in \mathbb{R}^{H_0 \times W_0 \times N_0}$ 其中 H_0，W_0 分别表示高度和宽度，$N_0 = 432$ 表示通道数。这些高光谱数据中包含非常微弱的 CH_4 特征，位于约 2100-2400nm，并与周围土地覆盖和背景杂波的辐射混杂在一起。单条航线可能长达数英里（其中一个维度约有 25K 个像素），传感器阵列以 1.5 米/像素的分辨率记录数据。图像在处理前会进行正射校正。

3.2. Technical Overview

参见图2，MM包含以下主要组件：(i)2个CNN骨干网络，用于从高光谱图像中提取感兴趣光谱区域的紧凑特征表示；(ii) 一个光谱特征生成器 (**SFG**)，以及 (iii) 一个查询细化器 (**QR**)，位于编码器 - 解码器对之间（受 GTNet [21], SSRT [20]启发）。高光谱图像首先经过两个独立的带通滤波器，以选择可见光（$400 - 700nm$）和短波红外（SWIR）（$2000 - 2500nm$）波长区域中的通道，然后输入 CNN 骨干网络。这些骨干网络的输出被拼接在一起，随后使用 Transformer 编码器进行编码。

该**SFG** (第 3.4节）接收高光谱图像的所有通道，并通过光谱线性滤波器对其进行处理。该**SFG** 利用光谱相关性生成甲烷候选特征图，并将其传递给**QR**。

该**QR** (第 3.5节）利用这些甲烷候选项来优化可学习查询。我们的高光谱解码器接收编码器输出的编码特征，以及来自**QR** 的优化查询，以生成嵌入。掩膜预测层将这些嵌入与骨干网络各层输出的特征金字塔一同处理，从而生成最终的甲烷羽流分割预测。

These individual blocks are discussed in more detail below.

3.3. Bandpass filtering for the Encoder

HSI由两个并行的带通滤波器处理；一个是可见波长（$400 - 700nm$）（RGB）和一个短波红外波长（$2000 - 2500nm$）（SWIR）带通滤波器。RGB滤波器输出3 个通道，分别对应正常的红、绿和蓝波长。SWIR 生成多个通道，相邻通道间隔约5nm。滤波后的输出为 $\mathbf{x}_{rgb} \in \mathbb{R}^{H_0 \times W_0 \times 3}$ 和$\mathbf{x}_{swir} \in \mathbb{R}^{H_0 \times W_0 \times 100}$。使用$\mathbf{x}^{rgb}$ 和 $\mathbf{x}_{swir}$，两个传统 CNN 骨干网络（例如 ResNet-50 [17, 27]）分别生成大小为 $\in \mathbb{R}^{H \times W \times N}$的两个特征图。这里 $H = \frac{H_0}{32}$，$W = \frac{W_0}{32}$和 $N = 2048$ 通常如此。我们沿通道维度拼接这些特征图，并通过一个 1×1 卷积层将通道维度保持为 N。最终输出为$f_{comb} \in \mathbb{R}^{H \times W \times N}$。

遵循先前工作中的Transformer 编码器标准架构 [4, 20, 21, 28, 48], ，我们使用 1×1 卷积将通道维度从f_{comb} 降低至$f_z \in \mathbb{R}^{H \times W \times d}$，并通过添加固定位置嵌入$p \in \mathbb{R}^{H \times W \times d}$来补充位置信息。编码器由多头自注意力模块和前馈网络（FFN）堆叠而成。编码后的特征图为$f_e \in \mathbb{R}^{H \times W \times d}$:

$$f_e = \text{Encoder}(f_z, p) \tag{1}$$

3.4. Spectral Feature Generator (SFG)

与此同时，输入的高光谱图像由 **SFG** 模块处理，以生成甲烷候选特征图 f_{mc}，为 **QR** 模块提供空间信息，帮助网络描绘甲烷羽流。

该**SFG** 由光谱线性滤波器（SLF）和特征提取器（例如 ResNet-50 [17]）组成。检测 CH_4 最常见的线性滤波方法，是从输入高光谱图像中提取每个像素$\{\mathbf{x}_{ij} \mid \mathbf{x}^{ij} \in \mathbb{R}^{1\times1\times N_0}\}^{H_0,\ W_0}_{i,\ j=1}$并将其投影到大小相同的 CH_4 光谱吸收特征向量上 [13]。这样做可以减少地面地形的干扰，并增强该像素中 CH_4的可见度。鉴于 SLF 旨在减少地面地形干扰，准确地建模 SLF 至关重要。为了建模**SLF**，我们采用信息论中最常见的匹配滤波方法 [47]。

光谱线性滤波器（SLF）： SLF的设计取决于 CH_4 气体的光谱吸收模式 [13]以及地面地形的分布。由于我们感兴趣的信号CH_4, 非常微弱，传统的线性滤波方法 [45, 46] 并不有效。对地面地形噪声进行白化的传统方法包括：通过选取一组相邻的 10-15 列$\{\mathbf{x}_i \mid \mathbf{x}_i \in \mathbb{R}^{1\times H_0\times N_0}\}_{i=1}^{W_0}$来计算背景的协方差（**Cov** $\in \mathbb{R}^{N_0\times N_0}$）。然而，在给定的飞行航线上，地形会频繁变化，从水体到裸土、植被、建筑物及其他城市结构。因此，单一的协方差近似无法正确估计 CH_4，而基于局部上下文的白化会更加有效。为了解决这一问题，我们采用了一种非常简单且有效的方法：进行土地覆盖分类和分割 [11, 35, 49],，然后根据土地覆盖计算每个类别的协方差。更多细节请参见补充材料。这提高了存在混淆因素（具有与CH_4相似光谱吸收模式的材料）时甲烷候选目标的质量，也适用于 CH_4 浓度较低的情况。最终采用每类协方差的**SLF** 设计如下：

$$\mathbf{SLF}(\mathbf{x}_{ij}) = \frac{(\mathbf{x}_{ij} - \mu_k)^T \mathbf{Cov}_k^{-1} t}{\sqrt{t^T \mathbf{Cov}_k^{-1} t}} \ \forall\ (i,j) \in \text{class } k \quad (2)$$

其中 t 表示CH_4 气体的光谱吸收模式 [13]，而$\mathbf{Cov}_k$、μ_k 分别是k^{th} 类的协方差和均值。$\mathbf{x}^{ij}$ 表示输入高光谱图像中位于(i,j)索引处 k^{th} 类的像素。此操作生成大小为$\mathbb{R}^{H_0\times W_0}$的二维空间$CH_4$ 候选图。接下来将该 CH_4 候选图馈送到特征提取器以生成CH_4 候选特征图f_{mc}。土地覆盖分割/分类和完整SLF推导的细节见补充材料。

$$f_{mc} = \text{特征提取器}(\ \mathbf{SLF}(\mathbf{x}^{ij})\ \ \forall\ i,j) \quad (3)$$

3.5. 查询细化器 (QR)

接下来，将甲烷候选特征图$f_{mc} \in \mathbb{R}^{H\times W\times d}$ 与一组100个可学习查询$Q \in \mathbb{R}^{100\times d}$一起馈送到**QR** 模块。$f_{mc}$ 通过交叉注意力机制细化可学习查询。此操作为查询提供了狭窄的搜索空间。**QR** 模块遵循受 [20,21]启发的类似Transformer解码器的架构。随机初始化的查询$Q \in \mathbb{R}^{100\times d}$首先通过自注意力层关注自身。接下来，这些查询通过交叉注意力层关注来自**SFG** 模块的甲烷候选特征图f_{mc}。甲烷候选特征图在我们的注意力架构中充当键值对。**QR** 的输出为Q^{ref}。

$$Q_{ref} = \mathbf{QR}(f_{mc}, Q) \quad (4)$$

3.6. 高光谱解码器

Q^{ref} 与编码器输出f_e 一起送入解码器模块以生成输出嵌入。我们的高光谱解码器遵循标准架构，仅有微小差异。没有自注意力层，仅堆叠多头交叉注意力层。细化后的查询被转换为输出嵌入$E_{out} \in \mathbb{R}^{100\times d}$。

$$E_{out} = \text{Decoder}(f_e, p, Q_{ref}) \quad (5)$$

3.7. 边界框与掩膜预测

解码器输出嵌入（E_{out}）被送入两个前馈网络（FFN）和一个掩码预测层。FFN的输出是覆盖每个 CH_4 羽流的边界框以及每个框对应的置信度分数。掩膜预测模块遵循DETR的标准分割头 [4]。它计算每个嵌入在f_e 上的多头注意力分数（公式1），为每个嵌入生成低分辨率热图。为了进行最终预测，使用了类似特征金字塔网络 [15]的结构。每个热图旨在捕获一个甲烷羽流。使用简单的阈值处理将热图合并为最终分割掩码。

$$mask = \text{Mask_pred}(E_{out}, f_e, f_{comb}) \quad (6)$$

3.8. Training and Inference

我们分两个阶段训练 MethaneMapper：首先训练与每个 CH_4 羽流对应的边界框检测；其次冻结边界框检测网络，仅训练掩码预测模块。我们还对边界框和掩膜预测模块进行端到端训练，取得了相近的性能。我们采用与 DETR [4] 类似的两阶段损失策略来训练 MethaneMapper：

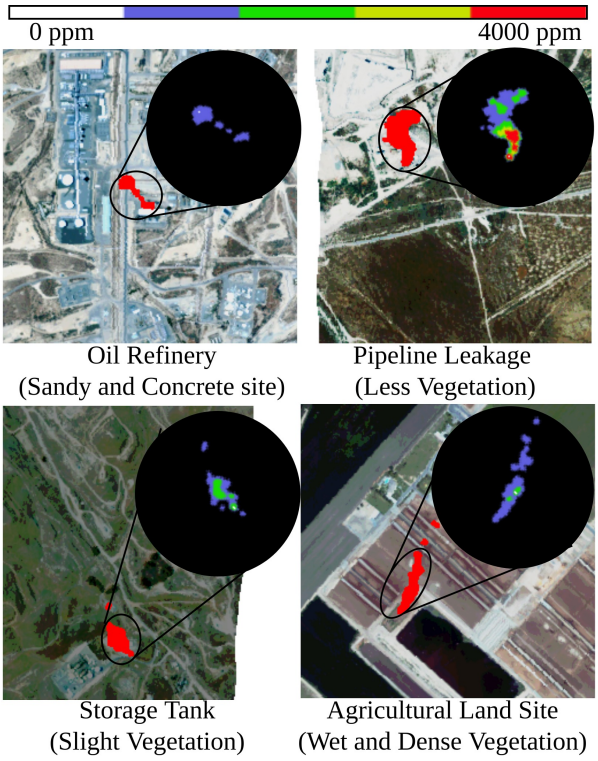

图 3. *MHS* 数据集中的示例图像。黑色圆圈内的色标显示了与红色所示羽流掩膜相对应的浓度图。图中展示了不同类型的泄漏源和土地覆盖类型。为便于观察，我们将二值掩膜叠加在由高光谱图像的可见光波段生成的彩色图像上。

第一阶段是在边界框和掩膜预测方面，对预测结果与真值进行二分图匹配；第二阶段则是计算匹配对的损失。二分图匹配采用匈牙利算法 [4]，以找出预测结果与真值之间的最优匹配。完成匹配后，每个预测结果都会与一个真值相关联。接下来，我们计算 l_1 和 $GIoU$ 边界框和掩膜预测的损失，以及类别预测的交叉熵损失 [4]。

推理： 推理管道与训练管道类似，可通过约 50 行代码实现。在推理过程中，我们首先过滤掉置信度低于 50% 的检测结果，并对每个像素取最大值，以确定哪些像素被预测为属于 CH4 羽流。

4. 甲烷热点（MHS）数据集

Another significant contribution of this work is a large scale curated MHS dataset. It contains the AVIRIS-NG spectral data with wavelength ranging from $380nm$ to $2510nm$, a $5nm$ sampling [22], and capturing 432 channels per pixel. The images from the flight-line are orthorectified

数据集	MHS（我们的方法） 数据集	JPL-CH4 detection-V1.0 [44]
# plume sites	**3961**	161
# flightlines	**1185**	46
# point source	**3675**	114
# diffused source	**286**	57
Time period	**2015 - 2022 (8 years)**	2015 (1 year)
Segmentation Mask	**Yes**	**Yes**
Bonding box	**Yes**	No
Concentration map	**Yes**	No
Number of Regions	**6**	1

Table 1. *Statistics shows MHS dataset comparison with JPL-CH4-detection-V1.0 [44] dataset. Each flightline have multiple large and small plume sites. Each flightline have atleast 4 plume sites. The Point Source represents high concentration ($300kg/hr$) to leakage from sources like pipeline leak, storage tanks, oil and gas refineries. Diffused Source represent low concentration leakages from sources like biomass degradation in landfills. Our dataset is covers more diverse type of terrain over 6 states.*

and of size $\sim 23K\times \sim 1.5K \times 432$. The only currently publicly-available dataset with methane plume segmentation masks is the JPL-CH4-detection-V1.0 [44] dataset released by JPL-NASA in 2017.

The MHS dataset has approximately 4000 plume sites corresponding to approximately 1200 AVIRIS-NG flightlines as shown in Table 1. MHS also has higher diversity data with flight lines spanning from 2015-2022 and covering terrain from 6 states– California, Nevada, New Mexico, Colorado, Midland Texas, and Virginia.

数据裁剪： 我们选择了覆盖不同区域的 AVIRIS-NG 飞行航线，因为其涵盖了多种 CH_4 羽流源，例如石油和天然气炼厂泄漏、石油和天然气开采点、自然渗漏、地下储- 存罐泄漏、煤矿、奶牛场、填埋场以及管道泄漏。除了不同的排放源外，我们还选择了具有不同类型地面地形的区域，例如裸土、岩石、山地、稀疏植被、水体和茂密植被，如图 3 所示。不同类型的地面地形表现出差异很大的反照率，因此会对 CH_4 检测质量产生重大影响，如图 4所示。因此，使用多样化的地面地形数据训练模型，可以得到更加稳健的模型。

4.1. Concentration map and Segmentation mask

浓度图 以与飞行航线相同的空间维度矩阵形式提供（$\sim 23k\times \sim 1.5k \times 1$）。每条飞行航线（经过正射校正）对应一张浓度图。它显示地面上每个像素的甲烷浓度，单位为百万分之一（ppm）。不含甲烷的像素区域设置为零。

分割掩码 以 *"PNG"*图像文件的形式提供，具有三个通道，空间维度与相应的飞行航线相同（$\sim 23k\times \sim 1.5k \times 3$）。分割掩码是通过将浓度掩膜文件中所有大于零的像素值设置为表示甲烷羽流而获得的。我们手动标注了 点源和 扩散源 基于地面地形类型和甲烷气体浓度。遵循基准数据集 [44], 使用三个通道进行颜色编码 点源 （红色）和 扩散源（绿色）。 点源 和 扩散源 的区分源自JPL-CH4-detection-V1.0基准数据集 [44]。我们的标注样式也与JPL-CH4-detection-V1.0基准数据集一致 [44],以便两个数据集可以无缝合并。

4.1.1 Constructing Concentration map

通过将专家标注的甲烷羽流浓度图映射到经过正射校正的AVIRIS-NG飞行航线，生成浓度图。这些甲烷羽流标注系统地收集自一家非营利 [3]实体。它们提供了甲烷排放的浓度掩膜，涵盖 150×150 尺寸的图块，并附带来自不同来源的位置信息（机载传感器 [6], 卫星 [38]）。为了将来自不同来源的这些图块映射到AVIRIS-NG飞行航线，我们使用标注和飞行航线所提供的像素坐标位置。我们利用这些信息创建单应性变换，将每个像素映射到飞行航线中的对应位置。图3 展示了不同类型地形的样例，其中CH_4 分割掩码以红色显示，浓度掩膜以黑色圆圈显示。关于分辨率匹配、正射校正和变换的详细信息，请参见补充材料。专家通过在排放实际地点进行现场访问，对图块标注进行了核验 [2]。加利福尼亚州的大多数区域均由加利福尼亚空气资源委员会通过现场访问进行核验 [2, 3]。

4.2. MHS 统计数据

MHS统计数据和属性总结于表1。**标注：** MHS 提供分割掩膜和浓度图，从而能够开发可同时生成 CH_4 羽流位置和浓度预测的深度学习算法。

多样性： MHS 数据集包含来自美国六个州的 AVARIS-NG 飞行航线，跨越 8 年（2015年 - 2022年）：加利福尼亚州、内华达州、新墨西哥州、科罗拉多州、得克萨斯州和弗吉尼亚州。

数据划分： 我们将 MHS 数据集划分为训练集和测试集，比例为 80：20，且两个集合涵盖重叠的时间段和地点。我们的数据集覆盖 6 个州。每个州包含多个子区域/地点（例如二叠纪盆地），这些区域由多条互不重叠的航线覆盖（$25k \times 1.5k \times 432$ 像素）。这些航线被划分到训练集和测试集中。在每个集合中，我们根据对应的航线创建图块（$256\times256\times432$ 像素）。从这些图块/切片中，我们选取所有正样本图块（甲烷（CH_4））， 并随机抽取数量相等的负样本（无 CH_4）图块。训练集和测试集分别执行此操作，以平衡数据；详细的消融实验请参见第6.2节。

5. 实验设置

评估指标：遵循 H-mrcnn 的评估协议 [31]我们以平均交并比（mIOU）报告性能。此处，mIOU 表示预测与真值 CH_4羽流掩膜之间的重叠。ED 表示羽流核心预测的准确性。此外，由于我们两阶段训练过程的第一阶段包含边界框预测，我们还以平均精度均值（mAP）报告羽流边界框的预测性能，这体现了 MethaneMapper 在消除羽流预测中假阳性方面的有效性。

Data Pre-Processing: Each input hyperspectral image is approximately of size $25000 \times 1500 \times 432$ taking up memory space of $55 - 60$ GB. We create tiles of each image in spatial domain, each tile is of size $256 \times 256 \times 432$ [31] with an overlap of 128. The CH_4 plume is available in very few pixels in the whole image, 90% of the tiles are negative samples (no methane, just ground terrain). We can not use the whole hyperspectral image because of GPU memory limitations

实现细节：波段选择器模块以432通道高光谱图像作为输入，RGB波段选择器从 $400nm - 700nm$ 波长范围中挑选60个通道并创建3通道RGB图像，SWIR波段选择器从波长范围$2000nm - 2500nm$中挑选100个通道。这些输入图像被传递到两个ResNet-50 [17]特征提取器骨干网络。骨干网络使用在COCO数据集 [32] 上训练的DETR [4] 进行初始化，输入层随机初始化 [16]。Transformer编码器-解码器和我们的查询细化器有6层和8个头。我们使用从DETR [4]模型中提取并剥离的权重初始化Transformer编码器-解码器。Transformer架构的维度为256，查询数量为100。**SFG** 模块接收所有432通道高光谱图像，并生成与输入相同空间维度的1通道输出图。**SFG** 中的特征提取器是ResNet-50 [17]，使用在COCO数据集 [32]上训练的DETR [4] 初始化。解码器输出嵌入的大小为512。掩码预测模块中的特征金字塔网络有3层。更多细节在补充材料中提到。

	方法	Back bone	SFG F.Ext.	#参数	mAP	mIOU
	JPL-CH4-detection-v1.0 数据集					
1	Hu 等人	R-50	-	75M	0.26	0.48
2	H-mrcnn	R-50	-	353M	0.53	0.86
3	MM	R-50	R-50	**80M**	**0.63**	**0.91**
	MHS (Ours) Dataset					
4	SpectralFormer	R-50	-	84M	0.33	0.41
5	UPSNet (stuff)	R-50	-	69M	0.32	0.38
6	UPSNet (stuff + things)	R-50	-	69M	0.29	0.35
7	DETR	R-18	*	33M	0.37	0.56
8	DETR	R-50	*	59M	0.44	0.59
10	MM	R-18	Linear Layer	39M	0.45	0.60
11		R-18	R-18	44M	0.52	0.63
12		R-50	R-50	**80M**	**0.59**	**0.68**

表2。 与基线方法的比较。*"-"*表示不适用，*"*"*表示无 ***SFG*** 模块，以及用于*Transformer*解码器的随机查询。顶部部分展示了 *JPL-CH*$_4$ 数据集 *[44]*上的性能。*MethaneMapper*以更少的参数取得了优于经过大量调优的*H-mrcnn*的结果。该数据集上的总体检测准确率更高，这是因为所有航线中的地面地形类型都具有统一性。在*MHS* 数据集上，*MM*的表现优于多个基线方法，如第*4*至第*12*行所示。*MM*在*MHS*中的准确率低于*JPL-CH*$_4$ 数据集，因为*MHS* 数据集包含分布于*6*个州的更多样化的地面地形。

Methods	mAP	mIOU
LogReg [5]	-	0.05
SVM [40]	-	0.29
PCA + LogReg	-	0.06
PCA + SVM	-	0.31
MM（R-50）	**0.63**	**0.91**

表 3. 与经典机器学习方法的比较。*"-"*表示不可用。经典机器学习方法不适用于 *CH*$_4$ 检测任务。*MethaneMapper* 在 *JPL* 数据集 *[44]*上的表现优于所有方法

6. 结果

本节将通过消融实验讨论并验证 MethaneMapper（MM）的所有设计选择。我们将展示，MM 的整体性能达到最先进水平，优于表2 & 3中展示的所有其他方法。

6.1. 性能比较

深度学习方法：我们使用 ResNet-50 [17]骨干网络训练 MM，所用数据集与 H-mrcnn [31](JPL-CH4-detection-V1.0 [44]) 的训练数据集相同，以便公平比较。为与 H-mrcnn 保持一致，我们使用了相同的数据划分和输入图像尺寸。拥有 80M 个参数的 MM 模型经过 250 个周期训练后，性能显著优于拥有352M 个参数的 H-mrcnn 模型。结果汇总于表 2 中，其中包括 MM 在新的更大规模 MHS 数据集上的性能。我们注意到，尽管 H-mrcnn 的代码可用，但其中许多模块已弃用，无法复现。"骨干网络"列 umn represents backbones used for feature extraction from input image,'SFG F.Ext.' represents the feature extractor in **SFG** module in MethaneMapper. We observed (qualitatively) that H-mrcnn fails to detect small CH_4 plumes with concentration lower than 100kg/hr while MM detects those.

我们通过实现 3 种基线模型 [4, 18, 50] 进行了评估，见表2第 4–8 行。这些方法并非为 CH_4 检测任务设计，因此我们需要调整其输入通道数。这些方法性能较差，可能是由于高维数据中目标信号较弱、混淆因素数量较多，以及标注数据有限。此外，唯一的高光谱基线方法 SpectralFormer [18]由于采用逐像素训练方案，效率较低。

经典机器学习方法： 我们训练并测试了多种现有的、基于机器学习的甲烷检测方法，其性能如表3所示。逻辑回归（LogReg） [5]和多项逻辑回归（MLR） [24]均未能实现有意义的检测，并产生了90% 的假阳性检测结果。我们还训练了一个支持向量机（SVM） [40,41]分类器，其表现略优于 LR 和 MLR 方法，IOU 达到 21%。与高斯混合模型一样，SVM 容易产生假阳性检测结果 [40]。我们观察到，所有传统方法都不适用于 CH_4 检测任务。我们还测试了使用主成分分析（PCA）降低维度，或仅选取显示出最大 CH_4 吸收的波段。在后一种情况下，传统方法的表现优于使用全部 432 个波段，这进一步支持了我们仅使用 SWIR 区域波段的想法。

Qualitative results. Fig. 5 shows comparison of MM's mask and bounding box prediction with ground truth mask on different ground terrains. The Leakages are from different type of sources such as, oil refinery, pipeline and storage tank. MM makes correct predictions in varying scenarios.

6.2. Ablation Studies

We did the experiments for ablation on MHS dataset with ResNet-50 as backbone and validate the design choices. One parameter is changed for each ablation and others kept at best settings. More ablations in Supplementary.

光谱特征生成模块： 在表2 下半部分中，我们展示了我们的 **光谱特征生成模块** 对查询细化模块的有效性。我们的基线是标准的 DETR 实现 [4] ，用于分割任务，由表2 下半部分的第 1 行和第 2 行表示。使用 CH_4来自 **光谱特征生成模块**的候选特征可将边界框检测性能提高 0.14 平均精度均值，并将掩膜预测提高 0.09 平均交并比。这表明，使用由 $_4$ 生成的CH**候选特征来引导查询，相较于随机查询可产生更好的嵌入**。

此外，我们还探索了在两个位置提供 CH_4 候选特征：（i）在输入层进行拼接。

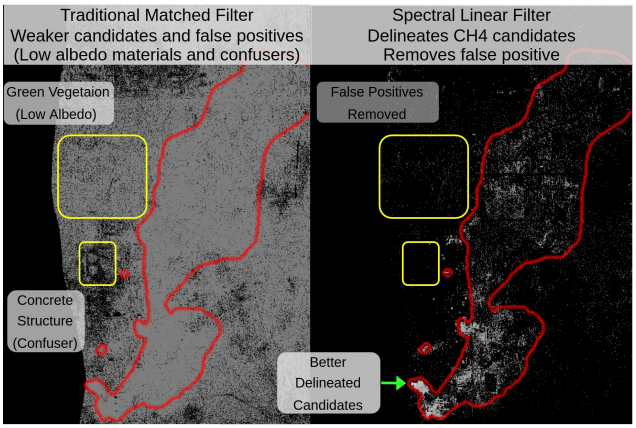

Figure 4. *Comparison of SLF with traditional filter in SFG module. White pixels represent methane and black no-methane. Red boundary represents ground-truth plume mask. SLF module generates better CH_4 candidates*

将其与 f_{comb} 结合；以及 (ii) 将其作为查询细化器的输入。当 **SFG** 模块的输出传递给查询细化器时，平均精度均值提高了 0.09，平均交并比提高了 0.08。我们推测，这是因为与输入特征拼接时，CH_4 候选特征信息会丢失；而使用查询进行交叉注意力则会缩小解码器的搜索空间，并生成更好的嵌入。

We also experimented with different types of feature extractors for **SFG** module, and observed that a Resnet18 or Resnet50 [17] is more effective than a 2 linear layer feature extractor as shown in Table 2.

光谱线性滤波器：我们实验了使用 SLF 计算协方差（*Cov*），采用输入高光谱图像中不同列子集。我们观察到，当基于土地覆盖按类别计算协方差时，SLF 最为有效。按类别 *Cov* 确保在计算 CH_4 增强时，地面地形的辐射吸收对所有像素都相同。如图 4 所示，SLF 放大了 CH_4 候选检测并减少假阳性。与传统滤波器相比，SLF 使检测的平均精度均值提高了 0.03。MM 的预测显示在图5 的第 1 行。

地理泛化： 为了评估 MM 的地理泛化能力，我们使用除加利福尼亚州以外所有州的 MHS数据对其进行训练，并在加利福尼亚州的航线数据上进行测试。我们观察到检测结果的平均精度均值略微下降了 0.04。然而，当使用除弗吉尼亚州以外的所有数据进行训练时，我们发现检测结果的平均精度均值显著下降了 0.09。我们认为，这是因为弗吉尼亚州的土地覆盖以茂密潮湿的植被为主，与加利福尼亚州、得克萨斯州和内华达州的干旱地区相比，其太阳反射率较低。

Temporal generalization: Testing MM on 2015 after training on data from 2016-2022 showed no performance drop.

Unbalanced test set: MM's performance dropped by 0.05 mAP on an unbalanced test set with only 10% positive samples (CH_4) and 90% negative samples (no-CH_4). This highlights the challenges in CH_4 detection. Future work will address this issue.

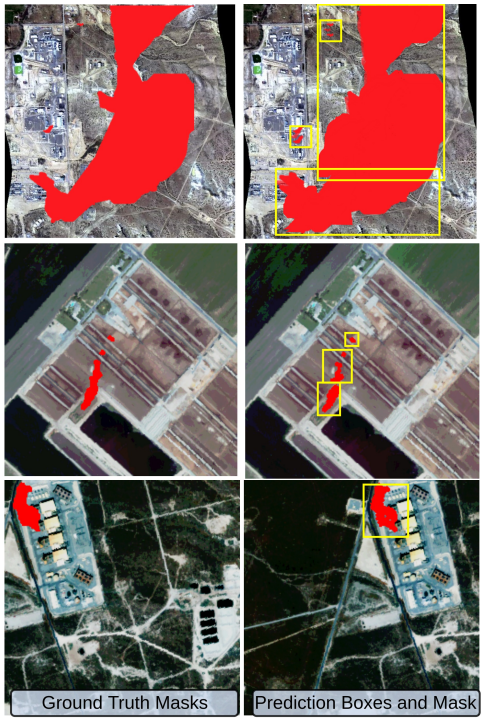

Figure 5. *Sample ground truths and predictions on MHS dataset. We show robustness of MethaneMapper predictions on different kind of ground terrain, rows 1 and 3 shows leakage at a refinery, row 2 shows leakage from pipeline in agricultural land, row 4 shows leakage from storage tank with concrete background.*

7. Conclusion

本文介绍了 MethaneMapper——一种用于甲烷羽流检测的高光谱Transformer。它利用光谱特征生成器和查询细化器来捕捉光谱和空间相关性，从而准确描绘 CH_4羽流。此外，我们为该任务整理了一个大规模数据集，这是同类数据集中的首个，未来将向所有研究人员开放。正如我们广泛的实验所表明的那样，所提出的 MethaneMapper在检测和定位精度方面显著优于当前方法。未来的工作将使用多光谱卫星成像数据将该模型扩展至全球监测 [29]。

8. Acknowledgments

本研究部分得到以下资助的支持：美国国家科学基金会（NSF）项目SI2-SSI #1664172，以及美国陆军研究实验室（ARL）协议编号 W911NF2020157。

References

[1] Alana K. Ayasse、Andrew K. Thorpe、Dar A. Roberts、Christopher C. Funk、Philip E. Dennison、Christian Frankenberg、Andrea Steffke 和 Andrew D. Aubrey。使用合成机载可见光/红外成像光谱仪下一代（AVIRIS-NG）图像评估地表属性对甲烷反演的影响。《环境遥感》，215:386–397，2018。1[2] 加利福尼亚空气资源委员会。加利福尼亚空气资源委员会发布的温室气体清单，2022。6[3] CARBON MAPPER 公司。Carbon Mapper 甲烷排放探测，2022。2, 6[4] Nicolas Carion、Francisco Massa、Gabriel Synnaeve、Nicolas Usunier、Alexander Kirillov 和 Sergey Zagoruyko。使用 Transformer 进行端到端目标检测。载于《欧洲计算机视觉会议》，第213–229页。Springer，2020。1, 3, 4, 5, 6, 7[5] Qi Cheng、Pramod K Varshney 和 Manoj K Arora。用于遥感数据特征选择和软分类的逻辑回归。《IEEE地球科学与遥感快报》，3(4):491–494，2006。7[6] E. Knapp David、Heckler Joseph、Seely Megs 和 P. Asner Gregory。全球机载观测站可见光至红外成像光谱仪报告，2020。2, 6[7] C Frankenberg、U Platt 和 T Wagner。用于反演强吸收痕量气体的迭代最大后验概率（IMAP）-DOAS：基于模型的研究，用于从搭载于 Envisat 的 SCIAMACHY 近红外光谱中反演 CH_4 和 CO_2。《大气化学与物理讨论》，4(5):6067–6106，2004。2[8] Christian Frankenberg、U Platt 和 T Wagner。用于反演强吸收痕量气体的迭代最大后验概率（IMAP）-DOAS：基于模型的研究，用于从搭载于 Envisat 的 SCIAMACHY 近红外光谱中反演 CH_4 和 CO_2。《大气化学与物理》，5(1):9–22，2005。1, 2[9] Christian Frankenberg、Andrew K Thorpe、David R Thompson、Glynn Hulley、Eric Adam Kort、Nick Vance、Jakob Borchardt、Thomas Krings、Konstantin Gerilowski、Colm Sweeney 等。机载甲烷遥感测量揭示四角地区的重尾通量分布。《美国国家科学院院刊》，113(35):9734–9739，2016。1, 2[10] Christopher C Funk、James Theiler、Dar A Roberts 和 Christoph C Borel。通过聚类改进高光谱热成像中弱气体羽流的匹配滤波检测。《IEEE地球科学与遥感汇刊》，39(7):1410–1420，2001。2, 12[11] Bo-Cai Gao。NDWI——用于从空间遥感植被液态水的归一化差异水体指数。《环境遥感》，58(3):257–266，1996。4, 13[12] Utsav B Gewali、Sildomar T Monteiro 和 Eli Saber。基于机器学习的高光谱图像分析：综述。arXiv 预印本 arXiv:1802.08701，2018。2[13] IE Gordon、LS Rothman、RJ Hargreaves、R Hashemi、EV Karlovets、FM Skinner、EK Conway、C Hill、RV Kochanov、Y Tan 等。HITRAN2020 分子光谱数据库数据库。《定量光谱学与辐射传输杂志》，277:107949，2022。4,11, 12[14] L.Hamlin、R. O.Green、P.Mouroulis、M.Eastwood、D. Wilson、M. Dudik 和 C.Paine。用于陆地生态学的成像光谱仪科学测量：AVIRIS 及其最新发展。载于2011年航空航天会议，第1–7页，2011年。11[15] KaimingHe、Georgia Gkioxari、Piotr Doll´ar 和 Ross Girshick。Mask R-CNN。载于IEEE国际计算机视觉会议论文集，第2961–2969页，2017年。4[16] Kaiming He、张祥雨、任少卿和孙剑。深入研究整流器：超越ImageNet分类的人类水平性能。载于IEEE国际计算机视觉会议论文集，第1026–1034页，2015年。6[17] Kaiming He、张祥雨、任少卿和孙剑。用于图像识别的深度残差学习。载于IEEE计算机视觉与模式识别会议论文集，第770–778页，2016年。3, 4,6, 7, 8[18] 洪丹枫、Zhu Han、Jing Yao、Lianru Gao、Bing Zhang、Antonio Plaza 和 Jocelyn Chanussot。SpectralFormer：利用Transformer重新思考高光谱图像分类。《IEEE地球科学与遥感汇刊》，60:1–15，2021年。2,7[19] 巴黎国际能源署。石油与天然气行业的甲烷排放，2020年。1[20] ASMIftekhar、Hao Chen、Kaustav Kundu、XinyuLi、JosephTighe和Davide Modolo。看什么以及在哪里看：用于检测人-物交互的语义与空间细化Transformer。载于IEEE/CVF计算机视觉与模式识别会议论文集，第5353–5363页，2022年。3,4[21] ASM Iftekhar、Satish Kumar、R Austin McEver、SuyaYou和 BS Manjunath。GTNet：用于检测人-物交互的引导式 Transformer网络。arXiv预印本，arXiv:2108.00596，2021年。3,4[22] 加州理工学院喷气推进实验室。机载可见光红外成像光谱仪-下一代（AVIRIS-NG）概述，2009年。1, 2, 5, 11[23] Siraput Jongaramrungruang、Andrew K Thorpe、GeorgiosMatheou和 Christian Frankenberg。MethaneT——一种人工智能驱动方法，用于从高分辨率二维羽流图像中量化甲烷点源排放。《环境遥感》，269:112809，2022年。1, 2[24] Mahdi Khodadadzadeh、Jun Li、Antonio Plaza和Jos´eMBioucas-Dias。用于高光谱图像分类的基于子空间的多项逻辑回归。《IEEE地球科学与遥感快报》，11 2109，2014 12 2105 7 ( ): – 。[25] Stefanie Kirschke、Philippe Bousquet、Philippe Ciais、Marielle Saunois、JosepG Canadell、EdwardJ Dlugokencky、Peter Bergamaschi、Daniel Bergmann、Donald RBlake、Lori Bruhwiler 等。三十年来的全球甲烷源汇。《自然·地球科学》，6(10):813–823，2013年。1[26] FJ Kriegler、WA Malila、RFNalepka 和 WRichardson。预处理变换及其对多光谱识别的影响。《环境遥感》，VI，第97页，1969年。13。

[27] Satish Kumar、ASM Iftekhar、Michael Goebel、Tom Bullock、Mary H MacLean、Michael B Miller、Tyler Santander、Barry Giesbrecht、Scott T Grafton 和 BS Manjunath。Stressnet：检测热成像视频中的压力。在《IEEE/CVF冬季计算机视觉应用会议》论文集，第999–1009页，2021年。3[28] Satish Kumar、ASM Iftekhar、EktaPrashnani 和 BS Man-junath。Locl：使用定位学习对象-属性组合。arXiv预印本 arXiv:2210.03780，2022年。3[29] Satish Kumar、William Kingwill、Orbio Earth、Rozanne Mouton、Wojciech Adamczyk、Robert Huppertz和EvanSherwin。用于检测Sentinel-2卫星图像中甲烷排放的引导式Transformer网络。8[30] Satish Kumar、Rui Kou、Henry Hill、Jake Lempges、Eric Qian和VikramJayaram。石油和天然气运营中的原位水质监测。arXiv预印本arXiv:2301 08800，2023年 2 。 .[31] Satish Kumar、Carlos Torres、Oytun Ulutan、Alana Ayasse、DarRoberts和BS Manjunath。用于从机载高光谱图像中检测甲烷的深度遥感方法。在《IEEE/CVF冬季计算机视觉应用会议》论文集，第17761785页，2020年 1,2, 6, 7, 12 –。[32] Tsung-Yi Lin、MichaelMaire、Serge Belongie、JamesHays、PietroPerona、Deva Ra manan、PiotrDoll´ar 和 CLawrence Zitnick。Microsoft coco：情境中的常见对象。在欧洲计算机视觉会议上，第740–755页。Springer，2014年。6[33] IlyaLoshchilov和 FrankHutter。解耦权重衰减正则化。arXiv预印本 arXiv:1711.05101，2017年。14[34] META-STANFORD。甲烷排放技术联盟（meta），斯坦福天然气倡议，2022年。2[35] MGISGeography。NDVI（归一化植被指数），1979年。4, 13[36] Gunnar Myhre、DrewShindell和JuliaPongratz。人为和自然辐射强迫。2014年。1[37] NathaliePettorelli。归一化植被指数。牛津大学出版社，2013年。13[38] Darius Phiri、Matamyo Simwanda、Serajis Salekin、Vin- cent R Nyirenda、Yuji Murayama 和Manjula Ranagalage。用于土地覆盖/土地利用制图的Sentinel-2数据：综述。《遥感》，12(14):2291，2020年。2, 6[39] Dar A Roberts、Eliza S Bradley、Ross Cheung、Ira Leifer、PhilipEDennison和 JackS Margolis。利用成像光谱学绘制海洋地质渗漏源的甲烷排放图。《环境遥感》，114(3):592–606，2010年。1, 2[40] CAShah、PKVarshney 和 MKArora。用于遥感图像无监督分类的ICA混合模型算法。《国际遥感杂志》，28(8):1711–1731，2007年。2, 7[41] David MJ Tax 和 Robert PW Duin。支持向量数据描述。《机器学习》，54(1):45–66，2004年。2, 7[42] James Theiler、BernardRFoy和AndrewM Fraser。超越自适应匹配滤波器：高维杂波中弱信号的非线性检测器。在《算法与《多光谱、高光谱和超光谱影像技术 XIII》，第6565卷，第26–37页。SPIE，2007年。12[43] DR Thompson、I Leifer、H Bovensmann、M Eastwood、M Fladeland、C Frankenberg、K Gerilowski、RO Green、S Kratwurst、T Krings等。机载成像光谱的实时遥感探测与测量：甲烷案例研究。《大气测量技术》，8(10):4383–4397，2015年。1，2[44] David R Thompson、Anuj Karpatne、Imme Ebert-Uphoff、Christian Frankenberg、Andrew K Thorpe、Brian D Bue和Robert O Green。ISGeo数据集 jpl-ch4-detection-2017-v1.0：基于成像光谱仪数据的甲烷源探测基准。2017年。2，5，6，7，11，12[45] Andrew K Thorpe、Christian Frankenberg、David R Thompson、Riley M Duren、Andrew D Aubrey、Brian D Bue、Robert O Green、Konstantin Gerilowski、Thomas Krings、Jakob Borchardt等。利用AVIRIS-NG在高空间分辨率下反演机载DOAS的甲烷、二氧化碳和水汽浓度：应用。《大气测量技术》，10(10):3833–3850，2017年。1，2，4[46] Andrew K Thorpe、Dar A Roberts、Eliza S Bradley、Christopher C Funk、Philip E Dennison和Ira Leifer。使用聚类调谐匹配滤波技术和成像光谱学高分辨率绘制海洋与陆地源的甲烷排放图。《环境遥感》，134:305–318，2013年。1，2，4[47] George Turin。匹配滤波器导论。《IRE信息论汇刊》，6(3):311–329，1960年。4[48] Oytun Ulutan、ASM Iftekhar和Bangalore S Manjunath。VSGNet：利用图卷积检测人体—物体交互的空间注意力网络。载于《IEEE/CVF计算机视觉与模式识别会议论文集》，第13617–13626页，2020年。3[49] 美国农业部外国农业服务局。归一化差异植被指数（NDVI），1969年。4，13[50] Yuwen Xiong、Renjie Liao、Hengshuang Zhao、Rui Hu、Min Bai、Ersin Yumer和Raquel Urtasun。UPSNet：统一的全景分割网络。载于《IEEE/CVF计算机视觉与模式识别会议论文集》，第8818–8826页，2019年。7[51] Xizhou Zhu、Weijie Su、Lewei Lu、Bin Li、Xiaogang Wang和Jifeng Dai。Defor mable DETR：用于端到端目标检测的可变形Transfor mer。arXiv预印本 arXiv:2010.04159，2020年。14

9. SUPPLEMENTARY MATERIALS

In supplementary section, we provide will all the details about the data collection and annotations creation process. We also provide with the complete derivation of Spectral Linear Filter (**SLF**) along with a pseudo implementation of **SLF** algorithm. Next in the document we provide some more qualitative examples of success and failure cases of MethaneMapper. Towards the end of the document we provide graph plots about training convergence of all the ablation experiments with Spectral Feature Generator (SFG) and Query Refiner (QR) module.

9.1. Dataset

9.1.1 AVIRIG-NG

AVIRIS-NG [22] 是*AirborneVisible In-fraRedImaging Spectrometer - NextGeneration* 的缩写，由喷气推进实验室（JPL）于 2009 年开发。过去7年中，JPL 使用AVIRIS-NG仪器开展了数千条飞行航线的数据记录。在AVIRIS-NG 仪器上，总计有598 个按推扫顺序排列的传感器，捕获空间维度为$\sim 23k \times 598$的未经正射校正的数据立方体，其中每个传感器记录的光谱波长范围为 $380nm - 2510nm$ [14]，形成432 个通道。其具有 34° 视场，瞬时视场为 1 mrad，可根据飞行高度生成 $1-8m$ 的空间分辨率。随后使用几何查找表对这些数据进行校正，所得数据立方体的大小为 $\sim 23k\times \sim 1.5k \times 432$。数据以按行交错（BIL）顺序提供。BIL顺序表示该 3D 矩阵首先按图像行索引，然后按通道索引，最后按图像列索引 [44]。每个飞行航线文件夹中的“README.txt”文件包含命名约定以及每个文件所含数据类型的详细信息。可以使用 Python 库轻松将数据加载到 $numpy$ 数组中。所有数据均经过正射校正。

9.1.2 Annotations

变换和正射校正。 第一步是读取大小为 150×150 的甲烷浓度掩膜标注 GeoTiff图块，并将其坐标参考系（CRT）转换为 AVIRIS-NG飞行航线的 CRT（EPSG4326）。接下来，我们使用对应的AVIRIS-NG飞行航线几何查找表和未经正射校正的地理像素位置，生成该航线的正射校正地理像素位置数据。然后，我们找出与甲烷浓度掩膜（标注 GeoTiff）的地理索引最接近的飞行航线地理索引。最后，我们使用这些对应像素计算单应性变换矩阵，将甲烷浓度掩膜（标注 GeoTiff）映射至 AVIRIS-NG飞行航线的空间维度。我们重复此过程

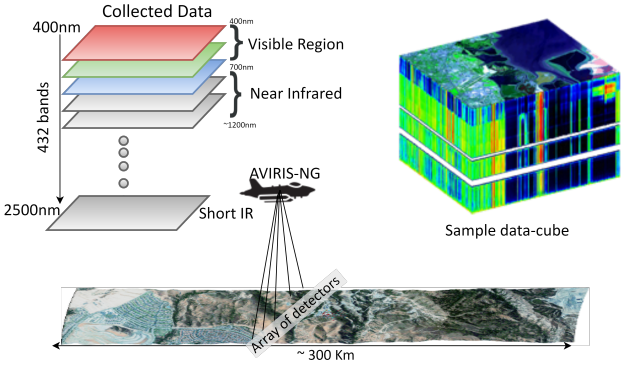

Figure 6. *Depiction of data collection process. Each flightline is* $\sim$ *300 kms long. An array of 598 sensors records data at 1.5m/pixel spatial resolution. All flightlines are ortho-corrected. Each data-cube is of dimension* $\sim 23k\times \sim 1.5k \times 432$.

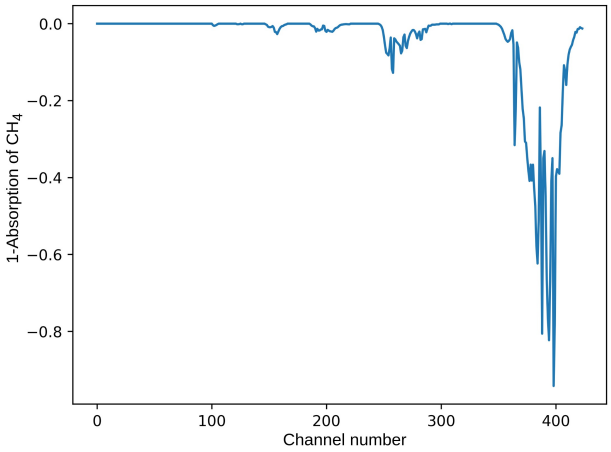

Figure 7. *Spectral absorption pattern of* CH_4 *gas. The x-axis show the channel number ranging from 0-400 corresponding to wavelength range* ($400nm - 2500nm$). *It is obtained from the public repository HITRAN [13].*

process for each plume in the flightline in order to generate the CH_4 concentration map for the entire flightline.

分辨率匹配。 为了使变换后的标注 GeoTiff补丁与AVIRIS-NG飞行航线的分辨率匹配，我们使用最近邻重采样。对于整条飞行航线，变换后的标注 GeoTiff补丁中的一个像素可能会在CH_4 浓度图中重复多次。

标注样式。点源和扩散源按照与JPL-CH4-detection-V1.0 [44] 数据集相同的标准进行编码。三通道的值范围为[0-255] 。

- 红色 (255,0,0)：羽流，被认为与一个点源有关
- 蓝色 (0,0,255)：羽流，据信与弥散源有关

- Black (0,0,0): no plume (or unlabeled)

我们使标注样式与 JPL-CH4- detection-V1.0基准数据集保持一致 [44] 以便将 JPL-CH4-detection-V1.0 和 MHS 数据集无缝合并。

9.2. Spectral Linear Filter(SFL)

9.2.1 Traditional Matched Filter

被动式高光谱成像传感器从 N_0 ($N_0 = 432$)个与波长范围为 $400nm - 2500nm$相对应的光谱通道中捕获光谱辐亮度值，如图6 所示，并给出了样本数据立方体。完整的高光谱图像表示为$\mathbf{x} \in \mathbb{R}^{H_0 \times W_0 \times N_0}$ 其中H_0、 W_0& N_0 分别为高度、宽度和通道数。在这些高光谱数据中，我们寻找隐藏在背景中的非常微弱的感兴趣特征。在本例中，感兴趣的特征是CH_4，而背景是地面地形。CH_4 在 $2100nm - 2500nm$ 波长附近显示出强烈的吸收模式。

用于查找CH_4 候选项的最常见线性方法是取一个 N_0维（与光谱通道数量相同）的向量 α，并将其与高光谱图像中的每个像素（N_0维）进行点积，以为每个像素生成一个标量输出。该操作旨在减少或去除地面地形、传感器噪声，并放大CH_4 特征。这个 α 向量称为“匹配滤波器”。因此，正确计算 α 对于生成更好的 CH_4 排放候选项至关重要。它取决于CH_4 的吸收模式以及地面地形的分布。为建模 α，设$\mathbf{r}_i \in \mathbb{R}^n$ 为高光谱图像中的一个 i^{th}像素，表示地面地形像素和传感器噪声，并设 $\mathbf{t}$为CH_4 吸收模式 [13]。这被建模为如下所示的加性扰动：

$$\mathbf{x}_i = \mathbf{r}_i + \mathbf{t}, \tag{7}$$

where $\mathbf{x}_i$ is the spectrum when CH_4 is present. The CH_4 absorption pattern $\mathbf{t}$ represents the change in radiance units of the background caused by adding a unit mixing ratio length of CH_4 absorption [10, 31]. Figure 7 shows the spectral absorption pattern of CH_4 per channel. In the ideal scenario where only CH_4 gas is present in signal (i.e. all white background), the matched filter output is $\alpha^T\mathbf{t}$. In case there is no gas and just ground terrain and sensor noise, the matched filter output is $\alpha^T\mathbf{r}_i$. The variance (Var) of $\alpha^T\mathbf{r}_i$ for latter is represented by:

$$Var(\alpha^T\mathbf{r}_i) = \langle(\alpha^T\mathbf{r}_i - \alpha^T\boldsymbol{\mu})^2\rangle = \alpha^T\mathbf{Cov}\alpha, \tag{8}$$

where $\mathbf{Cov}$ and μ are covariance and mean respectively computed for $\mathbf{r}_i$. Inspired from [10, 31] we define the Methane-to-Ground terrain Ratio (MGR) is:

$$\text{MGR} = \frac{|\alpha^T\mathbf{t}|^2}{\alpha^T\mathbf{Cov}\alpha}, \tag{9}$$

We can see that the magnitude of α does not affect MGR. According to [10, 31, 42], the MGR can be maximized subject to constraints(zero mean and $\alpha^T\mathbf{K}\alpha$ constraint to 1). The matched filter α is then represented by:

$$\alpha = \frac{\mathbf{Cov}^{-1}\mathbf{t}}{\sqrt{\mathbf{t}^T\mathbf{Cov}^{-1}\mathbf{t}}}. \tag{10}$$

In ideal instances when there is no background (i.e. all white background) and just CH_4 gas present. The matched filter in equation 11 is directly proportional to $\mathbf{t}$. This is just the target signature ($\mathbf{t}$) itself scaled so that the filtered output has variance of one. The methane enhancement per pixel can be computed as follows:

$$\hat{\alpha}(\mathbf{x}_i) = \frac{(\mathbf{x}_i - \mu)^T\mathbf{Cov}^{-1}\mathbf{t}}{\sqrt{\mathbf{t}^T\mathbf{Cov}^{-1}\mathbf{t}}}, \tag{11}$$

其中 $\hat{}\alpha(\mathbf{x}_i)$ 是每个像素的甲烷估计值，换句话说，是甲烷的柱浓度增强量。所使用的协方差矩阵（**Cov**）并非已知的 $prior$，而是根据数据估计得到的。它通过计算所有像素上经均值中心化的辐亮度的外积得到。换句话说，公式11 中的传统匹配滤波器计算的是地面地形的协方差（**Cov**），其基本假设是所有元素都具有相似的吸收模式。相同的协方差（**Cov**）矩阵用于对白化变化的地面地形，并放大其中存在的 CH_4。但在实际场景中，地面地形不断变化，地形类型经常改变；单幅图像中可能同时存在水体、裸土、植被、茂密植被、城市建筑结构、道路等。例如，水对太阳辐射具有很强的吸收作用，因此在此类背景上的甲烷可见性很弱。类似地，湿地和茂密植被也具有相似的行为。另一方面，裸土、岩石等的吸收较低，存在于此类背景中的甲烷具有很强的可见性。对地面分布的协方差（**Cov**）进行简单且单一的近似，无法对甲烷增强量作出正确而有效的估计。为解决这一局限性，我们开发了一种光谱线性滤波器（SLF），用于执行土地覆盖分类与分割，并降低噪声，具体内容将在后续章节中讨论。

9.2.2 Landcover Classification and Segmentation

在本节中，我们改进了上一节提到的局限性。我们首先选取可见光谱（$400nm - 700nm$）和近中红外区域（$800nm - 1350nm$）的高光谱波段。我们通过为每个颜色波段采用加权正态分布，重新构建了地面地形的RGB 表示。近红外区域也采用相同的处理方式。接下来，我们采用一种简单、非常有效且高效的方法进行土地覆盖分类—

检测与分割。我们计算归一化差异植被指数（NDVI）[35, 37] 和归一化差异水体指数（NDWI）[11]。NDVI 通过测量近红外（植被会强烈反射）与红光（植被会吸收）之间的差异来量化植被 [35]。其范围从 -1 到 $+1$。它是一种非常有效的指数，四十多年来一直见诸文献。 [11] 创建了 NDWI，并用它突出卫星图像中的开阔水体特征，使水体能够从土壤和植被中“凸显”出来。它使用绿光-近红外（可见绿光和近红外）计算，范围从 -1 到 $+1$。如今，它的主要用途是检测和监测水体含水量的细微变化。

$$ndvi = \frac{NIR - R}{NIR + R};\ \ ndwi = \frac{NIR - MIR}{NIR + MIR} \tag{12}$$

其中， NIR 是以$880nm$为中心归一化的近红外区域，MIR 是以 $1240nm$为中心归一化的中红外区域，而 R 是以 $660nm$为中心归一化的红色区域。我们利用这些指数创建了不同类型的植被、水体、裸土、岩石、山地、城市/城区、道路等的分割图。我们从 [26, 49]中获取分类阈值。为简化处理，我们还测试了将尺度 -1 至 $+1$ 划分为20个类别，每个类别的范围为 $< 0.1 >$。与使用 [26, 49]中的分类范围相比，我们获得了相当的结果。这种简单、有效且高效的方法使我们的光谱线性滤波器 CH_4 候选者估计提升了三倍。

9.2.3 Cov per class

我们采用上一步得到的分割图像，为简便起见，后文将分割图像称为分割掩码。实际处理中，我们有 20 个类别，每个类别对应一个分割掩码。如果某个类别中的像素数少于 10000，我们就将两个或更多相邻类别合并为一个。我们会确保每个类别中的像素数足够多，以便在计算协方差（**协方差**）矩阵时，甲烷信号不会产生影响，或产生的影响可以忽略不计。将相邻类别合并为一个是可行的，因为它们的辐亮度/反射率几乎相似，例如，稀疏植被和普通植被的反射率相似，等等。我们分别计算每个类别的均值和协方差矩阵。类别的协方差 **协方差**$_k$ of k^{th}计算如下：

$$\mathbf{Cov}_k = \frac{1}{N}\sum_{i=1}^{i=j}(\mathbf{x}_i - \mu_k)(\mathbf{x}_i - \mu_k)^T \ \forall\ j \in k, \tag{13}$$

where N is the number of pixels (> 10000) in k^{th} class and μ_k is the mean of kth class. For each class we compute the mean μ_k, covariance matrix $\mathbf{Cov}_k$. While iterating through each pixel of hyperspectral image, we check to which class k the pixel $\mathbf{x}_i$ belongs to and use those pre-computed values. The final Spectral Linear Fitler (**SLF**) is shown as below:

$$\mathbf{SLF}(\mathbf{x}_i) = \frac{(\mathbf{x}_i - \mu_k)^T\mathbf{Cov}_k^{-1}t}{\sqrt{t^T\mathbf{Cov}_k^{-1}t}}\ \forall\ (i) \in \text{class } k \tag{14}$$

其中$\mathbf{Cov}^{-1}$ 是协方差矩阵的逆。接下来，为了抑制传感器噪声，我们采用一种简单的方法来跟踪各个传感器。每个传感器都有不同的物理特性，这些特性可能会影响其采集的数据。我们在飞行航线上跟踪每个传感器。由于数据经过了校正，每个传感器的数据并非来自单独的一列，而是随机分布在所有列中。这取决于飞行路径以及飞机飞行过程中的运动。我们使用了深度优先搜索等简单的数据结构算法，跟踪每个边界像素，并将其分配给单个传感器。我们每次使用来自 10–15 个相邻传感器的数据，对其进行归一化，然后在前一步中结合分割掩码计算协方差矩阵。我们的方法非常简单直接。

算法1 展示了我们提出的光谱线性滤波器（**SLF**）的伪代码。

数据：MHS 数据集 结果：CH_4 浓度图初始化；对于 mhs中的每个 MHS，执行 1. 创建内存映射 mhs；2. 分割掩膜 = 计算分割掩码；对于分割掩膜中的掩膜，执行 data.append(mhs[掩膜]) 如果（len(data) $<$ 100000）：继续 Cov，μ = 计算统计量（data）；结束 3. 传感器阵列 = 各个传感器；对于传感器阵列中的数组，执行 data = mhs[数组]对于 $\mathbf{x}_i$ 中的每个 data，执行 k =分割掩膜[i]；SLF($\mathbf{x}_i$) = $(\mathbf{x}_{i-\mu_k})^T\mathbf{Cov}_k^{-1}$ t$\sqrt{}$t$^T\mathbf{Cov}_k^{-1}$ t 结束 结束 SLF($\mathbf{x}_i$) $\forall\ classes\ and\ i \in mhs$结束 算法 1：增强匹配滤波器

9.2.4 Training policy

我们以两种方式训练了 MethaneMapper，（i）首先预训练边界框和类别检测，然后冻结

预训练模型参数并仅训练掩码预测层；以及（ii）端到端训练整个管道，并在两种情况下都取得了相似的性能。

9.2.5 Qualitative Results

In this section we show few more qualitative examples of CH_4 plume mask prediction and few cases where MM failed to detect any CH_4 gas emission.Figure. 8 shows the CH_4 detections in different types of background terrain and different types of emission source.

Figure 9 shows some examples of missed CH_4 plume detections. We observed that going back to dataset samples and checking the timelines, these flightlines were recorded during the evening time. We believe that this might be because of evening time, the reflectance from the ground terrain is very weak and small. Hence we believe there is minimum absorption of reflected solar radiation by CH_4 gas present in the atmosphere and the plume goes undetected.

9.3. Ablations Studies

Attention Type: We also explored different attention mechanisms to encode and decode information. We replaced only the attention layers with deformable-attention [51] in the our architecture that resulted in a drop of 0.1 mAP in the baseline model.

9.3.1 Implementation details

The whole network is trained with AdamW [33] optimizer, batch size of 12, with initial learning rate for backbones set to 10^{-5} and for transformer the learning rate is set to 10^{-5} with a weight decay of 10^{-4}. The learning rate for mask prediction module is set to 10^{-4}. The learning rate is dropped at every 150 epochs, we train for 300 epochs. The baseline model is trained on 2 V100 GPUs.

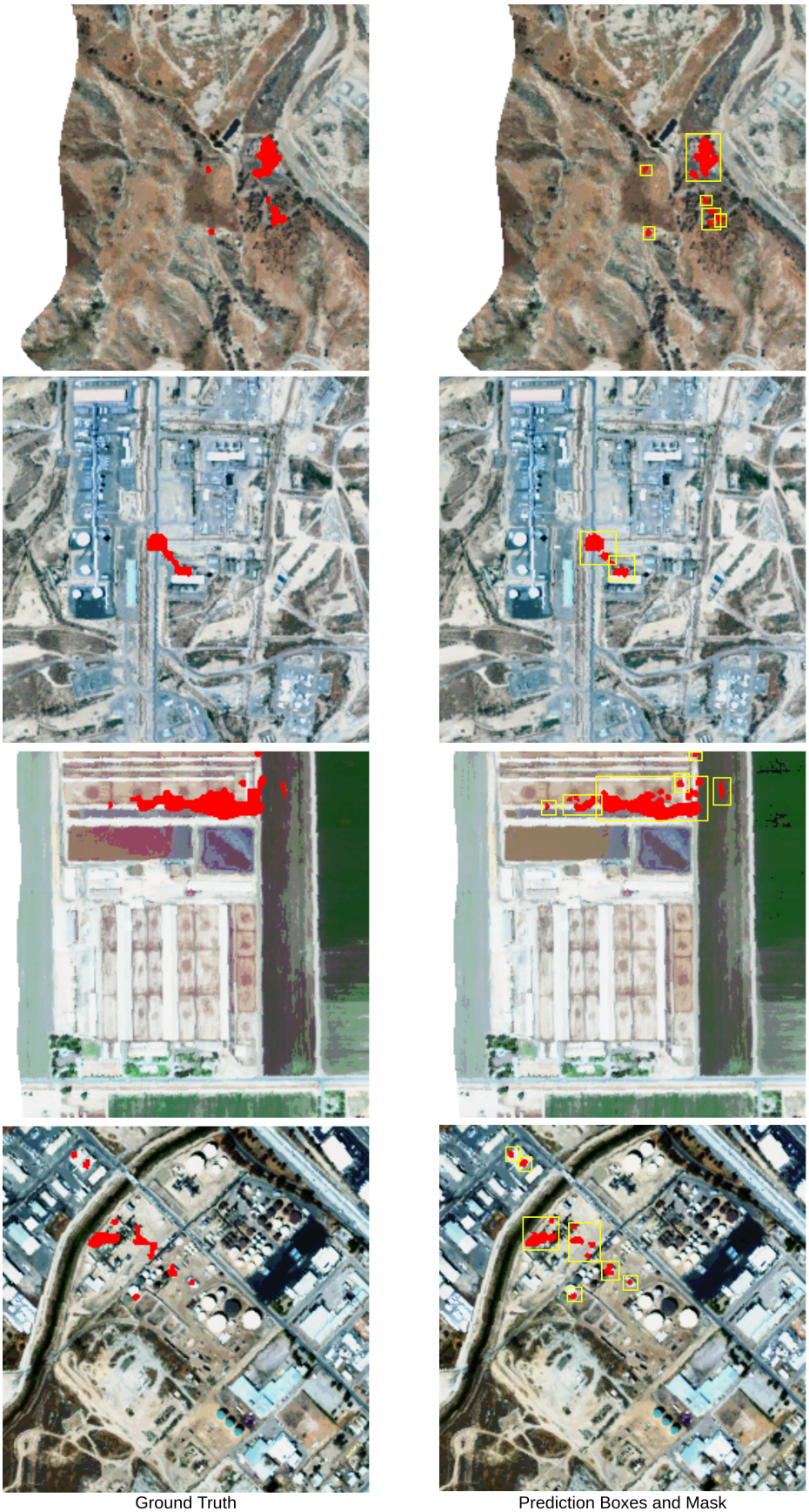

图 8. *MHS* 数据集上的样本真值与预测。我们展示了不同类型的地形以及 CH_4 在其上的预测。所有样本中的排放源类型也各不相同。

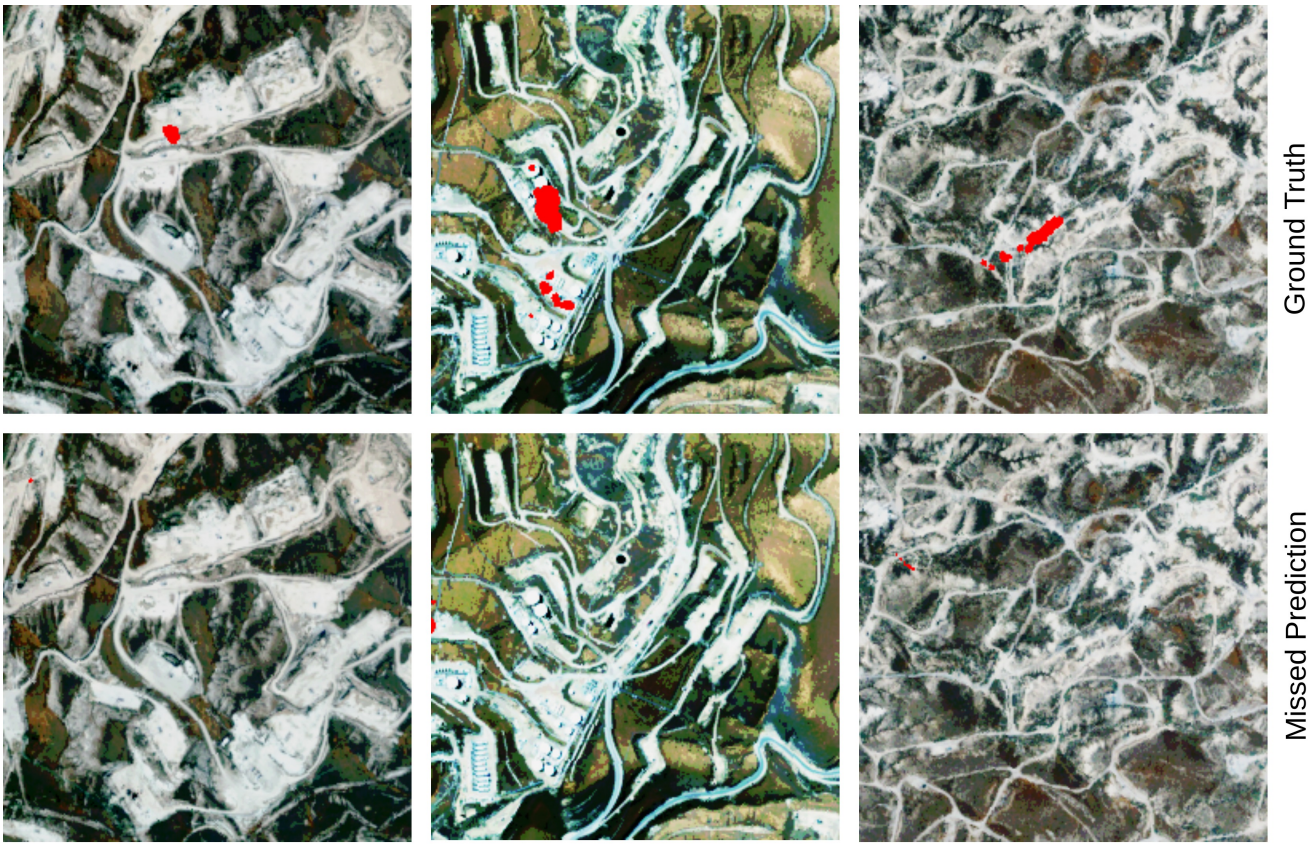

Figure 9. *Samples where MM fails to detect the CH_4 plume. We observed that these samples were recorded during the evening time and hence reflectance from the ground terrain is very weak. Therefore the absorption of reflected solar radiations by CH_4 is very low and hence the emissions goes undetected.*