A Data-Efficient Deep Transfer Learning Framework for Methane Super-Emitter Detection in Oil and Gas Fields Using Sentinel-2 Satellite

Shutao Zhao[1,2], Yuzhong Zhang[2,3]*, Shuang Zhao[2,3], Xinlu Wang[1,2]

1. Zhejiang University, Hangzhou, Zhejiang Province, 310058, China
2. Key Laboratory of Coastal Environment and Resources of Zhejiang Province, School of Engineering, Westlake University, Hangzhou, Zhejiang Province, 310024, China
3. Institute of Advanced Technology, Westlake Institute for Advanced Study, Hangzhou 310024, Zhejiang Province, China

Corresponding Author: Yuzhong Zhang, *Email: zhangyuzhong@westlake.edu.cn;

Text S1. Methane emission signal (ΔR) retrieval

Similar to (Ehret et al. 2022; Irakulis-Loitxate et al. 2022), we derived ΔR by comparing the ratio of band12 and band11 with a reference background without enhanced methane concentrations. The reference background is predicted by multivariate linear regression (MLR) models by pixel. For that, a sliding time window of T (60) dates was set and the patches in the time continuum were extracted (Ehret et al. 2022). To obtain optimal training set of MLR, we firstly introduced an image structural similarity index measure (SSIM) algorithm (Zhou et al. 2004) to discard the n (15) images that were most dissimilar to the date of interest t in the time series. Most of the discarded images contained opaque or circus clouds as shown in Fig. S1. SSIM estimated image distances considering the combination of structure, contrast, and luminance in band11. Band11 is ideal for comparison as it belongs to SWIR range like band12 but the methane absorption is much weaker to present anomalous absorption signal. Then, the proposed LRAD algorithm was deployed to detect and mask the potential artifacts in the SSIM-optimized data continuum. Within the data cube, patches of the past T-n-1 dates were employed to train MLR model and generate band11 and 12 references. If the coefficient of determination (R^2) of the MLR was lower than 0.5, the date of interest t was skipped and the S2L1C data was classified as cloudy observations. The calculation formular of ΔR is shown as follows:

$$\Delta R = \frac{\mathrm{band}_{12}^{t}/\mathrm{band}_{12}^{ref}}{\mathrm{band}_{11}^{t}/\mathrm{band}_{11}^{ref}}$$

Considering that band12 exhibits a significantly higher methane absorption capacity compared to band11, any pronounced methane emission event would lead to a noticeable reduction in the pixel values within the ΔR range of 0-1. Consequently, we applied a threshold range of [0, 1] to ΔR in order to mask anomalies and then a threshold of the 5th percentile value was applied to ΔR in order to remove background. In the end, we applied a colormap to map the unitless ΔR matrix into RGB imagery, so the input to the plume detection algorithm conforms to the structure of ResNet50 in order to use ImageNet-based pre-training parameters, and also can provide more hierarchical features to avoid potential accuracy degradation (Shorten and Khoshgoftaar 2019).

Text S2. Emission flux quantification and uncertainty estimation

Emission flux rates (Q, kg h^{-1}) are calculated for each detected plume-containing ΔR image. Firstly, we employed the radiative transfer model by (Varon et al. 2021) to convert unitless ΔR to methane column enhancements (mol m^{-2}). Secondly, we manually defined a plume mask based on the enhancement image. Background enhancement (mean enhancement outside the mask) is subtracted for pixels in the mask. Finally, the emission flux rate Q is computed using the integrated mass enhancement (IME) method (Frankenberg et al. 2016; Varon et al. 2018):

$$Q = \frac{IME \times U_{eff}}{L} \quad \text{(S1)}$$

where IME is computed as the sum of methane mass enhancements within the plume mask. U_{eff} (m s^{-1}) is the effective wind speed, which is computed based on the GEOS-FP 1 hour average 10-m wind speed U_{10} following the calibration equation developed in (Varon et al. 2021). L (m) is the plume length which is computed as the square root of the plume area.

To estimate the uncertainty for the emission flux rate, we consider three dominant error terms in Eqs. (S1). The random error of IME, mainly originated from retrieval noise, is estimated as the standard deviation of methane column mass enhancement outside the plume mask (Cusworth et al. 2020). The error of GEOS-FP U_{10} is assumed to be 50%, consistent with the ~1.5 m s^{-1} standard deviation given by (Varon et al. 2020). Following (Sánchez-García et al. 2022), an error of 0.01 is assumed for both the slope and intercept of the U_{eff} calibration function. We add the above errors in quadrature to derive the total uncertainty (1σ) of the emission flux.

Text S3. Labeling decision rule of ΔR imagery

We categorized the ΔR images into two classes, plume-containing and plume-free, following the procedure in Fig. 5. The determination is mainly based on visual inspection of ΔR images. We first look for the presence of methane plumes in ΔR images. If present, we then examine whether the potential methane plume signal is roughly aligned with wind direction and is free from surface and cloud interference. We use Goddard Earth Observing System-Fast Processing (GEOS-FP) 10 m wind reanalysis data as main information of wind direction (Varon et al. 2020). Since we find that the wind direction in the GEOS-FP 10-m wind data often does not align with the plume direction, the difference between plume and wind direction tolerated by this labeling process is less than 90°. Nonetheless, there are still a few cases where the plume morphology is distinct but the difference in wind direction is greater than 90°. To this end, the visual inspection of plume is supplemented by directions of nearby smoke plumes (if available) seen in RGB images. Subsequently, we use SWIR and RGB images to rule out potential interference by surface and cloud. It is noted that artifacts originate from low reflectivity surface features, so we focus on the least reflective pixels in the SWIR images. If the "plume" morphology in ΔR image presents to be the low-reflectivity region in SWIR images, then we discriminate it as a false signal.

Figures

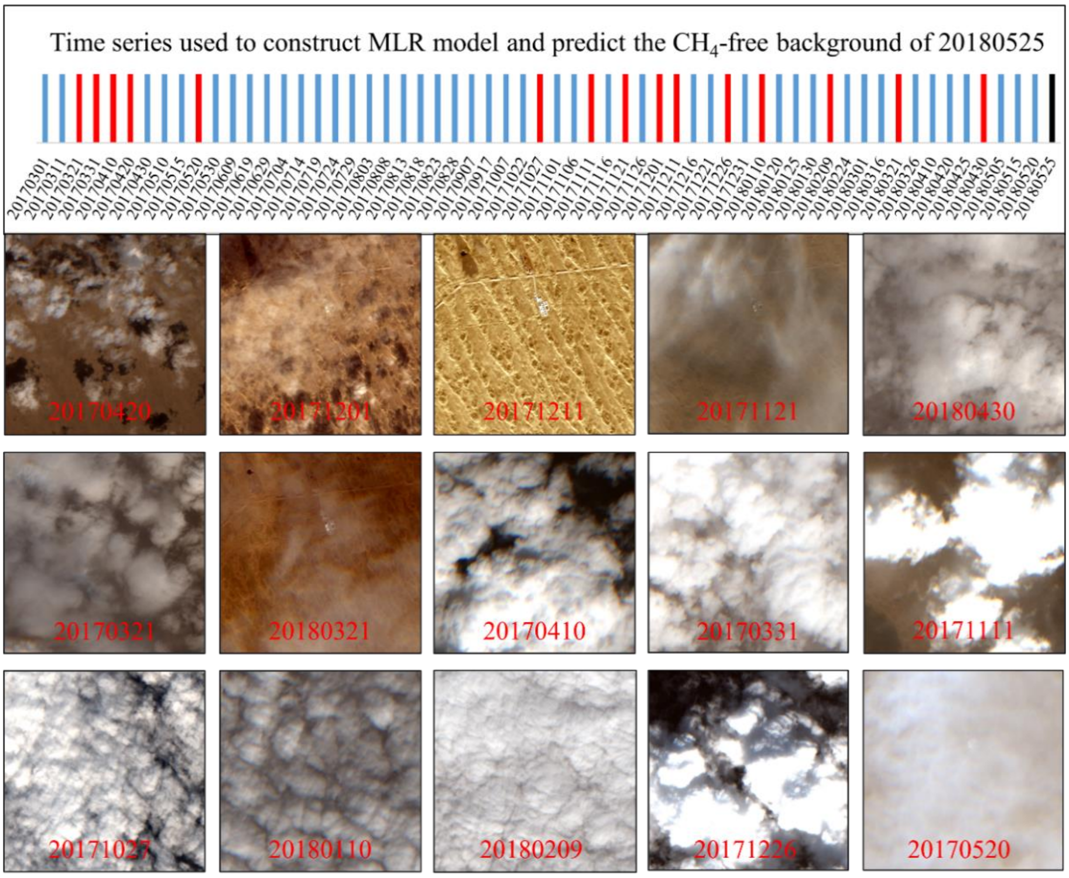

Fig. S1. RGB images of the S2L1C observations discarded by SSIM (take 20180525 as an example)

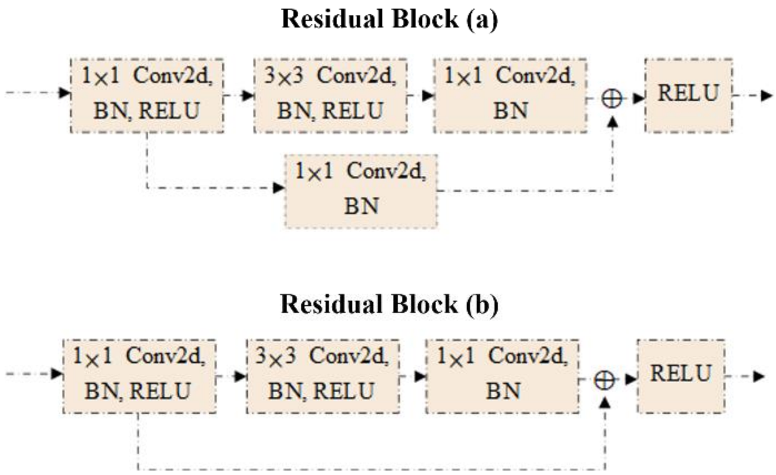

Fig. S2. Architecture of residual blocks in DSAN.

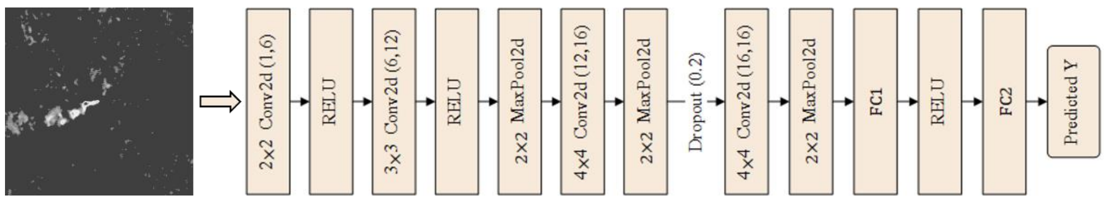

Fig. S3. Architecture of MethaNet proposed by (Jongaramrungruang et al. 2022).

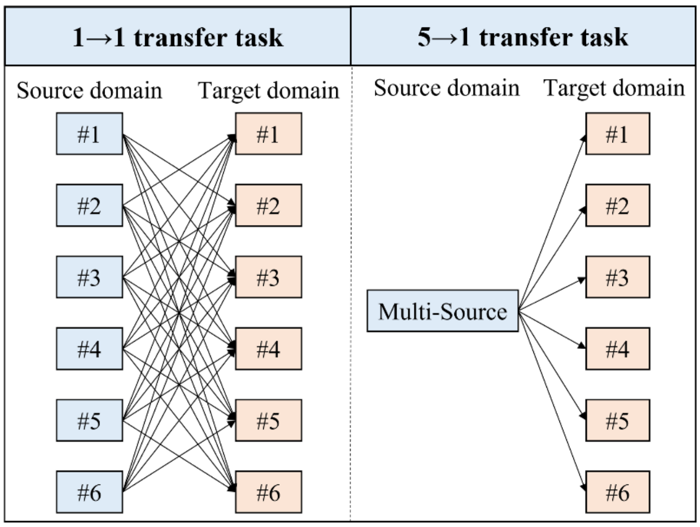

Fig. S4. Models were trained and assessed on two kinds of transfer tasks. "1→1": single source domain to single-target domain and "5→1": multi-source domain to single-target domain (multi-source domain is a fusion of all the datasets except for the target dataset).

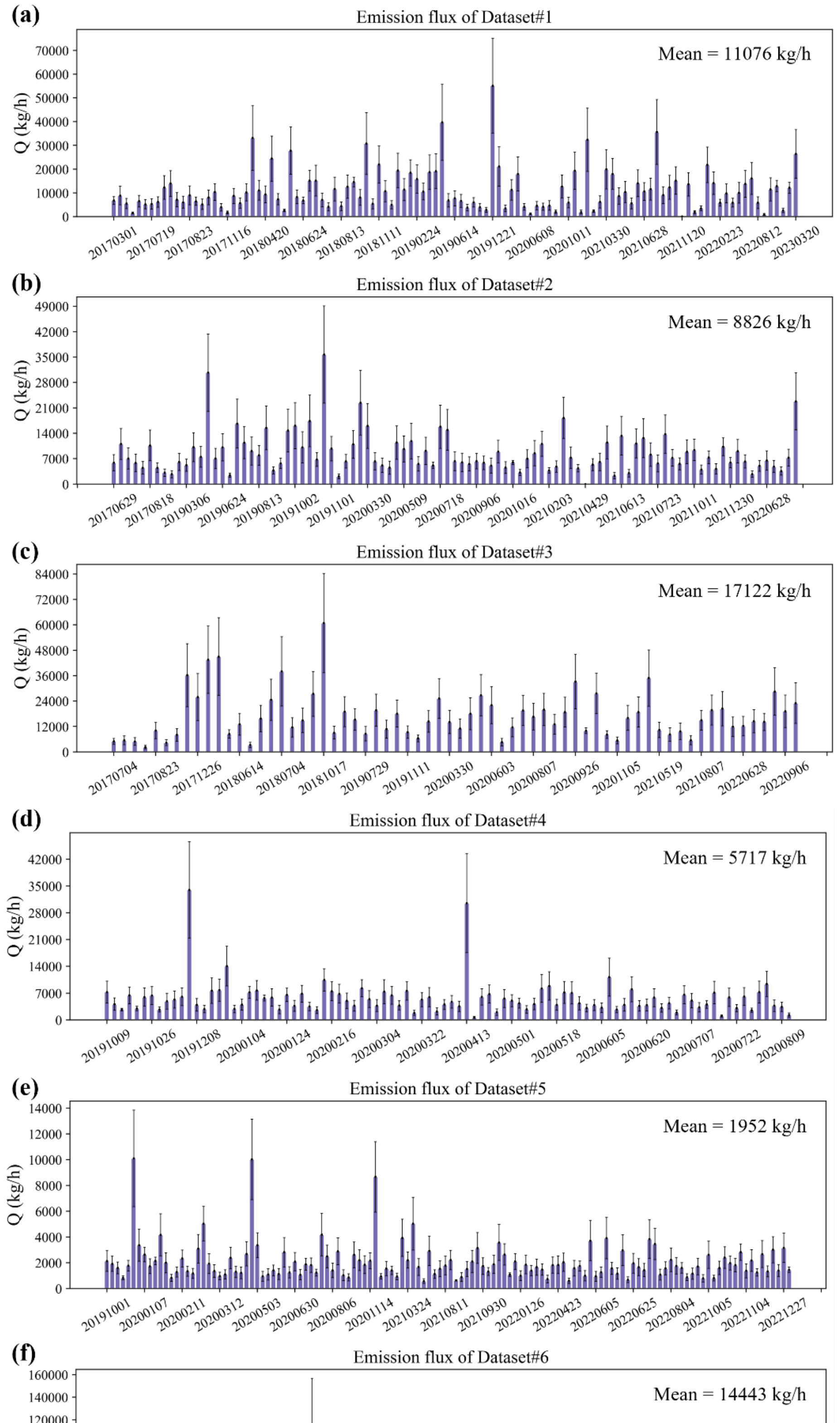

Fig. S5. Emission flux (kg/h) and uncertainty quantification of the methane plumes in Dataset#1-6

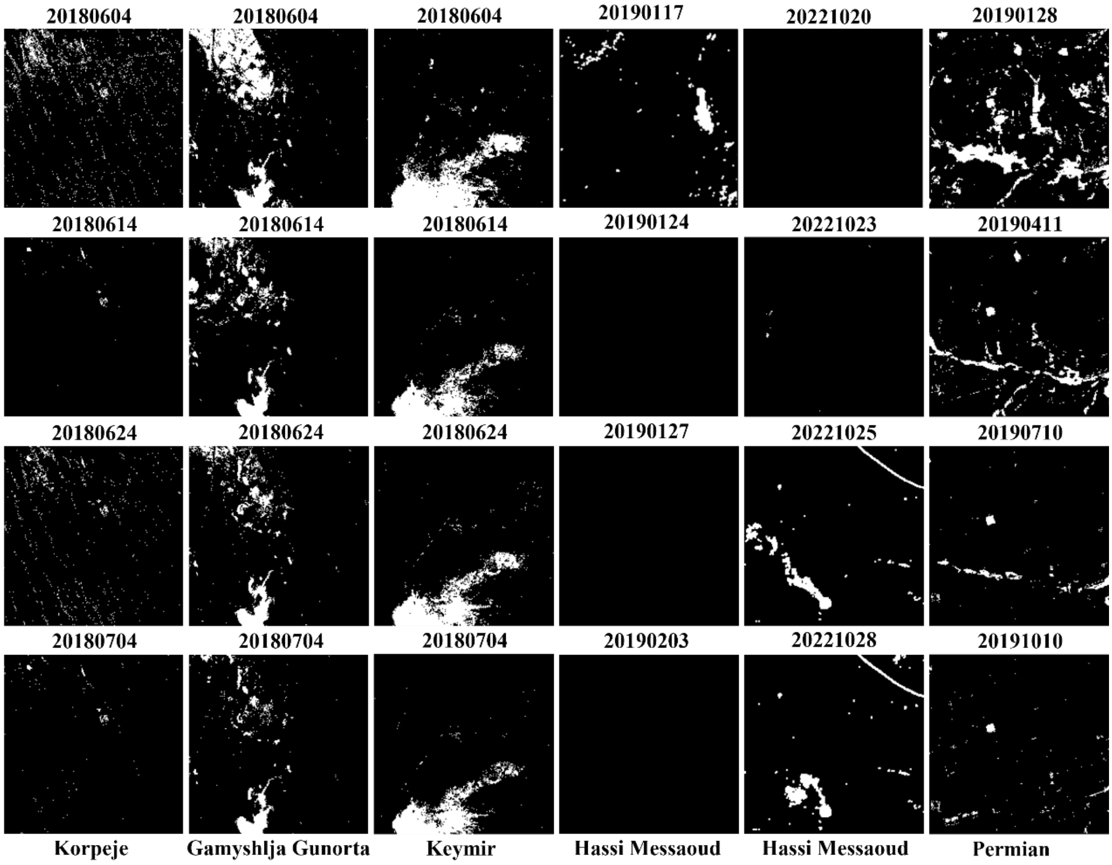

Fig. S6. Examples of the adaptive denoising masks over the six methane point sources generated by the LRAD algorithm. White color represents pixels that are filtered out as artifacts.

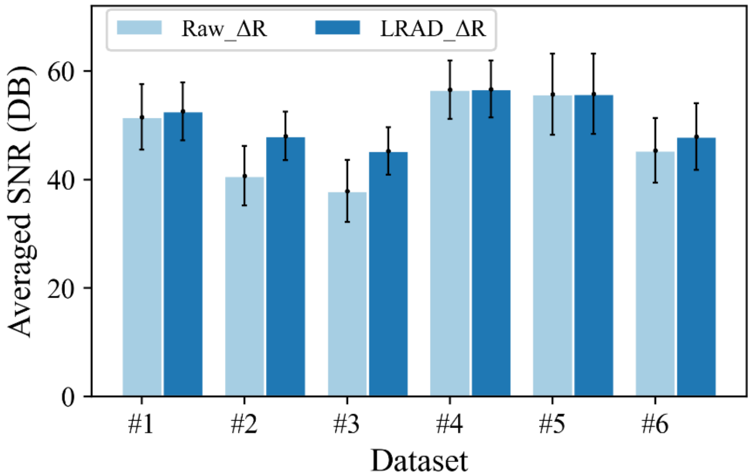

Fig. S7. Comparison of the averaged signal-to-noise ratios (SNRs) of the six ΔR datasets before and after deploying the LRAD algorithm.

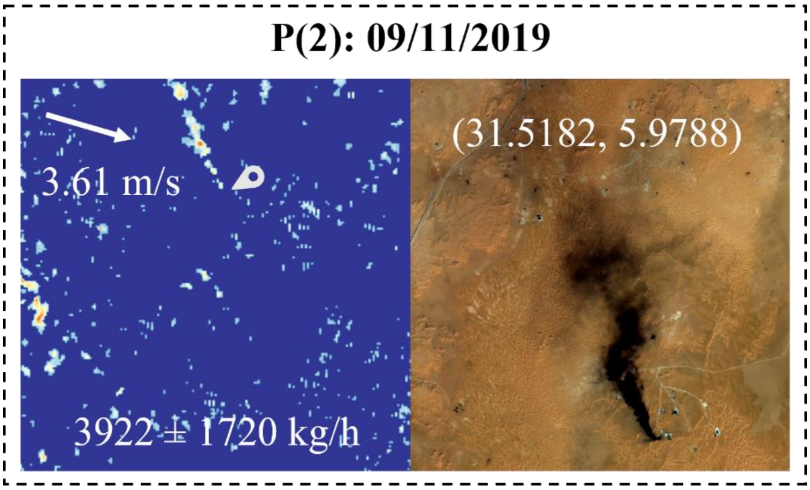

Fig. S8. False negative detection in source P(2), and the corresponding RGB images extracted from Sentinel-2 L1C product.

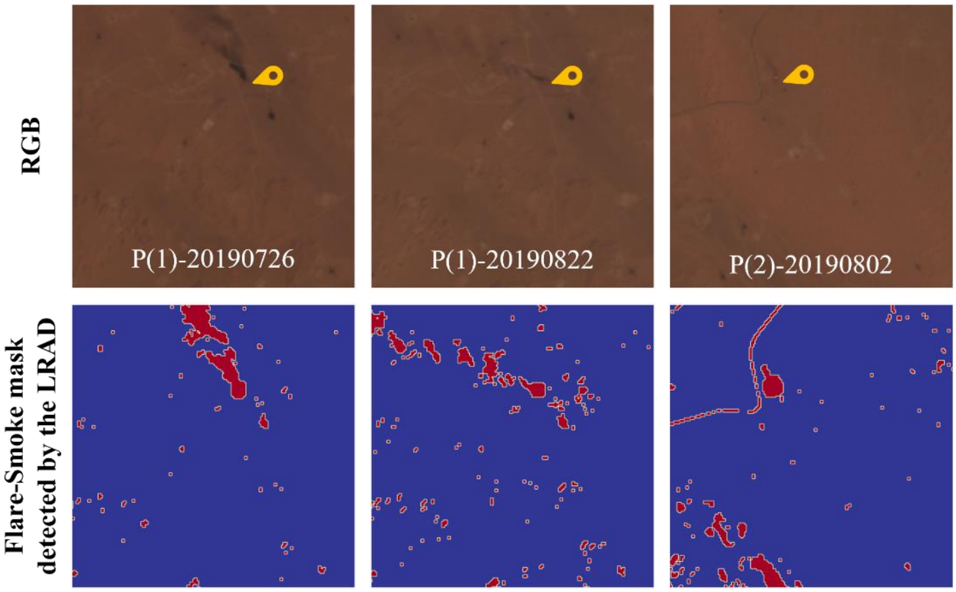

Fig. S9. Top row shows RGB images of flaring at P(1) and P(2), which are extracted from Sentinel-2 L1C product. Bottom row presents the flare and smoke (red pixels) masks detected by the LRAD algorithm. Yellow pins indicate the locations of flaring facilities.

Tables

Table S1 Performances of MethaNet, ResNet-50, and DSAN models on test sets of the '1→1' transfer tasks.

Task #	Class	MethaNet			ResNet-50			DSAN		
		Precision	Recall	Accuracy	Precision	Recall	Accuracy	Precision	Recall	Accuracy
1-2	contain	0.74	0.55	0.76	0.70	0.80	0.80	0.80	0.86	0.87
	free	0.77	0.89		0.87	0.80		0.92	0.87	
1-3	contain	0.85	0.59	0.87	0.64	0.89	0.84	0.76	0.97	0.91
	free	0.87	0.96		0.96	0.83		0.99	0.89	
1-4	contain	0.78	0.71	0.86	0.71	0.92	0.87	0.81	0.95	0.92
	free	0.89	0.92		0.97	0.85		0.98	0.91	
1-5	contain	0.76	0.20	0.65	0.78	0.78	0.81	0.86	0.87	0.89
	free	0.64	0.96		0.84	0.84		0.90	0.90	
1-6	contain	0.36	0.50	0.90	0.19	0.71	0.77	0.36	0.67	0.89
	free	0.96	0.93		0.97	0.77		0.97	0.91	
2-1	contain	0.74	0.84	0.80	0.80	0.82	0.83	0.93	0.90	0.93
	free	0.86	0.76		0.85	0.83		0.92	0.95	
2-3	contain	0.78	0.59	0.85	0.57	0.92	0.80	0.82	0.91	0.92
	free	0.87	0.94		0.97	0.75		0.97	0.93	
2-4	contain	0.87	0.73	0.89	0.77	0.92	0.90	0.84	0.90	0.92
	free	0.90	0.96		0.97	0.89		0.96	0.93	
2-5	contain	0.86	0.15	0.65	0.80	0.80	0.83	0.88	0.81	0.87
	free	0.63	0.98		0.85	0.85		0.87	0.92	
2-6	contain	0.38	0.28	0.91	0.18	0.71	0.75	0.45	0.72	0.91
	free	0.94	0.96		0.97	0.76		0.98	0.93	
3-1	contain	0.87	0.77	0.84	0.82	0.91	0.87	0.97	0.85	0.92
	free	0.83	0.90		0.92	0.83		0.89	0.98	
3-2	contain	0.84	0.57	0.80	0.68	0.79	0.79	0.81	0.83	0.87
	free	0.79	0.94		0.87	0.79		0.90	0.89	
3-4	contain	0.94	0.64	0.88	0.63	0.92	0.83	0.88	0.92	0.94
	free	0.87	0.98		0.96	0.79		0.97	0.95	
3-5	contain	0.77	0.30	0.68	0.77	0.76	0.81	0.91	0.81	0.88
	free	0.66	0.94		0.83	0.84		0.87	0.94	
3-6	contain	0.45	0.56	0.92	0.24	0.76	0.81	0.55	0.67	0.93
	free	0.96	0.95		0.98	0.81		0.97	0.95	
4-1	contain	0.85	0.78	0.84	0.87	0.94	0.91	0.94	0.84	0.90
	free	0.83	0.89		0.94	0.89		0.88	0.95	
4-2	contain	0.55	0.84	0.69	0.68	0.76	0.78	0.75	0.83	0.83
	free	0.87	0.60		0.85	0.79		0.90	0.84	
4-3	contain	0.66	0.73	0.83	0.68	0.85	0.86	0.85	0.76	0.90
	free	0.90	0.87		0.94	0.86		0.92	0.95	
4-5	contain	0.64	0.22	0.62	0.83	0.72	0.82	0.91	0.78	0.88

	free	0.62	0.91		0.82	0.89		0.86	0.94	
4-6	contain	0.33	0.44	0.89	0.21	0.71	0.79	0.53	0.56	0.93
	free	0.95	0.93		0.97	0.80		0.96	0.96	
5-1	contain	0.64	0.81	0.71	0.85	0.92	0.89	0.91	0.94	0.93
	free	0.80	0.63		0.93	0.86		0.95	0.92	
5-2	contain	0.49	0.79	0.63	0.62	0.84	0.76	0.75	0.89	0.85
	free	0.81	0.53		0.89	0.71		0.93	0.83	
5-3	contain	0.53	0.70	0.76	0.62	0.94	0.83	0.79	0.92	0.92
	free	0.88	0.78		0.97	0.80		0.97	0.91	
5-4	contain	0.48	0.84	0.70	0.73	0.90	0.88	0.80	0.94	0.91
	free	0.91	0.64		0.96	0.87		0.97	0.91	
5-6	contain	0.17	0.61	0.75	0.29	0.82	0.85	0.33	0.72	0.87
	free	0.96	0.76		0.98	0.85		0.98	0.88	
6-1	contain	0.88	0.53	0.76	0.88	0.56	0.77	0.99	0.67	0.85
	free	0.71	0.94		0.73	0.94		0.79	0.99	
6-2	contain	0.52	0.53	0.64	0.73	0.40	0.73	0.82	0.61	0.81
	free	0.72	0.71		0.73	0.91		0.80	0.92	
6-3	contain	0.84	0.24	0.79	0.75	0.37	0.80	0.90	0.67	0.89
	free	0.79	0.98		0.81	0.96		0.89	0.97	
6-4	contain	0.94	0.34	0.81	0.90	0.65	0.88	0.96	0.84	0.94
	free	0.79	0.99		0.88	0.97		0.94	0.99	
6-5	contain	0.96	0.18	0.67	0.90	0.48	0.76	0.97	0.55	0.81
	free	0.64	0.99		0.72	0.96		0.75	0.99	

Table S2 Performances of the MethaNet and ResNet-50 on test sets (30%) on validation sets of the non-transfer tasks.

Dataset#	Class	MethaNet			ResNet-50		
		Precision	Recall	Accuracy	Precision	Recall	Accuracy
1	contain	0.80	0.95	0.88	0.97	1.00	0.99
	free	0.96	0.82		1.00	0.98	
2	contain	0.76	0.72	0.83	0.85	1.00	0.94
	free	0.86	0.88		1.00	0.90	
3	contain	1.00	0.67	0.90	0.83	0.95	0.93
	free	0.88	1.00		0.98	0.93	
4	contain	0.95	0.87	0.94	0.90	1.00	0.97
	free	0.93	0.98		1.00	0.96	
5	contain	0.81	0.75	0.81	0.95	0.92	0.95
	free	0.81	0.85		0.95	0.96	
6	contain	1.00	0.80	0.98	0.83	1.00	0.99
	free	0.98	1.00		1.00	0.99	

Table S3 Performances of MethaNet, ResNet-50, and DSAN models on test sets of the '5→1' transfer tasks.

Task#	Class	MethaNet	ResNet-50	DSAN

		Precision	Recall	Accuracy	Precision	Recall	Accuracy	Precision	Recall	Accuracy
1	contain	0.79	0.87	0.84	0.83	0.93	0.88	0.93	0.92	0.93
	free	0.89	0.81		0.93	0.84		0.93	0.94	
2	contain	0.79	0.68	0.82	0.72	0.83	0.82	0.82	0.85	0.88
	free	0.83	0.90		0.89	0.81		0.91	0.89	
3	contain	0.84	0.71	0.89	0.68	0.83	0.85	0.89	0.94	0.95
	free	0.90	0.95		0.94	0.86		0.98	0.96	
4	contain	0.89	0.85	0.93	0.80	0.96	0.92	0.89	0.92	0.94
	free	0.94	0.96		0.98	0.91		0.97	0.95	
5	contain	0.74	0.15	0.62	0.86	0.74	0.84	0.95	0.81	0.90
	free	0.61	0.96		0.83	0.92		0.87	0.97	
6	contain	0.35	0.50	0.89	0.22	0.71	0.80	0.47	0.78	0.92
	free	0.96	0.92		0.97	0.81		0.98	0.93	

References

Cusworth, D.H., Duren, R.M., Thorpe, A.K., Tseng, E., Thompson, D., Guha, A., Newman, S., Foster, K.T., & Miller, C.E. (2020). Using remote sensing to detect, validate, and quantify methane emissions from California solid waste operations. *Environmental Research Letters, 15*, 054012

Ehret, T., De Truchis, A., Mazzolini, M., Morel, J.-M., d'Aspremont, A., Lauvaux, T., Duren, R., Cusworth, D., & Facciolo, G. (2022). Global Tracking and Quantification of Oil and Gas Methane Emissions from Recurrent Sentinel-2 Imagery. *Environmental Science & Technology, 56*, 10517-10529

Frankenberg, C., Thorpe, A.K., Thompson, D.R., Hulley, G., Kort, E.A., Vance, N., Borchardt, J., Krings, T., Gerilowski, K., Sweeney, C., Conley, S., Bue, B.D., Aubrey, A.D., Hook, S., & Green, R.O. (2016). Airborne methane remote measurements reveal heavy-tail flux distribution in Four Corners region. *Proceedings of the National Academy of Sciences, 113*, 9734-9739

Irakulis-Loitxate, I., Guanter, L., Maasakkers, J.D., Zavala-Araiza, D., & Aben, I. (2022). Satellites Detect Abatable Super-Emissions in One of the World's Largest Methane Hotspot Regions. *Environmental Science & Technology, 56*, 2143-2152

Sánchez-García, E., Gorroño, J., Irakulis-Loitxate, I., Varon, D.J., & Guanter, L. (2022). Mapping methane plumes at very high spatial resolution with the WorldView-3 satellite. *Atmos. Meas. Tech., 15*, 1657-1674

Shorten, C., & Khoshgoftaar, T.M. (2019). A survey on Image Data Augmentation for Deep Learning. Journal of Big Data, 6, 60

Varon, D.J., Jacob, D.J., Jervis, D., & McKeever, J. (2020). Quantifying Time-Averaged Methane Emissions from Individual Coal Mine Vents with GHGSat-D Satellite Observations. *Environmental Science & Technology, 54*, 10246-10253

Varon, D.J., Jacob, D.J., McKeever, J., Jervis, D., Durak, B.O.A., Xia, Y., & Huang, Y. (2018). Quantifying methane point sources from fine-scale satellite observations of atmospheric methane plumes. *Atmos. Meas. Tech., 11*, 5673-5686

Varon, D.J., Jervis, D., McKeever, J., Spence, I., Gains, D., & Jacob, D.J. (2021). High-frequency monitoring of anomalous methane point sources with multispectral Sentinel-2 satellite observations. *Atmos. Meas. Tech., 14*, 2771-2785

Zhou, W., Bovik, A.C., Sheikh, H.R., & Simoncelli, E.P. (2004). Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Image Processing, 13*, 600-612