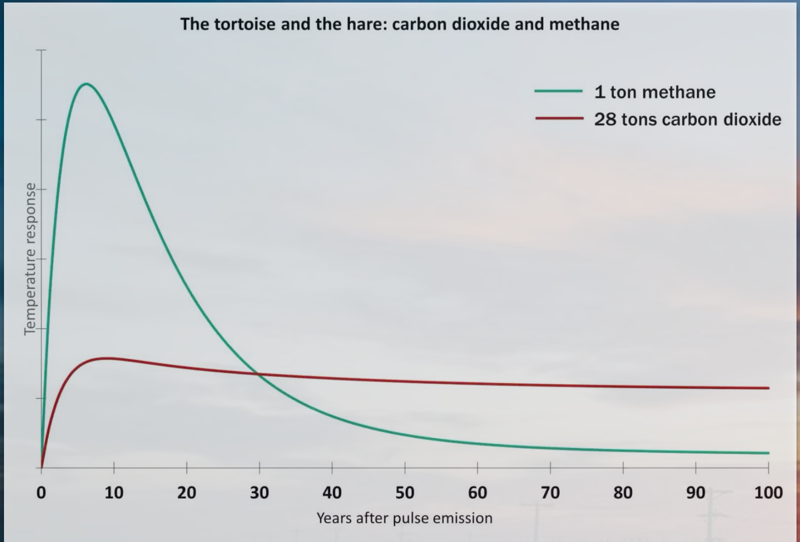
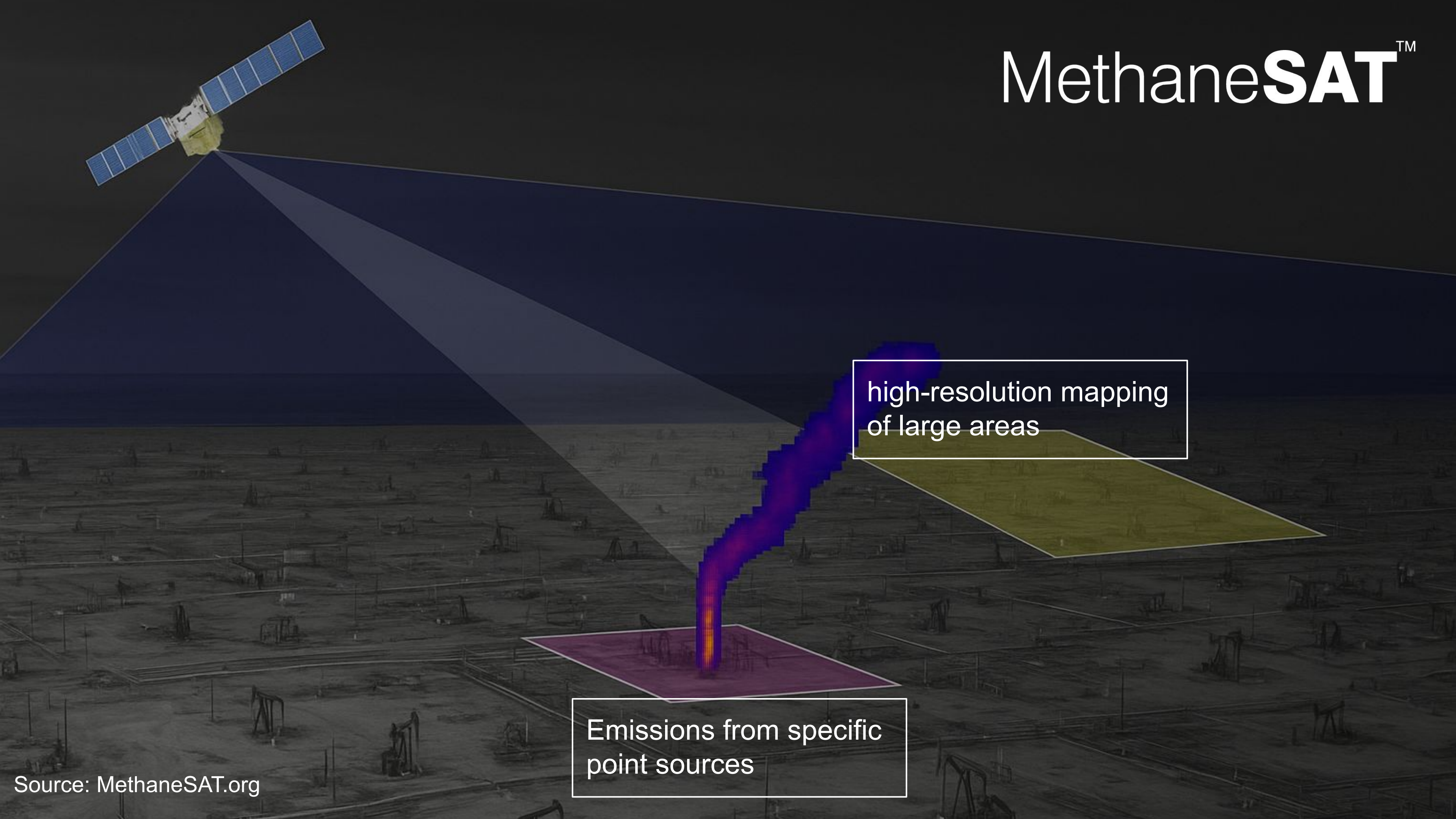
Spectral Channel Attention Network: A Method for Hyperspectral Semantic Segmentation of Cloud and Shadows

M. Pérez-Carrasco[1], M. Nasr[2 3], S. Roche[2 3], C. C. Miller[2 3], Z. Zhang[2 3], C. F. Park[4], E. Walker[3], C. Garraffo[1], D. Finkbeiner[4], R. Gautam[2], S. Wofsy[3]

[1] Center for Astrophysics | Harvard & Smithsonian
[2] Environmental Defense Fund
[3] Department of Earth And Planetary Sciences, Harvard University
[3] Department of Physics, Harvard University



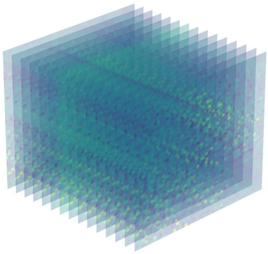
Source: Environmental Defense Fund
https://vitalsigns.edf.org/story/methane-its-not-your-grandparents-climate-pollution

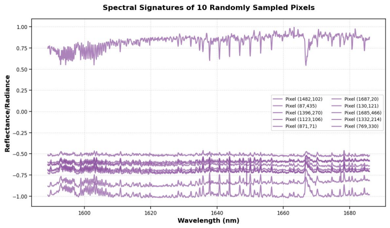


Source: MethaneSAT.org

Level 1B (L1B): Calibrated, corrected and georeferenced hyperspectral data

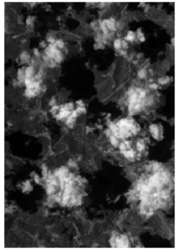
Conway et al. 2024

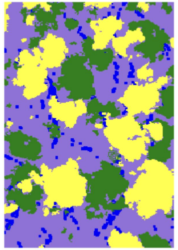




Level 2 (L2): Vertical column densities of CH_4 and CO_2, surface pressure, cloud products

Chan Miller et al. 2024

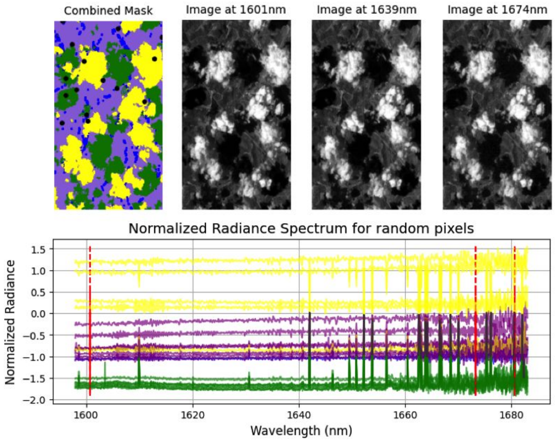




Dataset

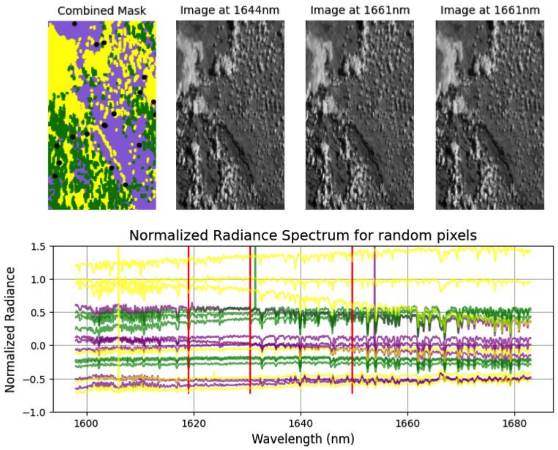
MethaneAIR: Airborne companion

- 508 hyperspectral samples
- Spectral range: 1592-1678 nm, 1024 spectral bins
- ~300×178 spatial soundings (along-track × across-track)
- 4 classes: cloud, shadows, dark surface and background



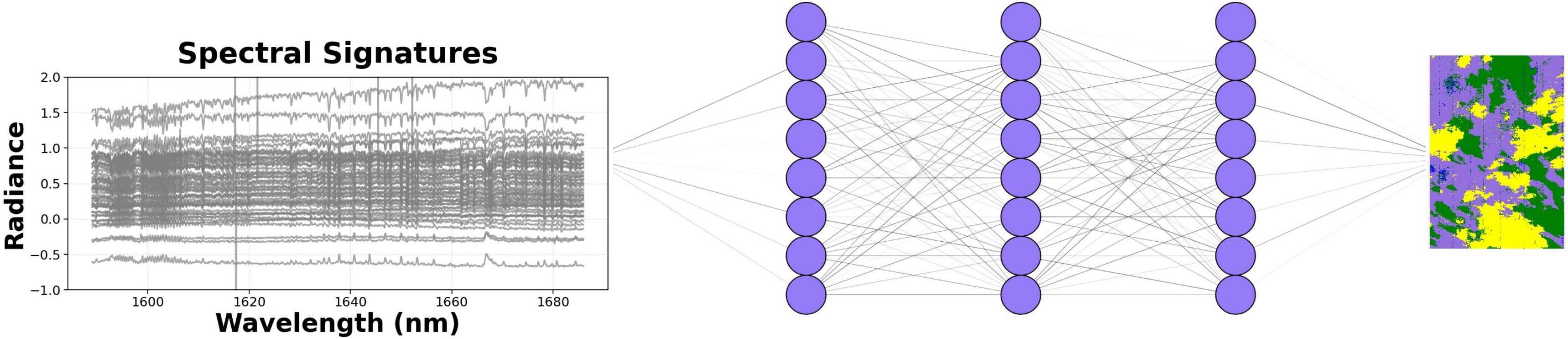
MethaneSAT: Satellite mission

- 262 hyperspectral samples
- Spectral range: 1589-1686 nm, 1080 spectral bins
- ~2200×500 spatial soundings (along-track × across-track)
- 3 classes: cloud, shadows and background



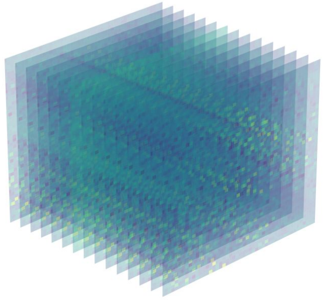
Methods

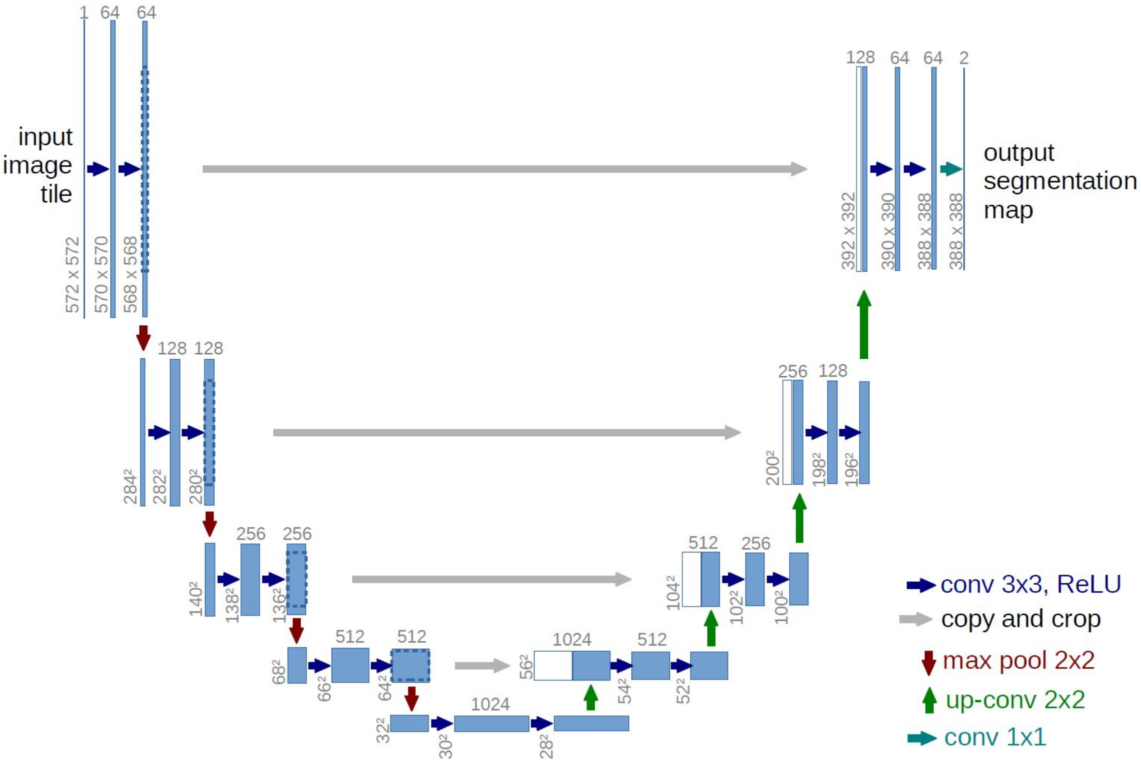
Semantic Segmentation: ILR and MLP

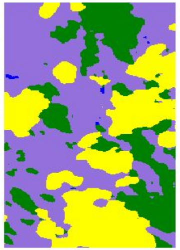


Park et al. 2023

Semantic Segmentation: U-Net

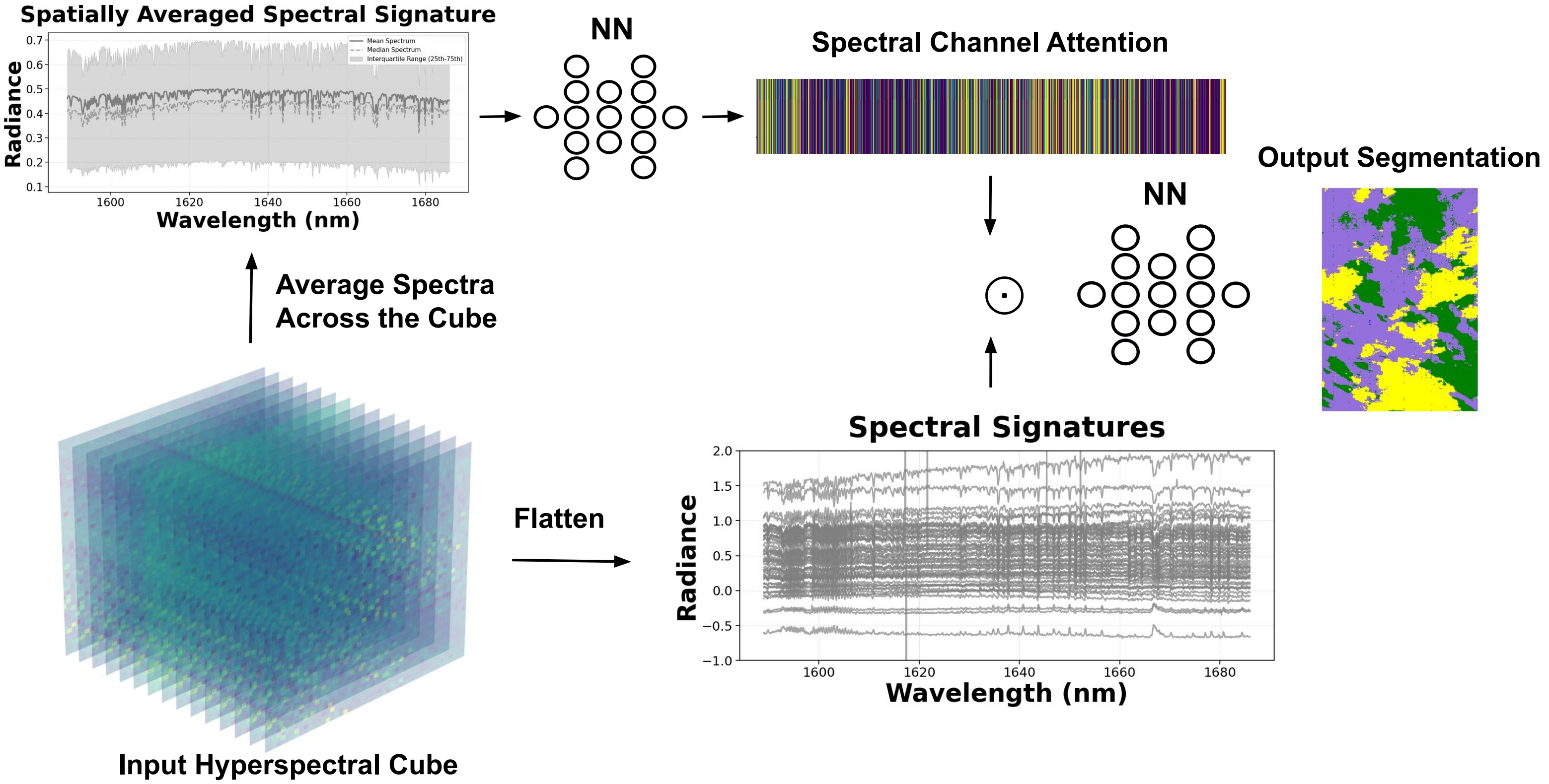




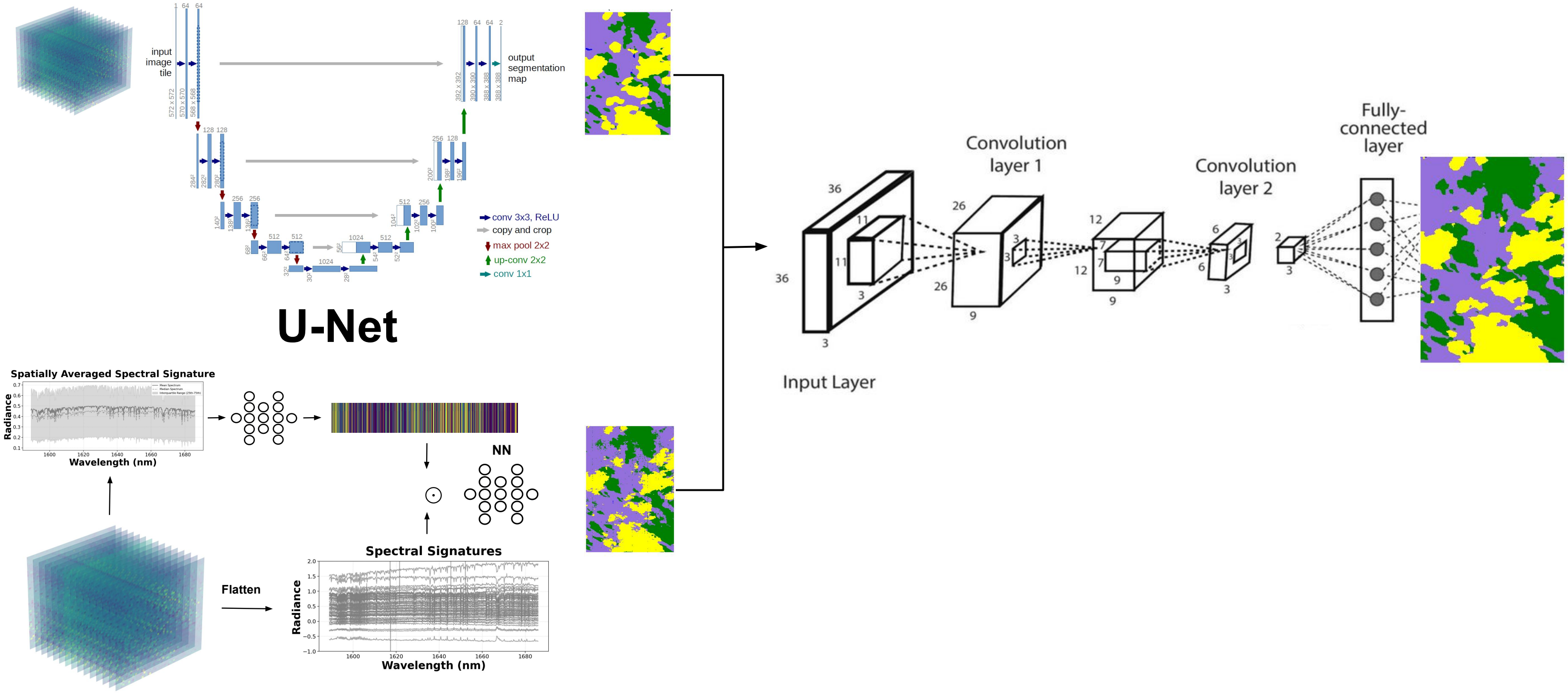


Ronnenberg et al. 2015

Spectral Channel Attention Network (SCAN)



Combined Approaches (ensemble)



SCAN

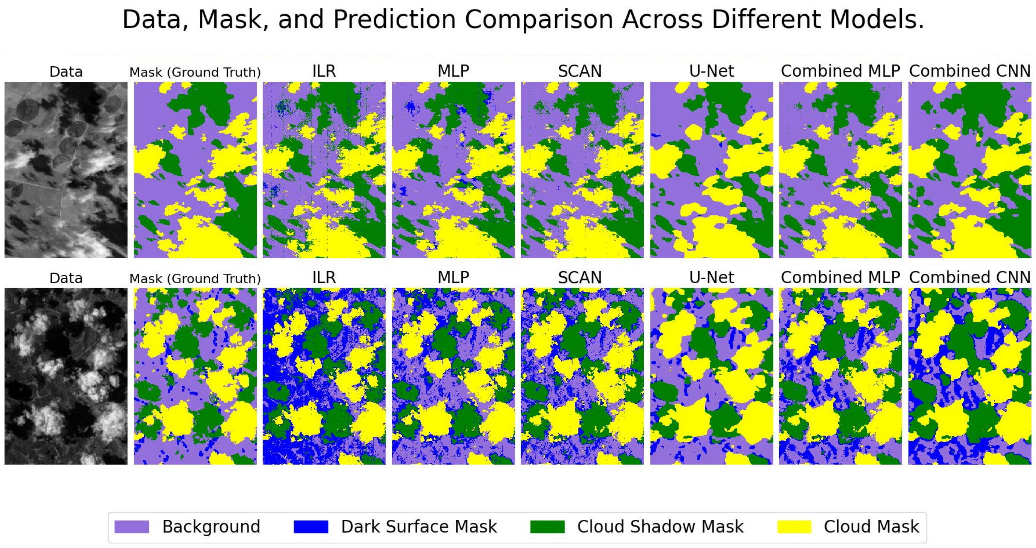
Results

Results

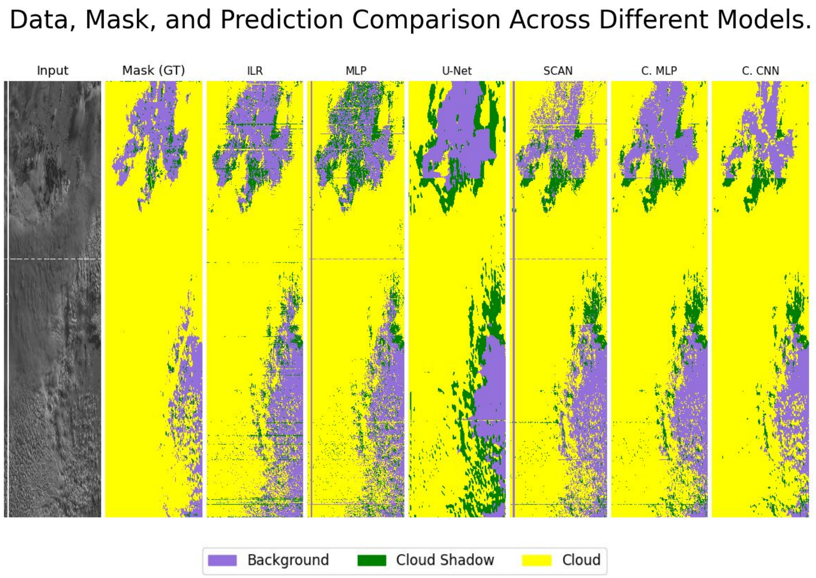
Table 1: Performance comparison across different models for both datasets.

Model	MethaneAIR Acc (%)	MethaneAIR F1 (%)	MethaneSAT Acc (%)	MethaneSAT F1 (%)
Individual Methods				
ILR	73.81±4.05	62.07±0.86	71.82±4.02	64.35±3.56
MLP	82.49±2.24	71.29±1.02	74.03±3.72	67.11±2.06
U-Net	**88.26±0.45**	**76.24±1.90**	78.73±3.23	68.56±0.36
SCAN	86.51±2.90	74.96±0.96	**80.33±3.43**	**71.53±0.75**
Ensemble Methods				
Combined MLP	88.92±1.80	76.99±6.78	81.32±1.28	78.10±1.72
Combined CNN	**89.42±1.20**	**78.50±3.08**	**81.96±1.45**	**78.80±1.28**

MAIR


Data, Mask, and Prediction Comparison Across Different Models.
Data
Mask (Ground Truth)
ILR
MLP
SCAN
U-Net
Combined MLP
Combined CNN
Background
Dark Surface Mask
Cloud Shadow Mask
Cloud Mask

MSAT


Data, Mask, and Prediction Comparison Across Different Models.
Input
Mask (GT)
ILR
MLP
U-Net
SCAN
C. MLP
C. CNN
Background
Cloud Shadow
Cloud

Conclusions

- Deep learning architectures achieve significantly higher performance, with U-Net and SCAN excelling in different aspects—U-Net in spatial coherence and SCAN in boundary precision.
- SCAN is better for shadows and dark surfaces. U-Net is better for clouds.
- Our ensemble approaches combine the complementary strengths of U-Net and SCAN, with the Combined CNN architecture achieving superior performance for both datasets.

Acknowledgements

Funding for MethaneSAT and MethaneAIR activities was provided in part by Anonymous, Arnold Ventures, The Audacious Project, Ballmer Group, Bezos Earth Fund, The Children's Investment Fund Foundation, Heising-Simons Family Fund, King Philanthropies, Robertson Foundation, Skyline Foundation and Valhalla Foundation. For a more complete list of funders, please visit www.methanesat.org. We thank the AstroAI and EarthAI institutes at the Center for Astrophyiscs | Harvard & Smithsonian for useful discussions and guidance. CG was supported by AstroAI at the Center for Astrophysics | Harvard and Smithsonian.

References

- G. Myhre et al. Anthropogenic and natural radiative forcing, pages 659–740. Cambridge University Press, Cambridge, UK, 2013.
- Maryam Etminan et al. Radiative forcing of carbon dioxide, methane, and nitrous oxide: A significant revision of the methane radiative forcing. Geophysical Research Letters, 43:12,614 – 12,623, 2016.
- R. R Rohrschneider et al. The methanesat mission. Small Satellite Conference.
- E. K. Conway et al. Level0 to level1b processor for methaneair. Atmospheric Measurement Techniques, 17(4):1347–1362, 2024.
- C. Chan Miller et al. Methane retrieval from methaneair using the co2 proxy approach: a demonstration for the upcoming methanesat mission. Atmospheric Measurement Techniques, 17(18):5429–5454, 2024.
- Apisada Chulakadabba et al. Methane point source quantification using methaneair: A new airborne imaging spectrometer. EGUsphere, 2023:1–22, 2023.
- L. Guanter et al. Remote sensing of methane point sources with the methaneair airborne spectrometer. EGUsphere, 2025:1–22, 2025.
- J. D. Warren et al. Sectoral contributions of high-emitting methane point sources from major u.s. onshore oil and gas producing basins using airborne measurements from methaneair. EGUsphere, 2024:1–22, 2024.
- Edisanter Lo et al. Target detection in hyperspectral Imaging using logistic regression. In Miguel Velez-Reyes and David W. Messinger, editors, Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XXII, volume 9840, page 98400W. International Society for Optics and Photonics, SPIE, 2016.
- Core Francisco Park et al. Hyperspectral shadow removal with iterative logistic regression and latent parametric linear combination of gaussians, 2023.
- Rajneesh Kumar Gautam et al. Hyperspectral image prediction using logistic regression model. In Arti Noor, Kriti Saroha, Emil Pricop, Abhijit Sen, and Gaurav Trivedi, editors, Proceedings of Emerging Trends and Technologies on Intelligent Systems, pages 283–293, Singapore, 2023. Springer Nature Singapore.
- P. H. SWAIN et al. Conjugate-gradient neural networks in classification of multisource and very-high-dimensional remote sensing data. International Journal of Remote Sensing, 14(15):2883–2903, 1993.
- H. Yang et al. A back-propagation neural network for mineralogical mapping from aviris data. International Journal of Remote Sensing, 20(1):97–110, 1999.
- Bin Tian et al. A study of cloud classification with neural networks using spectral and textural features. IEEE Transactions on Neural Networks, 10(1):138–151, 1999.
- Alireza Taravat et al. Multi-layer perceptron neural networks model for meteosat second generation seviri daytime cloud masking. Remote Sensing, 7(2):1529–1539, 2015.
- Luis Gómez-Chova et al. Cloud detection machine learning algorithms for proba-v. In 2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), pages 2251–2254, 2017.
- Zahid Hassan Tushar et al. Cloudunet: Adapting unet for retrieving cloud properties. In IGARSS 2024 - 2024 IEEE International Geoscience and Remote Sensing Symposium, pages 7163–7167, 2024.
- Olaf Ronneberger et al. U-net: Convolutional networks for biomedical image segmentation. In Nassir Navab, Joachim Hornegger, William M. Wells, and Alejandro F. Frangi, editors, Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015, pages 234–241, Cham, 2015. Springer International Publishing.
- Libin Jiao et al. Refined unet: Unet-based refinement network for cloud and shadow precise segmentation. Remote Sensing, 12:2001, 06 2020.
- Marc Wieland et al. Multi-sensor cloud and cloud shadow segmentation with a convolutional neural network. Remote Sensing of Environment, 230:111203, 2019.
- Shoukuan Miao et al. Cloud/shadow segmentation based on multi-level feature enhanced network for remote sensing imagery. International Journal of Remote Sensing, 43(15-16):5940–5960, 2022.
- Nicholas Wright et al. Clouds2mask: a novel deep learning approach for improved cloud and cloud shadow masking in sentinel-2 imagery. Remote Sensing of Environment, 306:114122, 2024.
- Xian Li et al. Gcdb-unet: A novel robust cloud detection approach for remote sensing images. Knowledge-Based Systems, 238:107890, 2022.
- Yuhao Tan et al. Cloud and cloud shadow detection of gf-1 images based on the swin-unet method. Atmosphere, 14(11), 2023.
- Dzmitry Bahdanau et al. Neural machine translation by jointly learning to align and translate. arXiv preprint arXiv:1409.0473, 2014.
- Paul Werbos et al. Applications of advances in nonlinear sensitivity analysis, volume 38, pages 762–770. 01 1970.
- Yann LeCun et al. Deep learning. Nature, 521(7553):436–444, 2015.
- Jie Hu et al. Squeeze-and-excitation networks. In 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, pages 7132–7141, 2018.
- Qilong Wang et al. Eca-net: Efficient channel attention for deep convolutional neural networks. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pages 11534–11542, 2020.
- Sanghyun Woo et al. Cbam: Convolutional block attention module. In Proceedings of the European conference on computer vision (ECCV), pages 3–19, 2018.
- Anita Thakur et al. Hyper spectral image classification using multilayer perceptron neural network & functional link ann. In 2017 7th International Conference on Cloud Computing, Data Science & Engineering - Confluence, pages 639–642, 2017.
- Alexey Dosovitskiy et al. An image is worth 16x16 words: Transformers for image recognition at scale. ArXiv, abs/2010.11929, 2020.
- Enze Xie et al. Segformer: Simple and efficient design for semantic segmentation with transformers. Advances in neural information processing systems, 34:12077–12090, 2021.
- Danfeng Hong et al. Spectralformer: Rethinking hyperspectral image classification with transformers. IEEE Transactions on Geoscience and Remote Sensing, 60:1–15, 2021.

Appendix

Ablation Study

- Benchmarks in **MethaneSAT** data: Vision Transformer (ViT-SegFormer) and SpectralFormer
- Both use standard self-attention using different tokenization approaches:
  * Standar ViT of *16x16xspectral_dim* non-overlapped patches projected to dim 768
  * Spectral patches (30 bands per patch) + class token for pixel classification

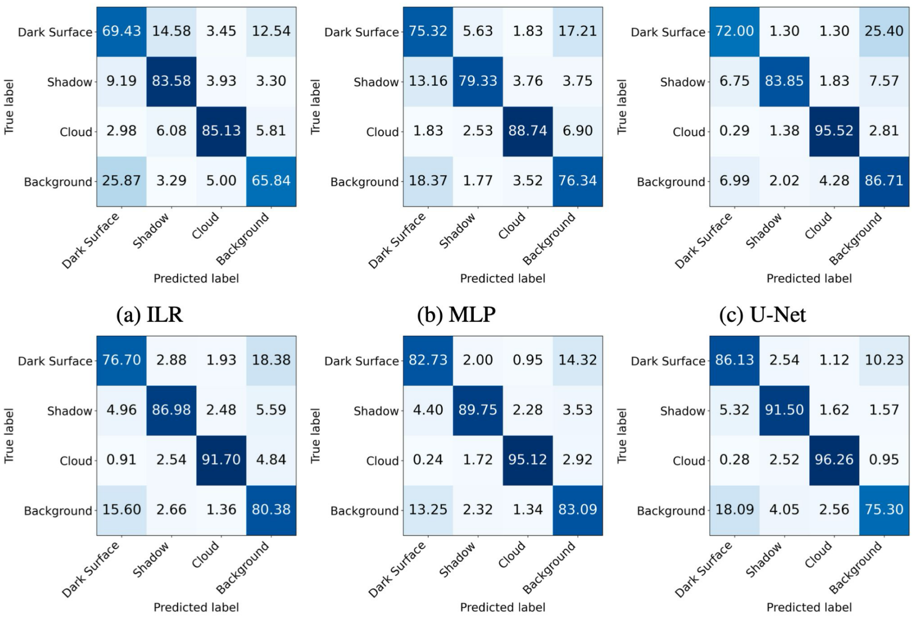
Method	Accuracy	F1-Score
SCAN	**80.33±3.43%**	**71.53±0.75%**
ViT-SegFormer	42.30+-17.4%	38.07+-12.3%
SpectralFormer	70.41+-2.94%	62.65+-1.69%
SEUNet	72.81+-8.24%	61.10+-14.2%

- SCAN achieves **11-13% higher accuracy** than self-attention methods

Computational Performance

Model	Parameters (M)	Training Time (s/epoch)	Inference Time (ms/1,000 km²)
MLP	0.022	245.4 ± 1.0	1.2 ± 0.0
U-Net	0.113	255.8 ± 7.2	2.1 ± 0.0
SCAN	0.168	290.7 ± 24.2	1.7 ± 0.0
Combined MLP	0.035	296.7 ± 12.4	4.2 ± 0.1
Combined CNN	0.026	326.6 ± 10.3	4.1 ± 0.0

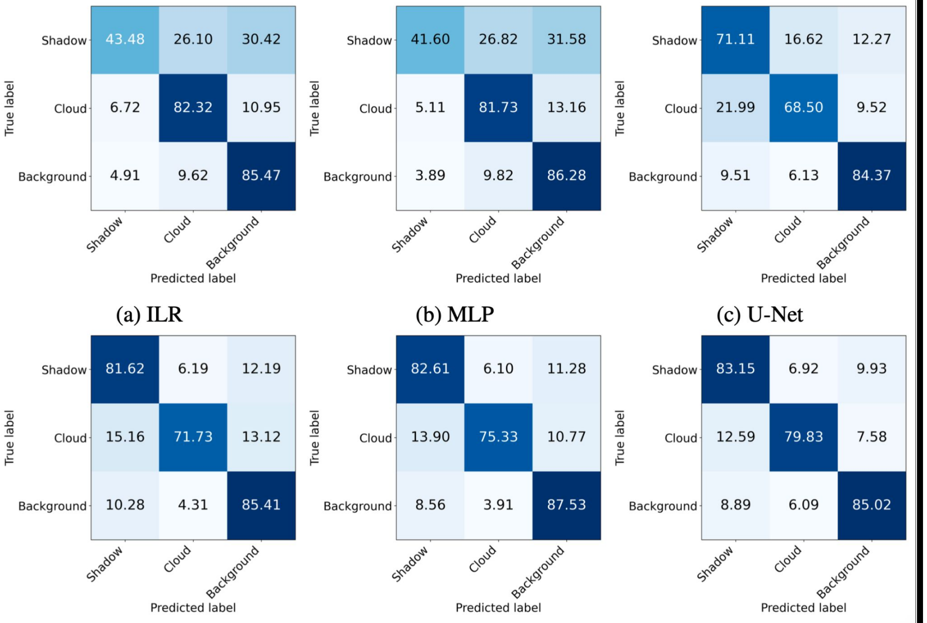
MAIR



(a) ILR (b) MLP (c) U-Net

(d) SCAN (e) Combined MLP (f) Combined CNN

MSAT



(a) ILR (b) MLP (c) U-Net

(d) SCAN (e) Combined MLP (f) Combined CNN