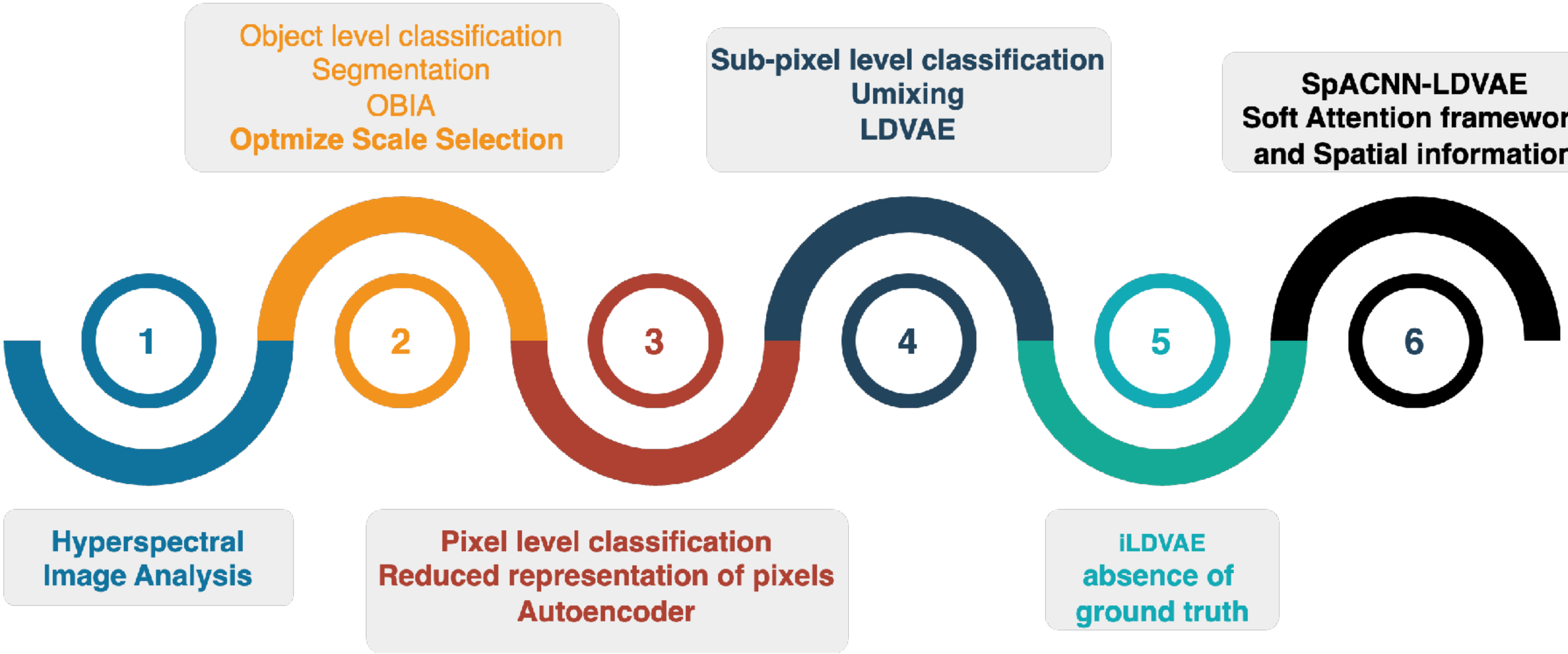
Hyperspectral Unmixing with Latent Dirichlet Variational Autoencoder

Segmentation, Classification, Dimensionality Reduction, Unmixing, Spectral Features, Spatial Features, Deep Variational Inference

Kiran Mantripragada
Advisor: Faisal Z. Qureshi
http://vclab.science.ontariotechu.ca

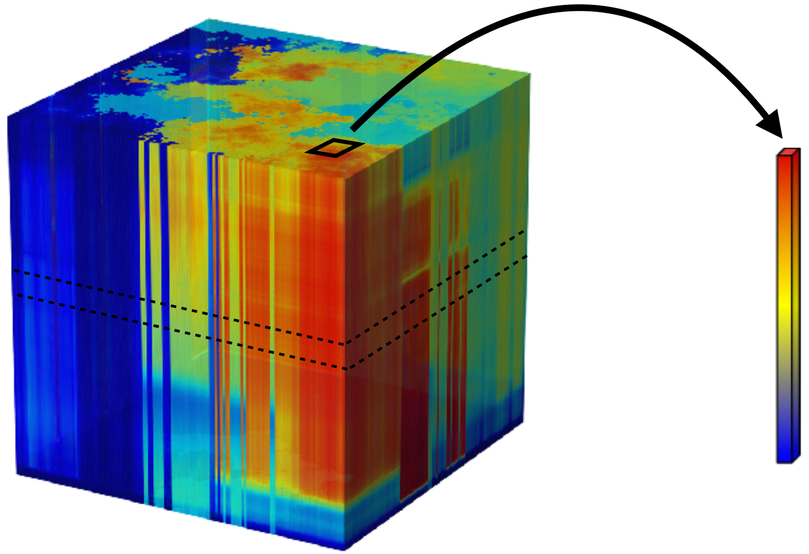
Introduction



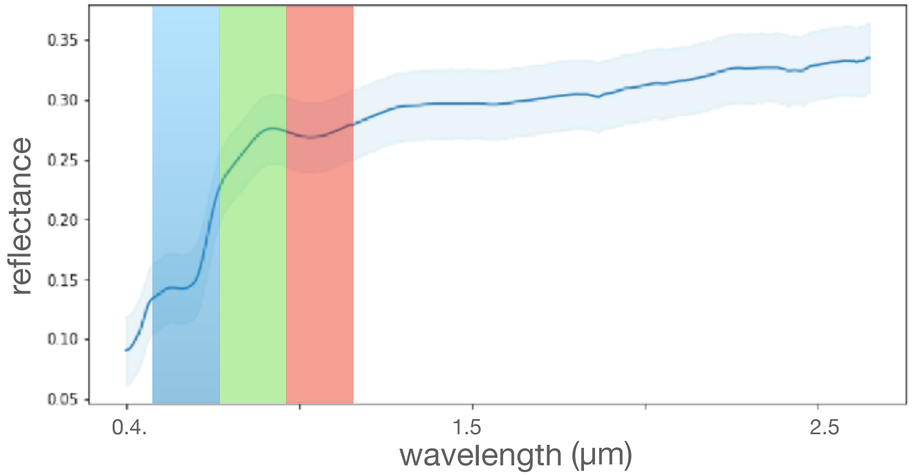
Hyperspectral Images

Image Acquisition

Satellite
Airborne
Ground based
Laboratory



Pixel spectrum



Spatial resolution

Surface area per pixel (e.g. 30 sq. Meters per pixel)

Spectral resolution

Bands per pixel (e.g. 400 bands, from 0.4 to 2.5 μm)

source: RESONON Hyperspectral Image Solutions [https://resonon.com/hyperspectral-airborne-remote-sensing-system]

Research Motivations

Hyperspectral Images

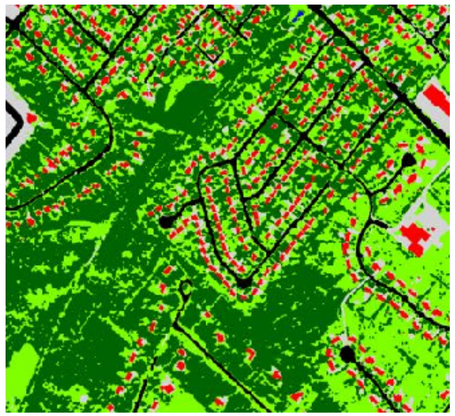
- Richness of information
- Richness of distinctive features
- Spectral and spatial data
- Remote Sensing
- Science of Materials
- Real world applications:
    - Agriculture, Forestry, Mineralogy, Urban Planning, Healthcare, and others.

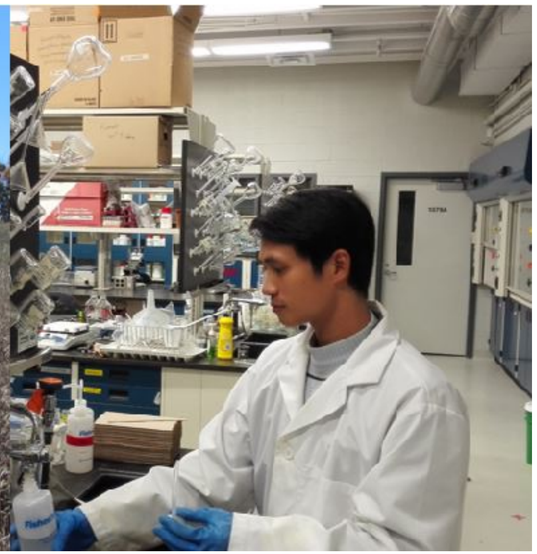
Deep Learning

- Complex classifications - Pixel Unmixing
- High dimensionality
- Computational complexity and costs
    - Storage, Transmission, Processing, Classification, model training
- Potential Data Fusion with spatial features

Hyperspectral Images and Ground-Truth

University of Toronto - Remote Sensing and Spatial Ecosystem Modeling Laboratory





source: University of Toronto - Remote Sensing and Spatial Ecosystem Modeling laboratory [http://sites.utm.utoronto.ca/yuhong/]

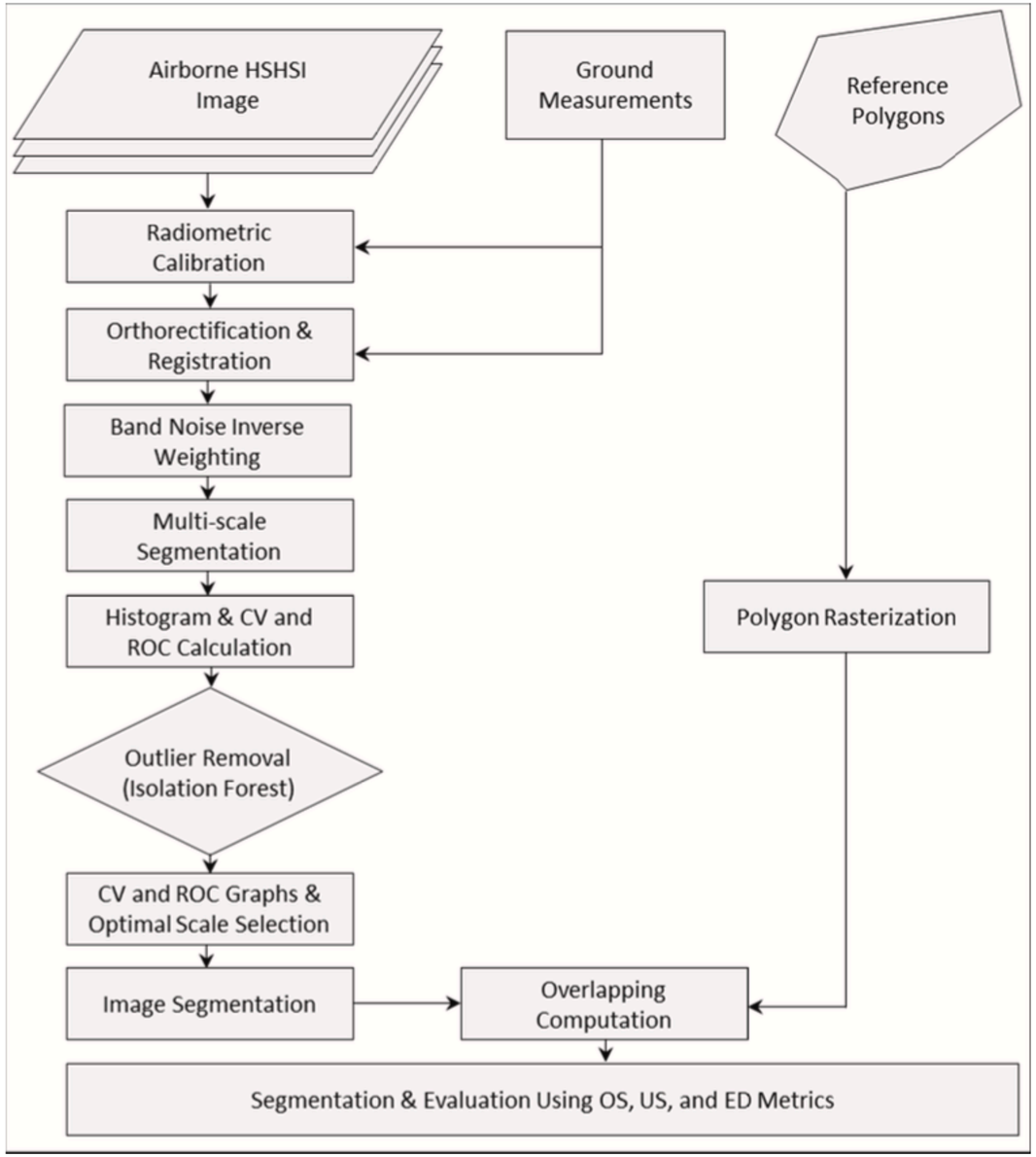
Segmentation and Classification

Dao, P. D.; Mantripragada, K.; He, Y.; and Qureshi, F. Z.; **Improving hyperspectral image segmentation by applying inverse noise weighting and outlier removal for optimal scale selection**; ISPRS Journal of Photogrammetry and Remote Sensing, 171: 348 - 366, 2021.

Segmentation

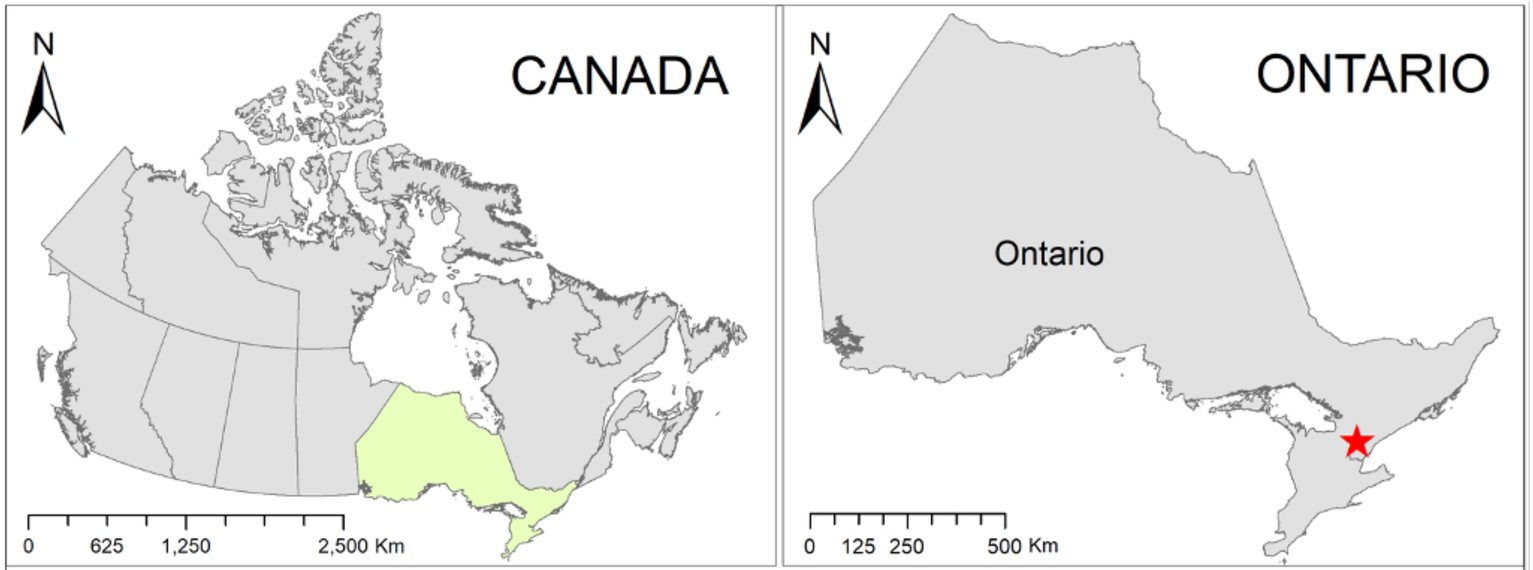
Method

- Inverse Noise Weighting
- Segmentation:
  - K-means (scale: seeds)
  - Mean-shift (scale: seeds)
  - Watershed (scale: markers)
- Coefficient of Variation
- Outlier removal
- Scale Selection
- Final segmentation

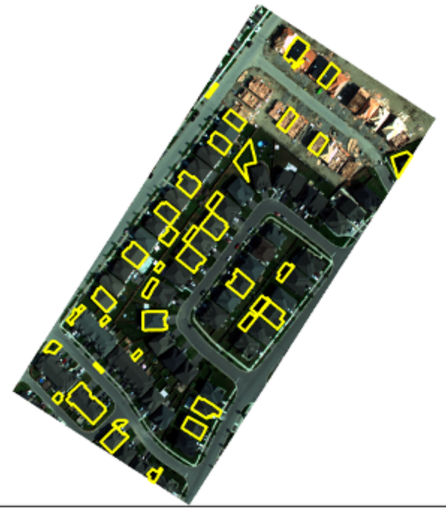


Segmentation

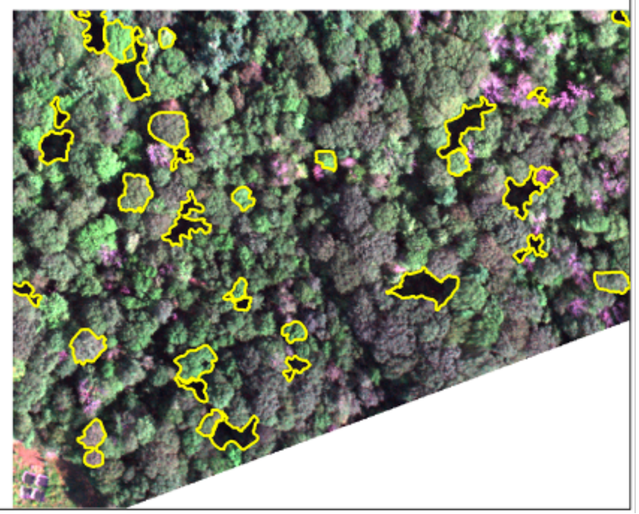
Datasets



Suburban
Bolton, Ontario



Urban
Bolton, Ontario



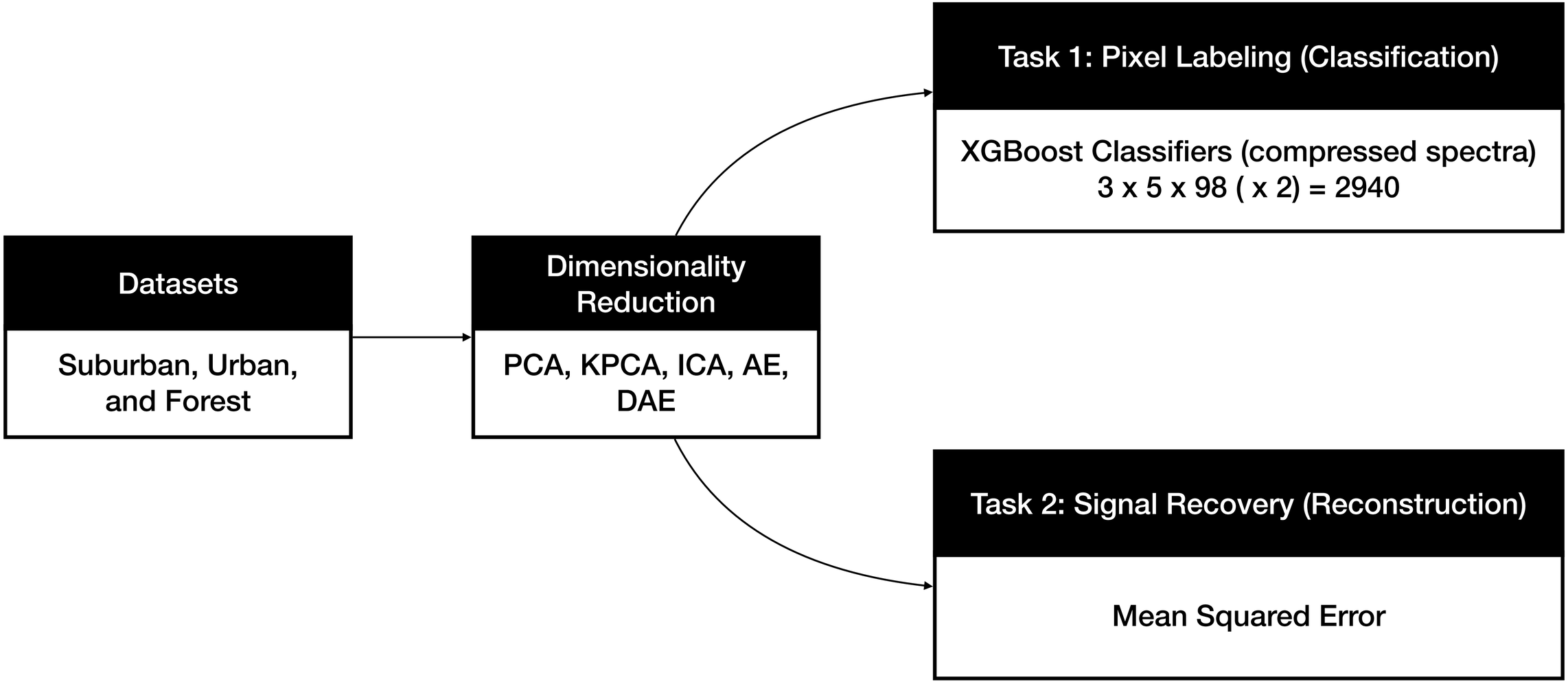
Forest
Bolton, Ontario

Dimensionality Reduction

Mantripragada K, Dao PD, He Y, Qureshi FZ (2022); **The Effects of Spectral Dimensionality Reduction on Hyperspectral Pixel Classification: A Case Study.** PLOS ONE 17(7): e0269174. https://doi.org/10.1371/journal.pone.0269174

Dimensionality Reduction

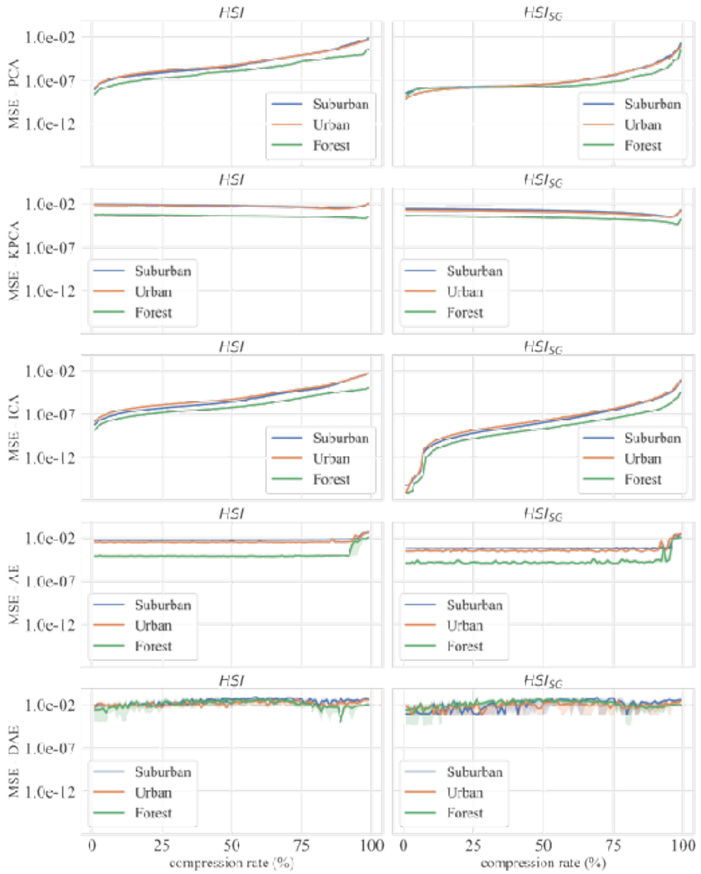
Development set up



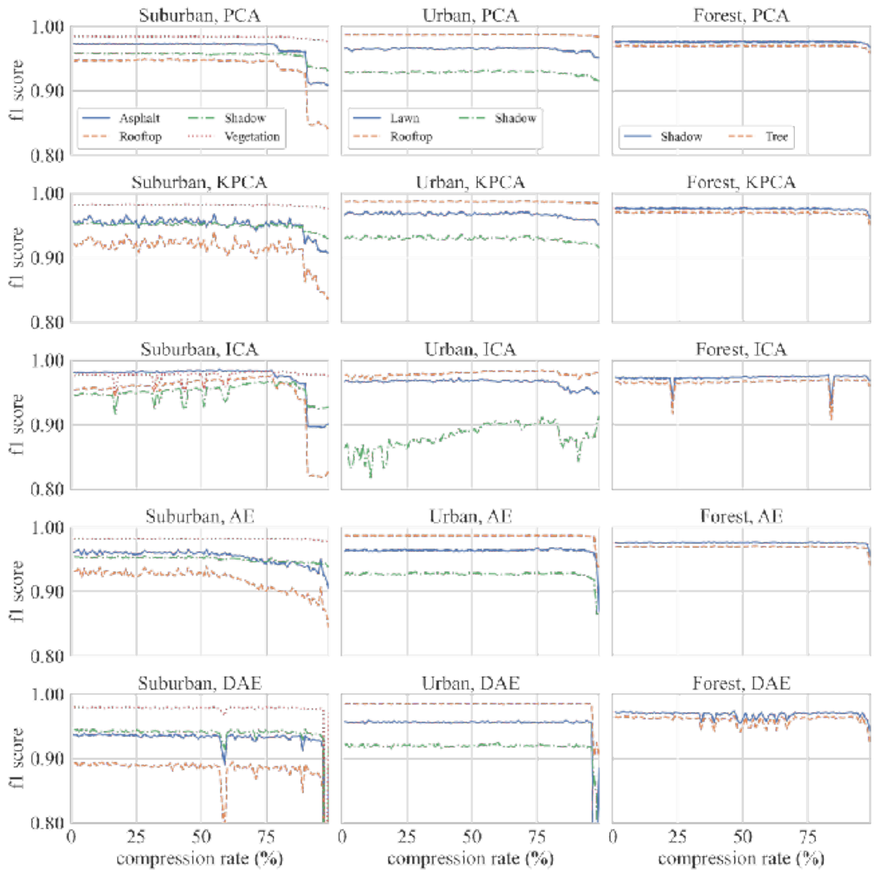
Dimensionality Reduction

Reconstruction & Classification

MSE



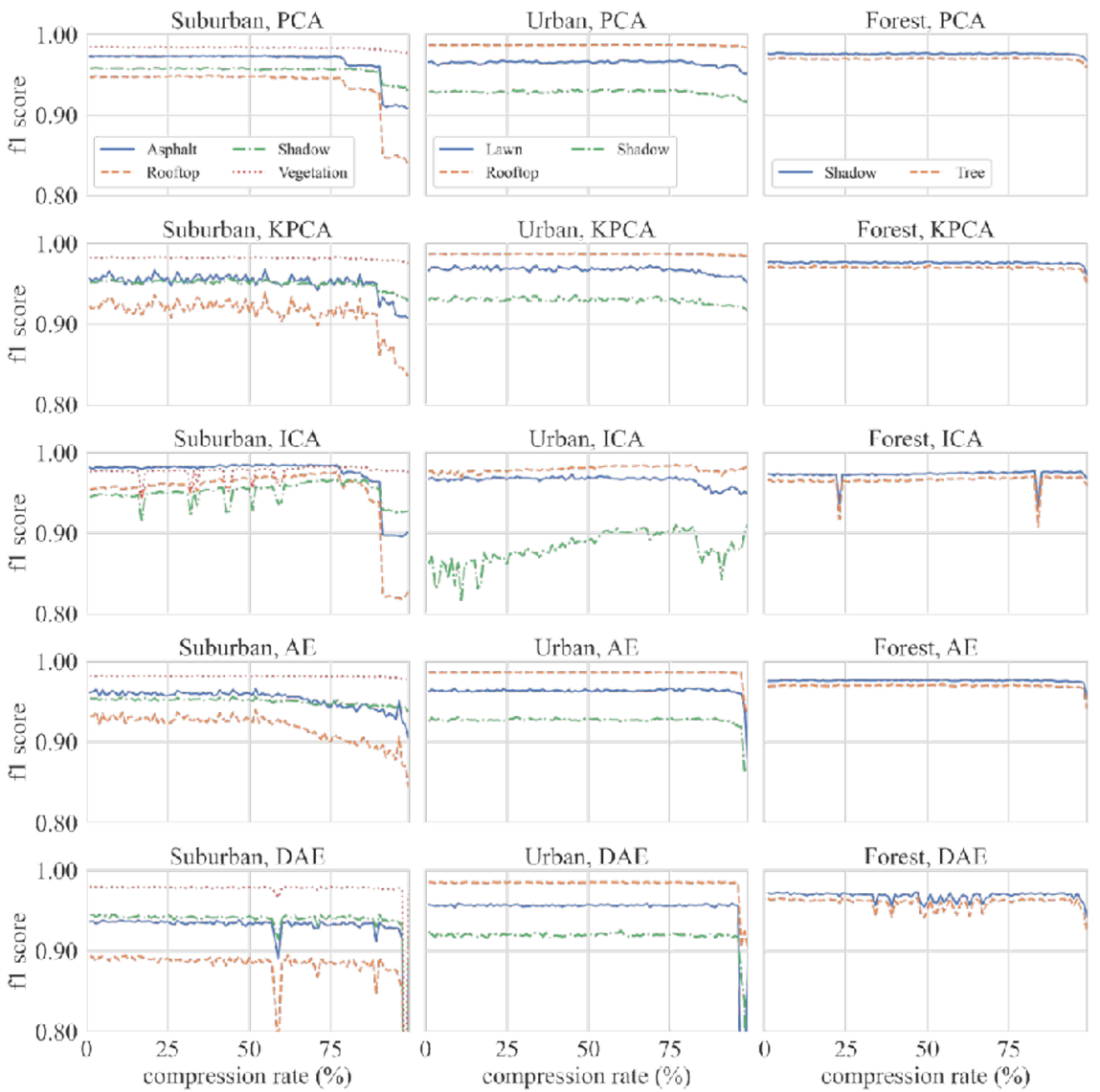
F1 scores



Dimensionality Reduction

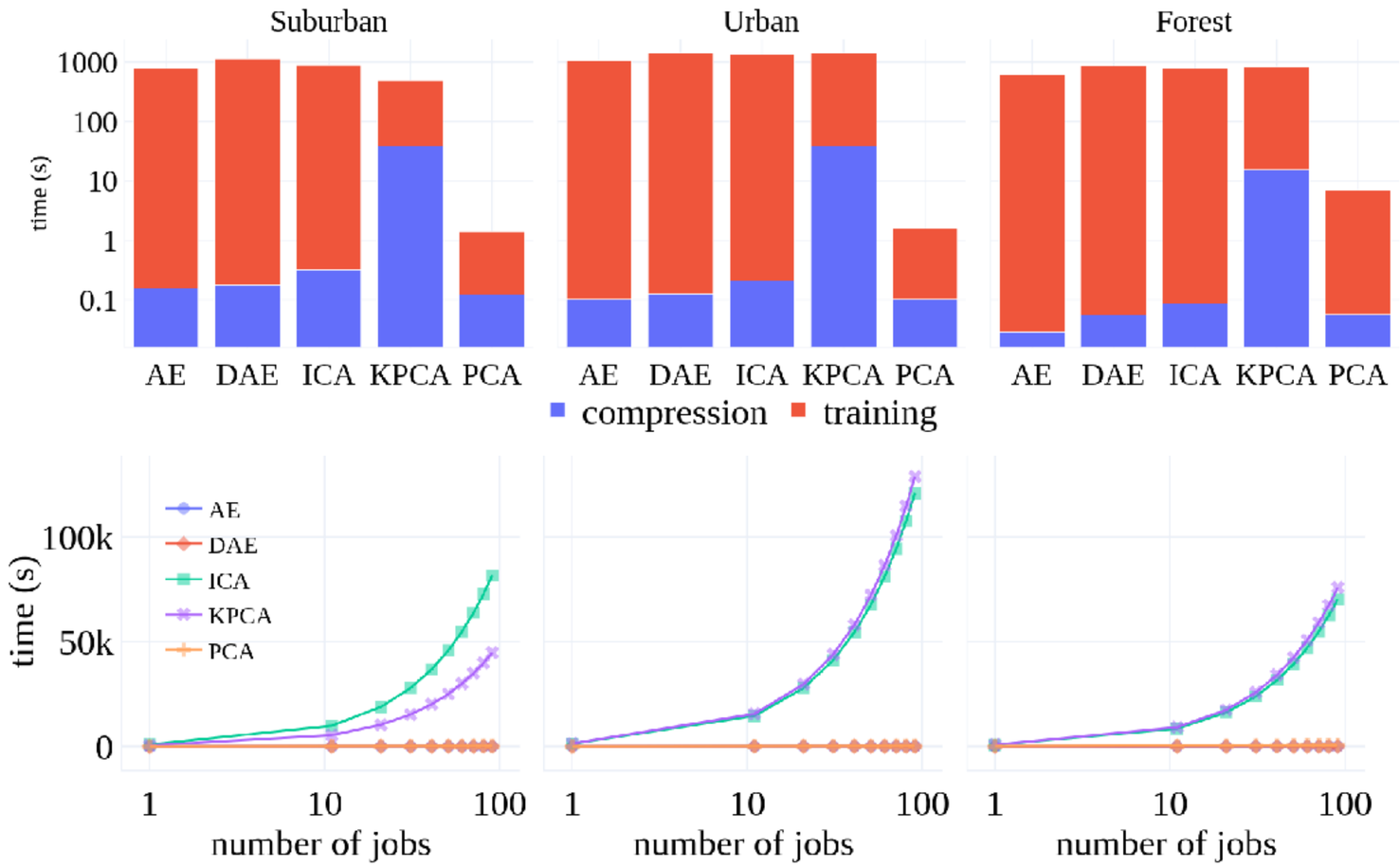
Classification

Classification f1 scores (using compressed data) vs. compression rates. The f1-scores are plotted for each label present in the dataset.



Dimensionality Reduction

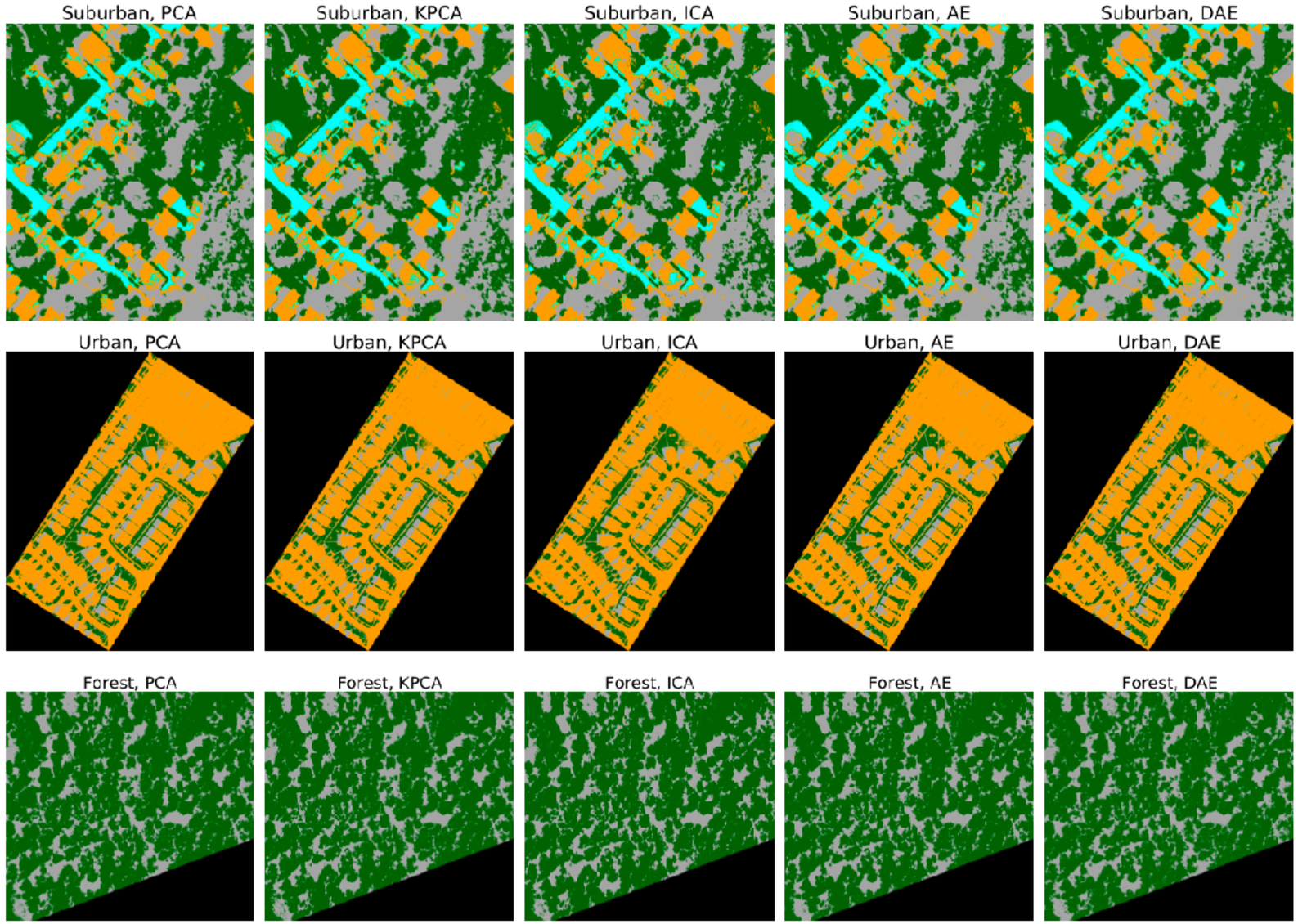
Computational Considerations



Dimensionality Reduction

Results

Classification results with spectra compressed to 95% of the original size.



Dimensionality Reduction

Conclusions

Hyperspectral data, even when compressed, outperforms RGB data for classification. This highlights the value of hyperspectral information for detailed land cover analysis.

Different compression methods are suited for different compression rates. PCA, KPCA, and ICA perform well at lower compression rates, while AE and DAE excel at higher compression rates (above 95%).

AE and DAE are particularly promising for high compression scenarios. They achieved the best classification scores when compressing data down to a very small fraction of its original size.

AE and DAE are suitable for in-situ applications. Their ability to be trained in batches makes them adaptable to resource-constrained environments.

Hyperspectral Pixel Unmixing

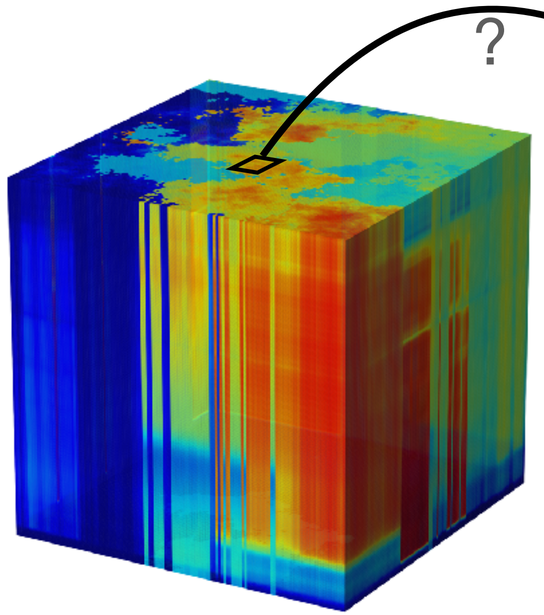
Latent Dirichlet Variational Autoencoder

Mantripragada K. and Qureshi F. Z., **Hyperspectral Pixel Unmixing With Latent Dirichlet Variational Autoencoder**, in IEEE Transactions on Geoscience and Remote Sensing, vol. 62, pp. 1-12, 2024, Art no. 5507112, doi: 10.1109/TGRS.2024.3357589.

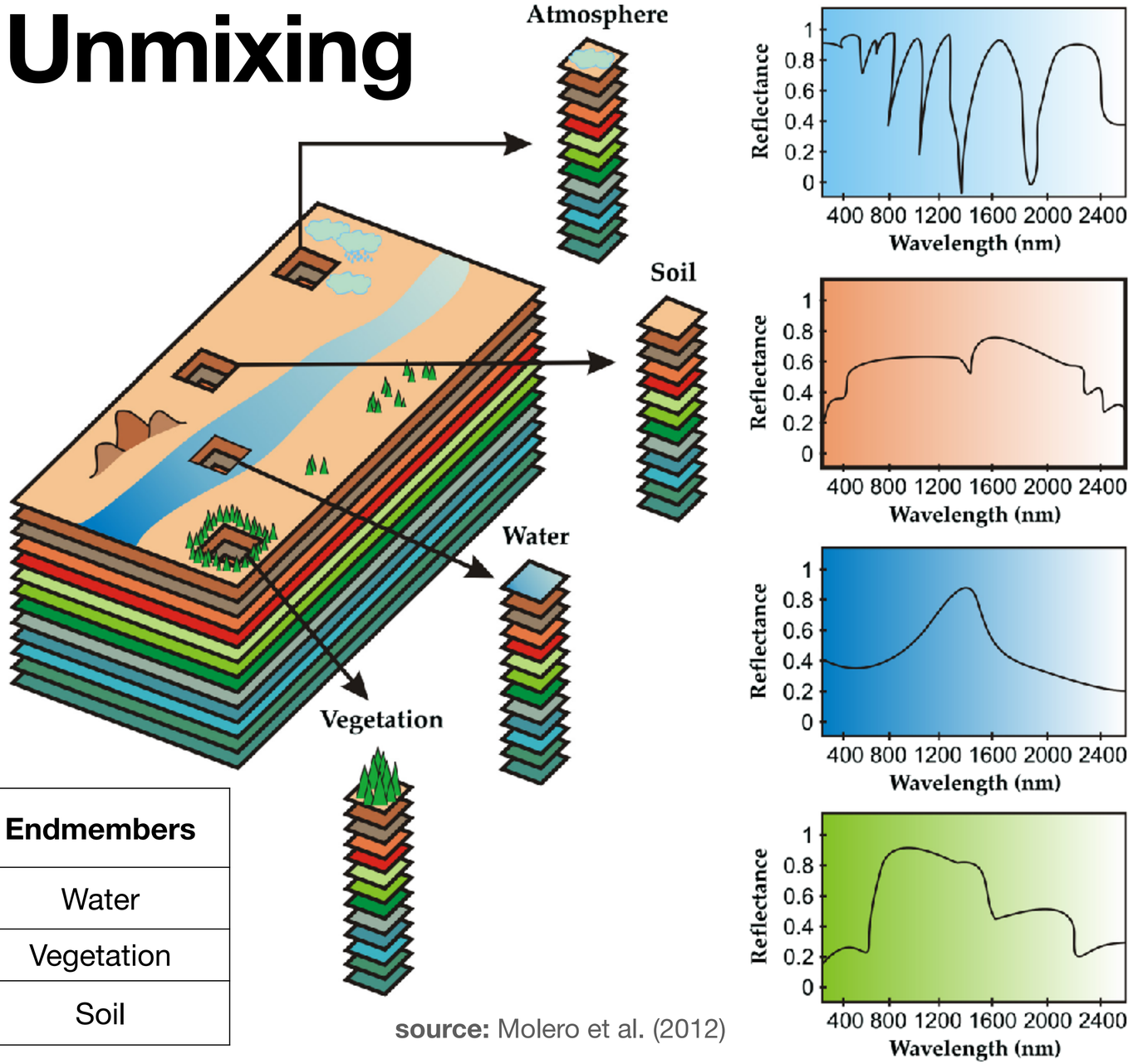
Hyperspectral Pixel Unmixing

The unmixing problem

- How to **extract the mixing ratios** of endmembers?
- How to **recover** the spectral signature of a **pure class**?



Mixing Ratios (abundances)	Endmembers
22%	Water
70%	Vegetation
8%	Soil



source: Molero et al. (2012)

Hyperspectral Pixel Unmixing
Motivation

- Spectral mixing is an important problem in Hyperspectral Image analysis
  - Earth Observation (Remote Sensing), agriculture, mineralogy, biology, botany, medicine, pharmaceutical, etc.
  - The assumption that observation is from one pure material is not realistic (Heylen, 2014)
- Linear unmixing methods are still widely used
- Physics-based models require intrinsic material properties
- We need to develop data driven approaches that eschew explicit modelling and linearity assumptions and offer advantages such as transfer learning and generalizability

Hyperspectral Pixel Unmixing

Abundances

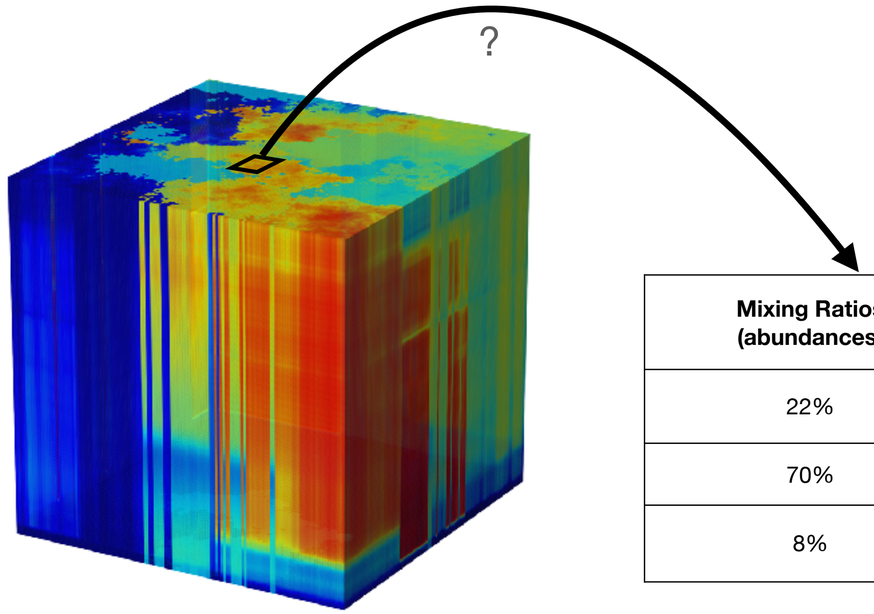
Abundances $\mathbf{z}$ are drawn from a Dirichlet distribution

$$\mathbf{z} \sim \mathrm{Dir}(\mathbf{z}; \alpha)$$

where

$$\mathbf{z} = \{z_1, z_2, z_3, ..., z_k\}$$

$$\sum_{i=1}^{k} z_i = 1 \quad \text{and} \quad z_i \geq 0$$



Mixing Ratios (abundances)	Endmembers
22%	Water
70%	Vegetation
8%	Soil

Hyperspectral Pixel Unmixing

Pixel Spectra

Also for the observed spectra $\mathbf{x}$:

$$\mathbf{x} \sim \text{Normal}(\mathbf{x}; \mu, \mathbf{\Sigma})$$

where

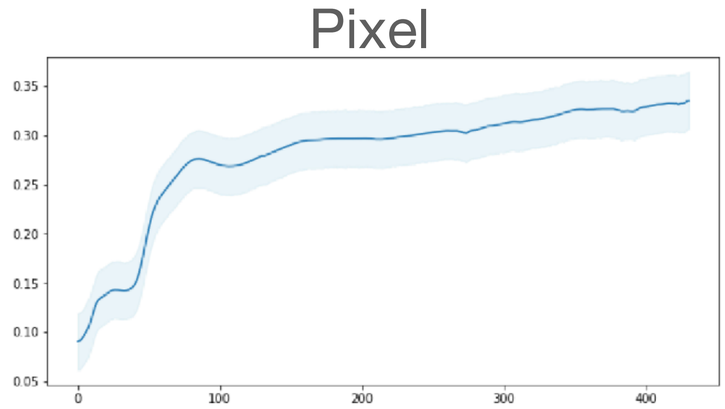
$$\mathbf{x} = \{x_1, x_2\, x_3, \ldots, x_k\}$$

$$\mu = \{\mu_1, \mu_2\, \mu_3, \ldots, \mu_k\}$$

$$\mathbf{\Sigma} = \text{diag}(\sigma_1^2, \sigma_2^2, \sigma_3^2, \ldots, \sigma_k^2)$$

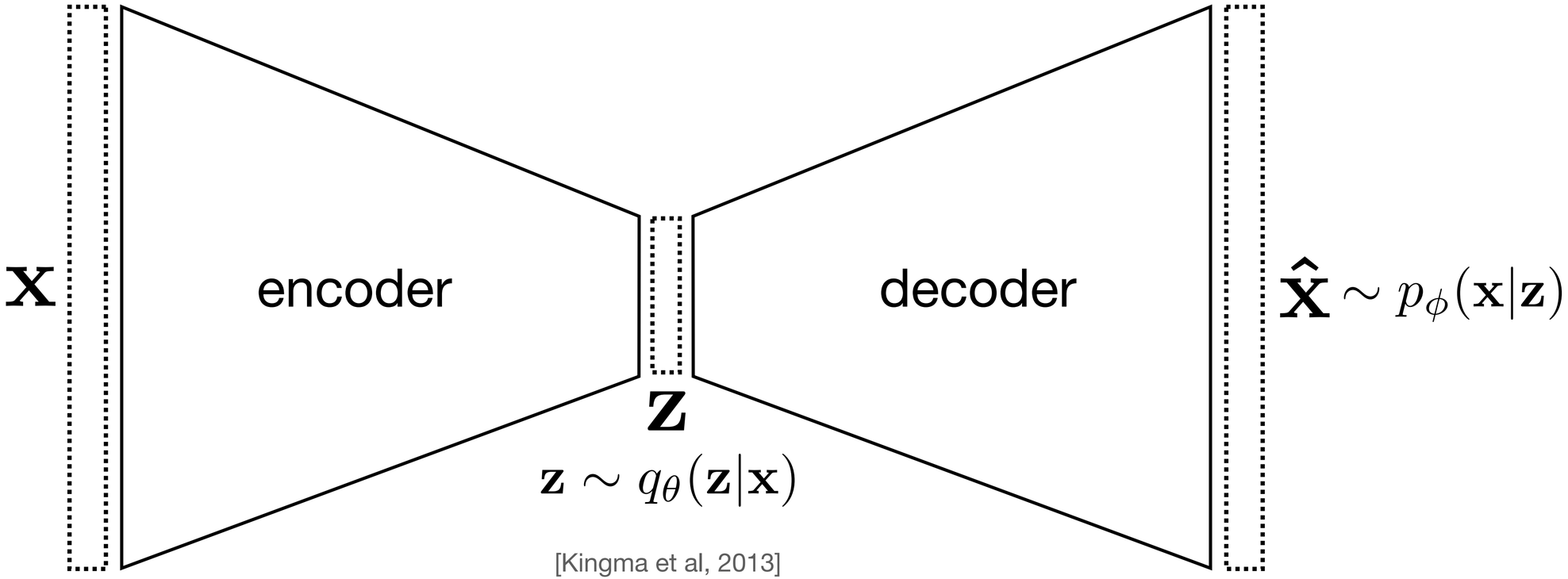
and each band value

$$\mathbf{x}_i \sim \text{Normal}(\mathbf{x}; \mu_i, \sigma_i^2)$$



Hyperspectral Pixel Unmixing

Variational Autoencoder setting

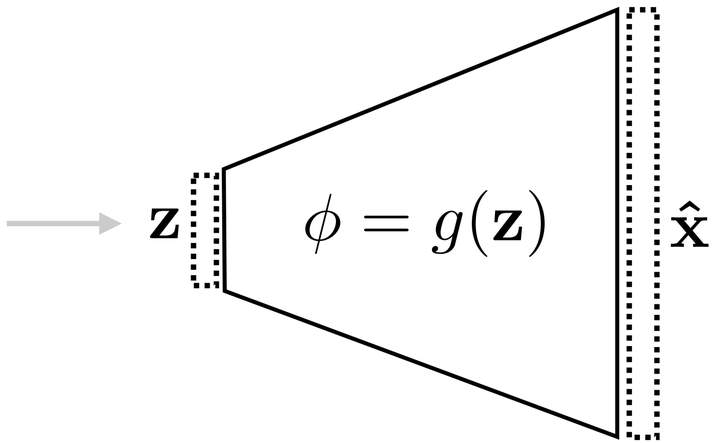


[Kingma et al, 2013]

Hyperspectral Pixel Unmixing

Decoder

- The decoder is set to:
  - **reconstruct** the spectra $\hat{\mathbf{x}} \sim p_\phi(\mathbf{x}|\mathbf{z})$
  - given the abundances $\mathbf{z}$
- where ϕ is a set of learnable parameters of the decoder



Hyperspectral Pixel Unmixing

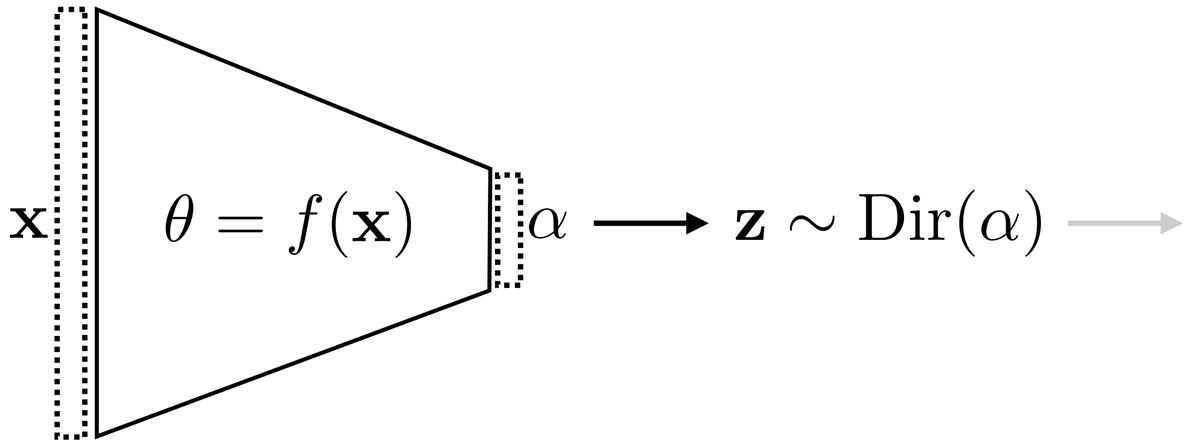
Encoder

The encoder is set to learn parameters of the distribution of abundances:

$$\mathbf{z} \sim q_\theta(\mathbf{z}|\mathbf{x}) \rightarrow \mathrm{Dir}(\alpha)$$

$$\theta = f(\mathbf{x})$$

where θ are the learnable parameters of the encoder



Hyperspectral Pixel Unmixing

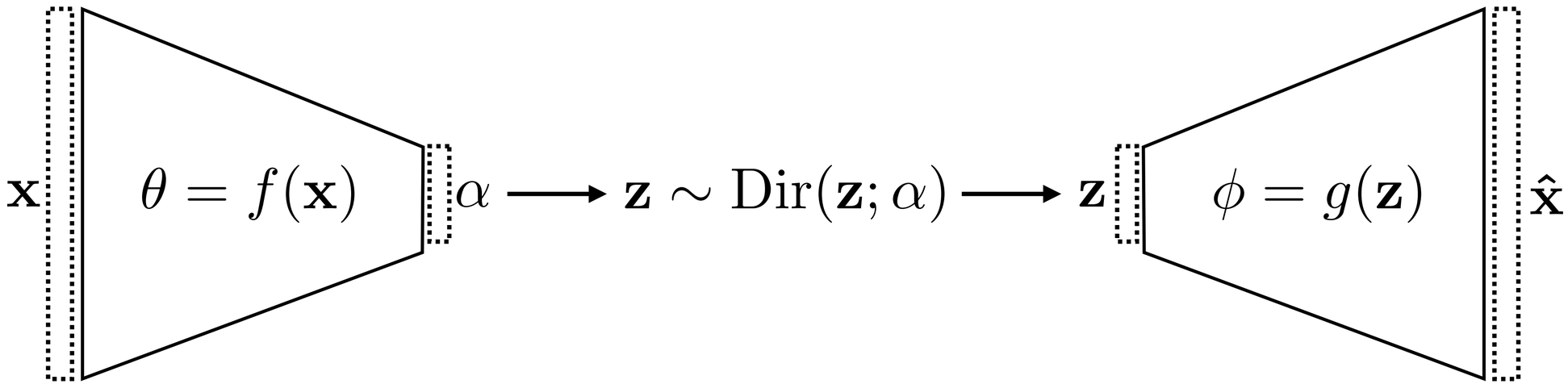
Variational Autoencoder setting - Latent Dirichlet

We aim to learn the encoder function such that:

$$\alpha = f(\mathbf{x}; \theta)$$

and

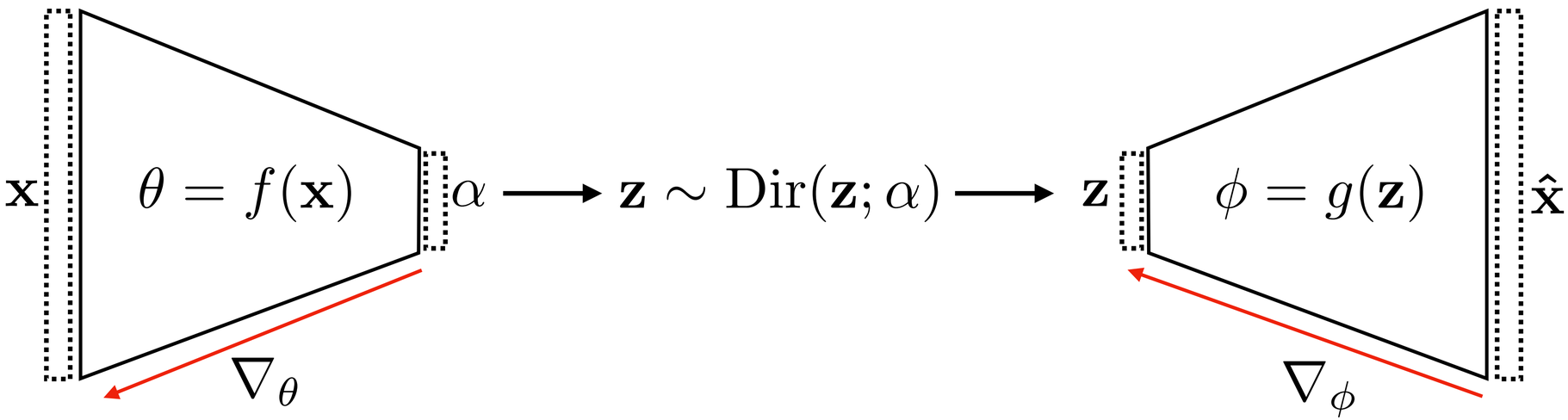
$$p(\mathbf{x}|\mathbf{z}) = g(\mathbf{z}; \phi)$$



Hyperspectral Pixel Unmixing

Reparametrization

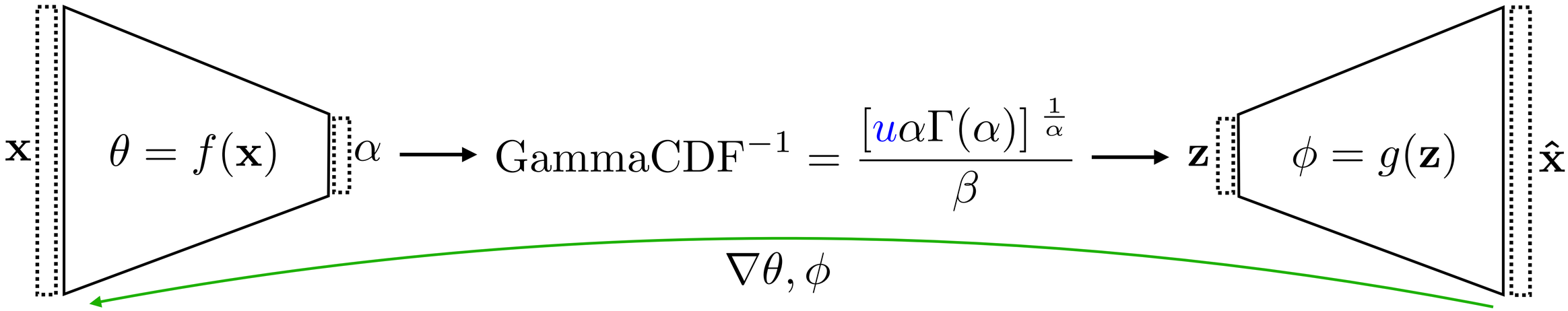
We cannot just calculate gradients and apply backpropagation due to the existence of the stochastic nodes, i.e, sampling stage [Kingma et al, 2013]



Hyperspectral Pixel Unmixing

Reparametrization

We decompose the components of the Dirichlet distribution into K Gamma random variables and use the Inverse Gamma CDF function [Joo et al, 2020]:



Hyperspectral Pixel Unmixing

KL for Dirichlet Distributions

$$\mathcal{L}(\mathbf{x};\theta,\phi) = \mathbb{E}_{q_\theta}\left[\log p_\phi(\mathbf{x}|\mathbf{z})\right] - KL(q_\theta(\mathbf{z}|\mathbf{x})\|p(\mathbf{z}))$$

We observe that a Dirichlet distribution can be expressed as a composition of multiple Gamma distributions [Joo et al, 2020]:

$$z_i \sim \Gamma(z_i;\alpha,\beta) = \text{Gamma}(z_i;\alpha,\beta) = \frac{\beta^\alpha z^{\alpha-1} e^{-\beta z_i}}{\Gamma(\alpha)}$$

and the KL divergence between MultiGamma distributions is:

$$\text{KL}\left[q(\mathbf{z},\mathbf{x};\hat{\alpha})\|p(\mathbf{z};\alpha)\right] = \sum \log\Gamma(\alpha_k) - \sum \log\Gamma(\hat{\alpha_k}) + \sum(\hat{\alpha_k}-\alpha_k)\frac{d}{dx}\ln\Gamma(\hat{\alpha_k})$$

Loss Function

$$\mathcal{L}(\mathbf{x};\theta,\phi) = \mathbb{E}_{q_\theta}\left[\log p_\phi(\mathbf{x}|\mathbf{z})\right] - KL(q_\theta(\mathbf{z}|\mathbf{x})\|p(\mathbf{z}))$$

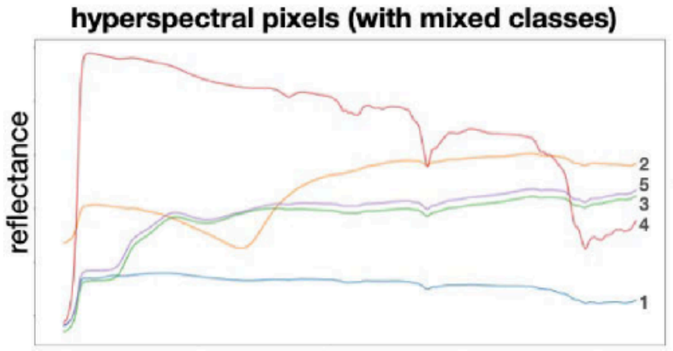
$$\mathrm{KL}\left[q(\mathbf{z},\mathbf{x};\hat{\alpha})\|p(\mathbf{z};\alpha)\right] = \sum \log\Gamma(\alpha_k) - \sum \log\Gamma(\hat{\alpha_k}) + \sum (\hat{\alpha_k} - \alpha_k)\frac{d}{dx}\ln\Gamma(\hat{\alpha_k})$$

Hyperspectral Pixel Unmixing
Benchmark datasets

- Cuprite dataset
  - 512-by-614, 188 channels, 12 endmembers
  - No training data available
    - Cuprite-Synthetic
      - Endmembers provided in Cuprite dataset
      - Random abundances
- Urban (HYDICE) dataset
  - 307-by-307, 162 channels, 6 endmembers
  - 50/50 train/test split
- Samson dataset
  - 95-by-95, 156 channels, 3 endmembers
  - 80/20 train/test split

Hyperspectral Pixel Unmixing

Abundance estimation, endmember extraction and pixel reconstruction



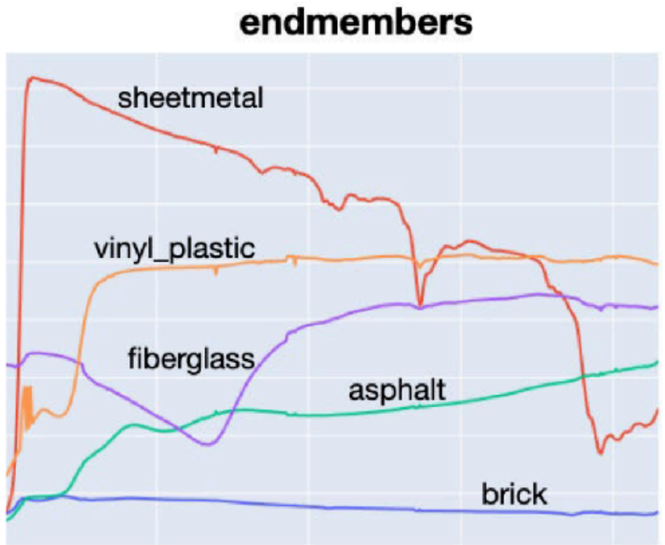
brick	sheetmetal	asphalt	fiberglass	vinyl_plastic
1	0	0	0	0
0	1	0	0	0
0	0	1	0	0
0	0	0	1	0
0	0	0	0	1

Encoder

	brick	sheetmetal	asphalt	fiberglass	vinyl_plastic
1	0.793	0.086	0.034	0.065	0.022
2	0.078	0.083	0.060	0.745	0.034
3	0.054	0.116	0.741	0.055	0.035
4	0.063	0.663	0.145	0.105	0.024
5	0.029	0.126	0.704	0.112	0.029

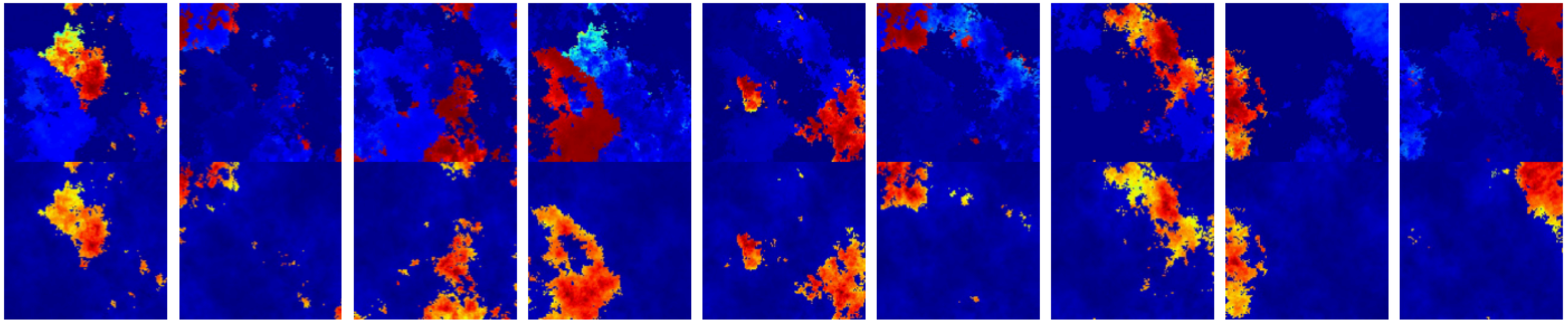
z

Decoder (generator)



Hyperspectral Pixel Unmixing

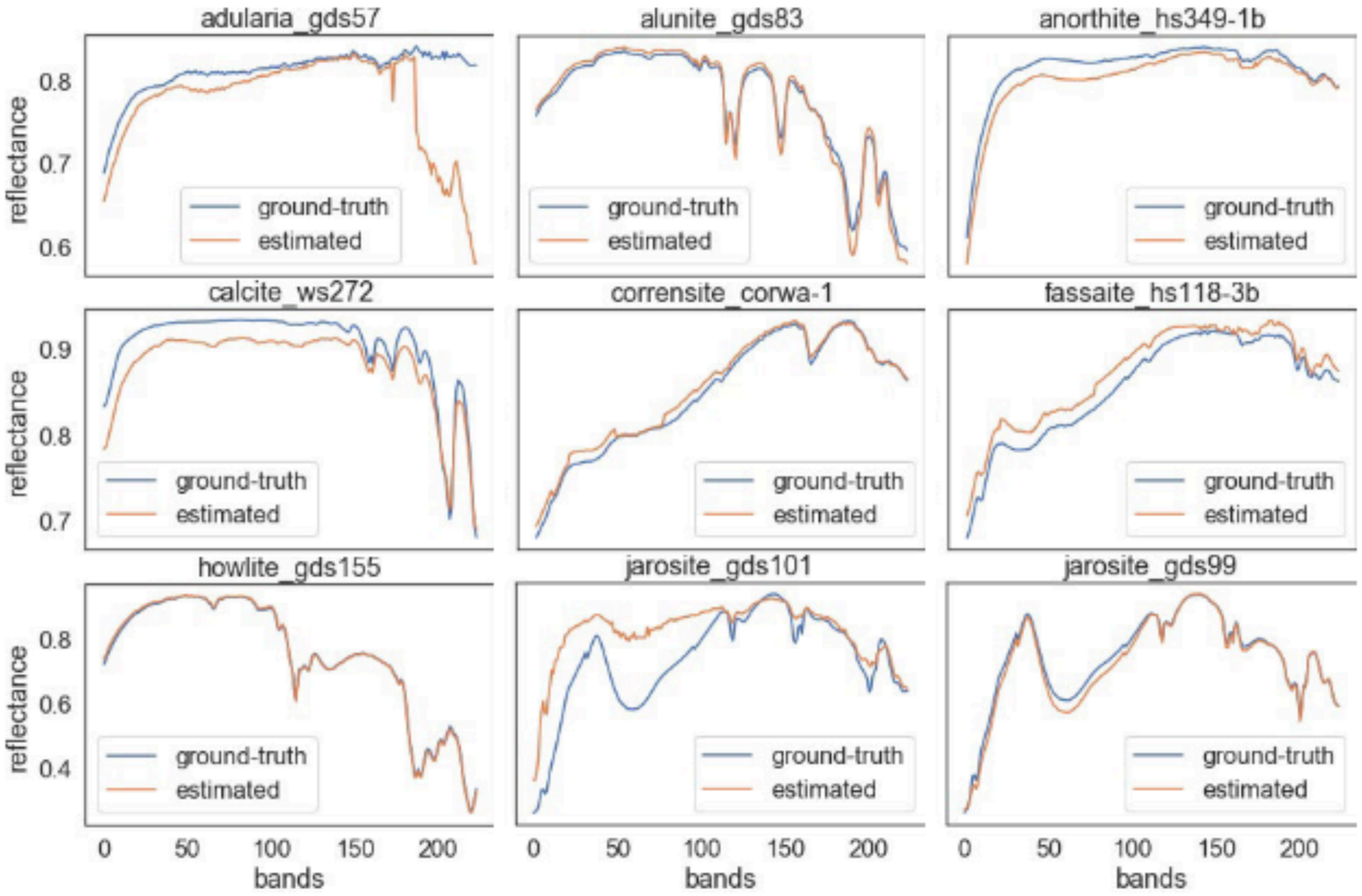
Abundance maps - Synthetic dataset



Abundance maps of the synthetic dataset (top row: LDVAE and bottom row: ground truth). From left to right: Adularia, Jarosite gds99, Jarosite gds101, Anorthite, Calcite, Alunite, Howlite, Corrensite, and Fassaite.

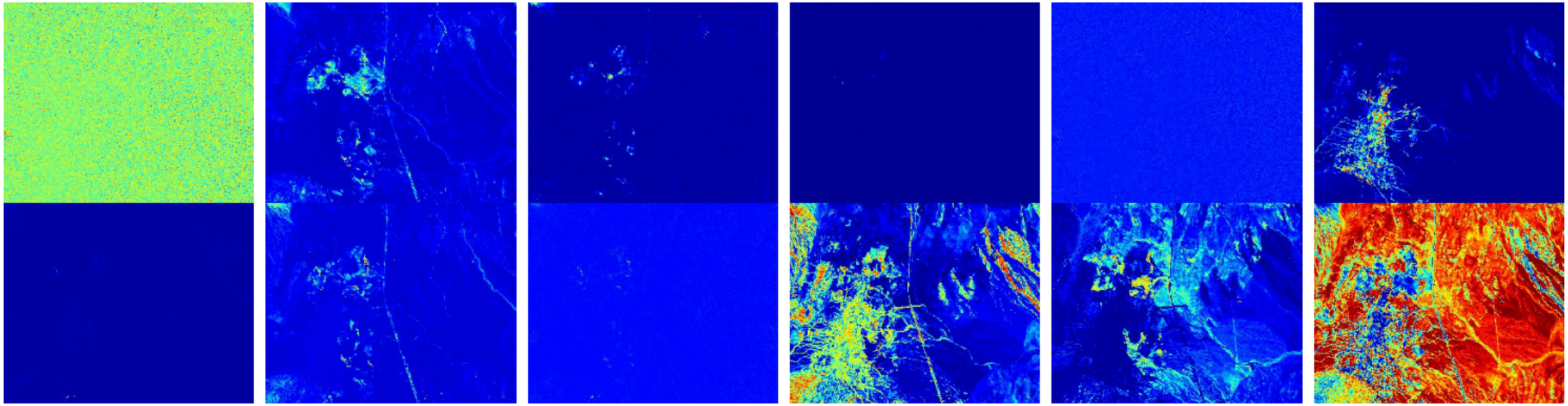
Hyperspectral Pixel Unmixing

Endmembers extraction - Synthetic dataset



Hyperspectral Pixel Unmixing

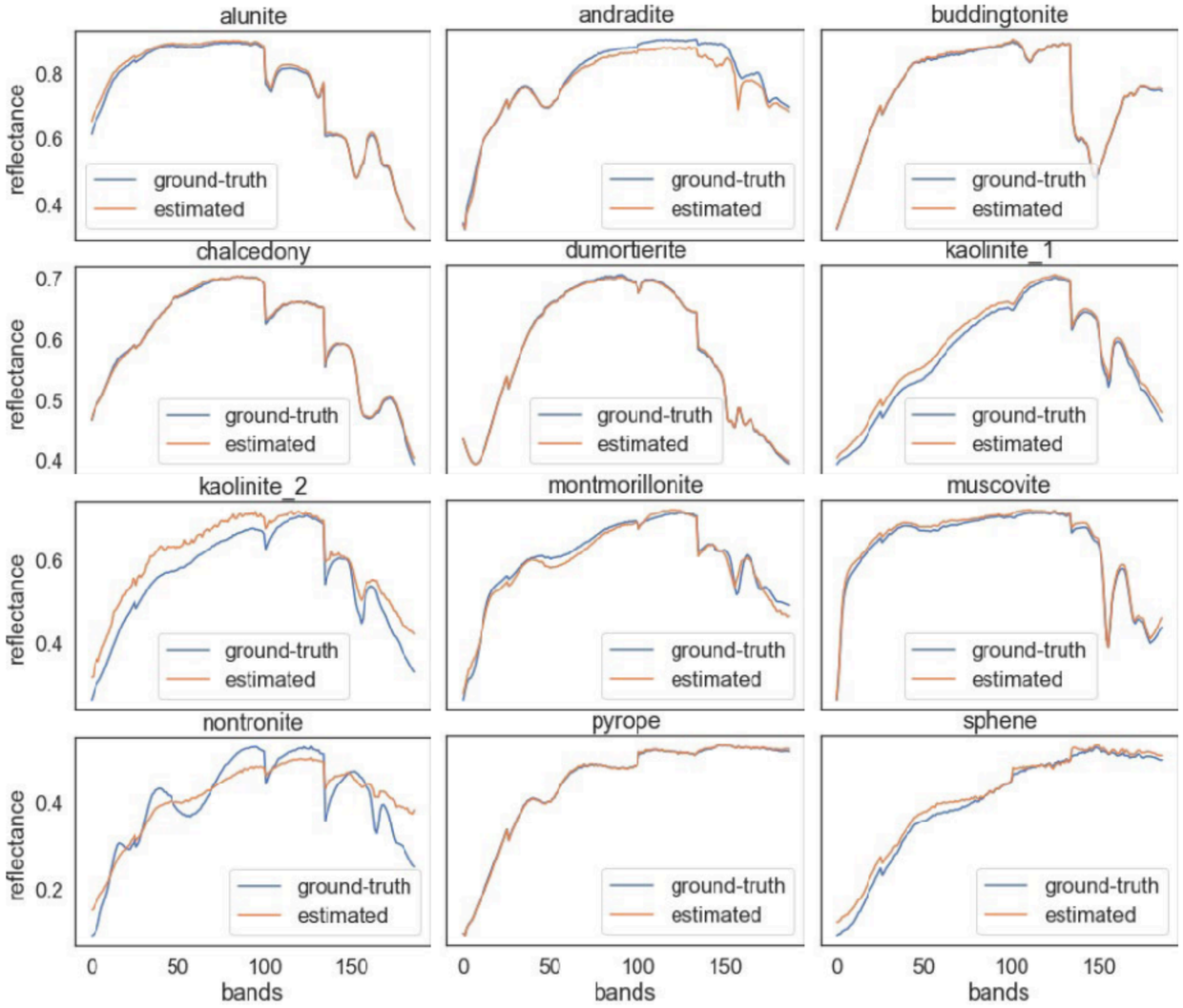
Abundance maps - Cuprite dataset



Abundance maps of the Cuprite dataset estimated by LDVAE (ground truth not available). From left to right: Alunite, Andradite, Buddingtonite, Chalcedony, Dumortierite, Kaolinite1, Kaolinite2, Montmorillonite, Muscovite Nontronite, Pyrope, and Sphene.

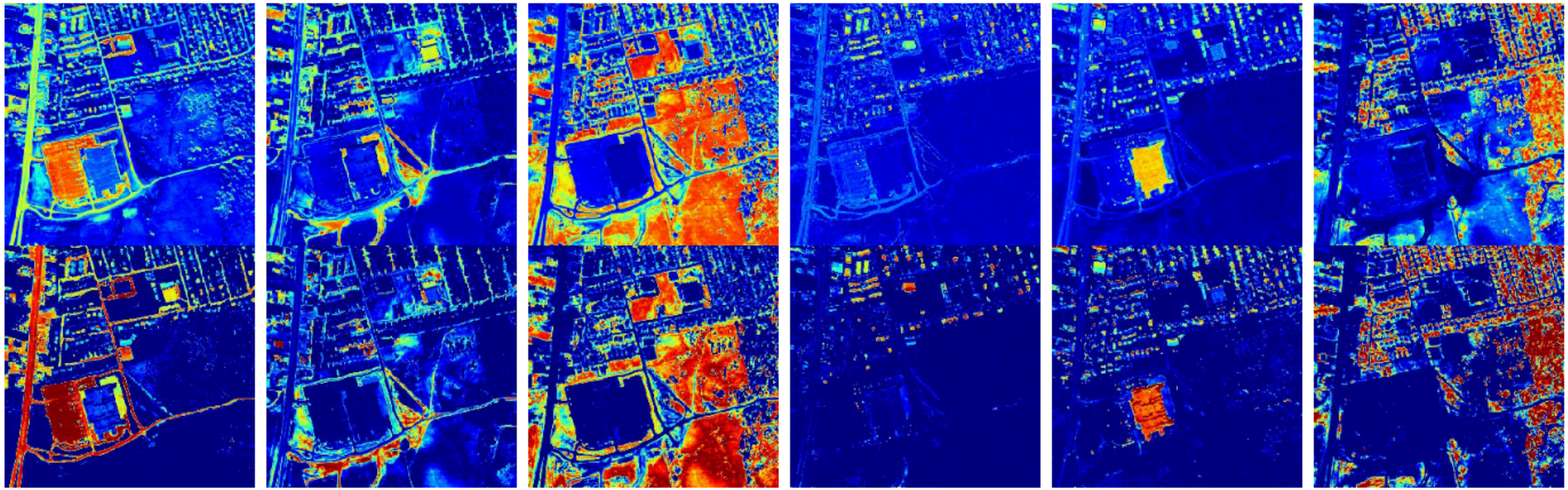
Hyperspectral Pixel Unmixing

Endmembers extraction - Cuprite dataset



Hyperspectral Pixel Unmixing

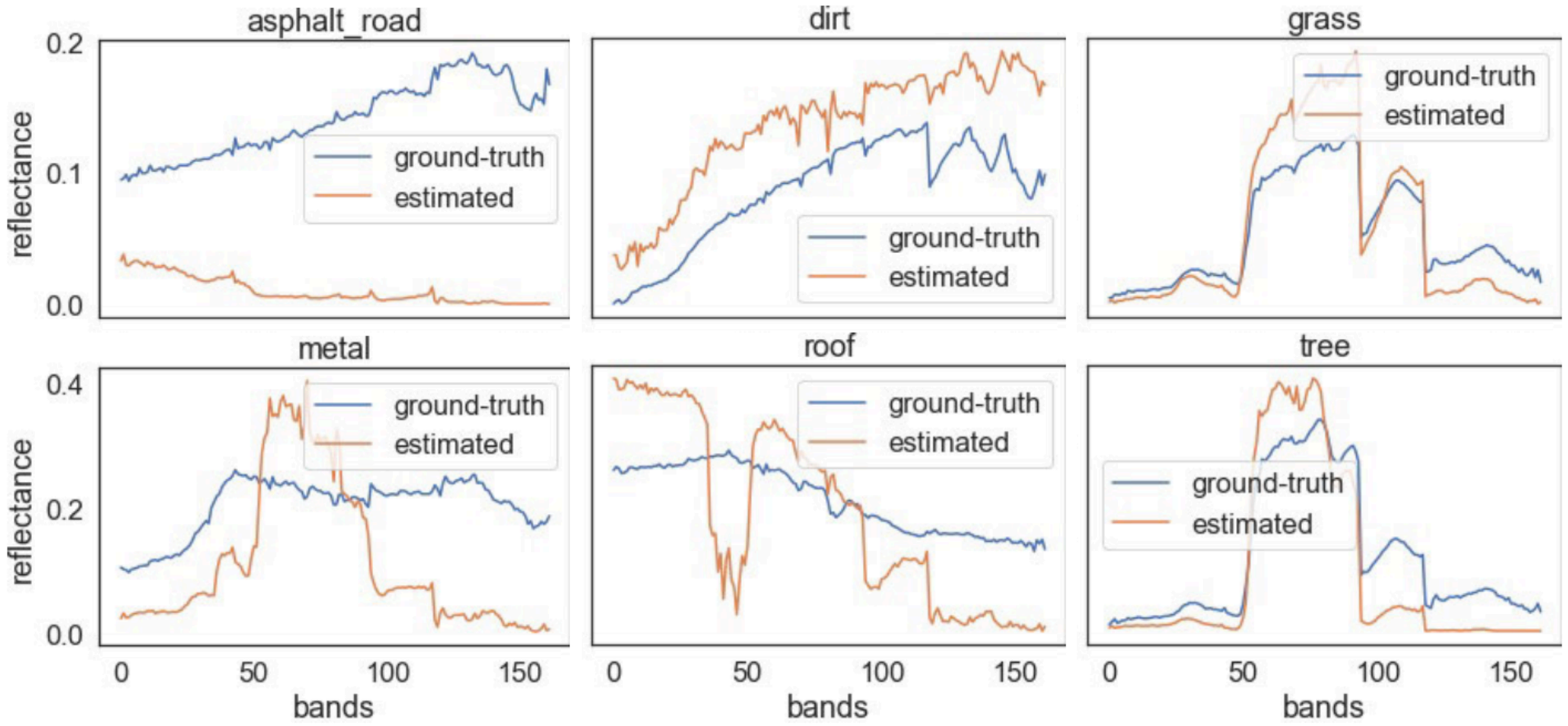
Abundance maps - HYDICE Urban dataset



Abundances maps of the HYDICE Urban dataset (top row: LDVAE and bottom row: ground truth). From left to right: asphalt, dirt, grass, and metal.

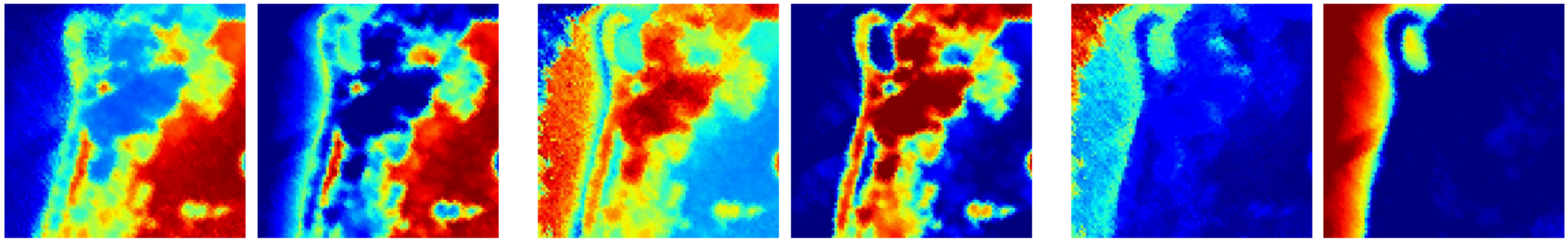
Hyperspectral Pixel Unmixing

Endmembers extraction - HYDICE Urban dataset



Hyperspectral Pixel Unmixing

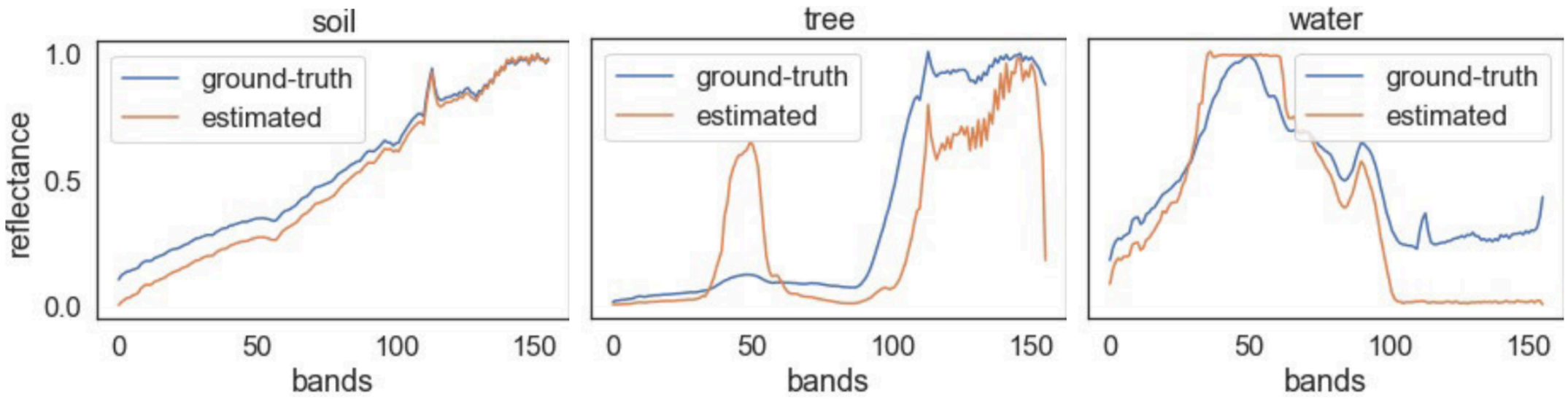
Abundance maps - Samson dataset



Abundances maps of the Samson dataset. From left to right, respectively: soil (LDVAE), soil (ground truth), tree (LDVAE), tree (ground truth), water (LDVAE), and water (ground truth).

Hyperspectral Pixel Unmixing

Endmembers extraction - Samson dataset



Hyperspectral Pixel Unmixing

Conclusions

Effective Unmixing: The LDVAE model effectively performs both endmember extraction and abundance estimation. It utilizes a Dirichlet distribution to represent abundances and a multivariate normal distribution for endmember spectra.

Spectral Reconstruction: The decoder stage of the LDVAE can reconstruct both mixed and pure pixel spectra accurately, demonstrating its ability to synthesize hyperspectral data.

Transfer Learning Capability: The model can be trained on synthetic data and successfully applied to real data lacking pixel-level abundance information. This was demonstrated with the Cuprite dataset.

Comparable Performance: The model achieves comparable and sometimes superior performance to other commonly used hyperspectral pixel unmixing techniques.

Overall, the paper demonstrates the effectiveness of the LDVAE model for hyperspectral pixel unmixing, showcasing its ability to learn from synthetic data and achieve strong performance on both synthetic and real-world datasets.

Iterative LDVAE

Overcoming the lack of ground truth

Mantripragada, K. and Adler, P. R. and Olsen P. A. and Qureshi F. Z.; **An Iterative Method for Hyperspectral Pixel Unmixing Leveraging Latent Dirichlet Variational Autoencoder**, IGARSS 2023 - 2023 IEEE International Geoscience and Remote Sensing Symposium, Pasadena, CA, USA, 2023, pp. 7527-7530, doi: 10.1109/IGARSS52108.2023.10282175.

Iterative LDVAE

Motivation

Overcome the limitations of labeled data in hyperspectral pixel unmixing: Traditional machine learning-based unmixing methods require large amounts of pixel-level labeled data, which is often costly and difficult to obtain.

Leveraging the power Latent Dirichlet Variational Autoencoder (LDVAE) on an iterative analysis-synthesis loop to learn endmember spectra and estimate their abundances in a self-supervised manner.

Developing a practical and effective unmixing solution: Create a method that can be applied to real-world hyperspectral analysis tasks, such as crop analysis, where obtaining labeled data is often impractical.

Providing a novel approach to hyperspectral unmixing: The iLDVAE method offers a new perspective on solving the unmixing problem by combining deep learning with an iterative approach, eliminating the reliance on labeled training data.

Iterative LDVAE

Motivation - Collaboration with USDA and Microsoft

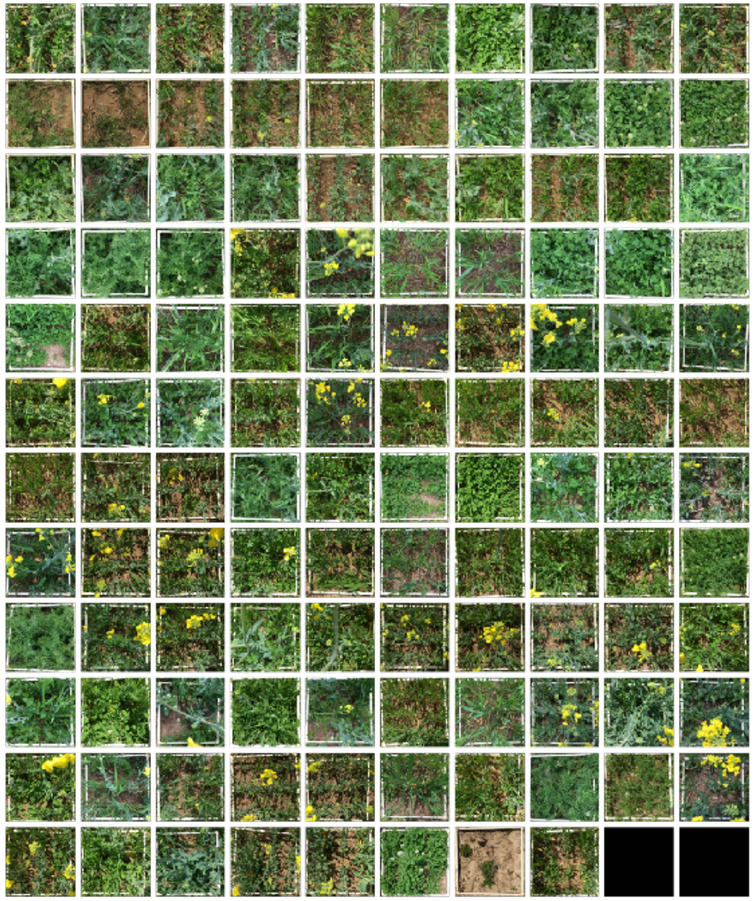
Since the traction of LDVAE after the pre-print submission to arXiv, we defined a research collaboration with United States Department of Agriculture (USDA) and Microsoft.

For this investigation, we explored the Cover Crop USDA dataset, which is a Hyperspectral imagery dataset collected from field plots using a DJI Matrice 600 Pro which collects data from 270 spectral bands over the range of 400nm to 1000nm with a spectral resolution of 2.2nm.

The imagery was collected by the United States Department of Agriculture (USDA) over 4 species of vegetation: canola, clover, triticale, and vetch.

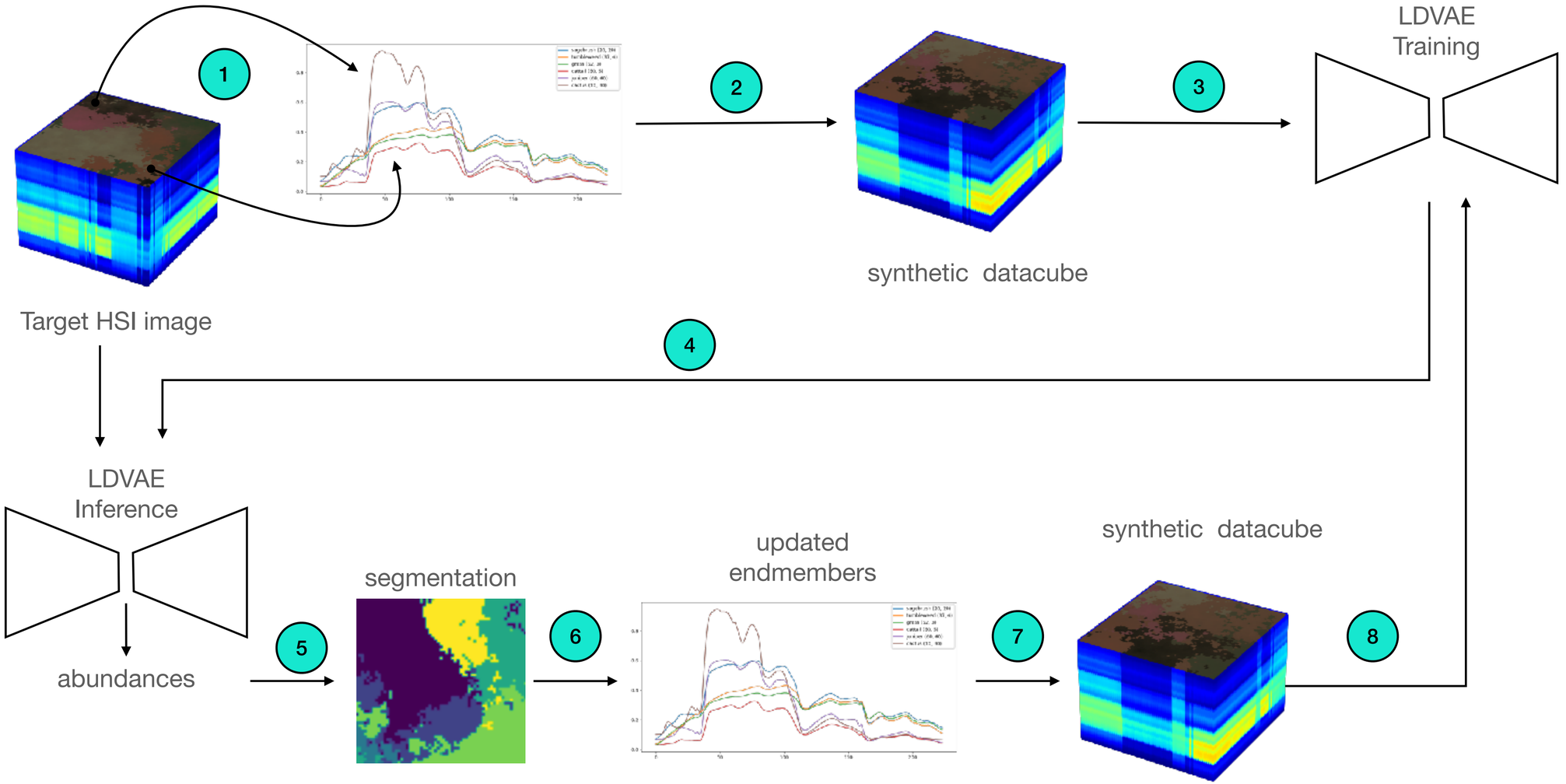
Additionally the class “soil” is also part of the categories to represent all the materials present in a scene.

There are 120 datacubes, 40 of them categorized as monocultures, i.e. , datacubes with only one of the species plus soil. This figure shows the 120 RGB images taken from regular iPhone cameras.



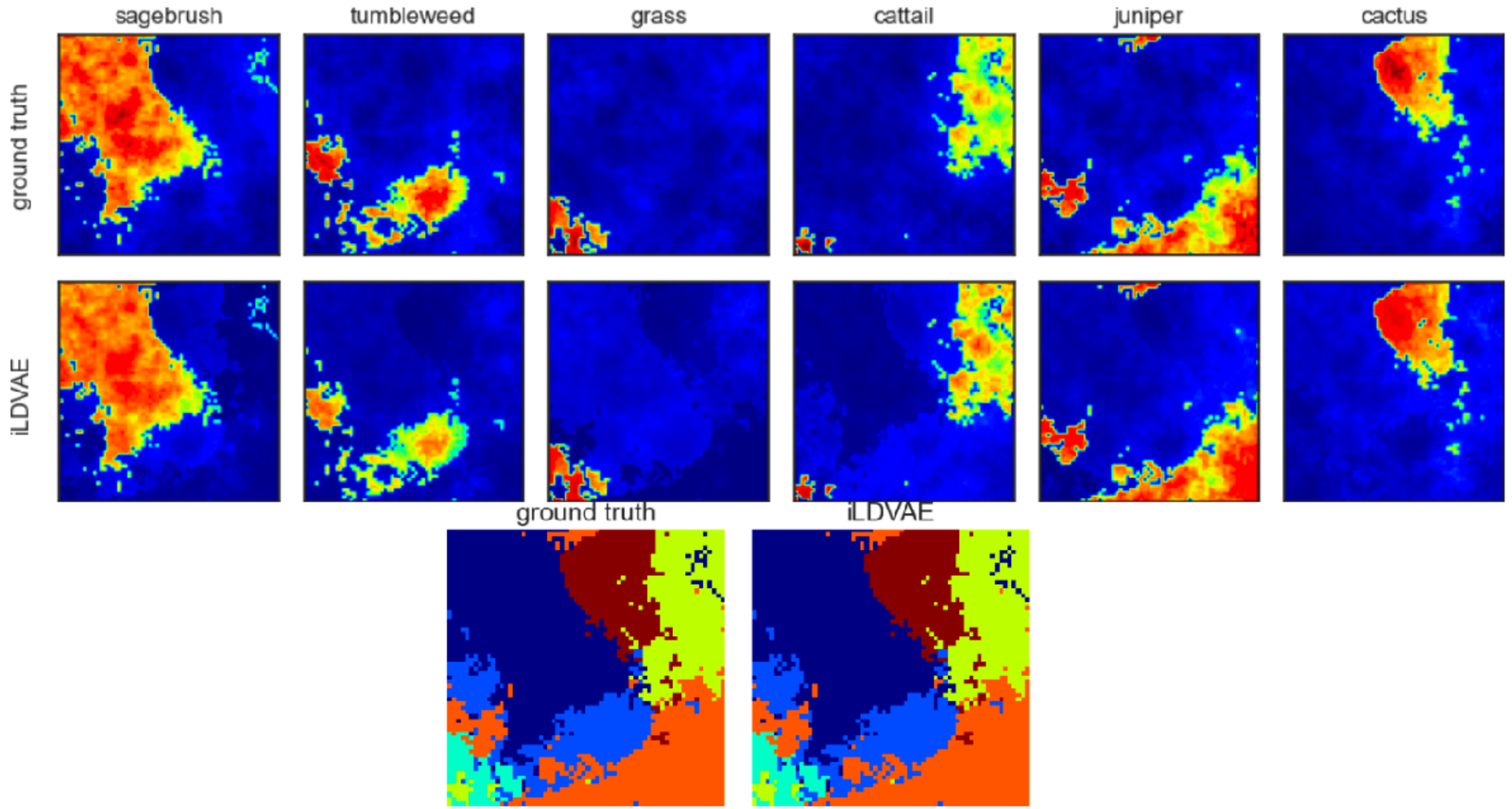
Iterative LDVAE

Method



Iterative LDVAE

Results - Synthetic dataset

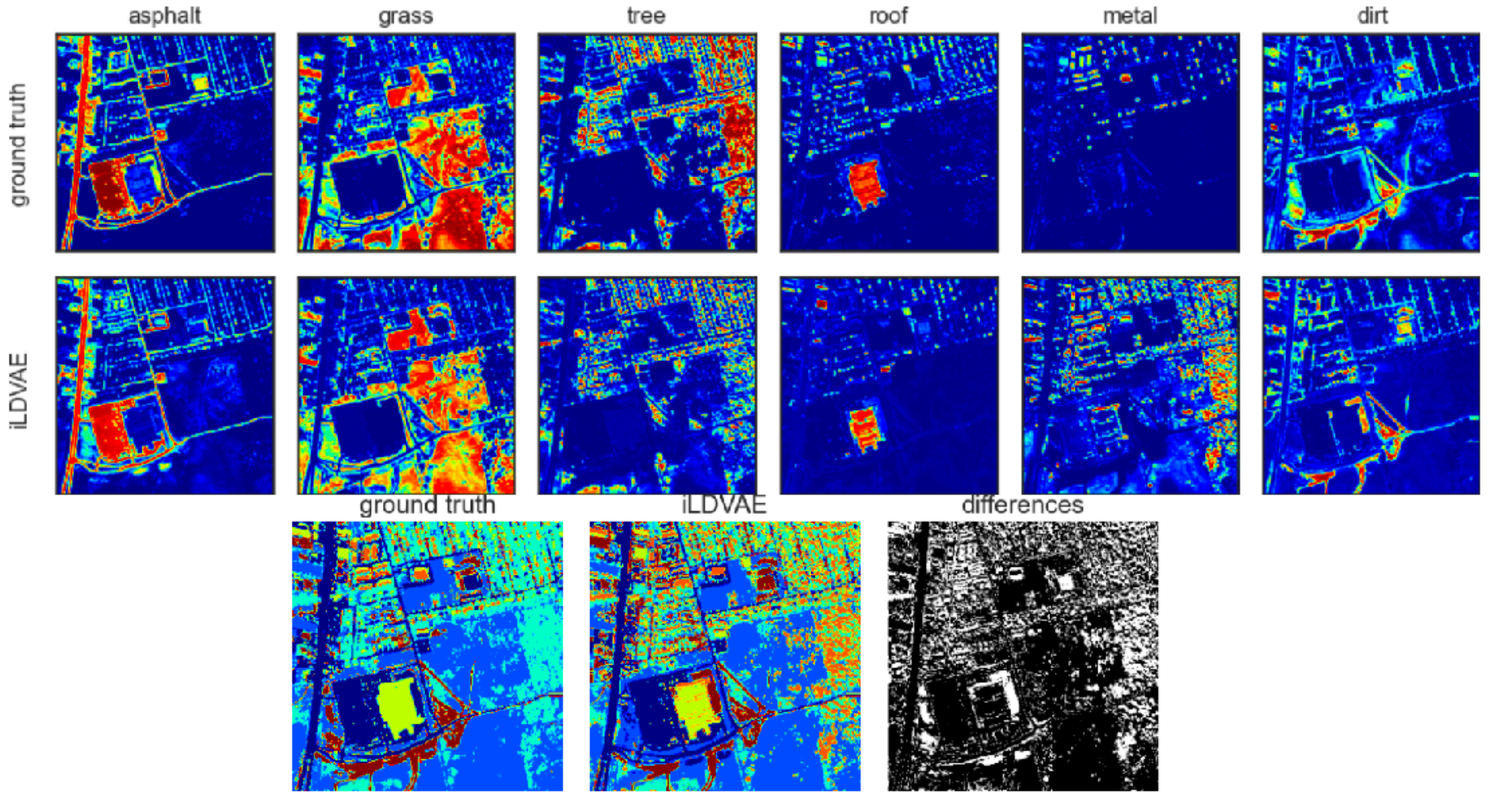


	SAD	RMSE
sagebrush	0.0447 ±0.0005	0.0664 ±0.0001
tumbleweed	0.0853 ±0.0005	0.0712 ±0.0002
grass	0.0233 ±0.0009	0.0342 ±0.0001
cattail	0.0721 ±0.0021	0.0899 ±0.0000
juniper	0.0429 ±0.0014	0.0796 ±0.0001
cactus	0.0810 ±0.0023	0.0758 ±0.0001
average	0.0582 ±0.0248	0.0695 ±0.0191

Abundances maps estimated by iLDVAE and Segmentation results (accuracy = 1.00)

Iterative LDVAE

Results - HYDICE Urban dataset

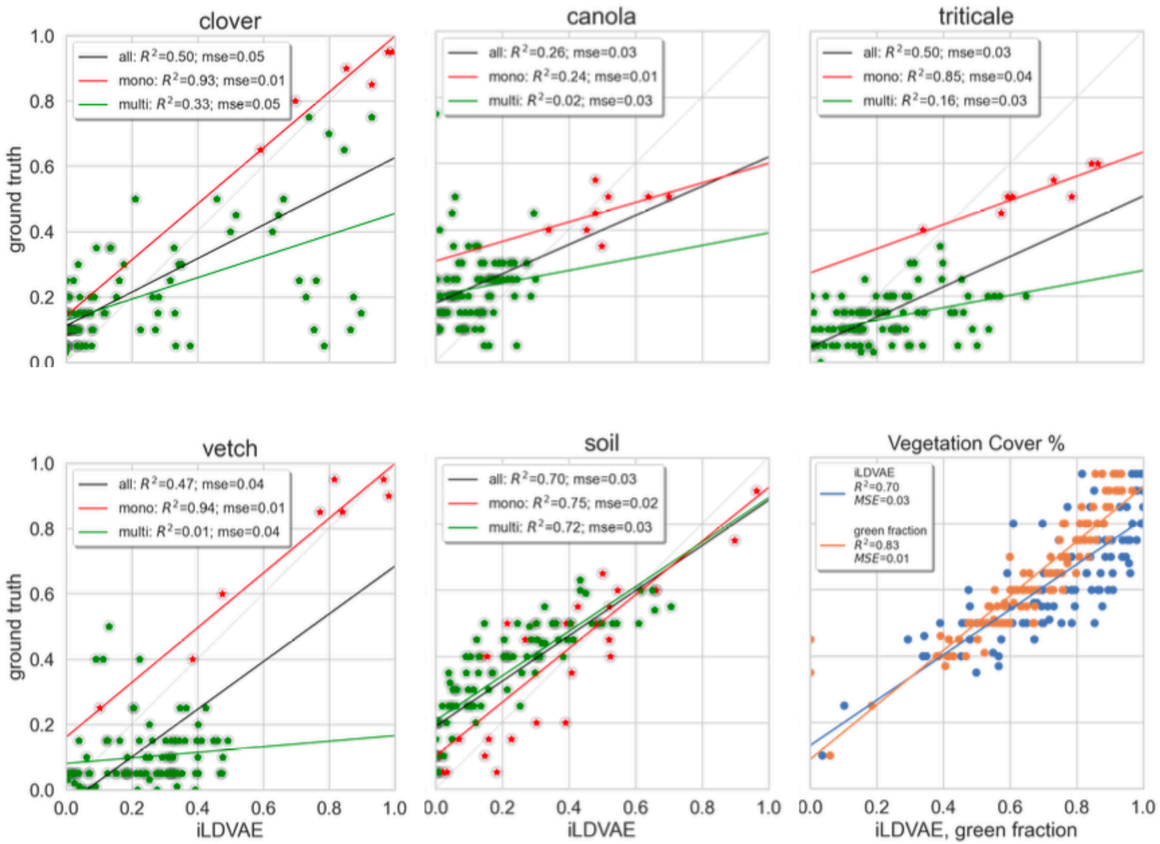


	SAD	RMSE
asphalt	0.0626 ±0.0029	0.1438 ±0.000 02
grass	0.1251 ±0.0010	0.1571 ±0.000 04
tree	0.1142 ±0.0016	0.2547 ±0.000 04
roof	0.0841 ±0.0017	0.1497 ±0.000 02
metal	0.5362 ±0.0047	0.3360 ±0.000 02
dirt	0.0793 ±0.0042	0.1490 ±0.000 01
average	0.1669 ±0.1824	0.1984 ±0.0795

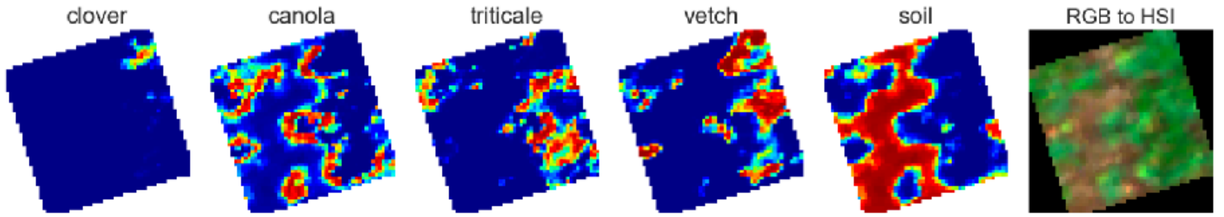
Abundances maps estimated by iLDVAE and Segmentation results (accuracy = 0.7238)

Iterative LDVAE

Results



Abundances of each species estimated by iLD- VAE (x-axis) vs. ground truth (y-axis); bottom-right: Percentage of vegetation cover: iLDVAE-estimated and Canopeo Green Fraction index [18, 19] vs. ground truth (y-axis). Each data point corresponds to one quadrat (total number of quadrats=120).



iLDVAE-estimated abundance maps and RGB image composite from HSI

Iterative LDVAE

Conclusions

Feasibility of self-supervised unmixing: iLDVAE successfully performs hyperspectral pixel unmixing without requiring labeled training data, achieving accurate results on both synthetic and benchmark datasets.

High purity pixels are crucial: The effectiveness of iLDVAE depends on the presence of pixels with high purity indices, as these serve as starting points for identifying endmembers.

Potential for real-world application: The application of iLDVAE on the Cover Crop USDA dataset demonstrates its potential for practical use in real-world scenarios, specifically in analyzing crop composition.

Competitive performance compared to existing methods: we highlight the results of iLDVAE, particularly its high correlation with ground truth in estimating vegetation cover compared to other methods like Canopeo Green Fraction.

Further research and comparison are needed: While iLDVAE shows promise, we acknowledge the need for further research, including a detailed comparison with other unmixing methods and exploration of techniques to address the limitation of requiring high purity pixels.

SpACNN-LDVAE

Combining Spatial and Spectral features into LDVAE for Hyperspectral Pixel Unmixing framework with Spatial Attention Convolutional

Chitnis, S. and Mantripragada K and Qureshi F.Z.; **SpACNN-LDVAE: Spatial Attention Convolutional Latent Dirichlet Variational Autoencoder for Hyperspectral Pixel Unmixing,** IGARSS 2024 - 2024 IEEE International Geoscience and Remote Sensing Symposium

SpACNN-LDVAE
Motivation

Limitations of existing methods: most well-known data-driven methods like BSS often don't consider spatial context.

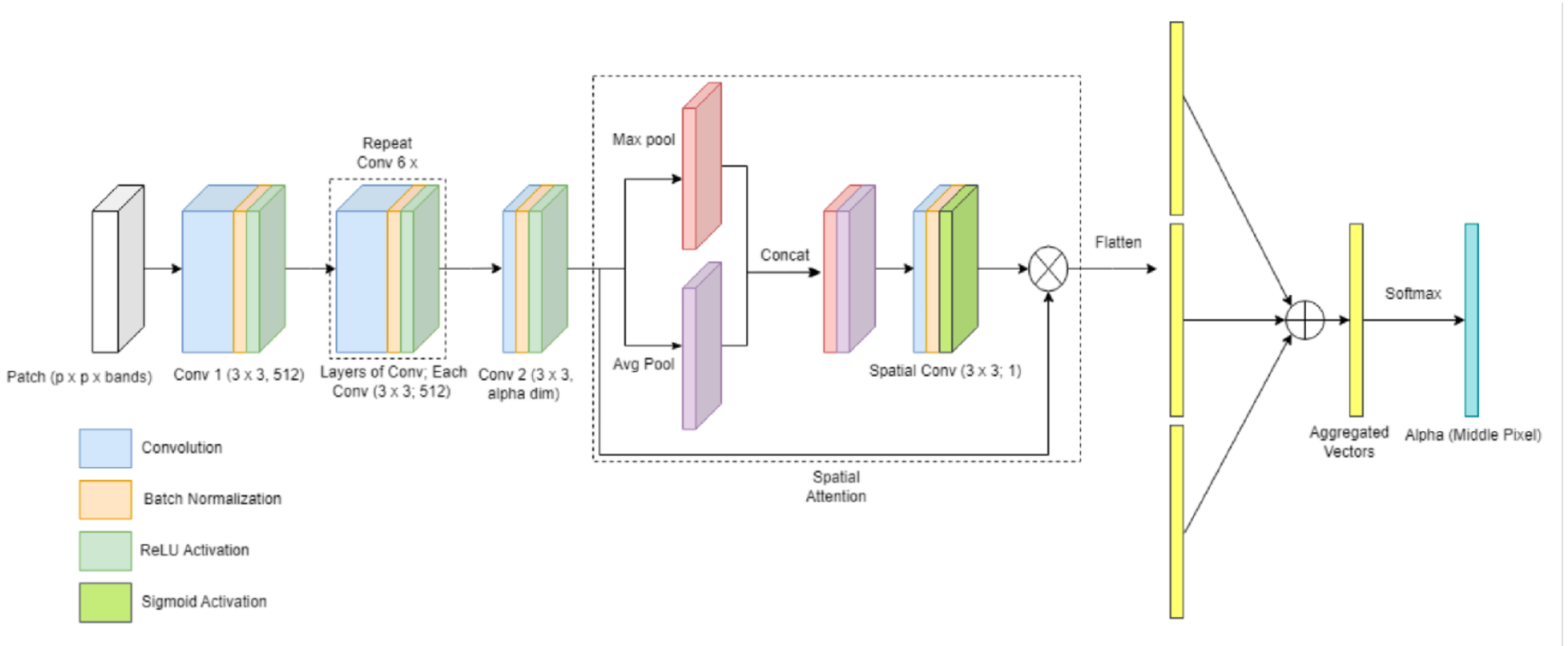
Exploiting spatial coherence: neighboring pixels often share similar endmembers and abundances. Incorporating this information can lead to better unmixing results.

Building upon LDVAE: The existing LDVAE approach successfully uses a probabilistic approach for unmixing but ignores spatial context.

Improving endmember extraction and abundance estimation: we investigate how the spatial context can improve results of unmixing, within the same LDVAE mathematical framework.

SpACNN-LDVAE

Method



Spatial Attention Convolutional Neural Network Encoder

SpACNN-LDVAE

Results

Table 1: Abundance Estimation and Endmember Extraction Results on Samson Dataset

		SpACNN-LDVAE	MLP-LDVAE [26]	VCA+FCLS [10]	PLMM [14]	ELMM [12, 13]	GLMM [39]	DeepGUn [24]	EACNN [22]
Soil	RMSE	0.2522 ± 0.00	0.2609 ± 0.00	-	-	-	-	-	-
	SAD	0.2097 ± 0.01	0.0959 ± 0.10	-	-	-	-	-	0.0328
Tree	RMSE	0.2614 ± 0.00	0.3431 ± 0.00	-	-	-	-	-	-
	SAD	0.5347 ± 0.03	1.2788 ± 1.28	-	-	-	-	-	0.0519
Water	RMSE	0.2098 ± 0.00	0.3165 ± 0.00	-	-	-	-	-	-
	SAD	0.8233 ± 0.04	0.4022 ± 0.40	-	-	-	-	-	0.1026
Average	RMSE	0.2412 ± 0.00	0.3078 ± 0.00	0.0545	0.0239	0.0119	0.0006	0.0862	0.0171
	SAD	0.5525 ± 0.03	0.5923 ± 0.59	-	-	-	-	-	0.0624

Table 2: Abundance Estimation and Endmember Extraction Results on HYDICE Urban Dataset

		SpACNN-LDVAE	MLP-LDVAE [26]	SSWNMF [17]	SGSNMF [18]	TV-RSNMF [19]	RSNMF [19]	GLNMF [20]	$L_{1/2}$NMF [21]	VCA+FCLS [10]
Asphalt road	RMSE	0.1566 ± 0.00	0.2889 ± 0.00	-	-	-	-	-	-	-
	SAD	0.2786 ± 0.02	0.4262 ± 0.43	0.0782 ± 3.29	0.0841 ± 4.01	0.0770 ± 2.97	0.0869 ± 3.81	0.1008 ± 3.19	0.0889 ± 2.88	0.2246 ± 3.44
Grass	RMSE	0.1977 ± 0.00	0.1832 ± 0.00	-	-	-	-	-	-	-
	SAD	0.1936 ± 0.01	0.3323 ± 0.33	0.1490 ± 3.58	0.1516 ± 3.25	0.1495 ± 3.54	0.1594 ± 3.62	0.1531 ± 3.06	0.1452 ± 3.57	0.1981 ± 3.39
Tree	RMSE	0.1632 ± 0.00	0.1737 ± 0.00	-	-	-	-	-	-	-
	SAD	0.4411 ± 0.04	0.3177 ± 0.32	0.1173 ± 3.46	0.1199 ± 3.36	0.1269 ± 4.02	0.1457 ± 4.29	0.1424 ± 3.79	0.1509 ± 3.18	0.2137 ± 2.41
Roof	RMSE	0.1283 ± 0.00	0.125 ± 0.00	-	-	-	-	-	-	-
	SAD	0.4502 ± 0.03	0.4393 ± 0.44	0.0713 ± 3.61	0.0731 ± 3.54	0.0746 ± 4.09	0.0849 ± 3.90	0.0986 ± 4.62	0.0863 ± 4.06	0.2673 ± 3.77
Metal	RMSE	0.0992 ± 0.00	0.2599 ± 0.00	-	-	-	-	-	-	-
	SAD	0.3241 ± 0.02	0.7004 ± 0.70	0.1241 ± 2.76	0.1250 ± 3.81	0.1247 ± 3.53	0.1324 ± 4.15	0.1370 ± 4.28	0.1334 ± 3.90	0.1848 ± 3.68
Dirt	RMSE	0.1894 ± 0.00	0.1334 ± 0.00	-	-	-	-	-	-	-
	SAD	0.2026 ± 0.01	0.2806 ± 0.28	0.0802 ± 3.17	0.0859 ± 3.91	0.0849 ± 3.92	0.0798 ± 3.77	0.1059 ± 3.96	0.1063 ± 3.54	0.1992 ± 3.43
Average	RMSE	0.1558 ± 0.00	0.1840 ± 0.00	0.0048 ± 0.72	0.0061 ± 0.67	0.0055 ± 0.81	0.0053 ± 0.98	0.0069 ± 0.85	0.0044 ± 0.76	0.0119 ± 0.66
	SAD	0.3151 ± 0.02	0.4161 ± 0.42	0.1034 ± 3.31	0.1060 ± 3.68	0.1063 ± 3.68	0.1148 ± 3.92	0.1230 ± 3.52	0.1185 ± 3.52	0.2142 ± 3.35

SpACNN-LDVAE
Conclusions

Spatial context matters: Incorporating spatial information within the LDVAE framework leads to better hyperspectral unmixing results.

Effectiveness of CNN with Spatial Attention: Spatial Attention mechanism successfully captures the local spatial structure of hyperspectral images, contributing to the improved unmixing performance.

Transfer learning has limitations: While the model shows promising results with transfer learning from synthetic to real-world data (Cuprite dataset), its performance in endmember extraction is comparable to MLP-LDVAE. This occurred because the training data lacks spatial coherence and it limited the ability of SpACNN-LDVAE to fully leverage its spatial awareness capabilities.

Lower standard deviations: The proposed method generally achieves lower standard deviations in its metrics compared to MLP-LDVAE, indicating more stable and consistent unmixing performance.

Papers and datasets

- Dao, P. D.; Mantripragada, K.; He, Y.; and Qureshi, F. Z.; **Improving hyperspectral image segmentation by applying inverse noise weighting and outlier removal for optimal scale selection**; ISPRS Journal of Photogrammetry and Remote Sensing, 171: 348 - 366, 2021.

- Mantripragada K, Dao PD, He Y, Qureshi FZ (2022);**The effects of spectral dimensionality reduction on hyperspectral pixel classification: A case study.** PLOS ONE 17(7): e0269174. https://doi.org/10.1371/journal.pone.0269174

- Mantripragada K. and Qureshi F. Z., **Hyperspectral Pixel Unmixing With Latent Dirichlet Variational Autoencoder**, in IEEE Transactions on Geoscience and Remote Sensing, vol. 62, pp. 1-12, 2024, Art no. 5507112, doi: 10.1109/TGRS.2024.3357589.

- Mantripragada, K. and Adler, P. R. and Olsen P. A. and Qureshi F. Z.; **An Iterative Method for Hyperspectral Pixel Unmixing Leveraging Latent Dirichlet Variational Autoencoder**, IGARSS 2023 - 2023 IEEE International Geoscience and Remote Sensing Symposium, Pasadena, CA, USA, 2023, pp. 7527-7530, doi: 10.1109/IGARSS52108.2023.10282175.

- Chitnis, S. and Mantripragada K and Qureshi F.Z.; **SpACNN-LDVAE: Spatial Attention Convolutional Latent Dirichlet Variational Autoencoder for Hyperspectral Pixel Unmixing,** IGARSS 2024 - 2024 IEEE International Geoscience and Remote Sensing Symposium

- Kiran Mantripragada, Faisal Z. Qureshi, **Evaluation of Dirichlet Process Gaussian Mixtures for Segmentation on Noisy Hyperspectral Images,** arXiv:2203.02820 [cs.CV] (tech report).

Datasets

http://vclab.science.uoit.ca/datasets/uoft-hsi301/uoft-hsi301.html

Q&A

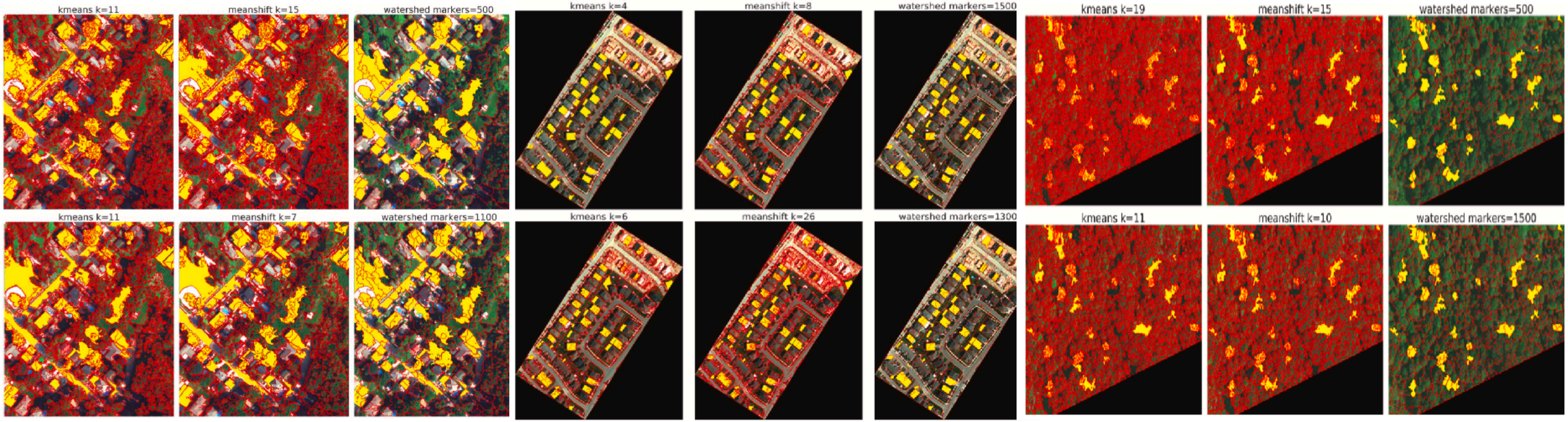
Kiran Mantripragada
kiran.mantripragada@ontariotechu.net

Faisal Z. Qureshi
faisal.qureshi@ontariotechu.ca

Appendix

Segmentation

Results



From top to bottom are the results of RoC and NN-nRoC methods, respectively. From left to right: Surburban, Urban, Forest datasets.

Segmentation

Conclusions

Impact of Noise and Outliers: The finer the segmentation, the more outliers are produced. Traditional approaches using RoC graphs without addressing noise and outliers lead to unreliable scale selection.

NN-nRoC Method's Effectiveness: Applying noise normalization and removing outlier segments in the calculation makes the NN-nRoC curves more robust.

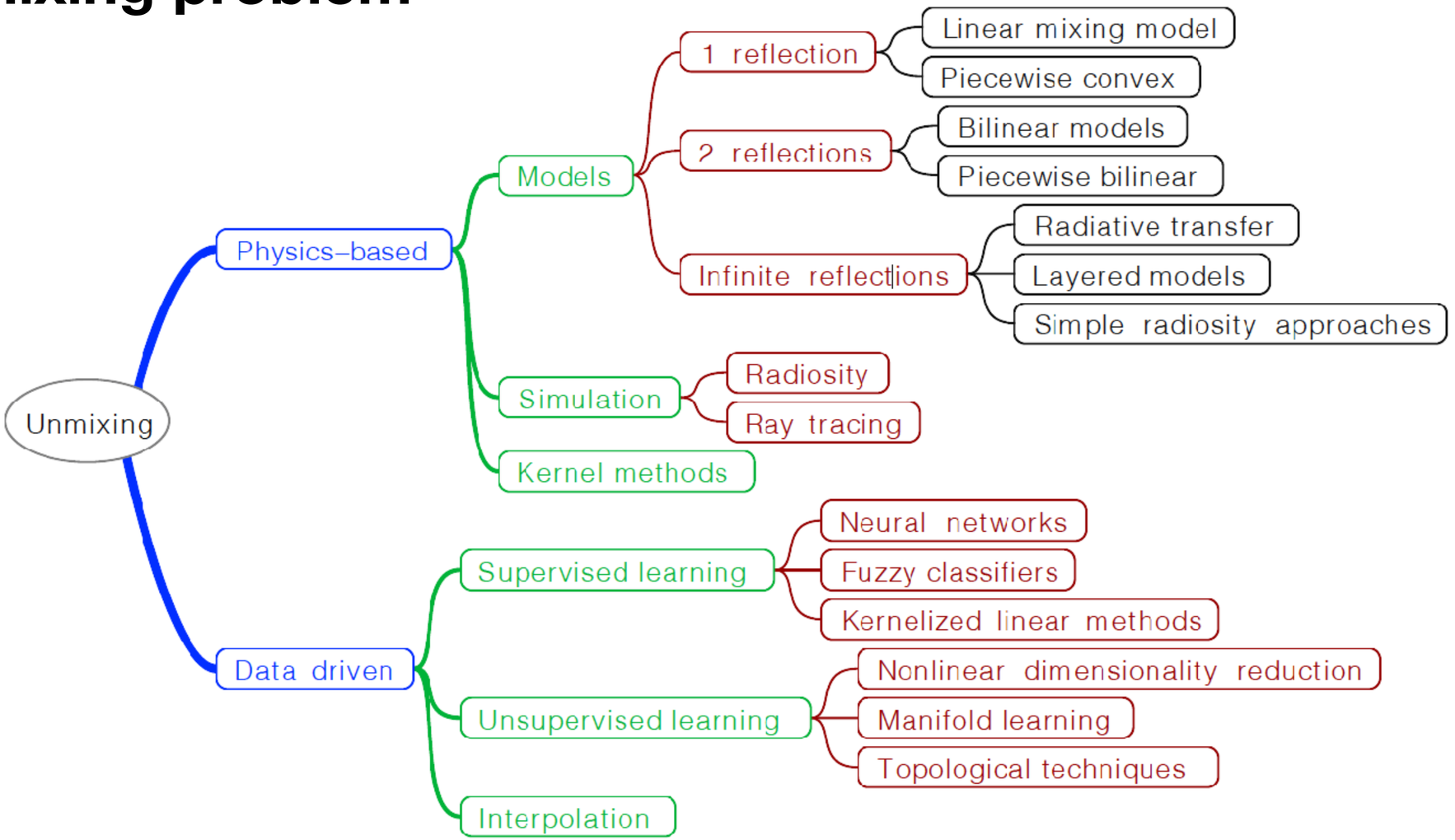
Improved Segmentation Results: Evaluation using OS, US, and ED metrics demonstrated that the NN-nRoC method produced better segmentation results for various segmentation algorithms.

Computational Efficiency: The NN-nRoC method offered a good balance between accuracy and computation efficiency, especially when compared to the ESP tool for the MRS algorithm.

Overall, the NN-nRoC method is a promising approach for hyperspectral image segmentation. It can be applied to other data types and has the potential to be automated for faster and more objective scale selection.

Hyperspectral Pixel Unmixing

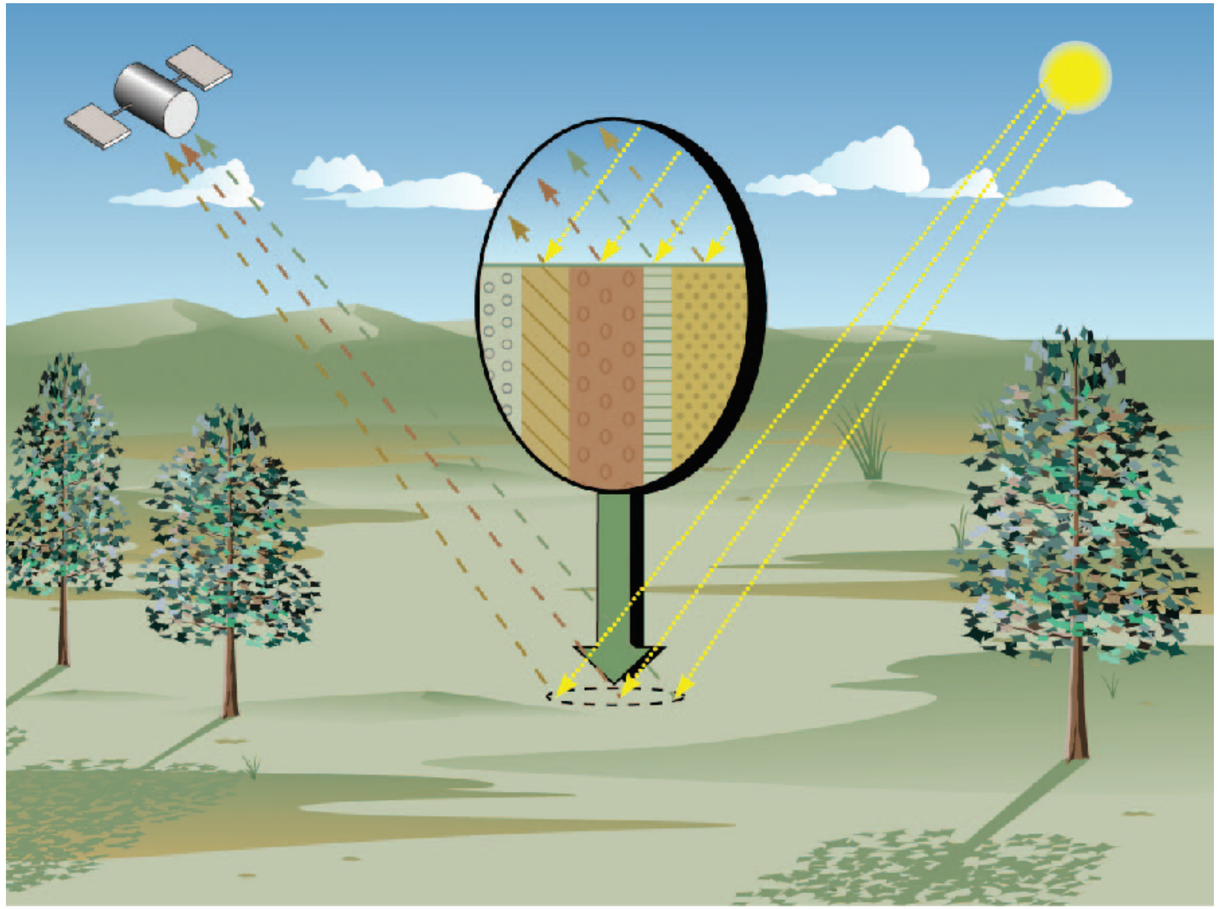
The unmixing problem

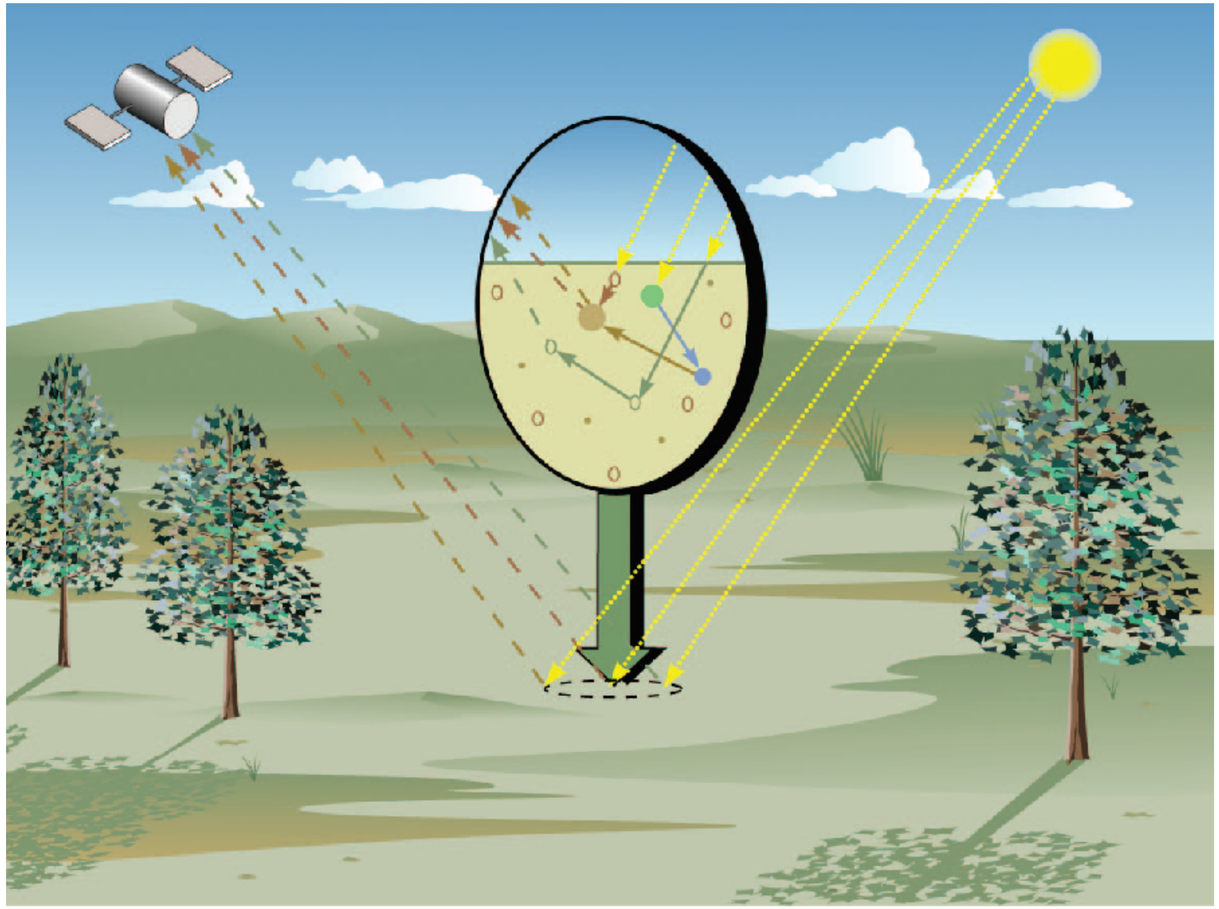


source: Heylen et al. (2014)

Hyperspectral Pixel Unmixing

Linear vs. Non-Linear





source: Keshava, 2003

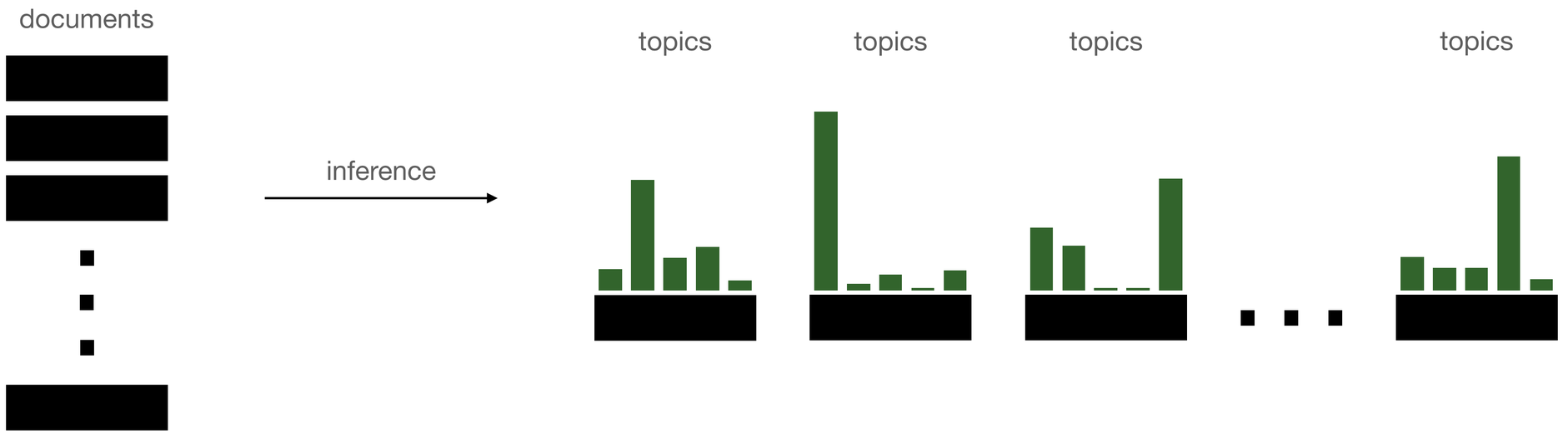
Hyperspectral Pixel Unmixing
Current methods

- **Physical models**
  - **Prior knowledge** of the constituents endmembers and surface
  - Models of **reflectance**, dispersion, absorption, scattering, refractance
  - [Hapke, 2012], [Heylen, 2014], [Sun and Lucey, 2021], [Janiczek et al., 2020], [Meganem et al., 2011], [Keshava, 2003]
- **Linear models**
  - Assume linear relationship between the observed and constituents spectra
  - [Khajehrayeni and Ghassemian, 2021], [Wei and Wang, 2020], [Keshava, 2003]
- **Abundance maps**
  - Given constituents endmembers, find mixing ratios (parameters of linear models)
  - Polinomial fitting through least squares (UCLS, NNLS, FCLS)
  - [Ibarrola et al., 2019]

Hyperspectral Pixel Unmixing

Key insight: Topic Analysis - Natural Language Processing

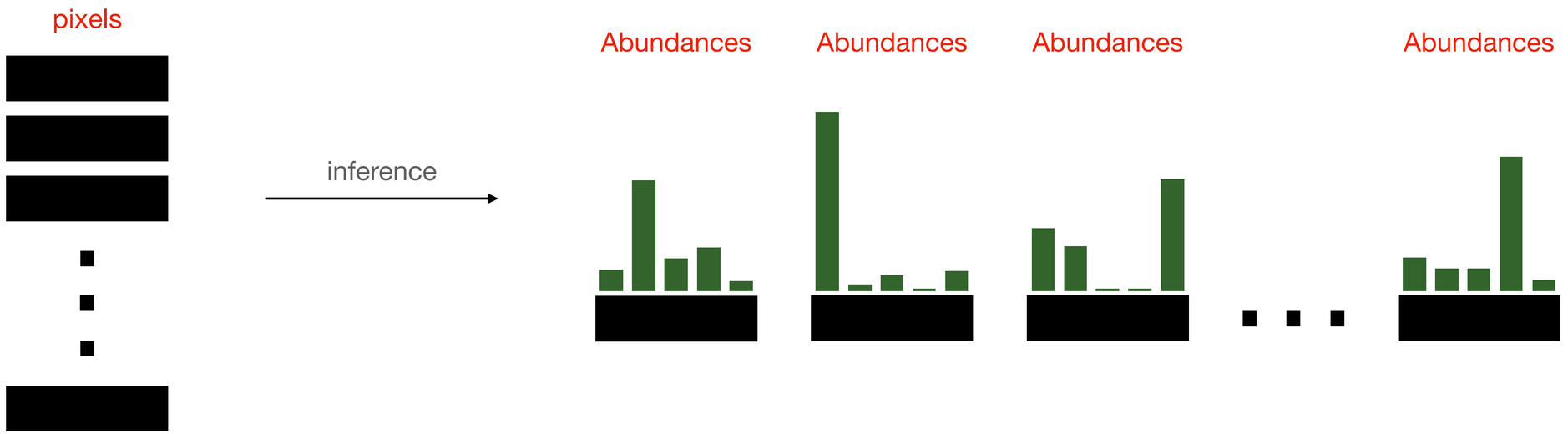
- Topic modelling framework: Latent Dirichlet Allocation (LDA) [Blei, 2003]
- The mixing ratio of topics is represented by a Dirichlet Distribution



Hyperspectral Pixel Unmixing

Key insight: Applying the concept of LDA to HSI Unmixing

- Reframe the unmixing problem within the Topic Modelling framework
- We model the abundances (mixing ratios) as Dirichlet distributions
- Learn Dirichlet distributions as latent space within a variational auto encoder setting



Hyperspectral Pixel Unmixing

ELBO (Evidence Lower Bound) Function

$$KL(q_\theta \| p_\phi) = \log p_\phi(\mathbf{x}) - \boxed{\mathcal{L}(\mathbf{x}; \theta, \phi)}$$

$$\boxed{\mathcal{L}(\mathbf{x}; \theta, \phi)} = \mathbb{E}_{q_\theta}\left[\log p_\phi(\mathbf{x}|\mathbf{z})\right] - KL(q_\theta(\mathbf{z}|\mathbf{x}) \| p(\mathbf{z}))$$

Hyperspectral Pixel Unmixing

Variational Autoencoder