GHENT UNIVERSITY

FACULTY OF ENGINEERING AND ARCHITECTURE

Sparse coding and deep learning in the analysis of hyperspectral images in remote sensing

Aleksandra Pižurica

Group for Artificial Intelligence and Sparse Coding (GAIM)
Department Telecommunications and Information Processing
Ghent University

MACLEAN 2020 – Machine Learning for Earth Observation
In Conjunction with ECML/PKDD
Ghent, Belgium, September 18 2020

Hyperspectral Imaging (HSI) in Earth observation

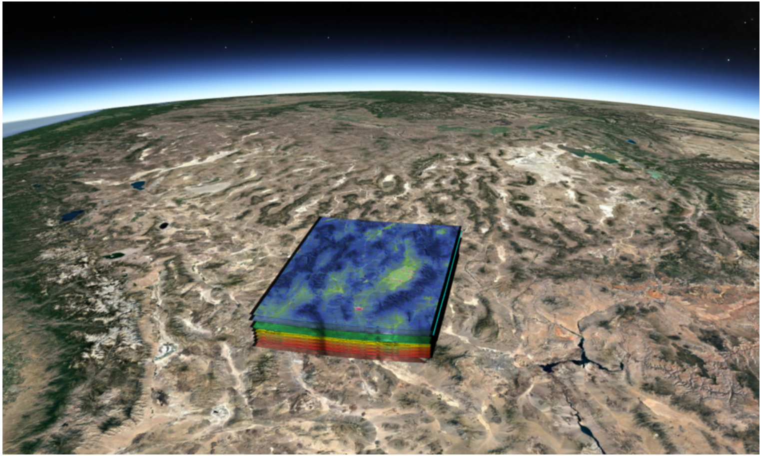
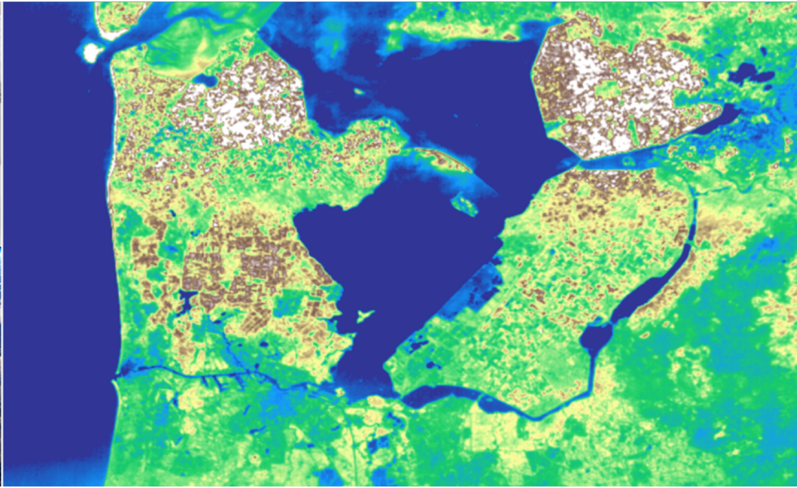


Image credit: Novus Light Technologies Today, December 2018.

HyperScout1 – the first miniaturized hyperspectral imager for space. Launched to an orbit 540km above the Earth. (ESA program, led by Cosine Measurement Systems)

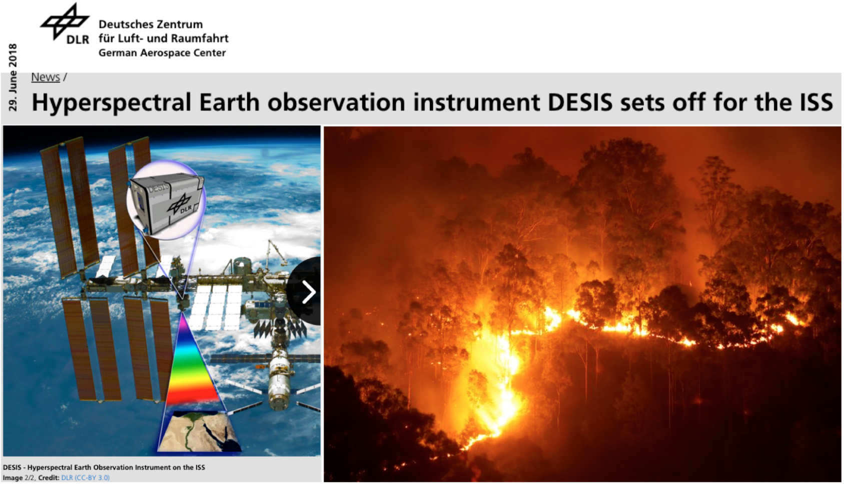
“Milk-carton-sized HyperScout making hyperspectral Earth views”

Space news feed, 20 May 2020



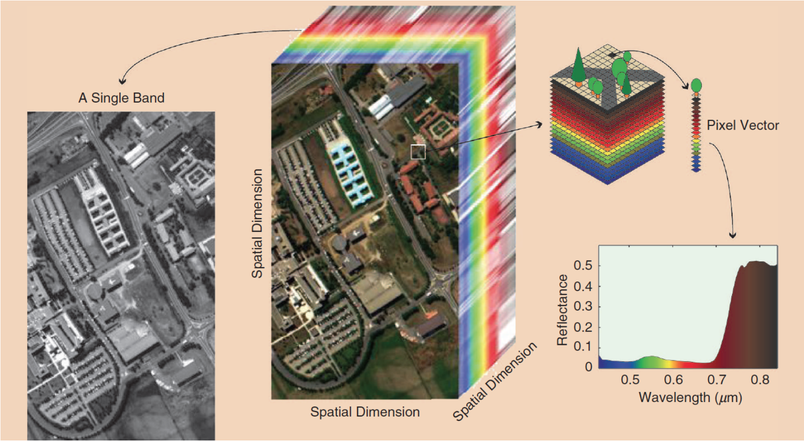
HyperScout view of Netherlands (courtesy: cosine)

HSI space technology - game changer in environmental monitoring



DLR Earth Sensing Imaging Spectrometer (DESIS) installed on the International Space Station (ISS). Monitors environmental changes on Earth.

HSI data cubes



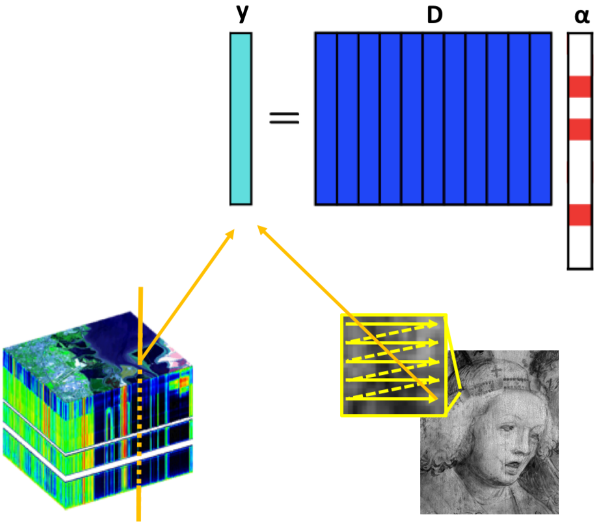
P. Ghamisi et al. Advances in Hyperspectral Image and Signal Processing. IEEE Geoscience and Remote Sensing Magazine, Dec 2017.

Outline

1. Sparse Coding of High-Dimensional Signals
   - Sparse representation
   - Sparse representation classification (SRC)
   - Sparse unmixing
   - Sparse subspace clustering (SSC)
2. Deep learning approaches
   - Deep learning in HSI classification
   - Some recent trends

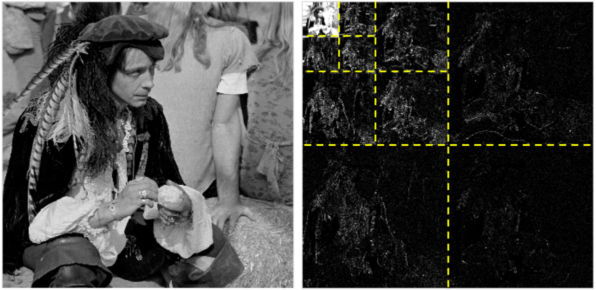
Outline

Sparse representation



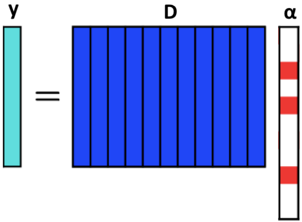
Designed vs. Learned Dictionaries

- **Designed dictionaries**: wavelets, curvelets, shearlets...
  - typically yield sparse representation of signals and images
  - advantages: generic, fast computation



- **Learned dictionaries**
  - trained on a set of representative examples
  - goal: optimally sparse representation for a given class of signals

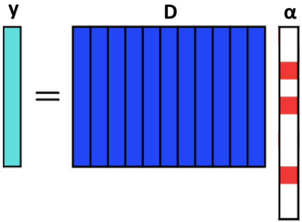
Sparse coding



$$\hat{\boldsymbol{\alpha}} = \arg\min_{\boldsymbol{\alpha}} \|\mathbf{y} - \mathbf{D}\boldsymbol{\alpha}\|_2^2 \quad \text{subject to} \quad \|\boldsymbol{\alpha}\|_0 \leq K$$

$$\hat{\boldsymbol{\alpha}} = \arg\min_{\boldsymbol{\alpha}} \|\boldsymbol{\alpha}\|_0 \quad \text{subject to} \quad \|\mathbf{y} - \mathbf{D}\boldsymbol{\alpha}\|_2^2 \leq \epsilon$$

Sparse coding



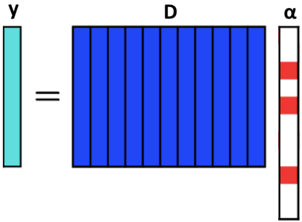
$$\hat{\boldsymbol{\alpha}} = \arg\min_{\boldsymbol{\alpha}} \|\mathbf{y} - \mathbf{D}\boldsymbol{\alpha}\|_2^2 \quad \text{subject to} \quad \|\boldsymbol{\alpha}\|_0 \leq K$$

$$\hat{\boldsymbol{\alpha}} = \arg\min_{\boldsymbol{\alpha}} \|\boldsymbol{\alpha}\|_0 \quad \text{subject to} \quad \|\mathbf{y} - \mathbf{D}\boldsymbol{\alpha}\|_2^2 \leq \epsilon$$

Greedy algorithms

- Matching Pursuit (MP) [Mallat and Zhang, '93]
- **OMP** [Tropp,'04], CoSaMP [Needell and Tropp, '09]
- IHT [Blumensath and Davies, 09]

Sparse coding



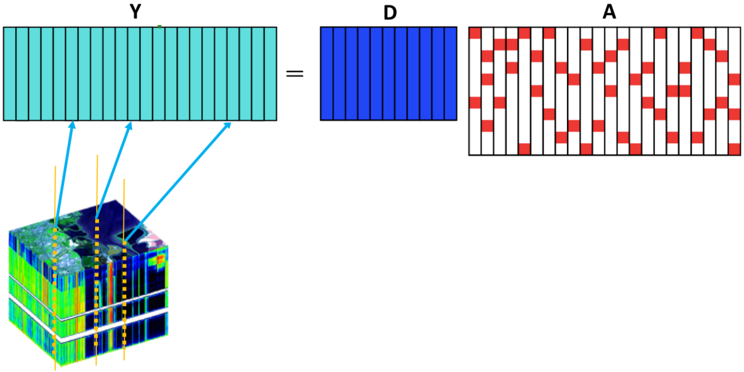
Convex relaxation:

$$\hat{\boldsymbol{\alpha}} = \arg\min_{\boldsymbol{\alpha}} \|\boldsymbol{\alpha}\|_1 \quad \text{subject to} \quad \|\mathbf{y} - \mathbf{D}\boldsymbol{\alpha}\|_2^2 \leq \epsilon$$

$$\hat{\boldsymbol{\alpha}} = \arg\min_{\boldsymbol{\alpha}} \|\mathbf{y} - \mathbf{D}\boldsymbol{\alpha}\|_2^2 + \lambda \|\boldsymbol{\alpha}\|_1$$

LASSO [Tibshirani, '96], BPDN [Chen et al, '01]

Sparse coding and dictionary learning

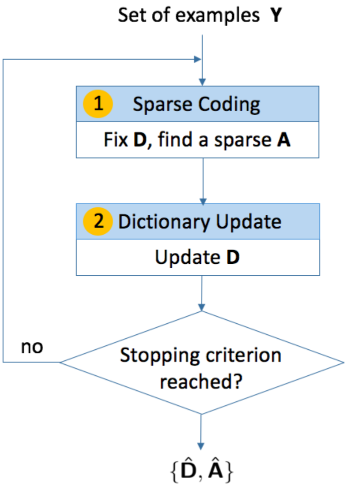


$$\{\hat{\mathbf{D}}, \hat{\mathbf{A}}\} = \arg\min_{\mathbf{D},\mathbf{A}} \left\{ \|\mathbf{Y} - \mathbf{D}\mathbf{A}\|_F^2 \right\} \quad \text{subject to} \quad \forall i, \|\boldsymbol{\alpha}_i\|_0 \leq K$$

A similar objective:

$$\{\hat{\mathbf{D}}, \hat{\mathbf{A}}\} = \arg\min_{\mathbf{D},\mathbf{A}} \sum_i \|\boldsymbol{\alpha}_i\|_0 \quad \text{subject to} \quad \|\mathbf{Y} - \mathbf{D}\mathbf{A}\|_F^2 \leq \epsilon$$

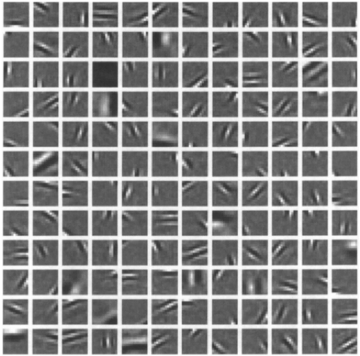
Iterate Two Steps: Sparse Coding and Dictionary Update

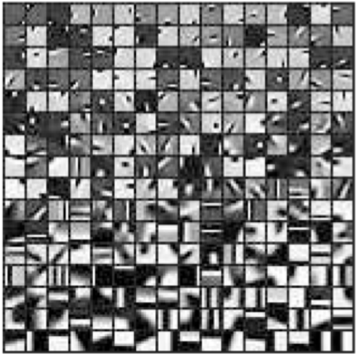


Dictionary update:

- Maximum likelihood method of [Olshausen and Field, 1997]
- MOD [Engan et al., 1999]
- K-SVD [Aharon et al., 2006]

Learned Dictionaries of Image Atoms - Examples



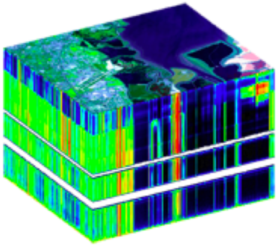


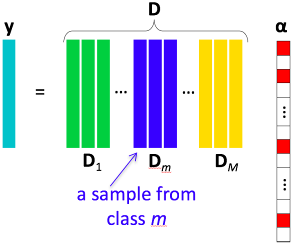
Examples of dictionaries trained by [Olshausen and Field, 1997] (left) and K-SVD [Aharon et al., 2006] (right)

Outline

Sparse Representation Classification

[Wright et al, 2009]





$$\hat{\boldsymbol{\alpha}} = \arg\min_{\boldsymbol{\alpha}} \|\mathbf{y} - \mathbf{D}\boldsymbol{\alpha}\|_2^2 \;\; \text{subject to } \|\boldsymbol{\alpha}\|_0 \leq K$$

$$r_m(\mathbf{y}) = \|\mathbf{y} - \mathbf{D}_m\hat{\boldsymbol{\alpha}}_m\|_2, \qquad m = 1, ..., M$$

$$class(\mathbf{y}) = \arg\min_{m=1,...,M} r_m(\mathbf{y})$$

Joint Sparsity Model

Collect pixels from a small neighbourhood $\mathcal{N}_\epsilon$ into $\mathbf{Y} = [\mathbf{y}_1, ..., \mathbf{y}_T] \in \mathbb{R}^{B \times T}$

$$\mathbf{Y} = \underbrace{[\mathbf{y}_1 \ \ldots \ \mathbf{y}_T]}_{\text{pixels from } \mathcal{N}_\epsilon} = [\mathbf{D}\boldsymbol{\alpha}_1 \ \ldots \ \mathbf{D}\boldsymbol{\alpha}_T] = \mathbf{D}\underbrace{[\boldsymbol{\alpha}_1 \ \ldots \ \boldsymbol{\alpha}_T]}_{\mathbf{A}} = \mathbf{DA}$$

Sparse codes $\{\boldsymbol{\alpha}_t\}_{t=1}^T$ share the same support $\implies$ **A** is sparse with only K non-zero rows, i.e., **A** is row sparse.

JSRC method [Chen et al., 2011]:

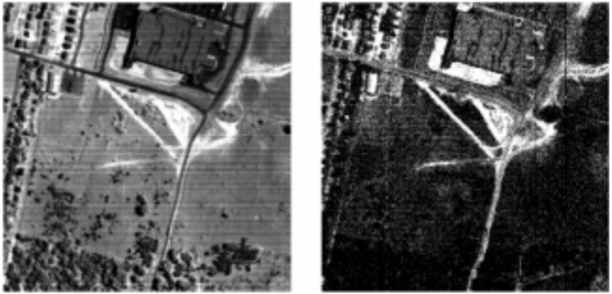
$$\hat{\mathbf{A}} = \arg\min_{\mathbf{A}} \|\mathbf{Y} - \mathbf{DA}\|_F^2 \ \text{ subject to } \|\mathbf{A}\|_{row,0} \le K$$

$$r_m(\mathbf{Y}) = \|\mathbf{Y} - \mathbf{D}_m\hat{\mathbf{A}}_m\|_F, \qquad m = 1, ..., M$$

$$class(\mathbf{y}_{central}) = \arg\min_{m=1,...,M} r_m(\mathbf{Y})$$

Robust SRC for Hyperspectral Image Classification

$$\mathbf{Y} = \underbrace{\mathbf{X}}_{\text{ideal image}} + \underbrace{\mathbf{N}}_{\text{Gaussian noise}} + \underbrace{\color{red}{\mathbf{S}}}_{\color{red}{\text{sparse noise}}}$$



Examples of stripe noise and mixed noise in a real hyperspectral image.

Huang, S., Zhang, H., Liao, W., and Pižurica, A. (2017). Robust joint sparsity model for hyperspectral image classification. In IEEE ICIP 2017.

Robust SRC for Hyperspectral Image Classification

$$\mathbf{Y} = \underbrace{\mathbf{X}}_{\text{ideal image}} + \underbrace{\mathbf{N}}_{\text{Gaussian noise}} + \underbrace{\mathbf{S}}_{\text{sparse noise}}$$

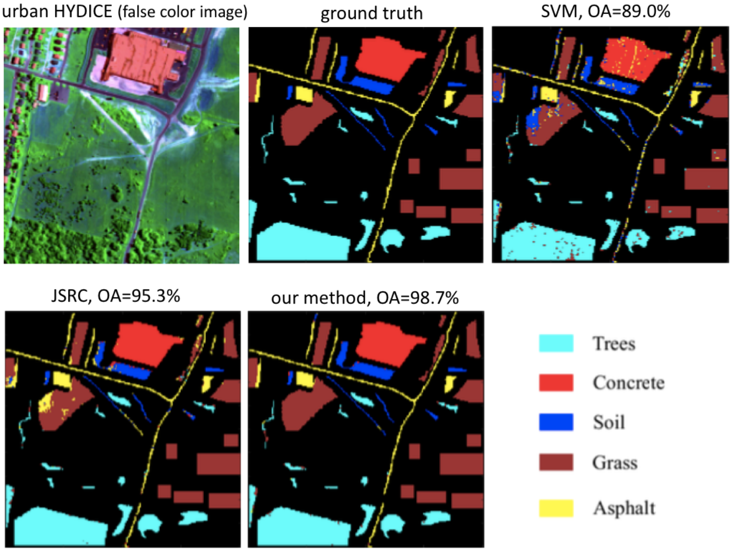
$$\{\hat{\mathbf{A}}, \hat{\mathbf{S}}\} = \arg\min_{\mathbf{A},\mathbf{S}} \|\mathbf{Y} - \mathbf{DA} - \mathbf{S}\|_F^2 + \lambda\|\mathbf{S}\|_1 \quad \text{subject to} \quad \|\mathbf{A}\|_{row,0} \leq K$$

$$r_m(\mathbf{Y}) = \|\mathbf{Y} - \mathbf{D}_m\hat{\mathbf{A}}_m - \hat{\mathbf{S}}\|_F, \qquad m = 1, \ldots, M$$

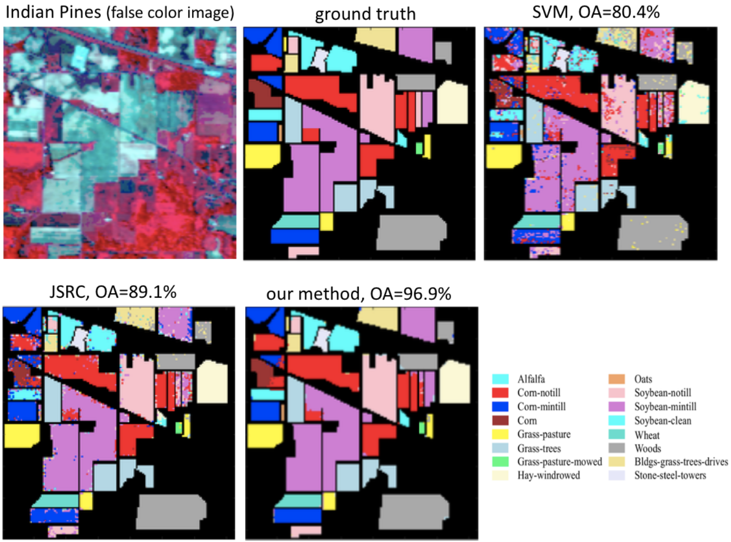
$$class(\mathbf{y}_{central}) = \arg\min_{m=1,\ldots,M} r_m(\mathbf{Y})$$

S. Huang, H. Zhang and A. Pižurica (2017). A Robust Sparse Representation Model for Hyperspectral Image Classification. Sensors.

Robust SRC for Hyperspectral Image Classification



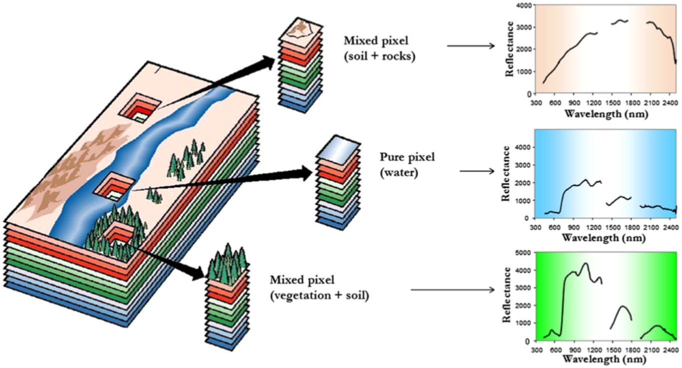
Robust SRC for Hyperspectral Image Classification



Outline

1. Sparse Coding of High-Dimensional Signals
   - Sparse representation
   - Sparse representation classification (SRC)
   - Sparse unmixing
   - Sparse subspace clustering (SSC)
2. Deep learning approaches
   - Deep learning in HSI classification
   - Some recent trends

Spectral Mixing



S.R Bijitha, P. Geetha and K.P. Soman (2016). Performance Analysis and Comparative Study of Geometrical Approaches for Spectral Unmixing. International Journal of Scientific and Engineering Research.

Sparse Unmixing

Ideal hyperspectral image reordered as a matrix $\mathbf{X} \in \mathbb{R}^{B \times MN}$

Linear mixing model:

$$\mathbf{X} = \mathbf{E}\mathbf{A}$$

$\mathbf{E} \in \mathbb{R}^{B \times K}$ – library of **endmembers**; $\mathbf{A} \in \mathbb{R}^{K \times MN}$ – **abundance**

$$\mathbf{Y} = \underbrace{\mathbf{E}}_{\text{library}} \underbrace{\mathbf{A}}_{\text{abundance}} + \underbrace{\mathbf{N}}_{\text{Gaussian noise}} + \underbrace{\mathbf{S}}_{\text{sparse noise}}$$

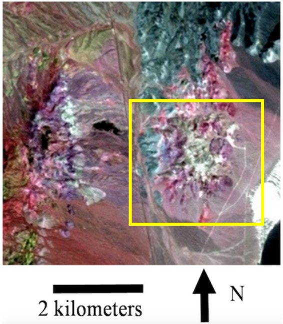
The approach of [Aggarval et al, 2016]:

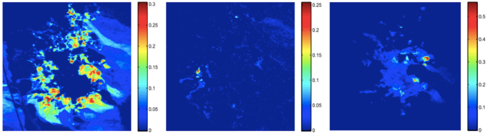
$$\min_{\mathbf{A},\mathbf{S}} \|\mathbf{Y} - \mathbf{E}\mathbf{A} - \mathbf{S}\|_F^2 + \lambda_1 \|\mathbf{A}\|_{2,1} + \lambda_2 \|\mathbf{S}\|_1$$

Many similar variants exist, also making use of low-rank assumption:

$$\min_{\mathbf{A}} \text{rank}\{\mathbf{A}\} \quad \text{subject to} \quad \|\mathbf{Y} - \mathbf{E}\mathbf{A} - \mathbf{S}\|_F^2 \leq \epsilon$$

Sparse Unmixing





Estimated fractional abundance maps (AVIRIS Cuprite subscene, USGS library).

R. Wang, H.-C. Li, A. Pižurica, J. Li, A. Plaza, and W. J. Emery (2017). Hyperspectral Unmixing Using Double Reweighted Sparse Regression and Total Variation. IEEE Geoscience and Remote Sensing Letters.

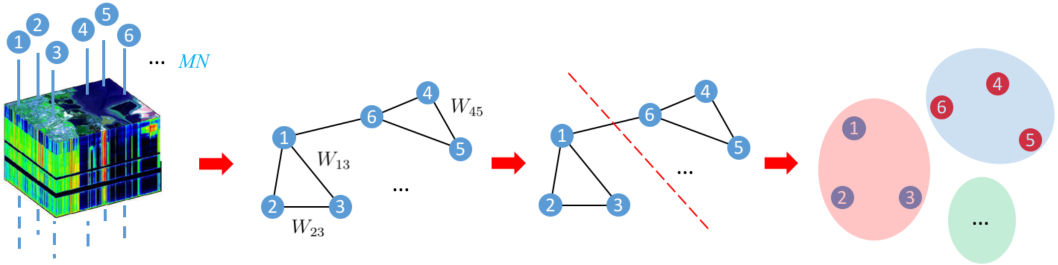
Outline

1. Sparse Coding of High-Dimensional Signals
   - Sparse representation
   - Sparse representation classification (SRC)
   - Sparse unmixing
   - Sparse subspace clustering (SSC)

2. Deep learning approaches
   - Deep learning in HSI classification
   - Some recent trends

Spectral clustering

No labelled data available → no supervised classification but instead clustering

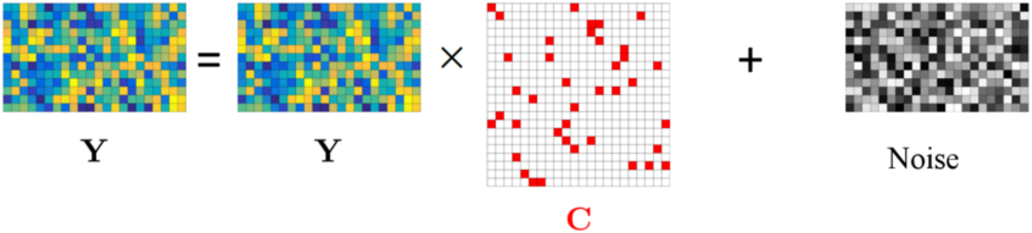


Similarity matrix: $\mathbf{W} \in \mathbb{R}^{MN \times MN}$

Sparse Subspace Clustering

[Elhamifar and Vidal, 2013]

Self-representation model: $\mathbf{Y} = \mathbf{Y}\mathbf{C} + \mathbf{N}$; $\quad \mathbf{Y} = [\mathbf{y}_1 ... \mathbf{y}_N] \in \mathbb{R}^{m \times N}$



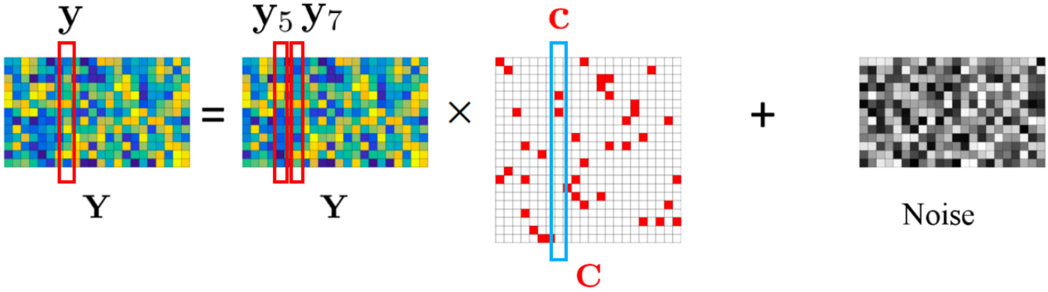
$C_{i,j} \neq 0 \rightarrow \mathbf{y}_i$ and $\mathbf{y}_j$ are in the same subspace.

Similarity matrix: $W = |C| + |C|^T$

Sparse Subspace Clustering

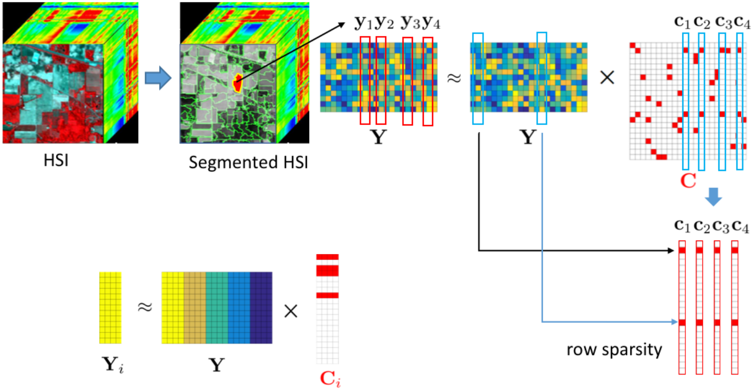
[Elhamifar and Vidal, 2013]

Self-representation model: $\mathbf{Y} = \mathbf{Y}\mathbf{C} + \mathbf{N}$; $\mathbf{Y} = [\mathbf{y}_1 ... \mathbf{y}_N] \in \mathbb{R}^{m \times N}$



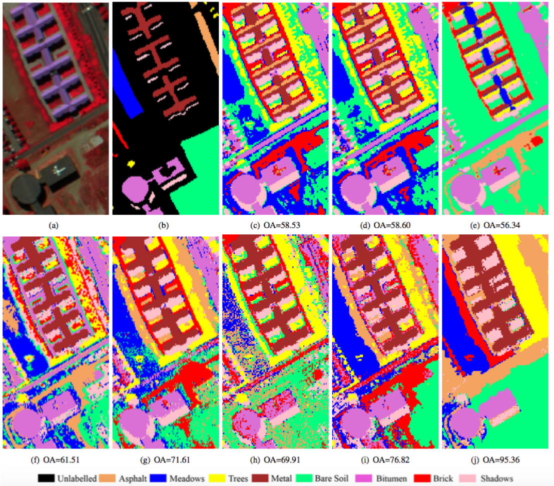
$\mathbf{y} \approx \mathbf{Y}\mathbf{c} = \sum_i \mathbf{y}_i c_i = c_5 \mathbf{y}_5 + c_7 \mathbf{y}_7$

Joint Sparse Subspace Clustering - JSSC



S. Huang, H. Zhang and A. Pižurica (2018). Semi-supervised Sparse Subspace Clustering Method With a Joint Sparsity Constraint for Hyperspectral Remote Sensing Images. IEEE J. Sel. Topics in Earth Observation and Remote Sens.

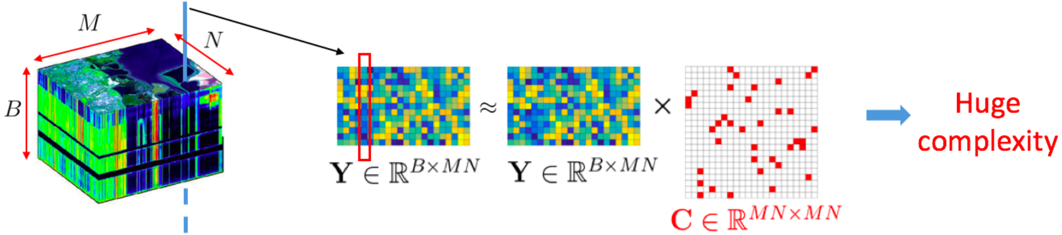
Joint Sparse Subspace Clustering - JSSC



Pavia University image. (a) False color image, (b) Ground truth (c) FCM, (d) k-means, (e) CFSFDP, (f) SSC, (g) L2-SSC, (h) CPPSSC (1 % labelled samples), (i) JSSC and (j) JSSC-L (1% labelled samples)

S. Huang, H. Zhang, A. Pižurica (2018). IEEE J. Sel. Topics in Earth Observ. Remote Sens.

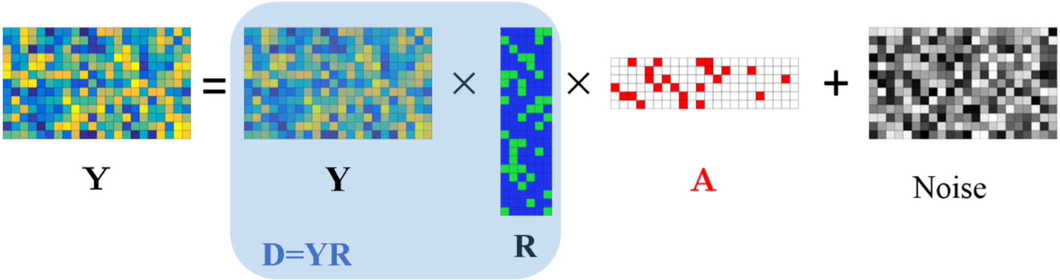
Nice, but ...



SSC becomes practically infeasible for very large scale data.

E.g. for the full *Pavia University* image 610 $\times$ 340, the size of **C** is 207400 $\times$ 207400 $\rightarrow$ 320,5 GB memory

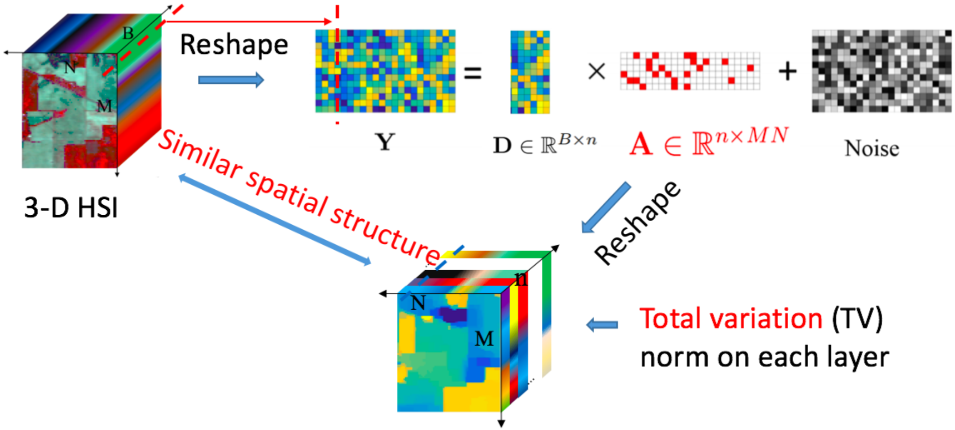
Sketching



Reduces greatly the size of the problem!

P. A. Traganitis and G. B. Giannakis (2018). Sketched subspace clustering. IEEE Trans. Signal Process.

Sketched Sparse Subspace Clustering for Hyperspectral Images

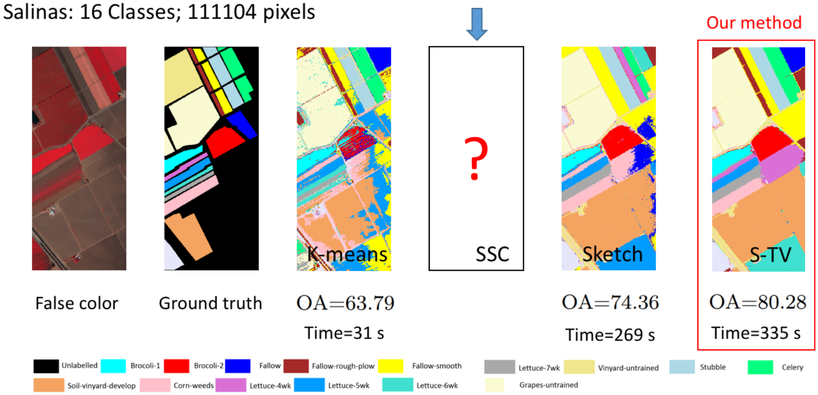


S. Huang, H. Zhang and A. Pižurica (2020).

Sketch-based Subspace Clustering of Hyperspectral Images. Remote Sensing.

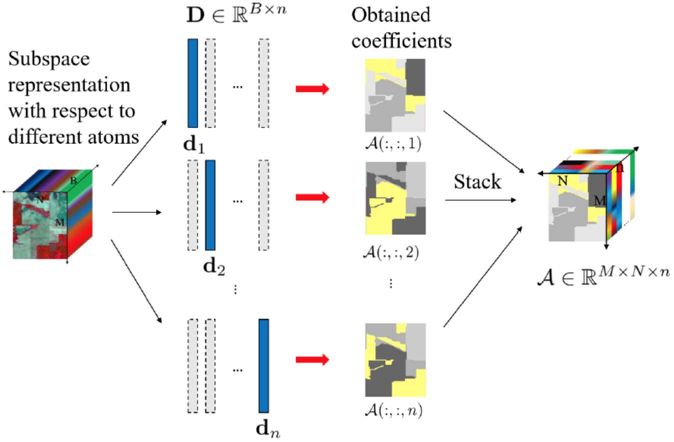
Sketched Sparse Subspace Clustering for Hyperspectral Images

$\mathbf{Y} \in \mathbb{R}^{204 \times 111104} \Longrightarrow \mathbf{C} \in \mathbb{R}^{111104 \times 111104}$



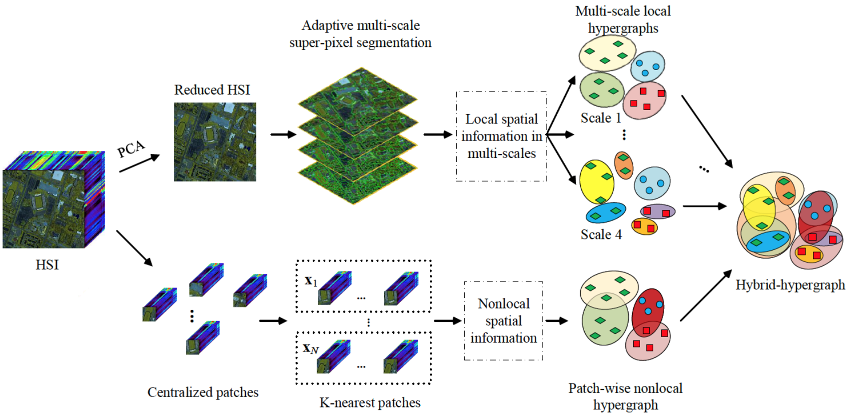
S. Huang, H. Zhang and A. Pižurica (2020).
Sketch-based Subspace Clustering of Hyperspectral Images. Remote Sensing.

Large-scale Sparse Subspace Clustering with Dictionary Learning



S. Huang, H. Zhang and A. Pižurica (2020). Subspace Clustering for Hyperspectral Images via Dictionary Learning with Adaptive Regularization (in review; IEEE Trans. Geosc. Remote Sens)

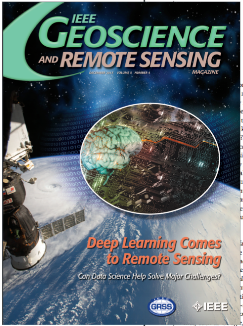
Multi-view subspace clustering



S. Huang, H. Zhang and A. Pižurica (2020). Hybrid-hypergraph Regularized Multi-view Subspace Clustering for Hyperspectral Images (in review; IEEE Trans. Geosc. Remote Sens)

Outline

Deep learning approaches



XIAO XIANG ZHU, DEVIS TUIA, LICHAO MOU, GUI-SONG XIA, LIANGPEI ZHANG, FENG XU, AND FRIEDRICH FRAUNDORFER

Deep Learning in Remote Sensing

A comprehensive review and list of resources

oming paradigm shift toward data-in-
machine-learning techniques are be-
important. In particular, deep learn-
e both a major breakthrough and an
tool in many fields. Shall we embrace
key to everything? Or should we resist
? These are controversial issues within
community. In this article, we analyze
ing deep learning for remote-sensing
recent advances, and provide resourc-
ake deep learning in remote sensing
mple. More importantly, we encourage
ntists to bring their expertise into deep
as an implicit general model to tackle
e-scale, influential challenges, such as
urbanization.

astest-growing trend in big data analysis
e of the ten breakthrough technologies
acterized by neural networks (NNs) in-
e than two hidden layers (for this rea-
deep). Like shallow NNs, deep NNs ex-
ntations learned exclusively from data,
ting features that are designed based
-specific knowledge. Deep learning
xtensively pushed by Internet compa-
e, Baidu, Microsoft, and Facebook, for several image analysis tasks, including image indexing, segmentation, and object detection.

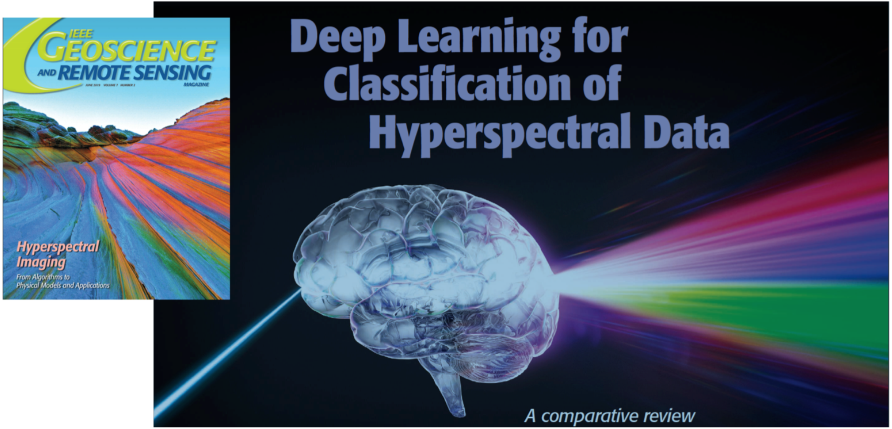
Based on recent advances, deep learning is proving to be a very successful set of tools, sometimes able to surpass even humans in solving highly computational tasks (consider, e.g., the widely reported Go match between Google's AlphaGo artificial intelligence program and the world Go champion Lee Sedol). Based on such exciting successes, deep learning is increasingly the model of choice in many application fields.

For instance, convolutional NNs (CNNs) have proven to be good at extracting mid- and high-level abstract features from raw images by interleaving convolutional and pooling layers (i.e., by spatially shrinking the feature maps layer by layer). Recent studies indicate that the feature representations learned by CNNs are highly effective in large-scale image recognition [2]–[4], object detection [5], [6], and semantic segmentation [7], [8]. Furthermore, recurrent NNs (RNNs), an important branch of the deep learning family, have demonstrated significant achievement on a variety of tasks involved in sequential data analysis, such as action recognition [9], [10] and image captioning [11].

In the wake of this success and thanks to the increased availability of data and computational resources, the use of deep learning is finally taking off in remote sensing as well. Remote-sensing data present some new challenges for deep learning, because satellite image analysis raises unique issues that pose difficult new scientific questions.

Digital Object Identifier 10.1109/MGRS.2017.2762307
Date of publication: 27 December 2017

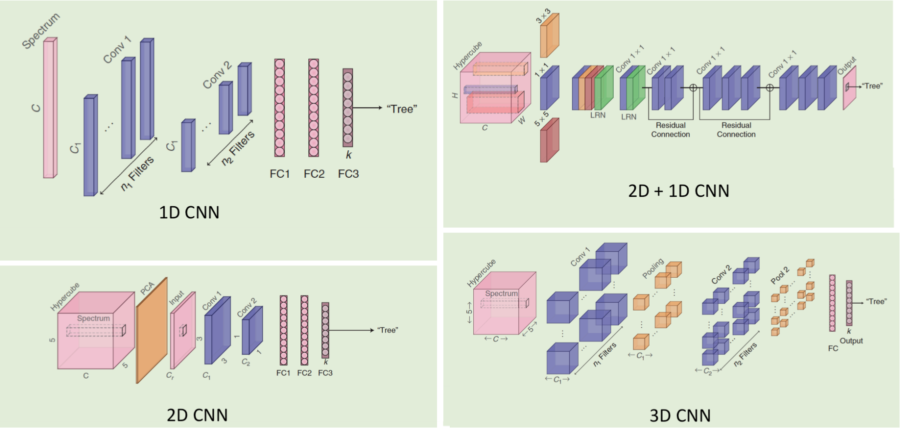
State-of-the-art in deep learning for HSI classification



NICOLAS AUDEBERT, BERTRAND LE SAUX, AND SÉBASTIEN LEFÈVRE

N. Audebert, B. Le Saux, and S. Lefèvre, IEEE Geosc. Remote Sens. Mag., June 2019.

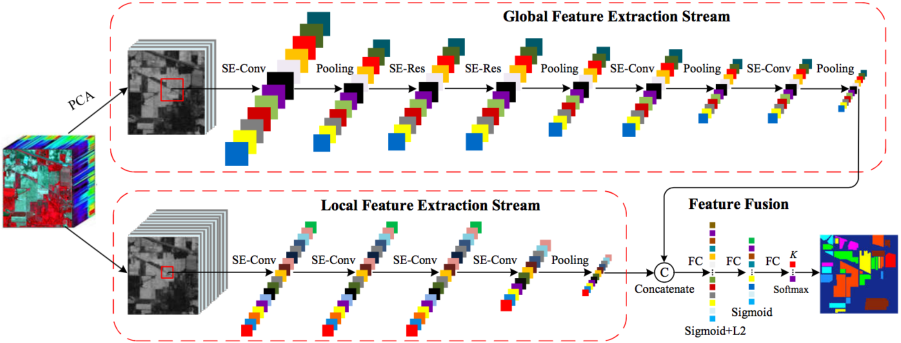
Deep learning in HSI classification



N. Audebert, B. Le Saux, and S. Lefèvre. Deep Learning for Classification of Hyperspectral Data - A comparative Review. IEEE Geosc. Remote Sens. Mag., June 2019.

Outline

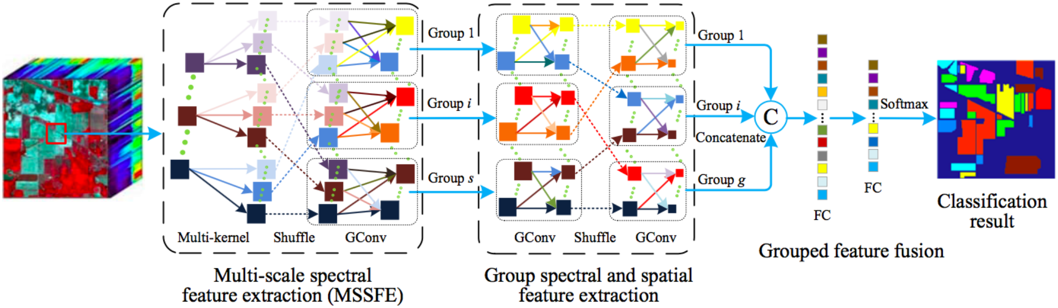
Spectral-spatial feature fusion with two-stream CNN



Improving the performance in the case of limited labelled data.

X. Li, M. Ding and A. Pižurica. Deep Feature Fusion via Two-Stream Convolutional Neural Network for Hyperspectral Image Classification, IEEE Transactions on Geoscience and Remote Sensing, in press (2019).

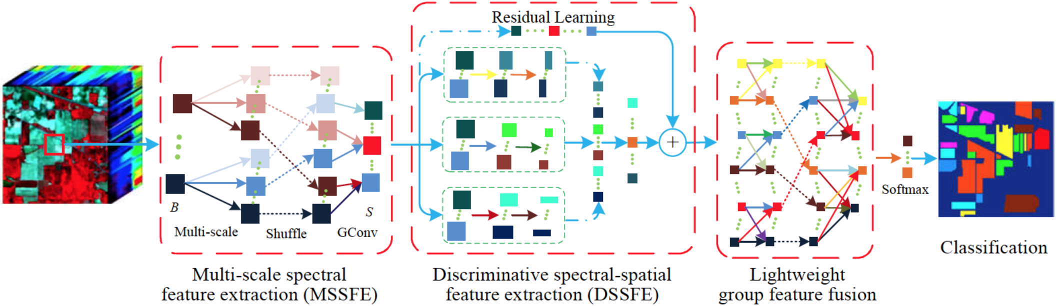
Group CNN for HSI classification



Reducing the computational complexity - applicability to large scale data.

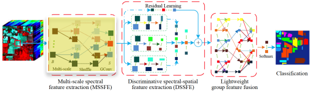
X. Li, M. Ding and A. Pižurica. Group Convolutional Neural Networks for Hyperspectral Image Classification, ICIP 2018.

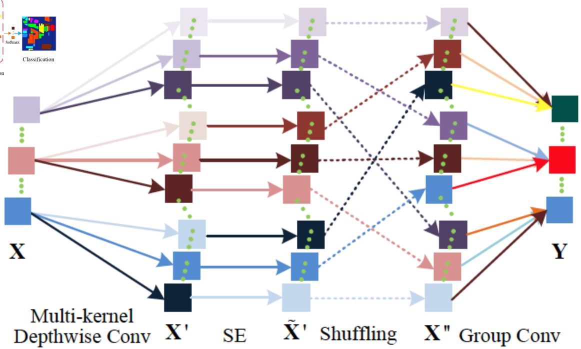
Full Group CNN (FGCNN)



X. Li, M. Ding and A. Pižurica. Full Group Convolutional Neural Networks for Robust Spectral-Spatial Feature Learning (2020). IEEE Trans. Image Process. (in review)

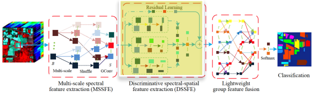
Full Group CNN (FGCNN)

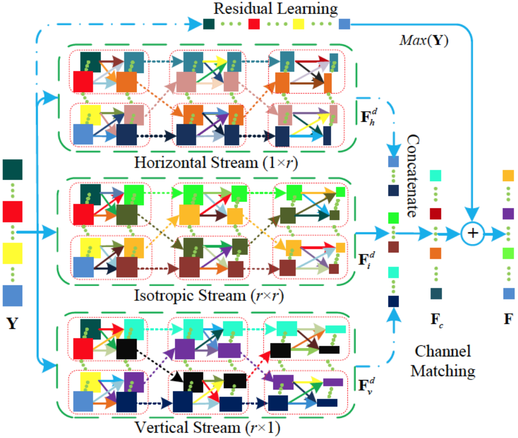




X. Li, M. Ding and A. Pižurica. Full Group Convolutional Neural Networks for Robust Spectral-Spatial Feature Learning (2020). IEEE Trans. Image Process. (in review)

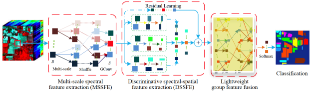
Full Group CNN (FGCNN)

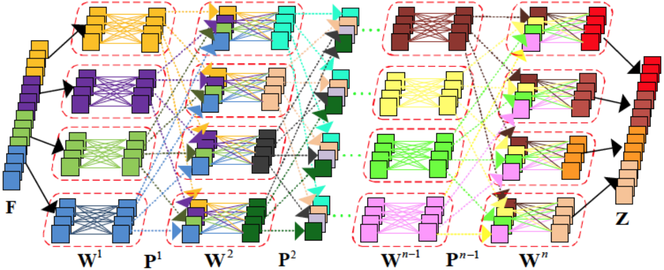




X. Li, M. Ding and A. Pižurica. Full Group Convolutional Neural Networks for Robust Spectral-Spatial Feature Learning (2020). IEEE Trans. Image Process. (in review)

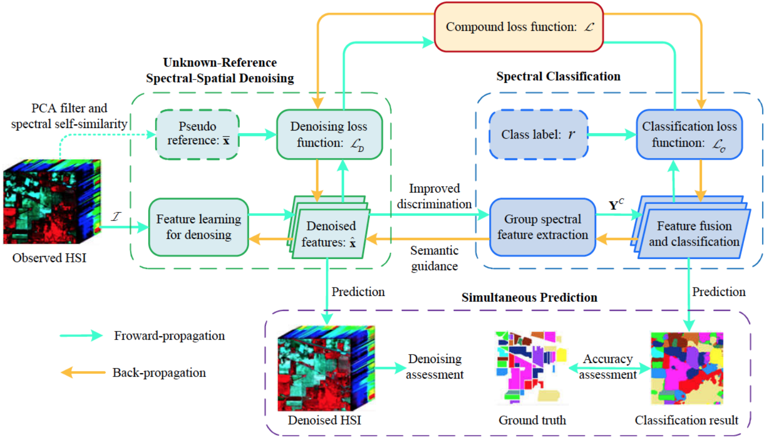
Full Group CNN (FGCNN)





X. Li, M. Ding and A. Pižurica. Full Group Convolutional Neural Networks for Robust Spectral-Spatial Feature Learning (2020). IEEE Trans. Image Process. (in review)

Joint Denoising and Classification



X. Li, M. Ding and A. Pižurica (2020). Integrated Networks for Simultaneous Denoising and Classification of Hyperspectral Images (in review, IEEE Trans. Neural Net. Learn. Syst.)

Special Issues

remote sensing

Submit to Special Issue

Submit Abstract to Special Issue

Review for *Remote Sensing*

Edit a Special Issue

Journal Menu

- *Remote Sensing* Home
- Aims & Scope
- Editorial Board

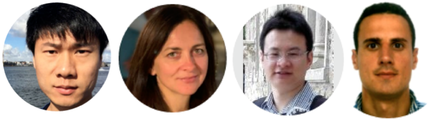
Special Issue "Spectral Unmixing of Hyperspectral Remote Sensing Imagery"

- **Special Issue Editors**
- **Special Issue Information**
- **Keywords**
- **Published Papers**

A special issue of ***Remote Sensing*** (ISSN 2072-4292). This special issue belongs to the section "**Remote Sensing Image Processing**".

Deadline for manuscript submissions: **20 March 2021**.

Guest Editors:
Shaoguang Huang, Aleksandra Pizurica, Hongyan Zhang and Mauro Dalla Mura



Special Issues

remote sensing

Submit to Special Issue

Submit Abstract to Special Issue

Review for *Remote Sensing*

Edit a Special Issue

Journal Menu

- *Remote Sensing* Home
- Aims & Scope
- Editorial Board
- Reviewer Board
- Topics Board

Special Issue "Remote Sensing Image Denoising, Restoration and Reconstruction"

- Special Issue Editors
- Special Issue Information
- Keywords
- Published Papers

A special issue of *Remote Sensing* (ISSN 2072-4292). This special issue belongs to the section "Remote Sensing Image Processing".

Deadline for manuscript submissions: **31 May 2021**.

Guest Editors:
Karen Egiazarian, Aleksandra Pizurica and Vladimir Lukin