CLASSIFICATION AND SEGMENTATION OF HYPERSPECTRAL IMAGES

Kemker, R., Gewali, U. B., Kanan, C. EarthMapper: A Tool Box for the Semantic Segmentation of Remote Sensing Imagery .IEEE Geoscience and Remote Sensing Letters (GRSL).

Reviewed by:

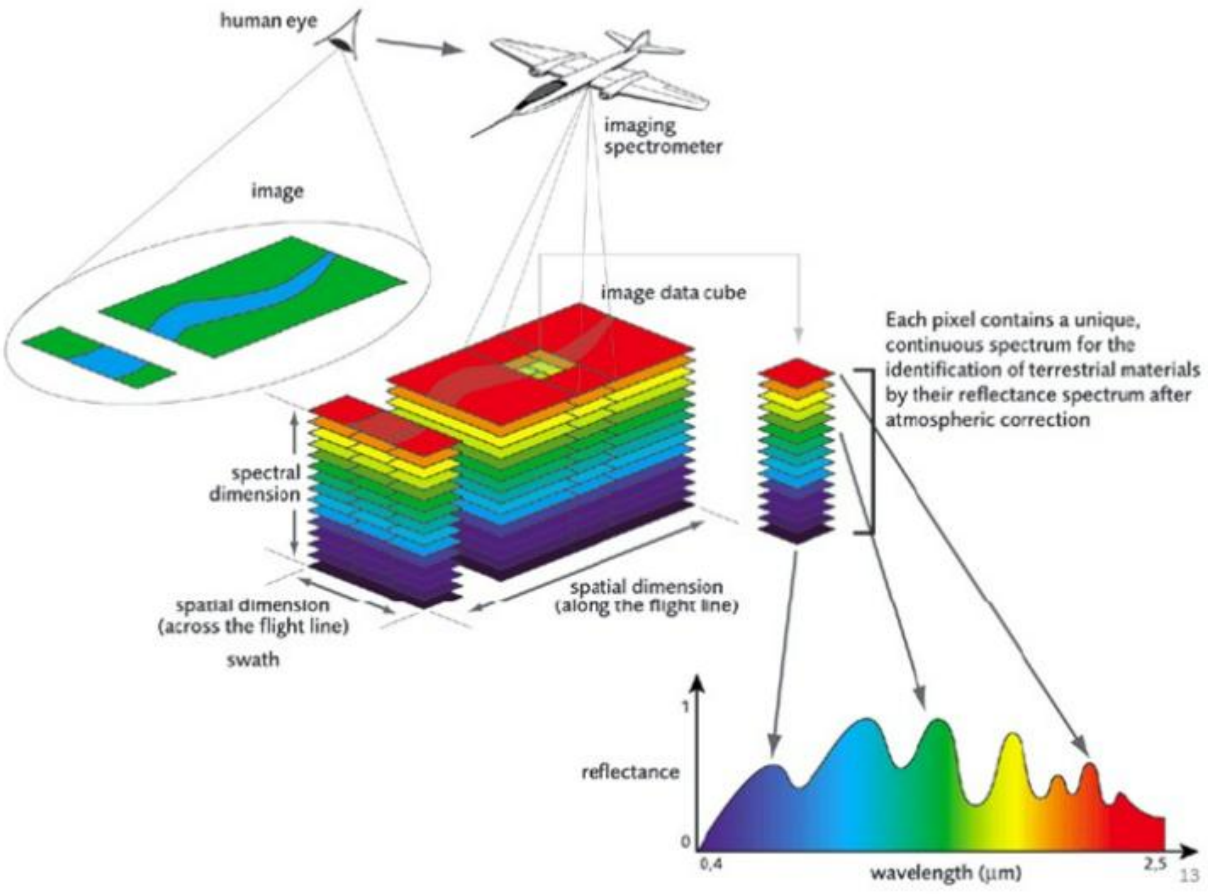
M Rupesh Kumar Yadav

Hyperspectral Imagery

Hyperspectral imaging is a area in remote sensing in which an imaging spectrometer collects hundreds of images at different wavelengths for the same spatial area

Hyperspectral Imagery & Limitations:

- Hyperspectral imagery is collected by Satellites, Aircraft, Drone platforms using special sensors like Hyperspectral Imaging Suite (HIS), Hyperion, AVAIS (Airborne Visible/Infrared Imaging Spectrometer)
- This practices also comes with certain limitations like need for atmospheric corrections, Noise in the data, high costs, data storage and transmission, etc.
- HSI data has mixed pixel issue



- Why Hyperspectral Data?

  - High Spectral & Spatial resolution.

  - other lower Alternatives Multispectral, Thermal, SAR, LiDAR.

- Why is it Interesting?

  - Combines both Spectroscopy & Imagery. Contiguous Spectral bands

  - Uses Environmental Modelling, Agriculture, Mineral Exploration, Military Surveillance, Urban Investigation

- How?

  - Using Classical Approaches like SVM, Random Forest, Spectral Angle Mapper(SAM)
  - Using CNN & Autoencoders

Datasets

1. Indian Pines
   - There are 16 classes in this dataset
   - Number of bands = 224

2. Pavia University
   - There are 9 classes in this dataset
   - Number of bands = 103

	Pavia University	**Indian Pines**
Sensor	ROSIS	AVIRIS
Location	Pavia, Italy	Northwest Indiana
Scene	Urban	Agricultural
Spectral Range	430 nm to 860 nm	400 nm to 2500 nm
Spatial Dimensions	610×340	145×145
Ground Sample Distance	1.3 m	20 m

ROSIS - Reflective Optics System Imaging Spectrometer
AVIRIS - Airborne Visible/Infrared Imaging Spectrometer

Datasets

Indian Pines

#	Class	Samples
1	Alfalfa	46
2	Corn-notill	1428
3	Corn-mintill	830
4	Corn	237
5	Grass-pasture	483
6	Grass-trees	730
7	Grass-pasture-mowed	28
8	Hay-windrowed	478
9	Oats	20
10	Soybean-notill	972
11	Soybean-mintill	2455
12	Soybean-clean	593
13	Wheat	205
14	Woods	1265
15	Buildings-Grass-Trees-Drives	386
16	Stone-Steel-Towers	93

Pavia University

#	Class	Samples
1	Asphalt	6631
2	Meadows	18649
3	Gravel	2099
4	Trees	3064
5	Painted metal sheets	1345
6	Bare Soil	5029
7	Bitumen	1330
8	Self-Blocking Bricks	3682
9	Shadows	947

LITERATURE REVIEW - Basic HSI Classification Methods

Dimensionality Reduction:

Dimensionality reduction methods aim to simplify data representation, which can lead to benefits such as improving efficiency of data analysis and modeling processes.

Principal Component Analysis (PCA):

- PCA is one of the most common dimensionality reduction techniques. It transforms a large set of variables into a smaller one that still contains most of the information from the large set, known as principal components.

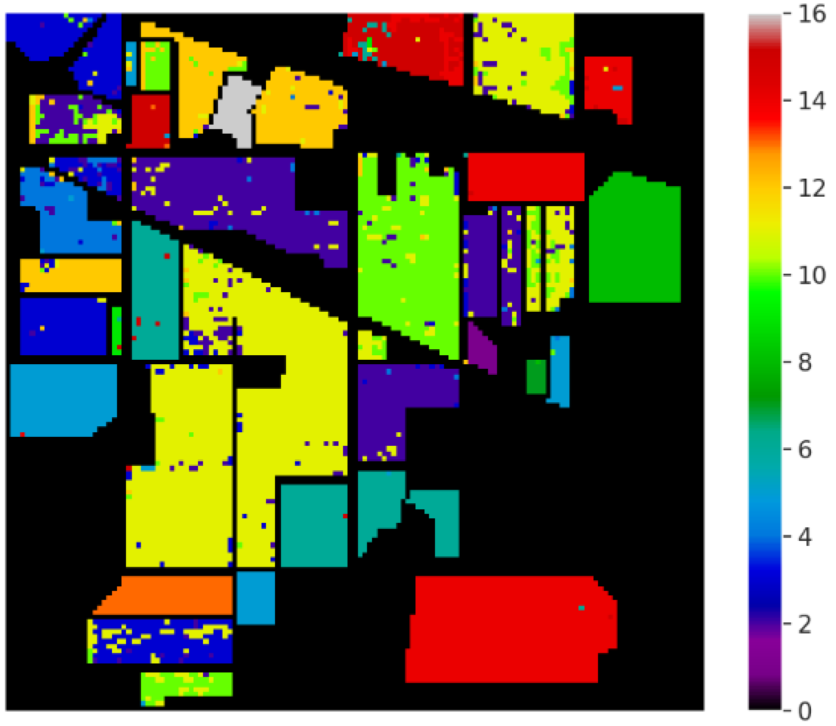
- Principal components are constructed in such a manner that the first principal component accounts for the largest possible variance in the data set.

LITERATURE REVIEW

Classification

Support Vector Machines(SVMs):

Support Vector Machines (SVMs) are commonly used for classification tasks in hyperspectral imagery data due to their ability to handle high-dimensional data even when the number of spectral bands is much larger than the number of samples (pixels) and their effectiveness in distinguishing between different classes.



LITERATURE REVIEW

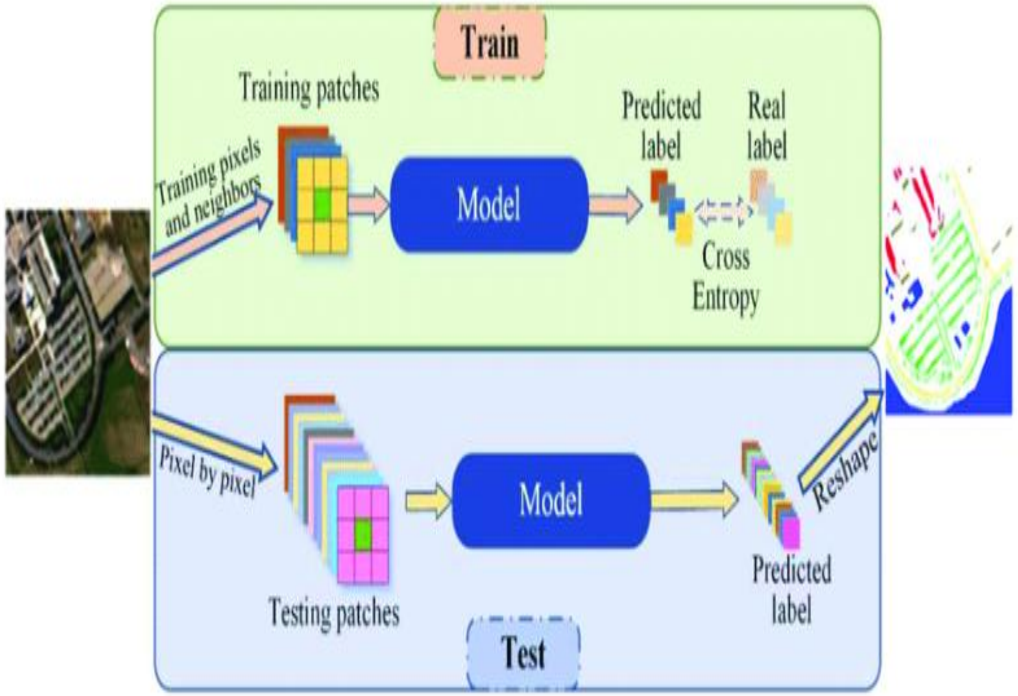
CONVOLUTIONAL NEURAL NETWORKS(CNN) Model:

The advancements of deep learning and computational power helped to develop effective deep learning models for hyperspectral image classification. Convolutional Neural Networks(CNN) are widely used to extract potential spectral and spatial features for the classification of the satellite imagery.

- Data Preprocessing: Preprocess the data, which includes loading the hyperspectral data, data analysis.
- Dimensionality Reduction using Principal component analysis(PCA)
- Padding: Zero-padding can help when applying 3D convolutions to the data, where the third dimension typically represents different spectral bands.

LITERATURE REVIEW

- Creating 3D patches: We need to break down the satellite image into patches, every patch will have a class.
- Splitting the data into training and test set.
- One-hot encoding: For representing categorical/class labels as binary vectors.
- CNN Model:3D CNN with multiple layers such as Convolution, Dropout, and Dense Layers are built according to our specific choice.
- Training the model and Validating it.



LITERATURE REVIEW

DRAWBACKS OF USING CNNs:

- CNNs are computationally complex when the size of the data (the number of spectral bands) increases. This results in longer training times and a higher demand for computational resources.
- CNNs does local operations and cannot model long range dependencies.
- Bound to noise in data.

PROBLEM STATEMENT

- Need for robust algorithms that can leverage the spectral and spatial information from Hyperspectral Data to enhance classification and segmentation performance.

PROPOSED METHODOLOGY

EarthMapper – ToolBox is a modular framework containing various pretrained self-taught feature learning frameworks, a classifier and a Fully connected undirected graph model for segmentation of Remote sensing Imagery.

There are three frameworks for self-taught feature learning. They are:

- Stacked convolutional autoencoder (SCAE)
- Stacked multi-loss convolutional autoencoder (SMCAE)
- Multi-scale independent component analysis (MICA)

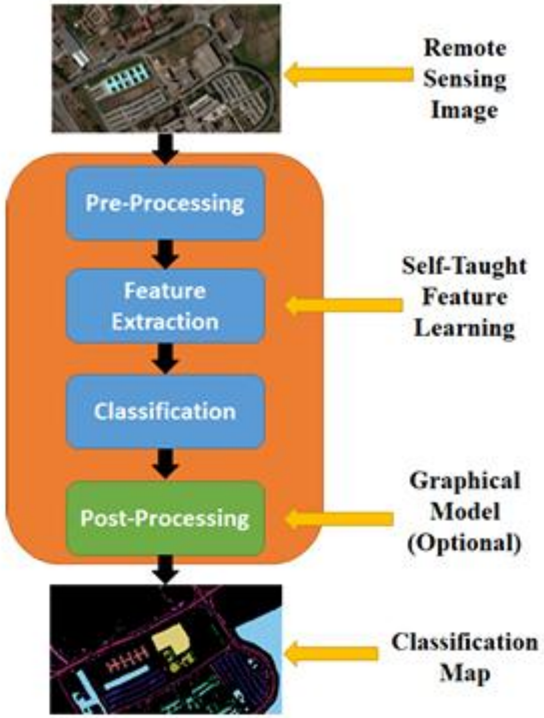


Fig4 - EarthMapper's pipeline

Stacked Convolutional Autoencoder (SCAE)

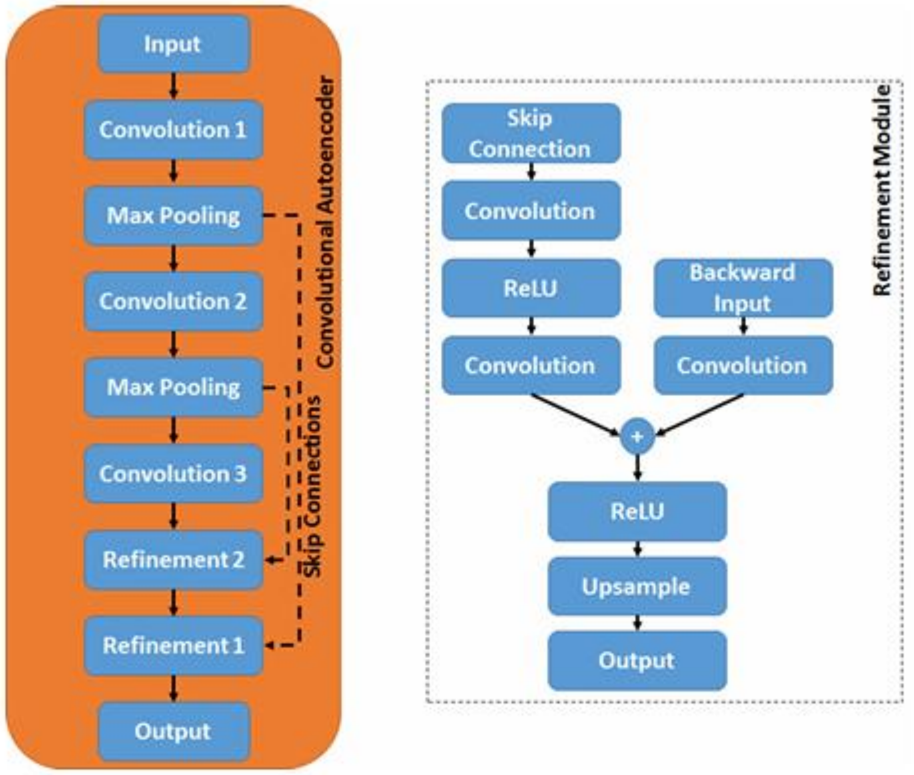


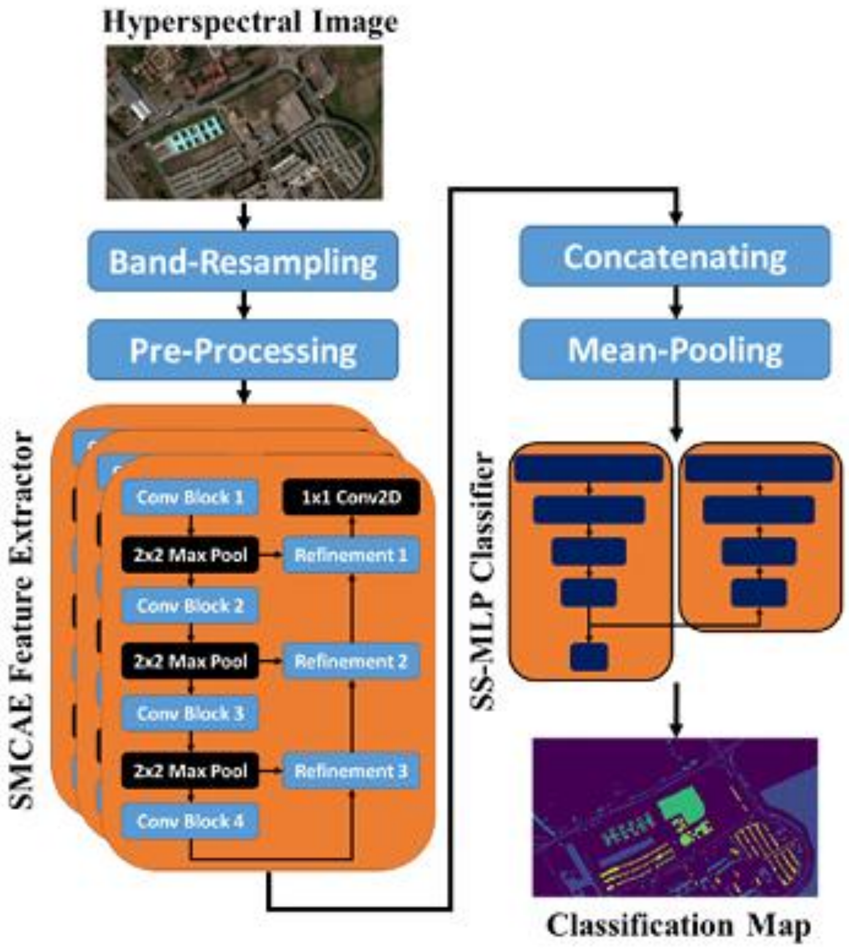
Fig5 -CAE modules & Refinement module

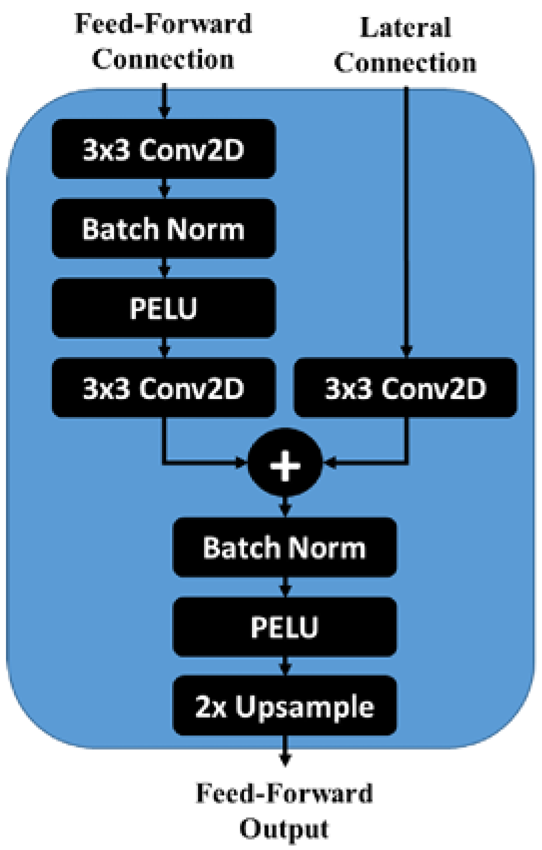
- Three Stacked Convolutional Autoencoder (CAE)
- Refinement Modules
- Loss function - Mean Square Error (L$_{mse}$)

$$\mathbf{H}^k = \sigma\left(\mathbf{W}^k * \mathbf{H}^{k-1}\right)$$

Where $H^0 = X$ is the input, k layers

Stacked multi-loss convolutional autoencoder (SMCAE)





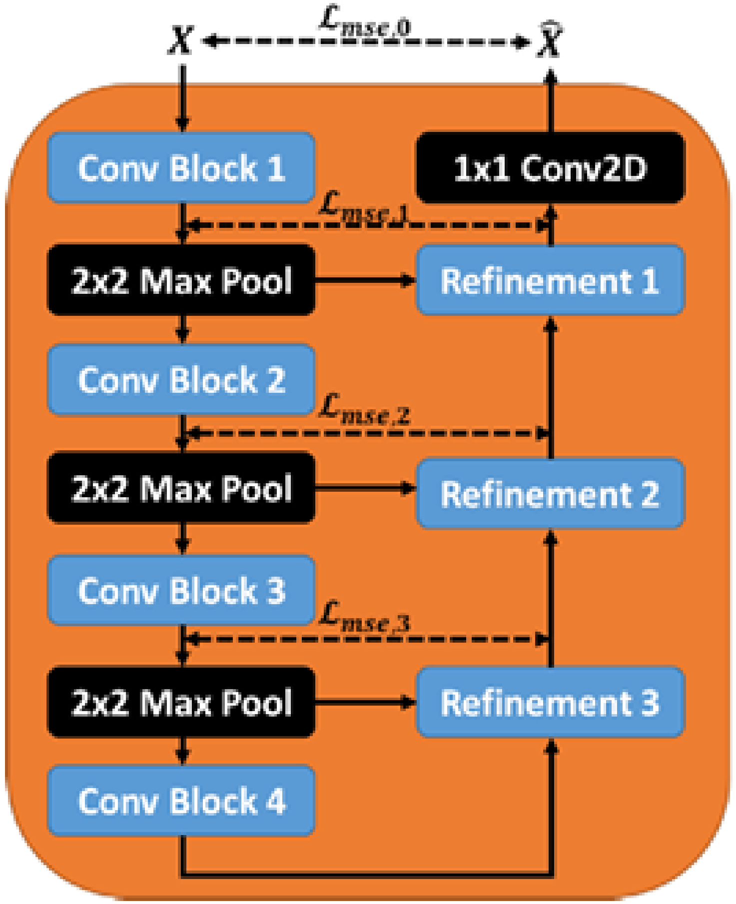
Refinement Layer

Stacked multi-loss convolutional autoencoder (SMCAE)

Training each CAE using a weighted sum of the reconstruction losses for each hidden layer.

$$\mathcal{L} = \sum_{j=1}^{M} \lambda_{mcae,j} \mathcal{L}_{mse,j}$$

Where, M is the number of hidden-layers, $L_{mse,j}$ is the MSE of the encoder and decoder at layer j, and $\lambda_{mcae,j}$ is the loss weight at layer j.



PROPOSED METHODOLOGY

Multi-scale independent component analysis (MICA):

- shallow feature extractor used for Thermal Imagery etc
- operates by convolving learned filters with HSI data to detect bar/edge structures, gradients, and blobs.
- resemble neurons in the primary visual cortex.
- The optimal receptive field size of MICA depends on the Ground Sample Distance
- Min-Max Scalar is used as a preprocessor

Semi-Supervised Multilayer Perceptron (SS-MLP)

SS-MLP neural network for classification.

Inputs: (1) a feature vector extracted from the raw HSI cube. or

(2) SMCAE features re-sampled to zero-mean/unit-variance

SSMLP assigns each pixel a set of probabilities that it belongs to a given class.

1) taking the index (i.e., argmax) of the maximum probability to give us our class label or
2) Passing those probabilities to the CRF to post process the classification map, which reduces salt-and-pepper classification errors.

SS-MLP is trained by minimizing the total supervised and unsupervised loss,

$$\mathcal{L} = \mathcal{L}_{class} + \sum_{j=1}^{M} \lambda_{recon,j} \cdot \mathcal{L}_{recon,j}$$

Semi-Supervised Multilayer Perceptron (SS-MLP)

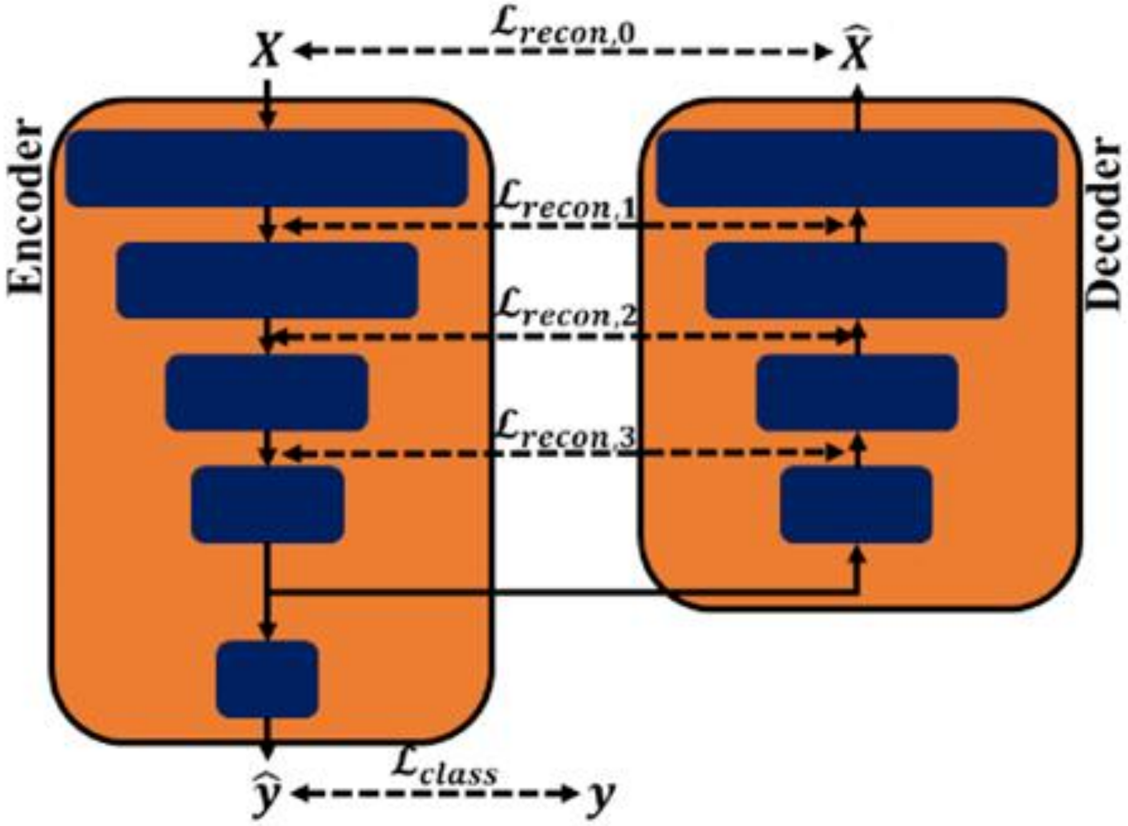


Fig - The semi-supervised multi-layer perceptron (SS-MLP) classification framework

Undirected Graphical Models for Postprocessing

The expressiveness of the UGMs is controlled by the structure of the graph and energy functions defined over the graph’s cliques.

$$p(\mathbf{y}) = \frac{1}{Z} \exp\left(-E\left(\mathbf{y}\right)\right)$$

Z = $\sum_{\mathbf{y}} \exp\left(-E\left(\mathbf{y}\right)\right)$

where y = [y1,..,yN]T is a vector containing labels of all N pixels in the image, E(y) is the total energy and Z is the partition function. The total energy is equal to the sum of unary energies and pairwise energies defined over all nodes and edges of the graph

$$E\left(\mathbf{y}\right) = \sum_{i \in V} E_i\left(y_i\right) + \sum_{(i,j) \in D} E_{ij}\left(y_i, y_j\right)$$

PROPOSED METHODOLOGY

The unary energy function is given by the negative log arithm of the class probability predicted by the classifier,

$$E_i\,(y_i) = -\log(P(y_i/x_i))$$

The pairwise energy function used depends on the spatial location of the pixels and is given by,

$$E_{ij}\,(y_i, y_j) = \begin{cases} 0, & \text{if } y_i = y_j \\ w_1 \exp\left(-\frac{|\mathbf{p}_i - \mathbf{p}_j|^2}{2\theta_\gamma^2}\right), & \text{otherwise,} \end{cases}$$

As The number of edges in a fully connected graph grows at O(n2) with the number of nodes, it is inefficient to perform inference in such models using standard algorithms, we use mean-field approximate inference.

References:

[1] U. B. Gewali and S. T. Monteiro. A tutorial on modeling and inference in undirected graphical models for hyperspectral image analysis.

[2] Self-taught feature learning for hyperspectral image classification. R. Kemker and C. Kanan.

[3] Low-shot learning for the semantic segmentation of remote sensing imagery. R. Kemker, R. Luu, and C. Kanan.

[4] Spectral-spatial classification of hyperspectral image using autoencoders. Z. Lin, Y. Chen, X. Zhao, and G. Wang.

[5] C. Tao, H. Pan, Y. Li, and Z. Zou. Unsupervised spectral–spatial feature learning with stacked sparse autoencoder for hyperspectral imagery classification.

References:

[6] P. Kr¨ahenb¨ uhl and V. Koltun. Efficient inference in fully connected CRFs with Gaussian edge potentials. In Advances in neural information processing systems.

[7] Y. Boykov, O. Veksler, and R. Zabih. Fast approximate energy min imization via graph cuts.

[8] H. Valpola. From neural PCA to deep unsupervised learning. Advances in Independent Component Analysis and Learning Machines,

Thankyou