Feature Extraction for Rasters Using Autoencoders

Miloš Manić Mladen Nikolić
Faculty of Mathematics, University of Belgrade, Serbia

GeoMLA, Belgrade, Serbia

June 23, 2016.

Overview

Overview

Feature extraction

- Informative features are essential for learning
- Features are often handcrafted, but automated feature extraction methods exist
- Neural networks and autoencoders
- Information retrieval
- Using compact representations is more efficient with respect to computing time

Contributions

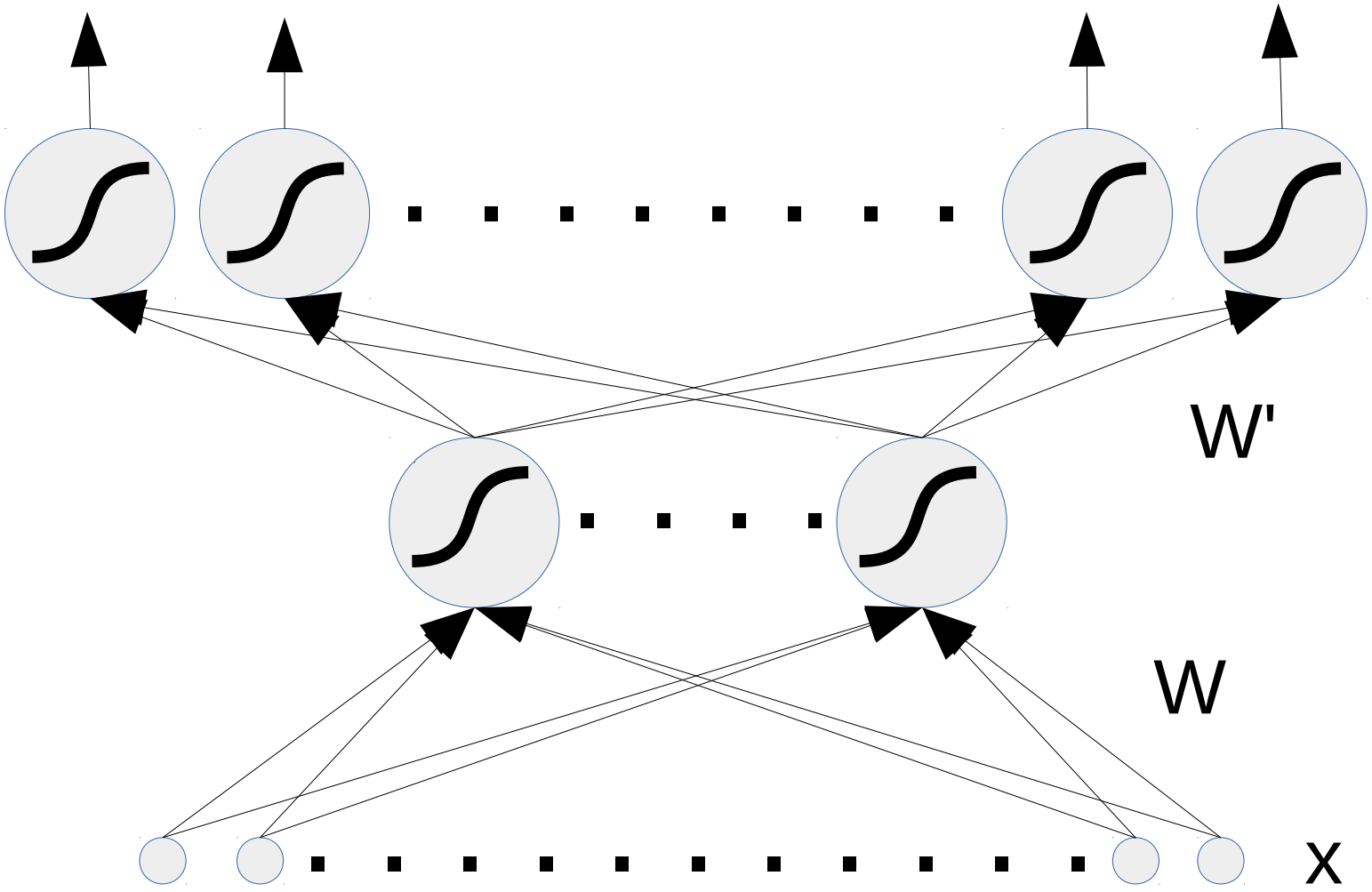
- We extract small number of informative features for temperature rasters of Serbia
- We show that distances over these features are highly correlated with distances over original rasters
- We propose procedure to speed up autoencoder training on rasters

Overview

Autoencoders

- Kind of unsupervised neural network
- Encoder $f : \mathbb{R}^N \to \mathbb{R}^M$
- Decoder $g : \mathbb{R}^M \to \mathbb{R}^N$
- $g(f(x)) \approx x$

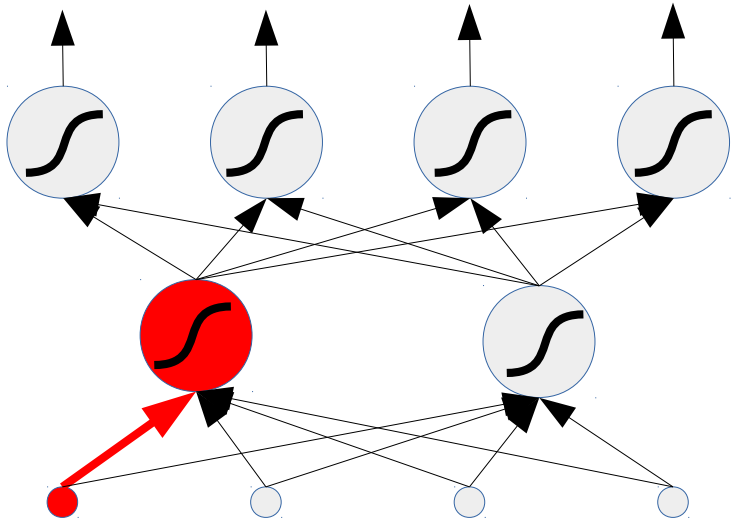
Simple Autoencoder Architecture

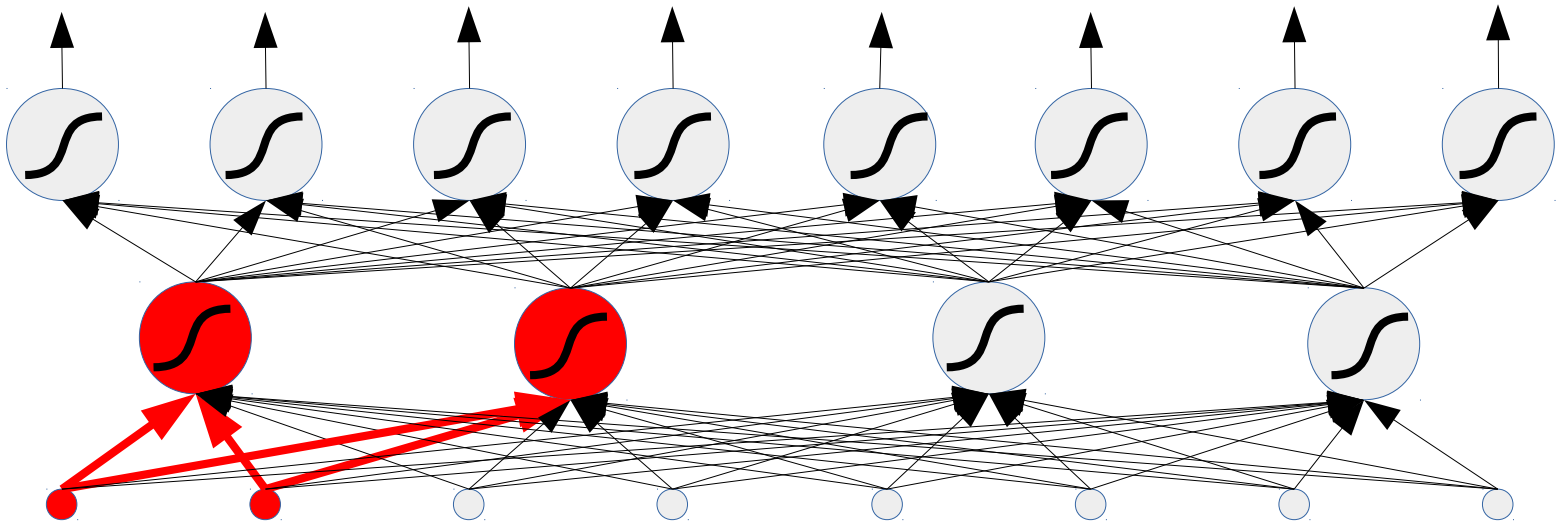


Overview

- Train small autoencoder on low resolution rasters
- Use its weights to initialize bigger autoencoder for higher resolution rasters
- Train bigger autoencoder
- Proceed to even higher resolutions

Incremental Training Idea





Overview

Data

- Maximal daily temperature rasters of Serbia in period 2006-2015
- Resolutions: $8km \times 8km$, $4km \times 4km$, $2km \times 2km$
- Pixels: 1973, 7888, 31645
- First 9 years used for training and last one for testing

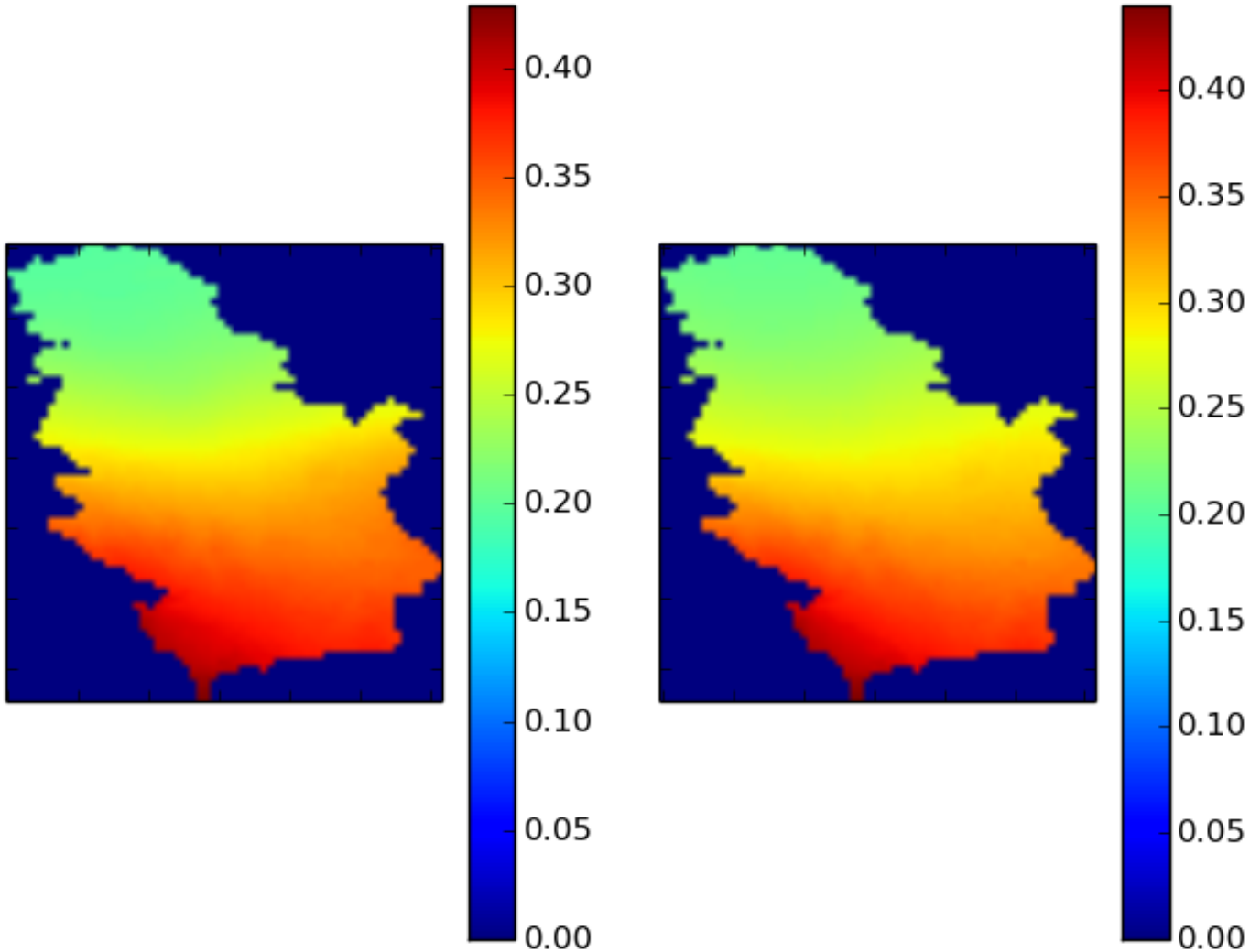
Experimental Setting

- Incremental training: $8km \times 8km$, $4km \times 4km$, $2km \times 2km$
- Non-incremental training: $2km \times 2km$
- 12 hidden units (features)
- Laptop computer Intel Core i7, 4 cores, 2.1GHz, 6GB RAM
- TensorFlow – Google platform for deep learning

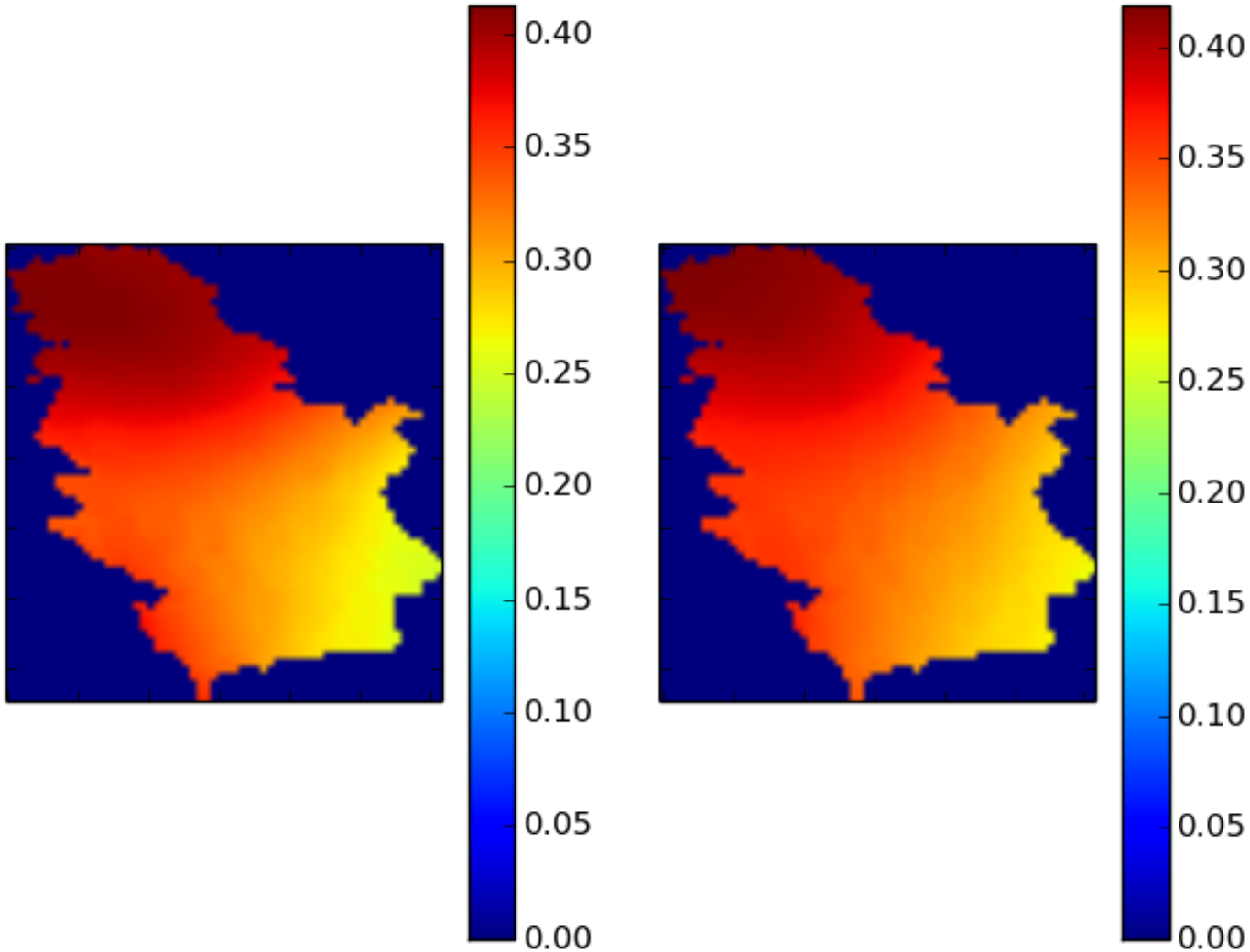
Reconstruction Quality

	Time (s)	Test error	Compression
Incremental	7643	0.32°C	99.65%
Non-incremental	37539	1.05°C	99.65%

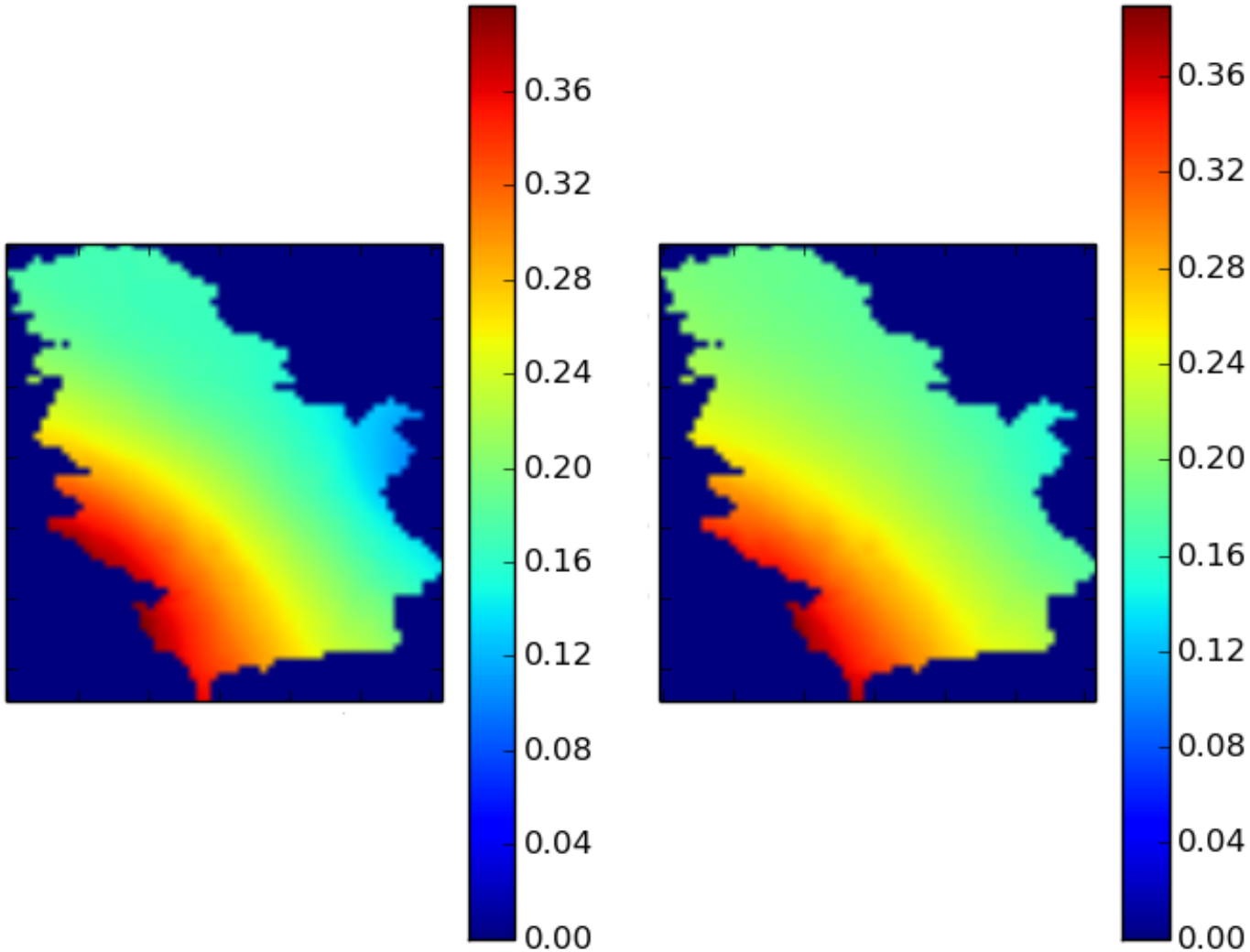
Reconstruction Example



Reconstruction Example



Reconstruction Example



Feature Quality

- Correlation between distances over extracted features and distances over high resolution rasters
- Average rank of truly nearest raster if rasters were retrieved based on distance over extracted features

	Correlation	Average rank
Incremental	0.95	5.68
Non-incremental	0.93	8.43

Overview

Future Work

- Details of incremental training procedure
- More thorough evaluation
- Raster search engine

Thank you!