Application of Softmax Regression and its Validation for Spectral-Based Land Cover Mapping

[a]Jason Wolfe, [a]Xiaoying Jin, [b]Thomas Bahr & [b]Nicolai Holzer

[a]Harris Corporation, Broomfield, Colorado, U.S.A. | (jwolfe05, xjin)@harris.com

[b]Harris Corporation, Gilching, Germany | (thomas.bahr, nicolai.holzer)@harris.com

ISPRS Hannover Workshop 2017

Hannover, Germany | June 6 – 9, 2017

SOFTMAX REGRESSION

SOFTMAX REGRESSION is a generalization of logistic regression that is used for multi-class classification where classes are mutually exclusive.

The **CLASSIFICATION WORKFLOW** allows training and evaluating the performance of a classifier once and then applying it to similar data multiple times.

The workflow uses the **ENVI APPLICATION PROGRAMMING INTERFACE** (API). This API in combination with the multiple image approach predisposes the workflow to be executed in both desktop and enterprise environments, e. g. to support time-series analysis of specific features.

A **CASE STUDY** is used to demonstrate and evaluate this Softmax Regression approach, exemplified by the classification of an urban environment with five simple classes: Asphalt, Concrete, Grass, Tree, and Building.

DATA PREPARATION

TEST SITE Urban area of Fruita, Colorado, USA.

DATA FILES

- Hyperspectral image, acquired 17 July 2013 with NEON imaging spectrometer (atmospherically corrected).
- Relative height image at 0.5 m resolution, derived from NEON point-cloud data.
- Digital orthorectified photograph from the NEON RGB camera at 0.25 m resolution.

MULTI-BAND ATTRIBUTE IMAGE, a layer stack of six attributes:

- Enhanced Vegetation Index (EVI)
- Height (m)
- Blue reflectance
- Green reflectance
- Red reflectance
- Near-infrared (NIR) reflectance

COLLECTION OF TRAINING DATA

EXTRACT EXAMPLES With supervised classification, one collects samples of pixels that belong to each class. This training data is used to train the classifier and classify pixels of unknown identity into known classes.

Training data for this case study was collected by drawing regions of interest (ROIs) for each of the five feature types:

- Asphalt,
- Concrete
- Grass,
- Tree,
- Building,

on the high-resolution orthorectified NEON photograph that was coregistered with the NEON reflectance image to be classified.

The one-dimensional array (or vector) of pixel values of a single pixel within the ROI is called an ***example***.

DEFINITION AND TRAINING OF THE CLASSIFIER

SHUFFLE EXAMPLES The examples were shuffled to create a random distribution of data.

SPLIT EXAMPLES Then the examples were split into two separate array elements. The first array element contained the examples that were used to train the classifier. The second array element contained the examples for the evaluation of the classifier.

DEFINE CLASSIFIER & TRAINER The Softmax Regression classifier should use a gradient descent trainer. A gradient descent algorithm iteratively updates the classifier according to the classifier's gradient for that iteration. It iterates until a loss function converges on a minimum value.

TRAIN CLASSIFIER The examples and class values are passed to a trainer to iteratively minimize the classification error. The classifier will be trained on one set of examples and evaluated with another set.

EVALUATION OF THE CLASSIFIER AND FINAL CLASSIFICATION

EVALUATE CLASSIFIER The performance of the Softmax Regression classifier was evaluated using the examples and corresponding truth class values that were not used to train the classifier.

Predicted class values from these input examples were calculated.

Finally accuracy metrics derived from a confusion matrix between the truth class values and the predicted class values were computed.

CLASSIFY RASTER The final step is to classify the attribute image. Prior to this step the gains and offsets were applied to the attribute image to assure a proper scaling of the pixel values among the different attributes.

After classifying the attribute image, we applied the trained classifier to a neighboring dataset of the same sensor, which had the same attributes and data representation.

The two classification images overlap precisely.

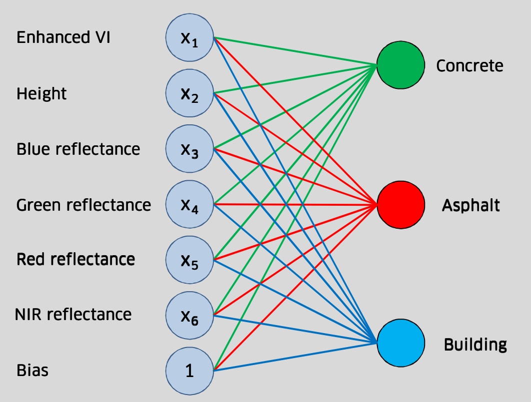
CONCLUSIONS

The Softmax Regression classifier can be created and trained on a reference dataset using spectral and spatial information and then applied to similar data multiple times. It can also be trained in stages where subsequent training will further refine the classifier.

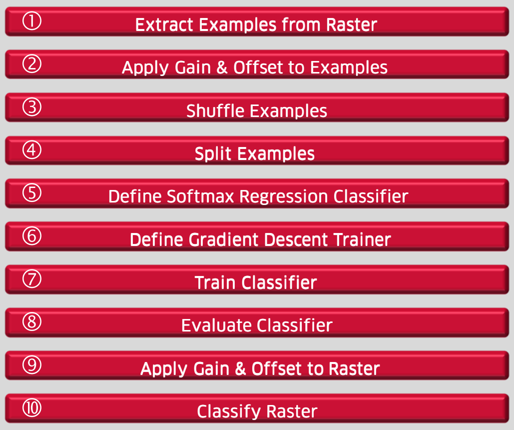
We evaluated this approach within a case study about the classification of an urban environment with five simple classes.

After classifying one attribute image derived from hyperspectral airborne and elevation imagery, we applied the trained Softmax Regression classifier successfully to a neighboring dataset from the same sensor.

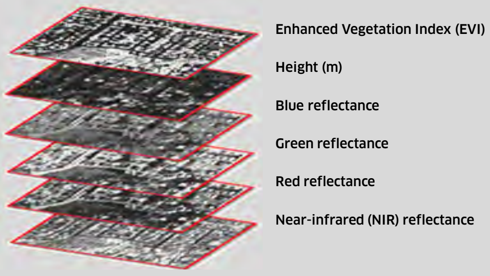
As the workflow is based on the ENVI API it can be embedded in any existing geospatial workflow for operational applications, including both desktop and enterprise environments. Integration options are e. g. the Harris Geospatial Services Framework, or ArcGIS®.



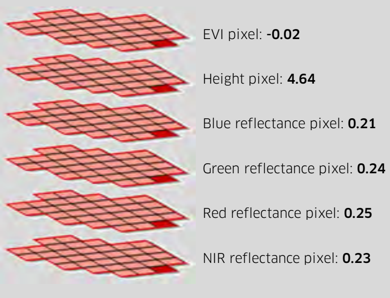
Example of a Softmax Regression model with 6 attributes and 3 classes. The weights were omitted from the diagram for clarity.



Softmax Regression classification workflow using ENVITasks.



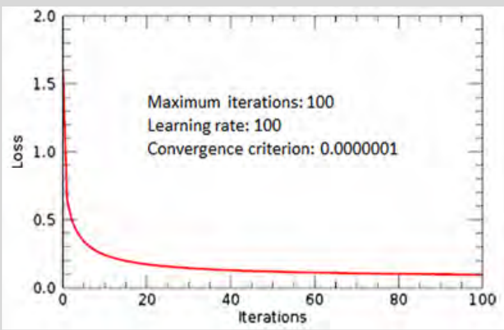
Example of a multi-band attribute image, where each band represents a different attribute.



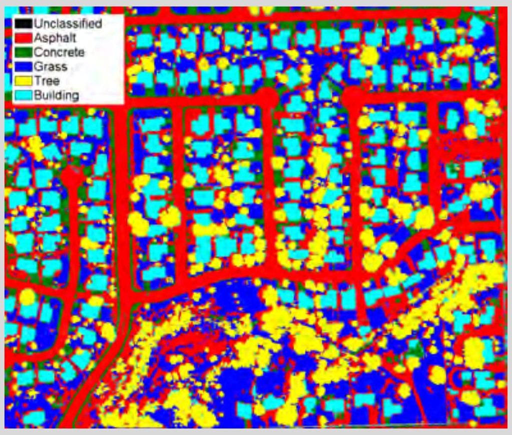
Example of the pixel values for six attributes in one ROI polygon using the building class.

True-color representation of the hyperspectral reflectance image, acquired 17 July 2013 with a NEON imaging spectrometer over Fruita, Colorado, USA.

The Regions of Interest overlaid refer to the classes below.



Loss profiles are used for the evaluation of the Softmax Regression classifier. The curve should flatten and approach a minimum value.

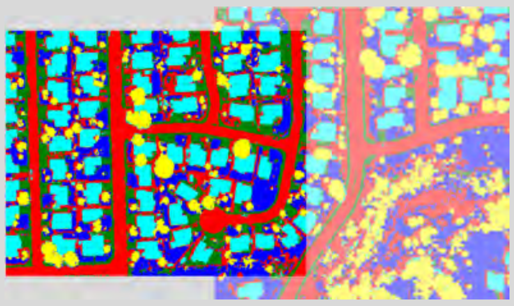


Classification result of the hyperspectral NEON image above.

	Asphalt	Concrete	Grass	Tree	Building
Error of commission	0.004	0.012	0.010	0.005	0.007
Error of omission	0.006	0.015	0.001	0.008	0.007
F1 value	0.995	0.986	0.994	0.993	0.993
Precision	0.996	0.988	0.990	0.995	0.993
Producer accuracy	0.994	0.985	0.999	0.992	0.993
Recall	0.994	0.985	0.999	0.992	0.993
User accuracy	0.996	0.988	0.990	0.995	0.993

Overall accuracy	0.994	Kappa coefficient	0.991

Accuracy metrics for the Softmax Regression classifier.



Merge of two classification images. The Softmax Regression classifier was trained on the attribute image corresponding to the right classification image and then applied on the attribute image corresponding to the left classification image.