Author: Adam Wehmann

Correspondence: adamwehmann+svmtool@gmail.com

More information: <http://www.adamwehmann.com/>

SVM Tools for ArcGIS 10.1+, version 1.5

This ArcGIS 10.1+ toolbox provides tool for the supervised classification of raster imagery by support vector machine using a Gaussian RBF kernel. Three tools are provided:

1. SVM Train and Classify – trains an SVM model from an input raster and associated training data. Optionally classifies that raster using the trained model. Additionally able to produce class conditional probability estimates.
2. SVM Resume – classifies an input raster using a previously trained SVM model. Also able to resume previously interrupted classification provided parameters are set appropriately to match those of the interrupted classification.
3. Assess Accuracy – performs an accuracy assessment of an output classification using associated testing data. Error matrix, overall accuracy, producer and user’s accuracies, and kappa statistic are produced.
4. Create Training Raster – creates a training raster from a vector dataset following conventions needed for input into other tools.

Training and testing data should be a single-band raster with known pixels give a value in the label set {1,...,n} and unknown pixels given a value of 0. These raster images should have the exact same extent and gridding as the input raster. This can be ensured by setting the Extent, Snap Raster, and Cell Size options appropriately in the environment settings during production.

All rasters used by this tool should have a pixel depth of at least 8 bits. Smaller pixels depths will be interpreted as Boolean array by ArcPy’s RasterToNumPyArray function which will result in misclassification. Pixel depth can be promoted using the Copy Raster tool provided in the ArcGIS Data Management toolbox.

The recommended method of producing training and testing data for this tool is to generate a distribution of random points over the image to be classified using the Create Random Points tool in the ArcGIS Data Management toolbox. Set the ‘constraining extent’ option to your image and set ‘number of points’ to some large number, such as 1000, 2000, or more depending on the size of study area. Create an integer datatype field in the attribute table of this vector point layer. Start an Edit Session and use the pan and zoom tools to inspect the location of each point, manually assigning a class label to it. Make sure to follow the labeling convention mentioned previously. Continue labeling until you’ve reached a sufficient number of samples for each class. Save your edits and exit the edit session. Now use the Point to Raster tool to convert this vector data to raster format, setting the Extent and Snap Raster environments to match the image you want to classify. Use the Reclassify tool to set any unlabeled pixel to a value of 0 and the Copy Raster tool, if necessary, to promote the pixel depth. Other methods of producing training samples are also possible.

Training and classification proceeds blockwise in order to deal with the memory requirements of large raster datasets. Changing the “blocksize” parameter in the tool options will change the size of the data block used. Larger blocks will take up more memory, but may result in faster tool execution. When classifying large images, it is recommended that an SVM be trained, but not used to classify until a high training accuracy can be achieved. Leaving the Output Classification option blank in the SVM Train and Classify tool will achieve this, causing the tool to train a SVM, but not perform classification with it. The training accuracy after cross-validation will be shown in the message window. If this percentage is not high, the training data used is poor quality and will not be able to produce a good classification. It is a waste of time and computation to classify when training accuracy is low.

When classifying an image, the tool will produce a number of temporary files in the same directory as the output path for the classification. For this reason, it’s recommended that output files be output to a new empty directory to avoid name conflicts. Additionally, the SVM Resume tool can resume the classifying an image if the tool’s execution was canceled or some other error occurred. To do this, you must set the same options as you used previously, which will include outputs that point to files (not yet created) in this directory containing temporary files. The temporary files are automatically cleaned up when the tool has finished executing successfully.

If you use these tools in scientific research, please cite:

Wehmann, A. 2013. *SVM Tools for ArcGIS 10.1+*. Columbus, OH. Software available at: <http://www.adamwehmann.com/>

and

Chang, C-C., Lin, C-J. 2011. LIBSVM: a library for support vector machines, ACM Transactions on Intelligent Systems and Technology, 2(27), pp. 1-27. Software available at: http://www.csie.ntu.edu.tw/~cjlin/libsvm

Send any bugs found, feature requests, questions to the email address provided at the top of this document.

---

LICENSING:

This tool is granted free of charge.

Because the tool is licensed free of charge, there is no warranty for the tool, to the extent permitted by applicable law except when otherwise stated in writing the copyright holders and/or other parties provide the program "as is" without warranty of any kind, either expressed or implied, including, but not limited to, the implied warranties of merchantability and fitness for a particular purpose. The entire risk as to the quality and performance of the tool is with you. Should the tool prove defective, you assume the cost of all necessary servicing, repair or correction.  
  
In no event unless required by applicable law or agreed to in writing will any copyright holder, or any other party who may modify and/or redistribute the tool, be liable to you for damages, including any general, special, incidental or consequential damages arising out of the use or inability to use the tool (including but not limited to loss of data or data being rendered inaccurate or losses sustained by you or third parties or a failure of the tool to operate with any other programs), even if such holder or other party has been advised of the possibility of such damages.

---

LIBSVM Library Notice:

Copyright (c) 2000-2012 Chih-Chung Chang and Chih-Jen Lin

All rights reserved.

Redistribution and use in source and binary forms, with or without

modification, are permitted provided that the following conditions

are met:

1. Redistributions of source code must retain the above copyright

notice, this list of conditions and the following disclaimer.

2. Redistributions in binary form must reproduce the above copyright

notice, this list of conditions and the following disclaimer in the

documentation and/or other materials provided with the distribution.

3. Neither name of copyright holders nor the names of its contributors

may be used to endorse or promote products derived from this software

without specific prior written permission.

THIS SOFTWARE IS PROVIDED BY THE COPYRIGHT HOLDERS AND CONTRIBUTORS

``AS IS'' AND ANY EXPRESS OR IMPLIED WARRANTIES, INCLUDING, BUT NOT

LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR

A PARTICULAR PURPOSE ARE DISCLAIMED. IN NO EVENT SHALL THE REGENTS OR

CONTRIBUTORS BE LIABLE FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL,

EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO,

PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR

PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF

LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING

NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY OUT OF THE USE OF THIS

SOFTWARE, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.