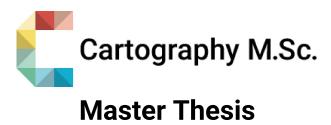
This is an excerpt from the thesis "Semi-Supervised Virtual Support Vector Machines with Self-Learning Constraint for Remote Sensing Image Classification".

Please contact Ozan Tuncbilek for a full version of the thesis.

Extended works on this topic are documented in: Geiß, C., Aravena Pelizari, P., Tuncbilek, O., and Taubenböck, H. (2023) Semi-supervised Learning with Constrained Virtual Support Vector Machines for Classification of Remote Sensing Image Data. International Journal of Applied Earth Observation and Geoinformation, 125, 103571.



Semi-Supervised Virtual Support Vector Machines with Self-Learning Constraint for Remote Sensing Image Classification

Ozan Tuncbilek

2021







Semi-Supervised Virtual Support Vector Machines with Self-Learning Constraint for Remote Sensing Image Classification

Submitted for the academic degree of Master of Science (M.Sc.) conducted at the Department of Aerospace and Geodesy Technical University of Munich

Author: Ozan Tuncbilek Study Course: Cartography M.Sc

Supervisors: Dr. Christian Murphy(TUM), Dr. Christian Geiß (DLR)

Prof. Dr.-Ing. Günter Strunz (DLR)

Reviewer: Mahdi Khodadadzadh(ITC)

Cooperation: German Aerospace Center (DLR)

Chair of the Thesis

Assessment Board: Prof. Dr. Ligiu Meng

Submission Date: 10.09.2021

Abstract

In real-world applications, it is difficult to collect labeled samples, and supervised learning methods rely on the quality of this labeled training data. Therefore, in this research, a semi-supervised learning approach is developed in order to benefit from the unlabeled samples that can be produced effortlessly. These semi-supervised methods are built on a popular machine learning technique called support vector machine, which is used to classify remote-sensing imagery in this thesis. Moreover, this work aims to enhance the accuracy of the methods in settings with very few labeled samples and deploy a constrained set of unlabeled samples with a self-learning strategy. Additionally, the aim includes model evaluation for existing support vectors and virtual samples. Moreover, the methodology is further extended with an active learning method. This extension involves uncertainty visualizations in order to increase the model accuracy by relabelling the uncertain samples in a prioritized way. To evaluate these models, experimental results were obtained over the city of Cologne, Germany, and the Hagadera Refugee Camp, Kenya from a very high spatial resolution multispectral data set. Results from newly proposed methods showed favorable performance properties, especially on the few labeled samples. Furthermore, the uncertainty of the models was compared with the active learning extension, and this extension also increased the accuracy.

Contents

1.1 1.2 1.3	Motivation and problem statement	9
		-
1.3	1.2.1 Research objectives	
1.3		9
	Innovation aimed at	10
The	oretical Background & Literature Review	11
2.1	Theoretical Background	11
	2.1.1 Support Vector Machines	11
	2.1.2 Semi-Supervised Support Vector Machines	14
	2.1.3 Uncertainty	15
	2.1.4 Support Vector Machines Settings	17
	2.1.5 Virtual Support Vector Machines	17
	2.1.6 Self-learning strategy	19
2.2	Literature Review	21
		21
	2.2.2 Uncertainty Visualization and Active Learning Methods	26
Met	hodology	28
3.1	Semi-Supervised Methods	28
		28
		30
	i S	31
3.2	Uncertainty Visualization	31
Res	ults & Discussion	34
4.1	Data	34
4.2	Results from Data set I: Cologne	36
4.3		48
4.4	Case Study of Extension on Active Learning	59
Con	clusion	64
Refe	arences	65
	Theo 2.1 2.2 Meti 3.1 3.2 Resi 4.1 4.2 4.3 4.4 Cond	Theoretical Background & Literature Review 2.1 Theoretical Background . 2.1.1 Support Vector Machines . 2.1.2 Semi-Supervised Support Vector Machines . 2.1.3 Uncertainty . 2.1.4 Support Vector Machines Settings . 2.1.5 Virtual Support Vector Machines . 2.1.6 Self-learning strategy . 2.2 Literature Review . 2.2.1 Theory and the Applications of Machine Learning Methods . 2.2.2 Uncertainty Visualization and Active Learning Methods . Methodology . 3.1 Semi-Supervised Methods . 3.1.1 VSVM-SL with Unlabeled Samples . 3.1.2 VSVM-SL with Virtual Unlabeled Samples . 3.1.3 Benchmark Method for Semi-Supervised Learning . 3.2 Uncertainty Visualization . Results & Discussion . 4.1 Data . 4.2 Results from Data set I: Cologne . 4.3 Results from Data set II: Hagadera . 4.4 Case Study of Extension on Active Learning . Conclusion

List of Figures

1	An example of linearly separable case. (Cortes and Vapnik, 1995)	12
2	Shattered three points by oriented lines. (Burges, 1998)	12
3	Transformation of input vectors to the higher dimensional space (Song	
	et al., 2012)	13
4	Example of mixed pixels causing uncertainty (Bastin et al., 2002)	15
5	Example of fuzzy classification (Bastin et al., 2002)	15
6	Working principle of Virtual Support Vector Machines (a),(b). (Geiß et al.,	
	2019)	18
7	Working principle of Virtual Support Vector Machines (c), (d). (Geiß et al.,	
	2019)	18
8	Working principle of Self-Learning strategy on VSV (Geiß et al., 2019)	19
9	Linear separating hyperplanes for the non-separable case (Burges, 1998)	22
10	Scheme of the proposed approach of multi-level context-based classifi-	
	cation strategy (Bruzzone and Carlin, 2006)	24
11	Main procedures of the proposed synergetic classification approach (Lu	
	et al., 2016)	25
12	Example showing the change in uncertainty (Kinkeldey, 2014)	26
13	Images into confidence map after 5 th iteration (Tuia and Munoz-Mari, 2012)	26
14	Supervised texture-based segmentation and related uncertainty for all	
	object building blocks (Lucieer, 2004)	27
15	Working principle of VSVM-SL with Unlabeled Samples	29
16	Illustration of distance to the hyperplane	32
17	Working principle of Active Learning Approach with Uncertainty on VSVM-	
	SL-U Method	33
18	(a)WorldView-II scene of Cologne, Germany; (b) Available labeled sam-	
	ples per class; (c) spatially disjoint training, testing, and validation areas.	35
19	(a)WorldView-II scene of Hagadera Refugee Camp, Kenya; (b) Available	
	labeled samples per class; (c) spatially disjoint training, testing, and	
	validation areas	36
20	Cologne- Binary classification setting- invariance to scale- Mean Accuracy	37
21	Cologne-Binary classification setting-invariance to shape- Kappa Statistics	38
22	Cologne- Binary classification setting- invariance to shape-Mean Accuracy	39
23	Cologne- Binary classification setting- invariance to shape-Kappa Statistics	39
24	Visualization of results from single realization for binary classification	
	setting for Cologne	41
25	Visualization of results from single realization for binary classification	
	setting for Cologne	42
26	Cologne- Multi-class classification setting- invariance to Scale- Mean	
	Accuracy	43
27	Cologne- Multi classification setting- invariance to scale- Kappa Statistics	43
28	Cologne- Multi-class classification setting- invariance to shape-Mean	
	Accuracy	44
29	Cologne- Multi-Class classification setting- invariance to shape-Kappa	
	Statistics	44
30	Visualization of results from single realization for multi-class classifica-	
	tion setting for Cologne	46

31	Visualization of results from single realization for multi-class classifica-	47
32	tion setting for Cologne	47
32	Accuracy	49
33	Hagadera- Binary- class classification setting- invariance to scale- Kappa Statistics	49
34	Hagadera- Binary- class classification setting- invariance to shape- Mean Accuracy	50
35	Hagadera- Binary- class classification setting- invariance to shape- Kappa Statistics	50
36	Visualization of results from single realization for binary classification setting for Hagadera	52
37	Visualization of results from single realization for binary classification setting for Hagadera	53
38	Hagadera- Multi-class classification setting- invariance to scale - Mean Accuracy	54
39	Hagadera- Multi-Class classification setting- invariance to scale - Kappa Statistics	54
40	Hagadera- Multi-class classification setting- invariance to shape - Mean Accuracy	55
41	Hagadera- Multi-Class classification setting- invariance to shape - Kappa Statistics	55
42	Visualization of results from single realization for multi-class classification setting for Hagadera	57
43	Visualization of results from single realization for multi-class classification setting for Hagadera	58
44	Visualization of uncertainty results on SVM-SL-U approach	60
45	Visualization of uncertainty results on VSVM-SL-U approach	61
46 47	Visualization of uncertainty results on VSVM-SL-VU approach Comparison of uncertainty visualizations for semi-supervised methods	62
	on Cologne data set	63

List of Tables

1	Cologne - Binary classification setting- Classification accuracies and	
	other configurations compared to benchmark semi-supervised method.	40
2	Cologne - Binary classification setting- Classification accuracies and	
	other configurations of semi-supervised methods	40
3	Cologne - Multi-class classification setting- Classification accuracies and	
	other configurations of benchmark semi-supervised method	45
4	Cologne - Multi-class classification setting- Classification accuracies and	
	other configurations of semi-supervised methods	45
5	Hagadera - Binary classification setting- Classification accuracies and	
	other configurations of benchmark semi-supervised method	51
6	Hagadera - Binary classification setting- Classification accuracies and	
	other configurations of semi-supervised methods	51
7	Hagadera - Multi-class classification setting- Classification accuracies	
	and other configurations of benchmark semi-supervised method	56
8	Hagadera - Multi-class classification setting- Classification accuracies	
	and other configurations of semi-supervised methods	56

List of Abbreviations

SVM Support Vector Machines

SV Support Vectors

VSVM Virtual Support Vector Machines

VSV Virtual Support Vectors

VSVM-SL Virtual Support Vector Machines with Self-Learning

QP Quadratic Programming

VC Vapnik-Chervonenkis

SRC Structural Risk Minimization

S3VM Semi-supervised Support Vector Machines

GIS Geographic Information Science

AL Active Learning

VSVM-SL-U Virtual Support Vector Machines with Self-Learning with Unlabeled Sam-

ples

VSVM-SL-VU Virtual Support Vector Machines with Self-Learning with Virtual Unla-

beled Samples

SVM-SL-U Support Vector Machines with Self-Learning with Unlabeled Samples

OA Overall Accuracy

AA Average Accuracy

1 Introduction

1.1 Motivation and problem statement

In the last decades, with the advent of high spatial and spectral resolution remote sensing data, land cover classification applications have become one of the main subjects in remote sensing (Lu et al., 2016). Consequently, it triggered the development of many methods to derive thematic classes from image data, and as an outcome, supervised methods became one of the most preferred classification approaches because of their robust and accurate information extraction properties (Geiß et al., 2019) Although it is overly challenging to determine the best method from numerous of existing approaches for a classification problem, Support Vector Machines (SVM) attracted attention regarding the classification of multispectral remote sensing images. As a working principle, SVM set suitable hyperplanes on different classes of labeled data and those samples are projected through a nonlinear transformation from input space to a higher-dimensional space. In that space, support vectors (SV), which are the samples closest to the separating surface, are determined in subject to the optimal hyperplane that maximizes the margin (Geiß et al., 2019; Burges, 1998). Therefore, SVM showed excellent performance due to their, (i) ability to manage high-dimensional feature space; (ii) relevant generalization properties (iii); the uniqueness of the solution (Tuia et al., 2009)

SVMs, as any other supervised method, rely on the quality of the labeled training data. However, this constrains the training set and requires extensive manual efforts regarding human-machine interaction. That is why active learning methods and semi-supervised learning approaches which use unlabeled samples will benefit the classification results, especially with respect to poorly sampled remote sensing applications (Izquierdo-Verdiguier et al., 2012). It is here where we combine self-learning constraints on Virtual Support Vector Machines (VSMV) with a semi-supervised approach to indicate useful information about the underlying data distribution which eventually achieves higher accuracies, especially with small amounts of training data.

1.2 Research identification

1.2.1 Research objectives

To eventually enhance the accuracy properties of the Virtual Support Vector Machines with the self-learning (VSVM-SL) method in settings with very few labeled samples, the goal is to deploy a constrained set of unlabeled samples for model learning and for very high spatial resolution multispectral remote sensing images. As a result, the training set which the model is learning will be enriched by informative unlabeled samples. Those are jointly evaluated and selected with respect to existing support vectors and virtual samples. In addition, the generation of spatial visualization for the uncertainty of results is done by checking the distance of SVMs hyperplane from the model and further monitoring how the uncertainty changes with the newly developed methods. The spatial visualization will be displayed as land cover classification maps

showing corresponding thematic uncertainty. Subsequently, these spatial visualizations will benefit the active learning process by providing human-machine interaction on relabeling uncertain samples in a prioritized way and will use those samples to relearn the model and eventually obtain higher accuracies.

RQ1: To what extent does the new Virtual Support Vector Machines with self-learning constraints on a semi-supervised scheme (VSVM-SL- Unlabeled Samples) method provide better classification accuracy with few labeled samples when compared to other/older methods such as SVM, VSVM, and VSVM-SL?

- Comparison analysis will be made between newly proposed semi-supervised methods to previous methods by comparing overall and average accuracies, kappa value, and F1 score.
- Line graphs will be used in order to see mean kappa values and overall accuracies
 of the methods.

RQ2: Does visualizing the uncertainties of the models improve human-monitored active learning approaches on relabeling uncertain samples?

- Model quantifies the certainty of unlabeled samples by checking the distance of SVM hyperplane and shows which land cover classes they belong to. Consequently, this helps the user to label those uncertain samples and bring them back to model.
- Therefore, a case study will be applied in order to assess the effects and overall
 performance of relabeling with visualization of uncertainty on active learning. Accuracy results of newly developed methods plus the uncertainty visualizations will
 be compared to the results of newly developed methods without the uncertainty
 visualizations.

1.3 Innovation aimed at

The innovation of the research aims at developing a semi-supervised classification method based on a self-learning strategy. This will provide results with higher accuracy on sparsely sampled remote sensing imageries and will be adaptable in the future to the classification of hyperspectral data. These innovations are aimed at bringing a new outlook with the extension and combination of methods on remote sensing and the cartography fields.

This is an excerpt from the thesis "Semi-Supervised Virtual Support Vector Machines with Self-Learning Constraint for Remote Sensing Image Classification".

Please contact Ozan Tuncbilek for a full version of the thesis.

Extended works on this topic are documented in: Geiß, C., Aravena Pelizari, P., Tuncbilek, O., and Taubenböck, H. (2023) Semi-supervised Learning with Constrained Virtual Support Vector Machines for Classification of Remote Sensing Image Data. International Journal of Applied Earth Observation and Geoinformation, 125, 103571.

In future works, the same methodology can be applied for classifying hyperspectral data within an adequate processing framework. Additionally, an active learning approach with uncertainty visualizations can be integrated into all settings of the data sets. Moreover, this approach can be further adapted to the supervised methods as well. At last, a combination of the semi-supervised methods and active learning approach can be integrated better in a collaborative learning scheme.

6 References

Bastin, L., P. F. Fisher, and J. Wood

2002. Visualizing uncertainty in multi-spectral remotely sensed imagery. *Computers* & *Geosciences*, 28(3):337–350.

Bennett, K., A. Demiriz, et al.

1999. Semi-supervised support vector machines. *Advances in Neural Information processing systems*, Pp. 368–374.

Blaschke, T.

2010. Object based image analysis for remote sensing. *ISPRS journal of photogrammetry and remote sensing*, 65(1):2–16.

Bruzzone, L. and L. Carlin

2006. A multilevel context-based system for classification of very high spatial resolution images. *IEEE transactions on Geoscience and Remote Sensing*, 44(9):2587–2600.

Bruzzone, L., M. Chi, and M. Marconcini

2006. A novel transductive sym for semisupervised classification of remote-sensing images. *IEEE Transactions on Geoscience and Remote Sensing*, 44(11):3363–3373.

Burges, C. J.

1998. A tutorial on support vector machines for pattern recognition. *Data mining and knowledge discovery*, 2(2):121–167.

Cortes, C. and V. Vapnik

1995. Support-vector networks. *Machine learning*, 20(3):273–297.

Demir, B., C. Persello, and L. Bruzzone

2010. Batch-mode active-learning methods for the interactive classification of remote sensing images. *IEEE Transactions on Geoscience and Remote Sensing*, 49(3):1014–1031.

Ding, S., Z. Zhu, and X. Zhang

2017. An overview on semi-supervised support vector machine. *Neural Computing and Applications*, 28(5):969–978.

Dópido, I., J. Li, P. R. Marpu, A. Plaza, J. M. B. Dias, and J. A. Benediktsson 2013. Semisupervised self-learning for hyperspectral image classification. *IEEE transactions on geoscience and remote sensing*, 51(7):4032–4044.

Fernández-Delgado, M., E. Cernadas, S. Barro, and D. Amorim 2014. Do we need hundreds of classifiers to solve real world classification problems? *The journal of machine learning research*, 15(1):3133–3181.

Foody, G. M.

2009. On training and evaluation of sym for remote sensing applications. In *Kernel methods for remote sensing data analysis*, Pp. 85–109. Wiley.

Foody, G. M. and A. Mathur

2004. A relative evaluation of multiclass image classification by support vector machines. *IEEE Transactions on geoscience and remote sensing*, 42(6):1335–1343.

Geiß, C., P. A. Pelizari, L. Blickensdörfer, and H. Taubenböck

2019. Virtual support vector machines with self-learning strategy for classification of multispectral remote sensing imagery. *ISPRS journal of photogrammetry and remote sensing*, 151:42–58.

Geiß, C., P. A. Pelizari, H. Schrade, A. Brenning, and H. Taubenböck

2017. On the effect of spatially non-disjoint training and test samples on estimated model generalization capabilities in supervised classification with spatial features. *IEEE Geoscience and Remote Sensing Letters*, 14(11):2008–2012.

Goodchild, M. F.

2008. Imprecision and spatial uncertainty.

Hughes, G.

1968. On the mean accuracy of statistical pattern recognizers. *IEEE transactions on information theory*, 14(1):55–63.

Izquierdo-Verdiguier, E., V. Laparra, L. Gómez-Chova, and G. Camps-Valls

2012. Encoding invariances in remote sensing image classification with svm. *IEEE Geoscience and Remote Sensing Letters*, 10(5):981–985.

Kinkeldey, C.

2014. A concept for uncertainty-aware analysis of land cover change using geovisual analytics. *ISPRS International Journal of Geo-Information*, 3(3):1122–1138.

Kinkeldey, C., J. Mason, A. Klippel, and J. Schiewe

2014. Evaluation of noise annotation lines: using noise to represent thematic uncertainty in maps. *Cartography and Geographic Information Science*, 41(5):430–439.

Li, Y.-F. and Z.-H. Zhou

2014. Towards making unlabeled data never hurt. *IEEE transactions on pattern analysis and machine intelligence*, 37(1):175–188.

Lu, X., J. Zhang, T. Li, and Y. Zhang

2016. A novel synergetic classification approach for hyperspectral and panchromatic images based on self-learning. *IEEE Transactions on Geoscience and Remote Sensing*, 54(8):4917–4928.

Lucieer, A.

2004. *Uncertainties in Segmentation and their Visualisation*. PhD thesis, Utrecht University and International Institute for Geo-Information Science

Melgani, F. and L. Bruzzone

2004. Classification of hyperspectral remote sensing images with support vector machines. *IEEE Transactions on geoscience and remote sensing*, 42(8):1778–1790.

Mountrakis, G., J. Im, and C. Ogole

2011. Support vector machines in remote sensing: A review. *ISPRS Journal of Photogrammetry and Remote Sensing*, 66(3):247–259.

Murphy, C. E.

2015. Intellectual highlighting of remote sensing imagery for better image map design. In *Proc. 27th Int. Cartographic Conf.*

Pan, C., J. Li, Y. Wang, and X. Gao

2018. Collaborative learning for hyperspectral image classification. *Neurocomputing*, 275:2512–2524.

Song, X., Z. Duan, and X. Jiang

2012. Comparison of artificial neural networks and support vector machine classifiers for land cover classification in northern china using a spot-5 hrg image. *International Journal of Remote Sensing*, 33:3301–3320.

Tuia, D. and J. Munoz-Mari

2012. Learning user's confidence for active learning. *IEEE Transactions on Geoscience and Remote Sensing*, 51(2):872–880.

Tuia, D., F. Ratle, F. Pacifici, M. F. Kanevski, and W. J. Emery

2009. Active learning methods for remote sensing image classification. *IEEE Transactions on Geoscience and Remote Sensing*, 47(7):2218–2232.

Tuia, D., M. Volpi, L. Copa, M. Kanevski, and J. Munoz-Mari

2011. A survey of active learning algorithms for supervised remote sensing image classification. *IEEE Journal of Selected Topics in Signal Processing*, 5(3):606–617.

Ul Hag, Q. S., L. Tao, F. Sun, and S. Yang

2011. A fast and robust sparse approach for hyperspectral data classification using a few labeled samples. *IEEE Transactions on Geoscience and Remote Sensing*, 50(6):2287–2302.

Volpi, M., D. Tuia, F. Bovolo, M. Kanevski, and L. Bruzzone

2013. Supervised change detection in vhr images using contextual information and support vector machines. *International Journal of Applied Earth Observation and Geoinformation*, 20:77–85.

Wang, X. and P. M. Pardalos

2014. A survey of support vector machines with uncertainties. *Annals of Data Science*, 1(3-4):293–309.