ADAPTIVE SAMPLING AND OBJECT FEATURE SELECTION FOR LANDSLIDE MAPPING USING RANDOM FORESTS

A. Stumpf [a, b, c,] *, N. Lachiche [d], N. Kerle [c], A. Puissant [a]

[a] Laboratoire Image, Ville, Environnement, CNRS ERL 7230, Université de Strasbourg, 3 rue de l'Argonne, F- 67083 Strasbourg, France
[b] Institut de Physique du Globe de Strasbourg, CNRS UMR 7516, Université de Strasbourg / EOST, 5 rue René Descartes, F-67084 Strasbourg, France
[c] ITC-Faculty of Geo-Information Science and Earth Observation, University of Twente, Department of Earth Systems Analysis, Hengelosestraat 99, P.O. Box 6, Enschede, 7500 AA, The Netherlands
[d] Image Sciences, Computer Sciences and Remote Sensing Laboratory, CNRS UMR 7005, Université de Strasbourg, Bd Sébastien Brant - BP 10413, F-67412 Illkirch, France

andre.stumpf[at]unistra.fr,
nicolas.lachiche[at]lsiit-cnrs.unistra.fr,
kerle[at]itc.nl,
anne.puissant[at]live-cnrs.unistra.fr

KEY WORDS: feature selection, object-oriented analysis, landslide mapping, active learning, class-imbalance

ABSTRACT: The selection of relevant object-features and cost-efficient sampling of training data are frequently encountered issues when applying supervised classification techniques in object-oriented image analysis. Additionally, when one of the targeted classes is rather sparse the classifications maybe considerably biased toward the majority class, and the class-imbalance may hinder the generation of accurate maps. This study targeted to address those issues in the context of landslide inventory mapping from very-high resolution (VHR) satellite images. To reduce the costs of sampling an active learning routine based on the Random Forest classifier was developed to focus the sampling efforts on few but relevant spatial subsets in the analysed scene. Subsequently backward feature selection was used to identify relevant features and reduce the complexity of the model without sacrificing accuracy, and finally iterative stratified bootstrap sampling is proposed to compensate the spurious effects of class-imbalance. The proposed approach was tested on two datasets comprising mono-temporal and bi-temporal VHR images of sites recently affected by large landslide events. Approximately balanced user's and producer's accuracies of up to 78% were achieved adopting only 6-10% of the data for training.

1. INTRODUCTION

Object-oriented image analysis has proven to be an efficient tool for image classification with many advantages among traditional pixel-based methods, especially when dealing with VHR remote sensing imagery. Much of the gains in accuracy must be attributed to an enriched feature space resulting from image segmentation, where a vast number of additional features to characterize texture, shape or contextual relationships of objects become available to potentially increase class separability.
However, for both rule-based classification and the training of supervised algorithms it remains crucial to identify truly relevant features, and disregard insignificant attributes that may deteriorate the classification. This task is commonly known as feature selection. It constitutes a fundamental step in most object-oriented studies, and a tedious one with subjective results if performed manually.
A variety of automated feature selection techniques have been developed and tested successfully for general machine learning problems (Guyon and Elisseeff, 2003) and in neighbouring fields of remote sensing (e.g. Serpico and Bruzzone, 2001; Benediktsson et al., 2003; Pal and Foody, 2010), but to date only few studies have addressed the use of such techniques for object-oriented analysis (Laliberte et al.; Nussbaum et al., 2006; Van Coillie et al., 2007; Laliberte and Rango, 2009; Stumpf and Kerle, 2011; Xi et al., 2011).
Besides the choice of optimal features, the quantity and quality of the adopted training samples is a fundamental factor for the performance of supervised classification algorithms. Since the collection/labelling of samples typically comprises considerable costs it is general desirable to collect few but representative training samples. Several recent pixel-based studies have demonstrated that active learning algorithms are capable to focus the search on interesting samples and thereby help to reduce the size of required training set without sacrificing the accuracy of remote sensing products (Tuia et al., 2011). On the other hand limited research has been dedicated to the integration of active learning and object-oriented studies (Michel et al., 2010) and with few exceptions (Liu et al., 2009; Pasolli et al., 2011) the significance of the spatial domain is generally ignored in development active learning heuristics.
For many applications it is furthermore desirable to, obtain equally good classification results for the different classes involved, or gain at least some control on the error distributions where costs for misclassification can be defined. This aspect becomes especially relevant when processing datasets with imbalanced class distribution (He and Garcia, 2009).
Considering such issues this study investigated different strategies to optimize the choice of features and training samples in the context of landslide inventory mapping. The main objective is thereby to support the elaboration of landslide inventory maps from VHR satellite and aerial images with a supervised object-oriented image analysis workflow that reduces the time and effort of manual image interpretation by

* corresponding author

experts. The developed workflow comprises image segmentation and object-feature extraction techniques commonly used for object-oriented image analysis, and makes use of feature selection and adaptive sampling strategies based on the ensemble decision tree algorithm commonly known as Random Forest (RF, Breiman, 2001). Two different problem settings are illustrated exemplarily with two datasets of VHR satellite images picturing sites in Italy and Brazil that have recently been affected by predominately flow type landslides triggered during heavy rainfall events.

2. METHODS AND DATA

2.1 Study Sites and Data

Test site A (~10 km^2) is located in the Serrana Mountains around Nova Friburgo (Brazil), which were affected by heavy thunderstorms on 11th and 12th of January 2011. Cumulative rainfalls in the area reached peak values of 200 mm in 24h and triggered thousands of debris flows. The event claimed more than 1.500 victims and caused severe damages to houses and infrastructure (Avelar et al., 2011). Geoeye-1 imagery was acquired on the 20 January 2011 and pre-event imagery from the same sensor was available for the 26 May 2010. The bi-temporal dataset was incorporated in this study together with the corresponding subset of the globally available Aster GDEM (ASTER-GDEM-VALIDATION-TEAM, 2011) at 30 m resolution.
Test site B (1 km^2) is located in the Peloritani Mountains that rise from sea level to about 700 m few kilometres south of the city of Messina (Italy). After a series of prolonged precipitation extraordinarily intense rainfall affected several catchments on the 1st of October 2009 and triggered a series of debris flows that killed 31 people and caused a direct economic loss of approximately US$ 825 million (Civil-Protection-Sicily, 2010). Quickbird imagery was recorded 7 days after the event and adopted in this study together with a post-landslide digital elevation model with 10 m resolution.
Those two test data sets resemble typical scenarios, where only post-event imagery or both pre- and post-event imagery may be available. Reference landslide inventories based on expert image interpretation and field surveys served for the labelling of training samples and where established as the reference to assess the accuracy assessment.

2.2 Image segmentation and feature extraction

The multi-resolution image segmentation algorithm (MRIS, Baatz and Schäpe, 2000) implemented in eCognition® 8.64 was adopted considering only spectral information of the post-event images and equal weights of all spectral bands. At test site B the scale factor of the region-growing algorithm, which is a threshold for the maximum allowed increase in the segment's variance, was set to 10. With reference to the landslides this corresponds to a strong over-segmentation which was found to typically yield higher accuracies in a supervised framework than a coarser segmentation (Stumpf and Kerle, 2011). The set of object features for the test site B comprised 96 features including spectral variables, texture measures, shape measures and terrain variables frequently mentioned as relevant for landslide mapping in the literature. In addition to traditional derivatives from Haralick's grey-level co-occurrence matrix (GLCM) this comprised topographically guided textures derivate with an increased sensitivity to scouring traces typical for landslides affected surfaces (Stumpf and Kerle, 2011).

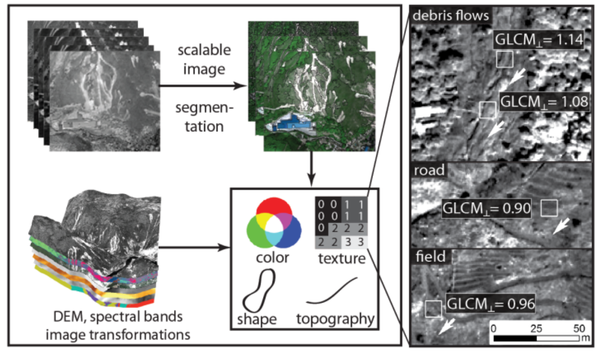

Fig. 1: Initial steps of the workflow including image segmentation and the extraction of a large set of potentially useful features such as topographically guided texture measures (GLCM⊥).

For the analysis of the images at test site A and over-segmentation was performed at a scale factor of 20. The feature set was further expanded to 105 attributes using additionally spectral variables extracted from the pre-event image, a topographic variable capturing the distance to the most proximate ridge crest (Tesfa et al., 2011), and neighbourhood relationships describing the gradient of spectral values between neighbouring objects.

2.3 Adaptive sampling and feature selection

The proposed strategy comprises three aspects. First, an iterative active learning routine is used to guide the user inputs toward subsets of the scene where the classification remains highly uncertain. This step targets to reduce the labelling costs and create a relatively small relevant training set. Second, a feature selection algorithm (Diaz-Uriarte, 2010) is employed to disregard non-informative features and obtain an overview of the relevance of different features. Third, in order to achieve a balance between user's and producer's accuracy the stratified bootstrap sampling is performed iteratively with changing class ratios.
In this study Random Forest was chosen as the basic learning algorithm because of its relatively fast and generally accurate performance on large and noisy datasets. The RF algorithm constructs an ensemble of classification trees from bootstrap samples of the original training data and assigns a label to new unknown samples based on the majority vote of trees within the ensemble.
Using the disagreement among ensembles members as a uncertainty measure and query the label of the most uncertain samples is a commonly used active learning strategy known as Query-by-committee (QBC, Seung et al., 1992). The uncertainty can be measured as the entropy H (Shannon, 1948) of the ensemble votes and for a binary decision is calculated by Eq.1:

$$H = -\left(\frac{v_1}{n} log \frac{v_1}{n} + \frac{v_2}{n} log \frac{v_2}{n}\right) \qquad \text{Eq.1}$$

Where v_1 is the number of votes for the first class, v_2 the number of votes for the second class and n corresponds to the total number of trees in the ensemble. Since the classifier re-training after each query is often the computational bottleneck of such approaches it is typically more efficient to query a batch of training samples after each iteration, and related RF-based approaches rank among the most accurate state-of-the-art active learning techniques (Borisov et al., 2011). For remote sensing application it is however important to consider that the labelling costs are not necessarily related to the number of samples but rather to the number of localities that need to be assessed (during field surveys or via image interpretation). For practical applications it seems therefore more relevant to identify compact sampling areas with high expected utility or in other words relevant batches within constrained spatial neighbourhoods. The proposed approach makes therefore use of a sliding window to assess the vote entropy of local neighbourhoods S_r within a defined radius *r*.

To initiate the routine one centre point from all objects belonging to class landslide is sampled randomly assuring that the first sampling area (S_{r1}) contains at least one positive training sample (Fig. 2b, d). The trained RF subsequently votes on the class membership of all unlabelled objects and a uncertainty map is generated calculating the mean Entropy μ_H in a sliding window with the radius *r*. Subsequently all objects located the neighbourhood with the highest μ_H ($S_{rMax}\mu_H$) are labelled and added to the training set to retrain the RF. Since it is a well know issue that uncertainty-based sampling is prone to query outliers and samples with redundant information (Brinker, 2003; Borisov et al., 2011; Demir et al., 2011) a second similar strategy was tested querying regions with the highest standard deviation of the vote entropy σ_H in order to encourage greater diversity within the batches. The radius *r* was set to 100 m for the large test site A (10 km^2) and to 40 m for the smaller test site B (1 km^2). Buffers around each sampling area *S* were used to avoid spatial overlap (Fig. 3a, c). Both active learning heuristics were executed over 20 iterations and repeated 10 times with different initial random seeds to assess the stability of the classification accuracies on the respective remaining test sets. The resulting learning curves were compared with the performance of spatial coverage sampling (SPCOSA, Walvoort et al., 2010), using circular sampling areas with an equal radius *r* at the centroid of 20 spatial clusters (Fig. 2a, c). Similarly SPCOSA was repeated 10 times to assess the stability of the results. All experiments were conducted using RFs with 500 trees and labels were obtained according to the same inventories elaborated by the experts.

With each bootstrap sample generated during the construction of a RF approximately one third of the cases remain unconsidered for training and can be used to aproximate the generalization or out-of-bag (OOB) error (Breiman, 2001).

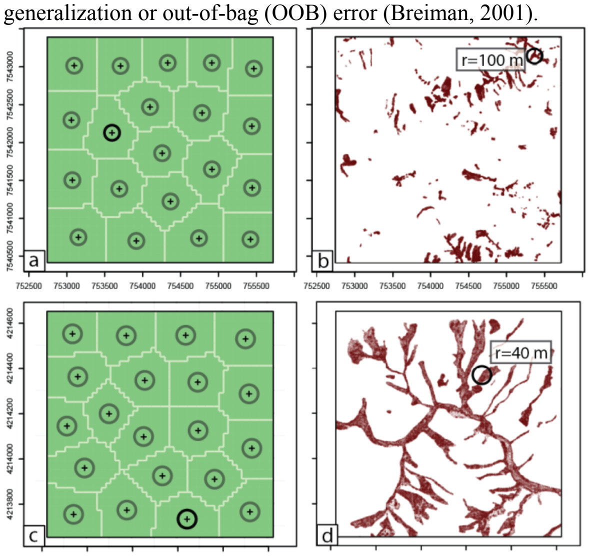

Fig. 2: Initial batches of 20 sampling areas for (a, b) spatial coverage sampling and (c, d) active learning routine at test site A and B, respectively.

Furthermore, when one variable is randomly permuted the corresponding increase in the OOB error provides an intrinsic measure for variable importance. The final training set resulting after 20 iterations was used for a backward feature selection based on the variable importance as described in Diaz-Uriarte and Alvarez de Andres (2006). After an initial ranking of all features, RF models are retrained repeatedly dropping each time a fraction of 20%, and the selection the final feature according to the RF model with the lowest OOB error.

Since landslides typically cover only minor fractions of the landscape class-imbalance is an intrinsic issue that generally needs special attention when applying supervised learning algorithms (He and Garcia, 2009). A previous study demonstrated the utility of an iterative resampling scheme to estimate class-ratios for the training sample yielding an approximate balance between user's and producer's accuracies on the test set (Stumpf and Kerle, 2011). However, this approach still comprises two limiting factors. First, the routine splits the training sample into further subsets. Those may become very small if the original training set already includes rather few samples and consequently yield unstable results. Second, in the final step the approach drops a fraction of samples from the majority class and thereby leaves some costly and potentially informative labelled samples unconsidered. A possible enhancement to resolve these issues is to artificially bias the class ratio of the bootstrap samples and estimate user's and producer's accuracies directly on the corresponding out-of-bad samples. To this end the bootstrap sampling is performed stratified for the minority and the majority class sampling n_{mi} (overall number of cases in the training sample belonging to the minority class) samples with replacement the from the minority class (landslides) and a number of cases n_{ma} from the majority class (non-landslides). Starting form $n_{ma} = n_{mi}$ the ratio $\beta = n_{ma} / n_{mi}$ is stepwise increased until the original class-balance of the training sample is reached. In each step the user's and producer's accuracies are assessed on the OOB sample and the final ratio β_n is determined as the class-ratio that yields the best balance. The approach was tested using in each iteration relatively small forests composed by 50 trees and increasing β by steps of 0.1. A final RF was built employing the stratified bootstrap sampling with the determined β_n to construct 500 trees and assess the accuracy on the test set.

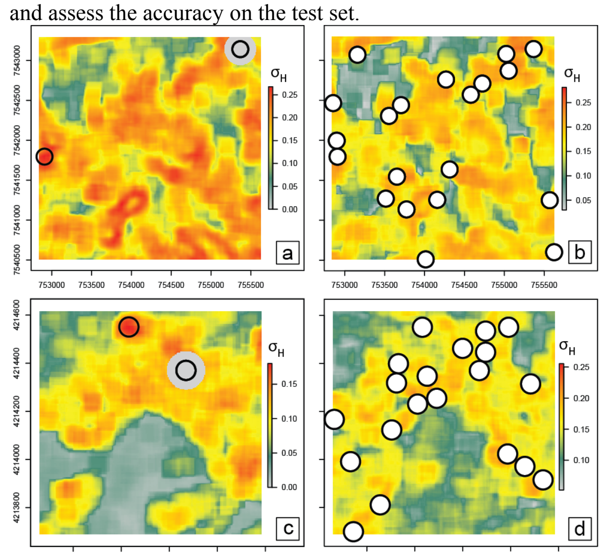

Fig. 3: (a, c) First and (b, d) last iteration of the QBC routine at test site A and B, respectively. The standard deviation of the vote entropy σ_H was used as query criterion.

3. RESULTS AND DISCUSSION

Fig. 3 illustrates exemplarily iterations obtained using the QBC sampling based on the standard deviation of the vote entropy. Generally, it was observed that areas affected by landslides obtained relatively high vote entropies and were therefore favoured when using μ_H as query criterion. This applied to some degree also when using σ_H, whereas landslide boundaries and heterogeneous areas such as human settlements gained greater importance. This positive bias toward the minority class resulted in a class ratio of the training sample that was generally more balanced than the underlying distributions (test site A: 16.4, test site B: β= 4.3) and the class ratios obtained via SPCOSA (Fig. 6a, c).

Comparing the learning curves obtained with the different sampling techniques (Fig. 4) it can be seen that the QBC sampling strategies indeed lead to a generally stronger enhancement of the overall accuracies (F-measure) with the size of the training area, and a better convergence between user's and producer's accuracies. The latter must be partially attributed to the more balanced training samples resulting from QBC. However, while the use of μ_H as a query criterion yields more balanced training samples than using σ_H (Fig. 4a, c), the latter still leads to better convergence of user's and producer's accuracies (Fig. 4c, f). This indicates that not only the number of samples per class but also their respective positions in feature space may influence the classification bias towards the majority class.

Besides performance differences among the tested sampling techniques a strong contrast between the accuracies at the test site A and B was observed. The lower accuracies at test site B thereby result from the fact that pre-event imagery was not integrated but also from the more complex scene characteristics where many of the landslides occurred in narrow channels and /or left the vegetation partially intact. On this rather noisy dataset the QBC sampling based on μ_H largely failed to query relevant batches (Fig. 4e), whereas σ_H-based QBC demonstrates a generally more robust performance (Fig. 4c, f). Training sets obtained via σ_H-based QBC sampling were consequently used for the subsequent steps of feature selection and accuracy balancing.

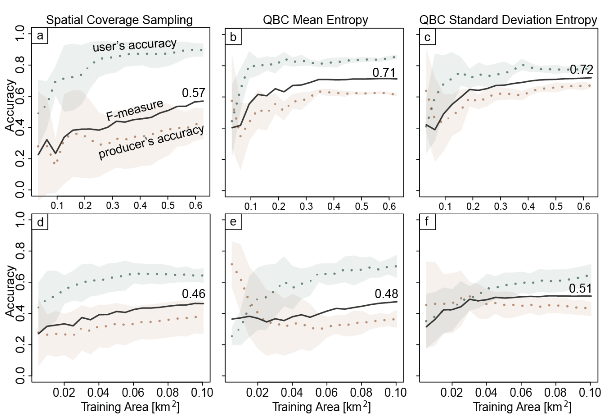

Fig. 4: Average learning curve obtained from 10 repeated runs at (a, b, c) test site A and (d, e f) test site B. The standard deviations of user's and producer's accuracies are indicated by semi-transparent bounds.

Fig. 5 displays the results of the RF-based backward feature selection with 69 out of 105 features selected at test site A, and 50 out 96 selected features at test site B. The top-ranked features at test site B (Fig. 5d) are largely consistent with the results obtained in an earlier study (Stumpf and Kerle, 2011) and illustrate the benefits of integrating spectral information with topographically texture measures and topographic information. Also at test site A the enhanced texture measures ranked among the most important variables and additional features resulting from the integration of pre-event imagery, neighbourhood relationships and the distance to the most proximate ridge proofed to enhance the classification accuracies. In this context feature ranking and selection not only provides the possibility to reduce the model complexity but also to identify generically relevant object features for further studies and future operational applications.

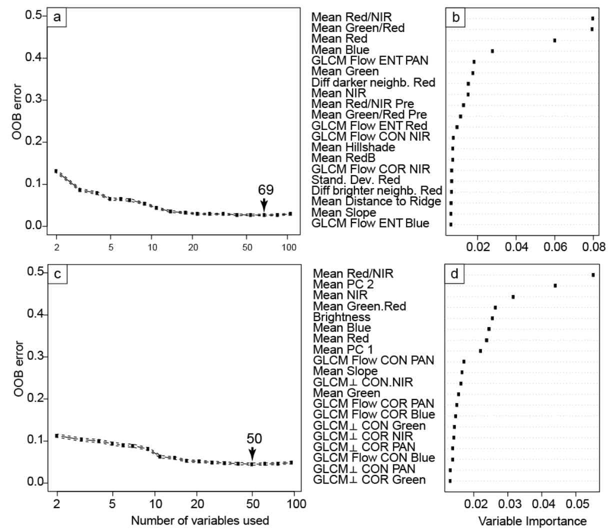

Fig. 5: (a, c) Results of the backward feature selection after σ_H-based QBC sampling and the (b, d) 20 highest ranked features at the test sites A and B, respectively.

The training sets resulting from σ_H-based QBC and the corresponding reduced feature set were introduced in the last step of the analysis to compensate the remaining classification bias toward the majority class.

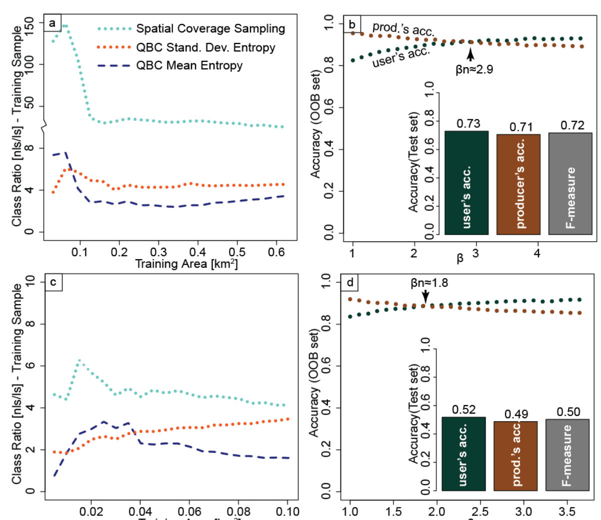

Fig. 6: (a, c) Development of the class ratio in the training sample over 20 iterations of the different sampling techniques and (b, d) results of the routine to balance user's and producer's accuracy for the two test sites, respectively.

Fig. 6 b and d display the OOB accuracy curves resulting from atrifically biasing the class ratios in the training sample via the stratified bootstrapping described above. The approach results in smooth accuracy curves with a clear crossing marking the estimate of *βn*. The barplots in Fig. 6 b and d also show the final accuracies on the test set obtained when using the estimate *βn* for stratified bootstrapping of a full RF with 500 trees. It can be seen that the the estimated *βn* provides indeed an aproximate balance of user's and producer's accuracies on the test set, and that the selected feature set provides equally good accuracies as the full initilally used feature set (Fig. 4c, f).

The final maps and corresponding accuracy measure integrate test set predictions and training data and are displayed in Fig. 7.

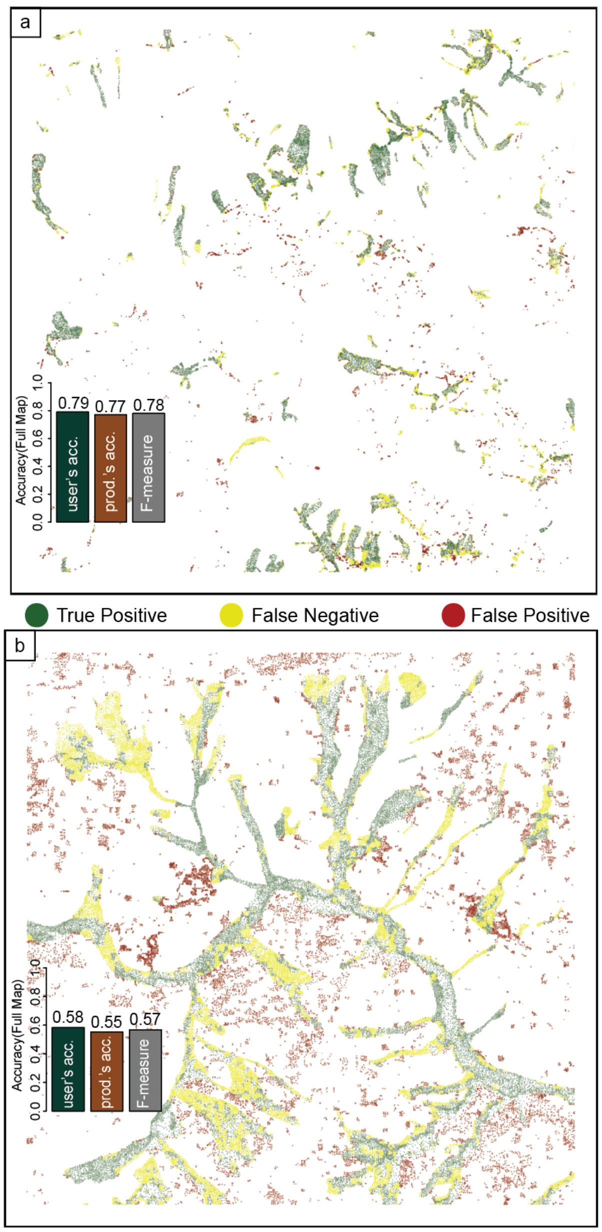

Fig. 7: Final classified maps (center points of objects) and accuracy measures resulting after σ_H-QBC sampling, feature selection and accuracy balancing at (a) test site A and (b) test site B.

At the test site A aproximately 6% of area were made available for training and resulted in nearly balanced accuracies between 77% and 79% (Fig. 7a). The final training set at test B comprised aproximately 10% of the area and resulted in nearly balanced but significantly lower accuracies between 55% and 58% (Fig. 7b). Although QBC sampling proofed to provide enhancements over spatial coverage sampling on both datasets, test site B still constitutes a challenging scenario. Complex scene characteristics and significant class-overlap resulting from small debris flows that were mapped during field surveys but are barely visible through the satellite images contribute to relatively low performance of the method. Under such conditions the use of post-event imagery alone appears insufficient for an accurate mapping and also a further integration of neighborhood relationships and topographic variables as used at test site A seems desireable. Interestingly, many false positives show a rather disperse distribution indicating that an adittional morphological filtering might be usefull to enhace the accuracy and visual appearance of the final maps.

4. CONCLUSIONS

The cost-efficient collection of training data, the selection of relevant features, and class-imbalance are issues frequently encountered in the application of supervised machine learning techniques and object-oriented analysis. This study adopted the Random Forest framework to address those three aspects in the context of landslide inventory mapping from VHR satellite images. It was demonstrated that an active learning heuristic (σ_H-based QBC) can yield significantly improvements above the performance of spatial coverage sampling. Unlike most previously developed active learning heuristics the proposed technique works explicitly in the spatial domain by selecting sample batches in spatial neighbourhoods with a high variance of the vote-entropy.

It was demonstrated that backward feature selection is a feasible instrument to identify the most relevant features for subsequent studies and future operational applications. Furthermore, a resampling scheme was developed to compensate class-imbalance and proofed useful to balance user's and producer's accuracies. Integrating QBC sampling, feature selection and the iterative routine to compensate class-imbalance it was possible to achieve map accuracies of up to 78% on a bi-temporal dataset using only 6% of the data for training. On the other hand, the analysis of a mono-temporal dataset with more challenging scene characteristics also revealed that further methodological enhancements are still desirable. Since historical imagery with coarser resolution is typically available at low costs further studies should target the integration of multi-modal sensor data and enhanced object features sets. The feature selection framework may thereby proof useful to quantify the relevance of further potentially useful object-features.

Focusing on areas with high standard deviations of the vote entropy the proposed active learning routine indirectly encourages diversity within the batch, whereas a more explicit integration of a criterion for sample diversity (e.g. Demir et al., 2011) could yield further improvements. For remote sensing applications batch diversity and batch size seem to be related with aspects such as spatial auto-correlation and the operating radius of users (during image interpretation or field surveys), and call for a closer integration of active learning techniques with geographic object-oriented analysis.

References

ASTER-GDEM-VALIDATION-TEAM, 2011. ASTER Global Digital Elevation Model Version 2 – Summary of Validation Results, METI/ERSDAC, NASA/LPDAAC, USGS/EROS.

Avelar, A.S., Netto, A.L.C., Lacerda, W.A., Becker, L.B., Mendonça, M.B., 2011. Mechanisms of the recent catastrophic landslides in the mountainous range of Rio de Janeiro, Brazil, Second World Landslide Forum, Rome, Italy, pp. 5.

Baatz, M., Schäpe, A., 2000. Multiresolution Segmentation – an optimization approach for high quality multi-scale image segmentation. In: J. Strobl, T. Blaschke, G. Griesebner (Eds.), Angewandte Geographische Informationsverarbeitung XII. Wichmann, Heidelberg, Salzburg, pp. 12-23.

Benediktsson, J.A., Pesaresi, M., Amason, K., 2003. Classification and feature extraction for remote sensing images from urban areas based on morphological transformations. Geoscience and Remote Sensing, IEEE Transactions on, 41(9), 1940-1949.

Borisov, A., Tuv, E., Runger, G., 2011. Active Batch Learning with Stochastic Query-by-Forest (SQBF), JMLR Workshop on Active Learning and Experimental Design, pp. 59-69.

Breiman, L., 2001. Random Forests. Machine Learning, 45(1), 5-32.

Brinker, K., 2003. Incorporating diversity in active learning with support vector machines, International Conference on Machine Learning Washington, DC, pp. 59–66.

Civil-Protection-Sicily, 2010. Landslide and mud food emergency Messina province, Itlay, October 1st 2009, Regione Siciliana - Presidenza, Dipartimento dellla Protezione Civile.

Demir, B., Persello, C., Bruzzone, L., 2011. Batch-Mode Active-Learning Methods for the Interactive Classification of Remote Sensing Images. Geoscience and Remote Sensing, IEEE Transactions on, 49(3), 1014-1031.

Diaz-Uriarte, R., 2010. varSelRF: Variable selection using random forests. R package version 0.7-3.

Diaz-Uriarte, R., Alvarez de Andres, S., 2006. Gene selection and classification of microarray data using random forest. BMC Bioinformatics, 7(1), 3.

Guyon, I., Elisseeff, A., 2003. An Introduction to Variable and Feature Selection. Journal of Machine Learning Research, 3, 1157-1182.

He, H., Garcia, E.A., 2009. Learning from imbalanced data. IEEE Transactions on Knowledge and Data Engineering, 21(9), 1263-1284.

Laliberte, A., Rango, A., 2009. Texture and scale in object-based analysis of subdecimeter resolution unmanned aerial vehicle (UAV) imagery. IEEE Transactions on Geoscience and Remote Sensing, 47(3), 761-770.

Laliberte, A.S., Browning, D.M., Rango, A., A comparison of three feature selection methods for object-based classification of sub-decimeter resolution UltraCam-L imagery. International Journal of Applied Earth Observation and Geoinformation(0).

Liu, A., Jun, G., Gho, J., 2009. Spatially Cost-sensitive Active Learning, Ninth SIAM International Conference on Data Mining Sparks, Nevada, pp. 814-825.

Michel, J., Malik, J., Inglada, J., 2010. Lazy yet efficient land-cover map generation for HR optical images, Geoscience and Remote Sensing Symposium (IGARSS), 2010 IEEE International, pp. 1863-1866.

Nussbaum, S., Niemeyer, I., Canty, M.L., 2006. SEaTH – A new tool for automated feature extraction in the context of object-based image analysis., 1st International Conference on Object-based Image Analysis, Salzburg, Austria.

Pal, M., Foody, G.M., 2010. Feature Selection for Classification of Hyperspectral Data by SVM. Geoscience and Remote Sensing, IEEE Transactions on, 48(5), 2297-2307.

Pasolli, E., Melgani, F., Tuia, D., Pacifici, F., Emery, W.J., 2011. Improving active learning methods using spatial information, Geoscience and Remote Sensing Symposium (IGARSS), 2011 IEEE International, pp. 3923-3926.

Serpico, S.B., Bruzzone, L., 2001. A new search algorithm for feature selection in hyperspectral remote sensing images. Geoscience and Remote Sensing, IEEE Transactions on, 39(7), 1360-1367.

Seung, H.S., Opper, M., Sompolinsky, H., 1992. Query by committee, Fifth Workshop on Computational Learning Theory. Morgan Kaufmann, San Mateo, CA, pp. 287–294.

Shannon, C.E., 1948. A Mathematical Theory of Communication. Reprinted with corrections from The Bell System Technical Journal, 27(379–423, 623–656).

Stumpf, A., Kerle, N., 2011. Object-oriented mapping of landslides using Random Forests. Remote Sensing of Environment, 115(10), 2564-2577.

Tesfa, T.K., Tarboton, D.G., Watson, D.W., Schreuders, K.A.T., Baker, M.E., Wallace, R.M., 2011. Extraction of hydrological proximity measures from DEMs using parallel processing. Environmental Modelling & Software, 26(12), 1696-1709.

Tuia, D., Volpi, M., Copa, L., Kanevski, M., Munoz-Mari, J., 2011. A survey of active learning algorithms for supervised remote sensing image classification. Selected Topics in Signal Processing, IEEE Journal of, PP(99), 1-1.

Van Coillie, F.M.B., Verbeke, L.P.C., De Wulf, R.R., 2007. Feature selection by genetic algorithms in object-based classification of IKONOS imagery for forest mapping in Flanders, Belgium. Remote Sensing of Environment, 110(4), 476-487.

Walvoort, D.J.J., Brus, D.J., de Gruijter, J.J., 2010. An R package for spatial coverage sampling and random sampling from compact geographical strata by k-means. Computers & Geosciences, 36(10), 1261-1267.

Xi, C., Tao, F., Hong, H., Deren, L., 2011. Graph-Based Feature Selection for Object-Oriented Classification in VHR Airborne Imagery. Geoscience and Remote Sensing, IEEE Transactions on, 49(1), 353-365.