

MULTITEMPORAL CROP TYPE CLASSIFICATION USING CONDITIONAL RANDOM FIELDS AND RAPIDEYE DATA

T. Hoberg*, S. Müller

IPI - Institute of Photogrammetry and GeoInformation, Leibniz Universität Hannover, Germany
(hoberg, mueller)@ipi.uni-hannover.de

KEY WORDS: Crop, Agriculture, Classification, Multitemporal, Satellite

ABSTRACT:

The task of crop type classification with multitemporal imagery is nowadays often done applying classifiers that are originally developed for single images like support vector machines (SVM). These approaches do not model temporal dependencies in an explicit way. Existing approaches that make use of temporal dependencies are in most cases quite simple and based on rules. Approaches that integrate temporal dependencies to statistical models are very rare and at an early stage of development. Here our approach CRF_{multi} , based on conditional random fields (CRF), should make a contribution. Conditional random fields consider context knowledge among neighboring primitives in the same way as Markov random fields (MRF) do. Furthermore conditional random fields handle the feature vectors of the neighboring primitives and not only the class labels. Additional to taking spatial context into account, we present an approach for multitemporal data processing where a temporal association potential has been integrated to the common CRF approach to model temporal dependencies. The classification works on pixel-level using spectral image features, whereas all available single images are taken separately. For our experiments a high resolution RapidEye satellite data set of 2010 consisting of 4 images made during the whole vegetation period from April to October is taken. Six crop type categories are distinguished, namely grassland, corn, winter crop, rapeseed, root crops and other crops. To evaluate the potential of the new conditional random field approach the classification result is compared to a manual reference on pixel- and on object-level. Additional a SVM approach is applied under the same conditions and should serve as a benchmark.

1. INTRODUCTION

In 2008 the German RapidEye system in its constellation of five satellites launched. Each satellite has five spectral bands (blue, green, red, red edge and near infrared) with a GSD of 6.5 m. This system and many other recently launched high resolution optical remote sensing satellites allow a higher availability of multitemporal image data. These data can be used for enhancing the classification accuracy and for analyzing land cover changes. While in the latter case, one has to deal with potential class transitions, in the first case the class remains unchanged. Nevertheless the appearance of the individual classes at several epochs might change and that can be used to achieve higher classification accuracy.

In this work we present two CRF-based approaches for multitemporal crop type classification of high resolution optical remote sensing data. The first one, called CRF_{all} , is a common CRF approach with a 2D-grid graph structure. For this approach, all extracted features of all images are concatenated in one feature vector per pixel. The second approach, CRF_{multi} , is based on an extension of the CRF concept by an additional temporal interaction potential. Here each pixel of each image is represented as one node in a 3D-grid. The decision for a pixel belonging to one class is based on the extracted features at pixel site and on spatial and temporal context. Although this approach was originally developed for change detection, in this work we investigate its potential for multitemporal crop type analysis where no class transitions occur during one vegetation period.

For both approaches no existing land cover map is required. Nevertheless, finally the results are overlaid to cropland objects

of an existing GIS to evaluate the geometric quality. The approaches are demonstrated for a set of four multispectral RapidEye-images. Additionally a SVM-classification that serves as a benchmark is performed.

In chapter 2 an overview on related work on multitemporal crop type classification and on CRF is given. Next the data and extracted features for our tests are described. Chapter 4 deals with the principle of CRF and our implementation. A brief overview on the applied SVM can be found in chapter 5. Finally chapter 6 shows the results of our experiments.

2. RELATED WORK

2.1 Multitemporal crop type classification

In the field of multitemporal crop type classification three main directions of approaches can be observed:

In the first category powerful classifiers or combinations of several classifiers that are developed originally for single image classification are used. The temporal dependencies are not modeled in any way and are only implicit contained. To use a time series of images simultaneously, the multitemporal images are simply stacked to one image. The contributions of Bruzzone et al. (2004), Gislason et al. (2006) and Waske and Braun (2009) as well as our approach CRF_{all} belong to this category. Bruzzone et al. (2004) propose a multi classifier system and apply three classifiers in parallel, one of them considering the k nearest neighbors of each pixel. Decision-tree-ensembles (Random-Forest) (Gislason et al., 2006) (Waske and Braun, 2009) that are based on several parallel decision

* Corresponding author.



Figure 1. RapidEye images 2010 (© 2010 RapidEye AG, Germany. All rights reserved.)

trees of weak classifiers are also discussed for some time in the field of crop classification and show good results.

The second category of approaches model temporal dependencies by rules or comparison strategies to typical phenological behavior. These models are explicit but in most cases quite simple to stay handable. To this category belong the contributions of Müller et al. (2010), Simonneaux et al. (2008), Lucas et al. (2007) and De Wit and Clevers (2004). Müller et al. (2010) use a classification based on weighting functions. Thereby feature vectors of different time are compared to typical phenological behavior that is learned from a learning sample. Similar Simonneaux et al. (2008) calculates NDVI-profiles that are evaluated using combinations of different thresholds. Lucas et al. (2007) proposes a rule based classification applying the software eCognition. The approach leads to a quite complex rule basis that is difficult to handle. De Wit and Clevers (2004) apply a pixel-wise Maximum Likelihood classification and combine the result with an evaluation of the NDVI profiles concerning the phenological behavior.

The third category of approaches integrates temporal dependencies to statistical models. This approach should be continued in our work using conditional random fields with the approach CRF_{multi} . To compare the results of the developed statistical methodology a SVM-classifier as state of the art is chosen. Examples for the third category are Feitosa et al. (2009) and Melgani and Serpico (2004). Feitosa et al. (2009) model temporal dependencies by Markov chains for detecting land cover transitions in Landsat images, but spatial context is not taken into account. In Melgani and Serpico (2004) the Markov Random Field (MRF) framework is extended by a temporal energy term based on a transition probability matrix in order to improve the classification results for two consecutive images.

2.2 Conditional random fields

The interaction between neighboring image sites (pixels or segments) in MRF is restricted to the class labels, whereas the features extracted from different sites are assumed to be conditionally independent. This restriction is overcome by Conditional Random Fields (CRF) that were introduced by Lafferty et al. (2001) for classifying one-dimensional data. CRF provide a discriminative framework that can also model dependencies between the features from different image sites and interactions between the labels and the features. They were first applied on image data by Kumar and Hebert (2003) for the detection of man-made objects in terrestrial images.

In remote sensing CRF have been used for the classification of settlement areas in high-resolution optical satellite images (Zhong and Wang, 2007) (Hoberg and Rottensteiner, 2010) and for generating a digital terrain model from LiDAR (Lu et al.,

2009). Roscher et al. (2010) classified crop types amongst other land cover classes in monotemporal Landsat data and achieved an accuracy of approx. 70% for the crop type classes. Hoberg et al. (2010) applied CRF on multitemporal high resolution data in order to enhance classification accuracy in no change areas as well as to detect changes. In most approaches using CRF on imagery the graph is constructed as a regular grid with nodes representing pixels or square patches. In contrast, Wegner et al. (2011) use an irregular graph derived from a mean-shift segmentation for their CRF-based approach for building detection based on features generated from aerial images and airborne InSAR data.

3. DATA DESCRIPTION

3.1 RapidEye data

RapidEye consists of a constellation of five satellites launched in 2008. Each satellite has five spectral bands (blue, green, red, red edge and near infrared) with a GSD of 6.5 m and a dynamic range of 12 bit. Parts of the used time series are depicted in Figure 1. For the experiments RapidEye orthophotos automatically preprocessed are taken. The preprocessing includes orthorectification, resampling to a pixel size of 5 m and applying an automatic atmospheric correction. The images are taken from April to October 2010.

3.2 Features

The spectral bands green, red, red edge and near infrared are crucial during classification of vegetation classes. Only these bands are taken as input for feature extraction. Inside a filter radius of 5 pixels resulting in a filter window of 11×11 pixels the mean value is calculated during feature extraction. Four available images of 2010 lead to a total of 16 features. For both approaches (CRF and SVM) the same features are taken as input to receive a direct comparison of both approaches. Because the focus of this work is not on feature selection and tests with gradient-based features did not improve our results, we decided to choose only these simple features.

4. CONDITIONAL RANDOM FIELDS

4.1 Principle

In many classification algorithms the decision for a class at a certain image site is just based on information derived at the regarded site, where a site might be a pixel, a square block of pixels in a regular grid or a segment of arbitrary shape. In fact, the class labels and also the data of spatially and temporally neighboring sites are often similar or show characteristic patterns, which can be modeled using CRF. In monotemporal

classification, we want to determine the vector of class labels \mathbf{x} whose components x_i correspond to the classes of image site $i \in S$ from the given image data \mathbf{y} by maximizing the posterior probability $P(\mathbf{x} | \mathbf{y})$ (Kumar and Hebert, 2006):

$$P(\mathbf{x} | \mathbf{y}) = \frac{1}{Z} \exp \left\{ \sum_{i \in S} A_i(x_i, \mathbf{y}) + \sum_{i \in S} \sum_{j \in N_i} I_{ij}(x_i, x_j, \mathbf{y}) \right\} \quad (1)$$

In Equation 1, N_i is the spatial neighborhood of image site i (thus, j is a spatial neighbor to i), and Z is a normalization constant, called the *partition function*. The *association potential* A_i links the class label x_i of image site i to the data \mathbf{y} , whereas the term I_{ij} , called *interaction potential*, models the dependencies between the labels x_i and x_j of neighboring sites i and j and the data \mathbf{y} . The model is very general in terms of the definition of the functional model for both A_i and I_{ij} ; refer to (Kumar and Hebert, 2006) for details.

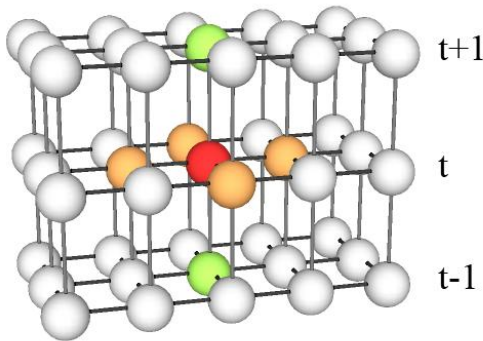


Figure 2. CRF_{multi} graph structure.

Red node: processed primitive; orange nodes: spatial neighbors; green nodes: temporal neighbors

In the multitemporal case, we have M co-registered images. The components of the image data vector \mathbf{y} are site-wise data vectors \mathbf{y}_i consisting of M components \mathbf{y}_i^t , where \mathbf{y}_i^t is the vector of the observed pixel values at image site i at epoch $t \in T$ and $T = \{1, \dots, M\}$. The components of \mathbf{x} are vectors $\mathbf{x}_i = [x_i^1, \dots, x_i^M]^T$, where x_i^t describes the class of image site i at epoch $t \in T$. For each image site we want to determine its class x_i^t for each time t from a set of C pre-defined classes. In order to model the mutual dependency of the class labels at an image site at different epochs, the model for $P(\mathbf{x} | \mathbf{y})$ in Equation 1 has to be expanded:

$$P(\mathbf{x} | \mathbf{y}) = \frac{1}{Z} \exp \left\{ \sum_{i \in S} \sum_{t \in T} A_i^t(x_i^t, \mathbf{y}^t) + \sum_{i \in S} \sum_{t \in T} \sum_{j \in N_i} IS_{ij}^t(x_i^t, x_j^t, \mathbf{y}^t) + \sum_{i \in S} \sum_{t \in T} \sum_{k \in K_i} IT_{i^t k}^t(x_i^t, x_i^k, \mathbf{y}^t, \mathbf{y}^k) \right\} \quad (2)$$

In Equation 2, \mathbf{y}^t and \mathbf{y}^k are the images observed at epochs t and k , respectively. The association potential A_i^t is identical to A_i for epoch t in Equation 1. The second term in the exponent, IS_{ij}^t , is identical to I_{ij} for epoch t in Equation 1, but it is called *spatial interaction potential* in order to distinguish it from the third term in the exponent, the *temporal interaction potential* $IT_{i^t k}^t$. K_i

is the *temporal neighborhood* of image site i at epoch t , thus k is the time index of an epoch that is a “temporal neighbor” of t . The temporal interaction potential models the dependency of class labels at consecutive epochs and the observed data. In our case the image sites are chosen to be pixels and thus are ordered in a regular grid (Figure 2). We model the CRF to be isotropic and homogeneous, hence the functions used for A_i^t , IS_{ij}^t and $IT_{i^t k}^t$ are independent of the location of image site i .

4.2 Implementation

The image data are represented by a site-wise feature vector $\mathbf{f}_i^t(\mathbf{y}^t)$ that may depend on the whole image at epoch t , e.g. by using features at different scales in scale space (Kumar and Hebert, 2006).

For our task of multitemporal crop type classification we apply two CRF-based approaches:

CRF_{all}: The extracted features \mathbf{f}_i^t of all images at each site are concatenated in one feature vector \mathbf{f}_i^{all} . So all information of one site is merged and no temporal structure (e.g. order of images) is existent any more. This allows the use of standard classification techniques, here the application of a common CRF approach with a 2D-grid structure as described in Equation 1. It should be noted that it would not be possible to detect class changes between epochs with this approach. Of course this is not wanted in this case.

CRF_{multi}: Implementation of the approach as described in Equation 2 and Figure 1. Each image site i at each time t is represented as one node in the graph with its feature vector \mathbf{f}_i^t . In general this approach allows class transitions between epochs, in this work we have to avoid this by appropriate definition of the temporal interaction potential.

The association potential $A_i^t(x_i^t, \mathbf{y}^t)$ is related to the probability of label x_i^t given the image \mathbf{y}^t at epoch t by $A_i^t(x_i^t, \mathbf{y}^t) = \log \{ P[x_i^t | \mathbf{f}_i^t(\mathbf{y}^t)] \}$. We use a simple Gaussian model for $P[x_i^t | \mathbf{f}_i^t(\mathbf{y}^t)]$ (Bishop, 2006):

$$A_i(x_i^t, \mathbf{y}^t) = -\frac{n}{2} \log(2\pi) - \frac{1}{2} \log \left[\det(\boldsymbol{\Sigma}_{fc}^t) \right] - \frac{1}{2} \left[\mathbf{f}_i^t(\mathbf{y}^t) - \mathbf{E}_{fc}^t \right]^T \cdot \boldsymbol{\Sigma}_{fc}^{-1} \cdot \left[\mathbf{f}_i^t(\mathbf{y}^t) - \mathbf{E}_{fc}^t \right] \quad (3)$$

In Equation 3, \mathbf{E}_{fc}^t and $\boldsymbol{\Sigma}_{fc}^t$ is the mean and co-variance matrix of the features of class c , respectively. For CRF_{multi} they are determined from the features $\mathbf{f}_i^t(\mathbf{y}^t)$ in training sites individually for each epoch t and each class c . With $t=1$ Equation 3 is also applied for CRF_{all} but using the concatenated feature vector \mathbf{f}_i^{all} . The spatial interaction potential IS_{ij}^t is a measure for the influence of the data \mathbf{y}^t and the neighboring labels x_j^t on the class x_i^t of site i at epoch t . For both approaches CRF_{all} (with $t=1$) and CRF_{multi} we applied the identical implementation and parameters to ensure comparability of the results. The data are represented by site-wise vectors of *interaction features* $\boldsymbol{\mu}_{ij}^t(\mathbf{y}^t)$ (Kumar and Hebert, 2006). We use the component-wise differences of the feature vectors \mathbf{f}_i^t respectively \mathbf{f}_i^{all} , i.e. $\boldsymbol{\mu}_{ij}^t(\mathbf{y}^t) = [\mu_{ij1}^t, \dots, \mu_{ijR}^t]^T$, where R is the dimension of the vectors

$\mathbf{f}_i^t / \mathbf{f}_i^{all}$, $\mu_{ijk}^t = |f_{ik}^t(\mathbf{y}^t) - f_{jk}^t(\mathbf{y}^t)|$, and $f_{ik}^t(\mathbf{y}^t)$ is the k^{th} component of $\mathbf{f}_i^t(\mathbf{y}^t)$. Introducing β as a sensitiveness factor for the effect of neighboring data being different, IS_{ij}^t is modeled as:

$$IS_{ij}^t(x_i^t, x_j^t, \mathbf{y}^t) = \begin{cases} \exp \left[-\beta \cdot \left\| \boldsymbol{\mu}_{ij}^t(\mathbf{y}^t) \right\|^2 \right] & \text{if } x_i^t = x_j^t \\ 1 - \exp \left[-\beta \cdot \left\| \boldsymbol{\mu}_{ij}^t(\mathbf{y}^t) \right\|^2 \right] & \text{if } x_i^t \neq x_j^t \end{cases} \quad (4)$$

In Equation 4, $\|\boldsymbol{\mu}_{ij}^t(\mathbf{y}^t)\|$ is the Euclidean norm of $\boldsymbol{\mu}_{ij}^t(\mathbf{y}^t)$. This is a very simple model that penalizes local changes of the class labels if the data are similar. The only model parameter is β . It could be determined from training data if fully labeled training images were available, but currently it is defined by the user. The neighborhood N_i of image site i over which IS_{ij}^t has to be summed in Equation 2 consists of the four neighboring image sites in a regular grid (Figure 1).

The temporal interaction potential IT_i^{tk} that we use for CRF_{multi} models the dependencies between the data \mathbf{y} and the labels x_i^t and x_i^k of site i at epochs t and k . We apply a bidirectional transfer of temporal information, so the temporal neighborhood K_i of x_i^t is chosen to contain the two elements x_i^{t-1} and x_i^{t+1} . Due to seasonal effects on the vegetation and due to different atmospheric and lighting conditions it would not be sufficient to derive IT_i^{tk} just from the difference of features vectors as for IS_{ij}^t . Instead this difference $\mathbf{d}_i^{tk}(\mathbf{y}^t, \mathbf{y}^k)$ is set in relation to the differences \mathbf{D}_{fc}^{tk} of the mean of the features of each class c (Equations 5-7).

$$\mathbf{d}_i^{tk}(\mathbf{y}^t, \mathbf{y}^k) = \mathbf{f}_i^t(\mathbf{y}^t) - \mathbf{f}_i^k(\mathbf{y}^k) \quad (5)$$

$$\mathbf{D}_{fc}^{tk} = \mathbf{E}_{fc}^t - \mathbf{E}_{fc}^k \quad (6)$$

$$\mathbf{v}_{ic}^{tk}(\mathbf{y}^t, \mathbf{y}^k) = \mathbf{D}_{fc}^{tk} - \mathbf{d}_i^{tk}(\mathbf{y}^t, \mathbf{y}^k) \quad (7)$$

Using the difference measure $\mathbf{v}_{ic}^{tk}(\mathbf{y}^t, \mathbf{y}^k)$ it is possible to support the decision for a site i to belong to a class c , if the features at that site show a typical development for that class between epochs t and k . This is realized in Equation 8 when computing IT_{ic}^{tk} .

$$IT_{ic}^{tk}(x_i^t, x_i^k, \mathbf{y}^t, \mathbf{y}^k) = \begin{cases} 1 - \varphi(c) & \text{if } x_i^t = x_i^k \text{ and } \varphi(c) < 0.99 \\ 0.01 & \text{if } x_i^t = x_i^k \text{ and } \varphi(c) \geq 0.99 \\ 0 & \text{if } x_i^t \neq x_i^k \end{cases} \quad (8)$$

$$\text{with } \varphi(c) = \left\| \mathbf{v}_{ic}^{tk}(\mathbf{y}^t, \mathbf{y}^k) \right\| / \varepsilon \quad (9)$$

The temporal interaction potential IT_i^{tk} is set to zero if the classes assigned to x_i^t and x_i^k differ (Equation 8), so class transitions can be avoided.

The only parameter of our temporal model is ε (Equation 9). Using ε the effect for one site having a different development than the average class development of any class can be adjusted. It could be estimated from training data, but currently it is set by the user.

Exact inference is computationally intractable for CRF (Kumar and Hebert, 2006). In (Vishwanathan et al., 2006), several methods for parameter learning and inference are compared. In this paper we use Loopy-Belief-Propagation (LBP) (Nocedal and Wright, 2006), which is a standard technique for performing probability propagation in graphs with cycles.

5. SVM

As support vector machines (SVM) are successfully applied in numerous remote sensing applications (Mountrakis et al., 2011) a SVM-classifier should serve as a reference classification operator.

5.1 Principle

In principle, the SVM is a binary classifier. Therefore, samples of two classes are used to train a model that separates these classes in feature space. Being a maximum margin classifier, the SVM maximizes the space between cluster borders. Based on this, separating hyperplanes are defined by support vectors which are a subset of the training vectors. In order to allow separation of non-linearly separable data, a common approach in machine learning is applied: The feature space is mapped to a higher dimensional space, where classes become linearly separable. This is done by applying kernel methods (Hofmann et al., 2008).

In our case, more than two classes have to be discerned. There are several strategies to transfer two-class problems to multi-class problems (Vapnik, 1998, Schölkopf and Smola, 2002). The most common are one vs. one and one vs. rest. The first one uses a voting scheme that accumulates the number of wins in a pairwise classification of each two classes. The second classifies each class against all others. The highest output determines the winning class. Here, the approach of (Chang and Lin, 2001) is applied with the one vs. one strategy using the concatenated feature vector \mathbf{f}_i^{all} . It performs similar to one vs. rest strategy for classification, but is faster in the training process.

The output of the classification process is a pixel-wise classification result. For more detail, in (Burges, 1998) a comprehensive tutorial about SVM is given.

6. RESULTS

Both CRF-based approaches as well as the SVM-classifier are applied to the data described in chapter 3.1.

6.1 Description of classes

The choice of crop type classes is based on expert knowledge about the most important categories of agricultural crops. The classes of interest are

- grassland,
- corn,
- winter crop,
- rapeseed,
- root crops and
- other crops.

Whereas the class winter crop consists in detail of the classes barley, oat, rye, wheat and triticale. Other crops include in our case the classes asparagus and strawberries.

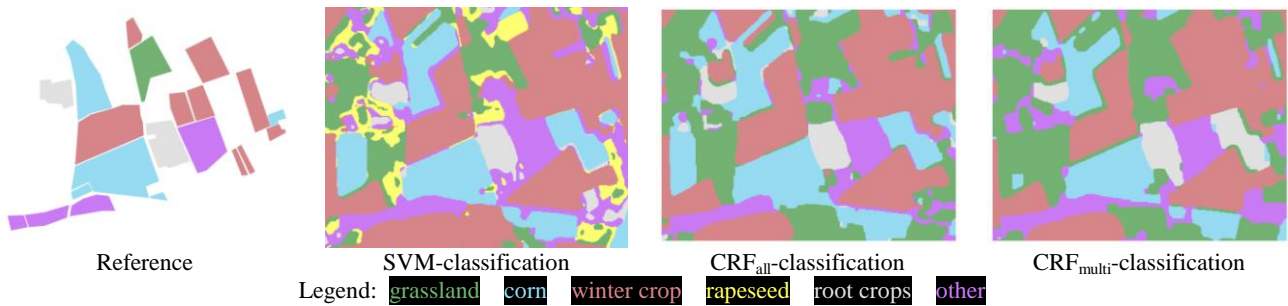


Figure 3. Reference data and classification results in comparison

6.2 Description of manual reference

A manual reference is used that was made in the same year 2010 like the satellite images. It consists of 121 separate fields of an area of about 322ha and has been acquired by field walking. The portion of the individual crop types of the whole area is 12% grassland, 21% corn, 28% winter crop, 11% rapeseed, 11% root crops and 17% other crops. In the process an existing GIS was used to define the borders of single fields. The manual reference builds the training and evaluation sample for the test classifications.

6.3 CRF vs. SVM

For our tests we applied the cross-validation method by separation the learning sample into two equal parts. Tables 1-4 show the results for the two CRF-approaches and the SVM-classification. Overall 129001 pixels were classified.

Both of the CRF-based approaches slightly outperform the SVM classification with CRF_{all} being best. For all approaches the overall accuracy is far over 80%, only the class grassland is classified with lower accuracy in each case. The classification results for a section of 21 fields are displayed in Figure 3.

In a next step, the majority of pixels belonging to a class in each reference fields was determined. This gives an idea of how good this approach is suited for classifying complete fields, ignoring classification errors at their borders. Applying CRF_{all} 108 of the 121 reference fields were classified correctly (89.3%), with CRF_{multi} 102 fields were correct (84.3%).

In general there are two main reasons for misclassifications: At first some fields show an “untypical” appearance for their class, e.g. most of the fields of one class are already harvested at one time of image acquisition but on some fields the crop is still present. Second some fields are very slender. So by using our feature extraction in an 11*11 window, their characteristics become blurred.

Cl Ref	Gra	Cor	Win	Rap	Roo	Oth
Gra	74.8	2.4	13.1	5.7	2.9	1.1
Cor	0.2	93.4	0	4.0	1.8	0.6
Win	1.0	3.3	88.7	2.7	1.9	2.3
Rap	0	4.5	0	95.5	0	0
Roo	0.8	2.3	0.5	0.1	95.2	1.0
Oth	2.7	6.8	4.0	0.4	9.4	76.7

Table 1. Confusion matrix for CRF_{all}-classification

Cl Ref	Gra	Cor	Win	Rap	Roo	Oth
Gra	63.2	11.1	13.8	7.1	3.2	1.6
Cor	0.1	90.6	0.1	4.5	4.0	0.7
Win	2.1	0	94.9	0	0	2.9
Rap	0.8	17.7	0.2	81.2	0	0.2
Roo	0.3	13.5	0.1	0	83.5	2.7
Oth	1.5	7.4	3.4	0.3	7.1	80.4

Table 2. Confusion matrix for CRF_{multi}-classification

Cl Ref	Gra	Cor	Win	Rap	Roo	Oth
Gra	70.9	0.5	8.9	12.9	0.3	6.6
Cor	1.8	76.6	0.1	12.3	0.4	8.9
Win	3.2	0.2	87.7	0.6	0.1	8.1
Rap	4.3	6.5	0.5	88.6	0	0.1
Roo	1.1	3.7	1.6	0.5	83.4	8.9
Oth	2.4	0.6	6.3	0.5	5.3	84.9

Table 3. Confusion matrix for SVM-classification (Cl=classification, Ref=Reference, Gra=Grassland, Cor=Corn, Win=winter crop, Rap=rapeseed, Roo=root crops, Oth=other crops.)

	overall accuracy	kappa
CRF _{all}	87.4	0.85
CRF _{multi}	84.2	0.81
SVM	82.7	0.79

Table 4. Overview on overall accuracy and kappa coefficient

6.4 Comparison to GIS data

To evaluate the results concerning geometric accuracy we superimposed them with a national GIS dataset more specifically the German Authoritative Topographic Cartographic Information System (ATKIS). Among other sources ATKIS data are collected using aerial photography with a resolution of 20cm or 40cm supported by ground truth data, and set to be used in scale between 1:10.000 and 1:25.000. Objects of interest are point, line and area based objects listed at (AdV, 1997) with a minimum mapping unit of 0.1 ha to 1 ha. The geometric accuracy is 3m. Figure 4 illustrates that the class borders of the CRF-based approaches fit to boundaries of the GIS-dataset quite well.



Figure 4. CRF_{multi}-classification superimposed by GIS cropland object borders (black borders)

6.5 Detailed results of CRF_{multi}

The association potential in CRF_{multi} is the context-free result of a separate Maximum-Likelihood-classification for each epoch (Equation 3). A section of the ML-classification result for the four individual epochs is displayed in Figure 5 a)-d). For most of the pixels the classification results for the epochs differ, sometimes three or even four different classes are assigned. Moreover the classification result within many fields is inhomogeneous. Overall the classification accuracy for the single epochs is 59.6%.

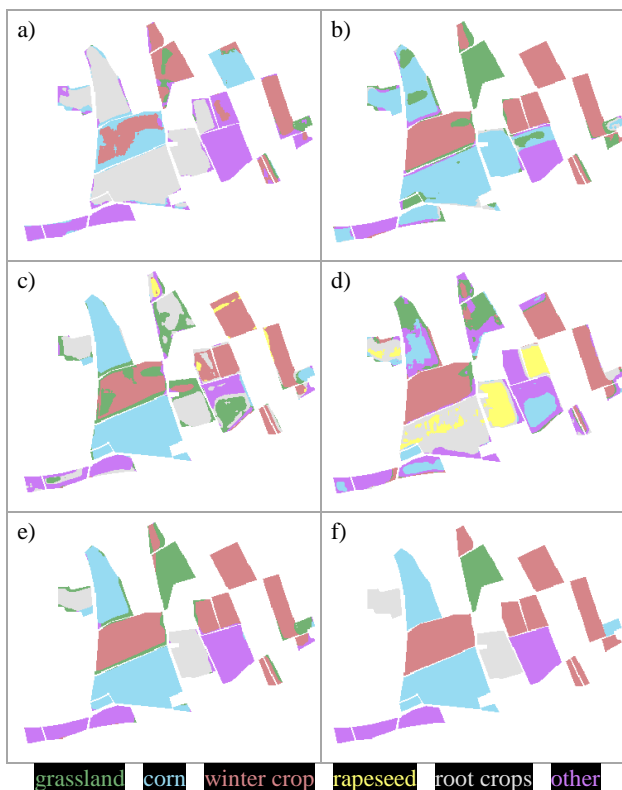


Figure 5 a)-d) Results of ML-classification for $t=1..4$;
 e) Result of CRF_{multi}-classification; f) Reference

By use of the temporal interaction potential (Equations 5-9) these results are set in temporal context and the overall accuracy is increased to 84.2%. Results of the CRF_{multi}-classification and the corresponding reference can be seen in Figure 5 e) and f).

7. CONCLUSION

In this work we presented two CRF-based approaches for multitemporal crop type analysis and tested them on RapidEye data of 4 epochs. Even with using just very few simple features, we achieved a classification accuracy of far over 80% for six crop type classes (grassland, corn, winter crop, rapeseed, root crops and other crops) with both approaches. Both of them performed better than a SVM-classification that served as a benchmark with the “classic approach” CRF_{all} being slightly better than CRF_{multi}. Nevertheless the CRF_{multi} approach generally has a higher potential for any kind of multitemporal analysis. Because of its flexibility in the definition of the temporal interaction potential, it is also applicable for tasks of change detection or multi-scale analysis..

ACKNOWLEDGEMENT

The implementation of the CRF-classification is based on “UGM: A Matlab toolbox for probabilistic undirected graphical models” by Mark Schmidt, <http://people.cs.ubc.ca/~schmidt/Software/UGM.html>.

The research is funded by the Federal Ministry of Economics and Technology (BMWi) via the German Aerospace Center (DLR e.V.) under the funding number 50EE0914 and by the German Science Foundation (Deutsche Forschungsgemeinschaft) under grant HE 1822/22-1.

REFERENCES

- AdV, Arbeitsgemeinschaft der Vermessungsverwaltungen der Länder der Bundesrepublik Deutschland, 1997. ATKIS – Amtlich Topographisch-Kartographisches Informationssystem, Germany. <http://www.atkis.de> (accessed 2011-04-01).
- Bishop, C. M., 2006. *Pattern recognition and machine learning*. 1st edition, Springer New York.
- Burges, C. J. C., 1998. A Tutorial on Support Vector Machines for Pattern Recognition. *Data Mining and Knowledge Discovery*, 2, pp. 121–167.
- Bruzzone, L., Cossu, R., Vernazza, G., 2004. Detection of land-cover transitions by combining multivariate classifiers. *Pattern Recognition Letters*, 25(13), pp. 1491-1500.
- Chang, C.C. and Lin, C.J., 2001 LIBSVM: a library for support vector machines. <http://www.csie.ntu.edu.tw/~cjlin/libsvm/>, accessed: 2011-04-01.
- De Wit, A. J. W. and Clevers, J. G. P. W., 2004. Efficiency and accuracy of per-field classification for operational crop mapping. *International Journal of Remote Sensing*, 25(20), pp. 4091–4112.
- Feitosa, R. Q., Costa, G. A. O. P., Mota, G. L. A., Pakzad, K., Costa, M. C. O., 2009. Cascade multitemporal classification

- based on fuzzy Markov chains. *ISPRS J. Photogrammetry Remote Sens.* 64(2), pp. 159-170.
- Gislason, P. O., Benediktsson, J. A., and Sveinsson, J. R., 2006. Random Forests for land cover classification. *Pattern Recognition Letters*, 27(4), pp. 294–300.
- Hoberg, T. and Rottensteiner, F., 2010. Classification of settlement areas in remote sensing imagery using Conditional Random Fields. *Int. Arch. Photogrammetry, Remote Sens., SIS XXXVIII (7A)*, pp. 53-58.
- Hoberg, T., Rottensteiner, F., Heipke, C., 2010. Classification of Multitemporal Remote Sensing Data Using Conditional Random Fields. *6. IAPR TC 7 Workshop on Pattern Recognition in Remote Sensing*, Istanbul.
- Hofmann, T., Schölkopf, B., Smola, A.J., 2008. Kernel methods in machine learning, *Annals of Statistics*, 36, No.3, pp. 1171–1220.
- Kumar, S. and Hebert, M., 2006. Discriminative Random Fields. *Int. J. Computer Vision*, 68(2), pp. 179-201.
- Kumar, S., Hebert, M., 2003. Discriminative Random Fields: A discriminative framework for contextual interaction in classification. *Proc. of IEEE International Conference on Computer Vision*, 2, pp. 1150-1157.
- Lafferty, J., McCallum, A., Pereira, F., 2001. Conditional Random Fields: Probabilistic Models for segmenting and labeling sequence data. *Proc. of International Conference on Machine Learning*, 8p.
- Lu, W.-L., Murphy, K. P., Little, J. J., Sheffer, A., Fu, H., 2009. A hybrid conditional random field for estimating the underlying ground surface from airborne Lidar data. *IEEE-TGARS*, 47(8/2), pp. 2913-2922.
- Lucas, R., Rowlands A., Brown A., Keyworth, S. and Bunting, P., 2007. Rule-based classification of multi-temporal satellite imagery for habitat and agricultural land cover mapping. *ISPRS Journal of Photogrammetry and Remote Sensing*, 62(3), pp.165–185.
- Melgani, F. and Serpico, S. B., 2003. A Markov Random Field approach to spatio-temporal contextual image classification. *IEEE-TGARS*, 41(11), pp. 2478-2487.
- Mountrakis, G., Im, J., Ogole, C., 2011. Support vector machines in remote sensing: A review. *ISPRS Journal of Photogrammetry and Remote Sensing*, 66(3), pp. 247-259.
- Müller, S., Heipke, C. and Pakzad, K., 2010. Classification of farmland using multitemporal aerial images. *International Archive of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. XXXVIII, part 4-8-2/W9, Haifa, pp. 70–74.
- Nocedal, J. and Wright, S. J., 2006. *Numerical Optimization*. 2nd edition, Springer New York.
- Roscher, R., Waske, B., Förstner, W., 2010. Kernel Discriminative Random Fields for Land Cover Classification. *6. IAPR TC 7 Workshop on Pattern Recognition in Remote Sensing*, Istanbul.
- Simonneaux, V., Duchemin, B., Helson, D., Er-Raki, S., Olios, A. and Chehbouni, A. G., 2008. The use of high-resolution image time series for crop classification and evapotranspiration estimate over an irrigated area in central Morocco. *International Journal of Remote Sensing*, 29(1), pp. 95–116.
- Schölkopf, B. and Smola, A.J., 2002. *Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond (Adaptive Computation and Machine Learning)*. The MIT Press, Cambridge, MA, USA.
- Vapnik, V.N., 1998. *Statistical Learning Theory*. Wiley, New York, USA.
- Vishwanathan, S., Schraudolph, N. N., Schmidt, M. W., Murphy, K. P., 2006. Accelerated training of conditional random fields with stochastic gradient methods. 23rd *International Conference on Machine Learning*, pp. 969-976.
- Waske, B. and Braun M., 2009. Classifier ensembles for land cover mapping using multitemporal SAR imagery. *ISPRS Journal of Photogrammetry and Remote Sensing*, 64(5), pp. 450–457.
- Wegner, J.D., Sörgel, U., Rosenhahn, B., 2011. Segment-based building detection with Conditional Random Fields, In: *U. Stilla, P. Gamba, C. Juergens, D. Maktav (Eds), Proceedings of 6th IEEE/GRSS/ISPRS Joint Urban Remote Sensing Event*, Munich, 2011, pp. 205-208.
- Zhong, P. and Wang, R., 2007. A multiple conditional random fields ensemble model for urban area detection in remote sensing optical images. *IEEE-TGARS*, 45(12), pp. 3978-3988.