ISPRS Journal of Photogrammetry and Remote Sensing 131 (2017) 65–76

Contents lists available at ScienceDirect

ISPRS Journal of Photogrammetry and Remote Sensing

journal homepage: www.elsevier.com/locate/isprsjprs

Unsupervised domain adaptation for early detection of drought stress in hyperspectral images

P. Schmitter [a,*], J. Steinrücken [a], C. Römer [a], A. Ballvora [b], J. Léon [b], U. Rascher [c], L. Plümer [a]

[a] Institute of Geodesy and Geoinformation, Department of Geoinformation, University of Bonn, Meckenheimer Allee 172, 53115 Bonn, Germany
[b] Institute of Crop Science and Resource Conservation, Plant Breeding and Biotechnology, University of Bonn, Katzenburgweg 5, 53115 Bonn, Germany
[c] Institute of Bio- and Geosciences, IBG-2: Plant Sciences, Forschungszentrum Jülich, Leo-Brandt-Str., 52425 Jülich, Germany

ARTICLE INFO

Article history:
Received 27 May 2017
Received in revised form 25 June 2017
Accepted 6 July 2017

Keywords:
Unsupervised domain adaptation
Machine learning
Support vector machine
Hyper spectral
Agriculture

ABSTRACT

Hyperspectral images can be used to uncover physiological processes in plants if interpreted properly. Machine Learning methods such as Support Vector Machines (SVM) and Random Forests have been applied to estimate development of biomass and detect and predict plant diseases and drought stress. One basic requirement of machine learning implies, that training and testing is done in the same domain and the same distribution. Different genotypes, environmental conditions, illumination and sensors violate this requirement in most practical circumstances. Here, we present an approach, which enables the detection of physiological processes by transferring the prior knowledge within an existing model into a related target domain, where no label information is available. We propose a two-step transformation of the target features, which enables a direct application of an existing model. The transformation is evaluated by an objective function including additional prior knowledge about classification and physiological processes in plants. We have applied the approach to three sets of hyperspectral images, which were acquired with different plant species in different environments observed with different sensors. It is shown, that a classification model, derived on one of the sets, delivers satisfying classification results on the transformed features of the other data sets. Furthermore, in all cases early non-invasive detection of drought stress was possible.

1. Introduction

In recent years, several research groups have successfully demonstrated the detection of plants' physiological processes (e.g. plant stress) from hyperspectral images by using (supervised) methods of Machine Learning (e.g. Karimi et al. (2006), Mucherino et al. (2009), Römer et al. (2011)). Typically, each approach has been developed from scratch, i.e. a new model has been derived from data, which were measured in a specific experiment. However, considering the effort of labelling training data, the question arises how far a model is transferable to other data sets obtained by different sensors observing different plants in different environments.

Generally, most methods of Machine Learning are based on the fundamental assumption that training and test data have the same underlying feature space and distribution (Pan and Yang, 2010). In most real world applications this constraint is violated by varying measuring setup, different sensors or changing environment (e.g. illumination or background).

Strategies of reusing knowledge from a *source domain* (consisting of a feature space and feature distribution) and a *source task* (consisting of a label space and a predictive function) for learning a predictive function in a *target domain* with a *target task* are addressed in *Transfer Learning* (Pan and Yang, 2010). Subject to different settings, specific subcategories have been identified. In *Transductive Transfer Learning* the domains of source and target are different, whereas the tasks are the same. Labelled data are only available in the source domain. Different domains are caused either by different feature spaces or by different feature distributions. The latter case is also known as *Domain Adaptation* (Arnold et al., 2007).

In this paper we propose a Domain Adaptation approach for the detection of water limitation based on hyperspectral images. We focus on stress responses which

- are in early states and cannot be perceived by the naked eye,
- progress continuously and can be characterised by a set of ordinal, subsequent stages.

* Corresponding author.
E-mail address: schmitter@igg.uni-bonn.de (P. Schmitter).

http://dx.doi.org/10.1016/j.isprsjprs.2017.07.003
0924-2716/

The detection of such physiological processes is challenging as the hyperspectral data sets are influenced by the environmental factors of real world applications, and the invisibility of processes prevents a labelling.

Based on the hypothesis that this knowledge allows the reuse of an existing classification model in a different target domain, we propose a transformation of the features of the target domain and formulate an objective function, which enables an evaluation of the transformation parameters without using labelled data in the target domain. The objective function is based on general characteristics about classification and the biological knowledge about the ordinal scale of the process.

In this study we focus on the automated detection of drought stress induced changes in leaf pigments and leaf structure. Drought stress, caused by water scarcity, is one of the biggest challenges in global crop production and it has been estimated that drought can cause a depreciation of crop yield up to 70% in conjunction with other abiotic stresses (Pennisi, 2008; Tuberosa and Salvi, 2006). Prolonged water scarcity initiates invisible biochemical and physiological processes and, subsequently leads to an impairment of crop growth and yield.

The recent rapid developments of destructive and non-destructive technologies offer new opportunities for an effective and high-throughput analysis of plant characteristics. The connectivity and flow of these data, towards molecular breeding and farming, has been hampered by a bottleneck at the level of phenotyping, the so called phenomics bottleneck (Tardieu and Schurr, 2009). Effective use of sensors could contribute towards a pinpoint accuracy in phenotyping, reduced experimental requirements and will enable multiple, simultaneous and objective data to be collected.

Drought is a spatial-temporal process, which triggers various reactions within the plants. As a first intermittent reaction many plants close their leaf stomata, limiting the transpiration loss in periods of limited soil water availability. This physiological reaction normally cannot be detected by hyperspectral imaging as no plant pigments are affected. If water however remains a limiting resource for a longer time period, plants react by changing growth patterns and resource allocation in their organs. Leaf pigments may be broken down, leaf surface may be come pigmented and leaf anatomy changes. These responses however are very variable. Depending of the ecological niche of each species and the developmental stage plants may react very different to a limitation of water. Often drought responses proceed from older to younger leaves and, within a leaf, from the tip towards the leaf base.

The derivation of a source model requires labelled data for the invisible stages of drought stress. However, while it is quite easy to manually identify perfectly healthy and dead leaf pigments, it is not possible to visually grade the stage of senescence for presymptomatic stress detection. Therefore, while the change from healthy to senescent pigments is a continuous process, it is not possible to manually get continuous labels from the image alone. Hence, either an unsupervised regression model or a discretization into ordinal classes are feasible options. Römer et al. (2012) have presented an unsupervised regression method for early detection of drought induced stress with hyperspectral images based on cluster analysis. However, as described in detail later in this manuscript, the mightiness of classes from similar spectra is an important optimization criteria. Therefore, discrete ordinal classes as used in Behmann et al. (2014b) are preferable for the presented method. In Behmann et al. (2014a) the authors show that this ordinal classification approach outperforms non-linear regression with regard to plant stress detection in hyperspectral data. This multi-class model is based on the previous knowledge that process the of the drought induced stress forms an ordinal order, mainly related to chlorophyll degradation (Merzlyak et al., 1999). The model classifies each pixel into a drought class, following an ordinal scale from healthy to stressed. The drought states are a discrete representation of the progressive process of the drought induced stress. The model provides a description of the drought state of a plant; the relative frequencies of the drought classes enable an early detection of drought induced stress (Behmann et al., 2014b).

Furthermore, we assume, that the spectral information of the hyperspectral images is adequately represented by a set of commonly used Vegetation Indices (VI, Table 1). However, although the indices are usually invariant to changes in the environment (Jensen, 2009), a simple reuse of the model does not provide sufficient results (Fig. 1).

The contribution of this paper is an unsupervised method, which allows the reuse of a classification model for the early detection of physiological plant processes from hyperspectral images. The method is characterized by

- the handling of data without any labels in the target domain, as well as without the need for labels at training time,
- the transformation of the feature space of the target domain,
- the use of biological knowledge for evaluating the transformation parameters, and
- the application of a source model which is not changed in the target domain.

We demonstrate the applicability of the proposed approach on three sets of hyperspectral images. Two sets have been measured in drought stress experiments on single barley plants. The experiments were set up in two consecutive years in *foliar tunnels* in Germany. The third data set was collected on maize, grown up in *the field* under different treatments in Italy. It is obvious that the transfer from barley in foliar tunnels in Germany to maize on the field in Italy is rather challenging. However, we will show that an ordinal classification model, derived from one of the sets, delivers satisfying classification results on the transformed features of the other data sets and that the classification result enables an early detection of drought effects.

2. Related work

Classic machine learning methods assume the same domain $\mathscr{D}$ and the same learning task $\mathscr{T}$ for training and test data. A domain consists of a feature space $\mathscr{X}$ and a probability distribution $P(X): \mathscr{D} = \{\mathscr{X}, P(X)\}$. For a given domain, a task consists of a label space $\mathscr{Y}$ and a predictive function $f(\cdot): \mathscr{T} = \{\mathscr{Y}, f(\cdot)\}$ (Pan and Yang, 2010).

Transfer Learning differentiates between a source domain $\mathscr{D}_{\mathscr{S}} = \{\mathscr{X}_{\mathscr{S}}, P(X_S)\}$ with a source task $\mathscr{T}_{\mathscr{S}} = \{\mathscr{Y}_{\mathscr{S}}, f_S(\cdot)\}$, and a target domain $\mathscr{D}_{\mathscr{T}} = \{\mathscr{X}_{\mathscr{T}}, P(X_T)\}$ with a target task $\mathscr{T}_{\mathscr{T}} = \{\mathscr{Y}_{\mathscr{T}}, f_T(\cdot)\}$, where $\mathscr{D}_{\mathscr{S}} \neq \mathscr{D}_{\mathscr{T}}$, or $\mathscr{T}_{\mathscr{S}} \neq \mathscr{T}_{\mathscr{T}}$ (Pan and Yang, 2010). The target predictive function $f_T(\cdot)$ is to be learned by reusing knowledge from $\mathscr{D}_{\mathscr{S}}$ and $\mathscr{T}_{\mathscr{S}}$.

Subject to the availability of labels and different relations between the source and target domain and the source and the target task, specific subcategories of Transfer Learning have been identified (for a full taxonomy see, e.g. Pan and Yang, 2010). Problems in which labelled data are only available in the source domain are addressed in Transductive Transfer Learning. Furthermore, the source and the target domain are different ($\mathscr{D}_{\mathscr{S}} \neq \mathscr{D}_{\mathscr{T}}$), while the source task and the target task are the same ($\mathscr{T}_{\mathscr{S}} = \mathscr{T}_{\mathscr{T}}$). The different domains are caused either by different feature spaces ($\mathscr{X}_{\mathscr{S}} \neq \mathscr{X}_{\mathscr{T}}$) or by different feature distributions ($\mathscr{P}(\mathscr{X}_{\mathscr{S}}) \neq \mathscr{P}(\mathscr{X}_{\mathscr{T}})$). In this paper, we focus the latter case, which is also known as unsupervised Domain Adaptation.