Time series-based classifier fusion for fine-grained plant species recognition☆

Fabio A. Faria [a,c,*], Jurandy Almeida [a,c], Bruna Alberton [b], Leonor Patricia C. Morellato [b], Anderson Rocha [c], Ricardo da S. Torres [c]

[a] Institute of Science and Technology, Federal University of São Paulo – UNIFESP, 12247-014 São José dos Campos, SP, Brazil
[b] Department of Botany, São Paulo State University – UNESP, 13506-900 Rio Claro, SP, Brazil
[c] Institute of Computing, University of Campinas – UNICAMP, 13083-852 Campinas, SP, Brazil

ARTICLE INFO

ABSTRACT

Global warming and its resulting environmental changes surely are ubiquitous subjects nowadays and undisputedly important research topics. One way of tracking such environmental changes is by means of phenology, which studies natural periodic events and their relationship to climate. Phenology is seen as the simplest and most reliable indicator of the effects of climate change on plants and animals. The search for phenological information and monitoring systems has stimulated many research centers worldwide to pursue the development of effective and innovative solutions in this direction. One fundamental requirement for phenological systems is concerned with achieving fine-grained recognition of plants. In this sense, the present work seeks to understand specific properties of each target plant species and to provide the solutions for gathering specific knowledge of such plants for further levels of recognition and exploration in related tasks. In this work, we address some important questions such as: (i) how species from the same leaf functional group differ from each other; (ii) how different pattern classifiers might be combined to improve the effectiveness results in target species identification; and (iii) whether it is possible to achieve good classification results with fewer classifiers for fine-grained plant species identification. In this sense, we perform different analysis considering RGB color information channels from a digital hemispherical lens camera in different hours of day and plant species. A study about the correlation of classifiers associated with time series extracted from digital images is also performed. We adopt a successful selection and fusion framework to combine the most suitable classifiers and features improving the plant identification decision-making task as it is nearly impossible to develop just a single "silver bullet" image descriptor that would capture all subtle discriminatory features of plants within the same functional group. This adopted framework turns out to be an effective solution in the target task, achieving better results than well-known approaches in the literature.

© 2015 Elsevier B.V. All rights reserved.

1. Introduction

Environmental changes have emerged as an important question in the global agenda [24,29]. This has spurred important research interest in phenology, the science of studying recurrent life cycles events and its relationship to climate [8,10,31]. To increase the range of study sites and species and the accuracy of phenological observations, digital cameras have been successfully applied as multi-channel imaging sensors, providing measures to estimate changes on phenological events, such as leaf flushing and senescence [1,26,27].

Previous work of our research group has monitored leaf changing patterns of a neotropical savanna (cerrado sensu stricto vegetation) based on daily acquired digital images [2]. We extracted image color information from the RGB (red, green, and blue) channels and correlated the changes in pixel levels over time with leaf phenology patterns [2]. The analysis was conducted after we defined regions of interest (ROI) based on the random selection of plant species crowns identified in the digital image [26]. Time series associated with each ROI have been obtained, raising the need of using appropriate tools for mining patterns of interest [3,4,6,28,30].

Fine-grained species identification in digital images is a key issue for the phenological observation of tree crowns, especially in

---

tropical vegetations where one single image may include a high number of species [2]. Usually, this task is very time consuming since it has to be done in the field, first by matching each crown in the image to the tree in the soil and then by identifying the tree at species level [5]. In this sense, we have designed and deployed machine learning methods to find similar patterns in the digital images and then we checked if they correspond to similar species or leaf functional groups [5].

Our first studies have focused on the intraspecies analysis, i.e., on detecting different individuals of the same species [5]. However, different species from the same leaf functional group may exhibit similar phenological traits, confusing the classifiers as discriminative features among the classes are more subtle, hardening the identification process. Hence many questions arise when considering interspecies interactions, i.e., the recognition of individuals from different species belonging to the same leaf functional group [2] spurring the need for the proper design and development of fine-grained recognition algorithms to tackle the problem. Therefore, in this paper, we aim at addressing the following questions: (i) how to distinguish different species from the same leaf functional group using a pattern classification scheme for proper fine-grained decision-making; (ii) how the individual responses of classifiers built upon different phenological features are correlated to each other; and (iii) how to combine such phenological features so as to improve the responsiveness of all plant species as it is unlikely that just one phenological treat would be enough for complete and proper identification.

We start by evaluating the performance of classification models built upon a single phenological feature. Thereafter we perform a correlation analysis in order to understand the responsiveness of each plant species regarding multiple phenological features. Based on the observations made, we adopted a successful fusion framework [12] to collect complementary features and better solve the multiclass classification problem. This kind of problem has never been addressed in our studies before. Finally, we analyze the impact of increasing the number of classifiers in the individual responses of each plant species.

The most important contributions of our work are: (1) A correlation analysis between different time series-based classifiers for each species. In this vein, we can identify correlations between classifiers and relationship intra/inter species, which show to be very important for the problem we are tackling in this paper; (2) The exploration and proper custom-tailoring of a classifier fusion framework [12] to improve the effectiveness results in a new application (species recognition); (3) The exploration of this classifier fusion framework for combining time series-based classifiers. Unlike the previous work of ours [12], which has used visual properties (e.g., color, texture, and shape) as input data to training base classifiers, our new proposal uses time series, a much different problem with its own intrinsic particularities and reduced information when compared to images; (4) Finally, in this work, we have considered a multi-class classification problem differently from the previous work.
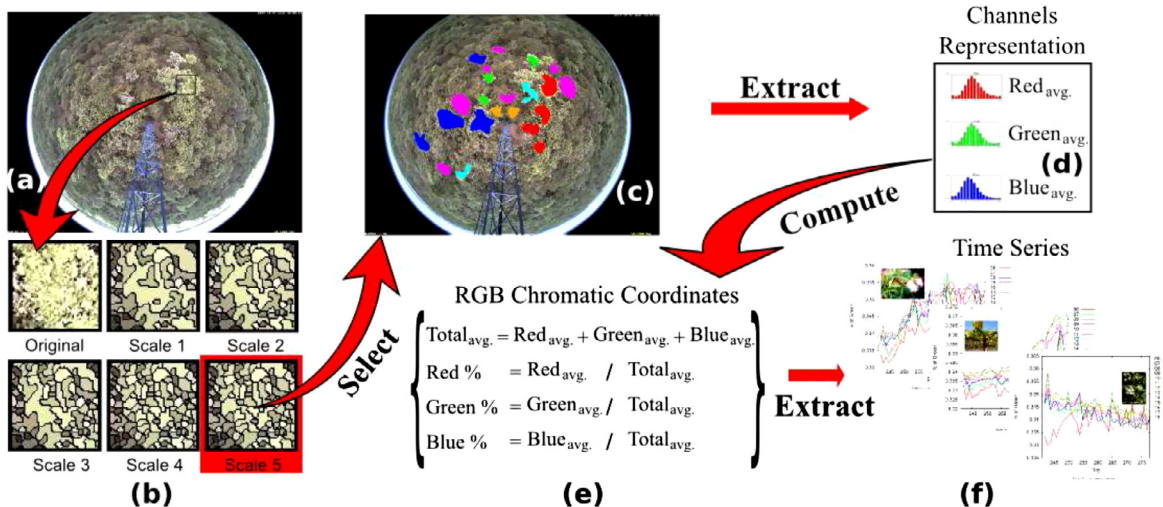
2. Time series acquisition

A digital hemispherical lens camera (Mobotix Q24) was set up in an 18 m-high tower in a Cerrado *sensu stricto*, a neotropical savanna vegetation located at Itirapina, São Paulo State, Brazil [2,25]. Fig. 1 shows all steps of the time series acquisition process used in our work.

First, we set up the RGB digital camera to take a daily sequence of five JPEG images (at 1280 × 960 pixels of resolution) per hour, from 6:00 to 18:00 h (UTC-3). The present study was based on the analysis of over 2700 images (Fig. 1(a)), recorded at the end of the dry season, between August 29 and October 3, 2011, day of year 241 to 278, during the main leaf flushing season [2].

Next, the image analysis has been conducted by defining different regions of interest (ROI), as described in [26] and defined by [2] for our target species. Then, we analyzed 22 ROIs (Fig. 1(b)) obtained from a random selection of six plant species identified manually by phenology experts in the hemispheric image [2]: (i) three regions associated with *Aspidosperma tomentosum* (green areas), (ii) four regions for *Caryocar brasiliensis* (blue areas), (iii) two regions for *Myrcia guianensis* (orange areas), (iv) six regions for *Miconia rubiginosa* (magenta areas), (v) two regions for *Pouteria ramiflora* (cyan areas), and (vi) four regions for *Pouteria torta* (red areas).

We analyzed each ROI in terms of the contribution of the primary colors (R, G, and B), as proposed by [27] and described in [2]. Initially, we analyze each color channel and compute the average value of the pixel intensity (Fig. 1(c)). After that, we compute the normalized brightness of each color channel (RGB Chromatic coordinates) (Fig. 1(d)). The normalization of those values reduces the influence of the incident light, decreasing the color variability due to changes on illumination conditions [9,34]. Finally, by computing those values along the whole period (August 28 to October 3, 2011), we obtained time series to use as input data for a learning method (Fig. 1(e)).

Fig. 1. The time series acquisition process pipeline. (a) Sample image of the Cerrado savanna; (b) Different segmentation scales are computed and the coarse scale is selected; (c) Hemispherical image with the selected ROI's species; (d) Channel representation are extracted from ROI's; (e) RGB chromatic coordinates are computed; (f) Phenological time series extracted from digital images. (For interpretation of the references to color in this figure text, the reader is referred to the web version of this article.)

According to the leaf exchange data from the on-the-ground field observations on leaf fall and leaf flush at our study site, those species were classified on three functional groups [2,23]: (i) deciduous, *A. tomentosum* and *C. brasiliensis*; (ii) evergreen, *M. guianensis* and *M. rubiginosa*; and (iii) semideciduous, *P. ramiflora* and *P. torta*.

3. Framework for time series-based classifier fusion

The objective of adopting a classifier fusion approach is to exploit the degree of agreement/disagreement among different classifiers, concept known as *diversity* and to improve the effectiveness results in the target task. In this sense, we adapt a successful selection and fusion framework, originally proposed for multimedia recognition, to be used as a combiner of times series-based classifier [12]. This framework selects the most suitable classifiers to be used in a meta-learning approach. In this work, we define a classifier as a tuple containing a simple learning method (e.g., k-Nearest Neighbors – kNN) and a description technique (e.g., a color channel from the RGB color channels).

3.1. Classifier fusion approach

Fig. 2 illustrates the used framework [12] for plant identification. First, classifiers learn patterns from a training set (T) that contains samples of time series. Next, $|\mathcal{C}|$ classification models are created. They are applied on a validation set V, resulting in a matrix of classifier outcomes M_V, where $|M_V| = |V| \times |\mathcal{C}|$ and $|V|$ is the number of time series extracted from a validation set V (Fig. 2(a)).

Thereafter, in the selection of the most appropriate time series-based classifiers to be combined, M_V is used to calculate different diversity measures (D). These measures compute the degree of agreement and disagreement all of $|\mathcal{C}|$ available classifiers [20].

The main objective of this selection process is to select a set $\mathcal{C}^* \subset \mathcal{C}$ of classifiers that are more suitable to be fused/combined (Fig. 2(b)). Note that a new matrix $M_V^* \subset M_V$ is computed. Finally, given a new time series extracted from I, a fusion technique (e.g., Support Vector Machines) uses the newly created matrix M_V^* to learn patterns and thus define the final class of I through a meta-learning approach (Fig. 2(c)).
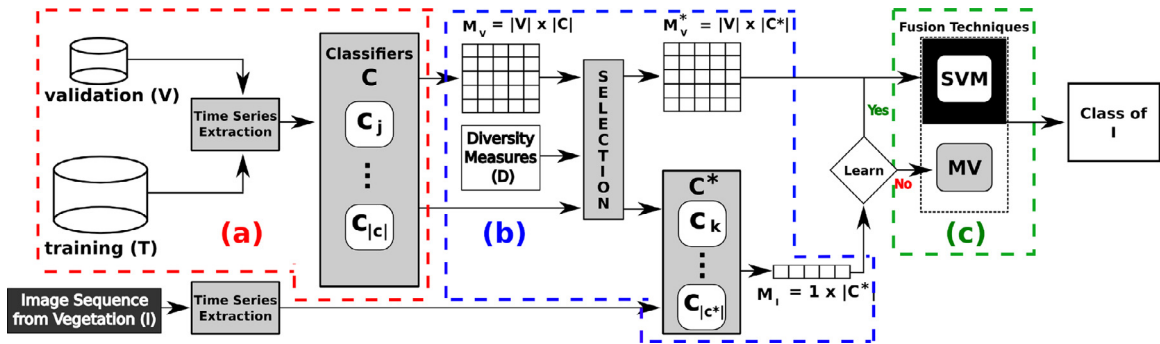
3.2. Classifier selection approach

Fig. 3 illustrates the adopted five-step approach for selecting classifiers based on diversity measures, previously introduced in [12].

First, set $\mathcal{D}$ of diversity measures are used to assess the degree of agreement among available classifiers in $\mathcal{C}$ by taking into account the M_V matrix previously computed. In this approach, five different measures have been used (*Correlation Coefficient p, Double-Fault Measure, Disagreement Measure, Interrater Agreement k*, and *Q-Statistic* [20,21]). That step is represented by arrow (a) in Fig. 3. Pairs of classifiers are then ranked according to their diversity score. Each diversity measure defines a different ranked list and, at the end of this step, a set $\mathcal{R}$ of ranked lists is produced (arrow (b)). In the following, a novel set of ranked lists $\mathcal{R}^t$ is computed by selecting the top t pairs of classifiers from each ranked list in $\mathcal{R}$ (arrow (c)), and a histogram H that counts the number of occurrences of a classifier in all ranked lists of $\mathcal{R}^t$ is computed (arrow (d)). Finally, the most frequent classifiers in H, whose accuracy is greater that a given threshold $\mathcal{T}$, are combined by a fusion approach (arrow (e)). $\mathcal{T}$ is a threshold defined in terms of the average accuracy among all classifiers using the validation set V.

4. Experimental protocol

This section presents the experimental protocol used in this work.

4.1. Dataset

In this work, we have applied the same evaluation method used in [2]. It relies on the classification of time series extracted from pixels associated with individuals of a same species. For that, we used the algorithm introduced by [15] to segment the hemispheric image into small polygons, obtaining 8849 segmented regions (SR). Then, we associated each SR with a single ROI aiming to label it. A labeled region is created if there is at least 80% of overlapped area between an SR and a ROI.

Finally, we extracted a time series from each labeled region using the approach described in Section 2. In this way, we built a dataset of 892 time series separated into six classes, one for each plant species: *A. tomentosum* (96), *C. brasiliensis* (346), *M. guianensis* (36), *M. rubiginosa* (195), *P. ramiflora* (50), and *P. torta* (169).

Fig. 2. Time series-based classifier selection and fusion framework adapted from [12]. In (a), given a classification problem with training examples, different classifiers are trained using data from training set T. In (b), the most discriminating classifiers are selected ($\mathcal{C}^*$) by taking into account diversity measures (D). Finally, in (c), classifiers are combined in a meta-level approach using any other classifier. In this particular example, both the SVM and Majority Voting (MV) techniques can be used as the classifier fusion technique.
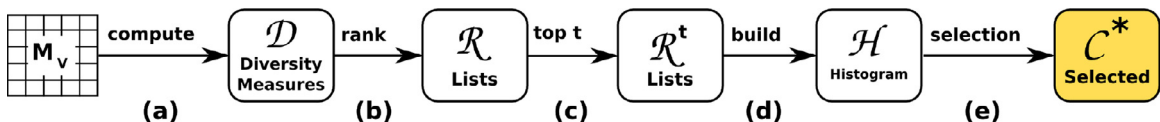

Fig. 3. The five steps for classifier selection are: (a) Computation of diversity measures from the validation matrix M_V; (b) Ranking of pairs of classifiers by their diversity measures scores; (c) Selection of the top t ranked pairs of classifiers; (d) Computation of a histogram H that counts the number of occurrences of each classifier; (e) Select the most appropriate classifiers $|\mathcal{C}^*|$ based on their occurrence in H and satisfy a defined threshold $\mathcal{T}$.

Table 1
Confusion matrix. TP, TN, FP, and FN stand for true positive, true negative, false positive, and false negative, respectively.

		Predicted	
		Class A	Class B
Real	Class A	TP	FN
	Class B	FP	TN

4.2. Evaluation measures

To report the effectiveness of each method in the experiments, we have used evaluation measures based on the confusion matrix: accuracy and average accuracy [32]. Given a confusion matrix as Table 1 shows, the measures can be calculated according to Eqs. (1–5).

$$\text{Total} = TP + FP + FN + TN \tag{1}$$

$$\text{Specificity} = \frac{TN}{(FP + TN)} \tag{2}$$

$$\text{Sensitivity} = \frac{TP}{(TP + FN)} \tag{3}$$

$$\textbf{Accuracy} = \frac{TP + TN}{\text{Total}} \tag{4}$$

$$\textbf{Average Accuracy} = \frac{\text{Specificity} + \text{Sensitivity}}{2} \tag{5}$$
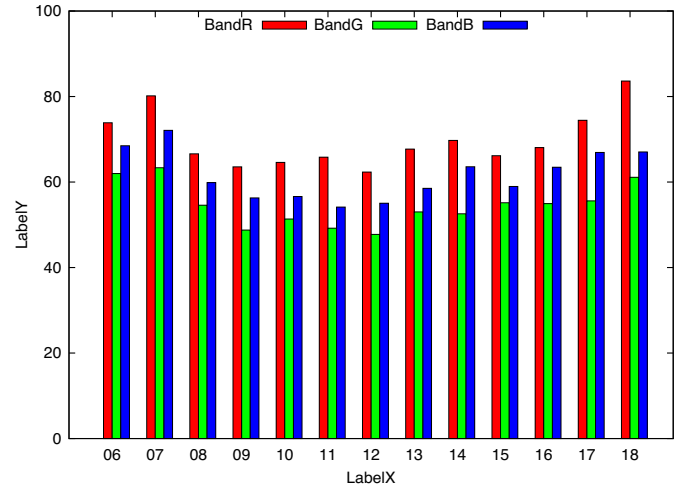
In the case of multi-class classification task with unbalanced datasets, the use of Average Accuracy avoids that the evaluation of learning methods are biased towards the majority class [32]. Since in our experiments, we have used the 5-fold cross-validation protocol, all reported results are in terms of Mean Accuracy or Mean Average Accuracy.

5. Results and discussion

This section presents five different performed experiments and discusses the obtained results. In Section 5.1, we compare different values for the k parameter of the kNN learning method that yields better results in the target problem. Since it is impracticable to use all and any learning method from the literature, we conduct a pre-processing of simple classifiers. In Section 5.2, we analyze the relationship between hours of the day and the six different plant species. This experiments is essential to identify differences between species. In Section 5.3, a correlation analysis between available time series-based classifiers is performed. Thus, we might measure the degree of agreement/disagreement between involved classifiers. In Section 5.4, we adapt the framework to combine different classifiers to consider the use of complementary information provided by RGB channels. Furthermore, we compare this framework to other well-known techniques from the literature (e.g., majority voting [22] and bootstrap aggregation [7]). Finally, in Section 3.2, we use the same framework with classifier selection process to reduce the number of time series-based classifiers used, while maintaining similar effectiveness results.

5.1. Finding the best kNN classifier

We have used eight k-Nearest Neighbors (kNN) methods [13], using $k \in \{1, 3, 5, 7, 9, 11, 13, 15\}$. Such methods are simple and fast, being suitable to be combined in a real-time recognition system. As in this paper we rely on the presence of several descriptors, using

Fig. 4. Mean accuracy results for all hours and RGB channels using kNN-1 as learning method.

all those learning methods might be costly. Therefore, we have conducted a study to find the best parameter k that yields good results in our approach. Through experiments, we observed that the best effectiveness performance was obtained for $k = 1$. From now on, all the experiments reported in this work considers kNN-1 as base classifiers inside the proposed framework. Fig. 4 shows the effectiveness results of the RGB channels for each hour of the day on the validation set V.

In these experiments, we observed that the best results were achieved for red channel (R) in all hours. We also observe that all classifiers performed better in extreme hours 6, 7, 17, and 18. These results are used in the next section to guide the correlation analysis between all 39 available time series-based classifiers (3 channels $\times$ 13 h = 39 classifiers).

5.2. Relation between hour and species

In these experiments, we have analyzed the behavior of each species (*A. tomentosum*, *C. brasiliensis*, *M. guianensis*, *M. rubiginosa*, *P. ramiflora*, and *P. torta*) throughout the day using the same validation set V used in the previous section. For these species, we use the same color patterns employed for their regions in Fig. 1(c). (*A. tomentosum* – green, *C. brasiliensis* – blue, *M. guianensis* – orange, *M. rubiginosa* – magenta, *P. ramiflora* – cyan, and *P. torta* – red).

Figs. 5 , 6, and 7 show the behavior of all species for each one of three different color channels (RGB). The x-axis refers to the hours of day (6, . . . , 18), while the y-axis refers to the mean average accuracy. As it can be observed, although all species have the same behavior with relation to the best results in the extreme hours, these behaviors might vary for each species. In Fig. 5, notice that the *C. brasiliensis* species (blue line) has a behavior more stable than the other curves. In contrast, the *M. guianensis* species (orange line) has the highest performance decrease between ranges 7–8 and the highest increase on the ranges 6–7 and 16–18.

In Fig. 6, we can observe that the curve for the *C. brasiliensis* species is more stable over the time of day. The curve of the *A. tomentosum* species (green line) differs from the other curves. It has a "U-like" shape, which indicates that time series-based classifiers yield better results for extreme hours.

In Fig. 7, we can see that the behavior of the curve of the *P. ramiflora* species (cyan line) has the highest performance decrease between ranges 8–9. Furthermore, we can observe that there are differences between R and B color channels with relation to the best results achieved at 6 and 18 h for all species. Classifiers using R color channel at 18 h achieved better results than those at 6 h. The contrary
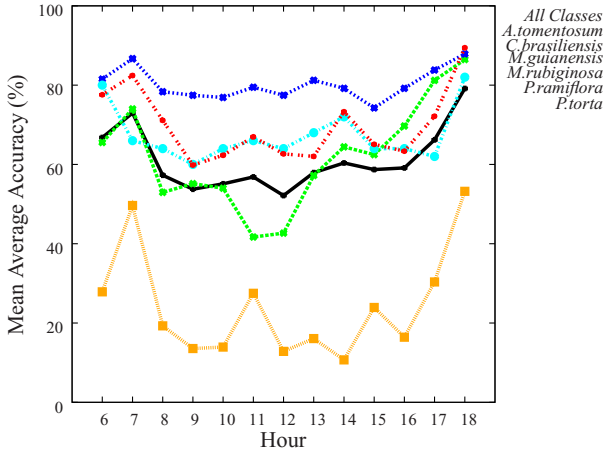
Fig. 5. Mean accuracy results for all hours and R color channel using *k*NN-1 as learning method for each class. (For interpretation of the references to color in this figure text, the reader is referred to the web version of this article.)

Fig. 7. Mean accuracy results for all hours and B channel using *k*NN-1 as learning method for each class. (For interpretation of the references to color in this figure text, the reader is referred to the web version of this article.)
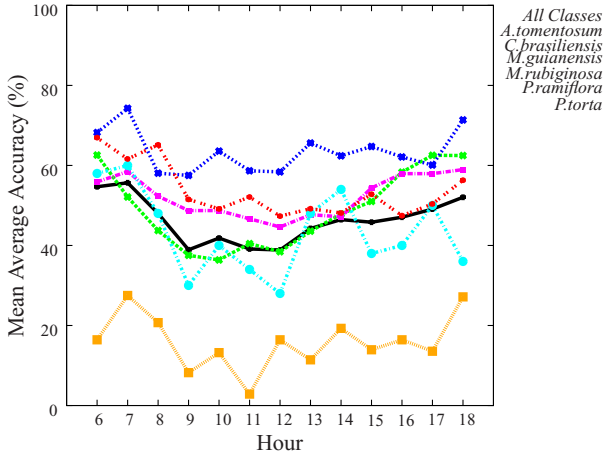

Fig. 6. Mean accuracy results for all hours and G channel using *k*NN-1 as learning method for each class. (For interpretation of the references to color in this figure text, the reader is referred to the web version of this article.)

was observed for the *G* channel. In this case, the best results were observed at 6 h. Note also that the curves of the species *M. guianensis* presents the lower mean average accuracy for all color channels. We might attribute these consistent difference to its leafing phenology: *M. guianensis* (orange line) is the only species reducing the percentage of green over the period of study [2].

In summary, each species has a particular behavior with regard to different RGB color channels throughout the day. This phenomenon might be justified by scattering of solar radiation and canopy reflectance [33]. Moreover, the leaf biochemical contents (e.g., chlorophyll, water, and dry matter) and canopy architecture (e.g., leaf area index, leaf angle distribution, and relative leaf size) might have impacted the spectral response of leaves [17].

This study might reinforce the importance of the *R* channel for the identification of these plant species as pointed out in [11]. However, the *G* channel has shown also to be important for some species (e.g., *M. guianensis*, *P. ramiflora*, and *P. torta*) for some hours of the day. The difference in behavior of the time series-based classifiers led us to consider their combination as a suitable alternative for improving the classification results in plant identification systems. We therefore performed a correlation analysis between all those classifiers to guide us in the time series-based classifier selection process.
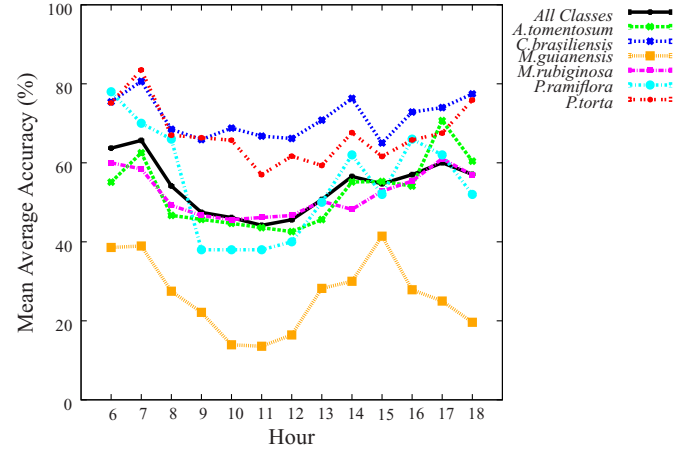
5.3. Correlation analysis between time series-based classifier

This section shows a correlation analysis of each pair of classifiers for all 39 available time series-based classifiers aiming at finding which of them might be combined by the framework described in Section 3.
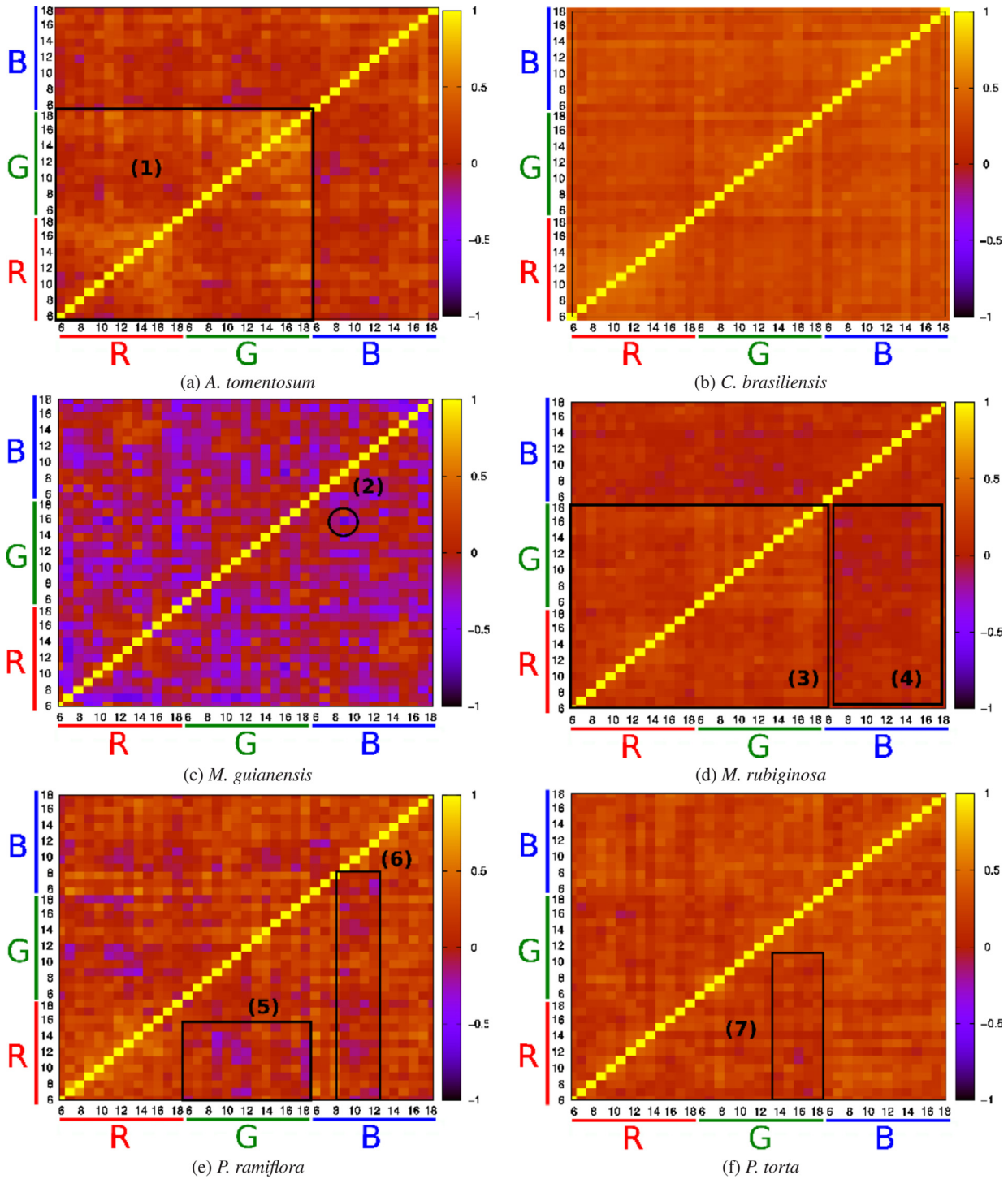
The *Correlation Coefficient* ρ (*COR*) [20] has been used to assess the correlation of two classifiers c_i and c_j:

$$COR(c_i, c_j) = \frac{ad - bc}{\sqrt{(a+b)(c+d)(a+c)(b+d)}}, \tag{6}$$

where a is the percentage of time series that both classifiers c_i and c_j classified correctly in a validation set V. Value b is the percentage of time series that c_j hit and c_i missed, c is the opposite of b. The value d is the percentage of time series that both classifiers missed. The pairs of classifiers with lower *COR* values have greater degree of complementarity and are more likely to yield better results when combined. Range of *COR* is in [−1, +1].

Fig. 8 presents the *COR* values for all possible combinations of pairs of classifiers considering the six classes. The lowest correlation coefficients are closer to the purple color (−1) and the highest coefficients are closer to the yellow color (+1). Furthermore, in this figure, there are seven important regions that have been highlighted and they are explained below.

As we can observe in Fig. 8(a), in region (1), there are few classifiers with high correlation between the channels *R* and *G* which are closer to their extreme hours (16–18 h). However, outside of region (1), the classifier from channels *R* and *G* are less correlated with channel *B*. We can see some points in purple, which means lower correlation coefficients. Fig. 8(b) shows a more homogeneous behavior of the classifiers, since instances of class *C. brasiliensis* are not difficult to be correctly classified. Furthermore, we can notice that there are a few yellow strips meaning high correlation of the same classifier with all other classifiers. Fig. 8(c) shows many purple points, which mean a low correlation between almost all classifiers used in this work. This phenomenon can be explained by the difficulty of classifying instances of the class *M. guianensis*. In region (2), we found the lowest correlation coefficient for all experiments. Fig. 8(d) shows a similar behavior to Fig. 8(a) with coefficients more homogeneous. We can observe in region (3) that classifiers of the same color channel *G* are more correlated between them. However, in region (4), *R* and *G*

Fig. 8. Correlation analysis considering all 39 available time series-based classifiers (3 channels × 13 h = 39 classifiers). The lowest correlation coefficients are closer to the purple color (−1) and the highest coefficients are closer to the yellow color (+1). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

are less correlated with B. Fig. 8(e) shows the second most difficult class to be classified: P. ramiflora. In regions (5) and (6), we found the lowest correlation coefficients between almost all classifiers of R × G and all classifiers of the color channels R and G with part of classifiers of channel B (9–12 h). Finally, in Fig. 8(f), which refers to the class P. torta, we can notice that the classifiers achieve similar behavior to Fig. 8(a) and (b). In region (7), there are the less correlated classifiers within Fig. 8(f).

5.4. Time series-based classifier fusion

In these experiments, 12 fusion techniques were compared: four techniques that use the adopted framework [12] (FSVM-*ALL*-39, FSVM-R-13, FSVM-G-13, and FSVM-B-13), four majority voting techniques [22] (MV-*ALL*-39, MV-R-13, MV-G-13, and MV-B-13), and four bootstrap aggregation approaches [7] (BAGG-MERGE-*ALL*-39, BAGG-MERGE-R-13, BAGG-MERGE-G-13, and BAGG-MERGE-B-13).
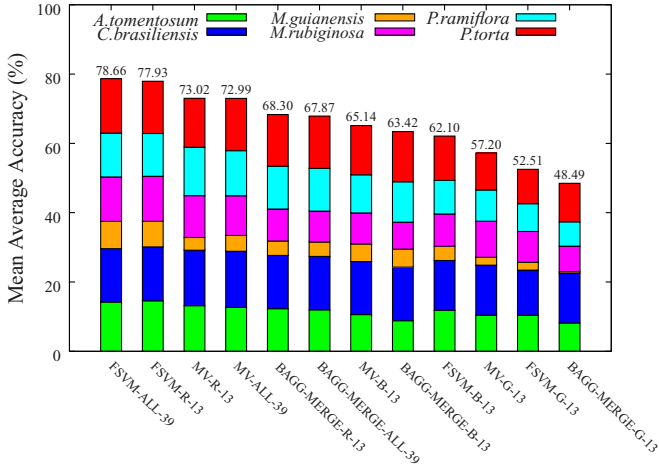
Fig. 9. Mean average accuracy of all fusion techniques.

Fig. 11. Behavioral analysis of the framework for different numbers of classifiers. On the x-axis are the numbers of selected classifiers $|\mathcal{C}^*| = \{2, \ldots, 39\}$ and y-axis are the mean accuracy achieved in the experiments. (For interpretation of the references to color in this figure text, the reader is referred to the web version of this article.)

ALL means that all color channels, hours, and features have been used. *R*, *G*, and *B* refer to the color channels that compose the RGB channel. Furthermore, the number after a name (e.g., 13 in FSVM-B-13) refers to amount of time series-based classifiers considered in the fusion process. MERGE denotes a binding of the all available channels, hours, and feature vectors.

The proposed framework aims at finding suitable combinations of time series-based classifiers formed by descriptors and learning methods. We have used the implementation of those learning methods available in the WEKA [16] data mining library. All learning methods were used with default parameters which means we did not optimize them.

Figs. 9 and 10 show the effectiveness results of all fusion techniques considered in this work. Two evaluation measures have been adopted, mean average accuracy (Fig. 9) e mean accuracy (Fig. 10). Balanced mean accuracy per class considers the mean accuracy of each class using the 5-fold cross validation protocol and the final effectiveness result is the average of these accuracies. Mean accuracy is the principal diagonal from confusion matrix, which counts the number of correct classification cases with respect to the total instances using the 5-fold cross validation protocol.

As it can be observed, two of our approaches (FSVM-*ALL*-39 and FSVM-R-13) achieved the best results among all involved fusion techniques. FSVM-*ALL*-39 approach obtained a mean average accuracy of
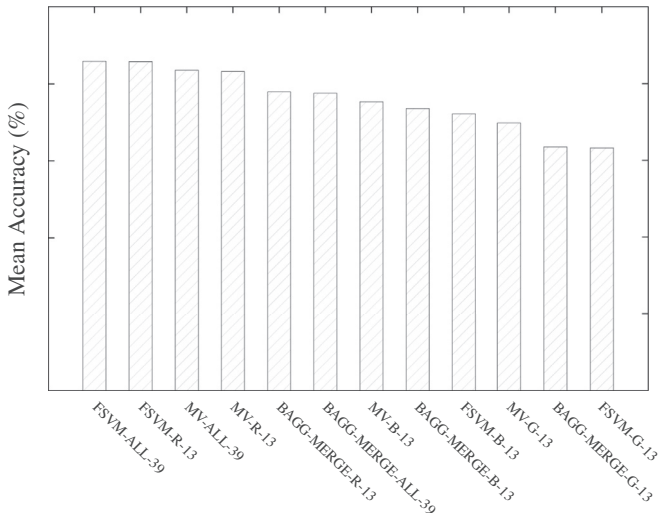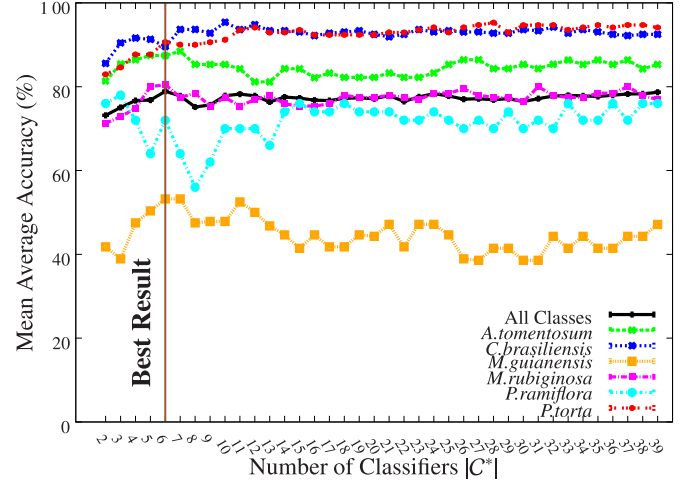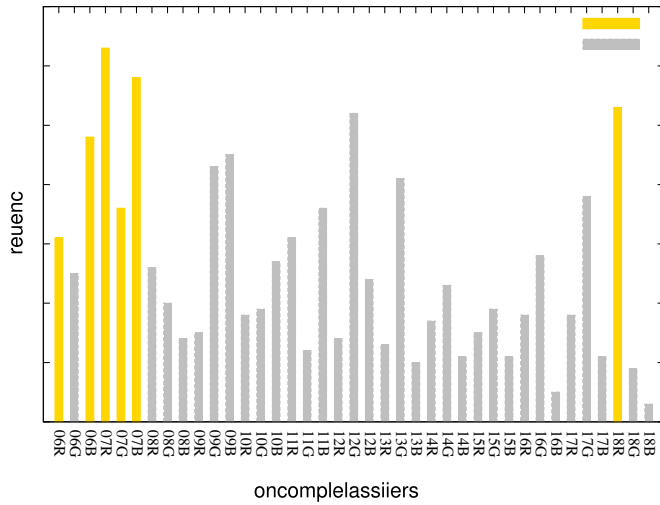
78.66% and mean accuracy of 85.87% against the best baseline MV-R-13 that yielded 73.02% and 83.18%, respectively. A fine-grained analysis considering mean average accuracy showed that FSVM-*ALL*-39 approach achieved better results in two classes (*M. guianensis* and *P. torta*), FSVM-R-13 approach is the best for other two classes (*A. tomentosum* and *M. rubiginosa*). Finally, MV-*ALL*-39 achieved better results in the classes *C. brasiliensis* and *P. ramiflora*. Thus, our approaches achieved better results in four out of six possible classes from the dataset used.

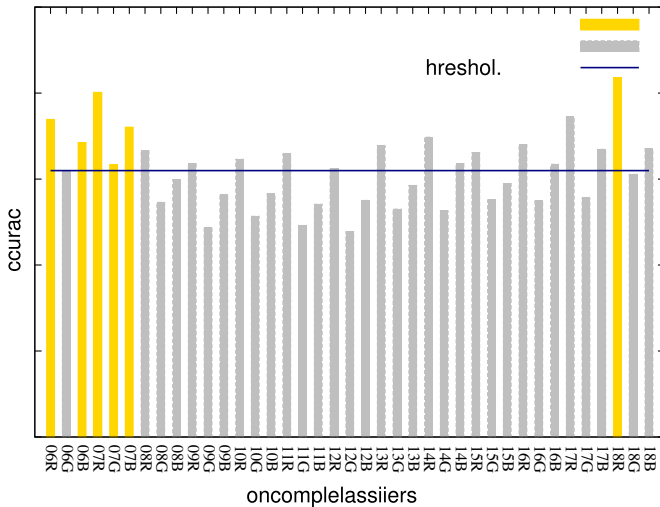### 5.5. Time series-based classifier selection

This section describes a behavioral analysis of the proposed framework using different numbers of time series-based classifiers $|\mathcal{C}^*|$ for the six classes from the dataset. Fig. 11 shows this analysis per class as well as all classes together.

In these experiments, we can notice that the classes *C. brasiliensis*, *A. tomentosum*, *M. rubiginosa*, and *P. torta* have a stable behavior with the increasing number of selected time series-based classifiers. This fact cannot be observed in the classes *M. guianensis* and *P. ramiflora*. The class *M. guianensis* achieved an increase of more than 10% of mean accuracy in the range [3, 7]. However, the curve of the class *P. ramiflora* decreases more than 15% for the same ranges. Furthermore, note that the framework using only six classifiers achieves similar results to those observed when 39 time series-based classifiers are used (see brown line in Fig. 11). In summary, the investigation for minimizing the misclassification rate per class seems to be a promising research venue and a classifier selection approach based on balanced classes might be a good solution to address this problem.

Fig. 12 depicts the histogram $\mathcal{H}$ created in the selection process, while Fig. 13 shows the accuracy performances of all simple/non-complex classifiers using the validation set *V*. We highlight in gold bars the six time series-based classifiers (see brown line in Fig. 11) that have been selected by our selection process. In Fig. 12, although classifiers *k*NN1-09G, *k*NN1-09B, and *k*NN1-12G have achieved the higher frequency than *k*NN1-06R and *k*NN1-07G, our selection approach does not choose any of those classifiers as candidate for fusion (gold bars). This is due to the policy of also considering the individual accuracy performance of classifiers in the selection process. As the accuracy performances of *k*NN1-09G, *k*NN1-09B, and *k*NN1-12G are below than the employed threshold values (dark blue line, $\mathcal{T} = 61.97\%$), these classifiers are not selected. Selected classifiers are *k*NN1-06R, *k*NN1-06B, *k*NN1-07R, *k*NN1-07G, *k*NN1-07B, and

Fig. 10. Mean accuracy of all fusion techniques.

Fig. 12. Histogram $\mathcal{H}$ related to the occurrence of classifiers in the selection process. (For interpretation of the references to color in this figure text, the reader is referred to the web version of this article.)

Fig. 13. Average accuracy performances of all non-complex classifiers used in our experiments. The dark blue line defines the employed threshold ($\mathcal{T}$) value. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

kNN1-18R. In these experiments, three selected classifiers use the *R* channel, two classifiers use the *B* channel, and one classifier uses the *G* channel. We can observe also a huge impact of the extreme hours (06, 07, and 18) in the selection process.

6. Conclusions

In this paper, we proposed a time series-based classifier selection and fusion framework to address problems in fine-grained plant identification tasks. We validated such framework with phenology studies. Recalling the objectives of our work, we addressed some very important questions, such as: (i) how species from the same leaf functional group differ from each other; (ii) how different pattern classifiers might be combined to improve the effectiveness results in target fine-grained species identification as just one classifier or feature would not be discriminatory enough for the task; and (iii) whether it is possible to achieve good classification results with fewer classifiers for plant identification improving the efficiency of the whole decision-making process. To answer such questions, we have performed five different analyses: (1) Investigation of the best

value for the k parameter of the kNN learning method that yields better results in the target problem; (2) An analysis of the relationship between hours of the day and the six different plant species considered in this study; (3) A correlation analysis measuring the degree of agreement/disagreement between available time series-based classifiers; (4) Use of a classifier selection and fusion framework aiming at exploring complementary information provided by different descriptions and learning methods, thus improving the effectiveness results; (5) Investigation of the impact of using the same framework with classifier selection process to decrease the number of time series-based classifiers used, while maintaining similar effectiveness results.

The experiments performed in this work confirm that there are some differences in terms of classification performance depending on the plant species considered, as well as, the correlation that exists between RGB channels and hours of the day using kNN-1 classifiers. Also, in the experiments, it was possible to note the importance of the red channel for plant species identification, in spite of the good contribution observed for the blue channel for some species (e.g., *M. guianensis*, *P. ramiflora*, and *P. torta*). Furthermore, the adopted framework achieved excellent results when compared with other well-known fusion techniques of the literature (Bootstrap Aggregation-BAGG [7] and Majority Voting-MV [22]). This framework, extended upon [12], brings an essential and important property different from those techniques in the literature: it is highly flexible and parallelizable. Regarding flexibility, our adopted approach can use any description technique or learning methods as base classifier. In addition, the method is able to use any learning method in the late fusion (meta-learning process), such as a simple majority voting or more complex techniques such as support vector machines (SVMs). As for parallelization, each combination of description and learning methods can be used in different parallel approaches (e.g., thread, processor, and GPU) due the framework structure being designed independently (in modules). With that, the more important modules are the training classifiers, the classifier selection, and the meta-learning. Furthermore, other important advantage of the adopted approach is that it considers each description and learning method as a single representation thus it does not incur in the common normalization problems and, consequently reduces the risk the "the curse of dimensionality" that may result of direct combination of features (e.g., by means of concatenating feature vectors).

A fine-grained analysis using time series-based classifier selection showed that the framework using fewer classifiers (i.e., six) achieved similar results to those observed when 39 time series-based classifiers are used. Therefore, our proposed method might be able to perform real-time plant identification in different domains such as when monitoring plants through intelligent autonomous vehicles (e.g., unmanned aerial vehicles and drones).

Another important contribution of this work is related to the use of multiclass classification approaches, which was not used in previous phenology work [5]. The main take-home message of our work is that the adopted framework might be a good solution to address complex problems such as the ones involving phenology derived from sequential digital images and fine-grained recognition of plants from the same functional leaf groups. This framework takes advantage of different and complementary information provided by RGB channels and combine them for a better decision-making process. In addition, it achieves good effectiveness results with fewer classifiers, which is paramount for efficient plant recognition.

Future work may consider the use of the proposed framework in semi-automatic plant identification tasks, in which users may provide relevance feedback that may tune the used classification models for better fine-grained categorization that always includes humans (specialists) in the loop. Also, we may consider the use of the proposed framework for real-time plant identification systems. Furthermore, other vegetation indices might be studied and combined

such as Color Index of Vegetation Extraction (CIVE) [19], Excess Green (ExG) [14], and the Plant Phenology Index (PPI) [18]. Regarding the generalization, with the proposed approach in this work for dealing with time-series based problems, we believe that other application domains, still in the realm of time series as input data, may include crop monitoring, recognizing different plant growth rates and also for plague control. Other applications also may tap on snowmelt and comparison of species or different kinds of vegetation.

Acknowledgments

References

[1] H. Ahrends, S. Etzold, W. Kutsch, R. Stoeckli, R. Bruegger, F. Jeanneret, H. Wanner, N. Buchmann, W. Eugster, Tree phenology and carbon dioxide fluxes: use of digital photography for process-based interpretation at the ecosystem scale, Clim. Res. 39 (2009) 261–274.

[2] B. Alberton, J. Almeida, R. Henneken, R. da S. Torres, A. Menzel, L.P.C. Morellato, Using phenological cameras to track the green up in a Cerrado savanna and its on-the-ground validation, Ecol. Inform. 19 (2014) 62–70.

[3] J. Almeida, J.A. dos Santos, B. Alberton, L.P.C. Morellato, R. da S. Torres, Plant species identification with phenological visual rhythms, in: Proceedings of the IEEE International Conference on eScience (eScience'13), 2013, pp. 148–154.

[4] J. Almeida, J.A. dos Santos, B. Alberton, L.P.C. Morellato, R. da S. Torres, Visual rhythm-based time series analysis for phenology studies, in: Proceedings of the IEEE International Conference on Image Processing (ICIP'13), 2013, pp. 4412–4416.

[5] J. Almeida, J.A. dos Santos, B. Alberton, R. da S. Torres, L.P.C. Morellato, Applying machine learning based on multiscale classifiers to detect remote phenology patterns in Cerrado savanna trees, Ecol. Inform. 23 (2014) 49–61.

[6] J. Almeida, J.A. Santos, W.O. Miranda, B. Alberton, L.P.C. Morellato, R.S. Torres, Deriving vegetation indices for phenology analysis using genetic programming, Ecol. Inform. 26 (2015) 61–69.

[7] L. Breiman, Bagging predictors, Mach. Learn. 24 (1996) 123–140.

[8] L.E. Chambers, R. Altwegg, C. Barbraud, P. Barnard, L. Beaumont, R. Crawford, J.M. Durant, L. Hughes, M.R. Keatley, M. Low, L.P.C. Morellato, E. Poloczanska, V. Ruoppolo, R.E.T. Vanstreels, E.J. Woehler, A. Wolfaardt, Changes in southern hemisphere phenology, PlosOne 8 (2013) e75514.

[9] H.D. Cheng, X. Jiang, Y. Sun, J. Wang, Color image segmentation: advances and prospects, Pattern Recognit. 34 (2001) 2259–2281.

[10] E.E. Cleland, I. Chuine, A. Menzel, H.A. Mooney, M.D. Schwartz, Shifting plant phenology in response to global change, Trends Ecol. Evol. 22 (7) (2007) 357–365.

[11] J.C. Conti, F.A. Faria, J. Almeida, B. Alberton, L.P.C. Morellato, L. Camolesi Jr., R. da S. Torres, Evaluation of time series distance functions in the task of detecting remote phenology patterns, in: Proceedings of the IEEE International Conference on Pattern Recognition (ICPR'14), 2014, pp. 3126–3131.

[12] F.A. Faria, J.A. dos Santos, A. Rocha, R. da S. Torres, A framework for selection and fusion of pattern classifiers in multimedia recognition, Pattern Recognit. Lett. 39 (2014) 52–64.

[13] J. Friedman, T. Hastie, R. Tibshirani, The Elements of Statistical Learning, first ed., Springer, 2001.

[14] C. Gée, J. Bossu, G. Jones, F. Truchetet, Crop/weed discrimination in perspective agronomic images, Comput. Electron. Agric. 60 (2008) 49–59.

[15] L. Guigues, J. Cocquerez, H. Le Men, Scale-sets image analysis, Int. J. Comput. Vis. 68 (2006) 289–317.

[16] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, I.H. Witten, The weka data mining software: an update, ACM SIGKDD Explor. Newsl. 11 (2009) 10–18.

[17] S. Jacquemoud, W. Verhoef, F. Baret, C. Bacour, P.J. Zarco-Tejada, G.P. Asner, C. Franois, S.L. Ustin, Prospect + sail models: a review of use for vegetation characterization, Remote Sens. Environ. 113 (Suppl. 1) (2009) S56–S66.

[18] H. Jin, L. Eklundh, A physically based vegetation index for improved monitoring of plant phenology, Remote Sens. Environ. 152 (2014) 512–525.

[19] T. Kataoka, T. Kaneko, H. Okamoto, S. Hata, Crop growth estimation system using machine vision, in: Proceedings of the 2003 IEEE/ASME International Conference on Advanced Intelligent Mechatronics, vol. 2, 2003, pp. b1079–b1083.

[20] L.I. Kuncheva, Combining Pattern Classifiers: Methods and Algorithms, Wiley-Interscience, 2004.

[21] L.I. Kuncheva, C.J. Whitaker, Measures of diversity in classifier ensembles and their relationship with the ensemble accuracy, Mach. Learn. 51 (2003) 181–207.

[22] L. Lam, S.Y. Suen, Application of majority voting to pattern recognition: an analysis of its behavior and performance, IEEE Trans. Syst., Man, Cybern., Part A: Syst. Hum. 27 (1997) 553–568.

[23] L.P.C. Morellato, R.R. Rodrigues, H.F. Leit ao Filho, C.A. Joly, Estudo comparativo da fenologia de espécies arbóreas de floresta de altitude e floresta mesófila semidecídua na serra do iapí, jundiaí, são paulo, Braz. J. Bot. 12 (1989) 85–98.

[24] M. Parry, O. Canziani, J. Palutikof, P. van der Linden, C. Hanson, IPCC (2007) Climate Change 2007: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge UK: Cambridge University Press, 2007.

[25] P. Reys, M.G.G. Camargo, A.P. Teixeira, M.A. Assis, M.T. Grombone-Guaratini, L.P.C. Morellato, Estrutura e composio florstica entre borda e interior de um cerrado sensu stricto e sua importância para propostas de recuperação, Hoennea 40 (2013) 437–452.

[26] A.D. Richardson, B.H. Braswell, D.Y. Hollinger, J.P. Jenkins, S.V. Ollinger, Near-surface remote sensing of spatial and temporal variation in canopy phenology, Ecol. Appl. 19 (2009) 1417–1428.

[27] A.D. Richardson, J.P. Jenkins, B.H. Braswell, D.Y. Hollinger, S.V. Ollinger, M.L. Smith, Use of digital webcam images to track spring greep-up in a deciduous broadleaf forest, Oecologia 152 (2007) 323–334.

[28] R. da S. Torres, M. Hasegawa, S. Tabbone, J. Almeida, J.A. dos Santos, B. Alberton, L.P.C. Morellato, Shape-based time series analysis for remote phenology studies, in: Proceedings of the IEEE International Geoscience and Remote Sensing Symposium (IGARSS'13), 2013, pp. 3598–3601.

[29] R. da S. Torres, C.B. Medeiros, M.A. Gonçalves, E.A. Fox, A digital library framework for biodiversity information systems, Int. J. Digit. Libr. 6 (2006) 3–17.

[30] L.C.B. Santos, J. Almeida, J.A. dos Santos, S.J.F. Guimar aes, A.A. Araújo, B. Alberton, L.P.C. Morellato, R. da S. Torres, Phenological event detection by visual rhythm dissimilarity analysis, in: Proceedings of the IEEE International Conference on eScience (eScience'14), 2014, pp. 263–270.

[31] M.D. Schwartz, Phenology: An Integrative Environmental Science, Springer, 2013.

[32] M. Sokolova, G. Lapalme, A systematic analysis of performance measures for classification tasks, Inf. Process. Manag. 45 (2009) 427–437.

[33] M.M. Verstraete, Radiation transfer in plant canopies: scattering of solar radiation and canopy reflectance, J. Geophys. Res.: Atmos. 93 (1988) 9483–9494.

[34] D.M. Woebbecke, G.E. Meyer, K. Von Bargen, D.A. Mortensen, Color indices for weed identification under various soil, residue, and lighting conditions, Trans. ASAE 38 (1995) 259–269.