A Framework for Pattern Classifier Selection and Fusion

Fabio A. Faria*, Anderson Rocha, and Ricardo da S. Torres

Institute of Computing, University of Campinas (Unicamp)
Cidade Universitária "Zeferino Vaz",
Campinas, SP, Brazil – CEP 13083-852
{ffaria,anderson.rocha,rtorres}@ic.unicamp.br
http://www.ic.unicamp.br/~ffaria

Abstract. In this work, we propose a framework for classifier selection and fusion. Our method seeks to combine image characterization and learning methods by means of a meta-learning approach responsible for assessing which methods contribute more towards the solution of a given problem. The framework uses three different strategies of classifier selection that pinpoint the less correlated, yet effective, classifiers through a series of diversity measure analysis. The experiments show that the proposed approaches yield comparable results to well-known algorithms from the literature on many different applications but using less learning and description methods as well as not incurring in the curse of dimensionality and normalization problems common to some fusion techniques. Furthermore, our approach yields effective classification results using very reduced training sets.

Keywords: meta-learning; diversity measure; rank aggregation; kendall correlation

1 Motivation

The huge amount of visual data created through the popularization of mobile devices (e.g., cell phone, camera, and tablet), makes us face many new challenges unthinkable two decades ago. Image and video classification tasks have been inserted in different and complex applications and the use of machine learning-based solutions has become the most popular approach to several applications. However, there is no single solution (learning or extraction technique) that solves all the problems. Depending on the extraction and learning methods used might create different classifiers that provide complementary information. One common strategy that has been used to take advantage of these complementary information and improve classification results is the Multiple Classifier System (MCS). In MCS, the diversity of classifiers is an essential factor to reach better effectiveness

* The author thank São Paulo Research Foundation - FAPESP (grant 2010/14910-0) 2010-2014 and CAPES (grant 1260-12-0) for the financial support.

results [17]. Diversity measures assess the degree of agreement/disagreement between classifiers and might identify potential classifiers for fusion. In this sense, Kuncheva and Whitaker [18] studied different diversity measures as well as discussed their impacts on the final accuracy of ensemble systems. Different works have been using diversity measures to select appropriate high-performance classifiers, but the challenge of finding the optimal number of classifiers for a target task has not been properly addressed yet. In general, the proposed solutions rely on the a priori use of ad hoc strategies for selecting classifiers, followed by the evaluation of their effectiveness results during training. Searching by the optimal number of classifiers, however, makes the selection process more expensive.

Currently, some of the most important challenges in MCS involve: choosing the best diversity measure to be used; combining different available measures for classifier selection in an ensemble system; and finding out whether or not the existing measures describe the "real" diversity within the ensemble systems [4]. Typically, works in the literature have adopted a single diversity measure or combined different measures using simple strategies (e.g., based on average of the classification scores [5]). However, the aforementioned methods might not take full advantage of the different opinions provided by all of the available diversity measures. Moreover, another problem in MCS approaches is how to combine different and non-correlated extraction and learning methods automatically.

In the literature, many works have been proposed to try sorting out problems cited previously as for example, the well-known AdaBoost [13] and Bagging [2] approaches. AdaBoost and Bagging ensemble approaches have been used in several works in the literature due to their good results achieved in diverse applications. However, previous work has also shown their limitations in terms of efficiency, normalization, overfitting, and feature dimensionality problems. In [21], for example, training time has been a concern when more features were used to train an AdaBoost algorithm for face localization. The same problem has been reported in [19], which trained an AdaBoost algorithm for tracking indoor soccer players using videos. In [16], the authors discuss another problem: the sensitivity of the classical AdaBoost algorithm to noisy datasets. They have proposed different solutions to reduce the overfitting effect caused in those cases. In [20], the authors discuss the problems of feature normalization in the context of combining classifiers. More detail about tracking down fusion and classification problems can be found in [6].

The combination of multiple feature vectors defined by different image descriptors in AdaBoost and Bagging approaches is usually based on their concatenation (feature binding). Usually, when performing feature binding of different nature/domain, normalization techniques should be applied to standardize all feature values in the same range, which is a very challenging task [20]. Another common problem faced when features are concatenated refers to the "curse of dimensionality" [22]. The curse of dimensionality problem is related to the fact that the dimension of the feature space increases in such a way that the available training instances become indistinguishable and it is not enough for allowing the definition of a good decision hyperplane [1].

2 Objective and Contributions

In this work, we seek an alternative to AdaBoost and Bagging ensembles, which might suffer curse of dimensionality and normalized problems. Our objective is to propose a stacking framework, able to perform automatic fusion of different visual properties (color, texture, and shape) and learning methods in existence in the literature for different multimedia recognition tasks.

The framework assesses several descriptors and learning methods performing fusion in a final stage (late fusion) using a low-dimension feature vector and simple (fast) classifiers. Another difference of the proposed method, when compared to AdaBoost and Bagging techniques, is that the proposed framework seeks greater diversity between the simple classifiers being able to choose only the ones that effectively contribute to the solution of the classification problem of interest.

Diversity may be obtained in different ways such as using: (a) different learning methods and the same training set; (b) the same learning method and different training samples; (c) different methods using different types of classifier outcomes during the combination; and (d) predictions as new attributes to train some learning method (meta-learning). In this work, we use two out of four ways (a and d). We also use different visual properties (color, texture, and shape) to each of the learning methods chosen to be simple classifiers. We follow the concept that two instances of the same class have similar classification outputs for the same set of classifiers [14].

In this regard, in this work, we investigate the combination of several learning methods and image descriptors aiming at creating more effective classifiers. We propose a framework for automatically combining the most discriminative classifiers using the support vector machine (SVM) technique, as well as exploring the use of diversity measures to select the less-correlated, yet effective, classifiers in three different selection strategies. We have performed experiments that demonstrate that the proposed framework for classifier fusion yields comparable results to the traditional fusion approaches but using less learning and description methods as well as not incurring in the curse of dimensionality problems, which are common to some fusion techniques. Another major advantage of the proposed method is that it yields good classification results using small training examples being more robust to the small sample size problem common in many classification techniques [1].

Our research hypothesis is that: *Appropriate classifier selection approaches can take advantage of classifier diversity to improve the accuracy performance of multiple classifier systems.*

The contributions and publications directly related to this thesis are: a framework for classifier fusion through a meta-learning approach using Support Vector Machines techniques [11]; a new classifier selection approach based on diversity measures consensus [9, 10]; a new classifier selection approach based on *Kendall* correlation analysis [12]; and a new classifier selection approach based on rank aggregation techniques [8]; a multimodal framework for automatic identification of fruit flies (Diptera: Tephritidae) [7].

3 The Classifier Fusion Framework

The objective of the fusion framework [10] is to exploit the degree of agreement/disagreement among classifiers, concept known as diversity, to select the most suitable ones to be used in a combination scheme.

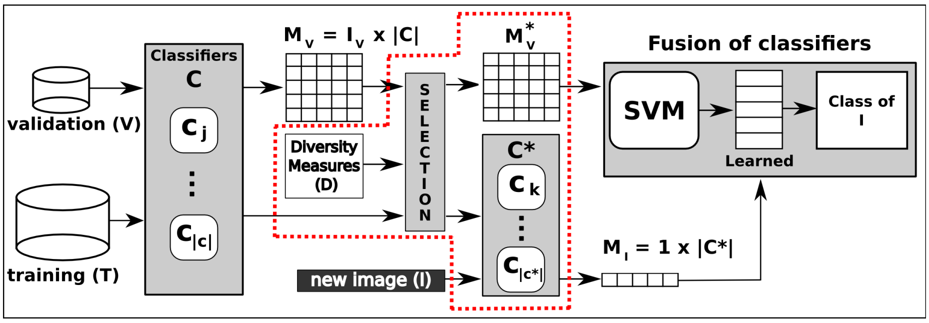

Fig. 1: Framework for classifier selection and fusion [10]. The classifier selection process is delimited by the dashed red line.

3.1 Formalization

Let $\mathcal{C}$ be the set of classifiers $\mathcal{C} = \{c_{11}, c_{12}, \ldots, c_{22}, \ldots, c_{|\mathcal{L}||\mathcal{F}|}\}$, with $c_{ij} = (l_i, f_j)$, where l_i is a learning method (e.g., Decision Tree, Naïve Bayes, kNN, etc.), and f_j is an image descriptor (e.g., Color Histogram). $|\mathcal{C}| = |\mathcal{L}| \times |\mathcal{F}|$, where $\mathcal{L}$ and $\mathcal{F}$ are sets of available learning methods and image descriptors, respectively. Initially, all classifiers $c_k \in \mathcal{C}$ $(1 < k \leq |\mathcal{C}|)$ are trained on a training set T. Next, classifier results on a validation set V are computed and stored into a matrix M_V, where $|M_V| = |V| \times |\mathcal{C}|$ and $|V|$ is the number of regions in a validation set V. The actual classes of training and validation data points are known *a priori*.

The objective of our framework is to select a set $\mathcal{C}^* \subset \mathcal{C}$ of classifiers that are good candidates to be combined. $\mathcal{C}^*$ is determined by using M_V as input in an approach that exploits diversity measures (see Section 3.2). Note that $\mathcal{C}^*$ can be used to compute a new matrix $M_V^* \subset M_V$. Each selected classifier in $\mathcal{C}^*$ is used to determine the class of an unknown instance. The outcomes of those classifiers are later combined by a novel fusion technique (majority voting, SVM, etc.), which is responsible for defining the class of the unknown instance. Fig. 1 illustrates the framework FSVM for combining classifiers.

3.2 Selection based on Consensus

Fig. 2 illustrates the adopted five-step approach for selecting classifiers based on diversity measures, previously introduced in [10]. First, diversity measures (set $\mathcal{D}$ in Fig. 2) are used to assess the degree of agreement among available classifiers in $\mathcal{C}$ by taking into account the M_V matrix previously computed. That step is represented by arrow (a) in Fig. 2. 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 $\mathcal{H}$ 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 $\mathcal{H}$, 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.

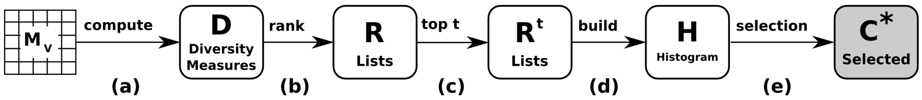

Fig. 2: The five steps for classifier selection are: (a) Computation of diversity measures from the validation matrix M_V; (b) Ranking of pairs of classifiers by their diversity measure scores; (c) Selection of the top $t = 100$ ranked pairs of classifiers; (d) Computation of a histogram $\mathcal{H}$ that counts the number of occurrences of each classifier; (e) Selection of classifiers $|\mathcal{C}^*|$ based on their occurrence in $\mathcal{H}$ and on a defined threshold $\mathcal{T}$ [10].

3.3 Selection based on Kendall Correlation

Let $\mathcal{C}$ be the set of classifiers created by the combination of learning methods and image descriptors. Let $\mathcal{P} = \{p_1, p_2, \ldots, p_{|\mathcal{C}\times\mathcal{C}|}\}$ be a set of all possible pairs of classifiers, i.e., $p_l = (c_i, c_j)$, where $(c_i, c_j) \in \mathcal{C} \times \mathcal{C}$. Let $\mathcal{D} = \{d_1, d_2, \ldots, d_{|\mathcal{D}|}\}$ be a set of diversity measures, such that each diversity measure $d_k \in \mathcal{D}$ defines a distance function $\rho : \mathcal{P} \to \mathbb{R}$, where $\mathbb{R}$ denotes real numbers. Equations described in paper [18] that define different criteria for implementing the function ρ. Consider $\rho(p_l) \geq 0$ for all $p_l \in \mathcal{P}$ and $\rho(p_l) = 0$, with $p_l = (c_i, c_j)$, if $c_i = c_j$. The distance $\rho(p_l)$ among all pairs of classifiers $p_l = (c_i, c_j) \in \mathcal{C} \times \mathcal{C}$ can be computed to obtain a $|\mathcal{C}| \times |\mathcal{C}|$ distance matrix A. Given a diversity measure $d_k \in \mathcal{D}$, we can compute a ranked list $\mathcal{R}_{d_l}$ by taking into account the distance matrix A. The ranked list $\mathcal{R}_{d_l}$=$\{p_1, p_2, \ldots, p_{|\mathcal{C}\times\mathcal{C}|}\}$ (where $p_l = (c_i, c_j)$ is a pair of classifiers) can be defined as a permutation of the collection $\mathcal{P}$, such that, if p_l is ranked at lower positions than p_m, i.e., p_l is ranked before p_m, then $\rho(p_l)$ $< \rho(p_m)$. In this way, pairs of classifiers are ranked according to their agreement score defined in terms of a diversity measure.

We exploit the correlation of ranked lists of pairs of classifiers to select the more appropriate ones to be combined. In this thesis, we use the *Kendall* tau rank correlation coefficient (τ) [15] to measure the degree of concordance between two different ranked lists of the same set of observed samples. The *Kendall* correlation $\tau(\mathcal{R}_{d_i}, \mathcal{R}_{d_j})$ between two ranked lists $\mathcal{R}_{d_i}$ and $\mathcal{R}_{d_j}$ is defined in terms of the number of concordant pairs NC in $\mathcal{R}_{d_i}$ and $\mathcal{R}_{d_j}$, the number of discordant pairs ND, and the number of positions n in the ranked lists.

We propose a novel strategy, named *Kendall* classifier selection (KCS), to define appropriate classifiers to be used in the classification framework presented in [10]. KCS makes use of the degree of agreement of different diversity measures.

This agreement is measured in terms of the *Kendall* correlation among ranked lists of classifiers.

Let d_{H_1} and d_{H_2} be the diversity measures with the highest correlation scores, which are defined by the *Kendall* correlation. Let $\mathcal{R}_{d_{H_1}}$ and $\mathcal{R}_{d_{H_2}}$ be the ranked lists of pairs of classifies defined by d_{H_1} and d_{H_2}, respectively. KCS defines the top-ranked pairs of classifiers in $\mathcal{R}_{d_{H_1}}$ and $\mathcal{R}_{d_{H_2}}$ as the most appropriate ones to be used in the classification framework presented in [10]. We also tested in our experiments selected classifiers defined in terms of the lowest correlated diversity measures (d_{L_1} and d_{L_2}). In this case, we use classifiers defined in the top-ranked positions of $\mathcal{R}_{d_{L_1}}$ and $\mathcal{R}_{d_{L_2}}$.

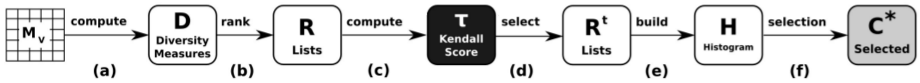

Fig. 3: The six steps for new classifier selection are: (a) Compute diversity measures from the validation matrix M_V; (b) Sort $\mathcal{R}$ lists by diversity measure scores; (c) Compute *Kendall* correlation coefficients among all ranked lists of classifiers $\mathcal{R}$; (d) Select $\mathcal{R}_{d_{H_1}}$ and $\mathcal{R}_{d_{H_2}}$ or $\mathcal{R}_{d_{L_1}}$ and $\mathcal{R}_{d_{L_2}}$ ranked lists to be used in the next step; (e) $\mathcal{R}^t$ lists with top $t = 100$; (f) Compute a histogram $\mathcal{H}$ that counts the number of occurrences of each classifier; (g) Select the most appropriate classifiers $|\mathcal{C}^*|$ based on their occurrence in $\mathcal{H}$ and a defined threshold $\mathcal{T}$ [12].

Figure 3 summarizes in six steps the new approach for selecting classifiers based on *Kendall* correlation. It is important to highlight that all steps regarding the selection of classifiers for fusion are performed during the training phase of the decision-making framework. Using a validation set separated during training allows us to evaluate different descriptors and learning techniques, assess their outcomes when classifying the validation examples, and properly selecting, by means of the proposed *Kendall*-based methodology, the most suitable classifiers for deployment during testing.

3.4 Selection based on Rank Aggregation

We propose to use multiple diversity and evaluation measures (Kappa, Tau, and accuracy) to determine which classifiers should be combined to improve the classification results in a given problem. Recall that different diversity measures would rank pairs of classifiers differently. In many situations, *rank aggregation* methods have been used as a way of obtaining a consensus ranking when multiple ranked lists are computed by different approaches. Rank aggregation has also been treated as the task of combining different ranked lists (or scores) in order to obtain a single, and more accurate, ranked list. For classification tasks, the combination with the lowest error occurs when the classifiers being combined are non-correlated (high diversity) and yields high accuracy rate [3]. In our approach, each considered measure (both diversity and evaluation measures) produces a ranked list of pairs of classifiers. A rank aggregation method combines all ranked lists, producing a single combined ranked list, which is used to identify pairs of classifiers with good classification performance and high diversity. In the next section, we formally define the proposed rank aggregation approach.

Figure 4 summarizes the six-step approach for selecting classifiers based on rank aggregation.

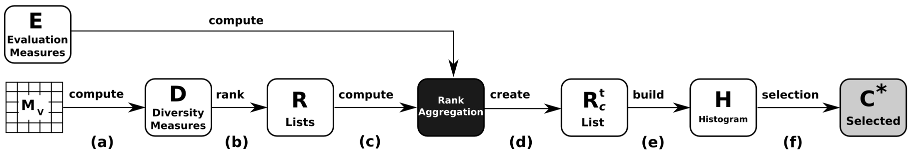

Fig. 4: The six steps of the new classifier selection are: (a) Compute diversity measures from the validation matrix M_V; (b) Sort $\mathcal{R}$ lists according to scores of diversity measures; (c) Compute rank aggregation using all ranked lists of classifiers ($\mathcal{R}$) and evaluation measures (E); (d) Create a single list $\mathcal{R}_c^t$, which list has the top $t = 100$; (e) Compute a histogram $\mathcal{H}$ that counts the number of occurrences of each classifier; (f) Select the most appropriate classifiers $|\mathcal{C}^*|$ that satisfy a defined threshold $\mathcal{T}$ [8].

4 Experiments and Discussion

This section presents some of several experiments that we performed to evaluate the robustness of our fusion framework with each selection process [8, 10, 12].

4.1 Effectiveness Analysis

In these experiments, six fusion techniques were compared: our approach using SVM (FSVM-PK-49) considering $|C| = 49$, two Adaboost approaches (BOOST-DEFAULT and BOOST-49), Bagging (BAGG-49), and Majority Voting (MV-49). Recall that using $|C| = 49$ means that all available classifiers (7 learning methods $\times$ 7 image descriptors) are employed in the fusion process. FSVM-PK means the SVM technique uses a polynomial kernel to combine different simple classifiers in our approach. Furthermore, we have included the best single classifier (no fusion) between all tested learning methods.

Table 1 presents the results obtained for each fusion technique and the best single classifier using one of four datasets considered in the work and considering three different evaluation measures (Accuracy, Kappa, and Tau). Notice that BOOST and BAGG techniques show up with the suffix ALL, which means the concatenation of the feature vectors produced by the seven different image descriptors considered. Thus BAGG-49-ALL and BOOST-49-ALL techniques refer to the use of 49 iterations and seven image descriptors (hybrid fusion).

In these experiments, our late fusion approach (FSVM-PK-49), which uses meta-learning on the outputs of all available classifiers yielded a slightly better classification result considering the three evaluation measures, when compared to other techniques in any tested datasets. However, the achieved results when considering the selection of the most appropriate descriptors and learning methods automatically during the fusion process are more interesting.

Important to note that BOOST and BAGG techniques use a fusion hybrid (feature and decision level fusion) to achieve similar results to our framework that uses only decision level fusion. We also computed the confidence intervals to verify if the results obtained by the proposed framework differ from those observed for the baselines. We could observe that FSVM-PK-49 has no statistical difference between the best baseline in the Caltech dataset. Furthermore, our late fusion approach (FSVM-PK-49) achieves similar results to BOOST-49-ALL (hybrid fusion).

Dataset	**Techniques**	**Measures**		
		Accuracy	**Kappa**	**TAU**
Caltech	**FSVM-PK-**49	**47.05%±1.77**	**0.45±0.02**	**0.46±0.02**
	BOOST-49-ALL	46.90%±0.63	0.45±0.01	0.46±0.01
	BAGG-49-ALL	43.01%±1.38	0.41±0.01	0.42±0.01
	SVM-PK-LAS	41.30%±0.41	0.39±0.00	0.40±0.00
	MV-49	41.02%±0.46	0.38±0.00	0.40±0.00
	BOOST-DEFAULT-ALL	39.92%±0.57	0.38±0.01	0.39±0.01

Table 1: Classification effectiveness of the proposed framework and baselines [10].

4.2 Training Set Size Impact

This section shows a behavioral study among the classifiers compared in Table 1 using reduced training sets. In our experiments, we conducted a study considering five different sizes for the training set (T): $8\%, 16\%, 33\%, 67\%, 100\%$, which represents 5%, 10%, 20%, 40% and 60% of the entire datasets, respectively. These subsets have been selected from original training set. We use again the 5-fold cross-validation protocol previously adopted in our experiments.

Fig. 5 shows the results for one of four datasets (Caltech) used in our work. The x-axis denotes the number of images in the training set and the y-axis represents the average accuracy in the testing set. The FSVM-PK-49 approach using a subset of 8% of training set achieves 39.52% of accuracy. In the same training set, BOOST-49-ALL yields 32.33%, which means that our approaches have a gain of more than 19% compared to the best baseline. In the subset 16%, our approaches are still better and achieve accuracy results of 40.67% (FSVM-PK-49) against 37.24% of the BOOST-49-ALL. That represents a gain of more than 7% in classification accuracy. From the subset 33% to 100%, the best baseline yields similar performance to our approach. In summary, we can see that the proposed approach are able to learn from small training sets.

4.3 Classifier Selection Approaches

This section discusses the results regarding the effectiveness and efficiency of the proposed framework using one of the two different datasets performed in our work. In this case, the Urban dataset has been used. In our experiments, we have used *Double-Fault Measure* (DFM), *Q-Statistic* ($QSTAT$), *Interrater Agreement k* (IA), *Correlation Coefficient* ρ (COR), and *Disagreement Measure* (DM). Our framework is denoted as FSVM-NORM-$|\mathcal{C}^*|$, where NORM means the normalized polynomial SVM kernel used in our experiments and $|\mathcal{C}^*|$ is number of simple classifiers that will be combined by the SVM-based meta-learning technique.

Table 2 shows the average kappa indices for all performed experiments with Urban dataset. The columns refer to the number of classifiers $|\mathcal{C}^*|$, which have values range from 5 to 36, where 5 is the lowest number of classifiers selected and 36 is the total amount of possible classifiers that can be selected (six image descriptors and six learning methods result in 36 different simple classifiers).

In these experiments, we compare three selection strategies: *Consensus*, *Kendall*, and Rank Aggregation. *Consensus* refers to the strategy described in Section 3.2, which uses all the five diversity measures in the selection process. *Kendall*, in turn, refers to the strategy described in Section 3.3, which uses the two less correlated diversity measures (in the case, IA and $QSTAT$) in the selection process. These diversity measures were defined according to an *a priori* correlation analysis. Finally, Rank Aggregation refers to the use of the rank aggregation strategy described in Section 3.4. In Table 2, we highlight in blue the number minimum of classifier that each approach needs to achieve similar result than the FSVM-NORM-$|\mathcal{C}^*|$ using all classifiers ($|\mathcal{C}^*| = 36$). *Consensus* approach need to use $|\mathcal{C}^*| = 15$ classifiers. *Kendall* approach achieves similar result using $|\mathcal{C}^*| = 10$ classifiers. Finally, the rank aggregation approach with configuration *Kappa+DFM+IA+QSTAT* is able to yield very effective results with only $|\mathcal{C}^*| = 5$ classifiers.

We also computed the confidence intervals to verify if the results obtained by the proposed fusion approach differ from those observed for the baselines. We could observe that our approach achieves similar results to those observed for almost all baselines compared, but with fewer classifiers. Please, refer to the associated thesis[1] for more details regarding performed experiments.

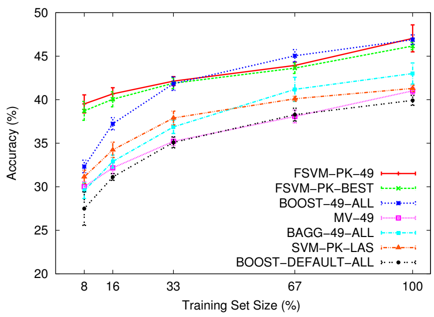

Fig. 5: Accuracy scores of all classifiers using training sets with different sizes [10].

Urban Dataset				
Approaches	Number of Classifiers $\|\mathcal{C}^*\|$			
	5	**10**	**15**	**36**
***Consensus* [10]**	0.564	0.570	**0.594**	0.612
***Kendall* [12]**	0.566	**0.592**	0.604	0.612
Rank Agg. [8]	**0.593**	0.592	0.602	0.612

Table 2: Kappa indices computed for each selection approach using different number of classifiers ($|\mathcal{C}^*|$) in the Urban dataset.

5 Conclusion

This work presented a framework for selection and fusion of simple classifiers using diversity measures and meta-learning on top of classifier outcomes. The

[1] `www.ic.unicamp.br/~ffaria/ffaria_final_thesis.pdf`

main novelty of this work relies on the use of diversity measures to determine which learning and image descriptor methods are more suitable to be combined in a given classification problem. Thus, three different strategies for classifier selection have been proposed (*Consensus*, *Kendall* correlation, and Rank Aggregation). This work resulted in papers in three important international journals [7,8,10] and three conference papers [9,11,12]. In addition, two articles have been submitted to international journals.For future work, we plan to investigate additional strategies and metrics for improving the classifier selection process, find the optimal diversity measures set for each application, test non-pairwise diversity measures, and perform experiments in other application domains.

References

1. Bishop, C.M.: Pattern Recognition and Machine Learning. Springer (2006)
2. Breiman, L.: Bagging predictors. Mach. Learn. 24(2), 123–140 (1996)
3. Croft, W.B.: Combining approaches to information retrieval. In: Croft, W.B. (ed.) Advances in Information Retrieval, vol. 7, pp. 1–36. Springer US (2002)
4. Didaci, L., Fumera, G., Roli, F.: Diversity in classifier ensembles: Fertile concept or dead end? In: MCS. pp. 37–48 (2013)
5. Du, P., Xia, J., Zhang, W., Tan, K., Liu, Y., Liu, S.: Multiple classifier system for remote sensing image classification: a review. Sensors 12(4), 4764 (2012)
6. Duin, R.P.W., Pekalska, E.: Open issues in pattern recognition. In: CORE. vol. 30, pp. 27–42 (2005)
7. Faria, F.A., P. Perre, P., Zucchi, R.A., Jorge, L.R., Lewinsohn, T.M., Rocha, A., Torres, R.d.S.: Automatic identification of fruit flies (diptera: Tephritidae). JVCIR. (2014)
8. Faria, F.A., Pedronette, D.C.G., dos Santos, J.A., Rocha, A., Torres, R.d.S.: Rank aggregation for pattern classifier selection in remote sensing images. JSTARS 7(4), 1103–1115 (April 2014)
9. Faria, F.A., dos Santos, J.A., Rocha, A., da S. Torres, R.: Automatic classifier fusion for produce recognition. In: SIBGRAPI. pp. 252–259 (2012)
10. Faria, F.A., dos Santos, J.A., Rocha, A., da S. Torres, R.: A framework for selection and fusion of pattern classifiers in multimedia recognition. PRL 39(0), 52 – 64 (2014)
11. Faria, F.A., dos Santos, J.A., da S. Torres, R., Rocha, A., Falcão, A.X.: Automatic fusion of region-based classifiers for coffee crop recognition. In: IGARSS. pp. 2221–2224 (2012)
12. Faria, F.A., dos Santos, J.A., Sarkar, S., Rocha, A., da S. Torres, R.: Classifier selection based on the correlation of diversity measures: When fewer is more. In: SIBGRAPI. pp. 16–23. Arequipa, Peru (august 2013)
13. Freund, Y., Schapire, R.E.: Experiments with a new boosting algorithm. ICML pp. 148–156 (1996)
14. Jurek, A., Bi, Y., Wu, S., Nugent, C.: Classification by cluster analysis: a new meta-learning based approach. In: MCS. pp. 259–268 (2011)
15. Kendall, M.G.: A New Measure of Rank Correlation. Biometrika 30(1/2), 81–93 (Jun 1938)
16. Kim, D., Baek, Y., Kim, W.: Reducing overfitting of adaboost by clustering-based pruning of hard examples. In: ICUIMC. p. 90 (2013)
17. Kuncheva, L.I.: Combining Pattern Classifiers: Methods and Algorithms. Wiley-Interscience (2004)
18. Kuncheva, L.I., Whitaker, C.J.: Measures of diversity in classifier ensembles and their relationship with the ensemble accuracy. Mach. Learn. 51, 181–207 (2003)
19. Morais, E., Goldenstein, S., Ferreira, A., Rocha, A.: Automatic tracking of indoor soccer players using videos from multiple cameras. In: SIBGRAPI. pp. 174–181 (aug 2012)
20. Rocha, A., Papa, J.P., Meira, L.A.A.: How far do we get using machine learning black-boxes? IJPRAI 26(2), 1261001–1–1261001–23 (2012)
21. Viola, P., Jones, M.: Rapid object detection using a boosted cascade of simple features. In: CVPR. pp. I–511–I–518 (2001)
22. Weber, R., Schek, H., Blott, S.: A quantitative analysis and performance study for similarity-search methods in high-dimensional spaces. In: VLDB. pp. 194–205 (1998)