A Meta-learning Approach for Recommendation of Image Segmentation Algorithms

Gabriel F. C. Campos, Sylvio Barbon Jr
Department of Computer Science
Londrina State University (UEL)
Londrina - PR, Brazil
Email: camposg@uel.com, barbon@uel.br

Rafael G. Mantovani
Sciences Institute of Mathematics and Computers (ICMC)
University of São Paulo (USP)
São Carlos - SP, Brazil
Email: rgmantov@icmc.usp.br

***Abstract*—There are many algorithms for image segmentation, but there is no optimal algorithm for all kind of image applications. To recommend an adequate algorithm for segmentation is a challenging task that requires knowledge about the problem and algorithms. Meta-learning has recently emerged from machine learning research field to solve the algorithm selection problem. This paper applies meta-learning to recommend segmentation algorithms based on meta-knowledge. We performed experiments in four different meta-databases representing various real world problems, recommending when three different segmentation techniques are adequate or not. A set of 44 features based on color, frequency domain, histogram, texture, contrast and image quality were extracted from images, obtaining enough discriminative power for the recommending task in different segmentation scenarios. Results show that Random Forest meta-models were able to recommend segmentation algorithms with high predictive performance.**

***Keywords*-Segmentation algorithm recommendation; meta-learning; image processing.**

I. Introduction

Image segmentation is one of the most difficult tasks in image processing [1] and also one of the most studied problems in the image analysis and computer vision fields [2]. There are many algorithms for image segmentation, but most of them are not equally good for a particular type of image. Furthermore, following something common on the Machine Learning (ML) field by the '*No free lunch theorem*' [3], no algorithm is the best choice for all kind of images. The best solution is to select different algorithms to segment different images types, but recommending them is a difficult duty [4].

Several different approaches for image segmentation have been emerging in the past years, and by evaluation of the segmentation itself, it is possible to choose suitable algorithms to a specific type or related group of images [2]. The features from an image provide useful information for automatic classification [5]. Thus, a suitable segmentation algorithm to a specific problem domain may be recommended by a representative features subset and learning-based methods, [4], [6].

Recently, an approach used to solve the algorithm selection problem is the Meta-learning (MtL) [7]. The main concept is to employ the knowledge acquired from previous similar problems to recommend one or more techniques successfully applied, now for the new dataset. The knowledge is represented by a meta-model, in image segmentation scenario responsible for mapping the dataset characteristics (meta-features) from each image problem to the performance of segmentation techniques. MtL has been employed in different contexts, for example in ML tasks to select [8], rank [9], or predict [10] the algorithms performance for employ on a new dataset.

MtL has also been employed in image processing and computer vision fields, for example, to retrieve the best set of parameter values to perform the watershed segmentation [11], to object detection and localization [12] and to search similar images [13].

In this paper, we have dealt with the recommending problem of segmentation algorithm via MtL. In other words, for a given image, several meta-features are extracted to provide information to recommend the appropriate segmentation algorithm.

Different from Yong et al. in [6], our proposed approach will consider color information for both image segmentation and description, as well real world images to build our meta-databases. In [6], was applied a learning-based approach for selection of a segmentation algorithm over synthetic/simulated gray-scale images composed of simple geometric forms.

This paper is organized as follows. Section II reviews related work. The experimental methodology is described in Section III. Section IV presents results and a discussion of our findings. Finally, the main conclusions and future work suggestions are presented in Section V.

II. Related work

Meta-learning (MtL), introduced by [7], brings the concept of *learning about learning*. It has a major intersection between machine learning and data mining techniques, frequently used to solve algorithm recommendation problems by suggesting algorithms and its hyper-parameters. The basic idea of MtL methods is to employ the knowledge and experience previously acquired from the application of various ML techniques on different problems to recommend one of these algorithms for a new dataset/problem. The *meta-knowledge* is represented by a meta-model that captures the relation between a set of characteristics (also called *meta-features*) from these datasets and the performance of ML algorithms.

Meta-learning have been used to recommend [14], [15], or optimize ML algorithms hyper-parameters [10], [16], [17]. In [18], [19] the authors used MtL to recommend the initial settings for search-based methods to optimize ML hyper-parameters. Some studies were also concerned in predict ML algorithms runtime using MtL [20], [21]. Other studies have been using MtL to select the number of clusters [22] or clustering algorithm [23], or even for periodic algorithm selection in data stream environments [9].

Regarding image processing, in the past few years, MtL have been applied to solve problems in image processing and computer vision domains. In [11], the authors applied MtL to relate image features with the results of the watershed algorithm and they were able to identify the best set of parameter to perform the watershed segmentation.

A solution for object detection and localization system was presented in [12]. In this work was proposed the use of AdaBoost on MtL task. Experimental results on three different datasets achieved comparable results with the state of the art methods but proved its generality.

The similar image search problem was discussed in [13]. The authors created a variant of the Particle Swarm Optimization (PSO) algorithm and a MtL approach to solving the search problem. The proposed approach was called Transfer Learning and achieved advantages when compared to the traditional PSO algorithm.

The segmentation algorithm can be selected by evaluating the segmentation performance itself. In other words, by the use of features extracted from an image and learning methods based on segmentation results, it is possible to recommend a specific algorithm. In this way, Fernandez et al. [2] evaluated segmentation methods from a higher level perspective. Results were promising, obtaining a good segmentation performance and also detecting image regions where segmentation was less accurate.

A learning-based approach for image segmentation selection was proposed in [6]. Results achieved high accuracy, but only gray-scale synthetic images were examined in the experiments. These synthetic images were composed of simple geometric forms.

In [4] a novel framework for intracellular image segmentation based on effective algorithm selection was proposed. The algorithm selection was conducted by measuring similarities between the user-supervised region and the automatically segmented regions. Experimental results showed that the framework could select an optimal algorithm to segment a region that has similar characteristics to the user-supervised region.

III. Experimental Methodology

In this paper, we propose the use of Meta-learning to recommend an image segmentation algorithms. In order to evaluate the capability of our proposed approach, were conducted experiments with different image databases. This section presents a detailed description of the experimental setup, including all the meta-databases, algorithms and evaluation methods adopted.

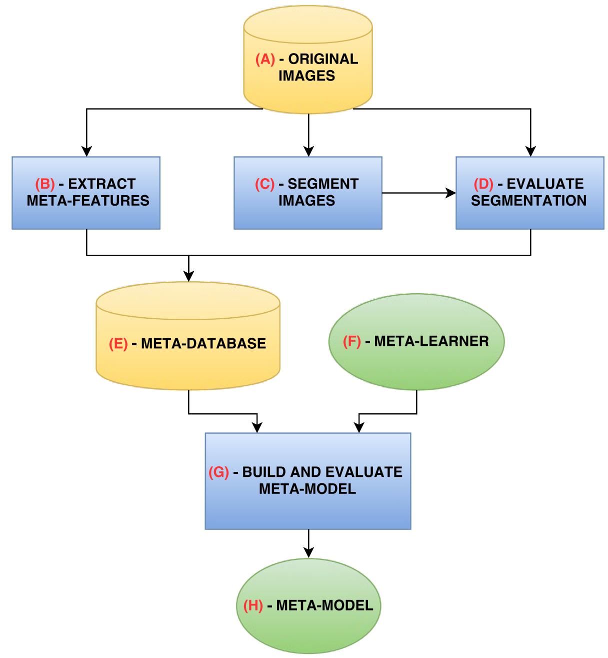


Fig. 1. Proposed meta-recommending system for image segmentation algorithms .

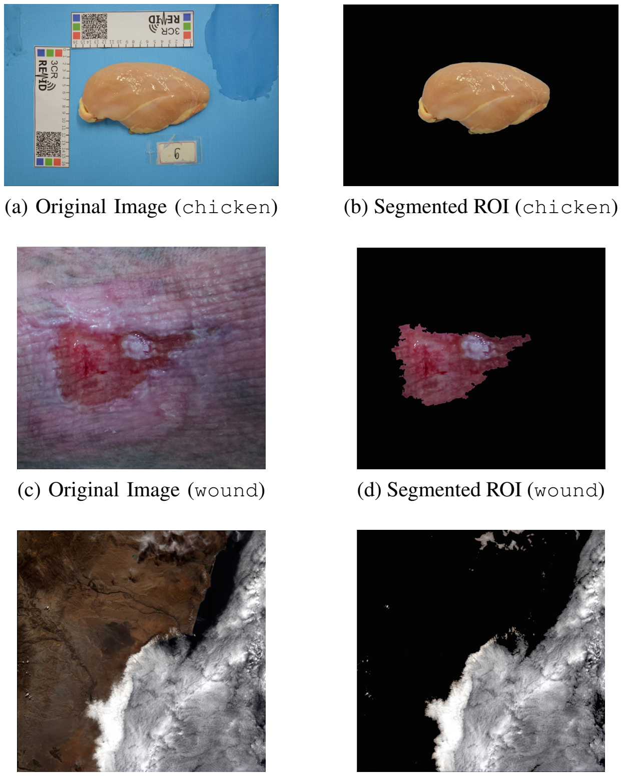
Figure 1 illustrates the proposed MtL framework. The flow starts with the original images **(A)**, from which the meta-knowledge is acquired with the extraction of meta-features **(B)**, the images segmentation **(C)** and the meta-target definition by evaluating the segmentation **(D)**. Once the meta-database **(E)** was obtained and the ML algorithm was selected as meta-learner **(F)**, the induction of the meta-model can be performed **(G)**. The meta-model **(H)** is the induced model ready to predict if a segmentation technique is suitable to a new image.

A. *Meta-databases*

Four meta-databases with different levels of complexity were used in the experiments. The first one (`chicken`), presented in [24], comprehends chicken breast samples, a simple segmentation task, which the Region of Interest (ROI) is the sample muscle itself extracted from the background. The second, `wound`, is composed of medical images which the ROI regards a wound region [25] and was considered a medium complexity task. The third, `cloud` (the more complex), consists of satellite images in which the ROI is cloud regions[1] [26]. We also performed experiments with a meta-database named as `All`, formed by the previous meta-databases. The image files were Portable Network Graphics (PNG). Examples of images and segmented ROIs of each meta-database are depicted in Figure 2.

It is important to note that the all the meta-databases represent different real world problems. The main difference among meta-databases focuses on ROI: size, colors, and background contrast. Another motivation to select these image scenarios

[1]Images from Landsat 8 are freely available on Amazon S3. We selected them to get only images with cloud regions.



(a) Original Image (`chicken`) (b) Segmented ROI (`chicken`)

(c) Original Image (`wound`) (d) Segmented ROI (`wound`)

(e) Original Image (`cloud`) (f) Segmented ROI (`cloud`)

Fig. 2. Examples of desirable segmentation of each meta-database: `chicken`, `wound` and `cloud`.

was the recent researches about: image segmentation in medical applications (`wound`) [25], [27]; food quality (`chicken`) [24], [28] and satellite/aerial imaging (`cloud`) [29], [30].

This way, our meta-databases were chosen to represent diverse segmentation problems. Note that different segmentation algorithms were used in those applications, endorsing our proposal of recommending an image segmentation algorithm to a given image.

B. Meta-features

One of the main steps in MtL is the meta-knowledge representation, made regarding meta-features and the meta-target attribute. In a meta-database, each dataset/problem is represented by a vector of characteristics called meta-features. According to [7], the meta-features must follow tree requirements: they need to have good discriminative power; the number of meta-features should not be too large to avoid overfitting, and their extraction should not be computational complex.

In this paper, for a given image, a set of 44 meta-features were explored. They include histogram-based [31], contrast and quality [32], gray-level co-occurrence matrix [33], Fast Fourier Transform (FFT) [34] and meta-features based on statistical information of colours [35], from 'Red, Green and Blue' (RGB) channels, and in the 'Hue, Saturation and Value' (HSV) channels. A complete list of all meta-features used in experiments is presented in Table I. The same set of meta-features was used by the meat-learners and to represent each image. No meta-features were used to represent the whole dataset in this experiment.

C. Meta-target

Three widely used segmentation methods, following different paradigms, were explored in this experiment: Otsu's thresholding [36], K-means [37], and SVM [38]. K-means and SVM were implemented for RGB values while Otsu's approach uses a gray-scale image as input. All images from the three meta-databases were segmented by these three algorithms.

The segmentation quality on each image was evaluated as *Adequate (AD)* or *Not Adequate (NAD)* with respect to each segmentation method. It was done by eight trained observers followed by a simple voting step. Samples with tied evaluations were considered undetermined and replaced by the statistical mode of the column.

Thus, each meta-database has 3 different binary meta-target attributes, one for each segmentation method. The classes' levels were: 'AD' indicating that the segmentation method is adequate for the meta-example (positive class); and 'NAD' when it is not adequate (negative class).

The main characteristics about the meta-databases are showed in Table II. The 'nExamp'-column indicates the number examples, while 'Class distr.' shows the class distribution for each one of the labels/segmentation methods.

D. Meta-learners

Due to the considered small number of meta-examples, the Leave-One-Out Cross-Validation (LOO-CV) resampling methodology was adopted to evaluate the predictive performance of the meta-learners. During the meta-learning level, we predicted the probability of a meta-example belongs to a specific meta-target value. Thus, seven ML classification algorithms, with different learning biases, were used as meta-learners: a linear classifier (LogReg), C4.5 Decision Tree (used through the J48 implementation), Naïve Bayes (NB), k-Nearest Neighbors (k-NN), Neural Networks Using Model Averaging (avNNet), Random Forest (RF) and Support Vector Machine (SVM). All of them were implemented in R environment using the `mlr` ('Machine Learning in R') package[2] and relying on the default parameters.

E. Performance measures

Each meta-learner was evaluated according to several performance measures. We modeled our problem as a set of binary problems, via a Binary Relevance (BR) strategy. BR is the most straightforward and arguably simplest approach to performing multi-label classification [39]. Thus, we extracted the simple Predictive Accuracy, the True Positive Rate (TPRate), the True Negative Rate (TNRate), the Area Under the ROC curve (AUC) and the F-Measure for each one of the binary classification problems. The Hamming Loss value is

[2] https://github.com/mlr-org/mlr

TABLE I
LIST OF ALL IMAGE FEATURES AS META-FEATURES IN PROPOSED MTL RECOMMENDING APPROACH.

No.	Type	Name	Description	No	Type	Name	Description
1	Color	meanB	Mean value of the B channel	23	IntenHist	nNzGIntHist	Amount of non-zero groups in intensity histogram
2	Color	stdB	Standard deviation of the B channel	24	IntenHist	kurIntHist	Kurtosis of intensity histogram
3	Color	varB	Variance value of the B channel	25	IntenHist	pLgNzGIntHist	Peak of the largest non-zero group in intensity histogram
4	Color	meanG	Mean value of the G channel	26	IntenHist	pSmNzGIntHist	Peak of the smaller non-zero group in intensity histogram
5	Color	stdG	Standard deviation of the G channel	27	IntenHist	skeIntHist	Skewness of intensity histogram
6	Color	varG	Variance value of the G channel	28	IntenHist	lenLgNzGIntHist	Length of the largest non-zero group in intensity histogram
7	Color	meanH	Mean value of the H channel	29	IntenHist	lenSmNzGIntHist	Length of the smaller non-zero group in intensity histogram
8	Color	stdH	Standard deviation of the H channel	30	IntenHist	meanIntHist	Mean of intensity histogram amplitude
9	Color	varH	Variance value of the H channel	31	IntenHist	mdIntHist	Median of intensity histogram amplitude
10	Color	meanR	Mean value of the R channel	32	IntenHist	sdIntHist	Standard deviation of intensity histogram amplitude
11	Color	stdR	Standard deviation of the R channel	33	IntenHist	varIntHist	Variance of intensity histogram amplitude
12	Color	varR	Variance value of the R channel	34	IntenHist	nNzIntHist	Amount of non-zero values on intensity histogram
13	Color	meanS	Mean value of the S channel	35	IntenHist	peakIntHist	Peak of intensity histogram
14	Color	stdS	Standard deviation of the S channel	36	CoMatrix	entGMat	Entropy of gray-level co-occurrence matrix
15	Color	varS	Variance value of the S channel	37	CoMatrix	homGMat	Homogeneity of gray-level co-occurrence matrix
16	Color	meanV	Mean value of the V channel	38	CoMatrix	ineGMat	Inertia of gray-level co-occurrence matrix
17	Color	stdV	Standard deviation of the V channel	39	CoMatrix	corGMat	Correlation of gray-level co-occurrence matrix
18	Color	varV	Variance value of the V channel	40	CoMatrix	engGMat	Energy of gray-level co-occurrence matrix
19	FFT	engFFT	FFT Energy	41	Contrast	entInt	Entropy of original intensity image
20	FFT	entFFT	FFT Entropy	42	Contrast	gConstFact	Global Contrast Factor
21	FFT	homFFT	FFT Homogeneity	43	Image Quality	stNatMeas	Statistical Naturalness Measure (SNM)
22	FFT	ineFFT	FFT Inertia	44	Image Quality	measEnhan	Measure of Enhancement (EME)

TABLE II
SPECIFICATION OF THE META-DATABASES USED IN EXPERIMENTS.

Dataset	nExamp	Class distr (AD-NAD).		
		Otsu	K-means	SVM
Chicken	142	139-3	15-127	13-129
Wound	133	56-77	4-129	33-100
Cloud	91	51-40	63-28	70-21
All	366	237-115	73-276	102-230

usually used to measure accuracy in a multi-label classification task. The inverse of the Hamming Loss value is used as Predictive Accuracy in this paper.

F. Baselines

A trivial meta-model (`SingleBest`), which always predict the majority class, and a random meta-model (`Random`) were used as baselines of our induced meta-models.

IV. RESULTS AND DISCUSSION

In this section, the main experiment results are presented and discussed. We ran seven meta-learners, with different learning biases in the four meta-databases. The performance measures for each scenario are depicted trough a radar chart in Figure 3. At the figure, each colored line represents a meta-model, and each polygon vertex accounts for a different performance measure. Greater the values/area better is the meta-model. The figure also shows the performance of the two baselines: `Random` (light gray line), and `SingleBest` (inner black line).

Looking to the `wound` results, one might see that the biggest area is generated by RF meta-model (red polygon). It has the best values for accuracy 0.907, AUC 0.942 and TNRate 0.961. The linear meta-model (LogReg) obtained the best results regarding F-Measure (0.569) and TPRate (0.629). No meta-model presented an F-Measure value greater than 0.6; this is directly related to the class distribution of the k-Means segmentation method: just 4 images have the label 'AD' while 129 have 'NAD'. Even that, in the overall scenario with all the measures, RF has a better performance than the others meta-models.

In the `chicken` meta-database, the SVM meta-model was the best one in terms of accuracy (0.960) and F-Measure (0.786), followed by the RF algorithm (accuracy = 0.955, F-Measure = 0.780). The linear meta-model (LogReg) presented the best results in terms of AUC (0.917) and TPRate (0.893) measures, and the biggest area over all meta-models (blue region). This is interesting since the best choice regarding all the measures for this meta-database is the simplest meta-model (LogReg).

Analyzing the `cloud` scenario we may state the meta-models performed similarly. RF and SVM were the best ones in terms of accuracy (0.743 both), AUC (0.675 both) with the NB, and F-Measure (0.817 both). The NB was the best regarding TNRate (0.619 - yellow line). The avNNet was the best in TPRate (0.994 - purple line) with the SingleBest baseline. So, the best meta-models for this meta-database are SVM and RF.

Finally, looking to the overall scenario (`all`), we can see that when inducing meta-models with all the available images RF reached the best values for all the evaluation measures. The biggest region is the red one and it contains all the other ones. The information provided by the complete meta-database improved the RF performance, fixing some previous wrong predictions (This is more discussed in the section IV-B).

RF meta-models performed well in 3/4 of the scenarios. Table III compares RF meta-models results with both baselines. One may note that RF meta-model outperforms both baselines regarding predictive accuracy, AUC and F-Measure. In all the cases, the areas generated by the baseline methods at Figure 3 (light gray and black ones) are contained by the

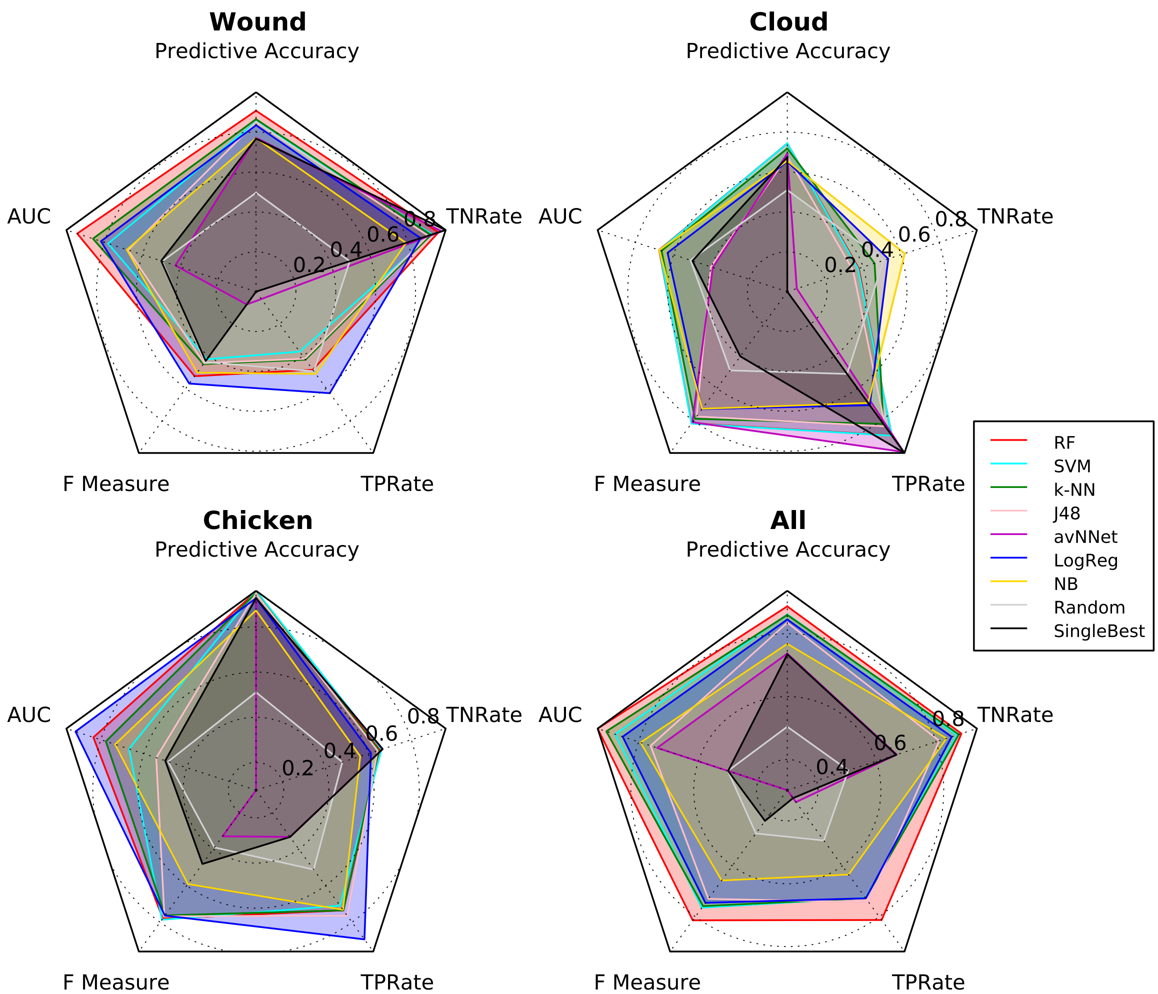


Fig. 3. Radar chart with the meta-models' performance according to five different measures: mean Predictive Accuracy, mean AUC, mean F-Measure, mean True Positive Rate (TPRate) and the mean True Negative Rate (TNRate).

best meta-models in all scenarios. Thus, only the RF results will be discussed in the next sections.

A. Random Forest Attributes Importance

We extracted the importance of each meta-feature at the RF meta-models using the Mean Decrease Gini value.

RF algorithm uses the Gini Index as a measure for the best split selection [40]. The Figure 4 depicts, for each meta-database, the main contribution made by each meta-feature during the recommendation.

In the `wound` meta-database, color features from the HSV color space had a high importance value with the Otsu and SVM segmentation algorithms, especially the Hue (H) and Saturation (S) channels. It makes sense, once there is a high difference in saturation between ROI and background as can be seen in Figure 2. The Measure of Enhancement (EME) also reached a high importance value in this meta-database

TABLE III
RANDOM FOREST META-MODEL'S PERFORMANCE COMPARED WITH BASELINES RANDOM AND SINGLEBEST.

Dataset	Algorithm	Accuracy	AUC	F-Measure	TPRate	TNRate
chicken	Random	0.51064	0.48669	0.39066	0.51154	0.47913
chicken	SingleBest	0.92723	0.50000	0.48093	0.33333	**0.66667**
chicken	RF	**0.95549**	**0.83452**	**0.77959**	**0.73504**	0.65367
cloud	Random	0.50952	0.51047	0.48811	0.50985	**0.50811**
cloud	SingleBest	0.67399	0.50000	0.40101	**1.00000**	0.00000
cloud	RF	**0.74359**	**0.67570**	**0.81606**	0.89035	0.37698
wound	Random	0.49599	0.49432	0.43637	**0.49347**	0.49402
wound	SingleBest	0.76692	0.50000	0.42941	0.00000	**1.00000**
wound	RF	**0.90727**	**0.94280**	**0.52381**	0.48413	0.96169
all	Random	0.50413	0.50079	0.47239	0.49992	0.50633
all	SingleBest	0.73406	0.50000	0.42285	0.33333	0.66667
all	RF	**0.88707**	**0.93658**	**0.81384**	**0.81264**	**0.88463**

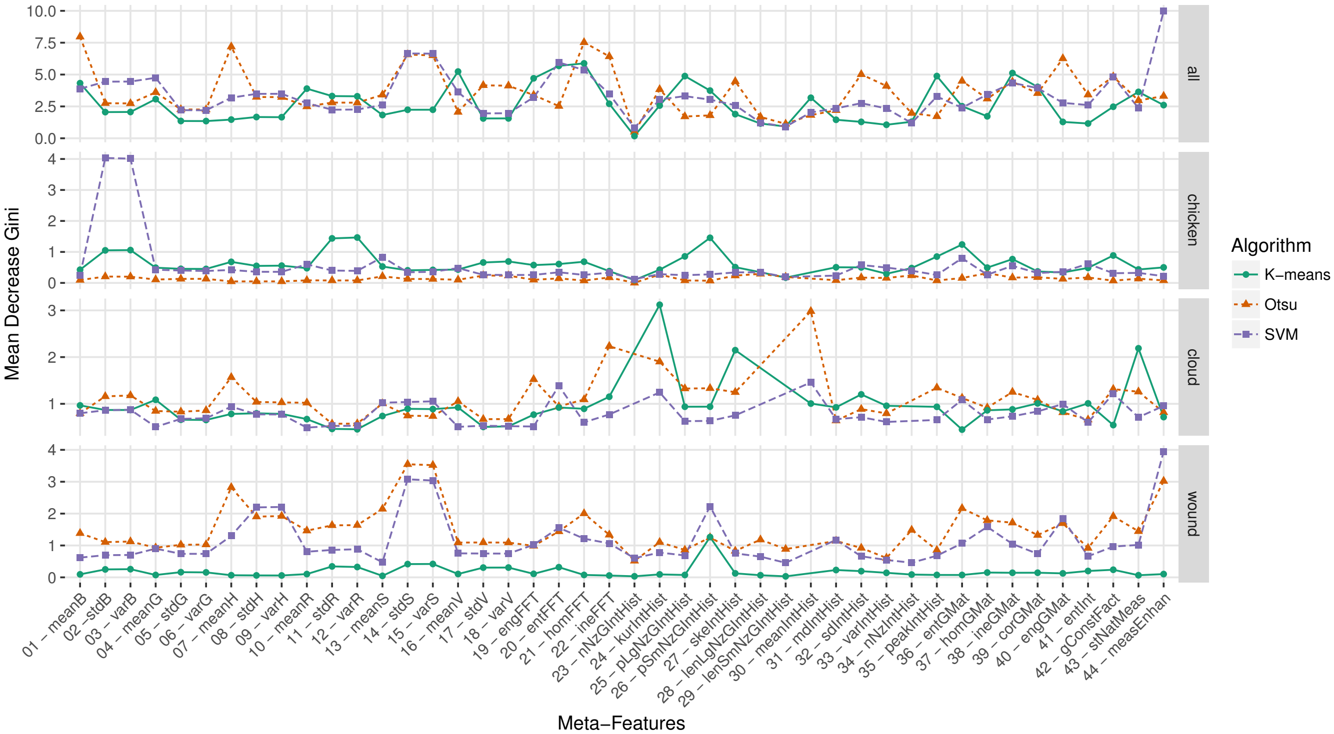


Fig. 4. Mean decrease Gini for all the meta-features extracted from RF meta-models.

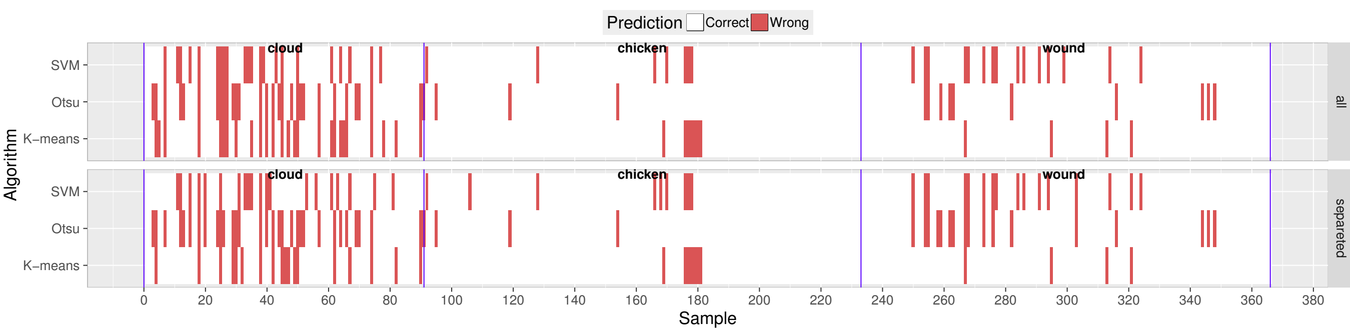


Fig. 5. Prediction for each sample regarding the full meta-database (top) and each meta-database separately (bottom)

with Otsu and SVM. Regarding the K-means algorithm, only the peak of the smaller non-zero group in intensity histogram outperformed.

Looking to the `chicken` results, blue channel features from the RGB color space performed better than the others, mainly with the SVM segmentation algorithm. It probably happened due to the blue background used to generate the images. There were also some small peaks in features like the red channel of the RGB color space, the peak of the smaller non-zero group in intensity histogram and the entropy of gray-level co-occurrence matrix, regarding the K-means segmentation algorithm. Once the Gini Index is related to the best split selection, the lack of an outperforming feature with the Otsu segmentation algorithm is related to the class distribution (139 NAD and 3 AD).

The `cloud` meta-database had a chaotic distribution of meta-feature importance, probably due its complexity. However, overall, to the K-means algorithm, the most relevant features were about the intensity histogram, specifically the kurtosis and the skewness values. The Statistical Naturalness Measure (SNM) was necessary to the K-means as well. To the Otsu and SVM algorithms, intensity histogram features were also relevant, followed by the Inertia, Entropy and Energy values of the FFT.

Finally, looking to the overall scenario (`all`), the most relevant feature was the Measure of Enhancement (EME) with the SVM algorithm, followed by color information in the Blue and Hue channels and the Homogeneity value of FFT (to the K-means algorithm). Generally speaking, nearly all features had their importance to the full meta-database and, although some peaks are presented, no feature outperformed though all meta-databases. It indicates that meta-feature selection may not be appropriate to this problem.

B. Random Forest Predictions

Figure 5 presents the predictions obtained by each RF meta-model. The x-axis shows each one of the samples (images) while y-axis presents all the segmentation algorithms (Otsu, SVM, and k-Means). At the figure, a red rectangle indicates a wrong prediction, while white ones indicate samples predicted correctly.

This way, problematic samples are recognized by a 'red line', which means that sample was wrongly predicted to all segmentation algorithms in that meta-database. No problematic samples were found in the `chicken` meta-database. In the `wound` meta-database, only one problematic sample (267) was found when the meta-model was build to each meta-database separately. This sample is showed in Figure 6. The segmentation of sample 267 was evaluated as AD to all the segmentation algorithms, but our meta-model failed by predicting it as NAD.

Most of the problematic samples were found in the `cloud` meta-database, the most complex meta-database to segment and evaluate. Three problematic samples were identified with $ids = 18, 25, 38$. When analyzing the whole meta-database (`all`) this number have grown to nine ($ids = 7, 18, 25, 26, 27, 38, 45, 50, 74$). The images $ids = 18, 25, 38$ appeared in both cases. They were evaluated as 'NAD' to all the segmentation algorithms, but the meta-models failed by predicting them as 'AD'. Those samples can also be seen in Figure 6.

V. Conclusion

In this study, a framework to recommend image segmentation algorithms using meta-learning was presented. A set of 44 image-based meta-features were used to characterize real-world images and provide information to predict the segmentation algorithm adequate to a specific problem.

Experiments performed using four meta-datasets and seven ML algorithms, even without meta-feature selection or hyper-parameter tuning, achieved high accurate meta-models. RF meta-models were able to recommend segmentation algorithms at the overall scenario with high predictive accuracy (0.887), AUC (0.936) and F-Measure (0.813) values. It indicates a good discriminative power of the chosen meta-features and represents a contribution to computer vision and image processing fields.

As future work, besides implementing more meta-features and segmentation algorithms, we intend to try unsupervised segmentation methods in the evaluation step of our framework. It would make possible to increase our meta-databases significantly.

Acknowledgment

The authors would like to thank FAPESP (Brazilian Agency) for their financial support, specially the grant #2012/23114-9, São Paulo Research Foundation (FAPESP).

References

[1] T. Sa and M. unka, "Color image segmentation based on multiobjective artificial bee colony optimization," *Applied Soft Computing*, vol. 34, no. 0, pp. 389 – 401, 2015.

[2] M. A. Fernandez, R. M. Lopes, and N. S. T. Hirata, "Image segmentation assessment from the perspective of a higher level task," in *2015 28th SIBGRAPI Conference on Graphics, Patterns and Images*, Aug 2015, pp. 111–118.

[3] D. H. Wolpert, "The lack of a priori distinctions between learning algorithms," *Neural Computation*, vol. 8, no. 7, pp. 1341–1390, Oct. 1996.

[4] S. Takemoto and H. Yokota, "Algorithm selection for intracellular image segmentation based on region similarity," in *2009 Ninth International Conference on Intelligent Systems Design and Applications*, Nov 2009, pp. 1413–1418.

[5] G. F. C. Campos, R. A. Igawa, J. L. Seixas, A. M. G. Almeida, R. C. Guido, and S. Barbon, "Supervised approach for indication of contrast enhancement in application of image segmentation," in *MMEDIA 2016, The Eighth International Conferences on Advances in Multimedia*, Feb 2016, pp. 12–18.

[6] X. Yong, D. Feng, Z. Rongchun, and M. Petrou, "Learning-based algorithm selection for image segmentation," *Pattern Recognition Letters*, vol. 26, no. 8, pp. 1059 – 1068, 2005.

[7] P. Brazdil, C. Giraud-Carrier, C. Soares, and R. Vilalta, *Metalearning: Applications to Data Mining*, 2nd ed. Springer Verlag, 2009.

[8] S. Ali and K. A. Smith-Miles, "A meta-learning approach to automatic kernel selection for support vector machines," *Neurocomputing*, vol. 70, no. 13, pp. 173–186, 2006.

[9] A. L. D. Rossi, A. C. P. de Leon Ferreira de Carvalho, C. Soares, and B. F. de Souza, "Metastream: A meta-learning based method for periodic algorithm selection in time-changing data," *Neurocomputing*, vol. 127, pp. 52 – 64, 2014.

[10] M. Reif, F. Shafait, and A. Dengel, "Meta-learning for evolutionary parameter optimization of classifiers," *Machine Learning*, vol. 87, pp. 357–380, 2012.

[11] A. Attig and P. Perner, "A study on the case image description for learning the model of the watershed segmentation," *Transactions on Case-Based Reasoning*, vol. 2, no. 1, pp. 41–53, 2009.

[12] H. A. Marakeby, M. Zaki, and I. S. Samir, "A generalized object detection system using automatic feature selection," in *2010 10th International Conference on Intelligent Systems Design and Applications*, Nov 2010, pp. 839–844.

[13] T. Yamaguchi, K. Mori, K. J. Mackin, and Y. Nagai, "Application of particle swarm optimization to similar image search on satellite sensor data," in *Soft Computing and Intelligent Systems (SCIS) and 13th International Symposium on Advanced Intelligent Systems (ISIS), 2012 Joint 6th International Conference on*, Nov 2012, pp. 1573–1577.

[14] C. Soares, P. B. Brazdil, and P. Kuba, "A meta-learning method to select the kernel width in support vector regression," *Machine Learning*, vol. 54, no. 3, pp. 195–209, 2004.

[15] C. Soares and P. B. Brazdil, "Selecting parameters of svm using meta-learning and kernel matrix-based meta-features," in *Proceedings of the 2006 ACM symposium on Applied computing*, ser. SAC'06. ACM Press, 2006, pp. 564–568.

[16] T. A. F. Gomes, R. B. C. Prudêncio, C. Soares, A. L. D. Rossi, and nd André C. P. L. F. de Carvalho, "Combining meta-learning and search techniques to select parameters for support vector machines," *Neurocomputing*, vol. 75, no. 1, pp. 3–13, Jan. 2012.

[17] P. B. C. Miranda, R. B. C. Prudêncio, A. C. P. L. F. de Carvalho, and C. Soares, "An experimental study of the combination of meta-learning with particle swarm algorithms for svm parameter selection," *Lecture Notes in Computer Science*, vol. 7335 LNCS, no. PART 3, pp. 562–575, 2012.

[18] P. B. Miranda, R. B. Prudêncio, A. P. de Carvalho, and C. Soares, "A hybrid meta-learning architecture for multi-objective optimization of {SVM} parameters," *Neurocomputing*, vol. 143, pp. 27–43, 2014.

[19] M. Feurer, J. T. Springenberg, and F. Hutter, "Initializing bayesian hyperparameter optimization via meta-learning," in *Twenty-Ninth AAAI Conference on Artificial Intelligence*, 2015, pp. 1128–1135.

[20] M. Reif, F. Shafait, and A. Dengel, "Prediction of classifier training time including parameter optimization," in *Proceedings of the 34th Annual German conference on Advances in artificial intelligence*, ser. KI'11. Springer-Verlag, 2011, pp. 260–271.

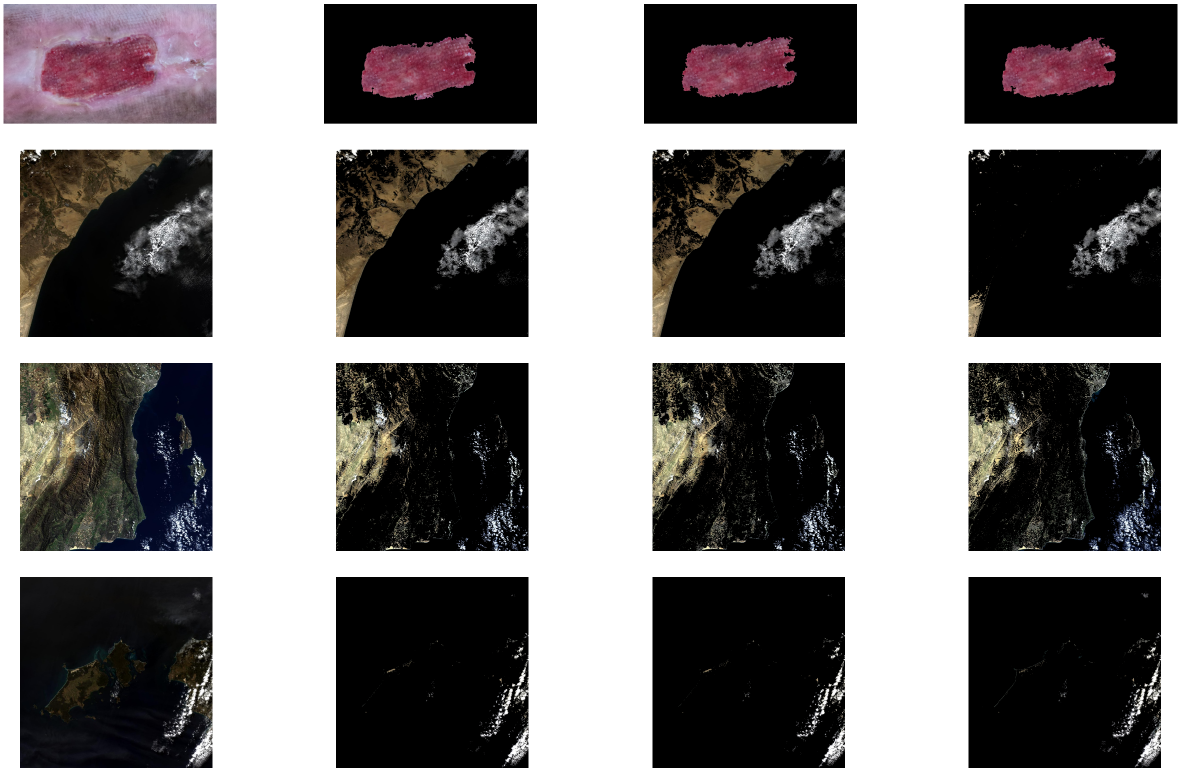


Fig. 6. Problematic Samples

[21] R. Priya, B. F. De Souza, A. L. D. Rossi, and A. C. P. L. F. de Carvalho, "Using genetic algorithms to improve prediction of execution times of ML tasks," in *Lecture Notes in Computer Science*, ser. Lecture Notes in Computer Science, vol. 7208 LNAI, no. PART 1, 2012, pp. 196–207.

[22] J.-S. Lee and S. Olafsson, "A meta-learning approach for determining the number of clusters with consideration of nearest neighbors," *Information Sciences*, vol. 232, pp. 208–224, 2013.

[23] D. G. Ferrari and L. N. de Castro, "Clustering algorithm selection by meta-learning systems: A new distance-based problem characterization and ranking combination methods," *Information Sciences*, vol. 301, pp. 181–194, 2015.

[24] D. F. Barbin, S. M. Mastelini, S. B. Jr., G. F. Campos, A. P. A. Barbon, and M. Shimokomaki, "Digital image analyses as an alternative tool for chicken quality assessment," *Biosystems Engineering*, vol. 144, pp. 85 – 93, 2016.

[25] J. Seixas *et al.*, "Color energy as a seed descriptor for image segmentation with region growing algorithms on skin wound images," in *e-Health Networking, Applications and Services (Healthcom), 2014 IEEE 16th International Conference on*, Oct 2014, pp. 387–392.

[26] "Landsat on aws," https://aws.amazon.com/public-data-sets/landsat/, accessed: 2016-03-14.

[27] M. Avendi, A. Kheradvar, and H. Jafarkhani, "A combined deep-learning and deformable-model approach to fully automatic segmentation of the left ventricle in cardiac {MRI}," *Medical Image Analysis*, vol. 30, pp. 108 – 119, 2016.

[28] M. K. Dutta, A. Issac, N. Minhas, and B. Sarkar, "Image processing based method to assess fish quality and freshness," *Journal of Food Engineering*, vol. 177, pp. 50 – 58, 2016.

[29] A. Bhandari, A. Kumar, and G. Singh, "Tsallis entropy based multilevel thresholding for colored satellite image segmentation using evolutionary algorithms," *Expert Systems with Applications*, vol. 42, no. 22, pp. 8707 – 8730, 2015.

[30] S. Suresh and S. Lal, "An efficient cuckoo search algorithm based multilevel thresholding for segmentation of satellite images using different objective functions," *Expert Systems with Applications*, vol. 58, pp. 184 – 209, 2016.

[31] G. Balaji, T. Subashini, and N. Chidambaram, "Automatic classification of cardiac views in echocardiogram using histogram and statistical features," *Procedia Computer Science*, vol. 46, no. 0, pp. 1569–1576, 2015, proceedings of the International Conference on Information and Communication Technologies, {ICICT} 2014, 3-5 December 2014 at Bolgatty Palace Camp; Island Resort, Kochi, India.

[32] H. Yeganeh and Z. Wang, "Objective quality assessment of tone-mapped images," *IEEE Transactions on Image Processing*, vol. 22, no. 2, pp. 657–667, Feb 2013.

[33] S. Chowdhury, B. Verma, and D. Stockwell, "A novel texture feature based multiple classifier technique for roadside vegetation classification," *Expert Systems with Applications*, vol. 42, no. 12, pp. 5047–5055, 2015.

[34] H.-K. Shen, P.-H. Chen, and L.-M. Chang, "Automated steel bridge coating rust defect recognition method based on color and texture feature," *Automation in Construction*, vol. 31, no. 0, pp. 338–356, 2013.

[35] D. Li, N. Li, J. Wang, and T. Zhu, "Pornographic images recognition based on spatial pyramid partition and multi-instance ensemble learning," *Knowledge-Based Systems*, vol. 84, pp. 214 – 223, 2015.

[36] N. Otsu, "A Threshold Selection Method from Gray-level Histograms," *IEEE Transactions on Systems, Man and Cybernetics*, vol. 9, no. 1, pp. 62–66, Jan. 1979.

[37] H. Li, H. He, and Y. Wen, "Dynamic particle swarm optimization and k-means clustering algorithm for image segmentation," *Optik - International Journal for Light and Electron Optics*, vol. 126, no. 24, pp. 4817 – 4822, 2015.

[38] X.-Y. Wang, T. Wang, and J. Bu, "Color image segmentation using pixel wise support vector machine classification," *Pattern Recogn.*, vol. 44, no. 4, pp. 777–787, Apr. 2011.

[39] E. Montaes, R. Senge, J. Barranquero, J. R. Quevedo, J. J. del Coz, and E. Hllermeier, "Dependent binary relevance models for multi-label classification," *Pattern Recognition*, vol. 47, no. 3, pp. 1494 – 1508, 2014.

[40] V. Rodriguez-Galiano, B. Ghimire, J. Rogan, M. Chica-Olmo, and J. Rigol-Sanchez, "An assessment of the effectiveness of a random forest classifier for land-cover classification," {*ISPRS*} *Journal of Photogrammetry and Remote Sensing*, vol. 67, pp. 93 – 104, 2012.