Accepted version of paper published in the 15th International Workshop on Content-Based Multimedia Indexing (CBMI), 2017.
DOI: https://doi.org/10.1145/3095713.3095741

Selection and Combination of Unsupervised Learning Methods for Image Retrieval

Lucas Pascotti Valem and Daniel Carlos Guimarães Pedronette
Department of Statistics, Applied Math. and Computing, São Paulo State University (UNESP)
Rio Claro, Brazil
lucasvalem@rc.unesp.br,daniel@rc.unesp.br

ABSTRACT

The evolution of technologies to store and share images has made imperative the need for methods to index and retrieve multimedia information based on visual content. The CBIR (Content-Based Image Retrieval) systems are the main solution in this scenario. Originally, these systems were solely based on the use of low-level visual features, but evolved through the years in order to incorporate various supervised learning techniques. More recently, unsupervised learning methods have been showing promising results for improving the effectiveness of retrieval results. However, given the development of different methods, a challenging task consists in to exploit the advantages of diverse approaches. As different methods present distinct results even for the same dataset and set of features, a promising approach is to combine these methods. In this work, a framework is proposed aiming at selecting the best combination of methods in a given scenario, using different strategies based on effectiveness and correlation measures. Regarding the experimental evaluation, six distinct unsupervised learning methods and two different datasets were used. The results as a whole are promising and also reveal good perspectives for future works.

CCS CONCEPTS

• **Information systems → Multimedia and multimodal retrieval**;

KEYWORDS

content-based image retrieval; unsupervised learning; re-ranking; rank-aggregation; combination; effectiveness; correlation

1 INTRODUCTION

Due to the continuously and consistent technological advances in the late years, the development of approaches to index and retrieve information has become indispensable, specially for images and visual contents. Content-based image retrieval tasks are of extreme importance, with different applications available [1] (diagnosis of diseases, facial recognition, remote sensing, object identification, and others). The CBIR (Content-Based Image Retrieval) systems are the main solution for retrieval based on visual content. The main objective of these systems consists in, given a query image, retrieve the most similar images of the dataset in a descending order of similarity. For a given image, its features can be extracted by a visual descriptor. These descriptors extract feature vectors and, subsequently, compute the distance between images [2]. Different approaches have been exploited by diverse visual descriptors: global (color, texture, shape) [3, 4], local [5], and deep learning [6]. However, this strategy has an intrinsic problem related to its representation, that considers low-level features instead of the semantic content, deepening the well-known semantic gap problem [7].

As an approach to reduce the semantic gap impact, supervised learning methods have been included in the retrieval pipeline. However, in some scenarios, the use of supervised methods are limited due to the requirement of training procedures and labeled data. Recently, unsupervised learning methods have been proposed, substituting pairwise measures by global affinity ones [8–10].

The results of unsupervised learning methods have been very promising, mainly in effectiveness aspects. In order to further improve the accuracy, approaches to aggregate complementary results of distinct retrieval methods have been proposed [11]. Since different methods present distinct and often complementary results for a same dataset, an interesting strategy is to perform a combination of such methods. However, the selection of the best combination strongly depends on the dataset and descriptors being used. Some recent works present approaches to select and combine methods using supervised techniques [12] and genetic algorithms [13].

In this work, we present an investigation which aims to determine if the best combination of unsupervised learning methods can also be selected in an unsupervised way, without utilizing training data or supervised approaches. To perform the analysis, different effectiveness and correlation measures were used to compose a selection strategy, which aims at defining the best combination in a given retrieval scenario. The experiments were conducted using six distinct unsupervised learning methods and two different datasets. In most cases, the results have shown that the best combinations can be selected using our presented approach, indicating promising results.

This paper is organized as follows: Section 2 discusses related work. Section 3 presents the formulation of the image retrieval model considered, while Section 4 presents the proposed selection framework. Section 5 reports the experimental evaluation results. Finally, Section 6 draws the conclusions of this work.

2 RELATED WORK

Unsupervised learning methods have been proposed [8, 14] aiming at minimizing the semantic gap impact through contextual

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.
CBMI '17, June 19-21, 2017, Florence, Italy
© 2017 Association for Computing Machinery.
ACM ISBN 978-1-4503-5333-5/17/06…$15.00
https://doi.org/10.1145/3095713.3095741

analysis. The main idea of these post-processing methods is to replace pairwise measures by global affinity values capable of taking into account the dataset manifold [15]. In the literature, different unsupervised learning methods can be found based on different approaches: diffusion processes [16, 17], graphs [8, 15], clustering [18], frequency of patterns [19], rank-analysis [20], and others.

Over the recent years, unsupervised learning methods based on rank-analysis have been achieving high and promising effectiveness results. Additionally, these methods require low computational efforts, since they do not need all the positions of the ranked lists [10, 21–23] to perform an execution. This research considers six different unsupervised learning methods: ContextRR [9], Cartesian Product of Ranking References (CPRR) [21], Ranked List Graph Distance [23], ReckNNGraph [22], RL-Recommmendation [10], and RL-Sim* [24, 25]. The methods have results comparable with the state-of-the-art and are based on rank-analysis, but use distinct techniques (graphs, Cartesian product, recommendations, and others). All of them are open-source and public available [26].

Different retrieval methods are capable of providing distinct information about a given dataset. Therefore, several unsupervised learning methods, specially the ones based on rank-analysis, use rank-aggregation approaches for exploiting such complementarity. A rank-aggregation method consists in an algorithm that receives two or more different inputs provided by distinct retrieval methods, with the purpose of aggregate complementary information to achieve more effective results [27]. In most cases, the rank-aggregation is seen as a way to create consensual ranked lists [28]. The aggregation exploits information of different ranked lists. For example, if an image appears in the top positions of two different ranked lists, it is inferred that this image must be placed in the top positions. Different approaches have been presented with this objective, among which we can cite genetic programming, support vector machines (SVM), and association rules [11, 29]. Strategies to combine multiple heterogeneous methods for image retrieval in annotated collections have also been proposed [12].

Considering the promising gains obtained for the unsupervised learning methods, models to perform the combinations of these methods have been proposed [30]. In the same way that descriptors can have their results aggregated, the same can be done with the output of unsupervised learning methods. Recently, a framework that combines different methods using distinct supervised approaches and genetic algorithms has been proposed [13].

3 IMAGE RETRIEVAL MODEL

This section formally describes the retrieval model used along this paper. Let C=$\{img_1, img_2, \dots, img_n\}$ be an image collection (dataset), where n denotes the collection size.

Let $\mathcal{D}$ be an image descriptor, which can be defined as a tuple (ϵ, ρ), where $\hat{I} \rightarrow \mathbb{R}^m$ is a function that extracts the feature vector $v_{\hat{I}}$ of a given image in $\hat{I}$; and ρ: $\mathbb{R}^m \times \mathbb{R}^m \rightarrow \mathbb{R}$ is a distance function that provides the distance between two images according to their corresponding feature vectors. The distance between two images img_i and img_j is given by the value of $\rho(\epsilon(img_i), \epsilon(img_j))$. For readability purposes, the notation $\rho(i, j)$ is used to denote the distance between two images.

The distance $\rho(i, j)$ between all images $img_i, img_j \in C$ can be computed to obtain a square distance matrix A, such that $A_{ij} = \rho(i, j)$. The matrix A is used as the input for the majority of the unsupervised learning methods.

An alternative representation of the retrieval results is the ranked lists. Based on the distance function ρ, a ranked list can be computed for a query image img_q. A ranked list can contain information of the whole dataset; but, since the most important information is available at the top positions, a generally adopted strategy is to consider only a subset of the top-L positions, which contains the L most similar images to the given query image. This is an useful approach used to accelerate the retrieval process, since $L \ll N$ for larger datasets.

Therefore, the ranked list τ_q=$(img_1, img_2, \dots, img_L)$ can be defined as a permutation of an image collection $C_s \subset C$, which contains the most similar images to a given query image img_q, such that $|C_s| = L$. The permutation τ_q is a bijection from the set C_s onto the set $[L] = \{1, 2, \dots, L\}$. For a permutation τ_q, the $\tau_q(i)$ notation denotes the position (or rank) of image img_i in the ranked list τ_q. We can say that, if image img_i is ranked before image img_j in the ranked list of image img_q, that is, $\tau_q(i) < \tau_q(j)$, then $\rho(q, i) \le \rho(q, j)$. Taking every image $img_i \in C$ as a query image img_q, the set of ranked lists $\mathcal{R} = \{\tau_1, \tau_2, \dots, \tau_n\}$ is obtained.

In general, unsupervised learning methods consider as input a distance matrix A or a set of ranked lists $\mathcal{R}$ and perform the execution to provide a more effective set of ranked lists $\mathcal{R}_f$ as output. Aggregation occurs when a method receives as input two or more distance matrices A or ranked lists $\mathcal{R}$ and provides a single set of ranked lists $\mathcal{R}_{fa}$ as output.

4 PROPOSED SELECTION FRAMEWORK

The most effective combinations of retrieval methods occur when the inputs being combined present high effectiveness results and are complementary to each other. Our proposed framework consists in the use of both effectiveness and correlation measures aiming at selecting the best methods to be combined. Figure 1 illustrates the main steps of the proposed framework, which are identified by the black circles and described in the following subsections.

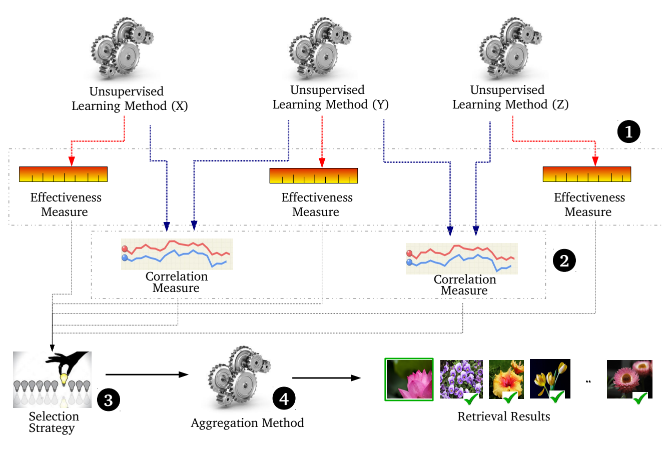

Figure 1: The proposed selection framework.

4.1 Effectiveness Measures

Effectiveness results are computed for all the available methods and are used to estimate the accuracy of the results. In this work, two effectiveness measures are considered.

4.1.1 Precision. In the image retrieval scenario, the precision is defined as $P_n = c/n$, where n is the number of retrieved images and c is the number of relevant images retrieved. The precision is computed for all the ranked lists and the arithmetic mean is subsequently calculated.

4.1.2 Mean Average Precision (MAP). The MAP value is provided by the arithmetic mean of the Average Precisions (AP) computed for all the ranked lists. Let q be a query item and let N_r be the number of relevant items in a collection for a given query q. Let $\langle r_i | i = 1, 2, \ldots, d \rangle$ be a ranked relevance vector of depth d, where r_i indicates the relevance of the ith ranked document scored as either 0 (not relevant) or 1 (relevant), the AP is defined in the Equation 1.

$$AP = \frac{1}{N_r} \sum_{i=1}^{d} \left(\frac{r_i}{i} \sum_{j=1}^{i} r_j \right) \qquad (1)$$

4.2 Correlation Measures

Correlation measures are computed for all the possible pairs of methods aiming at determining the less correlated results, which have more potential for combination. This is performed calculating the correlation for all the ranked lists available. Let two ranked lists τ_i and τ_j, a similarity value is computed in the interval $[0, 1]$. The higher the value, the higher the similarity of the lists. Three different correlation measures are considered in this work.

4.2.1 Jaccard. The Jaccard index is a statistic measure that computes the correlation between two ranked lists and is defined as shown in the Equation 2.

$$J(\tau_i, \tau_j) = \frac{|\tau_i \cap \tau_j|}{|\tau_i \cup \tau_j|} \qquad (2)$$

In this case, the idea is to offer a similarity score based on the number of elements in common of two ranked lists.

4.2.2 Kendall Tau. Given two ranked lists, the measure computes the number of discordant pairs between them. Let the image pair (img_x, img_y) and two ranked lists τ_i and τ_j. If $\tau_i(x) > \tau_j(y)$ and $\tau_i(y) > \tau_j(x)$, the pair (img_x, img_y) is a discordant pair regarding τ_i and τ_j.

Let K_d be the number of discordant pairs (Kendall Tau distance) and n be the size of these lists. To obtain a similarity score in the interval $[0, 1]$ (K_s), the following operation is performed.

$$K_s(\tau_i, \tau_j) = 1 - \frac{K_d(\tau_i, \tau_j)}{n_d}, \text{ where } n_d = \frac{n \times (n-1)}{2} \qquad (3)$$

4.2.3 Spearman. The Spearman similarity considers the difference between the position of an image in two ranked lists. Let τ_i and τ_j be two ranked lists and let n be the size of these lists, the Spearman similarity is defined by the Equation 4.

$$S(\tau_i, \tau_j) = 1 - \frac{\sum_{img_x \in \tau_i} |\tau_i(x) - \tau_j(x)|}{n \times (n+1)} \qquad (4)$$

4.3 Selection Strategy

Given a set of retrieval methods, a selection strategy aims to distinguish among high and low-effective combinations. In this work, a score is proposed to estimate the quality of a combination among the possibilities available. Each combination considers a pair of retrieval methods.

Let UL_1 and UL_2 be two different unsupervised learning methods, the measure provides a score based on the output results of the methods. The first hypothesis is that a good combination is given by two methods with high effectiveness results. Let M_{eff} be a selection measure and $eff(UL)$ be a function that returns the effectiveness of a method UL. Equation 5 summarizes this reasoning.

$$M_{eff}(UL_1, UL_2) = 1 + eff(UL_1) \times eff(UL_2) \qquad (5)$$

Another hypothesis is to consider that the best results are provided by the combination of methods that offer highly complementary results. Let M_{cor} be a selection measure and $cor(UL_1, UL_2)$ be a function that returns the correlation value of two ranked lists of two different unsupervised learning methods. The selection by correlation measure is computed as follows.

$$M_{cor}(UL_1, UL_2) = \frac{1}{1 + cor(UL_1, UL_2)} \qquad (6)$$

In order to combine the two previously presented hypotheses, a final selection measure M_{sel} is defined considering both effectiveness and correlation, like is shown in the Equation 7.

$$M_{sel}(UL_1, UL_2) = (M_{eff})^{\alpha} \times (M_{cor})^{\beta} \qquad (7)$$

The coefficients α and β are used aiming at applying different weights for each one of the considered measures.

4.4 Aggregation Method

In this paper, the term aggregation refers to an execution where a method receives more than one input providing a single output. Originally, the idea of aggregation is to combine complementary results offered by image descriptors [11]. However, this idea can also be applied to aggregate complementary information of different unsupervised learning methods. Figure 2 illustrates the combination model described.

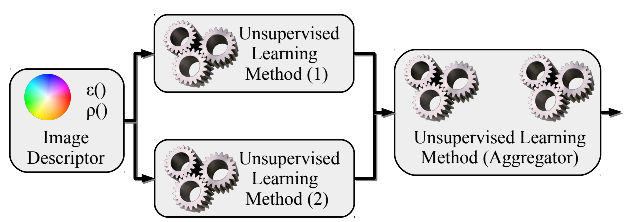

Figure 2: Aggregation method workflow.

Initially, a descriptor extracts the features from images and computes the distances between them, providing retrieval results for two different unsupervised learning methods. These methods output two ranked lists that are used together as the input of a third method called aggregator. The higher the complementarity of the outputs (1) and (2), better the output result of the aggregator. The aggregator can be one of the methods previously considered (1 and 2) or any other method. In this work, the method that achieved the best isolated effectiveness result was considered as aggregator.

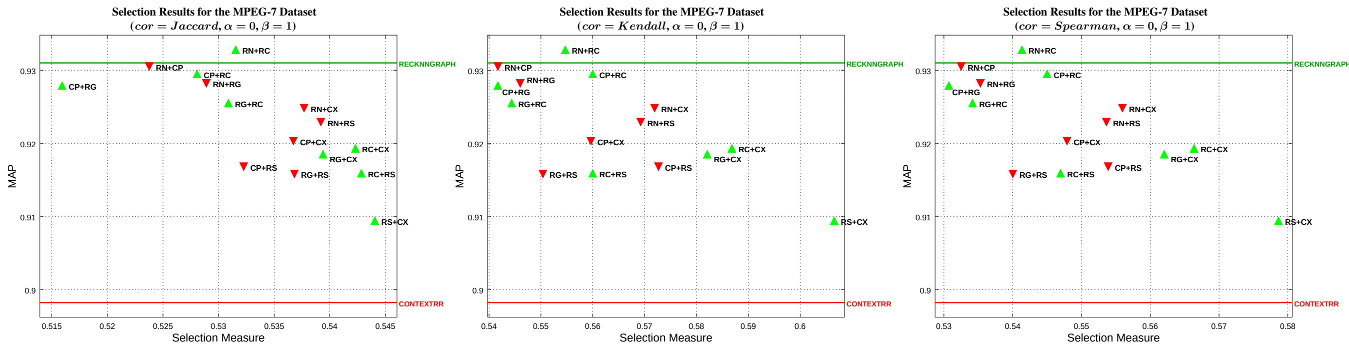

Figure 3: Selection results comparing three different correlation measures on the MPEG-7 dataset.

5 EXPERIMENTAL EVALUATION

This section describes the experimental evaluation conducted in order to assess the effectiveness of the proposed approach. Section 5.1 discusses the experimental protocol and Section 5.2 presents the obtained results.

5.1 Experimental Protocol

Two different datasets were considered in the experiments as shown by Table 1. We used the descriptors that achieved the higher effectiveness values for each dataset and do not offer saturated results. A comparison among descriptors can be found in the literature [9, 10, 21–25]. The number of images per class of each dataset was considered as the depth to compute the precision and correlation measures.

Table 1: Datasets and descriptors used in the evaluation.

Dataset	Descriptor	Description
Brodatz [31]	LAS (Local Activity Spectrum) [32]	Composed of 1776 images of 111 distinct textures divided into 16 blocks.
MPEG-7 [33]	ASC (Aspect Shape Context) [34]	A popular dataset composed of 1400 shapes divided into 70 classes.

A briefly description of the unsupervised learning methods used is presented in the Table 2. The combined methods are identified by mnemonics.

Table 2: Unsupervised methods used in the evaluation.

Name (Mnemonic)	Description
ContextRR (CX) [9]	Based on contextual information extracted from ranked lists, a more effective distance measure is computed.
CPRR (CP) [21]	Perform Cartesian product operations in kNN and reverse kNN neighborhood sets obtained from the ranked lists.
Ranked List Graph Distance (RG) [23]	Each ranked list is represented as a weighted sub-graph based on its top-k positions. The weight of the edges is used to increment the similarity scores between images.
ReckNNGraph (RN) [22]	A graph considering the reciprocal references in the top-k positions of the ranked lists is built. The graph structure is exploited through scores in order to analyze the dataset.
RL-Recommmendation (RC) [10]	Similarity information available in the ranked lists can be used to recommend images among themselves.
RL-Sim* (RS) [24, 25]	If two images are similar, their ranked lists should also be similar. Therefore, the distance between them is decreased.

It is expected that, higher the value provided by the selection measure, higher the effectiveness of the combination result. Therefore, graphs were built, where each point represents a different combination of methods. The final effectiveness value provided by the combination is evaluated using the MAP. The top and bottom lines indicate the highest and lowest isolated values, respectively. If a combination has an effectiveness value that is lower than one of the combined methods, the point is represented by an upside down red triangle. Otherwise, the combination is represented by an upside up green triangle. High-effective combination results are expected to produce points in the right-top side.

5.2 Experimental Results

Figure 3 presents three graphs for the MPEG-7 dataset considering different correlation measures. In this experiment, the effectiveness measures were not used ($\alpha = 0$). As can be noticed, the correlation by itself is not able to properly estimate the best and the worst combinations. The three measures are comparable, since the combinations selected as best and worst are the same. However, the Jaccard index seems to be slightly better than the others, once it places the best result (RN+RC) in the most right position among the graphs. Therefore, for the remaining experiments, we considered the Jaccard index as the default correlation measure.

The selection measure is evaluated considering different parameters (weights) for effectiveness and correlation. Figure 4 illustrates six different graphs, each row presents three graphs for a different dataset. The leftmost graphs consider the same value for the coefficients ($\alpha = \beta = 1$). These parameters have not achieved the best results, although several upside up green triangles indicate the effectiveness potential of combinations. Another hypothesis is that the effectiveness by itself is able to provide a better estimation than the previous approaches. As an indication, good results can be seen in the middle graphs ($\alpha = 1$, $\beta = 0$). However, the results can be further improved by incorporating correlation information. The results can be seen in the rightmost graphs which consider both effectiveness and correlation with double weight for effectiveness ($\alpha = 2$, $\beta = 1$).

For comparison purposes, the same experiments were reproduced changing eff from precision to MAP and are presented by Figure 5. Although the results show that the effectiveness measures are comparable, the MAP has presented a negative impact for the MPEG-7 dataset, since the best combination is not in the most right-top position like in the results obtained considering precision.

In order to illustrate the visual results of the combinations compared with the original descriptor and isolated methods, Figure 6 shows a ranked list example for the *ReckNNGraph + RL-Recom* combination, which presented the best result compared to the available methods. The query images are illustrated in green borders and the incorrect images in red borders. An incremental and noticeable gain can be seen from the descriptor until reaching the combination.

Figure 4: Selection results comparing different weights for effectiveness and correlation (α and β) using $eff = Precision$.

Figure 5: Selection results comparing different weights for effectiveness and correlation (α and β) using $eff = MAP$.

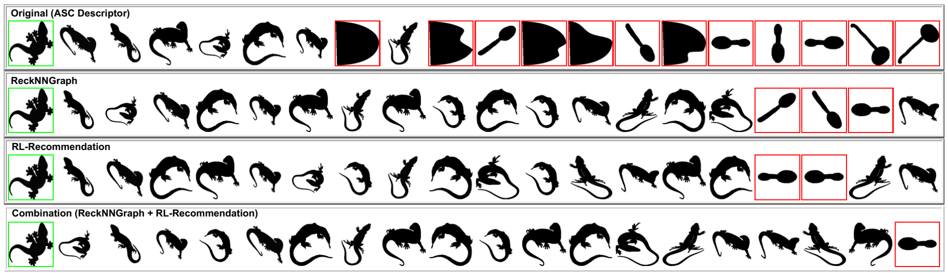

Figure 6: Example of image retrieval results for the MPEG-7 dataset showing the impact of the method combination.

6 CONCLUSIONS

In this paper, a strategy to select and aggregate combinations of unsupervised learning methods for image retrieval was presented. The results demonstrated that the best combination can be selected using an approach that combines correlation and effectiveness measures. Although the correlation measures are unsupervised, the effectiveness measures use labeled data. As future work, we intend to use unsupervised effectiveness estimations instead, allowing the selection to be performed in a fully unsupervised way. Another way to expand the evaluation is to combine not only pairs of methods, but larger sets. We also intend to evaluate combinations in other multimedia scenarios (audio, video, and others).

ACKNOWLEDGMENT

The authors are grateful to the São Paulo Research Foundation - FAPESP (grants 2017/02091-4 and 2013/08645-0).

REFERENCES

[1] R. Datta, D. Joshi, J. Li, and J. Z. Wang, "Image retrieval: Ideas, influences, and trends of the new age," *ACM Comput. Surv.*, vol. 40, no. 2, pp. 5:1–5:60, May 2008.
[2] R. D. S. Torres and A. X. Falcão, "Content-based image retrieval: Theory and applications," *Revista de Informática Teórica e Aplicada*, vol. 13, pp. 161–185, 2006.
[3] J. Huang, S. R. Kumar, M. Mitra, W.-J. Zhu, and R. Zabih, "Image indexing using color correlograms," in *CVPR*, 1997, pp. 762–768.
[4] H. Ling and D. W. Jacobs, "Shape classification using the inner-distance," *IEEE TPAMI*, vol. 29, no. 2, pp. 286–299, 2007.
[5] H. Bay, A. Ess, T. Tuytelaars, and L. Van Gool, "Speeded-up robust features (surf)," *Computer vision and image understanding*, vol. 110, no. 3, pp. 346–359, 2008.
[6] Y. Jia, E. Shelhamer, J. Donahue, S. Karayev, J. Long, R. Girshick, S. Guadarrama, and T. Darrell, "Caffe: Convolutional architecture for fast feature embedding," *arXiv preprint arXiv:1408.5093*, 2014.
[7] S. C. Hoi, W. Liu, and S.-F. Chang, "Semi-supervised distance metric learning for collaborative image retrieval and clustering," *ACM Transactions on Multimedia Computing and Communication Applications*, vol. 6, no. 3, pp. 18:1–18:26, August 2010.
[8] P. Kontschieder, M. Donoser, and H. Bischof, "Beyond pairwise shape similarity analysis," in *ACCV*, 2009, pp. 655–666.
[9] D. C. G. Pedronette and R. da S. Torres, "Exploiting contextual information for image re-ranking and rank aggregation," *International Journal of Multimedia Information Retrieval*, vol. 1, no. 2, pp. 115–128, 2012.
[10] L. P. Valem, D. C. G. Pedronette, R. d. S. Torres, E. Borin, and J. Almeida, "Effective, efficient, and scalable unsupervised distance learning in image retrieval tasks," *ICMR*, 2015.
[11] F. F. Faria, A. Veloso, H. M. Almeida, E. Valle, R. d. S. Torres, M. A. Gonçalves, and W. Meira, Jr., "Learning to rank for content-based image retrieval," in *Proceedings of the International Conference on Multimedia Information Retrieval*, ser. MIR '10. New York, NY, USA: ACM, 2010, pp. 285–294.
[12] H. J. Escalante, C. A. Hérnadez, L. E. Sucar, and M. Montes, "Late fusion of heterogeneous methods for multimedia image retrieval," in *Proceedings of the 1st ACM International Conference on Multimedia Information Retrieval*, ser. MIR '08. New York, NY, USA: ACM, 2008, pp. 172–179.
[13] J. A. Vargas Muñoz, R. da Silva Torres, and M. A. Gonçalves, "A soft computing approach for learning to aggregate rankings," in *Proceedings of the 24th ACM International on Conference on Information and Knowledge Management*, ser. CIKM '15. New York, NY, USA: ACM, 2015, pp. 83–92.
[14] D. C. G. Pedronette and R. da S. Torres, "Exploiting clustering approaches for image re-ranking," *Journal of Visual Languages and Computing*, vol. 22, no. 6, pp. 453–466, 2011.
[15] X. Yang and L. J. Latecki, "Affinity learning on a tensor product graph with applications to shape and image retrieval," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR'2011)*, 2011, pp. 2369–2376.
[16] X. Yang, S. Koknar-Tezel, and L. J. Latecki, "Locally constrained diffusion process on locally densified distance spaces with applications to shape retrieval." in *CVPR*, 2009, pp. 357–364.
[17] X. Yang, X. Bai, L. J. Latecki, and Z. Tu, "Improving shape retrieval by learning graph transduction," in *ECCV*, vol. 4, 2008, pp. 788–801.
[18] G. Park, Y. Baek, and H.-K. Lee, "Re-ranking algorithm using post-retrieval clustering for content-based image retrieval," *Information Processing and Management*, vol. 41, no. 2, pp. 177–194, 2005.
[19] W. Voravuthikunchai, B. Crémilleux, and F. Jurie, "Image re-ranking based on statistics of frequent patterns," in *Proceedings of International Conference on Multimedia Retrieval*, ser. ICMR '14. New York, NY, USA: ACM, 2014, pp. 129:129–129:136.
[20] D. C. G. Pedronette and R. da S. Torres, "Image re-ranking and rank aggregation based on similarity of ranked lists," in *Computer Analysis of Images and Patterns (CAIP'2011)*, vol. 6854, 2011, pp. 369–376.
[21] L. P. Valem and D. C. G. Pedronette, "Unsupervised similarity learning through cartesian product of ranking references for image retrieval tasks," *Conference on Graphics, Image and Patterns (SIBGRAPI)*, 2016.
[22] D. C. G. Pedronette, O. A. Penatti, and R. da S. Torres, "Unsupervised manifold learning using reciprocal knn graphs in image re-ranking and rank aggregation tasks," *Image and Vision Computing*, vol. 32, no. 2, pp. 120 – 130, 2014.
[23] D. C. G. Pedronette, J. Almeida, and R. da S. Torres, "A graph-based ranked-list model for unsupervised distance learning on shape retrieval," *Pattern Recognition Letters*, vol. 83, Part 3, pp. 357 – 367, 2016, efficient Shape Representation, Matching, Ranking, and its Applications.
[24] D. C. G. Pedronette and R. da S. Torres, "Image re-ranking and rank aggregation based on similarity of ranked lists," *Pattern Recognition*, vol. 46, no. 8, pp. 2350–2360, 2013.
[25] C. Y. Okada, D. C. G. a. Pedronette, and R. da S. Torres, "Unsupervised distance learning by rank correlation measures for image retrieval," in *Proceedings of the 5th ACM on International Conference on Multimedia Retrieval*, ser. ICMR '15. New York, NY, USA: ACM, 2015, pp. 331–338.
[26] L. P. Valem and D. C. G. Pedronette, "An unsupervised distance learning framework for multimedia retrieval," ser. ICMR '17, 2017.
[27] E. A. Fox and J. A. Shaw, "Combination of multiple searches," in *The Second Text REtrieval Conference (TREC-2)*, ser. NIST Special Publication, vol. 500-215. NIST, 1994, pp. 243–252.
[28] R. Fagin, R. Kumar, M. Mahdian, D. Sivakumar, and E. Vee, "Comparing and aggregating rankings with ties," in *23th ACM SIGMOD Symposium on Principles of Database Systems (PODS'04)*, 2004, pp. 47–58.
[29] Y.-T. Liu, T.-Y. Liu, T. Qin, Z.-M. Ma, and H. Li, "Supervised rank aggregation," in *International Conference on World Wide Web (WWW'2007)*, 2007, pp. 481–490.
[30] D. C. G. Pedronette and R. d. S. Torres, "Combining re-ranking and rank aggregation methods for image retrieval," *Multimedia Tools and Applications*, vol. 75, no. 15, pp. 9121–9144, 2016.
[31] P. Brodatz, *Textures: A Photographic Album for Artists and Designers.* Dover, 1966.
[32] B. Tao and B. W. Dickinson, "Texture recognition and image retrieval using gradient indexing," *JVCIR*, vol. 11, no. 3, pp. 327–342, 2000.
[33] L. J. Latecki, R. Lakmper, and U. Eckhardt, "Shape descriptors for non-rigid shapes with a single closed contour," in *CVPR*, 2000, pp. 424–429.
[34] H. Ling, X. Yang, and L. J. Latecki, "Balancing deformability and discriminability for shape matching," in *ECCV*, vol. 3, 2010, pp. 411–424.