

# An Unsupervised Genetic Algorithm Framework for Rank Selection and Fusion on 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  
[lucas.valem@unesp.br](mailto:lucas.valem@unesp.br), [daniel.pedronette@unesp.br](mailto:daniel.pedronette@unesp.br)

## ABSTRACT

Despite the major advances on feature development for low and mid-level representations, a single visual feature is often insufficient to achieve effective retrieval results in different scenarios. Since diverse visual properties provide distinct and often complementary information for a same query, the combination of different features, including handcrafted and learned features, has been establishing as a relevant trend in image retrieval. An intrinsic difficulty task consists in selecting and combining features that provide a high-effective result, which is often supported by supervised learning methods. However, in the absence of labeled data, selecting and fusing features in a completely unsupervised fashion becomes an essential, although very challenging task. The proposed genetic algorithm employs effectiveness estimation measures as fitness functions, making the evolutionary process fully unsupervised. Our approach was evaluated considering 3 public datasets and 35 different descriptors achieving relative gains up to +53.96% in scenarios with more than 8 billion possible combinations of rankers. The framework was also compared to different baselines, including state-of-the-art methods.

## CCS CONCEPTS

• Information systems → Multimedia and multimodal retrieval;

## KEYWORDS

content-based image retrieval; genetic algorithm; unsupervised learning; re-ranking; rank-aggregation; effectiveness estimation

### ACM Reference Format:

Lucas Pascotti Valem and Daniel Carlos Guimarães Pedronette. 2019. An Unsupervised Genetic Algorithm Framework for Rank Selection and Fusion on Image Retrieval. In *International Conference on Multimedia Retrieval (ICMR '19)*, June 10–13, 2019, Ottawa, ON, Canada. ACM, New York, NY, USA, 5 pages. <https://doi.org/10.1145/3323873.3325022>

## 1 INTRODUCTION

Technologies to produce, acquire, and share images have been consistently and significantly improved for decades. Such advancements have triggered the use of content-based image retrieval (CBIR) as something indispensable in the modern world [11]. In general, CBIR systems rely on the extraction of visual features. Mainly due to the diverse aspects involved in the human visual

perception, a single feature is generally not enough to encode all the relevant aspects in a retrieval scenario [19, 37]. Despite the significant attention that deep learning approaches have been receiving due to their generally high-effective results, the state-of-art is not achieved in all cases [48], turning the selection and combination of visual features an attractive and interesting alternative.

In order to reach more effective retrieval results, a considerable amount of fusion methods have been recently proposed [37, 46, 52]. The fusion strategies are broadly divided into two categories [1, 39]: early and late fusion. Methods based on early fusion usually combine raw feature vectors, different from the late fusion ones that combine any representation obtained from the feature vectors (e.g. ranked lists, distance or similarity matrices).

Another task of crucial relevance consists in selecting the visual features to be combined. In unsupervised late fusion scenarios [34, 52, 54], the selection of visual features is often performed *ad hoc* or by the attribution of different weights for each rank. In fact, the task of choosing the features that best fit the needs for a given problem is very challenging [37]. Therefore, optimization strategies as genetic algorithms (GA) [4, 14, 20] and genetic programming (GP) approaches [9, 13] represent very attractive solutions. Genetic algorithms are a specific category of evolutionary algorithms inspired in the biological theories of natural selection and evolution [16]. They are often used for optimization operations, since they do not require exhaustive searches, which can be too time demanding in many real-world applications. The use of genetic approaches for rank selection and fusion in unsupervised scenarios is still very challenging, since most of fitness functions are supervised and require labeled data [9, 13].

In this work, the Unsupervised Genetic Algorithm Framework for Rank Selection and Fusion (UGAF-RSF) is proposed. A rank-based strategy is used for modeling, selecting and fusing the retrieval results of different features. The proposed genetic algorithm employs an unsupervised fitness function based on effectiveness estimation measures. Such measures do not require any training data or previous information about the features, making the evolutionary process fully unsupervised. A wide experimental analysis was conducted considering 3 public datasets and 35 different features (including deep learning, local, and global descriptors). High-effective results were achieved with relative gains up to +53.96% in relation to the best isolated feature, considering scenarios with more than 8 billion possible combinations, where the complete brute-force search would be impracticable. The proposed approach was also compared to different baselines, including state-of-the-art methods.

The paper is organized as follows. Section 2 describes the proposed unsupervised genetic algorithm framework for selective rank fusion. Section 3 shows and discusses the experiments. Finally, Section 4 states conclusions and possible future works.

## 2 PROPOSED FRAMEWORK

This section presents the Unsupervised Genetic Algorithm Framework for Rank Selection and Fusion (UGAF-RSF) proposed in this

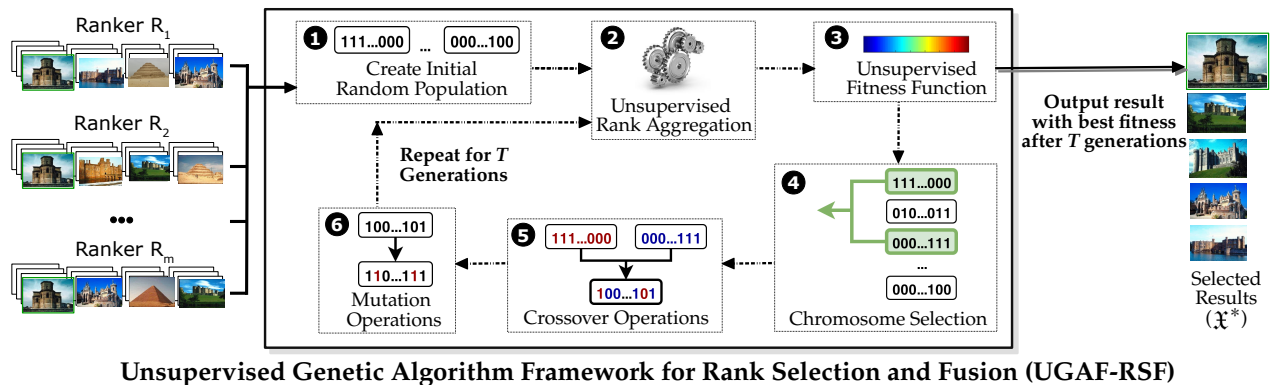
Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components 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](mailto:permissions@acm.org).

ICMR '19, June 10–13, 2019, Ottawa, ON, Canada

© 2019 Association for Computing Machinery.

ACM ISBN 978-1-4503-6765-3/19/06...\$15.00

<https://doi.org/10.1145/3323873.3325022>



Unsupervised Genetic Algorithm Framework for Rank Selection and Fusion (UGAF-RSF)

Figure 1: General view of the framework.

paper. Figure 1 illustrates the main steps which define the framework. Each feature is extracted by a descriptor (as defined in [42]) and modeled as a ranker, which computes a set of ranked lists used as the input of our approach. The possible combinations among the  $m$  available rankers are modeled as binary chromosomes. The combination defined by each chromosome is computed through an unsupervised late fusion based on rank aggregation. The retrieval results are evaluated by effectiveness estimation measures, which are used to perform a selection step. Based on the selected individuals, a next generation is created through crossover and mutation operations. After  $T$  generations, the fusion results defined by the selected rankers ( $\mathcal{X}^*$ ) are taken as output. The main steps of the framework are described as follows.

**1. Chromosome Representation:** each individual is described as a chromosome which is represented by a bit-string. The size of the string is equal to the number of rankers available. For each ranker, there is a bit (genome) that indicates if it is part of the combination ("1") or not ("0"). Before the iterations start, a initial population is randomly generated.

**2. Unsupervised Rank Aggregation:** for every iteration, the combination defined by each of the chromosomes is computed using an unsupervised rank aggregation algorithm. In this work, the CPRR (Cartesian Product of Ranking References) [43] method is used. The CPRR method is a recent unsupervised rank aggregation method, which employs Cartesian product operations for maximizing the relationships among images ranked at top- $k$  positions, achieving effective and efficient results. Additionally, it is worth mentioning that the proposed approach is flexible and can use different aggregation methods for the fusion of rankers results.

**3. Unsupervised Fitness Function:** the fitness functions are defined by effectiveness estimation measures, which allow a completely unsupervised selection along iterations. Such measures do not require any labeled data and exploit the cluster hypothesis [18] that is based on the idea that similar items tend to be relevant among themselves and, therefore, they should have elements in common in the top- $k$  positions. Each chromosome is executed with the rank aggregation method (CPRR in this case) and the output is submitted to the fitness function.

The effectiveness measure is used with the intent of selecting ranked lists with high effectiveness. This work considers two different effectiveness measures: Authority Measure [36] and Reciprocal Density [33]. The Authority Score uses a graph representing the ranking references at top positions of ranked lists and estimates the effectiveness according to the density of the graph. The Reciprocal

Density considers a similar idea, but a weight is assigned to the occurrence of each reciprocal neighbor according to its position in the ranked lists. In this work we propose a Hybrid Score, which combines both the Authority and Reciprocal density measures as follows:  $Hyb(q, k) = (Auth(q, k) + 1) \times (Rec(q, k) + 1)$ .

**4. Chromosome Selection:** the selection stage filters a set of individuals that are destined for crossover or to be kept in the population of the next generation. For this work, the tournament selection was used to select the best individuals. This method works by randomly picking three different chromosomes and selecting the one with the highest fitness. The elitism is applied in this step, where a percentage of the individuals with best fitness of the entire population are obligatory kept for the next iteration.

**5. Crossover Operations:** reproduction among chromosomes is crucial to keep the diversity of the population and, at the same time, improve the convergence and stability of the algorithm. There are different types of crossover available. In this work, the uniform crossover is applied to generate new individuals. In this form of crossover, each bit is independently chosen from any of the two parents. Each bit is chosen from either parent with equal probability.

**6. Mutation Operations:** the mutation is of critical importance to keep the diversity of the population and also to prevent local maximum. For the mutation, each bit has a mutation probability that defines the probability of each bit (genome) being flipped (i.e. if the bit is "0", it is changed to "1" and vice versa).

### 3 EXPERIMENTAL EVALUATION

The experimental analysis considered three different datasets:

- **Flowers** [27] (1,360 images): composed of 17 species of flowers with 80 images of each presenting pose and light variations. This dataset is distributed by the University of Oxford;
- **Corel5k** [22] (5,000 images): composed of 50 categories with 100 images for each class, including diverse scene content such as fireworks, bark, microscopy images, tiles, trees, etc;
- **UKBench** [28] (10,200 images): composed of 2,550 objects or scenes. Each object/scene is captured 4 times from different view-points, distances, and illumination conditions.

Most of them are frequently used for image retrieval tasks. The MAP was computed considering every image in the dataset as a query. Besides the MAP, N-S Score was also used for the UKBench dataset with the intent to facilitate the comparison of our results with state-of-the-art baselines.

A wide variety of visual features were considered, including different categories: global, local, and deep learning. Table 1 presents

**Table 1: Descriptors used in the experimental evaluation.**

Category	Type	Descriptor	Original MAP (%)		
			Flowers	Core5k	UKbench
Global	Color	ACC [17]	18.99	23.44	87.72
		SPACC [17, 25]	19.20	23.86	85.30
		CLD [8]	18.54	17.86	59.58
		SCD [8]	10.25	14.56	83.04
		SCH [8]	13.43	17.56	48.98
		FOH [25, 44]	11.42	15.87	57.05
		BIC [40]	25.56	---	80.46
	Shape	PHOG [10, 25]	14.74	15.80	41.60
	Texture	LBP [29]	10.34	14.83	47.19
		SPLBP [25, 29]	10.92	15.41	52.14
		EHD [26]	12.46	16.80	44.10
	Color and Texture	CEDD [6]	20.48	23.00	70.45
		SPCEDD [6, 25]	21.94	28.70	74.98
		FCFH [7]	20.56	23.93	73.70
SPFCTH [7, 25]		21.73	26.43	77.78	
JCD [51]		20.89	24.73	74.85	
SPJCD [25, 51]		22.56	28.02	76.67	
COMO [45]		21.83	21.05	79.77	
Holistic	GIST [30]	9.82	15.98	45.44	
Local	Bag of Words	SIFT [24]	28.47	12.60	74.52
		VOC [47]	---	---	91.14
Deep Learning	Convolut. Neural Networks trained on Imagenet	CNN-SENet <sup>1</sup>	43.16	56.92	92.15
		CNN-ResNet <sup>1</sup>	51.83	64.81	94.54
		CNN-FBResNet <sup>1</sup>	<b>52.56</b>	64.21	93.88
		CNN-ResNeXt <sup>1</sup>	51.91	62.39	93.67
		CNN-DPNet <sup>1</sup>	50.93	<b>65.15</b>	90.47
		CNN-VGGNet <sup>1</sup>	39.05	47.85	87.99
		CNN-BnVGGNet <sup>1</sup>	41.87	52.72	89.24
		CNN-InceptionV4 <sup>1</sup>	42.35	58.66	86.82
		CNN-InceptionResNet <sup>1</sup>	42.20	61.17	87.23
		CNN-BnInception <sup>1</sup>	46.58	46.60	91.84
		CNN-NASnet-Large <sup>1</sup>	40.74	53.55	86.90
		CNN-AlexNet <sup>1</sup>	46.04	37.67	85.57
		CNN-Xception <sup>1</sup>	47.31	54.44	90.83
		Pooled CNN	CNN-OLDFP [38]	---	---

each of them followed by their respective types, references, and the MAP obtained for each dataset.

The deep learning results were obtained using PyTorch [31]<sup>1</sup>, one of the most popular open-source frameworks for machine learning. All of the networks were trained on ImageNet [12], a dataset commonly used for training general purpose convolutional neural networks. The distances were extracted from the feature vectors obtained from the last layer before the classification layer. For the CNN-OLDFP, the feature vectors are provided by [38], the network mixes techniques of deep learning and bag of words aiming at optimizing the effectiveness of the results.

The parameters used in most of the conducted experiments are described as follows: 100 individuals for the population size; 30 generations per execution; hybrid fitness function; 70% of crossover rate; 0.08% of mutation rate; 0.05% of elitism rate;  $k = 50$  for Flowers and Core5k, and  $k = 4$  for UKBench [43]. Since genetic algorithms are stochastic, all the results reported in this paper are given by the arithmetic mean of 10 executions.

All the results were obtained through a completely unsupervised evolutionary process. For evaluation purposes, the results in the graphs report the highest MAP of the population along the generations. The results reported in the tables are from the combination with the highest fitness in the last population ( $\mathfrak{X}^*$ ). For this reason, the results from graphs, which show the evaluation of parameters, may differ from the ones in the tables, which show the final selection results.

The Flowers dataset was chosen to evaluate the parameters, since it is the smallest dataset in the experimental protocol. Along

<sup>1</sup> [github.com/Cadene/pretrained-models.pytorch](https://github.com/Cadene/pretrained-models.pytorch)

the experimental evaluation, it can be seen that our results are far higher than the best isolated descriptor (CNN-FBResNet with MAP of **52.56%** for Flowers dataset).

- **Fitness function:** an experiment was conducted for comparing the results of each of the three effectiveness estimation measures when considered as fitness function of the genetic programming. Figure 2 presents the results by reporting the highest MAP of the population along 30 generations. The results reveal that Hybrid measure presents the best results being slightly above Authority. Therefore, we considered our proposed Hybrid measure as the fitness function for all the remaining experiments;

- **Population size and number of generations:** Figure 3 presents the results of the genetic algorithm along 60 generations for five different population sizes (25, 50, 100, 150, and 200). As can be seen, there is few gain in executing more than 30 generations, so we considered this value for all the remaining executions. The higher the population size, higher the results obtained. However, there is a trade-off between efficiency and effectiveness in this regard. Therefore, for all the remaining experiments, we adopted 100 as the default population size, since it presented results slightly close to 150 and 200;

- **Crossover rate:** Figure 4 presents an analysis of the crossover rate. As can be seen, the variation of the crossover rate does not have a great impact in the results. We considered a crossover rate of 70% for all the remaining experiments, once this value presented slightly better results;

- **Mutation rate:** Figure 5 presents an analysis of the mutation rate. As can be seen, the variation of the mutation rate also does not have a great impact in the results. However, 8% revealed the best results and we decided to keep it for all the remaining experiments;

- **Evolutionary strategy:** three different strategies are considered aiming at improving the results of the genetic approach: (1) in this strategy, the elitism increases along the generations; (2) crossover and mutation decrease along the generations; (3) both strategies 1 and 2 are used together. Figure 6 evaluates the impact of the proposed strategies on the results. The strategy (1) revealed the best results and is used for the final results that are reported in the tables ( $\mathfrak{X}^*$ );

- **Maximum elitism rate:** the maximum elitism rate determines the maximum value that the elitism can reach when using the strategy (1). This is evaluated by the Figure 7. We considered 20% for all the remaining experiments;

The results obtained by our proposed framework is presented in Table 2, where ( $\mathfrak{X}^*$ ) indicates the most frequently selected combination and the average MAP of 10 executions. The relative gain is computed in relation to the best isolated descriptor. Notice that our method achieved positive gains in all of the circumstances.

**Table 2: The most frequently selected combination and the average MAP for each dataset.**

Dataset	Frequently Selected Combination ( $\mathfrak{X}^*$ )	MAP (%)	Relative Gain
Flowers	ACC + AlexNet + BnInception + ResNeXt + ResNet + SIFT	<b>80.92 ± 0.6005</b>	<b>+53.96%</b>
Core5k	CEDD + BnInception + FBResNet + InceptionV4 + NASnet-Large + ResNeXt + GIST + LBP + SCD	<b>91.45 ± 0.3738</b>	<b>+40.37%</b>
UKBench	AlexNet + BnInception + InceptionV4 + OLDFP + ResNeXt + SENet + VGGNet + COMO + FOH + GIST + SCD + SIFT + SPACC + VOC	<b>99.25 ± 0.0573</b>	<b>+1.54%</b>

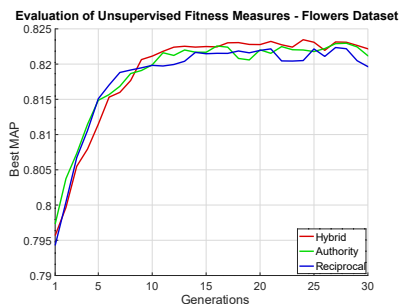


Figure 2: Fitness functions.

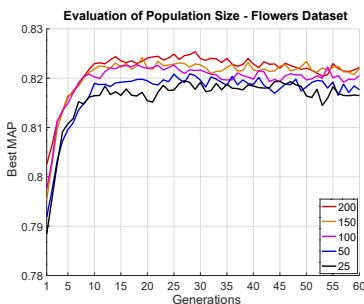


Figure 3: Populations size.

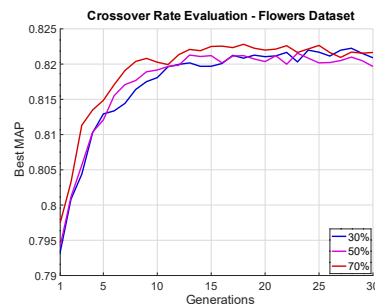


Figure 4: Crossover rate.

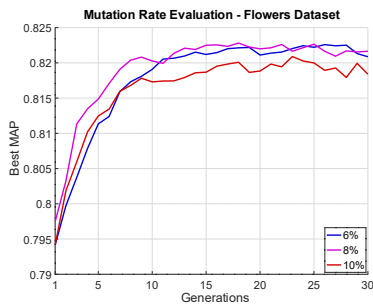


Figure 5: Mutation rate.

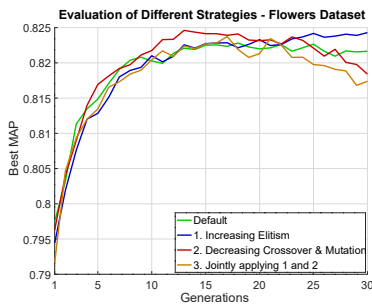


Figure 6: Evolutionary strategies.

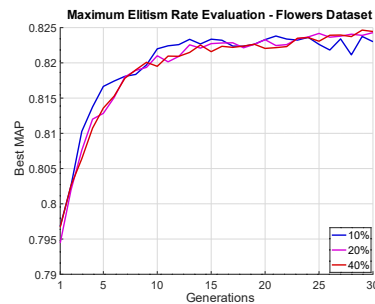


Figure 7: Maximum elitism rate.

The obtained results were compared to different methods presented in the literature. We compared our approach with feature selection techniques. After applying PCA and L2-norm regularization for all the feature vectors, the feature selection methods were employed to select the top-100 most relevant features among the available ones. From the euclidean distance of the new vectors, we computed the ranked lists for the results and submitted them as input for the CPRR re-ranking algorithm. Besides that, as these methods are very expensive for a large number of features, the results are presented just for the Flowers and Corel5k datasets. A comparison of the proposed approach with feature selection methods based on early fusion is presented in Table 3.

Table 3: Comparison with feature selection techniques.

Dataset	MAP (%)						Proposed Framework
	Best Desc.	Laplace [15]	SPEC [53]	MCFS [5]	UDFS [50]	NDFS [21]	
Flowers	52.56	61.28	49.67	55.95	63.97	65.41	<b>80.92 ± 0.6005</b>
Corel5k	65.15	78.40	63.99	84.74	78.20	87.17	<b>91.45 ± 0.3738</b>

Table 4: State-of-the-art in the UKBench dataset (N-S Score).

N-S scores for the state-of-the-art methods					
Zheng et al. [55]	Wang et al. [46]	Sun et al. [41]	Paulin et al. [32]	Zhang et al. [52]	Zheng et al. [54]
3.57	3.68	3.76	3.76	3.83	3.84
Bai et al. [2]	Xie et al. [49]	Liu et al. [23]	Pedronette et al. [35]	Bai et al. [3]	
3.86	3.89	3.92	3.93	3.94	
<b>Proposed Framework</b>			<b>3.95 ± 0.3206</b>		

We also compared our results with state-of-the-art methods for the UKBench dataset. Our proposed method considers all of the rankers as input and performs a completely unsupervised fusion, while other methods consider a small set of them, that is generally selected *ad hoc*. Several state-of-the-art methods are included in the comparison. Table 4 presents a comparison on the UKBench dataset. Our proposed method achieved a N-S Score (equivalent to P@4) of **3.95**. Notice that UGAF-RSF achieved the best results when compared to the baselines.

## 4 CONCLUSION

In this work, we have presented a fully unsupervised genetic algorithm framework for content-based image retrieval that innovates by using effectiveness estimation measures as fitness functions. A wide experimental analysis was conducted for evaluating the parameters and each of the datasets. The results reveal the great potential of the framework with high-effective values (relative gains up to +53.96%) in scenarios with billions of combinations available, where the complete brute-force search would be totally impracticable. In addition to that, the achieved scores are closer or higher to the baselines and the state-of-the-art in the majority of the cases.

As future work, we intend to evaluate the approach in other multimedia scenarios (e.g. textual, sound, video). We also intend to analyse the impact of different rank-aggregation methods and chromosome representations (e.g. tree structures).

## 5 ACKNOWLEDGMENTS

The authors are grateful to the São Paulo Research Foundation - FAPESP (grants #2017/02091-4, #2018/15597-6, and #2017/25908-6), the Brazilian National Council for Scientific and Technological Development - CNPq (grant #308194/2017-9), and Petrobras (grant #2017/00285-6).

## REFERENCES

- [1] Pradeep K. Atrey, M. Anwar Hossain, Abdulmotaleb El Saddik, and Mohan S. Kankanhalli. 2010. Multimodal fusion for multimedia analysis: a survey. *Multimedia Systems* 16, 6 (01 Nov 2010), 345–379.
- [2] S. Bai and X. Bai. 2016. Sparse Contextual Activation for Efficient Visual Re-ranking. *IEEE Trans. on Image Processing (TIP)* 25, 3 (2016), 1056–1069.
- [3] Song Bai, Xiang Bai, Qi Tian, and Longin Jan Latecki. 2017. Regularized Diffusion Process for Visual Retrieval. In *Conf. on Artificial Intelligence (AAAI)*. 3967–3973.
- [4] Waad Bouaguel, Afef Ben Brahim, and Mohamed Limam. 2013. Feature Selection By Rank Aggregation and Genetic Algorithms. In *KDIR/KMIS 2013 - Proceedings of the International Conference on Knowledge Discovery and Information Retrieval and the International*.
- [5] Deng Cai, Chiyuan Zhang, and Xiaofei He. 2010. Unsupervised Feature Selection for Multi-cluster Data. In *Proceedings of the 16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '10)*. 333–342.
- [6] Savvas A. Chatzichristofis and Yiannis S. Boutalis. 2008. CEDD: color and edge directivity descriptor: a compact descriptor for image indexing and retrieval. In *Proceedings of the 6th international conference on Computer vision systems (ICVS'08)*. 312–322.
- [7] Savvas A. Chatzichristofis and Yiannis S. Boutalis. 2008. FCTH: Fuzzy Color and Texture Histogram A Low Level Feature for Accurate Image Retrieval. In *Int. Workshop on Image Analysis for Multimedia Interactive Services*. 191–196.
- [8] Leszek Cieplinski. 2001. MPEG-7 Color Descriptors and Their Applications. In *Computer Analysis of Images and Patterns*, Wladyslaw Skarbek (Ed.). Springer Berlin Heidelberg, Berlin, Heidelberg, 11–20.
- [9] Ricardo da S. Torres, Alexandre X. Falcão, Marcos A. Gonçalves, João P. Papa, Baoping Zhang, Weiguo Fan, and Edward A. Fox. 2009. A genetic programming framework for content-based image retrieval. *Pattern Recognition* 42, 2, 283 – 292. Learning Semantics from Multimedia Content.
- [10] N. Dalal and B. Triggs. 2005. Histograms of oriented gradients for human detection. In *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, Vol. 1. 886–893 vol. 1.
- [11] Ritendra Datta, Dhiraj Joshi, Jia Li, and James Z. Wang. 2008. Image retrieval: Ideas, influences, and trends of the new age. *Comput. Surveys* 40, 2 (2008), 5:1–5:60.
- [12] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei. 2009. ImageNet: A Large-Scale Hierarchical Image Database. In *CVPR09*.
- [13] Cristiano D. Ferreira, Jeferson Alex dos Santos, Ricardo da S. Torres, Marcos André Gonçalves, R. C. Rezende, and Weiguo Fan. 2011. Relevance feedback based on genetic programming for image retrieval. *Pattern Recognition Letters* 32, 1 (2011), 27–37.
- [14] R. Gali, M. L. Dewal, and R. S. Anand. 2012. Genetic Algorithm for Content Based Image Retrieval. In *2012 Fourth International Conference on Computational Intelligence, Communication Systems and Networks*. 243–247.
- [15] Xiaofei He, Deng Cai, and Partha Niyogi. 2005. Laplacian Score for Feature Selection. In *Proceedings of the 18th International Conference on Neural Information Processing Systems (NIPS'05)*. 507–514.
- [16] John H. Holland. 1975. *Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control and Artificial Intelligence*. University of Michigan Press, Ann Arbor, MI.
- [17] Jing Huang, S. Ravi Kumar, Mandar Mitra, Wei-Jing Zhu, and Ramin Zabih. 1997. Image Indexing Using Color Correlograms. In *CVPR'97*. 762–768.
- [18] N. Jardine and C.J. van Rijsbergen. 1971. The use of hierarchic clustering in information retrieval. *Information Storage and Retrieval* 7, 5 (1971), 217 – 240.
- [19] M. L. Kherfi and D. Ziou. 2006. Relevance feedback for CBIR: a new approach based on probabilistic feature weighting with positive and negative examples. *IEEE Transactions on Image Processing* 15, 4 (2006), 1017–1030.
- [20] Preeti Kushwaha and Rashmi R. Welekar. 2016. Feature Selection for Image Retrieval based on Genetic Algorithm. *International Journal of Interactive Multimedia and Artificial Intelligence* 4, 2, 16–21.
- [21] Zechao Li, Yi Yang, Jing Liu, Xiaofang Zhou, and Hanqing Lu. 2012. Unsupervised Feature Selection Using Nonnegative Spectral Analysis. In *Proceedings of the Twenty-Sixth AAAI Conference on Artificial Intelligence*.
- [22] Guang-Hai Liu and Jing-Yu Yang. 2013. Content-based image retrieval using color difference histogram. *Pattern Recognition* 46, 1 (2013), 188 – 198.
- [23] Z. Liu, S. Wang, L. Zheng, and Q. Tian. 2017. Robust ImageGraph: Rank-Level Feature Fusion for Image Search. *IEEE Transactions on Image Processing* 26, 7 (2017), 3128–3141.
- [24] D.G. Lowe. 1999. Object recognition from local scale-invariant features. In *IEEE International Conference on Computer Vision (ICCV)*. 1150–1157.
- [25] Mathias Lux. 2011. Content Based Image Retrieval with LIRE. In *Proceedings of the 19th ACM International Conference on Multimedia (MM '11)*.
- [26] B. S. Manjunath, J. R. Ohm, V. V. Vasudevan, and A. Yamada. 2001. Color and texture descriptors. *IEEE Transactions on Circuits and Systems for Video Technology* 11, 6 (2001), 703–715.
- [27] M-E. Nilsback and A. Zisserman. 2006. A Visual Vocabulary for Flower Classification. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Vol. 2. 1447–1454.
- [28] D. Nistér and H. Stewénius. 2006. Scalable Recognition with a Vocabulary Tree. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR'2006)*, Vol. 2. 2161–2168.
- [29] Timo Ojala, Matti Pietikäinen, and Topi Mäenpää. 2002. Multiresolution Gray-Scale and Rotation Invariant Texture Classification with Local Binary Patterns. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 24, 7 (2002), 971–987.
- [30] Aude Oliva and Antonio Torralba. 2001. Modeling the Shape of the Scene: A Holistic Representation of the Spatial Envelope. *IJCV* 42, 3 (May 2001), 145–175.
- [31] Adam Paszke, Sam Gross, Soumith Chintala, Gregory Chanan, Edward Yang, Zachary DeVito, Zeming Lin, Alban Desmaison, Luca Antiga, and Adam Lerer. 2017. Automatic differentiation in PyTorch. In *NIPS-W*.
- [32] Mattis Paulin, Julien Mairal, Matthijs Douze, Zaid Harchaoui, Florent Perronnin, and Cordelia Schmid. 2017. Convolutional Patch Representations for Image Retrieval: An Unsupervised Approach. *Int. Journal of Computer Vision* (2017).
- [33] D. C. G. Pedronette and R. d. S. Torres. 2015. Unsupervised Effectiveness Estimation for Image Retrieval Using Reciprocal Rank Information. In *2015 28th SIBGRAPI Conference on Graphics, Patterns and Images*. 321–328.
- [34] Daniel Carlos Guimarães Pedronette and Ricardo da S. Torres. 2016. A correlation graph approach for unsupervised manifold learning in image retrieval tasks. *Neurocomputing* 208, Sup C (2016), 66 – 79.
- [35] Daniel Carlos Guimarães Pedronette, Filipe Marcel Fernandes Gonçalves, and Ivan Rizzo Guilherme. 2018. Unsupervised manifold learning through reciprocal kNN graph and Connected Components for image retrieval tasks. *Pattern Recognition* 75 (2018), 161 – 174. Distance Metric Learning for Pattern Recognition.
- [36] Daniel Carlos Guimarães Pedronette, Otávio A.B. Penatti, and Ricardo da S. Torres. 2014. Unsupervised manifold learning using Reciprocal kNN Graphs in image re-ranking and rank aggregation tasks. *Image and Vision Computing* 32, 2 (2014), 120 – 130.
- [37] Luca Piras and Giorgio Giacinto. 2017. Information fusion in content based image retrieval: A comprehensive overview. *Information Fusion* 37, Supplement C (2017), 50 – 60.
- [38] Konda Reddy Mopuri and R. Venkatesh Babu. 2015. Object Level Deep Feature Pooling for Compact Image Representation. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*.
- [39] Cees G. M. Snoek, Marcel Worring, and Arnold W. M. Smeulders. 2005. Early Versus Late Fusion in Semantic Video Analysis. In *Proceedings of the 13th Annual ACM International Conference on Multimedia (MULTIMEDIA '05)*. 399–402.
- [40] Renato O. Stehling, Mario A. Nascimento, and Alexandre X. Falcão. 2002. A compact and efficient image retrieval approach based on border/interior pixel classification. In *CIKM'2002*. 102–109.
- [41] Shaoyan Sun, Ying Li, Wengang Zhou, Qi Tian, and Houqiang Li. 2017. Local residual similarity for image re-ranking. *Information Sciences* 417, Sup. C (2017), 143 – 153.
- [42] Ricardo Da Silva Torres and Alexandre Xavier Falcão. 2006. Content-Based Image Retrieval: Theory and Applications. *Revista de Informática Teórica e Aplicada* 13 (2006), 161–185.
- [43] Lucas Pascotti Valem, Daniel Carlos Guimarães Pedronette, and Jurandy Almeida. 2018. Unsupervised similarity learning through Cartesian product of ranking references. *Pattern Recognition Letters* 114 (2018), 41 – 52. Data Representation and Representation Learning for Video Analysis.
- [44] K. E. A. van de Sande, T. Gevers, and C. G. M. Snoek. 2010. Evaluating Color Descriptors for Object and Scene Recognition. *PAMI* 32, 9 (2010), 1582–1596.
- [45] Sotiris A. Vassou, Nektarios Anagnostopoulos, Angelos Amanatiadis, Klitos Christodoulou, and Savvas A. Chatzichristofis. 2017. CoMo: A Compact Composite Moment-Based Descriptor for Image Retrieval. In *Proceedings of the 15th International Workshop on Content-Based Multimedia Indexing (CBMI '17)*. Article 30, 5 pages.
- [46] Bo Wang, Jiayan Jiang, Wei Wang, Zhi-Hua Zhou, and Zhuowen Tu. 2012. Unsupervised Metric Fusion by Cross Diffusion. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR'2012)*. 3013–3020.
- [47] Xiaoyu Wang, Ming Yang, T. Cour, Shenghuo Zhu, Kai Yu, and T.X. Han. 2011. Contextual weighting for vocabulary tree based image retrieval. In *IEEE International Conference on Computer Vision (ICCV'2011)*. 209–216.
- [48] J. Xiao, J. Hays, K. A. Ehinger, A. Oliva, and A. Torralba. 2010. SUN database: Large-scale scene recognition from abbey to zoo. In *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. 3485–3492.
- [49] Lingxi Xie, Richang Hong, Bo Zhang, and Qi Tian. 2015. Image Classification and Retrieval Are ONE. In *ACM Int. Conf. on Multimedia Retrieval (ICMR)*. 3–10.
- [50] Yi Yang, Heng Tao Shen, Zhigang Ma, Zi Huang, and Xiaofang Zhou. 2011. L2,1-norm Regularized Discriminative Feature Selection for Unsupervised Learning. In *Proceedings of the Twenty-Second International Joint Conference on Artificial Intelligence (IJCAI'11)*. 1589–1594.
- [51] K. Zagoris, S.A. Chatzichristofis, N. Papamarkos, and Y.S. Boutalis. 2010. Automatic Image Annotation and Retrieval Using the Joint Composite Descriptor. In *14th Panhellenic Conference on Informatics (PCI)*. 143–147.
- [52] Shaoting Zhang, Ming Yang, T. Cour, Kai Yu, and D.N. Metaxas. 2015. Query Specific Rank Fusion for Image Retrieval. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 37, 4 (April 2015), 803–815.
- [53] Zheng Zhao and Huan Liu. 2007. Spectral Feature Selection for Supervised and Unsupervised Learning. In *Proceedings of the 24th International Conference on Machine Learning (ICML '07)*. 1151–1157.
- [54] Liang Zheng, Shengjin Wang, Lu Tian, Fei He, Ziqiong Liu, and Qi Tian. 2015. Query-Adaptive Late Fusion for Image Search and Person Re-identification. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.
- [55] Liang Zheng, Shengjin Wang, and Qi Tian. 2014. Lp-Norm IDF for Scalable Image Retrieval. *IEEE Trans. on Image Processing* 23, 8 (Aug 2014), 3604–3617.