

## RESSALVA

Atendendo solicitação do autor, o texto completo desta dissertação será disponibilizado somente a partir de 01/03/2021.



UNIVERSIDADE ESTADUAL PAULISTA  
"JÚLIO DE MESQUITA FILHO"  
Câmpus de Rio Claro

Lucas Pascotti Valem

# Combinação Seletiva Não Supervisionada de Listas Ranqueadas Aplicada à Busca de Imagens pelo Conteúdo

Rio Claro - SP

2019

Lucas Pascotti Valem

**Combinação Seletiva Não Supervisionada de Listas  
Ranqueadas Aplicada à Busca de Imagens pelo  
Conteúdo**

Orientador: Prof. Dr. Daniel Carlos Guimarães Pedronette

Dissertação apresentada como parte dos requisitos para obtenção do título de Mestre em Ciência da Computação, junto ao Programa de Pós-Graduação em Ciência da Computação, do Instituto de Geociências e Ciências Exatas da Universidade Estadual Paulista “Júlio de Mesquita Filho”, Câmpus de Rio Claro.

Financiadora: FAPESP – Proc. 2017/02091-4.

Rio Claro - SP

2019

V151c

Valem, Lucas Pascotti

Combinação Seletiva Não Supervisionada de Listas Ranqueadas Aplicada à Busca de Imagens pelo Conteúdo / Lucas Pascotti Valem. -- Rio Claro, 2019  
106 f. : il., tabs.

Dissertação (mestrado) - Universidade Estadual Paulista (Unesp), Instituto de Geociências e Ciências Exatas, Rio Claro

Orientador: Daniel Carlos Guimarães Pedronette

1. Ciência da Computação. 2. Recuperação de Informações. 3. Recuperação de Imagens por Conteúdo. 4. Seleção e Combinação de Listas Ranqueadas. 5. Aprendizado Não Supervisionado. I. Título.

Sistema de geração automática de fichas catalográficas da Unesp. Biblioteca do Instituto de Geociências e Ciências Exatas, Rio Claro. Dados fornecidos pelo autor(a).

Essa ficha não pode ser modificada.

## ERRATA

VALEM, Lucas Pascotti. **Combinação Seletiva Não Supervisionada de Listas Ranqueadas Aplicada à Busca de Imagens pelo Conteúdo**. 2019. 106 f. Dissertação (Mestrado em Ciência da Computação) - Instituto de Biociências Letras e Ciências Exatas, Universidade Estadual Paulista (Unesp), São José do Rio Preto, 2019.

| Folha | Linha | Onde se lê                                                                                                                                            | Leia-se                                                                                                                                                                                                                      |
|-------|-------|-------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 1     | 14    | Financiadora: FAPESP – Proc. 2017/02091-4.                                                                                                            | Financiadores: FAPESP e CAPES – Proc. FAPESP nº 2017/02091-4.                                                                                                                                                                |
| 3     | 13    | Financiadora: FAPESP – Proc. 2017/02091-4.                                                                                                            | Financiadores: FAPESP e CAPES – Proc. FAPESP nº 2017/02091-4.                                                                                                                                                                |
| 4     | 7     | Agradeço à FAPESP pela concessão da bolsa de pesquisa, sob o processo nº 2017/02091-4, Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP). | Agradeço à FAPESP – Fundação de Amparo à Pesquisa do Estado de São Paulo e a CAPES – Coordenação de Aperfeiçoamento de Pessoal de Nível Superior pela concessão da bolsa de pesquisa, sob o processo FAPESP nº 2017/02091-4. |

IBILCE/UNESP

SEÇÃO DE PÓS-GRADUAÇÃO

Aprovado pelo Conselho do Programa em  
reunião de 1/1 Encaminha-se a  
STPG para providências

São José do Rio Preto, 19 de 07 de 2019

x 

Prof. Dr. Adriano Mauro Cansia  
Vice - Coordenador do Programa de  
Pós-Graduação em Ciências da  
Computação

Lucas Pascotti Valem

# **Combinação Seletiva Não Supervisionada de Listas Ranqueadas Aplicada à Busca de Imagens pelo Conteúdo**

Dissertação apresentada como parte dos requisitos para obtenção do título de Mestre em Ciência da Computação, junto ao Programa de Pós-Graduação em Ciência da Computação, do Instituto de Geociências e Ciências Exatas da Universidade Estadual Paulista “Júlio de Mesquita Filho”, Câmpus de Rio Claro.

Financiadora: FAPESP – Proc. 2017/02091-4.

## **Comissão Examinadora**

- Prof. Dr. Daniel Carlos Guimarães Pedronette (Orientador)  
Departamento de Estatística, Matemática Aplicada e Computação  
Universidade Estadual Paulista - UNESP
- Prof. Dr. Fabio Augusto Faria  
Instituto de Ciência e Tecnologia (ICT)  
Universidade Federal de São Paulo (UNIFESP)
- Prof. Dr. Fabricio Aparecido Breve  
Departamento de Estatística, Matemática Aplicada e Computação  
Universidade Estadual Paulista - UNESP

Rio Claro - SP

01 de Março de 2019

# Agradecimentos

Agradeço primeiramente a Deus pela vida, toda saúde e disposição.

Aos meus pais e familiares, pelo amor e apoio incondicional.

Ao meu orientador, pela orientação, apoio e confiança.

Agradeço à universidade e a todo o corpo docente do programa de pós-graduação.

Aos amigos e colegas, pelo incentivo e apoio.

Agradeço à FAPESP pela concessão da bolsa de pesquisa, sob o processo nº 2017/02091-4, Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP).

# Resumo

Em virtude do grande crescimento das coleções multimídia e da evolução das tecnologias no decorrer dos últimos anos, buscar por imagens considerando o conteúdo visual das mesmas se tornou uma tarefa de extrema importância. Os sistemas de Recuperação de Imagens pelo Conteúdo (CBIR) são a principal solução dentro desse cenário e representam um campo de pesquisa cada vez mais ativo. Atualmente, esses sistemas empregam diferentes abordagens para execução de suas tarefas, dentre elas estão diferentes descritores e métodos de aprendizado. Contudo, mensurar a similaridade de maneira eficaz ainda é um desafio, principalmente devido ao conhecido problema da lacuna semântica. Como diferentes listas ranqueadas possuem resultados distintos e muitas vezes complementares, uma estratégia que tem ganhado significativa atenção é a de combinar essas listas. Este trabalho propõe três diferentes métodos completamente não supervisionados para seleção e combinação de listas ranqueadas. As abordagens foram avaliadas em cinco coleções de imagens e diversos descritores. Foram obtidos resultados comparáveis ou superiores ao estado-da-arte na maioria dos cenários analisados.

**Palavras-chave:** Recuperação de Informações, Recuperação de Imagens por Conteúdo, Seleção e Combinação de Características, Seleção e Combinação de Listas Ranqueadas, Reconhecimento de Padrões, Aprendizado Não Supervisionado, Rerankeamento.

# Abstract

Due to the evolution of technologies to store and share images, finding effective methods to index and retrieve this type of information is indispensable. The CBIR (Content-Based Image Retrieval) systems are the main solution for image retrieval tasks. These systems consist in the use of different descriptors, supervised learning methods, and more recently unsupervised learning methods. However, accurately estimating the similarity between images is still a challenging task, mainly due to the well-known semantic gap problem. As different ranked lists present different effectiveness results, an interesting methodology is to combine these ranked lists. This work proposes three different unsupervised methods for selecting and combining ranked lists. The approaches were evaluated in five different image collections and several descriptors. The methods achieved results comparable or superior to the state-of-the-art in most of the evaluated scenarios.

**Keywords:** Information Retrieval, Content-Based Image Retrieval, Feature Selection and Fusion, Pattern Recognition, Unsupervised Learning, Rank-aggregation, Re-ranking.

# Lista de Figuras

|                                                                                                                                                                                       |    |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| Figura 1 – Arquitetura típica de um sistema CBIR. . . . .                                                                                                                             | 18 |
| Figura 2 – Ilustração do processo de extração de características e da lacuna semântica. . . . .                                                                                       | 20 |
| Figura 3 – Esquema de um sistema CBIR considerando a combinação de caracte-<br>rísticas. . . . .                                                                                      | 27 |
| Figura 4 – Exemplo de etapas com presença de <i>early fusion</i> . . . . .                                                                                                            | 28 |
| Figura 5 – Exemplo de etapas com presença de <i>late fusion</i> . . . . .                                                                                                             | 28 |
| Figura 6 – Exemplo de gene na programação genética. . . . .                                                                                                                           | 37 |
| Figura 7 – Ilustração da coocorrência de vizinhos recíprocos no grafo de vizinhança. . . . .                                                                                          | 43 |
| Figura 8 – Etapas do algoritmo CPRR. . . . .                                                                                                                                          | 46 |
| Figura 9 – Visão geral do método USRF. . . . .                                                                                                                                        | 47 |
| Figura 10 – Ilustração das etapas que compõem o USRF. . . . .                                                                                                                         | 49 |
| Figura 11 – Ilustração do algoritmo de seleção proposto. . . . .                                                                                                                      | 51 |
| Figura 12 – Impacto dos parâmetros $k$ e $\beta$ na média aritmética ponderada consi-<br>derando o MAP dos cinco primeiros resultados selecionados para cada<br>coleção. . . . .      | 54 |
| Figura 13 – Avaliação do parâmetro $\beta$ considerando cenários aleatórios com diferen-<br>tes números de descritores. . . . .                                                       | 56 |
| Figura 14 – Avaliação do parâmetro $L_R$ nos resultados de seleção. . . . .                                                                                                           | 57 |
| Figura 15 – Distribuição dos pares considerando a métrica de seleção comparada ao<br>MAP nos cenários personalizados. . . . .                                                         | 58 |
| Figura 16 – Avaliação da seleção proposta (USRF) para pares de ranqueadores nos<br>cenários personalizados. . . . .                                                                   | 59 |
| Figura 17 – Avaliação da seleção proposta (USRF) para pares de ranqueadores nos<br>cenários completos. . . . .                                                                        | 60 |
| Figura 18 – Avaliação da seleção proposta (USRF) para pares de ranqueadores em<br>diferentes cenários. . . . .                                                                        | 61 |
| Figura 19 – Análise do tamanho das combinações em relação aos resultados de<br>seleção para cada uma das coleções considerando todos os ranqueadores<br>(cenários completos). . . . . | 63 |
| Figura 20 – Principais etapas que compõem o método RGSF. . . . .                                                                                                                      | 64 |
| Figura 21 – Ilustração das iterações realizadas para a seleção dos ranqueadores pelas<br>componentes conexas. . . . .                                                                 | 66 |
| Figura 22 – Impacto do parâmetro $k$ no MAP da combinação selecionada para as<br>coleções com maior número de imagens por classe. . . . .                                             | 68 |
| Figura 23 – Impacto do parâmetro $k$ no MAP da combinação selecionada para as<br>coleções com menor número de imagens por classe. . . . .                                             | 69 |

|                                                                                                                               |    |
|-------------------------------------------------------------------------------------------------------------------------------|----|
| Figura 24 – Avaliação do método RGSF nos cenários personalizados. . . . .                                                     | 70 |
| Figura 25 – Visão geral do <i>framework</i> UGAF-RSF. . . . .                                                                 | 72 |
| Figura 26 – Exemplo de cromossomo binário. . . . .                                                                            | 73 |
| Figura 27 – Comparação entre as diferentes funções de aptidão. . . . .                                                        | 75 |
| Figura 28 – Análise do tamanho da população. . . . .                                                                          | 75 |
| Figura 29 – Avaliação da taxa de cruzamento. . . . .                                                                          | 76 |
| Figura 30 – Avaliação da taxa de mutação. . . . .                                                                             | 76 |
| Figura 31 – Comparação das estratégias evolutivas. . . . .                                                                    | 76 |
| Figura 32 – Análise da taxa de elitismo máxima. . . . .                                                                       | 76 |
| Figura 33 – Análise do tamanho da vizinhança $k$ . . . . .                                                                    | 77 |
| Figura 34 – Dois exemplos de resultados visuais que ilustram o impacto da seleção<br>e combinação na coleção UKBench. . . . . | 83 |

# Lista de Tabelas

|           |                                                                                                                                             |    |
|-----------|---------------------------------------------------------------------------------------------------------------------------------------------|----|
| Tabela 1  | – Exemplos de métricas de distância. . . . .                                                                                                | 19 |
| Tabela 2  | – Exemplos de operações aritméticas simples para obtenção de $F'$ . . . . .                                                                 | 29 |
| Tabela 3  | – Algoritmos de combinação da família Comb*. . . . .                                                                                        | 32 |
| Tabela 4  | – Coleções de imagens utilizadas na avaliação experimental. . . . .                                                                         | 40 |
| Tabela 5  | – Descritores de imagens utilizados na avaliação experimental. . . . .                                                                      | 40 |
| Tabela 6  | – Descritores considerados nos cenários personalizados. . . . .                                                                             | 41 |
| Tabela 7  | – Métodos considerados como <i>baselines</i> para <i>late fusion</i> . . . . .                                                              | 42 |
| Tabela 8  | – Análise do tamanho dos conjuntos de combinações. . . . .                                                                                  | 48 |
| Tabela 9  | – Média ponderada do MAP para os cinco primeiros pares selecionados considerando diferentes combinações de métricas e $\beta = 1$ . . . . . | 55 |
| Tabela 10 | – MAP das três primeiras combinações selecionadas considerando diferentes tamanhos nos cenários personalizados e completos. . . . .         | 62 |
| Tabela 11 | – Especificação das combinações selecionadas para cada uma das coleções nos cenários personalizados e completos. . . . .                    | 63 |
| Tabela 12 | – MAP da combinação selecionada para as coleções em diferentes cenários. . . . .                                                            | 70 |
| Tabela 13 | – Combinações mais frequentemente selecionadas e a média do MAP para cada uma das coleções. . . . .                                         | 77 |
| Tabela 14 | – Comparação com métodos de <i>late fusion</i> nos cenários personalizados. . . . .                                                         | 79 |
| Tabela 15 | – Comparação com métodos de <i>early fusion</i> . . . . .                                                                                   | 80 |
| Tabela 16 | – Comparação com diversos métodos de pós-processamento na coleção MPEG-7 ( <i>Bull's eye score</i> - <i>Recall@40</i> ). . . . .            | 81 |
| Tabela 17 | – Comparação com o estado-da-arte na coleção Holidays (MAP). . . . .                                                                        | 82 |
| Tabela 18 | – Comparação com o estado-da-arte na coleção UKBench ( <i>N-S score</i> ). . . . .                                                          | 83 |
| Tabela 19 | – Métodos de aprendizado não supervisionado utilizados para o experimento. . . . .                                                          | 84 |
| Tabela 20 | – Resultados de MAP para combinação de métodos de aprendizado não supervisionado. . . . .                                                   | 85 |
| Tabela 21 | – Tempo de execução dos métodos propostos nos cenários completos. . . . .                                                                   | 85 |
| Tabela 22 | – Análise de eficácia e eficiência dos métodos propostos. . . . .                                                                           | 86 |

# Lista de Abreviaturas

|       |                                                                                                                       |
|-------|-----------------------------------------------------------------------------------------------------------------------|
| ACC   | <i>Color Autocorrelogram Descriptor</i>                                                                               |
| AIR   | <i>Articulation-Invariant Representation Descriptor</i>                                                               |
| AP    | <i>Average Precision</i>                                                                                              |
| ASC   | <i>Aspect Shape Context Descriptor</i>                                                                                |
| BAS   | <i>Beam Angle Statistics Descriptor</i>                                                                               |
| BIC   | <i>Border/Interior Pixel Classification Descriptor</i>                                                                |
| BoW   | <i>Bag of Words</i>                                                                                                   |
| BoVW  | <i>Bag of Visual Words</i>                                                                                            |
| BRIEF | <i>Binary Robust Independent Elementary Features</i>                                                                  |
| CBIR  | <i>Content-Based Image Retrieval</i>                                                                                  |
| CEDD  | <i>Color and Edge Directivity Descriptor</i>                                                                          |
| CLD   | <i>Color Layout Descriptor</i>                                                                                        |
| COMO  | <i>Compact Composite Moment-Based Descriptor</i>                                                                      |
| CPRR  | <i>Cartesian Product of Ranking References</i>                                                                        |
| CCA   | <i>Canonical Correlation Analysis</i>                                                                                 |
| CNN   | <i>Convolutional Neural Networks</i>                                                                                  |
| COMO  | <i>Compact Composite Moment-Based Descriptor</i>                                                                      |
| DCA   | <i>Discriminant Correlation Analysis</i>                                                                              |
| DCCA  | <i>Deep Canonical Correlation Analysis</i>                                                                            |
| DCNN  | <i>Deep Convolutional Neural Networks</i>                                                                             |
| EHD   | <i>Edge Histogram Descriptor</i>                                                                                      |
| FCTH  | <i>Fuzzy Color and Texture Histogram</i>                                                                              |
| FOH   | <i>Fuzzy Opponent Histogram</i>                                                                                       |
| GIST  | <i>Global Image Descriptor for low-dimensional features</i>                                                           |
| HoG   | <i>Histogram of Oriented Gradients</i>                                                                                |
| HSV   | Sistema de cores formado pelas componentes <i>hue</i> (matiz), <i>saturation</i> (saturação) e <i>value</i> (brilho). |
| HSVD3 | <i>Three-dimensional HSV Histogram</i>                                                                                |
| IDSC  | <i>Inner Distance Shape Context Descriptor</i>                                                                        |
| JCD   | <i>Joint Composite Descriptor</i>                                                                                     |
| LASSO | <i>Least Absolute Shrinkage and Selection Operator</i>                                                                |
| LBP   | <i>Local Binary Patterns</i>                                                                                          |
| LDA   | <i>Linear Discriminant Analysis</i>                                                                                   |
| MAP   | <i>Mean Average Precision</i>                                                                                         |
| MCFS  | <i>Muti-cluster Feature Selection</i>                                                                                 |
| MKL   | <i>Multiple Kernel Learning</i>                                                                                       |
| MRA   | <i>Median Rank Aggregation</i>                                                                                        |
| NDFS  | <i>Nonnegative Discriminative Feature Selection</i>                                                                   |

|          |                                                                               |
|----------|-------------------------------------------------------------------------------|
| OPF      | <i>Optimum-Path Forest</i>                                                    |
| ORB      | <i>Oriented Fast and Rotated BRIEF</i>                                        |
| PCA      | <i>Principal Component Analysis</i>                                           |
| PCC      | <i>Particle Competition and Cooperation</i>                                   |
| PG       | Programação Genética                                                          |
| PHOG     | <i>Pyramidal Histogram of oriented gradients</i>                              |
| RBO      | <i>Rank-biased Overlap</i>                                                    |
| RGB      | <i>Red, Green, Blue</i>                                                       |
| RI       | Recuperação de Informações                                                    |
| SCD      | <i>Scalable Color Descriptor</i>                                              |
| SCH      | <i>Simple Color Histogram</i>                                                 |
| SIFT     | <i>Scale-Invariant Feature Transform</i>                                      |
| SPACC    | <i>Spatial Pyramid Color Autocorrelogram Descriptor</i>                       |
| SPCEDD   | <i>Spatial Pyramid CEDD</i>                                                   |
| SPEC     | <i>Spectral Regression</i>                                                    |
| SPFCTH   | <i>Spatial Pyramid FCTH</i>                                                   |
| SPJCD    | <i>Spatial Pyramid JCD</i>                                                    |
| SPLBP    | <i>Spatial Pyramid LBP</i>                                                    |
| SS       | <i>Segment Saliences Descriptor</i>                                           |
| SURF     | <i>Speeded-Up Robust Features</i>                                             |
| SVM      | <i>Support Vector Machines</i>                                                |
| UDFS     | <i>Unsupervised Discriminative Feature Selection</i>                          |
| UDLF     | <i>Unsupervised Distance Learning Framework</i>                               |
| USRF     | <i>Unsupervised Selective Rank Fusion</i>                                     |
| RGSF     | <i>Ranker Graph for Selection and Fusion</i>                                  |
| UGAF-RSF | <i>Unsupervised Genetic Algorithm Framework for Rank Selection and Fusion</i> |
| VOC      | <i>Vocabulary Tree Descriptor</i>                                             |

# Lista de Símbolos

|                              |                                                                                            |
|------------------------------|--------------------------------------------------------------------------------------------|
| $\mathcal{C}$                | Coleção de imagens.                                                                        |
| $\mathcal{N}$                | Tamanho da coleção de imagens.                                                             |
| $\mathcal{A}$                | Matriz quadrada de distância ou similaridade.                                              |
| $\mathcal{T}$                | Conjunto das listas ranqueadas de todas as imagens de uma coleção.                         |
| $L$                          | Tamanho das listas ranqueadas das imagens de consulta.                                     |
| $\tau_i$                     | Lista ranqueada da imagem de índice $i$ .                                                  |
| $\tau_q(i)$                  | Posição da imagem $i$ na lista ranqueada da imagem de consulta $q$ .                       |
| $f_s$                        | Função de seleção.                                                                         |
| $R_i$                        | Ranqueador de índice $i$ .                                                                 |
| $\tau_{i,q}$                 | Lista ranqueada da imagem de índice $q$ calculada pelo ranqueador $i$ .                    |
| $\mathcal{T}_i$              | Conjunto de listas ranqueadas produzidas pelo ranqueador $R_i$ .                           |
| $\mathfrak{R}$               | Conjunto de ranqueadores fornecidos como entrada do método.                                |
| $m$                          | Tamanho do conjunto $\mathfrak{R}$ .                                                       |
| $\gamma$                     | Função que calcula a estimativa de eficácia.                                               |
| $\lambda$                    | Função que calcula a correlação.                                                           |
| $th_\gamma$                  | Limiar ( <i>threshold</i> ) para estimativa de eficácia.                                   |
| $th_\lambda$                 | Limiar ( <i>threshold</i> ) para correlação.                                               |
| $n$                          | Número de ranqueadores que compõem uma combinação.                                         |
| $\mathfrak{X}_n$             | Combinação de $n$ ranqueadores.                                                            |
| $\mathfrak{X}_{cc_i}$        | Combinação fornecida pela componente conexa de índice $i$ .                                |
| $\mathfrak{X}'$              | Combinação de ranqueadores após a etapa de agregação.                                      |
| $\mathfrak{X}_n^*$           | Combinação de tamanho $n$ selecionada pelo método.                                         |
| $\mathfrak{X}^*$             | Combinação selecionada pelo método.                                                        |
| $C_n$                        | Conjunto de combinações em que cada combinação é de tamanho $n$ .                          |
| $\tau_n^R$                   | Lista ordenada das combinações de tamanho $n$ . Também referenciada como lista de seleção. |
| $\tau_n^R(\mathfrak{X}_n^i)$ | Posição da combinação $\mathfrak{X}_n^i$ na lista de seleção $\tau_n^R$ .                  |
| $L_R$                        | Tamanho das listas de seleção $\tau_n^R$ .                                                 |
| $\mathcal{S}$                | Conjunto de seleção de todas as possíveis combinações de ranqueadores.                     |
| $\mathcal{S}_p$              | Conjunto de seleção de pares de ranqueadores.                                              |
| $S$                          | Lista de combinações selecionadas pelo RGSF.                                               |
| $S'$                         | Lista de combinações selecionadas pelo RGSF após a etapa de agregação.                     |
| $w$                          | Métrica de seleção de combinações de ranqueadores.                                         |
| $w_p$                        | Métrica de seleção de pares de ranqueadores.                                               |
| $\beta$                      | Peso ou relevância da correlação na métrica de seleção.                                    |
| $k$                          | Tamanho do conjunto de vizinhança.                                                         |
| $T$                          | Número de iterações.                                                                       |

# Sumário

|            |                                                   |           |
|------------|---------------------------------------------------|-----------|
| <b>1</b>   | <b>INTRODUÇÃO</b>                                 | <b>13</b> |
| <b>2</b>   | <b>ASPECTOS TEÓRICOS</b>                          | <b>17</b> |
| <b>2.1</b> | <b>Recuperação de Informações</b>                 | <b>17</b> |
| 2.1.1      | Sistemas de Recuperação de Imagens pelo Conteúdo  | 17        |
| 2.1.2      | Descritores de Imagens                            | 19        |
| 2.1.3      | Definição Formal                                  | 22        |
| 2.1.4      | Métricas de Eficácia                              | 23        |
| <b>2.2</b> | <b>Aprendizado de Máquina</b>                     | <b>24</b> |
| 2.2.1      | Aprendizado Supervisionado                        | 25        |
| 2.2.2      | Aprendizado Não Supervisionado                    | 25        |
| 2.2.3      | Aprendizado Semi-Supervisionado                   | 26        |
| <b>3</b>   | <b>TRABALHOS RELACIONADOS</b>                     | <b>27</b> |
| <b>3.1</b> | <b>Categorização das Abordagens de Combinação</b> | <b>27</b> |
| 3.1.1      | <i>Early Fusion</i>                               | 27        |
| 3.1.2      | <i>Late Fusion</i>                                | 28        |
| <b>3.2</b> | <b>Abordagens de Combinação</b>                   | <b>29</b> |
| 3.2.1      | Combinação Aritmética Não Treinável               | 29        |
| 3.2.2      | Combinação Conservativa                           | 29        |
| 3.2.3      | Agregação de Listas Ranqueadas                    | 31        |
| 3.2.4      | Combinação de Núcleos                             | 33        |
| 3.2.5      | Abordagens Estatísticas                           | 35        |
| 3.2.6      | Combinação baseada em Programação Genética        | 36        |
| <b>3.3</b> | <b>Abordagens Recentes</b>                        | <b>37</b> |
| <b>4</b>   | <b>PROTOCOLO EXPERIMENTAL</b>                     | <b>39</b> |
| <b>4.1</b> | <b>Coleções e Descritores de Imagens</b>          | <b>39</b> |
| <b>4.2</b> | <b><i>Baselines</i></b>                           | <b>41</b> |
| <b>4.3</b> | <b>Métricas para Seleção</b>                      | <b>42</b> |
| 4.3.1      | Estimativas de Eficácia                           | 42        |
| 4.3.2      | Métricas de Correlação                            | 44        |
| <b>4.4</b> | <b>Método de Agregação</b>                        | <b>46</b> |
| <b>5</b>   | <b>SELEÇÃO POR MÉTRICAS PAR-A-PAR</b>             | <b>47</b> |
| <b>5.1</b> | <b>Descrição do Método</b>                        | <b>48</b> |

|            |                                                                          |           |
|------------|--------------------------------------------------------------------------|-----------|
| 5.1.1      | Estratégia de Seleção . . . . .                                          | 48        |
| 5.1.1.1    | Seleção de Pares de Ranqueadores . . . . .                               | 49        |
| 5.1.1.2    | Seleção de Conjuntos de Ranqueadores . . . . .                           | 51        |
| <b>5.2</b> | <b>Avaliação Experimental . . . . .</b>                                  | <b>53</b> |
| 5.2.1      | Análise dos Parâmetros . . . . .                                         | 53        |
| 5.2.2      | Resultados de Seleção . . . . .                                          | 57        |
| 5.2.2.1    | Seleção de Pares . . . . .                                               | 57        |
| 5.2.2.2    | Seleção de Conjuntos de Ranqueadores . . . . .                           | 61        |
| <b>6</b>   | <b>SELEÇÃO POR GRAFOS E COMPONENTES CONEXAS . . . . .</b>                | <b>64</b> |
| <b>6.1</b> | <b>Combinação de Ranqueadores pelas Componentes Conexas . . . . .</b>    | <b>65</b> |
| <b>6.2</b> | <b>Avaliação Experimental . . . . .</b>                                  | <b>68</b> |
| 6.2.1      | Análise dos Parâmetros . . . . .                                         | 68        |
| 6.2.2      | Resultados de Seleção . . . . .                                          | 69        |
| <b>7</b>   | <b>SELEÇÃO POR ALGORITMOS GENÉTICOS . . . . .</b>                        | <b>72</b> |
| <b>7.1</b> | <b>Descrição do Método . . . . .</b>                                     | <b>72</b> |
| 7.1.1      | Representação dos Cromossomos . . . . .                                  | 73        |
| 7.1.2      | Agregação Não Supervisionada . . . . .                                   | 73        |
| 7.1.3      | Função de Aptidão Não Supervisionada . . . . .                           | 73        |
| 7.1.4      | Seleção de Cromossomos . . . . .                                         | 73        |
| 7.1.5      | Operações de Cruzamento . . . . .                                        | 74        |
| 7.1.6      | Operações de Mutação . . . . .                                           | 74        |
| <b>7.2</b> | <b>Avaliação Experimental . . . . .</b>                                  | <b>74</b> |
| 7.2.1      | Análise dos Parâmetros . . . . .                                         | 74        |
| <b>8</b>   | <b>DISCUSSÕES, COMPARAÇÕES E ANÁLISES . . . . .</b>                      | <b>79</b> |
| <b>8.1</b> | <b>Avaliação Comparativa . . . . .</b>                                   | <b>79</b> |
| 8.1.1      | Comparação com <i>Baselines</i> de <i>Late Fusion</i> . . . . .          | 79        |
| 8.1.2      | Comparação com <i>Baselines</i> de <i>Early Fusion</i> . . . . .         | 80        |
| 8.1.3      | Comparação com Estado-da-arte . . . . .                                  | 80        |
| <b>8.2</b> | <b>Análise Visual . . . . .</b>                                          | <b>82</b> |
| <b>8.3</b> | <b>Combinação de Métodos de Aprendizado Não Supervisionado . . . . .</b> | <b>83</b> |
| <b>8.4</b> | <b>Discussões . . . . .</b>                                              | <b>85</b> |
| <b>8.5</b> | <b>Contribuições e Publicações . . . . .</b>                             | <b>86</b> |
| <b>9</b>   | <b>CONCLUSÕES . . . . .</b>                                              | <b>88</b> |
| <b>9.1</b> | <b>Trabalhos Futuros . . . . .</b>                                       | <b>88</b> |
|            | <b>REFERÊNCIAS . . . . .</b>                                             | <b>90</b> |

# 1 Introdução

O grande crescimento na quantidade de dispositivos móveis disponíveis aos usuários e suas capacidades de não somente consumir como também produzir conteúdo geram um aumento crescente na quantidade de imagens e das coleções multimídia em geral [1]. Os sistemas de recuperação de informações são baseados na ideia de recuperar as informações mais similares a partir de um objeto de entrada, fornecendo uma lista ranqueada (*ranked list*) com os elementos mais relevantes no topo da lista. A recuperação de informações se iniciou considerando busca textual e tem evoluído de maneira considerável nos últimos anos, fornecendo hoje resultados já bastante satisfatórios em grande parte dos cenários existentes. Apesar da eficácia e desempenho dos sistemas de recuperação textual terem atingido um bom grau de maturidade, ainda há desafios na área, especialmente em aplicações de recuperação de imagens e conteúdos multimídia, devido a riqueza de conteúdo e subjetividade na interpretação dos mesmos [2].

Devido ao surgimento de novas tecnologias, a busca de imagens se tornou uma necessidade imprescindível. Inicialmente, a metodologia para recuperação de conteúdo multimídia consistia no uso de metadados (*metadata*) associados aos arquivos como por exemplo, geolocalização, tempo (*timestamp*), palavras-chave, marcações, ou até mesmo descrições. A partir desses metadados, uma busca textual era aplicada [3]. Catalogar manualmente uma grande coleção de imagens, além de ser um trabalho extremamente custoso, muitas vezes não é eficaz devido a grande quantidade de informações e até mesmo ao grau de subjetividade do conteúdo envolvido. Além disso, é importante considerar que os metadados associados a uma imagem muitas vezes não descrevem de maneira completa o que está sendo representado, o que pode ocasionar em ambiguidades.

Dessa forma, a fim de explorar o conteúdo visual das coleções de imagens e amenizar o impacto dos problemas citados, os Sistemas de Recuperação de Imagem pelo Conteúdo (CBIR) são propostos e possuem as mais diferentes aplicações (diagnóstico de doenças [4], reconhecimento facial [5], sensoriamento remoto [6], identificação de objetos [7], entre outras). Originalmente, esses sistemas são unicamente baseados no uso de descritores. Um descritor tem como objetivo definir uma função, capaz de extrair um vetor numérico a partir de uma imagem de entrada. Tal vetor atua como uma forma de representação ou identificação (*“fingerprint”*) para aquela imagem. Os descritores visuais consideram características visuais como cor, textura, contorno e outros. Esse tem sido um campo de pesquisa muito ativo nos últimos 20 anos [8, 9, 1].

Uma das maiores limitações do uso de descritores se deve a lacuna semântica (*“semantic gap”*) [10] presente entre as características de baixo e alto nível. Com a

representação do vetor de características, conceitos e significados originalmente contidos na imagem podem ser perdidos ou apenas parcialmente caracterizados. Atualmente, há sistemas de recuperação de imagens eficazes para finalidades específicas. No entanto, um sistema de busca de imagens para propósito geral que seja maduro constitui uma tarefa desafiadora, mesmo considerando as técnicas de estado-da-arte [1]. Na maioria dos casos, pode-se dizer que quanto mais específico o domínio do sistema, mais fácil de ser desenvolvido [8]. Dessa forma, diferentes métodos de aprendizado têm sido propostos com a finalidade de minimizar o problema da lacuna semântica.

Técnicas supervisionadas, como a realimentação de relevância (*relevance feedback*) [11], surgiram como uma forma de aprimorar os resultados. Tais técnicas são baseadas na interação com o usuário, o qual submete ao sistema uma imagem de interesse e o sistema retorna as primeiras  $k$  imagens consideradas mais relevantes. A partir disso, o usuário informa quais das imagens retornadas são relevantes e quais não são. Dessa forma, o sistema CBIR pode atribuir (“aprender”) novos pesos de similaridade ou distância entre as imagens para uma próxima busca.

Recentemente, uma nova categoria de descritores foi proposta e vem sendo continuamente utilizada e desenvolvida através do uso de *Deep Convolutional Neural Networks* (CNN) [12]. São utilizadas diversas camadas de informação interconectadas para tarefas de reconhecimento de padrões e aprendizado de características (*feature learning tasks*) [13]. O aspecto mais interessante desse paradigma de *deep learning* é que as características não precisam necessariamente ser extraídas a partir dos dados originais (*raw data*) de antemão, mas os dados originais são processados pela rede produzindo uma representação interna dos dados que pode ser utilizada como uma representação de características. Um dos trabalhos pioneiros em CNN para recuperação e classificação de imagens [12], mostra que as camadas superiores da rede podem ser utilizadas como descritores visuais em diferentes cenários. Alguns autores [14] revelam, inclusive, que dados de treinamento para uma determinada coleção de imagens apresentam resultados consideravelmente satisfatórios em outras bases, mesmo sendo relativamente diferentes. Apesar disso, os resultados podem ser geralmente melhorados retreinando para a base em questão. No entanto, apesar de promissores, os resultados nem sempre atingem o estado-da-arte quando comparados a outras estratégias tradicionais, como pode ser visto em [15].

Além da diversidade de métodos para extração de características, outro fenômeno notável é o crescimento da quantidade e tamanho das coleções de imagens, geralmente em cenários em que dados rotulados são inexistentes. Dessa forma, com grandes coleções de imagens, nem sempre uma quantidade considerável de dados rotulados está disponível. Sendo assim, recentemente, métodos de aprendizado não supervisionado vêm sendo propostos e têm se mostrado bastante promissores, principalmente no que se refere aos ganhos de eficácia obtidos. Esses métodos não necessitam de nenhuma intervenção do

usuário, dados rotulados ou treinamento. A estratégia desses métodos reside em substituir medidas de distância par a par por medidas de afinidade globais, realizando uma análise da informação contextual entre as imagens [16]. Na literatura, são encontrados métodos de aprendizado não supervisionado baseados em diferentes abordagens (processos de difusão [17], grafos [18], análises das listas ranqueadas [19], etc.). No decorrer dos últimos anos, os métodos de aprendizado não supervisionado baseados na análise das listas ranqueadas ganharam um grande destaque. Isso se deve ao fato desses métodos serem eficientes e de baixo custo computacional [20]. A fim de melhorar os resultados de eficácia desses métodos, foram propostas abordagens com o objetivo de realizar a agregação de resultados complementares providos pelos descritores [21].

Embora os métodos de aprendizado não supervisionado operem de forma independente à intervenção do usuário, a seleção dos descritores utilizados consiste em uma tarefa desafiadora. Geralmente, costuma ser difícil selecionar o melhor tipo de descritor (cor, forma, textura, etc.) para um determinado cenário e as características extraídas por ele não definem ou representam a imagem como um todo. Se compararmos, por exemplo, a imagem de um limão com a de uma laranja [22] considerando apenas o contorno, diremos que as imagens são muito semelhantes. No entanto, se fizermos o mesmo utilizando um descritor de cor, teremos que as imagens são consideravelmente diferentes. Considerando esses desafios, pesquisas vêm sendo realizadas visando a combinação de diferentes descritores [23]. Em alguns cenários, pode ser necessário ainda considerar graus de relevância para diferentes informações, de tal forma que possam ser atribuídos pesos diferentes para cada tipo de característica.

Como diferentes métodos apresentam resultados distintos e, muitas vezes complementares, para uma mesma coleção de imagens, uma estratégia de grande importância consiste em realizar a combinação desses métodos. No entanto, a escolha da melhor combinação depende muito da coleção de imagens e do descritor a ser utilizado. Trabalhos recentes apresentam propostas de seleção de combinação de métodos usando técnicas supervisionadas e algoritmos genéticos [24].

Este trabalho propõe três diferentes métodos para seleção e combinação de listas ranqueadas completamente não supervisionados. Cada um deles consiste de diferentes estratégias, cuja aplicação pode ser mais ou menos indicada a depender do cenário de uso.

O USRF (*Unsupervised Selective Rank Fusion*) faz a seleção através da atribuição de valores para cada um dos possíveis pares de ranqueadores (nesse caso descritores) pela utilização de métricas de correlação e estimativas de eficácia. O método aplica a seleção para combinações de qualquer tamanho através da união dos pares mais relevantes, o que é imprescindível para viabilizar a seleção para grandes conjuntos de ranqueadores.

Enquanto isso, o RGSF (*Ranker Graph for Selection and Fusion*) explora um grafo de ranqueadores em que valores de estimativa de eficácia são atribuídos aos vértices e

## 9 Conclusões

A dissertação apresentou como contribuição três novos métodos para seleção e combinação de ranqueadores completamente não supervisionados que consideram técnicas distintas. Dentre os principais conceitos empregados, fez-se o uso de métricas de estimativa de eficácia, métricas de correlação, grafos, abordagens de reranqueamento e também algoritmos genéticos. Todos os métodos propostos foram avaliados em diferentes cenários demonstrando resultados muito promissores, os quais são próximos ou superiores ao estado-da-arte.

Diversos desafios foram encontrados durante a execução deste trabalho. Apesar do grande número de métodos desenvolvidos na área de aprendizado de máquina e CBIR em geral, métodos robustos capazes de selecionar um grande conjunto de ranqueadores de maneira completamente não supervisionada são, pelo que se tem conhecimento, escassos na literatura. Como a maior parte dos métodos utilizam de conjuntos pequenos e pré-selecionados de descritores, encontrar *baselines* adequados foi uma das tarefas mais difíceis durante a pesquisa.

A análise experimental foi conduzida em diferentes contextos, considerando cinco coleções de imagens e uma grande diversidade de descritores (mais de trinta) de diversas categorias e tipos. O que demonstrou que nossos métodos não são só capazes de explorar informações complementares entre diversos ranqueadores, mas também de filtrar os ranqueadores menos relevantes tanto em pequenos como em grandes conjuntos.

Os métodos propostos foram comparados considerando o mesmo conjunto de descritores tanto em *baselines* de *early* quanto *late fusion*, se revelando superiores ou equiparáveis na maioria dos casos. A execução de métodos de terceiros é uma tarefa complexa pelo fato da grande parte dos métodos não possuírem código aberto e serem supervisionados ou semi-supervisionados. Manteve-se a maior fidelidade possível aos parâmetros utilizados pelos autores originais.

As comparações com o estado-da-arte foram possíveis em virtude da grande diversidade de coleções de imagens avaliadas e pelo fato destas serem frequentemente utilizadas como *benchmarks* em CBIR. Nota-se que os métodos propostos são comparáveis ou superiores aos demais métodos nos casos apresentados.

### 9.1 Trabalhos Futuros

Dentre os diversos trabalhos futuros possíveis, é apresentada a seguir uma lista de alguns tópicos que podem ser considerados para uma continuação desse trabalho:

- **Utilização de outros métodos de agregação:** Como as abordagens propostas para seleção e combinação são muito flexíveis, novos métodos de agregação podem ser utilizados para agregar os ranqueadores. Pode-se, por exemplo, avaliar os disponíveis no UDLF [173], uma plataforma de código aberto com diversos métodos de reranqueamento não supervisionados;
- **Proposta de novas estimativas de eficácia:** As estimativa de eficácia utilizadas neste trabalho se baseiam na vizinhança recíproca de elementos nas listas ranqueadas. Dessa forma, é possível que sejam atribuídos valores altos de estimativa de eficácia caso os ranqueadores errem de maneira consistente. Portanto, dentre os trabalhos futuros está propor e avaliar novas abordagens para estimar a eficácia das listas ranqueadas;
- **Alterações no grafo do RGSF:** O método RGSF considera apenas a correlação para formular o peso das arestas. No entanto, é possível incorporar a estimativa de eficácia e outras métricas de diversidade;
- **Programação genética:** Para a seleção e combinação de ranqueadores, o UGAF-RSF utiliza de um algoritmo genético. No entanto, é possível fazer uma modelagem de programação genética. Em tal abordagem, cada indivíduo da população pode ser representado como uma árvore na qual os nós folhas são ranqueadores e os nós externos são métodos de pós-processamento não supervisionados;
- **Outras modalidades de recuperação:** Há também a possibilidade de aplicar e avaliar os métodos propostos em outros cenários de recuperação como som, vídeo e texto, por exemplo.

# Referências

- [1] R. Datta, D. Joshi, J. Li, and J. Z. Wang, “Image retrieval: Ideas, influences, and trends of the new age,” *ACM Computing Surveys*, vol. 40, no. 2, pp. 5:1–5:60, 2008.
- [2] C. Kofler, M. Larson, and A. Hanjalic, “User intent in multimedia search: A survey of the state of the art and future challenges,” *ACM Comput. Surv.*, vol. 49, pp. 36:1–36:37, Aug. 2016.
- [3] K.-P. Yee, K. Swearingen, K. Li, and M. Hearst, “Faceted metadata for image search and browsing,” in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI ’03, pp. 401–408, 2003.
- [4] M. Agarwal and J. Mostafa, “Content-based image retrieval for alzheimer’s disease detection,” in *2011 9th International Workshop on Content-Based Multimedia Indexing (CBMI)*, pp. 13–18, June 2011.
- [5] M. Sultana and M. Gavrilova, *A Content Based Feature Combination Method for Face Recognition*, pp. 197–206. Heidelberg: Springer International Publishing, 2013.
- [6] M. Wang and T. Song, “Remote sensing image retrieval by scene semantic matching,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 51, pp. 2874–2886, May 2013.
- [7] J. pierre Schober, T. Hermes, and O. Herzog, “Content-based image retrieval by ontology-based object recognition,” in *in KI-2004 Workshop on Applications of Description Logics*, pp. 61–67, 2004.
- [8] A. W. M. Smeulders, M. Worring, S. Santini, A. Gupta, and R. Jain, “Content-based image retrieval at the end of the early years,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, pp. 1349–1380, Dec 2000.
- [9] M. S. Lew, N. Sebe, C. Djeraba, and R. Jain, “Content-based multimedia information retrieval: State of the art and challenges,” *ACM Trans. Multimedia Comput. Commun. Appl.*, vol. 2, pp. 1–19, Feb. 2006.
- [10] 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, pp. 18:1–18:26, August 2010.
- [11] B. Thomee and M. S. Lew, “Interactive search in image retrieval: a survey,” *International Journal of Multimedia Information Retrieval*, vol. 1, pp. 71–86, Jul 2012.

- [12] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “Imagenet classification with deep convolutional neural networks,” in *Proceedings of the 25th International Conference on Neural Information Processing Systems - Volume 1*, NIPS’12, pp. 1097–1105, 2012.
- [13] L. Deng, “A tutorial survey of architectures, algorithms, and applications for deep learning,” *APSIPA Transactions on Signal and Information Processing*, vol. 3, 2014.
- [14] A. Babenko, A. Slesarev, A. Chigorin, and V. Lempitsky, *Neural Codes for Image Retrieval*, pp. 584–599. Cham: Springer International Publishing, 2014.
- [15] J. Xiao, J. Hays, K. A. Ehinger, A. Oliva, and A. Torralba, “Sun database: Large-scale scene recognition from abbey to zoo,” in *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 3485–3492, 2010.
- [16] 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)*, pp. 2369–2376, 2011.
- [17] 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*, pp. 357–364, 2009.
- [18] P. Kotschieder, M. Donoser, and H. Bischof, “Beyond pairwise shape similarity analysis,” in *ACCV*, pp. 655–666, 2009.
- [19] 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.
- [20] L. P. Valem and D. C. G. Pedronette, “Unsupervised similarity learning through cartesian product of ranking references for image retrieval tasks,” *SIBGRAPI*, 2016.
- [21] 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, MIR ’10*, pp. 285–294, 2010.
- [22] M. L. Kherfi and D. Ziou, “Relevance feedback for cbir: a new approach based on probabilistic feature weighting with positive and negative examples,” *IEEE Transactions on Image Processing*, vol. 15, no. 4, pp. 1017–1030, 2006.
- [23] J. A. Shaw, E. A. Fox, J. A. Shaw, and E. A. Fox, “Combination of multiple searches,” in *The Second Text REtrieval Conference (TREC-2)*, pp. 243–252, 1994.

- [24] 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*, CIKM '15, pp. 83–92, 2015.
- [25] X. He, D. Cai, and P. Niyogi, “Laplacian score for feature selection,” in *Proceedings of the 18th International Conference on Neural Information Processing Systems*, NIPS'05, pp. 507–514, 2005.
- [26] Z. Zhao and H. Liu, “Spectral feature selection for supervised and unsupervised learning,” in *Proceedings of the 24th International Conference on Machine Learning*, ICML '07, pp. 1151–1157, 2007.
- [27] D. Cai, C. Zhang, and X. He, “Unsupervised feature selection for multi-cluster data,” in *Proceedings of the 16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '10, pp. 333–342, 2010.
- [28] Y. Yang, H. T. Shen, Z. Ma, Z. Huang, and X. Zhou, “L2,1-norm regularized discriminative feature selection for unsupervised learning,” in *Proceedings of the Twenty-Second International Joint Conference on Artificial Intelligence - Volume Volume Two*, IJCAI'11, pp. 1589–1594, 2011.
- [29] Z. Li, Y. Yang, J. Liu, X. Zhou, and H. Lu, “Unsupervised feature selection using nonnegative spectral analysis,” in *Proceedings of the Twenty-Sixth AAAI Conference on Artificial Intelligence*, 2012.
- [30] B. R.-N. Ricardo Baeza-Yates, *Recuperação de Informação: Conceitos e Tecnologia das Máquinas de Busca*. Editora Bookman, 2013.
- [31] L. Piras and G. Giacinto, “Information fusion in content based image retrieval: A comprehensive overview,” vol. 37, pp. 50 – 60, 2017.
- [32] S. V. Shriram K. Vasudevan, P.L.K. Priyadarsini, *Content Based Image Retrieval (CBIR): A deeper insight*. LAP LAMBERT Academic Publishing, 2012.
- [33] 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.
- [34] N. Arica and F. T. Y. Vural, “BAS: a perceptual shape descriptor based on the beam angle statistics,” *Pattern Recognition Letters*, vol. 24, no. 9-10, pp. 1627–1639, 2003.
- [35] J. Huang, S. R. Kumar, M. Mitra, W.-J. Zhu, and R. Zabih, “Image indexing using color correlograms,” in *CVPR*, pp. 762–768, 1997.

- [36] S. Patil and S. Talbar, *Content Based Image Retrieval Using Various Distance Metrics*, pp. 154–161. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012.
- [37] R. O. Stehling, M. A. Nascimento, and A. X. Falcão, “A compact and efficient image retrieval approach based on border/interior pixel classification,” in *CIKM*, pp. 102–109, 2002.
- [38] M. J. Swain and D. H. Ballard, “Color indexing,” *International Journal on Computer Vision*, vol. 7, no. 1, pp. 11–32, 1991.
- [39] R. da S. Torres and A. X. Falcão, “Contour Saliency Descriptors for Effective Image Retrieval and Analysis,” *Image and Vision Computing*, vol. 25, no. 1, pp. 3–13, 2007.
- [40] D. C. G. Pedronette and R. da S. Torres, “Shape retrieval using contour features and distance optimization,” in *VISAPP*, vol. 1, pp. 197 – 202, 2010.
- [41] H. Ling, X. Yang, and L. J. Latecki, “Balancing deformability and discriminability for shape matching,” in *ECCV*, vol. 3, pp. 411–424, 2010.
- [42] T. Ojala, M. Pietikäinen, and T. Mäenpää, “Multiresolution gray-scale and rotation invariant texture classification with local binary patterns,” *PAMI*, vol. 24, no. 7, pp. 971–987, 2002.
- [43] V. Kovalev and S. Volmer, “Color co-occurrence descriptors for querying-by-example,” in *ICMM*, p. 32, 1998.
- [44] B. Tao and B. W. Dickinson, “Texture recognition and image retrieval using gradient indexing,” *JVCIR*, vol. 11, no. 3, pp. 327–342, 2000.
- [45] N. Dalal and B. Triggs, “Histograms of oriented gradients for human detection,” in *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR’05)*, vol. 1, pp. 886–893 vol. 1, 2005.
- [46] D. Lowe, “Object recognition from local scale-invariant features,” in *ICCV*, pp. 1150–1157, 1999.
- [47] 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.
- [48] M. Calonder, V. Lepetit, C. Strecha, and P. Fua, *BRIEF: Binary Robust Independent Elementary Features*, pp. 778–792. Berlin, Heidelberg: Springer Berlin Heidelberg, 2010.
- [49] E. Rublee, V. Rabaud, K. Konolige, and G. Bradski, “Orb: An efficient alternative to sift or surf,” in *2011 International Conference on Computer Vision*, pp. 2564–2571, Nov 2011.

- [50] S. A. Vassou, N. Anagnostopoulos, A. Amanatiadis, K. Christodoulou, and S. A. Chatzichristofis, “Como: A compact composite moment-based descriptor for image retrieval,” in *Proceedings of the 15th International Workshop on Content-Based Multimedia Indexing*, CBMI '17, pp. 30:1–30:5, 2017.
- [51] L. Zheng, Y. Yang, and Q. Tian, “SIFT meets CNN: A decade survey of instance retrieval,” *CoRR*, vol. abs/1608.01807, 2016.
- [52] L. I. Kuncheva, *Combining Pattern Classifiers: Methods and Algorithms*. Wiley, 2004.
- [53] 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.
- [54] T. S. Huang and X. S. Zhou, “Image retrieval with relevance feedback: from heuristic weight adjustment to optimal learning methods,” in *Proceedings 2001 International Conference on Image Processing (Cat. No.01CH37205)*, vol. 3, pp. 2–5 vol.3, 2001.
- [55] S. Tong and E. Chang, “Support vector machine active learning for image retrieval,” in *Proceedings of the Ninth ACM International Conference on Multimedia*, MULTIMEDIA '01, pp. 107–118, 2001.
- [56] C. D. Ferreira, J. A. dos Santos, R. da S. Torres, M. A. Gonçalves, R. C. Rezende, and W. Fan, “Relevance feedback based on genetic programming for image retrieval,” *Pattern Recognition Letters*, vol. 32, no. 1, pp. 27–37, 2011.
- [57] R. da S. Torres, A. X. Falcão, M. A. Gonçalves, J. P. Papa, B. Zhang, W. Fan, and E. A. Fox, “A genetic programming framework for content-based image retrieval,” vol. 42, pp. 283 – 292, 2009. Learning Semantics from Multimedia Content.
- [58] J. Papa, A. Falcão, and C. Suzuki, “Supervised pattern classification based on optimum-path forest,” *Int. J. Imaging Syst. Technol.*, vol. 19, pp. 120–131, June 2009.
- [59] J. Wan, D. Wang, S. C. H. Hoi, P. Wu, J. Zhu, Y. Zhang, and J. Li, “Deep learning for content-based image retrieval: A comprehensive study,” in *Proceedings of the 22Nd ACM International Conference on Multimedia*, MM '14, pp. 157–166, 2014.
- [60] X. Yang, X. Bai, L. J. Latecki, and Z. Tu, “Improving shape retrieval by learning graph transduction,” in *ECCV*, vol. 4, pp. 788–801, 2008.
- [61] 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.

- [62] W. Voravuthikunchai, B. Crémilleux, and F. Jurie, “Image re-ranking based on statistics of frequent patterns,” in *Proceedings of International Conference on Multimedia Retrieval*, ICMR '14, pp. 129:129–129:136, 2014.
- [63] Y. Chen, X. Li, A. Dick, and R. Hill, “Ranking consistency for image matching and object retrieval,” *Pattern Recognition*, vol. 47, no. 3, pp. 1349 – 1360, 2014.
- [64] X. Bai, S. Bai, and X. Wang, “Beyond diffusion process: Neighbor set similarity for fast re-ranking,” *Information Sciences*, vol. 325, pp. 342 – 354, 2015.
- [65] D. C. G. Pedronette and R. d. S. Torres, “A correlation graph approach for unsupervised manifold learning in image retrieval tasks,” *Neurocomputing*, vol. 208, no. Supplement C, pp. 66 – 79, 2016.
- [66] D. C. G. Pedronette, J. Almeida, and R. d. 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.
- [67] 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.
- [68] F. Breve, L. Zhao, M. Quiles, W. Pedrycz, and J. Liu, “Particle competition and cooperation in networks for semi-supervised learning,” *IEEE Transactions on Knowledge and Data Engineering*, vol. 24, pp. 1686–1698, Sept 2012.
- [69] W. P. Amorim, A. X. Falcão, and M. H. d. Carvalho, “Semi-supervised pattern classification using optimum-path forest,” in *2014 27th SIBGRAPI Conference on Graphics, Patterns and Images*, pp. 111–118, Aug 2014.
- [70] S. C. H. Hoi, R. Jin, J. Zhu, and M. R. Lyu, “Semi-supervised svm batch mode active learning for image retrieval,” in *2008 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1–7, June 2008.
- [71] N. Bhowmik, V. R. González, V. Gouet-Brunet, H. Pedrini, and G. Bloch, “Efficient fusion of multidimensional descriptors for image retrieval,” in *2014 IEEE International Conference on Image Processing (ICIP)*, pp. 5766–5770, Oct 2014.
- [72] C. G. M. Snoek, M. Worring, and A. W. M. Smeulders, “Early versus late fusion in semantic video analysis,” in *Proceedings of the 13th Annual ACM International Conference on Multimedia*, MULTIMEDIA '05, pp. 399–402, 2005.

- [73] P. K. Atrey, M. A. Hossain, A. El Saddik, and M. S. Kankanhalli, “Multimodal fusion for multimedia analysis: a survey,” *Multimedia Systems*, vol. 16, pp. 345–379, Nov 2010.
- [74] J. Yu, Z. Qin, T. Wan, and X. Zhang, “Feature integration analysis of bag-of-features model for image retrieval,” *Neurocomputing*, vol. 120, no. Supplement C, pp. 355 – 364, 2013. Image Feature Detection and Description.
- [75] J. Yue, Z. Li, L. Liu, and Z. Fu, “Content-based image retrieval using color and texture fused features,” *Mathematical and Computer Modelling*.
- [76] H. J. Escalante, C. A. Hérnandez, 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*, MIR ’08, pp. 172–179, 2008.
- [77] W. Zhang, Z. Qin, and T. Wan, “Image scene categorization using multi-bag-of-features,” in *2011 International Conference on Machine Learning and Cybernetics*, vol. 4, pp. 1804–1808, July 2011.
- [78] L. Piras, R. Tronci, and G. Giacinto, *Diversity in Ensembles of Codebooks for Visual Concept Detection*, pp. 399–408. Berlin, Heidelberg: Springer Berlin Heidelberg, 2013.
- [79] L. Wang, Y. Wang, and Q. Chang, “Feature selection methods for big data bioinformatics: A survey from the search perspective,” *Methods*, vol. 111, no. Supplement C, pp. 21 – 31, 2016. Big Data Bioinformatics.
- [80] B. Ionescu, J. Benois-Pineau, T. Piatrik, and G. Qunot, *Fusion in Computer Vision: Understanding Complex Visual Content*. Springer Publishing Company, Incorporated, 2014.
- [81] J. Yang, J.-y. Yang, D. Zhang, and J.-f. Lu, “Feature fusion: parallel strategy vs. serial strategy,” *Pattern Recognition*, vol. 36, no. 6, pp. 1369 – 1381, 2003.
- [82] H. Young, “An axiomatization of borda’s rule,” *Journal of Economic Theory*.
- [83] G. V. Cormack, C. L. A. Clarke, and S. Buettcher, “Reciprocal rank fusion outperforms condorcet and individual rank learning methods,” in *Proceedings of the 32Nd International ACM SIGIR Conference on Research and Development in Information Retrieval*, SIGIR ’09, pp. 758–759, 2009.
- [84] M. Farah and D. Vanderpooten, “An outranking approach for rank aggregation in information retrieval,” in *Proceedings of the 30th Annual International ACM*

- SIGIR Conference on Research and Development in Information Retrieval*, SIGIR '07, pp. 591–598, 2007.
- [85] R. Fagin, R. Kumar, and D. Sivakumar, “Comparing top k lists,” in *Proceedings of the Fourteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '03, pp. 28–36, 2003.
- [86] T.-Y. Liu, “Learning to rank for information retrieval,” *Found. Trends Inf. Retr.*, vol. 3, pp. 225–331, Mar. 2009.
- [87] T. Joachims, “Training linear svms in linear time,” in *Proceedings of the 12th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '06, pp. 217–226, 2006.
- [88] C. Burges, T. Shaked, E. Renshaw, A. Lazier, M. Deeds, N. Hamilton, and G. Hullender, “Learning to rank using gradient descent,” in *Proceedings of the 22Nd International Conference on Machine Learning*, ICML '05, pp. 89–96, 2005.
- [89] Z. Cao, T. Qin, T.-Y. Liu, M.-F. Tsai, and H. Li, “Learning to rank: From pairwise approach to listwise approach,” in *Proceedings of the 24th International Conference on Machine Learning*, ICML '07, pp. 129–136, 2007.
- [90] J. Xu and H. Li, “Adarank: A boosting algorithm for information retrieval,” in *Proceedings of the 30th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, SIGIR '07, pp. 391–398, 2007.
- [91] S. S. Bucak, R. Jin, and A. K. Jain, “Multiple kernel learning for visual object recognition: A review,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 36, no. 7, pp. 1354–1369, 2014.
- [92] B. Scholkopf and A. J. Smola, *Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond*. Cambridge, MA, USA: MIT Press, 2001.
- [93] F. R. Bach, G. R. G. Lanckriet, and M. I. Jordan, “Multiple kernel learning, conic duality, and the smo algorithm,” in *Proceedings of the Twenty-first International Conference on Machine Learning*, ICML '04, pp. 6–, 2004.
- [94] S. Sonnenburg, G. Rätsch, and C. Schäfer, “A general and efficient multiple kernel learning algorithm,” in *Proceedings of the 18th International Conference on Neural Information Processing Systems*, NIPS'05, pp. 1273–1280, 2005.
- [95] M. Kowalski, M. Szafranski, and L. Ralaivola, “Multiple indefinite kernel learning with mixed norm regularization,” in *Proceedings of the 26th Annual International Conference on Machine Learning*, ICML '09, pp. 545–552, 2009.

- [96] A. Pinar, T. C. Havens, D. T. Anderson, and L. Hu, “Feature and decision level fusion using multiple kernel learning and fuzzy integrals,” in *2015 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*, pp. 1–7, Aug 2015.
- [97] M. Haghigat, M. Abdel-Mottaleb, and W. Alhalabi, “Fully automatic face normalization and single sample face recognition in unconstrained environments,” *Expert Systems with Applications*, vol. 47, no. Supplement C, pp. 23 – 34, 2016.
- [98] M. Haghigat, M. Abdel-Mottaleb, and W. Alhalabi, “Discriminant correlation analysis: Real-time feature level fusion for multimodal biometric recognition,” *IEEE Transactions on Information Forensics and Security*, vol. 11, pp. 1984–1996, Sept 2016.
- [99] F. R. Bach, “Consistency of the group lasso and multiple kernel learning,” *J. Mach. Learn. Res.*, vol. 9, pp. 1179–1225, June 2008.
- [100] S. Zhang, M. Yang, T. Cour, K. Yu, and D. Metaxas, “Query specific rank fusion for image retrieval,” *IEEE TPAMI*, vol. 37, pp. 803–815, April 2015.
- [101] B. Wang, J. Jiang, W. Wang, Z.-H. Zhou, and Z. Tu, “Unsupervised metric fusion by cross diffusion,” in *CVPR*, pp. 3013 –3020, 2012.
- [102] J. Shao, L. Wang, Z. Zhao, F. Su, and A. Cai, “Deep canonical correlation analysis with progressive and hypergraph learning for cross-modal retrieval,” *Neurocomputing*, vol. 214, no. Supplement C, pp. 618 – 628, 2016.
- [103] W. He, X. Cheng, R. Hu, Y. Zhu, and G. Wen, “Feature self-representation based hypergraph unsupervised feature selection via low-rank representation,” *Neurocomput.*, vol. 253, pp. 127–134, Aug. 2017.
- [104] X. Cheng, Y. Zhu, J. Song, G. Wen, and W. He, “A novel low-rank hypergraph feature selection for multi-view classification,” *Neurocomputing*, vol. 253, no. Supplement C, pp. 115 – 121, 2017. Learning Multimodal Data.
- [105] L. Zhang, Y. Gao, C. Hong, Y. Feng, J. Zhu, and D. Cai, “Feature correlation hypergraph: Exploiting high-order potentials for multimodal recognition,” *IEEE Transactions on Cybernetics*, vol. 44, pp. 1408–1419, Aug 2014.
- [106] M. Akbari, L. Nie, and T.-S. Chua, “amm: Towards adaptive ranking of multi-modal documents,” *International Journal of Multimedia Information Retrieval*, vol. 4, pp. 233–245, Dec 2015.
- [107] H. Jegou, M. Douze, and C. Schmid, “Hamming embedding and weak geometric consistency for large scale image search,” in *European Conference on Computer Vision, ECCV ’08*, pp. 304–317, 2008.

- [108] M.-E. Nilsback and A. Zisserman, “A visual vocabulary for flower classification,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, vol. 2, pp. 1447–1454, 2006.
- [109] L. J. Latecki, R. Lakmper, and U. Eckhardt, “Shape descriptors for non-rigid shapes with a single closed contour,” in *CVPR*, pp. 424–429, 2000.
- [110] G.-H. Liu and J.-Y. Yang, “Content-based image retrieval using color difference histogram,” *Pattern Recognition*, vol. 46, no. 1, pp. 188 – 198, 2013.
- [111] D. Nistér and H. Stewénus, “Scalable recognition with a vocabulary tree,” in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR’2006)*, vol. 2, pp. 2161–2168, 2006.
- [112] M. Lux, “Content based image retrieval with lire,” in *Proceedings of the 19th ACM International Conference on Multimedia*, MM ’11, pp. 735–738, 2011.
- [113] L. Cieplinski, “Mpeg-7 color descriptors and their applications,” in *Computer Analysis of Images and Patterns* (W. Skarbek, ed.), (Berlin, Heidelberg), pp. 11–20, Springer Berlin Heidelberg, 2001.
- [114] K. E. A. van de Sande, T. Gevers, and C. G. M. Snoek, “Evaluating color descriptors for object and scene recognition,” *PAMI*, vol. 32, no. 9, pp. 1582–1596, 2010.
- [115] R. Gopalan, P. Turaga, and R. Chellappa, “Articulation-invariant representation of non-planar shapes,” in *ECCV*, vol. 3, pp. 286–299, 2010.
- [116] H. Ling and D. W. Jacobs, “Shape classification using the inner-distance,” *IEEE TPAMI*, vol. 29, no. 2, pp. 286–299, 2007.
- [117] B. S. Manjunath, J. R. Ohm, V. V. Vasudevan, and A. Yamada, “Color and texture descriptors,” *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 11, no. 6, pp. 703–715, 2001.
- [118] S. A. Chatzichristofis and Y. S. Boutalis, “Cedd: color and edge directivity descriptor: a compact descriptor for image indexing and retrieval,” in *ICVS*, pp. 312–322, 2008.
- [119] S. A. Chatzichristofis and Y. S. Boutalis, “Fcth: Fuzzy color and texture histogram - a low level feature for accurate image retrieval,” in *WIAMIS*, pp. 191–196, 2008.
- [120] K. Zagoris, S. Chatzichristofis, N. Papamarkos, and Y. Boutalis, “Automatic image annotation and retrieval using the joint composite descriptor,” in *PCI*, pp. 143–147, 2010.
- [121] A. Oliva and A. Torralba, “Modeling the shape of the scene: A holistic representation of the spatial envelope,” *IJCV*, vol. 42, pp. 145–175, May 2001.

- [122] X. Wang, M. Yang, T. Cour, S. Zhu, K. Yu, and T. Han, “Contextual weighting for vocabulary tree based image retrieval,” in *ICCV’2011*, pp. 209–216, Nov 2011.
- [123] J. Hu, L. Shen, and G. Sun, “Squeeze-and-excitation networks,” in *2018 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2018.
- [124] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 770–778, 2016.
- [125] S. Xie, R. Girshick, P. Dollár, Z. Tu, and K. He, “Aggregated residual transformations for deep neural networks,” in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 5987–5995, 2017.
- [126] Y. Chen, J. Li, H. Xiao, X. Jin, S. Yan, and J. Feng, “Dual path networks,” in *Advances in Neural Information Processing Systems 30* (I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett, eds.), pp. 4467–4475, Curran Associates, Inc., 2017.
- [127] S. Liu and W. Deng, “Very deep convolutional neural network based image classification using small training sample size,” in *2015 3rd IAPR Asian Conference on Pattern Recognition (ACPR)*, pp. 730–734, 2015.
- [128] C. Szegedy, S. Ioffe, V. Vanhoucke, and A. Alemi, “Inception-v4, inception-resnet and the impact of residual connections on learning,” 2017.
- [129] S. Ioffe and C. Szegedy, “Batch normalization: Accelerating deep network training by reducing internal covariate shift,” in *Proceedings of the 32Nd International Conference on International Conference on Machine Learning - Volume 37*, ICML’15, pp. 448–456, 2015.
- [130] B. Zoph, V. Vasudevan, J. Shlens, and Q. V. Le, “Learning transferable architectures for scalable image recognition,” *CoRR*, vol. abs/1707.07012, 2017.
- [131] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “Imagenet classification with deep convolutional neural networks,” in *Proceedings of the 25th International Conference on Neural Information Processing Systems - Volume 1*, NIPS’12, pp. 1097–1105, 2012.
- [132] F. Chollet, “Xception: Deep learning with depthwise separable convolutions,” in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1800–1807, 2017.

- [133] K. Reddy Mopuri and R. Venkatesh Babu, “Object level deep feature pooling for compact image representation,” in *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, June 2015.
- [134] A. Paszke, S. Gross, S. Chintala, G. Chanan, E. Yang, Z. DeVito, Z. Lin, A. Desmaison, L. Antiga, and A. Lerer, “Automatic differentiation in pytorch,” in *NIPS-W*, 2017.
- [135] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, “ImageNet: A Large-Scale Hierarchical Image Database,” in *CVPR09*, 2009.
- [136] D. C. G. Pedronette and R. da S. Torres, “A correlation graph approach for unsupervised manifold learning in image retrieval tasks,” *Neurocomputing*, vol. 208, no. Sup C, pp. 66 – 79, 2016.
- [137] L. Zheng, S. Wang, L. Tian, F. He, Z. Liu, and Q. Tian, “Query-adaptive late fusion for image search and person re-identification,” in *CVPR*, 2015.
- [138] L. Page, S. Brin, R. Motwani, and T. Winograd, “The pagerank citation ranking: Bringing order to the web,” 1999.
- [139] J. Li, K. Cheng, S. Wang, F. Morstatter, R. P. Trevino, J. Tang, and H. Liu, “Feature selection: A data perspective,” *arXiv preprint arXiv:1601.07996*, 2016.
- [140] A. Y. Ng, “Feature selection, l1 vs. l2 regularization, and rotational invariance,” in *ICML '04*, 2004.
- [141] N. Jardine and C. van Rijsbergen, “The use of hierarchic clustering in information retrieval,” *Information Storage and Retrieval*, vol. 7, no. 5, pp. 217 – 240, 1971.
- [142] D. C. G. Pedronette, O. A. Penatti, and R. d. S. Torres, “Unsupervised manifold learning using reciprocal knn graphs in image re-ranking and rank aggregation tasks,”
- [143] D. C. G. Pedronette and R. d. S. Torres, “Unsupervised effectiveness estimation for image retrieval using reciprocal rank information,” in *2015 28th SIBGRAPI Conference on Graphics, Patterns and Images*, pp. 321–328, 2015.
- [144] C. Y. Okada, D. C. G. Pedronette, and R. da S. Torres, “Unsupervised distance learning by rank correlation measures for image retrieval,” in *ACM International Conference on Multimedia Retrieval (ICMR'2015)*, 2015.
- [145] W. Webber, A. Moffat, and J. Zobel, “A similarity measure for indefinite rankings,” *ACM Transactions on Information Systems*, vol. 28, no. 4, pp. 20:1–20:38, 2010.

- [146] 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*, ICMR ’15, (New York, NY, USA), pp. 331–338, ACM, 2015.
- [147] R. Fagin, R. Kumar, and D. Sivakumar, “Comparing top k lists,” in *ACM-SIAM Symposium on Discrete algorithms (SODA’03)*, pp. 28–36, 2003.
- [148] L. P. Valem and D. C. G. a. Pedronette, “Selection and combination of unsupervised learning methods for image retrieval,” in *Proceedings of the 15th International Workshop on Content-Based Multimedia Indexing*, CBMI ’17, pp. 27:1–27:6, 2017.
- [149] H. Jegou, C. Schmid, H. Harzallah, and J. Verbeek, “Accurate image search using the contextual dissimilarity measure,” *PAMI*, vol. 32, no. 1, pp. 2–11, 2010.
- [150] J. Jiang, B. Wang, and Z. Tu, “Unsupervised metric learning by self-smoothing operator,” in *ICCV*, pp. 794–801, 2011.
- [151] S. Bai and X. Bai, “Sparse contextual activation for efficient visual re-ranking,” *IEEE Trans. on Image Processing (TIP)*, vol. 25, no. 3, pp. 1056–1069, 2016.
- [152] S. Bai, S. Sun, X. Bai, Z. Zhang, and Q. Tian, “Smooth neighborhood structure mining on multiple affinity graphs with applications to context-sensitive similarity,” in *European Conference on Computer Vision (ECCV)*, pp. 592–608, 2016.
- [153] D. C. G. Pedronette, F. M. F. Gonçalves, and I. R. Guilherme, “Unsupervised manifold learning through reciprocal kNN graph and Connected Components for image retrieval tasks,” *Pattern Recognition*, 2017. On-line, To appear.
- [154] D. C. G. Pedronette, J. Almeida, and R. da S. Torres, “A scalable re-ranking method for content-based image retrieval,” *Information Sciences*, vol. 265, no. 1, pp. 91–104, 2014.
- [155] 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.
- [156] M. Donoser and H. Bischof, “Diffusion processes for retrieval revisited,” in *Conf. on Computer Vision and Pattern Recognition (CVPR)*, pp. 1320–1327, 2013.
- [157] X. Yang, L. Prasad, and L. Latecki, “Affinity learning with diffusion on tensor product graph,” *IEEE TPAMI*, vol. 35, no. 1, pp. 28–38, 2013.
- [158] X. Bai, S. Bai, and X. Wang, “Beyond diffusion process: Neighbor set similarity for fast re-ranking,” *Information Sciences*, vol. 325, pp. 342 – 354, 2015.

- [159] Z. Liu, S. Wang, L. Zheng, and Q. Tian, “Robust imagegraph: Rank-level feature fusion for image search,” *IEEE Transactions on Image Processing*, vol. 26, no. 7, pp. 3128–3141, 2017.
- [160] G. Tolias, Y. Avrithis, and H. Jégou, “To aggregate or not to aggregate: Selective match kernels for image search,” in *IEEE International Conference on Computer Vision (ICCV’2013)*, pp. 1401–1408, Dec 2013.
- [161] M. Paulin, J. Mairal, M. Douze, Z. Harchaoui, F. Perronnin, and C. Schmid, “Convolutional patch representations for image retrieval: An unsupervised approach,” *Int. Journal of Computer Vision*, 2017.
- [162] D. Qin, C. Wengert, and L. V. Gool, “Query adaptive similarity for large scale object retrieval,” in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR’2013)*, pp. 1610–1617, June 2013.
- [163] L. Zheng, S. Wang, and Q. Tian, “Coupled binary embedding for large-scale image retrieval,” *IEEE Transactions on Image Processing (TIP)*, vol. 23, no. 8, pp. 3368–3380, 2014.
- [164] S. Sun, Y. Li, W. Zhou, Q. Tian, and H. Li, “Local residual similarity for image re-ranking,” *Information Sciences*, vol. 417, no. Sup. C, pp. 143 – 153, 2017.
- [165] L. Zheng, S. Wang, Z. Liu, and Q. Tian, “Packing and padding: Coupled multi-index for accurate image retrieval,” in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR’2014)*, pp. 1947–1954, June 2014.
- [166] X. Li, M. Larson, and A. Hanjalic, “Pairwise geometric matching for large-scale object retrieval,” in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR’2015)*, pp. 5153–5161, June 2015.
- [167] S. Bai, X. Bai, Q. Tian, and L. J. Latecki, “Regularized diffusion process for visual retrieval,” in *Conf. on Artificial Intelligence (AAAI)*, pp. 3967–3973, 2017.
- [168] L. Zheng, S. Wang, and Q. Tian, “Lp-norm idf for scalable image retrieval,” *IEEE TIP*, vol. 23, pp. 3604–3617, Aug 2014.
- [169] L. Xie, R. Hong, B. Zhang, and Q. Tian, “Image classification and retrieval are one,” in *ACM ICMR’2015*, pp. 3–10, 2015.
- [170] J. Almeida, L. P. Valem, and D. C. G. a. Pedronette, “A rank aggregation framework for video interestingness prediction,” in *Proceedings of the 9th International Conference on Image Analysis and Processing, ICIAP ’17*, 2017.

- 
- [171] L. P. Valem, D. C. G. Pedronette, F. Breve, and I. R. Guilherme, “Manifold correlation graph for semi-supervised learning,” in *2018 International Joint Conference on Neural Networks (IJCNN)*, pp. 1–7, 2018.
  - [172] L. P. Valem, C. R. D. Oliveira, D. C. G. a. Pedronette, and J. Almeida, “Unsupervised similarity learning through rank correlation and knn sets,” vol. 14, (New York, NY, USA), pp. 80:1–80:23, ACM, Oct. 2018.
  - [173] L. P. Valem and D. C. G. a. Pedronette, “An unsupervised distance learning framework for multimedia retrieval,” in *Proceedings of the 2017 ACM on International Conference on Multimedia Retrieval*, ICMR '17, (New York, NY, USA), pp. 107–111, ACM, 2017.

## TERMO DE REPRODUÇÃO XEROGRÁFICA

Autorizo a reprodução xerográfica do presente Trabalho de Conclusão, na íntegra ou em partes, para fins de pesquisa.

São José do Rio Preto, 15/03/2019



---

Assinatura do autor