

UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ
CÂMPUS CORNÉLIO PROCÓPIO
DIRETORIA DE PESQUISA E PÓS-GRADUAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM INFORMÁTICA

RAFAEL STAIGER BRESSAN

**APRENDIZADO ATIVO PARA RECUPERAÇÃO E CLASSIFICAÇÃO
DE IMAGENS**

DISSERTAÇÃO – MESTRADO

CORNÉLIO PROCÓPIO

2018

RAFAEL STAIGER BRESSAN

**APRENDIZADO ATIVO PARA RECUPERAÇÃO E CLASSIFICAÇÃO
DE IMAGENS**

Dissertação apresentada ao Programa de Pós-Graduação em Informática da Universidade Tecnológica Federal do Paraná – UTFPR como requisito parcial para a obtenção do título de “Mestre Profissional em Informática”.

Orientadora: Profa. Dra. Priscila Tiemi Maeda Saito

Co-orientador: Prof. Dr. Pedro Henrique Bugatti

CORNÉLIO PROCÓPIO

2018

Dados Internacionais de Catalogação na Publicação

B843 Bressan, Rafael Staiger

Aprendizado ativo para recuperação e classificação de imagens/ Rafael Staiger Bressan. – 2018.
86 f. : il. ; 31 cm.

Orientador: Priscila Tiemi Maeda Saito .
Dissertação (Mestrado) – Universidade Tecnológica Federal do Paraná.
Programa de Pós-Graduação em Informática. Cornélio Procópio, 2018.
Bibliografia: p. 77-86.

1. Banco de dados. 2. Processamento de imagens. 3. Sistemas de reconhecimento de padrões. 4. Informática – Dissertações. I. Saito, Priscila Tiemi Maeda , orient. II. Universidade Tecnológica Federal do Paraná. Programa de Pós-Graduação em Informática. III. Título.

Biblioteca da UTFPR - Câmpus Cornélio Procópio

Bibliotecário/Documentalista responsável:
Simone Fidencio de Oliveira Guerra – CRB9/1276



Título da Dissertação Nº 52:

“APRENDIZADO ATIVO PARA RECUPERAÇÃO E CLASSIFICAÇÃO DE IMAGENS”.

por

Rafael Staiger Bressan

Orientadora: **Profa. Dra. Priscila Tiemi Maeda Saito**

Co-orientador: **Prof. Dr. Pedro Henrique Bugatti**

Esta dissertação foi apresentada como requisito parcial à obtenção do grau de MESTRE EM INFORMÁTICA – Área de Concentração: Computação Aplicada, pelo Programa de Pós-Graduação em Informática – PPGI – da Universidade Tecnológica Federal do Paraná – UTFPR – Câmpus Cornélio Procópio, às 15h do dia 15 de agosto de 2018. O trabalho foi _____ pela Banca Examinadora, composta pelos professores:

Prof. Dr. Danilo Sipoli Sanches
(Presidente – UTFPR-CP)

Prof. Dr. Alexandre Rossi Paschoal
(UTFPR-CP)

Prof. Dr. Sylvio Barbon Junior
(DC-UEL)

Profa. Dra. Priscila Tiemi Maeda Saito
(UTFPR-CP)

Participação à distância via _____

Visto da coordenação:

Danilo Sipoli Sanches
Coordenador do Programa de Pós-Graduação em Informática
UTFPR Câmpus Cornélio Procópio

A Folha de Aprovação assinada encontra-se na Coordenação do Programa.

“Cada pessoa que passa em nossa vida, passa sozinha, é porque cada pessoa é única e nenhuma substitui a outra! Cada pessoa que passa em nossa vida passa sozinha e não nos deixa só porque deixa um pouco de si e leva um pouquinho de nós. Essa é a mais bela responsabilidade da vida e a prova de que as pessoas não se encontram por acaso.”

Charles Chaplin

AGRADECIMENTOS

Agradeço cada pessoa que passou pela minha vida, pois sem elas eu não seria quem sou. Manifesto os meus sinceros e reconhecidos agradecimentos, em particular aos Professores do PPGI da UTFPR pela confiança e pelos conhecimentos passados em aula e fora dela. A minha orientadora e coorientador Profa. Dra. Priscila Tiemi Maeda Saito e Prof. Dr. Pedro Henrique Bugatti, que não mediram esforços para sanar minhas dúvidas. Às instituições de fomento e apoio CNPq (*grants* #472625/2014-4, #384034/2015-2, #431668/2016-7), CAPES, Fundação Araucária, FAPESP, SETI, Belagrícola e UTFPR. Por último, a minha esposa Kaline Flauzino Maciel e minha família que eu amo e a qual dedico todos os meus esforços.

”Só existem dois dias no ano em que nada pode ser feito. Um se chama **ontem** e o outro se chama **amanhã**, portanto hoje é o dia certo para **AMAR, ACREDITAR, FAZER** e principalmente **VIVER.**”

Dalai Lama

RESUMO

BRESSAN, Rafael S.. APRENDIZADO ATIVO PARA RECUPERAÇÃO E CLASSIFICAÇÃO DE IMAGENS. 86 f. Dissertação – Mestrado – Programa de Pós-Graduação em Informática, Universidade Tecnológica Federal do Paraná. Cornélio Procópio, 2018.

Atualmente, bancos de dados de imagens vêm crescendo, surgindo a necessidade de otimização e de aceleração dos processos de recuperação e classificação de imagens, em conjunto com a melhoria da qualidade dos resultados retornados. Neste contexto, este trabalho propõe a utilização de estratégias de aprendizado ativo para classificação e recuperação de imagens, de forma a selecionar amostras mais informativas e minimizar a interação do especialista durante o processo de aprendizado. Além disso, novas estratégias de aprendizado ativo são propostas para as tarefas de classificação e de recuperação de imagens baseadas em conteúdo. Para validação das propostas, foram realizados experimentos utilizando conjuntos de dados de diferentes domínios de aplicação. A partir dos resultados obtidos, é possível observar ganhos significativos apresentados pelas propostas em relação às estratégias amplamente utilizadas na literatura.

Palavras-chave: CBIR, aprendizado ativo, realimentação de relevância, reconhecimento de padrões, classificação de imagens.

ABSTRACT

BRESSAN, Rafael S.. ACTIVE LEARNING FOR IMAGE RETRIEVAL AND CLASSIFICATION. 86 f. Dissertação – Mestrado – Programa de Pós-Graduação em Informática, Universidade Tecnológica Federal do Paraná. Cornélio Procópio, 2018.

Currently, image databases have been growing, resulting in the need for optimization and acceleration of the image retrieval and classification processes, together with the improvement of the quality of the returned results. In this context, this work proposes the use of active learning strategies for image classification and retrieval, in order to select more informative samples and to minimize the interaction of the specialist during the learning process. In addition, new active learning strategies are proposed for classification and content-based image retrieval tasks. To validate the proposals, experiments were performed using datasets from different application domains. From the obtained results, it is possible to observe significant gains presented by the proposals in relation to the strategies widely used in the literature.

Keywords: CBIR, active learning, relevance feedback, pattern recognition, image classification.

LISTA DE FIGURAS

FIGURA 1	– Processo de recuperação de imagens baseado em conteúdo.	17
FIGURA 2	– Processo de CBIR com Realimentação de Relevância	20
FIGURA 3	– Técnica de realimentação de relevância <i>Query Point Movement</i> - QPM .	21
FIGURA 4	– Técnica de realimentação de relevância <i>Query Expansion</i> - QEX	21
FIGURA 5	– Processo de Aprendizado Ativo.	23
FIGURA 6	– Exemplos de estratégias de seleção, com base em critérios como: (a) aleatoriedade, (b) diversidade, (c) representatividade e (d) incerteza.	24
FIGURA 7	– Exemplo de separação de classes utilizando o classificador SVM.	26
FIGURA 8	– Pipeline da estratégia de avaliação de desempenho de diferentes classificadores em diferentes conjuntos biomédicos.	28
FIGURA 9	– Pipeline da abordagem de aprendizado ativo proposta para classificação.	35
FIGURA 10	– Acurácia média por iteração dos métodos de aprendizado ativo no conjunto de dados <i>MLL Leukemia</i> utilizando diferentes classificadores.	41
FIGURA 11	– Acurácia média por iteração dos métodos de aprendizado ativo no conjunto de dados <i>Breast Cancer</i> utilizando diferentes classificadores.	43
FIGURA 12	– Acurácia média por iteração dos métodos de aprendizado ativo no conjunto de dados <i>Central Nervous System</i> utilizando diferentes classificadores.	44
FIGURA 13	– Acurácia média por iteração dos métodos de aprendizado ativo no conjunto de dados <i>Colon Tumor</i> utilizando diferentes classificadores.	45
FIGURA 14	– Acurácia média por iteração dos métodos de aprendizado ativo no conjunto de dados <i>Lung Cancer</i> utilizando diferentes classificadores.	46
FIGURA 15	– Acurácia média por iteração dos métodos de aprendizado ativo no conjunto de dados <i>Subtypes of Acute Lymphoblastic Leukemia</i> utilizando diferentes classificadores.	47
FIGURA 16	– Pipeline da metodologia proposta para avaliação de desempenho de descritores.	50
FIGURA 17	– Resultados obtidos para o conjunto de dados B_1 , utilizando os melhores descritores tradicional e profundos, ao longo das consultas por similaridade: (a) consulta inicial e, refinamentos das consultas: (b) primeira iteração; (c) segunda iteração; e (d) terceira iteração.	56
FIGURA 18	– Resultados obtidos para o conjunto de dados B_2 , utilizando os melhores descritores tradicional e profundos, ao longo das consultas por similaridade: (a) consulta inicial e, refinamentos das consultas: (b) primeira iteração; (c) segunda iteração; e (d) terceira iteração.	57
FIGURA 19	– Exemplos de amostras informativas (incertas) de duas classes diferentes (a) lesões benignas e (b) lesões malignas.	58
FIGURA 20	– Pipeline da abordagem de aprendizado ativo proposta para realimentação de relevância em CBIR.	59
FIGURA 21	– Curvas P&R obtidas pelas abordagens, utilizando o conjunto de dados ETH-80, considerando: (a) primeira, (b) terceira, (c) quinta e (d) oitava iteração.	69

FIGURA 22 – Curvas P&R obtidas pelas abordagens, utilizando o conjunto de dados Cevada ₁ , considerando: (a) primeira, (b) terceira, (c) quinta e (d) oitava iteração.	70
FIGURA 23 – Curvas P&R obtidas pelas abordagens, utilizando o conjunto de dados Cevada ₂ , considerando: (a) primeira, (b) terceira, (c) quinta e (d) oitava iteração.	71
FIGURA 24 – Curvas P&R obtidas pelas abordagens, utilizando o conjunto de dados Soja ₁ , considerando: (a) primeira, (b) terceira, (c) quinta e (d) oitava iteração.	72
FIGURA 25 – Curvas P&R obtidas pelas abordagens, utilizando o conjunto de dados Soja ₂ , considerando: (a) primeira, (b) terceira, (c) quinta e (d) oitava iteração.	72
FIGURA 26 – Curvas P&R obtidas pelas abordagens, utilizando o conjunto de dados Soja ₃ , considerando: (a) primeira, (b) terceira, (c) quinta e (d) oitava iteração.	73
FIGURA 27 – Curvas P&R obtidas pelas abordagens, utilizando o conjunto de dados Soja ₄ , considerando: (a) primeira, (b) terceira, (c) quinta e (d) oitava iteração.	74

LISTA DE TABELAS

TABELA 1	– Funções de Distância.	19
TABELA 2	– Acurácias médias obtidas por cada classificador em cada conjunto de dados (CAMARGO et al., 2017).	31
TABELA 3	– Tempos médios de treinamento obtidos por cada classificador em cada conjunto de dados (CAMARGO et al., 2017).	31
TABELA 4	– Tempos médios de teste obtidos por cada classificador em cada conjunto de dados (CAMARGO et al., 2017).	32
TABELA 5	– Desempenho médio para cada classificador. (CAMARGO et al., 2017).	32
TABELA 6	– Média harmônica obtida para cada classificador em cada conjunto de dados (CAMARGO et al., 2017).	33
TABELA 7	– Acurácia média ao longo das iterações obtidas pelos métodos de aprendizado ativo, utilizando cada um dos classificadores, em cada conjunto de dados.	48
TABELA 8	– Descrição do conjunto de imagens VIENNA - B_1	52
TABELA 9	– Descrição do conjunto de imagens MINI-MIAS - B_2	53
TABELA 10	– Extratores de características tradicionais (BRESSAN et al., 2018a). ...	53
TABELA 11	– Melhores descritores para o conjunto de dados B_1 . Extratores de características e funções de distância em negrito compreendem os melhores pares para caraterísticas tradicionais e profundas (BRESSAN et al., 2018a).	55
TABELA 12	– Melhores descritores para o conjunto de dados B_2 . Extratores de características e funções de distância em negrito compreendem os melhores pares para caraterísticas tradicionais e profundas (BRESSAN et al., 2018a).	55
TABELA 13	– Descrição dos conjuntos de imagens de semente de soja utilizados nos experimentos	65
TABELA 14	– Propriedades de cada extrator de características.	67
TABELA 15	– Melhor descritor (par extrator de característica e função de distância) para o conjunto de imagens ETH-80.	68
TABELA 16	– Melhor descritor (par extrator de característica e função de distância) para os conjuntos de imagens Cevada ₁ e Cevada ₂	70
TABELA 17	– Melhor descritor (par extrator de característica e função de distância) para os conjuntos de imagens Soja ₁ , Soja ₂ , Soja ₃ e Soja ₄	71
TABELA 18	– Precisões médias gerais obtidas por cada abordagem nos conjuntos de imagens, considerando todas as iterações de realimentação.	74

LISTA DE SIGLAS

AL	<i>Active Learning</i>
AM	<i>Aprendizado de Máquina</i>
BIC	<i>Border/Interior Classification</i>
CBIR	<i>Content-Based Image Retrieval</i>
CBVIR	<i>Content Visual Information Retrieval</i>
CSS	<i>Curvature Scala Space</i>
GCH	<i>Global Color Histogram</i>
HC	<i>Hierarchical Cluster</i>
IA	<i>Inteligência Artificial</i>
<i>k</i> -NN	<i>k-Nearest Neighbor</i>
LAS	<i>Local Activity Spectrum</i>
LBP	<i>Local Binary Patterns</i>
LCH	<i>Local Color Histogram</i>
LRGOD	<i>Locally Regressive G-Optimal Design</i>
MAP	<i>Mean Average Precision</i>
MINI-MIAS	<i>Mammographic Image Analysis Society</i>
MS	<i>Margin Sample</i>
NB	<i>Naive Bayes</i>
OPF	<i>Florestas de Caminho Ótimo</i>
PHOG	<i>Generates Pyramid Histograms of Oriented Gradients</i>
QBIC	<i>Query By Image Content</i>
QEX	<i>Query Expansion</i>
QPM	<i>Query Point Movement</i>
RF	<i>Random Forest</i>
RGB	<i>Red Green e Blue</i>
ROIs	<i>Regions Of Interests</i>
SVM	<i>Support Vector Machine</i>

SUMÁRIO

1	INTRODUÇÃO	13
1.1	OBJETIVOS	15
1.1.1	Objetivo Geral	15
1.1.2	Objetivos Específicos	15
1.2	ORGANIZAÇÃO DO TEXTO	15
2	FUNDAMENTAÇÃO TEÓRICA	16
2.1	RECUPERAÇÃO DE IMAGENS BASEADA EM CONTEÚDO	16
2.1.1	Descritores	17
2.1.2	Realimentação de Relevância	19
2.2	APRENDIZADO DE MÁQUINA	22
2.2.1	Aprendizado Ativo	23
2.2.1.1	Estratégias de Seleção	24
2.3	COMBINAÇÃO DE APRENDIZADO ATIVO E CBIR	25
3	ABORDAGEM DE APRENDIZADO PARA CLASSIFICAÇÃO	27
3.1	ESTRATÉGIA PARA AVALIAÇÃO DE DESEMPENHO DO APRENDIZADO	27
3.1.1	Descrição dos conjuntos de dados	28
3.1.2	Descrição dos experimentos	30
3.1.3	Resultados e Discussões	30
3.2	ESTRATÉGIA DE APRENDIZADO ATIVO	33
3.2.1	Descrição dos Conjuntos de Dados	39
3.2.2	Descrição dos Experimentos	40
3.2.3	Resultados e Discussões	40
4	ABORDAGEM DE APRENDIZADO PARA CBIR	49
4.1	ESTRATÉGIA PARA AVALIAÇÃO DE DESEMPENHO DE DESCRITORES	49
4.1.1	Descrição dos Conjuntos de Dados	52
4.1.2	Descrição dos Experimentos	53
4.1.3	Resultados e Discussões	54
4.2	ESTRATÉGIAS DE APRENDIZADO ATIVO	57
4.2.1	Estratégia de Seleção ourAL-2	62
4.2.2	Estratégia de Seleção ourAL-3	63
4.2.3	Descrição dos Conjuntos de Dados	65
4.2.4	Descrição dos Experimentos	66
4.2.5	Resultados e Discussões	67
5	CONSIDERAÇÕES FINAIS	75
	REFERÊNCIAS	77

1 INTRODUÇÃO

Atualmente, bases de dados (textos, imagens, áudios, vídeos, entre outros) tornam-se cada vez maiores com os avanços nos dispositivos de aquisição e armazenamento de dados. O desenvolvimento de abordagens automáticas de recuperação e classificação tornou-se necessário para gerenciamento e organização desses dados.

A maioria das abordagens, geralmente, lida com dados textuais e numéricos. No entanto, em alguns casos, é possível a associação das informações textuais com informações gráficas, provenientes de imagens. Técnicas de Recuperação de Imagens Baseada em Conteúdo (*Content-Based Image Retrieval - CBIR*) têm sido amplamente utilizadas e bem sucedidas em diferentes domínios de imagens (ULUWITIGE et al., 2016; MORITA et al., 2016; AMATO et al., 2016; MALODE; GUMASTE, 2015; BUGATTI, 2008; TORRES; FALCAO, 2006; EAKINS et al., 1991), possibilitando a realização de consultas e recuperação de tais imagens baseando-se apenas no conteúdo das mesmas.

Na área médica, por exemplo, sistemas CBIR têm sido uma ferramenta bastante eficaz no auxílio ao diagnóstico médico (PALLE; KULKARNI, 2016; BUGATTI, 2008; ZHOU et al., 2008; MÜLLER et al., 2004), nos quais é possível encontrar registros de pacientes cujas imagens associadas apresentam aspectos de similaridade, utilizando padrões gráficos, bem como obter correlações utilizando sintomas previamente registrados, tratamentos efetuados, entre outros. Sistemas CBIR podem então ser capazes de responder a consultas por similaridade, tais como: “quais são as 20 imagens mais similares (próximas) de uma dada imagem de consulta?” ou “quais são as imagens que diferem de até 10 unidades de uma imagem padrão?”

No entanto, o conjunto de características extraídas das imagens, a função de distância e a interação de um especialista no processo são aspectos chaves para obter resultados mais precisos e correspondentes às expectativas do especialista, a partir das consultas por similaridade realizadas. Existe na literatura (BUGATTI, 2008; TORRES; FALCAO, 2006; OJALA et al., 1996; SWAIN; BALLARD, 1991; IRONS; PETERSEN, 1981; SAGHRI; FREEMAN, 1981; HARALICK et al., 1973) uma grande diversidade de extratores de características e funções de

distância. Além de técnicas que analisam o inter-relacionamento entre as funções de distância e os extratores de características, selecionando os mais adequados a serem integrados e aplicados a um domínio específico.

Pesquisadores têm explorado as áreas de classificação (MORENTE-MOLINERA et al., 2016; BATISTA et al., 2003; BREIMAN, 2001; DIETTERICH, 1998; KOHAVI et al., 1995; QUINLAN, 1993; AHA et al., 1991; RIJSBERGEN, 1979) e recuperação de imagens por conteúdo (AMATO et al., 2016; YANG et al., 2015; KASAUDHAN et al., 2015; WAN et al., 2014a; BUGATTI, 2012; ZHA et al., 2011; WU et al., 2010; ZHANG et al., 2009; HOI et al., 2009; HE et al., 2007; TONG; CHANG, 2001a), propondo técnicas de aprendizado de máquina e realimentação de relevância para amenizar alguns problemas intrínsecos dessas áreas. Como por exemplo, reduzindo problemas como o *gap* semântico em sistemas CBIR, minimizando a difícil tarefa do especialista de ajuste manual de valores de parâmetros de baixo nível na especificação das consultas, bem como o refinamento das mesmas, envolvendo, de forma transparente, combinações complexas de parâmetros.

Considerando que: (i) existe uma grande quantidade de imagens não rotuladas disponíveis; (ii) o processo de anotação é crucial para aprendizado efetivo da informação relacionada às imagens; (iii) a anotação manual é inviável em grandes conjuntos de dados, torna-se essencial o desenvolvimento de métodos robustos para lidar com essas questões.

Neste contexto, técnicas de aprendizado ativo têm sido propostas na área de *classificação*, para selecionar, a partir de um grande conjunto de dados, uma pequena quantidade de amostras mais informativas para o aprendizado, com uma sugestão de rótulo que pode ser confirmado ou corrigido por um especialista. Em se tratando da área de *recuperação de imagens por conteúdo*, métodos embasados em realimentação de relevância envolvem o aprendizado a partir da informação fornecida pela interação do usuário com o sistema. Trata-se de uma estratégia de aprendizado em tempo real (*on the fly*), em que ocorre o refinamento da consulta inicial realizada por determinado especialista, visando torná-la mais precisa, retornando somente respostas relevantes e descartando as irrelevantes.

Neste sentido, trabalhos na literatura têm explorado a combinação de técnicas de aprendizado ativo e técnicas de recuperação por conteúdo. Apesar de alguns esforços, ainda existem necessidades de otimizações, de forma a prover melhorias no processo de aprendizado e de recuperação das imagens. Portanto, o presente trabalho visa diminuir essa lacuna, propondo técnicas mais efetivas e eficientes para integração e viabilizando o desenvolvimento de um arcabouço robusto para classificação e recuperação de imagens por conteúdo.

1.1 OBJETIVOS

1.1.1 OBJETIVO GERAL

Este trabalho tem o objetivo de realizar o estudo, o desenvolvimento e a validação de técnicas efetivas e eficientes de aprendizado ativo para classificação e recuperação de imagens, de forma a acelerar os processos de aprendizado e recuperação, obtendo amostras mais relevantes, bem como minimizar a interação do especialista nestes processos.

1.1.2 OBJETIVOS ESPECÍFICOS

- Aquisição e seleção de conjuntos de imagens de diferentes domínios de aplicação;
- Avaliação de descritores (forma, cor e textura) apropriados para as aplicações de interesse;
- Desenvolvimento de estratégias de seleção de amostras mais informativas ao aprendizado;
- Comparação das estratégias propostas, utilizando diferentes técnica de agrupamento e de classificação, abordagens do estado da arte e conjuntos de imagens selecionados;
- Análise e utilização de técnicas CBIR com realimentação de relevância;
- Investigação de viabilidade de integração das técnicas de seleção propostas juntamente com as técnicas CBIR com realimentação de relevância;
- Desenvolvimento do arcabouço para classificação e recuperação de imagens por conteúdo.

1.2 ORGANIZAÇÃO DO TEXTO

Este trabalho apresenta a seguinte organização: no Capítulo 2 são introduzidos os principais conceitos e trabalhos envolvidos nos processos de CBIR e de aprendizado de máquina, bem como os principais trabalhos relacionados à integração de técnicas CBIR e de aprendizado ativo apresentados na literatura. No Capítulo 3 são abordadas as principais contribuições realizadas ao longo desta pesquisa direcionadas para a tarefa de classificação, as quais incluem uma metodologia para avaliação de desempenho do aprendizado de classificadores e uma estratégia de aprendizado ativo. No Capítulo 4 são apresentadas as contribuições na área CBIR, as quais incluem uma estratégia para avaliação de desempenho de descritores, bem como duas estratégias de aprendizado ativo para realimentação de relevância. Por fim, no Capítulo 5, são apresentadas as considerações finais, bem como os trabalhos futuros.

2 FUNDAMENTAÇÃO TEÓRICA

Neste capítulo serão apresentados os principais conceitos, necessários para entendimento do presente trabalho. Os conceitos encontram-se divididos em três tópicos: recuperação de imagens baseada em conteúdo (Seção 2.1), aprendizado de máquina (Seção 2.2), combinação de recuperação de imagens e aprendizado ativo (Seção 2.3).

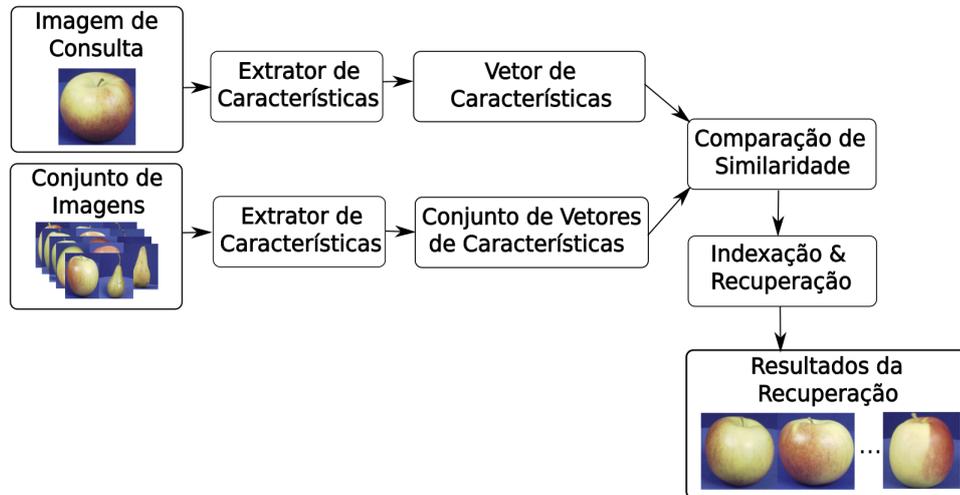
2.1 RECUPERAÇÃO DE IMAGENS BASEADA EM CONTEÚDO

O processo de recuperação de imagens baseada em conteúdo (*Content-Based Image Retrieval* - CBIR), visa a recuperação fundamentada na similaridade (ou dissimilaridade) entre as imagens de um conjunto e uma dada imagem de consulta. Para computar tal similaridade, geralmente, características de baixo nível, baseadas em cor, textura e/ou forma, são extraídas das imagens (MÜLLER et al., 2004).

O processo básico do CBIR é ilustrado na Figura 1. Tem-se como entrada uma imagem de consulta e um conjunto de imagens, realiza-se a descrição das imagens por meio de suas características intrínsecas, utilizando um dado método de extração, como cor, textura ou forma (ou até mesmo uma combinação de vários métodos de extração) e avalia-se a similaridade (ou dissimilaridade) entre as amostras, tornando possível o retorno de amostras mais similares (menos dissimilares) a imagem de consulta.

Existem algumas possibilidades de consulta, cujas principais são: as consultas por abrangência (Rq - limitam a lista de imagens recuperadas segundo um dado valor máximo de dissimilaridade, de acordo com uma dada imagem de consulta) e as consultas aos k -vizinhos mais próximos (k -NN - limitam a busca dada uma quantidade k de imagens menos dissimilares à imagem de consulta). Um exemplo de consulta por abrangência seria algo como: “Recupere todas as cidades que estão até 50 km de Cornélio Procópio”. Já um exemplo de consulta aos k -vizinhos mais próximos seria: “Recupere as 5 cidades mais próximas a Cornélio Procópio”, sendo nesse caso, $k = 5$.

Figura 1: Processo de recuperação de imagens baseado em conteúdo.



Fonte: Autoria própria.

É importante salientar que apesar de atualmente o processo CBIR tenha ganhado grande atenção por parte da comunidade acadêmica ainda existem inúmeros problemas em aberto. Isso ocorre principalmente devido à enorme lacuna semântica (*semantic gap*) existente entre as características de baixo nível extraídas das imagens e a representação de dissimilaridade adotada pelo ser humano, a qual baseia-se em processos de percepção intrínsecos e subjetivos. Dessa forma, o processo CBIR tende a retornar imagens divergentes à esperada pela percepção humana.

2.1.1 DESCRITORES

Para viabilizar a recuperação de imagens por conteúdo é necessária a avaliação e utilização de técnicas que melhor as descrevem. Os descritores de imagem consistem no par extrator de características e função de distância (TORRES; FALCAO, 2006). Existem na literatura inúmeras técnicas para extração de características e para avaliação de similaridade (KAWANO; YANAI, 2014; BUGATTI, 2012). As técnicas de extração de características podem ser divididas de acordo com as propriedades intrínsecas às imagens com base em textura, cor e forma.

Os extratores de textura podem mensurar algumas propriedades da imagem, como direcionalidade, orientação, regularidade, suavidade, entre outras. Esta é a medida do arranjo estrutural dos *pixels* em uma imagem, ou seja, são as diferenças locais nos níveis de intensidade. As abordagens baseadas em textura podem ser *top-down* (abordagem estrutural) ou *bottom-up* (abordagem estatística) (GONZALEZ; WOODS, 2010). Nesse contexto, diferentes extratores de textura podem ser aplicados, como por exemplo, *Local Binary Patterns* - LBP (OJALA et al.,

1996), *Local Activity Spectrum* - LAS (TAO; DICKINSON, 2000), Matrizes *Run-length* (GALLOWAY, 1975), Tamura (TAMURA et al., 1978), Gabor (ZHANG et al., 2000) e *Generates Pyramid Histograms of Oriented Gradients* - PHOG (BOSCH et al., 2007).

As estatísticas de primeira ordem, obtidas a partir de níveis de cinza/cores de uma imagem, podem também descrever uma textura (IRONS; PETERSEN, 1981; GUPTA; JAIN, 1997). Acredita-se que medidas estatísticas similares tenham texturas similares. Algumas medidas de primeira ordem são: média, desvio padrão, grau de assimetria, curtose, energia e entropia. Para obter os extratores de textura a partir das medidas de segunda ordem são considerados os relacionamentos entre os grupos de dois *pixels*. Para tanto é necessária a construção da matriz de coocorrência, a qual baseia-se no relacionamento espacial entre os *pixels* que compõem uma textura (GONZALEZ; WOODS, 2010). Com o objetivo de descrever as propriedades contidas nas texturas, Haralick et al. (1973) propõem 14 medidas estatísticas a serem calculadas a partir das matrizes de coocorrência.

Extratores de características baseados em cor também são amplamente utilizados na literatura, como por exemplo para resolver problemas de categorização de imagens naturais (ALMEIDA et al., 2008). Sendo o histograma, o qual descreve o conteúdo global de uma imagem de acordo com o percentual de *pixels* de cada cor, um dos extratores mais comumente utilizado no processo de extração Swain e Ballard (1991).

Alguns exemplos de extratores baseados em cor são: *Global Color Histogram* - GCH (PENATTI et al., 2012; YUE et al., 2011; ZHAO; GROSKY, 2002; JEONG, 2001) e *Local Color Histogram* - LCH (WEIJER; SCHMID, 2006; WANG, 2001; STEHLING et al., 2000; SETHI et al., 1997), *Border/Interior Classification* - BIC (STEHLING et al., 2002), dentre outras propostas (HUANG et al., 1997; CHATZICHRISTOFIS; BOUTALIS, 2008a; STRICKER; ORENGO, 1995; SMITH; CHANG, 1996; KRIEGEL et al., 2011; CHATZICHRISTOFIS; BOUTALIS, 2008b).

Os extratores de características baseados em forma consideram diferentes perceptivas da forma, com base em contorno ou em região Zhang e Lu (2004). Normalmente, as características são extraídas a partir de uma imagem binarizada. Sendo desejável que a extração seja invariante à rotação, escala e translação (TORRES; FALCAO, 2006).

Para os extratores de forma baseados em contorno pode-se citar algumas técnicas, tais como: assinatura (*Curvature Scale Space* - CSS) (GONZALEZ; WOODS, 2010), representação de cadeia de códigos (SAGHRI; FREEMAN, 1981), curvatura do espaço de escala (MOKHTARIAN; MACKWORTH, 1986) e o modelo elástico de forma (BIMBO; PALA, 1997). Para os extratores de forma baseados em região, pode-se citar as medidas globais (PEURA; IIVARI-

NEN, 1997) e os momentos (LEU, 1991).

Conforme mencionado, descritores têm um papel fundamental no processo de CBIR, sendo responsáveis por codificarem as propriedades das imagens. Sendo assim, além dos diferentes tipos de extratores de características apresentados, é importante analisar as funções de similaridade mais adequadas e que melhor descrevem as imagens (BUGATTI, 2012; TORRES; FALCAO, 2006).

Alguns exemplos de medidas de similaridade podem ser analisadas na Tabela 1, onde, x e y são as coordenadas dos *pixels*, n é o número de dimensões e $m_i = \frac{y_i + x_i}{2}$.

Tabela 1: Funções de Distância.

Nomes	Equações
L_1 (SOKAL; MICHENER, 1967)	$d_{L_1}(x, y) = \sum_{i=1}^n x_i - y_i $
L_2 (SNEATH et al., 1973)	$d_{L_2}(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$
L_∞ (KLOVE et al., 2010)	$d_{L_\infty}((x_1, \dots, x_n), (y_1, \dots, y_n)) = \max_{i=1}^n x_i - y_i $
X^2 (SKOPAL; BUSTOS, 2011)	$d_{X^2}(x, y) = \sum_{i=1}^n \frac{(x_i - m_i)^2}{m_i}$
<i>Canberra</i> (JURMAN et al., 2009)	$d_{Canberra}(x, y) = \sum_{i=1}^n \frac{ x_i - y_i }{ x_i + y_i }$
<i>Jeffrey Divergence</i> (KIM et al., 2010)	$d_{Jeffrey}(x, y) = \sum_{i=1}^n \left(y_i \log \frac{y_i}{m_i} + x_i \log \frac{x_i}{m_i} \right)$
$dLog$ (STEHLING et al., 2002)	$d_{dLog}(x, y) = \sum_{i=1}^n f(x_i) - f(y_i) $
	$f(m) = \begin{cases} 0 & \text{se } x = 0 \\ 1 & \text{se } 0 < x \leq 1 \\ \log_2 x + 1 & \text{senão} \end{cases}$

2.1.2 REALIMENTAÇÃO DE RELEVÂNCIA

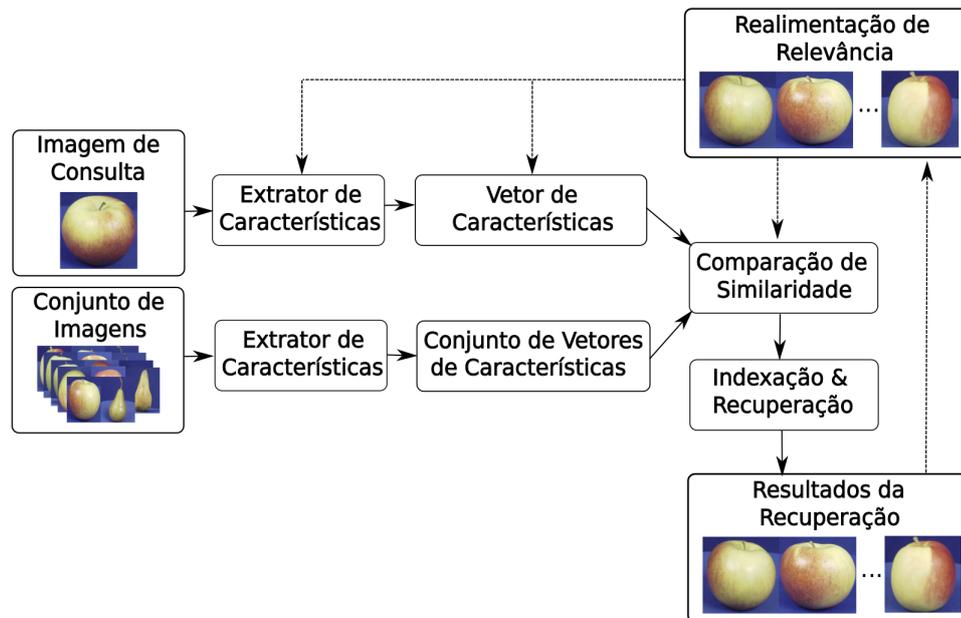
Um dos maiores problemas em CBIR refere-se ao *gap* semântico Wu et al. (2010), Hoi et al. (2009), Huiskes e Lew (2008), Zhou et al. (2006), He et al. (2004), Tong e Chang (2001a). Sendo assim, a realimentação de relevância foi introduzida para reduzir a distância que há entre o conceito semântico de alto nível das características da imagem de baixo nível, em relação à percepção do especialista. A utilização da realimentação de relevância e do conhecimento do especialista permitem uma atuação de forma ativa com resultados mais promissores, reduzindo o *gap* semântico.

A realimentação de relevância tem como princípio permitir que o especialista realize a anotação das imagens retornadas como relevantes ou irrelevantes a cada iteração do sistema,

possibilitando a recuperação de imagens mais similares de acordo com a intenção do mesmo (SILVA et al., 2010). Dessa forma, é possível obter uma personalização da consulta, de acordo com a percepção do especialista em questão.

Pode-se observar na Figura 2 o processo de CBIR com realimentação de relevância. Nesse processo tem-se como entrada inicial uma imagem de consulta, em que realiza-se a extração das características da imagem de consulta e das imagens do conjunto de imagens. O especialista realiza a indicação das imagens relevantes e irrelevantes retornadas pelo sistema, possibilitando o ajuste do aprendizado, de acordo com suas anotações. O processo é repetido até que uma dada condição de parada desejada seja obtida, como por exemplo, um número x de iterações ou até que o especialista esteja satisfeito com o resultado obtido.

Figura 2: Processo de CBIR com Realimentação de Relevância

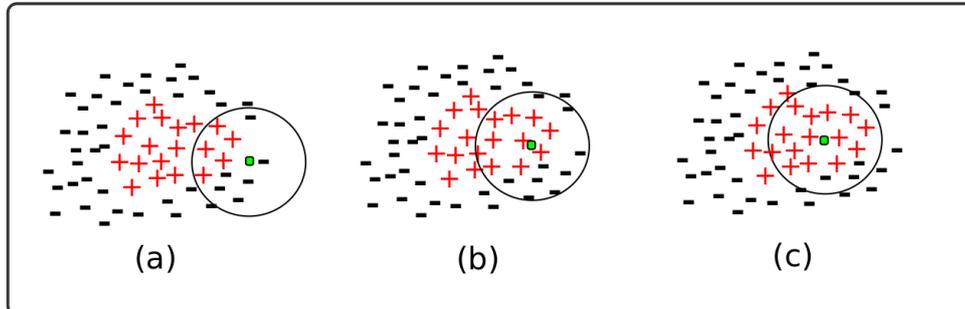


Fonte: Autoria própria.

Algumas técnicas são propostas na literatura para a realimentação de relevância em CBIR. Entre elas, o *Query Point Movement* - QPM (LIU et al., 2009), o qual movimentava o centro geométrico de busca de acordo com as amostras relevantes (positivas) retornadas a cada iteração. O centro geométrico pode ser calculado a partir da média das características das imagens retornadas. A Figura 3 ilustra o funcionamento da técnica QPM. Na Figura 3(a) é apresentada a consulta inicial, a partir de uma dada imagem de consulta (representada pelo quadrado), e as imagens retornadas (representadas pelos símbolos positivos e negativos delimitados pelo círculo). Após a consulta inicial, amostras relevantes e irrelevantes são retornadas e o especialista seleciona as amostras relevantes (amostras de interesse). Calcula-se o centro de massa das

amostras anotadas como relevantes e realiza-se nova busca a partir do centro de massa obtido (Figura 3(b)). O processo é repetido até que o especialista esteja satisfeito (Figura 3(c)).

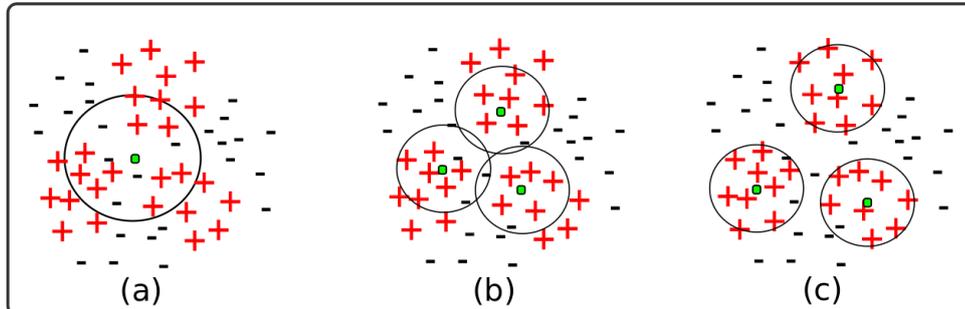
Figura 3: Técnica de realimentação de relevância *Query Point Movement* - QPM



Fonte: Autoria própria.

Outra técnica que move o centro geométrico de busca refere-se a *Query Expansion* - QEX (PORKAEW; CHAKRABARTI, 1999), a qual utiliza mais de um ponto de busca no processo. Este processo é apresentado na Figura 4.

Figura 4: Técnica de realimentação de relevância *Query Expansion* - QEX



Fonte: Autoria própria.

Na primeira iteração (Figura 4(a)), realiza-se o mesmo processo que apresentado pela técnica QPM. Após obtenção das mostras retornadas na primeira iteração, são utilizadas as amostras consideradas pelo especialista como relevantes para gerar k grupos. Para cada grupo formado é localizado o seu centro de massa (centroide). Após obter os centroides de cada grupo, localiza-se as n amostras mais próximas de cada centroide para serem avaliadas pelo especialista (Figura 4(b)). O processo é repetido até que o especialista esteja satisfeito (Figura 4(c)) com as imagens retornadas.

2.2 APRENDIZADO DE MÁQUINA

O aprendizado de máquina (AM), subárea da inteligência artificial (IA), tem como foco desenvolver sistemas capazes de aprender um determinado padrão ou comportamento automaticamente, a partir de exemplos ou observações. Os métodos utilizados em AM visam a automatização de novos conhecimentos para evolução do aprendizado e/ou a reorganização do conhecimento já existente (MITCHELL et al., 1997; MONARD; BARANAUSKAS, 2003).

Nesse contexto, diferentes abordagens de aprendizado não supervisionadas e supervisionadas podem ser consideradas. No aprendizado não supervisionado Jain et al. (1999), objetiva-se encontrar no espaço de características m -dimensional, agrupamentos (*clusters*) de dados de acordo com determinados critérios. Sendo assim, um dado *cluster* agrupa amostras que sejam similares entre si e dissimilares em relação a amostras que pertençam a outros *clusters* (LINDEN, 2009). É possível descrever as características inerentes de cada um dos *clusters* encontrados no processo de agrupamento, fornecendo um melhor entendimento do conjunto de dados e possibilitando a classificação de novos dados.

Na literatura, existem várias técnicas de agrupamento. Os algoritmos baseados em métodos particionais dividem a base de dados em k *clusters*, de acordo com uma medida de proximidade adotada, sendo k um valor informado pelo usuário. Cada amostra é agrupada segundo a minimização das distâncias entre os amostras e/ou centro de cada *cluster*. O k -means (JAIN et al., 1999), algoritmo amplamente conhecido, é utilizado no presente trabalho, como uma etapa de pré-processamento para posterior etapa de aprendizado de outros algoritmos de classificação (Subseção 3.2, 4.2.1 e 4.2.2)

No aprendizado supervisionado, o modelo de aprendizado é gerado por meio de um conjunto de (treinamento de) dados previamente rotulado. Várias propostas de algoritmos de aprendizado supervisionados são encontradas na literatura, dentre elas: *k-Nearest Neighbor - k-NN* (AHA et al., 1991), *Árvore de Decisão J48* (WITTEN et al., 1999; QUINLAN, 1993), *Random Forest - RF* (BREIMAN, 2001), *Support Vector Machines - SVM* (JOACHIMS, 1999), *Optimum-Path Forest - OPF* (PAPA et al., 2012), *Naive Bayes - NB* (BAYES, 1763). Tais classificadores supervisionados foram considerados no presente trabalho, por apresentarem bom desempenho e possuírem ampla utilização em diversos domínios de aplicação.

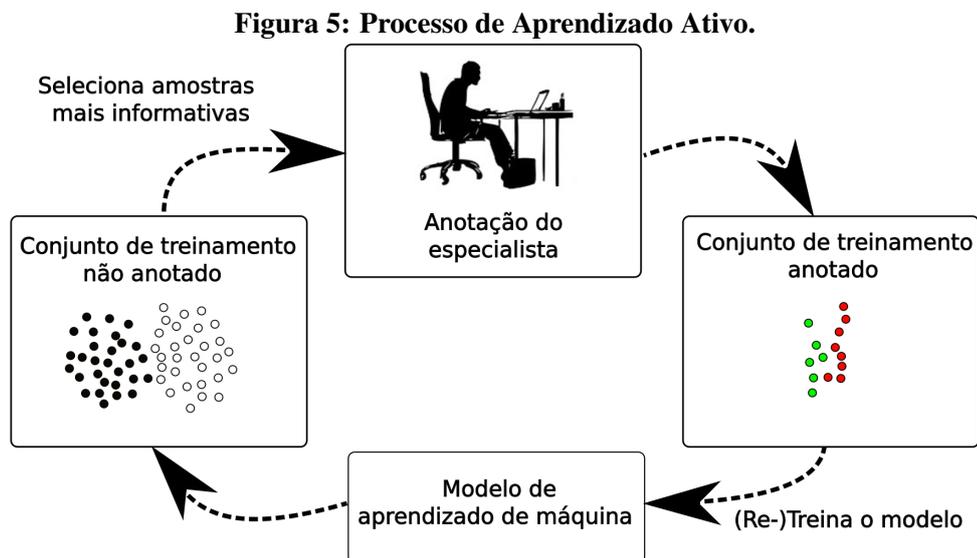
No entanto, alguns desafios são encontrados no processo de aprendizado supervisionado, tais como: desbalanceamento das amostras em cada uma das classes na base de dados; presença de ruídos na base, como dados imperfeitos ou dados fora da distribuição normal esperada (*outliers*); ajuste excessivo dos dados (*overfitting*) ou generalização insuficiente dos dados

(*underfitting*), ausência de valores (*missing values*) para determinadas características (DUDA et al., 2012). Tais desafios dependem diretamente dos dados utilizados no treinamento do algoritmo de aprendizado. Sendo assim, é fundamental selecionar a quantidade adequada de amostras, bem como analisar sua qualidade e relevância para o aprendizado, conforme discutido na Seção 2.2.1.

2.2.1 APRENDIZADO ATIVO

O processo de aprendizado ativo é iterativo, em que o classificador participa ativamente do seu processo de aprendizado, auxiliando, juntamente com o critério de seleção adotado, na seleção das amostras mais informativas a serem consideradas no processo. A Figura 5 ilustra o processo de aprendizado ativo.

Após a seleção do conjunto de amostras mais informativas em uma dada iteração, tal conjunto é exibido a um especialista ou outra fonte de informação (oráculo), para confirmação dos rótulos fornecidos pelo classificador. Posteriormente, após a confirmação/correção dos rótulos, as amostras são incorporadas ao conjunto de treinamento, sendo criada uma nova instância do classificador. O processo se repete até que o especialista esteja satisfeito com o desempenho do classificador.



Fonte: Adaptado de (SETTLES, 2009)

No entanto, Settles (2009) apresenta algumas problemáticas no processo de aprendizado ativo, tais como: i-) como selecionar as amostras mais informativas? ii-) como garantir a qualidade da anotação fornecida pelo especialista?

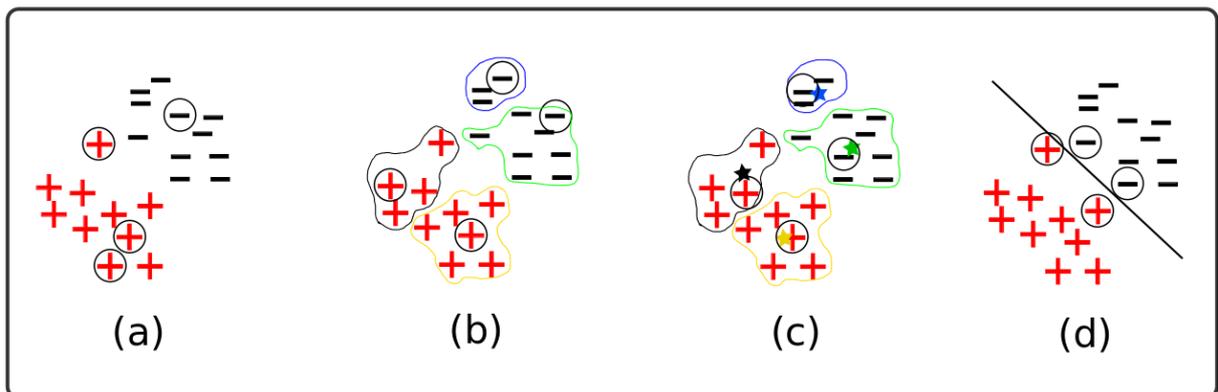
As amostras mais informativas podem ser obtidas, de acordo com alguns critérios de seleção a serem adotados, com base em incerteza (amostras mais difíceis de serem classificadas), representatividade (amostras representativas de cada uma das classes), diversidade (amostras de classes distintas), dentre outros critérios.

Em relação à qualidade da anotação fornecida pelo especialista, uma alternativa para aumentar a qualidade seria considerar mais de um especialista no processo. No entanto, não é trivial obter a disponibilidade de mais de um especialista. Uma solução seria selecionar um conjunto pequeno de amostras mais significativas (informativas) para solicitar a anotação dos especialistas. Portanto, o critério de seleção das amostras mais informativas é fundamental, para acelerar o processo de aprendizado e minimizar o esforço e a interação do especialista no processo de anotação.

2.2.1.1 ESTRATÉGIAS DE SELEÇÃO

As heurísticas utilizadas no processo de aprendizado ativo consistem em como localizar as amostras mais significativas para o treinamento de um classificador. Para realizar tal tarefa encontram-se na literatura diversas estratégias (SETTLES, 2012). A Figura 6 ilustra 4 exemplos de estratégias amplamente utilizadas.

Figura 6: Exemplos de estratégias de seleção, com base em critérios como: (a) aleatoriedade, (b) diversidade, (c) representatividade e (d) incerteza.



Fonte: Autoria própria.

A estratégia mais simples consiste na seleção *aleatória* (Figura 6 (a)) das amostras do conjunto. Técnicas de agrupamento também têm sido aplicadas (ALAJLAN et al., 2014; SETTLES, 2012; KANG et al., 2004; NGUYEN; SMEULDERS, 2004), adotando critérios de seleção com base em *diversidade* (Figura 6 (b)). A partir do agrupamento das amostras do conjunto em k grupos, um conjunto de amostras pode ser selecionado aleatoriamente a partir de

cada um dos grupos, de forma a obter amostras diversas (as quais podem constituir amostras de classes distintas, contribuindo melhor para o aprendizado do classificador). Alajlan et al. (2014) apresenta comparações entre essas estratégias, nomeadas como HR e HCLU, respectivamente. Além de obter amostras aleatórias a partir de cada grupo, outros critérios podem ser adotados com base em *representatividade* (Figura 6 (c)), obtendo amostras representativas de cada um dos grupos, com o intuito de obter amostras de cada uma das classes do conjunto. Além disso, o agrupamento pode ser realizado iterativamente, de forma a obter melhores refinamentos a cada iteração do aprendizado.

Outras estratégias de seleção exploram critérios com base em *incerteza* (Figura 6 (d)), apresentada por Settles (2012), Alajlan et al. (2014), propõem a seleção das amostras próximas ao limiar de decisão do classificador, demonstrando que são as amostras (mais difíceis) com maior indecisão e, conseqüentemente, proporcionando maior benefício para o aprendizado do modelo. Várias propostas são encontradas na literatura, dentre elas, utilizando cálculos de probabilidade (LEWIS; GALE, 1994), comitês de modelos treinados em conjuntos rotulados que apresentam hipóteses concorrentes (SEUNG et al., 1992; MUSLEA et al., 2000), classificador SVM, o qual utiliza métricas para localizar as amostras mais próximas do hiperplano criado pelo classificador (TONG; KOLLER, 2001; WANG, 2018).

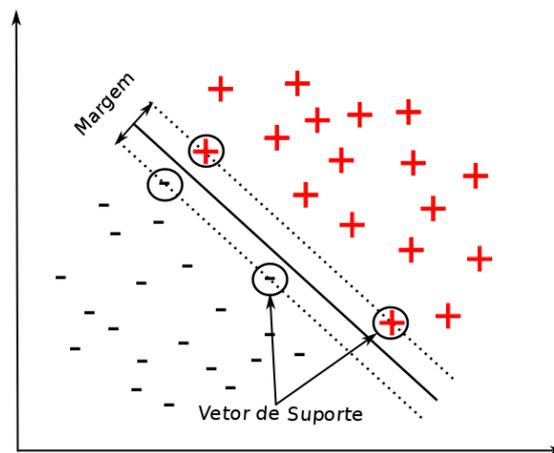
2.3 COMBINAÇÃO DE APRENDIZADO ATIVO E CBIR

Nas últimas décadas muitos autores propuseram a combinação de aprendizado ativo com CBIR no processo de realimentação de relevância na tentativa de diminuir o problema do *gap* semântico (RAO et al., 2018; SILVA et al., 2011, 2011; ZHA et al., 2011; WU et al., 2010; HOI et al., 2009; ZHANG et al., 2009; HE et al., 2004; TONG; CHANG, 2001a). Uma das primeiras abordagens de aprendizado ativo aplicado ao processo CBIR foi com o auxílio do classificador Bayesiano, utilizando a probabilidade na escolha das imagens mais informativas para treinamento do classificador (COX et al., 2000; KING; JIN, 2003).

Uma das estratégias mais utilizadas é baseada no classificador SVM. Tong e Chang (2001b) propuseram o uso do classificador SVM para localização do hiperplano, separando as classes relevantes e irrelevantes em um espaço multidimensional (Figura 7). A cada iteração, as imagens mais próximas ao hiperplano são apresentadas ao especialista para anotação. Por fim, as imagens (relevantes) mais afastadas do hiperplano apresentadas para o especialista, como sendo as imagens mais relevantes em relação à imagem de consulta.

Trabalhos recentes que utilizam SVM buscam otimizar esse método (PUTZU et al.,

Figura 7: Exemplo de separação de classes utilizando o classificador SVM.



Fonte: Autoria própria.

2018; RAO et al., 2018; WU et al., 2010; HOI et al., 2009; HE et al., 2004). No entanto, a proposta ainda é considerada *baseline* para comparação com novas abordagens (TAVARES, 2011), sendo assim considerada no presente trabalho.

Vale ressaltar que outras propostas com outros classificadores também são encontradas na literatura, tais como, baseada em floresta de caminhos ótimos (SILVA et al., 2011), baseadas em redes neurais artificiais (RAO; REDDY, 2018; ANH et al., 2010), assim como baseada na ideia de que para cada amostra pode-se estimar seu rótulo com base em seus vizinhos, escolhendo as amostras que podem minimizar a variação máxima de predição para serem exibidas e anotadas pelo especialista (ZHA et al., 2011).

3 ABORDAGEM DE APRENDIZADO PARA CLASSIFICAÇÃO

Este capítulo aborda as principais contribuições realizadas ao longo desta pesquisa direcionadas para a tarefa de classificação. Para melhor organização e entendimento das abordagens desenvolvidas, dividiu-se esse capítulo em duas seções principais. A Seção 3.1 aborda uma metodologia (CAMARGO et al., 2017) para avaliação de desempenho do aprendizado de alguns classificadores em diferentes conjuntos biomédicos. A Seção 3.2.2 descreve uma proposta de aprendizado ativo para classificação de conjuntos biomédicos (BRESSAN et al., 2018c).

3.1 ESTRATÉGIA PARA AVALIAÇÃO DE DESEMPENHO DO APRENDIZADO

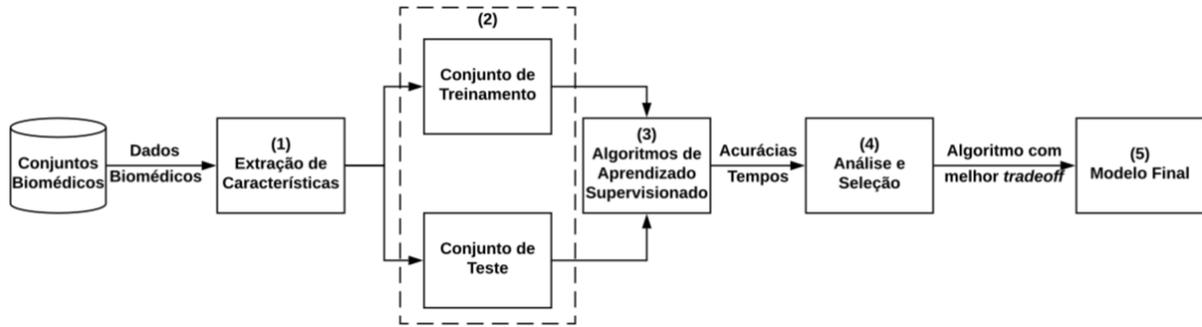
A análise da eficiência e da eficácia de um classificador não é uma tarefa trivial. Muitos parâmetros devem ser levados em consideração para obter um modelo apropriado para um conjunto de dados. Assim, para realizar uma análise refinada de desempenho de alguns classificadores, foi proposta uma metodologia de avaliação (CAMARGO et al., 2017) que leva em consideração regras de relação de compromisso na escolha do melhor modelo para conjuntos biomédicos.

A Figura 8 ilustra as etapas desenvolvidas para a avaliação de desempenho de diferentes classificadores em diferentes conjuntos biomédicos. Na Etapa 1, realizam-se as extrações das características de cada conjunto de dados.

Na Etapa 2, os conjuntos são divididos em dois subconjuntos (conjuntos de treinamento e conjuntos de teste). Nessa divisão, os conjuntos são balanceados, mantendo assim a mesma quantidade de amostras por classe em cada conjunto. Tal procedimento pode ser realizado n vezes.

Na Etapa 3, são utilizados os conjuntos de treinamento para aprendizado dos classificadores supervisionados. Após os treinamentos para cada divisão, os conjuntos de teste são submetidos à avaliação de cada classificador, obtendo assim as acurácias (Acc_{ij}) e tem-

Figura 8: Pipeline da estratégia de avaliação de desempenho de diferentes classificadores em diferentes conjuntos biomédicos.



Fonte: (CAMARGO et al., 2017)

pos computacionais (T_{ij}), onde i representa o i -ésimo conjunto de dados e j corresponde o j -ésimo modelo de aprendizado supervisionado. Os resultados de Acc_{ij} e de T_{ij} são então utilizados na Etapa 4 para análise de políticas de relação de compromisso, de forma a selecionar o melhor modelo de classificação supervisionado (M_i^Ω), em que Ω representa a melhor seleção. Para possibilitar a avaliação e comparação entre os resultados, duas listas são criadas, uma de acurácias médias (μAcc_i) e outra de tempo computacional médio (μT_i) obtidas a partir de cada classificador, considerando todas as divisões realizadas nos conjuntos de dados. O cálculo de compromisso entre tempo computacional e acurácia é dado pela média harmônica (Equação 1). Assim, quanto menor for o valor de H , melhor será o modelo de aprendizado (Equação 2).

$$H_{ij} = 2 \times \frac{\mu Acc_{ij} \times \mu T_{ij}}{\mu Acc_{ij} + \mu T_{ij}} \quad (1)$$

$$\Omega_i = \arg \min \{H_{ij}\} \quad (2)$$

Na etapa 5, o modelo final (M_i^Ω) é apresentado, sendo o modelo de aprendizado mais efetivo e eficiente dentre os classificadores avaliados neste contexto.

3.1.1 DESCRIÇÃO DOS CONJUNTOS DE DADOS

Para validação da metodologia proposta foram utilizados 9 conjuntos de dados biomédicos públicos, com diferentes quantidades de amostras, classes e características, os quais são descritos com mais detalhes a seguir.

- **ALL-AML Leukemia** (GOLUB et al., 1999): composto por 72 samples, 2 classes e 7.129

características. As classes correspondem às amostras de *Acute Lymphocytic Leukemia* (ALL) e *Acute Myeloid Leukemia* (AML) da medula óssea e sangue periférico.

- **Breast Cancer** (VEER et al., 2001): composto por 97 amostras, 2 classes e 24.481 características. As amostras representam pacientes que desenvolveram metástases em um intervalo de 5 anos (rotuladas como “*relapse*”) e pacientes que permaneceram saudáveis após o diagnóstico inicial por um intervalo de pelo menos 5 anos (rotulado como “*non-relapse*”).
- **Central Nervous System** (POMEROY et al., 2002): composto por 60 amostras, 2 classes e 7.129 características. É usado para analisar o resultado do tratamento do tumor embrionário do sistema nervoso central. As amostras são distribuídas em 21 sobreviventes (pacientes que estão vivos após o tratamento) e 39 fracassos (pacientes que sucumbiram à doença).
- **Colon Tumor** (ALON et al., 1999): composto de 62 amostras, 2 classes e 2.000 características. As amostras foram coletadas de pacientes com câncer de cólon, sendo 40 amostras de biópsias de tumores (rotulados como “*negative*”) e 22 amostras de biópsias normais, ou seja partes saudáveis (rotuladas como “*positive*”).
- **Lung Cancer** (GORDON et al., 2002): composto de 181 amostras, 2 classes e 12.533 características. É dividido em 31 amostras de mesotelioma maligno da pleura (MPM) do pulmão e 150 amostras de adenocarcinoma (ADCA) do pulmão.
- **MLL Leukemia** (ARMSTRONG et al., 2002): composto por 72 amostras, 3 classes e 12.582 características. As classes são organizadas em 24 amostras linfoblásticas agudas (LLA), 20 amostras de leucemias de linhagens mistas (MLL) e 28 amostras de leucemias mielogênicas agudas (AML).
- **Ovarian Cancer** (PETRICOIN et al., 2002): composto por 253 amostras, 2 classes e 15.154 características. Inclui 91 amostras que não contém câncer e 162 amostras de pessoas com câncer de ovário.
- **Prostate Cancer** (SINGH et al., 2002): composto por 136 amostras, 2 classes e 12.600 características. É dividido em 77 amostras rotuladas como “tumor” e 59 amostras de próstata não tumorais (rotuladas como “normais”).
- **Subtypes of Acute Lymphoblastic Leukemia** (YEOH et al., 2002): composto por 327 amostras, 7 classes e 12.558 características. Contém todos os subtipos de leucemia linfoblástica aguda conhecidos, incluindo células T (T-ALL), E2APBX1, TEL-AML1,

MLL, BCR-ABL e hiperdiplóides (hiperdiplóides > 50), bem como uma classe que contém amostras que não se encaixavam em nenhuma das classes mencionadas (rotuladas como “Others”).

3.1.2 DESCRIÇÃO DOS EXPERIMENTOS

Para avaliação experimental, cada conjunto foi dividido em 80% para o conjunto de treinamento e 20% para o conjunto de teste. Esse procedimento foi realizado 10 vezes, mantendo a distinção entre eles. Para resumir os resultados, abstraíram-se as acurácias médias obtidas por cada classificador, bem como o tempo computacional obtido para as etapas de treinamento e de teste.

A metodologia de avaliação proposta foi aplicada nos conjuntos biomédicos, apresentados em 3.1.1, utilizando os classificadores *k-Nearest Neighbors* - *k*-NN (AHA et al., 1991) considerando $k = 1$, *Support Vector Machines* - SVM (CSVM e NuSVM) (BOSER et al., 1992), *Random Forest* - RF (BREIMAN, 2001), *Optimum-Path Forest* - OPF (PAPA et al., 2012) e *Multilayer Perceptron* - MLP (MINSKY; PAPERT, 1969). Para realizar os experimentos considerando SVM e OPF foram utilizados libSVM (CHANG; LIN, 2011) e libOPF (PAPA et al., 2012), respectivamente, mantendo a configuração padrão de cada biblioteca.

Os experimentos foram realizados em um computador com processador Intel Core i3 de 2,3 GHz e 3 GB de RAM, com o uso do sistema operacional Ubuntu 14,04 de 64 bits.

3.1.3 RESULTADOS E DISCUSSÕES

As acurácias médias obtidas por cada classificador em cada conjunto podem ser verificadas na Tabela 2. Destacam-se em negrito os melhores resultados, i.e. os classificadores que apresentaram as melhores acurácias médias para cada conjunto de dados.

Observa-se que o classificador RF obteve as melhores acurácias médias, com exceção dos conjuntos de dados *AML-ALL* e *Ovarian Cancer*, em que o MLP e o NuSVM, respectivamente, apresentaram uma acurácia maior. Considerando o classificador RF, a melhor acurácia obtida foi de 99,07% para o conjunto de dados *Lung Cancer*.

Os piores resultados foram observados com as duas variações do SVM (CSVM e NuSVM). Esses resultados estão relacionados à ausência de otimização dos parâmetros do SVM. Nesse trabalho não foi aplicada técnicas de otimização, tal como *gridsearch*, devido a restrições de tempo, uma vez que as avaliações dos classificadores têm como foco não apenas a eficácia, mas também a eficiência dos mesmos.

Tabela 2: Acurácias médias obtidas por cada classificador em cada conjunto de dados (CAMARGO et al., 2017).

	<i>k</i> -NN	CSVM	NuSVM	RF	OPF	MLP
AML-ALL	92,88	64,28	64,28	95,24	91,48	97,61
Breast Cancer	87,71	56,14	56,14	89,47	51,85	66,66
Central Nervous	86,11	66,66	66,66	97,22	60,41	69,44
Colon Tumor	88,89	66,66	66,66	97,22	72,92	78,55
Lung Cancer	95,37	83,33	83,33	99,07	96,66	83,33
MLL	91,11	40,00	40,00	95,55	94,89	77,77
Ovarian Cancer	98,00	86,66	100,00	98,66	91,03	97,33
Prostate Cancer	93,82	55,55	55,55	97,53	76,94	58,02
Subtypes of ALL	89,90	24,24	24,24	92,92	87,84	24,24

Pode-se observar na Tabela 3 o tempo médio em segundos para o treinamento de cada classificador, considerando cada conjunto de dados.

Tabela 3: Tempos médios de treinamento obtidos por cada classificador em cada conjunto de dados (CAMARGO et al., 2017).

	<i>k</i> -NN	CSVM	NuSVM	RF	OPF	MLP
AML-ALL	—	14,70	13,22	3,91	1,22	697,52
Breast Cancer	—	8,60	8,19	3,54	0,19	266,64
Central Nervous	—	1,52	1,54	1,09	0,02	18,12
Colon Tumor	—	0,43	0,49	0,70	0,01	4,74
Lung Cancer	—	9,13	7,78	1,99	0,39	154,21
MLL	—	3,00	3,08	1,72	0,06	70,72
Ovarian Cancer	—	16,51	12,88	3,19	0,93	322,80
Prostate Cancer	—	6,13	5,85	2,15	0,22	117,84
Subtypes of ALL	—	1,90	1,71	1,16	0,04	21,66

O classificador OPF obteve melhores resultados (menores tempos) de treinamento em todos os conjuntos avaliados. Nota-se que o OPF obteve um menor tempo de treinamento (0,01 segundos) no conjunto de dados *Colon Tumor*.

O classificador que atingiu o segundo melhor tempo foi o RF, com valores entre 0,69 e 3,91 segundos. O classificador que necessitou de mais tempo para treinamento foi o MLP, com valores entre 4,74 e 697,52 segundos. Tal resultado é esperado, dado que o mesmo requer várias iterações para regulagens dos pesos sinápticos.

Vale ressaltar que o classificador *k*-NN não apresenta valores de tempo de treinamento, dado que não gera um modelo de aprendizagem.

A avaliação de tempo médio obtido para o teste de cada classificador em cada conjunto de dados pode ser observada na Tabela 4. Nota-se novamente que o classificador OPF obteve menor tempo em todos os conjuntos de dados e o segundo melhor foi o classificador RF. O

classificador que obteve o pior desempenho foi o k -NN, obtendo valores entre 0,11 a 8,04 segundos.

Tabela 4: Tempos médios de teste obtidos por cada classificador em cada conjunto de dados (CAMARGO et al., 2017).

	k -NN	CSVM	NuSVM	RF	OPF	MLP
AML-ALL	8,04	2,58	2,75	0,99	0,24	1,50
Breast Cancer	2,06	2,98	2,85	0,61	0,04	0,83
Central Nervous	0,31	0,42	0,34	0,20	0,01	0,21
Colon Tumor	0,11	0,11	0,12	0,08	0,00	0,08
Lung Cancer	2,76	2,51	2,57	0,63	0,07	0,76
MLL	0,71	1,06	0,79	0,36	0,01	0,33
Ovarian Cancer	4,61	5,47	4,46	0,87	0,14	1,11
Prostate Cancer	1,28	1,97	2,19	0,48	0,04	0,55
Subtypes of ALL	0,34	0,40	0,39	0,13	0,00	0,16

A Tabela 5 apresenta o desempenho médio obtido por cada classificador, considerando valores de acurácia e tempos de treinamento e de teste em todos os conjuntos avaliados. Ao avaliar a eficácia (ou seja, acurácia) e eficiência (ou seja, restrições de tempo de treinamento e teste), o classificador OPF alcançou um bom equilíbrio entre eficiência e eficácia, atingindo até 80,44% de acurácia média e valores entre 0,34 e 0,06 segundos para treinamento e teste, respectivamente.

Tabela 5: Desempenho médio para cada classificador. (CAMARGO et al., 2017).

	k -NN	CSVM	NuSVM	RF	OPF	MLP
Acurácia	91,53	60,39	61,88	95,88	80,45	72,55
Tempo treinamento	—	6,88	6,08	2,16	0,34	186,03
Tempo teste	2,25	1,94	1,83	0,48	0,06	0,61

Nota-se também que o classificador RF atingiu notáveis resultados em termos de eficiência e eficácia (Tabela 5), resultando em uma acurácia média de 95,87% (acurácia maior em relação à obtida pelo OPF). No entanto, atingiu 2,16 e 0,48 segundos para treinamento e teste, respectivamente. O classificador k -NN apresentou uma boa acurácia média, mas apresentou o pior tempo de teste (2,24 segundos).

A partir dos desempenhos obtidos por cada classificador em cada conjunto de dados, geraram-se as médias harmônicas que podem ser analisadas na Tabela 6.

Como já mencionado na Seção 3.1, quanto menor o valor das médias harmônicas melhor será o modelo de aprendizado. Assim, ao analisarmos a Tabela 6, percebe-se que o classificador OPF apresentou melhores médias harmônicas em todos os conjuntos de dados. É possível observar que o classificador k -NN obteve, em 9 conjuntos de dados, melhores resultados quando

Tabela 6: Média harmônica obtida para cada classificador em cada conjunto de dados (CAMARGO et al., 2017).

	<i>k</i> -NN	CSVM	NuSVM	RF	OPF	MLP
AML-ALL	14,80	27,24	25,57	9,30	2,83	171,28
Breast Cancer	4,02	19,18	18,45	7,93	0,46	106,72
Central Nervous	0,60	3,75	3,62	2,53	0,05	29,00
Colon Tumor	0,22	1,07	1,21	1,53	0,02	9,06
Lung Cancer	5,35	20,41	18,41	5,10	0,88	108,38
MLL	1,41	7,35	7,04	4,03	0,12	74,26
Ovarian Cancer	8,81	35,05	29,54	7,80	2,10	149,68
Prostate Cancer	2,53	14,12	14,03	5,12	0,48	77,87
Subtypes of ALL	0,68	4,20	3,83	2,54	0,07	25,46

comparado com o classificador RF. Por exemplo, considerando o conjunto de dados *Breast Cancer*, *k*-NN apresentou uma média harmônica de até 4,03 em relação a 7,93 apresentada pelo RF. Analisando o conjunto *AML-ALL*, RF apresentou uma média harmônica de até 9,30 em relação a 14,79 obtida por meio do classificador *k*-NN.

A partir dos resultados obtidos, percebe-se que o classificador OPF é o modelo mais eficiente e eficaz, considerando a metodologia de avaliação de desempenho proposta baseada em análise das estratégias que fornecem um melhor *trade-off* entre acurácias e restrições de tempo. Sendo assim, o OPF pode ser útil em aplicações que requerem classificadores mais rápidos, considerando os tempos de aprendizado e teste apresentados pelo mesmo. Além disso, os classificadores *k*-NN e RF são opções a serem consideradas.

Analisando apenas as acurácias obtidas, o classificador RF apresenta-se como a melhor escolha. No entanto, a escalabilidade do classificador deve ser analisada e considerada, de forma a possibilitar a manipulação do crescimento acelerado dos dados (por exemplo, imagens médicas, genes, entre outros) gerados diariamente, especialmente em dados biomédicos. Além disso, algumas aplicações podem apresentar determinadas restrições de tempo e recursos computacionais. Portanto, a metodologia proposta apresenta-se como uma contribuição valiosa, a medida que possibilita analisar o desempenho do aprendizado, levando-se em consideração regras de relação de compromisso para escolha do melhor modelo de classificação.

3.2 ESTRATÉGIA DE APRENDIZADO ATIVO

Conforme mencionado, uma grande quantidade de dados biomédicos são gerados e têm aumentado exponencialmente. Dessa forma, automatizar e otimizar a análise e classificação desses grandes conjuntos de dados torna-se uma tarefa extremamente importante.

Em geral, a classificação e o diagnóstico baseados na análise de dados biomédicos exigem uma anotação precisa, para representar o conteúdo de interesse dos dados e treinar um modelo de classificação.

No entanto, em problemas reais, é comum ter que lidar com conjuntos de dados que apresentam apenas algumas amostras rotuladas. Isso ocorre devido aos elevados esforços e custos envolvidos no processo de anotação manual, uma vez que especialistas precisam executar tal tarefa. Outra desvantagem está relacionada às inconsistências que podem ocorrer nas anotações fornecidas pelos especialistas, pois trata-se de um processo cansativo que leva à fadiga gerada pela análise de uma grande quantidade de amostras em um período de tempo longo.

Apesar de alguns esforços (SAITO et al., 2015; CAMARGO et al., 2017) lidarem com restrições relacionadas a eficácia e eficiência requeridas por determinadas aplicações, ainda assim é necessária uma grande quantidade de dados anotados para treinamento do classificador. Nesse sentido, deve-se estabelecer um novo *trade-off* entre a necessidade de uma quantidade considerável de dados anotados para obter elevadas acurácias e o esforço e custo da anotação realizada por um especialista. Nesse caso, estratégias de aprendizado ativo podem ser exploradas para lidar com esse *trade-off*.

Como já mencionado, as técnicas de aprendizado ativo permitem que um pequeno conjunto de amostras de aprendizado não anotado seja selecionado e exibido, iterativamente, para anotações de especialistas. Em seguida, o conjunto anotado é utilizado para treinar o classificador. Durante o processo de aprendizado, o classificador participa ativamente, selecionando as amostras mais informativas (SETTLES, 2009).

Portanto, propõe-se o uso de estratégias de aprendizado ativo para alavancar o processo de aprendizagem. Além disso, um melhor *trade-off* pode ser obtido, uma vez que um pequeno subconjunto de amostras é utilizado no processo de aprendizado do classificador (menor custo e esforço de anotação pelos especialistas) e maiores acurácias podem ser alcançadas (uso das amostras mais informativas). Com base em tais propriedades, a presente seção estende os estudos realizados na seção anterior (Seção 3.1), propondo e descrevendo duas novas contribuições.

Inicialmente, estratégias de aprendizado ativo são inseridas ao processo de aprendizado do classificador, o qual na medida em que aprende, seleciona e sugere rótulos para a anotação das amostras que são mais difíceis de classificação, aliviando o esforço de anotação manual de todo o conjunto de dados fornecido pelo especialista, conforme requerido no processo de aprendizado supervisionado tradicional.

Além disso, uma nova estratégia de aprendizado ativo é proposta, capaz de explorar

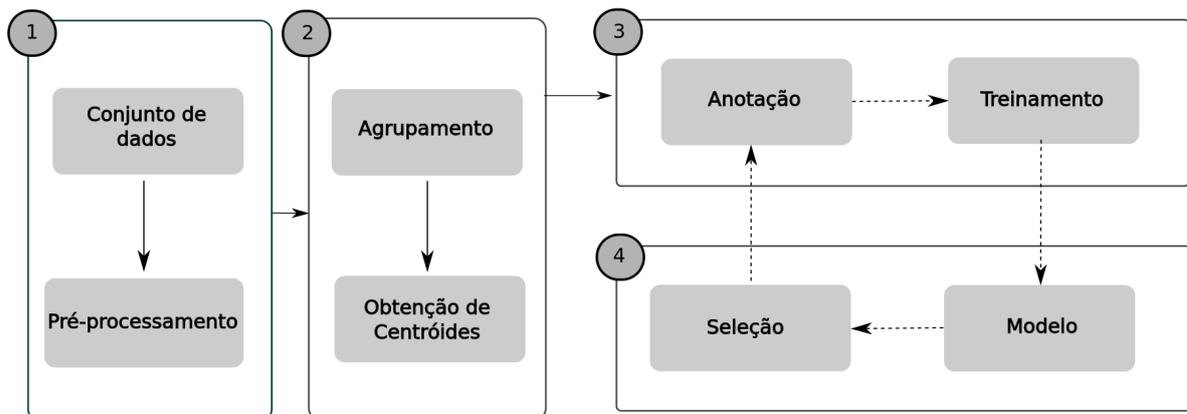
melhor as amostras mais informativas para o aprendizado do classificador, por meio dos critérios de seleção propostos com base na representatividade e na incerteza das amostras. A estratégia de seleção é crucial para atingir elevadas acurácias mais rapidamente (ou seja, nas primeiras iterações de aprendizado).

Para validação da estratégia proposta foi realizada uma avaliação extensiva, de forma a realizar comparações entre estratégias de aprendizado da literatura (descritos na Seção 3.2.2), bem como utilizando diferentes técnicas de classificação (Seção 3.2.2), em diferentes conjuntos de dados biomédicos públicos (apresentados na Seção 3.2.1). Os resultados dos experimentos empíricos (apresentados na Seção 3.2.3) mostram que a estratégia proposta, em muitos casos, supera estratégias de aprendizado da literatura.

A estratégia posposta (BREISSAN et al., 2018c) tem como premissa minimizar o envolvimento de especialistas, sem perder o controle sobre o processo de aprendizado do classificador. O esforço e o custo da anotação são reduzidos, uma vez que: i-) um pequeno subconjunto de amostras mais informativas requer anotação; ii-) o especialista precisa apenas corrigir os rótulos classificados erroneamente, uma vez que o classificador fornece previamente um rótulo para as amostras. Além disso, iii-) como o classificador aprende, ocorrem menos erros de classificação e, portanto, o especialista realiza menos correções, ao longo das iterações do aprendizado.

A abordagem proposta de aprendizado ativo para classificação é dividida em 4 etapas principais. A Figura 9 ilustra o *pipeline* da abordagem proposta, bem como cada etapa do processo de aprendizado. O Algoritmo 1 descreve detalhadamente a abordagem proposta.

Figura 9: Pipeline da abordagem de aprendizado ativo proposta para classificação.



Fonte: (BREISSAN et al., 2018c)

Na Etapa 1, alimenta-se o sistema com um conjunto de dados Z , que é então normalizado na etapa de pré-processamento por uma dada técnica (por exemplo, min-max, z-score,

Algoritmo 1: Abordagem proposta para classificação

Entrada : conjunto de dados Z , número N de amostras selecionadas em cada iteração, número k de clusters desejados para o agrupamento.

Saida : modelo final de aprendizado M^Ω .

auxiliares: conjunto de aprendizado Z_2 , conjunto de raízes R , conjunto temporário de treinamento Z'_1 , conjunto de treinamento Z_1 , e instâncias do modelo de aprendizado M .

- 1 $Z_2 \leftarrow \text{preProcessamento}(Z)$;
- 2 $\text{agrupamento}(Z_2, k)$;
- 3 $R \leftarrow \text{obtencaoRaizes}(Z_2, k)$;
- 4 $Z'_1 \leftarrow R$;
- 5 $Z_1 \leftarrow \text{anotacao}(Z'_1)$;
- 6 $M \leftarrow \text{treinamento}(Z_1)$;
- 7 **repeat**
- 8 $Z'_1 \leftarrow \text{selecao}(Z_2 \setminus Z_1, N, M)$;
- 9 $Z_1 \leftarrow Z_1 \cup \text{anotacao}(Z'_1)$;
- 10 $M \leftarrow \text{reTreinamento}(Z_1, M)$;
- 11 **until** *satisfeito*;
- 12 $M_\Omega \leftarrow M$;

entre outras) (Linha 1). Em seguida, o conjunto Z é dividido em dois subconjuntos, um conjunto de aprendizado Z_2 e um conjunto de testes Z_3 . O conjunto de testes Z_3 é utilizado apenas para avaliar o desempenho do classificador em um conjunto de dados desconhecido durante o processo de aprendizado. Além disso, inicialmente, dois conjuntos vazios também são criados Z_1 e Z'_1 . O conjunto Z_1 será utilizado para treinar o classificador incrementalmente ao longo das iterações e o conjunto Z'_1 será utilizado como um conjunto de treinamento temporário.

Na Etapa 2, após o pré-processamento, as amostras de Z_2 são agrupadas com uma determinada abordagem de agrupamento (Linha 2). Uma lista R é criada, recebendo as k amostras mais próximas de cada centróide (amostras raízes) a partir de Z_2 (Linha 3). O conjunto de treinamento auxiliar Z'_1 armazena temporariamente as amostras contidas em R (Linha 4).

Na Etapa 3, cada amostra de Z'_1 é apresentada ao especialista para ser anotada. As amostras anotadas constituem o conjunto de treinamento Z_1 (Linha 5). Após o processo de treinamento, na Etapa 4, uma instância M do modelo de aprendizado é criada (Linha 6). Tal modelo participa ativamente do processo de seleção das amostras mais informativas.

Os critérios de seleção (Linha 8) baseados na representatividade e na incerteza das amostras são aplicados, obtendo o conjunto temporário Z'_1 de N amostras de treinamento, de acordo com o atual modelo de aprendizado M e as amostras restantes em Z_2 ($Z_2 \cap Z_1 = \emptyset$, em uma notação simplificada $Z_2 \setminus Z_1$).

O conjunto de amostras selecionado Z'_1 é exibido ao especialista (Linha 9). A partir da primeira iteração de aprendizado, tal conjunto de amostras selecionadas são rotuladas previamente pela instância atual do classificador. Sendo assim, o especialista precisa apenas corrigir os rótulos de amostras erroneamente classificadas. As amostras confirmadas e corrigidas apropriadamente pelo especialista são adicionadas ao conjunto de treinamento anterior Z_1 . O treinamento é executado novamente e uma nova instância do classificador M é gerada (Linha 10). As Etapas 3 e 4 são repetidas até que se esteja satisfeito com o processo de aprendizado.

Os critérios de seleção propostos são descritos no Algoritmo 2. Inicialmente, um conjunto de amostras candidatas SC_i ($0 < i \leq nc$) é criado para receber as amostras candidatas a serem exibidas para o especialista, em que i denota a classe i e nc o número de classes (Linha 1).

Conforme já mencionado, a ideia é explorar melhor o conhecimento do classificador. Sendo assim, o conjunto de aprendizado $Z_2 \setminus Z_1$ é classificado pelo modelo de aprendizado atual (M) (Linha 2), para avaliar quais amostras são mais informativas para o seu próprio aprendizado.

Após a classificação, as amostras são separadas de acordo com os rótulos das classes fornecidos pelo modelo. Cada amostra $s \in Z_2 \setminus Z_1$ é armazenada em uma lista de rótulos (L_i) correspondente ao seu respectivo rótulo da classe i ($i = s.labelid$). A amostra (1-NN) mais próxima à amostra s , nomeada como amostra adjacente ($adjs$), é avaliada. Se os rótulos das amostras s e $adjs$, fornecidos pelo classificador, são diferentes (ou seja, $s.labelid \neq adjs.labelid$), essas amostras constituem amostras informativas, pois podem ser de classes distintas. Tais amostras são armazenadas em uma lista de candidatas SC_i , correspondente ao respectivo rótulo de classe i (Linhas 3 a 10).

Caso existam amostras armazenadas nas listas de candidatas SC_i , os centros de massa com_i do espaço de características das amostras de cada classe i são localizados e armazenados em sua lista correspondente de centros de massa $Lcom_i$. Para identificar cada centro de massa com_i , um cálculo simples é realizado considerando a média das amostras da classe i . Posteriormente, as amostras candidatas são organizadas em suas respectivas listas SC_i , em ordem decrescente, de acordo com as distâncias entre as amostras e os centros de massa com_i a partir de $Lcom_i$ (Linhas 11 a 14).

No entanto, devido à distribuição do espaço de características, a técnica proposta pode não ser satisfeita, ou seja, $SC_i = \emptyset$, dado que não há $s.labelid \neq adjs.labelid$. Para resolver este cenário, calcula-se o centro de massa com_i a partir de cada lista de amostras de aprendizado L_i , sendo armazenados em suas respectivas listas de centros de massa $Lcom_i$. Com essas informações, as amostras candidatas SC_i são organizadas em ordem decrescente, da mesma

Algoritmo 2: Estratégia de Seleção

Entrada : conjunto de aprendizado $Z_2 \setminus Z_1$, numero N de amostras selecionadas em cada iteração para anotação do usuário, instância atual M do modelo de aprendizado.

Saída : conjunto SS de amostras selecionadas.

Auxiliares: conjuntos SC_i de amostras candidatas, listas de aprendizado L_i organizadas com base em rótulos, amostra vizinha mais próxima $adjs$, e listas $Lcom_i$ de amostras correspondentes ao centro de massa de cada rótulo.

```

1  $SC_i \leftarrow \emptyset, i = 1, 2, \dots, nc;$ 
2  $classificacao(Z_2 \setminus Z_1, M);$ 
3 for each  $s \in Z_2 \setminus Z_1$  do
4    $L_i \leftarrow L_i \cup \{s\}, i = s.labelid;$ 
5    $adjs \leftarrow obtencao1NN(s);$ 
6   if  $s.labelid \neq adjs.labelid$  then
7      $SC_i \leftarrow SC_i \cup \{s\}, i = s.labelid;$ 
8      $SC_i \leftarrow SC_i \cup \{adjs\}, i = adjs.labelid;$ 
9   end
10 end
11 if  $SC_i \neq \emptyset$  then
12    $Lcom_i \leftarrow calculoCentroMassa(SC_i), i = 1, 2, \dots, nc;$ 
13    $SC_i \leftarrow organizacao(SC_i, Lcom_i), i = 1, 2, \dots, nc;$ 
14 end
15 else
16    $Lcom_i \leftarrow obtencaoCentroMassa(L_i), i = 1, 2, \dots, nc;$ 
17    $SC_i \leftarrow organizacao(L_i, Lcom_i), i = 1, 2, \dots, nc;$ 
18 end
19 while  $i \leq N, i = 1, 2, \dots, nc$  do
20    $SS \leftarrow SS \cup \{s\}, s \in SC_i,$  seguindo a ordem dada por  $SC_i;$ 
21 end

```

forma que as mencionadas anteriormente, no entanto a partir de amostras de L_i (Linhas 15 a 18). Finalmente, a estratégia de seleção retorna o conjunto de amostras selecionado SS , composto de N amostras mais informativas (representativas e incertas), correspondentes às primeiras amostras obtidas a partir de cada conjunto de amostras candidatas SC_i (Linhas 19 a 21).

A estratégia de seleção proposta possibilita obter um conjunto de amostras mais representativas e incertas, uma vez que tais amostras são selecionadas com base nas classificações fornecidas pelo classificador. Além disso, os critérios de seleção estão relacionados aos centros de massa a partir de diferentes rótulos das classes (representatividade), bem como levam em consideração as amostras mais próximas que podem ser de classes diferentes (incerteza).

3.2.1 DESCRIÇÃO DOS CONJUNTOS DE DADOS

Para avaliação da estratégia proposta foram utilizados 6 conjuntos de dados biomédicos públicos. Informações como quantidades de amostras, classes, características e descrições mais detalhadas sobre esses conjuntos de dados são descritas a seguir.

- ***Breast Cancer*** (VEER et al., 2001): composto por 97 amostras, 2 classes e 24.481 características. As amostras representam pacientes que desenvolveram metástases em um intervalo de 5 anos (rotuladas como “relapse”) e pacientes que permaneceram saudáveis após o diagnóstico inicial por um intervalo de pelo menos 5 anos (rotulado como “non-relapse”).
- ***Central Nervous System*** (POMEROY et al., 2002): composto por 60 amostras, 2 classes e 7.129 características. É usado para analisar o resultado do tratamento do tumor embrionário do sistema nervoso central. As amostras são distribuídas em 21 sobreviventes (pacientes que estão vivos após o tratamento) e 39 fracassos (pacientes que sucumbiram à doença).
- ***Colon Tumor*** (ALON et al., 1999): composto de 62 amostras, 2 classes e 2.000 características. As amostras foram coletadas de pacientes com câncer de cólon, sendo 40 amostras de biópsias de tumores (rotulados como “*negative*”) e 22 amostras de biópsias normais, ou seja partes saudáveis (rotuladas como “*positive*”).
- ***Lung Cancer*** (GORDON et al., 2002): composto de 181 amostras, 2 classes e 12.533 características. É dividido em 31 amostras de mesotelioma maligno da pleura (MPM) do pulmão e 150 amostras de adenocarcinoma (ADCA) do pulmão.
- ***MLL Leukemia*** (ARMSTRONG et al., 2002): composto por 72 amostras, 3 classes e 12.582 características. As classes são organizadas em 24 amostras linfoblásticas agudas (LLA), 20 amostras de leucemia de linhagens mistas (MLL) e 28 amostras de leucemias mielogênicas agudas (AML).
- ***Subtypes of Acute Lymphoblastic Leukemia*** (YEOH et al., 2002): composto por 327 amostras, 7 classes e 12.558 características. Contém todos os subtipos de leucemia linfoblástica aguda conhecidos, incluindo células T (T-ALL), E2APBX1, TEL-AML1, MLL, BCR-ABL e hiperdiploides (hiperdiploides > 50), bem como uma classe que contém amostras que não se encaixavam em nenhuma das classes mencionadas (rotuladas como “Others”).

3.2.2 DESCRIÇÃO DOS EXPERIMENTOS

Na avaliação experimental, cada conjunto de dados (descrito em 3.2.1) foi dividido em dois subconjuntos, um conjunto de aprendizado contendo 80% das amostras e um conjunto de teste contendo 20% das amostras. Este procedimento de divisão foi realizado 10 vezes de maneira estratificada. Foram gerados 10 conjuntos diferentes de treinamento e teste com interseção vazia entre eles e divididos proporcionalmente de acordo com o número de amostras em cada classe. Para todos os experimentos, a mesma divisão de cada conjunto de dados foi mantida, permitindo uma comparação justa entre os métodos.

Para validação da estratégia proposta para classificação, denominada como ourAL-1 (Algoritmo 1), foram realizadas comparações, utilizando dois métodos da literatura como baselines HCLU e HR (ALAJLAN et al., 2014) (descritos na Seção 2.2.1.1).

Diferentes técnicas de classificação supervisionada também foram analisadas, a fim de avaliar e comparar o desempenho dos métodos utilizando cada um deles. Nesse caso, foram considerados os classificadores: *Support Vector Machines* - SVM (JOACHIMS, 1999) com o kernel baseado em funções polinomiais, *k-Nearest Neighbor* - *k*-NN (AHA et al., 1991) utilizando $k = 1$, *Naive Bayes* - NB (BAYES, 1763), *Optimum-Path Forest* - OPF (PAPA et al., 2012) e *Random Forest* - RF (BREIMAN, 2001).

Para resumir os resultados, foi utilizada a média das acurácias obtidas pelos classificadores, a cada iteração de aprendizado, de cada método de aprendizado ativo. Utilizou-se também a configuração padronizada para todos os experimentos, onde na etapa de pré-processamento utilizou-se a normalização dos dados com MIN-MAX, na etapa de agrupamento o algoritmo *k*-means (JAIN et al., 1999), sendo o valor de *k* igual a quantidade de classes existentes no conjunto de dados. Para as *N* amostras exibidas a cada iteração para serem anotadas pelo especialista também utilizou-se o mesmo valor *k* do agrupamento.

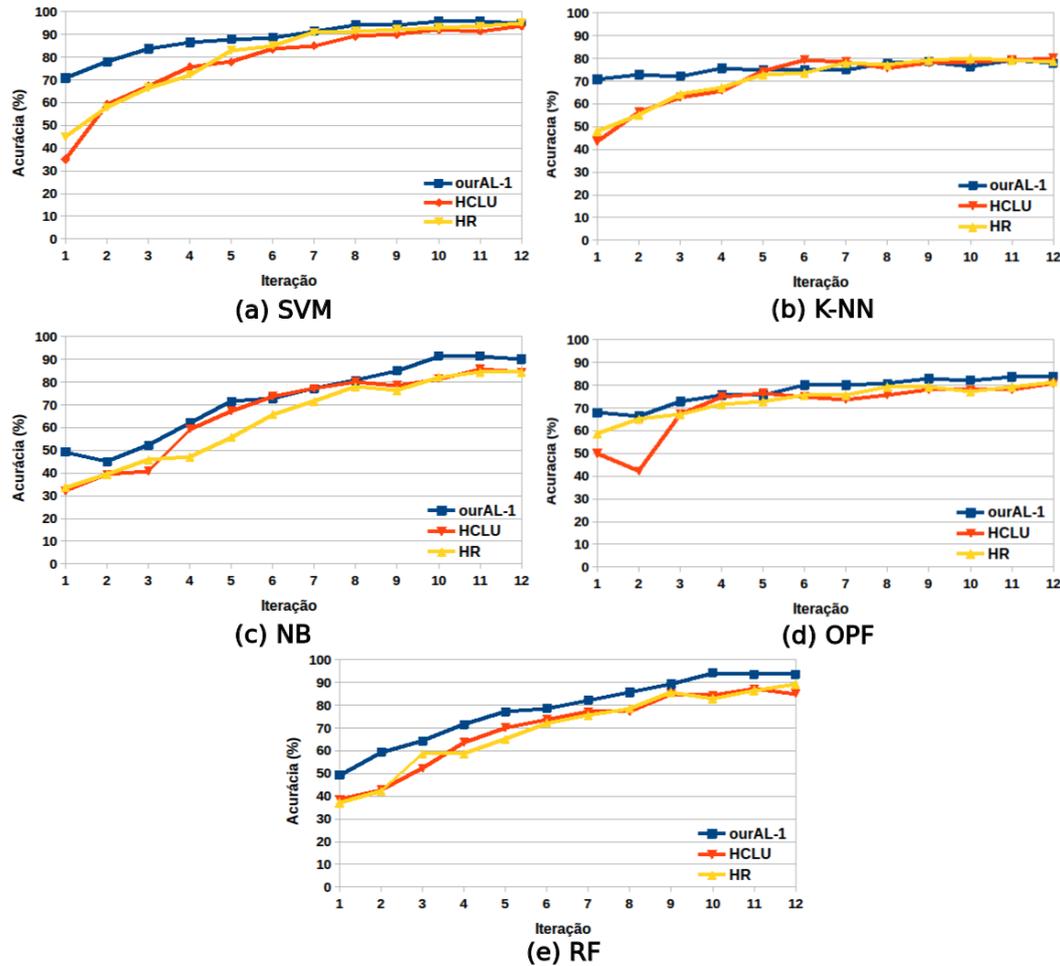
3.2.3 RESULTADOS E DISCUSSÕES

Pode-se observar nas Figuras 10 a 15 as acurácias médias obtidas a cada iteração pelos métodos de aprendizado (ourAL-1, HCLU e HR) em cada conjunto de dados (*MLL Leukemia*, *Breast Cancer*, *Central Nervous System*, *Colon Tumor*, *Lung Cancer*, *Subtypes of Acute Lymphoblastic Leukemia*, respectivamente).

Para facilitar o entendimento e análise dos dados, o eixo *x* representa cada iteração de aprendizado e o eixo *y* as acurácias obtidas. As letras de (a)-(e) representam os classificadores utilizados no processo de aprendizado, sendo: SVM, *k*-NN, NB, OPF e RF, respectivamente.

Considerando o conjunto de dados *MLL Leukemia* (Figura 10), de forma geral, a técnica proposta ourAL-1 apresentou um melhor desempenho, atingindo acurácias elevadas mais rapidamente, em relação às estratégias de aprendizado ativo HCLU e HR, utilizando cada um dos classificadores.

Figura 10: Acurácia média por iteração dos métodos de aprendizado ativo no conjunto de dados *MLL Leukemia* utilizando diferentes classificadores.



Fonte: Autoria própria.

Por exemplo, para o classificador SVM (Figura 10(a)), a partir da primeira até a quinta iteração, pode-se observar ganhos da estratégia ourAL-1 em relação à estratégia HCLU de 102,04%, 31,33%, 24,47%, 14,15% e 12,84%, respectivamente. Em relação à estratégia HR, os ganhos de ourAL-1 foram de 57,14%, 34,57%, 25,81%, 19,80% e 6,03%.

Utilizando o *k*-NN (Figura 10(b)), os ganhos apresentados pela estratégia ourAL-1 em relação à HCLU foram de 62,30%, 29,11%, 14,77%, 15,22% e 0,96%, considerando da primeira a quinta iteração, respectivamente. Em relação à HR, os ganhos foram de 47,76%, 32,47%, 12,22%, 12,77% e 2,94%.

Na Figura 10(c) os ganhos obtidos pela ourAL-1 utilizando o classificador NB foram de 53,33%, 14,55%, 28,07%, 4,82% e 6,38% (primeira a quinta iteração) em relação à HCLU e de 46,81%, 14,55%, 14,06%, 31,82% e 28,21% em comparação à HR.

Para o classificador OPF (Figura 10(d)), os ganhos de ourAL-1 em comparação à HCLU foram de 35,71%, 57,63%, 8,51% e 0,95%, considerando da primeira a quarta iteração. Os ganhos apresentados pela ourAL-1 em relação à HR foram de 15,85%, 2,20%, 8,51%, 6,00% e 3,92% considerando da primeira até a quinta iteração.

Por fim, a técnica ourAL-1 com o uso classificador RF (Figura 10(e)) obteve ganhos em todas as iterações, entre 38,33% a 10,20% considerando da primeira a quinta iteração em relação à HCLU, bem como valores de ganhos entre 40,68% a 21,95% em relação à HR.

A Figura 11 ilustra os resultados considerando o conjunto de dados *Breast Cancer*. Neste conjunto de dados, percebe-se que, desde as primeiras iterações, a abordagem proposta ourAL-1 obteve ganhos com os classificadores SVM e OPF quando comparados com HCLU e HR. Por exemplo, ao analisar a Figura 11(a), ourAL-1 obteve 65,79% de acurácia na terceira iteração, utilizando SVM. Enquanto HCLU e HR obtiveram 53,95% e 57,89% de acurácia.

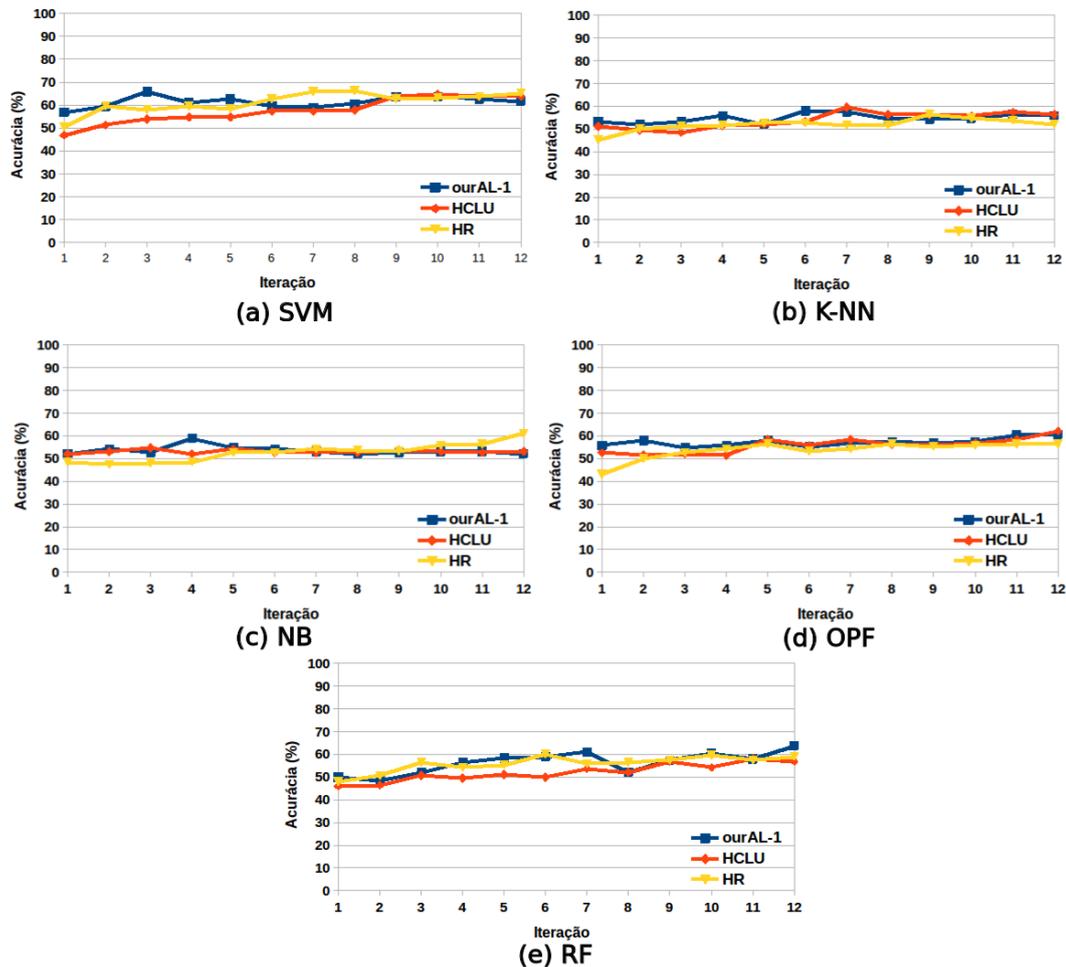
Em alguns casos, as estratégias atingiram certa estabilidade mais rapidamente, apresentando resultados equivalentes, como por exemplo, utilizando os classificadores NB e RF (Figuras 11(c) e 11(e)). No entanto, vale destacar que, em alguns cenários (especialmente Figuras 11(a), 11(b) e 11(d)), ourAL-1 apresenta um melhor desempenho, analisando as primeiras iterações, as quais são consideradas mais relevantes dado que ao longo das iterações do aprendizado todas as estratégias tendem a atingir certa estabilidade. A estratégia proposta ourAL-1 apresenta ganhos em relação às demais estratégias, atingindo acurácias elevadas mais rapidamente, com menos iterações e interações de aprendizado.

Para o conjunto de dados *Central Nervous System* (Figura 12), a estratégia proposta ourAL-1 apresenta, de forma geral, ganhos em relação às estratégias HCLU e HR, utilizando cada um dos classificadores.

A estratégia ourAL-1 apresenta ganhos em relação ao HCLU: de até 51,11% e 17,65% (na primeira iteração) utilizando, respectivamente, os classificadores OPF e RF; 45,28% (na segunda iteração) utilizando k -NN e, de até 14,29% e 9,59% (na terceira iteração) utilizando SVM e NB, respectivamente. Os ganhos de ourAL-1 em relação à HR foram: de até 47,83% e 25% (na primeira iteração) utilizando, respectivamente, OPF e RF; 14,93% (na segunda iteração) utilizando k -NN e, de até 23,08% e 9,59% utilizando SVM e NB, respectivamente.

Vale também destacar que ourAL-1 atinge acurácias elevadas mais rapidamente em

Figura 11: Acurácia média por iteração dos métodos de aprendizado ativo no conjunto de dados *Breast Cancer* utilizando diferentes classificadores.



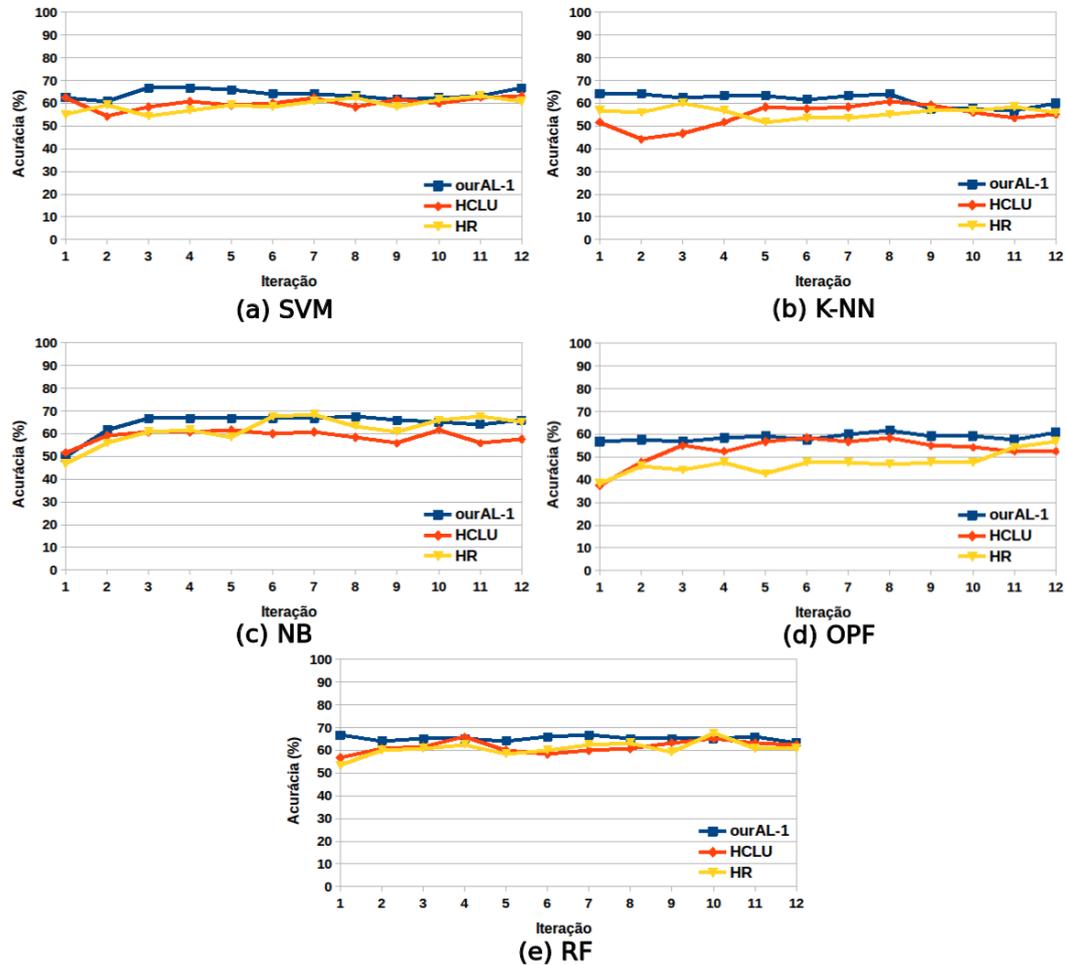
Fonte: Autoria própria.

relação às demais estratégias. Pode-se observar, por exemplo, na Figura 12 (e), que ourAL-1 atinge 66,67% de acurácia na primeira iteração, enquanto que HCLU e HR atingem tal valor de acurácias apenas após a oitava iteração.

Analisando os resultados obtidos para o conjunto *Colon Tumor* (Figura 13), ourAL-1 apresenta ganhos utilizando os classificadores k -NN, OPF e RF (Figuras 13(b), 13(d) e 13(e), respectivamente).

A estratégia proposta ourAL-1 possibilita a seleção de amostras de fato mais informativas, obtendo acurácias elevadas mais rapidamente, com menos iterações e interações de aprendizado. Por exemplo, utilizando o classificador k -NN (Figura 13(b)), ourAL-1 atinge 78% de acurácia na sexta iteração, enquanto as demais técnicas atingem valores de acurácias equivalentes apenas ao final (contendo todas as amostras) do aprendizado. O mesmo comportamento

Figura 12: Acurácia média por iteração dos métodos de aprendizado ativo no conjunto de dados *Central Nervous System* utilizando diferentes classificadores.



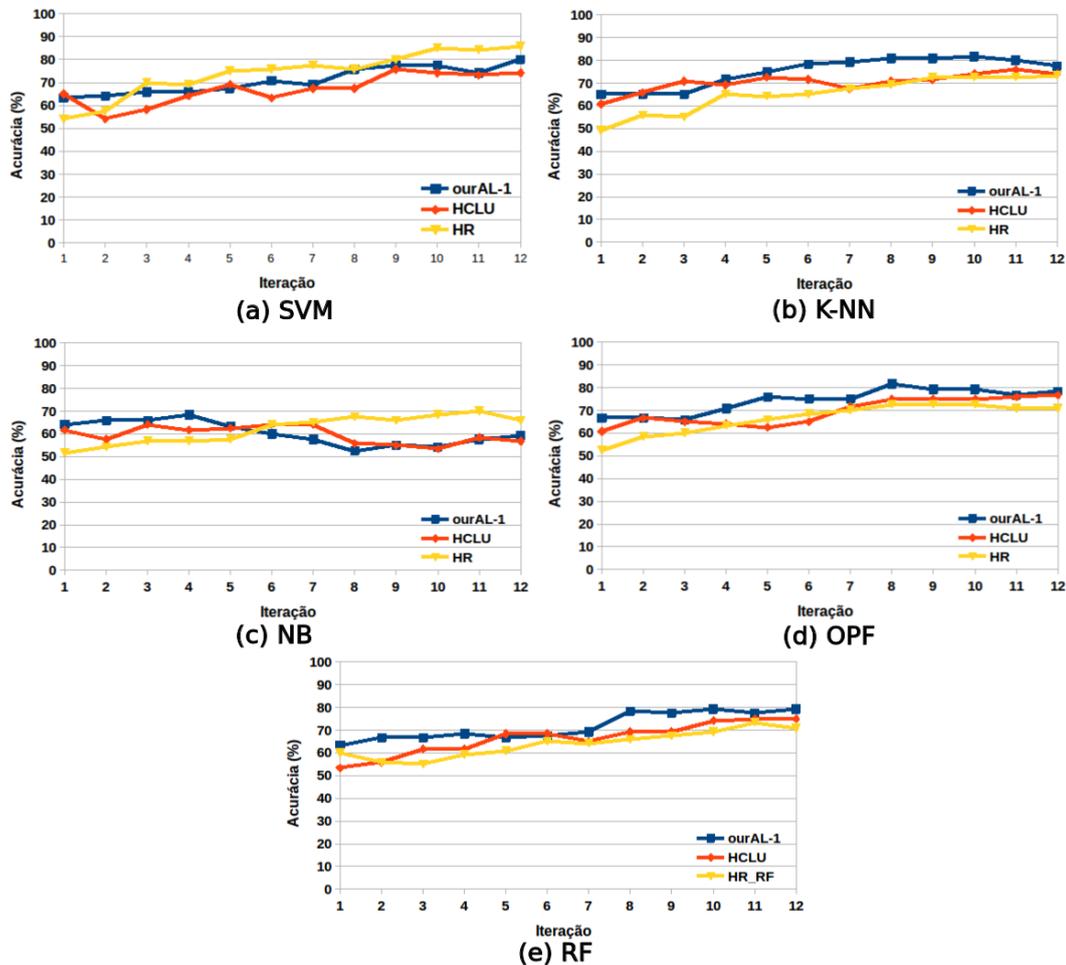
Fonte: Autoria própria.

pode ser observado com a utilização dos classificadores OPF e RF (Figuras 13(d) e 13(e), respectivamente).

Para o conjunto *Lung Cancer* (Figura 14), pode-se observar comportamento semelhante, em que ourAL-1 apresenta, de forma geral, melhores acurácias desde as primeiras iterações, utilizando cada um dos classificadores. Por exemplo, na Figura 14(d), utilizando o classificador OPF, ourAL-1 atinge 81,89% de acurácia já na primeira iteração do aprendizado, enquanto HCLU e HR atingem 73,78% e 23,24%, respectivamente.

Por fim, considerando o conjunto de dados *Subtypes of Acute Lymphoblastic Leukemia* (Figura 15), observa-se que a estratégia ourAL-1 apresenta desempenho superior em relação às demais estratégias (HCLU e HR) utilizando o classificador *k*-NN (Figura 15(b)) em todas as iterações do aprendizado. Além disso, ourAL-1 apresenta acurácias maiores nas primeiras iterações, utilizando os classificadores SVM e NB (Figuras 15(a) e 15(c)). Considerando os

Figura 13: Acurácia média por iteração dos métodos de aprendizado ativo no conjunto de dados *Colon Tumor* utilizando diferentes classificadores.

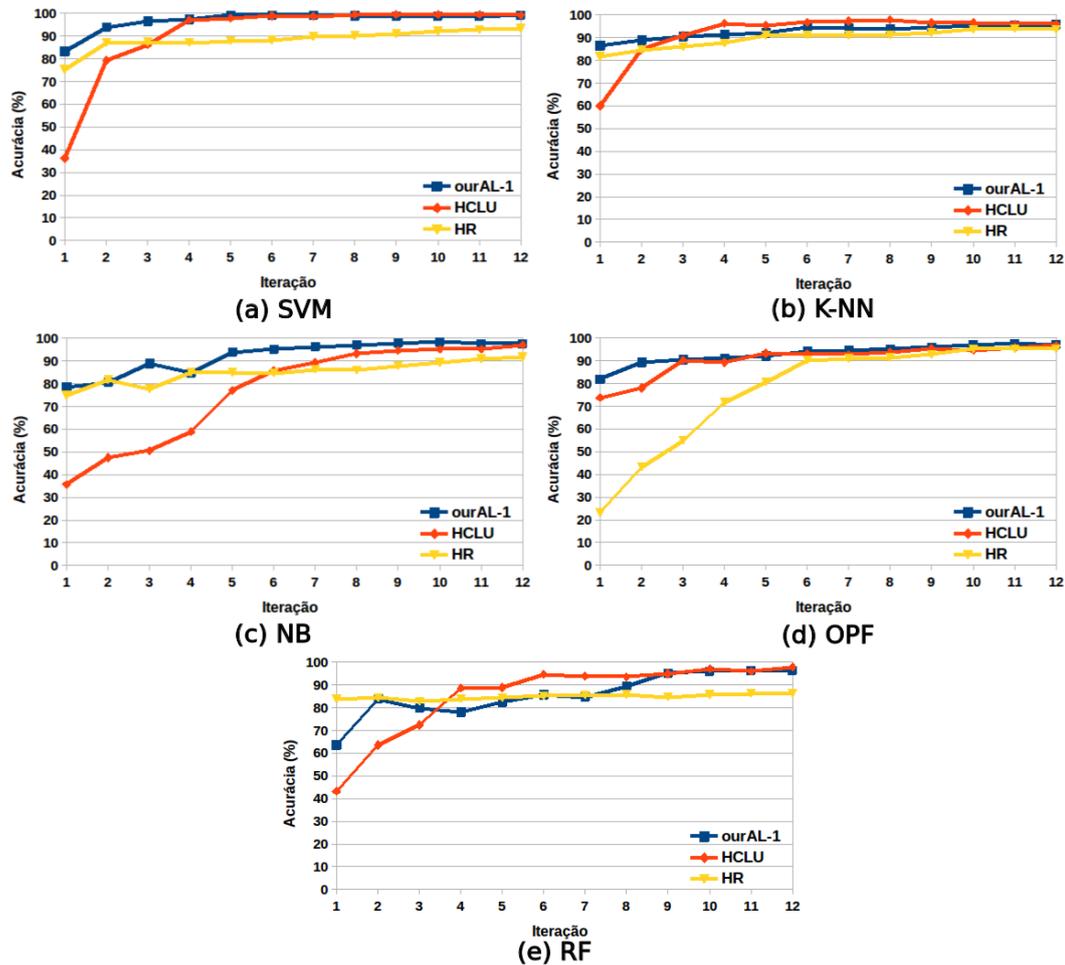


Fonte: Autoria própria.

classificadores OPF e RF, ourAL-1 e HCLU apresentam resultados equivalentes. Já a estratégia HR apresentou os piores valores de acurácia ao longo das iterações de aprendizado, utilizando cada um dos classificadores. Acredita-se que os resultados estão relacionados com o desbalanceamento de amostras em cada uma das (7) classes do conjunto de dados, aliada à forma de seleção aleatória de amostras a cada iteração do aprendizado.

Na Tabela 7, pode-se observar um resumo dos resultados obtidos pelas técnicas de aprendizado, utilizando cada um dos classificadores em cada conjunto de dados. Os valores apresentados são gerados a partir da média das acurácias considerando as 12 primeiras iterações. Nota-se que a estratégia proposta ourAL-1 obteve excelentes resultados em praticamente todos os conjuntos avaliados, considerando cada um dos classificadores e técnicas de aprendizado baselines HCLU e HR. Dentre as estratégias de aprendizado comparadas, os maiores valores de acurácia obtidos encontram-se destacados em negrito, considerando cada um dos classificadores

Figura 14: Acurácia média por iteração dos métodos de aprendizado ativo no conjunto de dados *Lung Cancer* utilizando diferentes classificadores.



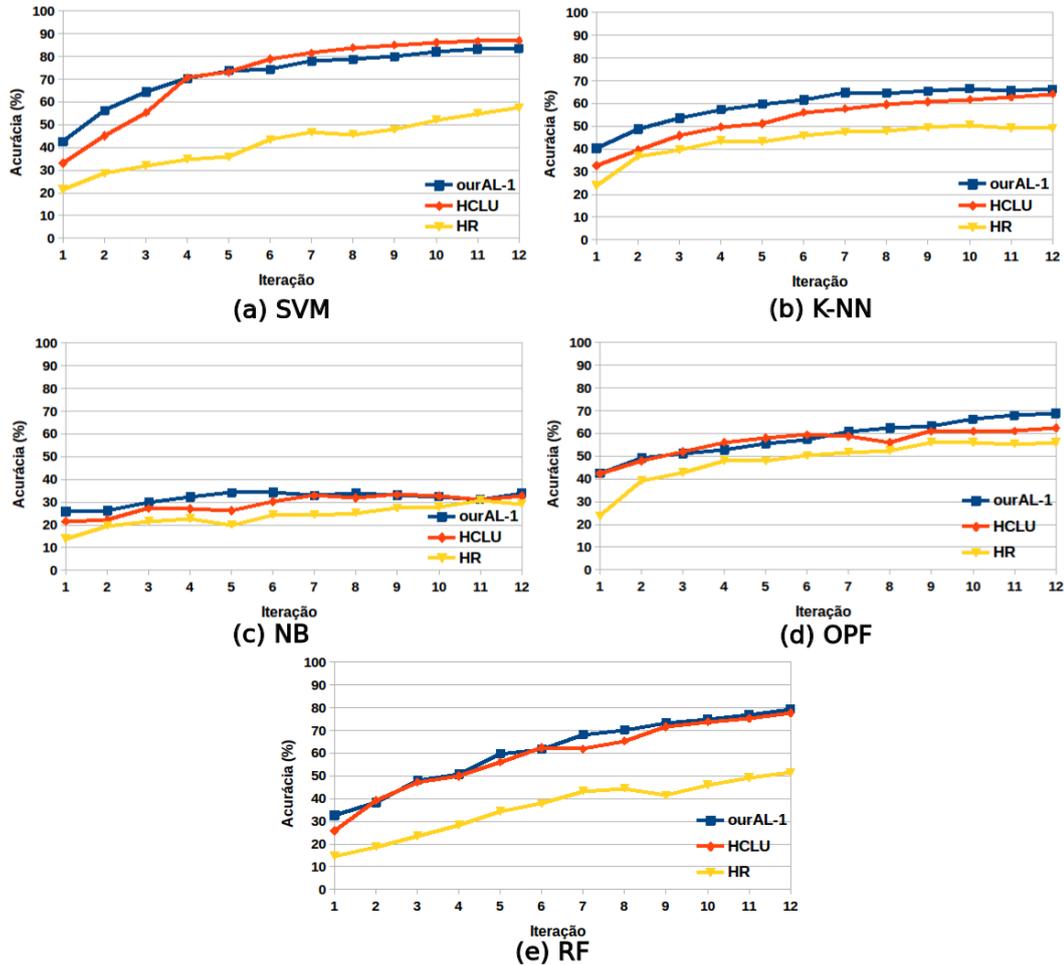
Fonte: Autoria própria.

e conjunto de dados.

Dentre os classificadores, as maiores acurácias (as quais encontram-se destacadas em sublinhado na Tabela 7), de forma geral, foram obtidas pelo SVM para todas as estratégias de aprendizado e conjuntos de dados utilizados.

Vale ressaltar que o total de amostras avaliadas a cada iteração é equivalente à quantidade de classes existentes em cada conjunto de dados. Sendo assim, analisando os conjuntos utilizados percebe-se que poucas amostras foram avaliadas a cada iteração. Como exemplo, considerando o conjunto *Lung Cancer*, o qual contém 2 classes, a técnica ourAL-1, utilizando SVM, obtém 96,87% de acurácia média (Tabela 7) com apenas 24 amostras avaliadas e anotadas ao longo das (12) iterações do aprendizado. É possível observar também que a técnica ourAL-1 consegue atingir valores de acurácias similares entre a segunda e terceira iteração (ver Figura 14(a)), ou seja, com cerca de 4 a 6 amostras.

Figura 15: Acurácia média por iteração dos métodos de aprendizado ativo no conjunto de dados *Subtypes of Acute Lymphoblastic Leukemia* utilizando diferentes classificadores.



Fonte: Autoria própria.

A partir dos resultados obtidos por meio da abordagem proposta de aprendizado para fins de classificação, a qual inclui estratégias de aprendizado ativo no processo, pode-se argumentar que a aplicação de uma análise detalhada da incerteza e representatividade das amostras no espaço de características pode melhorar consideravelmente o processo de aprendizagem do classificador. A abordagem proposta permite alcançar acurácias elevadas de maneira mais rápida, em menos iterações e interações de aprendizado. Portanto, possibilitando ganhos tanto em eficácia e eficiência, bem como minimizando o esforço de um especialista no processo de anotação.

Tabela 7: Acurácia média ao longo das iterações obtidas pelos métodos de aprendizado ativo, utilizando cada um dos classificadores, em cada conjunto de dados.

<i>MLL Leukemia</i>			
	ourAL-1	HCLU	HR
SVM	*88,45%	78,33%	80,42%
<i>k</i> -NN	75,54%	71,01%	71,07%
NB	72,38%	66,61%	63,63%
OPF	77,62%	70,77%	73,57%
RF	78,21%	69,70%	69,35%
<i>Breast Cancer</i>			
SVM	*61,36%	57,50%	61,27%
<i>k</i> -NN	54,74%	53,95%	51,93%
NB	53,60%	53,07%	52,63%
OPF	57,24%	55,83%	53,64%
RF	56,39%	52,12%	55,79%
<i>Central Nervous System</i>			
SVM	64,03%	60,28%	59,17%
<i>k</i> -NN	61,53%	54,37%	55,83%
NB	64,44%	58,68%	61,81%
OPF	58,68%	53,06%	47,15%
RF	*65,14%	61,53%	60,76%
<i>Colon Tumor</i>			
SVM	70,97%	67,22%	74,17%
<i>k</i> -NN	*75,00%	70,42%	65,14%
NB	60,28%	59,58%	61,94%
OPF	74,24%	69,44%	66,46%
RF	71,67%	66,39%	63,89%
<i>Lung Cancer</i>			
SVM	*96,87%	90,92%	88,42%
<i>k</i> -NN	92,73%	92,00%	89,71%
NB	92,18%	76,67%	85,00%
OPF	93,02%	90,54%	77,00%
RF	85,90%	85,36%	84,82%
<i>Subtypes of Acute Lymphoblastic Leukemia</i>			
SVM	*72,21%	72,16%	41,64%
<i>k</i> -NN	59,49%	53,45%	43,76%
NB	31,64%	29,08%	23,88%
OPF	58,14%	56,31%	48,18%
RF	61,15%	58,84%	36,05%

4 ABORDAGEM DE APRENDIZADO PARA CBIR

Este capítulo apresenta as contribuições deste trabalho direcionadas para a tarefa de recuperação de imagem por conteúdo. Para melhor organização, o capítulo encontra-se dividido em duas seções principais. A Seção 4.1 apresenta a descrição de uma estratégia para avaliação de desempenho de descritores (BRESSAN et al., 2018a), explorando o papel das características (*deep features*) obtidas por meio de aprendizado profundo na recuperação de imagens médicas baseadas em conteúdo. Na Seção 4.2, duas novas estratégias de aprendizado ativo são propostas para realimentação de relevância em CBIR (BRESSAN et al., 2018d, 2018b). As propostas têm como objetivo a obtenção de características e amostras mais informativas a cada iteração de realimentação do sistema, de forma a minimizar a interação do usuário e a acelerar o processo de aprendizado.

4.1 ESTRATÉGIA PARA AVALIAÇÃO DE DESEMPENHO DE DESCRITORES

Nos últimos anos, técnicas de aprendizado profundo (*deep learning*) têm sido empregadas em diferentes contextos para melhorar vários problemas de visão computacional. Este paradigma foi empregado com sucesso em muitas áreas (RAZZAK et al., 2018; GARCIA et al., 2017; WAN et al., 2014b). Uma grande desvantagem dessas abordagens é que elas precisam de um grande volume de dados rotulados para serem capazes de generalizar o modelo de aprendizado. Isso também se aplica às arquiteturas de última geração, como *Inception* (SZEGEDY et al., 2016), *ResNet* (HE et al., 2016), *NASNet* (ZOPH et al., 2017), entre outras.

No entanto, sabe-se que uma vez pré-treinadas por meio de um grande volume de dados (por exemplo, *ImageNet* (RUSSAKOVSKY et al., 2015)), essas arquiteturas podem ser eficientemente transferidas em subproblemas, bem como ajustadas aos mesmos. No entanto, sua capacidade de alcançar bons resultados é saturada quando aplicada a problemas (por exemplo, no contexto médico) extremamente diferentes das arquiteturas pré-treinadas.

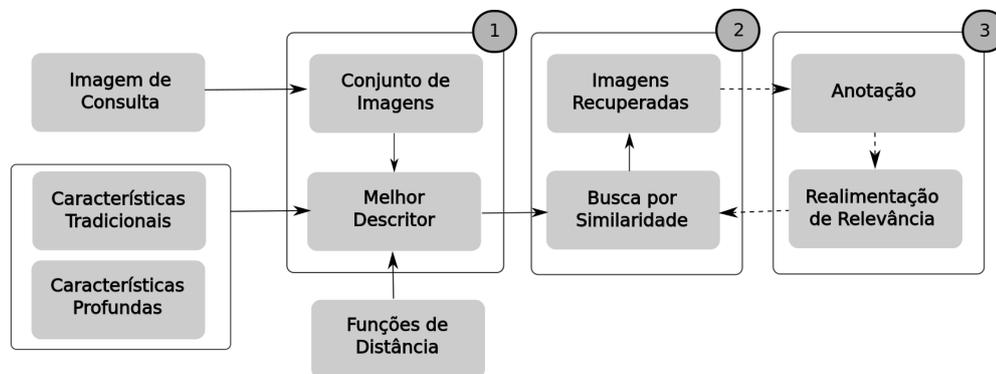
Apesar da grande aplicação atual de arquiteturas de aprendizagem profunda em vários problemas médicos, a maioria dos trabalhos concentram-se no treinamento das arquiteturas a

partir “do zero”, ou seja, não utilizam arquiteturas pré-treinadas. No entanto, além de requerer um elevado custo computacional, exige também um grande volume de dados para apresentar uma boa generalização.

Nesse sentido, estratégias de *transfer learning* (YOSINSKI et al., 2014) podem ser exploradas. A ideia das estratégias de *transfer learning* consiste em treinar as *Convolutional Neural Networks* - CNNs em um domínio geral e utilizar esse conhecimento prévio para resolver um novo ou o mesmo problema em um novo domínio específico. Essas técnicas têm sido utilizadas com sucesso no contexto de CNNs (YOSINSKI et al., 2014), em que pode-se fixar um determinado número de camadas de um treinamento prévio e utilizá-las para extrair características ou ajustar a rede, acelerando o tempo de treinamento e minimizando a probabilidade de *overfitting*, ao lidar com pequenos conjuntos de dados.

Nesse contexto, as contribuições propostas envolvem uma metodologia para avaliação de desempenho de descritores, de forma a obter as melhores características a serem consideradas no processo CBIR. A metodologia proposta (BRESSAN et al., 2018a) consiste em três etapas principais (Figura 16, em que as linhas tracejadas representam o ciclo de aprendizado incremental). O Algoritmo 3 descreve detalhes dessa metodologia.

Figura 16: Pipeline da metodologia proposta para avaliação de desempenho de descritores.



Fonte: (BRESSAN et al., 2018a)

Na Etapa 1, dada uma imagem de consulta q e um conjunto de imagens B , é realizada a seleção do melhor descritor (ou seja, o melhor par extrator de características e função de distância). Selecionar o melhor descritor é fundamental para o processo de recuperação. O processo de seleção do descritor é apresentado no Algoritmo 4. Nesse processo, analisam-se vários conjuntos de extratores de características E_j e funções de distância F_n .

Para obter as características que melhor descrevem as imagens foram consideradas as estratégias de extração de características tradicionais (HandCrafted $_j$) e profundas (Deep $_l$).

Algoritmo 3: Metodologia Proposta

Entrada : conjunto de imagens B e imagem de consulta q .
Saida : lista R_{ik} de k imagens recuperadas, considerando o i -ésimo extrator
Auxiliares: conjunto de extratores de características E_j , conjunto de funções de distância F_n , lista dos melhores descritores D_i^Ω , número nd de descritores desejados, número k de imagens recuperadas e estratégia de refinamento de consulta RR

- 1 $D_i^\Omega \leftarrow \text{obtencaoMelhorDescritor}(B, E_j, F_n, nd), i = 1, 2, \dots, nd;$
- 2 $R_{ik} \leftarrow \text{buscaPorSimilaridade}(q, D_i^\Omega, k);$
- 3 **repeat**
- 4 $R_{ik} \leftarrow \text{anotacao}(R_{ik});$
- 5 $R_{ik} \leftarrow \text{refinamentoConsulta}(q, R_{ik}, k, RR);$
- 6 **until** *insatisfeito*;

Com relação às características obtidas por meio de aprendizado profundo, propõe-se o uso de estratégias de *transfer learning* T_l .

Após a obtenção das características tradicionais e profundas (features_m), para cada par descritor de características e função de distância, consultas por similaridade são realizadas utilizando todas as imagens de B como imagem de consulta, para verificar previamente a precisão (P_{mn}) obtida por cada descritor (ou seja, cada um (m -ésimo) dos extratores de características e cada uma (n -ésima) das funções de distância). Em seguida, é possível obter os nd melhores descritores D_i^Ω , os quais são aqueles que apresentam os melhores resultados de precisão.

Algoritmo 4: Obtenção do Melhor Descritor

Entrada : conjunto de imagens B
Saida : lista dos nd melhores descritores D_i^Ω
Auxiliares: conjunto de extratores de características E_j , conjunto de características tradicionais HandCrafted_j , conjunto de características profundas Deep_l , conjunto de estratégias de *transfer learning* T_l , conjunto de características tradicionais e profundas features_m , conjunto de funções de distância F_n , número de descritores desejados nd e lista de valores de precisão P_{nm}

- 1 $\text{HandCrafted}_j \leftarrow \text{ExtracaoTradicional}(B, E_j);$
- 2 $\text{Deep}_l \leftarrow \text{ExtracaoDeep}(B, T_l);$
- 3 $\text{features}_m \leftarrow \{\text{HandCrafted}_j\} \cup \{\text{Deep}_l\};$
- 4 **for** *each* $m \in \text{features}_m$ **do**
- 5 **for** *each* $n \in F_n$ **do**
- 6 $P_{mn} = \text{buscaPorSimilaridade}(B, \text{features}_m, F_n);$
- 7 **end**
- 8 **end**
- 9 $D_i^\Omega \leftarrow \text{selecaoDescritores}(P_{mn}, nd), i = 1, 2, \dots, nd;$

Na Etapa 2, os melhores descritores (obtidos pelo Algoritmo 4), podem iniciar o pro-

cesso de recuperação (Algoritmo 3), retornando as k imagens mais próximas R_{ik} da imagem de consulta q para cada i -ésimo descritor. O processo de recuperação é repetido até que os resultados retornados sejam satisfatórios. Neste caso, na Etapa 3, o processo de anotação é realizado, na qual as imagens relevantes são indicadas pelo especialista. Em seguida, a consulta é refinada por meio da estratégia de realimentação de relevância RR (Etapa 3).

4.1.1 DESCRIÇÃO DOS CONJUNTOS DE DADOS

Os conjuntos de imagens (VIENNA e MINI-MIAS) considerados nos experimentos são subconjuntos do MAMMOSET (OLIVEIRA et al., 2017), o qual é composto por regiões de interesse (*Regions Of Interests* - ROIs) de mamografias.

O conjunto de dados VIENNA, criado pelo Departamento de Radiologia da Universidade de Viena, é composto de imagens de mamografias coletadas a partir do *Breast Imaging Reporting and Data System* - BI-RADS), realizado na mesma universidade. O repositório inclui 447 imagens de ROIs, contendo informações relacionadas ao tipo de lesão (calcificação ou massa) de tecidos tumorais e de mamografias mistas.

O repositório da Sociedade de Análise de Imagens Mamográficas (*Mammographic Image Analysis Society* - MINI-MIAS) é uma versão reduzida do conjunto de dados MIAS, contendo 118 imagens. Esses conjuntos de dados contêm informações sobre o tecido de fundo, classe e gravidade da anormalidade, bem como as coordenadas para o centro da anormalidade e o raio aproximado de um círculo que a envolve.

As Tabelas 8 e 9 apresentam as classes e quantidades de amostras em cada classe dos conjuntos de imagens (VIENNA e MINI-MIAS) considerados nos experimentos, os quais são denominados como B_1 e B_2 , respectivamente.

Tabela 8: Descrição do conjunto de imagens VIENNA - B_1

Classes	Numero de amostras
birads3_calcification	33
birads3_mass	61
birads4_calcification	113
birads4_mass	72
birads5_calcification	32
birads5_mass	49
normal	41

Tabela 9: Descrição do conjunto de imagens MINI-MIAS - B_2

Classes	Numero de amostras
calcification_b	14
calcification_m	12
mass_b	38
mass_m	20

4.1.2 DESCRIÇÃO DOS EXPERIMENTOS

Como uma primeira análise foram realizadas consultas por similaridade, utilizando cada imagem dos conjuntos de dados descritos na Seção 4.1.1 e variando o extrator de características e a função de distância, a fim de estabelecer os melhores descritores (pares de extrator de características e função de distância).

A Tabela 10 apresenta as estratégias de extração de características tradicionais, suas respectivas categorias (cor, textura, forma) e quantidade de características consideradas. As características profundas foram obtidas a partir da camada convolucional final das CNNs (*Inception-v3*, *Inception-ResNet-v2* e *NASNet-Large*) por meio do processo de *transfer learning*. Essas CNNs foram pré-treinadas por meio do conjunto de dados do *ImageNet*.

Tabela 10: Extratores de características tradicionais (BREISSAN et al., 2018a).

Extrator	Categoria	#Características
BIC Histogram (STEHLING et al., 2002)	Cor	512
Edge Histogram (WON et al., 2002)	Cor	150
Normalized Histogram (NIXON; AGUADO, 2012)	Cor	256
Haralick (HARALICK et al., 1973)	Textura	24
Rotation Invariant LBP (GUO et al., 2010)	Textura	108
Texture Spectrum (HE; WANG, 1990)	Textura	8
Daubechies (WANG et al., 1998)	Forma	16
Haar (WANG et al., 1998)	Forma	16
Zernike (KHOTANZAD; HONG, 1990)	Forma	36

As funções de distância consideradas foram L_1 , L_2 , L_∞ , χ^2 , Canberra, Jeffrey Divergence e dLog (Seção 2.1.1). Sendo assim, é realizado um produto cartesiano entre os conjuntos de características e funções de distância, gerando um total de 84 (12×7) análises para cada conjunto de dados. Para obter os melhores pares é considerada a precisão média (*Mean Average Precision* - MAP), a partir das consultas por similaridade, conforme descrito por Baeza-Yates e Ribeiro-Neto (2011).

Para o método de realimentação de relevância, foi utilizada a abordagem *Query-Point Movement* - QPM, amplamente utilizada na literatura, conforme descrita em detalhes na Seção 2.1.2. Para simular o processo de anotação foi realizada a consulta de k vizinhos mais

próximos, sendo utilizadas todas as imagens do conjunto como imagem de consulta. O valor de k foi definido como 30, uma vez que valores maiores geralmente não são empregados na prática diária de ambientes médicos. Quando a quantidade de imagens (nc) de uma determinada classe do conjunto de dados é menor que 30, o valor de k é definido como sendo nc ($k = nc$). Dessa forma, é possível recuperar as k imagens, as quais são anotadas e utilizadas no processo de realimentação de relevância a cada iteração.

Para avaliar os resultados obtidos pelas consultas por similaridade são apresentados gráficos de Precisão e Revocação (curvas P&R), utilizando os melhores descritores de características tradicionais (nomeados como “Tradicional”) e os melhores descritores de características profundas, obtidas pelas arquiteturas (nomeadas “Inception-v3” (SZEGEDY et al., 2016), “Inception-ResNet-v2” (HE et al., 2016) e “NASNet-Large” (ZOPH et al., 2017)). A precisão de uma consulta é a fração das imagens recuperadas que são relevantes e a revocação é a proporção de imagens relevantes que foram recuperadas pela consulta, considerando todas as imagens relevantes no conjunto de dados. Assim, a precisão atinge 100%, quando cada imagem recuperada é relevante para a consulta e a revocação atinge 100%, quando todas as imagens relevantes do conjunto de dados foram retornadas (BAEZA-YATES; RIBEIRO-NETO, 2011). Como regra geral, quanto mais próxima a curva P&R do topo, melhor é o método.

4.1.3 RESULTADOS E DISCUSSÕES

Inicialmente, de acordo com a análise dos melhores descritores, obtiveram-se os melhores pares destacados em negrito nas Tabelas 11 e 12 para os conjuntos de dados B_1 e B_2 , respectivamente.

Após a análise dos melhores descritores, nas Figuras 17 e 18 são apresentados os resultados obtidos a partir das consultas por similaridade (curvas P&R) ao longo das iterações de realimentação de relevância, utilizando o melhor descritor tradicional e os melhores descritores profundos nos conjuntos de dados B_1 e B_2 , respectivamente.

A partir das Figuras 17(a) e 17(b), referentes à consulta inicial e à primeira iteração, respectivamente, é possível observar que os dois tipos de características (tradicionais e profundas) apresentaram resultados equivalentes.

No entanto, ao longo das iterações de realimentação de relevância, esses resultados apresentaram diferenças consideráveis, ou seja, quanto maior a iteração, maior diferença entre eles. Por exemplo, na Figura 17 (c) (segunda iteração de realimentação de relevância), as características profundas, obtidas a partir do processo de *transfer learning* utilizando Inception-v3,

Tabela 11: Melhores descritores para o conjunto de dados B_1 . Extratores de características e funções de distância em negrito compreendem os melhores pares para características tradicionais e profundas (BREISSAN et al., 2018a).

Extrator	Função de Distância
BIC Histogram	L_∞
Edge Histogram	<i>Jeffrey Divergence</i>
Normalized Histogram	L_∞
Haralick	<i>Jeffrey Divergence</i>
Rotation Invariant LBP	<i>Jeffrey Divergence</i>
Texture Spectrum	<i>Canberra</i>
Daubechies	<i>Jeffrey Divergence</i>
Haar	L_1
Zernike	L_∞
Inception-v3	<i>Canberra</i>
Inception-ResNet-v2	L_∞
NASNet-Large	<i>Jeffrey Divergence</i>

Tabela 12: Melhores descritores para o conjunto de dados B_2 . Extratores de características e funções de distância em negrito compreendem os melhores pares para características tradicionais e profundas (BREISSAN et al., 2018a).

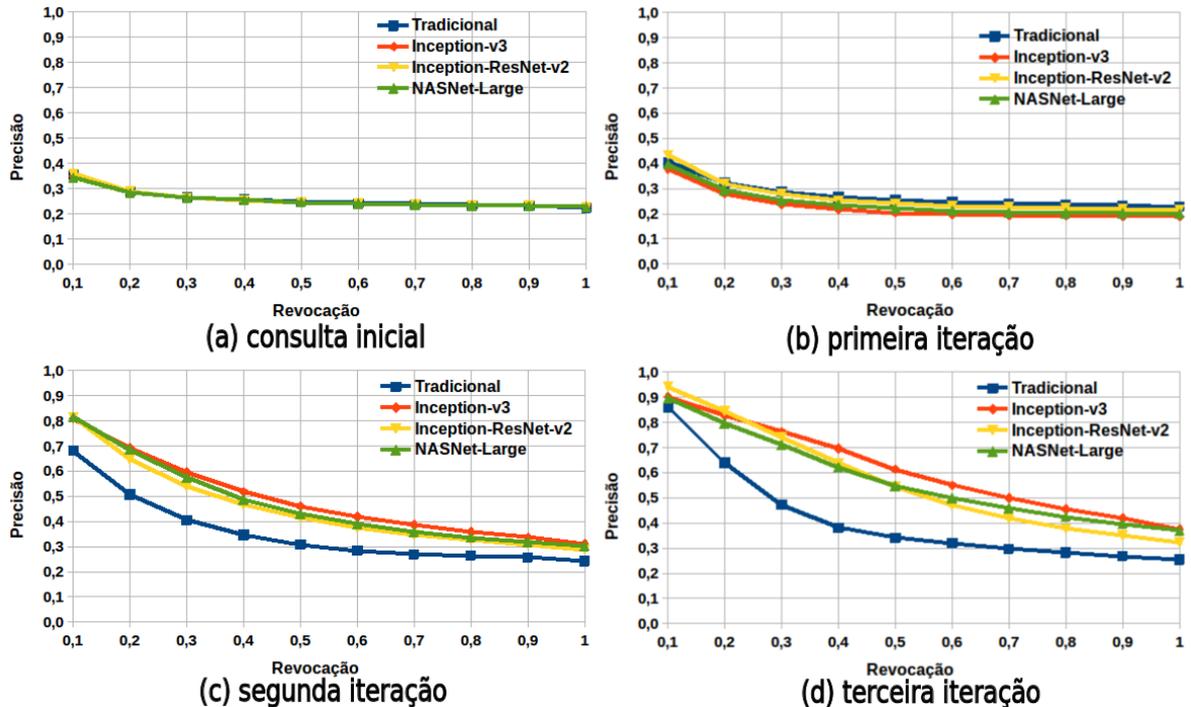
Extrator	Função de Distância
BIC Histogram	L_2
Edge Histogram	L_2
Normalized Histogram	L_2
Haralick	<i>Canberra</i>
Rotation Invariant LBP	L_2
Texture Spectrum	<i>Canberra</i>
Daubechies	L_2
Haar	<i>Canberra</i>
Zernike	L_1
Inception-v3	L_2
Inception-ResNet_v2	<i>dLog</i>
NASNet-Large	L_2

Inception-ResNet-v2 e NASNet-Large, apresentaram ganhos notáveis quando comparadas com as tradicionais.

As características profundas atingiram ganhos de precisão de até 49,3% em relação ao tradicional. Em uma análise geral, realizando comparações de desempenho dentre as arquiteturas profundas, o Inception-v3 apresentou precisões um pouco melhores. Vale mencionar que, de acordo com (ZOPH et al., 2017), o Inception-v3 e o Inception-ResNet-v2 requerem aproximadamente 2,5 vezes menos recursos computacionais em relação ao NASNet-Large.

Na Figura 17 (d) os ganhos de precisão apresentados pelas arquiteturas profundas aumentaram ainda mais, atingindo ganhos de até 82,6%. Ao considerar 20% e 30% de valores

Figura 17: Resultados obtidos para o conjunto de dados B_1 , utilizando os melhores descritores tradicional e profundos, ao longo das consultas por similaridade: (a) consulta inicial e, refinamentos das consultas: (b) primeira iteração; (c) segunda iteração; e (d) terceira iteração.



Fonte: (BRESSAN et al., 2018a)

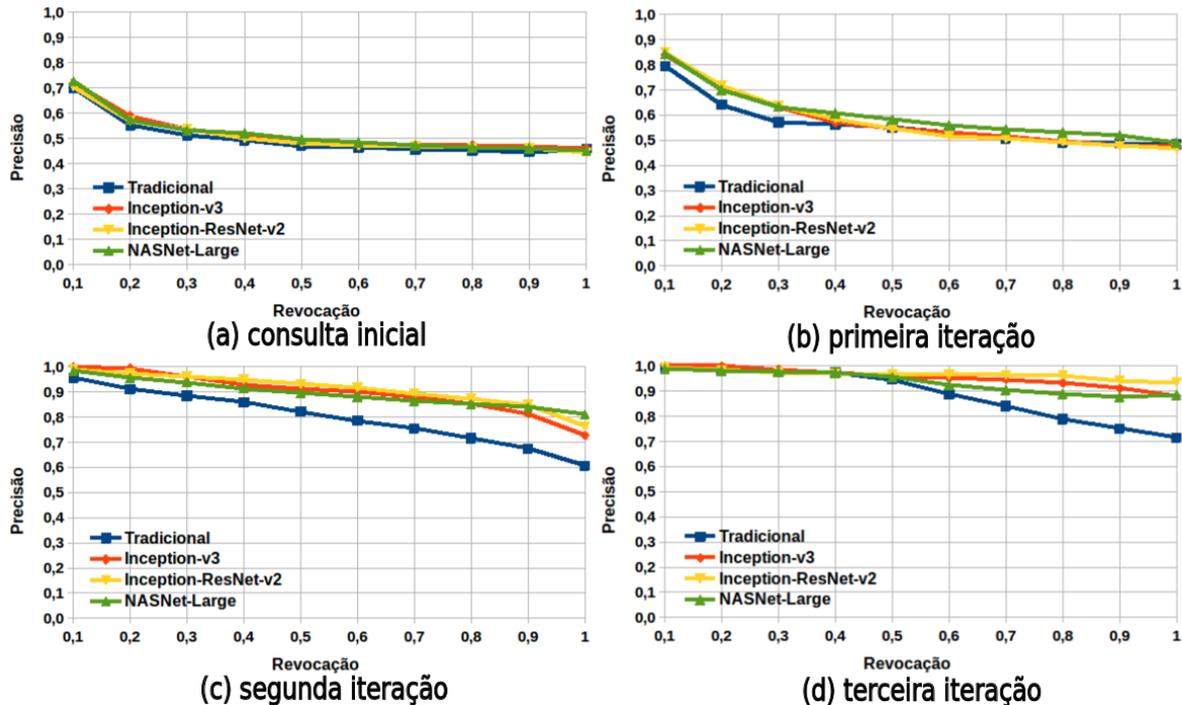
de revocação, as características obtidas a partir do Inception-v3 apresentaram ganhos de 30% e 61,5%, respectivamente. Para os mesmos valores de revocação, o Inception-ResNet-v2 atingiu ganhos de 32% e 57,1%, respectivamente.

As consultas por similaridade iniciais (Figura 18 (a)) apresentaram valores de precisão baixos, utilizando ambos descritores tradicionais e arquiteturas profundas. No entanto, já no primeiro refinamento da consulta (Figura 18 (a)), as arquiteturas profundas apresentam os melhores desempenhos. Considerando níveis de revocação de 20% e 30%, Inception-ResNet-v2 apresentou ganhos de 11,7% e 11,3%, respectivamente. Inception-v3 e NASNet-Large também apresentaram resultados equivalentes.

Analisando os níveis de revocação de 10% e 20% (Figura 18), Inception-ResNet-v2 apresentou precisões de 100% e 99%, respectivamente. Inception-v3 apresentou uma precisão de 100% para os mesmos níveis de revocação. As arquiteturas profundas apresentaram ganhos de precisão de até 33,7% em relação ao descritor tradicional, bem como mantiveram os melhores resultados, atingindo ganhos de até 12,89%. Além disso, dentre as características profundas, Inception-v3 e Inception-ResNet-v2 apresentaram os melhores resultados.

A partir dos resultados obtidos, pode-se atestar que é possível fazer uso do aprendi-

Figura 18: Resultados obtidos para o conjunto de dados B_2 , utilizando os melhores descritores tradicional e profundos, ao longo das consultas por similaridade: (a) consulta inicial e, refinamentos das consultas: (b) primeira iteração; (c) segunda iteração; e (d) terceira iteração.



Fonte: (BRESSAN et al., 2018a)

zado por transferência de um contexto geral para um contexto médico específico. No entanto, para aproveitar ao máximo esse princípio, o processo de recuperação deve ser realizado com características gerais profundas em um determinado processo de refinamento da consulta. Esses resultados também são interessantes, considerando a rotina real dos radiologistas, os quais desejam ter determinado controle sobre o processo para melhor aceitação do sistema CBIR. Sendo assim, a utilização do refinamento da consulta com as características profundas não apenas aumentou em grande parte a precisão das consultas por similaridade, mas também possibilita uma melhor aceitação dos sistemas CBIR na prática real em ambientes de saúde. A Seção 4.2 apresenta algumas estratégias que podem melhorar o processo de refinamento das consultas.

4.2 ESTRATÉGIAS DE APRENDIZADO ATIVO

Embora haja uma variedade de métodos de realimentação de relevância na literatura, a maioria deles renuncia a definição do grau de relevância e irrelevância para o usuário, descartando essa informação ao longo das iterações. No entanto, uma vez que uma consulta é refinada, a cada iteração, mais e mais imagens relevantes são retornadas, o que não contribui para o processo de aprendizado. Isso ocorre dado que é necessário analisar o melhor *trade-*

off entre similaridade e diversidade, de forma a permitir que o processo escape de possíveis armadilhas locais (mínimo local).

Alguns trabalhos tentam unir métodos de aprendizado de máquina e técnicas de realimentação de relevância para mitigar melhor o problema do *gap* semântico em sistemas CBIR (ZHOU et al., 2008) e realizar uma captura mais refinada da intenção do especialista. No entanto, esses trabalhos não utilizam tal intenção capturada durante o processo de realimentação de relevância para melhorar gradativamente o modelo de aprendizado.

Sendo assim, diferentemente dos trabalhos anteriores, é proposta a inclusão de estratégias de aprendizado ativo, a fim de obter melhorias em consultas por similaridade em sistemas CBIR. Neste caso, as imagens mais informativas a serem selecionadas são aquelas que apresentam o melhor balanceamento entre a similaridade (com a imagem de consulta) e determinados graus de diversidade e incerteza. O critério de seleção adotado deve permitir selecionar imagens de classes distintas, bem como que sejam difíceis de diferenciar quando comparada à semântica da imagem de consulta e a semântica da imagem recuperada (por exemplo, imagens na fronteira de duas classes diferentes/sobrepostas).

A Figura 19 apresenta um exemplo de duas amostras informativas de classes distintas, lesões benignas e lesões malignas, respectivamente. De acordo com o critério adotado de seleção, tais amostras podem ser selecionadas e apresentadas ao especialista, diferentemente do processo CBIR tradicional, que apresentaria apenas imagens próximas ao centro de consulta.

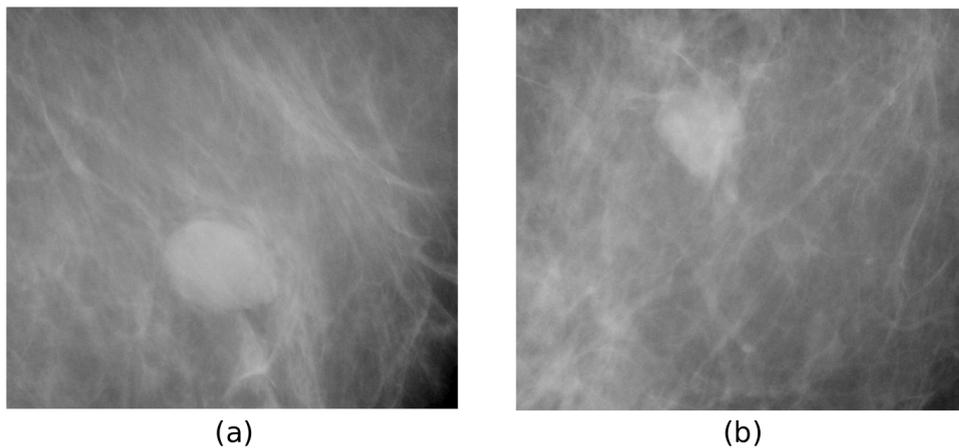


Figura 19: Exemplos de amostras informativas (incertas) de duas classes diferentes (a) lesões benignas e (b) lesões malignas.

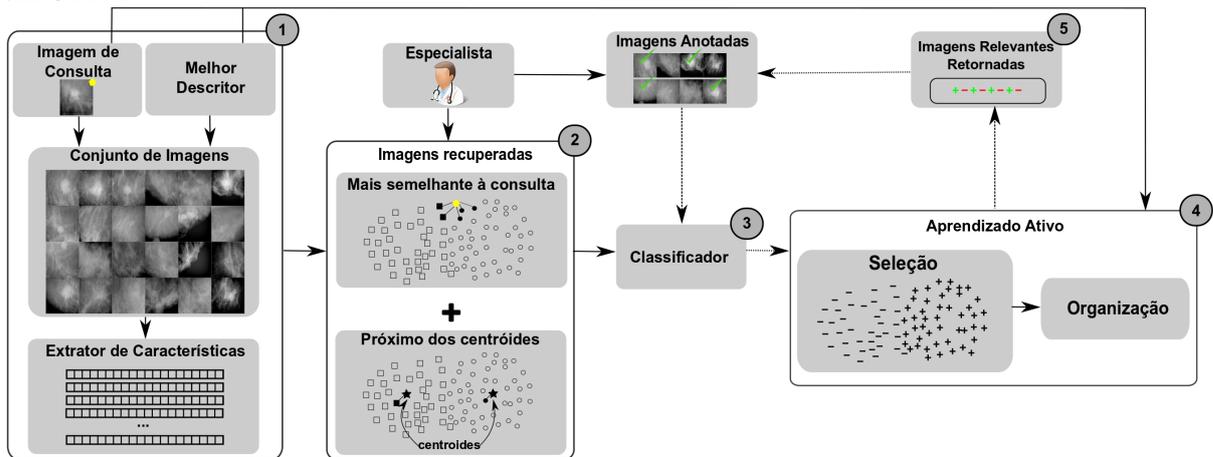
Fonte: Autoria própria.

A estratégia proposta possibilita equilibrar o processo de aprendizado com as imagens mais informativas (similares e incertas) que melhor contribuem para obter um classificador

robusto mais rapidamente, e conseqüentemente, melhorar a qualidade das imagens retornadas, mais similares em relação à imagem de consulta.

As principais etapas da abordagem proposta são apresentadas na Figura 20. As linhas tracejadas representam o ciclo do processo de aprendizado incremental. O Algoritmo 5 descreve detalhes da abordagem proposta.

Figura 20: Pipeline da abordagem de aprendizado ativo proposta para realimentação de relevância em CBIR.



Fonte: Autoria própria.

Na Etapa 1, dada uma imagem de consulta q , o processo de recuperação é iniciado. Inicialmente, considerando q e um conjunto de dados de imagem I , é realizada a seleção do melhor descritor (melhor par extrator de característica e função de distância). Como esse processo é fundamental para o processo de recuperação, diferentes extratores de características F_i e funções de distância D_j foram avaliados. Em seguida, as características de baixo nível são extraídas de I (Linha 1), utilizando o melhor extrator de características (*BestFD.FeatureExtractor*) e gerando o conjunto de aprendizado Z_2 (Linha 2). Também são extraídas características de q , utilizando o mesmo extrator.

Considerando a Etapa 2, o conjunto de aprendizado Z_2 é particionado em k grupos utilizando um algoritmo de agrupamento (Linha 3). De acordo com a estratégia de agrupamento aplicada nesta tarefa, o centro de cada grupo pode ser um medoide, sendo assim não é necessário obter a imagem mais próxima do centroide, mas a própria imagem, gerando o conjunto de centroides/medoides C . Além das imagens de C , as imagens mais similares em relação à q são selecionadas. Para tanto, ns imagens desejadas são obtidas de L_S , ordenadas por uma ordem crescente de distância em relação à q (Linhas 4 e 5). Essas imagens são apresentadas ao especialista para anotação (Linha 6). Sendo assim, o conjunto de treinamento Z_1 é composto pelas imagens mais próximas de cada centroide C e as L_{S_i} , $i = 1, \dots, ns$ imagens mais próximas

Algoritmo 5: Abordagem Proposta

Entrada : imagem de consulta q

Saida : instância final do modelo de aprendizado M^Ω e uma lista L_R^Ω ordenada por relevância em relação à q , sendo $L_R \subset I$

Auxiliares: conjunto de imagens I , conjuntos de extratores de características F_i e de funções de distância D_j , conjunto de aprendizado Z_2 , número de *clusters* k , conjunto de centroides/medoides C , lista L_S ordenada das imagens mais próximas a q , número de amostras desejadas ns , conjunto de treinamento temporário Z'_1 , conjunto de treinamento Z_1 , instância do modelo de aprendizado M , número de amostras a serem selecionadas nu .

- 1 $BestFD \leftarrow obtencaoMelhorDescritor(I, F_i, D_j)$;
- 2 $Z_2 \leftarrow extracaoCarateristicas(BestFD.FeatureExtractor)$;
- 3 $C \leftarrow obtencaoCentroides(Z_2, k)$;
- 4 $L_S \leftarrow buscaPorSimilaridade(q, Z_2, BestFD.Distance)$;
- 5 $Z'_1 \leftarrow C \cup s_i \in L_S, i = 1, 2, \dots, ns$;
- 6 $Z_1 \leftarrow anotacao(Z'_1)$;
- 7 $M \leftarrow treinamento(Z_1)$;
- 8 **repeat**
- 9 $Z'_1 \leftarrow estrategiaAprendizadoAtivo(nu, M, Z_2 \setminus Z_1, q, BestFD.Distance)$;
- 10 $Z_1 \leftarrow Z_1 \cup anotacao(Z'_1)$;
- 11 $M \leftarrow reTreinamento(Z_1, M)$;
- 12 **until** *satisfeito*;
- 13 $M^\Omega \leftarrow M$;
- 14 $L_R^\Omega \leftarrow ordenacao(Z, M^\Omega), L_R^\Omega \subset I$;

à q . As imagens anotadas constituem o conjunto de treinamento inicial Z_1 , o qual é utilizado no processo de treinamento (Linha 7), gerando a primeira instância do modelo de aprendizado M (Etapa 3).

Na Etapa 4, o atual modelo de aprendizado M participa ativamente do processo de seleção das imagens mais informativas a serem utilizadas em seu próprio treinamento, a fim de melhorar o resultado da consulta, retornando imagens mais similares. Para tanto, os critérios de seleção propostos são baseados na incerteza e similaridade em relação à imagem de consulta q . Assim, o conjunto temporário de treinamento Z'_1 recebe as nu imagens mais informativas em relação à q , de acordo com o modelo de aprendizado atual M , as amostras restantes em Z_2 ($Z_2 \cap Z_1 = \emptyset$ ou em uma notação simplificada $Z_2 \setminus Z_1$), assim como a imagem de consulta q e a melhor função de distância $BestFD.Distance$ (Linha 9).

Na Etapa 5, o conjunto de imagens selecionado Z'_1 é exibido para anotação (como relevante ou irrelevante) pelo especialista. A partir da primeira iteração, esse conjunto já é rotulado previamente pela instância atual do modelo. Sendo assim, o especialista precisa apenas

corrigir os rótulos de imagens classificadas incorretamente. As imagens confirmadas e corrigidas pelo especialista são adicionadas ao conjunto de treinamento Z_1 (Linha 10). Em seguida, o modelo M é treinado novamente e uma nova instância é gerada (Linha 11). O especialista pode continuar o processo de aprendizado, Etapas 3 e 4 (Linhas 8-12), até que esteja satisfeito com os resultados retornados. Uma vez satisfeito, pode-se obter um modelo final de aprendizado M^Ω (Linha 13) e uma lista final L_R^Ω (Linha 14) ordenada por relevância (da mais similar para a menos similar) em relação à imagem de consulta q .

A estratégia de aprendizado ativo proposta para selecionar as amostras mais informativas é descrita em detalhes no Algoritmo 6. Inicialmente, as Linhas 1-3 referem-se apenas a um controle para verificar se existem nu amostras desejadas para serem selecionadas no conjunto de aprendizado $Z_2 \setminus Z_1$, dado que trata-se de um processo iterativo, no qual um conjunto de imagens é selecionado a cada iteração.

Algoritmo 6: Estratégia de Aprendizado Ativo

Entrada : número de amostras a serem selecionadas nu , conjunto de aprendizado $Z_2 \setminus Z_1$, imagem de consulta q e melhor função de distância $BestFD.Distance$

Saida : lista de amostras relevantes selecionadas L_R .

Auxiliares: conjunto de aprendizado $Z_2' \setminus Z_1$, lista ordenada de candidatas L_C

```

1 if tamanho( $Z_2 \setminus Z_1$ ) <  $nu$  then
2   |  $nu \leftarrow$  tamanho( $Z_2 \setminus Z_1$ );
3 end
4  $Z_2' \setminus Z_1 \leftarrow$  classificacaoImagens( $Z_2 \setminus Z_1, M$ );
5  $L_C \leftarrow$  estrategiaSelecao( $nu, Z_2' \setminus Z_1, q, BestFD.Distance$ );
6 while  $i \leq nu$  do
7   |  $L_R \leftarrow L_R \cup s_i \in L_C$ ;
8 end

```

O intuito é explorar e utilizar o classificador, construído a partir das amostras mais informativas, para melhorar o processo de recuperação de imagens. Sendo assim, o conjunto de aprendizado $Z_2 \setminus Z_1$ é classificado pelo modelo de aprendizado atual M , gerando o conjunto de aprendizado rotulado $Z_2' \setminus Z_1$ (Linha 4). Após o processo de classificação, obtém-se a lista das imagens candidatas mais informativas L_C (Linha 5), de acordo com os critérios de seleção propostos (ourAL-2 e ourAL-3), os quais são descritos nas Seções 4.2.1 e 4.2.2, respectivamente. Em seguida, obtém-se o conjunto de amostras (relevantes) selecionado L_R , composto de nu amostras mais informativas (incertas e similares) a partir da lista de candidatas L_C (Algoritmo 6, Linhas 6-8).

4.2.1 ESTRATÉGIA DE SELEÇÃO ourAL-2

Para seleção das amostras mais informativas, a estratégia ourAL-2 (BRESSAN et al., 2018d) considera amostras, tanto relevantes como irrelevantes, as quais foram classificadas pela instância atual do classificador. Na primeira iteração do aprendizado, o especialista realiza o processo de anotação (indicando as imagens relevantes e irrelevantes, de acordo com uma dada imagem de consulta q), conforme apresentado no Algoritmo 5 (Linhas 1-7). Em seguida, a partir da primeira iteração, as imagens recuperadas devem ser aquelas que mais contribuirão para o processo de aprendizado do classificador, o qual participa ativamente do processo, auxiliando na seleção das amostras mais informativas. O Algoritmo 7 descreve o funcionamento da seleção realizada pela estratégia ourAL-2.

Algoritmo 7: Estratégia de Seleção - ourAL-2

Entrada : número de amostras a serem selecionadas nu , conjunto de aprendizado $Z'_2 \setminus Z_1$, imagem de consulta q e melhor função de distância $BestFD.Distance$

Saida : lista de amostras candidatas selecionadas L_C

Auxiliares: listas temporárias de imagens relevantes e irrelevantes L'_P e L'_N , centro de massa das amostras relevantes C_P , listas de imagens relevantes e irrelevantes L_P e L_N

- 1 $L'_P \leftarrow \text{selecaoRelevantes}(Z'_2 \setminus Z_1)$;
 - 2 $L'_N \leftarrow \text{selecaoIrrelevantes}(Z'_2 \setminus Z_1)$;
 - 3 $C_P \leftarrow \text{localizacaoCentroMassa}(L'_P)$;
 - 4 $L_P \leftarrow \text{ordenacaoDecrescente}(L'_P, BestFD.Distance, C_P)$;
 - 5 $L_N \leftarrow \text{ordenacaoCrescente}(L'_N, BestFD.Distance, C_P)$;
 - 6 $L_C \leftarrow \text{concatenacao}(L_P, L_N)$;
-

Inicialmente, as amostras do conjunto $Z'_2 \setminus Z_1$, classificadas como relevantes ou irrelevantes a partir do modelo de aprendizado M (Algoritmo 6, Linha 4), são organizadas e separadas em listas. Sendo assim, a lista L'_P recebe as amostras classificadas como relevantes a partir de $Z'_2 \setminus Z_1$ e L'_N recebe as classificadas como irrelevantes (Linhas 1-2). Em seguida, localiza-se o centro de massa de L_P , armazenando-o em C_P (Linha 3). O centro de massa pode ser calculado simplesmente pela média das características das amostras existentes na lista.

As imagens em L'_P são ordenadas em relação à distância do centro de massa C_P , sendo priorizadas as imagens de maior distância, ou seja, das mais distantes para as mais próximas de C_P (Linha 4). As imagens em L'_N também são ordenadas em relação à distância do centro de massa C_P . No entanto, nesse caso, as amostras são organizadas de forma inversa, sendo das mais próximas para as mais distantes de C_P (Linha 5). Após a ordenação, as amostras relevantes e irrelevantes de L_P e L_N são concatenadas formando a lista de candidatas L_C (Linha 6), a qual

é retornada e exibida ao especialista para anotação (confirmação dos rótulos).

4.2.2 ESTRATÉGIA DE SELEÇÃO ourAL-3

Para a estratégia de seleção ourAL-3 (Algoritmo 8), inicialmente, uma lista de candidatas L_C é criada para receber as amostras candidatas a serem exibidas para o especialista (Linha 1). As listas de aprendizado organizadas com base nos rótulos L_i também são criadas (Linha 1), onde i denota a i -ésima classe ($i = \{1, 2\}$, ou seja, classes relevante e irrelevante, respectivamente).

Algoritmo 8: Estratégia de Seleção - ourAL-3

Entrada : número de amostras a serem selecionadas nu , conjunto de aprendizado $Z'_2 \setminus Z_1$, imagem de consulta q e melhor função de distância $BestFD.Distance$

Saida : lista de amostras candidatas selecionadas L_C

Auxiliares: listas L_i organizadas com base em rótulos, vizinho adjs mais próximo, listas temporárias de imagens relevantes e irrelevantes L'_P e L'_N , listas ordenadas de imagens relevantes e irrelevantes L_P e L_N , e lista $Lcom_i$ com base no centro de massa de cada rótulo

```

1  $L_C, L_i \leftarrow \emptyset, i = \{1, 2\}$ ;
2 for cada  $s \in Z'_2 \setminus Z_1$  do
3    $L_i \leftarrow L_i \cup \{s\}, i = s.labelid$ ;
4    $adjs \leftarrow obtencaoVizinhaMaisProxima(s)$ ;
5   if  $s.labelid \neq adjs.labelid$  then
6      $L_C \leftarrow L_C \cup \{s, adjs\}$ ;
7   end
8 end
9 if  $L_C \neq \emptyset$  then
10   $L'_P \leftarrow selecaoRelevantes(L_C)$ ;
11   $L'_N \leftarrow selecaoIrrelevantes(L_C)$ ;
12   $L_P \leftarrow ordenacaoSimilaridade(q, L'_P, BestFD.Distance)$ ;
13   $L_N \leftarrow ordenacaoSimilaridade(q, L'_N, BestFD.Distance)$ ;
14   $L_C \leftarrow concatenacao(L_P, L_N)$ ;
15 end
16 if  $tamanho(L_C) < nu$  then
17   $Lcom_i \leftarrow localizacaoCentroMassa(L_i), i = \{1, 2\}$ ;
18   $L_P \leftarrow ordenacaoDecrescente(Lcom_1, L_1)$ ;
19   $L_N \leftarrow ordenacaoDecrescente(Lcom_2, L_2)$ ;
20  repeat
21     $L_C \leftarrow L_C \cup \{s_j, t_k\}, s_j \in L_P \text{ e } t_k \in L_N$ ;
22  until  $tamanho(L_P) > j$  e  $tamanho(L_N) > k$ ;
23 end

```

Em seguida, cada amostra s do conjunto de aprendizado $Z'_2 \setminus Z_1$, previamente rotulada

pelo modelo atual de aprendizado M , é analisada para avaliar quais são as candidatas mais informativas (Linhas 2 a 8). Todas as amostras são separadas de acordo com os rótulos das classes fornecidos pelo modelo. Sendo assim, cada amostra $s \in Z'_2 \setminus Z_1$ é armazenada em uma lista de rótulos L_i , correspondente ao respectivo rótulo da classe i ($i = s.labelid$) (Linha 3).

Obtém-se a imagem vizinha mais próxima (1-NN) em relação à imagem s , nomeada como imagem adjacente $adjs$ (Linha 4). Se os rótulos (fornecidos por M) das imagens s e $adjs$ são diferentes (isto é, $s.labelid \neq ads.labelid$, Linha 5), ambas imagens são inseridas na lista de candidatas (Linha 6). Estas imagens são consideradas as candidatas mais informativas (incertas), uma vez que são próximas umas das outras e têm sido rotuladas pela instância atual do modelo como sendo de classes distintas. Além disso, amostras de classes distintas podem fornecer melhorias (diversidade) significativas e acelerar o aprendizado do classificador.

Se o critério de seleção anterior é satisfeito, o conjunto de amostras candidatas L_C não está vazio (isto é, $L_C \neq \emptyset$, Linha 9). Sendo assim, amostras de L_C são separadas em duas listas temporárias L'_P e L'_N , contendo as imagens classificadas como relevantes (positivas) e irrelevantes (negativas), respectivamente (Linhas 10-11). Em seguida, as amostras de L_P e L_N são ordenadas pela distância em relação à imagem de consulta q (Linhas 12 a 13). Após a ordenação, L_C recebe a concatenação de amostras de L_P e L_N (Linha 14).

Se o número desejado de imagens nu não é obtido (por exemplo, tamanho ($L_C \leq nu$), Linha 16), outro critério é considerado para selecionar as imagens (Linhas 16 a 23). Os centros de massa com_i a partir do espaço de características das amostras de cada classe i são localizados e armazenados em suas listas correspondentes de centros de massa $Lcom_i$ (Linha 17). Para identificar cada com_i , realiza-se um cálculo considerando a média das imagens da i -ésima classe. Posteriormente, amostras de L_i , $i = \{1, 2\}$ (isto é, amostras positivas e negativas) são organizadas em suas respectivas listas (L_P e L_N), em ordem decrescente, de acordo com as distâncias entre elas e o com_i , a partir de $Lcom_i$, (Linhas 18-19).

Obtém-se, então, uma lista de amostras candidatas L_C , selecionando alternadamente uma imagem de cada lista L_P e L_N , respectivamente, até que as nu imagens sejam obtidas (Linhas 20 a 22). Finalmente, a estratégia de seleção ourAL-3 retorna a lista L_C contendo as imagens candidatas mais informativas (mais incertas e similares) para anotação do especialista e posterior utilização no treinamento do classificador.

4.2.3 DESCRIÇÃO DOS CONJUNTOS DE DADOS

Para avaliação das estratégias propostas foram considerados conjuntos de imagens de diferentes domínios de aplicação. Descrições mais detalhadas de cada conjunto são apresentadas a seguir.

- **ETH-80** (LEIBE; SCHIELE, 2003): composto por 2.384 imagens e 8 classes, dentre elas: maçãs, peras, tomates, vacas, cães, cavalos, copos e carros. O conjunto é balanceado, contendo 298 imagens em cada classe.
- **Cevada**: composto por 2.200 imagens de cevada. As imagens foram adquiridas pelo Departamento de Ciência e Tecnologia de Alimentos na Universidade Estadual de Londrina (UEL). Para análise da qualidade da cevada, são considerados dois conjuntos (Cevada₁ e Cevada₂) nos experimentos realizados. No conjunto Cevada₁ são consideradas 2 classes (“Cervejeira” e “SementeNua”), sendo 1.400 amostras para a classe Cervejeira e 800 amostras para a classe SementeNua. No conjunto Cevada₂ são consideradas 22 classes com 100 amostras cada, sendo divididas em 14 classes de tipos de cervejeira e 8 classes de tipos de semente nua.
- **Soja**: composto por imagens de sementes de soja submetidas ao teste de tetrazólio para avaliação do vigor das mesmas. Para os experimentos, foram considerados 4 conjuntos de imagens de soja. A Tabela 13 apresenta a descrição de cada um deles, incluindo a quantidade de imagens pertencentes a cada conjunto, quantidade de classes e intervalo de variação de quantidade de imagens por classe. Notam-se os desafios e desbalanceamentos característicos de problemas e dados reais.

Tabela 13: Descrição dos conjuntos de imagens de semente de soja utilizados nos experimentos

Conjunto	Imagens	Classes	Imagens/classes
Soja ₁	576	12	15-74
Soja ₂	350	7	5-103
Soja ₃	691	7	6-199
Soja ₄	1.333	10	7-529

O conjunto Soja₁ foi obtido a partir de Santanna et al. (2014), considera-se neste conjunto cada parte de uma semente sendo uma amostra de uma classe de dano. Os conjuntos Soja₂, Soja₃ e Soja₄ foram criados por meio de captura in-loco em um laboratório de análise de sementes, pertencente a uma unidade de beneficiamento de sementes da empresa Belagrícola. Os conjuntos Soja₂ e Soja₃ foram criados considerando a junção das 4 partes de uma semente como sendo uma amostra e atribuindo-na um rótulo referente a

uma classe de dano. O conjunto Soja₄ considera cada parte de uma semente como sendo uma amostra associada a uma classe de dano.

4.2.4 DESCRIÇÃO DOS EXPERIMENTOS

Com o intuito de analisar e validar a abordagem de aprendizado, incluindo as estratégias de seleção propostas, ourAL-2 (Algoritmo 7) e ourAL-3 (Algoritmo 8), foram realizados testes comparativos utilizando três abordagens do estado da arte amplamente utilizadas na literatura: movimentação de centro de consulta (*Query Point Movement Strategy* - QPM) (LIU et al., 2009), expansão de consulta (*Query Expansion* - QEX) (PORKAEW; CHAKRA-BARTI, 1999), e aprendizado ativo baseado em SVMs (SVM-AL) (TONG; CHANG, 2001b). Além disso, como comparativo basal para todas as abordagens elencadas, foram também explicitados os resultados obtidos por meio do processo CBIR tradicional. Para facilitar a leitura e interpretação dos resultados, as abordagens da literatura citadas foram nomeadas para: QPM, QEX, SVM-AL e CBIR-T, respectivamente.

A partir dos conjuntos de imagens descritos na Seção 4.2.3, foram extraídas características de cor e textura, por meio de diferentes extratores, detalhados na Tabela 14. Os extratores MPO e MPOC consideram valores estatísticos, tais como, média, desvio padrão, variância, grau de assimetria, curtose, energia e entropia. Foram considerados diferentes extratores específicos para cada conjunto de imagens (apresentados na Seção 4.2.5). Cada extrator foi avaliado com diferentes funções de distância, dentre elas: L_1 , L_2 , L_∞ , X^2 , *Canberra*, *Jeffrey Divergence* e *dLog* apresentadas em detalhe na Seção 2.1.1. Sendo assim, os melhores descritores (pares extrator de característica e função de distância) são obtidos e utilizados nos processos de aprendizado e consulta por similaridade, considerando os respectivos conjuntos de imagens.

Para o processo de aprendizado do modelo foi utilizado o classificador baseado em k -vizinhos mais próximos (k -NN). Em relação ao processo de agrupamento foi utilizada a técnica k -means. Vale ressaltar que a abordagem proposta pode ser instanciada considerando qualquer classificador supervisionado ou técnica de agrupamento. Com relação à quantidade de amostras selecionadas nu , para a primeira iteração, o valor de k definido para o agrupamento consiste em $1/3$ de nu , já para as amostras mais próximas da imagem de consulta, o valor definido foi de $2/3$ de nu , formando assim a primeira instância do classificador.

Para avaliação dos resultados foram gerados gráficos de Precisão e Revocação (P&R). A precisão de uma consulta pode ser entendida como a fração de imagens relevantes recuperadas, já a revocação como a proporção de imagens relevantes que foram recuperadas por meio da mesma, considerando todas as imagens relevantes no conjunto de dados. Dessa

Tabela 14: Propriedades de cada extrator de características.

Extrator	Categoria	#Caraterísticas
F ₁ - AutoColorCorrelogram (HUANG et al., 1997)	Cor	768
F ₂ - BIC (STEHLLING et al., 2002)	Cor	128
F ₃ - CEDD (CHATZICHRISTOFIS; BOUTALIS, 2008a)	Cor	144
F ₄ - FCTH (CHATZICHRISTOFIS; BOUTALIS, 2008b)	Cor e Textura	192
F ₅ - Gabor (ZHANG et al., 2000)	Textura	60
F ₆ - GCH (STRICKER; ORENGO, 1995)	Cor	255
F ₇ - Haralick (HARALICK et al., 1973)	Textura	14
F ₈ - JCD (União CEDD + FCTH)	Cor e Textura	336
F ₉ - LBP (GUO et al., 2010)	Textura	256
F ₁₀ - LCH (SMITH; CHANG, 1996)	Cor	135
F ₁₁ - Moments (GRAF, 2015)	Textura	4
F ₁₂ - MPO (momentos estatísticos)	Textura	6
F ₁₃ - MPOC (momentos estatísticos - RGB)	Textura	18
F ₁₄ - PHOG (BOSCH et al., 2007)	Textura	40
F ₁₅ - ReferenceColorSimilarity (KRIEGEL et al., 2011)	Cor	77
F ₁₆ - Tamura (TAMURA et al., 1978)	Textura	18

forma, a precisão atinge 100% quando cada imagem recuperada é relevante para a consulta, e a revocação atinge 100% quando todas as imagens relevantes do conjunto de dados foram retornadas (BAEZA-YATES; RIBEIRO-NETO, 2011). Como regra geral, quanto mais próxima a curva de P&R do topo do gráfico, melhor é a abordagem.

Para construir os gráficos P&R foram realizadas diversas consultas por similaridade, utilizando o operador baseado em k -vizinhos mais próximos. Vale ressaltar que todas as imagens de cada um dos conjuntos foram utilizadas como imagens de consulta. A quantidade de imagens retornadas pelas consultas por similaridade foi definida como 30, ou seja, . Tal valor foi especificado para estar em concordância com a prática de análise diária realizada em laboratórios e amplamente adotada em trabalhos da literatura. Podem ocorrer casos nos quais uma dada classe de imagens não possua 30 amostras. Dessa forma, em tal ocorrência, o valor nu é definido de acordo com o número de amostras dessa classe. Com o intuito de sumarizar os resultados obtidos, foram gerados gráficos de P&R por meio do conceito de *Mean Average Precision* - MAP, conforme definido em (BAEZA-YATES; RIBEIRO-NETO, 2011).

4.2.5 RESULTADOS E DISCUSSÕES

A análise do extrator e função de distância que melhor descreve cada conjunto de dados é realizada e apresentada (Tabelas 15-17). Após tal análise, são apresentados os resultados obtidos a partir das consultas por similaridade (curvas P&R) de cada abordagem, considerando cada conjunto de dados e seu respectivo melhor descritor (Figuras 21-27).

Cada gráfico das Figuras representa as iterações (primeira, terceira, quinta e oitava) de realimentação de relevância das abordagens propostas ourAL-2 e ourAL-3 em comparação com as abordagens tradicionais CBIR-T, QPM, QEX e SVM-AL. Os resultados para a abordagem CBIR tradicional foram apresentados apenas como comparativo base em todos os gráficos, uma vez que não passa pelos refinamentos da consulta inicial.

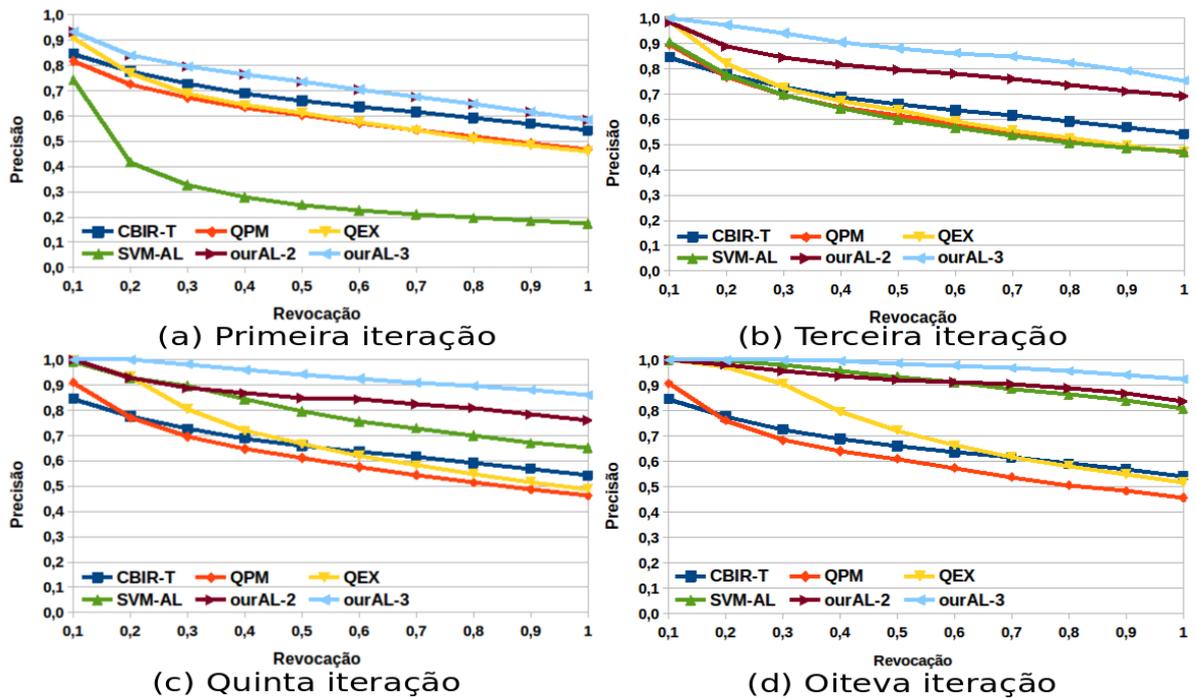
Para o conjunto de imagens ETH-80 foram utilizados 16 extratores de características apresentados na Tabela 14. A Tabela 15 apresenta a melhor função de distância para cada extrator de característica. Destaca-se também em negrito o melhor descritor, sendo considerado para esse conjunto, como melhor extrator JCD e melhor função de distância L_2 .

Tabela 15: Melhor descritor (par extrator de característica e função de distância) para o conjunto de imagens ETH-80.

Extrator	Função de distância
F ₁ - AutoColorCorrelogram	<i>Canberra</i>
F ₂ - BIC	<i>Canberra</i>
F ₃ - CEDD	L_2
F ₄ - FCTH	L_2
F ₅ - Gabor	<i>Canberra</i>
F ₆ - GCH	<i>Canberra</i>
F ₇ - Haralick	<i>Jeffrey Divergence</i>
F₈ - JCD	L_2
F ₉ - LBP	L_1
F ₁₀ - LCH	<i>Canberra</i>
F ₁₁ - Moments	<i>Jeffrey Divergence</i>
F ₁₂ - MPO	<i>Canberra</i>
F ₁₃ - MPOC	<i>dLog</i>
F ₁₄ - PHOG	<i>Canberra</i>
F ₁₅ - ReferenceColorSimilarity	L_∞
F ₁₆ - Tamura	<i>Canberra</i>

Analisando os resultados (curvas P&R) obtidos pelas abordagens, utilizando o conjunto de imagens ETH-80 (Figura 21), pode-se observar que as abordagens ourAL-3 e ourAL-2 apresentam as melhores precisões em todas as iterações, quando comparadas com as outras abordagens em análise. ourAL-3 obteve ganhos (2, 3 vezes) em relação a SVM-AL na primeira iteração, bem como obteve ganhos de precisão de até 60,01%, 37,77% e 14,33%, considerando, respectivamente, a terceira, quinta e oitava iterações, quando a revocação é igual a 100%. Em relação à abordagem QEX, ourAL-3 apresentou ganhos de até 27,19%, 60,21%, 75,78% e 78,44% para as mesmas iterações. Ao comparar ourAL-3 com a abordagem QPM, novamente foram obtidos ganhos consideráveis de até 25,06%, 60,93%, 86,25% e 102% na precisão das consultas.

Figura 21: Curvas P&R obtidas pelas abordagens, utilizando o conjunto de dados ETH-80, considerando: (a) primeira, (b) terceira, (c) quinta e (d) oitava iteração.



Fonte: Autoria própria.

Para o conjunto de imagens de cevada, considerando os dois experimentos ($Cevada_1$ e $Cevada_2$), os melhores descritores (destacados em negrito) são apresentados na Tabela 16. Para ambos, $Cevada_1$ e $Cevada_2$, a melhor combinação consiste na utilização do extrator *Reference Color Similarity* e da função de distância L_∞ . Sendo assim, tal combinação foi aplicada nos experimentos (Figuras 22 e 23). É possível observar que para os conjuntos de imagens de cevada, ourAL-2 e ourAL-3 apresentarem melhores precisões quando comparadas com as demais abordagens. Em geral, ourAL-3 apresentou os melhores resultados.

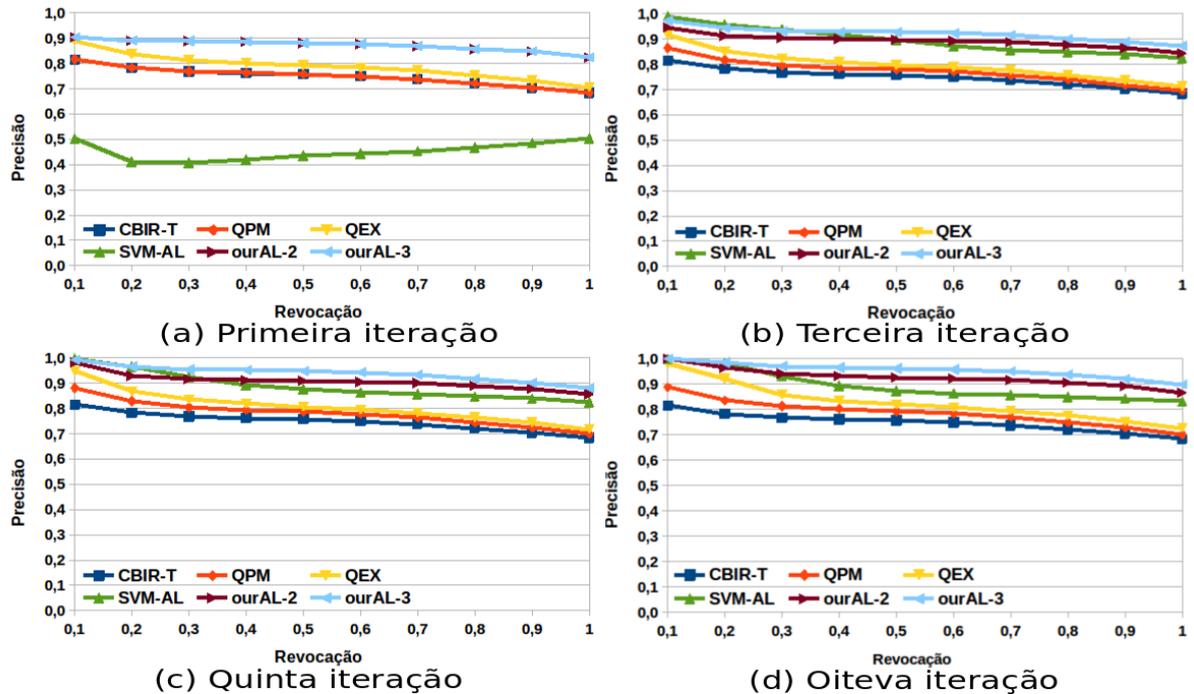
Nas Figuras 22(a)-22(d), correspondentes à primeira, terceira, quinta e oitava iterações, respectivamente, com revocação 100%, ourAL-3 obtém ganhos de 63,28%, 5,93%, 6,85% e 8,02% em relação a SVM-AL; de 16,75%, 22,63%, 22,73% e 23,75% em relação a QEX; de 20,26%, 25,69%, 26,23% e 28,07% em relação a QPM e, de 0%, 3,53%, 2,76%, 3,75% em relação a ourAL-2. Nas Figuras 23(a)-23(d), pode-se observar ganhos de até 75,25% em relação a SVM-AL.

Considerando os conjuntos de imagens de soja ($Soja_1$, $Soja_2$, $Soja_3$ e $Soja_4$), a Tabela 17 apresenta os melhores descritores para cada um deles, sendo BIC e Jeffrey Divergence; LBP e dLog; GCH e L_∞ e, BIC e Canberra, respectivamente. De maneira geral, o mesmo comportamento apresentado pelas abordagens pode ser observado (Figura 24(a)-(d) a Figura

Tabela 16: Melhor descritor (par extrator de característica e função de distância) para os conjuntos de imagens Cevada₁ e Cevada₂.

Extrator	Cevada ₁	Cevada ₂
F ₁ - AutoColorCorrelogram	<i>Jeffrey Divergence</i>	X^2
F ₂ - BIC	<i>Jeffrey Divergence</i>	L_1
F ₃ - CEDD	<i>Jeffrey Divergence</i>	L_1
F ₄ - FCTH	L_∞	L_∞
F ₅ - Gabor	L_2	<i>dLog</i>
F ₆ - GCH	<i>Jeffrey Divergence</i>	<i>Canberra</i>
F ₇ - Haralick	L_∞	L_∞
F ₈ - JCD	L_∞	L_1
F ₉ - LBP	<i>Canberra</i>	<i>Canberra</i>
F ₁₀ - LCH	<i>Jeffrey Divergence</i>	<i>Canberra</i>
F ₁₁ - Moments	L_2	L_2
F ₁₂ - MPO	L_∞	L_∞
F ₁₃ - MPOC	L_∞	L_∞
F ₁₄ - PHOG	L_∞	<i>Canberra</i>
F₁₅ - ReferenceColorSimilarity	L_∞	L_∞
F ₁₆ - Tamura	L_1	<i>Jeffrey Divergence</i>

Figura 22: Curvas P&R obtidas pelas abordagens, utilizando o conjunto de dados Cevada₁, considerando: (a) primeira, (b) terceira, (c) quinta e (d) oitava iteração.

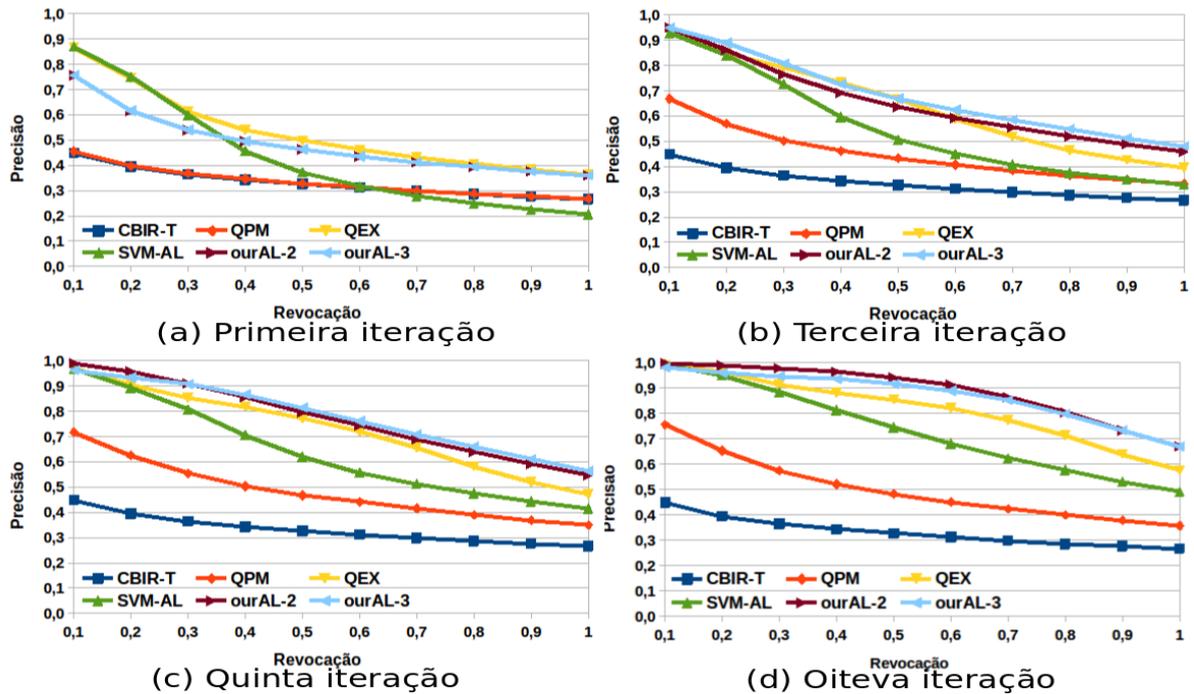


Fonte: Autoria própria.

27(a)-(d)), considerando os conjuntos de imagens de soja.

As abordagens propostas ourAL-2 e ourAL-3 (Figuras 24 e 25) apresentaram um de-

Figura 23: Curvas P&R obtidas pelas abordagens, utilizando o conjunto de dados Cevada₂, considerando: (a) primeira, (b) terceira, (c) quinta e (d) oitava iteração.



Fonte: Autoria própria.

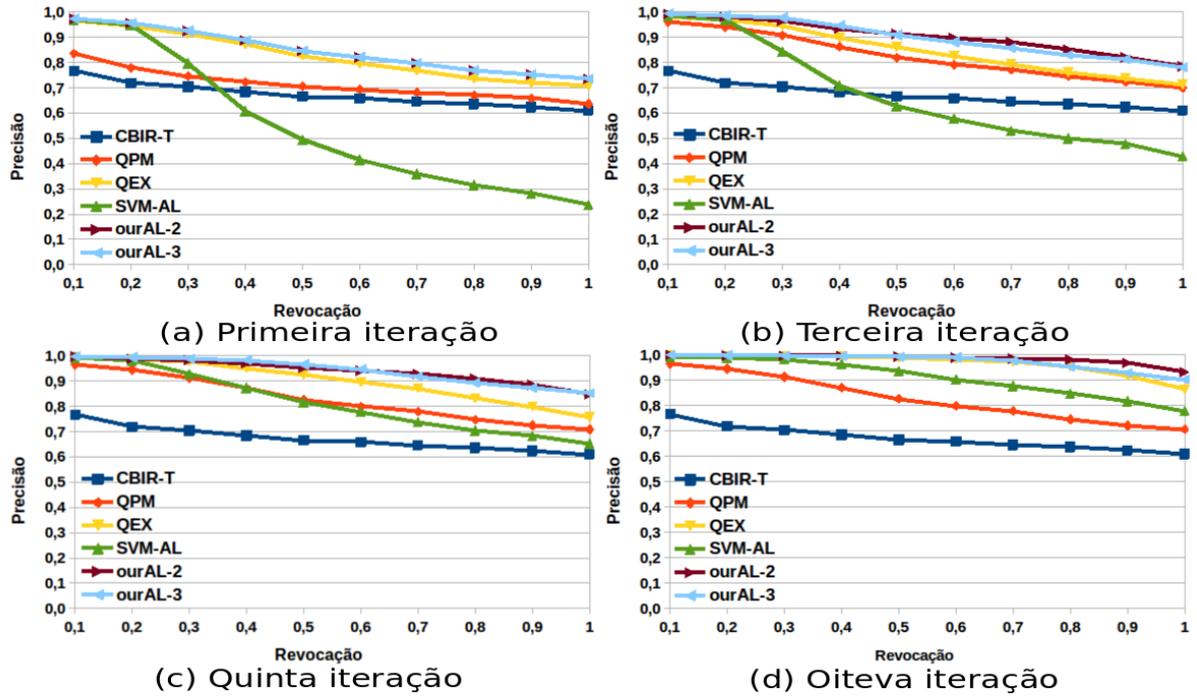
sempenho melhor tanto em relação às abordagens tradicionais (CBIR-T, QPM e QEX), as quais não consideram estratégias de aprendizado ativo no processo de realimentação, como também em relação à SVM-AL, a qual considera seleção de amostras mais informativas no processo. No entanto, nesse caso, o desempenho de ourAL-2 e ourAL-3 são equivalentes.

Tabela 17: Melhor descritor (par extrator de característica e função de distância) para os conjuntos de imagens Soja₁, Soja₂, Soja₃ e Soja₄.

Extrator	Soja ₁	Soja ₂	Soja ₃	Soja ₄
F ₂ - BIC	<i>Jeffrey Divergence</i>	<i>dLog</i>	<i>dLog</i>	<i>Canberra</i>
F ₆ - GCH	X^2	L_∞	L_∞	<i>Jeffrey Divergence</i>
F ₉ - LBP	<i>dLog</i>	<i>dLog</i>	<i>dLog</i>	L_2
F ₁₀ - LCH	X^2	<i>Canberra</i>	<i>Canberra</i>	<i>Canberra</i>

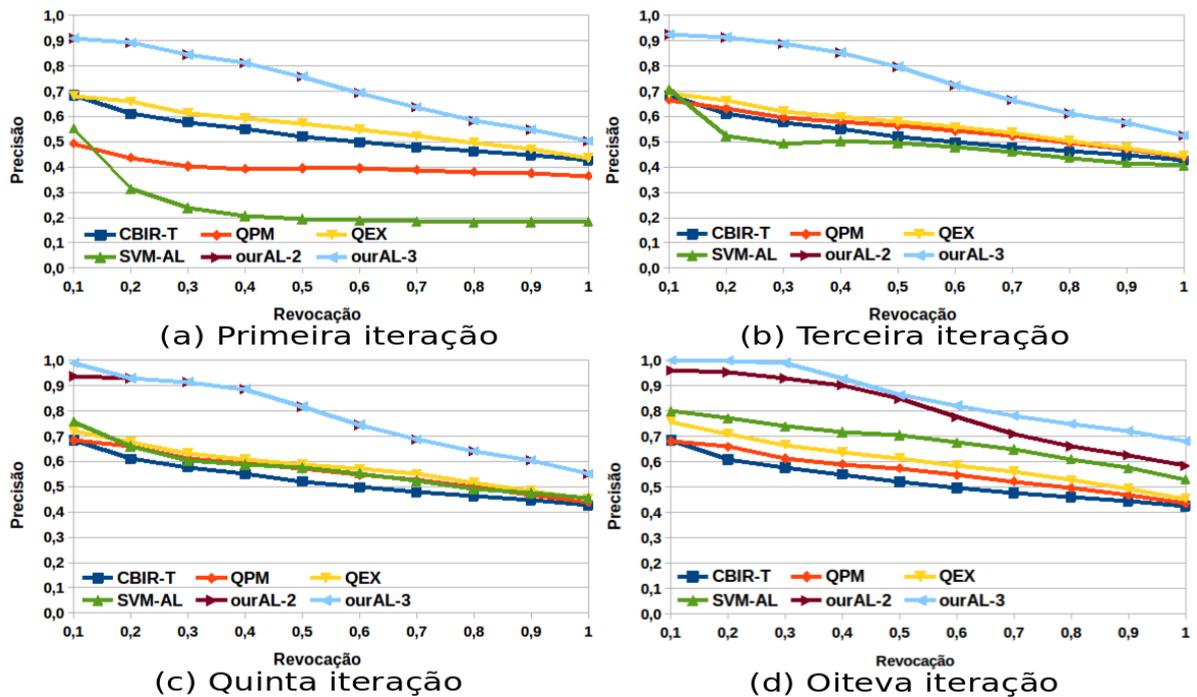
Na Figura 26, para o conjunto de imagens Soja₃, as abordagens propostas ourAL-2 e ourAL-3 também obtiveram bons resultados ao longo das oito iterações de aprendizado. Nesse cenário, ourAL-3 apresentou melhores precisões dentre todas as abordagens. Ao se analisar a revocação em 10%, nas sucessivas iterações (Figuras 26(a)-(d)), ourAL-3 atingiu ganhos de no mínimo 3.98% quando comparada com QEX na primeira iteração e no máximo 66,27% quando comparada com QPM na oitava iteração. Vale enfatizar que níveis de revocação mais baixos são de maior importância, uma vez que aproximam-se mais de análises cotidianas realizadas. Com

Figura 24: Curvas P&R obtidas pelas abordagens, utilizando o conjunto de dados Soja₁, considerando: (a) primeira, (b) terceira, (c) quinta e (d) oitava iteração.



Fonte: Autoria própria.

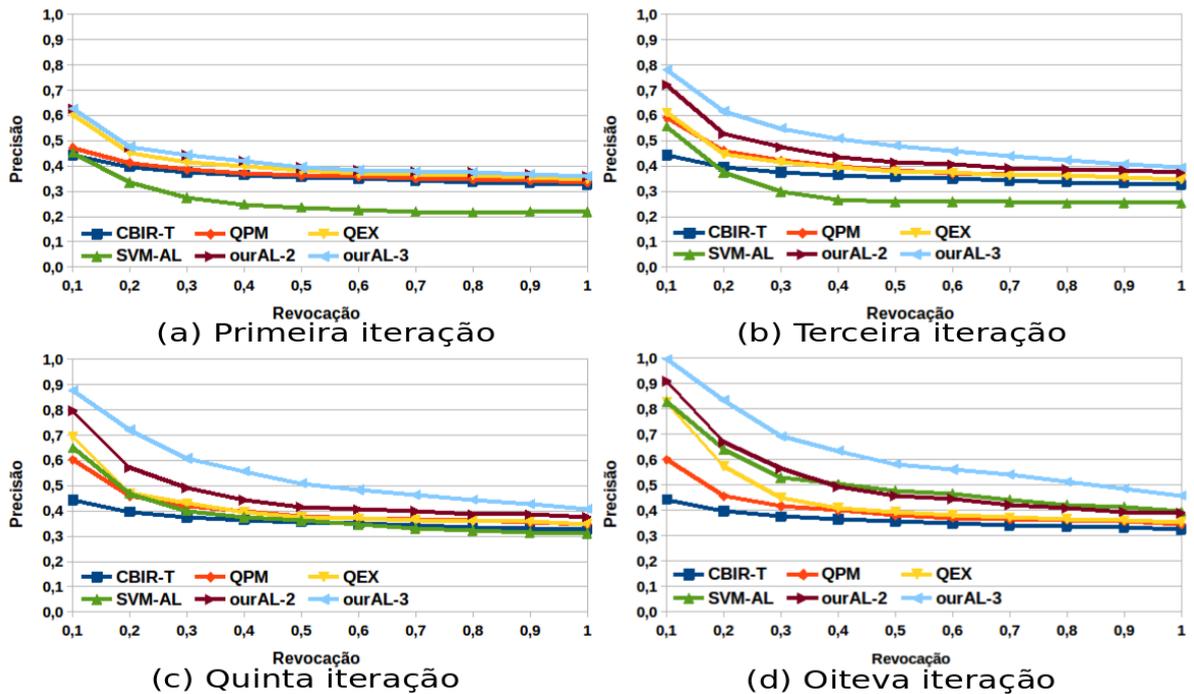
Figura 25: Curvas P&R obtidas pelas abordagens, utilizando o conjunto de dados Soja₂, considerando: (a) primeira, (b) terceira, (c) quinta e (d) oitava iteração.



Fonte: Autoria própria.

relação à abordagem que seleciona amostras mais informativas SVM-AL, ourAL-3 apresenta ganhos de 38,29%, 40,29%, 34,93% e 20,03% ao longo da primeira, terceira, quinta e oitava iteração, respectivamente. ourAL-3 também apresenta melhores precisões na terceira, quinta e oitava iteração em relação a ourAL-2 com valores de 8,37%, 10,09% e 9,64% respectivamente.

Figura 26: Curvas P&R obtidas pelas abordagens, utilizando o conjunto de dados Soja₃, considerando: (a) primeira, (b) terceira, (c) quinta e (d) oitava iteração.

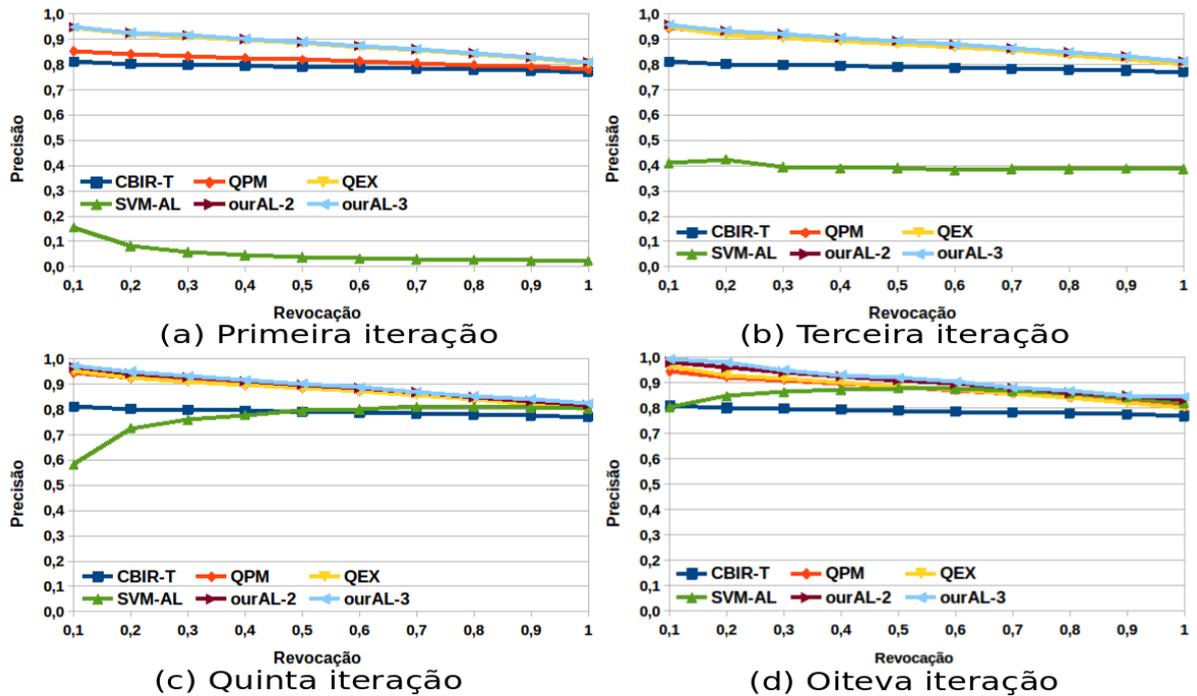


Fonte: Autoria própria.

Para o conjunto Soja₄ (Figura 27), as abordagens ourAL-2 e ourAL-3 novamente apresentaram ganhos significativos de precisão em comparação com as demais abordagens. Por exemplo, considerando um nível de revocação de 10%, a abordagem ourAL-3 foi 5,09 vezes melhor que a abordagem SVM-AL na primeira iteração. Em relação a QPM obteve um ganho de até 11,07%, considerando a primeira iteração. No entanto, nesse cenário, em relação a QEX apresentou desempenho equivalente nas primeiras iterações e ganhos de até 2,87% na oitava iteração.

A partir dos resultados obtidos, pode-se concluir que as abordagens propostas apresentaram consideráveis ganhos de precisão em relação às consultas por similaridade em conjuntos de imagens provenientes de domínios de aplicação distintos. Além disso, a abordagem ourAL-3 também apresenta um grau menor de saturação no decorrer das subsequentes iterações de realimentação em todos os conjuntos de imagens analisados. Com o intuito de resumir os resultados, na Tabela 18 são apresentadas as precisões médias gerais obtidas por cada abordagem

Figura 27: Curvas P&R obtidas pelas abordagens, utilizando o conjunto de dados Soja₄, considerando: (a) primeira, (b) terceira, (c) quinta e (d) oitava iteração.



Fonte: Autoria própria.

e conjunto de dados utilizados, considerando todas as iterações de realimentação (i.e. da 1^a até a 8^a iteração). Para a obtenção de tal precisão média geral foi calculada a integral das curvas de P&R para cada abordagem em cada iteração e realizada a média de tais integrais, como definido em (BAEZA-YATES; RIBEIRO-NETO, 2011). Ao analisar os resultados sumarizados na Tabela 18, nota-se que, em geral, a abordagem proposta ourAL-3 atingiu os melhores níveis de precisão no decorrer das iterações, mantendo maior consistência e robustez.

Tabela 18: Precisões médias gerais obtidas por cada abordagem nos conjuntos de imagens, considerando todas as iterações de realimentação.

Conjunto	QPM	QEX	SVM-AL	ourAL-2	ourAL-3
ETH-80	61,86	67,63	70,30	83,46	89,53
Cevada ₁	77,44	80,59	83,49	90,23	92,81
Cevada ₂	45,35	69,04	60,30	72,01	72,59
Soja ₁	80,88	88,79	76,40	92,28	92,29
Soja ₂	53,08	57,71	51,80	76,11	77,75
Soja ₃	40,07	41,84	37,44	46,48	53,54
Soja ₄	86,69	57,58	60,20	89,02	89,32

5 CONSIDERAÇÕES FINAIS

Esta pesquisa possibilitou o estudo, o desenvolvimento e a validação de técnicas efetivas de aprendizado ativo para classificação e recuperação de imagens, de forma a acelerar os processos de aprendizado e recuperação, obtendo amostras mais relevantes, bem como minimizando a interação do especialista nestes processos.

Para a tarefa de classificação, duas principais contribuições foram apresentadas (Capítulo 3). A primeira, uma metodologia (CAMARGO et al., 2017) para avaliação de desempenho do aprendizado de alguns classificadores em diferentes conjuntos biomédicos. Nesta proposta são consideradas regras de relação de compromisso na escolha do melhor modelo de classificação, levando-se em consideração valores de acurácias e de tempos computacionais. Para avaliação experimental, diferentes classificadores supervisionados foram aplicados em nove conjuntos biomédicos públicos. Dentre eles, o classificador OPF foi o modelo mais eficaz e eficiente, apresentando os melhores resultados (menores valores de médias harmônicas) em todos os conjuntos avaliados, ou seja, apresentou elevadas acurácias de classificação e menores tempos de treinamento e de teste. A metodologia proposta mostra-se como uma contribuição relevante, dado que possibilita analisar o desempenho de cada modelo de classificação em termos de escalabilidade, bem como determinadas restrições de tempo e de recursos computacionais requeridas por aplicações reais.

A segunda contribuição, na área de classificação, refere-se à inclusão de estratégias de aprendizado ativo no processo de aprendizado do classificador, o qual a medida que aprende, seleciona e sugere rótulos para a anotação das amostras que são mais difíceis de classificação, minimizando o esforço de anotação manual de todo o conjunto de dados fornecido pelo especialista, conforme requerido no processo de aprendizado supervisionado tradicional. Uma nova estratégia de aprendizado ativo é proposta (BRESSAN et al., 2018c) (Seção 3.2), a qual possibilita explorar melhor as amostras mais informativas para o aprendizado do classificador, por meio de critérios de seleção com base em representatividade e incerteza das amostras. Para validação da estratégia foi realizada uma avaliação extensiva, realizando comparações entre estratégias de aprendizado da literatura, utilizando diferentes técnicas de classificação e conjuntos

de dados. A partir dos resultados obtidos, é possível observar que a estratégia de seleção proposta, de forma geral, atinge elevadas acurácias mais rapidamente, nas primeiras iterações de aprendizado em relação às demais estratégias comparadas.

Para a área de recuperação de imagens por conteúdo são apresentadas três principais contribuições (Capítulo 4). A primeira inclui a proposta de uma estratégia para avaliação de desempenho de descritores (BRESSAN et al., 2018a), avaliando o papel e o impacto das características profundas, obtidas por meio de aprendizado profundo, na recuperação de imagens médicas baseadas em conteúdo. Analisando os resultados, é possível utilizar técnicas de *transfer learning* para realizar o aprendizado a partir de um contexto geral para um contexto médico específico. No entanto, para aproveitar ao máximo esse princípio, o processo de recuperação deve ser realizado com características gerais profundas em um determinado processo de refinamento da consulta.

Sendo assim, para melhorar o processo de refinamento das consultas, são apresentadas duas novas contribuições, as quais referem-se a estratégias de aprendizado ativo para realimentação de relevância em CBIR. As propostas (BRESSAN et al., 2018d, 2018b) têm como objetivo a obtenção de características e amostras mais informativas (similares e incertas) a cada iteração de realimentação do sistema, de forma a minimizar a interação do especialista e a acelerar o processo de aprendizado, obtendo um classificador robusto mais rapidamente, e conseqüentemente, melhorando a qualidade das imagens retornadas, mais similares em relação à imagem de consulta e à expectativa do especialista.

Em trabalhos futuros pretende-se realizar a análise de diferentes descritores de imagens, de forma a prover melhorias na descrição das mesmas, obtidas e relacionadas às contribuições apresentadas em Bressan et al. (2018a, 2018d, 2018b, 2018c), Junior et al. (2018), Camargo et al. (2017). Pretende-se também realizar a avaliação de diferentes classificadores supervisionados e técnicas de agrupamentos, gerando diferentes instâncias das propostas. Além disso, vislumbra-se também o desenvolvimento de outras estratégias de aprendizado ativo para otimizar o processo de seleção de amostras mais informativas, de forma a possibilitar a obtenção de melhorias na qualidade dos resultados retornados a partir das consultas por similaridade.

REFERÊNCIAS

- AHA, D. W.; KIBLER, D.; ALBERT, M. K. Instance-based learning algorithms. **Machine learning**, Springer, v. 6, n. 1, p. 37–66, 1991.
- ALAJLAN, N. et al. Large-scale image classification using active learning. **IEEE Geosci. Remote Sensing Lett.**, v. 11, n. 1, p. 259–263, 2014. Disponível em: <<http://dblp.uni-trier.de/db/journals/lgrs/lgrs11.html>>.
- ALMEIDA, J. et al. Making colors worth more than a thousand words. In: ACM. **Proceedings of the 2008 ACM symposium on Applied computing**. [S.l.], 2008. p. 1180–1186.
- ALON, U. et al. Broad patterns of gene expression revealed by clustering analysis of tumor and normal colon tissues probed by oligonucleotide arrays. **National Academy of Sciences USA**, v. 96, n. 12, p. 6745–6750, 1999.
- AMATO, G. et al. Yfcc100m hybridnet fc6 deep features for content-based image retrieval. In: **Proceedings of the 2016 ACM Workshop on Multimedia COMMONS**. New York, NY, USA: ACM, 2016. (MMCommons '16), p. 11–18. ISBN 978-1-4503-4515-6. Disponível em: <<http://doi.acm.org/10.1145/2983554.2983557>>.
- ANH, N. D. et al. A new cbir system using sift combined with neural network and graph-based segmentation. In: SPRINGER. **Asian Conference on Intelligent Information and Database Systems**. [S.l.], 2010. p. 294–301.
- ARMSTRONG, S. A. et al. Mll translocations specify a distinct gene expression profile that distinguishes a unique leukemia. **Nature Genetics**, Nature Publishing Group, v. 30, n. 1, p. 41, 2002.
- BAEZA-YATES, R.; RIBEIRO-NETO, B. **Modern Information Retrieval: The Concepts and Technology Behind Search**. [S.l.]: Addison-Wesley Publ. Comp., 2nd ed., 2011.
- BATISTA, G. E. d. A. P. et al. **Pré-processamento de dados em aprendizado de máquina supervisionado**. Tese (Doutorado) — Universidade de São Paulo, 2003.
- BAYES, T. An essay towards solving a problem in the doctrine of chances. **Phil. Trans. of the Royal Society of London**, v. 53, p. 370–418, 1763.
- BIMBO, A. D.; PALA, P. Visual image retrieval by elastic matching of user sketches. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, IEEE, v. 19, n. 2, p. 121–132, 1997.
- BOSCH, A.; ZISSERMAN, A.; MUNOZ, X. Representing shape with a spatial pyramid kernel. In: ACM. **Proceedings of the 6th ACM international conference on Image and video retrieval**. [S.l.], 2007. p. 401–408.

- BOSER, B. E.; GUYON, I. M.; VAPNIK, V. N. A training algorithm for optimal margin classifiers. In: **Fifth Annual Workshop on Computational Learning Theory**. [S.l.: s.n.], 1992. p. 144–152.
- BREIMAN, L. Random forests. **Machine learning**, Springer, v. 45, n. 1, p. 5–32, 2001.
- BRESSAN, R. S. et al. Doctor: The role of deep features in content-based mammographic image retrieval. In: **2018 IEEE 31st International Symposium on Computer-Based Medical Systems (CBMS)**. [S.l.: s.n.], 2018. p. 158–163.
- BRESSAN, R. S.; BUGATTI, P. H.; SAITO, P. T. M. Marrow: Active learning strategy in content-based image retrieval for breast cancer diagnosis (under review). **Journal of Biomedical Informatics (JBI)**, Elsevier Science, 2018.
- BRESSAN, R. S. et al. Exploring active learning based on representativeness and uncertainty for biomedical data classification (under review). **IEEE Journal of Biomedical and Health Informatics**, 2018.
- BRESSAN, R. S. et al. Soyretrieval - técnicas de aprendizado e recuperação de imagens para análise do vigor de sementes de soja. In: **45º Seminário Integrado de Software e Hardware 2018 (SEMISH)**. [s.n.], 2018. v. 45, n. 1/2018. ISSN 2595-6205. Disponível em: <<http://portaldeconteudo.sbc.org.br/index.php/semish/article/view/3434>>.
- BUGATTI, P. H. **Análise da influência de funções de distância para o processamento de consultas por similaridade em recuperação de imagens por conteúdo**. Dissertação (Mestrado) — Universidade de São Paulo, 2008.
- BUGATTI, P. H. **Arcabouço para recuperação de imagens por conteúdo visando à percepção do usuário**. Tese (Doutorado) — Universidade de São Paulo, 2012.
- CAMARGO, G. et al. Towards an effective and efficient learning for biomedical data classification. In: **IEEE. Computer-Based Medical Systems (CBMS), 2017 IEEE 30th International Symposium on**. [S.l.], 2017. p. 13–18.
- CHANG, C.-C.; LIN, C.-J. LIBSVM: A library for support vector machines. **ACM Transactions on Intelligent Systems and Technology**, v. 2, p. 27:1–27:27, 2011. Library available at <http://www.csie.ntu.edu.tw/~cjlin/libsvm>.
- CHATZICHRISTOFIS, S. A.; BOUTALIS, Y. S. Cedd: color and edge directivity descriptor: a compact descriptor for image indexing and retrieval. In: **SPRINGER. International Conference on Computer Vision Systems**. [S.l.], 2008. p. 312–322.
- CHATZICHRISTOFIS, S. A.; BOUTALIS, Y. S. Fcth: Fuzzy color and texture histogram—a low level feature for accurate image retrieval. In: **IEEE. Image Analysis for Multimedia Interactive Services, 2008. WIAMIS'08. Ninth International Workshop on**. [S.l.], 2008. p. 191–196.
- COX, I. J. et al. The bayesian image retrieval system, pichunter: theory, implementation, and psychophysical experiments. **IEEE transactions on image processing**, IEEE, v. 9, n. 1, p. 20–37, 2000.
- DIETTERICH, T. G. Approximate statistical tests for comparing supervised classification learning algorithms. **Neural computation**, MIT Press, v. 10, n. 7, p. 1895–1923, 1998.

DUDA, R. O.; HART, P. E.; STORK, D. G. **Pattern classification**. [S.l.]: John Wiley & Sons, 2012.

EAKINS, J.; GRAHAM, M.; FRANKLIN, T. **Content-based image retrieval**. [S.l.]: Library and Information Briefings, 1991. 85 1-15 p.

GALLOWAY, M. M. Texture analysis using gray level run lengths. **Computer Graphics and Image Processing**, v. 4, n. 2, p. 172 – 179, 1975. ISSN 0146-664X. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0146664X75800086>>.

GARCIA, E. et al. Automatic lymphocyte detection on gastric cancer ihc images using deep learning. In: **IEEE CBMS**. [S.l.: s.n.], 2017. p. 200–204.

GOLUB, T. R. et al. Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. **Science**, v. 286, p. 531–537, 1999.

GONZALEZ, R. C.; WOODS, R. E. **Processamento Digital de Imagens**. 3. ed. [S.l.]: Pearson, 2010. 624 p.

GORDON, G. J. et al. Translation of microarray data into clinically relevant cancer diagnostic tests using gene expression ratios in lung cancer and mesothelioma. **Cancer Research**, v. 62, n. 17, p. 4963–4967, 2002.

GRAF, F. **JFeatureLib v1.6.3**. out. 2015. Disponível em: <<http://dx.doi.org/10.5281/zenodo.31793>>.

GUO, Z.; ZHANG, L.; ZHANG, D. Rotation invariant texture classification using LBP variance (LBPV) with global matching. **PR**, Elsevier, v. 43, n. 3, p. 706–719, 2010.

GUPTA, A.; JAIN, R. Visual information retrieval. **Commun. ACM**, ACM, New York, NY, USA, v. 40, n. 5, p. 70–79, maio 1997. ISSN 0001-0782. Disponível em: <<http://doi.acm.org/10.1145/253769.253798>>.

HARALICK, R. M.; SHANMUGAM, K.; DINSTEN, I. Textural features for image classification. **IEEE Transactions on Systems, Man, and Cybernetics**, SMC-3, n. 6, p. 610–621, Nov 1973. ISSN 0018-9472.

HE, D.-C.; WANG, L. Texture unit, texture spectrum, and texture analysis. **IEEE transactions on Geoscience and Remote Sensing**, IEEE, v. 28, n. 4, p. 509–512, 1990.

HE, J. et al. Mean version space: A new active learning method for content-based image retrieval. In: **Proceedings of the 6th ACM SIGMM International Workshop on Multimedia Information Retrieval**. ACM, 2004. (MIR '04), p. 15–22. ISBN 1-58113-940-3. Disponível em: <<http://doi.acm.org/10.1145/1026711.1026715>>.

HE, K. et al. Deep residual learning for image recognition. In: **IEEE CVPR**. [S.l.: s.n.], 2016. p. 770–778.

HE, X. et al. Laplacian optimal design for image retrieval. In: **Proceedings of the 30th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval**. New York, NY, USA: ACM, 2007. (SIGIR '07), p. 119–126. ISBN 978-1-59593-597-7. Disponível em: <<http://doi.acm.org/10.1145/1277741.1277764>>.

- HOI, S. C. H. et al. Semisupervised svm batch mode active learning with applications to image retrieval. **ACM Trans. Inf. Syst.**, ACM, New York, NY, USA, v. 27, n. 3, p. 16:1–16:29, maio 2009. ISSN 1046-8188. Disponível em: <<http://doi.acm.org/10.1145/1508850.1508854>>.
- HUANG, J. et al. Image indexing using color correlograms. In: IEEE. **Computer Vision and Pattern Recognition, 1997. Proceedings., 1997 IEEE Computer Society Conference on**. [S.l.], 1997. p. 762–768.
- HUISKES, M. J.; LEW, M. S. Performance evaluation of relevance feedback methods. In: **Proceedings of the 2008 International Conference on Content-based Image and Video Retrieval**. ACM, 2008. (CIVR '08), p. 239–248. ISBN 978-1-60558-070-8. Disponível em: <<http://doi.acm.org/10.1145/1386352.1386387>>.
- IRONS, J. R.; PETERSEN, G. W. Texture transforms of remote sensing data. **Remote Sensing of Environment**, v. 11, p. 359–370, 1981. ISSN 0034-4257. Disponível em: <<http://www.sciencedirect.com/science/article/pii/003442578190033X>>.
- JAIN, A. K.; MURTY, M. N.; FLYNN, P. J. Data clustering: a review. **ACM computing surveys (CSUR)**, Acm, v. 31, n. 3, p. 264–323, 1999.
- JEONG, S. Histogram-based color image retrieval. **Psych221/EE362 Project Report**, 2001.
- JOACHIMS, J. **Making large-scale SVM learning practical**. Scholkopf, C. JC Burges and AJ Smola, editors, **Advances in Kernel Methods-Support Vector Learning, 169-184**. [S.l.]: MIT Press, Cambridge, MA, 1999.
- JUNIOR, M. et al. Rumo à melhoria de produtividade e sustentabilidade agrícola por meio da classificação automática do vigor de sementes de soja. In: **45º Seminário Integrado de Software e Hardware 2018 (SEMISH)**. [s.n.], 2018. v. 45, n. 1/2018. ISSN 2595-6205. Disponível em: <<http://portaldeconteudo.sbc.org.br/index.php/semish/article/view/3428>>.
- JURMAN, G. et al. Canberra distance on ranked lists. In: CITESEER. **Proceedings of Advances in Ranking NIPS 09 Workshop**. [S.l.], 2009. p. 22–27.
- KANG, J.; RYU, K. R.; KWON, H.-C. Using cluster-based sampling to select initial training set for active learning in text classification. In: SPRINGER. **Pacific-Asia Conference on Knowledge Discovery and Data Mining**. [S.l.], 2004. p. 384–388.
- KASAUDHAN, R. et al. Similarity measurement with mesh distance fourier transform in 2d binary image. In: **Proceedings of the 2015 Conference on Research in Adaptive and Convergent Systems**. New York, NY, USA: ACM, 2015. (RACS), p. 183–187. ISBN 978-1-4503-3738-0. Disponível em: <<http://doi.acm.org/10.1145/2811411.2811506>>.
- KAWANO, Y.; YANAI, K. Food image recognition with deep convolutional features. In: ACM. **Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication**. [S.l.], 2014. p. 589–593.
- KHOTANZAD, A.; HONG, Y. H. Invariant image recognition by zernike moments. **IEEE TPAMI**, IEEE, v. 12, n. 5, p. 489–497, 1990.
- KIM, R. et al. An investigation into the relationship between semantic and content based similarity using lidc. In: ACM. **Proceedings of The International Conference on Multimedia Information Retrieval**. [S.l.], 2010. p. 185–192.

- KING, I.; JIN, Z. Integrated probability function and its application to content-based image retrieval by relevance feedback. **Pattern Recognition**, Elsevier, v. 36, n. 9, p. 2177–2186, 2003.
- KLOVE, T. et al. Permutation arrays under the chebyshev distance. **IEEE Transactions on Information Theory**, IEEE, v. 56, n. 6, p. 2611–2617, 2010.
- KOHAVI, R. et al. A study of cross-validation and bootstrap for accuracy estimation and model selection. In: STANFORD, CA. **Ijcai**. [S.l.], 1995. v. 14, n. 2, p. 1137–1145.
- KRIEGEL, H.-P.; SCHUBERT, E.; ZIMEK, A. Evaluation of multiple clustering solutions. In: CITESEER. **MultiClust@ ECML/PKDD**. [S.l.], 2011. p. 55–66.
- LEIBE, B.; SCHIELE, B. Analyzing appearance and contour based methods for object categorization. In: IEEE. **Computer Vision and Pattern Recognition, 2003. Proceedings. 2003 IEEE Computer Society Conference on**. [S.l.], 2003. v. 2, p. II–409.
- LEU, J.-G. Computing a shape's moments from its boundary. **Pattern recognition**, Elsevier, v. 24, n. 10, p. 949–957, 1991.
- LEWIS, D. D.; GALE, W. A. A sequential algorithm for training text classifiers. In: SPRINGER-VERLAG NEW YORK, INC. **Proceedings of the 17th annual international ACM SIGIR conference on Research and development in information retrieval**. [S.l.], 1994. p. 3–12.
- LINDEN, R. Técnicas de agrupamento. **Revista de Sistemas de Informação da FSMA**, v. 1, n. 4, p. 18–36, 2009.
- LIU, D. et al. Fast query point movement techniques for large cbir systems. **IEEE TKDE**, IEEE, v. 21, n. 5, p. 729–743, 2009.
- MALODE, P.; GUMASTE, S. V. A review paper on content based image retrieval. **Intl. Research Journal of Engineering and Technology**, v. 20, p. 883–885, 2015.
- MINSKY, M.; PAPERT, S. **Perceptrons : an introduction to computational geometry**. Cambridge (Mass.), London: The MIT Press, 1969.
- MITCHELL, T. M. et al. **Machine learning. WCB**. [S.l.]: McGraw-Hill Boston, MA., 1997.
- MOKHTARIAN, F.; MACKWORTH, A. Scale-based description and recognition of planar curves and two-dimensional shapes. **IEEE transactions on pattern analysis and machine intelligence**, IEEE, n. 1, p. 34–43, 1986.
- MONARD, M. C.; BARANAUSKAS, J. A. Conceitos sobre aprendizado de máquina. **Sistemas Inteligentes-Fundamentos e Aplicações**, v. 1, n. 1, 2003.
- MORENTE-MOLINERA, J. A. et al. Improving supervised learning classification methods using multi-granular linguistic modelling and fuzzy entropy. **IEEE Transactions on Fuzzy Systems**, IEEE, 2016.
- MORITA, H.; ABE, K.; HAYASHI, T. A similarity retrieval of trademark images considering similarity for local objects using vector images. In: **Proceedings of the 8th International Conference on Signal Processing Systems**. New York, NY, USA: ACM, 2016. (ICSPS 2016), p. 79–82. ISBN 978-1-4503-4790-7. Disponível em: <<http://doi.acm.org/10.1145/3015166.3015194>>.

- MÜLLER, H. et al. A review of content-based image retrieval systems in medical applications-clinical benefits and future directions. **International journal of medical informatics**, Elsevier, v. 73, n. 1, p. 1–23, 2004.
- MUSLEA, I.; MINTON, S.; KNOBLOCK, C. A. Selective sampling with redundant views. In: **AAAI/IAAI**. [S.l.: s.n.], 2000. p. 621–626.
- NGUYEN, H. T.; SMEULDERS, A. Active learning using pre-clustering. In: **ACM. Proceedings of the twenty-first international conference on Machine learning**. [S.l.], 2004. p. 79.
- NIXON, M. S.; AGUADO, A. S. **Feature extraction & image processing for computer vision**. [S.l.]: Academic Press, 2012.
- OJALA, T.; PIETIKAINEN, M.; HARWOOD, D. A comparative study of texture measures with classification based on featured distributions. **Pattern Recognition**, v. 29, n. 1, p. 51–59, 1996. ISSN 0031-3203. Disponível em: <<http://www.sciencedirect.com/science/article/pii/0031320395000674>>.
- OLIVEIRA, P. H. et al. MAMMOSET: An Enhanced Dataset of Mammograms. In: **Satellite Events of the Brazilian Symp. on Databases**. [S.l.]: SBC, 2017. p. 256–266.
- PALLE, A.; KULKARNI, R. B. Classification of medical mri brain images based on hadoop. In: **Proceedings of the Second International Conference on Information and Communication Technology for Competitive Strategies**. New York, NY, USA: ACM, 2016. (ICTCS '16), p. 95:1–95:4. ISBN 978-1-4503-3962-9. Disponível em: <<http://doi.acm.org/10.1145/2905055.2905154>>.
- PAPA, J. P. et al. Efficient supervised optimum-path forest classification for large datasets. **Pattern Recognition**, Elsevier Science Inc., v. 45, p. 512–520, 2012.
- PENATTI, O. A.; VALLE, E.; TORRES, R. d. S. Comparative study of global color and texture descriptors for web image retrieval. **Journal of Visual Communication and Image Representation**, Elsevier, v. 23, n. 2, p. 359–380, 2012.
- PETRICOIN, E. F. et al. Use of proteomic patterns in serum to identify ovarian cancer. **Lancet**, v. 359, n. 9306, p. 572–577, 2002.
- PEURA, M.; IIVARINEN, J. Efficiency of simple shape descriptors. **Aspects of Visual Form**, Citeseer, p. 443–451, 1997.
- POMEROY, S. L. et al. Prediction of central nervous system embryonal tumour outcome based on gene expression. **Nature**, v. 415, p. 436–442, 2002.
- PORKAEW, K.; CHAKRABARTI, K. Query refinement for multimedia similarity retrieval in mars. In: **ACM Intl. Conference on Multimedia**. [S.l.: s.n.], 1999. p. 235–238.
- PUTZU, L.; PIRAS, L.; GIACINTO, G. Ten years of relevance score for content based image retrieval. In: **SPRINGER. International Conference on Machine Learning and Data Mining in Pattern Recognition**. [S.l.], 2018. p. 117–131.
- QUINLAN, J. R. **C4.5: Programs for Machine Learning**. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 1993. ISBN 1-55860-238-0.

- RAO, T. Y. S.; REDDY, P. C. Content and context based image retrieval classification based on firefly-neural network. **Multimedia Tools and Applications**, Springer, p. 1–22, 2018.
- RAO, Y. et al. A novel relevance feedback method for cbir. **World Wide Web**, Springer, p. 1–18, 2018.
- RAZZAK, M. I.; NAZ, S.; ZAIB, A. Deep learning for medical image processing: Overview, challenges and the future. In: **Classification in BioApps**. [S.l.]: Springer, 2018. p. 323–350.
- RIJSBERGEN, C. J. V. **Information Retrieval**. 2nd. ed. Newton, MA, USA: Butterworth-Heinemann, 1979. ISBN 0408709294.
- RUSSAKOVSKY, O. et al. ImageNet Large Scale Visual Recognition Challenge. **Intl. Journal of Computer Vision**, v. 115, n. 3, p. 211–252, 2015.
- SAGHRI, J. A.; FREEMAN, H. Analysis of the precision of generalized chain codes for the representation of planar curves. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, PAMI-3, n. 5, p. 533–539, 1981. ISSN 0162-8828.
- SAITO, P. T. M. et al. Choosing the most effective pattern classification model under learning-time constraint. In: **PlosOne**. [S.l.: s.n.], 2015. v. 10, p. e0129947.
- SANTANNA, M. G. F.; SAITO, P. T. M.; BUGATTI, P. H. Content-based image retrieval towards the automatic characterization of soybean seed vigor. In: **ACM Symposium on Applied Computing**. [S.l.: s.n.], 2014. p. 964–969.
- SETHI, I. K. et al. Color-wise: A system for image similarity retrieval using color. In: INTERNATIONAL SOCIETY FOR OPTICS AND PHOTONICS. **Photonics West'98 Electronic Imaging**. [S.l.], 1997. p. 140–149.
- SETTLES, B. Active learning literature survey. n. 1648, 2009. Disponível em: <<http://burrsettles.com/pub/settles.activelearning.pdf>>.
- SETTLES, B. Active learning. **Synthesis Lectures on Artificial Intelligence and Machine Learning**, Morgan & Claypool Publishers, v. 6, n. 1, p. 1–114, 2012.
- SEUNG, H. S.; OPPER, M.; SOMPOLINSKY, H. Query by committee. In: ACM. **Proceedings of the fifth annual workshop on Computational learning theory**. [S.l.], 1992. p. 287–294.
- SILVA, A. T. D.; FALCÃO, A. X.; MAGALHÃES, L. P. Active learning paradigms for cbir systems based on optimum-path forest classification. **Pattern Recognition**, Elsevier, v. 44, n. 12, p. 2971–2978, 2011.
- SILVA, A. T. da; MAGALHÃES, L. P.; FALCÃO, A. X. Uma nova abordagem cbir baseada em realimentação de relevância e classificação por opf. 2010. Disponível em: <<http://www.dca.fee.unicamp.br/portugues/pesquisa/seminarios/2010/anais/EADCA/2010.pdf>>.
- SINGH, D. et al. Gene expression correlates of clinical prostate cancer behavior. **Cancer Cell**, v. 1, n. 2, p. 203–209, 2002.
- SKOPAL, T.; BUSTOS, B. On nonmetric similarity search problems in complex domains. **ACM Computing Surveys (CSUR)**, ACM, v. 43, n. 4, p. 34, 2011.

- SMITH, J. R.; CHANG, S.-F. Local color and texture extraction and spatial query. In: **Intl. Conference on Image Processing**. [S.l.: s.n.], 1996. v. 3, p. 1011–1014.
- SNEATH, P. H.; SOKAL, R. R. et al. **Numerical taxonomy. The principles and practice of numerical classification**. [S.l.: s.n.], 1973.
- SOKAL, R. R.; MICHENER, C. D. The effects of different numerical techniques on the phenetic classification of bees of the hoplitis complex (megachilidae). In: WILEY ONLINE LIBRARY. **Proceedings of the Linnean society of London**. [S.l.], 1967. v. 178, n. 1, p. 59–74.
- STEHLING, R. O.; NASCIMENTO, M. A.; FALCAO, A. X. On shapes of colors for content-based image retrieval. In: ACM. **Proceedings of the 2000 ACM workshops on Multimedia**. [S.l.], 2000. p. 171–174.
- STEHLING, R. O.; NASCIMENTO, M. A.; FALCÃO, A. X. A compact and efficient image retrieval approach based on border/interior pixel classification. In: **Intl. Conference on Information and Knowledge Management**. [S.l.: s.n.], 2002. p. 102–109.
- STRICKER, M. A.; ORENGO, M. Similarity of color images. In: INTERNATIONAL SOCIETY FOR OPTICS AND PHOTONICS. **Storage and Retrieval for Image and Video Databases III**. [S.l.], 1995. v. 2420, p. 381–393.
- SWAIN, M. J.; BALLARD, D. H. Color indexing. **International Journal of Computer Vision**, v. 7, n. 1, p. 11–32, 1991. ISSN 1573-1405. Disponível em: <<http://dx.doi.org/10.1007/BF00130487>>.
- SZEGEDY, C. et al. Rethinking the inception architecture for computer vision. In: **IEEE CVPR**. [S.l.: s.n.], 2016. p. 2818–2826.
- TAMURA, H.; MORI, S.; YAMAWAKI, T. Textural features corresponding to visual perception. **IEEE Transactions on Systems, man, and cybernetics**, IEEE, v. 8, n. 6, p. 460–473, 1978.
- TAO, B.; DICKINSON, B. W. Texture recognition and image retrieval using gradient indexing. **Journal of Visual Communication and Image Representation**, v. 11, n. 3, p. 327–342, 2000. ISSN 1047-3203. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S1047320300904486>>.
- TAVARES, A. **Recuperação de imagens por conteúdo baseada em realimentação de relevância e classificador por floresta de caminhos ótimos**. Tese (Doutorado) — Tese de Doutorado. Universidade Estadual de Campinas, 2011.
- TONG, S.; CHANG, E. Support vector machine active learning for image retrieval. In: ACM. **Proceedings of The Ninth ACM International Conference on Multimedia**. [S.l.], 2001. p. 107–118.
- TONG, S.; CHANG, E. Support vector machine active learning for image retrieval. In: **ACM International Conference on Multimedia**. [S.l.: s.n.], 2001. p. 107–118.
- TONG, S.; KOLLER, D. Support vector machine active learning with applications to text classification. **Journal of Machine Learning Research**, v. 2, n. Nov, p. 45–66, 2001.

- TORRES, R. S.; FALCAO, A. X. Content-based image retrieval: Theory and applications. **Revista de Informatica Teorica e Aplicada**, v. 13, p. 161–185, 2006.
- ULUWITIGE, D. C. N. W. et al. Effective user relevance feedback for image retrieval with image signatures. In: **Proceedings of the 21st Australasian Document Computing Symposium**. New York, NY, USA: ACM, 2016. (ADCS '16), p. 49–56. ISBN 978-1-4503-4865-2. Disponível em: <<http://doi.acm.org/10.1145/3015022.3015034>>.
- VEER, L. J. van 't et al. Gene expression profiling predicts clinical outcome of breast cancer. **Nature**, v. 415, p. 530–536, 2001.
- WAN, J. et al. Deep learning for content-based image retrieval: A comprehensive study. In: **Proceedings of the 22Nd ACM International Conference on Multimedia**. New York, NY, USA: ACM, 2014. (MM '14), p. 157–166. ISBN 978-1-4503-3063-3. Disponível em: <<http://doi.acm.org/10.1145/2647868.2654948>>.
- WAN, J. et al. Deep learning for content-based image retrieval: A comprehensive study. In: **ACM Intl. Conference on Multimedia**. [S.l.: s.n.], 2014. p. 157–166.
- WANG, J. Z. et al. Content-based image indexing and searching using daubechies' wavelets. **Intl. Journal on Digital Libraries**, Springer, v. 1, n. 4, p. 311–328, 1998.
- WANG, S. **A robust CBIR approach using local color histograms**. [S.l.]: University of Alberta, 2001.
- WANG, T. Musical classification method research based on svm active learning. **Journal Of Advanced Oxidation Technologies**, SYCAMORE GLOBAL PUBLICATIONS LLC 6815 E MANOR DR, TERRE HAUTE, IN 47802 USA, v. 21, n. 2, 2018.
- WEIJER, J. V. D.; SCHMID, C. Coloring local feature extraction. In: SPRINGER. **European Conference on Computer Vision**. [S.l.], 2006. p. 334–348.
- WITTEN, I. H. et al. Weka: Practical machine learning tools and techniques with java implementations. 1999.
- WON, C. S.; PARK, D. K.; PARK, S.-J. Efficient use of mpeg-7 edge histogram descriptor. **ETRI journal**, Wiley Online Library, v. 24, n. 1, p. 23–30, 2002.
- WU, J.; LIN, Z.-K.; LU, M.-Y. Asymmetric semi-supervised boosting for svm active learning in cbir. In: **Proceedings of the ACM International Conference on Image and Video Retrieval**. ACM, 2010. (CIVR '10), p. 182–188. ISBN 978-1-4503-0117-6. Disponível em: <<http://doi.acm.org/10.1145/1816041.1816070>>.
- YANG, Y. et al. Multi-class active learning by uncertainty sampling with diversity maximization. **International Journal of Computer Vision**, Springer, v. 113, n. 2, p. 113–127, 2015.
- YEOH, E. J. et al. Classification, subtype discovery, and prediction of outcome in pediatric acute lymphoblastic leukemia by gene expression profiling. **Cancer Cell**, v. 1, n. 2, p. 133–143, 2002.
- YOSINSKI, J. et al. How transferable are features in deep neural networks? In: **Advances in Neural Information Processing Systems**. [S.l.: s.n.], 2014. p. 3320–3328.

YUE, J. et al. Content-based image retrieval using color and texture fused features. **Mathematical and Computer Modelling**, Elsevier, v. 54, n. 3, p. 1121–1127, 2011.

ZHA, Z.-J. et al. Locally regressive g-optimal design for image retrieval. In: **Proceedings of the 1st ACM International Conference on Multimedia Retrieval**. New York, NY, USA: ACM, 2011. (ICMR '11), p. 59:1–59:8. ISBN 978-1-4503-0336-1. Disponível em: <<http://doi.acm.org/10.1145/1991996.1992055>>.

ZHANG, D.; LU, G. Review of shape representation and description techniques. **Pattern Recognition**, v. 37, n. 1, p. 1–19, 2004. ISSN 0031-3203. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0031320303002759>>.

ZHANG, D. et al. Content-based image retrieval using gabor texture features. **IEEE Transactions PAMI**, p. 13–15, 2000.

ZHANG, L. et al. Convex experimental design using manifold structure for image retrieval. In: **Proceedings of the 17th ACM International Conference on Multimedia**. New York, NY, USA: ACM, 2009. (MM '09), p. 45–54. ISBN 978-1-60558-608-3. Disponível em: <<http://doi.acm.org/10.1145/1631272.1631282>>.

ZHAO, R.; GROSKY, W. I. Negotiating the semantic gap: from feature maps to semantic landscapes. **Pattern Recognition**, Elsevier, v. 35, n. 3, p. 593–600, 2002.

ZHOU, X. S. et al. Semantics and cbir: A medical imaging perspective. In: **Proceedings of the 2008 International Conference on Content-based Image and Video Retrieval**. ACM, 2008. (CIVR '08), p. 571–580. ISBN 978-1-60558-070-8. Disponível em: <<http://doi.acm.org/10.1145/1386352.1386436>>.

ZHOU, Z.-H.; CHEN, K.-J.; DAI, H.-B. Enhancing relevance feedback in image retrieval using unlabeled data. **ACM Trans. Inf. Syst.**, ACM, New York, NY, USA, v. 24, n. 2, p. 219–244, abr. 2006. ISSN 1046-8188. Disponível em: <<http://doi.acm.org/10.1145/1148020.1148023>>.

ZOPH, B. et al. Learning transferable architectures for scalable image recognition. **arXiv**, abs/1707.07012, 2017.