

# SoyRetrieval - Técnicas de Aprendizado e Recuperação de Imagens para Análise do Vigor de Sementes de Soja

Rafael S. Bressan<sup>1</sup>, Marcelo de Souza Junior<sup>1</sup>, Douglas F. Pereira<sup>1</sup>,  
Pedro H. Bugatti<sup>1</sup>, Priscila T. M. Saito<sup>1,2</sup>

<sup>1</sup>Departamento de Computação Universidade Tecnológica Federal do Paraná  
(UTFPR) – Cornélio Procópio – PR – Brasil

<sup>2</sup>Instituto de Computação Universidade de Campinas  
(UNICAMP) – Campinas, SP – Brasil

{rafaelbressan, marcelojunior.2015, douglaspereira}@alunos.utfpr.edu.br  
{pbugatti, psaito}@utfpr.edu.br

**Abstract.** *The production and selection of high quality seeds are fundamental factors for the success of the soybean crop. To evaluate such quality the seed industry in Brazil has been adopting the so-called tetrazolium test. However, a problem that can be observed in this testing process and analysis of seeds is the time and effort required by the specialist, making the process extremely laborious. Therefore, this work aims at developing and making available to the specialists content-based image retrieval and classification tools. The results of the experiments show that the proposed approach presents significant gains compared to state-of-the-art techniques.*

**Resumo.** *A produção e seleção de sementes de alta qualidade são fatores primordiais para o sucesso da cultura da soja. O teste de tetrazólio tem se destacado e sido adotado pela indústria de sementes no Brasil. No entanto, uma problemática que pode ser observada nos processos de testes e análises das sementes é o tempo e esforço do especialista, tornando o processo extremamente laborioso. Portanto, o objetivo desse trabalho é desenvolver e disponibilizar para os especialistas ferramentas de classificação e de recuperação de imagens baseada em conteúdo. Os resultados dos experimentos mostram que a abordagem proposta apresenta ganhos significativos em comparação a técnicas amplamente utilizadas na literatura.*

## Introdução

A produção e seleção de sementes de alta qualidade são fatores primordiais para o sucesso da cultura da soja. Dessa forma, o processo de análise da qualidade das sementes torna-se essencial. Tal processo deve ser confiável e ágil, de forma a fornecer resultados precisos e de maneira rápida. A demora para obtenção de resultados em alguns testes de avaliação tradicional pode impactar negativamente a tomada de decisões relacionada aos processos de colheita, processamento, armazenagem e comercialização, resultando em prejuízos aos produtores de sementes [França-Neto et al. 2007].

O teste de tetrazólio tem se destacado e tem sido adotado pela indústria de sementes no Brasil, dada à sua rapidez, precisão, bem como pelas inúmeras informações

fornecidas. O teste possibilita avaliar a viabilidade e o vigor dos lotes de sementes, além de fornecer as causas da redução de qualidade, tais como: danos mecânicos, por umidade ou por percevejo. No entanto, uma problemática que pode ser observada nos processos de teste e análise das sementes é o tempo e esforço do especialista, pois a quantidade de lotes para análise pode chegar a milhares de sementes, tornando o processo extremamente laborioso.

Portanto, o objetivo desse trabalho é desenvolver e disponibilizar para os especialistas ferramentas de recuperação de imagens baseada em conteúdo e de classificação, de forma a auxiliá-los no processo de análise do vigor de sementes de soja, e que consequentemente possibilitem aumento na produtividade e na sustentabilidade da cultura da soja.

Para tanto, o presente trabalho propõe uma nova abordagem explorando estratégias de aprendizado ativo de forma a selecionar imagens mais relevantes para ambos os processos de aprendizado do classificador e do sistema de recuperação de imagens. Além disso, uma nova estratégia de aprendizado ativo específica para o problema é apresentada. Para validar a abordagem e estratégia propostas foram realizados experimentos utilizando diferentes subconjuntos de imagens de sementes de soja reais obtidas por meio do teste de tetrazólio, bem como técnicas de aprendizado do estado da arte. Para a descrição das imagens foram considerados diferentes extratores, bem como obtidos os melhores descritores, pares de extratores de características e funções de distância para cada um dos conjuntos avaliados.

## **Fundamentações**

A utilização de sementes de qualidade é um fator primordial para o sucesso da cultura da soja. No entanto, para que o sucesso seja alcançado, o controle de qualidade deve ser rápido e preciso, disponibilizando resultados plausíveis, de forma clara e objetiva [França-Neto et al. 1998]. A demora na obtenção de resultados interfere diretamente no processo de tomada de decisões na indústria de sementes, resultando em sérios prejuízos aos produtores de sementes, dificultando a tomada de decisão sobre a lavoura, a colheita e a armazenagem das sementes.

Devido à sua agilidade, o teste de tetrazólio se destacou na indústria de sementes, por trazer um maior número de informações sobre as mesmas, fornecendo resultados precisos sobre o vigor dos lotes de sementes, detectando as causas responsáveis pela redução da qualidade das sementes. A qualidade das sementes pode variar em função de alguns fatores, como o ataque de percevejos, danos mecânicos e deterioração por umidade. Dessa maneira, as sementes devem ser manuseadas de forma cuidadosa em todas as etapas da colheita, visando sua qualidade [Moreano 2012]. Apesar do teste de tetrazólio ser simples e não utilizar equipamentos e reagentes caros, [França-Neto et al. 1998] destaca que para a execução do mesmo é necessário que o analista de sementes seja bem treinado nas técnicas do teste, possuindo conhecimento das estruturas anatômicas da semente. Dessa forma, o conhecimento para realização do mesmo não é trivial.

Nesse sentido, técnicas de recuperação de imagem baseada em conteúdo (*Content-based Image Retrieval* - CBIR) podem ser interessantes, dado que possibilitam a recuperação de imagens com base na similaridade entre uma determinada imagem de consulta e um conjunto de dados. Para calcular essa similaridade, carac-

terísticas de baixo nível baseadas em cor, textura e/ou forma são extraídas das imagens [Malode and Gumaste 2015]. Quando o processo CBIR retorna as imagens similares, de acordo com uma determinada imagem de consulta, o processo pode ser alimentado com o grau de relevância de cada imagem retornada. Assim, essas informações são agregadas com às características da imagem e a função de distância para realizar uma nova consulta mais próxima da intenção do usuário. Esse processo de realimentação de relevância (*Relevance Feedback* - RF) pode ser realizado até que o usuário esteja satisfeito com as imagens retornadas.

De maneira geral, as técnicas de RF permitem realizar o refinamento de uma dada consulta por similaridade baseado na intenção do usuário. No entanto, apesar da existência de diversas técnicas de RF na literatura, a maioria dessas solicita do especialista a anotação das imagens visando apenas a relevância e irrelevância, relegando assim o conceito de diversidade. Tal fato, não contribui de maneira satisfatória para o aprendizado de um classificador ao utilizar tal conjunto de imagens baseado na experiência do especialista. Isso ocorre pois, com o refinamento da consulta ao longo das iterações, mais imagens relevantes são retornadas, e em muitos casos apenas imagens relevantes começam a ser retornadas. Em um primeiro instante, pode parecer que tal fato é extremamente satisfatório. No entanto, ao retornar apenas imagens relevantes em relação a uma dada consulta isso caracteriza-se como uma ênfase exacerbada a uma classe de imagens específica.

Caso um classificador seja treinado com tais imagens, o mesmo será enviesado a classificar corretamente apenas imagens pertencentes a tal classe, ou até mesmo, gerar um modelo extremamente superajustado que consiga apenas classificar corretamente as imagens pertencentes ao próprio conjunto de treinamento sem possibilidades de generalização. Portanto, para que um processo de aprendizado do classificador seja realizado de maneira condizente por meio de tais informações obtidas pelo processo de anotação do usuário, é necessário construir um conjunto de treinamento composto de imagens que possuam um balanceamento entre similaridade e incerteza com uma dada imagem de consulta. Isso permite que o processo não fique preso em soluções locais e restrito às mesmas.

Nesse contexto, estratégias de aprendizado ativo (*Active Learning* - AL) [Settles 2009] podem ser incorporadas ao processo de CBIR. Tais estratégias consistem em selecionar as amostras mais informativas para o treinamento do classificador, o qual participa do seu processo de aprendizado, auxiliando ativamente na seleção das amostras. Um oráculo (por exemplo, especialista) é considerado para confirmar/corrigir os rótulos das amostras.

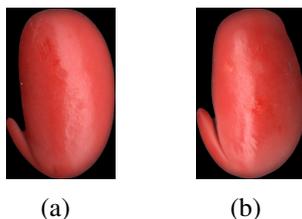
A heurística usada no processo de aprendizado ativo é baseada em como localizar as amostras mais informativas para treinamento do classificador. Para realizar essa tarefa alguns trabalhos propõem diferentes estratégias, dentre elas a seleção das amostras mais próximas à fronteira de decisão de um classificador [Alajlan et al. 2014, Tong and Chang 2001]. O objetivo consiste em selecionar as amostras mais diversas e incertas, próximas à fronteira de decisão do classificador, como sendo as amostras mais informativas, de forma a proporcionar um maior benefício ao modelo de aprendizado.

## Abordagem Proposta

A abordagem proposta agrega o paradigma de aprendizado ativo ao processo de CBIR. Para tanto, uma nova estratégia de aprendizado ativo é proposta e incorporada ao processo CBIR juntamente com RF. Considerando o processo de realimentação de relevância tradicional, na primeira iteração, o usuário realiza o processo de anotação (indicando as imagens relevantes e irrelevantes, de acordo com uma dada imagem de consulta). Em seguida, a partir da segunda iteração, as imagens recuperadas devem ser aquelas que mais contribuirão para o processo de aprendizado do classificador, o qual será treinado em um processo de aprendizado ativo.

Neste cenário, as imagens mais informativas a serem selecionadas são aquelas que apresentam o melhor balanceamento não somente entre a similaridade com a imagem de consulta, mas também certo grau de incerteza. O critério de seleção adotado deve permitir selecionar imagens de classes distintas, bem como que sejam difíceis de diferenciar quando comparada a semântica da imagem da consulta e a semântica da imagem recuperada (por exemplo, imagens na fronteira de duas classes diferentes/sobrepostas).

A Figura 1 apresenta um exemplo de duas amostras informativas de classes distintas, uma amostra sem dano e uma amostra com dano por percevejo, respectivamente. De acordo com o critério adotado de seleção, tais amostras podem ser selecionadas e apresentadas ao usuário, diferentemente do processo CBIR tradicional, que apresentaria apenas imagens mais próximas ao centro de consulta. 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 consequentemente, melhorar a qualidade das imagens retornadas, mais similares em relação à imagem de consulta.



**Figura 1. Exemplos de amostras informativas de duas classes diferentes (a) sem dano e (b) com dano por percevejo nível 3.**

As principais etapas da abordagem proposta são apresentadas na Figura 2. As linhas tracejadas representam o ciclo do processo de aprendizado incremental. O Algoritmo 1 descreve detalhes da abordagem proposta.

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  $\mathcal{I}$ , é realizada a seleção do melhor descritor (melhor par extrator de característica e função de distância). Como este 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  $\mathcal{I}$  (Linha 1), utilizando o melhor extrator de características (BestFD.FeatureExtractor) e gerando o conjunto de aprendizado  $\mathcal{Z}_2$  (Linha 2). Também são extraídas características de  $q$ , utilizando o mesmo extrator.

Na etapa 2, o conjunto de aprendizado  $\mathcal{Z}_2$  é particionado em  $k$  grupos utilizando

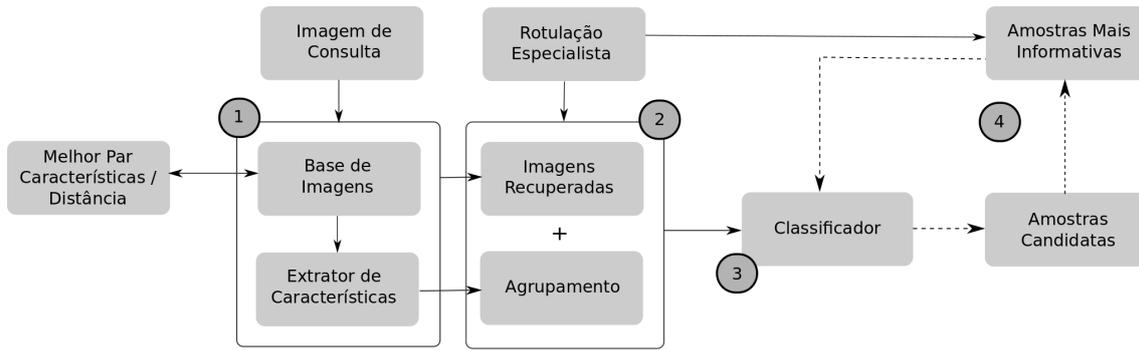


Figura 2. Pipeline da abordagem proposta.

---

**Algoritmo 1:** Abordagem proposta

---

**Entrada** : imagem de consulta  $q$

**Saida** : instância final do modelo de aprendizado  $\mathcal{M}^\Omega$  e uma lista  $\mathcal{L}_R^\Omega$  ordenada por relevância em relação à  $q$ , sendo  $\mathcal{L}_R \subset I$

**Auxiliares:** conjunto de imagens  $I$ , conjuntos de extratores de características  $F_i$  e de funções de distância  $D_j$ , conjunto de aprendizado  $\mathcal{Z}_2$ , número de clusters  $k$ , conjunto de centroides/medóides  $\mathcal{C}$ , lista ordenada  $\mathcal{L}_S$  das imagens mais próximas a  $q$ , número de amostras  $ns$  desejadas, conjunto de treinamento temporário  $\mathcal{Z}'_1$ , conjunto de treinamento  $\mathcal{Z}_1$ , instância do modelo de aprendizado  $M$ , número de amostras a serem selecionadas  $nu$ .

- 1  $BestFD \leftarrow \text{MelhorDescritor}(I, F_i, D_j)$ ;
  - 2  $\mathcal{Z}_2 \leftarrow \text{extraicaoCarateristicas}(BestFD.FeatureExtractor)$ ;
  - 3  $\mathcal{C} \leftarrow \text{localizacaoCentroides}(\mathcal{Z}_2, k)$ ;
  - 4  $\mathcal{L}_S \leftarrow \text{ordenacaoSimilaridade}(q, \mathcal{Z}_2, BestFD.Distance)$ ;
  - 5  $\mathcal{Z}'_1 \leftarrow \mathcal{C} \cup \mathcal{L}_{S_i} \in \mathcal{Z}_2, i = 1, 2, \dots, ns$ ;
  - 6  $\mathcal{Z}_1 \leftarrow \text{anotacao}(\mathcal{Z}'_1)$ ;
  - 7  $M \leftarrow \text{treinamento}(\mathcal{Z}_1)$ ;
  - 8 **repeat**
  - 9      $\mathcal{Z}'_1 \leftarrow \text{selecao}(q, nu, BestFD.Distance, \mathcal{Z}_2 \setminus \mathcal{Z}_1, M)$ ;
  - 10     $\mathcal{Z}_1 \leftarrow \mathcal{Z}_1 \cup \text{anotacao}(\mathcal{Z}'_1)$ ;
  - 11     $M \leftarrow \text{reTreinamento}(\mathcal{Z}_1, M)$ ;
  - 12 **until** *satisfeito*;
  - 13  $\mathcal{M}^\Omega \leftarrow M$ ;
  - 14  $\mathcal{L}_R^\Omega \leftarrow \text{ordenar}(\mathcal{Z}, \mathcal{M}^\Omega), \mathcal{L}_R^\Omega \subset I$ ;
- 

um algoritmo de agrupamento (Linha 3). De acordo com a estratégia de agrupamento aplicada nesta tarefa, o centro de cada grupo pode ser um medóide, sendo assim não é necessário obter a imagem mais próxima do centroide, mas a própria imagem, gerando o conjunto de centróides/medóides  $\mathcal{C}$ . Além das imagens de  $\mathcal{C}$ , as imagens mais similares em relação à  $q$  são selecionadas. Para tanto,  $ns$  imagens desejadas são obtidas de  $\mathcal{L}_S$ , ordenadas por uma ordem crescente de distância em relação à  $q$  (Linhas 4 e 5). Essas imagens são apresentadas ao usuário para anotação (Linha 6). Sendo assim, o conjunto de treinamento  $\mathcal{Z}_1$  é composto pelas imagens mais próximas de cada centróide  $C$  e as  $\mathcal{L}_{S_i}$ ,

$i = 1, \dots, ns$  imagens mais próximas à  $q$ . As imagens anotadas constituem o conjunto de treinamento inicial  $\mathcal{Z}_1$ , o qual é utilizado no processo de treinamento (Linha 7), gerando a primeira instância do modelo de aprendizado  $\mathcal{M}$  (Etapa 3).

Finalmente, na etapa 4, o atual modelo de aprendizado  $\mathcal{M}$  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  $\mathcal{Z}'_1$  recebe as  $nu$  imagens mais informativas em relação à  $q$ , de acordo com o modelo de aprendizado atual  $\mathcal{M}$ , as amostras restantes em  $\mathcal{Z}_2$  ( $\mathcal{Z}_2 \cap \mathcal{Z}_1 = \emptyset$  ou em uma notação simplificada  $\mathcal{Z}_2 \setminus \mathcal{Z}_1$ ), assim como a imagem de consulta  $q$  e a melhor função de distância *BestFD.Distance* (Linha 9). O conjunto de imagens selecionado  $\mathcal{Z}'_1$  é exibido para o usuário. A partir da primeira iteração, esse conjunto já é rotulado previamente pela instância atual do modelo. Assim, o usuário precisa apenas corrigir os rótulos de imagens classificadas incorretamente. As imagens confirmadas e corrigidas pelo usuário são adicionadas ao conjunto de treinamento  $\mathcal{Z}_1$ . O treinamento é realizado novamente e uma nova instância do modelo  $\mathcal{M}$  é gerada. As etapas 3 e 4 (Linhas 8 – 12) são repetidas até que o usuário esteja satisfeito com o processo de aprendizado.

O Algoritmo 2 descreve a estratégia de seleção proposta. 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  $\mathcal{Z}_2 \setminus \mathcal{Z}_1$ , dado que trata-se de um processo iterativo, no qual um conjunto de imagens é selecionado a cada iteração. Em seguida, 4 conjuntos ou listas para controle e armazenamento de amostras são criados (Linha 4). As amostras do conjunto  $\mathcal{Z}_2 \setminus \mathcal{Z}_1$  são classificadas como relevantes ou irrelevantes a partir do modelo de aprendizado  $\mathcal{M}$  (Linha 5). Sendo assim, a lista  $\mathcal{L}_P$  recebe as amostras classificadas como relevantes a partir de  $\mathcal{Z}_2 \setminus \mathcal{Z}_1$  e  $\mathcal{L}_N$  recebe as classificadas como irrelevantes (Linhas 6-7). Em seguida, localiza-se o centro de massa de  $\mathcal{L}_P$ , armazenando-o em  $\mathcal{C}_P$  (Linha 8). O centro de massa pode ser calculado simplesmente pela média das características das amostras existentes na lista. As imagens em  $\mathcal{L}_P$  são ordenadas em relação à distância do centro de massa  $\mathcal{C}_P$ , sendo priorizadas as imagens de maior distância, ou seja, das mais distantes para mais próximas de  $\mathcal{C}_P$  (Linha 9). As imagens em  $\mathcal{L}_N$  também são ordenadas em relação à distância do centro de massa  $\mathcal{C}_P$ . No entanto, nesse caso as amostras são organizadas de forma inversa, sendo das mais próximas para mais distantes de  $\mathcal{C}_P$  (Linha 10). As amostras de  $\mathcal{L}_P$  e  $\mathcal{L}_N$  são concatenadas formando a lista de candidatas  $\mathcal{L}_C$ .

Finalmente, a estratégia de seleção retorna as  $nu$  primeiras amostras mais informativas de  $\mathcal{L}_{C_i}$ ,  $i = 1, \dots, nu$  (Algoritmo 1, Linha 9). Tais amostras são apresentadas para a anotação (como relevante ou irrelevante) pelo usuário. Após a anotação, tais amostras são incorporadas ao conjunto de treinamento  $\mathcal{Z}_1$  (Linha 10). Em seguida, o modelo  $\mathcal{M}$  é treinado novamente e uma nova instância é gerada (Linha 11). O usuário pode continuar o processo até que esteja satisfeito com os resultados retornados. Uma vez satisfeito, pode-se obter um modelo final de aprendizado  $\mathcal{M}^\Omega$  (Linha 13) e uma lista final  $\mathcal{L}_R^\Omega$  (Linha 14) ordenada por relevância (da mais similar para a menos similar) em relação a amostra de consulta  $q$ .

---

**Algoritmo 2:** Estratégia de seleção das imagens mais informativas

---

**Entrada** : imagem de consulta  $q$ , número de amostras a serem selecionadas  $nu$ , melhor função de distância  $BestFD.Distance$ , conjunto de aprendizado  $\mathcal{Z}_2 \setminus \mathcal{Z}_1$  e instância do modelo de aprendizado  $\mathcal{M}$

**Saida** : lista de amostras selecionadas  $\mathcal{L}_{\mathcal{R}}$ .

**Auxiliares:** lista de imagens candidatas  $\mathcal{L}_{\mathcal{C}}$ , lista de imagens relevantes  $\mathcal{L}_{\mathcal{P}}$ , lista de imagens irrelevantes  $\mathcal{L}_{\mathcal{N}}$  e centro de massa das amostras relevantes  $\mathcal{C}_{\mathcal{P}}$ .

```
1 if tamanho( $\mathcal{Z}_2 \setminus \mathcal{Z}_1$ ) <  $nu$  then
2   |  $nu \leftarrow$  tamanho( $\mathcal{Z}_2 \setminus \mathcal{Z}_1$ );
3 end
4  $\mathcal{L}_{\mathcal{P}}, \mathcal{L}_{\mathcal{N}}, \mathcal{C}_{\mathcal{P}}, \mathcal{L}_{\mathcal{C}} \leftarrow \emptyset$ ;
5 classificacaoImagens( $\mathcal{Z}_2 \setminus \mathcal{Z}_1, \mathcal{M}$ );
6  $\mathcal{L}_{\mathcal{P}} \leftarrow$  amostrasRelevantes( $\mathcal{Z}_2 \setminus \mathcal{Z}_1$ );
7  $\mathcal{L}_{\mathcal{N}} \leftarrow$  amostrasIrrelevantes( $\mathcal{Z}_2 \setminus \mathcal{Z}_1$ );
8  $\mathcal{C}_{\mathcal{P}} \leftarrow$  localizacaoCentroMassa( $\mathcal{L}_{\mathcal{P}}$ );
9 ordenacaoDecrescente( $\mathcal{L}_{\mathcal{P}}, BestFD.Distance, \mathcal{C}_{\mathcal{P}}$ );
10 ordenacaoCrescente( $\mathcal{L}_{\mathcal{N}}, BestFD.Distance, \mathcal{C}_{\mathcal{P}}$ );
11  $\mathcal{L}_{\mathcal{C}} \leftarrow$  concatenacao( $\mathcal{L}_{\mathcal{P}}, \mathcal{L}_{\mathcal{N}}$ );
12 while  $i \leq nu$  do
13   |  $\mathcal{L}_{\mathcal{R}} \leftarrow \mathcal{L}_{\mathcal{C}_i} \cup \mathcal{Z}_2 \setminus \mathcal{Z}_1$ ;
14 end
```

---

## Experimentos

A presente seção descreve os experimentos realizados para validação, abrangendo a descrição dos conjuntos de dados utilizados, os cenários de análise, bem como os resultados obtidos por meio de comparações com técnicas do estado da arte na área.

### Conjuntos de Dados

Para os experimentos foram considerados 4 conjuntos de imagens reais de sementes de soja. A Tabela 1 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. Nota-se os desafios e desbalanceamentos característicos de problemas e dados reais. O conjunto  $I_1$  foi obtido a partir de [Santanna et al. 2014]. Já os conjuntos  $I_2$ ,  $I_3$  e  $I_4$  foram criados por meio de captura in-loco em um laboratório de análise de sementes, pertencente a uma unidade de beneficiamento de sementes.

**Tabela 1. Descrição de cada conjunto de dados público usado nos experimentos**

Nome	Imagens	Classes	Imagens/classes
$I_1$	576	12	15-74
$I_2$	350	7	5-103
$I_3$	691	7	6-199
$I_4$	1333	10	7-529

## Cenários

Com o intuito de analisar e validar a abordagem proposta, a qual foi nomeada como AL-NEW, foram realizados testes comparativos utilizando 3 diferentes 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 and Chakrabarti 1999], e aprendizado ativo baseado em SVMs (SVM-AL) [Tong and Chang 2001]. 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, respectivamente, para: QPM, QEX, SVM-AL e CBIR-T.

A partir dos conjuntos de imagens descritos na Seção 4.1 foram extraídas características de cor e textura. Os métodos utilizados para extração de tais características são detalhados na Tabela 2 e nomeados como  $F_i$ , onde  $0 < i < 5$ . Cada tipo de característica extraída foi também comparado com diferentes funções de distância, obtendo o melhor descritor (par extrator de características/função de distância) em cada conjunto de imagens, para realizar as consultas por similaridade. Para tais comparações, foram utilizadas 7 diferentes funções de distância, tais como:  $L_1$ ,  $L_2$ ,  $L_\infty$ ,  $X_2$ , *Canberra*, *Jeffrey Divergence* e *dLog*. A Tabela 3 apresenta a melhor função de distância para cada extrator de característica  $F_i$  e conjunto de dados. Além disso, é possível observar o melhor descritor para cada conjunto (por meio dos extratores destacados em negrito).

**Tabela 2. Propriedades de cada extrator de características.**

Extratores	Categoria	#Caraterísticas
$F_1$ - BIC [Stehling et al. 2002]	Cor	128
$F_2$ - GCH [Stricker and Orengo 1995]	Cor	255
$F_3$ - LBP [Guo et al. 2010]	Textura	256
$F_4$ - LCH [Smith and Chang 1996]	Cor	135

**Tabela 3. Melhores descritores para cada conjunto de imagens  $I$ .**

	$F_1$	$F_2$	$F_3$	$F_4$
$I_1$	<b>Jeffrey</b>	$X_2$	dLog	$X_2$
$I_2$	dLog	$L_\infty$	<b>dLog</b>	Canberra
$I_3$	dLog	$L_\infty$	dLog	Canberra
$I_4$	<b>Canberra</b>	Jeffrey	$L_2$	Canberra

Para o processo de aprendizado do modelo  $\mathcal{M}$ , 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 instancia do classificador.

Para a comparação da abordagem proposta com os outros métodos do estado da arte foram gerados gráficos de Precisão e Revocação (P&R). A precisão de uma con-

sulta pode ser entendida como a fração de imagens relevantes recuperadas, já a revocação é a proporção de imagens relevantes que foram recuperadas por meio da mesma, considerando todas as imagens relevantes no conjunto de dados. Dessa 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 and 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,  $nu = 30$ . Tal valor foi especificado para estar em concordância com a prática de análise diária realizada no processo de análise de sementes. Com o intuito de resumir os resultados obtidos, foram gerados gráficos de P&R por meio do conceito de *Mean Average Precision* (MAP), conforme definido em [Baeza-Yates and Ribeiro-Neto 2011]. Além disso, 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.

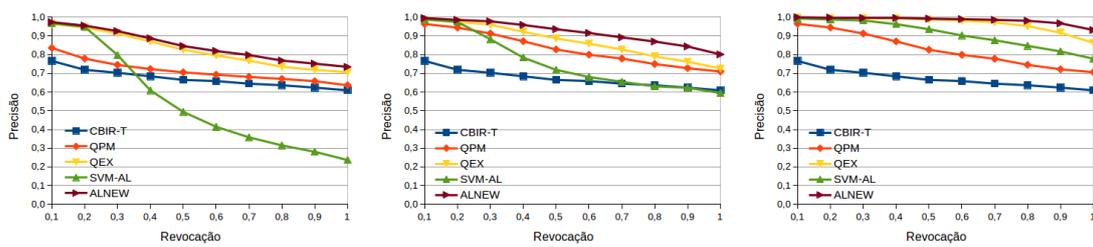
## Resultados

As Figuras 3 a 6 ilustram os resultados obtidos pela abordagem proposta ALNEW em comparação com as abordagens tradicionais CBIR-T, QPM, QEX e SVM-AL, na primeira, quarta e oitava iterações de realimentação, utilizando os conjuntos de imagens  $I_1$  a  $I_4$  (ver Seção 4.1), respectivamente. Em tais gráficos de P&R, a precisão obtida pela abordagem CBIR tradicional foi ilustrada como comparativo base em todos os gráficos, uma vez que não passa pelo refinamento da consulta inicial.

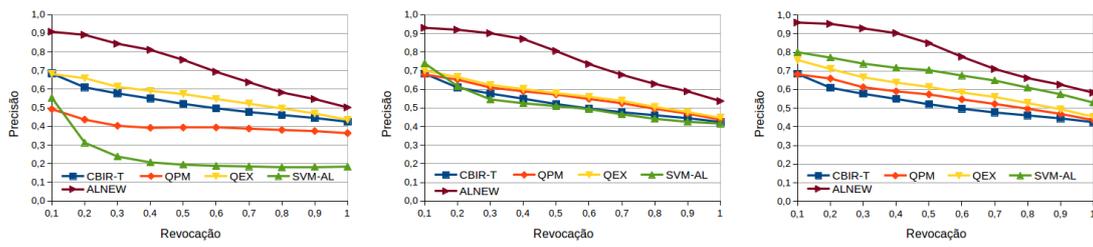
Analisando o conjunto  $I_1$  (Figura 3) pode-se observar que a abordagem ALNEW apresenta as melhores precisões em todas as iterações, quando comparada com as outras abordagens em análise. ALNEW foi até 2,1 vezes melhor que SVM-AL, bem como obteve ganhos de precisão de até 34,66% e 19,68%, considerando, respectivamente, a primeira, quarta e oitava iterações. Em relação à abordagem QEX, ALNEW apresentou ganhos de até 3,99%, 10,34% e 7,87% para as mesmas iterações. Ao comparar ALNEW com a abordagem QPM, novamente foram obtidos ganhos consideráveis de até 15,38%, 12,89% e 31,90% na precisão das consultas.

Comportamento semelhante foi observado ao analisar os resultados obtidos pela abordagem proposta ALNEW, quando aplicada ao conjunto de imagens  $I_2$  (Figura 4). ALNEW apresentou maiores valores de precisão em comparação às demais abordagens. ALNEW obteve uma precisão de 76%, 81% e 85% para a primeira, quarta e oitava iterações, respectivamente. Considerando um nível de revocação de 50%, ALNEW apresenta um ganho mínimo de precisão de 20% e um ganho máximo de até 2,89 vezes em relação às demais técnicas.

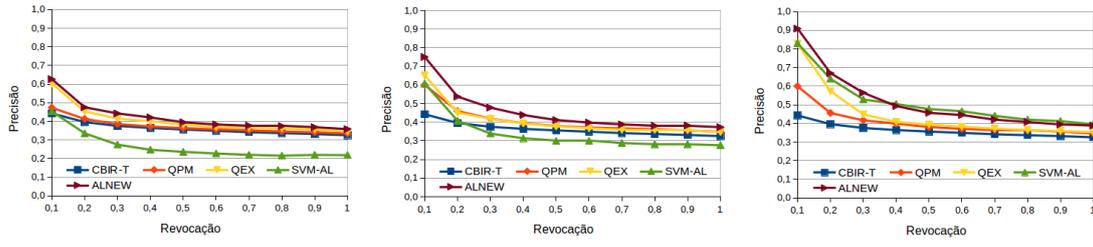
Para o conjunto  $I_3$  (Figura 5), nota-se novamente que ALNEW obteve um ganho de precisão considerável. Por exemplo, para um nível de revocação de 10% nas sucessivas iterações (Figuras 5(a)-(c)), a abordagem proposta atingiu ganhos de no mínimo 3,98% e no máximo 105,22%. Vale enfatizar que níveis de revocação mais baixos são de maior



(a) (b) (c)  
**Figura 3. Curvas P&R obtidas pelas abordagens, utilizando o conjunto de dados I<sub>1</sub>, considerando: (a) primeira, (b) quarta e (c) oitava iteração.**



(a) (b) (c)  
**Figura 4. Curvas P&R obtidas pelas abordagens, utilizando o conjunto de dados I<sub>2</sub>, considerando: (a) primeira, (b) quarta e (c) oitava iteração.**



(a) (b) (c)  
**Figura 5. Curvas P&R obtidas pelas abordagens, utilizando o conjunto de dados I<sub>3</sub>, considerando: (a) primeira, (b) quarta e (c) oitava iteração.**

importância, uma vez que aproximam-se mais de análises cotidianas realizadas.

Por fim, para o conjunto I<sub>4</sub> (Figura 6), a abordagem ALNEW novamente apresentou ganhos consideráveis de precisão em comparação com as outras abordagens. Por exemplo, considerando um nível de revocação de 10%, a mesma foi 5,09 vezes melhor que a abordagem SVM-AL na primeira iteração. Além disso, para o mesmo nível de revocação, considerando a quarta e a oitava iterações, ALNEW apresentou ganhos de precisão de 63% e 22% em relação à SVM-AL. Em relação a QPM obteve um ganho de até 21%, considerando a primeira iteração. Já ao comparar com a QEX, o ganho na precisão foi de até 2% na oitava iteração.

A partir dos resultados obtidos, pode-se concluir que a abordagem proposta apresentou consideráveis ganhos de precisão em relação às consultas por similaridade em conjuntos provenientes de imagens de soja submetidas ao teste de tetrazólio. Além disso, a abordagem ALNEW 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 4 são apresentadas as precisões médias

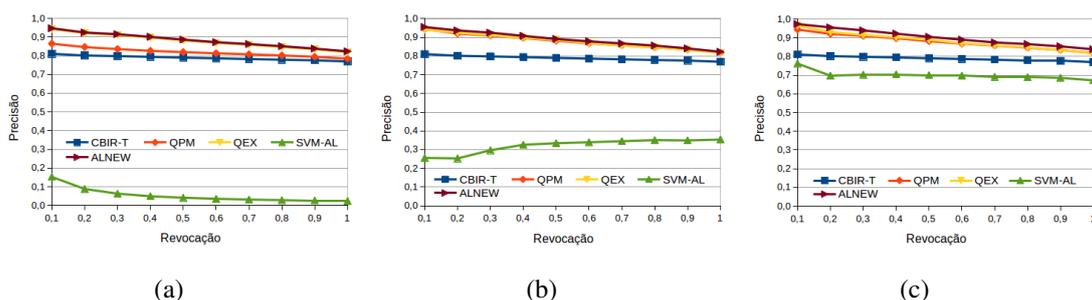


Figura 6. Curvas P&R obtidas pelas abordagens, utilizando o conjunto de imagens  $I_4$ , considerando: (a) primeira, (b) quarta e (c) oitava iteração.

Tabela 4. Precisões médias gerais obtidas por cada abordagem e conjunto de imagens, considerando todas as iterações de realimentação.

Base	QPM	QEX	SVM-AL	ALNEW
$I_1$	80,9	88,8	76,4	<b>92,3</b>
$I_2$	53,1	57,7	51,8	<b>76,1</b>
$I_3$	40,1	41,8	37,4	<b>46,5</b>
$I_4$	86,7	87,6	60,2	<b>89,0</b>

gerais obtidas por cada abordagem e conjunto de dados utilizados no presente trabalho, considerando todas as iterações de realimentação (i.e. da 1<sup>a</sup> até a 8<sup>a</sup> 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 and Ribeiro-Neto 2011]. Ao analisar os resultados sumarizados na Tabela 4, nota-se claramente que a abordagem ALNEW atingiu os melhores níveis de precisão no decorrer das iterações, mantendo maior consistência e robustez.

## Conclusão

No presente trabalho é apresentada uma nova abordagem de aprendizado e recuperação de imagens, incluindo uma nova estratégia de aprendizado ativo capaz de ser integrada de maneira trivial ao processo CBIR, de forma a tratar problemas intrínsecos ao mesmo, bem como prover maior eficácia ao processo.

Por meio dos experimentos, foi possível constatar que a abordagem proposta provê melhorias consideráveis em relação à precisão de consultas por similaridade no domínio de análise de sementes. Dessa forma, os resultados demonstram que a abordagem é altamente viável de aplicação no contexto de análise do vigor de sementes de soja. Tal contribuição pode gerar impacto considerável e melhorias na produtividade de tal cultivar e sustentabilidade como um todo, ao prover um processo mais robusto, com maior rapidez e eficácia.

Em trabalhos futuros pretende-se realizar a avaliação de diferentes classificadores supervisionados e técnicas de agrupamentos gerando diferentes instâncias da abordagem proposta. Além disso, vislumbra-se também a proposta de outras estratégias de aprendizado ativo para a seleção de amostras informativas, conseqüentemente, levando à melhoria da precisão de consultas por similaridade.

## Agradecimentos

O presente trabalho recebeu apoios oriundos do CNPq: grants #472625/2014-4, #431668/2016-7, #422811/2016-5; CAPES; Fundação Araucária de Apoio ao Desenvolvimento Científico e Tecnológico do Paraná; SETI; Belagrícola; e UTFPR.

## Referências

- Alajlan, N., Pasolli, E., Melgani, F., and Franzoso, A. (2014). Large-scale image classification using active learning. *IEEE GRSL*, 11(1):259–263.
- Baeza-Yates, R. and Ribeiro-Neto, B. (2011). *Modern Information Retrieval: The Concepts and Technology Behind Search*. Addison-Wesley Publ. Comp., 2nd ed.
- França-Neto, J. B., Krzyzanowski, F. C., and da Costa, N. P. (1998). *O Teste de Tetrazólio em Sementes de Soja*. Embrapa - Centro Nacional de Pesquisa de Soja, Londrina.
- França-Neto, J. B., Krzyzanowski, F. C., Henning, A. A., de Pádua, G. P., et al. (2007). Tecnologia da produção de semente de soja de alta qualidade. *Embrapa Soja-Artigo em periódico indexado (ALICE)*.
- Guo, Z., Zhang, L., and Zhang, D. (2010). Rotation invariant texture classification using LBP variance (LBPV) with global matching. *PR*, 43(3):706–719.
- Liu, D., Hua, K. A., Vu, K., and Yu, N. (2009). Fast query point movement techniques for large cbir systems. *IEEE TKDE*, 21(5):729–743.
- Malode, P. and Gumaste, S. V. (2015). A review paper on content based image retrieval. *Intl. Research Journal of Engineering and Technology*, 20:883–885.
- Moreano, T. B. (2012). *Evolução das Qualidades Física e Fisiológica da Semente de Soja no Beneficiamento*. Dissertação de mestrado, UEM, Maringá.
- Porkaew, K. and Chakrabarti, K. (1999). Query refinement for multimedia similarity retrieval in mars. In *ACM Intl. Conference on Multimedia*, pages 235–238.
- Santanna, M. G. F., Saito, P. T. M., and Bugatti, P. H. (2014). Content-based image retrieval towards the automatic characterization of soybean seed vigor. In *ACM Symposium on Applied Computing*, pages 964–969.
- Settles, B. (2009). Active learning literature survey. Computer Sciences Technical Report 1648, University of Wisconsin–Madison.
- Smith, J. R. and Chang, S.-F. (1996). Local color and texture extraction and spatial query. In *Intl. Conference on Image Processing*, volume 3, pages 1011–1014.
- Stehling, R. O., Nascimento, M. A., and Falcão, A. X. (2002). A compact and efficient image retrieval approach based on border/interior pixel classification. In *Intl. Conference on Information and Knowledge Management*, pages 102–109.
- Stricker, M. A. and Orengo, M. (1995). Similarity of color images. In *Storage and Retrieval for Image and Video Databases III*, volume 2420, pages 381–393. International Society for Optics and Photonics.
- Tong, S. and Chang, E. (2001). Support vector machine active learning for image retrieval. In *ACM International Conference on Multimedia*, pages 107–118.