



10º Congresso Interinstitucional de Iniciação Científica – CIIC 2016
02 a 04 de agosto de 2016 – Campinas, São Paulo
ISBN 978-85-7029-135-6

Um Estudo de Descritores e Classificadores para Reconhecimento de Bagas de Café

Pedro Henrique Ferreira **Stringhini**¹; Thiago Teixeira **Santos**²

Nº 16605

RESUMO – Métodos de reconhecimento de frutos baseados na utilização de diferentes descritores e classificadores foram estudados. Foi utilizada uma base de dados de 3.393 imagens de café e não-café anteriormente criada e rotulada manualmente. Testes quantitativos demonstraram a identificação de bagas com 93% de precisão e 77% de cobertura utilizando descritores HoG adicionados a mediana dos componentes de cor do formato La^*b^* , aliados ao classificador Gradient Boosting. Esses resultados melhoram o método anteriormente proposto por Santos (2015), e demonstram a possibilidade de evolução de métodos que podem ser aplicados em metodologias de agricultura de precisão, monitoramento e predição de safra.

Palavras-chaves: Aprendizado de Máquina, Reconhecimento de Imagens, Cafeicultura.

ABSTRACT – Methods for fruit recognition based on different descriptors and classifiers were studied. A database of 3,393 images of Coffee and Not-Coffee, previously created and labeled manually was used. Quantitative tests demonstrate the identification of coffee berries with precision of 93% and recall of 77% using HoG descriptors with the median of La^*b^* space color components, classified by Gradient Boosting. These results improve the method proposed by Santos (2015), and demonstrate the possibility of evolution for methods that can be applied in precision farming systems, crop prediction and monitoring systems.

Keywords: Machine Learning, Image Recognition, Coffee Production.

¹ Autor, Bolsista CNPq (PIBIC): Graduação em Engenharia de Computação, UNICAMP, Campinas-SP; phstring95@gmail.com;

² Orientador: Pesquisador da Embrapa Informática, Campinas-SP; thiago.santos@embrapa.br



INTRODUÇÃO

A automatização do reconhecimento de frutos é um aspecto importante para o aumento da captação de informações agrícolas, a exemplo da predição de safra e medição da eficiência de técnicas agrícolas utilizadas sobre um determinado conjunto de amostras. Além disso, técnicas de reconhecimento são essenciais para a automação de outras atividades agrícolas, a exemplo de técnicas de colheita automática (LEE et. al., 2010).

Técnicas propostas com a utilização de imagens e métodos de visão computacional e aprendizado de máquina não são invasivas por padrão. Elas se baseiam no sensoriamento proximal, dada a inviabilidade de um sensoriamento remoto por conta do tamanho dos frutos (SANTOS, 2015). Além disso, elas permitem transmissão e análise de grandes quantidades de dados utilizando padrões semelhantes, permitindo sua utilização em larga escala em contraste com classificações e contagens manuais.

O artigo proposto tem como objetivo expandir os estudos realizados por Santos (2015), explorando o uso de descritores de características em técnicas de aprendizado de máquina para melhorar o desempenho de classificadores de bagas de café.

MATERIAL E MÉTODOS

1 Aquisição da Base de Dados

A base de dados utilizada foi obtida de cafeeiros na Fazenda Tozan (Monte d'Este), em abril de 2015, por uma câmera digital SLR (Canon® EOS Rebel T3i) de lentes 18-55 milímetros configurada em modo automático. As imagens foram adquiridas de uma distância de aproximadamente 1 metro. As imagens são de plantas do cultivar Catuí, com frutos de café em diferentes estágios de desenvolvimento, e são coloridas, de 8 bits e possuem 5184 x 3456 pixels. A base de dados foi apresentada pela primeira vez por Santos (2015).

As imagens originais foram subdivididas em 6276 regiões candidatas de 32 x 32 pixels por meio de um detector que apresentou cobertura³ de 97% nos testes realizados por Santos (2015). Essas regiões foram anotadas manualmente chegando ao número de 510 amostras da classe 1 (Café) e 2.883 amostras da classe 2 (Não-Café).

³ Estamos utilizando *cobertura* como tradução para o termo *recall*.



2 Descritores Utilizados

Para elaboração dos testes foram construídos vetores de característica baseados nos pixels puros da representação em escala de cinza das imagens 32 x 32 pixels na base de dados, e também outros baseados no Histograma de Gradientes Orientados.

O Histograma de Gradientes Orientados, ou HoG (DALAL; TRIGGS, 2005), é um descritor que representa um objeto por meio da distribuição da direção de suas bordas. Ele apresenta robustez quanto a diferenças de luminosidade e geométricas, e foi elaborado convertendo as imagens em 16 células com 64 pixels cada uma, sendo avaliadas 9 direções diferentes e concatenados os histogramas gerados por cada uma delas.

3 Descritores de Cor

Um bom classificador possui como característica essencial a utilização de bons descritores. A qualidade dos descritores se deve a quantidade de informações que estão contidas, e dessa maneira, dado que as outras características utilizadas não levam em conta os aspectos de cor dos objetos, inserção de informações relativas a cor foram testadas.

Para fazer isso, foi realizada a conversão das imagens utilizadas de RGB para o formato La^*b^* , que apresenta como componentes o brilho e duas dimensões de cor baseadas em coordenadas não-linearmente comprimidas. Após a conversão, foram calculadas as medianas para os valores de cor a^* e b^* na imagem, de cada amostra, que foram em seguida anexadas ao seu respectivo vetor de características (imagem em escala de cinza ou HoG).

Testes foram realizados com e sem as características de cor, sendo avaliadas as diferenças no resultado.

4 Classificadores e Avaliação

Para a construção dos classificadores, foram utilizadas Máquinas de Vetores de Suporte (*Support Vector Machines*, ou SVMs), e *Gradient Boosting*.

Os dados foram aleatoriamente separados uma vez em um conjunto de treinamento (70% do total) e de testes (30% do total). Os número de amostras de teste e de treinamento de cada classe estão indicados na tabela 1.

Tabela 1. Separação de dados de teste e treinamento realizada



Conjunto	Amostras de café	Amostras Não-Café	Total
Treinamento	350	4043	4393
Teste	160	1753	1883

Deles foram construídos os descritores utilizados, como indicado na sessão 2.2. Várias combinações de parâmetros dos classificadores foram testadas sobre o conjunto de treinamento, sendo a com melhores resultados selecionada para a avaliação final sobre o conjunto de testes. Em seguida foi realizado o treinamento seguido do teste e decorrente computação dos dados de precisão, cobertura e F1-score adquiridos.

Precisão aqui corresponde a razão entre o número de positivos corretamente indicados em relação a anotação manual sobre o total de positivos (que engloba os falsos positivos para uma classe). Já cobertura indica a razão entre os valores corretamente taxados como positivos e o total de positivos na base de dados. F1-score⁴, por sua vez, é a média harmônica entre as taxas de precisão e cobertura de uma dada classe, e funciona geralmente como um bom indicador para uma avaliação geral de um método de classificação.

5 Implementação

Para o descritor HoG e a descrição das imagens em formato La*b* , foi utilizada as implementações encontradas na biblioteca *scikit-image* (WALT *et. al.*, 2014). As implementações de SVM e *Gradient Boosting* utilizadas estão disponíveis na biblioteca *scikit-learn* (PEDREGOSA *et. al.*, 2011).

RESULTADOS E DISCUSSÃO

Os resultados da classificação de frutos pela utilização dos pixels puros em escala de cinza com Gradient Boosting e *Support Vector Classifier*, com e sem a inserção de descritores de cor, estão indicados nas tabelas de 2 a 5, a seguir.

Tabela 2. Resultados da classificação das amostras de teste para Pixels + SVM

4 F1-score pode ser calculada por
$$F1 = \frac{2 * \text{precisão} * \text{cobertura}}{\text{precisão} + \text{cobertura}}$$



Classe	Precisão	Cobertura	F1-Score	Suporte
Não-café	0,96	0,95	0,96	1723
Café	0,56	0,62	0,59	160
Média/total	0,93	0,93	0,93	1883

Tabela 3. Resultados da classificação das amostras de teste para Pixeis + Cor + SVM

Classe	Precisão	Cobertura	F1-Score	Suporte
Não-café	0,96	0,98	0,97	1723
Café	0,77	0,56	0,64	160
Média/total	0,94	0,95	0,94	1883

Tabela 4. Resultados da classificação das amostras de teste para Pixeis + Gradient Boosting

Classe	Precisão	Cobertura	F1-Score	Suporte
Não-café	0,96	0,99	0,98	1723
Café	0,90	0,56	0,69	160
Média/total	0,96	0,96	0,95	1883

Tabela 5. Resultados da classificação das amostras de teste para Pixeis + Cor + Gradient Boosting

Classe	Precisão	Cobertura	F1-Score	Suporte
Não-café	0,97	0,99	0,98	1723
Café	0,93	0,65	0,76	160
Média/total	0,96	0,97	0,96	1883

Os dados apresentados indicam melhorias na performance advindas da inserção dos componentes de mediana de cores La^*b^* nas *features* da base de dados. Isso é percebido na utilização de *Support Vector Classifier* com um aumento de 21% na precisão e, por consequência, 5% na F1-score da classe café, apesar de haver uma perda de cobertura de 6%. Uma melhora também pode ser notada nos resultados do classificador *Gradient Boosting*, com um aumento de 3% na precisão, 9% em cobertura e 7% na F1-score da classe café.

As melhorias de performance são reforçadas pela avaliação qualitativa da Figura 1, que representa os valores da mediana da componentes de cor a^* e b^* de cada amostra da base de dados utilizada.

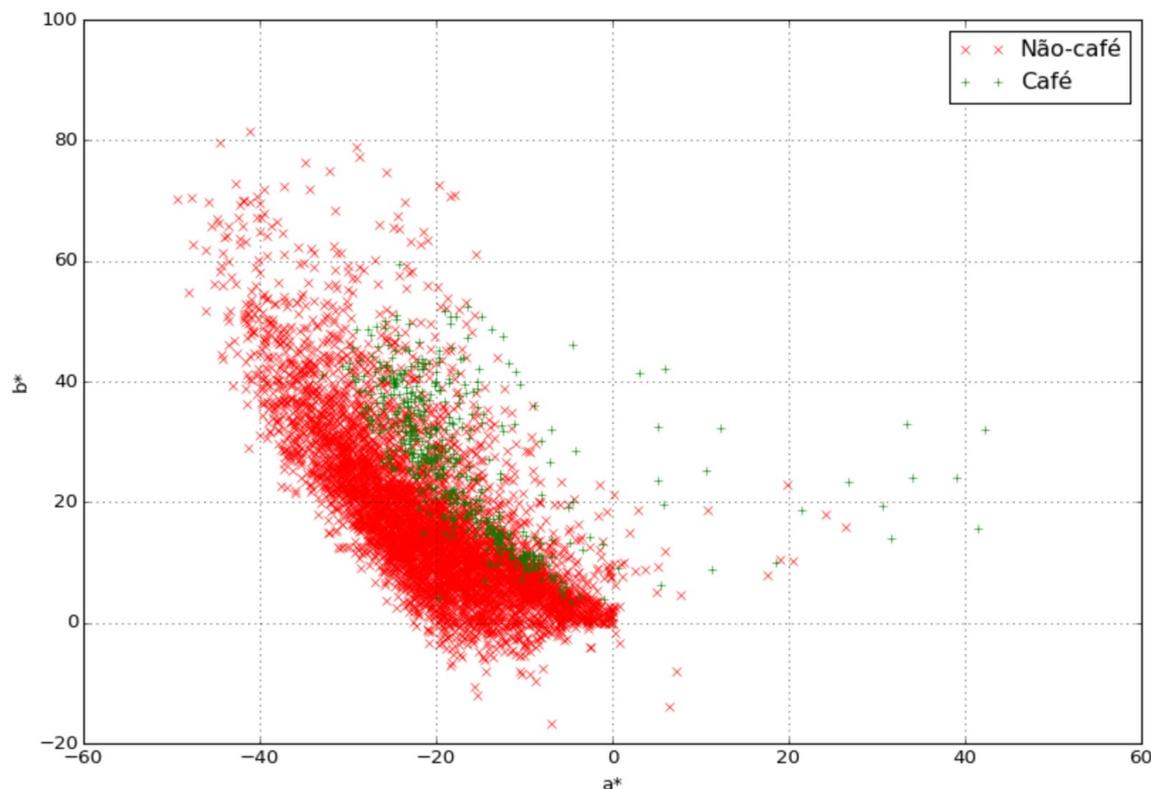


Figura 1: Medianas dos componentes a^* e b^* das amostras da base de dados

É possível observar uma separação nítida de grande parte das amostras da classe 0, Não-Café (em vermelho no gráfico), que se encontram com medianas bem menores que todas as amostras de Café (em verde). Isso é uma boa indicação de que o descritor de cor é útil na classificação das amostras. No entanto, dado que a classe café não é facilmente separável apenas com esses dados, percebe-se que esse descritor não é o suficiente para uma boa classificação, apesar de ser um bom auxílio.

Os resultados da avaliação com os descritores HoG se encontram nas tabelas de 6 a 9 abaixo.

Tabela 6. Resultados da classificação das amostras de teste para HoG + SVM

Classe	Precisão	Cobertura	F1-Score	Suporte
Não-café	0,97	0,98	0,98	1723
Café	0,80	0,70	0,75	160
Média/total	0,96	0,96	0,96	1883



Tabela 7. Resultados da classificação das amostras de teste para HoG + Cor + SVM

Classe	Precisão	Cobertura	F1-Score	Suporte
Não-café	0,97	0,99	0,98	1723
Café	0,88	0,72	0,79	160
Média/total	0,97	0,97	0,97	1883

Tabela 8. Resultados da classificação das amostras de teste para HoG + Gradient Boosting

Classe	Precisão	Cobertura	F1-Score	Suporte
Não-café	0,97	0,99	0,98	1723
Café	0,92	0,71	0,80	160
Média/total	0,97	0,97	0,97	1883

Tabela 9. Resultados da classificação das amostras de teste para HoG + Cor + Gradient Boosting

Classe	Precisão	Cobertura	F1-Score	Suporte
Não-café	0,98	0,99	0,99	1723
Café	0,93	0,77	0,84	160
Média/total	0,97	0,98	0,97	1883

Com esse comparativo, é possível notar o ganho de desempenho obtido com a utilização do Histograma de Gradientes Orientados em relação aos pixels originais, com um aumento sólido de precisão e cobertura para a classe café na maioria das configurações, e sem redução de performance em nenhuma delas.

Além disso, novamente os descritores de cor auxiliaram na melhoria da classificação, com aumento de 8% na precisão e 2% na cobertura do classificador SVC para a classe café, e aumento de 1% na precisão e 6% na cobertura da classe café no classificador *Gradient Boosting*. Esses resultados mostram uma melhoria em relação aos métodos de classificação propostos por Santos (2015), que utilizou os mesmos classificadores com HoG, sem a utilização dos descritores de cor. Especialmente no âmbito de cobertura da classe café, o uso de HoG com classificadores SVC e *Gradient Boosting* apresentou taxas 63 e 65% para a mesma base de dados, enquanto o método aqui utilizado com cores apontou 72 e 77%.

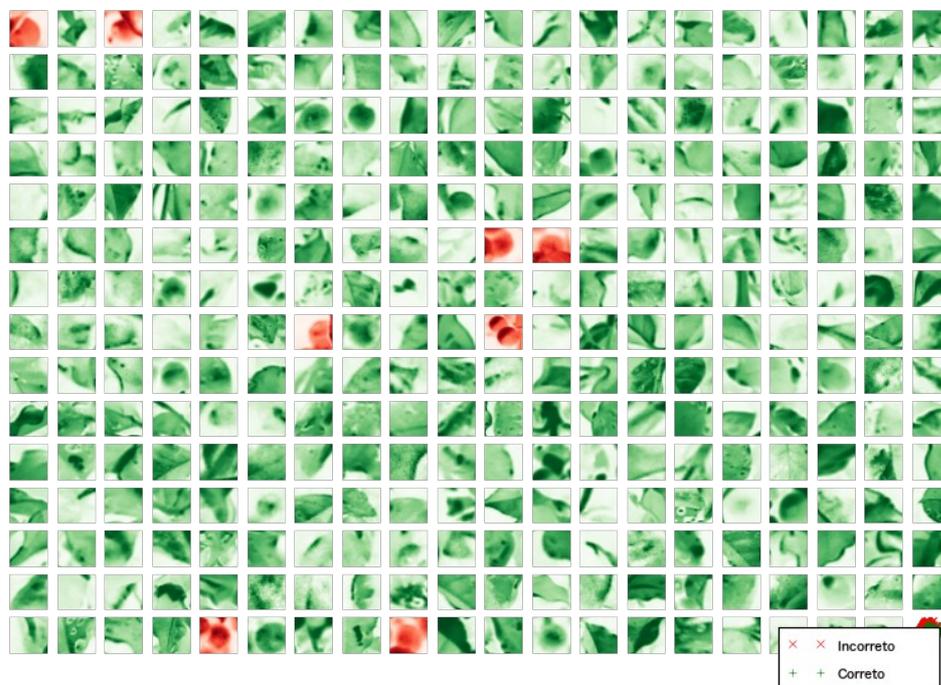


Figura 2: Resultado da classificação de 300 elementos aleatórios da base de testes com descritores HoG+Cor e classificador Gradient Boosting. Em verde a classificações bem-sucedidas, e em vermelho as incorretas.

AGRADECIMENTOS

O autor agradece ao CNPq pela concessão da bolsa para a realização dessa pesquisa de Iniciação Científica. O autor agradece também aos administradores da Fazenda Monte d'Este (Tozan) por autorizarem a aquisição de imagens para a base de dados utilizada por Santos (2015) e nessa pesquisa.

REFERÊNCIAS

DALAL, N.; TRIGGS, B. Histograms of oriented gradients for human detection. *Proceedings - 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR 2005*, I, p. 886–893, 2005. ISSN 1063-6919.

LEE, W. S. et al. Sensing technologies for precision specialty crop production. *Computers and Electronics in Agriculture*, v. 74, n. 1, p. 2–33, 2010. ISSN 01681699.



10º Congresso Interinstitucional de Iniciação Científica – CIIC 2016
02 a 04 de agosto de 2016 – Campinas, São Paulo
ISBN 978-85-7029-135-6

SANTOS, T. T. Detecção automática de bagas de café em imagens de campo. In: X CONGRESSO BRASILEIRO DE AGROINFORMÁTICA, 10., 2015, Ponta Grossa. Anais... Ponta Grossa: UNIVERSIDADE ESTADUAL DE PONTA GROSSA, 2015. p. 1264-1263.

PEDREGOSA, F. et al. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, v. 12, p. 2825–2830, 2011.

WALT, S. van der et al. Scikit-image: image processing in Python. *PeerJ*, v. 2, p. e453, 6 2014. ISSN 2167-8359.