

Segmentação de amêndoas de cacau em imagens digitais de tábuas de corte

Patrick S. Ferraz¹, Prabhát K. de Oliveira¹, Marta Magda Dornelles²

¹Ciência da Computação – Universidade Estadual de Santa Cruz (UESC)
45662-900 – Ilhéus – BA – Brasil

²Departamento de Ciências Exatas e Tecnológicas
Universidade Estadual de Santa Cruz (UESC) – 45662-900 – Ilhéus – BA – Brasil

patrick.ferraz@outlook.com, bartkoliveira@gmail.com, mmbertoldi@uesc.br

Abstract. This paper describes the process of segmentation of the different classes of cocoa beans with the development of an algorithm using the concepts of computer vision with the OpenCV open source library.

Resumo. A prova de corte é uma inspeção visual e olfativa realizada sobre 300 amêndoas cortadas longitudinalmente, escolhidas ao acaso em um lote de amêndoas de cacau, dispostas individualmente em uma tábua. O resultado final desta análise qualifica o lote de cacau, indicando, por exemplo, se o mesmo pode ser usado para a fabricação de cacau gourmet. Neste artigo é apresentada uma metodologia para segmentação de amêndoas de cacau em tábuas de corte utilizando os conceitos de processamento de imagens. As amêndoas segmentadas poderão ser utilizadas em um processo posterior de classificação e qualificação. A metodologia foi implementada e testada mostrando-se promissora.

1. Introdução

O cacau se tornou um fruto popular, sendo um dos alimentos mais apreciados do mundo [Anuário Brasileiro do Cacau 2016]. Uma das causas desta popularidade é a comercialização de chocolate, em que sua qualidade pode ser sentida ao mordê-lo e depende de alguns fatores relacionados às amêndoas de cacau, nomenclatura dada às sementes do cacau, como as características físicas de salubridade, produção de material comestível e características da manteiga de cacau [Ferreira et al. 2013]. Seu notório sabor, apesar dos atributos genéticos, devem-se também às modificações que ocorrem durante sua primeira etapa de produção, o beneficiamento, que se estende desde o seu cultivo, colheita, abertura dos frutos, retirada das sementes, extração da polpa ou do mel, fermentação das sementes, secagem e armazenamento das amêndoas [Efraim et al. 2010].

Existem metodologias que podem ser aplicadas em cada fase do beneficiamento para melhorar a qualidade da amêndoa final. Essa qualidade pode ser verificada pela prova de corte, que é a inspeção visual e olfativa sobre amêndoas de cacau extraídas de um lote de amêndoas. O objetivo é avaliar se o beneficiamento foi realizado corretamente e em que fase do processo ocorreram problemas [Feitosa 2016a]. Além disso, a prova de corte fornece uma classificação para o respectivo lote, indicando a qualidade das amêndoas.

Em Ilhéus-BA A CEPLAC, Comissão Executiva do Plano da Lavoura Cacaueira, órgão do Ministério da Agricultura, Pecuária e Abastecimento, realiza a prova de corte

que é executada de forma manual em que a inspeção visual demora em torno de 90 minutos para ser realizada. O resultado está sujeito a erros humanos e percepções individuais do especialista.

Existe um equipamento que realiza a medição da cor de alimentos, é o espectrofotômetrocolorímetro, que poderia auxiliar na inspeção visual das amêndoas, mas o seu custo é elevado (valor R\$ 166.796,42 em 2017) e a CEPLAC não dispõem de recursos para sua aquisição.

Com a intenção de auxiliar o especialista na análise visual sem a necessidade de equipamento específico para tal, dando maior confiabilidade ao resultado do processo, está em desenvolvimento uma proposta de sistema computacional para a inspeção visual da prova de corte. A primeira etapa deste projeto é a aquisição de uma imagem digital, mediante celular, das amêndoas disposta sobre uma tábua de corte e, a partir dessa imagem, deve-se realizar a extração de cada amêndoas para posterior classificação. O objetivo deste texto é propor uma metodologia para a segmentação das amêndoas.

2. Referencial teórico

2.1. Prova de corte

Normalmente a prova de corte é utilizada para: mensurar o grau de fermentação das amêndoas baseado na coloração; avaliar o aroma das amêndoas antes do corte como fumaça, por exemplo, que pode desqualificar a utilização das amêndoas; verificar infestação de pragas, presença de fungos, quantidade de amêndoas germinadas, achacadas, entre outros. A avaliação resulta na classificação do lote, do qual as amêndoas foram extraídas, em tipo 1, tipo 2, tipo 3 ou fora de tipo [Silva et al. 2013a], de acordo com a instrução normativa nº 38/2008 do Ministério da Agricultura, Pecuária e Abastecimento (MAPA, 38/2008), que define o padrão oficial de classificação de amêndoas de cacau. Tipo 1 é a melhor classificação dada a um lote.

A prova de corte envolve a inspeção visual e olfativa sobre 300 amêndoas de cacau, selecionadas ao acaso em um lote de amêndoas, cortadas longitudinalmente. Após o corte, uma metade é descartada e a outra é inspecionada e disposta em uma tábua contendo 300 orifícios 6.



Figura 1. Exemplo de tábua de corte sem amêndoas.

Na inspeção das amêndoas elas são classificadas nas categorias a seguir:

Amêndoa tipo achatada - caracterizada por seu tamanho pequeno e, geralmente, a amêndoa é seca.

Amêndoa tipo ardósia - possui cor purpúrea (cinza-enverdeado) e geralmente é chapaada, sem compartimentação. Não possui sabor característico de chocolate sendo normalmente muito adstringente.

Amêndoa albina - caracterizada por apresentar uma cor marmorizada, próxima de branca.

Amêndoa tipo germinada - apresentam a casca furada pelo desenvolvimento do embrião.

Amêndoa tipo inseto/brocada - característica de amêndoa estocada em local infestado de insetos. Seu interior apresenta-se carcomido pela ação dos insetos.

Amêndoa tipo marrom - possui cor marrom escura . Normalmente é proveniente de cacau bem fermentado.

Amêndoa mofada - caracterizada por mofo.

Amêndoa tipo parcialmente marrom - possui cor amarronzada. Normalmente é proveniente de cacau bem fermentado, mas a variação na intensidade do cacau pode caracterizar uma variação no grau de fermentação.

Amêndoa tipo violeta - parte interna da amêndoa com cor escura, resistente ao corte. Não possui sabor característico de chocolate sendo normalmente muito adstringente .

Um exemplo para cada categoria pode ser observado na Figura 2. A porcentagem de cada uma dessas categorias de amêndoa na tábua de corte estabelece a qualidade do lote, de acordo com os tipos 1, 2, 3 e sem tipo. Quando a amêndoa apresentar mais de um defeito, será computado apenas o mais prejudicial de acordo com a ordem de gravidade estabelecida na IN-38/2008.



Figura 2. Exemplo para cada categoria de amêndoa.

2.2. Trabalhos correlatos

Muitos trabalhos relacionados ao cacau focam a caracterização das amêndoas analisando seus compostos físico-químicos, como é o caso de [Cruz 2012] e [Silva et al. 2013b].

O trabalho de [Feitosa 2016b] realiza uma investigação visual microscópica sobre as amêndoas cortadas. A ideia é interessante, mas levaria um tempo superior aos 90 minutos da prova de corte, além da necessidade de um equipamento específico.

Não foi encontrada nenhuma pesquisa que se assemelhasse ao projeto em que este trabalho está inserido.

2.3. OpenCV

O OpenCV (Open Source Computer Vision) é uma biblioteca de código aberto e multiplataforma desenvolvida inicialmente pela Intel, utilizada para visão computacional, processamento de imagem, aprendizagem de máquina e operações em tempo real com aceleração de GPU. É liberada sob uma licença de BSD e foi escolhida devido a grande presença em uso acadêmico e comercial. Embora tenha sido inicialmente desenvolvida em C/C++, também oferece suporte para utilização nas linguagens Java, Python e sistemas operacionais Windows, Mac OS, iOS e Android [Passarelli 2017].

3. Metodologia proposta para segmentação das amêndoas

Após estudos e experimentos realizados com alguns algoritmos de segmentação, é proposta uma metodologia para extrair amêndoas de uma tábua de corte. Esta metodologia foi dividida em passos como apresentado na Figura 3.

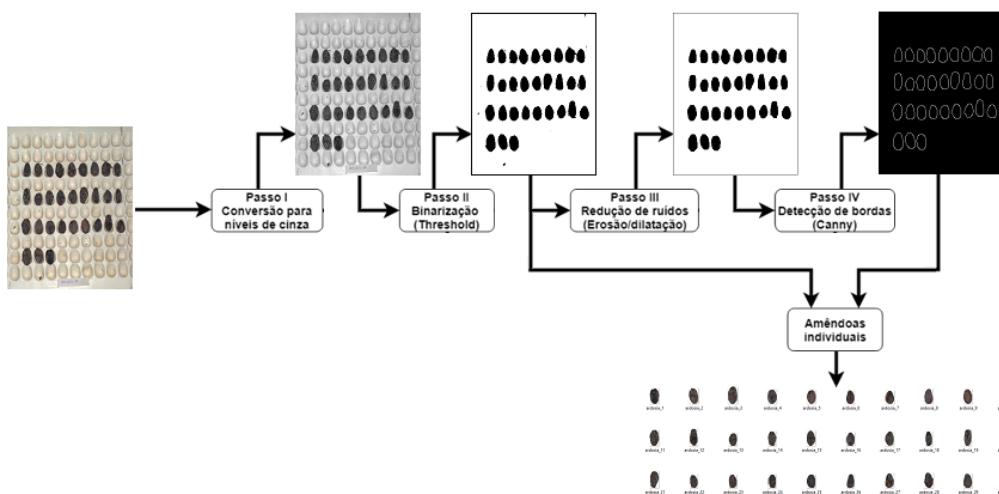


Figura 3. Passos da metodologia proposta.

O primeiro passo pega a imagem de entrada, que é colorida, e a converte para níveis de cinza. Esta conversão foi realizada pela função cv2.cvtColor do pacote OpenCV. No passo II, a imagem em nível de cinza é binarizada, usando um threshold com limiar de 127.

No terceiro passo foi necessário a remoção de ruídos que interferem no próximo passo: detecção de bordas. Esta remoção ocorreu com a utilização dos operadores morfológicos de erosão e dilatação [Dougherty and Lotufo 2003].

No passo IV, o algoritmo Canny [Canny 1986] foi utilizado para a detecção de bordas na imagem resultante do passo três.

De posse da imagem do passo IV e da imagem do passo III, as amêndoas são extraídas uma a uma e arquivos individuais são gerados. Esses arquivos podem ser então utilizados em qualquer processo para análise e classificação das amêndoas.

4. Testes

4.1. Implementação

A metodologia foi implementada em Python 3.5 [Python Software Foundation 2018], utilizando a biblioteca OpenCV 3.4.1 [OpenCV 2018] e o sistema operacional Debian GNU/Linux 9 Stretch 64bits.

A escolha do Python deve-se a rapidez na implementação, mas já está em andamento uma implementação em C++. Posteriormente, uma versão para Android será desenvolvida.

4.2. Imagens de entrada

Uma imagem de entrada é capturada por celular, com resolução de 3072x1728, no formato jpg. Dois exemplos são apresentados na Figura 4.



Figura 4. Exemplos de tábuas de corte capturadas com celular.

A metodologia implementada foi aplicada sobre 8 imagens. Uma imagem da tábua toda preenchida, contemplando todas as classes de amêndoas e 7 imagens (como as da Figura 4.b) das seguintes classes cada uma: achatada, ardósia, parcialmente marrom, marrom, inset, germinada e violeta.

5. Resultados

5.1. Testes de parâmetros e algoritmos

As amêndoas estão divididas em 7 classes, sendo elas: achatada, ardósia, germinada, inseto, marrom, parcialmente marrom e violeta, mostrada na Figura 5.

Foram realizados testes com diferentes algoritmos de segmentação, e para cada um, uma variação nos parâmetros para comprovar quais se mostrariam as melhores escolhas para a segmentação de amêndoas.

Para estes testes foi utilizado a Figura 6. O primeiro passo foi testar o algoritmo *threshold* com diferentes estilos de limiar da biblioteca OpenCV:

- cv.THRESH_BINARY
- cv.THRESH_BINARY_INV
- cv.THRESH_TRUNC

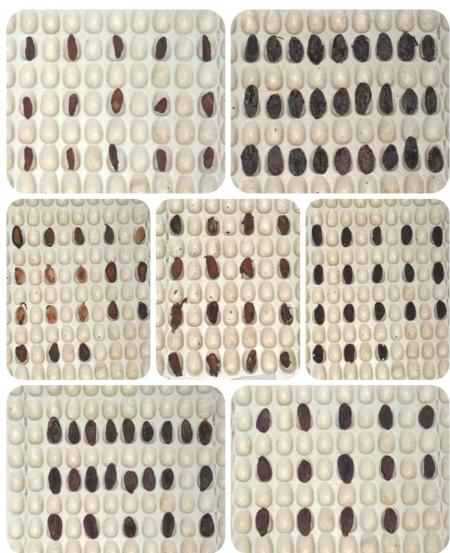


Figura 5. Amêndoas divididas em classes. Achatada, ardósia, germinada, inseto, marrom, parcialmente marrom e violeta, respectivamente.



Figura 6. Exemplares na tábua de corte para testes dos algoritmos.

- cv.THRESH_TOZERO
- cv.THRESH_TOZERO_INV

Os parâmetros de valor limite para classificar os valores de *pixel* do *threshold* foram variados de 0 a 255, em passos de 17, gerando assim 15 imagens. Porém, as imagens com o valor acima de 200 não apresentaram variação visual significativa, então foram mantidas apenas 12 imagens (variação de 0 a 204). Resultados presentes na Figura 7.

Foram testados também algoritmos de detecção de borda, como: Canny, Sobel e Laplacian, mas o último não apresentou resultados significativos para esta segmentação.

5.2. Passo a passo do algoritmo de segmentação

1. O algoritmo inicia o processamento em uma imagem de escolha, para este caso, uma das imagens apresentadas na Figura 5, e faz uma transformação para escala de cinza utilizando a função cv2.cvtColor do pacote OpenCV. Esta

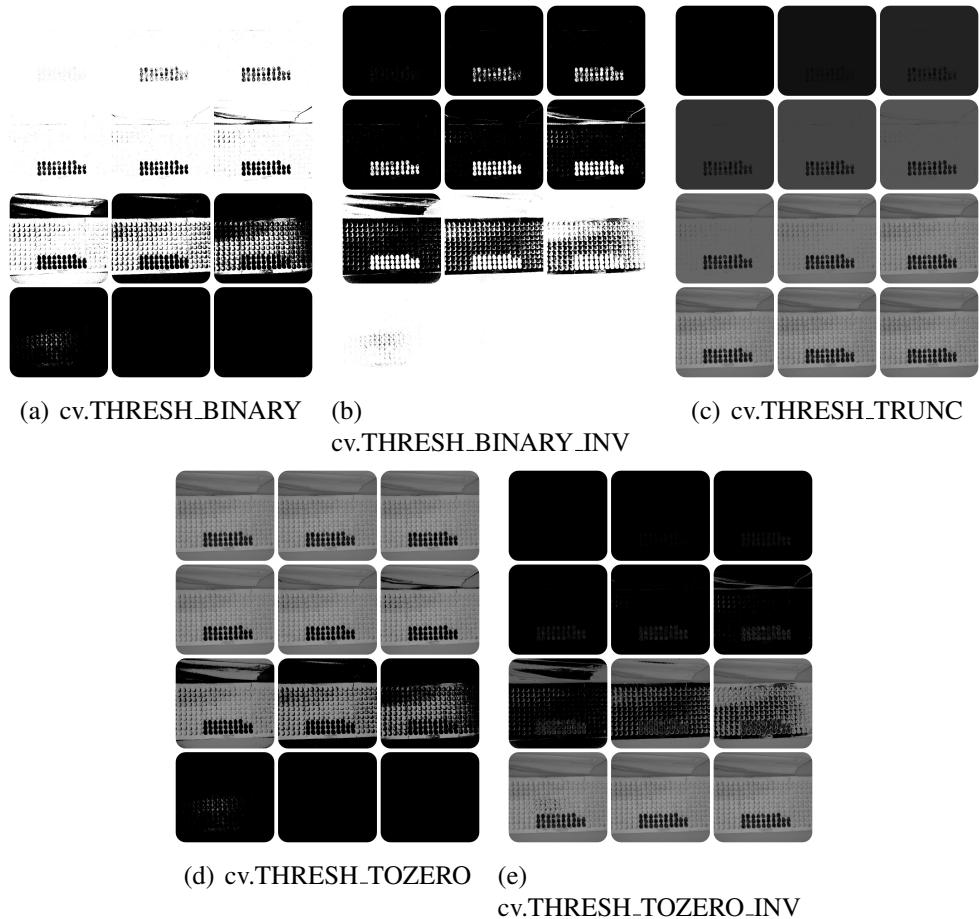


Figura 7. Resultados dos valores de pixels para threshold com diferentes limiar.

função recebe como parâmetro a imagem de entrada e o modo de cor de saída (cv2.COLOR_BGR2GRAY nesse caso). Resultados em Figura 8.

2. É realizado, a partir das imagens obtidas em Figura 8, a técnica do *thresholding*, mas constatou-se que a imagem apresentava “corpos estranhos” que não faziam parte das amêndoas, então viu-se a necessidade de utilizar uma técnica para reduzir ou remover esses “ruídos” da imagem. Primeiramente foi utilizado a função GaussianBlur, que recebe a imagem e o *kernel* como parâmetro. Para testar qual o melhor *kernel* (sendo necessário positivo e ímpar), foram variados de 1 a 5, respectivamente. A Figura 9 ilustra os resultados obtidos com a classe inseto e marrom.

Nota-se que o *kernel* (5,5) foi o que apresentou melhores resultados, para todos os casos, mas ainda assim não eliminou todos os ruídos. Portanto, foi utilizado as funções de erosão e dilatação, respectivamente, com variação na quantidade de iterações. As iterações variam de 1 a 3, onde primeiro fixa-se 1 para erosão, depois fixa-se 2 para erosão, e 3, variando os 3 valores para dilatação em cada caso. A Figura 10 expõe os resultados obtidos no pior e melhor caso das classes inseto e marrom.

Foi escolhido 3 iterações para erosão, e 3 para dilatação, que foram os valores que apresentaram os melhores resultados para todos os casos.

3. O passo seguinte foi utilizar o algoritmo de detecção de borda Canny, utilizando,

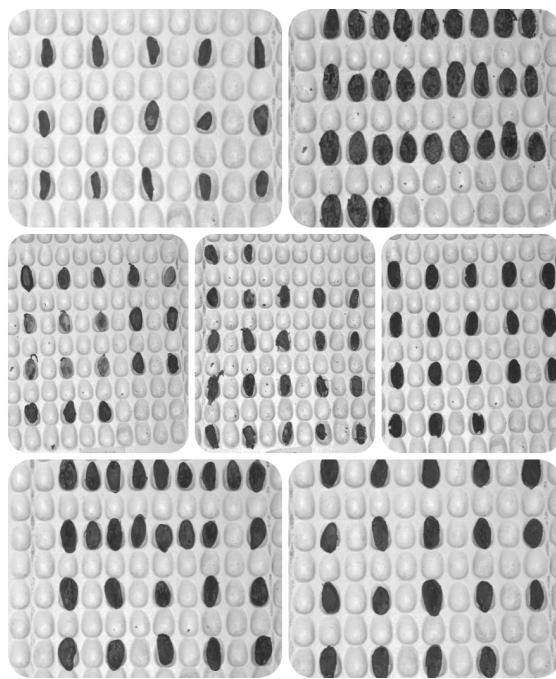


Figura 8. Conversão para escala de cinza.

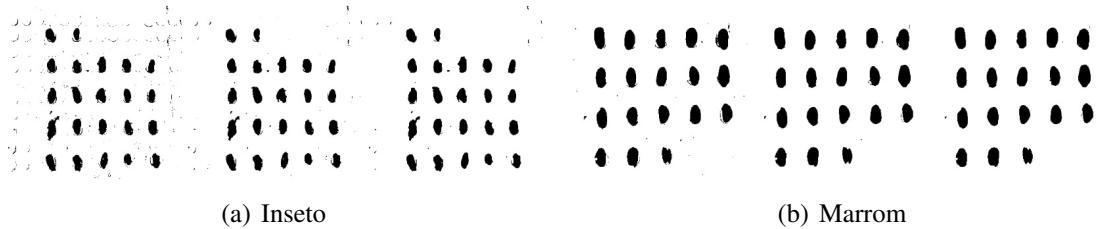


Figura 9. Resultados obtidos após aplicação do Gaussian blur com kernel 1, 3 e 5, respectivamente.

como entrada, as imagens obtidas na Figura 10. Como já foi visto, o algoritmo de Canny possui parâmetros variáveis, e foi realizado testes com estes parâmetros para identificar a melhor opção. A seguir será apresentado apenas os testes para classe achatada, devido a semelhança entre os resultados das outras classes, e o resultado das demais classes, para o melhor caso, que foi escolhido para ser usada no passo seguinte. Os dois parâmetros variam de 0 a 255, sendo assim, o passo de 17, gera 16 imagens para apenas um parâmetro fixo, com um total de 256 imagens. Para exemplificar os testes, serão utilizadas apenas 16 dessas.

4. Posteriormente, os resultados de borda obtidos da Figura 11(b) foram aplicados como máscara na função cv2.findContours, que analisa a imagem e encontra os contornos, para assim destacar as amêndoas encontradas na imagem original, conforme demonstrado na Figura 12.

Embora as amêndoas tenham sido destacadas na imagem original, ainda é necessário a segmentação completa e individual das mesmas, para que possa ser extraído, posteriormente, suas características. Foi utilizado então as imagens resultantes da Figura 10, invertendo-se as cores, onde a silhueta das amêndoas

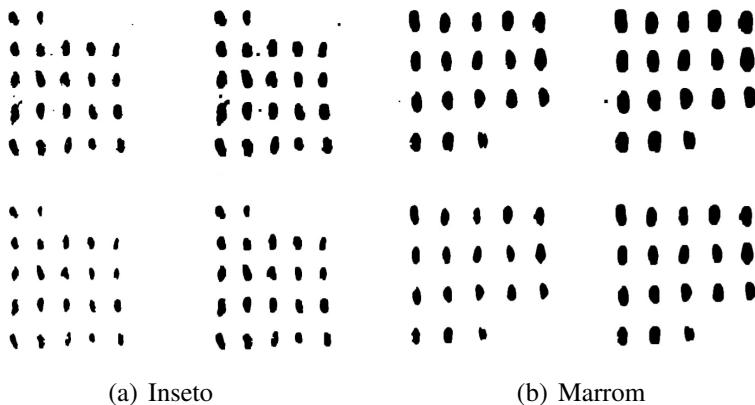


Figura 10. Resultados obtidos após aplicação da erosão e dilatação para cada classe, no pior caso (cima), erosão 1 e dilatação 1, e no melhor caso (baixo), erosão 3 e dilatação 3, respectivamente.

passariam a ter a cor branca, e o plano de fundo a cor preta. Assim, com essa máscara obtida, aplicou-se a função cv2.bitwise_and junto a imagem original, resultando em amêndoas destacadas com as cores originais e todo o fundo convertido a branco (Figura 13).

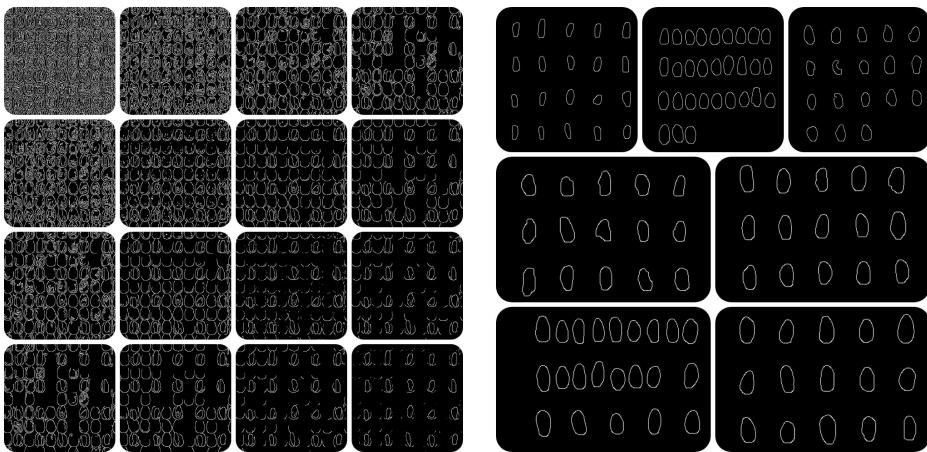
Por fim, foi utilizado os contornos obtidos na Figura 12 para extrair cada amêndoia individualmente, gerando amostras para posterior extração das características e podendo ser utilizadas para classificação das amêndoas.

6. Conclusão

Conforme discutido, o cacau é um fruto de vital importância para economia, principalmente por conta da produção do chocolate. A prova de corte para classificação das amêndoas, apesar de ser um procedimento trivial, evolve inúmeros cuidados para atestar a qualidade da amêndoas. Ressalta-se que, embora os resultados obtidos através da visão computacional com OpenCV, que originou o algoritmo de segmentação aplicado aos processos descritos, tenham trazido os resultados esperados, gerando imagens individuais e de qualidade de cada amêndoas, estamos apenas no início do objetivo que se deseja alcançar. Tais resultados proporcionarão a implementação de métodos de baixo custo, também utilizando a visão computacional, para classificação dessas amêndoas, acarretando na autonomia e maior produtividade para as empresas de classificação de médio e pequeno porte, que muitas vezes, ainda realizam as atividades de forma manual ou não possuem recursos para compra de equipamentos que garantam a sua presença competitiva no mercado.

Referências

- Anuário Brasileiro do Cacau (2016). Anuário brasileiro do cacau 2016. Editora Gazeta.
- Canny, J. (1986). A computational approach to edge detection. *IEEE Trans. Pattern Anal. Mach. Intell.*, 8(6):679–698.
- Cruz, J. F. M. (2012). Caracterização das sementes de variedades à vassoura de bruxa durante a fermentação e após a secagem. Master's thesis, Universidade Federal da Bahia, Salvador, BA.



(a) Testes na variação dos parâmetros do Canny Edge na classe achatada, sendo o primeiro par (0,0) e o último par (255,255). (b) Resultados para o melhor caso na classe achatada, sendo o primeiro par (255,255), nas classes achatada, ardósia, germinada, inseto, marrom, parcialmente marrom e violeta, respectivamente.

Figura 11. Aplicação do Canny Edge na classe achatada (esquerda) e do melhor caso nas demais classes (direita).

Dougherty, E. R. and Lotufo, R. A. (2003). *Hands-on morphological image processing*, volume 59. SPIE press.

Efraim, P., Pezoa-García, N. H., Jardim, D. C. P., Nishikawa, A., Haddad, R., and Eberlin, M. N. (2010). Influência da fermentação e secagem de amêndoas de cacau no teor de compostos fenólicos e na aceitação sensorial. In *Ciência e Tecnologia de Alimentos*. SciELO Analytics.

Feitosa, R. (2016a). Avaliação microscópica de características internas em amêndoas de cacau. Universidade Estadual de Santa Cruz, Ilhéus - BA.

Feitosa, R. (2016b). Avaliação microscópica de características internas em amêndoas de cacau.

Ferreira, A. C. R., Ahnert, D., de Melo Neto, B. A., and Mello, D. L. N. (2013). Guia de beneficiamento de cacau de qualidade. Instituto Cabruca.

OpenCV (2018). Open source computer vision library. <https://opencv.org>. Online; acessado 18 de Junho de 2018.

Passarelli, L. (2017). Aplicação de visão computacional com opencv. <https://www.embarcados.com.br/aplicacao-de-visao-computacional-com-opencv/>. Online; acessado 18 de Junho de 2018.

Python Software Foundation (2018). Python 3.5.5 documentation. <https://docs.python.org/3.5/>. Online; acessado 18 de Junho de 2018.

Silva, A. R. d. A. et al. (2013a). Caracterização de amêndoas e chocolate de diferentes variedades de cacau visando a melhoria da qualidade tecnológica.

Silva, A. R. d. A. et al. (2013b). Caracterização de amêndoas e chocolate de diferentes variedades de cacau visando a melhoria da qualidade tecnológica.



Figura 12. Amêndoas segmentadas e destacadas na imagem original.

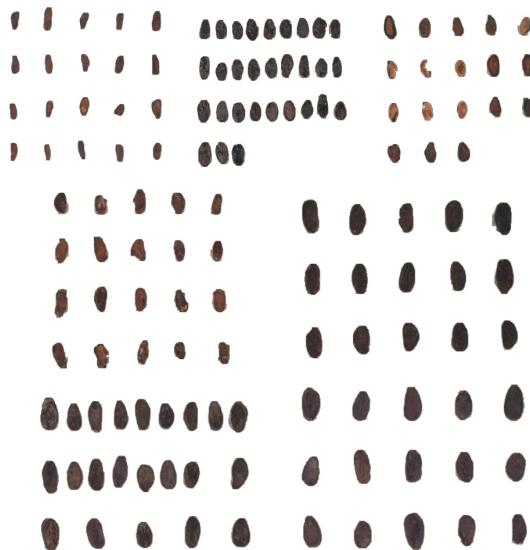


Figura 13. Amêndoas segmentadas e destacadas na imagem original com fundo branco.

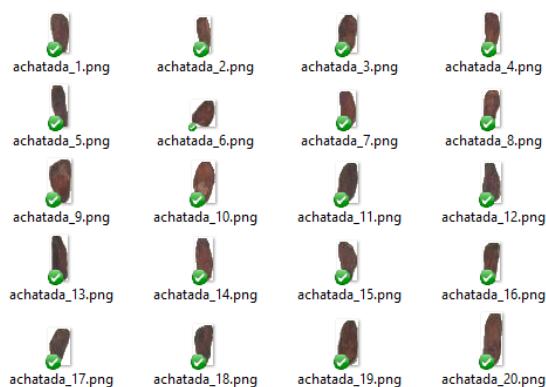


Figura 14. Exemplo de pasta contendo amêndoas extraídas da Figura 13, da classe achatada.