

Identificação da condição física da amêndoa da castanha do Brasil utilizando processamento de imagens

Raphael Gava de Andrade¹; Valentin Obac Roda²; José Dalton Cruz Pessoa³

¹Aluno de mestrado em Engenharia Elétrica, Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos, SP, raphael@cnpdia.embrapa.br;

²Professor Doutor do Departamento de Engenharia Elétrica, Universidade de São Paulo, São Carlos, SP;

³Pesquisador, Embrapa Instrumentação Agropecuária, São Carlos, SP.

As amêndoas das castanhas do Brasil (*Bertholletia excelsa* H.B.K.) têm alto valor nutritivo e também alto valor comercial. Delas são possíveis extrair o óleo ou ainda serem vendidas na forma de sub-produtos (castanhas processadas). Em média, as amêndoas valem 6 vezes mais do que as castanhas em natura e o seu principal exportador é a Bolívia. A Bolívia chegou a essa posição no mercado, devido o uso de tecnologia em todas as etapas do processamento das castanhas. No Brasil sua produção é distribuída em toda a região norte, sendo que este sistema ainda é precário. Este estudo teve como objetivo a identificação da condição da amêndoa (íntegra ou não) através de imagens digitais, utilizando para isso conceitos de visão computacional e redes neurais artificiais (RNA), uma vez que a seleção das amêndoas ainda é feita por inspeção visual. Imagens de amêndoas boas e ruins foram digitalizadas por um scanner HP Scanjet 8250 e passaram por alguns algoritmos bem conhecidos de processamento de imagens: conversão para níveis de cinza, suavização utilizando um *template* de 25x25 e binarização. Com as imagens tratadas dessa forma, foi possível obter dados de alguns descritores (área, perímetro, eixo maior, eixo menor, excentricidade e circularidade) e também suas assinaturas, ou seja, distância entre o centro de massa até a borda. Esses algoritmos foram desenvolvidos em C++ Builder 6 com a biblioteca OpenCV e tiveram seus dados processados por uma *Multi Layer Perceptron* (MLP), arquitetada pelo software Weka 3.6.0. Foram feitos basicamente 3 análises: utilizando os descritores, as assinaturas e as assinaturas divididas pelo descritor de eixo maior. As RNAs foram testadas por validação cruzada e os dados foram testados com apenas 30% do espaço amostral. No teste com os descritores a MLP foi configurada com uma taxa de aprendizagem (TA) de 0.2, fator de *momentum* (FM) de 0.7, 200 épocas, 4, 12 e 2 neurônios nas camadas de entrada, intermediária e de saída assim respectivamente, obteve-se 85% de acerto na validação cruzada e 91,6% de acerto com os dados. Já a rede configurada com TA de 0.5, FM de 0.8, 300 épocas, 18 neurônios, 18 neurônios e 2 neurônios em suas camadas, se conseguiu acerto de 67,5% na validação cruzada e 83,3% com os dados, isso utilizando as assinaturas como entrada. Finalmente para o teste das assinaturas divididas pelo eixo maior a rede configurada com a TA de 0.1, FM de 0.7, 300 épocas, 18 neurônios, 9 neurônios e 2 neurônios em suas camadas, resultou em 75% de acerto na validação cruzada e 83,3% para os dados. Isso mostra que os descritores identificaram mais corretamente as amêndoas boas e ruins. Porém, ao dividir as assinaturas pelo eixo maior, criou-se um tipo de invariância ao tamanho, o que resultou em um melhor resultado no teste de validação cruzada. Assim, tanto os descritores quanto as assinaturas podem ser utilizadas para a avaliação da amêndoa, porém no caso das assinaturas não bastam somente elas.

Apoio financeiro: Embrapa.

Área: Instrumentação.