

Classificação de Imagens de Laranjas Pelo Uso de Histogramas de Gradientes Orientados e Aprendizado de Máquina

João Pedro Paes Landim Alkamim (PIC/DIN-UEM), Maysa Domingues Rodrigues (PIC/DCI-UEM), Franklin César Flores (Orientador/DIN-UEM), Cid Marcos Gonçalves Andrade (Coorientador/DEQ-UEM), e-mail: jpalkamim@gmail.com e maysah.ms2@gmail.com.

Ciências Exatas e da Terra. Ciência da Computação.

Palavras-chave: Classificação de imagens de laranjas, Histogramas dos Gradientes Orientados, Processamento Digital de Imagens, Aprendizado de Máquina

Resumo

O Brasil é o maior produtor mundial de laranjas. A tarefa de se avaliar a qualidade de uma laranja na produção é realizada usualmente por avaliadores cujos parâmetros de qualidade podem variar, o que pode trazer imprecisão ao sistema. Assim, novos sistemas automatizados de classificação para os frutos são desejáveis. O objetivo deste trabalho é o estudo e aplicação de técnicas de aprendizado de máquina e processamento de imagens para a classificação de imagens de laranjas. Em particular, neste trabalho, as imagens são caracterizadas pelos seus histogramas de gradientes orientados. Estes histogramas irão compor o vetor de características que será submetido a dois métodos de classificação: distância mínima ao protótipo e k-NN. Os métodos aqui propostos e experimentados poderão fazer parte de um sistema maior de classificação de laranjas com imagens capturadas em tempo real por uma câmera posicionada em uma linha de produção.

Introdução

O Brasil é o maior país produtor de laranjas do mundo. O processo de classificação das laranjas na produção citrícola de acordo com sua qualidade pode não ser preciso na avaliação dos parâmetros de qualidade do produto, pois depende do julgamento de avaliadores. Diante disso, tecnologias e o conhecimento de computação podem ajudar neste processo. Uma das soluções é o uso de métodos de Processamento de Imagens (Gonzalez, 2010). Processar uma imagem consiste em transformá-la várias vezes com o objetivo de extrair mais facilmente a informação que ela guarda e com ela, reconhecer um padrão para classificação – no caso deste trabalho, a qualidade da laranja. Normalmente o processamento é dividido em etapas, iniciando pelo processo de aquisição das imagens. Em seguida, as imagens são submetidas às etapas de pré-processamento, segmentação e extração de características. Estas características serão utilizadas no processo de classificação das imagens em laranjas. Como representações de características a serem extraídas e representadas, citamos o Local Binary Pattern (LBP) (Armi, 2019) e o Histograma de Gradientes Orientados (HOG) (Dalal, 2005). Tais descritores podem ser submetidos a diversos métodos de classificação tais como k-NN e SVM (Lorena, 2000).

Objetivos

O objetivo geral deste trabalho foi colaborar no desenvolvimento e testes de um algoritmo classificador de laranjas, quanto à sua qualidade, usando técnicas de processamento de imagens e aprendizado de máquina. Foi utilizado neste trabalho o descritor de característica conhecido como HOG (Histogram Oriented Gradient), ou Histograma de Gradiente Orientado (Dalal, 2005). Para o processo de classificação foram utilizados os métodos de distância mínima ao protótipo e o K-NN.

Materiais e Métodos

Para realização do projeto foi usada a linguagem Python 3.x, bem como as bibliotecas Numpy, scikit-image, e OpenCV. Foi montada uma base de dados, estudando e utilizando o descritor *hog* e dois classificadores. Para a validação dos resultados, foi utilizado o método de validação cruzada.

Base de Dados

A base de dados criada para os testes foi feita com mais de 300 laranjas separadas em vários tipos: Boa, Casca Grossa, Podre, Dano por Praga e Verde. Para melhora dos testes e um banco maior, foi utilizado a técnica *Data Augmentation*, que usa-se de técnicas para aumentar a quantidade de dados, adicionando cópias um pouco modificadas. Há no banco de dados um total de vinte e dois mil e setecentas laranjas de cada tipo.

Hog

Histograma de Gradientes Orientados (*hog*) (Dalal, 2005) é um descritor de gradientes, usado no processamento de imagens na parte de detecção de objetos e texturas. O *hog* determina os gradientes horizontalmente (G_x) e verticalmente (G_y) para cada *pixel* da imagem. Com esses valores, é possível calcular a magnitude e o ângulo gerando uma imagem resultante, a qual é capaz de identificar informações importantes de mudança de textura, detecção de objetos e eliminar partes desnecessárias.

A imagem então é dividida em células 8x8 e os gradientes são calculados para cada uma delas. A combinação de todos os *pixels* nas células são somados e distribuídos por um padrão de 9 *bins* formando o histograma. Combinações destes blocos compõem blocos maiores onde vários histogramas são agregados e normalizados. Um vetor unidimensional é formado com todos os valores do histograma e posteriormente é usado para os classificadores.

Classificadores

Neste trabalho, foram utilizados dois classificadores: distância mínima ao protótipo e o *K-Nearest Neighbors* (K-NN). Em ambos os classificadores, foram utilizadas as distâncias L1 e L2.

Distância mínima ao protótipo

O classificador de distância mínima ao protótipo (Gonzalez, 2010), calcula a distância entre cada um dos vetores previamente feitos no banco de dados com algum novo vetor desconhecido. É definido que o protótipo de classe de padrões é o vetor médio dos padrões dessa classe. Para definir a propriedade de um

determinado vetor y de padrão ainda misterioso, é relacioná-lo à classe dos vetores médios do protótipo mais próximo, ou seja, a menor distância entre eles representa o melhor resultado para a classe desconhecida.

K-NN

O classificador *Nearest Neighbors* (vizinhos próximos) (Lorena, 2000) calcula a partir de uma votação de maioria simples dos vizinhos mais próximos de cada ponto. Um ponto de consulta desconhecido é atribuído à classe de dados que possui o maior número de representantes nos vizinhos mais próximos desse ponto. A classificação *K-Nearest Neighbors* é feita por um valor de k arbitrário, escolhido conforme o banco de dados e suas necessidades. O k representa o maior número de representantes que serão selecionados ao redor de um ponto desconhecido para classificá-lo em alguma classe.

Validação Cruzada

Para validar os resultados dos testes, o método de validação cruzada (Lorena, 2000) foi usado em 10 pastas, ou seja, o banco de dados foi dividido nestas pastas contendo o mesmo número de imagens de cada classe de laranja. A técnica de distância ao protótipo foi usada dez vezes de forma que em cada vez fossem utilizadas 9 pastas para treino e uma para teste, sendo a de teste diferente em cada execução, o mesmo método foi utilizado para verificar com o K-NN.

Resultados e Discussão

Ao fazer os teste foram feitos *downsampling* nas fotos das laranjas, tornando-as em imagens de 32x32 *pixel*, 64x64 *pixel* e 128x128 *pixel*. Cada classificador utilizou dessas imagens para comparação entre laranjas do tipo boa, casca grossa, dano por praga, podre e verde.

Foram feitos 2 experimentos. O primeiro experimento comparou variações dos classificadores estudados para a classificação dos cinco tipos citados de laranjas. A comparação se deu pela avaliação das acurácias dos classificadores. No segundo experimento, os mesmos classificadores foram comparados em um esquema “um contra um” para avaliar a discriminação entre os tipos de laranjas comparadas duas a duas. Novamente, a acurácia de cada classificador foi o critério avaliado. No primeiro experimento, para a distância ao protótipo os resultados mostraram que quanto maior a imagem (128x128) melhores são os resultados da acurácia usando a norma L2, sendo eles em média, laranja boa: 70,48%; laranja de casca grossa: 60,79%; laranja podre: 41,41%; laranja com dano por praga: 23,79% e laranja verde: 72,69%. Para o KNN os resultados mostraram que quanto maior a imagem (128x128) melhores são os resultados de acurácia usando a norma L1 e $K = 3$, sendo eles em média, laranja boa: 85,46%; laranja de casca grossa: 55,95%; laranja podre: 39,21%; laranja com dano por praga: 53,30% e laranja verde: 91,63%.

No segundo experimento foram testadas diversas configurações dos classificadores e parâmetros para o cálculo do hog nos testes “um contra um” nas comparações entre os tipos de laranjas. Em seguida, foi calculada a acurácia média para cada tipo de laranja: a laranja boa acurácia média de 82,70%; casca grossa, acurácia média de 78,52%; laranja podre, 56,38%; dano por praga, 42,40% e laranja verde, 10,23% de acurácia média.

Analisando os resultados, as laranjas boas e verdes obtiveram as melhores acurácias no primeiro experimento. No segundo experimento, as laranjas boas e com casca grossa obtiveram boas taxas de acurácia em comparação com as demais. Em particular, a laranja verde obteve uma taxa ruim de acurácia.

Conclusões

Este trabalho apresenta a aplicação do descritor de imagens conhecido com histogramas dos gradientes orientados (HOG) para a classificação de imagens de laranjas. Foram testadas diversas configurações do HOG, bem como dois classificadores: distância mínima ao protótipo e k-NN. Nos experimentos realizados, as laranjas boas tiveram as maiores taxas de acurácias médias, seguidas pelas laranjas de casca grossa. Laranjas verdes obtiveram uma alta taxa média de acurácia no modelo “todos contra todos”. Os resultados encorajam a expansão do estudo com a aplicação de outros descritores e de outros classificadores. Estes estão entre os trabalhos futuros.

Agradecimentos

Os dois primeiros autores agradecem aos orientadores prof. Franklin e prof. Cid pela dedicação e confiança depositadas e ao Programa de Educação Tutorial (MEC/SESu) pelo estímulo à pesquisa, ensino e extensão.

Referências

- Dalal, N. and Triggs, B., “Histograms of Oriented Gradients for Human Detection,” IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2005, San Diego, CA, USA.
- GONZALEZ, Rafael C.; WOODS, Richard E. Digital Image Processing. 3.ed. Londres: Pearson, 2007.
- LORENA, Ana Carolina; GAMA, João; FACELI, Katti. Inteligência Artificial: Uma abordagem de aprendizado de máquina. Grupo Gen-LTC, 2000.
- ARMI, Laleh; FEKRI-ERSHAD, Shervan. Texture image analysis and texture classification methods-A review. arXiv preprint arXiv:1904.06554, 2019.