

CÓDIGO VEGETAL

Um sistema para identificação automática de frutas e legumes

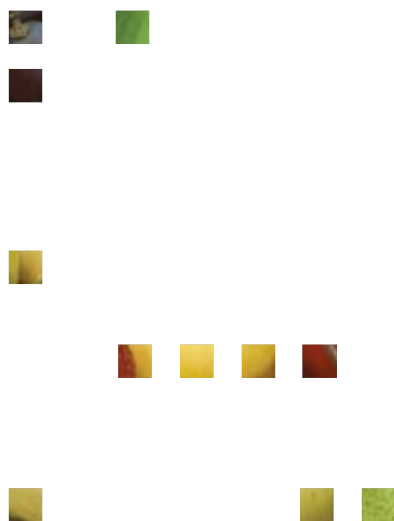
EVANILDO DA SILVEIRA

O tempo gasto pelo caixa de um supermercado de Campinas para localizar numa lista impressa os códigos referentes a frutas e legumes chamou a atenção do professor de ciência da computação Anderson de Rezende Rocha, da Universidade Estadual de Campinas (Unicamp). Enquanto os produtos com códigos de barra em suas embalagens eram registrados rapidamente, a identificação daqueles vegetais atravancava o andamento da fila. Surgiu ali a ideia de desenvolver um sistema capaz de distinguir esse tipo de produto vendido a granel, difícil de ser identificado pelo leitor eletrônico do caixa porque não possui um código.

A solução encontrada por Rocha em conjunto com os pesquisadores Daniel Hauagge, Jacques Wainer e Siome Goldenstein, também do Instituto de Computação da Unicamp, foi desenvolver um sistema, com uma câmera instalada sobre a balança do caixa, pa-

ra analisar imagens do produto a ser classificado. A invenção leva em consideração diversos tipos de informação, como, por exemplo, cor, forma, textura, silhueta, aparência de frutas e legumes, e combina-os de maneira a criar um discriminador poderoso para cada um desses produtos. O *software* desenvolvido por eles é capaz de diferenciar os vários vegetais a partir da combinação de características de cada um.

Rocha explica que a câmera captura apenas a imagem do produto. As informações são extraídas pelo sistema utilizando-se de algoritmos (cálculos matemáticos) de processamento de imagens e reconhecimento de padrões. Apesar de parecer complicado, o funcionamento da invenção desenvolvida durante o doutorado de Rocha – orientado pelo professor Siome Goldenstein e com bolsa da FAPESP – é simples. Ele possui dois estágios: treinamento e teste. Durante o treinamento, várias imagens de produtos vendidos no supermerca-



MONTAGEM SOBRE FOTOS DE EDUARDO CESAR

do são fornecidas ao sistema de modo que ele possa aprender as características descritivas de cada um. Isso é feito identificando-se especificidades de cada tipo de fruta ou legume. “Em seguida, cada tipo de produto é treinado [comparado] contra um outro, em vez de treinar todos contra todos de uma vez.”

O sistema utiliza um método que divide o problema de categorizar muitos produtos diferentes em problemas menores e mais tratáveis. “Isso pode ser mais bem entendido se for considerada uma situação com três classes, como três frutas diferentes, por exemplo, laranja, maçã e abacaxi”, diz Rocha. Nesse exemplo, pode-se definir duas classes de cada vez e dizer que uma delas será considerada como classe virtual positiva e a outra como negativa, ou, simplificando, que uma é maçã e a outra laranja. Isso é feito para as diversas combinações de produ-

tos, tomando-se dois de cada vez. No caso, pode ser “laranja x maçã”, “laranja x abacaxi” e “maçã x abacaxi”. Diversas outras possibilidades podem existir. Por exemplo, o sistema poderia ser treinado para comparar um tipo de produto contra todos os outros. Nesse cenário seria possível ter “laranja x resto”, “maçã x resto” e “abacaxi x resto”. “O importante aqui é tratar o problema dividindo-o em partes menores”, diz Rocha.

Candidatos na balança - Quando o sistema entra em operação, começa a fase de teste. A cada imagem capturada e fornecida para classificação ele extrai o mesmo conjunto de características do vegetal. Elas são comparadas com aquelas armazenadas previamente na etapa de treinamento. Com isso, ele poderá fornecer ao operador do caixa uma lista de candidatos prováveis a ser uma determinada fruta ou legume. Após a confirmação do funcionário, basta verificar o preço do quilograma do produto e multiplicar pelo seu peso. Essa forma de resolver o problema é a grande inovação do sistema, que resultou em uma patente depositada no Instituto Nacional de Propriedade Industrial (INPI). Segundo Rocha, os sistemas existentes são diferentes e menos precisos. Há, por exemplo, um nos Estados Unidos, chamado VeggieVision. “Esse sistema extrai o fundo, identifica o tamanho dos objetos independentemente do número deles na cena e compara com as referências”, explica Rocha. “Ele se baseia em propriedades de cor, textura e densidade, o que requer informações extras da balança.”

Na comparação do índice de acerto na classificação de frutas e legumes, o

VeggieVision perde para o sistema brasileiro. “O índice de acerto do similar americano, mostrando os quatro produtos mais prováveis ao caixa, é de 95%”, diz Rocha. “Enquanto o nosso, mostrando as duas respostas mais prováveis, é de 99%.” Para Rocha, uma comparação mais completa também deveria levar em conta outros fatores. Outra desvantagem do VeggieVision é que ele incorpora no equipamento de aquisição dos dados mecanismos especiais para lidar com variações em iluminação e supressão de reflexões provocadas pela luz na balança e nos sacos plásticos. Em um cenário real, tais mecanismos podem encarecer a adoção do produto pelo supermercado.

O próximo passo é desenvolver um protótipo físico. Por enquanto, o que foi desenvolvido é um *software* e algoritmos para a identificação de frutas e legumes. Para testar a eficiência desse sistema, Rocha e a equipe usaram uma câmera digital para capturar 2.633 imagens de 15 diferentes espécies, entre as quais cebola, laranja, limão, melancia, pera, maçã, caju, quiuí e batata, expostas para venda na Central de Abastecimentos de Campinas (Ceasa). “No momento estamos negociando uma parceria com uma empresa americana para dar continuidade ao projeto”, revela Rocha.

O objetivo agora é melhorar as etapas de separação de variedades dentro de um mesmo tipo de produto, tornando o sistema capaz de diferenciar, por exemplo, dois tipos de banana, como a nanica e a prata. Além disso, os pesquisadores querem incorporar o aprendizado em tempo de operação, ou seja, que a cada resposta confirmada pelo operador do caixa o sistema aprenda com essa confirmação, de modo a ter mais qualidade em classificações futuras. “A última etapa do projeto será integrar o nosso sistema aos existentes nos supermercados baseados em códigos de barra e conectados às impressoras fiscais”, explica Rocha. ■

O PROJETO

Classificadores e aprendizado em processamento de imagem e visão computacional - nº 05/58103-3

MODALIDADE

Bolsa de doutorado

ORIENTADOR

Siome Klein Goldenstein - Unicamp

BOLSISTA

Anderson de Rezende Rocha - Unicamp

INVESTIMENTO

R\$ 95.443,92 (FAPESP)



Artigo científico

ROCHA, ANDERSON; HAUAGGE, DANIEL, C.; WAINER, JACQUES; GOLDENSTEIN, SIOME. Automatic fruit and vegetable classification from images. *Computer and Electronics in Agriculture (Compag)*. v. 70, n. 1, p. 96-104. 2010.