

Estudo de métodos de aprendizagem profunda para reconhecimento de bagas de uva

Andreza Aparecida dos Santos¹

Thiago Teixeira Santos²

Resumo: Dois métodos de aprendizagem profunda (*deep learning*) para reconhecimento de frutos foram estudados. Foi utilizada uma base de dados de 1.830 imagens contando exemplos de bagas de uva e não uva manualmente anotadas. Os testes realizados demonstraram a identificação de bagas com 85% de precisão e de cobertura (*recall*) utilizando redes neurais convolutivas. Esses resultados melhoraram o método anteriormente proposto por Santos e Santos (2017) e demonstram a possibilidade de evolução de métodos que podem ser aplicados em campo.

Palavras-chave: Aprendizagem Profunda, Visão Computacional, Viticultura, Reconhecimento de Padrões.

¹Estudante de Engenharia da Computação da Universidade Estadual de Campinas (Unicamp), bolsista da Embrapa Informática Agropecuária, Campinas, SP.

²Cientista da computação, doutor em Ciências da Computação, pesquisador da Embrapa Informática Agropecuária, Campinas, SP.

Introdução

O reconhecimento automático de frutos se configura em uma etapa importante para a aplicação de métodos de predição de safra, além de ser essencial para automação de outras atividades agrícolas. Técnicas baseadas em imagens e métodos de visão computacional, aprendizado de máquina e aprendizagem profunda não são invasivas e se baseiam em sensoriamento proximal, devido à inviabilidade do sensoriamento remoto por conta do tamanho dos frutos (SANTOS, 2015), e possibilitam a análise de grandes quantidades de dados.

O presente artigo tem como objetivo expandir os estudos realizados por Santos e Santos (2017), explorando o uso de técnicas de aprendizagem profunda para avaliar seu desempenho na classificação de bagas de uva.

Materiais e Métodos

A base de dados utilizada nesse experimento foi obtida na Fazenda Guaspari (Espírito Santo do Pinhal, SP, em abril de 2017). As imagens foram adquiridas de uma distância de aproximadamente 1 metro, e apresentavam frutos de uva da variedade Sirah em diferentes estágios de desenvolvimento. São imagens RGB apresentando 8 bits por canal e possuem 5.184 x 3.456 píxeis. A Transformada Circular de Hough (YUEN et. al., 1990) foi utilizada para identificar 1.830 regiões candidatas, posteriormente escaladas como amostras de 32 x 32 píxeis. Essas regiões foram anotadas manualmente, chegando a um número de 608 amostras da classe 1 ("Uva") e 1.232 amostras da classe 0 ("Não-Uva"). A base de dados é a mesma utilizada por Santos e Santos (2017). O número de amostras estão indicados na Tabela 1.

Tabela 1. Desempenho do classificador J48 para série temporal do EVI.

Conjunto	Amostras de Uva	Amostras de Não-Uva	Total
Treinamento	980	978	1958
Teste	236	254	490

Para a classificação, foram testadas as técnicas de Rede Neural Sequencial e Rede Neural Convolutiva, ambas utilizando o *backend* Theano (BERGSTRA et. al., 2010), e cujas implementações estão disponíveis na biblioteca Keras (CHOLLET, 2017).

Os dados das amostras foram aleatoriamente separados em conjunto de treinamento (80% do total) e conjunto de teste (20% do total).

Em seguida foram construídas as redes. A rede neural sequencial é composta de duas camadas densas e utiliza as funções de ativação Rectified Linear Unit (ReLU) e Softmax. A rede neural convolutiva é composta por seis camadas, sendo: a) uma camada convolutiva de duas dimensões utilizando função de ativação ReLU; b) uma camada de *max pooling* máximo para dados espaciais; c) uma camada *dropout*; uma camada *flatten*; d) duas camadas densas, que utilizam as funções de ativação ReLU e Softmax. As duas redes foram compiladas utilizando a função de *cross entropy* como *loss function*, otimização "adam" e acurácia como métrica.

Para avaliação dos métodos, foram utilizados os indicadores de precisão, cobertura e F1-score, que é a média harmônica entre as taxas de precisão e cobertura de uma classe e funciona como um indicador para uma avaliação geral do método de classificação.

Resultados e Discussão

Os resultados da classificação de frutos utilizando a Rede Neural Sequencial e a Rede Neural Convolutiva estão indicados nas Tabelas 2 e 3 abaixo.

Os dados apresentados indicam um melhor desempenho da Rede Neural Convolutiva, que atingiu taxas 9% maiores na precisão e no F1-score, 7% na cobertura e 10% na acurácia em relação à Rede

Tabela 2. Resultados da classificação das amostras de teste para Rede Neural Sequencial.

Classe	Precisão	Cobertura	F1-Score	Suporte
Uva	0.76	0.78	0,77	236
Não-uva	0.79	0.77	0,78	254
Média/total	0.76	0.78	0,76	490

Tabela 3. Resultados da classificação das amostras de teste para Rede Neural Convolutiva.

Classe	Precisão	Cobertura	F1-Score	Suporte
Uva	0.83	0.86	0,85	236
Não-uva	0.87	0.83	0,85	254
Média/total	0.85	0.85	0,85	490

Neural Sequencial. A Figura 1 exibe alguns dos resultados da classificação no melhor caso. Em relação aos métodos baseados em descritores de características e aprendizado de máquina (SANTOS; SANTOS, 2017), a Rede Neural Convolutiva apresentou melhoras na classificação, com um aumento total de 6% em todos os indicadores para o mesmo conjunto de teste. Avaliando a melhoria por classe, temos, para a classe uva, uma melhora de 11% e de 8% de F1-Score, e para classe não uva, uma melhora de 9% de precisão e 5% de F1-score.

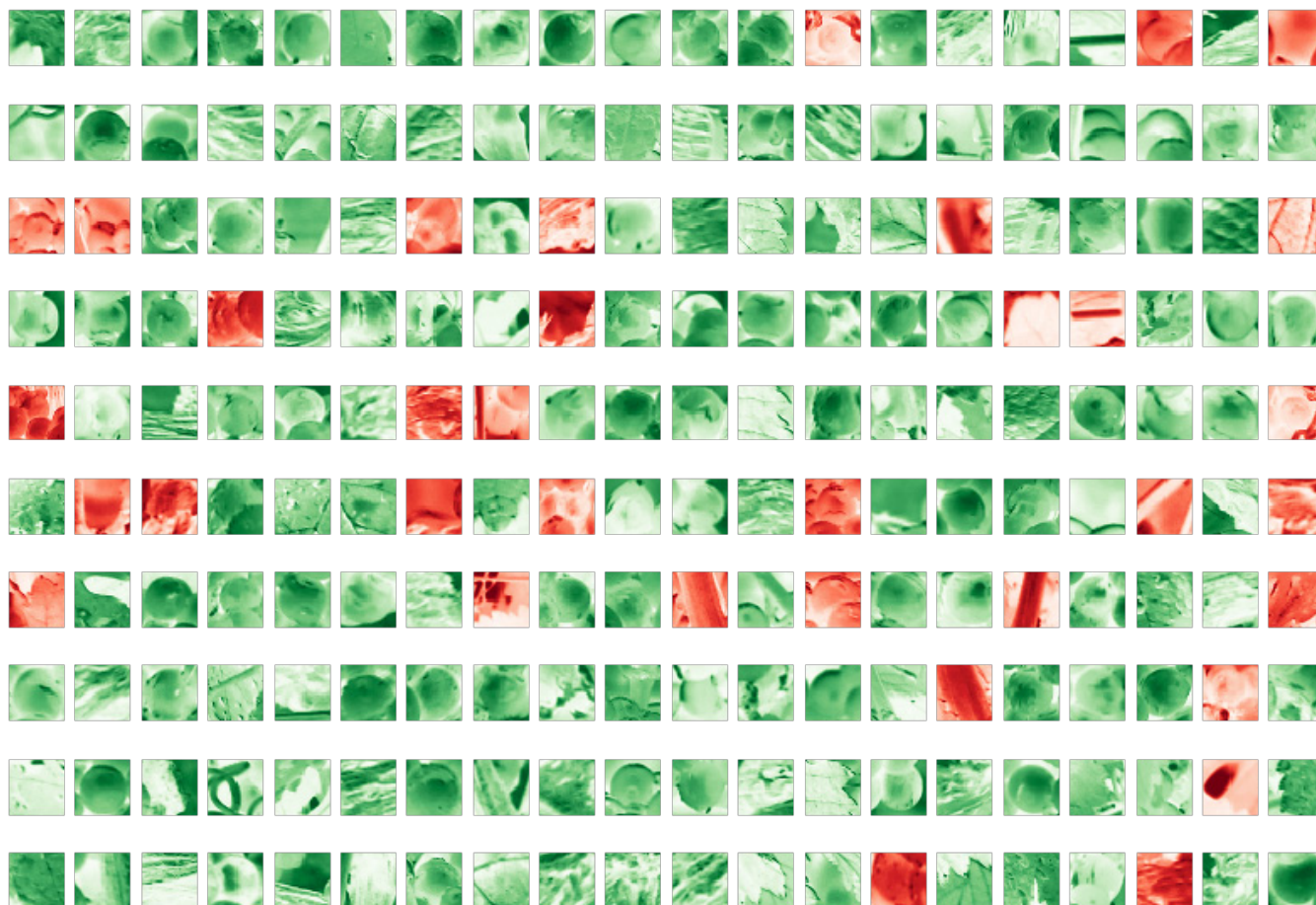


Figura 1. Resultado da classificação de 200 elementos aleatórios do conjunto de teste. Em verde as classificações bem-sucedidas e em vermelho as incorretas.

Agradecimentos

A autora agradece ao CNPq e à Embrapa Informática Agropecuária pela concessão da bolsa para a realização dessa pesquisa de Iniciação Científica. A autora agradece também à equipe da Fazenda Guaspari por autorizarem a aquisição de imagens para a base de dados utilizada nessa pesquisa.

Referências

- BERGSTRA, J.; BREULEUX, O.; BASTIEN, F.; LAMBLIN, P.; PASCANU, R.; DESJARDINS, G.; TURIAN, J.; WARDE-FARLEY, D.; BENGIO, Y. Theano: a CPU and GPU math compiler in Python. In: PYTHON IN SCIENCE CONFERENCE, 9., 2010, Austin. **Proceedings...** [S.l.: s.n.], 2010. p. 3-10.
- CHOLLET, François. Deep Learning with Python. Editora Manning Publications. Publication in October 2017.
- SANTOS, A. A. dos; SANTOS, T. T. Detecção de frutos em campo por aprendizado de máquina. In: CONGRESSO INTERINSTITUCIONAL DE INICIAÇÃO CIENTÍFICA, 11., 2017, Campinas. **Anais...** [S.l.: s.n.], 2017.
- SANTOS, T. T. Detecção automática de bagas de café em imagens de campo. In: CONGRESSO BRASILEIRO DE AGROINFORMÁTICA, 10., 2015, Ponta Grossa. **Uso de VANTs e sensores para avanços no agronegócio: anais.** Ponta Grossa: Universidade Estadual de Ponta Grossa, 2015. Não paginado. SBIAgro 2015.
- YUEN, H.; PRINCEN, J.; JILLINGWORTH, J.; KITTLER, J. Comparative study of Hough Transform methods for circle finding. **Image and Vision Computing**, Amsterdam, v. 8, n. 1, p. 71–77, 1990.