

IDENTIFICAÇÃO DE DANOS EM SEMENTES DE SOJA NO TESTE DE TETRAZÓLIO USANDO MACRO IMAGENS

MECHELN, L.H.M.V.¹; PAULA FILHO, P.L.¹; FRANÇA-NETO, J.B.²; KRZYZANOWSKI, F.C.²; HENNING, A.A.² ¹Universidade Tecnológica Federal do Paraná - UTFPR, Campus Medianeira, Medianeira-PR, luishenrique.si@gmail.com; ²Embrapa Soja.

A produção de soja se consolidou nas últimas décadas como a principal atividade do agronegócio brasileiro, colocando o país em papel de destaque no mercado mundial do grão, sendo assim o segundo maior produtor do mundo e o maior da América Latina. Tal avanço se deu no início pelo aumento da área cultivada e atualmente pela introdução tecnológica nas lavouras (SILVA et al., 2011). As tecnologias em questão são o desenvolvimento de variedades adaptadas para cada região do país e resistentes a pragas e doenças por meio do melhoramento genético e o desenvolvimento de testes capazes de avaliar a qualidade das sementes utilizadas (TECNOLOGIAS..., 2013). Um dos testes mais usados nos laboratórios de sementes é o de tetrazólio, o qual analisa uma amostra de um lote de sementes e determina se a semente é viável para a semeadura e se terá vigor para enfrentar as intempéries da lavoura.

Segundo França-Neto et al. (1998), o teste de tetrazólio é realizado pela análise das características físicas visíveis na semente e a análise consiste na identificação quanto ao dano: o seu tipo, o local da semente onde ocorreu e quais estruturas da semente foram afetadas por ele. A relação entre estas três variáveis permite classificar a semente em uma escala de 1 a 8, sendo 1 a semente sem danos até 8 uma semente completamente danificada (morta). Durante o teste pode ser encontrado um ou mais danos na mesma semente, sendo mais comuns os danos causados por intempéries (dano por umidade), danos mecânicos e danos causados por picadas de percevejos. As limitações do teste de tetrazólio, citadas por França-Neto et al. (1998), incluem a exigência de um treinamento especial sobre a estrutura embrionária da semente, experiência e paciência, pois a análise é relativamente tediosa. Atualmente o mercado exige cada vez mais profissionais capacitados em realizar o teste, por isso, o uso de visão computacional pode tornar o processo de análise mais rápido.

Uma vez que o princípio do teste de tetrazólio é realizar a análise das características físicas visíveis na semente, a hipótese deste trabalho é propor o uso de visão computacional na identificação dos três principais tipos de danos (Umidade, Percevejo e Mecânico) presentes na semente. Para tal objetivo, foi apresentado um processo de captura e catalogação de imagens das amostras das sementes usadas no teste de tetrazólio.

A partir da construção deste acervo, usando técnicas de processamento digital, são extraídas das imagens informações de cor e textura, classificando-as como possuindo dano de umidade, percevejo, mecânico ou não possuindo nenhum tipo de dano.

Para a captura das imagens, cada semente de soja passou pelos processos de pré-condicionamento e coloração padrão do teste e, no momento após o corte e análise visual, as duas metades foram posicionadas sobre uma superfície azul de fundo, que contrasta com a semente, pois é a única cor não encontrada usando o sal de tetrazólio, que resulta na coloração vermelho carmim, que é o padrão usado do teste (FRANÇA-NETO et al., 1998).

Para cada semente foram capturadas duas imagens, uma exibindo a face exterior e outra exibindo a face interior das metades. Apesar de ser fundamental

para o teste de tetrazólio que o analista observe o interior e o exterior de cada metade da semente, este trabalho se limita em propor um método de identificação do tipo de dano na face externa da semente. Para o trabalho apresentado, foi criado um banco de imagens que conta com 90 imagens cuidadosamente escolhidas por representarem os danos encontrados com maior frequência (Umidade, Percevejo e Mecânico). Junto a estas imagens, foi criada uma planilha que indica o número da amostra, nome do arquivo da imagem, quais tipos de danos encontrados na amostra e qual é o dano presente com maior relevância na amostra. A base de imagens e a planilha foram criadas pela Embrapa Soja, utilizando uma câmera Canon, modelo EOS Digital Rebel XT e uma estrutura de PVC com dois anéis de LEDs que produzem uma iluminação circular indireta sem gerar reflexo na superfície úmida da semente e eliminando a sombra gerada pela própria estrutura de iluminação (Figura 1).

Após a criação do banco de imagens, as mesmas passaram por quatro etapas até o final do processamento: remoção de fundo; segmentação das sementes de soja; extração de características e reconhecimento (treinamento e classificação). Após a remoção do fundo azul, a imagem foi convertida em níveis de cinza e processada pelo método de extração de características CLBP (GUO et al., 2010), com três configurações diferentes para comparação dos resultados, sendo elas: R1-P8; R2-P8; e R3-P16, onde R é o raio ou distância e P a vizinhança a partir do pixel analisado. O resultado de método CLBP retorna matrizes dos valores referentes ao sinal (S), magnitude (M) e o valor (C) do pixel central em relação à imagem inteira, que foram combinadas gerando outros produtos derivados destas três matrizes de resultados, sendo eles:

- MC - Multiplicação entre a matriz M e C;
- S_MC - Multiplicação entre a matriz S e MC;
- SM - Multiplicação entre a matriz S e M;
- SMC - Multiplicação entre a matriz S e a Soma das matrizes M e C;

Contudo a matriz CLBP_C é usada apenas para calcular outras matrizes e não serviu de entrada para os classificadores, gerando assim seis arquivos contendo os descritores de todas as imagens para cada uma das três configurações de vizinhança, totalizando dezoito vetores de texturas. Os arquivos de saída nesta etapa foram adequados para o formato ARFF para ser usado na plataforma WEKA (Waikato Environment for Knowledge Analysis) (WEKA, 2014), onde foi configurado um experimento usando os classificadores KNN (*k-Nearest Neighbor*), RNA (Rede Neural Artificial) e SVM (*Support Vector Machines*) para classificar os seis resultados do extrator de características CLBP em três configurações de raio e vizinhança diferentes, gerando 54 classificações diferentes a serem comparadas. O método de extração de características CLBP se mostrou robusto, pois obteve em seu melhor caso 648 descritores, mas poderia gerar 118.098 se não fosse optado pela configuração compacta do método.

Os resultados obtidos pelos classificadores indicam que a melhor configuração do CLBP, para este caso, é o uso de raio 3 e vizinhança 16, obtendo 56,56% de acerto usando o classificador KNN, porém classificando somente a matriz de sinal que possui apenas 18 descritores. A comparação dos resultados pode ser vista na Tabela 1, onde observa-se que o classificador KNN teve a maior diferença entre o pior e o melhor resultado com 23,67 pp enquanto o RNA obteve a diferença de 18 pp. O classificador SVM manteve a mesma taxa de acerto, 39,33%, em todos os arquivos analisados e o vetor que obteve a menor diferença entre os três classificadores foi a configuração que representa o CBPL R2-P8 e o produto MC.

O resultado do estudo mostra que é possível realizar a identificação dos danos na semente de soja por meio do processamento de imagens, contudo a metodologia

utilizada obteve uma taxa de acerto considerada estatisticamente baixa, mostrando que é preciso ser feita uma reavaliação e aprimoramento das técnicas empregadas.

Para se obter um resultado melhor, inicialmente será preciso uma base de imagens maior e com segmentação por dano, uma vez que a semente pode conter mais de um tipo dano, desta forma a imagem é tratada por parte e a identificação será determinada por segmento. No trabalho atual foi utilizado apenas o dano mais latente na amostra inteira, porém os outros danos contribuíram na extração de características.

Podem ser exploradas mais configurações do CLBP, variando o tamanho do raio de análise e o número de vizinhos, com o propósito de encontrar uma configuração mais adequada que descreva melhor a textura da imagem.

Referências

SILVA, A.C.; LIMA, E.P.C.; BATISTA, H.R. A IMPORTÂNCIA DA SOJA PARA O AGRONEGÓCIO BRASILEIRO: UMA ANÁLISE SOB O ENFOQUE DA PRODUÇÃO, EMPREGO E EXPORTAÇÃO. In: V Encontro de Economia Catarinense, 2011, Florianópolis - SC.

FRANÇA-NETO, J.B.; KRZYZANOWSKI, F.C.; DA COSTA, N.P. **O teste de tetrazólio em sementes de soja**. Londrina: EMBRAPA-CNPSO, 1998. 72p. (EMBRAPA-CNPSO, Documentos, 116).

Guo, Z.; Zhang, D.; Zhang, D. . A completed modeling of local binary pattern operator for texture classification, Image Processing, IEEE Transactions on, v.19, n.6, p.1657–1663, 2010.

TECNOLOGIAS de produção de soja - Região Central do Brasil 2014. Londrina: Embrapa Soja, 2013. 265p. (Embrapa Soja. Sistemas de Produção, 16).

WEKA. Waikato Environment for Knowledge Analysis. Disponível em <<http://www.cs.waikato.ac.nz/ml/weka/>>. Acesso em set/2014.

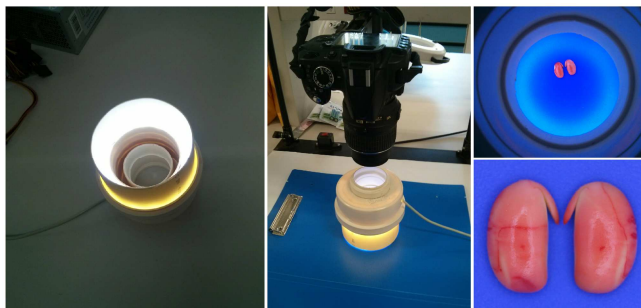


Figura 1. Estrutura do equipamento utilizado para capturar as fotografias

Tabela 1. Resultado do experimento de classificação.

R – P	1 – 8			2 – 8			3 – 16		
Matriz	Classificadores			Classificadores			Classificadores		
	KNN	SVM	RNA	KNN	SVM	RNA	KNN	SVM	RNA
S	45.89	39.33	37.33	53.22	39.33	44.67	56.56^{2 3}	39.33	44.78
M	<u>32.89¹</u>	39.33	37.00	41.56	39.33	37.22	39.44	39.33	42.89
MC	52.00	39.33	40.33	38.33	39.33	39.33	51.11	39.33	38.11
S_MC	46.89	39.33	40.33	40.33	39.33	<u>34.00¹</u>	45.89	39.33	47.00
SM	37.22	39.33	42.56	49.11	39.33	<u>52.00²</u>	47.78	39.33	44.67
SMC	41.44	39.33	41.56	46.89	39.33	41.22	45.89	39.33	47.78

¹Pior resultado do classificador.

²Melhor resultado do classificador.

³Melhor resultado entre os classificadores.