



QUANTIFICAÇÃO DE GERMINAÇÃO DE MUDAS DE CANA-DE-AÇÚCAR POR MEIO DE PROCESSAMENTO DE IMAGENS

Discente: João Vítor Fiolo Pozzuto – RA: 218943

Orientador: Prof. Dr. Lucas Rios do Amaral

Coorientadora: Isabela O. P. da S. Simões (Doutoranda)

Faculdade de Engenharia Agrícola da Unicamp (FEAGRI/Unicamp)

Campinas, outubro de 2020.

1 RESUMO

A avaliação do potencial germinativo das sementes provindas de um cruzamento de cana-de-açúcar é uma etapa crucial para os programas de melhoramento genético vegetal. Essa avaliação é costumeiramente realizada por estimativa visual de avaliadores humanos, e por consequência, sujeita a subjetividade de cada avaliador. Assim, esta pesquisa teve como objetivo o desenvolvimento de uma metodologia para quantificar as plântulas emergidas com base no processamento de imagens coloridas (RGB). Para tanto, analisou-se dois métodos de contagem individual de plântulas, o primeiro baseado na porcentagem de área que as folhas ocupam na foto, denominado grau de cobertura foliar, e o segundo, no número de pontos angulares, a partir da detecção de Quinas de Harris, avaliando-se ainda uma terceira metodologia que utiliza dos dois métodos citados, em conjunto. Verificou-se que o método das Quinas de Harris possui melhor resultado dentre os estudados, apresentando R^2 de 0,863 na predição, contra 0,788 para o método do grau de cobertura. Assim, acredita-se que o processamento digital de imagens oferece suficiente precisão para substituir a avaliação visual humana, uma vez que esta se mostrou bastante subjetiva nas avaliações. O método que utilizou de ambas variáveis não apresentou ganhos de precisão.

2 INTRODUÇÃO

A cana-de-açúcar (*Saccharum spp.*) é uma das principais culturas agrícolas no mundo, cultivada principalmente em países tropicais, como o Brasil, sendo neste, o segundo mais lucrativo produto agrícola, movimentando, segundo IBGE (2017), cerca de R\$ 54 bilhões em 2017, ou, aproximadamente, 2% do P.I.B. nacional, o que mostra a importância do setor sucroenergético. Assim, há grande interesse comercial por cultivares mais produtivos, adaptados às condições edafoclimáticas e mais resistentes a pragas e doenças (CAIEIRO et al., 2010).

Dentro dos programas de melhoramento genético vegetal, uma das etapas é denominada de “teste de germinação”, onde, assim como descrito por Brunkhorst et al. (2000), semeia-se uma pequena amostra de sementes de cada cruzamento em caixas retangulares, que contém um substrato (uma mistura de cascalho, areia e turfa) especialmente desenvolvido para esta aplicação. Estas caixas passam, após aproximadamente 6 dias, por uma avaliação visual humana, que visa caracterizar o potencial germinativo das sementes, com notas de 1 a 5, sendo nota 1 uma caixa com quase total germinação, e 5, uma que apresenta quase nenhuma (ou nenhuma) germinação, sendo essa informação usada para escolha dos cultivares para as próximas etapas do programa.¹

Entretanto, tal avaliação pode apresentar erro de viés humano. De forma alternativa, Brunkhorst et al. (2000) relatam que pode ser realizada contagem por meio de amostras e a metodologia se baseia no uso de uma placa acrílica sobre a caixa, que contém furos, juntos correspondentes a 10% da área, a fim de se contar e, posteriormente, estimar quantas plântulas ali existem. Embora tal método possua boa exatidão nas estimativas de germinação, demanda tempo em demasia, ainda mais em casos reais, onde é necessário avaliar centenas de caixas em um único dia.

Cada vez mais é utilizado o processamento computacional de imagens para os mais diversos fins, desde detecção de rostos por inteligência artificial (YAMAGUCHI et al., 1998), até aplicações agrícolas, como a detecção de falhas de plantio em cafezais (OLIVEIRA et al., 2018). Portanto, é de se

¹ Informações obtidas a partir de comunicação pessoal com o engenheiro agrônomo Danilo Eduardo Cursi, pesquisador do programa de melhoramento genético de cana-de-açúcar da RIDESA.

entender que uma abordagem computacional se enquadre como forma de avaliar o potencial germinativo dos cruzamentos de cana, melhorando a precisão do método visual atual e não demandando tempo em excesso, como a metodologia de Brunkhorst et al. (2000).

Logo, o objetivo desta pesquisa foi avaliar três metodologias de análise digital de imagem para identificar o potencial germinativo dos cruzamentos. A primeira metodologia será baseada na área que as folhas das plântulas ocupam na foto, denominada de “grau de cobertura foliar” (*Coverage Degree - Co*). O segundo método é baseado na quantidade de pontos angulares, a partir do método de detecção de Quinas de Harris (HARRIS, 1998). Por fim, o terceiro método se trata de um modelo matemático que utiliza dos dois primeiros métodos, em conjunto, para contagem de plântulas. Estes três modelos têm como objetivo estimar a quantidade individual de plântulas.

3 MATERIAIS E MÉTODOS

3.1 Obtenção das imagens

Para a aquisição das imagens RGB utilizou-se uma câmera Nikon D7000, com resolução de 16.2Mp e ISO dinâmico com alcance de 100 a 6400 [1/s], acoplada a uma objetiva 18-135mm NIKKORDX VR. Como a lente da câmera foi fixada a aproximadamente 35 cm da superfície do substrato das caixas, o GSD (tamanho médio dos pixels) ficou em torno de 0,13 mm.

Era necessário que as imagens fossem obtidas em uma mesma posição em relação às caixas, perpendicular ao substrato e centralizada, assegurando assim semelhantes regiões fotografadas, além de condições iguais de foco e exposição, que garantem maior uniformidade e padronização das fotos e uma consequente melhor análise. Para isso construiu-se um suporte (Figura 1), que se encaixasse nas caixas de germinação de forma precisa e rápida e ainda garantisse a desejada uniformidade na posição da câmera.

As caixas de germinação encontravam-se em uma casa-de-vegetação de vidro (Figura 1), que garante condições edafoclimáticas ótimas ao desenvolvimento das plântulas de cana-de-açúcar. Foram coletadas imagens em dois dias consecutivos, 09 e 10 de setembro (de 2019), sendo registradas uma imagem para cada uma das 359 caixas avaliadas. Pelas fotos terem sido obtidas em iluminação natural (embora dentro da casa-de-vegetação), é importante ressaltar que ambos os dias apresentavam grande insolação e não possuíam nuvens, além de que os registros foram feitos entre 10h e 17h, sempre com grande luminosidade.



Figura 1 - Câmera acoplada ao suporte, destacando a casa-de-vegetação e a organização das caixas ao fundo.

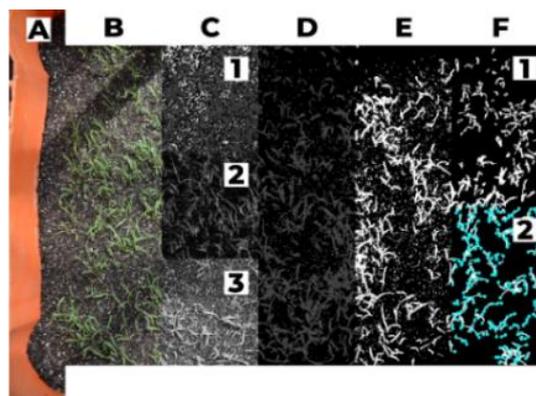


Figura 2 - Resumo visual da metodologia empregada para o processamento das imagens.

3.2 Processamento das imagens

De modo a deixar mais fácil o entendimento das etapas do processamento das imagens, apresenta-se, na Figura 2, a sequência das metodologias do processamento, de forma visual, e que serão explicadas ao decorrer do texto.

Embora as fotos (parte A, Figura 2) tenham sido registradas com a câmera acoplada ao suporte, ainda há certa movimentação que acarretaram em uma discreta diferença na posição das caixas e, por isso, a primeira etapa do tratamento de imagens foi o corte das imagens (parte B, figura 2), em posições fixas, a fim de garantir uma mesma área para análise das caixas, assegurando iguais condições de processamento aos métodos abordados. Em sequência, pelo fato das imagens terem sido obtidas em

iluminação natural, não se garante total homogeneidade da luz, além da ocorrência de sombras. Para isso, torna-se necessário um espaço de cores mais robusto que o RGB, podendo assim eliminar os efeitos da sombra e brilho, que poderiam interferir no processamento. Assim, converteu-se as imagens para o espaço de cores HSV (bandas: *Hue* - Matiz; *Saturation* - Saturação; *Value* - valor (ou brilho)) que, segundo Yang et al. (2015), apresenta maior robustez em relação às mudanças de ambiente, reduzindo a influência da luminosidade. Podemos ver tais bandas, respectivamente, nas seções C1, C2 e C3, da Figura 2. Então, com as três matrizes H, S e V, plotou-se histogramas para encontrar os limiares que diferenciavam as plântulas do substrato e, com esses limiares, comparou-se, pixel a pixel, as imagens em escala de cinza, a fim de obter uma imagem em escala de cinza com as plântulas isoladas do fundo (YANG et al., 2015) (Seção D, Figura 2). Então, converteu-se esta imagem em imagem binária, ou seja, contem apenas valores de 0 e 1 (Seção E, Figura 2), sendo 0 para substrato e 1 para plântulas.

Contudo, mesmo com a comparação do H, S e V, ainda permaneceram na foto pequenos pontos de substrato, em função da sílica refletir de forma semelhante à partes das folhas de cana. Logo, utilizou-se as funções *remove_small_objects* e *remove_small_holes* do pacote *Skimage* do Python, tendo como parâmetros o tamanho dos objetos e furos a se remover, 125 e 30 pixels, respectivamente, obtendo assim uma imagem binária sem outliers (Seção F1, Figura 2), e que será base para os métodos de estimativa de plântulas utilizados.

O método do grau de cobertura foliar se refere a porcentagem de área que as folhas ocupam na foto, ou seja, é a razão entre a área que contém folhas e a área total da imagem, ou, ainda, divisão entre a quantidade de pixels com valor um (referentes às plântulas) e a quantidade total.

O segundo método consiste na contagem de plântulas a partir da detecção de pontos angulares, com base no método de detecção de Quinas de Harris (HARRIS, 1998). A partir de uma matriz de variação, detecta-se pixels onde haja grande variação de valor frente a seus vizinhos, identificando assim quinas e pontos angulares em geral. Para tanto, uma função do pacote *Skimage*, *corner_harris* (baseada nessa matriz), identifica as quinas (seção F2, Figura 2) na imagem binária e então é obtido a quantidade destas. Por fim, o terceiro método é um modelo matemático que utiliza como variáveis o grau de cobertura foliar e a quantidade de Quinas de Harris, descritos acima, e para isso, se utiliza de uma regressão linear múltipla, correlacionando ambos valores.

3.3 Calibração e validação dos métodos

Para geração dos modelos de predição é necessário correlacionar o número de plântulas reais com os resultados obtidos pelos métodos (Quinas de Harris e/ou grau de cobertura foliar). Para isso, utiliza-se de regressões lineares simples, para os dois primeiros métodos e uma regressão linear múltipla para o terceiro, que se utiliza das duas variáveis. Para o número de plântulas real (medido), optou-se por realizar amostragens em 1% da área das caixas, visto que tal contagem é inviável de ser feita para a caixa toda, devido à grande quantidade de plântulas que podem emergir. Portanto, foram coletadas 100 amostras aleatórias para contagem de plântulas. Os valores obtidos foram relacionados aos resultados dos métodos testados.

Para validar a qualidade dos modelos preditivos gerados na calibração foram realizadas mais 50 amostras, seguindo a mesma metodologia daquelas utilizadas para a calibração, ou seja, amostras com 1% da área da caixa e escolhidas aleatoriamente, contando de forma visual quantas plântulas ali haviam. Com esses dados relacionou-se o número de plântulas predito com o real (medido)(Figura 3).

4 RESULTADOS E DISCUSSÕES

O terceiro método, que utiliza de ambas variáveis, não apresentou desempenho superior aos demais, o que pode ser observado pelo R^2 da validação na Figura 3. Isso ocorreu porque o grau de cobertura foliar apresentou p-valor igual a 0,254, ou seja, maior que 0,05 na análise de regressão múltipla, não agregando informações ao modelo. Ademais, essa não melhora do modelo quando se utiliza as duas variáveis, nos indica que as mesmas possuem alta relação entre si, fato comprovado pelo ajuste de curva entre as variáveis que apresentou $R^2 = 0,901$ (Figura 4), e por isso, ao se enquadrarem como variáveis dependentes, não se acrescenta precisão ao método. Portanto, o método que utiliza ambas variáveis se provou inadequado e não será mais discutido neste relatório.

O modelo que utilizou-se do Grau de Cobertura Foliar apresentou resultado inferior ao que se baseia nas Quinas de Harris, fato evidenciado pela menor qualidade preditiva do modelo, ilustrado pelo R^2 (Figura 3).

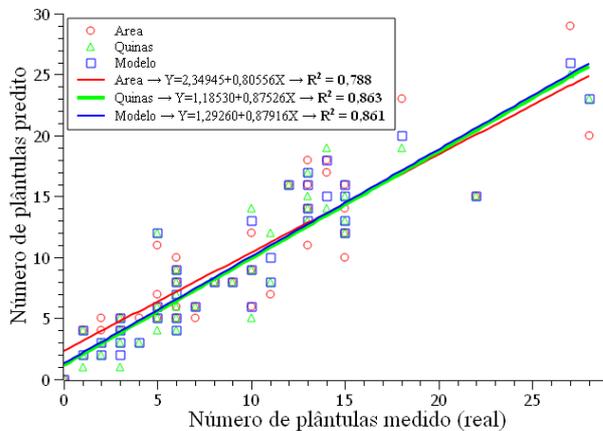


Figura 3 - Gráfico de correlação entre o número de plântulas predito e real para as 50 amostras de validação conforme os 3 métodos estudados.

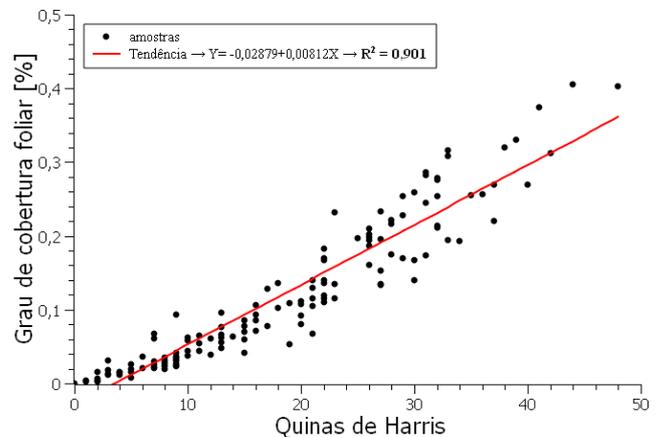


Figura 4 - Gráfico de correlação entre as duas variáveis estudadas, evidenciando a alta correlação destas.

A fim de entender se era possível classificar as caixas de germinação da mesma forma que o avaliador humano do programa de melhoramento genético, ou seja, atribuir notas de 1 a 5, fora feito o gráfico da Figura 5, que nos permite visualizar que não há intervalos de valores bem definidos para essa classificação, por exemplo, não pode-se dizer que caixas nota 4 possuem entre 100 e 500 plântulas (valores apenas ilustrativos). Tal impossibilidade é causada pelo erro de viés humano ocorrido quando o avaliador, de forma subjetiva, classifica as caixas, em notas, somente pela rápida visualização das mesmas, e não a partir de parâmetros mais precisos, como a contagem de plântulas. Também não se leva em conta fatores como diferentes níveis de desenvolvimento das plântulas e diferenças morfológicas das folhas, ambas causadas pela diferença genética dos cruzamentos, podendo enviesar a avaliação humana.

Na Figura 5 pode-se ainda notar outliers marcantes, circutados em vermelho e suas referentes caixas apresentadas (Figura 6), onde é possível confrontar a nota da avaliação visual e o número estimado de plântulas pelo método das Quinas de Harris. Pode-se verificar que essas caixas, com notas, respectivamente, 3, 2, 1 e 1, são caixas onde não houve avaliações condizentes com seus devidos potenciais; esperar-se-ia que fossem classificadas como nota 4 ou 5, uma vez que possuíam poucas plantas emergidas. Isso reforça a subjetividade da avaliação visual humana.

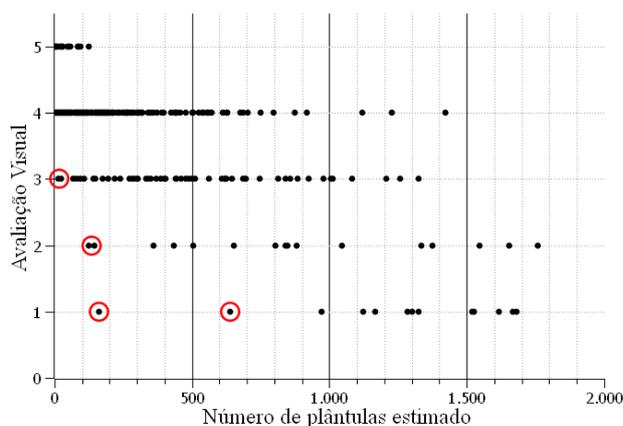


Figura 5 -Relação entre as avaliações visuais e número estimado de plântulas pelo modelo das Quinas de Harris. Ressalta-se os pontos circutados em vermelho, referentes aos outliers

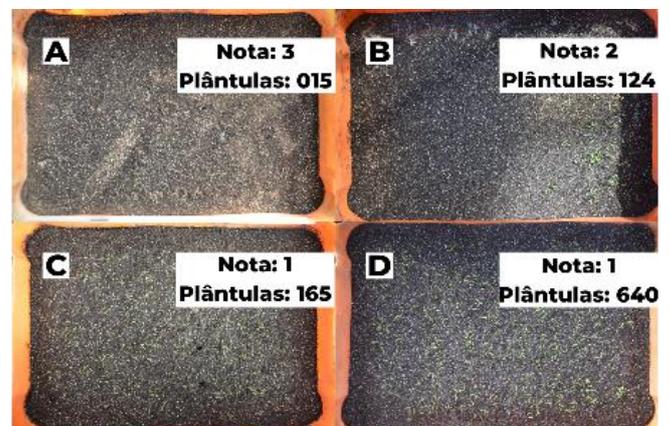


Figura 6 -Caixas de germinação tidas como outliers, confrontando a avaliação visual humana com o número estimado de plântulas pelo método das Quinas de Harris.

Grande parte da sílica presente no substrato, provinda da utilização de areia na confecção do mesmo, possui grande poder refletivo, sendo fácil de se remover no tratamento de imagem, pela diferenciação de brilho do espaço de cores HSV. Entretanto, uma parte menor, reflete de forma muito similar às folhas da cana-de-açúcar e com isso se tornam de difícil remoção pela comparação das bandas, então para contornar tal efeito, fora utilizado o algoritmo citado de remoção de pequenos objetos. Mas por este não distinguir qual objeto está removendo, acaba por entender algumas plântulas de tamanho reduzido (costumeiramente recém emergidas) como substrato e as remove da imagem. Este é um fator que pode causar, em alguns casos, subestimação na quantidade de plântulas pelo programa, causando assim pequena diminuição na estimativa do potencial germinativo.

5 CONCLUSÃO

O método que estima com maior precisão a quantidade de plântulas presentes nas caixas de germinação é o das Quinas de Harris, apresentando R^2 igual a 0,863. Assim, entende-se que tal metodologia apresenta precisão suficiente para substituir a subjetiva avaliação visual humana na avaliação do potencial germinativo dos cruzamentos de cana-de-açúcar.

6 AGRADECIMENTOS

Agradeço a todos funcionários da Ridesa/UFSscar (Rede Interuniversitária para o Desenvolvimento do Setor Sucroenergético) que auxiliaram no projeto, em especial ao Danilo Eduardo Cursi, pesquisador do programa de melhoramento genético de cana-de-açúcar da Ridesa, que além de surgir com a necessidade de um projeto que abordasse tal problemática, prestou auxílio com todas atividades que envolviam a Rede, emprestando ainda a câmera utilizada para obtenção das imagens.

7 REFERÊNCIAS BIBLIOGRÁFICAS

BRUNKHORST, M. J.; COETZEE, N.A.; NUSS, K.J. EFFICIENCY OF THE GERMINATION TEST FOR PREDICTING SUGARCANE SEEDLING NUMBERS AT MOUNT EDGEcombe. **PROC. SOUTH AFRICAN SUGAR TECHNOLOGY ASSOCIATION**, v. 74, p.234-237, 2000.

CAIEIRO, J. T.; PANOBIANCO, M.; BESPALHOK FILHO, J.C.; OHLSON, O.C. PHYSICAL PURITY AND GERMINATION OF SUGARCANE SEEDS (CARYOPSES) (SACCHARUM SPP.). **REVISTA BRASILEIRA DE SEMENTES**, v. 32, n. 2, p. 140-145, 2010 .

HARRIS, C. G. **A COMBINED CORNER AND EDGE DETECTOR**. DISPONÍVEL EM < [HTTP://WWW.BMVA.ORG/BMVC/1988/AVC-88-023.PDF](http://www.bmva.org/bmvc/1988/avc-88-023.pdf) >, DATA DE ACESSO 10/04/2019.

IBGE 2017. **PRODUÇÃO AGRÍCOLA MUNICIPAL : CULTURAS TEMPORÁRIAS E PERMANENTES, INFORMATIVO TÉCNICO 2017**, DISPONÍVEL EM < [HTTPS://BIBLIOTECA.IBGE.GOV.BR/INDEX.PHP/BIBLIOTECA-CATALOGO?VIEW=DETALHES&ID=766](https://biblioteca.ibge.gov.br/index.php/biblioteca-catalogo?view=detalhes&id=766) >, DATA DE ACESSO 07/04/2019.

OLIVEIRA, H. C.; GUIZILINI, V. C.; NUNES, I. P.; SOUZA J. R.. FAILURE DETECTION IN ROW CROPS FROM UAV IMAGES USING MORPHOLOGICAL OPERATORS. **IEEE GEOSCIENCE AND REMOTE SENSING LETTERS**, VOL. 15, NO. 7, P. 991-995, 2018, DOI: 10.1109/LGRS.2018.2819944.

YAMAGUCHI, O.; FUKUI, K.; MAEDA, K.I-. FACE RECOGNITION USING TEMPORAL IMAGE SEQUENCE. **PROCEEDINGS THIRD IEEE INTERNATIONAL CONFERENCE ON AUTOMATIC FACE AND GESTURE RECOGNITION**, NARA, P. 318-323, 1998, DOI: 10.1109/AFGR.1998.670968

YANG, W.; WANG, S.; ZHAO, X.; ZHANG, J.; FENG, J. GREENNESS IDENTIFICATION BASED ON HSV DECISION TREE. **INFORMATION PROCESSING IN AGRICULTURE**, VOL.2, IS. 3-4. P.149-160, 2015.