

## MACHINE LEARNING E PROCESSAMENTO DIGITAL DE IMAGENS UAV: UMA ABORDAGEM PARA ESTIMAR DISTRIBUIÇÃO LONGITUDINAL DE PLANTAS DE SOJA\*

FLÁVIA LUIZE PEREIRA DE SOUZA<sup>1</sup>, JOÃO RICARDO FAVAN<sup>2</sup>, JOSÉ RAIMUNDO DE SOUZA PASSOS<sup>3</sup>, MAURÍCIO ACCONCIA DIAS<sup>4</sup>, SÉRGIO CAMPOS<sup>5</sup>

\* Artigo extraído da dissertação do primeiro autor.

<sup>1</sup> Louisiana State University – LSU, School of Plant, Environmental, and Soil Sciences, 104 MB Sturgis Hall, LSU Campus, 70801, Baton Rouge, LA, Estados Unidos, Doutorado Universidade Estadual Paulista - UNESP, Faculdade de Ciências Agrônômicas, Departamento de Engenharia Rural, Av. Universitária, 3780, Altos do Paraíso, 18610-034, Botucatu, SP, Brasil, e-mail: flavia.luize@unesp.br

<sup>2</sup> Faculdade de Tecnologia de Pompéia - Shunji Nishimura - Fatec, Av. Shunji Nishimura, Departamento de Big Data no Agronegócio, 605, Distrito Industrial, 17580-000, Pompéia, SP, Brasil, e-mail: joao.favan@fatecpompeia.edu.br

<sup>3</sup> Universidade Estadual Paulista – UNESP, Instituto de Biociências, Departamento de Biodiversidade e Bioestatística, R. Prof. Dr. Antônio Celso Wagner Zanin, 250, Distrito de Rubião Junior, 18618-689, Botucatu, SP, Brasil, e-mail: jr.passos@unesp.br

<sup>4</sup> Fundação Hermínio Ometto - FHO, Departamento de Engenharia, Av. Dr. Maximiliano Baruto, 500, Jardim Universitário, 13607-339, Araras, SP, Brasil, e-mail: macdias@fho.edu.br

<sup>5</sup> Universidade Estadual Paulista - UNESP, Faculdade de Ciências Agrônômicas, Departamento de Engenharia Rural, Av. Universitária, 3780, Altos do Paraíso, 18610-034, Botucatu, SP, Brasil, e-mail: sergio.campos@unesp.br

**RESUMO:** É possível alcançar altas produtividades na lavoura de soja, por meio da semeadura com distribuição espacial adequada e uniforme das sementes. Com isso se faz importante o uso de tecnologias, como a aplicação do Processamento Digital de Imagens, que permite tratar as imagens coletadas e aperfeiçoá-las para a interpretação humana e, em seguida, a análise automática pelo computador, a partir da classificação do reconhecimento de padrões. O objetivo desta pesquisa foi testar métodos de Machine Learning para estimar a distribuição das plantas na linha de plantio da lavoura de soja. O modelo Random Forest apresentou melhor resultado com acurácia em torno de 65% em média, porém o algoritmo não obteve um resultado considerado satisfatório. É possível concluir que a dificuldade de classificação das distâncias entre plantas de soja com o modelo utilizado pode estar associada às variáveis qualidade de imagem, sobreposição das plantas de soja e a precisão do modelo.

**Palavras-chave:** Aeronave remotamente pilotada; Agricultura de precisão; Inteligência Artificial; Visão Computacional.

## MACHINE LEARNING AND DIGITAL UAV IMAGE PROCESSING: AN APPROACH TO ESTIMATE THE LONGITUDINAL DISTRIBUTION OF SOYBEAN PLANTS

**ABSTRACT:** It is possible to achieve high productivity in soybean crops, by sowing with adequate and uniform local distribution of the seeds. To attain this, the use of technologies is important - for example the application of Digital Image Processing that allows treatment of collected images and improvement for human interpretation, and then the automatic analysis by the computer, based on pattern recognition classification. The objective of this research was to test Machine Learning methods to estimate the distribution of plants in the soybean planting line. The Random Forest model showed the best result, where the resulting accuracy was 65% on average, but the algorithm did not obtain a good result. It is possible to conclude that the difficulty of classifying the distances between soybean plants with the model used may be associated with the variables of image quality, overlapping of soybean plants and the precision of the model.

**Keywords:** Artificial intelligence; Computer vision; Precision agriculture; Remotely piloted aircraft.

## 1 INTRODUÇÃO

A crescente demanda da soja no Brasil resultou na expansão das áreas agricultáveis. Na safra 2020/2021, a cultura mantém a tendência de crescimento com uma estimativa de produção de 133,7 milhões de toneladas, o equivalente a um crescimento de 7,1% em relação ao ciclo passado e mantém o Brasil como o maior produtor mundial da soja (SOJA, 2021).

Fatores como número de plantas e distanciamento entre elas afetam a qualidade e produtividade das culturas agrícolas. Garantir que a semeadura das sementes seja realizada com distribuição espacial adequada e uniforme, influencia na competitividade, eficácia do uso dos recursos ambientais e consequentemente na possibilidade de alcançar altas produtividades na lavoura (FISS *et al.*, 2018; CALONEGO *et al.*, 2011).

Ovanessoff e Plastino (2017) preveem que a Inteligência Artificial (I.A.) vai gerar um impulso no VAB (Valor Agregado Bruto) de 432 bilhões de dólares até 2035 para a Economia Brasileira.

Considerando este cenário e a importância agrícola e econômica da soja, a utilização de ferramentas como a I.A. proporciona otimização de processos, serviços e produtos, além de maximização na produção (ALVES *et al.*, 2018). Em virtude disto, o objetivo desta pesquisa foi testar métodos de *Machine Learning* para estimar a distribuição das plantas na linha de plantio da cultura de soja.

## 2 REVISÃO DE LITERATURA

### 2.1 Espaçamento entre plantas de soja

De acordo com Endres (1996), espaçamentos maiores ou menores do que o recomendado resulta em um aproveitamento ineficiente de recursos como água, luz e nutrientes, com desenvolvimento irregular e acamamento causado pela competição intraespecífica. Além disso, pode haver também o aparecimento de outras plantas que não são de interesse, e, portanto, a distribuição

longitudinal desuniforme da cultura pode provocar alterações na produtividade.

### 2.2 Técnicas para mensuração da distribuição longitudinal da soja de forma automática

A utilização de plataformas não tripuladas para a obtenção de imagens permitiu a evolução do sensoriamento remoto (mensuração e mapeamento à distância). Com as imagens aéreas obtidas pela Aeronave Remotamente Pilotada (UAV do inglês *Unmanned Aerial Vehicle*) é possível construir mosaicos para a realização de análises, uma ferramenta que auxiliará na agricultura de precisão (ANAC, 2015). Deste modo, mapas de recomendações podem ser elaborados para auxiliar no planejamento da área de plantio (LELONG *et al.*, 2008).

### 2.3 Visão Computacional

Nas últimas décadas, o objetivo da visão computacional tem sido a utilização de técnicas de aprendizado de máquina para problemas de visão de computador, com a utilização da visão computacional, informações são extraídas das imagens para diferentes finalidades (SZELISKI, 2010).

De forma geral, os sistemas de visão computacional abrangem o reconhecimento de objetos em imagens e transformações dos objetos em informações para serem processadas e empregadas em algum sistema, se aproximando da inteligência humana (SZELISKI, 2010).

### 2.4 Aprendizado de máquina

O aprendizado de máquina é um subcampo da Inteligência Artificial, um ramo da ciência da computação que projeta dispositivos para simular a capacidade humana de raciocinar, perceber, tomar decisões e resolver problemas (MONARD; BARANAUKAS, 2000). Abrange um campo de estudo amplo, com teorias, métodos e tecnologias (MONARD; BARANAUKAS, 2000). Portanto, com a I.A. é possível sistematizar e automatizar tarefas intelectuais e

permitir o maior controle dos processos envolvidos na agricultura (RUSSEL; NORVIG, 2013).

Existem três tipos de aprendizagem do modelo: aquele que possui informações de dados rotulados (*label*) previamente conhecidas, chamado de aprendizagem supervisionada; outro com dados não rotulados, chamado de aprendizagem não supervisionada e por último o processo de aprendizagem por reforço, que é um misto entre os aprendizados anteriores, a cada sinal de saída o modelo é recompensado conforme o acerto (SILVA *et al.*, 2010).

## 2.5 Trabalhos correlatos

As pesquisas direcionadas para automatizar tarefas no campo permitem a geração de mapas de densidade de plantas daninhas em culturas de cereais e detectar linhas de colheita com o uso de imagens de câmeras colocadas em trator.

Ji (2011) avaliou algoritmo para detectar a linha central da linha da cultura, para a localização de plantas em tempo real de sistema de pulverização de taxa variável. Marchant e Brivot (1995) utilizou esse método para identificar as linhas da cultura em tempo real, e os erros encontrados foram de 12,5 mm. O algoritmo RHT (*Randomized Hough Transform*) foi utilizado para identificar a linha de colheita, porém não é habilitado para identificar culturas com alta quantidade de ervas daninhas e perdas na cultura analisada (XU, 1990).

O terceiro método embasado em pontos de fuga é utilizado por Pla *et al.* (1997) para a orientação de um veículo de navegação na linha de cultivo. O método de transformação de Hough e o ponto de fuga são integrados no projeto para detectar as linhas de plantio do trigo durante o início do estágio de crescimento, por meio da transformação da imagem colorida para escala de cinza, em seguida a segmentação da imagem para delimitar as plantas e por último a retirar as plantas daninhas (JIANG *et al.*, 2016).

O artigo de Zhang *et al.* (2018) apresentaram um novo método (um algoritmo de processamento) para estimar a distância

entre plantas de milho, com o uso do UAV para a coleta de imagens de experimentos com plantas artificiais e reais. Os procedimentos utilizados incluem a segmentação do milho para a retirada dos ruídos das imagens, como plantas daninhas, o ajuste dos pontos centrais no milho referentes à segmentação e a mensuração das distâncias entre plantas.

Para o cálculo da distância entre plantas, os algoritmos provavelmente necessitam da identificação da mesma primeiro e o estudo da identificação de mudas de *Eucalyptus* mostrou que é possível a aplicação de redes neurais artificiais com o *software Neurodic* em ortomosaicos adquiridos pelo processamento das imagens do UAV, sendo necessário no experimento o pós-processamento para a correção de erros (RUZA *et al.*, 2017). Hentz *et al.* (2018) também utilizaram algoritmos para a detecção de árvores com UAV de forma automática, a ferramenta construída é chamada de *TreeDetection*.

Então, é possível ver nos trabalhos apresentados algoritmo para detectar a linha central da linha da cultura, orientação de um veículo de navegação na linha de cultivo, método de transformação de Hough e o ponto de fuga para detectar as linhas de plantio do trigo, segmentação e a mensuração das distâncias entre plantas de milho e a aplicação de redes neurais artificiais para a identificação de mudas de *Eucalyptus*.

Apesar dos vários métodos e ferramentas apresentados, estimar a distribuição das plantas na linha de plantio da soja ainda é feita de forma manual através do uso de trenas. Neste ponto o trabalho procura explorar de forma diferente o tema por meio de visão computacional e processamento de imagens coletadas pelo UAV utilizando *Machine Learning*.

## 3 MATERIAL E MÉTODOS

O experimento foi instalado na Fazenda Experimental Lageado, pertencente à Faculdade de Ciências Agrônômicas da UNESP, situada no município de Botucatu – SP.

Implantou-se a cultura da soja (*Glycine max* (L) Merr.) TMG INTACTA RR IIPRO

(germinação de 80% e pureza de 99%), com cerca de 16 sementes por metro linear, supondo uma população de 355.000 plantas ha<sup>-1</sup>.

O delineamento experimental foi dividido em faixas de 50 metros cada, em sistema de plantio direto sobre palha de aveia preta, com velocidades de plantio de 4, 6 e 8 km·h<sup>-1</sup>, executando 4 repetições para cada velocidade de plantio, totalizando 12 parcelas.

A porcentagem de espaçamentos entre plantas (EEP) normais (N), falhos (F) e duplos (D), foi obtida de acordo com Kurachi *et al.* (1989) (Tabela 1). Os espaçamentos são classificados como normais quando situados entre 0,5 e 1,5 vezes o espaçamento médio esperado (neste caso, 6,25 cm).

**Tabela 1.** Índice de espaçamentos Aceitáveis, duplos e falhos (%).

Classificação	Espaçamentos EEP
Aceitáveis	3,1 cm < EEP < 9,4 cm
Falhas	EEP > 9,4 cm
Duplos	EEP < 3,1 cm

Fonte: Adaptado, Kurachi *et al.* (1989).

### 3.4 Distribuição Longitudinal estimada com I.A.

As imagens foram obtidas utilizando uma Aeronave Remotamente Pilotada (UAV) de mapeamento, cujo método possibilita maior resolução se comparado à obtida pelos métodos tradicionais de sensoriamento remoto, como por exemplo as imagens de satélite. O UAV utilizado para a área foi o modelo Dji Phantom 3 PRO e as imagens foram coletadas no dia 10 de dezembro de 2019 em altura de 10m e com sobreposição de 70%.

As imagens resultantes do processo de recorte dos ortomosaicos foram processadas a fim de identificar os conjuntos de plantas, denominadas de grupos, e posteriormente estimar as distâncias entre eles, para classificação de possível falha. Este processamento foi realizado por meio do desenvolvimento de um *script* em linguagem de programação Python 3.8.2 (ROSSUM; DRAKE, 2009), com as bibliotecas *SciKit Learn* e *OpenCV* 4.4.0.44 de processamento de imagens (BRADSKI, 2000).

As imagens foram separadas em seus canais vermelho (R), verde (G) e azul (B), utilizando a técnica *Split* (BRADSKI, 2000). Foi selecionado para o processamento o canal verde, que apresentou maior distinção entre as plantas de soja e o solo. A técnica de *Edge* (BRADSKI, 2000) foi utilizada para detecção dos contornos que formam os grupos, e esses contornos detectados foram preenchidos com a

técnica *Erode*. Essas são técnicas de processamento de imagens (BRADSKI, 2000).

Cada grupo foi referenciado na imagem e suas distâncias euclidianas (em pixels) mensuradas. A medida de distância foi realizada por meio da atribuição de um grupo como referência em relação aos demais grupos considerados como alvos da medida. O objetivo do algoritmo foi encontrar a menor distância entre o grupo referência e o grupo alvo, considerado o grupo próximo. O grupo considerado alvo de uma rodada de medidas de distâncias recebe o papel de referência para a próxima tomada de medidas, conforme adaptação da metodologia proposta por Othman *et al.* (2018).

Durante as rodadas de tomada de medidas, os dados obtidos dos grupos foram armazenados em arquivos *comma separated values* (CSV), sendo que cada linha de dados era formada pela posição vertical (RefX) e horizontal (RefY) do grupo de referência, posição vertical (TargetX) e horizontal (TargetY) do grupo alvo e a distância em pixels entre esses grupos (D), conforme adaptado de Othman *et al.* (2018). Foram gerados 12 arquivos CSVs, correspondendo a cada bloco de cada velocidade de plantio experimentado.

Os dados presentes nesses arquivos CSVs foram submetidos a análise para identificação manual das distâncias medidas, do grupo de referência diretamente para o grupo mais próximo. As distâncias que escolheram o grupo mais próximo corretamente foram

assinaladas com valor verdadeiro (1), enquanto as medidas lateralizadas e medição de mesmo bloco foram assinaladas com o valor falso (0), conforme adaptação para o projeto atual das metodologias propostas nos trabalhos de Vink e Hann (2015) e Chicco (2017).

Todos os arquivos de dados gerados foram agregados em um único arquivo, do qual foram selecionadas aleatoriamente 129 amostras positivas (classe 1) e 129 amostras negativas (classe 0). Essas amostras eram formadas pelos atributos descritos anteriormente para a formação do arquivo CSV (RefX, RefY, TargetX, TargetY, D) e definido como rótulo, atributo utilizado na classificação, a classe da amostra como positiva (1) e negativa (0). Posteriormente, os dados foram normalizados pelo processo de normalização por máximo e mínimo (CASTRO; FERRARI, 2016), formando assim o conjunto de dados utilizado no processo de treinamento da inteligência artificial.

Foram treinados os modelos provenientes de quatro algoritmos de aprendizado de máquinas, sendo eles: *Dummy Classifier*, posicionado como referencial negativo (BAUMANN, 2016); *Logistic Regression* (ALMQUIST; BUTTS, 2013);

*Random Forest* (BREIMAN, 2001) e *Support Vector Machine* (SVM) (SMOLA; SCHÖLKOPF, 2004). Foi utilizado o processo de validação cruzada, com dez partições, para treinamento e validação de todos os algoritmos e esses foram avaliados segundo suas métricas de Acurácia, Precisão, Revocação e Média F (SHARMA *et al.*, 2020).

O modelo com melhores métricas de classificação foi implementado ao *script* principal de análise de imagens para definir se a medida de distância obtida era aceitável (amostra positiva) ou não aceitável (amostra negativa).

#### 4 RESULTADOS E DISCUSSÃO

Para o cálculo de distribuição da soja na linha de plantio com I.A., os modelos utilizados para a solução do problema foram: *Random Forest*, *Logistic Regression*, SVM e *Dummy Classifier*. Na Tabela 2, estão apresentados os resultados obtidos com os modelos descritos anteriormente, permitindo a comparação entre eles. Os modelos não conseguiram estimar a distâncias entre plantas individuais, mas as distâncias entre conjunto de plantas (grupo).

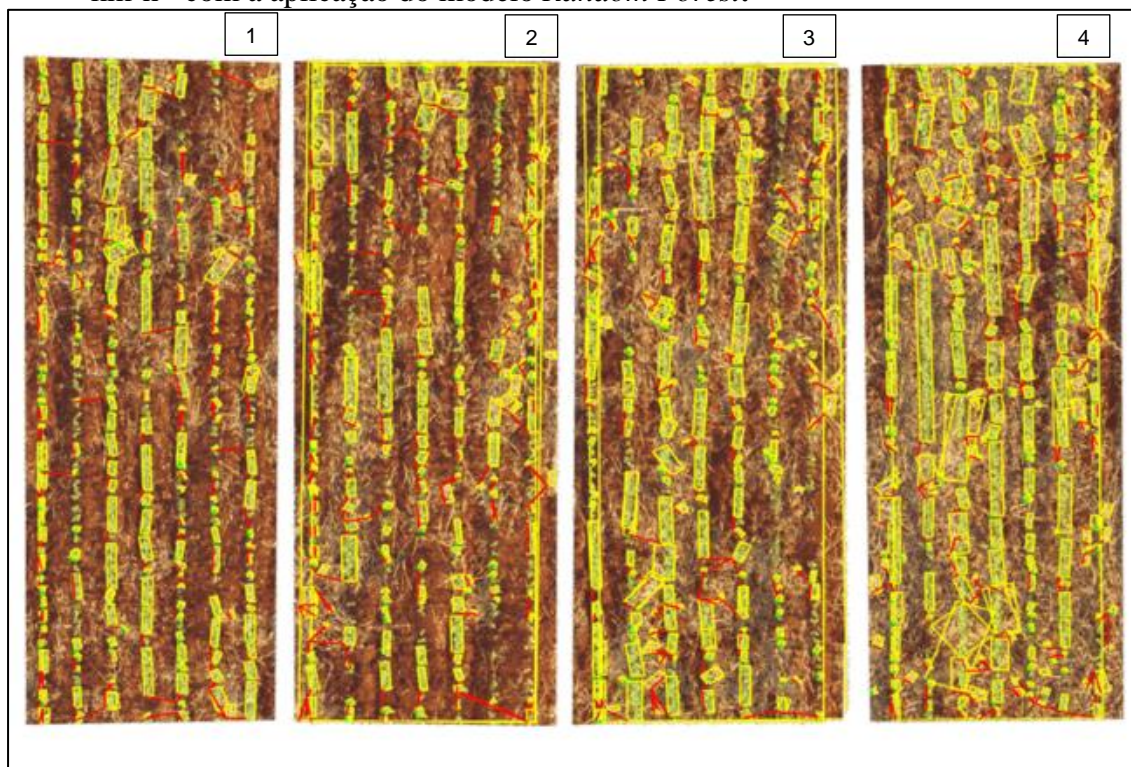
**Tabela 2.** Resultados da validação cruzada para os modelos de *Random Forest*, *Logistic Regression*, SVM e *Dummy Classifier*, avaliados segundo suas métricas de Acurácia, Valor-F, Precisão e Revocação.

Modelo	Acurácia	Valor-F	Precisão	Revocação
<i>Random Forest</i>	0,655	0,655	0,655	0,655
<i>Logistic Regression</i>	0,581	0,581	0,582	0,581
SVM	0,539	0,538	0,550	0,539
<i>Dummy Classifier</i>	0,496	0,496	0,489	0,496

O modelo *Random Forest* apresentou melhor resultado, com acurácia resultante em torno de 65% em média; todavia, o algoritmo não obteve um bom resultado, tendo apresentado distâncias lateralizadas entre os grupos de plantas de soja.

Na velocidade de 4 km·h<sup>-1</sup>, o modelo *Random Forest* identificou grupos de plantas de soja que estavam muito próximas como um único grupo, como visto nas linhas amarelas da Figura 11; devido à sobreposição das plantas, sua identificação individual foi dificultada.

**Figura 1.** Recorte do ortomosaico em blocos (1, 2, 3 e 4) do plantio da soja em velocidade de 4 km·h<sup>-1</sup> com a aplicação do modelo *Random Forest*.



Linhas amarelas - conjunto de plantas de soja; linhas vermelhas – distância entre conjunto de plantas acima de 9,4cm; linhas azuis- distâncias entre conjunto de plantas abaixo de 3,1cm e linhas verdes – distância entre conjuntos de plantas entre 3,1cm e 9,4 cm.

**Fonte:** Souza (2019).

A distância entre os grupos permite a identificação de falhas, mas por se tratar de imagens tiradas por um UAV, com uma altura de 10 m, a qualidade diminui conforme aumenta a altura do voo para o caso da soja - uma cultura que é de dimensões pequenas, vista de cima e com sobreposição, a perda da qualidade da imagem causou uma lateralização e contagem de distâncias do mesmo grupo.

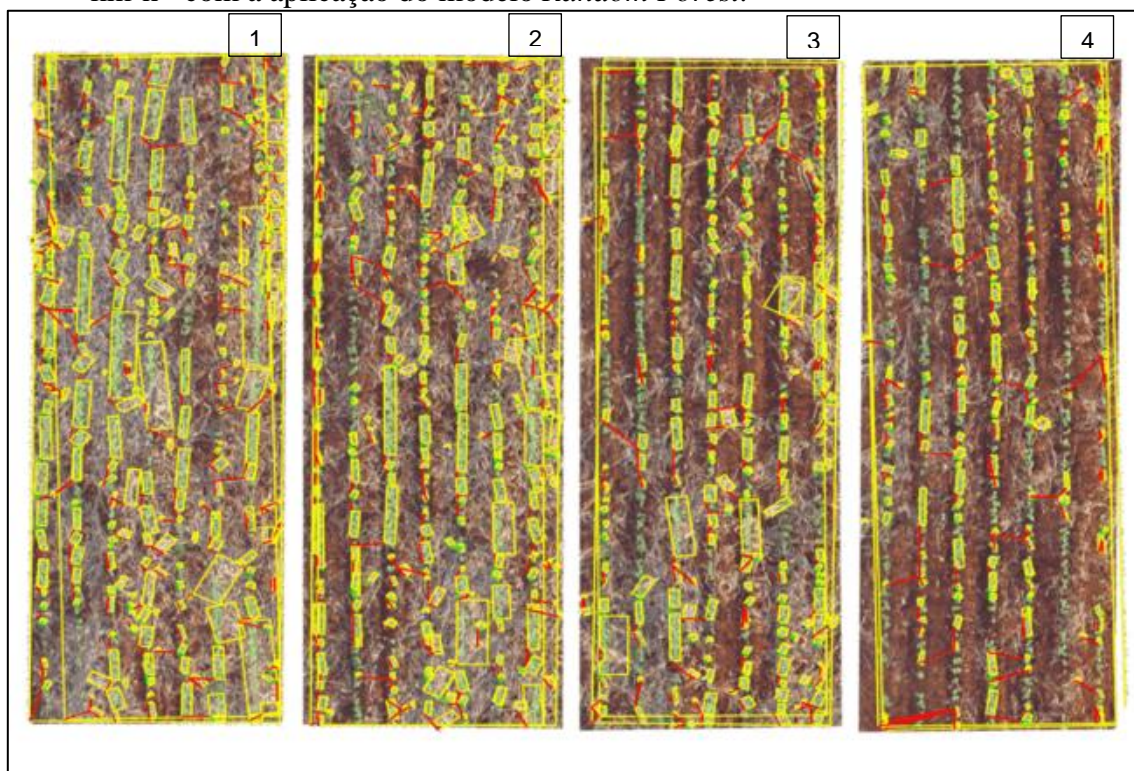
As fotos com o UAV foram tiradas quando as plantas estavam no período vegetativo (V4). O processo não foi feito na emergência, porque a soja estaria muito

pequena pra um voo a 10 m de altitude, e também não foram realizadas em um estágio fenológico após o V4 para evitar a sobreposição, porém, mesmo assim na altura de voo utilizada houve sobreposição da soja.

Na Figura 12, os grupos dos ortomosaicos com velocidades de 6 km·h<sup>-1</sup> apresentaram medidas do mesmo grupo e lateralizadas, e algumas plantas de soja não foram identificadas pelo algoritmo, o que levou ao valor de 65% de acurácia das falhas entre grupo de plantas.



**Figura 2.** Recorte do ortomosaico em blocos (1, 2, 3 e 4) do plantio da soja em velocidade de 6  $\text{km}\cdot\text{h}^{-1}$  com a aplicação do modelo *Random Forest*.



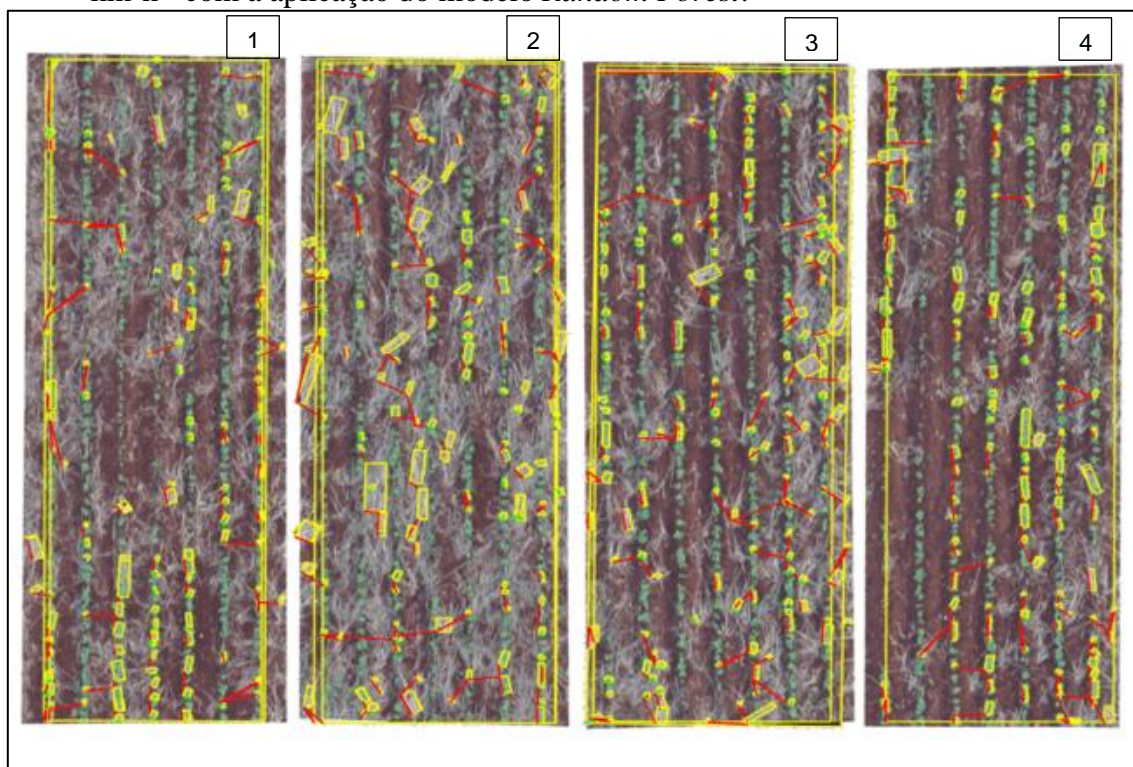
Linhas amarelas - conjunto de plantas de soja; linhas vermelhas – distância entre conjunto de plantas acima de 9,4cm; linhas azuis- distâncias entre conjunto de plantas abaixo de 3,1cm e linhas verdes – distância entre conjuntos de plantas entre 3,1cm e 9,4 cm.

**Fonte:** Souza (2019).

Na velocidade de 8  $\text{km}\cdot\text{h}^{-1}$  (Figura 13), o modelo *Random Forest* também estimou a distância entre conjuntos de plantas, uma porcentagem de medidas lateralizadas e do mesmo bloco e algumas plantas de soja não identificadas conforme as velocidades 4 e 6  $\text{km}\cdot\text{h}^{-1}$ . Contudo, comparando os recortes das três velocidades, muitas plantas de soja não

foram identificadas na maior velocidade como visto na Figura 13; o ortomosaico apresenta uma coloração mais escura, porque no momento em que o UAV estava passando pela área, provavelmente o sol estava atrás das nuvens, ocasionando uma qualidade de imagem diferente das outras velocidades.

**Figura 3.** Recorte do ortomosaico em blocos (1, 2, 3 e 4) do plantio da soja em velocidade de 8 km·h<sup>-1</sup> com a aplicação do modelo *Random Forest*.



Linhas amarelas - conjunto de plantas de soja; linhas vermelhas – distância entre conjunto de plantas acima de 9,4cm; linhas azuis- distâncias entre conjunto de plantas abaixo de 3,1cm e linhas verdes – distância entre conjuntos de plantas entre 3,1cm e 9,4 cm.

**Fonte:** Souza (2019).

No experimento, o espaçamento médio esperado foi de 6,25 cm de acordo com a população de 16 sementes por metro. Fatores como a distância curta entre as plantas, período de sobreposição, e a altitude de voo a 10 m (altura mínima para realizar o planejamento do voo automático pelo programa DroneDeploy) podem ter interferido no resultado do modelo *Random Forest*.

O trabalho de Zhang *et al.* (2018) mostram que o cálculo das distâncias entre plantas de milho é possível por imagens de UAV e a precisão das distâncias está relacionada com a altura das imagens coletadas pelo UAV. A sobreposição de duas plantas de milho de acordo com o ponto de vista da câmera utilizada gerou valores superestimados das distâncias do intervalo do milho.

Para de Zhang *et al.* (2018) , os fatores que podem ter favorecido a sobreposição das plantas de milho foram as distâncias mais próximas entre plantas ou altitudes de voo mais baixas. As maiores porcentagens de erro foram para a menor distância do milho (9 cm) em

todas as variações de altitude do UAV. A altura do voo do experimento da soja foi 10 m, maior que a utilizada por de Zhang *et al.* (2018), porém a distância esperada entre plantas é menor (6,25 cm), o que pode ter facilitado a sobreposição.

A altitude do voo, se comparada ao experimento de de Zhang *et al.* (2018) , foi alta; porém, para o planejamento do voo a altura mínima necessária era de 10m e a área total do experimento de soja foi mais extensa. Quanto maior for a altura de voo, maior a área coberta e maior o GSD (*Ground Sample Distance*), distância no solo representado por cada pixel.

O detalhamento da imagem aumenta conforme a altitude do voo diminui, mas a área de cobertura será menor e consequentemente em áreas extensas mais voos serão necessários. Portanto, vários pontos precisam ser levados em consideração, como se o GSD utilizado é o suficiente para atingir o objetivo do trabalho, qual o tempo disponível para a quantidade de voos, a autonomia de voo do UAV, e se for um serviço de campo prestado para um cliente



quantas diárias serão necessárias para cobrir toda a área com detalhes.

No modelo *Random Forest* para o teste Wilcoxon (não paramétrico), todos os tratamentos são diferentes estatisticamente e para o teste de Tukey (paramétrico - mais rigoroso), somente o recorte do ortomosaico de velocidade  $8 \text{ km} \cdot \text{h}^{-1}$  é diferente estatisticamente da velocidade  $4 \text{ km} \cdot \text{h}^{-1}$ .

Os algoritmos de aprendizado de máquina não apresentaram um bom desempenho com o método utilizado para o cálculo de distância das plantas. Um sistema automatizado que faça uso de outros métodos, como diferentes alturas ou a utilização de imagens de celular, precisam ser estudados para melhorar a aplicabilidade do método proposto.

O modelo *Random Forest* identificou na velocidade de  $4 \text{ km} \cdot \text{h}^{-1}$  valores de 55,26% de falhas, 29,67% de aceitáveis e 15,07% de espaçamentos duplos; na velocidade de  $6 \text{ km} \cdot \text{h}^{-1}$  foram encontrados 61,73% de falhas, 25,49% de aceitáveis e 12,79% de duplas; e na velocidade de  $8 \text{ km} \cdot \text{h}^{-1}$  as falhas foram de 61,60%, aceitáveis com 23,67% e 14,73 de duplas. O resultado da porcentagem de espaçamentos entre plantas do modelo *Random Forest* foi obtido de acordo com a classificação de Kurachi *et al.* (1989) (Tabela 1).

O algoritmo apresentou valores altos de falhas e espaçamentos duplos quando comparados com o cálculo das distâncias entre plantas realizada com trena no campo, e consequentemente o modelo *Random Forest* constatou uma porcentagem baixa de espaçamentos aceitáveis.

O estudo contribui para a literatura, visto que orienta a busca por outros métodos para o cálculo de distâncias entre plantas após a emergência com inteligência artificial, que é um fator auxiliador na compreensão dos motivos da variabilidade das distâncias entre plantas e sua relação com a produtividade da cultura da soja.

## 5 CONCLUSÕES

Neste trabalho, foi apresentado um procedimento que foi desenvolvido para estimar a distância do intervalo entre plantas de soja usando um UAV. Foi verificado que o

modelo *Random Forest* teve um bom desempenho em identificar distância entre grupo de plantas, no entanto identificar a distância entre plantas foi uma limitação do estudo e precisa ser investigada em trabalhos futuros. A dificuldade de classificação das distâncias entre plantas de soja pode estar associada às variáveis qualidade de imagem, sobreposição das plantas de soja e a precisão do modelo. Há possibilidade de utilizar outras metodologias para aumentar a precisão do modelo, como a utilização de imagens de alta resolução.

## 6 AGRADECIMENTOS

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

## 7 REFERÊNCIAS

ALMQUIST, Z. W.; BUTTS, C. T. Dynamic network logistic regression: a logistic choice analysis of inter- and intra-group blog citation dynamics in the 2004 us presidential election. **Political Analysis**, Cambridge, v. 21, n. 4, p. 430-448, 2013.

ALVES, G. R.; TEIXEIRA, I. R.; MELO, F. R.; SOUZA, R.T.G.; SILVA, A.G. Estimating soybean yields with artificial neural networks. **Acta Scientiarum - Agronomy**, Maringá, v. 40, n. 1, p. 1-9, 2018.

ANAC. **RPAS - Sistemas de aeronaves remotamente pilotadas**. Brasília, DF: Anac, 2015. Disponível em: <https://www2.anac.gov.br/rpas/>. Acesso em: 02 jun. 2020.

BAUMANN, P.; LANGHEINRICH, M.; DEY, A.; SANTINI, S. Quantifying the Uncertainty of Next-Place Psplirictions. In: EAI INTERNATIONAL CONFERENCE ON MOBILE COMPUTING, APPLICATIONS AND SERVICES, 8., 2016, Cambridge. **Proceedings** [...]. Cambridge: ACM, 2016. Disponível em: <http://eudl.eu/doi/10.4108/eai>.

30-11-2016.2267147. Acesso em: 24 nov. 2020.

BRADSKI, G. The OpenCV Library. **Dr Dobb's journal: software tools for the professional programmer**, San Mateo, v. 25, n. 11, p. 120 -125, 2000.

BREIMAN, L. Random Forests. **Machine Learning**, Norwell, v. 45, n. 1, p. 5-32, 2001.

CALONEGO, J. C.; POLETO, L. C.; DOMINGUES, F. N.; TIRETEN, C. S. Produtividade e crescimento de milho em diferentes arranjos de plantas. **Agrarian**, Dourados, v. 4, n. 12, p. 84-90, 2011.

CASTRO, L. N. FERRARI, D. G. **Introdução à mineração de dados**. 1. ed. São Paulo: Saraiva, 2016.

CHICCO, D. Ten quick tips for machine learning in computational biology. **BioData Mining**, Filadélfia, v. 10, n. 35, p. 1-17, 2017.

ENDRES, V. C. Espaçamento, densidade e época de semadura. *In*: EMBRAPA. Centro de Pesquisa Agropecuária do Oeste. **Soja: Recomendações técnicas para Mato Grosso do Sul e Mato Grosso**. Dourados: Embrapa, 1996. p. 82-85. (Circular técnica, n. 3).

FISS, G.; SCHUCH, L. O. B.; PESKES, S. T.; CASTELLANOS, C. I. S.; MENEGHELLO, G. E.; AUMONDE, T. Z. Produtividade e características agrônomicas da soja em função de falhas na semeadura. **Ciências agrárias**, Belém, PA, v. 61, p. 1-7, 2018.

HENTZ, A. M. K.; CORTE, A. P. D.; PÉLLICO NETO, S.; STRAGER, M. P.; SCHOENINGER, E. R. Treedetection: automatic tree detection using UAV-based data. **Revista Floresta**, Curitiba, v. 48, n. 3, p. 393-402, 2018.

JI, R.; QI, L. Crop-row detection algorithm based on Random Hough Transformation. **Mathematical and Computer Modelling**, Guildford, v. 54, n. 3/4, p. 1016-1020, 2011.

JIANG, G.; WANG, W.; WANG, X.; LIU, H. Wheat rows detection at the early growth stage based on Hough transform and vanishing point. **Computers and Electronics in Agriculture**. Nova York, v. 123, p. 211-223, 2016.

KURACHI, S. A. H.; COSTA, J. A. S.; BERNARDI, J. A.; COELHO, J. L. D.; SILVEIRA, G. M. Avaliação tecnológica de semeadoras e/ou adubadoras: tratamento de dados de ensaios e regularidade de distribuição longitudinal de sementes. **Bragantia**, Campinas, v. 48, n. 2, p. 249-262, 1989.

LELONG, C. C.; BURGER, P.; JUBELIN, C.; ROUSC, B.; LABLLÉ, S.; BARET, F. Assessment of unmanned aerial vehicles imagery for quantitative monitoring of wheat crop in small plots. **Sensors**, Basel, v. 8, n. 5, p. 3557-3585, 2008.

MARCHANT, J. A.; BRIVOT, R. Real-time tracking of plant rows using a Hough transform. **Real-Time Imaging**. Londres, v. 1, n. 1, p. 363-371, 1995.

MONARD, M. C.; BARANAUKAS, J. A. **Aplicações de Inteligência Artificial: uma visão geral**. São Carlos: Instituto de Ciências Matemáticas e de Computação de São Carlos, 2000.

OTHMAN, N. A.; SALUR, M. U.; KARAKOSE, M.; AYDIM, I. An Embedded Real-Time Object Detection and Measurement of its Size. *In*: INTERNATIONAL CONFERENCE ON ARTIFICIAL INTELLIGENCE AND DATA PROCESSING, 2018. Malatya. **Anais [...]**. Malatya: IEEE, 2018. p. 1-4. Disponível em: <https://doi.org/10.1109/IDAP.2018.8620812>. Acesso em: 3 jan. 2021.

OVANESSOFF, A.; PLASTINO, E. **Como a inteligência artificial pode acelerar o crescimento da América do Sul**. [S. l.]: Accenture, 2017.

PLA, F; SANCHIZ, J.M; MARCHANT, J.A.; BRIVOT, R. Building perspective models to

guide a row crop navigation vehicle. **Image and Vision Computing**, Guildford, v. 15, n. 6, p. 465-473, 1997.

ROSSUM, G. V.; DRAKE, F. L. **Python 3 Reference Manual**. Scotts Valley: CreateSpace, 2009.

RUSSEL, S.; NORVIG, P. **Inteligência Artificial**. 2. ed. Rio de Janeiro: Campos, 2013.

RUZA, M. R.; CORTE, A. P. D.; HENTZ, A. M. K.; SANQUETA, C. R.; SILVA, C. A.; SCHOENINGER, E. R. Inventário de Sobrevivência de povoamento de Eucalyptus com uso de Redes Neurais Artificiais em Fotografias obtidas por VANTs. **Advances in Forestry Science**, Cuiabá, v. 4, n. 1, p. 83-88, 2017.

SHARMA, R.; RIBEIRO, B.; PINTO, A. M.; CARFDOSO, F. A. Exploring geometric feature hyper-space in data to learn representations of abstract concepts. **Applied Sciences**, Basel, v. 10, n. 6, p. 1-28, 2020.

SILVA, I. N.; SPATTI, D. H.; FLAUZINO, R. A. **Redes Neurais Artificiais para Engenharia e Ciências Aplicadas**. São Paulo: Artliber, 2010.

SMOLA, A. J.; SCHÖLKOPF, B. A tutorial on support vector regression. **Statistics and Computing**, Heidelberg, v. 14, n. 3, p. 199-222, 2004.

SOJA. **Acompanhamento da Safra Brasileira: grãos**, Brasília, DF, v. 7, n. 4, p. 1-85, 2021.

SOUZA, F. L. P. **Inteligência Artificial utilizada na determinação longitudinal da cultura da soja**. Dissertação (Mestrado em Agronomia/Energia na Agricultura) – Faculdade de Ciências Agrônômicas, Universidade Estadual Paulista “Júlio de Mesquita Filho”, Botucatu, 2021.

SZELISKI, R. **Computer Vision: Algorithms and Applications**. Seattle, Springer, 2010.

VINK, J. P.; DE HAAN, G. Comparison of machine learning techniques for target detection. **Artificial Intelligence Review**, Oxford, v. 43, n. 1, p. 125-139, 2015.

XU, L.; OJA, E.; KULTANEN, P. A new curve detection method: Randomized hough transform (rht). **Pattern Recognition Letters**. Amsterdã, v. 11, n. 5, p. 331-338, 1990.

ZHANG, J.; BASSO, B.; PRICE, R. F.; PUTMAN, G.; SHUAI, G. Estimating plant distance in maize using Unmanned Aerial Vehicle (UAV). **PLOS ONE**, San Francisco, v. 13, n. 4, p. 1-22, 2018.

## APÊNDICE A – Código do trabalho

O *script* do algoritmo do trabalho pode ser encontrado em:

[https://gitlab.com/projetosoja/distsoja/-/blob/1338daedc263adff79e3f951687a994af11ee0af/oficial/distance\\_soja\\_teste3.py](https://gitlab.com/projetosoja/distsoja/-/blob/1338daedc263adff79e3f951687a994af11ee0af/oficial/distance_soja_teste3.py)