

FLÁVIA LUIZE PEREIRA DE SOUZA

**INTELIGÊNCIA ARTIFICIAL UTILIZADA NA DETERMINAÇÃO LONGITUDINAL
DA CULTURA DA SOJA**

Botucatu

2021

FLÁVIA LUIZE PEREIRA DE SOUZA

**INTELIGÊNCIA ARTIFICIAL UTILIZADA NA DETERMINAÇÃO LONGITUDINAL
DA CULTURA DA SOJA**

Dissertação apresentada à Faculdade de Ciências Agronômicas da Unesp, Campus de Botucatu, para obtenção do título de Mestre em Agronomia – Energia na Agricultura.

Orientador: Prof. Dr. Sérgio Campos

Coorientador: Prof. Dr. Anderson Antonio da Conceição Sartori

Botucatu

2021

S729i Souza, Flávia Luize de
Inteligência artificial utilizada na determinação longitudinal da cultura da soja / Flávia Luize de Souza. --Botucatu, 2021
53 p.

Dissertação (mestrado) - Universidade Estadual Paulista (Unesp), Faculdade de Ciências Agrônômicas, Botucatu
Orientador: Sérgio Campos
Coorientador: Anderson Antonio da Conceição Sartori

1. Engenharia Agrônômica. 2. Inteligência Artificial. 3. *Machine Learning*. 4. Plantabilidade. 5. Soja. I. Título.

Sistema de geração automática de fichas catalográficas da Unesp. Bibliotecada Faculdade de Ciências Agrônômicas, Botucatu. Dados fornecidos pelo autor(a).

Essa ficha não pode ser modificada.



UNIVERSIDADE ESTADUAL PAULISTA

Câmpus de Botucatu



CERTIFICADO DE APROVAÇÃO

TÍTULO DA DISSERTAÇÃO: INTELIGENCIA ARTIFICIAL UTILIZADA NA DETERMINAÇÃO LONGITUDINAL DA CULTURA DA SOJA

AUTORA: FLÁVIA LUIZE PEREIRA DE SOUZA

ORIENTADOR: SÉRGIO CAMPOS

COORIENTADOR: ANDERSON ANTONIO DA CONCEIÇÃO SARTORI

Aprovada como parte das exigências para obtenção do Título de Mestra em AGRONOMIA (ENERGIA NA AGRICULTURA), pela Comissão Examinadora:

Prof. Dr. SÉRGIO CAMPOS (Participação Virtual)
Engenharia Rural e Socioeconomia / Faculdade de Ciências Agrônomicas de Botucatu - UNESP

Prof. Dr. OSVALDO CESAR PINHEIRO DE ALMEIDA (Participação Virtual)
Análise e Desenvolvimento de Sistemas / Faculdade de Tecnologia de Botucatu

Prof. Dr. JOÃO RICARDO FAVAN (Participação Virtual)
Big Data no Agronegócio / Faculdade de Tecnologia de Pompeia

Botucatu, 25 de fevereiro de 2021

AGRADECIMENTOS

Ao Prof. Dr. Sérgio Campos, pela orientação e ensinamentos.

Aos professores que me acompanharam durante a pós-graduação, em especial ao Prof. Dr. José Raimundo de Souza Passos e ao Prof. Dr. João Ricardo Favan pela orientação e atenção durante as etapas desenvolvidas, e ao Prof. Dr. Paulo Roberto Arbex Silva por disponibilizar a área experimental para a realização do projeto.

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

Aos grupos de pesquisas LINEAR (Grupo de Pesquisas Avançadas em Inteligência Artificial no Setor Agroflorestal e ao GPD (Grupo de Plantio Direto), pelo apoio e incentivo.

A todos os envolvidos direta e indiretamente nesta pesquisa, meus cordiais agradecimentos.

RESUMO

Em virtude da importância agrícola da soja no contexto econômico, faz-se necessário o seu acompanhamento com o uso de técnicas buscando a maximização da produtividade e qualidade do produto. É possível alcançar altas produtividades na lavoura, 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 desenvolver um novo método para estimar a distribuição das plantas na linha de plantio da lavoura de soja, por meio de visão computacional e processamento de imagens coletadas pelo RPA (sigla em inglês para “Aeronave Remotamente Pilotada”) utilizando *Machine Learning*. O estudo foi realizado na Fazenda Experimental Lageado, município de Botucatu, SP. O delineamento experimental utilizado foi em faixas, variando-se a velocidade de semeadura (4, 6 e 8 km·h⁻¹), com quatro repetições. A distribuição longitudinal da soja foi obtido em campo de forma convencional, inicialmente, para posterior análise do cálculo de distância entre conjuntos de plantas de soja com Inteligência Artificial (I.A.) com a utilização de RPA para a detecção das imagens. Os resultados obtidos das distâncias entre plantas com trena demonstraram que o aumento da velocidade de semeadura proporcionou menos espaçamentos normais e mais espaçamentos duplos e falhos, enquanto menores velocidades de semeadura favoreceram maior regularidade. Entretanto, nas velocidades de 6 e 8 km·h⁻¹ os resultados não diferiram estatisticamente. 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. Foi verificado que o modelo *Random Forest* não é a melhor tecnologia para a estimativa.

Palavras-chave: Agricultura de precisão. Plantabilidade. Aeronave remotamente pilotada. Machine learning. Análise automática.

ABSTRACT

In view of the importance of soy farming in an economic context, it is necessary to monitor agriculture with the use of techniques to maximize productivity and product quality. It is possible to achieve high yields in the field 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 develop a new method to estimate the distribution of plants in the soybean planting line, through computer vision and processing of images collected by RPA (Remotely Piloted Aircraft) using Machine Learning. The study was carried out at Fazenda Experimental Lageado, municipality of Botucatu, SP. The experimental design used was in bands, varying sowing speed (4, 6 and 8 km·h⁻¹) with four replications. The longitudinal distribution of soybeans was initially obtained in the field in a conventional way, with further analysis of the distance between sets of soybean plants with Artificial Intelligence (A.I.) with the use of RPA for image detection and A.I. application. The results obtained from the distances between plants with measuring tape showed that the increase in sowing speed provided less normal spacing and more double and flawed spacing while lower sowing speed improved regularity. However, for 6 and 8 km·h⁻¹ sowing speeds the results did not differ statistically. 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. It was verified that the Random model is not the best technology for the estimation.

Keywords: Precision agriculture. Plantability. Remotely Piloted Aircraft. Machine Learning. Automatic analysis.

LISTA DE ILUSTRAÇÕES

Figura 1 – Velocidades de semeadura mencionadas pelos entrevistados do sul do Matopiba, março de 2018.....	22
Figura 2 – Fluxograma das principais etapas usadas para descrever o processo de cálculo da distância do intervalo entre conjunto de plantas de soja.....	28
Figura 3 – Representação esquemática do delineamento experimental da área de estudo.....	30
Figura 4 – Coleta de campo com trena das informações de distribuição de plantas.....	31
Figura 5 – Tela do programa DroneDeploy na etapa de configuração do plano de voo.....	33
Figura 6 – Índices de espaçamentos aceitáveis médios de soja em função das velocidades ($\text{km}\cdot\text{h}^{-1}$).....	36
Figura 7 – Índices de espaçamentos duplos médios de soja em função das velocidades ($\text{km}\cdot\text{h}^{-1}$).....	37
Figura 8 – Índices de espaçamentos falhos médios de soja em função das velocidades ($\text{km}\cdot\text{h}^{-1}$).....	38
Figura 9 – Índices de coeficiente de variação (CV) médios de soja em função das velocidades ($\text{km}\cdot\text{h}^{-1}$).....	39
Figura 10 – Índices de produtividade médios de soja em função das velocidades ($\text{km}\cdot\text{h}^{-1}$).....	40
Figura 11 – Recorte do ortomosaico em blocos (1, 2, 3 e 4) do plantio da soja em velocidade de $4 \text{ km}\cdot\text{h}^{-1}$ com a aplicação do modelo <i>Random Forest</i>	41
Figura 12 – 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 <i>Random Forest</i>	42
Figura 13 – Recorte do ortomosaico em blocos (1, 2, 3 e 4) do plantio da soja em velocidade de $8 \text{ km}\cdot\text{h}^{-1}$ com a aplicação do modelo <i>Random Forest</i>	43

LISTA DE TABELAS

Tabela 1 – Índice de espaçamentos aceitáveis, duplos e falhos (%).....	32
Tabela 2 – Representa os resultados da validação cruzada para os modelos de <i>Random Forest</i> , <i>Logistic Regression</i> , <i>SVM</i> e <i>Dummy Classifier</i> , avaliados segundo suas métricas de Acurácia, Valor-F, Precisão e Revocação.....	40

SUMÁRIO

1	INTRODUÇÃO.....	17
2	REVISÃO DE LITERATURA.....	19
2.1	Importância da cultura da soja para o Brasil.....	19
2.2	Fatores que interferem na boa plantabilidade.....	19
2.2.1	Espaçamento entre plantas de soja.....	20
2.2.2	Velocidade de operação da semeadora.....	21
2.3	Avaliação da distribuição longitudinal da soja no campo.....	23
2.4	Técnicas para mensuração da distribuição longitudinal da soja de forma automática.....	23
2.4.1	Visão Computacional.....	24
2.4.2	Aprendizado de máquina.....	25
2.5	Trabalhos correlatos.....	26
3	MATERIAL E MÉTODOS.....	28
3.1	Área Experimental.....	29
3.2	Delineamento Experimental.....	29
3.3	Distribuição Longitudinal estimada com trena.....	30
3.4	Distribuição Longitudinal estimada com I.A.....	32
4	RESULTADOS E DISCUSSÃO.....	36
5	CONCLUSÕES.....	46
	REFERÊNCIAS.....	47

1 INTRODUÇÃO

Uma das culturas que representa a agricultura do Brasil e do mundo é a soja (*Glycine max* (L.) Merrill), por apresentar potencial produtivo, composição química e valores nutritivos que proporcionam uma diversidade de utilidades para a alimentação humana e animal. Os dados da CONAB (2021) demonstram que, na safra 2020/2021, a cultura continuará com o crescimento de área em relação às últimas safras, com uma estimativa de produção recorde de 133,7 milhões de toneladas.

O setor agroindustrial corresponde a uma parte significativa da economia do Brasil, além de ser um dos setores que mais investe em tecnologias e vem crescendo com a aplicação da Inteligência Artificial (I.A.) e a utilização de Aeronaves Remotamente Pilotadas (RPA, na sigla em inglês) como ferramentas para captura de dados.

A Accenture prevê que a I.A. vai gerar um impulso no VAB (Valor Agregado Bruto) de 432 bilhões de dólares até 2035 para a Economia Brasileira (ACCENTURE, 2017). De acordo com a empresa de pesquisa *Markets and Markets*, a I.A., no mercado agrícola, foi avaliada em 600 milhões de dólares em 2018 e deve atingir 2,6 bilhões de dólares em 2025, com taxa de crescimento anual de 22,5% durante o período de previsão (CLAVER, 2019).

Muitos são os fatores que podem afetar a qualidade e produtividade das culturas agrícolas; de acordo com Brachtvogel et al. (2009), entre eles estão os números de plantas e as distâncias entre as mesmas (CALONEGO et al., 2011). Por meio da semeadura com distribuição espacial adequada e uniforme das sementes, é possível alcançar altas produtividades na lavoura (FISS et al., 2018), uma vez que a distribuição uniforme da cultura nas linhas influencia a competitividade e eficácia do uso dos recursos do ambiente (TOURINO et al., 2009).

Os espaçamentos falhos na linha da cultura da soja podem acarretar o surgimento de plantas com tamanho reduzido e, conseqüentemente, é favorecido o crescimento de plantas daninhas diminuindo a produtividade. De maneira oposta, o agrupamento da cultura na linha também pode diminuir a produtividade por resultar em indivíduos menos ramificados, mais altos, diâmetros dos caules reduzidos e suscetíveis ao acamamento.

Perante este cenário, em virtude da importância agrícola da soja no contexto econômico, é necessário o acompanhamento da lavoura com o uso de técnicas para obter uma maximização da produtividade e qualidade do produto. Desta forma, faz-se importante o uso de tecnologias, sendo a Inteligência Artificial não apenas uma tecnologia, mas um conjunto de soluções que busca melhorar e otimizar processos, serviços e produtos.

O objetivo desta pesquisa foi desenvolver um novo método para estimar a distribuição das plantas na linha de plantio da lavoura de soja, por meio de visão computacional e processamento de imagens coletadas pelo RPA utilizando *Machine Learning*.

2 REVISÃO DE LITERATURA

2.1 Importância da cultura da soja para o Brasil

A soja é uma das principais culturas agrícolas do Brasil (CONAB, 2021). Segundo a Companhia Nacional de Abastecimento, em 2017, os produtos mais exportados no país foram a soja em grãos e seus derivados (óleo e farelo de soja), representando 14,10% da exportação brasileira (o equivalente US\$ 30,69 bilhões) sendo o Brasil considerado o maior exportador de soja mundial. Internamente, foram utilizadas em torno de 46,10 milhões de toneladas de soja em grãos, para a obtenção de 8,15 milhões de toneladas de óleo de soja, utilizados para o consumo humano e fabricação de biodiesel (CONAB, 2018).

Nas últimas três décadas, a soja obteve um dos maiores crescimentos entre as culturas, cultivada principalmente nas regiões Centro Oeste e Sul do Brasil, sendo considerada a principal *commodity* da agricultura brasileira (COLUSSI et al., 2016).

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 (CONAB, 2021).

2.2 Fatores que interferem na boa plantabilidade

A plantabilidade corresponde a uma distribuição uniforme de sementes ao longo do sulco, na densidade de plantas e população desejada e na profundidade correta. Alguns fatores podem interferir em uma boa plantabilidade, como por exemplo o tipo de preparo do solo, compactação, teor de água, qualidade e quantidade de insumos, regulagem da máquina, velocidade de semeadura, distribuição longitudinal da semente e a manutenção da máquina (EMBRAPA, 2018).

No momento da semeadura o rendimento da lavoura é estabelecido e a plantabilidade é essencial para o crescimento e desenvolvimento das plantas nas lavouras com condições adequadas. Os espaçamentos entre e dentre linhas são determinados pela arquitetura da planta, pelo potencial genético das culturas e pelas condições edafoclimáticas locais, e as lavouras dependem desse arranjo (EMBRAPA,

2000).

As recomendações de semeadura para a população de soja são em torno de 400 mil plantas por hectare, o equivalente a 40 plantas por m². Sem alterar significativamente o rendimento de grãos as variações podem ser para mais ou para menos entre 20 a 25% para a maioria dos casos, porém com plantas distribuídas uniformemente (EMBRAPA, 2000).

Segundo Rambo et al. (2003), algumas cultivares da soja apresentam capacidade de se adaptar às condições ambientais e de manejo (alta plasticidade), com modificações dos componentes do rendimento, na morfologia e arquitetura da planta.

Mediante essa característica da soja, Barni et al. (1985) e Gaudêncio et al. (1990) afirmam que se altera mais a morfologia da planta do que o rendimento de grãos. Porém, os resultados de muitos trabalhos não confirmam essa hipótese (RIGSBY; BOARD, 2003; REZENDE et al., 2004; HEITHOLT et al., 2005; LUDWIG et al., 2007).

Foram coletados dados de produtividade, variando entre 32 e 70 sacas ha⁻¹, para um projeto de pesquisa entre a Embrapa Milho e Sorgo e a Fazenda Trijunção. Visitaram-se 18 propriedades rurais da região do Oeste baiano, nos municípios de Cocos e Jabor. Com a amplitude deste resultado, conclui-se que é possível melhorar a produtividade com o manejo adequado para a cultura e o ambiente, trazendo redução de influência para a compra de novas áreas e aumento do rendimento em um mesmo local (EMBRAPA, 2018).

2.2.1 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 plantas invasoras, e, portanto, a distribuição longitudinal desuniforme da cultura pode provocar alterações na produtividade.

Pontos da lavoura com acúmulo de indivíduos ou espaçamentos duplos causam a ocorrência de plantas mais altas e um baixo número de ramificações, o que aumenta o risco de acamamento pela haste apresentar uma baixa resistência, plantas

dominadas e dificuldade no controle de doenças. A situação contrária, com espaçamentos falhos, pode trazer o aparecimento de plantas daninhas, afetar a produtividade e uma diminuição na altura das plantas (TOURINO et al., 2002).

Kurachi et al. (1989) dizem que as distâncias entre plantas precisam ser uniformes e seguir as recomendações agrônômicas ao número de plantas por hectare. Peske e Schuch (2008) afirmam que a soja é capaz de compensar variações no número de plantas por área.

2.2.2 Velocidade de operação da semeadora

Para uma deposição correta e uma distribuição uniforme de sementes, as semeadoras e os mecanismos dosadores de sementes são fundamentais (DIAS, 2012). Correia et al. (2014) afirmam que a operação de semeadura deve evitar falhas e duplas para uma distribuição longitudinal uniforme. Fatores podem interferir na homogeneidade entre plantas, como a velocidade de deslocamento da semeadora, velocidade periférica do disco dosador e a uniformidade entre as sementes e o disco dosador (KURACHI et al., 1989; MAHL et al., 2004).

O excesso de velocidade durante a semeadura da soja pode provocar a mudança da trajetória da semente dentro do condutor até o solo, devido ao “repique” na trajetória, podendo originar um atraso ou antecipação da queda da semente no sulco, ou ainda o rolamento das sementes nos sulcos (COPETTI, 2003).

De acordo com a Circular Técnica nº 51, a velocidade adequada para a semeadura está entre $4 \text{ km} \cdot \text{h}^{-1}$ e $6 \text{ km} \cdot \text{h}^{-1}$. A escolha da velocidade dentro dessa faixa depende da uniformidade da superfície da área de plantio (EMBRAPA, 2007).

Produtores da região do sul do Matopiba, região territorial que incorpora todo o Cerrado dentro dos limites do Maranhão, Tocantins, Piauí e Bahia, foram entrevistados sobre as operações mecânicas do plantio. Destes, 64,7% utilizam velocidades próximas ou acima de $6 \text{ km} \cdot \text{h}^{-1}$ (Figura 1), velocidades não recomendadas porque podem comprometer a produtividade. A população de plantas de soja adequada para cada condição climática e cultivar depende de um estande com baixa ocorrência de falhas e duplas (singulação), corretos espaçamentos e distribuição vertical (EMBRAPA, 2018).

Figura 1 – Velocidades de semeadura mencionadas pelos entrevistados do sul do Matopiba, março de 2018



Fonte: EMBRAPA (2018, p.16).

A velocidade do plantio tem influência direta sobre a uniformidade da cultura, a frequência de ocorrência de falhas ou sementes duplas, e assim interfere na produtividade da lavoura. Conforme se aumenta a velocidade de deslocamento da máquina, a regularidade de distribuição de plantas é influenciada de forma que a porcentagem de espaçamentos normais diminui (FURLANI, 2010). Esse parâmetro é um dos mais críticos, porém em muitos casos é negligenciado. Dados de pesquisa e observações em campo demonstram que semeaduras com velocidades acima de 8 $\text{km} \cdot \text{h}^{-1}$ comprometem a precisão dos equipamentos (MIALHE, 2012).

Mialhe (2012) constatou que o limite adequado de velocidade de semeadura relativo a mecanismos dosadores pneumáticos é recomendado em até 8 $\text{km} \cdot \text{h}^{-1}$, enquanto para mecanismos de funcionamento mecânico recomendam-se velocidades entre 4 e 6 $\text{km} \cdot \text{h}^{-1}$. Conforme há o aumento da velocidade da semeadora, a velocidade em que a semente chega no sulco de plantio é prejudicada, revolve mais o solo e consequentemente influencia na variação de profundidade da semente.

Em estudo realizado por Chaves (2015), com uma semeadora pneumática em diferentes sistemas de plantio e com velocidades que variaram entre 4 e 8 $\text{km} \cdot \text{h}^{-1}$, os resultados demonstraram a diminuição da produtividade com o aumento da velocidade sendo o melhor resultado referente à velocidade de 5,5 $\text{km} \cdot \text{h}^{-1}$, por apresentar a maior quantidade de vagens.

2.3 Avaliação da distribuição longitudinal da soja no campo

A avaliação da qualidade da semeadura pode ser realizada por meio da distribuição horizontal, que é a distância entre plantas na linha da semeadura, é possível avaliar a regularidade da distribuição longitudinal pelo coeficiente de variação (CV) (KURACHI et al., 1989). A distribuição de plantas pode ser classificada em três classes: dupla, quando a distância entre as plantas for menor que 0,5 vez a distância ideal; falha, quando o valor da distância entre plantas estiver 1,5 vezes maior que a distância ideal; e aceitável ou normal, quando a distância for maior que 0,5 e menor que 1,5 vezes a distância ideal (KURACHI et al., 1989).

A medição do CV é feita após o plantio, com abertura do sulco e exposição das sementes para as mensurações ou entre os estágios iniciais de desenvolvimento V2-V3. Para o cálculo do CV são necessários a média dos dados e o desvio padrão, e os dados são coletados pelas mensurações das distâncias entre plantas na linha de plantio em 5-10m, utilizando apenas parte da população. Deste ponto de vista, o cálculo está sujeito a críticas, visto que a amostra analisada da lavoura não reflete a variação total (KURACHI et al., 1989).

Para Kachman e Smith (1995) o valor máximo do CV é de 29%, e acima desse valor a precisão é considerada inadequada para avaliar o desempenho de semeadoras. Portanto, quanto menor o CV, maior a qualidade de semeadura. Segundo a literatura, é comum pesquisadores da área agônômica compararem os dados com os resultados descritos por Pimentel-Gomes (2009), onde o coeficiente é classificado como baixo quando inferior a 10%, médio se estimado entre 10% e 20%, alto quando entre 20 e 30% e muito alto se superior a 30%.

2.4 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). O popularmente conhecido “*drone*”, que significa zangão na língua inglesa, é controlado à distância por meios eletrônicos e computacionais, não demanda piloto embarcado e desempenha muitas funções (MEDEIROS et al., 2007).

O termo para melhor distinguir suas categorias e finalidades, utilizado pelo

Regulamento Brasileiro da Aviação Civil da ANAC (Agência Nacional de Aviação Civil), é Aeronave Remotamente Pilotada (RPA do inglês *Remotely Piloted Aircraft*). Além dessa ferramenta, a utilização de programas para o processamento das imagens geradas pelo RPA desde o planejamento do voo até o produto final do mapeamento é crucial para a execução do processo (ANAC, 2015).

Com as imagens aéreas obtidas pelo RPA é possível construir mosaicos para a realização de análises, uma ferramenta que auxiliará na agricultura de precisão. De acordo com Chon, Kim e Lin (2010), mosaicos representam a junção de uma quantidade de imagens sobrepostas entre si em uma nova imagem, e podem ser utilizados em áreas agrícolas para vistorias de campo antes, durante e após o ciclo de cultivo (LEBOEUF, 2000); também são utilizados mosaicos quando é necessária mais de uma imagem para cobrir toda a área (WOLF; DEWITT, 2000). Deste modo, mapas de recomendações podem ser elaborados para auxiliar nas providências da área, como no controle de falhas de plantio (LELONG et al., 2008).

2.4.1 Visão Computacional

Por meio da visão computacional, é possível processar e interpretar dados para extrair informações do processamento de imagens e tentar alcançar resultados semelhantes à visão humana. Uma imagem digital é interpretada com base nas características extraídas, diferente do ser humano que é capaz de identificar uma cena e perceber a estrutura tridimensional (SZELISKI, 2010; RUDEK et al., 2008).

Segundo Milano e Honorato (2010), a visão computacional é a ciência responsável pela extração de informações relevantes de imagens registradas por dispositivos imageadores, tais como a câmera digital, RPA, *scanner*, entre outros.

Portanto, visão computacional representa a maneira como um computador processa o meio à sua volta, e mediante o processamento de uma imagem é possível gerar uma outra, como o ortomosaico; com a utilização da visão computacional, informações são extraídas das imagens para diferentes finalidades (MILANO; HONORATO, 2010).

No início da década de 1970, o objetivo da visão computacional era imitar a inteligência humana objetivando a construção de robôs com comportamentos inteligentes. Na década de 1980, o propósito era voltado para técnicas matemáticas mais aprimoradas para a análise de imagens e cenários. Ocorreu um aprofundamento

da relação entre a visão do computador com a computação gráfica na década de 1990 e 2000 (DAVIES, 2018).

Nas últimas décadas, o objetivo tem sido a utilização de técnicas de aprendizado de máquina para problemas de visão de computador (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.

Um sistema de visão computacional tem como principais etapas: aquisição de imagens, pré-processamento, detecção e segmentação, extração de atributos ou características e processamento de alto nível (MILANO; HONORATO, 2010).

2.4.2 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, 2004).

O objetivo do aprendizado de máquina é desenvolver algoritmos capazes de fazer o computador aprender com a experiência, métodos que permitem aos computadores, além de aprender, melhorar o desempenho de aprendizagem com a experiência (CAPELA; GEORGIEVA, 2012). Segundo Annes (1999), a aprendizagem automática é utilizada para a prova de teoremas, processamento de linguagem natural e robótica, para solução de problemas e comprovação de proposições.

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 (HAYKIN, 2001; KOVÁCS, 2006; SILVA; SPATTI; FLAUZINO, 2010).

2.5 Trabalhos correlatos

Existem pesquisas direcionadas para automatizar tarefas no campo, como o estudo de Hague (2006) para 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 (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 (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, 2016).

O artigo de Zhang (2018) apresentou um novo método (um algoritmo de processamento) para estimar a distância entre plantas de milho, com o uso do RPA 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 RPA, sendo necessário no experimento o pós-processamento para a correção de erros (RUZA et al., 2017).

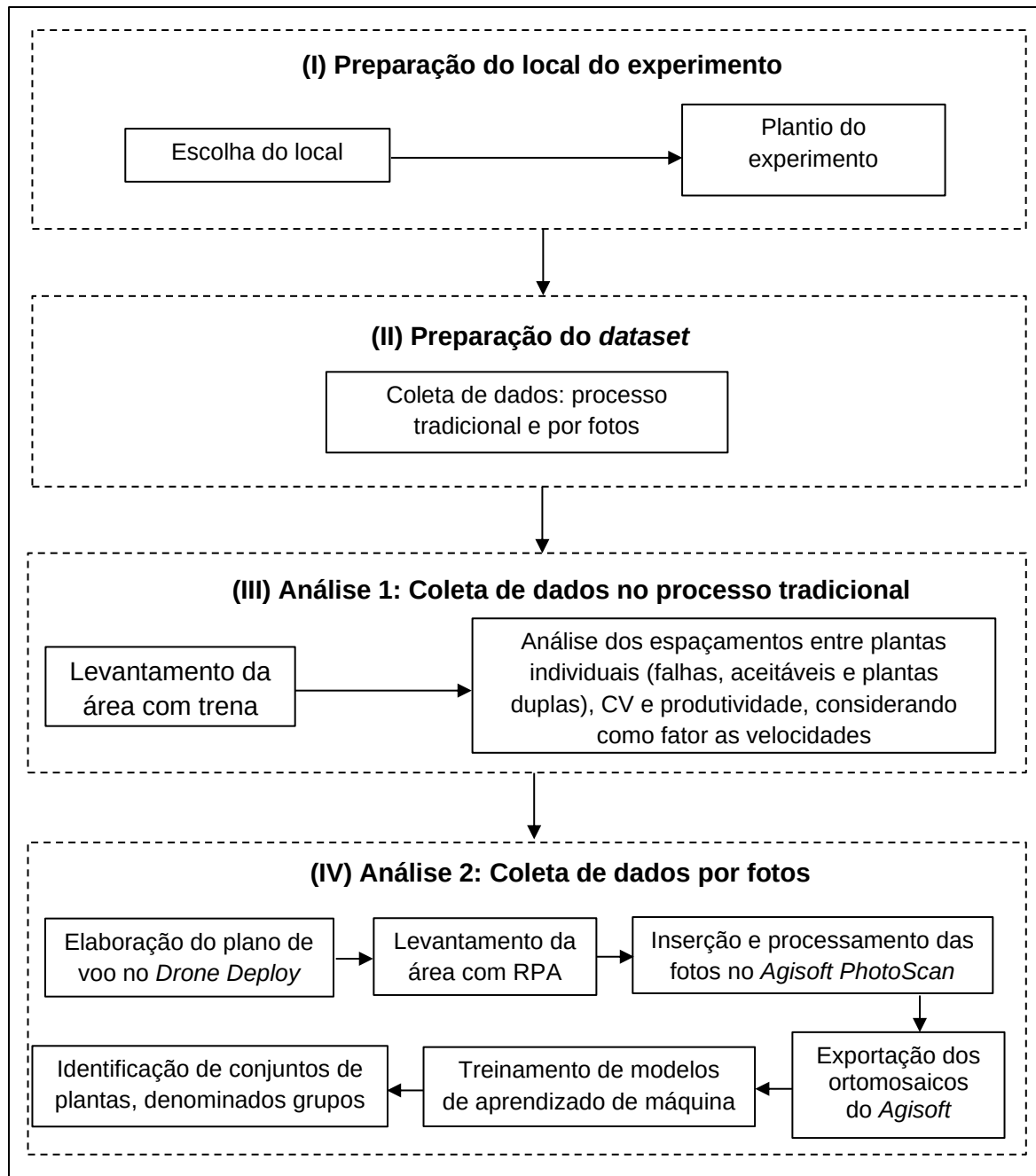
Hentz et.al. (2018) também utilizaram algoritmos para a detecção de árvores com RPA de forma automática, a ferramenta construída é chamada de TreeDetection.

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*. Porém, para estimar a distribuição das plantas na linha de plantio da lavoura de soja, é feito manualmente por técnicos de campo, que usam trenas. Sendo assim, este trabalho visa contribuir para a área visto que explora de forma diferente o tema por meio de visão computacional e processamento de imagens coletadas pelo RPA utilizando *Machine Learning*.

3 MATERIAL E MÉTODOS

A Figura 2 apresenta de forma resumida as etapas da metodologia utilizada para o desenvolvimento da dissertação.

Figura 2 – Fluxograma das principais etapas das metodologias utilizadas



3.1 Área Experimental

O experimento foi instalado na Fazenda Experimental Lageado, pertencente à Faculdade de Ciências Agronômicas da UNESP, situada no município de Botucatu - SP, em solo classificado pelo SBCS (2018) como Latossolo Vermelho – Amarelo distroférico. A área está localizada na latitude 22°48'59.83"S e longitude 48°25'43.00"O, com plantio realizado no dia 21 de novembro de 2019 e a colheita no dia 23 de março de 2020.

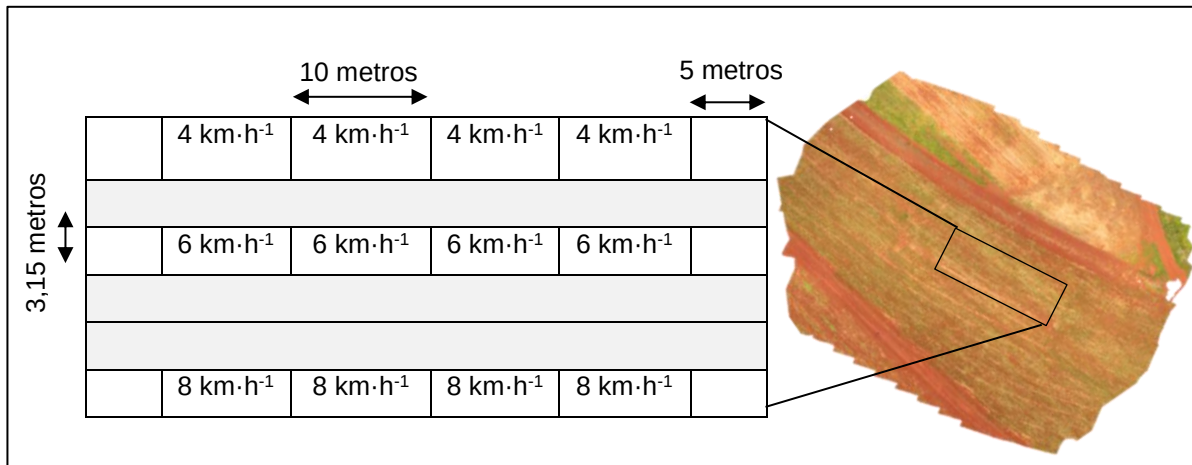
3.2 Delineamento Experimental

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⁻¹. As sementes foram tratadas com a adição de 4g de grafite por kg de sementes, após a secagem prévia do tratamento químico das sementes.

O delineamento experimental foi dividido em faixas de 50 metros cada, em sistema de plantio direto sobre palha de aveia preta, com semeadora-adubadora da marca Semeato, modelo PD17R, de 7 linhas com de 0,45 m de espaçamento, sulcador de fertilizante do tipo facão-guilhotina, dosadores de sementes mecânico de disco horizontal (modelo *Titanium* – J. ASSY Agrícola), discos com tecnologia *Rampflow* de 9mm com 90 furos e anel de anteparo com 1mm de rebaixo.

A semeadora foi tracionada com o trator de marca John Deere, modelo 6600 (4x2 TDA), com 121CV (89 Kw) de potência no motor, com velocidades de plantio de 4, 6 e 8 km·h⁻¹, executando 4 repetições para cada velocidade de plantio. As áreas de plantio equivaleram a blocos com dimensões 10 x 3,15 m, além de 5 metros no início e no final de cada faixa que foram descartados, totalizando 12 parcelas (Figura 3).

Figura 3 – Representação esquemática do delineamento experimental da área de estudo



A distribuição longitudinal da soja foi estimada em campo de forma convencional (com auxílio de trena). A identificação da soja e a distribuição da cultura na linha de plantio com aprendizado de máquina utilizou ortomosaicos gerados no *Agisoft PhotoScan* (Licença educacional), pelas imagens coletadas pelo RPA *Dji Phantom 3 PRO*, considerando posterior aplicação de algoritmos.

3.3 Distribuição Longitudinal estimada com trena

A avaliação manual da distribuição longitudinal de plantas da área experimental ocorreu especificamente no estágio vegetativo V4 da escala fenológica da cultura (terceira folha trifoliolada completamente desenvolvida), com o uso de uma trena graduada em milímetros. Em cada ponto amostral foi efetuada a mensuração da distância entre plantas presentes, em cinco linhas de semeadura (com a eliminação das linhas de bordadura) de três metros de comprimento, totalizando quinze metros em cada ponto amostral.

As distâncias foram mensuradas com uma trena estendida na linha de semeadura, de tal forma que fosse possível abranger 3 metros. Em seguida, procedeu-se à medição contínua entre plantas, anotando as informações em uma planilha (Figura 4).

Figura 4 – Coleta de campo com trena das informações de distribuição de plantas



Para o cálculo do CV (%), todos os espaçamentos entre plantas de soja coletados em campo foram somados, para em seguida determinar a média (cm) e o desvio padrão de cada tratamento. A Equação 1 mostra a fórmula utilizada para estimar os valores de CV:

$$CV = \left(\frac{DP}{\bar{x}} \right) \times 100 \quad (1)$$

Onde:

CV = Coeficiente de Variação (%).

DP= Desvio Padrão.

$\bar{x}$ = Média dos espaçamentos (cm).

Desta forma, foi possível determinar a porcentagem de espaçamentos entre plantas (EEP) normais (N), falhos (F) e duplos (D), obtida de acordo com Kurachi et al. (1993) (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. (1993).

Na análise da falha, aceitáveis, plantas duplas, CV e produtividade, considerando como fator as velocidades, foram ajustados modelos lineares generalizados com distribuição gama e função de ligação logarítmica (NELDER; WEDDERBURN, 1972). Em todos os ajustes, foi utilizado o procedimento *genmod* (do Programa estatístico SAS – *Free Statistical Statistical Software*, SAS University Edition), enquanto para comparações entre tratamentos foi utilizado o teste de Tukey-Kramer (WESTFALL, et al., 1999).

Para a avaliação da produtividade de soja foi contado o número de plantas por 10 metros em linha aleatória dentro do bloco com 4 repetições em cada velocidade de plantio, ao final do ciclo da cultura. Foram coletadas 10 plantas de cada linha amostral para calcular o número de vagens por planta, grãos por vagem e o peso de mil de grãos e estimou-se a produtividade de soja final em sacos por hectare (sc/ha).

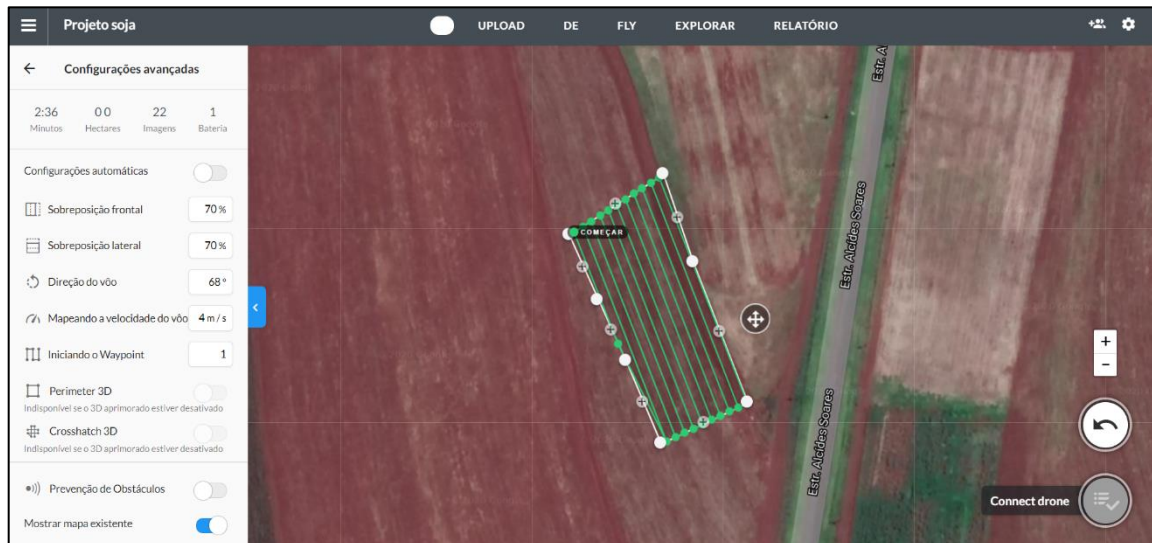
3.4 Distribuição Longitudinal estimada com I.A.

As imagens foram obtidas utilizando uma Aeronave Remotamente Pilotada (RPA) 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 RPA utilizado para a área foi o modelo *Dji Phantom 3 PRO*.

Para a obtenção das fotos da área experimental, foram elaboradas configurações para os planos de voo autônomos no programa *DroneDeploy*, onde é

possível definir a altura, direção, porcentagem de sobreposições laterais e frontais e resolução das fotos capturadas. Em seguida, o programa fornece a área mapeada, tempo de voo, número de fotos capturadas e número de baterias necessárias para a missão (Figura 5).

Figura 5 - Tela do programa DroneDeploy na etapa de configuração do plano de voo



Depois de configurar a missão em laboratório, a missão foi visualizada pelo programa DroneDeploy para a telemetria do voo em campo. O esquema de voo foi totalmente autônomo e realizado em conformidade com as exigências legais vigentes (Norma ICA 100-40 do DECEA), em altura de 10m e com sobreposição de 70% (DECEA, 2020).

As fotos do voo obtidas foram processadas no programa *Agisolt Photoscan* Profissional para obter os ortomosaicos da área do voo. As etapas utilizadas foram: *Add Photos*; *Align Photos*; *Build Dense Cloud*; *Build Mesh*; *Build DEM*; *Generate Contours* e *Build Ortomosaic*.

As imagens resultantes do processo de recorte dos ortomosaicos foram processadas a fim de identificar os conjuntos de plantas, denominadas de grupos, e posteriormente mensurar 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 (VAN 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 medida.

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 distância em pixels entre esses grupos (D). 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).

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 (KOHAVI et al., 1995) e esses foram avaliados segundo suas métricas de Acurácia, Precisão, Revocação e Média F (SHARMA et al., 2020).

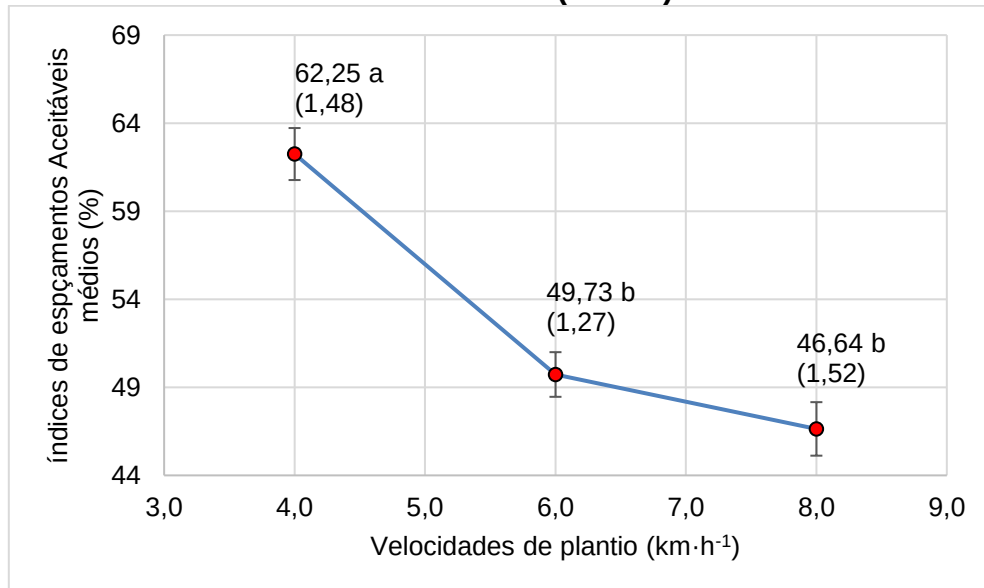
O modelo com melhores métricas de classificação foi implementado ao *script* principal de análise de imagens para definir se a medida de distância obtida era aceitável (amostra positiva) ou não aceitável (amostra negativa).

4 RESULTADOS E DISCUSSÃO

O cálculo da distância entre plantas feito manualmente no campo com o uso de trenas resultou na classificação de índices de espaçamentos médios aceitáveis, duplos e falhos; índices de CV e produtividade. Os resultados do teste de médias para o índice de espaçamentos aceitáveis são apresentados na Figura 6. Conforme o teste Tukey-Kramer (WESTFALL, et al., 1999), foram encontradas significâncias e interação para as variáveis índice de espaçamentos aceitáveis e velocidades de operação.

Os índices de espaçamentos aceitáveis médios estatisticamente foram maiores utilizando velocidade de semeadura de $4 \text{ km}\cdot\text{h}^{-1}$ (Figura 6) quando comparados às velocidades 6 e $8 \text{ km}\cdot\text{h}^{-1}$. Entre estes dois, não houve diferença significativa. Conforme o aumento da velocidade houve a diminuição dos espaçamentos considerados aceitáveis entre as velocidades $4 \text{ km}\cdot\text{h}^{-1}$ e $6 \text{ km}\cdot\text{h}^{-1}$, e consequentemente o aumento do surgimento de duplas e falhas (Figura 7 e 8).

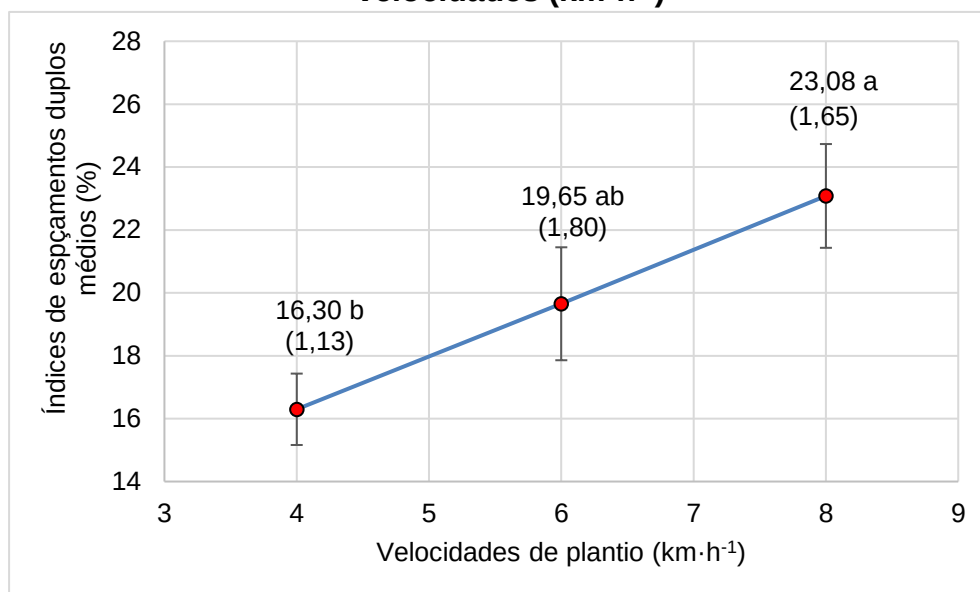
Figura 6 - Índices de espaçamentos aceitáveis médios de soja em função das velocidades ($\text{km}\cdot\text{h}^{-1}$)



Médias seguidas de mesma letra não diferem estatisticamente pelo teste de Tukey-Kramer ($p < 0,05$). Erro padrão da média entre parênteses.

Na velocidade de $8 \text{ km}\cdot\text{h}^{-1}$, verificou-se que os índices de espaçamentos duplos médios diferem de $4 \text{ km}\cdot\text{h}^{-1}$ a nível de 95% de confiança, porém não apresentou diferença significativa quando comparado com $6 \text{ km}\cdot\text{h}^{-1}$ (Figura 7).

Figura 7 - Índices de espaçamentos duplos médios de soja em função das velocidades ($\text{km}\cdot\text{h}^{-1}$)

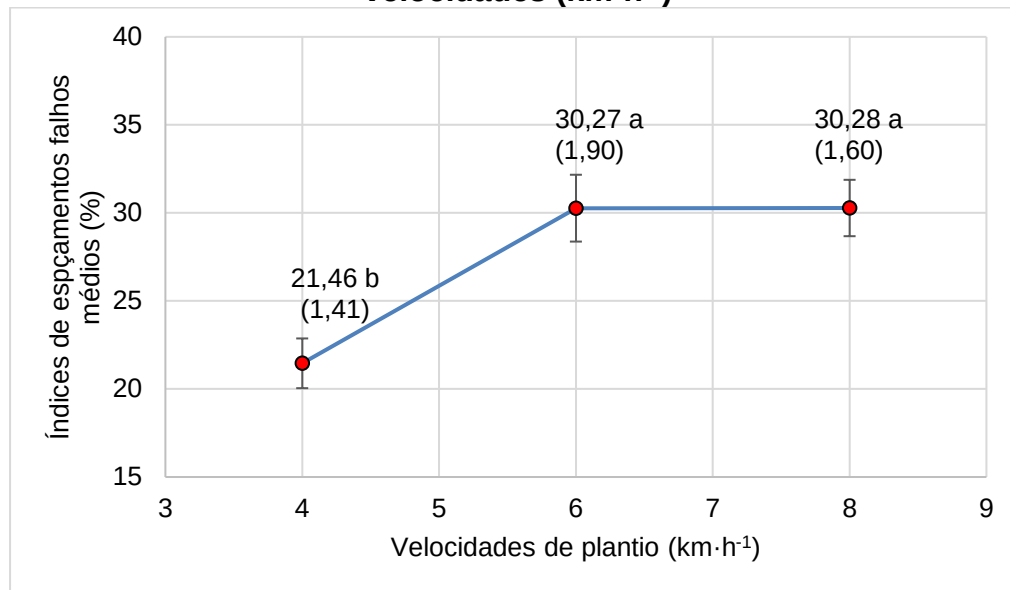


Médias seguidas de mesma letra não diferem estatisticamente pelo teste de Tukey-Kramer ($p < 0,05$). Erro padrão da média entre parênteses.

Os índices de espaçamentos falhos médios são apresentados na Figura 8. As velocidades de 6 e 8 $\text{km}\cdot\text{h}^{-1}$ possuem os maiores valores e não diferem estatisticamente. Na menor velocidade (4 $\text{km}\cdot\text{h}^{-1}$), verificaram-se os menores índices de espaçamentos falhos médios, com diferença estatisticamente significativa quando comparado com 6 e 8 $\text{km}\cdot\text{h}^{-1}$.

O aumento da velocidade de semeadura proporcionou menos espaçamentos normais e mais espaçamentos duplos e falhos, enquanto as menores velocidades de semeadura favoreceram a maior regularidade (CORREIA et al. 2020; ANGHINONI, 2019). Entretanto, nas velocidades de 6 e 8 $\text{km}\cdot\text{h}^{-1}$ os resultados não diferiram estatisticamente.

Figura 8 - Índices de espaçamentos falhos médios de soja em função das velocidades ($\text{km}\cdot\text{h}^{-1}$)

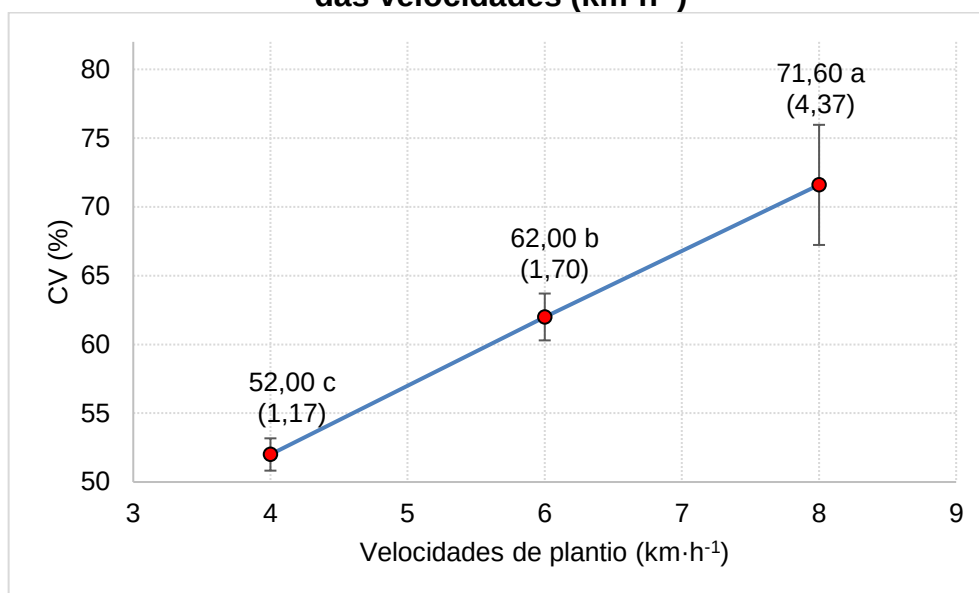


Médias seguidas de mesma letra não diferem estatisticamente pelo teste de Tukey-Kramer ($p < 0,05$). Erro padrão da média entre parênteses.

Os CVs obtidos variaram da menor para a maior velocidade de deslocamento, entre 52% e 71,6% sendo considerados muito altos segundo Pimentel-Gomes (2009), além de favorecer a desuniformidade da semeadura (Figura 9). Com o aumento da velocidade de semeadura, o CV aumentou.

Segundo Ceolin (2015), com o aumento da velocidade de plantio há diminuição da uniformidade dos espaçamentos entre plantas, portanto a velocidade pode influenciar o CV. Na Figura 9, é possível verificar que as menores velocidades de deslocamento (4 e 6 $\text{km}\cdot\text{h}^{-1}$) apresentaram menores coeficientes de variação, proporcionando a estas velocidades uma semeadura mais uniforme que as verificadas nas demais.

Figura 9 - Índices de coeficiente de variação (CV) médios de soja em função das velocidades ($\text{km}\cdot\text{h}^{-1}$)



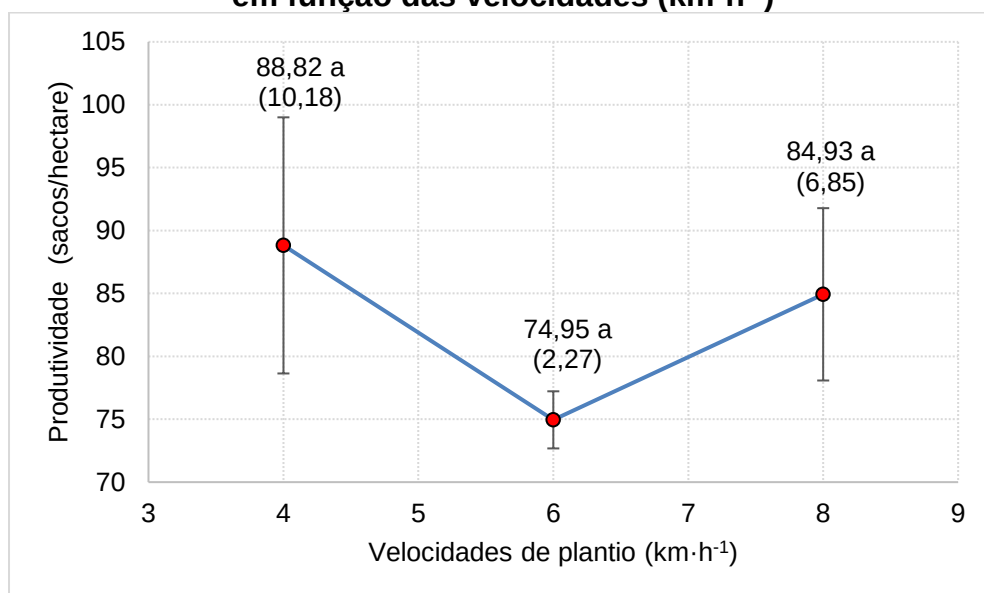
Médias seguidas de mesma letra não diferem estatisticamente pelo teste de Tukey-Kramer ($p < 0,05$). Erro padrão da média entre parênteses.

Os índices de produtividade médios de soja (Figura 10) não diferiram estatisticamente em função das velocidades ($4, 6$ e $8 \text{ km}\cdot\text{h}^{-1}$). Os valores de amostragem de produtividade coletados flutuaram muito nas velocidades de 4 e $8 \text{ km}\cdot\text{h}^{-1}$, impedindo que os índices diferissem entre si, porém não significa que não sejam diferentes. Provavelmente, deve haver outro fator que interferiu na produtividade do experimento.

Estudos demonstram que o número de plantas por metro é um dos pontos que podem afetar o rendimento final de grãos. Um número menor de população de soja perde mais água por evaporação e possibilita a invasão de plantas daninhas (MARTIN et al., 2005; COULTER et al., 2011; KOVÁCS, 2014; WOLI et al., 2014). Porém, na Figura 10, a produtividade do estudo estatisticamente não apresentou diferença, porque outros aspectos podem ter interferido no resultado.

Durante o plantio até a colheita, muitos fatores podem interferir no potencial produtivo da lavoura, como manchas de solo, declividade, teor de argila, profundidade do solo, disponibilidade de água (tanto o déficit hídrico quanto o excesso da mesma podem prejudicar a cultura), compactação do solo e luminosidade (BRACHTVOGEL et al., 2009).

Figura 10 - Índices de produtividade médios de soja em função das velocidades ($\text{km}\cdot\text{h}^{-1}$)



Médias seguidas de mesma letra não diferem estatisticamente pelo teste de Tukey-Kramer ($p < 0,05$). Erro padrão da média entre parênteses.

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

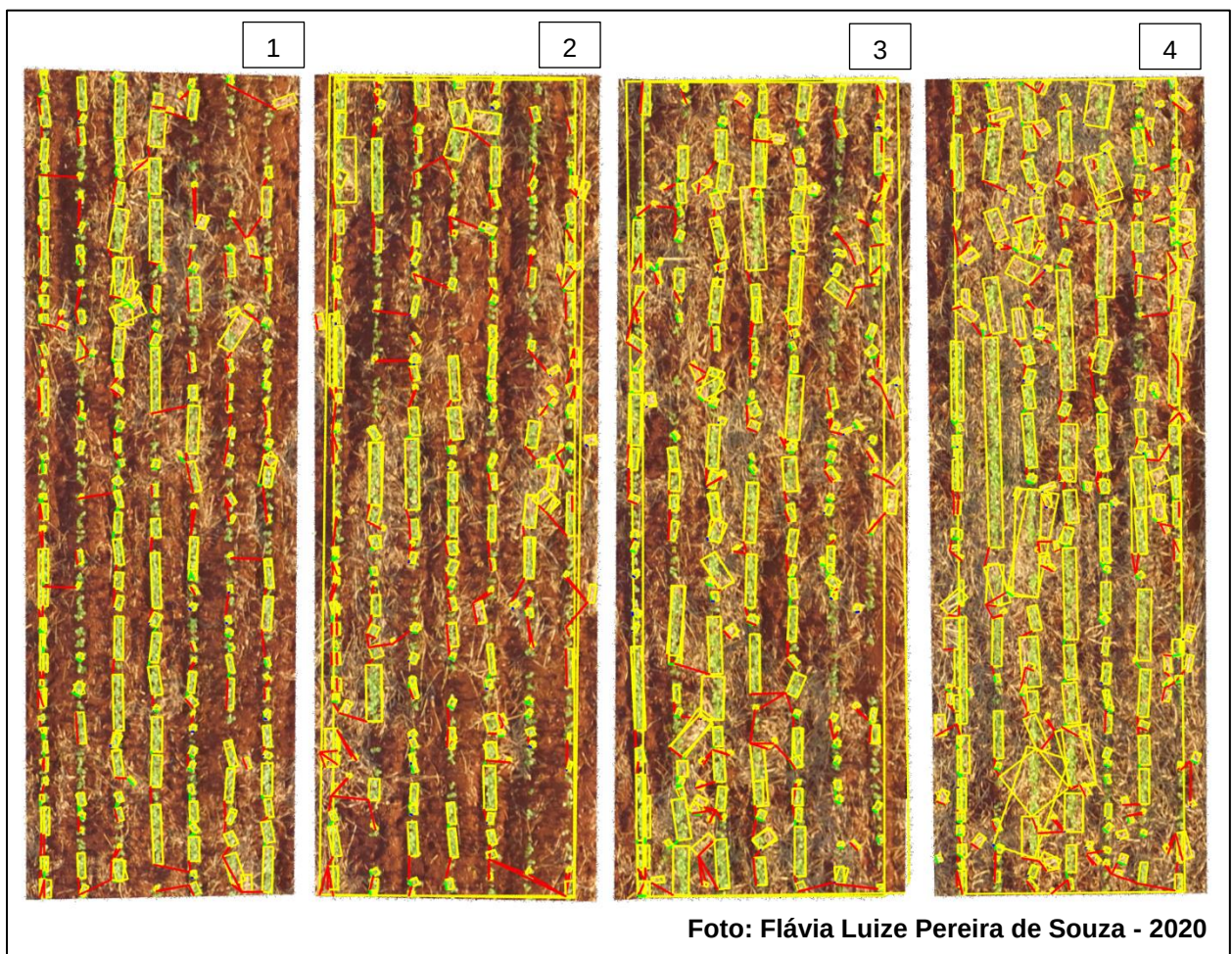
Modelo	Acurácia	Valor-F	Precisão	Revocação
<i>Random Forest</i>	0,655	0,655	0,655	0,655
<i>Logistic Regression</i>	0,581	0,581	0,582	0,581
SVM	0,539	0,538	0,550	0,539
<i>Dummy Classifier</i>	0,496	0,496	0,489	0,496

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 $4 \text{ km} \cdot \text{h}^{-1}$, 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 11 – Recorte do ortomosaico em blocos (1, 2, 3 e 4) do plantio da soja em velocidade de $4 \text{ km} \cdot \text{h}^{-1}$ 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.

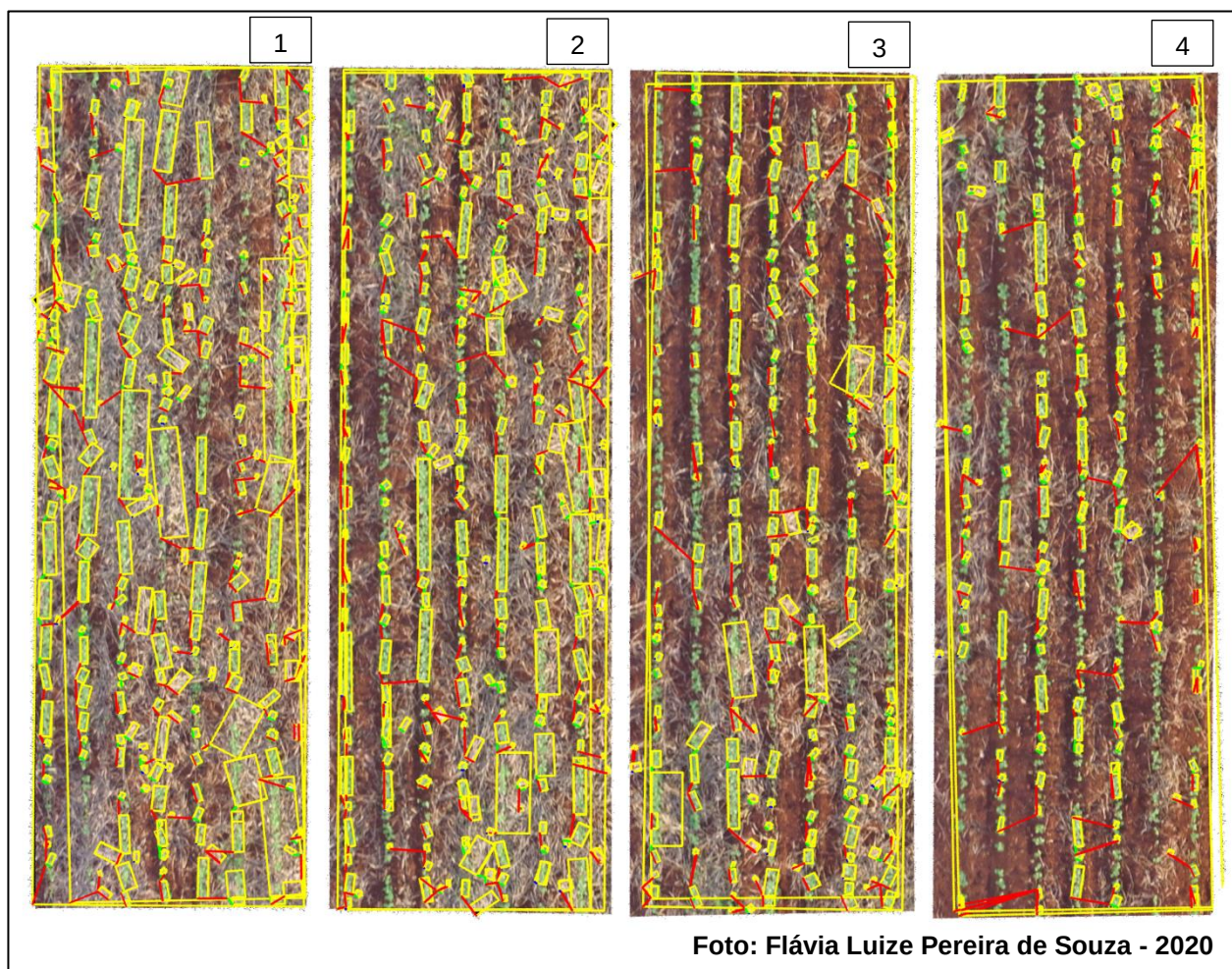
A distância entre os grupos permite a identificação de falhas, mas por se tratar de imagens tiradas por um RPA, com uma altura de 10 m, a qualidade diminui conforme aumenta a altura do voo para o caso da soja - uma cultura que é de

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 RPA 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 $6 \text{ km}\cdot\text{h}^{-1}$ 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.

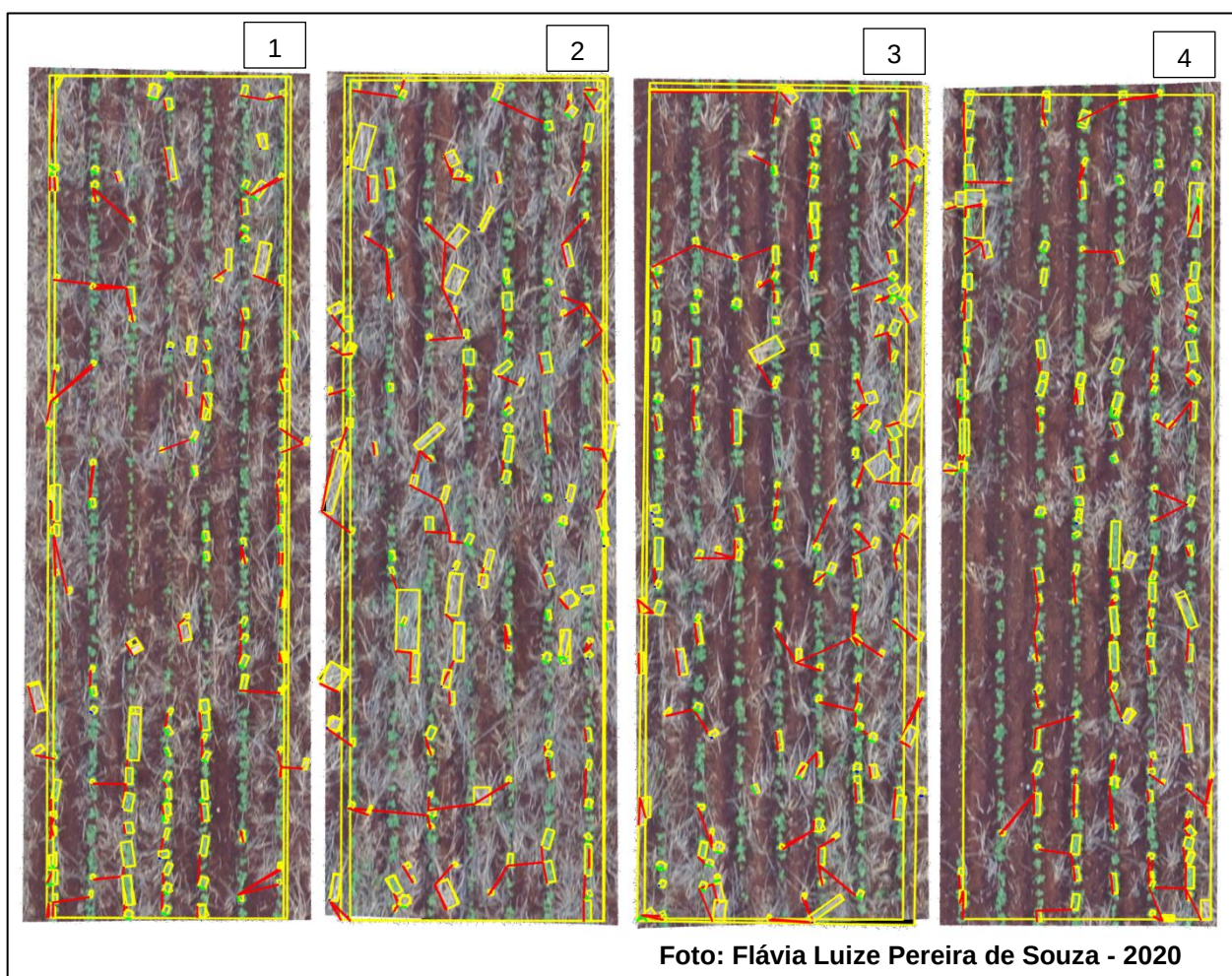
Figura 12 – Recorte do ortomosaico em blocos (1, 2, 3 e 4) do plantio da soja em velocidade de $6 \text{ km}\cdot\text{h}^{-1}$ com a aplicação do modelo *Random Forest*



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.

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 RPA estava passando pela área, provavelmente o sol estava atrás das nuvens, ocasionando uma qualidade de imagem diferente das outras velocidades.

Figura 13 – Recorte do ortomosaico em blocos (1, 2, 3 e 4) do plantio da soja em velocidade de $8 \text{ km}\cdot\text{h}^{-1}$ 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.

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 (2018) mostra que o cálculo das distâncias entre plantas de milho é possível por imagens de RPA e a precisão das distâncias está relacionada com a altura das imagens coletadas pelo RPA. 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 Zhang (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 RPA. A altura do voo do experimento da soja foi 10 m, maior que a utilizada por Zhang (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 Zhang (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 RPA, 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. No ano de 2021 a Syngenta Digital, em parceria com o grupo O Telhar, lançou uma nova ferramenta de Inteligência Artificial (plataforma *Cropwise*) que promete realizar a contagem de plantas em estágio inicial da lavoura de soja por imagem de celular e obter a estimativa de população de plantas final (SYNGENTADIGITAL, 2021).

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. (1993) (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 conseqüentemente 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 RPA. Foi verificado que o modelo *Random Forest* não é a melhor tecnologia para a estimativa. 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 fotos de celular em alturas menores que 10 m.

REFERÊNCIAS

- ACCENTURE. **Inteligência Artificial para acelerar o crescimento da América do Sul**. 2017. Disponível em: < <https://www.accenture.com/br-pt/insight-artificial-intelligence-south-america>>. Acesso em: 02 jun. 2020.
- 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.
- ANAC. **RPAS - Sistemas de aeronaves remotamente pilotadas**. 2015. Disponível em: < <https://www2.anac.gov.br/rpas/>>. Acesso em: 02 jun. 2020.
- ANGHINONI, M. **Mecanismos dosadores e velocidade de deslocamento do conjunto trator semeadora nos componentes agronômicos do milho**. Dissertação 44p. 2019.
- ANNES, R.. **Abordagem Multiagente para Processamento de Múltiplos Corpora: uma arquitetura multiagente de rotuladores de categorias gramaticais**. Dissertação (mestrado). Faculdade de Informática da Pontifícia Universidade Católica do Rio Grande do Sul, Faculdade de Informática, Porto Alegre, 1999. 79p.
- BARNI, N.A.; GOMES, J.E.S.; GONÇALVES, J.C. Efeito da época de semeadura, espaçamento e população de plantas sobre o desempenho da soja [*Glycine max* (L.) Merrill], em solo hidromórfico. **Agronomia Sulriograndense**, v.21, n.2, p. 245-296, 1985.
- BAUMANN, P. et al. **Quantifying the Uncertainty of Next-Place Psplirictions**. Proceedings of the The 8th EAI International Conference on Mobile Computing, Applications and Services. Anais... In: THE 8TH EAI INTERNATIONAL CONFERENCE ON MOBILE COMPUTING, APPLICATIONS AND SERVICES. Cambridge, Great Britain: ACM, 2016. Disponível em: <<http://eudl.eu/doi/10.4108/eai.30-11-2016.2267147>>. Acesso em: 24 nov. 2020.
- BRACHTVOGEL, E.L. et al. Densidades populacionais de milho em arranjos espaciais convencional e eqüidistante entre plantas. **Ciência Rural**, Santa Maria, v.39, n.8, p. 2334-2339, 2009.
- BRADSKI, G.. The OpenCV Library. **Dr Dobb's Journal of Software Tools**, 2000.
- BREIMAN, L. **Random Forests**. **Machine Learning**, v. 45, n. 1, p. 5-32, 1 out. 2001.
- CALONEGO, J.C. et al. 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.

CAPELA, N.F, GEORGIEVA P. **Aplicação de técnicas de aprendizagem automática para classificação de emoções humanas com sinais de EEG.** *Electrónica e Telecomunicações*, v. 5, n. 4, 2012.

CEOLIN, G. **Qualidade da semeadura da soja em função da velocidade e do sistema de distribuição.** Dissertação (Mestrado- Ciência e Tecnologia de sementes), Faculdade de Agronomia Eliseu Maciel, Universidade Federal de Pelotas. Pelotas-RS, 43p. 2015.

CHAVES, R. G. **Sistemas de Manejo do solo e velocidade de semeadura da soja.** Dourados/UFGD-MS, 46p. 2015.

CHON, J.; KIM, H.; LIN, C.-S. Seam-line determination for image mosaicking: A technique minimizing the maximum local mismatch and the global cost. *ISPRS Journal of Photogrammetry and Remote Sensing*, v. 65, n. 1, p. 86-92, 2010.

CLAVER, H. **Rapid adoption of artificial intelligence in agriculture.** 2019. Disponível em: https://www.futurefarming.com/Smart-farmers/Articles/2019/8/Rapid-adoption-of-artificial-intelligence-in-agriculture-461266E/?cmpid=NLC%7Cfuturefarm ing%7C2019-08-15%7CRapid_adoption_of_artificial_intelligence_in_agriculture. Acesso em: 02 jun. 2020.

COLUSSI, J. et al. O agronegócio da soja: Uma análise da rentabilidade do cultivo da soja no Brasil. **Revista Espacios** v. 37, n.16. 2016.

CONAB. Companhia Nacional de Abastecimento. **Acompanhamento da safra brasileira - grãos.** Brasília, v.7, Safra 2020/21, n.4, Quarto levantamento, p. 1-85, jan. 2021.

CONAB. Companhia Nacional de Abastecimento. **Perspectivas para agropecuária.** Brasília, v.6, Safra 2018/19, p. 1-112, ago. 2018.

COPETTI, E. Plantadoras: culpa de quem? **Cultivar Máquinas**, v. 3, n. 18, p. 14-17, 2003.

CORREIA, T. P. S.; LOPES, A. G. C.; FAGGION, F.; SILVA, P. R. A.; SOUSA, S. F. G. Semeadura de soja em função de mecanismos dosadores e velocidade operacional. **Energia na Agricultura**, Botucatu, v. 35, n. 2, p. 190-198, abril-junho, 2020.

COULTER, J.A. et al. Respostas agronômicas do milho à redução do estande em estádios de crescimento vegetativo. **Agronomia**. v.103, p. 577–583. 2011.

CRIST, J. **Dask amp; Numba: Simple libraries for optimizing scientific python code.** 2016 IEEE International Conference on Big Data (Big Data). **Anais...** In: 2016 IEEE INTERNATIONAL CONFERENCE ON BIG DATA (BIG DATA). dez. 2016

CORREIA, T.P.S. et al. Deposição e danos mecânicos em sementes de sorgo utilizando um mecanismo dosador de fluxo contínuo em ensaio de bancada. **Energia na agricultura**, Botucatu, v.29, n.1, p.22-26, 2014.

DAVIES, E.R. **Computer Vision. Principles, Algorithms, Applications, Learning**. 5. ed., p. 900, 2018.

DECEA. **Aeronaves não tripuladas e o acesso ao espaço aéreo brasileiro**. 2020. Disponível em: < <https://publicacoes.decea.mil.br/publicacao/ica-100-40>>. Acesso em: 02 dez. 2020.

DIAS, V.O. Tamanho amostral para ensaios em esteira de distribuição longitudinal de sementes. 110 f. **Tese** (Doutorado em Engenharia Agrícola) apresentado a Universidade Federal de Santa Maria, Santa Maria. 2012.

EMBRAPA. A Agropecuária do Sul do Matopiba em Perspectiva Circuito Solos Arenosos da Região Cocos-Jaborandi. **Embrapa Milho e Sorgo Sete Lagoas**, MG, 2018.

EMBRAPA. **Instalação da lavoura de soja: época, cultivares, espaçamento e população de plantas**. Circular Técnica 51, Londrina, 2007.

EMBRAPA. **Recomendações técnicas para a cultura da soja na região central do Brasil 2000/01**. Embrapa Soja, Londrina, 2000.

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

SBCS. **Sistema Brasileiro de Classificação de Solos**. Brasília, 5. ed., rev. e ampl.,Brasília: DF, Embrapa. p. 356, 2018.

FISS, G. et al. Produtividade e características agronômicas da soja em função de falhas na semeadura. **Revista Ciências Agrárias**, v. 61, 2018.

FURLANI, C. E. A. et al. Influência do manejo da cobertura vegetal e da velocidade de semeadura no estabelecimento da soja (Glycine max). **Engenharia na Agricultura**, Viçosa, v.18, n.3, p. 227-233, 2010.

GAUDÊNCIO, C.A.A. et al. População de plantas de soja no sistema de semeadura direta para o Centro-Sul do Estado do Paraná. Londrina: **Embrapa**, CNPSo, 4p. (Comunicado Técnico, 47). 1990.

HAGUE T, TILLET ND, WHEELER H. Automated crop and weed monitoring in widely spaced cereals. **Precision Agriculture**. v.7, p. 21–32. 2006.

HAYKIN, S. **Redes Neurais – Princípios e Práticas**. Bookman. 2. ed.Porto Alegre, 2001.

HEITHOLT JJ; FARR JB; EASON R. Planting Configuration x Cultivar Effects on Soybean Production in Low-Yield Environments. **Crop Science**, v. 45, p. 1800-1808. 2005.

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-PR, v. 48, n. 3, p. 393-402, jul-set, 2018.

JIANG G. et al. Wheat rows detection at the early growth stage based on Hough transform and vanishing point. **Comput Electron Agric.** v.123, p. 211–223.2016.

JI R.; QI, L. Crop-row detection algorithm based on Random Hough Transformation. **Math Comput Model.** v.54, p.1016–1020. 2011.

KACHMAN, S.D.; SMITH, J.A. Alternative measures of accuracy in plant spacing for planters using single seed metering. **Transactions of the ASAE**, St Joseph, v. 38, n. 2, p.379-387, 1995.

KOVÁCS, P.; VYN, T.J. Retrospectivas de estação completa sobre as causas da variabilidade planta a planta na resposta do rendimento de grãos de milho ao nitrogênio e ao preparo do solo. **Agronomia.** v. 106, p. 1746-1757. 2014.

KOVÁCS, Z. L. **Redes Neurais Artificiais - Fundamentos e Aplicações**. Editora Livraria da Física. 4ª edição. São Paulo, 2006.

KURACHI, S.A.H et al. 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.

LEBOEUF, J. Practical applications of remote sensing technology-an industry perspective. HortTechnology, **American Society for Horticultural Science**, v. 10, n. 3, p. 475-480, 2000.

LELONG, C. C. et al. Assessment of unmanned aerial vehicles imagery for quantitative monitoring of wheat crop in small plots. Sensors, **Molecular Diversity Preservation International**, v. 8, n. 5, p. 3557-3585, 2008.

LUDWIG M. P. et al. Efeito da densidade de semeadura e genótipos no rendimento de grãos e seus componentes na soja semeada após a época indicada. **Revista da FZVA**, Uruguaiana, v.14, n.2, p. 13-22. 2007.

MARTIN K.L. et al. Variabilidade planta a planta na produção de milho. **Agronomia.** v. 97, p. 1603-1611. 2005.

MAHL, D. et al. Demanda energética e eficiência da distribuição de sementes de milho sob variação de velocidade e condição de solo. **Engenharia Agrícola**, Jaboticabal, v. 24, n.1, p. 150-157, 2004.

MARCHANT JA, BRIVOT R. Real-time tracking of plant rows using a Hough transform. **Real-Time Imaging.** v.1, p. 363–371. 1995.

- MEDEIROS, F. A. et al. **Desenvolvimento de um veículo aéreo não tripulado para aplicação em agricultura de precisão**. Universidade Federal de Santa Maria, 2007.
- MIALHE, L. G. Máquinas agrícolas para plantio. Campinas: **Millennium**, 2012. 648 p.
- MILANO, D. D.; HONORATO, L. B. **Visão computacional**. – Universidade Estadual de Campinas– Faculdade de Tecnologia. 2010.
- 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.
- NELDER, J. A.; WEDDERBURN, R. W. M. Generalized linear models. **Journal of the Royal Statistical Society Series A**. v.135, n. 3, p. 370–384. 1972.
- PESKE, S. T.; SCHUCH, L. O. B. Falhas e duplos na produtividade. **Seed News**. Pelotas, v. 12, p. 22-27, 2008.
- PIMENTEL-GOMES, F. **Curso de estatística experimental**. 15. ed., Piracicaba: Fealq, 2009, 451 p.
- PLA F. et al. Building perspective models to guide a row crop navigation vehicle. **Image Vision Comput**. v.15, p. 465–473. 1997.
- RAMBO, L. et al. Rendimento de grãos da soja em função do arranjo de plantas. **Ciência Rural**, Santa Maria, v. 33, n. 3, p. 405-411, 2003.
- REZENDE P.M. et al. Efeito da semeadura a lanço e da população de plantas no rendimento de grãos e outras características da soja [*Glycine max* (L.) Merrill]. **Ciências Agrotécnicas**, vol.28, n.3, p. 499-504. 2004.
- RIGSBY B.; BOARD J.E. Identification of soybean cultivars that yield well at low plant populations. **Crop Science**, Madison v. 43, n. 1, p. 234-239, jan./feb. 2003.
- RUDEK, M.; COELHO, L. dos S.; CANCEGLIERI, O. **Visão Computacional Aplicada a Sistemas Produtivos: Fundamentos e Estudo de Caso**. 2008. Disponível em: <http://www.abepro.org.br/biblioteca/ENEGEP2001_TR10_0917.pdf>. Acesso em: 12 jun. 2019.
- RUSSEL, S.; NORVIG, P. **Inteligência Artificial**. 2. Ed. Rio de Janeiro: Campos, 2004.
- RUZA, M.R. et al. Inventário de Sobrevivência de povoamento de Eucalyptus com uso de Redes Neurais Artificiais em Fotografias obtidas por VANTs. **Advances in Forestry Science**, v.4, n.1. 2017.
- SHARMA, R. et al. **Exploring Geometric Feature Hyper-Space in Data to Learn Representations of Abstract Concepts**. MDPI, 2020.

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

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

SYNGENTADIGITAL. **Cropwise**, 2021. Disponível em: <<https://www.syngentadigital.com.br/#cropwise-lavoura>>. Acesso em: 3 jan. 2021.

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

TOURINO, M. C. C. et al. Semeadoras-adubadoras em semeadura convencional de soja. **Ciência Rural**, Santa Maria, v.39, n.1, p.241-245, 2009.

TOURINO, M. C. C.; REZENDE, P. M.; SALVADOR, N. Espaçamento, densidade e uniformidade de semeadura na produtividade e características agronômicas da soja. **Pesquisa Agropecuária Brasileira**, Brasília, v. 37, n. 8, p.1071-1077, ago. 2002.

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

XU L.; OJA, E.; KULTANEN, P. A new curve detection method: randomized Hough transform (RHT). **Pattern Recognit Lett**. v.11, p. 331–338. 1990.

WESTFALL, P. H. et al. **Multiple Comparisons and Multiple Tests Using the SAS System**, Cary, NC: SAS Institute Inc. 1999.

WOLF, P. R.; DEWITT, B. A. Elements of Photogrammetry: With Applications in GIS, 3ed. U.S.A.: **McGraw-Hill Book Company**. 2000.

WOLI, K.P. et al. Otimizando as taxas de semeadura de milho usando a classificação de aptidão de milho de um campo. **Agronomia**. v.106, p. 1523-1532. 2014.

ZHANG J. et al. **Estimating plant distance in maize using Unmanned Aerial Vehicle (UAV)**. PLoS ONE. v. 13. 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