

FLÁVIA LUIZE PEREIRA DE SOUZA

**CONTAGEM AUTOMÁTICA DE PLANTAS DE SOJA POR MEIO DE
INTELIGÊNCIA ARTIFICIAL E DADOS DE AERONAVE REMOTAMENTE
PILOTADA - RPA**

Botucatu

2024

FLÁVIA LUIZE PEREIRA DE SOUZA

**CONTAGEM AUTOMÁTICA DE PLANTAS DE SOJA POR MEIO DE
INTELIGÊNCIA ARTIFICIAL E DADOS DE AERONAVE REMOTAMENTE
PILOTADA - RPA**

Tese apresentada à Faculdade de Ciências Agrônômicas da Unesp Câmpus de Botucatu, para obtenção do título de Doutor em Engenharia Agrícola.

Orientador: Sérgio Campos

Coorientador: Luciano Shozo Shiratsuchi;
Mauricio Acconcia Dias e João Ricardo Favan

Botucatu

2024

S729c

Souza, Flávia Luíze Pereira de

Contagem automática de plantas de soja por meio de Inteligência Artificial e dados de Aeronave Remotamente Pilotada - RPA / Flávia Luíze Pereira de Souza. -- Botucatu, 2024

110 p. : il., tabs., fotos, mapas

Tese (doutorado) - Universidade Estadual Paulista (UNESP), Faculdade de Ciências Agrônômicas, Botucatu

Orientador: Sérgio Campos

Coorientador: Luciano Shozo Shiratsuchi; Mauricio Acconcia Dias e João Ricardo Favaretto

1. Agricultura de Precisão. 2. Inteligência Artificial. 3. Machine Learning. 4. Sensoriamento Remoto. 5. Contagem de plantas de soja. I. Título.

Sistema de geração automática de fichas catalográficas da Unesp. Biblioteca da Universidade Estadual Paulista (UNESP), 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:

CONTAGEM AUTOMÁTICA DE PLANTAS DE SOJA POR MEIO DE INTELIGÊNCIA ARTIFICIAL E DADOS DE AERONAVE REMOTAMENTE PILOTADA – RPA

AUTORA: FLAVIA LUIZE PEREIRA DE SOUZA

ORIENTADOR: SÉRGIO CAMPOS

COORIENTADOR: JOÃO RICARDO FAVAN

COORIENTADOR: MAURICIO ACCONCIA DIAS

COORIENTADOR: LUCIANO SHOZO SHIRATSUCHI

Aprovada como parte das exigências para obtenção do Título de Doutora em Engenharia Agrícola, pela Comissão Examinadora:

Prof. Dr. SÉRGIO CAMPOS (Participação Virtual)
Engenharia Rural e Socioeconomia / FCA UNESP Botucatu

Documento assinado digitalmente
SERGIO CAMPOS
Data: 08/03/2024 21:21:48-0300
Verifique em <https://validar.itl.gov.br>

Prof. Dr. LUCIANO NARDINI GOMES (Participação Virtual)
Geociências / Universidade Estadual de Londrina

Documento assinado digitalmente
LUCIANO NARDINI GOMES
Data: 08/03/2024 14:09:50-0300
Verifique em <https://validar.itl.gov.br>

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

Documento assinado digitalmente
OSVALDO CESAR PINHEIRO DE ALMEIDA
Data: 11/03/2024 13:56:36-0300
Verifique em <https://validar.itl.gov.br>

Dr. ANDERSON ANTONIO DA CONCEIÇÃO SARTORI (Participação Virtual)
. / .

Documento assinado digitalmente
ANDERSON ANTONIO DA CONCEIÇÃO SARTORI
Data: 10/03/2024 20:53:25-0300
Verifique em <https://validar.itl.gov.br>

Pesquisador Dr. BRUNO TIMOTEO RODRIGUES (Participação Virtual)
Geotecnologias / Parque Tecnológico de Itaipú

Documento assinado digitalmente
BRUNO TIMOTEO RODRIGUES
Data: 08/03/2024 18:16:32-0300
Verifique em <https://validar.itl.gov.br>

Botucatu, 23 de fevereiro de 2024.

AGRADECIMENTOS

Aos meus queridos pais e irmãos pelo apoio incondicional.

Ao meu orientador Prof. Dr. Sérgio Campos, que tenho grande consideração pelo seu trabalho. Por sua confiança e pela oportunidade de poder trabalhar com algo novo.

Ao Prof. Dr. Luciano Shozo Shiratsuchi, Profa. Dr. Haiying Tao, Prof. Dr. Maurício Acconcia Dias, Prof. Dr. Tri Deri Setiyono e novamente ao Prof. Dr. Sergio Campos pela orientação, ensinamentos e ajuda.

A todos os professores e funcionários da Faculdade de Ciências Agrônômicas (FCA) da Universidade Estadual Paulista "Júlio de Mesquita Filho" (UNESP), da "Louisiana State University" (LSU) e da "University of Connecticut" (UConn).

À FCA da UNESP, Botucatu, por sua generosidade ao fornecer terreno e gerenciamento de local.

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

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), pela bolsa de estudos concedida para estágio de doutoramento nos Estados Unidos pelo Programa Institucional de Internacionalização na modalidade de Doutorado Sanduíche no exterior (Nº do Processo: 88887.695839/2022-00).

À Universidade dos Estados Unidos Louisiana State University (LSU), na pessoa do Prof. Dr. Luciano Shozo Shiratsuchi, pela orientação e ensinamentos durante todo meu período de estágio de doutoramento no exterior.

À Universidade dos Estados Unidos University of Connecticut (UConn), na pessoa da Profa. Dr. Haiying Tao, pela orientação e ajuda no meu crescimento acadêmico e profissional, durante o meu período final de estágio de doutoramento no exterior.

Ao "Grupo de Plantio Direto" (GPD), e ao grupo de inteligência artificial intitulado "(Grupo de Pesquisas Avançadas em Inteligência Artificial no Setor Agroflorestal - LINEAR)", dentro do Departamento de Engenharia Rural e Socioeconomia da UNESP, Brasil, pelo suporte infraestrutural e conhecimento adquirido.

Ao Laboratório de Solo do Departamento de Ciência Vegetal e Arquitetura Paisagística da UConn, EUA, e ao Laboratório de Agricultura de Precisão da Escola de Ciências Vegetais, Ambientais e do Solo da LSU, EUA, por seu contínuo suporte infraestrutural e disposição em compartilhar conhecimentos.

Quero estender meus agradecimentos a todos que, de forma direta ou indireta, foram parte fundamental na realizaço e concluso deste trabalho. Agradeço sinceramente por todo o apoio e contribuiçes.

RESUMO

Soja, um componente vital da agricultura global, possui uma importância econômica significativa no Brasil e Estados Unidos, juntos produzem 69% da produção global de soja. As sementes representaram 26% dos custos operacionais associados à produção de soja nos EUA em 2022, e seus custos aumentaram em mais de 230% entre 1997 e 2022; no contexto brasileiro os custos operacionais associados à produção de soja também têm aumentado, principalmente devido ao aumento dos preços dos agrotóxicos e sementes usados em cultivos. Assim, a semente, que há dez anos representava menos de 10% do custo de aquisição de insumos, agora representa entre 10 e 20% no Brasil. A falta de sementes capazes de iniciar a germinação, emergir de forma adequada ou progredir para o desenvolvimento pode causar uma grande perda econômica. Além disso, o espaçamento inadequado nas linhas de cultivo de soja pode resultar em vários desafios que impactam negativamente a produtividade geral. Entretanto a evolução contínua do conhecimento e os avanços tecnológicos desempenham papel fundamental no impulsionamento da produtividade da soja. O mapeamento preciso da população de plantas por imagens é um desafio, especialmente quando não há informações sobre a qualidade das imagens e o melhor momento para capturá-las. Diante deste cenário o estudo foi realizado no Brasil na Fazenda Experimental Lageado, município de Botucatu, SP e nos Estados Unidos, na fazenda Ben Hur, município de Baton Rouge, LA. Neste estudo foi avaliado os impactos de diferentes algoritmos, sensores de captura de imagens, estágio da cultura, velocidade de semeadura e processos de análises de imagens e algoritmos de *machine learning* em relação a densidade de plantas de soja. Os resultados desta pesquisa destacam o potencial significativo das Redes Neurais na classificação de falhas de soja em imagens de Aeronaves Remotamente Pilotada (RPA). Ao analisar a contagem de plantas, o estágio de cotilédone (VC) e velocidade de 1,1 m/s (aproximadamente 4 km/h) apresentaram melhor acurácia, devido a sobreposição de plantas aumentar conforme o crescimento da planta, o que resulta em uma redução na precisão da estimativa. Em relação à velocidade de semeadura, foi observado que a precisão do modelo diminuiu conforme aumenta a velocidade de semeadura, pois a velocidade de deslocamento da semeadora pode prejudicar a homogeneidade entre as plantas e aumentar o número de sobreposições. As análises do número de plantas com o uso de recortes de

ortomosaicos obtidos por câmeras equipadas com sensores RGB (Red, Green, Blue) e a aplicação de Redes Neurais demonstraram alta precisão, em contraste, a análise utilizando processamento de imagens na linguagem R demonstrou alta precisão apenas em cenários de ausência de sobreposição de plantas, dessa forma, para solucionar essa limitação em estudos futuros, outras estratégias técnicas podem investigadas e avaliadas. Esses resultados servem como uma base sólida para o desenvolvimento futuro de algoritmos destinados à detecção de falhas na semeadura por meio da análise de imagens e fornece diretrizes para escolher estágio, sensores de captura de imagens e algoritmo para estimar a densidade de plantas por meio do processamento de imagens. Este avanço pode ter implicações significativas para a agricultura de precisão e o manejo de culturas.

Palavras-chave: falhas de semeadura; número de plantas; estágio fenológico da soja; machine learning; RPA.

ABSTRACT

Soybeans, a vital component of global agriculture, are of significant economic importance in Brazil and the United States, which together produce 69% of global soybean production. Seeds represented 26% of the operating costs associated with soybean production in the US in 2022, and their costs increased by more than 230% between 1997 and 2022; in the Brazilian context, the operating costs associated with soybean production have also increased, mainly due to the increase in the prices of agrochemicals and seeds used in crops. Thus, seed, which ten years ago represented less than 10% of the cost of purchasing inputs, now represents between 10 and 20% in Brazil. The lack of seeds capable of starting germination, emerging properly, or progressing to development can cause economic loss. In addition, inadequate spacing in soybean crop rows can result in challenges that negatively impact overall yield. However, the continuous evolution of knowledge and technological advances play a fundamental role in boosting soybean yield. Accurate mapping of plant populations using images is a challenge, especially when there is no information on the quality of the images and the best time to capture them. Given this scenario, the study was carried out in Brazil at the Lageado Experimental Farm in Botucatu, SP and in the United States at the Ben Hur Farm in Baton Rouge, LA. This study evaluated the impact of different algorithms, image capture sensors, crop stage, sowing speed, image analysis processes and machine learning algorithms in relation to soybean plant density. The results of this research highlight the significant potential of Neural Networks in classifying soybean flaws in Unmanned Aerial Vehicle (UAV) images. When analyzing the plant count, the cotyledon stage (CT) and speed of 1.1 m/s (approximately 4 km/h) showed better accuracy, due to the overlapping of plants increasing as the plant grows, which results in a reduction in the accuracy of the estimate. In relation to sowing speed, it was observed that the model's accuracy decreased as sowing speed increased, as the speed at which the seeder moves can damage the homogeneity between plants and increase the number of overlaps. The analysis of the number of plants using orthomosaic cutouts obtained by cameras equipped with RGB (Red, Green, Blue) sensors and the application of Neural Networks showed high accuracy, in contrast, the analysis using image processing in the R language showed high accuracy only in scenarios where there was no overlapping of plants, so to solve this limitation in future studies, other technical strategies can be

investigated and evaluated. These results serve as a solid basis for the future development of algorithms aimed at detecting sowing faults through image analysis and provide guidelines for choosing the stadium, image capture sensors and algorithm for estimating plant density through image processing. This breakthrough could have significant implications for precision agriculture and crop management.

Keywords: sowing gaps; number of plants; soybean phenological stage; machine learning; UAV.

LISTA DE ILUSTRAÇÕES

Capítulo 1 – Identificação de falhas na semeadura de soja utilizando *machine learning*

- Figura 1 - Visualização esquemática das localizações geográficas (A), seguida do layout do campo (B), coleta de dados (C), ortomosaicos criados a partir de imagens capturadas pelo Phantom 3 Pro (D) comparando o desempenho de três algoritmos: Árvore de Decisão, SVM e Rede Neural, na detecção de falhas na semeadura da soja em recortes dos ortomosaicos (E).....26
- Figura 2 - Ilustração da análise dos 200 recortes dos ortomosaicos retratando a semeadura de soja, incluindo 100 recortes com falhas na área de semeadura e 100 recortes sem falhas.....28
- Figura 3 - Curva ROC dos modelos de Rede Neural (azul), SVM (laranja) e Árvore de Decisão (verde) dos recortes com falhas de semeadura.....35
- Figura 4 - Curva ROC dos modelos de Rede Neural (azul), SVM (laranja) e Árvore de Decisão (verde) dos recortes sem falhas de semeadura.....35

Capítulo 2 – Contagem de plantas de soja com imagens RGB de RPA: uma abordagem eficaz durante mudanças fenológicas

- Figura 1 - Visualização esquemática da localização geográfica da área experimental, destacando as amostras selecionadas para a aplicação do modelo proposto. A obtenção de imagens RGB por meio de RPA.....53
- Figura 2 - Padrão do número de plantas de soja contadas manualmente no campo (laranja) e nos recortes dos ortomosaicos (azul) em diferentes velocidades de semeadura.....59
- Figura 3 - Índice de Vegetação.....61
- Figura 4 - Gráficos de dispersão comparando o número de plantas de campo com o número estimado a partir dos recortes dos ortomosaicos. A comparação considera o estágio de crescimento da planta (VC, V1 e V2), a contagem manual no campo e a contagem manual nos recortes63
- Figura 5 - Gráficos de dispersão comparando o número de plantas contadas manualmente no campo com o número estimado a partir dos ortomosaicos, com o uso do pacote R FIELDImageR na linguagem de programação R. A comparação considera o estágio de crescimento (VC, V1 e V2) e a velocidade de semeadura da soja (1,1 m/s, 1,6m/s e 2,2 m/s).....65

Figura 6 - Gráficos de dispersão comparando o número de plantas contadas manualmente nos recortes dos ortomosaicos com o número estimado dos mesmos recortes, com o uso do pacote R FIELDImageR na linguagem de programação R. A comparação considera o estágio de crescimento (VC, V1 e V2) e a velocidade de semeadura da soja (1,1 m/s, 1,6m/s e 2,2 m/s).....	66
Figura 7 - Exemplo de sobreposição de soja.....	70

Capítulo 3 – Contagem de plantas de soja baseado em imagens multissensor

Figura 1 - Delimitação das amostras em campo.....	84
Figura 2 - RPA DJI Matrice 300 utilizado no estudo.....	85
Figura 3 - Tipos de sensores utilizados.....	85
Figura 4 - Ortomosaico RGB (A); Ortomosaico multiespectral (B)).....	87
Figura 5 - Gráficos de dispersão comparando o número de plantas de campo com o número estimado com a linguagem R. A comparação considera o estágio de crescimento da planta (VC, V1.....	96
Figura 6 - Gráficos de dispersão comparando o número de plantas contados manualmente nas imagens com o número estimado com a linguagem R. A comparação considera o estágio de crescimento da planta (VC, V1 e V2), e o sensores (RGB e multiespectral).....	98

LISTA DE TABELAS

Capítulo 1 – Identificação de falhas na semeadura de soja utilizando *machine learning*

Tabela 1 – Resultados dos modelos de Rede Neural, SVM e Árvore de Decisão.....	32
Tabela 2 – Matriz de Confusão - Rede Neural, Árvore de Decisão e SVM.....	34
Tabela 3 – Resultados das Configurações do Modelo de Rede Neural.....	37
Tabela 4 – Matriz de Confusão da Rede Neural Configurada.....	40

Capítulo 2 – Contagem de plantas de soja com imagens RGB de RPA: uma abordagem eficaz durante mudanças fenológicas

Tabela 1 – Descrição dos estádios vegetativos da soja.....	55
Tabela 2 – Índices utilizados neste estudo. G - em inglês Green, R em inglês RED, B - em inglês Blue.....	57
Tabela 3 – Média do número de plantas contados manualmente em campo e nos recortes dos ortomosaicos nos estádios VC, V1 e V2 da soja.....	60
Tabela 4 – Erro Absoluto Médio (MAE) dos índices testados para remoção de ruído das imagens.....	61
Tabela 5 – Erro Absoluto Médio (MAE) entre o número de plantas estimado em comparação com a contagem manual no campo e nos recortes dos ortomosaicos para cada estádio vegetativo.....	64
Tabela 6 – O Erro Absoluto Médio (MAE) entre o número de plantas estimado nos recortes dos ortomosaicos em comparação com o número de plantas contados manualmente no campo e nos recortes dos ortomosaicos para cada estádio vegetativo.....	67

Capítulo 3 – Contagem de plantas de soja baseado em imagens multissensor

Tabela 1 – Especificações dos sensores e informações da aquisição de imagens RGB e multiespectral utilizados para este estudo.....	86
Tabela 2 – Índices utilizados neste estudo. G - em inglês Green, R - em inglês RED, B - em inglês Blue, NIR - em inglês Near Infrared.....	89

Tabela 3 – Resultado de classificação da Rede Neural.....	92
Tabela 4 – Matriz de Confusão da Rede Neural (RGB e multiespectral).....	93
Tabela 5 – Erros Médios Absolutos (MAE) dos índices testados para remoção de ruído das imagens (RGB e multiespectral).....	94
Tabela 6 – Média do número de plantas contados manualmente em campo e média do número de plantas estimado nos estádios VC, V1 e V2 e para cada sensor (RGB e multiespectral).....	97
Tabela 7 – Erros absolutos médios (MAE) entre o número de plantas estimado em comparação com o número de plantas no campo e nos recortes dos ortomosaicos para cada estágio vegetativo e para cada sensor (RGB e multiespectral).....	99

SUMÁRIO

INTRODUÇÃO GERAL.....	17
CAPÍTULO 1 - IDENTIFICAÇÃO DE FALHAS NA SEMEADURA DE SOJA UTILIZANDO MACHINE LEARNING	19
RESUMO.....	19
ABSTRACT.....	20
1.1 INTRODUÇÃO	21
1.2 MATERIAL E MÉTODOS	24
1.2.1 Modelos	24
1.2.2 Área experimental e coleta de dados	24
1.2.3 Análise com algoritmos	28
1.3 RESULTADOS.....	32
1.4 DISCUSSÕES.....	40
1.5 CONCLUSÕES.....	42
REFERÊNCIAS.....	43
CAPÍTULO 2 – CONTAGEM DE PLANTAS DE SOJA COM IMAGENS RGB DE RPA: UMA ABORDAGEM EFICAZ DURANTE MUDANÇAS FENOLÓGICAS.....	47
RESUMO.....	47
ABSTRACT.....	49
2.1 INTRODUÇÃO.....	50
2.2 MATERIAL E MÉTODOS.....	52
2.2.1 Área e delineamento experimental.....	52
2.2.2 Estádios de crescimento da soja.....	54
2.2.3 RPA e aquisição de dados.....	55
2.2.4 Processamento das imagens e contagem da população de plantas.....	56
2.3 RESULTADOS.....	58
2.3.1 Número de plantas contadas manualmente no campo e nos recortes dos ortomosaicos em diferentes estádios de crescimento da soja.....	58
2.3.2 Estimativa do número de plantas.....	60
2.4 DISCUSSÕES.....	68
2.5 CONSLUSÕES.....	71

REFERÊNCIAS	73
CAPÍTULO 3 – CONTAGEM DE PLANTAS DE SOJA BASEADO EM	77
IMAGENS MULTISSENSOR.....	
RESUMO.....	77
ABSTRACT.....	79
3.1 INTRODUÇÃO.....	80
3.2 MATERIAL E MÉTODOS.....	83
3.2.1 Área de estudo e projeto experimental.....	83
3.2.2 Aquisição das imagens.....	84
3.2.3 Processamento de dados.....	87
3.2.4 Estudo de validação.....	89
3.3 RESULTADOS.....	91
3.3.1 Classificação do número de plantas com Rede Neural comparando as imagens adquiridas pelos diferentes sensores.....	91
3.3.2 Índices para o processamento de imagens RGB e multiespectral em linguagem R.....	94
3.3.3 Estimativa do número de plantas com a técnica de processamento de imagens RGB e multiespectral.....	95
3.4 DISCUSSÕES.....	99
3.5 CONSLUSÕES.....	101
REFERÊNCIAS.....	102
CONSIDERAÇÕES FINAIS.....	107
REFERÊNCIAS.....	109

INTRODUÇÃO GERAL

A leguminosa *Glycine max* (Linnaeus) Merrill, reconhece-se como tendo sua origem estimada entre 6.000 e 9.000 anos atrás, situada primariamente na Ásia Oriental, onde passou por um processo de domesticação a partir de uma variante selvagem, Glicina soja Sieb. e Zucc. (Sedivy; Wu; Hanzawa, 2017). Em 1740, a soja foi transportada da China para a Europa e, em 1804, foi introduzida nos Estados Unidos da América (Sediyama, 2009). No contexto brasileiro, o primeiro registro remonta a 1882, marcando a realização de um cultivo experimental de soja. Entretanto, essa iniciativa revelou-se improdutivo devido à inadequação do material genético testado para as condições regionais. A bem-sucedida incorporação da soja no Brasil teve início nos anos 40, especialmente no estado do Rio Grande do Sul (RS) (Dall'agnol, 2016). No cenário atual, a soja desempenha um papel preponderante como componente na produção de ração animal e é também consumida diretamente pela população (USDA, 2023a).

A significativa parcela dos custos associados à produção de soja nos Estados Unidos em 2022 é direcionada aos seguintes insumos: sementes, representando 26% do total dos custos operacionais; fertilizantes, contribuindo com 24%; e químicos (fungicidas, herbicidas e inseticidas), responsáveis por 20% dos custos totais. O componente de sementes, como o insumo mais oneroso, registrou um incremento de mais de 230% no custo entre 1997 e 2022 (USDA, 2023b). Esse aumento é atribuído, em parte, à inclusão de características protegidas por patentes e direitos intelectuais (Turnipseed, 2019; Chen; Wiatrak, 2011), melhorias na qualidade das sementes e potencial genético (Lee *et al.*, 2008; Rincker *et al.*, 2014; Gaspar *et al.*, 2020), bem como às pressões decorrentes da alta inflação (Schnitkey *et al.*, 2022). Os custos operacionais no Brasil têm aumentado, de acordo com especialistas, principalmente devido ao aumento dos custos com sementes e agrotóxicos usados em cultivos. Assim, a semente, que há dez anos representava menos de 10% do custo de aquisição de insumos, conforme enfatizado pelos especialistas, agora representa entre 10 e 20% (Hirakuri, 2020).

A presença de sementes incapazes de iniciar a germinação, de emergir de maneira adequada ou, ao emergirem, não progredirem para o desenvolvimento de sementes, resulta em uma considerável perda econômica, como indicado por Board *et al.* (2013). As fontes de prejuízo entre a taxa de semeadura da soja e a população

final de plantas podem ser classificadas em três categorias principais: germinação, emergência e atrito. A germinação, como única forma de perda relativamente previsível, uma vez que a porcentagem estimada de germinação é indicada nas etiquetas das sementes, é abordada por Licht *et al.* (2020). O teste de germinação a quente, que avalia a porcentagem de plântulas normais desenvolvidas após exposição constante à umidade a 25 °C ao longo de 7 dias (Nafziger, 2022), é uma métrica frequentemente mencionada nessas etiquetas.

Por outro lado, a taxa de emergência é menos previsível, dadas as variáveis ambientais, como seca ou excesso de umidade, baixa temperatura do solo, profundidade inconsistente de semeadura, doenças de sementes/mudas ou alimentação por insetos, que podem resultar em perda significativa no estande de culturas jovens de soja (Board *et al.*, 2013; Licht *et al.*, 2020). Qualquer redução na produção de soja após o estabelecimento inicial do povoamento é considerada atrito (Setiyono *et al.*, 2010). Atualmente a principal fonte de atrito é o estresse hídrico: a falta de água durante períodos críticos de desenvolvimento da soja tem levado a uma redução na produção (Xu *et al.*, 2021).

Progressos tecnológicos na agricultura contemporânea têm despertado interesse em abordagens inovadoras para a avaliação de lavouras, incluindo o uso de imagens digitais obtidas por Aeronaves Remotamente Pilotadas (RPAs). Desta forma o objetivo do estudo foi quantificar e avaliar a densidade populacional de soja em ambientes agrícolas por meio da aplicação de algoritmos destinados à segmentação de campos de soja, nosso intuito é Nesse contexto, a presente tese está subdividida em três capítulos, delineando os seguintes objetivos de pesquisa: i) avaliar a eficácia de diferentes modelos e identificar a configuração mais adequada para detecção de falhas na semeadura de soja por meio de recortes dos ortomosaicos; ii) determinar o período fenológico ótimo, desde o estágio VC até o V2, para a análise da população de plantas de soja; iii) verificar a precisão de algoritmos utilizando imagens de câmeras com sensores Vermelho, Verde e Azul (RGB), multiespectral e térmico na contagem de plantas.

CAPÍTULO 1

IDENTIFICAÇÃO DE FALHAS NA SEMEADURA DE SOJA UTILIZANDO MACHINE LEARNING

Autora: Flávia Luize Pereira de Souza

RESUMO

A soja desempenha um papel crucial na agricultura brasileira e global, sendo uma cultura de alto valor econômico e nutricional. Avanços tecnológicos e conhecimento contínuo têm impulsionado sua produtividade e importância, especialmente considerando seu impacto nos setores econômico, comercial, nutricional e ambiental. Este estudo se concentra na detecção de falhas na semeadura de soja usando técnicas de aprendizado de máquina. Árvores de decisão, máquinas de vetor de suporte e redes neurais de perceptron multicamadas (MLP) são três modelos usados para analisar imagens de Aeronave Remotamente Pilotada (RPA) e encontrar falhas na plantação de soja. A pesquisa foi conduzida na Fazenda Experimental Lageado, localizada em Botucatu, São Paulo (SP), Brasil. Os resultados indicam que a Rede Neural MLP, após ajustes em suas configurações (função de ativação, a regularização, o número de neurônios nas camadas ocultas e o número de iterações), demonstrou ser o algoritmo mais promissor para caracterizar áreas de cultivo de soja e classificar com precisão as áreas de semeadura, identificando falhas. Esse desempenho reforça o potencial das Redes Neurais na resolução de desafios específicos, como a classificação de falhas de soja em imagens obtidas por Aeronaves Remotamente Pilotadas (RPAs). O estudo enfatiza a importância da tecnologia agrícola de precisão, impulsionada por avanços em aprendizado de máquina, para aumentar a eficiência, a sustentabilidade e a segurança alimentar global.

Palavras-chave: falhas de semeadura; RGB, Inteligência Artificial; sensoriamento remoto; RPA.

*Submetido na revista "Computers and Electronics in Agriculture"

ABSTRACT

Soybeans play a crucial role in Brazilian and global agriculture, being a crop of high economic and nutritional value. Technological advancements and continuous knowledge have driven its productivity and importance, especially considering its impact on economic, commercial, nutritional, and environmental sectors. This study focuses on detecting sowing gaps in soybeans using machine learning techniques. Decision trees, support vector machines, and multilayer perceptron (MLP) neural networks are three models used to analyze images from unmanned aerial vehicle (UAV) and identify gaps in soybean cultivation. The research was conducted at the Experimental Farm Lageado, located in Botucatu, São Paulo (SP), Brazil. The results indicate that the MLP Neural Network, after adjustments in its settings (activation function, regularization, number of neurons in hidden layers, and number of iterations), proved to be the most promising algorithm for characterizing soybean cultivation areas and accurately classifying planting areas, identifying gaps. This performance reinforces the potential of Neural Networks in addressing specific challenges, such as classifying soybean gaps in images obtained from UAVs. The study emphasizes the importance of precision agriculture technology, driven by advances in machine learning, to increase efficiency, sustainability, and global food security.

Keywords: seeding gaps; RGB, Artificial Intelligence; remote sensing; RPA.

1.1 INTRODUÇÃO

Glycine max (Linnaeus) Merrill, comumente conhecida como soja, ocupa posição de destaque na agricultura não apenas no Brasil, mas também em escala global. A soja, cultura vital no Brasil, tem alto significado econômico, conforme respaldado por dados da Companhia Nacional de Abastecimento (CONAB, 2023). A evolução contínua do conhecimento e os avanços tecnológicos desempenham papel fundamental no impulsionamento da produtividade da soja, resultando em maiores rendimentos na mesma área de cultivo e influenciando, conseqüentemente, os esforços de conservação ambiental. Seu potencial produtivo, composição química única e ricos valores nutricionais tornam-na altamente versátil em aplicações tanto para alimentação humana quanto animal (Liu, 1997). A soja tem um impacto econômico de alcance global, influenciando o comércio internacional, a agricultura, a nutrição, a energia renovável e os setores industriais. Sua versatilidade e importância tornam a soja uma cultura vital para a sustentabilidade econômica global. O crescimento do setor foi ainda mais acelerado pela integração da Inteligência Artificial (IA) e pela adoção de Aeronaves Remotamente Pilotadas (RPAs) como ferramentas essenciais para aquisição de dados (Habibi, 2021).

No contexto da produção de soja, manter uma alta densidade de plantas é reconhecido como uma abordagem primária para otimizar o rendimento da colheita (Ball *et al.*, 2000; Gan *et al.*, 2002). O conceito de plantabilidade envolve a distribuição uniforme de sementes ao longo do sulco para alcançar a densidade e a população desejadas na profundidade certa, o que é essencial para o cultivo de soja bem-sucedido. Preparo do solo, compactação, teor de água, quantidade e qualidade de insumos, manutenção da máquina, velocidade de semeadura, distribuição longitudinal de sementes e outros fatores podem afetar a plantabilidade (Dos Santos *et al.*, 2018).

O espaçamento inadequado nas linhas de cultivo de soja pode resultar em vários desafios que impactam negativamente a produtividade geral. Falhas nas linhas de cultivo podem levar ao surgimento de plantas de tamanho reduzido, proporcionando condições favoráveis para o crescimento de ervas daninhas e, por fim, causando uma redução no rendimento. Por outro lado, agrupar as plantas muito próximas pode restringir a ramificação, resultando em indivíduos mais altos com diâmetros de caule reduzidos e altamente suscetíveis a quedas (Fiss *et al.*, 2018). Endres (1996) destaca as conseqüências de desvios de espaçamento, enfatizando o

uso ineficiente de recursos essenciais, como água, luz e nutrientes, levando ao desenvolvimento irregular das plantas e à competição intensificada entre elas. As variações de espaçamento também podem contribuir para a proliferação de plantas invasoras, e a distribuição longitudinal irregular da cultura também pode influenciar ainda mais a produtividade final (Tourino, 2002).

O método manual de contagem de plantas de soja é uma abordagem tradicional usada para avaliar a distribuição e a densidade de plantas no campo. Esse método envolve a contagem física do número de plantas de soja dentro de uma área designada ou ao longo de uma linha de cultivo. Ele tem sido amplamente utilizado para avaliar a densidade de plantas e o estabelecimento em práticas agrícolas (Egli, 1988). Uma vantagem do método manual é sua simplicidade e natureza de baixo custo, exigindo equipamento e conhecimento mínimo e fornece observação direta, o que pode ajudar a identificar a saúde individual das plantas e padrões de distribuição específicos.

No entanto, o método manual tem várias desvantagens, como ser intensivo em mão de obra, demorado e propenso a erros humanos em extensões territoriais, devido à sua cobertura limitada. A precisão do método manual pode variar dependendo da habilidade e fadiga do observador. Consequentemente, o método manual pode não fornecer dados abrangentes e precisos para avaliações de culturas em larga escala, tornando-o menos eficiente nas abordagens modernas de agricultura de precisão. (Corassa *et al.*, 2018; Gaspar *et al.*, 2020).

A agricultura de precisão revolucionou o cenário agrícola, integrando os RPAs (Lelong *et al.*, 2008), visão computacional e algoritmos de aprendizado de máquina surgiu como uma abordagem poderosa para avaliar a distribuição de plantas de soja no campo. RPAs equipados com câmeras de alta resolução e sensores multiespectrais capturam imagens aéreas detalhadas, possibilitando o monitoramento abrangente da cultura (Geipel *et al.*, 2014). Técnicas de visão computacional processam essas imagens para identificar e segmentar precisamente as plantas de soja (Davies, 2017; De Milano; Honorato, 2014).

Posteriormente, algoritmos de aprendizado de máquina são aplicados para avaliar estratégias de manejo da cultura (Habibi *et al.*, 2021). Essa combinação de métodos de agricultura de precisão fornece aos agricultores informações valiosas

sobre a saúde da cultura, facilitando intervenções direcionadas para uma gestão otimizada de recursos e aumento da produtividade. Utilizando imagens de RPAs e técnicas de aprendizado de máquina, pesquisadores conseguiram contar com precisão plantas para diversas culturas, incluindo soja, arroz, trigo, milho, colza e algodão (Jiang *et al.*, 2019; Jin *et al.*, 2017; Randelović *et al.*, 2020; Wu *et al.*, 2019; Zan *et al.*, 2020; Zhang *et al.*, 2020). Além disso, a eficácia desses métodos avançados se estende à gestão de ervas daninhas, pois demonstraram capacidade em detectar plântulas de ervas daninhas em campos agrícolas (Peña *et al.*, 2015).

A aprendizagem de máquina tem sido aplicada com sucesso em diversos campos, incluindo a avaliação do manejo da soja. Com a crescente disponibilidade de dados e avanços na tecnologia de inteligência artificial, pesquisadores e agricultores têm aproveitado o potencial da aprendizagem de máquina para avaliar e otimizar o cultivo de soja. Técnicas de aprendizagem de máquina, como Redes Neurais, Árvore de Decisão e Máquinas de Vetores de Suporte (SVM), têm sido empregadas para abordar questões complexas relacionadas à produção de soja, como previsão de rendimento, identificação de padrões de doenças, análise de padrões de crescimento e detecção de falhas na cultura (Liakos, 2018; Rajput; Shukla; Thakur, 2020).

Com base nesse contexto, o uso de *machine learning* para avaliar a semeadura da soja pode resultar em mapeamento de falhas preciso. Baseando-se em pesquisas anteriores e reconhecendo a importância econômica do cultivo de soja, tanto para o Brasil quanto para o mercado global, este estudo teve como objetivo comparar o desempenho de algoritmos de Árvore de Decisão, SVM (Smola; Schölkopf, 2009), e Rede Neural Perceptron Multicamadas (MLP) (Haykin, 2009) usando o software Orange para desenvolver um algoritmo capaz de detectar com precisão falhas na semeadura de soja a partir de imagens e classificá-las adequadamente em categorias distintas com base na presença ou ausência de falhas (Demšar *et al.*, 2013).

1.2 MATERIAL E MÉTODOS

1.2.1 Modelos

Foram comparados três algoritmos: Árvore de Decisão, Máquinas de Vetores de Suporte (SVM) (Smola; Schölkopf, 2009), e Rede Neural Perceptron Multicamadas (MLP) (Haykin, 2009). A escolha desses algoritmos foi baseada na disponibilidade de implementação deles na ferramenta utilizada. O algoritmo de Árvore de Decisão é um método simples que divide os dados em nós com base na pureza da classe, utilizando ganho de informação para variáveis-alvo categóricas e erro médio quadrático (MSE) para variáveis numéricas. Ele serve como precursor do Random Forest. A Árvore internamente projetada no software Orange pode lidar com conjuntos de dados tanto categóricos quanto numéricos e pode ser usada para tarefas de classificação e regressão (Demšar *et al.*, 2013).

SVM é uma técnica de aprendizado de máquina que separa o espaço de atributos com um hiperplano, maximizando a margem entre instâncias de diferentes classes ou valores de classe. É conhecido por seu desempenho preditivo superior (Smola; Schölkopf, 2009). O software Orange incorpora uma implementação popular de SVM no pacote LIBSVM, que é acessível por meio de sua interface gráfica.

A Rede Neural utilizada neste estudo utiliza o algoritmo de perceptron de múltiplas camadas do sklearn, capaz de aprender modelos lineares e não lineares (Demšar *et al.*, 2013). A Rede Neural foi configurada com base no número de neurônios nas camadas ocultas, função de ativação, regularização e iterações. O Perceptron Multicamadas é uma Rede Neural que possui uma ou mais camadas ocultas e contém uma quantidade não especificada de neurônios. Como a saída desejada não pode ser prevista nessas camadas intermediárias, o termo "camada oculta" é usado.

1.2.2 Área experimental e coleta de dados

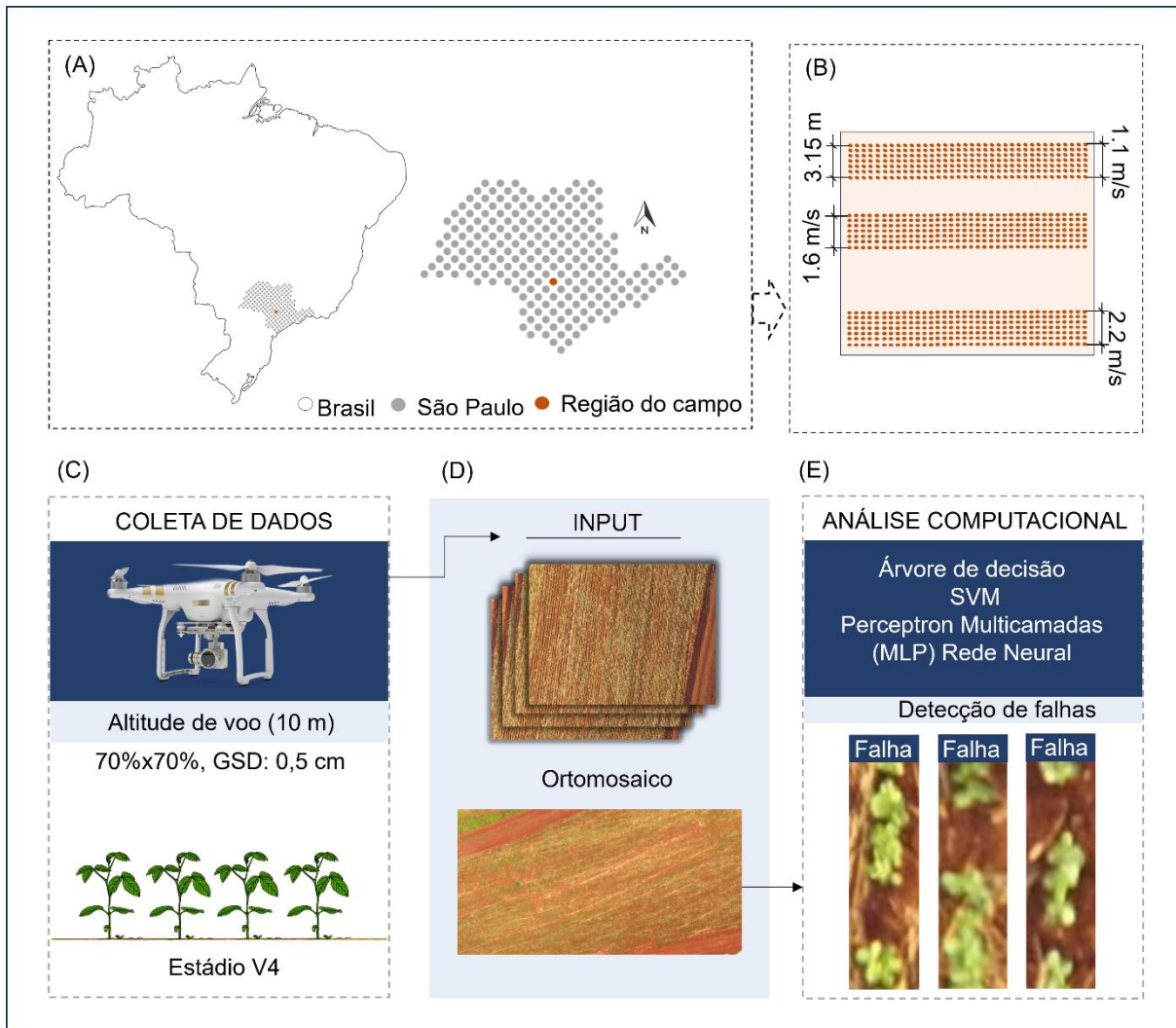
A Fazenda Experimental Lageado, pertencente à Faculdade de Ciências Agronômicas da UNESP, está localizada no município de Botucatu, São Paulo (SP), Brasil (Figura 1A). O solo nesta área é classificado de acordo com o sistema brasileiro

de classificação de solo como Latossolo Vermelho - Amarelo Distrófico (EMBRAPA, 2013). As coordenadas da fazenda são aproximadamente latitude 22°48'59,83"S e longitude 48°25'43,00"W. O Stand da cultura de soja (*Glycine max (Linnaeus) Merrill*) variedade TMG INTACTA RR IIPRO foi estabelecida com uma taxa de germinação de 80% e pureza de 99%, utilizando uma média de 16 sementes por metro linear, assumindo uma densidade populacional de 355.000 plantas por hectare. Antes de passar por tratamento químico, as sementes passaram por um processo de secagem e foram posteriormente tratadas com a adição de 4g de grafite por quilograma de sementes.

O delineamento experimental foi implementado em faixas de 50 metros para um sistema de plantio direto em uma superfície coberta com palha de aveia-preta. A semeadura foi realizada um semeador-fertilizador da marca Semeato, especificamente o modelo PD17R, que apresentava 7 linhas espaçadas a intervalos de 0,45 metros. O semeador era movido por um trator John Deere, modelo 6600 (4x2 TDA), com potência de motor de 121CV (89 Kw).

As velocidades de semeadura utilizadas foram 1,1 m/s (aproximadamente 4km/h), 1,6m/s (aproximadamente 6 km/h) e 2,2 m/s (aproximadamente 8 km/h), com 4 repetições para cada velocidade (Figura 1B). As áreas eram equivalentes a blocos medindo 10 x 3,15 metros, com mais 5 metros no início e no final de cada faixa, que foram descartados, totalizando 12 parcelas. O fertilizante foi aplicado usando um sulcador tipo guilhotina, e a dosagem de sementes foi realizada por dosadores de disco horizontal mecânicos (modelo Titanium - J. ASSY Agrícola). Esses discos estavam equipados com tecnologia Rampflow, apresentando 90 furos com um diâmetro de 9 mm e um anel de anteparo com um recesso de 1 mm. A semeadura da soja foi realizada em 21 de novembro de 2019.

Figura 1 - Visualização esquemática das localizações geográficas (A), seguida do layout do campo (B), coleta de dados (C), ortomosaicos criados a partir de imagens capturadas pelo Phantom 3 Pro (D) comparando o desempenho de três algoritmos: Árvore de Decisão, SVM e Rede Neural, na detecção de falhas na semeadura da soja em recortes de ortomosaicos (E)



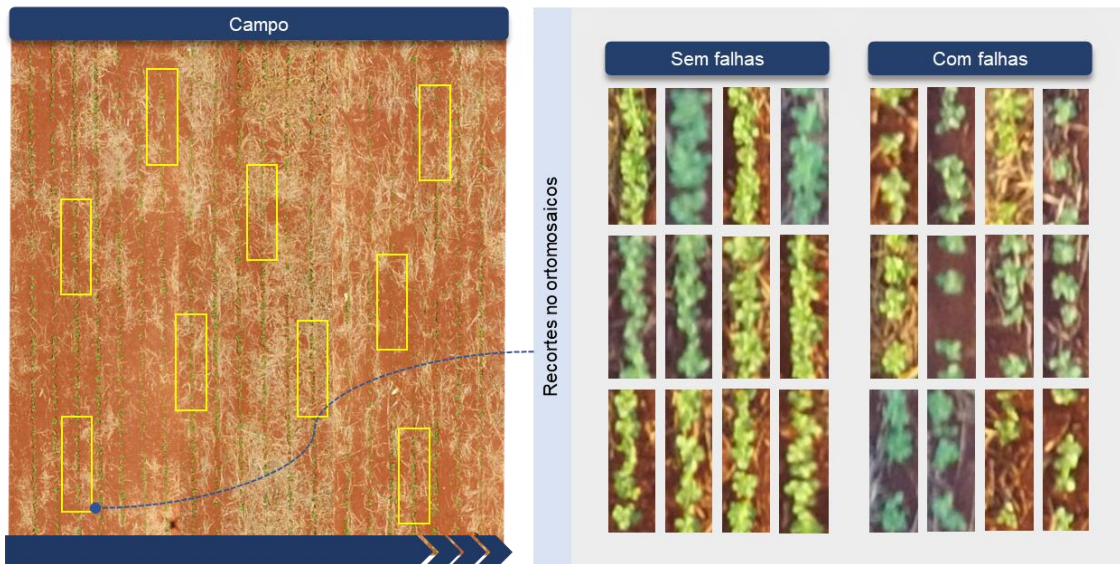
As imagens foram capturadas usando um RPA de modelo específico Phantom 3 PRO, obtidas durante um projeto de mestrado em 10 de dezembro de 2019 (Souza *et al.*, 2022) (Figura 1C). Essas imagens foram processadas para gerar ortomosaicos da área de voo (Figura D) e posteriormente o modelo de inteligência artificial foi desenvolvido para identificar falhas na semeadura da soja nos recortes dos ortomosaicos, classificando áreas com e sem falhas de semeadura (Figura E).

O Phantom 3 Pro foi utilizado, equipado com sua câmera gimbal integrada de fábrica com capacidade de 12 megapixels, com resolução de imagem de 4000x3000 pixels com sensor Vermelho, Verde e Azul (RGB). Essas imagens aéreas foram capturadas usando a Aeronave Remotamente Pilotada (RPA) durante o estágio de crescimento vegetativo das plantas, precisamente no estágio V4 (Terceira folha trifoliolada completamente desenvolvida), tempo suficiente para a emergência e o crescimento inicial, mas não em uma fase avançada que complique a identificação das falhas (Randelović et al., 2020). Configurações foram criadas para os planos de voo autônomos do programa DroneDeploy para que pudessem ser tiradas imagens da área experimental. O aerolevanteamento foi realizado de forma totalmente autônoma a uma altitude de 10m, assegurando assim garantindo assim um GSD (*Ground Sample Distance*) de 0,5 cm, e um total de 343 imagens foram coletadas. Para a variável Sobreposição de Imagens, que engloba as configurações de sobreposição Frontal (*Overlap*) e sobreposição lateral (*Sidelap*) para as imagens coletadas em voo, foi adotado um planejamento e execução de voo para captura de dados fotogramétricos, esse segmento foi 70%x70%.

As imagens do voo obtidas foram processadas com o auxílio do Agisoft Metashape 2.1.1 para obter os ortomosaicos da área do voo. As etapas que foram empregadas foram: adicionar imagens (*Add Photos*); alinhar imagens (*Align Photos*); construir uma nuvem densa (*Build Dense Cloud*); construir uma rede (*Build Mesh*); construir um Modelo de Elevação Digital (*Build DEM*); criar contornos (*Generate Contours*) e construir um ortomosaico (*Build Ortomosaic*).

Posteriormente, recortes do ortomosaico foram gerados para todas as três velocidades de semeadura de soja, resultando em um total de 230 recortes, para o treino do modelo foram usados 100 recortes rotulados (*label*) com falhas de semeadura dentro da área de soja e outros 100 recortes rotulados como semeadura sem falhas, e para o teste outros 30 recortes com falhas e sem falhas na semeadura, utilizados sem informação adicional (sem *label*) (Figura 2). Cada recorte de 40x130 pixels, sendo 3 camadas RGB (15.600 *inputs*) e o *output* (com e sem falha).

Figura 2 - Ilustração da análise dos 200 recortes dos ortomosaicos retratando a semeadura de soja, incluindo 100 recortes com falhas na área de semeadura e 100 recortes sem falhas



1.2.3 Análise com algoritmos

As análises dos algoritmos foram realizadas nos laboratórios de Sensoriamento Remoto e Geoprocessamento na UNESP, Brasil; no laboratório de Agricultura de Precisão da Escola de Ciências Vegetais, Ambientais e do Solo da LSU, EUA, e no laboratório de Solo do Departamento de Ciência Vegetal e Arquitetura Paisagística da UConn, EUA.

Após avaliar os resultados, a Rede Neural foi explorada mais a fundo devido ao seu desempenho superior. Para melhorar o desempenho da Rede Neural, foi realizado um processo de configuração. Inicialmente, valores aleatórios foram atribuídos aos pesos e limites. O comportamento da rede é afetado pela escolha dos valores iniciais, portanto, na ausência de familiaridade com os pesos e limites, seus valores iniciais devem ser aleatórios e distribuídos uniformemente, o que é essencial para uma inicialização adequada. As principais métricas de avaliação usadas para avaliar o desempenho do modelo incluem AUC (Área Sob a Curva) na Equação 1, Precisão na Equação 2, Acurácia (CA) na Equação 3, *Recall* conforme a Equação 4 e Pontuação F1 na Equação 5 (Demšar *et al.*, 2013).

$$AUC \approx \sum_{i=1}^{n-1} \frac{(y_i + y_{i+1} + 1)}{2} \times (x_{i+1} - x_i) \quad (1)$$

$$\text{Precisão} = \frac{VP}{VP + FP} \quad (2)$$

$$\text{Acurácia} = \frac{VP + VN}{VP + VN + FP + FN} \quad (3)$$

$$\text{Recall} = \frac{VP}{VP + FN} \quad (4)$$

$$F1 = 2 \times \frac{\text{Precisão} \times \text{Recall}}{\text{Precisão} + \text{Recall}} \quad (5)$$

Onde:

x_i = Taxa de falsos positivos;

y_i = Taxa de verdadeiros positivos;

VP = Verdadeiros positivos;

FP = Falsos positivos;

FN = Falsos negativos.

Além disso, utilizamos para avaliar o desempenho dos modelos as curvas de características de operação do receptor, um gráfico conhecido como curva de ROC (em inglês *Receiver Operating Characteristic*). As curvas são feitas ao calcular a taxa de verdadeiros positivos (TVP) conforme Equação 6 e a taxa de falsos positivos (TFP) para cada limite de classificação possível, de acordo com equação 7 a seguir (TING, 2011):

$$TPR = \frac{VP}{VP + FN} \quad (6)$$

$$FPR = \frac{FP}{FP + VN} \quad (7)$$

Onde:

VP = Verdadeiros positivos;

FP = Falsos positivos;

FN = Falsos negativos;

VN = Verdadeiros negativos.

A realização da classificação dos recortes por meio de Árvore de Decisão, SVM e Redes Neurais envolveu a integração eficiente de *widgets* do software Orange, tais como "*import images*", "*image views*" e "*image embedding*". Este processo resultou na criação de uma tabela de dados contendo descritores dos recortes. Adicionalmente, a ferramenta "*image embedding*" desempenhou um papel crucial na conversão de cada pixel do recorte num número binário, otimizando a representação para posterior análise, onde foram utilizadas redes neuronais pré-treinadas Inception v3 para reconhecimento da imagem. O comando do *widget* "*data table*" recebeu o conjunto de dados, apresentando-o numa folha de cálculo organizada para uma análise mais aprofundada.

Quatro características da Rede Neural foram configuradas na seguinte ordem: Ativação, Regularização, Neurônios em camadas ocultas e Número máximo de iterações. A Rede Neural começou com a configuração padrão do software Orange, que consistia em 100 neurônios em camadas ocultas com 2048 neurônios na camada de entrada e 2 na camada de saída, função de ativação Relu, parâmetro de regularização de 0,0001 e 200 iterações. Na etapa de inicialização (primeira etapa), valores aleatórios foram atribuídos aos pesos e limiares. Valores iniciais aleatórios, uniformemente distribuídos e pequenos, foram escolhidos na ausência de conhecimento prévio. A etapa de ativação (segunda etapa) envolveu o cálculo dos valores dos neurônios na camada oculta e na camada de saída. A etapa de treinamento (terceira etapa) focou no ajuste dos pesos, calculando os erros dos neurônios nas camadas de saída e oculta, computando correções de peso e atualizando os pesos. A etapa final envolveu a definição do processo de iteração para

se repetir a partir da etapa de ativação até que o critério de erro desejado fosse atendido. A rede foi então ajustada em quatro etapas para otimizar seu desempenho. A primeira etapa envolveu a comparação das funções de ativação (*Tanh*, *Logistic* e *Relu*). Em seguida, na segunda etapa, a estabilidade dos pesos dos neurônios foi melhorada aplicando regularização com valores de 0,05, 0,5 e 5. Para a terceira etapa, foram testados diferentes números de neurônios na camada oculta, incluindo 10, 25, 50, 150 e 350. Finalmente, na quarta etapa, o número de iterações na Rede Neural foi variado, especificamente 500 e 1000.

A plataforma Orange utiliza a validação cruzada como método de amostragem para avaliar algoritmos de aprendizado. A validação cruzada divide o conjunto de dados em um número específico de subconjuntos, chamados “*folds*”, geralmente 5 ou 10, como recomendado pelo programa Orange. Para esta análise, foram selecionados 10 *folds* devido ao desempenho superior na análise dos modelos SVM, Árvore de Decisão e Rede Neural. Durante a validação cruzada, o algoritmo é testado retendo exemplos de um *fold* de cada vez enquanto constrói o modelo usando os outros *folds*. Os exemplos retidos são então usados para avaliar o desempenho do modelo. Esse processo é repetido para todos os *folds* no conjunto de dados. Ao adotar essa abordagem, o Orange testa e analisa efetivamente o desempenho de algoritmos de aprendizado, garantindo resultados confiáveis para a avaliação do modelo.

A validação cruzada é especialmente útil quando o conjunto de dados é limitado em tamanho, pois permite o melhor uso dos dados disponíveis maximizando as iterações de treinamento e teste. Ajuda a mitigar problemas relacionados ao *overfitting* e *underfitting*, pois o modelo é exposto a diferentes partes dos dados durante cada *fold*. Além disso, a validação cruzada de 10 *folds* é amplamente utilizada como um método padrão para a avaliação de modelos devido ao seu equilíbrio entre eficiência computacional e estimativa confiável de desempenho (Favero; Belfiore; De Freitas, 2023). Ao configurar sistematicamente a Rede Neural, o desempenho foi melhorado e as configurações mais ótimas foram determinadas para cada uma das quatro características. Esse processo permitiu uma Rede Neural mais poderosa e eficiente, capacitando-a a lidar efetivamente com uma ampla gama de tarefas. A implementação desses algoritmos foi realizada usando o programa Orange, escolhido por fornecer os algoritmos necessários. O objetivo era testar a eficácia da Árvore de Decisão, SVM e Rede Neur.

1.3 RESULTADOS

Os resultados de avaliação para três modelos distintos de aprendizado de máquina estão apresentados na Tabela 1. A Rede Neural exibe alta capacidade de classificação, como evidenciado por sua pontuação AUC de 0,986. O valor AUC significa a notável capacidade da Rede Neural em distinguir entre instâncias positivas e negativas. Além disso, esse modelo alcança um F1 de 0,945, indicando um equilíbrio harmonioso entre precisão e *Recall* em tarefas de classificação binária. A precisão e o *Recall* do modelo atingem impressionantes 94,5%, destacando sua capacidade de identificar com precisão instâncias positivas. No geral, a Rede Neural emerge como um classificador altamente eficaz com desempenho notável em várias métricas de avaliação.

Tabela 1 - Resultados dos modelos de Rede Neural, SVM e Árvore de Decisão

Modelos	AUC	CA	F1	Precisão	<i>Recall</i>
Rede Neural	0.986	0.945	0.945	0.945	0.945
SVM	0.986	0.940	0.940	0.940	0.940
Árvore de Decisão	0.744	0.765	0.765	0.765	0.765

Da mesma forma, o modelo SVM também apresenta notáveis capacidades de classificação. Com uma pontuação AUC de 0,986, o SVM demonstra excelente capacidade de discriminação entre instâncias positivas e negativas, semelhante à Rede Neural. Embora seu F1 de 0,940 seja ligeiramente inferior ao da Rede Neural, ainda representa um equilíbrio louvável entre precisão e *Recall*. O SVM atinge uma precisão de 94,0%, indicando sua capacidade de identificar corretamente uma parte significativa das instâncias positivas previstas. Da mesma forma, o valor de *Recall* de 94,0% enfatiza a proficiência do modelo em identificar instâncias positivas reais. Embora o desempenho do SVM seja ligeiramente inferior ao da Rede Neural, ele permanece um classificador potente com resultados impressionantes nas métricas de avaliação.

Por outro lado, o modelo de Árvore de Decisão, embora competente, apresenta desempenho comparativamente inferior ao da Rede Neural e do SVM. Sua pontuação AUC de 0,744 indica uma capacidade razoavelmente boa de distinguir entre instâncias positivas e negativas, mas fica aquém dos resultados dos outros modelos. Da mesma forma, o F1 de 0,765 sugere um equilíbrio decente entre precisão e *Recall*, proporcionando previsões razoavelmente precisas. A precisão e o *Recall* do modelo atingem ambos 76,5%, indicando sua capacidade de identificar corretamente uma parte considerável das instâncias positivas previstas e instâncias positivas reais. Embora o modelo de Árvore de Decisão demonstre um desempenho respeitável, ele fica um pouco atrás dos modelos de Rede Neural e SVM em termos de métricas de avaliação, sugerindo áreas potenciais para otimização e aprimoramento.

A Tabela 2 apresenta as matrizes de confusão obtidas de três classificadores distintos: Rede Neural, Árvore de Decisão e SVM aplicados a um problema de classificação binária com o objetivo de discernir sojas com falhas daquelas sem falhas. Os resultados revelam diferenças notáveis no desempenho dos classificadores. A Rede Neural exibe uma alta precisão, pois classificou incorretamente apenas 11 imagens no total, resultando em uma taxa de erro impressionantemente baixa de 5,5%. Sua superior capacidade de discernir padrões e relações nos dados contribui para sua classificação precisa de ambas as classes.

Por outro lado, a Árvore de Decisão, com um erro de 47 imagens (23,5%), demonstra uma taxa mais alta de classificação incorreta. Essa limitação pode ser atribuída à dificuldade inerente da Árvore de Decisão em capturar padrões intrincados de dados, o que prejudica seu poder discriminativo. O SVM, por sua vez, entrega um desempenho moderado com um erro de 24 imagens (12%). Embora o SVM supere a Árvore de Decisão, ele fica aquém em comparação com a Rede Neural em termos de precisão e robustez. Essa discrepância sugere que a Rede Neural possui uma melhor capacidade de identificar sojas com e sem falhas, tornando-a o classificador mais promissor para aplicações agrícolas do mundo real. Os resultados da Tabela 2 destacam a eficácia da Rede Neural em caracterizar precisamente plantações de soja, mostrando seu potencial para auxiliar na tomada de decisões agrícolas.

Tabela 2 - Matriz de Confusão - Rede Neural, Árvore de Decisão e SVM

Rede Neural				Árvore de Decisão				SVM						
		Predito			Predito			Predito						
		com falhas	sem falhas	Σ			Σ			Σ				
Real	com falhas	93	7	100	Real	com falhas	77	23	100	Real	com falhas	89	11	100
	sem falhas	4	96	100		sem falhas	24	76	100		sem falhas	13	87	100
Σ		104	96	200	Σ		101	99	200	Σ		101	99	200

De acordo com as curvas ROC (*Receiver Operating Characteristic*) representadas nas Figuras 3 e 4, é evidente que a Rede Neural superou os outros modelos em termos de precisão. No software foi gerado as curvas ROC para os modelos testados e suas respectivas envoltórias convexas. Isso permite uma comparação entre modelos de classificação. A curva representa a relação entre a taxa de falso positivo (1 para especificidade, que é a probabilidade de o alvo ser 1 quando o valor verdadeiro é 0) no eixo x e a taxa de verdadeiro positivo (sensibilidade, que é a probabilidade de o alvo ser 1 quando o valor verdadeiro é 1) no eixo y. Um classificador é considerado mais preciso se sua curva seguir perto a borda esquerda e, em seguida, a borda superior do espaço ROC (Hanley; McNeil, 1982).

Figura 3 - Curva ROC dos modelos de Rede Neural (azul), SVM (laranja) e Árvore de Decisão (verde) dos recortes com falhas de semeadura

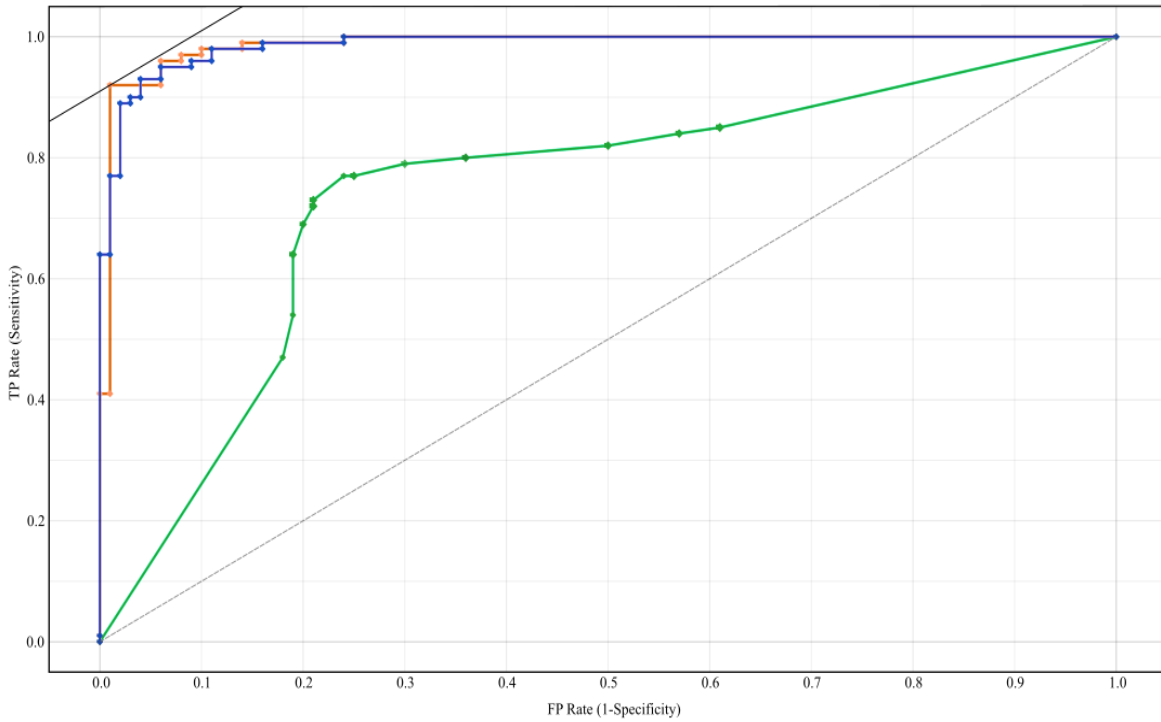
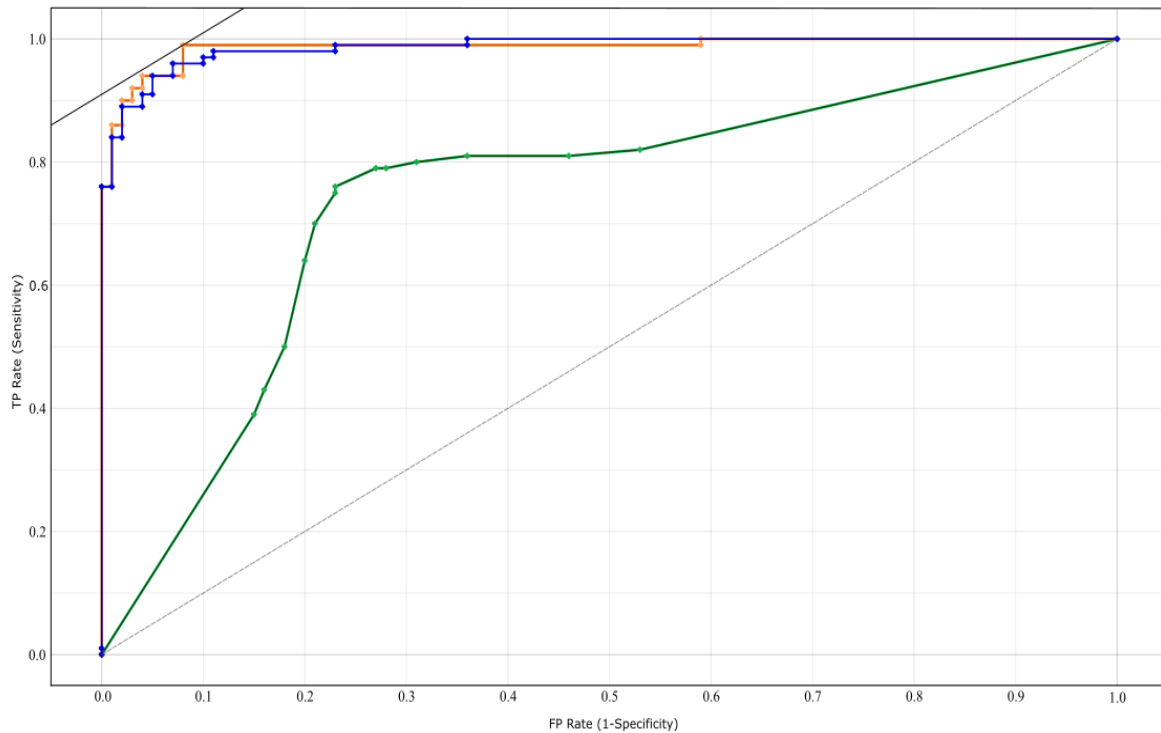


Figura 4 - Curva ROC dos modelos de Rede Neural (azul), SVM (laranja) e Árvore de Decisão (verde) dos recortes sem falhas de semeadura



Com base nos dados apresentados na Tabela 1, é evidente que a Rede Neural emergiu como o claro líder, entregando uma taxa de precisão de 98%. Além disso, uma análise detalhada da matriz de confusão (Tabela 2) revela uma incidência incrivelmente baixa de erros para este algoritmo, que classificou incorretamente apenas 11 recortes no total. Como resultado, com a identificação do algoritmo superior, a fase subsequente envolveu uma exploração detalhada com o objetivo de descobrir a configuração ótima para a Rede Neural, maximizando assim seu potencial e refinando seu desempenho. Uma limitação da Rede Neural é que aumentar o número de camadas e neurônios pode levar ao aumento do número de parâmetros, tornando-o computacionalmente pesado e potencialmente causando problemas de convergência.

A Rede Neural foi configurada (no número de neurônios nas camadas ocultas, função de ativação, regularização e iterações) e os valores testados foram escolhidos para abranger uma variedade de opções. Como resultado, a Rede Neural final utilizou a função de ativação Logística, 10 neurônios por camada oculta, um fator de regularização de 5 e um máximo de 500 iterações (Tabela 3) (Demšar *et al.*, 2013).

Tabela 3 - Resultados das Configurações do Modelo de Rede Neural

Resultados de Avaliação - Rede Neural					
Função de ativação	AUC	CA	F1	Precisão	Recall
Relu	0,986	0,945	0,945	0,945	0,945
Logistic	0,993	0,950	0,950	0,950	0,950
Tanh	0,990	0,945	0,945	0,945	0,945
Regularização	AUC	CA	F1	Precisão	Recall
0.05	0,979	0,930	0,930	0,930	0,930
0.5	0,979	0,930	0,930	0,930	0,930
5	0,984	0,935	0,935	0,935	0,935
Neurônios nas camadas ocultas	AUC	CA	F1	Precisão	Recall
10	0,984	0,940	0,940	0,940	0,940
25	0,983	0,935	0,935	0,935	0,935
50	0,983	0,940	0,940	0,940	0,940
150	0,983	0,940	0,940	0,940	0,940
350	0,984	0,945	0,945	0,945	0,945
Número máximo de iterações	AUC	CA	F1	Precision	Recall
500	0,984	0,945	0,945	0,945	0,945
1000	0,984	0,945	0,945	0,945	0,945

Os resultados da avaliação do modelo de Rede Neural com várias configurações são apresentados na Tabela 3. Esses resultados são baseados em diferentes funções de ativação, forças de regularização, número de neurônios em camadas ocultas e número máximo de iterações. Em relação às funções de ativação, o modelo foi testado com Relu, Logistic e Tanh. Entre essas, a função de ativação Logistic demonstrou a maior AUC de 0,993, sugerindo habilidades superiores de separação de classes. Alcançou também os maiores valores para CA, F1, Precisão e *Recall*, todos iguais a 0,950, indicando seu desempenho ótimo na classificação precisa de instâncias.

As forças de regularização examinadas foram 0,05, 0,5 e 5. Embora todas as três configurações de regularização tenham gerado valores altos de AUC (variando de 0,979 a 0,984), a força de regularização de 5 alcançou 0,935 de CA, F1, Precisão e *Recall* indicando melhor desempenho geral e generalização para dados não vistos.

Ao determinar o número de neurônios nas camadas ocultas, a seleção empírica é tipicamente utilizada. É importante equilibrar e evitar o uso de muitas unidades, o que pode levar a que a rede memorize os dados de treinamento (*overfitting*), em vez de extrair características gerais para uma generalização eficaz. Por outro lado, usar muitos poucos neurônios pode fazer com que a rede gaste tempo excessivo procurando uma representação ótima. Considerando esses desafios, geralmente é recomendado manter o número de neurônios ocultos relativamente baixo, mas não a ponto de restringir sua capacidade (Favero; Belfiore; De Freitas, 2023). Em relação ao número de neurônios nas camadas ocultas, o modelo foi avaliado com 10, 25, 50, 150 e 350 neurônios. A configuração com 10 neurônios demonstrou a maior AUC de 0,984, juntamente com os maiores valores de CA, F1, Precisão e *Recall*, todos iguais a 0,940. Isso sugere que usar 10 neurônios nas camadas ocultas resultou na representação mais eficaz dos padrões subjacentes nos dados e aumentar os neurônios para 350 pode não fornecer melhorias substanciais no desempenho, enquanto aumenta o risco de *overfitting*.

Além disso, o número máximo de iterações testadas foi 500 e 1000. Ambas as iterações alcançaram uma AUC de 0,984, indicando forte convergência e capacidade de generalização. As métricas de avaliação (CA, F1, Precisão e *Recall*) foram constantes em 0,945, sugerindo desempenho consistente e estável do modelo, independentemente do número de iterações.

Com base nos resultados da Tabela 3, a configuração ótima do modelo de Rede Neural envolve o uso da função de ativação logística, definindo a força de regularização como 5, utilizando 10 neurônios nas camadas ocultas e limitando o número máximo de iterações para 500. Essa configuração garante uma separação superior de classes, excelente generalização para dados não vistos e desempenho consistente em todas as métricas de avaliação, tornando-a a escolha mais eficaz para a tarefa de classificação específica.

A decisão de selecionar a melhor configuração para um modelo de Rede Neural depende de vários fatores, incluindo o problema específico (no caso deste estudo a falha na semente), as características dos dados e o equilíbrio entre a complexidade do modelo e a generalização (Haykin, 2009). Embora um maior número de iterações possa parecer benéfico para refinar os pesos do modelo e alcançar maior precisão, também pode levar ao *overfitting*, especialmente se os dados forem insuficientes para suportar tal complexidade. *Overfitting* ocorre quando o modelo se torna muito especializado nos dados de treinamento, resultando em desempenho reduzido em dados não vistos. Neste caso, a configuração com 500 iterações já alcançou uma forte AUC de 0,984 e métricas de avaliação consistentes com CA, F1, Precisão e *Recall* todos em 0,945. Portanto, aumentar as iterações para 1000 pode não levar a melhorias significativas e pode introduzir desnecessariamente complexidade e *overfitting* (Favero; Belfiore; De Freitas, 2023).

Com relação a configuração de 1000 iterações e 350 neurônios em camadas ocultas pode não ser a melhor escolha devido ao potencial de *overfitting*. As configurações com 500 iterações e 10 neurônios alcançaram um desempenho forte com risco mínimo de *overfitting*. Encontrar um equilíbrio entre a complexidade do modelo e a generalização é crucial, e, neste caso, a configuração com 500 iterações e 10 neurônios parece efetivamente atingir esse equilíbrio, tornando-a a escolha melhor para a tarefa de classificação específica. A Matriz de Confusão serve como uma representação visual da distribuição de instâncias entre classes previstas e reais. Conforme mostrado na Tabela 4, é evidente que 4 recortes representando áreas com falhas foram classificadas incorretamente como sem falhas, enquanto 5 dos recortes sem falhas foram erroneamente identificadas como com falhas, resultando em um total de 9 imagens classificadas incorretamente, equivalente a 4,5%.

Tabela 4 - Matriz de Confusão da Rede Neural Configurada

Rede Neural		Predito		Σ
		com falha	sem falha	
Real	com falha	96	4	100
	sem falha	5	95	100
Σ		101	99	200

1.4 DISCUSSÕES

É importante observar que o desempenho do modelo de Árvore de Decisão obteve resultado inferior da Rede Neural e do SVM, indicando áreas com potencial para otimização e aprimoramento adicionais. Em um estudo conduzido por Lu *et al.* (2023), o foco foi principalmente na otimização de hiperparâmetros para modelos de classificação de culturas. No entanto, este estudo centrou-se principalmente na configuração de hiperparâmetros para tarefas específicas de classificação de culturas. Em contraste, a pesquisa expandiu seu escopo para abranger a avaliação e otimização abrangentes de três modelos distintos de aprendizado de máquina: Redes Neural, SVM e Árvore de Decisão. Por meio desta análise comparativa extensiva, o estudo facilitou a seleção do modelo mais adequado para a classificação de falhas de soja, oferecendo uma perspectiva mais ampla sobre a aplicabilidade de modelos de aprendizado de máquina em ambientes agrícolas.

Kumar e Shukla (2022) mergulharam no domínio de métodos de aprendizado em conjunto para a detecção de doenças em culturas no contexto da agricultura de precisão. Embora seu foco principal tenha sido em uma aplicação agrícola distinta, seu trabalho é relevante ao considerar a avaliação de modelos de aprendizado de máquina. Patel *et al.* (2023) também conduziram avaliações de vários modelos de aprendizado de máquina. No entanto, sua ênfase foi em um objetivo agrícola diferente, especificamente na identificação de doenças em culturas. Apesar das tarefas

agrícolas diferentes, ambos os estudos destacam a importância crítica da seleção e otimização meticolosas do modelo como fatores cruciais para obter resultados confiáveis no domínio do aprendizado de máquina para fins agrícolas.

Os resultados deste estudo destacam o potencial substancial das redes neurais para aprimorar o manejo da semeadura da soja e facilitar a previsão de rendimento. Explorações adicionais podem ser conduzidas para avaliar o desempenho do método em condições ambientais e sazonais diversas. Além disso, é essencial avaliar sua escalabilidade no tratamento de fazendas e conjuntos de dados maiores, garantindo sua viabilidade prática em aplicações agrícolas. Com pesquisas contínuas e aprimoramentos, as redes neurais têm a perspectiva de revolucionar as práticas de cultivo de soja e contribuir para o cultivo de safras mais eficiente e sustentável no futuro.

A tecnologia de visão computacional possui potencial para auxiliar os agricultores na detecção de falhas em áreas de soja, otimizando assim o rendimento da colheita e reduzindo o desperdício ao identificar com precisão regiões com densidade de plantas inadequada ou excessiva. Essa capacidade não apenas capacita os agricultores a tomar medidas proativas, mas também lhes permite tomar decisões baseadas em dados para aprimorar o manejo da colheita. À medida que a tecnologia continua a avançar a passos largos, podemos antecipar o surgimento de sistemas de visão computacional altamente avançados capazes de monitorar em tempo real a densidade da colheita, revolucionando a maneira como os agricultores gerenciam suas plantações. A integração de tais sistemas abre novas possibilidades para estratégias de manejo de colheitas mais precisas e eficientes, revolucionando a paisagem agrícola e maximizando a produtividade agrícola (Gonzalez; Woods, 2010).

Ao aproveitar o poder da visão computacional, os agricultores podem otimizar o cultivo de soja. Por meio da detecção de falhas em áreas de soja, a tecnologia de visão computacional ajuda os agricultores a tomar decisões para melhorar o rendimento da colheita e minimizar o desperdício, como redistribuir recursos ou implementar medidas corretivas, para garantir condições de crescimento ideais. À medida que a tecnologia avança, o desenvolvimento de sistemas de visão computacional capazes de monitorar a densidade da colheita em tempo real torna-se cada vez mais viável. Essa inovação capacitaria os agricultores com insights em tempo real e lhes permitiria implementar estratégias de manejo de colheitas altamente adaptadas e eficientes. Ao aproveitar o potencial da visão computacional na

agricultura, os agricultores podem maximizar o rendimento da colheita, minimizar o desperdício e alcançar práticas agrícolas sustentáveis e eficientes (Gonzalez; Woods, 2010).

Esses resultados servem como uma base sólida para o desenvolvimento futuro de algoritmos destinados à detecção de falhas por meio de análise de imagens e podem ser um ponto de partida para a construção de modelos de redes neurais mais sofisticados. Projetos futuros podem explorar outras plataformas para análise de dados a fim de expandir as capacidades dos algoritmos para abordar diferentes funções. Embora este estudo tenha utilizado principalmente imagens RGB, outros estudos podem se beneficiar do uso de dados de várias fontes para analisar os resultados de maneira mais abrangente (Maes e Steppe, 2019; Pathak *et al.*, 2022).

Ampliar o escopo de projetos futuros para incluir dados de sensores diversos contribuirá para uma compreensão mais holística do problema, abrindo caminho para algoritmos e modelos aprimorados de detecção de falhas. Ao incorporar dados de fontes diferentes, os pesquisadores podem obter insights valiosos sobre as complexidades da detecção de falhas em campos agrícolas, abordando os desafios de maneira mais eficaz. A integração de dados de vários sensores promoverá uma perspectiva abrangente, permitindo que os pesquisadores tomem decisões informadas e desenvolvam soluções inovadoras para otimizar o manejo de culturas e a utilização de recursos (Valente *et al.*, 2020; Yang, 2020).

1.5 CONCLUSÕES

A Rede Neural despontou como o algoritmo mais promissor para caracterizar áreas de cultivo de soja, motivando ajustes em suas configurações para aprimorar a classificação precisa de áreas de semeadura com e sem falhas. Os resultados desta pesquisa destacam o potencial significativo das Redes Neurais na resolução de desafios de classificação de falhas de soja em imagens de RPA. Ao aproveitar uma variedade de fontes de dados e empregar algoritmos avançados, os pesquisadores podem impulsionar a tecnologia agrícola, tornando as práticas de agricultura de precisão mais eficientes, sustentáveis e, em última análise, contribuindo para a segurança alimentar global.

REFERÊNCIAS

- BALL, R.A.; PURCELL, L.C.; VORIES, E.D. Optimizing Soybean Plant Population for a Short-Season Production System in the Southern USA. **Crop Science**, 40, 757–764, 2000.
- CONAB, 2023. Companhia Nacional de Abastecimento. **Acompanhamento da safra brasileira de grãos**. Brasília, v.7, Safra 2022/23, n.10, Décimo levantamento, p. 1-110, 2023.
- CORASSA, G.M. *et al.* Optimum soybean seeding rates by yield environment in southern Brazil. **Agronomy Journal**, v. 110, n. 6, p. 2430-2438, 2018.
- DAVIES, E.R. **Computer vision: principles, algorithms, applications, learning**. 5. ed., p. 900, 2017.
- DE BRUIN, J.L.; PEDERSEN, P. New and old soybean cultivar responses to plant density and intercepted light. **Crop science**, v. 49, n. 6, p. 2225-2232, 2009.
- 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. 2023.
- DE MILANO, D.; HONORATO, L.B. **Visão computacional**. UNICAMP Universidade Estadual de Campinas FT Faculdade de Tecnologia, 2014.
- DEMŠAR, J *et al.* **Orange: Data Mining Toolbox in Python**, Journal of Machine Learning Research 14(Aug): 2349–2353, 2013.
- EGLI, D. B. Plant density and soybean yield. **Crop Science**, v. 28, n. 6, p. 977-981, 1988.
- DOS SANTOS, F. C. *et al.* **A agropecuária do sul do Matopiba em perspectiva: Circuito solos arenosos da região Cocos-Jaborandi**. 2018.
- EMBRAPA. **Sistema brasileiro de classificação de solos**. Centro Nacional de Pesquisa de Solos: Rio de Janeiro, v. 3, 2013.
- ENDRES, V. C. Espaçamento, densidade e época de semeadura. In: EMBRAPA. Centro de Pesquisa Agropecuária do Oeste (Dourados, MS). **Soja: recomendações técnicas para Mato Grosso do Sul e Mato Grosso**. Dourados, 1996. p. 82-85. (Circular Técnica, 3).
- FAVERO, L. P.; BELFIORE, P.; DE FREITAS S., R. **Data science, analytics and machine learning with R**. Academic Press, 2023.

FISS, G. *et al.* Produtividade e características agronômicas da soja em função de falhas na semeadura. **Revista de Ciências Agrárias Amazonian Journal of Agricultural and Environmental Sciences**, v. 61, 2018.

GAN, Y. *et al.* Physiological response of soybean genotypes to plant density. **Field Crops Research**, v. 74, n. 2-3, p. 231-241, 2002.

GASPAR, A.P. *et al.* Defining Optimal Soybean Seeding Rates and Associated Risk across North America. **Agronomy Journal**, 112, 2928-2943, 2020.

GEIPEL, J.; LINK, J.; CLAUPEIN, W. Combined Spectral and Spatial Modeling of Corn Yield Based on Aerial Images and Crop Surface Models Acquired with an Unmanned Aircraft System. **Remote Sens.** 6, 10335-10355, 2014.

GONZALEZ, R.C.; WOODS, R. E. Digital Image Processing, 3.ed., p. 624, 2009.

HABIBI, L.N. *et al.* 2021. Machine Learning Techniques to Predict Soybean Plant Density Using UAV and Satellite-Based Remote Sensing. **Remote Sensing**, 13, 2548, 2010.

HANLEY, J. A.; MCNEIL, B. J. The meaning and use of the area under a receiver operating characteristic (ROC) curve. **Radiology**, v. 143, n. 1, p. 29-36, 1982.

HAYKIN, S. **Neural networks and learning machines**, 3/E. Pearson Education India, 2009.

HUANG, C.; XIE, J. Rapeseed Stand Count Estimation at Leaf Development Stages with UAV Imagery and Convolutional Neural Networks. **Front. Plant Sci.**11, 617, 2020.

JIANG, Y.; LI, C.; PATERSON, A.H.; ROBERTSON, J.S. DeepSeedling: Deep convolutional network and Kalman filter for plant seedling detection and counting in the field. **Plant Methods**, 15, 141, 2019.

JIN, X. *et al.* Estimates of plant density of wheat crops at emergence from very low altitude UAV imagery. **Remote Sensing of Environment**, v. 198, p. 105-114, 2017.

KUMAR, S.; SHUKLA, A. Automatic Grading of Potato Leaf using Machine learning & Computer Vision. **Research Square**, 2022.

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

LIAKOS, K. G. *et al.* Machine learning in agriculture: A review. **Sensors**, v. 18, n. 8, p. 2674, 2018.

LIU, K. S. **Soybeans: Chemistry, Technology, and Utilization**. International Thomson Publishing, New York.1997.

LU, T., GAO, M.; WANG, L. Crop classification in high-resolution remote sensing images based on multi-scale feature fusion semantic segmentation model. **Frontiers in Plant Science** , v. 14, p. 1196634, 2023.

MAES, W. H.; STEPPE, K. Perspectives for remote sensing with unmanned aerial vehicles in precision agriculture. **Trends in plant science**, v. 24, n. 2, p. 152-164, 2019.

PATEL A; MISHRA R; SHARMA A. Maize Plant Leaf Disease Classification Using Supervised Machine Learning Algorithms. **Fusion: Practice and Applications**. v.13, n. 2, p. 08-21, 2023.

PATHAK, H. *et al.* A review of unmanned aerial vehicle-based methods for plant stand count evaluation in row crops. **Computers and Electronics in Agriculture**, v. 198, p. 107064, 2022.

PEÑA, J. M. *et al.* Quantifying efficacy and limits of unmanned aerial vehicle (UAV) technology for weed seedling detection as affected by sensor resolution. **Sensors**, v. 15, n. 3, p. 5609-5626, 2015.

RANDELOVIĆ, P. *et al.* Prediction of soybean plant density using a machine learning model and vegetation indices extracted from RGB images taken with a UAV. **Agronomy**, v. 10, n. 8, p. 1108, 2020.

RAJPUT, A. S.; SHUKLA, S.; THAKUR, S. S. Soybean leaf diseases detection and classification using recent image processing techniques. **International Journal of Students' Research in Technology & Management** eISSN, p. 2321-2543, 2020.

SHARMA, N. **XGBoost. The extreme gradient boosting for mining applications**. Munich: GRIN Verlag, 2018.

VALENTE, J. *et al.* Automated crop plant counting from very high-resolution aerial imagery. **Precision Agriculture**, v. 21, p. 1366-1384, 2020.

SOUZA, F. L. P. DE .; FAVAN, J. R. ., PASSOS, J. R. DE S. ., DIAS , M. A. ., CAMPOS, S., 2022. Machine learning e processamento digital de imagens UAV: uma abordagem para estimar distribuição longitudinal de plantas de soja. **Energia na agricultura**, v. 37, n. 3, p. 1-11, 2022.

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

TOURINO, M. C.C.; REZENDE, P.M.; SALVADOR, N, 2002. Espaçamento, densidade e uniformidade de semeadura na produtividade e características agrônômicas da soja. **Pesquisa agropecuária brasileira**, v. 37, p. 1071-1077, 2002.

WU, J. *et al.* Automatic counting of in situ rice seedlings from UAV images based on a deep fully convolutional neural network. **Remote Sensing**, v. 11, n. 6, p. 691, 2019.

ZAN, X. *et al.* Automatic detection of maize tassels from UAV images by combining random forest classifier and VGG16. **Remote Sensing**, v. 12, n. 18, p. 3049, 2020

ZHANG, J. *et al.* Rapeseed stand count estimation at leaf development stages with UAV imagery and convolutional neural networks. **Frontiers in plant science**, v. 11, p. 617, 2020.

CAPÍTULO 2

CONTAGEM DE PLANTAS DE SOJA COM IMAGENS RGB DE RPA: UMA ABORDAGEM EFICAZ DURANTE MUDANÇAS FENOLÓGICAS

Autora: Flávia Luize Pereira de Souza

RESUMO

A produção de soja (*Glycine max* (Linnaeus) Merrill) nos Estados Unidos desempenha um papel crucial na agricultura, ocupando uma extensão considerável de área cultivada. Entretanto, os custos associados à produção de soja têm demonstrado um aumento notável nos últimos anos, com as despesas relacionadas às sementes representando uma proporção significativa do total. Este aumento de custos é atribuído a diversos fatores, incluindo a introdução de características genéticas patenteadas e protegidas, bem como pressões inflacionárias. A contagem precisa do número de plantas de soja por unidade de área é essencial para monitorar a emergência e as perdas na densidade populacional de plantas, desempenhando um papel crítico na ciência e prática agrícola. No entanto, os métodos tradicionais de contagem manual são ineficientes e sujeitos a imprecisões, devido a variáveis como a densidade das plantas, limitações na percepção visual humana e a representatividade das amostras coletadas. Este estudo investigou a viabilidade da utilização de Aeronaves Remotamente Pilotadas (RPAs) para monitorar a densidade populacional de plantas de soja em diferentes estádios fenológicos. O experimento foi conduzido na Ben Hur Research Farm, localizada no sul dos Estados Unidos, utilizando imagens de RPA e técnicas avançadas de processamento de imagens para a contagem de plantas. Os resultados demonstraram que o estágio fenológico VC apresentou a maior acurácia, devido ao aumento na visibilidade das plantas à medida que estas crescem. Além disso, a influência da velocidade de semeadura na precisão da contagem de plantas foi examinada. Observou-se que o aumento da velocidade de semeadura resultou em uma diminuição na precisão do modelo, devido à maior sobreposição entre as plantas. No entanto, ao excluir as plantas sobrepostas, a acurácia da contagem manual pela imagem de RPA foi alta no estágio VC. Este

proeijo contribui para o avanço do conhecimento científico ao fornecer insights sobre a dinâmica entre estádios fenológicos da soja, velocidade de semeadura e técnicas de contagem de plantas. A abordagem integrada utilizando tecnologias de RPA e processamento de imagens oferece uma maneira eficaz de otimizar a contagem de plantas na cultura da soja, fornecendo informações valiosas para melhorar as práticas agrícolas e maximizar a produção. Estes resultados não apenas enriquecem a literatura científica, mas também têm implicações práticas significativas para a agricultura de precisão e o manejo de culturas.

Palavras-chave: contagem de plantas; estágio fenológico; velocidade de semeadura; processamento de imagens; RPA.

ABSTRACT

Soybean (*Glycine max* (Linnaeus) Merrill) production in the United States plays a crucial role in agriculture, occupying a considerable amount of cultivated land. However, the costs associated with soybean production have shown a notable increase in recent years, with seed-related expenses accounting for a significant proportion of the total. This increase in costs is attributed to a number of factors, including the introduction of patented and protected genetic traits, as well as inflationary pressures. Accurate counting of the number of soybean plants per unit area is essential for monitoring emergence and losses in plant population density, playing a critical role in agricultural science and practice. However, traditional manual counting methods are inefficient and subject to inaccuracies, due to variables such as plant density, limitations in human visual perception and the representativeness of the samples collected. This study investigated the feasibility of using Remotely Piloted Aircraft (RPAs) to monitor the population density of soybean plants at different phenological stages. The experiment was conducted at Ben Hur Research Farm, located in the southern United States, using RPA images and advanced image processing techniques to count plants. The results showed that the VC phenological stage had the highest accuracy, due to the increased visibility of the plants as they grow. In addition, the influence of sowing speed on plant counting accuracy was examined. It was observed that increasing the sowing speed resulted in a decrease in the model's accuracy, due to the greater overlap between plants. However, when excluding overlapping plants, the accuracy of manual counting using the RPA image was high at the VC stage. This project contributes to the advancement of scientific knowledge by providing insights into the dynamics between soybean phenological stages, sowing speed and plant counting techniques. The integrated approach using RPA and image processing technologies offers an effective way to optimize plant counting in the soybean crop, providing valuable information to improve agricultural practices and maximize production. These results not only enrich the scientific literature, but also have significant practical implications for precision agriculture and crop management.

Keywords: plant count; phenological stage; sowing speed; image processing; RPA.

2.1 INTRODUÇÃO

A soja (*Glycine max (Linnaeus) Merrill*) desempenha um papel significativo na agricultura dos EUA, cobrindo uma área de 36,8 milhões de hectares para produção em 2022 (USDA-NASS, 2022). Nos últimos anos, o custo da produção de soja aumentou consideravelmente, sendo a semente o fator mais caro, representando 32% do total das despesas operacionais dos produtores de soja dos EUA. Esse aumento de custo, de mais de 200% entre 2000 e 2021, é atribuída em grande parte à inclusão de características patenteadas e protegidas por propriedade intelectual, melhorias na qualidade genética das sementes e pressões inflacionárias (Bergada *et al.*, 2015; Chen; Wiatrak, 2011; Lee *et al.*, 2008; USDA-NASS, 2022).

O número de plantas de soja por unidade de área fornece informações sobre a emergência e as perdas na densidade das plantas, o que é importante para a ciência e produção agrícola. A contagem convencional de plantas é normalmente feita manualmente por uma equipe de campo, leva muito tempo e não é escalonável. Este método é sujeito a imprecisões devido a variáveis como a densidade das plantas, limitações na percepção visual humana, a representatividade das amostras colhidas, as metodologias usadas na amostragem e manter um monitoramento espaço-temporal contínuo de campos de grande escala utilizando esses métodos é um desafio. (Li *et al.*, 2019; Xu *et al.*, 2020; Wei e Molin, 2020). Portanto, a implementação de tecnologias como as Aeronaves Remotamente Pilotadas (RPAs) surge como uma solução oportuna para monitorar a taxa de população de plantas (Li, Wang e Huang, 2022).

A avaliação de povoamentos por meio de imagens digitais exhibe variações significativas entre distintas culturas (Pathak *et al.*, 2022). Embora a maioria das pesquisas sobre a contagem de povoamentos em culturas em linha tenha se concentrado, em grande parte, em milho (Gnädinger e Schmidhalter, 2017; Varela *et al.*, 2018; Pang *et al.*, 2020; Shirzadifar *et al.*, 2018; Garcia-Martinez *et al.*, 2020), algodão (Butler *et al.*, 2019; Feng *et al.*, 2020; Oh *et al.*, 2020), espinafre (Valente *et al.*, 2020), sorgo granífero (Ribera *et al.*, 2017), tabaco (Fan *et al.*, 2018; Kumar; Domic, 2019), e pomares de frutas ou nozes (Neupane *et al.*, 2019; Haddadi *et al.*, 2020), existem falhas consideráveis em abordagens para determinar contagens de

povoamentos em culturas como arroz, trigo (Jin *et al.*, 2017), canola (Zhao *et al.*, 2018; Zhang *et al.*, 2020), e soja (Randelović *et al.*, 2020) (Pathak *et al.*, 2022).

A tecnologia de sensoriamento remoto RPA, está se tornando cada vez mais importante na agricultura de precisão moderna. Analisando a literatura sobre a contagem de plantas de soja usando imagens de RPA, encontramos algumas contribuições científicas em potencial. Por exemplo, Yang *et al.* (2024) usaram imagens RGB entre o estágio VC (estádio cotilédone) e o estágio V2 (segundo estágio trifoliolado) analisaram os dados por meio do método de detecção rápida baseado no modelo YOLOv5s aprimorado. Seus resultados foram promissores e conseguiram detecção rápida de mudas densas de soja em ambientes de campo. Randelović *et al.*, (2020) apresentaram evidências da capacidade de estimar a população de plantas de soja, utilizando Índices de Vegetação extraídos de imagens de câmera com sensor vermelho, verde e azul (RGB) e empregando uma abordagem de aprendizado de máquina por meio de Random Forest, as plantas foram avaliadas durante as fases de desenvolvimento da soja V4 (terceira folha trifoliolada completamente desenvolvida) e R3 (vagem inicial). Eles obtiveram melhores resultados quando as plantas estavam em V4. No entanto, o hábito de crescimento da soja é complexo e leva a uma rápida sobreposição das folhas de plantas adjacentes, especialmente em densidades de semeadura mais altas, o que torna a contagem de plantas um desafio. Portanto, novas abordagens devem ser feitas.

Tradicionalmente, muitas configurações são usadas para a semeadura de soja, inclusive a velocidade de semeadura. O aumento da velocidade de semeadura tem o potencial de diminuir a uniformidade da distribuição das plantas na linha; por outro lado, a semeadura em velocidades mais altas pode aumentar a capacidade operacional e aumentar a área semeada (Bortoli, 2021), além de afetar a deposição de sementes no sulco de semeadura e o espaçamento entre plantas (Bertelli *et al.*, 2016). O processo de semeadura faz com que os mecanismos da máquina liberem as sementes em dois níveis de velocidade: primeiro, devido à queda livre da semente; e segundo, devido ao deslocamento da máquina (Pacheco *et al.*, 1996). O componente horizontal da velocidade direciona a semente para impactar o solo, ricocheteando no local de destino. Isso altera o espaço entre as sementes e pode resultar em falhas na semeadura ou em plantas sobrepostas.

Novos métodos de observação e avaliação de povoamentos ganham relevância, especialmente considerando a densidade de semeadura da soja em

comparação com outras culturas (Pathak *et al.*, 2022). Enquanto a contagem de povoamentos em culturas como milho e algodão tem sido extensivamente estudada, a soja apresenta desafios únicos devido à sua densidade de semeadura mais alta. A utilização de RPA para avaliação de povoamentos na soja demanda a superação de desafios como a sobreposição entre plantas, impactando a precisão das estimativas (Randelović *et al.*, 2020).

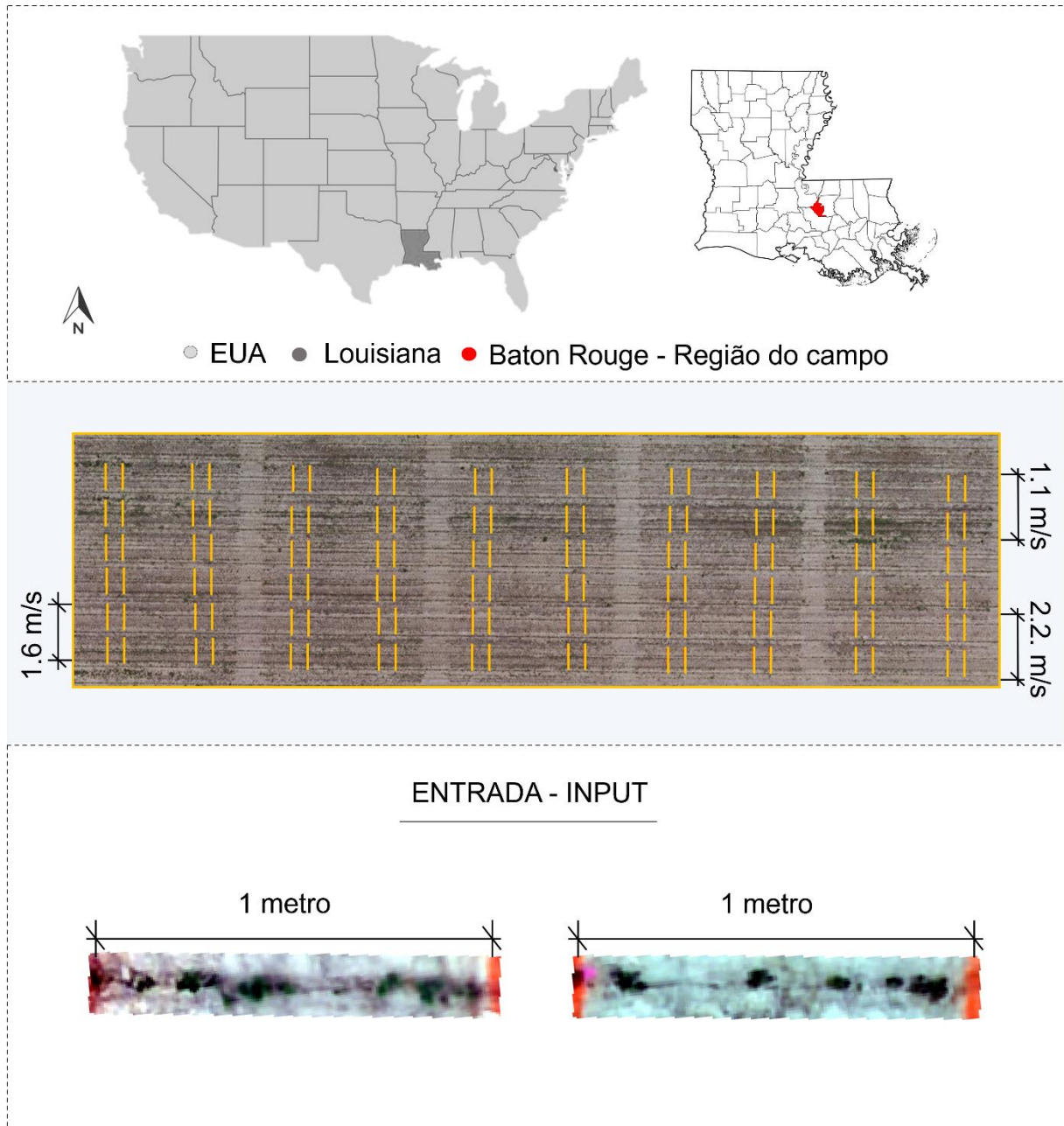
Diante desse cenário, este trabalho buscou contribuir para o avanço da compreensão da interdependência dinâmica entre o estágio fenológico da soja, a velocidade de semeadura, a captura de imagens por RPA e as técnicas de contagem de plantas, com o objetivo de identificar o período ideal para a contagem de soja com imagens de RPA e processamento de imagens. A combinação de conhecimento fenológico, tecnologias de RPA e técnicas de processamento de imagens representa uma abordagem integrada para otimizar a eficiência da contagem de plantas na cultura da soja, que pode fornecer informações valiosas para melhorar as práticas agrícolas e maximizar a produção.

2.2 MATERIAL E MÉTODOS

2.2.1 Área e delineamento experimental

O experimento (Figura 1) foi realizada na Ben Hur Research Farm, situada no município de Baton Rouge, Estado da Louisiana, Sul dos Estados Unidos. Esta fazenda, ocupada com culturas agrícolas e pastagens, está situada a 8000m ao sul do campus da Louisiana State University (LSU), proporcionando um ambiente propício para estudos agrícolas. A região tem características climáticas classificadas como subtropicais húmidas, com verões quentes e húmidos seguidos de invernos suaves. A época de semeadura da soja em Baton Rouge, alinhada com a primavera e o início do verão, coincide com o período de maior concentração de precipitação na região. A data de semeadura de 19 de maio de 2023 alinha-se com a prática comum no Meio-Oeste dos EUA, onde a soja é tradicionalmente plantada do final de abril a meados de junho (De Bruin; Pedersen, 2008; Robinson *et al.*, 2009; Vossenkemper *et al.*, 2015).

Figura 1 – Visualização esquemática da localização geográfica da área experimental, destacando as amostras selecionadas para a aplicação do modelo proposto. A obtenção de imagens RGB por meio de RPA



A semeadura foi realizada em três velocidades diferentes, foram 1,1 m/s (aproximadamente 4km/h), 1,6m/s (aproximadamente 6 km/h) e 2,2 m/s (aproximadamente 8 km/h), com a implementação de 4 linhas de semeadura para cada velocidade. As áreas de semeadura, totalizando 12 linhas e cobrindo uma área total de 0,13 hectares, foram dispostas em um desenho aleatório para a contagem de plantas. Esse delineamento, composto por 60 amostras e com 1 metro de largura

cada, garante uma representação estatística da população de plantas nas diversas condições de semeadura. A variedade de soja plantada para este experimento foi a Roundup Ready 2 Xtend (*Glycine max (Linnaeus) Merrill*). A taxa de semeadura foi fixada aproximadamente em 326.000 sementes por hectare, com linhas espaçadas a 0,99 m, após uma cultura anterior de milho.

2.2.2 Estádios de crescimento da soja

O estágio de crescimento da soja durante a coleta de dados foi em diferentes estádios de desenvolvimento vegetativo, no estágio de cotilédone (VC), estágio de primeiro nó (V1) e o estágio de segundo nó (V2), conforme estabelecido por Fehr e Caviness (1977) (Tabela 1). Os estádios selecionados para avaliação no estudo foram escolhidos com base em trabalhos anteriores, abrangendo estádios nos quais a emergência completa das plantas foi alcançada, ao mesmo tempo em que evitava estádios com alta sobreposição de plantas para a identificação individual de plantas (Matias; Caraza-Harter; Endelman, 2020; Randelović *et al.*, 2020; Yang *et al.*, 2024).

O modelo de Fehr e Caviness divide o desenvolvimento da soja em estádios vegetativos (V) e reprodutivos (R). Além dos estádios VE (emergência) e VC (cotilédone), as letras V e R são seguidas por índices numéricos que identificam estádios específicos durante as fases específicas. A fenologia descritiva da soja, voltada para a análise do ciclo de desenvolvimento, permite identificar e classificar os estádios de crescimento, associando-os a exigências específicas ao longo desse ciclo.

Os estádios vegetativos, representados pela letra V, narram o avanço do desenvolvimento vegetativo da soja. Quando os cotilédones estão acima do solo e formam um ângulo de 90° ou mais com seus respectivos hipocótilos, o estágio VE marca a emergência dos cotilédones. O estágio VC indica a fase em que os cotilédones estão expandidos e totalmente abertos. As folhas jovens começam com folíolos cilíndricos que se desenrolam e separam à medida que crescem até formar uma abertura completa. Uma folha é considerada completamente desenvolvida quando está totalmente aberta e os bordos dos folíolos no nó superior não se tocam mais. Os estádios vegetativos do VC são numericamente subdivididos em estádios

sequenciais (V1, V2, V3, V4, V5, V6,... Vn), onde n é o número de nós acima do nó cotiledonar, com folha completamente desenvolvida (Tabela 1).

Tabela 1 - Descrição dos estádios vegetativos da soja

Estádio	Denominação	Descrição
VE	Emergência	Cotilédones acima da superfície do solo
VC	Cotilédone	Cotilédones completamente abertos
V1	Primeiro nó	Folhas unifolioladas completamente desenvolvidas
V2	Segundo nó	Primeira folha trifoliolada completamente desenvolvida
V3	Terceiro nó	Segunda folha trifoliolada completamente desenvolvida
V4	Quarto nó	Terceira folha trifoliolada completamente desenvolvida
V5	Quinto nó	Quarta folha trifoliolada completamente desenvolvida
V6	Sexto nó	Quinta folha trifoliolada completamente desenvolvida
V...
Vn	Enésimo nó	Ante-enésima folha trifoliolada completamente desenvolvida

Fonte: Adaptado de Fehr e Caviness, 1977

2.2.3 RPA e aquisição de dados

Os conjuntos de dados de imagens foram capturados utilizando um RPA (Matrice 300, DJI, Shenzhen, China) equipado com uma câmera zoom RGB de 20 megapixels (MP) DJI Zenmuse H20, capturando imagens de 5184 x 3888 pixels para cada cena. As imagens foram captadas a 15 m acima da superfície, com um GSD (*Ground Sample Distance*) de 0,5 cm e a trajetória do RPA foi definida de modo a garantir uma sobreposição de 80% das imagens frontais e laterais com plano de voo autônomo. As primeiras imagens do RPA e as contagens de plantas no campo foram adquiridas em 29 de maio de 2023, quando a soja estava no estágio VC (um total de 500 imagens). A segunda coleta foi feita 6 dias depois no estágio V1 (com 500

imagens), em 02 de junho de 2023, e a coleta final foi feita após 4 dias no estádio V2 (total de 500 imagens), em 06 de junho de 2023. As imagens foram obtidas entre 9h30 da manhã e 14h30 da tarde, quando não havia nuvens.

Para realizar uma avaliação abrangente da densidade populacional das plantas de soja, foram coletados dados sobre o número de plantas ao longo do experimento, visando uma comparação direta com o método manual de campo. Essas avaliações foram sincronizadas com a aquisição das imagens do RPA, iniciando aos 10 dias após a semeadura, no estádio VC, e continuando até 20 dias após a semeadura, no estádio V2.

2.2.4 Processamento das imagens e contagem da população de plantas

O processamento das imagens foi realizado nos laboratórios de Agricultura de Precisão da Escola de Ciências Vegetais, Ambientais e do Solo da LSU, EUA, e no laboratório de Solo do Departamento de Ciência Vegetal e Arquitetura Paisagística da UConn, EUA.

Iniciando com o processamento das imagens brutas para construção do ortomosaico e delineamento da área de interesse, foi empregado o software Agisoft Metashape 2.1.1 (AGISOFT, 2023) para processamento de três ortomosaicos. Simultaneamente, o software QGIS, versão 3.28.11 (QGIS Development Team, 2023), foi utilizado para delinear o polígono de 60 amostras de cada estádio vegetativo em um arquivo vetorial (*shapefile*) e cada amostra com linha de 1 metro. A extração dos recortes das áreas de interesse foi posteriormente realizada no ambiente do software R, versão 4.3.1, desenvolvido pela equipe Team (2010).

A fase inicial do processamento envolve adicionar todas as imagens no Agisoft Metashape para em seguida o alinhamento de imagens, utilizando informações homólogas de várias imagens sobrepostas e em mais alta acurácia (*highest*). Posteriormente, foi gerado o modelo digital de elevação (DEM) e o ortomosaico com base no DEM. Em uma etapa subsequente, a análise da densidade populacional de plantas foi conduzida utilizando o pacote R *FIELDImageR* na linguagem de programação R. *FIELDImageR* usa pacotes de análise de imagens existentes, como

EImage (Pau, Fuchs, Sklyar, Boutros e Huber, 2010) que permitem a análise de imagens ortomosaicadas em experimentos de campo agrícola.

Durante a preparação das imagens para a extração de dados, uma função foi utilizada para remover elementos indesejados (todos objetos que não fossem a cultura da soja), especialmente o solo, por meio da função `fieldMask`. Índices de Vegetação utilizados em estudos similares (Randelović et al., 2020; Matias; Caraza-Harter; Endelman, 2020) foram testados nesse processo para mitigar possíveis interferências que poderiam comprometer a integridade dos dados extraídos (Tabela 2).

Tabela 2 - Índices utilizados neste estudo. G - em inglês Green, R em inglês RED, B - em inglês Blue

Descrição	Índice	Fórmula	Características relacionadas	Referências
Green leaf index	GLI	$(2 \times G - R - B) / (2 \times G + R + B)$	Cobertura vegetal	Hunt <i>et al.</i> (2013)
Blue Green Pigment Index	BGI	B/G	Clorofila, LAI	Zarco-Tejada <i>et al.</i> (2005)
Overall Hue Index	HUE	$\text{atan}(2 \times (B - G - R) / 30.5 \times (G - R))$	Cor do solo	Escadafal <i>et al.</i> (1994)
Primary Colors Hue Index	HI	$(2 \times R - G - B) / (G - B)$	Cor do solo	Escadafal <i>et al.</i> (1994)
Triangular greenness index	TGI	$G - 0.39 \times R - 0.61 \times B$	Cobertura vegetal	Hunt <i>et al.</i> (2011)

A etapa subsequente envolveu a contagem automática de plantas nas imagens analisadas, realizada pelo pacote *Fieldobject* em linguagem R. A validação dos resultados foi executada por meio da comparação entre a estimativa populacional obtida pelas contagens manual nos recortes dos ortomosaicos e a contagem direta da população de plantas realizada no campo. Essa análise foi repetida para as imagens nos diferentes estádios vegetativos analisados no projeto. Os resultados de estimativa dos modelos foram avaliados por meio do coeficiente de determinação (R^2 ; Equação 1), Erro Médio Absoluto (MAE; Equação 2), para comparar os valores estimados de densidade de plantas em recortes de ortomosaico e os dados coletados em campo de forma manual.

$$R^2 = \frac{\sum(\hat{y}_i - \bar{y})^2}{(y_i - \bar{y})^2} \quad (1)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (2)$$

Onde:

y_i refere-se aos valores observados no campo,

\hat{y}_i representa a estimativa populacional derivada das imagens,

\bar{y} é a média das observações no campo,

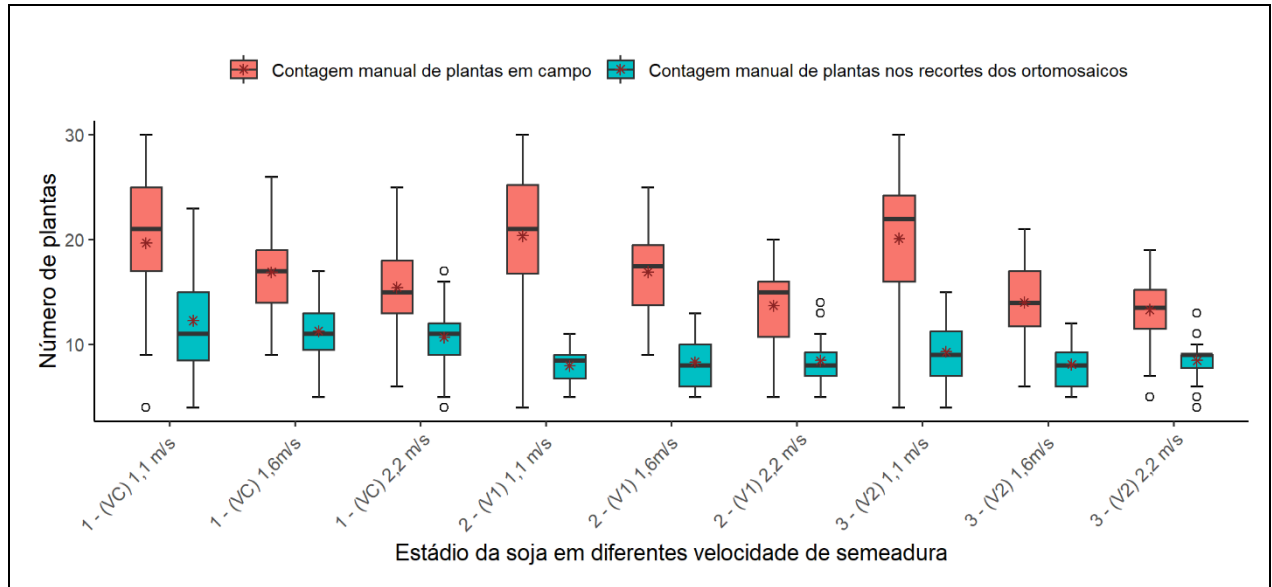
n é o número total de amostras.

2.3 RESULTADOS

2.3.1. Número de plantas contadas manualmente no campo e nos recortes dos ortomosaicos em diferentes estádios de crescimento da soja

Foram obtidos resultados por meio da análise dos recortes dos ortomosaicos distintos para cada combinação das variáveis analisadas (estádio da soja e velocidade de semeadura). As sementes de soja foram semeadas em diferentes velocidades para obter variabilidade nas situações para o treinamento do modelo. Na Figura 2, pode-se observar que, quando a velocidade de semeadura aumenta, o número de plantas no campo é menor, e em todos os estádios a uma velocidade de 1,1 m/s a variação foi entre 4 e 30 plantas em casa amostra, maior do que os estádios em velocidades de 1,6m/s e 2,2 m/s, indicando uma diminuição no número de plantas em velocidades mais altas. Para contar o número de plantas manualmente pelos recortes dos ortomosaicos, apenas as plantas visíveis foram contadas, de modo que o número de plantas teve menor variabilidade em comparação com a quantidade no campo devido à existência de sobreposição na soja (Figura 2).

Figura 2 - Padrão do número de plantas de soja contadas manualmente no campo (laranja) e nos recortes dos ortomosaicos (azul) em diferentes velocidades de semeadura



O número médio de plantas em campo na velocidade de 1,1 m/s foi superior em todos os estádios de crescimento da cultura em comparação com as velocidades de 1,6 m/s e 2,2 m/s. Contudo, a contagem manual nos recortes dos ortomosaicos revelou uma média influenciada pela sobreposição, resultando em uma média menor e mais estável nas diferentes velocidades em comparação com a média no campo, conforme evidenciado nos dados da Tabela 3. Isso se deve ao fato de que, na contagem manual nos recortes dos ortomosaicos, plantas sobrepostas foram consideradas como indivíduos únicos. Quando as plantas estão agrupadas, não é possível identificar visualmente no recorte quantas plantas existem, portanto, são consideradas como um único indivíduo.

A homogeneidade das plantas pode ser afetada por variáveis como a velocidade de deslocamento da semeadora (Kurachi *et al.*, 1989). O excesso de velocidade usado durante a semeadura de soja pode alterar a rota da semente do condutor para o solo. Isso pode fazer com que a semente caia mais tarde ou mais cedo no sulco, ou ambos (Mahl *et al.*, 2004).

Tabela 3 – Média do número de plantas contados manualmente em campo e nos recortes dos ortomosaicos nos estádios VC, V1 e V2 da soja

Estádio da planta	Velocidades de semeadura (m/s)	Média (número de plantas no campo)	Média (número de plantas nos recortes dos ortomosaicos)
VC	1,1	20	12
VC	1,6	17	12
VC	2,2	15	10
V1	1,1	20	8
V1	1,6	16	8
V1	2,2	13	8
V2	1,1	20	9
V2	1,6	14	8
V2	2,2	13	8

2.3.2 Estimativa do número de plantas

A coleta de dados simultânea em diferentes estádios de desenvolvimento proporcionou uma análise abrangente da densidade populacional de plantas de soja, e permitiu correlacionar as informações obtidas pelo método manual no campo e nos recortes dos ortomosaicos com os dados estimados pelo modelo. Essa abordagem integrada contribuiu para uma compreensão mais completa e comparativa da distribuição das plantas ao longo do ciclo de crescimento da cultura.

Para a estimativa da densidade de plantas pelo modelo foram testados Índices de Vegetação, e o que apresentou melhor resultado para contagem de plantas foi o índice GLI (Figura 3), conforme observado na Tabela 4 o MAE foi menor para todos os estádios da soja VC (MAE = 0,057), V1 (MAE = 0,096) e V2 (MAE = 0,195). Esse índice conseguiu remover melhor os ruídos da imagem, como solo, plantas daninhas e algumas palhas que estavam no solo da safra anterior de milho.

Figura 3 – Índice de Vegetação





Recorte de entrada	Índice GLI	Soil mask	Recorte do resultado
			

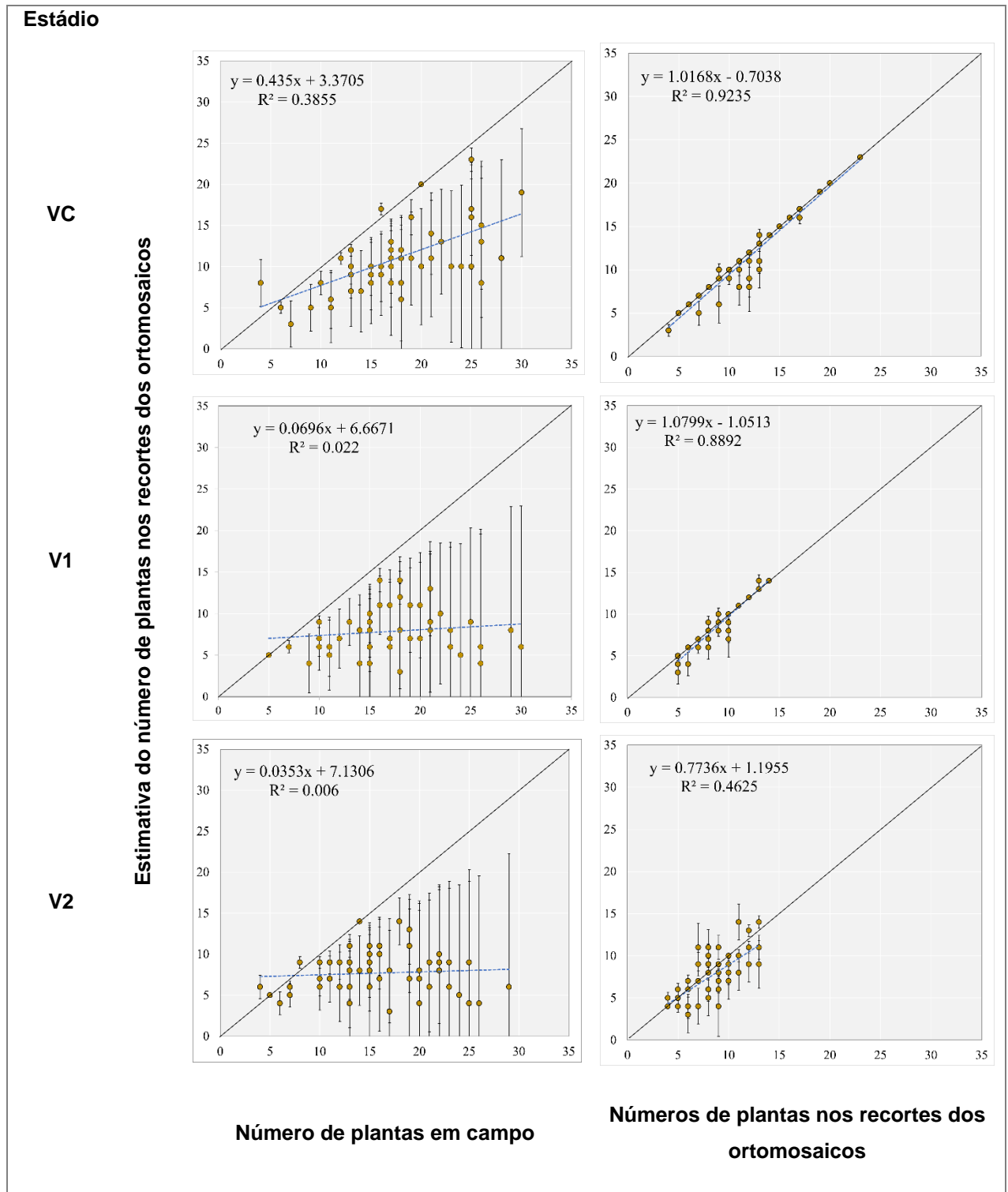
Tabela 4 - Erros Médios Absolutos (MAE) dos índices testados para remoção de ruído das imagens

Descrição	Índices	MAE		
		VC	V1	V2
Green Leaf Index	GLI	0,057	0,096	0,195
Blue Green Pigment Index	BGI	0,273	0,542	0,525
Overall Hue Index*	HUE	0,353	0,282	0,268
Primary Colors Hue Index	HI	0,450	0,542	0,475
Triangular greenness index	TGI	0,479	0,208	0,230

A Figura 4 compara o número estimado de plantas a partir de imagens com o número de plantas contados manualmente no campo e nos recortes dos ortomosaicos com base no estágio vegetativo da soja. Cada gráfico é distribuído por estágio vegetativo nas linhas e colunas representam os números de plantas em duas situações (contadas no campo e nos recortes dos ortomosaicos), sendo que cada gráfico de dispersão fornece uma comparação para cada estimativa de número de plantas. Observa-se que a quanto menor o estágio da planta analisado, melhor a estimativa do número de plantas, e o algoritmo apresentou melhor desempenho comparado ao número de plantas contados manualmente nos recortes dos ortomosaicos (situação em que as plantas sobrepostas não são consideradas, apenas as plantas visíveis individualmente).

A estimativa do número de plantas foi mais precisa para o estágio VC, diminuindo à medida que o estágio de crescimento aumenta. Entretanto, a estimativa da quantidade de plantas foi melhor para o VC quando comparado com o número contados de plantas na imagem ($R^2 = 0,92$). No geral, os resultados fornecem fortes evidências de que o algoritmo apresenta excelente desempenho para estimar o número de plantas de soja por meio de imagens de RPA quando não há sobreposição de plantas.

Figura 4 - Gráficos de dispersão comparando o número de plantas de campo com o número estimado a partir dos recortes dos ortomosaicos. A comparação considera o estágio de crescimento da planta (VC, V1 e V2), a contagem manual no campo e a contagem manual nos recortes dos ortomosaicos



Como visto nos gráficos de dispersão, foi possível estimar os números de plantas a partir dos recortes, quando não apresentar sobreposição de plantas e com melhor desempenho nos estádios VC e V1 ($R^2= 0,92$ e $R^2= 0,88$ respectivamente) exceto no estádio V2 ($R^2= 0,46$). A estimativa da quantidade de plantas seguiu um padrão; à medida que o estádio vegetativo da soja aumentou, houve uma diminuição na precisão da estimativa do número de plantas. Em geral, nas condições em que o estádio vegetativo é menor (VC e V1), a precisão é melhor, indicando que quando as plantas têm a primeira folha trifoliada totalmente desenvolvida, a precisão da estimativa do número de plantas diminui.

Mais detalhes sobre a diferença entre o número de plantas contados manualmente em campo e o número estimado a partir dos recortes para cada condição são apresentados na Tabela 5. O Erro Absoluto Médio (MAE) quanto mais próximo de 0 melhor é a precisão da estimativa e na tabela é possível observar que o erro foi aumentando conforme aumentou o estádio da cultura tanto para a comparação da contagem manual em campo quanto para a realizada manualmente pelos recortes. No estádio VC as duas comparações apresentaram o menor erro, sendo 0,382 no campo e 0,057 no recorte.

Tabela 5 - Erro Absoluto Médio (MAE) entre o número de plantas estimado em comparação com a contagem manual no campo e nos recortes dos ortomosaicos para cada estádio vegetativo

Estádio da cultura	MAE (número de plantas no campo)	MAE (número de plantas nos recortes dos ortomosaicos)
VC	0,382	0,057
V1	0,464	0,096
V2	0,471	0,195

Ao avaliar a precisão da estimativa em condições de diferentes velocidades e estádios de crescimento foi possível constatar que ao aumentar a velocidade de semeadura a precisão diminuiu tanto para a contagem manual em campo (Figura 5)

quanto para a realizada manualmente nos recortes dos ortomosaicos (Figura 6). No estádio VC com velocidade de 1,1 m/s para a contagem manual nos recortes a precisão foi alta ($R^2 = 0,96$) enquanto na velocidade de 2,2 m/s a precisão diminuiu ($R^2 = 0,81$). Conforme aumenta a velocidade de semeadura, a homogeneidade da distribuição entre plantas diminui podendo acarretar o aumento da sobreposição (Kurachi *et al.*, 1989).

Figura 5 - Gráficos de dispersão comparando o número de plantas contadas manualmente no campo com o número estimado a partir dos recortes dos ortomosaicos, com o uso do pacote R FIELDImageR na linguagem de programação R. A comparação considera o estágio de crescimento (VC, V1 e V2) e a velocidade de semeadura da soja (1,1 m/s, 1,6m/s e 2,2 m/s)

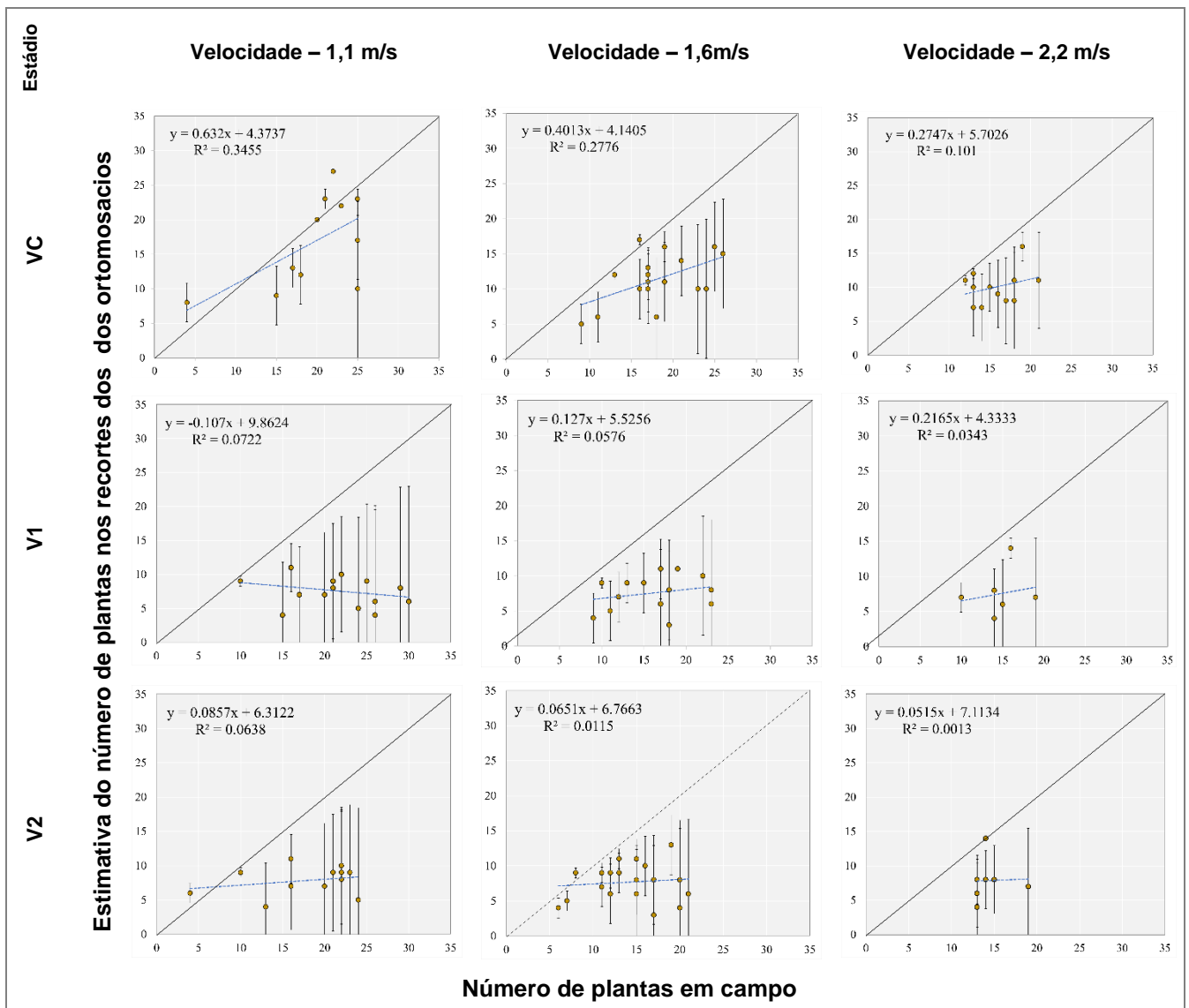
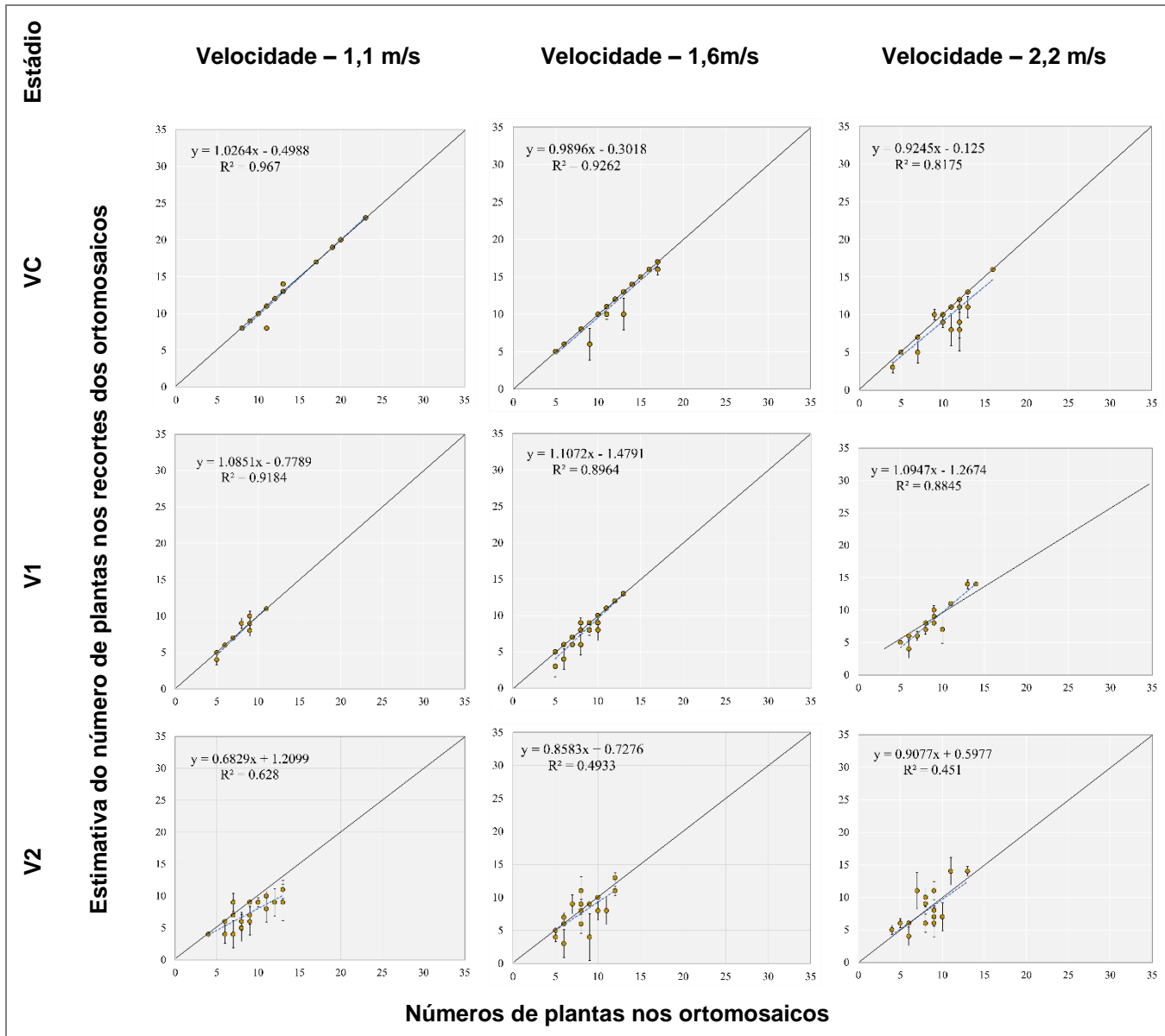


Figura 6 - Gráficos de dispersão comparando o número de plantas contadas manualmente nos recortes dos ortomosaicos com o número estimado dos mesmos recortes, com o uso do pacote R FIELDImageR na linguagem de programação R. A comparação considera o estágio de crescimento (VC, V1 e V2) e a velocidade de semeadura da soja (1,1 m/s, 1,6m/s e 2,2 m/s)



Como demonstrado no gráfico de dispersão na Tabela 6 o MAE aumenta conforme o aumento da velocidade de semeadura e do estágio de crescimento da planta. O MAE para a estimativa comparada ao número de plantas em campo inicia com um erro de 0,333 no estágio VC com velocidade de 1,1 m/s e alcança 0,659 de

erro no V2 com velocidade 2,2 m/s. O aumento do erro também ocorre na comparação com o número de plantas contados manualmente pelos recortes, VC com velocidade de 1,1 m/s apresenta um erro de 0,025 e V2 com velocidade 2,2 m/s o erro aumenta para 0,239.

Tabela 6 – O Erro Absoluto Médio (MAE) entre o número de plantas estimado nos recortes dos ortomosaicos em comparação com o número de plantas contados manualmente no campo e nos recortes dos ortomosaicos para cada estágio vegetativo

Estádio da cultura	Velocidade de semeadura (m/s)	MAE	
		(número de plantas no campo)	(número de plantas nos recortes dos ortomosaicos)
VC	1,1	0,333	0,025
VC	1,6	0,366	0,038
VC	2,2	0,382	0,102
V1	1,1	0,550	0,069
V1	1,6	0,561	0,113
V1	2,2	0,590	0,116
V2	1,1	0,598	0,176
V2	1,6	0,629	0,186
V2	2,2	0,659	0,239

2.4 DISCUSSÕES

Devido a sobreposição e emergência desigual de plantas, obter o número exato de contagem de plantações é um desafio (García-Martínez *et al.*, 2020). Portanto, o melhor estágio de crescimento da planta para se realizar a coleta dos dados precisa ser investigada. Este estudo analisou a influência do estágio fenológico e da velocidade de semeadura no processo de estimativa do número de plantas de soja por meio de recortes dos ortomosaicos. Até o presente momento, a literatura carece de uma abordagem abrangente acerca dos efeitos específicos do estágio de crescimento da soja e da velocidade de semeadura nesse procedimento, destacando a relevância desta investigação (Pathak *et al.*, 2022). A obtenção de resultados precisos, visando minimizar os erros na estimativa do número de plantas, requer a aquisição de imagens em momentos ideais durante o ciclo fenológico da soja e com velocidade de semeadura recomendada para a sua situação no campo (Luna e Lobo, 2016; De Souza *et al.*, 2017).

Em nossa pesquisa, identificamos uma clara influência do estágio fenológico da soja, sendo este um dos fatores preponderantes. À medida que a planta se desenvolve, ocorre um aumento no comprimento e no número de folhas, o que culmina na sobreposição das plantas, impactando diretamente na precisão das estimativas. Tal fenômeno é mais acentuado em estádios avançados de desenvolvimento. Além disso a homogeneidade entre as plantas pode ser prejudicada por fatores como a velocidade de deslocamento da semeadora, podendo aumentar o número de sobreposições (Kurachi *et al.*, 1989). Considerando que os voos foram realizados até o estágio V2 da soja devido aos próximos estádios estarem com alta sobreposição, sendo objetivo contar plantas a sobreposição é um problema para identificação individual de planta de soja. E mesmo assim o trabalho encontrou problemas de sobreposição no primeiro estágio avaliado o que interferiu na acurácia da análise.

O espaçamento inadequado entre plantas de soja pode criar vários desafios que impactam negativamente a produtividade geral. Falhas nas linhas de cultivo podem criar áreas com densidade de semeadura inadequada, condições favoráveis para o crescimento de ervas daninhas, resultando variações no tamanho das plantas,

consequentemente à redução da produtividade (Endres, 1996). Por outro lado, plantas muito próximas competem entre si por recursos essenciais, como luz solar, água e nutrientes do solo. Isso pode levar a um crescimento desigual e resultar em plantas mais fracas e menos produtivas (Fiss *et al.*, 2018).

A determinação precisa da população de plantas a partir de imagens digitais coletadas por Aeronave Remotamente Pilotada (RPA) é comprometida pela oclusão, conforme observado por Keller *et al.* (2018). A interferência do estágio fenológico na estimativa do número de plantas também é um fenômeno observado em estudos de mapeamento (García-Martínez *et al.*, 2020). À medida que as plantas alcançam estádios mais avançados, a sobreposição foliar e o contato entre as plantas reduzem a área visível para a identificação individual. Ao analisar as estimativas para diferentes estádios fenológicos, observou-se que o menor estágio avaliado (VC) apresentou menor percentual de erro, indicando maior precisão. Em contrapartida, em condições desafiadoras, especialmente no estágio vegetativo V2 e com maior sobreposição, os erros foram mais significativos. As estimativas mostraram uma tendência à subestimação do número real de plantas, atribuída principalmente à sobreposição das plantas nesses estádios avançados.

Um aspecto a ser considerado é a relação intrínseca entre o número de plantas e o estágio fenológico ideal. A menor distância entre as plantas em estádios mais avançados pode dificultar a identificação individual, tornando crucial a escolha criteriosa do estágio fenológico para otimizar a precisão (Figura 7). Além disso, destaca-se que a seleção de um estágio fenológico excessivamente precoce pode resultar em menor emergência de plantas, associada a tamanhos reduzidos, o que torna a identificação por imagens mais desafiadora. Em síntese, esta pesquisa contribui para a compreensão crítica da influência do estágio fenológico da soja nas estimativas do número de plantas por meio dos recortes dos ortomosaicos, proporcionando insights valiosos para aprimorar a eficácia das práticas agrícolas e a gestão de cultivos.

Figura 7 – Exemplo de sobreposição de soja

A superestimação do número real de plantas de soja, evidenciada em campo, pode ser atribuída à possível interferência de plantas daninhas, fenômeno destacado em estudos anteriores, como o de Randelović *et al.* (2020). No entanto, em nosso estudo, adotamos uma abordagem que incorporou Índices de Vegetação na modelagem utilizando a função FIELDmask, para contribuir com a redução do impacto das plantas daninhas na precisão da estimativa (Randelović *et al.*, 2020). A aquisição de dados de imagens por meio de um sensor RGB a bordo de um RPA representa uma inovação valiosa com o uso de processamento de imagens (Li *et al.*, 2019; Varela *et al.*, 2018). No entanto, é imperativo considerar os desafios associados a essa abordagem, pois erros podem surgir de diversas fontes, como sobreposições, conforme identificado neste estudo e em estudos anteriores (Yang *et al.*, 2024; Zhao *et al.*, 2018).

A estimativa precisa do número de plantas de soja é fundamental para avaliar a uniformidade de emergência, fator intrinsecamente ligado à produtividade. Além disso, é fundamental reconhecer que as informações derivadas das estimativas do número de plantas podem ter aplicações não apenas na safra atual, mas também no planejamento da semeadura seguinte. Esse entendimento oferece uma perspectiva estratégica para o uso contínuo desses dados ao longo do tempo. Estimar o número de plantas é fundamental para entender os impactos na produtividade da soja.

Embora seja aconselhável avaliar a contagem de plantas nesta fase da soja (VC), é fundamental considerar a população de plantas utilizada. Isso porque quanto maior o número de plantas por metro, mais cedo pode ocorrer a sobreposição, limitando assim a contagem por imagem de RPA. Essas observações reforçam a

necessidade contínua de melhorar as metodologias de estimativa, adaptando-as às complexidades específicas do ambiente de cultivo da soja.

Esta constatação destaca uma limitação significativa deste estudo, que enfrentou desafios na precisão da estimativa quando confrontado com sobreposição de plantas em campo. Consideramos esses resultados como um indicativo de que, embora nossa abordagem seja robusta, existem limitações específicas a serem consideradas, especialmente em distâncias entre plantas menores (população de plantas maiores) que podem ocorrer uma sobreposição em um estágio mais cedo.

A aplicabilidade mais ampla dessa pesquisa ultrapassa o âmbito da soja, sugerindo que a abordagem desenvolvida pode ser adaptada e potencialmente aplicada a outras culturas. Em contextos em que fatores como o número de plantas e o estágio de crescimento da planta influenciam a coleta precisa de informações, nossa metodologia pode oferecer contribuições valiosas. No entanto, ressaltamos a importância de considerar as nuances específicas de cada cultura para otimizar a precisão e relevância das estimativas.

2.5 CONCLUSÕES

Neste estudo, foram investigados os efeitos do estágio fenológico da soja e da velocidade de semeadura na estimativa do número de plantas, utilizando uma abordagem que propôs uma interação entre esses fatores para alcançar uma correlação entre as estimativas obtidas nos recortes dos ortomosaicos e o número real de plantas em campo, assim como os encontrados manualmente nos recortes dos ortomosaicos. Os resultados enfatizam a necessidade crucial de selecionar meticulosamente o estágio fenológico da soja e de definir o momento ideal para a captura das imagens, levando em consideração as condições específicas do ambiente de cultivo, como a velocidade de semeadura.

O estágio VC e a velocidade de 1,1 m/s (aproximadamente 4 km/h) demonstraram superioridade em termos de acurácia devido ao aumento na sobreposição de vegetação à medida que as plantas crescem, resultando em uma redução na precisão da estimativa. Com o aumento da velocidade de semeadura, foi observada uma diminuição na homogeneidade, afetando a cobertura da área e levando a possíveis sobreposições de plantas. Isso resultou em dificuldades na contagem de plantas, especialmente na presença de sobreposição. No entanto, ao

desconsiderar as plantas sobrepostas na contagem manual realizada pelos recortes dos ortomosaicos, a acurácia foi alta no estágio VC e na velocidade de 1,1 m/s. Em resumo, esta pesquisa não apenas contribui para uma compreensão aprofundada da dinâmica entre o estágio fenológico da soja, velocidade de semeadura e estimativa do número de plantas em ortomosaicos, mas também abre caminho para futuras investigações e aplicações práticas em diferentes contextos agrícolas.

REFERÊNCIAS

- AGISOFT. **Agisoft Metashape User Manual Professional Edition**. Version 2.0.3, 2023.
- BERGADA, P., M. *et al.* 2015. **Generating value in the soybean an international royalty collection: chain through study**. Int. Seed Fed.: 1–56.
- BERTELLI, Guilherme Augusto *et al.* Plantability performance of pneumatic seeders in the soybean culture implantation in the Piauí cerrado-Brasil. **Applied Research & Agrotechnology**, v. 9, n. 1, p. 91-103, 2016.
- BORTOLI, Luan Felipe *et al.* Sowing speed can affect distribution and yield of soybean. **Australian Journal of Crop Science**, v. 15, n. 1, p. 16-22, 2021.
- BUTLER, S. **Making the Replant Decision: Utilization of an Aerial Platform to Guide Replant Decisions in Tennessee Cotton**. Ph.D. dissertation. The University of Tennessee, Knoxville: Knoxville, TN 37996, 73 pp, 2019.
- DE BRUIN, J. L.; PEDERSEN, P. Soybean seed yield response to planting date and seeding rate in the Upper Midwest. **Agronomy Journal**, v. 100, n. 3, p. 696-703, 2008.
- DE SOUZA, C. H. W. *et al.* Mapping skips in sugarcane fields using object-based analysis of unmanned aerial vehicle (UAV) images. **Computers and Electronics in Agriculture**, v. 143, p. 49-56, 2017.
- ENDRES, V. C. Espaçamento, densidade e época de semeadura. In: EMBRAPA. Centro de Pesquisa Agropecuária do Oeste (Dourados, MS). **Soja: recomendações técnicas para Mato Grosso do Sul e Mato Grosso**. Dourados, 1996. p. 82-85. (Circular Técnica, 3).
- FAN, Z. *et al.* Automatic tobacco plant detection in UAV images via deep neural networks. **IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing**, v. 11, n. 3, p. 876-887, 2018.
- FEHR, W. R.; CAVINESS, C. E. **Stages of soybean development**. Ames: Iowa State University of Science and Technology, 1977. 11 p. (Special report, 80).
- FENG, A. *et al.* Evaluation of cotton emergence using UAV-based imagery and deep learning. **Computers and Electronics in Agriculture**, v. 177, p. 105711, 2020.
- GARCÍA-MARTÍNEZ, H. *et al.* Digital count of corn plants using images taken by unmanned aerial vehicles and cross correlation of templates. **Agronomy**, v. 10, n. 4, p. 469, 2020.
- GITELSON, Anatoly A. *et al.* Novel algorithms for remote estimation of vegetation fraction. **Remote sensing of Environment**, v. 80, n. 1, p. 76-87, 2002.
- GNÄDINGER, F.; SCHMIDHALTER, U. Digital counts of maize plants by unmanned aerial vehicles (UAVs). **Remote sensing**, v. 9, n. 6, p. 544, 2017.

HADDADI, A.; LEBLON, B.; PATTERSON, G. Detecting and Counting Orchard Trees on Unmanned Aerial Vehicle (UAV)-Based Images Using Entropy and Ndvi Features. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, v. 43, p. 1211-1215, 2020.

JIN, Xiuliang *et al.* Estimates of plant density of wheat crops at emergence from very low altitude UAV imagery. *Remote Sensing of Environment*, v. 198, p. 105-114, 2017.

KELLER, Kevin *et al.* Soybean leaf coverage estimation with machine learning and thresholding algorithms for field phenotyping. In: **Proceedings of the British Machine Vision Conference, Newcastle, UK**. BMVA Press, 2018. p. 3-6.

Li, H.; Wang, P.; Huang, C. Comparison of deep learning methods for detecting and counting sorghum heads in UAV imagery. *Remote Sensing*, v.14, n. 13, p.3143, 2022.

LI, Y., *et al.* Soybean seed counting based on pod image using two-column convolution neural network. *IEEE Access*, v.7, p. 64177-64185, 2019.

LUNA, I.; LOBO, A. Mapping crop planting quality in sugarcane from UAV imagery: A pilot study in Nicaragua. *Remote Sensing*, v. 8, n. 6, p. 500, 2016.

MATIAS, F. I.; CARAZA-HARTER, M. V.; ENDELMAN, J. B. FIELDimageR: An R package to analyze orthomosaic images from agricultural field trials. *The Plant Phenome Journal*, v. 3, n. 1, p. e20005, 2020.

WRIGLEY, C W. *et al.* (Ed.). **Encyclopedia of food grains**. Academic Press, 2015.

KUMAR, J. P.; DOMNIC, S. Image based leaf segmentation and counting in rosette plants. *Information processing in agriculture*, v. 6, n. 2, p. 233-246, 2019.

NEUPANE, B.; HORANONT, T.; HUNG, N. D. Deep learning based banana plant detection and counting using high-resolution red-green-blue (RGB) images collected from unmanned aerial vehicle (UAV). *PloS one*, v. 14, n. 10, p. e0223906, 2019.

OH, S. *et al.* Plant counting of cotton from UAS imagery using deep learning-based object detection framework. *Remote Sensing*, v. 12, n. 18, p. 2981, 2020.

PACHECO, Edson Patto *et al.* Avaliação de uma semeadora-adubadora de precisão. *Pesquisa Agropecuária Brasileira*, v. 31, n. 3, p. 209-214, 1996.

PANG, Y. *et al.* Improved crop row detection with deep neural network for early-season maize stand count in UAV imagery. *Computers and Electronics in Agriculture*, v. 178, p. 105766, 2020.

PATHAK, H. *et al.* A review of unmanned aerial vehicle-based methods for plant stand count evaluation in row crops. *Computers and Electronics in Agriculture*, v. 198, p. 107064, 2022.

QGIS DEVELOPMENT TEAM. QGIS **Geographic Information System**. Versão 3.28.11- "Firenze". 2023. Disponível em: <http://qgis.osgeo.org>. Acesso em: 11 out. 2023.

RANDELOVIĆ, P. *et al.* Prediction of soybean plant density using a machine learning model and vegetation indices extracted from RGB images taken with a UAV. **Agronomy**, v. 10, n. 8, p. 1108, 2020.

TEAM, R. Development Core. R: **A language and environment for statistical computing**, 2010.

RIBERA, J. *et al.* Counting plants using deep learning. In: **2017 IEEE global conference on signal and information processing (GlobalSIP)**. IEEE, 2017. p. 1344-1348.

ROBINSON, A. P. *et al.* Analysis of high yielding, early-planted soybean in Indiana. **Agronomy Journal**, v. 101, n. 1, p. 131-139, 2009. Shirzadifar, A., M. Maharlooei, S.G. Bajwa, P.G. Oduor, and J.F. Nowatzki. 2020. Mapping crop stand count and planting uniformity using high resolution imagery in a maize crop. *Biosyst. Eng.* 200: 377–390. doi: 10.1016/j.biosystemseng.2020.10.013.

VOSENKEMPER, Jacob P. *et al.* Early planting, full-season cultivars, and seed treatments maximize soybean yield potential. **Crop, Forage & Turfgrass Management**, v. 1, n. 1, p. 1-9, 2015.

USDA. **National Agriculture Statistics Services Quick Stats**. 2023a. USDA-NASS, Washington D.C. Disponível em: https://quickstats.nass.usda.gov/results/27B99EF0-5739-34F3-9263-C4320EC2552D?pivot=short_desc. Acesso em: 11 nov. 2023.

USDA. **National Agriculture Recent Cost and Returns**. 2023b. USDA-NASS, Washington D.C. Disponível em: <https://www.ers.usda.gov/data-products/commodity-costs-and-returns/commodity-costs-and-returns/#Historical%20Costs%20and%20Returns:%20Soybeans>. Acesso em: 11 nov. 2023.

VARELA, Sebastian *et al.* Early-season stand count determination in corn via integration of imagery from unmanned aerial systems (UAS) and supervised learning techniques. **Remote Sensing**, v. 10, n. 2, p. 343, 2018.

VALENTE, J. *et al.* Automated crop plant counting from very high-resolution aerial imagery. **Precision Agriculture**, v. 21, p. 1366-1384, 2020.

WEI, M.C.F.; MOLIN, J.P. Soybean yield estimation and its components: A linear regression approach. **Agriculture**, v. 10, n. 8, p.348, 2020.

YANG, Z. *et al.* Fast and Precise Detection of Dense Soybean Seedlings Images Based on Airborne Edge Device. **Agriculture**. <https://doi.org/10.3390/agriculture14020208>. v.14, p. 208, 2024.

ZHANG, J. *et al.* Rapeseed stand count estimation at leaf development stages with UAV imagery and convolutional neural networks. **Frontiers in plant science**, v. 11, p. 617, 2020.

ZHAO, B. *et al.* Rapeseed seedling stand counting and seeding performance evaluation at two early growth stages based on unmanned aerial vehicle imagery. **Frontiers in Plant Science**, v. 9, p. 1362, 2018.

CAPÍTULO 3

CONTAGEM DE PLANTAS DE SOJA BASEADO EM IMAGENS MULTISSENSOR

Autora: Flávia Luize Pereira de Souza

RESUMO

A agricultura de precisão (AP) tem se destacado como uma abordagem fundamental na gestão agrícola contemporânea, visando maximizar a eficiência no uso de recursos e aprimorar a produtividade das culturas. A transição para a chamada "agricultura 4.0" representa uma revolução na forma como a tecnologia é aplicada no campo, com ênfase em soluções digitais e automatizadas, como os RPAs (Aeronaves Remotamente Pilotadas). Esses dispositivos oferecem novas capacidades na captura de imagens de alta resolução, possibilitando uma análise detalhada das variáveis agronômicas em nível de talhão. Este estudo concentrou-se na avaliação da precisão da contagem de plantas de soja (*Glycine max* (Linnaeus) Merrill) em diferentes estádios de desenvolvimento e velocidades de semeadura, utilizando imagens obtidas por RPAs equipados com sensores RGB e multiespectrais. O estudo foi desenvolvido na Ben Hur Research Farm, localizada nos Estados Unidos no município de Baton Rouge, LA, para avaliar o desempenho dos sensores em condições reais, e determinar o sensor mais eficiente para a contagem precisa de plantas de soja. Por meio de análises avançadas, incluindo Redes Neurais e processamento de imagens com a Linguagem R e o pacote FIELDimageR, este estudo identificou a tecnologia RGB como altamente precisa na classificação do número de plantas de soja. No entanto, o processamento de imagens apresentou desafios em ambientes com alta sobreposição de plantas, resultando em uma precisão reduzida na estimativa do número de plantas. Esses resultados destacam a importância da seleção adequada do sensor e do momento da captura das imagens para obter estimativas precisas de contagem de plantas em diferentes condições agronômicas. Além disso, sugerem a necessidade de explorar novas técnicas e abordagens para melhorar a precisão do processamento de imagens em ambientes desafiadores, como áreas com alta densidade populacional de plantas. Essas descobertas têm implicações significativas para a prática da agricultura de precisão, fornecendo insights valiosos para o manejo eficiente de culturas, a otimização do uso de recursos e o aumento da produtividade

nas lavouras de soja. O estudo contribui para o avanço do conhecimento científico nesse campo em constante evolução, abrindo caminho para futuras pesquisas e aplicações práticas na agricultura moderna.

Palavras-chave: densidade de plantas; inteligência artificial; machine learning; multiespectral; RGB.

ABSTRACT

Precision agriculture (PA) has emerged as a fundamental approach in contemporary agricultural management, aimed at maximizing efficiency in the use of resources and improving crop productivity. The transition to so-called "agriculture 4.0" represents a revolution in the way technology is applied in the field, with an emphasis on digital and automated solutions such as UAVs (Unmanned Aerial Vehicles). These devices offer new capabilities for capturing high-resolution images, enabling detailed analysis of agronomic variables at plot level. This study focused on evaluating the accuracy of counting soybean plants (*Glycine max* (Linnaeus) Merrill) at different stages of development and sowing speeds, using images obtained by RPAs equipped with RGB and multispectral sensors. The project was carried out at Ben Hur Research Farm, located in the United States in Baton Rouge, LA, to evaluate the performance of the sensors under real conditions, and determine the most efficient sensor for accurately counting soybean plants. Through advanced analysis, including Neural Networks and image processing with R Language and the FIELDimageR package, this study identified RGB technology as highly accurate in classifying the number of soybean plants. However, image processing presented challenges in environments with high plant overlap, resulting in reduced accuracy in estimating the number of plants. These results highlight the importance of proper sensor selection and timing of image capture to obtain accurate plant count estimates in different agronomic conditions. Furthermore, they suggest the need to explore new techniques and approaches to improve the accuracy of image processing in challenging environments, such as areas with high plant population density. These findings have significant implications for the practice of precision agriculture, providing valuable insights for efficient crop management, optimizing the use of resources and increasing productivity in soybean crops. The study contributes to the advancement of scientific knowledge in this constantly evolving field, paving the way for future research and practical applications in modern agriculture.

Keywords: plant density; artificial intelligence; machine learning; multispectral; RGB.

3.1 INTRODUÇÃO

A Sociedade Brasileira de Ciência do Solo (SBCS) define a agricultura de precisão (AP) como uma abordagem agronômica que utiliza tecnologias de informação e comunicação para avaliar, monitorar e gerenciar de forma eficiente as variáveis nos talhões. O objetivo é otimizar o uso dos recursos e maximizar a produtividade e a qualidade das culturas (SBCS, 2021, p. 1). A agricultura passou de 1.0 (tradicional) para 4.0 (nova era) e continuará evoluindo no futuro. O termo "agricultura 4.0" refere-se a um modelo de produção agrícola em que tecnologias digitais, como inteligência artificial, RPAs (Aeronaves Remotamente Pilotadas), sistemas de informação geográfica e sensores, são usadas para monitorar e controlar os processos agrícolas (Araújo *et al.*, 2021).

A implementação de práticas de AP permitiu que os agricultores reduzissem os custos, o cultivo aplicado, a mão de obra necessária e as horas de maquinários gastos, bem como a quantidade de sementes utilizadas (Wrigley, 2015). Uma variedade de parâmetros agrícolas pode ser monitorados para melhorar o rendimento das colheitas, reduzir custos e otimizar os insumos agrícolas de acordo com diferentes condições ambientais, obter uma melhor compreensão do estado de crescimento das culturas e otimizar o manejo de nutrientes, pragas, ervas daninhas e falhas de semeadura (Nukala *et al.*, 2016).

A obtenção de informações provenientes de sensoriamento remoto tem sido uma prática consolidada na AP, sendo tradicionalmente realizada por meio de imagens obtidas por satélites e aeronaves tripuladas. Recentemente, as Aeronaves Remotamente Pilotadas (RPAs), equipados com câmeras digitais, têm emergido como uma ferramenta proeminente na gestão da agricultura de precisão (Maes e Steppe, 2019; Valente *et al.*, 2020; Yang, 2020), especialmente na fenotipagem de culturas (Torres-Sanchez, 2014; Vogel *et al.*, 2021; Li *et al.*, 2019; Moreira *et al.*, 2019) e proporciona a avaliação por meio de técnicas de extração de características baseadas em processamento digital de imagens.

Após o estabelecimento da cultura de soja, recomenda-se que os produtores realizem avaliações dos povoamentos para assegurar a presença de uma população de plantas adequada, capaz de garantir um rendimento de grãos lucrativo na colheita.

Além de verificar a qualidade das sementes e a eficiência da semeadura (Pathak *et al.*, 2022), essa avaliação, conhecida como "contagem de stand", permite uma detecção precoce de emergências abaixo do ideal, possibilitando uma resposta rápida por meio de replantio (Staton, 2013; Carver *et al.*, 2018). Recomenda-se realizar essas avaliações nos estádios iniciais de crescimento, do VC ao V3 (Licht, 2020). Essa gestão decisiva é crucial para proporcionar tempo suficiente ao desenvolvimento da soja replantada ou semeada em excesso, contribuindo assim para o rendimento geral de grãos.

O potencial de rendimento de grãos da soja é comprometido com o atraso na data de semeadura (Egli; Cornelius, 2009; Salmeron *et al.*, 2014; Nleya *et al.*, 2020). Se uma população mínima aceitável de plantas for atingida, nenhuma ação será necessária. No entanto, se a população for insuficiente para maximizar o rendimento de grãos, o replantio pode ser considerado (Olechowski, 1983; Pathak *et al.*, 2022), desde que a diferença no rendimento entre o povoamento inicial e o replantado compense os custos adicionais associados ao replantio (sementes, combustível, mão de obra e tempo).

Diversos estudos indicaram que a população ótima de plantas agrônômicas (AOPP) na soja é aproximadamente 247.000 plantas por hectare (Epler; Staggenborg, 2008; Gaspar; Conley, 2015; Lee *et al.*, 2008; Gaspar *et al.*, 2020). A população para iniciar o replantio deve ser inferior à AOPP, pois é necessário considerar se a perda de receitas devido à diminuição do rendimento de cereais de uma cultura mal estabelecida é mais onerosa ou não do que o replantio. Em Kentucky, Lee e Herbek (2011) sugeriram que populações de 124.000 plantas por hectare ou mais não justificam replantio para a soja de safra completa com data de semeadura oportuna, devido à perda relativamente baixa de rendimento de grãos e aos custos adicionais associados ao replantio (herbicidas, combustível, sementes, mão de obra).

Os métodos convencionais de avaliação de povoamentos de soja envolvem a visita de um indivíduo a vários locais representativos dentro de um campo, contando manualmente o número de plantas vivas de soja em uma área específica e, em seguida, dimensionando a contagem para plantas por hectare. A fita métrica e o método do bambolê são exemplos comuns dessas estratégias (Whigham *et al.*, 2000; Robinson; Conley, 2007). Embora essas técnicas de medição possam oferecer estimativas precisas do desempenho da população de plantas, apresentam potenciais desvantagens (Pathak *et al.*, 2022), incluindo a possibilidade de áreas exploradas não

serem representativas da população real de plantas em todo o campo. O erro humano também pode resultar em contagens imprecisas de plantas, e a combinação desses fatores pode induzir a decisões equivocadas de gestão, como supersemeadura ou replantio excessivo.

Em conjunto com os métodos tradicionais, o cenário contemporâneo da avaliação de povoamentos vegetais destaca o advento de abordagens inovadoras, utilizando sensores e técnicas de processamento digital. Estas novas metodologias buscam proporcionar medições mais precisas, minimizando a subjetividade inerente às práticas manuais. Por exemplo, Shi *et al.* (2013) empregaram um carrinho movido a energia humana para contar talos de milho (*Zea mays* L.) utilizando a tecnologia LIDAR e varredura de linha laser. Outras técnicas terrestres para a avaliação de plantações de milho incluíram a coleta e o processamento de imagens digitais de alta resolução (Jia; Krutz, 1992; Shrestha; Steward, 2003; Tang; Tian, 2008). A aplicação de tecnologias emergentes, como o uso de smartphones, também tem sido explorada para monitorar o desenvolvimento da copa da soja. Smith *et al.* (2019) utilizaram o aplicativo Canopeo (Patrignani; Ochsner, 2015) para rastrear o crescimento da copa da soja plantada em fileiras duplas, sujeitas a erros de semeadura. Essa abordagem inovadora não apenas diversifica os métodos de avaliação, mas também introduz uma dimensão prática, uma vez que o uso de smartphones permite uma análise em tempo real (Colorado *et al.*, 2020).

Em particular, a ascensão dos RPAs revolucionou a capacidade de obtenção de imagens de alta resolução espacial e espectral a baixa altitude. Essa flexibilidade permite a análise e gestão de campos agrícolas em menor escala, proporcionando uma alternativa eficaz às imagens de satélite ou aeronaves tripuladas (Hunt *et al.*, 2005). Esses RPAs são frequentemente equipados com sensores, incluindo câmeras com sensor Vermelho, Verde e Azul (RGB) e multiespectrais. Enquanto as câmeras RGB capturam dados no espectro visível, as multiespectrais registram informações em uma gama mais ampla do espectro eletromagnético, incluindo bandas além do visível, como infravermelho próximo e médio (Prakash, 2000).

Essa variedade de sensores possibilita a aplicação de diversos Índices de Vegetação, como o Índice de Vegetação por Diferença Normalizada (NDVI), que é amplamente utilizado na agricultura para avaliar a saúde e o desenvolvimento das

culturas (Rouse *et al.*, 1974). Além disso, índices derivados de imagens RGB, apresentam-se como ferramentas valiosas para o monitoramento de vegetação, com aplicabilidade prática na segmentação de imagens e contagens de povoamentos no início da safra (Woebbecke *et al.*, 1995; Vong *et al.*, 2021; Fan *et al.*, 2018). A integração desses métodos avançados representa não apenas um avanço tecnológico, mas também uma evolução prática na avaliação de povoamentos de soja.

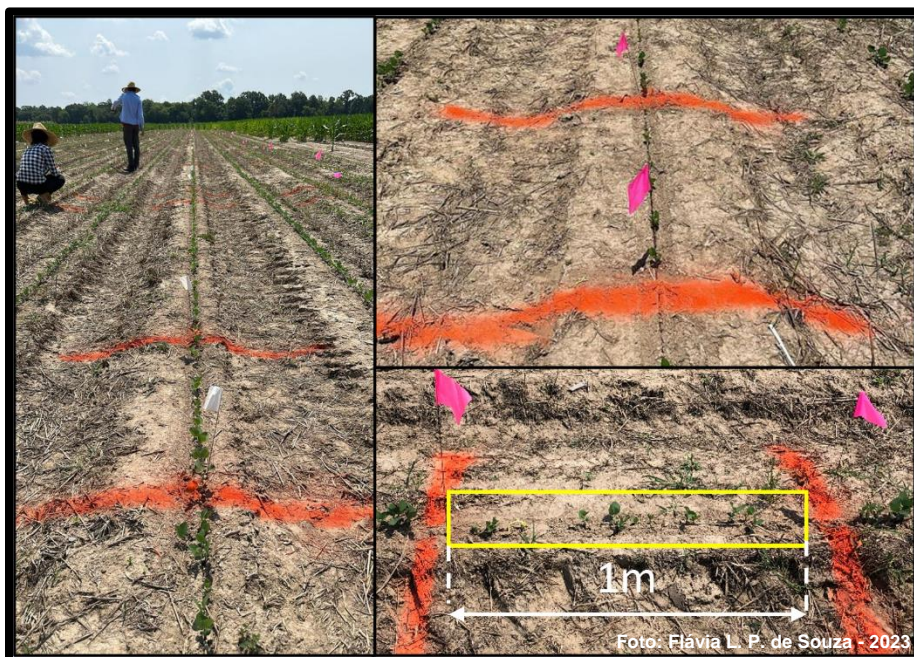
Em decorrência desse contexto, acreditamos que o sensor de captura de imagens ideal resultará na estimativa do número de plantas preciso. O propósito deste artigo é investigar e comparar a precisão da contagem de plantas de soja, empregando imagens RGB e multiespectrais. O estudo busca oferecer uma análise da eficácia desses métodos, identificando suas potenciais contribuições para a otimização do manejo agrícola e aprimoramento da produtividade nas lavouras de soja.

3.2 MATERIAL E MÉTODOS

3.2.1 Área de estudo e projeto experimental

O experimento teve início em 19 de maio de 2023, na Ben Hur Research Farm, localizada no município de Baton Rouge, LA. A escolha estratégica deste local visava atender ao propósito central de determinar o sensor mais eficiente para a contagem precisa de plantas jovens de soja (*Glycine max (Linnaeus) Merrill*) Roundup Ready 2 Xtend. Essa localização específica foi selecionada como parte de um estudo abrangente que objetivava não apenas identificar o estágio fenológico ideal para a contagem populacional de plantas de soja, mas também para garantir precisão na estimativa da população da cultura.

A área experimental foi planejada, abrangendo dimensões totais de 90 x 15 metros e compreendendo 12 linhas de semeadura, totalizando uma área de 0,13 hectares. Para assegurar a viabilidade do experimento, a taxa de semeadura foi estabelecida em 326.000 sementes por hectare, com espaçamento entre linhas de 0,99 m, e a semeadura foi realizada em três velocidades diferentes, 1,1 m/s (aproximadamente 4km/h), 1,6m/s (aproximadamente 6 km/h) e 2,2 m/s (aproximadamente 8 km/h). Foram demarcadas em campo 60 amostras de 1 metro cada, com tinta spray e bandeiras para a melhor visualização conforme a Figura 1.

Figura 1 – Delimitação das amostras em campo



3.2.2 Aquisição das imagens

As imagens foram obtidas, em práticas de levantamentos de voos em campo, por meio de um RPA do tipo multirrotor (Quadrirrotor), modelo DJI Matrice 300 (Figura 2) embarcado com uma câmera DJI Zenmuse H20 e Micasense Red-Edge-MX. Para a aquisição e processamento de imagens nas distintas bandas espectrais, empregou-se uma abordagem visando capturar dados abrangentes e de alta resolução (Figura 3). Os sensores acoplados para aquisição das imagens possuem resolução de 5184 x 3888 (RGB) e 1280 x 960 (multiespectral) (Figura 3). Eles são projetados para capturar segmentos específicos do espectro eletromagnético, constituindo segmentos infravermelhos e visíveis. Este estudo se concentrou nos sensores de imagem mais utilizados e em suas aplicações agrícolas devido à grande variedade de sensores remotos disponíveis (Zhu, 2018).

Figura 2 - RPA DJI Matrice 300 utilizado no estudo



Figura 3 - Tipos de sensores utilizados

Tipo	Espectro eletromagnético	
	DJI Zenmuse H20 - Zoom	Luz visível
	Micasense Red-Edge-MX	Luz visível, perto do infravermelho

A coleta de dados de sensoriamento remoto foi realizada por meio de RPA DJI Matrice 300 em três estádios da soja, incluindo o estádio de cotilédone (VC), o estádio de primeiro nó (V1) e o estádio de segundo nó (V2), conforme estabelecido por Fehr e Caviness (1977). Dois sensores foram acoplados para aquisição das imagens, um sensor possui resolução de 5184 x 3888, com 20 megapixels câmera zoom e comprimento focal de 24 mm (DJI Zenmuse H20) e o outro sensor com resolução de 1280 x 960, com 6 megapixels e comprimento focal de 5,4 mm (Micasense Red-Edge-MX). Foi necessário um (1) voo para cada estádio da soja e para ambos os sensores, totalizando seis (6) voos. Especificações detalhadas dos sensores utilizados para a coleta de dados são fornecidas na Tabela 1.

Tabela 1 – Especificações dos sensores e informações da aquisição de imagens RGB e multiespectral utilizados para este estudo

Especificação	DJI Zenmuse H20 (RGB)	Micasense Red-Edge- MX (Multiespectral)
Resolução do sensor	5184 x 3888 pixels	1280 x 960 pixels
Resolução Espacial a 15 metros	0,5 cm	1 cm
Banda Espectral	8 – 14 μ m	
Comprimento focal	24 mm	
Resolução Espectral	3 bandas	5 bandas
Blue		475nm
Green		560nm
Red		668nm
Red edge		717 nm
NIR (Near infrared)		842 nm
Tamanho do sensor (mm)	150x114x151 mm	4,8 mm x 3,6 mm
Número de imagens coletadas	VC – 500 imagens V1 – 500 imagens V2 – 500 imagens	VC – 925 imagens V1 – 925 imagens V2 – 925 imagens
Data		VC - 29/05/2023 V1 - 02/06/2023 V2 - 06/06/2023
Horário da coleta de imagens	Entre 9h30 da manhã e 14h30 da tarde	

Os planejamentos dos voos autônomos foram feitos com os dois tipos distintos de sensores (RGB e multiespectral), onde foi configurado e delimitado a área de sobrevoo de forma a garantir que fosse recoberto toda a área de estudos, o RPA foi empregado a uma altitude de voo de 15 metros e sobreposição de 80% frontal e lateral, garantindo assim um GSD (*Ground Sample Distance*) de 0,5 cm para as imagens RGB e 1,0 cm para imagens multiespectral.

A coleta de dados foi executada visando avaliar a eficiência de diversos sensores na contagem precisa de plantas de soja. Para garantir a representatividade estatística e a abrangência na análise, foi realizado o levantamento de 60 amostras cada uma abrangendo uma área de 1 m² e foram realizadas contagens manuais em campo das 60 amostras e a contagem manual diretamente nos recortes dos

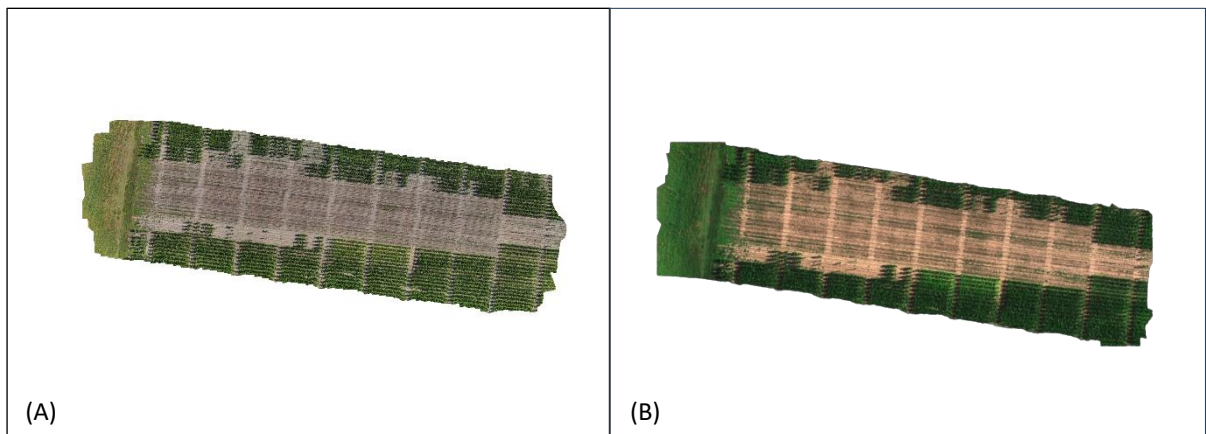
ortomosaicos, devido a possibilidade de sobreposição de plantas interferir na precisão da estimativa realizada pelo pacote FIELDImageR e fieldObject em linguagem R. Essa abordagem manual em campo proporcionou uma verdade terrestre para as tecnologias empregadas e a abordagem manual nos recortes permitiu uma verdade nas imagens capturadas.

3.2.3 Processamento de dados

Os dados foram processados no laboratório de Agricultura de Precisão da Escola de Ciências Vegetais, Ambientais e do Solo da LSU, EUA, e no laboratório de Solo do Departamento de Ciência Vegetal e Arquitetura Paisagística da UConn, EUA.

As imagens multiespectrais e RGB foram usadas para criar ortomosaicos da área de pesquisa com o software Agisoft Metashape (Agisoft Metashape Professional 2.0.4, Agisoft, St. Petersburg, Russian) para gerar 6 ortomosaicos, sendo 3 ortomosaicos RGB e 3 multiespectrais, onde foi utilizado as 3 bandas da RGB e as 5 bandas da multiespectral para a composição (Figura 4).

Figura 4 - Ortomosaico RGB (A); Ortomosaico multiespectral (B)



As etapas usadas para processar os ortomosaicos RGB foram: adicionar imagens (*Add Photos*); alinhar imagens (*Align Photos*) com acurácia alta (*highest*); construir Modelo de Elevação Digital (*Build DEM*) e construir ortomosaico (*Build Ortomosaic*) pelo DEM. Enquanto para os ortomosaicos multiespectrais foram utilizados os seguintes passos: adicionar imagens (*multi-camera system*); calibrar refletância (*calibrate reflectance*); alinhar imagens em alta acurácia; criar DEM e o ortomosaico com o DEM.

As amostras foram recortadas após a criação do ortomosaico, resultando em um total de 360 recortes obtidos de todos os sensores. Para a análise utilizando Redes Neurais, foram selecionadas 220 imagens. Dessas, 110 foram provenientes de cada sensor de captura, cobrindo todos os estádios da soja em todas as classes, e 10 imagens foram utilizadas em cada classe de número de plantas, todas as imagens acompanhadas por uma marcação manual do número de plantas feita no campo.

Inicialmente para análise das imagens capturada pelos sensores (RGB e multiespectral) foi empregado a Rede Neural para classificar o número de plantas em soja (Haykin, 2009). O Perceptron Multicamadas é uma Rede Neural composta por uma ou mais camadas ocultas e um número indeterminado de neurônios. O processo de classificação utilizou widgets do software Orange, como "*import images*", "*image views*" e "*image embedding*" para a conversão de cada pixel da imagem em um número binário com o uso de redes neurais pré-treinadas chamada Inception v3. A Rede Neural foi avaliada por meio da validação cruzada e configurada com as seguintes características: função de ativação Logística, 100 neurônios por camada oculta (com 2048 neurônios na camada de entrada e 11 na camada de saída), um fator de regularização de 5 e um máximo de 500 iterações (Demšar *et al.*, 2013).

Na etapa subsequente, os sensores foram submetidos à análise adicional por meio da técnica de processamento de imagens utilizando a linguagem R e o pacote FIELDImageR. O ambiente de programação R foi utilizado para manipular os dados obtidos a partir do sensoriamento remoto realizado pelo RPA (Aeronave Pilotada Remotamente). Esta ferramenta possibilitou uma análise minuciosa das imagens capturadas, permitindo a extração de informações fundamentais para compreender as condições das plantas de soja em diferentes faixas espectrais.

Para a contagem de plantas, o pacote FIELDImage R (Matias; Caraza-Harter; Endelman, 2020), se mostrou fundamental, foi utilizado para realizar uma contagem das plantas em cada uma das 60 amostras previamente demarcadas (60 amostras de cada sensor, em três estádios da soja, um total de 360 imagens). Essa etapa foi para avaliar a distribuição populacional das plantas e extrair dados significativos para a análise. FIELDImageR é baseado em pacotes de análise de imagens existentes, como *raster* (Hijmans, 2019), *sp* (Bivand, Pebesma, Gómez-Rubio, 2013) e *EImage* (Pau,

Fuchs, Sklyar, Boutros e Huber, 2010) para análise de imagens ortomosaicas de experimentos de campo agrícola.

Durante o processamento das imagens dos sensores RGB e multiespectral, uma função foi empregada para remover o solo por meio da função fieldMask e foram incluídos Índices de Vegetação no processo para a remoção do solo (consulte Tabela 2).

Tabela 2 - Índices utilizados neste estudo. G - em inglês Green, R - em inglês RED, B - em inglês Blue, NIR - em inglês Near Infrared.

Sensor	Descrição	Índice	Formúla	Características relacionadas	Referências
RGB	Green leaf index	GLI	$(2 \times G - R - B) / (2 \times G + R + B)$	Vegetation coverage	Hunt <i>et al.</i> (2013)
RGB	Blue Green Pigment Index	BGI	B/G	Chlorophyll, LAI	Zarco-Tejada <i>et al.</i> (2005)
RGB	Overall Hue Index	HUE	$\text{atan}(2 \times (B-G-R) / 30.5 \times (G-R))$	Soil color	Escadafal <i>et al.</i> (1994)
RGB	Primary Colors Hue Index	HI	$(2 \times R - G - B) / (G - B)$	Soil color	Escadafal <i>et al.</i> (1994)
RGB	Triangular greenness index	TGI	$G - 0.39 \times R - 0.61 \times B$	Vegetation coverage	Hunt <i>et al.</i> (2011)
Multiespectral	Normalized Difference Vegetation Index	NDVI	$(\text{NIR} - \text{Red}) / (\text{NIR} + \text{Red})$	Vegetation coverage	(Rouse <i>et al.</i> , 1974)
Multiespectral	Normalized Difference Red Edge Index	NDRE	$(\text{NIR} - \text{RedEdge}) / (\text{NIR} + \text{RedEdge})$	Chlorophyll	(Gitelson; Merzlyak, 1994)
Multiespectral	Chlorophyll index — green	CIG	$(\text{NIR}/G) - 1$	Chlorophyll	(Gitelson <i>et al.</i> , 2003)

3.2.4. Estudo de validação

A AUC (Área Sob a Curva) na Equação 1, Precisão na Equação 2, Acurácia (CA) na Equação 3, *Recall* conforme na Equação 4 e Pontuação F1 na Equação 5 são as principais métricas usadas para avaliar o desempenho do modelo da Rede Neural (Demšar *et al.*, 2013). O número de plantas classificadas pelo modelo foi comparado com o número de plantas encontradas no campo. Além disso, as imagens foram

replicadas nos espectros RGB (3 bandas - *Red*, *Green* e *Blue*) e multiespectral (5 bandas – *Red*, *Green*, *Blue*, *NIR* e *RedEdge*).

(1)

$$AUC \approx \sum_{i=1}^{n-1} \frac{(y_i + y_{i+1})}{2} \times (x_{i+1} - x_i)$$

(2)

$$\text{Precisão} = \frac{VP}{VP + FP}$$

(3)

$$\text{Acurácia} = \frac{VP + VN}{VP + VN + FP + FN}$$

(4)

$$\text{Recall} = \frac{VP}{VP + FN}$$

(5)

$$F1 = 2 \times \frac{\text{Precisão} \times \text{Recall}}{\text{Precisão} + \text{Recall}}$$

Onde:

x_i = Taxa de falsos positivos;

y_i = Taxa de verdadeiros positivos;

VP = Verdadeiros positivos;

FP = Falsos positivos;

FN = Falsos negativos.

A validação dos resultados do processamento de imagens com FIELDimageR em linguagem R foi executada por meio da comparação entre a estimativa populacional obtida pelas imagens de RPAs e a contagem direta da população de plantas realizada no campo e manualmente nos recortes dos ortomosaicos. Essa análise foi repetida para as imagens nos espectros RGB (3 bandas – *Red*, *Green* e *Blue*) e multiespectral (5 bandas – *Blue*, *Green*, *Red*, *Red edge* e *NIR*). Utilizamos o coeficiente de determinação (R^2 ; Equação 7) para avaliar a precisão e o Erro Absoluto

Médio (MAE; Equação 8) como métrica para avaliar a acurácia. Essas métricas compararam os valores de cada estimativa populacional, derivados das imagens, com os dados obtidos diretamente no campo e pelas imagens.

$$R^2 = \frac{\sum(\hat{y}_i - \bar{y})^2}{(y_i - \bar{y})^2} \quad (7)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (8)$$

Onde:

y_i refere-se aos valores observados no campo,

\hat{y}_i representa a estimativa populacional derivada das imagens,

\bar{y} é a média das observações no campo,

n é o número total de amostras.

3.3 RESULTADOS

3.3.1 Classificação do número de plantas com Rede Neural comparando as imagens adquiridas pelos diferentes sensores.

Os resultados de avaliação com a Rede Neural para dois sensores distintos estão apresentados na Tabela 3. A Rede Neural exibe capacidade de classificação para imagens RGB e multiespectral, como evidenciado por sua pontuação AUC de 0,969 (RGB) e 0,942 (multiespectral), indicando uma forte capacidade em discriminar as diferentes classes. O modelo para a RGB tem uma acurácia de 0,8, indicando que acerta 80% das predições e 71% (0,718) para multiespectrais. Além disso, o modelo com imagens RGB alcança um F1 de 0,803 e 0,713 para multiespectral, indicando um bom equilíbrio entre precisão e recall. A precisão do modelo atingiu 82,5% (RGB) e 71,8% (multiespectral) e um recall de 80% para RGB e 70% para multiespectral destacando sua capacidade de identificar com precisão instâncias positivas. No geral, a Rede Neural emerge como um classificador altamente eficaz com desempenho notável em várias métricas de avaliação para os sensores.

Tabela 3 – Resultado de classificação da Rede Neural

Modelos	AUC	CA	F1	Precisão	Recall
RGB	0,969	0,800	0,803	0,827	0,800
Multiespectral	0,942	0,718	0,713	0,718	0,705

Em nosso estudo para a classificação com Rede Neural utilizamos a mesma quantidade de imagens para cada classe de número de plantas, sendo 10 imagens em cada para que o número entre as classes fossem o mesmo, sendo o total de 110 imagens por sensor. Em projetos futuros a quantidade de imagens em cada classe pode ser aumentada para auxiliar na melhoria do algoritmo em classificar diferentes números de plantas.

A Tabela 4 apresenta as matrizes de confusão obtidas dos dois sensores (RGB e multiespectral) aplicados a classificação do número de plantas de soja com Rede Neural. A Rede Neural exibe uma precisão alta para a RGB, pois classificou incorretamente apenas 22 imagens no total, resultando em uma taxa de erro baixa de 20%. Enquanto para a multiespectral o total de erros são 31 imagens, o que representa 28%.

Tabela 4 - Matriz de Confusão da Rede Neural (RGB e multiespectral)

(A) RGB

Predito

	10	11	12	13	15	16	17	21	22	23	25	Σ
10	9	0	0	0	0	0	0	1	0	0	0	10
11	0	9	0	0	0	0	0	0	0	1	0	10
12	0	1	7	0	0	0	0	2	0	0	0	10
13	0	0	0	8	0	1	0	0	0	1	0	10
15	0	0	0	0	8	1	0	0	0	1	0	10
16	0	0	0	0	0	7	0	1	1	1	0	10
17	0	0	0	0	0	1	7	0	2	0	0	10
21	0	0	0	0	0	0	0	10	0	0	0	10
22	1	0	0	0	0	0	0	0	9	0	0	10
23	0	0	1	1	1	0	0	0	0	7	0	10
25	0	0	0	0	0	1	0	0	0	2	7	10
Σ	10	10	8	9	9	11	7	14	12	13	7	110

Real

(B) MULTIESPECTRAL

Predito

	10	11	12	13	15	16	17	21	22	23	25	Σ
10	8	0	1	0	0	1	0	0	0	0	0	10
11	0	10	0	0	0	0	0	0	0	0	0	10
12	0	3	6	0	0	0	1	0	0	0	0	10
13	1	0	1	6	0	0	0	0	2	0	0	10
15	1	0	0	0	9	0	0	0	0	0	0	10
16	1	0	0	0	0	4	2	0	0	1	2	10
17	0	0	0	0	0	1	8	0	1	0	0	10
21	0	0	0	0	1	0	1	7	0	0	1	10
22	0	0	0	1	0	0	0	0	7	1	1	10
23	1	0	0	1	0	0	0	0	0	8	0	10
25	0	0	0	2	0	1	0	1	0	0	6	10
Σ	10	10	8	9	9	11	7	14	12	13	7	110

Real

3.3.2 Índices para o processamento de imagens RGB e multiespectral em linguagem R

Foram testados Índices de Vegetação usados em pesquisas semelhantes com finalidade para contagem de plantas de soja, os Índices de Vegetação que apresentaram melhores resultados foram os GLI e CIG, obtidos respectivamente pelas imagens RGB e multiespectral (Randelović et al., 2020; Matias; Caraza-Harter; Endelman, 2020). Conforme observado na Tabela 5 o MAE foi menor para os estádios VC, um valor de 0,057 para o índice GLI (imagens RGB) e 0.292 para o índice CIG (imagens multiespectrais). Esses índices conseguiram remover melhor os ruídos da imagem.

Tabela 5 - Erros Médios Absolutos (MAE) dos índices testados para remoção de ruído das imagens (RGB e multiespectral)

Sensores	Descrição	Índices	MAE		
			VC	V1	V2
RGB	Green Leaf Index	GLI	0,057	0,096	0,195
RGB	Blue Green Pigment Index	BGI	0,273	0,542	0,525
RGB	Overall Hue Index*	HUE	0,353	0,282	0,268
RGB	Primary Colors Hue Index	HI	0,450	0,542	0,475
RGB	Triangular greenness index	TGI	0,479	0,208	0,230
Multiespectral	Chlorophyll index — green	CIG	0,292	0,293	0,380
Multiespectral	Normalized Difference Red Edge Index	NDRE	0,306	0,316	0,444
Multiespectral	Normalized Difference Vegetation Index	NDVI	0,342	0,348	0,452

3.3.3 Estimativa do número de plantas com a técnica de processamento de imagens RGB e multiespectral

A Figura 5 mostra uma comparação entre o número de plantas estimado com processamento de imagens em linguagem R a partir de recortes de ortomosaicos e o número de plantas contadas manualmente no campo, com base no estágio vegetativo da soja. Cada gráfico de dispersão fornece uma comparação para cada estimativa do número de plantas. Observa-se que a estimativa do número de plantas diminui com o avanço do estágio da planta analisada, e o modelo apresentou maior precisão com o sensor de captura RGB no estágio VC ($R^2 = 0,38$), porém devido a presença de sobreposição a precisão do modelo foi baixa para todas as situações da RGB e multiespectral. No geral, os resultados fornecem evidências de que o modelo apresenta dificuldade de precisão de desempenho para estimar o número de plantas de soja quando há sobreposição em campo, o modelo conta uma menor média de plantas que a constatada na contagem em campo, indicando a sobreposição das plantas conforme na Tabela 6.

Figura 5 - Gráficos de dispersão comparando o número de plantas de campo com o número estimado com a linguagem R. A comparação considera o estágio de crescimento da planta (VC, V1 e V2), e o sensores (RGB e multiespectral)

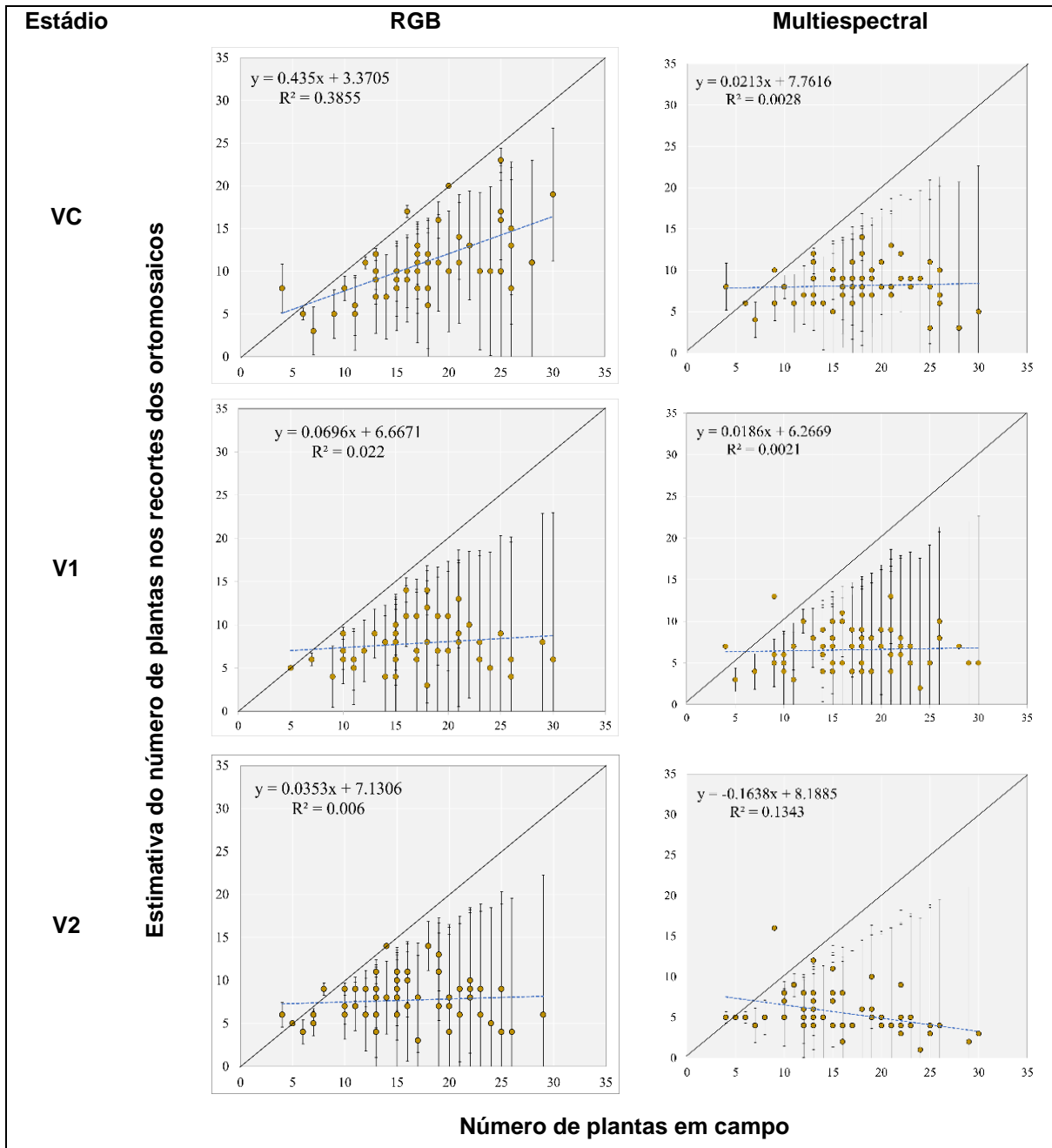
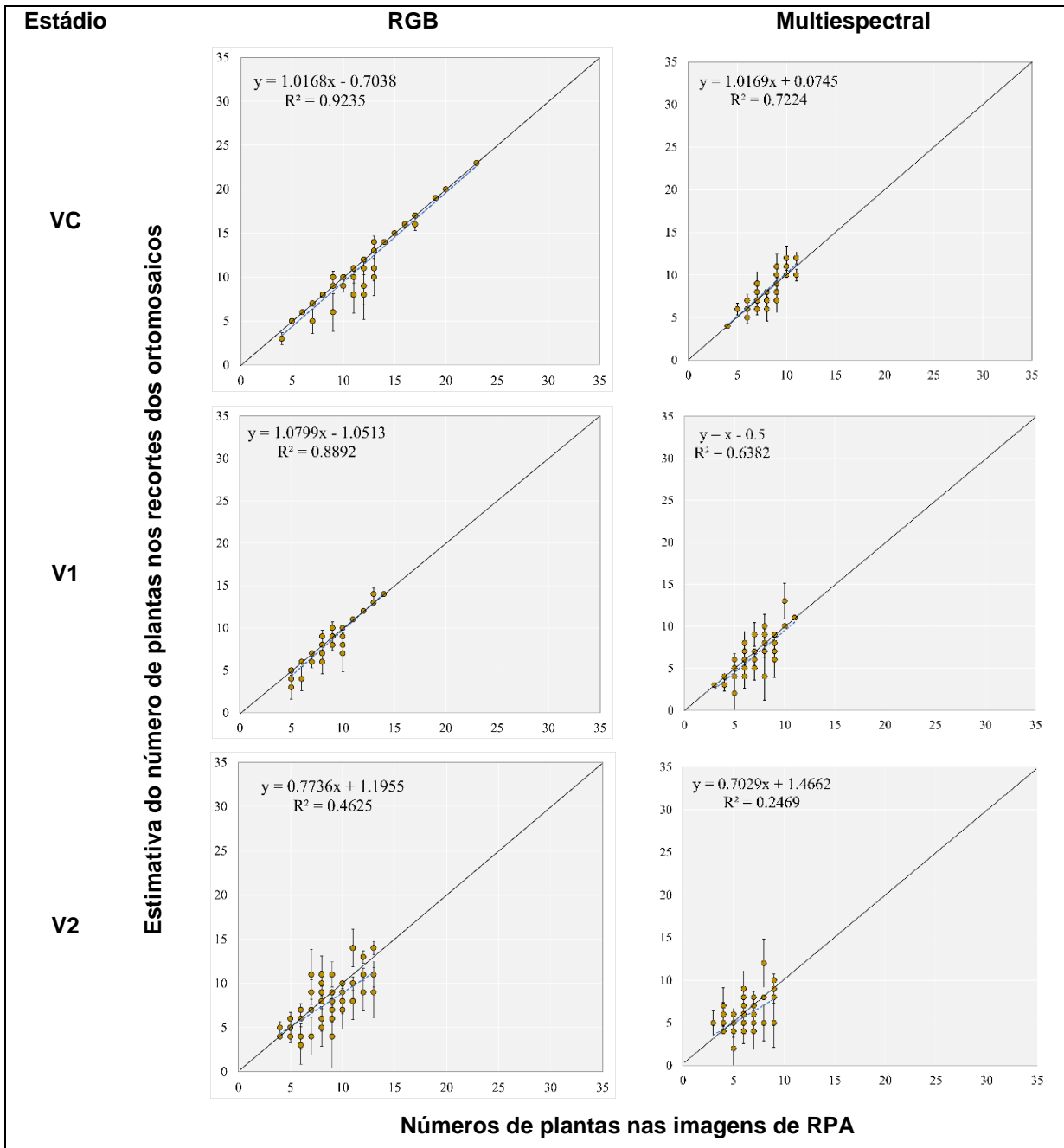


Tabela 6 – Média do número de plantas contados manualmente em campo e média do número de plantas estimado nos estádios VC, V1 e V2 e para cada sensor (RGB e multiespectral)

Média do número de plantas			
Real		Previsto	
Estádio	Em campo	RGB	Multiespectral
VC	17	12	8
V1	17	8	7
V2	16	8	6

No gráfico de dispersão na Figura 6, na estimativa de plantas comparada a contagem manual realizada diretamente nos recortes dos ortomosaicos, portanto eliminando as plantas sobrepostas, os modelos apresentaram alta precisão no estádio VC para imagens RGB e multiespectral ($R^2= 0,92; 0,72$). Foi possível estimar os números de plantas a partir das imagens, quando não apresentam sobreposição de plantas e com melhor desempenho nos estádios VC e V1 ($R^2= 0,92$ e $0,88$ respectivamente) com imagens RGB exceto no estádio V2 ($R^2= 0,46$).

Figura 6 - Gráficos de dispersão comparando o número de plantas contados manualmente nas imagens com o número estimado com a linguagem R. A comparação considera o estágio de crescimento da planta (VC, V1 e V2), e o sensores (RGB e multiespectral)



A estimativa da quantidade de plantas seguiu um padrão; à medida que o estágio vegetativo da soja aumentou, houve uma diminuição na precisão da estimativa do número de plantas. Em geral, nas condições em que o estágio vegetativo é menor (VC e V1), a precisão é melhor, indicando que quando as plantas têm a primeira folha trifoliada totalmente desenvolvida, a precisão da estimativa do número de plantas diminui. Conforme observado no gráfico de dispersão é possível verificar na Tabela 7 que a estimativa da contagem de plantas com menor MAE foi com o sensor RGB no estágio vegetativo VC (MAE = 0.057) quando eliminado as condições de sobreposição (realizado na contagem manual na imagem).

Tabela 7 – Erros absolutos médios (MAE) entre o número de plantas estimado em comparação com o número de plantas no campo e nos recortes dos ortomosaicos para cada estágio vegetativo e para cada sensor (RGB e multiespectral)

Estádio da soja	MAE			
	RGB		Multiespectral	
	Campo	Ortomosaico	Campo	Ortomosaico
VC	0,382	0,057	0,519	0,156
V1	0,464	0,096	0,572	0,152
V2	0,471	0,195	0,596	0,321

3.4 DISCUSSÃO

Os resultados da pesquisa foram influenciados tanto pelos tamanhos de pixel quanto pelos estádios de desenvolvimento das plantas. Isso implica que a agrupação de pixels dificulta a distinção entre solo e vegetação, especialmente devido à redução na resolução das imagens conforme o tamanho do pixel aumenta, o que torna mais difícil identificar plantas individuais. A qualidade das imagens é diretamente impactada pelo tamanho do pixel, sendo que o modelo da câmera e a altura do voo para captura das imagens também afetam essa medida (Hu et al., 2019).

Contudo, o estudo demonstrou que a técnica de aprendizado de máquina com Rede Neural pode ser aplicada mesmo em casos de sobreposição de plantas, para estimar a contagem de estandes de plantas. Essa abordagem mostrou alta precisão

em comparação com técnicas de processamento de imagens, mesmo diante da presença de sobreposição de plantas e com diferentes resoluções nas imagens RGB e multiespectral. Devido à sobreposição e ao espaçamento variável entre as plantas, a contagem de plantas é complexa em culturas como a colza, o trigo e a soja (Zhao et al., 2018), refletindo desafios semelhantes em todas essas culturas. Plantas como milho e algodão são mais facilmente contadas, pois possuem espaçamento quase uniforme e menor sobreposição em estádios iniciais de desenvolvimento (Varela et al., 2018).

O sensor RGB demonstrou superioridade nas duas técnicas empregadas neste estudo, devido à sua resolução mais alta de 0,5 cm por pixel, enquanto a câmera multiespectral operava com 1 cm por pixel. Essa diferença de resolução teve um impacto notável na precisão da contagem de plantas, resultando em uma precisão maior para o sensor RGB ($R^2 = 0,92$) e um erro menor na contagem ($MAE = 0,057$) em comparação com a câmera multiespectral. Além disso, as câmeras RGB têm a capacidade de capturar imagens de alta resolução, superiores a 10 megapixels, o que permite a criação de ortomosaicos mais detalhados, melhorando assim o processamento de imagens para a contagem de plantas (Li et al., 2019). Outras pesquisas confirmam a eficácia das imagens RGB, demonstrada pela habilidade de um algoritmo de segmentação por cluster k-means em identificar 91% das plantas de milho no campo (Shirzadifar et al., 2020).

Os resultados evidenciaram a precisão da Rede Neural, destacando sua capacidade como ferramenta para avaliação do rendimento da colheita e da semeadura por meio da contagem de plantas. Essa eficácia decorre da correlação direta entre a densidade populacional de plantas por unidade de área e a produtividade da colheita (Abuzar et al., 2011). Diversos fatores, como a disponibilidade de nutrientes e água (incluindo a textura do solo, o teor de umidade e a temperatura), a profundidade de semeadura, os efeitos dos herbicidas, as condições climáticas, as infestações por pragas e a exposição à luz solar, influenciam significativamente a densidade e a uniformidade das plantas (Al-kaisi; Hanna, 2006). Em última análise, esses elementos têm um impacto direto sobre o rendimento global e a qualidade da produção. A contagem precisa dos estandes de plantas é

fundamental para a determinação da produtividade, a otimização do processo de semeadura e a garantia da qualidade da semente (Ampatzidis; Partel, 2019).

3.5 CONCLUSÕES

O estudo investigou o impacto das condições ideais e da sobreposição de plantas na qualidade das estimativas de vegetação. Os resultados confirmaram a importância de selecionar o sensor apropriado e o momento adequado para a captura de imagens. Observou-se que a estimativa das plantas até o estágio VC reduziu consideravelmente o erro. A tecnologia RGB demonstrou alta precisão na classificação do número de plantas de soja utilizando Redes Neurais, enquanto com o processamento de imagens a precisão da estimativa do número de plantas apresentou baixa precisão, comprometida em ambientes de alta sobreposição de plantas, sugerindo a exploração de outras técnicas.

REFERÊNCIAS

ABBAS, A. *et al.* Characterizing soil salinity in irrigated agriculture using a remote sensing approach. **Physics and chemistry of the Earth, Parts A/B/C**, v. 55, p. 43-52, 2013.

ABUZAR, M. R. *et al.* Effect of plant population densities on yield of maize. **The Journal of Animal & Plant Sciences**, v. 21, n. 4, p. 692-695, 2011.

AL-KAISI, M.; HANNA, H. M. **Field soil variability and its impact on crop stand uniformity**. 2006.

AMPATZIDIS, Y.; PARTEL, V.; COSTA, L. Agroview: Cloud-based application to process, analyze and visualize UAV-collected data for precision agriculture applications utilizing artificial intelligence. **Computers and Electronics in Agriculture**, v. 174, p. 105457, 2020.

ARAÚJO, S.O.; PERES, R.S.; BARATA, J.; LIDON, F.; RAMALHO, J.C. Characterising the agriculture 4.0 landscape - emerging trends, challenges and opportunities. **Agronomy**, v. 11, 2021.

BROGE, Niels Henrik; LEBLANC, Eric. Comparing prediction power and stability of broadband and hyperspectral vegetation indices for estimation of green leaf area index and canopy chlorophyll density. **Remote sensing of environment**, v. 76, n. 2, p. 156-172, 2001.

CARVER, S. M. **Impact of Planting Strategies on Soybean (Glycine max L.) Growth, Development and Yield**. Mississippi State University, 2018.

CHEN, G.; WIATRAK, P. Seeding rate effects on soybean height, yield, and economic return. **Agronomy journal**, v. 103, n. 5, p. 1301-1307, 2011.

COLORADO, J. D. *et al.* A novel NIR-image segmentation method for the precise estimation of above-ground biomass in rice crops. **PloS one**, v. 15, n. 10, p. e0239591, 2020.

DEMŠAR, J *et al.* **Orange: Data Mining Toolbox in Python**, Journal of Machine Learning Research 14(Aug): 2349–2353, 2013.

EGLI, D. B.; CORNELIUS, P. L. A regional analysis of the response of soybean yield to planting date. **Agronomy journal**, v. 101, n. 2, p. 330-335, 2009.

EPLER, M.; STAGGENBORG, S. Soybean yield and yield component response to plant density in narrow row systems. **Crop management**, v. 7, n. 1, p. 1-13, 2008.

FAN, Z. *et al.* Automatic tobacco plant detection in UAV images via deep neural networks. **IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing**, v. 11, n. 3, p. 876-887, 2018

GASPAR, A. P.; CONLEY, S. P. Responses of canopy reflectance, light interception, and soybean seed yield to replanting suboptimal stands. **Crop Science**, v. 55, n. 1, p. 377-385, 2015.

GASPAR, A.P. *et al.* Defining optimal soybean seeding rates and associated risk across North America. **Agronomy Journal**, v. 112, n. 3, p. 2103-2114, 2020.

GITELSON, A. A.; GRITZ, Y.; MERZLYAK, M. N. Relationships between leaf chlorophyll content and spectral reflectance and algorithms for non-destructive chlorophyll assessment in higher plant leaves. **Journal of plant physiology**, v. 160, n. 3, p. 271-282, 2003.

GITELSON, A. A.; KAUFMAN, Y. J.; MERZLYAK, M. N. Use of a green channel in remote sensing of global vegetation from EOS-MODIS. **Remote sensing of Environment**, v. 58, n. 3, p. 289-298, 1996.

HAGEN, N.; KUDENOV, M. W. Review of snapshot spectral imaging technologies. **Optical Engineering**, v. 52, n. 9, p. 090901-090901, 2013.

HU, P. *et al.* Pixel size of aerial imagery constrains the applications of unmanned aerial vehicle in crop breeding. **ISPRS Journal of Photogrammetry and Remote Sensing**, v. 154, p. 1-9, 2019.

HUNT, E. Raymond *et al.* Evaluation of digital photography from model aircraft for remote sensing of crop biomass and nitrogen status. **Precision Agriculture**, v. 6, p. 359-378, 2005

JIA, J.; KRUTZ, G. W. Location of the maize plant with machine vision. **Journal of agricultural engineering research**, v. 52, p. 169-181, 1992.

LI, B. *et al.* The estimation of crop emergence in potatoes by UAV RGB imagery. **Plant Methods**, v. 15, n. 1, p. 1-13, 2019.

LEE, C. D.; EGLI, D. B.; TEKRONY, D. M. Soybean response to plant population at early and late planting dates in the Mid-South. **Agronomy Journal**, v. 100, n. 4, p. 971-976, 2008.

LEE, C.; HERBEK, J. Soybean planting in Kentucky. **Retrieved from University**, 2011.

LI, B. *et al.* The estimation of crop emergence in potatoes by UAV RGB imagery. **Plant Methods**, v. 15, n. 1, p. 1-13, 2019.

LICHT, M. **Time for early season stand assessments**. 2020. Iowa State Univ. Extension and Outreach: Integrated Crop Management. Disponível em: <https://crops.extension.iastate.edu/blog/mark-licht/time-early-season-stand-assessments>. Acesso em: 22 ago. 2023.

MAES, W. H.; STEPPE, K. Perspectives for remote sensing with unmanned aerial vehicles in precision agriculture. **Trends in plant science**, v. 24, n. 2, p. 152-164, 2019.

MATIAS, F. I.; CARAZA-HARTER, M. V.; ENDELMAN, J. B. FIELDimageR: An R package to analyze orthomosaic images from agricultural field trials. **The Plant Phenome Journal**, v. 3, n. 1, p. e20005, 2020.

MENDONÇA, G. V.; SILVA, E. A. B. **The electrical engineering handbook**, p. 891, 2004.

MERZLYAK, Mark N. *et al.* Non-destructive optical detection of pigment changes during leaf senescence and fruit ripening. **Physiologia plantarum**, v. 106, n. 1, p. 135-141, 1999.

MOREIRA, F. F. *et al.* Improving the efficiency of soybean breeding with high-throughput canopy phenotyping. **Plant Methods**, v. 15, p. 1-9, 2019.

NLEYA, T. *et al.* Planting date, cultivar, seed treatment, and seeding rate effects on soybean growth and yield. **Agrosystems, Geosciences & Environment**, v. 3, n. 1, p. e20045, 2020.

NUKALA, Revathi *et al.* Internet of Things: A review from 'Farm to Fork'. In: **2016 27th Irish signals and systems conference (ISSC)**. IEEE, 2016. p. 1-6.

OLECHOWSKI, H.T. **Soybean production**. 1983. Ontario Ministry of Agriculture. Disponível em: http://eap.mcgill.ca/CPSO_2.htm. Acesso em: 11 jul. 2023.

PATHAK, H. *et al.* A review of unmanned aerial vehicle-based methods for plant stand count evaluation in row crops. **Computers and Electronics in Agriculture**, v. 198, p. 107064, 2022.

PATRIGNANI, Andres; OCHSNER, Tyson E. Canopeo: A powerful new tool for measuring fractional green canopy cover. **Agronomy journal**, v. 107, n. 6, p. 2312-2320, 2015.

PRAKASH, Anupma. Thermal remote sensing: concepts, issues and applications. **International Archives of Photogrammetry and Remote Sensing**, v. 33, n. B1; PART 1, p. 239-243, 2000.

ROBINSON, A.P.; CONLEY, S.P. **Soybean production systems**. 2007. Purdue Ext. p. 1–3. Disponível em: <https://www.extension.purdue.edu/extmedia/ay/ay-217-w.pdf>. Acesso em: 11 jul. 2023.

ROUSE JR, J. W. *et al.* **Monitoring the vernal advancement and retrogradation (green wave effect) of natural vegetation**. Texas: NTRS - NASA Technical Reports Server. 1974.

SALMERON, M. *et al.* Soybean maturity group choices for early and late plantings in the Midsouth. **Agronomy Journal**, v. 106, n. 5, p. 1893-1901, 2014.

SHIRZADIFAR, A. *et al.* Mapping crop stand count and planting uniformity using high resolution imagery in a maize crop. **Biosystems Engineering**, v. 200, p. 377-390, 2020.

SHI, Y. *et al.* Automatic corn plant location and spacing measurement using laser line-scan technique. **Precision Agriculture**, v. 14, p. 478-494, 2013.

STOLL, M.; JONES, H. G. Thermal imaging as a viable tool for monitoring plant stress. **OENO One**, v. 41, n. 2, p. 77-84, 2007.

SHRESTHA, D. S.; STEWARD, B. L. Automatic corn plant population measurement using machine vision. **Transactions of the ASAE**, v. 46, n. 2, p. 559, 2003.

SMITH, R. M. *et al.* Evaluation of planter errors associated with twin-row soybean production in Mississippi. **Agronomy Journal**, v. 111, n. 4, p. 1643-1649, 2019.

STATON, M. **Assessing soybean emergence**. 2013. Michigan State Univ. Extension Soybeans. Disponível em: https://www.canr.msu.edu/news/assessing_soybean_emergence. Acesso em: 11 jul. 2023.

TANG, L.; TIAN, L. F. Plant identification in mosaicked crop row images for automatic emerged corn plant spacing measurement. **Transactions of the ASABE**, v. 51, n. 6, p. 2181-2191, 2008.

TORRES-SÁNCHEZ, J. *et al.* Multi-temporal mapping of the vegetation fraction in early-season wheat fields using images from UAV. **Computers and Electronics in Agriculture**, v. 103, p. 104-113, 2014.

VALENTE, J. *et al.* Automated crop plant counting from very high-resolution aerial imagery. **Precision Agriculture**, v. 21, p. 1366-1384, 2020.

VARELA, Sebastian *et al.* Early-season stand count determination in corn via integration of imagery from unmanned aerial systems (UAS) and supervised learning techniques. **Remote Sensing**, v. 10, n. 2, p. 343, 2018.

VOGEL, J.T. *et al.* Soybean yield formation physiology—a foundation for precision breeding based improvement. **Frontiers in plant science**, v. 12, p. 719706, 2021.

VONG, C.N. *et al.* Early corn stand count of different cropping systems using UAV-imagery and deep learning. **Computers and Electronics in Agriculture**, v. 186, p. 106214, 2021.

WHIGHAM, K. *et al.* **Soybean replant decisions**. Iowa State University, University Extension, Ames, IA. 2000.

WOEBBECKE, David M. *et al.* Color indices for weed identification under various soil, residue, and lighting conditions. **Transactions of the ASAE**, v. 38, n. 1, p. 259-269, 1995.

WRIGLEY, Colin W. *et al.* (Ed.). **Encyclopedia of food grains**. Academic Press, 2015.

YANG, C. Remote sensing and precision agriculture technologies for crop disease detection and management with a practical application example. **Engineering**, v. 6, n. 5, p. 528-532, 2020.

ZHAO, B. *et al.* Rapeseed seedling stand counting and seeding performance evaluation at two early growth stages based on unmanned aerial vehicle imagery. **Frontiers in Plant Science**, v. 9, p. 1362, 2018.

ZHU, Lingli *et al.* A review: Remote sensing sensors. **Multi-purposeful application of geospatial data**. p. 19-42, 2018.

CONSIDERAÇÕES FINAIS

Os estudos referentes a classificação de falhas de semeadura apontam para avanços notáveis na caracterização de áreas de cultivo de soja, utilizando Redes Neurais como algoritmo promissor. A otimização das configurações das Redes Neurais resultou em uma classificação mais precisa das áreas de semeadura, tanto com falhas quanto sem, em imagens de RPA. Este estudo revelou o potencial das Redes Neurais na solução de desafios específicos, como a classificação de falhas de soja, contribuindo para o avanço da agricultura de precisão.

Foram investigados os impactos do estágio fenológico da soja e da velocidade de semeadura na estimativa do número de plantas. Os resultados indicaram a importância de uma interação estratégica com esses fatores para alcançar uma correlação precisa entre as estimativas obtidas nos ortomosaicos e o número real de plantas em campo. Foi observado que a escolha criteriosa do estágio fenológico da soja e do momento ideal para a captura de imagens são determinantes para a acurácia das estimativas. Destaca-se que o estágio VC e a velocidade 1,1 m/s demonstraram maior precisão, relacionada à crescente sobreposição de vegetação à medida que as plantas crescem e a velocidade aumenta, resultando em uma redução na precisão das estimativas.

Em ambientes de alta sobreposição de plantas, o processamento de imagens apresentou limitações na estimativa do número de plantas. Isso ressalta a necessidade de explorar outras técnicas e estratégias para aprimorar a precisão nessas condições específicas.

Por fim, no âmbito da tecnologia RGB combinada com Redes Neurais, evidenciou-se a alta precisão na classificação do número de plantas de soja. A seleção cuidadosa do sensor adequado e a consideração da sobreposição de plantas foram determinantes para alcançar resultados satisfatórios nesse aspecto.

REFERÊNCIAS

BOARD, J. *et al.* **Can Soybean Seeding Rates Be Reduced Without Affecting Yields in Louisiana.** 2013.

CHEN, G.; WIATRAC, P. Seeding rate effects on soybean height, yield, and economic return. **Agronomy journal**, v. 103, n. 5, p. 1301-1307, 2011.

DALL'AGNOL, A.; **A Embrapa Soja no contexto do desenvolvimento da soja no Brasil: histórico e contribuições.** Brasília, DF: Embrapa, 2016.

GASPAR, A.P. *et al.* Defining optimal soybean seeding rates and associated risk across North America. **Agronomy Journal**, v. 112, n. 3, p. 2103-2114, 2020.

HAYKIN, S. **Neural networks and learning machines**, 3/E. Pearson Education India, 2009.

HIRAKURI, M. H. O contexto econômico da produção de soja. **Tecnologias de Produção de Soja. Londrina: Embrapa Soja.** p. 15-32, 2020.

LEE, C. D.; EGLI, D. B.; TEKRONY, D. M. Soybean response to plant population at early and late planting dates in the Mid-South. **Agronomy Journal**, v. 100, n. 4, p. 971-976, 2008.

LICHT, M. **Soybean plant population.** 2016. Science for Success Soybean: United Soybean Board. Disponível em: http://crops.extension.iastate.edu/soybean/production_plantpopulation.html. Acesso em: 11 dez. 2022.

NAFZIGER, E.D. **Soybean.** 2022. Illinois Agronomy Handbook. Crop Sci. Dept.: 282–319. doi: 10.1016/B978-0-12-819194-1.00008-6. Disponível em: <http://extension.cropsciences.illinois.edu/handbook/pdfs/chapter03.pdf>. Acesso em: 12 set. 2022.

RINCKER, K. *et al.* Genetic improvement of US soybean in maturity groups II, III, and IV. **Crop science**, v. 54, n. 4, p. 1419-1432, 2014.

ROUSE JR, J. W. *et al.* **Monitoring the vernal advancement and retrogradation (green wave effect) of natural vegetation.** Texas: NTRS - NASA Technical Reports Server. 1974.

SCHNITKEY, G. *et al.* **Weekly Farm Economics: 2022 break-even prices for corn and soybeans.** farmdocDAILY Univ. Illinois Urbana-Champaign 11(168), p. 1–5, 2022.

SEDIVY, E.J.; WU, F.; HANZAWA, Y. Soybean domestication: the origin, genetic architecture and molecular bases. **New Phytologist**, v. 214, n. 2, p. 539-553, 2017.

SEDIYAMA, T. **Tecnologia de produção e usos da soja.** Londrina: Mecenias, 2009. 314p.

SETIYONO, T.D. **SoySim: A simulation model for soybean growth and yield.** 2010. Dept. of Agronomy and Horticulture, Institute of Agricultural and Natural Resources by University of Nebraska-Lincoln. Disponível em: https://soysim.unl.edu/soysim_manual.html. Acesso em: 20 ago. 2022.

TURNIPSEED, B. **Chapter 9: Soybean seed testing and seed quality.** SDSU Ext. Best Manag. Pract. Ch. 9, p. 79–84, 2019.

USDA. **National Agriculture Statistics Services Quick Stats.** 2023a. USDA-NASS, Washington D.C. Disponível em: https://quickstats.nass.usda.gov/results/27B99EF0-5739-34F3-9263-C4320EC2552D?pivot=short_desc. Acesso em: 11 nov. 2023.

USDA. **National Agriculture Recent Cost and Returns.** 2023b. USDA-NASS, Washington D.C. Disponível em: https://quickstats.nass.usda.gov/results/27B99EF0-5739-34F3-9263-C4320EC2552D?pivot=short_desc. Acesso em: 11 nov. 2023.

ZHAO, B. *et al.* Rapeseed seedling stand counting and seeding performance evaluation at two early growth stages based on unmanned aerial vehicle imagery. **Frontiers in Plant Science**, v. 9, p. 1362, 2018.

XU, C. *et al.* High density and uniform plant distribution improve soybean yield by regulating population uniformity and canopy light interception. **Agronomy**, v. 11, n. 9, p. 1880, 2021.