

ORIGINAL PAPER

Aprendizado de máquina na agricultura indoor: um mapeamento sistemático e taxonomia

Machine learning in indoor agriculture: a systematic mapping and taxonomy

Bruno Guilherme Martini ¹ and Jorge Luis Victória Barbosa ¹

¹Universidade do Vale do Rio dos Sinos – UNISINOS, Programa de Pós-Graduação em Computação Aplicada – PPGCA

*brunomartin@edu.unisinos.br; jbarbosa@unisinos.br

Recebido: 11/05/2024. Revisado: 11/11/2024. Aceito: 30/11/2024.

Resumo

O presente artigo descreve um mapeamento sistemático dos trabalhos relacionados à aplicação de Aprendizado de Máquina na Agricultura Indoor. Esta pesquisa abrange buscas realizadas até março de 2024 nas bases IEEE Xplore, ACM Digital Library, Springer Library, Science Direct, Scopus, MDPI, Wiley e Taylor & Francis. A busca inicial resultou em 10.149 artigos, dos quais foram selecionados 76 estudos para leitura completa após a aplicação dos critérios de inclusão e exclusão. Essa análise resultou na seleção de 36 artigos que foram estudados com o intuito de responder às 9 questões de pesquisa juntamente com uma proposta de taxonomia. Os principais resultados revelam que os artigos selecionados empregaram 43 técnicas diferentes de aprendizado de máquina. Além disso, constatou-se que nenhum estudo foi publicado antes do ano de 2018. Foram identificadas 26 informações distintas coletadas dos cultivos, observou-se o cultivo de 17 culturas diferentes e o emprego de 4 dispositivos para coleta das informações, entre outras informações relevantes. A análise dos artigos evidenciou uma clara tendência no uso de Aprendizado de Máquina na Agricultura Indoor.

Palavras-Chave: Agricultura Indoor; Aprendizado de máquina; Fazenda Indoor; Inteligência Artificial

Abstract

The present article describes a systematic mapping of works related to the application of Machine Learning in Indoor Agriculture. This research encompasses searches conducted up to March 2024 in the IEEE Xplore, ACM Digital Library, Springer Library, Science Direct, Scopus, MDPI, Wiley, and Taylor & Francis databases. The initial search resulted in 10,149 articles, of which 76 studies were selected for full reading after applying the inclusion and exclusion criteria. This analysis led to the selection of 36 articles that were studied with the aim of answering 9 research questions along with a taxonomy proposal. The main results reveal that the selected articles employed 43 different machine learning techniques. Additionally, it was found that no study was published before the year 2018. Twenty-six distinct pieces of information were identified from the crops, with the cultivation of 17 different crops and the use of 4 devices for information collection, among other relevant information. The analysis of the articles evidenced a clear trend in the use of machine learning in Indoor Agriculture.

Keywords: Artificial Intelligence; Indoor Agriculture; Indoor Farming; Machine Learning

1 Introdução

O avanço da tecnologia vem ampliando a modernização da agricultura. Essa modernização vem ocorrendo principalmente através do uso de computadores embarcados, do surgimento de sensores para as mais diversas finalidades relacionadas à agricultura e do aumento da eficácia e confiabilidade na comunicação entre os dispositivos eletrônicos. Com isso a agricultura está adotando recursos tecnológicos projetados para aumento do rendimento desde o plantio até a colheita através de monitoramento, recomendações de melhorias do cultivo (Rodríguez et al., 2017) e suporte no cultivo através de auxílio em dificuldades dos agricultores (Sisyanto et al., 2017).

A oferta de monitoramento, recomendações e suporte para a agricultura desde o preparo do solo até a venda do produto pode ser uma tarefa inviável, mas oferecer esses recursos em ambientes dotados de tecnologias, tais como rede de sensores, informações de cultivo, geolocalização, mapeamento de recursos, e sinal de Internet, torna-se possível em um local fechado como estufas e/ou pavilhões. O aproveitamento das informações geradas durante o monitoramento de uma plantação, por exemplo, passa a ser um aspecto importante para oferecer indicações para outros cultivos.

Os avanços tecnológicos ocasionaram o aumento na produtividade e na imunidade das culturas plantadas. Um dos fatos que gera a busca pelo aumento da produtividade agrícola é que aproximadamente 70% da água doce disponível no planeta é consumida pelo setor agrícola (FAO, 2017). Assim, o aumento da eficiência da agricultura torna-se estratégica, fomentando investimentos públicos e privados neste setor. Devido ao crescimento da população mundial e consequentemente maior consumo de comida, a produção de alimentos deverá aumentar em 70% até o ano de 2050 (Saath e Fachinello, 2018). Isso mostra a necessidade do aumento da eficiência da agricultura.

Como as informações relativas às lavouras estão cada vez mais acessíveis, estudos sobre a agricultura *indoor* (Martini et al., 2021) e também sobre aprendizado de máquina aplicado à agricultura (Altalak et al., 2022; Anbananthen et al., 2021; Vadivel et al., 2019) estão sendo elaborados com maior frequência pela comunidade científica como mostra esse mapeamento sistemático. Além da utilização de sensores para obter informações relativas à produtividade, estudos estão utilizando câmeras dispostas ao longo do cultivo (Joshi et al., 2017; Tan et al., 2016) para viabilizar o processamento de imagens para monitoramento da produção. Esse processamento normalmente ocorre junto a um computador com capacidade de processamento e não diretamente nos módulos que capturam as imagens.

Plantas de espécies que normalmente cresciam em campos ao ar livre estão cada vez mais sendo cultivadas em locais protegidos, como estufas e pavilhões. Devido a essas mudanças, surgiu uma variação do termo agricultura, conhecida como agricultura *indoor*, também chamada de cultivo protegido no Brasil. Essa mudança ocorre porque dentro das estufas é possível criar um microclima favorável para as plantas, o que resulta em melhores resultados na redução de pragas, menor necessidade de aplicação de pesticidas, aumento da produção, entre outros benefícios.

O cultivo protegido utiliza o conceito de *Smart Farming* (Karunathilake et al., 2023; Doshi et al., 2019) (Fazenda Inteligente) que representa a utilização dos sistemas de tecnologia da informação e comunicação (ICT do inglês *Information and Communication Technology*) aplicados na agricultura, levando ao que pode ser chamado de Terceira Revolução Verde (Akis, 2017).

Os modelos e estudos aplicados à agricultura podem utilizar dados de contexto para melhorar na tomada de decisões. Contexto é um conjunto de informações que podem caracterizar circunstâncias de um lugar, objeto, pessoa ou até mesmo entidades. Normalmente, o contexto aborda a localização, identidade e estado de pessoas, grupos e objetos computacionais e físicos (Dey, 2001). Para a agricultura o contexto geralmente utiliza as informações disponíveis do cultivo, como quantidade de irrigações, temperatura do ambiente, umidade do solo, umidade do ambiente e outras informações que podem ser obtidas com sensores embarcados nos cultivos. Como metodologia este artigo utiliza o mapeamento sistemático, que de acordo com Petersen et al. (2015), elimina ou reduz as divergências que possam existir quando comparado com revisões que não possuam um método determinado. Portanto, este estudo permite mostrar as técnicas mais utilizadas na agricultura *indoor* com a aplicação de aprendizado de máquina. O trabalho visa também, com base na literatura analisada, apresentar lacunas e desafios de pesquisa. Deste modo, o trabalho apresenta uma visão dos estudos publicados nos últimos anos suportando futuras pesquisas.

O artigo está organizado da seguinte forma. A seção 2 apresenta os estudos relacionados. A seção 3 aborda como o mapeamento foi elaborado e executado. Na mesma seção, os artigos utilizados no mapeamento sistemático estão organizados em ordem cronológica juntamente de seus autores. Os resultados obtidos referentes a cada pergunta do mapeamento são apresentados na seção 4. A seção 5 expõe as ameaças à validade do trabalho. Na seção 6 é apresentada uma discussão sobre os trabalhos encontrados, e finalmente a seção 7 contém a conclusão e sugestão de trabalhos futuros.

2 Trabalhos Relacionados

Durante a execução desse mapeamento não foram encontrados trabalhos de revisões voltados para agricultura *indoor* e aprendizado de máquina. Devido a isso os trabalhos citados são revisões encontradas no decorrer desse mapeamento e que possuem alguma relação da agricultura com aprendizado de máquina. Um dos trabalhos está interligado a Era 4.0 (Martinho, 2022), onde a Agricultura 4.0 faz uso de tecnologias modernas, como o Controle de Sensoriamento Remoto, *Machine Learning*, *Big Data*, *Deep Learning*, Imagens Térmicas e Veículos Aéreos Não Tripulados (VANTs). Outro estudo apresenta técnicas de *deep learning* junto a imagens de satélites para utilização na agricultura extensiva (Victor et al., 2023). O estudo de Helfer et al. (2020) apresenta um modelo que usa aprendizado de máquina para prever a fertilidade do solo onde as plantas são cultivadas. Por fim, um quarto estudo traz as estratégias avançadas de controle e monitoramento na Agricultura Inteligente (Hassan et al., 2021).

Também foram encontrados três estudos (Ojha et al., 2015; Wu et al., 2017) com objetivos semelhantes a este mapeamento, dois trabalhos são sobre o estado da arte dos sensores, comunicação para transmitir as informações e outras tecnologias que estão sendo aplicadas na agricultura. O trabalho de Wu et al. (2017) apresentou as tecnologias que auxiliam a agricultura, como IoT (*Internet of Things*), tecnologia dos sensores, computação em nuvem, entre outros aspectos. Além disso, Wu et al. (2017) trataram da agricultura inteligente (*smart agriculture*) e sua interligação com os requisitos do *smart planet*.

Já Ojha et al. (2015) abordaram o estado da arte dos sensores de rede sem fio, comumente chamados de *Wireless Sensor Networks* (WSNs). Os WSNs impulsionaram a pesquisa no domínio agrícola, pois não há mais a necessidade de cabeamento para receber as informações. O estudo de Avila e Barbosa (2024) revisa o uso de ambientes inteligentes na agricultura digital. A revisão incluiu 51 artigos entre 2019 e abril de 2023, destacando tecnologias como sensores, *Big Data* e aprendizado de máquina.

Este mapeamento visa apresentar como os estudos publicados utilizaram o aprendizado de máquina aplicado na agricultura *indoor*. O aprendizado de máquina pode ser aplicado para diversas finalidades, como: obter melhores resultados de produção, detectar pragas no cultivo, gerar informações através de imagens e outros serviços. O mapeamento sistemático, por possuir uma metodologia bem explicada, de boa aceitação e com várias etapas, torna-se mais valioso do que revisões ou pesquisas simples voltadas a temas específicos, devido à falta de demonstração da metodologia e de resultados completos sobre a área. Além disso, o estudo tem como motivação identificar os principais autores da área, os trabalhos mais importantes, os sensores mais utilizados, as bases de dados mais relevantes para a área, as deficiências que os estudos possuem, o índice de estudos ao longo do ano e os desafios da área, visando auxiliar futuras pesquisas.

3 Metodologia

Com a utilização da metodologia do mapeamento sistemático é possível conduzir uma revisão da literatura sobre como o aprendizado de máquina está auxiliando a agricultura *indoor*. Assim, o mapeamento sistemático é capaz de fazer a ligação entre a coleta de dados, os locais onde o estudo foi publicado, qual a área de aplicação na agricultura, qual o ano de publicação dos estudos, tecnologias que foram utilizadas, qual metodologia empregada no desenvolvimento, entre outras informações. O mapeamento se concentra em artigos publicados em periódicos, conferências e workshops. As etapas para o mapeamento são:

- Elaboração das questões de pesquisa.
- Elaboração do processo de busca.
- Definição dos critérios de filtragem dos resultados.

3.1 Elaboração das questões de pesquisa

Foram definidas oito questões de pesquisa para este mapeamento. As oito questões foram organizadas em: uma questão geral (QG), cinco questões específicas (QEP) e duas

questões estatísticas (QET). O objetivo da questão geral é entender como os estudos estão utilizando o aprendizado de máquina na agricultura *indoor*. As questões específicas concentram-se em desafios específicos, tecnologias, técnicas, dados e recursos adicionais utilizados nos estudos. Por fim, o objetivo das questões estatísticas é reunir dados sobre a área em estudo e mostrar o comportamento das publicações ao longo dos últimos anos, finalizando em março de 2024. As questões de pesquisa são apresentadas na Tabela 1.

Tabela 1: Questões de Pesquisa

Tipo	Questões
Questão Geral	
QG	Com qual finalidade os estudos estão utilizando o aprendizado de máquina na agricultura <i>indoor</i> ?
Questões Específicas	
QEP1	Quais as técnicas de aprendizado de máquina estão sendo utilizadas??
QEP2	Quais os dados estão sendo utilizados nos estudos?
QEP3	Os trabalhos utilizam estratégias de fusão de informações de datasets?
QEP4	Quais estudos utilizam informações de contexto?
QEP5	Quais foram os dispositivos utilizados para obter a coleta de informações dos cultivos?
QEP6	Quais as culturas de plantas estão sendo cultivadas nos estudos?
Questões Estatísticas	
QET1	Onde as pesquisas foram publicadas?
QET2	Quantas publicações ocorreram por ano?

3.2 Elaboração do processo de busca

O processo de busca foi organizado em três etapas: (1) especificar a *string* de pesquisa, (2) selecionar as bases de dados e entender o método de pesquisa de cada uma delas, e por fim, (3) obter os resultados das bases. Para a primeira etapa, foi adotado o processo de pesquisa proposto por Petersen et al. (2015). A primeira etapa identifica os termos principais e seus sinônimos mais relevantes. Neste estudo foram selecionados como termos principais os seguintes: *Machine Learning* e *Indoor Agriculture*.

Na busca dos estudos também são aplicados os termos secundários, que são sinônimos dos termos primários, e que também são utilizados na *string* de pesquisa. O termo principal *Machine Learning* possui os sinônimos *Artificial Intelligence*, *Neural Networks* e *Deep Learning*. É possível notar que o termo secundário *Artificial Intelligence* não é um sinônimo de *Machine Learning*, porém podem existir trabalhos de aprendizado de máquina que utilizam o termo *Artificial Intelligence* e não o termo *Machine Learning*, apesar de serem estratégias de aprendizado de máquina.

O termo *Indoor Agriculture* possui os termos secundários *Urban Agriculture*, *Indoor Farming*, *Indoor Farm* e *Vertical Farming*. Os termos secundários relacionados a *Machine*

Learning são estratégias utilizadas no aprendizado de máquina, com exceção do termo *Artificial Intelligence*. Já os termos secundários da *Indoor Agriculture* são termos também utilizados nos estudos que condizem com cultivos em áreas internas, sendo esses cultivos em solo, vasos e perfis de PVC. Os termos são apresentados na [Tabela 2](#).

Tabela 2: String de Pesquisa

Termos da string de pesquisa	
Termos Principais	Termos Secundários
<i>Machine Learning</i>	"machine learning" OR "artificial intelligence" OR "neural networks" OR "deep learning"
<i>Indoor Agriculture</i>	"indoor agriculture" OR "urban agriculture" OR "indoor farming" OR "indoor farm" OR "vertical farming"

Uma vez definidos os termos e sinônimos, foi elaborada a seguinte string de pesquisa:

((("machine learning" OR "artificial intelligence" OR "neural networks" OR "deep learning")) AND ("indoor agriculture" OR "urban agriculture" OR "indoor farming" OR "indoor farm" OR "vertical farming"))

Uma vez definida a string de pesquisa, a segunda etapa foi selecionar bases de dados relevantes para a área de estudo. Cada base possui seu próprio método de pesquisa. Após a escolha delas, foi necessário determinar os parâmetros de pesquisa a serem usados em cada uma. Oito bases de dados foram utilizadas, incluindo ACM Digital Library, IEEE Xplore, Science Direct, Springer Link, Scopus, MDPI, Taylor and Francis e Wiley.

A pesquisa na base de dados ACM Digital Library exigiu o uso de recursos de pesquisa avançada, nos quais a string foi inserida no campo de edição da própria base. Para pesquisar na base de dados da IEEE Xplore, também foi utilizado o recurso de pesquisa avançada. O processo de busca no repositório da Science Direct envolveu a aplicação da string de consulta ao título, resumo e palavras-chave. Na Springer Library, além de usar a string de pesquisa como a consulta, também foi necessária a remoção dos documentos categorizados como "Preview Only". Após a exclusão desses documentos, foi selecionado o filtro de pesquisa intitulado "Computer Science" inicialmente, efetuiu-se a pesquisa e posteriormente selecionou-se o filtro "Life Science" para obter resultados que também possam ter ligação com o tema pesquisado.

Na base de dados Scopus foi inserida a string do mapeamento juntamente dos seguintes filtros: artigos de periódico, artigos de conferência, revisões de conferências e revisões. Na base Wiley o único filtro utilizado junto da string de pesquisa foi a marcação de "artigos" na pesquisa avançada. A base Taylor and Francis teve a pesquisa somente com a string de pesquisa, já na base MDPI além da inserção da string também foram selecionados os filtros artigos e artigos de conferências.

3.3 Definição dos critérios de filtragem dos resultados

A obtenção dos estudos através da pesquisa nas bases de dados passou por Critérios de Inclusão (CI) e Critérios de Exclusão (CE). A seguir são apresentados os CI:

- CI 1: O estudo deverá estar publicado em periódico, conferência ou workshop.
- CI 2: O estudo deverá ser um full paper.
- CI 3: O estudo deve conter termos que se relacionem com a string de busca.
- CI 4: O estudo deve ser escrito em inglês.
- CI 5: O estudo deverá ter sido publicado até março de 2024.

Os Critérios de Exclusão (CE) são os seguintes:

- CE 1: O estudo consiste em uma revisão de literatura ou revisão sistemática.
- CE 2: O estudo não possui foco na agricultura indoor.
- CE 3: O estudo não apresenta uma aplicação de aprendizado de máquina.

Os critérios de inclusão e exclusão auxiliam durante o processo de filtragem a fim de obter os estudos mais aderentes ao mapeamento e eliminar discrepâncias geradas na pesquisa. A quantidade de trabalhos obtidos em cada etapa da filtragem, assim como as porcentagens de trabalhos retirados em cada uma delas, são mostradas na [Fig. 1](#).

Após os artigos terem sido filtrados na busca inicial através da string e depois de terem passado pelos cinco primeiros critérios de inclusão, os mesmos foram submetidos à terceira filtragem. O terceiro filtro do mapeamento consistiu na retirada dos estudos devido ao primeiro critério de exclusão. Em seguida, foram filtrados por título e palavras-chave. Na quinta etapa de filtragem, os estudos foram removidos após a leitura do resumo (*abstract*). Por fim, os artigos duplicados foram eliminados.

No final do processo de filtragem, foi realizada uma leitura completa de cada artigo, observando os dois últimos critérios de exclusão (CE 2 e CE 3). A Figura 1 apresenta o processo de filtragem, com os Critérios de Inclusão e Critérios de Exclusão sendo aplicados em cada etapa. A base de dados IEEE Xplore foi a que resultou em menos trabalhos. Entretanto, essa base foi mais assertiva, considerando a porcentagem entre o número de artigos encontrados na pesquisa inicial até a junção deles após a leitura por *abstract*.

Quatro bases de dados tiveram resultados semelhantes na etapa da busca inicial, onde a quantidade de estudos variou entre 70 e 153 e ao longo da filtragem os números também foram próximos na porcentagem de estudos que permaneceram após as etapas. Considerando que as bases de dados MDPI e Scopus tiveram mais de 99,8% dos resultados filtrados ao longo do processo, conclui-se que, embora as bases tenham trazido mais resultados, quando comparadas com as demais bases de dados, as buscas originaram artigos menos aderentes ao tema proposto.

A [Tabela 3](#) apresenta a lista final dos 36 estudos analisados na última fase deste estudo. Os estudos estão dispostos em ordem cronológica, com o seu ID, seus autores junto de

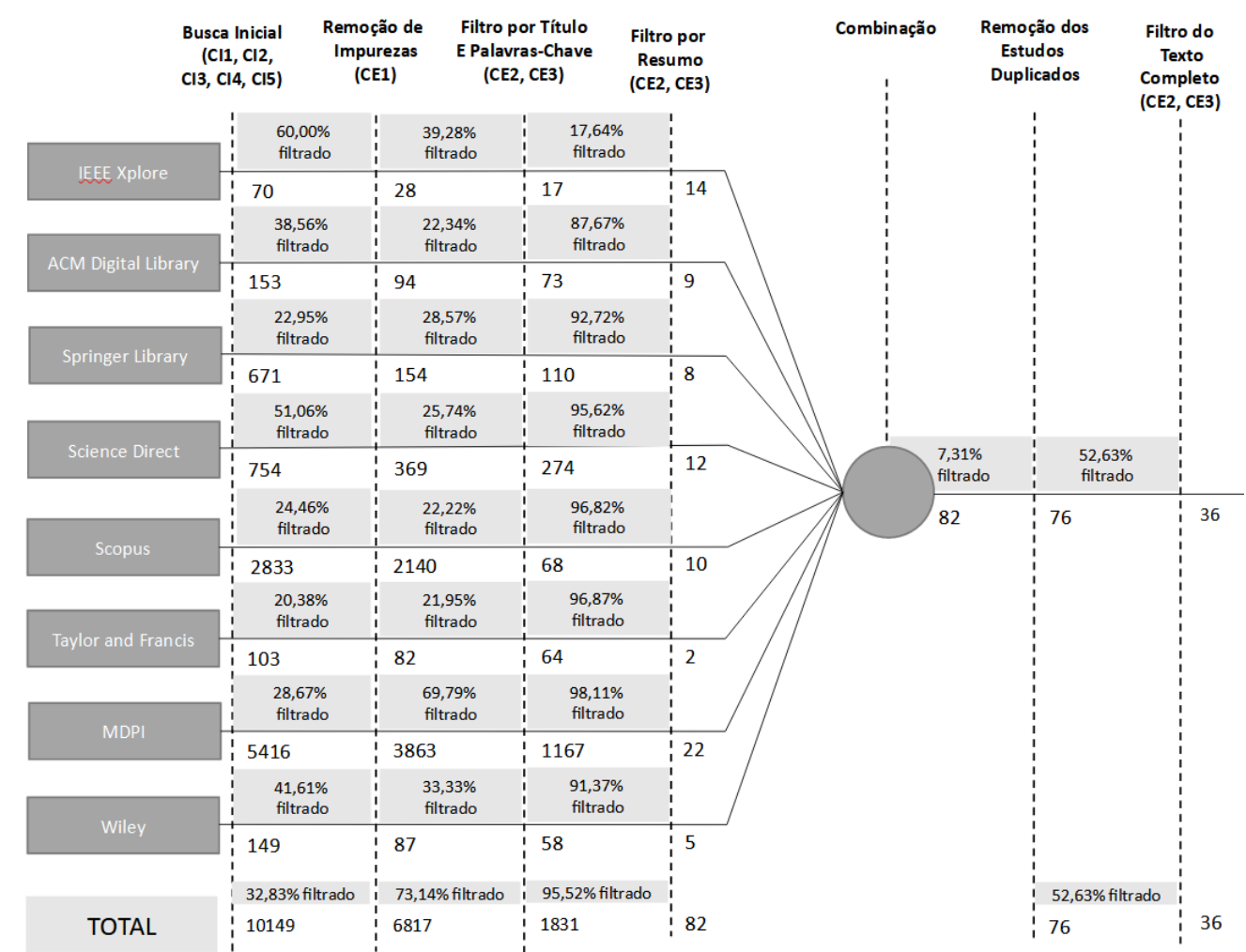


Figura 1: Processo de Filtragem.

suas respectivas referências e o H-Index, que é uma métrica utilizada para descrever o impacto dos periódicos e conferências na literatura. Essa disposição na [Tabela 3](#) facilita a compreensão da evolução da quantidade de trabalhos publicados até o final do ano de 2023. Os IDs dos artigos serão utilizados posteriormente nas figuras e tabelas para facilitar a compreensão de qual estudo o mapeamento está se referindo. A busca ocorreu até o fim do mês de março de 2024, porém não foi encontrado nenhum estudo que se adequasse ao mapeamento nos três primeiros meses desse ano.

4 Resultados

Nesta seção são apresentados os resultados obtidos através da leitura e análise dos 36 estudos selecionados no mapeamento. A seguir as questões de pesquisa foram respondidas através de considerações e análises dos trabalhos selecionados.

4.1 QG 1: Com qual finalidade os estudos estão utilizando o aprendizado de máquina na agricultura indoor?

Após a seleção dos 36 estudos do mapeamento, é possível constatar que o aprendizado de máquina é utilizado em diferentes áreas e com objetivos distintos para a agricultura indoor. O uso do aprendizado de máquina ocorre para a redução no consumo de energia ([Ajagekar e You, 2022](#)) dentro das estufas e também para prever o clima ([Li et al., 2022](#))) dentro delas 72 horas à frente, visando proporcionar melhores condições para o desenvolvimento das plantas e detectar doenças com certa antecedência. Assim como o estudo de [García-Vázquez et al. \(2023\)](#), que prevê a temperatura interna da estufa onde o cultivo está inserido. O aprendizado de máquina também é utilizado para auxiliar na melhor utilização das terras agrícolas, como mostra o estudo de [Yang et al. \(2018\)](#). O estudo de [Sharma et al. \(2021\)](#) fornece sugestões de qual cultura se adapta melhor às condições encontradas no ambiente e no solo da estufa.

A detecção de folhas queimadas ([Hamidon e Ahamed,](#)

Tabela 3: Relação dos Trabalhos Filtrados

ID	Ano	Autores e Referência	H-Index
A01	2018	Mehra et al. (2018)	149
A02	2018	Yang et al. (2018)	92
A03	2019	Vadivel et al. (2019)	10
A04	2020	Sadik et al. (2020).	1
A05	2020	Lauguico et al. (2020)	11
A06	2020	Wickramaarachchi et al. (2020)	26
A07	2021	Mya et al. (2021)	6
A08	2021	Ghandar et al. (2021)	204
A09	2021	Sharma et al. (2021)	14
A10	2021	Amir et al. (2021)	204
A11	2021	Chowdhury et al. (2021).	14
A12	2021	Martini et al. (2021)	219
A13	2021	Büyükoçkan et al. (2021)	148
A14	2021	Chang et al. (2021)	187
A15	2022	Liu et al. (2022)	149
A16	2022	Hwang et al. (2022).	52
A17	2022	Hamidon e Ahamed (2022)	219
A18	2022	Li et al. (2022)	149
A19	2022	Ajagekar e You (2022)	86
A20	2022	Dhal et al. (2022)	11
A21	2022	Franchetti e Pirri (2022)	187
A22	2022	Wang et al. (2022)	282
A23	2022	Soheli et al. (2022).	75
A24	2022	Zhang e Chen (2022)	98
A25	2022	Cui et al. (2022)	44
A26	2023	Kalavathi et al. (2023)	81
A27	2023	Abbasi et al. (2023)	18
A28	2023	Rajendiran e Rethnaraj (2023)	24
A29	2023	Parsa et al. (2023)	108
A30	2023	Mohmed et al. (2023)	282
A31	2023	Prasad et al. (2023)	51
A32	2023	Moorthi et al. (2023)	13
A33	2023	García-Vázquez et al. (2023)	101
A34	2023	Sikati e Nouaze (2023)	35
A35	2023	Wang et al. (2023)	219
A36	2023	Maya et al. (2023)	219

2022; Franchetti e Pirri, 2022) nas plantas junto ao cultivo e a detecção de doenças nas plantas (Chowdhury et al., 2021; Li et al., 2022; Liu et al., 2022; Sadik et al., 2020; Zhang e Chen, 2022; Abbasi et al., 2023) também ocorrem com a utilização do aprendizado de máquina. O estudo de Ghandar et al. (2021) prevê o crescimento das plantas semanalmente e dos peixes diariamente, já que o cultivo ocorre através da aquaponia e o sistema interliga o tanque dos peixes com o cultivo das plantas. Dez estudos controlam os cultivos, sendo seis deles com a utilização de aprendizado de máquina através de um ou mais índices nas estufas (Mehra et al., 2018; Wickramaarachchi et al., 2020; Martini et al., 2021; Kalavathi et al., 2023; Rajendiran e Rethnaraj, 2023; Maya et al., 2023) e outros quatro estudos na quantidade de nutrientes enviados nas irrigações para as plantas (Cui et al., 2022; Dhal et al., 2022; Mya et al., 2021; Vadivel et al., 2019).

O estudo de Martini et al. (2021), além de indicar melhores índices para o cultivo, também busca melhorar a produtividade dos cultivos ao tentar diminuir o tempo de cultivo das plantas e/ou aumentar o peso delas se comparado ao cultivo tradicional. Esse ganho de produtividade também é buscado pelos estudos de Chang et al. (2021);

Soheli et al. (2022); Cui et al. (2022). A estimativa de produção e contagem de frutos é desenvolvida no trabalho de Wang et al. (2022) através do crescimento natural das plantas, sem um controle sobre irrigações ou algo do gênero. O gerenciamento da planta através de informações obtidas por sensores é apresentado por Prasad et al. (2023), assim como Moorthi et al. (2023) fornecem a estrutura para o cultivo de cogumelos através da Internet das Coisas (Internet of Things (IoT)) e aprendizado de máquina.

Amir et al. (2021) utilizaram o aprendizado de máquina para prever o fluxo da seiva do tomate relacionando diferentes climas e índices de irrigação, já Lauguico et al. (2020) otimizaram a produção de alface prevenindo as características fitomorfológicas da planta enquanto Sikati e Nouaze (2023) buscam identificar as deficiências nutricionais nos alfaces. O estudo de Mohmed et al. (2023) propõe um modelo que prevê a resposta das plantas devido às condições ambientais em que se encontram, buscando melhorar o desempenho e a eficiência dos recursos.

Um estudo está focado no melhor aproveitamento das estufas (Hwang et al., 2022) e o estudo de Büyükoçkan et al. (2021) em avaliar as tecnologias instaladas nelas. A criação de robôs autônomos para executarem a colheita diretamente nos cultivos apareceu nos estudos de Parsa et al. (2023) e Wang et al. (2023), esses estudos tiveram seu foco na colheita de morango e de tomates, respectivamente.

4.2 QEP 1: Quais as técnicas de aprendizado de máquina estão sendo utilizadas?

A Fig. 2 apresenta uma taxonomia que mostra a variedade de algoritmos de aprendizado de máquina utilizados nos estudos selecionados. Foram encontradas 43 técnicas de aprendizado de máquina aplicadas à agricultura indoor. Esta taxonomia organiza as técnicas em suas principais categorias que são: métodos tradicionais, redes neurais e aprendizagem profunda. Junto do nome de cada técnica utilizada, são mostrados os IDs dos estudos que foram listados na Tabela 3. Os métodos tradicionais representam o maior grupo e, portanto, esta categoria contém subcategorias baseadas na estrutura das técnicas, como Baseado em Árvore (*Tree Based*), Linear, *Clustering*, *Ensemble*, Lógica Fuzzy (*Fuzzy Logic*), Aprendizado Supervisionado (*Supervised Learning*) e *Bayes*.

Entre os algoritmos mais utilizados em cada subcategoria de métodos tradicionais, a Árvore de Decisão (*Decision Tree* (DT)) foi a mais comum nos algoritmos baseados em Árvore, pois foi utilizado em seis estudos, sendo quatro utilizando diretamente DT (Amir et al., 2021; Liu et al., 2022; Sharma et al., 2021; Yang et al., 2018) e mais dois estudos (Dhal et al., 2022; Ghandar et al., 2021) que utilizaram a subcategoria CART. A técnica Máquina de Vetores de Suporte (*Support Vector Machine* (SVM)) foi a mais utilizada na categoria Linear por ser utilizada em 7 estudos, sendo seis deles (Amir et al., 2021; Dhal et al., 2022; Ghandar et al., 2021; Yang et al., 2018; Zhang e Chen, 2022; Kalavathi et al., 2023) diretamente com SVM e o estudo de García-Vázquez et al. (2023) que utilizou a técnica de Suporte à Regressão Vetorial (*Support Vector Regression* (SVR)).

Importante salientar que a SVM foi seguida de perto pela Regressão Logística (*Linear Regression* (LR)) que foi utili-

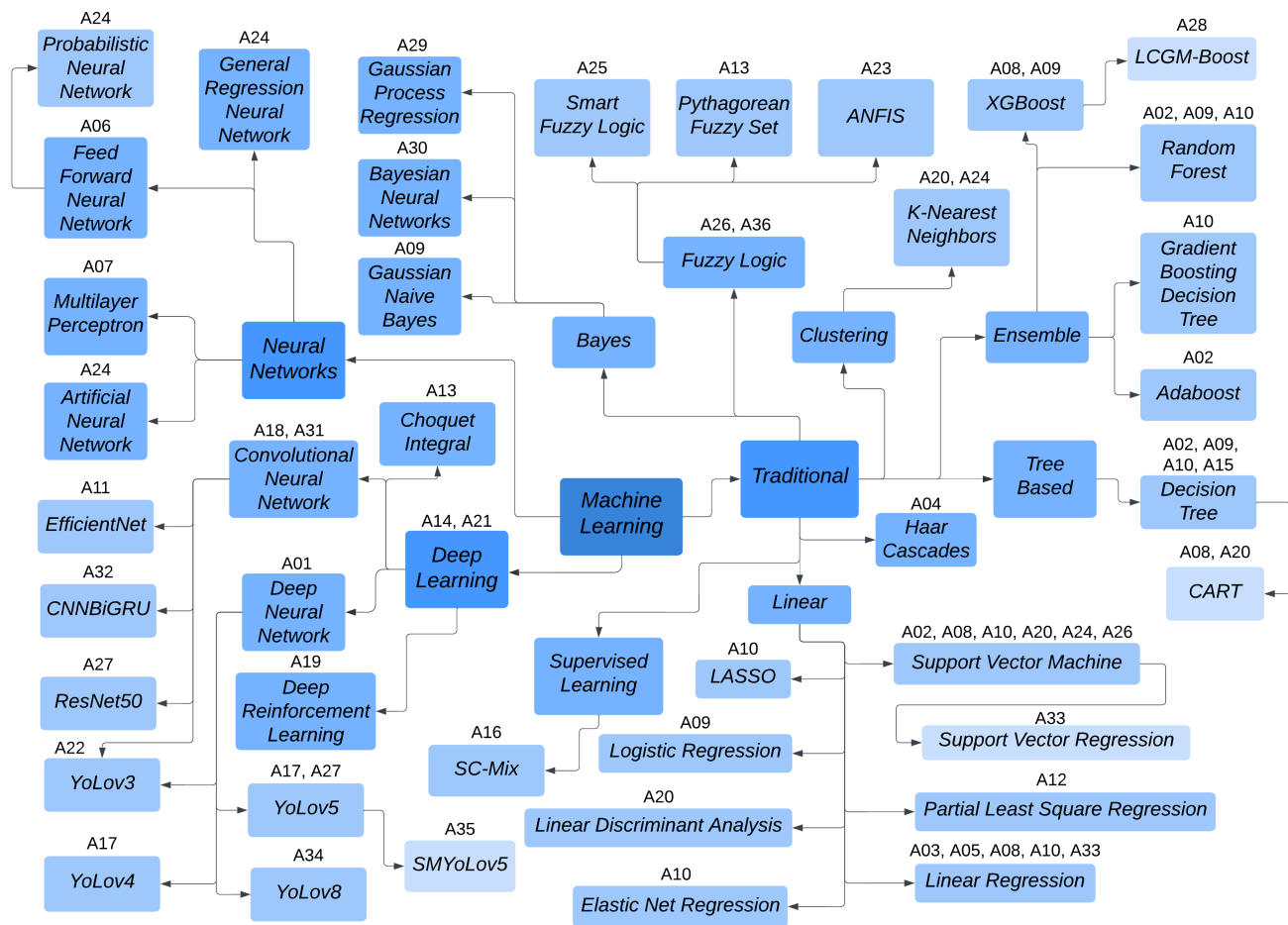


Figura 2: Técnicas utilizadas de aprendizado de máquina.

zado em 5 estudos (Amir et al., 2021; Ghandar et al., 2021; Lauguico et al., 2020; Vadivel et al., 2019; García-Vázquez et al., 2023). A subcategoria, *K-Nearest Neighbours* (KNN) foi a única utilizada na subcategoria *Clustering*, Floresta Aleatória (*Random Forest* (RF)) apareceu em mais estudos referentes aos algoritmos *Ensemble* seguido pela técnica XGBoost que foi utilizada em três estudos, sendo dois diretamente (Ghandar et al., 2021; Sharma et al., 2021) e o terceiro através de uma melhoria do XGBoost através do trabalho de Rajendiran e Rethnaraj (2023) que utilizou a técnica denominada monitoramento acelerado do crescimento da colheita de alface (*Lettuce Crop Growth Monitoring-Boost* (LCGM-Boost)). Na categoria Redes Neurais a subcategoria *Feed Forward Neural Network* foi utilizada duas vezes, visto que ela foi aplicada por Wickramaarachchi et al. (2020) e a sua subcategoria Redes Neurais Probabilísticas (*Probabilistic Neural Network*) foi utilizada no estudo de Zhang e Chen (2022).

Existe uma estreita relação entre as categorias redes neurais e aprendizagem profunda, uma vez que os modelos de aprendizagem profunda consistem em redes neurais com diversas camadas. No entanto, esta taxonomia separa essas duas categorias para destacar os grupos de técnicas

de aprendizado de máquina mais utilizados nestes campos de pesquisa. Em relação aos algoritmos de *Deep Learning*, Redes Neurais Profundas (*Deep Neural Network* (DNN)) apareceram em 6 estudos, porém em 6 subcategorias. A utilização de DNN ocorreu com um estudo (Mehra et al., 2018) utilizando diretamente DNN e outros cinco estudos (Hamidon e Ahamed, 2022; Wang et al., 2022; Abbasi et al., 2023; Sikati e Nouaze, 2023; Wang et al., 2023) que utilizaram as técnicas YoLov3, YoLov4, YoLov5 e seu aprimoramento SMYoLov5 e YoLov8.

A Fig. 2 mostra que alguns estudos utilizam mais de uma técnica de aprendizado de máquina, como o estudo de Amir et al. (2021) que utiliza sete categorias de aprendizado de máquina sendo elas: *Linear Regression*, *Elastic Net Regression*, *Least Absolute Shrinkage and Selection Operator* (LASSO), SVM, DT, *Gradient Boosting Decision Tree*, *Random Forest*. Outros estudos que também utilizaram uma maior quantidade de técnicas são os estudos de Sharma et al. (2021) com cinco técnicas e Ghandar et al. (2021) com quatro técnicas.

4.3 QEP 2: Quais os dados estão sendo utilizados nos estudos?

A [Tabela 4](#) apresenta as informações obtidas nos cultivos e utilizadas nos estudos. Os tipos de informações obtidas são bastante diversificados, totalizando 26 tipos. A informação mais utilizada nos cultivos foi a temperatura do ambiente, aplicada em 52,77% (19/36) dos estudos. A segunda informação mais utilizada foram as imagens/fotos do cultivo, aplicadas em 14 estudos, ou seja, 38,88% dos trabalhos obtêm informações adicionais por meio de imagens dos cultivos.

A terceira informação mais coletada é a umidade do ambiente, utilizada em 36,11% (13/36) dos trabalhos. É importante salientar que, para evitar dúvidas ou discrepâncias entre os estudos que utilizavam irrigação apenas com água e aqueles que utilizavam solução nutritiva (água junto com os nutrientes das plantas), todos os estudos são indicados na tabela como "água", referindo-se, por exemplo, ao pH ou temperatura da água. Os dois trabalhos que mais utilizaram diferentes informações foram os estudos de [Amir et al. \(2021\)](#) e [Martini et al. \(2021\)](#). O primeiro estudo utilizou 7 informações do cultivo que foram pH da água, temperatura e umidade do ambiente, nível de CO₂, Condutividade Elétrica (*Electric conductivity* (EC)) da Água, fluxo da seiva e temperatura da planta. Já o estudo de [Martini et al. \(2021\)](#) utilizou 8 informações que foram: pH da água, temperatura da água, temperatura e umidade do ambiente, nível do tanque, EC da água, tempo de cultivo e quantidade de plantas.

4.4 QEP 3: Os trabalhos utilizam estratégias de fusão de informações de datasets?

Entre os 36 estudos desse mapeamento não foi possível constatar trabalhos que suportem fusões com *datasets* de outros cultivos e/ou estudos. Os trabalhos sempre são voltados para um *dataset* próprio e normalmente apenas para um tipo de cultivo. O estudo de [Franchetti e Pirri \(2022\)](#) utilizou *datasets* de outros cultivos para testes de seu sistema, mas não é especificado que o sistema suporte automaticamente outros *datasets*, assim como o estudo de [Abbasi et al. \(2023\)](#) que utilizou três *datasets* diferentes, porém apenas para treinamento e testes do seu modelo. Com isso é necessário implantar e treinar manualmente qualquer dataset novo. O importante nessa situação seria se o sistema como um todo aceitasse *datasets* de outros cultivos. Além disso seria relevante se o sistema treinasse com novos *datasets* automaticamente e já fornecesse indicações de melhorias caso não seja seguro aplicar diretamente nos cultivos.

Quanto a suportar mais de um tipo de cultivo, apenas dois estudos são capazes disso. O estudo de [Martini et al. \(2021\)](#) suportou os cultivos de hidroponia e de semi-hidroponia, também conhecido por gotejamento em alguns locais. Já o estudo de [Vadivel et al. \(2019\)](#) suportou os cultivos em hidroponia e o cultivo tradicional no solo, porém esse cultivo no solo como já mencionado anteriormente é dentro de estufas.

4.5 QEP 4: Quais estudos utilizam informações de contexto?

Contexto e históricos de contextos ([Aranda et al., 2021](#); [Filippetto et al., 2021](#)) estiveram presentes em apenas um estudo encontrado nesta revisão. O estudo de [Martini et al. \(2021\)](#) além de utilizar o contexto, também montou sua arquitetura voltada para ele, em sua base de dados existe um módulo para os históricos dos contextos e em seu servidor existe um módulo para comparar a similaridade dos contextos e não enviar informações repetidas e/ou desnecessárias para o seu banco de dados.

Apenas o estudo de [Martini et al. \(2021\)](#) utilizou o conceito de computação ubíqua por meio do uso de sensores nas estufas para coleta de informações dos cultivos. Essa coleta acontece de forma automática e a cada 15 minutos. [Martini et al. \(2021\)](#) desenvolveu esse módulo por se tratar de cultivo em estufa, onde o clima não muda tão rapidamente como a céu aberto, já que com as cortinas e plásticos fechados o clima muda gradativamente.

Mais investigações são necessárias para avaliar como as informações de contexto podem melhorar o desempenho dos modelos já que apenas um estudo utilizou o contexto aplicado ao cultivo. O armazenamento dessas informações permite a previsão do contexto ([Rosa et al., 2015](#)), assim, os modelos podem avaliar os momentos em que os cultivos estão atualmente e quais índices estão atuando sobre eles. O estudo de [Liu et al. \(2022\)](#) prevê o clima dentro das estufas 72 horas a frente, porém a predição é utilizada apenas para detectar possíveis doenças nas plantas, o estudo não utiliza os conceitos de contexto e históricos de contextos.

4.6 QEP 5: Quais os dispositivos utilizados para obter a coleta de informações dos cultivos?

Os artigos foram classificados de acordo com o tipo de dispositivos que utilizam em seus sistemas para obtenção das informações dos cultivos. Nos estudos, os dois microcontroladores mais utilizados são Raspberry Pi e Arduino. No entanto, o dispositivo mais utilizado foi a câmera para obter imagens do cultivo. A [Tabela 5](#) apresenta que 36,1% dos estudos (13/36) utilizaram câmeras para obter imagens dos cultivos e, posteriormente, aplicaram visão computacional para extrair informações importantes. Além disso, a [Tabela 5](#) também mostra que o Arduino foi utilizado em 22,2% (8/36) dos trabalhos, enquanto o Raspberry Pi foi utilizado em 3 trabalhos, representando 8,33% do total.

Alguns estudos ([Ajagekar e You, 2022](#); [Amir et al., 2021](#); [Büyükközkcan et al., 2021](#); [Dhal et al., 2022](#); [Liu et al., 2022](#); [Sharma et al., 2021](#); [Yang et al., 2018](#); [Wickramaarachchi et al., 2020](#); [Kalavathi et al., 2023](#); [Rajendiran e Rethnaraj, 2023](#); [García-Vázquez et al., 2023](#)) não informaram quais os dispositivos foram utilizados para obter as informações. Por outro lado, os estudos de [Cui et al. \(2022\)](#), [Mehra et al. \(2018\)](#) e [Mya et al. \(2021\)](#) que utilizaram Raspberry sempre aplicaram Arduino junto para obtenção das informações.

Tabela 4: Relação das Informações Utilizadas nos Estudos

Informações	Qtd. de Estudos (%)	ID dos Estudos
PH da Água	10 (27,7%)	A01, A03, A07, A08, A09, A10, A12, A20, A28, A36
Temperatura da Água	4 (11,1%)	A03, A08, A12, A28
Temperatura do Ambiente	19 (52,7%)	A01, A02, A03, A06, A07, A08, A10, A12, A15, A19, A23, A24, A25, A26, A30, A31, A32, A33, A36
Umidade do Ambiente	13 (36,1%)	A01, A02, A03, A09, A10, A12, A23, A24, A25, A26, A31, A32, A33
Nível do Tanque	3 (8,3%)	A01, A07, A12
Luminosidade	4 (11,1%)	A01, A02, A23, A30
Umidade do Solo	9 (25%)	A03, A23, A24, A25, A26, A30, A31, A32, A36
CO2	4 (11,1%)	A06, A10, A30, A32
Imagens(Fotos)	14 (38,8%)	A04, A05, A11, A14, A16, A17, A18, A21, A22, A27, A29, A31, A34, A35
Altura da Planta	1 (2,7%)	A06
EC da Água	6 (16,6%)	A07, A10, A12, A20, A28, A36
TDS da Água	1 (2,7%)	A07
ORP da Água	1 (2,7%)	A07
Ração	1 (2,7%)	A08
Composição do Solo	1 (2,7%)	A09
Chuva	1 (2,7%)	A09
Fluxo da Seiva	1 (2,7%)	A10
Temperatura da Planta	1 (2,7%)	A10
Tempo de Cultivo	1 (2,7%)	A12
Quantidade de Plantas	1 (2,7%)	A12
Dados Econômicos	1 (2,7%)	A13
Umidade das Folhas	1 (2,7%)	A15
PPM da Água	1 (2,7%)	A28
Turbidez da Água	1 (2,7%)	A28
Ponto de Orvalho	1 (2,7%)	A33
Radiação Solar	1 (2,7%)	A33

Tabela 5: Relação dos Dispositivos Utilizadas nos Estudos

Dispositivos	Quantidade de Estudos (%)	ID dos Estudos
Arduino	8 (22,2%)	A01, A03, A07, A12, A23, A24, A31, A36
Raspberry	3 (8,3%)	A01, A07, A12
ESP8266	3 (8,3%)	A08, A25, A32
Câmera	13 (36,1%)	A04, A05, A11, A14, A16, A17, A18, A21, A22, A27, A29, A34, A35

4.7 QEP 6: Quais as culturas de plantas estão sendo cultivadas nos estudos?

A Tabela 6 apresenta todas as culturas que foram cultivadas nos estudos do mapeamento. É possível ver que duas culturas se destacam como as mais utilizadas que são o tomate e a alface. O tomate como mostra a Tabela 6 foi

utilizado por 10 estudos, ou seja, 27,7% de todos os trabalhos. Já a alface foi cultivada em 9 estudos, totalizando 25% dos trabalhos. Importante também notar que dois estudos cultivaram quatro tipos de plantas, o estudo de Abbasi et al. (2023) cultivou alface, manjerição, espinafre e salsa, enquanto o estudo de Prasad et al. (2023) cultivou uva, tomate, maçã e batata. Já o estudo de Martini et al. (2021) cultivou 3 culturas de plantas que foram: rúcula, radiche e morango.

4.8 QET 1: Onde as pesquisas foram publicadas?

A Fig. 3 apresenta a distribuição dos artigos selecionados quanto à fonte de publicação. Os periódicos representam 75% (27/36) do total, enquanto as conferências representam os outros 25% (9/36). A publicação dos primeiros estudos sobre aprendizado de máquina na agricultura *indoor* ocorreu em periódicos em 2018, e em 2019 surgiram os primeiros estudos publicados em conferências. Nos anos de 2019 e 2020, não foram publicados estudos em periódicos; no entanto, em 2021 e 2022, houve um crescimento consi-

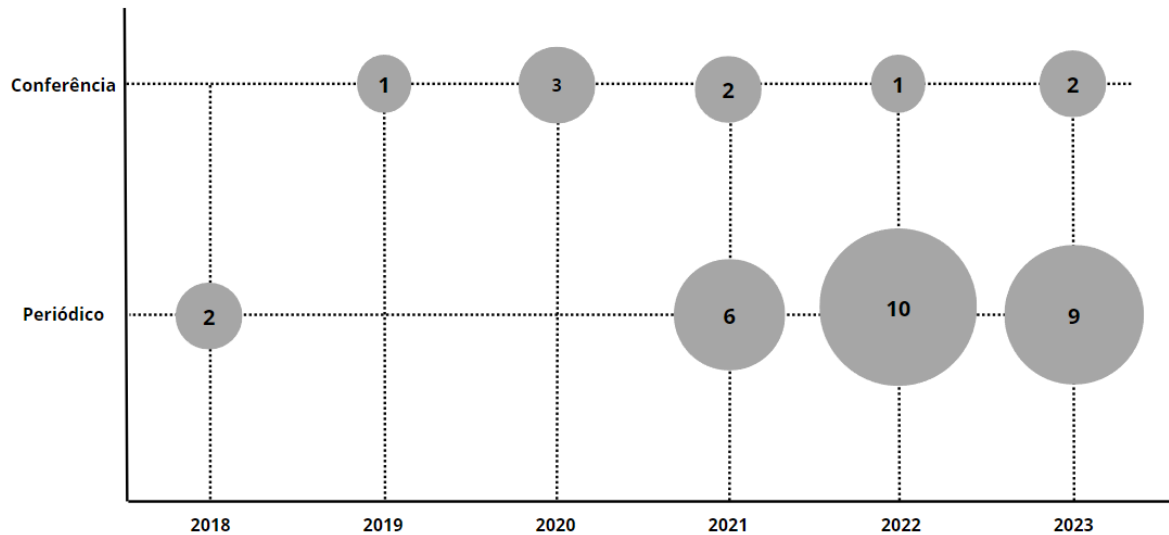


Figura 3: Total de Artigos por Tipo de Publicação e Ano.

Tabela 6: Relação das Culturas das Plantas Cultivadas nos Estudos

Culturas	Quantidade de Estudos (%)	ID dos Estudos
Pepino	1 (2,7%)	A15
Alface	9 (25,0%)	A05, A07, A16, A17, A20, A21, A27, A30, A34
<i>Tuberosa Bulbes</i>	1 (2,7%)	A08
Coentro	1 (2,71%)	A03
Tomate	10 (27,7%)	A01, A04, A06, A10, A11, A22, A24, A31, A33, A35
Rúcula	1 (2,7%)	A12
Morango	3 (8,3%)	A12, A29, A36
Radiche	1 (2,7%)	A12
Arabidopsis	1 (2,7%)	A14
Pimenta	1 (2,7%)	A25
Manjeriço	1 (2,7%)	A27
Espinafre	1 (2,7%)	A27
Salsa	1 (2,7%)	A27
Uva	1 (2,7%)	A31
Maça	1 (2,7%)	A31
Batata	1 (2,7%)	A31
Cogumelo	1 (2,7%)	A32

derável de publicações. Esse crescimento não continuou de 2022 para 2023, quando a quantidade de publicações em periódicos diminuiu de 10 para 9. A quantidade de publicações em conferências não variou consideravelmente ao longo dos anos. Além da primeira publicação ter ocorrido em 2019, houveram ainda três publicações em 2020, duas

em 2021, uma em 2022 e duas em 2023.

4.9 QET 2: Quantas publicações ocorreram por ano?

O mapeamento utilizou os estudos que possuíssem conexão de aprendizado de máquina com a agricultura *indoor* sem uma data limítrofe inicial para as buscas, porém após o processo de filtragem ficou constatado que somente a partir do ano de 2018 ocorreram estudos nessa área. A Fig. 4 apresenta a distribuição dos artigos de acordo com o ano de publicação e base de dados. É possível constatar que a quantidade de estudos teve um crescimento relevante nos anos de 2021 e 2022. No ano de 2023 a quantidade de artigos publicados se manteve o que mostra que a área é promissora e está em desenvolvimento no momento.

Os trabalhos foram agrupados na Fig. 4 de acordo com as oito bases utilizadas na pesquisa e o ano de publicação. A base de dados que obteve o maior número de trabalhos foi a IEEE Xplore, com 9 artigos de um total de 36, ou seja, 25%. Em segundo lugar, houve um empate entre duas bases de dados, Science Direct e MDPI, ambas as bases retornaram 8 estudos cada, o que significa que cada uma delas contribuiu com 22,2% da totalidade de estudos do mapeamento. A base Springer Library retornou 5 estudos, os quais correspondem a 13,8% do total. A base Scopus resultou em 3 estudos (8,3%) selecionados, já a base Wiley resultou em 2 estudos (5,5%) e a base de dados ACM Digital Library retornou apenas 1 estudo (2,7%). Por fim, a base Taylor & Francis não retornou nenhum estudo após passar por todos os critérios de inclusão e exclusão e leituras, por isso não foi inserida a legenda da base na Fig. 4.

Além de apresentar quais bases de dados se destacaram mais no mapeamento ao resultarem numa quantidade maior de estudos selecionados, também é importante informar quais periódicos se sobressaíram com uma quantidade maior de estudos na área. Em primeiro lugar ficou o periódico *Sensors* com 4 estudos publicados (Martini et al.,



Figura 4: Quantidade de publicações a cada ano e suas respectivas bases.

2021; Hamidon e Ahamed, 2022; Wang et al., 2023; Maya et al., 2023). No segundo lugar com mais estudos selecionados ficou o periódico *Computers and Electronics in Agriculture* com 3 estudos (Mehra et al., 2018; Li et al., 2022; Liu et al., 2022).

5 Ameaças à Validade do Trabalho

Os trabalhos de mapeamento sistemático estão expostos a riscos que podem invalidar o estudo. Esses riscos podem ter origem em decisões tomadas equivocadamente durante o mapeamento sistemático. Com o objetivo de garantir melhores resultados, foram selecionadas oito bases de dados, conhecidas na área acadêmica por sua relevância na área de ciência da computação e afins. Portanto, procurou-se mitigar a possibilidade das bases de dados impactarem nos resultados. Foram utilizados dois termos principais classificados como termos primários e seus respectivos termos secundários. Como resultado da string buscou-se obter estudos que tivessem relação com um dos oito termos secundários relacionados com os termos primários *Machine Learning* e *Indoor Agriculture*. Assim, procurou-se a obtenção do maior número de estudos nos resultados de busca. O mapeamento não considerou aspectos conceituais da área de Agricultura, como estudos aplicados somente na área de ciências da natureza, atendo-se exclusivamente na aplicação do Aprendizado de Máquina na Agricultura Indoor. Esta decisão contribui com a filtragem, direcionando a busca apenas a um contexto específico dentro da área da "Agricultura, Aprendizado de Máquina e Ciência da Computação".

O processo de filtragem pode restringir os trabalhos

de forma que algum estudo relevante possa ser removido. Para mitigar esse risco, o processo de filtragem foi baseado em uma técnica amplamente adotada em trabalhos acadêmicos de mapeamento sistemático (Petersen et al., 2015). Além disso, foram considerados aspectos de processos de revisão já utilizados por outros autores (Heckler et al., 2021; Micol et al., 2021).

6 Discussão

O limite de terras agrícolas e a escassez de comida são dois fatores importantes a serem considerados junto do aumento populacional do mundo. Por isso, a aplicação de aprendizado de máquina na agricultura indoor tem potencial para ajudar nesse cenário por meio do aumento de produtividade. Assim, o ganho de produtividade permite que mais alimentos sejam produzidos nas mesmas áreas de cultivos.

Foram encontradas 43 técnicas de aprendizado de máquina sendo aplicadas na agricultura indoor buscando diferentes objetivos. Entre os resultados que são buscados com a aplicação dessas técnicas é possível citar: a detecção de pragas nos cultivos, redução do consumo de energia, prever o clima 72 horas a frente, controlar os índices das estufas e irrigações entre outras possibilidades.

Apesar de terem bons desempenhos, os modelos de aprendizado de máquina são treinados com *datasets* específicos de um certo tipo de informações. Dificilmente ocorre um treinamento do aprendizado de máquina com *datasets* que possuam informações diferentes das que o sistema já utilizou. Isso pode limitar a utilização dos sistemas para funções específicas, já que o sistema não compreende

e não sabe como utilizar informações de outros tipos.

Quanto aos tipos de cultivo é interessante compreender que os estudos foram aplicados para 5 diferentes tipos de cultivo, que são: hidroponia, semi-hidroponia, cultivo no solo, aquaponia e aeroponia. A hidroponia é um modelo de cultivo da agricultura que funciona através da recirculação de água que possui os nutrientes necessários para as plantas. As plantas não ficam em contato com nada além dos perfis de PVC e da solução nutritiva que passa pelos perfis, ou seja, as plantas não ficam em contato com a terra e/ou substrato em nenhum momento. No total nove estudos, ou seja, 25% de todos os artigos foram aplicados na hidroponia, sendo o estudo de Mehra et al. (2018) o mais antigo desse mapeamento. A semi-hidroponia é o modelo de cultivo que as plantas são cultivadas em vasos ou plásticos e a irrigação delas não retorna, a água e seus nutrientes ficam depositados no substrato onde as plantas são cultivadas.

O cultivo no solo que é o mais tradicional na agricultura convencional esteve presente em 52,77% dos estudos, 19 dos 36 estudos tiveram as plantas sendo cultivadas no solo. O último tipo de cultivo e diferente do método tradicional do solo foi a aquaponia, que possui o mesmo princípio que a hidroponia, porém, em vez de serem adicionados nutrientes na água de circulação existe uma criação de peixes que trabalham em conjunto com o cultivo das plantas. Os peixes ficam localizados em um tanque separado e geram os nutrientes para as plantas através de seus excrementos. Esse tipo de cultivo foi encontrado em três estudos (Dhal et al., 2022; Ghandar et al., 2021; Lauguico et al., 2020). Por fim, no trabalho de Rajendiran e Rethnaraj (2023) o tipo de cultivo utilizado foi a aeroponia. A aeroponia é quando as raízes das plantas ficam suspensas dentro de um reservatório onde no fundo do mesmo existe uma solução nutritiva e devido a umidade da solução existente as raízes e as plantas vão crescendo.

Faltam modelos mais gerais que suportem diferentes informações e que sejam aplicados a uma quantidade maior de cultivos. Apenas 3 estudos suportam 2 tipos de cultivos (Martini et al., 2021; Vadivel et al., 2019; Soheli et al., 2022). A integração de outros datasets para treinamento dos estudos também passa a ter importância, já que as técnicas de aprendizado de máquina normalmente são treinadas com poucas informações.

7 Conclusão

Este estudo apresentou um mapeamento sistemático da literatura sobre estudos que utilizam aprendizado de máquina na agricultura indoor. A lista final contém 36 estudos, selecionados por um processo de filtragem que iniciou com 10149 trabalhos de oito bases de dados. Foi realizada a leitura completa desses artigos para explorar as nove questões de pesquisa. Este mapeamento abrangeu domínios de aplicação, bem como quais dados dos cultivos são utilizados, técnicas de aprendizado de máquina, dispositivos utilizados para obtenção das informações, quais informações são utilizadas dos cultivos, os tipos de cultivo que estão sendo aplicados nos estudos, quais as culturas de plantas estão sendo produzidas e outras informações. Por fim, avaliou-se os desafios e direções futuras para pesquisas na área do aprendizado de máquina na agricultura

indoor.

A utilização do aprendizado de máquina na agricultura indoor é uma tendência nos estudos atualmente já que é possível criar o microclima para as plantas dentro das estufas e além disso as plantas não sofrem com as variações climáticas que podem ocorrer se estivessem sendo cultivadas a céu aberto. De forma geral, os estudos abordaram a aplicação do aprendizado de máquina em tipos de cultivos específicos e realizaram as análises voltadas apenas para as informações que eles possuíam de cada cultivo para obter uma certa finalidade após o processamento das informações.

Trabalhos futuros deverão investigar modelos genéricos e proativos que não sejam focados apenas em um certo cultivo. A fusão de informações com datasets de outros cultivos ou com informações diferentes das já conhecidas é um desafio de pesquisa já que os modelos atuais combinam dados diferentes apenas do seu cultivo. Novos modelos poderão explorar a fusão com outros datasets e com outras informações de cultivos, para que seja possível o sistema fornecer informações baseados em uma quantidade maior de cultivos e até com informações que antes nem eram cogitadas nos cultivos.

Outro ponto a ser levado em consideração como trabalho futuro é a possibilidade de criar um sistema que execute um autotreinamento para sempre ficar atualizado e fornecer as melhores diretrizes para o cultivo. Um ponto preocupante e que deve ser considerado é o sistema tomar decisões de mudanças automaticamente já que as tecnologias e informações geradas ainda são recentes, porém seria interessante se o sistema sugerisse melhorias e o agrônomo responsável aprovasse elas. Revisões futuras poderão ampliar este estudo considerando outras técnicas de inteligência artificial e o cultivo na agricultura outdoor. Finalmente existem oportunidades para analisar quais são os atuais desafios na aplicação de aprendizado de máquina na agricultura indoor.

Agradecimentos

Os autores agradecem a CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior, Código de Financiamento 001) e CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico). Agradecemos também à Unisinos (Universidade do Vale do Rio dos Sinos) pelo apoio a este trabalho.

Referências

- Abbasi, R., Martinez, P. e Ahmad, D. R. (2023). Crop diagnostic system: A robust disease detection and management system for leafy green crops grown in an aquaponics facility, *SSRN Electronic Journal* pp. 1–12. <https://doi.org/10.2139/ssrn.4188680>.
- Ajagekar, A. e You, F. (2022). Deep reinforcement learning based automatic control in semi-closed greenhouse systems, *IFAC-PapersOnLine* 55: 406–411. <https://doi.org/10.1016/j.ifacol.2022.07.477>.
- Akis, S. (2017). What is smart farming?, *Technical report*.

- Available at <https://www.smart-akis.com/index.php/network/what-is-smart-farming/>.
- Altalak, M., uddin, M., Alajmi, A. e Rizg, A. (2022). Smart agriculture applications using deep learning technologies: A survey, *Applied Sciences* 12: 5919. <https://doi.org/10.3390/app12125919>.
- Amir, A., Butt, M. e Kooten, O. (2021). Using machine learning algorithms to forecast the sap flow of cherry tomatoes in a greenhouse, *IEEE Access* 9: 1–1. <https://doi.org/10.1109/ACCESS.2021.3127453>.
- Anbananthen, K., S, S., Chelliah, D., Sivakumar, P., Somasundaram, V., Velshankar, K. e Khan, A. (2021). An intelligent decision support system for crop yield prediction using hybrid machine learning algorithms, *F1000Research* 10: 1143. <https://doi.org/10.12688/f1000research.73009.1>.
- Aranda, J., Barbosa, J., Bavaresco, R., Carvalho, J., Tavares, M. e Yamin, A. (2021). A computational model for adaptive recording of vital signs through context histories, *Journal of Ambient Intelligence and Humanized Computing*. <https://doi.org/10.1007/s12652-021-03126-8>.
- Avila, F. R. d. e Barbosa, J. L. V. (2024). Ambientes inteligentes na agricultura digital: uma revisão sistemática e taxonomia, *Revista Brasileira de Computação Aplicada* 16(1): 11–25. <https://doi.org/10.1016/j.compap.2020.105602>.
- Büyüközkan, G., Göçer, F. e Uztürk, D. (2021). A novel pythagorean fuzzy set integrated choquet integral approach for vertical farming technology assessment, *Computers and Industrial Engineering* 158: 107384. <https://doi.org/10.1016/j.cie.2021.107384>.
- Chang, S., Lee, U., Hong, M., Jo, Y. e Kim, J.-B. (2021). Time-series growth prediction model based on u-net and machine learning in arabidopsis, *Frontiers in Plant Science* 12. <https://doi.org/10.3389/fpls.2021.721512>.
- Chowdhury, M., Rahman, T., Khandakar, A., Ayari, M., Khan, A., Khan, M. S., Al-Emadi, N., Reaz, M. B. I., Islam, M. e Ali, S. (2021). Automatic and reliable leaf disease detection using deep learning techniques, *AgriEngineering* 3: 294–312. <https://doi.org/10.3390/agriengineering3020020>.
- Cui, H., Hanafi, M., Ilahi, W., Zamri, M., Shafie, S. e Mashohor, S. (2022). The effect of smart fertigation systems on chilli grown in a greenhouse for urban farming, *Irrigation and Drainage* 71. <https://doi.org/10.1002/ird.2709>.
- Dey, A. K. (2001). Understanding and using context, *Personal and Ubiquitous Computing* 5: 4–7. <https://doi.org/10.1007/s007790170019>.
- Dhal, S., Bagavathiannan, M., Braga-Neto, U. e Kalafatis, S. (2022). Nutrient optimization for plant growth in aquaponic irrigation using machine learning for small training datasets, *Artificial Intelligence in Agriculture* 6: 68–76. <https://doi.org/10.1016/j.aiaa.2022.05.001>.
- Doshi, J., Patel, T. e Bharti, D. (2019). Smart farming using iot, a solution for optimally monitoring farming conditions, *Procedia Computer Science* 160: 746–751. <https://doi.org/10.1016/j.procs.2019.11.016>.
- FAO (2017). Water for sustainable food and agriculture—a report produced for the g20 presidency of germany, *Technical report*, Berlim, Alemanha. Available at <http://www.fao.org/3/i7959e/i7959e.pdf>.
- Filippetto, A., Lima, R. e Barbosa, J. (2021). A risk prediction model for software project management based on similarity analysis of context histories, *Information and Software Technology* 131: 106497. <https://doi.org/10.1016/j.infsof.2020.106497>.
- Franchetti, B. e Pirri, F. (2022). Detection and localization of tip-burn on large lettuce canopies, *Frontiers in Plant Science* 13: 874035. <https://doi.org/10.3389/fpls.2022.874035>.
- García-Vázquez, F., Ponce-González, J., Guerrero-Osuna, H., Carrasco-Navarro, R., Luque-Vega, L., Mata-Romero, M., Martínez-Blanco, M. D. R., Castañeda-Miranda, C. e Díaz-Flórez, G. (2023). Prediction of internal temperature in greenhouses using the supervised learning techniques: Linear and support vector regressions, *Applied Sciences* 13: 1–23. <https://doi.org/10.3390/app13148531>.
- Ghandar, A., Ahmed, A., Zulfiqar, S., Hua, Z., Hanai, M. e Theodoropoulos, G. (2021). A decision support system for urban agriculture using digital twin: A case study with aquaponics, *IEEE Access* PP: 1–1. <https://doi.org/10.1109/ACCESS.2021.3061722>.
- Hamidon, M. H. e Ahamed, T. (2022). Detection of tip-burn stress on lettuce grown in an indoor environment using deep learning algorithms, *Sensors* 22: 7251. <https://doi.org/10.3390/s22197251>.
- Hassan, S., Alam, M., Illahi, U., Al Ghamdi, M., Almotiri, S. e Mazliham, M. (2021). A systematic review on monitoring and advanced control strategies in smart agriculture, *IEEE Access* PP: 1–1. <https://doi.org/10.1109/ACCESS.2021.3057865>.
- Heckler, W., Carvalho, J. e Barbosa, J. (2021). Machine learning for suicidal ideation identification: A systematic literature review, *Computers in Human Behavior* 128: 107095. <https://doi.org/10.1016/j.chb.2021.107095>.
- Helfer, G., Barbosa, J., Santos, R. e Costa, A. (2020). A computational model for soil fertility prediction in ubiquitous agriculture, *Computers and Electronics in Agriculture* 175: 105602. <https://doi.org/10.1016/j.compap.2020.105602>.
- Hwang, Y., Lee, S., Kim, T., Baik, K. e Choi, Y. (2022). Crop growth monitoring system in vertical farms based on region-of-interest prediction, *Agriculture* 12: 656. <https://doi.org/10.3390/agriculture12050656>.
- Joshi, J., Polepally, S., Kumar, P., Samineni, R., Rahul, S., Sumedh, K., Tej, D. e Rajapriya, V. (2017). Machine

- learning based cloud integrated farming, *ICMLSC '17: Proceedings of the 2017 International Conference on Machine Learning and Soft Computing*, pp. 1–6. <https://doi.org/10.1145/3036290.3036297>.
- Kalavathi, D., Seerengasamy, U., Palaniappan, S. e Sekar, R. (2023). Prediction of factors for controlling of green house farming with fuzzy based multiclass support vector machine, *Alexandria Engineering Journal* 62. <https://doi.org/10.1016/j.aej.2022.07.016>.
- Karunathilake, E. M., Le, A. T., Heo, S., Chung, Y. S. e Mansoor, S. (2023). The path to smart farming: Innovations and opportunities in precision agriculture, *Agriculture* 13: 26. <https://doi.org/10.3390/agriculture13081593>.
- Lauguico, S., Concepcion II, R., Tobias, R. R., Alejandrino, J., De Guia, J., Guillermo, M., Sybingco, E. e Dadios, E. (2020). Machine vision-based prediction of lettuce phytomorphological descriptors using deep learning networks, *2020 IEEE 12th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment, and Management (HNICEM)*, pp. 1–6. <https://doi.org/10.1109/HNICEM51456.2020.9400103>.
- Li, J., Qiao, Y., Liu, S., Zhang, J., Yang, Z. e Wang, M. (2022). An improved yolov5-based vegetable disease detection method, *Computers and Electronics in Agriculture* 202: 107345. <https://doi.org/10.1016/j.compag.2022.107345>.
- Liu, R., Wang, H., Guzmán, J. e Li, M. (2022). A model-based methodology for the early warning detection of cucumber downy mildew in greenhouses: An experimental evaluation, *Computers and Electronics in Agriculture*. <https://doi.org/10.1016/j.compag.2022.106751>.
- Martinho, V. (2022). *Systematic Review of Agriculture and Era 4.0: The Most Relevant Insights*, pp. 49–64. https://doi.org/10.1007/978-3-030-98959-0_2.
- Martini, B., Helfer, G., Barbosa, J., Espinosa Modolo, R., Rosa da Silva, M., Figueiredo, R., Mendes, S., Silva, L. e Leithardt, V. (2021). Indoorplant: A model for intelligent services in indoor agriculture based on context histories, *Sensors* 21: 1631. <https://doi.org/10.3390/s21051631>.
- Maya, E., Lopez, A., Zambrano, M., Domínguez, H., Díaz-Iza, H. e Vasquez, C. (2023). Fuzzy control application to an irrigation system of hydroponic crops under greenhouse: Case cultivation of strawberries (fragaria vesca), 23. <https://doi.org/10.3390/s23084088>.
- Mehra, M., Saxena, S., Sankaranarayanan, S., Tom, R. e Veeramaniandan, M. (2018). Iot based hydroponics system using deep neural networks, *Computers and Electronics in Agriculture* 155: 473–486. <https://doi.org/10.1016/j.compag.2018.10.015>.
- Micol, L., Barbosa, J., Righi, R., Monteiro, E. e Alberti, A. (2021). Blockchain in the reverse agrochemical supply chain: a systematic mapping study., *International Journal of Business Information Systems* 1: 1. <https://doi.org/10.1504/IJBIS.2021.10044483>.
- Mohmed, G., Heynes, X., Naser, A., Sun, W., Hardy, K., Grundy, S. e Lu, C. (2023). Modelling daily plant growth response to environmental conditions in chinese solar greenhouse using bayesian neural network, *Scientific Reports* 13. <https://doi.org/10.1038/s41598-023-30846-y>.
- Moorthi, K., Sureshu, S., Vasanti, G., Mishra, A., Gandhewar, N. e Kalra, G. (2023). Iot enabled mushroom farm automation based on cnn-bigru-crf method, *2023 7th International Conference on Electronics, Communication and Aerospace Technology (ICECA)*, pp. 1312–1317. <https://doi.org/10.1109/ICECA58529.2023.10395856>.
- Mya, K., Sein, M., Nyunt, T., Chong, Y. e Zainal, R. (2021). Automatic data-driven agriculture system for hydroponic farming, pp. 6–11. <https://doi.org/10.1145/3449365>.
- Ojha, T., Misra, S. e Raghuwanshi, N. (2015). Wireless sensor networks for agriculture: The state-of-the-art in practice and future challenges, *Computers and Electronics in Agriculture* 118. <https://doi.org/10.1016/j.compag.2015.08.011>.
- Parsa, S., Debnath, B., Khan, M. e Ghalamzan, A. (2023). Autonomous strawberry picking robotic system (robofruit). <https://doi.org/10.48550/arXiv.2301.03947>.
- Petersen, K., Vakkalanka, S. e Kuzniarz, L. (2015). Guidelines for conducting systematic mapping studies in software engineering: An update, *Information and Software Technology* 64. <https://doi.org/10.1016/j.infsof.2015.03.007>.
- Prasad, S., Umme Haniya, K., Zeenathunnisa, S., Tahir, M. e Sulaiman, M. (2023). Deep learning based automatic indoor plant management system using arduino, *2023 4th IEEE Global Conference for Advancement in Technology (GCAT)*, pp. 1–5. <https://doi.org/10.1109/GCAT59970.2023.10353399>.
- Rajendiran, G. e Rethnaraj, J. (2023). Smart aeroponic farming system: Using iot with lcgmm-boost regression model for monitoring and predicting lettuce crop yield, *International Journal of Intelligent Engineering and Systems* 16: 252–262. <https://doi.org/10.22266/ijies2023.1031.22>.
- Rodríguez, S., Gualotuña, T. e Grilo, C. (2017). A system for the monitoring and predicting of data in precision agriculture in a rose greenhouse based on wireless sensor networks, *Procedia Computer Science* 121: 306–313. <https://doi.org/10.1016/j.procs.2017.11.042>.
- Rosa, J., Barbosa, J. e Ribeiro, G. (2015). Oracon: An adaptive model for context prediction, *Expert Systems with Applications* 45. <https://doi.org/10.1016/j.eswa.2015.09.016>.
- Saath, K. e Fachinello, A. (2018). Crescimento da demanda mundial de alimentos e restrições do fator terra no brasil, *Revista de Economia e Sociologia Rural* 56: 195–212. <https://doi.org/10.1590/1234-56781806-94790560201>.

- Sadik, M., Anubhove, M. S. T., Akter, M., Ashrafi, N., Saif, S. e Saleque, A. (2020). Machine learning algorithm based disease detection in tomato with automated image telemetry for vertical farming, *2020 International Conference on Computational Performance Evaluation (ComPE)*. <https://doi.org/10.1109/ComPE49325.2020.9200129>.
- Sharma, A., Bhargava, M. e Khanna, A. (2021). Ai-farm: A crop recommendation system, *2021 International Conference on Advances in Computing and Communications (ICACC)*, pp. 1–7. <https://doi.org/10.1109/ICACC-202152719.2021.9708104>.
- Sikati, J. e Nouaze, J. C. (2023). Yolo-npk: A lightweight deep network for lettuce nutrient deficiency classification based on improved yolov8 nano, p. 31. <https://doi.org/10.3390/ecsa-10-16256>.
- Sisyanto, R., Suhardi, S. e Budi Kurniawan, N. (2017). Hydroponic smart farming using cyber physical social system with telegram messenger, *2017 International Conference on Information Technology Systems and Innovation (ICITSI)*, pp. 239–245. <https://doi.org/10.1109/ICITSI.2017.8267950>.
- Soheli, S. J., Jahan, N., Hossain, M., Adhikary, A., Khan, D. M. A. e Wahiduzzaman, M. (2022). Smart greenhouse monitoring system using internet of things and artificial intelligence, *Wireless Personal Communications* **124**. <https://doi.org/10.1007/s11277-022-09528-x>.
- Tan, W., Zhao, C. e Wu, H. (2016). Intelligent alerting for fruit-melon lesion image based on momentum deep learning, *Multimedia Tools and Applications* **75**. <https://doi.org/10.1007/s11042-015-2940-7>.
- Vadivel, R., Parthasarathi, R., Navaneethraj, A., Sridhar, P. e Karan, S. (2019). Hypaponics - monitoring and controlling using internet of things and machine learning, *2019 1st International Conference on Innovations in Information and Communication Technology (ICIICT)*, pp. 1–6. <https://doi.org/10.1109/ICIICT1.2019.8741487x>.
- Victor, B., He, Z. e Nibali, A. (2023). A systematic review of the use of deep learning in satellite imagery for agriculture, *Arxiv, Cornell University*. <https://doi.org/10.48550/arXiv.2210.01272>.
- Wang, X., Vladislav, Z., Viktor, O., Wu, Z. e Zhao, M. (2022). Online recognition and yield estimation of tomato in plant factory based on yolov3, *Scientific Reports* **12**: 8686. <https://doi.org/10.1038/s41598-022-12732-1>.
- Wang, X., Wu, Z., Jia, M., Xu, T., Pan, C., Qi, X. e Zhao, M. (2023). Lightweight sm-yolov5 tomato fruit detection algorithm for plant factory, *Sensors* **23**: 1–19. <https://doi.org/10.3390/s23063336>.
- Wickramaarachchi, P., Balasooriya, N., Welipenne, L., Gunasekara, S. e Jayakody, A. (2020). Real-time greenhouse environmental conditions optimization using neural network and image processing, *2020 20th International Conference on Advances in ICT for Emerging Regions (ICTer)*, pp. 232–237. <https://doi.org/10.1109/ICTer51097.2020.9325472>.
- Wu, Q., Liang, Y., Li, Y. e Liang, Y. (2017). Research on intelligent acquisition of smart agricultural big data, *2017 25th International Conference on Geoinformatics*, pp. 1–7. <https://doi.org/10.1109/GEINFORMATICS.2017.8090913>.
- Yang, J., Liu, M., Lu, J., Miao, Y., Hossain, M. e Alhamid, M. (2018). Botanical internet of things: Toward smart indoor farming by connecting people, plant, data and clouds, *Mobile Networks and Applications* **23**. <https://doi.org/10.1007/s11036-017-0930-x>.
- Zhang, Y. e Chen, M. (2022). An iot-enabled energy-efficient approach for the detection of leaf curl disease in tomato crops, *Wireless Networks* **29**. <https://doi.org/10.1007/s11276-022-03071-0>.