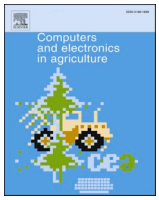
Computers and Electronics in Agriculture 232 (2025) 110048

Contents lists available at ScienceDirect

Computers and Electronics in Agriculture

journal homepage: www.elsevier.com/locate/compag



Original papers

# Agricultural data privacy and federated learning: A review of challenges and opportunities☆

Rahool Dembani [a,b,c,*], Ioannis Karvelas [a], Nur Arifin Akbar [c], Stamatia Rizou [a], Domenico Tegolo [c], Spyros Fountas [d]

[a] R&D and Innovation Department, SingularLogic, Athens, Greece
[b] Department of Mathematics and Computational Sciences, University of Messina, Messina, Sicily, Italy
[c] Department of Mathematics and Computer Science, University of Palermo, Palermo, Italy
[d] Department of Natural Resources Development and Agricultural Engineering, Agricultural University of Athens, Athens, Greece

ARTICLE INFO

*Keywords:*
Federated learning
Data privacy
Agriculture
Privacy-preserving techniques
Smart farming
Data governance

ABSTRACT

The rapid digitalization of agriculture has resulted in an unprecedented surge in data collection, necessitating this way the privacy protection in innovative data analytics solutions. Federated Learning emerges as a promising solution since it allows for collaborative model training across decentralized data sources without sharing raw data. This review explores the use of Federated Learning in agriculture, focusing on privacy-preserving methods. We thoroughly reviewed a large corpus of relevant research, examining several Federated Learning types and their application to agricultural scenarios, such as pest and disease detection, crop yield prediction, and resource management. Our findings underscore Federated Learning's potential to revolutionize privacy-preserving data analysis in agriculture by enabling better decision-making through aggregated insights from various farms, while retaining data confidentiality. At the same time, a number of technical complications arise, including data heterogeneity, communication impediments, and limited computational capabilities in rural areas. Data ownership, fairness, and stakeholder trust are significant barriers to widespread use in practice. The present study provides research gaps that need to be addressed to fully use the potential of Federated Learning in agriculture. Tailoring the design of Federated Learning algorithms and adhering to the nature of agricultural data and its peculiarities can promote the enhancement of agriculture-friendly frameworks to ensure privacy-preserving mechanisms for agriculture-oriented applications, and the development of frameworks that bear ethical issues in mind and facilitate farmers-based equitable benefit distribution. Since Federated Learning can potentially change the landscape of data-driven agriculture by allowing collaborative data analytics without compromising privacy, it is highly important to overcome the technological and ethical barriers demonstrated in this study, maximizing its impact on sustainable farming practices and innovations.

## 1. Introduction

Due to the increasing digitalization of agriculture, there has been an increase in data generation, necessitating advanced analytical techniques that are meant to utilize the available data, while protecting sensitive information. Several studies (Haseeb et al., 2020; Huang et al., 2016; Ongadi, 2024; Rathod & Shinde, 2023) state that the increase of data arising from sensors, drones, and satellite images has a lot of potential to enhance yield, resource utilization, and sustainability. Recent research has explored fuzzy deep learning and optimization techniques to address the challenges of limited and imbalanced datasets in agricultural image analysis, demonstrating high accuracy in tasks, such as citrus fruit disease detection (Shah et al., 2024). However, it has been stated that there are adverse effects associated with the use of technology today in terms of data privacy and security being compromised (Jayashankar et al., 2018; Raturi et al., 2022; Runck et al., 2022; Wilgenbusch et al., 2022). In agricultural science, data sharing is valued as it leads to enhanced research, increased productivity, and improved sustainability of the industry. This allows scientists to check the accuracy of findings, advance science for other purposes, and contribute to

☆ This article is part of a special issue entitled: 'Sust. Agri. 4.0' published in Computers and Electronics in Agriculture.

* Corresponding author at: R&D and Innovation Department, SingularLogic, Athens, Greece.
*E-mail address:* rdembani@singularlogic.eu (R. Dembani).

https://doi.org/10.1016/j.compag.2025.110048
Received 19 October 2024; Received in revised form 16 January 2025; Accepted 26 January 2025
Available online 10 February 2025

the global body of knowledge (Tenopir et al., 2011). Nevertheless, Wilgenbusch et al. (2022) argue that there have been concerns about the publication of farming data for a number of years, surfacing during the 1970 s and early 1980 s with the commercialization of biotechnological advancements.

In the digital age, data privacy issues have become even more complex. Farmers now face challenges related to data asymmetry and uncertainties about how Agricultural Technology Providers (ATPs) store and use their data (Jayashankar et al., 2018). Additionally, Shepherd et al. (2020) highlight that intellectual property rights linked to the increasing economic value of data have further heightened the need for robust data governance in agriculture, a sentiment echoed by Ashworth et al. (2023) as well. Actual events have also shown that agriculture is faced with significant privacy concerns. Bergstrom et al. (2022) illustrate how the deployment of large unmanned aircraft systems (UAS) in agriculture can raise such concerns, thereby underscoring the need for secure rural broadband infrastructure. The proliferation of digital agriculture platforms and the influx of agriculturally relevant data further amplify these privacy and security concerns (Runck et al., 2022). Safeguarding sensitive agricultural data is a collective obligation among various stakeholders in the agricultural ecosystem, emphasizing the need for collaborative efforts to protect farmers' information (Kaur et al., 2022).

Traditional machine learning models usually aggregate data from various sources into a central database for training purposes. Although effective, this method may also present major threats to privacy and security. When sensitive information is collected in one centralized location, it becomes susceptible to breaches and abuse. For instance, the scenario where someone utilizes a trained model to retrieve sensitive information is an inference attack, which is of great concern. Another threat presented by this method is that of a single point of failure created by the necessity to adhere to a central server, rendering the system vulnerable to disruptions and attacks. In light of these issues, Federated Learning (FL) is poised as promising direction in machine learning, as it resolves the privacy and security aspects of the data by allowing model training without using the actual dataset. This approach has received considerable popularity, as it allows multiple clients to train a model in a decentralized manner and addresses key privacy concerns by allowing sensitive data to remain within the borders of the owner. Similar to how 6G networks are envisioned to leverage pervasive Artificial Intelligence (AI) for data-driven machine learning applications in heterogeneous networks, FL offers a way to address the privacy and communication limitations of traditional centralized machine learning (Hasan et al., 2024). Instead of sending raw data to a central server, clients send updates about the model based on their local data, which can be combined to improve the global model. This decentralized approach is especially advantageous in regions like agriculture, where data is sensitive, vulnerable, diverse, and dispersed in many locations. Additionally, studies that adopt advanced deep neural networks for remote sensing-based land use classification (Albarakati et al., 2024) and novel bag-of-features methods for infected leaf detection (Vijh et al., 2023) demonstrate how large-scale agricultural datasets can significantly benefit from increased security measures and privacy-preserving approaches within machine learning frameworks.

Data in federated learning is protected significantly with secure aggregation, differential privacy, and homomorphic encryption techniques. These security measures can augment even more privacy and security (Aledhari et al., 2020; Kairouz et al., 2021; Lim et al., 2020). While allowing for efficient and secure model training in a world that is becoming more and more data-centric, FL also provides an answer to the problem of data centralization.

Several contributions with different methodological approaches have been proposed in the literature (Dwarampudi & Yogi, 2024; Konečný et al., 2017; C. Ma et al., 2020; Sattler et al., 2019; Shanmugam et al., 2023; Z. Sun et al., 2021) in which it is highlighted how FL allows training a joint model across multiple companies or agricultural organizations while keeping their local data local and promoting collaboration and innovation without compromising the confidentiality of the data itself. Konečný et al. (2017) describe FL as a machine learning environment in which models are trained centrally while training data resides on multiple clients with different network characteristics. FL eliminates the need to centralize private information and is therefore suitable for privacy-sensitive agricultural applications (C. Ma et al., 2020; Sattler et al., 2019; J. Yang et al., 2024). Furthermore, it is worth noting that FL addresses data privacy and security concerns by ensuring that sensitive information, such as crop yields, soil conditions, and agricultural management practices, remains under the control of individual farmers or organizations.

Data security and privacy are highly important concerns when sharing farmer information for research and decision-making. Wirth et al. (2021) emphasize the need for robust data-sharing architectures that protect sensitive information while enabling collaboration and research. Additionally, Mamba Kabala et al. (2023) demonstrate that FL has been applied to enhance agricultural practices, such as image-based crop disease detection while ensuring data privacy. As digital technology is increasingly adopted in agriculture, FL is not limited to traditional machine learning applications. Its versatility extends across various sectors, such as innovative healthcare, where AI approaches based on FL demonstrate its adaptability (Rahman et al., 2023). Dwarampudi & Yogi (2024), He & Zhao (2022), Li et al. (2020), and Sattler et al. (2019) indicate that FL enables multiple stakeholders in the agricultural ecosystem, including farmers, researchers, and technology providers, to collaborate and learn from each other without compromising data confidentiality. By allowing computations on local devices, FL reduces the need to transfer large datasets, improving efficiency and scalability, especially in agricultural settings with limited internet connectivity (Konečný et al., 2017; Woubie & Bäckström, 2021; J. Xu et al., 2021).

The present paper aims to provide a literature review that incorporates a comprehensive analysis of FL's application in agriculture. It examines the range of current FL practices and the challenges specific to the agricultural industry, shedding light on how FL can address data privacy issues without compromising the rigor of data-driven agricultural decision-making.

### 1.1. Objectives and research questions

This literature review examines FL's role in data privacy issues relevant to the agricultural sector. By exploring the range of current FL practices and the problem characteristics of the agriculture industry, it aims to shed light on how FL can allay data privacy issues without undermining the rigor implied when making data-backed agricultural decisions. It is important to note that farming is becoming increasingly data-driven with crop health monitoring, yield forecasting, resource utilization, and data analytics to search for cost-effective and environmentally sound solutions, bringing data privacy to the forefront. Although in other sectors, such as health care and finance, the potential of FL in practice has been investigated to a degree, in agriculture, its application has been underexplored. The main contribution of this review is filling the gap in the current literature concerning the application of FL to agriculture by examining the adoption of FL in an agricultural context. The inquiry in the present paper is framed using the following research questions:

RQ1. How does Federated Learning counter the issue of data privacy?
RQ2. Which Federated Learning approach is most suitable for different agricultural applications and why?
RQ3. What are the opportunities and challenges from a technical perspective in utilizing Federated Learning frameworks for agricultural settings?

## 2. Methodology

The literature search was conducted across three major databases: (1) Scopus, (2) Web of Science, and (3) IEEE Xplore (see Fig. 1). These databases were selected since they collectively cover a broad spectrum of high-impact journals, conference proceedings, and technical reports in computer science, engineering, and multidisciplinary research. Moreover, Scopus and Web of Science are renowned for their extensive indexing of peer-reviewed literature, while IEEE Xplore is a focal source for scholarly works in computing and information technology, including privacy-preserving machine learning. Although several other databases (e.g., ACM Digital Library, ScienceDirect) also contain relevant research, we aimed to prevent overlap where possible. Since Scopus often indexes many titles from publishers like Elsevier and ACM, plus conference publications that also appear on IEEE Xplore, this combination of repositories ensures the inclusion of the most relevant studies.

The search strategy was designed to capture the intersection of privacy-preserving technologies and FL within the agricultural domain. The query strings that were employed for the search were the following:

- Query1: (“homomorphic encryption” OR “secure multi-party computation” OR “differential privacy”) AND “federated learning” AND “agriculture”
- Query2: “federated learning” AND “agriculture” AND (“challenges” OR “limitations” OR “future directions”)
- Query3: “federated learning” AND (“pest detection” OR “disease detection” OR “crop yield prediction” OR “precision agriculture”)
- Query4: (“privacy-preserving” AND “federated learning” AND “agriculture”)

The key terms included in these queries (homomorphic encryption, secure multi-party computation, and differential privacy) reflect three core privacy-preserving technologies commonly employed in FL. Meanwhile, keywords like “pest detection,” “disease detection,” and

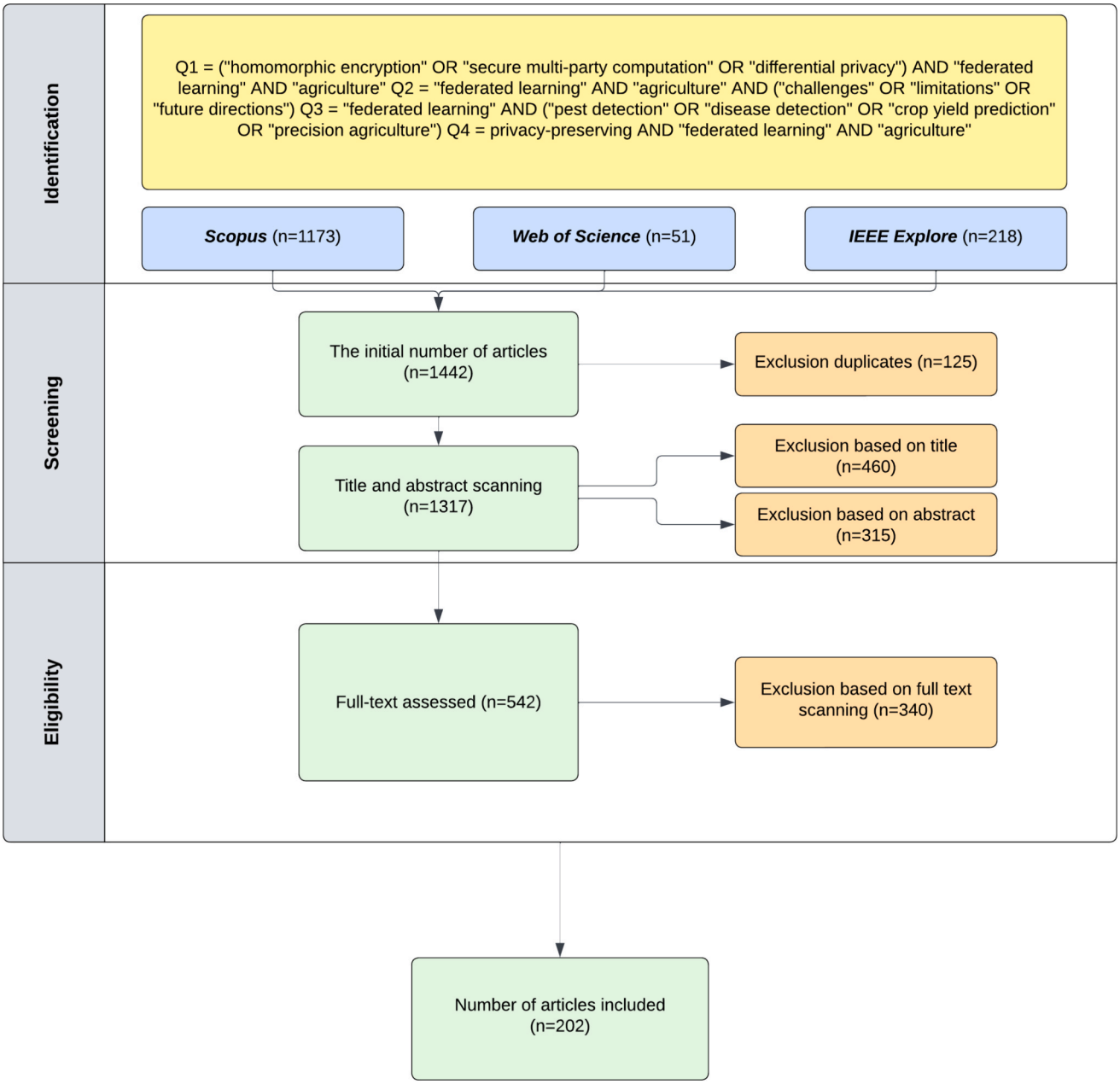



**Fig. 1.** Paper selection and reviewing process used to identify relevant literature.

"crop yield prediction" were selected to identify specific agricultural applications of FL. The publication range that was targeted was between January 2014 and August 2024 to include both foundational and recent developments in FL and privacy-preserving techniques. Only English-language articles (journal and conference) were considered.

Following the database searches, a total of 1,442 articles were initially retrieved (see Fig. 1). We proceeded to the removal of duplicate articles from the initial list and after 125 duplicates were identified, the list was sized down to 1,317 articles. The list was subsequently filtered based on article relevance. All titles and abstracts were examined, resulting in the exclusion of 460 articles based on titles and 315 based on abstracts. This process left 542 articles eligible for a full-text screening. At this stage, two additional layers of review were applied: (a) thematic inclusion/exclusion and (b) quality assessment. Thematic inclusion was applied using the following inclusion/exclusion criteria:

- The article must explicitly address Federated Learning (FL) methods or frameworks.
- The article must clearly focus on agriculture (e.g., pest detection, resource optimization, crop yield prediction, precision farming).
- The article must discuss or implement at least one privacy-preserving technique (e.g., Homomorphic Encryption, Secure Multi-Party Computation, or Differential Privacy).
- Studies purely conceptual with no application details, or papers lacking meaningful discussion of privacy or agricultural context were excluded.

During the quality assessment, the included articles were examined for clear descriptions of the study design, data sources, and evaluation metrics, ensuring that each study provided enough detail on how FL was implemented, covering algorithmic choices and data handling. Relevance and originality were also assessed by requiring each study to address our core research questions, as outlined in Section 1.1, and to contribute more than superficial discussions of FL or data privacy.

As a result of this more detailed examination and quality filtering, 340 papers were excluded, primarily due to insufficient specificity (no dedicated privacy-preserving framework) or minimal discussion of agricultural use cases. Ultimately, the present study incorporated a total of 202 articles into its literature search. The articles in this collection provided insights on how privacy preservation techniques are implemented in FL systems for pest and disease detection, crop yield forecasting, and operations in precision agriculture. Moreover, through the review, the lack of coverage in some aspects was identified and areas for potential research on privacy-preserving FL in agriculture were provided as suggestions. Thus, the present study broadens the scope of understanding and encourages further exploration of secure and collaborative data mining techniques for agriculture to expedite the development of the important sector.

## 3. Privacy-preserving techniques in federated learning

Data privacy is a prevalent concern in FL, especially in the agricultural sector where farm-sensitive data, such as yield potential and soil types, are involved. In this section, we address RQ1 by showcasing how FL mitigates data privacy challenges in agriculture through advanced techniques. Specifically, we present several measures to ensure agricultural data security in FL scenarios. Three core techniques, Homomorphic Encryption (HE), Secure Multi-Party Computation (SMPC), and Differential Privacy (DP), are discussed. These techniques represent diverse approaches to privacy preservation within FL and are particularly relevant to the agricultural domain due to the sensitivity of the data involved. With HE, it is possible to perform calculations on encrypted data and complete the analysis while still working with encrypted text, maintaining confidentiality throughout the training process. SMPC allows two or more parties to work together without revealing their individual data inputs to each other or the other parties in a collaborative situation. While DP suppresses updates that would result in a high change of the overall model functionality such that a single data point presence or absence does not materially influence the final functionality of the model, the firm changes the parameters by increasing noise. Other privacy-preserving strategies include: "Federated Learning with Secure Aggregation", "Private Set Intersection," and "k-anonymity" but for this discussion, HE, SMPC, and DP were selected because of their strengths in privacy assurance and their promise in practical segments of FL including health care and finance, which are associated with privacy risks that are also present in agricultural applications. Moreover, these techniques present a range of trade-offs in terms of their privacy protection, accuracy, and computational cost, which is beneficial for the adaptation to the different needs of application in agriculture.

In order to provide the distinctions and uses of the discussed privacy-preserving mechanisms in FL, a comparative study matrix is provided in Table 1, featuring HE, SMPC, and DP, their basic principles, merits, limitations, functional domains in FL, prospects in agriculture, privacy level, and computational complexity. These methods are more thoroughly described in Sections 3.1-3.3.

### 3.1. Homomorphic encryption

Homomorphic encryption (HE) has become an effective mechanism for enhancing data security in FL as it allows data encryption during operations without needing the data to be decrypted. This protects the data exchange from being accessed by unauthorized actors during training. For example, Pan et al. (2024) proposed FedSHE, which enhanced the accuracy, efficiency, and security of FL models through adaptive segmented homomorphic encryption, which enabled the aggregation of models without centralizing sensitive information. Likewise, Li et al. (2024) also explain how HE can combine updates from multiple clients with different keys to allow them to train a model together without breaching any user's privacy. In the context of healthcare, Wang et al. (2023) discusses PPFLHE, which employs HE in FL to provide security for big data. This guarantees the confidentiality of model sharing between users and avoids the risk of privacy leakage during training. In another study (Rieyan et al., 2024), the application of partial HE in a data fabric architecture that facilitates secure FL-based medical image analysis is presented. This also enables healthcare organizations to jointly build models in a secure environment without sharing actual health images, which are sensitive to patients. Further enhancing the use of HE in FL, Fan et al. (2022) presented an ID-based multi-receiver Homomorphic Proxy Re-Encryption algorithm. This method enhances privacy in FL settings with multiple participants, since secure model aggregation can be performed on encryption models for more than one receiver and vice versa. In the same way, Ma et al. (2022) advance the MK-CKKS multi-key homomorphic encryption framework to safeguard privacy in model updates that are exchanged during FL training. This approach guarantees that even when updates to the model are combined across clients, the updates have not left the encryption level, rendering it as a suitable approach for privacy-preserving FL.

### 3.2. Secure multi-party computation

Secure Multi-Party Computation (SMPC) is a crucial cryptographic protocol that allows different parties to jointly evaluate a function on the shareholders' private inputs, while keeping those inputs concealed from each other. In the context of FL, SMPC improves privacy by facilitating secure interaction between decentralized actors without sharing sensitive information. Many researchers have sought to determine how integrating SMPC with FL is beneficial in different settings. Elfares et al. (2024) merge SMPC with FL in their study, focusing on a privacy-enhancing training approach (PrivatEyes) that utilizes gaze interactive systems to improve privacy. This method guarantees the protection of private information even from malicious servers, while achieving reasonable accuracy and computation efficiency. Muazu et al. (2024)

**Table 1**
Comparative analysis of privacy-preserving techniques in FL.

| Technique | Principle | Advantages | Disadvantages | Applications in FL | | Potential in Agriculture | Privacy Level | Computational Overhead |
|---|---|---|---|---|---|---|---|---|
| HE | Allows computations on encrypted data without needing decryption. | Strong data confidentiality | High computational cost | (Pan et al., 2024) | Enhances efficiency and security | Secure collaborative training on encrypted farm data | High | High |
| | | No raw data exposure | Increased latency | (B. Wang et al., 2023) | Protects healthcare data privacy | Protecting sensitive crop data | | |
| | | Secure model aggregation | Complex key management | (J. Ma et al., 2022) | Prevents model update leakage | | | |
| SMPC | Computes functions over private inputs without revealing them to others | No need for a trusted aggregator | Communication overhead | (Elfares et al., 2024) | Enhances gaze estimation privacy | Collaborative optimization of resources among farms | High | High |
| | | Data remains local | Scalability issues | (Arora et al., 2023) | Financial anomaly detection | Secure sharing of soil data | | |
| | | Collaborative model building | Protocol complexity | (Hosseini et al., 2022) | Cluster-based approach in healthcare | | | |
| DP | Introduces noise to data or computations to prevent disclosure of individual data points | Quantifiable privacy guarantees | Potential loss of accuracy | (Wu et al., 2021) | Adaptive gradient descent with DP | Sharing aggregated data insights | Moderate to High | Low to Moderate |
| | | Balances privacy and utility | Noise calibration is critical | (Truex et al., 2020) | Provides formal privacy guarantees | Protecting farm-specific practices | | |
| | | Resistant to certain attacks | Complex implementation | (Adnan et al., 2022) | Medical image analysis | | | |

propose a secure FL system that employs SMPC and additive secret sharing to secure shape gradient parameters during model updates, achieving data privacy and high accuracy.

In distributed machine learning systems context, Yang et al. (2023) discuss how FL incorporates SMPC to safeguard data security throughout the model's lifecycle. This technique emphasizes privacy protection and intellectual property rights by ensuring that sensitive information is never exposed during training. Similarly, in the financial sector, Arora et al. (2023) have used SMPC to implement privacy-preserving FL for financial anomaly detection. This approach allows multiple banks to jointly train accurate machine learning models while ensuring that customer data remains private. Further enhancements to FL with SMPC have been explored in healthcare. For example, Hosseini et al. (2022) propose a cluster-based approach where SMPC securely shares model weights among hospitals in clusters. This method prevents individual hospitals from accessing others' data while still improving overall model accuracy, although it introduces higher communication overhead. Another innovative approach, CE-Fed, combines SMPC with FL to enable communication-efficient collaborative model training among multiple parties (Kanagavelu et al., 2022). This approach improves privacy without compromising communication efficiency. While SMPC enhances privacy in FL, it introduces challenges such as increased computational complexity and communication overhead. Overcoming these challenges is essential for practical applications in agriculture.

### 3.3. *Differential privacy*

FL has emerged as a transformative approach in machine learning, particularly in sectors such as agriculture, where data privacy is paramount. Integrating Differential Privacy (DP) into FL frameworks enhances the protection of sensitive data while enabling collaborative model training across decentralized data sources. This synthesis explores the intersection of FL and differential privacy within the agricultural context, highlighting the benefits, challenges, and potential applications. Initially introduced in 2006, DP offers a mathematical guarantee that the output of a data analysis remains nearly indistinguishable, regardless of whether any individual's data is included in the dataset (Chen & Liu, 2022; Hardt & Talwar, 2010). This approach addresses significant limitations of previous privacy models, such as k-anonymity, which can be vulnerable to various forms of re-identification attacks (Gotz et al., 2012; Jiang et al., 2018). The key principle of differential privacy is the introduction of randomness into the output of queries made on a dataset. The Laplace mechanism (Dwork, 2006) often achieves this by adding noise to the results. According to Al-Hussaeni et al. (2018) and Li et al. (2010) the scale of the noise added is determined by the query's sensitivity, which is how much the output changes with the addition or removal of a single data point. By adjusting the noise level, differential privacy provides a quantifiable trade-off between privacy and utility, enabling valuable insights from data while safeguarding individual privacy (Liu et al., 2018; Roth & Roughgarden, 2010). When combined with differential privacy, FL offers a robust solution for maintaining data privacy while enabling collaborative model training across different parties. This technique is being explored in various domains to enhance privacy and performance.

Wu et al. (2021) proposed a combinatoric approach consisting of modeling and DP, which seeks to enhance the efficiency of a multi-party collaborative model while reducing messages sent in an efficient way of communication. Likewise, LDP-Fed (Truex et al., 2020) employs local differential privacy to guarantee formal privacy for model training parameters, which solves the high dimensionality of large target neural network catalogs containing continuous data with fundamental image features. In the field of medicine, Adnan et al. (2022) show that differentially private FL is very promising in medical image interpretation of histopathology images. Their method provides results at the same level as those from conventional training approaches but with strong privacy protection. To work with realistic health data, Choudhury et al. (2020) designed an FL architecture that favors privacy by not allowing the movement or dissemination of raw data, using DP techniques as a safety measure. In the context of personalized FL, Yang et al. (2023) offer FedDPA, which overcomes the disadvantages of non-personalized personalization and the problems that arise during the convergence of their approach. It uses layer-wise Fisher information to personalize the model, increasing its flexibility and adjustable constraint strategies to improve convergence and clipping. Further, Sun et al. (2021) introduced

an innovative approach to enhancing privacy in FL, focusing on reducing the recognizability and reusability of sensitive data and implementing vertical and horizontal mixups of model updates to achieve better performance in deep learning, while still attaining high levels of privacy. Shi et al. (2021) propose a privacy-preserving approach for Hierarchical Federated Learning (HFL). This method ensures order preservation for shared model parameters by adding noise via local differential privacy before they are sent to the edge and cloud servers. This technique has been evaluated on image classification tasks and has been shown to perform well regarding privacy protection.

### 3.4. Leveraging privacy-preserving federated learning for agricultural data security

Even though techniques ensuring the users' privacy have been successfully applied in FL contexts in the medical and financial sectors, their use in agriculture is underexplored. Farmers hold large volumes of information that could be classified as sensitive, such as expected yields, soil and climatic conditions, that could be greatly benefitted by the use of FL systems. Still, it is important to note that there is a lack of research emphasizing on the application of HE, SMPC or DP specifically in FL for agriculture, providing ample room for more research and development in the area. This underscores the need for further exploration and development in the field of encryption. HE in FL allows multiple farms that possess a machine learning model to collaborate for training using encrypted data, while preventing exposure of raw data to each other or to a central server. For instance, one related case would be that of several farms wanting to devise a predictive model to early identify crop diseases using specific soil moisture level, weather conditions and past records of disease. Each farm's data would be encoded via HE so that, during the FL process, all computations occur on this encrypted data at a local level (i.e., client). The model could potentially predict the outbreaks of the disease efficiently, while allowing preventive strategies to be used by farmers and keeping their sensitive data safe.

Integrating SMPC into agricultural FL frameworks allows multiple parties to collaborate securely in developing predictive models or analyzing trends while maintaining the confidentiality of their individual data. Consider a group of farmers that seeks to optimize fertilizer application by analyzing soil nutrient data in relation to crop requirements. Incorporating SMPC within a FL framework enables each farmer to contribute to a model that forecasts optimal fertilizer application rates, while maintaining the confidentiality of their individual soil data. This collaboration may result in improved models that increase crop yields and minimize environmental impact, benefiting all stakeholders involved. Such secure collaboration highlights the potential of SMPC in boosting both confidentiality and productivity.

The implementation of DP in FL offers a significant opportunity for the agricultural sector. Applications may encompass crop yield prediction, pest and disease management, precision agriculture for resource optimization, and supply chain optimization for agricultural cooperatives. Research into FL with DP is essential due to the growing digitization of agriculture and the sensitivity of farm data. This approach could enable collaborative model development while protecting individual farm data. Multiple farms collaborating to enhance pest resistance strategies can utilize DP to ensure that the contributions from each farm do not disclose specific information regarding their pest occurrences or management practices. This method facilitates the collaborative creation of efficient pest management models while maintaining the privacy of each participant.

In conclusion, integrating HE, SMPC, and DP into FL frameworks presents considerable potential for the advancement of the agriculture sector. Privacy-preserving techniques facilitate secure and collaborative data analysis, resulting in enhanced efficiency and sustainability in agricultural practices. As FL continues to evolve, adopting these methods in agriculture may significantly improve data confidentiality, promote collaboration among farmers, and drive innovation in agricultural technologies.

## 4. Federated learning applications in agriculture

Building upon the concepts of FL and the privacy-preserving techniques discussed in Section 3. In this section, we tackle RQ2 by examining various FL approaches applied to agricultural challenges. we provide a comparative analysis of FL applications and models within the agricultural sector. By examining specific use cases, such as pest and disease detection, crop yield prediction, and precision agriculture, we aim to offer insight into how FL tackles specific challenges within the agriculture sector through the case studies focusing on pest and disease detection, predicting crop yields, and precision farming, while upholding data privacy principles. This type of analysis not only represents the level of progress of FL implementations but also demonstrates areas that need further research and development. FL has the advantage of fostering collaboration without putting vulnerable nodes at risk, which is highly important for the farming industry. In the next sections an analysis of the different uses of FL in agriculture is presented and it is demonstrated how in some special cases this technology changes agricultural processes.

### 4.1. Potential pest and disease detection

One of the critical challenges in agriculture is the timely detection of pests and diseases, which can significantly impact crop yields and food security. FL offers innovative solutions for enhancing pest and disease detection systems while preserving data privacy. Table 2 provides a comprehensive demonstration of the way in which FL can be used to detect pests and diseases in different agricultural crops. Crops like apples, rice, wheat, bananas, strawberries, and potatoes are included, but it should be noted that each one of these crops poses a different challenge due to their distinct pests and diseases. Deng et al. (2022) highlight several diseases in apple crops, including anthracnose, bitterpox, ring rot, and fruit rust. Similarly, banana crops suffer from diseases such as Black Sigatoka, Fusarium Wilt, Banana Bunchy Top Virus, Moko disease (Suryavanshi et al., 2024), while in the case of rice crops, bacterial leaf blight, blast, brown spot, and tungro pose significant threats (Aggarwal et al., 2024). Thapliyal et al. (2024) highlighted the diseases of the strawberry leaf, while (Mehta et al., 2023) studied wheat's resistance to several diseases, including Fusarium Head Blight, Stem Rust, Leaf Rust, Septoria Leaf Blotch, and Powdery Mildew. Furthermore, potato crops are seriously affected by early and late blight diseases (Kamal et al., 2024).

Based on the FL approaches utilized in these studies, a distinction may be made between centralized and decentralized systems of FL. Centralized approaches are adopted in the work of Deng et al. (2022), where a Faster Region Convolutional Neural Network (R-CNN) embedded with ResNet-101 and the FedAvg algorithm is a common approach in apple pest detection. Moreover, decentralized methods are utilized, among others, in Suryavanshi et al. (2024) for banana plants or in Thapliyal et al. (2024) for strawberries, where CNN models with FL are employed. A centralized federated architecture and decentralized methods were also combined in the research by Aggarwal et al. (2024), who worked on the problem of rice crop disease detection.

Comparatively, the models used across these studies also included Convolutional Neural Networks (CNNs), EfficientNet architectures, and the faster R-CNN models. For instance, Efficient Net B03, a 344 layer CNN based architecture, was employed by Ahmad et al. (2022) to parasitic insects and pests such as beetles, mosquitoes, and aphids in various crops. For instance, in this particular case, transfer learning is used in the study of Kamal et al. (2024) who worked on cultivating potato crops using pre-trained models such as Inception-V3 and VGG16.

The number of images ranges from 4463 in the study of strawberry leaf disease by Thapliyal et al. (2024) to 57,849 in the study on banana crops (Suryavanshi et al., 2024). This illustrates the scope and depth of

**Table 2**
Applications for pest and disease detection in agricultural crops.

| Study | Crop Type | Disease/Pest Detected | FL Approach | Model Used | Technique | Dataset Size | Model Accuracy | Challenges & Limitations |
|---|---|---|---|---|---|---|---|---|
| (F. Deng et al., 2022) | Apple | Various orchard pests include apple diseases like anthracnose, bitter pox, ring rot, and fruit rust. | Centralized | Faster Region Convolutional Neural Network (Faster R-CNN) | R-CNN network is enhanced by using ResNet-101 with FedAvg algorithm for FL | 15,522 images after data augmentation | 89.34 % | Unbalanced and insufficient data from different orchards, diversity of pests and diseases, and complex detection environments |
| (Ahmad et al., 2022) | Various crops | beetle, mosquito, aphids, armyworm, grasshopper, bollworm, stem borer, sawfly, and mites | Centralized | Convolutional Neural Network (CNN) based on EfficientNet B03 architecture with 344 layers and 7 MBConv blocks | FL – EfficientNet B03 | 5400 images (600 images per class for 9 pest classes) | 99.55 % | Computational resource constraints and the complexity of real-time detection in varied environmental conditions |
| (Aggarwal et al., 2024) | Rice | bacterial leaf blight, blast, brown spot, and tungro | Centralized & decentralized | EfficientNetB3 | Federated Transfer Learning | 5,932 images | 99 % | Resource constraints on IoAT devices, heterogeneous data distribution |
| (Thapliyal et al., 2024) | Strawberry | Strawberry Leaf Disease | Decentralized | CNN | FL with Convolutional Neural Networks (CNN) | 4663 images | 95.49 % to 96.91 % | Data heterogeneity, ensuring model generalization across different environments |
| (Suryavanshi et al., 2024) | Banana | Black Sigatoka, Fusarium Wilt, Banana Bunchy Top Virus (BBTV), Moko disease, and Cigar end rot | Decentralized | CNN | FL approach combined with CNN | involved 5 clients (labeled kx_1 to kx_5) – 57,849 samples | 0.98 to 0.99 % | Data heterogeneity and the need for high computational resources for CNNs |
| (Mehta et al., 2023) | Wheat | Wheat diseases include Fusarium Head Blight, Stem Rust, Leaf Rust, Septoria Leaf Blotch, Powdery Mildew, and Stripe Rust. | Decentralized | CNN | FL with CNNs, using federated averaging for wheat disease detection | 9,876 images | 90.27 % | Data heterogeneity, communication overhead, and potential issues with model convergence in FL. |
| (Kamal et al., 2024) | Potato | Early Blight & Late Blight | Not mentioned | Inception-V3 | Transfer Learning, leveraging pre-trained models (VGG16, VGG19, InceptionV3, EfficientNet B1) | Two datasets (from Bangladesh and Pakistan). Not mention the exact number | 88.46 % | Data Heterogeneity |

complex and scalable machine learning techniques that have been utilized. The performance of these models also differs, the best of which was given by Ahmad et al. (2022), with 99.55 % accuracy for pest detection using EfficientNet B03. Other studies have also reported high accuracy, such as the rice disease detection model (Aggarwal et al., 2024), achieving an accuracy of 99 % and the strawberry disease detection model with 95.49 %-96.91 % (Thapliyal et al., 2024).

The problem of data heterogeneity appears in studies conducted across several disciplines, including decentralized systems, specifically in the works of Suryavanshi et al. (2024) on bananas and Mehta et al. (2023) on wheat crops. Other shortcomings include commitment of computational resources and difficulties of real-time detection in different environmental conditions (Ahmad et al., 2022). There is also concern over the amount of processing power that is needed, especially for CNNs (Suryavanshi et al., 2024). Other limitations include communication overhead and model convergence issues in FL models (Mehta et al., 2023).

### 4.2. Crop yield prediction

Accurate crop yield prediction is vital for planning, marketing, and ensuring food security. Traditional centralized models often face challenges due to data privacy concerns and the heterogeneity of agricultural data. FL provides a framework for collaborative prediction models that respect data privacy. This section delves into studies that have applied FL to enhance the precision and reliability of crop yield predictions across different regions and crops. Table 3 provides various applications of FL for predicting crop yields. This table demonstrates how FL models improve yield forecasts for different crops in terms of precision and dependability while maintaining the confidentiality and privacy of sensitive agricultural data.

Durrant et al. (2022) and T et al. (2022) studied soybean yield forecast using cross-silo and horizontal FL, respectively. Other studies employed a variety of FL techniques integrating neural network types such as CNN, RNN, and ResNet with the federated average approaches. The conclusion has been drawn that it is possible to use FL models without sacrificing the performance level, which is typical only for the centralized learning approach.

Towards advancing the scope of use of FL beyond soybeans, Q. Zhang et al. (2023) investigated maize yield estimation based on federated random forest model. Their method involved encrypted features and a secure model-sharing approach, which they tested to offer improved accuracy and cost efficiency under data-scarce situations. In further research, Idoje et al. (2023) shifted the focus to decentralized FL models, applying them to the yield estimation of multiple crops, including chickpeas, rice, and maize. Their work, in which Gaussian Naive Bayes was deployed along with various optimizers, focused on the prospects of FL in realizing the accuracy and speed of convergence without the need

**Table 3**
Studies that focus on applications for crop yield prediction.

| Study | Crop Type | FL Approach | Model Used | Techniques | Data Privacy | Key Benefits | Outcomes |
|---|---|---|---|---|---|---|---|
| (Durrant et al., 2022) | Soybean | Cross-silo | CNN, RNN, FedAvg | FedAvg, Model Sharing | Train the model locally | Improved performance, data privacy | Effective, privacy-preserving, near-baseline results |
| (T et al., 2022) | Soybean | Horizontal | ResNet-16, ResNet-28 | Federated Averaging with ResNet; compared with centralized learning (RF, LASSO) | Preserved through decentralized model training | Data privacy, decentralized training, improved performance | Federated models perform comparably to centralized ones and are suitable for privacy-constrained environments. |
| (Q. Zhang et al., 2023) | Maize | Horizontal | Federated Random Forest | FL with random forest | Local data, encrypted features, secure model | Improved accuracy, data privacy, cost-effective | Nearly lossless accuracy benefits small datasets |
| (Idoje et al., 2023) | Chickpea, Rice, Maize | Decentralized | Gaussian Naïve Bayes | FL with Gaussian Naïve Bayes, SGD, Adam optimizers | High – raw data remains on edge devices | Data privacy, accuracy, faster convergence | Adam optimizer provides high accuracy and privacy |

for data.

The above study results further support the growing importance of FL technology, especially for agricultural applications, as it can protect data, improve the models, and enable unsupervised learning for different types of crops. The development of this trend is related to the increasing interest in fitting various machine learning algorithms onto the FL schema for implementation in agricultural predictive tasks.

### 4.3. Precision resource management

Precision agriculture aims to optimize resource use and enhance agricultural productivity by leveraging detailed field data. However, the collection and analysis of such data raise privacy issues. FL enables the development of models for resource optimization without compromising individual data ownership. This section explores how FL has been applied to various aspects of precision agriculture, including irrigation management, energy consumption, and nutrient optimization. Table 4 describes various studies that focus on optimizing resource management in precision agriculture utilizing a FL approach. Akbari et al. (2023) concentrate on the issues of energy efficiency and low age of information in active or real-time applications such as monitoring different environmental parameters and smart irrigation. With the aid of FL and reinforcement learning (Deep Q-Network), their system is capable of combining IoT devices and UAVs with mobile edge computing (MEC) servers using dispersed UAVs to allow system expansion. The study achieved real-time data processing, despite challenges such as virtual network functions (VNF) placement, energy efficiency, and real-time processing demands.

Other works also stress energy management and security of heterogeneous data (Yu et al., 2022; Kumar et al., 2022). Within edge IoT systems, Yu et al. (2022) utilize a joint FL and greedy scheduling algorithm to improve the level of precision when predicting crop growth or diagnosing pests. Privacy is preserved as there is only the sharing of model parameters and intermediate results, although the focus is on energy efficiency and not real-time processing to ensure the availability of the system across large datasets. On the other hand, Kumar et al. (2022) attempted to improve privacy in IoT networks for smart agriculture by using a two-tier privacy management system with GRU and LSTM-AE models. Their investigation points out several issues including the need to maintain not only data security but also introduce real-time intrusion detection system while dealing with the problem of computation and energy.

The remaining studies discuss other aspects of using FL to optimize agricultural processes and the distribution of water in urban areas. Siniosoglou et al. (2023) present the FL and LSTM models for real-time forecasts focusing on animal and crop production. Even though the system is scalable and offers real-time insights, communication overhead and data variability present some of the biggest challenges. Similarly, Elhachmi & Kobbane (2022) offer an irrigation distribution system based on FL, with the objective of enhancing water resource management. Even though their system can provide the necessary ecosystem to promote distributed data processing and scalability across many edge nodes, problems regarding synchronization and computational load still exist. As was noticed in all investigated studies, FL is quite effective in resource optimization. At the same time, there is a great deal of communication overhead and non-uniform data treatment, which needs to be addressed.

## 5. Challenges and opportunities in implementing federated learning in agriculture

Building upon the insights gained from the comparative analysis of FL applications and techniques, this section answers RQ3 by presenting an in-depth analysis of the technical challenges and opportunities in deploying FL solutions in agricultural settings. We delves into the limitations and challenges that currently hinder the widespread adoption of FL in agriculture. By acknowledging these technical, social, and ethical obstacles, we can better understand the complexities of implementing FL and identify avenues for future research to address these critical issues.

### 5.1. Technical challenges

As it has been emphasized multiple times in the present paper, FL is a technique that can potentially ensure privacy in the agricultural domain, among others. Nonetheless, its implementation poses various technical difficulties that need to be overcome if its full potential in agriculture is to be realized. These issues are diverse and include heterogeneity in data and concerns for privacy, among others. In Table 5, the technical problems prevalent in the implementation of FL in agriculture are presented, as well as proposed solutions, challenges recognized, and goals for the next stages in research. Data heterogeneity, communication efficiency, model convergence, computational resources, privacy and security, and scalability, which are included in the challenges summarized in the table, are crucial for implementing FL technology in agriculture.

For example, one such cause of data heterogeneity is the nature of the data being considered as non-IID, data gathered from different farms is unique, like different crops, land size, etc. Due to different data classifications and methods of collection (Jiang et al., 2020; Wang, 2024; Wang, 2022). This obstacle presents a great practical opportunity, as FL algorithms can be adapted to deal with such diversity. Future studies should aim to develop FL frameworks specific to the agricultural domain to address this situation effectively.

Communication efficiency, particularly in villages with insufficient bandwidth, is another significant problem in the FL framework. In multiple studies, the authors outline the applications of gradient compression methods and efficient communication protocols (Konečný et al., 2017; Ni et al., 2023; Wang et al., 2023; Yu et al., 2022). Integrated edge computing and developing efficient communication algorithms are mentioned as promising directions towards combatting this problem.

**Table 4**
Studies focusing on applications of precision resource management.

| Study | Objective | Application Area | FL Approach | Data Sources | Privacy Concerns | Computational Resources | Scalability | Real-time Capabilities | Resource Optimization Focus | Challenges |
|---|---|---|---|---|---|---|---|---|---|---|
| (Akbari et al., 2023) | Minimize energy use and ensure low AoI in real-time agriculture applications. | Environmental monitoring, precision farming, and smart irrigation. | FL and reinforcement learning (Deep Q-Network). | IoT devices in agricultural fields. | FL to protect local data. | UAVs and MEC server; optimized via NFV. | scalable with distributed UAVs; challenges in resource orchestration. | Supports real-time monitoring with strict AoI requirements. | CPU, memory, bandwidth, and energy optimization. | Balancing energy, real-time processing, and VNF placement. |
| (Yu et al., 2022) | Optimize energy use in FL for Edge-IoAT. | Crop growth prediction, pest diagnosis, IDC prediction in soybean crops. | Joint FL with a greedy scheduling algorithm. | Farm edge nodes (soil moisture, crop images) and server (satellite images). | Only model parameters and intermediate results are shared, preserving raw data privacy. | Edge nodes (drones, iPads) and a central server optimized for energy efficiency. | Scales well with optimized device scheduling across large farm datasets. | Delay-tolerant, focusing on energy efficiency rather than real-time processing. | Energy consumption and communication resource (spectrum) optimization. | Energy constraints, communication limits, non-i.i.d. data handling. |
| (Kumar et al., 2022) | Enhance data security and privacy in smart agriculture using PEFL. | Intrusion detection in smart agriculture IoT networks | FL, GRU, LSTM-AE | IoT sensors and ToN-IoT dataset. | Two-level privacy: perturbation-based encoding and LSTM-AE transformation | Requires strong computing power, tested on Intel Xeon with 128 GB RAM | Scales well with large IoT networks using edge devices and FL | Supports real-time intrusion detection and optimization | Secures and efficiently processes IoT data | Ensuring data privacy, detecting intrusions, managing computational load |
| (Siniosoglou et al., 2023) | Optimize forecasting models in smart agriculture using FL | Animal welfare prediction, crop production optimization | LSTM models combined with FL | Sensor data from stables and fields | Data remains local; only model updates are shared | Moderate; local training and communication required | Highly scalable with increased nodes; some communication overhead | Capable of real-time forecasting and insights | Optimizing crop yield and animal welfare | Communication overhead, non-uniform data, computational load |
| (Elhachmi & Kobbane, 2022) | Optimize water distribution minimize loss using FL | Water distribution systems, optimizing urban water supply | Federated Averaging (FedAvg), Linear Regression | Sensor data from water distribution networks | Data stays decentralized; only model parameters are shared | Local computation at gateways; central server for model aggregation | Scales across multiple edge devices (gateways) | Potential for near real-time updates; not explicitly detailed | Water resources, reducing waste, and optimizing distribution | Data heterogeneity, synchronization, computational load on devices |

**Table 5**
Technical challenges in the implementation of FL in agriculture.

| Studies | Description | Possible Solutions | Challenges | Future Directions |
|---|---|---|---|---|
| (D. Jiang et al., 2020; S. Wang, 2024; X. Wang, 2022) | Non-IID data across farms, diverse data types, and collection methods | Adaptive FL algorithms, robust algorithms for handling diverse data | Data Heterogeneity | Develop domain-specific FL models for agriculture |
| (Konečný et al., 2017; Ni et al., 2023; Y. Wang et al., 2023; Yu et al., 2022) | Limited bandwidth in rural areas, high communication overhead | Gradient compression, efficient aggregation methods, lightweight FL models | Communication Efficiency | Edge computing integration, development of communication-efficient FL algorithms |
| (Q. Li et al., 2021; Rana et al., 2023; Y. Wang et al., 2023) | Slow convergence due to data diversity and system heterogeneity | Hierarchical FL, adaptive optimization techniques | Model Convergence | Personalized FL for farm-specific models, development of convergence-guaranteeing algorithms |
| (Kumar et al., 2022; Ni et al., 2023; Park & Lee, 2022; T. Zhang et al., 2021) | Limited processing power of agricultural IoT devices | Lightweight FL models, efficient training algorithms | Computational Resources | Hardware-software co-design for agricultural IoT, development of resource-aware FL frameworks |
| (Kumar et al., 2022; Reisizadeh et al., 2020; Zhu et al., 2023) | Protecting sensitive farm data, ensuring data confidentiality | Differential privacy, secure aggregation, blockchain integration | Privacy and Security | Quantum-resistant cryptography for long-term security, advanced privacy-preserving techniques tailored for agricultural data |
| (Ni et al., 2023; Yu et al., 2022) | Challenges in scaling FL across numerous heterogeneous agricultural devices | Efficient device scheduling, semi-FL approaches | Scalability | Development of scalable FL architectures for large-scale agricultural IoT networks |

Another issue that emerged in the literature was the slow convergence of FL models arising out of data and system heterogeneities (Li et al., 2021; Rana et al., 2023; Wang et al., 2023). To address this, hierarchical FL models and adaptive optimization methodologies have been proposed. Table 5 outlines future directions for personalized FL solutions in farms to achieve better FL model convergence in various agricultural settings.

From a technical standpoint, IoT devices often face constraints related to cost, energy, and limited computational resources, which can impede their performance and security. To address these challenges, Kumar et al. (2022) leveraged a deep privacy-encoding-based FL approach, enhancing data security without imposing significant hardware demands. Building on this, Ni et al. (2023) integrated centralized and federated paradigms to manage imbalanced data and device diversity in large-scale IoT networks, thereby improving scalability and resource utilization. Towards further refining FL frameworks for resource-constrained environments, Park & Lee (2022) introduce rate-splitting transmission alongside optimized training parameters and fronthaul quantization, significantly reducing model training time and enhancing efficiency for agricultural IoT devices. Furthermore, Zhang et al. (2021) present the FedIoT platform and FedDetect algorithm, which facilitate on-device anomaly detection and utilize adaptive optimization techniques to minimize communication and storage overhead, ensuring effective learning processes even with limited computational and energy resources. These studies provide scaffolding to overcome the hardware and data challenges in deploying FL for agricultural IoT systems, ensuring improved performance, security, and operational efficiency.

Privacy and security need to be protected in any situation that involves sensitive data, including farms. Multiple studies suggest the use of techniques such as differential privacy, secure multi-party computation, and blockchain to solve this problem (Kumar et al., 2022; Reisizadeh et al., 2020; Zhu et al., 2023). While the presented future challenges are based on practical realizations, quantum-resistant encryption and superior privacy mechanism design are emphasized in terms of future research efforts for long-term data protection.

Finally, scalability is a challenge stemmed from the large number of diverse devices in agricultural networks. Strategies such as device scheduling and semi-FL approaches could be potential alternatives, and deploying them on a large farm infrastructure, even without increasing FL capacity, presents an opportunity to develop suitable architecture for agricultural (IoT) systems (Ni et al., 2023; Yu et al., 2022).

### 5.2. Social and ethical challenges

Many technical challenges in implementing FL in agriculture have some social and ethical concerns as well. Ethical and social problems surrounding data privacy, farmer autonomy, and the broader social impacts of advanced agricultural technology complicate this matter further. Table 6 shows major ethical and social issues related to FL in agriculture, their ramifications, and how they can be addressed. These problems are complicated and need a multi-faceted approach provided in this comprehensive overview to guide all the individuals involved.

In the first row of Table 6, the focus is on data ownership and approval. Raturi et al. (2022) and Wilgenbusch et al. (2022) emphasize concerns regarding data ownership that may deter farmers from collecting agricultural data via FL systems. Well-defined consent requirements and fair usage policies, which will enhance trust and openness, could potentially mitigate this issue. Xu et al. (2019) highlight that while FL allows participants to collaborate on model training without sharing their actual data, which is crucial for maintaining data ownership and privacy, existing methods often lead to significant communication overhead and slower training times. This trade-off can deter adoption in agriculture, where efficiency and timely data processing are vital. Dwarampudi & Yogi (2024) further discuss how FL addresses data ownership and privacy concerns in agriculture by enabling collaborative model training across decentralized data sources. This approach allows farmers to share their data for machine learning purposes without risking their privacy, thus facilitating the adoption of FL in the agricultural sector.

Furthermore, Table 6 signifies focus on the issues of equitable access (e.g., access to FL models by minority or fragmentation farms and their inclusion in modeling). The works of Macaulay & Butsic (2017) and Raturi et al. (2022) provide an example of such problems when FL models are designed without minimizing risks, benefiting mostly larger and more technologically mature farms. To rectify this, it is sufficient to apply approaches ensuring representative sampling from various categories of farms.

Another pressing issue is the digital divide as well as accessibility in important societies. Bergstrom et al. (2022) and Ongadi (2024) developed a case that underscores how FL could create an imbalance in the farms that are technologically advanced and those that operate using more conventional means. Limited access to computers and to network connectivity in the non-urban areas – a common problem in many countries – is a critical issue. Measures must be put to curb these inequalities to avoid the small or low-end-technology farms being sidelined.

Equally important are issues of privacy and trust, as there is a need to avoid too much sharing of data that can disadvantage farmers versus the need for data sharing in agriculture (Jayashankar et al., 2018; Kaur et al., 2022). FL offers a solution by enabling local data processing and only sharing model updates, thereby addressing privacy concerns (Lim

**Table 6**
Social and ethical challenges for implementing FL for agriculture.

| Studies | Description | Implications | Challenges | Potential Mitigation Strategies |
|---|---|---|---|---|
| (Dwarampudi & Yogi, 2024; Raturi et al., 2022; Wilgenbusch et al., 2022; R. Xu et al., 2019) | Concerns about who owns agricultural data collected through FL systems | Farmers may be hesitant to participate in FL networks | Data Ownership and Consent | Develop precise consent mechanisms and data ownership policies |
| (Macaulay & Butsic, 2017; Raturi et al., 2022) | Ensuring smaller farms and underrepresented practices are fairly represented in FL models. | Potential bias in model outcomes favoring larger farms | Fair Representation and Equity | Implement techniques to balance data representation across different farm types and sizes. |
| (Bergstrom et al., 2022; Ongadi, 2024) | Potential exacerbation of the gap between technologically advanced and traditional farming practices | Unequal access to benefits of FL in agriculture | Digital Divide and Accessibility | Develop strategies to overcome limited computational resources and network connectivity in rural areas. |
| (Jayashankar et al., 2018; Kaur et al., 2022; Lim et al., 2020; Mugunthan et al., 2020) | Balancing data sharing needs with farmers' privacy concerns | Reluctance to share sensitive agricultural data | Privacy and Trust | Implement robust privacy-preserving techniques and build trust through transparency. |
| (Ashworth et al., 2023) | Making FL models interpretable for farmers and policymakers | Lack of trust in "black box" FL models | Transparency and Explainability | Develop methods to make FL models more explainable and transparent |
| (Bongiovanni & Lowenberg-Deboer, 2004) | Considering the environmental impact of increased computational demands | Potential increase in energy consumption | Environmental and Sustainability Implications | Balance the benefits of data-driven agriculture with sustainability concerns. |
| (Ashworth et al., 2023) | Ensuring responsible use of aggregated insights | Potential for unfair market advantages or exploitation | Ethical Use of Agricultural Data | Develop ethical frameworks for data use in agricultural FL |
| (Wilgenbusch et al., 2022) | Navigating complex data protection regulations | Compliance issues in implementing FL across different jurisdictions | Regulatory and Legal Challenges | Develop industry standards and best practices aligned with relevant laws and regulations. |

et al., 2020). Additionally, implementing robust privacy-preserving algorithms helps prevent information leakage from model parameters, fostering trust among agricultural stakeholders (Mugunthan et al., 2020). These privacy-preserving approaches and trust-building measures are essential for tackling the challenges of data sharing faced by farmers.

Such FL models are usually looked down upon as "black box" models which lack transparency (Ashworth et al., 2023). Trust in researchers' results may be undermined among farmers and decision-makers due to the absence of the models' ease of understanding in result interpretation. Hence it becomes important to develop strategies that will improve the details in terms of explainability in order to ensure confidence in such practices.

From an environmental perspective, Bongiovanni & Lowenberg-Deboer (2004) elaborated on how FL has environmental and sustainability consequences bearing in mind the likely ramifications on energy use trends from implementing FL. Reconciling the promises of data-centric agriculture with sustainability principles is crucial since unbridled escalation of computational power may have adverse effects on the ecosystem.

Some ethical issues also emerge from the use of aggregated data. For instance, Ashworth et al. (2023) mentioned issues of unfair advantages on the market or even exploitations. Oversight principles must be created to ensure that the information gained from FL is not misused.

Lastly, Table 6 provides the regulatory and legal challenges, noting that FL implementations must navigate complex data protection regulations across various jurisdictions (Wilgenbusch et al., 2022). Developing industry standards and best practices that align with the relevant legal frameworks can address compliance issues.

## 6. Discussion and future work

Considering the identified technical, social, and ethical challenges, future research must focus on developing solutions that address these barriers. In this section, we discuss potential research directions that can advance the field of FL in agriculture, including algorithmic innovations, privacy-preserving techniques, and collaborative frameworks that promote equitable benefit-sharing. The study of current literature on FL in agriculture has revealed several promising areas for further research. Designing customized FL algorithms specific to the peculiarities inherent in agricultural data such as seasonal variations, geographical diversity, and multi-modal data types has been described as highly important (Jiang et al., 2020; Wang, 2024). Li et al. (2021) and Wang (2022) emphasize that more research is needed to improve model performance on non-independent and identically distributed (non-IID) data, which is often observed in farming systems due to differences in farming practices, soil conditions, and climate across different regions. Given the unique nature of farm-specific information and its economic implications if leaked, Kumar et al. (2022) and Zhu et al. (2023) assert that it is mandatory to investigate and develop sophisticated privacy-preserving techniques applicable to agricultural data.

Ni et al. (2023) and Yu et al. (2022) highlight that future research should investigate methods to improve the scalability and efficiency of FL systems used in rural areas with limited computational resources and network connectivity. Kamilaris & Prenafeta-Boldú (2018) suggest that integrating FL with edge computing technologies is a promising research direction to enable real-time decision-making and reduce latency in agricultural IoT applications. Additionally, Konečný et al. (2017) emphasize the importance of determining whether cross-silo FL can promote collaboration between different agricultural organizations without compromising data privacy. To ensure the trust and adoption of AI-driven agricultural technologies by farmers, Kairouz et al. (2021) and Li et al. (2020) stress the need for interpretable FL models that offer insights into decision processes. Li et al. (2020) indicate that there is significant potential in utilizing knowledge from data-rich areas in agriculture to enhance the performance of models in less data-abundant regions through federated transfer learning. Furthermore, Zhu et al. (2023) suggest combining blockchain technologies with FL, which warrants further research to ensure reliability and transparency when sharing agricultural data.

Deng et al. (2020) and Li et al. (2022) point out that developing adaptive FL approaches capable of handling the dynamic nature of agricultural environments, including seasonal changes and long-term climate trends, is an important area for future research. Future work should prioritize several key areas to fully realize the potential of FL in agriculture. One crucial direction is developing robust FL frameworks capable of handling the heterogeneous nature of agricultural data. This involves investigating *meta*-learning techniques to train adaptable global models and incorporating transfer learning to leverage pre-trained models for related agricultural tasks. Performance evaluation should consider accuracy, communication efficiency, and robustness to distribution shifts. Simultaneously, enhancing privacy preservation is paramount. Exploring advanced privacy-enhancing technologies (PETs) like homomorphic encryption and secure multi-party computation is

crucial for developing secure aggregation protocols that prevent data leakage during model training. Research should evaluate the trade-off between privacy guarantees and model performance using different privacy budgets, while also analyzing the suitability of various PETs for diverse agricultural applications. Furthermore, designing efficient and scalable FL systems for resource-constrained agricultural environments is essential. This necessitates developing lightweight FL frameworks that operate efficiently on edge devices, investigating model compression techniques, and exploring integration with edge and fog computing architectures to optimize resource utilization and reduce latency. The development of a blockchain-based framework for secure and transparent data sharing is another promising avenue. This involves designing permissioned blockchain networks to record data contributions and model updates, implementing smart contracts for data sharing agreements, and investigating zero-knowledge proofs for verifiable data integrity. Finally, enhancing the interpretability and explainability of FL models for agricultural applications is crucial for building trust and acceptance among farmers. This requires investigating explainable AI (XAI) techniques, developing visualization methods for feature contributions, and focusing on generating understandable and actionable explanations for non-technical users.

By addressing these research directions, future studies can contribute to the advancement of FL in agriculture, resulting in more efficient, secure, and privacy-preserving data analysis and decision-making in the agricultural sector. In summary, while FL presents transformative opportunities for agriculture, it is accompanied by a set of complex challenges that require comprehensive solutions. The concerted efforts to address these limitations will not only enhance the applicability of FL but also contribute to the broader goal of sustainable and equitable agricultural practices.

## 7. Conclusion

In the present work, we have investigated the emerging subject of FL and how it can transform agriculture, especially regarding privacy-protecting data analytics and anonymization methods. We found that FL is an efficient way to handle sensitive agricultural data, an indispensable feature in today's data-driven agriculture, where data sharing is essential for optimizing yields, managing resources, and ensuring sustainable practices. Toward this end, we have explored three main types of FL: horizontal, vertical, and federated transfer learning, highlighting their respective advantages in a range of agricultural scenarios. We have also surveyed various applications of FL in agriculture, including pest and disease detection, precision farming, resource optimization, and crop yield prediction, illustrating FL's practical benefits for improving farming methods while maintaining data privacy.

The present review highlights several challenges in implementing FL in agriculture. These include technical hurdles such as heterogeneous data, limited computational resources, and bandwidth constraints, especially in rural areas, as well as broader social and ethical challenges like data ownership, equitable access to technology, and fairness in model outputs. Despite these complications, FL has significant potential for advancing agricultural innovation by fostering stronger collaboration among farmers, researchers, and other stakeholders while safeguarding confidential data. The ongoing development of more robust privacy-preserving techniques, aligned with user-friendly FL frameworks, will further enable data-driven agriculture to reach its full potential.

Future work should focus on developing more specialized FL algorithms tailored to agricultural data, improving privacy-preserving mechanisms, and tackling social and ethical concerns for equitable technology distribution. Further integration of FL with edge computing and blockchain technology could also facilitate broader deployment and enhance trust, transparency, and operational efficiency across the industry. In short, FL heralds a transformative shift in agricultural data analysis by making collaborative learning possible while preserving individual data privacy. By overcoming present limitations and encouraging responsible innovation, FL can help farmers, researchers, and policymakers fully unlock the power of data-centric agriculture for a more sustainable future.

## CRediT authorship contribution statement

**Rahool Dembani:** Writing – review & editing, Writing – original draft, Visualization, Software, Methodology, Formal analysis, Data curation, Conceptualization. **Ioannis Karvelas:** Writing – review & editing, Supervision. Conceptualization. **Nur Arifin Akbar:** Data curation. **Stamatia Rizou:** Supervision, Conceptualization. **Domenico Tegolo:** Supervision, Conceptualization. **Spyros Fountas:** Supervision.

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Acknowledgments

This work was funded by the European Commission under the Doctoral Networks Programme (MSCA-DN-101073381–EnTrust) within the Horizon Europe (HORIZON) Marie Skłodowska-Curie Actions.

## Data availability

Data will be made available on request.

## References

Adnan, M., Kalra, S., Cresswell, J.C., Taylor, G.W., Tizhoosh, H.R., 2022. Federated learning and differential privacy for medical image analysis. Sci. Rep. 12 (1), 1953. https://doi.org/10.1038/s41598-022-05539-7.

Aggarwal, M., Khullar, V., Goyal, N., Prola, T.A., 2024. Resource-efficient federated learning over IoAT for rice leaf disease classification. Comput. Electron. Agric. 221, 109001. https://doi.org/10.1016/j.compag.2024.109001.

Ahmad, R., Khan, Engr. Dr. F., Khan, S., Haji Mohd, M. N., Mohd, H., Waseem, A., Khan, M. N. A., Ali, S., Jamal, A., & Khan, H. (2022). Federated learning-based UAVs for the diagnosis of Plant Diseases. doi: 10.1109/ICEET56468.2022.10007133.

Akbari, M., Syed, A., Kennedy, W.S., Erol-Kantarci, M., 2023. Constrained federated learning for AoI-limited SFC in UAV-aided MEC for smart agriculture. IEEE Trans. Mach. Learn. Commun. Networking 1, 277–295. https://doi.org/10.1109/TMLCN.2023.3311749.

Albarakati, H., Khan, M., Hamza, A., Khan, F., Kraiem, N., Jamel, L., Almuqren, L., Alroobaea, R., 2024. A novel deep learning architecture for agriculture land cover and land use classification from remote sensing images based on network-level fusion of self-attention architecture. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. PP 1–16. https://doi.org/10.1109/JSTARS.2024.3369950.

Aledhari, M., Razzak, R., Parizi, R. M., Saeed, F., 2020. Federated Learning: A Survey on Enabling Technologies, Protocols, and Applications. IEEE Access, 8, 140699–140725. IEEE Access. doi: 10.1109/ACCESS.2020.3013541.

Al-Hussaeni, K., Fung, B.C.M., Iqbal, F., Liu, J., Hung, P.C.K., 2018. Differentially private multidimensional data publishing. Knowl. Inf. Syst. 56 (3), 717–752. https://doi.org/10.1007/s10115-017-1132-3.

Arora, S., Beams, A., Chatzigiannis, P., Meiser, S., Patel, K., Raghuraman, S., Rindal, P., Shah, H., Wang, Y., Wu, Y., Yang, H., Zamani, M., 2023. Privacy-Preserving Financial Anomaly Detection via Federated Learning & Multi-Party Computation (arXiv:2310.04546). arXiv. doi: 10.48550/arXiv.2310.04546.

Ashworth, A.J., Marshall, L., Volenec, J.J., Casler, M.D., Berti, M.T., van Santen, E., Williams, C.L., Gopakumar, V., Foster, J.L., Propst, T., Picasso, V., Su, J., 2023. Framework to develop an open-source forage data network to improve primary productivity and enhance system resiliency. Agron. J. 115 (6), 3062–3073. https://doi.org/10.1002/agj2.21441.

Bergstrom, A., Nowatzki, J., Witt, T., Barnhart, I., Krueger, J., Askelson, M., Barnhart, K., Desell, T., 2022. Protecting farm privacy while researching large-scale unmanned aircraft systems platforms for agricultural applications. Agron. J. 114 (5), 2700–2714. https://doi.org/10.1002/agj2.21054.

Bongiovanni, R., Lowenberg-Deboer, J., 2004. Precision agriculture and sustainability. Precis. Agric. 5 (4), 359–387. https://doi.org/10.1023/B:PRAG.0000040806.39604.aa.

Chen, J., Liu, Y., 2022. Research and Application Path Analysis of Deep Learning Differential Privacy Protection Method Based on Multiple Data Sources. 299–310. doi: 10.2991/978-94-6463-064-0_34.

Choudhury, O., Gkoulalas-Divanis, A., Salonidis, T., Sylla, I., Park, Y., Hsu, G., Das, A., 2020. Differential Privacy-enabled Federated Learning for Sensitive Health Data (arXiv:1910.02578). arXiv. doi: 10.48550/arXiv.1910.02578.

Deng, Y., Kamani, M.M., Mahdavi, M., 2020. Adaptive Personalized Federated Learning (arXiv:2003.13461). arXiv. doi: 10.48550/arXiv.2003.13461.

Deng, F., Mao, W., Zeng, Z., Zeng, H., Wei, B., 2022. Multiple Diseases and Pests Detection Based on Federated Learning and Improved Faster R-CNN. IEEE Transactions on Instrumentation and Measurement, 71, 1–11. IEEE Transactions on Instrumentation and Measurement. doi: 10.1109/TIM.2022.3201937.

Durrant, A., Markovic, M., Matthews, D., May, D., Enright, J., Leontidis, G., 2022. The role of cross-silo federated learning in facilitating data sharing in the agri-food sector. Comput. Electron. Agric. 193, 106648. https://doi.org/10.1016/j.compag.2021.106648.

Dwarampudi, A., Yogi, M.K., 2024. Application of federated learning for smart agriculture system. Int. J. Inform. Technol. & Comput. Eng. (IJITC) ISSN : 2455-5290, 4(03), Article 03. doi: 10.55529/ijitc.43.36.48.

Dwork, C., 2006. Differential Privacy. In: Bugliesi, M., Preneel, B., Sassone, V., Wegener, I. (Eds.), Automata, Languages and Programming. Springer, pp. 1–12. https://doi.org/10.1007/11787006_1.

Elfares, M., Reisert, P., Hu, Z., Tang, W., Küsters, R., Bulling, A., 2024. PrivatEyes: appearance-based gaze estimation using federated secure multi-party computation. Proc. ACM Hum.-Comput. Interact. 8 (ETRA), 232:1–232:23. https://doi.org/10.1145/3655606.

Elhachmi, J., Kobbane, A., 2022. A Federated Learning Approach for Water Distribution Networks Monitoring. In: 2022 9th International Conference on Wireless Networks and Mobile Communications (WINCOM), pp. 1–6. https://doi.org/10.1109/WINCOM55661.2022.9966455.

Fan, C.-I., Hsu, Y.-W., Shie, C.-H., Tseng, Y.-F., 2022. ID-based multireceiver homomorphic proxy re-encryption in federated learning. ACM Trans. Sen. Netw. 18 (4), 55:1–55:25. https://doi.org/10.1145/3540199.

Gotz, M., Machanavajjhala, A., Wang, G., Xiao, X., Gehrke, J., 2012. Publishing search logs—a comparative study of privacy guarantees. IEEE Trans. on Knowl. and Data Eng. 24 (3), 520–532. https://doi.org/10.1109/TKDE.2011.26.

Hardt, M., Talwar, K., 2010. On the geometry of differential privacy. In: Proceedings of the Forty-Second ACM Symposium on Theory of Computing, pp. 705–714. https://doi.org/10.1145/1806689.1806786.

Hasan, M.K., Habib, A.K.M.A., Islam, S., Safie, N., Ghazal, T.M., Khan, M.A., Alzahrani, A.I., Alalwan, N., Kadry, S., Masood, A., 2024. Federated learning enables 6 G communication technology: Requirements, applications, and integrated with intelligence framework. Alex. Eng. J. 91, 658–668. https://doi.org/10.1016/j.aej.2024.02.044.

Haseeb, K., Ud Din, I., Almogren, A., Islam, N., 2020. An energy efficient and secure IoT-based WSN framework: an application to smart agriculture. Sensors 20 (7). https://doi.org/10.3390/s20072081. Article 7.

He, W., Zhao, L., 2022. Application of federated learning algorithm based on K-means in electric power data. J. New Media 4, 191–203. https://doi.org/10.32604/jnm.2022.032994.

Hosseini, S.M., Sikaroudi, M., Babaei, M., Tizhoosh, H.R., 2022. Cluster based secure multi-party computation in federated learning for histopathology images (arXiv:2208.10919). arXiv. doi: 10.48550/arXiv.2208.10919.

Huang, T., Yan, S., Yang, F., Liu, J., 2016. Multi-domain SDN survivability for agricultural wireless sensor networks. Sensors 16 (11). https://doi.org/10.3390/s16111861. Article 11.

Idoje, G., Dagiuklas, T., Iqbal, M., 2023. Federated learning: crop classification in a smart farm decentralised network. Smart Agric. Technol. 5, 100277. https://doi.org/10.1016/j.atech.2023.100277.

Jayashankar, P., Nilakanta, S., Johnston, W.J., Gill, P., Burres, R., 2018. IoT adoption in agriculture: The role of trust, perceived value and risk. J. Bus. Ind. Mark. 33 (6), 804–821. https://doi.org/10.1108/JBIM-01-2018-0023.

Jiang, B., Li, M., Tandon, R., 2018. Context-aware Data Aggregation with Localized Information Privacy (arXiv:1804.02149). arXiv. doi: 10.48550/arXiv.1804.02149.

Jiang, D., Shan, C., Zhang, Z., 2020. Federated Learning Algorithm Based on Knowledge Distillation (p. 167). doi: 10.1109/ICAICE51518.2020.00038.

Kairouz, P., McMahan, H.B., Avent, B., Bellet, A., Bennis, M., Bhagoji, A.N., Bonawitz, K., Charles, Z., Cormode, G., Cummings, R., D'Oliveira, R.G.L., Eichner, H., Rouayheb, S.E., Evans, D., Gardner, J., Garrett, Z., Gascón, A., Ghazi, B., Gibbons, P. B., Zhao, S., 2021. Advances and open problems in federated learning. foundations and trends®. Mach. Learn. 14 (1–2), 1–210. https://doi.org/10.1561/2200000083.

Kamal, M., Rifat, R., Shruti, A., Haque, M., Gupta, K. D., George, R., 2024. DetectPLD: A Federated CNN Approach for Collaborative Potato Leaf Disease Detection (p. 0191). doi: 10.1109/AIIoT61789.2024.10578980.

Kamilaris, A., Prenafeta-Boldú, F.X., 2018. Deep learning in agriculture: A survey. Comput. Electron. Agric. 147, 70–90. https://doi.org/10.1016/j.compag.2018.02.016.

Kanagavelu, R., Wei, Q., Li, Z., Zhang, H., Samsudin, J., Yang, Y., Goh, R.S.M., Wang, S., 2022. CE-Fed: Communication efficient multi-party computation enabled federated learning. Array 15, 100207. https://doi.org/10.1016/j.array.2022.100207.

Kaur, J., Hazrati Fard, S.M., Amiri-Zarandi, M., Dara, R., 2022. Protecting farmers' data privacy and confidentiality: Recommendations and considerations. Front. Sustainable Food Syst. 6. https://doi.org/10.3389/fsufs.2022.903230.

Konečný, J., McMahan, H. B., Yu, F. X., Richtárik, P., Suresh, A. T., Bacon, D., 2017. Federated Learning: Strategies for Improving Communication Efficiency (arXiv:1610.05492). arXiv. doi: 10.48550/arXiv.1610.05492.

Kumar, P., Gupta, G. P., Tripathi, R., 2022. PEFL: deep privacy-encoding-based federated learning framework for smart agriculture. IEEE Micro, 42(1), 33–40. IEEE Micro. doi: 10.1109/MM.2021.3112476.

Li, Q., He, B., Song, D., 2021. Model-Contrastive Federated Learning (arXiv:2103.16257). arXiv. doi: 10.48550/arXiv.2103.16257.

Li, Y., Qin, X., Chen, H., Han, K., Zhang, P., 2022. Energy-Aware Edge Association for Cluster-based Personalized Federated Learning (arXiv:2202.02727). arXiv. doi: 10.48550/arXiv.2202.02727.

Li, L., Fan, Y., Tse, M., Lin, K.-Y., 2020a. A review of applications in federated learning. Comput. Ind. Eng. 149, 106854. https://doi.org/10.1016/j.cie.2020.106854.

Li, C., Hay, M., Rastogi, V., Miklau, G., McGregor, A., 2010. Optimizing linear counting queries under differential privacy. In: Proceedings of the Twenty-Ninth ACM SIGMOD-SIGACT-SIGART Symposium on Principles of Database Systems, pp. 123–134. https://doi.org/10.1145/1807085.1807104.

Li, Y., Lai, J., Zhang, R., Sun, M., 2024. Secure and efficient multi-key aggregation for federated learning. Inf. Sci. 654, 119830. https://doi.org/10.1016/j.ins.2023.119830.

Li, T., Sahu, A.K., Talwalkar, A., Smith, V., 2020b. Federated learning: challenges, methods, and future directions. IEEE Signal Process Mag. 37 (3), 50–60. https://doi.org/10.1109/MSP.2020.2975749.

Lim, W.Y.B., Luong, N.C., Hoang, D.T., Jiao, Y., Liang, Y.-C., Yang, Q., Niyato, D., Miao, C., 2020. Federated learning in mobile edge networks: a comprehensive survey. IEEE Commun. Surv. Tutorials 22 (3), 2031–2063. https://doi.org/10.1109/COMST.2020.2986024.

Liu, H., Wu, Z., Zhou, Y., Peng, C., Tian, F., Lu, L., 2018. Privacy-preserving monotonicity of differential privacy mechanisms. Appl. Sci. 8 (11). https://doi.org/10.3390/app8112081. Article 11.

Ma, C., Li, J., Ding, M., Yang, H. H., Shu, F., Quek, T.Q.S., Poor, H.V., 2020. On safeguarding privacy and security in the framework of federated learning. IEEE Network, 34(4), 242–248. IEEE Network. doi: 10.1109/MNET.001.1900506.

Ma, J., Naas, S.-A., Sigg, S., Lyu, X., 2022. Privacy-preserving federated learning based on multi-key homomorphic encryption. Int. J. Intell. Syst. 37 (9), 5880–5901. https://doi.org/10.1002/int.22818.

Macaulay, L., Butsic, V., 2017. Ownership characteristics and crop selection in California cropland. Calif. Agric. 71 (4), 221–230. https://doi.org/10.3733/ca.2017a0041.

Mamba Kabala, D., Hafiane, A., Bobelin, L., Canals, R., 2023. Image-based crop disease detection with federated learning. Sci. Rep. 13 (1), 19220. https://doi.org/10.1038/s41598-023-46218-5.

Mehta, S., Kukreja, V., Vats, S., 2023. Empowering Farmers with AI: Federated Learning of CNNs for Wheat Diseases Multi-Classification. In: 2023 4th International Conference for Emerging Technology (INCET), pp. 1–6. https://doi.org/10.1109/INCET57972.2023.10170091.

Muazu, T., Mao, Y., Muhammad, A.U., Ibrahim, M., Kumshe, U.M.M., Samuel, O., 2024. A federated learning system with data fusion for healthcare using multi-party computation and additive secret sharing. Comput. Commun. 216, 168–182. https://doi.org/10.1016/j.comcom.2024.01.006.

Mugunthan, V., Peraire-Bueno, A., Kagal, L., 2020. PrivacyFL: A Simulator for Privacy-Preserving and Secure Federated Learning. In: Proceedings of the 29th ACM International Conference on Information & Knowledge Management, pp. 3085–3092. https://doi.org/10.1145/3340531.3412771.

Ni, W., Zheng, J., Tian, H., 2023. Semi-federated learning for collaborative intelligence in massive IoT networks. IEEE Internet Things J. 10 (13), 11942–11943. https://doi.org/10.1109/JIOT.2023.3253853.

Ongadi, P.A., 2024. A comprehensive examination of security and privacy in precision agriculture technologies. GSC Adv. Res. Rev. 18 (1). https://doi.org/10.30574/gscarr.2024.18.1.0026. Article 1.

Pan, Y., Chao, Z., He, W., Jing, Y., Hongjia, L., Liming, W., 2024. FedSHE: Privacy preserving and efficient federated learning with adaptive segmented CKKS homomorphic encryption. Cybersecurity 7 (1), 40. https://doi.org/10.1186/s42400-024-00232-w.

Park, S.-H., Lee, H., 2022. Completion time minimization of fog-RAN-assisted federated learning with rate-splitting transmission (arXiv:2206.01373). arXiv. doi: 10.48550/arXiv.2206.01373.

Rahman, A., Hossain, M.S., Muhammad, G., Kundu, D., Debnath, T., Rahman, M., Khan, M.S.I., Tiwari, P., Band, S.S., 2023. Federated learning-based AI approaches in smart healthcare: Concepts, taxonomies, challenges and open issues. Clust. Comput. 26 (4), 2271–2311. https://doi.org/10.1007/s10586-022-03658-4.

Rana, O., Spyridopoulos, T., Hudson, N., Baughman, M., Chard, K., Foster, I., Khan, A., 2023. Hierarchical and Decentralised Federated Learning (arXiv:2304.14982). arXiv. doi: 10.48550/arXiv.2304.14982.

Rathod, P.D., Shinde, G.U., 2023. Autonomous aerial system (UAV) for sustainable agriculture: a review. International Journal of Environment and Climate Change 13 (8), 1343–1355. https://doi.org/10.9734/ijecc/2023/v13i82080.

Raturi, A., Thompson, J.J., Ackroyd, V., Chase, C.A., Davis, B.W., Myers, R., Poncet, A., Ramos-Giraldo, P., Reberg-Horton, C., Rejesus, R., Robertson, A., Ruark, M.D., Seehaver-Eagen, S., Mirsky, S., 2022. Cultivating trust in technology-mediated sustainable agricultural research. Agron. J. 114 (5), 2669–2680. https://doi.org/10.1002/agj2.20974.

Reisizadeh, A., Farnia, F., Pedarsani, R., Jadbabaie, A., 2020. Robust Federated Learning: The Case of Affine Distribution Shifts (arXiv:2006.08907). arXiv. doi: 10.48550/arXiv.2006.08907.

Rieyan, S.A., News, M.R.K., Rahman, A.B.M.M., Khan, S.A., Zaarif, S.T.J., Alam, M.G.R., Hassan, M.M., Ianni, M., Fortino, G., 2024. An advanced data fabric architecture leveraging homomorphic encryption and federated learning. Inf. Fusion 102, 102004. https://doi.org/10.1016/j.inffus.2023.102004.

Roth, A., Roughgarden, T., 2010. Interactive privacy via the median mechanism. In: Proceedings of the Forty-Second ACM Symposium on Theory of Computing, pp. 765–774. https://doi.org/10.1145/1806689.1806794.

Runck, B.C., Joglekar, A., Silverstein, K.A.T., Chan-Kang, C., Pardey, P.G., Wilgenbusch, J.C., 2022. Digital agriculture platforms: Driving data-enabled agricultural innovation in a world fraught with privacy and security concerns. Agron. J. 114 (5), 2635–2643. https://doi.org/10.1002/agj2.20873.

Sattler, F., Wiedemann, S., Müller, K.-R., Samek, W., 2019. Robust and communication-efficient federated learning from non-IID data (arXiv:1903.02891). arXiv. doi: 10.48550/arXiv.1903.02891.

Shah, A., Attique Khan, M., Ibrahim Alzahrani, A., Alalwan, N., Hamza, A., Manic, S., Zhang, Y., Damaševičius, R., 2024. FuzzyShallow: A framework of deep shallow neural networks and modified tree growth optimization for agriculture land cover and fruit disease recognition from remote sensing and digital imaging. Measurement 237, 115224. https://doi.org/10.1016/j.measurement.2024.115224.

Shanmugam, L., Tillu, R., Tomar, M., 2023. Federated learning architecture: design, implementation, and challenges in distributed AI systems. J. Knowledge Learning and Sci. Technol. ISSN: 2959-6386 (Online), 2(2), Article 2. doi: 10.60087/jklst.vol2.n2.p384.

Shepherd, M., Turner, J.A., Small, B., Wheeler, D., 2020. Priorities for science to overcome hurdles thwarting the full promise of the “digital agriculture” revolution. J. Sci. Food Agric. 100 (14), 5083–5092. https://doi.org/10.1002/jsfa.9346.

Shi, L., Shu, J., Zhang, W., Liu, Y., 2021. HFL-DP: hierarchical federated learning with differential privacy (p. 7). doi: 10.1109/GLOBECOM46510.2021.9685644.

Siniosoglou, I., Xouveroudis, K., Argyriou, V., Lagkas, T., Margounakis, D., Boulogeorgos, A.-A., Sarigiannidis, P., 2023. Applying federated learning on decentralized smart farming: a case study (p. 1300). doi: 10.1109/ICCWorkshops57953.2023.10283681.

Sun, L., Qian, J., Chen, X., 2021. LDP-FL: practical private aggregation in federated learning with local differential privacy (arXiv:2007.15789). arXiv. doi: 10.48550/arXiv.2007.15789.

Sun, Z., Feng, J., Yin, L., Zhang, Z., Li, R., Hu, Y., Na, C., 2021b. Fed-DFE: A decentralized function encryption-based privacy-preserving scheme for federated learning. Computers, Materials & Continua 71, 1867–1886. https://doi.org/10.32604/cmc.2022.022290.

Suryavanshi, A., Kukreja, V., Mehta, S., Malhotra, S., Manwal, M., 2024. Agriculture Advances: Federated Learning CNN’s for Combatting Banana Leaf Diseases. In: 2024 5th International Conference for Emerging Technology (INCET), pp. 1–6. https://doi.org/10.1109/INCET61516.2024.10592894.

T, M., Makkithaya, K., & G, N.V., 2022. A federated learning-based crop yield prediction for agricultural production risk management. 2022 IEEE Delhi Section Conference (DELCON), 1–7. doi: 10.1109/DELCON54057.2022.9752836.

Tenopir, C., Allard, S., Douglass, K., Aydinoglu, A.U., Wu, L., Read, E., Manoff, M., Frame, M., 2011. Data sharing by scientists: practices and perceptions. PLoS One 6 (6), e21101. https://doi.org/10.1371/journal.pone.0021101.

Thapliyal, N., Kukreja, V., Garg, N., Mehta, S., Anand, J., 2024. Empowering farmers: federated learning CNN for accurate strawberry leaf disease (p. 6). doi: 10.1109/ICRITO61523.2024.10522397.

Truex, S., Liu, L., Chow, K.-H., Gursoy, M. E., Wei, W., 2020. LDP-fed: federated learning with local differential privacy (arXiv:2006.03637). arXiv. doi: 10.48550/arXiv.2006.03637.

Vijh, S., Gaurav, P., Kumar, S., Bansal, P., Singh, M., Khan, M., Palade, V., 2023. USMA-BOF: A novel bag-of-features algorithm for classification of infected plant leaf images in precision agriculture. IEEE Rob. Autom. Mag. PP 2–12. https://doi.org/10.1109/MRA.2023.3315929.

Wang, S., 2024. A study of data heterogeneity in federated learning. Appl. Comput. Eng. 40, 162–167. https://doi.org/10.54254/2755-2721/40/20230643.

Wang, Y., Lin, L., Chen, J., 2023. Communication-efficient adaptive federated learning (arXiv:2205.02719). arXiv. doi: 10.48550/arXiv.2205.02719.

Wang, B., Li, H., Guo, Y., Wang, J., 2023a. PPFLHE: A privacy-preserving federated learning scheme with homomorphic encryption for healthcare data. Appl. Soft Comput. 146, 110677. https://doi.org/10.1016/j.asoc.2023.110677.

Wang, X., 2022. An efficient federated learning optimization algorithm on non-IID data. International Conference on Cloud Computing, Performance Computing, and Deep Learning (CCPCDL 2022), 12287, 87–91. doi: 10.1117/12.2640939.

Wilgenbusch, J.C., Pardey, P.G., Hospodarsky, N., Lynch, B.J., 2022. Addressing new data privacy realities affecting agricultural research and development: A tiered-risk, standards-based approach. Agron. J. 114 (5), 2653–2668. https://doi.org/10.1002/agj2.20968.

Wirth, F.N., Meurers, T., Johns, M., Prasser, F., 2021. Privacy-preserving data sharing infrastructures for medical research: Systematization and comparison. BMC Med. Inf. Decis. Making 21 (1), 242. https://doi.org/10.1186/s12911-021-01602-x.

Woubie, A., Bäckström, T., 2021. Federated learning for privacy preserving on-device speaker recognition. 1–5. doi: 10.21437/SPSC.2021-1.

Wu, X., Zhang, Y., Shi, M., Li, P., Li, R., Xiong, N., 2021. An adaptive federated learning scheme with differential privacy preserving. Futur. Gener. Comput. Syst. 127. https://doi.org/10.1016/j.future.2021.09.015.

Xu, R., Baracaldo, N., Zhou, Y., Anwar, A., Ludwig, H., 2019. HybridAlpha: An Efficient Approach for Privacy-Preserving Federated Learning. In: Proceedings of the 12th ACM Workshop on Artificial Intelligence and Security, pp. 13–23. https://doi.org/10.1145/3338501.3357371.

Xu, J., Glicksberg, B.S., Su, C., Walker, P., Bian, J., Wang, F., 2021. Federated learning for healthcare informatics. J. Healthcare Informatics Res. 5 (1), 1–19. https://doi.org/10.1007/s41666-020-00082-4.

Yang, Q., Huang, A., Fan, L., Chan, C.S., Lim, J.H., Ng, K.W., Ong, D.S., Li, B., 2023a. Federated learning with privacy-preserving and model IP-right-protection. Mach. Intell. Res. 20 (1), 19–37. https://doi.org/10.1007/s11633-022-1343-2.

Yang, X., Huang, W., Ye, M., 2023b. Dynamic personalized federated learning with adaptive differential privacy. Adv. Neural Inf. Proces. Syst. 36, 72181–72192.

Yang, J., Zhang, W., Guo, Z., Gao, Z., 2024. TrustDFL: A blockchain-based verifiable and trusty decentralized federated learning framework. Electronics 13 (1). https://doi.org/10.3390/electronics13010086. Article 1.

Yu, C., Shen, S., Zhang, K., Zhao, H., Shi, Y., 2022. Energy-aware device scheduling for joint federated learning in edge-assisted internet of agriculture things (p. 1145). doi: 10.1109/WCNC51071.2022.9771547.

Zhang, T., He, C., Ma, T., Gao, L., Ma, M., Avestimehr, S., 2021. Federated Learning for Internet of Things: A Federated Learning Framework for On-device Anomaly Data Detection. In: Proceedings of the 19th ACM Conference on Embedded Networked Sensor Systems, pp. 413–419. https://doi.org/10.1145/3485730.3493444.

Zhang, Q., Zhao, X., Han, Y., Yang, F., Pan, S., Liu, Z., Wang, K., Zhao, C., 2023. Maize yield prediction using federated random forest. Comput. Electron. Agric. 210, 107930. https://doi.org/10.1016/j.compag.2023.107930.

Zhu, J., Cao, J., Saxena, D., Jiang, S., Ferradi, H., 2023. Blockchain-empowered federated learning: challenges, solutions, and future directions. ACM Comput. Surv. 55(11), 240: 1 –240, 31. https://doi.org/10.1145/3570953.