Combating Agricultural Challenges with Secure Digital Farming

Cheikhou Akhmed Kane[1][a] and Pascal Francois Faye[2][b]
[1] *Université Rose Dieng France-Sénégal, Dakar, Senegal*
[2] *Department of Mathematics and Computer Science, Université du Sine Saloum El Hadj Ibrahima NIASS, Kaolack, Senegal*

Keywords: Digital Farming, Machine Learning, Distributed Control.

Abstract: This paper introduces Secure Digital Farming, a comprehensive approach to enhancing farm security and optimizing crop yield. SDF addresses critical challenges faced by modern agriculture, including climate change, pest control, rural crime, and demographic pressures, all of which threaten agricultural perimeters and impact yield. Our SDF-based solution leverages deep-learning algorithms to analyze sensor data and video streams from security cameras, enabling intelligent access control, pest detection, and yield estimation. This paper outlines the implementation framework for SDF, highlighting its feasibility for real-life testing and validation. We plan to conduct field tests on our educational farm in the peanut basin of Senegal to evaluate the efficacy and practicality of SDF in a real-world setting.

1 INTRODUCTION

Global agricultural yields are increasingly compromised by climate change, input taxes, government restrictions without farmer consensus, and various security issues within farming environments. Figure 1 illustrates a comparison between a set of security issues and the corresponding security solutions necessary to balance performance and social welfare in farming. Additionally, farmers face arduous labor, high equipment acquisition costs, risks associated with pests (insects, wild animals, domestic animals, etc.), rising rural crime rates, and demographic expansion that threatens agricultural perimeters. This balance is delicate, as attackers—including humans, animals, insects, bacteria, nematodes, and termites—consistently devise methods to circumvent security measures.

Farmers typically secure their farms based on recurrent risks in their area using techniques familiar to aggressors. Therefore, it is crucial to categorize all potential risks, even the least likely ones, to develop comprehensive solutions. However, relying on real-time monitoring by guards or watchdogs has physiological limitations such as sleep, hunger, and exhaustion, which can hinder effective performance. A well-trained distributed artificial intelligence (AI) system can process large volumes of data and make better decisions without experiencing fatigue or biased judgment, offering a more reliable alternative for securing farms.

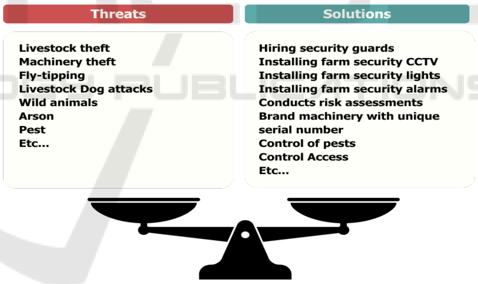

Figure 1: Farm's security solutions versus threats that affects farm effectiveness.

Fig. 2 below shows different field in order to ensure a secured digital farming. Despite technological advances, the management of agricultural perimeters remains a persistent challenge, particularly in regions like the peanut basin of Senegal. In this study, we advocate for the implementation of an artificial intelligence (AI) system to address critical security issues in agricultural fields. The integration of AI has proven beneficial and sustainable across various sectors, including arboricultural farming.

An arboricultural farm is defined by its capacity to produce fruit, operate independently day-to-day, and meet specified thresholds in both area and fruit production. The suitability of soil types, such as clay-sandy, airy, supple, fertile, and cool, varies across

[a] https://orcid.org/0009-0000-9341-2466
[b] https://orcid.org/0000-0002-2078-5891

Kane, C. and Faye, P.
Combating Agricultural Challenges with Secure Digital Farming.
DOI: 10.5220/0013011800003822

In *Proceedings of the 21st International Conference on Informatics in Control, Automation and Robotics (ICINCO 2024) - Volume 1*, pages 376-383
ISBN: 978-989-758-717-7; ISSN: 2184-2809

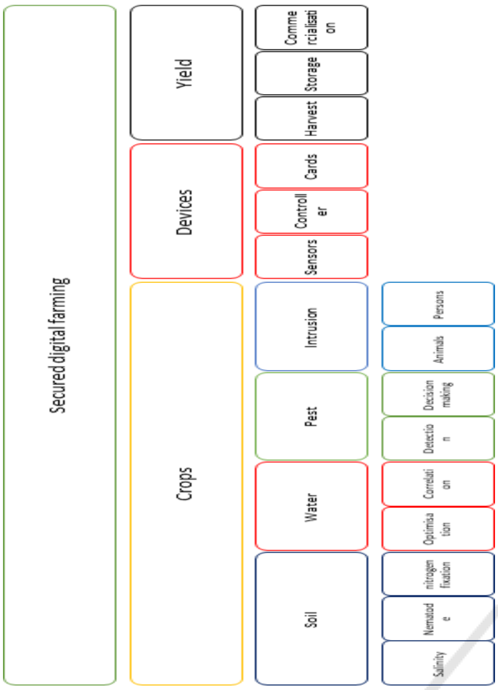

Figure 2: The main components for secure agricultural spaces with the contribution of digital.

the Senegalese peanut basin, influenced further by the challenges posed by climate change. Each type of soil presents unique advantages and challenges. Arboricultural environments are complex due to the size of trees, their foliage, and the diverse activities occurring within them, such as pest monitoring and potential intrusions by animals or humans. Therefore, aiding farmers in identifying these activities, situations, or specific trees on their farms is crucial. This identification process spans various stages of tree growth and fruit development, with particular emphasis on pre-harvest planning. AI is particularly suitable for this task due to its capability to analyze large volumes of data efficiently, identify patterns, and provide predictive insights, thereby enabling farmers to optimize resource allocation and anticipate yield outcomes effectively.

The objective of this work is to develop an application employing various machine learning techniques, such as deep learning and K-nearest neighbors, to enhance farm management and security. This application will be capable of detecting and recognizing activities, machinery, and agronomic equipment on farms. Additionally, it will monitor the farm environment, assist farmers in predicting yields, and help anticipate attacks or intrusions. Our proposed approach focuses on securing access to different categories of crops and trees, thereby preserving yield.

2 RELATED WORK

(Sellam and Poovammal, 2010) investigates the impact of environmental factors, such as annual rainfall and cultivated area, on crop yield using multivariate regression analysis over a 10-year period. The study finds a strong correlation between these factors and yield, as indicated by R^2 value. While the authors acknowledge the potential influence of economic factors (e.g., minimum support price, cost price index, etc..), they do not consider crop security aspects like pest or intrusion detection. (Zhang et al., 2010) compares the accuracy of linear regression and a spatial autoregressive model for predicting corn yield in Iowa. The study demonstrates the superiority of the autoregressive model, which accounts for spatial autocorrelation not captured by OLS, leading to improved predictions. While highlighting the importance of NDVI (Normalized Difference Vegetation Index) and precipitation as predictors, the study does not incorporate other factors like soil health or pest infestations. (Zingade et al., 2018) introduces an Android application and website that leverage machine learning to recommend the most profitable crops based on current weather, soil, and environmental conditions. This system aids farmers in crop selection for long-term profitability but overlooks crop security considerations. (Sun et al., 2022b) focuses on a different aspect of crop management, proposing an improved density peak clustering algorithm for RGB images, incorporating depth information to locate and recognize target fruit (green apples) for harvesting or yield estimation. However, this approach also neglects security factors. (Feng et al., 2022a) focuses on disease identification in crops using image processing techniques, employing morphological operations, contrast stretching, and image scaling for preprocessing, followed by circle-fitting-based segmentation of leaf lesions. Classification is then performed using SVM and random forest models trained on LBP histogram features. While this approach addresses disease detection, it does not consider broader security aspects. (Fu et al., 2022) introduces a rapeseed dataset and a target-dependent neural architecture search for analysis. Addressing the challenges of data loss and misrepresentation in smart agriculture, (Cheng et al., 2022) proposes an anomaly detection model for multidimensional time series data. (Uyeh et al., 2022) employs a multi-objective machine learning approach to optimize sensor placement in a protected cultivation system, utilizing a gradient boosting model with observed and derived environmental variables. (Maia et al., 2022) analyzes sensor data from Australian cotton fields, revealing a correlation between soil matric potential and satellite-derived cumulative crop evapotranspiration, but does not provide yield comparisons. (Feng et al., 2022b) focuses on disrupting the mating of brown planthoppers (BPH), a significant rice pest. They develop a system to record, monitor, and playback BPH courtship vibrations, identifying key frequencies for potential disruption strategies. (Sun et al., 2022a) tackles the problem of monitoring migratory rice pests by developing an intelligent system. This system utilizes a searchlight trap to capture insects, which are then automatically identified using computer vision, providing real-time monitoring data. These studies highlight various aspects of smart agriculture, but a holistic approach that integrates farm security considerations into yield estimation has not been fully addressed. In this work, our goal is to use a set of AI algorithms for the detection and the recognition of objects (person, animals, truck, car, etc.), trees, and fruits in a farm for safety and harvest prediction in a farm environment.

3 SYSTEM DESCRIPTION

Knowledge on farm security and its effects on the various sectors of the national economy is a major challenge for the country's policymakers. Various initiatives are therefore being developed to better identify the implications of risks in the agricultural sector. However, while a causal relationship has clearly been established between the vulnerability of the agricultural sectors and a set of risks on crops, livestock, machinery, etc. Supporting farmers to better manage the risks associated with security is a major necessity. All economic activities which promote food security and suitable agriculture must incorporate the risks of farm management into their planning. The aims of this work are proposing :

1. An intrusion detection system that automates the intrusion detection process.
2. An intrusion prevention system that can detect and also attempt to stop possible incidents.
3. An network topology for a secure digital farming

This by using a full mesh sensor network, a set of IOT card for local and remote control. In order to improve decision making we store information in a local database server and on a ThingSpeak platform. In this way, a farmer can make decisions based on a real-time view of his or her farm. This work combines several Machine learning algorithms (ML) like:

1. Deep Learning (see Fig 3) ;
2. K-Nearest Neighbours (see Fig 4).

This combination delivers a solution that addresses well the dynamism and uncertainty challenges targeted in this work.

3.1 Preliminaries

In this part, we will give a set of analytic view for understanding of our solution. This work is based on AI algorithms like Deep Learning and K-Nearest Neighbours. The Deep Learning is a subcategory of neural networks. Indeed, it is a set of tools and methods of machine learning based on the use of neural networks. We use the words Deep in reference to the number of layers of neurons that make up these networks: the greater the number of layers the deeper the network and the more it allows to treat complex learning problems, but the harder it is to train.

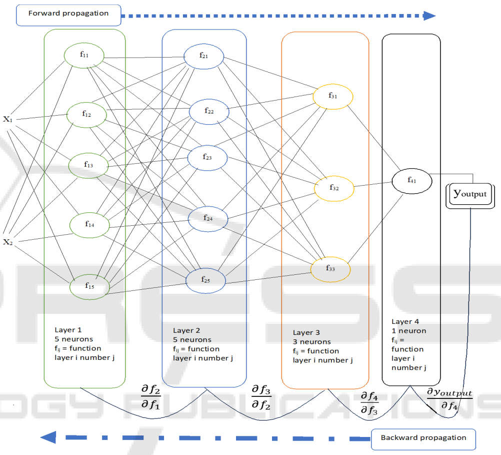
Forward propagation
X₁
X₂
f₁₁
f₁₂
f₁₃
f₁₄
f₁₅
f₂₁
f₂₂
f₂₃
f₂₄
f₂₅
f₃₁
f₃₂
f₃₃
f₄₁
youtput
Layer 1 5 neurons fij = function layer i number j
Layer 2 5 neurons fij = function layer i number j
Layer 3 3 neurons fij = function layer i number j
Layer 4 1 neuron fij = function layer i number j
∂f₂/∂f₁
∂f₃/∂f₂
∂f₄/∂f₃
∂youtput/∂f₄
Backward propagation

Figure 3: Visual neural network with 4 layers.

In these mathematical architectures, each neuron performs simple calculations but the input data passes through several layers of computation before producing an output. The results of the first layer of neurons serve as input to the calculation of the next layer and so on. It is possible to switch on different parameters of the network architecture, like: the number of layers, the type of each layer, the number of neurons that make up each layer. A Deep Learning algorithm can be summarised by the steps:

- Forward propagation
- Cost function
- Backward propagation
- Gradient descendent

Further information about Deep Learning can be found in (Schmidhuber, 2015).

The K-Nearest Neighbours (kNN) algorithm is an ML algorithm that belongs to the class of unsupervised learning algorithms that can be used to solve classification and regression problems. In unsupervised learning, an algorithm receives a data set that is tagged with corresponding output values on which it will be able to train and define a prediction model. This algorithm can subsequently be used on new data in order to predict their output values.

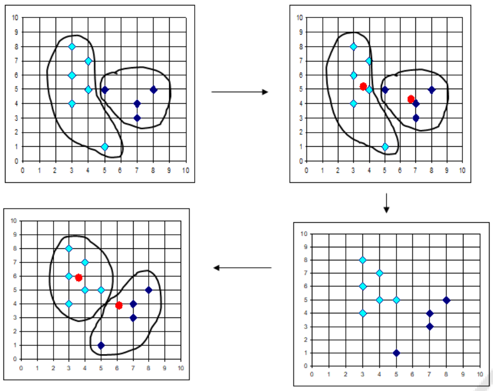

Figure 4: Visual cluster assignment.

The main steps of the KNN algorithm are:

- Step 1: Find the appropriate number of K, by using Elbow method (Umargono et al., 2020);
- Step 2: Select the number K of neighbours;
- Step 3: Calculate Euclidian distance

$$\sum_{i=1}^{n} |X_i - Y_i| \tag{1}$$

or Manhattan distance;

$$\sqrt{\sum_{i=1}^{n} (X_i - Y_i)^2} \tag{2}$$

- Step 4: Take the K-Nearest Neighbours according to the calculated distance.
- Step 5: Count among these K neighbours, the number of points belonging to each cluster.
- Step 6: Assign the new point to the category most present among these K neighbours.
- Step 7: Our model will be ready when all points are in a cluster

3.2 Main Steps of Our Mechanism

In this section, we present the operational architecture and working principles of our mechanism as illustrated in Figures 5 and 6, respectively.

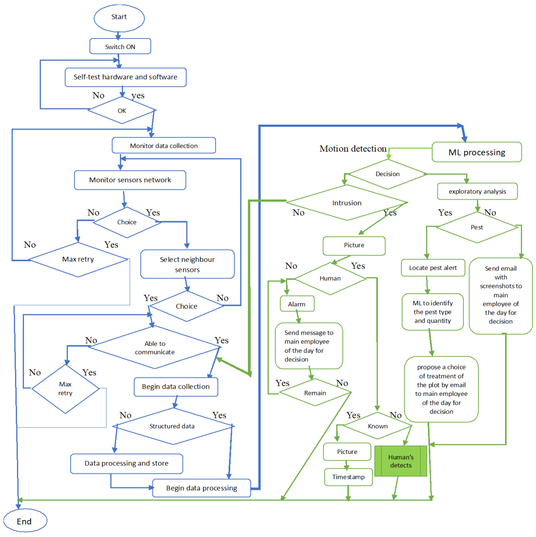

Figure 5: Visual illustration of our mechanism.

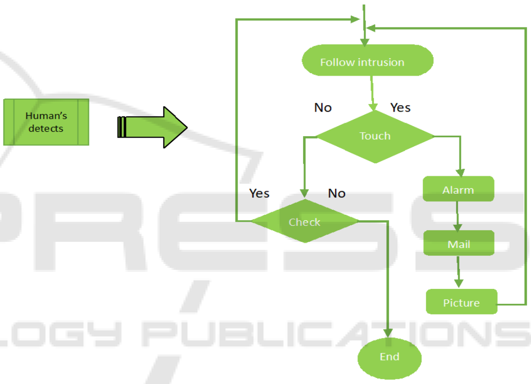

Figure 6: Main steps when a human is detected.

4 IMPLEMENTATION

In this section, we explain the steps we follow for implementation and tests in a real-life environment. Our environment is one hectare with sides of the same size. We started with the virtual grid to identify the number of optimal access points we will need to cover our field. This, to ensure that, there is no white space (without Wi-Fi coverage) that can prevent sensors from transmitting their data. We have used access points with 20-60 meters of reach for a mesh network that seamlessly connects machinery, sensors, etc. ensuring uninterrupted harvesting and other agriculture operations by establishing a self-configuring network.

Thus, the maximum diameter is $D = 120\ meters$ as shown in Fig 8.

- Access point ray coverage

$$20\ meters < r < 60\ meters \tag{3}$$

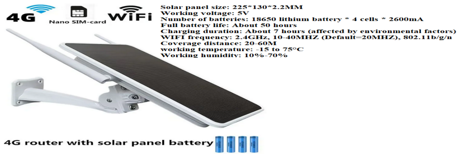
4G
Nano SIM-card
WiFi
Solar panel size: 225*130*2.2MM
Working voltage: 5V
Number of batteries: 18650 lithium battery * 4 cells * 2600mA
Full battery life: About 50 hours
Charging duration: About 7 hours (affected by environmental factors)
WIFI frequency: 2.4GHz, 10-40MHZ (Default=20MHZ), 802.11b/g/n
Coverage distance: 20-60M
working temperature: -15 to 75°C
Working humidity: 10%-70%
4G router with solar panel battery

Figure 7: Access point : 6W 4G solar router;WiFi repeater;4G router solar powered all in one;IP66 Waterproof;Customized according to different regions.

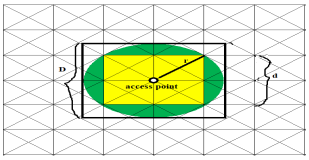

Figure 8: Virtual grid to identify the required access point required for 1 hectare. This to set up a sensor network for data collection. This proposition for 1 hectare is a sampling in order to show how to cover a farm.

- We can have a inside square

$$d = \sqrt{r^2 + r^2} = r\sqrt{2} \tag{4}$$

- On one hectare we will have

$$X = \frac{1\ hectare}{d^2} = \frac{1\ hectare}{2r^2} \tag{5}$$

In this work, we assume that, due to trees and other farm's component which block the signal, the better coverage distance we can have is $r = 35\ meters$. we need $X = \frac{1\ hectare}{2r^2} = 4$ access points in our farm's mesh network with an overlapping signals range (see Fig. 9) to prevent a sensor from being without a Wi-Fi signal. This implementation produce a set of data collected through an IOT network which can be represented by the Fig 10.

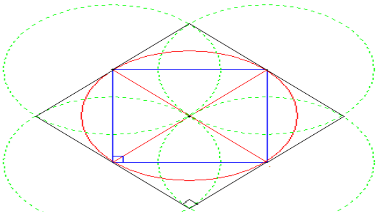

Figure 9: Overlapping signals range to avoid white space.

The data are collected via:

- sensors : ultrasonic, infra-red, thermal, pressure, RGB, UV, rain, temperature, accelerometer, PIR motion, ultrasonic transducer, obstacle detector, vibration detectors.
- cards : ESP 8266, ESP CAM, ARDUINO UNO, servomotors to control equipment (access, sensor, fence, pump, tractor, irrigation, etc.), to have visual data, sound data and create a mesh network of field sensors that will serve as a medium for transmitting data from sensors but also user controls.
  - The different cards are connected to the access point networks to transmit their data. The data is sent simultaneously on the ThingSpeak platform and on our local server via https (Hypertext Transfer Protocol Secure) and SQL (Structured Query Language) requests.
    - ∗ In our local server, we have a web application that implements our SDF algorithm and creates data visualization to support decision making.
    - ∗ ThingSpeak (MathWorks, 2024) is a cloud IoT analytics platform service that allows to aggregate, visualize, and analyze live data streams. We send data to ThingSpeak from our devices, create instant visualization of live data, and send alerts. We could set up more features such as turning on a motor when the water level in the water tank drops below a specified limit er even remotely control devices, such as battery-operated door locks.
      - · Send sensor data privately to our cloud account.
      - · Analyze and visualize data with MATLAB.
      - · Trigger a reaction after a threshold value or a queue up command for a device to execute.
    - ∗
- Viewing data or launching commands can be done via smart-phone or computer.

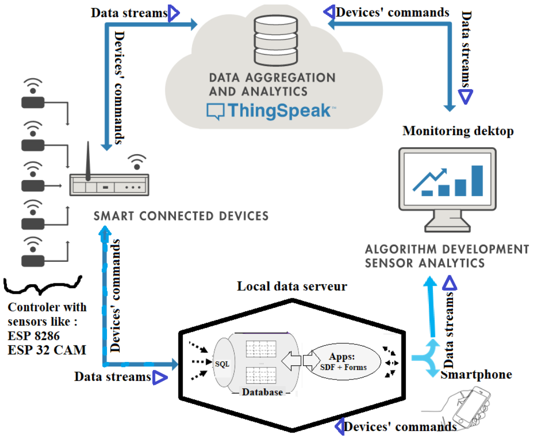

Figure 10: Data network of our Secure Digital Farming Environments.

5 RESULTS

In this section, we give a set of theorems to prove the reliability and feasibility of our system. As we use a set of IOT device, we define a sensor $\sigma \in \Sigma$ which is constrained by the parameters: $\{R\sigma_i, \vartheta^t_{\sigma_i}, L_{Net}\}$. $R\sigma_i$ is its resource(s). A *view* $\vartheta^t_{\sigma_i}$ is the set of sensors in its neighborhood with whom it can directly communicate at time t. L_{Net} defines the dependence level between the received data in a given sensors network (Net).

Corollary 1. *For each Access Point (AP) coverage, it exists an inside square.*

Proof. Let AP be an access point centered at $A = (x,y)$ with ray coverage r. Consider a square with center A and side length $d = r$. The distance from A to any vertex is $\frac{r}{\sqrt{2}}$. Since $\frac{r}{\sqrt{2}} < r$, the vertices of the square lie within the circle.
This proves our corollary. □

Theorem 1. $\forall$ *S as a farm' size, it exists a finite number X of access points (AP) for sensor data transmission.*

Proof. Let r an access point (AP) ray coverage and S a farm' size. Thus, $\forall$ AP coverage $\exists$ $(d*d)$ as a square include in $S = \pi r^2$ ($(d*d) \subset S/d$ is the side of the square (cf. corollary 1).
This proves our theorem, $\forall S, \exists X = \frac{S}{d^2} \in \mathbb{N}$. □

Lemma 1. $\forall$ *S as a farm' size, a sensor belongs at least in one AP' signal range.*

Proof. Let $\sigma_j \in \Sigma$ be a sensor in this network deployed in a farm of size S. Due to our overlapping AP topology signal range, it exists at least one AP which can allow data transmission. In addition, $\forall$ AP it transmits data with MIMO (Multiple Input Multiple Output) strategy and may allows fifty (50) simultaneous sensor data transmission. □

Theorem 2. $\forall$ *S as a farm' size, if an intrusion happen, it exists at least one sensor which can detect the event.*

Proof. Let $\sigma_j \in \Sigma$ a sensor in this network deployed in a farm of size S. Due to our overlapping AP topology signal range and MIMO strategy, there exists at least one AP which can allow data transmission for Cap (cf. Lemma 1). Therefore, if an intrusion happens anywhere within the farm, it will be within the detection range of at least one sensor. Thus, for any farm size S, if an intrusion happens, there exists at least one sensor which can detect the event. □

Lemma 2. *The message costs are bounded and polynomial for each sensor* σ *until the convergence of network.*

Proof. Let $\bar{M}sg$ be the message cost of our sensor network and $\sigma_j \in \Sigma$ be a sensor in this network. For each $\sigma_j \in \Sigma$, if it probes its neighborhood, it sends 2 messages (one for the probe and another for the acknowledgment from the receiver). Additionally, σ_j may receive messages from its neighboring sensors. The number of these neighbors is denoted by $\mathrm{Card}(\vartheta^t_{\sigma_j})$, where $\vartheta^t_{\sigma_j}$ represents the set of neighbors at time t. Thus, $2 \leq \bar{M}sg \leq \mathrm{Card}(\vartheta^t_{\sigma_j})$. For each $\sigma_j \in \Sigma$, if a network layer L_{Net} exists, σ_j sends 1 message to its non-zero neighbor set $\vartheta^t_{\sigma_j} \neq 0$. Additionally, σ_j must send 1 response message after receiving input from its neighbor sensors in $\vartheta^t_{\sigma_j}$. Thus, initially, $\bar{M}sg = 2$ for the probe and acknowledgment. If additional information is received after detection, an extra message is needed, making $\bar{M}sg = 3$. Since the number of sensors is limited in any neighborhood $\vartheta^t_{\sigma_j}$, the value of $\mathrm{Card}(\vartheta^t_{\sigma_j})$ is bounded. Therefore, the message cost $\bar{M}sg$ is not exponential but polynomial in the number of sensors. This proves our lemma. □

Lemma 3. $\forall$ *sensor* σ_j*, the energy cost is optimized.*

Proof. Let $\bar{M}sg$ be the message cost of our sensor network and $\sigma_j \in \Sigma$ be a sensor in this network. $\bar{M}sg$ is bounded and polynomial (cf. Lemma 2). $\sigma_j \in \Sigma$ consumes energy during data transmission, i.e the energy cost is proportional to communication, the number of messages sent. Therefore, it is also bounded and polynomial. This ensures controlled and optimized energy consumption. This proves our lemma. □

Theorem 3. *For each decision-making, it is the better and rational we can do without human interactions.*

Proof. Let σ_j be a sensor in the network. Our overlapping AP topology and MIMO strategy ensure reliable data transmission. All intrusions are detected by at least one sensor σ_j (cf. Theorem 2). This guarantees that sensors have access to complete and accurate data for decision-making. Since the network's communication framework ensures data completeness and reliability, sensors can make well-informed and rational decisions autonomously. The bounded and polynomial nature of the message cost $\bar{M}sg$ and optimized energy use further support this capability by ensuring efficient communication. Thus, the decision-making process within the network is both optimized and rational, operating effectively without human interactions. This proves our theorem. □

Lemma 4. *The near and most reliable sensors of the neighbourhood $\vartheta^t_{\sigma_j}$ of the sensor σ_j which initiate the network have always more ability to be selected as a data relay.*

Proof. Let σ_j be the initiating sensor in the network, and let $\vartheta^t_{\sigma_j}$ denote its neighborhood view. Consider two random sensors σ_k and σ_l from $\vartheta^t_{\sigma_j}$. Let $d(\sigma_j,\sigma_k)$ and $d(\sigma_j,\sigma_l)$ be the distances from σ_j to σ_k and σ_l, respectively. Let $\rho(\sigma_k)$ and $\rho(\sigma_l)$ represent the reliabilities of σ_k and σ_l, respectively. The reliability $\rho(\sigma_i)$ of a sensor σ_i is computed using the Poisson Law (Yates and Goodman, 2005). Suppose the number of times a sensor fails to transmit data follows a Poisson distribution with an average rate of failure occurrences λ per unit time. The number of failures N_f in a time interval t is given by:

$$P(N_f = k) = \frac{(\lambda)^k e^{-\lambda}}{k!}$$

where λ is the expected number of failures in time t. The reliability $\rho(\sigma_i)$ is then defined as:

$$\rho(\sigma_i) = 1 - \frac{\text{Number of failures}}{\begin{array}{c}\text{Total number of}\\ \text{opportunities for transmission}\end{array}}$$

which, in terms of the Poisson parameter λ, becomes:

$$\rho(\sigma_i) = 1 - \frac{\lambda}{\text{Total number of opportunities}}$$

Assume $d(\sigma_j,\sigma_k) < d(\sigma_j,\sigma_l)$, indicating σ_k is closer to σ_j than σ_l, and $\rho(\sigma_k) \geq \rho(\sigma_l)$, implying σ_k has equal or higher reliability compared to σ_l.

Since the reliability $\rho(\sigma_i)$ is inversely related to the number of failures, sensors with a lower rate of failures (higher ρ) are more reliable. The probability of a sensor being selected as a data relay is positively correlated with its reliability and effective communication role. Thus, sensors that are both closer to σ_j and have a higher reliability ρ are more likely to be selected as data relays compared to those further away or with lower reliability. Therefore, sensors closer to σ_j and with higher ρ are more likely to be selected as data relays.

This proves our lemma. □

Theorem 4. *The protocol SDF, is auto-stabilizing if a full connected network is reached.*

Proof. Let Σ be the set of all sensors in the network, and *SDF* be our secure digital farming protocol. We aim to show that *SDF* is auto-stabilizing, meaning it can recover from any transient faults and eventually reach a stable state where all sensors are correctly functioning in a fully connected network.

In a fully connected network, every sensor $\sigma_i \in \Sigma$ can directly communicate with every other sensor $\sigma_j \in \Sigma$. This ensures that each sensor has complete and accurate data about the network.

The message costs $\bar{M}sg$ are bounded and polynomial for each sensor σ_i until the network converges (cf. Lemma 2). This ensures that communication overhead remains manageable and efficient.

From our earlier results, for any farm size S, if an intrusion occurs, there exists at least one sensor that can detect the event. This implies that the protocol can maintain security by ensuring that intrusions are always detected.

Lemma 4 shows that the nearest and most reliable sensors in the neighborhood of any given sensor σ_i have a higher probability of being selected as data relays. Thus, sensors with higher reliability are more likely to be selected, which contributes to the stability of the network.

Given that *SDF* ensures reliable communication and intrusion detection in a fully connected network, the protocol can recover from any transient faults. Even if some sensors temporarily fail or produce incorrect data, the protocol's design ensures that the network will eventually stabilize. The bounded message costs and efficient selection of reliable sensors contribute to this auto-stabilization. Thus, the protocol *SDF* is auto-stabilizing in a fully connected network.

This proves our theorem. □

6 CONCLUSIONS

This paper presents a comprehensive approach to enhancing farm security and optimizing crop yield, introducing the Secure Digital Farming protocol. We have defined the feasibility of our SDF-based solution, which addresses critical challenges faced by modern agriculture, including climate change, pest control, rural crime, and demographic pressures. SDF leverages deep-learning algorithms to analyze sensor data and video streams from security cameras, enabling intelligent access control, pest detection, and yield estimation. We present an implementation framework ready for real-life testing and validation.

To further validate our solution's efficacy and practicality, we will conduct field tests on our educational farm. This controlled environment will allow us to rigorously evaluate its performance and fine-tune its components. Future work will focus on assessing the scalability of our SDF-based approach for application in diverse farming contexts. We will also explore potential collaborations to refine and implement the solution on a larger scale and evaluate the social impact of our technology, particularly its potential to improve farmer livelihoods and promote sustainable agricultural practices.

REFERENCES

Cheng, W., andXiaoting Wang, T. M., and Wang, G. (2022). Anomaly detection for internet of things time series data using generative adversarial networks with attention mechanism in smart agriculture. *Frontiers in Plant Science*, page https://doi.org/10.3389/fpls.2022.890563.

Feng, Q., Wang, S., Wang, H., Qin, Z., and Wang, H. (2022a). Circle fitting based image segmentation and multi-scale block local binary pattern based distinction of ring rot and anthracnose on apple fruits. *Frontiers in Plant Science*, 13:doi: 10.3389/fpls.2022.884891.

Feng, Z., Wei, Q., Ye, Z., Yang, B., Gao, Y., Lv, J., Dai, Y., Bao, J., and Yao, Q. (2022b). Vibrational courtship disruption of nilaparvata lugens using artificial disruptive signals. *Frontiers in Plant Science*, 13:https://doi.org/10.3389/fpls.2022.897475.

Fu, L., Li, S., Rao, Y., Liang, J., Teng, J., and He, Q. (2022). A novel heuristic target-dependent neural architecture search method with small samples. *Frontiers in Plant Science*, page 10.3389/fpls.2022.897883.

Maia, R. F., Lurbe, C. B., and Hornbuckle, J. (2022). Machine learning approach to estimate soil matric potential in the plant root zone based on remote sensing data. *Frontiers in Plant Science*, page https://doi.org/10.3389/fpls.2022.931491.

MathWorks, I. (2024). Data collection in the cloud with advanced data analysis using matlab.

Schmidhuber, J. (2015). Deep learning in neural networks: An overview. neural networks. *Official Journal of the International Neural Network Society*, 61, https://doi.org/10.1016/j.neunet.2014.09.003:pages 85–117.

Sellam, V. and Poovammal, E. (2010). Prediction of crop yield using regression analysis. *IEEE Trans. Knowl. Data Eng.*, pages vol. 23, no. 10, pp. 1498–1512.

Sun, G., Liu, S., Luo, H., Feng, Z., Yang, B., Luo, J., Tang, J., Yao, Q., and Xu, J. (2022a). Intelligent monitoring system of migratory pests based on searchlight trap and machine vision. *Frontiers in Plant Science*, 13:https://doi.org/10.3389/fpls.2022.897739.

Sun, M., Xu1, L., Luo, R., Lu, Y., and Jia, W. (2022b). Fast location and recognition of green apple based on rgb-d image. *Frontiers in Plant Science*, 13:doi: 10.3389/fpls.2022.864458.

Umargono, E., Suseno, J. E., and K., V. G. S. (2020). K-means clustering optimization using the elbow method and early centroid determination based-on mean and median. *International Conferences on Information System and Technology (CONRIST 2019*, DOI: 10.5220/0009908402340240:pages 234–240.

Uyeh, D. D., Iyiola, O., Mallipeddi, R., Asem-Hiablie, S., Amaizu, M., Ha, Y., and Park, T. (2022). Grid search for lowest root mean squared error in predicting optimal sensor location in protected cultivation systems. *Frontiers in Plant Science*, page https://doi.org/10.3389/fpls.2022.920284.

Yates, R. D. and Goodman, D. J. (2005). *Probability and Stochastic Processes: A Friendly Introduction for Electrical and Computer Engineers*. John Wiley and Sons, INC, Rutgers, The State University of New Jersey.

Zhang, L., Lei, L., and Yan, D. (2010). Comparison of two regression models for predicting crop yield. *IEEE International Symposium on Geoscience and Remote Sensing (IGARSS)*.

Zingade, P. D. S., Buchade, O., Mehta, N., Ghodekar, S., and Mehta, C. (2018). Machine learning-based crop prediction system using multi-linear regression. *International Journal of Emerging Technology and Computer Science (IJETCS)*, pages Vol 3, Issue 2.