

Article

Design of an IOTA Tangle-Based Intelligent Food Safety Service Platform for Bubble Tea

Hao-Hsiang Ku ^{1,2} , Ching-Ho Chi ³ and Min-Pei Ling ^{1,2,*}¹ Institute of Food Safety and Risk Management, National Taiwan Ocean University, Keelung City 202301, Taiwan; kuhh@email.ntou.edu.tw² Department of Food Science, National Taiwan Ocean University, Keelung City 202301, Taiwan³ Institute of Clinical Pharmacy and Pharmaceutical Sciences, National Cheng Kung University, Tainan City 701401, Taiwan; s68091029@gs.ncku.edu.tw

* Correspondence: mpling@email.ntou.edu.tw

Abstract: Many food safety incidents have occurred in the world in the past 20 years, causing major threats and harm to human life and health. Each country or region has established different food safety management systems (FSMSs) in response, to increase food safety and to reduce food safety risks. Hence, it is important to develop an FSMS service platform with convenience, consistency, effectiveness, scalability, and lightweight computing. The aim of this study is to design and propose an IOTA Tangle-based intelligent food safety service platform for bubble tea—called IF4BT—which modularizes and integrates hazard analysis and critical control point (HACCP) principles to increase data transparency. The deep learning inference engine is based on long short-term memory and Siamese networks to check and extract significant rare data of high-risk factors, exception factors, and noises, depending on daily check and audit. IF4BT can ensure the correctness of the information of food manufacturers, so as to increase food safety and to reduce food safety issues such as allergen cross-contamination, food expiration, food defense, and food fraud.



Citation: Ku, H.-H.; Chi, C.-H.; Ling, M.-P. Design of an IOTA Tangle-Based Intelligent Food Safety Service Platform for Bubble Tea. *Processes* **2021**, *9*, 1937. <https://doi.org/10.3390/pr9111937>

Academic Editor: Jer-An Lin

Received: 20 September 2021

Accepted: 26 October 2021

Published: 28 October 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: bubble tea; hazard analysis and critical control points; tangle; deep learning; risk factor

1. Introduction

In the past 20 years, many major food safety incidents have occurred in the world, e.g., the melamine incident, plasticizer incident, toxic starch incident, toxic soy sauce incident, gutter oil incident, etc. [1–3]. Food safety incidents can present major threats to human life and health. Governments and industries of various countries are actively investing in the field of food safety [4,5]. Each country or region has formulated different food safety management systems (FSMSs) to increase food safety and reduce food safety risks [6]. Each national accreditation body (AB) and certification body (CB) issues a certification after certifying that the production line or the product meets food safety standards. Manufacturers need to sell goods internationally, so they must have different food safety certifications. The CB needs to audit production status of the manufacturer from time to time. In fact, many problems can be solved by the manufacturer's own internal audit [5]. However, there is a lack of professional confirmation of each process. Because it is uncertain whether or not the internal audit information has been tampered with, a third-party verification agency is still required in order to ensure that all processes are correct.

The latter must conduct internal and external audits and prepare verification procedures and related documents for different certification units. It takes a lot of time to prepare many kinds of documents. It is not easy to save the collected data and to find the missing process. Some unscrupulous manufacturers even have inconsistent documents with the production line, which may cause food safety issues such as food expiration, food defense, or food fraud. Therefore, effective tracking and tracing, and verifying the correctness of the

data—preventing unscrupulous manufacturers from tampering with the content—is an important issue.

Bubble tea—milk tea with tapioca balls—is a very popular ethnic local and international beverage. The ingredients of a cup of bubble tea are tapioca balls, milk, tea, and sugar. Each of its ingredients may have food safety issues, food defense issues, and food fraud issues. Hence, to use IT techniques to construct a standard operating procedure (SOP)-based service platform for food safety is an important issue. Consequently, five problems must be considered: convenience, consistency, effectiveness, scalability, and lightweight computing.

- (1) **Convenience:** Convenience is defined as records being quickly and automatically convertible into officially checkable documents for auditing, applications, or examinations. This can save a lot of labor time and manual input errors. To tackle this problem, a web-based service platform using Extensible Markup Language (XML) and an intelligent mapper can effectively interchange document formats. Extensible Markup Language (XML) is a tag-based language, which can easily migrate, integrate, and transfer data with other services [7];
- (2) **Data consistency:** Consistency is defined as each user seeing a consistent view of the data. Data cannot be deliberately modified. To tackle this problem, the blockchain-based framework is a decentralized architecture. When data are uploaded on the distributed ledger, it records all collected data, transactions, and documents via a tamper-resistant mechanism [8]. To tackle this problem, the distributed ledger ensures that the raw materials, production processes, processing, sales, and audit data in the food safety supply chain are consistent;
- (3) **Effectiveness:** Effectiveness is defined as the ability of participants to effectively track and trace goods in the food safety supply chain. Traditional tracking and tracing of goods require a lot of data to consult. However, the data are often broken, and the paper data and electronic data are interlaced and coexist, so they are not easy to acquire. To tackle this problem, transactions are electronically stored in blockchain or Tangle. Participants can effectively track and trace authorized data and transactions via authorities [9];
- (4) **Scalability:** Scalability is defined as a manufacturer being able to use low-cost computer equipment or a low-power IoT device to create a new node of the blockchain, reducing the manufacturer's construction cost. To tackle this problem, devices can participate in the editing of the distributed ledger. IOTA Tangle is helpful for data transmission and connection by multiple IoT devices [9]. Allowing devices to directly participate in data and message transmission can reduce the chance of data tampering [10];
- (5) **Lightweight computing.** Lightweight computing is defined as a low-power device being able to quickly analyze and confirm exceptional data or conditions. To tackle this problem, a participant uses lightweight equipment—e.g., Raspberry Pi—to automatically set multiple triggers, thresholds, and rules by the server for detecting and computing exceptions of ingredients and foods [11].

Regarding the aforementioned five design issues, this study designs and proposes an IOTA Tangle-based intelligent food safety service platform for bubble tea—called IF4BT—which modularizes and integrates hazard analysis and critical control point (HACCP) principles to increase data transparency. IF4BT is composed of a risk management service platform and a cloud-based deep learning service platform; it can check daily transactions and exceptions. Users can easily trace and track goods. IF4BT can ensure the correctness of the information available to food manufacturers, so as to increase food safety and reduce food safety issues, such as allergen cross-contamination, food expiration, food defense, and food fraud.

The remainder of this paper is organized as follows: Section 2 presents the literature review. Section 3 describes the research methodology. Section 4 contains the results and discussion. Finally, Section 5 presents the conclusion.

2. Literature Review

In recent years, people have experienced various food safety incidents, and the requirements for food safety have become greater and greater. It is important to prevent food safety problems before they happen. Hence, many researchers have devoted efforts to designing and developing various information technology (IT)-based applications to help, check, and reduce food safety risks, which can be roughly divided into the following four categories:

(1) The Internet of Things (IoT)-based Applications:

To deploy various IoT devices—such as RFID, environmental sensors, or GPS to monitor and control food safety risks in the food safety supply chain—Raju et al. [12] proposed near-field and ultrahigh-frequency (UHF) wireless passive sensors for monitoring of food quality indices and food spoilage indicators. Zheng et al. proposed a meat food traceability system based on RFID technology to achieve full process supervision on a food supply chain; the authors built a fault-tolerant RFID mechanism to ensure the practicability of the system [13]. Prajwal et al. designed a device to check the quality of food, using pH sensors, gas sensors, temperature sensors, etc.; this helped them to identify food conditions, preventing the consumption of rotten food [11]. Pal et al. proposed an IoT-based sensing and communications infrastructure for the fresh food supply chain to help reduce food waste; it efficiently improved transportation and distribution, and quickly removed contaminated or spoiled products from the fresh food supply chain [14].

(2) Informational assistance:

To assist manufacturers in inspections and audits during production processes via hazard analysis and critical control point (HACCP)-based food safety standards, Dusenko et al. proposed a system to ensure the safety of food services based on the HACCP method. The proposed system uses digital technology to monitor products and processes during manufacturing, storage, transportation, sales, and disposal [15]. Jing discussed applications of different links in the grain storage and transportation process; the author proposed a food storage and transportation safety information management platform to enhance the informatization process of food enterprises [16]. Gallo et al. proposed a traceability support system to control safety and sustainability indicators in food distribution, aiding practitioners in the monitoring of unsupervised operations throughout the supply chain [17]. Todorović et al. proposed the use of augmented reality to provide relevant information on food production and packaging systems. The authors tried to apply mobile augmented reality in the food market. This can help consumers to obtain information on goods quickly and in a user-friendly manner [18].

(3) To build a blockchain:

To use the decentralized, tamperproof, trackable, and traceable characteristics of the blockchain, all data are stored in the blockchain to ensure that the data will not be arbitrarily tampered with. Yu et al. proposed an architecture based on blockchains and RFID. By scanning the traceable QR code of the foot ring with a smart device, all the data of each process in the chicken life cycle can be obtained [19]. Tao et al. proposed a food safety supervision system based on a hierarchical multidomain blockchain network and a secondary check mechanism. The proposed system can support the timely correction and replacement of malicious nodes through the joint governance of regional nodes, auxiliary verification of supervisory nodes, and arbitration of higher level regions [20]. Tian proposed a supply chain traceability system for food safety based on HACCP, blockchain, and Internet of Things; the information collection stage and food transaction stage are the purpose of the proposed system. All supply chain members are open, transparent, neutral, reliable, and secure [21].

(4) To design a data analysis technology:

To design a technology to analyze data in the supply chain and predict possible food safety incidents, Wahyuni et al. used a Bayesian network to measure and analyze

food safety risks. The proposed method was used to identify risk events in a case study, determine their probability, develop a Bayesian network structure, calculate condition probability tables, and analyze of food safety [22]. Dong et al. proposed a hazard analysis and critical control point (HACCP)-based system combined with fault tree analysis (FTA) to build a traceability model for the production and sales processes of vegetables. The proposed system can ensure food safety and hygiene by controlling key factors in the supply chain [23]. Hebbar used IoT technology to detect the gas released by food, and used deep learning to estimate the probability of food spoilage; if an abnormality was found, an alarm was issued [24].

Of the above-mentioned related studies, most discussed food safety- and risk-related topics using information technology. Table 1 shows a summary of the findings of the literature review.

Table 1. The summary of the findings of the literature review.

Technology	Convenience	Data Consistency	Effectiveness	Scalability	Lightweight Computing
IoT	V [11–14]			V [11–14]	V [12,14]
Informational Assistance			V [15–18]		
Blockchain		V [19–21]	V [19–21]		
Data Analysis			V [22–24]		V [24]

3. Methods

This study designs and proposes an IOTA Tangle-based intelligent food safety service platform for bubble tea, which is called IF4BT. IF4BT modules and divides hazard analysis and critical control points (HACCP), and increases data transparency [25]. IF4BT aims to control food risk, including four aspects:

- (1) Food safety: To list all hazards in the processes, including acceptance testing of raw materials, processing, manufacturing, packaging, storage, and transportation. This ensures food hygiene and food safety, which can reduce the pathogenic bacteria and prevent food poisoning;
- (2) Allergens: In addition to the outer packaging being clearly marked, allergens must be controlled, and sources of allergens must be prevented from cross-contaminating the product;
- (3) Food defense: Prevent food from being physically, chemically, biologically, or radioactively contaminated in a deliberate manner;
- (4) Food fraud: Manufacturers are prohibited from making fakes, counterfeiting, diluting, false labeling, using raw materials from unknown sources, or making false or misleading claims about products.

According to the above-mentioned purpose and design concept, the scenario of IF4BT is shown in Figures 1 and 2: (1) the ingredients are tapioca balls, milk, tea, and sugar; (2) the service platform is based on IOTA Tangle; (3) the food safety management system is based on hazard analysis and critical control point (HACCP) principles; and (4) reports are in the form of audit forms, warning messages, analysis reports, and XML reports.

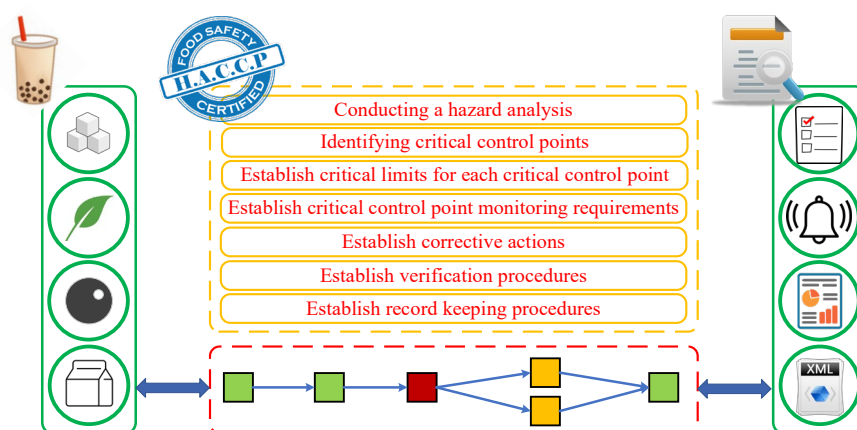


Figure 1. The scenario of IF4BT.

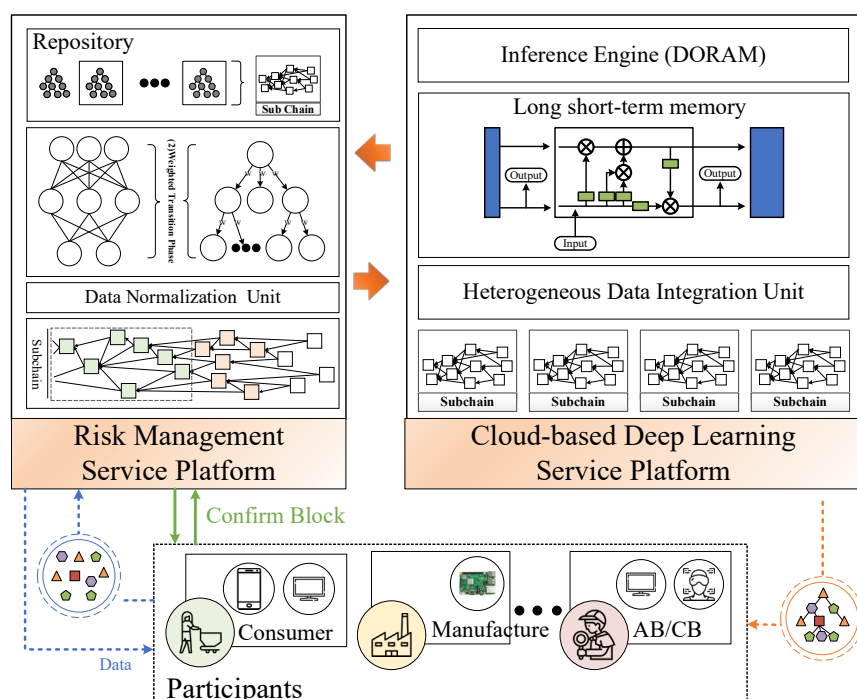


Figure 2. IF4BT framework.

3.1. Participants

The participants in the HACCP-based Tangle system are consumers, manufacturers, accreditation bodies, and certification bodies.

3.1.1. Consumer

Consumers use QR codes to inquire about the purchased products, from raw materials to product sales across the entire food supply chain for the processing and production of the relevant details of the products. Consumers can inquire about product-related information and passed certification marks on the risk management service platform. The data and records of Tangle should comply with the General Data Protection Regulation (GDPR) principles and related Computer-Processed Personal Data Protection Law [26]. If a problematic product is purchased, it can also help consumers report to it to the manufacturer, and quickly notify consumers who have purchased the same batch of products. This can effectively accelerate the entire response to the food safety incident.

3.1.2. Manufacturer

Manufacturers upload relevant documents such as product information, production and sales history, inspection reports, etc., to Tangle. IOTA Tangle distributes and synchronizes ledgers of data in secure, distributed, decentralized, and permissionless environments [9]. After data, transactions, and documents have been uploaded to the IF4BT, they cannot be tampered with. Hence, this information on data, transactions and documents is helpful for audit and improvement processes. When performing an internal audit, it can be supplemented with other equipment information to perform the audit based on HACCP. Furthermore, it allows third-party auditors to quickly understand the on-site situation, and the certification body can continue to understand whether the company's audit behavior is correct. In this way, its production process can be confirmed, as can its coaching improvement status.

3.1.3. Accreditation Body and Certification Body

The accreditation body and certification body mainly store the third-party audit status, missing content, missing correction status, and actual factory visits of the manufacturer to Tangle. In addition, allowing the certification body to objectively understand the internal audit situation of the manufacturer requires (1) reconfirming the internal audit of the manufacturer based on the collected data; (2) allowing the certification body to objectively observe production status; (3) pre-preparation before the actual visit to the factory; and (4) helping auditors to perform key audits in accordance with the established important audit procedures.

Therefore, in order to enable effective communication and sharing of information, relevant information on manufacturers is uploaded to the main chain of the risk management service platform for all relevant parties to view relevant reports.

3.2. Risk Management Service Platform

The risk management service platform is responsible for constructing and maintaining subchains, allowing users to upload raw data, documents, and reports. After transactions are uploaded to the subchain, data of transactions are normalized by the data normalization unit and evaluated through the constructed rules. If an exception occurs, warning messages will be issued. All uploaded data are formatted and normalized by the data normalization unit, providing it to the deep-learning-oriented food safety, risk, reasoning, and analysis model for reasoning out rules. The working flow of the data normalization unit is shown in Figure 3. After data being collected from the supply chain, they are manipulated by the normalization phase and Shapes Constraint Language (SHACL) validation. Data are tagged by XML tags. These data are changed to resource description framework (RDF) serialization formats, which can be stored in Tangle using the same data format.

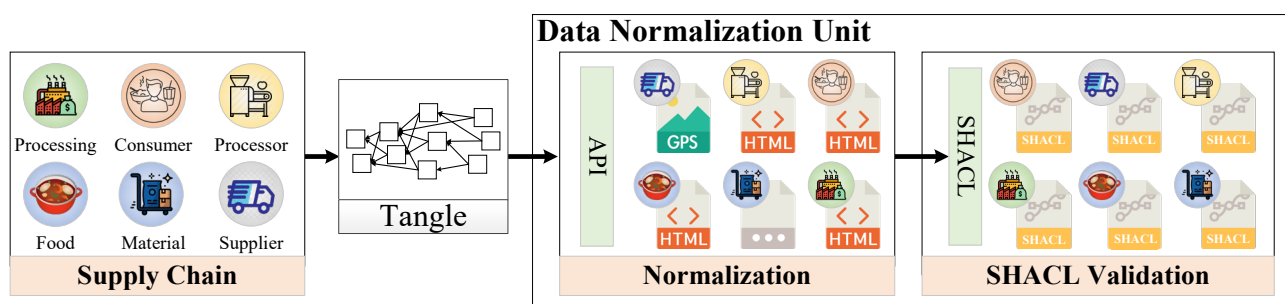


Figure 3. The working flow of the data normalization unit.

Tangle uses decentralized application (DApp) to ensure that the decentralization of the data provided by the participants cannot be controlled or tampered with by a partner. Therefore, Tangle can be used to improve the speed and reliability of tracking and tracing

when food safety incidents occur, and to assist in the construction of inference models based on collected data. The backbone of IF4BT is based on IOTA Tangle—a distributed ledger based on Tangle, which is illustrated in Figure 4. IOTA Tangle is based on a directed acyclic graph (DAG) [9,27]. IOTA Tangle is an improved technology based on blockchain; it is also a distribute-and-synchronize ledger platform that uses blocks and connects the blocks in series [11]. IOTA improves the transaction speed and power consumption; hence, IOTA Tangle is helpful for data transmission and connection by multiple IoT devices [28–30]. Buffered blocks can be quickly proofed and added to Tangle [31,32]. Furthermore, the unspent transaction output (UTXO) model is responsible for making transactions with the accounting balance, where the indicated total inputs are equal to the total outputs and total unspent outputs. The unspent transaction output model is illustrated in Figure 5. UTXO can prevent transaction conflicts, making transactions more effectively.

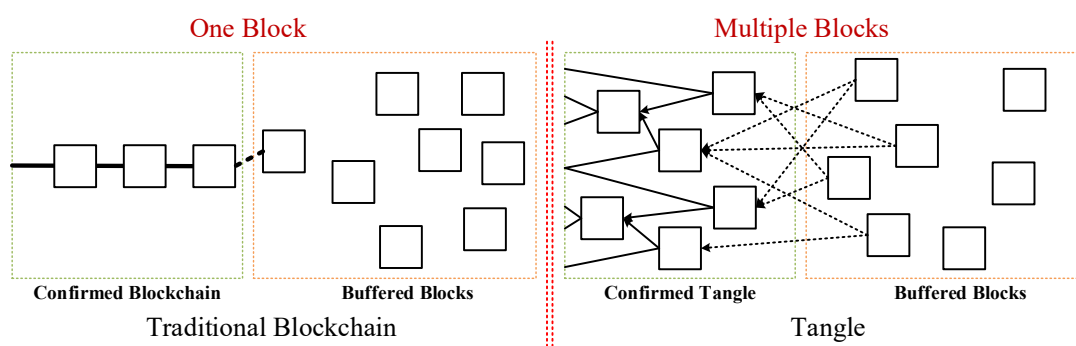


Figure 4. The difference between a traditional blockchain and Tangle.

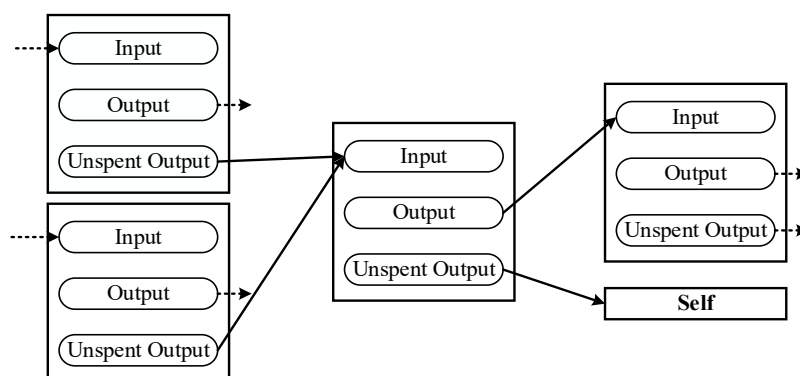


Figure 5. The unspent transaction output model.

3.3. Cloud-Based Deep Learning Service Platform

A cloud-based deep learning service platform is responsible for identifying the influential factors on food safety and risk management. The inference engine is called the deep-learning-oriented food safety, risk, reasoning, and analysis model (DORAM), which is composed of a heterogeneous data integration unit, a weight-added class-imbalance learning unit, and a deep learning training unit, and is used to identify significant rare data. Significant rare data are defined as rarely occurring, and being a serious matter once they do. Figure 6 illustrates the working flow of the deep-learning-oriented food safety, risk, reasoning, and analysis model.

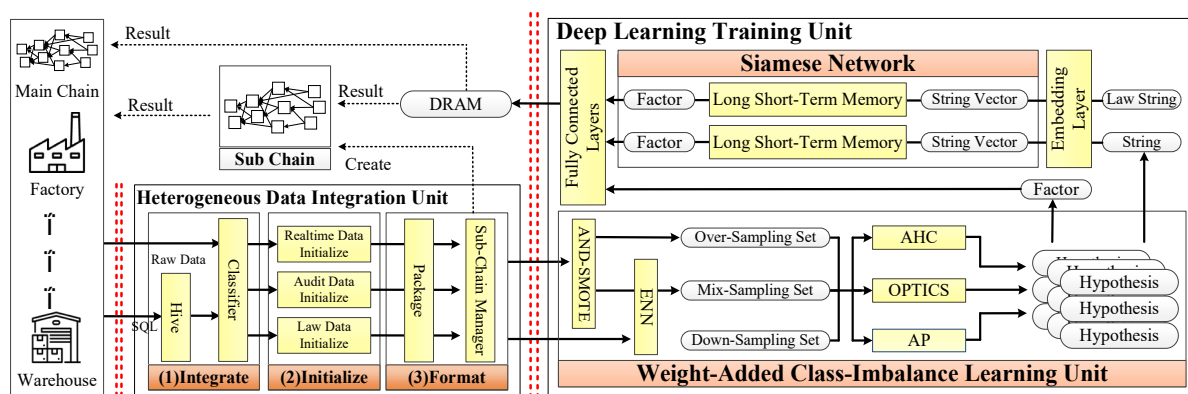


Figure 6. The working flow of the deep-learning-oriented food safety, risk, reasoning, and analysis model.

3.3.1. Heterogeneous Data Integration Unit

The heterogeneous data integration unit is responsible for data normalization and integration [33]. Figure 7 illustrates a detailed flowchart of the heterogeneous data integration unit. The data formats generated by each platform are different from the table defined in each database. If the raw data are not filtered and formatted, the cost of subsequent operations will be greatly increased in data allocation, formatting, and acquisition. Data are tagged as three different types, including real-time data, audit data, and law data. After tagging, data are packaged and then delivered to subchains and the weight-added class-imbalance learning unit.

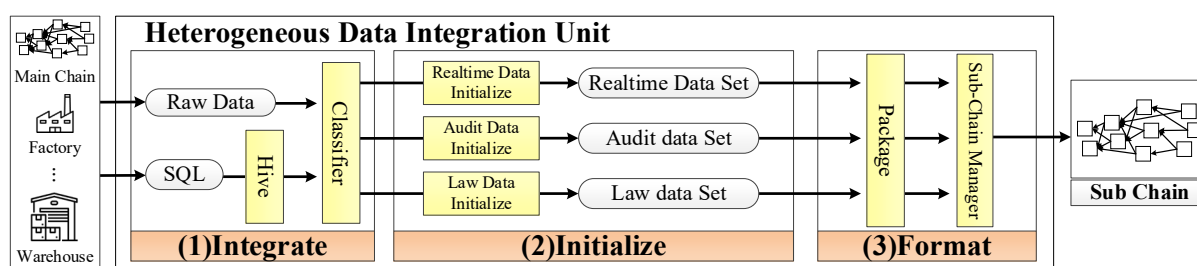


Figure 7. The flowchart of the heterogeneous data integration unit.

3.3.2. Weight-Added Class-Imbalance Learning Unit

The weight-added class-imbalance learning unit is responsible for observing and extracting the significant rare data of high risk. The main purpose of this unit is to establish hypotheses for observing and filtering significant rare data via automatic neighborhood size determination–synthetic minority oversampling technique (AND–SMOTE), edited nearest neighbors (ENN), agglomerative hierarchical clustering (AHC), ordering points to identify the clustering structure (OPTICS), and affinity propagation (AP). The significant rare data are not produced easily, but have a high effect on the food safety. When a manufacturer performs production-related activities, they follow standard procedures. IF4BT monitors exceptional events, and uses the weight-added mode to accurately analyze the occurrence of food safety risk factors. Therefore, this study proposes a weight-added class-imbalance learning unit to solve this problem. The original data use AND–SMOTE to generate new data points, strengthening the weight value of all data for analysis, such that the information will not be ignored because some data are too small [34]. After generating new relevant data points, the weight values of all data are readjusted, using feature extraction algorithms to reduce the amount of data that need to be calculated and to maintain the original appearance of the data [35]. Figure 8 illustrates the flowchart of the weight-added

class-imbalance learning unit, including the oversampling phase, the downsampling phase, and the create hypothesis phase.

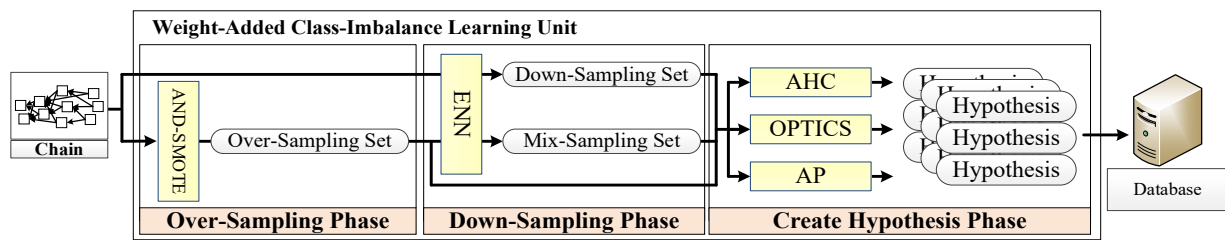


Figure 8. The working flowchart of the weight-added class-imbalance learning unit.

(1) Oversampling phase

In the oversampling phase—although the data have first been processed by the heterogeneous data integration unit—in order to increase effects of the significant rare data, the AND-SMOTE algorithm is responsible for increasing the balance and weight of significant rare data. The AND-SMOTE algorithm is composed of the automatic neighborhood size determination (AND) and the synthetic minority oversampling technique (SMOTE). The AND uses characteristic values to divide data into appropriate adjacent area sizes, as seen in Equations (1)–(3):

$$region_{ik} = \bigcap_{j=1}^m \left[\min(O_{ij}, N_{kj}), \max(O_{ij}, N_{kj}) \right] \quad (1)$$

$$Union_{ik} = \bigcup_{j=1}^k region_{ik} \quad (2)$$

$$label_{ik} = (\hat{y}_1, \hat{y}_2, \dots, \hat{y}_m)^T, \text{ if } \begin{cases} \hat{y}_k = \text{Positive}, \text{ if } x \in Union_{ik} \\ \hat{y}_k = \text{Negative}, \text{ else} \end{cases} \quad (3)$$

where i and j are counters, m is the number of samples, k is the neighborhood size, and T is number of minority class samples.

After determining the neighborhood size k for SMOTE, x_{ref} is a randomly selected object from the minority neighborhood of x by AND. The magnification α is a random variable between 0 and 1, generated from the neighbor x_{new} randomly selected from the overall data object. Equation (4) is used to determine the x_{new} . Equation (5) is used to produce a new dataset S to increase the weighting of the original x .

$$x_{new} = x + (x_{ref} - x) \times \alpha \quad (4)$$

$$S = S \cup \{x_{new}\} \quad (5)$$

(2) Downsampling phase

This phase uses the edited nearest neighbors (ENN) algorithm to clean the database by removing samples close to the decision boundary, based on the heterogeneous value difference metric (HVDM) [36,37]. HVDM considers the similarity of each value to calculate the distances between these values. The ENN algorithm weighs the weight of neighboring points based on the distance between each point and target point, S_q . The weight value is calculated by the ENN algorithm, as described in Equation (6).

$$h(S_q) = \arg \max_{c \in C} \sum_{i=1}^k \frac{1}{d(S_q, \hat{S}_i)^2} \delta(c, f(\hat{S}_i)) \quad (6)$$

C represents 1 to k datasets with Sq . $f(x)$ is a parameter depending on the x type. \hat{S} is the closest k dataset with Sq , and the distance function d calculates its similarity. $\delta(a, b)$ is a parameter depending on a and b ; if a is equal to b , $\delta(a, b) = 1$; otherwise $\delta(a, b) = 0$.

(3) Create hypothesis phase

This phase is responsible for identifying the high-effect significant rare data via a constructed hypothesis, avoiding the wrong results. Hence, this phase is composed of three algorithms: agglomerative hierarchical clustering (AHC), ordering points to identify the clustering structure (OPTICS), and affinity propagation (AP).

Agglomerative hierarchical clustering (AHC) is used to group objects in clusters based on their similarity [38]. Through repeated iterative evolutions, all data are aggregated and classified according to the weights of each layer and, finally, a tree structure is produced. This study used the minimum variance method and Ward's method to construct the agglomerative hierarchical clustering, as shown in Formula (7). The data form clusters, which combine according to their total variance. Ward's method merges the data with the smallest differences in total variance.

$$AHC_{A,B} = n_A \|\bar{x}_A - \bar{x}\|^2 + n_B \|\bar{x}_B - \bar{x}\|^2 \quad (7)$$

Ordering points to identify the clustering structure (OPTICS) is an algorithm for finding density-based clusters in spatial data [39]. The OPTICS algorithm needs to define two parameters: ε and $MinPts$. ε is the maximum distance (radius) centered on this point, while $MinPts$ is the minimum point within the range that needs to be included. ε and $MinPts$ define the core point p , as shown in Equation (8):

$$N_\varepsilon(p) \geq MinPts \quad (8)$$

To calculate the distance from p to each $MinPts_{TH}$, and to calculate the reachability—the distance that can reach p from another point o —the expressions of core distance and reachability distance are shown in Equations (9) and (10), respectively.

$$coreDist_{\varepsilon, MinPts}(p) = \begin{cases} UNDEFINED & , \text{ if } |N_\varepsilon(p)| < MinPts \\ MinPts_{TH} \text{ distance to } N_\varepsilon(p) & , \text{ othrise} \end{cases} \quad (9)$$

$$reachDist_{\varepsilon, MinPts}(o, p) = \begin{cases} UNDEFINED & , \text{ if } |N_\varepsilon(p)| < MinPts \\ \max(coreDist_{\varepsilon, MinPts}(p), dist(p, o)) & , \text{ othrise} \end{cases} \quad (10)$$

According to the constant iteration of the above formula, each o point is divided into the group with the most similar p points. The algorithm continuously calculates and divides the weight value of each candidate p point to each remaining o point, outputting the result when there is no independent or unclassified o point in the space. The OPTICS algorithm does not need to specify the number of clusters in advance; it is a clustering algorithm divided by data density, so it can find highly unique cluster centers that are suitable for constructing hypotheses.

The affinity propagation (AP) algorithm uses the similarity and attractiveness of each point to group data [40]. The best clustering result is based on multiple iterations and weight correction. The calculation process is to set all sample points as the center point of the candidate group, in order to find the exemplar point that best represents the center of the group, and to establish the similarity between all points. Equation (11) describes the similarity:

$$s(i, k) = -||x_i - x_k||^2 \quad (11)$$

After the similarity is calculated, the iterative method is used to continuously adjust the information on the degree of attraction between the sample points—responsibility and availability—as shown in Formulae (12) and (13).

$$r(i, k) \leftarrow s(i, k) - \max_{k' \neq k} \{a(i, k') + s(i, k')\} \quad (12)$$

$$a(i, k) \leftarrow \min\left(0, r(k, k) + \sum_{i' \notin \{i, k\}} \max(0, r(i', k))\right) \text{ for } i \neq k \text{ and } a(k, k) \leftarrow \sum_{i' \notin k} \max(0, r(i', k)) \quad (13)$$

where $r()$ is responsibility and $a()$ is availability. Continuously correcting the $r(i, k)$ and $a(i, k)$ and the position of the group center, until all of the data stop changing, the final group center position can be found.

3.3.3. Deep Learning Training Unit

The deep learning training unit is based on long short-term memory mechanisms, and is responsible for observing the hypotheses of the reasoned effect factors and significant rare data. A traditional artificial neural network can manipulate the data at single point in time. The long short-term memory combines the data of past events in order to comprehensively consider the correlation between time and data. Using the Siamese network, the past section time data and the new section time data can be converted into comparable values in order to further understand whether the production situation has changed. Hence, the key advantage of the deep learning training unit is that it does not use data at a single point in time as the basis for judgment, but combines historical data to analyze new event conditions and understand the current status by comparing the data of the period of time. Figure 9 illustrates a detailed flowchart of the deep learning training unit.

Long short-term memory networks are used to classify, process, and make predictions based on time-series data [41]; they can handle the unknown relationship between important events in a time series, and work via multiple cells, each with an input gate and an output gate. Hence, this study developed them to manipulate the significant rare data and effects. Figure 10 illustrates the working flow of the long short-term memory network.

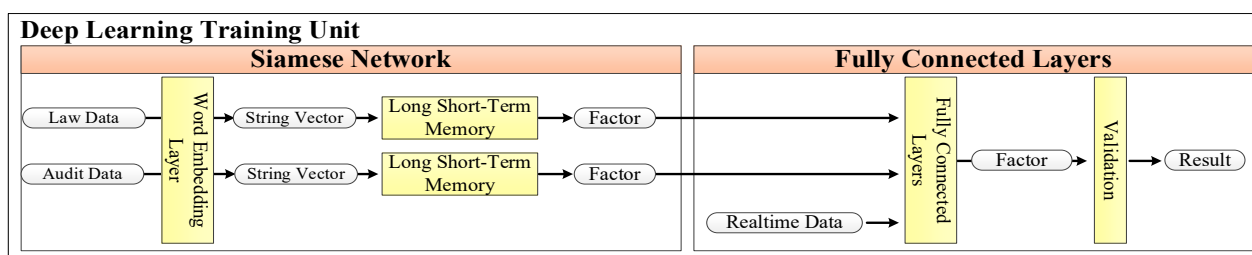


Figure 9. The flowchart of the deep learning training unit.

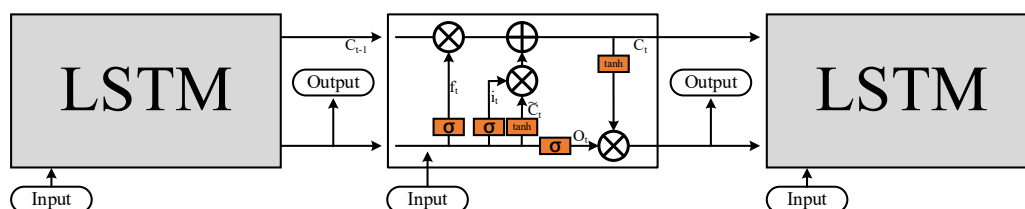


Figure 10. The working flow of the long short-term memory network.

- (1) The first step is to decide whether the information needs to be forgotten. This decision is determined by σ —the forget gate—as shown in Equation (14). The value of σ between the previous output and the current input will be between 0 and 1. If f_t is 1, it means that the memory is completely retained; if f_t is 0, it means that the memory is completely forgotten;

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (14)$$

- (2) Data will be recorded and updated by the value of σ between 0 and 1 from the input gate. If i_t is 1, it means complete replacement; if i_t is 0, it means no replacement is allowed. These processes are shown in Equations (15) and (16). After determining i_t , \tilde{C}_t is based on $\tanh()$ to create a new candidate value and to update the cell state;

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (15)$$

$$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C) \quad (16)$$

- (3) C_t is a new state that combines C_{t-1} and \tilde{C}_t , as described in Equation (17). f_t is used to determine the influence of previous memories from the forget gate, as well as to determine the current state of influence from the input gate;

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t \quad (17)$$

- (4) Finally, the hypothesis value $h(t)$ is used to evaluate whether the significant rare data are memorized or forgotten. The detailed formulae are described in Equations (18) and (19):

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \quad (18)$$

$$h_t = o_t * \tanh(C_t) \quad (19)$$

After constructing the long short-term memory network, this study used a Siamese network to evaluate the status and to check the consistency of auditing and HACCP [42,43]. A schematic of the Siamese network is shown in Figure 11.

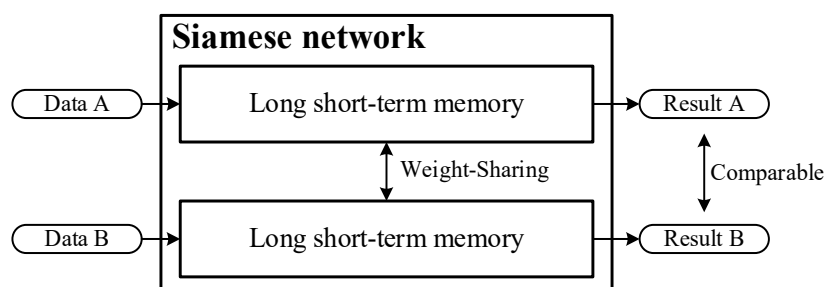


Figure 11. Siamese network used to evaluate the results.

After factors are calculated via the Siamese network formed by the long short-term memory model, the fully connected layer evaluates food safety factors, allergen factors, food defense factors, and food fraud factors. Figure 12 illustrates the fully connected layer used to evaluate the scenario of food safety. Figure 13 illustrates each node of the fully connected layer in detail.

The function of the fully connected layer is to multiply the value of all connected upper-layer nodes by the weight of their edge, then combine all of the products and bias to obtain the value of the node [44]. Finally, the result of the tanh function, with a value between 1 and -1 , is output, as shown in Equation (20).

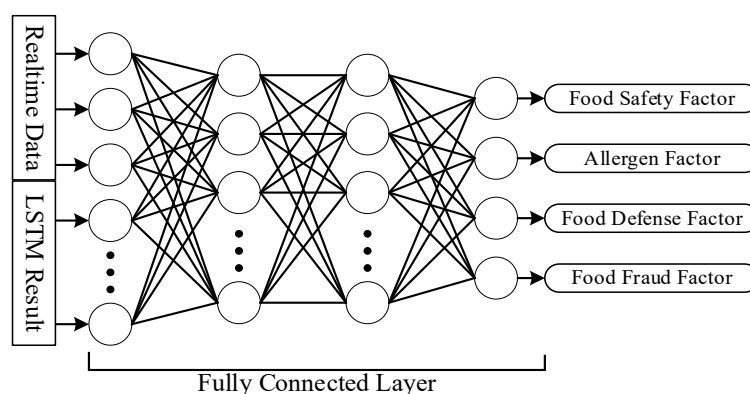


Figure 12. The working flow of the fully connected layer.

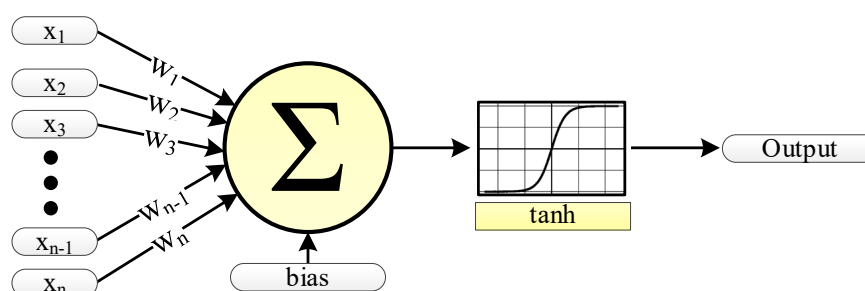


Figure 13. The working flow of each node in the fully connected layer.

$$Output = \tanh\left(\sum (X_n * W_n) + bias\right) \quad (20)$$

Hence, this study designed an inference engine to collect and classify relevant data using heterogeneous data integration units. The significant rare data balance unit is used to find important significant rare data. The deep learning training unit is used to extract the significant rare data and to generate a deep learning model that identifies factors related to food safety risks

4. Results and Discussion

This section presents the IOTA Tangle-based intelligent food safety service platform for bubble tea, including the significant rare data and system interfaces. The current status of each batch number can be seen in the system, and the boxes show the current status of the Tangle system. Figure 14 illustrates the states of Tangle. The green box means that the data are uploaded to Tangle, the orange box means that the data are missing, and the gray box means that is the data are scheduled for production jobs. Each circle represents an HACCP audit point. Green indicates processes that have been completed and passed the CCP, blue indicates current working processes, and gray indicates waiting processes. Figure 15 illustrates the detailed hazard analysis and critical control point of the tea. The production CCPs is checked and recorded. Table 2 shows an example of the HACCP for bubble tea ingredients.

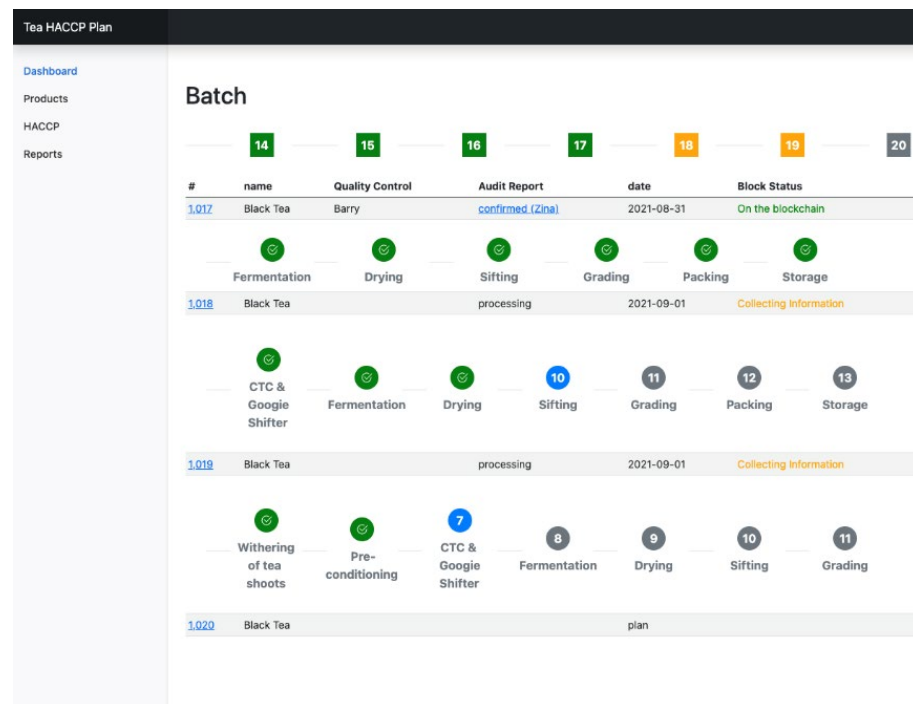


Figure 14. The system implementation of batch tracking.

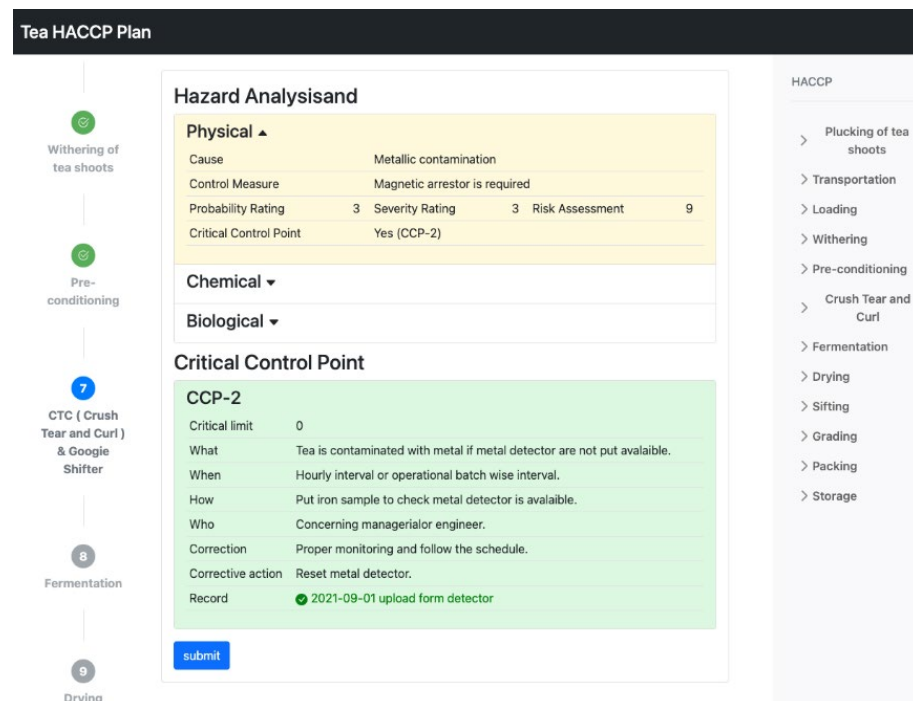


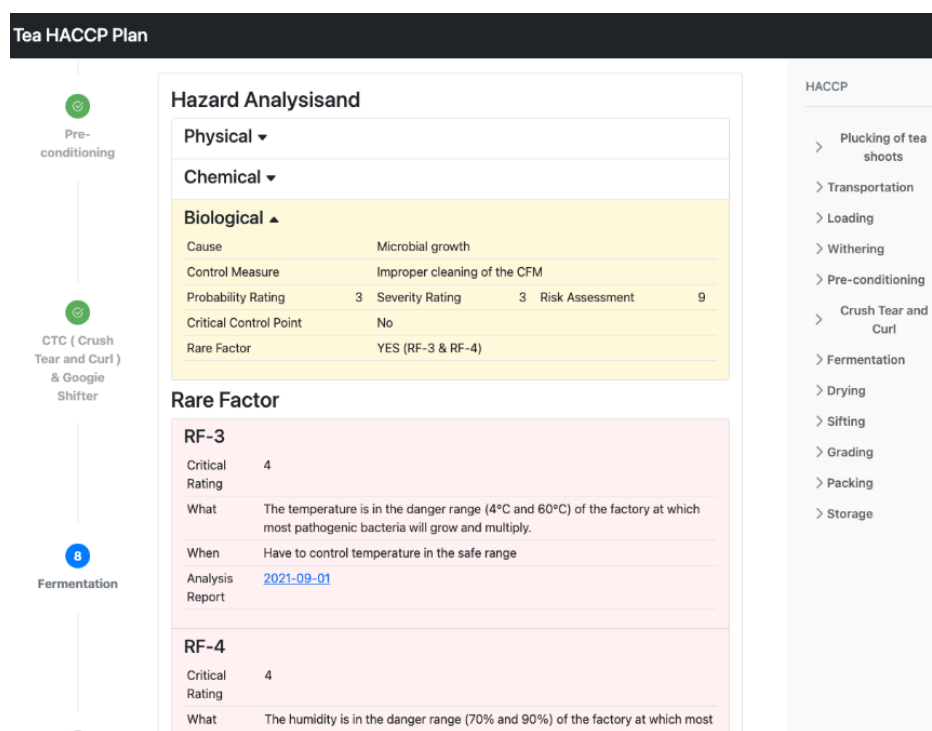
Figure 15. The system implementation of the HACCP audit.

IF4BT observes and extracts significant rare data that are high risk via the deep learning inference engine. Figure 16 illustrates the significant rare data of high risk. When a rare factor is found, IF4BT informs the manager. Although this point is not included in the CCP, it should be considered a factor of the CCP.

Table 2. An example of the HACCP for bubble tea ingredients.

Ingredient	Steps	Hazard	Cause/Source	Control Measure
Tea [45]	Cultivation	Biological	Disease and insect damage; infections	GLOBAL G.A.P.
	CTC and Google Shifter	Physical	Metallic contamination	Magnetic arrestor is required
Milk [46]	Pasteurization	Biological	Incorrect temperature or time range in this step	Using thermometer; stainless steel dryer.
Sugar [47]	Extraction	Chemical	Residual disinfectant and antifoam to the next stage.	Chemical analysis via GC
	Crystallization and Centrifugation	Chemical	Heavy metal contamination	Chemical analysis via GC
	Drying and Cooling	Chemical	Heavy metal contamination	Chemical analysis via GC
Tapioca balls [48]	Drying	Biological	Insufficient drying leading to the development of microorganisms.	Using thermometer; stainless steel dryer.

Figure 17 illustrate the system interface of each CCP. Each CCP is listed by the batch number. The green cycle is the process is completed and passed the CCP. The blue cycle is the current working process. The gray process is the waiting process. If the internal check is finished, the status, the system time and the signature are recorded. Figure 18 illustrates a detailed report of a CCP. If the manager wants to know the detailed information on each check point, they can click the check point to receive it. This will show who performed the check, the detailed time stamp of the check, and the signature of the checker. Finally, the result of the check will be displayed.

**Figure 16.** The system implementation of rare factors.

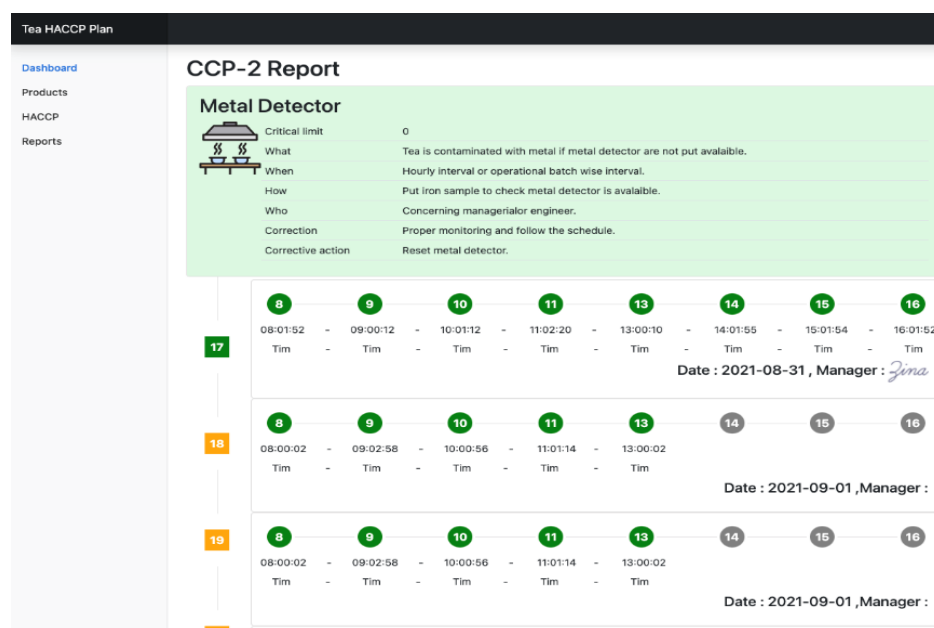


Figure 17. The system implementation of CCP reports for each batch.

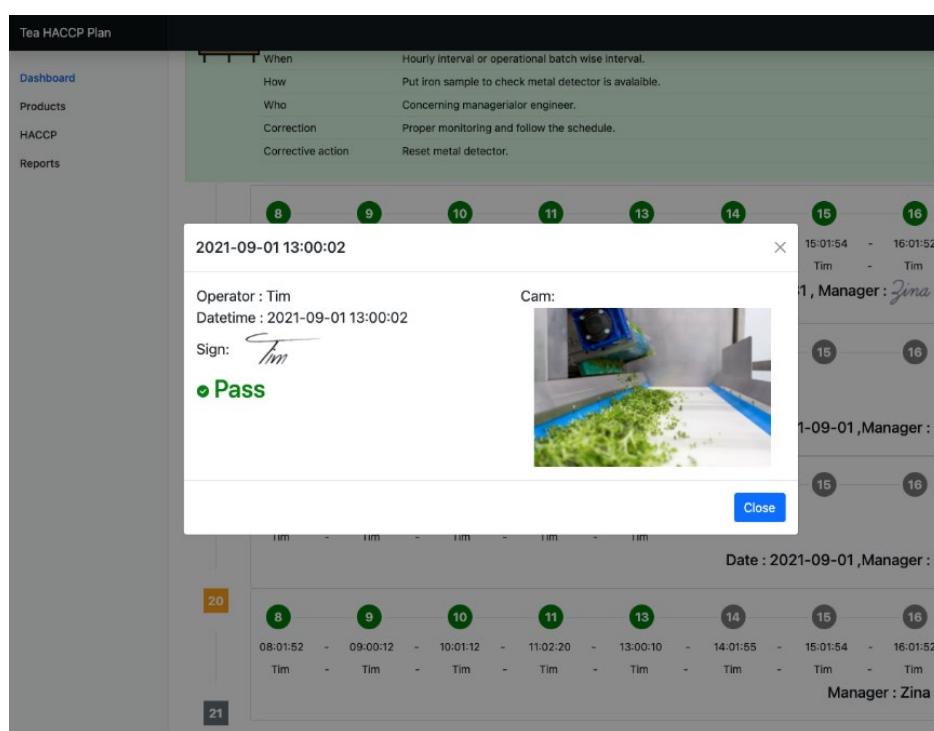


Figure 18. The system implementation of audit detail for CCP reports.

5. Conclusions

This study presents an IOTA Tangle-based intelligent food safety service platform for bubble tea (IF4BT), which modularizes and integrates hazard analysis and critical control point (HACCP) principles to increase data transparency. The inference engine is based on a deep learning mechanism composed of long short-term memory, a Siamese network, and a fully connected layer. IF4BT can enable inspectors to quickly carry out daily inspections in accordance with the HACCP procedures. All data are stored in Tangle for fast tracking and tracing. Five problems are considered and overcome: convenience,

consistency, effectiveness, scalability, and lightweight computing—(1) IF4BT uses a web-based service platform to overcome the problem of convenience; (2) IF4BT uses distributed ledger techniques to overcome the problem of data consistency; (3) IF4BT uses IOTA Tangle to overcome the problems of effectiveness and scalability; (4) IF4BT uses lightweight equipment (a Raspberry Pi) to automatically set reasoned rules via the cloud-based deep learning service platform to overcome the problem of lightweight computing. Hence, the four contributions of this study:

- (1) To develop a verification unit for food safety risk factors. IF4BT is responsible for collecting the manufacturer's usual production information and internal and external audit information in order to analyze the risk factors;
- (2) To construct an HACCP-based food safety management system. It has highly flexible for ISO 22000, FSSC 22000, SQF, and BRCS;
- (3) To design a deep learning inference engine. The inference engine is responsible for identifying the significant rare data of the risk factors for food safety;
- (4) To maintain data consistency and transparency. Using the tamperproof mechanism of the IOTA Tangle platform, the internal inspection data are input to the food safety supply chain—including raw materials, production processes, processing, sales, and audit data—on a daily basis in order to ensure the correctness of the information of the food manufacturer. Each check is an audit to increase food safety and reduce food safety issues such as food expiration, food defense, and food fraud.

Three areas of study are proposed for future researchers: (1) to design and to develop a knowledge base—this increases the HACCP domain knowhow of other products, and can improve the flexibility of the proposed intelligent food safety service platform; (2) to increase other food safety certifications—this increases the ability of data compatibility; and (3) to integrate with the business flow, logistics flow, information flow, and money flow in a service platform. Finally, IF4BT can be a reference model for researchers and engineers.

Author Contributions: Conceptualization, H.-H.K. and C.-H.C.; methodology, H.-H.K. and M.-P.L.; software, H.-H.K. and C.-H.C.; validation, H.-H.K., C.-H.C. and M.-P.L.; formal analysis, H.-H.K. and M.-P.L.; investigation, H.-H.K.; resources, H.-H.K.; data curation, C.-H.C.; writing—original draft preparation, H.-H.K.; writing—review and editing, H.-H.K.; visualization, C.-H.C.; supervision, M.-P.L.; project administration, H.-H.K.; funding acquisition, H.-H.K. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Ministry of Science and Technology of the R.O.C. under the grant MOST 110-2221-E-019 -057-MY2 in support of this research.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: The physical design of this research was supported by MI SHENG ENTERPRISE CO., LTD., Taiwan (<https://www.ms-tw.com/>).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Anal, A.K.; Perpetuini, G.; Petchkongkaew, A.; Tan, R.; Avallone, S.; Tofalo, R.; Nguyen, H.V.; Ky, S.C.; Ho, P.H.; Phan, T.T.; et al. Food safety risks in traditional fermented food from South-East Asia. *Food Control* **2020**, *109*, 106922. [[CrossRef](#)]
2. Walsh, C.; Leva, M.C. A review of human factors and food safety in Ireland. *Saf. Sci.* **2019**, *119*, 399–411. [[CrossRef](#)]
3. Duffy, E.W.; Hall, M.G.; Carpentier, F.R.D.; Musicus, A.A.; Meyer, M.L.; Rimm, E.; Taillie, L.S. Nutrition Claims on Fruit Drinks Are Inconsistent Indicators of Nutritional Profile: A Content Analysis of Fruit Drinks Purchased by Households with Young Children. *J. Acad. Nutr. Diet.* **2021**, *121*, 36–46. [[CrossRef](#)]
4. Evans, E.W.; Lacey, J.; Taylor, H.R. Development and piloting of a support package to enable small and medium sized food and drink manufacturers to obtain third party food safety certification. *Food Control* **2021**, *127*, 108129. [[CrossRef](#)]
5. Dzwolak, W. Assessment of HACCP plans in standardized food safety management systems—the case of small-sized Polish food businesses. *Food Control* **2019**, *106*, 106716. [[CrossRef](#)]

6. Chen, H.; Liou, B.K.; Dai, F.J.; Chuang, P.T.; Chen, C.S. Study on the risks of metal detection in food solid seasoning powder and liquid sauce to meet the core concepts of ISO 22000: 2018 based on the Taiwanese experience. *Food Control* **2020**, *111*, 107071. [CrossRef]
7. Bray, T.; Paoli, J.; Sperberg-McQueen, C.M.; Maler, E.; Yergeau, F. *Extensible Markup Language (XML) 1.0*; World Wide Web Consortium: Cambridge, MA, USA, 2008; p. 5.
8. Popov, S. *The Tangle*; IOTA Foundation: Berlin, Germany, 2018; p. 1.
9. Niya, S.R.; Dordevic, D.; Nabi, A.G.; Mann, T.; Stiller, B. A Platform-Independent, Generic-Purpose, and Blockchain-Based Supply Chain Tracking. In Proceedings of the 2019 IEEE International Conference on Blockchain and Cryptocurrency, Seoul, Korea, 14–17 May 2019; pp. 11–12.
10. Popov, S.; Moog, H.; Camargo, D.; Caposelle, A.; Dimitrov, V.; Gal, A.; Greve, A.; Kusmierz, B.; Mueller, S.; Penzkofer, A.; et al. *The Coordicide*; IOTA Foundation: Berlin, Germany, 2020; p. 1.
11. Prajwal, A.; Vaishali, P.; Payal, Z.; Sumit, D. Food Quality Detection and Monitoring System. In Proceedings of the 2020 IEEE International Students' Conference on Electrical, Electronics and Computer Science, Bhopal, India, 22–23 February 2020; p. 4.
12. Raju, R.; Bridges, G.E.; Bhadra, S. Wireless Passive Sensors for Food Quality Monitoring: Improving the Safety of Food Products. *IEEE Antennas Propag. Mag.* **2020**, *62*, 76–89. [CrossRef]
13. Zhang, Y.; Ruan, Y.; Liu, F.; Shang, J.; Liu, S. Research on Meat Food Traceability System Based on RFID Technology. In Proceedings of the 2019 IEEE 3rd Information Technology, Networking, Electronic and Automation Control Conference, Chengdu, China, 15–17 March 2019; pp. 2172–2175.
14. Pal, A.; Kant, K. IoT-Based Sensing and Communications Infrastructure for the Fresh Food Supply Chain. *Computer* **2018**, *51*, 76–80. [CrossRef]
15. Dusenko, S.V. Digital Technology in Ensuring the Safety of Food Services in the Hospitality Industry. In Proceedings of the 2018 IEEE International Conference Quality Management, Transport and Information Security, Information Technologies, St. Petersburg, Russia, 24–28 September 2018; pp. 618–619.
16. Jing, Z. Application of Information Technology in Food Storage and Transportation Safety Management and Establishment of Information Network Integration Platform for Food Storage and Transportation Safety Management. In Proceedings of the 2018 International Conference on Information Management and Processing, London, UK, 12–14 January 2018; pp. 125–129.
17. Gallo, A.; Accorsi, R.; Goh, A.; Hsiao, H.; Manzini, R. A traceability-support system to control safety and sustainability indicators in food distribution. *Food Control* **2021**, *124*, 107866. [CrossRef]
18. Todorović, V.; Milić, N.; Lazarević, M. Augmented Reality in Food Production Traceability—Use Case. In Proceedings of the IEEE EUROCON 18th International Conference on Smart Technologies, Novi Sad, Serbia, 1–4 July 2019; p. 5.
19. Yu, W.; Huang, S. Traceability of Food Safety Based on Block Chain and RFID Technology. In Proceedings of the 11th International Symposium on Computational Intelligence and Design, Hangzhou, China, 8–9 December 2018; pp. 339–342.
20. Tao, Q.; Cui, X.; Huang, X.; Leigh, A.M.; Gu, H. Food Safety Supervision System Based on Hierarchical Multi-Domain Blockchain Network. *IEEE Access* **2019**, *7*, 51817–51826. [CrossRef]
21. Tian, F. A Supply Chain Traceability System for Food Safety Based on HACCP, Blockchain & Internet of Things. In Proceedings of the 2017 International Conference on Service Systems and Service Management, Dalian, China, 16–18 June 2017; p. 6.
22. Wahyuni, H.C.; Vanany, I.; Ciptomulyono, U. Application of Bayesian Network for Food Safety Risk in Cattle Slaughtering Industry. In Proceedings of the 2019 IEEE International Conference on Industrial Engineering and Engineering Management, Macao, China, 15–18 December 2019; pp. 450–454.
23. Dong, Y.; Fu, Z.; Stankowski, S.; Wang, S.; Li, X. Nutritional quality and safety traceability system for China's leafy vegetable supply chain based on fault tree analysis and QR code. *IEEE Access* **2020**, *8*, 161261–161275. [CrossRef]
24. Hebbar, N. Freshness of Food Detection Using IoT and Machine Learning. In Proceedings of the 2020 International Conference on Emerging Trends in Information Technology and Engineering, Vellore, India, 24–25 February 2020; p. 3.
25. Hazard Analysis Critical Control Point (HACCP). Available online: <https://www.fda.gov/food/guidance-regulation-food-and-dietary-supplements/hazard-analysis-critical-control-point-haccp> (accessed on 20 September 2021).
26. Finck, M. *Blockchain and the General Data Protection Regulation*; European Parliamentary Research Service: Brussels, Belgium, 2019; p. 1.
27. Ferraro, P.; King, C.; Shorten, R. On the stability of unverified transactions in a dag-based distributed ledger. *IEEE Trans. Autom. Control.* **2019**, *65*, 3772–3783. [CrossRef]
28. Kelly, F.P.; Maulloo, A.K.; Tan, D.K.H. Rate Control for Communication Networks: Shadow Prices, Proportional Fairness and Stability. *J. Oper. Res. Soc.* **1998**, *49*, 16. [CrossRef]
29. Abdullah, M.A.; Draief, M. Global Majority Consensus by Local Majority Polling on Graphs of a Given Degree Sequence. *Discret. Appl. Math.* **2015**, *180*, 1–10. [CrossRef]
30. Benjamini, I.; Chan, S.O.; O'Donnell, R.; Tamuz, O.; Tand, L.Y. Convergence, Unanimity and Disagreement in Majority Dynamics on Unimodular Graphs and Random Graphs. *Stoch. Process. Appl.* **2016**, *126*, 2719–2733. [CrossRef]
31. Gärtner, B.; Zehmakan, A.N. Majority Model on Random Regular Graphs. In Proceedings of the 13th Latin American Symposium on Theoretical Informatics, Istanbul, Turkey, 13–15 May 2009; pp. 572–583.
32. Neudecker, T.; Hartenstein, H. Network Layer Aspects of Permissionless Blockchains. *IEEE Commun. Surv. Tutor.* **2018**, *21*, 838–857. [CrossRef]

33. Szekely, G.; Rizzo, M. Hierarchical Clustering via Joint Between-Within Distances: Extending Ward's Minimum Variance Method. *J. Classif.* **2005**, *22*, 151–183. [\[CrossRef\]](#)
34. Yun, J.; Ha, J.; Lee, J.S. Automatic Determination of Neighborhood Size in SMOTE. In Proceedings of the 10th International Conference on Ubiquitous Information Management and Communication, Danang, Vietnam, 4–6 January 2016; p. 8.
35. Batista, G.E.A.P.A.; Prati, R.C.; Monard, M.C. A Study of the Behavior of Several Methods for Balancing Machine Learning Training Data. *SIGKDD Explor.* **2004**, *6*, 20–29. [\[CrossRef\]](#)
36. Wilson, D.R.; Wilson, T.; Martinez, T.R. Reduction Techniques for Exemplar-Based Learning Algorithms. *Mach. Learn.* **2000**, *38*, 257–286. [\[CrossRef\]](#)
37. Wilson, D.L. Asymptotic Properties of Nearest Neighbor Rules Using Edited Data. *IEEE Trans. Syst. Man Cybern.* **1972**, *SMC-2*, 408–421. [\[CrossRef\]](#)
38. Day, W.H.; Edelsbrunner, H. Efficient algorithms for agglomerative hierarchical clustering methods. *J. Classif.* **1984**, *1*, 7–24. [\[CrossRef\]](#)
39. Ankerst, M.; Breunig, M.M.; Kriegel, H.P.; Sander, J. OPTICS: Ordering points to identify the clustering structure. *ACM Sigmod Rec.* **1999**, *28*, 49–60. [\[CrossRef\]](#)
40. Frey, B.J.; Dueck, D. Clustering by passing messages between data points. *Science* **2007**, *315*, 972–976. [\[CrossRef\]](#) [\[PubMed\]](#)
41. Hochreiter, S.; Schmidhuber, J. Long Short-Term Memory. *Neural Comput.* **1997**, *9*, 1735–1780. [\[CrossRef\]](#) [\[PubMed\]](#)
42. Chopra, S.; Hadsell, R.; LeCun, Y. Learning a Similarity Metric Discriminatively, with Application to Face Verification. In Proceedings of the 2015 Computer Vision and Pattern Recognition, Boston, MA, USA, 7–12 June 2015; p. 8.
43. Norouzi, M.; Fleet, D.J.; Salakhutdinov, R. Hamming Distance Metric Learning. In Proceedings of the Advances in Neural Information Processing Systems, Lake Tahoe, NV, USA, 3–6 December 2012; p. 9.
44. Hagan, M.T.; Demuth, H.B.; Beale, M. *Neural Network Design*, 2nd ed.; PWS Publishing Co.: Stillwater, OK, USA, 2014.
45. Kotzamanidis, C.Z.; Arvanitoyannis, I.S.; Skaracis, G.N.; Hadjiantoniou, D.C. Implementation of hazard analysis critical control point (HACCP) to a production line of beet sugar, molasses and pulp: A case study. *Zuckerindustrie* **2000**, *125*, 970–977.
46. Ahmad, I.; Chowdhury, R.S.; Uddin, R.; Shakawat, A.; Rahman, W.U.; Zzaman, W. HACCP Plan and Adoption of HACCP Metasystem in the Tea Industries of Bangladesh. *Asian Food Sci. J.* **2019**, *13*, 1–11. [\[CrossRef\]](#)
47. Tabeen, J.; Yadav, K.C.; Sujit, B. Study of HACCP implementation in milk processing plant at Khyber Agro Pvt. Ltd in Jammu & Kashmir. *J. Food Process Preserv.* **2016**, *7*, 610.
48. Thao, N.T.; Tu, N.T.M.; Liem, N.T.; Huong, N.M.; Pintado, M.; Anh, T.K. Investigation on food safety programs in cassava production at SME level in Vietnam. In Proceedings of the VBFoodNET 2013 Conference, Hanoi, Vietnam, 11 November 2013; p. 7.