Article

An Adaptive Control Algorithm Based on Q-Learning for UHF Passive RFID Robots in Dynamic Scenarios

Honggang Wang [1], **Ruixue Yu** [1], **Ruoyu Pan** [1], **Peidong Pei** [1], **Zhao Han** [1], **Nanfeng Zhang** [2] **and Jingfeng Yang** [3,*]

[1] School of Communications and Information Engineering and School of Artificial Intelligence, Xi'an University of Posts and Telecommunications, Xi'an 710121, China
[2] Huangpu Customs District Technology Center, Guangzhou 510730, China
[3] Guangzhou Institute of Industrial Intelligence, Guangzhou 511458, China
* Correspondence: jingfengyang@126.com

Abstract: The Identification State (IS) of Radio Frequency Identification (RFID) robot systems changes continuously with the environment, so improving the identification efficiency of RFID robot systems requires adaptive control of system parameters through real-time evaluation of the IS. This paper first expounds on the important roles of the real-time evaluation of the IS and adaptive control of parameters in the RFID robot systems. Secondly, a method for real-time evaluation of the IS of UHF passive RFID robot systems in dynamic scenarios based on principal component analysis (PCA)-K-Nearest Neighbor (KNN) is proposed and establishes an experimental scene to complete algorithm verification. The results show that the accuracy of the real-time evaluation method of IS based on PCA-KNN is 92.4%, and the running time of a single data is 0.258 ms, compared with other algorithms. The proposed evaluation method has higher accuracy and shorter running time. Finally, this paper proposes a Q-learning-based adaptive control algorithm for RFID robot systems. This method dynamically controls the reader's transmission power and the robot's moving speed according to the IS fed back by the system; compared with the default parameters, the adaptive control algorithm effectively improves the identification rate of the system, the power consumption under the adaptive parameters is reduced by 36.4%, and the time spent decreases by 29.7%.

Keywords: RFID robots; dynamic scenarios; identification state; adaptive control; Q-learning

MSC: 68T05

1. Introduction

UHF passive RFID technology has been widely used in various industries due to its advantages of low cost, long-distance, and rapid batch identification [1–3]. In large-scale application scenarios such as unmanned warehouses, clothing retail, and file management, the model of fixed parameters and the traditional statically deployed RFID systems no longer meet the performance requirements. In recent years, industrial applications have applied RFID technology to mobile robots, drones, or conveyor belts, and the identification scenarios have changed from static to dynamic [4–6]. The development of intelligent mobile identification requires the continuous integration of RFID technology with new technologies such as automatic control, 5G technology, intelligent computing, and deep learning. It is a new trend in RFID technology for dynamic application scenarios.

The combination of mobile robots and RFID technology has become an important way of mobile identification. AdvanRobots is an autonomously moving UHF passive RFID robot that is equipped with six RFID antennas on each side, which can automatically count goods inventory in a given space, which is more accurate than RFID handheld devices to count inventory [7]. Equipped with a reader and multiple antennas, the UHF RFID robot can move within the target scene to detect multiple tags and locate the tags

through a synthetic array method [8]. Robots combine RFID technology to achieve path navigation, query, and positioning of target objects [9–11]. Ref. [12] presents a cloth-dressing robot system, which uses RFID as key elements for data management to command an adaptive cloth-dressing robot control with a fuzzy-PID controller, which is used to adjust the robot's posture. [13] proposes a UHF-RFID mobile robot platform, which uses eight parallel channels for multiple-input multiple-output localization and uses three-dimensional product maps for inventory counting but does not consider the identification efficiency and time. In addition, more RFID robots are used for positioning, rather than for object identification in dynamic scenarios [14–16].

The identification efficiency of RFID systems will be affected by the constant changes of the identification environment and the quantity and medium of tags in dynamic scenarios. In order to ensure high efficiency and stability of the identification efficiency, the RFID systems must dynamically adjust parameters according to the real-time IS. There have been many studies on the rate of identification (RoI) of RFID systems deployed in static scenarios or when only tags are moving [17], but less studies on the IS and adaptive control of RFID systems in dynamic scenarios.

Aiming at the above problems, this paper first expounds on the effect of real-time evaluation of the IS and adaptive control in the RFID robot systems and divides the IS according to real-time RoI and the difference between theoretical value and actual value of speed of identification (SoI). Secondly, a real-time evaluation method of the IS for UHF passive RFID robot systems in dynamic scenarios is proposed. Finally, this paper proposes a Q-learning-based adaptive control algorithm for RFID robot systems.

The remainder of this paper is organized as follows. Section 2 introduces the RFID systems in dynamic scenarios, Section 3 proposes the real-time evaluation model and theoretical algorithm of the IS, and Section 4 presents the dynamic scene test and the calculation and analysis of the real-time evaluation of the IS for the RFID robots. A Q-learning-based adaptive control algorithm for RFID robot systems is proposed in Section 5, and the paper is concluded in Section 6.

2. The IS of RFID Systems in Dynamic Scenarios

2.1. RFID Identification in Dynamic Scenarios

Typical RFID applications are usually static identification, such as fixed bayonet, channel, or using handheld devices to identify tags. The air interface parameter Q of the static RFID systems dynamically adjusts the frame length according to the collision of tags within the identification range of the reader to improve the system throughput. The RFID applications in dynamic scenarios are no longer limited to a fixed pattern, as readers are being mounted on mobile robots, drones, AGVs, and other mobile devices. It is foreseeable that RFID robots will replace static identification in warehousing, logistics, and other application scenarios to complete mobile intelligent inventory.

In dynamic identification scenarios, the reader and tags are always relatively moving. The low identification efficiency of RFID systems is caused by problems such as random post-identification and missed reading of tags. The complexity of RFID dynamic identification application scenarios limits the performance improvement of the traditional Q algorithm, and the improvement of system identification efficiency is limited. A new direction for RFID technology is to apply machine learning or reinforcement learning methods to cluster, evaluate, and predict the IS and environment of RFID systems in real-time, and realize adaptive control of parameters.

2.2. Real-Time Evaluation of the IS in RFID Robot Systems

The UHF RFID robot systems with automatic control capability, local computing, and cloud-based remote communication was designed and implemented in a previous study [18].

In this paper, the existing RFID robot hardware is redesigned and implemented. The hardware is divided into five modules: the MT7620-based main control, the algorithm

module, the robot chassis, the M6 four-channel RFID Reader, and the two-degree-of-freedom steering group. The structure uses the robot chassis to load the RFID antenna array, where each antenna is mounted on two-degree-of-freedom steering. In terms of function, it first has the basic function of an RFID application system, which can read and write tags. Additionally, antennas can be adjusted in attitude in response to steering tilt and heading adjustments. Furthermore, the robot chassis will allow automatic path planning and adaptive movement control in indoor environments. Figure 1 is the hardware topology of the RFID robot systems, and Figure 2 is the physical picture of the RFID robot.
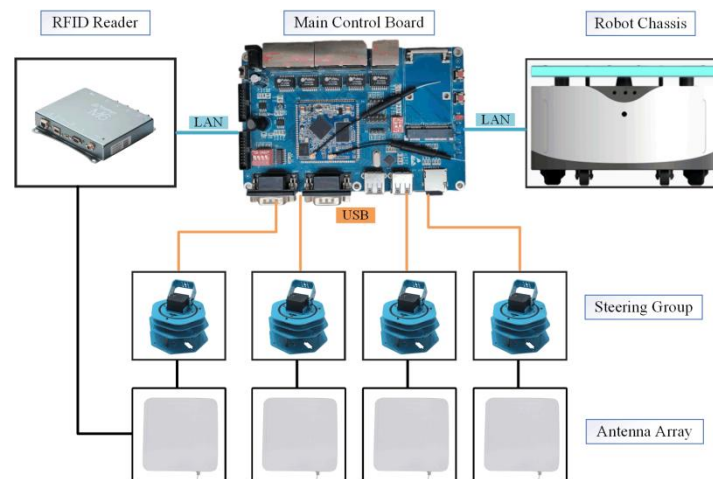

Figure 1. Hardware topology of RFID robot systems.

Figure 2. The physical picture of RFID robot.

In this paper, the software architecture of the RFID robot systems is optimized, and an independent algorithm module and computing unit are deployed on the Raspberry Pi 4b based on the ARM architecture, which is mainly used for real-time analysis and perception of system state and running the adaptive algorithm of the RFID robot systems. The algorithm module is divided into a real-time state sensing module and an intelligent control module. The former is used for sensing the IS and operation state of the robot systems, which communicates with the intelligent control module through an interface call. The computational process required for the intelligent control module is sent to the computing module via the serial port.

The software architecture of the RFID robot systems is shown in Figure 3. The specific adaptive control flow is as follows: When the RFID robot is performing mobile inventory, the system parameters and tags information are fed back to the algorithm module in

real-time through the main control module and middleware of the local platform, and the real-time state perception unit in the algorithm module is used according to the received information. The real-time state evaluation model evaluates the IS, and then the evaluation results are fed back to the intelligent control unit, which uses the adaptive algorithm to calculate the adaptive strategy and sends the results to the middleware, and the main control board receives information of the middleware to realize the adaptive control of RFID robot systems.
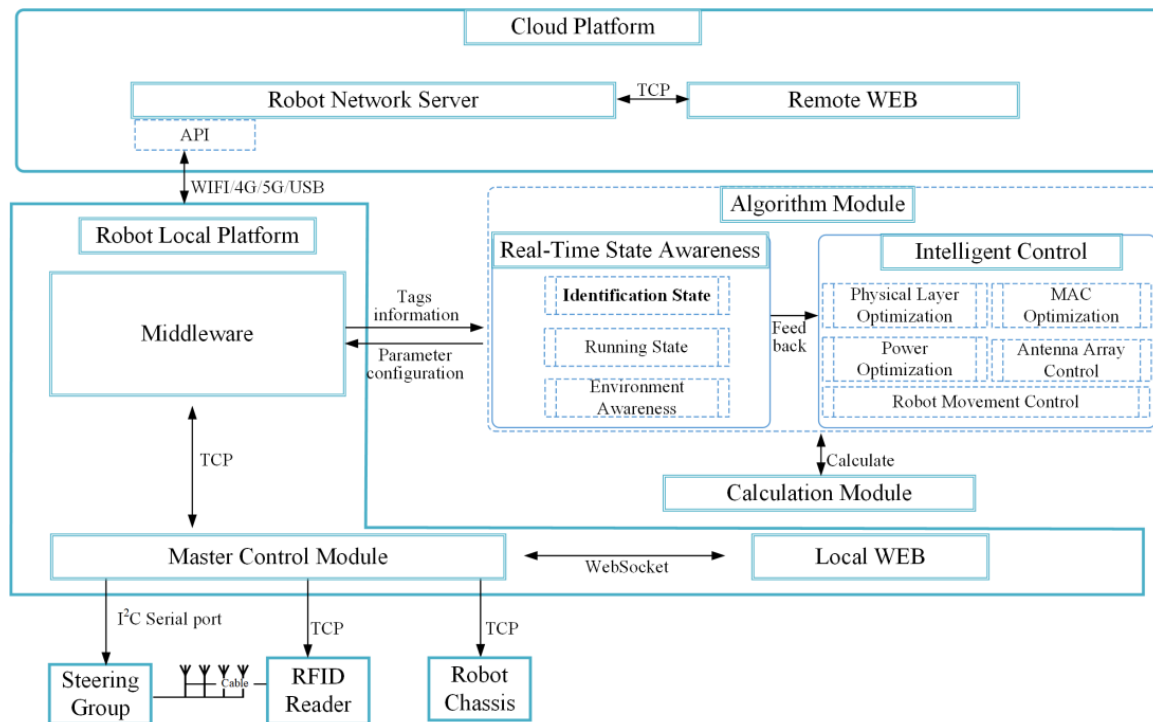

Figure 3. The software architecture of the RFID robot systems.

2.3. Real-Time Evaluation of the IS

Static RFID systems use fixed-position and fixed-parameter readers, and the IS can only be evaluated by RoI. In dynamic scenarios, the identification range and environment of RFID readers are constantly changing, so it is inaccurate to use RoI alone to evaluate the IS. In this paper, the RoI of the system is calculated based on the total number of tags identified by the RFID systems in real-time, and the theoretical value of the current SoI is calculated. The IS is evaluated in real-time using the difference between the theoretical and actual value of the SoI and the real-time RoI.

2.3.1. Real-Time RoI

Since the RFID robot performs mobile identification, the identification range of the reader is constantly changing, and the change of the identification range is shown in Figure 4. The identification probability of tags directly in front of the reader is high, while the RoI of tags on both sides is low. As a result, the reader's identification range is defined as the rectangular area formed by the two dashed lines of the same color in Figure 5, ignoring the tags on both sides [19].
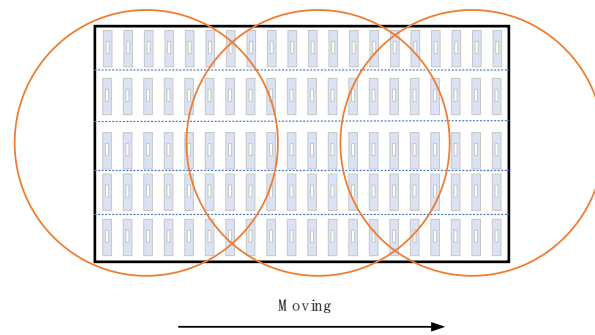
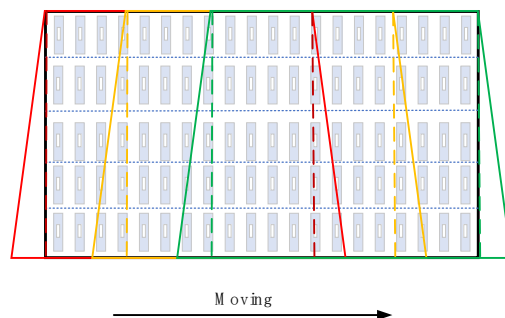**Figure 4.** The change in the identification range.

Figure 5. Optimization of identification range.

The dynamic identification scenario can be assumed as follows: m_{total} tags are evenly distributed on a bookshelf of length l, so the number of tags per unit length is $\frac{l}{m_{total}}$; the robot chassis moves at a constant speed of v m/s, so the number of tags entering the reader's identification range per second is $\frac{l}{m_{total}} \cdot v$, which represents the number of tags that should be successfully identified theoretically; then the number of tags that should be successfully identified m_t in the t second is:

$$m_t = \frac{l}{m_{total}} \cdot v \cdot t \tag{1}$$

The actual number of tags that have successfully identified can be obtained by RFID systems is $m_{success}$, so the RoI at the t second is:

$$\text{RoI} = \frac{m_{success}}{m_t} = \frac{m_{success} \cdot m_{total}}{l \cdot v \cdot t} \tag{2}$$

2.3.2. Real-Time Theoretical Value of the SoI

In the ISO/IEC 18000-6C protocol, the reader sends query/queryAdjust/queryRepeat commands to identify tags. In the process of mobile identification, the total number m of these commands in a second can be obtained, and each command has $L_i(i = 1, 2, 3 \ldots m)$ slots. Since the tags are evenly distributed, at the second t, the number of tags entering the reader's identification range is the same as the number of tags leaving the reader's identification range, so the number of tags per second within the reader's range is constant n. It is known that the system throughput is highest when the number of slots in a frame is the same as the number of tags. So, assuming the number of slots per command $L_i = L = n$, when there are n tags in L slots, r tags in the same slot obey the binomial distribution [20], that is:

$$B(r) = C_n^r \left(\frac{1}{L}\right)^r \left(1 - \frac{1}{L}\right)^{n-r} \tag{3}$$

Then, the expectation of successful slots in a frame is:

$$S = L * B(1) = L * C_n^1 \left(\frac{1}{L}\right)^1 \left(1 - \frac{1}{L}\right)^{n-1} = n * \left(1 - \frac{1}{L}\right)^{n-1} \tag{4}$$

The number of successful slots in each frame is $n * \left(1 - \frac{1}{n}\right)^{n-1}$, so the theoretical value of the speed of identification (TSoI) is equal to the number of successful slots per unit time, that is:

$$\text{TSoI} = m * n * \left(1 - \frac{1}{n}\right)^{n-1} \tag{5}$$

2.3.3. Classification of the IS

The purpose of mobile identification is to identify tags as quickly as possible while avoiding the missed reading of tags and maintaining a high RoI. Therefore, the RoI represents the pros and cons of the IS of RFID systems. The SoI measures the number of successfully identified tags per unit time. The more tags that are identified, the better the current IS. To determine the IS of RFID systems, the SoI and the difference between TSoI and SoI are key parameters. In this paper, the IS is divided into three classes, as shown in Figure 6.
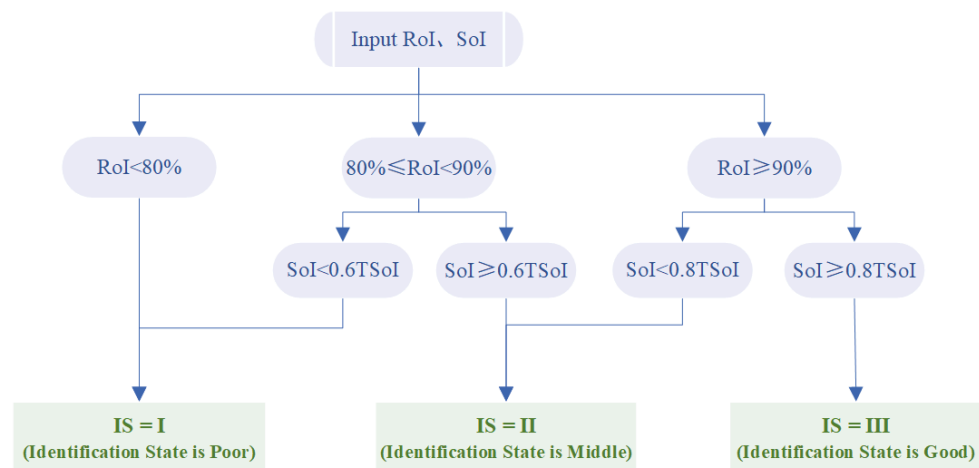

Figure 6. Classification of the IS.

3. Real-Time Evaluation Model and Theoretical Algorithm of the IS

3.1. Evaluation Model

The real-time evaluation model of the IS proposed in this paper is shown in Figure 7. In the mobile scenarios, sample data is normalized to eliminate the influence of different dimensions. The PCA reduces the complexity of the data by selecting the important influence parameters of RFID systems. Using a 3-Class KNN model and cross-validation to optimize parameters, an evaluation model of the IS of RFID systems for dynamic scenarios is constructed.
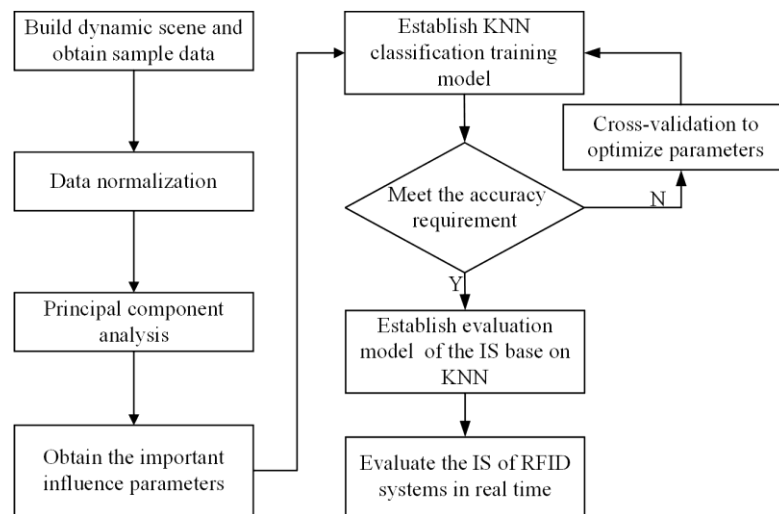
Figure 7. The evaluation model of the IS.

3.2. Theory of Parameter Selection Based on PCA

PCA is an unsupervised learning method, which uses orthogonal transformation to convert the observation data represented by linearly related into a few data represented by linearly independent variables. The linearly independent variables are called principal components [21–23].

The data obtained from the RFID systems in the dynamic scenarios, due to the parameters of the RFID systems being in different dimensions, directly seeking the principal components will produce unreasonable results, so the parameters need to be normalized (mean value 0 and variance 1). The steps to obtain the important influence parameters of the RFID systems by using the eigenvalue decomposition covariance matrix are as follows:

1. Normalize the $m \times n$ dimensional random variables representing the influence parameters of the IS of RFID systems to obtain a normalized data matrix X, and calculate the sample correlation matrix R.

$$R = \frac{1}{n-1} X X^{\mathrm{T}} \tag{6}$$

2. Calculate the k principal components according to the k eigenvalues of the sample matrix R and the corresponding unit eigenvectors $\alpha_i = (\alpha_{1i}, \alpha_{2i}, \cdots, \alpha_{mi})^{\mathrm{T}}, i = 1, 2, \cdots, k$.

$$y_i = a_i^{\mathrm{T}} x, i = 1, 2, \cdots, k \tag{7}$$

3. Calculate the correlation coefficient $\rho(x_i, y_j)$ of the k principal components y_j and the original variable x_i, and the contribution rate v_i of the k principal components to the original variable x_i.

4. Substitute the normalized data into (7) to obtain k principal component values of n samples. The *i-th* principal component value of the *j-th* variable $x_j = (x_{1j}, x_{2j}, \cdots, x_{mj})^{\mathrm{T}}$ is:

$$y_{ij} = (a_{1i}, a_{2i}, \cdots, a_{mi})(x_{1j}, x_{2j}, \cdots, x_{mj})^{\mathrm{T}} = \sum_{l=1}^{m} a_{li} x_{lj} \tag{8}$$

This is according to the correlation between each principal component and the influence parameters of the IS of RFID systems to obtain the important influence parameters.

3.3. The Classification of IS Based on KNN

KNN is a data mining classification algorithm, which belongs to supervised learning methods. The distance between the unknown data and the data points in the training set of the known category is calculated through all the features of the data, and the calculated distance represents the similarity between the features of the unknown data and the features of each data in the training set. The smaller the distance, the greater the similarity, and the greater the probability that the unknown data will become the corresponding category of the data [24–26]. After the calculation, the top K data with the smallest distance are selected. Among the K data, the number of occurrences of each data is recorded, and the category corresponding to the data with the most occurrences is the category of the unknown data.

There are three elements of the KNN: distance metric, K, and classification decision rule. The commonly used distance metric is Euclidean distance, the K value is determined according to cross-validation, and the classification decision adopts majority voting. When the training set and the above parameters are determined, the classification result is uniquely determined. The steps of the KNN algorithm are as follows:

The input to the KNN algorithm:

The training set $T = \{(x_1, y_1), (x_2, y_2), \cdots, (x_N, y_N)\}$; Among them, $x_i \in \chi \subseteq R^n$ is the important influence parameters of the RFID systems, and $y_i \in \gamma = \{c_1, c_2, \cdots, c_K\}$ is the classification of the IS of the RFID systems, $i = 1, 2, \cdots, N$.

1. Calculate the distance between two sample points x_i and x_j according to the distance metric.

$$L_2(x_i, x_j) = \left(\sum_{l=1}^{n} \left| x_i^{(l)} - x_j^{(l)} \right|^2 \right)^{\frac{1}{2}} \tag{9}$$

2. Find the k points closest to x in the training set T, and the neighborhood of x covering these k points is denoted as $N_k(x)$.
3. Determine the category y of x according to the classification decision rule in $N_k(x)$:

$$y = \arg\max_{c_j} \sum_{x_i \in N_k(x)} I(y_i = c_j), i = 1, 2, \cdots, N; j = 1, 2, \cdots, K \tag{10}$$

4. I is an indicator function, that is, I is 1 when $y_i = c_j$, otherwise I is 0.

The output to KNN: The class y to which x belongs.

4. Calculation and Analysis of Real-Time Evaluation of the IS of RFID Robots

4.1. Experimental Scene and Method

The experimental scene was selected in an open room, and 100 file boxes with UHF passive RFID tags are evenly placed on the bookshelf with a length of 2.5 m and a height of 1.8 m. Figure 8 shows the experimental scene, and Table 1 shows the experimental devices.

Table 1. Experimental devices.

Device	Parameter
RFID tags	900 MHz UHF Passive tags
Reader	ThingMagic-Mercury6
Antenna	9 dBi circular polarization
robot	Water robot

Figure 8. Experimental scene.

The experimental process is as follows: the robot equipped with the reader moves at a constant speed and performs tag identification through the bookshelf. During the moving process, the orthogonal combination test is carried out using the parameters of RFID systems in Table 2, and the average value is obtained after each group of parameters is tested 50 times.

Table 2. The parameters of RFID systems.

Names of Parameters	Value	Unit of Parameters
Reader Power	5–30	dBm
Robot Speed	0.3–1.0	m/s
BLF	250, 640	KHz
Tari	6.25, 12.5, 25 μs	μs
Q	5, 10, 15, Dynamic Q	/
Encoding of Tags	FM0, M2, M4, M8	/

4.2. The Selection of Important Influence Parameters Based on PCA

The parameters of RFID systems were tested by the orthogonal combination, and 8320 groups of data were obtained after eliminating abnormal data. Each group is composed of the influence parameters and the IS. In order to prevent information overlap and redundancy between parameters, this paper uses PCA to eliminate redundancy for the influence parameters [27]. Figure 9 shows the cumulative variance of the principal components, which shows that the first four principal components can represent more than 90% of the variance of the entire data, that is, the first four principal components can represent most of the information in the data. Figure 10 shows the cumulative sum of the correlations between all the influence parameters and the first four principal components. Due to the low correlation between the Tari, the encoding of tags, and principal components, these two parameters have little effect on the IS, so they are ignored in the subsequent data processing.
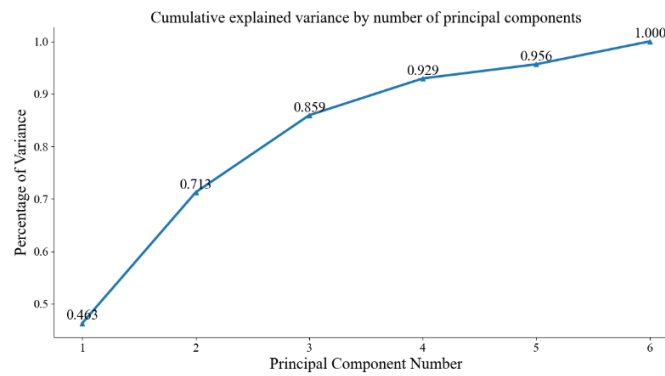
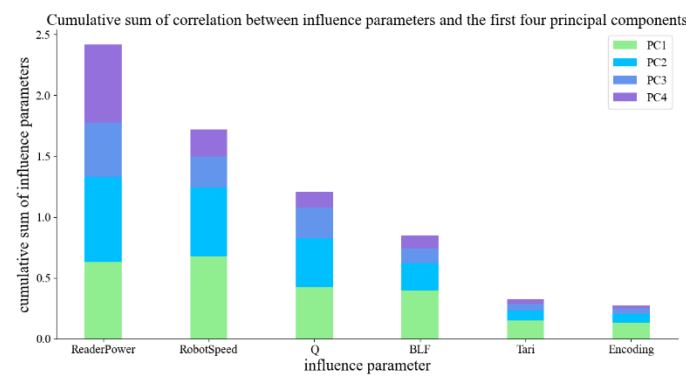**Figure 9.** Cumulative explained variance by number of principal components.

Figure 10. Cumulative sum of the correlations between influence parameters and the first four principal components.

4.3. Real-Time Evaluation of IS based on PCA-KNN

The schematic diagram of real-time evaluation modeling is shown in Figure 11.
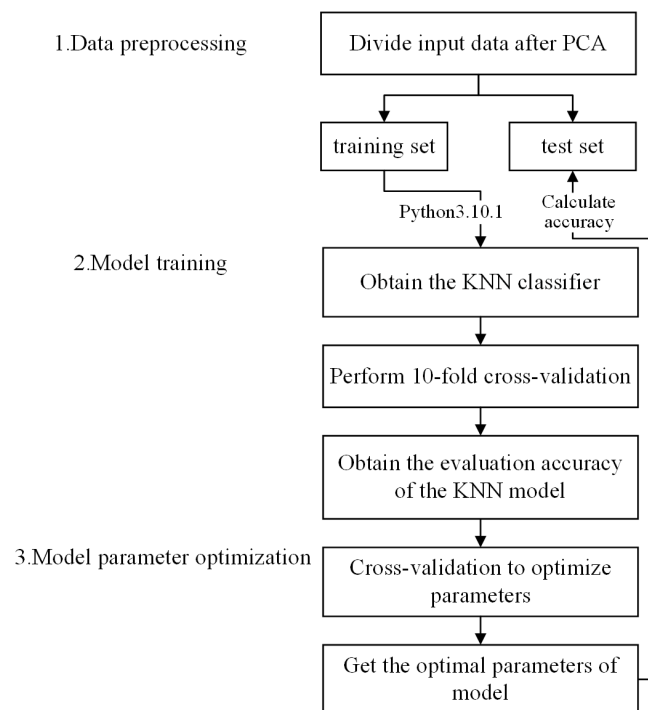

Figure 11. The schematic diagram of real-time evaluation modeling.

The specific steps of real-time evaluation modeling are as follows:

1. Data preprocessing

The set of influence parameters obtained by PCA:
{Reader Power, Robot Speed, Q, BLF}
The 8320 × 4 groups of experimental data composed as the input data of the KNN algorithm model, and the input data is divided into training set and test set.

2. Model training

Model training and parameter optimization use Python (Guido van Rossum, 1990, Amsterdam, Netherlands) with software version 3.10.1 and use the KNN classification algorithm in the sklearn.neighbors package. The basic steps of the algorithm are as follows:
Step 1 Enter the experimental data;
Step 2 Obtain the classifier using the function KNeighborsClassifier() in the package sklearn.neighbors;
Step 3 Use the function cross_val_score() to perform 10-fold cross-validation on the training set and test set [28];
Step 4 Obtain the evaluation accuracy of the KNN model.

3. Model parameter optimization

In the KNN algorithm, K represents the tradeoff between approximation error and estimation error [29], and distance weight must also be chosen carefully when building the model. In this paper, the cross-validation method is used to optimize the parameters of the KNN model. The value range of K is [1,14]; the distance weight can choose *uniform* and *distance*, *uniform* means that the distance weight is not considered, and *distance* means that the weight and distance are inversely relationship. Bring these two parameters into the above algorithm to obtain the optimal parameter combination. Figure 12 shows the cross-validation parameter optimization diagram of the KNN algorithm. The optimal parameter combination is K = 11, regardless of the distance weight, and classification accuracy rate of the training set is 92.2%. Figure 12 shows the optimization diagram of the cross-validation parameters of the KNN algorithm.
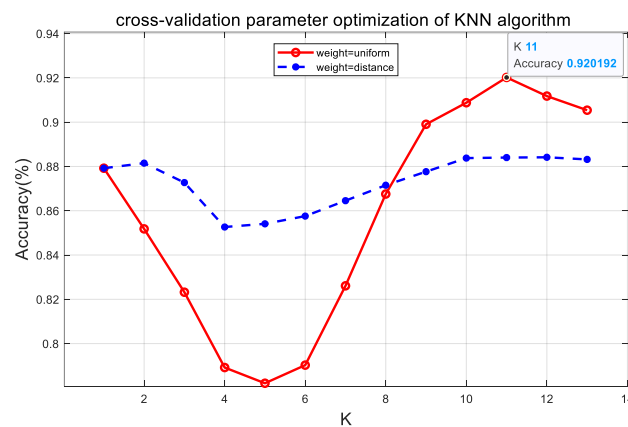

Figure 12. Cross-validation parameter optimization of KNN algorithm.

4.4. Classification Result and Analysis

In Section 4.3, we obtain the optimal parameter combination of the KNN algorithm as K = 11, regardless of the distance weight. Using the optimal parameter combination to classify the test set, the final classification accuracy is shown in Table 3. The overall accuracy of the test set classification is 92.4%. The classification accuracy of class I and class III is higher, and the classification accuracy of class II is lower. The sample data of class II in the test set is less, and in the middle of the class I and class III when the algorithm is classified according to the distance. It may be closer to the other two classes, resulting

in classification errors. Figure 13 shows the actual distribution of the IS, in which the red marks are the misclassification samples. It can be seen from the figure that classification accuracy of class II is low. The overall classification accuracy of the algorithm is higher than class II and class III.

Table 3. The classification accuracy of test set.

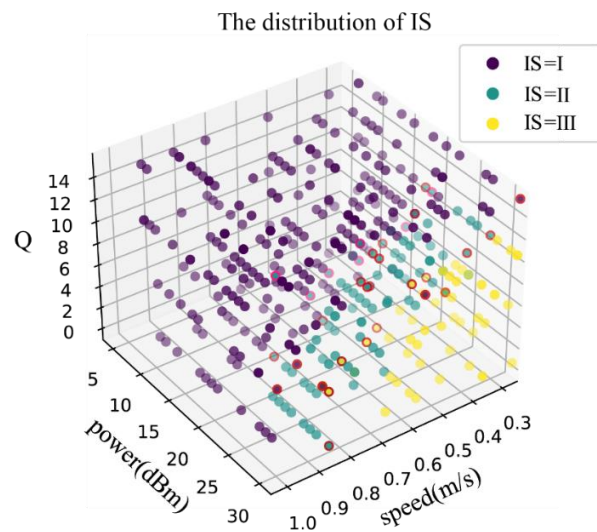| Classification of IS	Accuracy of Test Set
Class I	93.6%
Class II	84.7%
Class III	89.1%
overall	92.4%

Figure 13. The actual distribution of the IS.

4.5. Compare with Other Algorithms

The random forest, support vector machine, and decision tree are selected to compare with the evaluation algorithm of the IS for RFID systems based on PCA-KNN proposed in this paper. If the classification accuracy is higher, the algorithm running time is shorter, which proves that the model performance is better. The above algorithms use Python3.10.1 to optimize the parameters, the random forest takes n_estimators = 150, the support vector machine is set to C = 2.643, g = 0.167, the decision tree uses the CART algorithm, and the Gini coefficient is used as the feature selection criterion.

Using the same data trained in the above algorithm models, 1000 groups of test data were used for classification. Table 4 shows the classification accuracy results. The algorithm processing of a single data was performed on a Raspberry Pi with a main frequency of 1.5G Hz, a 4-core CPU, and a memory of 2 GB. The comparison result of the running time is shown in Table 5. The result shows that the evaluation method of IS based on PCA-KNN proposed in this paper has a shorter running time and a higher classification accuracy.

Table 4. Comparison of classification accuracy of different algorithms.

Algorithm	Number of Correct Classifications	Number of Error Classifications	Accuracy
Random Forest	917	83	91.7%
Support Vector Machine	904	96	90.4%
Decision Tree	871	129	87.1%
PCA-KNN	938	62	92.4%

Table 5. Comparison of running time of different algorithms.

Algorithm	Running Time (ms)
Random Forest	1.225
Support Vector Machine	0.376
Decision Tree	0.482
PCA-KNN	0.258

5. Adaptive Control for RFID Robot Systems Based on Q-Learning

Section 4 proposes a real-time evaluation method for the IS of RFID robot systems in dynamic scenarios. In the PCA-based analysis of important influence parameters, the reader power (P) and the robot speed (S) have the highest correlation with the IS. Therefore, this section combines the Q-learning to adjust the P and S in the RFID robot systems according to the real-time evaluation result of the IS, so as to improve the identification efficiency of the RFID robot systems.

5.1. The Model of Adaptive Control

Q-learning is a kind of reinforcement learning, which emphasizes exploring actions and learning based on the environment in order to maximize the expected benefit Q [30]. The model of adaptive control of parameters for RFID robot systems based on Q-learning is shown in Figure 14, where different parts represent different structures in Q-learning.
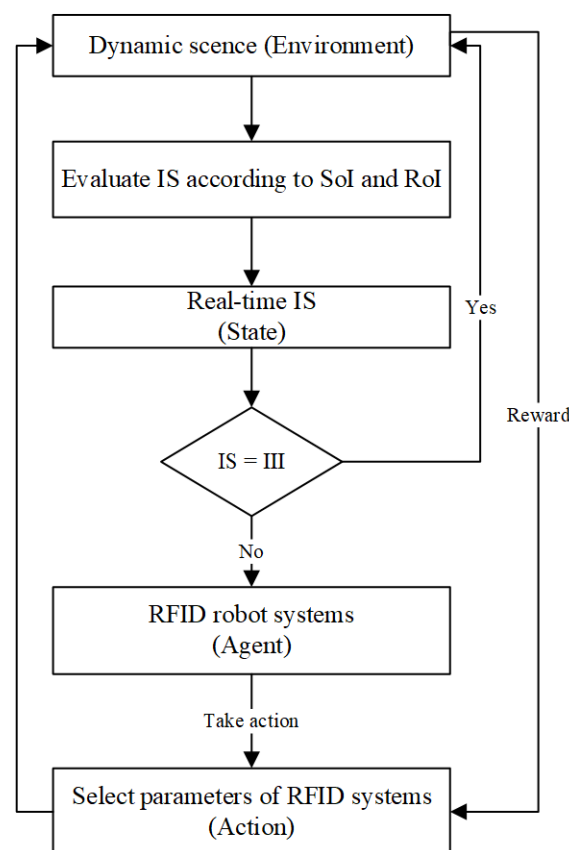

Figure 14. The model of adaptive control of parameters for RFID robot systems.

The actions Q-learning that can be taken are represented in action space δ, which contains six actions and are shown in Table 6. The states of Q-learning are shown by state space ζ: [I, II, III], which represent three classes of the IS.

Table 6. Action space δ.

Action	Explanation	Range
$P + 1$	P increased by 1 dBm	
$P + 2$	P increased by 2 dBm	$P_{\text{range}} = [1,30]$
$P - 1$	P decreased by 1 dBm	
$S - 0.1$	S decreased by 0.1 m/s	
$S - 0.2$	S decreased by 0.2 m/s	$S_{\text{range}} = [0,1.0]$
$S + 0.1$	S increased by 0.1 m/s	

The rewards obtained by taking different actions in different states are different. When the IS is poor, it is necessary to increase P or decrease S to ensure the reliability of the identification. When the IS is good, the P can be decreased, and the S can be increased to improve the identification efficiency. The reward matrix is R, where the rows and columns represent actions and states, respectively. R represents the reward value that can be obtained when an action is taken in a certain state.

$$R = \begin{pmatrix} 1 & 2 & -1 & 1 & 2 & -1 \\ 1 & 2 & -1 & 1 & 2 & -1 \\ -1 & -2 & 1 & -1 & -2 & 1 \end{pmatrix} \tag{11}$$

Q table is used to record the estimated Q value of different actions in different states. $Q(s, a)$ is the expectation that taking action $a(a \in \delta)$ can obtain reward under $s(s \in \zeta)$. When the agent explores the environment, it will use the Ballman equation to iteratively update $Q(s, a)$ until it converges or reaches the set number of iterations. The updated formula of Q-learning is as follows:

$$NewQ(s, a) = Q(s, a) + \alpha \left[R(s, a) + \gamma max Q'(s', a') - Q(s, a) \right] \tag{12}$$

α represents learning efficiency, $R(s, a)$ represents real-time reward, γ represents the decay of future reward, and $\gamma max Q'(s', a')$ represents future long-term reward.

The adaptive control algorithm of parameters for RFID robot systems proposed in this paper based on Q-learning is shown in Algorithm 1.

Algorithm 1 The adaptive control algorithm of parameters for RFID robot systems

---

Input: Real-Time IS, P, S, R, ζ, δ
Initialize $\alpha = 0.5$, $\gamma = 0.9$
repeat
Initialize s, $Q(s, a) = 0$
repeat
Choose a from δ
If $a \in P_{\text{range}}$ or S_{range} **then**
Take action a, observe $R(s, a)$, s'
    $Q(s, a) \leftarrow Q(s, a) + \alpha [R(s, a) + \gamma max Q'(s', a') - Q(s, a)]$
    $s \leftarrow s'$
End if
until s is terminal
Output: Q table, strategy $\pi(s) = \text{argmax}_{a \in \delta} Q(s, a)$

5.2. Results and Analysis

The final Q table is obtained by simulating the algorithm, and the parameters can be adaptively adjusted according to the Q table to improve the efficiency of the RFID robot systems in dynamic scenarios.

$$Q\ table = \begin{pmatrix} 8.18 & 10.00 & 0 & 8.18 & 10.00 & 0 \\ 8.20 & 10.00 & 0 & 8.20 & 10.00 & 0 \\ 0 & 0 & 8.99 & 0 & 0 & 8.99 \end{pmatrix} \qquad (13)$$

In order to verify the effectiveness of the proposed adaptive control algorithm of parameters, the experimental verification is carried out under the same experimental scene as in Section 4. The default parameters are set to $P = 23$, $S = 0.5$. Comparing the RoI with the default parameters and the adaptive control parameters under different tag densities, the results are shown in Figure 15. As the tag density increases, the RoI decreases, but the adaptive control algorithm is always better than the default parameters.
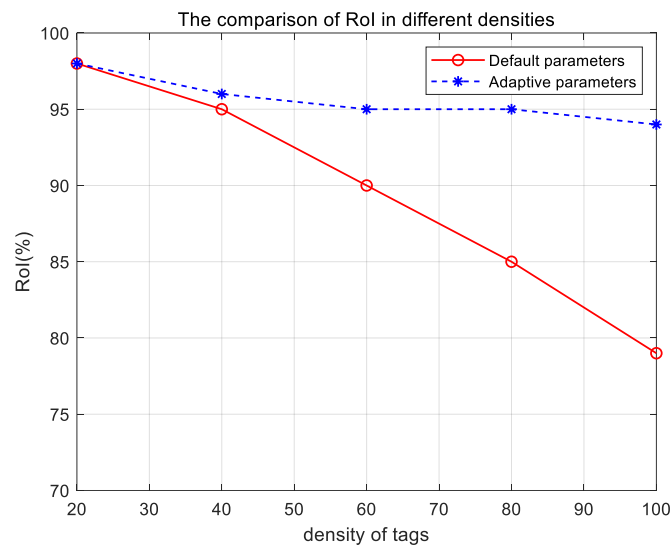

Figure 15. The RoI in different tag density.

A comparison of the power consumption and the reading time for all tags identified under the default parameters and adaptive parameters is presented in Figure 16. The power consumption under the adaptive parameters is reduced by 36.4%, and the time spent decreases by 29.7%. So, the proposed algorithm can improve the efficiency of the RFID robot systems in dynamic scenarios.
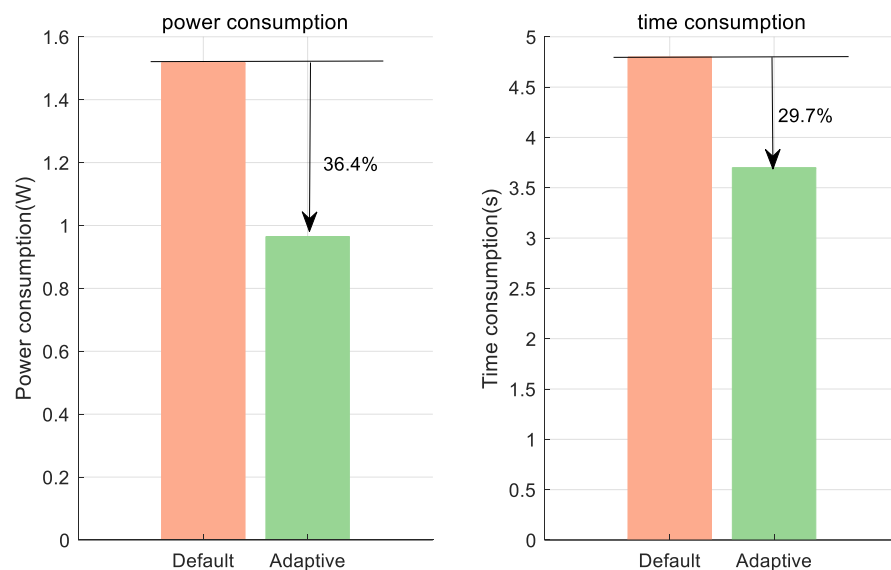

Figure 16. Power consumption and time consumption.

6. Conclusions

This paper firstly presents the important roles of the real-time evaluation of the IS and adaptive control of parameters in the RFID robot systems and proposes the main division method of the IS. Secondly, a real-time evaluation method of the IS of UHF passive RFID robot systems in dynamic scenarios based on PCA-KNN is proposed. PCA is used to select the important influence parameters of the IS, and a 3-Class KNN evaluation model of the IS is established based on the selected parameter set. Compared with other algorithms, the result shows that the accuracy of the evaluation method of the IS proposed in this paper is 92.4%, and the running time of a single data is 0.258 ms, which is better than other algorithms. Finally, this paper proposes a Q-learning-based adaptive control algorithm for RFID robot systems. This algorithm can dynamically control the reader's transmission power and the robot's moving speed. The results show that, compared with the default parameters in RFID robot systems, the algorithm effectively improves the identification rate of the system, the power consumption under the adaptive parameters is reduced by 36.4%, and the time spent decreases by 29.7%. Therefore, the adaptive control algorithm can be applied to RFID robot systems in dynamic scenarios to improve system efficiency.

Author Contributions: Conceptualization, H.W.; methodology, R.P. and Z.H.; software, R.Y.; validation, R.Y.; formal analysis, H.W. and P.P.; investigation, R.P. and N.Z.; writing—original draft preparation, R.Y. and J.Y.; writing—review and editing, H.W.; All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

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

References

1. Marín Díaz, G.; Carrasco, R.A.; Gómez, D. RFID: A Fuzzy Linguistic Model to Manage Customers from the Perspective of Their Interactions with the Contact Center. *Mathematics* **2021**, *9*, 2362. [CrossRef]
2. Guchhait, R.; Pareek, S.; Sarkar, B. How Does a Radio Frequency Identification Optimize the Profit in an Unreliable Supply Chain Management? *Mathematics* **2019**, *7*, 490. [CrossRef]
3. Wu, H.; Tao, B.; Gong, Z.; Yin, Z.; Ding, H. A Standalone RFID-Based Mobile Robot Navigation Method Using Single Passive Tag. *IEEE Trans. Autom. Sci. Eng.* **2020**, *18*, 1529–1537. [CrossRef]
4. Khan, Z.; Chen, X.; He, H.; Xu, J.; Wang, T.; Cheng, L.; Ukkonen, L.; Virkki, J. Glove-Integrated Passive UHF RFID Tagsabrication, Testing and Applications. *IEEE J. Radio Frequency Identif.* **2019**, *3*, 127–132. [CrossRef]
5. Zhao, N.; Zhang, L.; Lei, L.; Cai, S. Dynamic Query Tree Anti-Collision Protocol for RFID Systems. In Proceedings of the 2019 IEEE 25th International Conference on Parallel and Distributed Systems, Tianjin, China, 4–6 December 2019.
6. Montanaro, T.; Sergi, I.; Motroni, A.; Buffi, A.; Nepa, P.; Pirozzi, M.; Catarinucci, L.; Colella, R.; Chietera, F.P.; Patrono, L. An IoT-Aware Smart System Exploiting the Electromagnetic Behavior of UHF-RFID Tags to Improve Worker Safety in Outdoor Environments. *Electronics* **2022**, *11*, 717. [CrossRef]
7. Keonn. Available online: https://www.keonn.com/systems/view-all-2/inventory-robotss.html (accessed on 15 June 2022).
8. Bernardini, F.; Motroni, A.; Nepa, P.; Tripicchio, P.; Buffi, A.; Del Col, L. The MONITOR Project: RFID-based Robots enabling real-time inventory and localization in warehouses and retail areas. In Proceedings of the 2021 6th International Conference on Smart and Sustainable Technologies, Bol and Split, Croatia, 8–11 September 2021; pp. 1–6.
9. Giannelos, E.; Andrianakis, E.; Skyvalakis, K.; Dimitriou, A.G.; Bletsas, A. Robust RFID Localization in Multipath with Phase-Based Particle Filtering and a Mobile Robot. *IEEE J. Radio Freq. Id.* **2021**, *5*, 302–310. [CrossRef]
10. DiGiampaolo, E.; Martinelli, F. A Robotic System for Localization of Passive UHF-RFID Tagged Objects on Shelves. *IEEE Sens J.* **2018**, *18*, 8558–8568. [CrossRef]

11. Madanian, S.; Parry, D. Identifying the Potential of RFID in Disaster Healthcare: An International Delphi Study. *Electronics* **2021**, *10*, 2621. [CrossRef]

12. Kantawong, S. Development of RFID dressing robot using DC servo motor with fuzzy-PID control system. In Proceedings of the 13th International Symposium on Communications and Information Technologies (ISCIT), Surat Thani, Thailand, 4–6 September 2013; pp. 14–19.

13. Gareis, M.; Hehn, M.; Stief, P.; Körner, G.; Birkenhauer, C.; Trabert, J.; Mehner, T.; Vossiek, M.; Carlowitz, C. Novel UHF-RFID Listener Hardware Architecture and System Concept for a Mobile Robot Based MIMO SAR RFID Localization. *IEEE Access* **2021**, *9*, 497–510. [CrossRef]

14. Kammel, C.; Kögel, T.; Gareis, M.; Vossiek, M. A Cost-Efficient Hybrid UHF RFID and Odometry-Based Mobile Robot Self-Localization Technique with Centimeter Precision. *IEEE J. Radio Freq. Id.* **2022**, *6*, 467–480. [CrossRef]

15. Wu, C.; Tao, B.; Wu, H.; Gong, Z.; Yin, Z. A UHF RFID-Based Dynamic Object Following Method for a Mobile Robot Using Phase Difference Information. *IEEE Trans. Instrum. Meas.* **2021**, *70*, 8003611. [CrossRef]

16. Motroni, A.; Bernardini, F.; Buffi, A.; Nepa, P.; Tellini, B. A UHF-RFID Multi-Antenna Sensor Fusion Enables Item and Robot Localization. *IEEE J. Radio Freq. Id.* **2022**, *6*, 456–466. [CrossRef]

17. Imran, M.A.; Sharif, A.; Yan, Y.; Ouyang, J.; Tariq Chattha, H.; Arshad, K.; Assaleh, K.; Alhumaidi Alotabi, A.; Althobaiti, T.; Ramzan, N.; et al. Uniform Magnetic Field Characteristics Based UHF RFID Tag for Internet of Things Applications. *Electronics* **2021**, *10*, 1603.

18. Wang, H.; Yu, R.; Pan, R.; Liu, M.; Yang, J. Fast tag identification for mobile RFID robots in manufacturing environments. *Assem. Autom.* **2021**, *3*, 292–301. [CrossRef]

19. Xie, L.; Li, Q.; Wang, C.; Chen, X.; Lu, S. Exploring the Gap between Ideal and Reality: An Experimental Study on Continuous Scanning with Mobile Reader in RFID Systems. *IEEE Trans. Mob. Comput.* **2015**, *14*, 2272–2285. [CrossRef]

20. Chen, W.-T. Optimal Frame Length Analysis and an Efficient Anti-Collision Algorithm with Early Adjustment of Frame Length for RFID Systems. *IEEE Trans. Veh. Technol.* **2016**, *65*, 3342–3348. [CrossRef]

21. Hadri, A.; Chougdali, K.; Touahni, R. Intrusion detection system using PCA and Fuzzy PCA techniques. In Proceedings of the 2016 International Conference on Advanced Communication Systems and Information Security, Marrakesh, Morocco, 17–19 October 2016.

22. Xia, Z.; Chen, Y.; Xu, C. Multiview PCA: A Methodology of Feature Extraction and Dimension Reduction for High-Order Data. *IEEE Trans Cybern.* **2021**, *52*, 11068–11080. [CrossRef] [PubMed]

23. Wang, L. Research on Distributed Parallel Dimensionality Reduction Algorithm Based on PCA Algorithm. In Proceedings of the 2019 IEEE 3rd Information Technology, Networking, Electronic and Automation Control Conference, Chengdu, China, 15–17 March 2019; pp. 1363–1367.

24. Bajpai, D.; He, L. Evaluating KNN Performance on WESAD Dataset. In Proceedings of the 2020 12th International Conference on Computational Intelligence and Communication Networks, Bhimtal, India, 25–26 September 2020; pp. 60–62.

25. Chethana, C. Prediction of Heart Disease using Different KNN Classifier. In Proceedings of the 2021 5th International Conference on Intelligent Computing and Control Systems, Madurai, India, 6–8 May 2021; pp. 1186–1194.

26. Huang, J.; Wei, Y.; Yi, J.; Liu, M. An Improved KNN Based on Class Contribution and Feature Weighting. In Proceedings of the 2018 10th International Conference on Measuring Technology and Mechatronics Automation, Changsha, China, 10–11 February 2018; pp. 313–316.

27. Luo, Y.; Wang, B. Prediction of Negative Conversion Days of Childhood Nephrotic Syndrome Based on PCA and BP-AdaBoost Neural Network. *IEEE Access* **2019**, *7*, 151579–151586. [CrossRef]

28. Pal, K.; Patel, B.V. Data Classification with k-fold Cross Validation and Holdout Accuracy Estimation Methods with 5 Different Machine Learning Techniques. In Proceedings of the 2020 Fourth International Conference on Computing Methodologies and Communication, Erode, India, 11–13 March 2020; pp. 83–87.

29. Li, H. *Statistical Learning Methods*; Tsinghua University Press: Beijing, China, 2019; pp. 52–53.

30. Ye, W.; Chen, D. Analysis of Performance Measure in Q Learning with UCB Exploration. *Mathematics* **2022**, *10*, 575. [CrossRef]