

## Article

# Modelling the Trust Value for Human Agents Based on Real-Time Human States in Human-Autonomous Teaming Systems

Chin-Teng Lin <sup>1,2,\*</sup> , Hsiu-Yu Fan <sup>2</sup>, Yu-Cheng Chang <sup>1</sup> , Liang Ou <sup>1</sup>, Jia Liu <sup>1</sup> , Yu-Kai Wang <sup>1</sup>   
and Tzyy-Ping Jung <sup>3</sup> 

<sup>1</sup> CIBCI Lab, Australian Artificial Intelligence Institute, University of Technology Sydney, Ultimo, NSW 2007, Australia

<sup>2</sup> Institute of Imaging and Biomedical Photonics, National Yang Ming Chiao Tung University, Hsinchu City 30010, Taiwan

<sup>3</sup> Institute of Engineering in Medicine and Institute for Neural Computation, University of California San Diego, La Jolla, CA 92093, USA

\* Correspondence: chin-teng.lin@uts.edu.au; Tel.: +61-2-95142000

**Abstract:** The modelling of trust values on agents is broadly considered fundamental for decision-making in human-autonomous teaming (HAT) systems. Compared to the evaluation of trust values for robotic agents, estimating human trust is more challenging due to trust miscalibration issues, including undertrust and overtrust problems. From a subjective perception, human trust could be altered along with dynamic human cognitive states, which makes trust values hard to calibrate properly. Thus, in an attempt to capture the dynamics of human trust, the present study evaluated the dynamic nature of trust for human agents through real-time multievidence measures, including human states of attention, stress and perception abilities. The proposed multievidence human trust model applied an adaptive fusion method based on fuzzy reinforcement learning to fuse multievidence from eye trackers, heart rate monitors and human awareness. In addition, fuzzy reinforcement learning was applied to generate rewards via a fuzzy logic inference process that has tolerance for uncertainty in human physiological signals. The results of robot simulation suggest that the proposed trust model can generate reliable human trust values based on real-time cognitive states in the process of ongoing tasks. Moreover, the human-autonomous team with the proposed trust model improved the system efficiency by over 50% compared to the team with only autonomous agents. These results may demonstrate that the proposed model could provide insight into the real-time adaptation of HAT systems based on human states and, thus, might help develop new ways to enhance future HAT systems better.

**Keywords:** trust modelling; information fusion; human-autonomous teaming



**Citation:** Lin, C.-T.; Fan, H.-Y.; Chang, Y.-C.; Ou, L.; Liu, J.; Wang, Y.-K.; Jung, T.P. Modelling the Trust Value for Human Agents Based on Real-Time Human States in Human-Autonomous Teaming Systems. *Technologies* **2022**, *10*, 115. <https://doi.org/10.3390/technologies10060115>

Academic Editor: Mohammed Mahmoud

Received: 16 September 2022

Accepted: 30 October 2022

Published: 8 November 2022

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



**Copyright:** © 2022 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/>).

## 1. Introduction

The emerging cooperation of artificial intelligence and advanced automation systems provides an opportunity to ease the requirements of human labor and minimise risk in various tasks. In many instances, human and autonomous agents are coupled in a human-autonomous teaming (HAT) system to address complex problems where the tasks could be either unreachable or dangerous for humans or not suitable for autonomous agents with conventional automation [1–5]. Such problems often contain a series of factors that can easily cause mistakes and result in a high cost, including, but not limited to, navigation, patrolling, medical health insurance, rescue and scientific research [6–9].

As a critical factor in coordinating agents or allocating tasks, the evaluation of trust values for human agents becomes an essential issue in the cooperation of human and autonomous agents [10]. Previous studies proposed trust-based approaches to explore

either human or teammate trust for the optimization of interactions among agents in specified tasks [5,11–20]. The trust in autonomous agents can be well-modelled based on their previous experience, states, and actions, where humans can judge the trustworthiness of autonomous agents by observing their actions, e.g., whether they can act as expected. Additionally, measuring human cognitive states may benefit the identification of under what circumstances and contexts autonomous agents' performance can be higher or lower than expected [21,22]. However, it is challenging to fairly measure an individual's states, as cognitive states, such as mental stress and attention, are easily affected by human behaviours, which could cause human cognitive states to change from time to time [5,23]. Therefore, in an attempt to properly evaluate the trustworthiness of human agents, this study captured the human state in real-time and investigated the distributed human trust dynamics in an HAT system. We considered human trust to be affected by human psychological state and situational awareness as factors that indicated individuals' bias when making decisions during human-autonomous interactions. This definition aligns with the concept proposed by Guo et al. [24] and Azevedo-Sa et al. [25].

In this study, we introduced a fusion mechanism in the proposed trust model to estimate human trust values by fusing multiple pieces of information from human agents. To obtain an adaptive fusion mechanism for the human trust model, we leveraged a reinforcement learning (RL) algorithm to learn fusion weights from an external reward via a simulation-based training process. One advantage of using RL is that it can learn without prior knowledge, which avoids bias based on forepassed data [26,27]. However, a mathematical equation to describe reward values is still difficult to define for a system with multiple sources with RL. Moreover, uncertainty and noise are additional issues, as external or ineffective information could confuse rewards with reinforcement learning. To overcome these issues, we applied the fuzzy inference system (FIS) in our model. FIS is well known as an effective method for generating rewards for complex scenarios and deal with uncertainty from the environment. Several recent studies present the implementation of FIS-based reward structures for different complex scenarios [28–30]. Evidence shows that with the aid of its member functions and If-Then-Rule structure, FIS has an inherent capability to overcome uncertainty and noise from the environment [23,31–33]. Therefore, we used FIS in this study to generate rewards for the proposed trust model to overcome the above issues.

To verify the effectiveness of the proposed trust model, we use a robot simulator to design a ball collection task scenario that includes an HAT team working together. The HAT team involves a human agent who has to cooperate with one or two robot agents to collect balls with collision-free movements while performing the task. The human agent's sight is restricted; the environment can only be observed through a fixed monitoring camera in the simulator. Robot agents can determine whether they follow human commands or not, based on the human trust values evaluated by the proposed trust model. We used a training scenario to learn the fusion method with the Q-learning algorithm and tested it in three test scenarios with different settings. We further compare the performance of the HAT, only human agents, and only robot agents. The comparison results demonstrate that the proposed trust model can improve coordination in the HAT teams with different human participants in all test scenarios, which also suggests that the proposed model can adapt to various levels of human performance and generate reliable trust values via the Q-learning algorithm.

The main contributions of this research are three-fold:

- This paper proposes a trust model to estimate human trust value in real-time. The proposed trust model was applied to a ball collection task with robot agents, which presents uses of the proposed trust model in the human-autonomous teaming framework.
- The proposed trust model considers multiple pieces of information from a human agent, e.g., attention level, stress index and situational awareness, by leveraging a fuzzy fusion model. In this research, the attention level and the stress index are

evaluated based on pupil response and heart rate variability, respectively; situational awareness is measured from the environment through visual perception.

- We further use a Q-learning algorithm with a fuzzy reward to adaptively learn the fusion weight of the fusion model. The fuzzy reward is generated by a TSK-type fuzzy inference system, which facilitates the defending reward for complex scenarios and is able to handle the uncertainty of human information.

This paper is organised as follows. Section 2 introduces related work on human trust modelling. Section 3 describes the proposed Multi-Human-Evidence Based Trust Evaluation Model. This section first describes the details of trust evaluation metrics, followed by the details of the Trust Metric Fusion Model and the Reinforcement Learning Algorithm. Section 4 presents the experimental methods, including scenario design, human agent setup and recording and experimental procedures. Section 5 presents the experimental results. Section 6 shows the discussion based on the experimental results. Finally, Section 6 presents conclusions.

## 2. Related Works

Several studies have modelled human trust by analysing human physiological signals and behaviours. Sadrfaridpour et al. proposed a mutual trust model between human and autonomous agents to coordinate collaboration [14]. The authors defined human performance based on muscle fatigue and the recovery dynamics of the human body when performing repetitive tasks, and the performance of autonomous agents was evaluated using the difference between human and autonomous agents' behaviours. Mahani, Jiang and Wang applied a Bayesian mechanism to predict human-to-autonomy trust based on human trust feedback to each individual robot and human intervention [16]. With the aid of a data-driven approach, Hu et al. proposed a human trust model that classified an individual's trust and untrust with electroencephalography (EEG) signals and galvanic skin response (GSR) data [15]. Similar work is also presented in [34]; the researchers exploited human cognitive states extracted from EEG data to model human workers' trust in Collaborative Construction Robots. In addition, the pupillary response is observed to be an effective index for human trust estimation. Lu and Sarter exploited eye tracking metrics to infer human trust in real time [17]. Alves et al. [18] consider kinesic courtesy cues from human to machine as an important factor in establishing human trust in HAT collaboration. Apart from pure human factors, some work considers human behaviour and machine performance when modelling the mutual trust between human users and machines. Inspired by human social behaviour, Jacovi et al. [19] proposed a formalisation method to model mutual trust between a human user and a machine. Furthermore, some researchers proposed computational trust models for HAT systems. In [11], the model of pupil dilation is extended as a computational trust model to facilitate the interaction between humans and robots. The computational model of human-robot mutual trust presented in [20] considers multiple pieces of information in physical human-robot collaboration, such as robot motion, robot safety, robot singularity, and human performance. Although all of the above studies provide valuable perspectives on human trust modelling, some trust models consider subjective feedback or the historical behaviour of humans, which is not reliable enough and may lead to bias in the evaluation of present status and condition. Other approaches that generate trust based on a single human cognitive state might also result in a miscalibration of trust in complex scenarios. Thus, this study attempted to remedy the lack of multievidence in current trust models by modelling information from multiple sources that can be optimised without prior knowledge and provides a comprehensive evaluation of human trust.

## 3. Multi-Human-Evidence-Based Trust Evaluation Model

This section introduces the proposed trust evaluation framework used to generate human trust values based on real-time human cognition signals. The objective of the proposed framework is to enable autonomous agents to be aware of the real-time human

states and, therefore, to make a decision of the proper action based on the current human conditions. By generating a single trust value without historical data, the proposed model could reduce the complexity of the cooperation task and surely eliminate the bias from previous behaviour. Figure 1 shows the structure of our model. The proposed model combines multiple human evidence to estimate a single human trust value. The evidence contains three human states, including attention level, stress index and human perception. The information fusion block is responsible for combining the three pieces of evidence with sorting and weight learning via fuzzy Q-learning. The final output of the framework is the human trust value. Note that the human trust value is produced in real time, although the learning of the fusion weights requires offline training. Details of the components of the framework are presented in the following subsections.

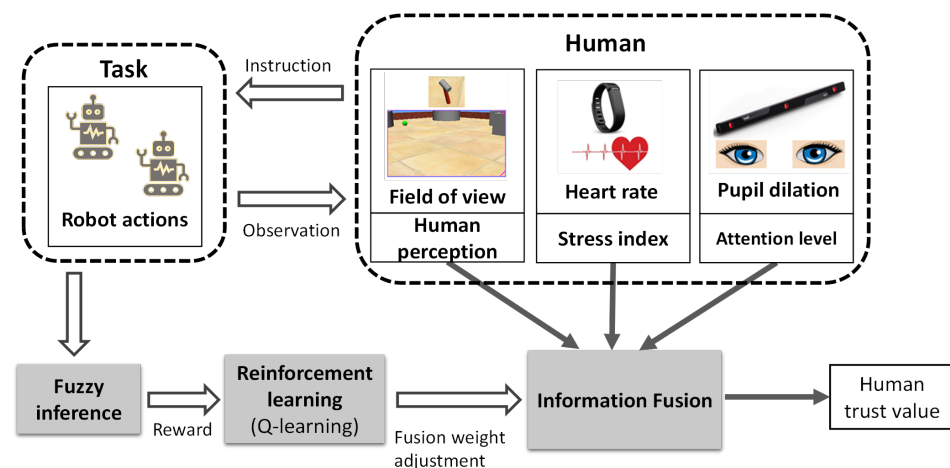


Figure 1. Structure of the proposed model.

### 3.1. Trust Evaluation Metrics

#### 3.1.1. Attention Level

As one of our three evaluation metrics for human performance, the attention level is calculated based on pupil response. Research evidence has shown that the dynamics of pupil response are an effective characteristic for estimating the human state of concentration or distraction [35]. The attention level is computed based on Equation (1) proposed by Hoeks and Levelt [35]:

$$y(t) = h(t) * x(t), \quad (1)$$

where  $y(t)$  is the pupillary response,  $h(t)$  is a system constant called the impulse response,  $x(t)$  is the attention level and  $*$  is the convolution operator. The variables  $y$ ,  $h$ , and  $x$  indicate the functions for the independent variable, time  $t$ .

The impulse response  $h(t)$ , derived from the approach introduced by Hoeks and Levelt [35], is set to represent the relation between attention and the pupillary response. The computing equation of impulse response  $h(t)$  is presented in (2).

$$h(t) = s \times (t^n) \times e^{\left(\frac{-n \times t}{t_{\max}}\right)}, \quad (2)$$

where  $n$  is the number of layers that is set to 10.1,  $t_{\max} = 5000$  ms is the maximum response time of participants,  $s = \frac{1}{10^{33}}$  is a constant used to scale the impulse response, and  $t$  is the response time, which are the same settings as in the cited equation [35].

#### 3.1.2. Stress Index

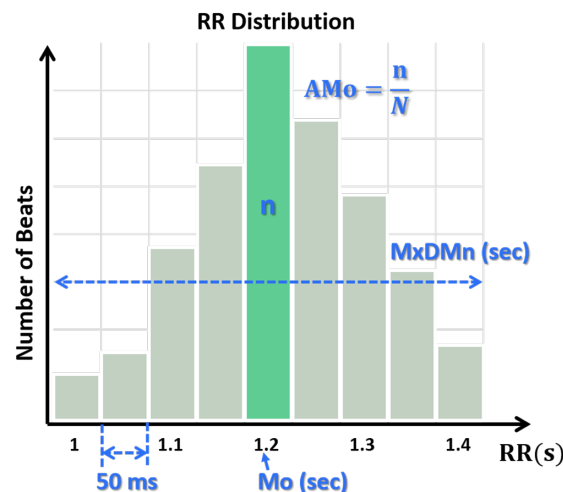
It has been well accepted that the stress level of individuals can affect the corresponding performance in a task. Thus, we also calculated human stress as one metric used for human trust evaluation. In this paper, we used heart rate variability (HRV) as the stress index of the participants. HRV is computed based on the measurement of the duration



from a series of continuous heart cycles, known as the interbeat interval (IBI), which is used to evaluate the human body's autonomic regulation. A normal heart rate is in the range of 60–70 beats/min, which is controlled by the parasympathetic nervous system. During a cognitive state with stress, human sympathetic nervous system activity increases, which affects the duration of IBI and heart rate. To quantitatively evaluate the stress level, we applied geometric methods to analyze the distribution and shapes of the IBI. The stress index (*SI*) was computed with the IBI data by means of Baevsky's equation in (3) [36].

$$SI = \frac{AMo}{2 \times Mo \times MxDMn}, \quad (3)$$

where *AMo* is the pattern amplitude expressed as a percentage, *Mo* is the mode that represents the most frequently occurring RR interval (the interval between successive heartbeats), and *MxDMn* is the variation range, which reflects the variability degree of the RR interval, as shown in Figure 2. The mode of *Mo* is simply taken as the median of the RR intervals, and *AMo* is the height of the histogram of the normalized RR interval (the width is 50 msec). *MxDMn* represents the difference between the shortest and longest RR intervals for each participant.



**Figure 2.** Histogram of the RR distribution for Baevsky's stress index, in which *n* is the number of beats, *N* is successive beat intervals, *AMo* is the height of the histogram of the normalised RR interval, *MxDMn* represents the difference between the shortest and longest RR intervals, and *Mo* is the median of RR intervals.

### 3.1.3. Human Perception

Human perception measures confidence in the decision-making of the human agent based on the situational awareness of the human in a HAT environment through visual perception. The simulation task applied in this study is a target (ball)-collection task. As shown in Table 1, based on human perceptions of the autonomous agent and target, four situations could occur during the task. The first situation is that the human can see both the autonomous agent and target; the second and third are that the human only observes either the target or the agent, respectively; and last, the human can see neither the autonomous agents nor the target. We use four factors, including the indexes to indicate position (*S1*), orientation (*S2*), distance (*S3*) and view angle (*S4*), to estimate human perception ability, where the human perception evaluation level equals  $f(S1) + f(S2) + f(S3) + f(S4)$ . More details of our experiment and indexes are elaborated in Section 4.

**Table 1.** Four situations of human perception.

	Human Perception
First situation	Agent + Target
Second situation	No Agent + Target
Third situation	Agent + No Target
Fourth situation	No Agent + No Target

### 3.2. Trust Metric Fusion Model

The fusion model combines three pieces of evidence from human states in real time to produce a trust value. The function is defined as  $F : [0, 1]^n \rightarrow [0, 1]$ , through which multiple values located in the interval  $[0, 1]$  are assigned to a single final value. We use the Hamacher product [37] to implement the fusion for the proposed trust model. The Hamacher product is a nonlinear transformation operation that uses confidence values from detectors to produce the final confidence for the fusion. If the evaluated single input values improve, the final trust value will increase such that  $F(0, 0, \dots, 0) = 0$  and  $F(1, 1, \dots, 1) = 1$ . When all single input values evaluated are zero, the final trust value falls to the minimum; in other words, the human is completely untrustworthy. However, if all the values of the evaluated input are 1, the upper limit of the trust value is 1. Assuming that the result of the fusion  $F(E)$  satisfies the constraint  $\min(E^1, E^2, \dots, E^n) \leq F(E) \leq \max(E^1, E^2, \dots, E^n)$ , we can define an aggregation function as follows:

$$F(E) = \sum_{i=1}^n f(E_i - E_{i-1}, w_i), \quad (4)$$

where  $E = (E_1, E_2, \dots, E_n) \in [0, 1]^n$  is an increasing permutation of evaluations such that  $0 \leq E_1 \leq E_2 \leq \dots \leq E_n$ ,  $w = [w_1, w_2, \dots, w_n]$  is the fusion weight vector, and  $w_1 + \dots + w_n = 1$ .

We use the Hamacher product to fuse each pair of evidence. The Hamacher  $t$ -norm involves the use of a fuzzy measure [35]. Therefore, the fusion model that produces the trust value based on the three pieces of evidence is defined as follows:

$$F_h(E) = \frac{g(E) \times w_i}{g(E) + w_i - g(E) \times w_i}, \quad (5)$$

where  $F_h(E)$  represents the human trust value, and  $g(E) = \sum_{i=1}^n (E_i - E_{i-1})$ . The fusion weights  $w = [w_1, w_2, w_3]$  are learnt by Q-learning based on collective human state data, including the pupil, HRV and human perception signals. In addition, we used the min-max normalisation for the pupil and HRV data to normalise the value in the range of  $[0, 1]$ .

#### 3.2.1. Reinforcement Learning

This section discusses the Q-learning method used to update the fusion weights. As a model-free, off-policy reinforcement learning method, Q-learning tracks what has been learnt and finds the best course of action for the agent to gain the greatest reward [38,39]. As discussed above, the final trust value is calculated by multiplying the evaluated values of three human states by the corresponding weights and then summing the results. Weights represent the relative importance of each individual evaluation, and the vector of weights is initialised randomly. Thus, since Q-learning is capable of transferring functions or reward functions with random factors [40], we applied its algorithm to determine which weight vector is used to fuse the estimated trust values from a random perspective. The equation to update the Q-value with action  $i$  and state  $s$  in each step is shown below:

$$Q(s, i) \leftarrow Q(s, i) + \alpha \times \left( w(s, i) \times \nabla + \gamma \times \sum_{j=1}^n (w(s', j) \times Q(s', j)) - Q(s, i) \right) \quad (6)$$

where  $\alpha$  is a fixed value used as the learning rate that satisfies the condition  $0 < \alpha \leq 1$ ,  $w(s, i)$  represents the value of the weight in state  $s$  and action  $i$ , the parameter  $\gamma$  is a temporal discount factor that satisfies the condition  $0 < \gamma \leq 1$ ,  $s'$  is the state after performing the action under state  $s$ , and  $\nabla$  is the reward. Here, we set  $\alpha$  to 0.1 and  $\gamma$  to 0.2. Additionally, we update the weight vector based on the Q-values after updating the Q-tables. Two conditions are used when applying the value of weight  $w(s, i)$ : 1) The summation of  $w(s, i)$  is normalized to 1. 2) Parameter  $\delta$  is within the range (0, 1].

Formula (7) shows the weight updating rule:

$$w'(s, i) \leftarrow w(s, i) + \begin{cases} (1 - w(s, i)) \times \delta \times \left( \frac{1}{1 + e^{-a \times Q(s, i) + b}} \right), & \text{if } i = \arg \max_j Q(s, j) \\ (0 - w(s, i)) \times \delta \times \left( \frac{1}{1 + e^{-a \times Q(s, i) + b}} \right), & \text{otherwise.} \end{cases} \quad (7)$$

Then, we normalize the weight value in (8) so that  $\sum_{i=1}^n w(s, i) = 1$ :

$$w(s, i) \leftarrow \frac{w'(s, i)}{\sum_{j=1}^n w'(s, j)}, \quad (8)$$

### 3.2.2. Fuzzy Reward

This section presents the FIS used to produce fuzzy rewards for RL to learn fusion weights and adjust the Q-learning reward. The FIS is composed of a zero-order Takagi–Sugeno–Kang (TSK) fuzzy system [41–43], which can be defined as

$$\begin{aligned} R_i : & \text{If } x_1(k) \text{ is } A_{i1} \text{ And } \dots \text{ And } x_n(k) \text{ is } A_{in} \\ & \text{Then } y_1(k) \text{ is } a_i, \end{aligned} \quad (9)$$

where  $x_1(k), \dots, x_n(k)$  represents the input variables at time  $k$ ,  $A_{i1}, \dots, A_{in}$  are the fuzzy sets, and  $a_i$  represents the singleton consequence. Moreover,  $\mu_{A_{ij}}$  is the membership value of  $A_{ij}$ , and  $\Phi_i$  is the firing strength of rule  $R_i$ . We use algebraic multiplication to implement the fuzzy AND operation. Then,  $\Phi_i$  with input data set  $\vec{x}(k) = [x_1(k), \dots, x_n(k)]$  can be described as

$$\Phi_i(\vec{x}(k)) = \prod_{j=1}^n \mu_{A_{ij}}(x_j(k)) \quad (10)$$

Supposing the FIS consists of  $r$  rules, the output of the FIS  $y(k)$  can be calculated by the weighted average defuzzification method in (11).

$$y(k) = \frac{\sum_{i=1}^r \Phi_i(\vec{x}(k)) a_i}{\sum_{i=1}^r \Phi_i(\vec{x}(k))} \quad (11)$$

To properly score the relationship between the generated trust value and human performance, we used the fuzzy reward to feed back the score to the proposed trust model to tune the fusion weights. We defined four rules for reward evaluation based on human performance for the Q-learning algorithm. There are two input variables for each fuzzy rule: human reaction time  $\tau_h$  and human trust value  $F_h$ . Here, the reaction times are divided into fast and slow camps, and the trust values also contain high and low levels. Thus, four combinations exist. The rules are defined as follows.

$R_1$ : If  $\tau_h(k)$  is  $A_{fast}$  and  $F_h(k)$  is  $B_{high}$ , then  $r(k) = 1$ .

$R_2$ : If  $\tau_h(k)$  is  $A_{slow}$  and  $F_h(k)$  is  $B_{low}$ , then  $r(k) = 1$ .

$R_3$ : If  $\tau_h(k)$  is  $A_{fast}$  and  $F_h(k)$  is  $B_{low}$ , then  $r(k) = -1$ .

$R_4$ : If  $\tau_h(k)$  is  $A_{slow}$  and  $F_h(k)$  is  $B_{high}$ , then  $r(k) = -1$ .

where  $A_{fast}$  and  $A_{slow}$  are fuzzy sets describing fast and slow human reaction times, respectively. Specifically, under  $R_1$ , humans make decisions faster and are trustworthy, the trust value of the fusion result is high, which represents a positive result, and the reward value of  $R_1$  is 1. Under  $R_2$ , humans are slow to make decisions and, thus, untrustworthy. In addition, the fusion result of the trust value is low. The two input variables of  $R_2$  both

show a consistently negative result, so the reward value of  $R_2$  is 1. However, if the trend is inconsistent, as in  $R_3$  and  $R_4$ , the reward is  $-1$ .

The membership value of  $A_{fast}$  and  $A_{slow}$  is computed by the membership function, as follows

$$\mu_{A_{fast}} = \begin{cases} f(\tau_h | m_{fast}, \sigma_{fast}), & \tau_h > m_{fast}, \\ 1, & otherwise \end{cases} \quad (12)$$

and

$$\mu_{A_{slow}} = \begin{cases} f(\tau_h | m_{slow}, \sigma_{slow}), & \tau_h > m_{slow}, \\ 1, & otherwise \end{cases} \quad (13)$$

where  $f(x|m, \sigma) = \exp\left[-\frac{(m-x)^2}{\sigma^2}\right]$ ,  $B_{high}$  and  $B_{low}$  are the fuzzy sets describing high and low human trust values, respectively. The membership value of  $B_{high}$  and  $B_{low}$  is computed by the membership function as follows:

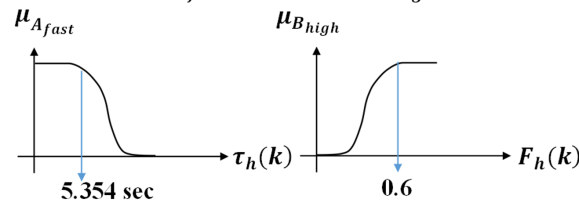
$$\mu_{B_{high}} = \begin{cases} f(F_h | m_{high}, \sigma_{high}), & F_h > m_{high}, \\ 1, & otherwise \end{cases} \quad (14)$$

and

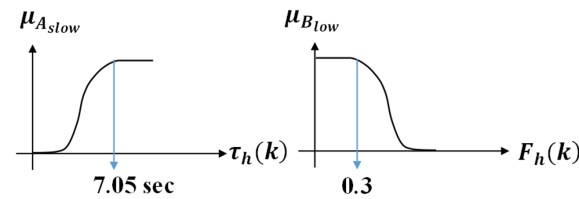
$$\mu_{B_{low}} = \begin{cases} f(F_h | m_{low}, \sigma_{low}), & F_h > m_{low}, \\ 1, & otherwise \end{cases} \quad (15)$$

where  $m_{high}$  and  $m_{low}$  of the reaction time are calculated using the average reaction time and standard deviation of reaction times from all participants. The slow reaction time is defined as the time values that are twice the standard deviation more than the average reaction time, and the fast reaction time is twice the standard deviation less than the average reaction time. Figure 3 shows a schematic diagram of the four rules.

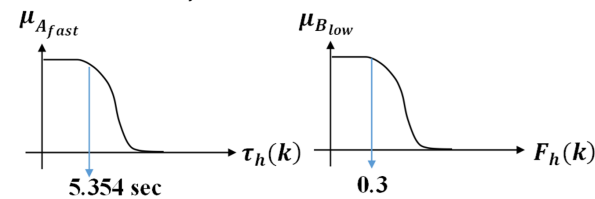
**$R_1$ : If  $\tau_h(k)$  is  $A_{fast}$  And  $F_h(k)$  is  $B_{high}$ , Then  $r(k) = 1$**



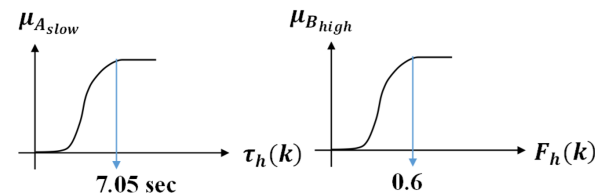
**$R_2$ : If  $\tau_h(k)$  is  $A_{slow}$  And  $F_h(k)$  is  $B_{low}$ , Then  $r(k) = 1$**



**$R_3$ : If  $\tau_h(k)$  is  $A_{fast}$  And  $F_h(k)$  is  $B_{low}$ , Then  $r(k) = -1$**



**$R_4$ : If  $\tau_h(k)$  is  $A_{slow}$  And  $F_h(k)$  is  $B_{high}$ , Then  $r(k) = -1$**



**Figure 3.** Four rules of the fuzzy neural network.

## 4. Methods

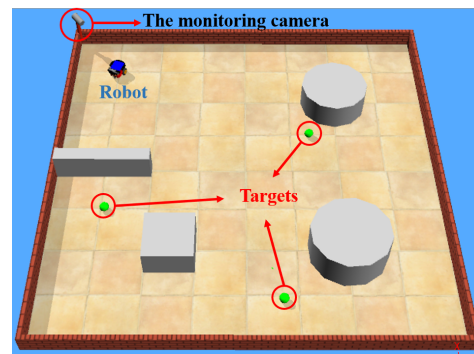
### 4.1. Participants

Six healthy male participants aged 21 to 24 years participated in this study. Following an explanation of the experimental procedure, all participants received an informed consent form and signed before participating in the study. This study received the approval of the Institute's Human Research Ethics Committee of National Chiao Tung University, Hsinchu, Taiwan. None of the participants reported a history of psychological disorder, which could have affected the experimental results.

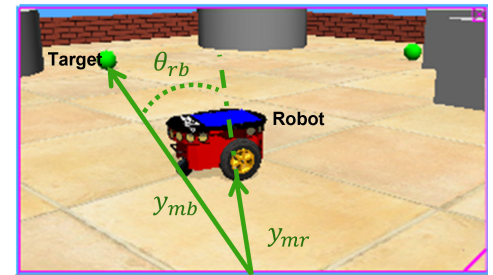
### 4.2. Scenario Design

The built simulation scenario of ball collection is designed by a professional robot simulator, Webots 8.6.2 (Cyberbotics Ltd., Lausanne, Switzerland). As shown in Figure 4, the environment is fenced with several balls and obstacles inside. A human agent and a robot agent are expected to work together to collect the balls without collision between the robot and the wall or obstacles. Humans can only observe the entire environment with restricted sight through a monitoring camera located in the top left corner of the scenario. On the basis of the observation, the human can instruct the robot to search for the ball, and the robots are also allowed to explore the environment by themselves when there is no instruction from the human or the human trust values are not high enough to be trusted.





(a) Scenario with one robot and three balls.



(b) Participant's view via the monitoring camera during the experiment.

**Figure 4.** Scenario for training data collection (Scenario\_1).

#### 4.3. Human-Agent Setup and Recording

The design of the whole scenario is affected by not only the autonomous agents' actions but also human physiological factors. Here, we use two instruments, eye-tracking and a heart rate monitoring watch, to measure the human physiological state in real time, as shown in Figure 1. Eye tracking data were recorded using the Tobii Pro X2-30 screen-based eye tracker (Tobii AB Corp., Stockholm, Sweden). We corrected the pupil data and gaze location of each participant to monitor their concentration level and fixation pathways. Heart rate data was recorded using the Empatica-E4 wristband (Empatica Inc., Cambridge, MA, USA). We used the real-time heart rate of each participant to estimate the current stress level of the human agent while performing the task.

Human perception ability was identified by the monitoring camera used to provide sight of the scenario situation for the human agent, as presented in Figure 4a. Following our definition of human perception in Section Trust Evaluation Metrics, we categorised human perception into four classes (see Table 1). Real-time perception ability is calculated according to the following formula:

$$E^a = k_1 \times \left(1 - \frac{y_{mr}}{y}\right) + k_2 \times \left(1 - \frac{y_{mb}}{y}\right) + k_3 \times \left(1 - \frac{\theta_{rb}}{\pi}\right), \quad (16)$$

where  $E^a$  is the value of current perception ability,  $k_1, k_2, k_3$  are predefined weights,  $y_{mr}$  is the distance between the monitoring camera and the robot,  $y_{mb}$  is the distance between the monitor and ball,  $\theta_{rb}$  is the deviation value between the robot and ball, and  $y$  is the distance the monitor can measure, as shown in Figure 4b.

We set  $k_3$  as the largest weight because it represents the situation in which both the robot and ball can be perceived, in which humans have the best chance of completing the task successfully.  $k_2$  is given the second highest weight that represents the situation in which the human knows the exact position of the ball, although the robot is not visible. Finally,  $k_1$  has the smallest weight, which indicates the situation in which the human does not know the location of the ball, although the robot is visible. Every situation is transformed into a corresponding evaluation value, which ranges in value between 0 and 1. Here, if the human cannot see the ball or robot, the corresponding terms are set to zero in the equation. Then, if neither the robot nor ball can be seen, the third term in the equation is set to 0. Furthermore, if an object does not exist in some conditions, the value of that item is also set to 0.

#### 4.4. Experimental Procedures

During the experiment, participants sat in front of the computer screen while performing the task. Each participant first performed a calibration for the eye tracker was performed by each participant first. Next, we introduced the operation and process of the whole experiment, including how the robot is controlled and various other considerations in the experiment. While conducting the ball-collection task, participants used two keys on

the keyboard to control the clockwise or anticlockwise rotation of the robot, following our instructions. The balls were scattered around the environment, including invisible or blind areas. The participant can only monitor the scenario from a fixed perspective, as mentioned above, and an example of the participant's view via the monitoring camera is shown in Figure 4b. The task would be completed once all the balls have been found by the robot.

For the cooperative work between human and robot agents, we define the process of their interaction in each trial. First, the human has eight seconds to set the facing direction for the robot through rotation control. Then, the robot moves in the direction the human agent has selected for 15 seconds. To discard the direction setting for the robot, the human agent is allowed to select robot self-exploration. The robot may move for more than fifteen seconds if it detects a target by itself during self-exploration. A schematic diagram of each trial is shown in Figure 5. Here,  $t_p$  represents the time that the human controls the direction of the robot,  $t_{p_{max}}$  represents the maximum time for the human to control the robot, and  $t_r$  represents the time that the robot acts and follows the command from the human.

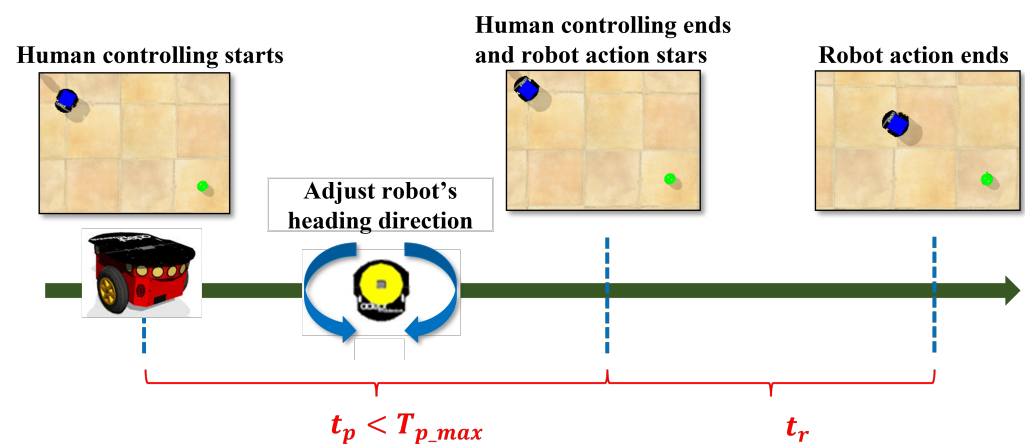


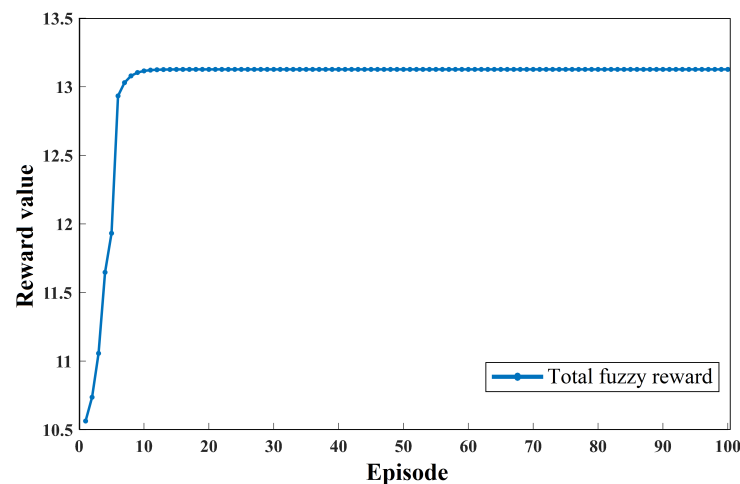
Figure 5. Robot and human interaction in each trial.

## 5. Results

This section describes the results of our simulation experiment with our proposed trust model in the human-autonomous cooperation task scenario. We first discuss the training results by presenting the convergence of the fuzzy reward and its standard deviation. Then, we provided our testing results based on three different scenarios with our trained trust model.

### 5.1. Training Results

The training process inputs the data collected in Figure 4 into the Q-learning. In (6), we set the learning rate,  $\alpha$ , to 0.1 and the discount factor,  $\gamma$ , to 0.2. These settings are commonly applied to various scenarios with the fuzzy neural network to obtain the reward value [43–45]. We implemented 100 episodes to train our model to eliminate the impact of the unstable reward in the first ten episodes. The visualized convergence result of the reward values from each episode is shown in Figure 6. The data recorded in Figure 5 were used to train the reinforcement learning method, including three pieces of human evidence signals and the reaction time of humans  $t_p$ . The training results combined cross-subject data. One of the best-performing weights with the best reward values is  $[w_1, w_2, w_3] = [0.2440, 0.3688, 0.3872]$ .



**Figure 6.** Convergence of the fuzzy reward during the training process of the fusion mechanism.

### 5.2. Testing Results

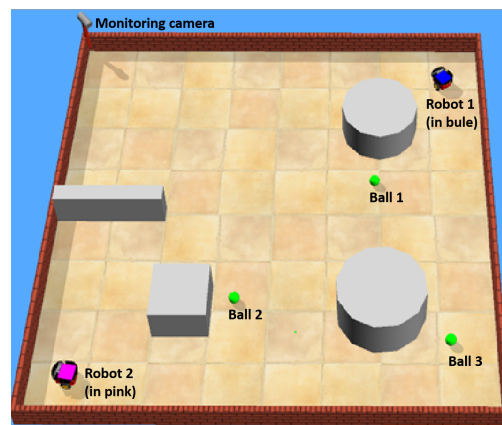
We used the trained weights to fuse the three human states. The fusion results from all six participants are used to assess the human trust value, which ranges from 0 to 1. Then, we conducted the tests in one training (Scenario\_1) and three test scenarios (Scenario\_2–4). Figure 7 presents the three test scenarios, which we refer to as Scenario\_2, Scenario\_3 and Scenario\_4. Tables 2 and 3 provide all the test results. We statistically analyzed the execution time of three task-performing modes, including human instruction (robot follows human instruction only to find the targets), a collaboration of human and robot agents (HAT) and robot random search in Table 2. The results contain both the completion time and decision time. Additionally, the number of operations in the collaboration and human instruction modes are presented in Table 3.

**Table 2.** Completion time under Scenario\_1–4, HAT represents experiments with control switching between the human and robot.

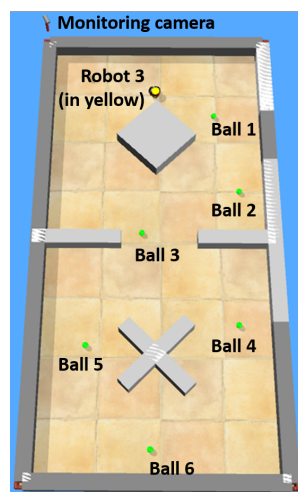
Evaluation of		Participant					
Completion Time		1	2	3	4	5	6
Scenario	Setting	Time (s)					
Scenario_1	human instruction	223	175	194	196	177	216
	HAT	194	138	171	140	131	143
	random search	287					
Scenario_2	human instruction	165	181	168	121	132	129
	HAT	117	119	123	98	100	90
	random search	185					
Scenario_3	human instruction	408	428	407	469	427	450
	HAT	372	343	371	403	380	359
	random search	573					
Scenario_4	human instruction	209	237	248	229	222	242
	HAT	178	208	195	201	197	213
	random search	330					

**Table 3.** Number of decisions made in Scenarios\_1–4.

Evaluation of Number of Decisions Made		Participant					
		1	2	3	4	5	6
Scenario	Setting	Number of Decisions					
Scenario_1	human instruction	6	6	6	6	6	6
	human/robot	4 / 2	4 / 2	4 / 2	4 / 2	5 / 1	4 / 2
Scenario_2	human instruction	6 / 6	5 / 6	6 / 6	4 / 5	4 / 6	4 / 5
	human/blue	4 / 2	4 / 1	5 / 1	3 / 1	3 / 1	2 / 2
	human/pink	2 / 4	3 / 3	4 / 2	3 / 2	5 / 1	2 / 3
Scenario_3	human instruction	12	12	13	12	12	11
	human/robot	9 / 3	9 / 3	7 / 6	7 / 5	8 / 4	6 / 5
Scenario_4	human instruction	5 / 7	6 / 7	6 / 8	6 / 7	6 / 7	8 / 7
	human/blue	3 / 2	3 / 3	2 / 4	2 / 4	4 / 2	4 / 2
	human/pink	7 / 0	7 / 0	5 / 3	5 / 2	4 / 3	4 / 3



(a) Small scenario with two robots and three balls. (Scenario\_2)



(b) Large scenario with one robot and six balls. (Scenario\_3)



(c) Large scenario with two robots and six balls. (Scenario\_4)

**Figure 7.** Scenarios for testing.

Overall, the completion time in the collaboration mode is always the shortest compared to the other two modes for all participants. Specifically, in Scenario\_1 shown in Table 2, Participant 5 spent the shortest time completing the task in collaboration mode, and the proportion of manipulation by humans was also the highest, indicating that a high level of trust was evaluated for this participant while performing the ball collection task.

In Scenario\_2, Participant 3 took the longest time to complete the task in collaboration mode. However, the proportion of manipulations by humans was also the highest in this case. This may be because Participant 3 maintained a high level of trust, but did not control the robot well, which led to a longer time required to complete the task. In Scenario\_3, the second participant took the shortest time to achieve the task in collaboration mode, and the robot was involved in the least amount of autonomous control. This result may suggest that the second participant was trustworthy and able to identify the shortest route to save a considerable amount of time during the experiment. In Scenario\_4, the result of the decisions indicates that Participants 1 and 2 controlled the robot all by themselves without robot intervention, and the completion time varied greatly. This may suggest that both participants were trustworthy, but the first participant could identify a better path than the second.

## 6. Discussion

This section discusses the improvement of efficiency among the three modes (robot random search, human instruction and collaboration/HAT) and explores the reasons for the improved efficiency. The magnitude of the improvement for each scenario is shown in Table 4.

**Table 4.** A Comparison of improvement rate in Scenarios\_1–4. HAT represents experiments with control switching between the human and robot, H represents experiments with only human instructions, and RS represents experiments with only robot random search

Scenario	Setting	Participant						Avg
		1	2	3	4	5	6	
		Improvement Rates						
Scenario_1	H vs. RS	22.29	39.02	32.4	31.71	38.33	24.74	31.42
	HAT vs. RS	32.4	51.92	40.42	51.22	54.36	50.17	46.75
	HAT vs. H	13.01	21.14	11.86	28.57	25.99	33.79	22.39
Scenario_2	H vs. RS	10.81	2.16	9.19	34.59	28.65	30.27	19.28
	HAT vs. RS	36.76	35.68	33.51	47.02	45.95	51.35	41.71
	HAT vs. H	29.09	34.25	26.79	19.01	24.24	30.23	27.27
Scenario_3	H vs. RS	28.8	25.31	28.97	18.15	25.48	21.47	24.69
	HAT vs. RS	35.08	40.14	35.25	29.67	33.68	37.35	35.19
	HAT vs. H	8.82	19.86	8.85	14.07	11.01	20.22	13.81
Scenario_4	H vs. RS	36.67	28.18	24.85	30.61	32.73	26.67	29.95
	HAT vs. RS	46.06	36.97	40.91	39.09	40.3	35.45	39.79
	HAT vs. H	14.83	12.24	21.37	12.22	11.26	11.98	13.98

In Scenario\_2, Participant 3 conducted the task with the largest number of decisions made by humans and the longest completion time, and, on the contrary, Participant 6 had the largest number of decisions made by two robots and the shortest completion time. The two robots follow Participant 3's instructions nine times in 12 trials, which is 75% of decisions made by the human, and follow Participant 2's instructions seven times in 11 trials, which is 63.6% of decisions made by the human. Whereas, Participant 6 made decisions four times for the two robots in nine trials, which is 44.4% of decisions made



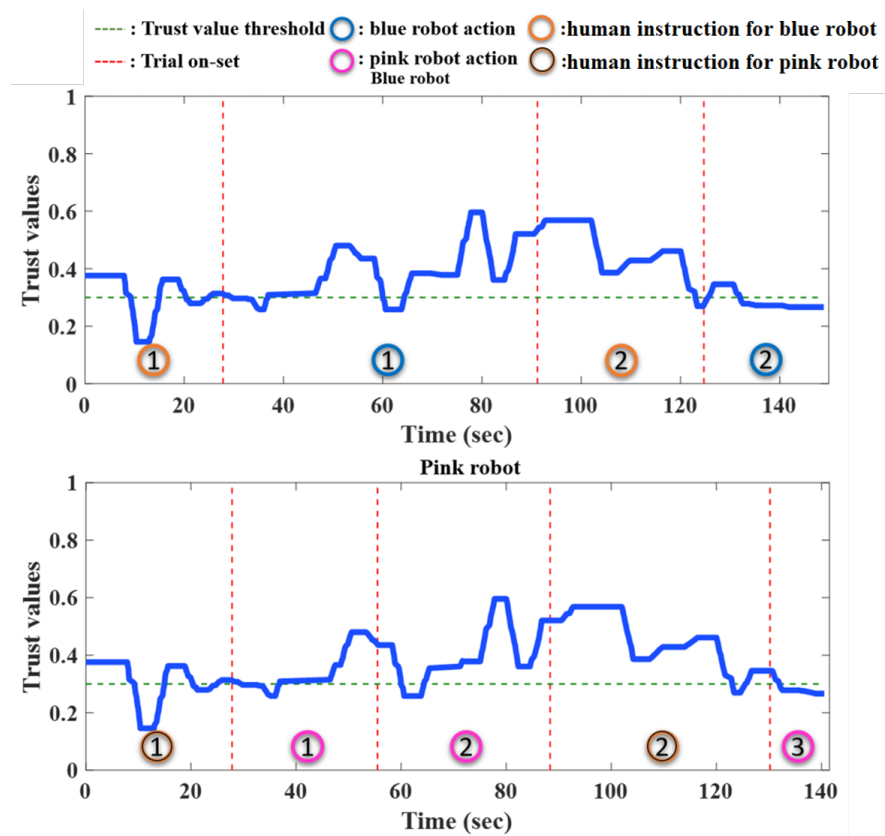
by the human. To visualise how these two participants conducted the tasks, we present the robot paths and control decisions in Figures 8 and 9 for each participant, respectively. As revealed in Figure 8, Participant 3 controlled the pink robot in the third trial, but did not adjust it in the right direction, causing the pink robot to take a long detour to find the ball and wasted a substantial amount of time. In contrast, the pathways of robots in Figure 9 indicate that Participant 6 could well control both robots and guide them on a shorter route to find the balls. Furthermore, as the greatest improvement achieved between human instruction and HAT, we also visualised the route of the task of Participant 2, as shown in Figure 10. In the human instruction condition, Participant 2 failed to adjust the robot in the right direction in the fifth trial, which resulted in a miss-out of the targets for the robot. However, in the collaboration condition, due to the lower trust value in the fifth trial for Participant 2 than the threshold value, the pink robot did not receive human instructions and proceeded forward to collect the ball. In other words, along with the successful awareness of human states, the robot made the decision itself and achieved better performance through an efficient evaluation of human states by our model.

In addition to Scenario\_2, the collaboration controlled by our proposed model also greatly improves in the more complicated scenarios. The performance of Participant 2 in Scenario\_3 indicates that the human decision was estimated to be trustworthy to make the shortest choice of path through the aid of our evaluation model on real-time human states, which achieved an optimal decision on the route to save a lot of time to complete the task. Similarly, in Scenario\_4, Participant 1 and Participant 2 were successfully evaluated as trustworthy agents, although Participant 1 could choose the better path. The test results shown in Table 2 suggest that the fusion weights trained with the Q-learning algorithm in Scenario\_1 can be directly applied to more complicated scenarios without retraining. The fusion weights were trained with collected cross-subject human data and can be used in real-time scenarios, implying that the FIS can overcome the subject difference in human data and compute appropriate rewards for the Q-learning algorithm to tune the fusion weights.

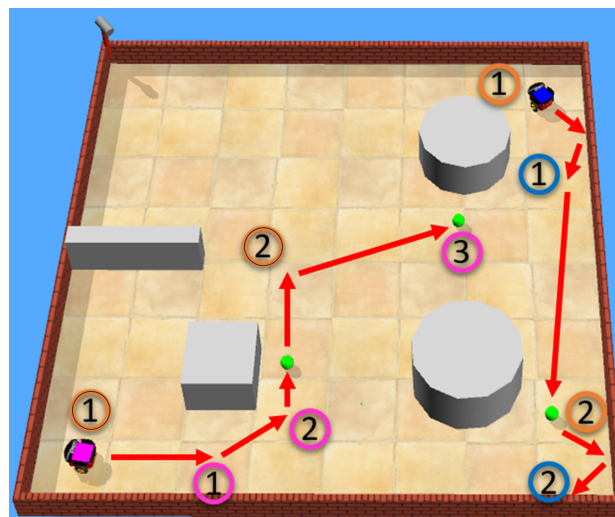
In summary, the proposed multievidence-based trust evaluation model could generate a trust-considering value for human agents that reflects the dynamics of human states in real-time. The comparison among robot random search, human instruction only, and collaboration modes demonstrates that the collaboration between human and autonomous robots controlled by the proposed trust model has adaptability and robustness for the ball-collection task under different levels, which greatly improves the task completion time compared to the other two modes.



**Figure 8.** Experimental results made by Participant 3.

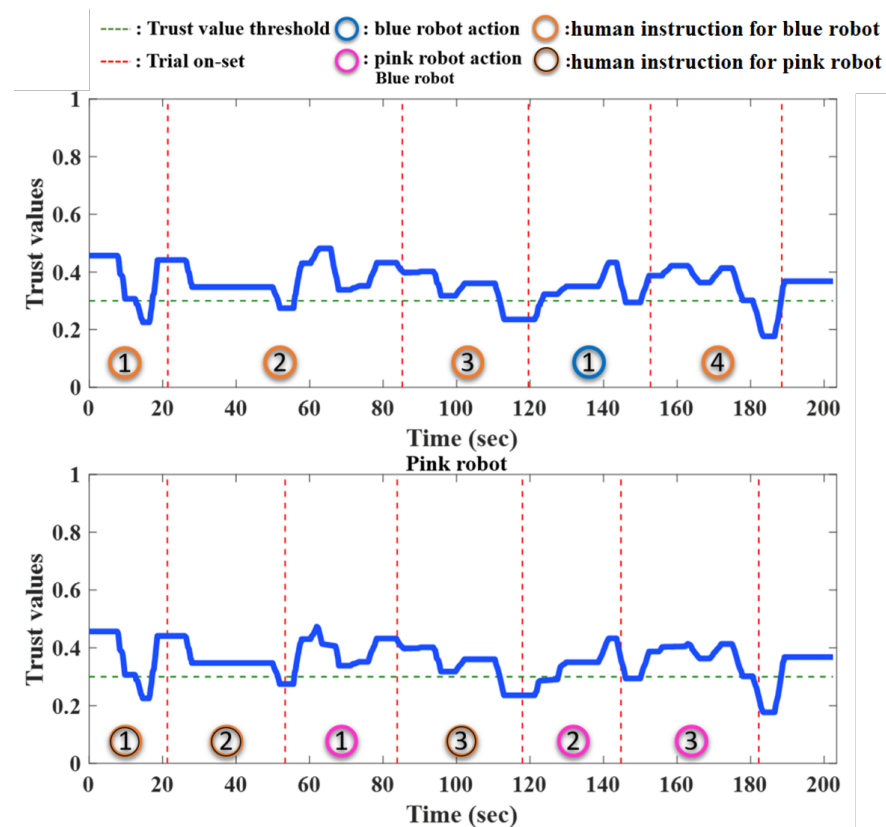


(a) Control switch between human and robot.

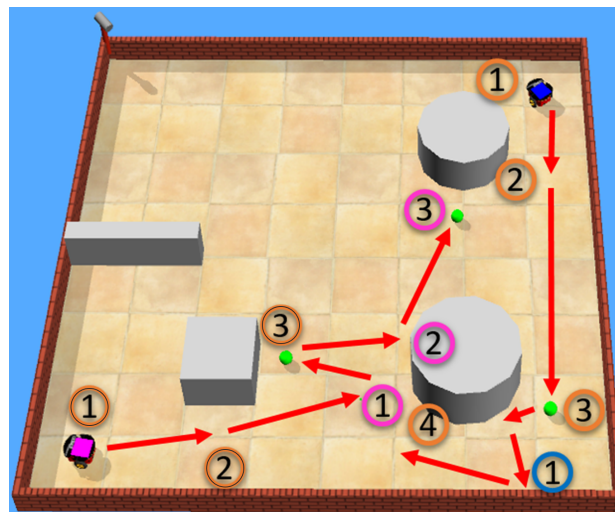


(b) Robot path trajectory.

**Figure 9.** Experimental results made by Participant 6.



(a) Control switch between human and robot.



(b) Robot path trajectory.

**Figure 10.** Experimental results of Participant 2.

## 7. Conclusions

This study proposed an adaptive trust model considering multiple real-time human cognitive states. The proposed trust model uses a fusion mechanism to combine various types of information, namely human attention level, stress index, and human perception. To verify the performance of the proposed trust model, we implemented four environmental settings, including different types of obstacles and different numbers of robot agents. We compare the performance of the HAT with those of pure human agents and those of robot agents. The results of the comparison show that the HAT team with the proposed trust

model can improve the efficiency of the given task by at least 13% in different scenario settings; The HAT team coordinated by the proposed trust model can complete the given task faster than others. Our results also suggest that the trust value generated based on these three pieces of evidence can reflect the performance of a human agent more accurately, which contributed to an improvement in efficiency for the cooperation between human and autonomous robot agents in all test scenarios. These results demonstrate that the proposed model can adapt to various levels of human performance and generate reliable trust values via the reinforcement learning algorithm. The main limitation of this study is our participant pool; only male participants were involved in our experiments. For future works, we will enlarge the participant pool and consider gender balance to conduct more comprehensive research. Furthermore, we will develop trust modelling to assess the trust of robot agents and then create a mutual trust model to provide more informatic reasoning for interaction in the HAT systems.

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

**Funding:** This work was supported in part by the Australian Research Council (ARC) under discovery grants DP180100670, DP180100656 and DP210101093. The research was also sponsored in part by the Australia Defence Innovation Hub under Contract No. P18-650825, US Office of Naval Research Global under Cooperative Agreement Number ONRG - NICOP - N62909-19-1-2058, and AFOSR – DST Australian Autonomy Initiative agreement ID10134, and AFOSR Grant No. FA2386-22-1-0042. We also thank the NSW Defence Innovation Network and NSW State Government of Australia for financial support in part of this research through grant DINPP2019 S1-03/09.

**Institutional Review Board Statement:** The study was conducted in accordance with the Declaration of Helsinki and approved by the Institutional Human Research Ethics Committee of National Chiao Tung University, Hsinchu, Taiwan.

**Informed Consent Statement:** Informed consent was obtained from all subjects involved in the study.

**Data Availability Statement:** Not applicable.

**Conflicts of Interest:** The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analysis, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

## References

1. Shneiderman, B. Human-centered artificial intelligence: Reliable, safe & trustworthy. *Int. J. Hum. Comput. Interact.* **2020**, *36*, 495–504.
2. Doroodgar, B.; Ficocelli, M.; Mobedi, B.; Nejat, G. The search for survivors: Cooperative human-robot interaction in search and rescue environments using semi-autonomous robots. In Proceedings of the 2010 IEEE International Conference on Robotics and Automation, Anchorage, AK, USA, 3–7 May 2010; pp. 2858–2863.
3. Söffker, D. From human-machine-interaction modeling to new concepts constructing autonomous systems: A phenomenological engineering-oriented approach. *J. Intell. Robot. Syst.* **2001**, *32*, 191–205.
4. Benderius, O.; Berger, C.; Lundgren, V.M. The best rated human-machine interface design for autonomous vehicles in the 2016 grand cooperative driving challenge. *IEEE Trans. Intell. Transp. Syst.* **2017**, *19*, 1302–1307.
5. Demir, M.; McNeese, N.J.; Gorman, J.C.; Cooke, N.J.; Myers, C.W.; Grimm, D.A. Exploration of teammate trust and interaction dynamics in human-autonomy teaming. *IEEE Trans. Hum. Mach. Syst.* **2021**, *51*, 696–705.
6. Dorronzoro Zubiete, E.; Nakahata, K.; Imamoglu, N.; Sekine, M.; Sun, G.; Gomez, I.; Yu, W. Evaluation of a home biomonitoring autonomous Mobile Robot. *Comput. Intell. Neurosci.* **2016**, *2016*, 9845816.
7. Robinette, P.; Howard, A.M.; Wagner, A.R. Effect of robot performance on human-robot trust in time-critical situations. *IEEE Trans. Hum. Mach. Syst.* **2017**, *47*, 425–436.
8. Pippin, C.; Christensen, H. Trust modeling in multi-robot patrolling. In Proceedings of the 2014 IEEE International Conference on Robotics and Automation (ICRA), Hong Kong, China, 31 May–7 June 2014; pp. 59–66.



9. Holbrook, J.; Prinzel, L.J.; Chancey, E.T.; Shively, R.J.; Feary, M.; Dao, Q.; Ballin, M.G.; Teubert, C. Enabling urban air mobility: Human-autonomy teaming research challenges and recommendations. In Proceedings of the AIAA AVIATION 2020 FORUM, Virtual, 15–19 June 2020; p. 3250.
10. Huang, L.; Cooke, N.J.; Gutzwiller, R.S.; Berman, S.; Chiou, E.K.; Demir, M.; Zhang, W. Distributed dynamic team trust in human, artificial intelligence, and robot teaming. In *Trust in Human-Robot Interaction*; Elsevier: Amsterdam, The Netherlands, 2021; pp. 301–319.
11. Tjøstheim, T.A.; Johansson, B.; Balkenius, C. A computational model of trust-, pupil-, and motivation dynamics. In Proceedings of the 7th International Conference on Human-Agent Interaction, Kyoto, Japan, 6–10 October 2019; pp. 179–185.
12. Pavlidis, M.; Mouratidis, H.; Islam, S.; Kearney, P. Dealing with trust and control: a meta-model for trustworthy information systems development. In Proceedings of the 2012 Sixth International Conference on Research Challenges in Information Science (RCIS), Valencia, Spain, 16–18 May 2012; pp. 1–9.
13. Kaniarasu, P.; Steinfeld, A.M. Effects of blame on trust in human robot interaction. In Proceedings of the The 23rd IEEE International Symposium on Robot and Human Interactive Communication, Edinburgh, UK, 25–29 August 2014; pp. 850–855.
14. Sadrifaridpour, B.; Saeidi, H.; Burke, J.; Madathil, K.; Wang, Y. Modeling and control of trust in human-robot collaborative manufacturing. In *Robust Intelligence and Trust in Autonomous Systems*; Springer: Berlin/Heidelberg, Germany, 2016; pp. 115–141.
15. Hu, W.L.; Akash, K.; Jain, N.; Reid, T. Real-time sensing of trust in human-machine interactions. *IFAC-PapersOnLine* **2016**, *49*, 48–53.
16. Mahani, M.F.; Jiang, L.; Wang, Y. A Bayesian Trust Inference Model for Human-Multi-Robot Teams. *Int. J. Soc. Robot.* **2020**, *13*, 1951–1965.
17. Lu, Y.; Sarter, N. Modeling and inferring human trust in automation based on real-time eye tracking data. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*; SAGE Publications Sage CA: Los Angeles, CA, USA, 2020; Volume 64, pp. 344–348.
18. Alves, C.; Cardoso, A.; Colim, A.; Bicho, E.; Braga, A.C.; Cunha, J.; Faria, C.; Rocha, L.A. Human–Robot Interaction in Industrial Settings: Perception of Multiple Participants at a Crossroad Intersection Scenario with Different Courtesy Cues. *Robotics* **2022**, *11*, 59.
19. Jacovi, A.; Marasović, A.; Miller, T.; Goldberg, Y. Formalizing trust in artificial intelligence: Prerequisites, causes and goals of human trust in AI. In Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency, Virtual, 3–10 March 2021; pp. 624–635.
20. Wang, Q.; Liu, D.; Carmichael, M.G.; Aldini, S.; Lin, C.T. Computational Model of Robot Trust in Human Co-Worker for Physical Human-Robot Collaboration. *IEEE Robot. Autom. Lett.* **2022**, *7*, 3146–3153.
21. Xing, Y.; Lv, C.; Cao, D.; Hang, P. Toward human-vehicle collaboration: Review and perspectives on human-centered collaborative automated driving. *Transp. Res. Part C Emerg. Technol.* **2021**, *128*, 103199.
22. Liu, Y.; Habibnezhad, M.; Jebelli, H. Brainwave-driven human-robot collaboration in construction. *Autom. Constr.* **2021**, *124*, 103556.
23. Chang, Y.C.; Wang, Y.K.; Pal, N.R.; Lin, C.T. Exploring Covert States of Brain Dynamics via Fuzzy Inference Encoding. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2021**, *29*, 2464–2473.
24. Guo, Y.; Yang, X.J. Modeling and Predicting Trust Dynamics in Human–Robot Teaming: A Bayesian Inference Approach. *Int. J. Soc. Robot.* **2021**, *13*, 1899–1909.
25. Azevedo-Sa, H.; Jayaraman, S.K.; Esterwood, C.T.; Yang, X.J.; Robert, L.P.; Tilbury, D.M. Real-time estimation of drivers’ trust in automated driving systems. *Int. J. Soc. Robot.* **2021**, *13*, 1911–1927.
26. Nian, R.; Liu, J.; Huang, B. A review on reinforcement learning: Introduction and applications in industrial process control. *Comput. Chem. Eng.* **2020**, *139*, 106886.
27. Joo, T.; Jun, H.; Shin, D. Task Allocation in Human–Machine Manufacturing Systems Using Deep Reinforcement Learning. *Sustainability* **2022**, *14*, 2245.
28. Yang, Y.; Li, Z.; He, L.; Zhao, R. A systematic study of reward for reinforcement learning based continuous integration testing. *J. Syst. Softw.* **2020**, *170*, 110787.
29. Chen, M.; Lam, H.K.; Shi, Q.; Xiao, B. Reinforcement learning-based control of nonlinear systems using Lyapunov stability concept and fuzzy reward scheme. *IEEE Trans. Circuits Syst. II Express Briefs* **2019**, *67*, 2059–2063.
30. Kofinas, P.; Vouros, G.; Dounis, A.I. Energy management in solar microgrid via reinforcement learning using fuzzy reward. *Adv. Build. Energy Res.* **2018**, *12*, 97–115.
31. Jafarifarmand, A.; Badamchizadeh, M.A.; Khanmohammadi, S.; Nazari, M.A.; Tazehkand, B.M. A new self-regulated neuro-fuzzy framework for classification of EEG signals in motor imagery BCI. *IEEE Trans. Fuzzy Syst.* **2017**, *26*, 1485–1497.
32. Lin, F.C.; Ko, L.W.; Chuang, C.H.; Su, T.P.; Lin, C.T. Generalized EEG-based drowsiness prediction system by using a self-organizing neural fuzzy system. *IEEE Trans. Circuits Syst. I Regul. Pap.* **2012**, *59*, 2044–2055.
33. Zhang, L.; Shi, Y.; Chang, Y.C.; Lin, C.T. Hierarchical Fuzzy Neural Networks With Privacy Preservation for Heterogeneous Big Data. *IEEE Trans. Fuzzy Syst.* **2020**, *29*, 46–58.
34. Shayesteh, S.; Ojha, A.; Jebelli, H. Workers’ Trust in Collaborative Construction Robots: EEG-Based Trust Recognition in an Immersive Environment. In *Automation and Robotics in the Architecture, Engineering, and Construction Industry*; Springer: Berlin/Heidelberg, Germany, 2022; pp. 201–215.

35. Hoeks, B.; Ellenbroek, B.A. A neural basis for a quantitative pupillary model. *J. Psychophysiol.* **1993**, *7*, 315–315.
36. Baevsky, R.M.; Chernikova, A.G. Heart rate variability analysis: Physiological foundations and main methods. *Cardiometry* **2017**, 66–76, doi:10.12710/cardiometry.2017.10.6676.
37. Silambarasan, I.; Sriram, S. Hamacher sum and Hamacher product of fuzzy matrices. *Intern. J. Fuzzy Math. Arch.* **2017**, *13*, 191–198.
38. Watkins, C.J.; Dayan, P. Q-learning. *Mach. Learn.* **1992**, *8*, 279–292.
39. Clifton, J.; Laber, E. Q-learning: theory and applications. *Annu. Rev. Stat. Its Appl.* **2020**, *7*, 279–301.
40. Lin, J.L.; Hwang, K.S.; Wang, Y.L. A simple scheme for formation control based on weighted behavior learning. *IEEE Trans. Neural Networks Learn. Syst.* **2013**, *25*, 1033–1044.
41. Qin, B.; Chung, F.L.; Wang, S. KAT: A Knowledge Adversarial Training Method for Zero-Order Takagi-Sugeno-Kang Fuzzy Classifiers. *IEEE Trans. Cybern.* **2020**, *52*, 6857–6871.
42. Tkachenko, R.; Izonin, I.; Tkachenko, P. Neuro-Fuzzy Diagnostics Systems Based on SGTm Neural-Like Structure and T-Controller. In *Proceedings of the International Scientific Conference “Intellectual Systems of Decision Making and Problem of Computational Intelligence”*; Springer: Berlin/Heidelberg, Germany, 2021; pp. 685–695.
43. Lin, C.J.; Lin, C.T. Reinforcement learning for an ART-based fuzzy adaptive learning control network. *IEEE Trans. Neural Networks* **1996**, *7*, 709–731.
44. Xie, J.; Xu, X.; Wang, F.; Liu, Z.; Chen, L. Coordination Control Strategy for Human-Machine Cooperative Steering of Intelligent Vehicles: A Reinforcement Learning Approach. *IEEE Trans. Intell. Transp. Syst.* **2022**, 1–15, doi:10.1109/TITS.2022.3187016.
45. Lin, C.T.; Kan, M.C. Adaptive fuzzy command acquisition with reinforcement learning. *IEEE Trans. Fuzzy Syst.* **1998**, *6*, 102–121.