machines

MDPI

Article

Online Hand Gesture Detection and Recognition for UAV Motion Planning

Cong Lu [1], Haoyang Zhang [2,3,*], Yu Pei [2,3], Liang Xie [2,3], Ye Yan [2,3], Erwei Yin [2,3] and Jing Jin [4,5]

[1] School of Information Science and Engineering, East China University of Science and Technology, Shanghai 200237, China
[2] National Innovation Institute of Defense Technology, Academy of Military Sciences, Beijing 100071, China
[3] Tianjin Artificial Intelligence Innovation Center, Tianjin 300450, China
[4] Key Laboratory of Smart Manufacturing in Energy Chemical Process, Ministry of Education, East China University of Science and Technology, Shanghai 200237, China
[5] Shenzhen Research Institute of East China University of Science and Technology, Shenzhen 518063, China
* Correspondence: haoyang@tju.edu.cn

Abstract: Recent advances in hand gesture recognition have produced more natural and intuitive methods of controlling unmanned aerial vehicles (UAVs). However, in unknown and cluttered environments, UAV motion planning requires the assistance of hand gesture interaction in complex flight tasks, which remains a significant challenge. In this paper, a novel framework based on hand gesture interaction is proposed, to support efficient and robust UAV flight. A cascading structure, which includes Gaussian Native Bayes (GNB) and Random Forest (RF), was designed, to classify hand gestures based on the Six Degrees of Freedom (6DoF) inertial measurement units (IMUs) of the data glove. The hand gestures were mapped onto UAV's flight commands, which corresponded to the direction of the UAV flight.The experimental results, which tested the 10 evaluated hand gestures, revealed the high accuracy of online hand gesture recognition under asynchronous detection (92%), and relatively low latency for interaction (average recognition time of 7.5 ms; average total time of 3 s).The average time of the UAV's complex flight task was about 8 s shorter than that of the synchronous hand gesture detection and recognition. The proposed framework was validated as efficient and robust, with extensive benchmark comparisons in various complex real-world environments.

Keywords: IMU data glove; hand gesture detection; hand gesture recognition; UAV motion planning; interaction efficiency

check for updates

Citation: Lu, C.; Zhang, H.; Pei, Y.; Xie, L.; Yan, Y.; Yin, E.; Jin, J. Online Hand Gesture Detection and Recognition for UAV Motion Planning. *Machines* **2023**, *11*, 210. https://doi.org/10.3390/machines11020210

Academic Editor: Ning Sun

Received: 18 December 2022
Revised: 20 January 2023
Accepted: 27 January 2023
Published: 1 February 2023

Copyright: © 2023 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

Hand gestures can interact with human beings or machines in the real world. Compared to speech interaction [1] and gaze-based interaction [2], gesture interaction is relatively natural, private, and intuitive. There are numerous applications in the area of hand gesture recognition, such as sign languages [3], interaction in the metaverse [4], and UAV flights [5] and so on. In addition, hand gestures are usually used to interact with unmanned systems in rescue operations, which contributes to efficient human–machine collaboration in certain environments [6].

The latest development in hand gesture recognition is primarily based on multi-modal sensing, which is called MGR (Multi-modal Hand Gesture Recognition). MGR can improve the performance of hand gesture recognition by incorporating multisensor data. A prototype system, including a three-axis accelerometer and four-surface electromyographic (sEMG), has been developed, to realize gesture-based real-time interaction [7]. Accelerometers can measure acceleration (ACC) from vibrations and gravity. In addition, sEMG signals can indicate the activities of related muscles. Using the fusion of ACC and sEMG signals, response time and recognition accuracy can be improved. Moreover, a real-time hand gesture tracking and recognition system, based on inertial sensors and visual sensors, has

been proposed [8]. Long-term error growth in the position and orientation estimation of IMU can be compensated by the vision sensor. IMU can track fast motions much better than the vision sensor does. The recognition time has been improved greatly. Traditional hand gesture recognition is mostly based on unimodal sensing, which is called UGR (Unimodal Hand Gesture Recognition). UGR uses a single sensor (e.g., IMU, sEMG, visual sensor) and the data resource is unitary; therefore, the data alignment is not considered to fuse the multichannel data. For example, using the sEMG signal on the forearm, a real-time hand gesture recognition system was proposed [9]. Based on the best feature, the classification accuracy was satisfying, but the sEMG signal was susceptible to noise interference. Through the camera Kinect, the Histogram of Oriented Gradient (HOG) was used, to recognize the palm gestures [10]: however, the vision sensor was limited by illumination, occlusion, and space [11]; therefore, wearable IMU devices were designed, so that the sensor could capture the motion of the fingertip and the postures of the hand [11]. Although hand gesture interaction technology based on the IMU is maturing, efficient and robust recognition of dynamic hand gestures from the continuous data stream is a challenging task.

The main problems solved in this paper are as follows:

(i) poor adaptability of hand gestures design can lead to a high rate of nonrecognition;
(ii) unsatisfying hand gesture detection and recognition results affect the real-time performance and stability of the system;
(iii) UAV's complex flight tasks assisted by hand gesture interaction in an unknown and cluttered environment have not been considered well.

To address problem (i), a novel hand gesture set was designed. Hand gestures of high accuracy were considered for further combination, and the adaptability of hand gestures was improved. For problem (ii), a hierarchical structure of hand gesture detection and recognition was adopted. The hand gesture detection classifier was used, to identify *gesture* and *non-gesture*. When the hand gesture was detected as *gesture*, the specific label was given by the hand gesture recognition classifier. Problem (iii) was overcome by integrating online hand gesture detection and recognition into UAV motion planning. The commands of online hand gesture detection and recognition could be abstracted as a series of discrete points that indicated the direction of UAV flights; therefore, a novel method to guide UAV motion planning using hand gesture interaction is proposed.

According to the core idea of closed-loop feedback in control theory, the proposed system is presented in Figure 1.

A real-time hand gesture was generated by the user with the IMU data glove. The hand gesture command was taken as the set point. The continuous data stream of dynamic hand gestures was segmented by the hand gesture segmentation module. The segmented data were sent to the hand gesture detection module and hand gesture recognition module. The hand gesture was recognized and mapped onto the control command. The control command was transferred to the navigation point. The UAV was considered as a controlled object, which flew along navigation points. Through the service and topic communication of the Robot Operating System (ROS), waypoint information was sent to the MAVROS node, which translated into understandable orders for the UAV movement. Then, MAVROS communicated with PX4 firmware based on the MAVLink (Micro Air Vehicle Link), and the UAV flew in the direction provided by the hand gesture command. The mapping (completed by D435i) and positioning (completed by T265) information of the UAV served as feedback to the user's front-end. The feedback information on mapping and positioning was used for the PID operation. Thus, a closed-loop negative feedback system was constructed.

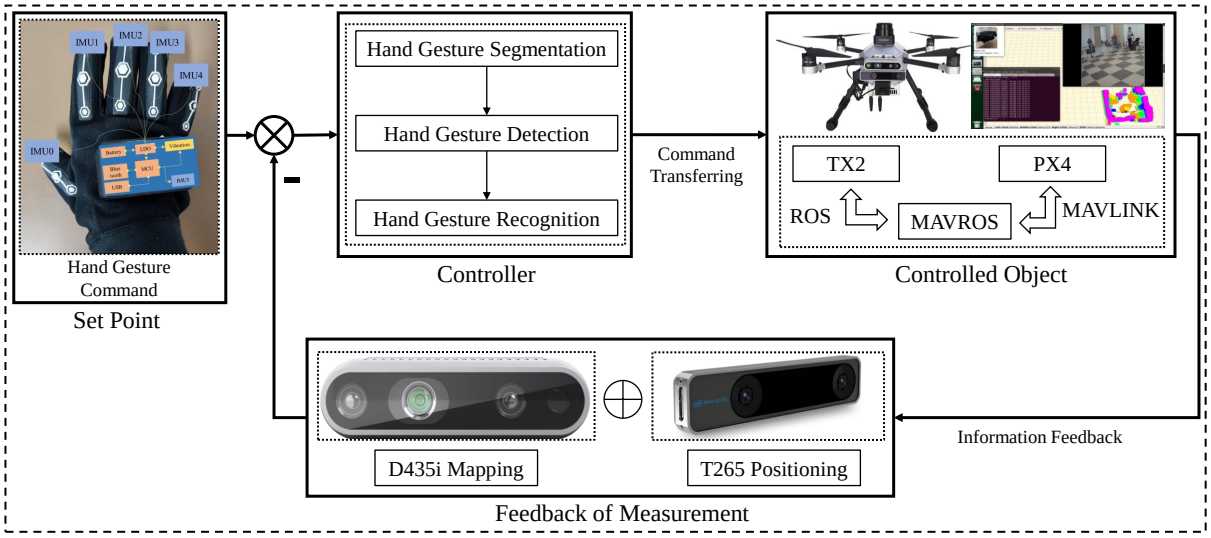


Figure 1. The proposed system. Set Point: hand gesture command with IMU data glove. Controller: hand gesture segmentation algorithm; hand gesture detection algorithm; hand gesture recognition algorithm; the recognized result is transferred into the navigation point. Controlled Object: UAV integrated with the motion planning framework of the ego-planner. Feedback of Measurement: the depth camera (D435i) for mapping and the binocular camera (T265) for localization.

The main contribution of this paper is as follows:

- To improve the adaptability of hand gestures, an IMU data glove with a high signal-to-noise ratio and high transmission rate was developed, and a public hand gesture set was designed for interaction between hand gestures and UAV.
- To enhance the effectiveness and robustness of the system, a new asynchronous hand gesture detection and recognition method was proposed, which cascaded two high-precision classifiers.
- To overcome the problem of UAV's complex flight tasks in unknown and cluttered environments, an online hand gesture detection and recognition method was innovatively applied to UAV motion planning, which realized complex flight tasks asynchronously.

The remainder of this paper is organized as follows. In Section 2, we discuss the work as it relates to three aspects of the data glove and hand gestures, i.e., hand gesture segmentation and recognition, and the interaction between hand gestures and UAV. Then, in Section 3, the hardware structure of the IMU data glove, and the design method of the hand gesture set is introduced. An online hand gesture detection and recognition method is proposed for asynchronously interacting with the UAV. Extensive real-world experiments are conducted, and the experimental results are analyzed in Section 4. The proposed work is discussed in Section 5. Finally, in Section 6, the research work is summarized, and the in-depth research ideas are proposed.

2. Related Work

Our research work is mainly related to three aspects: (1) data glove and hand gesture set construction; (2) hand gesture segmentation and recognition; (3) hand gesture and UAV interaction.

Recently, various wearable data gloves have been developed. For example, Mummadi et al. introduced a new data glove equipped with Intel's Edison module, which uses Bluetooth or WiFi protocol, and connects serially with five IMUs at all fingers through the multiplexer [12]. While the hardware components were designed to be miniaturized, the cost of the entire equipment was relatively high, and the time of hand gesture detection was relatively long; therefore, another low-cost and real-time embedded device, based on Arduino Nano 33 BLE, was developed, and a hand gesture set was built. In a decod-

ing task of eight hand gestures, online recognition accuracy of 95% was achieved [13]. Moreover, based on IMU, electromyographic (EMG), and finger and palm pressure data, Zhang et al. [14] built a relatively larger database of 10 sequential hand gestures, which consisted of 5 dynamic gestures and 5 air gestures collected from 10 participants. The classification accuracy for the dynamic hand gestures and air hand gestures was 89.28% and 76.86%, respectively, by long short-term memory (LSTM). Furthermore, Jiang et al. [15] designed eight air gestures and four surface gestures with two distinct force levels, based on a real-time gesture recognition wristband with sEMG and IMU sensing fusion. The classification accuracy for the initial experiment was 92.6% and 88.8%, respectively, for air and surface gestures. With respect to the first aspect, the current study proposes a data glove integrating with six-axis IMU sensors for hand gesture design. The movement of fingers can be detected by five relevant IMUs. Previous research on hand gesture recognition has typically classified single hand gestures, and the extensibility of the hand gesture set has not been well-considered. This work proposes a multiple-hand-gestures set, which corresponds to the commands set of UAV flights.

For hand gesture segmentation and recognition, a cascaded artificial neural network (ANN) structure has been proposed, to recognize interactive and non-interactive gestures. In addition, a high recognition rate of over 96% for 30 hand gestures was achieved [16]. However, the ANNs method requires more training time and computational resources; therefore, Simao et al. [17] introduced an arm or gesture segmentation method, based on an unsupervised threshold value: it could accurately divide a continuous data stream into dynamic and static segments, which reduced the number of misclassified hand gestures with an average over-segmentation error of 2.70%, and improved the quality of hand gesture segmentation. In addition, Li Q et al. proposed a dynamic hand gesture spotting algorithm based on the evidence theory, by which accurate and real-time dynamic hand gesture spotting was realized. The experimental results showed that the accuracy of recognition after spotting was higher than that of simultaneous recognition and spotting [18]. However, by the method based on the experience and threshold, it was hard to obtain the best effect; therefore, a dynamic hand gesture spotting algorithm based on deep learning was proposed [19]. The start and end of the hand gesture sequence in a continuous data stream were detected by a scalar value. The proposed system took about 12 ms to recognize the complete hand gesture in real-time. Regarding the second aspect, based on threshold detection, a certain interval of inactivity was considered, to ensure the clarity of hand gesture segmentation. In addition, our proposed work considered cascading two high-precision models, to complete hand gesture detection and recognition. Compared with previous studies, the proposed system could realize lower recognition time, and be more robust and efficient.

In terms of hand gesture recognition interacting with UAV, most studies have only verified the interaction performance between hand gesture recognition and simple UAV flight. For example, the wearable device's 9-axis IMU is used to recognize the user's hand gesture commands, and to control UAV flight in real-time [20]. In addition, using the open-source visual library, a robust algorithm was designed to distinguish nine hand gestures for multi-UAV control [21]. Recently, to manipulate the UAV speedily and flexibly with the hand gesture, Yu C et al. proposed an end-side hand gesture recognition for real-time and stable control of the UAV. The trained model was deployed on a wireless data glove based on the STM32 microcontroller, in which the back propagation (BP) network was used for static gesture recognition, and the bidirectional gated recurrent unit (Bi-GRU) network was used for dynamic gesture recognition [22]. However, most of the research has not considered the scenario of complex flight tasks, and the UAV motion planning methods have only used the onboard computing resources to complete real-time trajectory planning and optimization, without considering external intervention to assist flight. For example, in order to solve the problem of autonomous navigation of UAV in an unknown and complex environment, a robust and efficient quadrotor motion planning system was proposed [23]. To find a safe, kinodynamic-feasible and collision-avoidance trajectory, a Euclidean Signed Distance

Field (ESDF)-free gradient-based planning framework was proposed, which significantly reduced computation time [24]. About the third aspect, this research was innovative, in that hand gestures were used to guide the UAV motion planning, rather than simply controlling the UAVs, as in most previous studies. The collision avoidance of the UAV flight was considered throughout the hand gesture interaction.

3. Materials and Methods

3.1. IMU Data Glove

The IMU data glove is designed to consider the requirements of limb movement and practical applications [13]. As shown in Figure 2, the hardware structure of the IMU data glove is divided into the following modules.

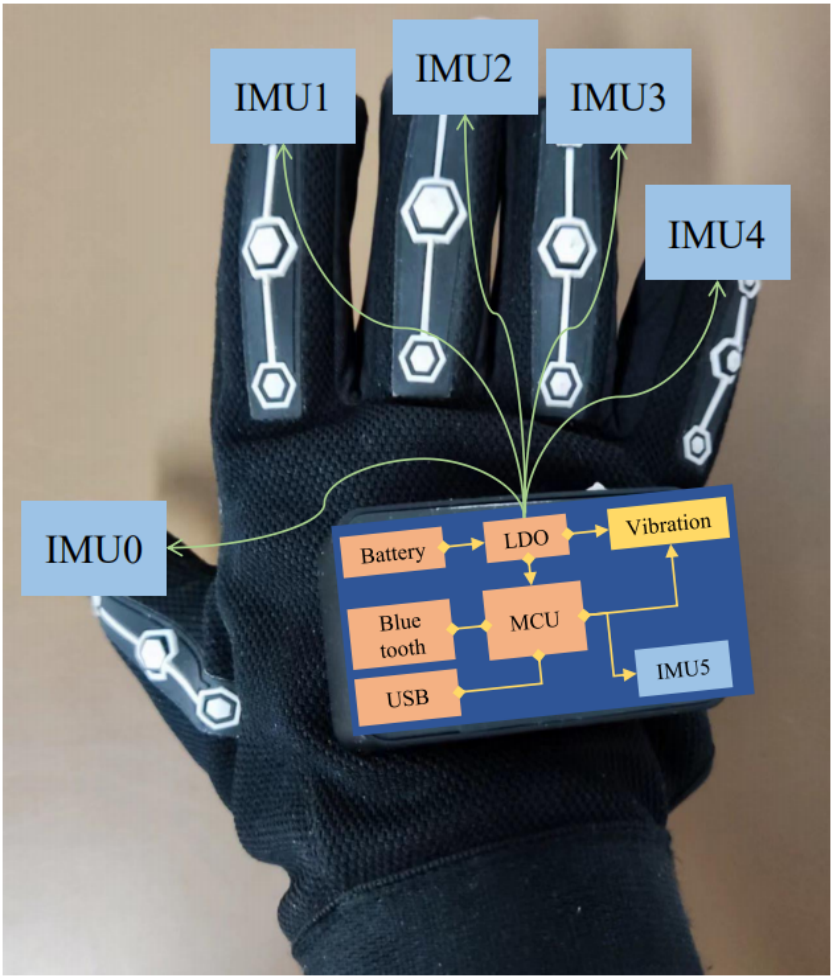


Figure 2. The hardware structure of the IMU data glove. Five IMU modules (IMU0∼IMU4) are located at the distal fingertips, and one IMU module (IMU5) is located at the metacarpal of the back of the hand. A battery supplies power. MCU is the microcontroller. Bluetooth and USB are the communication interfaces. Vibration provides force feedback.

3.1.1. Central Control and Wireless Transmission Module

The IMU data glove mainly includes a microcontroller (MCU, STM32L151CCT), a Bluetooth module (FSC-BT826B), a power management unit (AP2112-3V3 and LP5907), and a Universal Serial Bus (USB)-to-Serial Peripheral Interface (SPI) module. The battery supplies power to the regulator (LDO), which then supplies power to other modules. The battery can be charged via USB. The microcontroller writes and reads the register of the sensor, and communicates with the IMU and vibration sensor through the SPI protocol. The collected data are transmitted to the external personal computer (PC) through Bluetooth or

USB in hexadecimal formats. The start point and end point are distinguished by the 5-byte header and the 1-byte flag of each frame, respectively. Consequently, the output data of the IMU data glove are 78 bytes.

3.1.2. Distributed Multi-Node IMU Module

It is integrated with six 6-axis IMU modules (ICM20648). Five of the IMU modules are located at the distal fingertips, and one IMU module is located at the metacarpal of the back of the hand. Each IMU combines a 3-axis accelerometer (range: ±2 g; resolution: 16,384 LSB/g) and 3-axis angular velocity (range: ±2000 deg/s; resolution: 16.4 LSB/(deg/s)). The internal data transmission adopts an SPI interface. The IMU data of the wrist and each finger include 3-axis acceleration and 3-axis angular velocity values of 12 bytes.

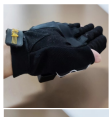
3.1.3. Vibration Motor Module

The module consists of a vibration motor and a driver chip A1442. The A1442 is a full-bridge motor driver for driving a low-voltage bipolar brushless direct current (DC) motor. The vibration motor can provide force feedback for each hand gesture.

3.2. Hand Gesture Set Design

Considering users' habits and the accessibility of the UAV control, a single-hand-gestures set was designed for UAV's complex flight tasks, which is shown in Table 1. All hand gestures corresponded to the direction of the UAV flight as closely as possible, and were easy to perform and remember.

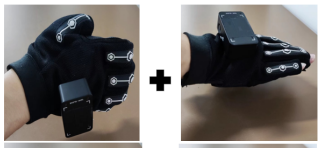
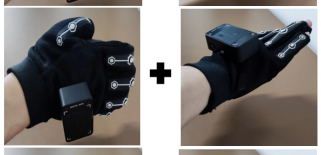
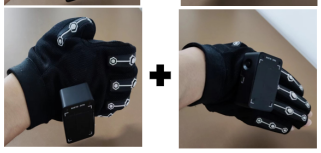
Table 1. Single-hand-gestures set.

ID	Hand Gesture	Comment	ID	Hand Gesture	Comment
S1	LD [1]		S2	RD [2]	
S3	UM [3]		S4	DM [4]	
S5	HF [5]		S6	CF [6]	
S7	TU [7]		S8	OS [8]	
S9	LF [9]		S10	RS [10]	

[1] LD: Left Deflection; [2] RD: Right Deflection; [3] UM: Upward Diagonal Move; [4] DM: Downward Diagonal Move; [5] HF: Half-Clenched Fist; [6] CF: Clenched Fist; [7] TU: Thumbs Up; [8] OS: OK Sign; [9] LF: Little Finger; [10] RS: Rock Sign.

Furthermore, the multiple-hand-gestures set was designed, as shown in Table 2, in order to reduce the rate of hand gesture misoperation [25], and to improve the adaptability of hand gesture interaction. A number of practical hand gestures were selected and combined into the multiple-hand-gestures set. The initial hand gesture of HF was performed first, followed by a direction hand gesture (i.e., LD, RD, UM, DM), to determine the direction of the UAV flight. The extensibility and adaptability of hand gestures can be improved by the way of combination.

Table 2. Multiple-hand-gestures set.

ID	Hand Gesture	Comment
C1	HF+LD [1]	
C2	HF+RD [2]	
C3	HF+UM [3]	
C4	HF+DM [4]	

[1] HF+LD: Half-Clenched Fist and Left Deflection; [2] HF+RD: Half-Clenched Fist and Right Deflection; [3] HF+UM: Half-Clenched Fist and Upward Diagonal Move; [4] HF+DM: Half-Clenched Fist and Downward Diagonal Move.

3.3. Dataset Generation

Ten healthy subjects (5 male/5 female, average age: 25 years) participated in this experiment. During the collection, the image of each hand gesture was shown on the screen for 2 s, to prompt the subjects to perform the corresponding hand gesture, followed by a rest of 2 s. Considering the time difference of the hand gesture data set, session 2 and session 3 were collected one hour later and one day later, respectively [15]. Each subject performed each gesture ten times for three sessions, of which 3300 samples (10 subjects * 3 sessions * 10 trials * 11 hand gestures) of hand gestures were collected. The experiment protocol is shown in Figure 3.

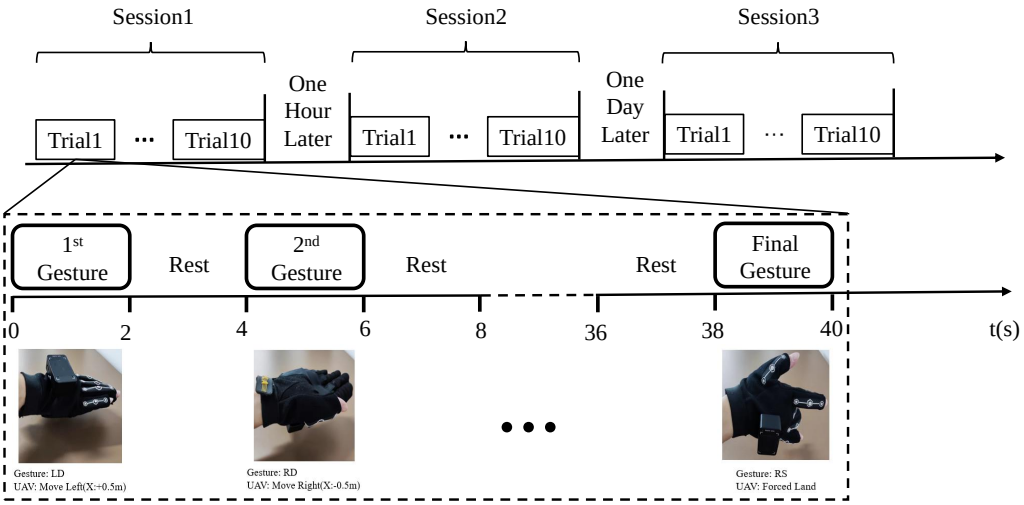


Figure 3. Experiment protocol of hand gesture data collection.

A right-hand IMU data glove was worn by each subject, and connected to the PC through a USB port. The subjects maintained a standing position, with their palms perpendicular to the ground, and their arms parallel to the ground. After turning on the glove, the hand gesture data collection program started to run, and the subjects were required to perform hand gestures according to the prompts on the screen. Afterwards, an IMU hand gesture data set was generated, by extracting both accelerated velocity and angular velocity at a sampling rate of 200 Hz.

According to the requirements of the UAV's complex flight task, the hand gesture commands were mapped onto the UAV flight motions, as listed in Table 3.

Table 3. Hand gestures mapped onto UAV flight motions.

ID	Hand Gesture	Flight Motion
S1	LD	Move Left
S2	RD	Move Right
S3	UM	Move Up
S4	DM	Move Down
S5	HF	Wait for Combination
S6	CF	Disarm
S7	TU	Take off on High
S8	OS	Arm
S9	LF	Hover
S10	RS	Forced Land
C1	HF+LD	Move Forward
C2	HF+RD	Move Backward
C3	HF+UM	Turn Left
C4	HF+DM	Turn Right

3.4. Hand Gesture Detection and Recognition

In Figure 4, the details of model training and online recognition are presented. In terms of the model training of hand gesture detection, ten dynamic hand gestures were labeled as 0, and one static hand gesture was labeled as 1. For the hand gesture recognition, ten hand gestures were labeled as 0 to 9. Due to the fact that the IMU signals were subjected to interference and extra slight movement of the hand, a Butterworth filter was designed during the data processing. The variance was extracted as the feature to characterize the hand gesture movement information. Finally, the feature matrix was readjusted to adapt to the training model. In order to distinguish the *gesture* and *non-gesture*, an imbalanced two-classifier was trained. Afterwards, a balanced ten-classifier was trained to recognize the specific categories of hand gestures. In the online recognition phase, the hand gestures were recognized based on the cascaded model, and the recognized results were transferred into control commands.

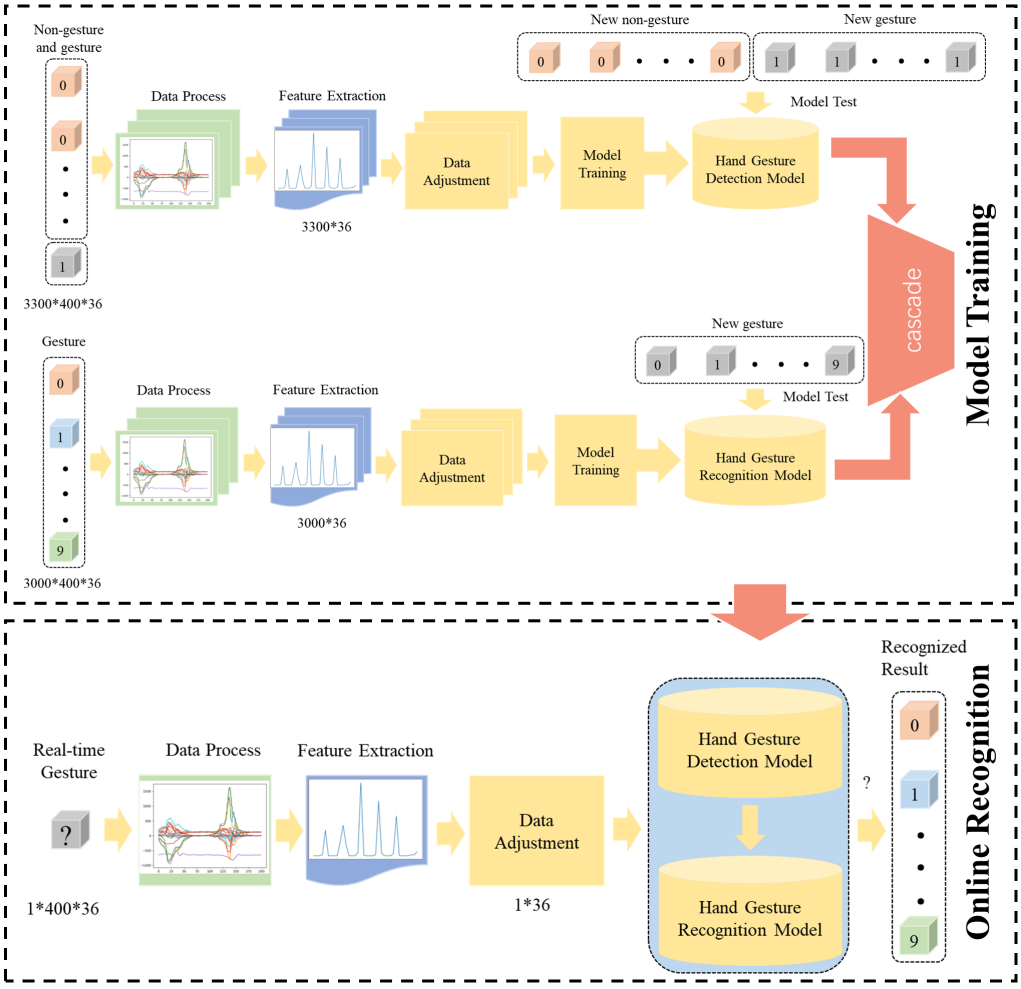


Figure 4. The proposed model architecture. In the model training phase, the model of hand gesture detection was trained to detect *gesture* and *non-gesture*. The model of hand gesture recognition was trained to recognize specific categories of predefined hand gestures. In the online recognition phase, the hand gestures were recognized in real-time by cascading hand gesture detection model and hand gesture recognition model.

The key steps of the online hand gesture detection and recognition are presented in Algorithm 1. Multithreading was used to implement online hand gesture detection and recognition. In the child thread (lines 1–16), the sliding window data were the global variable W, with the fixed frames of f, and the IMU observation data O were received from the IMU data glove per frame. When the child thread was running, W was initialized by the O of f frames (lines 2–4). Then, the sum of angular velocity and the difference of $T(j)$ and $T(j-1)$ were always calculated (lines 6 and 7). If the $GS(i)$ was greater than the threshold value of the angular velocity GT, and the time interval between the two hand gestures was greater than the threshold value TT, then the current time was updated (line 9) and W was returned (line 10). GT was set to segment the continuous data stream of dynamic hand gestures, and TT was set to ensure that there was only one hand gesture in a time window; otherwise, W would be continuously updated (lines 12–13). In the parent thread (lines 17–30), updated W was filtered by the Butterworth filter, and the variance was extracted as the feature (lines 19–21). $\tilde{W}$ after feature extraction was inputted into the hand gesture detection classifier (line 22). If the output of the hand gesture detection classifier was *gesture*, then $\tilde{W}$ was inputted into the hand gesture recognition classifier, and the specific label of the hand gesture was returned; otherwise, the *non-gesture* was returned (lines 23–28).

Algorithm 1 Online Hand Gesture Detection and Recognition

Input: O: IMU observation data; f: fixed frame length; W: sliding window data; T: current time clock; GT: threshold value of angular velocity; TT: threshold value of time interval between two hand gestures;
Output: recognized result m;

function CHILDPTHREAD(W, O)
  for $i \in [0, f-1]$ **do**
    $W(i) \leftarrow O(i)$;
  end for
  while TRUE **do**
    $GS(i) \leftarrow$ **SUM OF ANGULAR VELOCITY**;
    $\Delta T \leftarrow T(j) - T(j-1)$;
    if $GS(i) > GT$ **and** $\Delta T > TT$ **then**
      $T(j-1) = T(j)$;
      return W;
    else
      $W(0 : f-1) \leftarrow W(1 : f)$;
      $W(f-1) \leftarrow O$;
    end if
  end while
end function
function PARENTPTHREAD(W)
  while TRUE **do**
    process the data W with the Butterworth filter;
    compute the variance of the filtered data;
    $\tilde{W} \leftarrow$ **DATA OF FEATURE EXTRACTION**;
    input $\tilde{W}$ into the hand gesture detection classifier;
    if the output of hand gesture detection classifier is *gesture* **then**
      input the $\tilde{W}$ into the hand gesture recognition classifier;
      return $m \leftarrow specific\ gesture$;
    else
      return $m \leftarrow$ *non-gesture*;
    end if
  end while
end function

4. Experimental Results and Analysis

To verify the performance of the proposed asynchronous detection method, many comparative experiments were conducted. For the synchronous detection method, the subjects were required to perform hand gestures according to the screen prompts, which were used as the starting flags as well. The data were recorded for 2 s after the starting points, and then directly fed into the hand gesture recognition classifier. For the asynchronous detection method, the subjects were required to perform hand gestures according to the task requirement. The proposed hand gesture segmentation algorithm automatically detected the starting points according to the threshold value of angular velocity and the time interval of two hand gestures. The data for 2 s after the starting points were captured, to decide whether the current hand gesture was *gesture* or *non-gesture*. When it was detected as *gesture*, the specific category of the current hand gesture was determined by the hand gesture recognition classifier.

4.1. Comparison of Classification Accuracy for Different Classifiers and Hand Gestures

Model selection is of great importance for the performance of hand gesture detection and recognition. The Naive Bayes model is based on the Bayes theorem and the independence assumption of feature condition; therefore, it can be assumed that the variables (IMU data) follow the Gaussian distribution. Consequently, the GNB model for hand gesture

detection was adopted. During the online hand gesture detection, the label probabilities of the input sample were calculated, and the label corresponding to the maximum probability was used as the basis for the differentiation between *gestures* and *non-gestures*. As shown in Table 4, compared with the performance of Random Forest (RF), Support Vector Machine (SVM) [26,27], k-Nearest Neighbor (kNN) [28], and Linear Discriminant Analysis (LDA) [29], the average classification accuracy of 10 hand gestures by the GNB model achieved 98.45 %, leading other methods by 3.15 % on average.

Table 4. The balance accuracy for *gesture* and *non-gesture* by different classifiers. GNB: Gaussian Native Bayes; RF: Random Forest; SVM: Support Vector Machine; KNN: K-Nearest Neighbor; LDA: Linear Discriminant Analysis.

Subject	GNB	RF	SVM	KNN	LDA
Subject 1	100.00	100.00	99.14	100.00	99.14
Subject 2	96.55	90.83	84.91	100.00	55.39
Subject 3	100.00	100.00	100.00	100.00	100.00
Subject 4	99.14	82.50	80.39	99.14	96.55
Subject 5	100.00	100.00	100.00	100.00	100.00
Subject 6	92.24	91.67	72.41	88.58	93.10
Subject 7	100.00	100.00	100.00	100.00	99.14
Subject 8	98.28	100.00	98.28	98.28	97.41
Subject 9	98.28	98.33	93.75	98.28	96.55
Subject 10	100.00	98.33	100.00	100.00	100.00
Avg.	**98.45**	**96.17**	**92.89**	**98.43**	**93.73**

On the other hand, the RF [30] model, which originates from ensemble learning, integrates a number of decision trees into a random forest. There are N classification results (corresponding to N trees) for one input sample. Through a multi-class voting strategy, the category with the most votes is determined as the final recognition result. In addition, the RF model is suitable for small hand gesture datasets, and is adaptable to the noise of IMU data during online experiments. Therefore, the RF [31] model was used for hand gesture recognition. As shown in Table 5, compared with the performance of the SVM [32], kNN, LDA [33], and GNB, the average classification accuracy of 10 hand gestures by the RF model achieved 95.83%, leading other methods by 5.29 % on average.

Table 5. The accuracy for 10 specific hand gestures by different classifiers. RF: Random Forest; SVM: Support Vector Machine; KNN: K-Nearest Neighbor; LDA: Linear Discriminant Analysis; GNB: Gaussian Native Bayes.

Subject	RF	SVM	KNN	LDA	GNB
Subject 1	98.33	93.33	71.67	93.33	88.33
Subject 2	85.00	91.67	91.67	88.33	83.33
Subject 3	100.00	100.00	96.67	96.67	100.00
Subject 4	95.00	86.67	83.33	90.00	90.00
Subject 5	98.33	98.33	98.33	95.00	98.33
Subject 6	93.33	81.67	86.67	86.67	80.00
Subject 7	98.33	95.00	90.00	95.00	95.00
Subject 8	100.00	95.00	88.33	96.67	95.00
Subject 9	96.67	90.00	83.33	78.33	78.33
Subject 10	93.33	93.33	90.00	96.67	91.67
Avg.	**95.83**	**92.50**	**88.00**	**91.67**	**90.00**

The raw IMU dataset of *gesture* and *non-gesture*, collected by 10 subjects, were divided into the training set and test set. The offline classification accuracy of *gesture* and *non-gesture* was compared, based on the GNB model, as shown in Figure 5a. It can be seen that the average hand gesture detection accuracy for *gesture* was 100.00%, and that for *non-gesture* was 96.51% on the test set. Similarly, the confusion matrix of 10 specific hand gestures was obtained. As shown in Figure 5b, the average classification accuracy for 10 kinds of hand gestures (LD, RD, UM, DM, HF, CF, TU, OS, LF, RS) was over 90% on

the test set, except for the hand gesture of RD, which was only 89.39%. Furthermore, the classification accuracy of the hand gesture of RS could reach 100%. The results indicate that the design of the hand gesture set was relatively reasonable, and the adaptability was relatively strong.

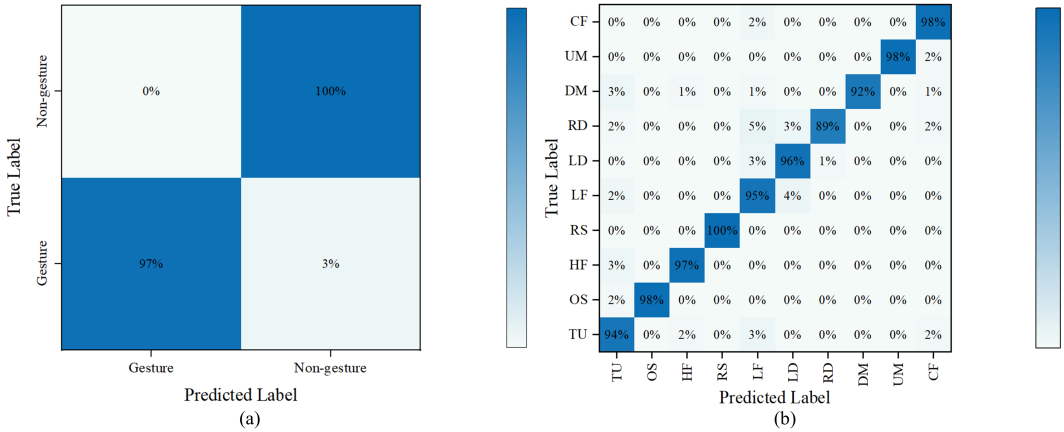


Figure 5. The confusion matrix of hand gesture detection model and hand gesture recognition model: (**a**) the confusion matrix of *gesture* and *non-gesture*; (**b**) the confusion matrix of 10 specific hand gestures (CF, UM, DM, RD, LD, LF, RS, HF, OS, TU).

4.2. Comparison of Online Hand Gesture Recognition Performance under Different Hand Gesture Detection Methods

To verify the performance of the proposed asynchronous detection method, the online recognition accuracy, recognition time, and total time were tested. The recognition time refers to the respective times taken to recognize the hand gesture by the two hand gesture detection methods. The total time (from the execution of hand gestures to the response of the UAV) was calculated, to test the performance of the two hand gesture detection methods respectively in a systematic way.

As shown in Table 6, the average recognition accuracy under the asynchronous hand gesture detection method was 92%, which was higher than that under synchronous hand gesture detection. Moreover, the average recognition time was 7.5 ms, and the total time of one hand gesture was about 3 s of the asynchronous detection method, both of which were shorter than the respective times of the synchronous detection method. This means that the time interval between the publication of the hand gesture command and the UAV's response was about 1 s.

Table 6. The comparison of online recognition accuracy, recognition time, and total time for different hand gesture detection methods.

Detection Method	Subject	Recognition Accuracy (%)	Recognition Time (ms)	Total Time (s)
Syn [1]	Subject 1	87	8.0	3.0488
	Subject 2	87	7.9	3.0533
	Subject 3	100	8.0	3.0485
	Subject 4	63	7.9	3.0485
	Subject 5	80	7.3	3.0119
	Avg.	**83**	**7.8**	**3.0422**
Asyn [2]	Subject 1	83	7.6	3.5267
	Subject 2	77	7.5	2.4560
	Subject 3	100	8.5	2.8597
	Subject 4	100	6.8	3.1599
	Subject 5	100	7.2	2.8525
	Avg.	**92**	**7.5**	**2.9710**

[1] Syn: synchronous hand gesture detection and recognition. [2] Asyn: asynchronous hand gesture detection and recognition.

The differences in average recognition time and total time of different hand gestures under two hand gesture detection methods were compared, as shown in Figure 6. Under the asynchronous detection method, the variations of average recognition time and total time were more dramatic than under the synchronous hand gesture detection method. Paired *t*-test was performed for the two hand gesture detection methods ($p > 0.05$). The results indicated that there were no significant differences in the average recognition time and total time between the two hand gesture detection methods.

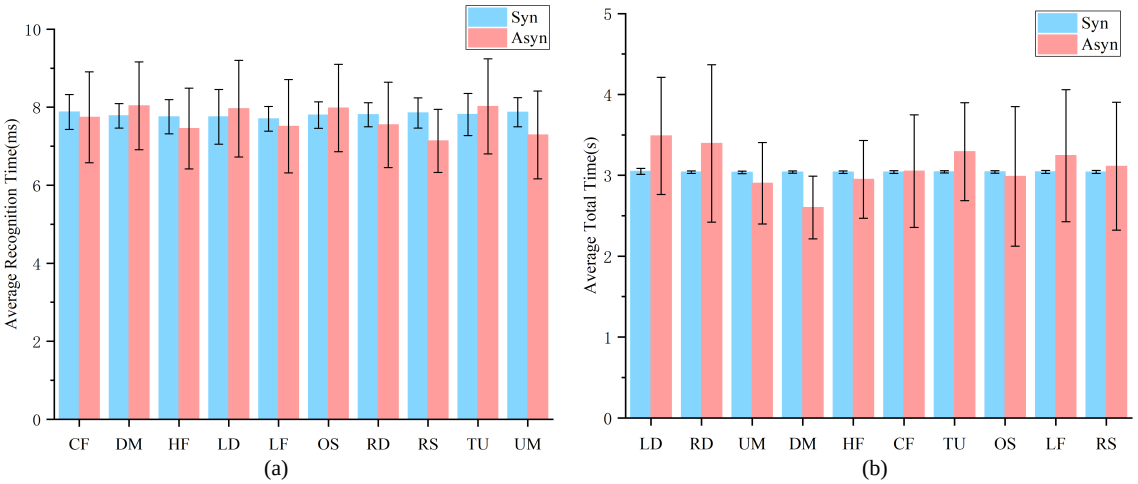


Figure 6. The differences in average recognition time and total time of different hand gestures under the synchronous and asynchronous detection methods: (**a**) the difference in the average recognition time of different hand gestures; (**b**) the difference in the average total time of different hand gestures. Syn: synchronous hand gesture detection and recognition. Asyn: asynchronous hand gesture detection and recognition.

4.3. Comparison of Online Interaction Performance under Different Hand Gesture Detection Methods

To verify the online interaction performance using the asynchronous hand gesture detection and recognition method, UAV flight experiments were conducted in a real-world environment. Users performed different hand gestures to publish the flight direction, and then the mobile workstation would process the IMU data of real-time hand gestures. The recognized result would be transferred into UAV flight waypoint information, and the initial optimal path would be searched by the UAV. During the flight, the UAV would replan and optimize the trajectory in real time, according to the distribution of surrounding obstacles, and would finally fly to the target waypoint. The computing resource included a mobile workstation with a dual-core 2.6 GHz Intel i7-9750H processor, 16 GB RAM, and 1TB SSD. The online hand gesture detection and recognition module was implemented in Python3.8, and the UAV interaction module was implemented in C++11. The interaction experiments were implemented on Ego-Planner, which is an open-source UAV motion planning framework, and were conducted in a closed environment of 5.4 m * 5.4 m * 2.5 m.

Five subjects with no flight experience were invited to participate in our experiment, which was a square loop flight task consisting of 12 navigation points (corresponding to 12 hand gesture commands). The number of errors was recorded for both synchronous and asynchronous interaction experiments. As shown in Figure 7, the 3D scatter figure was plotted to show the number of errors when different hand gesture commands were executed by different subjects. The axial of *Subject* presents the different subjects, and the axial of *Waypoint* presents the take-off point (S) and 11 navigation points (N0, . . . , N10). It manifests that the asynchronous method can effectively reduce the number of errors.

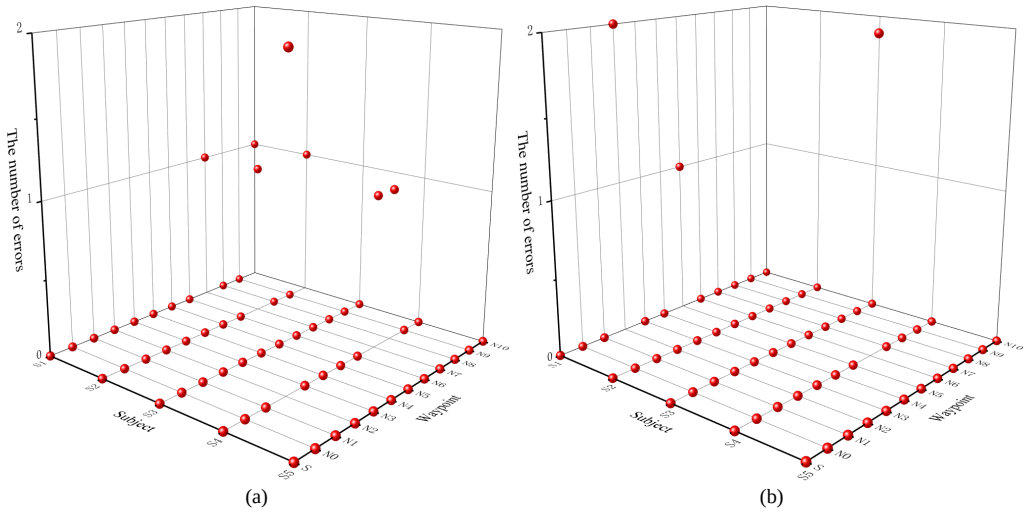


Figure 7. The number of errors in the UAV's complex flight experiments: (**a**) the number of errors in the synchronous interaction experiment; (**b**) the number of errors in the asynchronous interaction experiment.

The time consumed for the square loop flight task is recorded in Figure 8. The time was calculated from the take-off of S to the land-point S, which passed through 11 navigation points (N0, . . . , N10).

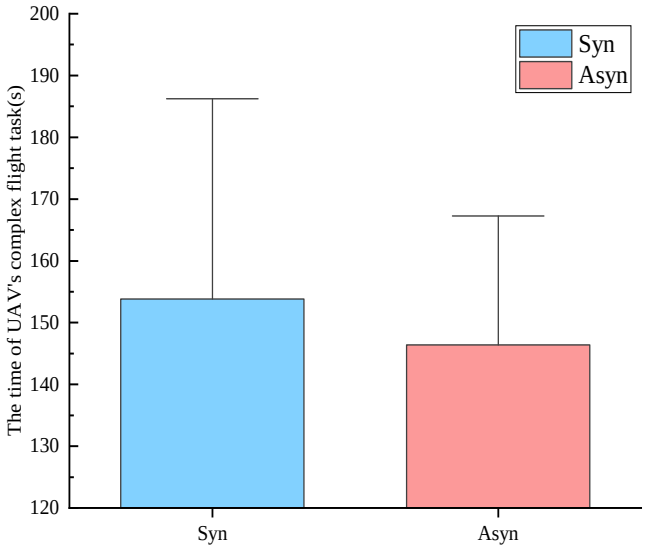


Figure 8. The time of the UAV's complex flight task. Syn: synchronous interaction experiment. Asyn: asynchronous interaction experiment.

In the synchronous interaction experiment, the mean flight time was 153.8 s, which was longer than that of the asynchronous interaction experiment (146.4 s). The standard error (SE) of the asynchronous interaction was smaller, especially in the case of complex flight tasks. This implies that the real-time performance and stability of the asynchronous interaction were better than those of the synchronous interaction.

The efficiency of interaction E of each UAV flight task t_i was defined as follows [34]:

$$E(t_i) = Eval(T, R) = [0.5 \times \frac{R}{100} + (1 - 0.5) \times \frac{1}{T + 1}] \quad (1)$$

where T (the difference between the end time and the start time of the UAV's flight task) was the required time of completing the UAV's flight task, R (the ratio of the number

of successful experiments to the total number of experiments) was the success rate of completing the UAV's flight task, and R and T were given the same weight value of 0.5. $Eval(T, R) \in (0, 1)$ was the evaluation of the interaction performance. As shown in Table 7, the interaction evaluation of the UAV's complex flight task was computed for the five participants. In the square loop flight task, the mean E under the asynchronous interaction was 0.4692, which was a little higher than that of the synchronous interaction (0.4564).

Table 7. The interaction evaluation of the UAV's complex flight task.

Subject	The Efficiency of Synchronous Interaction	The Efficiency of Asynchronous Interaction
Subject 1	0.4323	0.4027
Subject 2	0.4655	0.5037
Subject 3	0.5038	0.5043
Subject 4	0.3771	0.4314
Subject 5	0.5036	0.5039
Avg.	**0.4564**	**0.4692**

Under asynchronous hand gesture detection and recognition, the interaction effect in the real-world experiment is presented in Figure 9. The figure shows that a UAV can complete a complex flight task according to the direction provided by hand gestures, and that discrete hand gesture commands can be successfully transferred into a series of waypoints. The flight trajectory can be constantly replanned and optimized. A relative video has been released to the community (https://www.bilibili.com/video/BV14g411W7xq/, accessed on 1 December 2022). The hand gesture data set and implemented codes used in this work have been released as an open-source package (https://github.com/flytohign/HandGestureEgoPlanner.git).

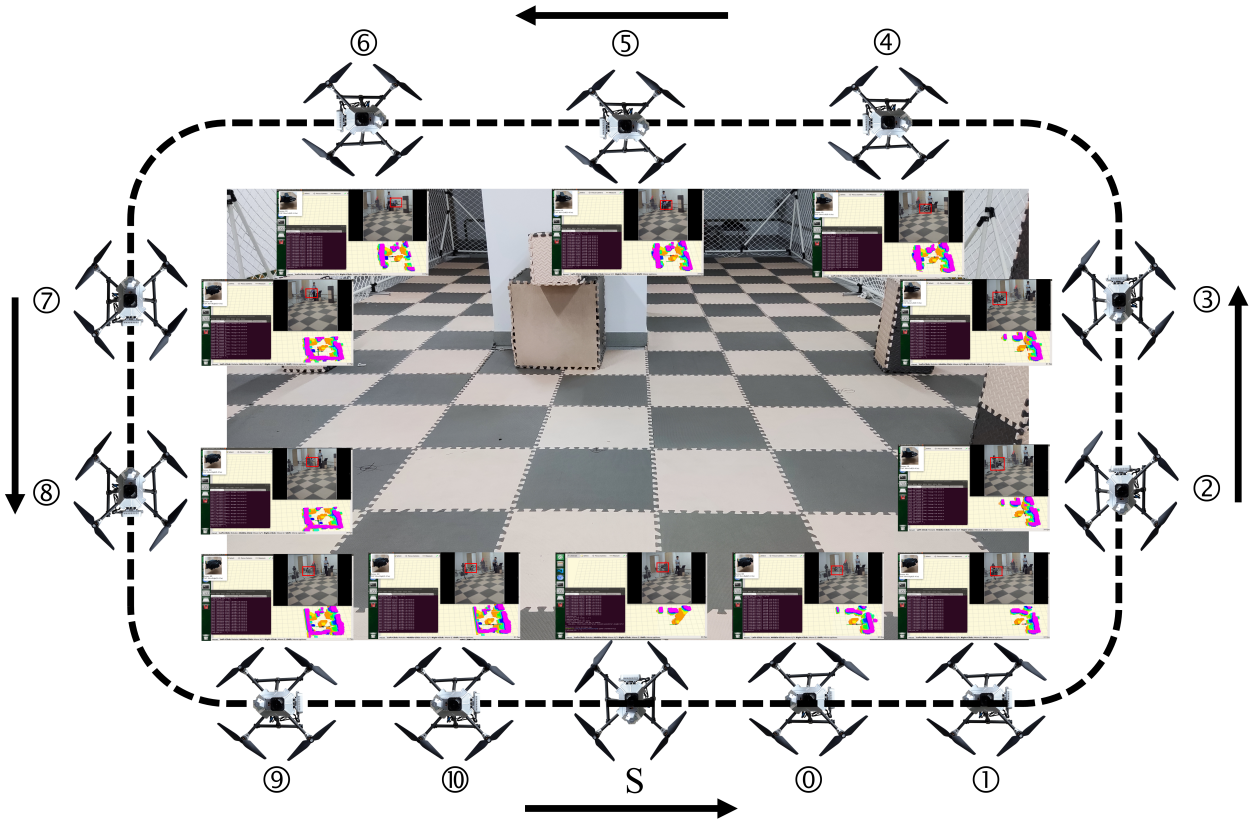


Figure 9. Real-world square loop experiments. The UAV takes off at S, passes through 11 navigation points ⓪, ..., ⑩(N0, ..., N10) and lands at S again. The arrows represent the flight directions. The red boxes represent that the UAV has arrived at the navigation point.

5. Discussion

This study proposes an IMU data glove integrated with six-axis IMU sensors for hand gesture recognition. Ten single hand gestures and four combined hand gestures were designed to interact with a UAV. An asynchronous hand gesture detection and recognition method is proposed, to complete the interaction. In the study, if the current action was recognized as a hand gesture, further recognition would be performed. Meanwhile, the online recognition performance of our proposed method was validated. The results demonstrated the adaptability of the hand gesture set, and showed high online recognition accuracy and real-time performance. In addition, the proposed asynchronous hand gesture detection and recognition method was integrated into the framework of the UAV motion planning. By establishing a wireless data channel between hand gesture and UAV, efficient and robust interaction was realized. Furthermore, in actual scenarios, the results of various complex flight experiments and interaction efficiency evaluation verified the feasibility of the framework.

Detailed comparison of our data glove to previous studies is listed in Table 8. With the increase in the number of IMUs, and the implementation of deep learning, the recognition performance will be significantly improved. For example, Lin et al. [35] proposed a data glove system integrating 27 IMUs for hand function evaluation: the average hit rate was 70.22%, based on the machine learning algorithm. Xu et al. [36] proposed a convolutional neural network (CNN) model on the terminal, which realized 98.79% for 53 hand gestures. The reasons for the equivalent or even better performance of the proposed data glove with fewer IMUs were probably the design of the hand gestures and the application of the cascading structure of the algorithm.

Table 8. Comparison of the proposed data glove to previous ones. MLP: multilayer perceptron; RNN: recurrent neural networks; CNN: convolutional neural network.

Research	[35]	[12]	[13]	[36]	This Work
Components	MSP430, 27 IMUs, Bluetooth	Intel's Edison, 5 IMUs, Bluetooth and WiFi	Arduino Nano 33BLE, 5 IMUs, Bluetooth and USB	STM32F103RCT6, 15 IMUs, Bluetooth	STM32L151CCT, 5 IMUs, USB and Bluetooth
Model	k-means clustering	Naïve Bayes, MLP, RF	RNN	CNN	GNB + RF
Offline average recognition accuracy	70.22% for three tasks from 15 healthy subjects and 15 stroke patients	92% for 22 distinct hand gestures from 57 participants	95% for 8 classes decoding task from 3 subjects	98.79% for 53 hand gestures from 22 subjects	98.45% for gesture and non-gesture; 95.83% for 10 hand gestures from 10 subjects

Table 9 presents the comparison of methods and performance for hand gesture recognition. Previous studies reported using data frame analysis, unsupervised threshold detection, and scalar value estimation for hand gesture segmentation: however, there were disadvantages, such as long computational time, large segmentation delay, and long recognition time, and it was difficult to balance between real-time performance and recognition accuracy. Although previous studies applied hand gesture segmentation based on threshold detection, most of them considered the difference in amplitude variation of IMU signals, and ignored the impact of the threshold setting on segmentation performance. In addition, in the current study, a certain interval of inactivity was considered in threshold detection, which contributed to better real-time performance and higher online recognition accuracy, along with the cascading structure of algorithm. A balance between real-time and recognition accuracy was achieved, to ensure that the proposed system was efficient and robust.

Table 9. Comparison of methods and performance. ANN: artificial neural network; HMM: hidden Markov model; LSTM: long short-term memory.

Research	[16]	[17]	[18]	[19]	**This Work**
hand gesture segmentation	analyze each frame from the glove sensors	unsupervised threshold-based segmentation	based on Dempster–Shafer theory	detect the start/end of a hand gesture sequence by estimating a scalar value	threshold detection and interval of inactivity
hand gesture recognition	two ANNs in series		HMM	based on two LSTM layers	GNB cascading with RF
real-time performance	computational time is about 9 min for 10 hand gestures	average segmentation delay is 263 ms	based on the evidence reasoning, the delay between spotting and recognition is eliminated	no more than 12 ms to recognize the completed hand gesture in real-time	about 7.8 ms to recognize the completed hand gesture
segmentation or recognition accuracy	over 99% for a library of 10 gestures and over 96% for a library of 30 gestures	segmentation accuracy can rise to 100% at a window size of 24 frames and average oversegmentation error is 2.70%	the recognition accuracy (95.2%) after spotting is higher than the accuracy (96.7%) of simultaneous recognition with spotting	offline recognition accuracy is 100%	online recognition accuracy up to 92%

Previous studies have generally focused on basic UAV flight commands based on hand gesture recognition, and the collision avoidance problems have not been well-considered. The comparison of hand gesture commands and UAV flight actions is shown in Table 10. Hu et al. [37] developed a real-time dynamic hand gesture recognition system for UAV flight control, and a Leap Motion Controller to collect the skeleton data of dynamic hand gestures. Based on the 8-layer convolutional neural network, an average accuracy of 96.9% was achieved on non-scaled datasets. Yu et al. [22] designed basic control commands and mode-switching commands for UAV control. Wireless data gloves incorporating flex sensors and inertial sensors were used to build hand gesture datasets. The average recognition accuracy of 10 static hand gestures and 5 dynamic hand gestures was 100% and 98.4%, respectively. Although research progress has been made regarding the interaction of hand gestures and UAV flights, most dynamic hand gestures are essentially a combination of static hand gestures and transition processes, which affects robustness and real-time performance. Therefore, the designed hand gesture set takes full account of hand features, and is reasonably mapped to UAV flight motions. Single hand gestures of high accuracy are combined, to ensure stability of interaction. Additionally, various studies have only used hand gestures to avoid obstacles during UAV flights, rather than to achieve collision avoidance in complex environments by integrating a motion planning framework.

Undoubtedly, limitations exist in the current study. The IMU-based data glove communicates with the remote processing terminal via wired links, while the terminal communicates with the UAV wirelessly within a Local Area Network (LAN). The communication mode between the glove and UAV is expected to be improved, and outdoor scenarios will be developed for more applications.

Table 10. Comparison of hand gesture commands and UAV flight actions.

Research	[20]	[21]	[37]	[22]	This Work
hand gesture command	forward, backward, left, right, ascent, descent, hovering, rotating clockwise, rotating anticlockwise	take off, land, height down, hover, height up, pilot	move forward, move backward, turn left, turn right, move up, move down, turn clockwise, turn anticlockwise, special movement 1, special movement 2	basic commands: throttle up, throttle down, pitch down, pitch up, roll left, roll right, yaw left, yaw right, flag, no command; mode switching commands: arm, disarm, position mode, hold mode, return mode	move left, move right, move up, move down, wait for combination, disarm, take off on high, hover, forced land, move forward, move backward, turn left, turn right
UAV flight	basic action flight	basic action flight	basic action flight	simple flight mission, artificial collision avoidance	complex flight task, automatic collision avoidance

6. Conclusions and Future Work

In this paper, a novel hand gesture detection and recognition method was proposed, based on the IMU data glove for UAV motion planning. The cascaded classifier was designed to recognize the generated hand gestures asynchronously, through which, higher online recognition accuracy, shorter online recognition time, and the shorter total time of one hand gesture were obtained. Afterwards, the online gesture detection and recognition module was integrated with the open-source UAV motion planning module, and extensive real-world experiments of complex flight tasks were conducted. The real-time performance and stability of the asynchronous interaction were better than those of the synchronous interaction. The experimental results show that the proposed method can support UAV motion planning efficiently and robustly.

In future work, in-depth research will be conducted on the following three aspects: (i) to adapt to the requirements of different scenarios, cross-scene hand gesture recognition can be proposed with the transfer learning method, and the cost of data acquisition will be reduced; (ii) to realize the fine-grained finger gesture recognition onboard, the construction of a hand gesture network structure could be considered, which would abstract the relative position between different fingers into '0'/'1' (char type) values, with each hand gesture corresponding to a unique '0'/'1' matrix; (iii) to solve complex collision avoidance problems, hand gesture interaction could be integrated into every control point of UAV motion planning. Based on the task learning method, the trajectory of the UAV flight would be modified by the hand gesture interaction, to adapt to unknown and complex environments.

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

Funding: This work was supported by Science and Technology Innovation 2030 major projects 2022ZD0208900 and the National Natural Science Foundation of China under Grant 62176090, Grant 62076250, Grant 61703407, and Grant 61901505; by the Shanghai Municipal Science and Technology Major Project under Grant 2021SHZDZX; by the Program of Introducing Talents of Discipline to Universities through the 111 Project under Grant B17017; by the ShuGuang Project supported by the Shanghai Municipal Education Commission and the Shanghai Education Development Foundation under Grant 19SG25; by the Ministry of Education and Science of the Russian Federation under Grant 14.756.31.0001; and by the Polish National Science Center under Grant UMO-2016/20/W/NZ4/00354. This research was also supported by National Government Guided Special Funds for Local Science and Technology Development (Shenzhen, China) (No. 2021Szvup043) and by the Project of Jiangsu Province Science and Technology Plan Special Fund in 2022 (Key Research and Development Plan Industry Foresight and Key Core Technologies) under Grant BE2022064-1.

Institutional Review Board Statement: Ethical review and approval were waived for this study, due to hand gesture data collection, which does not need approval from an ethical committee.

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

Data Availability Statement: The original contributions presented in the study are included in the article; further inquiries can be directed to the corresponding author.

Acknowledgments: The authors would like to thank the users who took part in the study, and the devices provided by Tianjin Artificial Intelligence Innovation Center.

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

References

1. Oneata, D.; Cucu, H. Kite: Automatic speech recognition for unmanned aerial vehicles. *arXiv* **2019**, arXiv:1907.01195.
2. Smolyanskiy, N.; Gonzalez-Franco, M. Stereoscopic first person view system for drone navigation. *Front. Robot. AI* **2017**, *4*, 11. [CrossRef]
3. How, D.N.T.; Ibrahim, W.Z.F.B.W.; Sahari, K.S.M. A Dataglove Hardware Design and Real-Time Sign Gesture Interpretation. In Proceedings of the 2018 Joint 10th International Conference on Soft Computing and Intelligent Systems (SCIS) and 19th International Symposium on Advanced Intelligent Systems (ISIS), Toyama, Japan, 5–8 December 2018; pp. 946–949.
4. Ilyina, I.A.; Eltikova, E.A.; Uvarova, K.A.; Chelysheva, S.D. Metaverse-Death to Offline Communication or Empowerment of Interaction? In Proceedings of the 2022 Communication Strategies in Digital Society Seminar (ComSDS), Saint Petersburg, Russia, 13 April 2022; pp. 117–119.
5. Serpiva, V.; Karmanova, E.; Fedoseev, A.; Perminov, S.; Tsetserukou, D. DronePaint: Swarm Light Painting with DNN-based Gesture Recognition. In Proceedings of the ACM SIGGRAPH 2021 Emerging Technologies, Virtual Event, USA, 9–13 August 2021; pp. 1–4.
6. Liu, C.; Szirányi, T. Real-time human detection and gesture recognition for on-board uav rescue. *Sensors* **2021**, *21*, 2180. [CrossRef] [PubMed]
7. Lu, Z.; Chen, X.; Li, Q.; Zhang, X.; Zhou, P. A hand gesture recognition framework and wearable gesture-based interaction prototype for mobile devices. *IEEE Trans. Hum. Mach. Syst.* **2014**, *44*, 293–299. [CrossRef]
8. Zhou, S.; Zhang, G.; Chung, R.; Liou, J.Y.; Li, W.J. Real-time hand-writing tracking and recognition by integrated micro motion and vision sensors platform. In Proceedings of the 2012 IEEE International Conference on Robotics and Biomimetics (ROBIO), Guangzhou, China, 11–14 December 2012; pp. 1747–1752.
9. Yang, K.; Zhang, Z. Real-time pattern recognition for hand gesture based on ANN and surface EMG. In Proceedings of the 2019 IEEE 8th Joint International Information Technology and Artificial Intelligence Conference (ITAIC), Chongqing, China, 24–26 May 2019; pp. 799–802.
10. Krisandria, K.N.; Dewantara, B.S.B.; Pramadihanto, D. HOG-based Hand Gesture Recognition Using Kinect. In Proceedings of the 2019 International Electronics Symposium (IES), Surabaya, Indonesia, 27–28 September 2019; pp. 254–259.
11. Li, J.; Liu, X.; Wang, Z.; Zhang, T.; Qiu, S.; Zhao, H.; Zhou, X.; Cai, H.; Ni, R.; Cangelosi, A. Real-Time Hand Gesture Tracking for Human–Computer Interface Based on Multi-Sensor Data Fusion. *IEEE Sens. J.* **2021**, *21*, 26642–26654. [CrossRef]
12. Mummadi, C.K.; Philips Peter Leo, F.; Deep Verma, K.; Kasireddy, S.; Scholl, P.M.; Kempfle, J.; Van Laerhoven, K. Real-time and embedded detection of hand gestures with an IMU-based glove. *Proc. Inform.* **2018**, *5*, 28. [CrossRef]
13. Makaussov, O.; Krassavin, M.; Zhabinets, M.; Fazli, S. A low-cost, IMU-based real-time on device gesture recognition glove. In Proceedings of the 2020 IEEE International Conference on Systems, Man, and Cybernetics (SMC), Toronto, ON, Canada, 11–14 October 2020; pp. 3346–3351.
14. Zhang, X.; Yang, Z.; Chen, T.; Chen, D.; Huang, M.C. Cooperative sensing and wearable computing for sequential hand gesture recognition. *IEEE Sens. J.* **2019**, *19*, 5775–5783. [CrossRef]
15. Jiang, S.; Lv, B.; Guo, W.; Zhang, C.; Wang, H.; Sheng, X.; Shull, P.B. Feasibility of wrist-worn, real-time hand, and surface gesture recognition via sEMG and IMU sensing. *IEEE Trans. Ind. Inform.* **2017**, *14*, 3376–3385. [CrossRef]
16. Neto, P.; Pereira, D.; Pires, J.N.; Moreira, A.P. Real-time and continuous hand gesture spotting: An approach based on artificial neural networks. In Proceedings of the 2013 IEEE International Conference on Robotics and Automation, Karlsruhe, Germany, 6–10 May 2013; pp. 178–183.
17. Simão, M.A.; Neto, P.; Gibaru, O. Unsupervised gesture segmentation by motion detection of a real-time data stream. *IEEE Trans. Ind. Inform.* **2016**, *13*, 473–481. [CrossRef]
18. Li, Q.; Huang, C.; Yao, Z.; Chen, Y.; Ma, L. Continuous dynamic gesture spotting algorithm based on Dempster–Shafer Theory in the augmented reality human computer interaction. *Int. J. Med. Robot. Comput. Assist. Surg.* **2018**, *14*, e1931. [CrossRef] [PubMed]
19. Lee, M.; Bae, J. Deep learning based real-time recognition of dynamic finger gestures using a data glove. *IEEE Access* **2020**, *8*, 219923–219933. [CrossRef]
20. Choi, Y.; Hwang, I.; Oh, S. Wearable gesture control of agile micro quadrotors. In Proceedings of the 2017 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI), Daegu, Republic of Korea, 16–18 November 2017; pp. 266–271.

21. Yu, Y.; Wang, X.; Zhong, Z.; Zhang, Y. ROS-based UAV control using hand gesture recognition. In Proceedings of the 2017 29th Chinese Control And Decision Conference (CCDC), Chongqing, China, 28–30 May 2017; pp. 6795–6799.
22. Yu, C.; Fan, S.; Liu, Y.; Shu, Y. End-Side Gesture Recognition Method for UAV Control. *IEEE Sens. J.* **2022**, *22*, 24526–24540. [CrossRef]
23. Zhou, B.; Gao, F.; Wang, L.; Liu, C.; Shen, S. Robust and efficient quadrotor trajectory generation for fast autonomous flight. *IEEE Robot. Autom. Lett.* **2019**, *4*, 3529–3536. [CrossRef]
24. Zhou, X.; Wang, Z.; Ye, H.; Xu, C.; Gao, F. Ego-planner: An esdf-free gradient-based local planner for quadrotors. *IEEE Robot. Autom. Lett.* **2020**, *6*, 478–485. [CrossRef]
25. Han, H.; Yoon, S.W. Gyroscope-based continuous human hand gesture recognition for multi-modal wearable input device for human machine interaction. *Sensors* **2019**, *19*, 2562. [CrossRef] [PubMed]
26. Chen, W.; Zhang, Z. Hand gesture recognition using sEMG signals based on support vector machine. In Proceedings of the 2019 IEEE 8th Joint International Information Technology and Artificial Intelligence Conference (ITAIC), Chongqing, China, 24–26 May 2019; pp. 230–234.
27. Chen, Y.; Luo, B.; Chen, Y.L.; Liang, G.; Wu, X. A real-time dynamic hand gesture recognition system using kinect sensor. In Proceedings of the 2015 IEEE International Conference on Robotics and Biomimetics (ROBIO), Zhuhai, China, 6–9 December 2015; pp. 2026–2030.
28. Lian, K.Y.; Chiu, C.C.; Hong, Y.J.; Sung, W.T. Wearable armband for real time hand gesture recognition. In Proceedings of the 2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC), Banff, AB, Canada, 5–8 October 2017; pp. 2992–2995.
29. Wang, N.; Chen, Y.; Zhang, X. The recognition of multi-finger prehensile postures using LDA. *Biomed. Signal Process. Control* **2013**, *8*, 706–712. [CrossRef]
30. Joshi, A.; Monnier, C.; Betke, M.; Sclaroff, S. A random forest approach to segmenting and classifying gestures. In Proceedings of the 2015 11Th IEEE International Conference And Workshops on Automatic Face and Gesture Recognition (FG), Ljubljana, Slovenia, 4–8 May 2015; Volume 1, pp. 1–7.
31. Jia, R.; Yang, L.; Li, Y.; Xin, Z. Gestures recognition of sEMG signal based on Random Forest. In Proceedings of the 2021 IEEE 16th Conference on Industrial Electronics and Applications (ICIEA), Chengdu, China, 1–4 August 2021; pp. 1673–1678.
32. Liang, X.; Ghannam, R.; Heidari, H. Wrist-worn gesture sensing with wearable intelligence. *IEEE Sens. J.* **2018**, *19*, 1082–1090. [CrossRef]
33. Dellacasa Bellingegni, A.; Gruppioni, E.; Colazzo, G.; Davalli, A.; Sacchetti, R.; Guglielmelli, E.; Zollo, L. NLR, MLP, SVM, and LDA: A comparative analysis on EMG data from people with trans-radial amputation. *J. Neuroeng. Rehabil.* **2017**, *14*, 1–16. [CrossRef] [PubMed]
34. Zhang, S.; Liu, X.; Yu, J.; Zhang, L.; Zhou, X. Research on Multi-modal Interactive Control for Quadrotor UAV. In Proceedings of the 2019 IEEE 16th International Conference on Networking, Sensing and Control (ICNSC), Banff, AB, Canada, 9–11 May 2019; pp. 329–334.
35. Lin, B.S.; Hsiao, P.C.; Yang, S.Y.; Su, C.S.; Lee, I.J. Data glove system embedded with inertial measurement units for hand function evaluation in stroke patients. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2017**, *25*, 2204–2213. [CrossRef] [PubMed]
36. Xu, P.F.; Liu, Z.X.; Li, F.; Wang, H.P. A Low-Cost Wearable Hand Gesture Detecting System Based on IMU and Convolutional Neural Network. In Proceedings of the 2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), Virtual, 1–5 November 2021; pp. 6999–7002.
37. Hu, B.; Wang, J. Deep learning based hand gesture recognition and UAV flight controls. *Int. J. Autom. Comput.* **2020**, *17*, 17–29. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.