Passive User Authentication Utilizing Screen-Touch Trajectory for Unmanned Aerial Vehicle Systems

Guozhu Zhao[1], Jiasheng He[1,*], Yaxin Wang[1], Zixian Du[2], Xiaolan Liu[3], Chenxi Yang[1], Yichen Li[1], and Renzhi Ji[1]
[1]School of Computer and Information Engineering, Chuzhou University, Chuzhou, Anhui, 239000, China
[2]School of Intelligent Engineering, Xi'an Jiaotong-Liverpool University, Suzhou, Jiangsu, 215123, China
[3]School of Basic Medicine, Shandong Second Medical University, Weifang, Shandong, 261053, China
[*]Corresponding author

User identity verification is of great significance for the secure operation of UAV systems. This article exploits the spatial-temporal features of user screen-touch trajectory (STT) to develop a passive user authentication framework for unmanned aerial vehicle (UAV) systems. We first design a multi-dimensional STT data structure to accurately record continuous touch interactions and extract 21 discriminative spatial-temporal features for behavioral characterization. We then construct a classifier based on Random Forest (RF) for robust nonlinear feature learning and also a classifier based on Radial Basis Function Neural Network (RBFNN) for modeling fine-grained touch behavior patterns. By combining the two classifiers and assigning each classifier an appropriate weight, we develop a fusion-based passive user authentication framework. The new framework has the potential to significantly enhance the security of UAV operation systems by offering a flexible, continuous, and non-intrusive authentication solution, and it also can serve as a complementary security layer or an enhancement to traditional authentication mechanisms in UAV ground control systems.

***Index Terms*—Unmanned Aerial Vehicle Security, Passive Authentication, Screen-touch Trajectory, Spatial-temporal Features, Machine Learning.**

I. Introduction

UNMANNED Aerial Vehicle (UAV) technology represents a new generation of aerospace operational platforms, scenario-driven application solutions, and cross-domain industrial ecosystems that deeply integrate information and communication technologies with modern industrial development. Emerging UAV technologies are expected to establish a new paradigm of aerial operations and multi-scenario service systems, enabling comprehensive support and coordinated interaction across domains such as agricultural crop protection, logistics and transportation, energy infrastructure inspection, emergency response, urban governance, and geological surveying [8], [12]. The deep convergence of UAV systems with 5G communications will further accelerate transformative upgrades across vertical industries—expanding application scenarios, enhancing real-time responsiveness and operational precision, reducing overall operational costs, and strengthening large-scale collaborative management capabilities.

At present, UAV technology has been widely adopted in critical domains such as agricultural crop protection, logistics transportation, energy infrastructure inspection, emergency response, urban governance, geological surveying, environmental monitoring, forest fire prevention, marine mapping, bridge inspection, security patrols, aerial cinematography, meteorological sensing, border control, cold-chain delivery, and structural inspection [8], [13]. In these mission-critical UAV application scenarios, the Ground Control Station (GCS) is typically required to store sensitive information—including flight parameters, mission planning data, confidential operational commands, and device access credentials—while also interfacing with sensitive cloud-based scheduling platforms and classified mission databases.

With the rapid proliferation of operational terminals and the increasing integration of UAVs across diverse industry systems, precise verification of operator identity and access privileges has become increasingly essential. Such verification is critical for ensuring strong protection of classified operational data and core industry information within the broader UAV application ecosystem [15]. Consequently, secure and reliable operator access control is of paramount significance for maintaining the safe and trustworthy operation of UAV systems.

Identity and access verification within UAV operation systems serves as an indispensable security service for validating the legitimacy of UAV operators and their associated privilege levels. Modern UAV operation systems typically consist of large numbers of UAV devices, ground control terminals, and cross-industry access nodes deployed across heterogeneous and multi-scenario working environments. This diversity introduces substantial complexity and challenges to the large-scale deployment and governance of identity and access control mechanisms [9]. Passive authentication verifies user identity by analyzing behavioral characteristics exhibited during the operational process, making it particularly suitable for scenarios requiring continuous authentication. Compared with traditional active authentication methods, passive authentication integrates seamlessly into the user's workflow without interrupting the operational process. This property makes it an ideal approach for achieving continuous authentication in UAV missions.

It is worth noting that in the complex operational scenarios of UAV applications, passive authentication technologies play a critical role in enabling continuous and non-intrusive verifi-

Manuscript received September 9, 2025; revised December 13, 2025.
Corresponding author: Jiasheng He (email: 1791732955@qq.com).

cation of operator identity and device privileges. First, active authentication requires the operator's explicit participation, which is difficult to enforce in large-scale, multi-scenario UAV operations or remote-control environments. Second, to ensure flight security and mission accuracy, operators must continuously perform demanding tasks—such as flight control, data acquisition, and emergency response—leaving little opportunity for frequent additional active verification procedures [14], [4], [3]. Furthermore, UAV operations continuously generate large volumes of sensitive mission data, necessitating periodic verification of both operator identity and task legitimacy to strengthen access control over sensitive information [5]. Motivated by these considerations, we aim to design a flexible and cost-effective passive authentication scheme that enables continuous and non-intrusive identity and mission verification within UAV systems.

Observations show that different operators exhibit stable and distinctive touch interaction trajectories, shaped by their operational preferences and long-term muscle memory [7], [17], [2]. In UAV operation scenarios, operators naturally leave interaction traces on the touchscreen interface. To investigate the practical differences in touch behaviors among operators, we designed a controlled experiment. Four participants were instructed to perform the same UAV inspection task, with the flight route, mission objectives, testing environment, and operating device fully standardized. A fixed 10-step sampling strategy was used to capture the STT. Fig. 1 presents the visualized trajectories of the four participants, where each color represents one continuous touch gesture without finger lift. The visual results demonstrate that despite identical task conditions, the operators exhibit significant variations in both spatial distribution features (e.g., trajectory span) and dynamic interaction characteristics (e.g., turning frequency of trajectories).

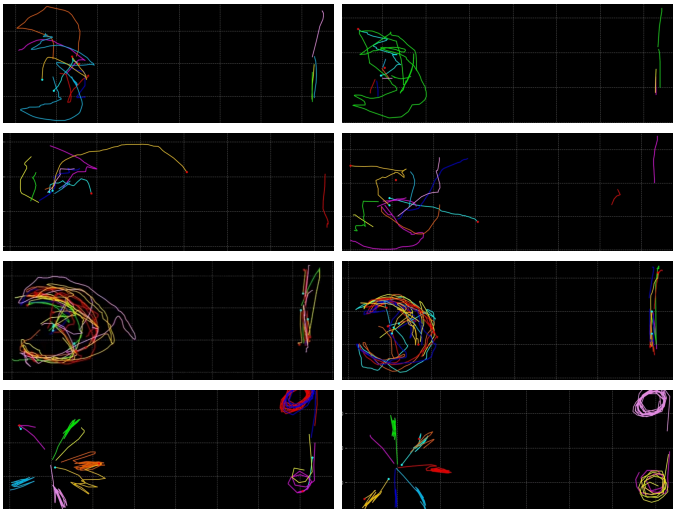

Fig. 1. Differences in STT among four users under the same UAV operation task.

This experiment confirms that touch interaction patterns show clear individual uniqueness across operators and can serve as a critical basis for identity verification and differentiation, further validating the feasibility of the proposed approach.

Existing passive authentication techniques are primarily based on spatial or temporal features. However, in UAV touch operation scenarios, as the complexity of UAV mission tasks increases, these two approaches become increasingly inadequate in accurately and deeply describing operator behaviors [1], [10], [6]. This results in a significant reduction in the ability to distinguish operational features, making it difficult to meet the stringent identity verification and differentiation requirements. While these methods represent major breakthroughs in passive authentication, relying solely on spatial or temporal features is unlikely to effectively address the practical challenges posed by more complex and dynamic UAV operation environments [18]. For example, factors such as wind interference in outdoor environments, body vibration during flight, sudden changes in operation pace under emergency conditions, operator fatigue from long-duration tasks, and temperature-induced sensitivity variations in touchscreen can all disrupt the spatial distribution and temporal stability of touch trajectories. These disturbances may directly compromise the distinguishability of an operator's touch characteristics, resulting in degraded performance of passive authentication systems in UAV operations.

However, the results of this study indicate that by jointly leveraging the spatial-temporal features of STT, we are able to not only provide a complete spatial-temporal representation of operator identity but also significantly improve the performance of passive user authentication in UAV mission scenarios. The key contributions of this work are summarized as follows:

1) We provide extensive experimental results demonstrating that, in UAV operational scenarios, an operator's continuous touchs behavior can be represented as both spatial and temporal biometric features. These spatial-temporal features exhibit good stability for a specific operator and show high distinguishability across different operators.

2) We propose a novel method for describing the spatial-temporal characteristics of user-related touch behavior. Specifically, we define a new data structure that supports the recording of multi-finger touch data, which allows for the complete storage and description of touch operation behavior. We then compute 21 derived features, including position, velocity, direction, rate of change, time, and multi-touch correlation features, forming a touch feature array (TFA) to characterize the spatial-temporal properties of touch behavior.

3) We further utilize two machine learning algorithms—Random Forest (RF) and Radial Basis Function Neural Network (RBFNN)—to learn the TFA, forming two classifiers. By assigning appropriate weights to each classifier, we developed a new passive authentication framework suitable for UAV operational systems.

II. Proposed Authentication Framework

A. Overview of Our Approach

In this section, we design a new identity verification method based on the spatial-temporal features of STT, as illustrated in Fig. 2. First, STT data are collected from UAV control operations and preprocessed to suppress noise. A moving

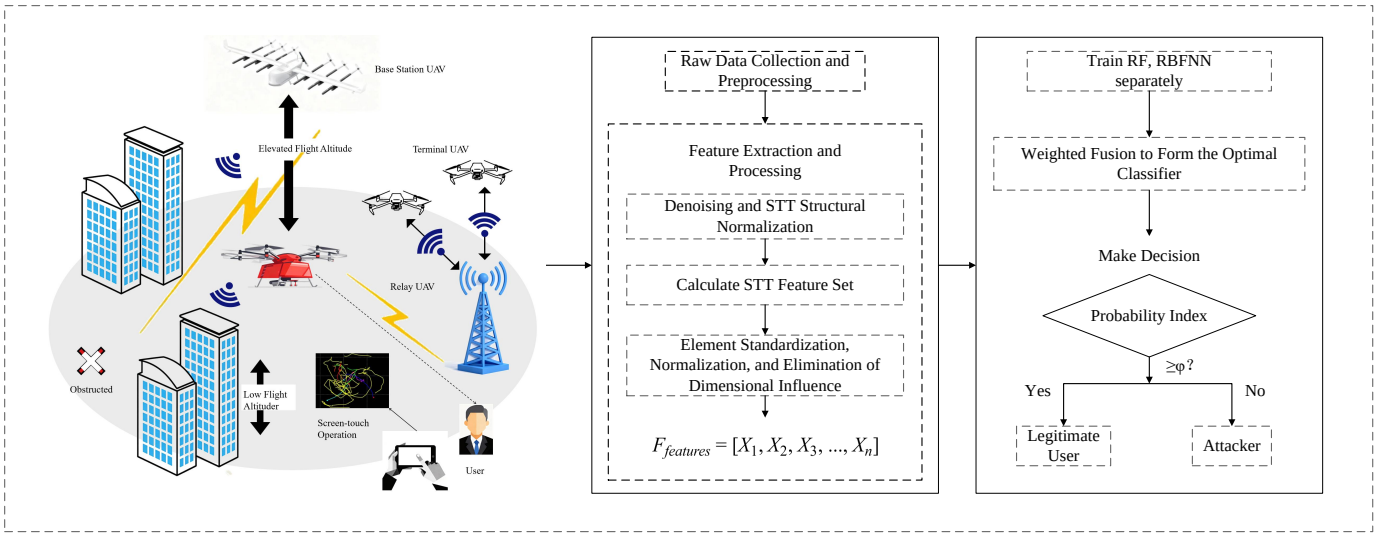

Fig. 2. Workflow of the proposed authentication approach for UAV systems.

average filter is employed to smooth raw trajectories while preserving their essential temporal characteristics. Second, 21-dimensional spatial-temporal features are extracted from the denoised STT signals and organized into a unified feature set. To eliminate scale inconsistencies and dimensional bias, the extracted features are further standardized and normalized prior to model training. Third, the processed feature set is fed into the classification module. Two machine learning models, namely RF and RBFNN, are adopted to independently learn the STT feature representations and construct two binary classifiers. Each classifier is trained exclusively on positive samples from legitimate users, whereas all other samples are implicitly treated as negative. This one-class–oriented binary classification strategy effectively addresses the challenge of inexhaustible negative samples and significantly reduces the reliance on large-scale labeled datasets. Finally, the outputs of the trained classifiers are combined through weighted fusion to construct an optimal ensemble classifier. During inference, the ensemble classifier evaluates incoming STT samples and produces a detection probability. If the probability exceeds a predefined threshold, the user is authenticated as legitimate; otherwise, the operation is identified as an illegal intrusion. In such cases, the current control permissions are immediately revoked, and active authentication or additional identity verification mechanisms are triggered.

B. Data Acquisition Module and SCTO Modeling

We define one complete touch operation as: the process from the first finger touch screen to all finger contacts leaving is recorded as a single complete touch operation (SCTO), which will be used for feature calculation and machine learning in subsequent studies. In order to comprehensively capture touch data and model the SCTO, we designed a new continuous touch data storage structure, which is used to record touch data and fundamental elements of feature calculation, and can support multi touch recording of multi finger operation, calculate the correlation characteristics between multi contacts, and realize accurate digital storage and description of SCTO. This data structure records the coordinate set (X, Y), instantaneous speed set (V), instantaneous direction set (D), and corresponding time frame set (T) for each touch point at the specified sampling frequency. These sets represent the data collection of the SCTO accurate to time frames. The spatial-temporal features of each SCTO will be calculated based on these touch data tuples.

The higher the sampling frequency, such as once every 0.01s, the more precise the data description is. Meanwhile, this will also increase memory usage for the same time sample, which can be adjusted according to the specific needs in practical applications. In this design, a sampling frequency of 0.1s is used, and the size of a single SCTO sample is only 500KB to 1MB, which is lightweight, and can achieve good classification and identity verification performance. Additionally, the required training dataset size is only 200MB to 400MB, consisting of 400 SCTO samples, which is highly acceptable.

Based on the SCTO sample data, highly discriminative spatial-temporal features can be calculated. As shown in Table 1, we calculate 21 types of SCTO features. Depending on the actual performance, all or a subset of these features are selected to form the final feature arrays of the SCTO samples. These arrays are then input into machine learning algorithms to form the classifiers, which are subsequently evaluated in terms of accuracy, detection time, and other performance metrics.

C. Machine Learning Algorithms and Classifier Module

Multi identity classification in this scenario will have the challenge that negative samples can not exhaust all the operation modes of potential attackers. To address this limitation, this study adopts binary classification framework, where legitimate users serve as the sole positive sample, and other touch operations deviating from this sample's features are classified as negative samples. This approach significantly reduces the sample collection scale, eliminating the need to collect extensive data from illegal users for classifier training. It avoids

TABLE I
DESCRIPTION OF SCTO FEATURES.

Feature Category	Feature Name	Description
Position Features	Average X Coordinate	The mean position of all points in the X direction, reflecting the horizontal concentration trend.
	X Coordinate Standard Deviation	The dispersion of X-direction positions, with a larger value indicating a more scattered distribution.
	X Coordinate Range	The maximum span of X-direction positions, reflecting the horizontal operation range.
	Average Y Coordinate	The mean position of all points in the Y direction, reflecting the vertical concentration trend.
	Y Coordinate Standard Deviation	The dispersion of Y-direction positions, with a larger value indicating a more scattered distribution.
	Y Coordinate Range	The maximum span of Y-direction positions, reflecting the vertical operation range.
	Total Movement Distance	The total length of the touchpoint trajectory, reflecting the spatial span of the operation.
	Average Step Length	The average displacement between adjacent frames, reflecting the average distance of a single movement.
Speed Features	Average Speed	The ratio of total movement distance to duration, reflecting the overall speed of the operation.
	Speed Standard Deviation	The standard deviation of speed at each moment, reflecting the degree of speed fluctuation; a smaller value indicates more stable speed.
	Average Speed Change Rate	The average absolute change in speed between adjacent moments, reflecting the average magnitude of speed increase or decrease.
	Speed Range	The difference between the maximum and minimum speed, reflecting the extreme fluctuations in speed.
	Maximum Speed	The highest speed during the operation, reflecting the speed upper bound of the operation.
Direction Features	Average Direction	The average direction of all moments, reflecting the main orientation of the operation.
	Average Direction Change Rate	The average absolute change in direction angle between adjacent moments, reflecting the average magnitude of direction change.
	Direction Change Standard Deviation	The standard deviation of direction angle changes, reflecting the stability of direction change; a smaller value indicates more stable direction.
Change Rate Features	Average X Direction Position Change	The mean change in X-coordinate between adjacent points, reflecting the average trend of X-direction position change (positive for rightward movement, negative for leftward movement).
	Average Y Direction Position Change	The mean change in Y-coordinate between adjacent points, reflecting the average trend of Y-direction position change (positive for upward movement, negative for downward movement).
Multi-Touch Features	Multi-Touch Start Time Difference	The time difference between the start times of multiple touchpoints, reflecting the synchronization of multi-touch activation.
	Relative Position of Two Touchpoints	The coordinate difference between any two touchpoints, reflecting the spatial distribution relationship of multi-touch points.
Time Features	Duration	The total time from the start to the end of a touchpoint, reflecting the time span of the operation.

dependence on negative samples, effectively enhancing the system's universality and resistance to interference in multi-touch and multi-scenario environments.

We trained multiple machine learning frameworks using the SCTO feature arrays and found that the two with the best comprehensive performance in this scenario were Random Forest (RF) and Radial Basis Function Neural Network (RBFNN), forming two classifiers. To further enhance authentication performance, we assign appropriate weights to each classifier and form the weighted fusion classifier [18], [16].

In order to explore the optimal judgement performance of this binary classification framework, we conducted performance evaluations of each classifier in a predefined multi-touch operation scenario. We use classification performance and time efficiency as two key evaluation metrics. Classification performance was measured by accuracy (the proportion of correctly classified samples to total samples) and recall rate (the ability to correctly identify authorized user operations), which respectively reflect the overall reliability of classification results and the completeness of identifying positive samples. Time efficiency focuses on the detection time, reflecting whether the system can timely, rapidly, and effectively identify and prevent illegal intrusions. This is directly related to the effective protection ability of the device.

Subsequently, this study compares accuracy, recall, detection time and other key metrics of each classifier through experiments, and determines the optimal classifier, which provided solid technical support for enhancing the reliability of identity verification of UAV touch devices. The detailed working principles of two classifiers (RF and RBFNN) in this scenario are elaborated on in the following subsection.

1) Random Forest

Random Forest (RF) transforms identity authentication into a multidimensional feature-voting process through ensemble learning with multiple decision trees. Its core workflow consists of three stages: construction of positive and negative samples, mixed-feature learning, and ensemble decision-making.

RF constructs positive samples labeled as 1 using the legitimate user's SCTO data. Negative samples labeled as 0 are generated by injecting Gaussian noise into the feature matrix of the positive samples. After noise injection, each feature follows the probability density function

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right), \tag{1}$$

where x denotes the perturbed feature value, μ is the perturbation mean (set to 0 so that features fluctuate around the original positive-sample values), and σ is the perturbation standard deviation, controlled at 50% of the standard deviation of the corresponding positive-sample feature. This approach simulates both the similarity between legitimate and illegitimate operations and the characteristics of spoofing behaviors. It addresses the difficulty of obtaining large-scale illegitimate-operation data, enabling the generation of suitable negative samples using only positive samples and reducing dependence on extensive training datasets.

This study employs a forest of 100 decision trees, where each tree independently learns a different subset of the SCTO feature group. The ensemble of decision trees effectively reduces the risk of overfitting and prevents the model from learning noise. During the learning process, each tree ranks the importance of features, with higher-contributing features to identity differentiation being prioritized for splitting. The parallel decision-making of multiple trees enables the capture of complex relationships between features.

Random Forest (RF) performs feature importance evaluation through random selection of feature subsets. During the training of each decision tree, a subset of features is randomly chosen to build the split nodes, preventing a single feature from dominating the classification result. This approach allows the model to deeply explore the distinguishing value of multiple feature combinations for identity recognition, such as the correlation between the x coordinate concentration and velocity stability, which could be characteristic of a user's typical operation behavior. The node-splitting mechanism of decision trees is based on information gain and the Gini index. The Gini index measures node impurity and is computed as

$$G(t) = 1 - \sum_{k=1}^{K} p(k|t)^2, \tag{2}$$

where $G(t)$ represents the Gini index of node t (a lower value indicates higher purity), K is the total number of classes (in this case, $K = 2$ corresponding to positive sample $k = 1$ and negative sample $k = 0$), and $p(k|t)$ is the proportion of samples of class k at node t. Information gain is based on entropy, which is calculated as

$$H(t) = -\sum_{k=1}^{K} p(k|t) \log_2 p(k|t), \tag{3}$$

where $H(t)$ is the entropy of node t. Information gain is then given by

$$IG(t,a) = H(t) - \sum_{v \in \text{values}(a)} \frac{|t_v|}{|t|} H(t_v), \tag{4}$$

where $IG(t,a)$ is the information gain of feature a at node t, and a higher value indicates a better split. The term a represents the feature to be evaluated, and values(a) is the set of all possible values for feature a. The subscript t_v corresponds to the child node when feature a takes value v, and $|t_v|/|t|$ represents the proportion of child node samples to the total number of parent node samples. This splitting mechanism adapts automatically to different feature distributions, finding the optimal threshold for continuous features and capturing associations in discrete features by combining different intervals. The frequency of a feature being selected as a split node in each tree is used to quantify the feature's importance, which is computed as

$$I(a) = \frac{1}{T} \sum_{t=1}^{T} \left(\Delta G(t,a) \cdot \frac{|t|}{N}\right), \tag{5}$$

where $I(a)$ is the importance score of feature a, T is the number of decision trees in the forest (here $T = 100$), $\Delta G(t,a)$ is the decrease in Gini index when feature a is used to split node t, $|t|$ is the number of samples at node t, and N is the total number of samples in the training set. This quantifies the feature's contribution to the classification process and ensures that critical features are not omitted.

In the ensemble decision stage, RF outputs the probability that an SCTO sample belongs to a legitimate user. This probability is aggregated from the voting results of multiple decision trees and can be expressed as

$$P(x = 1) = \frac{1}{T} \sum_{t=1}^{T} I(h_t(x) = 1), \tag{6}$$

where $P(x = 1)$ is the probability that sample x is a legitimate user, and $h_t(x)$ is the detection result of tree t for sample x. A value of 1 indicates the sample is classified as a legitimate user, while 0 indicates an illegitimate user. $I()$ is the indicator function, which returns 1 if the condition is true, and 0 otherwise. A higher detection probability indicates that the touch characteristics match more closely with the legitimate user's pattern. Further, the algorithm divides the feature distribution of a single sample into several sub-samples for further detection. The proportion of valid sub-samples is determined by

$$R = \frac{1}{M} \sum_{m=1}^{M} I\left(P(x_m = 1) > \theta\right), \tag{7}$$

where R is the proportion of sub-samples classified as legitimate, M is the total number of sub-samples, x_m is the m-th sub-sample, and θ is the detection probability threshold. If the probability of more than a certain proportion of sub-samples exceeds the threshold, the sample is classified as legitimate, following the majority voting principle to determine whether the overall touch operation is from a legitimate user.

Specifically, during the detection process, RF aggregates all sub-samples of a single sample, computes the average prediction probability, the maximum probability, and the proportion of samples exceeding the threshold, and ultimately uses the proportion of samples exceeding the detection threshold to determine the SCTO sample classification.

This aggregation strategy helps resist "occasional operational variations," such as coordinate deviations caused by hand tremors, thus improving the stability of identity authentication. However, experimental studies show that when each SCTO sample is divided into only one sub-sample, the classification accuracy of the model is highest during both training and detection.

2) Radial Basis Function Neural Network

We first preprocess the SCTO data by applying Z-score normalization to map all features to a distribution with a mean of 0 and a variance of 1, expressed as

$$x' = \frac{x - \mu}{\sigma}, \tag{8}$$

where x is the original feature value, μ is the mean of the feature in the training set, and σ is the standard deviation of the feature. This ensures that all features contribute equally to the RBFNN learning process, enabling nonlinear fitting. All training samples are labeled as positive, allowing the RBFNN to focus on learning the distribution patterns of legitimate user touch features, thus avoiding classifier bias due to uneven distribution of negative samples. The core advantage of RBFNN stems from the radial basis functions in the hidden layer and the adaptive learning of parameters. The process involves using K-means clustering to determine centers, calculating β based on neighborhood distance, and solving for weights using regularized least squares, thereby modeling the SCTO features.

The hidden layer of the RBFNN maps the low-dimensional nonlinear features of the touch operation to a high-dimensional linearly separable space. The center parameters and β values are automatically learned through data-driven processes. The centers are determined using the K-means clustering algorithm applied to the training set features. The optimization objective of the K-means algorithm is to minimize the sum of squared distances within clusters, expressed as

$$J = \sum_{k=1}^{K} \sum_{x \in C_k} \|x - \mu_k\|^2, \tag{9}$$

where K is the pre-set number of cluster centers, i.e., the number of RBFNN neurons, C_k is the set of samples belonging to the k-th cluster, and μ_k is the center vector of the k-th cluster. $|x - \mu_k|^2$ is the squared Euclidean distance between sample x and the center μ_k. Each cluster center corresponds to an activation point of a radial basis function. This process captures the legitimate user's touch features, ensuring that all typical feature clusters of touch operations are covered.

The width parameter β determines the activation range of the radial basis function. It is computed by calculating the distance between each center and the nearest neighboring center. The Euclidean distance between two neighboring centers is expressed as

$$\text{min\_dist}_k = \min_{\substack{j=1,2,\ldots,K \\ j \neq k}} \sqrt{\sum_{d=1}^{D} (\mu_{k,d} - \mu_{j,d})^2}, \tag{10}$$

where μ_k, d is the value of the k-th center in the d-th feature dimension, and D is the total number of feature dimensions. Combining the formula for β, the closer the centers are, the larger the β value will be, resulting in a smaller activation range. This enables RBFNN to produce strong activations for samples close to the center of the legitimate user feature and weak activations for samples that deviate from the center. Regularization weights need to be solved to avoid overfitting. We construct the design matrix and solve for the weights using regularized least squares. The design matrix G is defined as

$$G = \begin{bmatrix} \phi_1(x_1) & \phi_2(x_1) & \ldots & \phi_K(x_1) & 1 \\ \phi_1(x_2) & \phi_2(x_2) & \ldots & \phi_K(x_2) & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \phi_1(x_N) & \phi_2(x_N) & \ldots & \phi_K(x_N) & 1 \end{bmatrix}, \tag{11}$$

where N is the number of training samples, and $\phi_k(x_i)$ is the activation value of the k-th radial basis function for the i-th sample. The last column represents the bias term. Each row in the design matrix G corresponds to a training sample, and each column corresponds to the output of a hidden layer neuron, i.e., the activation value of the radial basis function for that sample. The last column, filled with 1s, ensures that RBFNN is resilient to incidental noise, such as tremors or mis-touches, when learning touch features. To solve for regularized weights, a unit matrix is added to suppress overly large weights and improve the generalization ability of the classifier.

RBFNN classifies touch operation features through the nonlinear mapping of the radial basis function (RBF). The radial basis function is described as

$$\phi_k(x_i) = \exp\left(-\beta_k \cdot \|x_i - \mu_k\|^2\right), \tag{12}$$

where the similarity between sample x and center μ is measured. The kernel width parameter β is dynamically determined by the minimum Euclidean distance between centers. For each center μ, the minimum distance to other centers is computed as min_dist, and the kernel width is adapted by

$$\beta = \frac{1}{2 \cdot \text{min\_dist}^2}. \tag{13}$$

When the touch features are dense, the distance between centers is small, leading to a large β value. This results in a function that is sensitive to fine differences near the centers. When the touch features are sparse, the distance between centers is large, leading to a smaller β value to ensure coverage of peripheral samples. This adaptive sparsity feature perfectly matches the distribution patterns of touch operations, where conventional patterns are concentrated and peripheral variations are scattered. On this basis, high-dimensional feature mapping is achieved by constructing the design matrix G.

For each input sample, the RBF outputs (i.e., the similarities) to all centers are computed and form matrix column vectors. Each row of the matrix intuitively reflects how well

the sample matches each of the touch feature centers. To guard against noise in touch operations, regularized least squares are used to solve for the weights, which are expressed as

$$\text{weights} = (G^T G + \lambda I)^{-1} G^T y, \tag{14}$$

where G^T is the transpose of the design matrix G, λ is the regularization parameter used to control the strength of regularization, I is the identity matrix of the same dimensions as $G^T G$, and y is the label vector of the training samples. Since all training samples are positive labels, y is a vector of all 1s. This method helps reduce the impact of noise and ensures that RBFNN focuses on stable touch patterns, ultimately achieving robust linear combinations and classification of nonlinear features.

Once the weights in (14) are determined, we use the RBFNN classification to determine whether the STT sample is positive or negative. The decision function can be expressed as

$$f(x) = \sum_{k=1}^{K} w_k \phi_k(x) + b, \tag{15}$$

where $f(x)$ denotes the decision value output by RBFNN for the STT sample x, w_k represents the trained weight associated with the kth radial basis function, $\phi_k(x)$ is the activation value of the kth RBF corresponding to sample x, and b is the bias term. During classification, a threshold θ is introduced to separate authorized and unauthorized operations. If $f(x) \geq \theta$, the sample is identified as originating from a legitimate user; otherwise, it is classified as an illegal intrusion.

D. Data Integration and Decision Output Module

This module is used to collect and integrate the large volume of data generated by previous processes, further formulate the decision rules, and output the final judgement results. This judgement result will be read by the operation authority control program. If identified as an unauthorized user, the operation authority will be terminated, and a series of input-based identity information (such as passwords, account binding, verification codes, or biometric information) will be re-verified, thereby ensuring the security of critical devices.

For detection results with a large sample size, we adopt a proportion-threshold-based decision method. Specifically, when the proportion of positive samples exceeds the preset threshold, the user is ultimately classified as a legitimate user, otherwise it is deemed an unauthorized intrusion. The higher the threshold, the stricter the judgment becomes. It enhances the flexibility and adjustability of the system in various UAV operational scenarios. This module serves as the final refinement of our passive identity verification system.

III. EXPERIMENT AND ANALYSIS

A. Data Acquisition and Performance Metrics

To evaluate the performance of the proposed passive identity authentication method, we applied it to an industrial scenario involving UAV inspection tasks for potential safety hazards within a factory and conducted a series of experiments. A specific inspection mission was selected as the experimental scene, ensuring that the UAV followed the same route, while the study focused solely on the differences in operational behavior among different users.

In this experiment, 400 touch data samples from one authorized user were collected as the classifier training set. Additionally, 151 samples from the same authorized user were used as the positive test set, and 173 samples from five non-authorized users were used as the negative test set. To evaluate the classifier's resistance to imitation attacks, the negative samples were collected by non-authorized users who attempted to replicate the authorized user's operation behavior through observation and imitation. This design allowed for the verification of whether the classifier could distinguish between genuine samples and imitation-based negative samples. In total, 324 test samples were used in this evaluation.

Each sample represented a continuous touch operation lasting up to 5 minutes. The data acquisition frequency was set to 0.1 seconds, resulting in a maximum sequence length of 3000, although actual lengths varied between 1 and 3000 depending on individual operating habits. The maximum number of simultaneous touch points was limited to three fingers.

Performance of the passive authentication method was evaluated using False Acceptance Rate (FAR) and False Rejection Rate (FRR). Specifically, FAR measures the proportion of imposters or unauthorized users incorrectly accepted as legitimate users, while FRR measures the proportion of legitimate users incorrectly rejected by the system. The Equal Error Rate (EER) represents the critical point at which FAR equals FRR. In addition, the Receiver Operating Characteristic (ROC) curve and the Area Under the Curve (AUC) were used to illustrate the trade-off relationship between system sensitivity and specificity. $Table2$ summarizes the AUC values for individual classifiers and the optimal fusion classifier, providing a direct comparison of their performance.

TABLE II
AUC VALUES OF SINGLE CLASSIFIERS AND OPTIMAL FUSION CLASSIFIER

Classifier	AUC
RF	0.922
RBFNN	0.917
RF + RBFNN	0.955

B. Performance Analysis of Individual Classifiers

Two machine learning algorithms, RBFNN and RF, were used to learn from the aforementioned training set to construct two classifiers. These classifiers were then tested on the collected positive and negative sample datasets to evaluate their discrimination capability and conduct comparative performance analysis. As shown in Fig. 3, The ROC curves and the corresponding AUC values of the two classifiers were plotted as evaluation metrics.

We can see from Fig. 3 that both the RBFNN and the RF classifiers demonstrated excellent performance, with AUC values exceeding 0.91. This indicates that both RBFNN and RF exhibit high discrimination accuracy for touch sequences

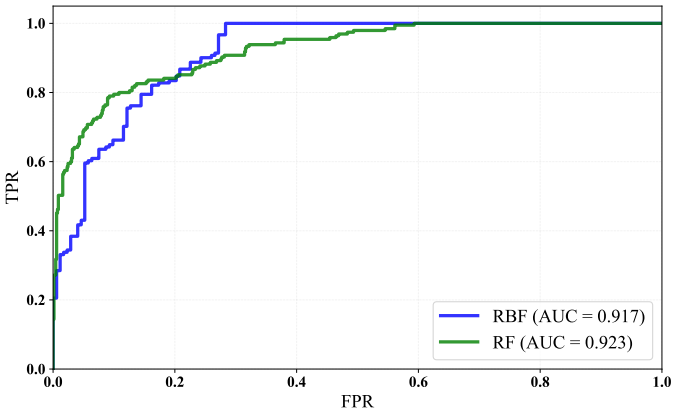

Fig. 3. Comparison of ROC curves between RF and RBFNN classifiers.

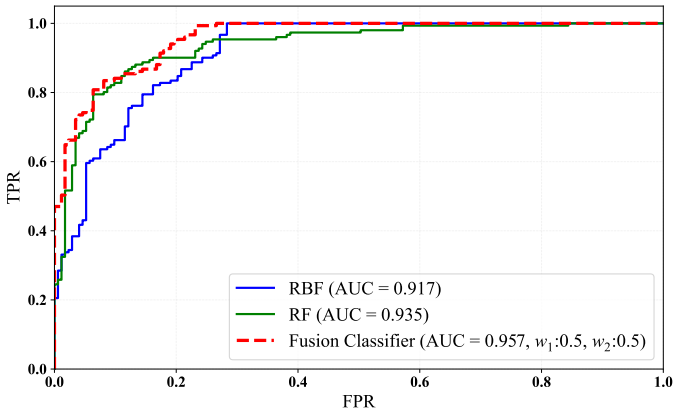

Fig. 4. ROC curve comparison among RBFNN, RF, and the weighted fusion classifier.

under the proposed method, and both classifiers show strong resistance to imitation-based attacks. These findings suggest that the passive user authentication approach based on spatial-temporal feature extraction from continuous touch trajectories has significant potential for adaptation to various complex UAV application scenarios.

C. Performance Analysis of the Fusion Classifier

To further improve binary classification accuracy, reduce the risk of misclassification caused by single-model bias, and explore the potential of classifier integration, a weighted fusion approach was applied to the detection results of individual samples.

We conducted a weighted allocation experiment using the RF and RBFNN classifiers. Specifically, the weighted detection probability for each sample was defined as follows. Let q_1 denote the detection probability of RF for a given sample, and q_2 denote that of RBFNN. Assign the weight of RF as w_1 and the weight of RBFNN as $w_2 = 1-w_1$. The weighted detection probability q for a single sample can then be expressed as

$$q = w_1 q_1 + w_2 q_2. \tag{16}$$

In this way, each sample obtains a new weighted probability value. Combined with the original positive and negative labels, a new ROC curve and AUC value for the dual-classifier weighted fusion can be obtained. This represents the fusion performance under a specific weight allocation scheme.

We performed weighted allocation experiments using RF and RBFNN, and Fig. 4 illustrates the comparison between the fusion classifier with equal weights ($w_1 = w_2 = 0.5$) and the original single classifiers. As shown in Fig. 4, the ROC curve of the fusion classifier lies entirely above those of the single classifiers, with an AUC value reaching 0.95—higher than either individual model. This indicates that the weighted fusion classifier achieves significantly improved performance compared to single classifiers, demonstrating the potential for highly accurate, imitation-resistant binary classification. Such performance suggests that the proposed method can better adapt to various complex Industrial Internet of Things (IIoT) application scenarios.

D. Detailed Analysis of the Relationship Between Weighting Scheme and Performance

Considering the vast number of possible weight distributions in practice, we aimed to identify the optimal weight allocation and explore the relationship between weight distribution and performance (AUC value), as well as its trend. To achieve this, we conducted an interval precision segmentation experiment. The weight range of 0 to 1 was evenly divided into N intervals. One weight, w_1, was then iteratively assigned to each point within this range, and the second weight, w_2, was determined as $1 - w_1$. This process generated N AUC values corresponding to these different weight allocations. Fig. 5 shows the relationship between weight allocation and the fusion classifier's performance as the weight range [0, 1] was evenly divided into 1000 parts. Each allocation corresponds to an AUC value, reflecting the trend of fusion performance.

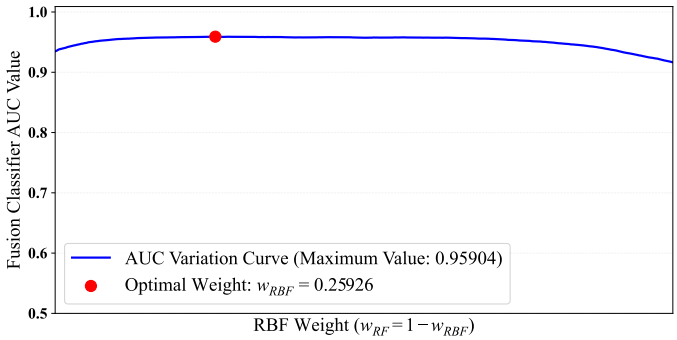

Fig. 5. Usability to weights of two classifiers.

From Fig. 5, a peak is observed, indicating that the optimal weight distribution occurs at RBFNN : RF = 0.25926 : 0.74074, where the fusion classifier achieves the highest AUC value of 0.95904, representing the optimal performance. Additionally, this method establishes a universal optimization framework, enabling the determination of the optimal weight distribution for any future application scenario or newly constructed training and test dataset.

This adaptability verifies the scalability and robustness of the proposed dual-classifier fusion method, ensuring it maintains excellent performance across different data distributions

and application scenarios. It also demonstrates that by appropriately adjusting the weights (ω_1 and ω_2) of the two classifiers (RF and RBFNN), the proposed method can flexibly control the authentication performance, making it suitable for various industrial Internet of Things (IIoT) scenarios.

E. Detection Time Analysis

The detection time of a classifier is directly related to the ability to detect intruders in a timely manner and effectively intercept them. The longer the detection time, the longer the intruder's illegal operation, which increases the potential damage to relevant devices and information. Therefore, it is essential to validate the detection time of each classifier.

We recorded the detection times for 30 trials of 5 samples using RF and RBFNN, and compared the differences in their detection times. In practical applications, only about 5 samples (with a sample group size of 5 steps) are sufficient to determine whether the user is legitimate, thus significantly reducing the detection time. As shown in Fig. 6, the detection times for the two classifiers (RF and RBFNN) over 30 trials for 5 samples were recorded. It can be observed that the detection times for both RF and RBFNN are in the millisecond range, indicating their fast response capabilities and strong timeliness and security.

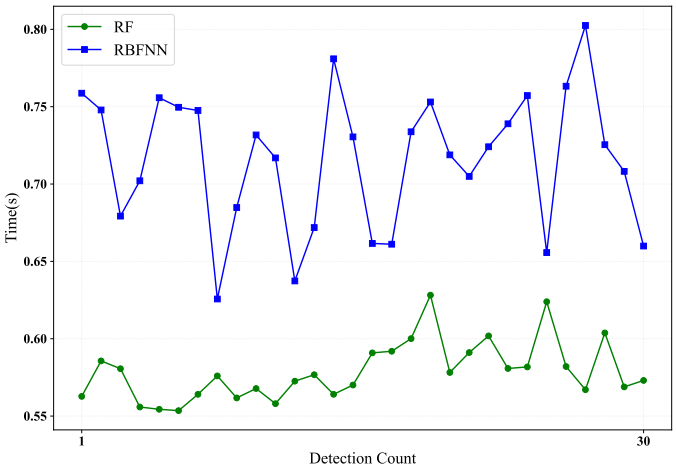

Fig. 6. Comparison of millisecond-level detection times for RF and RBFNN.

Box plots are a statistical visualization method that intuitively presents the central tendency, dispersion, and outliers of data. They consist of five key statistics: minimum, first quartile, median, third quartile, and maximum. By using a box plot, one can easily compare the central position, degree of dispersion, and distribution differences of various data groups.

Fig. 7 provides a more detailed analysis of the detection time distribution for RF and RBFNN using box plots. The horizontal axis represents the classifier type, and the vertical axis represents the detection time. The red horizontal line represents the median, the box represents the interquartile range (IQR), and the whiskers represent the maximum and minimum detection times. The experimental results show that the IQR for RBFNN is relatively longer, with greater fluctuations in detection time, and the overall detection time is longer, ranging from 0.60s to 0.80s. In contrast, RF has the shortest IQR, with the data most concentrated and the most stable detection time, completing detection between 0.55s and 0.60s.

Although there are differences in the data distributions, these differences are minimal, with all times being within the millisecond range and not exceeding 0.8 seconds. Such small time differences provide almost no additional operational space in practical applications. Moreover, the longest detection time does not exceed 0.8 seconds, and the shortest is 0.5 seconds, meaning detection can be completed almost instantaneously after sample collection. This ensures timely and fast identification of legitimate users and illegal intrusions, providing solid and stable support for effective interception.

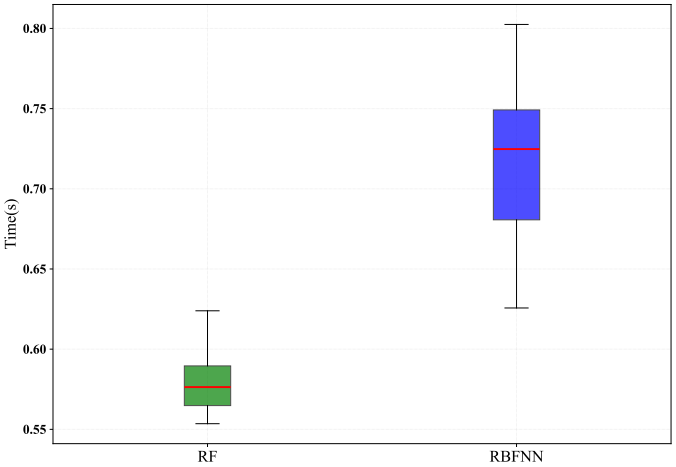

Fig. 7. Box plot of 30 trials of detection times for RF and RBFNN.

F. Security Analysis of Anti-Impersonation Attacks

Impersonation attacks refer to the scenario where an attacker forges or mimics the biometric behavior features (such as screen-touch operation habits) of a legitimate user in order to gain unauthorized access [11]. These attacks pose a direct threat to the security of biometric recognition systems, making it crucial to evaluate the system's robustness under such conditions. In our anti-impersonation attack experiment, four unauthorized users are instructed to observe the legitimate user's touch interaction habits during UAV control operations. These unauthorized users attempt to imitate the operations of legitimate users under performing the same UAV control tasks.

Our experimental results show that the recognition accuracy of single classifiers is only between 0.87 and 0.92, indicating that their classification accuracy and security are relatively weak and more susceptible to deceptive attacks. In contrast, the weighted fusion strategy significantly improves the ability to resist impersonation attacks. The AUC of the RF+RBFNN fusion classifier exceeds 0.95, outperforming single classifiers and demonstrating strong anti-fraud capabilities.

The experimental detection set's negative samples were data collected from unauthorized users who observed and deliberately mimicked the actions of authorized users. Despite this, the UAV's inspection tasks and flight routes were consistent,

leading to excellent identification and classification results. This is directly reflected in the classifier's AUC value, which consistently remained above 0.95, indicating that the series of implicit operational habit features calculated by the proposed method are difficult to observe and mimic by the naked eye. This demonstrates the method's ability to resist impersonation attacks, highlighting the uniqueness and non-reproducibility of these habitual features.

These results suggest that the anti-imitation performance of the method can significantly enhance the security of operational authority for important devices in fields such as drones and Industrial Internet of Things (IIoT). Furthermore, it provides a reliable supplementary identity verification solution, strengthening passive identity authentication and security defense.

G. Experimental Analysis and Summary

In this experiment, we evaluated the performance of the RF and RBFNN classifiers using a series of key metrics to identify the optimal single-classifier model for the UAV industrial inspection scenario. The selected classifier demonstrated sufficient discriminative capability in distinguishing between authorized and unauthorized screen-touch operations. Building on this, we assigned appropriate weights to each classifier and constructed a fused classifier that effectively combines their complementary strengths. Experimental results show that the AUC of the fused classifier consistently rises above 0.95, representing a significant improvement over individual classifiers. This enhancement further strengthens its adaptability to various complex operational conditions, including routine authorized actions, imitation attacks, and noisy or disturbed touch interactions, thereby improving both decision accuracy and robustness.

Applying this authentication technique to UAV security verification effectively addresses several limitations of traditional UAV control authentication mechanisms. First, it mitigates the risks associated with hardware tokens (e.g., USB keys) that can be lost and passwords that can be leaked, by binding identity to a user's unique screen-touch interaction patterns—serving as a behavioral fingerprint. Second, compared with biometric authentication methods (e.g., face recognition) that are sensitive to lighting conditions and occlusions, touch behavior-based authentication remains stable in dynamic UAV operational environments and requires no additional hardware modules. Third, the fused classifier, while maintaining a high decision accuracy ($AUC > 0.95$), benefits from the fast training and detection characteristics of individual classifiers, enabling real-time identity verification during UAV inspection tasks without causing delays that could undermine operational efficiency. These advantages highlight the practical value of this research in enhancing passive identity verification security and anti-spoofing resilience in UAV operations.

In summary, the passive authentication technique based on touch behavioral patterns offers a secure, timely, and convenient identity verification solution for UAV systems. It holds significant practical relevance for safeguarding operational permissions and ensuring secure access to critical UAV equipment and mission workflows.

IV. Conclusion

This study proposes a passive identity authentication framework based on touch interaction behavioral biometrics, aiming to enhance the security of operation permissions in UAV systems. We investigated whether SCTO can serve as a reliable behavioral biometric for continuous authentication, and developed an effective scheme for SCTO data acquisition and structured storage. Two machine learning algorithms—RF and RBFNN—were trained on SCTO feature array to construct classifiers. A weighted fusion mechanism was further introduced to boost performance. Experimental evaluations, including tests with spoofed SCTO samples, demonstrated that the proposed framework achieves stable and high authentication accuracy while maintaining strong resistance against mimicry attacks. Moreover, the optimally selected scheme reaches millisecond-level detection latency, verifying its capability for fast, timely, and effective intrusion prevention.

In addition, the proposed method requires no extra hardware modules. It leverages the native touch interactions of UAV control terminals to bind a user's identity with their "behavioral fingerprint," ensuring both stability during mission execution and a practical balance between real-time responsiveness and security. This provides a lightweight, reliable, and deployable solution for operation permission control, offering a robust intrusion prevention mechanism for touchscreen-based UAV equipment. The framework thus holds substantial practical value for UAV operation security and exhibits promising application potential. Future work will focus on three directions:

1) Optimizing SCTO feature engineering to extract more latent and discriminative patterns from users' operational habits;

2) Exploring adaptive weight fusion strategies to enable dynamic parameter adjustment across different application scenarios;

3) Extending the system to broader UAV tasks, such as power-line inspection and agricultural operations, promoting the transition from experimental validation to practical deployment and further enhancing security across touchscreen-based UAV interaction scenarios.

Acknowledgment

This work was supported in part by the Anhui Province Outstanding Young Teacher Training Project (YQYB2023048), in part by the New Generation Information Technology Innovation Project of the 2023 China University Industry Research Innovation Fund (2023IT010), in part by Chuzhou University Research Launch Fund Project (2024qd11), and in part by the Chuzhou Science and Technology Plan Project (2023ZD028, 2023ZD029).

References

[1] Peter Aaby et al. An omnidirectional approach to touch-based continuous authentication. *arXiv*, 2023.

[2] Ritu Agrawal et al. A study of touch dynamics biometrics authentication. *International Journal of Scientific Research in Engineering and Management (IJSREM)*, 6:181–185, 2022.

[3] Anonymous. A survey on uavs security issues: attack modelling, security requirements and future research directions. *Computers and Communications*, 2023.

[4] N. Bai et al. A survey on unmanned aerial systems cybersecurity. *Computers & Electrical Engineering*, 97:107544, 2024.

[5] R. Choudhary et al. Intrusion detection systems for networked unmanned aerial vehicles: A survey. In *Proceedings of the 2018 International Conference on Advances in Computing, Communication and Engineering*, pages 14–20, 2018.

[6] O. L. Finnegan et al. The utility of behavioral biometrics in user authentication and screen-time measurement: a scoping review. *Systematic Reviews*, 13:68, 2024.

[7] Mario Frank et al. Touchalytics: On the applicability of touchscreen input as a behavioral biometric for continuous authentication. *IEEE Transactions on Information Forensics and Security*, 8(1):136–148, 2013.

[8] S. M. Gilani et al. A robust internet of drones security surveillance communication network based on iota. *Internet of Things*, 19:101066, 2024.

[9] H. J. Hadi et al. A comprehensive survey on security, privacy issues and countermeasures for uavs. *Computers & Security*, 115:102511, 2023.

[10] S. Miri Kelaniki et al. A study on iot device authentication using artificial intelligence. *Sensors*, 24:1060, 2024.

[11] H. Khan et al. Targeted mimicry attacks on touch input based implicit authentication systems. In *Proceedings of the 2016 ACM on Interactive Smart Products and Services*, pages 155–162, 2016.

[12] Yassine Mekdad et al. A survey on security and privacy issues of uavs. *Computer Networks*, 224:109626, 2023.

[13] V. Sihag et al. Cyber security and forensics for unmanned aerial vehicles: A survey. *Sensors*, 7:430, 2023.

[14] J. P. Yaacoub et al. Security analysis of drones systems: Attacks, limitations and recommendations. *Computer Networks*, 174:107243, 2020.

[15] D. Zhao et al. Security situation assessment in uav swarm networks. *Computers & Security*, 113:102498, 2024.

[16] Guozhu Zhao et al. Passive user authentication utilizing behavioral biometrics for iiot systems. *IEEE Internet of Things Journal*, 9(14):12783–12796, 2022.

[17] Guozhu Zhao et al. Passive user authentication utilizing two-dimensional features for iiot systems. *IEEE Transactions on Cloud Computing*, 11(3):2770–2783, 2023.

[18] Guozhu Zhao et al. Exploiting screen-touch trajectory for passive user authentication in industrial internet of things systems. *IEEE Transactions on Industrial Informatics*, 20(7):9098–9110, 2024.