ARTICLE

Dynamic Hand Gesture-Based Person Identification Using Leap Motion and Machine Learning Approaches

Jungpil Shin[1,*], Md. Al Mehedi Hasan[2], Md. Maniruzzaman[1], Taiki Watanabe[1] and Issei Jozume[1]

[1]School of Computer Science and Engineering, The University of Aizu, Aizuwakamatsu, Fukushima, 965-8580, Japan

[2]Department of Computer Science & Engineering, Rajshahi University of Engineering & Technology, Rajshahi, 6204, Bangladesh

*Corresponding Author: Jungpil Shin. Email: jpshin@u-aizu.ac.jp

ABSTRACT

Person identification is one of the most vital tasks for network security. People are more concerned about their security due to traditional passwords becoming weaker or leaking in various attacks. In recent decades, fingerprints and faces have been widely used for person identification, which has the risk of information leakage as a result of reproducing fingers or faces by taking a snapshot. Recently, people have focused on creating an identifiable pattern, which will not be reproducible falsely by capturing psychological and behavioral information of a person using vision and sensor-based techniques. In existing studies, most of the researchers used very complex patterns in this direction, which need special training and attention to remember the patterns and failed to capture the psychological and behavioral information of a person properly. To overcome these problems, this research devised a novel dynamic hand gesture-based person identification system using a Leap Motion sensor. This study developed two hand gesture-based pattern datasets for performing the experiments, which contained more than 500 samples, collected from 25 subjects. Various static and dynamic features were extracted from the hand geometry. Random forest was used to measure feature importance using the Gini Index. Finally, the support vector machine was implemented for person identification and evaluate its performance using identification accuracy. The experimental results showed that the proposed system produced an identification accuracy of 99.8% for arbitrary hand gesture-based patterns and 99.6% for the same dynamic hand gesture-based patterns. This result indicated that the proposed system can be used for person identification in the field of security.

KEYWORDS

Person identification; leap motion; hand gesture; random forest; support vector machine

1 Introduction

Person identification is one of the most vital security tasks. Passwords and biometrics are widely used for user authentication. Traditionally, login names and passwords were used more frequently for user authentication. Although text-based passwords are a primitive technique, one can suffer from various well-known problems [1]. An existing study illustrated that text passwords needed to be longer and more complex to be secure [2]. Long or strong passwords are very challenging to remember due to their complex structures. Besides, text passwords can be easily copied or stolen by attackers [3].

Furthermore, people cannot prevent password leakage from the corporation by cracking [4]. Physical and behavioral characteristics were widely used to determine and discriminate between one person and others [5]. Voice [6], fingerprint [7], and facial-based [8] recognitions were more popular and widely used as identification techniques. The attackers can easily be enabled to deceive the facial and iris-based identification systems by screening photos. Although the face and fingerprints of each person are distinct, an attacker can easily make fake fingerprints or faces by taking a snapshot of your finger or face. An existing study illustrated that 54.5% of pattern-based authentications were enabled to be recovered by attackers [9]. As a result, it should be noted that the fingerprint-based recognition system necessitated a touching device.

To solve these problems, we designed a simple dynamic hand gesture-based person identification system. This system had a lower chance of information leakage and the possibility of attacks due to collecting information from hand gestures. Moreover, our invented system does not require touching devices due to hand gestures in mid-air. This system was invented based on hand movement patterns to identify a person. It was not only used for identifying a person but also used without thinking about it. We employed our proposed system using a 3D motion sensor, called Leap Motion (LM) controller. Using two infrared cameras and the infrared light-emitting diode (LED), the LM-based sensor recorded users' movements of hands above the camera and converted them into a 3D input [10].

Recently, a lot of existing studies have been conducted on user or person identification using different hand gesture patterns such as air-handwriting [11–13], drawing circles [14], drawing signatures [15,16], alphabet strings [17] and so on. For example, Saritha et al. [14] proposed an LM-based authentication system based on circle-drawing, swipe, screen-tap, and key-tap-based gestures. The limitation of this study was that sometime fingers were hidden from LM-based camera and the LM-based sensor may provide ambiguous data. Xiao et al. [15] developed a 3D hand gesture-based signature system for user authentication. Chahar et al. [17] collected alphabet string-based hand gestures for making passwords using an LM-based sensor. The main limitation of alphabet string gesture-based system was that it was more difficult to remember the pattern. Xu et al. [11] introduced an authentication system based on air handwriting style. They randomly chose a string (a few words) and asked users to write this string in the air. The users' writing movements and writing styles were captured by an LM-based sensor. The limitations of their proposed approach was that it was difficult to write the string in the air and that every time person could create a different shape of the string. Wong et al. [18] also proposed stationary-based hand gestures (arms, wrists, arms with hands, and fingers) patterns to develop a new biometric authentication approach using an LM-based sensor. The limitations of stationary-based hand gestures patterns were that need a special training to know the gesture of corresponding number and sometime difficult to remember the patterns.

In a summary, we accumulated the limitations of these existing studies were as follows: (i) did not use comparatively simple gestures and used gestures with complex motions of fingers; (ii) it was difficult to make these gestures, such as writing signatures, specific words, or drawing patterns in the air. Moreover, every time person could be created different shapes of the circle or different types of signatures; (iii) it was difficult to remember the pattern of passwords, such as showing a specific number using gestures; (iv) although Wong et al. [18] used more complex hand gestures but they still did not implement hands motions; and (v) Sometimes fingers were hidden from the LM-based camera and as a result, the LM-based sensor may provide ambiguous data.

To overcome all of these limitations, this current study proposed a novel simple dynamic hand gesture-based person identification approach. Physiological as well as behavioral information of a person was captured by an LM-based device during the time of performing hand gestures. Any

permutation of the finger can be possible (any combination of 4 digits is possible). Furthermore, these gesture patterns were very simple and easy to remember, easy to train, and touchless, but had strong discriminative power. It was impossible to create a fake gesture. To do these hand gesture patterns, first we need to install an LM-based sensor, which easily captures all information from the hand. Whereas every part of the hand was more clearly visible to LM using these gestures. In this work, we collected hand gestures from 25 subjects for two hand gesture-based patterns: Arbitrary hand gesture-based patterns and the same dynamic gesture-based patterns. Each subject had hand gestures about 20 times. In this work, we devised a system that used LM-based sensors to capture physiological as well as behavioral information of a person and performed classification using a support vector machine (SVM). The main contributions of this study are as follows:

- We proposed a novel simple dynamic hand gesture-based person identification approach. This system used physiological as well as behavioral information of a person captured by an LM-based device.
- The proposed dynamic hand gestures are easy to remember, easy to train, and touchless.
- We made two gesture datasets (one for arbitrary hand gesture-based patterns and another for the same dynamic gesture-based patterns).
- We extracted various types of static and dynamic features from hand geometry using statistical techniques. A random forest (RF)-based model was used to select the important features to reduce the complexity of the system.
- Performed person classification using SVM and assessed its performance using identification accuracy.

The layout rest of the paper is as follows: Section 2 represents related work; Section 3 presents materials and methods that include the proposed methodology, measurement device, used 3D hand gesture, data collection procedure, data formation, feature extraction, and feature normalization, feature important measure using RF, classification using SVM, device, experimental setup and performance evaluation metrics of classification model are more clearly explained. Section 4 presents experimental protocols and their correspondence results are discussed in Section 5. Finally, the conclusion and future work direction is more clearly discussed in Section 6.

2 Related Work

A lot of existing works have already been performed on persons or user's identification. In this section, we reviewed these existing works and tried to briefly discuss the summarization of these existing works which are presented in Table 1. Chahar et al. [17] collected alphabet string-based hand gestures for making passwords using the LM sensor. They developed a leap password-based hand gesture system for biometric authentication and obtained an authentication accuracy of more than 80.0%. Wong et al. [18] also developed a biometric authentication approach using an LM-based sensor. They utilized stationary hand gestures, like arms, wrists, arms with hands, and fingers. They illustrated that their proposed authentication system obtained an average equal error rate (EER) of 0.10% and an accuracy of 0.90%. Xiao et al. [15] developed a 3D hand gesture-based authentication system. They collected experimental data from 10 subjects to check the efficiency of their proposed system and evaluated the performance of the proposed system using EER. They also collected the behavioral data measured by an LM-based sensor. They showed that their proposed authentication system obtained an average EER of 3.75%.

Table 1: Overview of existing studies for person identification

Authors	Subjects	Proposed method	PR (in %)
Chahar et al. [17]	150	NB, RF, NN	ACC: 81.0
Wong et al. [18]	10	ED-FPDI	ACC 80.0
			EER: 0.10
Xiao et al. [15]	10	DTW	EER: 3.75
Xu et al. [11]	24	SVM	EER: 1.18
Kamaishi et al. [19]		SOM	ACC: 86.57
Satoru et al. [20]	9	ED	ACC: 98.8
Manabe et al. [21]	21	RF	ACC: 94.98
Malik et al. [22]	40	CNN	ACC: 67.6
			EER: 0.06
Rzecki [23]	10	k-NN	ACC: 95.0
Jung et al. [24]	100	CNN	
Guerra-Segura et al. [25]	100	LSSVM	EER: 1.20
Kabisha et al. [26]		CNN, VGG16	More than 98.0
Maruyama et al. [27]	20	Distance-based	ACC: 84.56
Shin et al. [28]	20	SVM	ACC: 98.70

Note: ACC: Accuracy; ED: Euclidean distance; and PR: Performance rate.

Xu et al. [11] introduced a challenge-response authentication system based on air handwriting style. They randomly chose a string (a few words) and asked the users to write this string in the air. The users writing movements and writing styles were captured by an LM-based sensor. They evaluated their proposed system by collecting samples from 24 subjects and obtained an average EER of 1.18% (accuracy: 98.82%). Kamaishi et al. [19] also proposed an authentication system based on handwritten signatures. The shapes of fingers, velocity, and directions of fingers were measured by an LM-based device. They adopted Self-organizing maps (SOMs) for user identification and obtained an authentication accuracy of 86.57%. Imura et al. [20] also proposed 7 types of 3D hand gestures for user authentication. They made gesture samples from 9 participants. Each participant performed each of the 7 hand gestures 10 times and got 630 samples ($9 \times 7 \times 10$). They computed the similarity using Euclidean distance (ED) and obtained a true acceptance rate (TAR) of more than 90.0% and EER of less than 4%.

Manabe et al. [21] proposed a two-factor user authentication system combining password and biometric authentication of hand. The geometry of hands and movement of hand data were measured by the LM device. Then, they extracted 40 features from raw features, which were used for user authentication. Subsequently, an RF-based classifier was implemented to classify whether the users were genuine or not. At the same time, their proposed authentication system was evaluated by 21 testers and got an authentication accuracy of 94.98%.

Malik et al. [22] developed a novel user identification approach based on air signature using a depth sensor. They collected 1800 signatures from 40 participants. They adopted a deep learning-based approach and obtained an accuracy of 67.6% with 0.06% EER. Rzecki [23] introduced a machine learning (ML)-based person identification and gesture recognition system using hand gestures in the

case of small training samples. About 22 hand gestures were performed by 10 subjects. Each subject was asked to execute 22 gestures 10 times. A total of 2200 gesture samples (22 × 10 × 10) were used to evaluate the efficiency of their proposed system. When they considered only five gesture samples as a training set, their proposed system obtained an error rate of 37.0% to 75.0% than other models. However, the proposed system achieved an error rate of 45.0% to 95.0% when the training set had only one sample per class.

Jung et al. [24] proposed a user identification system using a Wi-Fi-based air-handwritten signature. They engaged 100 subjects to collect air-handwritten signature signals. They collected 10 samples from each of the two sitting positions, each of which had four signatures. A total of 8,000 (2 × 4 × 10 × 100) samples were used for the experiments. They showed that their proposed system obtained a promising identification accuracy. Guerra-Segura et al. [25] explored a robust user identification system based on air signature using LM-based controller. To perform experiments, they constructed a database that consisted of 100 users with 10 genuine and 10 fake signatures for each user. They used two sets of impostor samples to validate their proposed system. One was a zero-effort attack imposter and another one was an active imposter. Moreover, the least square SVM was implemented for classification and obtained an EER of 1.20%. Kabisha et al. [26] proposed a convolutional neural network (CNN)-based model for the identification of a person based on hand gestures and face. The proposed system had two models: One was a VGG16 model for the recognition of face and another was a simple CNN with two layers for hand gestures. They showed that the face recognition-based model obtained an identification accuracy of 98.0% and the gesture-based model obtained an identification accuracy of 98.3%.

Maruyama et al. [27] also proposed user authentication system with an LM-based sensor. They identified a user by comparing the inter-digit and intra-digit feature values, which were computed from the 3D coordinates of finger joints. They enrolled 20 subjects to do this experiment and measured the hand gestures of each subject 30 times. Finally, they used 600 gesture samples (20 × 30) and obtained an authentication rate of 84.65%. Shin et al. [28] developed a user authentication approach based on acceleration, gyroscope and EMG sensors. They collected 400 samples from 20 subjects. Each subject performed gestures 20 times and then, trained multiclass SVM and obtained an authenticate rate of 98.70%.

3 Materials and Methods

3.1 Proposed Methodology

The steps involved in our proposed methodology for person identification are more clearly explained in Fig. 1. The first step was to divide the dataset into two phases: (i) a training phase and (ii) a testing phase. The second step was to extract some static and dynamic features from the hand geometry. Then, min-max normalization was performed to transform the feature values of a dataset to a similar scale, usually between 0 and 1. The fourth step was to select relevant or important features using the RF model. The next step was to tune the hyperparameters of SVM by Optuna and select the hyperparameters based on the classification accuracy. Subsequently, we retrained the SVM-based classifier, used for person identification and evaluated its performance using identification accuracy.
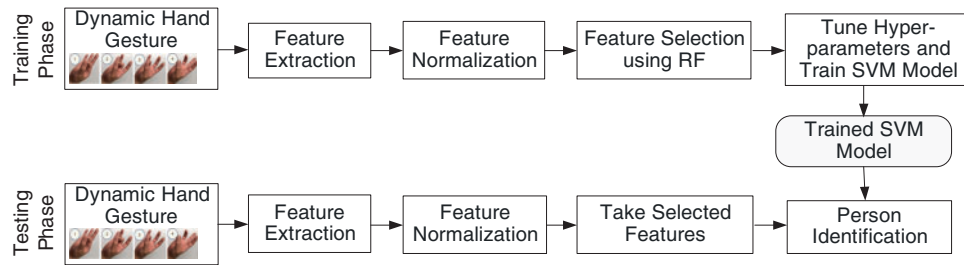
Figure 1: Proposed methodology for person identification

3.2 Measurement Device

In this work, a Leap Motion (LM) device was used to measure hand gestures. LM is an interactive device that can be used to capture all information from the hand through infrared cameras. Recently, LM has gained more popularity as a virtual reality controller. The recognition procedure of LM is illustrated in Fig. 2, whereas the hand is represented as a Skelton. Usually, we can easily obtain the data points of the hand based on the API of the LM controller. According to officials, LM was a high-performance device whose tracking accuracy was 0.01 mm. Whereas, from the existing studies, the accuracy was 0.2 mm for normal accuracy and 1.2 mm for dynamic hand gestures [29,30]. LM photographs of fingers and hands were lit up using the infrared LEDs of two monochromatic IR cameras and sent at almost 120 frames (per second) to the host computer. It had a typical model of human hands as well as the ability to treat different hand motions by connecting sensor data to the model.
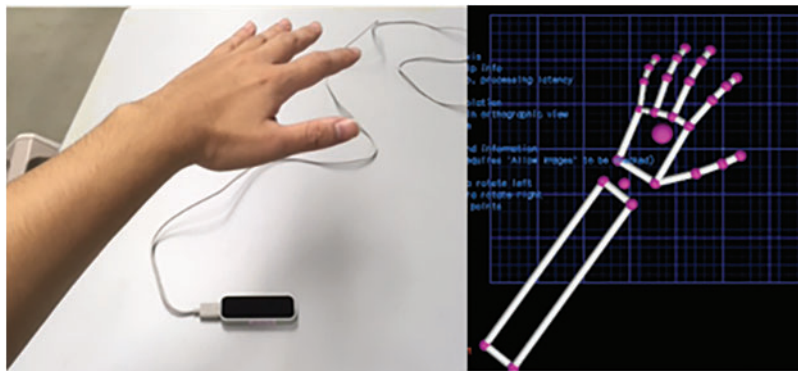

Figure 2: Capturing procedure of hand geometry by leap motion

3.3 Used 3D Hand Gesture

It was essential to define a hand gesture pattern to perform biometric authentication based on geometry and 3D hand gestures. In this work, we moved the hand gestures to get the geometry and position of the hand using the LM device. This current study proposed two types of dynamic hand gesture patterns for person identification. One was arbitrary hand gesture patterns, and another was a specific order or the same dynamic hand gesture patterns, which are illustrated in Fig. 3. Person identification was performed by finger-tapping-like gestures. Any permutation of the finger can be possible (4 any digit is possible). Physiological as well as behavioral information of a person was

captured by an LM-based device during the time of performing hand gestures. This proposed novel simple dynamic hand gesture pattern moved all fingers without hiding any.
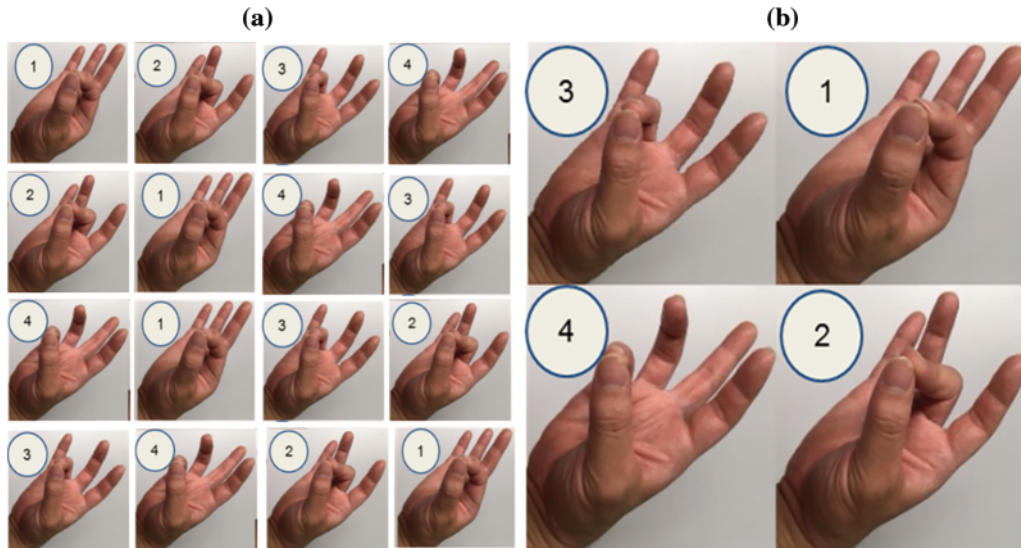

Figure 3: Flow of Gestures: (a) Arbitrary patterns and (b) Specific order of same dynamic gesture patterns

3.4 Data Collection Procedure

In this work, 25 subjects aged 18–30 years were used to measure hand gestures. These hand gesture measurements were taken from the right hand, regardless of the dominant hand. While performing hand gestures, we considered position/inclination variability conditions and followed certain rules: (i) keep the hand directly on LM; (ii) keep the hand more than 15 cm and less than 40 cm from LM; (iii) begin measuring hand gestures after opening the hand; (iv) after the measurement, open your hands; (v) subjects naturally execute hand gestures; (vi) measure hand gestures about 20 times from each subject. The camera of the LM-based device was faced directly above. For this reason, we chose these gestures in such a way that all fingers were unable to hide from cameras. We can easily identify the hand gestures in which each finger touched the other fingers. We hope that our new proposed dynamic gesture pattern approach will make it very simple to measure hand gestures by moving all fingers.

3.5 Dataset Formation

In this paper, we enrolled 25 subjects and collected their hand gesture samples using LM-based device. Each subject was performed by finger tapping and required in the 20 and 30 s to perform hand gestures. Each subject performed finger tapping with one hand. Tapping was done by touching the tip of one finger to the tip of the thumb (see Fig. 4). We performed this movement 4 times and used this as one password. The collected dataset was divided into two parts: The first was a collection of arbitrary hand gesture patterns and the second was a collection of the same hand gesture patterns. We took these hand gestures about 20 times for each subject. As a result, we got a total of 502 gesture samples for arbitrary hand gesture patterns and 517 gesture samples for the same dynamic gesture patterns.
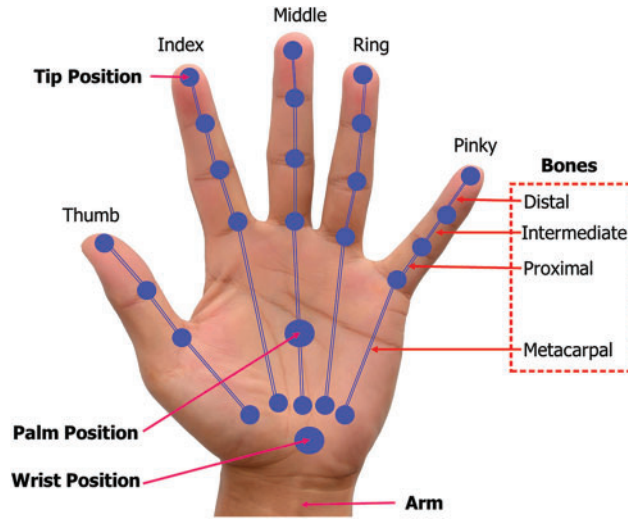
Figure 4: Different parts of a hand used for feature extraction

3.6 Feature Extraction

In this work, we collected over 500 samples that involved timing information. To keep time-oriented behavioral information and the meaning of passwords separately, we divided the data of each dynamic gesture pattern into quarters. Then, for each segmentation, we calculated rotation-type-based features from the bones, palm, and arm; distance to the palm and angle-type-based features from the fingertips and finger joint, and angle and distance to the thumb tip-type-based features from the finger and fingertips. The rest of the feature types, like length, width, velocity, direction, PinchDis, and GrabAngle were calculated from the whole dataset. In this work, we extracted two types of features, such as (i) static features and (ii) dynamic features. A static feature did not change over time, whereas a dynamic feature changed over time. The mean of length and width of bone, palm, and arm-based features were extracted for the static features, whereas the maximum (Max), minimum (Min), mean, variance (Var), and standard deviation (Std) were calculated for rotation, angle, velocity, direction, pinch distance, grab angle, and the distance between each fingertip and thumb tip for the dynamic features. Finally, a total of 1,195 features were used for feature selection and classification, which is presented in Table 2.

3.7 Feature Normalization

Feature normalization is required when working with datasets that contain features that have various ranges and units of measurement. In this case, the variation in feature values may bias the performance of models cause difficulties during the training process or even reduce the performance of some models. In this study, we perform min-max normalization to transform the feature values of a dataset to a similar scale, usually between 0 and 1. It ensures that all features contribute equally to the model and avoids the dominance of features with larger values. It is mathematically defined as:

$$Z = \frac{X - X_{min}}{X_{max} - X_{min}} \tag{1}$$

where, X_{max} and X_{min} are the maximum and the minimum values of the feature, respectively.

Table 2: List of various parts of a hand, the corresponding feature type, the feature name, and the number of extracted statistical features

Part of the hand	Feature types	Feature name	No. of feature
Bones, Palm, Arm	Length, width	Mean	47
Bones, Palm, Arm	Rotation	Max, Min, Mean, Var, Std	440
Fingertip, Finger joint	Distance to Palm (4 divisions)	Max, Min, Mean	508
	Angle (4 divisions)	Max, Min, Mean, Var, Std	
Finger	Angle	Max, Min, Mean, Var, Std	80
Finger tips	Distances to thumb tip		80
Palm	Velocity, direction		30
Arm	PinchDis, GrabAngle		10

3.8 Feature Important Measure Using RF

Feature selection is a technique that is known as "variable selection" and is widely used for the dimension reduction technique in various fields [31]. It improved the performance of a predictive model by reducing irrelevant features. Generally, the potential features helped us to reduce model computational time and costs [32] and improved the predictive model performance [33]. We studied the existing studies and got RF as a potential feature selection approach over various domains [31]. From this viewpoint, We implemented an RF-based model as a feature selection method [34] to measure the discriminative power of the consistency features [35]. The results of the contained feature selection of RF were visualized by the "Gini Index" [36], which was used to measure the importance of features. RF with the Gini Index was widely used to measure the relevant features over different fields [37–39]. In the case of our work, we measured the importance of features using the following steps:

- Divide the dataset into training and test sets.
- Trained RF-based model with 5-fold CV on the training dataset.
- Compute the Gini impurity based on the following formula:

$$GI\,(S) = \sum_{i=1}^{K} P_i\,(1 - P_i) = 1 - \sum_{i=1}^{K} P_i^{2} \tag{2}$$

where, K is the number of classes, P_i is the probability of picking a data point with class i. The value of GI ranged from 0 to 1.

- Sort the features based on the scores of GI in descending order.

3.9 Classification Using SVM

Support vector machines (SVM) is a supervised learning method that is widely used to solve problems in classification and regression [40]. The main purpose of SVM is to find boundaries (hyperplanes) with maximum margins to separate the classes [41]. It requires to solve the following constraint problem:

$$\max_{\alpha} \sum_{i=1}^{n} \alpha_i - \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \alpha_j \mathbf{y_i y_j} \mathbf{K(z_i, z_j)} \tag{3}$$

Subject to

$$\sum_{i=1}^{n} \mathbf{y_i}^T \alpha_i = 1, 0 \le \alpha_i \le C, i = 1, \dots, n \,\&\, \forall\, i = 1, 2,, 3, \dots, n \tag{4}$$

The final discriminate function takes the following form:

$$f(z) = \sum_{i=1}^{n} \alpha_i K(z_i, z_j) + b \tag{5}$$

where, b is the bias term. A kernel function with its parameters need to be chosen for building an SVM classifier [42]. In this paper, the radial basis function (RBF) is used to build SVM classifiers. The RBF kernel which is defined as follows:

$$K(z_i, z_j) = \exp\left(-\gamma \|z_i - z_j\|^2\right) \tag{6}$$

The value of cost (C) and gamma (γ) are tuned using the Optuna library in Python [43]. To tune these parameters, we set the value of C and γ as follows: C: {1e–5 to 1e5} and γ: {1e$-$5 to 1e5}. We repeated the optimization process in 50 trials. We choose the values of C and γ at which points SVM provides the highest classification accuracy. In the case of our study, we used SVMs for identifying persons (classification problems) based on hand gesture authentication.

3.10 Device, Experimental Setup and Performance Evaluation Metrics

In this paper Python version 3.10 was used to perform all experiments. To perform all experiments, we used Windows 10 Pro operating version 21H2 (64-bit) system, Processor Intel(R) Core (TM) i5-10400 CPU @ 2.90 GHz with 16 GB RAM. In this work, SVM with a 5-fold CV was adopted for person identification and its performance was evaluated using identification accuracy (IAC), which is computed using the following formula:

$$IAC = \frac{n\left(\hat{y} \bigcap y\right)}{\text{Total number of samples (n)}} \times 100 \tag{7}$$

where, $\hat{y}$ is the predicted class label and y is the ground true label.

4 Experimental Protocols

In this work, we implemented an RF-based classifier to measure the feature importance. At the same time, we also implemented SVM for person identification. This study performed three experiments: (i) RF-based measure of feature importance; (ii) All feature-based performance analysis for person identification; (iii) Significant feature-based performance analysis for person identification. The results of the three experiments are explained more clearly in the following subsections.

4.1 Experiment-I: Measure of Feature Importance Using RF

The purpose of this experiment was to measure the feature's importance for arbitrary gesture-based and same gesture-based person identification. To measure feature importance, we adopted an RF-based classifier and computed the scores of the Gini Index for each feature during the training

phase. Then, we took $-\log_{10}$ of the Gini Index value and then sorted them in ascending order of magnitude.

4.2 Experiment-II: All Features Based Performance Analysis for Person Identification

The objective of this experiment was to check the performance of arbitrary hand gestures and the same dynamic hand gestures pattern-based person identification system by considering all features (n = 1,195). We implemented SVM with RBF kernel and computed the identification accuracy for both arbitrary gesture-based and same dynamic hand gesture-based pattern datasets.

4.3 Experiment-III: Significant Features Based Performance Analysis for Person Identification

The main goal of this experiment was to display the effect of significant features on the arbitrary hand gesture-based patterns and the same dynamic hand gesture-based patterns. It was essential to understand the identification accuracy of arbitrary hand gesture-based patterns and the same dynamic hand gesture pattern-based person identification systems separately with an increasing number of features. To do that, we took the top 10, top 20, top 30, top 40, top 50, top 60, top 70, top 80, top 90, top 100, top 110, top 120, top 130, top 140, top 150, top 160, top 170, top 180, top 190, and top 200 features and trained SVM with RBF kernel on the training set for each combination of these features and computed the identification accuracy on the test set. This partition was repeated five times to reduce the biasness of the training and test set. Then, the average identification accuracy was calculated for each combination of features. This same procedure was done for both two gesture pattern datasets. At the same time, we plotted them in a diagram of identification accuracy (on the y-axis) *vs.* a combination of features (on the x-axis). Generally, the identification accuracy of our proposed system was increased by increasing the number of feature combinations. While the identification accuracy of our proposed system arrived at flatness or began to slightly decrease, these combinations of feature sets were considered as the more potential features for person identification.

5 Experimental Results and Discussion

The experimental results of three experimental protocols (see Sections 4.1 to 4.3) were more clearly explained in this section. In the 1st experiment, we measured the important features using the Gini Index. The 2nd experiment was to understand the performance of the proposed system for both arbitrary and same dynamic hand gesture-based patterns considering all features. The 3rd experiment was to evaluate the performance of the proposed system for arbitrary hand gesture patterns and the same dynamic hand gesture-based patterns by considering the top prominent features. The results of these three experiments were more clearly explained in the following Sections 5.1–5.3, respectively.

5.1 Results of Experiment-I: Measure of Feature Importance Using RF

We implemented an RF-based model on both arbitrary hand gesture patterns and the same dynamic hand gesture-based pattern datasets to compute the score of the Gini Index for each feature and also analyzed the relevancy of features. The calculation procedure of feature importance measurement using an RF-based Gini Index was more clearly explained in Section 3.8. At the same time, we took $-\log_{10}$ of the Gini Index value for each feature and then sorted them in ascending order of magnitude. Subsequently, these features were fed into the SVM model for classification.

5.2 Results of Experiment-II: All Features Based Performance Analysis for Person Identification

As we previously mentioned, we extracted 1,195 features (both static and dynamic) from different parts of the hands for both arbitrary hand gestures and the same dynamic hand gesture-based patterns datasets (see in Table 2). Then, we normalized these extracted features using the min-max normalization. At the same time, we trained SVM with RBF kernel and then, we computed the identification accuracy. This process was repeated five times and then, we computed the average identification accuracy, which is presented in Table 3. Here, five optimal values of C and gamma for training SVM with rbf kernel five times were as follows: {93920.810, 5.040, 83775.892, 356.517, 7.715} and {0.000631492, 0.027126667, 0.015974293, 0.000742292, 0.016296782} respectively for arbitrary dynamic gesture pattern dataset. Whereas, five values of C and gamma were {1684.125, 96186.042, 20.816, 16152.232, 5517.891} and {0.001739006, 0.013045585, 0.001492318, 0.010034647, 0.000144104} respectively for same dynamic gesture pattern dataset. As shown in Table 3, the SVM-based classifier provided a person identification accuracy of more than 99.0% for both arbitrary hand gesture-based pattern and the same dynamic hand gesture-based pattern datasets.

Table 3: Person identification accuracy (in %) of our proposed system by considering all features

Datasets	Accuracy
Same dynamic hand gesture patterns	99.6
Arbitrary dynamic gesture patterns	99.8

5.3 Results of Experiment-III: Significant Features-Based Performance Analysis for Person Identification

This experiment displayed the performance analysis of our proposed system by taking only significant features. To perform this experiment, we took the top 200 features out of 1,195 features to show the discriminative power of our proposed system. At the same time, we took the top 10, top 20, top 30, top 40, top 50, top 60, top 70, top 80, top 90, top 100, top 110, top 120, top 130, top 140, top 150, top 160, top 170, top 180, top 190, and 200 features. Every time, the combination of features was fed into an SVM-based classifier, and computed the identification accuracy of each trail. After that, the average identification accuracy was computed for each combination of features, which is presented in Table 4 and their correspondence results are also presented in Fig. 5.
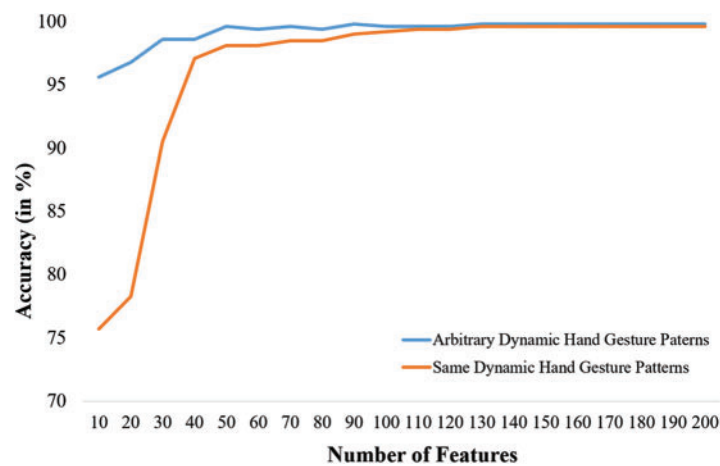
Table 4: Person identification accuracy (in %) of our proposed system by considering the top 200 features

Combination of features	Arbitrary hand gesture patterns	Same dynamic gesture patterns
10	95.6	75.7
20	96.8	78.3
30	98.6	90.5
40	98.6	97.1
50	99.6	98.1
60	99.4	98.1
70	99.6	98.5

(Continued)

Table 4 (continued)

Combination of features	Arbitrary hand gesture patterns	Same dynamic gesture patterns
80	99.4	98.5
90	99.8	99.0
100	99.6	99.2
110	99.6	99.4
120	99.6	99.4
130	99.8	99.6
140	99.8	99.6
150	99.8	99.6
160	99.8	99.6
170	99.8	99.6
180	99.8	99.6
190	99.8	99.6
200	99.8	99.6

Figure 5: Effect of changing combination of features *vs*. person identification accuracy

As shown in Fig. 5, we observed that the identification accuracy was increased by increasing the combination of features. For arbitrary hand gesture-based patterns, it was observed that the identification accuracy remained unchanged after obtaining an average identification accuracy of 99.8% for the combination of 130 features. On the other hand, it was also observed that the identification accuracy remained unchanged after the combination of 130 features. At that combination of 130 features, our proposed dynamic hand gesture-based system produced an identification accuracy of more than 99.0%. Here, we used the five optimal values of C and gamma for training SVM were {20.28, 32.43, 82246.46767, 27.28, 947.99}, and {0.214411977, 0.135645423, 5.94E-05, 0.390716241, 0.023743738} respectively for arbitrary hand gesture patterns dataset and {1461.915188, 70.14861802, 21.79148782, 55.84993545, 415.7135579} and {0.002513596, 0.04267012, 0.943202675, 0.514050892, 0.005184438} for same hand gesture patterns dataset. The recall, precision, and F1-score of our

proposed system for the top 130 features are also presented in Table 5. We observed that our proposed system produced higher performance scores (recall: 99.8%, precision: 98.8%, and F1-score: 99.7%) for arbitrary dynamic hand gesture patterns compared to the same dynamic hand gesture patterns.

Table 5: Performance scores (in %) of our proposed system by considering top 130 features

Datasets	Recall	Precision	F1-score
Same dynamic hand gesture patterns	99.6	99.7	99.5
Arbitrary dynamic hand gesture patterns	99.8	99.8	99.7

5.4 Comparison between Our Study and Similar Existing Studies

The comparison of our proposed dynamic gesture-based person identification system against similar existing approaches is presented in Table 6. Nigam et al. [44] proposed a new drawing signature-based bio-metric system for user authentication. They collected 900 gesture samples from 60 participants. They asked each participant to write their signature on the LM-based device and also collected their facial images while performing gestures. Moreover, they extracted the Histogram of Oriented Trajectories (HOT) and Histogram of Oriented Optical Flow (HOOF)-based features from the leap signature. They used 40% of gesture samples as a training set and 60% of samples as a test set. Two classification methods (naive Bayes (NB) and SVM) were trained on the training samples and predicted users based on test samples and then, the identification accuracy was computed. Following that, the computed identification accuracy was fused with a four-patch local binary pattern-based face recognition system and improved the identification accuracy to more than 91.0%.

Table 6: Comparison of identification accuracy of our proposed system against similar existing methods

Authors	Gesture type	Subjects	Classifiers	IAC (%)
Nigam et al. [44]	Drawing signature	60	NB, SVM, NN	91.4
Chahar et al. [17]	Finger taps	150	NB, RF, NN	81.0
Atas [5]	Hand tremor		NB, MLP	Above 95.0
Wong et al. [18]	Hand gesture	10	ED-FPDI	90.0
Zhao et al. [45]	Simple gesture	4	HMM	91.4
	Complicated gesture			95.2
Imura et al. [20]	Hand gesture	9	Euclidean distance	90.0
Our proposed system	Arbitrary gesture patterns	25	SVM	99.8
	Same gesture patterns			99.4

Chanar et al. [17] proposed a novel user authentication approach using an LM-based sensor. The LM-based device was used to measure a set of string gestures to make a leap password that contained both physiological as well as behavioral information. They collected more than 1700 samples from 150 participants. They extracted the following three features from raw features: (i) measurement of fingers and palms of the users' hand, (ii) time taken between successive gestures, and (iii) similarity measure for gesture passwords. The conditional mutual information maximization-based method was

used to select the potential features. These potential features were fed into three ML-based classifiers (NB, neural network (NN), and RF) for authentication, and computed the performance scores of each classifier. A fusion of scores was obtained to merge information from these three classifiers. They showed that their proposed authentication system achieved an identification accuracy of more than 81.0%.

Atas [5] introduced a hand tremor-based biometric identification system using a 3D LM-based device. They extracted different types of features, including statistical, fast Fourier transform, discrete wavelet transforms, and 1D local binary patterns. Two popular ML-based techniques, like NB and multilayer perceptron (MLP) with 10-fold CV were implemented for user identification and got an identification accuracy of more than 95.0%. Wong et al. [18] also developed a stationary hand gesture-based biometric authentication approach using an LM-based sensor. They utilized stationary hand gestures, like hands, wrists, arms with hands, and fingers. They collected 350 samples from 10 subjects. An edit distance on finger-pointing direction interval (ED-FPDI) was adopted for user authentication and obtained an identification accuracy of 90.0%.

Zhao et al. [45] developed a dynamic gesture-based user authentication system using an LM-based sensor. They used dynamic gestures because dynamic gesture provides more information compared to static. They used two hand gestures such as simple gestures and complicated gestures. They enrolled four subjects and asked each subject to register a gesture (simple and complicated) making it 20 trials. Some features were extracted based on clustering, and then we adopted a Hidden Markov model (HMM) to identify users. HMM provided identification accuracy of 91.4% for simple gestures and 95.2% for complicated gestures. Imura et al. [20] proposed a new gesture-based approach for user authentication using LM-based controllers. They used seven 3D gestures to collect 630 samples from 9 subjects. They got over 90.0% TAR and less than 4% EER. In our work, we used physiological as well as behavioral information of a person that were captured by an LM-based device during the time of performing hand gestures. Since our all of the fingers of our proposed dynamic gesture clearly visible to the LM-based camera, it has the capability to produce more accurately physiological and behavioral features of a person. As a result, our proposed dynamic hand gesture patterns for person identification achieved an identification accuracy of more than 99.0%.

6 Conclusion and Future Work Direction

This work introduced a novel simple dynamic hand gesture-based person identification approach. Our proposed system consisted of several steps: (i) collection of gesture samples using an LM-based sensor by performing two dynamic hand gesture patterns: Arbitrary hand gesture patterns and same hand gesture patterns; (ii) extraction of static and dynamic features from hand geometry; (iii) adoption of RF with based algorithm using Gini Index to sort the features based on classification accuracy and select top 200 features; and (iii) implementation of SVM with RBF model for person identification. The experimental results showed that our proposed system produced an identification accuracy of over 99.0%. Therefore, our invented hand gesture-based person identification system can provide a more reliable identification approach than the traditional identification approach. We expect that this study will provide more evidence of the possibility of person identification. In the future, we will collect more gesture samples by adding more subjects and considering the weight and fat variation conditions to check the precision of our proposed system. We will try to implement different feature selection approaches and compare their performances.

Author Contributions: All listed authors participated meaningfully in the study, and they have seen and approved the submission of this manuscript. Conceptualization, J.S., M.A.M.H., T.W.; Methodology, M.A.M.H., M.M., T.W., J.S.; Data collection and curation: M.A.M.H., T.W., I.J., J.S.; Interpreted and analyzed the data, M.A.M.H., M.M., J.S., T.W., I.J., J.S.; Writing—original draft preparation, M.A.M.H., M.M., T.W., J.S.; Writing—review and editing, M.A.M.H., M.M., J.S., I.J.; Supervision, J.S.; Project administration and funding, J.S.

Availability of Data and Materials: Not applicable.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

[1]   A. Aftab *et al.*, "Hand-based multibiometric systems: State-of-the-art and future challenges," *PeerJ. Comp. Sci.*, vol. 7, no. 5, pp. e1–47, 2021. doi: 10.7717/peerj-cs.707.

[2]   W. Melicher *et al.*, "Usability and security of text passwords on mobile devices," in *Proc. 2016 CHI Conf. Hum. Factors Comput. Syst.*, San Jose, California, USA, 2016, pp. 527–539.

[3]   A. J. Aviv, K. Gibson, E. Mossop, M. Blaze, and J. M. Smith, "Smudge attacks on smartphone touch screens," in *4th USENIX Workshop Offensive Technol. (WOOT 10)*, Washington DC, USA, 2010, pp. 1–7.

[4]   J. Shin, M. A. M. Hasan, and M. Maniruzzaman, "Hand gesture authentication using optimal feature selection and dynamic time warping based k-nearest neighbor," in *2022 5th Int. Conf. Electron., Commun. Control Eng.*, Higashi-Ku, Japan, 2022, pp. 22–26.

[5]   M. Atas, "Hand tremor based biometric recognition using leap motion device," *IEEE Access*, vol. 5, pp. 23320–23326, 2017. doi: 10.1109/ACCESS.2017.2764471.

[6]   A. R. Raoof, "Security of iris recognition and voice recognition techniques," *IBN AL-Haitham J. Pure Appl. Sci.*, vol. 26, no. 1, pp. 346–357, 2017.

[7]   A. K. Jain, S. S. Arora, K. Cao, L. Best-Rowden, and A. Bhatnagar, "Fingerprint recognition of young children," *IEEE Trans. Inform. Forensic. Secur.*, vol. 12, no. 7, pp. 1501–1514, 2016. doi: 10.1109/TIFS.2016.2639346.

[8]   L. Best-Rowden and A. K. Jain, "Longitudinal study of automatic face recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 40, no. 1, pp. 148–162, 2017. doi: 10.1109/TPAMI.2017.2652466.

[9]   P. Andriotis, T. Tryfonas, G. Oikonomou, and C. Yildiz, "A pilot study on the security of pattern screen-lock methods and soft side channel attacks," in *Proc. Sixth ACM Conf. Secur. Privacy Wirel. Mobile Netw.*, 2013, pp. 1–6.

[10] M. Gorman, "Leap motion controller review," 2013. Accessed: Dec. 22, 2022. [Online]. Available: http://www.engadget.com/2013/07/22/leap-motioncontroller-review/

[11] W. Xu, J. Tian, Y. Cao, and S. Wang, "Challenge-response authentication using in-air handwriting style verification," *IEEE Trans. Depend. Secur. Comput.*, vol. 17, no. 1, pp. 51–64, 2017. doi: 10.1109/TDSC.2017.2752164.

[12] D. Lu, K. Xu, and D. Huang, "A data driven in-air-handwriting biometric authentication system," in *2017 IEEE Int. Joint Conf. Biometrics (IJCB)*, Denver, CO, USA, 2017, pp. 531–537.

[13] D. Lu, D. Huang, and A. Rai, "FMHash: Deep hashing of in-air handwriting for user identification," in *ICC 2019-2019 IEEE Int. Conf. Commun. (ICC)*, Shanghai, China, 2019, pp. 1–7.

[14] L. Saritha, D. Thomas, N. Mohandas, and P. Ramnath, "Behavioral biometric authentication using leap motion sensor," *Int. J. Latest Trends Eng. Technol.*, vol. 8, no. 1, pp. 643–649, 2017.

[15] G. Xiao, M. Milanova, and M. Xie, "Secure behavioral biometric authentication with leap motion," in *2016 4th Int. Symp. Digit. Forensic Secur. (ISDFS)*, Little Rock, AR, USA, 2016, pp. 112–11.

[16] M. M. Taher and L. E. George, "A digital signature system based on hand geometry," *J. Algebr. Stat.*, vol. 13, no. 3, pp. 4538–4556, 2022.

[17] A. Chahar, S. Yadav, I. Nigam, R. Singh, and M. Vatsa, "A leap password-based verification system," in *2015 IEEE 7th Int. Conf. Biom. Theory, Appl. Syst. (BTAS)*, Arlington, VA, USA, 2015, pp. 1–6.

[18] A. M. H. Wong and D. K. Kang, "Stationary hand gesture authentication using edit distance on finger pointing direction interval," *Sci. Programming-NETH*, vol. 2016, no. 8, pp. 1–16, 2016. doi: 10.1155/2016/7427980.

[19] S. Kamaishi and R. Uda, "Biometric authentication by handwriting with single direction using self-organizing maps," in *Proc. 11th Int. Conf. Ubiquitous Inf. Manag. Commun.*, Beppu, Japan, 2017, pp. 1–6.

[20] S. Imura and H. Hosobe, "A hand gesture-based method for biometric authentication," in *Int. Conf. Hum.-Comput. Interact.*, Nevada, USA, 2018, pp. 554–566.

[21] T. Manabe and H. Yamana, "Two-factor authentication using leap motion and numeric keypad," in *Int. Conf. Hum.-Comput. Interac.*, Orlando, FL, USA, 2019, pp. 38–51.

[22] J. Malik, A. Elhayek, S. Guha, S. Ahmed, A. Gillani and D. Stricker, "DeepAirSig: End-to-end deep learning based in-air signature verification," *IEEE Access*, vol. 8, pp. 195832–195843, 2020. doi: 10.1109/ACCESS.2020.3033848.

[23] K. Rzecki, "Classification algorithm for person identification and gesture recognition based on hand gestures with small training sets," *Sens.*, vol. 20, no. 24, pp. 1–14, 2020. doi: 10.3390/s20247279.

[24] J. Jung, H. C. Moon, J. Kim, D. Kim, K. A. Toh, "Wi-Fi based user identification using in-air handwritten signature," *IEEE Access*, vol. 9, pp. 53548–53565, 2021. doi: 10.1109/ACCESS.2021.3071228.

[25] E. Guerra-Segura, A. Ortega-Perez, and C. M. Travieso, "In-air signature verification system using leap motion," *Expert Syst. Appl.*, vol. 165, no. 2, pp. 113797–113811, 2021. doi: 10.1016/j.eswa.2020.113797.

[26] M. S. Kabisha, K. A. Rahim, M. Khaliluzzaman, and S. I. Khan, "Face and hand gesture recognition-based person identification system using convolutional neural network," *Int. J. Intell. Syst. Appl. Eng.*, vol. 10, no. 1, pp. 105–115, 2022. doi: 10.18201/ijisae.2022.273.

[27] K. Maruyama, J. Shin, C. M. Kim, and C. L. Chen, "User authentication using leap motion," in *Proc. of Int. Conf. Res. Adapt. Converg. Syst.*, Krakow, Poland, 2017, pp. 213–216.

[28] J. Shin, M. R. Islam, M. A. Rahim, and H. J. Mun, "Arm movement activity-based user authentication in P2P systems," *Peer Peer Netw. Appl.*, vol. 13, no. 2, pp. 635–646, 2020. doi: 10.1007/s12083-019-00775-7.

[29] F. Weichert, D. Bachmann, B. Rudak, and D. Fisseler, "Analysis of the accuracy and robustness of the leap motion controller," *Sens.*, vol. 13, no. 5, pp. 6380–6393, 2013. doi: 10.3390/s130506380.

[30] J. Guna, G. Jakus, M. Pogačnik, S. Tomažič, and J. Sodnik, "An analysis of the precision and reliability of the leap motion sensor and its suitability for static and dynamic tracking," *Sens.*, vol. 14, no. 2, pp. 3702–3720, 2014. doi: 10.3390/s140203702.

[31] M. A. M. Hasan, M. Nasser, S. Ahmad, and K. I. Molla, "Feature selection for intrusion detection using random forest," *J. Inform. Sec.*, vol. 7, no. 3, pp. 129–140, 2016. doi: 10.4236/jis.2016.73009.

[32] F. Jalali-Najafabadi *et al.*, "Application of information theoretic feature selection and machine learning methods for the development of genetic risk prediction models," *Sci. Rep.*, vol. 11, no. 1, pp. 1–14, 2021. doi: 10.1038/s41598-021-00854-x.

[33] H. Naseri, E. O. D. Waygood, B. Wang, Z. Patterson, and R. A. Daziano, "A novel feature selection technique to better predict climate change stage of change," *Sustainability*, vol. 14, no. 1, pp. 1–23, 2021. doi: 10.3390/su14010040.

[34] Y. Qi, "Random forest for bioinformatics," in *Ensemble Mach. Learn.*, pp. 307–323, 2012. doi: 10.1007/978-1-4419-9326-7_11.

[35] L. Breiman, "Consistency for a simple model of random forests. Statistical department," *Technical Report*, University of California at Berkeley, vol. 670," 2004.

[36] L. Breman, "Random forests," *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, 2001. doi: 10.1023/A:1010933404324.

[37] B. H. Menze *et al.*, "A comparison of random forest and its gini importance with standard chemometric methods for the feature selection and classification of spectral data," *BMC Bioinform.*, vol. 10, no. 1, pp. 1–16, 2009. doi: 10.1186/1471-2105-10-213.

[38] E. A. Algehyne, M. L. Jibril, N. A. Algehainy, O. A. Alamri, and A. K. Alzahrani, "Fuzzy neural network expert system with an improved gini index random forest-based feature importance measure algorithm for early diagnosis of breast cancer in Saudi Arabia," *Big Data Cogn. Comput.*, vol. 6, no. 1, pp. 1–29, 2022. doi: 10.3390/bdcc6010013.

[39] J. Wang *et al.*, "Prediction of large vessel occlusion for ischaemic stroke by using the machine learning model random forests," *Stroke Vasc. Neurol.*, vol. 7, no. 2, pp. e001096, 2022. doi: 10.1136/svn-2021-001096.

[40] V. Cherkassky and Y. Ma, "Practical selection of SVM parameters and noise estimation for SVM regression," *Neural Netw.*, vol. 17, no. 1, pp. 113–126, 2004. doi: 10.1016/S0893-6080(03)00169-2.

[41] V. Jaiganesh and P. Sumathi, "Intrusion detection using kernelized support vector machine with leven-bergmarquardt learning," *Int. J. Eng. Sci. Technol.*, vol. 4, no. 3, pp. 1153–1160, 2012.

[42] M. A. M. Hasan, M. Nasser, B. Pal, and S. Ahmad, "Support vector machine and random forest modeling for intrusion detection system (IDs)," *J. Intell. Learn. Syst. Appl.*, vol. 6, no. 1, pp. 45–52, 2014. doi: 10.4236/jilsa.2014.61005.

[43] T. Akiba, S. Sano, T. Yanase, T. Ohta, and M. Koyama, "Optuna: A next-generation hyperparameter optimization framework," in *Proc. 25th ACM SIGKDD Int. Conf. Knowl. Discov. Data Mining*, Anchorage, AK, USA, 2019, pp. 2623–2631.

[44] I. Nigam, M. Vatsa, and R. Singh, "Leap signature recognition using hoof and hot features," in *2014 IEEE Int. Conf. Image Process. (ICIP)*, Paris, France, 2014, pp. 5012–5016.

[45] J. Zhao and J. Tanaka, "Hand gesture authentication using depth camera," in *Adv. Inf. Commun. Netw.: Proc. 2018 Future Inf. Commun. Conf. (FICC)*, Singapore, 2019, pp. 641–654.