aerospace

MDPI

Article

Flight Trainee Performance Evaluation Using Gradient Boosting Decision Tree, Particle Swarm Optimization, and Convolutional Neural Network (GBDT-PSO-CNN) in Simulated Flights

Lei Shang [1], Haibo Wang [1], Haiqing Si [1,*,†], Yonghu Wang [2,3,*,†], Ting Pan [1], Haibo Liu [1] and Yixuan Li [1]

[1] College of General Aviation and Flight, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China; shanglei@nuaa.edu.cn (L.S.); nhwanghaibo@nuaa.edu.cn (H.W.); nuaapanting@nuaa.edu.cn (T.P.); liuhaibo@nuaa.edu.cn (H.L.); lyx_richard@163.com (Y.L.)
[2] School of Aeronautics, Chongqing Jiaotong University, Chongqing 400074, China
[3] Chongqing City Vocational College, Chongqing 402160, China
* Correspondence: sihaiqing@126.com (H.S.); wangyonghu@cqjtu.edu.cn (Y.W.); Tel.: +86-137-7031-7772 (H.S.); +86-023-4721-9765 (Y.W.)
† These authors contributed equally to this work.

Abstract: Flight simulation training is one of the most important methods in early-stage civil aviation flight training. In this regard, flight simulation competitions are effective tools for evaluating the flight skills of trainees. In this study, a model is developed for evaluating the flight skills of trainees by integrating GBDT (Gradient Boosting Decision Tree), PSO (Particle Swarm Optimization), and CNNs (Convolutional Neural Networks). Flight data from simulations is employed for model training. Initially, performance data and scores are gathered from a simulated flight competition platform. The GBDT algorithm is then applied to filter and identify essential flight parameters from the collected data. Subsequently, the PSO-CNN model is utilized to train on the extracted flight parameters. The proposed GBDT-PSO-CNN model achieves a recognition rate of 93.8% on the test dataset. This assessment system is of significant importance for improving the specific maneuvering skill level of flight trainees.

Keywords: flight simulation; flight skills; GBDT; PSO-CNN

Citation: Shang, L.; Wang, H.; Si, H.; Wang, Y.; Pan, T.; Liu, H.; Li, Y. Flight Trainee Performance Evaluation Using Gradient Boosting Decision Tree, Particle Swarm Optimization, and Convolutional Neural Network (GBDT-PSO-CNN) in Simulated Flights. *Aerospace* **2024**, *11*, 343. https://doi.org/10.3390/aerospace11050343

Academic Editor: Konstantinos Kontis

Received: 20 February 2024
Revised: 15 April 2024
Accepted: 23 April 2024
Published: 25 April 2024

1. Introduction

With the rapid advancement of the civil aviation industry, there is a growing demand for expert pilots who bear significant responsibility for ensuring safety [1,2]. Flight trainees, in their early stage of learning to fly, require training to develop their flying skills. The conventional approach to flight training involves utilizing flight simulators for simulation training. Therefore, accurate and efficient evaluation of flight trainees during flight simulation is of significant importance.

Reviewing the literature indicates that numerous investigations have been conducted in the field of flight training using flight data [3,4]. For instance, Kim et al. [5] proposed an improved algorithm for quick access recorder (QAR) data recording and decoding, covering the entire process of conversion, design, development, and validation of data. Cohen et al. [6] proposed a risk evaluation model for aircraft performance that integrates exceedance information, environmental information, and expert experience to assess daily flight performance. Moreover, Wang et al. [7] utilized QAR data and the Analytic Hierarchy Process TOPSIS (AHP-TOPSIS) method to calculate subjective and objective weights to effectively overcome the challenges of incorporating expert knowledge and experience in flight quality evaluation. Gorinevsky et al. [8] applied data mining techniques to distributed fleet monitoring (DFM) and Flight Operations Quality Assurance (FOQA) data collected from a commercial fleet to detect anomalies in aircraft data such as weight and angle of attack, as well as faults and deviations. Gómez et al. [9] introduced an integrated flight data recorder analysis tool, FRiiDA, for military transport aircraft. This tool provides an integrated flight data recorder enabling preventive measures to ensure the quality and safety of aircraft. Yao [10] designed an assessment index system using pilot physiological signals and flight operation parameters. In this regard, a pilot flight training quality assessment model was developed based on the backpropagation (BP) neural network, and the influence of each index on the assessment score was analyzed through the weight distribution of the trained network. The reliability of the model was verified through mathematical analysis. Zhang [11] proposed a BP neural network assessment method for objectively evaluating test pilot piloting skills. The results demonstrated that the BP neural network method improves assessment accuracy and resolves the subjectivity issues of conventional assessment models. Yang [12] established a pilot safety performance indicator preference model for flight school pilots based on grey entropy correlation analysis. The model was applied to analyze the impact of overrun events on safety performance and identify the main factors affecting pilot safety performance indicators.

In terms of flight safety evaluation based on data, numerous studies have been conducted worldwide. For instance, the U.S. Department of Defense [13] established a system for assessing the operational quality of military pilots, based on established rules and standards. Furthermore, Malhotra [14] presented the concept of clearing the Control Laws (CLAW) by Handling Qualities (HQ) assessment using a simulator, discussing challenges and the outcome based on flight test results of the actual air-to-air refueling (AAR) task. Sun [15] demonstrated that the evaluation of flight maneuvering quality level is a crucial metric for the development of civil aircraft flight control systems. Cooper and Harper [16] introduced the Cooper–Harper aircraft handling qualities scale, which rates the operational skills of pilots based on the mastery of operational control and stability. Among various indicators for assessing the handling skills of pilots, this scale focuses on the optimal choices chosen by pilots. Payne and Harris [17] selected 12 key technical indicators and developed a pilot operational skill evaluation system to evaluate pilot operational skills in five aspects, including lift control, tilt control, balance ability, yaw control, and speed control. Finally, the system generates a 10-level skill rating for evaluating behavioral parameters at the operational level. You Xuqun et al. [18] analyzed the requirements for pilot work, proposed a hypothetical model for evaluating the technical skills of pilots, validated the proposed model, and finally derived a four-dimensional evaluation model for airline flight skills to comprehensively evaluate the operational skills of pilots. Liu Zhongqi [19] utilized pilot eye movement data to develop a BP neural network model for predicting and evaluating the flight performance of pilots with different skill levels during training and simulation. Mickael Rey et al. [20] proposed a data-driven approach for classifying the safety or risk of civil aviation flights using data analysis methods and machine learning tools.

The performed literature survey indicates that most investigations in this field have focused on flight data, pilot physiological [21,22] and subjective data, and objective scales to assess the flight quality of civil aviation pilots. However, there is no comprehensive understanding of the flight quality of civil aviation pilots during the initial stages of flight training. This issue is particularly more pronounced during flight simulation. The present study introduces a novel approach to evaluating flight performance utilizing GBDT-PSO-CNN. The assessment is based on scoring data from simulated flights, integrated with real-time flight data. This combined approach ensures a precise evaluation of the pilot's maneuvering performance, contributing to the enhancement of flight skill levels among trainees.

The terminology used in the literature in the introduction section is shown in the Table 1 below:

Table 1. Nomenclature of introductory terms.

Nomenclature	
QAR	Quick access recorder, used to record various parameters during aircraft flight.
AHP-TOPSIS	An AHP-TOPSIS integrated evaluation method
DFM	Distributed fleet monitoring
FOQA	Flight Operations Quality Assurance
BP	Backpropagation, this is a feedforward neural network based on the error backpropagation algorithm.
CLAW	Control Laws
HQ	Handling Qualities
AAR	Air-to-Air Refueling

2. Methods and Materials

The experimental process involves the utilization of a flight simulation trainer to provide a training experience for flight trainees. Data collected from the flight simulation trainer was utilized to organize the National Collegiate Flight Simulation Championship, encompassing a wide range of challenging scenarios. This championship provides a platform for students specializing in flight techniques to enhance their technical proficiency and gain practical experience in performing complex flight operations. Moreover, it offers participants a valuable opportunity to develop essential practical skills that are crucial to their future careers in aviation.

2.1. Participants

In this research, a total of forty male students who are pursuing a degree in flight technology from the College of General Aviation and Flight at Nanjing University of Aeronautics and Astronautics (NUAA) participated in the tests. Participants were aged from 20 to 22 years, with a mean of 21.2 years (SD = 0.68). The flight trainees participating in the experiment were members of the school's flight simulation team, each with prior experience of over 40 h in flight simulation training. Additionally, these trainees were familiar with simulation software such as P3D and Xplane. All participants received training before the flight simulation.

The flight trainees admitted to the Flight Technology Program at NUAA undergo a rigorous selection process. They are admitted based on a physical examination conducted by the Civil Aviation Administration of China (CAAC) and the examination of higher education institutions. Upon admission, a flight trainee must progress through three key stages to become a civilian pilot as follows:

(1) Theoretical study phase at Nanjing University of Aeronautics and Astronautics;

(2) Sending pilot trainees to CCAR-141 flight schools approved or recognized by the Civil Aviation Administration of China for flight license training;

(3) After completing their flight license training, flight trainees return to the university to continue their relevant course work, theoretical training, and examination for the Airline Transport Pilot License, and defend their undergraduate graduation design (thesis).

After completing the aforementioned stages, the flight trainees are eligible to join corresponding airlines as official civil aviation pilots. Those who fulfill the graduation requirements receive an undergraduate diploma and bachelor's degree certificate in flight technology from NUAA.

2.2. Simulator and Scenarios

2.2.1. Experimental Equipment

The flight trainer is equipped with several advanced features, including a mechanical joystick and foot rudder, providing a realistic and tactile flight control experience. Moreover, it consists of a flight operation panel that integrates various components such as the ignition key, light switch panel, throttle lever, fuel–air mixing lever, leveling wheels, flaps, and fuel tank selector switch. These components control the simulated flight operations, ensuring a comprehensive and immersive training experience for trainees. The flight simulation equipment is depicted in Figure 1.

Figure 1. Flight simulation equipment.

Additionally, the simulation flight trainer is equipped with an integrated avionics system (GARMIN 1000, Kansas City, MO, USA) and a primary flight display (PFD) panel. This avionics system can accurately monitor instruments and navigation data during the flight simulation. The designed avionics system helps flight trainees in developing their skills during simulated flight scenarios. Overall, the incorporation of these advanced features within the flight trainer enables trainees to acquire hands-on experience with the operational aspects of flight controls and avionics systems typically used in aircraft.

2.2.2. Experimental Flight Subject

Based on the comprehensive literature survey and consultation with senior flight instructors, this project selected airfield traffic patterns for flight simulation. It is worth noting that the airfield traffic pattern is a vital component of pilot training, as it involves flying maneuvers around airports. In this training scenario, pilots acquire skills such as takeoff, climb, turning, leveling off, descent, and landing. Figure 2 illustrates the schematic of an airfield traffic pattern.

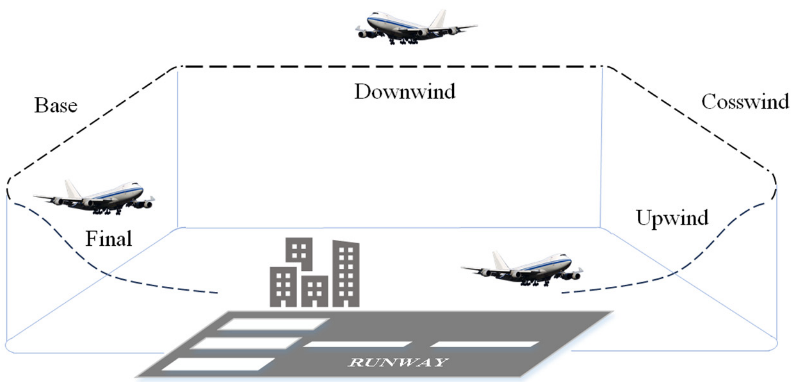

Figure 2. Schematic of an airfield traffic pattern.

In order to evaluate the operational accuracy of the flight simulation, the operational accuracy score of each pilot throughout the flight was obtained based on the simulated airfield traffic pattern rating scale, with a full item score of 98. This score not only reflects the pilot's operational level for given scenarios of a flight but also enables training focusing on the weak points.

2.3. Procedure

The experiment was conducted using a Cessna C172SP Skyhawk airplane. The tests were carried out on runway 19 at Beijing Capital International Airport. During the tests, the sky was clear with increasing side winds between the ground and 3000 feet, ranging from 5 to 15 knots. To prevent competitors from using abnormal operational methods, the following actions should be implemented: disabling the auto rudder in hard mode (which controls the course), enabling the display flight tips, activating the auto mixer (to adjust the mixture ratio of oil and gas), and enabling the spiral effect (which is a side effect of the propeller, causing left yaw moment due to the propeller turning to the right, particularly noticeable at low speed). Other options and functions were disabled during the test. With the evaluation software turned on, the pilot flew the airplane through an airfield traffic pattern to the left before landing and stopping on the runway. The maximum flight time was 12 min, and the scoring software discarded the results if the flight time was exceeded or if pilots slipped off the runway. If the system determined that the aircraft was damaged, or it could not operate normally due to structural damage, the competition result was discarded and a score of 0 points was recorded. The scoring items and codes are presented in Table 2.

Table 2. Airfield traffic pattern scoring items.

Project Code	Project Name	Standard Values	Unit	Score
A	Track deviation on upwind	≤1	°	5
B	Lift wheel gauge speed	55	knots	5
C	Maximum climbing rate on upwind	≤500	ft	5
D	First turn height	≥800	ft	5
E	First turn gradient	≤30	°	5
F	Track deviation on crosswind	0	°	5
G	Second turn slope	≤30	°	5
H	Track deviation on downwind	0	°	5
I	Height deviation on downwind	0	ft	5
J	Third turn gradient	≤30	°	5
K	Fourth turn gradient	≤30	°	5
L	Approach to the maximum rate of decline	≥−500	ft	5
M	Approach track deviation	≤50	m	6
N	Runway entrance height (radio height)	≤50	ft	6
O	Ground position deviation	0	m	−10
P	Grounding rate	≥−40	ft	8
Q	Grounding overload	≤1.2	G	6
R	Distance between landing glide and centerline	≤0.2	m	6
S	Maximum flight overload	≤1.5	G	6

3. Methodology

The present study proposes a new hybrid algorithm for the feature selection of civil aviation flight training, which incorporates the gradient boosting decision tree and multilayer perceptron. First, the training data are screened to select feature importance using the gradient boosting decision tree (GBDT) model. Then, the screened features are introduced to the multilayer perceptron model for training. Finally, flight training data are classified for each trainee.

3.1. Gradient Boosting Decision Tree Algorithm

GBDT is an integrated learning model, which is classified as a decision tree algorithm based on iterative accumulation. This model constructs a set of weak learning trees and accumulates the results of multiple decision trees as the final prediction result. The role of integrated learning is to enhance these multiple weak learners into one strong learner to minimize error rates. A weak learner typically refers to a single decision tree model constructed in each iteration, while a strong learner refers to the decision tree model constructed at each stage. The model structure of integrated learning is shown in Figure 3:

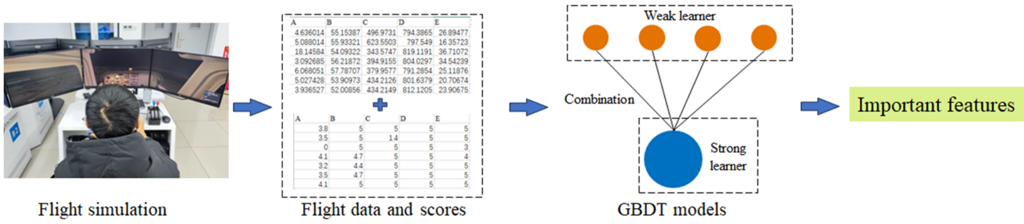

Figure 3. Structure of integrated learning.

When establishing a GBDT model, an initialized regression tree is initially trained on the training dataset, and more regression trees are generated by iterating on this base. Finally, all data are combined into a strong learner.

The main steps of the GBDT model construction are as follows:

(1) Weak learner initialization

$$f_0(x) = \text{argmin}_c \sum_{i=1}^{N} L(y_i, c) \tag{1}$$

(2) Calculate the negative gradient for each sample

$$r_{im} = -\frac{\partial L(y_i, f(x_i))}{\partial f(x_i)}, i = 1,2,3,\ldots,N \tag{2}$$

(3) The negative gradient obtained in step (2) is used as the new true value of the sample, and $(x_i, r_{im}), i = 1,2,3,\ldots,N$ is used as the training data for the next tree. Then, a new regression tree $f_m(x)$ is obtained and its corresponding leaf node region is $R_{jm}, j = 1,2,3,\ldots,J$, where J is the number of leaf nodes of regression tree t.

(4) Calculate the best-fit value for the leaf region

$$g_{jm} = \text{argmin}_g \sum_{xi \in R_{jm}} L(y_i, f_{m-1}(x_i) + g) \tag{3}$$

(5) Update to the strong learner

$$f_m(x) = f_{m-1}(x) + \sum_{j=1}^{J} g_{jm} I(x \in R_{jm}) \tag{4}$$

(6) Finally, obtain a strong learner

$$f(x) = f_M(x) = f_0(x) + \sum_{m=1}^{M} \sum_{j=1}^{J} g_{jm} I(x \in R_{jm}) \tag{5}$$

3.2. PSO-CNN Model

3.2.1. Particle Swarm Optimization

Particle swarm optimization (PSO) originates from the observation of bird flock feeding behavior, and its main idea involves seeking the optimal solution through collaboration and information sharing among individuals within a group. The initial state of PSO is a group of random particles. Through iterative processes, the algorithm refines and converges towards the optimal solution. PSO can be mathematically expressed in the form below:

$$v_i = v_i + c_1 \times rand(\ ) \times (pbest_i - x_i) \\ +c_2 \times rand(\ ) \times (gbest_i - x_i) \quad (6)$$

$$x_i = x_i + v_i \quad (7)$$

where i = 1,2,3. . .N, N is the total number of particles; v_i is the velocity of the particle; rand() is a random number between (0,1); x_i is the current position of the particle; c_1 and c_2 are the learning factors, which are usually set to 2 [21]; and pbest and gbest are the two extremes that the particle needs to track.

The particle determines its next movement based on its individual experience and insights gained from the best-performing peers. The combination of these principles constitutes the standard form of PSO.

3.2.2. Convolutional Neural Network

A convolutional neural network is a type of feedforward neural network with a network computing function, which mainly comprises the convolutional layer, pooling layer, and fully connected layer.

3.3. GBDT-PSO-CNN Model

The present study performs feature selection on flight training data using the GBDT-PSO-CNN hybrid model. In the model, the important features in flight training data by GBDT are trained and predicted by PSO-CNN. The structure of the model is presented in Figure 4.

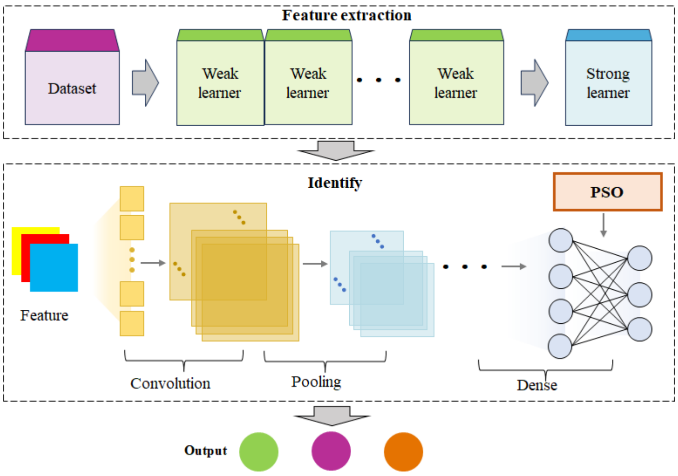

Figure 4. GBDT-PSO-CNN hybrid model.

In the GBDT-PSO-CNN model, feature importance selection is initially performed using the GBDT model through the following steps:

(1) Calculate the reduction in weighted impurity of all non-leaf nodes at splitting. The higher the reduction degree, the higher the importance of the feature.

(2) The reduction in weighted impurity is the gain of that node for this split. Therefore, the greater the benefit of node splitting, the higher the importance of the corresponding characteristics of the node.

After selecting the important features, they are trained using a PSO-CNN model. The convolutional layer contains multiple convolutional kernels, which cover an area called the "sensory field"; the pooling layer is used for feature selection to reduce the number of features in the input data; and the fully connected layer is used to nonlinearly fit the extracted features to the output.

In summary, the model introduced in this study initially filters out the important features of the data using the GBDT algorithm and subsequently employs a convolutional neural network optimized by the PSO algorithm to classify the target values. Applying the GBDT-PSO-CNN hybrid model effectively reduces the number of features of the data, thereby improving faster training and more accurate predictions.

4. Analysis

4.1. Raw Data

During the tests, each participant undertook three airfield traffic pattern experiments. Prior to the experiments, participants received a verbal explanation regarding the study's purpose, as well as an introduction to the nature, objectives, and relevant flight rules of the experiment.

Forty trainees participated in the experiments and a total of 114 pieces of appropriate data was gathered. It should be indicated that inappropriate data were mainly generated in experiments, in which the aircraft crashed, ran off the runway, and timeouts occurred. Among the recorded data, the highest and lowest total scores, and standard deviation were 89.1, 42.2, and 11.15, respectively. A portion of the collected data is presented in Table 3.

Table 3. A portion of raw data.

No.	A	B	C	D	E	F	G	H	I	J
1	4.64	55.15	496.97	794.39	26.89	7.45	18.43	9.11	63.32	13.09
2	5.09	55.93	623.55	797.55	16.36	7.11	22.44	5.84	48.23	21.58
3	18.15	54.09	343.57	819.12	36.71	19.67	27.99	10.64	53.14	13.97
4	3.09	56.22	394.92	804.03	34.54	24.33	13.89	6.43	62.12	20.98
5	6.07	57.79	379.96	791.29	25.12	5.87	17.40	7.06	63.97	17.75
No.	**K**	**L**	**M**	**N**	**O**	**P**	**Q**	**R**	**S**	**total points**
1	10.96	−622.61	44.62	28.64	218.20	−245.06	1.07	5.29	1.40	77.4
2	32.76	−403.50	15.14	33.38	18.12	−124.04	1.25	8.56	1.88	78
3	16.64	−742.02	162.90	68.45	32.71	−403.13	2.72	2.78	1.59	86.9
4	19.85	−1028.97	132.73	37.67	13.27	−403.41	1.15	5.81	1.48	76
5	14.68	−571.95	94.66	64.91	12.06	−111.29	1.34	2.72	1.78	72.8

The units for the various data categories in Table 3 are listed in Table 2.

4.2. Data Pre-Processing

Data Analysis

The initial analysis of the collected raw flight data and scores is essential to examine the correlation between various types of flight data and the overall scores.

In order to better analyze the correlation between the data, a correlation analysis was performed on airfield traffic pattern data. Figure 5 shows a correlation coefficient, which is a measure that falls within the range of $(-1, 1)$.

In the correlation graph, the lighter the color, the stronger the positive correlation between the two pieces of data, while a darker color reflects a stronger negative correlation. The last row shows the correlation between each score and the total score. It is observed that the total score has the strongest positive correlation with I, L, M, N, and O.

The 114 pieces of raw data had a wide range of total scores. In order to better use the hybrid model for processing and prediction, it is necessary to classify the data according to the total score. Experiments reveal that in the simulated airfield traffic pattern test, a total score of 65 points is an important threshold for distinguishing between acceptable and non-acceptable levels of skills. The conclusions are derived from simulating the training process of the flying team and the selection of new members. Additionally, this scoring system helps to distribute the data evenly. It is worth noting that the higher the overall score, the more skillful the flight student. Meanwhile, a score of 75 points is the median value of the collected data, which is a crucial index in training and prediction.

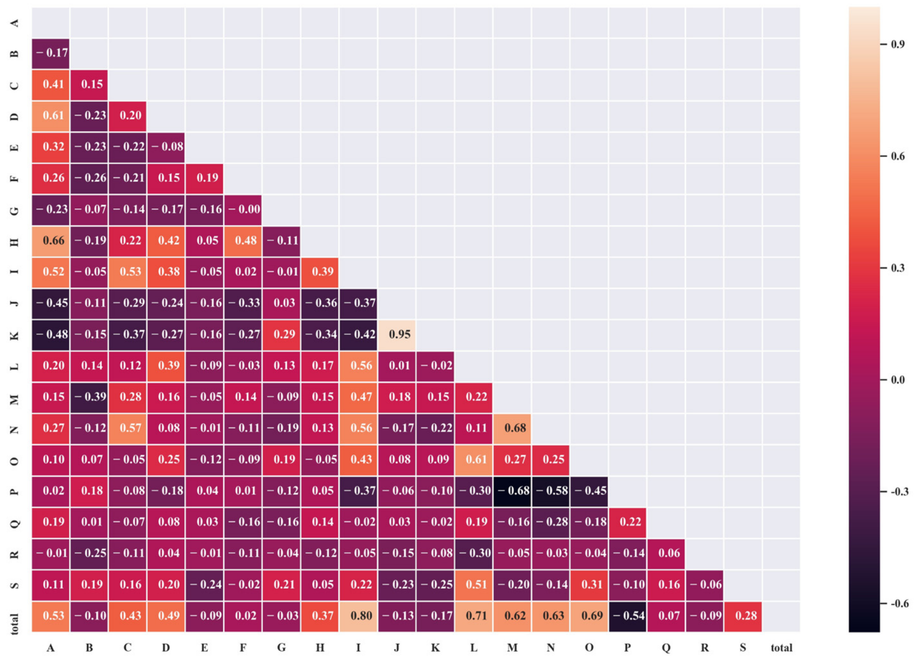

Figure 5. Correlation coefficient graph.

4.3. Data Analysis Using the GBDT-PSO-CNN Hybrid Model

4.3.1. Software Environment

The training process of the network model utilizes the Keras architecture based on the TensorFlow deep learning framework to establish a runtime environment. The parameters for the experimental setup are specified as follows: CPU, Intel i5-10400f 3.2 GHz; operating system, Win10 64-bit; programming language, Python 3.7.7; deep learning architecture, TensorFlow 2.3.0 and Keras 2.4.3.

4.3.2. Data Training and Prediction

The hybrid model initially derives feature importance indicators through a GBDT, and indicators are ranked according to importance in Table 4:

Table 4. Importance ranking of scoring indicators.

Order of Importance	Project Code	Project Name
1	M	Approach track deviation
2	O	Ground position deviation
3	L	Approach to the maximum rate of decline
4	C	Maximum climbing rate on upwind
5	A	Track deviation on upwind

The top three rows are indicators with feature importance values higher than 0.1. Accordingly, M, O, and L are selected as the feature indicators for the next input in the hybrid model.

In this article, the PSO-CNN network undergoes a sequence of four convolutional and pooling operations. To prevent overfitting, a Dropout layer is incorporated. The final architecture includes three fully connected layers (Dense) to produce the model output. The activation function for each convolutional layer and the first two fully connected layers is ReLU, while the activation function of the last fully connected layer is Softmax. Meanwhile, the number of neurons in the second fully connected layer is optimized using PSO to obtain the optimal model parameters.

In the PSO optimization process, the optimal neuron interval is set to [32, 64], with a maximum of 1 iteration, 2 particles, and 300 training cycles during the optimization.

Following 4 optimization cycles, the optimal number of neurons in the second fully connected layer is determined to be 38. The structure and parameters of the whole network are detailed in Table 5.

Table 5. PSO-CNN model parameters.

Layers	Output Size	Parameters
Conv1D	(None, 3, 64)	768
Conv1D	(None, 3, 64)	45,120
MaxPooling	(None, 1, 64)	0
Dropout	(None, 1, 64)	0
Conv1D	(None, 1, 128)	90,240
Conv1D	(None, 1, 128)	180,352
MaxPooling	(None, 1, 128)	0
Dropout	(None, 1, 128)	0
Conv1D	(None, 1, 256)	360,704
Conv1D	(None, 1, 256)	721,152
MaxPooling	(None, 1, 256)	0
Dropout	(None, 1, 256)	0
Conv1D	(None, 1, 512)	1,442,304
Conv1D	(None, 1, 512)	2,884,096
Dropout	(None, 1, 512)	0
Global average pooling	(None, 512)	0
Dropout	(None, 512)	0
Dense	(None, 256)	131,328
Dropout	(None, 256)	0
Dense	(None, 38)	9766
Dense	(None, 3)	117
Total	/	5,865,947

Figure 6 illustrates the model structure.

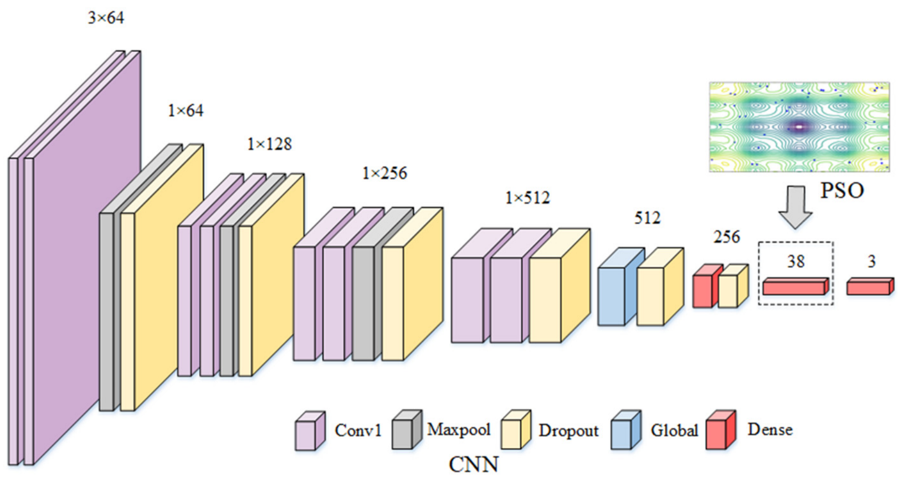

Figure 6. Definition of the hybrid model.

The model utilizes the RMSprop optimized function and the hyperparameters are set as follows: learning rate = 0.001, batch = 16, and epochs = 300. Meanwhile, the model divides the dataset into training and test datasets with a ratio of 4:1. Figure 7 shows the accuracy of the GBDT-PSO-CNN model and the GBDT-CNN model on the train and test datasets.

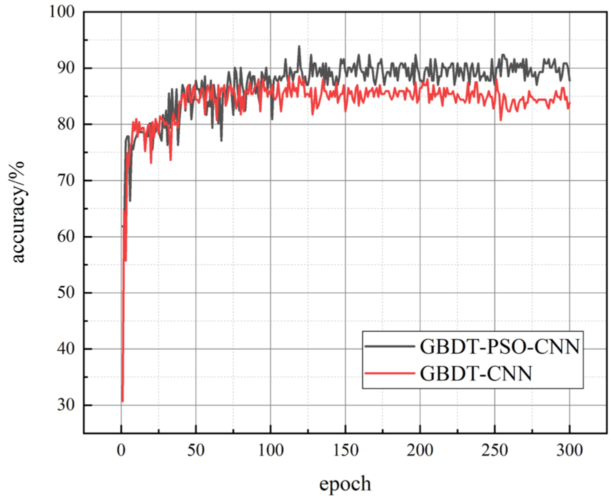

Figure 7. Accuracy of the two models.

Figure 8 reveals that the GBDT-PSO-CNN model has a higher recognition rate on the test set compared with the GBTD-CNN algorithm; the highest recognition rates of the two models are 93.8% and 88.5%, respectively. It is observed that the GBDT-PSO-CNN model has higher recognition accuracy.

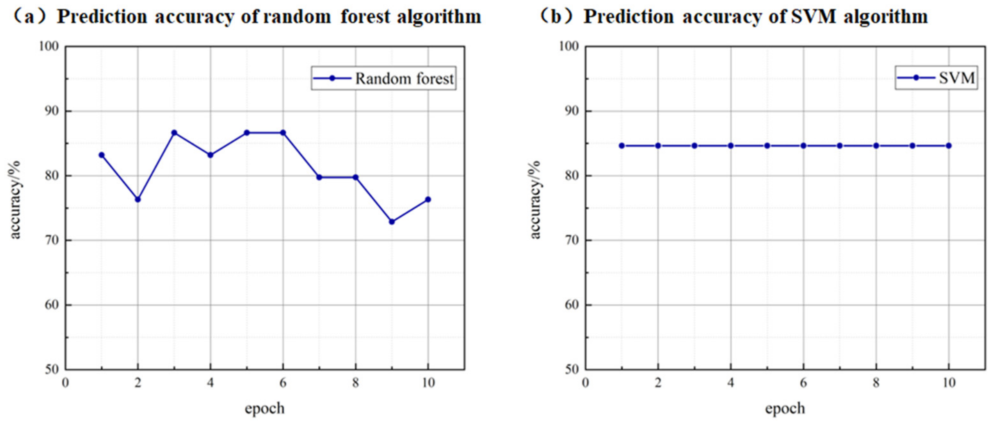

Figure 8. Prediction accuracy of random forest and support vector machine algorithms on the collected dataset.

4.3.3. Analysis of Other Algorithms

To verify the reliability of the hybrid model, the random forest algorithm and the support vector machine (SVM) algorithm, which are two widely used machine learning algorithms, were selected for comparative validation. These algorithms were used to train and predict data after a 10-time binary classification process. Figure 9 shows the prediction accuracy of these models.

The comparison results of each algorithm are presented in the following Table 6.

It is observed that the prediction accuracy of the random forest algorithm ranges from 72.9% to 86.7% with an average accuracy of 82%. On the other hand, the prediction accuracy of the support vector machine algorithm is maintained at 84.7%. The results demonstrate that both algorithms exhibit high prediction accuracy on the prepared dataset. In comparison, the hybrid model proposed in this paper has higher recognition accuracy.

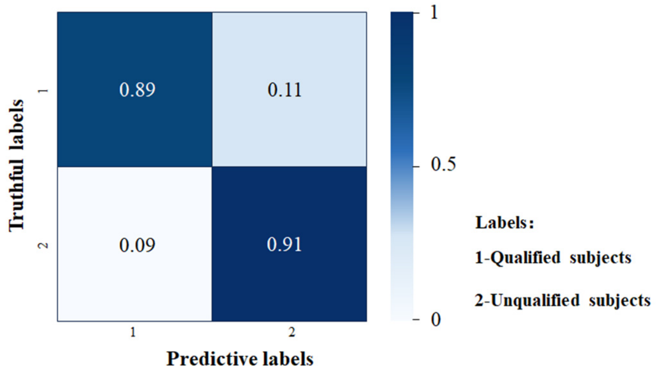

Figure 9. Confusion matrix.

Table 6. Comparison of algorithms.

Algorithms	Maximum Accuracy
GBDT-PSO-CNN	93.8%
GBDT-CNN	88.5%
Random Forest	86.7%
SVM	84.7%

4.3.4. Verification and Validation of Hybrid Model

In order to deeply verify the accuracy of the model, we collected some additional data to verify and validate the model. We searched for five flight trainees majoring in flight technology according to the same criteria, and also recruited five graduate students in transportation engineering from the College of General Aviation and Flight. Participants were aged from 20 to 26 years, with a mean of 22.6 years (SD = 1.86). Following five hours of flight simulation training, all of the graduate students included in the experiment were able to successfully complete the simulated flights in accordance with the requirements. Each of the 10 subjects performed two complete simulated flights and collected flight data and scoring data. After the extraction of the key indicators from the 20 flight datasets, the flight scoring data were processed accordingly, and some flight characteristics data are shown in Table 7:

Table 7. Partial flight characterization data.

No.	A	C	L	M	O	Total
1	5.575437	439.1993	−1565.07	56.84234	389.4412	53.5
2	2.635527	518.8191	−849.453	45.70391	696.9078	50.7
3	5.990701	492.3047	−657.651	38.27551	689.6252	66.2
4	6.137057	450.0083	−1097.32	156.3077	320.2329	68.4
5	2.929803	512.7805	−530.846	35.36362	398.5504	73.1

The flight data that were processed were inputted into the hybrid model proposed in this study for prediction, resulting in the acquisition of the confusion matrix depicted in Figure 9.

The results of the confusion matrix show that the recognition accuracy of the hybrid model on the new data reaches 90%, which can effectively screen out the flight trainees who meet the requirements, and verifies the reliability and accuracy of the proposed model.

Meanwhile, statistical analysis was performed on the significant indicators found in the individuals' flight data, and the statistical results shown in Table 8 were obtained:

Table 8. Statistics on key flight indicators.

Category	Statistical Term	A	C	L	M	O
Qualified subjects	Mean	4.017794	492.3807	−623.144	122.6385	379.2011
	Max	6.137057	609.8047	−459.882	228.3594	689.6252
	Min	1.877831	424.8352	−1097.32	16.09314	304.8896
Unqualified subjects	Mean	7.037137	521.1375	−997.435	591.6879	689.0239
	Max	24.14047	760.1803	−545.219	5001.496	2526.933
	Min	2.458178	410.5572	−2255.74	16.37023	293.9267

Through statistical analysis, it can be found that qualified and unqualified subjects have significant differences in the flight data of important features, thus verifying the accuracy of this paper for flight data feature extraction. Additionally, the deficiencies of the subjects may be more precisely ascertained by identifying subpar individuals and integrating this with the flight data.

Overall, for flight trainees participating in flight simulation competitions, they can be evaluated faster through important flight characteristics by using the hybrid model proposed in this paper. At the same time, when the result of a particular flight simulation is not satisfactory, the weak flight link can be found and targeted for training and development.

5. Conclusions

In this research, a hybrid GBDT-PSO-CNN model is proposed. Initially, the GBDT algorithm was employed to extract features from the flight data related to the airport traffic patterns in simulated flights. By analyzing 114 simulated flight data points, crucial features of the airport traffic patterns were identified. Subsequently, a convolutional neural network optimized by the PSO algorithm was used to train the processed data. The recognition accuracy on the test dataset was 93.9% after 300 training cycles on the training set. This performance outperforms similar algorithms in terms of recognition accuracy and resilience, effectively assessing the trainees' flight simulation competency. The model proposed in this paper mainly implements the following features:

(1) The model has the ability to evaluate the flying performance of flight trainees using flight data with significant features, allowing for the rapid and accurate selection of appropriate flight trainees.

(2) The study's findings offer empirical evidence to improve trainees' simulated flight training, and by using the methodology proposed in this study, weaknesses in flight can be quickly and effectively identified for targeted training.

The hybrid model proposed in this study exhibits promising performance for effectively screening suitable flight trainees such as recruiting new members for a flight simulation team or selecting participants for flight simulation competitions. Through the proposed method, flight trainees' simulation performance weaknesses may be found, enabling more focused instruction. This identification approach effectively evaluates the flight skills of trainees, which in turn guarantees the security and dependability of their subsequent flights.

Author Contributions: Methodology, L.S. and H.W.; investigation, validation, formal analysis, and writing—original draft, L.S.; software, H.W. and T.P.; writing—review and editing, H.W.; funding acquisition, H.W., H.S. and Y.W.; supervision and project administration, H.S. and Y.W.; investigation and editing, H.L.; editing, Y.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Joint Fund of National Natural Science Foundation of China and Civil Aviation Administration of China (No. U2033202), the Fundamental Research Funds for the Central Universities (No. NS2022094), the first batch of industry–university–research cooperative collaborative education projects of the Ministry of Education in 2021 (No. 202101042005), and the Experimental technology research and development project of Nanjing University of Aeronautics and Astronautics Project (SYJS202207Y), the Nanjing University of Aeronautics and Astronautics PhD short-term visiting scholar project (No. ZDGB2021024), the Special Key Projects of Technological Innovation and Application Development of Chongqing (grant no. cstc2019jscx-fxydX0036), and the Research and Innovation Program for Graduate Students in Chongqing (2022S0085).

Data Availability Statement: The data used to support the findings of this study are available from the corresponding author upon request.

Acknowledgments: All authors agreed to publish this study, and thank all authors for their contributions. We are grateful to the editors and the reviewers for their insightful comments and suggestions.

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

References

1. Li, Y.; Chen, H.; Xin, X.; Ji, M. The influence of mindfulness on mental state with regard to safety among civil pilots. *J. Air Transp. Manag.* **2020**, *84*, 101768. [CrossRef]
2. Du, H.B.; Zhang, Q.H. Statistics and Causes Analyses of Flight Accidents. *Saf. Environ. Prot.* **2016**, *42*, 17–20.
3. Chen, N.T.; Li, J.H.; Sun, Y.C.; Wang, G.X. Progress of research on the application of flight data in the field of aerospace ergonomics. *Chin. J. Ergon.* **2020**, *26*, 81–86.
4. Ebbatson, M.; Harris, D.; Huddlestone, J.; Sears, R. The relationship between manual handling performance and recent flying experience in air transport pilots. *Ergonomics* **2010**, *53*, 268–277. [CrossRef] [PubMed]
5. Kim, J.H.; Lyou, J. Enhanced QAR Flight Data Encoding and Decoding Algorithm for Civil Aircraft. In Proceedings of the International Joint International on SICE-ICASE IEEE, Busan, Republic of Korea, 18–21 October 2007; pp. 5169–5173.
6. Cohen, B.; Cassell, R.; Smith, A. Development of an aircraft performance risk assessment model. In Proceedings of the Digital Avionics Systems Conference IEEE, St. Louis, MO, USA, 24–29 October 1999; Volume 1, pp. 5.A.2-1–5.A.2-8.
7. Wang, B.; Du, J.; Ding, C.; Wu, Z.; Zhang, S. Flight quality evaluation of aircraft take-off stage based on AHP-TOPSIS method. *Flight Dyn.* **2019**, *37*, 80–84+88.
8. Matthews, B.L.; Gorinevsky, D.; Martin, R. Aircraft anomaly detection using performance models trained on fleet data. In Proceedings of the Intelligent Data Understanding (CIDU), 2012 Conference on IEEE, Boulder, CO, USA, 24–26 October 2012; pp. 17–23.
9. Gomez, Z.C.; Lora, E.M.; Mestre, A.B. FRiiDA: An integrated flight data recorder analysis tool for Airbus Defence and Space. In Proceedings of the 2015 IEEE International Conference on Industrial Technology (ICIT) IEEE, Seville, Spain, 17–19 March 2015; pp. 2004–2009.
10. Yao, Y.; Kaijun, X.U. Quality assessment of flight training based on BP neural network. *Acta Aeronaut. Astronaut. Sin.* **2017**, *38*, 24–32.
11. Zhang, T.B.; Ti-Fang, L.I.; Sheng, Y.W. Test Pilot Driving Skill Assessment Based on BP Neural Network. *Comput. Simul.* **2012**, *29*, 110–113.
12. Yang, X.Q.; Lin, X.Y.; Jiang, L.; Lu, Z.C.; Cao, H. Research on the Influence of Aviation School Pilot's Safety Performance Index Based on Grey Entropy Correlation Analysis. *J. Civ. Aviat. Flight Univ. China* **2021**, *32*, 33–37.
13. US Department of Defense. *Flying Qualities of Piloted Airplanes (MIL-F-8785C)*; US Department of Defense: Washington, DC, USA, 1980.
14. Malhotra, M.K.; Rohith, M.; Myala, J.; Patel, V. Perfecting Handling Qualities Assessment Task on Simulator and Outcome of Actual Flight Test for Air to Air Refueling. In Proceedings of the 2019 Sixth Indian Control Conference (ICC), Hyderabad, India, 18–20 December 2019.
15. Sun, S.D.; Qiu, Z.F.; Sun, Y.C.; Zhang, Y.J.; Guo, Y.Y. Research on evaluation of handling qualities rating method for civil aircraft. *Flight Dyn.* **2019**, *37*, 83–87.
16. Cooper, G.E.; Harper, R.P. The use of pilot rating in the evaluation of aircraft handling qualities. *Epigenetics Off. J. Dna Methylation Soc.* **1969**. [CrossRef]
17. Payne, K.H.; Harris, D. A psychometric approach to the development of multidimensional scale to assess aircraft. *Int. J. Aviat. Psychol.* **2000**, *10*, 343–362. [CrossRef]
18. You, X.Q.; Yan, B.H.; Li, Y.; Gu, X.H.; Yang, S.Y.; Tu, J.L. The Effects of Flight Management Attitude on the Standardization of the Navigation Behavior of Airline Pilots. *Acta Psychol. Sin.* **2008**, *40*, 466–473. [CrossRef]
19. Liu, Z.Q.; Yuan, X.G.; Fan, Y.B. Pilot performance evaluation model based on BP neural networks. *J. Beijing Univ. Aeronaut. Astronaut.* **2010**, *36*, 403–406.
20. Rey, M.; Aloise, D.; Soumis, F.; Pieugueu, R. A data-driven model for safety risk identification from flight data analysis. *Transp. Eng.* **2021**, *1*, 100087. [CrossRef]

21. Pan, T.; Wang, H.; Si, H.; Liu, H.; Xu, M. Research on the Identification of Pilots' Fatigue Status Based on Functional Near-Infrared Spectroscopy. *Aerospace* **2022**, *9*, 173. [CrossRef]
22. Amodio, A.; Ermidoro, M.; Maggi, D.; Formentin, S.; Savaresi, S.M. Automatic Detection of Driver Impairment Based on Pupillary Light Reflex. *IEEE Trans. Intell. Transp. Syst.* **2018**, *20*, 3038–3048. [CrossRef]