Nanotechnology Perceptions
ISSN 1660-6795
www.nano-ntp.com

Machine Learning-Based Spatial Disorientation Detection In Rotary-Wing Aircraft

Yim Sehoon[1], Kim Gyuwang[2], Kim Sejun[3], Park Cheol[4], Cho Youngjin[5*]

[1]*Professor, Department of Helicopter Operation, Hanseo University, Taean 32158, Korea; E-mail: SHprof@hanseo.ac.kr*
[2]*Professor, Department of Flight Operation, Hanseo University, Taean 32158, Korea; E-mail: kimchi1230@hanseo.ac.kr*
[3]*Assistant Professor, Department of Flight Operation, Hanseo University, Taean 32158, Korea; E-mail: junkim0118@hanseo.ac.kr*
[4]*Associate Professor, Department of Helicopter Operation, Hanseo University, Taean 32158, Korea; E-mail: heli69@hanseo.ac.kr*
[5*]*Associate Professor, Department of Helicopter Operation, Hanseo University, Taean 32158, Korea*
E-mail: speedshock@hanseo.ac.kr
[*]*Corresponding author*

Despite considerable advancements in aviation technology and systemic improvements in aviation safety management, accidents caused by spatial disorientation in helicopters continue to occur. A flight illusion refers to a phenomenon wherein a pilot misjudges the aircraft attitude. Most spatial disorientations fall under Type I, wherein the pilot is unaware of the illusion. This has been a significant challenge for spatial-disorientation detection through physiological responses in past studies. Therefore, this study developed a new machine-learned-based tool for detecting spatial disorientation in helicopters by collecting flight data from a simulator mimicking real flight conditions. The collected data were subjected to feature selection and labeling during preprocessing, which were then used for training four machine learning models capable of classifying spatial-disorientation occurrences. The results indicated that the random forest model demonstrated the best performance. The tool developed in this study has considerable potential for application in real operational settings and can be integrated into automated flight safety monitoring systems in the future.

Keywords: spatial disorientation; spatial-disorientation detection; rotary-wing aircraft; helicopter; machine learning; flight simulation; accidents

1. Introduction

Recent advancements in aviation technology have gradually decreased aircraft-accident rates. However, accidents caused due to spatial disorientation (SD) have increased in recent years

(Newman and Rupert, 2020). SD is defined as "the failure of a pilot to accurately perceive their own or the aircraft's precise position, motion, and attitude relative to the ground" (Navathe and Singh, 1994). The Federal Aviation Administration (FAA) defines SD as "a state in which a pilot perceives the flight situation differently from reality due to errors arising from physical sensations in the visual, vestibular, and proprioceptive systems" (Federal Aviation Administration, 2009). The International Civil Aviation Organization (ICAO) defines the term as "Illusion in Flight" and describes it as "a situation where a pilot misinterprets information received from the visual, vestibular, and proprioceptive senses as being different from the actual situation due to environmental conditions or circumstances" (ICAO, 2009). Although the definitions of SD vary, they generally refer to the phenomenon in which a pilot incorrectly perceives the aircraft's attitude, direction, and speed. This concept has evolved from the past notion of simply making incorrect judgments regarding stimuli during flight to encompassing impairments in judging one's own or the aircraft's relative position, movement, and attitude with respect to the ground.

Accidents caused by SD in rotary-wing aircraft continue to occur, primarily at night and under low-visibility conditions, with 90% of them resulting in fatalities (Vreeken, 2013). The high SD-related accident rate in rotary-wing aircraft can be attributed to the unique characteristics of their missions. Rotary-wing aircrafts are often operated at low altitudes and maneuvered manually. Therefore, they require a high level of pilot attention, which can increase pilot fatigue during flights of the same duration. Consequently, higher fatigue can affect situational awareness and decision making capabilities of pilots, potentially increasing the likelihood of human error. Over 80% of pilots have reported that they have experienced SD at least once (Tu et al., 2021). Additionally, it is most likely to occur during night flights or under instrument meteorological conditions (IMCs), which involve significant visual constraints (Gu, 1994).

Research in this field has focused on comprehending the causes, mechanisms, and measurement methods, and developing countermeasures for SD. In particular, the use of flight simulators to expose pilots to SD situations and train them to respond to abnormal aircraft attitudes has been effective (Gibb, Ercoline and Scharff, 2011). However, despite such extensive research, SD-related accidents continue to occur. The fundamental reason for this is the lack of an objective SD-detection method. Most studies have focused on Type-II SD (perceived SD), whereas research on Type-I SD (unrecognized SD), which is the primary cause of fatal accidents, is relatively rare. If pilots are made aware of the SD situation, the problem-solving process becomes relatively straightforward. They can remove their hands from aircraft controls, switch to automatic flight mode, or hand control over to a fellow pilot (US Helicopter Safety Team, 2020). Although these response-training methods are emphasized in SD-recovery education, their limitations become apparent during solo flights or situations in which both pilots experience SD or fail to monitor each other.

Therefore, this study aimed to develop a tool capable of detecting SD occurrences in rotary-wing aircraft. The integration of the SD detection and response methods developed in this study into actual flight environments is expected to considerably improve rotary-wing aircraft safety. Furthermore, the risk of fatal aviation accidents can be reduced by effectively detecting and responding to SD situations that pilots are unaware of.

2. Related Work

SD in pilots can lead to instability in the attitude of the aircraft and potentially exceeding its performance limits, which can result in a total loss of control that often leads to crashes and fatal accidents. Owing to the unique characteristics of their missions, rotary-wing aircraft pilots frequently encounter IMCs. Therefore, the SD problem is more prevalent in rotary-wing aircraft and poses a serious threat to their flight safety. Various factors that affect SD have been identified, including human sensory (vision, vestibular system, etc.), environmental (clouds, fog, etc.), and cognitive (fatigue, experience, stress, etc.) factors. However, well-established methods or solutions that can effectively prevent SD are lacking (US Helicopter Safety Team, 2020).

Previous research has primarily focused on human factors related to SD such as fatigue, stress, and situational awareness. Despite gaining a deeper understanding of the various causes of SD and the development of better response methods, accidents caused by SD continue to increase, which may also be related to the introduction of advanced technologies, such as night vision goggles. While advanced technologies, such as autopilot systems, are helpful in low-visibility and SD situations, studies have shown that the inappropriate use of autopilot can increase the risk of SD in IMC situations as it may overwhelm the pilots (Johnson and Wiegmann, 2011). These technologies allow pilots to fly under challenging or previously impossible environmental conditions; however, they have also introduced new types of risks. Over the past 20 years, researchers have focused on studying human cognitive processes and behavioral responses to identify the causes of SD and found that SD involves the interpretation and response to information through the human visual, vestibular, and proprioceptive systems (Newman et al., 2014). Additionally, extensive research has been conducted on the interplay between human factors, such as fatigue, stress, sleep, and various environmental factors, which are employed in aeromedical training of pilots (Headquarters Department of the Army, 2000; Ayiei, Murray and Wild, 2020).

Most SD-related research has focused on reviewing various databases using qualitative or mixed methodologies to identify causes, such as pilot cognitive characteristics, weather conditions, aircraft type, and accident time. However, qualitative studies that incorporate pilot experiences are scarce. A phenomenological approach was used to explore the SD experiences of general aviation pilots and yielded themes of "weather and expectations," "thoughts and actions," and "post-flight experiences" (Gallo et al., 2015). Although this study made important contributions to exploring the root causes of SD using qualitative methods, its conclusion focused solely on the widely known importance of training, which limited its ability to provide more innovative or specific solutions. However, given that SD reflects the complex interaction between human and technical factors, conducting in-depth research on specific cases or experiences through qualitative studies is a valid approach. Moreover, as generalizing SD is difficult, it is necessary to focus on specific cases or aircraft types.

Any pilot can experience SD with varying degrees of individual susceptibility. Sleep deprivation has been found to significantly affect the occurrence of SD (Previc et al., 2007). Additionally, SD can negatively impact pilots' cognitive processing (Gresty and Golding, 2009) and increase their workload during flights (Webb et al., 2010). Although these research findings can be considered important for SD detection, no specific measures exist to effectively detect SD. A more concrete approach for detecting SD has been studied using human physiological responses. Li et al. (2015) demonstrated that electroencephalogram (EEG) signals of pilots are related to SD, and subsequently, Williams et al. (2018) proposed a noninvasive EEG monitoring helmet to detect SD. However, these studies focused on the technical complexities of EEG monitoring and prediction algorithms, as well as their limitations in fully understanding the complex cognitive functions of the human brain. Additionally, test methods have been developed to assess pilots' physiological data, such as electrocardiogram (ECG), heart rate (HR), blood pressure (BP), and eye movement activity, revealing that their gaze processing can offer an important clue for detecting SD (Lewkowicz et al., 2015). However, if the pilots are unaware that they are experiencing SD, they may not feel any discomfort or identify the abnormal state of the aircraft. Therefore, behavioral responses and physiological data may have limitations in clearly detecting SD situations below a certain threshold.

A crucial aspect of SD is whether pilots can identify that they are experiencing it. SD can be classified into three types based on pilot awareness (Heinle and Ercoline, 2003; Hao et al., 2020). In Type I, the pilots do not identify that they are experiencing SD and continue flying without realizing that they are unaware of the exact position or state of the aircraft. This is the most dangerous type of SD and can easily result in accidents. Type II is a state in which pilots recognize that they are experiencing SD and have difficulty maintaining stable aircraft control. Although pilots realize that their perception is incorrect and they attempt corrective actions, they are not performed as intended. Type III is a state in which pilots are overwhelmed by SD, experiencing extreme confusion regarding their position or orientation, which can lead to severe confusion, stress, performance degradation, and even incapacitation, similar to the "Giant Hand" phenomenon. In such cases, pilots generally identify the difficulties they are facing but find it challenging to take appropriate actions due to extreme stress.

Currently, the use of simulators to train pilots to recognize and recover from SD is considered the most definitive method for preventing SD-related accidents (Gibb, Ercoline and Scharff, 2011). However, most SD accidents occur when pilots are unaware of their SD state. In fact, approximately 80–85% of SD-related accidents occur in situations wherein pilots do not identify their SD state (Holmes et al., 2003). Therefore, the ability to recognize SD or an effective method for identifying them is critical.

As humans interact with an aircraft, their responses are transferred as inputs to control devices, and the changes are reflected in the attitude, performance, or specifications of the aircraft. It is already known that pilot performance can be evaluated through flight-data analysis (Tu et al., 2021). In an SD state, pilot performance can drastically deteriorate, whether intentionally or unintentionally, which can be identified through flight data. Crognale and Krebs (2011) used pilot-error occurrence indicators to confirm that pilot performance deteriorates under low-visibility conditions. Although error-occurrence rate is the most accurate indicator for measuring pilot performance, the reference values differ for each aircraft type, necessitating clear definitions of error criteria. Another study used quick access recorder flight data and analysis of variance and discovered that the flare operation of fixed-wing aircraft affects their landing distance and vertical acceleration. During SD, specifically in the Black Hole Illusion condition, an analysis of 14 flight parameters in a simulator revealed a tendency for higher descent rates during the initial approach phase (Huang et al., 2023).

Additionally, research has been conducted to develop SD-detection algorithms using flight data. A study proposed an algorithm that employs parameters such as aircraft heading, glide-path deviation, and aircraft tilt by analyzing flight data recorded in flight experiments (Frantis and Petru, 2018). However, its application is limited because the algorithm was developed considering deviations in the roll and course deviation indicator without including pilot response data (control input values). Moreover, such analysis methods rely on strict assumptions and may not fully analyze the complex relationships between each parameter, potentially reducing the accuracy of their results. In the past, the complexity and multidimensionality of data were not sufficiently considered and the detection of hidden patterns and correlations was limited. However, recent advancements in artificial intelligence technology and computing power have enabled the learning of complex nonlinear patterns from large-scale flight datasets, thereby enabling precise anomaly detection and prediction. In aviation, research using machine learning (ML) or deep-learning methods has been conducted in various areas such as aircraft delay rates and landing anomaly detection (Fernández et al., 2019; Timothy, Peng and Jung, 2019; Chahine, Hasan and Iddin, 2023; e Silva and Murça, 2023). These techniques are useful for accurately analyzing correlations in complex and vast flight datasets. Foucher et al. (2022) developed ML and deep-learning models that can detect SD caused by vestibular stimulation, suggesting the potential utility of these techniques for SD detection. However, they acquired data using a virtual device rather than an actual flight simulator, thereby lacking validation in a real flight environment. Additionally, they labeled SD based on the pilots' response speed (changes in control displacement values) for feature classification, which may have reduced their detection accuracy for Type-I SD.

3. Proposed Method

Acquiring high-quality flight data is the most critical challenge for developing SD-detection tools based on flight-data analysis. The following characteristics were considered to obtain high-quality flight data for this study:

First, to ensure the accuracy and reliability of the data, scenarios for the SD simulator experiments were designed using phenomenological analysis. Phenomenological research focuses on understanding the essence of human experiences and allows for their in-depth exploration from the participants' perspectives (Patton, 2002; Johnson and Christensen, 2012). Because vulnerability to SD varies depending on the aircraft type and individual pilot conditions, one-on-one in-depth interviews were conducted with 26 experienced pilots, and scenarios were designed considering the characteristics of rotary-wing aircraft.

Second, we employed a 6-axis motion flight training simulator located in a university certified as a professional training institution. It employed Prepar3D, a simulation program developed by Lockheed Martin for educational purposes, and the data were saved in real time in a CSV file on the instructor's computer. As shown in Fig. 1, severe turbulence conditions were applied when clouds were encountered during flight to induce anxiety, pressure, and stress in the pilots and increase their workload. Turbulence makes it difficult for pilots to control the aircraft, and severe shaking increases their stress and anxiety levels. Additionally, the instructor assigned them with a mission that involved finding a target object during a low-altitude flight. While the pilot focused on this task, low-visibility conditions were suddenly introduces to observe how they would overcome the situation, and the flight data were recorded.

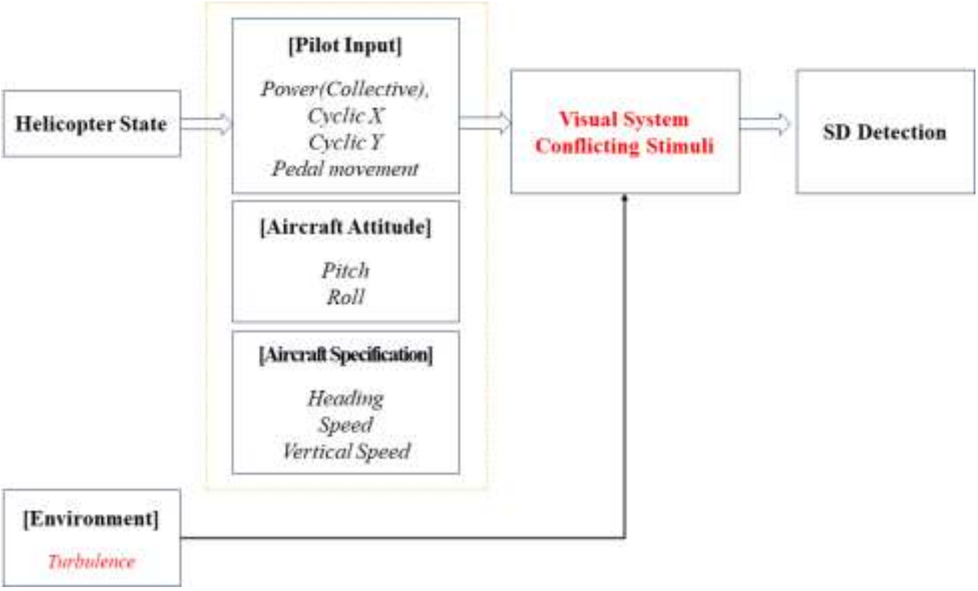

Figure 1. Simulator design diagram

3.1. Simulator scenario

The experiment included 65 pilots: 22 student, 13 private, and 28 commercial. The pilot qualifications, flight phases, and visibility conditions, which are the important factors that influence the occurrence of SD, are listed in Table 1.

Category	Flight Conditions
Pilot Qualification	Student Pilot, Private Pilot, Commercial Pilot
Flight Phase	Climbing Flight, Level Flight, Descending Flight
Visibility Conditions	10km or more (Good), 3sm (Normal), Obscured (Bad)

Table 1. Flight Conditions

Additionally, we employed Bell-206 as the rotary-wing aircraft, and before the actual flight simulation, each pilot underwent familiarization training for at least 9 min to adapt to the simulator. The instructor observed the participants' reactions to determine whether they were fully familiarized with the controls. The flight scenario began at Incheon Airport, which is adjacent to the coast. This airport was selected because most environments that induce SD have been identified as those involving flights from land to sea. While the pilot maintained a normal flight, the instructor introduced deteriorating weather conditions that were not anticipated by the pilot. Rotary-wing aircraft have high accident rates while transitioning from visual flight rules (VFR) to instrument flight rules (IFR) during visual flights (Ayiei, Murray and Wild, 2020). Rotary-wing aircraft pilots often enter IMCs unintentionally or unexpectedly during visual flights (Gallo et al., 2015), a situation defined as inadvertent entry into IMCs. Initially, favorable visibility of over 10 km was maintained, which was gradually deteriorated to simulate worsening weather conditions. Finally, the flight characteristics and tendencies of the pilot were observed under obscured visibility conditions. To extract standardized data, the accurate times of flight phase and visibility changes were recorded and stored. After the flight, interviews were conducted to distinguish pilots who experienced SD from those who did not.

3.2. Data acquisition and preprocessing

The pilot boarded the flight simulator and conducted a flight in a rotary-wing aircraft-based SD scenario. The data were collected every 0.25 s, resulting in 175,500 data points for the 65 participants. The flight parameters; pilot input values, which refer to the displacement values of the aircraft-control devices, such as the cyclic, collective, and pedal devices; aircraft attitude values, indicating aircraft movements; and aircraft performance values are listed in Table 2. Aircraft control began with the pilot's control input, which changed the aircraft's attitude values and ultimately its performance values.

The flight data were preprocessed as follows: First, to reflect the characteristics of each flight phase accurately, the flight phases were distinguished by coding them into three stages: climb (0), level (1), and descent (2). Next, to reflect the influence of environmental conditions on the model, visibility was categorized into "very good (0)," "average (1)," and "very poor (2)," and corresponding coding was applied. Finally, pilot qualifications were classified as "student pilot (0),” private pilot (1),” and “commercial pilot (2)."

After the flight experiment, interviews were conducted with the participants and the instructor pilot to determine whether they had experienced SD, and the data of the pilots who experienced SD were separated from those of who did not and labeled. The SD labeling method is discussed in detail in Section 3.3.

Finally, each data point was assigned labels for the aircraft attitude (pitch, bank), aircraft specifications (heading, speed, vertical speed, altitude), and pilot input (slider, x, y, rx). All data were stored in CSV files, and rows and columns with missing values were removed.

The dataset comprised 175,500 data points, of which 91,800 were from pilots who experienced SD, whereas the remaining 83,700 were from those who did not. To address the class imbalance, an oversampling technique was applied to increase the number of data points in the minority class. This prevented performance degradation owing to class imbalance during model learning and validation and balanced the two classes.

Category	Parameters
Time	· Time(T) : Hour, Minute, Second
Pilot Input Values	· Power(Slider) : Collective Displacement · Lateral cyclic movement(x) : Cyclic Left-Right Displacement · Fore/aft cyclic movement(y) : Cyclic Up-Down Displacement · Pedal movement(rx) : Pedal Displacement
Aircraft Attitude Values	· Pitch : Aircraft Pitch Attitude

	· Bank : Aircraft Roll Attitude
Aircraft Specification Values	· Heading : Aircraft Heading · Speed : Aircraft Indicated Airspeed · VS(Vertical speed) : Rate of Climb, Rate of Descent

Table 2. Flight Parameters

3.3. ML analysis

The SD dataset was analyzed using an ML-based supervised learning classification methodology. Python was used as the analysis tool and feature selection was conducted by setting an error threshold to maximize the efficiency and accuracy of the model. This process focused on determining the optimal value that minimized the error rate of the model while maintaining important information by setting an error threshold. The error criteria varied based on the aircraft category (airplanes, helicopters, etc.) and operating conditions (normal operations, qualification training, and abnormal attitudes) (ICAO, 2014). Abnormal attitudes include turbulence, instrument failure, and poor piloting skills, encompassing all unintentionally induced attitudes. Therefore, it is not desirable to apply these to rotary-wing aircraft. Consequently, as listed in Table 3, the operationally defined characteristic values of rotary-wing aircraft based on consultations with helicopter safety experts from the FAA were used as the criteria (Crognale and Krebs, 2011).

Parameters	**Error Threshold**
Pitch	0 ± 5°
Bank	0 ± 18°
Heading	340° ± 10°
Speed	80knots ± 10knots
VS(vertical speed)	Climb : +500FPM ± 1,000FPM
	Level : 0FPM ± 1,000FPM
	Descent : -500FPM ± 1,000FPM

Table 3. Error Threshold

Data labeling involves assigning labels that identify the category to which a particular data point belongs during the training of an ML model. The aim was to accurately classify the presence or absence of SD and generate a training dataset that detects and predicts the corresponding state. To identify SD situations that may occur during flight, errors related to the pilot's spatial perception and recognition or judgment of aircraft controls were included. The classification criteria were established based on flight data, pilot behavioral responses, and aircraft performance indicators. Based on these, an SD detection and classification algorithm was designed, as shown in Fig. 2. Additionally, the professional judgment of the instructor pilot was essential to distinguish between SD-induced errors and those caused by the poor manipulation skills of the pilot. The instructor observed the pilots' responses, aircraft specifications, and consistencies of the three control inputs (pedal, cyclic, and collective) during the flight experiment. The coordination of the three control inputs refers to the synchronization of the operation of the three control sticks (pedal, cyclic, and collective) used to control the aircraft attitude. As stated previously, SD can be broadly classified into Type I, wherein the pilot is unaware that they are experiencing SD, and Types II and II, wherein they are aware but unable to take appropriate actions owing to a conflict with the illusion. To accurately classify these SD states, additional interviews were conducted with the participants and instructor pilot after the experiment. They included in-depth discussions regarding the abnormal flight situation at a particular time, and the presence or absence of SD was determined based on the results. The collected data points were manually labeled by classifying them into SD and non-SD situations. The labeled data were then separated into training and validation datasets for use in subsequent analyses and modeling processes.

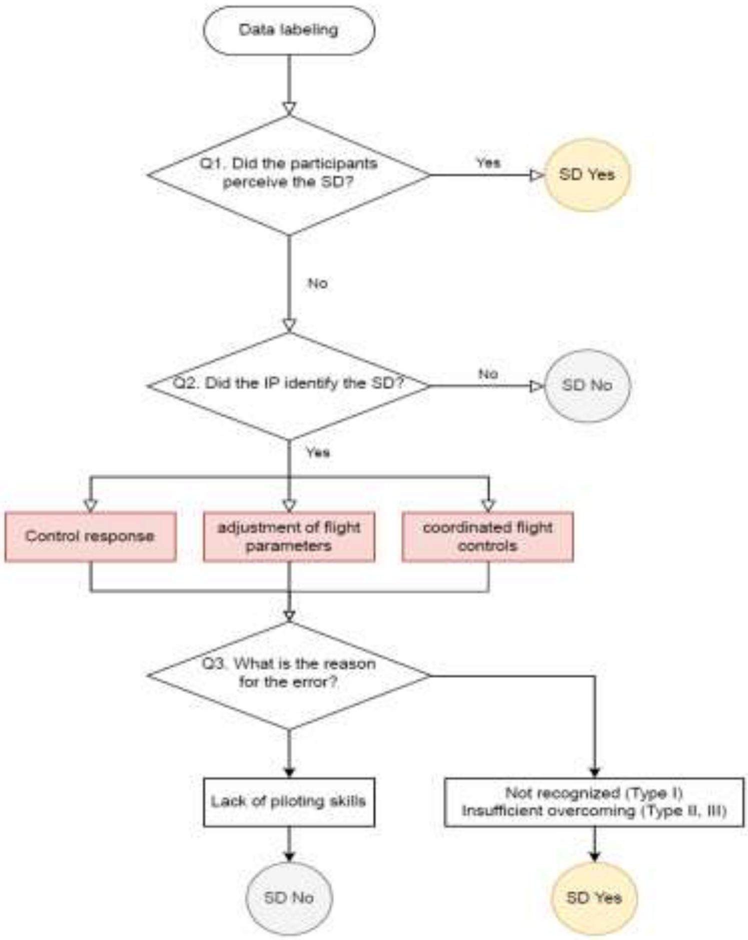

Figure 2. SD classification algorithm

4. Results

4.1 Analysis of pilot errors

The results of analyzing the differences in pilot-error rates based on the presence or absence of SD are presented in Table 4. Heading and Vs were not statistically significant ($p > 0.05$), indicating that there was no significant difference in the median between the two conditions. However, significant differences were observed in the p-values for the pitch, bank, and speed data. These results imply that when visibility decreased, the pilot switched to IFR and focused primarily on the aircraft pitch and bank to maintain a stable attitude. This is because all pilots are trained to prioritize the stabilization of pitch and bank as controlling the basic attitude of an aircraft is the most crucial factor for ensuring safety. Therefore, only after securing basic stability can the pilot control the aircraft to maintain its path or altitude, enabling them to respond effectively to emergencies. However, the situation is different in SD cases, wherein even if the pilot wants to control the basic flight attitude, they are unable to take appropriate actions owing to a conflict with their sensory organs. Therefore, the possibility of pitch, bank, and speed errors increases, and their inability to secure basic stability of the aircraft can lead to a crash.

Category	n	Mann-Whitney U	p-value
Pitch	65	1542.0	0.003
Bank	65	1982.5	0.043
Heading	65	1768.0	0.080
Speed	65	1666.5	0.010
Vs	65	1830.5	0.075

Table 4. Mann-Whitney U Test Results by SD

The results of the error-rate analysis based on the pilots' qualification level (student (Certi = 1), private (Certi = 2), and commercial (Certi = 3)) are presented in Table 5. Under the SD condition, the error-occurrence rate differed significantly based on pilots' qualifications (p-value <.001). The regression coefficient of student pilots (Certi = 1) was 0.426 and their Exp(B) value was estimated to be 1.530, which indicates an increase of approximately 53% in the error-occurrence rate compared with commercial pilots. The regression coefficient and Exp(B) value of private pilots (Certi = 2) were 0.382 and 1.465, respectively, indicating and approximately 46.5% increase in the error-occurrence rate compared with commercial pilots. Therefore, a pilot's qualification level is related to the occurrence rate of errors, and student and private pilots tend to exhibit higher error-occurrence rates than commercial pilots. Thus, as expected, the more experienced the pilot, the fewer the errors, verifying the reliability and validity of the simulation flight data employed in this study.

Parameter Estimates

Parameter	B	Std. Error	95% Wald CI		Hypothesis Test			Exp(B)	95% CI for Exp(B)	
			Lower	Upper	Wald Chi-Square	Degrees of Freedom	Probability		Lower	Upper
(Intercept)	6.055	.0092	6.037	6.073	437430.651	1	<.001	426.143	418.565	433.858
[Certi=1]	.426	.0122	.402	.449	1225.962	1	<.001	1.530	1.494	1.567
[Certi=2]	.382	.0144	.353	.410	703.377	1	<.001	1.465	1.424	1.507
[Certi=3]	0a	.	.	.	.	.	.	1	.	.
(Scale)	1b									
Dependent Variable : total error Model : (Intercept), Certi										

Table 5. Poisson regression analysis by Certification

4.2. Performance evaluation of the SD-detection tool

After conducting an initial performance evaluation through five-fold cross-validation, the performances of the ML models were further evaluated using the area under the receiver operating characteristic curve (AUROC) metric and compared with the results obtained in previous studies. The AUROC is an important metric for evaluating the performance of binary classification models. Additionally, true positive rate (TPR) indicates the ratio of actual positive samples predicted as positive, whereas the false positive rate (FPR) indicates the ratio of actual negative samples incorrectly predicted as positive. An AUROC value close to 1 indicates excellent performance and good class discrimination, whereas that close to 0.5 indicates a performance akin to random guessing.

The tool developed in this study exhibited AUROC scores of 0.944 for decision tree, 0.953 for random forest, 0.949 for extra trees, and 0.951 for gradient boosting models, as shown in Table 6 and Fig. 3. Thus, random forest exhibited the best performance, indicating that it can effectively handle complex SD situations. Additionally, these results demonstrated a performance improvement of approximately 13.3% compared to recently employed random forest (AUROC = 0.82) and long short-term memory (AUROC = 0.84) models [45]. Furthermore, the flight data in this study were generated using a flight simulator featuring an environment identical to that of an actual flight, and by employing a labeling algorithm that additionally considered the error threshold and three-way consistency, a tool that can effectively detect SD was developed.

Decision Tree (DT)	Random Forest (RF)	Extra Trees (ET)	Gradient Boosting (GBC)
0.944	0.953	0.949	0.951

Table 6. ROC-AUC comparison

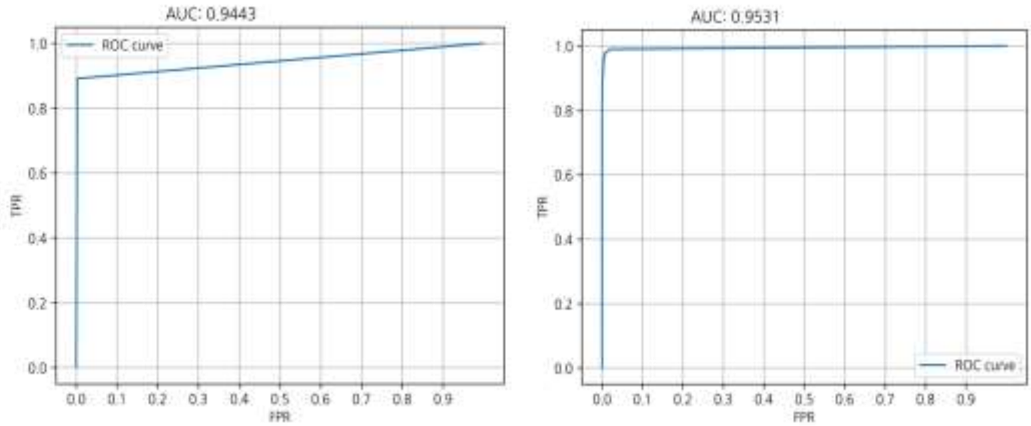

Figure 3. ROC-AUC Curve(Left DT, Right RF)

4.3. Importance of flight characteristics

The influence of each flight parameter on inducing SD was analyzed by extracting the importance of flight characteristics using the trained SD-detection tool. The results indicated that the aircraft heading was the most important parameter across all models. Additionally, although the order of the top four important characteristics differed among the models, aircraft heading, pilot pedal input (rx), power (slider), and aircraft speed were consistently selected as important parameters.

The results of this study indicate that pitch and bank, which have been considered crucial for detecting SD occurrences, may not be reliable indicators of SD because pilots make every effort to maintain the aircraft's attitude, even in SD situations.

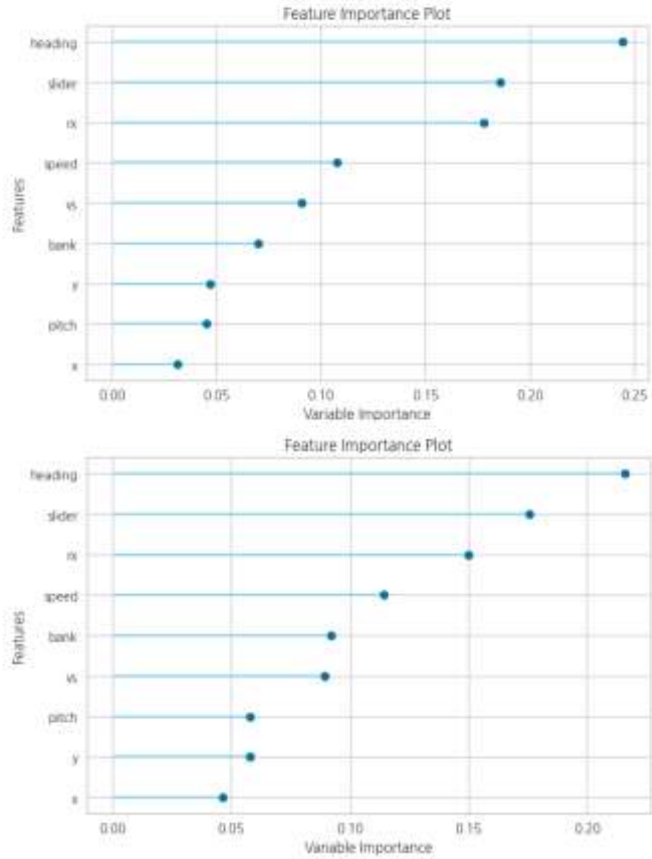

Fig. 4. Random Forest(left) & Extra Trees(right) feature importance

5. Discussion

This study did not to develop a generalized SD-detection tool but one based on optimized scenarios that are compatible with the mission environments and flight characteristics of rotary-wing aircraft. Unlike previous studies that mainly focused on identifying the causes of SD from a human behavioral or cognitive perspective, this study developed an SD-detection tool using various ML algorithms. However, the accuracy and reliability of the model may be affected if high-quality data are not guaranteed. Therefore, high-quality data collected from various mission environments and conditions are required to ensure its scalability. Moreover, ML models excel at recognizing patterns in complex flight data and can reveal relationships that traditional statistical methods may not. In this study, the prediction of SD occurrence and its characteristic importance were analyzed using a simulator. However, broader insights can be obtained by employing actual flight data in the future. However, realistically reproducing SD situations during an actual flight is a dangerous and challenging endeavor. Therefore, a certified simulator of the same type can be useful for reproducing SD situations without the additional complexities and risks involved in actual flight situations. Thus, the proposed simulator-based approach can increase the practical applicability of ML-based SD detection tools and contribute significantly to aviation safety research and development.

6. Conclusion

This study collected high-quality data from flight scenarios specifically designed to induce and detect SD occurrences in rotary-wing aircraft and developed a new ML-based SD detection tool using these data. Considering that the performance of ML models depends heavily on the data quality, this study succeeded in improving the detection accuracy in three key aspects.

First, unlike previous studies that did not consider various flight conditions to detect general SD occurrences, this study generated a more realistic flight dataset using scenarios designed to reflect actual flight situations. This approach allowed the ML model to detect and predict SD occurrences in real flight situations more accurately by precisely capturing the flight data and situational responses of a rotary-wing aircraft.

Second, in the flight-data selection process, the control input values were directly related to the pilot's response and the aircraft's attitude values, and the specification values provided more accurate criteria reflecting the pilot's control state and SD situation. In particular, the aircraft heading was found to be crucial for identifying the pilot's intended control state and SD situation, and the feature selection process through the error-rate threshold setting contributed to improving the accuracy of the data for SD detection.

Third, data labeling was performed using an SD detection and classification algorithm, enabling the distinction between SD and poor manipulation by pilots. The classification criteria for this were set as the control response, specification correction, and three-way consistency, which allowed identifying and classifying the pilot’s SD state more accurately.

Thus, we developed a unique ML-based method for detecting SD occurrences. The random forest algorithm exhibited the best performance, demonstrating a performance improvement of approximately 13.3% compared to other recently proposed models. In addition, important new parameters were discovered to detect SD. Pitch and bank, which are traditionally considered to be closely related to SD, were identified as major indicators of loss of control, and it was confirmed that heading, pedal input (rx), power (slider), and speed were important factors for SD detection.

Acknowledgments This study was supported by a Hanseo University Research Grant in 2023.

References

[1] Administration, F.A. (2009) Pilot’s handbook of aeronautical knowledge. Skyhorse Publishing Inc.

[2] Ayiei, A., Murray, J. and Wild, G. (2020) ‘Visual flight into instrument meteorological condition: A post accident analysis’, Safety, 6(2), p. 19.

[3] Chahine, M.H.A.A., Hasan, R. and Iddin, W.A. (2023) ‘Flight Delays Prediction by using Machine Learning’, GSJ, 11(6).

[4] Crognale, M.A. and Krebs, W.K. (2011) ‘Performance of Helicopter Pilots During Inadvertent Flight Into Instrument Meteorological Conditions’, The International Journal of Aviation Psychology, 21(3), pp. 235–253. Available at: https://doi.org/10.1080/10508414.2011.582443.

[5] Fernández, A. et al. (2019) ‘Flight data monitoring (FDM) unknown hazards detection during approach phase using clustering techniques and AutoEncoders’, Proceedings of the Ninth SESAR Innovation Days, Athens, Greece, pp. 2–5.

[6] Foucher, J. et al. (2022) ‘Simulation and Classification of Spatial Disorientation in a Flight Use-Case Using Vestibular Stimulation’, IEEE Access, 10, pp. 104242–104269. Available at: https://doi.org/10.1109/ACCESS.2022.3210526.
[7] Frantis, P. and Petru, A. (2018) ‘Automatic Evaluation of Spatial Disorientation’, Advances in Military Technology, 13(2), pp. 185–192.
[8] Gallo, M.A. et al. (2015) ‘Inadvertent VFR-into-IMC flights: A qualitative approach to describing GA pilots’ first-hand experiences’, The Collegiate Aviation Review International, 33(2).
[9] Gibb, R., Ercoline, B. and Scharff, L. (2011) ‘Spatial disorientation: decades of pilot fatalities’, Aviation, space, and environmental medicine, 82(7), pp. 717–724.
[10] Gresty, M.A. and Golding, J.F. (2009) ‘Impact of Vertigo and Spatial Disorientation on Concurrent Cognitive Tasks’, Annals of the New York Academy of Sciences, 1164(1), pp. 263–267. Available at: https://doi.org/10.1111/j.1749-6632.2008.03744.x.
[11] Gu, B.-S. (1994) ‘A Search on the Spatial Orientation in Flight’, Journal of the Korean Society for Aviation and Aeronautics, 2, pp. 139–182.
[12] Hao, C. et al. (2020) ‘A classification method for unrecognized spatial disorientation based on perceptual process’, Ieee Access, 8, pp. 140654–140660.
[13] Headquarters Department of the Army (2000) ‘Aeromedical Training for Flight Personnel’. Washington DC.
[14] Heinle, T.E. And Ercoline, W.R. (2003) ‘Spatial Disorientation: Causes, Consequences And Countermeasures For The Usaf’, In The Rto Hfm Symposium, Air Force Research Lab Wright-Patterson Afb Oh Human Effectiveness Directorate, La Coruña, Spain.
[15] Holmes, S.R. et al. (2003) ‘Survey of spatial disorientation in military pilots and navigators’, Aviation, space, and environmental medicine, 74(9), pp. 957–965.
[16] Huang, L. et al. (2023) ‘How the black hole illusion environment affects operational performance at different flight phases in aviation’, Applied Ergonomics, 113, p. 104048.
[17] ICAO (2009) ‘Manual of criteria for the qualification of flight simulation training devices’, International Civil Aviation Organization, Montreal, volume I—aeroplanes.
[18] ICAO (2014) ‘Manual on Aeroplane Upset Prevention and Recovery Training’. INTERNATIONAL CIVIL AVIATION ORGANIZATION.
[19] Johnson, B. and Christensen, L. (2012) ‘Quantitative, qualitative, and mixed approaches’, Educational Research, University of South Alabama: SAGE [Preprint].
[20] Johnson, C.M. and Wiegmann, D.A. (2011) ‘Pilot Error During Visual Flight Into Instrument Weather: An Experiment Using Advanced Simulation and Analysis Methods’, Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 55(1), pp. 138–142. Available at: https://doi.org/10.1177/1071181311551029.
[21] Lewkowicz, R. et al. (2015) ‘Flights with the risk of spatial disorientation in the measurements of oculomotor activity of pilots’, The Polish Journal of Aviation Medicine and Psychology, 21(3), p. 23.
[22] Li, Y. et al. (2015) ‘EEG functional network properties related to visually induced unrecognized spatial disorientation’, Bio-medical materials and engineering, 26(s1), pp. S1115–S1124.
[23] Navathe, P.D. and Singh, B. (1994) ‘An operational definition for spatial disorientation’, Aviation Space and Environmental Medicine, 65, pp. 1153–1153.
[24] Newman, M. et al. (2014) ‘Perceptual Modeling as a Tool to Prevent Aircraft Upset Associated with Spatial Disorientation’, in AIAA Guidance, Navigation, and Control Conference. AIAA Guidance, Navigation, and Control Conference, National Harbor, Maryland: American Institute of Aeronautics and Astronautics. Available at: https://doi.org/10.2514/6.2014-0443.
[25] Newman, R.L. and Rupert, A.H. (2020) ‘The Magnitude of the Spatial Disorientation Problem in Transport Airplanes’, Aerospace Medicine and Human Performance, 91(2), pp. 65–70. Available at: https://doi.org/10.3357/AMHP.5442.2020.

[26] Patton, M. (2002) ‘Qualitative Research and Evaluation Methods. London: Sage Publications’. Inc.
[27] Previc, F.H. et al. (2007) ‘Simulator-induced spatial disorientation: effects of age, sleep deprivation, and type of conflict’, Aviation, space, and environmental medicine, 78(5), pp. 470–477.
[28] e Silva, L.C. and Murça, M.C.R. (2023) ‘A data analytics framework for anomaly detection in flight operations’, Journal of Air Transport Management, 110, p. 102409.
[29] Timothy, K.K., Peng, C. and Jung, J.J. (2019) ‘Spatial-Temporal Long Short-Term Memory (ST-LSTM) Modelling for Flight Departure Delay Prediction: Preliminary study’, Proceedings of the Korea Information and Communications Society Conference, 23(2), pp. 265–268.
[30] Tu, M.-Y. et al. (2021) ‘Analysis of in-flight spatial disorientation among military pilots in Taiwan’, Journal of Medical Sciences, 41(1), pp. 22–28.
[31] US Helicopter Safety Team (2020) ‘Spatial Disorientation Induced by a Degraded Visual Environment’. Available at: https://ushst.org/wp-content/uploads/2021/02/Rec-Prac-HSE-127-Spatial-Disorientation.pdf (Accessed: 28 April 2024).
[32] Vreeken, J. (2013) ‘Helicopter flight in a degraded visual environment’. National Aerospace Laboratory NLR.
[33] Webb, C.M. et al. (2010) ‘The effect of spatial disorientation on working memory and mathematical processing’, Aviation Space and Environmental Medicine, 81, p. 313.
[34] Williams, Q.L. et al. (2018) ‘PW 2658 Human machine interaction system for providing neural interface based model for alerting and mitigation of spatial disorientation for pilots’, Injury Prevention, 24(Suppl 2), pp. A218–A218. Available at: https://doi.org/10.1136/injuryprevention-2018-safety.603.