

비행데이터를 활용한 머신러닝 기반 비행착각 탐지 알고리즘 성능 분석

Performance Analysis of Machine Learning Based Spatial Disorientation Detection Algorithm Using Flight Data

Yim Se-Hoon¹ · Park Chul^{2*} · Cho Young jin^{3*}

Department of Helicopter Operation, Hanseo University, Taeon 32158, Korea

[Abstract]

Helicopter accidents due to spatial disorientation in low visibility conditions continue to persist as a major issue. These incidents often stem from human error, typically induced by stress, and frequently result in fatal outcomes. This study employs machine learning to analyze flight data and evaluate the efficacy of a flight illusion detection algorithm, laying groundwork for further research.

This study collected flight data from approximately 20 pilots using a simulated flight training device to construct a range of flight scenarios. These scenarios included three stages of flight: ascending, level, and descent, and were further categorized into good visibility conditions and 0-mile visibility conditions. The aim was to investigate the occurrence of flight illusions under these conditions.

From the extracted data, we obtained a total of 54,000 time-series data points, sampled five times per second. These were then analyzed using a machine learning approach.

Key word : Accidents, Helicopter, Machine learning, Pilot performance, Spatial disorientation.

1. 서 론

1-1 연구 배경

항공산업 기술의 발달로 항공기 사고율은 점차 감소하는 추세를 보이고 있지만 사고원인 중 인적요인으로 인한 사고의 감소율은 점차 둔화되고 있다. 특히, 국내 회전익항공기의 사고는 지속적으로 발생하고 있으며 대부분 시계비행 기상조건에서 다양한 저고도 비행임무를 수행하기 때문에 인적요인의 중요성이 고정익항공기에 비해 매우 크다. 아울러, 회전익 항공기는 산불진화, 인명구조 등과 같은 임무 특성상 조종사들은 임무에 대한 책임감과 외부 환경적으로 압력을 받게 되므로 악기상에서 무리하게 비행하는 경우가 많다.

회전익항공기 사고 중 야간, 저시정 환경에서 비행착각으로 인한 사고는 지속적으로 발생하고 있고 비행착각으로 인한 사고의 90%는 대부분 치명적인 사망사고로 연결된다[1]. 조종사의 80%이상은 비행착각을 경험하였고[2], 비행착각이 가장 잘 발생하는 경우는 야간, 계기비행 기상조건과 같이 시야 제약이 많이 받을 때 발생한다[3].

시뮬레이터를 활용하여 비행착각을 경험하고 극복할 수 있는 비행기술은 대부분 일회성 교육으로 끝나고 법적으로 의무화 되어 있지 않으며, 시뮬레이터가 없는 기관이 많다. 이와 같이 회전익항공기 조종사 교육환경은 상대적으로 열악한 실정이다.

인간은 감각 매커니즘과 정보처리를 할 수 있는 양적 한계로 인하여 공간적 방향감각 상실이 발생할 수 밖에 없으며, 외부 시각적 참조물이 잘 보이지 않는 경우 자주 발생하게 된다. 비행 중 공간정위상실에 빠졌을 경우 항공기의 비정상적인 자세가 나타날 수 있으며 조종사의 성능은 급격히 떨어지게 된다. 본 연구에서는 머신러닝 기법을 적용하여 저시정 상태의 비행착각이 발생한 비행데이터를 분석하고 사전에 비행착각을 탐지할 수 있는 기초 연구를 수행하고자 한다.

<http://dx.doi.org/10.12673/jant.2023.27.4.391>



This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Received 4 August 2023; Revised 10 August 2023
Accepted (Publication) 29 August 2023 (30 August 2023)

*Corresponding Author ; Cho Young jin

Tel: +82-41-671-6271

E-mail: newjxfree@gmail.com

1-2 사고사례

최근 10년간 비행착각으로 발생한 국내 회전익항공기 사고는 총 네건이 발생하였다. 첫째는 2013년 11월에 LG 소속 헬리콥터는 짙은 안개 속에서 아파트와 충돌하여 조종사 2명이 사망하였다. 둘째 2014년 7월 소방청 소속 헬리콥터가 계기비행 방식(IFR; instrument flight rules)으로 이륙 중 구름 속에서 의도하지 않은 우선회 반응으로 인해 추락하여 5명이 사망하였다. 셋째는 2015년 3월 해양경찰 소속 헬리콥터가 가거도 헬리패드에 착륙 하던 중 기상악화로 인하여 추락한 사고이다. 넷째, 2019년 10월 중앙119 소속 헬리콥터가 무월광 야간비행 기상 상태에서 독도 남방 600 m 지점에서 추락하여 3명이 사망하고 4명이 실종되었다.

미국의 대표적인 회전익항공기 사고는 2015년 환자를 병원으로 이송하는 응급의료 비행 복귀 중 추락한 사례이다. 당시 야간 계기기상조건(IMC; instrument meteorological condition)에서 선회 후 공간정위상실로 인한 지형에 충돌한 사례이다. 둘째는 2015년 낮은 시정 조건에서 이륙한 헬리콥터는 강 수면 위로 낮게 비행 중에 고도를 유지하지 못하고 비행착각으로 인하여 추락하였다. 셋째는 계기 자격이 없는 조종사가 알래스카 지역에서 착륙 하는 동안 화이트아웃으로 인하여 시정이 감소되었고 공간정위상실로 인하여 추락하였다. 미연방교통안전위원회(NTSB; national transportation safety board)의 사고조사 보고서에 따르면, 그림 1과 같이 2016년을 제외하고 매년 지속적으로 비행착각으로 인한 회전익항공기 사고가 발생하고 있다.

이러한 사고의 원인의 공통점은 시정이 좋지 않았고, 특히 회전익항공기는 낮은 고도에서 임무를 수행하므로 조치할 시간이 부족하므로 대부분 사망사료로 연결된다는 점이다.

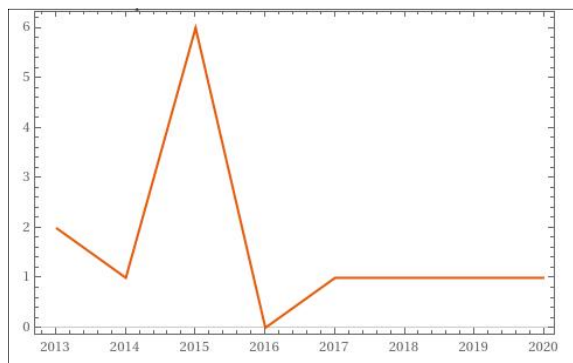


그림 1. 미국 비행착각 사고발생 건수

Fig. 1. Number of Spatial Disorientation Accidents in the U.S.

1-3 연구동향

비행 데이터는 시간의 흐름에 따라 구성되므로 시계열 분석이 필수적이다. 시계열 데이터 분석 및 예측은 모터 이상 감지, 전력 사용량 예측, 차량 경로 예측 등 많은 분야에서 중요하게 다루어지고 있다. 비행 데이터의 복잡하고 다양한 특성으로 인해 체계적인 분석 및 모델링의 적용이 요구된다. 머신러닝 또는 딥러닝을 활용한 항공분야 연구는 자율비행, 항공영상 분석 등 다양하게 이루어지고 있으나, 조종사의 비행 데이터를 분석하여 성능을 평가하거나 이상치를 감지하는 연구는 많지 않다.

2021년에는 한국전자통신연구원에서 딥러닝 기반 항공안전 이상치 탐지 기술 동향을 연구하여 이상치 탐지 분야에 대한 다양한 방법론이 제안되고 있다[4].

항공기 지연 여부의 정확도와 시간 예측에 가장 좋은 성능을 보인 것은 딥러닝 모델이었고, 랜덤 포레스트 모델이 F1-score를 고려했을 때, 가장 좋은 성능을 보였다[5].

비행 데이터 관련해서는 장단기기억신경망(LSTM; long short-term memory) Encoder-Decoder 모델을 사용하여 비행 데이터를 예측하였다. 해당 모델은 과거부터 현재까지의 연속적인 항공기 상태변수 값을 바탕으로 원하는 길이만큼 항공기의 상태변수를 예측하였다[6].

해외에서는 비행 데이터 모니터(FDM; flight data monitor)을 활용하여 분석하는 연구가 진행되고 있다. 과거의 전통적인 분석방법은 비행데이터를 전문가에 의존하여 사전에 정의된 임계값을 초과하는 특정 매개변수를 기반으로 하는 경우가 많다. 그러나 이러한 패턴의 복잡성으로 인해 회귀 이벤트나 이상 현상의 탐지는 까다로울 수 있다. 이러한 어려움은 안전 사고에 더 취약한 접근 및 이륙과 같은 중요한 비행 단계에서 특히 많이 발생한다. 클러스터링 기술은 이러한 데이터 내에서 패턴, 상관 관계 및 이벤트를 발견하는 데 도움이 된다. 그런 다음 예측 분석, 특히 인공신경망(ANN; artificial neural network) 및 오토인코더(AutoEncoders)와 같은 딥러닝 기술을 사용하여 이러한 비정상적인 이벤트를 감지할 수 있다[7].

비행 데이터 예측과 관련해서는 기존에 장단기기억신경망(LSTM)을 기반으로 항공기의 4D 경로를 예측하는 연구가 수행되었다[8].

II. 본 론

2-1 시계열 데이터(Time series)

시계열 데이터는 시간의 흐름에 따라 관측된 값들로 구성되

어 있으며, 경제나 자연과 같은 다양한 현상을 분석하기 위해 사용된다. 이러한 데이터는 그 이전의 자료들과 연관성을 가지고 있어 분석이 복잡하며 예측이 어렵다. 전통적인 시계열 데이터 분석 모델인 ARIMA는 모형의 제약적인 가정과 복잡한 모형 적합, 그리고 예측의 부정확성 등의 문제점을 가지고 있다.

그러나, 시계열 데이터 분석은 연구자들에게 지속적인 관심을 받아오며, 더 높은 정확도의 예측을 위한 연구가 계속되고 있다. 파라웨이(Faraway)의 신경망 모델 연구나, 박스 젠킨스(Box-Jenkins)의 ARIMA와의 비교 분석은 그 예시 중 하나이다. 인공 신경망과 딥러닝을 활용한 시계열 데이터 분석 연구가 확산되면서, 다차원 공간에서 복잡한 비선형 분류가 가능한 순환신경망(RNN; recurrent neural network)은 특히 주목을 받고 있다.

순환신경망(RNN)은 내부의 정보가 지속되도록 구조화되어 있으며, 시계열 자료에 적합하게 설계되었다. 하지만, 기술기 유실 문제와 장기 의존성 문제가 발생할 수 있는데, 이를 해결하기 위해 장단기기억신경망(LSTM)이 유르겐 슈미트후버(Jürgen Schmidhuber)에 의해 제안되었다. 본 연구에서는 우선 지도학습의 방식인 랜덤 포레스트와 엑스트라 트리 모델을 활용하여 분석하였다.

2-2 랜덤 포레스트(Random Forest)

랜덤 포레스트는 다수의 의사결정나무(Decision Trees)를 결합하여 작동하는 알고리즘으로, 일반화 성능을 향상시키기 위해 배깅(Bagging) 기법과 특징 임의성(Random Subspaces)을 활용한다. 해당 알고리즘의 특성은 다수의 결정 트리가 예측을 수행하고, 그 결과를 투표하여 최종 예측을 도출함으로써 오버피팅 문제를 완화하고, 높은 분류 정확도와 안정성을 보장하며, 변수 중요도를 측정하는 능력이 있다.

간편한 훈련, 누락된 데이터의 효율적 처리, 높은 예측 성능 등이 장점이지만 모델의 복잡성으로 인한 해석의 어려움, 무작위성으로 인한 재현성의 제한 그리고 높은 차원의 희소한 데이터 처리에 대한 제약이 존재하는 단점이 있다. 현재의 랜덤 포레스트 개념은 레오(Leo)의 논문에서 만들어 졌다[9].

2-3 엑스트라 트리(Extra Trees)

엑스트라 트리는 랜덤 포레스트의 변형 중 하나로 의사결정 나무의 분할에 더 많은 무작위성을 도입하는 방법이다. 엑스트림 랜덤 트리에서는 노드의 각 분할에서 변수를 완전히 무작위로 선택하고 이 중에서 최적의 분할을 선택한다. 이로 인해, 엑스트림 랜덤 트리는 계산 효율성이 더욱 높으며 빠른 훈련 시간을 가진다.

또한, 이 추가적인 무작위성은 일반적인 랜덤 포레스트보다 오버피팅을 더욱 줄일 수 있게 해주지만 이런 무작위성이 모델의 해석을 더 어렵게 만들 수도 있으며, 때로는 랜덤 포레스트보다 성능이 낮을 수도 있다[10].

III. 실험

3-1 실험환경

비행 데이터 추출을 위해 회전익항공기 시뮬레이션 환경을 구축하였다. 시뮬레이터는 조종사 비행작각 훈련을 적용하기 위해 6축 모션 제어시스템과 화면, 교관석으로 구성되고 상용 비행 시뮬레이터 프로그램인 Prepar 3D와 Bell-206 헬리콥터 기종으로 실험을 실시하였다. 항공기 움직임에 따른 계기 정보를 계기정보 패널로 내보내며 모션 시스템은 FlyPT Mover 프로그램을 활용하여 비행 데이터를 모션 데이터로 바꾸게 된다. 이러한 모션 기반 헬리콥터 시뮬레이터는 조종사에게 더욱 현실감 있는 조종환경을 만들어주고 모든 비행데이터는 CSV 파일 형태로 교관석 컴퓨터에 실시간으로 저장된다.

3-2 비행 시나리오

표준화된 비행시나리오를 구성하여 실험에 적합한 환경을 구축하였다. 실험에 참가한 조종사는 총 20명으로 자가용 조종 자격이 있는 조종사와 학생 조종사 그룹으로 나누었다.

기종은 Bell-206 회전익항공기로 실시하였고 그림 2와 같이 인천공항에서 340° 방향으로 이륙하여 시뮬레이터에 적응할 수 있도록 약 10분간의 장주비행을 실시하였다.

친숙화 훈련 후에는 개인별로 약 9분간 비행을 실시하면서 상승, 수평, 강하 비행단계를 지시하였다. 각각의 비행제원은 속도 80 kts, 상승·강하율 500 fpm, Heading 340°를 유지하도록 지시하였다.

조종사가 정상 비행 상태를 유지하는 동안 교관은 조종사가 예측하지 못한 악화된 기상 조건을 부과하였다. 가시거리는 최초 10 km 이상의 양호한 상태를 유지하다가 점차 기상이 안 좋아지는 상황을 모사하였으며 최종적으로는 차폐의 시정 상태에서 조종사의 비행특성 및 경향성을 확인하였다.

정형화된 표준데이터를 추출하기 위해 각 비행단계나 가시 거리가 바뀌는 시점의 시간은 정확히 기록하고 유지하였다.

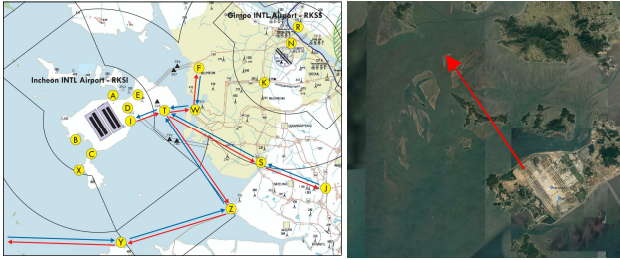


그림 2. 비행 시나리오
Fig. 2. Flight Scenario

3-3 데이터 수집 및 전처리

본 연구에서 비행 시뮬레이션을 수행한 결과를 CSV 파일로 추출하여 결측치가 있는 행이나 열을 모두 제거하였고 비행데이터 원자료는 시간(Time), 피치(Pitch), 뱅크(Bank), 방향(Heading), 속도(Speed), 고도(Alt) 등의 파라미터로 구성된다.

비행단계는 상승비행, 수평비행, 강하비행 세 단계로 구분하였고, 가시거리는 매우 좋음, 보통, 매우 나쁨의 3단계로 분류하여 코딩하였다. 마지막으로 비행 후에 설문을 실시하여 비행착각에 들어간 조종사와 들어가지 않은 조종사를 구분하여 코딩을 수행하였다. 지도학습을 위해 각 데이터에 대응하는 레이블은 항공기 자세값(피치, 뱅크, 방향, 속도, 고도, 수직속도), 조종사 입력값(x, y, rx, power)으로 할당하였다.

그림 3은 피치, 뱅크 등에 대하여 결측치 처리를 한 후 그래프를 시각화한 내용으로 이 외 모든 비행데이터의 시계열 데이터도 추세나 계절성 패턴은 관찰되지 않았고, 가시거리가 점차 악화되는 시점에서는 변동성이 커지는 것을 시각적으로 확인할 수 있었다. 추출된 모든 비행데이터는 총 54,000개의 시계열 데이터로 비행착각을 경험한 데이터는 51,300개, 비행착각을 경험하지 않은 데이터는 2,700개로 데이터가 분류되었다. 두 개 클래스의 데이터 비율 불균형으로 인해 발생하는 모델의 성능 저하를 방지하기 위하여 두 클래스의 비율을 유지하여 학습 데이터와 검증 데이터를 구성하였다.

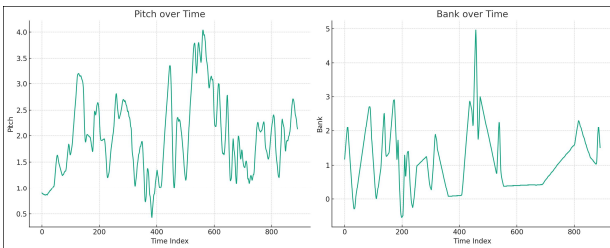


그림 3. 피치와 뱅크 데이터 시각화
Fig. 3. Pitch and bank data visualization

3-4 분석결과

표 1과 같이 엑스트라 트리의 정확도는 99.6%이고 정밀도와 재현율도 높은 비율로 측정되었다. 조화평균은 99.3%로 전반적으로 비행착각에 빠진 경우와 빠지지 않은 경우를 잘 분류하고 있음을 나타낸다. 랜덤 포레스트도 높은 성능을 보이고 있으나 엑스트라 트리에 비해 약간 낮았다.

표 1. 분류성능평가지표

Table 1. Classification Performance Metrics

Algorithm	Accuracy	Precision	Recall	F1Score
Extra tree	99.6%	99.4%	99.1%	99.3%
Random forest	99.3%	98.6%	99.1%	98.8%

두 개의 모델 모두 성능이 너무 높아 과적합의 우려가 있어, K겹 교차 검증(K-fold cross validation)으로 전체 데이터 세트를 10개의 동일한 부분으로 나누고 교차 검증을 통해 일반화 성능을 향상시켰다. 그 결과 표 2와 같이 엑스트라 트리 모델은 전체적으로 성능이 향상되었고, 랜덤 포레스트 모델은 재현율에서 약간의 하락이 있었다.

표 2. 교차검증 후 분류성능평가지표

Table 2. Classification Performance Metrics after K-fold cross validation

Algorithm	Accuracy	Precision	Recall	F1Score
Extra tree	99.7% (+0.1)	99.5% (+0.1)	99.5% (+0.4)	99.5% (+0.2)
Random forest	99.4% (+0.1)	99.1% (+0.5)	98.9% (-0.2)	99.0% (+0.2)

그림 4의 특성 중요도는 각각의 모델이 판단 과정에서 중요하게 생각하고 있는 주요 특징들을 보여준다. 랜덤 포레스트 모델의 특성 중요도는 방향(Heading), 파워(Power), 페달 변위(Rx), 속도(Speed), 뱅크(Bank), 수직속도(Vs), 피치(Pitch), X축(X) 순으로 나타났다. 엑스트라 트리 모델의 특성 중요도는 방향(Heading), 파워(Power), 페달 변위(Rx), 속도(Speed), 뱅크(Bank), 수직속도(Vs), 피치(Pitch), Y축(Y), X축(X) 순으로 나타났다.

상위 3개의 중요한 특성은 두 개의 모델 동일하게 방향(Heading), 파워(Power), 페달 변위(Rx)로 나타났다. 방향(Heading)은 비행 경로와 연관이 깊으며, 잘못된 방향으로의 비행은 항로를 벗어나기 때문에 장애물과 충돌 사고로 이어질 수 있는 중요한 요소이다. 파워(Power)와 페달 변위(Rx)는 회전의

항공기 토류와 꼬리날개 제어에 직접적인 영향을 미치는 요인으로 항공기 안정성과 관련이 깊다고 할 수 있다. 따라서 항공기 방향과 토류, 조종사 페달 입력량은 실제 비행착각과 관련된 중요한 변수로 판단된다.

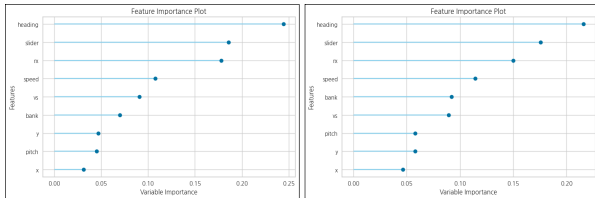


그림 4. 랜덤 포레스트(좌) 및 엑스트라 트리(우) 특성중요도
Fig. 4. Random Forest(left) & Extra Trees(right) feature importance

V. 결 론

회전의 항공기는 비행착각으로 인한 사고가 지속적으로 발생하고 있고, 그 결과는 대부분 치명적인 사망사고로 연결된다. 항공기 사고는 대부분 인적요인에 의한 것으로 조종사의 능력이 중요하지만 비행착각에 의한 사고는 스스로가 인지하지 못한 상태에서 발생하는 경우가 많다.

이와 같이 조종사의 능력이나 성능을 정확하게 측정할 수 있다면 이를 적용하여 교육훈련 방법이나 운항품질을 향상시킬 수 있을 것이다. 과거의 전통적인 교육방법은 교관이 관찰하고 정성적으로 판단하는 방법에 대부분 의존하여 왔고 비행데이터를 활용한 정량적인 평가 방법은 제한적이었다.

본 연구에서는 시뮬레이션을 통해 항공기 자세와 조종사 입력값에 대한 비행데이터를 추출하고 가시거리에 따른 비행착각 유무를 분류하였다. 이후 데이터 전처리와 교차 검정을 통해 비행착각 유무를 식별할 수 있도록 랜덤 포레스트와 엑스트라 트리 모델로 학습을 시켰다. 그 성능을 비교한 결과 모두 높은 성능을 보였지만 엑스트라 트리 모델이 약간의 우위를 보였다.

한편, 두 모델에서 모두 방향(Heading), 파워(Power), 페달 변위(Rx)의 순으로 특성 중요도가 유사하게 나타났는데, 이는 비행착각과 관련된 중요한 변수로 항공기 방향과 토류, 조종사 페달 입력량이 비행착각 발생의 결정적인 역할을 하는 것으로 판단된다.

본 연구에서 진행된 데이터는 시뮬레이터에서 획득한 자료이므로 추후에는 실제 비행데이터로 더욱 다양하게 분석할 필요가 있다.

Acknowledgments

본 과제(결과물)는 2023년도 교육부의 재원으로 한국연구재단의 지원을 받아 수행된 지자체-대학 협력기반 지역혁신 사업의 결과입니다.(2021RIS-004)

- [1] Vreeken, J.: Helicopter flight in a degraded visual environment, European Aviation Safety agency, Cologne Germany: Research Project EASA, 2011.02.
- [2] Tu, Min. Yu., Cheng, C. C., Hsin, Y. H., Huang, W. W., Li, F. L., Hu, J. M., ... & Lai, "Analysis of in-flight spatial disorientation among military pilots in Taiwan," *Journal of Medical Sciences*, Vol. 41, No. 1, pp. 22-28, Feb. 2021.
- [3] B. S. Gu, "A Search on the Spatial Orientation in Flight," *Journal of the Korean Society for Aeronautical Science and Flight operation*, Vol. 2, pp. 139-182, Jun. 1994.
- [4] N. S. Park, "Research trends on deep learning for anomaly detection of aviation safety," *Electronics and Telecommunications Trends*, Vol. 36, No. 5, PP. 82-91, 2021
- [5] G. H. Lee, S. K. Song, "Predicting Flight Delays based on Deep Neural Network," in *Proceedings of the Korea Information Science Society*, Korea, pp. 1075-1076, Jun. 2021.
- [6] T. H. Kim, M. J. Song, E. J. Choi, B. S. Kim, Y. H. Moon, "Flight data prediction method using LSTM based-deep learning model," in *Proceedings of Korean Society for Aeronautical and Space Sciences*, Korea, PP. 968-969, 2021.
- [7] Fernández, A., Martinez, D., Hernández, P., Cristóbal, S., Schwaiger, F., Nunez, J. M., & Ruiz, J. M, "Flight data monitoring(FDM) unknown hazards detection during approach phase using clustering techniques and AutoEncoders," in *Proceedings of the Ninth SESAR Innovation Days*, Athens, Greece, PP. 2-5, 2019.
- [8] Z. Shi, M. Xu, Q. Pan, B. Yan and H.Zhang, "LSTM-based Flight Trajectory Prediction," *International Joint Conference on Neural Networks (IJCNN)*, Rio de janeiro, Brazil, pp. 1-8, 2018.
- [9] Breiman, L, "Random forests," *Machine learning*, Vol. 45, pp. 5-32, Oct. 2001.
- [10] Geurts, P., Ernst, D., & Wehenkel, L, "Extremely randomized trees," *Machine learning*, Vol. 63, pp. 3-42, Mar. 2006.