

Detection of Pilot's Cognitive Workload Levels using Flight Performance and Heart Rate Variability Features

Sungho Kim, Dongwoo Kang

R.O.K. Air Force Academy, Department of Systems Engineering, Chungcheongbuk-do, 28187

비행 수행도와 심박변이도 특징을 활용한 조종사 인지부하 수준 감지

김성호, 강동우

공군사관학교 시스템공학과

Corresponding Author

Sungho Kim

R.O.K. Air Force Academy, Department of Systems Engineering, Chungcheongbuk-do, 28187

Mobile: +82-10-4956-7866

Email: dilemma37@naver.com

Received: October 15, 2018

Revised: October 16, 2018

Accepted: October 31, 2018

Copyright©2018 by Ergonomics Society of Korea. All right reserved.

© 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.

Objective: The purpose of this study is to analyze characteristics of flight performance and heart rate variability features and develop cognitive workload classification model in order to detect cognitive workload levels of pilots.

Background: Pilots can experience high cognitive workload conditions by performing various tasks simultaneously during air operation mission. Accurate detection of the cognitive workload state is required to improve the survivability of pilots and ensure flight safety.

Method: Five student pilots in 20s performed dual task consisting of simulated flight task with baseline task and three levels of N-back task (0-back, 1-back, 2-back). Three flight performance features (standard deviation of altitude, heading, airspeed) and two heart rate variability features (standard deviation of NN (SDNN), Low Frequency to High Frequency ratio (LF/HF)) were measured from the flight simulation experiment with photoplethysmogram (PPG) sensor. Cognitive workload classification model was developed by pre-processing input features and output classes, applying five classifiers (Decision tree (DT), Support vector machine (SVM), K-nearest neighbor (KNN), Bayesian, Ensemble), and validating the model performance.

Results: High cognitive workload level (2-back) was 327%, 161%, and 1.2% higher than baseline in terms of standard deviation of altitude, standard deviation of airspeed, and LF/HF respectively. DT, SVM, KNN, Bayesian, and Ensemble classifier performance was 94.4%, 91.1%, 96.6%, 93.3%, and 96.6% with two output classes, 62%, 51.4%, 64.6%, 57.1%, and 64.6% with three output classes, 36.3%, 36%, 48%, 36.3%, and 49% with four output classes respectively.

Conclusion: The best classifier performance for detecting cognitive workload levels of pilots was acquired from an ensemble classifier with two output classes (baseline, 2-back) with an accuracy of 96.6%.

Application: The cognitive workload classification model using the ensemble classifier in this study can contribute to the development of a system capable of providing warning signal in real time under pilot's cognitive overload situation.

Keywords: Cognitive workload, Flight performance, Heart rate variability, Ensemble classifier

1. Introduction

최근 전투기 성능이 급격히 발전함에 따라 성공적인 항공작전 임무 수행을 위한 조종사들의 역할이 증대되고 있다. 조종사들은 Figure 1과 같이 외부 환경에 대한 상황인식 유지와 내부 계기 정보에 대한 모니터링를 통해 상당한 양의 정보를 연속적으로 수용 및 기억하고, 정보처리과정을 거쳐 미래를 계획하여야 하며, 동시에 작전 수행을 위해 타 항공기, 작전지휘부, 항공통제전력 등과 통신을 유지하여야 한다. 즉, 조종사들은 항공작전 임무 수행 중 다양한 과업을 동시에 수행하여야 함에 따라 높은 인지부하 상태에 도달할 수 있으며, 이는 임무의 집중도 및 효율성을 저하시킬 뿐만 아니라 상황인식을 상실하게 하여 치명적인 사고로 이어질 수 있다. 따라서, 조종사들의 임무목표 달성, 전투원 생존성 증진, 비행안전 보장 등을 위해서는 인지부하의 개념을 이해하고, 인지부하 상태를 파악하는 연구가 필요하다.



Figure 1. Pilot's mission environment

인지부하는 일반적으로 과제 해결에 요구되는 인지자원의 양이 인지구조가 보유하고 있는 자원의 용량을 초과할 때 발생한다(Sweller et al, 1998). 또한, 인지부하 이론에 따르면 인지부하는 과업 자체가 가지고 있는 난이도와 복잡성에 따라 발생하는 내재적 인지부하 (Intrinsic cognitive workload), 과업의 전달방식에 따라 발생하는 외재적 인지부하(Extrinsic cognitive workload), 그리고 작업기억(Working memory) 범위 내에서 과업과 관련 있는 새로운 정보를 기존 체계에 통합하기 위한 정신적 노력에 따라 발생하는 본질적 인지부하 (Germane cognitive workload)로 분류될 수 있다(Coyne et al, 2009). 특히, 이중 과업과 같은 비효율적인 과업 수행은 내재적 인지부하를 증가시켜 인지조절 기능을 위축시키고 스트레스 및 피로를 초래하여 상황 인지의 오류를 유발시킬 수 있다.

한편, 비행 수행도와 심박변이도를 활용하여 조종사들의 인지부하 상태를 측정하기 위한 다양한 연구들이 수행되었으나, 인지부하의 정도를 정확히 감지하기에는 미흡한 측면이 있었다. 비행 수행도에 의한 인지부하 측정은 조종간 조작 정도, 기준 고도 대비 이탈 정도 등을 통해 조종사의 인지부하 상태를 측정하는 방법으로, 비침습적이나 정확성이 낮다는 단점을 가지고 있다. 예를 들어, Gawron (2002)은 비정상 비행자세 회복훈련(Upset recovery training) 수행도를 평가하기 위해 비행 수행도 기준들을 개발하였다. 심박변이도에 의한 인지부하 측정은 Electrocardiogram (ECG), Photoplethysmogram (PPG) 센서 등으로 획득된 생체신호 데이터를 통해 조종사의 인지부하 상태를 판단하는 기술로, 다른 감지 방법에 비해 신뢰성 및 정확성이 높지만 침습적이고 항공기 내부와 같은 제한된 환경에서 생체신호를 획득하기 어렵다는 단점을 가지고 있다. Lahtinen et al. (2007)은 조종사에게 모의비행훈련 시뮬레이터 환경 하에서 다수의 적기 출현 상황을 부여하고 ECG 센서를 통해 측정된 심박수와 비행구간별 수행도 간의 관계를 규명하였다. 그러나, 조종사들의 인지부하 상태를 정확히 파악하기 위해서는 인지부하 수준별 상태를 반영하는 유효 입력 특징들을 선정하고, 고성능의 모형을 적용하여 인지부

하 수준을 효과적으로 분류하는 것이 요구된다.

본 연구는 조종사들의 인지부하의 정도를 감지하기 위해 비행 수행도 및 심박변이도 척도들의 특성을 분석하고, 인지부하 분류 모형을 개발하였다. 조종사들의 인지부하 상태는 조종사들이 이중 과업을 수행하면서 획득된 비행 수행도 및 심박변이도 데이터를 비교 분석하여 파악되었다. 인지부하 분류 모형은 인지부하 상태를 반영하는 비행 수행도 및 심박변이도 특징들을 선별하고, 모형의 종류 및 출력 조건에 따른 성능 비교를 통해 최종 선정되었다.

2. Method

2.1 Participants

본 연구는 20대 학생 조종사 5명이 참여하였다. 실험 참여자들의 평균 연령은 23.8세(SD=0.84)이고, 모의비행훈련 시뮬레이터를 통한 평균 비행경력 40.1시간(SD=2.11)이었다. 실험 참여자들은 원활한 실험 수행을 위하여 심혈관계 질환이 없고 인지 능력에 이상이 없는 자로 선정하였으며, 실험 전날 실험 결과에 영향을 미칠 수 있는 활동을 금하였다. 한편, 본 연구의 실험은 실험 참여자들의 동의하에 진행되었으며, 소정의 실험 참여비가 지급되었다.

2.2 Apparatus

본 연구는 비행 수행도와 심박변이도를 측정하기 위해 Figure 2와 같이 모의비행훈련 시뮬레이터와 PPG 센서를 각각 활용하였다. 실험 참여자의 비행 수행도는 Pilot Aptitude Research Equipment (PARE) 시뮬레이터(Dodaam Systems Ltd., Korea)를 활용하였다. PARE 시뮬레이터는 학생 조종사들이 훈련 Cockpit에서 모의 비행을 수행하는 동안 운용 Console에서 실시간으로 모의비행의 수행도를 평가하고, 학생 조종사들의 고도, 속도, 방향 등 항공기 제원과 관련된 평가 요소 데이터를 초 단위로 추출 가능하며, 기준 값 범위 내 유지하는 정도를 점수화하여 제공한다(R.O.K Air Force, 2012). 예를 들어, 조종사의 수평 비행 훈련 시 속도의 기준 값이 160km/h일 경우 속도처리 점수는 ± 10 km/h 범위 이내는 10점, $\pm 10\sim 15$ km/h 범위는 8점, $\pm 15\sim 20$ km/h 범위는 6점, $\pm 20\sim 25$ km/h 범위는 4점, $\pm 25\sim 30$ km/h 범위는 2점, 그리고 ± 30 km/h 이상 범위는 0점으로 정량적으로 채점할 수 있다.

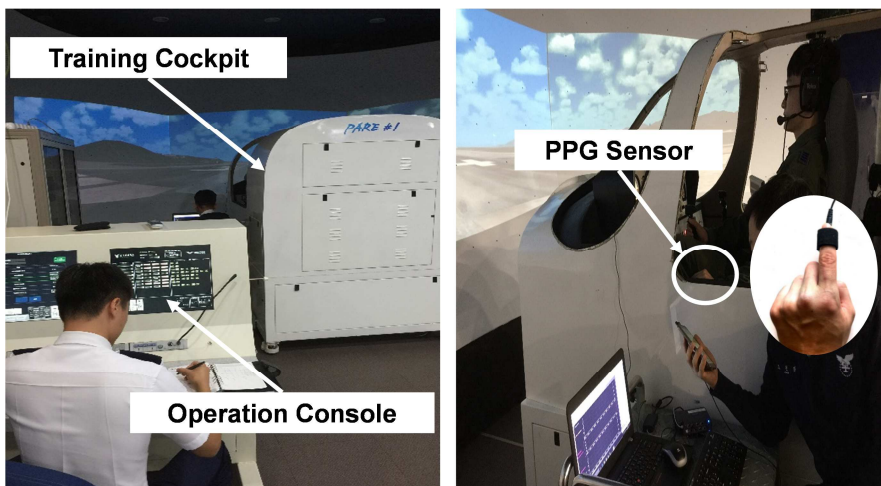


Figure 2. Pilot Aptitude Research Equipment (PARE) simulator and Photoplethysmogram (PPG) sensor

실험 참여자의 심박변이도는 BioGraph Infiniti (Thought Technology Ltd., Canada)의 PPG 센서를 활용하여 측정되었다. PPG 센서는 LED

광원을 말초모세혈관에 조사하여 PPG 센서 내 포토다이오드(Photodiode)에서 피부, 조직, 혈액 등에 의해 일부 흡수 및 반사되어 광량이 감소된 광원을 수신하는데, 이때 심장박동에 의한 혈류변화를 제외하고 피부, 조직 등은 변하지 않는 성분이기 때문에 수신된 광원의 양을 혈류파형 신호로 변환하여 심박변이도를 간접적으로 산출할 수 있다(Kim et al., 2016). 본 연구에서는 PPG 센서를 Figure 2와 같이 실험 참여자의 왼손 검지에 부착하였다.

2.3 Cognitive workload tasks and measures

본 연구는 가상의 인지부하 상황을 재현하기 위해 학생 조종사들이 모의 비행 중 N-back 과업을 수행하는 이중 과업 방법을 활용하였다. 이중 과업 방법은 일차 과제에 의해 소비된 인지적 자원이 많을수록 이차 과제의 수행도는 잉여 인지적 자원이 부족하기 때문에 저하될 것이라고 가정하는 방법이다(Wickens et al., 2003). 일차 과제인 모의비행 과업은 학생 조종사들이 고도 200m, 수평속도 160km/h, 그리고 방향 340°의 기준 비행제원을 지속 유지하도록 수평 비행을 부여하였다. 이차 과제인 N-back 과업은 Figure 3과 같이 학생 조종사들이 음성으로 연속적으로 제시되는 임의의 숫자를 기억하여 대답하도록 요구하였다(Park et al., 2010; Son and Park, 2013). N-back 과업에서 제시되는 숫자는 1~9 사이 정수를 활용하였으나, 숫자 4는 숫자 3과의 한글 발음 유사성을 고려하여 제외하였다. N-back 과업은 각 난이도별 30초 동안 10개의 숫자를 제시하였으며, N-back 난이도를 임의적으로 제시하여 N-back 난이도에 대한 학습효과를 최소화하였다. 또한, N-back 과업의 개별 숫자의 음성제시 소요시간은 1초였고 각 개별 숫자의 음성제시 간 시간간격은 3초였으며 각각 다른 난이도의 N-back 과업을 수행한 후 학생 조종사들이 인지 능력을 회복할 수 있도록 2분의 휴식을 포함하였다.

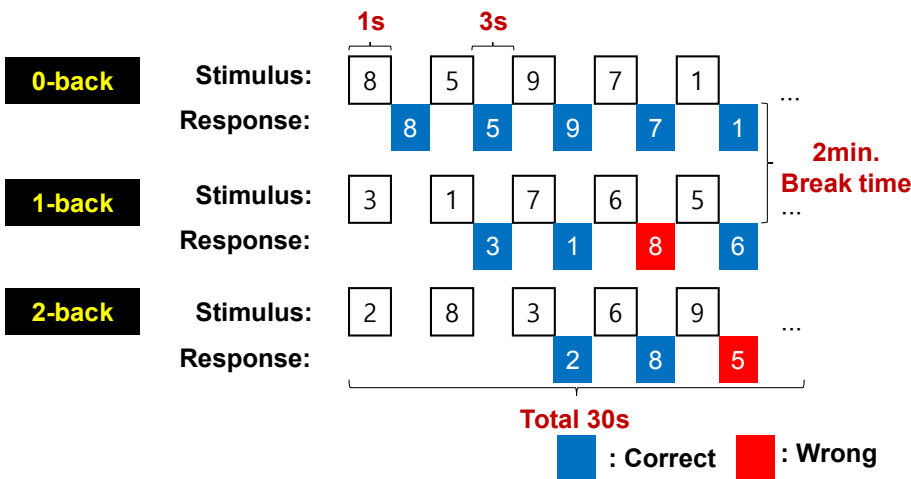


Figure 3. N-back task design

본 연구는 인지부하 수준을 독립변수로, 3가지 비행 수행도 척도와 2가지 심박변이도 척도를 종속변수로 활용하였다. 본 연구에서는 학생 조종사의 인지부하를 3수준으로 발생시키기 위해 제시되는 숫자를 듣고 그 숫자를 말로 대답하는 0-back을 인지부하 저수준, 한 단계 전에 제시되는 숫자를 기억하여 대답하는 1-back을 인지부하 중수준, 그리고 두 단계 전에 제시되는 숫자를 기억하여 대답하는 2-back을 인지부하 고수준으로 활용하였다. 한편, 본 연구에서는 비행 수행도 척도로 고도, 속도, 그리고 방향의 표준편차를, 심박변이도 척도로 심박간격표준편차(Standard Deviation of NN interval, SDNN), 저주파수 대 고주파수 비율(Low Frequency to High Frequency ratio, LF/HF)을 활용하였다. 고도, 속도, 그리고 방향의 표준편차는 학생 조종사들의 항공기 3축 방향 제어 능력을 각각 의미한다. 또한, SDNN은 시간 영역의 척도로 체내 스트레스 지수 및 저항도를, LF/HF는 주파수 영역의 척도로 자율신경계의 전체적인 균형 정도를 반영하는 것으로 알려져 있다(Kim and Min, 2015).

2.4 Experimental procedure

본 연구의 실험은 실험 준비, 예비 실험, 본 실험, 그리고 설문 조사의 네 단계로 진행되었다. 실험 준비 단계에서는 실험 참여자에게 실험 소개 및 실험 참여 동의를 구하였다. 또한, 실험 참여자가 PARE 시뮬레이터에 착석 후 PPG 센서를 부착하고 생체신호 안정화를 위해 5분 간 대기하도록 하였다. 예비 실험 단계에서는 실험 참여자들이 안정적으로 이중 과업을 수행하도록 수평 비행 및 N-back 과업을 연습할 수 있는 기회를 3회 제공하였다. 본 실험 단계에서는 인지부하 수준별 비행 수행도 및 심박변이도 데이터를 학생 조종사 1명당 총 3회 반복 실험하여 추출하였다. 이때, 인지부하가 부여된 상황과 인지부하가 부여되지 않은 상황을 비교하기 위해 N-back 과업을 수행하지 않고 수평 비행만 수행하는 기준 과업(Baseline task)을 병행하여 실시하였다. 설문 조사 단계에서는 실험 간 특이사항 발생 여부 및 실험진행 관련 의견을 조사하였다.

2.5 Data analysis

비행 수행도 및 심박변이도 데이터는 데이터 획득, 분석구간 결정, 데이터 산출, 통계 분석의 네 단계를 통해 분석되었다. 데이터 획득 단계에서는 비행 수행도의 경우 1Hz 샘플링 주파수로, 심박변이도의 경우 256Hz 샘플링 주파수로 데이터가 수집되었다. 분석구간 결정 단계에서는 수평 비행 중 인지부하가 부여되지 않은 기준 과업 상황과 N-back 과업을 통해 인지부하가 수준별로 부과되는 상황의 시작 지점과 종료 지점 간 30초 구간이 추출되었다. 데이터 산출 단계에서는 비행 수행도의 경우 분석구간 내 고도 200m, 수평속도 160km/h, 그리고 방향 340°의 기준 비행제원과 대비하여 각각의 실제 비행제원 값과의 편차를 산출하였고, 심박변이도의 경우 개인별 PPG 신호 데이터의 크기를 비교할 수 있도록 기존 연구논문을 참고하여 중간값(Median)으로 각각의 데이터를 나눠 표준화하였다 (Tjolleng et al., 2017). 통계 분석 단계에서는 학생 조종사들의 인지부하 수준별 유의한 차이가 있는지 파악하기 위해 개별 학생 조종사들의 인지부하 수준별 30초 분석구간을 10초 윈도우 크기로 분할하여 윈도우별 평균값들을 산출한 후 유의수준 0.05에서 일원분산 분석(One-way ANOVA)을 수행하였다. 인지부하 수준별 효과에 대해서는 유의 수준 0.05에서 Scheffe 검정을 통해 다중 비교를 수행하였고, SPSS (Ver. 21)가 통계 분석에 활용되었다.

3. Results

3.1 Comparison of flight performance and heart rate variability features

본 연구는 조종사들의 인지부하 수준에 따른 비행 수행도 및 심박변이도 특성을 비교 분석하였다. 통계 분석 결과, 인지부하 요인은 Table 1과 같이 유의수준 0.05에서 비행 수행도 측면에서 고도와 수평속도의 표준편차가, 심박변이도 측면에서 LF/HF가 유의한 차이가 있는 것으로 파악되었다. 먼저, 비행 수행도 측면에서 학생 조종사들은 Figure 4.a와 같이 인지부하 수준이 증가할수록 고도와 수평속도의 표준편차가 대체적으로 증가하는 것으로 나타났다. 고도의 표준편차는 인지부하 고수준(2-back)이 기준 과업(Baseline) 대비 327% (MD=5.78)만큼 높게 나타났고, 수평속도의 표준편차는 인지부하 고수준(2-back)이 기준 과업(Baseline) 대비 161% (MD=2.60)만큼 높

Table 1. ANOVA Summary Table

Dependent variables	Independent variables	Cognitive workload levels
Flight performance	Standard deviation of altitude	$F(3, 56) = 6.608^*$
	Standard deviation of heading	$F(3, 56) = 0.445$
	Standard deviation of airspeed	$F(3, 56) = 2.989^*$
Heart rate variability	Standard deviation of NN interval	$F(3, 56) = 0.509$
	Low frequency to high frequency ratio	$F(3, 56) = 7.086^*$

* $p < .05$

게 나타났다. 한편, 방향의 표준편차는 인지부하 고수준(2-back)이 기준 과업(Baseline) 대비 22% (MD=0.70)만큼 높게 나타났으나, 통계적으로 유의하지는 않았다. 다음으로, 심박변이도 측면에서 학생 조종사들은 Figure 4.b와 같이 인지부하 수준이 증가할수록 LF/HF가 증가하는 경향이 있는 것으로 나타났다. LF/HF는 인지부하 고수준(2-back)이 기준 과업(Baseline) 대비 1.2% (MD=0.01)만큼 높게 나타났다. 한편, SDNN은 인지부하 고수준(2-back)이 기준 과업(Baseline) 대비 1.8% (MD=-0.02)만큼 낮게 나타났으나, 통계적으로 유의하지는 않았다.

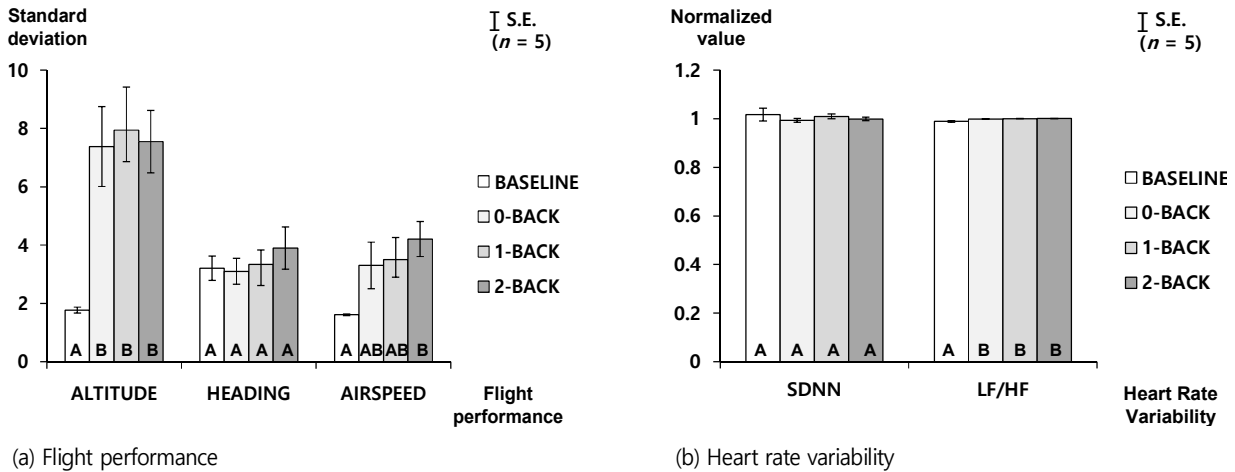


Figure 4. Comparison of flight performance and heart rate variability by pilot's cognitive workload levels (Alphabet shows statistical significance ($\alpha=0.05$))

3.2 Development of classification model for cognitive workload levels

본 연구는 학생 조종사의 인지부하 수준을 효과적으로 파악하기 위해 데이터 전처리, 분류기 적용, 분류기 성능 검증의 세 단계 절차를 통해 분류 모형을 개발하였다. 먼저, 데이터 전처리 단계에서는 실험의 종속변수로 활용되었던 3가지 비행 수행도 척도(고도, 속도, 방향의 표준편차)와 2가지 심박변이도 척도(SDNN, LF/HF)의 총 5가지 척도 중 불필요한 입력 특징의 포함 여부를 확인하였다. 5가지 척도에 대해 개별 척도의 분산이 0에 가까워 종속변수로 의미를 가지지 못하는지 판단하는 영분산분석과 각 변수 사이에 상관관계가 1에 가까워 종속변수로 의미가 없는지 판단하는 상관 분석을 수행한 결과, 입력 특징으로 포함되기에 불필요한 척도는 없는 것으로 파악되었다. 또한, 분류 Class 개수에 따른 모형 성능 비교를 위해 총 3가지 분류 Class 조건(2개 분류 Class: 기준 과업, 2-back 과업; 3개 분류 Class: 기준 과업, 0-back 및 1-back 과업, 2-back 과업; 4개 분류 Class: 기준 과업, 0-back 과업, 1-back 과업, 2-back 과업)에 대해 실험을 수행하였다.

다음으로, 분류기 적용 단계에서는 실험을 통해 획득된 데이터를 기반으로 R (Ver 3.5.1) 프로그램을 활용하여 각각 5가지 분류 방법론(의사결정나무, 서포트 벡터 머신, K-근접 이웃 분류기, 베이지안 분류기, 앙상블 분류기)을 이용하여 분류 모형을 구축하였다. 의사결정나무(Decision Tree, DT) 기법은 대표적인 규칙 기반의 분류기로 본 연구에서는 Gini 인덱스를 이용하여 분류변수를 정하고 사전가지치기를 통해 분류기를 구축하였다. 서포트 벡터 머신(Support Vector Machine, SVM) 기법은 대표적인 kernel 기반의 분류기로 본 연구에서는 볼록최적화(Convex optimization) 문제를 풀어 분류기를 구축하였다. K-근접 이웃(K-Nearest Neighbor, KNN) 기법은 분류하고자 하는 새로운 데이터들과 근접한 K개의 주변 학습 데이터들을 탐색하여 그 중 가장 많은 데이터가 속한 Class로 배정하는 분류기로 본 연구에서는 근접 5개(K=5)의 데이터를 탐색하여 분류기를 구축하였다. 베이지안(Bayesian) 분류 기법은 모든 변수가 조건부 독립임을 가정하고 데이터가 속할 확률이 가장 높은 Class로 배정하는 분류기로 본 연구에서는 가장 단순한 Naive Bayesian 분류기를 활용하여 구축하였다. 앙상블(Ensemble) 분류 기법은 학습 데이터에 대해 의사결정나무, 서포트 벡터 머신, K-근접 이웃, 베이지안 분류 방법론을 결합한 분류기로 본 연구에서는 4가지 방법론을 통해 분류한 결과를 비교하여 가장 많은 분류기가 예측하는 Class로 최종 의사결

정 하도록 구축하였다.

마지막으로, 분류기 성능 검증 단계에서는 5중 교차 검증(5-fold cross validation)을 통해 학습 데이터의 정확도(Accuracy)를 계산하여 결과를 산출하였다. 분류기별 성능 검증 결과, 앙상블 분류기, K-근접 이웃 분류기, 의사결정나무 분류기, 베이지안 분류기, 서포트 벡터 머신 분류기 순으로 정확도가 높은 것으로 파악되었다. 한편, 분류 Class의 개수가 증가할수록 분류기의 정확도는 점차 감소하는 것으로 나타났다. Table 2는 3가지 분류 Class 조건에 따른 5가지 분류기별 정확도 산출 결과를 보여준다.

Table 2. Accuracy of each classifier by classification class conditions

Accuracy of classifier	Number of classification classes		
	2	3	4
Decision tree	94.4±0.079%	62.0±0.113%	36.3±0.143%
Support vector machine	91.1±0.145%	51.4±0.249%	36.0±0.083%
K-nearest neighbor	96.6±0.075%	64.6±0.124%	48.0±0.187%
Bayesian	93.3±0.091%	57.1±0.218%	36.3±0.073%
Ensemble	96.6±0.075%	64.6±0.187%	49.0±0.243%

M ± SD

4. Discussion

본 연구는 비행 수행도 및 심박변이도 척도들을 입력 특징으로 하는 앙상블 분류기를 적용하여 인지부하 상태를 수준별로 높은 정확도로 분류하는 모형을 개발하였다. 본 연구에서 사용된 앙상블 분류기는 인지부하 상황 미부여 상태를 나타내는 기준 과업과 고수준의 인지부하 상황이 부여된 상태를 나타내는 2-back 과업의 2개 분류 Class 기준일 경우 96.6%의 높은 정확도로 인지부하 수준을 분류하는 것으로 파악되었다. 기존 연구들(Katisis et al., 2008; Noel et al., 2005; Solovey et al., 2014)은 서포트 벡터 머신, 인공신경망 등과 같은 특정 모형만을 활용하여 조종사 또는 운전자들의 상태 수준을 분류해왔다. 그러나, 의사결정나무, 서포트 벡터 머신, K-근접 이웃 분류기, 그리고 베이지안 분류기 등과 같은 기존 분류기들은 높은 분류 및 예측 정확도를 나타내는 데이터의 형태가 상이한 것으로 알려져 있다(Dietterich, 2000). 따라서, 앙상블 분류기를 활용하면 상대적으로 낮은 정확도를 나타내는 분류기들의 의견을 포함하면서도 분류기별 성능 비교를 통해 가장 높은 정확도의 결과를 반영할 수 있어 특정 데이터 형태에 민감하지 않게 보다 강력한 분류 결과를 획득할 수 있다.

또한, 조종사들의 인지부하 상태를 효과적으로 감지하기 위한 인지부하 수준별 분류 Class는 인지부하 상황 미부여 상태를 나타내는 기준 과업과 고수준의 인지부하 상황이 부여된 상태를 나타내는 2-back 과업의 2개 분류 Class인 것으로 파악되었다. 인지부하 수준 분류 Class 조건이 2개, 3개, 그리고 4개일 때 각각 평균 94.4%, 59.9%, 그리고 41.1%의 정확도 순으로 개별 인지부하 수준을 분류 가능한 것으로 나타났다. 이는 분류 Class 개수가 증가할수록 인접 분류 Class간 입력 특징들의 특성 차이가 불명확해지고 분류 Class별 데이터의 크기가 불충분해짐에 따라 분류기의 정확도가 점차 감소하는 것으로 판단된다. 따라서, 본 연구에서 활용된 5가지 입력 특징(고도 표준편차, 수평속도 표준편차, 방향 표준편차, SDNN, LF/HF) 외에 인지부하 수준 분류에 관련성 높은 변수들을 추가하고 다수의 조종사들에 대한 데이터를 수집한다면, 보다 세분화된 인지부하 수준을 정확히 분류하는데 도움을 줄 수 있다.

비행 수행도 및 심박변이도 특성 분석 결과, 고도의 표준편차와 수평속도의 표준편차는 인지부하 고수준이 기준 과업에 비해 각각 327%와 161% 정도 유의하게 높은 것으로 파악되었는데, 이는 인지 과부하에 따른 학생 조종사들의 항공기 자세제어 능력이 감소한 것을 반영하는 것으로 판단된다. 한편, 본 연구에서는 정상 상황 하 수평 비행을 가정하여 고도, 속도, 그리고 방향의 표준편차를 비행 수행도 척도로 활용하였으나, Gawron (2002)에서 활용된 Throttle 또는 Rudder의 정상 제어 최초 시간(Time to first throttle or rudder

input), 회복시간(Time to recover), 고도손실(Altitude loss) 등과 같은 비행 수행도 척도를 적용하면 비정상 상황 하 비행자세 회복훈련(Upset recovery training) 시 조종사들의 인지부하 수준을 판단하는데 기여할 수 있다. 또한, LF/HF는 인지부하 고수준이 기준 과업에 비해 1.2% 정도 유의하게 높은 것으로 나타났는데, 이는 불안, 공포, 긴장 상태 등의 교감신경계 활성화도가 부교감신경계보다 우세한 것을 반영하는 것으로 사료된다. SDNN은 통계적으로 유의하지는 않았으나 인지부하 고수준이 기준 과업에 비해 1.8% 정도 낮은 것으로 확인되었는데, 이는 학생 조종사들의 이종 과업에 대한 스트레스 대처 능력이 감소되면서 자율신경계 조절 능력이 저하된 것으로 해석할 수 있다.

추후 조종사의 인적 정보, 환경 정보 등의 변수들을 포함하여 인지부하 분류 모형을 개발하고, 신규 학생 조종사들에 대한 데이터를 활용하여 개발된 모형의 성능을 검증하며, 보다 침습적인 생체신호 센서를 활용하는 것이 필요하다. 먼저, 본 연구에서는 비행 수행도와 생체신호 척도를 기반으로 조종사의 인지부하 수준을 분류하는 연구를 수행하였으나, 모형의 분류 정확도를 향상시키기 위해서는 조종사의 성별, 나이 등과 같은 인적 정보와 기종, 비행경력 등과 같은 환경 정보를 변수로 포함시켜 종합적으로 고려하는 것이 요구된다. 다음으로, 본 연구에서는 5명의 학생 조종사들을 대상으로 실험을 수행하여 획득한 학습 데이터를 활용하여 인지부하 분류 모형을 개발하고 5중 교차 검증(5-fold cross validation)을 통해 성능을 검증하였으나, 분류 모형의 부족적합(Underfitting)을 방지하고 효용성을 보다 정확히 검증하기 위해서는 추가적인 학생 조종사들을 모집하여 실험을 통해 수집된 데이터를 활용하여 모형의 성능을 검증할 필요가 있다. 마지막으로, 본 연구에서는 실험 참여자들의 심박변이도를 측정하기 위해 PPG 센서를 활용하였으나, 실험 결과의 신뢰성 향상을 위해서는 손목밴드형 센서, 심장부위 직접 부착용 센서 등과 같이 침습성을 최소화할 수 있는 인터페이스를 활용할 필요가 있다. 본 연구의 앙상블 분류기를 활용한 인지부하 분류 모형이 보다 개선된다면, 인지과부하 상태를 조종사에게 실시간으로 제공할 수 있는 인지과부하 경고시스템에 적용될 수 있을 것으로 기대한다.

Acknowledgement

This research was supported by Agency for Defense Development.

References

- Coyne, J.T., Baldwin, C., Cole, A., Sibley, C. and Roberts, D.M., "Applying Real Time Physiological Measures of Cognitive Load to Improve Training", *Proceedings of HCI International Conference on Foundations of Augmented Cognition: Neuroergonomics and Operational Neuroscience*, (pp. 469-478), San Diego, USA, 2009.
- Dietterich, T.H., "Ensemble Methods in Machine Learning", *Proceedings of First International Workshop on Multiple Classifier Systems*, Cagliari, Italy, 2000.
- Gawron, V.J., Airplane Upset Training Evaluation Report, NASA/CR-2002-211405, Moffett Field, USA, National Aeronautics and Space Administration Ames Research Center, 2002.
- Katisis, C.D., Katertsidis, N., Ganiatsas, G. and Fotiadis, D.I., Toward Emotion Recognition in Car-Racing Drivers: A Biosignal Processing Approach, *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*, 38(3), 502-512, 2008.
- Kim, B.S. and Min, J.A., *Application and Interpretation of HRV in Stress Clinic*, First ed., Panmun Education, Korea, 2015.
- Kim, S.H., Choi, B.Y., Cho, T.H., Lee, Y.K., Koo, H.J. and Kim, D.S., Development of a Classification Model for Driver's Drowsiness and Waking Status Using Heart Rate Variability and Respiratory Features, *Journal of the Ergonomics Society of Korea*, 35(5), 371-381, 2016.
- Lahtinen, T.M., Koskelo, J.P., Laitinen, T. and Leino, T.K., Heart Rate and Performance during Combat Missions in a Flight Simulator,

Aviation Space and Environmental Medicine, 78(4), 387-391, 2007.

Noel, J.B., Bauer Jr, K.W. and Lanning, J.W., Improving Pilot Mental Workload Classification through Feature Exploitation and Combination: A Feasibility Study, *Computers and Operations Research*, 32(10), 2713-2730, 2005.

Park, Y.S., Park, S.W., Lee, T.Y. and Son, J.W., Impact of Cognitive Demand on Drivers' Visual Attention in Driving Simulator, *Proceedings of the Ergonomics Society of Korea Conference*, (pp. 180-183), Pyeongchang, Korea, 2010.

R.O.K. Air Force, *Operational Manual for Pilot Aptitude Research Equipment*, R.O.K. Air Force Headquarters, Korea, 2012.

Solovey, E.T., Zec, M., Perez, E.A.G., Reimer, B. and Mehler, B., Classifying Driver Workload Using Physiological and Driving Performance Data: Two Field Studies, *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*, (pp. 4057-4066), Toronto, Canada, 2014.

Son, J.W. and Park, M.O., The Impact of Cognitive Workload on Driving Performance and Visual Attention in Younger and Older Drivers, *Transaction of the Korean Society of Automotive Engineers*, 21(4), 62-69, 2013.

Sweller, J., Merrienboer, Jeroen J.G.V. and Paas, Fred G.W.C., Cognitive Architecture and Instructional Design, *Educational Psychology Review*, 10(3), 251-296, 1998.

Tjolleng, A., Jung, K.H., Hong, W.G., Lee, W.S., Lee, B.H., You, H.C., Son, J.W. and Park, S.K., Classification of a Driver's Cognitive Workload Levels Using Artificial Neural Network on ECG Signals, *Applied Ergonomics*, 59, Part A, 326-332, 2017.

Wickens, C.D., Lee, J.D., Liu, Y. and Becker, S.E.G., *An Introduction to Human Factors Engineering*, second ed., PrenticeHall, USA, 2003.

Author listings

Sunggho Kim: dilemma37@naver.com

Highest degree: M.S., Department of Industrial and Management Engineering, POSTECH

Position title: Assistant Professor, Department of Systems Engineering, R.O.K. Air Force Academy

Areas of interest: Human Factors in Aviation and Aerospace, Human Performance & Workload Assessment, User Interface Design & Evaluation, Biosignal Processing

Dongwoo Kang: towarmer@gmail.com

Highest degree: Ph.D. Candidate, School of Industrial Management Engineering, Korea University

Position title: Instructor, Department of Systems Engineering, R.O.K. Air Force Academy

Areas of interest: Mathematical Optimization, Mixed Integer and Nonlinear Programming, Combinatorial Optimization, Machine Learning