

An Efficient Approach for Constructing Intelligent Naval Combat System Dataset based on Combat Structural Data Cleansing Techniques

Hyeon-Mo Kim*, Won-Seok Jang*, Hyoung-Jo Huh**, Ju-Mi Park*, Woo-Hyeon Moon*,
Seo-Ho Lee*, Ji-Seok Yoon*

*Engineer, Naval System 1 Team, Hanwha Systems, Pangyo, Korea

**Engineer, R&D QA Team, Hanwha Systems, Yongin, Korea

[Abstract]

The combat management system(CMS) integrates sensors and weapon systems to effectively perform a series of procedures from detection of targets to engagement. In this paper, we propose a method of improving the generalization performance of the CMS core functions by refining the message-based structural data. The proposed method refines data contamination, including outliers, by employing publicly available cleansing algorithms and cluster-based re-sampling techniques. Two artificial intelligence tasks were defined for the engagement function of CMS, and models trained on contaminated and refined datasets were compared and analyzed. When generating data, data based on sensor failure and mission bias scenario were collected and used as experimental data. The performance differences before and after data refinement across various models were analyzed, and effective combinations of cleansing algorithms were estimated. In addition, the experimental evaluation using the sum of weighting scenarios was conducted to verify the performance improvement after data refinement, and statistical tests were performed in order to confirm its significance. Through comprehensive experiments, it was verified that the pair of LightGBM and AutoEncoder outperformed other approaches in both regression and classification tasks. This study is expected that will be usefully used for developing AI-based intelligent combat management system and reviewing the applicability of new technologies in a maritime domain.

▶ **Key words:** Combat Management System(CMS), Data cleansing, Tabular data, Engagement dataset, Intelligent combat system

-
- First Author: Hyeon-Mo Kim, Corresponding Author: Ji-Seok Yoon
 - *Hyeon-Mo Kim (hyeonmo1227@hanwha.com), Naval System 1 Team, Hanwha Systems
 - *Won-Seok Jang (cws0714g@gmail.com), Naval System 1 Team, Hanwha Systems
 - **Hyoung-Jo Huh (welchs@naver.com), R&D QA Team, Hanwha Systems
 - *Ju-Mi Park (jumipark1126@hanwha.com), Naval System 1 Team, Hanwha Systems
 - *Woo-Hyeon Moon (moonstar@hanwha.com), Naval System 1 Team, Hanwha Systems
 - *Seo-Ho Lee (leesh8426@naver.com), Naval System 1 Team, Hanwha Systems
 - *Ji-Seok Yoon (jsyoon2118@hanwha.com), Naval System 1 Team, Hanwha Systems
 - Received: 2025. 09. 30, Revised: 2025. 10. 27, Accepted: 2025. 11. 24.

[요 약]

함정 전투체계는 센서와 무장체계를 통합하여 표적의 탐지부터 교전까지 절차를 수행하는 시스템이다. 본 연구는 수상함 전투체계의 메시지 기반 정형데이터를 정제하여 전투체계 핵심 기능의 일반화 성능을 향상시키는 방법을 제안한다. 전투체계의 기능 중 교전에 대한 두 가지 인공지능 과업 설정하여 오염 데이터와 정제 데이터 기반으로 모델을 비교 및 분석한다. 데이터 생성시 센서 고장과 임무 편향 시나리오 기반의 데이터를 수집하여 실험 데이터로 활용하였으며, 다양한 인공지능 모델별 데이터 정제 전과 후의 성능 차이, 효과적인 정제 알고리즘의 조합을 추정하였다. 추가로 정제 후의 성능 개선 효과를 확인하기 위하여 가중합 시나리오를 활용한 평가를 수행하였으며, 통계 검정을 통해 유의성을 검증하였다. 광범위한 실험을 통해 LightGBM 과 AutoEncoder의 조합이 회귀와 분류 과업에서 모두 우수한 것을 확인하였다. 본 연구는 인공지능 기반 함정 전투체계 지능화와 신기술 적용성 검토에 유용하게 활용될 것으로 사료된다.

▶ **주제어:** 함정 전투체계, 데이터 정제, 정형데이터, 교전 데이터셋, 지능형 전투체계

I. Introduction

함정 전투체계(CMS, Combat Management System)는 항해/탐색/추적 레이더, 전자광학장비, 피아식별 장비와 같은 센서체계, 전술링크를 포함한 전술통신체계, 미사일, 함포, 근접방어무기, 기만기와 같은 무장체계, 그리고 항해체계로 구성된 무기체계이다. 전투체계 내부의 센서와 무장모듈에서 발생하는 다양한 이벤트 로그와 장비 정보는 DDS(Data Distribution Service) 미들웨어를 통해 전투체계 내에서 수집가능한 메시지 형태로 부유한다. 이에 따라, 임무 수행을 위한 표적 및 교전 정보를 내포하고 있으므로 4차 산업혁명의 여파로 인한 인공지능 적용성 연구가 활발하게 수행되고 있다[1-4]. 또한, 함정 전투체계에 적용하기 위한 연구 또한 활발하게 진행 중이다[5-8].

한편, 전투체계 센서 및 교전 데이터는 실운용 환경에서 다양한 원인에 의해서 오염이 불가피하며, 이는 고품질 인공지능 모델 구축에 영향을 미친다. 장비 고장으로 인한 결측값 발생, 외부 환경에 의한 잡음, 인간 중심의 장비운용 혹은 한정된 환경에서 운용에 의한 라벨 데이터 편향과 같이 다양한 원인을 예로 들 수 있다. 이에 따라, 실제 전장 운용 시 함정 승조원 혹은 인공지능 학습 전문가에게 기술적 상한을 형성하게 된다. 이와 달리, 함정 전투체계의 데이터를 정제하여 인공지능 성능을 개선하기 위한 연구는 미비한 실정이다.

최근 데이터 중심(data-centric)의 접근법을 통하여 모델 설계의 복잡도를 높이기보다 학습 데이터의 품질을 체계적으로 개선하기 위한 노력이 이루어지고 있다[9]. 또한, 주어진 과업의 높은 달성도를 위해서는 모델의 아키텍처

와 파라미터를 수정하는 것보다 일반적으로 높은 성능을 보이는 모델을 베이스라인(baseline)으로 설정한 후 데이터 정제에 신중을 가하는 것이 과업 달성 및 비용 절감에 효과적일 수 있다.

본 연구는 전투체계 도메인에 특화된 데이터 정제 파이프라인을 정립하고, 모델과 검증 조건을 통제된 실험에서 트리 기반의 머신러닝 모델과 신경망 기반 모델의 성능과 안정성을 공정하게 비교 및 분석한다. 또한, 앞선 전투체계 데이터의 오염을 해소하기 위한 효과적인 정제 절차를 설계하고, 결과 트리 계열과 신경망 계열 모델에게 미치는 영향을 광범위한 실험을 통하여 철저히 검토한다. 본 연구에서 제안하는 전투체계 학습용 데이터셋 구축 파이프라인은 다음과 같다:

- (1) 교전 모사 시뮬레이터 기반의 위협평가 및 무장할당 학습용 데이터셋 구축
- (2) 실제 전장환경의 외부영향을 고려하기 위한 데이터 오염 적용
- (3) 이상치 검출 기반의 표본 제거 및 데이터 보충을 통한 정제 데이터셋 구축
- (4) 인공지능 모델 성능평가를 활용한 정제 방법 적용된 정제 데이터셋과 미적용된 비정제 데이터셋 검증

결과적으로 본 연구의 기여는 다음과 같이 요약된다.

- 함정 전투체계 기능 중심의 주요 과업을 설정하고, 관련 데이터의 오염을 해소하기 위한 정제 절차를 제안한다.

이때, 데이터의 오염을 모사하기 위한 전투체계 교전 시뮬레이터를 구축하여 원시 데이터를 수집한다.

- 데이터 정제 기법을 적용하여 데이터셋(이하 정제 데이터셋) 과 그렇지 않은 데이터셋(이하 비정제 데이터셋)을 동일한 검증환경에서 비교 및 분석하여 본 연구의 제안 방식의 정량적 효과를 확인한다.

이후 논문의 구성은 다음과 같다. 2장에서 CMS와 데이터 특성, 트리 및 신경망 알고리즘의 배경을 설명한다. 3장에서 제안 파이프라인과 데이터 중심 검증 방법을 기술하고, 4장에서 실험 설정 및 실험 결과를 기술한다. 5장에서는 결론과 함께 향후 연구를 제시한다.

II. Related works

1. Combat management system

함정 전투체계는 해군 핵심 무기체계로, 공중/해상/수중 등으로부터 복합적인 위협에 대응하기 위해 센서, 무장, 통신, 지휘통제 기능을 통합한 시스템이다. 센서 체계를 통해 표적을 획득 및 분석하여 탑재 무장체계를 통한 교전을 수행한다. Fig. 1은 수상함 전투체계 구성을 보여준다. Fig. 2는 전투체계의 핵심 교전 절차를 나타내며, 최근에는 기능 고도화와 인공지능 적용을 고려한 다양한 연구가 수행되고 있다.

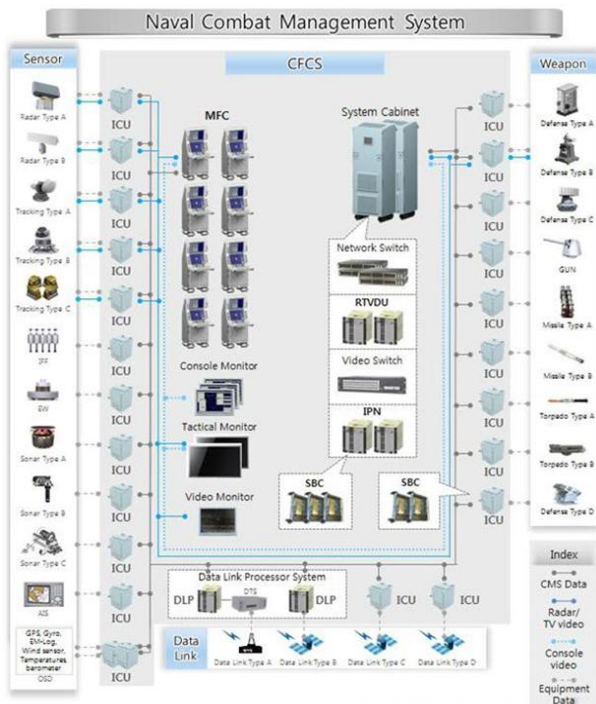


Fig. 1. Configuration of Naval Combat Management System[8]

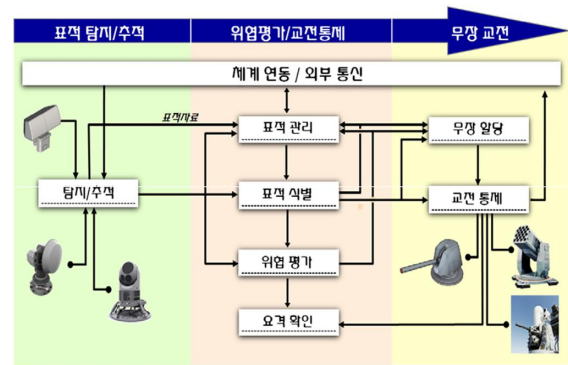


Fig. 2. Engagement Flow of CMS[10]

2. Structural data on CMS

정형데이터는 행과 열로 구성된 테이블 형태의 데이터로, 금융, 의료, 제조 등 다양한 산업 분야에서 광범위하게 활용된다. 한편, 함정 전투체계 내부 네트워크에서 DDS 기반 발행/구독(Pub/Sub)되는 메시지는 대부분 정해진 포맷과 필드 구조를 갖는다. 특히, 센서/표적/무장/교전 등의 정보를 포함한 정형화된 데이터 구조로 인하여 함정 전투체계 메시지 데이터는 정형데이터로 분류될 수 있고, 데이터 포맷을 후처리할 경우 머신러닝 모델에 학습을 위한 데이터셋 형태로 가공 가능하다.

3. Algorithms for processing tabular data

전통적으로 정형데이터 처리에 있어 트리 기반의 앙상블 알고리즘이 높은 성능을 보여왔다. Random Forest[11]는 대표적인 앙상블 알고리즘으로, 여러 개의 의사결정나무를 무작위로 학습시킨 뒤 결과를 투표 또는 평균 방식으로 결합한다. 이 과정에서 데이터 샘플과 특성을 무작위로 선택하여 모델의 다양성을 확보하여 일반화 성능을 높인다. 이후 제안된 XGBoost[12]는 그래디언트 부스팅을 최적화한 구현으로, 정규화 기법, 병렬 학습, 메모리 최적화 등을 통해 대규모 데이터셋에서도 빠르고 정확한 성능을 제공한다. 그러나, 여전히 학습에 필요한 소요시간이 높다는 단점이 있다. LightGBM[13]은 고성능 그래디언트 부스팅 프레임워크로, GOSS(Gradient-based One-Side Sampling)와 EFB(Exclusive Feature Bundling) 전략을 채택하였다. 이로 인해서 대규모 데이터 포인트와 고차원 특성을 처리하는 데 강점을 가진다. 한편, LightGBM은 리프 중심의 성장으로 인해서 불균형한 데이터에서 성능이 저하될 수 있다. CatBoost[14]는 범주형 변수를 자동 처리할 수 있으며, 순서형 부스팅(Ordered Boosting) 기법을 통해 정보 손실을 최소화하고 과적합을 억제한다. 최근에는 TabPFN이 제안되면서 표형식 데이터 분석의 새로운 패러다임이 등장하

었다. TabPFN[15]은 트랜스포머 기반 아키텍처를 활용하여 사전 학습된 네트워크가 데이터셋을 인컨텍스트(In-context) 학습 방식으로 처리한다. 이를 통해 별도의 학습 과정 없이도 회귀 및 분류 문제를 빠른 속도로 해결할 수 있다. 상기는 정형 데이터 분석을 위한 모델의 발전이 전통적인 앙상블 기법에서 출발하여, 최적화된 그래디언트 부스팅으로 이어지고, 최근에는 딥러닝 기반 트랜스포머 모델로 확장되는 과정을 보여주며, 지속적인 연구를 통해 발전되어 오고 있음을 알 수 있다. 본 연구에서는 네 가지의 트리 기반 앙상블 모델과 한 가지 신경망 모델을 활용하여 베이스라인으로 채택한다.

III. The Proposed Scheme

1. Simulator for constructing datasets

본 연구에서는 전투체계 교전 데이터의 정제 효과를 검증하기 위해 자체 개발한 시뮬레이터를 활용하여 데이터를 수집하였다. 해당 시뮬레이터는 자함과 다수 표적 간의 교전 상황을 단순화하여 모사하며, 위협평가 및 무장할당 모델 학습에 필요한 대규모 데이터셋 생성을 목적으로 설계되었다. 본 시뮬레이터는 실제 수상함 전투체계의 운용 논리를 교전 절차에 따라 단순화하였으며, 표적 획득 등의 절차는 생략된 교전 중심의 모델링에 집중한다. 시뮬레이터의 주요 구성은 하위 절에서 설명한다.

1.1 Own ship and target modeling

1) 자함의 모델링: 자함은 6종 무장(대공/대함 미사일, 127mm/76mm/40mm 함포, CIWS)을 보유하며, 각 무장의 사거리, 명중률, 탄약 수, 교전 방식(단발/일제사격), 재장전 시간 등이 반영되었다. 자함 위치는 (0,0)으로 고정되며, 교전 지속성을 위해 체력 요소가 포함된다.

2) 표적의 모델링: 표적은 고속정(FAC), 구축함(DDG), 대함미사일, 전투기(Fighter) 등 4종으로 구성되며, 자함 기준 임의의 방위와 거리에서 랜덤 생성된다. 기동 패턴은 직선 접근 및 회피 선회 등으로 단순화되었고, 속도는 랜덤 범위 내에서 결정되어 등속 운동을 수행한다. 교전 결과 판정을 위해 표적별 체력도 고려되었다.

1.2 Weapon modeling

무장은 자함의 함형에 따라 탑재 가능한 무장(Hard Kill, Soft Kill)이 상이하다. 그러나, 시나리오 단순화를 위하여 획득된 표적을 무력화할 수 있는 하드킬 무장만을 고려하며, 무장은 다음과 같다: 대함 및 대공 미사일, 구경별 함포, 기

관포, 근접방어무기체계(CIWS). 또한, 무장의 속성에는 사거리, 명중률, 기본 탄약 수, 살보당 탄 소모량, 재장전 시간을 포함하여 해상 교전시 고려 가능한 특성을 반영한다.

1.3 Threat-score modeling

위협평가는 표적별로 추론된 위협지수를 활용하여 표적에 적절한 무장을 할당하는 함정 전투체계의 주요 핵심 기능이다. 임의로 생성된 각 표적은 개체로 취급되어 시뮬레이터 내에서 독립적으로 활동한다. 그러므로, 표적별로 서로 다른 위협지수가 산출될 수 있으며, 이것은 아래 수식 (1)에 의해 계산된다. 수식 (1)에서 TTG 는 위협지수, d 는 자함과 표적(T)간 거리, v 는 표적의 속도를 나타낸다.

$$TTG = \frac{d_T}{v_T} \quad (1)$$

한편, 표적별 위협지수는 탐지 표적과 자함의 탑재 무장이 갖는 속성을 고려하여 산출된다. 위협지수는 경험적으로 얻어진 지식에 기반한 임의의 수식을 채택하여 계산한다. 그 후, 수식 (2)을 통해 최대 999값을 갖도록 정규화하여 최종적인 표적별 위협지수를 계산한다. 수식 (2)에서

$$Threat\ Score = Norm(999 \times \frac{score}{score + 10}, 0, 999) \quad (2)$$

$score$ 는 앞선 임의의 수식에 의해 계산된 위협지수를 의미하며, Threat Score는 0과 999 사이로 정규화된 최종 위협지수를 나타낸다.

결과적으로 위협지수가 큰 값일수록 위협도가 높은 표적으로 간주 된다. 또한, 시뮬레이터 상의 획득 표적의 위협지수가 높은 표적부터 정렬하여 최우선으로 무장할당이 수행된다.

1.4 Data labeling using weapon assignment rule

본 연구의 대상 인공지능 기능인 위협평가와 무장할당을 라벨링하기 위해 위협지수를 시뮬레이션간 고려하였다.

1) 위협지수: 각 표적의 교전 우선순위를 결정하기 위해 위협지수를 계산하였다. 위협지수는 거리, Time-to-Go (TTG), 속도, 공격력, 유형별 가중치, 탄약 소모 상황 등을 종합적으로 반영하며, 0~999의 범위로 정규화된다. 값이 클수록 자함에 대한 위협도가 크다고 판단한다.

2) 무장할당: 항공 표적(대함미사일, 전투기)에는 대공 미사일을 우선 고려하며, 사용 불가 시 함포로 대체된다. 무장 선택은 명중률 \times 기대 피해량 \times 위협지수를 기준으로 수행되며, 발사 시 탄약 차감, 쿨다운 적용, 피해 판정이 이루어진다. 라벨링 결과는 6종 무장 할당 외에도 무장 불가 상황(사거리 밖, 탄약 부족, 재장전 중) 3종을 포함한다.

Table 1. Definitions of Input Feature for Threat assessment AI and Weapon assignment AI

Input	Unit	Type(data)	Task(AI)
location_X	km	Numeric (Continuous)	Threat, Weapon
location_Y	km	Numeric (Continuous)	
target_speed	m/s	Numeric (Continuous)	
target_type	-	Categorical	
distance	km	Numeric (Continuous)	
TTG	sec	Numeric (Continuous)	
is_in_range	-	Categorical	Weapon
is_ammo	-	Categorical	
is_reload_unlock	-	Categorical	
is_avail_fire	-	Categorical	
reload_ratio	%	Numeric (Continuous)	
margin_range	km	Numeric (Continuous)	
ammo_ratio	%	Numeric (Continuous)	
remain_range	sec	Numeric (Continuous)	
remain_fire	sec	Numeric (Continuous)	
remain_reload	sec	Numeric (Continuous)	

2. Model tasks in the proposed scheme

앞서 언급되었듯이, 함정 전투체계의 교전 절차에는 탐지된 표적의 위협지수를 정량적으로 산출하고, 위협 우선순위가 높은 순서대로 적절한 무장을 할당한다. 본 논문에서는 위협평가 인공지능 및 무장할당 인공지능이라 명명한다. 또한, 두 인공지능 기능에 과업을 다음과 같이 설정한다. 위협평가 인공지능은 연속적인 실수값을 추론값으로 나타내기 때문에, 회귀 문제로 설정한다. 무장할당 인공지능은 탑재 무장 중 가용 가능한 무장을 위협표적에 할당하므로, 여러 범주 중 하나를 결정하는 분류 문제로 정의한다. Table 1에서는 각 인공지능의 입력데이터를 명세한다. 이때, Table 1의 모든 Input은 본 연구의 6종 무장과 4종 표적에 동일하게 적용되었다.

3. Data contamination

데이터셋 생성을 위해 개발한 시뮬레이터의 경우 이상치 데이터가 발생하지 않는 환경이다. 따라서, 수집된 데이터셋에서 정제 전과 후의 데이터셋을 이용한 비교 분석 실험을 진행하기 위해 데이터를 오염시켰다. 오염시키는 방법은 위협평가 인공지능과 무장할당 인공지능 학습 데이터셋을 오염시키게 되며 입력데이터에 이상치를 발생시키는 방법으로 진행하였다.

Algorithm 1. Tabular data contamination

Parameters	1) sim_data : data to be contaminated 2) feature_cols : features of tabular data 3) policy : a contamination policy 4) contam_ratio : a ratio to the degree of contamination ($\text{contam_ratio} \in \mathbb{R}$)
N = total number of rows in sim_data // 각 행 = 1개 교전 상황 샘플 n_contam = int(N * contam_ratio): if n_contam == 0: return sim_data contam_idx = randomly select n_contam rows for each feature in feature_cols: feature_policy = policy.get(feature, "all") if feature_policy == "class": for row in contam_idx: how = random choice ["none", "nan", "random_class"] case "nan": set feature[row] = missing (NaN) case "random_class": replace with random different class value // "none": 변화 없음 else if feature_policy == "all": compute stats (mean, std, quartiles, IQR, median) stuck_val = random pick of an existing feature value for row in contam_idx: how = random choice ["none", "nan", "outlier", "zero", "stuck"] case "nan": feature[row] = missing case "zero": feature[row] = 0 case "stuck": feature[row] = stuck_val case "outlier": if random < 0.5: feature[row] = Q3 + 3*IQR + noise else: feature[row] = Q1 - 3*IQR - noise case "none": skip return contaminated sim_data	

이상치는 센서에서 발생하기 쉬운 오류인 결측값(NaN, Not a Number), 스파이크 노이즈, 0 값 발생, 이전 타임스탬프의 센서 데이터가 갱신되지 않고 반복되는 스택 노이즈가 고려되었으며, 추가적으로 표적을 식별하는 과정에서 센서나 휴먼 오류로 인한 class 오류가 고려되어 최종적으로 5가지의 오류가 발생할 수 있도록 알고리즘을 설계하였다. 이상치 오염 로직은 Algorithm 1과 같다.

4. Outlier detection algorithm

본 연구는 오염된 데이터를 정제하기 위해서 이상치 탐지 알고리즘을 적용한다. 정제 알고리즘은 다음 총 10가지이다: LOF(Local Outlier Factor)[16,17], Isolation Forest[18,19], COPOD(Copula-Based Outlier Detection)[20], ECOD(Empirical CDF-based Outlier Detection)[21], HBOS(Histogram-Based Outlier Score)[22], KNN(K-Nearest Neighbor)[23,24], LODA(Lightweight On-line Detector of

Anomalies)[25], OCSVM(One-Class Support Vector Machine)[26], AutoEncoder[27], SUOD(Scalable Unsupervised Outlier Detection)[28]. LOF는 국소 밀도 대비 희소 영역의 샘플을 이상치로 판단한다. Isolation Forest는 무작위 분리 트리에서 짧은 분리 경로를 갖는 샘플을 이상치로 본다. COPOD는 copula 기반 주변분포를 추정하여 꼬리 확률로 이상도를 계산하고, ECOD는 경험적 누적분포를 이용해 비모수적으로 이상도를 산출한다. HBOS는 독립 가정 하의 히스토그램 밀도를 사용해 저밀도 구간을 이상치로 본다. KNN은 근접 이웃과의 거리/밀도를 이상도 척도로 변환한다. LODA는 무작위 사영과 희소 히스토그램을 결합하여 경량의 밀도 근사를 수행한다. OCSVM은 정상 영역을 둘러싸는 결정 경계를 학습해 경계 밖을 이상치로 간주한다. AutoEncoder는 encoder-decoder 구조의 딥러닝 모델로서, 입력층과 출력층을 제외하고 4개의 은닉층을 가진다. 이때, 입력층과 출력층의 차원은 동일하며, 이로써 재구성 오차가 큰 샘플을 이상치로 가정할 수 있다. SUOD는 여러 이상치 탐지 알고리즘을 앙상블하여 이상치를 감지한다. 본 연구에서는 각 알고리즘의 주요 파라미터를 다음과 같이 설정하여 데이터의 오염을 이상치로 탐지한다: LOF(n_neighbors=15, 20, 25, 35), Isolation Forest(n_estimators=100, 200), AutoEncoder(hidden_neuron_list=[features], 64, 32, 32, 64, [features])). 상기 파라미터 설정은 PyOD[29] 라이브러리에서 제공하는 기본 설정값을 사용한다. 위 알고리즘을 통해 식별된 이상치는 데이터셋에서 제거되고, 동일한 수만큼의 데이터를 재수집하여 보충된다.

5. Baseline models

앞서 설명하였듯이, 본 연구에서는 정형데이터 학습에서 높은 성능을 보인다고 알려진 대표 모델 4종을 채택하여 활용한다. 모델들은 배깅 기반 앙상블, 그라디언트 부스팅, 그리고 트랜스포머를 기반으로 삼는다. 또한, 다양한 연구에서 베이스라인으로 활용되고 있다. 결과적으로 Table 2는 실험 간에 활용된 5가지 모델들의 주요 하이퍼 파라미터와 특징에 대한 요약을 보여준다.

IV. Experimental results

해당 장에서는 함정 전투체계 교전 데이터 정제 효과 검증 위하여 세 가지 실험들이 수행되었다. 각 실험에서는 최신의 트리 기반 모델과 딥러닝 모델을 포함한 5가지 알

고리즘과 정제 방법 10종이 활용되었다. 실험 종래에는 세 실험의 결과를 종합하여 높은 과업 달성도를 위한 알고리즘과 정제 방법 페어(pair)를 산출한다.

Table 2. Models and Key Parameters

Model	Summary	Key Parameter
Random Forest [11]	<ul style="list-style-type: none"> • Bagging based Decision Tree Ensemble • Robust to non-linearity and interactions • Relatively easy to tune 	<ul style="list-style-type: none"> • n_estimators=100 • max_depth=None • random_state=42
XGBoost [12]	<ul style="list-style-type: none"> • Level-wise growth GBDT (Gradient Boosting Decision Tree) • Strongly regularized • Widely used baseline model 	<ul style="list-style-type: none"> • booster=gdbt • eta=0.3 • max_depth=6 • random_state=42
LightGBM [13]	<ul style="list-style-type: none"> • Leaf-wise growth GBDT(Gradient Boosting Decision Tree) • Robust for massive datasets 	<ul style="list-style-type: none"> • boosting_type=gdbt • num_leaves=31 • learning_rate=0.1 • random_state=42
CatBoost [14]	<ul style="list-style-type: none"> • Oblivious Tree • Ordered Boosting • Robust against overfitting and target leakage 	<ul style="list-style-type: none"> • iterations=1000 • depth=6 • learning_rate=0.03, • l2_leaf_reg=3.0, • random_state=42, • task_type="GPU"
TabPFN [15]	<ul style="list-style-type: none"> • Pretrained transformer for tabular data inference • Efficient construction of baseline with minimal tuning 	<ul style="list-style-type: none"> • n_estimators=4(Regressor) or 8(Classifier) • softmax_temperature=0.9 • device="cuda" • seed=42

Table 3. Label distribution of the classification's dataset

Class	Desc.	Train size	Test size
0	Missile (Maritime)	2223	6667
1	Missile (Air)	2222	6667
2	GUN ^(127mm)	2222	6667
3	GUN ^(76mm)	2223	6667
4	GUN ^(40mm)	2222	6666
5	CIWS	2222	6667
6	사거리 밖	2222	6666
7	잔탄 부족	2222	6667
8	재장전	2222	6666
Total		20,000	60,000

1. Experimental environment

본 절에서는 실험 환경을 상세히 설명한다. 실험의 신뢰성과 객관성을 높이기 위하여 앞서 언급되었던 정형데이터 처리를 위한 인공지능 모델 4종을 활용하였다. 또한, 데이터 오염은 주어진 전체 데이터 수의 0%에서부터 49% 까지 1% 간격으로 변화시켜 적용하였다. 데이터셋 구성의 경우, 오염 데이터 10,000개와 이를 정제한 10,000개의 데이터셋을 구성하고, 학습에 포함되지 않은 60,000개의 독립 데이터셋을 이용하여 평가(test)를 수행하였다. 특히, 무장할당 기능을 위한 데이터셋의 라벨 분포는 Table 3 에서 설명된다. 회귀 모델과 분류 모델의 평가 지표는 각각 평균절대오차(MAE)와 1-F1 점수를 사용하였으며, 이들은 대표적인 회귀 및 분류 모델의 평가 지표로 알려져있다.

실험 간 하드웨어 환경은 다음과 같다: Windows 11(OS), AMD Ryzen 9 7950X@4.50GHz(CPU), 64GB(RAM), 그리고 NVIDIA RTX 4090(GPU). 소프트웨어의 경우는 오픈소스 라이선스 기반의 라이브러리와 파이썬을 사용하였으며 상세 내용은 다음과 같다: Python 3.11.7, PyTorch 2.4.1, Scikit-learn 1.2.2, PyOD 2.0.5, LightGBM 4.5.0, XGBoost 3.0.2, CatBoost 1.2.8, NumPy 1.26.4, Pandas 1.5.3.

2. Model performance test

본 절에서는 오염 수준의 변화에 따른 데이터 정제 전과 후의 모델 성능 변화를 확인한다. 모델별로 모든 정제 방법에 대한 평균 값을 이용하여 비교 분석이 진행되었다. 범례의 ‘Unclean’은 정제하지 않은 비정제 데이터 기반 인공지능의 성능을 나타내며, ‘Clean’은 정제 데이터 기반의 것을 의미한다.

2.1 Result of threat assessment model

Fig. 3은 오염 정도에 따른 위협평가 인공지능의 MAE 결과를 보여준다. 오염 정도가 증가함에 따라 비정제 기반의 알고리즘 성능은 전반적으로 MAE가 상승하는 추세를 보였다. 반면에, 정제 기반의 성능은 모든 오염 구간에서 상대적으로 낮은 수준을 보였다. 특히, LightGBM과 XGBoost 모델의 성능은 오염 정도 3% 시점 이후로 비정제와 정제 데이터 기반의 성능이 명확하게 차이를 보였다. 또한, CatBoost의 경우 오염이 17% 시점 이후로부터 정제와 비정제 데이터 간의 성능 차이가 명확하게 나타났다. 한편, TabPFN은 오염 정도 증가에도 불구하고 두 데이터셋의 성능 차이가 유의미하지 않았으나, 오염에 대해 우수한 성능을 보였다.

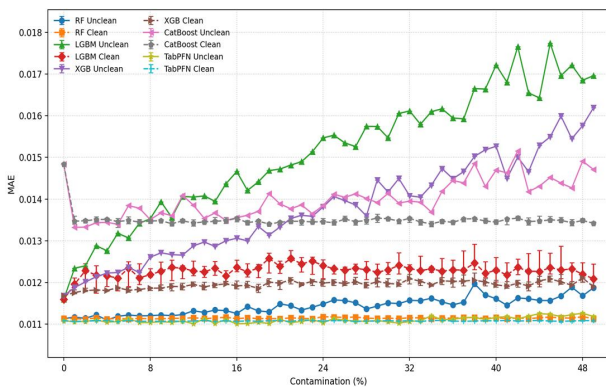


Fig. 3. Performance result of contamination test(MAE)

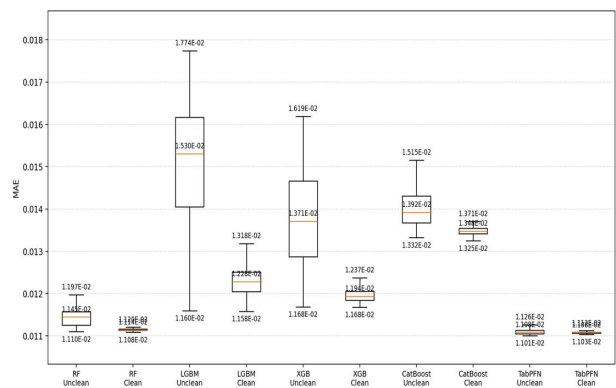


Fig. 5. Box plot-based stability analysis under data contamination(MAE)

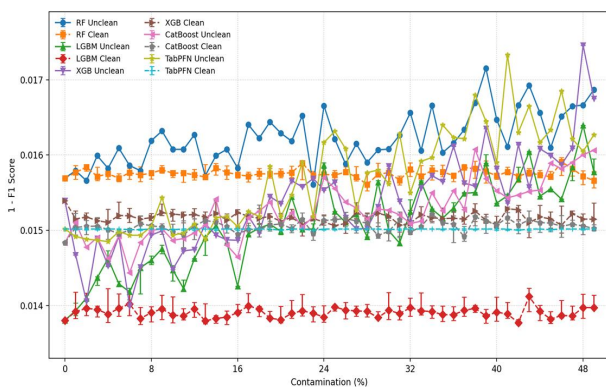


Fig. 4. Performance result of contamination test(F1-score)

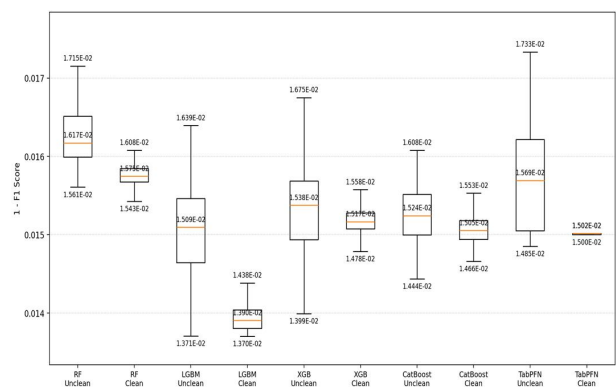


Fig. 6. Box plot-based stability analysis under data contamination(F1-score)

2.2 Result of weapon assignment model

Fig. 4 은 오염 정도에 따른 무장할당 인공지능의 1-F1 score 결과를 보여준다. 오염 정도가 증가할수록 비정제 데이터셋에 대한 성능은 상기 결과 같이 증가하는 경향이 나타났다. 반면에, 정제 데이터셋에 대한 결과는 모든 오염 정도 구간에 대해서 비정제 대비 낮은 수준을 유지하는 안정적인 결과가 확인되었다. 게다가, Random Forest와 XGBoost는 X축 30% 이상의 구간에서 정제와 비정제 간의 높은 성능 격차가 나타났다. 해당 관찰은 높은 오염 정도의 환경에서 정제의 효과성이 높음을 나타낸다. 이와 더불어 CatBoost와 TabPFN은 오염 정도 증가에도 불구하고 정제와 비정제 데이터셋 기반의 성능 차이가 작았으나, 이 결과는 오염 환경에서도 일관성 있는 성능을 유지한다는 것을 의미한다.

2.3 Stability evaluation under data contamination

본 절에서는 2.1과 2.2 실험 결과를 기반으로 오염 정도에 따른 인공지능 성능의 안정성을 박스 플롯(box-and-whisker plot)을 활용하여 분석한다. Fig. 5의 박스 플롯 분석 결과, Random Forest 모델은 MAE 중앙값이 0.011447에서 0.011142로 약 2.7% 개선되었다. LightGBM의 경우, 0.015304에서 0.012281로 약 19.8% 감소하였으며, XGBoost 또한 0.01304에서 0.0117로 약

10.3% 감소하였다. 이들의 IQR가 모두 감소하였는데, 이것은 안정성도 함께 개선되었음을 시사한다. 그러나, TabPFN은 중앙값 개선과 IQR 감소의 변화가 미미하였으나, 이것은 비정제 데이터 즉, 오염된 데이터에서도 정제 데이터만큼의 성능을 확보할 수 있다는 것을 의미한다. 마찬가지로, Fig. 6의 박스 플롯 결과에 따르면 Random Forest, LightGBM, CatBoost, 그리고 XGBoost의 1-F1 점수는 전체적으로 감소하였으며, IQR 감소도 동일하게 확인되었다. 한편, XGBoost의 경우 1-F1 점수의 중간값이 정제 전 0.015375에서 정제 후 0.015167로 1.4% 감소하였고, CatBoost의 경우도 0.015244에서 0.015054로 1.2%의 작은 변화가 실험 결과 확인되었다.

3. Cleansing method test

앞선 장에서 설명된 정제 알고리즘은 오염된 데이터를 정제하기 위해 사용된다. 한편, 각 정제 알고리즘의 개선 폭을 정량적으로 측정하기 위하여, 아래 수식 (3)을 활용한다. s 는 비정제 데이터셋에서의 성능을, s' 는 정제 데이터셋에 대한 성능을 나타낸다.

$$Improvement = \frac{s - s'}{s} \times 100 \quad (3)$$

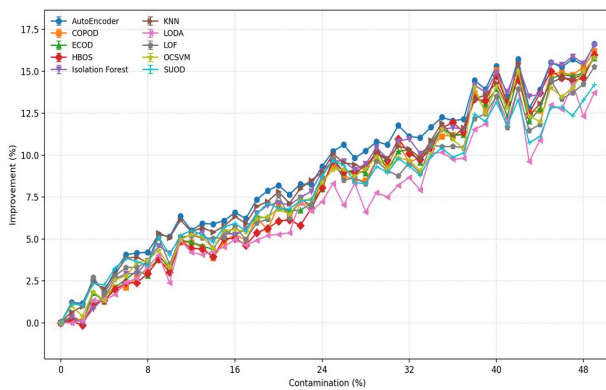


Fig. 7. Performance result of cleansing test(MAE)

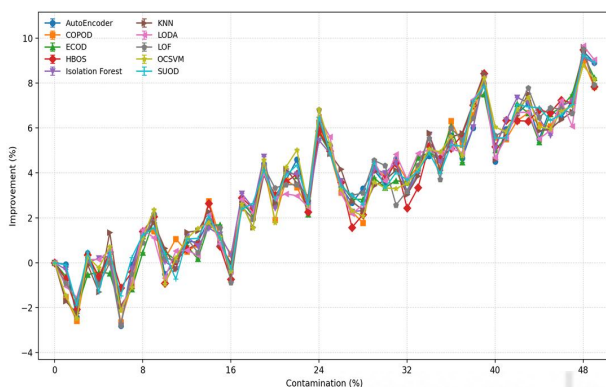


Fig. 8. Performance result of cleansing test(F1-score)

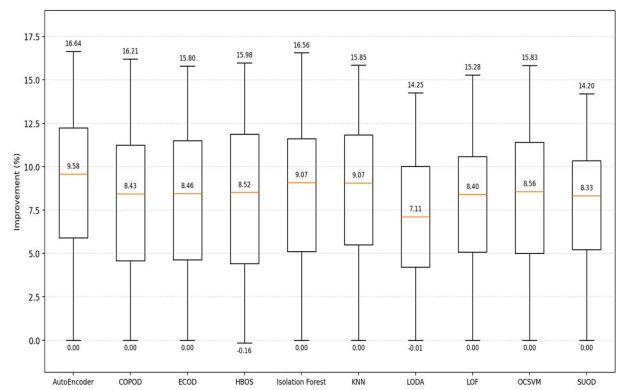


Fig. 9. Box plot-based stability analysis under data cleansing(MAE)

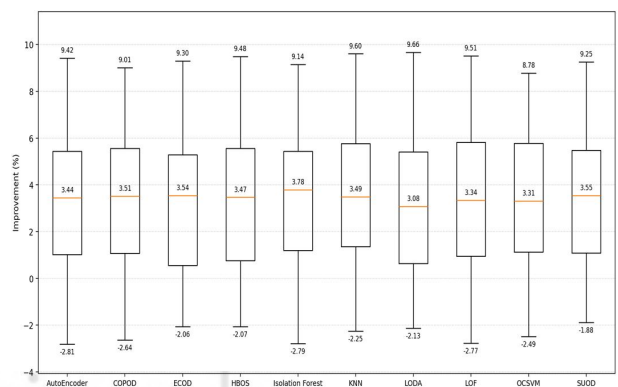


Fig. 10. Box plot-based stability analysis under data cleansing(F1-score)

그러나, 위 수식에 의해서 계산된 개선율은 단일 모델에서의 정제 알고리즘의 개선율을 나타내기 때문에, 본 논문 3장 5절에 언급된 모든 모델에서의 개선율을 측정하기 위한 수식이 요구된다.

$$Improvement = \sum_{m \in M} \frac{s(m) - s'(m)}{s(m)} \times 100 / |M| \quad (4)$$

수식 (4) 는 각 모델의 개선율에 대한 평균을 계산한다. M 은 실험에서 활용된 인공지능 모델을 원소로 하는 집합을 의미한다. 즉, $m \in M$ 이다. 해당 지표는 데이터 정제 실험을 위한 정량적 수치로서 활용되었다.

3.1 Result of threat assessment model

Fig. 7은 10가지 정제 알고리즘의 개선율을 보여준다. AutoEncoder와 KNN은 오염 정도가 증가할수록 개선율이 뚜렷하게 상승하였고, 40% 이상 구간에서는 평균 약 14.6% 이상의 개선을 확인하였다. 또한, Isolation Forest는 전반적으로 높은 개선율을 보인 반면에, 오염 정도가 높은 40% 이상의 구간에서 LODA와 SUOD는 개선 효과가 상대적으로 낮음을 확인했다. 본 실험을 통해 데이터 오염 환경에서 정제 알고리즘들은 분명한 효과를 나타냈지만, 오염에 영향을 주는 불확실한 외부 간섭이 빈번한 전장 환경에서 LODA와 SUOD는 적용을 위해 충분한 고려가 필요하다는 것을 암시한다.

3.2 Result of weapon assignment model

분류 과업에 대한 정제 알고리즘의 개선율 실험 결과가 Fig. 8에 나타난다. 전체적으로 성능 개선율이 높아진 것은 오염 정도가 낮을 때보다 높은 경우에서 오염이 비교적 많이 판별되었음을 알 수 있다. 그런 이유로, 전반적으로 우상향하는 결과를 확인할 수 있다. 그중에 OCSVM 알고리즘은 오염 정도가 10%와 20% 사이 구간에서 때때로 비교적 낮은 개선율을 보인다. 또한, 마지막 시점인 49%에서는 LODA의 개선율이 13.7%로 가장 낮았으나, 다른 알고리즘에과 비교했을 때 차이가 미미했다. 따라서, 본 절에서는 분류 과업에서의 정제 알고리즘별 개선율은 높은 오염이 발생하는 환경에서 뚜렷한 차이가 없음을 알 수 있다.

3.3 Stability evaluation under data cleansing

Fig. 9과 Fig. 10을 통하여 데이터 정제 개선율에 대한 성능 안정성을 박스 플롯 기반으로 분석한다. Fig. 8의 중앙값에 따르면, AutoEncoder 9.581%, Isolation Forest 9.074%, 그리고 KNN 9.068% 순서대로 높았으나, 이 중에서 AutoEncoder의 박스 크기가 가장 작은 것으로 미루

어 보아 값의 분포가 일관되고 안정적이라는 것을 알 수 있다. 또한, Isolation Forest와 LOF 그리고 SUOD는 중앙값이 박스의 위쪽으로 치우쳐져 있으므로, 대부분의 오염 구간에서는 높은 개선율을 보였으나, 일부 경우에서 낮은 개선율이 나타난 것을 의미한다. 결과적으로 박스의 크기와 중앙값의 크기를 조화롭게 고려했을 때, AutoEncoder와 Isolation Forest 그리고 KNN이 회귀 과업에서 안정적인 정제 개선율을 보인다고 할 수 있다. 이와 달리, Fig. 9의 분류 과업에 대한 결과에 의하면 AutoEncoder가 KNN에 비해 박스 크기는 유사하나, 중앙값이 AutoEncoder는 3.44% KNN은 3.49%로 상대적으로 낮음을 확인할 수 있다. 그 외에 전반적인 분류 과업에 대한 정제 알고리즘 전체를 비교하면 중앙값과 개선율의 분포가 전반적으로 안정적임을 확인할 수 있다.

4. Analysis of cleansing methods and models

4.1 Optimal pairs of models and cleansing methods

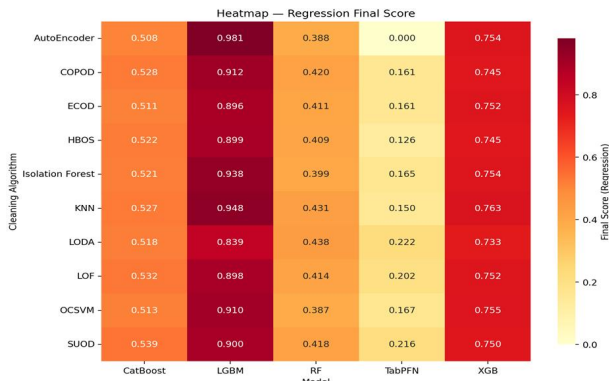
본 절에서는 앞선 실험 결과들을 바탕으로 최적의 정제 알고리즘과 모델의 조합을 찾아내어 오염 환경에 효과적으로 대응하기 위한 방법을 제시한다. 이를 위해, 세 가지(총 개선량, 안정성, 일관성) 기준을 수립하여, 각 정제 알고리즘과 모델들의 성능 실험결과를 종합한다. 아래의 수식들은 최적의 정제 알고리즘과 모델의 조합을 발견하기 위하여 앞서 제시된 세 가지 지표의 산출 방법을 나타낸다.

$$I_{a,m} = \sum_{c \in C} \{s_{a,m}(c) - s'_{a,m}(c)\} \quad (5)$$

$$S_{a,m} = \frac{1}{\sigma(s'_{a,m}(c))} \quad (6)$$

$$K_{a,m} = \frac{1}{|K|} \sum_{c \in C} I\{s'_{a,m}(c) < s_{a,m}(c)\} \quad (7)$$

수식 (5)는 총 개선량(I)에 대한 수식이며, 모든 오염 정도 집합(C)이 주어졌을 때, 특정 정제 알고리즘(a)와 모델(m)에 대한 정제 전과 후의 성능 차이의 총 합을 나타낸다. 즉, 정제 데이터에 대한 성능 개선율을 전 구간에서 누적한다. 수식 (6)의 안정성(S)은 정제 후 데이터셋에 대한 모델 성능의 표준편차를 나타낸다. 이를 통해 표준편차가 낮을수록 성능 측면에서 안정성이 높게 계산된다. 마지막으로 수식 (7)의 일관성(K)은 정제 후 모델의 성능이 정제 전보다 우세한 횟수를 비율로 계산한다. 이때, 해당 수식의 I 는 지시함수를 나타낸다. 결과적으로 각 결과는 0과 1 사이로 정규화되며, 각 지표에 대한 가중치에 따라 아래 수식 (8)과 같이 가중합($Final Score$)이 계산된다.



(a) Cleansing methods and regression models

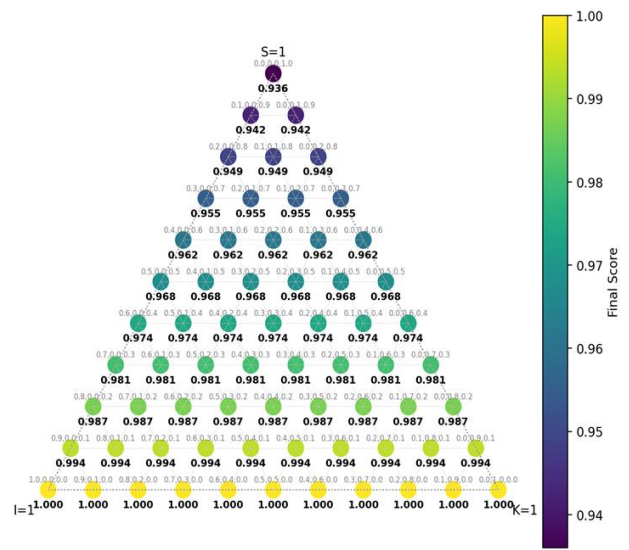


(b) Cleansing methods and classification models

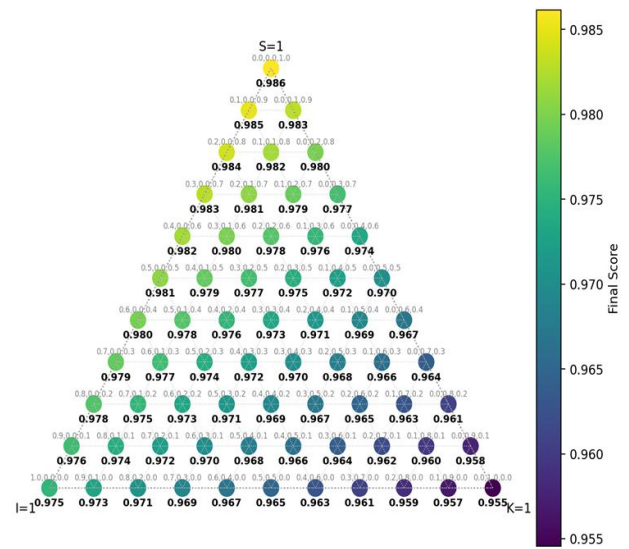
Fig. 11. Optimal pairs of cleansing methods and models by task ($\alpha=0.5, \beta=0.3, \gamma=0.2$)

$$Final\ Score = w_I \sum_{\forall a, \forall m} \tilde{I}_{a,m} + w_S \sum_{\forall a, \forall m} \tilde{S}_{a,m} + w_K \sum_{\forall a, \forall m} \tilde{K}_{a,m} \quad (8)$$

Fig. 11은 총개선풀 0.5, 일관성 0.3, 안정성 0.2 로 가정 하였을 때의 정제 알고리즘과 모델의 조합의 가중합 결과를 보여준다. 본 연구에서는 성능의 총량과 정제로 인한 개선 여부에 관심이 있다는 점을 고려했을 때, 상기 가중치 설정은 연구 목적에 부합한다. Fig. 11의 (a)는 회귀 과업의 최종 점수 히트맵을 보여준다. 또한, Fig. 11의 (b)는 분류 과업의 최종 점수 히트맵을 도시한다. 색이 진할수록 높은 값을 의미하고, 이는 정제 알고리즘과 모델의 조합이 종합적으로 우수하다는 것을 나타낸다. 결과에 따르면, LightGBM 알고리즘은 정제 알고리즘과 조합되었을 때, 산출된 최종 점수가 Random Forest, XGBoost, CatBoost 그리고 TabPFN 대비하여 높았다. 게다가, 위협평가 인공지능과 무장할당 인공지능의 과업에서 전반적으로 AutoEncoder와의 조합이 가장 높은 점수를 획득하였다.



(a) Sensitivity result of the regression task



(b) Sensitivity result of the classification task

Fig. 12. Sensitivity results with different weights on the pair of LightGBM and AutoEncoder

4.2 Sensitivity analysis with weighted metrics

Fig. 12는 가중합을 계산하기 위한 지표 3개, 즉 총 개선풀, 안정성, 그리고 일관성에 대한 가중치를 w_I, w_S, w_K 를 0.1씩 증가하며 변화하는 가중합(Final Score)을 삼원도(ternary plot)을 기반으로 가시화한 결과이다. 각 그림에서 삼각형 하나의 변은 두 가중치 변화를 동시에 표현할 수 있으며, 삼각형 내부는 세 가지 가중치의 조합에 따른 가중합 결과를 색상으로 구분할 수 있다. Fig. 12(a)는 회귀 과업에서 각 지표의 가중치 변화에 따른 가중합의 변화를 나타낸다. 가중합은 좌표 조합에 따라 0.936에서 1.0 범위에서 변화하며, 가중합은 w_K 증가와 w_S 가 감소

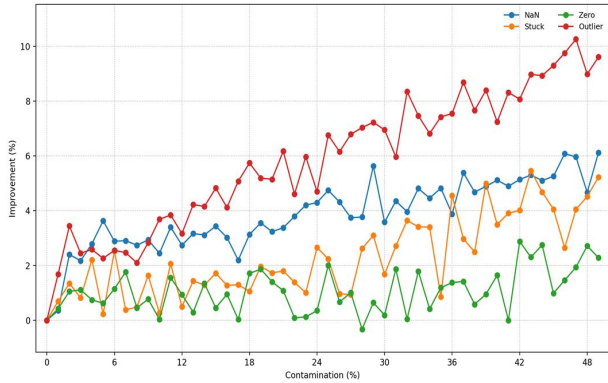


Fig. 13. Performance result of cleansing test(MAE) under single-contamination scenarios

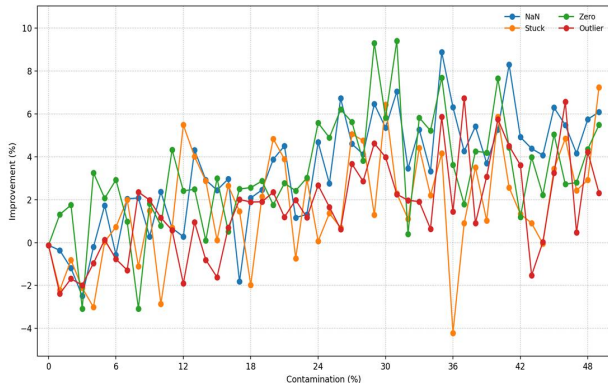


Fig. 14. Performance result of cleansing test(F1-score) under single-contamination scenarios

에 따라 개선되는 경향을 보였다. 반면에 w_I 변화에는 상대적으로 미미한 영향을 받았으며, 이는 가중합의 민감도는 w_K 와 w_S 에 집중되어 있음을 시사한다. Fig. 12(b)는

분류 과업에서의 가중치 변화에 따른 가중합을 보여준다. w_S 가 높은 구간에서는 가중합이 전반적으로 높았다. 특히, w_K 가 증가하고 w_S 가 감소하는 경우 가중합이 감소하는 경향을 확인할 수 있다. 결과적으로 w_I 의 변화보다 다른 두 가중치 변화가 가중합에 상대적으로 많은 영향을 준다는 것을 알 수 있다.

4.3 Sensitivity analysis with different contamination scenarios

본 절에서는 가장 뛰어난 성능을 나타낸 AutoEncoder 기반의 정제 알고리즘과 LightGBM 모델을 활용하여 단일 오염별로 정제 전과 후의 성능 향상을 비교 분석하였다. 분석하는 과정에서 활용된 이상치는 센서에서 발생하는 오류 (NaN, Zero, Stuck, Outlier)이 활용되었다. 먼저, Fig. 13는 회귀 과업에 관한 결과이다. 회귀 과업의 경우 전반적으로 Outlier로 인해 오염된 데이터를 정제할 경우 성능변화가 크고 Zero 값으로 인한 오염에 대한 영향이 가장 작은 것이 확인되었다. 그리고 Fig. 14은 분류 과업에 관한 결과로 NaN으로 인한 오염에 대한 영향이 가장 크고, Stuck으로 인한 오염의 영향이 가장 작은 것으로 확인되었다.

4.4 Statistical test for the pairs of cleansing methods and models

본 절에서는 데이터 정제 알고리즘이 비정제 데이터셋 대비 인공지능 모델의 성능을 통계적으로 유의하게 향상시키는지 검증하기 위한 쌍체 t-검정 결과를 보여준다. 분석 대상은 회귀 과업과 분류 과업으로, 전술된 10개의 정

Table 4. Paired t-test Results using Cleansing Methods and LightGBM ($n = 49$)

Task	Cleansing Method	S_d	SE	t-value	CI	p-value
Regression	AutoEncoder	0.001431	0.000204	16.637090	(0.00299,0.00381)	7.28612E-22
	COPOD	0.001554	0.000222	12.837743	(0.00240,0.00329)	1.94596E-17
	ECOD	0.001460	0.000209	13.468022	(0.00239,0.00322)	3.1831E-18
	HBOS	0.001576	0.000225	12.447815	(0.00235,0.00325)	6.11682E-17
	Isolation Forest	0.001601	0.000229	13.453647	(0.00261,0.00353)	3.31543E-18
	KNN	0.001301	0.000186	16.665655	(0.00272,0.00347)	6.79061E-22
	LODA	0.001251	0.000179	13.282481	(0.00201,0.00273)	5.39588E-18
	LOF	0.001195	0.00017	15.969889	(0.00238,0.00306)	3.87022E-21
	OCSVM	0.001360	0.000194	14.558741	(0.00243,0.00321)	1.55799E-19
	SUOD	0.001049	0.000150	18.160471	(0.00242,0.00302)	1.91296E-23
Classification	AutoEncoder	0.000590	8.437E-05	13.604305	(0.00097,0.00131)	2.16608E-18
	COPOD	0.000579	8.278E-05	13.411811	(0.00094,0.00127)	3.73331E-18
	ECOD	0.000605	8.638E-05	13.295931	(0.00097,0.00132)	5.19259E-18
	HBOS	0.000611	8.730E-05	13.129739	(0.00097,0.00132)	8.35914E-18
	Isolation Forest	0.000542	7.739E-05	14.947065	(0.00100,0.00131)	5.50843E-20
	KNN	0.000626	8.941E-05	12.828924	(0.00096,0.00132)	1.9966E-17
	LODA	0.000600	8.577E-05	13.413435	(0.00097,0.00132)	3.71613E-18
	LOF	0.000622	8.886E-05	12.922803	(0.00096,0.00132)	1.51967E-17
	OCSVM	0.000613	8.755E-05	12.954361	(0.00095,0.00131)	1.38679E-17
	SUOD	0.000568	8.113E-05	14.452836	(0.00100,0.00133)	2.07515E-19

제 알고리즘과 5개의 모델 조합에 대한 성능지표(MAE, 1-F1 score)를 선정하였다. 각 조합별로 오염 정도 1%에서 49%까지 총 49개 구간에 대한 성능 결과를 표본으로 채택하였으며, 이 과정에서 유의미한 분석을 위하여 오염 정도 0%에서의 성능은 표본에서 제외하였다. 귀무가설(H_0)은 정제 전과 후의 성능 차이가 없다고 설정하였고, 이에 대한 대립가설(H_1)은 정제 후에 대한 모델의 성능이 개선되었다고 설정하였다. Table 4는 앞선 실험 결과에서 우수한 성능을 보인 LightGBM에 대하여 유의수준 5%에 대한 통계 검정한 결과를 보여준다. 회귀 및 분류 과업에서 정제 알고리즘 10가지에 대한 LightGBM의 MAE와 1-F1 score에 대하여 오염 정도 총 49개에 대한 계산값들이 나열되어 있으며, 신뢰구간 95%에 대한 범위를 함께 보여준다. 유의수준 5%에 대해서 p-value 값이 0.05보다 작기 때문에, 두 과업에서 모든 정제 알고리즘에 대해 정제된 데이터셋이 정제 전의 데이터셋보다 성능개선이 유의미하다는 것을 알 수 있다. 또한, 신뢰구간 95%에 대한 최소 및 최대 범위를 고려했을 때, 성능 개선이 유의미하다는 것을 확인할 수 있다.

V. Conclusions

본 연구에서는 불확실한 데이터 오염을 고려한 실전장 환경의 함정 전투체계 정형데이터를 정제하여 강건한 교전(위협평가, 무장할당) 인공지능의 성능 개선을 위한 데이터셋 구축 방안을 제안하였다. 게다가, 제안된 방법은 실 전장환경의 불확실한 외부요인을 고려하기 위해서 정형데이터의 오염을 모사하는 시뮬레이터를 포함하였다. 또한, 정밀하게 설계된 환경에서 광범위한 실험이 수행되었으며, MAE, F1-score, 정제 개선율이 고려되었다. 더불어, 본 논문에서 자체적으로 설계 및 채택한 최종 점수를 활용하여 최적의 인공지능-정제방법 조합을 분석하였다. 이를 통해, 과업과 무관하게 오염 정도가 높을수록 인공지능의 성능은 점진적으로 낮아진 반면에 반대의 경우 성능은 우상향 함을 확인하였다. 특히, 결과 종합 시 LightGBM과 AutoEncoder 기반의 정제 방법이 위협평가 및 무장할당 인공지능 과업에서 전반적으로 높은 효과를 보였으며 Outlier 값과 NaN으로 인한 데이터 오염이 모델의 성능에 영향을 많이 끼친다는 것이 확인되었다.

본 연구는 함정 전투체계의 지능화를 위한 최초의 실험적 논문이다. 그러나, 실제 해상에서 운용된 함정 내의 장비, 전술 등의 실제 데이터를 활용하지 않아 후속 연구를

통하여 실질적인 실험 결과를 확보할 필요가 있다. 구체적으로, 논문에서는 다양한 오염을 고려하기 위해서 로직을 활용하였으나, 실제 오염의 패턴이나 분포를 고려할 필요가 있다. 또한, 함정 플랫폼 별로 탑재된 무장에 따라 전투체계 데이터의 구조가 상이하기 때문에, 인공지능 학습용 데이터셋 구축 시 고려할 필요가 있다. 게다가, 동역학(dynamics) 요소를 고려한 확장 연구가 필요하다. 향후 연구로서 실 체계 기반의 실험데이터 혹은 훈련 체계 기반의 가상데이터 활용한 확장 연구와 멀티모달을 위한 EO/IR 영상 및 이미지 데이터 기반의 확장 연구를 진행할 계획이며, Group K-Fold 등 그룹 인지형 교차검증을 통해 결과의 불확실성과 일반화 가능성을 정량화해 신뢰성 높은 연구를 진행할 예정이다. 이를 통해 지능형 함정 전투체계에 대한 적합성과 실질적인 적용 가능성 검증이 기대된다.

ACKNOWLEDGEMENT

This work was supported by Korea Research Institute for defense Technology planning and advancement (KRIT) - Grant funded by Defense Acquisition Program Administration (DAPA) (KRIT-CT-23-030)

REFERENCES

- [1] S. Lee, "Research on the application of Machine Learning to threat assessment of combat systems," *Journal of The Korea Society of Computer and Information*, Vol. 28, No. 7, pp. 47-55, July 2023. DOI: 10.9708/jksoci.2023.28.07.047
- [2] Y. Shin, J. Roh, S. Ahn, H. Jang, and H. Jeon, "Probabilistic AI-Based Prediction of Missile Target Selection, Launch Intent and Post-Engagement Behavior," *Journal of The Korea Society of Computer and Information*, Vol. 30, no.5, pp. 59-68, May 2025. DOI: 10.9708/jksoci.2025.30.05.059
- [3] M. Lee, J. Oh, C. Kim, J. Bae, Y. Kim, and C. Jee, "The Development of Rule-based AI Engagement Model for Air-to-Air Combat Simulation," Vol. 25, No. 6, pp. 637-647, December 2022. DOI: 10.9766/KIMST.2022.25.6.637
- [4] J. Kim and J. Yang, "A Study on Mission Engineering-based Development Methodology for AI Unmanned Systems: Focusing on Explainability and Controllability," Vol. 8, No. 3, pp. 345-353, September 2025. DOI: 10.31818/JKNST.2025.9.8.3.345

- [5] P. Kwon, K. Jang, S. Kim, J. Kim, W. Yun and K. Rhee, "A Study on the Implementation Method of Artificial Intelligence Shipboard Combat System," *Journal of convergence security*, Vol. 20, No. 2, pp. 123-136, June 2020. DOI: 10.33778/kcsa.2020.20.2.123
- [6] G. Do, J. Park, W. Jang, Y. Yang, and J. Yoon, "A Study on Deep Learning-based Automatic Target Recognition System in IR Image for Intelligent Combat Management System," *Journal of the Korea Society of Computer and Information*, Vol. 30, No. 1, pp. 33-40, 2025. DOI: 10.9708/jksci.2025.30.01.033
- [7] C. Hong and H. Baek, "A Study of the Concept of AI-based Unmanned System Maintenance for the Efficient Operation of Naval Manned-unmanned Teaming," *Journal of the KNST*, Vol. 8, No. 3, pp. 592-602, 2025. DOI: 10.31818/JKNST.2025.9.8.3.592
- [8] H. Na, "Research on AI-based anomaly detection for ship combat system weapon system interlocking control," *Journal of The Korea Society of Computer and Information*, Vol. 30, No. 1, pp. 23-32, January 2025. DOI: 10.9708/jksci.2025.30.01.023
- [9] D. Zha, Z. P. Bhat, K.-H. Lai, F. Yang, Z. Jiang, S. Zhong, and X. Hu, "Data-centric Artificial Intelligence: A Survey," *arXiv preprint arXiv:2303.10158*, March 2024. DOI: 10.48550/arXiv.2303.10158
- [10] Y. Jeong, C. Kim, W. Han, J. Kim, and H. Kim, "A Study on the Design and Verification-Validation of the Supportive Equipment for Shipyard Test of Naval Combat System," *Journal of the Korea Institute of Military Science and Technology*, Vol. 17, No. 3, pp. 318-326, June 2014. DOI: 10.9766/KIMST.2014.17.3.318
- [11] L. Breiman, "Random Forests," *Machine Learning*, Vol. 45, pp. 5-32, October 2001. DOI: 10.1023/A:1010933404324
- [12] T. Chen and C. Guestrin, "XGBoost: A Scalable Tree Boosting System," In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 785-794, San Francisco, United States, August 2016. DOI: 10.1145/2939672.2939785
- [13] G. Ke, Q. Meng, T. Finley, T. Wang, W. Chen, W. Ma, Q. Ye, and T.-Y. Liu, "LightGBM: a highly efficient gradient boosting decision tree," In *Proceedings of the 31st International Conference on Neural Information Processing Systems (NIPS'17)*, pp. 3149-3157, Long Beach, United States, December 2017. DOI: 10.5555/3294996.3295074
- [14] L. Prokhorenkova, G. Gusev, A. Vorobev, A. V. Dorogush, and A. Gulin, "CatBoost: unbiased boosting with categorical features," In *Proceedings of the 32nd International Conference on Neural Information Processing Systems (NIPS'18)*. pp. 6639-6649, Curran Associates Inc., Red Hook, USA, December 2018. DOI: 10.5555/3327757.3327770
- [15] N. Hollmann, S. Müller, L. Purucker, A. Krishnakumar, M. Körfer, S. B. Hoo, R. T. Schirmeister, and F. Hutter, "Accurate predictions on small data with a tabular foundation model," *Nature*, Vol. 637, pp. 319-326, January 2025. DOI: 10.1038/s41586-024-08328-6
- [16] M. M. Breunig, H.-P. Kriegel, R. T. Ng, and J. Sander, "LOF: identifying density-based local outliers," In *Proceedings of the 2000 ACM SIGMOD international conference on Management of data (SIGMOD '00)*, 273-278, Association for Computing Machinery, New York, USA, May 2000. DOI: <https://doi.org/10.1145/342009.335388>
- [17] M. Kim, J. Sunghwan, J. Kim, and S. Kim, "A Study on Fault Detection Method Using Weight-based Local Outlier Factor," Vol. 31, No. 4, pp. 273-278, 2021. DOI: 10.5391/JKIIS.2021.31.4.273
- [18] F. T. Liu, K. M. Ting, and Z. -H. Zhou, "Isolation Forest," In *Proceedings of the 2008 Eighth IEEE International Conference on Data Mining*, 413-422, Pisa, Italy, December 2008. DOI: 10.1109/ICDM.2008.17
- [19] Y. Ko, S. Kim, H. Lee, J. Park, and H. Park, "Anomaly Trajectory Detection Using Clustering and Isolation Forest," *Journal of Korean Institute of Intelligent Systems*, Vol. 34, No.1, pp. 25-35, February 2024. DOI: 10.5391/JKIIS.2024.34.1.25
- [20] Z. Li, Y. Zhao, N. Botta, C. Ionescu, and X. Hu, "COPOD: Copula-Based Outlier Detection," In *Proceedings of the 2020 IEEE International Conference on Data Mining (ICDM)*, 1118-1123, Sorrento, Italy, November 2020. DOI: 10.1109/ICDM50108.2020.00135
- [21] Z. Li, Y. Zhao, X. Hu, N. Botta, C. Ionescu, and G. H. Chen, "ECOD: Unsupervised Outlier Detection Using Empirical Cumulative Distribution Functions," *IEEE Transactions on Knowledge and Data Engineering*, Vol. 35, No. 12, pp. 12181-12193, December 2023. DOI: 10.1109/TKDE.2022.3159580
- [22] M. Goldstein and A. Dengel, "Histogram-based Outlier Score (HBOS): A fast Unsupervised Anomaly Detection Algorithm," In *Proceedings of the 35th German Conference on Artificial Intelligence*, 59-63, Saarbrücken, Germany, September 2012. DOI: 10.48550/arXiv.1207.4005
- [23] A. Mucherino, P. J. Papajorgji, P. M. Pardalos, "k-Nearest Neighbor Classification," In *Proceedings of the Data Mining in Agriculture*, 83-106, Springer, New York, NY, September 2009. DOI: 10.1007/978-0-387-88615-2_4
- [24] W. Han, S. Lee, G. Kim, and K. Park, "Lightweight Classification Model for Underwater Mines and Rocks Using Sonar Signals in Military Environment," *Journal of The Korea Society of Computer and Information*, Vol. 30, No. 5, pp. 131-140, May 2025. DOI: 10.9708/jksci.2025.30.05.131
- [25] T. Pevný, "Loda: Lightweight on-line detector of anomalies," *Machine Learning*, Vol. 102, pp. 275-304, 2016. DOI: 10.1007/s10994-015-5521-0
- [26] L. M. Manevitz and M. Yousef, "One-class svms for document classification," *The Journal of Machine Learning Research*, Vol.

- 2, pp. 139-154, March 2002. DOI: 10.5555/944790.944808
- [27] S. Yoon and N. Kim, "Document Classification Methodology Using Autoencoder-based Keywords Embedding," *Journal of The Korea Society of Computer and Information*, Vol. 28, No. 9, pp. 35-46. September 2023. DOI: 10.9708/jksci.2023.28.09.035
- [28] Y. Zhao, X. Hu, C. Cheng, C. Wang, C. Wan, W. Wang, J. Yang, H. Bai, Z. Li, C. Xiao, Y. Wang, Z. Qiao, J. Sun, and L. Akoglu, "SUOD: Accelerating Large-Scale Unsupervised Heterogeneous Outlier Detection," In *Proceedings of Machine Learning and Systems 3*, 463-478, San Jose, CA, USA, 2021. DOI: 10.48550/arXiv.2002.03222
- [29] S. Chen, Z. Qian, W. Siu, X. Hu, J. Li, S. Li, Y. Qin, T. Yang, Z. Xiao, W. Ye, Y. Zhang, Y. Dong, and Y. Zhao, "PyOD 2: A Python Library for Outlier Detection with LLM-powered Model Selection," *arXiv preprint arXiv:2412.12154*, December 2024, DOI: 10.48550/arXiv.2412.12154

Authors



Hyeon-Mo Kim received the B.S. and M.S. degrees in Computer Science and Engineering from Sejong University, South Korea, in 2022 and 2024. He is currently with Hanwha Systems Co. from 2024.

He is currently with Hanwha Systems Co. from 2024. His research interests include Intelligent Naval Combat Management System, Data Mining, Ontology Engineering, Large Language Model, and Artificial Intelligence.



Won-Seok Jang received M.S degree in Computer Engineering from Chungnam National University, South Korea and completed Ph.D. program from Kyungpook National University, Korea.

He is currently working in Hanwha Systems Co. from 2017. He is interested in Intelligent Naval Combat Management System using Artificial Intelligence, Naval Unmanned System and Naval Safety System.



Hyoung-Jo Huh received the B.S., M.S. Ph.D. degrees in Electronic Communication Engineering from Hanyang University, Korea, in 1996 and 1998, respectively and completed Ph.D. program in Management from Hanyang University, Korea, in 2025.

He is currently working in Hanwha Systems Co. from 2010. He is interested in Intelligent Naval Combat Management System Digital Quality Management using Artificial Intelligence.



Ju-Mi Park received the B.S degree in Division of Media, Culture and Design Technology from Hanyang University ERICA, South Korea and the M.S. degree in School of Integrated Technology from Gwangju

Institute of Science and Technology, Korea. She is currently working in Hanwha Systems Co. from 2023. She is interested in Intelligent Naval Combat System using Artificial Intelligence.



Woo-Hyeon Moon received the B.S. degrees in Electronic Engineering from Jeonbuk National University and M.S. degrees in Artificial Intelligence(Reinforcement Learning, Natural Language Processing)-based Robotics

from Korea Advanced Institute of Science and Technology, South Korea, in 2022 and 2024. He is currently with Hanwha Systems Co. from 2024. His research interests include Intelligent Combat Management System, Reinforcement Learning, Multi-Agent Reinforcement Learning, Natural Language Processing, Large Language Model, and Artificial Intelligence.



Seo-Ho Lee received the B.S. degrees in Mechanical Engineering and Unmanned Vehicle Engineering from Sejong University, Korea, in 2022, and the M.S. degrees in Intelligent Mechatronics from Sejong

University, Korea, in 2024. He is currently with Hanwha Systems Co. from 2024. His research interests include Intelligent Naval Combat Management System, Control Engineering, and Reinforcement Learning.



Ji-Seok Yoon received the B.S degree in Electronic Engineering from Kumoh National Institute of Technology, South Korea, in 2012, and the M.S. and Ph.D. degrees in mechatronics from Gwangju Institute of

Science and Technology (GIST), Gwangju, South Korea, in 2014 and 2021. From 2021 to 2023, he was a Postdoctoral Researcher with the Vision and Image Processing Laboratory, Tech University of Korea. Since 2023, he has been the CTO of IKLAB Inc. Currently, he is working in Hanwha Systems Co. from 2023. He is interested in Intelligent Naval Combat Management System using Artificial Intelligence.