AI 기반 지능형 CCTV 이상행위 탐지 성능 개선 방안

류 동 주*, 김 승 회**

요 약

최근 생성형 Artificial Intelligence(이하 AI)와 인공지능에 대한 수요가 높아짐에 따라, 오남용에 대한 심각성이 대두되고 있다. 그러나, 이상행위 탐지를 극대화한 지능형 CCTV는 군과 경찰에서 범죄 예방에 큰 도움이 되고 있다. AI는 인간이 가르쳐준 대로 학습을 수행한 후, 자가 학습을 진행한다. AI는 학습된 결과에 따라 판단을 하기 때문에, 학습 시 특징을 명확하게 이해해야만 한다. 그러나, 인간이 판단하기에도 모호한 이상한 행위와 비정상 행위의 시각적 판단이 어려운 경우가 많다. 이것을 인공지능의 눈으로 학습하기란 매우 어렵고, 학습을 한 결과는 오탐, 미탐 그리고 과탐이 매우 많아진다. 이에 대해 본 논문에서는 AI의 이상한 행위와 비정상 행위의 학습을 명확하게 하기 위한 기준과 방법을 제시하고, 지능형 CCTV의 오탐, 미탐 그리고 과탐에 대한 판단 능력을 최대화 하기 위한 학습 방안을 제시하였다. 본 논문을 통해, 현재 활용 중인 지능형 CCTV의 인공지능 엔진 성능을 극대화가 가능하고, 오탐율과 미탐율의 최소화가 가능할 것으로 기대된다.

AI-Based Intelligent CCTV Detection Performance Improvement

Dongju Ryu*, Kim Seung Hee**

ABSTRACT

Recently, as the demand for Generative Artificial Intelligence (AI) and artificial intelligence has increased, the seriousne ss of misuse and abuse has emerged. However, intelligent CCTV, which maximizes detection of abnormal behavior, is of great help to prevent crime in the military and police. AI performs learning as taught by humans and then proceeds with self-learning. Since AI makes judgments according to the learned results, it is necessary to clearly understand the charact eristics of learning. However, it is often difficult to visually judge strange and abnormal behaviors that are ambiguous even for humans to judge. It is very difficult to learn this with the eyes of artificial intelligence, and the result of learning is very many False Positive, False Negative, and True Negative. In response, this paper presented standards and methods for clarifying the learning of AI's strange and abnormal behaviors, and presented learning measures to maximize the judgment ability of intelligent CCTV's False Positive, False Negative, and True Negative. Through this paper, it is expected that the artificial intelligence engine performance of intelligent CCTV currently in use can be maximized, and the ratio of False Positive and False Negative can be minimized.

Key words: Artificial Intelligence(AI), Anormaly Behavior Detection, Data for Training, Machine Learning, Intelligent CCTV, Quality Inspection

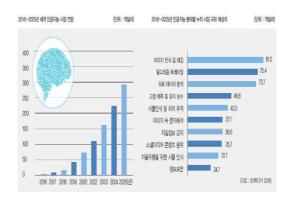
^{**} 극동대학교/인공지능보안학과 박사과정 (공동저자)

1. 서 론

최근 인공지능이 탑재된 지능형 CCTV가 범죄예 방에 상당한 성과를 이루고 있다. 그리고 생성형 AI의 돌풍이 사회를 다시 한번 뒤흔들고 있다. 하 지만, 현재까지의 인공지능과 생성형 AI 역시 학 습결과의 치명적인 약점과 개선점을 가지고 있는 데 그 부분이 바로 학습 데이터 부분이다. 인공지 능은 사람의 시각을 완벽하게 재현하기는 오랜 기 간이 걸리는데 이를 극복하기 위한 방법이 주기적 인 학습과 추가적인 특징점을 재학습하는 방법이 다. 이는 학습 시 편향성을 극복하고 인간의 판단 에 의한 라벨링 등의 오류 수정을 통해 인공지능 이 정확한 학습을 수행하도록 하는데 매우 중요한 영역이다[1,2,3,4]. 본 논문은 지능형 CCTV와 AI 의 시각적 판단 능력을 확보하기 위한 장비에서의 검출 목적성에 맞게 해당 이상 행위와 비정상 행 위의 학습 기준을 제시하고자 한다. 이는 사람이 느끼는 현장의 분위기와 감성, 시각적 요소등과 결합하여 이를 학습에 반영하고, 특히, 기계적인 영상에서의 문제점을 개선하기 위한 학습 방법을 제시함으로써 왜곡된 학습 결과를 바로잡아 판단 오류율을 최소화하기 위한 방안을 제시하였다. 본 논문에서는 이상행위와 비정상 행위를 사전에 제 시하고 본론에서는 오탐과 미탐 문제점이 발생될 가능성이 높은 다양한 현상과 상황을 해결하기 위 한 방안을 제공하여 Ai-Hub를 통해 학습 데이터 로 활용 가능하도록 하였다[5]. 마지막으로 결론을 통해 학습 시 휴먼 피처링을 통한 특장점 추출의 중요성을 강조하였다.

2. 연구 동향

최근의 지능형 CCTV의 현황을 살펴보면, 그림 1과 같다[6]. 그림 1을 살펴보면, 사회 전반에 걸 친 기술 발전과 영상 정보의 활용이 증가하고 있 음을 보여준다. 특히, 한국인터넷진흥원(이하 KIS A)에서의 지능형 CCTV 인증 기준이 지속적으로 업데이트 되면서 단순한 R&D의 형태에서 실생활 에 적극 활용 가능한 범위로 발전하고 있다[7].



(그림 1) 지능형 CCTV 시장 동향

2023년 10월 현재 약 140개 업체가 지능형 CC TV인증을 획득한 상태이다[7]. 해당 인증 기준이 되는 이상행위에 대해서 간략하게 도표로 정리하 면 표 1과 같다[7]. 한국인터넷진흥원이 보유하고 있는 영상 데이터베이스(DB)의 물리적인 구성은 10분 이내의 H.264로 인코딩된 비디오파일로 해 상도는 1280×730이상이며, 민간분야의 영상 DB 는 맑은날, 눈과 비, 안개 등 다양한 날씨 상황에 서 촬영하고, 시간은 분야에 관계없이 일출과 9시, 19시, 일몰, 야간 등 6가지로, 카메라는 근거리와 중거리, 원거리 등 총 3대로 이벤트를 발생시켜 촬영을 원칙으로 구성하여 인증을 수행한다. 이에 대한 기준은 한국지능정보사회진흥원에서 품질과 구축 가이드라인을 제작하면서 제시한 내용을 반 영한 것으로 판단된다.

<표 1> KISA 지능형 CCTV 성능시험 인증 분야

분야	항목	계
일반	배회, 침입, 유기, 싸움, 방화, 쓰러짐, 마케팅, 익수자 탐색, 실종자 수색, 화재 감시	10
안전	교통사고, 화재, 범죄, 생활안전, 자살, 감염병	6

그러나, 침입은 특정 건물이나 지역에서의 행위 가 보다 명확하게 정의를 할수 있지만, 배회의 경 우 단순하게 돌아다니거나 특정 지역을 자주 출몰 한다고 해서 배회가 될수는 없다. 이에 대한 좀더 명확한 행위 영상을 확보하여 학습이 되어야만 단 편 반복 동작의 기법으로 무조건 탐지, 인지되는 것이 아니기 때문이다. 또한, 화재의 경우 다양한 유형의 데이터를 보유한 소방청에서 제공되는 데 이터를 이용하여 단순 시나리오 기반의 학습 과정 이 아니라. 실질적인 발화 과정과 연기 등 다양한 환경적 조건이 같이 구성이 되어야 한다. 민간의 경우, 직접 촬영 시 각종 법적 규정을 준수해야만 하고 시나리오 기반이거나, 직접 촬영에서 수집된 데이터는 개인정보보호 및 영상정보보호법에 저촉 되지 않는 한도내에서 프레임별로 잘라서 사용해 야만 한다. 이를 통해, 실질적으로 업체들과 소통 후 개선 방안을 제시하여 촬영을 진행하였다. 그 결과, 실질적인 AI 학습 결과가 개선됨을 확인하 였다. 개선된 학습 데이터 결과는 한국지능정보사 회진흥원의 Ai-Hub에서 개방 중이다[5].

3. 본 론

지능형 CCTV에서 확보해야 할 학습 고려사항 은 표 1과 같다. 관련 판단 기준은 2020년부터 20 22년 NIA의 Ai-Hub 데이터와 학습 데이터 구축 업체와의 소통을 통해 검증한 내용이다[1,2,3,4]. 표 2에서 제시하고 있는 것은 실제 수행기관들과 협의하면서 실질적인 데이터의 모델 검증 결과를 토대로 수정 보완한 고려사항을 제시한 것이다. 업체들의 인공지능 데이터 사업을 진행하면서 발 생하는 많은 문제점들을 검토하고 분석한 결과, 수집 데이터의 문제가 대부분이며 라벨링에서의 정확도와 재현율이 심각한 성능을 좌우 한다는 것 을 발견하였다. 이에 대해서 지능형 CCTV를 이 용한 학습 시 고려해야 할 사항들을 간략하게나마 제시하고 이를 근거로 데이터를 수집하면 성능이 향상됨을 확인할 수 있다.

<표 2> 지능형 CCTV 데이터 학습 고려사항

목적성

- 이상행위(비정상)의 명확하고 객 관적인 정의 필요
- 정의된 행위에 대한 데이터의 특 징 반복 추출
- 이상행위 검출 환경과의 목적성,

일치성 여부 확인

- 사전 정의된 1개 행위에 대한 다 양한 시나리오 확보 필요
- 이상 행위와 관련된 반복적 학습 최소 시간 확보(개당 최소 30초에서 1분 이상 영상 필요)
- 이상행동을 명확히 확인 가능한 영상 프레임 시점에서 약 30초 이상 추적 대상 객체가 존재하는 경우만 선정 진행
- 흔들림, 초점불일치, 폐색 외 렌즈 이물질로 인한 오류 데이터 배제
- 재난, 화재, 사회 안전, 추적 등 목 적성에 맞는 휴먼 피처링을 반영하 여 학습 필요

편향성

- 편향성 배제 필요
- 특정 반복 구간과 중복 사람에 대 한 학습 배제
- 등장 인물의 객체 다변화
- 계절, 날씨, 장소, 시간대 등을 샘 플화 하여 선정
- 우천 시 빛 번짐과 카메라 빗방울 맺음 현상 등으로 인한 오류 제거 필요
- 아침과 저녁의 붉은색으로 인한 변색 반영(사람이 색을 표기)
- 어둡거나 너무 밝은 장소의 경우 빛 감도에 따른 색의 변화와 음영 고려 필요

환경적 요소

- 그림자는 학습 불가
- 조도에 따른 색의 변화(색감의 오 판) 학습 필요

예)검은색이 시간과 조도에 따라 짙은 남색이나 회색으로 변질됨(일 명 폐색 현상)(빨간색->검은색, 흰 색->회색)

- 해상도 : 1080(FHD)이상 4k급의 해상도를 가지고 학습 권고
- CCTV의 화각과 구도, 방향 등을

확인 후 화각별 레이블링 필요		
Re-ID (재식별)	■ Re-ID(재식별)을 위한 연동 학습 반영 ■ 시계열 화면 연계 학습 ■ ID 부여를 통한 환경적, 편향성 배제, 추적 영상 학습 필요 - 연속성에 따른 학습 프레임(예 30 초) 등 판단 가능한 영상 기준 필 요 - 얼굴 인식 불가 시 걸음걸이, 행 동등의 동일한 시그니처 관련 휴 먼 피처링 동반 학습 필요	
자연 재해 (산불)	■ 산불 감시에서 구름과 굴뚝 연기, 실제 화재에 대한 구별 후 학습 필 요(실제 측정 위치에 따라 구름과 안개 구별 불가) ■ 화재 발화 시점에 대한 지속적인 현실 데이터 적용 필요(발화시작과 과정을 원,중,근거리별로 시간대별 로) ■ ToF 외 열화상 등 특수 카메라의 경우 야간 식별 반영을 통한 RGB 와 쌍으로 1:1 학습 필요	
안전 분야	● 안전 분야에서 범죄 행위의 구별 과 식별, 이에 대한 특징 추출을 통한 반복적 학습 필요 ● 안전 분야 항목의 행위에 대해 보다 명확한 라벨링 특징 추출이 절대적으로 필요(전문가 반영) ● 작업 시설 및 정상적 장비 착용후 사전 학습 필요 ● 정상적 행위를 선행 학습 후 이상행위 변경 학습 필요	
기타	■ 전체 데이터 중 약 1~5% 이내로 노이즈 데이터(흐림, 객체 판별 불 가, 색 판정불가, 비, 눈, 시각적 노 이즈, 빛번짐)등 데이터 학습을 통 해 오버피팅(과적합)을 회피 가능함.	

산불 감시에서 사용되는 학습 데이터를 확인해 서 모델을 실행해 보면, 그림2와 같이 탐지가 두 개를 하게 된다. 이유는 구름과 연기의 구별이 용 이하지 않는 각도에서의 수집 데이터이기 때문이 다. 그림2에서와 같이 햇빛을 등진 상태에서 구름 이 전봇대와 겹치는 현상으로 음영값이 전봇대의 뒷 구름을 연기로 표현한 것이다. 이는 학습 시 이에 대한 반복적 상황을 제시해야만 하거나, 이 와 같은 역광 이미지를 일부 약 5%만 학습을 시 키고 나머지 95%는 역광이 아닌 각도의 사진을 학습 시켜야만 한다. 즉. 편향성을 고려한 다양한 각도의 사진을 학습해야만 정상적인 역광의 모습 까지도 인공지능이 학습이 가능하다는 결과를 말 해 준다.



(그림 2) 산불 학습 데이터 탐지 오류 결과

표2에서 제시한 바와 같이, 학습 시 수집 데이 터의 고려사항을 최대한 반영하여 수집을 해도 탐 지율이 70% 이상이 나올 수 없는 조건임을 확인 하였다. 그러므로 지능형 CCTV의 탐지율을 향상 시키기 위해서는 편향성과 조도, 역광 등을 고려 한 수집 설계가 매우 중요하다. 또한, 학습 결과를 확인 후, 라벨링 즉, 휴먼 피처링을 통한 특장점 추출을 좀더 명확하게 프레임별로 제시함으로써 탐지율과 학습 효율을 극대화할 수 있음을 확인하 였다.

현재까지 파악된 지능형 CCTV의 학습 관점에 서의 오류는 여러 가지 장애 요소를 가지고 있다. 그중, 객체의 포즈 변화에 대한 문제, 조명 변화 (낮/밤)에 따른 영상의 조도 변화문제, 가림(Occlu sion)에 의한 불확실성 문제, 카메라 해상도에 따 른 객체 선명도, 다양한 종류의 카메라(RGB, 적

외선, ToF 등)의 연동 문제, 일관되지 않는 카메 라 관점(View)에 대한 문제 등 장애 요소를 해결 하기 위해서는 각각의 항목에 대해 정확한 기준과 제약사항을 통한 휴먼 피처링을 반복하여 학습 시 켜야 한다. 특히, 생성형 AI와의 결합을 통한 인 공지능 알고리즘의 발전으로 입력 영상에서 특징 을 추출하는데 많은 초점을 맞추고 있다. 생성형 AI의 경우 오류율이 많은 이유는 너무나 많은 데 이터를 무작위로 휴먼 피처링 없이 즉, 걸러내는 작업을 제대로 수행하지 않고 학습을 수행하기 때 문에 발생되는 오류들이다. 또 다른 이유는 최근 정보의 시간차에 의해 발생되는 시점별 학습의 시 점왜곡 현상으로 발생되는 정보의 시간차 문제점 이 존재한다. 지속적 학습을 수행하는 모델을 장 착해야만 되는 것이다. 최근에 지능형 CCTV에서 사람이 직접 색인하여 추적할 때의 문제점을 해결 하기 위해 사용되는 인공지능 역시 학습에서의 오 류로 인한 문제점을 가지고 있다. 그러므로 이러 한 문제점을 해결하고 인공지능이 사람의 눈으로 확인하는 객체 추적 기술에서의 Re-ID(재식별) 기술 핵심요소는 특징 추출, 거리 측정에 따른 지 속적인 객체의 시각적 관점과 시계열의 연속성과 일관성 확보가 매우 중요하다. 또한, 정상적인 행 위와 비정상적인 이상행위를 구분하고 탐지하기 위해서 무조건 정상 외엔 이상행위라고 판단하여 학습시키는 방식은 학습된 결과를 확인하는 관제 자들의 업무 가중과 알람의 최대치에 따른 문제로 이어지므로 좀더 명확한 인간의 식별 능력을 유사 하게 판단하게끔 학습 시키는 것이 매우 중요함을 확인하였다. 개인정보와 관련된 법적 기준과 영상 데이터의 종류도 중요하다. 이외에도 각종 행위와 영상별 데이터 라벨링 기준들이 필요하다.

그림3은 실제 CCTV에서 사용자를 추적하기 위한 데이터를 학습하여 판단하도록 한 결과인데, 검출된 결과는 3인 유형 모두 검은색이 아니고 추 적 대상자가 아님에도 불구하고 그림 3과 같은 탐 지 결과를 출력하였다. 가장 좌측의 객체는 짙은 회색 계열로 빛반사에 의한 번집과 반팔티셔츠와 팔을 그림자로 인해 판단하지 못하는 상황이다., 그림 3의 가운데 객체는 짙은 남색으로 이 역시

빛의 산란 현상(빛의 퍼짐)으로 인해 해상도와 빛 의 각도에 의한 색의 왜곡 현상이 발생하여 오류 로 식별한 것이다. 마지막 세 번째 우측 객체는 가방을 메고 있지만 가방을 식별하지 못한 것이 고, 긴팔 티셔츠를 입었지만 빛의 역광으로 인한 가방과 옷의 색 판단이 불가능하였고, 긴팔과 반 팔을 빛의 그림자로 구분하지 못하고 주변의 조명 과 빛의 조도 영향으로 인한 객체 식별 불가한 상 황으로 탐지 결과를 확인할 수 있다.







(그림 3) 지능형 CCTV 데이터 동일인 식별 추적 오탐 결과

원본 사진에서의 조건은 동일 인물인지를 확인 하고 추적하기 위한 데이터인데, 3인이 모두 동일 인물이 아님에도 불구하고 동일 인물로 검출된 것 이고, 좌측 1개의 객체는 가방을 메고 있지 않음 에도 가방이 있다고 판단한 결과이다. 특히, 가방 의 유무도 추적 시 조건중 하나 였음에도 불구하 고, 가방의 유무와 관계없이 판단한 결과이다. 이 에 대한 원인은 검은색을 학습할 때, 조도와 각도, 음영등이 상당한 영향을 미침을 알수 있다. 또한, 가방이 검은색일 경우 몸체와 붙거나 조도가 매우 낮아서 확인이 안되는 경우에 이러한 결과를 발생 함을 알수 있었다. 따라서, 학습 시 휴먼 피처링을 통해 외곽을 만들어 주거나 음영을 변경해서 학습 해야만 한다. 다른 해결 방법 하나는 사람이나 자 동차 객체를 라벨링 할 경우, 대부분 바운딩 박스 를 사용하는 경우가 많지만 이러한 색 번짐과 색 이 객체와 일원화된 형태일 경우에는 폴리곤이나 세그멘테이션 기법의 외곽 라인을 이용한 라벨링 방법을 사용하면 검출과 식별률이 높아 짐을 알수 있었다. 그 외에 다양한 각도로 해당 대상자를 학 습하고, 편향성과 색감에 대한 빛의 변화를 학습

하는 것도 매우 중요하다. 그리고 가방을 별도로 학습을 수행하게 하는 방법 역시 매우 효과적인 학습 결과가 도출됨을 확인하였다. Re-ID 식별의 경우 수행기관별로 탐지 성능이 매우 저조하므로 탐지 결과에 사용하는 알고리즘을 TOP 5, 혹은 TOP 10처럼 5개나 10개를 탐지한것으로 출력하 면 성능이 우선인것처럼 표현한다. 그러나 본 논 문에서 제시한 특장점 수정 보완 및 정제 방법을 이용한 결과, 어노테이션 값을 제대로 입력하여 꾸준한 학습을 수행하면 5가지 선택형 답이 아니 라 1개의 객체에 대해 70% 이상의 확률을 나타냄 을 확인하였다. 정확도, 정밀도, 재현율을 기준으 로 표기하는 F1 Score나 AP(Average Precisio n), mAP(mean Average Precision)를 사용하는 객체 검출과 정확도에서 높은 탐지율을 확인하였 다. 이에 대한 최종 결과는 제3자 검증으로 TTA 를 통해 확인되었다. 본 논문에서 제시하는 방법 을 사용한 결과, 수행 기관별로 수용 범위가 매우 다양하기 때문에 예시와 결과를 수치로 표현하지 못하는 점이 존재하지만 탐지 성능과 학습 성능이 목표치 평균 70% 대비 최소 10~20% 이상 목표치 에서 향상됨을 확인하였다.

본 논문에서 제시한 기준들을 이용하여 인공지 능 모델의 탐지 성능 개선을 위해서는 많은 시간 과 노력이 필요하다. 첫째, 시간적 요소인 날씨와 계절 변화에 따른 영상 결과물의 변화 둘째, 디지 털 장비에 출력되는 영상물들의 특징을 명확히 인 지한 후 시각적 왜곡 현상을 극복하기 위한 대안 을 반영해야만 한다. 마지막으로 목적에 맞는 행 위 분석 정의와 시나리오등을 통한 부족한 데이터 의 확보 방안등이 절대적으로 필요하다. 특히, 부 족한 데이터를 위해 시나리오 중심의 다양한 영상 과 행위에 대한 빅데이터를 구현하여 탐지 성능과 시험을 반복 수행할 필요가 있다. 데이터 확보와 탐지 성능 개선을 위해서는 지방자치단체의 CCT V 데이터와 국방 관련 데이터, 경찰청, 소방청 등 정부 기관에서 제공하는 데이터를 사전 정제 작업 과 지속적인 행위 분석에 따른 특장점 분류와 분 석의 세밀함이 더해져야만 좀더 명확한 인공지능 판단 능력이 개선됨을 알 수 있었다.

4. 결 론

본 논문은 AI 기반 지능형 CCTV를 이용한 이상 행위와 비정상 행위를 탐지할 때. 오류율을 최소화 가 능한 학습 방법을 제시하였고, 이를 통해 향후 지능형 CCTV의 행위 분석 시 오탐율과 미탐율을 개선가능 하도록 학습 방안을 제시하였다. 실제 인공지능 학습 에서의 최적화를 위한 학습 데이터 개선은 모델 학습 에 지대한 영향을 미친다는 사실을 확인하였다. 생성 형 AI 데이터 결과를 확인해 보면, 휴먼 피처링의 프 로세스가 존재함에도 오류가 발생하는 것은 제대로 이 오류 검증 프로세스를 활용하지 못한 것이고, 학습 용 데이터의 명확한 특징점 추출 학습은 모델의 올바 른 학습 결과를 반영한다는 사실이 입증되었다. 따라 서 본 연구는 향후, 지능형 CCTV 자동 탐지 성능에 서 학습 시 수집 데이터의 오류를 최소화하고, 휴먼 피처링을 통한 특장점 추출 학습이 반영된다면, 오탐 율 극복화 방안에 활용 가능할 수 있을것으로 기대된 다.

참고문헌

- [1] 김승희, 류동주, "비정형데이터의 AI학습을 위한 영상/이미지 데이터 품질 향상 방법", 융합보안 논문지, 제23권 2호, 2023.
- [2] 과학기술정보통신부, 한국지능정보사회진흥원, 한국정보통신기술협회, 인공지능 학습용 데이터 품질관리 가이드라인 v3.0, 2023.
- [3] 과학기술정보통신부, 한국지능정보사회진흥원, 한국정보통신기술협회, 인공지능 학습용데이터 품질관리 안내서 v1.0 제1권 품질관리구축 안 내서, 2021.
- [4] 한국지능정보사회진흥원(NIA), 인공지능 학습용 데이터 구축 사업관리 매뉴얼, 2021.
- [5] 과학기술정보통신부, 한국지능정보사회진흥원 Ai -Hub (www.aihub.or.kr).
- [6] "2022년 국내 영상보안 시장 및 이슈 결산", 보안 뉴스, 2022.12.
- [7] 한국인터넷진흥원(www.ksecurity.or.kr) 지능형 CCTV 인증 현황.

------- [저 자 소 개] -------



류 동 주 (Ryu-Dong Ju) 2009년 2월 전남대학교 정보보안협동 과정 박사 2020년 2월 극동대학교 산업보안학과 조교수 역임 현 재 극동대학교 해킹보안학과 겸임 교수 2017년 ~ 현재 비트레스(주) 대표 email: ryu@btress.com



김 승 회 (Kim-Seung Hee) 2021년 2월 극동대학교 산업기술보안 학과 석사 2021년 9월 ~ 현재 극동대학교대학원 인공지능보안학과 박사과정 email: zelssen@btress.com