

2023 (사)ICT플랫폼학회 추계학술대회 자료집

행사일시 2023년 12월 7일(목) 14시

행사장소 동국대 충무로영상센터 7층

주최/주관 (사)ICT플랫폼학회
동국대학교 인공지능협업센터, 강남대 AI.SW융합연구소

후원 (주)LG CNS, (주)SK브로드밴드, 대신정보통신(주), 세림티에스지(주),
(주)스페이스앤빈, (주)시아인사이트, 쌍용정보통신(주), 아이씨티웨이(주),
(주)올포랜드, (주)이액티브, (주)인라스, (주)진인프라, (주)한국IT컨설팅

10주년 기념

(사)ICT플랫폼학회 2023 추계학술대회 논문집

사단법인 아이씨티플랫폼학회
서울 서초구 서초중앙로 63 (서초동) 15
ictps.staff@gmail.com

2023.07.27

PTL Volume 10-1 | ISSN 2288-8195

Editor-in-Chief

Dae-Sik Ko

MOKWON University,
Daejeon, Republic of Korea,
kds@mokwon.ac.kr

Managing Editor

Bongen Gu

KOREA NATIONAL UNIVERSITY OF TRANSPORTATION,
Chungju-si, Chungcheongbuk-do, Republic of Korea
bggoo@ut.ac.kr

세부 프로그램(안)

2023년 12월 7일(목) 13:30 ~ 18:00

시간	프로그램	
13:30~	등록	
13:30~15:10	연구윤리교육 : 더욱 강화되는 연구윤리	박동희 (한국교통대 교수)
	미래형자동차 핵심기술 R&D 전문 인력 양성센터 워크숍 진행 : 김시호 (연세대 교수)	
	1. 주변 차량과 상호 작용을 고려한 경로 예측 네트워크 이예린, 김시호 (연세대)	2. 악천후 날씨 환경 종류 인식 네트워크 주형진, 송한빈, 김시호 (연세대)
	논문 발표 세션 진행 : 김 현 (부천대 교수)	
	1. 차원축소와 클러스터링을 동시에 적용한 데이터 분석 방법에 대한 연구 신병철, 최승빈, 김보성, 배성민, 송성민, 장혁수, 한동근, 박세진 (계명대)	7. Teachable Machine을 이용한 가죽 결함 검출 방법 연구 권소영, 이종윤, 이용환 (금오공대)
	2. AI 기반 V2X 통신 구간 대상 이상 탐지모델 제안 이진규, 지일환, 전규현, 서정택 (가천대)	8. 커넥티드 자율주행차의 합류구간 속도제어를 위한 지능형 운전자 모델 파라미터의 강화학습 기반 최적화 신성재, 최경환 (광주과학기술원)
	3. 인공지능 할루시네이션에 대응하는 질의개선 이현정, 홍세영, 박찬솔, 김영철 (홍익대)	9. 도메인 일반화 영상 분할을 위한 클래스 기반 대조 학습 양근영, 임묘택 (고려대)
	4. 두 번째 로짓을 활용한 머신러닝 김효서, 최준석 (서강대)	10. 리소스가 제한된 환경에서 CNN 모델에 대한 주파수 관점의 분석 윤지애, 김의환 (광주과학기술원)
	5. 다수 자율주행차량의 이동 경로를 고려한 작업 할당 기술 이세빈, 남창주 (서강대)	11. 반도체 칩 기능 시험의 오류 디버깅 시 Single Long Scan Chain 활용 방안 박종수 (목원대)
	6. 사이버 킬 체인 기반 악성코드 공격 사례 분석 전규현, 전승호, 서정택 (가천대)	12. 지도 정보를 융합한 멀티뷰 카메라 객체인식 모델 개발 장민철, 문석하, 김진규 (고려대)
논문발표 및 우수논문 선정평가		
2023년 정기총회		
15:10~15:20	휴식	
15:20~18:00	기술세미나 진행	
18:00~18:10	대한민국ICT플랫폼대상 및 우수논문상 시상식	서정택 (가천대 교수)
18:10~18:20	경품 추첨 및 공지 ★ 사전등록 후 당일 현장참석자에 한해 추첨을 통한 경품 지급 ★	
18:20~	저녁 만찬 및 정보교류	

•
2023
(사)ICT플랫폼학회
추계학술대회
•

◎ 연구윤리교육 1
박동희 교수(한국교통대학교)

◎ 미래형자동차 핵심기술 R&D전문인력양성센터 워크숍 7
김시호 교수(연세대학교)

* * *

- 1. 주변 차량과 상호 작용을 고려한 경로 예측 네트워크 9
이예린, 김시호 (연세대학교)
- 2. 악천후 날씨 환경 종류 인식 네트워크 12
주형진, 송한빈, 김시호(연세대학교)

◎ 논문 발표 15
김 현 교수(부천대학교)

* * *

- 1. 차원축소와 클러스터링을 동시에 적용한 데이터 분석 방법에 대한 연구 17
신병철, 최승빈, 김보성, 배성민, 송성민, 장혁수, 한동근, 박세진 (계명대학교)
- 2. AI 기반 V2X 통신 구간 대상 이상 탐지모델 제안 25
이진규, 지일환, 전규현, 서정택 (가천대학교)
- 3. 인공지능 할루시네이션에 대응하는 질의개선 31
이현정, 홍세영, 박찬솔, 김영철 (홍익대학교)
- 4. 두 번째 로짓을 활용한 머신러닝 36
김효서, 최준석 (서강대학교)
- 5. 다수 자율주행차량의 이동 경로를 고려한 작업 할당 기술 40
이세빈, 남창주 (서강대학교)
- 6. 사이버 킬 체인 기반 악성코드 공격 사례 분석 43
전규현, 전승호, 서정택 (가천대학교)

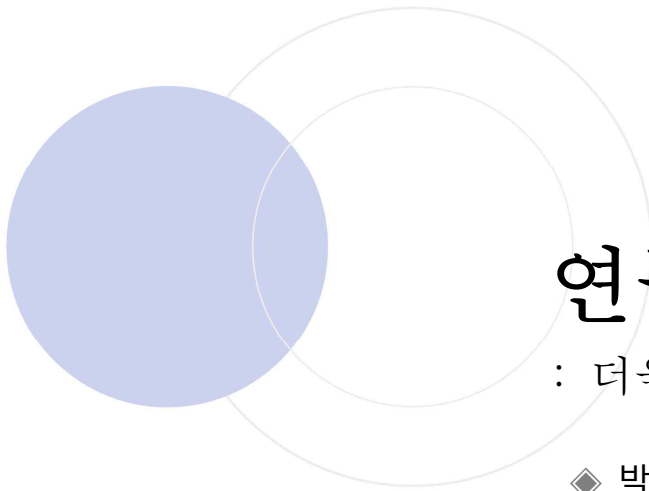
•

2023

(사)ICT플랫폼학회 추계학술대회

•

7. Teachable Machine을 이용한 가족 결함 검출 방법 연구	49
권소영, 이종윤, 이용환 (금오공과대학교)	
8. 커넥티드 자율주행차의 합류구간 속도제어를 위한 지능형 운전자 모델 파라미터의 강화학습 기반 최적화	52
신성재, 최경환 (광주과학기술원)	
9. 도메인 일반화 영상 분할을 위한 클래스 기반 대조 학습	56
양근영, 임묘택 (고려대학교)	
10. 리소스가 제한된 환경에서 CNN 모델에 대한 주파수 관점의 분석	60
윤지애, 김의환 (광주과학기술원)	
11. 반도체 칩 기능 시험의 오류 디버깅 시 Single Long Scan Chain 활용 방안	63
박종수 (목원대학교)	
12. 지도 정보를 융합한 멀티뷰 카메라 객체인식 모델 개발	67
장민철, 문석하, 김진규 (고려대학교)	



연구윤리교육

: 더욱 강화되는 연구윤리

◆ 박동희 (한국교통대학교 교수)

제1조(목적)

본 규정은 ICT플랫폼학회(이하 "학회"라 함) 회원들의 연구윤리를 확보하는데 필요한 의무와 책임에 관하여 기본적인 원칙과 방향을 제시함을 목적으로 한다.

제2조(적용대상)

본 규정은 학회의 모든 회원 학회 주관의 학술활동에 참여하는 모든 연구 관련자에 대하여 적용한다. 단, 학술활동이라 함은 학회주관의 학술행사, 출판사업, 교육사업 등을 포함한 연구와 관련한 제반활동을 의미한다.

제3조(적용범위)

특정 연구 분야의 윤리 및 진실성에 관하여 다른 법적 규정이 있는 경우를 제외하고는 이 규정에 의한다.

제4조(연구부정행위의 범위)

본 규정에서 제시하는 연구부정행위라 함은 논문지, 학술지, 학회와 연관된 연구개발의 수행 등에서 행하여진 위조, 변조, 표절, 중복게재, 부당한 저자 표시행위 등을 말하며 다음의 각 호와 같다.

- ① "위조"라 함은 존재하지 않는 데이터 또는 연구내용, 결과 등을 정당한 승인 또는 인용없이 행하는 학문적 행위를 말한다.
- ② "변조"라 함은 연구에 필요한 재료와 장비, 그리고 연구과정 및 결과 등을 인위적으로 조작하거나 데이터를 임의로 변형, 삭제함으로써 연구 내용 혹은 결과를 왜곡하는 행위를 말한다.
- ③ "표절"이라 함은 타인의 아이디어, 연구내용, 결과 등을 정당한 승인 또는 인용 없이 도용하는 행위를 말한다.
- ④ "중복게재"라 함은 연구자 자신의 이전 연구결과와 동일 또는 실질적으로 유사한 학술적 저작물을 처음 게재한 학술지 편집자나 저작물 저작권자의 허락 없이 또는 적절한 출처표시 없이 다른 학술지나 저작물에 상용하는 학문적 게재 행위를 말한다.
- ⑤ "부당한 논문저자 표시"라 함은 연구 내용 또는 결과에 대하여 논문, 저술, 보고서 등에 과학적이거나 기술적인 기여를 한 자에게 저자 자격을 부여하지 않거나, 이러한 기여를 하지 않은 자에게 저자 자격을 부여하는 행위를 말한다.
- ⑥ 본인 또는 타인의 부정행위의 의혹에 대한 조사를 고의로 방해하거나 제보자에게 위해를 가하는 행위를 말한다.
- ⑦ 과학기술계에서 통상적으로 용인되는 범위를 심각하게 벗어난 행위 등을 말한다.

제5조(제보자의 권리 보호)

- ① 제보자는 부정행위를 인지한 사실 또는 관련 증거를 학회에 알린 자를 말한다.
- ② 제보자는 구술, 서면, 전화, 전자우편 등 가능한 모든 방법으로 제보할 수 있으며 실명으로 제보함을 원칙으로 한다. 단, 익명의 제보라 하더라도 서면 또는 전자우편으로 연구과제명 또는 논문명 등 구체적인 부정행위의 내용과 증거를 포함하여 제보한 경우에는 이를 실명 제보에 준하여 처리한다.
- ③ 학회는 제보자의 신원에 관한 사항이 공개 않도록 제보자의 신상정보를 보호하여야 한다. 제보자가 부정행위 신고를 이유로 징계 등 신분상 불이익, 근무조건상의 차별, 부당한 압력 또는 위해 등을 받지 않도록 보호해야 할 의무를 지니며 이에 필요한 시책을 마련하여야 한다.
- ④ 제보자가 부정행위의 신고 이후에 진행되는 조사 절차 및 일정 등에 대하여 알고자 할 경우에 학회는 이에 성실히 응하여야 한다.
- ⑤ 제보 내용이 허위인 줄 알았거나 알 수 있었음에도 불구하고 이를 신고한 제보자는 보호 대상에 포함되지 않는다.

제6조(피조사자의 권리 보호)

- ① 피조사자는 제보 또는 학회의 인지에 의하여 부정행위의 조사 대상이 된 자 또는 조사 수행 과정에서 부정행위에 가담한 것으로 추정되어 조사의 대상이 된 자를 말하며, 조사과정에서의 참고인이나 증인은 이에 포함되지 아니한다.
- ② 학회는 검증과정에서 피조사자의 명예나 권리가 부당하게 침해되지 않도록 주의하여야 한다.
- ③ 부정행위에 대한 의혹은 판정 결과가 확정되기 전까지 외부에 공개되어서는 아니 된다.
- ④ 피조사자는 부정행위 조사 처리절차 및 처리일정 등에 대해 학회에 알려줄 것을 요구할 수 있으며, 학회는 이에 성실히 응하여야 한다.

제7조(진실성 검증 시효)

제보의 접수일로부터 만 5년 이전의 행위에 대해서는 접수를 하였더라도 처리하지 않는다. 단 5년 이전의 부정행위라 하더라도 피조사자가 그 결과를 직접 재인용하여 후속연구가 진행될 경우에는 예외조항으로 한다.

제8조(진실성 검증 원칙)

- ① 부정행위의 사실 여부를 입증할 책임은 학회와 연구윤리위원회에 있다. 단, 피조사자가 연구윤리위원회에서 요구하는 자료를 고의로 훼손하였거나 제출을 거부하는 경우에는 요구자료에 포함되어 있다고 인정되는 내용의

사실여부를 입증할 책임은 피조사자에게 있다.

- ② 연구윤리위원회는 제보자와 피조사자에게 의견진술, 이의제기 및 변론의 권리와 기회를 동등하게 보장하여야 하며 관련 절차를 사전에 알려주어야 한다.
- ③ 학회 회장단과 이사진은 연구윤리위원회가 부당한 압력이나 간섭을 받지 않고 독립성과 공정성을 유지할 수 있도록 노력하여야 한다.

제9조(진실성 검증절차)

- ① 연구부정행위의 조사는 접수일로부터 10일 이내에 수행되어야 한다.
- ② 모든 조사 일정은 접수일로부터 6개월 이내에 종료되어야 한다.
- ③ 조사 착수 후 1년 이내에 이에 대한 판정 및 조치가 이루어져야 한다.
- ④ 연구부정행위라는 판정결과를 받은 자는 이로부터 1개월 이내에 재심을 요구하여야 하며, 의의 신청을 받은 날로부터 1개월 이내에 윤리위원회는 재심결과를 도출하여 해당 재심 신청자에게 통보하여야 한다.

제10조(윤리위원회의 구성 및 권한)

- ① 본 학회의 논문에 대한 부정행위 문제가 제기되는 경우, 윤리위원회를 5인 이내로 위원장을 포함하여 2주 이내로 구성하며, 위원장은 학회장이 지명한다.
- ② 위원이 피조사자가 되었을 때는 관련자를 제외하고 구성한다.
- ③ 연구윤리위원회는 조사과정에서 제보자, 피조사자, 증인 및 참고인에 대하여 진술을 위한 출석을 요구할 수 있으며 이 경우 피조사자는 반드시 응하여야 한다.
- ④ 부정행위의 판정은 윤리위원 과반수의 의결로 이루어진다.

제11조(연구부정행위 조사결과 보고)

연구윤리위원회는 연구부정행위를 조사하고 그 결과를 조사종료 후 10일 이내에 회장에게 보고하여야 한다. 결과보고서에는 다음 각항이 반드시 포함되어야 한다.

- ① 제보의 내용
- ② 조사의 대상이 된 부정행위
- ③ 조사위원회의 조사위원 명단
- ④ 본 조사 실시 여부 및 판단의 근거
- ⑤ 해당 연구에서의 피조사자의 역할과 부정행위의 사실 여부
- ⑥ 관련 증거 및 증인

제12조(조사결과 보고에 대한 후속조치)

- ① 연구부정행위를 하여 본 학의 명예를 손상시킨 자에 대해서는 그 정도에 따라

최고 5년까지 본 학회의 활동을 중지시킨다.

- ② 연구부정행위에 대한 판정결과는 5년 이상 본 학회에서 보존하여야 한다.

제13조(심사의 객관성)

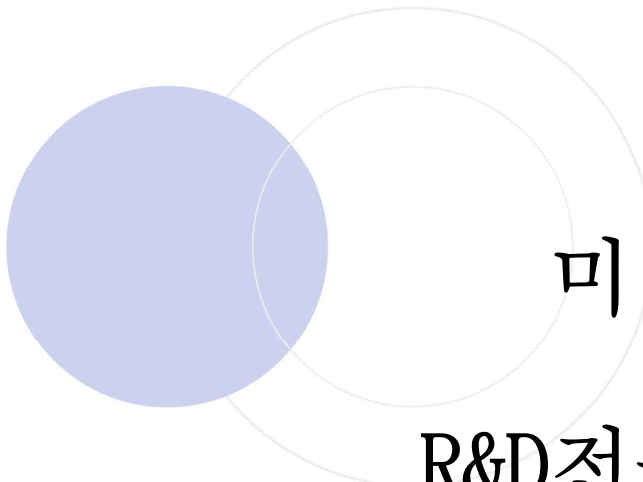
논문 심사위원 선정시 투고자와 동일기관 소속의 심사위원은 선정하지 않는다.

제14조(기타)

이 규정에서 정하지 않은 사항은 학회의 관련 내규를 따르거나 학회이사회에서 논의하여 결정한다.

부 칙

- ① 이 규정은 2013년 9월 1일부터 시행한다.
- ② 이 규정은 2014년 10월 1일부터 일부 수정하여 시행한다.
- ③ 이 규정은 2015년 3월 1일부터 일부 수정하여 시행한다.
- ④ 이 규정은 2018년 1월 1일부터 일부 수정하여 시행한다.



미래형자동차
핵심기술
R&D전문인력양성
센터 워크숍

◆ 김시호 (연세대학교 교수)

주변 차량과 상호 작용을 고려한 경로 예측 네트워크

¹이예린, ^{2*} 김시호

A Network for Trajectory Prediction based on Interaction with Surrounding Vehicles

¹Ye-rin Lee and ^{2*}Shiho Kim

요 약

자율주행 기술의 발전은 현대 교통 시스템에 근본적인 변화를 가져오고 있지만, 주변 환경과의 상호 작용을 충분히 반영하는 데에는 여전히 도전이 존재한다. 본 연구는 자율주행 차량이 다른 차량 및 보행자와의 상호 작용을 고려하여 보다 정확한 주행 경로를 예측하고 결정할 수 있도록 하는 것을 목표로 한다. 이를 위해, CARLA 시뮬레이터를 사용하여 다양한 운전 시나리오에서 데이터를 수집하고, VAE(Variational AutoEncoder)를 기반으로 한 경로 예측 모델을 개발했다. 또한, 운전자의 'Impatience'를 분석하여 차량의 차선 변경 행동을 예측하는 방법론을 적용했다. 이 연구는 자율주행 시스템의 정확성과 안전성을 향상시키는 데 중요한 기여를 할 것으로 기대된다.

Key words

Trajectory Prediction, Autonomous Driving, Simulation In the Loop, Driving Style

I. 서 론

자율주행 기술의 발전은 현대 교통 시스템의 근본적인 변화를 약속하고 있으나, 이 기술은 여전히 중요한 도전 과제에 직면해 있다. 특히, 자율주행 차량이 주변 환경과의 상호 작용을 충분히 반영하여 안전하고 효율적인 주행 결정을 내리는 능력은 인간 운전자에 비해 상대적으로 미흡한 상태이다. 이 연구는 자율주행 차량이 다른 차량 및 보행자와 같은 외부 요소와의 상호 작용을

고려하여 보다 정확한 주행 경로를 예측하고 결정할 수 있도록 하는 것을 목표로 한다.

본 연구의 핵심은 딥러닝과 시뮬레이션 기술을 활용하여 레벨4 이상의 자율주행 기술을 위한 경로 예측 모델을 개발하는 것이다. 이 모델은 주변 차량과의 상호 작용을 고려하여, 자율주행 차량이 다양한 주행 환경에서 보다 정확하게 주변 환경을 예측하고 적절한 주행 전략을 수립할 수 있도록 지원한다.

본 연구에서는 가상 자율주행 시뮬레이션인 CARLA[1]를 활용해 다양한

¹ 연세대학교 IT융합공학과 (BK21 지능형반도체IT융합전공) (yrin0703@yonsei.ac.kr)

² 교신저자 연세대학교 IT융합공학과, 교수 (shiho@yonsei.ac.kr)

운전 시나리오를 설계하고, 이를 통해 수집된 데이터를 기반으로 주변 차량의 행동을 분석하고 예측하는 모델을 제안한다. 이 모델은 VAE(Variational AutoEncoder)[2]를 기반으로 구성되어 데이터의 숨겨진 표현을 학습해 미래 행동을 예측하는 데 중점을 두었다.

통해 차량의 차선 변경 행동을 예측한다. 가상 자율주행 환경을 구성한 후, 5,000,000 프레임에 해당하는 차량의 데이터를 수집했다. 150대 차량의 위치, 속도, 운전자 성향 반영을 위한 Impatience 수치, 신호 정보 등을 수집해 경로 예측 네트워크의 학습 데이터로 활용했다.

II. 본 론

2.1 가상 자율주행 환경 구성

자율주행 시스템의 정확성과 안전성을 보장하기 위해, 다양한 차량과 참여자(예: 보행자, 자전거 등) 간의 복잡한 공간적 및 시간적 상호 작용을 이해하는 것이 중요하다. 우리는 실도로에서 자율주행 차량을 검증하기에 앞서, SIL (Simulation In the Loop)을 통해 네트워크를 검증하고자 했다. 이를 위해, 오픈소스 자율주행 차량 시뮬레이터인 CARLA 시뮬레이터를 사용했으며, 이는 Unreal Engine을 기반으로 하여 실제 도로 상황을 재현한다. CARLA에서는 다양한 도시 맵, 날씨 조건, 교통 상황 등을 설정할 수 있으며, 지원하는 다양한 주행 API를 통해 데이터를 수집했다.

2.2 운전자 성향 반영

본 연구에서는 자율주행 차량이 주변 차량과의 상호 작용을 보다 정확하게 이해하고 예측하기 위해 운전자 성향 분석[3]에 중점을 두었다. 이 분석의 핵심 요소는 ‘Impatience’, 즉 운전자의 답답함 정도[4]로, 이는 차선 변경, 급가속, 급제동, 높은 평균 주행 속도 등 운전 행동을 통해 나타난다. 우리는 각 차량의 Impatience를 그 차량의 현재 속도가 권장 속도보다 낮을 때 증가시키는 방식으로 계산하고, 이를

2.3 Variational AutoEncoder

수집한 데이터를 기반으로 미래의 주행 경로 예측을 위해 VAE 기반의 네트워크를 구성했다. VAE는 딥러닝과 확률론적 접근을 결합한 생성 모델로, 데이터의 숨겨진 표현을 학습하는 데 활용된다. VAE의 핵심 구조는 인코더, 잠재 공간, 디코더로 나눌 수 있다. 인코더는 입력 데이터 x 를 받아 평균과 표준편차의 출력을 만든다. 이 두 출력값은 잠재 공간에 있는 데이터의 확률 분포를 파라미터로서 표현한다. 잠재 공간은 입력 데이터의 압축된 표현을 담고 있는 VAE의 중심 부분이다. 잠재 공간의 각 포인트는 정규분포로부터 샘플링된 값을 바탕으로 평균과 표준편차를 사용하여 생성된다. 디코더는 잠재 공간에서의 포인트를 입력받아 $t+1$ 에 대한 예측된 경로를 출력한다.

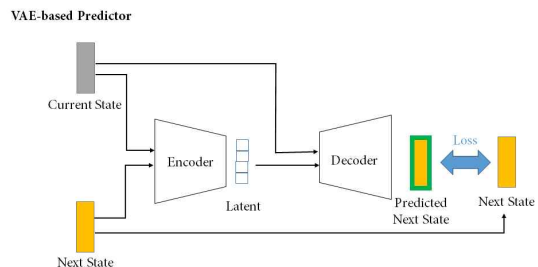


그림1. VAE 기반 경로 예측 네트워크

VAE 기반의 네트워크를 기반으로, 차량이 주행할 수 있는 경로를 3가지로 구분하여 학습했다. 차선 유지, 좌측 차선

변경, 우측 차선 변경의 경우로 나누어 각각의 행동에 대한 확률을 추론했다.

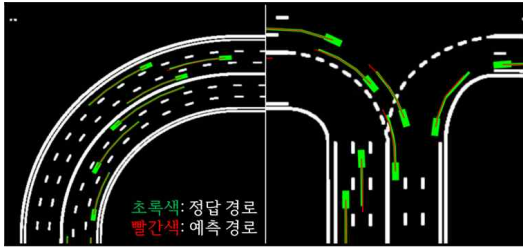


그림2 주변 차량과 상호작용을 고려한 경로 예측 결과

예측된 경로의 결과는 그림2와 같이 확인할 수 있었다. 각 차량마다 2.25초 후까지의 미래 경로를 예측하였고, 오차의 경우 평균 횡방향 0.161m, 평균 종방향 0.345m로 확인되었다.

Ⅲ. 결 론

본 연구에서는 주변 차량과의 상호 작용을 고려해 주행 경로를 예측하는 네트워크를 제안했다. 주변 차량의 주행 스타일을 분석하는 것이 상황에 맞는 주행 결정을 내리는 데 중요한 요소로 선택되었고, 운전자 성향 분석에 도움이 되는 답답함의 정도를 정의하여 입력 데이터로 사용하였다. 이후 입력 데이터의 내재된 표현을 학습하기 위해 VAE 기반의 네트워크를 구성하였고, 차선 변경 확률을 추론해 차량이 주행 가능한 경우를 고려하였다.

본 연구는 자율주행 차량이 인간 운전자와 유사한 수준으로 주변 환경과 상호작용하며 안전하고 효율적인 주행 결정을 내릴 수 있는 미래로 한 걸음 더 나아가는 데 중요한 기여를 할 것으로 기대된다.

감사의 말

이 논문은 2023년도 정부 (과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구임 (2021-0-01352, 자율주행 관련 법규 및 규제 대응 서비스 시나리오 실효성 검증 기술 개발)

참 고 문 헌

- [1] Dosovitskiy, Alexey, et al. "CARLA: An open urban driving simulator." Conference on robot learning. PMLR, 2017.
- [2] Kingma, Diederik P., and Max Welling. "Auto-encoding variational bayes." arXiv preprint arXiv:1312.6114 (2013).
- [3] Sagberg, Fridulv, et al. "A review of research on driving styles and road safety." Human factors 57.7 (2015): 1248-1275
- [4] Danaf, M., Abou-Zeid, M., & Kaysi, I. (2015). Modeling anger and aggressive driving behavior in a dynamic choice-latent variable model. Accident Analysis & Prevention, 75, 105-118.

악천후 날씨 환경 종류 인식 네트워크

¹주형진, ²송한빈, ^{3*} 김시호

A Network for Recognizing Environmental types of Adverse Weather

¹Hyeong-Jin Ju, ²Han-Bin Song and ^{3*}Shiho Kim

요 약

최근 인공지능, 머신러닝, 센서 기술, 자동차 공학 등 다양한 분야에서의 발전으로 자율주행 기술이 급부상하고 있다. 본 연구에서는 안정적인 자율주행을 위해 악천후 환경변화에서 상황을 인지하는 악천후 날씨 환경 종류 인식 네트워크의 개발 방안을 제시한다. 본 연구에서는 주행 환경에서의 악천후 데이터를 수집하고 재분류하여 학습데이터를 구성하고, 비전 트랜스포머를 기반으로 네트워크를 모델링한다. 이를 통해 악천후 영상 데이터의 다양한 특징을 효과적으로 추출하고 종류를 인식할 수 있을 것으로 기대된다. 본 연구에서 제시하는 방법과 결과는 안정적인 자율주행을 위한 AI SW 핵심기술로써 자율주행 상용화를 위한 기반을 마련하는데 실용적인 기여로 평가될 수 있다.

Key words

Adverse Weather environment, Image Classification, Autonomous Driving

I. 서 론

최근 인공지능, 머신러닝, 센서 기술, 자동차 공학 등 다양한 분야에서의 수년간의 연구, 개발, 혁신의 결정체로 자율주행 기술이 관심을 받고 상용화하기 위해 많은 연구가 이루어지고 있다. 자율주행을 상용화하기 위해선 일반적인 도로 상황뿐만 아니라, 비나 눈이 내리는 악천후 상황에서도 안정적으로 자율주행이

이루어지는 것이 중요하다[1]. 따라서 본 연구에서는 악천후 환경변화에서 상황을 인지하는 악천후 날씨 환경 종류 인식 네트워크의 개발을 목표로 한다.

본 연구에서는 공개 및 실도로 데이터 세트를 활용한 악천후 환경(눈/비/안개/조도 변화 등) 종류 인식 네트워크를 모델링한다. 인공지능 네트워크를 모델링하기 위해서는 관련 데이터가 필수적이다. 하지만 주행환경 영상 공개 데이터에는 맑은 환경에서의 주행 영상이

¹ 연세대학교 IT융합공학과/BK21 지능형반도체IT융합전공 (wngudwls000@yonsei.ac.kr)

² 연세대학교 IT융합공학과/BK21 지능형반도체IT융합전공 (thgksqls369@yonsei.ac.kr)

³ 교신저자 연세대학교 IT융합공학과, 교수 (shihoj@yonsei.ac.kr)

대부분이고, 약천후 환경에서의 데이터[2]가 부족하다. 효과적인 학습을 위해 서로 다른 환경에서 취득한 데이터들을 하나로 모아 네 개의 클래스로 재분류하는 작업을 거쳐 약천후 데이터를 확보한다. 학습을 위한 네트워크 구조로는 트랜스포머(Transformer) 기반의 아키텍처를 사용한다. 트랜스포머[3]는 영상 처리 네트워크중 SOTA(State-Of-The-Art) 성능을 보이는 네트워크 구조로써 강인한 학습 능력을 바탕으로 가상 데이터뿐만 아니라 실제 데이터에 대한 실용적 성능을 보유하고 있다. 또한 여러 약천후 상황의 영상 데이터의 특징을 추출하고 학습 개선을 위한 최적화 손실함수를 적용하여 기존 네트워크에 비해 성능이 개선되었다. 위의 방법으로 학습시킨 네트워크 모델을 통해 단일 영상 데이터를 입력으로 추론 함수를 수행한다.

II. 본 론

2.1 데이터 수집 및 전처리

약천후 관련 주행 데이터세트를 마련하기 위해 여러 공개 데이터세트를 수집한다. 수집한 데이터들을 학습시키려는 네 클래스(맑음, 비, 눈, 흐림)로 재분류한다. 상대적으로 맑은 날씨에 비해 약천후 관련 데이터가 부족하므로 다양한 증강기법(Translation, Rotation, Scale 등)을 적용하여 새로운 약천후 데이터를 생성할 수 있다. 이를 통해 학습을 위한 데이터 세트를 구성한다.

2.2 Vision Transformer

비전 트랜스포머(ViT)는 이미지 인식 및 처리 작업에서 높은 성능을 발휘하는 딥러닝

모델로, 자연어 처리(Natural Language Processing) 분야에서 성공을 거둔 Transformer 구조를 컴퓨터 비전(Computer Vision) 분야에 적용한 모델이다. 기존에 가장 많이 사용되던 CNN(Convolutional Neural Networks) 구조[4] 대신 Transformer 구조를 사용함으로써 전역적인 이미지 정보를 적절하게 포착하고 처리할 수 있다.

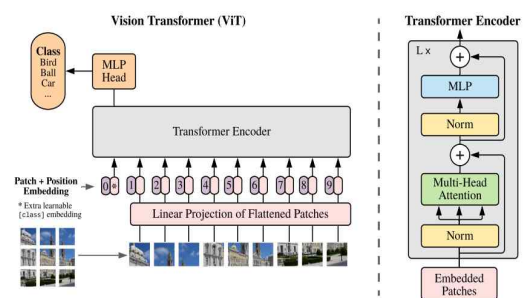


그림 1. Vision Transformer 개요도

비전 트랜스포머는 먼저 입력 이미지를 고정 크기의 작은 정사각형 패치로 나눈다. 각 패치는 1차원 벡터 형태로 변환된다. 변환 과정에서 Positional Embedding을 통해 각 패치의 위치에 대한 정보와 특별한 분류 토큰(Classification Token)이 추가된다. 벡터화하고 위치 및 분류 토큰 정보를 적용한 패치들을 비전 트랜스포머 인코더의 입력으로 사용한다. 인코더는 Self-Attention 및 Feed-Forward 신경망을 사용하여 이미지에서 복잡한 패턴과 문맥 정보를 추출한다. 인코더의 출력에서 분류 토큰의 정보를 사용하여 분류작업을 수행한다. 이를 위해 MLP(Multi-Layer Perceptron) 헤드가 사용되며, 벡터를 소프트맥스 활성화 함수가 적용된 최종 분류 결과로 변환한다.

비전 트랜스포머는 입력 영상 데이터의 다양한 특징을 집중할 수 있기 때문에, 중요한 정보를 더 잘 인지하고 활용할 수

있다. 이를 통해 기존 인공지능 모델을 사용할 때보다 더 높은 성능으로 맑음, 비, 눈, 흐림 등의 악천후 환경을 구분할 수 있게 된다.

Ⅲ. 결 론

본 연구에서는 악천후 환경 종류 인식 네트워크 모델링 방안에 대해 탐구했다. 먼저, 악천후 관련 공개 및 실도로 데이터 세트를 수집하고, 이를 재분류 및 전처리하여 입력 데이터를 생성한다. 입력 데이터는 네 가지 클래스(맑음, 비, 눈, 흐림)로 구분되어 비전 트랜스포머의 입력으로 사용된다. 비전 트랜스포머에서 입력 데이터의 복잡한 패턴과 문맥 정보를 추출하고 활용하여 결과를 출력한다. 이후 추론과 평가를 시행하여 모델의 성능을 파악한다. 이 과정에서 적절한 하이퍼파라미터를 설정하여 성능을 향상시킬 수 있다.

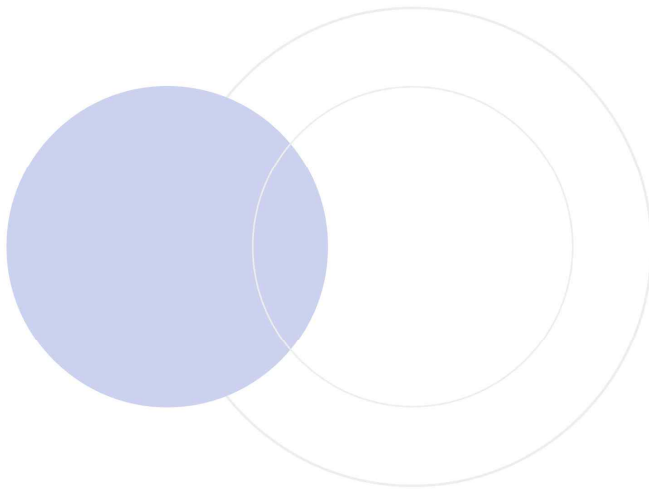
본 연구에서 제안된 악천후 환경 종류 인식 네트워크 모델은 안정적인 자율주행을 위한 하나의 AI SW 핵심기술로써 자율주행 상용화를 위한 기반을 마련하는데 기여할 것으로 기대된다.

감사의 글

이 논문은 2023년도 정부 (과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구임. (RS-2023-00236245, 악천후/비정형 환경변화에서의 Seamless 자율주행을 위한 인지/판단 AI SW 핵심기술 개발)

참 고 문 헌

- [1] Zhang, Yuxiao, et al. "Perception and sensing for autonomous vehicles under adverse weather conditions: A survey." *ISPRS Journal of Photogrammetry and Remote Sensing* 196, 2023.
- [2] Kenk, Mourad A., and Mahmoud Hassaballah. "DAWN: vehicle detection in adverse weather nature dataset." *arXiv*, 2020.
- [3] A. Dosovitskiy, et al. "An image is worth 16x16 words: Transformers for image recognition at scale." *ICLR*, 2021.
- [4] Q. A. Al-Haija, et al. "Multi-Class Weather Classification Using ResNet-18 CNN for Autonomous IoT and CPS Applications," *CSCI*, 2020.
- [5]



논문 발표

◆ 김 현 (부천대 교수)

차원축소와 클러스터링을 동시에 적용한 데이터 분석 방법에 대한 연구

⁴ 신병철, ⁵ 최승빈, ⁶ 김보성, ⁷ 배성민, ⁸ 송성민, ⁹ 장혁수, ¹⁰ 한동근, ^{11*} 박세진

A Study on Simultaneous Application of Dimensionality Reduction and Clustering in Data Analysis

¹Byung-Cheol Shin, ²Seung-Bin Choi, ³Bo-Sung Kim, ⁴Seong-Min-Bae,
⁵Seong-Min Song, ⁶Hyeok-Su Jang and ⁷Dong-Geun Han, ^{8*}Se-Jin Park

요 약

4차 산업혁명 시대로 넘어오며 데이터 분석이 중요해졌고 그 중 클러스터링과 같은 기술이 강조되고 있다. 본 논문에서 데이터 분석의 핵심요소인 클러스터링과 차원 분석 기법인 차원 축소에 대해 설명하고 클러스터링 차원에서 차원 축소를 적용했을 때 최대 약 2.95배 높은 실루엣 계수를 확인할 수 있다. 이를 통해 본 연구는 클러스터링과 차원 축소를 동시에 적용하였을 때 높은 성능 향상이 발생하고 저차원보다는 고차원에서 차원 축소가 더욱 효율적이라는 사실을 알 수 있다.

Key words

Dimensionality Reduction, Clustering Analysis, Clustering, PCA, t-SNE, UMAP

I. 서 론

4차 산업혁명 시대로 넘어오며 데이터 분석의 중요성이 점차 증가하고 있다. 데이터 분석이 중요해진 이유는 다양한 종류의 정보가 디지털화 되어 대용량이

되었고, 정형 또는 비정형 형태의 데이터를 활용하는 빅데이터 활용에서 그 필요성이 제기되었기 때문이다.

데이터분석의 방법에는 시각화와 공간 분석, 탐색적 자료 분석, 통계분석, 데이터 마이닝이 있다. 다양한 분석 방법 중 데이터

⁴ 계명대학교 수학과전공, 학부생 (cbzx28337q@naver.com)

⁵ 계명대학교 컴퓨터공학전공, 학부생 (csb8226@naver.com)

⁶ 계명대학교 컴퓨터공학전공, 학부생 (kimbosung22@naver.com)

⁷ 계명대학교 통계학전공, 학부생 (smin9e@naver.com)

⁸ 계명대학교 컴퓨터공학전공, 학부생 (song47166@naver.com)

⁹ 계명대학교 컴퓨터공학전공, 학부생 (hyeoksu459@naver.com)

¹⁰ 계명대학교 컴퓨터공학전공, 학부생 (ehdrms001030@naver.com)

¹¹ 교신저자 계명대학교 컴퓨터공학과 교수 (baksejin@kmu.ac.kr)

마이닝의 종류 중 하나인 클러스터링(Clustering)이 현재 강조되고 있는 모습을 보인다. 여기서 클러스터링이란[1] 각 객체의 유사성이 높은 대상 집단으로 분류하고, 군집에 속한 객체들의 유사성과 서로 다른 군집에 속한 객체간의 상이성을 규명하는 분석 방법을 의미한다. 클러스터링을 수행할 때 사용되는 데이터 객체는 객체가 가지고 있는 특징(feature)의 수가 많을 수 있으므로 많은 경우 유의미한 속성만을 가지는 차원 축소[2]를 수행한 후 가공된 데이터를 통해 데이터의 분석을 수행한다.

본 연구에서는 데이터 분석에서 강조되고 있는 방법인 클러스터링과 원활한 클러스터링을 위해 필요한 차원 축소에 대해 차원 축소의 효과와 중요성, 클러스터링과 차원 축소의 시너지 효과를 확인하기 위해 차원 축소의 유무에 따른 클러스터링 결과의 차이를 확인하며 클러스터링과 차원 축소의 시너지에 대해 논한다.

II. 차원 축소와 클러스터링

2.1 차원 축소의 이해

클러스터링과 차원 축소의 시너지에 관해 논하기 위해 차원 축소의 의미를 알 필요가 있다. 차원 축소란[2] 데이터 객체가 매우 많은 특징을 가질 경우 이를 축소하여 새로운 차원의 데이터 객체를 만들어 내는 것을 의미한다[3]. 데이터 객체에서는 일반적으로 차원이 증가 할수록 데이터 간의 거리가 기하급수적으로 증가하기 때문에 분석을 위한 알고리즘 실행이 매우 복잡하고 어려워진다.

이 문제를 해결하기 위해 데이터 차원 축소를 수행한다. 데이터 차원 축소는 [그림 1] 과 같이 분류된다.

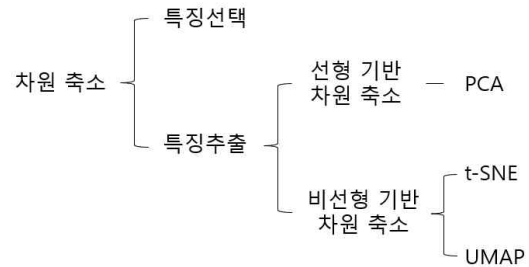


그림 1. 차원 축소 분류

데이터 차원 축소에는 유의미한 특징들만을 선택하는 특징 선택(feature selection), 상관관계가 높은 특징들을 합성하거나 변형하여 새로운 저차원의 특징을 만드는 특징 추출(feature extraction)로 분류할 수 있다[4]. 차원 축소는 이해하기 쉽고 빠르게 실행 가능한 데이터 모델을 만드는 데 그 목적을 두고 있다는 것이 특징이다. 두 방법 중 특징 선택은 질병 예측과 같은 의료 분야에서 데이터를 분석할 때 많이 사용되며, 특징 추출은 이미지 처리나 음성 데이터 분석에 사용하는 등 그 사용 범위가 더욱 넓다. 본 논문에서는 두 가지 방법 중 특징 추출에 대해 초점을 맞출 것이다.

데이터 차원 축소 방법 중 특징 추출에 속하는 기법에는 대표적으로 주성분 분석(PCA), t-SNE, UMAP 기법이 존재한다.

주성분 분석(PCA)[4]은 가장 대표적인 차원 축소 기법으로 여러 특성 간 존재하는 상관관계를 이용하여 대표하는 주 성분을 추출한 후 이를 이용해 차원을 축소하는 기법이다. 모든 차원의 축에 따른 값의 변화인 분산을 확인한 후 가장 높은 분산을 가지는 데이터 축을 찾아 차원을 축소하는 방법이 PCA이다.

t-SNE(t-Stochastic Neighbor Embedding)[4]는 고차원 상에 있는 이웃 간의 거리를 저차원 상에서도 최대한 보존하는

방향으로 차원을 축소하는 방법을 의미한다.

이를 위해서 목표 차원에 데이터를 무작위 배치 한 후 고차원 상에서의 배치와 비교하며 위치를 변경하는 방법이다.

UMAP(Uniform Manifold Approximation and Projection)[5]은 t-SNE와 유사한 기법으로, 데이터 포인트의 클러스터링이나 그룹을 시각화하는 데 효과적이다. 특히 각 데이터 포인트 그룹 간의 상대적인 근접성을 잘 보여주는 특징을 가지는 기법이다.

설명한 세 가지 차원 축소 기법은 PCA가 포함된 선형 기반 차원 축소 기법과 t-SNE와 UMAP 등이 포함된 비선형 기반 차원 축소 기법으로 분류할 수 있다[5]. 선형 기반 차원 축소[6]는 고차원의 데이터를 저차원으로 변환하는 과정에서 표현을 극대화하는 하나의 축을 기준으로 분산을 최대화하며 전체적인 데이터 구조를 보존하고자 하는 방법의 모음이다. 비선형 기반 차원 축소[6]는 선형 기반 차원 축소 방법의 한계인 복잡한 구조를 가진 다차원 데이터를 모델링함에 있어 어려움이 존재하기에 이 문제를 해결하고자 비선형 구조를 유지하면서 데이터 차원을 축소하는 방법의 모음이다.

하지만 이러한 비선형 기반 차원 축소도 강한 해석 가능성이 부족하고 전체적인 구조를 캡처하는 기능이 떨어진다는 단점[7]이 존재한다. 따라서 어떤 방법이 더 좋다고 보다는 각각의 기법이 특정 상황 또는 목적에 더 적합하며, 선택은 데이터의 특성과 분석 목표에 따라 달라진다고 할 수 있다.

2.2 클러스터 분석

클러스터링은[1] [그림 2]와 같이 객체의 유사성이 높은 대상 집단으로 분류하고, 군집에 속한 객체들의 유사성과 서로 다른 군집에 속한 객체간의 상이성을 규명하는 분석 방법을 의미한다.

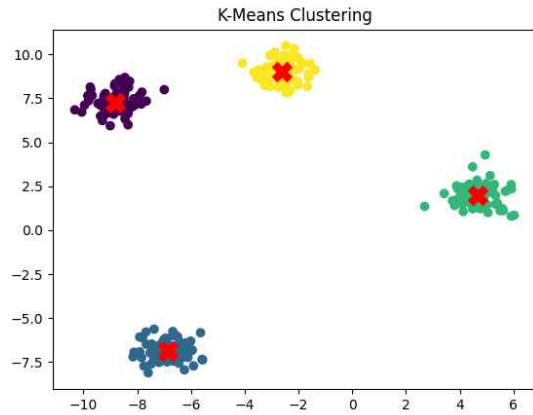


그림 2. K-Means Clustering

클러스터링의 방법에 따라 그 결과에서 차이가 발생할 수 있다. 클러스터링을 수행할 때는 데이터 간 유사성이나 근접성을 측정하여 유사한 경우 같은 군집으로 묶는다. 클러스터링에는 여러 가지 방법들이 존재하고 그 중 K-Means 클러스터링, Mean-Shift 클러스터링, DBSCAN에 대해 알아보겠다.

K-Means 클러스터링[8]은 가장 대중적인 클러스터링 방식이다. K-Means 클러스터링은 클래스 중심을 무작위로 초기화하고 각 데이터 포인트를 가장 가까운 중심에 할당하여 그룹을 형성한다. 반복적으로 그룹 중심을 업데이트하며 최적의 결과를 찾으므로 계산이 빠르고 효율적으로 수행되며, 설정된 반복 횟수나 중심의 큰 변화가 없을 때까지 계속 반복된다.

Mean-Shift 클러스터링은[8] 데이터 포인트가 밀집된 영역을 찾는 슬라이딩 윈도우 기반 클러스터링 방법이다. 해당방법은 중심점에 대한 후보를 슬라이딩 윈도우 내의 포인트의 평균을 가지고 각 그룹/클래스의 중심점을 찾는 것이 목표이다. 윈도우를 이용하여 후처리 단계에서 필터링하여 거의 중복을 제거해 최종 세트의 중심점과 그룹을 만들어낸다.

표 1. 클러스터링 기법의 활용분야

활용사례	알고리즘	연구내용
부식 모니터링[9]	K-Means 클러스터링	GrabCut을 활용한 강판 부식 모니터링 영상에서 K-Means 클러스터링을 사용하여 부식도 군집화
유전자 변이 분석[10]		닭의 피모색 픽셀을 K-Means 클러스터링으로 군집화하여 후보 유전자 변이에 대한 정확한 분류 기준 설정
얼굴 추적[11]	Mean-Shift 클러스터링	거리정보와 컬러정보를 이용하여 Mean-Shift 클러스터링 으로 얼굴 추적 및 사용자의 눈 추적 시스템 구현
눈 추적 시스템[12]		Mean-Shift 클러스터링과 신경망을 결합하여 사용자의 눈을 정확히 추출하고 추적하는 시스템 구현
대중교통 경로 선택[13]	DBSCAN	대중교통 정류장을 거리와 명칭 유사도를 기준으로 DBSCAN으로 군집화하여 최적의 경로 선택 지원
자율운항선박 분석[14]		항적데이터를 DBSCAN으로 군집화하여 선박의 입항패턴을 분석하고 자율운항선박의 자율적인 입항 및 접안 지원

DBSCAN은[8] Mean-Shift와 비슷한 밀도 기반 클러스터링 알고리즘으로. 시작 데이터 포인트를 기준으로 근방의 거리 ϵ 를 사용하여 클러스터를 형성한다. 충분한 근처 포인트가 있으면 클러스터링이 시작되고 해당 포인트는 새 클러스터의 첫 번째 포인트가 되고 충분한 포인트가 없으면 해당 포인트는 노이즈로 처리된다. 새 클러스터의 첫 번째 포인트에 대해 ϵ 거리 내의 점은 동일한 클러스터에 속하게 되며, 이 과정은 모든 점이 클러스터에 속할 때까지 반복된다. DBSCAN은 다른 클러스터링과 다르게 클러스터 집합이 필요하지 않다는 큰 장점이 존재한다.

소개한 클러스터링 방법들은 각각 다른 특성 및 장단점을 지니고 있는 만큼 활용하는 분야도 방법에 따라 상이하다. 각 분야별로 어떤 클러스터링 기법을 주로 사용하는지 알아보려고 한다.

[표 1]을 통해 클러스터링이 다양한 분야에서 사용되고 있다는 것을 알 수 있다.

가장 먼저 설명한 K-Means 클러스터링은 다른 클러스터링 대비 구현이 비교적 간단하다는 특징을 가지므로 이미지 분할이나 컴퓨터 비전, 지질통계학, 천문학 및 농업 등 다양한 분야에서 사용된다. 두 번째로 설명한 Mean-Shift 클러스터링은

연산의 단순성과 안정성이 뛰어나 영상 추적과 같은 실시간 처리에서 주로 사용된다. DBSCAN은 지리정보시스템, 이상치 감지 등 밀도와 관련한 분야에서 많이 사용된다.

이처럼 각 클러스터링 방식은 사용되는 분야가 각각 다르기 때문에 활용할 분야에 따라서 적절한 클러스터링 방법을 사용하여야 한다.

Ⅲ. 관련 연구

3.1 클러스터링 평가지표

차원 축소는 고차원 데이터를 저차원으로 투영하여 시각화가 가능하고 데이터의 패턴을 이해하고 해석하는데 도움이 된다. 또한 고차원 데이터에서 불필요한 정보를 제거함으로써 계산 효율성 또한 높다고 볼 수 있다. 하지만 원본데이터의 일부 정보를 손실시킬 수 있으며 해석에 어려움이 존재한다는 단점이 존재 한다.

클러스터링은 비슷한 특성을 가진 데이터를 그룹화하여 패턴을 식별할 수 있고 데이터의 불균형을 찾을 수 있다. 하지만 클러스터 중심에 위치에 따라 결과가 달라질 수 있고 고차원의 데이터를 다루기에는

버겁다는 단점이 있다.

하지만 이러한 차원 축소와 클러스터링을 병합하여 사용한다면 서로를 보완하는 효과가 발생할 수 있다. 고차원 데이터의 클러스터링 성능이 향상될 것이고 고차원의 데이터를 차원 축소를 통해 시각화하기 때문에 클러스터를 식별하고 해석하는 것도 쉬워질 것이다.

그렇다면 실제 실험을 통해 차원 축소와 클러스터링을 결합하여 사용하였을 때의 효과에 대해 알아보려고 한다. 우선 클러스터링이 성공적으로 진행되었는지는 확인할 수 있는 지표인 실루엣 계수[15]에 대해 알아보자.

실루엣 계수(Silhouette Coefficient)는 군집화 결과의 품질을 측정하는 지표 중 하나이다. 데이터 포인트 i 의 실루엣계수는 [그림 3]과 같이 계산된다.

$$s(i) = \frac{b(i) - a(i)}{\max(a(i), b(i))}$$

그림 3. 실루엣 계수

[그림 3]에서 $a(i)$ 는 데이터의 응집도를 나타내는 값이고 $b(i)$ 는 클러스터간의 분리도를 나타낸다. 실루엣계수는 -1에서 1까지의 값을 가지며, 높을수록 군집화의 품질이 더 좋다고 간주된다. 하지만 단순히 실루엣계수가 크다고 해서 군집화가 잘 됐다고 판단할 수 없다. 개별 군집의 실루엣 계수와 전체 실루엣 계수와 다르지 않아야 군집화가 잘 되었다고 판단할 수 있다. 전체 실루엣 계수가 크더라도 개별 군집의 실루엣 계수가 불규칙하다면 군집화가 잘되었다고 볼 수 없는 것이다.

3.2 차원 축소와 클러스터링의 결합

[표 2]의 데이터를 활용하여 클러스터링을

진행해보았다.

[표 2]는 각 이미지가 $28 * 28 = 784$ 개 차원의 벡터로 펼쳐진 데이터이다.

표 2. Fashion MNIST Dataset

Label	Class
0	T-shirt/top
1	Trouser
2	Pullover
3	Dress
4	Coat
5	Sanc
6	Shirt
7	Sneaker
8	Bag
9	Ankle boot

클러스터링을 진행하기에 앞서 각 클러스터링에서 필요한 값들을 구해야 한다. K-Means 클러스터링에서 k 값을 찾는 것이 중요하다. 최적의 k 값을 찾는 방법 중 하나는 Elbow Method[16]를 사용하는 것이다. Elbow Method는 군집화에서 적절한 k 값을 찾는 방법 중 하나이다. [그림 4]는 Elbow Method를 사용해 [표 2]의 최적의 k 값을 찾은 결과이다. [그림 4]를 통해 최적의 값이 $k=3$ 이라는 사실을 알 수 있다.

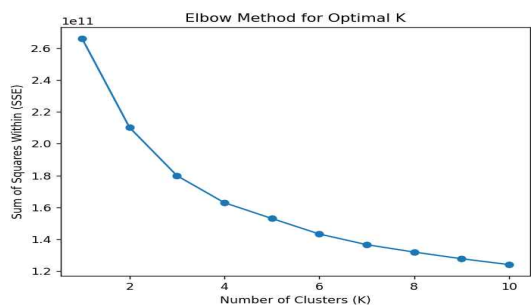


그림 4. Elbow Method

K-Means 클러스터링을 진행하였고 PCA, t-SNE, UMAP 총 3가지 차원 축소를 진행하여 실루엣 계수를 측정하였다. 실루엣 계수는 [표 3]와 같이 확인할 수 있다.

표 3. s.c (Silhouette Coefficient)

	K-Means	PCA	t-SNE	UMAP
s.c	0.186	0.445	0.408	0.549

[표 3]에서 고차원에서 클러스터링을 진행하였기 때문에 클러스터링만 진행하였을 때 상대적으로 실루엣 계수가 낮게 나온 것을 확인할 수 있다. 반면 클러스터링과 차원 축소를 병합하여 진행하였을 때 높은 실루엣 계수가 나타난 것을 확인할 수 있다.

실루엣 계수를 통해 UMAP이 대규모 데이터, 특히 고차원 데이터에 대해 효과적인 것을 확인할 수 있고 실험한 데이터에 대해 가장 적합한 차원 축소 방법인 것을 알 수 있다. Python의 Matplotlib을 사용하여 [그림 5]와 같이 2차원으로 시각화를 하였고 시각적으로 클러스터링이 잘 된 것을 확인할 수 있다.

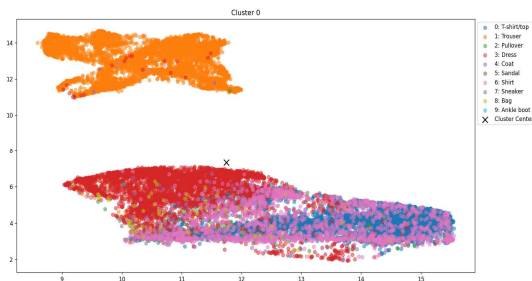


그림 5. UMAP을 사용하여 차원축소 후 시각화

따라서 실험을 통해 실루엣 계수를 도출해 낸 결과 고차원 데이터에 대해서는 어떤 차원 축소 방법을 사용하던 차원 축소가 모델의 성능을 개선한다는 사실을 도출해 낼 수 있다.

고차원의 데이터가 아니라 저차원의 데이터의 경우에도 차원 축소가 효과적인지에 대해서도 알아보았다. 아래와 같이 5차원으로 이루어진 [표 4]의 데이터를 활용하여 K-Means 클러스터링과 PCA

차원 축소방법을 진행했다.

표 4. Fish Dataset

	종	무게	길이	사선	높이	너비
1	잉어	242	25.4	30	11.5	4.02
2	잉어	290	26.3	31.2	12.4	4.3
3	잉어	340	26.5	31.1	12.3	4.69
...
158	빙어	12.2	13	13.8	2.27	1.25
159	빙어	19.7	14.3	15.2	2.87	2.06
160	빙어	19.9	15	16.2	2.93	1.87

표 5. Comparing Silhouette Coefficient

	Silhouette Coefficient
K-Means Clustering	0.5905
K-Means + PCA	0.5933

[표 5]와 같이 실루엣 계수를 확인할 수 있고 차원 축소를 하여도 크게 차이가 없는 모습을 확인할 수 있다. 따라서 차원축소는 저차원보다는 고차원 데이터에 있어서 효과적인 것을 알 수 있다.

3.3 도전과 고려사항

앞선 실험들을 바탕으로 클러스터링과 차원 축소를 결합하였을 때 저차원보다는 고차원에서 높은 성능 향상이 일어나고 좋은 시너지를 낸다는 것을 확인할 수 있었다. 하지만 차원 축소를 하였을 때 데이터의 정보가 일부 손실되거나 세부 사항이 누락되는 현상이 나타날 수 있다. 또한 축소된 차원에서의 각 차원이 원본 데이터의 의미를 명확하게 전달하지 않을 수 있고, 이는 클러스터링 결과의 해석을 어렵게 만들 수 있다. 즉 정보의 손실과 해석의 어려움이 발생할 수 있다.

이러한 문제점들을 해결하기 위해서는 PCA, t-SNE, UMAP 등의 차원 축소 기법을 사용할 때 목적에 맞는 적절한 차원을 선택하는 것이 중요하다. 주성분의 설명력이

충분한 차원을 선택한다면 정보의 손실을 줄일 수 있다.

또한, 차원 축소 기법 중 t-SNE 이나 UMAP 등 비선형 차원 축소 기법을 사용한다면 복잡한 데이터 구조를 비교적 잘 보존할 수 있다. 이외에도 정보의 손실과 해석의 어려움을 없애기 위한 해결 방향을 찾기 위해 다양한 연구를 진행할 필요가 있다.

IV. 결 론

본 연구는 클러스터링과 차원 축소에 대해 설명하고 K-Means 클러스터링에 PCA, t-SNE, UMAP이라는 3가지 차원축소 방법을 적용하여 차원축소 방법과 클러스터링을 결합하였을 때 높은 성능향상이 나타남을 보여주었다. 저차원보다 고차원에서 높은 성능향상이 나타났고 차원축소를 할 때 일부 데이터 손실과 해석의 어려움이 발생하는 것이 확인되었으며, 향후 이를 개선하기 위한 연구가 진행되기를 희망한다.

참 고 문 헌

[1] ‘군집분석’, 위키독스(WikiDocs) ‘개발자가 데이터 분석 준전문가 되기’, <https://wikidocs.net/48116>, chapter 3.4.5.

[2] ‘데이터 차원 축소’, 한국정보통신기술협회 정보통신용어사전, https://terms.tta.or.kr/dictionary/dictionaryView.do?word_seq=179641-1.

[3] 홍준호, 오민지, 조용빈, 이경희, 조완섭, 다차원 데이터의 군집분석을 위한 차원축소 방법: 주성분 분석 및 요인성분 분석 비교, 한국빅데이터학회지, 제5권, 제2호, 2020, page 135-137

[4] ‘차원축소’, 인코덤, <http://www.incodom.kr/차원축소>.

[5] Sivarajah, S, "Dimensionality reduction for data visualization: PCA vs TSNE vs UMAP vs LDA", Towards Data Science, 2020

[6] Laurens van der Maaten and Geoffrey Hinton, "Visualizing Data using t-SNE", Journal of Machine Learning Research, vol9, chapter1, 2008.

[7] McInnes, Leland, John Healy, and James Melville "UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction", Journal of Open Source Software, chapter6, pages 45, 2018.

[8] ‘데이터 과학자가 알아야 할 5가지 클러스터링 알고리즘’, Nextobe, <http://www.nextobe.com/2020/05/14/%EB%8D%B0%EC%9D%B4%ED%84%B0-%EA%B3%BC%ED%95%99%EC%9E%90%EA%B0%80-%EC%95%8C%EC%95%84%EC%95%BC-%ED%95%A0-5%EA%B0%80%EC%A7%80-%ED%81%B4%EB%9F%AC%EC%8A%A4%ED%84%B0%EB%A7%81-%EC%95%8C%EA%B3%A0%EB%A6%AC/>

[9] 김범수, 권재성, 최성웅, 노정필, 이경향, 양정현, K-Means 클러스터링을 이용한 강판의 부식 이미지 모니터링, 한국표면공학회지, 제54권, 제5호, 2021.10, pages 278-284.

[10] 박종호, 허선영, 김민준, 조은진, 차지혜, 진대혁, 고영준, 이승환, 이준현, K-Means Clustering으로 분류한 닭 깃털색 표현형의 분석, 한국가금학회지, 제49권, 제3호, 2022.09, pages 157-165.

[11] 최완용, 이윤형, 정문호, 양방향 필터 기반 Mean-Shift 기법을 이용한 강인한 얼굴추적, 한국전자통신학회 논문지, 제8권, 제9호, 2013.09, pages 1319-1324.

[12] 강신국, 김경태, 신윤희, 김나연, 김은이, 신경망과 Mean-shift를 이용한 눈 추적, 전자공학회논문지-CI, 제44권, 제1호, 2007.01, pages 56-63.

[13] 이민혁, 전인우, 전철민, 개선된 DBSCAN 알고리즘을 이용한 대중교통 정류장 군집화 기법, 한국지형공간정보학회지, 제25권, 제4호,

2017.12, pages 97-106.

- [14] 이형탁, 조익순, 양현, DBSCAN 알고리즘을 활용한 선박 입항패턴, 한국정보과학회 2020 한국소프트웨어종합학술대회 논문집, 2020.12, pages 433-435.
- [15] 김보선, 백의준, 지세현, 강민규, 신희종, 김명섭, 클러스터링 기반 비트코인 랜섬웨어 주소의 Ground-Truth 생성 방법, 한국통신학회학술대회논문집, 2021, pages 837.
- [16] Congming Shi, Bingtao Wei, Shoulin Wei, Wen Wang, Hai Liu and Jialei Liu, "A quantitative discriminant method of elbow point for the optimal number of clusters in clustering algorithm". EURASIP Journal on Wireless Communications and Networking, no. 31, chapter2, 2021.

AI 기반 V2X 통신 구간 대상 이상 탐지모델 제안

¹이진규, ²지일환, ³전규현, ^{4*}서정택

Proposal of AI-based anomaly detection model for V2X communication section

¹Jin Gyu Lee, ²Ilhwan Ji, ³GyuHyun Jeon and ^{4*}Jung Teak Seo

요 약

차량의 실시간 교통 정보나 위치 정보를 주고받기 위해 교통시스템에서는 V2X의 통신기술을 사용한다. 차량의 상태를 송수신하기 위해 단거리 전용 통신인 BSM(Basic Safety Message)을 사용한다. BSM은 설계당시 데이터 가용성을 위주로 설계되어 보안적 측면에서의 고려가 이루어지지 못했다. 이러한 이유로 사이버보안 취약점이 존재하며, 보안 위협이 발생 가능하다. 이에 대응하기 위하여, 최근 BSM에서 발생할 수 있는 사이버 공격을 대응하기 위해 AI기반의 이상탐지 시스템을 개발하고 있으나, 각 연구에 대한 체계적인 비교분석이 수행되지 못하고 있다. 본 논문에서는 최근 V2X 통신구간 대상 이상탐지 연구에서 제안하는 5가지의 모델에 대하여 훈련 데이터셋 수에 따른 이상탐지 비교 분석 실험을 진행하였으며, 테스트베드는 "VeReMi Dataset"을 사용하였다. 실험결과 99.6% Precision, 99.9% Recall, 99.7% F-1 Score를 도출하였다.

Key words

Transportation System, V2X, BSM, AI Anomaly Detection

I. 서 론

교통시스템은 V2X(vehicle to ever-ything)통신 기술을 통해 차량과 차량, 차량과 시설물 간의 위치나 속도와 같은 실시간 교통 정보를 주고받는다[1]. 교통 정보를 주고받을 때, BSM을 기반으로 통신하는데 BSM은 가용성 위주로 설계되어서 보안적 측면의 고려가 안 되어

다양한 사이버 보안 취약점이 존재하며, 데이터 변조와 DoS와 같은 보안 위협이 발생 가능하다. 이러한 보안 위협이 발생한다면, 교통 혼잡과 사고가 발생하여 금전적, 물리적, 인명 피해가 발생할 수 있다. 이러한 이유로 교통 시스템에서는 사이버 공격으로부터 사전에 이상을 탐지하고 대응하기 위하여 이상 탐지 연구가 수행되어야 한다. 하지만 기존에 수행되었던

¹ 가천대학교 기계·스마트·산업공학부 기계공학전공 학부생 (lee990107@gachon.ac.kr)

² 가천대학교 정보보호학과 일반대학원 석사과정 (ilhwan1013@gachon.ac.kr)

³ 가천대학교 정보보호학과 일반대학원 석사과정 (pengchan88@gachon.ac.kr)

⁴ 교신저자 가천대학교 컴퓨터공학부 컴퓨터공학부, 교수 (seojt@gachon.ac.kr)

규칙 기반 탐지방식은 빠르게 변화하는 공격방식에 대응하기 위한 규칙 업데이트 주기가 공격의 변화 속도를 따라갈 수 없고, 시그니처 기반 탐지방식은 알려지지 않은 공격에 대한 탐지를 수행할 수 없다는 문제점을 가진다[2][3]. 이러한 이유로 본 논문에서는 기존 이상 탐지 방식의 취약점을 보완하고 V2X 통신 구간을 대상으로 효과적인 이상탐지를 위하여 KNN(K-Nearest Neighbor), SVM(Support vector machine), LSTM(Long Short-Term Memory), GRU(Gated Re-current Unit), RF(Random Forest) 알고리즘을 이용한 AI 기반 이상탐지 방식을 제시한다. 또한, 제안하는 V2X 통신 구간 대한 이상탐지 방식의 성능 검증을 위하여 BSM에 대하여 정상데이터 및 5가지 종류의 공격 데이터를 포함하는 “VeReMi Dataset[4]”을 이용하였으며, 실험결과 99.6% 정밀도, 99.9% Recall, 99.7% F-1 Score를 도출하였다.

- 도로교통 환경에서의 V2X 단거리 전용 통신 대상 보안 위협 제시
- V2X 통신구간 대상 AI 기반 이상 탐지 방식 제안
- BSM(Basic Safety Message)에 대하여 정상데이터 및 5가지 종류의 공격 데이터를 포함하는 “VeReMi Dataset”을 사용하여 KNN, SVM, LSTM, GRU, RF 기반 이상탐지 모델의 성능 도출 및 비교분석

본 논문의 구성은 다음과 같다. 2장에서는 교통 시스템의 통신 기술에 대한 배경 지식을 제시하고 차량의 상태를 기반으로 교통 시스템 대상 이상탐지는 관련 연구를 분석한다. 3장에서는 공격 데이터셋과 이상탐지 모델의 기반이 되는 알고리즘을

제시 한다. 4장에서는 실험 결과에 대해 논의를 한 이후 5장에서 결론 및 향후 연구 방향을 제시한다.

II. 관련 연구 및 배경 지식

2.1 V2X(Vehicle-to-everything) 통신

도로에서 운행하는 차량 중 통신 기능이 있는 차량을 커넥티드 차량 (Connected Vehicle)이라고 부르며 커넥티드 차량의 구체적인 통신기술을 그림 1과 같이 V2X라고 정의한다[5].

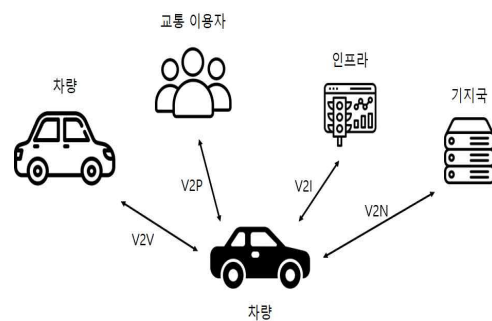


그림 1. V2X 통신 기술

V2X 통신기술은 미국 자동차공학회 (SAE)에서 제시한 J2735을 따라 메시지 형식을 갖춰 단거리 전용 통신을 한다. 단거리 전용 통신을 통해 OBU (On-Board Unit)와 기지국이나 교통 인프라에게 보내는 BSM(Basic Safety Message)을 전송한다[6]. BSM은 차량의 위치, 속도 및 가속도 등 현재 차량의 상태를 실시간으로 전송하는 메시지를 의미한다. BSM은 무결성을 유지하기 위해 수신자가 디지털 서명을 한다. 하지만 가용성을 위해 일반적으로 암호화가 되지 않으므로, 자격증명으로 인한 내부 공격이 발생할 수 있다. 내부 공격이 발생하면 위치 위조나 서비스 거부와 같은 사이버 보안 위협이

존재한다[7].

2.2 관련 연구

교통 시스템에서 이상 탐지와 관련된 연구는 표 1과 같다. 최근 연구에서 많이 사용된 모델로는 KNN과 SVM, Random Forest, LSTM 모델 등을 사용하여 이상 탐지를 제안하였다.

표 1. 관련 연구

저자	설명
Steven So (2018)[8]	KNN과 SVM을 사용하여 위치 스푸핑 공격을 탐지하였으며 각각 87%와 88%의 성능을 보였음
Sohan Gyawali (2019)[9]	로지스틱회귀, KNN, Bagging 등을 사용하여 이상 탐지를 제안하였으며, 각각 71%, 94%, 98%의 성능을 보였음
Prinkle Sharma(2019) [10]	로지스틱회귀, LSTM, Random Forest를 사용하여 오작동 탐지를 제안하였으며, 대부분 94%의 성능을 보였음
Secil Eran (2021) [11]	KNN과 Random Forest를 합친 앙상블 기법을 제안하였으며, 대부분 91%의 성능을 보였음
Harun Surej Ilango(2022) [12]	Random Forest를 변형하여 새로운 NPFADS 탐지 시스템을 제안하였으며, 98%의 성능을 보였음

Ⅲ. AI 기반 V2X 통신구간 대상 이상 탐지시스템 실험

3.1 이상 탐지(Anomaly Detection)

인공지능 기반 이상 탐지란 학습된 데이터를 기반으로 예상하지 않는 행위나 부적절한 행위를 탐지하는 것을 의미한다[13]. 인공지능 기반 이상탐지에는 알고리즘에 따라 SVM, K-NN, clustering 등의 머신러닝 기반의 이상탐지와 RNN,

LSTM등의 딥러닝 기반으로 나뉜다. 이상 탐지는 금융, 의료, 산업 등 다양한 분야에 적용할 수 있는데 최근 자율형 커넥티드카의 출현과 다양한 지능형 시스템의 발전으로 V2X 대상에 대한 이상탐지 연구도 활발히 이루어지고 있다[14]. 본 논문에서는 AI를 기반으로 차량의 BSM을 통해 위치와 동작 정보를 학습한다. 이후 AI가 예상하지 않은 위치와 동작에 대해 탐지한다.

3.2 V2X 통신구간 대상 이상탐지 알고리즘

본 논문에서는 KNN, SVM, Random Forest(RF) 3가지 머신러닝 알고리즘인과 LSTM, GRU 2가지 딥러닝 알고리즘 총 5가지의 모델을 학습하였다. KNN은 일반적인 이상 탐지 기법으로 구현이 간단하며 다중 클래스를 쉽게 처리할 수 있다. SVM은 평면의 포인트를 매핑하여 데이터를 처리하는 식별 분류기이다. 서프트 벡터와 커널 함수를 통해 선형 및 비선형 문제에 대해 효과적으로 해결할 수 있다[9]. Random Forest는 여러 의사 결정 트리를 무작위로 생성하고, 각 트리의 예측을 결합하여 안정적인 예측을 제공하는 앙상블 학습 모델이다. 무작위로 선택된 데이터 샘플과 특징으로 학습되므로 다양한 데이터 패턴에 대한 예측할 수 있는 성능이 뛰어나다[10]. 시계열 데이터의 분류를 위하여 사용되는 RNN의 기반의 LSTM과 GRU은 장기적인 종속성을 기억하는 장점을 가져 기존 RNN 알고리즘에 비해 성능이 뛰어나다[15][16].

3.3 VeReMi 데이터셋

본 논문에서 사용된 Dataset은 “VeReMi(Vehcular Reference Misbehavior) Dataset”[4]으로 차량의 OBU(On-Board Unit)에서 단거리 전용

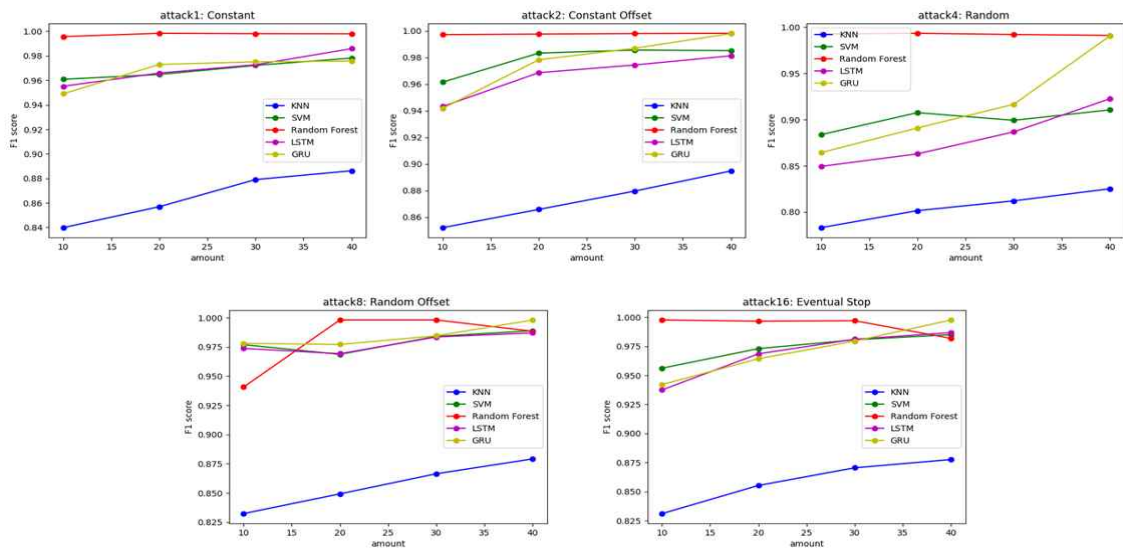


그림 2 공격유형별 학습 데이터양에 따른 F-1 score 성능

통신을 통해 주변의 차량이나 RSU(Road Side Unit)에게 수신된 BSM을 모두 포함하는 차량별 메시지 로그로 구성되어 있다. Dataset에는 약 300,000개의 정상 데이터와 약 100,000개의 공격 데이터가 존재한다. 공격 데이터는 표 2와 같이 5가지 종류로 이루어져 있으며, Dataset도 각각 공격 데이터에 따라 5가지로 나뉘어져 있다. 본 논문에서는 차량의 위치, 차량의 속도, 데이터 송수신하는 시간, RSSI(Received Signal Strength Indicator), AoA(Angle of Arrival) Feature를 학습 및 이상탐지를 위한 Feature로 사용하였으며, 데이터량의 변화에 따른 이상탐지 모델의 성능을 확인하기 위해 실험 시 약 400,000개의 데이터를 100,000개부터 400,000만개 까지 증가시키며 학습 및 검증을 진행하였다. 훈련데이터와 평가데이터는 75:25의 비율로 학습하였다.

표 2. VeReMi 공격 유형 설명[9]

공격 유형(ID)	설명
Constant (Type 1번)	고정된 위치를 전송하여 공격
Constant Offset (Type 2번)	고정된 위치를 일정값을 변경하며 공격
Random (Type 8번)	랜덤 위치를 전송하여 공격
Random Offset (Type 8번)	랜덤 위치에서 주변 차량 위치를 전송
Eventual Stop (Type 16번)	일정 시간 동안에 같은 위치를 반복하여 공격

3.4 성능 평가 지표

본 논문에서는 각각의 이상 탐지 모델에 대한 성능을 측정하기 위해 Precision(1), Recall(2) 및 F-1 Score(3)을 사용하였다. Precision은 이상데이터로 판단한 데이터 중 실제로 이상이 있는 데이터의 값의 비율을 측정하는 것을 의미하며 Recall은 예상되는 정상 데이터 중에 실제로 정상데이터의 값의 비율을 측정하는 것을 의미한다. 또한, Precision과 Recall 성능지표를 모두 반영하여 모델의 전체 정확도를 측정하는데

사용하는 평가지표인 F-Score를 사용하였다.

$$Precision = \frac{TP}{TP+FP} \quad (1)$$

$$Recall = \frac{TP}{TP+FN} \quad (2)$$

$$F-1(score) = \frac{2}{\frac{1}{Precision} + \frac{1}{Recall}} \quad (3)$$

IV. 실험 결과 및 논의

4.1 실험 결과 및 논의

그림 2는 각 이상탐지 모델에 대하여 공격유형별 학습 데이터양에 따른 F-1 Score 성능을 나타낸다. 대부분의 이상탐지 알고리즘에서 dataset의 데이터의 수가 증가함에 따라 공격 유형에 대한 이상 탐지 성능이 향상됨을 확인할 수 있다. 표 3은 실험을 통해 데이터의 양이 400,000개인 경우 이상탐지 모델의 성능을 나타내고 있으며 각 모델의 F-1 score 평균은 87% 이상의 성능을 보이고 있다. 또한, F-1 score 기준으로 공격자가 고정된 위치(Type 1번, 2번)를 공격할 때 Random Forest가 각각 97.79%와 99.81%로 가장 성능이 뛰어났으며, 공격자가 위치(Type 4번, 8번)를 바꿔가며 공격을 주입하거나 일정 시간동안 반복하여 공격할 때(Type 16번)에는 GRU모델이 각각 99.34%, 98.69%, 98.67%로 이상 탐지에 성능이 적합하였다. 또한, RNN의 기반의 LSTM과 GRU는 데이터양이 증가할수록 성능이 가장 많이 향상되었다.

표 3. 최대 데이터양에 대한 이상탐지모델의 탐지성능

ID	Model	precision	Recall	F-1 score
Constant (Type 1)	KNN	0.8875	0.8850	0.8862
	SVM	0.9772	0.9791	0.9781
	LSTM	0.9738	0.9778	0.9757
	GRU	0.9853	0.9863	0.9858
	RF	0.9982	0.9977	0.9979
Constant Offset (Type 2)	KNN	0.8952	0.8943	0.8947
	SVM	0.9851	0.9809	0.9829
	LSTM	0.9825	0.9805	0.9815
	GRU	0.9812	0.9862	0.9835
	RF	0.9979	0.9983	0.9981
Random (Type 4)	KNN	0.8251	0.8237	0.8244
	SVM	0.9104	0.8995	0.9039
	LSTM	0.9003	0.8774	0.8851
	GRU	0.9925	0.9944	0.9934
	RF	0.9905	0.9914	0.9909
Random Offset (Type 8)	KNN	0.8790	0.8790	0.8790
	SVM	0.9788	0.9788	0.9788
	LSTM	0.9783	0.9783	0.9783
	GRU	0.9869	0.9870	0.9869
	RF	0.9779	0.9776	0.9777
Eventual Stop (Type 16)	KNN	0.8775	0.8775	0.8775
	SVM	0.9849	0.9849	0.9849
	LSTM	0.9819	0.9817	0.9817
	GRU	0.9867	0.9867	0.9867
	RF	0.9775	0.9775	0.9775
Average	KNN	0.87286	0.8719	0.87236
	SVM	0.96928	0.96664	0.96772
	LSTM	0.96536	0.96114	0.96246
	GRU	0.97252	0.97412	0.97326
	RF	0.9964	0.99656	0.99646

V. 결론 및 향후 연구방향

본 논문에서는 교통 시스템 중 차량에서 위치 데이터를 보낼 때 사용되는 BSM이 위치 위조나 서비스 거부와 같은 사이버 보안 위협이 존재하는 것을 확인하였다. 이러한 위협을 탐지하기 위해 5가지 종류의 AI기반 이상탐지 모델을 제안하였다. 가장 성능이 좋은 모델을 선정하기 위해 "VeReMi Dataset"을 이용하여 각각의 모델을 비교하였다. 모델을 비교 및 분석한 결과

Random Forest와 GRU가 가장 성능이 좋다는 것을 확인하였고, [그림 2]를 통해 RNN 기반의 LSTM과 GRU가 실제 데이터에 대한 탐지 모델로 성능이 우수할 것으로 예상하였다. 향후 연구 방향으로는 평가를 통해 성능이 우수한 Random Forest와 GRU를 앙상블하여 새로운 프레임워크를 제안하고자 한다.

Acknowledgment

본 연구는 한국서부발전(주)의 과제 "AI(인공지능)를 이용한 신재생에너지 제어시스템 사이버공격 탐지기술 개발"의 지원을 받아 수행된 연구임 (2023-위탁-11).

참 고 문 헌

- [1] 김훈, "위치정보 산업 동향보고서", 한국인터넷진흥원, 2023 4월호
- [2] Koliadis, C., Kambourakis, G., Stavrou, A. and Voas, J. DDoS in the IoT: Mirai and other botnets. Computer, 2017.
- [3] Radoglou-Grammatikis, P. et al, "DIDEROT: An intrusion detection and prevention system for DNP3-based SCADA systems". ARES '20, 2020.
- [4] VeReMi dataset, <https://veremi-dataset.github.io/>
- [5] Raj Bridgelall, "Perspectives on Securing the Transportation System", MDPI Vehicles, 2022
- [6] J. B. Kenney "Dedicated Short-Range Communications(DSRC) Standards in the United States", IEEE Xplore, 2011
- [7] Aekta Sharma "Machine Learning Based Misbehaviour Detection in VANET Using Consecutive BSM Approach", IEEE Open Journal of Vehicular Technology, 2021
- [8] Steven So외 3명, "Integrating Plausibility Checks and Machine Learning for Misbehavior Detection in VANET", IEEE International Conference on Machine Learning and Applications, 2018
- [9] S. Gyawali외 1명, "Misbehavior Detection using Machine Learning in Vehicular Communication Networks.", IEEE International Conference on Communications (ICC), 2019
- [10] Prinkle Sharma외 2명, "Attacks on Machine Learning: Adversarial Examples in Connected and Autonomous Vehicles", IEEE International Symposium on Technologies for Homeland Security, 2019
- [11] Secil Ercan외 2명, "Misbehavior Detection for Position Falsification Attacks in VANETs Using Machine Learning", IEEE Access, 2022
- [12] Harun Surej Ilango외 2명, "A misbehavior detection system to detect novel position falsification attacks in the Internet of Vehicles", Engineering Applications of Artificial Intelligence, 2022
- [13] Mohammed Lamine Bouchouia, "A survey on misbehavior detection for connected and autonomous vehicles", Vehicular Communications, 2023
- [14] 박현목, "이상 탐지 1부-정상과 비정상, 그리고 이상 탐지", NHN 클라우드 MeetIP, <https://meetup.nhncloud.com/posts/362>
- [15] Franco van Wyk외 3명 "Real-Time Sensor Anomaly Detection and Identification in Automated Vehicles", IEEE Transactions on Intelligent Transportation Systems, 2020
- [16] Rui Fu외 2명, "Using LSTM and GRU neural network methods for traffic flow prediction", 2016 31st Youth Academic Annual Conference of Chinese Association of Automation (YAC), 2016

인공지능 할루시네이션에 대응하는 질의개선

¹ 이현정, ² 홍세영, ³박찬솔, ^{4*}김영철

Query Improvement against a Artificial Intelligence Hallucination in Large Language Model

¹Hyun Jeong Lee, ²Se Young Hong, ³Chansol Park and ^{4*}R. Young Chul Kim

요 약

현재 ChatGPT를 비롯한 인공지능 대화 시스템이 급격히 발전되고 있다. 이에 따라 교육, 창작, 검색 등의 다양한 분야에서 인공지능 대화 시스템의 사용이 활발하다. 하지만 대규모 언어 모델(Large Language Model, LLM)은 완벽하지 않다. 특히 LLM의 출력이 정확하지 검증되지 않은 것은 큰 문제이다. 실제로 인공지능 할루시네이션과 같이 실제 패턴이나 객체에 부합하지 않는 무의미하거나 부정확한 결과를 반환하는 경우가 빈번하다. 본 논문에서는 이러한 문제를 해결하기 위해 체계적인 프롬프트 작성 방법을 제안한다. 정확한 질의 프롬프트를 통해 인공지능이 정확한 답변을 할 수 있도록 유도한다. 이를 통해 일반적인 사용자도 쉽게 인공지능 할루시네이션을 회피할 수 있을 것으로 기대한다.

Key words

Artificial Intelligence, Large Language Model, ChatGPT, AI Hallucination

I. 서 론

본 논문은 2023년 홍익대학교 소프트웨어융합학과 종합 설계 과목에 결과물이다.

최근, LLM의 성장이 가파르다. 그 중에서도 생성형 언어 모델을 이용한 ChatGPT는 교육, 창작, 검색 등 다양한 분야에서 응용되고 있다. 하지만 LLM과 관련된 다양한 종류의 문제점들 또한

존재한다. 이러한 문제점들 중 하나는 생성된 답변의 정확도 문제이다. 그 중에서도 인공지능 할루시네이션 문제는 모델이 실제 패턴이나 객체에 부합하지 않는 무의미하거나 부정확한 결과를 생성하는 현상을 말한다. 본 논문에서는 인공지능 할루시네이션 문제를 해결하기 위해 체계적인 프롬프트 작성 방법을 제안 한다. 이 방법은 생성형 인공 지능의 출력이 사용자의 의도에 더 잘 부합하도록

¹ 홍익대학교 소프트웨어융합학과, 학부생 (hyunjlee@g.hongik.ac.kr)

² 홍익대학교 소프트웨어융합학과, 학부생 (C089084@g.hongik.ac.kr)

³ 홍익대학교 소프트웨어융합학과, 석사과정 (C2193102@g.hongik.ac.kr)

⁴ 교신저자 홍익대학교 소프트웨어융합학과, 교수 (bob@hongik.ac.kr)

설계되어, 생성형 인공 지능의 사용성과 효율성을 높인다. 또한 잘못된 프롬프트로 인한 할루시네이션 문제를 해결할 수 있다.

본 논문은 다음과 같이 구성된다. 2장에서는 LLM과 인공지능 대화 시스템 그리고 인공지능 할루시네이션의 문제에 대해 설명한다. 3장에서는 본 논문에서 제안하는 체계적인 프롬프트 작성 방법에 대해 소개한다. 마지막으로, 4장에서는 본 논문의 결론과 향후 연구에 대해 언급한다.

II. 관련연구

2.1 대규모 언어 모델

LLM은 대량의 텍스트 데이터를 학습하여 자연어 이해와 생성 작업을 수행할 수 있다[1]. LLM의 규모는 국내외에서 경쟁적으로 확장되고 있으며, 논문 작성, 프로그래밍, 언어 번역, 교정, 콘텐츠 생성과 같은 다양한 방면에 이용되고 있다.[2].

2.2 인공지능 대화 시스템의 현황

인공지능 대화 시스템은 음성 또는 문자를 사용해 사람과 컴퓨터 프로그램이 대화를 통하여 특정 작업을 수행하도록 만들어진 시스템이다. 현대의 인공지능 대화 시스템은 경쟁적으로 확장되며, ChatGPT와 같은 인공지능 대화 시스템은 다양한 분야에서 활용되고 있다. ChatGPT 이외에도 META의 LLAMA2, 구글 BARD와 같은 다양한 모델이 개발 되었다. 인공지능이 생성한 답변은 논리적으로 보일 수 있지만, 인터넷에서 수집된 데이터 셋으로 사용자와 관련된 편향성을 가지고 있을 가능성이 존재한다[2].

2.3 인공지능 할루시네이션

ChatGPT는 종종 사람들의 질문에 정확한

답을 알지 못하면서도 정확한 답을 찾는 것처럼 답변하는 경우가 있다. 이를 할루시네이션 현상이라고 한다. 대화형 인공지능의 할루시네이션은 편향되거나 부정확한 데이터, 입력에 대한 정확한 이해가 아닌 통계적 패턴을 기반으로 응답을 생성하는 모델의 경향으로 인해 발생 가능하다[3]. 다시 말해 훈련데이터의 문제, 확률적 추론의 문제, 진실 데이터의 부재, 모델의 복잡성 등 다양한 원인으로 인해 부정확한 답변을 제시하게 된다.

순다르 피차이 구글 최고경영자 (CEO)는 "바드 등 AI 챗봇에서 발생하는 할루시네이션 원인과 해결책을 그 누구도 제시하지 못하고 있다"며 "앞으로 AI 개발자 최우선 임무가 될 것"이라고 주장했다[4].

표 1은 LLM 할루시네이션 오류 사례이다. 현재로서 GPT-4가 인간과 비슷한 판단력을 지닌 범용 인공지능 (AGI) 수준에 도달하지 못하고 있다. 특히 문장의 이해 능력이 제한적이며, 정보가 부족하거나 사람들이 많이 물어보지 않은 질문에는 적절한 답을 생성하지 못하는 제한된 성능을 보인다.

표 1. LLM 할루시네이션 오류 사례

질문	공쥐팔쥐 이야기 알려줘
ChatGpt 대답	<p>"공쥐팔쥐"는 한국 전래 동화 중 하나로, 두마리의 작은 동물인 공쥐와 팔쥐가 등장하는 이야기입니다. 이 이야기는 어린 이들을 대상으로 한 교훈적인 내용을 담고 있습니다. 아래는 "공쥐팔쥐" 이야기의 일반적인 버전입니다.</p> <p>옛날 옛적에 공쥐와 팔쥐라는 두마리의 친구가 살고 있었습니다. 어느 날, 둘은 함께 밭에 나가 농사를 짓기로 결심했습니다. 공쥐는 콩을 심고, 팔쥐는 팥을 심었습니다.</p> <p>시간이 흘러 가을이 되었을 때, 공쥐의 밭에는 푸른 콩이 아름답게 자라고, 팔쥐의 밭에는 붉은 팥이 풍성하게 열렸습니다. 두마리는 자신의 농사 결과에 자랑스러워하며 서로에게 자랑을 했습니다.</p> <p>그러나 어느 날, 공쥐는 팔쥐를 자신의 집에 초대했습니다. 공쥐는 팔쥐에게 콩으로 만든 다양한 음식을 대접했습니다. 그리고 팔쥐는 공쥐를 자신의 집에 초대하여 팥으로 만든 다양한 음식을 대접했습니다.</p> <p>두마리는 서로의 집에서 즐거운 시간을 보내며 농사의 결과물을 나누었습니다. 이를 통해 둘은 서로의 노력과 공헌을 인정하고 협력의 중요성을 깨달았습니다.</p> <p>"공쥐팔쥐"는 이런 협력과 공유의 가치를 가르치며, 서로 다른 노력이 모여 풍성한 결과를 가져오는 소중함을 전하는 이야기입니다.</p>

표 2. 관련 배경정보를 통한 질의개선 예시

질문	<p>공쥐팔쥐는 공쥐와 팔쥐라는 두 자매의 이야기로 계모에 대한 모순된 대우와 학대로 자매간의 차별에 대한 비판적 시각을 제시하는 이야기야 계모 밑에서 갖은 학대를 받던 공쥐가 고귀한 인물과 혼인하게 되고, 공쥐를 괴롭히던 팔쥐와 계모는 처벌받는다는 내용의 설화이지 다시 한번 공쥐팔쥐 이야기를 알려줄래?</p>
ChatGpt 대답	<p>공쥐와 팔쥐가 자매인 버전의 "공쥐팔쥐" 이야기는 한국의 전래 동화 중 하나로, 주로 공쥐의 어머니와 계모에 대한 모순된 대우, 그리고 자매간의 협동과 도움을 강조하는 내용을 담고 있습니다. 이 이야기는 전통적인 가치관과 도덕적인 교훈을 전하는데 중점을 두고 있습니다.</p> <p>이야기의 주요 흐름은 다음과 같습니다: 옛날 옛적에 공쥐와 팔쥐는 어머니와 함께 살고 있었습니다. 그러나 어머니는 공쥐를 편애하며 팔쥐에게는 불공평한 대우를 했습니다. 어머니는 공쥐에게는 좋은 음식과 편안한 공간을 제공하면서도 팔쥐에게는 나쁜 대접을 했습니다.</p> <p>한 날, 공쥐는 산에서 우연히 선녀를 만나게 됩니다. 선녀는 공쥐에게 도움을 주고 감사의 표시로 보석들을 선물했습니다. 공쥐는 선녀에게 감사의 인사를 전하고, 선녀는 "어떤 어려움에 처했을 때, 내 이름을 불러 도움을 청하면 나타날 것이다" 라고 말합니다.</p> <p>이후, 공쥐는 어머니의 대우에 불만족하며 선녀를 불러 도움을 청합니다. 선녀는 공쥐를 도와 어머니에 대한 불공평한 대우를 바로잡아줍니다. 이에 어머니는 공쥐와 팔쥐에게 공평한 대우를 하게 되지만, 어머니의 태도는 변하지 않습니다.</p> <p>이후, 공쥐는 팔쥐를 돕기 위해 선녀를 다시 부르고, 선녀의 도움으로 팔쥐에게도 공평한 대우가 이루어집니다. 공쥐와 팔쥐는 서로 협동하여 어머니와 함께 행복한 삶을 살게 되며, 결국 공쥐는 선녀와의 인연을 통해 혼인하게 됩니다.</p> <p>이 이야기는 가치관의 변화와 협동의 중요성, 그리고 선행과 감사의 가치를 강조하는 전래동화로, 어머니와 계모에 대한 불공평한 대우를 극복하고 가족간의 협력을 통해 행복한 결말을 이루어냅니다.</p>

Ⅲ. 할루시네이션 해결을 위한 질의개선 연구

할루시네이션 오류를 해결하기 위해 모델을 개선하거나, 대화 시스템을 수정할 수 있다. 하지만 일반적인 사용자가 적용하기에는 어려운 방법이다.

일반적인 사용자가 인공지능의 할루시네이션을 회피하기 위한 방법으로 질의 프롬프트 개선을 제안한다. 질의 프롬프트 개선을 통해 인공지능 대화 시스템으로부터 정확한 답변을 유도할 수 있다.

3.1 관련 배경정보 추가

표 2는 관련 핵심 키워드를 프롬프트에 포함 하여 ChatGPT 답변의 정확도를 향상한 것이다. 배경정보를 프롬프트에 포함하면 모델이 대화의 맥락과 의도를 더 잘 이해할 수 있게 되어 모델은 더 일관된 답변을 생성하고, 사용자의 의도에 더 잘 부합하는 내용을 생성할 가능성이 높아진다. 예를 들어, "한국 전래동화 중에서 콩쥐팥쥐에 대해 알려줘"라는 프롬프트를 사용하면, 모델은 사용자가 전래동화에 대한 정보를 요청하는 것임을 명확히 이해하고 이에 맞는 정보를 생성한다.

3.2 관련 핵심 키워드 추가

표 3. 핵심 키워드를 통한 질의개선 예시

질문	콩쥐팥쥐는 콩쥐 와 팥쥐라는 사람의 이야기로 자매간의 불공평한 대우에 관한 전래동화야 다시한번 콩쥐팥쥐 이야기에 대해서 알려줘
ChatGpt 대답	"콩쥐팥쥐"는 한국의 전래동화 중 하나로, 자매 간의 불공평한 대우에 대한 이야기를 다룹니다. 이 이야기는 한국의 전통적인 가치관과 더불어 도덕적인 교훈을 담고 있습니다. 이 동화는 주로 두 자매, 콩쥐와 팥쥐에게 중심을 둡니다. 두 자매는 모두 열심히 노력하며 살아가지만, 환경이나 상황에 따라서 각자의 운명이 다르게 전개됩니다. 대개 콩쥐는 불운하게, 팥쥐는 행운을 맞이하는 경우가 많습니다. 이 동화는 주로 불행한 처지에서 희망을 잃지 않고 노력하는 데 중점을 두며, 어떤 상황에서도 긍정적으로 행동하고 인내심을 가지는 것의 중요성을 강조합니다. 또한, 자매 간의 차별에 대한 비판적 시각을 제시하여 공평하고 인간다운 사회를 지향하는 메시지를 전합니다.

표 3은 관련 핵심 키워드를 프롬프트에 포함하여 ChatGPT 답변의 정확도를 향상한 것이다. 프롬프트에 핵심 키워드를

포함시키면 ChatGPT는 사용자의 의도와 관련된 핵심 주제를 명확히 이해하고, 그에 따른 응답을 생성하는 데 도움이 된다. 특히 핵심 키워드는 모델이 어떤 정보를 포함해야 하는지에 대한 신호를 제공하며, 이를 통해 모델이 더 명확하고 정확한 답변을 생성할 수 있다. 예를 들어, "콩쥐팥쥐"가 사람인 것에 관한 정보를 얻고 싶다면, "콩쥐와 팥쥐라는 사람의 이야기로 자매 간의 협동과 도움을 강조한 전래동화"라는 키워드를 프롬프트에 포함시킬 수 있다. 이렇게 하면 모델은 주어진 키워드를 중심으로 정보를 제공하려고 시도하게 된다.

ChatGPT를 사용할 때 가장 중요한 것은 프롬프트 작성이다. 관련 배경 정보를 제공하면 ChatGPT가 질문의 맥락을 더 잘 이해 가능하다. 또한, 내가 원하는 바를 명확하게 질문해야 정확한 답변을 받을 수 있다.

IV. 결 론

본 논문에서는 인공지능 할루시네이션 문제를 해결하기 위해 프롬프트 질의 개선 방법을 제안했다. 프롬프트를 명확하게 주는 간단한 방법을 통해 인공지능 대화 시스템이 정확한 답변을 할 수 있도록 유도할 수 있다. 이를 통해 일반적인 사용자도 인공지능 대화 시스템의 할루시네이션 문제를 간단하게 회피할 수 있을 것으로 기대한다. 추후 연구 과제로는 할루시네이션 패턴을 분석하여 이를 해결할 수 있는 질의 프롬프트 구조를 설계할 예정이다.

참 고 문 헌

[1] 이세훈, 최정빈, 백영태, 윤선호. (2023.07).

대규모 언어 모델(LLM) 기반의 파이썬 입문자를 위한 코딩 도우미. 한국컴퓨터정보학회 학술발표논문집

- [2] 문창현, 권동현. (2023). 이미지 생성형 AI에 대한 사용자 수준의 이해와 애니메이션 제작 교육 적용 방안 연구. 만화애니메이션 연구, 213-262, 10.7230/KOSCAS.2023.72.213
- [3] S. Jha, S. K. Jha, P. Lincoln, N. D. Bastian, A. Velasquez and S. Neema, "Dehallucinating Large Language Models Using Formal Methods Guided Iterative Prompting," 2023 IEEE International Conference on Assured Autonomy (ICAA), Laurel, MD, USA, 2023, pp. 149-152, doi: 10.1109/ICAA58325.2023.00029.
- [4] 김미정, "구글 CEO "AI 발전하려면 '환각' 현상 극복 필수"", zdnet korea, 2023년 04월 18일자.

두 번째 로짓을 활용한 머신러닝

¹김효서, ^{2*}최준석

Second Logit Unlearning

¹Hyoseo Kim, ^{2*}Junsuk Choe

요 약

디지털 기술의 발전과 인터넷의 보급으로 인해 "잊혀질 권리"의 중요성이 부각되면서 머신러닝 개념이 더욱 중요해지고 있다. 머신러닝은 머신러닝의 하위 분야로서, 훈련된 모델에서 특정 부분 집합인 "잊혀질 집합"의 영향을 제거하려는 것을 목표로 하는 분야이다. 최근 연구에서는 DNN 모델의 매개변수를 조절하는 방식으로 머신 러닝을 수행한다. 그러나 이 방식은 매개변수 공간의 큰 차원 때문에 연산 비용이 지나치게 막대하여 계산하기 어렵다는 단점이 있다. 따라서 본 논문에서는 잊고자 하는 데이터를 다른 클래스에 레이블링하여 잊고자 하는 클래스를 빠르고 효과적으로 잊어버리는 방법으로 Second Logit Unlearning을 제안한다. 우리는 CIFAR-10 데이터셋에서 실험을 진행하였으며 결과는 Second Logit Unlearning이 이미지 분류 문제에서 망각하고자 하는 클래스를 효과적으로 잊어버릴 수 있음을 보여준다.

Key words

Machine Learning, Machine Unlearning, Deep learning, Classification

I. 서 론

개인정보를 보호하는 의식이 높아지면서 다양한 개인정보 보호 규정이 시행되고 있다. 예를 들어, 유럽 연합에서 시행한 일반 개인정보 보호 규정(GDPR)은 개인에게 "잊혀질 권리"를 부여하며, 정보주체가 요청하면 기업은 해당 개인 데이터를 지우도록 의무화하고 있다[1].

머신러닝 분야에서는 특정 데이터를 지우는 일이 개인정보 보호뿐만 아니라, 훈련된 모델에서 부정확하거나 오래된 정보

혹은 조작된 데이터를 삭제하기 위한 필수 작업으로 간주된다. 이러한 필요성에서 대두된 것이 머신러닝이라는 개념이다. 이는 머신러닝 모델이 학습 데이터의 일부를 잊어버리게 하는 기능을 제공한다.

DNNs에 대한 기존의 언러닝 방법은 주로 두 가지로 나뉜다: 재훈련 가속화와 매개변수 업데이트. 재훈련 가속화 방식[2][3]은 모델을 처음부터 다시 훈련하는 방식으로 재훈련 과정을 가속화하는데 중점을 두고 있다. 이 방식은 단순하고 확실하지만 비용이 지나치게 많이 발생하는 문제가

¹ 서강대학교, 석사과정 (hyoseo@sogang.ac.kr)

² 교신저자 서강대학교, 교수 (jschoe@sogang.ac.kr)

있다.

매개변수 업데이트 방식은 DNN 모델을 특정 데이터를 기반으로 매개변수를 업데이트하여 언러닝하려는 시도이다. 예를 들어 Fisher Information Matrix (FIM)[4]은 매개변수 공간에서 잊혀진 데이터의 영향을 찾는 데 사용된다. 그러나 매개변수 공간의 큰 차원 때문에 계산 비용이 크다는 한계를 지닌다.

본 논문에서는 재레이블링 방식을 활용하여 언러닝을 수행한다. 기존에도 nearest but incorrect class 로 재레이블링하는 방식[5]이 존재했으나, 해당 방법은 neighbor searching 하는 데 소요되는 시간이 많다는 단점이 있다. 반면에 본 논문에서는 second logit을 활용하여 클래스를 재레이블링하는 방식으로 시간과 비용이 적게 든다. 또한 CIFAR-10에서의 실험 결과, second logit 방식이 잊고자 하는 클래스를 더 효과적으로 잊어버리면서 잊지 말아야 할 나머지 클래스들은 더욱 잘 보존하는 것으로 나타났다.

II. 본 론

2.1 Second Logit Unlearning

본 논문에서는 특정 데이터를 잊어버리기 위해 해당 데이터에 틀린 레이블을 다시 매핑한다. 이때 어떤 틀린 레이블을 선택할지에는 여러 가지 방법이 있다. 가장 직관적인 방법은 랜덤하게 레이블을 부여하는 것이다. 그러나 랜덤하게 레이블을 부여하는 방법은 잊지 말아야 할 다른 클래스의 성능에 부정적인 영향을 미칠 수 있다. 본 논문에서 다루는 머신언러닝의 목표는 특정 클래스를 잊어버릴 때 해당 클래스의 성능이 감소하는 것뿐만 아니라,

나머지 잊지 말아야 할 클래스의 성능은 언러닝을 수행하기 전보다 높아지거나 최소한 낮아지지 않아야 한다는 점이다.

따라서 본 논문에서는 잊고자 하는 클래스의 각 데이터의 logit을 계산하고, 두 번째로 높은 logit에 해당하는 클래스로 매핑하는 방식을 제안한다. 이 방식은 잊고자 하는 데이터를 무작위로 선택된 클래스로 매핑하는 방식과 달리 결정 공간에서 가까울 것으로 예상되는, logit 값이 두 번째로 높은 클래스로 매핑한다.

이를 통해 나머지 잊지 말아야 할 클래스들의 결정 경계를 최대한 유지하면서도 잊어야 할 클래스들의 결정 경계를 움직일 수 있다. 이러한 방식으로 언러닝의 과정에서 원치 않는 영향을 최소화하면서도 잊고자 하는 클래스의 영향을 효과적으로 감소시킬 수 있다.

한편, Boundary Shift[5]는 결정 공간 상에서 가장 가까우나 틀린 클래스로 매핑하는 방식이다. 이 방식 또한 잊지 말아야 할 클래스들의 결정 경계에 미치는 영향을 최소화하는데 집중하지만, 가장 가까운 틀린 클래스를 찾는 과정에서 일부 시간이 소요될 수 있다. 반면에 본 논문에서 제안한 logit을 활용하는 방식은 두 번째로 큰 logit의 클래스를 뽑아내기만 하면 되기 때문에 틀린 클래스를 찾는 데에 거의 소요시간이 없다.

2.2 Performance Evaluation

본 논문에서는 CIFAR-10 데이터셋에 대해, 고정된 학습률 0.1, 모멘텀 0.9 및 배치 크기 64를 사용하여 SGD로 30 epoch 동안 All-CNN 모델을 처음부터 훈련한다. Second logit unlearning의 fine-tune 과정에서는 학습률을 10⁻⁵로 설정하여 10 epoch 동안 진행한다.

표 1. baseline들과 second logit 성능 비교

lr	random	boundary shift	logit second best
ACC on Dr	98.82	99.00	99.77
ACC on Df	8.06	8.16	5.16
ACC on Drt	80.77	81.76	83.22
ACC on Dft	6.20	6.40	4.10

표 1은 앞서 언급한 baseline들과 본 논문에서 제안하는 logit second 방식의 성능을 비교한 결과를 나타내고 있다. 여기서 Dr은 유지되어야 하는 클래스들의 훈련 데이터를 나타내며, Df는 잊어야 할 클래스의 훈련 데이터를 나타낸다. Drt는 유지되어야 할 클래스들의 실험 데이터를, 마지막으로 Dft는 잊어야 할 클래스의 실험 데이터를 나타낸다.

표 1을 통해 second logit best 방식이 유지되어야 할 클래스들의 성능인 Dr과 Drt에서 가장 높은 결과를 보이고, 잊어야 할 클래스의 성능인 Df와 Dft에서는 가장 낮은 결과를 나타내고 있음을 확인할 수 있다. 이는 second logit 방식이 유지되어야 할 클래스에 대한 영향을 최소화하면서 잊어야 할 클래스를 효과적으로 언러닝하는데 성공했음을 시사한다.

III. 결 론

본 논문에서는 결정 경계를 이동시켜 훈련된 DNN에서 잊고자 하는 클래스의 정보를 제거하는 데에 second logit을 제안한다. 이 방법은 재레이블링 방식을 활용하여 결정 경계를 효과적으로 이동시키는데 이 때 걸리는 시간이 매우 적다는 장점이 있다. 실험 결과는 second logit이 언러닝을 빠르고 효과적으로 수행한다는 것을 입증한다. 이로써 우리의 연구는 결정 경계 이동을 통한 언러닝이

효율적이며 시간적 비용이 낮다는 측면에서 실용적인 가치를 지닌다는 점을 강조한다. 앞으로의 연구에서는 결정 경계와 언러닝 간의 관계를 더 깊이 탐구하여 머신 언러닝 분야에 새로운 인사이트를 제공할 수 있을 것으로 기대된다.

감 사 의 글

이 논문은 2023년도 정부(산업통상자원부)의 재원으로 한국산업기술진흥원의 지원을 받아 수행된 연구임(P0020535, 2023년 산업혁신인재성장 지원사업)

참 고 문 헌

- [1] Paul Voigt and Axel Bussche, The EU General Data Protection Regulation (GDPR): A Practical Guide. Springer, 2017.
- [2] M. Muller, Computer Go, Artificial Intelligence, Vol. 134, Issues 1-2, 2002, pages 145-179. Lucas Bourtole, Varun Chandrasekaran, Christopher A Choquette-Choo, Hengrui Jia, Adelin Travers, Baiwu Zhang, David Lie, and Nicolas Papernot. Machine unlearning. In Proceedings of IEEE S&P, pages 141-159. IEEE, 2021.
- [3] Laura Graves, Vineel Nagisetty, and Vijay Ganesh. Amnesiac machine learning. In Proceedings of AAAI, volume 35, pages 11516-11524, 2021.
- [4] Aditya Golatkar, Alessandro Achille, and Stefano Soatto. Eternal sunshine of the spotless net: Selective forgetting in deep networks. In Proceedings of IEEE/CVF CVPR, pages 9304-9312, 2020.
- [5] M. Chen, W. Gao, G. Liu, K. Peng, and C.

Wang, “Boundary unlearning,” arXiv preprint arXiv:2303.11570, 2023.

다수 자율주행차량의 이동 경로를 고려한 작업 할당 기술

¹이세빈, ^{2*} 남창주

Task Allocation of Multiple Autonomous Vehicles considering Practical Navigation Costs

¹Seabin Lee and ^{2*}Changjoo Nam

요 약

다수의 자율주행 차량이 승객이송, 화물배송과 같은 작업을 최소의 비용을 통해 수행하기 위해서는 효율적인 배차 방법이 필요하다. 단순히 거리 상 가까운 작업을 순차적으로 할당하는 방법보다는 전체 차량과 작업을 고려한 조합 최적화 기법이 유리하다. 할당의 최적화 목표인 작업 수행비용은 차량과 작업 두 지점 사이의 직선거리를 기준으로 계산하는 것 보다 두 지점 사이 주행 경로를 기준으로 하는 것이 보다 현실적이다. 본 논문에서는 주어진 환경을 2D로 투영하여 차량이 통행할 수 있는 도로를 파악하고, 파악된 도로를 기반으로 로드맵을 구성한다. 이 로드맵으로부터 각 차량이 각 작업을 수행하는데 소요되는 비용을 산정하고 그를 바탕으로 작업 할당을 최적화하는 방법을 제안한다. 이렇게 생성된 작업 할당 및 이동 경로를 이용하면 차량이 장애물과 충돌하지 않고 작업을 수행할 수 있고 전체 차량이 작업을 수행하는 비용을 최소화하여 작업 수행의 안전성, 효율성을 높일 수 있다.

Key words

Multi-robot System, Multi-robot Coordination, Navigation

I. 서 론

현대 사회에서 자율주행 차량의 중요성이 날로 증가함에 따라, 이들 차량의 효율적인 경로 계획과 작업 할당은 매우 중요한 연구 분야로 자리잡고 있다. 특히, 도심과 같이 복잡하고 동적인 환경에서의 자율주행 차량 운용은 교통 체증 감소, 사고 위험 감소, 배송 및 운송 효율성 증가 등 다양한 사회적,

경제적 혜택을 가져올 수 있다. 본 연구는 이러한 배경 하에, 자율주행 차량들이 장애물을 피하며 효율적으로 목적지에 도달할 수 있는 작업을 할당하는 방법을 제시한다.

주어진 환경 내에 존재하는 장애물을 고려한 로드맵을 구축하기 위해 보로노이 다이어그램[1]을 이용한다. 이렇게 만들어진 로드맵은 그래프의 형태로 나타나며, 차량과 목적지들을 로드맵 위의

¹ 서강대학교 전자공학과, 석사과정 (123sebin@naver.com)

² 교신저자 서강대학교 전자공학과, 교수 (cinam@sogang.ac.kr)

노드에 결부시켜 그래프 상에서 각각의 차량에 대한 목적지를 할당한다. 이 방법은 차량 간의 충돌 가능성을 줄이고, 경로 효율성을 극대화하는데 중점을 둔다.

II. 다수 차량의 작업 할당

2.1 로드맵 구축

장애물들의 가장자리를 세밀하게 샘플링하여 일반화된 보로노이 다이어그램 (Generalized Voronoi Diagram, GVD) 기반의 로드맵을 생성한다. 샘플링된 지점을 기준으로 로드맵의 노드들이 불균일하게 생성되기 때문에 균일화를 위한 후처리 과정이 뒤따른다(Fig. 1). 환경 내 존재하는 차량 및 작업(목적지)들과 가장 가까운 로드맵 상 노드를 찾는 과정을 거쳐 차량과 작업들이 로드맵에 효과적으로 통합된다.

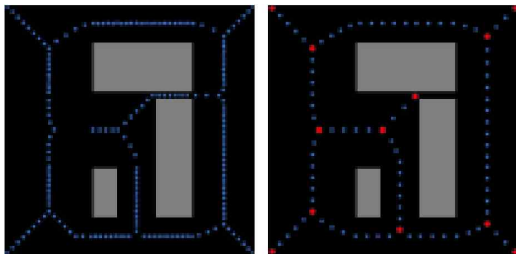


Fig. 1 전처리 전(좌) 후(우)의 로드맵 예시

2.2 다수 자율주행 차량의 작업 할당

본 연구의 핵심인 로드맵 기반의 작업 할당은 조합 최적화(combinatorial optimization)기법을 이용한다. 즉, 각각의 차량이 작업을 수행하는 비용(cost)의 합을 최소화하는 목표를 갖는다. 최적화를 위해서는 모든 차량과 모든 작업 사이의 비용이 필요하다. 본 연구에서는 비용을 이동 거리로 정의하는데 이전에 생성한 로드맵을 통해 쉽게 계산할 수 있다. 로드맵 상에 존재하는 모든 차량과 목적지 사이의

최단 경로를 A* 알고리즘 등과 같은 최단 경로 알고리즘을 통해 계산한다

계산된 비용을 바탕으로 헝가리안 알고리즘(Hungarian method)[2]을 이용해 최적화된 할당을 찾는다. 제안한 방법을 통해 차량들은 로드맵을 따라 이동할 수 있는 최적의 경로를 얻을 수 있고 전체 시스템은 차량들이 작업을 수행하기 위해 이동하는데 소요되는 비용을 최소화할 수 있다.

III. 실험 및 결론

두 지점 사이의 유클리디언 거리와 로드맵 상 거리를 이용한 작업 할당 방법을 비교하였다. Table 1에 요약된 실험 결과와 같이 로드맵 기반 방법은 유클리디언 거리 방식에 비해 최대 12.5% 시간 비용을 절감하는 효율성을 보였다. 이는 로드맵 기반 접근법이 실제 차량 이동의 효율성을 증대시킨다는 것을 시사한다.

장애물이 존재하는 환경에서는 로드맵 기반 거리를 사용하는 제안 방법이 유클리디언 거리를 이용하는 것에 비해 자율주행 차량의 작업 할당에 있어 효율적임을 확인 하였다.

Table. 1 비용 산정 방법에 따른 각 차량별 목적지에 도달하는데 소요된 시간의 합

차량의 대수	10	20	30	40	50
유클리디언 거리 이용	20.5	31.8	48.9	63.4	69.7
로드맵 위의 거리 이용	19.5	30.2	45.2	55.4	61.0

감사의 글

이 논문은 2023년도 정부 (산업통상자원부)의 재원으로 한국산업 기술진흥원의 지원을 받아 수행된 연구임 (P0020535, 2023년 산업혁신인재성장 지원사업)

참 고 문 헌

- [1] J. Yu and S. LaValle, "Structure and intractability of optimal multi-robot path planning on graphs," in Proceedings of the AAAI Conference on Artificial Intelligence, vol. 27, no. 1, 2013, pp. 1443-1449.
- [2] H. W. Kuhn, "The hungarian method for the assignment problem," Naval research logistics quarterly, vol. 2, no. 1-2, pp. 83-97, 1955.

사이버 킬 체인 기반 악성코드 공격 사례 분석

¹전규현, ²전승호, ^{3*}서정택

Analysis of Malware Attack cases based on Cyber Kill Chain

¹GyuHyun Jeon, ²Seungho Jeon, and ^{3*}Jung Taek Seo

요 약

사이버 보안은 지속적으로 진화하는 보안 도구와 기술로 인해 강화되고 있지만, 시스템 내에 단 하나의 취약점이 존재해도 공격자는 악성코드 등의 공격을 통해 침입할 수 있다. 따라서 시스템의 취약점을 발견하고 보완하며, 외부 접점에서 발생하는 공격에 대한 보호가 필요하다. 또한, 악성코드의 공격 과정을 식별하여 적합한 보안 대책을 수립하는 것이 중요하다. 이에, 본 논문에서는 사이버 킬 체인(Cyber Kill Chain, CKC) 프레임워크를 사용하여 악성코드 공격 사례 분석을 진행하고, 도출된 위협 요소에 대한 대응방안을 제시하였다.

Key words

Cyber Kill Chain, Malware, Case Analysis, Framework, Threat

I. 서 론

계속 진화하는 사이버 공격을 방어하기 위해 보안 수준은 지속적으로 높아지고 있다. 보안과 관련된 도구, 기술, 표준 등의 발전을 통해 보안 메커니즘이 정교해지면서 다양한 공격 위협으로부터 인프라를 보호할 수 있다. 또한, 악성코드를 이용한 공격으로부터 시스템을 보호하기 위한 방어 기법 역시 고도화되고 있다[1].

하지만, 공격자는 시스템 내 취약점이 개수와 상관없이 존재하기만 한다면, 인프라

내 시스템에 침입하여 악성코드 공격을 수행할 수 있다[2].

따라서, 시스템 내 취약점을 해결하기 위한 방법 및 시스템과 외부 접점에서 발생하는 공격에 대한 보호가 필요하다. 또한, 악성코드 공격 과정의 초기 감염, 전파 등 작동 방식을 식별하여 각 위협 요소에 대한 적합한 대응 방안을 도출할 수 있어야 한다.

이에, 본 논문에서는 사이버 킬 체인(Cyber Kill Chain, CKC) 프레임워크를 사용하여 악성코드 공격 사례 분석을 진행한다. 이후, 분석된 위협 요소에

¹ 가천대학교 정보보호학과 석사과정 (pengchan88@gachon.ac.kr)

² 가천대학교 컴퓨터공학부 스마트보안전공 연구교수 (shjeon90@gachon.ac.kr)

³ 교신저자 가천대학교 컴퓨터공학부 컴퓨터공학전공 교수 (seojt@gachon.ac.kr)

대해 적절한 대응 방안을 도출한다. 2장에서는 사이버 킬 체인 개요 및 특징, 3장에서는 악성코드 공격 사례 분석, 4장에서는 사이버 킬 체인 적용 및 대응 방안, 5장에서는 한계점, 6장에서는 결론 및 향후 연구방향에 대해 기술하였다.

II. 사이버 킬 체인

사이버 킬 체인에 대한 개념은 2009년 미국 Lockheed Martin社에서 제안하였으며, 다양한 유형의 위협, 지능형 지속 공격(Advanced Persistent Threat, APT) 및 악성코드를 식별하기 위한 인텔리전스 기반 위협 중심 접근방식의 프레임워크이다[3]. 총 7단계로 진행되며, 각 단계는 순서대로 Reconnaissance (정찰), Weaponization(무기화), Delivery (유포), Exploitation(악용), Installation (설치), Command and Control(명령 및 제어), Action on Objectives(목적 달성)로 구성된다. [표 1]은 사이버 킬 체인의 공격 진행 단계를 나타낸 것이다.

사이버 킬 체인의 특징은 각 공격 단계 중, 한 단계라도 기능을 수행하지 못하면 공격이 중단되거나 공격자가 의도한 목적을 수행할 수 없다. 따라서 방어자가 사이버 공격을 사이버 킬 체인을 기반으로 분석하여 각 단계별 위협을 식별한다면, 공격 실행 여부 및 공격이 실행될 단계를 알 수 있다. 이후, 적합한 대응 계획을 수립할 수 있다.

표 1. 사이버 킬 체인 공격 진행 단계

단계	설명
Reconnaissance (정찰)	<ul style="list-style-type: none"> 공격 표적 및 전술 식별 악용 가능한 취약점 식별
Weaponization (무기화)	<ul style="list-style-type: none"> 정찰 단계에서 수집된 정보 기반 악성 코드, DDoS, 봇넷 등 사이버 무기 준비 사이버 무기는 제로데이 취약점 또는 여러 개의 취약점을 조합하여 악용 가능
Delivery (전달, 유포)	<ul style="list-style-type: none"> 사이버 무기를 표적에게 전송 이메일, 웹사이트, USB 등
Exploitation (악용, 침투)	<ul style="list-style-type: none"> 전송된 사이버 무기를 사용 높은 접근 권한을 얻기 위해 시스템, 네트워크 등의 취약점을 악용
Installation (설치)	<ul style="list-style-type: none"> 공격 표적 시스템에 악성 프로그램 설치 공격 표적에서 공격 가능한 환경 구축 공격 탐지 회피, 권한 우회 및 상승, 암호 해시 추출 등
Command and Control (명령 및 제어)	<ul style="list-style-type: none"> 원격 공격 운영 및 모니터링을 위한 명령 및 제어 채널 설정 통신 및 명령 전달
Action on Objectives (목적 달성)	<ul style="list-style-type: none"> 공격 최종 목적 달성을 위한 조치 수행 데이터 수정/수집/유출/파괴/암호화

III. 악성코드 공격 사례 분석

본 논문에서 분석한 악성코드 공격 사례는 총 3건이며, 최근 1년 이내에 발생한 악성코드 공격 사례를 대상으로 하였다.

3.1 Hakuna Matata

2023년 07월, 국내 기업을 대상으로 공격하는 Hakuna Matata 랜섬웨어가 처음 발견되었다. 일반적인 랜섬웨어의 기능과 공격 대상의 클립보드를 모니터링하여 암호화폐 지갑 주소와 관련된 문자열이

탐지될 시, 공격자의 주소로 변경하는 ClipBanker 기능을 동시에 수행한다[4].

먼저, 외부에 노출 및 RDP가 활성화된 시스템을 대상으로 무차별 대입공격(Brute Force)을 수행한 후, 계정 정보를 획득한다. 이후, 원격으로 시스템 로그인을 수행하여 공격 대상 시스템의 제어를 탈취한다. 공격자는 “C:\Temp\” 등의 경로에 NirSoft社에서 제작한 계정 정보 탈취 프로그램 및 추가 악성파일을 설치했다. 계정 정보는 “\M\!logs\” 경로에 텍스트 파일로 생성했을 것으로 추정되며, 추가 악성파일들 중, “ver7.exe”가 Hakuna Matata 랜섬웨어이다.

해당 파일을 실행할 시, “%LOCALAPPDATA%\rundll32.exe” 경로에 파일을 복사 및 실행하여 안티바이러스 파일 탐지를 우회한다. 그리고 파일 암호화를 수행하며 Run Key에 파일을 등록하여 시스템 재부팅 이후에도 실행하도록 설정한다.

3.2 Scattered Spider(ALPHV/Blackcat)

2023년 09월, 카지노 리조트 운영 회사인 MGM Resorts International社는 Scattered Spider(ALPHV/Blackcat) 랜섬웨어로 인해 회사 IT 시스템 및 고객 대상 서비스가 36시간 이상 중단 및 약 1억 달러의 손해가 발생했다[5].

공격자는 MCM社의 IT 관리자를 표적으로 지정하였다. 이후, 클릭 시, 그룹에서 SIM 교환을 수행할 수 있는 SMS 피싱 메시지가 직원에게 전송되었다. 그리고 교체된 SIM으로 전화 통신을 캡처하여 대상 모바일 장치의 전화번호를 탈취하여 공격자에게 전송한다. 또한, IT 헬프데스크(Support)에 유선으로 연락하여 MFA 재설정 코드를 보내기 위한 세부 정보를 제공하고 클라우드 환경에 대한 액세스 권한을 얻었다. 그 후, MGM社가

Okta IAM 동기화 서버를 오프라인으로 전환한 후에도 네트워크에 액세스할 수 있었다. 그리고 도메인 컨트롤러(Domain Controller)에서 해시 덤프를 훔쳤다. 이를 통해, 공격자는 Okta Agent 서버에 액세스한 후, 해당 환경에서 비밀번호를 스니핑하였다. 이후, 타사 원격 제어 애플리케이션이나 Cobalt Strike, Brute Ratel 등과 같은 통째 백도어 설치 및 클라우드 서버를 통한 통신이 가능했을 것이라 추정했다. 해당 랜섬웨어로 인해 약 100대의 ESXi 하이퍼바이저 암호화 및 개인 식별 정보가 유출되었다.

3.3 DangerAds & AtlasAgent

2023년 09월, 피싱 문서를 기반으로 DangerAds 및 AtlasAgent 2개의 악성코드를 사용하는 APT 그룹 AtlasCross가 처음 발견되었다. 해당 공격은 미국 적십자와 관련된 인원을 표적으로 한 사이버 공격으로 추측된다[6]. 공격자는 “Blood Drive September 2023.docm”이라는 제목의 MS Word 문서를 작성했다. 문서 실행 시, 단어 편집 기능을 활성화 할 수 있는 악성 매크로가 활성화된다. 매크로 실행 시, “%APPDATA%\Microsoft\Word\“ 하위 경로의 임의의 숫자 이름을 가진 폴더에 “KB4495667.zip” 파일을 저장 및 “KB4495667.pkg“ 파일을 추출한다. 이후, 추출된 파일을 3일 동안 실행하도록 “Microsoft Office Update”라는 예약된 작업을 설정한다. 그리고 공격 대상 기본 정보를 업로드 하기 위해 AtlasCross 그룹이 관리하는 웹사이트와 통신을 시작한다.”KB4495667.pkg“ 파일이 DangerADs 악성코드(로더 트로이 목마)이며, 호스트 환경을 탐지 및 내장 셸코드를 실행한다. DangerADs는 최종 페이로드인 AtlasAgent 악성코드(x86 및

x64 버전의 DLL 트로이목마)를 메모리에 로드한다. AtlasAgent의 주요 기능은 호스트 정보 획득, 셸코드 실행, C2서버 간 통신, 파일 다운로드 및 실행이다. 해당 악성코드로 인해 데이터 유출, 시스템 손상, 개인정보 침해 등의 피해가 발생할 수 있다.

요소에 대해 적합한 대응방안을 도출한 것이다[7].

사이버 킬 체인 적용을 통해 사이버 공격의 전반적인 흐름 및 공격 목적을 파악할 수 있었지만, 단일 위협 요소가 여러 단계에 존재할 경우, 각 단계별 정확한 분류가 어려웠다. 특정 단계의 취약점을 사용하여 방어 기법이 적용된 단계를 우회할 가능성 역시 존재한다. 또한, 사이버 킬 체인에서는 공격 전술, 기법, 절차(Tactics, Techniques, and Procedures, TTP)에 대해 자세하게 알 수 없다[8].

IV. 사이버 킬 체인 적용 및 한계점

[표 2]는 분석한 공격 사례들을 사이버 킬 체인을 기반으로 위협 요소를 단계별 분류한 것이다. [표 3]은 [표 2]에서 분류한 위협

표 2. 공격 사례 사이버 킬 체인 분석

단계	공격 사례별 위협 요소		
	Hakuna Matata	Scattered Spider (ALPHV/Blackcat)	DangerAds & AtlasAgent
Reconnaissance	<ul style="list-style-type: none"> 외부 노출 및 RDP 활성화된 시스템 정보 수집 	<ul style="list-style-type: none"> IT 관리자 정보 수집 	<ul style="list-style-type: none"> 미국 적십자와 관련된 인원 정보 수집
Weaponization	<ul style="list-style-type: none"> 취약한 계정 정보 획득을 위한 무차별 대입 공격 도구 제작 	<ul style="list-style-type: none"> 클릭 시, SIM Swapping을 수행하는 SMS 피싱 메시지 제작 	<ul style="list-style-type: none"> 악성 매크로가 포함된 MS Word 문서 제작
Delivery	<ul style="list-style-type: none"> 제작한 도구를 사용하여 시스템 접근 후, 계정 정보 획득 시도 	<ul style="list-style-type: none"> 제작한 피싱 메시지를 직원에게 전송하여 클릭을 유도 	<ul style="list-style-type: none"> 제작한 문서를 피싱 메일에 첨부하여 표적 인원에게 전달 후, 매크로 활성화 유도
Exploitation	<ul style="list-style-type: none"> 원격으로 탈취한 계정 정보를 사용하여 시스템 로그인 및 시스템 제어 권한 탈취 	<ul style="list-style-type: none"> MFA의 재설정 코드를 얻어내 클라우드 환경에 액세스 	<ul style="list-style-type: none"> 문서 열람 시, 설정된 악성 행위 실행 DangerAds 악성코드 생성
Installation	<ul style="list-style-type: none"> NirSoft社의 계정 정보 탈취 프로그램 및 추가 악성 파일을 시스템에 설치 	<ul style="list-style-type: none"> 타사 원격 제어 애플리케이션, Cobalt Strike, Brute Ratel 등 도구를 사용하여 백도어 설치 	<ul style="list-style-type: none"> AtlasAgent 악성코드가 포함된 페이로드 실행
Command and Control	-	<ul style="list-style-type: none"> C2 서버(클라우드 서버)간 통신 및 시스템 제어 	<ul style="list-style-type: none"> AtlasCross 그룹이 관리하는 C2 서버와 통신 호스트 기본 정보가 업로드 및 추가 명령을 수신
Action on Objectives	<ul style="list-style-type: none"> 파일 암호화 암호화폐 지갑 주소 변경 	<ul style="list-style-type: none"> ESXi 하이퍼바이저 암호화 개인 식별 정보 유출 	<ul style="list-style-type: none"> 데이터 유출, 시스템 손상 및 개인정보 침해

표 3. 사이버 킬 체인 기반 공격 사례 분석에 대한 대응 방안

단계	공격 사례별 대응 방안		
	Hakuna Matata	Scattered Spider (ALPHV/Blackcat)	DangerAds & AtlasAgent
Reconnaissance	<ul style="list-style-type: none"> RDP 비활성화 네트워크 수준 인증(NLA) 활성화를 통한 추가 인증 	<ul style="list-style-type: none"> 개인 정보 관리 강화 	
Weaponization	-	-	-
Delivery	<ul style="list-style-type: none"> 비정상적 로그인 시도 차단 MFA 사용 네트워크 분리를 통한 중요 시스템에 대한 접근 차단 	<ul style="list-style-type: none"> 내부 보안 프로세스 강화 스팸 메일 및 메시지 분석 및 차단 솔루션 적용 	<ul style="list-style-type: none"> 신뢰할 수 없는 문서 실행 금지 MS Office 문서 매크로 자동 실행 비활성화
Exploitation	<ul style="list-style-type: none"> 비정상적인 계정 접근 및 행동 패턴 탐지 	<ul style="list-style-type: none"> MFA 재설정 프로세스 강화 	<ul style="list-style-type: none"> EDR 솔루션 적용을 통한 엔드포인트 보호
Installation	<ul style="list-style-type: none"> 승인되지 않은 도구의 사용을 차단 관리자 권한이 필요한 작업 모니터링 안티바이러스 및 안티멀웨어 사용 IDS 사용 취약점 패치 		
Command and Control	<ul style="list-style-type: none"> 네트워크 트래픽 모니터링 및 분석 네트워크 세분화 및 접근 제어 SSL/TLS 검사 등 암호화된 트래픽 분석 DNS 필터링 		
Action on Objectives	<ul style="list-style-type: none"> DLP(Data Loss Prevention) 솔루션 적용을 통한 데이터 손실 및 유출 방지 네트워크 접근 제어 강화 		

V. 결론 및 향후 연구방향

본 논문에서는 사이버 킬 체인 기반 악성코드 공격 사례를 분석을 수행하였다. 그 결과, 최근 1년 이내의 악성코드 공격 사례 3건을 확인하였으며, 각 공격 사례를 분석하였다. Hakuna Matata, Scattered Spider(ALPHV/Blackcat), DangerAds & AtlasAgent 공격 사례를 분석한 후, 사이버 킬 체인을 적용하여 각 단계별 위협 요소를 분류하였다. 그리고 위협 요소별 대응 방안을 도출하였다.

향후 연구로는 사이버 킬 체인 이외의 프레임워크들을 조사하고, 공격 사례에 해당 프레임워크들을 적용하여 도출된 결과를 기반으로 비교 분석 연구를 진행할 것이다.

Acknowledgement

이 논문은 2023년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구임 (No.2021-0-00493, 5G Massive 차세대 사이버공격 기만기술 개발)

참고 문헌

- [1] MIRZA, Qublai K. Ali, et al. Ransomware analysis using cyber kill chain. In: 2021 8th International Conference on Future Internet of Things and Cloud (FiCloud). IEEE, 2021. p. 58-65
- [2] NAMANYA, Anitta Patience, et al. Detection

- of malicious portable executables using evidence combinational theory with fuzzy hashing. In: 2016 IEEE 4th International Conference on Future Internet of Things and Cloud (FiCloud). IEEE, 2016. p. 91-98.
- [3] LOCKHEED MARTIN, Cyber Kill Chain, <https://www.lockheedmartin.com/en-us/capabilities/cyber/cyber-kill-chain.html>, 2014 [last access 2023/11/20]
- [4] ASEC, Hakuna Matata Ransomware Targeting Korean Companies, <https://asec.ahnlab.com/en/56010/>, 2023.08 [last access 2023/11/20]
- [5] MORPHISEC, Threat Analysis: MGM Resorts International ALPHV/Blackcat/Scattered Spider Ransomware Attack, <https://blog.morphisec.com/mgm-resorts-alphv-spider-ransomware-attack>, 2023.09 [last access 2023/11/20]
- [6] NSFOCUS, Warning: Newly Discovered APT Attacker AtlasCross Exploits Red Cross Blood Drive Phishing for Cyberattack, <https://nsfocusglobal.com/warning-newly-discovered-apt-attacker-atlascross-exploits-red-cross-blood-drive-phishing-for-cyberattack/>, 2023.09 [last access 2023/11/20]
- [7] DARKREADING, Are You Listening to Your Kill Chain?, <https://www.darkreading.com/vulnerabilities-threats/are-you-listening-to-your-kill-chain->, 2019.01 [last access 2023/11/20]
- [8] NAIK, Nitin, et al. An evaluation of potential attack surfaces based on attack tree modelling and risk matrix applied to self-sovereign identity. *Computers & Security*, 2022, 120: 102808.

Teachable Machine을 이용한 가죽 결함 검출 방법 연구

¹권소영, ²이종윤, ^{3*}이용환

Research on Methods for Detecting Defects in Leather using a Teachable Machine

¹So-Young Kwon, ²Jong-Youn Lee and ^{3*}Yong-Hwan Lee

요 약

현대 가죽 산업은 시장 경쟁력을 갖추기 위해 고품질 가죽 제품을 생산하는데 주력하고 있다. 소재 취급, 염색 등 제조 공정의 여러 단계에서 다양한 가죽 결함이 발생하고 있다. 가죽 결함 검사는 수동 검사를 통해 이루어지며 주관적이기 때문에 일관성이 없다. 본 논문에서는 Teachable Machine을 이용한 가죽 결함 자동 검사 비전 시스템을 제안한다. 모델 학습은 접힘 자국, growth marks, 가죽 벗겨짐, 핀홀, 정상 가죽 이미지를 데이터 셋으로 사용하여 학습하고, 카메라로 촬영한 실시간 영상을 통해 가죽 결함을 검출하도록 한다.

Key words

Teachable Machine, Leather, Defect, Machine Learning, Classification

I. 서 론

가죽은 옷, 신발, 가방 등에 주로 사용되는 소재이다[1]. 가죽 조각의 표면은 가죽의 등급을 결정하고 판매 가격에 영향을 미치는 중요한 지표이며 최근 가죽 산업에서 시장 경쟁력을 갖추기 위해 고품질 가죽 제품 생산의 중요성이 커지고 있다[2]. 가죽 제조 공정의 여러 단계에서 가죽 접힘, 벗겨짐 등 다양한 결함들이 발생하며 현재까지도 다양한 산업체의 가죽 제조 공정에서 가죽

결함을 수동으로 검사하는 방식을 사용하고 있다[3]. 이러한 수동 검사는 검사관에 따라 주관적인 특성이 있고 비효율적이며 일관성이 없다[4]. 이러한 문제점을 해결하기 위해 결함 검출을 위한 인력 소모를 줄일 수 있고 결함 검출 기준이 일관적인 가죽 결함 자동 검사 비전 시스템이 필요하다.

본 논문에서는 가죽 결함 검출을 위한 인력 소모와 일관성이 없는 문제점들을 해결하기 위하여 teachable machine을 사용하여 가죽 결함을 검출하고 결함을 접힘, 벗겨짐

¹ 금오공과대학교, 박사과정 (papaya4040@naver.com)

² 금오공과대학교, 학사과정 (lilyun0807@naver.com)

³ 교신저자 금오공과대학교, 교수 (yhlee@kumoh.ac.kr)

등 다양한 종류로 분류 시스템을 제안한다.

II. 관련 연구

Teachable Machine은 구글에서 만든 머신러닝 모델을 빠르고 간단하며 쉽게 만들 수 있도록 제작된 웹 기반 노코드 인공지능 학습 툴이다[5]. Teachable Machine은 배터리 성능, 컴퓨터 성능이 중요하게 생각되는 지점에서 사용될 목적으로 설계된 CNN(Convolutional Neural Network) 구조인 MobileNet으로 사전 훈련 모델을 생성하고, 학습되어 있는 모델에 추가 데이터를 학습시켜 모델을 만드는 기술인 전이 학습을 진행하는 방법을 사용한다[6].

일상적인 생활이 어려운 특정 장애우들을 위한 맞춤 모델, 자폐아들을 위한 감정 학습 프로그램 등 다양한 곳에서 Teachable Machine이 활용되고 있다.

III. 가죽 결함 검출 시스템

본 논문의 시스템은 가죽 결함 검출 및 분류를 위해 Kaggle에서 3,600개의 이미지를 수집하였다. 학습, 검증, 테스트 과정에서 folding marks, grain off, growth marks, loose grains, pinhole, non defective의 여섯 개의 class를 가지는 데이터 셋을 사용한다. 가죽 결함 검출 및 분류 시스템의 전체적인 구성도는 그림1과 같다.

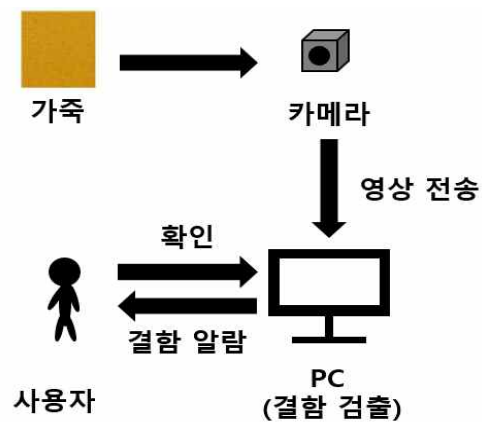


그림 1. 시스템 구성도

결함 검출 시스템은 가죽 결함 종류를 식별하고 결과를 화면에 출력한다. 결함이 없는 경우에는 non defective로 나타낸다.

그림 2는 가죽 결함 검출 시스템의 실행 화면이다. 영상의 밑에는 정상품, 결함의 종류들로 나열되어 있으며 영상에 들어온 이미지가 6가지의 클래스 중 어느 클래스에 가까운지 주황색으로 표시를 해준다.

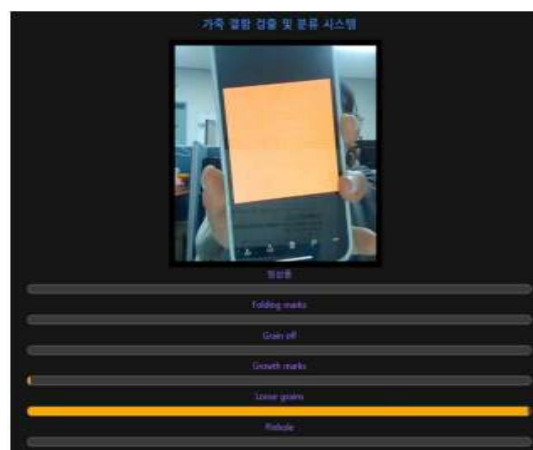


그림 2 가죽 결함 검출

IV. 결론

본 논문에서는 가죽 결함을 검출하고 결함을 분류하는 시스템을 제안하였다. 제안된 시스템은 웹 기반 인공지능 학습툴인

teachable machine을 사용하여 모델을 생성하였고 Goorm IDE를 활용하여 서버를 만들었으며 Netlify를 사용해 배포해 주었다. 이 시스템을 사용함으로써 다양한 가죽 산업체에서 실시간으로 가죽 결함을 검출 및 분류해 줄 수 있어 인력 소모를 줄여주고 육안 검사보다 더 정확하게 결함을 검출해 줄 것으로 기대된다. 향후 연구로 가죽 결함 검출 시 모든 가죽을 폐기할 수는 없으므로 가죽의 결함의 크기 및 심각도를 식별하는 방법에 대한 연구를 진행할 계획이다.

- Science, 2018.11, pages 300-304.
- [5] Teachable Machine, <https://teachablemachine.withgoogle.com/>.
- [6] A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto and H. Adam, MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications, Cornell University, 2017.04.

감사의 글

이 논문은 2023년도 정부 (산업통상자원부)의 재원으로 한국산업 기술진흥원의 지원을 받아 수행된 연구임. (P0017011, 2023년 산업혁신인재성장 지원사업)

참 고 문 헌

- [1] doopedia 두산백과, 가죽의 종류와 용도, <https://terms.naver.com/entry.naver?docId=1190134&cid=40942&categoryId=31891>.
- [2] 양나영, 황성걸, 소비자 관점에서의 가죽 인지 품질 연구, 조형미디어학, 제19권, 제2호, 2016.05, pages 157-166.
- [3] S. T. Lionga, Y. S. Ganb, Y. C. Huangb, C. A. Yuanc and H. C. Changa, Automatic Defect Segmentation on Leather with Deep Learning, Cornell University, 2019.03.
- [4] H. Q. Bong, Q. B. Truong, H. C. Nguyen and M. T. Nguyen, Vision-based Inspection System for Leather Surface Defect Detection and Classification, NAFOSTED Conference on Information and Computer

커넥티드 자율주행차의 합류구간 속도제어를 위한 지능형 운전자 모델 파라미터의 강화학습 기반 최적화

¹신성재, ^{2*}최경환

Reinforcement Learning-Based Optimization of Intelligent Driver Model Parameters for Speed Control of a Connected and Automated Vehicle at Merging Zones

¹Seong-Jae Shin, ^{2*}Kyung-Hwan Choi

요 약

교차로나 합류 구간과 같이 차량 간 상호작용이 필요한 도로에서는 차량 흐름이 느려지거나 사고로 이어질 수 있다. 본 연구에서는 다른 차량의 속도 및 위치 정보를 받아 올 수 있는 커넥티드 자율주행 차량의 합류 구간 속도제어 방법을 제시한다. 복잡한 차량 운행 환경 모델링에 의존하지 않는 강화학습과 차량 교통 흐름을 모델링하는데 사용되는 지능형 운전자 모델(Intelligent Driver Model, IDM)을 사용하여 차량의 속도를 제어함으로써, 충돌하지 않고 안전하게 합류 구간을 통과하도록 하는 알고리즘을 제안한다.

Key words

Connected and Automated Vehicle, Speed Control, Intelligent Driver Model, Reinforcement Learning, Merging Zone

I. 서 론

차량이 주행 중인 도로에 진입하는 합류 구간은 앞 차량의 움직임에 따라 연쇄적으로 차량이 정지하거나 충돌사고가 발생할 수 있다. 자율주행 기술의 발전을 통해 문제를 해결할 수 있을 것으로 기대하지만, 현재까지의 자율주행 기술은 기대에 미치지

못한다. 현재 시범 운행되고 있는 자율주행 차량들은 차량에 부착된 센서 정보만을 이용하여 주행하기 때문에, 차량 간 상호작용이 필요한 도로(교차로, 합류구간 등)에서 일반적인 운전자보다 방어적으로 주행하는 경향성을 보인다. 이는 곧 교통 흐름 저해로 이어질 가능성이 높다. 이러한 단점을 보완하기 위해 주변 차량과 V2X 통신을 활용해 정보를 주고받는 커넥티드

¹ 광주과학기술원 기계공학부, 석사과정 (ssjdooly@gm.gist.ac.kr)

² 교신저자 광주과학기술원 기계공학부, 조교수 (khchoi@gist.ac.kr)

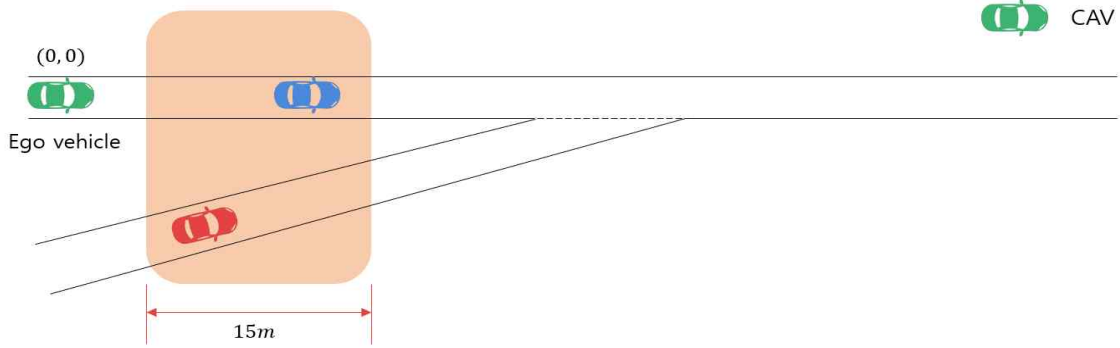


그림 1. 시뮬레이션 환경

자율주행차(Connected and Automated Vehicle, CAV)가 주목받고 있다.

본 논문에서는 복잡한 도로 환경 모델링 없이 솔루션이 도출 가능한 강화학습을 활용하여 합류 구간에서의 CAV 속도 제어 방법을 제안한다. 강화학습은 충분히 학습되더라도 충돌이 발생할 여지가 있기 때문에 차량 교통 흐름을 모델링한 지능형 운전자 모델(Intelligent Driver Model, IDM)을 활용한다. IDM의 파라미터를 강화학습으로 최적화하여 차량이 안전하고 빠르게 합류 구간을 통과하도록 한다. 제안한 방법은 MATLAB/Simulink를 활용하여 시뮬레이션 환경 구축 및 검증을 진행하였으며, 강화학습만 사용하여 훈련한 에이전트와 강화학습과 IDM을 사용하여 훈련시킨 에이전트를 비교한다.

이외의 파라미터는 표 1과 같으며 지능형 운전자 모델 (1)의 파라미터는 표 2와 같다.

$$\begin{aligned} \dot{x}_\alpha &= \frac{dx_\alpha}{dt} = v_\alpha \\ \dot{v}_\alpha &= \frac{dv_\alpha}{dt} = a \left(1 - \left(\frac{v_\alpha}{v_0} \right)^\delta - \left(\frac{s^*(v_\alpha, \Delta v_\alpha)}{s_\alpha} \right)^2 \right) \quad (1) \end{aligned}$$

where $s^*(v_\alpha, \Delta v_\alpha) = s_0 + v_\alpha \cdot T + \frac{v_\alpha \cdot \Delta v_\alpha}{2\sqrt{a \cdot b}}$

표 1. 시뮬레이션 파라미터

	v_{\min}	v_{\max}	a_{\min}	a_{\max}	$v_{0, \min}$	$v_{0, \max}$
RL	5	20	-3	3	-	-
RL + IDM	5	20	1	3	1	35

표 2. IDM 파라미터

모델 \ 파라미터	Delta	T	S_0	b
IDM	4	1.2	5	2

II. 시뮬레이션 모델

2.1 시뮬레이션 환경

시뮬레이션 환경은 그림 1과 같이 1차선 도로로 구성되며, 강화학습으로 제어하는 초록색 차량은 원점에 배치된다. 주변 차량은 다양성을 위해 주황색 영역 내에서 무작위로 배치된다. 초록색 차량은 초기속도 10m/s이며, 나머지 차량들의 속도는 8 ~ 21m/s 범위 내에서 무작위로 결정된다.

2.2 강화학습

강화학습은 에이전트가 현재 상태에서 어떤 행동을 하는 것이 최적인지를 학습하는 학습 방법으로, 이를 위해 상태(State), 행동(Action), 보상(Reward)을 정의해야 한다. 에이전트가 제어하는 차량을 커넥티드 자율주행 차량으로 가정하였기 때문에 자아 차량의 위치, 속도뿐만 아니라 다른 차량의 위치와 속도를 상태로 갖는다. 에이전트는

지능형 운전자 모델의 파라미터인 a 와 V_0 를 결정하며, 보상함수는 속도, 충돌, 추가적인 보상으로 이루어져 있다. 속도 보상은 제한된 속도 범위를 벗어나면 매 스텝마다 -10의 페널티를 받도록 설정하였으며, 속도 규정을 준수할 경우 속도가 높을수록 더 큰 보상을 받도록 (2)와 같이 설정하였다.

$$r_{speed} = \frac{v_{max} - \sqrt{(v - v_{max})^2}}{v_{max}} \quad (2)$$

충돌 보상은 (3)과 같이 유클리드 거리의 역수로 설정하여 차량 간 거리가 5m 이내일 때부터 페널티를 받도록 설정하였으며, 3m가 됐을 경우 충돌로 간주하였다.

$$r_{rear/lateral} = -\frac{1}{\text{euclidean distance}} \quad (3)$$

초록색 차량이 충돌하거나 300m에 도달하면 에피소드를 종료한다.

2.3 강화학습 알고리즘

에이전트 학습을 위해 강화학습 알고리즘인 DDPG를 사용하였으며, 모델 네트워크는 표 3과 같다.

표 3. 강화학습 모델

Num of Hidden Layers	3
Num of Nodes per Layer	64
Activation Function	ReLu
Sampling Time	200ms

III. 시뮬레이션 결과

학습된 강화학습 에이전트를 이용하여 총 1,000 에피소드에 대해 테스트를 수행했다. 결과는 표 4와 같으며, 지능형 운전자 모델을 사용함으로써 충돌 횟수는 9.8%에서 0%가

되었고, 평균 속도는 약 0.77% 증가, 차량 간 평균 거리는 48.81% 감소하였다. 평균 속도는 비슷하지만, 거리 차이가 큰 이유는 충돌이 발생하지 않은 에피소드에 대해서만 거리를 측정했기 때문에 강화학습만을 사용하여 테스트한 경우의 평균 거리가 더욱 크다.

표 4. 시뮬레이션 결과

모델	평가지표	충돌 횟수	평균속도 (m/s)	평균거리 (m)
RL		98	11.591	36.023
RL + IDM		0	11.680	18.439

IV. 결론

본 연구는 수학적 모델링에 의존하지 않는 강화학습과 차량 교통 흐름 모델인 지능형 운전자 모델을 사용하여 합류구간에서 커넥티드 자율주행 차량의 속도를 제어하였다. 강화학습만을 사용한 경우 충돌이 98회 발생한 반면, IDM과 강화학습을 함께 사용한 경우 충돌이 발생하지 않았다. 또한 선행 차량의 속도에 맞춰 차량 속도를 제어함으로써 차간 거리를 적정하게 유지하는 것을 확인할 수 있었다.

본 논문에서는 시뮬레이션 환경을 간단하게 구성하였지만, 추후 연구에서는 더욱 정밀한 시뮬레이션 환경 구성과 Observation 정보에 대한 더 많은 연구가 필요하다.

V. Acknowledgement

이 논문은 2022년도 정부 (산업통상자원부)의 재원으로 한국산업 기술진흥원의

지원을 받아 수행된 연구임 (P0020535,
2022년 산업혁신인재성장지원 사업)

참 고 문 헌

- [1] S. K. Sumanth Nakka, B. Chalagi and A. A. Malikopoulos, "A Multi-Agent Deep Reinforcement Learning Coordination Framework for Connected and Automated Vehicles at Merging Roadways," 2022 American Control Conference (ACC), Atlanta, GA, USA, 2022.
- [2] Dhawankar, P.; Agrawal, P.; Abderezzak, B.; Kaiwartya, O.; Busawon, K.; Raboacă, M.S. Design and Numerical Implementation of V2X Control Architecture for Autonomous Driving Vehicles. Mathematics 2021.

도메인 일반화 영상 분할을 위한 클래스 기반 대조 학습

¹양근영, ^{2*}임묘택

Class-based Contrastive Learning for Domain Generalized Semantic Segmentation

¹Geun-Yeong Yang and ^{2*}Myo-Taeg Lim

요 약

영상 분할을 위한 딥러닝 모델은 주로 훈련 단계에 포함되지 않은 새로운 타깃 도메인으로 배포될 때 성능 저하를 겪는다. 이는 안전성이 중요시되는 자율 주행과 로봇 시스템에 있어 심각한 문제로 이어질 수 있다. 본 논문은 도메인 일반화 영상 분할을 위하여 클래스 기반의 대조 학습을 도입하여 이를 해결하고자 한다. 데이터의 유사성 및 차이를 학습하는 대조 학습의 개념을 클래스에 대해 확장하여, 보다 차별적인 클래스 임베딩 공간을 구축한다. 이를 통하여 네트워크가 보다 강인한 영상 분할 능력을 갖도록 한다. 실험을 통해 제시한 기법의 정량적, 정성적 평가를 진행하였으며, 기반 모델 및 타 도메인 일반화 영상 분할 모델들보다도 뛰어난 성능을 보였다.

Key words

Deep Learning, Semantic Segmentation, Domain Generalization, Contrastive Learning

I. 서 론

딥 러닝 기술의 발전과 함께 영상 분할 기술들 또한 현저한 발전을 통해 뛰어난 결과[1]를 보이고 있다. 그러나 모든 픽셀을 분류해야 하는 작업 특성 상, 학습을 위한 라벨링에는 많은 노동과 시간, 그리고 비용을 필요로 한다. 이에 대처하기 위해 3D 그래픽을 활용하여 수동 라벨링의 필요성을 제거한 합성 데이터셋[2][3]이 등장하여 양질의 방대한 데이터를

제공하였다. 그러나 합성 데이터셋으로 훈련된 네트워크가 실제 환경에서 배포될 때 도메인 이동(domain shift)문제가 발생하여, 딥러닝 네트워크의 성능에 영향을 미치게 된다. 이러한 성능 저하는 영상 분할이 주로 적용되는 자율 주행[4], 로봇 시스템[5]에 있어, 신뢰도와 안전성과 직결되기 때문에 심각한 문제를 초래할 수 있다. 도메인 일반화는 이러한 도메인 이동으로 인한 문제를 해결하기 위한 방법으로, 학습을 위한 소스 도메인(source domain)에서만 학습된 네트워크가 학습된

¹ 고려대학교 전기전자공학과 일반대학원 석사과정 (hggofficial@korea.ac.kr)

² 교신저자 고려대학교 전기전자공학과 교수 (mlim@korea.ac.kr)

분포에 포함되지 않는 타깃 도메인(target domain)에서도 원활히 작동할 수 있게 하는 것을 목표로 한다.

본 논문은 대조 학습을 통한 도메인 일반화 영상 분할 기법을 제시한다. 클래스 임베딩 공간에서 같은 클래스에 속하는 특징은 가깝게, 다른 클래스에 해당하는 특징들은 멀어지게 학습하는 대조 학습의 개념을 도입하였다. 대조 학습을 통하여 차별적인 클래스 임베딩 공간을 구축하고, 이를 기반으로 네트워크가 타깃 도메인에서도 강인한 성능을 보이도록 하였다. 다양한 타깃 도메인에서의 정량적, 정성적 평가를 통해 본 알고리즘의 효과를 분석하고 입증하였다.

II. 본 론

2.1 클래스 기반 대조 학습

영상 분할에 있어, 영상의 모든 픽셀은 특정한 클래스에 속한다. 클래스에 대한 높은 이해도는 강인한 영상 분할 성능으로 이어진다는 개념을 바탕으로, 비지도 학습의 infoNCE[6] 손실 함수를 영상 분할의 영역으로 확장하여 적용하였다. 특징 맵의 픽셀 p 의 클래스가 c 라고 하였을 때, 클래스 c 에 해당하는 모든 픽셀을 p^+ , c 에 해당하지 않는 픽셀의 집합을 N 이라고 할 때, 다음과 같은 손실 함수를 정의한다.

$$\mathcal{L}_{NCE} = \frac{1}{P} \sum_{p \in P} -\log \left(\frac{e^{(p \cdot p^+ / \tau)}}{e^{(p \cdot p^+ / \tau)} + \sum_{p^- \in N} e^{(p \cdot p^- / \tau)}} \right)$$

$\tau > 0$ 은 temperature 하이퍼 파라미터이다. 이 손실 함수는 영상 분할 네트워크의 백본 네트워크(e.g. ResNet-50 [7]) 출력에만 적용된다. 영상 분할을 위한 Cross-entropy 손실 함수를 포함하여

학습에 적용되는 전체 손실 함수는 다음과 같다.

$$\mathcal{L}_{all} = \mathcal{L}_{CE} + w \cdot \mathcal{L}_{NCE}$$

w 는 가중치 하이퍼 파라미터로, 0.3을 적용하였다.

2.2 실험 결과

영상 분할 네트워크 학습엔 합성 데이터셋 GTAV[3]를 사용하였고, 평가는 실제 데이터셋인 Cityscapes[8], BDD -100K[9], Mapillary[10]에서 이루어졌다. 모든 데이터셋은 19개의 클래스를 공유한다. 실험은 영상 분할 모델인 DeepLabV3+[1]를 기반으로 제시한 손실 함수를 추가하여 진행하였으며, 백본 네트워크는 ResNet-50[7]을 사용하였다. 학습은 NVIDIA V100, Pytorch 환경에서 SGD 최적화 알고리즘을 사용하여 배치 크기 8로 40,000회 진행되었다.

표 1. 도메인 일반화 영상 분할에 대한 정량적 평가

Methods	C	B	M
Baseline	28.95	25.14	28.18
RobustNet[11]	37.31	35.20	27.90
SANSAW[12]	39.75	37.34	41.86
Ours	42.03	38.27	44.02

표 1.은 타 도메인 일반화 영상 분할 모델들과 제시한 기법의 정량적 성능을 비교한 것이다. 성능 지표로는 mIoU가 사용되었다. 표의 결과를 볼 때, 제시한 기법을 통해 기반 모델(Baseline) 대비 Cityscapes에서는 13.08%, BDD-100K 은 13.13%, Mapillary 에선 15.84%로, 높은 성능 향상을 보였다. 또한 타 도메인 일반화 영상 분할 모델들과 비교하여도 모든 데이터셋에서 보다 뛰어난 성능을 보였다.

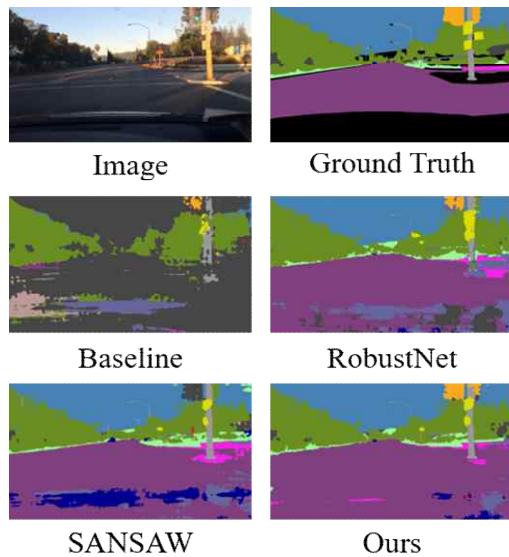


그림 1. 도메인 일반화 영상 분할에 대한 시각화

그림 1.은 타깃 도메인에서의 영상 분할 결과를 시각화하여 타 기법들과 비교한 것이다. 타깃 도메인에 대한 일반화 능력이 부족한 기반 모델의 경우 도로를 인식하는 것을 실패하였고, 타 기법들의 경우도 신호등과 표지판등의 클래스에 대한 인식을 원활히 하지 못하는 것을 볼 수 있다. 반면 제시한 기법을 적용한 경우, 각 클래스에 대한 높은 이해도를 바탕으로 타 기법 대비 뛰어난 결과를 보이고 있다.

그림 2.는 기반 모델과 제시한 기법을 적용한 모델의 클래스 임베딩 공간을 비교한 것이다. 클래스 임베딩 공간에 대한 시각화는 각 특징에 tSNE[13] 분석을 적용하여 진행하였다. 기반 모델의 경우 도로와 보도의 특징들에 큰 뒤얽힘이 발생한 것을 확인할 수 있으며, 타 클래스들도 명확한 구분이 이루어지지 않고 혼재되어 있는 것을 볼 수 있다. 제시한 기법을 적용하여 학습한 모델의 경우 각 클래스에 대한 균집화가 더욱 잘 진행되어 클래스 사이의 경계가 비교적 뚜렷하게 구분되어있는 것을 확인할 수 있다. 이는 제시한 기법이 차별적인 클래스 임베딩

공간을 구축하는데 기여하며, 이는 뛰어난 도메인 일반화 영상 분할 능력에 이어진다는 것을 보여준다.

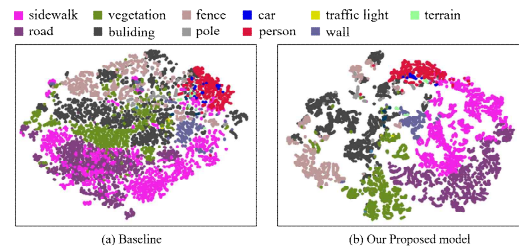


그림 2. 각 클래스 특징에 대한 tSNE 시각화

III. 결 론

본 연구는 도메인 일반화 영상 분할을 위하여 클래스 기반 대조 학습을 제시하였다. 타깃 도메인에서의 강인한 성능을 위하여, 영상 분할에 의미론적으로 직결되는 클래스에 대한 차별적인 임베딩 공간을 구축하도록 대조 학습을 도입하였다. 다양한 기법들과의 정성적, 정량적 평가를 통하여 제시한 기법의 효과를 확인하였으며, 보다 우수한 성능을 보임을 입증하였다.

참 고 문 헌

[1] Chen, Liang-Chieh, et al. "Encoder-decoder with atrous separable convolution for semantic image segmentation." Proceedings of the European conference on computer vision (ECCV). 2018. p. 801-818.

[2] Ros, German, et al. "The synthia dataset: A large collection of synthetic images for semantic segmentation of urban scenes." Proceedings of the IEEE conference on computer vision and pattern recognition.

2016. p. 3234-3243.
- [3] Richter, Stephan R., et al. "Playing for data: Ground truth from computer games." Computer Vision-ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11-14, 2016, Proceedings, Part II 14. Springer International Publishing, 2016. p. 102-118.
- [4] Hu, Yihan, et al. "Planning-oriented autonomous driving." Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2023. p. 17853-17862.
- [5] Onozuka, Yuya, Ryosuke Matsumi, and Motoki Shino. "Autonomous Mobile Robot Navigation Independent of Road Boundary Using Driving Recommendation Map." 2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). IEEE, 2021. p. 4501-4508.
- [6] Oord, Aaron van den, Yazhe Li, and Oriol Vinyals. "Representation learning with contrastive predictive coding." arXiv preprint arXiv:1807.03748 (2018).
- [7] He, Kaiming, et al. "Deep residual learning for image recognition." Proceedings of the IEEE conference on computer vision and pattern recognition. 2016. p. 770-778.
- [8] Cordts, Marius, et al. "The cityscapes dataset for semantic urban scene understanding." Proceedings of the IEEE conference on computer vision and pattern recognition. 2016. p. 3213-3223.
- [9] Yu, Fisher, et al. "Bdd100k: A diverse driving dataset for heterogeneous multitask learning." Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2020. p.2365-2645.
- [10] Neuhold, Gerhard, et al. "The mapillary vistas dataset for semantic understanding of street scenes." Proceedings of the IEEE international conference on computer vision. 2017. p. 4990-4999.
- [11] Choi, Sungha, et al. "Robustnet: Improving domain generalization in urban-scene segmentation via instance selective whitening." Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2021. p. 11580-11590.
- [12] Peng, Duo, et al. "Semantic-aware domain generalized segmentation." Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2022. p. 2594-2605.
- [13] Van der Maaten, Laurens, and Geoffrey Hinton. "Visualizing data using t-SNE." Journal of machine learning research 9.11 (2008).

리소스가 제한된 환경에서 CNN 모델에 대한 주파수 관점의 분석

¹윤지애, ^{2*}김의환

Frequency Perspective Analysis for CNN Models in Resource-Constrained Environments

¹Jiae Yoon and ^{2*}Ue-hwan Kim

요 약

데이터의 양과 모델의 크기가 제한된 환경에서 CNN 모델의 학습을 주파수 관점으로 분석한다. 푸리에 변환을 이용한 주파수 분리 실험으로 학습 데이터가 줄어들 때 경량 CNN 모델이 학습하는 주파수 성분을 측정하여 경량 모델의 경향성을 파악한다. 경량 모델은 데이터 수가 줄어들수록 저주파 성분에 대한 의존도가 올라간다. 이 분석을 통해 리소스가 제한된 환경에서 CNN 모델의 성능 저하의 원인을 분석한다.

Key words

Frequency Analysis, Lightweight Model, Data Augmentation

I. 서 론

최근 연구에서 이미지 데이터를 이용하는 CNN 모델이 학습하는 주파수 성분을 분석하였으며[1][2], 이러한 분석은 CNN 모델의 특성을 이해하고 성능을 향상시키는데 큰 도움을 주었다[3][4]. 특히 CNN 모델의 저주파 성분으로의 편향[5]과 고주파 성분이 성능 향상에 미치는 역할[6]은 적절한 주파수 성분 학습의 중요성을 지적한다.

컴퓨터 비전 분야의 적용 범위가 넓어지면서 모바일이나 임베디드 환경과

같은 리소스가 제한된 환경에서의 이미지 학습에 대한 관심이 높아지고 있다. 그러나 리소스가 제한된 환경에 적합한 경량 모델의 주파수 성분 학습에 대한 분석은 이루어지지 않았다. 본 연구에서는 리소스가 제한된 환경에서의 CNN 모델의 주파수 성분 학습에 대해 분석한다. 제한적인 환경에 적합한 모델을 채택하고 데이터 수를 줄여가면서 모델이 학습하는 주파수 성분에 어떠한 변화가 생기는지를 관찰한다. 우리의 분석은 리소스가 제한된 환경에서 사용하는 경량 CNN의 개선 방향에 대한 통찰을 준다.

¹ 광주과학기술원 AI 대학원, 석박통합과정 (jiaeyoon@gm.gist.ac.kr)

² 교신저자 광주과학기술원 AI 대학원, 조교수 (uehwan@gist.ac.kr)

II. 경량 모델 주파수 분석

2.1 주파수 분석 방법

임베디드 환경에 적합한 모델 분류에 대한 연구 결과[7]에 따라 선정한 ResNet18에서의 주파수 학습에 대해 분석하였다. 클래스 당 데이터의 수를 500, 100, 50, 25, 10, 5개로 제한하며 데이터의 수가 줄어들 때 경량 모델이 학습하는 주파수에 대해 분석하였다. 데이터셋으로 mini-ImageNet을 사용했다.

우리는 주파수 성분 분리를 통해 어떤 주파수 성분을 학습하는지 분석했다[6]. 푸리에 변환을 통해 주파수 도메인으로 변환한 이미지에서 고주파 성분과 저주파 성분을 분리하고 다시 이미지 도메인으로 재변환하여 각각 모델에 입력한다. 분리된 주파수 성분으로 학습했을 때와 원본 이미지로 학습했을 때의 정확도가 비슷하다면 그 주파수 성분을 주로 학습한 것이다.

2.1 주파수 분석 결과

표 1은 데이터의 수에 따른 원본 이미지로 학습한 모델, 저주파 성분으로 학습한 모델, 고주파 성분으로 학습한 모델의 정확도를 보여준다. 데이터의 수가 500개일 때는 원본 이미지로 학습한 모델과 저주파 성분으로 학습한 모델의 정확도 사이에 15.84% 차이가 있다. 이는 CNN 모델이 저주파 뿐만 아니라 고주파 성분의 정보를 학습하고 있는 것으로 해석할 수 있다. 데이터의 수가 100개, 50개, 25개일 때는 500개일 때보다 원본 이미지로 학습한 모델과 저주파 성분으로 학습한 모델의 차이가 더 줄어들었다(각각 9.32%, 5.52%, 2.66%). 그리고 데이터의 수가 10개와 5개일 때로 낮아지면 두 모델의 정확도 차이가 1%도 되지 않는다. 이는 데이터의 수가 줄어들수록

모델이 저주파 정보로만 학습하고 있다는 의미이다.

표 1. 데이터 수와 학습하는 주파수 성분에 따른 CNN 모델의 정확도 (%)

	원본	저주파	고주파
500	61.52	45.68	40.01
100	35.60	26.28	23.85
50	25.50	19.98	16.82
25	16.17	13.55	10.92
10	10.03	9.85	9.06
5	7.12	7.07	4.46

III. 결 론

우리는 데이터의 수에 따른 경량 모델이 학습하는 주파수 성분을 분석하였다. 데이터가 부족할수록 경량 모델은 저주파에 의지하며 고주파 성분을 거의 학습하지 못하였다. 우리는 리소스가 제한된 환경에서 CNN 모델의 성능 저하에 대한 원인 중 하나로 고주파 성분의 학습 불가가 있음을 밝혀냈다.

참 고 문 헌

- [1] Xu, Zhi-Qin John, Yaoyu Zhang, and Yanyang Xiao. "Training behavior of deep neural network in frequency domain." Neural Information Processing: 26th International Conference, ICONIP 2019, Sydney, NSW, Australia, December 12-15, 2019, Proceedings, Part I 26. Springer International Publishing, 2019.
- [2] Yin, Dong, et al. "A fourier perspective on model robustness in computer vision." Advances in Neural Information Processing Systems 32 (2019).
- [3] Xu, Kai, et al. "Learning in the frequency

- domain." Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2020.
- [4] Chen, Xiangyu, and Guanghui Wang. "Few-shot learning by integrating spatial and frequency representation." 2021 18th Conference on Robots and Vision (CRV). IEEE, 2021.
- [5] Rahaman, Nasim, et al. "On the spectral bias of neural networks." International Conference on Machine Learning. PMLR, 2019.
- [6] Wang, Haohan, et al. "High-frequency component helps explain the generalization of convolutional neural networks." Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2020.
- [7] Bianco, Simone, et al. "Benchmark analysis of representative deep neural network architectures." IEEE access 6 (2018): 64270-64277.

반도체 칩 기능 시험의 오류 디버깅 시 Single Long Scan Chain 활용 방안

¹박종수

The Method to Utilize Single Long Scan Chain for Debugging in Semiconductor Chip Function Test

¹Jongsu Park

요 약

반도체 칩의 불량 여부를 판별하기 위하여 후공정 단계에서 웨이퍼 테스트를 수행한다. 이때 대표적으로 반도체 기능 시험과 스캔 시험이 수행된다. 스캔 시험은 제조상의 불량 여부 판별을 전문적으로 수행하기 위한 시험 방법으로써, ASIC 설계 단계에서 디지털 로직 중 D-플립플롭을 스캔 시험을 지원하는 D-플립플롭으로 모두 변경 후 반도체의 기능 동작 외 별도의 스캔 시프트 동작이 추가되어 시험이 가능해진다. 반도체 칩 내부의 모든 플립플롭은 이 스캔 시프트 동작에 참여하게 되며, 이 스캔 시프트 동작을 지원하기 위해 모든 플립플롭은 스캔 시프트 체인에 연결된다. 일반적으로 스캔 시프트 체인은 플립플롭 수백개에서 수천개 단위로 분할된다. 본 논문에서는 분할되어 있는 스캔 시프트 체인을 하나의 긴 시프트 체인으로 연결하여 기능 시험 시 디버깅에 활용할 수 있는 방안을 제안한다.

Key words

Semiconductor Test, Scan Chain, Function Test, Single Long Scan Chain

I. 서 론

반도체 공정은 웨이퍼에 회로를 새기는 전공정 과정과 이후 칩을 패키징하기까지의 후공정 단계로 나뉜다. 반도체 패키징을 위해서 그 전에 반도체 칩의 불량을 정확하고 빠르게 검출하여야 제조비용을 줄일 수 있다. 반도체 후공정 단계에서 반도체 칩의 불량 여부를 판별하기 위하여 웨이퍼 상태 그대로 ATE (Automatic Test Equipment)

장비를 사용하여 여러 불량 검출 시험을 수행한다[1].

이때 대표적으로 수행되는 시험은 크게 반도체 칩 기능 시험(Function Test)과 스캔 시험(Scan Test)이다. 기능 시험은 설계자가 반도체 칩으로 구현하고자 하는 기능이 정상적으로 동작하는가를 시험하므로 ATE장비에서 사용되는 테스트 벡터 또한 RTL 설계자가 제공한다.

스캔 시험은 디지털 로직에 대하여

¹ 목원대학교 전기전자공학과 교수 (jspark@mokwon.ac.kr)

제조상의 불량 여부 판별을 전문적으로 수행하기 위한 시험 방법으로써 DFT (Design for Test) 기법 중 가장 대표적인 기법이다.

스캔 시험을 위해서, ASIC 설계 단계에서 디지털 로직 중 D-플립플롭을 스캔 시험을 지원하는 D-플립플롭으로 모두 변경 후 반도체의 기능 동작 외 별도의 스캔 시프트 동작이 추가되어 시험이 가능해진다. 반도체 칩 내부의 모든 플립플롭은 이 스캔 시프트 동작에 참여하게 되며, 이 스캔 시프트 동작을 지원하기 위해 모든 플립플롭은 스캔 시프트 체인에 연결된다. 일반적으로 스캔 시프트 체인은 플립플롭 수백개에서 수천개 단위로 분할된다. 이러한 과정은 시험은 그림 1에서 보여 지는 반도체 설계 플로우 중 Scan Insertion 단계를 통해서 수행 된다 [2].

이러한 스캔 시험은 반도체 설계 시 필수적이라고 할 수 있을 정도로 보편적으로 사용되므로, 본 논문에서는 별도의 큰 하드웨어 오버헤드 없이 분할되어 있는 스캔 시프트 체인을 하나의 긴 시프트 체인으로 연결하여 기능 시험 시 디버깅에 활용할 수 있는 방안을 제안한다.

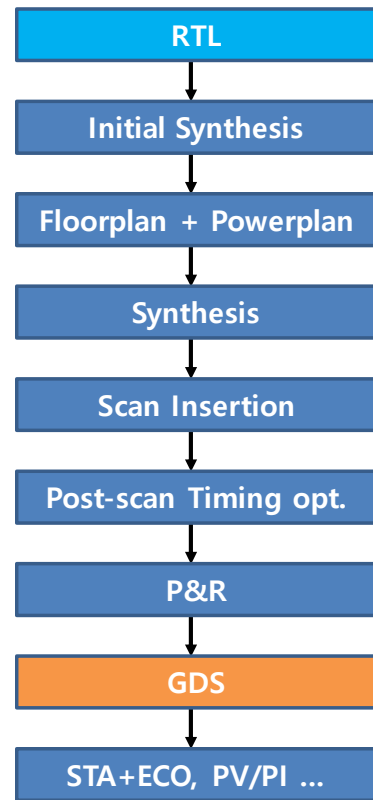


그림 1. 반도체 설계 플로우

II. 스캔 시프트 체인 및 기능 시험에 활용 방안

2.1 일반적인 스캔 시프트 체인

그림 1의 Synthesis 단계를 통하여 Verilog-HDL 혹은 VHDL 등의 하드웨어 기술 언어(Hardware Description Language)로 코딩된 RTL (Register Transfer Level) 코드는 게이트로 변환되며, 또한 모든 D-플립플롭은 스캔 시프트 체인을 구성할 수 있도록 스캔 D-플립플롭으로 변환된다. 스캔 D-플립플롭은 입력 핀 D 외에 SI (Scan Input) 핀과 SE (Scan Enable) 핀을 추가로 가지고 있으며, 멀티플렉서를 통하여 D와 SI를 선택하는 구조이다.

이후 Scan Insertion 단계에서 스캔

D-플립플롭의 SI 핀을 이용하여 반도체 칩 내부의 모든 D-플립플롭을 그림 2와 같이 약 1000~2000 단위의 스캔 체인으로 구성한다.

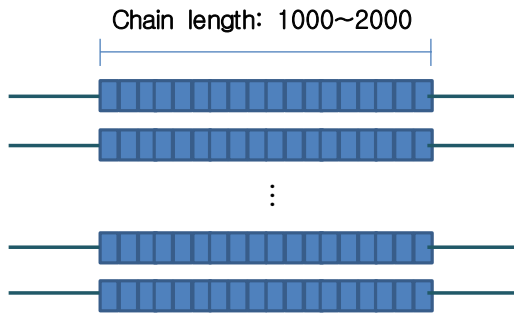


그림 2 일반적인 스캔 시프트 체인

2.2 싱글 롱 스캔 체인

스캔 시험은 최근의 고도화된 반도체 칩 설계 방법론에서는 빠질 수 없는 필수적인 과정으로서 모든 반도체 칩 설계에 적용되고 있다고 해도 과언이 아니다. 즉, 모든 D-플립플롭은 스캔 시프트 체인으로 구성되고 있으며, 이를 체인의 시작부와 종료부에 멀티플렉서를 추가하는 수준으로서 싱글 롱 스캔 체인을 구성할 수 있다. 이는 그림 3에서 보여진다.

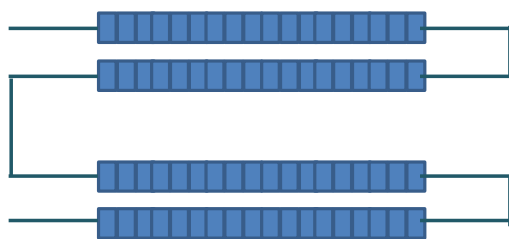


그림 3 싱글 롱 스캔 체인

2.3 기능 시험에 활용 방안

기능 시험은 반도체 칩의 동작으로 기대하는 기능에 문제가 있는지 없는지를 시험하는 것이기 때문에 ATE장비에서 사용되는 테스트 벡터 또한 RTL 설계자가 제공해야 한다.

스캔 시험에서의 불량은 대체로 제조 과정상에서의 문제일 가능성이 높기 때문에 디버깅의 난이도가 높지 않은 편이다. 그러나 기능 시험시 불량이 검출되는 경우는 제조 과정상의 문제일 수도 있고, 또 RTL설계 혹은 ASIC플로우 적용시의 문제 등 매우 다양한 영역에서 심도있는 분석이 필요하다.

기능 시험의 디버깅은 문제가 발생한 상황의 전후의 D-플립플롭 값을 외부로 출력하는 과정이 필요하다. 일반적으로는 ARM 코어 등에서 제공하는 디버깅 로직을 활용하는 방법을 사용하는데, 최근의 CPU 및 GPU는 매우 구조가 복잡하여 원하는 상황에서 원하는 플립플롭의 값을 제대로 출력시키는 것이 불편한 면이 있다.

이때 싱글 롱 체인을 활용한다면 매우 쉽게 이 문제를 해결할 수 있다. 디버깅이 필요한 상황에서 메인 클럭을 정지시키고, 기능모드에서 스캔모드로 동작모드 세팅 후 스캔 시프트 클럭을 토글 시험으로서 원하는 모든 플립플롭의 값을 외부로 출력시킬 수 있다. 이를 활용한다면 기능 시험의 디버깅에 매우 큰 도움이 될 것으로 예상된다.

Ⅲ. 결 론

최근의 고도화된 반도체 칩 설계 방법론에서는 스캔 시험이 빠질 수 없는 필수적인 과정이며, 모든 반도체 칩 설계에 적용되고 있다고 해도 과언이 아니다. 본 논문에서는 별도의 큰 하드웨어 오버헤드 없이 분할되어 있는 스캔 시프트 체인을 하나의 긴 시프트 체인으로 연결하여 기능 시험 시 디버깅에 활용할 수 있는 방안을 제안하였다.

참 고 문 헌

- [1] I. Pomeranz and S. M. Reddy, "Scan-Based Delay Test Types and Their Effect on Power Dissipation During Test," IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, Vol. 27, no. 2, pp. 398-403, Feb. 2008.
- [2] Y. Li, F. Yang, C. Yan and X. Zeng, "Efficient High-Level Synthesis of Approximate Computing Circuits via Multi-fidelity Modeling," 2021 IEEE 14th International Conference on ASIC (ASICON), Kunming, China, 2021, pp. 1-4

지도 정보를 융합한 멀티뷰 카메라 객체인식 모델 개발

¹장민철, ²문석하, ^{3*}김진규

Map-Fused Multi-View 3D Object Detection

¹Mincheol Chang, ²Seokha Moon and ^{3*}Jinkyu Kim

요 약

멀티뷰 카메라 영상을 사용한 자율주행 객체 탐지 알고리즘은 2차원의 정보를 3차원 조감도 (Bird's-Eye-View) 좌표계로 바꾸어 물체인식을 진행한다. 그러나, 카메라 기반 객체 탐지 모델은 2차원 영상의 깊이 정보를 추정하기 어려움이 있고 부정확한 조감도 정보가 생성되므로 심각한 물체 오탐지가 발생한다. 본 연구에서는 기존 멀티뷰 카메라 기반 객체인식 모델의 거리추정 정확성과 조감도 상의 특징정보를 향상시키고자 지도 정보를 2D에서 3D 전환 과정에 융합해 3D 객체 인지 성능을 향상시키는 AI 모델을 개발하였다. 본 연구는 2023년도 대한전자공학회 하계학술대회 및 IEEE/CVF Winter Conference on Applications of Computer Vision 2024에서 발표됨.

Key words

Autonomous Driving, Multi-view Camera 3D Object Detection, Multi-modal Fusion,

I. 서 론

일반적으로 자율 주행 차량에서는 라이다, 카메라, 레이더 등 다양한 센서를 사용해 데이터를 수집하고 딥러닝 모델이 3D 인식, 경로예측을 하는데 사용한다. 3차원 거리정보를 수집할 수 있는 라이다 센서를 기반한 객체인식 모델들의 정확도가 높지만, 라이다 센서의 비용이 높은 점과 날씨 상황에 따른 변동성 등으로 인하여 멀티뷰 카메라 영상정보만을 가지고 3차원 인지를 진행하는 모델 연구 또한 활발히 진행 되고

있다. 특히, 여러 각도에 있는 카메라 영상정보를 3차원 조감도 (Bird's Eye) 정보로 전환 후 객체 인식을 진행하는 방식은 많이 사용하고 라이다 기반 모델들과의 성능 격차도 서서히 줄고 있다. 그러나, 2차원 영상에서의 고질적인 깊이정보 추정 문제가 이미지 픽셀 정보를 조감도상으로 투영시킬 때 부정확한 조감도 정보의 생성을 야기하고, 이는 객체인지의 성능저하로 이어진다.

본 연구에서는 기존 카메라 기반 2D에서 3D 뷰 전환과정을 향상시키고자 지도

¹ 고려대학교 컴퓨터학과 석사과정 (m1ncheoree@korea.ac.kr)

² 고려대학교 컴퓨터학과 박사과정 (shmoon96@korea.ac.kr)

³ 교신저자 고려대학교 컴퓨터학과 교수 (jinkyukim@korea.ac.kr)

정보를 융합해 기존 모델의 깊이정보와 특징추출에 도움을 주므로 모델에게 3차원 공간과 도로 정보를 학습시키는 방법론을 제안한다. 공개 자율주행 데이터인 nuScenes [1] 에서 다양한 실험을 통해 지도와 카메라 융합 객체 인지 딥러닝 모델의 효율성과 파이프라인 성능을 검증하였다.

II. 본론

2.1 지도 데이터셋 생성

nuScenes와 같은 공개 데이터셋에서는 전체 도시에 대한 High Resolution Map을 제공한다. 본 연구에서는 nuScenes에서 제공하는 4개 도시의 도로 정보를 가져온 후 프레임 별 ego 차량의 주행 위치에 따라 50m 반경의 그리드를 자르면 프레임 별 BEV 지도가 생성된다. 또한 ego 좌표계에서 카메라 좌표계의 Rotation 및 Translation Matrix를 사용하여 생성한 BEV 지도를 여섯 개 카메라 좌표로 투영시켰다. 그림 1과 같이 학습과정에서 RGB+D의 투영된 지도와 BEV 지도를 여섯 개 카메라 영상과 같이 모델 입력으로 넣어준다.

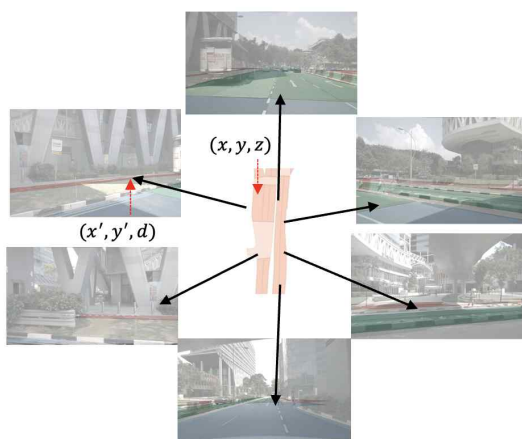


그림 1. BEV에서 PV 뷰 지도 투영 방식

2.2 카메라와 지도 정보 결합

기존 카메라 기반 2D-3D 뷰 전환 방법을 제시한 Lift-Splat-Shoot [2]에서는 EfficientNet 이미지 백본에서 이미지 피처를 뽑고 MLP를 사용해 3차원상에 깊이 (depth)와 특징정보(context)를 예측한다 (Lift 과정). 본 연구에서는 그림 2와 같이 지도 정보에서 CNN 인코더를 사용해 지도 피처를 뽑아 카메라 피처와 channel-wise로 결합하므로 MLP의 이미지 depth와 context 추정에 도움을 준다.

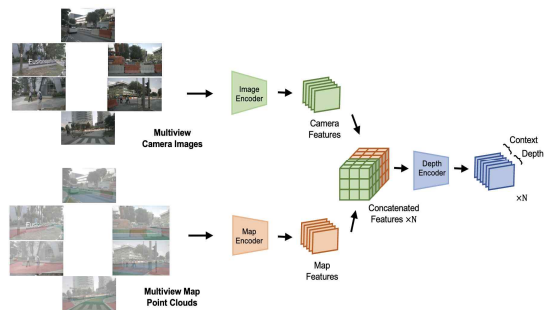


그림 2. Map-Fused Depth/Context Prediction

2.3 BEV 지도를 통한 Normalization

기존 Lift-Splat-Shoot (LSS) 파이프라인을 따라 예측한 depth 정보와 이미지 좌표 (x, y)를 가지고 context feature를 3차원 BEV 특성맵에 뿌려준다 (Splat 과정). 본 연구에서는 기존 연구인 SPADE [3]에 영감을 받아 BEV 조감도로 생성된 피쳐들에 BEV 지도 마스크를 사용하여 피쳐에 대한 spatial modulation을 진행한다.

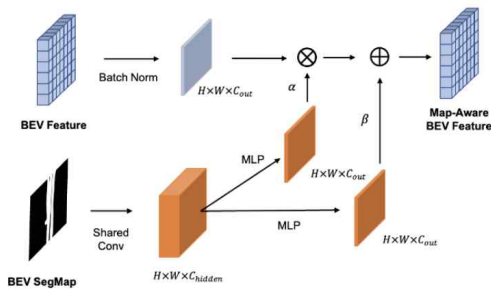


그림 3. BEV 지도 기반 Channel Modulation

Input BEV 지도 마스크를 MLP를 통해 Input Volume ($H \times W \times C$) 크기의 알파와 베타 값을 추출한 다음 이 파라미터들을 전 단계에서 생성된 BEV Feature를 그림 3과 같이 normalize하는데 사용한다.

2.4 성능 결과

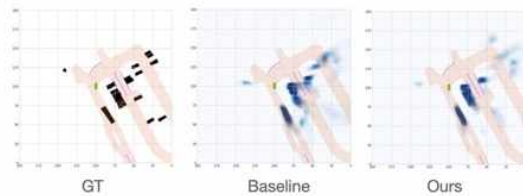


그림 4. BEV상 자동차 객체분할 시각화

먼저 기존 베이스라인 모델인 LSS에 본 연구에서 제안한 지도 융합구조와 BEV Normalization을 사용하는 모델구조를 자동차 객체 분할 실험에 적용 시켰다. nuScenes 데이터셋에서 30 epoch을 학습 시켰고 객체 분할 결과는 그림 4와 같다. 왼쪽 그림이 Ground Truth 자동차의 위치이고 중간 그림이 카메라 영상만 사용한 Baseline 모델, 오른쪽 그림이 지도를 융합한 모델의 시각화 결과이다. 지도를 카메라 피쳐와 융합했을 때 객체 분할이 카메라 영상만 사용한 모델보다 좀 더 깔끔하게 객체 분할이 되는 것을 확인할 수 있다. 정량적으로도 200×200 BEV 그리드에서의 Intersection over union

(IoU)가 32.1에서 33.3으로 성능 향상이 있었다.

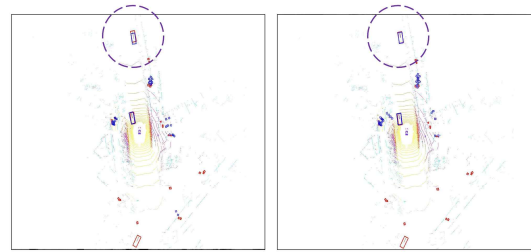


그림 5. BEV상 객체 인식 결과 시각화

두 번째 실험으로는 nuScenes 데이터셋에서 객체 인식 모델을 학습 시켰다. Baseline 모델로는 BEVDet [4]을 사용하였고 해당 모델은 RTX3090 GPU 8개로 24 epoch동안 학습시켰다. 또한, 기존 Baseline모델에 지도 융합 구조를 구현하여 똑같은 세팅으로 데이터를 학습 시켰다. 그림5에서 확인할 수 있듯이, 기존 baseline의 객체 인식 결과보다 지도를 사용한 모델이 더 정확하게 객체를 인지하고 있다. 정량적으로 nuScenes에서 제공하는 nuScenes Detection Score (NDS) 성능 또한 37.7에서 39.5의 향상이 있었다. 구체적으로 객체 박스에 대한 방향과 속도에 대한 예측이 크게 향상 되었는데 이는 모델이 도로정보를 뷰 전환 과정에서 지속적으로 참조하며 BEV 피쳐를 생성함을 보여준다.

III. 결론

멀티 뷰 카메라 기반 3차원 객체 인지 딥러닝 모델을 학습하기 위해서는 2D 정보를 3D BEV로 전환하기 위한 3차원 정보들이 모델에 주어져야 한다. 본 연구에서는 지도 정보가 2D 카메라 센서를 보완하여 카메라와 지도 영상의 결합을 통해 카메라 뷰에서의 거리 예측의 정확도를

높이고 도로 정보를 참조하며 false-positive 예측을 줄이는 효과가 있음을 실험적으로 확인하였다.

사사

본 연구는 2022년도 정부 (산업통상자원부)의 재원으로 한국산업 기술진흥원의 지원을 받아 수행된 연구임. (P0020535, 2023년 산업혁신인재성장 지원사업)

참 고 문 헌

- [1] H. Caesar, V. Bankiti, A. H. Lang, S. Vora, V. E. Liong, Q. Xu, A. Krishnan, Y. Pan, G. Baldan and O. Beijbom, nuScenes: A multimodal dataset for autonomous driving, In CVPR 2020.
- [2] J. Philion and S. Fidler, Lift, splat, shoot: Encoding images from arbitrary camera rigs by implicitly unprojecting to 3d, in European Conference on Computer Vision. Springer, 2020, pp. 194-210.
- [3] Taesung Park, Ming-Yu Liu, Ting-Chun Wang, and Jun-Yan Zhu. Semantic image synthesis with spatially-adaptive normalization. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pages 2337-2346, 2019.
- [4] Junjie Huang, Guan Huang, Zheng Zhu, and Dalong Du. Bevdet: High-performance multi-camera 3d object detection in bird-eye-view.arXiv preprint arXiv:2112.11790, 2021