모빌리티용 인공지능 가속기 설계와 객체 인식 응용

문정수*, 김성재*, 이현석*, 고국원**
*한라대학교 기계자동차로봇공학부
**한라대학교 미래모빌리티공학과
e-mail: 202014031@hu.halla.ac.kr, 201911047@hu.halla.ac.kr, 202125024@hu.halla.ac.kr, kukwon.ko@halla.ac.kr

Design of Artificial Intelligence Accelerator for Mobility and Its Applications to Object Recognitions

Seong Jae Kim*, Jeong Soo Moon*, Hyeon Seok Lee*, Kuk Won Ko**
*Dept. of Smart Automotive Engineering, Halla University
**Dept. of Future Mobility Engineering, Halla University

최근 자율주행 모빌리티 분야에서는 인공지능을 활용하는 사례가 증가하고 있다. 인공지능 사용 시 복잡한 연산처리를 위해서는 GPU가 필수적이나, 높은 전력 소모, 큰 크기 및 높은 가격대로 인한 배터리를 사용하는 모빌리티에 적용하기에는 어려움이 있다. 본 연구는 모빌리티에 적용 가능하도록 25W의 저전력, 저비용, 소형화, 고사양을 목표로 인공지능 가속기를 개발하였다. 개발된 가속기는 다양한 형태의 자율주행 모빌리티에 적용이 가능하는 것을 검증하였으며. 모빌리티 관련의 동적 및 정적 객체에 적용을 통해 26TOPs 이상의 연산 능력으로 처리속도 60fps의 성능을 얻었다.

1. 서론

최근 인공지능(AI) 기술의 발전은 모빌리티 분야에 혁신적 변화를 가지고 왔다. 특히, 자율주행 자동차와 같은 모빌리티 시스템에서 AI는 차량의 인지, 판단, 제어 과정을 획기적으로 개선하며, 불가피 요소로 여겨진다.[1] 그러나, 기술 발전과 더불어 나타나는 주요 문제점 중 하나는 고도화된 연산처리로 이를 위해 고성능의 GPU가 필수이다. 그러나 GPU 장착 시 공간 및 무게와 소비전력이 증가하는데, 이는 모빌리티의 인공지능 적용에서의 큰 어려운 점이며, 배달 로봇과 드론과 같은 소형 모빌리티에 있어서는 더 큰 문제로 나타난다. 또한 고성능 GPU는 높은 가격대를 형성하며, 전기차 기반의 자율주행 시스템을 개발하는 경우 많은 전력 소비로 인해 주행거리 감소를 초래한다.[2]

본 연구는 초소형 모빌리티에 적용하기 위해서는 초소형, 저전력, 저비용이면서 20TOPs 이상의 연산 능력을 가진 Edge AI embedded system을 개발하고 이를 사용하여 다양한 모빌리티 분야에 적용하였다.[3] 모빌리티 분야에서 사용되는 신호등 인식, 표지판 인식, 사물 인식, FOD(Foreign object debris)에 적용하여 효율성을 검증하였다.[4]

2. 모빌리티를 위한 초소형 인공지능 가속기

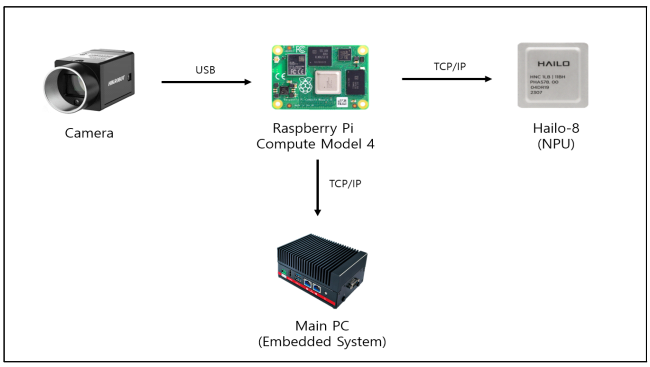
본 연구에서 가속기의 핵심 부품으로 사용된 NPU(Neural Processing Unit)는 Hailo 사의 H8 제품으로 성능은 표 1과 같다.

[표 1] Hailo H8 성능표

정격 전력	2.5W
연산 속도	26TOPs
크기	15mmx15mm
Host Architecture	X86, ARM
AI FrameWork	TensorFlow, Keras, Pytorch 및 ONNX

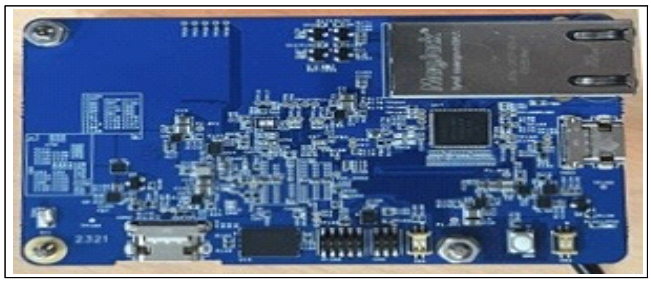
Hailo H8은 주요 AI Frame Work 형식인 TensorFlow, Kreas, Pytorch로 작성된 인공지능 모델을 ONNX(Open Neural Network Exchange)로의 변환 후에 Hailo Dataflow Compiler를 사용 가능하게 함으로써 기존에 개발한 신경회로망의 사용이 가능하다. 또한 TCP/IP 통신 방식을 이용하여 상위 제어기에서 제어가 가능하다.[5]

그림 1은 인공지능 가속기를 사용하기 위한 구조를 나타내고 있다. 기존 사용하고 있는 임베디드 시스템과 TCP/IP 방식을 사용하여 H8 AI Processor와 통신을 한다.[6]



[그림 1] 인공지능 가속기와 기존 상위제어기와의 통신 방법

그림 2는 최종적으로 개발된 Hailo H8을 이용한 인공지능 가속기 제작 사진으로 가속기의 크기는 라즈베리파이 4의 크기와 동일하게 설계함으로써 라즈베리파이와 스택으로 쌓을 수 있도록 제작되었다. 또한, 저전력 설계로 USB3.0으로 전력 공급이 가능하도록 하여 기존의 embedded system에 쉽게 추가 적용이 가능하도록 하였다.[7]



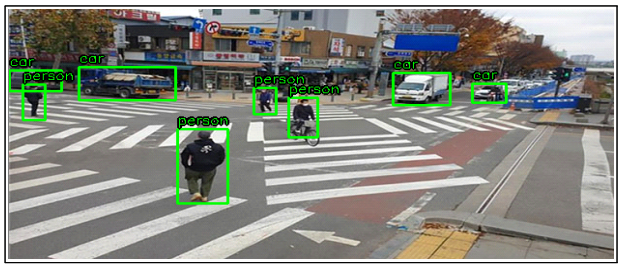
[그림 2] 개발된 인공지능 가속기 TCP/IP 타입

3. 인공지능 가속기 적용

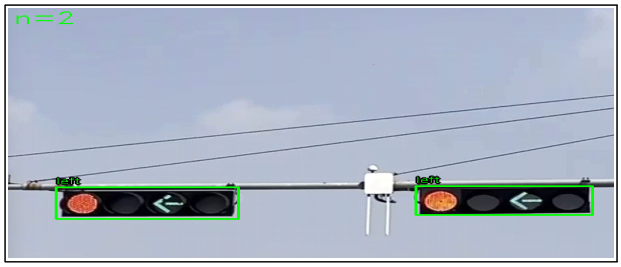
모빌리티용으로 개발한 인공지능 가속기의 효용성을 살펴보기 위하여 모빌리티 분야에서 필요한 객체 인식에 적용시켰다. 대표적인 자율주행에서의 필요한 신호등, 표지판, 다양한 객체 등에 적용하였다.[8]

3.1 학습데이터 구축

객체 인식을 위한 인공지능 학습데이터는 그림3, 그림4, 그림5, 그림6, 그림7과 같이 여러 분야에서 필요한 데이터를 사물 인식의 경우 20,000장, 신호등의 경우 10,000장, 표지판의 경우 6,000장. 라바콘의 경우, 3,000장, FOD의 경우 200,000장을 직접 취득하여 학습하였다. Bounding Box의 크기는 가능한 한 객체의 크기와 일치하도록 annotation을 진행하였다.[9]

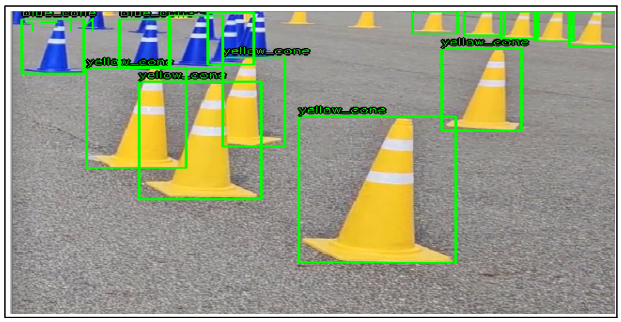


[그림 3] 사물 인식을 위한 학습 데이터

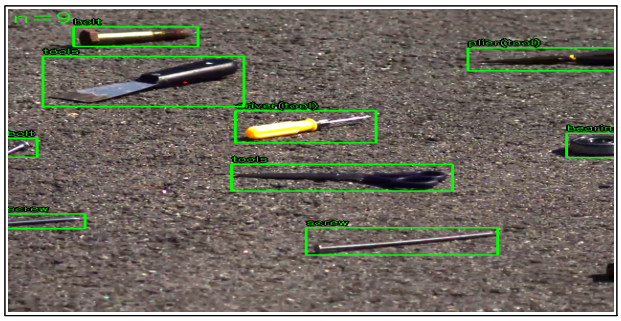


[그림 4] 신호등 인식을 위한 학습 데이터

[그림 5] 표지판 인식을 위한 학습 데이터



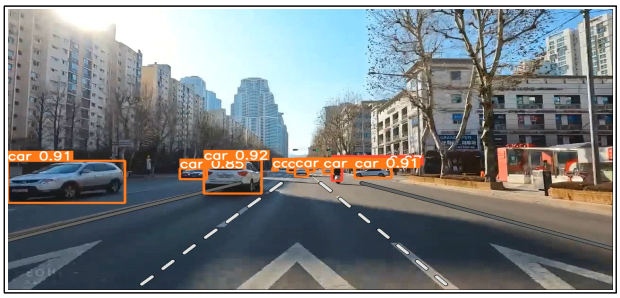
[그림 6] 라바콘 인식을 위한 학습 데이터



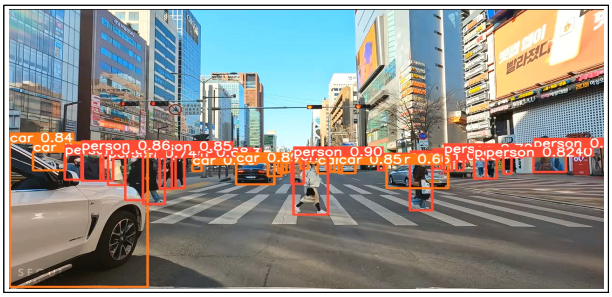
[그림 7] FOD 인식을 위한 학습 데이터

3.2 사물 인식 적용

Edge AI 시스템을 자율주행 청소차의 객체 인식에 적용하였다. 자율주행 청소차는 자동차, 그리고 보행자 인식의 위치 인식이 중요하다. 사용된 모델은 YOLOv5를 사용하여 학습하였다. 라즈베리파이에 연결하여 60 FPS의 속도로 인식을 하였으며, 객체 학습 후 confidence 값이 95% 이상 경우의 동적 객체 인식률을 recall(재현율)로 평가했다. 그 결과표2에 나타난 바와 같이 차량의 경우 95.3% 보행자의 경우 97.0%의 결과를 얻었다. 60fps의 성능을 얻기 위해서는 PC에서 그래픽 카드를 사용하는 경우 RTX3070 이상 급 또는 Jetson xavier nx 이상 급을 사용해야 하지만 라즈베리파이와 AI 가속기를 사용하여 동일한 성능을 얻었으며 소비전력은 25W로 PC에 비하여 월등히 뛰어남을 알 수 있었다.[10]



[그림 8] 동적 차량 인식



[그림 9] 차량 + 객체 인식

[표 2] 인공지능 가속기를 활용한 동적 객체 인식 결과

객체 종류		실험 결과		결과
동적 객체	차량	lane+car1.mp4	93.6	95.3
		lane+car2.mp4	96.5	
		Pedestrian+car1.mp4	95.7	
		Pedestrian+car2.mp4	95.5	
	보행자	lane+car1.mp4	97.2	97.0
		lane+car2.mp4	97.5	
		Pedestrian+car1.mp4	96.2	
		Pedestrian+car2.mp4	97.2	

3.3 신호등 인식 성능

Edge AI 시스템을 자율주행 모빌리티 플랫폼의 신호등 인식에 적용하였다. 자율주행 모빌리티는 교통신호를 정확하게 인식하여 빠른 결정을 내리는 것이 중요하다. 객체 학습 후 confidence 값이 90% 이상 경우의 신호등 인식률을 recall(재현율)로 평가했다. 그 결과 표3에 나타난 바와 같이 전체 학습 인식률은 94.7%의 결과를 얻었다.

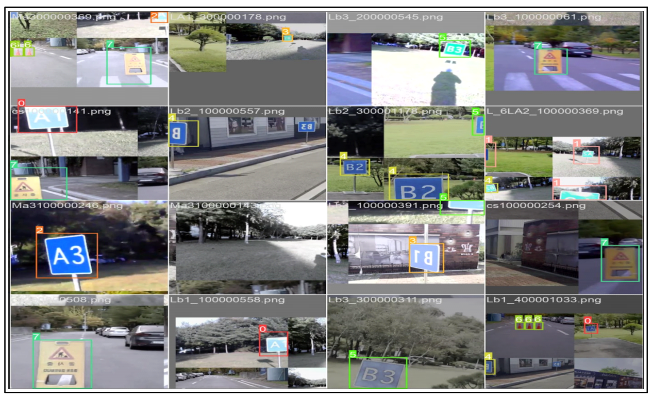
[그림 10] 인공지능 가속기를 활용한 신호등 인식

[표 3] 인공지능 가속기를 활용한 신호등 인식 결과

객체 종류	실험 결과		결과
정적 객체	Traffic_Sign_Red1.mp4	96.4	94.7
	Traffic_Sign_Red2.mp4	96.2	
	Traffic_Sign_Straight1.mp4	96.5	
	Traffic_Sign_Straight2.mp4	96.3	
	Traffic_Sign_Yellow1.mp4	93.8	
	Traffic_Sign_Yellow2.mp4	93.2	
	Traffic_Sign_Straight_Left_Arrow1.mp4	94.2	
	Traffic_Sign_Straight_Left_Arrow2.mp4	94.4	
	Traffic_Sign_Left_Arrow1.mp4	92.6	
	Traffic_Sign_Left_Arrow2.mp4	92.9	

3.4 표지판 인식 성능

표지판은 교통 규칙, 도로 상황 등 도로 주행 중 필수적인 정보를 제공하므로 표지판의 형태 및 색상, 문자를 정확하고 신속하게 인식하는 것이 중요하다. 객체 학습 후 confidence 값이 90% 이상 경우의 표지판 인식률을 recall(재현율)로 평가했다. 그 결과 표4에 나타난 바와 같이 전체 학습 인식율은 95.1%의 결과를 얻었다.



[그림 11] 인공지능 가속기를 활용한 표지판 인식

[표 4] 인공지능 가속기를 활용한 표지판 인식 결과

객체 종류	실험 결과		결과
정적 객체	Sign_A1_1.mp4	94.2	95.1
	Sign_B1_1.mp4	93.8	
	Sign_Construction_1.mp4	96.5	
	Sign_Stop_1.mp4	95.9	

3.5 라바콘 주행 적용

라바콘 주행은 자율주행 경진대회에서 주로 사용되는 종목으로 라바콘의 정확한 인식을 통해 양옆에 색깔이 다른 콘을 두어 좌, 우를 판단하여 그 중심점 위치를 검출하여 충돌 없이 주행하는 것이 중요하다. 객체 학습 후 confidence 값이 90% 이상 경우의 라바콘 인식률을 recall(재현율)로 평가했

다. 그 결과 표5에 나타난 바와 같이 전체 학습 인식율은 95.1의 결과를 얻었다.

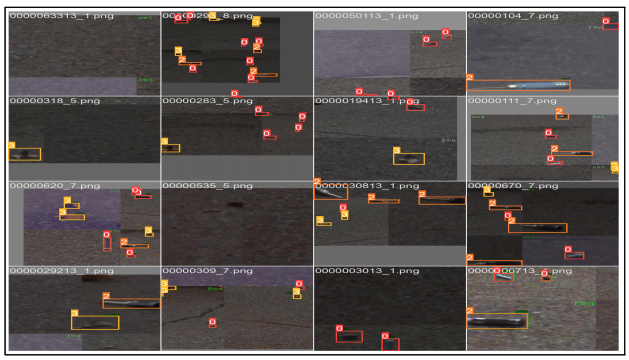
[그림 12] 인공지능 가속기를 활용한 라바콘 인식

[표 5] 인공지능 가속기를 활용한 라바콘 인식 결과

객체 종류	실험 결과		결과
정적 객체	yellow_cone.mp4	94.7	95.1
	blue_cone.mp4	95.4	

3.6 FOD 인식 적용

FOD(Foreign object damage)의 비 규격화 된 환경 특성을 고려할 때, 다양한 클래스와 데이터 셋의 중요성이 강조된다. 객체 학습 후 confidence 값이 90% 이상 경우의 FOD 인식률을 recall(재현율)로 평가했다. 그 결과 표6에 나타난 바와 같이 전체 학습 인식율은 93.1%의 결과를 얻었다.



[그림 13] 인공지능 가속기를 활용한 표지판 인식

[표 6] 인공지능 가속기를 활용한 FOD 인식 결과

객체 종류	실험 결과		결과
정적 객체	Tools_1.mp4	92.3	93.1
	Tools_2.mp4	92.2	
	Tools_3.mp4	94.5	
	Tools_4.mp4	93.2	

4. 결론

본 연구에서는 모빌리티용 인공지능 가속기를 개발하고 다양한 적용을 통하여 개발된 인공지능 가속기의 적용 가능성을 살펴보았다. 초소형 모빌리티에 사용되기 위해서는 소형화, 저전력, 저비용 및 고성능의 요구 조건을 만족하도록 설계하였으며 25W의 소비전력과 라즈베리파이4의 크기로 제작하였다. 인공지능은 26TOPS의 연산 능력을 가져 YoloV5 모델 기준 60FPS의 연산 능력을 보유하여 Nvidia 사의 GPU인 RTX 3070급 이상의 성능에 대응할 수 있었다. 적용된 모델의 경우 인식율은 다소 낮은편으로 이를 향상시키기 위해 학습 데이터 보강과 학습 횟수의 증가를 통해서 보완할 예정이며, 향후 모빌리티의 다양한 예제에 추가 적용할 계획이다.

Acknowledgement

"본 연구는 2024년 과학기술정보통신부 및 정보통신기획평가원의 SW 중심대학 사업 지원을 받아 수행되었음"

참고문헌

[1] 홍정하, 이강찬, 이승윤 "엣지 컴퓨팅 기술 동향" ETRI 일반논문 제35권 6호 pp. 78-87, 2020년

[2] 유병용"자율주행 차량 양산을 위한 핵심 이슈"2018대한전자공학회 하계종합학술집 pp. 1,688-1690 2018년

[3] 김동명, 이동기 "인공지능을 위한 Edge Computing 기술 동향" 한국통신학회지 (정보와통신) 제37권 제12호 pp. 28-34 2020년

[4] 박현문, 황태호 "엣지컴퓨팅기술의 변화와 동향" 한국통신학회지 (정보와통신) 제36권 제2호 pp. 41-47 2019

[5] 차명훈, 강기동, 안백송 외 3명 "AI 모델 파이프라인 병렬화 프레임워크의 설계 및 구현" 한국정보과학회 학술발표논문집 pp. 16-18, 2023년

[6] 장은영, 강래청, 이유식 "AI 서버와의 실시간 데이터 통신을 위한 차량모델 및 TCP/IP 모델 개발" 한국자동차공학회 추계학술대회 pp. 907-910 2021

[7] 조권능, 최도영, 정영우, 이승은 "독립운용이 가능한 임베디드 인공지능 프로세서 설계" 한국정보통신학회 2021년도 춘계종합학술대회 논문집 제25권 제1호 pp. 600-602 2021년

[8] 박용운, 남기문 "인공지능의 가속기 적용 연구" 한국진공학회 학술발표회 초록집 pp. 50-50 2019년

[9] 박태병, 엄주명 "ISO 14649 에 기반한 인공지능 학습데이터 구축 및 채터 감지 알고리즘 개발"전한국정밀공학회 2021년도 추계학술대회 논문집 pp. 688-688, 2021

[10] 서영호, 박성호, 박장호 "영상인식 및 분류용 인공지능 가속기의 최신 성능평가: MLPerf를 중심으로" 방송과 미디어 제25권 제1호 pp. 28-41 2020년