

구글 TPU 보드 기반 인공지능 알고리즘 적용 및 분석에 대한 연구

한광환*, 이창석*, 김도연*, 윤필상*, 가충희**, 정용범**, 정구민§*

*국민대학교 전자공학과

**㈜프리뉴

A Study on apply to AI algorithm using Google TPU Board

Kwang-Hwan Han*, Chang-Suk Lee*, Do-Yun Kim*, Pil-Sang Yoon*, Chung-Hee Ka**,
Yong-Bum Jung**, Gu-Min Jeong§*

*Dept. of Electronic Engineering, Kook-Min University

**Preneu

요약

본 논문에서는 최근 소개된 구글 TPU 보드를 사용하여 AI 알고리듬을 적용하고 성능 분석을 통하여 TPU를 통한 AI에 기반한 영상처리 시스템의 구현 가능성을 검증하고자 하였다. 구글 TPU 보드는 기계 학습에 특화된 Coral Dev 보드를 사용하였고. 수행하는 인공지능 알고리즘은 객체 인식 알고리즘인 SSD 알고리즘을 사용하였다. 이 후 동일한 알고리즘을 GPU가 장착되어 있는 고성능 데스크탑과 처리속도를 비교하여, TPU에 기반한 임베디드 AI 시스템의 활용 가능성을 검증하였다.

1. 서론

최근에 인공지능에 대한 관심이 높아지면서 여러분야에서 많이 사용되고 있다. 인공지능이란 인간이 지닌 지적 능력의 일부 또는 전체를 인공적으로 구현한 것을 뜻한다. 컴퓨터가 스스로 학습하여 결론을 내리는 것이다. 현재 인공지능은 사람 대신 위험한 일을 대체하거나 노동력 문제 해결을 위해서 산업 분야에서도 많은 사용이 되고 있다. 하지만 인공지능 알고리즘은 복잡하고 많은 연산이 필요하여 데스크탑 환경에서 많이 수행되었다. 산업현장에서는 데스크탑을 사용하기 어려운 환경이 많고, 이동하는 물체 같은 경우 이동에 제한이 적으려면 임베디드 환경이 구축되어야 한다. 이에 따라 많은 기업에서 인공지능용 보드를 출시하고 있는 현황이다. 2019년 3월에 구글에서 Coral Dev 보드를 출시하여 임베디드 환경에서 그 전보다 훨씬 좋은 성능으로 인공지능 알고리즘을 실행시킬 환경이 구축되었다.

2. 관련 기술

2.1 Coral Dev 보드

Coral Dev 보드는 구글에서 발표한 기계학습에 특화된 Edge TPU를 장착한 보드이다. 리눅스 환경에서 실행되고 Tensorflow Lite 라이브러리를 지원한다. Edge TPU란 동전 크기의 작은 칩으로, 구글의 Tensorflow AI 소프트웨어를 통해 머신러닝을 가속화하는 장치이다. Coral Dev 보드는 크기가 작은 것에 비해 인공지능

알고리즘을 수행하는 속도가 뛰어나기 때문에 임베디드 시스템에서 많은 활용이 될 것으로 보인다.

장치	사양
CPU	NXP i.MX 8M SOC (quad Cortex-A53, Cortex-M4F)
GPU	Integrated GC7000 Lite Graphics
ML accelerator	Google Edge TPU coprocessor
RAM	1 GB LPDDR4
Flash memory	8 GB eMMC
Wireless	Wi-Fi 2x2 MIMO (802.11b/g/n/ac 2.4/5GHz) Bluetooth 4.1
Dimensions	48mm x 40mm x 5mm

<표 1. Edge TPU 사양>

장치	사양
Flash memory	MicroSD slot
USB	Type-C OTG Type-C power Type-A 3.0 host Micro-B serial console
LAN	Gigabit Ethernet port
Audio	3.5mm audio jack (CTIA compliant) Digital PDM microphone (x2) 2.54mm 4-pin terminal for stereo speakers
Video	HDMI 2.0a (full size) 39-pin FFC connector for MIPI-DSI display (4-lane) 24-pin FFC connector for MIPI-CSI2 camera (4-lane)
GPIO	3.3V power rail 40 - 255 ohms programmable impedance ~82 mA max

	current
Power	5V DC (USB Type-C)
Dimensions	88 mm x 60 mm x 24mm

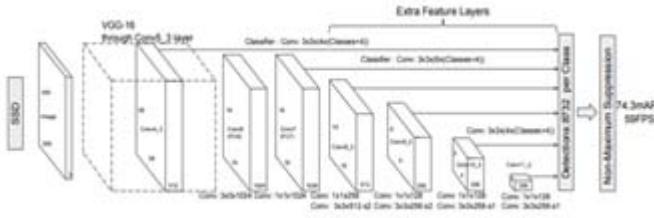
<표 2. Dev 보드 사양>

2.2 Tensorflow Lite

Tensorflow Lite 는 모바일 및 임베디드 시스템에서 인공지능 알고리즘을 저전력 모드 동작에 적합하도록 설계해주는 라이브러리이다. Tensorflow Lite 는 임베디드 환경에서 다음과 같은 동작으로 구동한다. 먼저 Tensorflow 로 학습모델을 정의 한 후 Tensorflow Lite 변환기를 이용하여 Tensorflow 모델을 Tensorflow Lite 포맷 형태로 변환해준다. 변환한 모델을 C++ API 를 통하여 불러오고 Interpreter 를 호출한다. Interpreter 는 Kernel set 을 이용하여 모델을 실행한다. 하드웨어에 맞게 CPU 또는 하드웨어 가속기를 이용하여 실행시킨다.[1]

2.3 SSD(Sinlge-Shot Detector)

Single-Shot Detector(SSD)는 사진의 변형없이 사진 한 장으로 훈련, 검출을 하는 detector 를 의미한다. One-Stage 구조의 객체 검출 기법인 SSD 는 후보영역을 생성하기 위한 RPN(Region Proposal Network)을 따로 학습시키지 않고 특징 피라미드에서 다양한 비율의 객체에 대해서 지역화를 수행하고 모든 객체의 종류에 따른 분류를 수행한다. 즉, 다양한 크기의 특징 맵을 이용하여 객체를 인식한다. CNN 기반의 기본 네트워크 모델인 VGG-16 으로부터 특징 피라미드를 만드는데, 특징 맵은 합성곱 층이 진행됨에 따라 크기가 줄어들게 된다. SSD 는 이 과정에서 추출된 모든 특징 맵들을 추론과정에 사용하여 객체를 인식한다. 입력 영상으로부터 가까이 있는 층에서 추출되어 크기가 큰 특징 맵은 작은 물체들을 검출할 수 있고 입력 영상으로부터 멀리있는 층에서 추출되어 크기가 작은 특징 맵은 큰 물체들을 검출 할 수 있다. SSD 는 RPN 을 제거함으로써 Two-Stage 방법들 보다 학습속도를 향상시켰으며, 다양한 크기의 특징 맵을 이용하여 보다 정확하게 객체를 인식할 수 있다.[2]



<그림 1. SSD 알고리즘 구조>

3. 실험

3.1 실험 환경

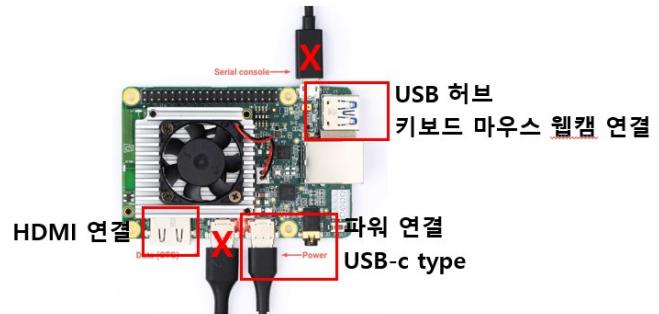
실험은 Coral Dev 보드와 데스크탑에서 진행하였으며 인공지능 모델은 Mobilenet ssd v2 coco 를 사용하였다. 비교 대상인 데스크탑 사양은 <표 3>과 같다.

장치	사양
CPU	Intel i5-7500

GPU	Nvidia GTX 1050 Ti
<표 3. 데스크탑 사양>	

3.2 Coral Dev 보드 환경 구축

Coral Dev 보드에서 Tensorflow Lite 라이브러리를 사용하기 위하여 Mendel Linux 시스템 이미지를 보드에 플래싱을 진행하였다. 또한 SSD 모델을 실행시키기 위하여 <그림 2>과 같이 하드웨어 설정을 진행하였다.



<그림 2. Coral Dev 보드 하드웨어 셋팅>

모니터와 연결을 하여 작업을 진행하기 때문에 HDMI 를 연결하였고 보드를 동작시키기 위하여 USB C-type 을 이용하여 파워를 연결하였다. 사용한 SSD 모델은 카메라로 입력을 받고 사람인지 사물인지 판정하는 모델이므로 USB 허브에 웹캠을 연결하였다.

3.3 실험 결과

장치	Coral Dev 보드	CPU(Intel i5-7500)	GPU(Nvidia GTX 1050 Ti)
프레임 속도(Fps)	20 ~ 30	12	8

<표 4. 프레임 속도 비교>



<그림 3. Coral Dev 보드에서 실행한 SSD 모델>

4. 결론

실험은 데스크탑과 Coral Dev 보드에서 진행되었다.

프레임 속도는 <표 4>와 같이 측정되었으며 Coral Dev 보드 환경에서 더 빠른 프레임 처리 속도를 보여주는 것으로 확인되었다. 실험결과 실험에 사용한 SSD Mobilenet은 핸드폰이나 임베디드 시스템에서 실행하기에 최적화된 모델이기 때문에 데스크탑과 동일한 조건에서 측정하였다고는 볼 수 없지만, Coral Dev 보드가 임베디드 환경의 소형 시스템인 것을 고려하고 TPU를 이용한다면 임베디드 환경하에서 실시간 AI 기반의 시스템의 구현이 충분히 가능할 것으로 판단된다. 적용 분야로는 산업현장에서 결함을 인식하여 판단하거나 사람을 판단하여 경보를 울려주는 경비 시스템과 같은 객체인식 분야에서 충분히 활용 될 수 있을 것으로 판단된다. 향후 Coral Dev를 드론이나 다른 산업분야에 활용하여 AI 시스템을 구현하는 연구를 진행 나갈 예정이다.

5. 감사의 글

이 논문은 2019년도 정부(중소기업기술정보진흥원)의 재원으로 산학연 Collabo R&D 사업의 지원을 받아 수행된 연구임(S2734736)

참고문헌

- [1]Oscar Alising, “Mobile Object Detection using TensorFlow Lite and Transfer Learning”, TRITA-EECS-EX ; 2018:535 ,2018.08
- [2] W. Liu, et al.,“SSD: Single Shot MultiBox Detecto,”In ECCV, 2016