

컴퓨팅 패러다임 개요

현대의 컴퓨팅 패러다임은 **규칙 기반 시스템** → **통계 기반 머신러닝** → **딥러닝** → **생성형 AI** 순으로 발전해왔습니다. 이러한 흐름 속에서 클라우드와 엣지 컴퓨팅은 각각 다른 요구사항과 환경에 최적화된 방식으로 등장했습니다.

Edge Impulse와 FOMO 기술

****FOMO (Faster Objects, More Objects)****는 제약이 있는 엣지 디바이스에서 객체 감지 모델을 실행하는 혁신적인 접근 방식입니다. FOMO는 마이크로컨트롤러에서 처음으로 실시간 객체 감지, 추적 및 카운팅을 가능하게 하는 획기적인 알고리즘입니다.

FOMO의 주요 특징

- **초고속 처리**: MobileNet SSD 대비 30배 빠른 속도
- **저메모리 사용**: 200K RAM 미만에서 동작
- **실시간 감지**: 마이크로컨트롤러에서 실시간 객체 감지
- **중심점 기반**: 바운딩 박스 대신 중심점(centroid) 위치 제공

FOMO의 제한사항

- 객체의 크기 정보는 제공되지 않음 (중심점만 제공)
- 비슷한 크기의 객체에서 최적 성능
- 객체들 간 거리가 너무 가까우면 성능 저하 (해상도 증가로 개선 가능)

실제 활용 사례

▲ FOMO를 활용한 주차장 점유율 실시간 모니터링

이 튜토리얼에서는 FOMO를 사용하여 주차 점유율을 추정하기 위해 자동차를 카운팅하는 방법을 설명합니다.

클라우드 컴퓨팅

클라우드 컴퓨팅은 중앙 데이터 센터에서 강력한 연산 자원을 제공하는 방식입니다. 사용자는 인터넷을 통해 접근하여 대규모 학습, 고성능 분석 등을 수행할 수 있습니다.

▲ 클라우드 기반 데이터 레이블링 시스템

- **특징**:
 - 인터넷 연결 필수
 - 중앙 집중형 처리
 - 확장성과 자원 효율성 우수
- **사용 예시**:
 - ChatGPT 같은 LLM
 - 빅데이터 분석 파이프라인
 - CI/CD 자동화 서버

엣지 컴퓨팅

엣지 컴퓨팅은 데이터를 생성하는 **장치 근처에서 직접 연산**을 수행합니다. 응답 속도가 빠르고, 오프라인 상황에서도 동작 가능하다는 장점이 있습니다.

- **특징:**
 - 로컬 장치에서 처리
 - 지연 시간 감소
 - 연결 없이도 작동 가능
- **사용 예시:**
 - IoT 센서 + ESP32 + 모델 추론
 - CCTV 내 실시간 이상 감지
 - Edge Impulse 기반 AIoT 장비

클라우드 vs 엣지 비교 요약

항목	클라우드 컴퓨팅	엣지 컴퓨팅
연산 위치	중앙 서버 (데이터 센터)	장치 자체 또는 근처 장치
응답 속도	느릴 수 있음	매우 빠름
연결 요구	인터넷 필수	필요하지 않을 수 있음
확장성	높음	제한적

결론

AIoT 시대에서는 클라우드와 엣지를 **상호보완적으로 활용**하는 전략이 필요합니다. 모델 학습은 클라우드에서, 추론(inference)은 엣지에서 수행하는 구조가 대표적입니다.

실제 구현 예시

Edge Impulse 플랫폼을 통해 다음과 같은 워크플로우를 구현할 수 있습니다:

1. **데이터 수집:** 엣지 디바이스나 모바일에서 센서 데이터 수집
2. **클라우드 학습:** Edge Impulse Studio에서 모델 학습 및 최적화
3. **엣지 배포:** 학습된 모델을 C++ 라이브러리나 펌웨어로 배포
4. **실시간 추론:** 엣지 디바이스에서 오프라인 추론 수행

이러한 접근 방식을 통해 **저지연, 고효율, 오프라인 동작**이 가능한 AIoT 시스템을 구축할 수 있습니다.