

지능형 홈케어를 위한 분할된 YOLO 모델과 잠재 코드 전송을 통한 효율적 데이터 통신

박민혁, 김동민*
순천향대학교

xlfksh9613@gmail.com, *dmk@sch.ac.kr

Efficient Data Communication using Split YOLO Model and Latent Code Transmission for Intelligent Homecare

Min Hyeok Park, Dong Min KIM*
Soonchunhyang University

요약

본 연구는 지능형 홈케어 시스템을 위한 분할된 YOLO 모델의 객체 감지 및 분류 기능을 향상시키기 위해 에지 컴퓨팅을 활용한다. 특히, 컴퓨팅 성능이 부족한 장치에서 처리하기 어려운 작업을 에지 서버로 위임할 때 에지 서버와의 데이터 통신 효율성을 개선하는데 중점을 둔다. 본 논문에서 제안하는 잠재 코드 전송 기법은 에지 컴퓨팅 환경에서 대규모 데이터를 효과적으로 관리하고 전송함으로써 딥러닝 모델의 추론 속도를 현저하게 향상시킨다. 이 기술은 특히 노인 감지 및 응급 상황 대응에 있어서 실시간 처리 능력을 요구하는 지능형 홈케어 응용 분야에서 매우 유용하다. 실험 결과는 제안된 시스템이 네트워크 상태에 관계없이 안정적인 성능을 보여주며, 이는 제한된 컴퓨팅 자원을 가진 장치에서도 효율적인 객체 감지가 가능함을 입증한다.

I. 서론

최근 고령화가 진행되는 사회에서는 노인들이 집에서 혼자 생활하는 경우가 많아지고 있어 이로 인해 생길 수 있는 위험 상황에 대한 대비가 중요해지고 있다. 본 논문에서는 객체 감지 모델을 활용하여 노인이 쓰러지거나 응급 상황에 처했을 때 상황을 감지하여 보호자에게 알림을 주는 서비스에서 모델 추론 부분에 에지 컴퓨팅 기반 하이브리드 추론 기법을 적용하여 구현하였다. 구현에 사용된 YOLO 모델[1] 시리즈는 실시간 처리 속도와 정확도 면에서 뛰어난 성능을 보여주는 대표적인 객체 감지 모델 중 하나이다. 하지만 라즈베리파이와 같은 에지 장치에서 YOLO 모델을 실행하는 경우, 컴퓨팅 성능 부족으로 인해 처리 속도가 느려지는 문제가 발생한다. 본 논문에서는 이러한 문제를 해결하기 위해 에지 컴퓨팅 기법을 활용한 Split YOLO 객체 감지 모델 가속화 방법을 제안한다. 제안하는 방법은 라즈베리파이 로컬 환경에서 모델 추론 과정의 일부를 수행하고, 나머지 추론 과정은 서버에서 처리하도록 하여 처리 속도를 향상시킨다. 이를 통해 실시간 객체 감지 시스템 구축에 필요한 빠른 처리 속도를 라즈베리파이와 같은 에지 장치에서도 달성할 수 있도록 한다.

II. 본론

지능형 노인 돌봄 CCTV의 전체적인 동작은 그림 1에 나타낸 것과 같이 동작하게 된다. STEP 1은 카메라로 집 내부를 촬영하는 것이다. STEP 2는 네트워크 컨디션에 따라 전처리 후 STEP 2-1로 전송할지 라즈베리파이에서 모델 추론을 모두 실행 후 STEP 3로 전송할지 결정한다. 만약 전처리 후 STEP 2-1로 중간 데이터를 보내게 됐을 경우 서버에서 남은 추론을 하게 되고 STEP 2-2에서는 결과 값만 라즈베리파이로 전송하게 된다. 마지막으로 STEP 3에서는 최종 결과를 보호자에게 전송하여 상태를 알려주게 된다. 본 논문에서는 STEP 1, 2, 2-1, 2-2에서 일어나는 추론 과정을 에지 컴퓨팅을

활용하여 처리하는 에지 컴퓨팅 기반 하이브리드 추론 기법을 구현하였다.

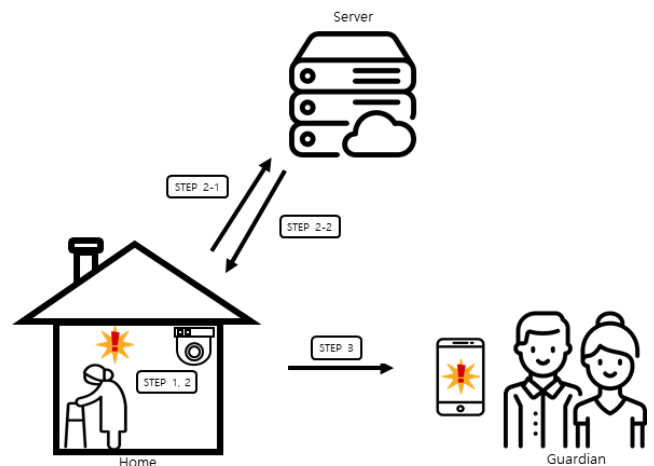


그림 1. 노인 돌봄을 위한 Split YOLO 동작 구조

신행연구[2]에서는 B5G 무선 네트워크 환경에서 네트워크 상태에 따라 에지 컴퓨팅 기반 VGG 모델의 전처리 및 추론 단계를 분할하는 전략을 제안했다. 이 전략은 네트워크 상태에 따라 가장 효율적인 추론 과정을 선택함으로써 에지 컴퓨팅 기반 객체 감지 모델의 성능을 향상하는 데 기여하였다.

최근 실시간 객체 감지 시스템에 대한 요구가 증가하고 있으며, 이러한 상황에서 빠른 추론 속도를 가진 YOLO 모델이 주목받고 있다. 따라서 본 논문에서는 YOLOv3 모델을 활용하여 에지 컴퓨팅 기반 하이브리드 추론 기법인 Split YOLO를 구현하였다. 에지 컴퓨팅 기반 하이브리드 추론 기법은 라즈베리파이 로컬 환경에서 모델 추론 과정의 일부를 수행하고, 나머지 추론 과정은 서버에서 처리하도록 하여 처리 속도를 향상시킨다. YOLOv3는[1] 객체 사이즈에 따른

객체 탐지를 총 세 차례 수행한다. 표 1의 Conv5, Conv6, Conv7 레이어 이후에 객체 탐지를 수행한다. 그에 맞게 본문에서도 네트워크 컨디션을 좋음, 중간, 나쁨으로 객체 사이즈 감지 레이어에 맞춰 총 3 단계로 파악하여 에지 컴퓨팅 기반 하이브리드 추론 기법 Split YOLO를 적용한다. 스타트 데이터로는 1920x1080에 컬러 이미지를 입력하였고, 이미지 형식을 CNN에 적합한 blob 형식으로 변환되어 본격적인 모델 입력으로 들어가게 된다. 하지만 이 전처리 과정부터 라즈베리파이 3 B+ 환경에서는 매 프레임마다 데이터 처리에 상당히 오랜 시간과 에너지가 소모된다. 따라서 이미지 전처리를 하기 전에 현재 서버와의 통신 채널 컨디션을 확인하여 80% 이상 통신 속도를 유지하고 있으면 원본 데이터를 서버로 보내어 상대적으로 빠르게 추론을 한 후 결과값만 받는다. 여기서 서버로부터 받게 되는 결과값은 객체의 좌표, 사이즈, 레이블 정보이다. 다음 50% ~ 80% 통신을 속도를 유지하고 있으면 표 1의 Conv5 레이블의 출력값을 서버로 보내서 첫 번째와 같이 객체들의 좌표, 사이즈, 레이블을 전송한다. 마지막으로 통신 속도가 50% 이하일 경우 전송 데이터 사이즈가 가장 작은 Conv6 레이블의 출력값을 서버로 보내 에지 컴퓨팅 기반 하이브리드 추론 기법을 수행한다.

Layer	Output Dimension	Data Size(MByte)	Latent Code Transmission
Original	1920,1080,3	5.933	Tx. in good channel
Conv1	416,416,32	21.125	-
Conv2	208,208,64	10.562	-
Conv3	-	-	-
Residual 1	208,208,64	10.562	-
Conv4	-	-	-
Residual 1	104,104,128	5.281	-
Conv5	-	-	-
Residual 1	52,52,256	2.641	Tx. in normal channel
Conv6	-	-	-
Residual 1	26,26,512	1.32	Tx. in bad channel
Conv7	-	-	-
Residual 1	13,13,1024	0.66	-

표 1. Split YOLO Architecture

표 2의 결과를 살펴보면 기존에 에지 디바이스에서만 처리한 경우 135 초가 소요되었지만 Lv1, Lv2, Lv3와 같이 현재 네트워크 상태를 확인한 후 에지 컴퓨팅 기반 하이브리드 추론 기법을 사용하게 되면 이전보다 훨씬 효율적인 데이터 처리를 수행할 수 있게 된다. 이를 통해 네트워크 컨디션이 에지 컴퓨팅 구현에 있어 매우 중요한 영향을 끼치고 리소스가 적은 에지 디바이스에서는 전처리 후 빠르게 에지 서버로 전송하여 수행 하는게 더 빠른 것으로 확인되었다. 이렇게 에지 컴퓨팅 하이브리드 추론 기법 Split YOLO를 사용함으로써 상대적으로 리소스가 적은 라즈베리파이 환경에서 YOLO를 사용한 객체 인식 기술을 사용할 수 있게 되었고, 처리량이 현저하게 줄어들어 에너지 효율 측면에서도 상당한 차이를 보여줄 것이다.

(단위 : 초(sec))

Condition	Process	RPi	Transfer	Server	Total
Lv0 (onDevice)		135	-	-	135
Lv1 (Good Channel)		-	0.7	0.68	1.38
Lv2 (Normal Channel)		1.276	1.1	0.623	2.999
Lv3 (Bad Channel)		8.89	0.51	0.45	9.85

표 2. Split YOLO를 활용한 단계별 수행 시간

III. 결론

본 논문에서는 에지 컴퓨팅 하이브리드 추론 기법을 활용하여 지능형 노인 돌봄 CCTV의 처리 과정을 능동적 추론이 가능하게 구현하였다. 이 기법은 에지 장치와 서버 간의 협력을 통해 모델 추론 과정을 분산하여 처리 속도를 향상시키는 것을 목표로 하였다. 제안된 기법은 네트워크 상태에 따라 임베디드 장치와 에지 컴퓨팅 서버로 추론 작업을 분할하여 최적의 성능을 얻을 수 있도록 하였다. 또한, 통신 속도에 따라 데이터 전송량을 조절하여 효율적인 처리를 이루었다. 실험 결과는 제안된 기법이 라즈베리파이와 같은 에지 장치에서 객체 감지 모델의 처리 속도를 현저하게 향상시키는 데 성공하였음을 보여주었다. 따라서, 본 연구에서 제안된 에지 컴퓨팅 기반 하이브리드 추론 기법은 실시간 객체 감지 시스템에 필수적인 빠른 처리 속도를 에지 장치에서도 달성할 수 있게 함으로써 지능형 노인 돌봄 CCTV 이외에도 네트워크 상태가 불안정하거나 라즈베리파이 같이 온디바이스 추론이 상대적으로 어려운 에지 디바이스에서 활용이 가능한 기술임을 보여준다.

ACKNOWLEDGMENT

이 논문은 2024년도 정부(산업통상자원부)의 재원으로 한국산업기술진흥원의 지원을 받아 수행된 연구임 (P0012724, 2024년 산업혁신인재성장지원사업). 이 논문은 2024년 대한민국 교육부와 한국연구재단의 인문사회분야 일반공동연구지원사업(융복합연구)의 지원을 받아 수행된 연구임(NRF-2022S1A5A2A03052880).

참고 문헌

- [1] J. Redmon, A. Farhadi, "YOLOv3: An Incremental Improvement," arXiv preprint arXiv:1804.02767, 2018.
- [2] D. M. Kim, Y. H. Suh and I. F. Ngwalo, "Efficient Data Communication for Deep Learning Application via Latent Code Transmission in B5G Wireless Networks," in Proc. IEEE Information and Communication Technology Convergence (IEEE ICTC 2023), Jan. 2023.