

영상 압축률을 이용한 네트워크 상태에 적응적인 MEC 기반 객체 검출 CNN 모델에 대한 성능 평가

김근용*, 김량수, 윤기하, 유학, 김성창

한국전자통신연구원

{*gykim, rskim, ygh, harkyoo, sungchang}@etri.re.kr

Performance evaluation for MEC-based object detection CNN model adaptable to network conditions using video compression rate

Geunyoung Kim*, Ryangsoo Kim, Giha Yoon, Hark yoo, and Sung Chang Kim

Electronics and Telecommunications Research Institute

요약

본 논문에서는 영상의 높은 압축률 변화로 인한 화질 열화에 대해서도 높은 객체 검출 정확도를 보일 수 있도록 CNN 기반의 객체 검출 인공지능 모델인 YOLOv4를 학습시키고 성능을 평가하였다. 실험을 통해 압축률을 높인 열화 영상과 원본 영상 데이터셋을 혼합하여 학습시키면 열화된 영상에 대해서도 객체 검출 정확도가 유지되는 것을 확인하였다. 본 논문의 결과를 활용하면 무선 네트워크 환경에서 MEC 기반의 객체 검출 결과를 바탕으로 임무를 수행하는 무인 이동체가 네트워크 상태에 따라 전송 영상의 압축률을 조절함으로써 전송 데이터의 크기를 줄여서 전력 소모뿐만 아니라, 다수의 개체로 인해 야기되는 통신 혼잡도로 인한 지연 시간을 감소시키는 효과를 얻을 수 있는 장점이 있다.

I. 서론

CNN 기반의 객체 검출 알고리즘은 기존 방법들에 비해 매우 높은 정확도를 제공하지만, 많은 연산량이 필요하기 때문에 인공지능 모델의 크기가 클 경우 고성능의 GPU가 필요하다. 따라서, 최근 현장에서 취득한 영상을 엣지 서버로 전송하여 객체 검출을 수행하는 MEC(Multi-access Edge Computing) 기반 offloading 기술들이 연구되고 있다[1]. 무인 이동체가 MEC 서버로 영상 데이터를 전송할 때 일정하지 않은 시간이 소요되며, 이는 무선 네트워크 환경에 따른 신호 감쇠, 통신 채널 혼잡도에 크게 영향을 받는다. 따라서, 가능한 일정한 지연 시간을 유지하면서 객체를 검출을 수행하기 위해서는 네트워크 상태가 열악할 경우 전송할 영상의 압축률을 높여 데이터의 크기를 줄여야 한다. 그러나, CNN 기반의 주요 객체 검출 모델의 정확도는 영상의 품질에 영향을 받는다[2]. 이는 객체 검출 모델의 필터들이 영상 데이터의 특징을 기반으로 학습되기 때문에 영상의 압축률을 높이면 영상의 열화로 인해 공간 특징이 변하기 때문이다. 정확도의 감소를 최소화하면서 엣지 서버로 전송되는 영상 데이터의 크기를 줄이면, 무인이동체의 통신모듈의 송신시간이 줄어들어 전력소모를 줄이는 효과가 있다. 또한, 무선 통신 혼잡도를 줄일 수 있기 때문에 다수의 이동체들이 통신할 경우 지연 시간이 줄어든다. 본 논문에서는 영상의 높은 압축률로 인한 화질 열화에 대해서도 정확한 객체 검출 정확도를 보일 수 있도록 CNN 기반의 객체 검출 인공지능 모델인 YOLOv4[3]를 학습시키고 성능을 평가하였다.

II. 본론

그림 1은 MEC기반 무인 이동체 객체 검출 서비스 구조도를 나타낸다. 그림 1에서 보는 바와 같이 무인 이동체는 임무 지역에서 영상 데이터를



그림 1. MEC 기반 무인 이동체 객체 검출 서비스 구조도

무선 네트워크를 통해 원격지의 엣지 서버로 전송하며, 엣지 서버는 CNN 기반 객체 검출 알고리즘을 구동하여 그 결과를 무인 이동체로 전송한다. 객체 인식 결과를 전달 받은 무인 이동체는 객체 추적 등의 임무를 수행한다. 무인 이동체를 위한 MEC기반 객체 검출 서비스에서 중요한 점은 적은 전력소모와 함께 저지연으로 서비스를 제공해야 한다는 것이다[4, 5].

그림 2의 점선 그래프는 원본 영상으로 학습한 모델의 성능 평가 영상의 품질에 따른 객체 검출 정확도의 변화를 나타낸다. 그림 2의 점선 그래프에서 객체 검출 모델은 COCO train2014 데이터셋에서 제공하는 원본 영상으로 학습한 YOLOv4를 사용하였다. 학습한 모델의 영상 품질에 따른 객체 검출 정확도의 변화를 살펴보기 위해 COCO valid2014 영상의 품질을 파이썬의 openCV2 라이브러리에서 제공하는 영상 저장 함수인 imwrite의 IMWRITE_JPEG_QUALITY 인자값을 변경하여 조절하였다. 그림 2에서 Q10 ~ Q90은 IMWRITE_JPEG_QUALITY의 인자값이 10 ~ 90임을 나타내며 Org는 원본 영상을 나타낸다. Q90에서 Q10으로 갈수록 압축률이

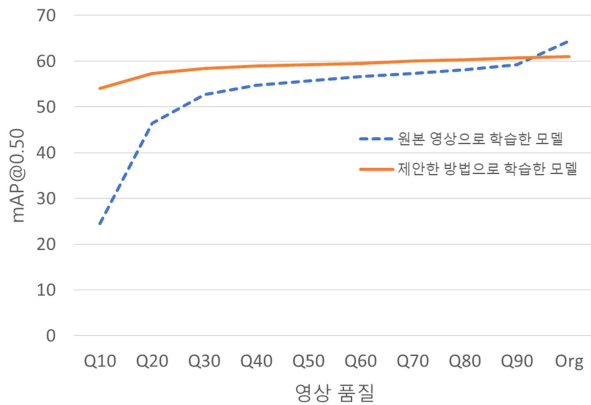


그림 2. 학습 영상의 품질에 따른 객체 검출 정확도 변화

높아짐에 따라 파일 크기가 작아지지만, 영상의 품질이 떨어진다. 일반적으로 동일한 영상 압축 알고리즘에서 압축률을 높이면 큰 양자화 계수가 사용되기 때문에 영상의 고주파 부분이 사라지며, 이는 영상의 공간적 특징을 필터를 통해 학습하여 객체를 검출하는 CNN 기반의 인공지능 모델의 객체 검출 정확도에 영향을 미친다[2]. 따라서, 높은 화질로 학습한 인공지능 모델을 통해 열화된 영상 내의 객체 검출을 수행하면 정확도가 떨어진 결과를 보인다. 그림 2에서 보는 바와 같이 영상의 압축률이 높아질수록 정확도가 떨어지는 것을 확인할 수 있다. 특히, Q10 영상에 대해 급격하게 정확도가 떨어지는데, 이는 원본 영상으로 학습된 YOLOv4 모델이 정상적으로 객체를 검출할 수 없음을 나타낸다. 본 논문에서는 이런 영상 압축 알고리즘과 CNN 기반 객체 검출 모델의 특징을 기반으로 동일한 영상에 대해 원본 영상과 압축률을 높은 영상을 데이터셋에 모두 포함시켜 인공지능 모델을 학습시킴으로써 열화 영상에 대해서도 높은 정확도를 보이는 YOLOv4를 학습시키고 성능을 평가하였다. YOLOv4 모델의 학습을 위해 NVIDIA의 T4 GPU 4대를 사용하였으며, COCO train 2014 데이터셋의 원본 영상과 Q10의 압축률로 품질을 열화시킨 데이터셋을 사용하였다. 학습을 위해 사용된 COCO train2014 영상 개수는 총 82738장이며, 정확도 측정을 위해 valid2014 영상 40504장을 사용하였다.

그림2의 실선은 제안한 방법으로 학습한 모델의 영상 품질 변화에 따른 객체 검출 정확도를 나타낸다. 그림 2에서 보는 바와 같이 제안한 방법으로 학습한 모델은 영상이 열화되어도 높은 정확도를 보이는 것을 확인할 수 있다. 그림 3은 학습 모델에 따른 객체 검출 정확도 비교 실험 결과를 나타낸다. 실험에서는 Q10 영상으로 학습한 모델, 원본 영상으로 학습한 모델, 제안한 방법으로 학습한 모델이 사용되었으며, 성능 평가를 위한 COCO train2014 영상 데이터셋은 Q10 열화된 영상과 원본 영상을 사용하였다. 그림 3에서 보는 바와 같이 Q10으로 열화된 영상으로 학습한 모델은 Q10 열화 영상과 원본 영상 모두에 대해 mAP@0.50의 정확도가 비슷한 객체 검출 정확도를 보였다. 원본 영상으로 학습한 모델은 원본 영상에 대해 가장 정확한 객체 검출 정확도를 보였지만, Q10 열화 영상에 대해 매우 부정확한 객체 검출 정확도를 보였다. 이는 Q10 열화 영상은 높은 압축을 위해 압축 알고리즘에서 큰 양자화 계수를 사용함에 따라 영상 내 객체의 엣지 부분의 데이터가 사라졌기 때문이다. 이런 단점을 보완하고자 원본 영상과 Q10 열화 영상을 함께 사용한 제안한 방법으로 학습한 모델은 Q10 열화 영상에 대해 가장 높은 객체 검출 정확도를 보였으며, 또한 원본 영상에 대해서도 높은 객체 검출 정확도를 보였다. 이는 YOLOv4 모델을 학습시킬 때, 고주파 성분인 엣지 부분을 포함한 원본 영상 뿐만 아니라, 높은 압축률로 인해 열화된 영상에 대해서도 학습시킴으로써, 고주파 부분이 사라진 블러링된 영상에 대해서도 높은 정확도를 갖게 필터가 학습

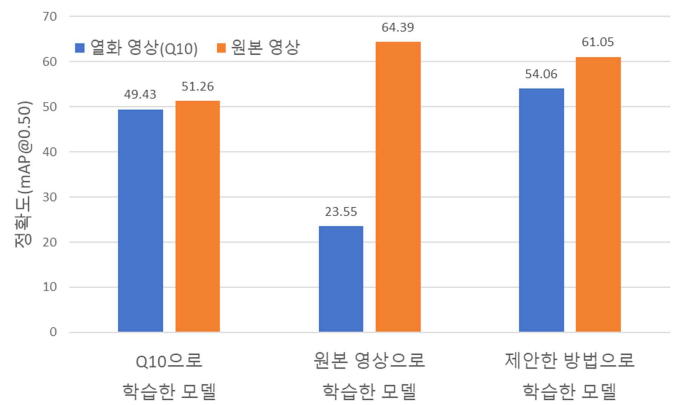


그림 3. 열화/원본 영상에 대한 학습 모델에 따른 객체 검출 정확도 비교

되었기 때문이다. 따라서, 본 논문에서 제안하는 방법을 활용하면 무선 네트워크상에서 무인 이동체가 MEC 기반 객체 검출 서비스를 제공할 때, 네트워크 상태에 따라 전송할 영상 압축률을 조절함으로써, 열악한 네트워크 환경에서도 일정한 품질의 객체 검출 정확도를 제공할 수 있다.

III. 결론

본 논문에서는 영상의 높은 압축률 변화로 인한 화질 열화에 대해서도 높은 객체 검출 정확도를 보일 수 있도록 CNN 기반의 객체 검출 인공지능 모델인 YOLOv4를 학습시키고 성능을 평가하였다. 실험을 통해 열화 영상과 원본 영상 데이터셋을 혼합하여 학습시키면 영상 품질에 변화가 있어도 객체 검출 정확도가 높아지는 것을 확인하였다. 본 논문의 결과를 활용하면 무선 네트워크 환경에서 MEC 기반의 객체 검출 결과를 바탕으로 임무를 수행하는 무인 이동체가 네트워크 상태에 따라 전송 영상의 압축률을 조절함으로써 전송 데이터의 크기를 줄여서 전력 소모뿐만 아니라, 다수의 개체로 인해 야기되는 통신 혼잡도로 인한 지연 시간을 감소시키는 효과를 얻을 수 있는 장점이 있다.

ACKNOWLEDGMENT

본 연구 논문은 한국전자통신연구원 연구운영지원사업의 일환으로 수행되었음. [22ZK1100, 호남권 지역산업 기반 ICT 융합기술 고도화 지원사업]

참 고 문 헌

- [1] Y. Matsubara, "Split computing and early exiting for deep learning applications: survey and research challenges," ACM computing surveys, March 2022.
- [2] M. Aqqa, "Understanding how video quality affects object detection algorithms," VISIGRAPP 2019, pp. 96-104, 2019.
- [3] YOLOv4, <https://github.com/pjreddie/darknet>
- [4] G. Kim, "DNN inference offloading for object detection in 5G multi-access edge computing," ICTC, pp.389-392, 2021.
- [5] G. Yoon, "Implementing practical DNN-based object detection offloading decision for maximizing detection performance of mobile edge devices," IEEE Access, v.9, pp.14099-140211, 2022.