

Custom Dataset 기반 Spiking YOLOv8n 구현 및 성능 분석

이승현, 오성현, 김정곤*

한국공학대학교 전자공학부

seng0424@tukorea.ac.kr, osh119@tukorea.ac.kr, jgkim@tukorea.ac.kr*

Spiking YOLOv8n Implementation and Performance Analysis Based on Custom Dataset

Seung Hyun Lee, Sung Hyun Oh, Jeong Gon Kim*

Dept. of Electronic Engineering, Tech University of Korea

요약

자율 주행 차량, 드론 및 무인 항공기 등 자원이 제한적인 시스템에서 실시간 객체 탐지는 높은 연산량과 전력 소모를 요구하는 ANN(Artificial Neural Network) 기반 모델의 제약으로 어려움을 겪고 있다. 본 연구에서는 기존 YOLOv8n 모델의 Convolution 계층에 LIF(Leaky Integrate-and-Fire) 뉴런을 결합한 SpikeConv 계층을 제안하여, 객체 탐지 성능은 유지하면서 저전력 이벤트 기반 연산을 실현한 Spiking YOLO 모델을 개발하였다. SpikeConv는 membrane potential을 순차적으로 갱신하고, 채널별 가중치 분포에 따른 adaptive threshold를 적용해 스파이크 소실을 최소화한다. 전동 키포드 불법 주행 단속용 Custom Dataset을 활용한 학습 및 추론 실험에서, 제안 모델은 mAP $\approx 99\%$, F1-score ≈ 0.99 를 유지하면서 추론 시간, 에너지 소모 및 GFLOPs를 모두 유의미하게 감소시켰다. 이러한 결과는 자원이 제한된 환경에서도 고성능·저전력을 동시에 달성할 수 있는 ANN to SNN 변환 기술의 실용 가능성을 제시한다.

I. 서론

현재 자율 주행 차량, 드론 및 무인 항공기 등의 지능형 이동체는 충돌 방지, 경로 계획, 환경 인식 등 자율 동작을 위해 실시간 객체 탐지 기능이 필수적이다. 특히 제한된 전력과 컴퓨팅 자원을 갖는 임베디드 시스템 환경에선, 높은 정확도를 보여주는 ANN(Artificial Neural Network)을 적용하면서 많은 연산 및 높은 전력 소모를 요구한다.

기존 YOLOv8n 모델은 경량화 및 고속 추론을 위해 설계된 구조로, 다양한 객체 탐지 응용 분야에서 우수한 성능을 입증하였다[1]. 그러나 YOLOv8n은 수백만 건의 부동소수점 연산(FLOPs)과 연산당 높은 전력을 필요로 하며, 배터리 구동 드론이나 차량 내 임베디드 보드 등 자원이 제한된 환경에서 실시간 처리에 한계가 있다[2].

SNN은 뉴런 간 스파이크 신호 전달을 기반으로 동작하며, 입력이 임계값을 초과한 순간에만 활성화되어 이벤트 기반 연산을 수행함으로써 불필요한 연산을 줄여 저전력 및 저지연 처리, 계산량 감소를 가능하게 한다[3]. 최근 들어 컴퓨팅 자원이 제한된 환경에서도 효율적인 추론 성능을 달성할 수 있는 SNN 모델에 대한 관심이 커지고 있으며, 특히 신경모방(Neuromorphic) 하드웨어와 결합하여 더욱 큰 효율성을 제공할 수 있음을 확인했다[4]. SNN 기반의 객체 탐지 시스템은 적은 전력과 낮은 하드웨어 리소스를 요구하므로, 자율 주행 차량이나 드론과 같이 연속 동작이 필수적인 분야에서 장기 운용 및 배터리 수명 연장에 기여할 수 있다. 또한, 기존 ANN 기반 객체 탐지 모델은 일반적으로 복잡한 환경이나 실제 환경에서의 성능 저하가 불가피하며, 특히 조도 변화, 다양한 각도의 객체 출현, 가림 현상 등 실외 환경에서 발생할 수 있는 여러 가지 상황에서의 탐지 정확도가 현저히 감소하는 경향이 있다. 이러한 한계를 극복하기 위한 다양한 방법 중 하나로 이벤트 기반 SNN 연산이 주목받고 있으며,

본 연구에서도 실제 환경에서 촬영된 다양한 각도의 객체 이미지를 포함한 전동 키포드 불법 주행 단속용 Custom Dataset을 사용하여 현실적이고 실용적인 SNN 객체 탐지 시스템을 구현하였다.

본 연구는 기존 ANN의 정확도를 유지하는 것을 넘어서, 실시간 객체 탐지 시나리오에서 발생할 수 있는 실제적인 제약 조건을 고려하여 전력 소모 및 계산 복잡성을 효과적으로 줄이고자 한다. 실험 결과를 통해 제안된 Spiking YOLO 모델은 실제 응용 환경에 적용되어 임베디드 하드웨어 플랫폼에서의 효율적인 추론 성능과 높은 신뢰성을 동시에 만족시킬 수 있는 실용적인 솔루션으로 활용 가능성을 갖는다.

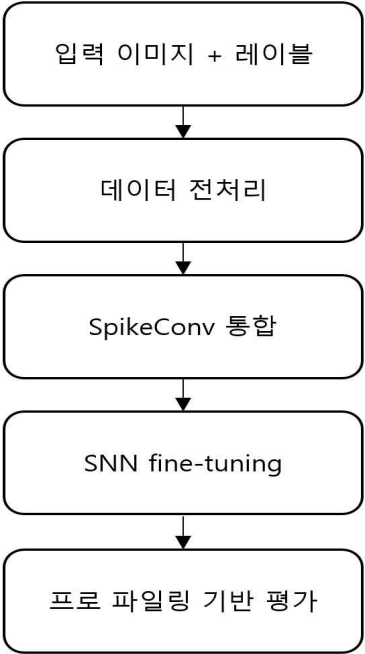
본 논문의 구성은 다음과 같다. 2절에서는 제안한 Spiking YOLO 모델의 데이터 전처리 파이프라인, SpikeConv 계층 설계, 그리고 fine tuning 학습 및 프로파일링 기반 성능 평가 방법을 상세히 기술하며 평가 결과를 분석한다. 끝으로 3절에서는 본 논문의 결론 및 향후 연구 방향에 대해 기술한다.

II. 본론

본 절에서는 [그림 2]의 5단계 파이프라인을 따라 Spiking YOLO 모델의 구현 및 평가 과정을 요약한다. Custom Dataset의 이미지와 텍스트 라벨을 매칭해 유효 샘플을 선별한 뒤 640×640 으로 조정하고 0~1로 정규화하여 채널 우선(C, H, W) 텐서로 변환하며, PyTorch DataLoader의 배치 크기, 셔플링, 다중 워커 설정으로 입력 일관성을 확보한다. YOLOv8n backbone의 모든 3×3 이상 Conv2d 계층을 SpikeConv 모듈로 교체한다. SpikeConv 내부에서는 convolution 연산으로 공간적 특징을 추출하고, 배치 정규화로 채널별 출력을 안정화한 뒤 Leaky Integrate & Fire Neuron이 membrane potential을 누적해 스파이크를 발생시킨다. 마지막으로 SignedNeuron이 양·음 스파이크를 생성하며, surrogate gradient 방법을 통해 미분 불가능 구간을 근사 미분함으로써 end-to-end 학습이 가

* : 교신저자

능하도록 설계하였다. 이 과정에서 원본 가중치·바이어스를 보존하고 채널별 가중치 분포에 기반한 adaptive threshold를 자동 조정해 초기 스파이크 소실을 억제하였다. 학습 설정은 배치 크기 32, 200 epochs 조건으로 fine-tuning을 수행해 최적 파라미터로 수렴시켰다.



[그림 1] Spiking YOLO 모델 구성 및 평가

학습 완료 후, 프로파일링 평가를 통해 GPU 전력 소비·스파이크 발생, GFLOPs, 평균 지연 시간·에너지 소비를 측정할 결과, [표 1]에서 볼 수 있듯이, Spiking YOLO는 기존 YOLO 모델 대비 전반적인 검출 성능을 유지하면서 연산량 및 전력 소모가 크게 절감되었다.

[표 1] YOLO 모델과 Spiking YOLO 모델 비교 결과

Model	mAP50(%)	Latency(ms)	Energy(mJ)	GFLOPs
YOLO	99.35	31.8	1464.8	0.85
Spiking YOLO	99.26	26.2	1217.0	0.72

먼저, mAP50의 경우는 두 모델 모두 99% 이상의 유사한 검출 정확도를 기록하였으며, F1-score는 두 모델 모두 0.99의 값을 가진다. 또한 추론 지연 시간 (Latency)은 Spiking YOLO 모델이 기존 YOLO 모델에 비해 약 17% 빠른 응답 속도를 보였다. GPU 전력 측정 결과, Spiking 모델은 단일 추론당 1217.0 mJ의 에너지를 소비하여 기존 YOLO 모델의 1464.8 mJ에 비해 약 17% 절감되었으며, 효율적 이벤트 기반 연산을 통해 GFLOPs 역시 0.72로 기존의 YOLO 모델의 0.85 대비 약 15% 감소하였다. 이처럼 Spiking YOLO 모델은 핵심 성능 지표인 mAP50 과 F1-score에 대해 유사한 성능을 보이면서 지연 시간, 에너지 소모, 연산량 측면에서 모두 기존 모델 대비 좋은 성능을 보인다. [그림 1]은 동일한 테스트 이미지에 대한 기존 YOLO 모델과 Spiking YOLO 모델의 검출 결과를 비교한 것이다. 기존 YOLO 모델은 제공된 테스트 이미지에서 ‘people’ 객체를 94.1%, ‘kickboard’ 89.1%, ‘nonhelmet’ 2개 객체에 대해 각각 99.7%, 87.8%, ‘person’ 2개의 객체에 대해 각각 90.5%, 88.7%의 신뢰도로 탐지되었다. Spiking YOLO 모델은 동일한 테스트 이미지에 대해서 ‘people’



[그림 2] 기존 YOLO 모델 (좌) 및 Spiking YOLO 모델 (우) 객체 검출 결과

객체를 94.9%, ‘kickboard’ 92.3%, ‘nonhelmet’ 2개 객체에 대해 각각 88.7%, 87.6%, ‘person’ 2개의 객체에 대해 각각 88.9%, 88.3%의 결과가 나타났다. 이처럼 Spiking YOLO는 주요 객체 클래스에 대해 기존 모델과 유사한 성능을 보이며, SpikeConv 대체가 탐지 안정성에 부정적 영향을 끼치지 않았음을 확인할 수 있다.

III 결론

본 연구에서는 YOLOv8n 모델의 Convolution 계층을 SpikeConv 모듈로 대체하고 fine tuning을 거쳐, 실용적 SNN 기반 객체 탐지기를 구현하였다. 제안된 Spiking YOLO는 기존 모델과 유사한 검출 정확도를 보장하면서도, 불필요한 연산과 전력을 효과적으로 줄여 자원 제한적 임베디드 환경에서의 실시간 객체 탐지에 적합함을 입증하였다. 향후에는 neuromorphic 칩 호환성 검증, 다양한 객체 클래스 및 실제 환경 적용 확대를 통해 제안 기법의 범용성과 실용성을 더욱 강화할 예정이다.

ACKNOWLEDGMENT

이 논문은 2025년도 정부(산업통상자원부)의 재원으로 한국산업기술진흥원의 지원을 받아 수행된 연구임 (RS-2024-00415938, 2024년 산업혁신인재성장지원사업)

참 고 문 헌

[1] Wang C. Y., Bochkovskiy A., Liao H Y. M., “YOLOv8n: Design and Performance of a Lightweight Real-Time Object Detector,” IEEE Trans. Circuits Syst. Video Technol., vol. 34, no. 5, pp. 1925 - 1938, May 2025.

[2] Luo W., Deng Y., Wu P., Zhao Q., Shi L., “Quantization-Aware ANN-to-SNN Conversion for High-Fidelity Object Detection,” Frontiers in Neuroscience, vol. 17, art. 1141701, Apr. 2023.

[3] Choi Y. S., Kang D. W., “A Survey of Surrogate-Gradient Training Methods for Spiking Neural Networks,” IEEE Trans. Neural Netw. Learn. Syst., vol. 35, no. 1, pp. 58 - 75, Jan. 2024.

[4] Kim G. H., Oh M. S., “Neuromorphic Object Detection for Low-Power Embedded Devices,” IEEE Embedded Systems Letters, vol. 16, no. 3, pp. 78 - 82, Sep. 2024.