

이벤트 데이터의 전처리 과정이 SNN의 학습 성능에 미치는 영향 분석

강희범^{1,3+0} 원현호^{2,3+} 박성식³

¹ 서울과학기술대학교 전기정보공학과

² 서울과학기술대학교 전자 IT 미디어공학과

³ 한국과학기술연구원 인공뇌융합연구단

hubo1024@gmail.com, dnjsgusgh@naver.com, seong.sik.park@kist.re.kr

An Analysis of Preprocessing Effect on SNN Training with Event Data

Huibeom Kang^{1,3+0} Hyunho Won^{2,3+} Seongsik Park³

¹Department of Electrical and Information Engineering, Seoul National University of Science and Technology

²Department of Electronic and IT Media Engineering, Seoul National University of Science and Technology

³Center for Neuromorphic Engineering, Korea Institute of Science and Technology

요 약

최근 높은 에너지 효율로 기대되는 spiking neural network (SNN) 모델에 대한 연구가 증가하고 있다. 이러한 SNN은 희소한 시공간 데이터인 이벤트 데이터를 처리하는 데 있어 효율적인 것으로 예상되고 있다. 이벤트 데이터를 활용하기 위해서는 전처리가 필요하며, 현재 voxel grid 기반의 전처리 방법이 많이 활용되고 있다. 본 논문에서는 voxel grid 기반 전처리의 중요한 요소인 time interval과 이벤트 누적 방법이 SNN의 학습 시간과 정확도에 미치는 영향을 알아본다.

1. 서 론

최근 기존 deep neural network (DNN)에 비해 높은 에너지 효율을 가질 것으로 기대되는 spiking neural network (SNN)에 대한 연구가 증가하고 있다. SNN은 적은 스파이크만으로 동작하기 때문에 소모 전력이 적어 에너지 효율이 좋고 시공간 데이터를 처리하는 데에 좋은 성능을 보일 것으로 예상된다 [1]. 하지만 스파이킹 함수가 미분이 불가능하여 DNN의 학습 방법인 역전파 알고리즘을 사용하여 학습하는 데 문제가 있다. 이를 해결하기 위해 surrogate gradient 방식을 사용하여 SNN을 학습하는 방법이 제안되었다 [2].

시공간 데이터의 일종인 이벤트 데이터는 이벤트 카메라를 이용하여 측정된다. 이벤트 카메라는 일반적인 프레임 기반 카메라와 다르게 이미지 픽셀의 밝기 변화를 감지하여 이벤트를 생성한다 [2]. 이때 밝기가 증가하였을 경우 양극성, 감소하였을 때는 음극성 이벤트를 생성한다. 이벤트 데이터는 시간 해상도가 높다는 장점이 있지만 사용하기 위해서는 전처리가 필요하다.

시간 해상도가 높은 이벤트 데이터는 시공간 데이터에 우수한 성능을 보이는 SNN에 적합할 것으로 기대되고 있다. 하지

만 이벤트 데이터는 전처리가 필요하고 전처리는 SNN 학습에 지대한 영향을 미치기 때문에 이벤트 데이터의 전처리 방법이 SNN의 학습 성능에 미치는 영향을 알아볼 필요가 있다. 따라서 본 논문에서는 현재 가장 많이 사용되고 있는 voxel grid 기반 전처리의 주요한 변수인 time interval(T_t)과 이벤트 누적 방법에 따른 학습 성능 변화를 분석하였다.

2. 전처리 방법

Voxel grid는 이벤트 데이터의 시간적 정보가 기록되는 3차원 구조이다 [3]. Voxel grid 기반의 전처리 과정은 아래의 순서로 수행된다. 먼저 식 $T'_t = T_t/T_i$ 와 같이 raw 이벤트 데이터의 time step길이(T_t)를 T_i 단위로 나누어 이벤트 데이터를 출력 time step(T'_t)개의 그룹으로 형성한다. 각 그룹의 이벤트 데이터는 시간(t)축을 기준으로 누적되어 출력된다.

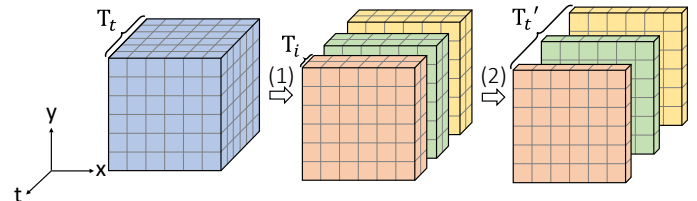


그림 1. 전처리 과정; (1) T'_t 개의 그룹으로 나뉘는 단계, (2)형성된 그룹별로 누적하는 단계

⁺ 해당 저자들은 기여도가 동일함

* 이 연구는 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원(NRF-2021R1C1C2010454)과, 한국과학기술연구원(KIST) 주요 사업인 '인공뇌융합연구사업(2E31550)'의 지원을 받아 수행되었음

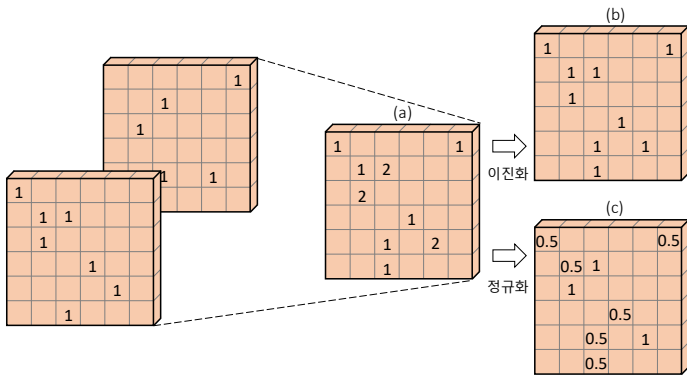


그림 2. 누적 방법; (a)이벤트 데이터를 픽셀 별로 누적하는 방법, (b)누적된 데이터를 이진화하는 방법, (c)누적된 데이터를 정규화하는 방법

T_i 가 커짐에 따라 시간적 정보가 손실되어 정확도가 낮아질 수 있고 동시에 T'_i 가 줄어들어 학습 시간과 계산량이 줄어든다. 본 논문에서는 T_i 가 학습 시간과 정확도에 미치는 영향을 실험을 통해 확인하였다.

이벤트 데이터들을 t축을 기준으로 누적하는 방법은 다양하다 [4]. 본 논문에서는 (a)이벤트 데이터를 픽셀별로 누적, (b)이벤트 데이터를 픽셀별로 누적 후 이진화, (c)이벤트 데이터를 픽셀 별로 누적 후 정규화까지 총 3가지 방법으로 분석한다. 이벤트 누적 방법에 따라 출력 데이터 값이 달라지기 때문에 정확도와 스파이크 수에 영향을 줄 수 있다. 본 논문에서는 이벤트 누적 방법이 정확도와 스파이크 수에 미치는 영향을 실험을 통해 확인하였다.

3. 실험

3.1 실험 환경

데이터 세트는 N-MNIST를 사용하였다. N-MNIST는 MNIST 데이터 세트를 이벤트 형식으로 변환한 데이터 세트이다 [5]. 옵티마이저는 MNIST 학습에서 우수한 성능을 보이는 NAdam을 사용하였다 [6].

이를 바탕으로 두가지 실험 환경을 세팅하였다. T_i 가 학습 성능에 미치는 영향을 확인하기 위해 첫 번째 실험 환경은 그림 2의 누적 방법(b)를 사용하고 T_i 는 100,000, 75,000, 50,000, 10,000, 5,000, 2,500, 1,000으로 설정하였다. 이벤트 누적 방법에 따른 학습 성능을 확인하기 위해 두 번째 실험 환경은 이벤트 데이터 누적 방법 3가지((a), (b), (c))를 사용하고 T_i 는 50,000, 10,000, 2,500으로 설정하였다.

3.2 실험 결과

그림 3은 첫 번째 실험 환경에서 모델이 충분히 수렴할 수 있도록 15 epoch(에포크)를 동일하게 진행한 결과이다. 테스트 정확도는 T_i 가 75,000인 경우 19.96%가 나오고 이후 급격히 증가하여 T_i 가 10,000인 경우 93.91%를 달성한다. T_i 가 10,000인 경우부터는 완만하게 증가하여 T_i 가 1,000에서 96.62%의 정확도를 보인다. T_i 가 1,000일 때의 정확도가 T_i 가 2,500일 때의 정확도보다 0.95% 감소한 이유는 옵티마이저가

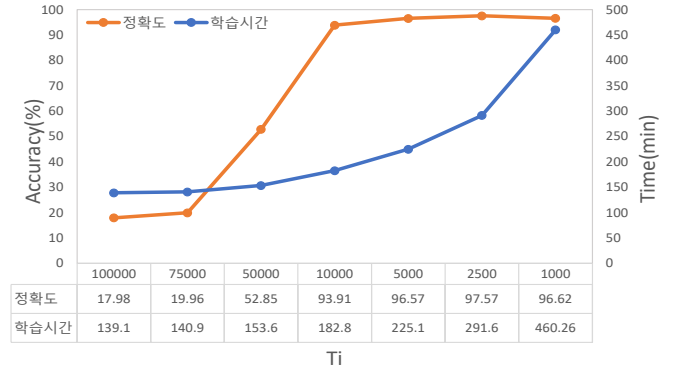


그림 3. 15 에포크 기준 테스트 정확도와 학습 시간

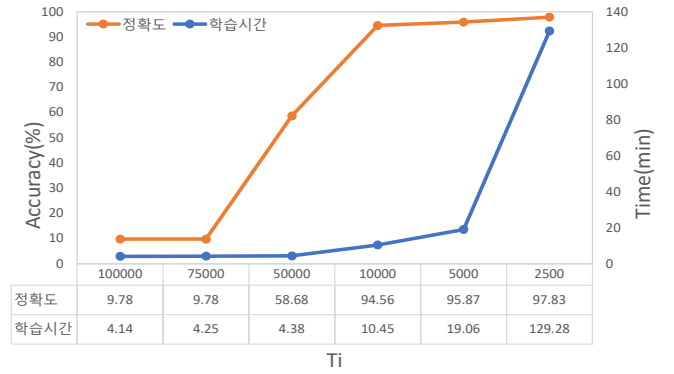


그림 4. 평균 손실 값 0.5 이하 기준 테스트 정확도와 학습 시간

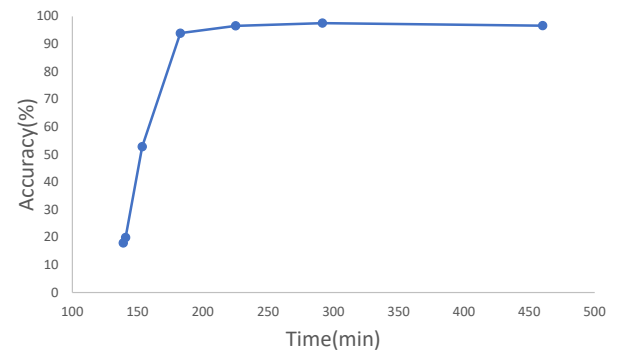


그림 5. 15 에포크 기준 학습 시간에 따른 테스트 정확도

더 이상 손실 값을 수렴시키지 못했기 때문으로 예상된다. 학습 시간은 T_i 가 증가함에 따라 점차 증가하였다.

그림 4는 첫 번째 실험 환경에서 0.5 에포크 마다 평균 손실 값이 0.5 이하일 경우 훈련을 조기 중단하였을 때의 결과이다. 테스트 정확도는 T_i 가 75,000일 때 9.78%를 보이고 이후 급격히 증가하여 T_i 가 10,000인 경우 94.56%를 달성한다. T_i 가 100,000일 때와 T_i 가 7,500일 때 9.78%의 정확도를 보이는 이유는 출력 스파이크가 생성되지 않았기 때문이다. T_i 가 10,000일 때부터 완만하게 증가하여 T_i 가 2,500일 때 97.83%를 달성하였다. T_i 가 1,000일 때는 손실 값이 0.5 이하로 수렴하지 않아 결과를 확인하지 못하였다. 학습 시간은 T_i 가 5,000에서 2,500으로 넘어갈 때 급격히 증가하는 모습을 보였다.

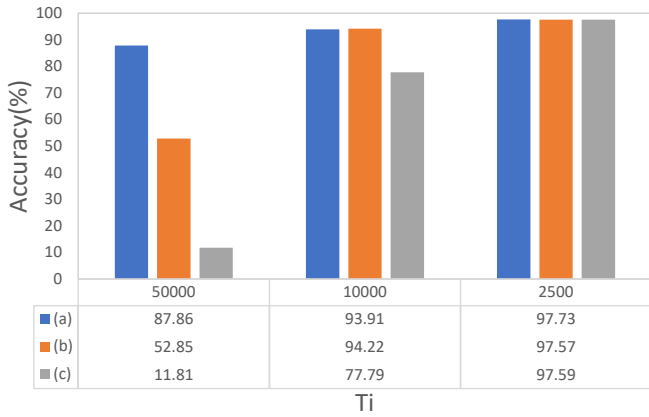


그림 6. 이벤트 데이터 누적 방법에 따른 정확도; (a)이벤트 데이터를 픽셀 별로 누적하는 방법, (b) 누적된 데이터를 이진화하는 방법, (c) 누적된 데이터를 정규화하는 방법

그림 5는 그림 3의 학습 시간에 따른 테스트 정확도를 나타낸 그래프이다. 이는 일반적인 DNN 테스트 정확도의 그래프 양상을 보여준다. 학습 시간 180분까지 정확도가 급격히 증가하고 이후 완만하게 증가하는 것을 볼 수 있다.

그림 6은 두 번째 실험 환경에서 진행한 정확도 결과이다. 누적 방법(a)는 T_i 가 50,000일 때 87.86%로 다른 누적 방법에 비해 상대적으로 높은 정확도를 보여주었다. 누적 방법(b)는 T_i 가 10,000일 때 다른 누적 방법보다 94.22%로 높은 정확도를 보여주었다. 누적 방법(c)는 T_i 가 2,500이 되어서야 97.59%로 다른 누적 방법과 동일한 수준의 정확도를 보여주었다.

누적 방법(a)는 세 가지 T_i 에서 안정적인 정확도를 보여 가장 T_i 영향을 적게 받는 것으로 보인다. 이는 누적된 이벤트 데이터에서 정규화와 이진화와 같은 변형이 이루어지지 않았기 때문으로 추측된다.

누적 방법(b)는 이진화를 하기 때문에 이벤트 수가 줄어든다. 임의로 원본 데이터의 이벤트 수를 추출해본 결과 원본 데이터의 이벤트 수는 1,975개이고 T_i 가 10,000일 때 이벤트 수는 1,951개로 큰 변화가 없었다. 그러나 T_i 가 50,000일 때 이벤트 수는 1,387개로 크게 줄어들었다. 이는 T_i 가 10,000인 경우를 기점으로 T_i 가 커질수록 정보 손실이 커지는 것으로 보인다. 이로 인해 T_i 가 50,000인 경우 낮은 정확도를 보인 것으로 판단된다.

누적 방법(c)가 다른 누적 방법과 동일한 수준의 정확도를 보여주지 못한 이유는 정규화를 하면서 스파이크 임계점을 넘지 못하는 이벤트 수가 많아지고 이로 인해 정보 손실이 발생했기 때문으로 보인다.

그림 7은 두 번째 실험 환경에서 진행한 스파이크 결과이다. 누적 방법에 따라 상이했던 정확도와 달리 스파이크 수는 누적 방법에 크게 영향을 받지 않고 T_i 에 의존적인 것으로 보인다.

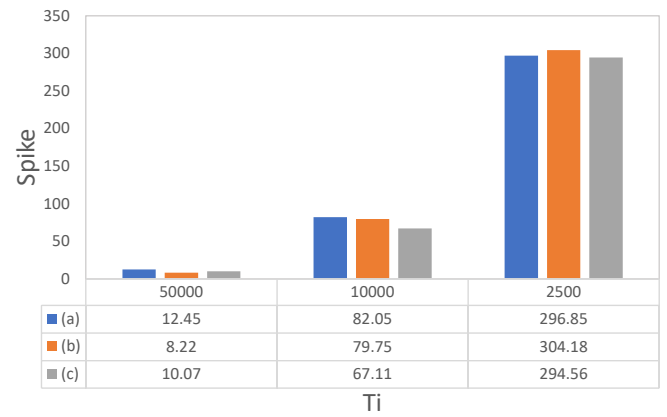


그림 7. 이벤트 데이터 누적 방법에 따른 스파이크 수; (a)이벤트 데이터를 픽셀 별로 누적하는 방법, (b) 누적된 데이터를 이진화하는 방법, (c) 누적된 데이터를 정규화하는 방법

4. 결 론

본 논문에서는 voxel grid 기반의 이벤트 데이터 전처리 과정에서 중요한 요소인 T_i 와 이벤트 데이터 누적 방법에 따른 SNN 학습 성능을 분석하는 실험을 진행하였다. 5,000에서 10,000 사이의 T_i 에서 높은 정확도와 적절한 학습 시간을 보여주었고 이벤트 데이터를 픽셀별로 누적하는 방법이 학습에서 상대적으로 안정적인 정확도를 보여주었다. SNN의 에너지 소모량과 비례하는 스파이크 수는 이벤트 데이터 누적 방법과 상관없이 T_i 에 따라 반비례하였다. 따라서 SNN의 에너지 소모량은 이벤트 데이터 누적 방법보다는 T_i 에 의존적일 것으로 예상된다. 본 연구에서는 상대적으로 간단한 N-MNIST 데이터 세트로만 SNN 훈련을 진행하였다는 한계가 있다. 후속 연구에서는 다양한 이벤트 데이터 세트에서도 동일한 방식의 실험을 진행해 볼 계획이다. 또한, 다른 이벤트 전처리 방식이 SNN의 학습 성능에 미치는 영향에 대해서도 분석할 계획이다.

참고 문헌

- [1] A. Tavanaei, et al., "Deep learning in spiking neural networks," *Neural Networks*, 2018
- [2] E. O. Neftci, H. Mostafa, F. Zenke, "Surrogate Gradient Learning in Spiking Neural Networks", *IEEE Signal Processing Magazine*, 2019
- [3] G. Gallego, et al., "Event-based Vision: A Survey," *IEEE TPAMI*, 2020
- [4] P. Bardow, et al., "Simultaneous optical flow and intensity estimation from an event camera," in *Proc. ICCV*, 2016
- [5] E. Stamatias, et al., "An Event-Driven Classifier for Spiking Neural Networks Fed with Synthetic or Dynamic Vision Sensor Data," in *Front. Neurosci.*, 2017
- [6] 주기현, 박치현, 임현승, "기계학습 오픈마이저 성능 평가," in *전기전자학회논문지*, 2020