(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2023-0114149
(43) 공개일자 2023년08월01일

(51) 국제특허분류(Int. Cl.)
G06N 3/04 (2023.01) *G06N 3/063* (2023.01)
G06N 3/08 (2023.01)
(52) CPC특허분류
G06N 3/049 (2023.01)
G06N 3/063 (2013.01)
(21) 출원번호 10-2022-0031156
(22) 출원일자 2022년03월14일
심사청구일자 2022년03월14일
(30) 우선권주장
1020220009753 2022년01월24일 대한민국(KR)

(71) 출원인
고려대학교 산학협력단
서울특별시 성북구 안암로 145, 고려대학교 (안암동5가)
(72) 발명자
박종선
서울특별시 서초구 신반포로 32
최성현
서울특별시 송파구 양재대로 1218(방이동, 올림픽선수기자촌아파트)
유동우
서울특별시 동대문구 안암로18가길 28(제기동)
(74) 대리인
특허법인주연케이알피

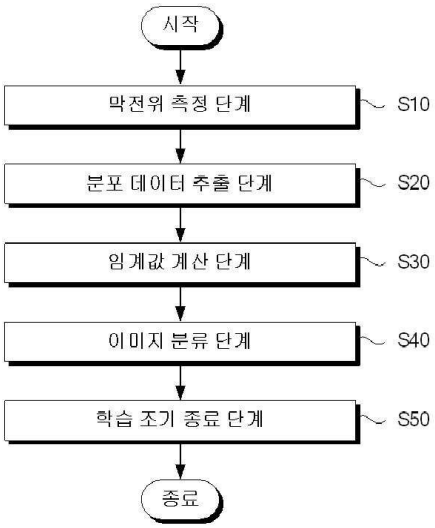
전체 청구항 수 : 총 13 항

(54) 발명의 명칭 **스파이킹 뉴럴 네트워크의 타임스텝 세분화 기반의 조기 판단 학습 장치 및 그 동작방법**

(57) 요 약

본 발명은 스파이킹 뉴럴 네트워크의 타임스텝 세분화 기반의 조기 판단 학습 장치 및 그 동작방법에 관한 것이다. 본 발명의 일 실시예에 있어서, 타임스텝 세분화 기반의 조기 판단 학습 가속 방법은 타임스텝을 세분화하는 타임스텝 세분화 단계, 현재 학습 과정 동안의 세분화된 타임스텝 별 제1 및 제2 막전위를 측정하는 막전위 측정 단계, 상기 제1 및 제2 막전위에 기초하여 다음 학습 과정에서 사용될 임계값을 계산하는 임계값 계산 단계, 및 상기 세분화된 타임스텝에서 상기 제1 및 제2 막전위의 차이가 상기 임계값보다 큰 경우, 상기 이미지에 상기 학습 기여도가 없다고 판단하여 상기 세분화된 타임스텝에서의 학습을 종료하는 학습 조기 종료 단계를 포함할 수 있다.

대 표 도 - 도2



(52) CPC특허분류
G06N 3/08 (2023.01)

이 발명을 지원한 국가연구개발사업
과제고유번호 1711128700
과제번호 2020R1A2C3014820
부처명 과학기술정보통신부
과제관리(전문)기관명 한국연구재단
연구사업명 개인기초연구(과기정통부)(R&D)
연구과제명 온-디바이스 개인화를 위한 에너지 효율적인 딥러닝 가속기 설계
기 여 율 1/1
과제수행기관명 고려대학교
연구기간 2021.03.01 ~ 2022.02.28

명 세 서

청구범위

청구항 1

타임스텝을 세분화하는 타임스텝 세분화 단계;

현재 학습 과정 동안의 세분화된 타임스텝 별 제1 및 제2 막전위를 측정하는 막전위 측정 단계;

상기 제1 및 제2 막전위에 기초하여 다음 학습 과정에서 사용될 임계값을 계산하는 임계값 계산 단계; 및

상기 세분화된 타임스텝에서 상기 제1 및 제2 막전위의 차이가 상기 임계값보다 큰 경우, 상기 이미지에 상기 학습 기여도가 없다고 판단하여 상기 세분화된 타임스텝에서의 학습을 종료하는 학습 조기 종료 단계를 포함하는, 타임스텝 세분화 기반의 조기 판단 학습 가속 방법.

청구항 2

제1항에 있어서,

상기 제1 막전위는 정답 뉴런의 막전위이고, 상기 제2 막전위는 복수 개의 막전위 중 상기 정답 뉴런의 막전위 값을 제외한 막전위 중 가장 큰 막전위인, 타임스텝 세분화 기반의 조기 판단 학습 가속 방법.

청구항 3

제2항에 있어서,

상기 임계값 계산 단계는 수학식 1에 기초하여 상기 임계값을 계산하고,

상기 수학식 1은 $a=m+2\sigma$이고,

상기 a는 상기 임계값이고, 상기 m은 상기 학습 기여도가 있는 이미지의 막전위 차 분포의 평균이며, 상기 σ는 상기 학습 기여도가 있는 이미지의 막전위 차 분포의 편차인, 타임스텝 세분화 기반의 조기 판단 학습 가속 방법.

청구항 4

제2항에 있어서,

상기 학습 조기 종료 단계는 수학식 2에 기초하여 상기 학습 기여도가 없는 이미지와 상기 학습 기여도가 있는 이미지의 경계를 설정하고,

상기 수학식 2는 $y=x-a$이고,

상기 a는 상기 이전 학습 과정에서 계산된 임계값이고, 상기 x는 상기 정답 뉴런의 막전위이고, 상기 y는 상기 복수 개의 막전위 중 상기 정답 뉴런의 막전위 값을 제외한 막전위 중 가장 큰 막전위인, 타임스텝 세분화 기반의 조기 판단 학습 가속 방법.

청구항 5

제4항에 있어서,

상기 학습 조기 종료 단계는 상기 수학식 2로부터 변형된 수학식 3에 기초하여 상기 학습 기여도가 없는 이미지를 판단하고,

상기 수학식 3은 $x-y \geq a$이고,

상기 a는 상기 이전 학습 과정에서 계산된 임계값이고, 상기 x는 상기 정답 뉴런의 막전위이고, 상기 y는 상기 복수 개의 막전위 중 상기 정답 뉴런의 막전위 값을 제외한 막전위 중 가장 큰 막전위인, 타임스텝 세분화 기반의 조기 판단 학습 가속 방법.

청구항 6

제5항에 있어서,

상기 학습 조기 종료 단계는 상기 세분화된 타임스텝 별 상기 제1 및 제2 막전위, 및 상기 임계값을 상기 수학식 3에 대입하여 상기 수학식 3이 성립될 경우에 학습 기여도가 없는 이미지로 판단하여 상기 세분화된 타임스텝에서의 학습을 종료하고,

상기 세분화된 타임스텝 별 상기 제1 및 제2 막전위, 및 상기 임계값을 상기 수학식 3에 대입하여 상기 수학식 3이 성립되지 않을 경우에 학습 기여도가 있는 이미지로 판단하여 학습을 진행하는, 타임스텝 세분화 기반의 조기 판단 학습 가속 방법.

청구항 7

제1항에 있어서,

상기 세분화된 타임스텝은 하나의 타임스텝을 2개 내지 16개 중 하나로 세분화한 타임스텝인, 타임스텝 세분화 기반의 조기 판단 학습 가속 방법.

청구항 8

스파이킹 뉴럴 네트워크의 입력 스파이크 신호가 입력되는 입력 레이어 모듈;

상기 입력 스파이크 신호를 입력받는 히든 레이어 모듈;

상기 히든 레이어 모듈로부터 입력 스파이크 신호를 전달받고, 학습 기여도가 없는 이미지를 판단하며, 상기 학습 기여도의 유무를 판단하기 위한 임계값을 계산하는 출력 레이어 모듈; 및

상기 출력 레이어 모듈이 판단한 상기 학습 기여도의 유무에 기초하여 학습 과정을 종료하는 글로벌 컨트롤러를 포함하는, 타임스텝 세분화 기반의 조기 판단 학습 가속 장치.

청구항 9

제8항에 있어서,

상기 히든 레이어 모듈은,

상기 입력 레이어 모듈로부터 입력 스파이크 신호를 전달받고, 입력 스파이크에 기초하여 복수 개의 뉴런들의 막전위를 계산하는 막전위 업데이트 모듈;

상기 입력 스파이크 신호의 입력에 기초하여 상기 입력 스파이크들의 가중치들을 상기 막전위에 더하는 가중치 업데이트 모듈;

상기 복수 개의 뉴런들의 막전위를 저장하는 막전위 버퍼; 및

상기 복수 개의 뉴런들의 스파이크 발생 시간을 저장하는 스파이크 시간 버퍼를 포함하는, 타임스텝 세분화 기반의 조기 판단 학습 가속 장치.

청구항 10

제8항에 있어서,

상기 출력 레이어 모듈은,

상기 히든 레이어 모듈로부터 상기 입력 스파이크 신호를 전달받는 스파이크 버퍼;

상기 스파이크 버퍼로부터 입력 스파이크 신호를 전달받고, 입력 스파이크에 기초하여 복수 개의 뉴런들의 막전위를 계산하는 막전위 업데이트 모듈;

상기 입력 스파이크 신호의 입력에 기초하여 상기 입력 스파이크들의 가중치들을 상기 막전위에 더하는 가중치 업데이트 모듈;

상기 복수 개의 뉴런들의 막전위를 저장하는 막전위 버퍼;

상기 복수 개의 뉴런들의 스파이크 발생 시간을 저장하는 스파이크 시간 버퍼;

학습 과정 중 순전파 과정이 모두 끝난 후의 상기 복수 개의 뉴런들의 스파이크 발생 시간과 목표하는 정답 신호와의 차이를 계산하는 에러 계산부;

상기 순전파 과정에서 각 타임스텝의 연산이 끝났을 때, 상기 학습 기여도가 없는 이미지를 판단하는 조기 판단부; 및

상기 순전파 과정이 모두 끝난 후에 다음 학습에서 사용될 임계값을 계산하는 임계값 계산부를 포함하는, 타임스텝 세분화 기반의 조기 판단 학습 가속 장치.

청구항 11

제10항에 있어서,

상기 조기 판단부는 타임스텝에서 제1 및 제2 막전위의 차이가 상기 임계값보다 큰 경우, 상기 이미지에 상기 학습 기여도가 없다고 판단하여 상기 타임스텝에서의 학습을 종료하는, 타임스텝 세분화 기반의 조기 판단 학습 가속 장치.

청구항 12

제10항에 있어서,

상기 임계값 계산부는 상기 복수 개의 뉴런들의 막전위 중 제1 및 제2 막전위의 분포 데이터에 기초하여 다음 학습 과정에서 사용될 상기 임계값을 계산하는, 타임스텝 세분화 기반의 조기 판단 학습 가속 장치.

청구항 13

제12항에 있어서,

상기 제1 막전위는 정답 뉴런의 막전위이고, 상기 제2 막전위는 복수 개의 막전위 중 상기 정답 뉴런의 막전위 값을 제외한 막전위 중 가장 큰 막전위인, 타임스텝 세분화 기반의 조기 판단 학습 가속 장치.

발명의 설명

기 술 분 야

[0001] 본 발명은 스파이킹 뉴럴 네트워크의 타임스텝 세분화 기반의 조기 판단 학습 장치 및 그 동작방법에 관한 것이다.

배 경 기 술

[0002] 스파이킹 뉴럴 네트워크는 0과 1의 spike 단위로 정보를 입력받고, 전달하므로 하드웨어 구현에 있어 저전력의 장점이 있다. 또한 temporal coding의 도입으로 추론 과정의 spike 수가 감소해 추론 과정의 에너지를 크게 감소시켰다. 하지만 뉴럴 네트워크의 학습은 매우 많은 수의 학습 이미지에 대해 추론, 역전파, 가중치 업데이트의 학습 과정이 반복되기 때문에 에너지 소모와 시간이 매우 크다는 문제점이 있다.

발명의 내용

해결하려는 과제

[0003] 본 발명은 타임스텝을 세분화하고 학습 이미지의 네트워크 기여도를 판단하여 학습 기여도가 없는 이미지 학습의 조기 종료를 통해 학습 에너지와 시간을 감소시킬 수 있는 타임스텝 세분화 기반의 조기 판단 학습 방법은 제공하는 것에 목적이 있다.

과제의 해결 수단

[0004] 본 발명의 일 실시예에 있어서, 타임스텝 세분화 기반의 조기 판단 학습 가속 방법은 타임스텝을 세분화하는 타임스텝 세분화 단계, 현재 학습 과정 동안의 세분화된 타임스텝 별 제1 및 제2 막전위를 측정하는 막전위 측정

단계, 상기 제1 및 제2 막전위에 기초하여 다음 학습 과정에서 사용될 임계값을 계산하는 임계값 계산 단계, 및 상기 세분화된 타임스텝에서 상기 제1 및 제2 막전위의 차이가 상기 임계값보다 큰 경우, 상기 이미지에 상기 학습 기여도가 없다고 판단하여 상기 세분화된 타임스텝에서의 학습을 종료하는 학습 조기 종료 단계를 포함할 수 있다.

[0005] 본 발명의 일 실시예에 있어서, 상기 제1 막전위는 정답 뉴런의 막전위이고, 상기 제2 막전위는 복수 개의 막전위 중 상기 정답 뉴런의 막전위 값을 제외한 막전위 중 가장 큰 막전위일 수 있다.

[0006] 본 발명의 일 실시예에 있어서, 상기 임계값 계산 단계는 수학식 1에 기초하여 상기 임계값을 계산하고, 상기 수학식 1은 $a=m+2\sigma$이고, 상기 a는 상기 임계값이고, 상기 m은 상기 학습 기여도가 있는 이미지의 막전위 차 분포의 평균이며, 상기 σ는 상기 학습 기여도가 있는 이미지의 막전위 차 분포의 편차일 수 있다.

[0007] 본 발명의 일 실시예에 있어서, 상기 학습 조기 종료 단계는 수학식 2에 기초하여 상기 학습 기여도가 없는 이미지와 상기 학습 기여도가 있는 이미지의 경계를 설정하고, 상기 수학식 2는 $y=x-a$이고, 상기 a는 상기 이전 학습 과정에서 계산된 임계값이고, 상기 x는 상기 정답 뉴런의 막전위이고, 상기 y는 상기 복수 개의 막전위 중 상기 정답 뉴런의 막전위 값을 제외한 막전위 중 가장 큰 막전위일 수 있다.

[0008] 본 발명의 일 실시예에 있어서, 상기 학습 조기 종료 단계는 상기 수학식 2로부터 변형된 수학식 3에 기초하여 상기 학습 기여도가 없는 이미지를 판단하고, 상기 수학식 3은 $x-y\geq a$이고, 상기 a는 상기 이전 학습 과정에서 계산된 임계값이고, 상기 x는 상기 정답 뉴런의 막전위이고, 상기 y는 상기 복수 개의 막전위 중 상기 정답 뉴런의 막전위 값을 제외한 막전위 중 가장 큰 막전위일 수 있다.

[0009] 본 발명의 일 실시예에 있어서, 상기 학습 조기 종료 단계는 상기 세분화된 타임스텝 별 상기 제1 및 제2 막전위, 및 상기 임계값을 상기 수학식 3에 대입하여 상기 수학식 3이 성립될 경우에 학습 기여도가 없는 이미지로 판단하여 상기 세분화된 타임스텝에서의 학습을 종료하고, 상기 세분화된 타임스텝 별 상기 제1 및 제2 막전위, 및 상기 임계값을 상기 수학식 3에 대입하여 상기 수학식 3이 성립되지 않을 경우에 학습 기여도가 있는 이미지로 판단하여 학습을 진행할 수 있다.

[0010] 본 발명의 일 실시예에 있어서, 상기 세분화된 타임스텝은 하나의 타임스텝을 2개 내지 16개 중 하나로 세분화한 타임스텝일 수 있다.

[0011] 본 발명의 일 실시예에 따른 타임스텝 세분화 기반의 조기 판단 학습 가속 장치는 스파이킹 뉴럴 네트워크의 입력 스파이크 신호가 입력되는 입력 레이어 모듈, 상기 입력 스파이크 신호를 입력받는 히든 레이어 모듈, 상기 히든 레이어 모듈로부터 입력 스파이크 신호를 전달받고, 학습 기여도가 없는 이미지를 판단하며, 상기 학습 기여도의 유무를 판단하기 위한 임계값을 계산하는 출력 레이어 모듈, 및 상기 출력 레이어 모듈이 판단한 상기 학습 기여도의 유무에 기초하여 학습 과정을 종료하는 글로벌 컨트롤러를 포함할 수 있다.

[0012] 본 발명의 일 실시예에 있어서, 상기 히든 레이어 모듈은, 상기 입력 레이어 모듈로부터 입력 스파이크 신호를 전달받고, 입력 스파이크에 기초하여 복수 개의 뉴런들의 막전위를 계산하는 막전위 업데이트 모듈, 상기 입력 스파이크 신호의 입력에 기초하여 상기 입력 스파이크들의 가중치들을 상기 막전위에 더하는 가중치 업데이트 모듈, 상기 복수 개의 뉴런들의 막전위를 저장하는 막전위 버퍼, 및 상기 복수 개의 뉴런들의 스파이크 발생 시간을 저장하는 스파이크 시간 버퍼를 포함할 수 있다.

[0013] 본 발명의 일 실시예에 있어서, 상기 출력 레이어 모듈은, 상기 히든 레이어 모듈로부터 상기 입력 스파이크 신호를 전달받는 스파이크 버퍼, 상기 스파이크 버퍼로부터 입력 스파이크 신호를 전달받고, 입력 스파이크에 기초하여 복수 개의 뉴런들의 막전위를 계산하는 막전위 업데이트 모듈, 상기 입력 스파이크 신호의 입력에 기초하여 상기 입력 스파이크들의 가중치들을 상기 막전위에 더하는 가중치 업데이트 모듈, 상기 복수 개의 뉴런들의 막전위를 저장하는 막전위 버퍼, 상기 복수 개의 뉴런들의 스파이크 발생 시간을 저장하는 스파이크 시간 버퍼, 학습 과정 중 순전파 과정이 모두 끝난 후의 상기 복수 개의 뉴런들의 스파이크 발생 시간과 목표하는 정답 신호와의 차이를 계산하는 에러 계산부, 상기 순전파 과정에서 각 타임스텝의 연산이 끝났을 때, 상기 학습 기여도가 없는 이미지를 판단하는 조기 판단부, 및 상기 순전파 과정이 모두 끝난 후에 다음 학습에서 사용될 임계값을 계산하는 임계값 계산부를 포함할 수 있다.

[0014] 본 발명의 일 실시예에 잇어서, 상기 조기 판단부는 타임스텝에서 제1 및 제2 막전위의 차이가 상기 임계값보다 큰 경우, 상기 이미지에 상기 학습 기여도가 없다고 판단하여 상기 타임스텝에서의 학습을 종료할 수 있다.

[0015] 본 발명의 일 실시예에 있어서, 상기 임계값 계산부는 상기 복수 개의 뉴런들의 막전위 중 제1 및 제2 막전위의

분포 데이터에 기초하여 다음 학습 과정에서 사용될 상기 임계값을 계산할 수 있다.

[0016] 본 발명의 일 실시예에 있어서, 상기 제1 막전위는 정답 뉴런의 막전위이고, 상기 제2 막전위는 복수 개의 막전위 중 상기 정답 뉴런의 막전위 값을 제외한 막전위 중 가장 큰 막전위일 수 있다.

발명의 효과

[0017] 본 발명의 타임스텝 세분화 기반의 조기 판단 학습 방법은 타임스텝을 세분화하고 학습 이미지의 네트워크 기여도를 판단하여 학습 기여도가 없는 이미지 학습의 조기 종료를 통해 학습 에너지와 시간을 감소시킬 수 있다.

도면의 간단한 설명

[0018] 도 1은 학습 과정에 따른 학습 기여도가 없는 이미지의 비율을 보여주는 그래프이다.

도 2는 본 발명의 일 실시예에 따른 타임스텝 세분화 기반의 조기 판단 학습 가속 방법을 보여주는 순서도이다.

도 3은 본 발명의 일 실시예에 따른 타임스텝 세분화 단계의 예를 보여주는 도이다.

도 4는 본 발명의 일 실시예에 따른 임계값을 계산하는 방법을 보여주는 도이다.

도 5는 본 발명의 일 실시예에 따른 학습 기여도가 없는 이미지를 판단하여 학습을 조기 종료하는 방법을 보여주는 도이다.

도 6은 본 발명의 일 실시예에 따른 세분화된 타임스텝에 대한 학습 조기 종료의 예를 보여주는 도이다.

도 7은 본 발명의 일 실시예에 따른 타임스텝 세분화 기반의 조기 판단 학습 가속 장치를 보여주는 도이다.

발명을 실시하기 위한 구체적인 내용

[0019] 이하, 본 출원의 기술적 사상을 본 출원의 기술분야에서 통상의 지식을 가진자가 용이하게 실시할 수 있을 정도로 상세히 설명하기 위하여, 본 출원의 실시 예를 첨부된 도면을 참조하여 설명하기로 한다.

[0020] 도 1은 학습 과정에 따른 학습 기여도가 없는 이미지의 비율을 보여주는 그래프이다.

[0021] 도 1을 참조하면, 복수 개의 데이터셋들을 이용하여 학습 과정에 따른 학습 기여도가 없는 이미지의 비율을 측정할 수 있다. 이때, x축은 학습 과정이고, y축은 복수 개의 데이터셋들 별로 학습 데이터 중 학습에 기여하지 않는 이미지의 비율일 수 있다. 복수 개의 데이터셋들은 MNIST, Fashion-MNIST, 및 ETH-80일 수 있다.

[0022] MNIST 데이터셋의 경우, 학습 세트의 크기가 60000개이고, 학습 과정에 따른 학습 기여도가 없는 이미지의 비율이 0.5라면 30000개의 학습 기여도가 없는 이미지가 존재할 수 있다.

[0023] Fashion-MNIST 데이터셋의 경우, 학습 세트의 크기가 60000개이고, 학습 과정에 따른 학습 기여도가 없는 이미지의 비율이 0.5라면 30000개의 학습 기여도가 없는 이미지가 존재할 수 있다.

[0024] ETH-80 데이터셋의 경우, 전체 3280개의 이미지로 이루어져 있다. 예를 들어, 이 중 80%인 2624개의 이미지가 학습 세트로 사용될 수 있고, 20%인 656개의 이미자가 추론 세트로 사용될 수 있다.

[0025] 상술한 바와 같이, 종래의 기술에 기반한 스파이킹 뉴럴 네트워크의 학습 과정은 기여도가 없는 이미지의 비율이 크고, 학습이 진행됨에 따라 기여도가 없는 이미지의 양이 더욱 증가하는 것을 알 수 있다. 이와 같이, 학습 과정 중 기여도가 없는 이미지의 양이 증가할 경우 에너지 소모량이 증가하고 연산 시간이 오래 걸린다는 문제점이 있다.

[0026] 도 2는 본 발명의 일 실시예에 따른 타임스텝 세분화 기반의 조기 판단 학습 가속 방법을 보여주는 순서도이다.

[0027] S10 단계에서, 타임스텝 세분화 기반의 조기 판단 학습 가속 장치는 타임스텝을 세분화할 수 있다. 예를 들어, 타임스텝 세분화 기반의 조기 판단 학습 가속 장치는 하나의 타임스텝을 2개 내지 16개 중 하나로 세분화할 수 있다.

[0028] S20 단계에서, 타임스텝 세분화 기반의 조기 판단 학습 가속 장치는 막전위를 측정할 수 있다. 예를 들어, 타임스텝 세분화 기반의 조기 판단 학습 가속 장치는 현재 학습 과정 동안의 세분화된 타임스텝 별 제1 및 제2 막전위를 측정할 수 있다.

[0029] 이때, 제1 막전위는 정답 뉴런의 막전위이고, 제2 막전위는 복수 개의 막전위 중 정답 뉴런의 막전위 값을 제외

한 막전위 중 가장 큰 막전위일 수 있다.

[0030] S30 단계에서, 타임스텝 세분화 기반의 조기 판단 학습 가속 장치는 임계값을 계산할 수 있다. 예를 들어, 타임스텝 세분화 기반의 조기 판단 학습 가속 장치는 현재의 제1 및 제2 막전위의 분포 데이터에 기초하여, 다음 학습 과정에서 사용될 임계값을 계산할 수 있다. 이때, 임계값은 수학식 1에 기초하여 계산될 수 있으며, 이와 관련하여 구체적인 내용은 도 4에서 후술될 것이다.

[0031] S40 단계에서, 타임스텝 세분화 기반의 조기 판단 학습 가속 장치는 세분화된 타임스텝에서 제1 및 제2 막전위의 차이가 상기 임계값보다 큰 경우, 이미지에 학습 기여도가 없다고 판단하여 세분화된 타임스텝에서의 학습을 종료할 수 있다. 이에 따라, 타임스텝 세분화 기반의 조기 판단 학습 가속 장치는 학습 기여도가 없는 이미지에 대한 학습을 종료시킴으로써, 정확도 감소없이 연산량과 연산시간을 감소시킬 수 있다.

[0032] 좀 더 자세하게, 타임스텝 세분화 기반의 조기 판단 학습 가속 장치는 하나의 타임스텝을 세분화하여 학습 기여도가 없는 이미지에 대한 학습을 종료함으로써, 하나의 타임스텝에서 학습 기여도가 없는 이미지에 대한 학습을 종료시키는 것 보다 연산량과 연산시간을 감소시킬 수 있다.

[0033] 도 3은 본 발명의 일 실시예에 따른 타임스텝 세분화 단계의 예를 보여주는 도이다.

[0034] 도 3을 참조하면, 하나의 타임스텝은 16칸으로 구성되고, 16칸은 중 적어도 하나는 입력 스파이크를 가지고 있을 수 있다. 예를 들어, 도 3과 같이 하나의 타임스텝은 제1 내지 제16번째 칸을 포함하고, 제1, 제4, 제6, 제11, 제13번째 칸은 입력 스파이크를 가지고 있을 수 있다.

[0035] 타임스텝 세분화 기반의 조기 판단 학습 가속 장치는 하나의 타임스텝을 4개로 세분화할 수 있다.

[0036] 이때, 제1 세분화 구역은 제1, 제3, 제9, 제11번째 칸을 포함하고, 제2 세분화 구역은 제2, 제4, 제10, 제12번째 칸을 포함하고, 제3 세분화 구역은 제5, 제7, 제13, 제15번째 칸을 포함하며, 제4 세분화 구역은 제6, 제8, 제14, 제16번째 칸을 포함할 수 있다.

[0037] 다만, 상술한 내용은 하나의 예시일 뿐이며, 타임스텝은 2개 내지 16개로 세분화될 수 있다.

[0038] 예를 들어, 타임스텝을 2개의 타임스텝으로 세분화할 경우에 제1 세분화 구역은 제1 내지 제8번째 칸을 포함하고, 제2 세분화 구역은 제9 내지 제16번째 칸을 포함할 수 있다.

[0039] 다른 예로, 타임스텝을 2개의 타임스텝으로 세분화할 경우에 제1 세분화 구역은 제1, 제3, 제5, 제7, 제9, 제11, 제13, 제15번째 칸을 포함하고, 제2 세분화 구역은 제2, 제4, 제6, 제8, 제10, 제12, 제14, 제16번째 칸을 포함할 수 있다.

[0040] 본 출원의 실시 예에 있어서, 설계자는 하드웨어의 구성에 따라 타임스텝의 세분화 개수를 적절히 설계할 수 있다. 예를 들어, 타임스텝의 세분화 개수를 늘리게 되면, 조기 판단을 더욱 수행할 수 있어 학습 가속 효과가 증가될 수 있다. 그러나, 타입스텝의 세분화 개수가 늘어날수록 하드웨어 측면에서는 조기 판단 모듈이 자주 동작하여야 하기 때문에, 추가적인 에너지 소모가 발생할 수 있다. 따라서, 설계자는 하드웨어 구성 및 성능을 고려하여 타임스텝의 개수를 적절히 설계할 수 있다.

[0041] 도 4는 본 발명의 일 실시예에 따른 임계값을 계산하는 방법을 보여주는 도이다.

[0042] 도 4를 참조하면, 타임스텝 세분화 기반의 조기 판단 학습 가속 장치는 학습 과정 동안 학습 기여도가 있는 이미지(NZL images)와 학습 기여도가 없는 이미지(ZL images)를 가우시안의 형태로 나타낼 수 있다.

[0043] 예를 들어, 타임스텝 세분화 기반의 조기 판단 학습 가속 장치는 학습 기여도가 있는 이미지(NZL images)와 학습 기여도가 없는 이미지(ZL images)를 x축이 막전위의 갭이고, y축이 해당 막전위 갭을 갖는 이미지의 개수인 가우시안의 형태로 나타낼 수 있다.

[0044] 타임스텝 세분화 기반의 조기 판단 학습 가속 장치는 가우시안 형태로 나타낸 데이터에 기초하여, 타임스텝에 따라 막전위의 갭과 해당 막전위 갭을 갖는 이미지의 개수에 기초하여 임계값(threshold)을 계산할 수 있다. 이때, 막전위의 갭은 정답 뉴런 막전위와 나머지 뉴런 막전위 중 최대값의 차일 수 있다.

[0045] 임계값은 수학식 1에 기초하여 계산될 수 있다. 수학식 1은 아래와 같다.

[0046] $$a=m+2\sigma$$ (수학식 1)

[0047] 이때, a는 임계값이고, m은 학습 기여도가 있는 이미지의 막전위 차 분포의 평균이며, σ는 학습 기여도가 있는

이미지의 막전위 차 분포의 편차일 수 있다.

[0048] 임계값은 경험적으로 찾지 않고 타임스텝과 학습과정 별로 계산됨에 따라, 데이터넷과 네트워크에 무관하게 자동적으로 계산할 수 있다.

[0049] 도 5는 본 발명의 일 실시예에 따른 학습 기여도가 없는 이미지를 판단하여 학습을 조기 종료하는 방법을 보여주는 도이다.

[0050] 도 5를 참조하면, 타임스텝 세분화 기반의 조기 판단 학습 가속 장치는 학습 기여도가 없는 이미지 및 학습 기여도가 있는 이미지의 분포를 x축이 정답 뉴런의 막전위(Vans)이고, y축이 정답 뉴런을 제외한 뉴런들의 막전위 중 최대값의 막전위(Vres,max)인 산점도의 형태로 나타낼 수 있다.

[0051] 타임스텝 세분화 기반의 조기 판단 학습 가속 장치는 산점도 형태로 나타낸 데이터에 기초하여, 타임스텝에 따라 학습 과정의 한 학습의 횟수동안 학습 이미지들에서 제1 및 제2 막전위의 관계를 판단할 수 있다.

[0052] 타임스텝 세분화 기반의 조기 판단 학습 가속 장치는 수학식 2에 기초하여, 학습 기여도가 없는 이미지와 학습 기여도가 있는 이미지의 경계를 설정할 수 있다.

[0053] 수학식 2는 y=x-a일 수 있다.

[0054] 이때, a는 이전 학습 과정에서 계산된 임계값이고, x는 정답 뉴런의 막전위인 제1 막전위이고, y는 복수 개의 막전위 중 정답 뉴런의 막전위 값을 제외한 막전위 중 가장 큰 막전위인 제2 막전위일 수 있다.

[0055] 타임스텝 세분화 기반의 조기 판단 학습 가속 장치는 수학식 2를 변환하여 생성된 수학식 3에 기초하여, 학습 기여도가 없는 이미지에 대해 학습을 종료할 수 있다.

[0056] 수학식 3은 $x-y \geq a$일 수 있다.

[0057] 이때, a는 이전 학습 과정에서 계산된 임계값이고, x는 정답 뉴런의 막전위인 제1 막전위이고, y는 복수 개의 막전위 중 정답 뉴런의 막전위 값을 제외한 막전위 중 가장 큰 막전위인 제2 막전위일 수 있다.

[0058] 타임스텝 세분화 기반의 조기 판단 학습 가속 장치는 수학식 3이 성립될 경우에 학습 기여도가 없는 이미지로 판단하여 진행 중인 학습을 종료할 수 있다. 반면, 타임스텝 세분화 기반의 조기 판단 학습 가속 장치는 수학식 3이 성립되지 않을 경우에 학습 기여도가 있는 이미지로 판단하여 학습을 진행할 수 있다.

[0059] 즉, 제1 막전위와 제2 막전위의 차이가 임계값보다 큰 경우(선의 우측 부분)에 타임스텝 세분화 기반의 조기 판단 학습 가속 장치는 이미지의 학습 기여도가 없음을 판단하고 학습을 종료할 수 있다.

[0060] 도 6은 본 발명의 일 실시예에 따른 세분화된 타임스텝에 대한 학습 조기 종료의 예를 보여주는 도이다.

[0061] 도 6을 참조하면, 타임스텝 세분화 기반의 조기 판단 학습 가속 장치는 세분화된 타임스텝에서, 제1 및 제2 막전위의 차(Vgap)가 임계값 이상일 경우에 해당 세분화된 타임스텝에서 학습을 종료시킬 수 있다.

[0062] 예를 들어, 하나의 타임스텝을 제1 내지 제4 타임스텝(T1, T2, T3, T4)으로 세분화하고, 제1 및 제2 타임스텝(T1, T2) 사이에서 제1 및 제2 막전위의 차(Vgap)가 임계값 이상일 경우, 타임스텝 세분화 기반의 조기 판단 학습 가속 장치는 제2 타임스텝(T2)에서 학습을 종료할 수 있다.

[0063] 상술한 바와 같이, 타임스텝 세분화 기반의 조기 판단 학습 가속 장치는 하나의 타임스텝의 전체에 대해 학습이 종료되는 것이 아닌 제1 및 제2 막전위 차(Vgap)와 임계값(a)에 기초하여 하나의 타임스텝의 중간에 학습이 종료될 수 있다, 이에 따라, 학습의 연산량 및 연산 시간을 감소시킬 수 있다.

[0064] `도 7은 본 발명의 일 실시예에 따른 타임스텝 세분화 기반의 조기 판단 학습 가속 장치를 보여주는 도이다.

[0065] 도 7을 참조하면, 타임스텝 세분화 기반의 조기 판단 학습 가속 장치(10)는 입력 레이어 모듈(200), 히든 레이어 모듈(300), 출력 레이어 모듈(400), 및 글로벌 컨트롤러(500)를 포함할 수 있다.

[0066] 입력 레이어 모듈(200)은 스파이킹 뉴럴 네트워크의 입력 스파이크 신호가 입력될 수 있다.

[0067] 히든 레이어 모듈(300)은 입력 레이어 모듈(200)로부터 입력 스파이크 신호를 입력받을 수 있다. 이때, 히든 레이어 모듈(300)은 막전위 업데이트 모듈(310), 가중치 업데이트 모듈(320), 막전위 버퍼(330), 및 스파이크 시간 버퍼(340)를 포함할 수 있다.

[0068] 막전위 업데이트 모듈(310)은 입력 레이어 모듈(200)로부터 입력 스파이크 신호를 전달받고, 입력 스파이크에

기초하여 복수 개의 뉴런들의 막전위를 계산할 수 있다.

[0069] 막전위 업데이트 모듈(310)은 학습 가중치 저장부(311) 및 추론 가중치 저장부(312)를 포함할 수 있다.

[0070] 학습 가중치 저장부(311)는 학습 과정에서 가중치 업데이트 모듈(320)로부터 가중치를 입력받아, 막전위를 업데이트 할 수 있다.

[0071] 추론 가중치 저장부(312)는 추론 과정에서 가중치 업데이트 모듈(320)로부터 가중치를 입력받아, 막전위를 업데이트 할 수 있다.

[0072] 가중치 업데이트 모듈(320)은 입력 스파이크 신호의 입력에 기초하여 입력 스파이크들의 가중치들을 막전위에 더할 수 있다. 가중치 업데이트 모듈(320)은 가중치 정보에 대해 학습 가중치 저장부(311) 및 추론 가중치 저장부(312)로 전송할 수 있다. 또한, 가중치 업데이트 모듈(320)은 가중치 정보를 출력 레이어 모듈(400)의 가중치 업데이트 모듈(430)로 전송할 수 있다.

[0073] 막전위 버퍼(330)는 복수 개의 뉴런들의 막전위를 저장할 수 있다. 막전위 버퍼는 막전위 업데이트 모듈(310)로 막전위 값을 전송하거나 전송받을 수 있다. 또한, 막전위 버퍼(330)는 스파이크 시간 버퍼(340)로 막전위에 대한 정보를 전송할 수 있다.

[0074] 스파이크 시간 버퍼(340)는 복수 개의 뉴런들의 스파이크 발생 시간을 저장할 수 있다.

[0075] 출력 레이어 모듈(400)은 히든 레이어 모듈(300)로부터 입력 스파이크 신호를 전달받고, 학습 기여도가 없는 이미지를 판단하며, 학습 기여도의 유무를 판단하기 위한 임계값을 계산할 수 있다. 이때, 출력 레이어 모듈(400)은 스파이크 버퍼(410), 막전위 업데이트 모듈(420), 가중치 업데이트 모듈(430), 막전위 버퍼(440), 조기 판단부(450), 임계값 계산부(460), 스파이크 시간 버퍼(470), 및 에러 계산부(480)를 포함할 수 있다.

[0076] 스파이크 버퍼(410)는 히든 레이어 모듈(300)의 막전위 업데이트 모듈로부터 입력 스파이크 신호를 전달받을 수 있다. 스파이크 버퍼(410)는 입력 스파이크 신호를 막전위 업데이트 모듈(420)로 전송할 수 있다.

[0077] 막전위 업데이트 모듈(420)은 히든 레이어 모듈(300)로부터 입력 스파이크 신호를 전달받고, 입력 스파이크에 기초하여 복수 개의 뉴런들의 막전위를 계산할 수 있다.

[0078] 막전위 업데이트 모듈(420)은 학습 가중치 저장부(421) 및 추론 가중치 저장부(422)를 포함할 수 있다.

[0079] 학습 가중치 저장부(421)는 학습 과정에서 가중치 업데이트 모듈(430)로부터 가중치를 입력받아, 막전위를 업데이트 할 수 있다.

[0080] 추론 가중치 저장부(422)는 추론 과정에서 가중치 업데이트 모듈(430)로부터 가중치를 입력받아, 막전위를 업데이트 할 수 있다.

[0081] 가중치 업데이트 모듈(430)은 입력 스파이크 신호의 입력에 기초하여 입력 스파이크들의 가중치들을 막전위에 더할 수 있다. 가중치 업데이트 모듈(430)은 가중치 정보에 대해 학습 가중치 저장부(421) 및 추론 가중치 저장부(422)로 전송할 수 있다. 또한, 가중치 업데이트 모듈(430)은 가중치 정보를 히든 레이어 모듈(300)의 가중치 업데이트 모듈(320)로 전송할 수 있다.

[0082] 막전위 버퍼(440)는 복수 개의 뉴런들의 막전위를 저장할 수 있다. 막전위 버퍼(440)는 막전위 업데이트 모듈(420)로 막전위 값을 전송하거나 전송받을 수 있다. 또한, 막전위 버퍼(440)는 스파이크 시간 버퍼(470) 및 조기 판단부(450)로 막전위에 대한 정보를 전송할 수 있다.

[0083] 조기 판단부(450)는 순전파 과정에서 각 타임스텝의 연산이 끝났을 때, 학습 기여도가 없는 이미지를 판단할 수 있다. 조기 판단부(450)는 판단 결과를 글로벌 컨트롤러(500)로 전송할 수 있다.

[0084] 예를 들어, 조기 판단부(450)는 타임스텝에서 제1 및 제2 막전위의 차이가 임계값보다 큰 경우, 이미지에 학습 기여도가 없다고 판단하여 타임스텝에서의 학습을 종료할 수 있다. 이때, 제1 막전위는 정답 뉴런의 막전위이고, 제2 막전위는 복수 개의 막전위 중 정답 뉴런의 막전위를 제외한 막전위 중 가장 큰 막전위일 수 있다.

[0085] 임계값 계산부(460)는 순전파 과정이 모두 끝난 후에 다음 학습에서 사용될 임계값을 계산할 수 있다. 또한, 임계값 계산부(460)는 이전 학습에서 계산된 임계값을 조기 판단부(450)로 전송할 수 있다.

[0086] 스파이크 시간 버퍼(470)는 복수 개의 뉴런들의 스파이크 발생 시간을 저장할 수 있다. 스파이크 시간 버퍼

(470)는 스파이크 발생 시간을 에러 계산부(480)로 전송할 수 있다.

[0087] 에러 계산부(480)는 순전파 과정이 모두 끝난 후의 복수 개의 뉴런들의 스파이크 발생 시간과 목표하는 정답 신호와의 차이를 계산하여 에러 여부를 계산할 수 있다. 에러 계산부는 계산 결과를 가중치 업데이트 모듈(430)로 전송할 수 있다.

[0088] 글로벌 컨트롤러(500)는 출력 레이어 모듈(400)의 조기 판단부(450)가 판단한 학습 기여도의 유무에 기초하여 학습 과정을 종료할 수 있다. 예를 들어, 조기 판단부(450)에서 학습 기여도가 없는 이미지라고 판단했다면, 글로벌 컨트롤러(500)는 학습 과정을 종료할 수 있다.

[0089] 본 명세서와 도면에 개시된 실시 예들은 본 개시의 내용을 쉽게 설명하고, 이해를 돕기 위해 특정 예를 제시한 것일 뿐이며, 본 개시의 범위를 한정하고자 하는 것은 아니다. 따라서 본 개시의 범위는 여기에 개시된 실시 예들 이외에도 본 개시의 기술적 사상을 바탕으로 도출되는 모든 변경 또는 변형된 형태가 본 개시의 범위에 포함되는 것으로 해석되어야 한다.

부호의 설명

[0090] 10 : 타임스텝 세분화 기반의 조기 판단 학습 가속 장치

200 : 입력 레이어 모듈

300 : 히든 레이어 모듈

310, 420 : 막전위 업데이트 모듈

311, 421 : 학습 가중치 저장부

312, 422 : 추론 가중치 저장부

320, 430 : 가중치 업데이트 모듈

330, 440 : 막전위 버퍼

340, 470 : 스파이크 시간 버퍼

400 : 출력 레이어 모듈

410 : 스파이크 버퍼

450 : 조기 판단부

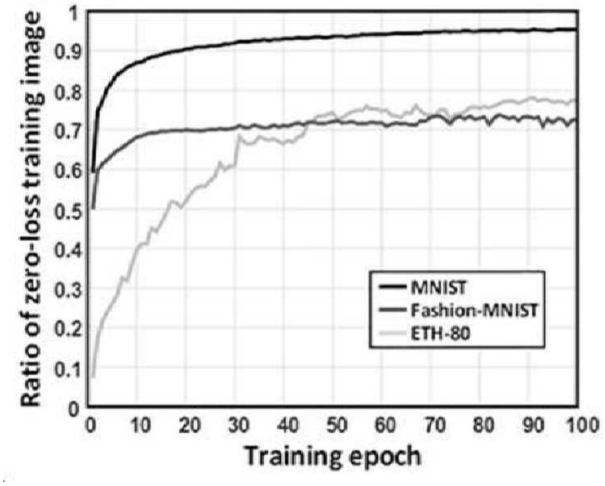
460 : 임계값 계산부

480 : 에러 계산부

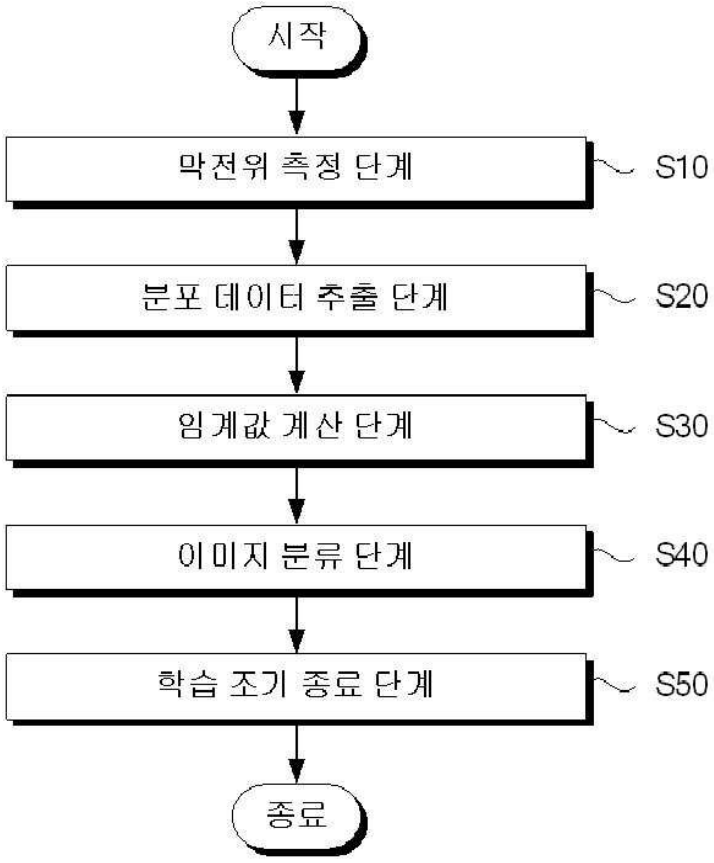
500 : 글로벌 컨트롤러

도면

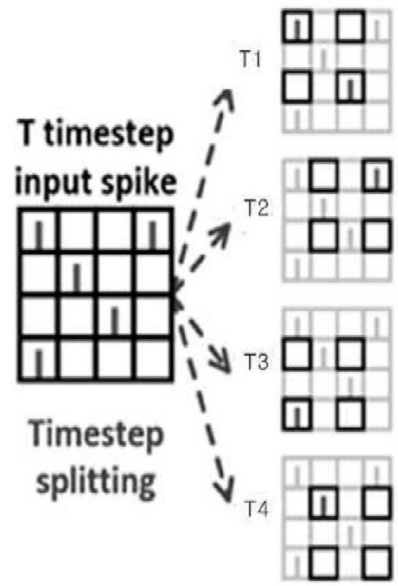
도면1


Ratio of zero-loss training image
Training epoch
MNIST
Fashion-MNIST
ETH-80
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
0
10
20
30
40
50
60
70
80
90
100

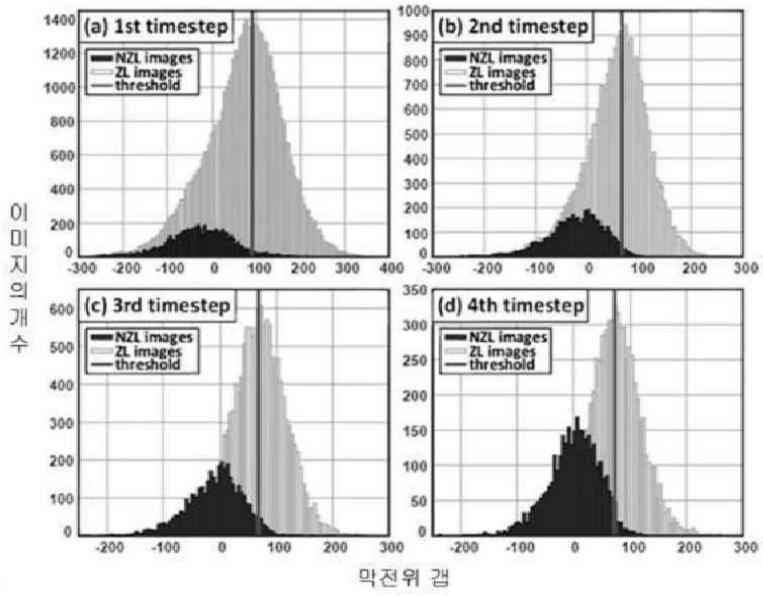
도면2


시작
막전위 측정 단계
S10
분포 데이터 추출 단계
S20
임계값 계산 단계
S30
이미지 분류 단계
S40
학습 조기 종료 단계
S50
종료

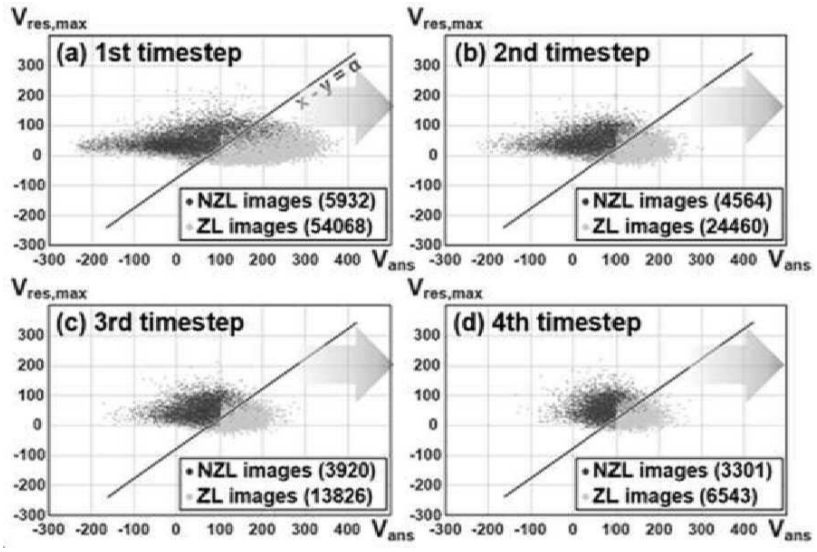
도면3


T timestep
input spike
T1
T2
T3
T4
Timestep
splitting

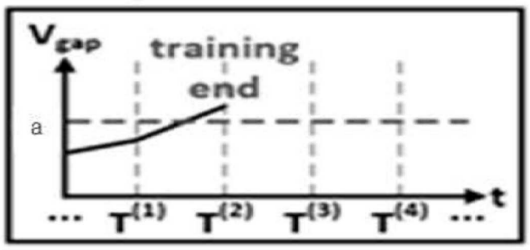
도면4


(a) 1st timestep
(b) 2nd timestep
(c) 3rd timestep
(d) 4th timestep
NZL images
ZL images
threshold
이미지의 개수
막전위 갭

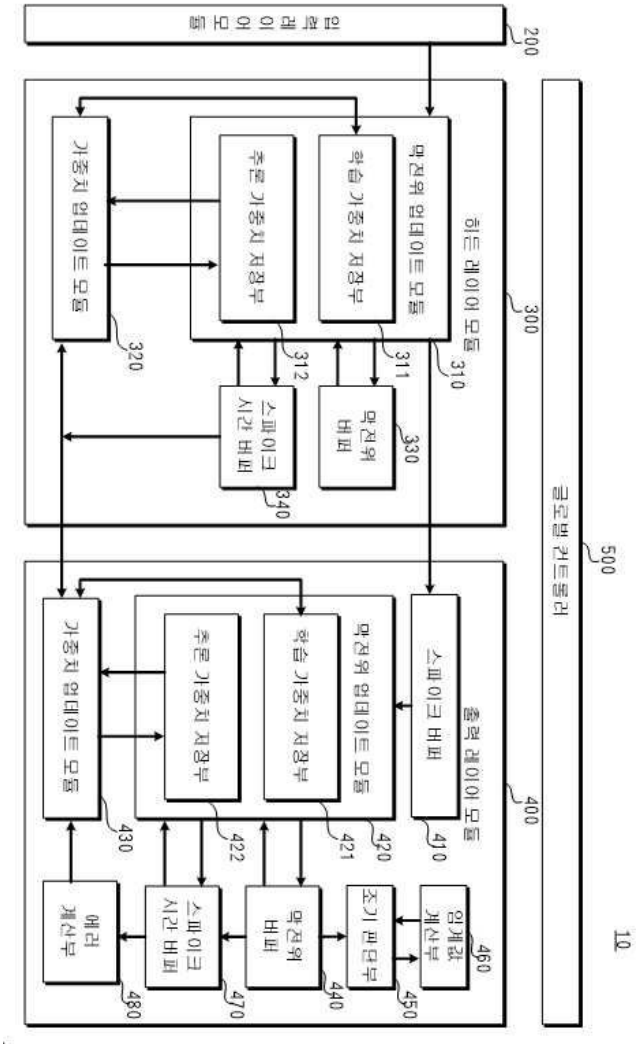
도면5


(a) 1st timestep
x - y = a
NZL images (5932)
ZL images (54068)
(b) 2nd timestep
NZL images (4564)
ZL images (24460)
(c) 3rd timestep
NZL images (3920)
ZL images (13826)
(d) 4th timestep
NZL images (3301)
ZL images (6543)
V res,max
V ans

도면6


training
end
a
t

도면7


입력 레이어 모듈
200
히든 레이어 모듈
300
가중치 업데이트 모듈
320
추론 가중치 저장부
312
학습 가중치 저장부
311
막전위 업데이트 모듈
310
스파이크 시간 버퍼
340
막전위 버퍼
330
글로벌 컨트롤러
500
출력 레이어 모듈
400
가중치 업데이트 모듈
430
추론 가중치 저장부
422
학습 가중치 저장부
421
막전위 업데이트 모듈
420
스파이크 버퍼
410
에러 계산부
480
스파이크 시간 버퍼
470
막전위 버퍼
440
조기 판단부
450
임계값 계산부
460
10