

신경회로망칩(ERNIE^{***})을 위한 학습모듈 설계

Learning Module Design for Neural Network Processor(ERNIE)

정 제 교*, 김 영 주*, 동 성 수**, 이 종 호*

(Je Kyo Jung, Yung Joo Kim, Sung Soo Dong, Chong Ho Lee)

* 인하대학교 정보통신대학원(전화:(032)860-7693, E-mail : L_nux@hotmail.com)

** 용인송담대학 디지털전자정보과(전화:(031)330-9292, E-mail : ssdong@ysc.ac.kr)

Abstract : In this paper, a Learning module for a reconfigurable neural network processor(ERNIE^{***}) was proposed for an On-chip learning. The existing reconfigurable neural network processor(ERNIE) has a much better performance than the software program but it doesn't support On-chip learning function. A learning module which is based on Back Propagation algorithm was designed for a help of this weak point. A pipeline structure let the learning module be able to update the weights rapidly and continuously. It was tested with five types of alphabet font to evaluate learning module. It compared with C programed neural network model on PC in calculation speed and correctness of recognition. As a result of this experiment, it can be found that the neural network processor(ERNIE) with learning module decrease the neural network training time efficiently at the same recognition rate compared with software computing based neural network model. This On-chip learning module showed that the reconfigurable neural network processor(ERNIE) could be a evolvable neural network processor which can fine the optimal configuration of network by itself

Keywords : Neural network, training, learning, synapsis, recognitio

*****ERNIE** : Expansible & Reconfigurable Neuro Informatics Engine

I. 서 론

신경망은 뇌의 정보처리 방식에서 얻은 아이디어를 모델화한 것으로 기존의 컴퓨터로 해결하기 어려웠던 추론, 인식, 연상 또는 복잡한 비선형 제어 등의 응용 분야에서 많이 사용되는 알고리즘이다. 신경망 구현에는 크게 범용 CPU를 사용한 프로그램 방식과 하드웨어 회로로서 구현한 전용 칩을 설계 방법이 있다. 프로그램 방식은 신경망 구조를 자유롭게 설계할 수 있으며 구현 방법이 비교적 간단하지만 앞서 설명한 대로 신경망이 어느 정도 복잡해지고 규모가 커지게 되면 병렬 및 분산 처리로 인해 기존의 소프트웨어 프로그램 방식으로는 연산 결과를 얻거나 시뮬레이션하기가 불가능해 질 정도로 계산량이 커지게 된다. 이를 해결하기 위해 다양한 종류의 하드웨어 신경망이 제안되어왔다.

1958년 F. Rosenblatt의 MARK I Perceptron[1] 이후로 하드웨어 신경망 구현이 많이 연구되었는데, 디지털 회로설계 기술의 발달로 최근에는 주로 디지털 신경회로망을 연구대상으로 한다. 이러한 노력의 결과 중 하나로 재구성 가능한 신경회로망(ERNIE)[2]이 개발되었다. 이 재구성 가능한 신경망(ERNIE)은 대규모

확장이 가능하고 병렬 연산 및 구조 변경이 유연하기 때문에 각종 신경망의 실험에 적합하지만 스스로 연결 강도를 변경하거나 구조를 바꿀 수 있는 기능은 아직 구현되지 않았다. 따라서 학습을 위해서는 외부에 따라 부가 장비가 필요하게 된다. 실험실 내에서는 크게 문제가 되지는 않았으나 앞으로 유비쿼터스 등 모바일 환경에서 신경망을 응용하기 위해서는 커다란 외부 시스템의 개입 없이도 칩 자체로 신경망의 주요 동작을 할 수 있어야 한다. 이를 위해 본 논문에서는 다양한 방면에 활용되는 Back Propagation 알고리즘을 기반으로 하는 학습 모듈을 구현하여 재구성 가능한 신경망(ERNIE)와 함께 실험을 하였다. ERNIE가 Verilog HDL로 코딩되었기 때문에 학습 모듈 역시 같은 방법으로 구현하였다.

II. 학습모듈

Back Propagation(BP) 알고리즘은 순방향 다층 신경망의 학습에 효과적으로 적용할 수 있어서 다양한 분야에서 보편적으로 사용되는 중요한 알고리즘이다. BP 알고리즘의 학습은 다음 3단계를 따라 이루어진다. 1

단계로 입력 패턴을 신경망에 입력해서 결과를 얻는다. 2단계에서 이 결과를 목표 값과 비교해서 오차를 얻는다. 3단계로 오차값에 따라 출력 층의 연결강도 및 은닉 층의 연결강도를 변경해준다.[3]

기존에는 학습 단계 자체를 프로그램으로 수행해서 얻은 연결강도를 신경망 칩에 전달해서 실험을 하거나 1~2단계를 신경망 칩이 수행하고 3단계를 프로그램으로 작성해서 학습을 진행 하는 방법을 썼다. 그러나 첫 번째 방법은 신경망을 하드웨어로 구현한 의미가 사라지게 되며 연산 시간이 많이 필요하게 된다. 두 번째 방법은 비록 앞의 방법에 비해서는 훨씬 빨리 결과를 얻을 수 있으나 학습을 위해서는 매번 연결강도를 변경할 때마다 호스트 프로그램과 데이터를 교환해야 하므로 잦은 데이터 입출력이 필요하게 되고, 이때 불필요한 오버헤드가 늘어나게 된다.

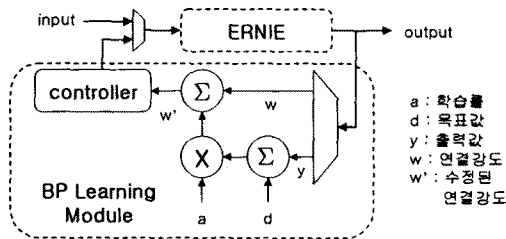


그림 1. 학습모듈의 구조
Fig. 1. Learning module structure

구현된 학습모듈은

그림 1과 같이 신경회로망에 연결된다. 신경회로망(ERNIE)의 출력이 학습 모듈의 입력으로 들어가서 목표값과의 차이를 구하고 그 차이에 따라 연속해서 들어오는 연결강도를 수정해 준다.

III. 설계 구조

신경회로망(ERNIE)에서 차례로 연결강도를 출력해 주면 학습 모듈에서는 수정된 연결강도를 다시 신경회로망(ERNIE)으로 입력해 준다. 학습 알고리즘은 일반적인 BP 알고리즘을 그대로 따른다. 개선된 연결강도를 신경회로망에 전달해주는 컨트롤러 부분이 연산 효율을 높여주게 된다. 이전 학습에서는 모든 연결강도를 구하고 나서야 비로소 신경망 칩에 연결강도를 전송해 줄 수 있었으나 칩 내부에 구현된 학습 모듈은 변경된 연결강도가 바로 MPU로 입력되는 파이프라인 형태를 갖추고 있다. 이로 인하여 따로 연결강도를 전송해 줄 필요가 없이 학습 모듈에서 연산이 끝나면 바로 다음번 학습 패턴을 신경망 칩으로 전달해 주면서 학습이 이루어지게 되므로 시스템 전체의 연산속도를 높일 수 있게 된다. 각 속도의 비교는 도표 1.에서 보는 바와 같다.

table 1. Speed comparison

도표 1. 속도 비교

	PC	ERNIE	ERNIE+Learning module
Learn	5	1	0.5
Run	4	1	1
total	9	2	1.5

(ERNIE의 연산 속도를 기준(1)으로 각 실험 대상의 상대속도를 비교. 실험 방법은 그림 5.의 인식 실험을 기준으로 했을 때의 연산 시간을 비교)

이러한 구성이 가능한 것은 신경회로망(ERNIE)의 구조적인 특징 때문이다. 그림 2에서와 같이 신경회로망(ERNIE)의 기본 단위는 MPU이다. 하나의 MPU 출력이 다른 MPU의 입력으로 연결됨으로 신경망의 확장이 가능하게 되고 자연스럽게 파이프라인이 구성된다. 신경회로망의 입력단과 출력단을 학습모듈에 연결해 주면 그림 4와 같은 순환 회로를 얻을 수 있다. 신경회로망(ERNIE)에서 나온 결과 값으로 연결강도 수정에 필요한 연산을 마치면 뉴런과 뉴런을 연결하는 연결강도 값이 순차적으로 신경회로망(ERNIE)에서 나오게 된다. 학습 모듈에서는 이 값을 받아서 연결강도를 수정해 주고 다시 신경회로망에 되돌려주게 된다. 이 구조의 장점으로는 연결강도를 수정하기 위해 따로 버퍼가 필요 없다는 점이다.

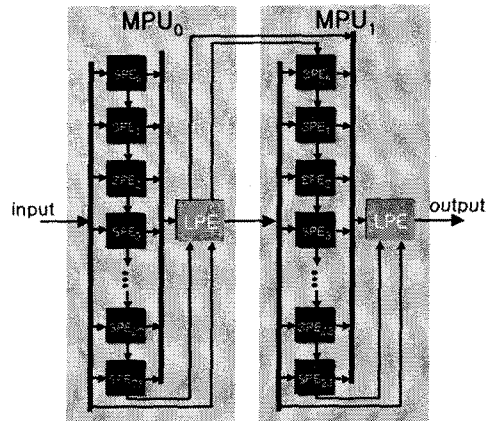


그림 2. MPU 연결 다이어그램
Fig. 2. MPU connectin diagram

또한, 앞에 있는 MPU는 다음 MPU로 데이터 및 제어 신호를 그대로 전달해 주므로 모든 MPU의 제어가 가능하게 된다. 따라서 MPU를 증가시켜 신경망이 확장 가능하게 되더라도 기존의 학습모듈은 변경할 필요가 없다.

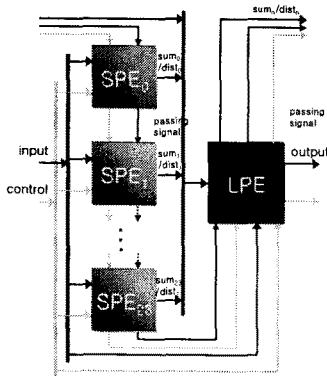


그림 3. MPU의 내부 구성
Fig. 3. Inside of MPU

따라서 그림 4에서 볼 수 있듯이 한 번 초기 연결강도를 입력 받으면 입력패턴이 들어올 때마다 연결강도가 MPU와 Controller를 계속 순환하면서 수정되면서 학습이 이루어지게 된다.

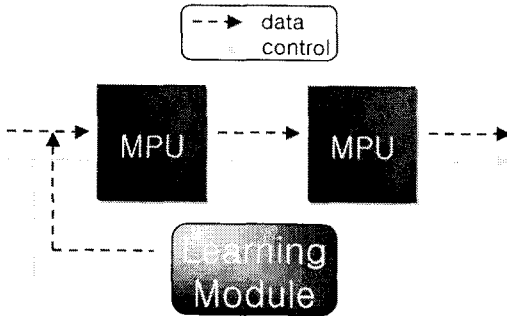


그림 4. 신경회로망과 학습모듈의 연결 구조
Fig. 4. The connection of ERNIE and Learning module

IV. 실험내용

학습 모듈의 동작 및 성능을 검증하기 위해 기존의 신경회로망(ERNIE)과 학습 모듈이 추가된 신경망(ERNIE), 그리고 C로 구현된 신경망 모델을 동일 조건으로 실행시켜서 그 결과를 비교 분석 하였다. 실험 방법은 참고 논문[1]과 동일하게 실행하였다.

그림 5와 같이 은닉층에 32개의 뉴런, 출력층에 5개의 뉴런으로 MLP를 구성하고 영문 26글자 다섯 가지 폰트의 정상 패턴으로 신경망을 학습 시킨 후 손상된 폰트의 인식능력을 검증하였다.

학습율은 0.05로 주었으며 학습에 걸린 반복 횟수는 약 2000회였다. 에러 임계값을 0.001로 주고 HDL 시뮬레이션과 C 모델의 결과를 봤을 때, C 모델과 연결강도를 PC에서 학습한 결과를 사용한 신경회로망

(ERNIE)에서는 동일하게 단층 퍼셉트론에서 0.075%, 다층 퍼셉트론에서 0.5%의 오차율이 확인됐다. 학습 모듈이 추가된 신경회로망(ERNIE)에서는 계단 함수를 활성화 함수로 사용했을 때는 비교 대상과 동일한 결과를 얻을 수 있었으나 시그모이드 함수를 적용했을 때는 0.1~0.02%정도 오차가 더 크다는 사실을 확인할 수 있었다. 평균 오차를 구하는 공식은 다음과 같으며 결과는 도표 2에 간단히 정리되어 있다.

$$error = \frac{\sum_{i=1}^T \sum_{j=1}^N |out_{e_{ij}} - out_{c_{ij}}|}{T \times N}$$

T : 테스트 패턴의 개수

N : 출력층 뉴런의 개수

$out_{e_{ij}}$: i 번째 테스트 패턴에 대한 j 번째 출력 뉴런 연산 결과(ERNIE)

$out_{c_{ij}}$: i 번째 테스트 패턴에 대한 j 번째 출력 뉴런 연산 결과(C 모델링)

도표 2. 오차 비교표

Table 2. Error comparison

(error)

	PC	ERNIE	+Learning Module
단층퍼셉트론	0.075%	0.075%	0.92%
다층퍼셉트론	0.5%	0.5%	0.57%
계단함수	0%	0%	0%

도표 2에서 보이는 오차는 곱셈 연산에서 유효범위를 조정함으로 충분히 수정할 수 있을 것으로 예상된다. 다만 이 때에는 메모리의 크기가 FPGA상에서 차지하는 면적을 고려해서 적절한 범위를 찾아야할 것이다. 신경회로망(ERNIE)의 구현은 Vertex2 V2C6000 위에 구현하였고 600만 게이트 중 약 86%의 면적을 사용하였다. 실제로 칩에 구현할 수 있는 최대 뉴런의 수는 120개이다. 이 때의 대략 20MHz의 속도로 동작이 가능하였다.

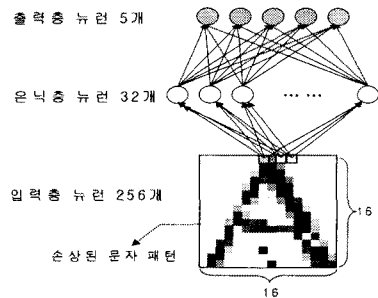


그림 5. 알파벳 인식을 위한 신경 회로망 구조

Fig. 5. Neural network structure for alphabet recognition

V. 결 론

본 논문에서는 기존에 개발된 신경회로망에 학습 기능을 구현함으로써 외부의 보조 없이도 학습이 가능함을 보이고 성능 또한 입증하였다. 학습에 필요한 연산 부분은 칩에 내장했기 때문에 칩 외부와 데이터 전송이 1/2로 줄어들었고 이로 인한 오버헤드의 감소로 전반적인 학습 시간도 줄일 수 있었다. 본 논문에서 BP 알고리즘의 동작이 가능함을 보였지만 추후 SOM, RBF 등의 학습 모듈도 설계가 된다면 신경회로망(ERNIE)의 활용도는 더욱 높아질 것이다. 현재 연결 강도 수정을 위해 부가적으로 설계한 컨트롤 모듈의 기능을 확장하면 신경회로망(ERNIE)의 재구성 기능도 수행할 수 있을 것으로 예상된다.

참 고 문 헌

- [1] 김영주, "대규모 확장이 가능한 범용 신경회로망 : ERNIE", Proceedings of IEEK Summer Conference 2003 제26권 제1호, pp.1263.
- [2] Robert J. Schalkoff, Artificial neural networks, McGraw-Hill, pp.1-2, pp.411, 1997.
- [3] 오창석, *뉴로컴퓨터*, 내하출판사, 성북, 서울, 1996. pp.194, ISBN 89-950557-5-8.