

NETLA를 이용한 이진 신경회로망의 최적 합성방법

성상규*, 김태우**, 박두환*, 조현우**, 하홍곤***, 이준택*
 *동아대학교 전기공학과, **한국전력공사, ***동의대학교 전자공학과

Optimal Synthesis Method for Binary Neural Network using NETLA

Sang-Kyu Sung*, Tae-Woo Kim**, Doo-Hwan Park*, Hyun-Woo Jo**, Hong-Gon Ha***, Joon-Tark Lee*
 *Dong-A Uni, **KEPCO, ***Dong-Eui Uni

Abstract-This paper describes an optimal synthesis method of binary neural network(BNN) for an approximation problem of a circular region using a newly proposed learning algorithm[7]

Our object is to minimize the number of connections and neurons in hidden layer by using a Newly Expanded and Truncated Learning Algorithm(NETLA) for the multilayer BNN. The synthesis method in the NETLA is based on the extension principle of Expanded and Truncated Learning(ETL) and is based on Expanded Sum of Product (ESP) as one of the boolean expression techniques. And it has an ability to optimize the given BNN in the binary space without any iterative training as the conventional Error Back Propagation(EBP) algorithm[6]

If all the true and false patterns are only given, the connection weights and the threshold values can be immediately determined by an optimal synthesis method of the NETLA without any tedious learning.

Futhermore, the number of the required neurons in hidden layer can be reduced and the fast learning of BNN can be realized. The superiority of this NETLA to other algorithms was proved by the approximation problem of one circular region.

1. 서 론

N-tuple 방법에 기반을 둔 Binary Neural Network (BNN)은 지식의 조작 내의 많은 잠재적인 이점을 가지고 있기 때문에 영상처리, 패턴인식 등에 이용되어 왔다. 그러나, 데이터의 거대한 양자화를 위해 BNN내의 실시간 처리 문제는 해결하지 못했다.

최근에, Expanded and Truncated Learning(ETL)라 불리는 기하학적 학습 알고리즘 binary to binary mapping의 일반화를 이용한 3층 BNN을 훈련시키기 위해 제안되었다. ETL은 hard-limiter 활성화함수를 이용하고, 알고리즘 내의 가중치와 임계치는 정수이다. 주어진 훈련 입력의 기하학적인 분석을 통해 요구된 separating hyperplane을 찾아내고, 그들 뉴런의 값을 결정하는 것이다. 이 알고리즘 역시, 모든 패턴의 순차적인 학습을 요구한다[6].

학습 알고리즘인 MSP Term Grouping Algorithm (MTGA)은 디지털 논리 합성방법 중의 하나인 MSP (Minimal Sum of Product)형태와 필요조건인 Unate의 특성을 만족하는 항만을 합성한다. 그러나, MTGA는 필요조건인 Unate의 특성을 만족하지 못하는 패턴은 합성되지 못하고, 하나의 은닉층 내의 뉴런으로 부가된다.

따라서, 본 논문에서 제안한 Newly ETL Algorithm (NETLA)는 unate의 특성을 만족하지 못하는 경우와 다른 형태의 boolean expression를 이용하여 적용한다. NETLA는 ETL 알고리즘을 이용하여 임계치와 가중치를 결정하고, 확장의 원리에 의해서 훈련 입력을 최적

합성한다. 원형 지역과 같은 물체 인식의 실질적인 적용을 통해 은닉층 내의 뉴런의 수를 감소시킴으로서, 학습 속도도 단축시킬 수 있는 우수성이 보이고자 한다.

2. 본 론

2.1 이진 신경회로망의 구조

n-비트의 이진 훈련 입력 패턴들이 (n-1)차원의 hyperplane에 의해 분리될 수 있다고 가정한다.

$$net(x, T) = w_1x_1 + w_2x_2 + \dots + w_nx_n - T = 0 \quad (1)$$

여기서, w_i : i번째 입력과 뉴런 사이의 연결 가중치

x_i : i번째 입력

T : 뉴런의 임계치

x : 입력 벡터 ; $x = [x_1, x_2, \dots, x_n]^T$

이 경우, 훈련입력의 집합은 Linearly Separable(LS) 선형적으로 분리가능하고, (n-1)차원의 hyperplane은 separating hyperplane을 아래의 hard-limiter 활성화함수를 가진 n-입력 뉴런에 의해 성립된다

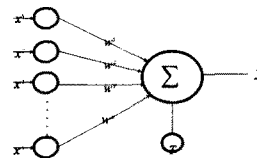


그림 1 이진 신경회로망의 구조
 Fig. 1 Structure of Binary Neural Network

$$y = \begin{cases} 1 : \sum_{i=1}^n w_i x_i - T \geq 0 \\ 0 : \text{otherwise} \end{cases} \quad (2)$$

여기서, y : 뉴런의 출력

2.2 가중치와 임계치의 결정

가중치와 임계치를 결정하기 위해서, 모든 true와 false 패턴들을 직접적으로 결정할 수 있는 ETL 알고리즘내의 RHS(Reference HyperSphere)방법을 이용한다. RHS방법은 reference hypersphere와 separating hyperplane 이용한다. reference hypersphere는 모든 훈련 입력을 포함하고, separating hypersphere는 오직 true 패턴만을 포함하는 것이다. 그리고, hyperplane은 두 hypersphere이 교차될 때 찾을 수 있다. 기하학적인 ETL방법을 이용함으로써, 훈련 입력들이 중심 $(1/2, 1/2, \dots, 1/2)$ 이고, 반경 $\sqrt{n}/2$ 인 n차원의 단위 hypercube를 분리함으로써, 필요한 hyperplane을 찾을

수 있다. RHS의 식은 아래와 같다.

$$(x_1 - \frac{1}{2})^2 + (x_2 - \frac{1}{2})^2 + \dots + (x_n - \frac{1}{2})^2 = \frac{n}{4} \quad (3)$$

그리고, RHS내의 중심 C_i/C_0 이고, 반경 R 인 separating hypersphere를 가정한다.

$$(x_1 - \frac{C_1}{C_0})^2 + (x_2 - \frac{C_2}{C_0})^2 + \dots + (x_n - \frac{C_n}{C_0})^2 = R^2 \quad (4)$$

$$(2c_1 - c_0)x_1 + (2c_2 - c_0)x_2 + \dots + (2c_n - c_0)x_n - T = 0 \quad (5)$$

여기서, C : separating hyperplane의 중심의 좌표의 반경

식(5)에 C_0 를 곱하면,

$$\sum_{i=1}^n w_i x_i - T = 0 \text{ 가 된다.}$$

여기서, $w_i = C_0 - 2C_i$

따라서, 가중치는 $w_i = C_0 - 2C_i$ 이고, 임계치는

$$T = R^2 C_0 - \sum_{i=1}^n C_i^2 / C_0 \text{로 계산된다.}$$

2.3 ETL 알고리즘의 확장의 원리

ETL 알고리즘의 확장의 원리를 예제를 통해 설명한다. 예를 들어, 입력 변수의 함수 $f(x_1, x_2, x_3)$ 인 6-훈련 입력을 고려한다.

만약, 입력이 {000, 010, 011, 111}일 때 '1', 입력들이 {001, 100}일 때 '0', 입력 vertex들이 {101, 110}일 때, $f(x_1, x_2, x_3)$ 가 "don't care"인 경우를 가정한다.

SITV는 {000, 010, 011}에 대해

$t_{\min} = \text{Minimum} [-3x_1 + x_2 - x_3]$ 이고, $f_{\max} = \text{Maximum} [-3x_1 + x_2 - x_3]$ 에서 $t_{\min} = 0$, $f_{\max} = -1$ 가 계산된다.

$t_{\min} > f_{\max}$ 와 $T=0$ 이기 때문에 hyperplane

$-3x_1 + x_2 - x_3 = 0$ 는 나머지 vertex들로부터 SITV {000, 010, 011}내에 vertex가 분리된다.

두 번째 요구된 hyperplane은 식(5)로부터 $C_0 = 5$, $C_1 = 1$, $C_2 = 2$, $C_3 = 2$ 를 구하고, $(2C_n - C_0)$ 이 가중치로 결정된다.

즉, $w_n = (2C_n - C_0)$ 이 되고, $-3x_1 - x_2 - x_3 - T = 0$

이 되고, $t_{\min} = -3$ 과 $f_{\max} = -5$ 와 $T = -4$ 이므로,

$t_{\min} > f_{\max}$ 의 조건을 만족함으로써, 두번째의 hyperplane은

$-3x_1 - x_2 - x_3 + 4 = 0$ 로 결정된다. 더 많은 true

vertex를 포함하기 위해서 다른 true vertex가 처음과 같이 사용, 선택되고, 만약, 새로운 trial vertex가 SITV에 부가할 수 있는지 테스트하고, 이 과정은 더 이상의 true vertex와 SITV내에 부가되지 않을 때까지 계속 반복한다.

따라서, 예제에서는 두 separating hyperplane을 찾을 수 있다. 즉, 은닉층 내의 2개의 뉴런을 구성한다.

2.4 실험결과

ETL, MTGA 그리고, NETLA 원의 근사 문제에 적용함으로써, 각각의 결과를 비교한다. 이 적용을 통해 은닉층의 뉴런의 수와 학습 속도가 단축됨을 보이고자 한다. 원은 정사각형 내에 위치하고 있고, 6bit로 양자화된 2차원 문제에 대하여 2차원의 각 축을 3 bit를 사용하여, 각축을 양자화하면 전체 공간이 64개의 사각형으로 구분되며, 가운데 중앙의 12개에 대해서는 '1', 나머지 52개에 대해서는 '0', 이를 구분하는 이진 신경회로망 합성 문제에 적용하였다.

아래의 함수는 원형 영역을 3bit로 양자화한 결과이다.

$$f = x_1 x_2 x_4 x_5 x_6 + x_1 x_2 x_3 x_4 x_5 + x_1 x_2 x_3 x_4 x_5 + x_1 x_2 x_4 x_5 x_6 + x_1 x_2 x_4 x_5 x_6 + x_1 x_2 x_3 x_4 x_5 + x_1 x_2 x_3 x_4 x_5 + x_1 x_2 x_4 x_5 x_6$$

2.4.1 ETL의 합성결과

기존의 학습 알고리즘인 ETL 학습 알고리즘은 그림 3과 같이 1개의 출력층과 은닉층내에 5개의 뉴런으로 구성된 구성된 회로망이 합성되었다.

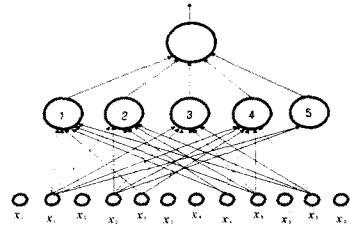


그림 3 ETL의 3층 신경회로망의 구조
Fig 3 Structure of Three Layer Neural Network for ETL

표 2 은닉층내의 가중치와 임계치
Table 2 Weight Value and Threshold Value in Hidden Layer

Neuron	Input						T
	W1	W2	W3	W4	W5	W6	
1	-3	3	1	-3	3	1	7
2	-29	-3	-1	-3	3	1	-5
3	-29	-3	-1	-3	3	1	-26
4	-19	-13	1	-3	3	1	-20
5	-16	-16	0	0	0	0	-24

2.4.2 MTGA의 합성결과

MTGA 학습 알고리즘은 MSP항으로 나타내고, 그 MSP항에 Unate의 특성을 적용하여, 합성한다.

여기서, Unate의 특성을 아래의 예와 같이 negative패턴과 positive패턴 형태가 동시에 존재하는 함수들은 Unate의 특성을 만족하지 못하고, 합성할 수 없다는 원리이다.

예를 들어, 함수 $f(x_1, x_2, x_3)$ 를 가정한다.

$f(x_1, x_2, x_3) = x_1 x_2 + x_2 x_3 + x_1 x_3$: Not-Unate

$f(x_1, x_2, x_3) = x_1 x_2 + x_2 x_3 + x_1 x_3$: Unate

MTGA 학습 알고리즘은 원형 영역을 3bit로 양자화한 결과에서 Unate의 특성을 이용하여 Unate의 특성을 만족하는 항으로 합성된 결과이다.

$$f = (x_1 x_2 x_4 x_5 x_6 + x_1 x_2 x_3 x_4 x_5) + (x_1 x_2 x_3 x_4 x_5 - x_1 x_2 x_4 x_5 x_6) + (x_1 x_2 x_4 x_5 x_6 + x_1 x_2 x_3 x_4 x_5) + (x_1 x_2 x_3 x_4 x_5 + x_1 x_2 x_4 x_5 x_6)$$

그림 4와 같이 1개의 출력층과 4개의 뉴런으로 구성된 은닉층으로 구성된 회로망이 합성되었다.

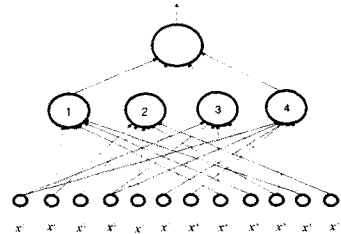


그림 4 MTGA의 3층 신경회로망의 구조
Fig 4 Structure of Three Layer Neural Network for MTGA

표 3 은닉층내의 가중치와 임계치
Table 3 Weight Value and Threshold Value
in Hidden Layer

Neuron	Input						
	W1	W2	W3	W4	W5	W6	T
1	-3	3	1	-3	3	-1	6.5
2	-3	3	1	3	-3	1	5.5
3	3	-3	-1	-3	3	1	5.5
4	3	-3	-1	3	-3	1	4.5

2.4.3 제안하는 NETLA의 최적 합성결과

본 논문에서 제안한 NETLA을 위의 문제에 적용한 결과이다. 이 알고리즘은 ETL의 확장의 원리와 boolean expression 중에 EPS(Expanded Product of Sums)를 이용하여 합성하였다. NETLA는 Unate의 특성을 만족하지 못한 경우와 MSP항으로 표현되지 않더라도, 합성할 수 있는 우수한 특성이 있다.

주어진 함수에서 x_5 과 x_6 으로 각각 분리하고,

$$f = (x_1x_2x_4x_6 + x_1x_2x_3x_4 + x_1x_2x_4x_6 + x_1x_2x_3x_4)x_5 + (x_1x_2x_3x_4 + x_1x_2x_4x_6 + x_1x_2x_3x_4 + x_1x_2x_4x_6)x_6$$

두 패턴에서

$$(x_1x_2x_4x_6 + x_1x_2x_3x_4 + x_1x_2x_4x_6 + x_1x_2x_3x_4) \cdot 1$$

$$(x_1x_2x_3x_4 + x_1x_2x_4x_6 + x_1x_2x_3x_4 + x_1x_2x_4x_6) \cdot 1$$

을 각각 확장하였다.

합성한 결과는 아래와 같다.

$$f = (x_1x_2x_4x_6 + x_1x_2x_3x_4 + x_1x_2x_4x_6 + x_1x_2x_3x_4) \cdot x_5 + (x_1x_2x_4x_6 + x_1x_2x_3x_4 + x_1x_2x_4x_6 + x_1x_2x_3x_4) \cdot 1 + (x_1x_2x_3x_4 + x_1x_2x_4x_6 + x_1x_2x_3x_4 + x_1x_2x_4x_6) \cdot x_6 + (x_1x_2x_4x_6 + x_1x_2x_3x_4 + x_1x_2x_4x_6 + x_1x_2x_3x_4) \cdot 1$$

따라서, MTGA에서 제안하는 MSP의 형태와는 다른 EPS를 적용하였다. 그림 5와 같이 1개의 출력층과 4개의 뉴런으로 구성된 은닉층으로 구성된 회로망이 합성되었다.

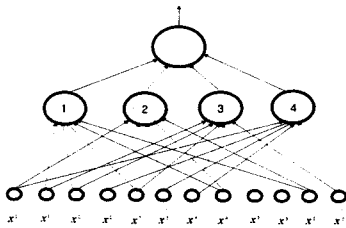


그림 5 NETLA의 3층 신경회로망의 구조
Fig 5 Structure of Three Layer Neural Network for NETLA

표 4 은닉층내의 가중치와 임계치
Table 4 Weight Value and Threshold Value in Hidden Layer

Neuron	Input					
	W1	W2	W3	W4	W5	T
1	-3	3	1	-3	-1	4.5
2	3	-3	-1	-3	1	3.5
3	-3	3	1	3	1	5.5
4	3	-3	-1	3	1	4.5

3. 결 론

본 논문에서 제시한 NETLA 학습 알고리즘은 직접적으로 false, true 패턴들로부터 가중치와 임계치를 계산할 수 있기 때문에 반복적인 학습을 요구하지 않는다.

ETL 학습알고리즘의 확장의 원리와 Boolean Expression을 이용하였고, MSP의 형태와 Unate의 특성을 만족하지 않은 형태

의 함수일지라도, 은닉층의 뉴런을 수를 같게 구성할 수 있었다. 각 은닉층 내의 뉴런의 수가 감소시킴으로써, 학습시간을 단축시킬 수 있었다. 이 NETLA 우수성을 원의 근사 문제에 적용에 의하여 증명하였다. 더 나아가, 이 알고리즘을 이용하여 지문인식, 패턴인식, 영상처리 등에 적용하려고 한다.

(참 고 문 헌)

- [1] P. L. Bartlett and T. Downs, "Using random weight to train multilayer network of hard-limiting units", IEEE Trans. Neural Networks, pp202-210, 1992.
- [2] M. L. Brady, R. Rayhavan, and J. Slawny, "Backpropagation fails to separate where perceptrons succeed", IEEE Trans. Circuits Systems., pp.665-674, May 1989.
- [3] S. Park, J. H. Kim, and H. Chang, "A learning algorithm for discrete multilayer perceptron", in Proc. Int. Syp. Cricuit Systems, Singapore, Jun.1991.
- [4] J. H. Kim, S. K. Park, Han, H. Oh, and M. S. Han, "The geometrical learning of binary neural network", IEEE Trans. Neural Networks, vol. 6.no. 1. pp.237-247, Jan. 1995.
- [5] Mitchell P. Marcus "Switching Circuits for Engineering", Prentice Hall, 1990.
- [6] Sung-Kwon Park "Geometrical learning Algorithm for Multiplayer Neural Networks in a Binary Field", IEEE Trans. on Computers, vol. 42 .no. 8. pp988-992, Aug. 1993.
- [7] John V.Kennedy, Jim Austin, Rick Pack and Bruce n Park "An advanced for the Parallel Processing of Binary Neural Networks", IEEE Trans. on Computers, vol. 52 .no. 4. pp444-451, Aug. 1995..