(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2020년07월21일
(11) 등록번호 10-2135632
(24) 등록일자 2020년07월14일

(51) 국제특허분류(Int. Cl.)
G06N 3/04 (2006.01) *G06N 3/08* (2006.01)
(52) CPC특허분류
G06N 3/04 (2013.01)
G06N 3/08 (2013.01)
(21) 출원번호 10-2018-0116612
(22) 출원일자 2018년09월28일
심사청구일자 2018년09월28일
(65) 공개번호 10-2020-0037028
(43) 공개일자 2020년04월08일
(56) 선행기술조사문헌
JP2006140952 A*
JP2016180839 A*
KR1020170044933 A*
*는 심사관에 의하여 인용된 문헌

(73) 특허권자
포항공과대학교 산학협력단
경상북도 포항시 남구 청암로 77 (지곡동)
(72) 발명자
이영주
경북 포항시 남구 지곡 155 4동
이승구
경북 포항시 남구 지곡로 155 9동
(뒷면에 계속)
(74) 대리인
특허법인 고려

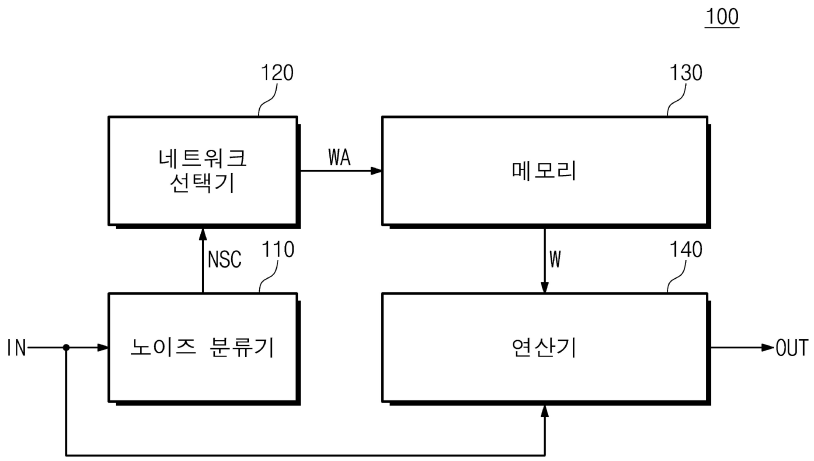
전체 청구항 수 : 총 16 항 심사관 : 이현중

(54) 발명의 명칭 **뉴럴 프로세싱 장치 및 그것의 동작 방법**

(57) 요 약

본 발명의 하나의 실시 예에 따른 뉴럴 프로세싱 장치는 입력 데이터에 대한 전처리를 수행하여 입력 데이터의 노이즈 특성을 판별하도록 구성된 노이즈 분류기, 노이즈 특성에 기초하여 복수의 뉴럴 네트워크들 중 하나를 선택하도록 구성된 네트워크 선택기 및 선택된 뉴럴 네트워크에 대응하는 선택된 가중치들을 기반으로 입력 데이터에 대한 추론을 수행하도록 구성된 연산기를 포함한다.

대 표 도 - 도3



(72) 발명자

하민호

경북 포항시 남구 연일읍 유강길9번길 16-1

변영훈

경북 포항시 남구 청암로 77 포항공과대학교 기숙사 10동

명 세 서

청구범위

청구항 1

입력 데이터에 대한 전처리를 수행하여 상기 입력 데이터의 노이즈 특성을 판별하도록 구성된 노이즈 분류기;

가중치 테이블에 기초하여 복수의 뉴럴 네트워크들 중 상기 판별된 노이즈 특성에 대응하는 뉴럴 네트워크를 선택하고, 상기 선택된 뉴럴 네트워크에 대응하는 가중치 주소를 출력하도록 구성된 네트워크 선택기; 및

상기 출력된 가중치 주소에 대응하는 선택된 가중치들을 기반으로 상기 입력 데이터에 대한 추론을 수행하도록 구성된 연산기를 포함하는 뉴럴 프로세싱 장치.

청구항 2

제 1 항에 있어서,

상기 노이즈 특성은 상기 입력 데이터의 노이즈 타입(type) 및 노이즈 강도(degree)를 포함하는 뉴럴 프로세싱 장치.

청구항 3

제 2 항에 있어서,

상기 노이즈 분류기는 상기 입력 데이터에 대한 푸리에(Fourier) 변환 값들에 기초하여 상기 노이즈 타입 및 상기 노이즈 강도를 판별하는 뉴럴 프로세싱 장치.

청구항 4

제 3 항에 있어서,

상기 노이즈 분류기는 상기 푸리에 변환 값들의 합이 미리 결정된 기준 푸리에 변환 값들의 합보다 크면 상기 노이즈 타입을 가우시안(Gaussian) 노이즈로 판별하고, 상기 푸리에 변환 값들의 합이 상기 미리 결정된 기준 푸리에 변환 값들의 합보다 작으면 상기 노이즈 타입을 블러(blur) 노이즈로 판별하는 뉴럴 프로세싱 장치.

청구항 5

제 3 항에 있어서,

상기 노이즈 분류기는 상기 푸리에 변환 값들의 합의 크기에 기초하여 상기 노이즈 강도를 판별하는 뉴럴 프로세싱 장치.

청구항 6

제 2 항에 있어서,

상기 노이즈 분류기는 뉴럴 네트워크를 기반으로 상기 노이즈 타입 및 상기 노이즈 강도를 판별하는 뉴럴 프로세싱 장치.

청구항 7

제 2 항에 있어서,

상기 복수의 뉴럴 네트워크들 각각은 상기 노이즈 타입 및 상기 노이즈 강도의 조합들 각각에 대응하는 노이즈를 포함하는 입력 데이터를 기반으로 학습된 네트워크인 뉴럴 프로세싱 장치.

청구항 8

삭제

청구항 9

제 1 항에 있어서,

상기 복수의 뉴럴 네트워크들의 가중치들을 저장하고, 상기 선택된 뉴럴 네트워크에 대응하는 상기 가중치 주소에 기초하여 상기 선택된 가중치들을 출력하도록 구성된 메모리를 더 포함하는 뉴럴 프로세싱 장치.

청구항 10

뉴럴 프로세싱 장치의 동작 방법에 있어서,

입력 데이터에 대한 푸리에(Fourier) 변환 값들에 기초하여 상기 입력 데이터의 노이즈 특성을 판별하는 단계;

가중치 테이블에 기초하여 복수의 뉴럴 네트워크들 중 상기 판별된 노이즈 특성에 대응하는 뉴럴 네트워크의 가중치 주소를 출력하는 단계; 및

상기 출력된 가중치 주소에 대응하는 선택된 가중치들을 기반으로 상기 입력 데이터에 대한 추론을 수행하는 단계를 포함하는 동작 방법.

청구항 11

제 10 항에 있어서,

상기 노이즈 특성은 상기 입력 데이터의 노이즈 타입(type) 및 노이즈 강도(degree)를 포함하는 동작 방법.

청구항 12

제 11 항에 있어서,

상기 노이즈 타입은 상기 푸리에 변환 값들의 합이 미리 결정된 기준 푸리에 변환 값들의 합보다 크면 가우시안(Gaussian) 노이즈로 판별되고, 상기 푸리에 변환 값들의 합이 상기 미리 결정된 기준 푸리에 변환 값들의 합보다 작으면 블러(blur) 노이즈로 판별되는 동작 방법.

청구항 13

제 11 항에 있어서,

상기 노이즈 강도는 상기 푸리에 변환 값들의 합의 크기에 기초하여 판별되는 동작 방법.

청구항 14

제 11 항에 있어서,

상기 노이즈 타입 및 상기 노이즈 강도의 조합들 각각에 대응하는 노이즈를 포함하는 입력 데이터를 기반으로 상기 복수의 뉴럴 네트워크들 각각에 대응하는 가중치들을 학습하는 단계를 더 포함하고,

상기 학습된 가중치들을 기반으로 상기 추론이 수행되는 동작 방법.

청구항 15

뉴럴 프로세싱 장치의 동작 방법에 있어서,

특정 뉴럴 네트워크를 기반으로 입력 데이터의 노이즈 특성을 판별하는 단계;

가중치 테이블에 기초하여 복수의 뉴럴 네트워크들 중 상기 판별된 노이즈 특성에 대응하는 뉴럴 네트워크의 가중치 주소를 출력하는 단계; 및

상기 출력된 가중치 주소에 대응하는 선택된 가중치들을 기반으로 상기 입력 데이터에 대한 추론을 수행하는 단계를 포함하는 동작 방법.

청구항 16

제 15 항에 있어서,

상기 노이즈 특성은 상기 입력 데이터의 노이즈 타입(type) 및 노이즈 강도(degree)를 포함하는 동작 방법.

청구항 17

제 15 항에 있어서,

상기 특정 뉴럴 네트워크는 컨볼루셔널 뉴럴 네트워크(CNN; Convolutional Neural Network)인 동작 방법.

발명의 설명

기 술 분 야

[0001] 본 발명은 프로세싱 장치에 관한 것으로써, 좀 더 상세하게는 노이즈가 포함된 입력 데이터에 대하여 뉴럴 네트워크를 기반으로 추론을 수행하는 뉴럴 프로세싱 장치 및 그것의 동작 방법에 관한 것이다.

배 경 기 술

[0002] 뉴럴 네트워크는 사람의 두뇌를 모방하여 데이터를 처리하기 위한 네트워크이다. 두뇌는 뉴런들 사이의 시냅스를 통해 하나의 뉴런으로부터 다른 뉴런으로 신호를 전달할 수 있다. 두뇌는 시냅스의 연결 강도를 조절하여 뉴런으로부터 다른 뉴런으로 전달되는 신호의 세기를 조절할 수 있다. 즉, 시냅스의 연결 강도가 조절됨으로써 정보의 학습 및 추론이 이루어질 수 있다.

[0003] 뉴럴 프로세싱 장치는 인공 신경망 알고리즘을 기반으로 입력 데이터를 처리하는 장치이다. 인공 신경망 알고리즘은 일반적으로 입력 데이터에 대한 학습 과정 및 추론 과정으로 구분될 수 있다. 뉴럴 프로세싱 장치는 학습 과정을 통해 획득된 학습 결과에 기초하여 입력 데이터에 대한 추론을 수행할 수 있다. 뉴럴 프로세싱 장치를 제한된 하드웨어 리소스(resource)로 구현하기 위해 근사 연산을 이용한 추론방식이 활용될 수 있다. 예를 들어, 뉴럴 프로세싱 장치는 단순화된 뉴럴 네트워크를 기반으로 추론을 수행할 수 있다. 이에 따라 연산량이 감소될 수 있다.

[0004] 입력 데이터를 추론함에 있어서, 뉴럴 프로세싱 장치가 근사 연산을 이용하여 추론을 수행하더라도, 추론 결과의 정확도가 높을 수 있다. 그러나, 입력 데이터에 노이즈(noise)가 포함되는 경우, 노이즈의 강도에 따라 추론 결과의 정확도가 급격하게 감소될 수 있다. 또한, 추론 결과의 정확도를 높이기 위해 더 많은 리소스가 할당되는 경우, 연산 시간이 증가될 수 있으며, 연산 수행에 따른 전력 소비가 증가될 수 있다.

발명의 내용

해결하려는 과제

[0005] 본 발명은 상술된 기술적 과제를 해결하기 위한 것으로써, 본 발명의 목적은 노이즈가 포함된 입력 데이터에 대하여 적은 전력을 소비하면서 추론 결과의 정확도를 향상시킬 수 있는 뉴럴 프로세싱 장치 및 그것의 동작 방법을 제공하는데 있다.

과제의 해결 수단

[0006] 본 발명의 하나의 실시 예에 따른 뉴럴 프로세싱 장치는 입력 데이터에 대한 전처리를 수행하여 상기 입력 데이터의 노이즈 특성을 판별하도록 구성된 노이즈 분류기, 상기 노이즈 특성에 기초하여 복수의 뉴럴 네트워크들 중 하나를 선택하도록 구성된 네트워크 선택기 및 상기 선택된 뉴럴 네트워크에 대응하는 선택된 가중치들을 기반으로 상기 입력 데이터에 대한 추론을 수행하도록 구성된 연산기를 포함한다.

[0007] 하나의 실시 예에 있어서, 상기 노이즈 특성은 상기 입력 데이터의 노이즈 타입(type) 및 노이즈 강도(degree)를 포함할 수 있다.

[0008] 하나의 실시 예에 있어서, 상기 노이즈 분류기는 상기 입력 데이터에 대한 푸리에(Fourier) 변환 값들에 기초하여 상기 노이즈 타입 및 상기 노이즈 강도를 판별할 수 있다.

[0009] 하나의 실시 예에 있어서, 상기 노이즈 분류기는 상기 푸리에 변환 값들의 합이 미리 결정된 기준 푸리에 변환 값들의 합보다 크면 상기 노이즈 타입을 가우시안(Gaussian) 노이즈로 판별하고, 상기 푸리에 변환 값들의 합이 상기 미리 결정된 기준 푸리에 변환 값들의 합보다 작으면 상기 노이즈 타입을 블러(blur) 노이즈로 판별할 수

있다.

[0010] 하나의 실시 예에 있어서, 상기 노이즈 분류기는 상기 푸리에 변환 값들의 합의 크기에 기초하여 상기 노이즈 강도를 판별할 수 있다.

[0011] 하나의 실시 예에 있어서, 상기 노이즈 분류기는 뉴럴 네트워크를 기반으로 상기 노이즈 타입 및 상기 노이즈 강도를 판별할 수 있다.

[0012] 하나의 실시 예에 있어서, 상기 복수의 뉴럴 네트워크들 각각은 상기 노이즈 타입 및 상기 노이즈 강도의 조합들 각각에 대응하는 노이즈를 포함하는 입력 데이터를 기반으로 학습된 네트워크일 수 있다.

[0013] 하나의 실시 예에 있어서, 상기 네트워크 선택기는 상기 노이즈 특성에 대응하는 뉴럴 네트워크의 가중치 주소를 저장하도록 구성된 가중치 테이블을 포함할 수 있다.

[0014] 하나의 실시 예에 따른 뉴럴 프로세싱 장치는 상기 복수의 뉴럴 네트워크들의 가중치들을 저장하고, 상기 선택된 뉴럴 네트워크에 대응하는 상기 가중치 주소에 기초하여 상기 선택된 가중치들을 출력하도록 구성된 메모리를 더 포함할 수 있다.

[0015] 본 발명의 하나의 실시 예에 따른 뉴럴 프로세싱 장치의 동작 방법은 입력 데이터에 대한 푸리에(Fourier) 변환 값들에 기초하여 상기 입력 데이터의 노이즈 특성을 판별하는 단계, 상기 노이즈 특성에 기초하여 복수의 뉴럴 네트워크들 중 하나를 선택하는 단계 및 상기 선택된 뉴럴 네트워크에 대응하는 선택된 가중치들을 기반으로 상기 입력 데이터에 대한 추론을 수행하는 단계를 포함한다.

[0016] 하나의 실시 예에 있어서, 상기 노이즈 특성은 상기 입력 데이터의 노이즈 타입(type) 및 노이즈 강도(degree)를 포함할 수 있다.

[0017] 하나의 실시 예에 있어서, 상기 노이즈 타입은 상기 푸리에 변환 값들의 합이 미리 결정된 기준 푸리에 변환 값들의 합보다 크면 가우시안(Gaussian) 노이즈로 판별되고, 상기 푸리에 변환 값들의 합이 상기 미리 결정된 기준 푸리에 변환 값들의 합보다 작으면 블러(blur) 노이즈로 판별될 수 있다.

[0018] 하나의 실시 예에 있어서, 상기 노이즈 강도는 상기 푸리에 변환 값들의 합의 크기에 기초하여 판별될 수 있다.

[0019] 하나의 실시 예에 따른 동작 방법은 상기 노이즈 타입 및 상기 노이즈 강도의 조합들 각각에 대응하는 노이즈를 포함하는 입력 데이터를 기반으로 상기 복수의 뉴럴 네트워크들 각각에 대응하는 가중치들을 학습하는 단계를 더 포함하고, 상기 학습된 가중치들을 기반으로 상기 추론이 수행될 수 있다.

[0020] 본 발명의 하나의 실시 예에 따른 뉴럴 프로세싱 장치의 동작 방법은 특정 뉴럴 네트워크를 기반으로 입력 데이터의 노이즈 특성을 판별하는 단계, 상기 노이즈 특성에 기초하여 복수의 뉴럴 네트워크들 중 하나를 선택하는 단계 및 상기 선택된 뉴럴 네트워크에 대응하는 선택된 가중치들을 기반으로 상기 입력 데이터에 대한 추론을 수행하는 단계를 포함할 수 있다.

[0021] 하나의 실시 예에 있어서, 상기 노이즈 특성은 상기 입력 데이터의 노이즈 타입(type) 및 노이즈 강도(degree)를 포함할 수 있다.

[0022] 하나의 실시 예에 있어서, 상기 특정 뉴럴 네트워크는 컨볼루셔널 뉴럴 네트워크(CNN; Convolutional Neural Network)일 수 있다.

발명의 효과

[0023] 본 발명의 실시 예에 따르면, 입력 데이터의 노이즈 특성에 따라 연산 리소스를 적게 사용하면서 원하는 정확도로 추론을 수행하는 뉴럴 프로세싱 장치를 제공할 수 있다.

[0024] 또한, 본 발명의 실시 예에 따르면, 저 전력으로 동작하는 뉴럴 프로세싱 장치를 제공할 수 있다.

도면의 간단한 설명

[0025] 도 1은 본 발명의 하나의 실시 예에 따른 뉴럴 프로세싱 장치를 나타내는 블록도이다.

도 2는 본 발명의 하나의 실시 예에 따른 뉴럴 네트워크를 나타내는 도면이다.

도 3은 도 1의 뉴럴 프로세싱 장치를 상세하게 나타내는 블록도이다.

도 4는 도 3의 노이즈 분류기의 하나의 예시를 나타내는 블록도이다.

도 5는 도 4의 노이즈 분류기의 동작에 따른 푸리에 변환의 예시를 나타내는 도면이다.

도 6은 도 3의 노이즈 분류기의 다른 예시를 나타내는 도면이다.

도 7은 도 3의 네트워크 선택기의 하나의 예시를 나타내는 도면이다.

도 8a 내지 도 8c는 도 7의 뉴럴 네트워크의 예시를 나타낸다.

도 9a 및 도 9b는 도 3의 뉴럴 프로세싱 장치의 동작의 예시를 나타내는 도면이다.

도 10은 도 1의 뉴럴 프로세싱 장치의 동작의 하나의 예시를 나타내는 순서도이다.

도 11은 도 1의 뉴럴 프로세싱 장치가 노이즈를 분류하는 동작의 하나의 예시를 나타내는 순서도이다.

도 12는 본 발명의 실시 예에 따른 추론 결과의 예시를 보여주는 도면이다.

도 13은 본 발명의 실시 예에 따른 추론 결과와 종래 기술에 따른 추론 결과의 비교 예시를 보여주는 도면이다.

도 14는 본 발명의 실시 예들에 따른 뉴럴 프로세싱 장치를 포함하는 컴퓨팅 시스템을 나타내는 블록도이다.

발명을 실시하기 위한 구체적인 내용

[0026] 이하에서, 본 발명의 기술 분야에서 통상의 지식을 가진 자가 본 발명을 용이하게 실시할 수 있을 강도로, 본 발명의 실시 예들이 명확하고 상세하게 기재될 것이다.

[0027] 도 1은 본 발명의 하나의 실시 예에 따른 뉴럴 프로세싱 장치를 나타내는 블록도이다. 도 1을 참조하면, 뉴럴 프로세싱 장치(100)는 입력 데이터(IN)를 수신할 수 있다. 뉴럴 프로세싱 장치(100)는 입력 데이터(IN)에 대한 추론을 수행할 수 있다. 뉴럴 프로세싱 장치(100)는 추론 결과로서 결과 데이터(OUT)를 출력할 수 있다. 예를 들어, 뉴럴 프로세싱 장치(100)는 입력 데이터(IN)로서 숫자가 포함된 이미지를 수신할 수 있다. 뉴럴 프로세싱 장치(100)는 이미지에 포함된 숫자를 식별할 수 있다. 뉴럴 프로세싱 장치(100)는 식별된 숫자를 결과 데이터(OUT)로서 출력할 수 있다.

[0028] 뉴럴 프로세싱 장치(100)는 뉴럴 네트워크를 이용하여 추론을 수행할 수 있다. 예시적으로, 뉴럴 프로세싱 장치(100)는 복수의 뉴럴 네트워크들 중 하나를 기반으로 추론을 수행할 수 있다. 예를 들어, 본 발명의 실시 예들에 따른 복수의 뉴럴 네트워크들은 컨볼루셔널 뉴럴 네트워크(CNN; Convolutional Neural Network)를 기반으로 구성될 수 있다. 그러나, 본 발명은 이에 한정되지 않으며, 복수의 뉴럴 네트워크들은 순환 뉴럴 네트워크(RNN; Recurrent Neural Network), 스파이킹 뉴럴 네트워크(SNN; Spiking Neural Network) 등 다양한 방식의 뉴럴 네트워크를 기반으로 구성될 수 있다.

[0029] 뉴럴 프로세싱 장치(100)는 노이즈 분류기(110) 및 네트워크 선택기(120)를 포함할 수 있다. 노이즈 분류기(110)는 입력 데이터(IN)에 대하여 전처리할 수 있다. 노이즈 분류기(110)는 전처리를 통해 입력 데이터(IN)의 노이즈 특성(NSC)을 판별할 수 있다. 노이즈 특성(NSC)은 입력 데이터(IN)에 포함된 노이즈의 다양한 정보를 포함할 수 있다. 예를 들어, 노이즈 특성(NSC)은 노이즈 타입(type) 및 노이즈 강도(degree)를 포함할 수 있다.

[0030] 네트워크 선택기(120)는 노이즈 특성(NSC)에 기초하여 복수의 뉴럴 네트워크들 중 하나를 선택할 수 있다. 복수의 뉴럴 네트워크들은 다양한 노이즈 특성(NSC)을 포함하는 입력 데이터(IN)를 기반으로 미리 학습될 수 있다. 네트워크 선택기(120)는 미리 학습된 뉴럴 네트워크들 중 하나를 선택할 수 있다.

[0031] 복수의 뉴럴 네트워크들 각각은 특정 노이즈 특성(NSC)을 가지는 입력 데이터(IN)에 대응하는 뉴럴 네트워크일 수 있다. 예를 들어, 제1 뉴럴 네트워크는 제1 노이즈 특성(NSC)을 가지는 입력 데이터(IN)에 대응하는 뉴럴 네트워크이고, 제2 뉴럴 네트워크는 제2 노이즈 특성(NSC)을 가지는 입력 데이터(IN)에 대응하는 뉴럴 네트워크일 수 있다. 즉, 네트워크 선택기(120)는 복수의 뉴럴 네트워크들 중 노이즈 특성(NSC)에 대응하는 뉴럴 네트워크를 선택할 수 있다.

[0032] 대응하는 뉴럴 네트워크를 이용하여 입력 데이터(IN)에 대한 추론이 수행되는 경우, 뉴럴 프로세싱 장치(100)는 연산 리소스(resource)를 적게 사용하면서 원하는 정확도로 추론을 수행할 수 있다. 즉, 뉴럴 프로세싱 장치(100)의 연산량이 감소될 수 있다.

[0033] 상술한 바와 같이, 뉴럴 프로세싱 장치(100)는 입력 데이터(IN)의 노이즈 특성(NSC)을 기반으로 추론을 수행함

으로써 원하는 정확도를 유지하면서 연산량을 감소시킬 수 있다. 따라서, 뉴럴 프로세싱 장치(100)는 저 전력으로 추론을 수행할 수 있다.

[0034] 도 2는 본 발명의 하나의 실시 예에 따른 뉴럴 네트워크를 나타내는 도면이다. 즉, 도 2는 본 발명의 실시 예에 따른 복수의 뉴럴 네트워크들 중 하나일 수 있다. 도 2를 참조하면, 뉴럴 네트워크는 제1 레이어(layer)(L1), 제2 레이어(L2) 및 시냅스들(S)을 포함할 수 있다. 제1 레이어(L1)는 제1 내지 제3 뉴런들(n1~n3)을 포함하고, 제2 레이어(L2)는 제4 내지 제6 뉴런들(n4~n6)을 포함할 수 있다. 제1 레이어(L1)의 뉴런들(n1~n3)은 시냅스들(S)을 통해 제2 레이어(L2)의 뉴런들(n4~n6)과 연결될 수 있다. 뉴런들(n1~n3)은 시냅스들(S)을 통해 뉴런들(n4~n6)로 신호를 전달할 수 있다.

[0035] 시냅스들(S) 각각은 서로 다른 연결 강도를 가질 수 있다. 시냅스들(S) 각각의 연결 강도는 가중치(weight)로 나타낼 수 있다. 예를 들어, 가중치(w14)는 제1 뉴런(n1)과 제4 뉴런(n4)을 연결하는 시냅스의 가중치 값이고, 가중치(w24)는 제2 뉴런(n2)과 제4 뉴런(n4)을 연결하는 시냅스의 가중치 값일 수 있다.

[0036] 시냅스들(S)에 대응하는 가중치들(W)은 뉴럴 네트워크의 연결 상태를 나타낼 수 있다. 예를 들어, 도 2에 도시된 바와 같이, 제1 레이어(L1)의 뉴런들(n1~n3)과 제2 레이어(L2)의 뉴런들(n4~n6)이 모두 연결된 경우, 가중치들(W)의 값은 모두 '0' 이 아닐 수 있다. 도 2에 도시된 바와 다르게, 제3 뉴런(n3)과 제6 뉴런(n6)이 연결되지 않은 경우, 가중치(w36)는 '0' 일 수 있다.

[0037] 따라서, 가중치들(W)은 뉴럴 네트워크의 구성을 나타내는 변수(parameter)일 수 있다. 즉, 서로 다른 가중치들(W)은 서로 다른 뉴럴 네트워크를 가리킬 수 있다. 즉, 네트워크 선택기(120)에 의해 뉴럴 네트워크가 선택되는 것은 가중치들(W)이 선택되는 것을 의미할 수 있다.

[0038] 도 2에 도시된 바와 같이, 본 발명의 실시 예에 따른 뉴럴 네트워크들은 2 개의 레이어들(L1, L2) 및 6 개의 뉴런들(n1~n6)을 포함할 수 있지만, 본 발명은 이에 한정되지 않는다. 예를 들어, 본 발명의 실시 예들에 따른 뉴럴 네트워크들은 다양한 개수의 레이어들 및 뉴런들을 포함할 수 있다. 이하에서는, 설명의 편의를 위해, 도 2와 같이 2 개의 레이어들(L1, L2) 및 6 개의 뉴런들(n1~n6)을 포함하는 뉴럴 네트워크들을 기반으로 뉴럴 프로세싱 장치(100)가 추론을 수행하는 것으로 가정한다.

[0039] 도 3은 도 1의 뉴럴 프로세싱 장치를 상세하게 나타내는 블록도이다. 도 1을 참조하면, 뉴럴 프로세싱 장치(100)는 노이즈 분류기(110), 네트워크 선택기(120), 메모리(130) 및 연산기(140)를 포함할 수 있다.

[0040] 노이즈 분류기(110)는 입력 데이터(IN)를 수신할 수 있다. 노이즈 분류기(110)는 입력 데이터(IN)를 전처리하여 입력 데이터(IN)에 포함된 노이즈 특성(NSC)을 판별할 수 있다. 예를 들어, 노이즈 특성(NSC)은 노이즈 타입(type) 및 노이즈 강도(degree)를 포함할 수 있다. 노이즈 분류기(110)는 판별된 노이즈 특성(NSC)을 네트워크 선택기(120)로 제공할 수 있다.

[0041] 네트워크 선택기(120)는 노이즈 특성(NSC)에 기초하여 복수의 뉴럴 네트워크들 중 하나를 선택할 수 있다. 네트워크 선택기(120)는 선택된 뉴럴 네트워크에 대응하는 가중치 주소(WA)를 출력할 수 있다. 가중치 주소(WA)는 뉴럴 네트워크의 가중치들(W)이 저장된 메모리(130)의 주소 정보일 수 있다. 예를 들어, 네트워크 선택기(120)가 도 2의 뉴럴 네트워크를 선택하는 경우, 가중치 주소(WA)는 가중치들(W)이 저장된 메모리(130)의 주소 정보일 수 있다. 네트워크 선택기(120)는 가중치 주소(WA)를 메모리(130)로 제공할 수 있다.

[0042] 메모리(130)는 가중치 주소(WA)에 기초하여 가중치들(W)을 출력할 수 있다. 출력된 가중치들(W)은 연산기(140)로 제공될 수 있다.

[0043] 연산기(140)는 가중치들(W) 및 입력 데이터(IN)를 수신할 수 있다. 연산기(140)는 가중치들(W)과 입력 데이터(IN)를 이용하여 연산을 수행할 수 있다. 예를 들어, 연산기(140)는 가중치들(W)과 입력 데이터(IN)의 곱셈을 수행할 수 있다. 연산기(140)는 가중치들(W)과 입력 데이터(IN)의 곱셈 결과들을 누적할 수 있다. 즉, 연산기(140)는 곱셈기 또는 누적기를 포함할 수 있다. 연산기(140)는 곱셈, 누적 등의 다양한 연산을 통해 입력 데이터(IN)에 대한 추론을 수행할 수 있다. 연산기(140)는 추론 결과로서 결과 데이터(OUT)를 출력할 수 있다.

[0044] 이하에서는 도 4 내지 도 9를 참조로 하여 도 3의 뉴럴 프로세싱 장치의 동작을 상세하게 설명할 것이다.

[0045] 도 4는 도 3의 노이즈 분류기의 하나의 예시를 나타내는 블록도이다. 도 5는 도 4의 노이즈 분류기의 동작에 따른 푸리에 변환의 예시를 나타내는 도면이다. 도 4를 참조하면, 노이즈 분류기(110)는 푸리에 변환기(111) 및 푸리에 변환 값 비교기(112)를 포함할 수 있다.

[0046] 푸리에 변환기(111)는 입력 데이터(IN)에 대하여 푸리에 변환(Fourier transform)을 수행할 수 있다. 푸리에 변환기(111)는 입력 데이터(IN)의 공간 영역(spatial domain)의 값으로부터 주파수 영역(frequency domain)의 값을 획득할 수 있다. 푸리에 변환기(111)는 입력 데이터(IN)의 주파수 영역의 값인 푸리에 변환 값(FTV)을 출력할 수 있다. 예를 들어, 푸리에 변환기(111)는 고속 푸리에 변환(FFT; Fast Fourier Transform)을 수행하여 푸리에 변환 값(FTV)을 출력할 수 있다.

[0047] 예를 들어, 도 5를 참조하면, 입력 데이터(IN)는 이미지일 수 있다. 이미지는 노이즈를 포함하거나 노이즈를 포함하지 않을 수 있다. 노이즈가 포함되지 않은 이미지는 클린(clean) 이미지로 지칭될 수 있다. 노이즈가 포함된 이미지는 가우시안(Gaussian) 노이즈 또는 블러(blur) 노이즈를 포함할 수 있다. 가우시안 노이즈는 원자의 열에너지로 인한 진동이나 열에너지로 인한 방사에 인해 발생될 수 있다. 블러 노이즈는 사진이나 영상 촬영 시 렌즈의 흔들림에 의해 발생될 수 있다. 그러나, 본 발명은 이에 한정되지 않으며, 이미지는 다양한 타입의 노이즈를 포함할 수 있다. 예를 들어, 이미지는 사진 촬영 시 물체가 가려짐에 의해 발생되는 노이즈를 포함할 수 있다. 도 5에 도시된 바와 같이, 클린 이미지의 물체는 명확하게 구분될 수 있으나, 노이즈가 포함된 이미지의 물체는 명확하게 구분되지 않을 수 있다.

[0048] 푸리에 변환기(111)는 이미지의 픽셀(pixel) 값들에 대하여 푸리에 변환을 수행할 수 있다. 푸리에 변환기(111)는 픽셀 값들에 대한 푸리에 변환 값들을 획득할 수 있다. 푸리에 변환 값들은 노이즈 타입에 따라 달라질 수 있다. 예를 들어, 가우시안 노이즈가 포함된 이미지는 블러 노이즈가 포함된 이미지와 비교하여 고 주파수 값을 더 포함할 수 있다. 이에 따라, 가우시안 노이즈가 포함된 이미지의 푸리에 변환 값들은 블러 노이즈가 포함된 이미지의 푸리에 변환 값들보다 클 수 있다.

[0049] 즉, 이미지의 특정 영역에서의 푸리에 변환 값들의 합은 노이즈 타입에 따라 달라질 수 있다. 도 5에 도시된 바와 같이, 클린 이미지의 특정 영역에서의 푸리에 변환 값들의 합(FS1), 가우시안 노이즈가 포함된 이미지의 특정 영역에서의 푸리에 변환 값들의 합(FS2) 및 블러 노이즈가 포함된 이미지의 특정 영역에서의 푸리에 변환 값들의 합(FS3)은 서로 다를 수 있다. 예를 들어, 클린 이미지의 푸리에 변환 값들의 합(FS1)은 가우시안 노이즈가 포함된 이미지의 푸리에 변환 값들의 합(FS2)보다 작을 수 있다. 클린 이미지의 푸리에 변환 값들의 합(FS1)은 블러 노이즈가 포함된 이미지의 푸리에 변환 값들의 합(FS3)보다 클 수 있다.

[0050] 다시 도 4를 참조하면, 푸리에 변환 값 비교기(112)는 푸리에 변환 값(FTV)을 수신할 수 있다. 푸리에 변환 값 비교기(112)는 푸리에 변환 값(FTV)들의 합을 기준 푸리에 변환 값들의 합(RFS)과 비교할 수 있다. 기준 푸리에 변환 값은 입력 데이터(IN)에 대한 학습 과정에서 미리 결정된 주파수 영역의 값일 수 있다. 예를 들어, 기준 푸리에 변환 값은 도 5의 클린 이미지에 대응하는 주파수 영역의 값일 수 있다. 기준 푸리에 변환 값들의 합(RFS)은 노이즈 분류기(110)의 내부 메모리(미도시) 또는 외부 메모리(예를 들어, 메모리(130))에 저장될 수 있다.

[0051] 예를 들어, 기준 푸리에 변환 값들의 합(RFS)이 도 5의 클린 이미지의 푸리에 변환 값들의 합(FS1)이고, 가우시안 노이즈가 포함된 이미지가 입력 데이터(IN)로 제공되는 경우, 푸리에 변환 값 비교기(112)는 푸리에 변환 값들의 합(FS1)과 푸리에 변환 값들의 합(FS2)을 비교할 수 있다.

[0052] 푸리에 변환 값 비교기(112)는 비교 결과에 기초하여 노이즈 특성(NSC)를 판별할 수 있다. 노이즈 특성(NSC)은 노이즈 타입(N_t) 및 노이즈 강도(N_d)를 포함할 수 있다. 예를 들어, 푸리에 변환 값(FTV)들의 합이 기준 푸리에 변환 값들의 합(RFS)과 동일한 경우, 푸리에 변환 값 비교기(112)는 입력 데이터(IN)의 노이즈 타입(N_t)을 클린 이미지로 판별(즉, 노이즈가 포함되지 않은 것으로 판별)할 수 있다. 푸리에 변환 값(FTV)들의 합이 기준 푸리에 변환 값들의 합(RFS)보다 큰 경우, 푸리에 변환 값 비교기(112)는 입력 데이터(IN)의 노이즈 타입(N_t)을 가우시안 노이즈로 판별할 수 있다. 푸리에 변환 값(FTV)들의 합이 기준 푸리에 변환 값들의 합(RFS)보다 작은 경우, 푸리에 변환 값 비교기(112)는 입력 데이터(IN)의 노이즈 타입(N_t)을 블러 노이즈로 판별할 수 있다.

[0053] 푸리에 변환 값 비교기(112)는 푸리에 변환 값(FTV)들의 합과 기준 푸리에 변환 값들의 합(RFS)의 차이에 기초하여 노이즈 강도(N_d)를 판별할 수 있다. 예를 들어, 푸리에 변환 값(FTV)들의 합과 기준 푸리에 변환 값들의 합(RFS)의 차이가 제1 임계값 이하인 경우, 푸리에 변환 값 비교기(112)는 입력 데이터(IN)의 노이즈 강도(N_d)를 제1 강도(first degree)로 판별할 수 있다. 푸리에 변환 값(FTV)들의 합과 기준 푸리에 변환 값들의 합(RFS)의 차이가 제1 임계값을 초과하고 제2 임계값 이하인 경우, 푸리에 변환 값 비교기(112)는 입력 데이터(IN)의 노이즈 강도(N_d)를 제2 강도(second degree)로 판별할 수 있다.

[0054] 즉, 푸리에 변환 값(FTV)들의 합이 기준 푸리에 변환 값들의 합(RFS)보다 크고, 차이가 제1 임계값 이하인

경우, 푸리에 변환 값 비교기(112)는 입력 데이터(IN)의 노이즈 타입(N_t)을 가우시안 노이즈로 판별하고, 노이즈 강도(N_d)를 제1 강도로 판별할 수 있다.

[0055] 도 6은 도 3의 노이즈 분류기의 다른 예시를 나타내는 도면이다. 도 6을 참조하면, 노이즈 분류기(110)는 뉴럴 네트워크를 이용하여 입력 데이터(IN)에 대한 노이즈 특성(NSC)을 판별할 수 있다. 노이즈 분류기(110)에서 이용되는 뉴럴 네트워크는 도 3의 연산기(140)에서 이용되는 뉴럴 네트워크와 다를 수 있다. 즉, 입력 데이터(IN)의 노이즈 특성(NSC)을 판별하기 위한 뉴럴 네트워크와 추론을 위한 뉴럴 네트워크는 다를 수 있다. 예를 들어, 노이즈 분류기(110)에서 이용되는 뉴럴 네트워크는 연산기(140)에서 이용되는 뉴럴 네트워크보다 더 적은 레이어들 및 뉴런들을 포함할 수 있다. 즉, 노이즈 분류기(110)에서 이용되는 뉴럴 네트워크의 사이즈는 연산기(140)에서 이용되는 뉴럴 네트워크의 사이즈보다 작을 수 있다.

[0056] 도 6에 도시된 바와 같이, 노이즈 분류기(110)는 컨볼루셔널 뉴럴 네트워크(CNN)를 기반으로 노이즈 특성(NSC)을 판별할 수 있다. CNN은 합성곱층(CL, convolution layer), 풀링층(PL, pooling layer) 및 전결합층(FCL, fully connected layer)을 포함할 수 있다.

[0057] 노이즈 분류기(110)는 합성곱층(CL)을 통해 입력 데이터(IN)로부터 특징 데이터(feature data)를 추출할 수 있다. 노이즈 분류기(110)는 풀링층(PL)을 통해 추출된 특징 데이터를 샘플링할 수 있다. 노이즈 분류기(110)는 전결합층(FCL)을 통해 샘플된 데이터로부터 노이즈 특성(NSC)을 판별할 수 있다.

[0058] 예를 들어, 입력 데이터(IN)로서 제2 강도의 블러 노이즈가 포함된 이미지가 제공되는 경우, 노이즈 분류기(110)는 뉴럴 네트워크를 기반으로 노이즈 특성(NSC)을 판별할 수 있다. 노이즈 분류기(110)는 입력 데이터(IN)의 노이즈 타입(N_t)이 블러 노이즈이고, 노이즈 강도(N_d)가 제2 강도인 것을 판별할 수 있다.

[0059] 도 6에는 합성곱층(CL), 풀링층(PL) 및 전결합층(FCL)이 각각 하나인 것으로 도시되었지만, 본 발명은 이에 한정되지 않는다. 예를 들어, 노이즈 분류기(110)는 다양한 개수의 합성곱층(CL), 풀링층(PL) 및 전결합층(FCL)을 이용하여 입력 데이터(IN)의 노이즈 특성(NSC)을 판별할 수 있다.

[0060] 도 7은 도 3의 네트워크 선택기의 하나의 예시를 나타내는 도면이다. 도 7을 참조하면, 네트워크 선택기(120)는 노이즈 특성(NSC)을 수신할 수 있다. 네트워크 선택기(120)는 가중치 테이블(WT)에 기초하여 수신된 노이즈 특성(NSC)에 대응하는 뉴럴 네트워크(NN) 및 가중치 주소(WA)를 결정할 수 있다. 네트워크 선택기(120)는 결정된 가중치 주소(WA)를 출력할 수 있다.

[0061] 도 7에 도시된 바와 같이, 네트워크 선택기(120)는 가중치 테이블(WT)을 포함할 수 있다. 가중치 테이블(WT)은 노이즈 특성(NSC) 필드, 뉴럴 네트워크(NN) 필드 및 가중치 주소(WA) 필드를 포함할 수 있다. 가중치 테이블(WT)은 네트워크 선택기(120) 내부의 메모리에 저장될 수 있지만, 본 발명은 이에 한정되지 않는다.

[0062] 노이즈 특성(NSC) 필드에는 노이즈 타입(N_t) 및 노이즈 강도(N_d)에 따라 제1 내지 제7 노이즈 특성들(NSC1~NSC7)이 포함될 수 있다. 예를 들어, 제1 노이즈 특성(NSC1)은 입력 데이터(IN)에 노이즈가 포함되지 않음을 나타낼 수 있다. 제2 노이즈 특성(NSC2)은 입력 데이터(IN)에 제1 강도(D1)의 가우시안 노이즈(Ng)가 포함됨을 나타낼 수 있다. 제7 노이즈 특성(NSC7)은 입력 데이터(IN)에 제3 강도(D3)의 블러 노이즈(Nb)가 포함됨을 나타낼 수 있다.

[0063] 뉴럴 네트워크(NN) 필드에는 제1 내지 제7 뉴럴 네트워크들(NN1~NN7)이 포함될 수 있다. 제1 내지 제7 뉴럴 네트워크들(NN1~NN7)은 제1 내지 제7 노이즈 특성들(NSC1~NSC7)에 각각 대응할 수 있다. 예를 들어, 제1 뉴럴 네트워크(NN1)는 제1 노이즈 특성(NSC1)에 대응하고, 제2 뉴럴 네트워크(NN2)는 제2 노이즈 특성(NSC2)에 대응할 수 있다.

[0064] 입력 데이터(IN)에 노이즈가 포함되지 않은 경우, 제1 뉴럴 네트워크(NN1)는 입력 데이터(IN)에 대응하는 뉴럴 네트워크일 수 있다. 즉, 제1 뉴럴 네트워크(NN1)를 이용하여 입력 데이터(IN)에 대한 추론이 수행되는 경우, 최소화된 연산량 및 원하는 정확도로 추론이 수행될 수 있다.

[0065] 입력 데이터(IN)에 제1 강도(D1)의 가우시안 노이즈(Ng)가 포함된 경우, 제2 뉴럴 네트워크(NN2)는 입력 데이터(IN)에 대응하는 뉴럴 네트워크일 수 있다. 즉, 제2 뉴럴 네트워크(NN2)를 이용하여 입력 데이터(IN)에 대한 추론이 수행되는 경우, 최소화된 연산량 및 원하는 정확도로 추론이 수행될 수 있다.

[0066] 이 경우, 제1 뉴럴 네트워크(NN1)는 제2 뉴럴 네트워크(NN2)보다 희소성(sparsity)이 더 클 수 있다. 예를 들어, 제1 뉴럴 네트워크(NN1)의 가중치들(W)은 제2 뉴럴 네트워크(NN2)의 가중치들(W)보다 '0' 인 가중치를 더 포함할 수 있다. 또는 제1 뉴럴 네트워크(NN1)의 가중치들(W)의 값은 제2 뉴럴 네트워크(NN2)의 가중치들(W)의

값보다 작을 수 있다. 이에 따라, 노이즈가 포함되지 않은 입력 데이터(IN)에 대한 추론은 노이즈가 포함된 입력 데이터(IN)에 대한 추론보다 더 근사적으로 수행될 수 있다. 이에 따라, 제1 뉴럴 네트워크(NN1)를 기반으로 수행되는 연산량이 제2 뉴럴 네트워크(NN2)를 기반으로 수행되는 연산량보다 적을 수 있다. 노이즈가 포함되지 않은 입력 데이터(IN)에 대한 추론은 입력 데이터(IN)에 노이즈가 포함된 경우보다 더 근사적으로 수행되더라도, 원하는 정확도를 가질 수 있다.

[0067] 제2 뉴럴 네트워크(NN2)를 이용하여 노이즈가 포함되지 않은 입력 데이터(IN)에 대한 추론이 수행되는 경우, 원하는 정확도에 의해 추론이 수행될 수 있으나, 연산량이 증가될 수 있다. 따라서, 노이즈가 포함되지 않은 입력 데이터(IN)(즉, 제1 노이즈 특성(NSC1))에 대하여 최소화된 연산량 및 원하는 정확도로써 추론이 수행되기 위해, 제1 뉴럴 네트워크(NN1)는 제1 노이즈 특성(NSC1)에 대응될 수 있다.

[0068] 마찬가지로, 제3 노이즈 특성(NSC3)의 입력 데이터(IN)에 대하여 최소화된 연산량 및 원하는 정확도로써 추론이 수행되기 위해, 제3 뉴럴 네트워크(NN3)는 제3 노이즈 특성(NSC3)에 대응될 수 있다. 제4 노이즈 특성(NSC4)의 입력 데이터(IN)에 대하여 최소화된 연산량 및 원하는 정확도로써 추론이 수행되기 위해, 제4 뉴럴 네트워크(NN4)는 제4 노이즈 특성(NSC4)에 대응될 수 있다. 이 경우, 제3 뉴럴 네트워크(NN3)는 제4 뉴럴 네트워크(NN4)보다 희소성이 더 클 수 있다. 예를 들어, 제3 뉴럴 네트워크(NN3)의 가중치들(W)은 제4 뉴럴 네트워크(NN4)의 가중치들(W)보다 '0' 인 가중치를 더 포함할 수 있다. 또는, 제3 뉴럴 네트워크(NN3)의 가중치들(W)의 값은 제4 뉴럴 네트워크(NN4)의 가중치들(W)들의 값보다 작을 수 있다. 이에 따라, 제2 강도의 노이즈가 포함된 입력 데이터(IN)에 대한 추론은 제3 강도의 노이즈가 포함된 입력 데이터(IN)에 대한 추론보다 더 근사적으로 수행될 수 있다. 즉, 노이즈의 강도(N_d)가 증가될수록 추론의 연산량이 증가될 수 있다.

[0069] 이와 같이, 가중치 테이블(WT)은 각각의 노이즈 특성(NSC)에 대응하는 뉴럴 네트워크(NN)에 대한 정보를 포함할 수 있다.

[0070] 도 7의 제1 내지 제7 뉴럴 네트워크들(NN1~NN7) 각각은 다양한 노이즈 특성(NSC)을 포함하는 입력 데이터(IN)를 기반으로 학습된 뉴럴 네트워크일 수 있다. 학습을 위해 제공되는 입력 데이터(IN)는 다양한 노이즈 타입(N_t) 및 노이즈 강도(N_d)의 조합들 각각에 대응하는 노이즈를 포함할 수 있다.

[0071] 가중치 주소(WA) 필드에는 제1 내지 제7 가중치 주소들(WA1~WA7)이 포함될 수 있다. 제1 내지 제7 가중치 주소들(WA1~WA7)은 제1 내지 제7 뉴럴 네트워크들(NN1~NN7)에 각각 대응될 수 있다. 예를 들어, 제1 뉴럴 네트워크(NN1)의 가중치들(W)은 메모리(130)의 제1 가중치 주소(WA1)에 저장될 수 이다. 제2 뉴럴 네트워크(NN2)의 가중치들(W)은 메모리(130)의 제2 가중치 주소(WA2)에 저장될 수 있다.

[0072] 예를 들어, 제1 노이즈 특성(NSC1)이 수신되는 경우, 네트워크 선택기(120)는 제1 노이즈 특성(NSC1)에 대응하는 제1 가중치 주소(WA1)를 출력할 수 있다. 제2 노이즈 특성(NSC2)이 수신되는 경우, 네트워크 선택기(120)는 제2 노이즈 특성(NSC2)에 대응하는 제2 가중치 주소(WA2)를 출력할 수 있다.

[0073] 도 7에는 가중치 테이블(WT)에 노이즈 특성(NSC) 필드, 뉴럴 네트워크(NN) 필드 및 가중치 주소(WA) 필드가 포함된 것으로 도시되었지만, 본 발명은 이에 한정되지 않는다. 예를 들어, 가중치 테이블(WT)은 노이즈 특성(NSC) 필드 및 가중치 주소(WA) 필드만을 포함할 수 있다.

[0074] 도 7의 가중치 테이블(WT)에는 제1 내지 제7 노이즈 특성들(NSC1~NSC7) 및 제1 내지 제7 가중치 주소들(WA1~WA7)이 포함된 것으로 도시되었지만, 본 발명은 이에 한정되지 않는다. 예를 들어, 가중치 테이블(WT)은 다양한 개수의 노이즈 특성(NSC)들과 이에 대응하는 가중치 주소(WA)들을 포함할 수 있다.

[0075] 도 8a 내지 도 8c는 도 7의 뉴럴 네트워크의 예시를 나타낸다. 구체적으로, 도 8a는 도 7의 제1 뉴럴 네트워크(NN1)의 예시를 나타낸다. 도 8b는 도 7의 제2 뉴럴 네트워크(NN2)의 예시를 나타낸다. 도 8c는 도 7의 제3 뉴럴 네트워크(NN3)의 예시를 나타낸다.

[0076] 도 8a를 참조하면, 제1 뉴럴 네트워크(NN1)는 제1 레이어(L1), 제2 레이어(L2) 및 시냅스들(S1)을 포함할 수 있다. 도 8a에 도시된 바와 같이, 제3 뉴런(n3)은 뉴런들(n4~n6)과 연결되지 않은 상태이고, 제5 뉴런(n5)은 뉴런들(n1~n3)과 연결되지 않은 상태일 수 있다. 이에 따라, 제1 뉴럴 네트워크(NN1)의 제1 가중치들(W1) 중 제3 뉴런(n3) 및 제5 뉴런(n5)에 대응하는 가중치들(w15, w25, w34, w35, w36)은 모두 ′ 0′ 일 수 있다. 따라서, 제1 뉴럴 네트워크(NN1)를 이용하여 추론이 수행되는 경우, 근사적으로 연산이 수행되어 연산량이 적을 수 있다.

[0077] 도 8b를 참조하면, 제2 뉴럴 네트워크(NN2)는 제1 레이어(L1), 제2 레이어(L2) 및 시냅스들(S2)을 포함할 수 있다. 도 8b에 도시된 바와 같이, 제3 뉴런(n3)은 뉴런들(n4~n6)과 연결되지 않은 상태일 수 있다. 이에 따라, 제

2 뉴럴 네트워크(NN2)의 제2 가중치들(W2) 중 제3 뉴런(n3)에 대응하는 가중치들(w34, w35, w36)은 모두 ′ 0′ 일 수 있다. 따라서, 제2 뉴럴 네트워크(NN2)를 이용하여 추론이 수행되는 경우, 제1 뉴럴 네트워크(NN1)를 이용하여 추론이 수행되는 경우보다 연산이 정확하게 수행될 수 있고, 연산량이 증가될 수 있다. 즉, 제2 노이즈 특성(NSC2)을 가지는 입력 데이터(IN)에 대한 추론 연산은 제1 노이즈 특성(NSC1)을 가지는 입력 데이터(IN)에 대한 추론 연산보다 정확하게 수행될 수 있다. 그러나, 제2 노이즈 특성(NSC2)을 가지는 입력 데이터(IN)에 대한 추론 연산량은 제1 노이즈 특성(NSC1)을 가지는 입력 데이터(IN)에 대한 추론 연산량보다 증가될 수 있다.

[0078] 도 8c를 참조하면, 제3 뉴럴 네트워크(NN3)는 제1 레이어(L1), 제2 레이어(L2) 및 시냅스들(S3)을 포함할 수 있다. 도 8c에 도시된 바와 같이, 제3 뉴런(n3)은 제6 뉴런(n6)과 연결되지 않은 상태일 수 있다. 이에 따라, 제3 뉴럴 네트워크(NN3)의 제3 가중치들(W3) 중 가중치(w36)는 ′ 0′ 일 수 있다. 따라서, 제3 뉴럴 네트워크(NN3)를 이용하여 추론이 수행되는 경우, 제2 뉴럴 네트워크(NN2)를 이용하여 추론이 수행되는 경우보다 정확하게 연산이 수행될 수 있고, 연산량이 증가될 수 있다. 즉, 제3 노이즈 특성(NSC3)을 가지는 입력 데이터(IN)에 대한 추론 연산은 제2 노이즈 특성(NSC2)을 가지는 입력 데이터(IN)에 대한 추론 연산보다 정확하게 수행될 수 있다. 그러나, 제3 노이즈 특성(NSC3)을 가지는 입력 데이터(IN)에 대한 추론 연산량은 제2 노이즈 특성(NSC2)을 가지는 입력 데이터(IN)에 대한 추론 연산량보다 증가될 수 있다.

[0079] 상술한 바와 같이, 노이즈의 강도(N_d)가 증가될수록 대응하는 뉴럴 네트워크의 희소성이 감소될 수 있다. 즉, 네트워크 선택기(120)는 노이즈 특성(NSC)에 따라 복수의 뉴럴 네트워크들 중 적합한 희소성을 가지는 뉴럴 네트워크를 선택할 수 있다.

[0080] 도 9a 및 도 9b는 도 3의 뉴럴 프로세싱 장치의 동작의 예시를 나타내는 도면이다. 구체적으로, 도 9a는 입력 데이터(IN)에 제1 강도(D1)의 가우시안 노이즈(Ng)가 포함된 예시를 나타낸다. 도 9b는 입력 데이터(IN)에 제3 강도(D3)의 블러 노이즈(Nb)가 포함된 예시를 나타낸다.

[0081] 도 9a 및 도 9b를 참조하면, 입력 데이터(IN)는 이미지일 수 있다. 입력 이미지는 3 채널의 컬러 이미지일 수 있다. 예를 들어, 입력 데이터(IN)로서 이미지의 RGB(red, green, blue) 3 채널의 픽셀 값들이 제공될 수 있다. 도 9a 및 도 9b에 도시된 바와 같이, 3 채널의 입력 데이터(IN)가 제공되는 경우, 뉴럴 네트워크의 가중치들(W)이 3 차원이므로, 입력 데이터(IN)의 추론에 이용되는 뉴럴 네트워크는 3 차원으로 도시될 수 있다.

[0082] 도 9a 및 도 9b의 뉴럴 네트워크들은 도 7의 제1 내지 제7 뉴럴 네트워크들에 대응할 수 있다. 예를 들어, 가중치들(C)에 대응하는 뉴럴 네트워크는 도 7의 제1 뉴럴 네트워크(NN1)일 수 있다. 가중치들(C, G1)에 대응하는 뉴럴 네트워크는 도 7의 제2 뉴럴 네트워크(NN2)일 수 있다.

[0083] 마찬가지로, 가중치들(C, G1, G2)에 대응하는 뉴럴 네트워크는 도 7의 제3 뉴럴 네트워크(NN3)일 수 있다. 가중치들(C, G1, G2, G3)에 대응하는 뉴럴 네트워크는 도 7의 제4 뉴럴 네트워크(NN4)일 수 있다. 가중치들(C, B1)에 대응하는 뉴럴 네트워크는 도 7의 제5 뉴럴 네트워크(NN5)일 수 있다. 가중치들(C, B1, B2)에 대응하는 뉴럴 네트워크는 도 7의 제6 뉴럴 네트워크(NN6)일 수 있다. 가중치들(C, B1, B2, B3)에 대응하는 뉴럴 네트워크는 도 7의 제7 뉴럴 네트워크(NN7)일 수 있다.

[0084] 도 9a 및 도 9b에 도시된 바와 같이, 복수의 뉴럴 네트워크들은 컨볼루셔널 뉴럴 네트워크(CNN) 기반의 뉴럴 네트워크일 수 있다. 컨볼루셔널 뉴럴 네트워크(CNN)는 합성곱층(CL), 풀링층(PL) 및 전결합층(FCL)을 포함할 수 있다. 합성곱층(CL)은 필터(filter)를 이용하여 입력 데이터(IN)에 대하여 합성곱 처리를 수행하고, 특징 데이터를 출력할 수 있다. 필터는 뉴럴 네트워크의 가중치들(W)에 대응될 수 있다. 풀링층(PL)은 합성곱층(CL)에서 출력된 특징 데이터에 대한 샘플링을 수행할 수 있다. 합성곱층(CL)과 마찬가지로 샘플링을 위해 필터가 이용될 수 있다. 예를 들어, 합성곱층(CL)은 필터를 이용하여 특징 데이터의 특정 영역에서 최대값을 추출할 수 있다. 필터는 뉴럴 네트워크의 가중치들(W)에 대응될 수 있다. 전결합층(FCL)은 가중치들(W)을 이용하여 샘플된 데이터로부터 결과 데이터(OUT)를 출력할 수 있다. 결과 데이터(OUT)는 입력 데이터(IN)에 대한 추론 결과일 수 있다. 예를 들어, 결과 데이터(OUT)는 이미지의 물체에 대한 식별 정보를 포함할 수 있다.

[0085] 도 3 및 도 9a를 참조하면, 노이즈 분류기(110)는 입력 데이터(IN)를 수신하고, 입력 데이터(IN)가 제1 강도(D1)의 가우시안 노이즈(Ng)를 포함하는 것으로 판별할 수 있다. 네트워크 선택기(120)는 판별된 노이즈 특성(NSC)(즉, 제1 강도(D1)의 가우시안 노이즈(Ng))에 대응하는 뉴럴 네트워크를 선택할 수 있다. 네트워크 선택기(120)는 가중치들(C, G1)에 대응하는 뉴럴 네트워크를 선택할 수 있다. 예를 들어, 네트워크 선택기(120)는 도 7의 제2 뉴럴 네트워크(NN2)를 선택할 수 있다. 네트워크 선택기(120)는 제2 뉴럴 네트워크(NN2)에 대응하는 제2 가중치 주소(WA2)를 출력할 수 있다. 메모리(130)는 제2 가중치 주소(WA2)에 응답하여 가중치들(C, G1)을 출

력할 수 있다. 연산기(140)는 가중치들(C, G1)을 기반으로 입력 데이터(IN)에 대한 추론을 수행할 수 있다.

[0086] 도 3 및 도 9b를 참조하면, 노이즈 분류기(110)는 입력 데이터(IN)를 수신하고, 입력 데이터(IN)가 제3 강도(D3)의 블러 노이즈(Nb)를 포함하는 것으로 판별할 수 있다. 네트워크 선택기(120)는 판별된 노이즈 특성(NSC)(즉, 제3 강도(D3)의 블러 노이즈(Nb))에 대응하는 뉴럴 네트워크를 선택할 수 있다. 네트워크 선택기(120)는 가중치들(C, B1, B2, B3)에 대응하는 뉴럴 네트워크를 선택할 수 있다. 예를 들어, 네트워크 선택기(120)는 도 7의 제7 뉴럴 네트워크(NN7)를 선택할 수 있다. 네트워크 선택기(120)는 제7 뉴럴 네트워크(NN7)에 대응하는 제7 가중치 주소(WA7)를 출력할 수 있다. 메모리(130)는 제7 가중치 주소(WA7)에 응답하여 가중치들(C, B1, B2, B3)을 출력할 수 있다. 연산기(140)는 가중치들(C, B1, B2, B3)을 기반으로 입력 데이터(IN)에 대한 추론을 수행할 수 있다.

[0087] 상술한 바와 같이, 뉴럴 프로세싱 장치(100)는 입력 데이터(IN)의 노이즈 특성(NSC)에 대응하는 뉴럴 네트워크를 기반으로 추론을 수행할 수 있다. 이 경우, 뉴럴 네트워크는 노이즈 타입(N_t) 및 노이즈 강도(N_d)에 따라 달라질 수 있다. 예를 들어, 노이즈 강도(N_d)가 클수록 대응하는 뉴럴 네트워크의 사이즈가 커질 수 있다. 즉, 뉴럴 프로세싱 장치(100)는 노이즈 강도(N_d)가 작은 입력 데이터(IN)에 대하여 사이즈가 작은 뉴럴 네트워크를 기반으로 추론을 수행할 수 있다. 따라서, 뉴럴 프로세싱 장치(100)는 최소화된 연산 리소스를 사용하면서 원하는 정확도로 추론을 수행할 수 있다.

[0088] 도 1 내지 도 9b에서 설명된 뉴럴 프로세싱 장치(100)는 사물인터넷(IoT) 센서, 스마트폰, 데스크탑, 서버 등에 탑재될 수 있다. 뉴럴 프로세싱 장치(100)는 중앙 처리 장치(CPU), 애플리케이션 프로세서(AP), 마이크로 컨트롤러(MCU) 또는 가속기 등에 포함될 수 있다. 그러나, 본 발명은 이에 한정되지 않으며, 뉴럴 프로세싱 장치(100)는 별도의 장치 또는 칩 형태로 구현될 수 있다.

[0089] 뉴럴 프로세싱 장치(100) 및 뉴럴 프로세싱 장치(100)의 구성 요소들은 임의의 적절한 하드웨어, 소프트웨어, 펌웨어 또는 하드웨어, 소프트웨어, 펌웨어의 조합을 활용하여 구현될 수 있다. 예시적으로, 소프트웨어는 기계 코드, 펌웨어, 임베디드 코드, 및 애플리케이션 소프트웨어일 수 있다. 예를 들어, 하드웨어는 전기 회로, 전자 회로, 프로세서, 컴퓨터, 집적 회로, 집적 회로 코어들, 압력 센서, 관성 센서, 멤즈(Micro Electro Mechanical System; MEMS), 수동 소자, 또는 그것들의 조합을 포함할 수 있다.

[0090] 도 10은 도 1의 뉴럴 프로세싱 장치의 동작의 하나의 예시를 나타내는 순서도이다. 도 1 및 도 10을 참조하면, S101 단계에서, 뉴럴 프로세싱 장치(100)는 입력 데이터(IN)의 노이즈 특성(NSC)을 판별할 수 있다. 예를 들어, 뉴럴 프로세싱 장치(100)는 노이즈 타입(N_t) 및 노이즈 강도(N_d)를 판별할 수 있다.

[0091] S102 단계에서, 뉴럴 프로세싱 장치(100)는 노이즈 특성(NSC)에 기초하여 복수의 뉴럴 네트워크들 중 하나를 선택할 수 있다. 예를 들어, 뉴럴 프로세싱 장치(100)는 노이즈 강도(N_d)가 작은 경우, 희소성이 큰(즉, 사이즈가 작은) 뉴럴 네트워크를 선택할 수 있다. 뉴럴 프로세싱 장치(100)는 노이즈 강도(N_d)가 큰 경우, 희소성이 작은(즉, 사이즈가 큰) 뉴럴 네트워크를 선택할 수 있다.

[0092] S103 단계에서, 뉴럴 프로세싱 장치(100)는 선택된 뉴럴 네트워크에 대응하는 가중치들(W)을 기반으로 입력 데이터(IN)에 대한 추론을 수행할 수 있다.

[0093] 뉴럴 프로세싱 장치(100)는 추론을 수행하기 전에, 학습을 통해 결정된 가중치들(W)을 미리 저장할 수 있다. 뉴럴 프로세싱 장치(100)는 다양한 노이즈를 포함하는 입력 데이터(IN)를 기반으로 학습을 수행할 수 있다. 학습을 위한 입력 데이터(IN)는 노이즈 타입(N_t) 및 노이즈 강도(N_d)의 조합들 각각에 대응하는 노이즈를 포함할 수 있다. 뉴럴 프로세싱 장치(100)는 학습을 통해 복수의 뉴럴 네트워크들 각각에 대응하는 가중치들(W)을 학습할 수 있다. 즉, 학습을 통해 가중치들(W)의 값이 결정될 수 있다. 이와 같이, 뉴럴 프로세싱 장치(100)는 학습된 가중치들(W)을 기반으로 추론을 수행할 수 있다.

[0094] 도 11은 도 1의 뉴럴 프로세싱 장치가 노이즈를 분류하는 동작의 하나의 예시를 나타내는 순서도이다. 도 1, 도 4 및 도 11을 참조하면, 뉴럴 프로세싱 장치(100)는 S111 단계에서, 입력 데이터(IN)를 수신할 수 있다. S112 단계에서, 뉴럴 프로세싱 장치(100)는 입력 데이터(IN)에 대한 푸리에 변환 값(FTV)을 산출할 수 있다.

[0095] S113 단계에서, 뉴럴 프로세싱 장치(100)는 푸리에 변환 값(FTV)들의 합이 기준 푸리에 변환 값들의 합(RFS)과 동일한지 여부를 판별할 수 있다. 동일한 경우, S114 단계에서, 뉴럴 프로세싱 장치(100)는 입력 데이터(IN)를 클린 데이터로 판별할 수 있다. 즉, 뉴럴 프로세싱 장치(100)는 입력 데이터(IN)가 노이즈를 포함하지 않는 것으로 판별할 수 있다.

[0096] 그렇지 않은 경우, S115 단계에서, 뉴럴 프로세싱 장치(100)는 푸리에 변환 값(FTV)들의 합이 기준 푸리에 변환 값들의 합(RFS)보다 큰지 여부를 판별할 수 있다. 큰 경우, S116 단계에서, 뉴럴 프로세싱 장치(100)는 입력 데이터(IN)의 노이즈 특성(NSC)을 가우시안 노이즈로 판별할 수 있다. 그렇지 않은 경우, S117 단계에서, 뉴럴 프로세싱 장치(100)는 입력 데이터(IN)의 노이즈 특성(NSC)을 블러 노이즈로 판별할 수 있다.

[0097] S118 단계에서, 뉴럴 프로세싱 장치(100)는 푸리에 변환 값(FTV)들의 합의 크기에 기초하여 노이즈 강도(N_d)를 판별할 수 있다. 예를 들어, 뉴럴 프로세싱 장치(100)는 푸리에 변환 값(FTV)들의 합과 기준 푸리에 변환 값들의 합(RFS)의 차이의 크기에 따라 노이즈 강도(N_d)를 판별할 수 있다. 또는, 뉴럴 프로세싱 장치(100)는 푸리에 변환 값(FTV)들의 합의 크기에 따라 노이즈 강도(N_d)를 판별할 수 있다.

[0098] 도 12는 본 발명의 실시 예에 따른 추론 결과의 예시를 보여주는 도면이다. 도 12의 가로축은 노이즈 특성(NSC)을 나타내고, 세로축은 추론의 정확도를 나타낸다. 도 12를 참조하면, 뉴럴 네트워크들(NN1~NN7)을 이용하여 도 7의 제1 내지 제7 노이즈 특성들(NSC1~NSC7)을 가지는 입력 데이터(IN)에 대한 추론 결과가 도시된다.

[0099] 제1 노이즈 특성(NSC1)을 가지는 입력 데이터(IN)에 대하여, 제1 뉴럴 네트워크(NN1)를 이용하여 추론이 수행되었을 때, 추론의 정확도가 가장 높을 수 있다. 즉, 도 7에 도시된 바와 같이, 제1 노이즈 특성(NSC1)에 대응하는 제1 뉴럴 네트워크(NN1)가 이용되는 경우, 추론의 정확도가 향상될 수 있다.

[0100] 마찬가지로, 각각의 노이즈 특성(NSC)을 가지는 입력 데이터(IN)에 대하여, 대응하는 뉴럴 네트워크를 이용하여 추론이 수행되었을 때, 추론의 정확도가 가장 높을 수 있다. 즉, 노이즈 특성(NSC)에 따라 대응하는 뉴럴 네트워크를 이용하여 추론이 수행되는 경우, 뉴럴 프로세싱 장치(100)는 원하는 정확도로서 추론을 수행할 수 있다.

[0101] 도 13은 본 발명의 실시 예에 따른 추론 결과와 종래 기술에 따른 추론 결과의 비교 예시를 보여주는 도면이다. 도 13의 가로축은 노이즈 특성(NSC)을 나타내고, 세로축은 추론의 연산량을 나타낸다. 본 발명의 실시 예에 따른 추론 결과는 노이즈 특성(NSC)에 따른 뉴럴 네트워크를 기반으로 수행된 추론 결과를 나타낸다. 종래 기술에 따른 추론 결과는 노이즈 특성(NSC)에 관계 없이 하나의 뉴럴 네트워크를 기반으로 수행된 추론 결과를 나타낸다.

[0102] 도 13에 도시된 바와 같이, 본 발명의 실시 예에 따르면, 제1 노이즈 특성(NSC1)을 가지는 입력 데이터(IN)에 대한 추론의 연산량은 다른 노이즈 특성(NSC)을 가지는 입력 데이터(IN)에 대한 추론의 연산량보다 작을 수 있다. 본 발명의 실시 예에 따르면, 노이즈 강도(N_d)가 작아질수록 추론의 연산량이 작아질 수 있다.

[0103] 반면에, 종래 기술에 따르면, 노이즈 특성(NSC)에 관계 없이 입력 데이터(IN)에 대한 추론의 연산량이 동일할 수 있다. 따라서, 동일한 입력 데이터(IN)에 대한 추론에 있어서, 종래 기술에 따른 추론의 연산량은 본 발명의 실시 예에 따른 추론의 연산량보다 클 수 있다. 예를 들어, 제1 노이즈 특성(NSC1)을 가지는 입력 데이터(IN)에 대한 추론에 있어서, 종래 기술에 따른 추론의 연산량은 본 발명의 실시 예에 따른 추론의 연산량보다 훨씬 클 수 있다.

[0104] 상술한 바와 같이, 본 발명의 실시 예에 따르면, 뉴럴 프로세싱 장치(100)는 최소화된 연산량 및 원하는 정확도로 추론을 수행할 수 있다. 따라서, 뉴럴 프로세싱 장치(100)는 저 전력으로 동작할 수 있다.

[0105] 도 14는 본 발명의 실시 예들에 따른 뉴럴 프로세싱 장치를 포함하는 컴퓨팅 시스템을 나타내는 블록도이다. 도 14를 참조하면, 컴퓨팅 시스템(1000)은 메인 프로세서(1100), 메모리(1200), 사용자 인터페이스(1300), 출력 장치(1400), 버스(1500) 및 뉴럴 프로세싱 장치(100)를 포함할 수 있다. 컴퓨팅 시스템(1000)은 모바일 장치, 데스크탑, 임베디드 시스템 및 서버로 구현될 수 있다.

[0106] 메인 프로세서(1100)는 컴퓨팅 시스템(1000)을 제어할 수 있다. 메인 프로세서(1100)는 메모리(1200), 뉴럴 프로세싱 장치(100) 및 출력 장치(1400)의 동작을 제어할 수 있다. 예를 들어, 메인 프로세서(1100)는 메모리(1200)에 저장된 데이터를 처리할 수 있다. 메인 프로세서(1100)는 처리된 데이터를 메모리(1200)에 저장할 수 있다. 메인 프로세서(1100)는 버스(1500)를 통해 컴퓨팅 시스템(1000)의 구성 요소들로 데이터 및 명령들을 전송할 수 있다.

[0107] 메모리(1200)는 데이터를 저장하거나 저장된 데이터를 출력할 수 있다. 메모리(1200)는 메인 프로세서(1100) 및 뉴럴 프로세싱 장치(100)에서 처리된 데이터 또는 처리될 데이터를 저장할 수 있다. 예를 들어, 메모리(1200)는 휘발성 메모리 또는 비휘발성 메모리로 구현될 수 있다.

[0108] 사용자 인터페이스(1300)는 사용자로부터 명령 및 데이터를 입력 받을 수 있다. 사용자 인터페이스(1300)는 다양한 형태의 입력을 수신할 수 있는 인터페이스로 구현될 수 있다. 예를 들어, 사용자 인터페이스(1300)는 문자

방식의 입력 사용자 인터페이스(CUI; Character based User Interface), 그래픽 기반의 입력 사용자 인터페이스(GUI; Graphic User Interface), 말과 행동 기반의 입력 인터페이스(NUI; Natural User Interface) 등으로 구현될 수 있다.

[0109] 출력 장치(1400)는 메인 프로세서(1100) 또는 뉴럴 프로세싱 장치(100)에서 처리된 데이터를 출력할 수 있다. 출력 장치(1400)는 다양한 형태의 출력을 사용자에게 제공할 수 있다. 예를 들어, 출력 장치(1400)는 디스플레이, 스피커 등으로 구현될 수 있다.

[0110] 버스(1500)는 컴퓨팅 시스템(1000)의 구성 요소들 사이에 명령들 및 데이터를 전달할 수 있다.

[0111] 뉴럴 프로세싱 장치(100)는 메인 프로세서(1100)의 제어에 따라 뉴럴 네트워크를 기반으로 추론 동작을 수행할 수 있다. 뉴럴 프로세싱 장치(100)는 사용자 인터페이스(1300)로부터 수신된 데이터 또는 메모리(1200)에 저장된 데이터에 대하여 추론을 수행할 수 있다. 뉴럴 프로세싱 장치(100)는 추론 결과를 메인 프로세서(1100)로 전달할 수 있다. 메인 프로세서(1100)는 추론 결과를 출력 장치(1400)를 통해 사용자에게 출력할 수 있다. 또는, 메인 프로세서(1100)는 추론 결과를 기반으로 컴퓨팅 시스템(1000)을 제어할 수 있다.

[0112] 도 1 내지 도 13에서 설명된 바와 같이, 뉴럴 프로세싱 장치(100)는 버스(1500)를 통해 전달된 입력 데이터(IN)의 노이즈 특성(NSC)을 판별할 수 있다. 뉴럴 프로세싱 장치(100)는 판별된 노이즈 특성(NSC)에 대응하는 뉴럴 네트워크를 기반으로 추론을 수행할 수 있다.

[0113] 상술된 내용은 본 발명을 실시하기 위한 구체적인 실시 예들이다. 본 발명은 상술된 실시 예들뿐만 아니라, 단순하게 설계 변경되거나 용이하게 변경할 수 있는 실시 예들 또한 포함할 것이다. 또한, 본 발명은 실시 예들을 이용하여 용이하게 변형하여 실시할 수 있는 기술들도 포함될 것이다. 따라서, 본 발명의 범위는 상술된 실시 예들에 국한되어 정해져서는 안되며 후술하는 특허청구범위뿐만 아니라 이 발명의 특허청구범위와 균등한 것들에 의해 정해져야 할 것이다.

부호의 설명

[0114] 100: 뉴럴 프로세싱 장치

110: 노이즈 분류기

111: 푸리에 변환기

112: 푸리에 변환 값 비교기

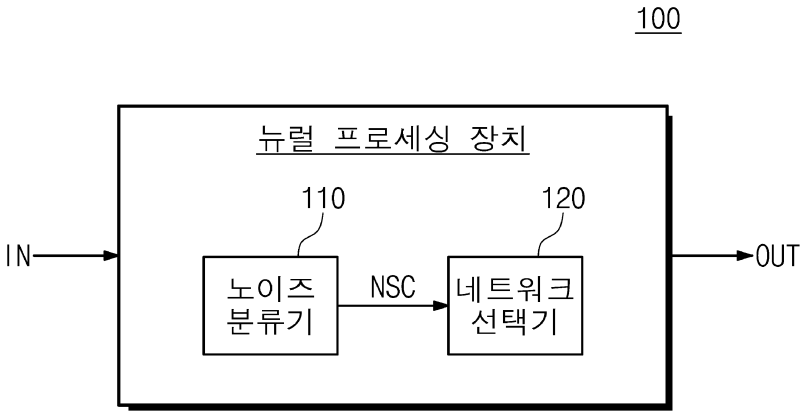
120: 네트워크 선택기

130: 메모리

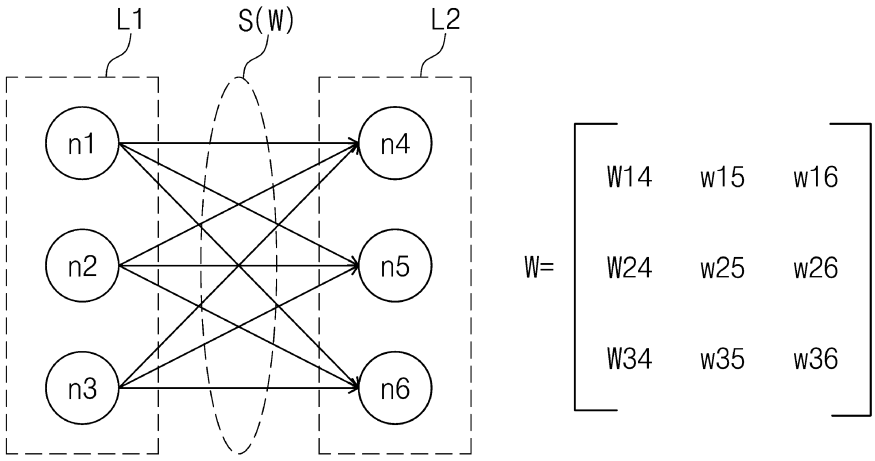
140: 연산기

도면

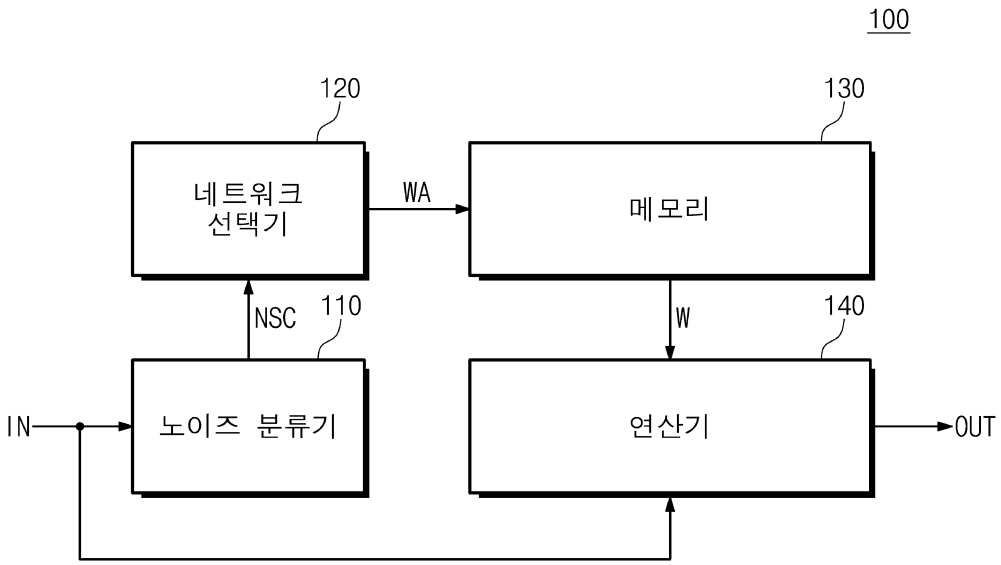
도면1


100
뉴럴 프로세싱 장치
110
120
IN
노이즈 분류기
NSC
네트워크 선택기
OUT

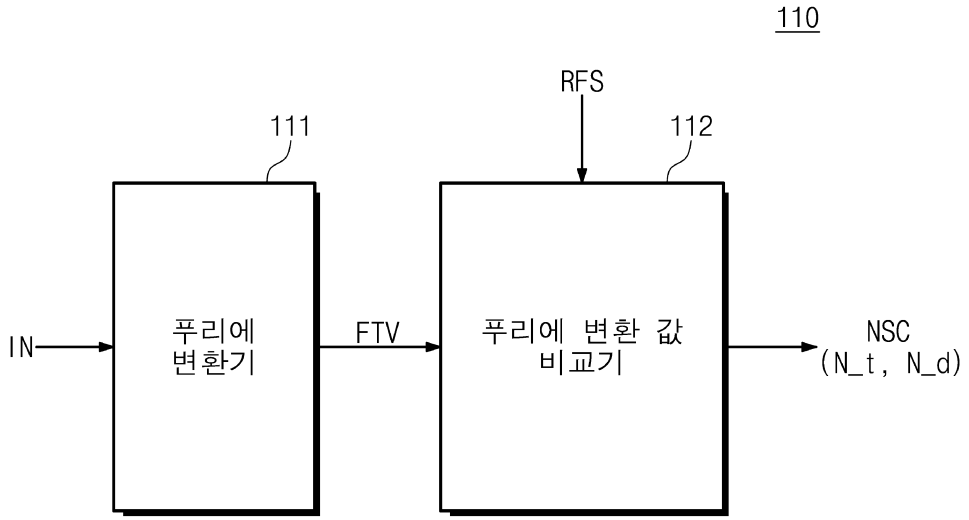
도면2


L1
S(W)
L2
n1
n2
n3
n4
n5
n6
W=
W14 w15 w16
W24 w25 w26
W34 w35 w36

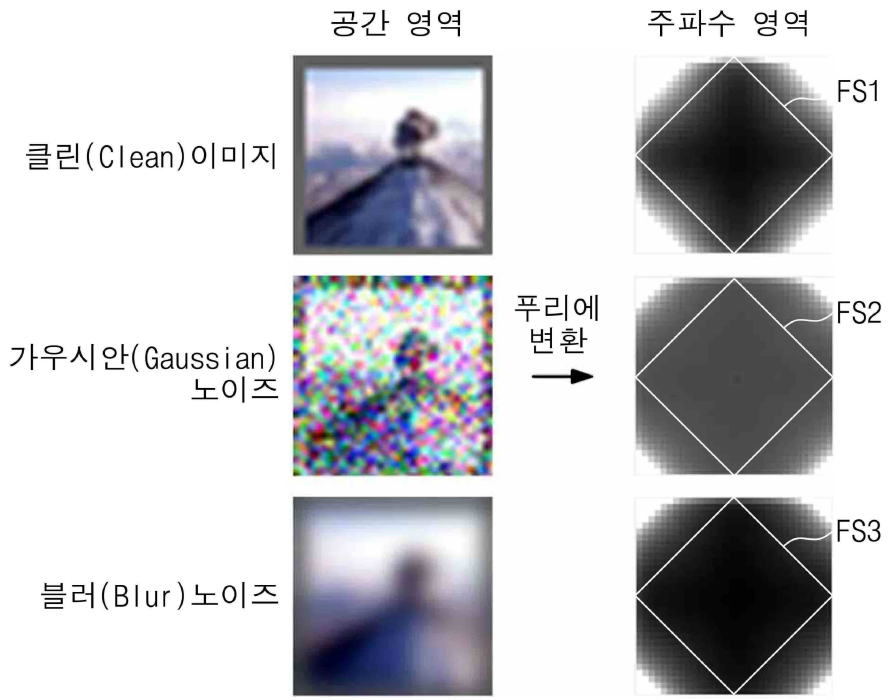
도면3


100
120
130
네트워크
선택기
WA
메모리
NSC
110
W
140
IN
노이즈 분류기
연산기
OUT

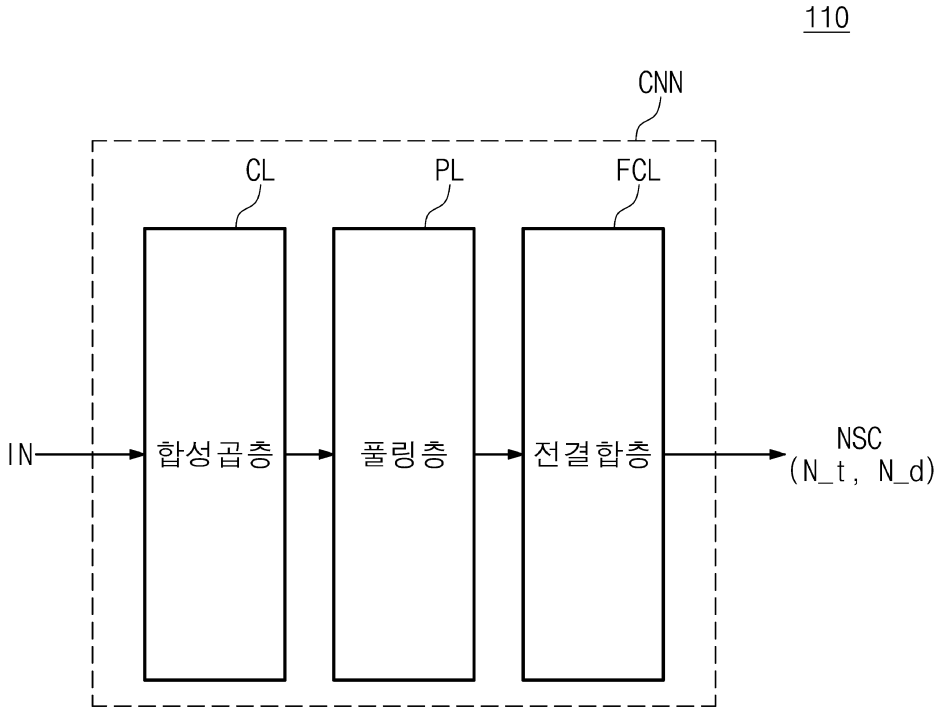
도면4


110
RFS
111
112
IN
푸리에
변환기
FTV
푸리에 변환 값
비교기
NSC
(N_t, N_d)

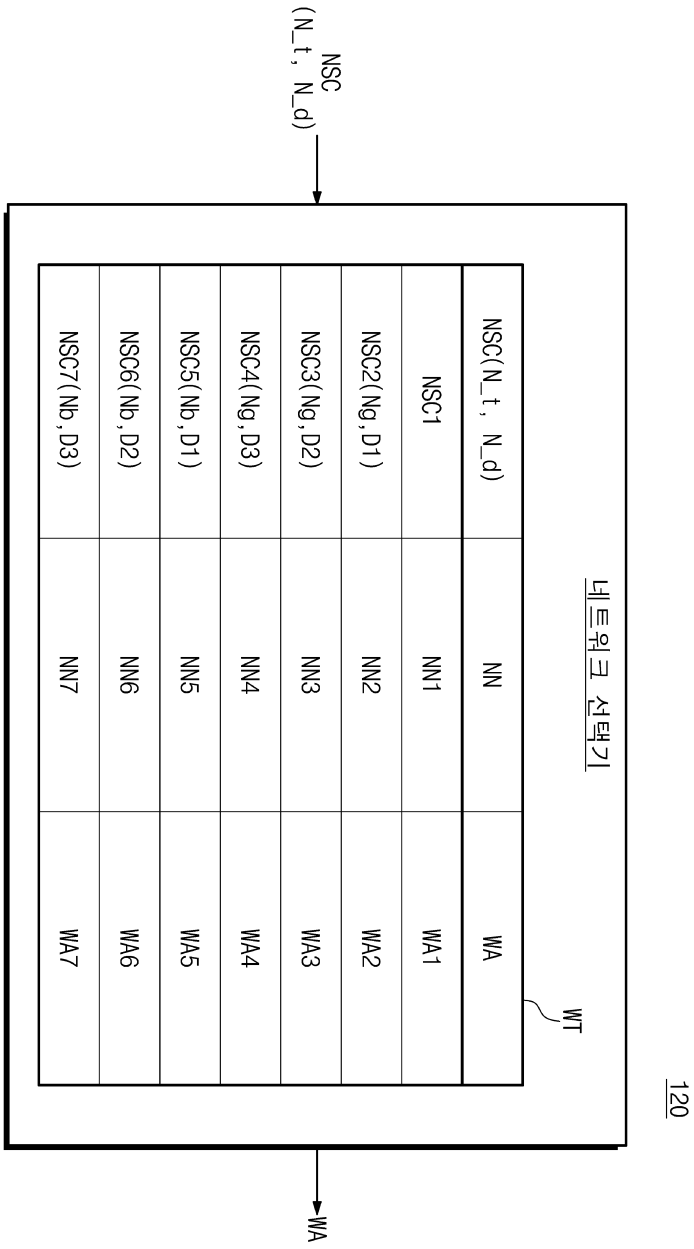
도면5


공간 영역
주파수 영역
클린(Clean)이미지
가우시안(Gaussian)
노이즈
블러(Blur)노이즈
푸리에
변환
FS1
FS2
FS3

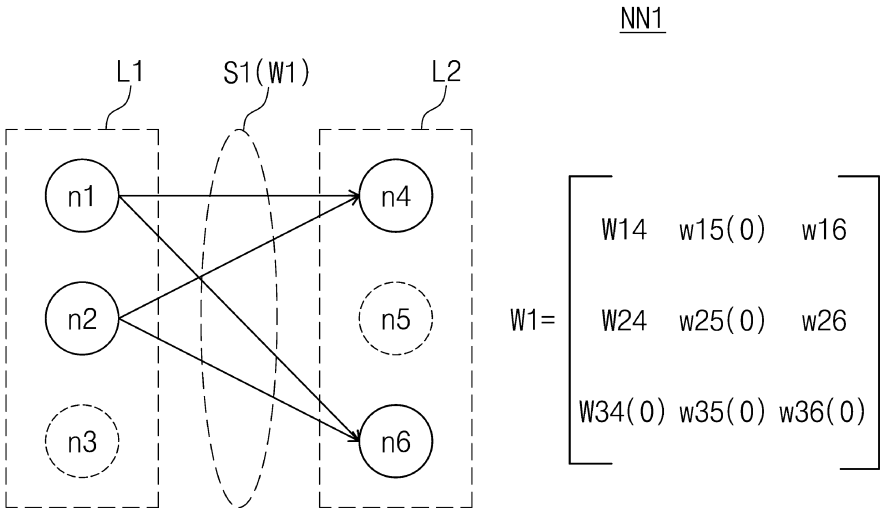
도면6


110
CNN
CL
PL
FCL
IN
합성곱층
풀링층
전결합층
NSC
(N_t, N_d)

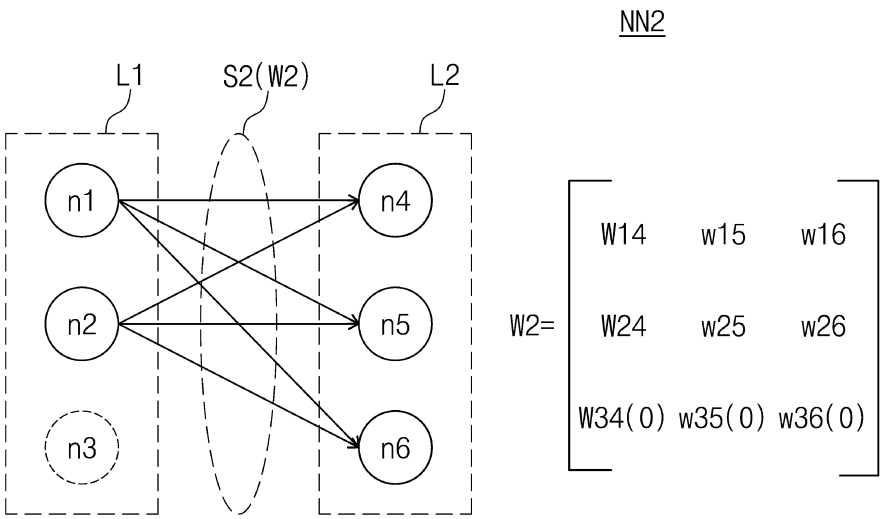
도면7


NSC
(N_t, N_d)
네트워크 선택기
120
WT
NSC(N_t, N_d)
NSC1
NSC2(Ng,D1)
NSC3(Ng,D2)
NSC4(Ng,D3)
NSC5(Nb,D1)
NSC6(Nb,D2)
NSC7(Nb,D3)
NN
NN1
NN2
NN3
NN4
NN5
NN6
NN7
WA
WA1
WA2
WA3
WA4
WA5
WA6
WA7
WA

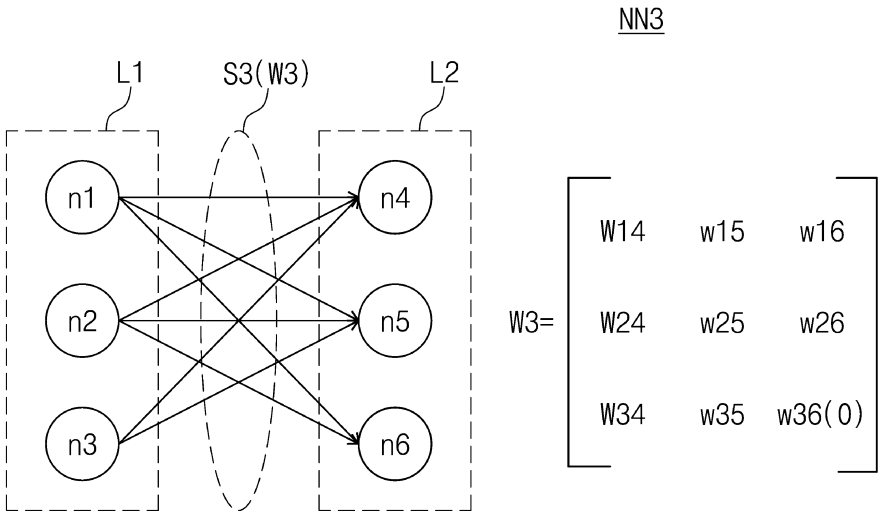
도면8a


NN1
L1
S1(W1)
L2
n1
n2
n3
n4
n5
n6
W1=
W14 w15(0) w16
W24 w25(0) w26
W34(0) w35(0) w36(0)

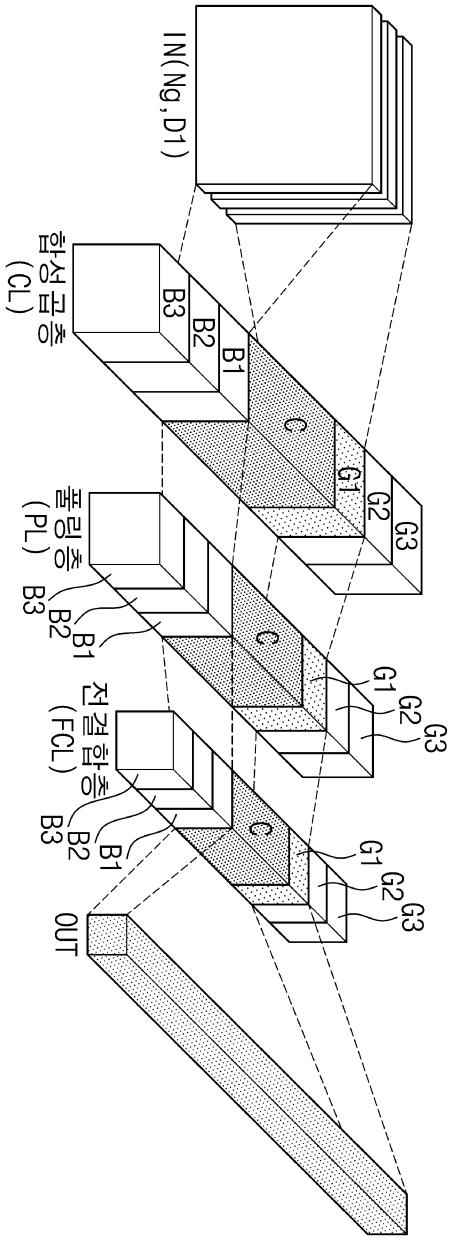
도면8b


NN2
L1
S2(W2)
L2
n1
n2
n3
n4
n5
n6
W2=
W14 w15 w16
W24 w25 w26
W34(0) w35(0) w36(0)

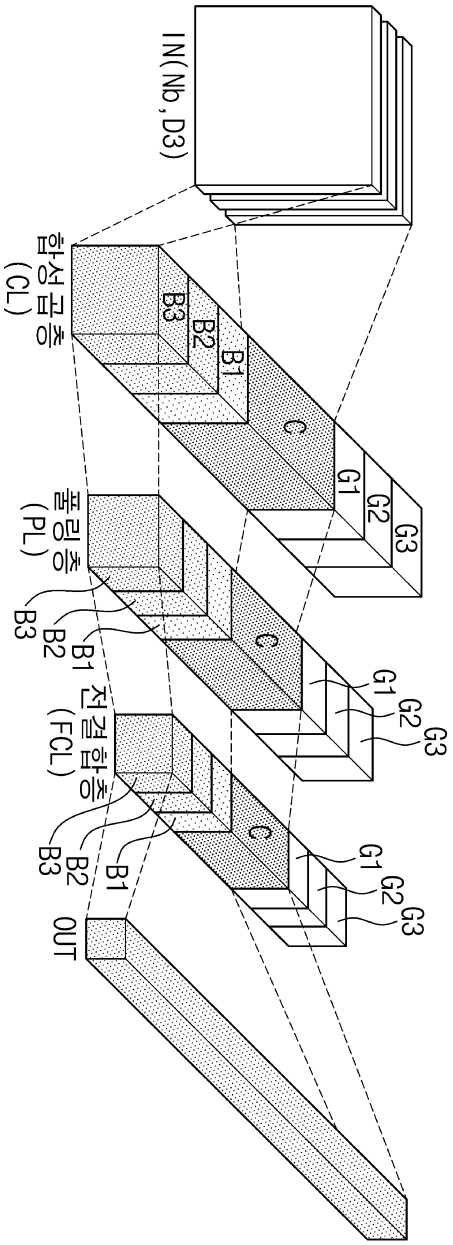
도면8c


NN3
L1
S3(W3)
L2
n1
n2
n3
n4
n5
n6
W3=
W14 w15 w16
W24 w25 w26
W34 w35 w36(0)

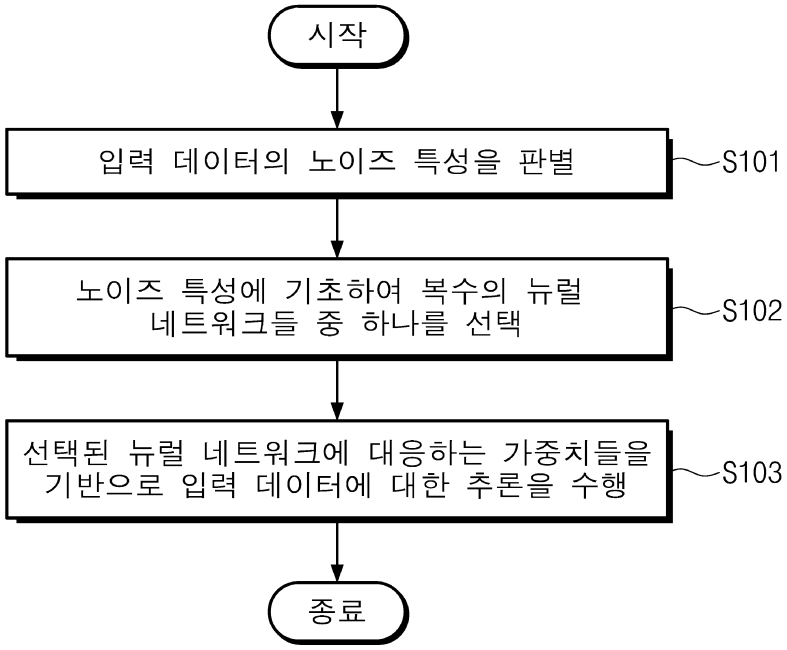
도면9a


IN(Ng,D1)
합성곱층
(CL)
B3
B2
B1
C
G1
G2
G3
풀링층
(PL)
B3
B2
B1
C
G1
G2
G3
전결합층
(FCL)
B3
B2
B1
C
G1
G2
G3
OUT

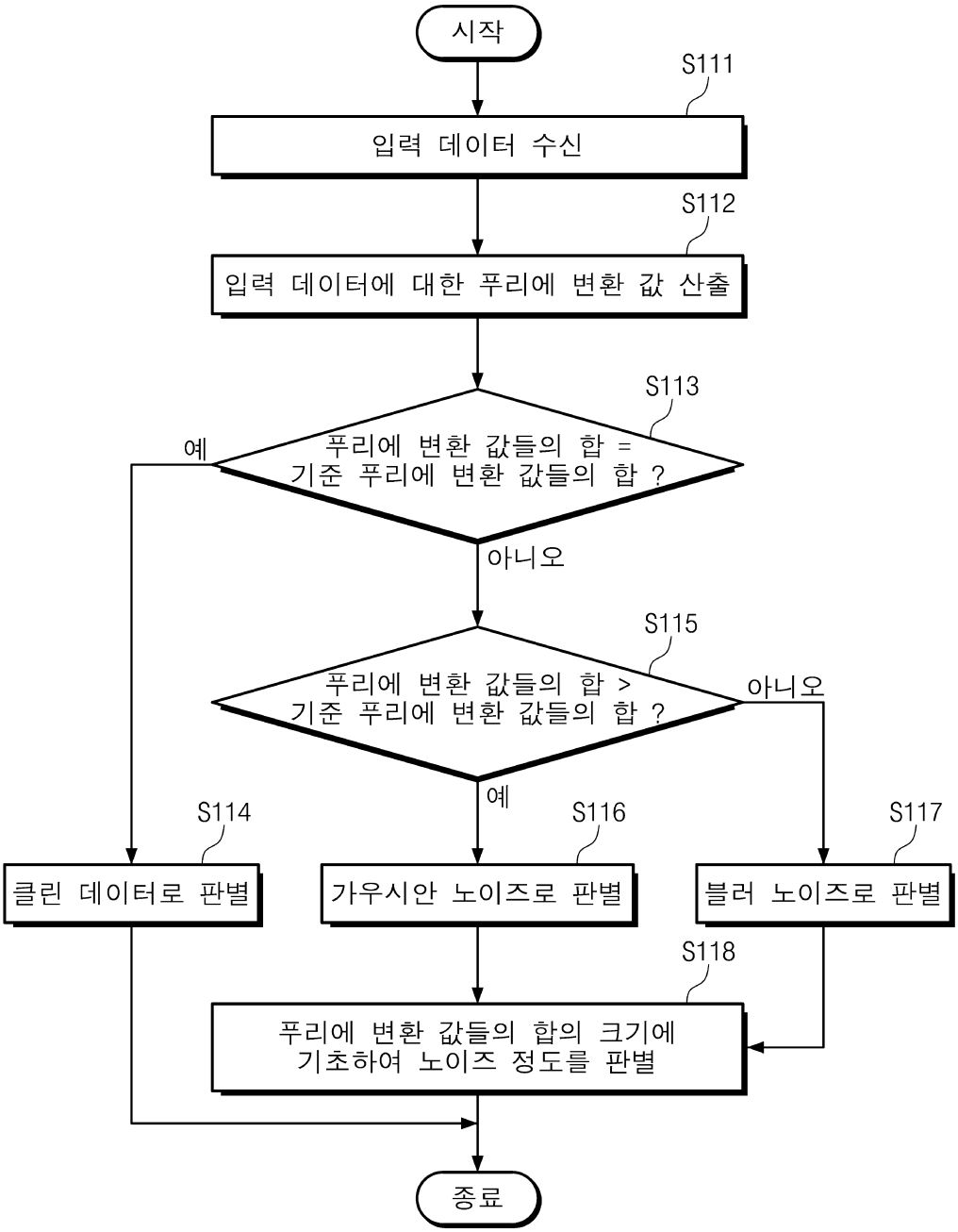
도면9b


IN(Nb,D3)
합성곱층
(CL)
B1
B2
B3
C
G1
G2
G3
풀링층
(PL)
B1
B2
B3
C
G1
G2
G3
전결합층
(FCL)
B1
B2
B3
C
G1
G2
G3
OUT

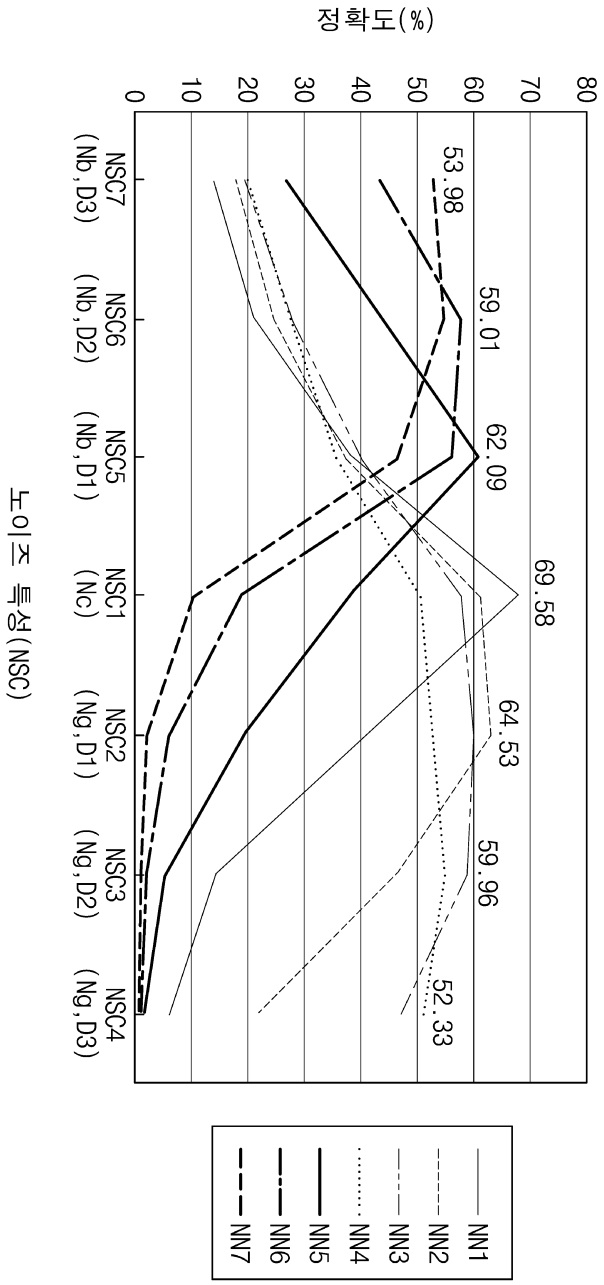
도면10


시작
입력 데이터의 노이즈 특성을 판별
S101
노이즈 특성에 기초하여 복수의 뉴럴 네트워크들 중 하나를 선택
S102
선택된 뉴럴 네트워크에 대응하는 가중치들을 기반으로 입력 데이터에 대한 추론을 수행
S103
종료

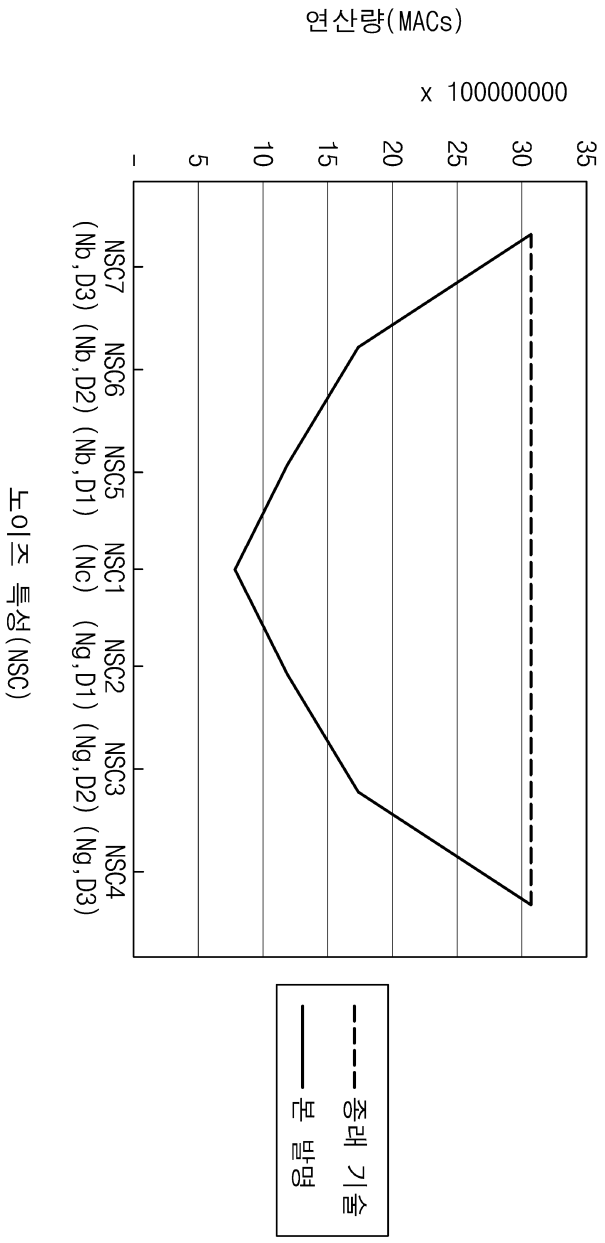
도면11


시작
S111
입력 데이터 수신
S112
입력 데이터에 대한 푸리에 변환 값 산출
S113
푸리에 변환 값들의 합 =
기준 푸리에 변환 값들의 합 ?
예
아니오
S115
푸리에 변환 값들의 합 >
기준 푸리에 변환 값들의 합 ?
아니오
예
S114
클린 데이터로 판별
S116
가우시안 노이즈로 판별
S117
블러 노이즈로 판별
S118
푸리에 변환 값들의 합의 크기에
기초하여 노이즈 정도를 판별
종료

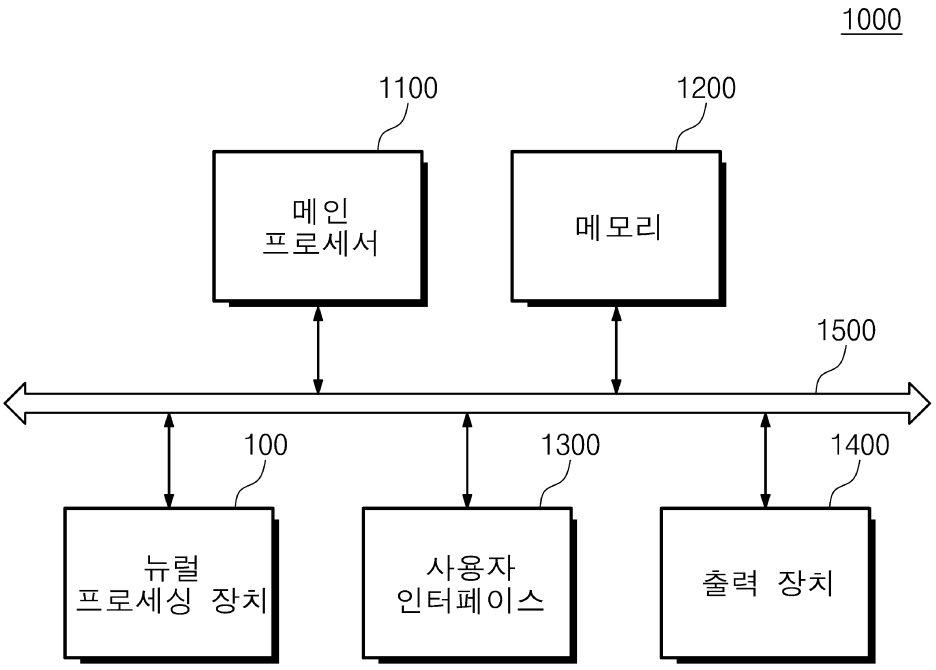
도면12


정확도(%)
0
10
20
30
40
50
60
70
80
53.98
59.01
62.09
69.58
64.53
59.96
52.33
NSC7 (Nb,D3)
NSC6 (Nb,D2)
NSC5 (Nb,D1)
NSC1 (Nc)
NSC2 (Ng,D1)
NSC3 (Ng,D2)
NSC4 (Ng,D3)
노이즈 특성(NSC)
NN1
NN2
NN3
NN4
NN5
NN6
NN7

도면13


연산량(MACs)
x 100000000
35
30
25
20
15
10
5
-
종래 기술
본 발명
NSC7 (Nb,D3)
NSC6 (Nb,D2)
NSC5 (Nb,D1)
NSC1 (Nc)
NSC2 (Ng,D1)
NSC3 (Ng,D2)
NSC4 (Ng,D3)
노이즈 특성(NSC)

도면14


1000
1100
메인
프로세서
1200
메모리
1500
100
뉴럴
프로세싱 장치
1300
사용자
인터페이스
1400
출력 장치