



(19) 대한민국특허청(KR)

(12) 등록특허공보(B1)

(45) 공고일자 2023년10월31일

(11) 등록번호 10-2597079

(24) 등록일자 2023년10월27일

(51) 국제특허분류(Int. Cl.)

G06N 3/063 (2023.01) G06F 17/15 (2006.01)

G06N 3/04 (2023.01)

(52) CPC특허분류

G06N 3/063 (2013.01)

G06F 17/153 (2013.01)

(21) 출원번호 10-2021-0117675

(22) 출원일자 2021년09월03일

심사청구일자 2021년09월03일

(65) 공개번호 10-2023-0034655

(43) 공개일자 2023년03월10일

(56) 선행기술조사문헌

KR1020200110165 A

(뒷면에 계속)

전체 청구항 수 : 총 8 항

심사관 : 이준상

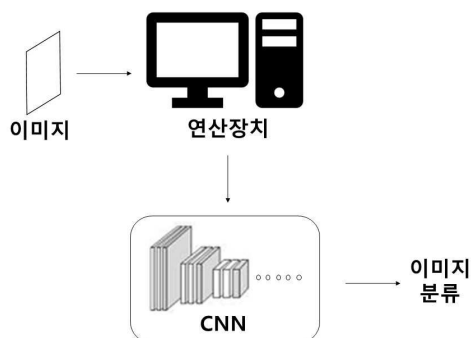
(54) 발명의 명칭 n차원 텐서 곱 연산을 이용한 합성곱 신경망 압축 방법 및 장치

## (57) 요약

개시된 기술은 n차원 텐서 곱 연산을 이용한 합성곱 신경망 압축 방법 및 장치에 관한 것으로, 연산장치가 합성곱 신경망(Convolution Neural Network, CNN)에 포함된 복수개의 계층들 중 특정 계층을 선택하는 단계; 상기 연산장치가 상기 특정 계층에 포함된 복수의 가중치 필터들 각각에 대한 n차원 텐서 곱 연산을 수행하여 상기 복수의 가중치 필터들 각각에 대한 중요도를 계산하는 단계; 및 상기 연산장치가 상기 계산 결과에 따라 상기 복수의 가중치 필터들 중 상기 중요도가 기준 미만인 일부의 가중치 필터를 제거하는 단계;를 포함한다.

대표도 - 도1

100



(52) CPC특허분류  
G06N 3/04 (2023.01)

(56) 선행기술조사문헌  
JP2021523468 A  
JP2021522565 A  
JP2021082289 A  
JP2021006980 A  
KR1020210046416 A  
KR1020210097382 A  
W02020105144 A1\*  
\*는 심사관에 의하여 인용된 문헌

이 발명을 지원한 국가연구개발사업

|             |                                       |
|-------------|---------------------------------------|
| 과제고유번호      | 1711126202                            |
| 과제번호        | 2021-0-00022-001                      |
| 부처명         | 과학기술정보통신부                             |
| 과제관리(전문)기관명 | 정보통신기획평가원                             |
| 연구사업명       | 정보통신방송연구개발사업                          |
| 연구과제명       | Edge 컴퓨팅 환경을 위한 AI 모델 최적화 및 경량화 기술 개발 |
| 기 여 율       | 1/1                                   |
| 과제수행기관명     | 한국전자기술연구원                             |
| 연구기간        | 2021.01.01 ~ 2021.12.31               |

공지예외적용 : 있음

---

## 명세서

### 청구범위

#### 청구항 1

연산장치가 합성곱 신경망(Convolution Neural Network, CNN)에 포함된 복수개의 계층들 중 특정 계층을 선택하는 단계;

상기 연산장치가 상기 특정 계층에 포함된 복수의 가중치 필터들 각각에 대한 n차원 텐서 곱 연산을 수행하여 상기 복수의 가중치 필터들 각각에 대한 중요도를 계산하는 단계; 및

상기 연산장치가 상기 계산 결과에 따라 상기 복수의 가중치 필터들 중 상기 중요도가 기준 미만인 일부의 가중치 필터를 제거하는 단계;를 포함하되,

상기 연산장치는 상기 n차원 텐서 곱 연산에 따라 계산된 상기 복수의 가중치 필터들의 텐서 데이터와 가중치 필터를 0으로 치환한 텐서 데이터를 뺀 결과가 기준 미만이면 해당 가중치 필터를 제거하는 합성곱 신경망 압축 방법.

#### 청구항 2

제 1 항에 있어서,

상기 특정 계층을 선택하는 단계는, 상기 합성곱 신경망에 포함된 복수개의 계층들 중 하나를 랜덤하게 선택하는 합성곱 신경망 압축 방법.

#### 청구항 3

삭제

#### 청구항 4

제 1 항에 있어서,

상기 중요도는 이하의 수학적 식 5에 따라 계산하는 합성곱 신경망 압축 방법.

[수학적 식 5]

$$I_{proposed}(\mathcal{F}_{i,j}) = \frac{\|\mathcal{F}_i \times_4^3 \mathcal{F}_{i+1} - \text{zero}(\mathcal{F}_i, j) \times_4^3 \mathcal{F}_{i+1}\|_1}{\|\mathcal{F}_i \times_4^3 \mathcal{F}_{i+1}\|_1},$$

(여기에서  $(\mathcal{F}_{i,j})$  는 i번째 계층의 j번째 가중치 필터의 중요도를 의미하고,  $\mathcal{F}_i \times_4^3 \mathcal{F}_{i+1}$  는 i번째 계층  $\mathcal{F}_i \in \mathbb{R}^{D_i \times D \times T_i \times T_{i+1}}$  과 i+1번째 계층  $\mathcal{F}_{i+1} \in \mathbb{R}^{D_i \times D \times T_{i+1} \times T_{i+2}}$  간의 n차원 곱셈을 의미함. 그리고  $D_i$  는 i번째 계층에 포함된 가중치 필터의 너비 및 높이를 의미하고  $T_i$  는 i번째 계층의 가중치 필터의 개수를 의미함. 그리고  $\text{zero}(\mathcal{F}_i, j)$  는 i번째 계층의 j번째 가중치 필터를 0으로 구성된 가중치 필터로 대체한 것을 의미함.)

#### 청구항 5

제 4 항에 있어서,

상기 n차원 곱셈은 이하의 수학적 식 6에 따라 계산되는 합성곱 신경망 압축 방법.

[수학식 6]

$$\begin{aligned}\mathcal{F}_i^{(4)} &= g(\mathcal{F}_i, T_{i+1} \times D_i D_i T_i) \in \mathbb{R}^{T_{i+1} \times D_i D_i T_i}, \\ \mathcal{F}_{i+1}^{(3)} &= g(\mathcal{F}_{i+1}, T_{i+1} \times D_{i+1} D_{i+1} T_{i+2}) \in \mathbb{R}^{T_{i+1} \times D_{i+1} D_{i+1} T_{i+2}}, \\ \mathcal{F}_i \times_4^3 \mathcal{F}_{i+1} &= \mathcal{F}_i^{(4)T} \mathcal{F}_{i+1}^{(3)} \in \mathbb{R}^{D_i D_i T_i \times D_{i+1} D_{i+1} T_{i+2}},\end{aligned}$$

(여기에서  $g(\mathcal{F}_i, T_{i+1} \times D_i D_i T_i)$  는  $i$  번째 계층  $\mathcal{F}_i \in \mathbb{R}^{D_i \times D_i \times T_i \times T_{i+1}}$  을  $T_{i+1} \times D_i D_i T_i$  의 형태로 변경하는 차원 변환 함수를 의미함.)

#### 청구항 6

복수개의 계층을 포함하는 합성곱 신경망(Convolution Neural Network, CNN)을 저장하는 저장장치; 및

상기 합성곱 신경망에 포함된 복수개의 계층들 중 특정 계층을 선택하여 상기 특정 계층에 포함된 가중치 필터에 대한  $n$ 차원 텐서 곱 연산을 수행하여 상기 가중치 필터에 대한 중요도를 계산하고, 상기 계산 결과에 따라 상기 가중치 필터를 제거하는 연산장치;를 포함하되,

상기 연산장치는 상기  $n$ 차원 텐서 곱 연산에 따라 계산된 상기 가중치 필터들의 텐서 데이터와 가중치 필터를 0으로 치환한 텐서 데이터를 뺀 결과가 기준 미만이면 상기 가중치 필터를 제거하는 합성곱 신경망 압축 장치.

#### 청구항 7

삭제

#### 청구항 8

제 6 항에 있어서,

상기 연산장치는 이하의 수학식 5에 따라 상기 중요도를 계산하는 합성곱 신경망 압축 장치.

[수학식 5]

$$I_{proposed}(\mathcal{F}_{i,j}) = \frac{\|\mathcal{F}_i \times_4^3 \mathcal{F}_{i+1} - \text{zero}(\mathcal{F}_{i,j}) \times_4^3 \mathcal{F}_{i+1}\|_1}{\|\mathcal{F}_i \times_4^3 \mathcal{F}_{i+1}\|_1},$$

(여기에서  $(\mathcal{F}_{i,j})$  는  $i$  번째 계층의  $j$  번째 가중치 필터의 중요도를 의미하고,  $\mathcal{F}_i \times_4^3 \mathcal{F}_{i+1}$  는  $i$  번째 계층  $\mathcal{F}_i \in \mathbb{R}^{D_i \times D_i \times T_i \times T_{i+1}}$  과  $i+1$  번째 계층  $\mathcal{F}_{i+1} \in \mathbb{R}^{D_{i+1} \times D_{i+1} \times T_{i+1} \times T_{i+2}}$  간의  $n$ 차원 곱셈을 의미함. 그리고  $D_i$  는  $i$  번째 계층에 포함된 가중치 필터의 너비 및 높이를 의미하고  $T_i$  는  $i$  번째 계층의 가중치 필터의 개수를 의미함. 그리고  $\text{zero}(\mathcal{F}_{i,j})$  는  $i$  번째 계층의  $j$  번째 가중치 필터를 0으로 구성된 가중치 필터로 대체한 것을 의미함.)

#### 청구항 9

제 8 항에 있어서,

상기 연산장치는 이하의 수학식 6에 따라 상기  $n$ 차원 곱셈을 계산하는 합성곱 신경망 압축 장치.

[수학식 6]

$$\begin{aligned}\mathcal{F}_i^{(4)} &= g(\mathcal{F}_i, T_{i+1} \times D_i D_i T_i) \in \mathbb{R}^{T_{i+1} \times D_i D_i T_i}, \\ \mathcal{F}_{i+1}^{(3)} &= g(\mathcal{F}_{i+1}, T_{i+1} \times D_{i+1} D_{i+1} T_{i+2}) \in \mathbb{R}^{T_{i+1} \times D_{i+1} D_{i+1} T_{i+2}}, \\ \mathcal{F}_i \times_4^3 \mathcal{F}_{i+1} &= \mathcal{F}_i^{(4)T} \mathcal{F}_{i+1}^{(3)} \in \mathbb{R}^{D_i D_i T_i \times D_{i+1} D_{i+1} T_{i+2}},\end{aligned}$$

(여기에서  $g(\mathcal{F}_i, T_{i+1} \times D_i D_i T_i)$  는  $i$  번째 계층  $\mathcal{F}_i \in \mathbb{R}^{D_i \times D_i \times T_i \times T_i + 1}$  을  $T_{i+1} \times D_i D_i T_i$  의 형태로 변경하는 차원 변환 함수를 의미함.)

## 청구항 10

제 6 항에 있어서,

상기 연산장치는 상기 특정 계층에 포함된 나머지 가중치 필터들에 대한 중요도를 각각 계산하여 중요도가 기준 미만인 적어도 하나의 가중치 필터를 제거하는 합성곱 신경망 압축 장치.

## 발명의 설명

### 기술 분야

[0001] 개시된 기술은 합성곱 신경망 내 특정 계층의 가중치 필터가 다음 계층에 미치는 영향을  $n$ 차원 텐서 곱 연산을 통해 계산하여 합성곱 신경망을 압축하는 방법 및 장치에 관한 것이다.

### 배경 기술

[0002] 최근 다양한 기계 학습 및 컴퓨터 비전 기술들이 심층 합성곱 신경망을 활용하여 매우 높은 성능을 달성하였다. 하지만, 합성곱 신경망의 구조적 특징으로 인해 다수의 기 개발된 심층 합성곱 신경망 기반의 기술들은 높은 수준의 연산 능력을 요구하는 필수 불가결한 문제를 보유하고 있다. 심층 합성곱 신경망의 높은 메모리 및 계산 복잡도 문제를 해결하기 위해, 합성곱 층을 구성하는 가중치 필터들 중에서 신경망의 성능에 영향을 미치지 않거나 미치는 영향이 미미한 잉여 가중치 필터들을 선별하고 제거하는 필터 선택법 기반의 신경망 압축 기법들이 제안되고 있다.

[0003] 기존 필터 선택법들은 특정 합성곱 층을 구성하는 가중치 필터들의 norm(노름)을 측정하여 낮은 노름을 가진 가중치 필터들을 제거하고, 반대로 높은 노름을 가지는 가중치 필터들을 유지하는 방법을 제안하였다. 이러한 신경망 압축 기법은 각각의 합성곱 층을 독립적으로 처리하므로, 가중치 필터의 중요도를 평가하는 과정에서 해당 필터가 후반 합성곱 층에서 미치는 영향을 고려하지 않는다. 다시 말해, 기존의 필터 선택법은 신경망 압축의 관점에서 국소 최적 문제(Local Optimal Problem)가 발생하는 문제가 있었다.

## 선행기술문헌

### 특허문헌

[0004] (특허문헌 0001) 한국 공개특허 제10-2020-0052200호

## 발명의 내용

### 해결하려는 과제

[0005] 개시된 기술은 합성곱 신경망 내 특정 계층의 가중치 필터가 다음 계층에 미치는 영향을  $n$ 차원 텐서 곱 연산을 통해 계산하여 합성곱 신경망을 압축하는 방법 및 장치를 제공하는데 있다.

### 과제의 해결 수단

[0006] 상기의 기술적 과제를 이루기 위하여 개시된 기술의 제 1 측면은 연산장치가 합성곱 신경망(Convolution Neural Network, CNN)에 포함된 복수개의 계층들 중 특정 계층을 선택하는 단계, 상기 연산장치가 상기 특정 계층에 포함된 복수의 가중치 필터들 각각에 대한  $n$ 차원 텐서 곱 연산을 수행하여 상기 복수의 가중치 필터들 각각에 대한 중요도를 계산하는 단계 및 상기 연산장치가 상기 계산 결과에 따라 상기 복수의 가중치 필터들 중 상기 중요도가 기준 미만인 일부의 가중치 필터를 제거하는 단계를 포함하는 합성곱 신경망 압축 방법을 제공하는데 있다.

[0007] 상기의 기술적 과제를 이루기 위하여 개시된 기술의 제 2 측면은 복수개의 계층을 포함하는 합성곱 신경망(Convolution Neural Network, CNN)을 저장하는 저장장치 및 상기 합성곱 신경망에 포함된 복수개의 계층들 중

특정 계층을 선택하여 상기 특정 계층에 포함된 가중치 필터에 대한  $n$ 차원 텐서 곱 연산을 수행하여 상기 가중치 필터에 대한 중요도를 계산하고, 상기 계산 결과에 따라 상기 가중치 필터를 제거하는 연산장치를 포함하는 합성곱 신경망 압축 장치를 제공하는데 있다.

### 발명의 효과

- [0008] 개시된 기술의 실시 예들은 다음의 장점들을 포함하는 효과를 가질 수 있다. 다만, 개시된 기술의 실시 예들이 이를 전부 포함하여야 한다는 의미는 아니므로, 개시된 기술의 권리범위는 이에 의하여 제한되는 것으로 이해되어서는 아니 될 것이다.
- [0009] 개시된 기술의 일 실시예에 따른  $n$ 차원 텐서 곱 연산을 이용한 합성곱 신경망 압축 방법 및 장치는 특정 가중치 필터에 대한  $n$ 차원 텐서 곱 연산을 수행하여 합성곱 신경망을 압축시키는 효과가 있다.
- [0010] 또한, 합성곱 신경망을 효율적으로 압축시켜서 메모리 소모량 및 계산 복잡도를 줄이는 효과가 있다.
- [0011] 또한, 특정 가중치 필터가 다음 계층에 미치는 영향을 계산하여 신경망의 국소 최적 문제가 발생하는 것을 방지하는 효과가 있다.

### 도면의 간단한 설명

- [0012] 도 1은 개시된 기술의 일 실시예에 따른  $n$ 차원 텐서 곱 연산을 이용하여 합성곱 신경망을 압축하는 과정을 나타낸 도면이다.
- 도 2는 개시된 기술의 일 실시예에 따른  $n$ 차원 텐서 곱 연산을 이용한 합성곱 신경망 압축 방법에 대한 순서도이다.
- 도 3은 개시된 기술의 일 실시예에 따른  $n$ 차원 텐서 곱 연산을 이용한 합성곱 신경망 압축 장치에 대한 블록도이다.
- 도 4는 종래 합성곱 신경망의 구조를 나타낸 도면이다.

### 발명을 실시하기 위한 구체적인 내용

- [0013] 본 발명은 다양한 변경을 가할 수 있고 여러 가지 실시예를 가질 수 있는 바, 특정 실시예들을 도면에 예시하고 상세한 설명에 상세하게 설명하고자 한다. 그러나, 이는 본 발명을 특정한 실시 형태에 대해 한정하려는 것이 아니며, 본 발명의 사상 및 기술 범위에 포함되는 모든 변경, 균등물 내지 대체물을 포함하는 것으로 이해되어야 한다.
- [0014] 제 1, 제 2, A, B 등의 용어는 다양한 구성요소들을 설명하는데 사용될 수 있지만, 해당 구성요소들은 상기 용어들에 의해 한정되지는 않으며, 단지 하나의 구성요소를 다른 구성요소로부터 구별하는 목적으로만 사용된다. 예를 들어, 본 발명의 권리 범위를 벗어나지 않으면서 제 1 구성요소는 제 2 구성요소로 명명될 수 있고, 유사하게 제 2 구성요소도 제 1 구성요소로 명명될 수 있다. 및/또는 이라는 용어는 복수의 관련된 기재된 항목들의 조합 또는 복수의 관련된 기재된 항목들 중의 어느 항목을 포함한다.
- [0015] 본 명세서에서 사용되는 용어에서 단수의 표현은 문맥상 명백하게 다르게 해석되지 않는 한 복수의 표현을 포함하는 것으로 이해되어야 한다. 그리고 "포함한다" 등의 용어는 실시된 특징, 개수, 단계, 동작, 구성요소, 부분품 또는 이들을 조합한 것이 존재함을 의미하는 것이지, 하나 또는 그 이상의 다른 특징들이나 개수, 단계, 동작, 구성요소, 부분품 또는 이들을 조합한 것들의 존재 또는 부가 가능성을 배제하지 않는 것으로 이해되어야 한다.
- [0016] 도면에 대한 상세한 설명을 하기에 앞서, 본 명세서에서의 구성부들에 대한 구분은 각 구성부가 담당하는 주기능 별로 구분한 것에 불과함을 명확히 하고자 한다. 즉, 이하에서 설명할 2개 이상의 구성부가 하나의 구성부로 합쳐지거나 또는 하나의 구성부가 보다 세분화된 기능별로 2개 이상으로 분화되어 구비될 수도 있다.
- [0017] 그리고 이하에서 설명할 구성부 각각은 자신이 담당하는 주기능 이외에도 다른 구성부가 담당하는 기능 중 일부 또는 전부의 기능을 추가적으로 수행할 수도 있으며, 구성부 각각이 담당하는 주기능 중 일부 기능이 다른 구성부에 의해 전담되어 수행될 수도 있음은 물론이다. 따라서, 본 명세서를 통해 설명되는 각 구성부들의 존재 여부는 기능적으로 해석되어야 할 것이다.
- [0018] 도 1은 개시된 기술의 일 실시예에 따른  $n$ 차원 텐서 곱 연산을 이용하여 합성곱 신경망(Convolution Neural Network, CNN)을 압축하는 과정을 나타낸 도면이다. 도 1을 참조하면 연산장치는 이미지 분석을 위한 합성곱 신

경망을 저장한다. 합성곱 신경망은 이미지 분류를 위해 사전에 학습된 것으로 다양한 구조의 합성곱 신경망을 탑재할 수 있다. 연산장치는 사용자가 입력하는 이미지를 합성곱 신경망에 입력하고 합성곱 신경망을 통해 출력되는 결과에 따라 이미지를 분류할 수 있다. 물론 이미지 분류 뿐만 아니라 텍스트나 음성 등을 입력받아 특징을 추출하는 것도 가능하다. 이하에서는 이미지 분류를 예시로 설명한다.

[0019] 한편, 합성곱 신경망은 복수개의 계층(Layer)들이 쌓여있는 구조이다. 네트워크의 깊이를 늘려서 특징 추출의 성능을 개선시킬 수 있다. 그러나 상술한 구조적 특징으로 인하여 합성곱 신경망을 탑재한 연산장치는 입력된 이미지를 분석할 때 메모리 복잡도와 계산 복잡도가 높아지는 문제가 발생할 수 있다. 이러한 문제를 해소하기 위해서 연산장치는 합성곱 신경망을 구성하는 복수의 계층들에 포함된 가중치 필터의 중요도를 계산하여 기준 미만의 가중치 필터를 제거함으로써 합성곱 신경망을 압축시킬 수 있다.

[0020] 한편, 합성곱 신경망을 압축하는 과정에 있어서, 종래에는 특정 계층에 포함된 가중치 필터들의 중요도를 독립적으로 판단하였다. 예컨대, 가중치 필터의 L1 놈(Norm) 또는 L2 놈을 측정하여 높은 놈을 가진 가중치 필터는 유지하고 낮은 놈을 가진 가중치 필터를 제거하는 방식으로 신경망을 압축하였다. 그러나 이러한 방식은 가중치 필터가 속한 계층에 대한 중요도만을 판단하는 것이어서 가중치 필터가 후속 계층에 미치는 영향력을 고려하지 않았기 때문에 국소 최적 문제가 발생하는 문제가 있었다. 따라서, 본 기술에서는 가중치 필터가 후속 계층에 미치는 영향력까지 고려하여 합성곱 신경망을 압축하는 기법을 개시한다.

[0021] 한편, 연산장치는 상술한 방식의 합성곱 신경망 압축을 수행하기 위하여 합성곱 신경망에 포함된 복수개의 계층들 중 특정 계층을 선택한다. 여기에서 특정 계층을 선택하는 과정은 합성곱 신경망을 구성하는 복수개의 계층들 중 랜덤한 방식으로 하나를 선택하는 방식으로 처리될 수 있다. 해당 계층에 포함된 복수의 가중치 필터들에 대한 중요도를 연산한 다음 다른 계층을 선택할 수 있다. 물론 최초 계층부터 순차적으로 선택하여 가중치 필터의 중요도를 계산할 수도 있고 합성곱 신경망의 계층의 순서와 가중치 필터의 순서를 모두 랜덤하게 선택할 수도 있다. 가령 i번째 계층의 j번째 가중치 필터를 선택하여 중요도를 계산한 다음 k번째 계층의 s번째 가중치 필터를 선택할 수도 있다.

[0022] 한편, 연산장치는 선택한 특정 계층에 포함된 복수의 가중치 필터들 각각에 대한 n차원 텐서 곱 연산을 수행하여 복수의 가중치 필터들 각각에 대한 중요도를 계산한다. 가중치 필터의 중요도를 계산하기에 앞서 이하의 수

학식 1에 따라 합성곱 신경망의 메모리 복잡도  $L_{memory}$  와 계산 복잡도  $L_{computation}$  를 정의하면 다음과 같다.

### 수학식 1

$$L_{memory} = \sum_{i=1}^{L-1} D_i D_i T_i T_{i+1},$$

$$L_{computation} = \sum_{i=1}^{L-1} D_i D_i T_i T_{i+1} W_{i+1} H_{i+1}$$

[0024]

[0025] 여기에서  $D_i$  는 합성곱 신경망의 i번째 계층  $\mathcal{F}_i \in \mathbb{R}^{D_i \times D_i \times T_i \times T_{i+1}}$  을 구성하는 가중치 필터의 너비 및 높이를 의미하고  $T_i$  는 가중치 필터의 개수를 의미한다. 그리고  $W_{i+1}$  는 i+1번째 특징맵의 너비를 의미하고  $H_{i+1}$  는 i+1번째 특징맵의 높이를 의미한다. 특정 가중치 필터를 제거하는 방식의 신경망 압축 기법에서는 가중치 필터의 개수  $T_i$  를  $t_i$  로 줄이는 것을 목표로 한다. 따라서, 압축된 합성곱 신경망의 메모리 복잡도

$\hat{L}_{memory}$  와 계산 복잡도  $\hat{L}_{computation}$  를 정의하면 다음과 같다.

## 수학식 2

$$\hat{L}_{memory} = \sum_{i=1}^{L-1} D_i D_i t_i t_{i+1},$$

$$\hat{L}_{computation} = \sum_{i=1}^{L-1} D_i D_i t_i t_{i+1} W_{i+1} H_{i+1}$$

[0027]

[0028]

한편, 신경망의 메모리 압축률은 원본 신경망의 메모리 복잡도 대비 압축된 신경망의 메모리 복잡도로 정의할 수 있다. 마찬가지로 신경망의 계산 복잡도 압축률도 원본 신경망의 계산 복잡도 대비 압축된 신경망의 계산 복잡도로 정의할 수 있다. 이와 같이 압축된 신경망의 메모리 압축률 및 계산 복잡도 압축률은 수학식 3으로 정의할 수 있다.

## 수학식 3

$$C_{memory} = \frac{\hat{L}_{memory}}{L_{memory}},$$

$$C_{computation} = \frac{\hat{L}_{computation}}{L_{computation}}$$

[0030]

[0031]

한편, 놈(Norm) 기반 필터 선택 기법에 있어서, i번째 계층에 포함된 j번째 가중치 필터  $\mathcal{F}_{i,j} \in \mathbb{R}^{D_i \times D_i \times T_i}$ 의 중요도는 아래의 수학식 4와 같이 표현할 수 있다.

## 수학식 4

$$I_{norm}(\mathcal{F}_{i,j}) = \|\mathcal{F}_{i,j}\|_n$$

[0033]

[0034]

여기에서 i번째 계층에 포함된 j번째 가중치 필터의 중요도  $I_{norm}(\mathcal{F}_{i,j})$ 는  $\|\mathcal{F}_{i,j}\|_n$ 과 같다. 즉,

$\|\mathcal{F}_{i,j}\|_n$ 는  $\mathcal{F}_{i,j}$ 의 놈 연산자를 의미한다. 종래 기법에서는 특정 가중치 필터의 놈이 높은지 낮은지에 대해서만 판단하기 때문에 해당 가중치 필터가 속한 계층에 대하여 독립적으로 중요도를 판단한다. 즉, 가중치 필터가 후속 계층에 어떤 영향력을 미칠 것인지에 대해서는 고려하지 않기 때문에 국소 최적 문제가 발생하는 것이다.

[0035]

이러한 문제점을 해소하기 위해서 연산장치는 아래의 수학식 5에 따라 n차원 텐서 곱 연산을 수행하여 가중치 필터의 중요도를 계산할 수 있다.

### 수학식 5

$$I_{proposed}(\mathcal{F}_{i,j}) = \frac{\|\mathcal{F}_i \times_4^3 \mathcal{F}_{i+1} - \text{zero}(\mathcal{F}_i, j) \times_4^3 \mathcal{F}_{i+1}\|_1}{\|\mathcal{F}_i \times_4^3 \mathcal{F}_{i+1}\|_1},$$

[0037]

[0038]

여기에서  $\mathcal{F}_i \times_4^3 \mathcal{F}_{i+1}$  는 i번째 계층  $\mathcal{F}_i \in \mathbb{R}^{D_i \times D \times T_i \times T_{i+1}}$  과 i+1번째 계층  $\mathcal{F}_{i+1} \in \mathbb{R}^{D_i \times D \times T_{i+1} \times T_{i+2}}$  간의 n차원 곱셈 과정을 의미한다. 그리고  $\text{zero}(\mathcal{F}_i, j)$  는 i번째 계층의 j번째 가중치 필터를 0으로 구성된 가중치 필터를 의미한다. n차원 곱셈은 이하의 수학식 6을 통해 계산된다.

### 수학식 6

$$\begin{aligned} \mathcal{F}_i^{(4)} &= g(\mathcal{F}_i, T_{i+1} \times D_i D_i T_i) \in \mathbb{R}^{T_{i+1} \times D_i D_i T_i}, \\ \mathcal{F}_{i+1}^{(3)} &= g(\mathcal{F}_{i+1}, T_{i+1} \times D_{i+1} D_{i+1} T_{i+2}) \in \mathbb{R}^{T_{i+1} \times D_{i+1} D_{i+1} T_{i+2}}, \\ \mathcal{F}_i \times_4^3 \mathcal{F}_{i+1} &= \mathcal{F}_i^{(4)T} \mathcal{F}_{i+1}^{(3)} \in \mathbb{R}^{D_i D_i T_i \times D_{i+1} D_{i+1} T_{i+2}}, \end{aligned}$$

[0040]

[0041]

여기에서  $g(\mathcal{F}_i, T_{i+1} \times D_i D_i T_i)$  는 i번째 계층  $\mathcal{F}_i \in \mathbb{R}^{D_i \times D_i \times T_i \times T_{i+1}}$  을  $T_{i+1} \times D_i D_i T_i$  의 형태로 변경하는 차원 변환 함수를 의미한다. 개시된 기술에서 제안하는 방법은 n차원 텐서 곱 연산에 따라 계산된 새로운 고차원 텐서의 재구성 손실 함수(Reconstruction loss)의 형태로 정의된다. n차원 텐서 곱 연산을 통해 구해진 새로운 텐서 데이터는 특정 계층이 후속 계층에 미치는 영향력을 표현한 것이며, 특정 가중치 필터를 0으로 치환한 것과의 재구성 손실값을 통해 해당 가중치 필터의 중요도를 정량화할 수 있다.

[0042]

한편, 연산장치는 상술한 계산 과정에 따라 특정 계층의 가중치 필터들 각각에 대한 중요도를 계산할 수 있다. 그리고 각각의 결과에 따라 중요도가 기준 미만인 일부의 가중치 필터를 제거할 수 있다. 그리고 이러한 과정을 합성곱 신경망 내 모든 계층들에 각각 적용함으로써 각 계층 별로 중요도가 낮은 가중치 필터를 제거하는 방식으로 합성곱 신경망을 압축시킬 수 있다. 이 과정에서 합성곱 신경망에 포함된 다수의 파라미터들 중 잉여 파라미터들이 제거되어 계산과정에 대한 복잡도가 줄어들 수 있고 계산 시 메모리 복잡도 또한 줄어들 수 있다.

[0043]

도 2는 개시된 기술의 일 실시예에 따른 n차원 텐서 곱 연산을 이용한 합성곱 신경망 압축 방법에 대한 순서도이다. 도 2를 참조하면 합성곱 신경망 압축 방법은 210 단계 내지 230 단계를 포함한다. 각 단계는 연산장치를 통해 순차적으로 수행될 수 있다.

[0044]

210 단계에서 연산장치는 합성곱 신경망(Convolution Neural Network, CNN)에 포함된 복수개의 계층들 중 특정 계층을 선택한다. 합성곱 신경망을 구성하는 복수개의 계층들 중 랜덤한 방식으로 하나를 선택할 수도 있다. 연산장치는 각 계층 별 가중치 필터들의 중요도를 알 수 없으므로 신경망에 포함된 모든 계층을 순차적으로 선택하여 각 계층 별로 n차원 텐서 곱 연산을 수행할 수도 있다.

[0045]

220 단계에서 연산장치는 특정 계층에 포함된 복수의 가중치 필터들 각각에 대한 n차원 텐서 곱 연산을 수행하여 복수의 가중치 필터들 각각에 대한 중요도를 계산한다. 중요도를 계산하는 과정은 앞서 도 1을 통해 설명한 바와 같이 각 계층별로 독립적으로 가중치 필터의 중요도를 계산하는 것이 아니라 후속 계층에 미치는 영향력을 고려하여 중요도를 계산한다.

[0046]

230 단계에서 연산장치는 계산 결과에 따라 복수의 가중치 필터들 중 중요도가 기준 미만인 일부의 가중치 필터를 제거한다. 연산장치는 210 내지 230 단계를 일정 횟수 반복하여 합성곱 신경망에서 중요도가 낮은 가중치 필터들을 제거할 수 있다.

[0047]

도 3은 개시된 기술의 일 실시예에 따른 n차원 텐서 곱 연산을 이용한 합성곱 신경망 압축 장치에 대한 블록도이다. 도 3을 참조하면 합성곱 신경망 압축 장치(300)는 저장장치(310) 및 연산장치(320)를 포함한다.

- [0048] 저장장치(310)는 합성곱 신경망을 저장한다. 합성곱 신경망은 계층 수를 늘려서 네트워크의 깊이를 깊게 형성한 것이므로 저장장치(310)는 이러한 합성곱 신경망을 저장할 수 있을 정도의 용량을 가진 메모리를 이용할 수 있다.
- [0049] 연산장치(320)는 합성곱 신경망에 포함된 복수개의 계층들 중 특정 계층을 선택하여 특정 계층에 포함된 가중치 필터에 대한  $n$ 차원 텐서 곱 연산을 수행하여 가중치 필터에 대한 중요도를 계산한다. 그리고, 계산 결과에 따라 가중치 필터를 제거한다. 연산장치(320)는 합성곱 신경망 압축 장치(300)의 프로세서 내지는 AP일 수 있다. 연산장치(320)는  $n$ 차원 텐서 곱 연산에 따라 계산된 가중치 필터들의 텐서 데이터와 가중치 필터를 0으로 치환한 텐서 데이터를 뺀 결과가 기준 미만이면 가중치 필터를 제거할 수 있다. 연산장치는 이러한 계산 과정을 특정 계층에 포함된 나머지 가중치 필터들에도 각각 적용하여 중요도가 기준 미만인 적어도 하나의 가중치 필터를 제거할 수 있다. 이러한 계산 과정을 반복함으로써 저장장치(310)에 저장된 합성곱 신경망을 전체적으로 압축시킬 수 있다.
- [0050] 한편, 상술한 합성곱 신경망 압축 장치(300)는 컴퓨터에서 실행될 수 있는 실행가능한 알고리즘을 포함하는 프로그램(또는 어플리케이션)으로 구현될 수도 있다. 상기 프로그램은 일시적 또는 비일시적 판독 가능 매체(non-transitory computer readable medium)에 저장되어 제공될 수 있다.
- [0051] 비일시적 판독 가능 매체란 레지스터, 캐쉬, 메모리 등과 같이 짧은 순간 동안 데이터를 저장하는 매체가 아니라 반영구적으로 데이터를 저장하며, 기기에 의해 판독(reading)이 가능한 매체를 의미한다. 구체적으로는, 상술한 다양한 어플리케이션 또는 프로그램들은 CD, DVD, 하드 디스크, 블루레이 디스크, USB, 메모리카드, ROM (read-only memory), PROM (programmable read only memory), EPROM(Erasable PROM, EPROM) 또는 EEPROM(Electrically EPROM) 또는 플래시 메모리 등과 같은 비일시적 판독 가능 매체에 저장되어 제공될 수 있다.
- [0052] 일시적 판독 가능 매체는 스태틱 램(Static RAM, SRAM), 다이내믹 램(Dynamic RAM, DRAM), 싱크로너스 디램(Synchronous DRAM, SDRAM), 2배속 SDRAM(Double Data Rate SDRAM, DDR SDRAM), 증강형 SDRAM(Enhanced SDRAM, ESDRAM), 동기화 DRAM(Synclink DRAM, SDRAM) 및 직접 램버스 램(Direct Rambus RAM, DRAM) 과 같은 다양한 RAM을 의미한다.
- [0053] 도 4는 종래 합성곱 신경망의 구조를 나타낸 도면이다. 도 4를 참조하면 합성곱 신경망의 한 종류인 VGG-16과 ResNet-18은 서로 다른 구조를 가지고 있다. 먼저 VGG-16는  $3 \times 3$  가중치 필터들을 쌓아서 신경망을 형성하는 구조이며 ResNet-18는 기본적으로는 VGG-16와 동일하되 스킵커넥션이라는 구조적 특징이 추가된 것이다. 아래 표 1을 참조하면 개시된 기술을 상기 두 종류의 합성곱 신경망에 적용한 결과 종래의 높 기반 필터 제거 기법 대비 압축 성능이 개선된 것을 확인할 수 있었다. 표 1은 공인 데이터셋인 CIFAR-100을 이용하여 이미지 분류 시 VGG-16과 ResNet-18의 성능, L1 Norm 필터 선택법을 이용한 압축 시 성능, L2 Norm 필터 선택법을 이용한 압축 시 성능, 개시된 기술을 적용하여 압축 시의 성능을 각각 나타낸 것이다.

표 1

|           | # of params         | Accuracy (%)     |
|-----------|---------------------|------------------|
| VGG-16    | 14723136            | 72.48            |
| L1 Norm   | 1833621<br>(12.45%) | 63.94<br>(-8.54) |
| L2 Norm   | 1833621<br>(12.45%) | 62.59<br>(-9.89) |
| Ours      | 1811760<br>(12.31%) | 65.73<br>(-6.75) |
| ResNet-18 | 6448192             | 76.38            |
| L1 Norm   | 929660<br>(14.4%)   | 68.17<br>(-8.21) |
| L2 Norm   | 929660<br>(14.4%)   | 67.29<br>(-9.09) |
| Ours      | 914923<br>(14.2%)   | 71.45<br>(-4.93) |

[0056]

[0057]

상기 표 1을 참조하면 VGG-16의 원본 신경망의 파라미터 개수(# of params)는 14723136개이고 정확도(Accuracy)는 72.48%를 나타냈다. 여기에서 L1 Norm 기반의 필터 선택법을 적용한 결과 파라미터 개수는 1833621개로 줄어들었고 정확도는 63.94%를 나타내었다. L2 Norm 기반 필터 선택법을 적용한 결과도 파라미터 개수 1833621개와 정확도 62.59%를 나타내었다. 두 기법 모두 비슷한 성능을 나타내었다. 그리고 개시된 기술을 VGG-16에 적용한 결과 파라미터 개수는 1811760개로 종래 기법 대비 더 많은 파라미터들이 줄어들었으며 정확도는 65.73%로 오히려 높아진 것을 확인할 수 있었다.

[0058]

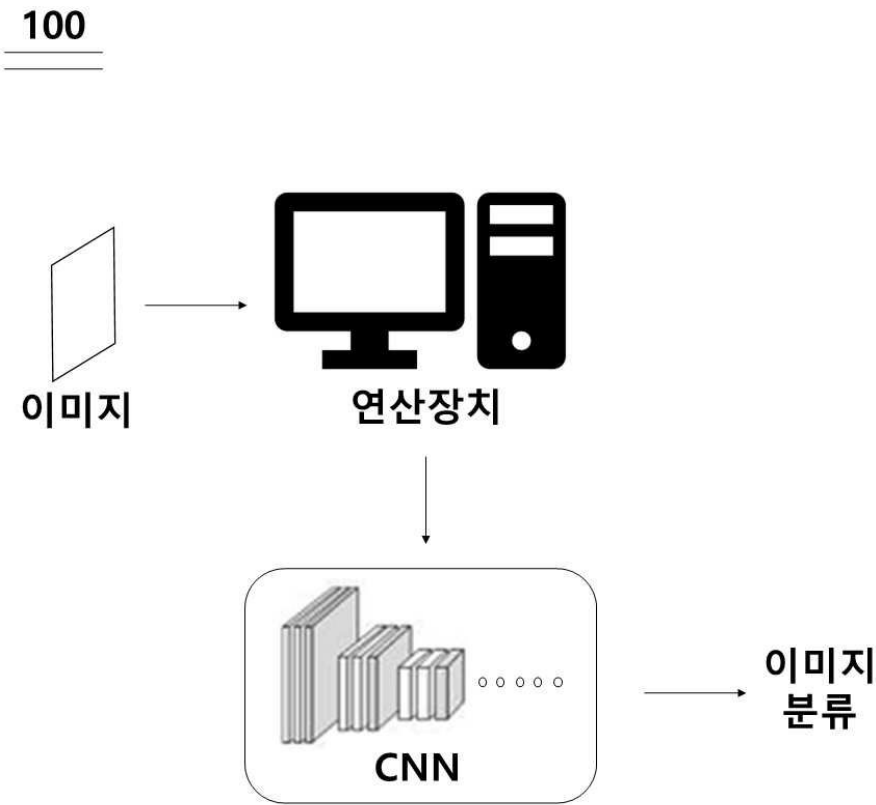
한편 ResNet-18의 원본 신경망의 파라미터 개수(# of params)는 6448192개이고 정확도(Accuracy)는 76.38%를 나타냈다. 여기에서 L1 Norm 기반의 필터 선택법을 적용한 결과 파라미터 개수는 929660개로 줄어들었고 정확도는 68.17%를 나타내었다. L2 Norm 기반 필터 선택법을 적용한 결과는 파라미터 개수 929660개와 정확도 67.29%를 나타내었다. 마찬가지로 두 기법 모두 비슷한 성능을 나타내었다. 그리고 개시된 기술을 ResNet-18에 적용한 결과 파라미터 개수는 914923개로 종래 기법 대비 더 많은 파라미터들이 줄어들었으며 정확도는 71.45%로 오히려 높아진 것을 확인할 수 있었다. 즉, 합성곱 신경망의 구조와 상관없이 개시된 기술을 적용하여 신경망을 압축한 결과가 종래의 L1, L2 Norm 기반 필터 선택법 대비 더 높은 성능을 나타내는 것을 확인할 수 있었다. 따라서, 개시된 기술을 합성곱 신경망의 종류나 구조에 구애받지 않고 적용시킴으로써 신경망을 효율적으로 압축시킬 수 있고 메모리 복잡도와 계산 복잡도를 줄일 수 있다.

[0059]

개시된 기술의 일 실시예에 따른 n차원 텐서 곱 연산을 이용한 합성곱 신경망 압축 방법 및 장치는 이해를 돕기 위하여 도면에 도시된 실시 예를 참고로 설명되었으나, 이는 예시적인 것에 불과하며, 당해 분야에서 통상적 지식을 가진 자라면 이로부터 다양한 변형 및 균등한 타 실시 예가 가능하다는 점을 이해할 것이다. 따라서, 개시된 기술의 진정한 기술적 보호범위는 첨부된 특허청구범위에 의해 정해져야 할 것이다.

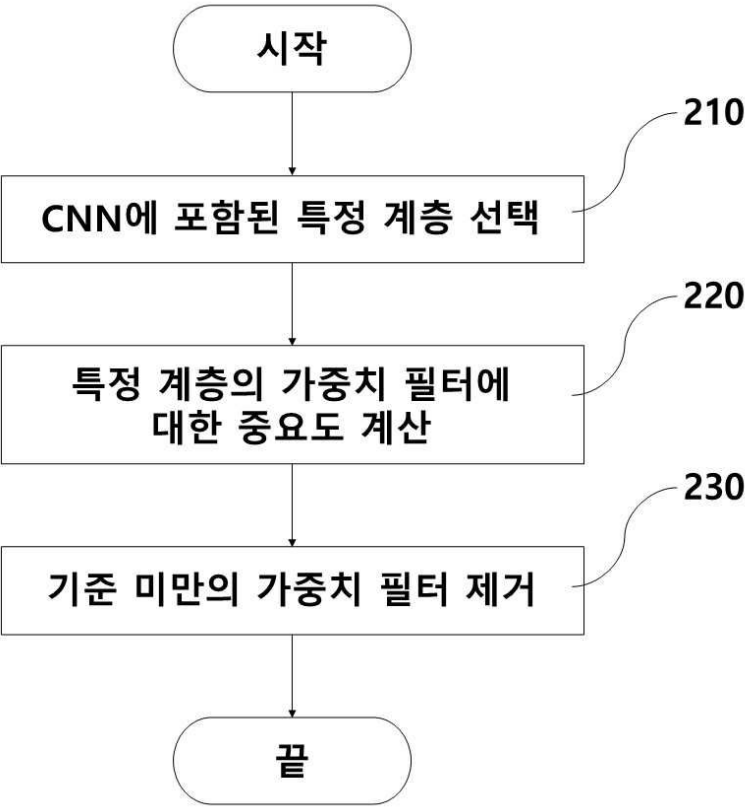
도면

도면1

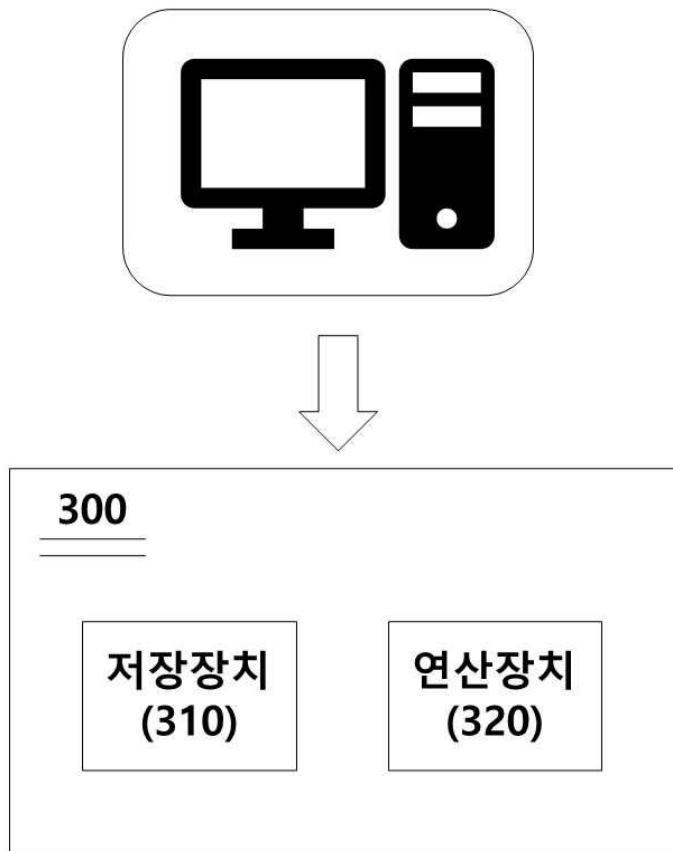


도면2

200



도면3



도면4

400

