



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2025-0029469
(43) 공개일자 2025년03월05일

(51) 국제특허분류(Int. Cl.)
G06N 3/0985 (2023.01) G06N 3/045 (2023.01)
(52) CPC특허분류
G06N 3/0985 (2023.01)
G06N 3/045 (2023.01)
(21) 출원번호 10-2023-0110427
(22) 출원일자 2023년08월23일
심사청구일자 2023년08월23일

(71) 출원인
세종대학교산학협력단
서울특별시 광진구 능동로 209 (군자동, 세종대학교)
(72) 발명자
구영현
서울특별시 광진구 능동로 209, 세종대학교
대양AI센터 717호
김동
서울특별시 광진구 능동로 209, 세종대학교
대양AI센터 410호
(74) 대리인
수안특허법인, 민영준

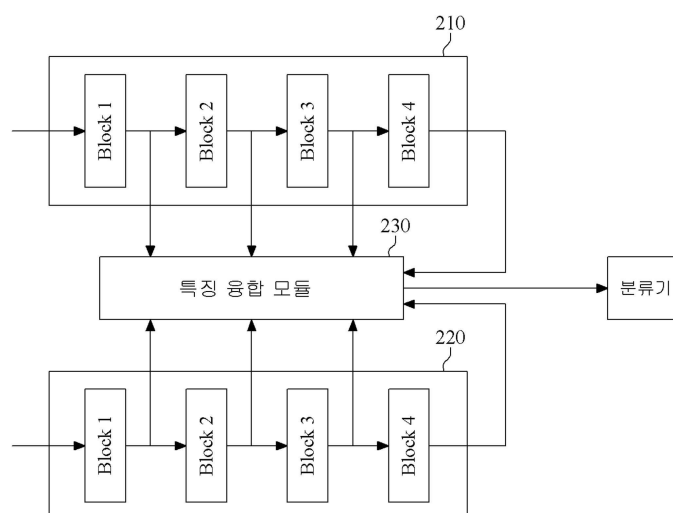
전체 청구항 수 : 총 11 항

(54) 발명의 명칭 멀티 모달 신경망 구조 탐색 방법

(57) 요약

멀티 모달리티 데이터를 이용하는 멀티 모달 신경망의 구조를 탐색하는 방법 및 장치가 개시된다. 개시된 멀티 모달 신경망 구조 탐색 방법은 제1신경망의 복수의 특징 추출 블록 각각과 제1특징 융합 모듈 사이의 제1링크에 할당된 가중치 중 하나와, 제2신경망의 복수의 특징 추출 블록 각각과 상기 제1특징 융합 모듈 사이의 제2링크에 할당된 가중치 중 하나를 선택하여, 제1가중치 조합을 생성하는 단계; 상기 제1가중치 조합에 포함된 가중치를 서로 곱하는 단계; 및 상기 제1 및 제2링크 중에서, 상기 곱한 값이 가장 큰 가중치 조합의 가중치에 대응되는 링크를 선택하는 단계를 포함하며, 상기 제1신경망은 제1모달리티 데이터로부터 특징값을 추출하는 신경망이며, 상기 제2신경망은 제2모달리티 데이터로부터 특징값을 추출하는 신경망이다.

대표도 - 도2



이 발명을 지원한 국가연구개발사업

과제고유번호 1711193685

과제번호 2022-0-00739-002

부처명 과학기술정보통신부

과제관리(전문)기관명 정보통신기획평가원

연구사업명 SW컴퓨팅산업원천기술개발(R&D, 정보화)

연구과제명 머신러닝 개발 전주기를 연결하고 쉽게 사용할 수 있는 자동화 MLOps 플랫폼 기술 개발

기 여 율 1/1

과제수행기관명 (주)와이즈넷

연구기간 2023.01.01 ~ 2023.12.31

명세서

청구범위

청구항 1

제1신경망의 복수의 특징 추출 블록 각각과 제1특징 융합 모듈 사이의 제1링크에 할당된 가중치 중 하나와, 제2신경망의 복수의 특징 추출 블록 각각과 상기 제1특징 융합 모듈 사이의 제2링크에 할당된 가중치 중 하나를 선택하여, 제1가중치 조합을 생성하는 단계;

상기 제1가중치 조합에 포함된 가중치를 서로 곱하는 단계; 및

상기 제1 및 제2링크 중에서, 상기 곱한 값이 가장 큰 가중치 조합의 가중치에 대응되는 링크를 선택하는 단계를 포함하며,

상기 제1신경망은 제1모달리티 데이터로부터 특징값을 추출하는 신경망이며,

상기 제2신경망은 제2모달리티 데이터로부터 특징값을 추출하는 신경망인,

멀티 모달 신경망 구조 탐색 방법.

청구항 2

제 1항에 있어서,

상기 특징값을 이용해 타겟 태스크를 수행하는 상기 멀티 모달 신경망의 손실값이 최소가 되도록, 상기 가중치에 대한 학습을 수행하는 단계

를 더 포함하는 멀티 모달 신경망 구조 탐색 방법.

청구항 3

제 2항에 있어서,

상기 가중치에 대한 학습을 수행하는 단계는

상기 가중치를 이용하여, 상기 제1 및 제2신경망의 특징 추출 블록에서 출력되는 특징값의 가중합을 계산하며,

상기 가중합된 값이, 상기 제1특징 융합 모듈로 입력되는

멀티 모달 신경망 구조 탐색 방법.

청구항 4

제 3항에 있어서,

상기 제1특징 융합 모듈로 입력되는 상기 가중합된 값의 개수는

상기 제1특징 융합 모듈이 융합하는 특징값의 개수에 대응되는

멀티 모달 신경망 구조 탐색 방법.

청구항 5

제 1항에 있어서,

상기 제1특징 융합 모듈은

상기 선택된 링크를 통해 전달되는 상기 제1 및 제2모달리티 데이터에 대한 특징값을 융합하는,
멀티 모달 신경망 구조 탐색 방법.

청구항 6

제 1항에 있어서,

상기 제1신경망의 복수의 특징 추출 블록 각각과 제2특징 융합 모듈 사이의 제3링크에 할당된 가중치 중 하나와, 상기 제2신경망의 복수의 특징 추출 블록 각각과 상기 제2특징 융합 모듈 사이의 제4링크에 할당된 가중치 중 하나와, 상기 제1특징 융합 모듈과 상기 제2특징 융합 모듈 사이의 제5링크에 할당된 가중치에 대한 제2가중치 조합을 생성하는 단계;

상기 제2가중치 조합에 포함된 가중치를 서로 곱하는 단계; 및

상기 제3 내지 제5링크 중에서, 상기 곱한 값이 가장 큰 가중치 조합의 가중치에 대응되는 링크를 선택하는 단계

를 더 포함하는 멀티 모달 신경망 구조 탐색 방법.

청구항 7

제 6항에 있어서,

상기 제2가중치 조합을 생성하는 단계는

상기 제3 내지 제5링크에 할당된 가중치 중에서, 서로 다른 2개의 링크의 가중치를 이용하여, 상기 제2가중치 조합을 생성하는

멀티 모달 신경망 구조 탐색 방법.

청구항 8

제1신경망의 복수의 특징 추출 블록 각각에서 출력되는 특징값 중 하나와, 제2신경망의 복수의 특징 추출 블록 각각에서 출력되는 특징값 중 하나를 선택하는 단계; 및

복수의 융합 연산 각각에 할당된 가중치 중, 최대 가중치에 대응되는 융합 연산을 이용하여, 상기 선택된 특징값을 융합하는 단계를 포함하며,

상기 제1신경망은 제1모달리티 데이터로부터 상기 특징값을 추출하는 신경망이며,

상기 제2신경망은 제2모달리티 데이터로부터 상기 특징값을 추출하는 신경망인,

멀티 모달 신경망 구조 탐색 방법.

청구항 9

제 8항에 있어서,

상기 특징값을 이용해 타겟 태스크를 수행하는 상기 멀티 모달 신경망의 손실값이 최소가 되도록, 상기 가중치에 대한 학습을 수행하는 단계

를 더 포함하는 멀티 모달 신경망 구조 탐색 방법.

청구항 10

제 9항에 있어서,

상기 가중치에 대한 학습을 수행하는 단계는

상기 복수의 융합 연산 각각을 이용하여, 상기 제1 및 제2신경망의 특징 추출 블록 각각에서 출력되는 특징값을 융합하는 단계;

상기 복수의 융합 연산 각각에 할당된 가중치를 이용하여, 융합된 특징값에 대한 가중합을 계산하는 단계; 및

상기 가중합된 값을 이용하여, 상기 가중치에 대한 학습을 수행하는 단계

를 포함하는 멀티 모달 신경망 구조 탐색 방법.

청구항 11

메모리; 및

상기 메모리와 전기적으로 연결된 적어도 하나의 프로세서를 포함하며,

상기 프로세서는

제1신경망의 복수의 특징 추출 블록 각각과 특징 융합 모듈 사이의 제1링크에 할당된 가중치 중 하나와, 제2신경망의 복수의 특징 추출 블록 각각과 상기 특징 융합 모듈 사이의 제2링크에 할당된 가중치 중 하나를 선택하여, 가중치 조합을 생성하고,

상기 제1 및 제2링크 중에서, 상기 가중치 조합에 포함된 가중치를 서로 곱한 값이 가장 큰 가중치 조합의 가중치에 대응되는 링크를 선택하고,

복수의 융합 연산 각각에 할당된 가중치 중, 최대 가중치에 대응되는 융합 연산을 이용하여, 상기 선택된 링크를 통해 전달되는 특징값을 융합하며,

상기 제1신경망은 제1모달리티 데이터로부터 특징값을 추출하는 신경망이며,

상기 제2신경망은 제2모달리티 데이터로부터 특징값을 추출하는 신경망인,

멀티 모달 신경망 구조 탐색 장치.

발명의 설명

기술 분야

[0001] 본 발명은 신경망 구조 탐색 방법 및 장치에 관한 것으로, 더욱 상세하게는 멀티 모달리티 데이터를 이용하는 멀티 모달 신경망의 구조를 탐색하는 방법 및 장치에 관한 것이다.

배경 기술

[0003] 현재 인공지능 분야에서는 VGG, MobileNet, ResNet, DenseNet 등 다양한 구조의 신경망 모델을 이미지 분류, 얼굴 인식, 자율주행 등 작업에 활용하고 있다. 높은 성능을 낼 수 있는 신경망 구조를 설계하고 하이퍼파라미터를 최적화하기 위해서는 많은 실험을 진행하고 다양한 시행착오를 거쳐야 하며, 이러한 과정은 상당한 시간과 노력을 필요로 한다.

[0004] 이에 주어진 데이터 셋 혹은 특정 태스크에 대해 자동으로 최적의 신경망 구조를 탐색하는 신경망 구조 탐색(neural architecture search, NAS) 기술이 연구되고 있다. 신경망 구조 탐색 기술에는 진화 알고리즘 기반 탐색, 강화학습 기반 탐색, 경사 하강법 기반 탐색 방법 등이 있으며, 신경망 구조 탐색 기술을 통해 신경망 모델의 가중치와 파라미터가 자동으로 최적화될 수 있다.

[0005] 기존 신경망 구조 탐색 기술은 주로, 단일 모달리티(modality) 데이터를 이용하는 신경망을 중심으로 연구되었다. 하지만 최근 이미지, 텍스트, 음성 데이터 등 다양한 모달리티 데이터를 동시에 여러 개를 이용해 특정 태스크를 수행하는 신경망 기술이 발달함에 따라서, 멀티 모달 신경망의 구조를 탐색하는 기술 개발이 요구되고 있다.

[0006] 관련 선행문헌으로 특허 문헌인 대한민국 등록특허 제10-2360246호, 대한민국 공개특허 제2023-0055937호, 제2023-0076799호가 있다.

발명의 내용

해결하려는 과제

[0008] 본 발명은, 멀티 모달리티 데이터를 입력받아 특정 태스크를 수행하는 멀티 모달 신경망의 구조를 자동으로 탐색하는 방법 및 장치를 제공하기 위한 것이다.

[0009] 또한 본 발명은 서로 다른 모달리티 데이터에 대한 특징값 중 융합할 특징값을 선택하고, 선택된 특징값을 융합하는 멀티 모달 신경망 구조를 탐색하는 방법 및 장치를 제공하기 위한 것이다.

과제의 해결 수단

[0011] 상기한 목적을 달성하기 위한 본 발명의 일 실시예에 따르면, 제1신경망의 복수의 특징 추출 블록 각각과 제1특징 융합 모듈 사이의 제1링크에 할당된 가중치 중 하나와, 제2신경망의 복수의 특징 추출 블록 각각과 상기 제1특징 융합 모듈 사이의 제2링크에 할당된 가중치 중 하나를 선택하여, 제1가중치 조합을 생성하는 단계; 상기 제1가중치 조합에 포함된 가중치를 서로 곱하는 단계; 및 상기 제1 및 제2링크 중에서, 상기 곱한 값이 가장 큰 가중치 조합의 가중치에 대응되는 링크를 선택하는 단계를 포함하며, 상기 제1신경망은 제1모달리티 데이터로부터 특징값을 추출하는 신경망이며, 상기 제2신경망은 제2모달리티 데이터로부터 특징값을 추출하는 신경망인, 멀티 모달 신경망 구조 탐색 방법이 제공된다.

[0012] 또한 상기한 목적을 달성하기 위한 본 발명의 다른 실시예에 따르면, 제1신경망의 복수의 특징 추출 블록 각각에서 출력되는 특징값 중 하나와, 제2신경망의 복수의 특징 추출 블록 각각에서 출력되는 특징값 중 하나를 선택하는 단계; 및 복수의 융합 연산 각각에 할당된 가중치 중, 최대 가중치에 대응되는 융합 연산을 이용하여, 상기 선택된 특징값을 융합하는 단계를 포함하며, 상기 제1신경망은 제1모달리티 데이터로부터 상기 특징값을 추출하는 신경망이며, 상기 제2신경망은 제2모달리티 데이터로부터 상기 특징값을 추출하는 신경망인, 멀티 모달 신경망 구조 탐색 방법이 제공된다.

[0013] 또한 상기한 목적을 달성하기 위한 본 발명의 또 다른 실시예에 따르면, 메모리; 및 상기 메모리와 전기적으로 연결된 적어도 하나의 프로세서를 포함하며, 상기 프로세서는 제1신경망의 복수의 특징 추출 블록 각각과 특징 융합 모듈 사이의 제1링크에 할당된 가중치 중 하나와, 제2신경망의 복수의 특징 추출 블록 각각과 상기 특징 융합 모듈 사이의 제2링크에 할당된 가중치 중 하나를 선택하여, 가중치 조합을 생성하고, 상기 제1 및 제2링크 중에서, 상기 가중치 조합에 포함된 가중치를 서로 곱한 값이 가장 큰 가중치 조합의 가중치에 대응되는 링크를 선택하고, 복수의 융합 연산 각각에 할당된 가중치 중, 최대 가중치에 대응되는 융합 연산을 이용하여, 상기 선택된 링크를 통해 전달되는 특징값을 융합하며, 상기 제1신경망은 제1모달리티 데이터로부터 특징값을 추출하는 신경망이며, 상기 제2신경망은 제2모달리티 데이터로부터 특징값을 추출하는 신경망인, 멀티 모달 신경망 구조 탐색 장치가 제공된다.

발명의 효과

[0015] 본 발명의 일 실시예에 따르면, 멀티 모달 신경망의 구조가 자동으로 탐색될 수 있으며, 따라서 멀티 모달 신경망의 설계 경험이 부족한 비전문가도 용이하게 멀티 모달 신경망을 설계할 수 있다.

[0016] 또한 본 발명의 일 실시예에 따르면, 학습 과정을 통해 서로 다른 모달리티 데이터에 대한 특징값 중 융합할 특징값이 선택되고, 선택된 특징값의 융합 연산이 선택됨으로써, 높은 성능이 발휘되는 멀티 모달 신경망 구조가

설계될 수 있다.

도면의 간단한 설명

- [0018] 도 1은 멀티 모달 신경망의 일예를 도시하는 도면이다.
- 도 2는 본 발명의 일실시예에 따른 멀티 모달 신경망을 도시하는 도면이다.
- 도 3은 특징 추출 블록과 특징 융합 모듈 사이의 링크에 할당된 가중치를 설명하기 위한 도면이다.
- 도 4는 본 발명의 일실시예에 따른, 멀티 모달 신경망 구조 탐색 방법을 설명하기 위한 도면이다.
- 도 5는 본 발명의 다른 실시예에 따른 멀티 모달 신경망을 도시하는 도면이다.
- 도 6은 본 발명의 다른 실시예에 따른 멀티 모달 신경망 구조 탐색 방법을 설명하기 위한 도면이다.
- 도 7은 본 발명의 일실시예에 따른 특징 융합 모듈을 도시하는 도면이다.

발명을 실시하기 위한 구체적인 내용

- [0019] 본 발명은 다양한 변경을 가할 수 있고 여러 가지 실시예를 가질 수 있는 바, 특정 실시예들을 도면에 예시하고 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. 각 도면을 설명하면서 유사한 참조부호를 유사한 구성요소에 대해 사용하였다.
- [0020] 이하에서, 본 발명에 따른 실시예들을 첨부된 도면을 참조하여 상세하게 설명한다.
- [0022] 도 1은 멀티 모달 신경망의 일예를 도시하는 도면이다.
- [0023] 도 1에 도시된 신경망은, 이미지와 텍스트라는 멀티 모달리티 데이터를 입력받아, 입력된 데이터를 미리 라벨링된 영화 클래스 중 하나로 분류하는 신경망이다.
- [0024] 제1특징 추출 네트워크(110)는 이미지 데이터를 입력받아, 이미지 데이터에 대한 특징값을 추출하고, 제2특징 추출 네트워크(120)는 텍스트 데이터를 입력받아, 텍스트 데이터에 대한 특징값을 추출한다. 추출된 특징값은 연결(concatenation)되어 분류기(130)로 입력되고, 분류기(130)는 입력 데이터를 미리 라벨링된 영화 클래스 중 하나로 분류한다.
- [0025] 제1 및 제2특징 추출 네트워크(110, 120)는 복수의 특징 추출 블록(Block1 내지 Block4)을 포함할 수 있으며, 특징 추출 블록 각각은 순차적으로 연결되어, 입력된 데이터에 대한 특징값을 출력한다. 예컨대, 제1 및 제2특징 추출 네트워크(110, 120)가 컨벌루션 기반의 네트워크라고 할 경우, 특징 추출 블록은 컨벌루션 레이어에 대응될 수 있다.
- [0026] 이러한 멀티 모달 신경망에서, 신경망 구조를 탐색하는 것은, 멀티 모달리티 데이터의 특징을 선택하는 문제와, 선택된 특징을 융합하는 방법을 설계하는 것에 대응되며, 본 발명은 멀티 모달 신경망에서, 융합할 멀티 모달리티 데이터의 특징을 선택하고, 선택된 특징을 융합하는 신경망 구조 탐색 방법을 제안한다.
- [0027] 본 발명의 일실시예는, 멀티 모달리티 데이터를 입력받는 특징 추출 네트워크 각각의 서로 다른 특징 추출 블록에서 출력되는 특징값 중에서 융합할 특징값을 선택하며, 복수의 융합 연산 중에서 선택된 융합 연산을 이용해 선택된 특징값을 융합한다.
- [0028] 융합할 특징값을 선택하고, 융합 연산을 선택하기 위해 가중치가 이용되며, 이러한 가중치는 신경망 구조 탐색 기법을 통해 학습될 수 있다. 일실시예로서, 전술된 가중치는 경사 하강법 기반의 Differentiable Architecture Search(DARTS) 알고리즘을 통해 학습될 수 있다.
- [0029] 본 발명의 일실시예에 따른 멀티 모달 신경망 구조 탐색 방법은, 메모리 및 메모리와 전기적으로 연결된 적어도 하나의 프로세서를 포함하는 컴퓨팅 장치에서 수행될 수 있으며, 프로세서는 멀티 모달 신경망 구조 탐색을 위한 일련의 과정을 수행한다. 컴퓨팅 장치는 데스크탑, 서버 등이거나 또는 멀티 모달 신경망 구조 탐색을 위한 별도의 장치일 수 있다.

- [0030] 이하에서는 본 발명의 일실시예에 따른 멀티 모달 신경망 구조 탐색 방법을, 특징 선택 방법과 특징 융합 방법으로 나누어 설명하기로 한다.
- [0032] **특징 선택 방법**
- [0033] 도 2는 본 발명의 일실시예에 따른 멀티 모달 신경망을 도시하는 도면이며, 도 3은 특징 추출 블록과 특징 융합 모듈 사이의 링크에 할당된 가중치를 설명하기 위한 도면이다. 그리고 도 4는 본 발명의 일실시예에 따른, 멀티 모달 신경망 구조 탐색 방법을 설명하기 위한 도면이다.
- [0034] 도 2에서는 4개의 특징 추출 블록과 1개의 특징 융합 모듈이 이용되는 실시예가 도시되나, 특징 추출 블록과 특징 융합 모듈의 개수는 실시예에 따라서 달라질 수 있다.
- [0035] 도 2를 참조하면, 본 발명의 일실시예에 따른 멀티 모달 신경망은 제1 및 제2신경망(210, 220)을 포함하며, 선택된 특징을 입력받아 융합하는 특징 융합 모듈(230)을 포함한다. 제1 및 제2신경망(210, 220)은 각각 제1 및 제2모달리티 데이터로부터 특징값을 추출하는 신경망으로서, 복수의 특징 추출 블록(Block1 내지 Block4)을 포함한다. 제1 및 제2신경망(210, 220)의 특징 추출 블록 각각과 특징 융합 모듈(230) 사이에서 특징값이 전달되는 링크에는 가중치가 할당되고, 가중치는 학습 과정을 통해 결정된다. 도 2에서, 특징 추출 블록의 출력단에서 특징 융합 모듈(230)을 연결하는 화살표는 링크를 나타낸다.
- [0036] 도 3에 도시된 바와 같이, 특징 추출 블록(Block1) 각각과 특징 융합 모듈(230) 사이의 링크에 할당되는 가중치는 일실시예로서, 아이덴티티(Identity) 가중치와, 제로(Zero) 가중치로 구분될 수 있다. 아이덴티티 가중치는 아이덴티티 연산에 할당되는 가중치이며, 제로 가중치는 제로 연산에 할당되는 가중치이다. 아이덴티티 연산은 특징 추출 블록의 특징값을 그대로 특징 융합 모듈(230)로 입력하는 연산이며, 제로 연산은 특징 추출 블록의 특징값 대신 0을 특징 융합 모듈(230)로 입력하는 연산이다.
- [0037] 제로 가중치는 실제 신경망 구조 탐색 과정에서 영향을 미치지 못하므로, 이하 언급되는 제로 또는 아이덴티티로 가중치로 구분되지 않고 사용되는 가중치는, 모두 아이덴티티 가중치를 의미한다.
- [0038] 도 2 내지 도 4를 참조하면, 본 발명의 일실시예에 따른 컴퓨팅 장치는 제1신경망(210)의 복수의 특징 추출 블록 각각과 특징 융합 모듈(230) 사이의 제1링크에 할당된 가중치 중 하나와, 제2신경망(220)의 복수의 특징 추출 블록 각각과 특징 융합 모듈(230) 사이의 제2링크에 할당된 가중치 중 하나를 선택하여, 가중치 조합을 생성(S410)한다. 제1링크는 제1신경망(210)의 복수의 특징 추출 블록 각각과 특징 융합 모듈(230) 사이의 4개의 링크를 나타내며, 제2링크는 제2신경망(220)의 복수의 특징 추출 블록 각각과 특징 융합 모듈(230) 사이의 4개의 링크를 나타낸다. 가중치 조합은 제1링크의 가중치 중 하나, 제2링크의 가중치 중 하나로 이루어진 2개의 가중치를 포함한다.
- [0039] 컴퓨팅 장치는 가중치 조합에 포함된 가중치를 서로 곱하는 곱셈 연산을 수행(S420)하고, 제1 및 제2링크 중에서, 곱한 값이 가장 큰 가중치 조합의 가중치에 대응되는 링크를 선택(S430)한다.
- [0040] 예컨대, 도 2와 같은 멀티 모달 신경망에서, 총 16개의 가중치 조합이 생성될 수 있으며, 컴퓨팅 장치는 16개의 가중치 조합 중, 가중치의 곱셈 값이 가장 큰 가중치 조합의 가중치에 대응되는 링크를 선택할 수 있다. 전술된 바와 같이, 가중치 조합은 제1링크의 가중치 중 하나와 제2링크의 가중치 중 하나로 이루어지므로, 선택되는 링크는, 제1링크 중에서 하나, 제2링크 중에서 하나이다.
- [0041] 선택된 링크를 통해 전달되는 제1 및 제2모달리티 데이터에 대한 특징값은, 특징 융합 모듈(230)에서 융합된다.
- [0042] 한편, 본 발명의 일실시예에 따른 컴퓨팅 장치는 특징값을 이용해 타겟 태스크를 수행하는 멀티 모달 신경망의 손실값이 최소가 되도록, 가중치에 대한 학습을 수행한다. 예컨대, 타겟 태스크가 입력 데이터의 클래스를 미리 라벨링된 영화 클래스 중 하나로 분류하는 태스크라고 할 때, 컴퓨팅 장치는 제1 및 제2신경망(210, 220)의 특징 추출 블록의 특징값 중에서 선택된 특징값을 융합하고, 융합된 특징값을 이용해 입력 데이터의 클래스를, 미리 라벨링된 영화 클래스 중 하나로 분류한다. 그리고 분류 데이터를 미리 설정된 손실 함수에 입력하여, 손실값이 최소가 되도록, 가중치에 대한 학습을 수행할 수 있다. 그리고 이와 같이 가중치를 결정하는 학습 과정에서, 전술된 DARTS 알고리즘이 이용될 수 있다.
- [0043] 또한 이러한 학습 과정에서 컴퓨팅 장치는, 제1 및 제2링크에 할당된 가중치를 이용하여, 제1 및 제2신경망(210, 220)의 특징 추출 블록에서 출력되는 특징값의 가중합을 계산하며, 가중합된 값이 특징 융합 모듈(230)로

입력된다. 특징 추출 블록의 특징값을 가중합하는 것은, 특징 융합 모듈(230)에서 융합되는 특징값의 개수가, 특징 추출 블록의 개수보다 적기 때문이다. 특징 융합 모듈(230)로 입력되는 가중합된 값의 개수는, 특징 융합 모듈(230)이 융합하는 특징값의 개수에 대응된다.

[0044] 도 2와 같은 멀티 모달 신경망에서, 컴퓨팅 장치는 [수학식 1]을 이용하여, 가중합(c^j)을 계산할 수 있다.

수학식 1

$$c^j = \sum_{k \in K} w_k^z * zero(O_k) + w_k^i * identity(O_k)$$

[0045] $K \in \{a1, a2, a3, a4, b1, b2, b3, b4\}$

[0046] 여기서, j는 특징 융합 모듈의 인덱스, w_k^z 는 제로 가중치, w_k^i 는 아이덴티티 가중치, $zero(O_k)$ 는 제로 연산 값, $identity(O_k)$ 는 아이덴티티 연산 값, a1 내지 a4는 제1링크, b1 내지 b4는 제2링크를 나타낸다.

[0047] 그리고, 특징 융합 모듈(230)이 융합하는 특징값의 개수는 2개이므로, 컴퓨팅 장치는 특징 추출 블록의 특징값을 가중합한 후, 가중합된 값을 특징 융합 모듈(230)에 대한 2개의 입력 데이터로 이용한다. 즉, 가중합된 값은 하나이지만, 2개의 동일한 가중합된 값이 특징 융합 모듈(230)로 입력되어, 특징 융합 모듈(230)은 2개의 동일한 가중합된 값을 융합한다.

[0048] 특징값은 특징맵 형태일 수 있으며, 특징 추출 블록에서 출력되는 특징맵의 사이즈는 모두 다를 수 있다. 이러한 특징맵의 사이즈는, 일실시예로서, patch wise global average pooling을 통해 동일해질 수 있으며, 특징맵의 사이즈 조정후 가중합이 계산될 수 있다.

[0050] 도 5는 본 발명의 다른 실시예에 따른 멀티 모달 신경망을 도시하는 도면이다.

[0051] 도 2의 신경망과 달리, 도 5의 신경망은 2개의 특징 융합 모듈인, 제1 및 제2특징 융합 모듈(230, 240)을 포함한다. 제1특징 융합 모듈(230)은, 전술된 특징 융합 모듈에 대응된다.

[0052] 이러한 실시예에서 컴퓨팅 장치는 제2특징 융합 모듈(240)을 위한 특징값 선택 과정을 추가로 수행한다. 컴퓨팅 장치는 제1신경망(210)의 복수의 특징 추출 블록 각각과 제2특징 융합 모듈(240) 사이의 제3링크에 할당된 가중치 중 하나와, 제2신경망(220)의 복수의 특징 추출 블록 각각과 제2특징 융합 모듈(240) 사이의 제4링크에 할당된 가중치 중 하나와, 제1특징 융합 모듈(230)과 제2특징 융합 모듈(240) 사이의 제5링크에 할당된 가중치에 대한 가중치 조합을 추가로 생성한다. 컴퓨팅 장치는 제3 내지 제5링크에 할당된 가중치 중에서, 서로 다른 2개의 링크의 가중치를 이용하여, 가중치 조합을 생성하며, 이러한 가중치 조합은 서로 다른 링크에 할당된 2개의 가중치로 이루어진다.

[0053] 즉, 컴퓨팅 장치는 제1 및 제2링크에 할당된 가중치로 이루어진 제1가중치 조합과 함께, 제3 내지 제5링크에 할당된 가중치로 이루어진 제2가중치 조합을 추가로 생성한다. 제3링크는 제1신경망(210)의 복수의 특징 추출 블록 각각과 제2특징 융합 모듈(240) 사이의 4개의 링크를 나타내며, 제4링크는 제2신경망(220)의 복수의 특징 추출 블록 각각과 제2특징 융합 모듈(240) 사이의 4개의 링크를 나타낸다. 그리고 제5링크는 제1특징 융합 모듈(230)과 제2특징 융합 모듈(240) 사이의 1개의 링크를 나타낸다.

[0054] 그리고 컴퓨팅 장치는 제2가중치 조합에 포함된 가중치를 서로 곱하고, 제3 내지 제5링크 중에서, 곱한 값이 가장 큰 제2가중치 조합의 가중치에 대응되는 링크를 추가로 선택한다. 선택된 링크는 제3 내지 제5링크 중 서로 다른 링크 2개이다.

[0055] 선택된 링크를 통해 전달되는 제1 및 제2모달리티 데이터에 대한 특징값은, 제2특징 융합 모듈(240)에서 융합된다.

[0057] 특징 융합 방법

[0058] 도 6은 본 발명의 다른 실시예에 따른 멀티 모달 신경망 구조 탐색 방법을 설명하기 위한 도면이며, 도 7은 본 발명의 일실시예에 따른 특징 융합 모듈을 도시하는 도면이다.

[0059] 도 6을 참조하면, 본 발명의 일실시예에 따른 컴퓨팅 장치는 제1신경망의 복수의 특징 추출 블록 각각에서 출력되는 특징값 중 하나와, 제2신경망의 복수의 특징 추출 블록 각각에서 출력되는 특징값 중 하나를 선택(S610)한다. 컴퓨팅 장치는, 전술된 특징 선택 방법과 같이 특징값을 선택할 수 있다.

[0060] 그리고 컴퓨팅 장치는 복수의 융합 연산 각각에 할당된 가중치 중, 최대 가중치에 대응되는 융합 연산을 이용하여, 단계 S610에서 선택된 특징값을 융합(S620)한다.

[0061] 특징값을 융합하는 연산에는 Sum 연산, ConcatAndFC 연산, ConcatAndPool 연산, Attention 연산 등이 있으며, 이러한 융합 연산 중 복수개가 단계 S620에서 이용될 수 있다.

[0062] 도 7에는 3개의 융합 연산(710, 720, 730)이 이용되는 특징 융합 모듈이 도시되어 있으며, 제1 내지 제3융합 연산(710, 720, 730)에는 가중치(w1, w2, w3)가 할당되어 있다. 특징 융합 모듈은 이러한 가중치 중 최대 가중치에 대응되는 융합 연산을 이용하여, 입력되는 특징값(a, b)을 융합하고, 융합된 특징값(c)을 출력한다. 예컨대, 제1가중치(w1)이 가장 크다면, 컴퓨팅 장치는 제1융합 연산(710)을 이용하여, 융합된 특징값(c)을 출력할 수 있다.

[0063] 한편 전술된 바와 같이, 융합 연산에 할당된 가중치는 학습을 통해 결정되며, 컴퓨팅 장치는 특징값을 이용해 타겟 태스크를 수행하는 멀티 모달 신경망의 손실값이 최소가 되도록, 가중치에 대한 학습을 수행할 수 있다. 그리고 학습 과정에서는 복수의 융합 연산이 모두 이용된다.

[0064] 즉, 컴퓨팅 장치는 복수의 융합 연산 각각을 이용하여, 제1 및 제2신경망의 특징 추출 블록 각각에서 출력되는 특징값을 융합한다. 이 때 전술된 바와 같이, 제1 및 제2신경망의 특징 추출 블록 각각에서 출력되는 특징값은 가중합되어 특징 융합 모듈로 입력되고, 가중합된 값이, 특징 융합 모듈로 입력되는 특징값(a, b)에 대응된다. 그리고 3개의 융합 연산(710, 720, 730)이 이용되는 경우, 3개의 융합 연산(710, 720, 730)으로부터 생성되는, 3개의 융합된 특징값이 출력된다.

[0065] 그리고 컴퓨팅 장치는 복수의 융합 연산 각각에 할당된 가중치를 이용하여, 융합된 특징값에 대한 가중합을 계산한다. 컴퓨팅 장치는 일실시예로서, [수학식 2]와 같이, 융합된 특징값에 대한 가중합(c)을 계산할 수 있다.

수학식 2

$$c = \sum_{f_i \in F} \frac{\exp(w_i)}{\sum_{f_i \in F} \exp(w_i)} f_i(a, b)$$

[0066]

[0067] 여기서, $f_i(a, b)$ 은 입력되는 특징값(a, b)에 대한 융합 연산을 나타내며, w_i 는 융합 연산에 할당된 가중치를 나타내며, i는 융합 연산의 인덱스를 나타내고, F는 융합 연산 집합을 나타낸다.

[0068] 그리고 컴퓨팅 장치는 [수학식 2]와 같이 가중합된 값을 이용하여, 가중치에 대한 학습을 수행한다. 전술된 바와 같이, DARTS 알고리즘이 학습에 이용될 수 있다.

[0070] 앞서 설명한 기술적 내용들은 다양한 컴퓨터 수단을 통하여 수행될 수 있는 프로그램 명령 형태로 구현되어 컴퓨터 판독 가능 매체에 기록될 수 있다. 상기 컴퓨터 판독 가능 매체는 프로그램 명령, 데이터 파일, 데이터 구조 등을 단독으로 또는 조합하여 포함할 수 있다. 상기 매체에 기록되는 프로그램 명령은 실시예들을 위하여

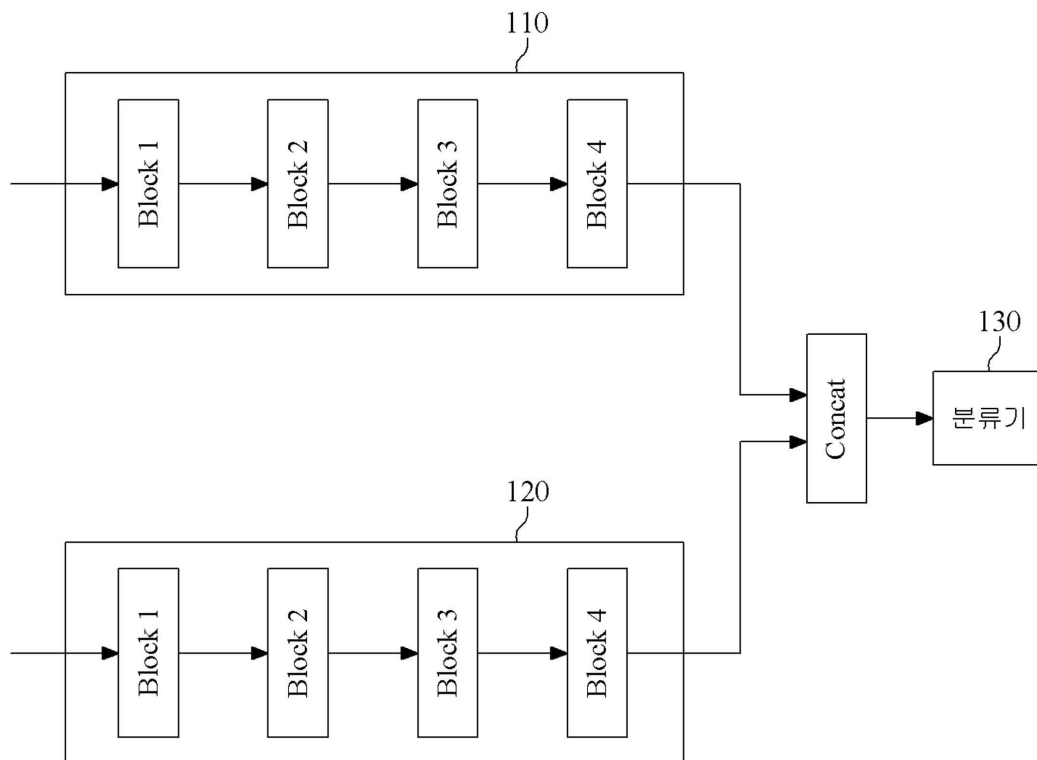
특별히 설계되고 구성된 것들이거나 컴퓨터 소프트웨어 당업자에게 공지되어 사용 가능한 것일 수도 있다. 컴퓨터 관독 가능 기록 매체의 예에는 하드 디스크, 플로피 디스크 및 자기 테이프와 같은 자기 매체(magnetic media), CD-ROM, DVD와 같은 광기록 매체(optical media), 플롭티컬 디스크(floptical disk)와 같은 자기-광 매체(magneto-optical media), 및 롬(ROM), 램(RAM), 플래시 메모리 등과 같은 프로그램 명령을 저장하고 수행하도록 특별히 구성된 하드웨어 장치가 포함된다. 프로그램 명령의 예에는 컴파일러에 의해 만들어지는 것과 같은 기계어 코드뿐만 아니라 인터프리터 등을 사용해서 컴퓨터에 의해서 실행될 수 있는 고급 언어 코드를 포함한다. 하드웨어 장치는 실시예들의 동작을 수행하기 위해 하나 이상의 소프트웨어 모듈로서 작동하도록 구성될 수 있으며, 그 역도 마찬가지이다.

[0072]

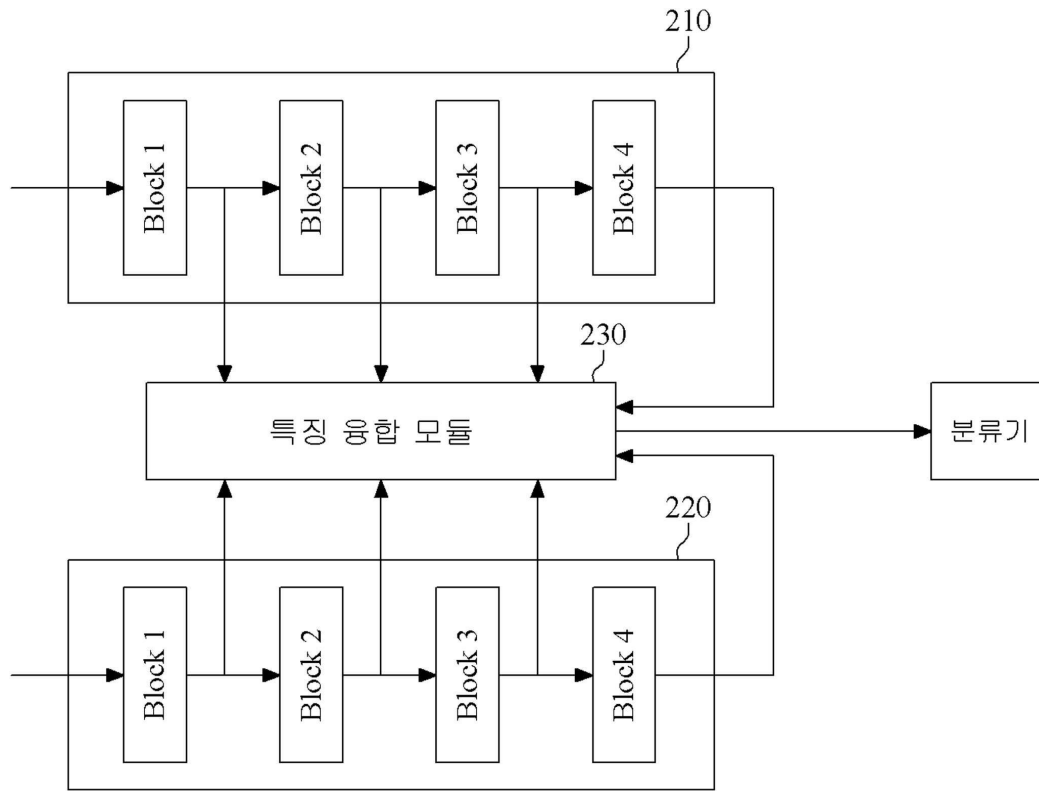
이상과 같이 본 발명에서는 구체적인 구성 요소 등과 같은 특정 사항들과 한정된 실시예 및 도면에 의해 설명되었으나 이는 본 발명의 보다 전반적인 이해를 돕기 위해서 제공된 것일 뿐, 본 발명은 상기의 실시예에 한정되는 것은 아니며, 본 발명이 속하는 분야에서 통상적인 지식을 가진 자라면 이러한 기재로부터 다양한 수정 및 변형이 가능하다. 따라서, 본 발명의 사상은 설명된 실시예에 국한되어 정해져서는 아니되며, 후술하는 특허청구범위뿐 아니라 이 특허청구범위와 균등하거나 등가적 변형이 있는 모든 것들은 본 발명 사상의 범주에 속한다고 할 것이다.

도면

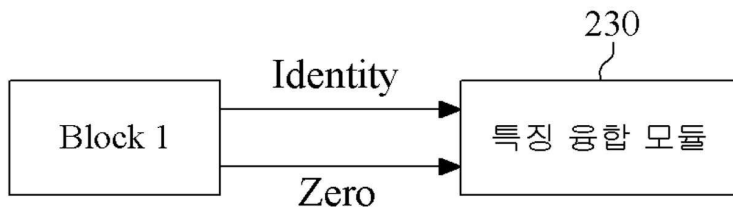
도면1



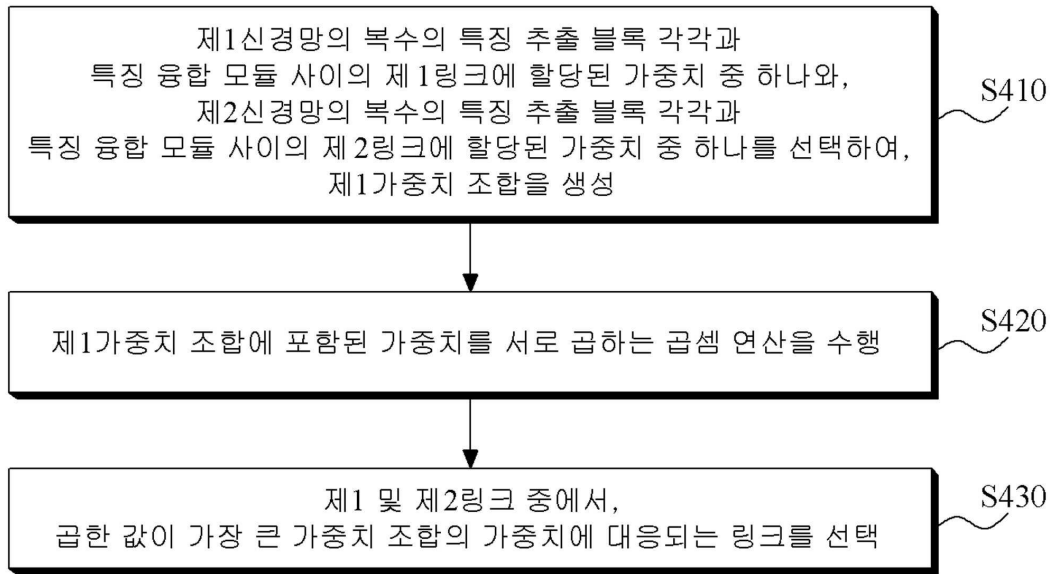
도면2



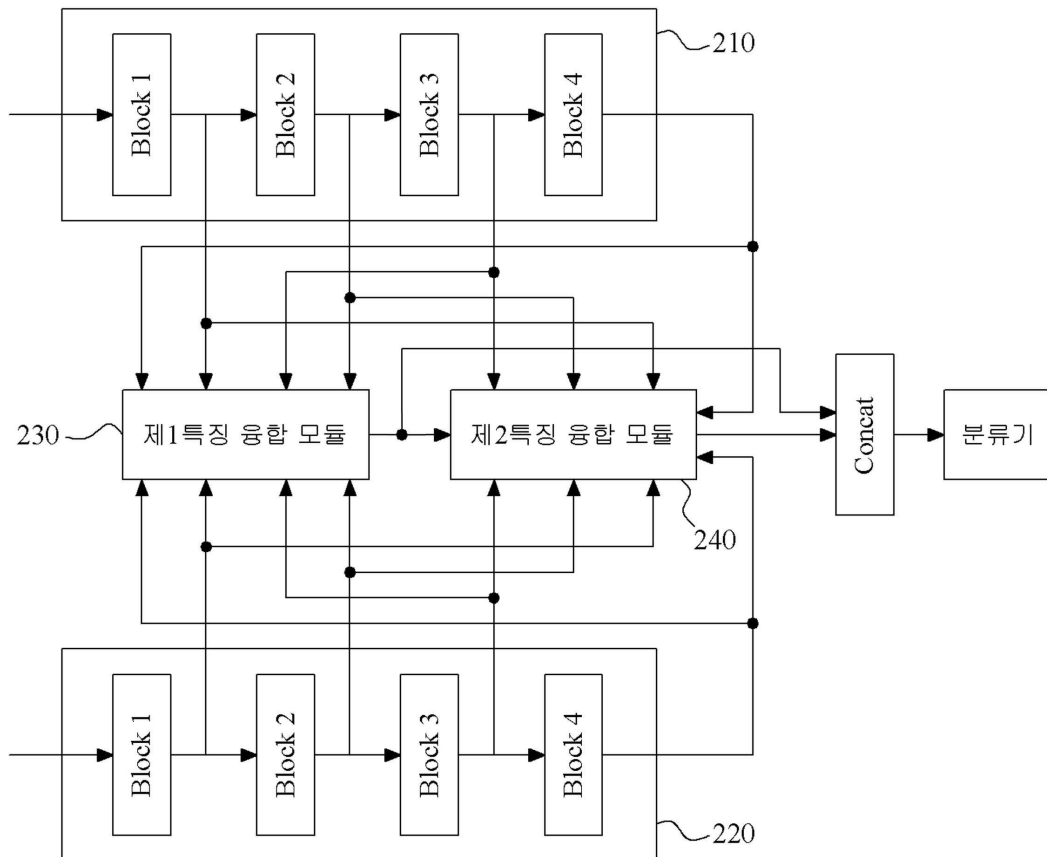
도면3



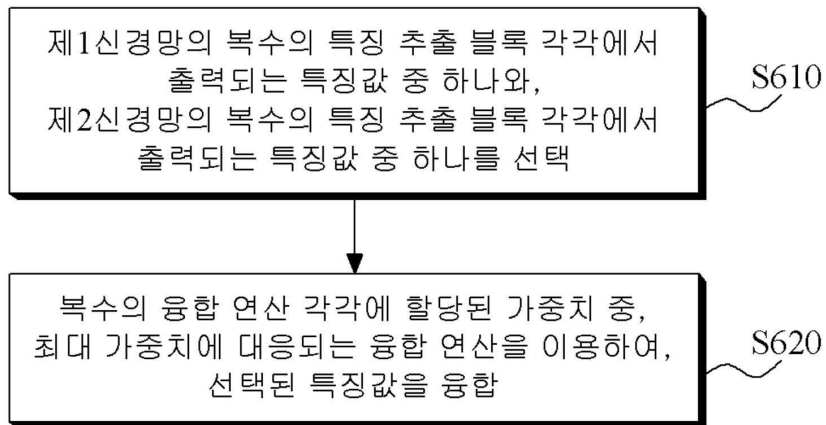
도면4



도면5



도면6



도면7

