



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2021-0121537
(43) 공개일자 2021년10월08일

(51) 국제특허분류(Int. Cl.)
G06N 3/08 (2006.01) G06N 3/04 (2006.01)
(52) CPC특허분류
G06N 3/08 (2013.01)
G06F 40/30 (2020.01)
(21) 출원번호 10-2020-0038369
(22) 출원일자 2020년03월30일
심사청구일자 2020년03월30일

(71) 출원인
고려대학교 산학협력단
서울특별시 성북구 안암로 145, 고려대학교 (안암동5가)
(72) 발명자
이상근
서울특별시 강남구 선릉로 221, 205동 1104호 (도곡동, 도곡테크아파트)
최병주
서울특별시 서초구 연남1길 19, 501호 (양재동)
박준형
서울특별시 동대문구 이문로1길 21, 5동 1102호 (회기동, 청량리신현대아파트)
(74) 대리인
특허법인엠에이피에스

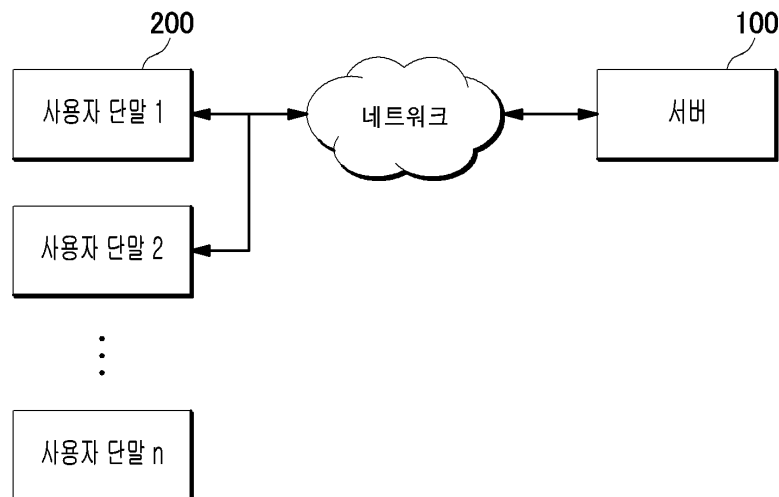
전체 청구항 수 : 총 17 항

(54) 발명의 명칭 텍스트 분류를 위한 적응형 콘볼루션 필터 생성 방법 및 서버

(57) 요약

본원의 일 측면에 따른 서버에 의해 수행되는 텍스트 분류를 위한 적응형 콘볼루션 필터 생성 방법은, 네트워크를 통해 수집된 원시 데이터에 전처리 프로세스를 수행하여 텍스트 데이터를 생성하는 단계; 상기 텍스트 데이터를 입력으로 필터 생성 모델을 통해 컨텍스트 벡터를 생성하는 단계; 및 상기 컨텍스트 벡터에 기초하여 상기 필터 생성 모델을 통해 콘볼루션 필터를 생성하는 단계;를 포함한다. 상기 필터 생성 모델은 상기 텍스트 데이터를 입력으로 인공신경망을 이용하여 적응형 콘볼루션 필터를 생성하는 것이다.

대표도 - 도1



(52) CPC특허분류

G06N 3/0454 (2013.01)

이 발명을 지원한 국가연구개발사업

과제고유번호	1711109835
과제번호	2018R1A2A1A05078380
부처명	과학기술정보통신부
과제관리(전문)기관명	한국연구재단
연구사업명	후속연구지원(도전)
연구과제명	온디바이스 텍스트 인공지능 개발
기 여 율	1/2
과제수행기관명	고려대학교
연구기간	2020.03.01 ~ 2021.02.28

이 발명을 지원한 국가연구개발사업

과제고유번호	1711116242
과제번호	2016-0-00464-005
부처명	과학기술정보통신부
과제관리(전문)기관명	정보통신기획평가원
연구사업명	연구지원
연구과제명	지능의 축적·확장·고정밀화를 위한 지능정보처리 SW 핵심기술 개발 및 글로벌 인
력양성	
기 여 율	1/2
과제수행기관명	고려대학교
연구기간	2016.07.01 ~ 2021.12.31

명세서

청구범위

청구항 1

서버에 의해 수행되는 텍스트 분류를 위한 적응형 콘볼루션 필터 생성 방법에 있어서,

(a) 네트워크를 통해 수집된 원시 데이터에 전처리 프로세스를 수행하여 텍스트 데이터를 생성하는 단계;

(b) 상기 텍스트 데이터를 입력으로 필터 생성 모델을 통해 컨텍스트 벡터를 생성하는 단계; 및

(c) 상기 컨텍스트 벡터에 기초하여 상기 필터 생성 모델을 통해 콘볼루션 필터를 생성하는 단계;를 포함하고,

상기 필터 생성 모델은 상기 텍스트 데이터를 입력으로 인공신경망을 이용하여 적응형 콘볼루션 필터를 생성하는 것인,

텍스트 분류를 위한 적응형 콘볼루션 필터 생성 방법.

청구항 2

제 1 항에 있어서,

상기 (b)단계는 양방향 GRU (Gated Recurrent Unit)를 통해 컨텍스트 벡터를 생성하는 것인,

텍스트 분류를 위한 적응형 콘볼루션 필터 생성 방법.

청구항 3

제 2 항에 있어서,

상기 양방향 GRU (Gated Recurrent Unit)는 상기 텍스트 데이터를 상관관계를 내포하는 벡터들의 집합으로 변형하고, 해당 시점에서 분류해야할 텍스트와 연관이 있는 텍스트 부분에 집중하는 어텐션 기법을 적용하여 가중치를 계산하여, 컨텍스트 벡터를 생성하는 것인,

텍스트 분류를 위한 적응형 콘볼루션 필터 생성 방법.

청구항 4

제 1 항에 있어서,

상기 (c)단계는 해싱 기법을 통해 콘볼루션 필터를 생성하는 단계를 더 포함하는 것인,

텍스트 분류를 위한 적응형 콘볼루션 필터 생성 방법.

청구항 5

제 4 항에 있어서,

상기 해싱 기법은 컴포넌트 필터를 포함하는 공유 행렬을 생성하고, 기설정된 해시 함수를 사용하여 후보 컴포넌트 필터를 선택하고, 상기 후보 컴포넌트 필터에 대해 가중치에 따라 선형 조합하는 것인,

텍스트 분류를 위한 적응형 콘볼루션 필터 생성 방법.

청구항 6

제 5 항에 있어서,

상기 가중치는 컨텍스트 벡터를 입력으로 받는 단층 신경망의 출력값인 것인,

텍스트 분류를 위한 적응형 콘볼루션 필터 생성 방법.

청구항 7

제 1 항에 있어서,

상기 (c) 단계 이후 생성된 상기 콘볼루션 필터를 기설정된 콘볼루션 신경망에 적용하는 단계를 더 포함하는 것인,

텍스트 분류를 위한 적응형 콘볼루션 필터 생성 방법.

청구항 8

텍스트 분류를 위한 적응형 콘볼루션 필터 생성 서버에 있어서,

텍스트 분류를 위한 적응형 콘볼루션 필터 생성 프로그램이 저장된 메모리;

상기 메모리에 저장된 프로그램을 실행하는 프로세서를 포함하며,

상기 프로세서는 상기 프로그램의 실행에 의해, 네트워크를 통해 수집된 원시 데이터에 전처리 프로세스를 수행하여 텍스트 데이터를 생성하고, 상기 텍스트 데이터를 입력으로 필터 생성 모델을 통해 컨텍스트 벡터를 생성하고, 상기 컨텍스트 벡터에 기초하여 상기 필터 생성 모델을 통해 콘볼루션 필터를 생성하고,

상기 필터 생성 모델은 상기 텍스트 데이터를 입력으로 인공신경망을 이용하여 적응형 콘볼루션 필터를 생성하는 것인,

텍스트 분류를 위한 적응형 콘볼루션 필터 생성 서버.

청구항 9

제 8 항에 있어서,

상기 필터 생성 모델은 양방향 GRU (Gated Recurrent Unit)를 포함하고,

상기 프로세서는 상기 양방향 GRU를 통해 컨텍스트 벡터를 생성하는 것이고,

상기 양방향 GRU (Gated Recurrent Unit)는 상기 텍스트 데이터를 상관관계를 내포하는 벡터들의 집합으로 변형하고, 해당 시점에서 분류해야 할 텍스트와 연관이 있는 텍스트 부분에 집중하는 어텐션 기법을 적용하여 가중치를 계산하여, 컨텍스트 벡터를 생성하는 것인,

텍스트 분류를 위한 적응형 콘볼루션 필터 생성 서버.

청구항 10

제 8 항에 있어서,

상기 프로세서는 해싱 기법을 통해 콘볼루션 필터를 생성하는 것이고,

상기 해싱 기법은 컴포넌트 필터를 포함하는 공유 행렬을 생성하고, 기설정된 해시 함수를 사용하여 후보 컴포넌트 필터를 선택하고, 상기 후보 컴포넌트 필터에 대해 가중치에 따라 선형 조합하는 것이고,

상기 가중치는 컨텍스트 벡터를 입력으로 받는 단층 신경망의 출력값인 것인,

텍스트 분류를 위한 적응형 콘볼루션 필터 생성 서버.

청구항 11

적응형 콘볼루션 필터를 이용한 텍스트 분류 서버에 있어서,

적응형 콘볼루션 필터를 이용한 텍스트 분류 프로그램이 저장된 메모리;

상기 메모리에 저장된 프로그램을 실행하는 프로세서를 포함하며,

상기 프로세서는 상기 프로그램의 실행에 의해, 네트워크를 통해 수집된 원시 데이터에 전처리 프로세스를 수행하여 텍스트 데이터를 생성하고, 상기 텍스트 데이터를 입력으로 필터 생성 모델을 통해 컨텍스트 벡터를 생성하고, 상기 컨텍스트 벡터에 기초하여 상기 필터 생성 모델을 통해 콘볼루션 필터를 생성하고, 상기 콘볼루션 필터를 기설정된 콘볼루션 신경망에 적용하고, 상기 콘볼루션 신경망을 통해 분류된 텍스트를 사용자 단말에 제공하는 것이고,

상기 필터 생성 모델은 상기 텍스트 데이터를 입력으로 인공신경망을 이용하여 적응형 콘볼루션 필터를 생성하는 것인,

적응형 콘볼루션 필터를 이용한 텍스트 분류 서버.

청구항 12

제 11 항에 있어서,

상기 필터 생성 모델은 양방향 GRU (Gated Recurrent Unit)를 포함하고,

상기 프로세서는 상기 양방향 GRU를 통해 컨텍스트 벡터를 생성하는 것인,

적응형 콘볼루션 필터를 이용한 텍스트 분류 서버.

청구항 13

제 12 항에 있어서,

상기 양방향 GRU (Gated Recurrent Unit)는 상기 텍스트 데이터를 상관관계를 내포하는 벡터들의 집합으로 변형하고, 해당 시점에서 분류해야할 텍스트와 연관이 있는 텍스트 부분에 집중하는 어텐션 기법을 적용하여 가중치를 계산하여, 컨텍스트 벡터를 생성하는 것인,

적응형 콘볼루션 필터를 이용한 텍스트 분류 서버.

청구항 14

제 11 항에 있어서,

상기 프로세서는 해싱 기법을 통해 콘볼루션 필터를 생성하는 것인,

적응형 콘볼루션 필터를 이용한 텍스트 분류 서버.

청구항 15

제 14 항에 있어서,

상기 해싱 기법은 킴포넨트 필터를 포함하는 공유 행렬을 생성하고, 기설정된 해시 함수를 사용하여 후보 킴포넨트 필터를 선택하고, 상기 후보 킴포넨트 필터에 대해 가중치에 따라 선형 조합하는 것인,

적응형 콘볼루션 필터를 이용한 텍스트 분류 서버.

청구항 16

제 15 항에 있어서,

상기 가중치는 컨텍스트 벡터를 입력으로 받는 단층 신경망의 출력값인 것인,

적응형 콘볼루션 필터를 이용한 텍스트 분류 서버.

청구항 17

제 1 항에 따르는 텍스트 분류를 위한 적응형 콘볼루션 필터 생성 방법을 수행하기 위한 컴퓨터 프로그램이 기록된 비일시적 컴퓨터 판독가능 기록매체.

발명의 설명

기술 분야

본 발명은 텍스트 분류를 위한 적응형 콘볼루션 필터 생성 방법에 관한 것으로, 보다 상세하게는 텍스트의 의미를 파악하는 데에 필수적인 콘볼루션 필터들만을 생성하고, 콘볼루션 연산을 수행함으로써 텍스트의 주제를 효과적으로 분류할 수 있도록 하는 방법에 관한 것이다.

배경 기술

[0001]

[0002] 컴퓨터비전시스템에서 콘볼루션 신경망 기술이 효과적으로 사용됨에 따라 이를 활용한 자연어처리 기술에 관한 연구도 급속도로 진행되고 있다. 그 중에서도 텍스트 분류 분야에서는 콘볼루션 신경망 기술이 집중적으로 연구되고 있다. 콘볼루션 신경망은 콘볼루션 필터를 이용하여 자연어에 내재되어 있는 특징을 추출하고 이를 바탕으로 텍스트의 주제를 분류한다. 그러나, 이러한 방법은 텍스트에 내재되어 있는 사전 정보를 활용하지 않은 상태로 항상 같은 필터 값만을 활용한다는 특징이 있다. 이 경우 주어진 텍스트의 의미를 분석하는데 도움이 되지 않는 필터들이 모델에 적용될 수 있다는 문제가 생기고, 이런 문제로 인해 모델이 텍스트의 미묘한 의미를 파악하는 데에 집중하지 못하게 되는 부정적인 효과가 나타날 수 있다.

[0003] 컴퓨터비전시스템에서는 여러 입력 이미지를 효과적으로 분석하기 위해 콘볼루션 필터를 생성하는 기술이 활발하게 연구되어 왔다. 이러한 방법은 하나의 입력 이미지를 처리하고, 그 처리 결과를 바탕으로 다른 이미지를 분석하는 콘볼루션 신경망의 콘볼루션 필터를 생성하고, 생성된 필터를 바탕으로 그 이미지를 분석하는 방법을 채택한다. 다만 이러한 컴퓨터비전시스템에서 활용되는 콘볼루션 필터 생성 기법을 텍스트 분류에 맞도록 변형 및 개선하는 연구는 부족한 실정이다.

선행기술문헌

특허문헌

[0004] (특허문헌 0001) 대한민국 공개특허공보 제 10-2019-0026529 호 (발명의 명칭: 문서 분류 시스템 및 그 방법)

발명의 내용

해결하려는 과제

[0005] 본 발명은 전술한 문제점을 해결하기 위한 것으로, 텍스트의 의미를 파악하는데 필수적인 콘볼루션 필터를 생성하고, 콘볼루션 필터를 보다 적은 가중치를 통해 생성하고, 생성된 콘볼루션 필터를 콘볼루션 신경망에 적용하는 방법을 제공하는 것을 기술적 과제로 한다.

[0006] 다만, 본 실시예가 이루고자 하는 기술적 과제는 상기된 바와 같은 기술적 과제로 한정되지 않으며, 또 다른 기술적 과제들이 존재할 수 있다.

과제의 해결 수단

[0007] 상술한 기술적 과제를 해결하기 위한 기술적 수단으로서, 본 개시의 제 1측면에 따른 서버에 의해 수행되는 텍스트 분류를 위한 적응형 콘볼루션 필터 생성 방법은, 네트워크를 통해 수집된 원시 데이터에 전처리 프로세스를 수행하여 텍스트 데이터를 생성하는 단계; 상기 텍스트 데이터를 입력으로 필터 생성 모델을 통해 컨텍스트 벡터를 생성하는 단계; 및 상기 컨텍스트 벡터에 기초하여 상기 필터 생성 모델을 통해 콘볼루션 필터를 생성하는 단계;를 포함한다. 상기 필터 생성 모델은 상기 텍스트 데이터를 입력으로 인공신경망을 이용하여 적응형 콘볼루션 필터를 생성하는 것이다.

[0008] 또한, 본 개시의 제 2 측면에 따른 텍스트 분류를 위한 적응형 콘볼루션 필터 생성 서버는 텍스트 분류를 위한 적응형 콘볼루션 필터 생성 프로그램이 저장된 메모리; 상기 메모리에 저장된 프로그램을 실행하는 프로세서를 포함한다. 상기 프로세서는 상기 프로그램의 실행에 의해, 네트워크를 통해 수집된 원시 데이터에 전처리 프로세스를 수행하여 텍스트 데이터를 생성하고, 상기 텍스트 데이터를 입력으로 필터 생성 모델을 통해 컨텍스트 벡터를 생성하고, 상기 컨텍스트 벡터에 기초하여 상기 필터 생성 모델을 통해 콘볼루션 필터를 생성한다. 상기 필터 생성 모델은 상기 텍스트 데이터를 입력으로 인공신경망을 이용하여 적응형 콘볼루션 필터를 생성하는 것이다.

[0009] 또한, 본 개시의 제 3 측면에 따른 적응형 콘볼루션 필터를 이용한 텍스트 분류 서버는 적응형 콘볼루션 필터를 이용한 텍스트 분류 프로그램이 저장된 메모리; 상기 메모리에 저장된 프로그램을 실행하는 프로세서를 포함한다. 상기 프로세서는 상기 프로그램의 실행에 의해, 네트워크를 통해 수집된 원시 데이터에 전처리 프로세스를 수행하여 텍스트 데이터를 생성하고, 상기 텍스트 데이터를 입력으로 필터 생성 모델을 통해 컨텍스트 벡터를 생성하고, 상기 컨텍스트 벡터에 기초하여 상기 필터 생성 모델을 통해 콘볼루션 필터를 생성하고, 상기 콘볼루션 필터를 기설정된 콘볼루션 신경망에 적용하고, 상기 콘볼루션 신경망을 통해 분류된 텍스트를 사용자 단말에

제공하는 것이다. 상기 필터 생성 모델은 상기 텍스트 데이터를 입력으로 인공지능망을 이용하여 적응형 콘볼루션 필터를 생성하는 것이다.

발명의 효과

- [0010] 전술한 본원의 과제 해결 수단 중 어느 하나에 의하면, 동적으로 콘볼루션 필터를 생성하며, 분류할 텍스트에 적합한 콘볼루션 필터들을 활용하므로 텍스트의 의미를 보다 명확하게 파악할 수 있다. 또한 이를 통해 적은 수의 콘볼루션 필터로도 효과적으로 텍스트를 분류할 수 있는 효과가 있다.
- [0011] 본 발명은 해싱 기법을 이용하여 콘볼루션 필터 생산 시 상대적으로 더 작은 가중치를 요구하며, 콘볼루션 필터를 생성하기 위한 신경망의 크기를 효과적으로 압축할 수 있다.

도면의 간단한 설명

- [0012] 도 1은 본 발명의 일 실시예에 따른 텍스트 분류를 위한 적응형 콘볼루션 필터 생성 시스템의 구성을 도시한 블록도이다.
- 도 2는 본 발명의 일 실시예에 따른 텍스트 분류를 위한 적응형 콘볼루션 필터 생성 서버의 구성을 도시한 블록도이다.
- 도 3은 본 발명의 일 실시예에 따른 텍스트 분류를 위한 적응형 콘볼루션 필터 생성 방법에서 필터 생성 신경망을 설명하는 도면이다.
- 도 4는 본 발명의 일 실시예에 따른 텍스트 분류를 위한 적응형 콘볼루션 필터 생성 방법을 설명하는 도면이다.
- 도 5는 본 발명의 일 실시예에 따른 텍스트 분류를 위한 적응형 콘볼루션 필터 생성 방법에서 해싱 기법을 설명하는 도면이다.
- 도 6은 본 발명의 일 실시예에 따른 텍스트 분류를 위한 적응형 콘볼루션 필터 생성 방법에서 생성된 적응형 콘볼루션 필터의 적용 결과를 설명하는 도면이다.
- 도 7은 본 발명의 일 실시예에 따른 텍스트 분류를 위한 적응형 콘볼루션 필터 생성 방법에서 생성된 적응형 콘볼루션 필터의 적용 결과를 설명하는 표이다.
- 도 8은 본 발명의 일 실시예에 따른 텍스트 분류를 위한 적응형 콘볼루션 필터 생성 방법에서 생성된 적응형 콘볼루션 필터의 적용 결과를 설명하는 표이다.
- 도 9는 본 발명의 일 실시예에 따른 텍스트 분류를 위한 적응형 콘볼루션 필터 생성 방법의 진행 단계를 나타낸 흐름도이다.

발명을 실시하기 위한 구체적인 내용

- [0013] 아래에서는 첨부한 도면을 참조하여 본원이 속하는 기술 분야에서 통상의 지식을 가진 자가 용이하게 실시할 수 있도록 본원의 실시예를 상세히 설명한다. 그러나 본원은 여러 가지 상이한 형태로 구현될 수 있으며 여기에서 설명하는 실시예에 한정되지 않는다. 그리고 도면에서 본원을 명확하게 설명하기 위해서 설명과 관계없는 부분은 생략하였으며, 명세서 전체를 통하여 유사한 부분에 대해서는 유사한 도면 부호를 붙였다.
- [0014] 본원 명세서 전체에서, 어떤 부분이 다른 부분과 "연결"되어 있다고 할 때, 이는 "직접적으로 연결"되어 있는 경우뿐 아니라, 그 중간에 다른 소자를 사이에 두고 "전기적으로 연결"되어 있는 경우도 포함한다.
- [0015] 본원 명세서 전체에서, 어떤 부재가 다른 부재 “상에” 위치하고 있다고 할 때, 이는 어떤 부재가 다른 부재에 접해 있는 경우뿐 아니라 두 부재 사이에 또 다른 부재가 존재하는 경우도 포함한다.
- [0016] 이하 첨부된 도면을 참고하여 본 발명의 일 실시예를 상세히 설명하기로 한다.
- [0017] 도 1은 본 발명의 일 실시예에 따른 텍스트 분류를 위한 적응형 콘볼루션 필터 생성 시스템의 구성을 도시한 블록도이다.
- [0018] 텍스트 분류를 위한 적응형 콘볼루션 필터 생성 방법은 서버(100) 또는 사용자 단말(200)에서 실시될 수 있다. 텍스트 분류를 위한 적응형 콘볼루션 필터 생성 방법에서 원시 데이터 수집은 사용자 단말(200)에서 이루어질 수 있다. 콘볼루션 필터 생성은 서버(100)에서 이루어지고, 생성된 콘볼루션 필터가 사용자 단말(200)에 제공

될 수 있다. 복수의 사용자는 각각 자신의 사용자 단말(200)을 통해 텍스트 분류를 위한 적응형 콘볼루션 필터 생성 방법을 실시할 수 있다.

- [0019] 적응형 콘볼루션 필터를 이용한 텍스트 분류 방법 또한 동일한 구성의 서버(100) 또는 사용자 단말(200)에서 실시될 수 있다.
- [0020] 도 2는 본 발명의 일 실시예에 따른 텍스트 분류를 위한 적응형 콘볼루션 필터 생성 서버의 구성을 도시한 블록도이다.
- [0021] 도시된 바와 같이 텍스트 분류를 위한 적응형 콘볼루션 필터 생성 서버(100)는 통신 모듈(110), 메모리(120), 프로세서(130), 데이터베이스(140) 및 입력모듈(150)을 포함할 수 있다.
- [0022] 통신모듈(110)은 접속된 사용자 단말(200)과 데이터를 송수신할 수 있다. 통신모듈(110)은 다른 네트워크 장치와 유무선 연결을 통해 제어 신호 또는 데이터 신호와 같은 신호를 송수신하기 위해 필요한 하드웨어 및 소프트웨어를 포함하는 장치일 수 있다.
- [0023] 메모리(120)에는 텍스트 분류를 위한 적응형 콘볼루션 필터 생성 프로그램이 저장된다. 텍스트 분류를 위한 적응형 콘볼루션 필터 생성 프로그램은 네트워크를 통해 수집된 원시 데이터에 전처리 프로세스를 수행하여 텍스트 데이터를 생성하고, 텍스트 데이터를 입력으로 필터 생성 모델을 통해 컨텍스트 벡터를 생성하고, 컨텍스트 벡터에 기초하여 필터 생성 모델을 통해 콘볼루션 필터를 생성하고, 콘볼루션 필터를 기설정된 콘볼루션 신경망에 적용하고, 콘볼루션 신경망을 통해 분류된 텍스트를 사용자 단말에 제공한다.
- [0024] 이러한 메모리(120)에는 텍스트 분류를 위한 적응형 콘볼루션 필터 생성 서버(100)의 구동을 위한 운영 체제나 텍스트 분류를 위한 적응형 콘볼루션 필터 생성 프로그램의 실행 과정에서 발생하는 여러 종류가 데이터가 저장된다.
- [0025] 이때, 메모리(120)는 전원이 공급되지 않아도 저장된 정보를 계속 유지하는 비휘발성 저장장치 및 저장된 정보를 유지하기 위하여 전력이 필요한 휘발성 저장장치를 통칭하는 것이다.
- [0026] 또한, 메모리(120)는 프로세서(130)가 처리하는 데이터를 일시적 또는 영구적으로 저장하는 기능을 수행할 수 있다. 여기서, 메모리(120)는 저장된 정보를 유지하기 위하여 전력이 필요한 휘발성 저장장치 외에 자기 저장 매체(magnetic storage media) 또는 플래시 저장 매체(flash storage media)를 포함할 수 있으나, 본 발명의 범위가 이에 한정되는 것은 아니다.
- [0027] 프로세서(130)는 메모리(120)에 저장된 프로그램을 실행하되, 텍스트 분류를 위한 적응형 콘볼루션 필터 생성 프로그램의 실행에 따르는 전체 과정을 제어한다. 프로세서(130)가 수행하는 각각의 동작에 대해서는 추후 보다 상세히 살펴보기로 한다.
- [0028] 이러한 프로세서(130)는 데이터를 처리할 수 있는 모든 종류의 장치를 포함할 수 있다. 예를 들어 프로그램 내에 포함된 코드 또는 명령으로 표현된 기능을 수행하기 위해 물리적으로 구조화된 회로를 갖는, 하드웨어에 내장된 데이터 처리 장치를 의미할 수 있다. 이와 같이 하드웨어에 내장된 데이터 처리 장치의 일 예로써, 마이크로프로세서(microprocessor), 중앙처리장치(central processing unit: CPU), 프로세서 코어(processor core), 멀티프로세서(multiprocessor), ASIC(application-specific integrated circuit), FPGA(field programmable gate array) 등의 처리 장치를 망라할 수 있으나, 본 발명의 범위가 이에 한정되는 것은 아니다.
- [0029] 데이터베이스(140)는 프로세서(130)의 제어에 따라, 텍스트 분류를 위한 적응형 콘볼루션 필터 생성 시스템에 필요한 데이터를 저장 또는 제공한다. 이러한 데이터베이스(140)는 메모리(120)와는 별도의 구성 요소로서 포함되거나, 또는 메모리(120)의 일부 영역에 구축될 수도 있다.
- [0030] 도 3은 본 발명의 일 실시예에 따른 텍스트 분류를 위한 적응형 콘볼루션 필터 생성 방법에서 필터 생성 모델에 포함된 필터 생성 신경망을 설명하는 도면이다.
- [0031] 텍스트 데이터, 또는 이전 레이어의 출력값은 필터 생성 신경망에서 사용하는 레이어의 입력이 될 수 있다. 예를 들어 레이어 1의 경우 텍스트 데이터를 입력으로 필터 생성 신경망을 통해 콘볼루션 필터 a를 생성하고, 레이어 2의 경우 레이어 1의 출력을 입력으로 필터 생성 신경망을 통해 콘볼루션 필터 b를 생성한다. 즉 각 레이어의 입력을 통해 콘볼루션 필터가 생성되고, 생성된 콘볼루션 필터는 각 레이어의 입력에 대한 콘볼루션 연산에 적용된다. 적응형 콘볼루션 필터 생성 방법은 필터 생성 신경망을 통해 동적으로 생성된 필터를 활용하여 텍스트를 효과적으로 분류할 수 있다.

[0032] 필터 생성 신경망의 구성요소에 대해서 컨텍스트 벡터를 생성하는 방법은 도4, 해싱 기법을 통해 적응형 콘볼루션 필터를 생성하는 방법은 도 5를 통해 상세히 설명한다.

[0033] 도 4는 본 발명의 일 실시예에 따른 텍스트 분류를 위한 적응형 콘볼루션 필터 생성 방법을 설명하는 도면이다.

[0034] 도 4는 콘볼루션 필터의 생성 과정을 도시한 것이다. 먼저 텍스트를 벡터 형태로 변환한 입력데이터 $x(x_1, \dots, x_m)$ 는 양방향 GRU를 거쳐 입력 간 상관관계를 내포하는 벡터들의 집합 $h(h_1, \dots, h_m)$ 로 변형된다. q 는 학습 가능한 쿼리 벡터이며, 이후 어텐션 기법을 통해 이들을 하나의 컨텍스트 벡터 c 로 축약한다. 어텐션 기법은 디코더에서 출력 단어를 예측하는 매 시점마다, 인코더에서의 전체 입력 문장을 다시 한 번 참고한다는 점을 기본 아이디어로 하며, 전체 입력 문장을 전부 다 동일한 비율로 참고하는 것이 아니라, 해당 시점에서 예측해야 할 단어와 연관이 있는 입력 단어 부분을 좀 더 집중(attention)해서 보는 특징이 있다.

[0035] 컨텍스트 벡터는 각 레이어의 입력값에 내포된 의미를 파악하고 축약한 것을 의미한다. 이는 하기 수식 1을 통해 계산할 수 있다.

[0036] [수식 1]

$$\begin{aligned}\vec{h}_t &= \overrightarrow{GRU}(x_t, \vec{h}_{t-1}), \\ \overleftarrow{h}_t &= \overleftarrow{GRU}(x_t, \overleftarrow{h}_{t+1}), \\ h_t &= [\vec{h}_t; \overleftarrow{h}_t] \\ a_j &= \frac{\exp(q^\top h_j)}{\sum_{k=1}^m \exp(q^\top h_k)} \\ c &= \sum_{j=1}^m a_j h_j\end{aligned}$$

[0037]

[0038] 상기 수식 1에서 레이어의 입력은 총 m 개의 벡터로 구성되어 있다. 이 중 t 번째 벡터를 x_t 라 할 때, x_t 는 양방향 GRU를 거쳐 다른 입력과 상관관계를 내포하는 h_t 라는 벡터를 형성한다. 모든 m 개의 입력 벡터가 양방향 GRU를 통해 처리되면 쿼리 벡터 q 와의 어텐션 기법을 적용하여 가중치 값 a_j 를 계산한다. 이때 q 는 학습 가능한 가중치 벡터이고 a_j 는 q 와 h_j 사이의 내적에 소프트맥스 함수를 적용하여 계산된다. 이후 총 m 개의 벡터와 가중치에 대해 각각의 벡터 h_j 에 가중치 값 a_j 를 곱하여 합산함으로써 컨텍스트 벡터 c 를 생성한다.

[0039] 생성된 컨텍스트 벡터는 두 가지 다른 방법을 통해 콘볼루션 필터를 생성한다. 첫 번째 방법은 컨텍스트 벡터를 하나의 단일층 신경망의 입력으로 전달하여 이 신경망의 출력값을 콘볼루션 필터로 사용하는 것이다. 이는 하기 수식 2를 통해 계산할 수 있다.

[0040] [수식 2]

$$f_i = W_i c$$

[0041]

[0042] 수식 2에서 컨텍스트 벡터 c 와 가중치 행렬 W_i 사이의 행렬곱 연산을 통해 필터 f_i 를 계산한다. 그러나 이러한 방법은 많은 가중치를 요구한다는 단점이 있다. 필터를 생성하기 위한 가중치는 컨텍스트 벡터의 크기와, 생성해야 할 콘볼루션 필터의 가중치 수에 정비례하는데, 생성해야 할 필터의 개수가 많을 경우 메모리에 신경망이 탑재되지 않는 문제가 발생할 가능성이 있다.

[0043] 도 5는 본 발명의 일 실시예에 따른 텍스트 분류를 위한 적응형 콘볼루션 필터 생성 방법에서 해싱 기법을 설명하는 도면이다.

[0044] 이러한 문제를 해결하기 위해 해싱 기법을 통해 콘볼루션 필터를 생성하는 방법을 사용한다. 도 5는 해싱 기법을 활용해 컨텍스트 벡터를 통해 콘볼루션 필터를 생성하는 방법을 도시화한 것이다.

[0045] 우선 모든 필터들이 공유하는 공유 행렬을 생성한다. 여기서 공유 행렬은 매 입력마다 동적으로 생성되는 것이

아니라 일반적인 가중치 행렬과 같이 정적인 행렬이다. 생성되어야 할 필터는 각기 다른 해싱 함수에 의해 해싱된다. 해싱 함수의 출력값을 인덱스로 활용해 공유 행렬에서 후보 벡터를 추출한다. 추출된 후보 벡터들은 선형 결합되어 하나의 벡터를 이루는데, 이것을 컨볼루션 필터로 사용한다. 이 때, 컨텍스트 벡터를 입력으로 받는 단층 신경망의 출력값을 추출 벡터들의 선형 조합을 위한 가중치로 사용한다.

[0046] 이 때, f_1 은 첫번째 레이어의 필터이며 H_n 은 해싱함수이며, 해싱함수의 파라미터에는 필터의 ID가 입력된다. 해싱함수를 통해 공유 행렬에서 n 개의 필터를 선택한다. 선택된 n 개의 후보벡터 b 는 가중치합 연산을 통해 컨볼루션 필터 a 를 생성한다. 이때 사용되는 중요도 값 p_n 은 단층 신경망에서 컨텍스트 벡터 c 와 가중치벡터의 내적을 통해 계산된다.

[0047] 이러한 일련의 과정은 하기 수식 3을 통해 계산할 수 있다.

[0048] [수식 3]

$$f_i = \sum_{j=1}^n p_{i,j} \mathcal{H}_j(\hat{f}_i)$$

$$p_{i,j} = w_{i,j}^T c$$

[0050] 수식 3에서 f_i 는 i 번째 필터를 의미하며 $\hat{f}_i$ 는 필터 f_i 의 ID를 의미한다. 총 n 개의 해싱 함수를 이용해 n 개의 후보 벡터(b)를 선택하고 이를 가중합하여 필터 f_i 를 생성한다. 이때 j 번째 해싱 함수 H_j 가 선택한 후보 벡터는 가중치 벡터 $w_{i,j}$ 와 컨텍스트 벡터 c 사이의 내적을 통해 계산된 중요도 값 $p_{i,j}$ 와 곱해져 합산된다. 즉 해싱 기법을 사용하면 공유 행렬의 후보 필터를 조합하는 방식을 입력에 따라 동적으로 결정하여 필터(a)를 생성하게 된다.

[0051] 도 6은 본 발명의 일 실시예에 따른 텍스트 분류를 위한 적응형 컨볼루션 필터 생성 방법에서 생성된 적응형 컨볼루션 필터의 적용 결과를 설명하는 도면이다.

[0052] 도 6은 적응형 컨볼루션을 일반 컨볼루션 신경망에 적용했을 경우 생성되는 컨볼루션 필터를 시각화 한 것이다. 일반 컨볼루션 신경망은 컨볼루션 필터가 공간 내 다양한 공간에 분포되어 있는 것을 확인할 수 있다. 반면, 적응형 컨볼루션이 적용되었을 경우, 생성된 컨볼루션 필터는 적용된 입력 텍스트에 따라 특정한 위치에 밀집되어 생성되었다. 이것은 생성된 컨볼루션 필터가 입력 텍스트의 의미를 보다 효율적으로 파악할 수 있는 위치에 집중적으로 생성된다는 것을 의미한다.

[0053] 적응형 컨볼루션은 현존하는 모든 컨볼루션 신경망에 적용할 수 있다. 이는 기존 컨볼루션 신경망의 모든 레이어에 필터 생성 기법을 적용하여, 레이어마다 필터를 생성하고 생성된 필터를 통해 컨볼루션을 적용하는 방식으로 작동한다.

[0054] 도 7은 적응형 컨볼루션을 기존 컨볼루션 신경망에 적용했을 때 얻을 수 있는 성능 향상 폭에 관한 것이다. 실험에 사용된 컨볼루션 신경망은 CNN, DPCNN, DenseCNN이다. CNN은 3, 4, 5의 윈도우 크기를 설정하였고, 각각 100개의 컨볼루션 필터를 사용하였다. DPCNN은 11개의 레이어를 사용하였고, 각 레이어는 윈도우 크기를 3으로 하는 150개의 컨볼루션 필터를 사용하였다. DenseCNN은 윈도우 크기를 3으로 하는 필터를 75개 사용하였으며, 레이어는 6개를 사용하였다. 실험을 위해 Yelp 데이터를 사용하였다. 적응형 컨볼루션은 세 가지 컨볼루션 신경망에 대해 모두 성능 향상을 가져왔다. 최대 성능 향상폭은 2.6 퍼센티지 포인트인 것을 확인하였으며, 평균 성능 향상 폭은 1.74 퍼센티지 포인트였다.

[0055] 도 8은 각 신경망에 사용된 가중치의 개수를 나타낸다. 적응형 컨볼루션은 기존 컨볼루션 신경망 대비 200배 가량 많은 가중치를 필요로 하였다. 그러나, 해싱 기법을 사용하면 최소한의 가중치를 사용하여 컨볼루션 필터를 생성할 수 있는 것을 확인하였다. 해싱 기법을 사용하면 기존 컨볼루션 신경망 대비 4배정도의 가중치만을 통해 컨볼루션 필터를 생성할 수 있다.

[0056] 도 9는 본 발명의 일 실시예에 따른 텍스트 분류를 위한 적응형 컨볼루션 필터 생성 방법의 진행 단계를 나타낸 흐름도이다.

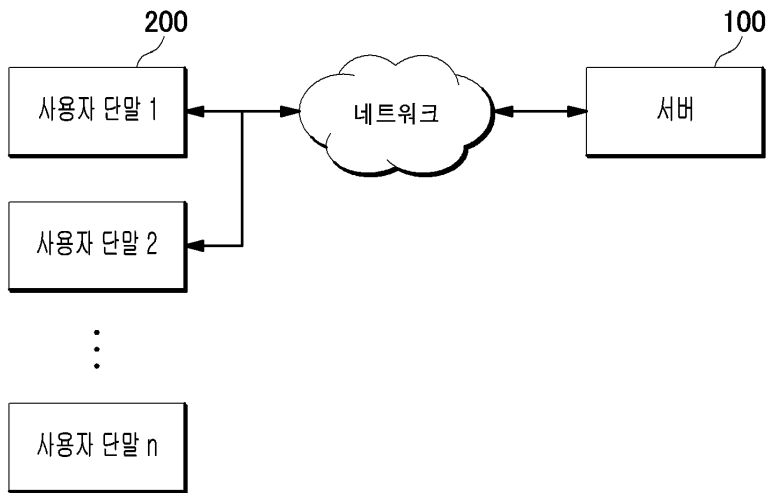
- [0057] 프로세서(130)는 네트워크를 통해 수집된 원시 데이터에 전처리 프로세스를 수행하여 텍스트 데이터를 생성한다(S110).
- [0058] 프로세서(130)는 상기 텍스트 데이터를 입력으로 필터 생성 모델을 통해 컨텍스트 벡터를 생성한다(S120).
- [0059] 필터 생성 모델은 양방향 GRU (Gated Recurrent Unit)를 통해 컨텍스트 벡터를 생성할 수 있다. 양방향 GRU (Gated Recurrent Unit)는 상기 텍스트 데이터를 상관관계를 내포하는 벡터들의 집합으로 변형하고, 해당 시점에서 분류해야 할 텍스트와 연관이 있는 텍스트 부분에 집중하는 어텐션 기법을 적용하여 가중치를 계산하여, 컨텍스트 벡터를 생성하는 것일 수 있다.
- [0060] 프로세서(130)는 상기 컨텍스트 벡터에 기초하여 상기 필터 생성 모델을 통해 콘볼루션 필터를 생성한다(S130).
- [0061] 필터 생성 모델은 해싱 기법을 통해 콘볼루션 필터를 생성할 수 있다. 상기 해싱 기법은 컴포넌트 필터를 포함하는 공유 행렬을 생성하고, 기설정된 해시 함수를 사용하여 후보 컴포넌트 필터를 선택하고, 상기 후보 컴포넌트 필터에 대해 가중치에 따라 선형 조합하는 것일 수 있다. 상기 가중치는 컨텍스트 벡터를 입력으로 받는 단층 신경망의 출력값일 수 있다.
- [0062] 프로세서(130)는 상기 콘볼루션 필터를 기설정된 콘볼루션 신경망에 적용한다(S140). 프로세서(130)는 상기 콘볼루션 필터가 적용된 콘볼루션 신경망을 통해 분류된 텍스트를 사용자 단말에 제공할 수 있다.
- [0063] 본 발명의 일 실시예는 컴퓨터에 의해 실행되는 프로그램 모듈과 같은 컴퓨터에 의해 실행가능한 명령어를 포함하는 기록 매체의 형태로도 구현될 수 있다. 컴퓨터 판독 가능 매체는 컴퓨터에 의해 액세스될 수 있는 임의의 가용 매체일 수 있고, 휘발성 및 비휘발성 매체, 분리형 및 비분리형 매체를 모두 포함한다. 또한, 컴퓨터 판독 가능 매체는 컴퓨터 저장 매체를 포함할 수 있다. 컴퓨터 저장 매체는 컴퓨터 판독가능 명령어, 데이터 구조, 프로그램 모듈 또는 기타 데이터와 같은 정보의 저장을 위한 임의의 방법 또는 기술로 구현된 휘발성 및 비휘발성, 분리형 및 비분리형 매체를 모두 포함한다.
- [0064] 본 발명의 방법 및 시스템은 특정 실시예와 관련하여 설명되었지만, 그것들의 구성 요소 또는 동작의 일부 또는 전부는 범용 하드웨어 아키텍처를 갖는 컴퓨터 시스템을 사용하여 구현될 수 있다.
- [0065] 진술한 본원의 설명은 예시를 위한 것이며, 본원이 속하는 기술분야의 통상의 지식을 가진 자는 본원의 기술적 사상이나 필수적인 특징을 변경하지 않고서 다른 구체적인 형태로 쉽게 변형이 가능하다는 것을 이해할 수 있을 것이다. 그러므로 이상에서 기술한 실시예들은 모든 면에서 예시적인 것이며 한정적이 아닌 것으로 이해해야만 한다. 예를 들어, 단일형으로 설명되어 있는 각 구성 요소는 분산되어 실시될 수도 있으며, 마찬가지로 분산된 것으로 설명되어 있는 구성 요소들도 결합된 형태로 실시될 수 있다.
- [0066] 본원의 범위는 상기 상세한 설명보다는 후술하는 특허청구범위에 의하여 나타내어지며, 특허청구범위의 의미 및 범위 그리고 그 균등 개념으로부터 도출되는 모든 변경 또는 변형된 형태가 본원의 범위에 포함되는 것으로 해석되어야 한다.

부호의 설명

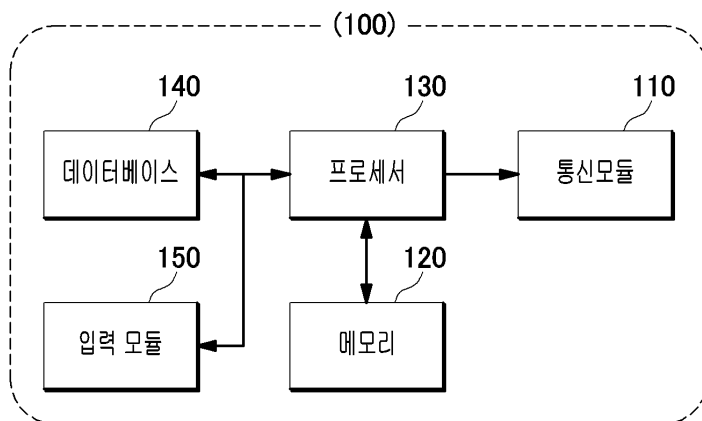
- [0067]
- 100: 서버
 - 110: 통신 모듈
 - 120: 메모리
 - 130: 프로세서
 - 140: 데이터베이스
 - 150: 입력모듈
 - 200: 사용자 단말

도면

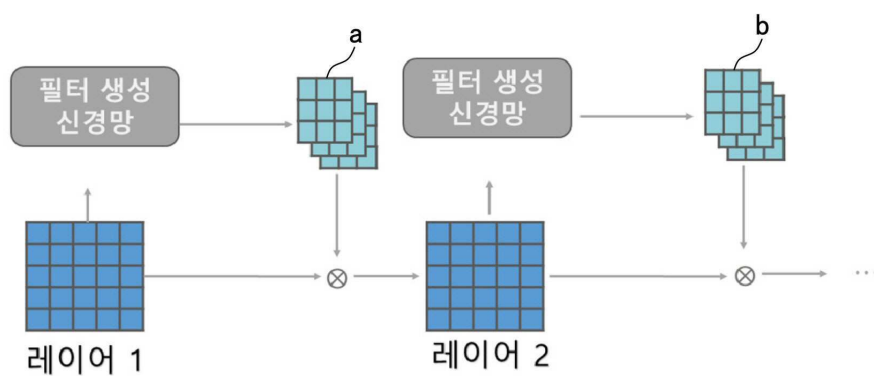
도면1



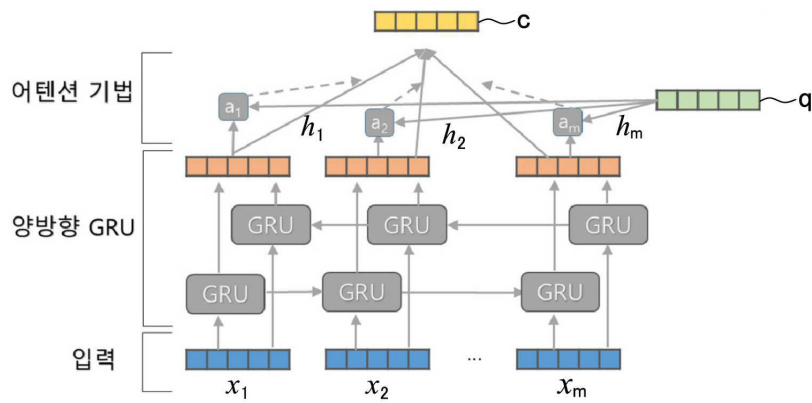
도면2



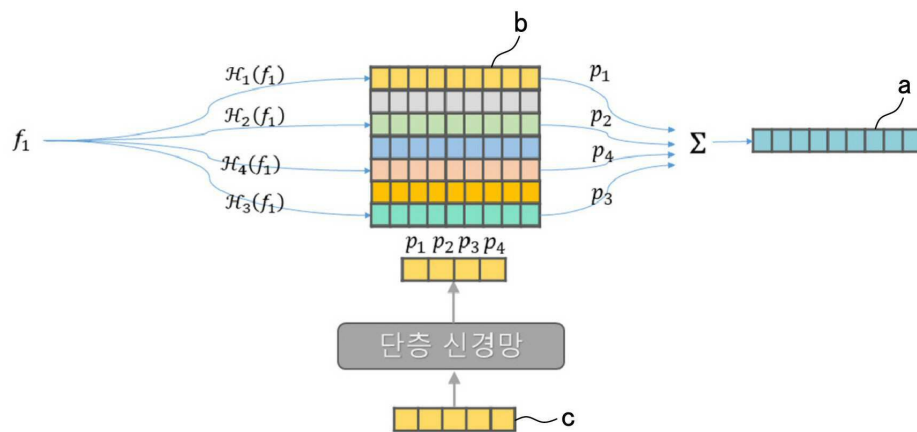
도면3



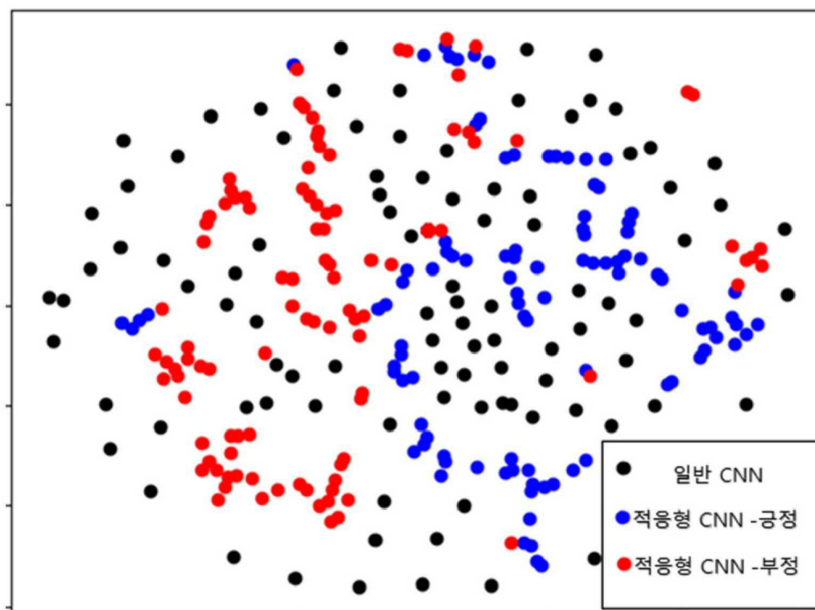
도면4



도면5



도면6



도면7

모델명	성능
일반 CNN	65.52
적응형 CNN	68.12
적응형 CNN - 해싱 기법 사용	68.17
일반 DPCNN	67.01
적응형 DPCNN	67.94
적응형 DPCNN - 해싱 기법 사용	67.98
일반 DenseCNN	66.02
적응형 DenseCNN	67.63
적응형 DenseCNN - 해싱 기법 사용	67.74

도면8

모델명	파라미터
일반 CNN	40만
적응형 CNN	2억1750만
적응형 CNN - 해싱 기법 사용	210만
일반 DPCNN	310만
적응형 DPCNN	2억2000만
적응형 DPCNN - 해싱 기법 사용	710만
일반 DenseCNN	220만
적응형 DenseCNN	2억6320만
적응형 DenseCNN - 해싱 기법 사용	1160만

도면9

