



(19) 대한민국특허청(KR)  
(12) 등록특허공보(B1)

(45) 공고일자 2023년09월20일  
(11) 등록번호 10-2580914  
(24) 등록일자 2023년09월15일

(51) 국제특허분류(Int. Cl.)  
G06F 16/532 (2019.01) G06F 16/538 (2019.01)  
G06F 16/56 (2019.01) G06F 16/583 (2019.01)  
G06N 3/08 (2023.01) G06T 5/00 (2019.01)  
G06T 7/11 (2017.01)  
(52) CPC특허분류  
G06F 16/532 (2019.01)  
G06F 16/538 (2019.01)  
(21) 출원번호 10-2020-0175093  
(22) 출원일자 2020년12월15일  
심사청구일자 2020년12월15일  
(65) 공개번호 10-2022-0085212  
(43) 공개일자 2022년06월22일  
(56) 선행기술조사문헌  
KR1020200002332 A  
(뒷면에 계속)

(73) 특허권자  
세종대학교산학협력단  
서울특별시 광진구 능동로 209 (군자동, 세종대학교)  
(72) 발명자  
유성준  
서울특별시 광진구 능동로 209, 대양AI센터 717호(군자동)  
구영현  
서울특별시 광진구 능동로 209, 대양AI센터 717호(군자동)  
(뒷면에 계속)  
(74) 대리인  
양성보

전체 청구항 수 : 총 4 항

심사관 : 박미정

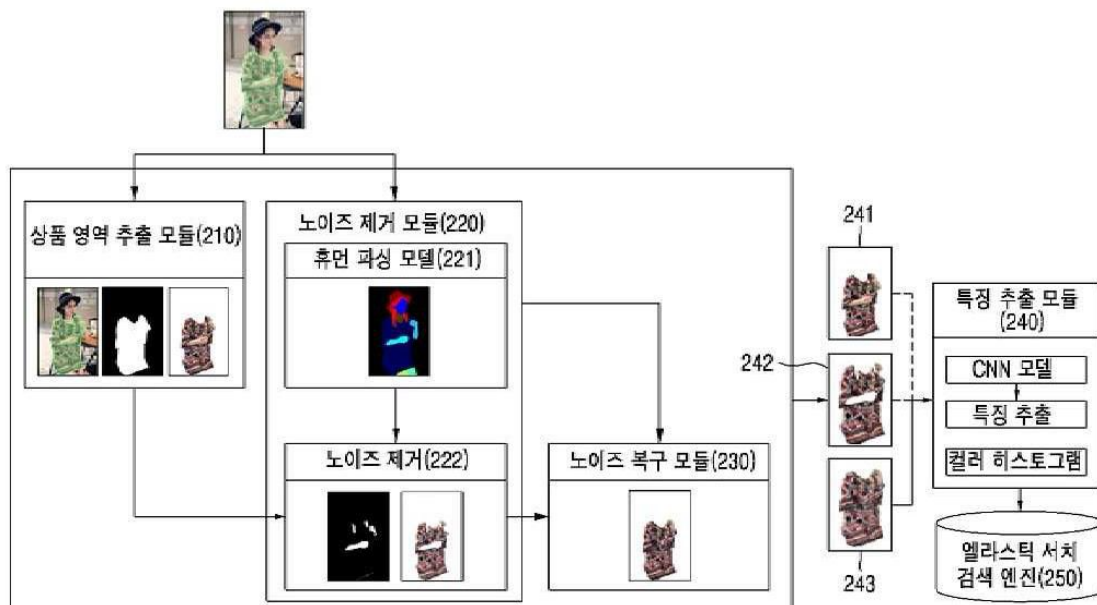
(54) 발명의 명칭 딥러닝 기반 이미지 복구 기술을 이용한 패션 이미지 검색 방법 및 장치

(57) 요약

딥러닝 기반 이미지 복구 기술을 이용한 패션 이미지 검색 방법 및 장치가 제시된다. 본 발명에서 제안하는 딥러닝 기반 이미지 복구 기술을 이용한 패션 이미지 검색 장치는 입력 이미지에서 검색을 위한 상품의 영역을 픽셀 단위로 추출하는 상품 영역 추출 모듈, 입력 이미지에서 휴먼 파싱(human parsing) 모델을 통해 사람의 팔,

(뒷면에 계속)

대표도



다리, 머리, 얼굴을 포함하는 인체 구성을 각각 분해하고, 분해된 인체 구성의 각각의 객체 영역을 픽셀 단위로 추출하고, 상품 영역 추출 모듈에서 추출된 상품 영역을 가진 객체 영역을 찾아 해당 노이즈 영역을 상품 영역에서 제거하는 노이즈 제거 모듈, 제거된 노이즈 영역에 의해 제거된 상품 영역을 복구하는 노이즈 복구 모듈, 복구된 상품 영역에서 상품의 모양, 스타일, 색상 정보를 포함하는 특징을 추출하는 특징 추출 모듈 및 추출된 특징을 벡터로 저장하고, 저장된 특징에 상응하는 상품을 검색하는 엘라스틱서치(Elasticsearch) 검색 엔진을 포함한다.

(52) CPC특허분류

**G06F 16/56** (2019.01)

**G06F 16/583** (2019.01)

**G06N 3/08** (2023.01)

**G06T 5/002** (2023.01)

**G06T 7/11** (2017.01)

(72) 발명자

**박철호**

서울특별시 광진구 능동로 209, 세종대학교 대양AI센터 717호(군자동)

**김해림**

서울특별시 광진구 능동로 209, 세종대학교 대양AI센터 410호(군자동)

**윤학림**

서울특별시 광진구 능동로 209, 대양AI센터 717호(군자동)

**정다운**

서울특별시 광진구 능동로 209, 대양AI센터 717호(군자동)

**정원희**

서울특별시 광진구 능동로 209, 대양AI센터 717호(군자동)

(56) 선행기술조사문헌

KR1020200080377 A\*

KR1020190105745 A

CN109978893 A

PATHAK Deepak, et al, "Context Encoders: Feature Learning by Inpainting", arXiv: 1604.07379v2 [cs.CV], pp. 1-12, 2016.11.21 1부.\*

\*는 심사관에 의하여 인용된 문헌

이 발명을 지원한 국가연구개발사업

과제고유번호	1711103369
과제번호	2019-0-00136-002
부처명	과학기술정보통신부
과제관리(전문)기관명	정보통신기획평가원
연구사업명	인공지능융합선도프로젝트(R&D)
연구과제명	스마트시티 산업 생산성 혁신을 위한 AI융합 기술 개발
기 여 율	1/1
과제수행기관명	세종대학교 산학협력단
연구기간	2020.01.01 ~ 2020.12.31

## 명세서

### 청구범위

#### 청구항 1

입력 이미지에서 검색을 위한 상품의 영역을 픽셀 단위로 추출하는 상품 영역 추출 모듈;

입력 이미지에서 휴먼 파싱(human parsing) 모델을 통해 사람의 팔, 다리, 머리, 얼굴을 포함하는 인체 구성을 각각 분해하고, 분해된 인체 구성의 각각의 객체 영역을 픽셀 단위로 추출하고, 상품 영역 추출 모듈에서 추출된 상품 영역을 가린 객체 영역을 찾아 해당 노이즈 영역을 상품 영역에서 제거하는 노이즈 제거 모듈;

제거된 노이즈 영역에 의해 제거된 상품 영역을 복구하는 노이즈 복구 모듈;

복구된 상품 영역에서 상품의 모양, 스타일, 색상 정보를 포함하는 특징을 추출하는 특징 추출 모듈; 및

추출된 특징을 벡터로 저장하고, 저장된 특징에 상응하는 상품을 검색하는 엘라스틱서치(Elasticsearch) 검색 엔진

을 포함하고,

상품 영역 추출 모듈은,

Deepfashion2 데이터셋에서 제공되는 상품 세그멘테이션 주석을 사용하여 모델을 훈련하고, 상기 Deepfashion2 데이터셋에서 제공되는 상품 세그멘테이션 주석은 상품의 모양에 기초하여 객체가 상품 영역을 가렸을 경우 객체를 상품 영역에 함께 포함하여, 상품 영역을 가린 객체 영역이 포함되어 있는 상품 영역을 추출하고, 상기 Deepfashion2 데이터셋을 이용하여 시멘틱 세그멘테이션 기법으로 추출한 상품의 영역에서 노이즈를 찾도록 하며,

상기 노이즈 제거 모듈은,

휴먼 파싱 모델을 통해 PGN(Part Grouping Network) 구조 및 LIP(Look Into Person) 데이터셋으로 미리 학습된 가중치를 사용하고, LIP 데이터셋은 복수의 시멘틱 인체 구성의 레이블을 표기한 복수의 이미지를 포함하고, LIP 데이터셋으로 미리 학습된 모델을 통해 입력 이미지에서 분해된 인체 구성의 각각의 객체 영역을 객체 별로 구분되고 픽셀 단위로 추출하고,

상품 영역 추출 모듈에서 추출된 상품 영역과 휴먼 파싱(human parsing) 모델을 통해 추출된 객체 영역을 이용하여, 입력 이미지로부터 상품의 영역과 배경을 구분하고 상품 영역을 변환하여 상품 마스크 이미지를 획득하고, 입력 이미지로부터 휴먼 파싱을 통해 각각의 객체를 픽셀 단위로 인식하고 객체 영역을 변환하여 노이즈 마스크 이미지를 획득하며, 획득된 상품 마스크 이미지와 노이즈 마스크 이미지를 이용하여 상품 영역에서 노이즈를 제거하는

딥러닝 기반 이미지 복구 기술을 이용한 패션 이미지 검색 장치.

#### 청구항 2

삭제

#### 청구항 3

삭제

#### 청구항 4

삭제

#### 청구항 5

제1항에 있어서,

노이즈 복구 모듈은,

노이즈 마스크 이미지 및 노이즈 제거 모듈의 출력인 상품 영역을 가린 객체 영역이 상품 영역에서 제거된 상품 이미지를 입력으로 받아 CRA(Contextual Residual Aggregation) 구조 및 미리 훈련된 훈련된 가중치를 사용하여 노이즈 영역에 의해 제거된 상품 영역을 복구하는

딥러닝 기반 이미지 복구 기술을 이용한 패션 이미지 검색 장치.

## 청구항 6

상품 영역 추출 모듈을 통해 입력 이미지에서 검색을 위한 상품의 영역을 픽셀 단위로 추출하는 단계;

노이즈 제거 모듈의 휴먼 파싱(human parsing) 모델을 통해 입력 이미지에서 사람의 팔, 다리, 머리, 얼굴을 포함하는 인체 구성을 각각 분해하고, 분해된 인체 구성의 각각의 객체 영역을 픽셀 단위로 추출하는 단계;

노이즈 제거 모듈을 통해 상품 영역 추출 모듈에서 추출된 상품 영역을 가린 객체 영역을 찾아 해당 노이즈 영역을 상품 영역에서 제거하는 단계;

제거된 노이즈 영역에 의해 제거된 상품 영역을 노이즈 복구 모듈을 통해 복구하는 단계;

복구된 상품 영역에서 특징 추출 모듈을 통해 상품의 모양, 스타일, 색상 정보를 포함하는 특징을 추출하는 단계; 및

추출된 특징을 엘라스틱서치(Elasticsearch) 검색 엔진을 통해 벡터로 저장하고, 저장된 특징에 상응하는 상품을 검색하는 단계

를 포함하고,

상품 상품 영역 추출 모듈을 통해 입력 이미지에서 검색을 위한 상품의 영역을 픽셀 단위로 추출하는 단계는,

Deepfashion2 데이터셋에서 제공되는 상품 세그멘테이션 주석을 사용하여 모델을 훈련하고, 상기 Deepfashion2 데이터셋에서 제공되는 상품 세그멘테이션 주석은 상품의 모양에 기초하여 객체가 상품 영역을 가렸을 경우 객체를 상품 영역에 함께 포함하여, 상품 영역을 가린 객체 영역이 포함되어 있는 상품 영역을 추출하고, 상기 Deepfashion2 데이터셋을 이용하여 시멘틱 세그멘테이션 기법으로 추출한 상품의 영역에서 노이즈를 찾으려 하며,

상기 노이즈 제거 모듈의 휴먼 파싱(human parsing) 모델을 통해 입력 이미지에서 사람의 팔, 다리, 머리, 얼굴을 포함하는 인체 구성을 각각 분해하고, 분해된 인체 구성의 각각의 객체 영역을 픽셀 단위로 추출하는 단계는,

휴먼 파싱 모델을 통해 PGN(Part Grouping Network) 구조 및 LIP(Look Into Person) 데이터셋으로 미리 학습된 가중치를 사용하고, LIP 데이터셋은 복수의 시멘틱 인체 구성의 레이블을 표기한 복수의 이미지를 포함하고, LIP 데이터셋으로 미리 학습된 모델을 통해 입력 이미지에서 분해된 인체 구성의 각각의 객체 영역을 객체 별로 구분되고 픽셀 단위로 추출되고,

상품 영역 추출 모듈에서 추출된 상품 영역과 휴먼 파싱(human parsing) 모델을 통해 추출된 객체 영역을 이용하여, 입력 이미지로부터 상품의 영역과 배경을 구분하고 상품 영역을 변환하여 상품 마스크 이미지를 획득하고, 입력 이미지로부터 휴먼 파싱을 통해 각각의 객체를 픽셀 단위로 인식하고 객체 영역을 변환하여 노이즈 마스크 이미지를 획득하며, 획득된 상품 마스크 이미지와 노이즈 마스크 이미지를 이용하여 상품 영역에서 노이즈를 제거되는

딥러닝 기반 이미지 복구 기술을 이용한 패션 이미지 검색 방법.

## 청구항 7

삭제

## 청구항 8

삭제

## 청구항 9

삭제

## 청구항 10

제6항에 있어서,

제거된 노이즈 영역에 의해 제거된 상품 영역을 노이즈 복구 모듈을 통해 복구하는 단계는,

노이즈 마스크 이미지 및 노이즈 제거 모듈의 출력인 상품 영역을 가진 객체 영역이 상품 영역에서 제거된 상품 이미지를 입력으로 받아 CRA(Contextual Residual Aggregation) 구조 및 미리 훈련된 훈련된 가중치를 사용하여 노이즈 영역에 의해 제거된 상품 영역이 복구되는

딥러닝 기반 이미지 복구 기술을 이용한 패션 이미지 검색 방법.

## 발명의 설명

### 기술 분야

[0001] 본 발명은 딥러닝 기반 이미지 복구 기술을 이용한 패션 이미지 검색 방법 및 장치에 관한 것이다.

### 배경 기술

[0002] 패션 이미지 분석은 최근 많은 관심을 받고 있다. 또한 인터넷 쇼핑의 발전으로 사용자는 선택의 범위가 많아지고, 사용자들의 인터넷 패션 쇼핑 요구는 점차적으로 세분화하고 정밀해지고 있다. 이러한 사용자의 쇼핑 요구를 만족하려면 패션 이미지 분석 연구는 많은 발전이 필요하다. 즉, 더 많은 연구자들이 컴퓨터 비전 기술을 이용해 소비자들에게 더 좋은 상품 또는 쇼핑 서비스를 제공하여야 한다.

[0003] 유사 패션 이미지 검색 또는 기타 패션 서비스를 진행할 때 이미지에서 패션의 영역을 추출하고 패션의 특징 값을 추출하는 것은 필수적인 작업이다. 더 정확한 패션 상품의 특징 값은 더 정확한 유사 패션 이미지 검색 또는 기타 패션 서비스를 제공할 수 있는 기초이다.

[0004] 패션 이미지 검색은 입력한 쿼리(query) 이미지의 상품과 가장 유사한 상품을 찾아 출력하는 기술이다. 패션 이미지 검색에서 상품을 제외한 배경 또는 상품을 가리고 있는 모델의 손, 팔, 액세서리(예를 들어, 가방, 모자 등)는 노이즈로 이미지 검색 정확도를 떨어뜨릴 수 있다. 종래기술에서는 객체 검출(object detection)[1] 또는 시멘틱 세그멘테이션(semantic segmentation)[2] 기법을 이용해 상품의 영역을 추출하는 방법으로 이미지의 배경 등 노이즈를 제거했다. 객체 검출은 이미지에서 바운딩 박스(bounding box)로 상품의 영역을 표기하는 방법으로 상품 주변에 조금의 배경 노이즈가 포함된다. 시멘틱 세그멘테이션은 이미지에서 상품 영역을 픽셀 단위로 추출하기 때문에 바운딩 박스 보다 노이즈가 적은 특징이 있다. 그러나 기타 객체가 상품의 영역을 가렸을 때 기존 시멘틱 세그멘테이션 방법으로 추출한 상품의 픽셀 영역은 상품의 모양 정보를 잃거나 기타 객체가 계속 상품의 영역에 노이즈로 존재하는 문제점이 있다.

[0006] 도 1은 종래기술에 따른 객체가 상품의 영역을 가린 경우와 가리지 않은 경우의 이미지 검출을 설명하기 위한 도면이다.

[0007] 이미지에서 패션 상품을 시멘틱 세그멘테이션 할 때 기타 객체가 상품의 영역을 가린 경우와 가리지 않은 경우의 2가지 경우로 나눌 수 있다. 도 1(a)에 도시된 "type-A"는 드레스 상품 영역을 가린 기타 객체가 없는 경우이다. 도 1(b)의 "type-B" 및 도 1(c)의 "type-C"는 기타 객체가 드레스 상품 영역을 가린 경우이다. "type-A"와 같이 상품 영역을 가린 기타 객체가 없을 경우는 시멘틱 세그멘테이션으로 정확한 상품의 픽셀 영역을 추출할 수 있다.

[0008] 하지만 기타 객체가 상품 영역을 가린 경우 시멘틱 세그멘테이션 기법을 사용하면 "type-B", "type-C"와 같은 결과를 출력하게 된다. "type-B"에서는 상품을 가린 손과 가방이 상품의 픽셀 영역에서 제거되었다. 이러한 경우 손과 가방이 상품에서 추출한 특징에 영향을 주는 노이즈는 제거하였지만 상품의 모양이 변형되어 상품 모양 정보를 잃게 된다. "type-C"에는 상품을 가린 모델의 손을 상품의 픽셀 영역에 포함하였다. 이는 종래기술 deepfashion2에서 사용하는 시멘틱 세그멘테이션 방법이다. Deepfashion2에서 사용한 시멘틱 세그멘테이션 기법은 상품의 모양 정보는 최대한 보존하였으나 상품의 픽셀 영역에 패션모델의 손이 노이즈로 남아 있는 문제가

존재한다. "type-B"와 "type-C"의 차이는 데이터셋을 구축할 때 상품 영역에 대한 정의가 다르기 때문이다. "type-B"에서는 이미지에서 픽셀 단위로 상품의 영역을 정확히 추출하는데 집중한다. "type-C"에서는 상품의 모양을 기반으로 이미지에서 상품의 가려진 픽셀도 상품의 영역에 포함한다. 일반적인 시멘틱 세그멘테이션의 데이터셋에서는 "type-B"와 같은 방법을 많이 사용한다.

## 선행기술문헌

### 특허문헌

[0009] (특허문헌 0001) 한국 공개특허공보 제10-2020-0002332호(2020.01.08.)

## 발명의 내용

### 해결하려는 과제

[0010] 본 발명이 이루고자 하는 기술적 과제는 휴먼 파싱(human parsing) 기법을 이용하여 시멘틱 세그멘테이션 기법으로 추출한 패션 상품 영역에 포함된 노이즈를 찾아 제거하고 이미지 복구 기술로 제거한 노이즈 영역을 복구하는 방법 및 장치를 제공하는데 있다.

### 과제의 해결 수단

[0011] 일 측면에 있어서, 본 발명에서 제안하는 딥러닝 기반 이미지 복구 기술을 이용한 패션 이미지 검색 장치는 입력 이미지에서 검색을 위한 상품의 영역을 픽셀 단위로 추출하는 상품 영역 추출 모듈, 입력 이미지에서 휴먼 파싱(human parsing) 모델을 통해 사람의 팔, 다리, 머리, 얼굴을 포함하는 인체 구성을 각각 분해하고, 분해된 인체 구성의 각각의 객체 영역을 픽셀 단위로 추출하고, 상품 영역 추출 모듈에서 추출된 상품 영역을 가린 객체 영역을 찾아 해당 노이즈 영역을 상품 영역에서 제거하는 노이즈 제거 모듈, 제거된 노이즈 영역에 의해 제거된 상품 영역을 복구하는 노이즈 복구 모듈, 복구된 상품 영역에서 상품의 모양, 스타일, 색상 정보를 포함하는 특징을 추출하는 특징 추출 모듈 및 추출된 특징을 벡터로 저장하고, 저장된 특징에 상응하는 상품을 검색하는 엘라스틱서치(Elasticsearch) 검색 엔진을 포함한다.

[0012] 상품 영역 추출 모듈은 시멘틱 세그멘테이션 모델을 통해 데이터셋에서 제공하는 상품의 시멘틱 세그멘테이션 주석을 사용하여 모델을 훈련하고, 상품의 모양에 기초하여 객체가 상품 영역을 가렸을 경우 객체를 상품 영역에 함께 포함하여, 상품 영역을 가린 객체 영역이 포함되어 있는 상품 영역을 추출한다.

[0013] 노이즈 제거 모듈은 휴먼 파싱 모델을 통해 PGN(Part Grouping Network) 구조 및 LIP(Look Into Person) 데이터셋으로 미리 학습된 가중치를 사용하고, LIP 데이터셋은 복수의 시멘틱 인체 구성의 레이블을 표기한 복수의 이미지를 포함하고, LIP 데이터셋으로 미리 학습된 모델을 통해 입력 이미지에서 분해된 인체 구성의 각각의 객체 영역을 객체 별로 구분되고 픽셀 단위로 추출한다.

[0014] 노이즈 제거 모듈은 상품 영역 추출 모듈에서 추출된 상품 영역과 휴먼 파싱(human parsing) 모델을 통해 추출된 객체 영역을 이용하여, 입력 이미지로부터 상품 영역을 변환하여 상품 마스크 이미지를 획득하고, 입력 이미지로부터 객체 영역을 변환하여 노이즈 마스크 이미지를 획득하며, 획득된 상품 마스크 이미지와 노이즈 마스크 이미지를 이용하여 상품 영역에서 노이즈를 제거한다.

[0015] 노이즈 복구 모듈은 노이즈 마스크 이미지 및 노이즈 제거 모듈의 출력인 상품 영역을 가린 객체 영역이 상품 영역에서 제거된 상품 이미지를 입력으로 받아 CRA(Contextual Residual Aggregation) 구조 및 미리 훈련된 훈련된 가중치를 사용하여 노이즈 영역에 의해 제거된 상품 영역을 복구한다.

[0016] 또 다른 일 측면에 있어서, 본 발명에서 제안하는 딥러닝 기반 이미지 복구 기술을 이용한 패션 이미지 검색 방법은 상품 영역 추출 모듈을 통해 입력 이미지에서 검색을 위한 상품의 영역을 픽셀 단위로 추출하는 단계, 노이즈 제거 모듈의 휴먼 파싱(human parsing) 모델을 통해 입력 이미지에서 사람의 팔, 다리, 머리, 얼굴을 포함하는 인체 구성을 각각 분해하고, 분해된 인체 구성의 각각의 객체 영역을 픽셀 단위로 추출하는 단계, 노이즈 제거 모듈을 통해 상품 영역 추출 모듈에서 추출된 상품 영역을 가린 객체 영역을 찾아 해당 노이즈 영역을 상품 영역에서 제거하는 단계, 제거된 노이즈 영역에 의해 제거된 상품 영역을 노이즈 복구 모듈을 통해 복구하는 단계, 복구된 상품 영역에서 특징 추출 모듈을 통해 상품의 모양, 스타일, 색상 정보를 포함하는 특징을 추출하

는 단계 및 추출된 특징을 엘라스틱서치(Elasticsearch) 검색 엔진을 통해 벡터로 저장하고, 저장된 특징에 상응하는 상품을 검색하는 단계를 포함한다.

### 발명의 효과

[0017] 본 발명의 실시예들에 따르면 휴먼 파싱(human parsing) 기법을 이용하여 시멘틱 세그멘테이션 기법으로 추출한 패션 상품 영역에 포함된 노이즈를 찾아 제거하고 이미지 복구 기술로 제거한 노이즈 영역을 복구하는 방법 및 장치를 제공하고, 복구한 상품 이미지에서 상품의 특징을 추출함으로써 정확한 특징을 추출할 수 있어 더 정확한 유사 이미지 검색을 진행할 수 있다.

### 도면의 간단한 설명

[0018] 도 1은 종래기술에 따른 객체가 상품의 영역을 가린 경우와 가리지 않은 경우의 이미지 검출을 설명하기 위한 도면이다.

도 2는 본 발명의 일 실시예에 따른 딥러닝 기반 이미지 복구 기술을 이용한 패션 이미지 검색 장치의 구성을 나타내는 도면이다.

도 3은 본 발명의 일 실시예에 따른 시멘틱 세그멘테이션 모델을 통한 상품 영역 추출을 설명하기 위한 도면이다.

도 4는 본 발명의 일 실시예에 따른 휴먼 파싱 모델을 통해 패션 상품 이미지에서 픽셀 별 카테고리를 탐지한 결과의 예시를 나타내는 도면이다.

도 5는 본 발명의 일 실시예에 따른 단계별 출력 이미지의 예시를 나타내는 도면이다.

도 6은 본 발명의 일 실시예에 따른 노이즈 복구 모듈의 입출력 예시를 나타내는 도면이다.

도 7은 본 발명의 일 실시예에 따른 딥러닝 기반 이미지 복구 기술을 이용한 패션 이미지 검색 방법을 설명하기 위한 흐름도이다.

### 발명을 실시하기 위한 구체적인 내용

[0019] 본 발명은 유사 패션 이미지 검색을 진행할 때 이미지에서 패션 상품을 추출하는 방법을 개선하여 패션 이미지 검색 정확도를 높이는 것에 관한 것이다. 기존 휴먼 파싱(human parsing) 기술을 이용해 이미지에서 시멘틱 세그멘테이션(semantic segmentation) 기법으로 추출한 패션 상품 영역에서 패션 상품 영역에 속하지 않는 모델의 손, 머리카락 등의 노이즈 영역을 찾아 제거하고 기존 다른 데이터셋에서 학습된 이미지 복구 모델로 해당 노이즈 영역을 패션의 원래 모양으로 복구하는 방법을 제안한다. 이후, 노이즈 영역을 복구한 상품 이미지로 유사 상품 이미지 검색을 진행하여 기존 시멘틱 세그멘테이션 기법으로 추출한 상품을 직접 사용하였을 때 보다 검색 정확도를 증가시킬 수 있다. 이하, 본 발명의 실시 예를 첨부된 도면을 참조하여 상세하게 설명한다.

[0021] 도 2는 본 발명의 일 실시예에 따른 딥러닝 기반 이미지 복구 기술을 이용한 패션 이미지 검색 장치의 구성을 나타내는 도면이다.

[0022] 제안하는 이미지 복구 기술을 이용한 패션 이미지 검색 장치는 상품 영역 추출 모듈(210), 노이즈 제거 모듈(220), 노이즈 복구 모듈(230), 특징 추출 모듈(240) 및 엘라스틱서치(Elasticsearch) 검색 엔진(250)을 포함한다.

[0023] 상품 영역 추출 모듈(210)은 입력 이미지에서 검색을 위한 상품의 영역을 픽셀 단위로 추출한다.

[0025] 도 3은 본 발명의 일 실시예에 따른 시멘틱 세그멘테이션 모델을 통한 상품 영역 추출을 설명하기 위한 도면이다.

[0026] 도 3(a)는 종래기술에 따른 이미지 매팅(image matting) 기법으로 추출한 결과, 도 3(b)는 종래기술에 따른 객체 검출(object detection) 기법으로 추출한 결과, 도 3(c)는 본 발명의 일 실시예에 따른 시멘틱 세그멘테이션(semantic segmentation) 기법으로 추출한 결과를 나타내는 도면이다.

[0027] 이미지에서 더 정확한 상품 영역을 추출하고 배경 등 노이즈 정보가 추출한 특징에 주는 영향을 최소화하는 방법에는 이미지 매팅[10], 객체 검출, 시멘틱 세그멘테이션 등의 기법이 있다. 도 3은 동일한 상품 이미지에서 3가지 기법으로 상품을 추출한 결과를 비교하기 위한 도면이다. 하였습니다. 도 3을 참조하면, 객체 검출 기법으



로 추출한 상품은 이미지 매칭 기법으로 추출한 상품보다 노이즈가 작고, 시멘틱 세그멘테이션 기법으로 추출한 상품은 객체 검출 기법으로 추출한 상품보다 노이즈가 작다. 하지만 시멘틱 세그멘테이션 기법으로 추출한 상품의 영역은 도 3(c)의 붉은색 원으로 표기한 영역과 같이 상품의 원래 모양 정보를 파괴하는 문제점이 있다. 즉, 추출한 특징은 상품 모양이 파괴된 특징 값이 된다.

[0028] 이미지 복구 기술은 전체 이미지의 구성에 근거하여 복구가 필요한 영역을 복구하는 기법이다. 하지만 이미지 복구 모델은 복구가 필요한 이미지와 복구할 마스크 영역을 동시에 입력 데이터로 입력하여야 고 화질의 복구 결과를 출력할 수 있다. 따라서, 세그멘테이션 결과에서 노이즈 영역을 찾아 표기하는 것이 문제의 핵심이다. Deepfashion2에서는 상품의 모양을 기반으로 한 패션 세그멘테이션 데이터를 제공한다. 즉, Deepfashion2에서 시멘틱 세그멘테이션 기법으로 추출한 상품 영역에는 패션모델의 손, 머리카락 등 노이즈가 존재한다.

[0029] 따라서, 본 발명에서는 휴먼 파싱 모델을 이용하여 Deepfashion2에서 시멘틱 세그멘테이션 기법으로 추출한 상품의 영역에서 노이즈를 찾는 방안을 제안한다.

[0030] 상품 영역 추출 모듈(210)은 시멘틱 세그멘테이션 모델을 통해 데이터셋에서 제공하는 상품의 시멘틱 세그멘테이션 주석을 사용하여 모델을 훈련하고, 상품의 모양에 기초하여 객체가 상품 영역을 가렸을 경우 객체를 상품 영역에 함께 포함하여, 상품 영역을 가린 객체 영역이 포함되어 있는 상품 영역을 추출한다.

[0031] 더욱 상세하게는, 상품 추출 모듈에서는 시멘틱 세그멘테이션 모델을 사용하여 입력 이미지에서 상품의 영역을 픽셀 단위로 찾고 상품의 카테고리를 분류한다. 본 발명의 실시예에 따르면, Johnson et al. [3]에서 제기한 Mask-CNN 구조를 사용할 수 있다. 그리고 deepfashion2 데이터셋에서 제공하는 패션 상품 세그멘테이션 주석을 사용해 모델을 훈련할 수 있다. deepfashion2에서 제공하는 패션 상품 세그멘테이션 주석은 상품의 모양을 기반으로 기타 객체가 상품을 가렸을 때 함께 상품의 영역에 포함한다. 따라서 시멘틱 세그멘테이션 기법으로 추출한 상품의 영역에는 상품을 제외한 패션모델의 손, 팔, 머리카락 등 불필요한 정보도 포함되어 있다.

[0032] 노이즈 제거 모듈(220)은 입력 이미지에서 휴먼 파싱(human parsing) 모델을 통해 사람의 팔, 다리, 머리, 얼굴을 포함하는 인체 구성을 각각 분해하고, 분해된 인체 구성의 각각의 객체 영역을 픽셀 단위로 추출한다.

[0033] 휴먼 파싱(human parsing) 모델(221)을 통해 PGN(Part Grouping Network) 구조 및 LIP 데이터셋으로 미리 학습된 가중치를 사용한다. LIP 데이터셋은 복수의 시멘틱 인체 구성의 레이블을 표기한 복수의 이미지를 포함하고, LIP 데이터셋으로 미리 학습된 모델을 통해 입력 이미지에서 분해된 인체 구성의 각각의 객체 영역을 객체 별로 구분되고 픽셀 단위로 추출한다.

[0034] 도 4는 본 발명의 일 실시예에 따른 휴먼 파싱 모델을 통해 패션 상품 이미지에서 픽셀 별 카테고리를 탐지한 결과의 예시를 나타내는 도면이다.

[0035] 휴먼 파싱은 사람을 팔, 다리, 머리, 얼굴 등 인체 구성으로 분해하여 표기하는 방법이다. 대표적인 데이터셋으로는 ATR, LIP(Look Into Person), Pascal-Person-Part 등이 있다. 본 발명의 실시예에 따르면, Gong et al.에서 제안하는 PGN(Part Grouping Network) 구조 및 LIP 데이터셋으로 미리 학습한 가중치를 사용할 수 있다.

[0036] LIP 데이터셋은 19개의 시멘틱 인체 구성 레이블(semantic human part labels)을 표기한 약 50,000장 이미지로 구성되어 있다. LIP 데이터셋에서 제공한 19개 시멘틱 인체 구성 레이블에는 모자, 머리, 장갑, 선글라스, 상의, 드레스, 코트, 양말, 바지, 몸통 피부, 스카프, 스커트, 얼굴, 왼쪽 팔, 오른쪽 팔, 왼쪽 다리, 오른쪽 다리, 왼쪽 신발, 오른쪽 신발이 있다. LIP 데이터셋으로 학습된 모델은 이미지에서 사람을 중심으로 사람의 피부 및 모자, 스카프 등의 아이템을 객체 별로 구분하고 픽셀 단위로 인식한다. 12는 human parsing 모델을 이용해 패션 상품 이미지에서 픽셀 별 카테고리를 탐지한 결과의 예시입니다.

[0037] 이후, 노이즈 제거 모듈(220)은 상품 영역 추출 모듈에서 추출된 상품 영역을 가린 객체 영역을 찾아 해당 노이즈 영역을 상품 영역에서 제거한다(222).

[0038] 상품 영역 추출 모듈에서 추출된 상품 영역과 휴먼 파싱 모델을 통해 추출된 객체 영역을 이용하여, 입력 이미지로부터 상품 영역을 변환하여 상품 마스크 이미지를 획득하고, 입력 이미지로부터 객체 영역을 변환하여 노이즈 마스크 이미지를 획득한다. 획득된 상품 마스크 이미지와 노이즈 마스크 이미지를 이용하여 상품 영역에서 노이즈를 제거한다.

[0039] 도 5는 본 발명의 일 실시예에 따른 단계별 출력 이미지의 예시를 나타내는 도면이다.

[0040] 노이즈란 패션 상품 검색 결과에 영향을 주는 배경, 기타 객체 등 상품을 제외한 불필요한 데이터들을



의미한다. 앞서 설명된 상품 추출 모듈에서는 배경, 기타 객체를 제외한 상품의 영역을 추출한다. 하지만 deepfashion2에서 제안한 패션 상품 모양을 기반으로 한 세그멘테이션 데이터셋을 사용하였기 때문에 상품 영역에는 패션모델의 손, 팔, 가방, 머리카락 등이 포함되어 있다. 이는 상품이 아니므로 유사 상품 이미지 검색에서 노이즈가 된다. 노이즈 제거 모듈에서는 픽셀 단위로 추출한 상품 영역에 포함된 이러한 노이즈를 찾아 상품 영역에서 제거한다.

[0041] 도 5는 노이즈 제거 절차를 설명하기 위해 단계별 출력한 샘플 이미지의 예시이다. 도 5(a)는 입력 이미지이고, 도 5(b)는 시멘틱 세그멘테이션 기법으로 추출한 상품 영역이다. 도 5(c)는 시멘틱 세그멘테이션 결과인 도 5(b)에서 상품의 영역을 흰색, 배경을 검은색으로 변경해 얻은 상품 마스크 이미지이다. 도 5(d)는 입력 이미지 도 5(a)에서 휴먼 파싱을 통해 각 객체를 픽셀 단위로 인식하고 배경을 검은색으로 표기한 이미지이다. 도 5(e)는 상품 이미지에서 찾은 노이즈 영역을 흰색으로 표기한 노이즈 마스크 이미지이다. 도 5(f)는 상품 이미지에서 노이즈를 제거한 최종 결과이다.

[0042] 노이즈 마스크 이미지 도 5(e)를 찾기 위해 본 발명에서는 상품 마스크 이미지 도 5(c)에서 흰색으로 표기한 상품의 영역과 휴먼 파싱 결과 도 5(d)에서 모자, 머리, 장갑, 선글라스, 스카프, 왼쪽 팔, 오른쪽 팔, 왼쪽 다리, 오른쪽 다리, 왼쪽 신발, 오른쪽 신발 등 카테고리화 겹치는 영역을 계산한다(이때, 두 이미지의 사이즈는 동일하다). 도 5(d)에서 선택한 11개 카테고리는 LIP 데이터셋에서 제공한 전체 카테고리에서 상품을 가릴 확률이 높은 카테고리만 선택한 결과이다. 도 5(d)에 있는 왼쪽 두 개의 이미지와 같이 패션모델(다시 말해, 사람)이 없을 경우 파싱 모델은 파싱이 실패하거나 잘못된 파싱 결과를 출력할 확률이 높다. 다음 단계로 상품 영역 도 5(b)에서 마이너스 노이즈 마스크 이미지 도 5(e)의 계산을 진행해 노이즈를 제거한 상품 영역 도 5(f)를 얻는다. 도 5(a) 내지 도 5(e)에서 앞쪽 두 라인은 상품을 가린 노이즈가 없는 케이스이고 뒤쪽 두 라인은 상품을 가린 노이즈가 있는 케이스이다.

[0043] 노이즈 복구 모듈(230)은 제거된 노이즈 영역에 의해 제거된 상품 영역을 복구한다. 노이즈 복구 모듈(230)은 노이즈 마스크 이미지 및 노이즈 제거 모듈의 출력인 상품 영역을 가린 객체 영역이 상품 영역에서 제거된 상품 이미지를 입력으로 받아 CRA(Contextual Residual Aggregation) 구조 및 미리 훈련된 훈련된 가중치를 사용하여 노이즈 영역에 의해 제거된 상품 영역을 복구한다.

[0044] 도 6은 본 발명의 일 실시예에 따른 노이즈 복구 모듈의 입출력 예시를 나타내는 도면이다.

[0045] 도 6(a)는 노이즈를 제거한 상품 이미지이고, 도 6(b)는 상품을 가린 노이즈 영역이고, 도 6(c)는 상품을 가린 노이즈 영역을 복구한 상품 이미지이다.

[0046] 노이즈 복구 모듈에서는 앞서 설명된 바와 같이 제거된 상품을 가린 노이즈 영역을 복구한다. 본 발명에서는 2020년 Yi et al. [11]에서 제안한 이미지 복구 모델 CRA(Contextual Residual Aggregation) 구조 및 미리 훈련된 가중치를 사용한다.

[0047] 노이즈 복구 모듈에서는 노이즈를 제거한 상품 이미지와 상품 이미지에서 노이즈 영역을 표기한 노이즈 마스크 이미지를 입력으로 받는다.

[0048] 노이즈 복구 모델은 노이즈를 제거한 상품 이미지에서 전반 이미지의 스타일 및 구조를 기반으로 노이즈 영역을 상품 원본과 동일하게 복구한다. 도 6(a)의 입력 데이터는 도 5(f)의 상품 영역에 있는 노이즈를 제거한 결과에 해당하고, 도 6(a)는 도 5(e)의 상품에 있는 노이즈 영역을 마스크로 표기한 결과에 해당한다. 도 6(c)는 노이즈 영역을 복구한 결과의 예시이다. 노이즈 복구 모델은 도 6(a)를 기반으로 eh 6(b)에 있는 노이즈 마스크(흰색으로 표기) 영역만 복구한다. 그러므로 입력 이미지에 상품을 가린 노이즈 영역이 없을 경우 노이즈 복구를 진행하지 않고 입력 이미지와 동일한 이미지를 출력한다(도 6에서 앞쪽 4 라인). 도 6에서 빨간색으로 표기한 부분은 상품 이미지의 노이즈 영역을 복구한 결과이다.

[0049] 특징 추출 모듈(240)은 복구된 상품 영역에서 상품의 모양, 스타일, 색상 정보를 포함하는 특징을 추출한다.

[0050] 특징 추출 모듈(240)은 상품 영역 추출 모듈의 출력 이미지(241), 노이즈 제거 모듈의 출력 이미지(230) 및 노이즈 복구 모듈의 출력 이미지(243)를 입력 받는다.

[0051] 패션 상품 특징 추출 모듈에서는 변환 학습(transfer learning) 및 OpenCV[12] 라이브러리(library)를 이용해 상품의 특징을 추출한다. 변환 학습[13]이란 한 도메인에서 학습한 모델을 다른 도메인에 적용하는 기술을 말한다. 미리 학습된 모델(pre-trained model)을 기반으로 일부분의 레이어 및 가중치를 재활용하는 특징이 있다. 본 발명에서는 ImageNet[14] 데이터셋으로 사전 학습된 InceptionResNetV2 모델을 사용한다. ImageNet은

WordNet 계층(다시 말해, 명사만)에 따라 구성된 이미지 데이터베이스로 1,000개의 카테고리 및 1,200,000장 이상의 이미지를 포함하고 있다. 이를 학습한 가중치는 일상생활에서 볼 수 있는 다수 객체의 일반적인 특징 정보를 포함하고 있다. InceptionResNetV2 모델을 선택해 keras[15]에서 제공하는 ImageNet 가중치를 로딩하고 'avg\_pool' 층을 출력 층으로 설정한다. 모델의 출력 값은 상품의 모양 및 스타일의 특징으로 사용한다. InceptionResNetV2 모델은 299 x 299 사이즈의 이미지를 입력으로 받고 길이가 1,536인 벡터를 출력한다.

[0052] 많은 패션 연구에서는 상품에서 색상 정보를 별도로 추출해 사용하였다[16, 17, 18]. 따라서 본 발명에서는 OpenCV 라이브러리를 이용해 이미지 히스토그램(histogram)을 추출하여 상품의 색상 특징으로 저장한다.

[0053] 종래기술에 따른 유사 이미지 계산을 위한 특징 추출 알고리즘이 존재 한다[19-21]. 본 발명에서는 변환 학습 기반 특징 추출 방법 및 OpenCV 라이브러리를 이용한 컬러 히스토그램(color histogram) 추출 방법을 사용한다.

[0054] 엘라스틱서치(Elasticsearch) 검색 엔진(250)은 추출된 특징을 벡터로 저장하고, 저장된 특징에 상응하는 상품을 검색한다.

[0055] 본 발명에서는 Elasticsearch[22] 기반 유사 이미지 검색 시스템을 구축한다. 엘라스틱서치는 Apache Lucene을 기반으로 구축한 텍스트, 숫자, 정형 및 비정형 데이터 등 모든 유형의 데이터를 위한 분산형 오픈소스 검색 엔진이다. 이미지 검색 시스템은 InceptionResNetV2로 추출한 특징을 벡터로 변경한 후 base64로 인코딩(encoding) 한 다음 v1로 엘라스틱서치 검색 엔진에 저장한다. OpenCV 라이브러리로 추출한 컬러 히스토그램도 마찬가지로 벡터로 표기하고 base64로 인코딩 한 다음 v2로 엘라스틱서치에 저장한다.

[0056] 유사 상품 이미지 검색 쿼리는 변환 학습 기반으로 추출한 상품의 모양 및 스타일 특징(v1) 및 컬러 히스토그램(v2)을 입력 데이터로 받는다. 변환 학습 기반으로 추출한 특징 (v1)은 v1 벡터 공간에서 상품의 모양 및 스타일 유사도를 계산한다. 컬러 히스토그램(v2)은 v2 벡터공간에서 상품의 색상 유사도를 계산한다. 단일 유사도 계산은 코사인 유사도(cosine similarity)를 사용한다. 이렇게 각자 계산된 유사도는 마지막 가중치 값에 근거해 조합되고 계산 방법은 식(1)과 같다. 딥러닝 모델의 가중치와 구분하기 위해 뒤에서는 식(1)의 "weight"를 유사도 가중치(similarity weight)라 정의한다. 식(1)에서 "weight" 값의 범위는 0~1 이다. "형태 유사도(shape similarity)"는 v1에서 계산한 상품의 모양 및 스타일 유사도이고 "색상 유사도(color similarity)"는 v2에서 계산한 상품의 색상 유사도이다. "유사도(통합)(similarity(total))"은 최종 상품의 유사도를 의미한다. 최종 검색 결과는 유사도가 높은 순으로 구분하여 출력한다.

[0057] 
$$\text{similarity}(\text{total}) = \text{weight} * \text{shape similarity} + (1 - \text{weight}) * \text{color similarity} \quad (1)$$

[0059] 도 7은 본 발명의 일 실시예에 따른 딥러닝 기반 이미지 복구 기술을 이용한 패션 이미지 검색 방법을 설명하기 위한 흐름도이다.

[0060] 제안하는 딥러닝 기반 이미지 복구 기술을 이용한 패션 이미지 검색 방법은 상품 영역 추출 모듈을 통해 입력 이미지에서 검색을 위한 상품의 영역을 픽셀 단위로 추출하는 단계(710), 노이즈 제거 모듈의 휴먼 파싱(human parsing) 모델을 통해 입력 이미지에서 사람의 팔, 다리, 머리, 얼굴을 포함하는 인체 구성을 각각 분해하고, 분해된 인체 구성의 각각의 객체 영역을 픽셀 단위로 추출하는 단계(720), 노이즈 제거 모듈을 통해 상품 영역 추출 모듈에서 추출된 상품 영역을 가린 객체 영역을 찾아 해당 노이즈 영역을 상품 영역에서 제거하는 단계(730), 제거된 노이즈 영역에 의해 제거된 상품 영역을 노이즈 복구 모듈을 통해 복구하는 단계(740), 복구된 상품 영역에서 특징 추출 모듈을 통해 상품의 모양, 스타일, 색상 정보를 포함하는 특징을 추출하는 단계(750) 및 추출된 특징을 엘라스틱서치(Elasticsearch) 검색 엔진을 통해 벡터로 저장하고, 저장된 특징에 상응하는 상품을 검색하는 단계(760)를 포함한다.

[0061] 제안하는 딥러닝 기반 이미지 복구 기술을 이용한 패션 이미지 검색 방법에서 입력 이미지는 상품 영역 추출 모듈과 노이즈 제거 모듈에 동시에 입력된다.

[0062] 단계(710)에서, 상품 영역 추출 모듈을 통해 입력 이미지에서 검색을 위한 상품의 영역을 픽셀 단위로 추출한다. 상품 영역 추출 모듈은 시멘틱 세그멘테이션 모델을 통해 데이터셋에서 제공하는 상품의 시멘틱 세그멘테이션 주석을 사용하여 모델을 훈련하고, 상품의 모양에 기초하여 객체가 상품 영역을 가렸을 경우 객체를 상품 영역에 함께 포함하여, 상품 영역을 가린 객체 영역이 포함되어 있는 상품 영역을 추출한다.

[0063] 단계(720)에서, 노이즈 제거 모듈의 휴먼 파싱(human parsing) 모델을 통해 입력 이미지에서 사람의 팔, 다리, 머리, 얼굴을 포함하는 인체 구성을 각각 분해하고, 분해된 인체 구성의 각각의 객체 영역을 픽셀 단위로 추출한다.

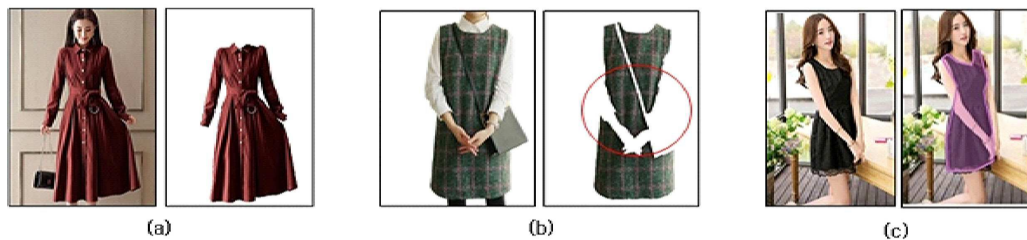
- [0064] 휴먼 파싱(human parsing) 모델을 통해 PGN(Part Grouping Network) 구조 및 LIP 데이터셋으로 미리 학습된 가중치를 사용한다. LIP 데이터셋은 복수의 시멘틱 인체 구성의 레이블을 표기한 복수의 이미지를 포함하고, LIP 데이터셋으로 미리 학습된 모델을 통해 입력 이미지에서 분해된 인체 구성의 각각의 객체 영역을 객체 별로 구분되고 픽셀 단위로 추출한다.
- [0065] 단계(730)에서, 노이즈 제거 모듈을 통해 상품 영역 추출 모듈에서 추출된 상품 영역을 가진 객체 영역을 찾아 해당 노이즈 영역을 상품 영역에서 제거한다.
- [0066] 상품 영역 추출 모듈에서 추출된 상품 영역과 휴먼 파싱 모델을 통해 추출된 객체 영역을 이용하여, 입력 이미지로부터 상품 영역을 변환하여 상품 마스크 이미지를 획득하고, 입력 이미지로부터 객체 영역을 변환하여 노이즈 마스크 이미지를 획득한다. 획득된 상품 마스크 이미지와 노이즈 마스크 이미지를 이용하여 상품 영역에서 노이즈를 제거한다.
- [0067] 단계(740)에서, 제거된 노이즈 영역에 의해 제거된 상품 영역을 노이즈 복구 모듈을 통해 복구한다.
- [0068] 노이즈 복구 모듈은 노이즈 마스크 이미지 및 노이즈 제거 모듈의 출력인 상품 영역을 가진 객체 영역이 상품 영역에서 제거된 상품 이미지를 입력으로 받아 CRA(Contextual Residual Aggregation) 구조 및 미리 훈련된 훈련된 가중치를 사용하여 노이즈 영역에 의해 제거된 상품 영역을 복구한다.
- [0069] 단계(750)에서, 복구된 상품 영역에서 특징 추출 모듈을 통해 상품의 모양, 스타일, 색상 정보를 포함하는 특징을 추출한다.
- [0070] 단계(760)에서 추출된 특징을 엘라스틱서치(Elasticsearch) 검색 엔진을 통해 벡터로 저장하고, 저장된 특징에 상응하는 상품을 검색한다.
- [0072] 이상에서 설명된 장치는 하드웨어 구성요소, 소프트웨어 구성요소, 및/또는 하드웨어 구성요소 및 소프트웨어 구성요소의 조합으로 구현될 수 있다. 예를 들어, 실시예들에서 설명된 장치 및 구성요소는, 예를 들어, 프로세서, 콘트롤러, ALU(arithmetic logic unit), 디지털 신호 프로세서(digital signal processor), 마이크로컴퓨터, FPA(field programmable array), PLU(programmable logic unit), 마이크로프로세서, 또는 명령(instruction)을 실행하고 응답할 수 있는 다른 어떠한 장치와 같이, 하나 이상의 범용 컴퓨터 또는 특수 목적 컴퓨터를 이용하여 구현될 수 있다. 처리 장치는 운영 체제(OS) 및 상기 운영 체제 상에서 수행되는 하나 이상의 소프트웨어 애플리케이션을 수행할 수 있다. 또한, 처리 장치는 소프트웨어의 실행에 응답하여, 데이터를 접근, 저장, 조작, 처리 및 생성할 수도 있다. 이해의 편의를 위하여, 처리 장치는 하나가 사용되는 것으로 설명된 경우도 있지만, 해당 기술분야에서 통상의 지식을 가진 자는, 처리 장치가 복수 개의 처리 요소(processing element) 및/또는 복수 유형의 처리 요소를 포함할 수 있음을 알 수 있다. 예를 들어, 처리 장치는 복수 개의 프로세서 또는 하나의 프로세서 및 하나의 콘트롤러를 포함할 수 있다. 또한, 병렬 프로세서(parallel processor)와 같은, 다른 처리 구성(processing configuration)도 가능하다.
- [0073] 소프트웨어는 컴퓨터 프로그램(computer program), 코드(code), 명령(instruction), 또는 이들 중 하나 이상의 조합을 포함할 수 있으며, 원하는 대로 동작하도록 처리 장치를 구성하거나 독립적으로 또는 결합적으로(collectively) 처리 장치를 명령할 수 있다. 소프트웨어 및/또는 데이터는, 처리 장치에 의하여 해석되거나 처리 장치에 명령 또는 데이터를 제공하기 위하여, 어떤 유형의 기계, 구성요소(component), 물리적 장치, 가상 장치(virtual equipment), 컴퓨터 저장 매체 또는 장치에 구체화(embodiment)될 수 있다. 소프트웨어는 네트워크로 연결된 컴퓨터 시스템 상에 분산되어서, 분산된 방법으로 저장되거나 실행될 수도 있다. 소프트웨어 및 데이터는 하나 이상의 컴퓨터 판독 가능 기록 매체에 저장될 수 있다.
- [0074] 실시예에 따른 방법은 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 실시예를 위하여 특별히 설계되고 구성된 것들이거나 컴퓨터 소프트웨어 당업자에게 공지되어 사용 가능한 것일 수도 있다. 컴퓨터 판독 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.

- [0075] 이상과 같이 실시예들이 비록 한정된 실시예와 도면에 의해 설명되었으나, 해당 기술분야에서 통상의 지식을 가진 자라면 상기의 기재로부터 다양한 수정 및 변형이 가능하다. 예를 들어, 설명된 기술들이 설명된 방법과 다른 순서로 수행되거나, 및/또는 설명된 시스템, 구조, 장치, 회로 등의 구성요소들이 설명된 방법과 다른 형태로 결합 또는 조합되거나, 다른 구성요소 또는 균등물에 의하여 대치되거나 치환되더라도 적절한 결과가 달성될 수 있다.
- [0076] 그러므로, 다른 구현들, 다른 실시예들 및 특허청구범위와 균등한 것들도 후술하는 특허청구범위의 범위에 속한다.
- [0078] <참고문헌>
- [0079] [1] E. Smirnov, A. Kulinkin, M. Pogrebnnyak, K. Ivanova, Deep Learning for Fast and Accurate Fashion Item Detection, KDD Fash. (2016) 1-4. <https://doi.org/10.475/123>.
- [0080] [2] H. Castro, M. Ramirez, Segmentation task for fashion and apparel, (2020). <http://arxiv.org/abs/2006.11375>.
- [0081] [3] J.W. Johnson, Adapting Mask-RCNN for Automatic Nucleus Segmentation, (2018) 1-7. <https://doi.org/10.1007/978-3-030-17798-0>.
- [0082] [4] W. Wang, Z. Zhang, S. Qi, J. Shen, Y. Pang, L. Shao, Learning compositional neural information fusion for human parsing, Proc. IEEE Int. Conf. Comput. Vis. 2019-October (2019) 5702-5712. <https://doi.org/10.1109/ICCV.2019.00580>.
- [0083] [5] K. Gong, X. Liang, D. Zhang, X. Shen, L. Lin, Look into Person: Self-supervised Structure-sensitive Learning and a new benchmark for human parsing, Proc. - 30th IEEE Conf. Comput. Vis. Pattern Recognition, CVPR 2017. 2017-January (2017) 6757-6765. <https://doi.org/10.1109/CVPR.2017.715>.
- [0084] [6] <http://sysu-hcp.net/lip/overview.php>. Accessed 2 Oct 2020
- [0085] [7] X. Zhang, Y. Chen, B. Zhu, J. Wang, M. Tang, Part-Aware Context Network for Human Parsing, (2020) 8968-8977. <https://doi.org/10.1109/cvpr42600.2020.00899>.
- [0086] [8] PASCAL-Part Dataset(2020) <http://roozbehm.info/pascal-parts/pascal-parts.html>. Accessed 5 Oct 2020
- [0087] [9] K. Gong, X. Liang, Y. Li, Y. Chen, M. Yang, L. Lin, Instance-Level Human Parsing via Part Grouping Network, Lect. Notes Comput. Sci. (Including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics). 11208 LNCS (2018) 805-822. [https://doi.org/10.1007/978-3-030-01225-0\\_47](https://doi.org/10.1007/978-3-030-01225-0_47).
- [0088] [10] R. Sharma, R. Deora, A. Vishvakarma, AlphaNet: An Attention Guided Deep Network for Automatic Image Matting, ArXiv. (2020).
- [0089] [11] Z. Yi, Q. Tang, S. Azizi, D. Jang, Z. Xu, Contextual Residual Aggregation for Ultra High-Resolution Image Inpainting, (2020) 7505-7514. <https://doi.org/10.1109/cvpr42600.2020.00753>.
- [0090] [12] <https://opencv.org/>. Accessed 16 Nov 2020
- [0091] [13] M. Claro, R. Veras, A. Santana, F. Araújo, R. Silva, J. Almeida, D. Leite, An hybrid feature space from texture information and transfer learning for glaucoma classification, J. Vis. Commun. Image Represent. 64 (2019). <https://doi.org/10.1016/j.jvcir.2019.102597>.
- [0092] [14] ImageNet(2020) <http://www.image-net.org/>. Accessed 5 Oct 2020
- [0093] [15] <https://keras.io/about/>. Accessed 13 Nov 2020
- [0094] [16] X. Guo, H. Wu, Y. Gao, S. Rennie, R. Feris, Fashion IQ: A New Dataset towards Retrieving Images by Natural Language Feedback, (2019). <http://arxiv.org/abs/1905.12794>.
- [0095] [17] W. Zhou, P.Y. Mok, Y. Zhou, Y. Zhou, J. Shen, Q. Qu, K.P. Chau, Fashion recommendations through cross-media information retrieval, J. Vis. Commun. Image Represent. 61 (2019) 112-120. <https://doi.org/10.1016/j.jvcir.2019.03.003>.

- [0096] [18] C. Reta, J.A. Cantoral-Ceballos, I. Solis-Moreno, J.A. Gonzalez, R. Alvarez-Vargas, N. Delgadillo-Checa, Color uniformity descriptor: An efficient contextual color representation for image indexing and retrieval, J. Vis. Commun. Image Represent. 54 (2018) 39-50. <https://doi.org/10.1016/j.jvcir.2018.04.009>.
- [0097] [19] L. Wang, D. Rajan, An image similarity descriptor for classification tasks, J. Vis. Commun. Image Represent. 71 (2020) 102847. <https://doi.org/10.1016/j.jvcir.2020.102847>.
- [0098] [20] G.M. Galshetwar, L.M. Waghmare, A.B. Gonde, S. Murala, Local energy oriented pattern for image indexing and retrieval, J. Vis. Commun. Image Represent. 64 (2019) 102615. <https://doi.org/10.1016/j.jvcir.2019.102615>.
- [0099] [21] S.G. Shaila, A. Vadivel, Indexing and encoding based image feature representation with bin overlapped similarity measure for CBIR applications, J. Vis. Commun. Image Represent. 36 (2016) 40-55. <https://doi.org/10.1016/j.jvcir.2016.01.003>.
- [0100] [22] Elasticsearch(2020) <https://www.elastic.co/kr/what-is/elasticsearch>. Accessed 3 Oct 2020

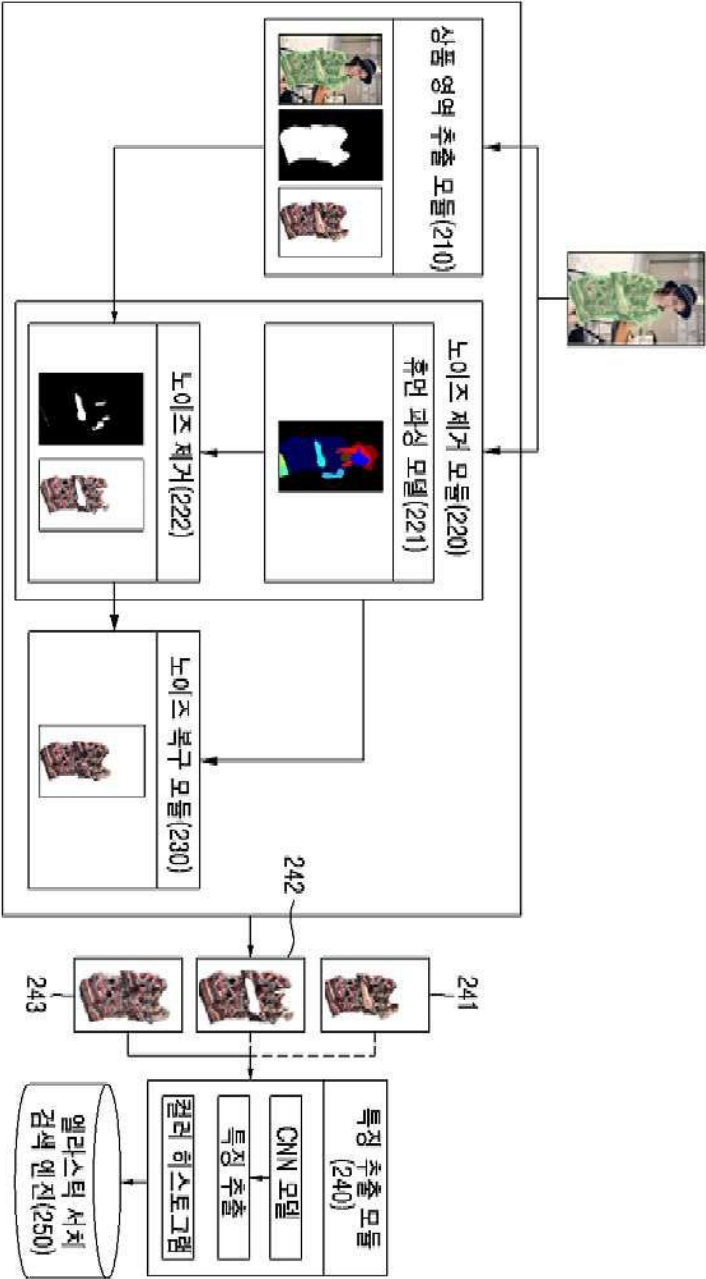
## 도면

### 도면1





도면2

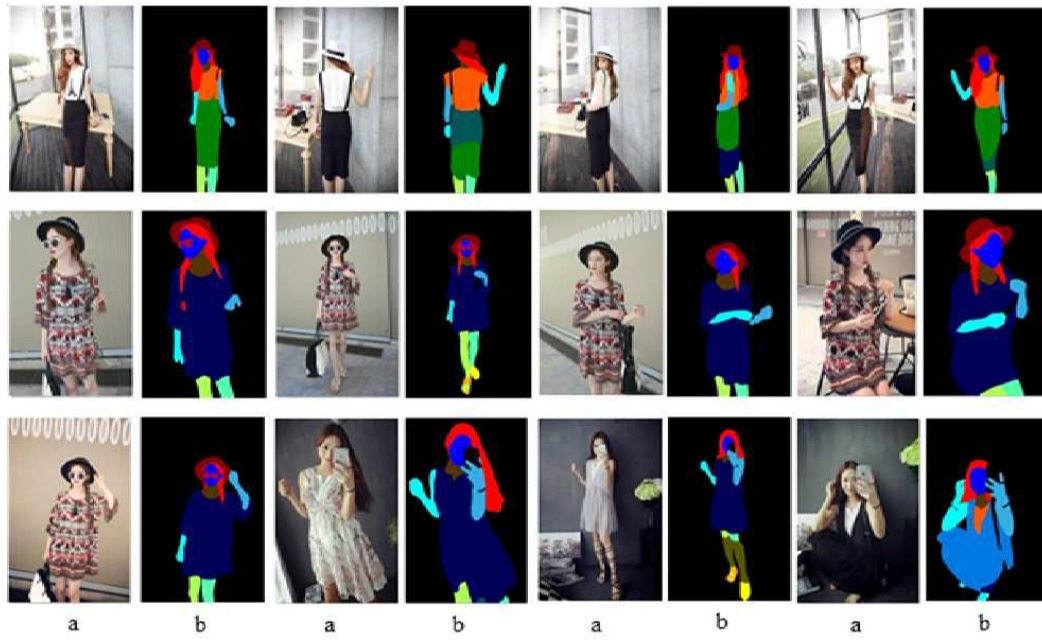


도면3





도면4



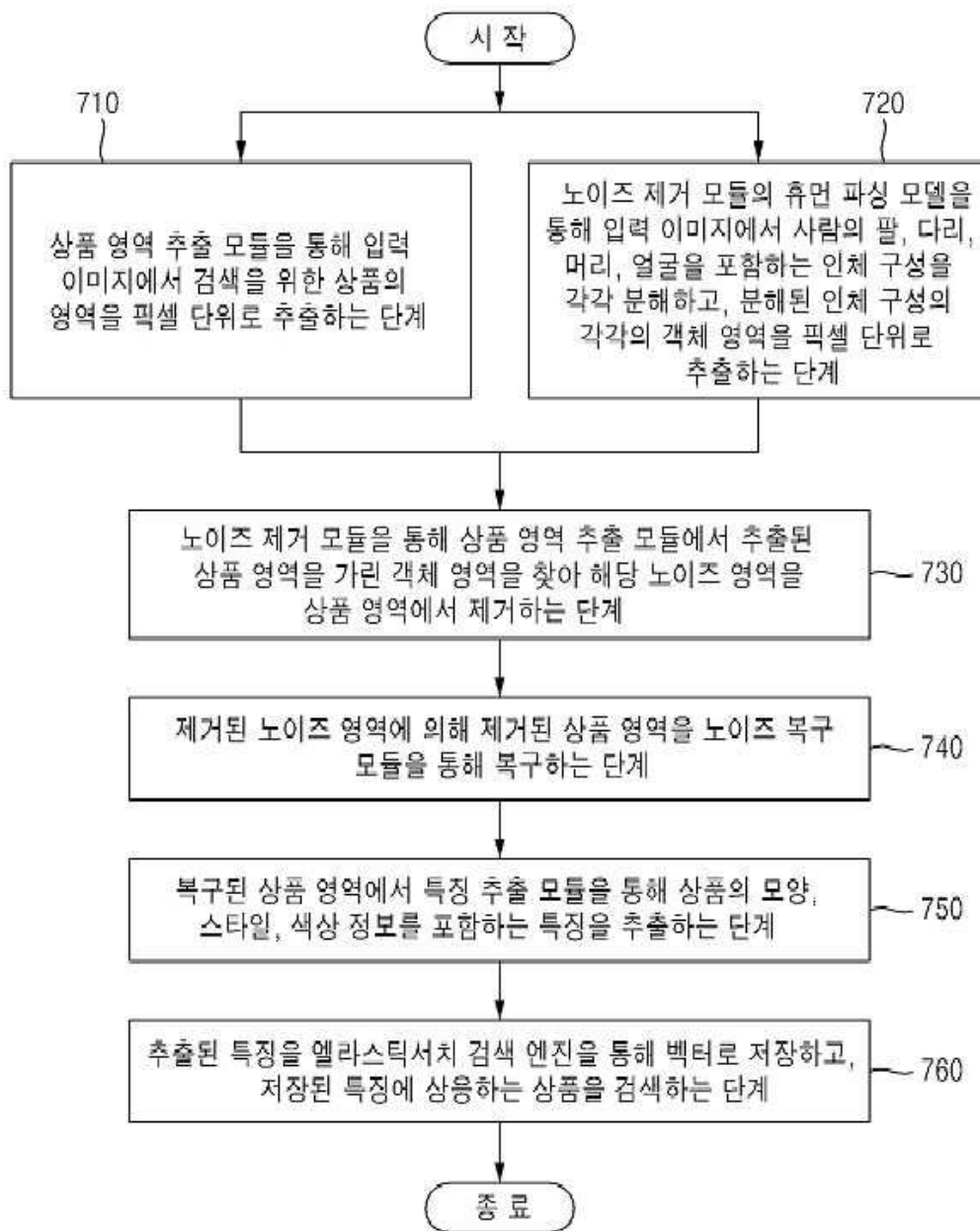
도면5



도면6



도면7



【심사관 직권보정사항】

【직권보정 1】

【보정항목】 청구범위

【보정세부항목】 청구항 1

【변경전】

입력 이미지에서 검색을 위한 상품의 영역을 픽셀 단위로 추출하는 상품 영역 추출 모듈;

입력 이미지에서 휴먼 파싱(human parsing) 모델을 통해 사람의 팔, 다리, 머리, 얼굴을 포함하는 인체 구성을 각각 분해하고, 분해된 인체 구성의 각각의 객체 영역을 픽셀 단위로 추출하고, 상품 영역 추출 모듈에서 추출된 상품 영역을 가린 객체 영역을 찾아 해당 노이즈 영역을 상품 영역에서 제거하는 노이즈 제거 모듈;

제거된 노이즈 영역에 의해 제거된 상품 영역을 복구하는 노이즈 복구 모듈;

복구된 상품 영역에서 상품의 모양, 스타일, 색상 정보를 포함하는 특징을 추출하는 특징 추출 모듈; 및

추출된 특징을 벡터로 저장하고, 저장된 특징에 상응하는 상품을 검색하는 엘라스틱서치(Elasticsearch) 검색 엔진

을 포함하고,

상품 영역 추출 모듈은,

Deepfashion2 데이터셋에서 제공되는 상품 세그멘테이션 주석을 사용하여 모델을 훈련하고, 상기 Deepfashion2 데이터셋 제공되는 상품 세그멘테이션 주석은 상품의 모양에 기초하여 객체가 상품 영역을 가렸을 경우 객체를 상품 영역에 함께 포함하여, 상품 영역을 가린 객체 영역이 포함되어 있는 상품 영역을 추출하고, 상기 Deepfashion2 데이터셋을 이용하여 시멘틱 세그멘테이션 기법으로 추출한 상품의 영역에서 노이즈를 찾도록 하며,

상기 노이즈 제거 모듈은,

휴먼 파싱 모델을 통해 PGN(Part Grouping Network) 구조 및 LIP(Look Into Person) 데이터셋으로 미리 학습된 가중치를 사용하고, LIP 데이터셋은 복수의 시멘틱 인체 구성의 레이블을 표기한 복수의 이미지를 포함하고, LIP 데이터셋으로 미리 학습된 모델을 통해 입력 이미지에서 분해된 인체 구성의 각각의 객체 영역을 객체 별로 구분되고 픽셀 단위로 추출하고,

상품 영역 추출 모듈에서 추출된 상품 영역과 휴먼 파싱(human parsing) 모델을 통해 추출된 객체 영역을 이용하여, 입력 이미지로부터 상품의 영역과 배경을 구분하고 상품 영역을 변환하여 상품 마스크 이미지를 획득하고, 입력 이미지로부터 휴먼 파싱을 통해 각각의 객체를 픽셀 단위로 인식하고 객체 영역을 변환하여 노이즈 마스크 이미지를 획득하며, 획득된 상품 마스크 이미지와 노이즈 마스크 이미지를 이용하여 상품 영역에서 노이즈를 제거하는

딥러닝 기반 이미지 복구 기술을 이용한 패션 이미지 검색 장치.

#### 【변경후】

입력 이미지에서 검색을 위한 상품의 영역을 픽셀 단위로 추출하는 상품 영역 추출 모듈;

입력 이미지에서 휴먼 파싱(human parsing) 모델을 통해 사람의 팔, 다리, 머리, 얼굴을 포함하는 인체 구성을 각각 분해하고, 분해된 인체 구성의 각각의 객체 영역을 픽셀 단위로 추출하고, 상품 영역 추출 모듈에서 추출된 상품 영역을 가린 객체 영역을 찾아 해당 노이즈 영역을 상품 영역에서 제거하는 노이즈 제거 모듈;

제거된 노이즈 영역에 의해 제거된 상품 영역을 복구하는 노이즈 복구 모듈;

복구된 상품 영역에서 상품의 모양, 스타일, 색상 정보를 포함하는 특징을 추출하는 특징 추출 모듈; 및

추출된 특징을 벡터로 저장하고, 저장된 특징에 상응하는 상품을 검색하는 엘라스틱서치(Elasticsearch) 검색 엔진

을 포함하고,

상품 영역 추출 모듈은,

Deepfashion2 데이터셋에서 제공되는 상품 세그멘테이션 주석을 사용하여 모델을 훈련하고, 상기 Deepfashion2 데이터셋에서 제공되는 상품 세그멘테이션 주석은 상품의 모양에 기초하여 객체가 상품 영역을 가렸을 경우 객체를 상품 영역에 함께 포함하여, 상품 영역을 가린 객체 영역이 포함되어 있는 상품 영역을 추출하고, 상기 Deepfashion2 데이터셋을 이용하여 시멘틱 세그멘테이션 기법으로 추출한 상품의 영역에서 노이즈를 찾도록 하며,

상기 노이즈 제거 모듈은,

휴먼 파싱 모델을 통해 PGN(Part Grouping Network) 구조 및 LIP(Look Into Person) 데이터셋으로 미리 학습된 가중치를 사용하고, LIP 데이터셋은 복수의 시멘틱 인체 구성의 레이블을 표기한 복수의 이미지를 포함하고, LIP 데이터셋으로 미리 학습된 모델을 통해 입력 이미지에서 분해된 인체 구성의 각각의 객체 영역을 객체 별로 구분되고 픽셀 단위로 추출하고,

상품 영역 추출 모듈에서 추출된 상품 영역과 휴먼 파싱(human parsing) 모델을 통해 추출된 객체 영역을 이용하여, 입력 이미지로부터 상품의 영역과 배경을 구분하고 상품 영역을 변환하여 상품 마스크 이미지를

획득하고, 입력 이미지로부터 휴먼 파싱을 통해 각각의 객체를 픽셀 단위로 인식하고 객체 영역을 변환하여 노이즈 마스크 이미지를 획득하며, 획득된 상품 마스크 이미지와 노이즈 마스크 이미지를 이용하여 상품 영역에서 노이즈를 제거하는

딥러닝 기반 이미지 복구 기술을 이용한 패션 이미지 검색 장치.

【식권보정 2】

【보정항목】 청구범위

【보정세부항목】 청구항 6

【변경전】

상품 영역 추출 모듈을 통해 입력 이미지에서 검색을 위한 상품의 영역을 픽셀 단위로 추출하는 단계;

노이즈 제거 모듈의 휴먼 파싱(human parsing) 모델을 통해 입력 이미지에서 사람의 팔, 다리, 머리, 얼굴을 포함하는 인체 구성을 각각 분해하고, 분해된 인체 구성의 각각의 객체 영역을 픽셀 단위로 추출하는 단계;

노이즈 제거 모듈을 통해 상품 영역 추출 모듈에서 추출된 상품 영역을 가린 객체 영역을 찾아 해당 노이즈 영역을 상품 영역에서 제거하는 단계;

제거된 노이즈 영역에 의해 제거된 상품 영역을 노이즈 복구 모듈을 통해 복구하는 단계;

복구된 상품 영역에서 특징 추출 모듈을 통해 상품의 모양, 스타일, 색상 정보를 포함하는 특징을 추출하는 단계; 및

추출된 특징을 엘라스틱서치(Elasticsearch) 검색 엔진을 통해 벡터로 저장하고, 저장된 특징에 상응하는 상품을 검색하는 단계

를 포함하고,

상품 상품 영역 추출 모듈을 통해 입력 이미지에서 검색을 위한 상품의 영역을 픽셀 단위로 추출하는 단계는,

Deepfashion2 데이터셋에서 제공되는 상품 세그멘테이션 주석을 사용하여 모델을 훈련하고, 상기 Deepfashion2 데이터셋 제공되는 상품 세그멘테이션 주석은 상품의 모양에 기초하여 객체가 상품 영역을 가렸을 경우 객체를 상품 영역에 함께 포함하여, 상품 영역을 가린 객체 영역이 포함되어 있는 상품 영역을 추출하고, 상기 Deepfashion2 데이터셋을 이용하여 시멘틱 세그멘테이션 기법으로 추출한 상품의 영역에서 노이즈를 찾도록 하며,

상기 노이즈 제거 모듈의 휴먼 파싱(human parsing) 모델을 통해 입력 이미지에서 사람의 팔, 다리, 머리, 얼굴을 포함하는 인체 구성을 각각 분해하고, 분해된 인체 구성의 각각의 객체 영역을 픽셀 단위로 추출하는 단계는,

휴먼 파싱 모델을 통해 PGN(Part Grouping Network) 구조 및 LIP(Look Into Person) 데이터셋으로 미리 학습된 가중치를 사용하고, LIP 데이터셋은 복수의 시멘틱 인체 구성의 레이블을 표기한 복수의 이미지를 포함하고, LIP 데이터셋으로 미리 학습된 모델을 통해 입력 이미지에서 분해된 인체 구성의 각각의 객체 영역을 객체 별로 구분되고 픽셀 단위로 추출되고,

상품 영역 추출 모듈에서 추출된 상품 영역과 휴먼 파싱(human parsing) 모델을 통해 추출된 객체 영역을 이용하여, 입력 이미지로부터 상품의 영역과 배경을 구분하고 상품 영역을 변환하여 상품 마스크 이미지를 획득하고, 입력 이미지로부터 휴먼 파싱을 통해 각각의 객체를 픽셀 단위로 인식하고 객체 영역을 변환하여 노이즈 마스크 이미지를 획득하며, 획득된 상품 마스크 이미지와 노이즈 마스크 이미지를 이용하여 상품 영역에서 노이즈를 제거되는

딥러닝 기반 이미지 복구 기술을 이용한 패션 이미지 검색 방법.

【변경후】

상품 영역 추출 모듈을 통해 입력 이미지에서 검색을 위한 상품의 영역을 픽셀 단위로 추출하는 단계;

노이즈 제거 모듈의 휴먼 파싱(human parsing) 모델을 통해 입력 이미지에서 사람의 팔, 다리, 머리, 얼굴을 포함하는 인체 구성을 각각 분해하고, 분해된 인체 구성의 각각의 객체 영역을 픽셀 단위로 추출하는 단계;

노이즈 제거 모듈을 통해 상품 영역 추출 모듈에서 추출된 상품 영역을 가린 객체 영역을 찾아 해당 노이즈 영



역을 상품 영역에서 제거하는 단계;

제거된 노이즈 영역에 의해 제거된 상품 영역을 노이즈 복구 모듈을 통해 복구하는 단계;

복구된 상품 영역에서 특징 추출 모듈을 통해 상품의 모양, 스타일, 색상 정보를 포함하는 특징을 추출하는 단계; 및

추출된 특징을 엘라스틱서치(Elasticsearch) 검색 엔진을 통해 벡터로 저장하고, 저장된 특징에 상응하는 상품을 검색하는 단계

를 포함하고,

상품 상품 영역 추출 모듈을 통해 입력 이미지에서 검색을 위한 상품의 영역을 픽셀 단위로 추출하는 단계는,

Deepfashion2 데이터셋에서 제공되는 상품 세그멘테이션 주석을 사용하여 모델을 훈련하고, 상기 Deepfashion2 데이터셋에서 제공되는 상품 세그멘테이션 주석은 상품의 모양에 기초하여 객체가 상품 영역을 가렸을 경우 객체를 상품 영역에 함께 포함하여, 상품 영역을 가린 객체 영역이 포함되어 있는 상품 영역을 추출하고, 상기 Deepfashion2 데이터셋을 이용하여 시멘틱 세그멘테이션 기법으로 추출한 상품의 영역에서 노이즈를 찾도록 하며,

상기 노이즈 제거 모듈의 휴먼 파싱(human parsing) 모델을 통해 입력 이미지에서 사람의 팔, 다리, 머리, 얼굴을 포함하는 인체 구성을 각각 분해하고, 분해된 인체 구성의 각각의 객체 영역을 픽셀 단위로 추출하는 단계는,

휴먼 파싱 모델을 통해 PGN(Part Grouping Network) 구조 및 LIP(Look Into Person) 데이터셋으로 미리 학습된 가중치를 사용하고, LIP 데이터셋은 복수의 시멘틱 인체 구성의 레이블을 표기한 복수의 이미지를 포함하고, LIP 데이터셋으로 미리 학습된 모델을 통해 입력 이미지에서 분해된 인체 구성의 각각의 객체 영역을 객체 별로 구분되고 픽셀 단위로 추출되고,

상품 영역 추출 모듈에서 추출된 상품 영역과 휴먼 파싱(human parsing) 모델을 통해 추출된 객체 영역을 이용하여, 입력 이미지로부터 상품의 영역과 배경을 구분하고 상품 영역을 변환하여 상품 마스크 이미지를 획득하고, 입력 이미지로부터 휴먼 파싱을 통해 각각의 객체를 픽셀 단위로 인식하고 객체 영역을 변환하여 노이즈 마스크 이미지를 획득하며, 획득된 상품 마스크 이미지와 노이즈 마스크 이미지를 이용하여 상품 영역에서 노이즈를 제거되는

딥러닝 기반 이미지 복구 기술을 이용한 패션 이미지 검색 방법.