

# 객체 이미지를 이용한 객체간 관계 기반의 내용기반 이미지 검색 기법

맹주형\*<sup>1</sup>, 김성종\*<sup>2</sup>, 송석일\*\*

## Content based Image Retrieval Method based on Relationship Between Objects using Object Images

Juhyeong Maeng\*<sup>1</sup>, Seongjong Kim\*<sup>2</sup>, and Seokil Song\*\*

---

This work is supported by the Korea Agency for Infrastructure Technology Advancement(KAIA) grant funded by the Ministry of Land, Infrastructure and Transport(Grant 22AMDP-C160549-02)

---

### 요 약

최근 내용 기반 이미지 검색(Content based image retrieval)을 위한 특징 벡터(Feature vector) 추출에 관한 연구는 주로 DCNN(Deep Convolution Neural Network) 기술을 기반으로 하고 있다. DCNN 기반의 특징 벡터 추출 방법들의 문제점은 이미지에 포함된 객체들의 클래스와 객체들 간의 위치 관계를 잘 반영하지 못한다는 것이다. 이 논문에서는 이미지 전체 특징(Global feature), 이미지의 객체 목록(Object list), 객체 관계 특징(Object relation feature)를 각각 추출하고 이 특징들을 결합하는 새로운 특징 추출 방법을 제안한다. 특히, 객체 관계 특징 추출을 위해 객체 이미지라는 새로운 개념을 제안한다. 또한, 이 논문에서는 10만 개의 MS COCO 데이터 집합을 이용해서 기존 방법과 제안하는 방법을 비교하는 실험을 수행한다.

### Abstract

Recently, feature extraction methods for contents-based image retrieval method is generally based on DCNN(Deep Convolution Neural Network) technologies. A problem of the existing DCNN based feature vector extraction methods is that they do not reflect the class of objects of the image and the positional relationship between the objects well. In this paper, a new feature extraction method that extracts global features, object lists, and object relation features from images, and combines these features is proposed. In particular, a new concept of object image to extract object relational features is also proposed. Finally, in this paper, an experiment comparing the existing method and the proposed method is performed using 100,000 MS COCO data sets.

### Keywords

contents based image retrieval, CNN, object image, image feature

---

\* 한국교통대학교 컴퓨터공학과  
- ORCID<sup>1</sup>: <https://orcid.org/0000-0001-5152-3213>  
- ORCID<sup>2</sup>: <https://orcid.org/0000-0003-1589-728X>  
\*\* 한국교통대학교 컴퓨터공학과 교수(교신저자)  
- ORCID: <https://orcid.org/0000-0002-0110-7155>

• Received: Dec. 19, 2022, Revised: Jan. 13, 2023, Accepted: Jan. 16, 2023  
• Corresponding Author: Seokil Song  
School of Computer, Korea National University of Transportation,  
Daehakro 50, Chungju, Chungbuk 27469, Korea  
Tel.: +82-43-841-5349, Email: [sisong@ut.ac.kr](mailto:sisong@ut.ac.kr)

## 1. 서 론

내용 기반 이미지 검색 방법은 질의 이미지의 특징과 데이터베이스에 저장된 이미지 특징들을 비교하여 질의 이미지와 가장 유사한 이미지를 검색하는 것이다[1]. 내용 기반 이미지 검색의 정확도는 질의 및 데이터베이스 이미지 특징 추출 방법과 이에 적합한 유사도 측정 방법에 의해 결정된다.

기존의 내용 기반 이미지 특징 추출 방법은 수작업 기반의 특징 추출 방법과 DCNN(Deep Convolutional Neural Network) 기반 특징 추출 방법으로 분류될 수 있다. 수작업 기반의 특징 추출 방법에는 SIFT(Scale-Invariant Feature Transform)[2]와 같은 방법이 대표적이다.

DCNN 기반의 특징 추출 방법에는 [3]-[8]와 같은 방법들이 존재한다. 이 방법들은 DCNN 모델의 특정 계층(Layer)의 출력을 이미지의 특징 벡터로 사용한다. [3]에서는 VGGNet[8]을 사용하여 이미지의 전체 특징(Global feature)을 추출한다. DCNN 기반 특징 추출 방법은 이미지 전체 특징을 추출할 때 VGGNet의 최대 풀링계층(Max pooling layer)과 활성화 함수 계층(Activation function layer)의 출력을 사용한다. [4]에서는 객체 탐지 모델을 통해 추출한 객체의 바운딩 박스(Bounding box) 부분만 DCNN 모델을 통해 특징 벡터를 추출하여 지역 특징(Local feature)으로 사용한다. SPoC[5]에서는 마지막 컨볼루션 계층(Convolution layer)과 연결된 활성화 함수 계층의 출력에 다양한 풀링(Pooling)을 수행하여 추출된 특징 벡터를 이미지의 전체 특징으로 사용한다. [6]에서도 마지막 컨볼루션 계층에 연결된 다음 계층, 활성화 함수 계층의 출력에 다양한 풀링을 수행한다. 풀링된 벡터를 다시 짧은 벡터로 변환하여 특징 벡터로 사용한다. [7]에서는 VGGNet의 완전 연결 계층에서 이미지 전체 특징을 추출하고 객체 탐지 모델들 [9][10] 중 Yolo v4[9]를 통해 객체 클래스 목록(Object class list)을 추출한다. 추출된 이미지 전체 특징과 객체 클래스 목록을 특징 벡터로써 사용한다.

이상 기술한 DCNN 기반의 이미지 특징 추출 방법들은 이미지 전체의 특징 및 객체의 형태, 객체의 클래스를 고려하고 있으나 객체 간의 관계나 객체

의 위치, 크기를 고려하지 못한다. 본 논문에서는 DCNN 기반의 이미지 분류기와 객체 탐지 기법을 이용하여 이미지 전체 특징, 객체 클래스 목록 그리고 객체 관계 특징을 결합하는 새로운 이미지 특징 추출 방법을 제안한다. 또한, 제안하는 이미지 특징 추출 방법에 적합한 유사도 비교 방법을 같이 제안한다.

이미지 전체 특징은 VGGNet에서 원본 이미지를 입력으로 받아 추출된다. 이때, 이미지 전체 특징은 VGGNet의 완전 연결 계층에서 추출되며 4,096의 크기를 가진다. 객체 클래스 목록은 Yolo v4에서 원본 이미지를 입력으로 받아 추출된다. Yolo v4의 출력 계층에서 바운딩 박스(Bounding box), 객체의 클래스(Object class), 신뢰도 점수(Confidence score)를 추출한다. 이 중, 객체의 클래스 값들이 객체 목록으로 사용된다. 이상 기술한 이미지 전체 특징과 객체 목록 추출 방법은 기존 방법[7]과 같다.

본 논문에서는 이미지 간 유사도 측정 시, 객체 간의 관계를 고려하기 위해서 객체 관계 특징을 추출한다. 객체 관계 특징은 추출된 객체 목록을 이용하여 객체 이미지를 생성한 후 이를 VGGNet을 이용해 생성한다.

객체 이미지는 객체의 바운딩 박스 안에 포함된 모든 픽셀(Pixel) 값을 객체의 클래스 값으로 채우고 배경은 0 값으로 채워진 원본 이미지와 동일한 크기의 이미지이다. 이때, 두 가지 이상의 객체 영역이 겹칠 때는 신뢰도 점수가 높은 객체의 클래스 값을 우선으로 사용한다. 이를 VGGNet에 입력으로 주어 완전 연결 계층에서 4,096차원의 객체 관계 특징을 추출한다. 또한, 추출한 특징 간의 유사도 측정 방법을 제안한다. 이미지 전체 특징 및 객체 관계 특징의 유사도는 RBF(Radial Basis Function)[11]를 통해 계산된다. 그리고 객체 관계 특징은 Jaccard[12]를 통해 계산된다. 그리고 계산된 각 유사도 값은 가중치와 결합하여 두 이미지 간의 최종 유사도 값을 계산한다.

그림 1은 (a)의 질의 이미지에 대해 (b) 이미지 전체 특징, (c) 객체 클래스 목록 특징을 이용해 이미지 데이터베이스에서 유사 이미지를 검색했을 때 (d) 가장 유사한 이미지가 검색된 결과를 보여준다.



(a) 질의 이미지 (b) 이미지 전체 특징  
(a) Query image (b) Global feature based query results



(c) 객체 클래스 목록 특징 (d) 가장 유사한 이미지  
(c) Object list based query results (d) Most similar image  
그림 1. 이미지 전체 특징과 객체 목록을 이용한 질의 비교

Fig. 1. Comparison of global feature based query and object list based query

질의 이미지의 전체 특징은 중앙을 가로지르는 벤치와 벤치 위의 개, 녹색의 잔디밭이라고 할 수 있다. 이 질의의 검색 결과인 그림 1의 (b)를 보면 테니스 코트의 잔디가 경계선과 경계선 위쪽의 사람이 유사한 특징으로 검색된 것으로 보인다. 그러나 이 결과는 검색된 이미지의 객체들이 서로 다르며 객체 간의 관계도 고려되지 않았다.

질의 이미지에 대한 객체 클래스 특징 질의는 이미지에 질의 이미지와 동일한 클래스의 객체가 많이 포함된 이미지를 유사 이미지로 검색한다. 그림 1의 (b)에서와 같이 동일한 클래스 비율이 높은 유사한 이미지를 잘 찾아내었으나 개와 벤치 간의 위치 관계가 다를 수 있다. 이미지 전체 특징, 객체 목록에 객체 들 간의 위치 관계를 추가로 고려해서 특징을 추출할 수 있다면 그림 1의 (d)와 같이 보다 정확한 내용 기반 유사 이미지 검색이 가능해질 수 있다.

이 논문에서는 이미지 내 객체 간의 관계를 반영할 수 있는 특징 추출 방법을 제안한다. 또한, 이미지 객체 관계 특징과 함께, 이미지 전체 특징, 이미지 객체 목록 특징을 결합하는 새로운 이미지 특징 추출 방법을 제안한다. 마지막으로, 제안하는 특징

추출 방법과 이미지 전체 특징 및 객체 목록을 이용한 유사 이미지 검색 결과를 비교하여 제안하는 방법이 보다 유사한 이미지를 잘 검색할 수 있음을 보인다.

## II. 관련 연구

여기에서는 DCNN을 이용한 특징 추출 방법 중 대표적인 연구들에 관해서 설명한다. 먼저, [4]에서는 이미지의 특징을 추출하기 위해 객체 인식 모델을 이용하여 객체를 추출한 후 객체 바운딩 박스를  $224 \times 224$  크기의 이미지를 생성한다. 생성된 이미지를 VGGNet의 입력으로 하여 완전 연결 계층에서 추출된 4,096 크기의 특징 벡터를 지역 특징으로 사용한다.

[7]에서 제안하는 방법은 VGGNet의 완전 연결 계층의 출력을 이미지 전체 특징으로 사용하고, Yolo의 출력 계층에서 추출한 객체 정보의 객체 클래스 목록을 이미지 전체 특징과 결합하는 이미지 특징 추출 방법을 제안하고 있다. 추출된 이미지 특징의 유사도는 RBF와 Jaccard를 이용하여 계산한다. [13]에서는 공간 풀링 계층(Spatial pooling layer)을 통해서 지역 특징에 해당하는 서브-패치(Sub-patch)를 생성한다. [13]에서는 이때 생성되는 서브-패치를 이미지의 특징벡터로 사용하며 특징 벡터간 유사도 계산은 유클리드 거리 측정 방법(Euclidean distance metric)을 사용한다. [14]에서는 이미지와 이미지에 대한 캡션(Caption) 데이터를 함께 사용하여 특징을 추출한다. 이 방법에서는 이미지에 대한 캡션 데이터를 스킵-그램 모델(Skip-gram model) 통해 임베딩한다. 그리고 원본 이미지를 컨볼루션 계층과 완전 연결 변환 계층(Fully connected transform layer)순으로 통과시켜 임베딩을 수행한다. 캡션 임베딩 값과 이미지 임베딩 값을 이용해 손실(Loss)을 계산하여 유사도를 학습한다.

## III. 제안하는 이미지 특징과 객체 관계를 고려한 내용 기반 검색 기법

본 논문의 내용 기반 이미지 검색을 위한 전체 과정은 그림 2와 같다.

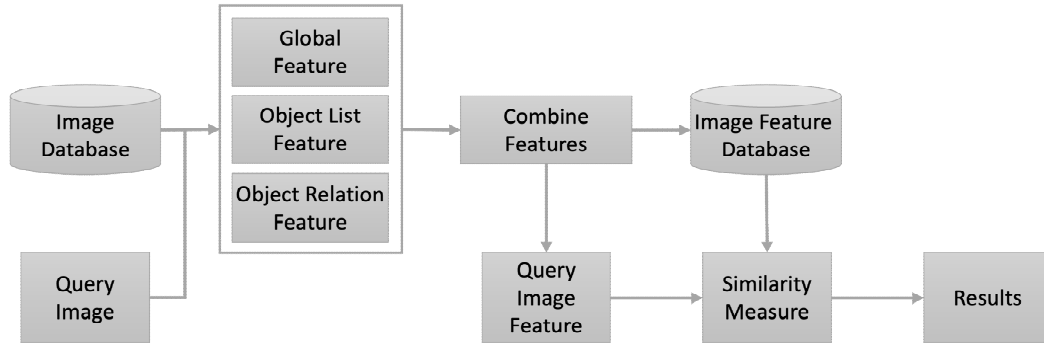


그림 2. 제안하는 내용 기반 이미지 검색 방법의 전체 구조  
 Fig. 2. Overall architecture of the proposed contents based image retrieval method

가장 먼저 이미지 데이터베이스(Image database)의 각 이미지에 대해서 전역 특징(Global feature), 객체 목록 특징(Object list feature), 객체 관계 특징(Object relation feature)를 추출하고 이를 결합하여 이미지 특징 데이터베이스(Image feature database)를 구축한다. 질의 처리를 위해서도 질의 이미지(Query image)에 대해서 3가지 특징을 각각 추출하고 이를 결합하여 질의 이미지 특징(Query image feature)을 추출한다. 추출한 질의 이미지 특징에 대해서 이미지 특징 데이터베이스와 비교를 수행하여 최종 결과를 도출한다.

제안하는 내용 기반 이미지 검색 방법은 이처럼 3가지 특징(전역특징, 객체 관계 특징, 객체 목록 특징)을 모두 반영하여 보다 정확하게 유사 이미지를 검색할 수 있도록 한다. 특히, 유사도를 비교할 때는 전역 특징, 객체 관계 특징, 객체 목록 특징 각각에 가중치를 부여하여 상황에 따라 더 우선해야 하는 특징을 선택할 수 있다.

### 3.1 특징 추출 방법

이미지 전체 특징은 기존에 잘 알려진 DCNN 방법인 VGG16을 통해 추출한다. 원본 이미지를 VGG16 입력으로 주어 완전 연결 계층에서 특징 벡터 (4,096 차원)를 추출하고 이를 이미지 전체 특징을 사용한다. VGG16 모델의 경우 이미지 분류기로써 사전 훈련된 모델을 사용한다.

객체 목록 특징은 객체 인식 모델인 Yolo의 헤드(Head)에서 추출한다. Yolo 헤드에서는 이미지에서 인식한 객체 클래스, 영역, 신뢰도 점수를 추출한다.

이미지에서 추출된 객체 클래스 값은 객체의 종류를 의미하며 일정 범위의 정수 값을 갖는다. 한 이미지에서 추출된 객체 클래스 목록을 객체 목록 특징으로 활용한다.

Yolo를 통해 추출한 각 객체의 바운딩 박스는 인식한 객체를 최소로 감싸는 사각형을 의미한다. Yolo의 바운딩 박스는 사각형의 가로 중심 값, 세로 중심 값, 넓이, 높이로 표현된다. 신뢰도 점수는 Yolo가 인식한 객체가 실제로 존재할 수 있는 확률을 의미한다. 신뢰도 점수는 0~1 사이의 값을 가지며 1에 가까울수록 신뢰도가 높음을 의미한다. Yolo를 통해 추출한 바운딩 박스, 신뢰도 점수, 클래스 값은 객체 관계 특징을 추출하기 위한 객체 이미지 생성에 사용된다. 객체 이미지는 객체의 바운딩 박스의 픽셀을 객체의 클래스 값에 대응하는 색으로 채우고 배경 픽셀은 0 으로 채운 원본 이미지와 동일한 크기의 이미지이다. 이때, 두 개 이상의 객체 바운딩 박스가 겹칠 때는 신뢰도 점수가 높은 객체의 클래스 값을 우선으로 사용한다.

객체 관계 특징은 객체 이미지를 VGG16의 입력으로 하여 추출된 특징 벡터이다. 객체 이미지 생성 및 객체 관계 특징 추출 과정을 그림 3에서 보여주고 있다. 이 그림에서 객체 이미지는 임의로 클래스 값에 100을 더하여 식별이 가능하도록 한 것이다.

추출된 세 가지 특징을 결합하여 최종 이미지 특징을 생성한다. 이 논문에서는 이미지 전체 특징 (4096), 객체 목록 특징(100), 객체 관계 특징(4096)이 연결되어 크기가 8,292인 1차원 벡터로 만들어진다. 그리고 이렇게 연결된 특징을 이미지의 특징으로 사용한다.

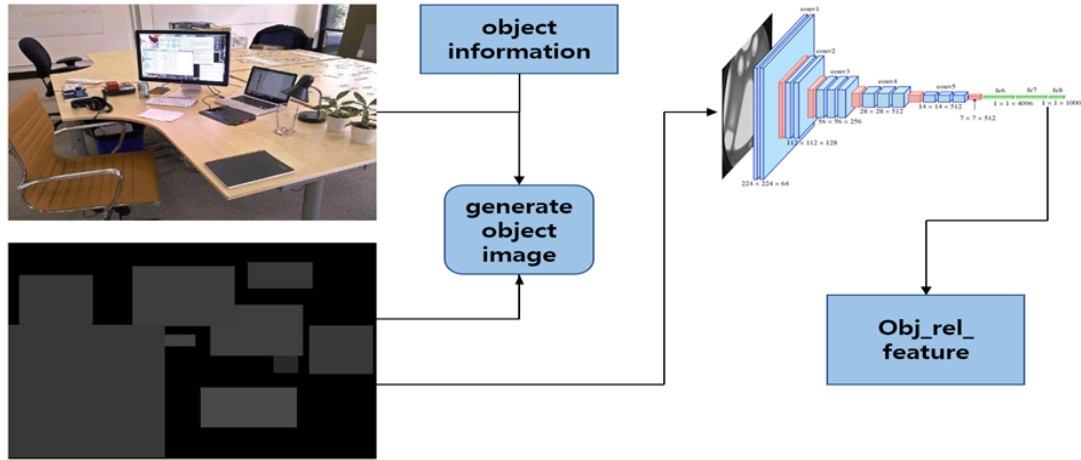


그림 3. 객체 관계 특징 추출 과정  
Fig. 3. Process of object relationship feature vector

### 3.2 추출된 특징의 유사도 측정 방법

두 이미지에서 추출된 특징 간의 유사도는 그림 4와 같이 계산된다.  $f1, f2$ 는 비교하려는 두 이미지의 특징 값을 의미한다. 그리고  $f1.global\_feature$ 는  $f1$ 의 이미지 전체 특징을 의미한다.  $f1.object\_list$ 는  $f1$ 의 객체 목록 특징을 의미한다.  $f1.obj\_relation\_feature$ 은  $f1$ 의 객체 관계 특징을 의미한다.  $W$ 는 가중치를 의미하며,  $W_a$ 는 이미지 전체 특징 가중치,  $W_b$ 는 객체 목록 특징 가중치,  $W_c$ 는 객체 관계 특징 가중치를 의미한다.

각 이미지의 특징 중 이미지 전체 특징과 객체 관계 특징 간의 유사도는 RBF를 이용해서 계산하며 객체 목록 특징 간의 유사도는 Jaccard를 통해 계산한다. 각 특징에 대한 유사도 점수는 0~1 사이의 단일 값으로 표현되며 1에 가까울수록 유사도가 높음을 의미한다. 특징별로 계산된 유사도 값들 각각에 가중치를 곱하고 합산하여 최종 이미지 간 유사도를 계산한다. 이때 세 가지의 가중치 합은 1이며 1에 가까울수록 유사도가 높음을 의미한다. 가중치를 설정에 따라서 두 이미지의 유사도 비교시 세 가지 특징 중 어떤 특징을 더 중요하게 여길지 결정할 수 있다.

$$similarity(f1, f2) = RBF(f1.global\_feature, f2.global\_feature) \times W_a + Jaccard(f1.object\_list, f2.object\_list) \times W_b + RBF(f1.obj\_relation\_feature, f2.obj\_relation\_feature) \times W_c$$

그림 4. 이미지의 특징 간 유사도 측정 방법  
Fig. 4. Proposed similarity measure method

## IV. 성능 평가

### 4.1 실험 환경

본 논문에서는 MS COCO[15] 데이터 집합 일부를 이용해서 이미지 검색 실험 및 성능 평가를 수행하였다. MS COCO 데이터 집합의 객체 클래스 수는 80개이며, 2017년 MS COCO 데이터 집합 기준으로 학습 데이터 118k건, 검증 데이터 5k건, 테스트 데이터 41k건으로 구성되어있다.

실험에서는 학습 데이터 이미지 집합에서는 총 100k건의 이미지(학습 데이터 59k건, 테스트 데이터 41k건)를 이용하였다. 실험은 이미지 전체 특징과 객체 목록 특징을 사용했을 때와 세 가지 특징 모두를 사용한 경우에 대해 비교하여 제안하는 방법이 정확도가 더 높음을 보인다. 표 1은 본 논문의 실험 환경을 보여준다.

표 1. 실험 환경  
Table 1. Experimental environment

Division	Content
Hardware	Intel(R) Core(TM) i7-7700 CPU @ 3.60GHz, GPU Titan XP, Memory DDR4 64GB
Software	CUDA 10.1, CUDNN 7.6.4, Python 3.6.9, Tensorflow 2.3.4
Data set	COCO 2017 train & test dataset, 10K images

### 4.2 실험 방법

실험은 다음과 같이 진행되었다. 총 100k건의 이미지 데이터 중에서 200건의 이미지를 임의로 선정하였다. 그리고 선정된 이미지 중에서 되도록 서로 다른 클래스의 객체들로 구성되는 이미지 50건을 질의 이미지로 선정하였다.

성능 평가를 위해 선정된 50개의 질의 이미지에 대해 3가지 종류의 k-NN(Nearest Neighbor) 질의 (5-NN, 10-NN, 20-NN)를 수행한다. k-NN 질의는 질의 이미지와 유사도가 가장 높은 k개의 이미지를 찾는 질의이다. 정확도는 각 질의 별로 상위 k개의 이미지와 질의 이미지의 특징을 직접 육안으로 확인하여 일치 여부를 확인하고 질의별로 정확도 평균을 계산한다.

정확도는 상위 k개의 검색 결과에 질의 이미지의 특징과 유사한 특징을 갖는 이미지가 몇 개 포함되어 있는지로 계산한다. 정확한 정확도 계산을 위해서는 질의 이미지와 유사한 실측(Ground truth) 이미지들을 사전에 알고 있어야 하나, 10만 개의 데이터에서 이들을 정확하게 분류해 내는 것이 어려워 배제하였다. 실험에서는 이미지의 전체 특징과 객체 목록 특징만을 이용한 기존 방법과 객체 관계 특징까지 더해 세 가지 특징을 활용하는 방법의 정확도를 비교하였다.

### 4.3 실험 결과

본 논문에서는 그림 4의  $W_a$ (이미지 전체 특징에 대한 유사도 가중치),  $W_b$ (객체 목록에 대한 유사도 가중치),  $W_c$ (객체 관계 특징에 대한 유사도 가중치)를 각각 0.2, 0.2, 0.6으로 설정하여 실험하였다. 이때의 정확도는 표 2에서처럼 이미지 전체 특징만을 사용했을 때보다 Top 5 기준 8.2%, Top 10 기준 8.6%, Top 20 기준 3.6% 정확도 향상을 볼 수 있다.

여러 번의 실험을 거쳐, 제안하는 방법에서는 유사한 이미지를 찾을 때 객체의 목록이 가장 중요하다고 판단하였다. 상위 20개 특징의 유사도 값을 고려하여서 객체의 클래스 유사도 값이 다른 특징의

유사도 값보다 우선으로 고려될 수 있으면서도 가능한 작은  $W_b$  값을 직접 설정하여 주었다. 또한 이미지 전체 특징에 대한 유사도 값보다 객체 관계 특징에 대한 유사도 값의 비중을 더 높이는 것이 높은 성능을 가지는 것을 확인하였다.

질의를 통해 얻은 이미지 20장은 질의 이미지마다 100k 장 각각이 유사한 이미지인지 판단하지 않아, 실제로 유사 이미지가 없어 검색되지 않은 것인지 유사한 이미지가 존재하는데도 검색되지 않은 것인지 정확하게 알 수 없다. 때문에 Top-5, Top-10, Top-20 순으로 정확도가 점점 내려가는 이유는 질의 이미지마다 유사한 이미지 개수 편차가 있음과 동시에 평균적으로 유사한 이미지의 수가 적은 것으로 유추하고 있다.

그림 4는 이 실험에서 사용한 질의의 예시를 보여주고 있다. 이 그림의 질의 이미지는 두 마리 이상 무리를 이룬 코끼리, 이미지 1/3 정도를 차지하는 코끼리 무리의 특징을 가진다. 이 실험에서는 그림 4의 질의 특징처럼 각 질의의 특징을 정하고 질의 결과에 대해 육안으로 이 특징과 부합하는지를 확인하였다.

표 2. 기존 방법과 제안하는 방법 질의 성능 비교  
Table 2. Comparison of query performance for the existing method and the proposed method

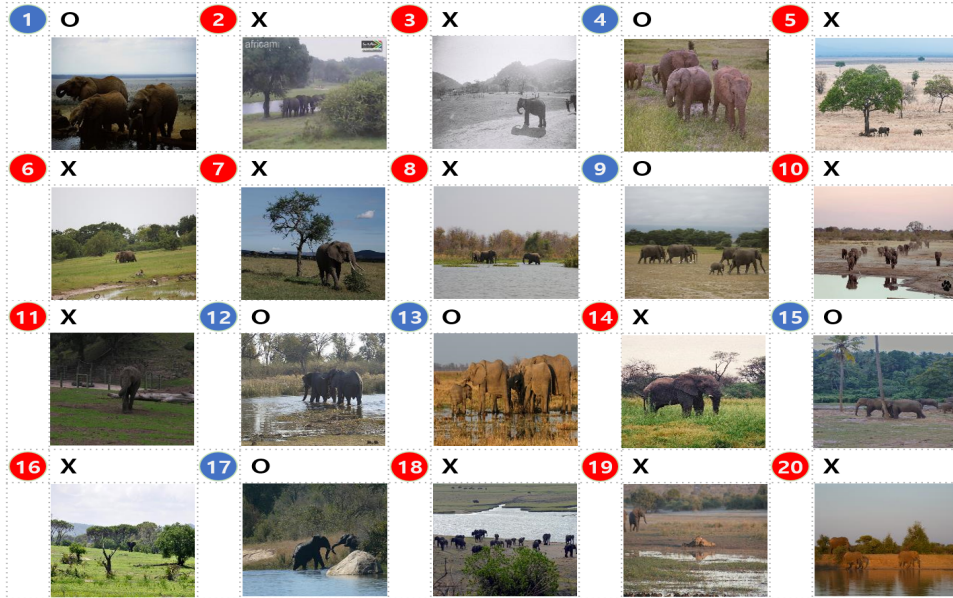
Division	Existing method	Proposed method
Top 5	58.2 %	66.4 %
Top 10	55.0 %	63.6 %
Top 20	53.4 %	57.0 %



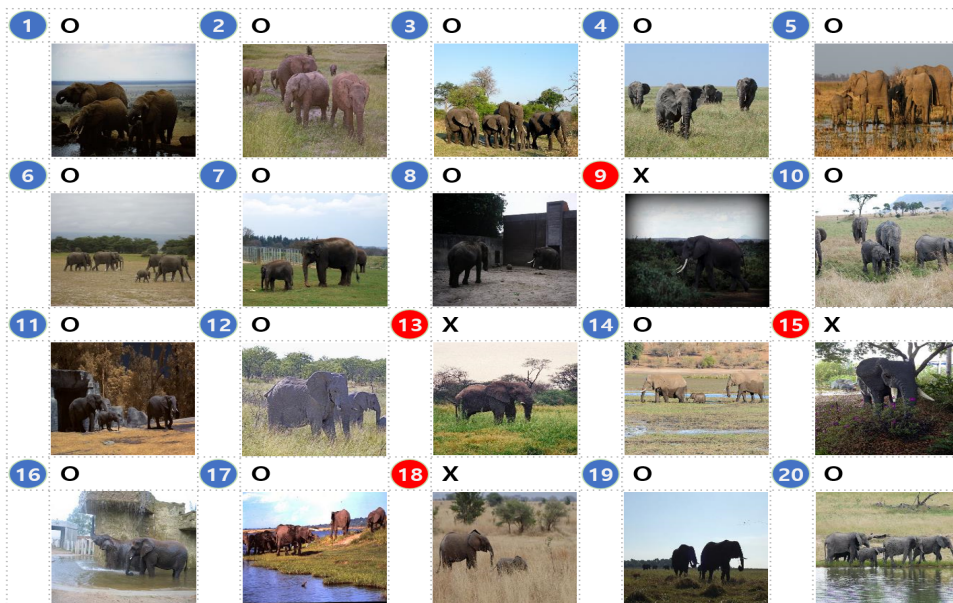
그림 4. 질의 이미지 예시  
Fig. 4. Example of query image

그림 5에서는 그림 4의 질의에 대한 기존 방법과 제안하는 방법의 검색 결과를 보여준다. 올바른 검색 결과 이미지의 경우 숫자의 바탕을 파란색('O')으로 잘못된 검색 결과 이미지의 경우 숫자의 바탕을 빨간색('X')으로 표현하였다. 그리고 각 번호는 유사 이미지 검색 시 부여된 순위와 동일하다. 즉, 1~20번의 번호는 유사도 순위를 의미한다.

질의 결과에서, 기존 방법의 결과 이미지는 전부 코끼리를 잘 찾아냈다. 하지만 코끼리 무리가 아닌 한 마리 혹은 이미지에서 코끼리가 찾아내는 비중이 적은 이미지들을 반 이상 검색한 것을 알 수 있다. 제안하는 방법에서는 이미지에서 차지하는 코끼리의 비중을 잘 고려하였으나, 한 마리의 코끼리만 존재하는 이미지를 결과로 가진 것이 일부 보인다.



(a) 기존 방법의 질의 결과  
(a) Query results of existing method



(b) 제안하는 방법의 질의 결과  
(b) Query results of the proposed method

그림 5. 그림 4의 질의 이미지의 검색 결과  
Fig. 5. Results for Fig. 4 query image

그림 5의 질의 결과는 제안하는 방법과 기존의 방법이 질의 이미지에 포함된 객체 클래스를 고려하여 유사 이미지를 검색하는 것을 확인할 수 있다. 그러나 기존의 방법은 객체의 위치와 관계 특징을 고려하지 못해 유사 이미지 검색의 성능이 제안하는 방법보다 낮은 것을 볼 수 있었다.

## V. 결 론

의학 이미지 분석, 원격 감지, 범죄 탐지 등의 응용 분야에 대한 필요성이 점차 증가하면서 해당 응용 분야의 핵심 기술인 내용 기반 유사 이미지 검색 기술(CBIR, Contents Based Image Retrieval)에 대한 중요성이 더욱더 주목받고 있다. CBIR의 핵심 기술은 각 응용 분야에서 처리해야 하는 이미지들의 특징을 잘 반영하는 특징 벡터 추출 기술이다.

기존에 수작업 기반 특징 벡터 추출 방법과 DCNN 기반의 특징 벡터 추출 방법이 다수 제안되었지만, 이 방법들은 이미지에 포함된 다양한 객체들 간의 관계를 잘 반영하는 특징 벡터 추출이 어렵다는 단점을 가지고 있다. 이 논문에서는 이 문제를 해결하기 위해서 이미지 전체 특징(Global feature), 이미지의 객체 목록 특징(Object list feature), 객체 관계 특징(Object relation feature)을 각각 추출하고 이 특징들을 결합하는 새로운 특징 추출 방법을 제안하였다.

특히, 객체 관계 특징 추출을 위해 객체 이미지 생성 방법을 추가로 제안하였다. 또한, 제안하는 특징 벡터 추출 방법에 적합한 유사도 계산 방법을 같이 제안하였다. 이 논문에서는 제안하는 방법을 구현하고 10만 개의 MS COCO 데이터 집합을 이용해서 기존방법과 비교하는 실험을 수행하였다. 실험 결과 약 8.2% (Top-5), 8.6% (Top-10), 3.6% (Top-20)의 성능향상이 있음을 확인하였다.

## Acknowledgement

본 논문은 한국교통대학교 석사학위 논문 'Contents based Image Retrieval Method using Global Image Features and Relationship among Objects'을 기

반으로 추가적인 실험을 수행하여 작성한 연구결과입니다.

## References

- [1] T. Deselaers, D. Keysers, and H. Ney, "Features for image retrieval: an experimental comparison", *Information retrieval*, Vol. 11, No. 2, pp. 77-107, Apr. 2008. <http://dx.doi.org/10.1007/s10791-007-9039-3>.
- [2] G. Tang, Z. Liu, and J. Xiong, "Distinctive image features from illumination and scale invariant keypoints", *Multimedia Tools and Applications*, Vol. 78, No. 16, pp. 23415-23442, May 2019. <https://doi.org/10.1007/s11042-019-7566-8>.
- [3] A. Babenko, A. Slesarev, A. Chigorin, and V. Lempitsky, "Neural codes for image retrieval", *Proc. of European conference on computer vision*, pp. 584-599, Sep. 2014.
- [4] S. Sun, W. Zhou, H. Li, and Q. Tian, "Search by detection: Object-level feature for image retrieval", *Proc. of international conference on internet multimedia computing and service*, pp. 46-49, Jul. 2014. <https://doi.org/10.1145/2632856.2632923>.
- [5] A. Babenko and V. Lempitsky, "Aggregating local deep features for image retrieval", *Proc. of the IEEE international conference on computer vision*, Santiago, Chile, pp. 1269-1277, Dec. 2015. <https://doi.org/10.1109/ICCV.2015.150>.
- [6] G. Toliás, R. Sicre, and H. Jégou, "Particular object retrieval with integral max-pooling of CNN activations", *ICLR*, pp. 1-12, May 2016. <https://doi.org/10.48550/arXiv.1511.05879>.
- [7] L. Gi-Won, M. Ju-Hyeong, and S. Seokil, "Content based Image Retrieval Method that Combining CNN based Image Features and Object Recognition Information", *Korean Institute of Information Technology/The, Journal of Korean Institute of Information Technology*, Vol. 20, No. 5, pp. 31-37, Jul. 2022. <http://dx.doi.org/10.14801/jkiit.2022.20.5.31>.



[8] W. Chen, Y. Liu, W. Wang, E. M. Bakker, T. Georgiou, and P. Fieguth, "Deep Learning for Instance Retrieval: A Survey", IEEE Transactions on Pattern Analysis and Machine Intelligence, pp. 1-20, Nov. 2022. <https://doi.org/10.1109/TPAMI.2022.3218591>.

[9] I. Kim, C. Jeong, and K. Jung, "AR Tourism Service Framework Using YOLOv3 Object Detection", The Journal of The Institute of Internet, Broadcasting and Communication(IIBC), Vol. 21, No. 1, pp. 195-200, Feb. 2021. <https://doi.org/10.7236/JIIBC.2021.21.1.195>.

[10] A. Bochkovskiy, C. Y. Wang, and H.-Y., M. Liao, "Yolov4: Optimal speed and accuracy of object detection", CoRR, Apr. 2020. <https://doi.org/10.48550/arXiv.2004.10934>.

[11] sklearn.metrics.pairwise.rbf\_kernel, [https://scikit-learn.org/stable/modules/generated/sklearn.metrics.pairwise.rbf\\_kernel.html](https://scikit-learn.org/stable/modules/generated/sklearn.metrics.pairwise.rbf_kernel.html) [accessed: Jul. 21, 2022]

[12] Calculate Jaccard Similarity in Python, <https://data-scienceparichay.com/article/jaccard-similarity-python> [accessed: Jul. 21, 2022]

[13] A. S. Razavian, J. Sullivan, S. Carlsson, and A. Maki, "Visual instance retrieval with deep convolutional networks", ITE Transactions on Media Technology and Applications, pp. 251-258, Jul. 2016. <https://doi.org/10.48550/arXiv.1412.6574>.

[14] Y. Cao, M. Long, J. Wang, and S. Liu, "Deep visual-semantic quantization for efficient image retrieval", Proc. of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, pp. 1328-1337, Jul. 2017. <https://doi.org/10.1109/CVPR.2017.104>.

[15] L. Tsung-Yi, "COCO" <https://cocodataset.org> [accessed: May 10, 2022]

[16] Juhyung Meang, "Contents based Image Retrieval Method using Global Image Features and Relationship among Objects", Master Thesis, Korea National University of Transportation, 2022.

저자소개

맹 주 형 (Juhyeong Maeng)



2020년 2월 : 한국교통대학교  
컴퓨터공학과(공학사)  
2020년 9월 ~ 2022년 8월 :  
한국교통대학교 컴퓨터공학과  
(공학석사)  
관심분야 : 인공지능, 데이터베이스,  
컴퓨터 비전, 이동객체

김 성 종 (Seongjong Kim)



2022년 2월 : 한국교통대학교  
컴퓨터공학과(공학사)  
2022년 3월 ~ 현재 :  
한국교통대학교 컴퓨터공학과  
(석사과정)  
관심분야 : 인공지능,  
데이터베이스, 컴퓨터 비전,

이동객체

송 석 일 (Seokil Song)



1998년 2월 : 충북대학교  
정보통신공학과(공학사)  
2000년 2월 : 충북대학교  
정보통신공학과(공학석사)  
2003년 2월 : 충북대학교  
정보통신공학과(공학박사)  
2003년 7월 ~ 현재 :

한국교통대학교 컴퓨터공학과 교수

관심분야 : 데이터베이스, 센서 네트워크, 스토리지  
시스템