

# 콘텐츠 기반의 이미지검색을 위한 분류기 접근방법

한 우 진\*, 손 경 아°

## Image Classification Approach for Improving CBIR System Performance

Woo-Jin Han\*, Kyung-Ah Sohn°

### 요 약

콘텐츠 기반 이미지 검색은 기존의 태그 또는 레이블이 있는 텍스트 기반의 검색이 아닌 이미지의 특징을 이용하여 검색하는 방법이다. 실생활 이미지 데이터는 태그나 레이블이 달려있는 경우가 많지 않기 때문에 텍스트 기반의 검색 방법을 사용하기 힘든 경우가 있다. 또한, 기존에 주로 사용되는 이미지 특징 벡터의 유사도를 사용하여 검색하는 방법은 추출 벡터의 유사도 기준으로 사용자가 의도한 결과가 나올지 확신할 수 없다. 예를 들어 사용자가 입력한 질의 이미지와 검색된 이미지들의 종류가 일치하는지의 문제가 있다. 본 논문에서는 사용자가 질의 이미지의 클래스를 예상하고 결과도 동일한 클래스를 원한다는 가정에 착안하여 이미지 검색 엔진의 성능을 개선하였다. 기존의 유사도 기반의 검색에 머신 러닝 기법을 사용한 이미지 분류기를 적용하여 질의와 동일한 클래스의 결과를 찾는 방법을 제안하였으며, 그 성능을 20개 카테고리에 속하는 11,530개의 이미지로 구성되어 있는 PASCAL VOC 공개 데이터를 이용하여 검증하였다.

**Key Words** : Content Based Image Retrieval, image classification, PASCAL VOC, SIFT, SVM, MLP

### ABSTRACT

Content-Based image retrieval is a method to search by image features such as local color, texture, and other image content information, which is different from conventional tag or labeled text-based searching. In real life data, the number of images having tags or labels is relatively small, so it is hard to search the relevant images with text-based approach. Existing image search method only based on image feature similarity has limited performance and does not ensure that the results are what the user expected. In this study, we propose and validate a machine learning based approach to improve the performance of the image search engine. We note that when users search relevant images with a query image, they would expect the retrieved images belong to the same category as that of the query. Image classification method is combined with the traditional image feature similarity method. The proposed method is extensively validated on a public PASCAL VOC dataset consisting of 11,530 images from 20 categories.

\* 이 논문은 2015년도 정부(미래창조과학부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(NRF-2014R1A1A3051169 & No. 2010-0028631)

• First Author : Ajou University Department of Software, data@ajou.ac.kr, 학생회원

° Corresponding Author : Ajou University Department of Software and Computer Engineering, kasohn@ajou.ac.kr, 정회원

논문번호 : KICS2016-05-091, Received May 11, 2016; Revised June 22, 2016; Accepted June 27, 2016

## I. 서론

스마트폰 기술의 발달과 대중화로 누구나 하나쯤은 손안의 작은 카메라를 휴대하고 있다. 스마트폰의 카메라 기능은 스마트폰 제조업체간 경쟁의 주요 이슈가 되었고, 그 어느 때보다 소비자는 편리하고 고성능 카메라를 휴대할 수 있었다. 자연스럽게 사람들은 사진을 많이 찍게 되었고 이미지를 손쉽게 저장하려는 요구사항에 힘입어 클라우드 환경의 스토리지 환경을 제공하는 업체가 생기고, 다른 사용자와 사진을 공유하는 소셜 플랫폼이 생김과 문자중심의 대화방식에서 이미지중심의 소통방법으로 점점 이동하고 있는 추세이다. 이러한 환경 속에서 기업들은 이미지를 이용한 서비스를 개발하고 출시하는데 콘텐츠 기반의 이미지 검색 서비스도 대표적인 서비스로 볼 수 있다.

콘텐츠 기반의 이미지검색 시스템(CBIR)<sup>[1]</sup>은 기존의 텍스트 기반의 검색방법이 아닌 이미지의 특징을 추출하여 유사한 이미지를 검색하는 방법이다. 하루에도 수많은 사용자가 업로드 하는 사진들 중 태그나 레이블이 달려 업로드 되는 수는 많지 않다. 수많은 사진들 속에서 다른 유저나 자신이 올린 사진의 유사도를 기반으로 사진 또는 사람을 추천해주는 서비스에 해당 할 수 있다. 이러한 과정은 이미지에서 특징들을 찾아 이미지벡터를 추출하여 유사도<sup>[2,3]</sup>를 계산 후 정렬하여 보여주는 방법을 사용하게 된다. 이러한 방법의 단점은 이미지 벡터의 유사도로 검색된 이미지가 사용자의 검색 의도와 일치하는 결과를 보이는지 명확하지 않다. 이러한 점을 보완하기 위해 본 연구에서는 사용자의 질의 이미지와 유사도를 기반으로 한 결과이미지에 이미지 분류기를 적용하여 질의 이미지와 결과 이미지가 동일한 분류항목일 때 검색 결과로 보여주는 접근방식을 제안한다.

## II. 시스템흐름 및 설계

### 2.1 유사도 방법과 분류기 적용 제한

기존의 이미지벡터 유사도를 바탕으로 한 방법은 3 가지 단계로 나눌 수 있다. 첫째로 이미지에서 특징벡터를 추출하는 과정, 둘째, 질의 이미지벡터와 데이터베이스에 저장된 이미지벡터사이의 유사도를 구하는 과정, 마지막으로 유사도를 바탕으로 정렬하여 질의 이미지와 유사한 이미지를 나타내는 과정으로 나눌 수 있다.

이미지검색의 성능향상을 위해 적합성 피드백(relevance feedback)을 적용하는 방법도 기존에 연구

되고 있다. 적합성 피드백은 사용자가 검색결과를 확인하여 결과들의 적합성 여부를 판단 후, 시스템이 질의에 반영하여 다시 검색하거나, 검색결과 상위에 있는 특징들을 추출하여 다시 질의에 추가하는 과정을 통해 성능을 높이는 방법으로 접근한다. 적합성 피드백과 분류기를 적용한 방법은 검색결과에 추가정보를 더하여 성능을 높이는 관점에서 관련성이 있다. 하지만, 적합성 피드백은 사용자가 시스템을 사용하면서 시스템의 성능이 발전되는 것이고, 분류기 적용방법은 사용자가 시스템 사용전 적합성 데이터를 사전에 학습시키는 것이 다른 특징이다. 또한, 적합성 피드백은 이미지당 여러 종류의 특징벡터를 필요로 하기 때문에, 과정이 복잡하거나 이미지를 대표하는 특징을 찾지 못하면 성능향상에 어려움이 있게 된다. 때문에 이번 연구에서는 분류기를 적용한 방법을 사용하여 연구를 진행하였다.

그림 1은 유사도 기반의 검색방법과 분류기를 결합하여 본 논문에서 제안하는 CBIR 흐름도를 나타낸다. 기존의 방법에 몇 단계가 추가 되는데, 먼저 분류기를 학습시키기 위해 레이블이 함께 있는 훈련데이터가 필요하다. 훈련데이터를 사용하고 분류알고리즘을 이용하여 분류기를 학습시킨다. 그 후, 질의 이미지가 입력되면 이미지 벡터를 추출하여 학습된 분류기를 사용하여 분류 확률이 상위 k개인 카테고리를 저장한다. 질의 이미지와 데이터베이스에 저장된 이미지간의 유사도를 계산하여 정렬 후, 상위 이미지들을 학습된 분류기를 사용하여 카테고리를 분류한다. 분류기로 예측된 카테고리가 질의 이미지의 상위 k개의 카테고리에 포함 될 때만 최종 결과에 포함한다. 결과적으로

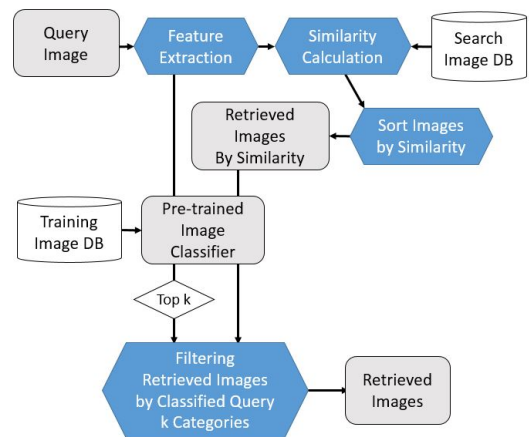


그림 1. 분류기를 적용한 CBIR 시스템 흐름도  
Fig. 1. CBIR System Flow with Classifier

질의 이미지와 같은 종류의 이미지만 검색되는 결과를 얻을 수 있다. 본 논문에서는 검색이미지의 수를 상위 50, 100, 150으로 설정하였고, 분류기의 상위 예측 카테고리는 3, 5개로 고정하여 각각의 성능을 측정하고 서로 다른 분류 알고리즘을 적용하여 성능을 비교, 분석한다.

### 2.2 데이터 및 이미지벡터추출

본 연구에서 사용한 데이터는 PASCAL VOC<sup>[4]</sup> 공개 데이터이다. PASCAL VOC는 20개의 클래스로 구성된 총 11,530개의 이미지 데이터이다. 클래스는 비행기, 자전거, 보트 등의 카테고리로 구성되어 있으며 한 개의 카테고리 당 평균 576.5개의 이미지가 포함되어 있다. 하지만 그림 2에서 실제 데이터의 분포를 확인하면 몇 개의 카테고리에 이미지가 몰려있는 불균형 분포를 확인할 수 있다.

이미지의 유사도를 구하기 위해 이미지의 특징벡터를 추출해야 하는데, 이 과정은 SIFT (Scale-invariant feature transform)<sup>[5,6]</sup> 알고리즘을 사용 하였고, BoVW(Bag of Visual Word)<sup>[7]</sup>를 이용하여 이미지들이 모두 같은 길이의 이미지 벡터로 표현되도록 하였다. BoVW는 하나의 이미지가 여러 시각단어의 묶음으로 표현될 수 있다는 이론을 바탕으로 벡터의 사이즈를 시각단어의 수로 고정할 수 있는 장점으로 SIFT와 함께 이미지 특징 벡터 추출에 많이 사용되는 알고리즘이다.

### 2.3 이미지 분류기 학습

이미지 분류 알고리즘으로는 SVM (Support Vector Machine)<sup>[8]</sup>, RF (Random Forest)<sup>[9]</sup>, SSL (Semi-Supervised Learning)<sup>[10]</sup> 그리고, MLP (Multilayer Perceptron)<sup>[11]</sup>을 사용 하였다. SVM은 전통적으로 머신러닝 분류 문제에서 높은 성능을 보여

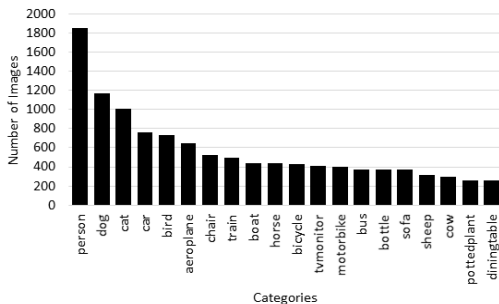


그림 2. 카테고리 당 이미지 개수 분포  
Fig. 2. The Number of Images per Category

왔던 강점이 있는 반면 분류하는 클래스의 개수가 많을 때 성능이 낮아지는 단점이 있다. SSL은 학습데이터가 적을 때 레이블이 없는 테스트 데이터도 학습에 포함 시킬 수 있는 장점이 있는 반면, 라플라시안 행렬 계산이 필요하기 때문에 데이터가 많을 경우 학습 시간이 오래 걸리거나 메모리를 많이 필요로 하는 단점이 있다. MLP는 신경망을 바탕으로 학습이 모델에 과 적합 하는 단점이 있지만 Dropout<sup>[12]</sup>의 방법으로 과적합을 피할 수 있는 방법들이 제안되고 있다. 마지막으로 랜덤 포레스트 (RF, Random Forest)는 앙상블 학습 방법의 하나로써 다수의 결정 트리를 이용하여 하나의 의견으로 수렴하는 방식을 취한다. 이러한 방법은 다수의 약한 분류기를 이용하여 하나의 강한 분류기를 만들어내는데 특정한 데이터에 편향 되지 않는 장점과 결정트리의 장점인 예측 클래스가 많은 경우 학습하기 좋은 강점이 있다.

## III. 실험방법과 결과

### 3.1 유사도를 사용한 이미지검색 성능

유사도를 사용하여 질의 이미지에 대한 검색성능은 코사인 거리 (Cosine Distance)와 유클리드 거리 (Euclidean Distance)를 사용하였다. 성능 평가는 질의 이미지 5개의 P@k (Precision at K)<sup>[13]</sup>와 MAP (Mean Average Precision)<sup>[13]</sup>을 측정하였다. P@k의 수식은 다음과 같다.

$$\text{Precision at K} = \text{rel}(k)/K \quad (1)$$

여기서 K는 검색된 결과의 상위 K값들을 의미하고 rel(k)는 검색된 결과 중 질의와 관련 있는 결과의 개수를 의미한다. AP(Average Precision)<sup>[13]</sup>과 MAP의 식은 다음과 같다.

$$ap@n = \sum_{k=1}^n P(k) / \min(m, n) \quad (2)$$

$$MAP@n = \sum_{i=1}^N ap@n_i / N \quad (3)$$

수식 (2)에서 P(k)는 상위 k의 precision을 의미하고 m은 질의와 관련 있는 이미지의 수, n은 검색된 이미지의 수이다. 수식 (3)은 MAP를 설명 하였는데, N은 질의의 수로 AP를 질의의 수로 평균한 것이다. P@k는 상위 50, 100, 150기준으로 계산하였고, MAP

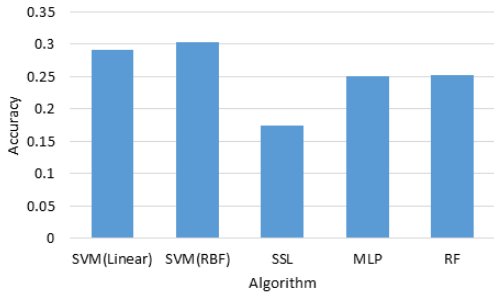


그림 3. 분류 알고리즘별 정확도 측정  
Fig. 3. Image Classifier Accuracy

도 50, 100, 150 기준으로 각 질의 이미지에 대한 Average Precision을 구한 후 평균값을 사용하였다. 이미지 특징 벡터를 추출할 때 사용된 BoVW는 벡터의 길이를 클러스터의 개수로 고정하여 만들 수 있어서 분류기에 학습시키기 좋은 이점이 있지만 비주얼 단어의 수인 클러스터 개수를 정하는데 있어서 별도의 정답이 없는 단점이 있다. 이번 실험에서는 클러스터의 개수를 100, 300, 500으로 설정하여 각각의 성능을 비교하였다. 유사도를 사용하여 검색한 상위 50, 100, 150개의 P@k를 비교해 보았을 때 BoVW 클러스터수가 100, 유클리드 거리를 사용한 결과가 비교적 높은 것으로 확인되었다. 또한, 100, 500개 클러스터의 유클리드를 이용한 결과가 성능이 높게 나왔다. 이를 바탕으로 분류기를 적용하는 단계에서는 유클리드 거리를 기준점(Baseline)으로 비교 분석 할 것이다.

### 3.2 이미지 분류기 성능 평가

이미지 분류기를 학습시켜 주어진 데이터의 예측 정확도를 평가 하였다. 데이터는 PASCAL VOC 데이터인 11,530개 데이터를 80%는 훈련데이터로 사용하고 20%는 테스트 데이터로 사용하여 성능평가를 진행하였다. SVM, SSL, MLP, RF 분류기를 따로 훈련하여 성능을 비교하였다. SVM은 다중클래스를 예측할 때 사용하는 방법인 One vs Rest 방법을 사용하여 분류기를 학습하였고, 커널은 선형커널, 방사 기초 함수(RBF, Radial basis function)를 이용하여 비교하였다. 그림 3에서 SVM의 RBF 커널을 사용한 모델의 분류 정확도가 다른 분류기보다 높은 것을 확인할 수 있었다. 전체적으로 BoVW의 클러스터 수는 500에서 좋은 성능을 보였다. SSL은 분류기 학습시에 레이블이 없는 데이터도 같이 학습하였지만 낮은 성능을 보인 것은 학습에 추가로 참여한 데이터의 수가 다소 적었던 이유로 예상된다. MLP은 은닉노드의 수를 바꿔

가며 2개의 은닉층으로 Input-layer, 2000 노드, 1000 노드, Output-layer, 50% Dropout으로 설계하였다.

### 3.3 분류기를 적용한 이미지검색 성능

기존의 유사도 기반검색에 분류기를 적용하는 방법은 미리 학습한 분류기를 유사도기반의 검색에 적용하여 질의이미지의 카테고리를 예측하는 것이다. 그리고, 유사도 상위이미지의 카테고리를 분류하여 질의 이미지와 같은 결과만 남기는 작업을 하게 된다. 이러한 과정을 통해 SVM (RBF, Linear), RF, MLP를 각각 적용한 검색성능을 비교분석 하였다. SSL은 이번 실험에서 분류 정확도가 다른 분류기에 비해 상대적으로 낮았기 때문에 적용하지 않았다.

#### 3.3.1 SVM을 적용한 이미지 검색 성능

SVM을 분류기로 적용하여 비교한 결과는 그림 4에서 확인 할 수 있다. 그림 4는 BoVW의 클러스터를 100, 300, 500으로 설정하여 추출한 벡터를 사용하고, k값의 50, 100, 150에서 가장 높은 성능을 보인 알고리즘의 P@k를 비교한 것이다. 또한 표 1은 각 클러스터의 개수와 알고리즘의 MAP의 성능을 비교하였다. 결과적으로 P@k는 SVM RBF커널과 분류기 예측 상

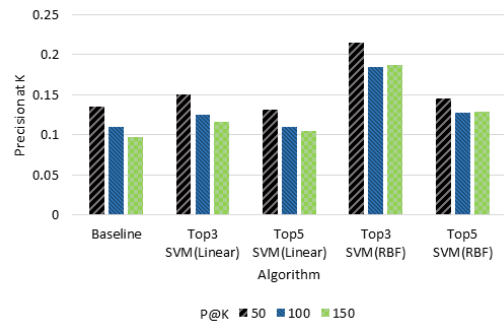


그림 4. SVM을 적용한 P@k 비교  
Fig. 4. P@k with SVM

표 1. SVM을 적용한 Mean Average Precision  
Table 1. Mean Average Precision with SVM

Algorithm	BoVW C100	BoVW C300	BoVW C500
Baseline	0.0911	0.0800	0.0678
Top3 SVM(Linear)	0.0873	0.0744	0.1044
Top5 SVM(Linear)	0.1158	0.0891	0.0938
Top3 SVM(RBF)	0.1444	0.1538	0.1507
Top5 SVM(RBF)	0.1069	0.0984	0.1031

위 k값이 Top3일 때 가장 높은 성능을 보였다. BoVW 클러스터를 100, 300, 500으로 설정한 표 1에서도 Top3의 SVM RBF커널이 가장 높은 성능을 보인 것을 확인하였다.

3.3.2 랜덤 포레스트를 적용한 이미지 검색 성능

랜덤 포레스트를 적용한 모델의 성능은 그림 5, 6에서 확인할 수 있다. 그림 5에서 앞의 3개 모델은 유사도를 사용하여 검색한 모델의 P@k이다. 다음 2개는 클러스터가 100개일 때 분류기 예측 상위 k값 3, 5에서의 성능 결과이다. 각 모델 중 가장 성능이 높은 모델의 결과만을 표시하였다. 그림 5에서 가장 높은 성능을 보인 모델은 BoVW 300개 클러스터와 분류기 상위 3개 예측과 결과 이미지 상위 50개를 측정 한 것이 가장 높은 성능을 보였다. 그림 6에서는 모든 질의 이미지에 대한 MAP 결과를 보여주는데, BoVW 100개 클러스터와 분류기 Top3에서 높은 성능을 보인 것을 확인할 수 있다.

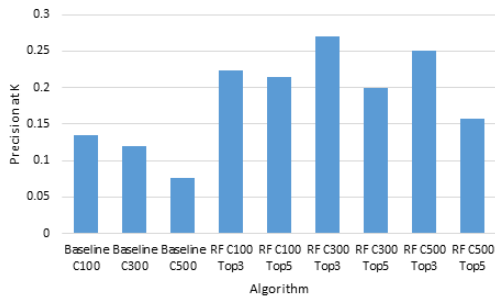


그림 5. 랜덤포레스트를 적용한 P@k 결과  
Fig. 5. P@k with Random Forest

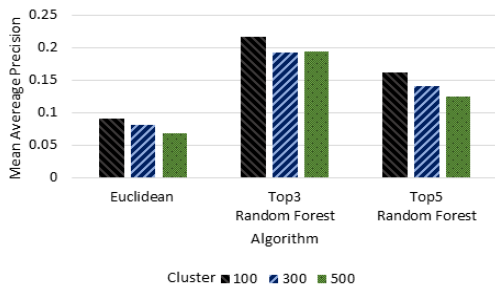


그림 6. 랜덤포레스트를 적용한 MAP결과  
Fig. 6. MAP with Random Forest

3.3.3 다층 퍼셉트론을 적용한 이미지검색 성능

다층 퍼셉트론을 적용한 모델의 성능은 그림 7, 8에서 확인할 수 있다. 그림 7또한 각각의 다른

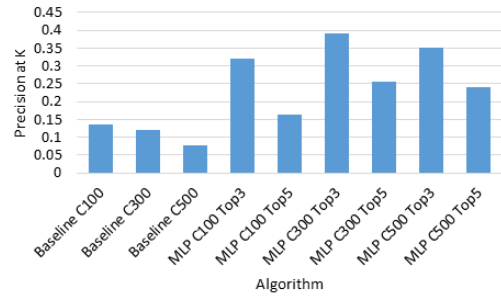


그림 7. 다층 퍼셉트론을 적용한 P@k결과  
Fig. 7. P@k with Multilayer Perceptron

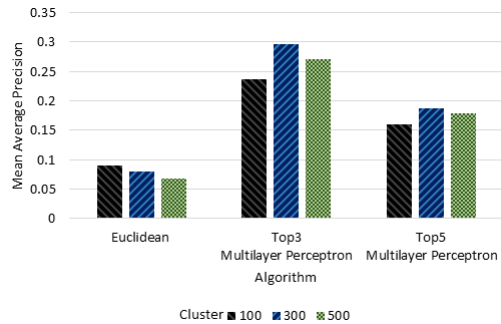


그림 8. 다층 퍼셉트론을 적용한 MAP결과  
Fig. 8. MAP with Multilayer Perceptron

BoVW 클러스터와 분류기 예측 상위 k 값을 변화시키며 적용한 모델들의 성능을 비교 하였고, 분류기에 예측 상위 Top3와 BoVW 클러스터 300개를 사용한 모델의 성능이 높은 결과를 보였다.

IV. 결론 및 향후 발전방향

기존의 유사도 접근의 방법에 분류기를 사용한 방법이 전체적인 성능향상이 있는 것을 확인 할 수 있었다. 그림 9는 지금까지 분류기를 적용한 모델들의 MAP 성능 값을 비교한 것이다. 가장 성능이 높은 모델은 분류기 상위 Top3의 MLP를 적용한 모델이었다. 상위 3개의 모델은 모두 MLP를 적용한 것이고, 뒤를 따르는 3개는 랜덤포레스트 모델을 적용한 모델이었다. 분류기의 성능만 비교 했을 때 SVM 성능이 가장 우수하였지만 상위 Top k 방법을 적용 하였을 때 상위권 확률에 포함되는 카테고리가 랜덤포레스트와 MLP가 우수한 것으로 분석된다. 향후 연구로서 유사도방법의 기본 성능을 매칭 방법<sup>[14]</sup>을 활용하여 향상시키고, 분류기 정확도를 높이는 연구를 진행하여 이미지 검색 엔진의 성능을 높이는 접근으로 연구가 진

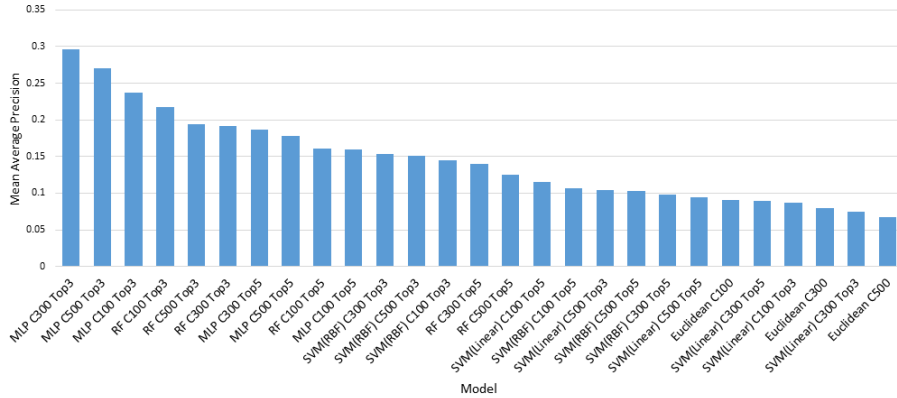


그림 9. 분류기 알고리즘을 적용한 MAP 성능 비교  
Fig. 9. MAP Comparison among the Models

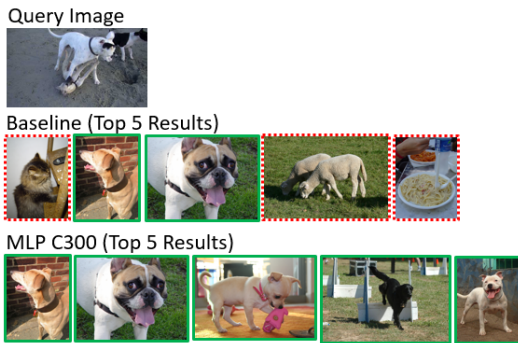


그림 10. 분류기적용 전후 상위 5개 이미지검색 결과  
Fig. 10. Top 5 results with Image Classifier

행 될 것이다.

### References

[1] P. S. Suhasini, K. Krishna, and I. M. Krishna, "CBIR using color histogram processing," *J. Theoretical & Appl. Inf. Technol.*, vol. 6, no. 1, 2009.

[2] W. J. Han and K. A. Sohn, "Image classification approach for Improving CBIR system performance," in *Proc. 2016 KICS Conf. Winter*, pp. 308-309, Jeongseon, Korea, Jan. 2016.

[3] J. S. Song, S. J. Hur, Y. W. Park, and J. H. Choi, "User positioning method based on image similarity comparison using single camera," *J. KICS*, vol. 40, no. 8, pp. 1655-1666, Aug. 2015.

[4] M. Everingham, et al., "The pascal visual object classes (voc) challenge," *Int. J. Computer Vision*, vol. 88, no. 2, pp. 303-338, 2010.

[5] Y. Ke and R. Sukthankar, "PCA-SIFT: A more distinctive representation for local image descriptors," in *Proc. IEEE Computer Soc. Conf. CVPR 2004*, vol. 2, 2004.

[6] H. J. Jung and J. S. Yoo, "Feature matching algorithm robust to viewpoint change," *J. KICS*, vol. 40, no. 12, pp. 2363-2371, Dec. 2015.

[7] J. Yang, Y. G. Jiang, A. G. Hauptmann, C. W. Ngo, "Evaluating bag-of-visual-words representations in scene classification," in *Proc. MIR '07*, pp. 197-206, Sept. 2007.

[8] S. Jain, "A machine learning approach: SVM for image classification in CBIR," *IJAIEM*, vol. 2, no. 4, Apr. 2013.

[9] A. Liaw and M. Wiener, "Classification and regression by randomForest" *R news*, vol. 2, no. 3, pp. 18-22, Dec. 2002.

[10] X. Zhu, *In Encyclopedia of Machine Learning*, Springer US, "Semi-supervised learning," pp. 892-897, 2011.

[11] B. B. Chaudhuri and Ujjwal Bhattacharya, "Efficient training and improved performance of multilayer perceptron in pattern classification," *Neurocomputing*, vol. 34, no.

1, pp. 11-27, 2000.

- [12] M. Lin, Q. Chen, and S. Yan. "Network in network," arXiv preprint arXiv: 1312.4400, Dec. 2013.
- [13] Tyler Byers, "Mean average precision" Retrieved Apr. 29, 2016. <https://www.kaggle.com/wiki/MeanAveragePrecision>.
- [14] G. R. Choi, H. W. Jung, J. H. Lee, "Contents-based image retrieval system design of shopping mall using SIFT matching," *KIIS Spring Conf.*, vol. 22, no. 1, pp. 161-163, Mokpo, Korea, Apr. 2012.

손 경 아 (Kyung-Ah Sohn)



2000년 2월 : 서울대학교 학사  
2003년 2월 : 서울대학교 석사  
2011년 12월 : Carnegie Mellon University 박사  
2013년 3월~현재 : 아주대학교 컴퓨터공학과 조교수

<관심분야> 머신러닝, 데이터마이닝, 의료정보학

한 우 진 (Woo-Jin Han)



2014년 2월 : 한양대학교  
ERICA 산업경영공학과 졸업  
2015년 3월~현재 : 아주대학교  
소프트웨어특성화 석사과정  
<관심분야> 데이터마이닝, 머신러닝, 컴퓨터비전