

다양한 딥 러닝 학습 환경하의 콘텐츠 기반 이미지 검색

장영균 / 서울대학교 정보신호처리연구실

방대한 데이터베이스에서 질의에 대한 관련 이미지를 찾는 콘텐츠 기반 이미지 검색은 컴퓨터 비전 분야의 근본적인 작업 중 하나이다[1]. 특히 빠르고 정확한 검색을 수행하기 위해 해싱(Hashing) 및 곱 양자화(Product Quantization, PQ)로 대표되는 근사최근접 이웃(Approximate Nearest Neighbor, ANN) 검색 방식이 이미지 검색 커뮤니티에서 주목받고 있다[2,3]. 신경망 기반 딥 러닝(CNN-based deep learning)이 많은 컴퓨터 비전 작업에서 우수한 성능을 보여준 이후로, 해싱 및 곱 양자화 기반 이미지 검색 시스템 모두 개선을 위해 딥 러닝을 채택하고 있다[4,5]. 본 학위 논문에서는 적절한 검색 시스템을 제안하기 위해 다양한 딥 러닝 학습 환경 아래에서 이미지 검색 방법을 제안하였다. 구체적으로, 이미지 검색의 목적을 고려하여 의미적으로 유사한 이미지를 검색하는 딥 러닝 해싱 시스템을 개발하기 위한 지도 학습 방법을 제안하고, 의미적, 시각적으로 모두 유사한 이미지를 검색하는 딥 러닝 곱 양자화 기반의 시스템을 구축하기 위한 준지도, 비지도 학습 방법을 제안하였다[6,7].

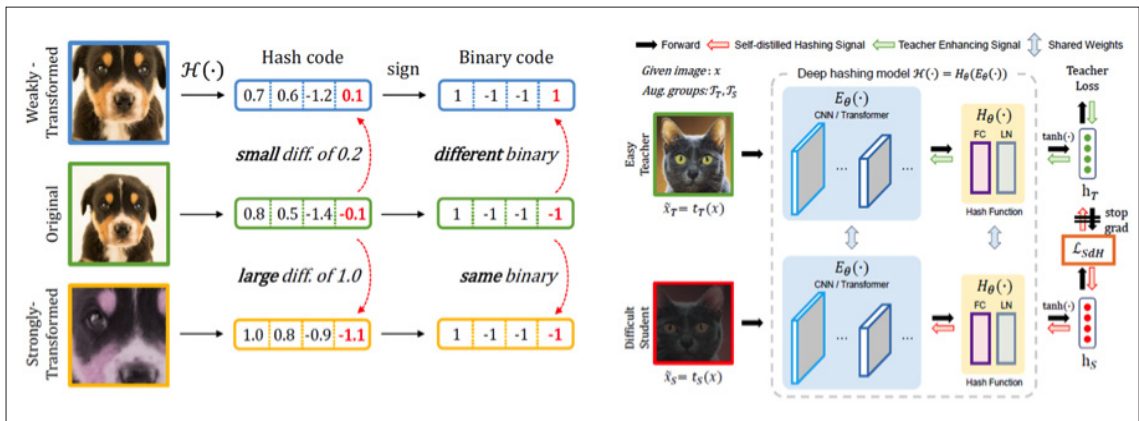
먼저 이미지에 부여된 의미론적 레이블을 사용하는 지

도 학습을 도입한 해싱 기반 검색 시스템에 대해 소개한다. 클래스 간 유사성(다른 사람 사이의 유사한 외모)과 클래스 내 변화(같은 사람의 다른 포즈, 표정, 조명)와 같은 얼굴 이미지 구별의 어려움을 해결하기 위해 각 이미지의 클래스 레이블을 사용한다. 얼굴 이미지 검색 품질을 더욱 향상시키기 위해 SGH(Similarity Guided Hashing) 방식을 제안하며(〈그림 1〉), 여기서 다중 데이터 증강 결과를 사용한 자기 유사성 학습이 훈련 중에 사용된다. 그리고 해싱 기반의 일반 이미지 검색 시스템을 구성하기 위해 DHD(Deep Hash Distillation) 방식을 제안하였다(〈그림 2〉).

둘째로, 레이블이 지정된 이미지 데이터와 레이블이 지정되지 않은 이미지 데이터를 모두 활용하는 준지도 학습을 조사하여 곱 양자화 기반 검색 시스템을 구축하였다. 지도 학습 딥 러닝 기반의 이미지 검색 방법들은 우수한 성능을 보이려면 값비싼 레이블 정보가 충분해야 한다는 단점이 있다. 이 문제를 해결하기 위해 벡터 양자화 기반 반지도 영상 검색 방식인 GPQ(Generalized Product Quantization) 네트워크를 제안하였다(〈그림 3〉).



<그림 1> Similarity Guided Hashing 방법

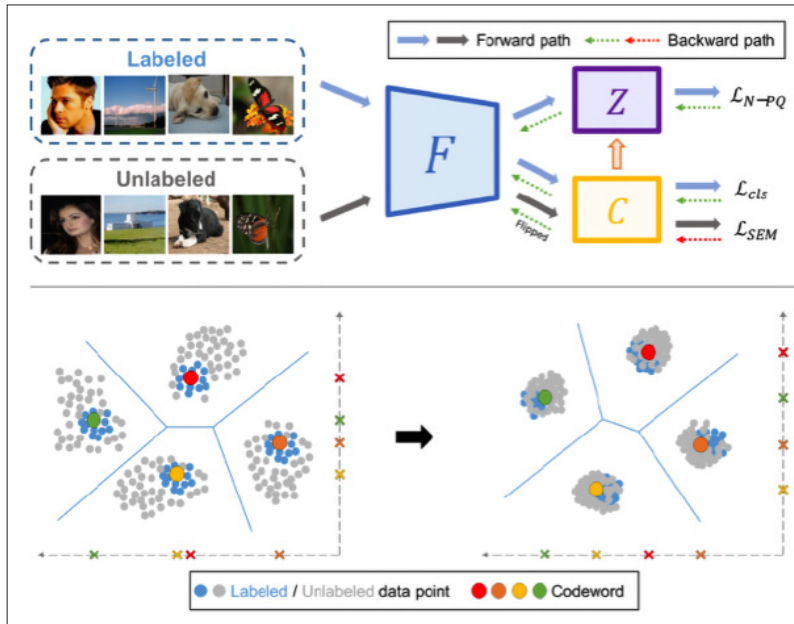


<그림 2> Deep Hash Distillation 방법

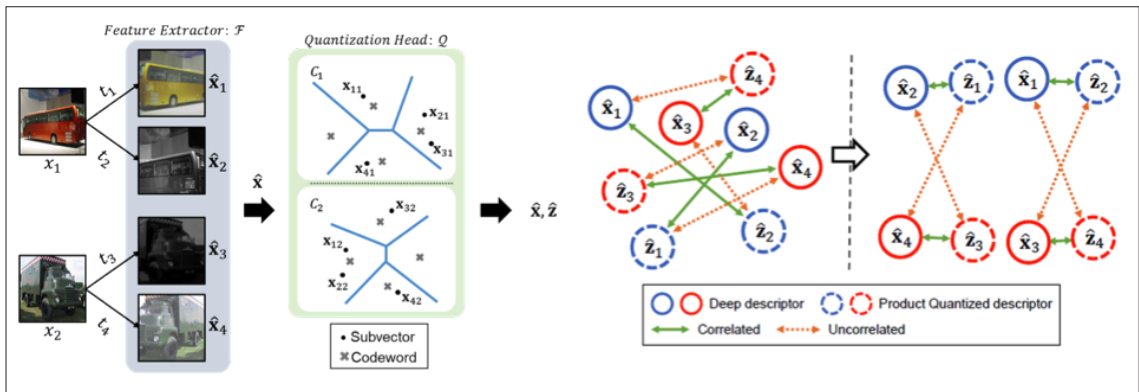
마지막으로, 딥 러닝 모델이 사람의 지도 없이 시각적으로 유사한 이미지 검색을 수행할 수 있도록 하기 위해 비지도 학습 알고리즘을 탐색한다. 비록 레이블 주석을 활용한 심층 지도 기반의 방법들이 기존 방법들 대비 우수한 검색 성능을 보일지라도, 방대한 양의 훈련 데이터에 대해

정확하게 레이블을 지정하는 것은 힘들고 주석에서 오류가 발생하기 쉽다는 한계가 있다. 이 문제를 해결하기 위해 레이블 없이 자체 지도 방식으로 훈련하는 SPQ(Self-supervised Product Quantization) 네트워크라는 심층 비지도 이미지 검색 방법을 제안하였다(그림 4).

졸업논문
소개



<그림 3> Generalized Product Quantization 방법



<그림 4> Self-supervised Product Quantization 방법

참 고 문 헌

-
- [1] Yong Rui, Thomas S. Huang, and Shih-Fu Chang. "Image retrieval: Past, present, and future." *Journal of Visual Communication and Image Representation* 10,1, 1999.
 - [2] Dong-ju Jeong, Sungkwon Choo, Wonkyo Seo, Nam Ik Cho, "Regional deep feature aggregation for image retrieval," *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2017.
 - [3] Albert Gordo, et al. "Deep image retrieval: Learning global representations for image search." *European Conference on Computer Vision (ECCV)*, 2022, 2016.
 - [4] Young Kyun Jang, Geonmo Gu, Byungsoo Ko, Isaac Kang, Nam Ik Cho, "Deep hash distillation for image retrieval," *European Conference on Computer Vision (ECCV)*, 2022.
 - [5] Young Kyun Jang, Nam Ik Cho, "Generalized product quantization network for semi-supervised image retrieval," *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2020.
 - [6] Young Kyun Jang, Nam Ik Cho, "Self-supervised product quantization for deep unsupervised image retrieval," *Proceedings of the IEEE/CVF International Conference on Computer Vision (CVPR)*, 2021.
 - [7] Haeyoon Yang, Young Kyun Jang, Isaac Kang, Nam Ik Cho, "Self-Supervised Pretraining for Deep Hash-Based Image Retrieval," *IEEE International Conference on Image Processing (ICIP)*, 2022.
-



장 영 균

- 2016년 : 서울대학교 전기정보공학부 학사
- 2022년 : 서울대학교 전기정보공학부 박사
- 2022년 ~ 현재 : *Postdoctoral Research Scientist at Meta*
- 주관심분야 : 멀티모달 표현학습, 정보 검색