

객체 탐지 알고리즘 기반 이미지 검색 시스템

안지현^o, 박승민^{*}

^o동서대학교 소프트웨어학과,

^{*}동서대학교 소프트웨어학과

e-mail: sminpark@dongseo.ac.kr^{*}

Image Search System Based on Object Detection Algorithm

Ji-Hyun Ahn^o, Seungmin Park^{*}

^oDept. of Software, Dongseo University,

^{*}Dept. of Software, Dongseo University

● 요약 ●

최근에 이르러 인공지능망의 발전은 CNN(Convolutional Neural Network) 알고리즘을 활용한 이미지 분석 및 검색 시스템에 비약적인 기여를 하고 있다. 이는 이미지를 입력으로 받아 유사한 이미지를 찾아내는 기능을 향상시키는 연구를 촉진시켰다. 이와 같은 기술의 실용화는 다양한 분야를 포괄하며, 대표적으로 쇼핑물의 상품 검색, 검색 엔진 등에 응용되어 사용자의 편의를 제고하고 있다. 이에 따라 상품명에 대한 정보가 없는 상황에서도 단순한 이미지 정보를 통해 원하는 상품을 검색하는 것이 가능해졌다. 그러나, 실제 세계의 이미지에는 다양한 객체들이 복잡하게 혼재하고 있어 CNN 알고리즘 단독으로는 이미지 내부의 객체를 정확히 분석하고, 그 객체가 포함된 다른 이미지들을 효과적으로 검색하는데 한계가 있음이 인지되고 있다. 본 연구는 이러한 문제점을 개선하기 위해 객체 탐지 알고리즘을 적용하는 방안을 모색하였다. 본 논문에서는 객체 탐지 알고리즘을 통해 이미지 내부의 객체를 분석하고, 그에 따른 유사 객체를 포함하는 이미지를 찾아내는 전략을 제시한다. 이를 통해 이미지 분석 및 검색의 정확성을 더욱 향상시킬 수 있는 가능성을 제안한다.

키워드: 객체 탐지(object detection), 이미지 검색(image search), 합성곱신경망(convolution network)

I. Introduction

인공지능과 디지털 기술의 발전은 정보의 수집, 처리, 그리고 활용에 대한 새로운 접근법을 제시하였다. 특히 이미지 검색 기술은 큰 변화를 겪었다. 이전까지 이미지 검색은 대부분 텍스트 메타데이터에 의존했다. 하지만 최근 들어서는 이미지의 시각적 특성을 분석하는 콘텐츠 기반 검색(CBIR), 그리고 딥러닝에 기반한 방법론이 크게 주목받고 있다. 이런 기술적 진보는 디지털 플랫폼에서의 상품 검색부터 이미지 검색 사이트까지, 다양한 분야에서 활용되고 있다.

그럼에도 불구하고, 현재의 이미지 검색 기술은 여전히 한계를 가지고 있다. 실제 세계의 이미지는 다양한 객체들이 복잡하게 혼재하고 있으며, 이를 정확하게 분석하고 객체를 포함하는 이미지를 효과적으로 검색하는 것은 쉽지 않다. 이를 극복하기 위해 본 논문에서는 이미지 내의 객체를 더욱 정확히 파악하고, 이를 통해 유사 객체를 포함하는 이미지를 효과적으로 찾아내는 방법에 대해 제안한다. 특히, 본 논문에서는 객체 탐지 알고리즘을 활용한 이미지 분석 및 검색

방법에 대해 다룬다. 본 연구가 제안하는 접근법은 객체 탐지 알고리즘과 CNN을 결합하여 이미지 내부의 복잡한 객체 구조를 더욱 정확히 파악하고, 이를 기반으로 유사한 객체를 포함하는 다른 이미지를 효과적으로 검색하는 방법론을 제시한다. 이는 이미지 검색의 정확성을 향상시킨다.

II. Related works

1. YOLO(You Only Look Once)

본 연구의 주요 기술적 요소 중 하나는 실시간 객체 인식 알고리즘인 YOLO(You Only Look Once)이다. YOLO는 그 이전의 탐지 방법들에 비해 뛰어난 실시간 처리 능력과 높은 인식 정확도를 가진 알고리즘으로, 이미지 내의 객체를 신속하게 인식하고 분류할 수 있다. YOLO는

기존의 R-CNN이나 Fast R-CNN 등과 같은 기존의 객체 탐지 방법과는 다르게 전체 이미지를 한 번에 처리하여 속도와 정확성을 동시에 증가시키는 방식을 제안하였다. 이러한 방식은 객체 탐지를 단일 회귀 문제로 바꾸어 해결하였다. 이후 YOLO v2 (Darknet-19)는 YOLO의 초기 버전의 문제점들을 개선하였다. 높은 해상도의 그리드에서 예측을 수행하여 작은 객체에 대한 탐지 성능을 향상시켰으며, 다양한 크기의 앵커 박스를 도입하여 다양한 비율의 객체를 처리할 수 있게 되었다. YOLO v3는 Darknet-53라는 새로운 컨볼루션 레이어를 사용하고, 세 가지 스케일에서 예측을 수행하여 다양한 크기의 객체를 탐지한다. 이와 함께, 다양한 가로세로 비율을 가진 세 가지 앵커 박스를 사용한다.

YOLO의 버전별 차이점은 Table 1이 보이는 바와 같다.

Table 1. YOLO versions

	YOLO v1	YOLO v2	YOLO v3
Pre-trained Network	Inception	Darknet-19	Darknet-53
Anchor Box 개수	2개	5개	9개
Feature Map Size	7x7	13x13	13x13 26x26 52x52

YOLO v3는 기본적으로 Darknet-53라는 컨볼루션 레이어를 사용한다. Darknet-53은 총 53개의 컨볼루션 레이어로 구성되어 있으며, 이는 ResNet과 비교해 비슷한 성능을 내면서도 더 적은 파라미터를 사용하여 더 빠르게 동작한다. YOLO v3는 다양한 크기의 객체를 탐지하기 위해 세 가지 스케일에서 예측을 수행한다. 이를 통해 알고리즘은 큰, 중간, 작은 크기의 객체들을 동시에 인식하고 분류할 수 있다. YOLO v3는 각각 다른 가로세로 비율을 가진 세 가지 앵커 박스를 사용한다. 앵커 박스는 초기 바운딩 박스 예측에 사용되며, 다양한 모양과 크기의 객체를 적절하게 캡처하는 데 도움을 준다.

YOLO v3의 이러한 기능들은 높은 정확도와 실시간 처리 능력을 통합하여 이미지 내의 객체 인식을 향상시키는 중요한 요소이다. 본 연구에서는 YOLO v3를 기반으로 한 객체 탐지와 피쳐맵 기반 유사도 계산을 통합하여 이미지 내의 복잡한 객체 구조를 세밀하게 분석하고 이를 통해 유사한 객체를 포함하는 다른 이미지를 효율적으로 검색할 수 있었다. 이로써 YOLO v3의 실용성과 강력함이 다시 한번 확인되었다.

III. The Proposed Scheme

본 연구는 입력 이미지 내의 객체를 신속하고 효과적으로 탐지하고, 이를 바탕으로 이미지를 검색하는 방법론을 제시한다. 본 연구에서 제안하는 접근법은 아래의 프로세스로 구성된다

1. 객체 인식

초기 단계에서, 입력 이미지에 대해 객체 인식을 수행한다. 이 때 적용하는 알고리즘은 실시간 객체 인식에서 뛰어난 성과를 보이는 YOLO(You Only Look Once)이다. 이 알고리즘은 각 객체를 정확하게 인식하고, 해당 객체를 바운딩 박스로 표현한다.

2. 객체 분할

YOLO를 통해 식별된 객체들은 그에 상응하는 바운딩 박스를 기반으로 이미지에서 분리된다. 이로써, 관련성이 높은 객체들에 대한 세부적인 분석이 가능해지며, 이후 단계에서의 유사도 측정에 필요한 정보를 제공한다.

3. 특징 맵 추출

분리된 객체에 대해, 특징 맵을 추출한다. 이 과정에서는 CNN(Convolutional Neural Network)을 사용하여 고차원의 이미지 특성을 저차원의 특징 맵으로 변환한다. 동일한 과정을 데이터베이스의 이미지에도 적용하여 각 이미지의 특징 맵을 사전에 계산한다.

4. 유사도 계산

입력 이미지에서 분리된 객체의 특징 맵과 데이터베이스의 이미지 특징 맵 간의 유사도를 계산한다. 이 때, 유사도 측정 방법으로는 L2 유사도를 채택하였다. 이 방식은 두 특징 맵 간의 거리를 계산하여 객체들 간의 유사성을 판정한다.

5. 이미지 검색

마지막으로, 계산된 유사도를 바탕으로, 입력 이미지와 가장 유사한 데이터베이스 내의 이미지를 반환한다.

본 논문에서 제안한 시스템의 구조는 Fig. 2와 같다.

본 연구가 제안하는 방법론은 객체 인식 알고리즘과 유사도 기반 이미지 검색의 강점을 통합하여, 이미지 내의 복잡한 객체 구조를 상세하게 인식하고 이를 이용해 유사한 이미지를 효율적으로 검색할 수 있게 한다.

IV. Conclusions

본 논문에서는 실제 세계의 복잡한 이미지에 대한 분석과 검색을 보다 효과적으로 수행하는 새로운 방법론을 제안하였다. 객체 인식 알고리즘인 YOLO와 피쳐맵 기반 유사도 계산을 결합함으로써, 이미지 내의 객체들을 더욱 정확하게 인식하고 이를 통해 유사한 객체를 포함하는 다른 이미지를 효율적으로 검색할 수 있었다.

본 연구에서 제안한 방법론은 이미지 내의 복잡한 객체 구조를 세밀하게 분석하며, 이를 통해 기존의 텍스트 메타데이터에 의존하는

검색 방법보다 훨씬 정확한 결과를 도출하는 것이 가능했다. 이는 다양한 분야, 특히 상품 검색, 이미지 검색 등에서의 활용 가능성을 보여준다. 또한, 본 연구는 이미지 분석과 검색에 관한 미래 연구 방향성을 제시하였다. 특히, 객체 인식과 유사도 계산을 활용한 이미지 분석은 실제 세계의 복잡한 이미지에 대한 더욱 깊이 있는 이해를 가능하게 하며, 이를 바탕으로 향후 보다 정교한 이미지 검색 알고리즘의 개발이 기대된다.

본 연구의 제안 방법론은 특정 상황에서만 유효한 방법이 아닌, 다양한 상황과 필드에서 활용 가능한 유연성을 가지고 있다. 이러한 점을 바탕으로, 본 연구에서 제시된 방법론이 이미지 분석 및 검색 분야의 기술 발전에 기여할 것으로 기대된다.

ACKNOWLEDGEMENT

본 연구는 과학기술정보통신부 및 정보통신기획평가원의 SW중심대학지원사업의 연구결과로 수행되었음(2019-0-01817)

REFERENCES

- [1] E. Smirni, and G. Ciardo, "Workload-Aware Load Balancing for Cluster Web Servers," IEEE Trans. on Parallel and Distributed Systems, Vol. 16, No. 3, pp. 219-232, March 2005.
- [2] Kdhong, "An Efficient Dynamic Workload Balancing Strategy," Journal of The Korea Society of Computer and Information, Vol. 15, No. 1, pp. 1-10, Nov. 2010.
- [3] D.H.Ballard, "Computer Vision," Prentice-Hall, pp.76-79, 1991.
- [4] Kdhong, "C Programming Language" Korea-Press, pp.100-120, 1991.
- [5] SIMGRID Project, <http://simgrid.gforge.inria.fr>