

장소와 유형 분리 신경망 기반의 고정형 대상 인식

이형석, 이재호*, 김도형***, 류철***
한국전자통신연구원

hyslee@etri.re.kr, *bigleap@etri.re.kr, **dhkim@etri.re.kr, ***ryuch@etri.re.kr

Fixed object recognition based on place and type separation neural network

Hyung-Seok Lee, Jae-Ho Lee*, Do-Hyung Kim**, Cheol Ryu***
Electronics and Telecommunications Research Institute

요 약

본 논문에서는 해결하고자 하는 문제는 위치 고정형 객체들이 서로 다른 장소에 존재할 때 유형이 비슷하다 하더라도 장소와 대상을 정확히 인식하는 것이다. 이미지 기반으로 고정형의 대상 객체를 인식하는 문제를 풀기위한 한 방법으로 장소 인식과 객체 유형 인식을 분리하여 수행한 다음 결과를 통합하여 최종 인식을 하는 방법을 제시한다.

I. 서 론

포스트 스마트폰으로 전망되어지는 스마트 글래스[1]에서 정면에 보이는 물체를 인식하는 것이 중요한 핵심 기능이다. 서로 다른 장소에서 같은 유형이나 비슷한 대상들이 다수가 있을 경우에 있어서 단순한 객체의 유형 분별을 한 딥러닝 신경망[2]을 단순히 사용하는 것은 적절치 않다. 객체가 존재하는 배경과 함께 촬영된 이미지들을 학습한 단일 신경망을 통하여 문제를 해결하는 방법을 먼저 생각해 볼 수 있다. 즉, 다른 장소에 있는 같은 유형의 대상 물체가 신경망에서 서로 다른 카테고리 이 방법은 대상을 특정 장소에 설치한 다음 학습할 이미지를 확보해야 하므로, 만일 장소를 이동한다면 새로운 학습 데이터를 확보하고 학습도 새로 수행해야 하는 문제가 있다. 그리고, 다른 장소에 같은 유형의 물체가 있는 경우 신경망에서 서로 다른 카테고리 로 분별하기 어렵게 된다.

본 논문에서는 이미지 기반으로 고정형의 대상 객체를 인식하는 문제를 풀기위해서 장소 인식과 객체 유형 인식을 분리하여 수행한 다음 결과를 통합하여 최종 인식을 하는 방법을 제시한다.

II. 본론

해결하고자 하는 문제는 위치 고정형 객체들이 서로 다른 장소에 존재할 때 유형이 비슷하다 하더라도 장소와 대상을 정확히 인식하는 것이다. 서로 다른 공간에서의 같은 유형을 분별할 수 있어야 한다. 이러한 문제를 해결하기 위해서 다중 구조 신경망을 활용하여 객체 종류 분류 뿐 만이 아니라 객체가 설치된 공간 정보도 동시에 사용할 필요가 있다. 이미지에 포함된 객체의 유형을 알고 그 객체가 존재하는 공간의 위치 정보를 알 수 있으면 대상을 더욱더 정확하게 식별할 수

있기 때문이다. 공간과 객체 이미지를 분리하여 각각을 인식하고 결과를 종합하여 대상을 추론함으로써, 공간에 대한 학습과정만 거치면, 객체의 유형에 대한 기존의 딥러닝 신경망을 그대로 활용할 수 있는 장점이 있다.

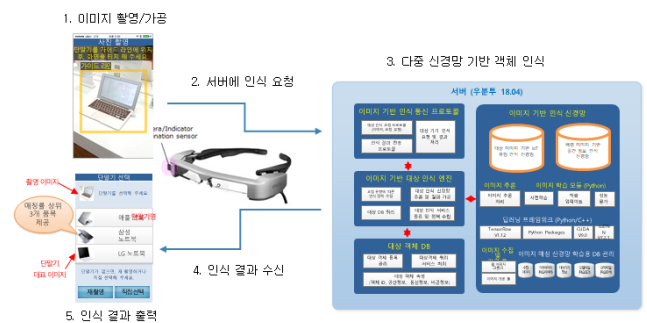


그림 1 이중 신경망 기반 객체 인식 구성

그림 3에서 공간과 객체를 각각 분별하는 이중 신경망 기반 객체인식 시스템의 구성을 보여준다. 단말은 안드로이드 기반으로 서버는 우분투 리눅스를 기반으로 구현하였다. 단말과 서버와의 통신은 웹서비스를 통한 REST 인터페이스를 사용하였다. 장소 및 기기를 분별하기 위한 기본 모델을 Inception v3[3]을 사용하였다. 기존 학습된 모델로부터 출력단만 본 시스템에 맞게 재 구성하고 재 학습하는 작업을 수행하였다.

안경형 단말(Moverio BT-350)에서 인식하고자 하는 물체를 카메라 중앙의 사각 박스에 두고 인식 버튼을 누르면, 인식 이미지가 서버로 전송되고 서버에서 인식 결과를 수신하는 구조이다. 배경을 포함한 전체 촬영 이미지를 사용하여 장소를 분별한다.

실험을 위하여 3 개 장소와 8 개의 종류의 기기를 대상으로 단말에서 촬영된 장소와 기기 이미지를 서버에

전송한 후 인식 결과를 단말이 수신하여 인식률과 처리 시간 정보를 기반으로 성능 분석을 하였다. 이를 위하여 장소와 기기들 각각 100 개 이상의 이미지를 학습하였다.

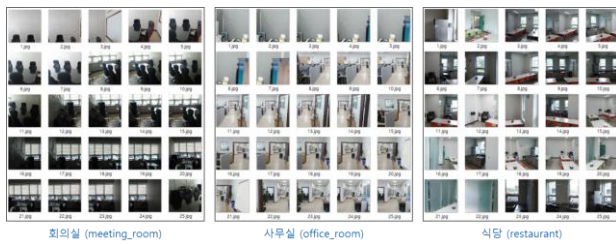


그림 2 장소 학습 데이터

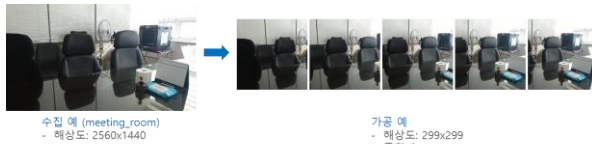


그림 3 장소 이미지 수집 및 가공 예

장소의 경우 배경 이미지를 활용하는 장소의 경우 신경망의 입력은 중형비 1:1 을 요구하므로 그림 3 와 같이 가공하여 사용하였다.



그림 4 학습에 사용된 8 개의 대상 기기 이미지 예

실험은 3 개 장소에서 10 개 기기를 대상으로 기기당 10 회 인식 실험을 수행하였다. 10 개 중 중복 유형의 기기는 2 종이어서 실제 기기 종류는 총 8 종이다. 대상 인식 시 장소 및 기기의 매칭 값이 모두 주어진 임계 값 이상일 경우 인식 성공으로 간주하였다. 3 개 장소에서 각각 등록되지 않은 기기 혹은 배경 등을 대상도 포함하여 전체 실험 수는 130 회이다.

	목표치 (Max=100)										평가 방법 (가산점 포함)									
	1	2	3	4	5	6	7	8	9	10	평가 방법 4 0 0 0 0 0 0 0 0 0	평가 방법 0 0 0 0 0 0 0 0 0 0								
01. 해외봉사 프로그램	0.80	0.95	0.88	0.88	1.00	0.79	1.00	0.82	0.88	0.99	1.00	0.96	0.88	0.95	1.00	0.75	9	9		
02. 해외봉사 실적 인정	0.99	0.92	0.99	0.99	0.98	0.91	0.78	0.98	0.94	0.92	0.89	0.73	0.82	0.88	0.96	0.84	9	8		
03. 해외봉사 실적 인정	0.90	0.90	0.87	0.90	0.90	0.90	0.78	0.79	0.88	0.84	0.88	0.88	0.88	0.88	0.88	0.88	8	8		
04. 해외봉사 실적 인정	0.96	0.96	0.96	0.96	0.96	0.96	0.96	0.96	0.96	0.96	0.96	0.96	0.96	0.96	0.96	0.96	10	10		
05. 자원봉사 프로그램 운영	0.99	0.93	0.95	0.82	0.76	0.82	0.93	0.93	0.85	0.93	0.96	0.76	0.72	0.75	0.67	0.75	6	7		
06. 자원봉사 실적	0.86	0.98	0.96	0.82	0.94	0.73	0.88	0.98	0.92	0.96	0.97	0.93	0.98	0.90	0.95	0.93	8	8		
07. 자원봉사 실적	0.86	0.95	0.87	0.95	0.98	0.87	0.86	0.91	0.88	0.92	0.86	0.85	0.88	0.87	0.87	0.87	0.90	10	10	
08. 자원봉사 실적	0.61	0.97	0.68	0.98	0.97	0.60	0.61	0.61	0.93	0.98	0.84	0.87	0.71	0.93	0.66	0.82	0.93	6	6	
09. 자원봉사 실적	0.99	0.77	0.98	0.99	0.97	0.99	0.98	0.98	0.99	0.93	0.85	0.88	0.88	0.84	0.99	0.96	9	8		
10. 자원봉사 실적	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	10	10		
11. 자원봉사 실적	0.79	0.25	0.65	0.91	0.40	0.93	0.42	0.97	0.37	0.98	0.34	0.25	0.62	0.41	0.73	0.61	0.51	5	5	
12. 자원봉사 실적	0.84	0.32	0.62	0.38	1.00	0.33	0.69	0.37	0.29	0.99	0.38	0.74	0.42	0.42	0.72	0.26	0.34	0.33	10	10
연속 달성률 (연속 3개월 연속 달성률) = 달성률 3 * 0.3333333333333333													100	100						
연속 달성률 (연속 3개월 연속 달성률) = 달성률 3 * 0.3333333333333333													100	100						
연속 달성률 (연속 3개월 연속 달성률) = 달성률 3 * 0.3333333333333333													1	0						
연속 달성률 (연속 3개월 연속 달성률) = 달성률 3 * 0.3333333333333333													20	25						
연속 달성률 (연속 3개월 연속 달성률) = 달성률 3 * 0.3333333333333333													83.85	80.77						
연속 달성률 (연속 3개월 연속 달성률) = 달성률 3 * 0.3333333333333333													9.77	9						
연속 달성률 (연속 3개월 연속 달성률) = 달성률 3 * 0.3333333333333333													15.39	15.23						

그림 5 인식률 성능 평가

그림 5 는 실험 결과를 보여준다. 인식률은 장소와 기기 모두 인식에 성공한 수치, 오인식률은 장소와 기기 중 하나라도 실패, 미인식률은 장소와 기기 중 하나라도 임계값을 만족시키지 못하여 인식에 실패한 비율이다.

실험 결과 인식률은 장소 매칭값 0.6 이상, 기기 매칭값 0.7 이상일 경우, 83.85%로 측정되었고, 장소 매칭값 0.6 이상, 기기 매칭값 0.75 이상일 경우, 80.77%로 측정되었다. 매칭값에 따라 인식률에 차이가

발생하며, 매칭값이 너무 낮을 경우 오인식률이 증가하여 인식률이 저하되었다.

수행 시간에 대한 성능 측정은 안경형 단말(Moverio BT-350) 뿐 만이 아니라 스마트폰에서 수행하여 시간 성능을 측정하였다. Moverio BT-350 에서는 약 6 초 정도의 인식 시간이 소요되었으며, 이는 최신 스마트폰에 비해 약 2 배 느린 속도를 보였다.

	이미지 촬영	인식 요청 및 결과	전체 인식 시간
안경형 단말(Moverio BT-350)	1.309초	4.903초	6.212초
스마트폰(Galaxy Note 9)	0.961초	3.031초	3.992초
스마트폰(Galaxy S10)	0.793초	2.317초	3.110초

III. 결론

본 논문에서는 스마트 글래스 같은 웨어러블 장치에서 사용자가 보이는 대상 객체에 대한 서비스를 제공하는 비주얼 서비스를 지원하기 위하여 이미지 기반으로 대상 객체를 인식하는 한 방법으로, 장소와 기기 유형을 각각 다른 신경망을 통하여 인식하고 결과를 종합하는 방식을 제안하고 실험을 통하여 성능을 분석하였다.

본 연구에서는 인식 대상을 포함하는 한 이미지로부터 장소를 인식 분류하는 기법을 사용하였는데 이 이미지의 배경만으로는 장소 인식에 다소 오류가 있음을 확인하였다. 이를 보완하는 차기 연구로 여러 개의 이미지로부터 인식을 한 후에 이 결과를 종합하여 장소를 판단하는 것을 생각해 볼 수 있다. 또한 본 논문에서는 대상 객체가 특정 장소에 고정되어 있는 것만 고려하였는데, 이동형 객체에 대한 문제에 대해서도 차기 연구과제가 될 수 있을 것이다.

ACKNOWLEDGMENT

이 논문은 2020 년도 정부(과학기술정보통신부)의 재원으로
정보통신기획평가원의 지원을 받아 수행된 연구임 (2018-0-
00226, 포스트 스마트폰 시대를 대비한 Trusted Reality
핵심기술 개발)

참 고 문 헌

- [1] K. Song, G. Kim, T. Kim, C. Ryu, S. Park, J.H. Lee, J.K. Lee, and S. Hwang, "Trusted Reality Technology, from a Post-Smartphone Perspective," *Electronics and Telecommunications Trend*, 2018.
- [2] Alex Krizhevsky, Ilya Sutskever, Geoffrey E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," *Proceedings of the 25th International Conference on Neural Information Processing Systems(NIPS) Volume 1* pp. 1097-1105, Dec. 20
- [2] Christian Szegedy, Vincent Vanhoucke, Sergey Ioffe, Jon Shlens, Zbigniew Wojna, "Rethinking the Inception Architecture for Computer Vision," 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)