

효율적인 사물 이미지 분류를 위한 계층적 이미지 분류 체계의 설계 및 구현*

유 태 우*, 김 윤 옥*, 정 하 민*, 유 현 수*, 안 용 학**

요 약

본 논문에서는 효율적인 사물 이미지 분류를 위한 계층적 이미지 분류 체계 방안에 대해 제안한다. 기존의 전체 이미지를 한 번에 분류하는 무 계층 이미지 분류에서는 상대적으로 유사한 모양을 가진 사물은 효율적으로 인식하지 못하는 모습을 보여줬다. 따라서 본 논문에서는 사물 이미지에 대해 계층적으로 분류를 시도하는 단계적 계층 구조에서의 이미지 분류 기법을 소개한다. 또한, 실제 시스템에 딥 러닝 이미지 분류가 적용되었을 때 발생할 수 있는 확장성에 대해서 고려하기 위해 확장성이 고려된 효율적인 클래스 구성 방식과 알고리즘도 소개한다. 이와 같은 방식은 상대적으로 유사한 형태를 보인 사물 이미지에 대해 더 높은 신뢰도로 이미지를 분류하는 것을 가능하게 한다.

Design and Implementation of Hierarchical Image Classification System for Efficient Image Classification of Objects

Taewoo You*, Yunuk Kim*, Hamin Jeong*, Hyunsoo Yoo*, Yonghak Ahn**

ABSTRACT

In this paper, we propose a hierarchical image classification scheme for efficient object image classification. In the non-hierarchical image classification, which classifies the existing whole images at one time, it showed that objects with relatively similar shapes are not recognized efficiently. Therefore, in this paper, we introduce the image classification method in the hierarchical structure which attempts to classify object images hierarchically. Also, we introduce to the efficient class structure and algorithms considering the scalability that can occur when a deep learning image classification is applied to an actual system. Such a scheme makes it possible to classify images with a higher degree of confidence in object images having relatively similar shapes.

Key words : 딥러닝(Deep Learning), 계층 구조(Hierarchy Structures), 이미지 분류(Image Classification), 이미지 인식(Image Recognition)

접수일(2018년 8월 11일), 게재확정일(2018년 9월 23일)

★ 본 연구는 과학기술정보통신부 및 정보통신기술진흥센터의 SW중심대학 사업의 연구결과로 수행되었음. (2015-0-00038)

* 세종대학교 디지털콘텐츠학과

** 세종대학교 컴퓨터공학과(교신저자)

1. 서 론

1.1 연구 배경 및 내용

최근 딥 러닝 연구가 활발히 진행됨에 따라 많은 기업에서 이미지 인식, 음성 인식 및 영상 인식 등을 이용한 새로운 형태의 서비스가 많이 출시되고 있다.[2][3]

딥 러닝을 활용한 다양한 서비스가 등장하고 있고, 이러한 서비스는 더욱 발전된 삶을 제공해준다. 하지만 사용자가 기대하는 결과를 정확하게 제시하지 못하는 경우 서비스의 사용성이 떨어지게 된다. 따라서 인공지능을 이용한 모든 서비스는 높은 정확도와 신뢰도를 요구하고, 이에 따라 인식률을 높이기 위한 다양한 연구가 시도되고 있다.

따라서 본 논문은 딥 러닝의 인식률을 향상하기 위한 연구 분야 중 하나인 이미지 분류에서 단계적 계층 구조를 활용하여 정확성을 높일 수 있는 방향을 제시한다.

1.2 관련 연구 및 동향

딥 러닝을 이용한 이미지 분류에서 이미지 인식률을 높이기 위한 여러 가지 방법이 시도되고 있다. 건물의 공간 객체를 인식해 사무 집기를 자동으로 관리하기 위한 연구[4]와 음식 사진을 찍음으로써 음식에 대한 정보를 조회하는 방법[5] 등의 연구에서는 일반적인 분류 방식을 사용하여 이미지를 분류했다. 이와 같은 방식은 항목에 따라 낮은 인식률이 나올 수 있기 때문에 사용성이 떨어질 수 있다.

따라서 위와 같은 문제를 보완하려는 방법으로 3단계 계층 구조를 이용한 이미지 분류에 관한 연구[6]가 진행되었고, 이 연구를 한 단계 발전시켜 더 세분된 특징을 이용한 다중 계층 이미지 분류[7]와 같은 방식이 제시되었다. 하지만 [6]과 [7]에서 클래스를 분류하기 위한 방식은 고정된 분류 체계를 사용하기 때문에 새로운 물체에 대한 확장성이 부족하다. 본 논문의 대

상이 되는 유실물 통합 관리 시스템에서 딥 러닝을 활용해 유실물 카테고리의 자동 분류를 구현한다면 새로운 객체가 추가될 때마다 클래스를 재구성해야 하는 문제점이 있다. 또한, 이는 새로운 물체를 추가하기 위한 유연성이 부족하고, 실제 서비스에 적용했을 때 클래스를 분류하기 위한 근본적인 수정이 필요하다는 단점이 있다.

따라서 본 논문에서는 효율적인 사물 이미지 분류를 위한 계층적 이미지 분류 체계 및 클래스 설계 방안을 소개한다. 이러한 계층적 이미지 분류 체계와 확장성을 고려한 클래스 설계는 새로운 물체를 유연하게 추가할 수 있고, 일반적인 휴대 물품처럼 상대적으로 유사한 사물 이미지에 대해 높은 신뢰도로 효율적인 이미지 분류를 가능하게 한다.

2. 단계적 계층 이미지 분류 구조

본 논문에서는 이미지 인식률 향상을 위해 새로운 학습 모델 제시가 아닌 Tensorflow와 Inception-v3 모델[1]을 활용한 단계적 계층 이미지 분류 체계를 제안한다.

이를 위해 첫째, 객체들의 공통된 특징을 기반으로 클래스를 나누는 것이 선행되고, 둘째, 단계적으로 클래스를 분류해 객체를 인식한다.

2.1 클래스 분류

본 논문에서는 <표 1>에 제시한 것처럼 총 36개의 항목에 대해 이미지 분류를 진행한다. 이 항목들은 경찰청에서 운영하는 lost112.go.kr의 유실물 카테고리를 토대로 사람들이 주로 분실하는 항목을 랜덤하게 선정한 것들로서 본 논문에서 제시한 방식을 토대로 이미지 분류의 정확성에 초점을 맞추고, 이미지 인식을 위한 단계적 구조의 이미지 분류 방식에 관해 서술한다.

<표 1> 기준 물품 목록

연번	국문	영문	연번	국문	영문
1	MP3	mp3	19	선글라스	sunglasses
2	거울	mirror	20	시계	watch
3	계산기	calculator	21	열쇠	key
4	귀걸이	earring	22	우산	umbrella
5	노트북	laptop	23	이어폰	earphone
6	담요	blanket	24	장갑	gloves
7	립밤/립스틱	lip balm	25	지갑	wallet
8	마우스	mouse	26	충전기	charger
9	머리끈	hair strap	27	카드	card/identification
10	머리띠	headband	28	카메라	camera
11	모자	hat	29	텀블러	tumbler
12	목걸이	necklace	30	파우치	pouch
13	목도리	scarf	31	펜	pen
14	반지	ring	32	필통	pencil case
15	보조배터리	supplementary battery	33	향수	perfume
16	빗	comb	34	헤드셋	headphone
17	생리대	sanitary napkin	35	현금	cash
18	서적	book/notebook	36	휴대폰	phone

2.2 클래스 분류 방식

본 논문에서 분류에 기준이 되는 대표적인 특징은 “재질, 유형, 용도, 범주”이다. 각 항목을 효율적으로 나눌 수 있는 특징들로 구성하였으며 특징을 추가하거나 변경해서 더욱 효율적으로 분류할 수 있는 특징을 정의하는 것이 가능하다.

특징이 정의된 후 특징별로 객체마다 적합한 세부 특징에 대해 기입한다. ‘재질’은 객체를 이루는 주된 성질을 의미하며 ‘유형’은 객체의 외형이나 생김새 등을 의미하고, ‘용도’는 객체가 주로 사용되는 용도를 의미하며 ‘범주’는 객체의 종류를 설명할 수 있는 카테고리 의미를 의미한다. 단, 세부 특징을 기입할 때는 해당 객체의 모든 아이টে姆에 통상적으로 적용할 수 있을 때만 기입한다. 예를 들어, 담요는 면이나 천처럼 직물로 이루어져 있기 때문에 담요를 분류할 때 직물이라는 세부 특징을 기입할 수 있다. 반면, 텀블러와 같이 한 가지 재질이 아닌 여러 재질로 만들어지는 경우 하나의 세부 특징을 선정할 수 없다. 이 같은 경우는 공백으로 표시한다.

2.3 특징을 이용한 클래스 분류

위 2.2절에서 제시한 방법으로 최종적으로 정해진

특징 및 세부 특징들을 이용해 객체들의 클래스를 정의한다. 우선 “재질, 유형, 용도, 범주”와 같은 특징들에 대해 우선순위를 선정한다. 그리고 우선순위를 기준으로 특징별로 같은 세부 특징을 가지는 항목들을 하나의 클래스로 묶어가며 하나씩 클래스를 구성한다. 하지만 무조건 우선순위가 더 높은 특징을 먼저 이용해 클래스를 구성할 경우 비효율적인 클래스가 구성되는 문제가 발생한다. 예를 들어 서적, 카드, 현금을 이용하여 클래스를 구성할 경우 ‘재질’의 우선순위가 ‘유형’보다 높으므로 서적과 현금이 하나의 클래스로 구성되고, 카드는 기타 클래스로 구성된다. 하지만 ‘유형’을 기준으로 분류하면 세 물품은 ‘평평함’이라는 한 클래스로 구성된다. 따라서 우선순위가 낮아도 기타 클래스에 속하는 항목들의 개수를 줄이는 방향으로 클래스를 구성한다.

실제로 시스템에 적용했을 때는 새로운 객체가 계속해서 추가될 수 있기 때문에 클래스의 확장성을 고려해야 한다. 제시한 내용을 토대로 실제 시스템에서 단소를 새롭게 추가하고자 할 때 ‘재질’은 나무 혹은 플라스틱이기 때문에 빈칸으로 두고 ‘유형’은 막대, ‘용도’는 음악 연주, ‘범주’는 악기라고 기입하면 새로운 클래스 생성 없이 기존 막대 클래스에 추가하게 된다. 하지만 확장성이 고려되지 않은 분류체계에서는 새로운 객체인 단소를 추가하기 위해 새로운 카테고리인 악기를 만든다. 이는 하나의 객체를 위해 새로운 클래스를 연쇄적으로 생성해야 함으로써 클래스의 활용성이 떨어진다는 문제가 있다.

2.4 클래스 생성을 위한 알고리즘

분류하고자 하는 항목의 개수와 특징 및 세부 특징이 많아질 경우 사람이 모든 경우의 수를 확인하고 클래스를 구성하는 것에 한계가 있기 때문에 클래스를 자동으로 구성하기 위한 알고리즘을 제시한다. 이는 깊이 우선 탐색(DFS: Depth-First Search)[13]으로 이루어져 있으며, 조합 가능한 모든 경우를 확인하지만 불필요한 경우의 수를 제거하는 방식으로 구성하였다. 사용자가 기입한 특징들을 추출한 후 클래스를 구성하는 최소 개수 이상으로 가지고 있는 특징들을 이용해 특징 목록을 구성한다. 그 후, 프로그램은 다음

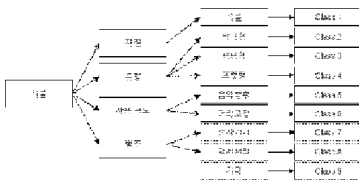
과 같은 순서로 탐색을 진행한다.

0/직물 0/종이 1/막대 1/평평 1/포켓
0/직물 1/막대 0/종이 1/평평 1/포켓
0/직물 1/막대 1/평평 0/종이 1/포켓
0/직물 1/막대 1/평평 1/포켓 0/종이
...
1/막대 0/직물 0/종이 1/평평 1/포켓
1/막대 0/직물 1/평평 0/종이 1/포켓
1/막대 0/직물 1/평평 1/포켓 0/종이
...
1/막대 1/평평 0/종이 0/직물 1/포켓
1/막대 1/평평 0/종이 1/포켓 0/직물
1/막대 1/평평 1/포켓 0/직물 0/종이

우선순위를 토대로 “0/직물 0/종이 1/막대 1/평평 1/포켓”을 탐색할 경우 1번 특징들은 순서를 바꿔도 같은 결과 값을 보여주므로 불필요한 반복을 하지 않도록 하여 시간을 단축한다. 해당 알고리즘은 <표 2>와 같은 결과를 출력하며 클래스를 분류하는 모습을 볼 수 있다. <표 2>과 (그림 1)은 2.1절에서 정의한 객체들을 특징별 분류 알고리즘을 이용해 클래스를 구성한 결과이다.

<표 2> 분류된 클래스

Class 1	단요 모자 목도리 생리대 장갑
Class 2	침방/침스틱 우산 랜
Class 3	시적 카드 현금
Class 4	지갑 파우치 필통
Class 5	이어폰 헤드셋
Class 6	머리끈 머리띠
Class 7	MP3 계산기 노트북 마우스 보조배터리 충전기 카메라 휴대폰
Class 8	가을 귀걸이 목걸이 만지 및 신글라스 시계
Class 9	열의 텀블러 향수



(그림 1) 최종 계층 구조

3. 실험 및 결과

3.1 실험 환경 및 실험 데이터

본 논문에서 이미지 학습 및 테스트에 사용된 환경은 <표 3>과 같다.

<표 3> 이미지 학습 및 테스트 환경

분류	학습 환경
Operating system	Windows 10 Pro
CPU	Intel Core i5-3570
RAM	8GB
Storage	SSD 512GB
GPU	NVIDIA GTX 1060 6GB

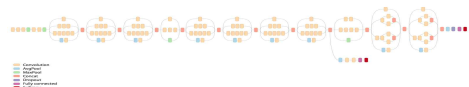
학습 및 테스트에 사용된 이미지는 각종 포털사이트에서 무작위로 수집하였으며 각 항목별 개수는 <표 4>와 같고, 총 143,136개(15.4GB)로 구성되어 있다.

<표 4> 이미지 개수 (단위: 개)

연번	국문	학습 이미지	테스트 이미지	총 이미지
1	MP3	2,070	690	2,760
2	거울	1,154	385	1,539
3	계산기	2,650	884	3,534
4	귀걸이	4,021	1,340	5,361
(중략)				
31	랜	3,198	1,066	4,264
32	필통	2,029	676	2,705
33	향수	3,790	1,263	5,053
34	헤드셋	2,432	811	3,243
35	현금	2,411	804	3,215
36	휴대폰	3,123	1,041	4,164
합계		107,351	35,785	143,136

3.2 이미지 학습 모델 및 방법

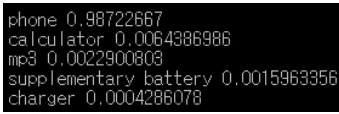
본 논문에서 이미지 학습은 Google의 Inception-v3[1] 모델과 Tensorflow를 적용한다[12]. 적용된 Inception-v3 모델은 (그림 2)와 같다.



(그림 2) Inception-v3 model architecture

3.3 실험 방법

객체별 전체 이미지 개수에서 75%의 이미지는 학습에 사용하고, 25%의 이미지는 테스트에 사용하였다. 무 계층에서 사용한 테스트 이미지와 단계적 계층 구조에서 사용된 테스트 셋은 동일하고, 객체별로 몇 개의 이미지가 정확히 선별되었는지에 대한 개수를 구한 뒤 모든 객체의 평균을 인식률로 사용한다. (그림 3)과 같이 최소 90% 이상 확률로 첫 번째로 이미지 분류에 성공하였을 때 이미지 인식에 성공한 것으로 정의한다.



(그림 3) 이미지 최종 인식 결과

3.4 무 계층에서의 실험 결과

단계적 계층 구조를 적용함에 따라 어느 정도의 인식이 상승하는지 비교를 위해 무 계층 구조에서 테스트를 진행하였다. 테스트 결과는 <표 5>와 같다.

<표 5> 무 계층 구조 테스트 결과 요약

신뢰도	분류 성공(개)	분류 실패(개)	비율(%)
90% 이상	22,272	13,513	62.24
95% 이상	17,412	18,373	48.66
99% 이상	8,800	26,985	24.59

이는 기준 정확도가 높아질수록 정확하게 분류하는 이미지의 개수가 대폭 감소하는 모습을 보여주며, 실제 이미지 분류를 이용하여 소비자에게 서비스를 제공하기 위한 시스템에 적용하기 힘든 수치를 보여준다.

3.5 단계적 계층에서의 실험 결과

단계적 계층 구조에서의 테스트 결과를 살펴보면 클래스를 정상적으로 분류한 확률은 96.69%로 나타났다. 이는 총 35,785장의 이미지 중 약 34,600장이 정상적으로 1계층에서 클래스 분류에 성공한 것을 의미한다.

최종 이미지 분류 테스트 결과는 <표 6>과 같다. 무 계층과는 달리 신뢰도가 증가하여도 인식률의 감소폭이 크지 않은 것을 볼 수 있다.

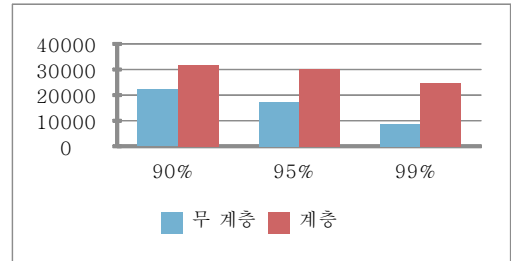
<표 6> 계층 구조 테스트 결과 요약

신뢰도	분류 성공(개)	분류 실패(개)	비율(%)
90% 이상	31,960	3,825	89.31
95% 이상	30,100	5,685	84.11
99% 이상	24,613	11,172	68.78

3.6 평가

앞선 결과에 따라 무 계층과 단계적 계층 구조에서의 이미지 인식 결과를 비교해보았을 때 Inception-v

3의 모델 자체가 훌륭하여서 기본적인 인식률이 높게 나오는 것을 볼 수 있다. 하지만 이미지 분류의 정확도에서 차이를 보인다.



(그림 4) 무 계층과 단계적 계층 구조의 인식률

(그림 4)에서 보는 것처럼 무 계층과 단계적 계층 구조의 인식률이 정확도가 높아질수록 차이가 벌어지는 모습을 보여준다. 90%의 정확도로 이미지 분류에 성공한 개수는 무 계층 구조에서는 22,272개이고 계층 구조에서는 31,960개로 약 44% 상승한 모습을 보여준다. 또한, 95%의 정확도로 이미지 분류에 성공한 개수는 각각 17,412개, 30,100개로 약 73% 더 높아졌다. 마지막으로 99%의 신뢰도에서는 무 계층 구조에서 8,800개, 계층 구조에서는 24,613개로 약 180% 상승한 결과로 3배 가까이 증가한 모습을 볼 수 있다. 이는 무 계층 구조에서 이미지를 분류하는 것보다 단계적 계층 구조를 이용하는 것이 더 높은 신뢰도로 이미지를 분류할 수 있다는 것을 의미한다. 따라서 실제 시스템에 적용하였을 때 사용자에게 더욱 높은 신뢰도로 인식 결과를 제공할 수 있기 때문에 효율적인 서비스 제공이 가능해진다.

4. 결론

기존의 전체 이미지를 한 번에 분류하는 무 계층 이미지 분류에서 상대적으로 유사한 모양을 가진 사물을 효율적으로 인식하지 못하는 문제점이 있다. 이를 해결하기 위해서 단계적 계층구조에서의 이미지 분류 체계를 제안하였고 이는 사물 이미지를 높은 인식률로 분류할 수 있는 효율적인 방안이다. 또한, 확장

성을 고려한 클래스 설계 방식을 통해 새로운 물체가 추가되었을 때도 전체 객체를 처음부터 재학습하지 않아도 된다는 점에서 효율성이 높다. 추후 사물에 대해 명확한 특징을 추출해 효율적으로 그룹화할 수 있는 방식에 관한 연구를 통해 더 높은 인식률과 확장성을 얻을 수 있을 것으로 기대한다.

참고문헌

- [1] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, Z. Wojna, "Rethinking the Inception Architecture for Computer Vision", Arxiv:1512.00567v3, pp. 2818-2826, 2015.
- [2] 김성우, "신 성장동력의 로봇개발 동향과 전망", 융합보안논문지 [KCI 등재], 제17권, 제2호, pp.153-158, 2017.
- [3] LG전자 뉴스룸 보도자료, "LG 로봇청소기, 67세 어린이 지능수준 갖춰", https://social.lge.co.kr/newsroom/lg_robotking_0717/, 2017.
- [4] 김진성, 송재열, 김하얀, 이진국, "딥러닝 기반 이미지 자동인식 기술을 활용한 사무집기 자동인식과 정보관리 시스템과의 연동방안", 한국퍼실리티매니지먼트학회논문집, 제12권, 제2호, pp. 73-80, 2017.
- [5] 정다운, 이승재, 이근동, 손형관, 고종국, 오원근, "딥러닝 기반 푸드 검색 시스템을 위한 데이터 베이스", 한국통신학회 학술대회논문집, 제2017권, 제11호, pp. 707-708, 2017.
- [6] 박미림, 용환승, "컨볼루션 신경망을 이용한 3단계 계층 구조에서의 사물인식 기법", 한국정보과학회 학술발표논문집, 제2016권, 제12호, pp.131-133, 2016.
- [7] 김보현, 용환승, "사물인식 시스템을 위한 클래스 집합의 다중 계층 사물인식구조 성능평가", 이화여자대학교대학원석사학위 청구논문, 2017.
- [8] 장현웅, 조수선, "CNN을 이용한 소셜 이미지 자동 태깅", 정보과학회논문지, 제43권, 제1호, pp.47-53, 2016.
- [9] 허지웅, 김영랑, 천현강, 이재환, "Text-CNN을 활용한 다중 감정 분류 모델 개발 및 챗봇 학습 데이터 연동 방안 연구", 한국정보과학회 학술발표논문집, 제2017권, 제12호, pp. 25-27, 2017.
- [10] 성재경, 박상민, 신상윤, 김영복, 김용국, "딥러닝 기반 쇼핑물 플랫폼용 상품 이미지 자동 분류 시스템 및 사용성 평가", 한국인터넷방송통신TV학회논문지, 제17권, 제3호, pp. 227-234, 2017.
- [11] 김태희, 강승호, "실시간 탐지를 위한 인공지능경망 기반의 네트워크 침입탐지 시스템", 융합보안논문지 [KCI 등재], 제12권, 제1호, pp.31-38, 2017.
- [12] Tensorflow tutorial, Image Recognition, https://www.tensorflow.org/tutorials/image_recognition.
- [13] DeniClassMaker; <https://github.com/pooi/DeniClassMaker>.

[저자 소개]



유 태 우 (Taewoo You)
2018년 6월 ~ 2018년 8월 삼성전자
무선사업부 SW개발 직무 인턴
2019년 2월 세종대학교
디지털콘텐츠학과 공학사 (졸업예정)
email : ldaytw@gmail.com



안 용 학 (Yonghak Ahn)
1997년 8월 경희대학교 컴퓨터공학과
공학석사
2005년 2월 경희대학교 컴퓨터공학과
공학박사
1999년 12월 한국통신정보기술 GIS
공학연구소 연구원
2006년 3월 카톨릭대학교
컴퓨터정보공학부 교수
2010년 3월 ~ 현재 세종대학교
컴퓨터공학과 교수
email : yohans@sejong.ac.kr



김 윤 욱 (Yunuk Kim)
2018년 7월 ~ 2018년 8월 에이택티엔
(T&) SW개발 직무 인턴
2020년 2월 세종대학교
디지털콘텐츠학과 공학사 (졸업예정)
email : goaldae@naver.com



정 하 민 (Hamin Jeong)
2018년 7월 ~ 2018년 8월 토마토시스
웹 개발 직무 인턴
2019년 2월 세종대학교
디지털콘텐츠학과 공학사 (졸업예정)
email : wjdgkals74@gmail.com



유 현 수 (Hyunsoo Yoo)
2018년 7월 ~ 2018년 8월 에이택티엔
(T&) SW개발 직무 인턴
2019년 8월 세종대학교
디지털콘텐츠학과 공학사 (졸업예정)
email : hs27951@gmail.com