Image Classification을 이용한 단계별 사진 정리

**요 약**

최근 데이터 기반 방법을 이용한 이미지 분류(Image classification)을 포함한 다양한 머신러닝 기능의 활용이 이루어지고 있고 Google Photo, Icloud 등 다양하게 이 기술이 적용되고 있다. 이 때 위에 나열한 플랫폼에서 해주는 이미지 분류는 대체로 단계별 classifier를 두지 않고 큰 카테고리, 예를 들어 인물 사진, 장소별, 사물별로 나뉘어 있다. 단계별로 classifier를 두어 Image classification을 했다 하더라도 사진을 분류하는 당사자 즉 사용자의 요구와는 다르게 분류를 해줄 때도 많다. 따라서 이 프로젝트에서는 사용자의 요구와 알맞게 단계별 classifier를 두어 사진을 정확하게 정리해주는 것을 목표로 한다.

**1. 서론**

**1.1 연구배경**

현재 다양한 사진들을 분류해주는 다양한 플랫폼과 어플리케이션이 존재한다. 이 때 여러가지 기술과 방법론들을 적용해서 Level별로 classifier를 나누고 사진들을 분류해주는데, 이렇게 classifier를 나누는 방식은 개발자의 결정에 달려있고 사용자가 원하는 방식으로는 사진을 분류하기가 힘들다. 따라서 사용자는 원하는 방식으로마다 폴더를 생성하고 그 안에 사진들을 분류해줘야하는 번거로움이 존재한다.

따라서 이런 모든 사용자의 요구를 충족시켜주기 위해서는 여러 개의 classifier를 만들고 그에 따라 사진들을 분류해줘야 하는데 이렇게 하게 되면 어플리케이션 내에 너무 많은 카테고리가 생성되어 UI가 복잡해질뿐더러 UX에도 더 안 좋은 영향을 끼칠 것이다. 또한 이렇게 많은 classifier를 두는 것에는 많은 시간과 비용이 들어가게 된다.

따라서 본 프로젝트에서는 사용자의 요구에 알맞게 high level로 인물 / 동물 / 풍경 / 음식 / 사물 / 기타 설정해서 각 카테고리마다 단계별로 classifier를 만들고 데이터를 학습시켜 편리하고 정확하게 Image classification을 하고자한다.

**1.2 연구목표**

이번 프로젝트의 목표는 사용자가 일일이 사진을 분류하지 않고도 사용자의 요구사항에 맞게 classifier를 두고 이미지를 학습시켜 다양한 환경과 상황에서 찍은 사진들을 편리하고 정확하게 자동으로 분류해주는 것이다.

**2. 관련 연구**

**2.1 이미지 분류 (Image Classificaion)**

**사람은 한 장의 이미지를 인식하고 분류하는데 큰 어려움이 없지만, 컴퓨터에게 있어서 이 문제는 굉장히 어려운 문제이다. 사람은 실체가 있는 물체로 인식하지만 컴퓨터에게 이 이미지는 0~255사이의 숫자로 표현되기 때문이다. 따라서 이를 해결하기 위해 여러가지 알고리즘들이 개발되어 왔고 이 알고리즘을 이용해 입력된 이미지가 어떤 클래스 라벨에 속하는지를 알려주는 것을 이미지 분류(Image classification)라고 한다.**

**2.2 데이터 기반 방법(Data-driven approach)**

이미지 분류를 위해 사람이 직접 알고리즘을 개발할 수도 있지만 데이터 기반 방법은 사람이 직접 알고리즘을 개발하는 것이 아니라 데이터를 기반으로 모델을 만들어 문제를 해결하는 것이다. 수많은 이미지와 레이블이 있는 데이터셋을 통해 모델을 학습하는데 이 때 모델을 학습하는 것을 머신 러닝(Machine Learning)이라고 한다. 이렇게 학습된 머신 러닝 모델은 새로운 이미지를 인풋으로 받아 그 이미지의 레이블을 예측한다. 위와 같은 방법은 사람이 학습하는 방법과도 비슷하기에 굉장히 효과적이다.

**2.3. Convolutional Neural Network (CNN)**

퍼셉트론은 인간의 뉴런 구조를 모방해 만든 알고리즘이지만 단순한 선형 분류기에 불과해 복잡한 문제를 풀 수 없다는 한계가 있었고, 이는 퍼셉트론을 여러 층으로 쌓아 올린 다층 퍼셉트론을 통해 해결했다. 하지만 이러한 다층 퍼셉트론도 고차원 데이터를 입력값으로 받을 때 파라미터의 수가 지나치게 많다는 한계가 있다.

따라서 이러한 다층 퍼셉트론의 한계를 보완하기 위해 Convolutional Neural Network (CNN)이 고안되었다. CNN은 인간의 시신경 구조를 모방한 Convolution과 Pooling을 반복해 특징을 추출하고, 완전연결계층(Fully connected layers)을 통해 입력된 이미지를 분류하기 위한 변별적 학습을 수행한다. CNN은 학습해야 할 전체 파라미터 수를 감소시켜 빠른 학습 속도와 우수한 일반화 능력을 가질 수 있도록 도와준다.

**3. 프로젝트 내용**

**3.1 학습 시나리오**

1. Pytorch vision model을 사용해 단계별 classifier를 만든다.
2. 데이터셋을 수집한다.
3. ImageNet, Food-101, CelebA, Places등의 데이터셋을 사용해 모델을 학습시킨다.
4. 학습시킨 모델을 이용해 Image classification을 하고 그 성능(정확도)을 측정한다.
5. 위의 정확도를 더욱 높일 수 있는 방법을 연구한다.

**3.2 요구사항**

1. Dataset은 ImageNet, Food-101, CelebA, Places 를 사용하고 추가로 더 적합한 데이터셋이 있으면 추가한다.
2. High level인 인물 / 동물 / 풍경 / 음식 / 기타 에 대한 단계별 classifier는 다음과 같다.

인물 -> 개인 / 단체

동물 -> 개 / 고양이 / 물고기 / 설치류(햄스터) / 새 / 기타

풍경 -> 실내 / 실외

음식 -> 한식 / 양식 / 중식 / 일식 / 기타

기타 -> 캡쳐화면 / 문서류 / 영수증 / 기타

**3.3 개발 환경 및 리소스**

CPU: 8 core

RAM: 16GB

GPU: Apple M1 Pro 14core GPU

사용언어: Python

딥러닝 프레임워크: Pytorch

IDE: goggle colab

**4. 향후 일정**

**4.1 향후 일정**

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| task | 4월 1주 | 4월 2주 | 4월 3주 | 4월 4주 | 5월 1주 | 5월 2주 | 5월 3주 | 5월 4주 | 5월 5주 | 6월 1주 |
| 단계별 classifier 만들기 |  |  |  |  |  |  |  |  |  |  |
| Dataset 모으기 |  |  |  |  |  |  |  |  |  |  |
| 토치비젼 모델에 학습시키기 |  |  |  |  |  |  |  |  |  |  |
| 중간보고서 |  |  |  |  |  |  |  |  |  |  |
| 사진 분류 테스트 정확도 측정 |  |  |  |  |  |  |  |  |  |  |
| 보완 및 정리 |  |  |  |  |  |  |  |  |  |  |
| 최종보고서 |  |  |  |  |  |  |  |  |  |  |
| 발표 |  |  |  |  |  |  |  |  |  |  |

**5. 결론 및 기대효과**

사용자가 직접 폴더를 생성하고 사진을 분류하지 않아도 사용자의 요구사항에 맞게 이미지를 자동으로 분류해줄 수 있도록 만들 것이다. 다양하고 많은 데이터셋을 학습시킴으로써 사진들을 정확하게 각 클래스에 맞게 분류될 것을 기대한다.

**6. 참고문헌**

[1] <https://bskyvision.com/413>

[2] <https://3months.tistory.com/512>

[3] <https://medium.com/ddiddu-log/%EC%9D%B4%EB%AF%B8%EC%A7%80-%EC%9D%B8%EC%8B%9D%EC%9D%98-%EC%A0%95%EC%9D%98%EC%99%80-%EC%A3%BC%EC%9A%94-%EB%AA%A8%EB%8D%B8-%EB%B9%84%EA%B5%90-1-%EC%9D%B4%EB%AF%B8%EC%A7%80-%EB%B6%84%EB%A5%98-image-classification-ae7a59bfaf65>