

MobileNet-V2를 이용한 스시 오마카세 코스 메뉴 자동 분류 시스템 (멀티미디어 특론 프로젝트 발표)

2020. 06. 26(금)

서강대학교 자연어처리연구실

320190046 허 윤 석

Introduction

● Motivation

- 오늘 날 젊은 세대의 음식 문화
 - 본인이 먹은 음식을 SNS에 포스팅 하거나 유튜브에 리뷰 영상을 제작
 - 최근에는 하이엔드 음식에 대한 관심도가 급증
- 스시 오마카세
 - 과거에는 높은 가격대로 인해 접하기가 어려웠음
 - 최근에는 다양한 가격대의 스시야가 확산됨에 따라 젊은 층도 쉽게 접할 수 있는 새로운 음식 문화로 자리매김
 - 스시에 대한 충분한 지식이 없는 젊은 세대는 여전히 높은 진입장벽
 - 일본어가 많고 메뉴 간 비슷하게 특성을 가지고 있음



청어(니싱)



흑점줄 전갱이(시마아지)

Introduction

● Goal

- 오마카세 코스 메뉴에 특화된 이미지 분류 모델을 개발하여 젊은 세대들에게 스시에 대한 정보를 제공

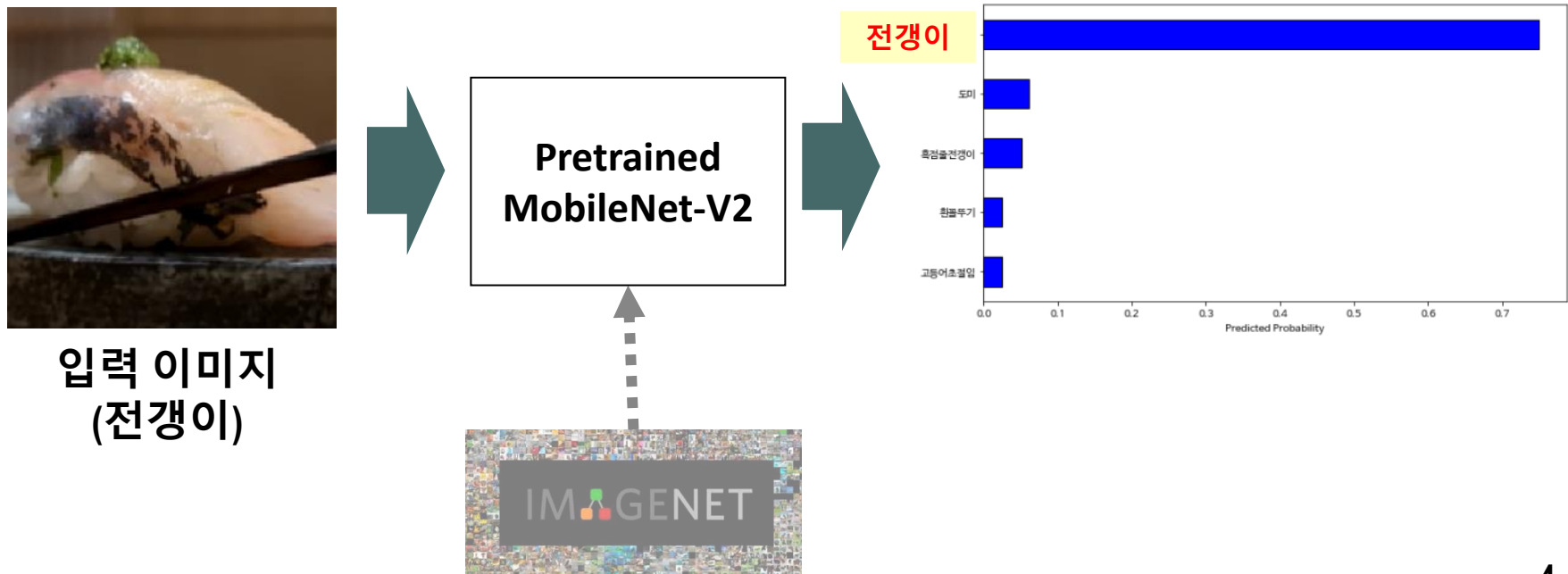
● Contribution

- 스시 오마카세 코스 메뉴에 특화된 Convolutional Neural Network 기반의 이미지 분류 모델을 제안
 - ImageNet과 같은 대용량 이미지들로 사전학습된 CNN 기반 네트워크에 직접 수집한 스시 오마카세 데이터로 Fine-tuning한 시스템을 제안
 - 본 시스템은 Mobile 환경과 같이 컴퓨팅 성능이 제한된 상황에서도 충분한 퍼포먼스를 보장할 수 있는 MobileNet-V2를 활용

Proposed Model

● 전체 구조

- ImageNet으로 사전학습 된 MobileNet-V2를 직접 수집한 오마카세 데이터로 Fine-tuning
- Fine-tuned 네트워크는 오마카세 메뉴 사진에 대하여 모든 메뉴 이름에 대한 생성확률을 생성
 - 이미지에 대한 최종 분류는 가장 높은 생성확률을 갖는 메뉴로 결정



Experiment

- **Dataset**
 - 메뉴 개수: 25개
 - 메뉴 이름은 3개 국어(한국어/영어/일본어) 모두 태깅
 - 총 이미지 수: 292장(70% / 10% / 20%)
- **Evaluation Metric: Accuracy**
- **Preprocessing**
 - Resizing
 - 이미지 중심을 기점으로 256 * 256으로 이미지 크기 통일
 - Data Augmentation
 - Rotation
 - Color Jittering

Experiment

- **Baseline : VGGNet-16**

- **Result**

- MobileNet-V2가 Baseline보다 더 경량화 된 모델이지만 더 높은 Accuracy를 보이고 있음
 - Parameter 수 :
 - MobileNet: 32K
 - VGGNet-16: > 1M
- MobileNet-V2의 파라미터 수가 VGGNet-16에 비해 약 20% 정도이기 때문에 제안 모델의 학습 시간이 훨씬 짧음

	Top-1 Accuracy	Top-5 Accuracy	Training Time
Baseline	62.39%	88.78%	15mins
MobileNet-V2	64.29%	91.84%	< 2mins

Discussion

- 제안 모델의 문제점
 - Fine-tuning에 사용한 데이터 수의 부족
 - 이미지 분류에 적합한 좋은 퀄리티의 이미지를 구하는 것이 어려움
 - 메뉴 별로 데이터 수의 차이가 존재 ➔ 성능 저하
 - 예시: 박고지 김밥
 - 데이터 수: 5개
 - 스시와는 달리 동그란 특성을 가지고 있지만 데이터 수가 부족하여 전혀 판별해 내지 못함.



Conclusion

- 모바일 환경에서 강건한 성능을 보이는 MobileNet-V2를 이용하여 스시 오마카세 메뉴를 자동으로 분류하는 시스템
- 향후 연구
 - 개인 사용자들이 실시간으로 찍는 스시 사진을 incremental learning 등의 방식을 이용하여 학습하여 실제로 Real-time으로 이미지를 분류할 수 있는 시스템을 개발