|  |
| --- |
| **음식 분류에 대한 베이스라인 모델 개발**  **2015104182 컴퓨터공학과 송영진**  **요 약**  최근 정부에서 공개한 170종의 인공지능 학습용 데이터셋 중 다양한 음식 사진을 활용하여, 이를 기반으로 음식의 종류를 구분할 수 있는 AI베이스라인 모델을 개발한다. |

**1. 서론**

**1.1. 개요**

최근 정부에서 170종의 인공지능 학습용 데이터셋을 공개하였다. 이 데이터셋에 대한 베이스라인 모델을 개발하는 것이 목적이다. 170종의 다양한 데이터셋 중 음식 이미지에 관한 데이터셋을 선정하여 이에 대한 반복학습을 통해 정확도를 확인하고 음식 이미지를 분별할 수 있는 베이스라인 모델을 개발하여 커뮤니티에 공개한다. 개발된 베이스라인 모델을 기반으로 많은 사람들이 이 데이터셋을 이용해 새로운 모델을 개발할 수 있도록 하는 것이 목적이다.

**1.2. 배경 지식**

케라스(Keras)에서 모델을 학습시킬 때에는 fit()함수를 사용한다. 이 fit()함수는 인자에 따라 학습과정 및 결과가 차이가 난다.



이와 같은 인자들을 가지게 되는데, 쉽게 말해 x는 문제이고, y는 문제에 대한 답이다. 그리고 batch\_size(배치사이즈)는 몇 개의 문제를 푸는지에 대한 값이다. 예를 들어, 100개의 문제가 있을 때, batch\_size가 100이면 전체를 다 풀고 난 뒤에 답을 맞춰보는 것이고, 이러한 경우에는 가중치 갱신이 한 번만 일어난다. 만약, 배치사이즈가 10이면 100개의 문제를 10문제씩 나눠서 10번 답을 맞춰보기 때문에 가중치 갱신이 10번 일어나게 된다.

Epoch는 반복 학습의 횟수를 의미한다. 같은 데이터셋으로 반복적으로 가중치를 갱신하면서 모델이 학습하기 때문에 반복을 하게 되면 틀리는 횟수가 적어지게 된다. 무조건적으로 많은 Epoch가 좋은 것은 아니다. 왜냐하면 과한 반복학습을 하면 Overfitting(과적합)이 일어나기 때문에 오버피팅이 일어나면 학습을 중단해야 한다.

**2. 관련연구**

**2.1. 딥러닝**

머신 러닝의 한 분야로 데이터를 컴퓨터가 처리 가능한 형태인 벡터나 그래프 등으로 표현하고 이를 학습하는 모델을 구축하는 연구를 포함한다. 얼굴이나 표정을 인식하는 등의 특정 학습 목표에 대해, 딥 러닝은 학습을 위한 더 나은 표현 방법과 효율적인 모델 구축에 초점을 맞춘다. 딥 러닝의 표현방법들 중 다수는 신경과학에서 영감을 얻었으며, 신경 시스템의 정보 처리나 통신 패턴에 기반을 두고 있다.

**2.2 인공지능**

인간의 경험과 지식을 바탕으로 문제를 해결하는 능력, 시각 및 음성 인식의 지각 능력, 자연 언어 이해 능력, 자율적으로 움직이는 능력 등을 컴퓨터나 전자 기술로 실현하는 것을 목적으로 하는 기술 영역. 인공 지능의 궁극적인 목표는 사람처럼 생각하고 행동까지 할 수 있는 기계를 개발하는데 있다.

**3. 프로젝트 내용**

**3.1. 음식 사진 데이터셋**

기계학습 중 컴퓨터 비전 분야의 기술 개발에 활용할 수 있는 학습 데이터셋으로 다음과 같이 구성하였다.

1. 서로 다른 인도 음식 20종에 대해 각 100개의 샘플 이미지 ( 총 2,000개의 이미지 )

컴퓨터 비전은 사진의 물체 인식부터 자율 주행에 이르기까지 다양한 목적으로 연구 개발되고 있는 기술이다. 본 데이터셋은 컴퓨터 비전 중 물체 인식과 이미지 분류를 응용한 음식 식별을 위해 구성되었다.

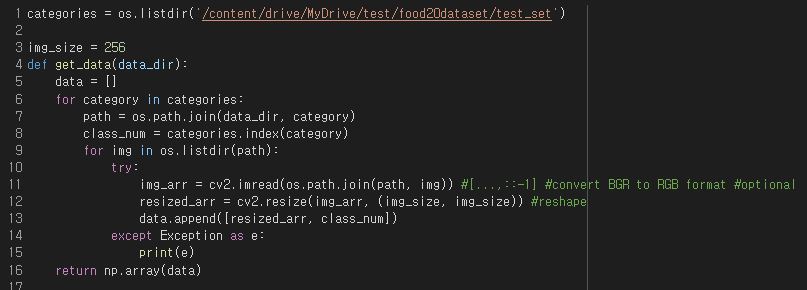
**3.2. 데이터셋의 구성**

본 데이터셋은 20개의 서로 다른 인도 음식의 이미지가 포함되어 있으며, 각 음식 항목에 대한 100개의 샘플 이미지가 있다. 데이터는 train-test(train-validation) 분할 형식으로 저장되어 있다. 열차 테스트 비율은 70:30이며, 이미지의 해상도는 200 \* 500 ~ 5760 \* 3840 픽셀 범위에서 다양하다.

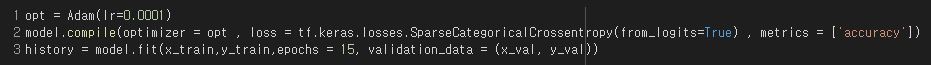
**3.3. 데이터셋의 설계 기준과 분포**

데이터셋을 설계할 때 가장 중요하게 고려했던 점은 음식 사진의 수집 가능성과 분포, 그리고 사진품질이다. 불특정 다수의 클라우드 워커를 활용하여 음식을 수집하는 만큼 음식을 접하는데 어려움이 없어야 다양한 사람들로부터 골고루 데이터가 수집되어 사진 데이터셋에 편향성을 최소화할 수 있을 것으로 예상되었기 때문이다. 사진 품질의 경우에는 명확한 기준을 제시하고, 수집된 사진에 대한 검수를 여러 단계에 걸쳐서 진행함으로써 낮은 품질의 사진을 최대한 걸러낼 수 있도록 했다.

**3.4. 프로젝트 세부 내용**

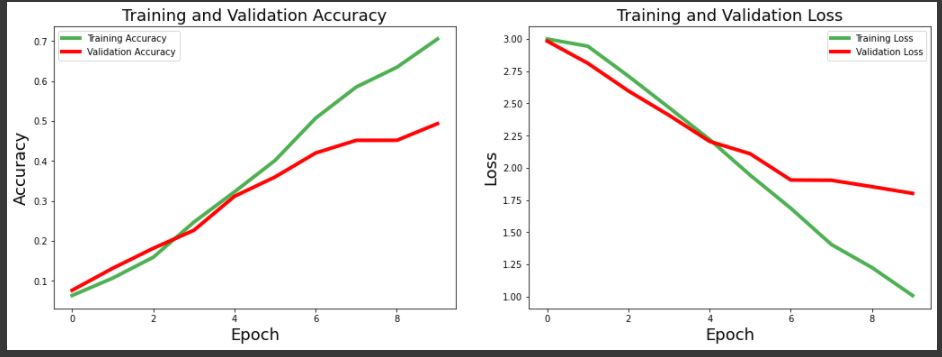
****

- 모든 이미지를 업로드 한 뒤, 각 이미지의 모양을 256 \* 256으로 변경하고 픽셀 값의 정규화를 수행한다.

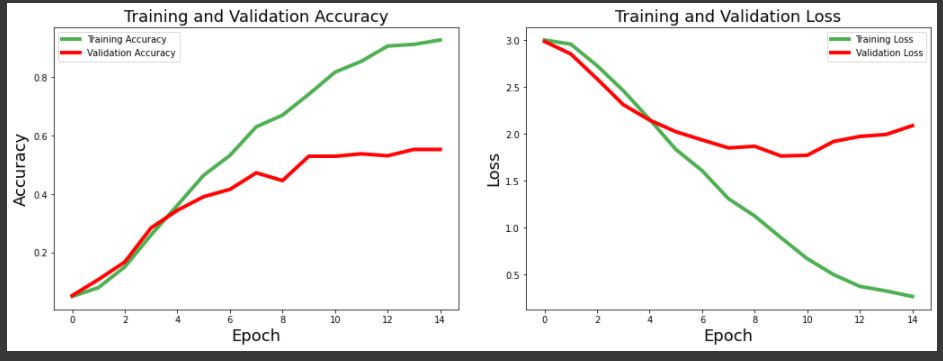


- epoch값을 설정하여 Overfitting(과적합)이 일어나는 지점을 찾아야한다.

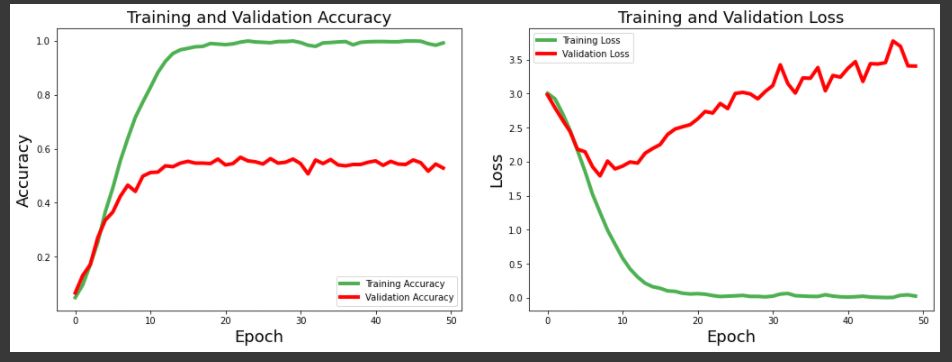
1. **epoch 10회 결과**

****

1. **epoch 15회 결과**



1. **epoch 50회 결과**



반복학습을 진행해보았을 때, 10회 지점에서 Overfitting(과적합)이 일어나는 것을 볼 수 있다.

과적합이 일어난 이후에 추가로 학습을 시키게 되면 시험 데이터에 대해서는 오차가 감소하지만 실제 데이터에 대해서는 오차가 증가하게 된다. 시험 데이터에 과도하게 최적화가 이루어지기 때문이다. 따라서 epoch값에 변동을 주어 과적합이 일어나는 지점을 찾은 뒤, 학습을 멈춰 베이스모델을 개발하여 사용하면 된다.

**4. 결론 및 기대효과**

시험 데이터셋을 통해 epoch값을 변화를 주어 모델을 학습시켜 Overfitting이 일어나는 지점을 찾을 수 있다. 이를 통해 과적합이 일어나기 전까지의 epoch값으로 베이스라인 모델을 개발할 수 있다. 이 베이스라인 모델을 커뮤니티에 공개하여 사람들로 하여금 처음 학습시킨 시험 데이터셋이 아닌 새로운 데이터셋을 가지고 학습을 진행하여 이 베이스라인 모델의 정확도를 확인할 수 있도록 한다.

**5. 참고문헌**

- Kaggle 플랫폼 : 데이터 세트에 존재하는 Food20의 인도 음식 20가지