

딥러닝을 이용한 암세포 식별 알고리즘 개발

과제번호 | 기업-19-04

● 팀 명 : C4

● 참여학생 : 박진형, 서동혁, 이연진

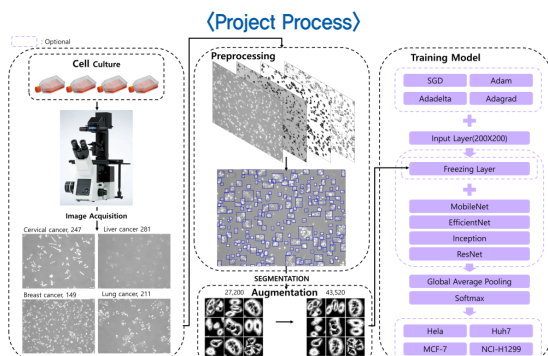
● 학부(과)명 : 메디컬IT융합공학과

● 지도교수 : 최세운

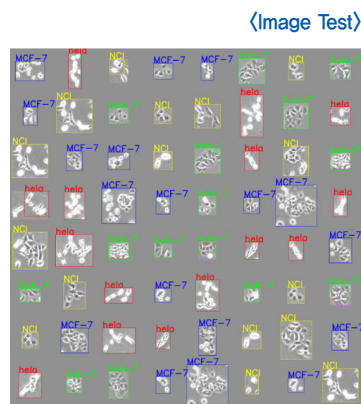
작품개요

- **목적** : 현재 암의 확진과 진행상태의 결정은 여러 가지 검사들을 종합하여 진단하게 된다. 의사의 진찰, 조직검사, 세포 검사, 내시경검사, 암 표지자 검사, 영상진단검사, 핵의학 검사 등이 있다. 하나의 검사로 암이 확진될 수 없으므로 암의 진단은 여러 검사를 복합적으로 실시하여 종합적으로 의사에 의해 신중히 판단된다. 다양한 검사를 하므로 검사에도 많은 시간적 소요와 여러 전문가가 필요하다. 게다가 이러한 생체 검사는 사람이 직접 어디서부터 발병된 암세포인지 현미경으로 확인하여 전이 여부를 판단하기 때문에 정확도가 떨어질 위험이 있다. 시간 및 인적 비용과 정확도 문제를 해결하기 위해 딥러닝을 이용하여 암세포를 식별하는 알고리즘을 개발하려 한다.
- **작품설명** : 총 네 가지의 암세포를 학습하여 분류하는 알고리즘을 개발하였다. 각각의 세포는 같은 환경에서 같은 배지 용액을 사용하여 3~4일 배양하였다. 그리고 적절히 자란 세포들을 대상으로 역상 광학현미경을 이용하여 자료를 수집하였다. OpenCV를 사용하여 획득한 세포 광학 영상 데이터에 적절한 Image Processing을 적용하였고, 세포의 윤곽선을 검출하였다. 이를 통해 별도의 라벨링 필요 없이 학습할 수 있도록 만들었다. 모델 학습은 Keras로 진행되었으며 ImageNet을 통해 Pretrained된 모델을 사용하였다. 여러 Optimizer와 Model을 비교하여 적절한 알고리즘을 개발하였다. 그리고 이러한 식별 알고리즘을 Django를 통한 웹 서비스화하여 어디서든지 바로바로 이미지를 업로드 하여 암세포를 식별 할 수 있도록 만들었다.
- **기대효과** : 해당 서비스를 진행함으로써 암세포 광학 영상을 획득하는 동시에 업로드 하면 해당 이미지에 존재하는 세포의 종류와 개수를 알 수 있어 암을 진단하는 데 있어 소모되는 인적, 시간적 자원을 줄일 수 있을 것이라 기대한다.

작품사진



프로젝트 전체의 실험 Process이다. 세포를 수집하고, 이미지에 대해 전처리 후 학습하는 방식으로 프로젝트가 진행되었다.

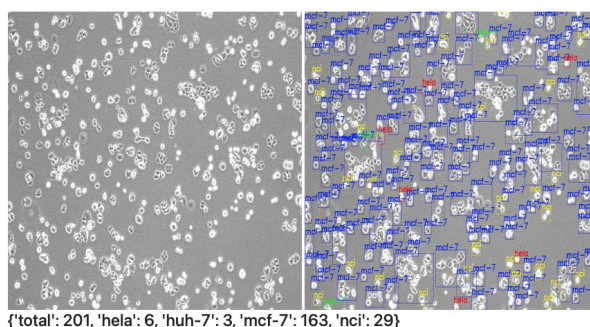


단일 Image에서의 model Inference 성능 확인을 위해 학습에 사용되지 않은 각각의 세포들을 사용하여 한 이미지에 무작위로 후 테스트를 진행하였다.

Confusion matrix

	hela	huh7	mcf7	nci
True labels	15	0	0	0
hela	0	15	1	0
huh7	0	0	17	0
mcf7	0	0	0	16
nci	0	0	0	16
Predicted labels	hela	huh7	mcf7	nci

Testing on Web



최종 웹에서 확인할 수 있는 결과로 이미지 업로드 시에 세포의 총개수와 각각의 개수, 원본 이미지 그리고 Model이 Inference 한 이미지를 같이 볼 수 있도록 하였다.