

유방 조직 이미지를 이용한 유방암 예측 모델

21800680 정하은

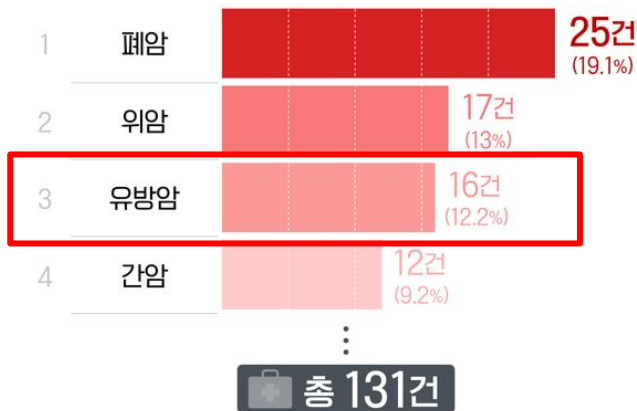
22000176 김지민

22100231 김현서

주제 선정 계기와 기대효과

암 종류별 오진 피해 사례

2017년 ~ 2021년 6월



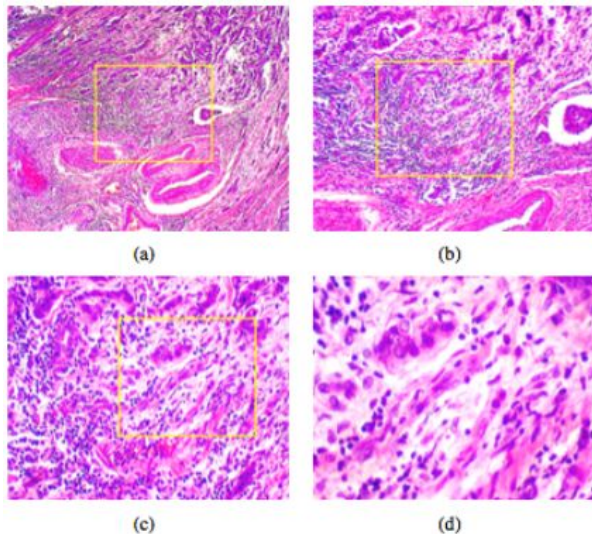
- 유방암은 우리나라 여성에게 가장 흔하게 발생하는 암이다.
- 2017년~2021년 암 종류별 오진율에서 유방암이 3위를 차지했다.

악성을 **양성**으로 오진할 경우 → 불필요한 조직검사로 인한 합병증 발생 가능성, 비용, 시간적 문제

양성을 **악성**으로 오진할 경우 → 수술가능시기 놓쳐서 생존에 심각한 문제가 될 가능성이 높아짐

- 기존의 검사보다 빠르고 정확하게 유방암 여부를 진단할 수 있는 딥러닝 기반 틀은 환자의 올바른 치료계획을 선택하고 생존율을 높이는 데에 도움이 될 것이라고 생각함.

데이터 출처와 정보



	fold	mag	grp	path	label	label_int	filename
0	1	100	train	BreaKHis_v1/histology_slides/breast/benign/SOB...	benign	0	SOB_B_A-14-22549AB-100-001.png
1	1	100	train	BreaKHis_v1/histology_slides/breast/benign/SOB...	benign	0	SOB_B_A-14-22549AB-100-002.png
2	1	100	train	BreaKHis_v1/histology_slides/breast/benign/SOB...	benign	0	SOB_B_A-14-22549AB-100-003.png
3	1	100	train	BreaKHis_v1/histology_slides/breast/benign/SOB...	benign	0	SOB_B_A-14-22549AB-100-004.png
4	1	100	train	BreaKHis_v1/histology_slides/breast/benign/SOB...	benign	0	SOB_B_A-14-22549AB-100-005.png
...
39540	5	400	test	BreaKHis_v1/histology_slides/breast/malignant/...	malignant	1	SOB_M_PC-15-190EF-400-011.png
39541	5	400	test	BreaKHis_v1/histology_slides/breast/malignant/...	malignant	1	SOB_M_PC-15-190EF-400-012.png
39542	5	400	test	BreaKHis_v1/histology_slides/breast/malignant/...	malignant	1	SOB_M_PC-15-190EF-400-013.png
39543	5	400	test	BreaKHis_v1/histology_slides/breast/malignant/...	malignant	1	SOB_M_PC-15-190EF-400-014.png
39544	5	400	test	BreaKHis_v1/histology_slides/breast/malignant/...	malignant	1	SOB_M_PC-15-190EF-400-015.png

39545 rows × 7 columns

- BreakHis는 82명의 환자로부터 각기 다른 확대 사진 (40X, 100X, 200X, 400X)를 사용하여 수집한 91097개의 유방 종양 조직 현미경 이미지로 구성됨
- 이미지는 2014년 1월부터 12월까지 임상 연구를 통해 수집됨
- 이 데이터 베이스는 브라질 파라나의 P&D 연구소 (병리해부학 및 세포병리학과)의 협력을 통해 수집됨

구현 알고리즘

전이학습

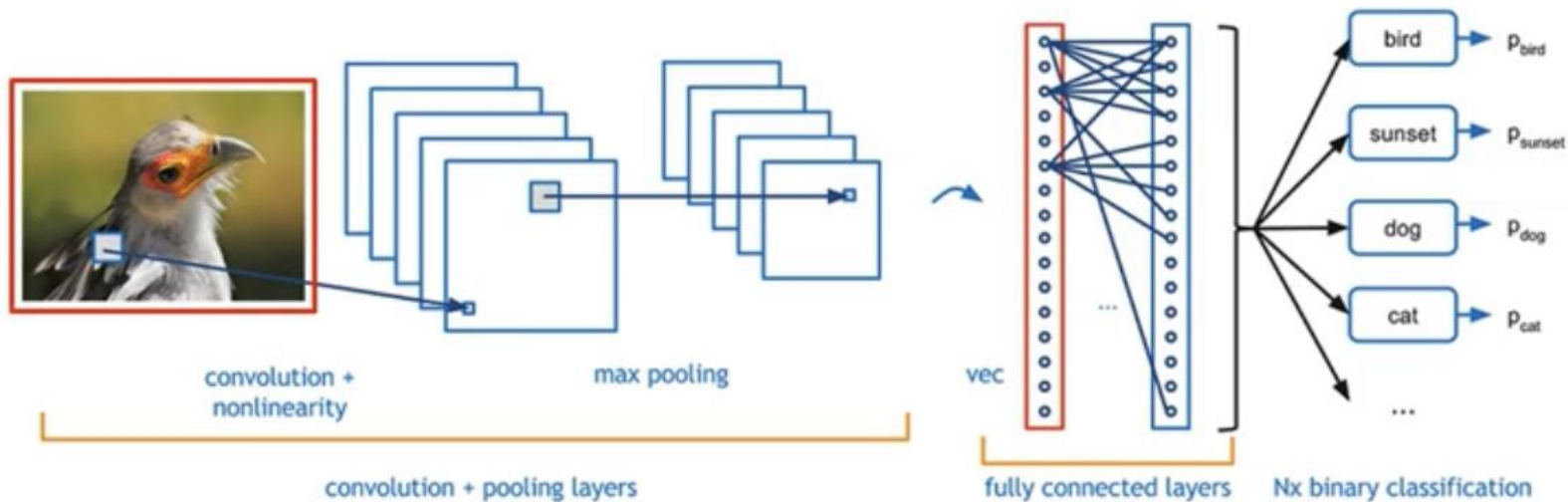
- 14만개가 넘는 이미지로 구성되어 있는 ImageNet 데이터를 사용해서 학습된 모델의 가중치를 가져와 해결하고자 하는 문제에 적용하여 새로운 모델을 만드는 방식

장점

- 데이터 부족 문제 해결. 비교적 적은 데이터로 높은 성능 결과
- 모델을 만들 때 layer 수, activation, hyperparameter 를 고려하는 데 걸리는 시간이 줄어듦

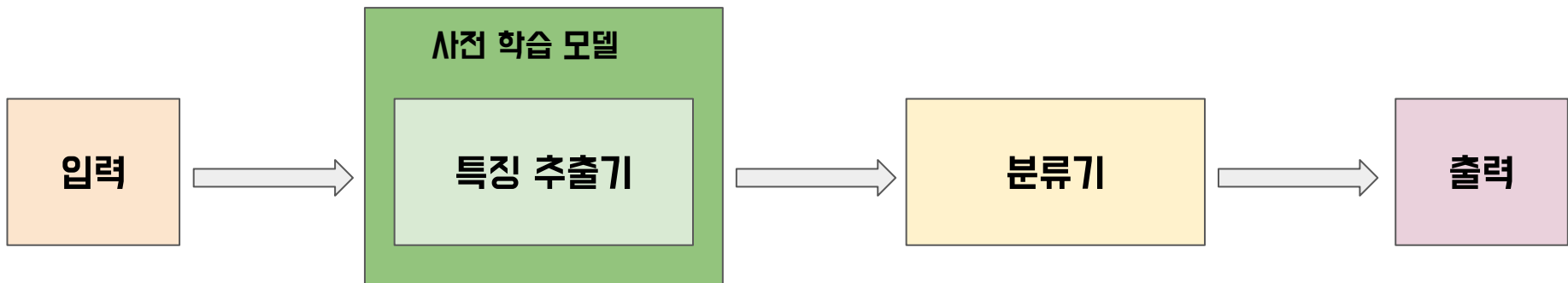
구현 알고리즘

CNN



구현 알고리즘

사전 학습 모델 구조



Efficientnetv2-b0

- 학습 효율성을 증가시키기 위해 만들어진 모델
- 비교적 적은 파라미터 수로 높은 성능을 낼 수 있다.