의료 데이터 이미지 전처리

의료용 데이터를 수집하는 목적 크게 3가지

분류 : 폐종약 약/악 성 진단 ,폐결핵 진단 등

검출 : 유방 종괴 검출 , 대장 용종 검출 등

분할 : 병변의 면적 , 형태 측정

데이터 augment

Flip, rotation, translation, scale, brightness, sharpening 등

* 불균형 해소해 줄 순 있으나 대표성이 낮고 모델의 일관성 떨어뜨림

GAN : 학습시킨 모델을 기반으로 진짜 데이터와 구분하기 힘든 가짜데이터를 생성해냄

* 가짜 데이터를 사용한 것 가설로 모델을 학습 시킨 것으로 실제로 적용하기에 무리가 있다고 함

의료 데이터에 사용되는 전처리 기법

비트변환 ,노이즈 제거, 정규화, 일괄적 영상 크기 조정

비트변환 : 기존 이미지는 대부분 8(0~255)bit x-ray, CT, MRI 같은경우 12bit(0~ 4095)

처리방법 연구, 보통 8bit 비트 변환을 통해 전처리

논문 찾아본 결과 8bit 16bit 차이 크게 없었음 비트에 알맞은 정규화만 진행해 주면 될 듯

Channel 축소 : 3channel -> 1channel로 변경

노이즈 제거 : 대비를 줄이는 (median gaussian) filter는 주변 구조물과 대비를 줄여 경계를 모호하게함 보통 선명도를 높이는 ( sharpening filter, 히스토그램 평활화 사용, 대비제한 적응 히스토그램 평활화 사용)

텍스트, 고양이이(가) 표시된 사진

자동 생성된 설명

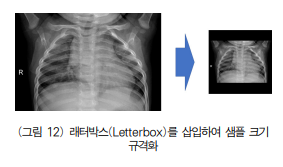
정규화 : 8bit ~ 12bit pixel 값을 - > 0 -1 이나 -1 – 1 사이로 줄임  
 <정규화 방법>

* (X – 최소값) / (최대값 – 최소값)
* 표준화 (평균과 표준편차를 이용)
* GCN(Global Contrast Normalization) 전체 이미지 픽셀 값의 평균 표준편차를 구한 후 표준화
* LCN(Local Contrast Normalization) 장마다 이미지 픽셀 값의 평균 표준편차를 구한 후 표준화
* MSCN(mean subtraction and contrast normalization) 장의 구역별 평균 표준편차를 구한 후 각 부분을 표준화

영상 크기 조정 : 영상의 가로 세로 비율은 유지한채 resize 진행하는 것이 좋음, 만약 데이터들의 크기가 상이하다면, 짧은 변을 기준으로 맞춘 후 긴변을 자르거나 , 긴변을 기준으로 맞춘 후 짧은 변을 자름 보통 긴변을 맞춘 후 제로패딩 진행  
또는 letter box를 사용하여 가로 세로 비율 맞춰준 후 resize 진행 하는 방법

보간 방법 또한 고려

* Bicubic
* Bilinear
* 영역 보간법



영상 해상도를 높이는 방법

* 히스토그램 평활화
* Srcnn 모델 사용

모델 구현시 고려

전이학습

하이퍼 파라미터 수정

Model

다양한 model 존재

Srcnn 거친 후 ouput으로 model 학습

잔류 밀집 블록(Residual Dense Block, RDB) 거친 후 ouput 으로 model 학습

생성적 적대 신경망 거친 후 output으로 model 학습

의문

Gray scale 로 할지 rgb로 할지

참고 문헌>

[:: JKSR :: Journal of the Korean Society of Radiology (jksronline.org)](https://jksronline.org/DOIx.php?id=10.3348/jksr.2020.0171)

[잔류 밀집 네트워크 기반 영상 해상도 향상 전처리 기법을 활용한 X-ray 폐렴 진단 연구 (kci.go.kr)](https://www.kci.go.kr/kciportal/ci/sereArticleSearch/ciSereArtiView.kci?sereArticleSearchBean.artiId=ART002643403)

[02본문 (kibme.org)](http://www.kibme.org/resources/journal/20200504094149078.pdf)

[42(4)-04(018김민정).fm (koreascience.or.kr)](https://www.koreascience.or.kr/article/JAKO202125761123569.pdf) (데이터 비트수가 모델학습에 끼치는 영향)

[42(3)-07(010김광기).fm](http://203.250.217.22/article/JAKO202121061501444.pdf) (IS ,HE , CLASHE 이미지 전처리시 정확도 차이)