

[RT704] Advanced Medical Image Processing

< REPORT >

(2021-12-11)

Robotics Engineering

202123008 Jinmin Kim

Phone: 010-6266-6099

Mail: rlawlsals@dgist.ac.kr



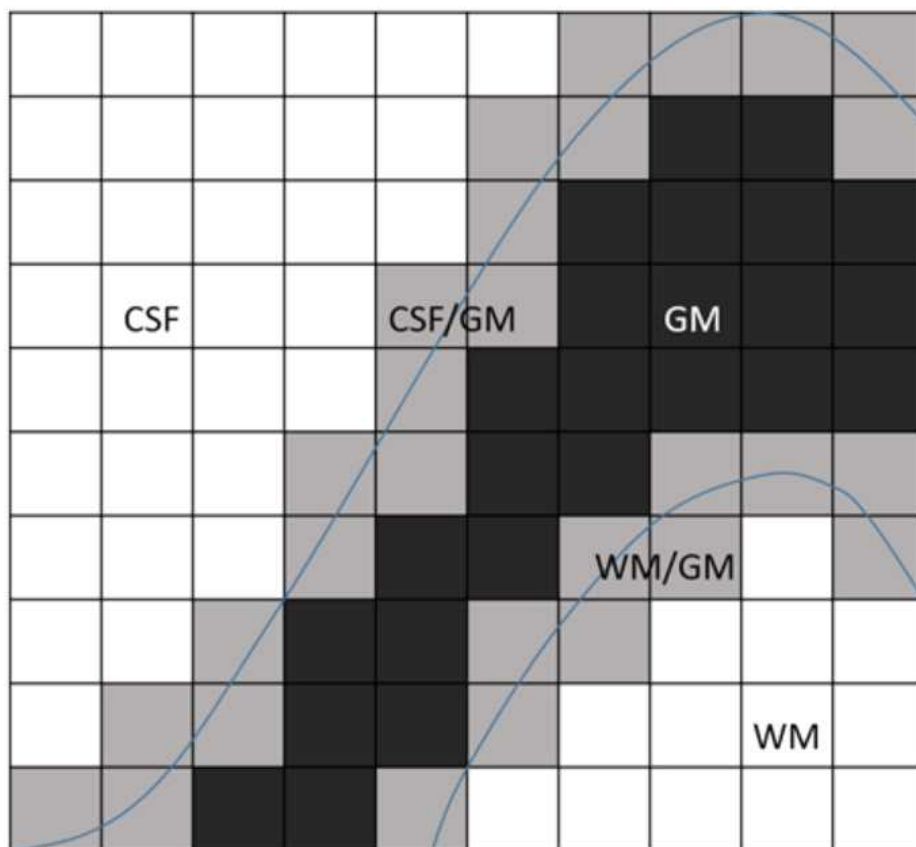
1. What is the partial volume effect and how does it affect the image?

Partial Volume Effect(PVE, 부분용적효과)는 한 복셀에 여러 tissue가 포함될 수 있는 확률을 말한다.

만약 $1\text{mm} \times 1\text{mm} \times 1\text{mm}$ 의 해상도를 가지는 영상이 있을 때, 1mm 크기의 복셀로는 정교하게 표현할 수 없는 상황이 발생하는 것이다.

아래 그림은 뇌 MR영상의 예시이다.

파란선의 경계에 따라 tissue가 구분된다고 하면, 회백질(GM)과 뇌척수액(CSF)간 경계를 따르는 복셀에 두 tissue가 모두 포함될 수 있다는 것이다.



PVE는 이미지 구별에 어려움을 야기하고 정확한 진단을 위해서는 PVE에 대한 보정이 고려되어야 한다. PVE를 제거하기 위한 방법으로는 Geometric transfer matrix (GTM), The fitting method, Reconstruction-based method, Image-based method, Mask-based method 등의 여러 가지 방법들을 사용할 수 있다.

References

[1]<https://beelinekim.tistory.com/38>

[2]Park, Minju and Lee, Sangbock, "Implementation of Filter for the Removal of Partial Volume Effect," Journal of the Korean Society of Radiology, vol. 9, no. 3, pp. 139-145, Apr. 2015.

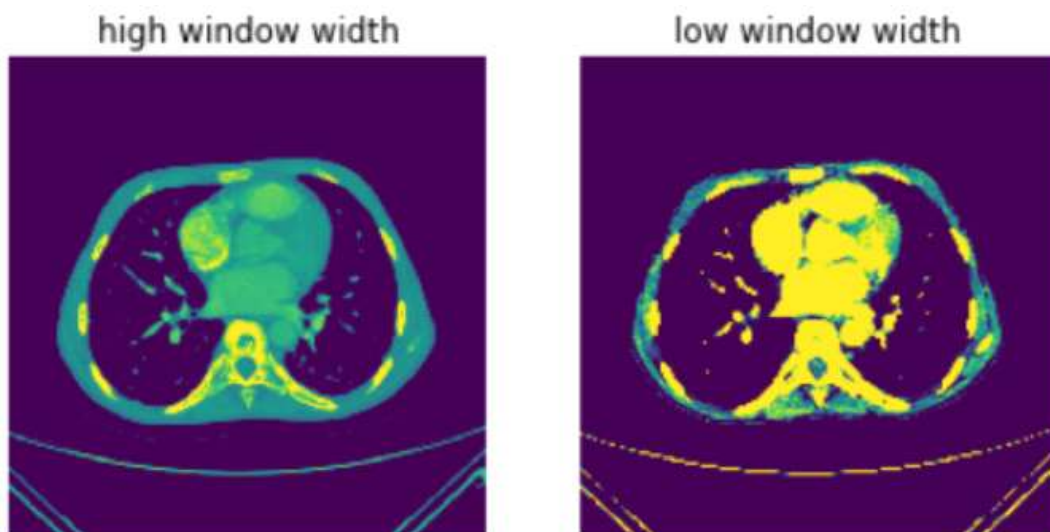
2. What is meant by window and level in displaying CT images? Why is it needed?

CT영상에서 Window width란 gray scale로 표현할 수 있는 범위를 의미하며, Window level은 gray scale의 중앙값을 의미한다.

window width는 영상의 contrast에 직접적인 영향을 미친다.

width를 높게 설정하면 표시되는 gray scale이 커지며 영상의 contrast가 나빠지고 Density scale이 커져서 영상의 분해능이 저하된다.

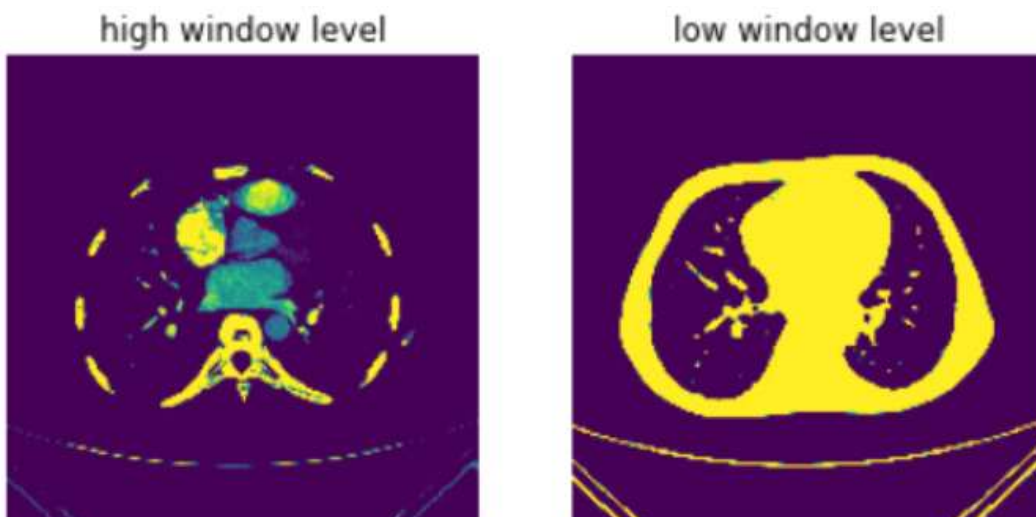
width를 좁게 설정하면 표시되는 gray scale 범위가 작아져 영상의 대조도가 좋아지고 Density scale이 작아져서 분해능이 좋아진다.

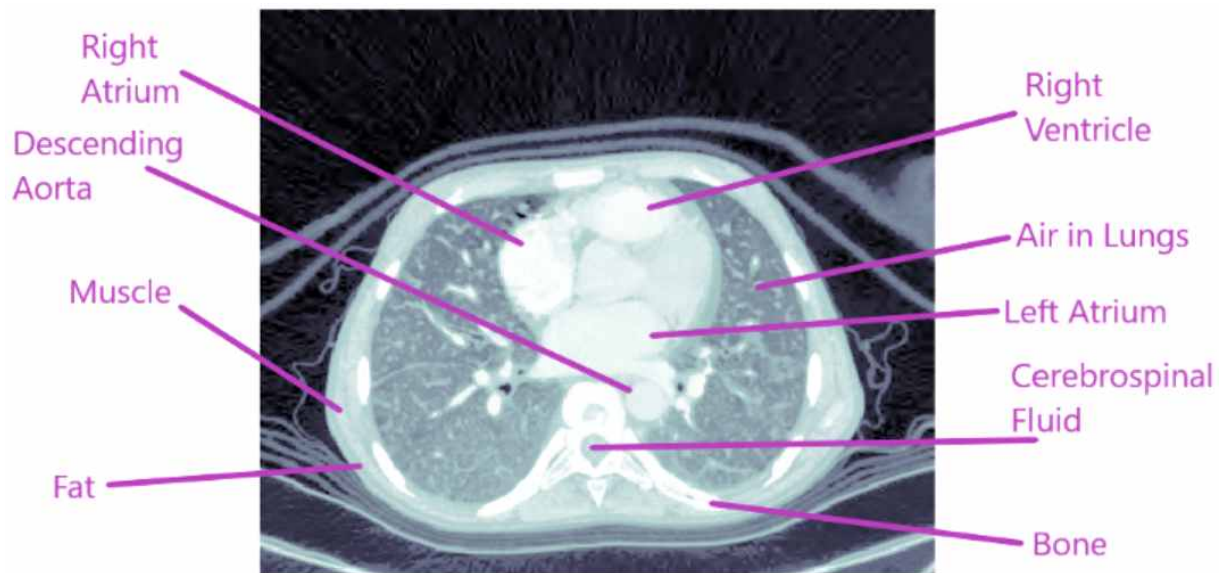


window level은 선감약계수, 흡수치, 농도와 직접적인 관련이 있다.

level을 최대한 낮추면 공기와 같이 흡수치가 적은 물질을 잘 나타낼 수 있다.

level을 최대한 높이면 뼈와 같이 흡수치가 큰 물질을 잘 나타낼 수 있다.





위 그림에서 보다시피, 의료 영상에는 다양한 흡수치를 가진 물질들이 있다.

정규화된 이미지는 광범위한 조직 밀도를 표시하는데, 일반 컴퓨터 화면은 256개의 gray scale만 표시할 수 있고 우리 눈은 gray scale의 약 6% 변화만 구분할 수 있다고 한다. 따라서 인간이 밀도차이를 감지하려면 정상 조직과 병리적으로 변경된 조직의 차이를 통해서 보아야 한다. 이것이 CT에서 window와 level을 사용하는 이유이다.

References

- [1]<https://blog.naver.com/PostView.nhn?isHttpsRedirect=true&blogId=simhs2017&logNo=221624960845&redirect=Dlog&widgetTypeCall=true&directAccess=false>
- [2]<https://ichi.pro/ko/geulei-seukeil-ui-munje-dicom-windows-ihae-29979749225149>
- [3]<https://www.youtube.com/watch?v=KZld-5W99cI>

3. What are the advantages and disadvantage of MRI opposed to X-ray CT?

MRI(자기공명영상)란 자기장을 발생하는 커다란 자석통 속에 인체를 들어가게 한 후, 인체에 고주파를 쏘아 인체에서 메아리 같은 신호가 발산되면 이를 되받아서 디지털 정보로 변환하여 영상화하는 장비이다.

- MRI의 장점

1. CT와 달리 방사선 노출이 없다.
2. CT에 비해 체내 연부조직간 대조도가 뛰어나다. (근육, 인대, 뇌, 종양 등)
3. 자세 변화 없이 3차원 입체 영상화가 가능하다. (CT는 한 단면만 가능)

- MRI의 단점

1. 20분~1시간 이상으로 촬영 시간이 길다. (CT는 5~10분)
2. 가격이 비싸다.
3. 폐쇄공포증이 있는 환자에게는 시행하기가 어렵다.
4. 금속성 인공치아나 척추 보형물 등이 진단에 방해가 될 수 있다.
5. 보청기나 구형 심박동기 등 작동에 장애를 초래할 수 있다.

4. What is the purpose of a PACS?

PACS(Picture Archiving and Communication System)는 의학용 영상 정보의 저장, 판독 및 검색 기능 등의 수행을 통합적으로 처리하는 시스템을 말한다.

PACS는 X-ray, CT, MRI, PET, SPECT 등에 의해 촬영된 모든 검사 결과를 디지털 이미지로 변환해주고, 촬영과 동시에 대용량 기억장치에 저장시켜 판독 전문의가 판독할 수 있도록 해준다.

- PACS의 특징

- DICOM 규격에 따라 이미지 데이터를 저장, 관리
- 의료영상획득장비와 진단방사선과와 임상 의사들을 하나로 연결
- 관계형 DB를 이용하여 의료영상을 저장하거나 요청에 따라 검색하여 전송
- 판독실이나 외래의사 및 진단방사선과 의사들은 GUI로 개발된 viewer를 제공하여 자신의 연구실에서 저장된 의료영상데이터를 전송받아 판독하고 저장할 수 있음.

- PACS의 목적

PACS를 통해 환자의 영상 판독을 필요로하는 임상 관련과로 직접 전송할 수 있기때문에, 환자에대한 결과를 신속하게 볼 수 있다. PACS가 없는 경우에 많은 시간에 걸쳐 전달되던 데이터가 촬영되는 즉시 수 초 내에 전달될 수 있어서, 환자에 대한 신속한 진료와 함께 병원 내에서의 정보흐름의 지연을 해소하여 환자의 진료시간 및 입원기간을 단축하는 효과를 얻을 수 있다. 그리고 PACS는 필름을 사용하지 않기 때문에 필름의 분실로 인한 데이터의 분실을 방지할 수 있다. 필름 창고에서 원하는 환자의 과거 영상을 찾아올 필요없이 PACS시스템의 화면에 나타난 메뉴를 통하여 선택하면 신속하게 찾고자 하는 영상이 모니터를 이용하여 출력된다.

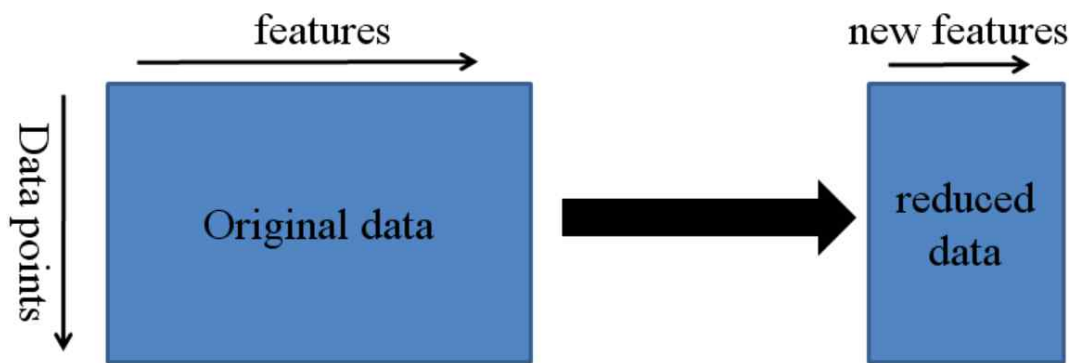
References

[1]<https://m.blog.naver.com/PostView.naver?isHttpsRedirect=true&blogId=ws7026&logNo=120056514260>

5. What is the difference between feature extraction and feature selection?
Find one or more feature selection techniques and describe how they work.

- Feature extraction (특징 추출)

: 기존 Feature들의 조합으로 유용한 Feature들을 새롭게 생성하는 과정.



- Feature selection (특징 선택)

: 관련없거나 중복되는 Feature들을 필터링하고 간결한 subset을 생성하는 과정.



- Feature extraction vs Feature selection

둘 다 차원 축소를 하는 방법이지만 다음과 같은 차이가 있다.

*Feature extraction - 기존 Feature에 기반하여 새로운 Feature들을 생성

*Feature selection - 기존 Feature들의 부분 집합(하위 집단)을 유지

- Feature selection method and techniques

1. Filter method



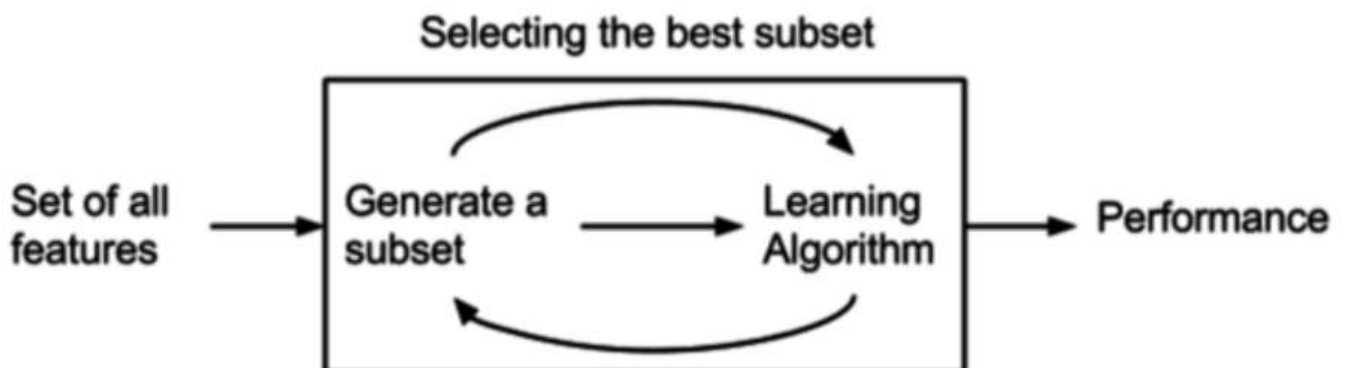
- : Feature의 조합을 정하고, 기계 학습을 돌리고, 성능을 평가함.
- : 조합을 바꿔가면서 계속 반복.
- : 가장 성능이 좋은 조합을 찾아냄.

- : 굉장히 많은 시간 비용이 필요함.
- : Feature가 너무 많으면 overfitting될 위험이 있음.
- : 단순해서 가장 많이 사용되는 방법임.

Filter method Techniques)

- information gain
- chi-square test
- fisher score
- correlation coefficient => 흔히 correlation을 heatmap으로 표현하여 시각화
- variance threshold

2. Wrapper method



- : 전처리 과정에서 미리 통계적 방법으로 Feature selection을 실행하고, 모델을 fitting.
- : 통계적 방법으로 종속변수와 독립변수 간의 피어슨 상관계수를 이용하는 것이 가장 많이 쓰임.
- : 예측 모델을 사용하여 Feature subset을 계속 테스트함.

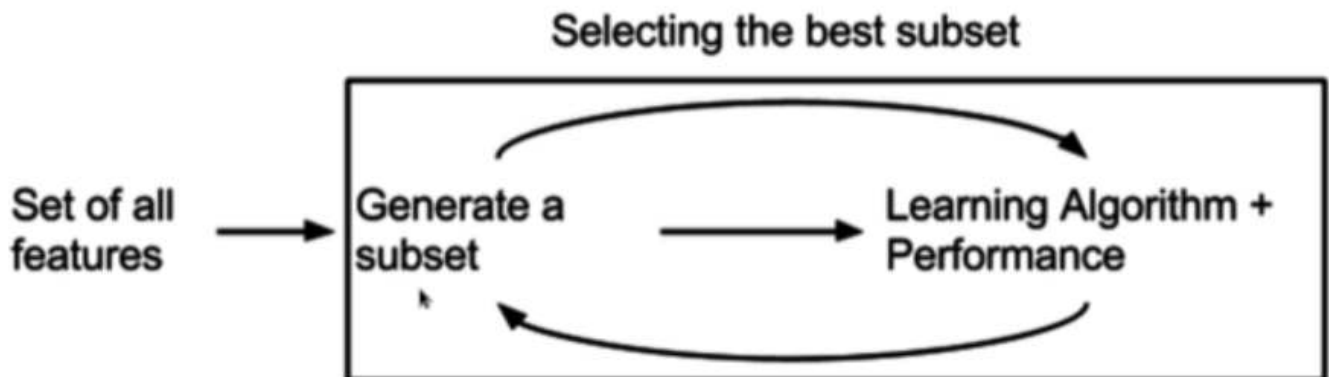
Wrapper method Techniques)

- recursive feature elimination(RFE)
 - : scikit-learn에 함수가 있음
 - : SVM을 사용하여 재귀적으로 제거하는 방법
 - : 유사한 방법으로 Backward elimination, Forward elimination, Bidirectional

elimination이 있음

- sequential feature selection(SFS)
 - :mlxtend에 함수가 있음
 - :그리디 알고리즘으로 빈 subset에서 피처를 하나씩 추가하는 방법
- genetic algorithm
- Univariate selection
- Exhaustive
- mRMR(Minimum Redundancy Maximum Relevance)
 - :피처의 중복성을 최소화하여 Relevancy를 최대화하는 방법

3. Embedded method



: 모델 자체에 Feature selection 기능이 추가되어 있음.

Embedded method Techniques)

- LASSO : L1-norm을 통해 제약 주는 방법
- Ridge : L2-norm을 통해 제약을 주는 방법
- Elastic Net : 위 둘을 선형결합한 방법
- SelectFromModel
 - decision tree 기반 알고리즘에서 피처를 뽑아오는 방법
 - scikit-learn에 함수가 있음

References

- [1]<https://bioinformaticsandme.tistory.com/188>
- [2]<https://subinium.github.io/feature-selection/>

6. What is the cross-validation? What is the leave-one-out validation?

— cross-validation

- 보통은 train set으로 모델을 훈련, test set으로 모델을 검증한다.
- 고정된 test set을 통해 모델의 성능을 검증하고 수정하는 과정만 반복하면, 결국 내가 만든 모델은 test set에만 잘 동작하는 모델이 된다.
- 즉, test set에 과적합(overfitting)하게 되므로, 다른 실제 데이터를 가져와 예측을 수행하면 엉망인 결과가 나와버리게 된다.
- 이를 해결하고자 하는 것이 교차 검증(cross validation)이다.
- 교차 검증은 train set을 train set + validation set으로 분리한 뒤, validation set을 사용해 검증하는 방식이다.

— leave-one-out validation

- 총 N(샘플 수)번의 model을 만들고, 각 모델을 만들 때에 하나의 샘플만 제외하면서 그 제외한 샘플로 test set performance를 계산하여 N개의 performance에 대해서 평균을 내는 방법이다.
- LOOCV의 장점
 - : 모든 샘플에 대해서 다 한번씩은 test하기 때문에 어떠한 randomness도 존재하지 않는다.
 - : validation set approach와는 다르게 굉장히 stable한 결과를 얻을 수 있다.
 - : 하나의 샘플만을 test set으로 사용하기 때문에 그만큼 많은 수의 training data를 활용하여 model을 만들어 볼 수 있다.
- LOOCV의 단점
 - : 많은 수의 model을 만들고 test해야하기 때문에 computing time이 오래 걸릴 수 있다.
 - : model의 다양성을 포함하기 어렵다.

References

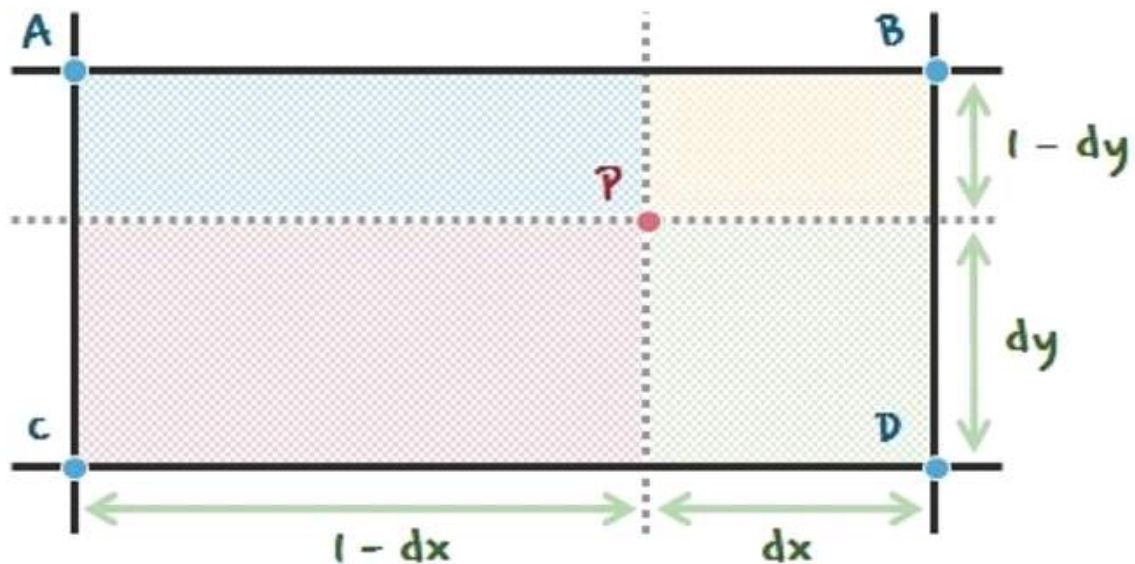
[1]<https://wooono.tistory.com/105>

[2]<https://smlee729.github.io/r/machine%20learning/2015/03/19/1-loocv.html>

7. Describe the bilinear interpolation and the bicubic interpolation, respectively.

7-1. Bilinear interpolation

: 1차원에서의 선형 보간법을 2차원으로 확장한 것으로, linear interpolation을 x축과 y축으로 두 번 적용하여 값을 유추한다. 4개의 인접한 점의 값들과 그에 따른 면적을 가중치(weight)로 하여 새로운 생성된 화소의 값을 얻게 된다.



적용방법 예시

1) 출력영상에서 (5, 5)좌표에 해당하는 값을 구해보고자 한다.

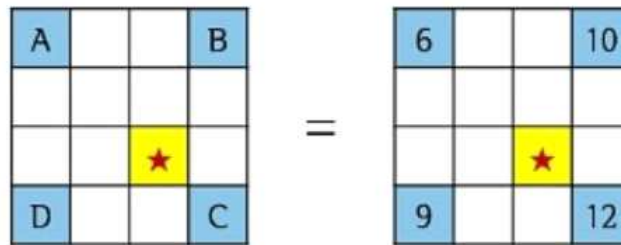
	0	1	2	3	4	5	6	
0	1			3			5	
1								
2								
3	4			6			10	...
4								
5						★		
6	7			9			12	

⋮

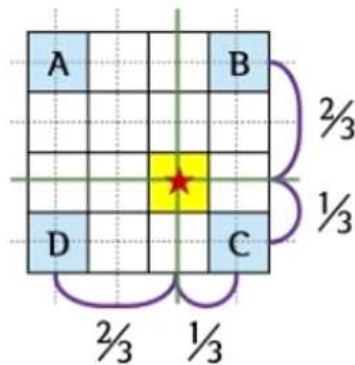
(5,5)의 값=?

2) 구하고자 하는 좌표 (5, 5)와 인접하며, 이미 알고있는 값 4개를 찾는다.

=> 6, 9, 10, 12

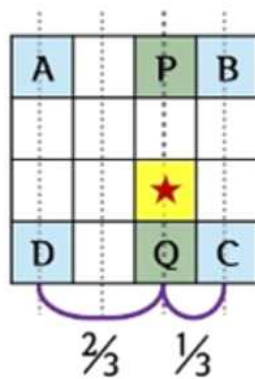


3) 네 점 6, 9, 10, 12와 (5, 5) 사이의 x축, y축 거리비를 구한다.



4) Linear interpolation을 x축에 대하여 적용하여 P, Q의 값을 구한다.

=> P = 약 8.67, Q = 11

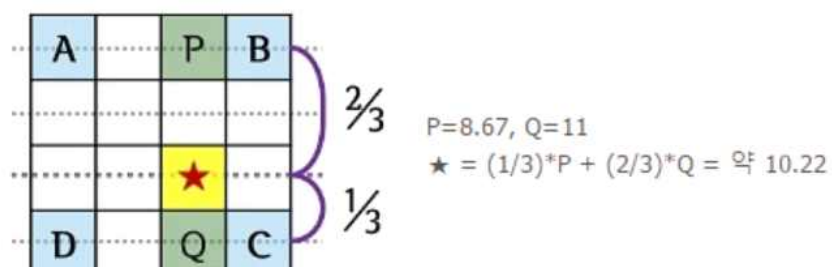


4-1.
A, B 값과 거리비로 P의 값을 구한다.
A=6, B=10
 $P = (1/3)*A + (2/3)*B = \text{약 } 8.67$

4-2.
C, D 값과 거리비로 Q의 값을 구한다.
D=9, C=12
 $Q = (1/3)*D + (2/3)*C = 11$

5) Linear interpolation을 y축에 적용하여 P, Q로부터 찾고자 하는 (5, 5)의 값을 구한다.

=> 약 10.22



7-2. Bicubic interpolation

: bilinear interpolation은 속도는 빠르지만 픽셀 값이 인접 그리드에서 값으로 부드럽게 이어지지 않는 단점이 있다. 인접 그리드 경계에서 픽셀 값이 부드럽게 이어지기 위해서는 적어도 1차 미분이 연속인 보간법을 사용해야 한다. 이러한 조건을 만족시키는 가장 낮은 차수의 다항식 보간법이 bicubic interpolation이다.

적용방법

4점 (0, 0), (0, 1), (1, 0), (1, 1)을 꼭짓점으로 하는 정사각형 내의 임의 지점

$D=\{(x, y) | 0 \leq x \leq 1, 0 \leq y \leq 1\}$ 에서 픽셀을 주는 보간 곡면 $f(x, y)$ 을 주변의 16개 점 $\{(i, j) | -1 \leq i \leq 2, -1 \leq j \leq 2\}$ 에서 픽셀 값을 이용해서 bicubic interpolation을 이용해서 추정하자. 곡면 $f(x, y)$ 은 x 와 y 의 3차 함수로 다음과 같이 쓸 수 있다.

$$f(x, y) = \sum_{i=0}^3 \sum_{j=0}^3 a_{ij} x^i y^j$$

문제는 계수 $\{a_{ij}\}$ 16개를 어떻게 찾을 것인지를이다. 이를 위해 16개의 조건이 필요한데 주변의 16개의 픽셀 값을 이용해서 만들 수 있다. 또한 픽셀 값이 인접 그리드 영역으로 smooth하게 연결되기 위해서는 $f(x, y)$ 의 미분도 고려해야 한다.

1) 4개의 점에서 값, 미분계수, 교차 미분계수를 구하면 다음과 같은 16개의 식을 얻는다.

$$f(0,0)=a_{00};$$

$$f(1,0)=a_{00}+a_{10}+a_{20}+a_{30};$$

$$f(0,1)=a_{00}+a_{01}+a_{02}+a_{03};$$

$$f(1,1)=a_{00}+a_{10}+a_{20}+a_{30}+a_{01}+a_{11}+a_{21}+a_{31}+a_{02}+a_{12}+a_{22}+a_{32}+a_{03}+a_{13}+a_{23}+a_{33};$$

$$f_x(0,0)=a_{10};$$

$$f_x(1,0)=a_{10}+2a_{20}+3a_{30};$$

$$f_x(0,1)=a_{10}+a_{11}+a_{12}+a_{13};$$

$$f_x(1,1)=a_{10}+2a_{20}+3a_{30}+a_{11}+2a_{21}+3a_{31}+a_{12}+2a_{22}+3a_{32}+a_{13}+2a_{23}+3a_{33};$$

$$f_y(0,0)=a_{01};$$

$$f_y(1,0)=a_{01}+a_{11}+a_{21}+a_{31};$$

$$f_y(0,1)=a_{01}+2a_{02}+3a_{03};$$

$$f_y(1,1)=a_{01}+a_{11}+a_{21}+a_{31}+2a_{02}+2a_{12}+2a_{22}+2a_{32}+3a_{03}+3a_{13}+3a_{23}+3a_{33};$$

$$f_{xy}(0,0)=a_{11};$$

$$f_{xy}(1,0)=a_{11}+2a_{21}+3a_{31};$$

$$f_{xy}(0,1)=a_{11}+2a_{12}+3a_{13};$$

$$f_{xy}(1,1)=a_{11}+2a_{21}+3a_{31}+2a_{12}+4a_{22}+6a_{32}+3a_{13}+6a_{23}+9a_{33};$$

2) 위의 식들은 16개의 미지수 $\{a_{ij}\}$ 를 가지는 16개의 연립 방정식이다.

3) 계수 벡터를 \mathbf{v} , 계수 벡터의 계수 행렬을 \mathbf{A} 로 나타내면
 $\mathbf{f} = \mathbf{A}\mathbf{v}$ 이고, \mathbf{A} 의 역행렬을 취해줌으로 \mathbf{v} 를 구할 수 있다.

4) 이렇게 구한 \mathbf{v} 를 이용하여 $f(x, y)$ 를 계산하면, $f(x, y)$ 는 $[0, 1] \times [0, 1]$ 인 사각형 구간에서 연속인 smooth한 곡면을 형성한다. 또한 인접하는 그리드 영역의 곡면과 경계에서 smooth하게 연결된다.

References

- [1]<https://m.blog.naver.com/PostView.naver?isHttpsRedirect=true&blogId=aorigin&logNo=220947541918>
- [2]<https://kipl.tistory.com/55>

8. Describe the role of the pooling layer in a convolutional neural network.
In U-net, explain the reason for reducing the feature map size with the pooling layer and then increasing it again.

– Pooling Layer

: Convolution Layer(CL)를 input으로 받아들인다.

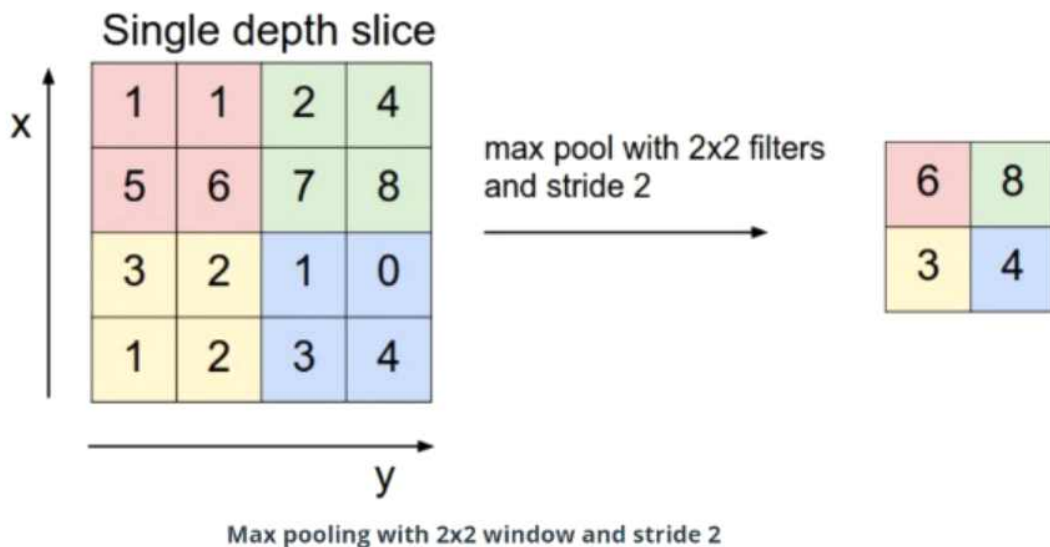
: CL은 각 필터 당 하나의 feature map이 형성되어 있다.

: 복잡한 데이터셋을 가지고 CNN을 구현하는 경우 많은 수의 필터를 필요로 한다.

: 필터가 많다는 얘기는 그만큼 CNN의 차원이 크다는 것이다.

: 고차원을 구현하려면 그에 상응하는 더 많은 수의 파라미터들을 필요로 하고, 그렇게 되면 over-fitting을 초래할 수 있다.

: 따라서 차원을 감소시킬 방법을 필요로 하는데, 이런 역할을 해 주는 레이어가 Pooling Layer이다.



대표적으로 Max pooling이 쓰이는데, 위 그림이 가장 많이 쓰이는 max pooling layer이다.
4x4크기의 window를 2x2로 줄여주고, 각 filter에서 가장 큰 값을 가져오게 된다.

- In U-net, the reason ...

: U-net은 적은 양의 학습 데이터로도 Data Augmentation을 활용해 여러 Segmentation 문제에서 우수한 성능을 보이고, 검증이 끝난 곳은 건너뛰고 다음 Patch부터 새 검증을 하기때문에 속도가 빠르다.

: U-net은 이미지의 전반적인 context 정보를 얻기 위한 네트워크(Contracting Path) + 정확한 지역화(localization)을 위한 네트워크(Expanding Path)로 구성된다.

- Contracting Path : 이미지의 context를 포착한다.

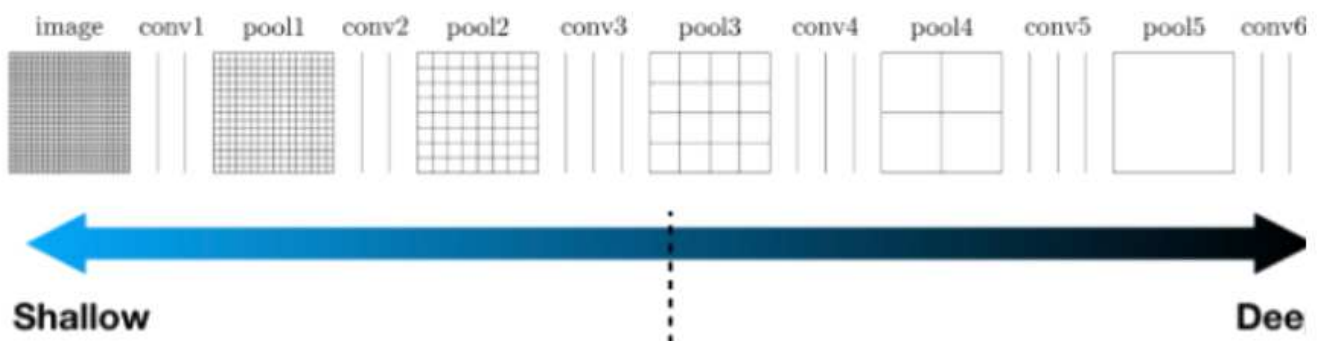
: CONV, POOL ...을 반복하며 Down-sampling을 수행한다.

: 각 step마다 Feature Map의 크기가 절반으로 줄어듦.

- Expanding Path : 특징맵을 확장하고 정확한 localization

: UP-CONV를 수행한다.

: 각 step마다 Feature Map의 크기가 두 배로 늘어남.



- U-net에서 피쳐 맵 크기를 줄인 다음 다시 증가시키는 이유

: 얇은 층은 국소적이고 세밀한 부분의 특징을 추출하고 깊은 층은 전반적으로 추상적인 특징을 추출한다. 서로 다른 특징을 추출하는 이 두 층을 결합해 줌으로서 국소적인 정보와 전역적인 정보를 모두 포함할 수 있게 되고, 이로 인해 성능이 더 좋아지게 된다.

Context 정보를 잘 사용하면서도 정확히 지역화해서, End-to-End 구조로 속도라 빨라지는 장점도 있다.

References

[1]<https://velog.io/@guide333/U-Net-%EC%A0%95%EB%A6%AC>

9. Describe measurement methods (or metrics) to evaluate classification, segmentation, image enhancement, and registration performances, respectively.

9-1. Classification Metrics

– Accuracy

: 전체 예측 개수 중에 맞춘 예측 개수로 나누어 100을 곱한 것.

: 흔히 말하는 얼마나 많이 맞췄냐만 생각하는 것.

– Precision

: 단순히 모든 결과를 X라고 예측하면 Accuracy가 높게 나올 수도 있기 때문에, 진짜로 맞춘 확률을 나타내는 방법.

: $\text{Precision} = \text{True_Positive} / (\text{True_Positive} + \text{False_Positive})$

– Recall

: 맞은 사진을 맞다고 판단하는 확률, 아닌 사진을 아니라고 판단하는 확률을 계산한 것.

9-2. Segmentation Metrics

– Precision / Recall / F1 score

$$\text{Precision} = \frac{TP}{TP+FP}, \text{Recall} = \frac{TP}{TP+FN} \quad F1 - \text{score} = \frac{2\text{Prec} * \text{Rec}}{\text{Prec} + \text{Rec}}$$

		Actual class	
		P	N
Predicted class	P	TP	FP
	N	FN	TN

– Pixel accuracy (PA)

: 분할된 픽셀 수를 전체 픽셀 수로 나눈 것.

– Intersection over Union (IoU)

: 가장 흔하게 쓰이는 metrics 중 하나이다.

: 예측된 분할 map과 ground truth 사이에 intersection 영역으로 정의.

– Dice coefficient

: Medical image analysis에서 유명한 metric

: 예측 및 gt maps의 겹쳐지는 영역에 2를 곱한 후, 두 이미지의 총 픽셀 수로 나누어 정의

$$Dice = \frac{2TP}{2TP+FP+FN} = F1$$

9-3. Enhancement Metrics

– Peak Signal to Noise Ratio (PSNR)

: 신호가 가질 수 있는 최대 전력에 대한 잡음의 전력을 나타낸 것.

: 주로 영상 또는 동영상 손실 압축에서 화질 손실 정보를 평가할 때 사용.

$$\begin{aligned} PSNR &= 10 \cdot \log_{10} \left(\frac{MAX_I^2}{MSE} \right) \\ &= 20 \cdot \log_{10} \left(\frac{MAX_I}{\sqrt{MSE}} \right) \\ &= 20 \cdot \log_{10}(MAX_I) - 10 \cdot \log_{10}(MSE) \end{aligned}$$

– Structural Similarity Index (SSIM)

: 두 영상 사이의 유사성을 측정하는 데 사용.

: 이미지 저하를 구조 정보의 인식 변화로 간주하는 동시에 휘도 마스킹과 대비 마스킹 향을 모두 포함하는 중요한 지각 현상을 통합하는 인식 기반 모델.

– Mean Opinion Score (MOS)

: 모집된 여러 명의 피실험자들이 점수를 매긴 것을 평균낸 것.

: 일반적으로 1은 가장 낮은 인식 품질이고 5는 가장 높은 인식 품질이다.

: 절대 범주 등급 척도는 1~5 사이의 숫자로 매핑하는 것이 매우 보편적으로 사용됨.

9-4. Registration Metrics

– Endpoint error (EPE)

: 두 지점의 Euclidean distance로 정의되는데, Moving image와 Fixed image의 포인트 간의 거리의 차이를 비교한 측정 기준이다.

– Landmark based validation

: Moving image의 Landmark 지점과 Fixed image의 Landmark 지점 간의 거리의 차이를 이용.

– Segmentation based validation

: Moving image의 Segmentation된 class와 Fixed image의 Segmentation된 class가 잘 매칭되었는지 비교하는 방법.

– Inverse consistency error

: Moving image의 점(x, y)에서 Fixed image로 변환하고 (x', y') 역변환을 통해 (x'', y'') 이 얻어졌을 때, (x, y)와 (x'', y'')들의 거리를 비교하는 방법.

– Intensity variance

: 정합된 영상의 겹쳐지는 곳들의 차이가 적으면 정합이 잘되었다고 평가하는 방법.

References

[1]<https://dgist.edwith.org/medical-20200327/lecture/63209?isDesc=false>

10. Find out how the p-value can be computed and explain what it means.

- p-value의 계산

p-value은 다음과 같은 3가지 확률의 합으로 구성된다.

1. 해당 사건이 발생할 확률
2. 해당 사건과는 다르지만 그만큼 드물게 일어나는 사건이 발생할 확률
3. 해당 사건보다 더 드물게 일어나는 사건의 확률

예를 들어 동전을 4번 던져서 3번 앞면이 나오는 경우 p-value는 다음과 같이 계산된다.

1. 동전을 4번 던져서 3번 앞면이 나오는 경우의 확률은 $4 / 16 = 0.25$
2. 동전을 4번 던져서 3번 앞면이 나오는 경우(16번 중 4번) 만큼 드문 사건은 동전을 4번 던져서 1번 앞면이 나오는 사건(16번 중 4번)이다. 이 사건이 발생할 확률은 역시나 $4 / 16 = 0.25$
3. 동전을 4번 던져서 3번 앞면이 나오는 경우보다 더욱 드물게 일어나는 사건은 4번 모두 앞면이 나오거나 0번 앞면이 나오는 사건의 확률의 합이다. $1 / 16 + 1 / 16 = 0.125$

위 확률을 모두 합하면 $0.25 + 0.25 + 0.125 = 0.625$ 이고, 이렇게 계산된 값이 동전을 4번 던져서 3번 앞면이 나올때의 p-value가 된다.

- p-value의 의미

: 일반적으로 p-value는 어떤 가설을 전제로, 그 가설이 맞다는 가정 하에, 내가 현재 구한 통계값이 얼마나 자주 나올 것인가를 의미한다고 할 수 있다.

: 가설검증이라는 것은 전체 데이터의 일부만을 추출하여 평균을 내고, 그 평균이 전체 데이터의 평균을 잘 반영한다는 가정 하에 전체 데이터의 평균을 구하는 작업인데, 아무리 무작위 추출을 잘 한다 하더라도 추출된 데이터의 평균은 전체 데이터의 평균에서 멀어질 수 있게 된다. 따라서, 내가 추출한 이 데이터의 평균이 원래의 전체 데이터의 평균과 얼마나 다른 값인지를 알 수 있는 방법이 필요하게 된다. 이와 같은 문제 때문에 나온 값이 p-value이다.

- p-value를 이용한 가설 검정의 활용

우리는 제약회사에서 바이러스 치료제를 만드는 사람들이라하자. 새로 개발한 치료제가 효과가 있는지 검정하고 싶다. 치료의 효과는 치유가 되는데 걸리는 시간으로 정량화를 하기로 하고, 치료 효과가 높다면 그 기간이 짧을 것이다. 여기서 입증해야 할 가설(“새로운 치료제는 효과적이다.”)은 대립가설이고, 귀무가설은 “새로운 치료제는 효과가 없다.”가 된다. 입증의 과정은 앞서 봤듯이, “새로운 치료제는 효과가 없다”는 가설을 참이라고 가정할 때, 예외적인 경우를 관찰하게 되면 “새로운 치료제의 효과는 없다”는 가설을 기각하고, “새로운 치료제는 바이러스 치료에 효과적이다.”라는 가설을 채택하게 된다. 그 예외적인 경우란 “새로운 치료제는 효과가 없다”는 가설로는 도저히 설명하기 힘든 자료, 그 치료제를 투여했더니 치유시간이 상당히 짧아진 데이터가 될 것이다.

References

- [1]http://www.dh-datascience.net/?page_id=23
- [2]<https://adnoctum.tistory.com/332>