



의료 분야 신경망 적용 기술

성명 석진욱

소속
한국전자통신연구원 (ETRI) 인공지능컴퓨팅연구소
초성능컴퓨팅 연구본부 AI컴퓨팅시스템SW연구실

인공지능 기술의 대중화
(AI Democratization)를 위한
제2회 탱고 커뮤니티 컨퍼런스





1	Introduction	00
	1. Requirement of AI for Medical Industry	
2	Classification	00
	1. Technology for Diagnosis based on Classification AI	
	2. Experimental Results (For Kaggle)	
3	Generative Algorithm	00
	1. DDPM 기반 흉부 X-ray 치료경과 영상 생성	
	2. Class Activation Mapping (CAM)	
	3. ETRI 개발 치료경과 생성을 위한 DDPM 모델	
	4. 기존 방법과 ETRI 개발 DDPM + CAM 기반 흉부 X-ray 치료경과 비교	
4	Conclusion and Future Work	00



Requirement of AI for Medical Industry

- 공공 의료 분야의 경우, 소수의 의료 전문가가 대량의 데이터 처리
 - 대량의 데이터중 소수 유증 데이터의 자동 선별 ⇒ 의료 전문가 노동 압박 경감.
 - 인적 오류 가능성을 줄여, 보다 정확한 진단 가능
 - 인공지능의 Classification 응용 대상
- 치료에 대한 가상 경과 영상 생성
 - 치료 경과의 검증 및 효과 예측
 - 생성형 AI 적용으로 가상 치료 경과 영상 생성
 - 정확한 생성을 위한 비지도 학습 기반 ROI(Region of Interest)추출 필요



Classification

- 흉부 X선 폐렴진단 문제에 신경망 적용
- 기존, 의료용 Classification 신경망의 경우 진단 정확도는 90% 초반 수준
- 폐렴 진단에 한하여, 최소 97% 이상의 폐렴 진단 정확도를 목표

Technology for Diagnosis based on Classification AI

- Appropriate AI Algorithm : DenseNet-121
- Appropriate Data Enhancement : Image Augmentation
- Appropriate Hyper Parameters : Grid Search
- Based on Neural Tangent Kernel



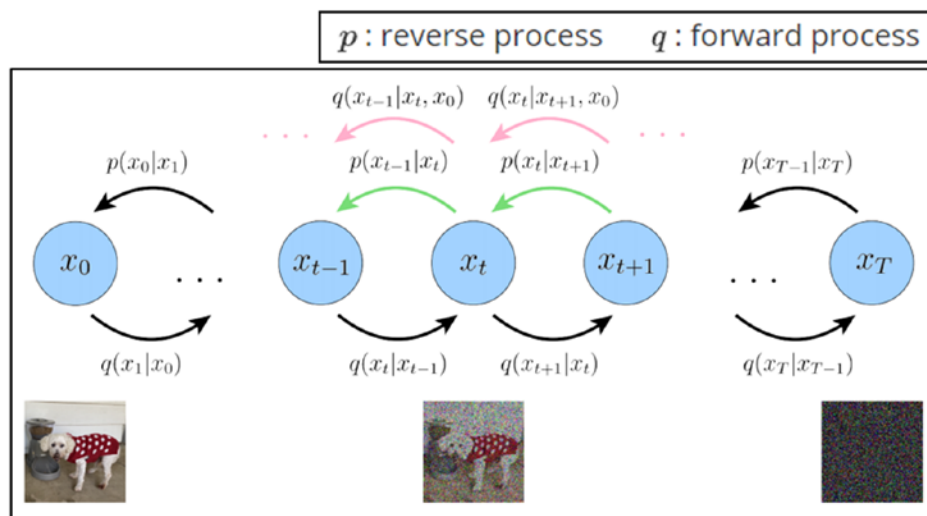
Experimental Results (For Kaggle)

- Data Set : Kaggle Data Set, 서울대 병원 데이터 (IRB 준수)
- Environment : python 3.8.12/Pytorch 1.11/ CUDA 11.5/ Ubuntu 20.04/ NVIDIA A100

Technology	max Accuracy (%)	
Base	86.41	Kaggle Data
Hyper Pram. Optim	89.94	Optimizer, Learning rate, Momentum
Algorithm Optimi	90.89	ResNet-152→ Densent-121
Augmentation	95.60	Randomly Select 4 Augamentation
Applying NTK	97.12	increasing Num. of FC (1024 → 2048)

- 서울대 데이터의 경우 **Training/ 99.5%, Testing / 98.2%**

- Stochastic Analysis를 기반으로 Whitening Transform과 Auto-Encoder/Optimal Controller의 결합으로, 최적화된 정보를 생성해내는 인공지능 알고리즘
- DDPM(Denosing Diffusion Probabilistic Model) 기반 영상 생성



* Calvin Luo. Understanding Diffusion Models, 2022

- Forward Process : 데이터에 노이즈를 점진적으로 더하여 Whitening 변환 수행
- Reverse Process는 학습을 통해 노이즈 영상으로 부터 이미지를 생성

ETRI. All copyrights reserved

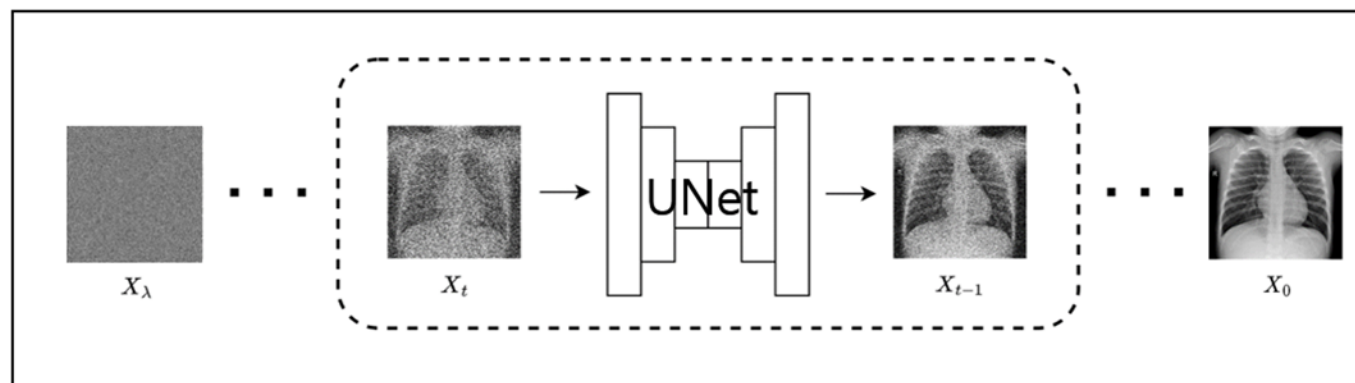


DDPM 기반 흉부 X-ray 치료경과 영상 생성

• 개념과 알고리즘

Algorithm 1 Training

```
1: repeat  
2:  $\mathbf{x}_0 \sim q(\mathbf{x}_0)$   
3:  $t \sim \text{Uniform}(\{1, \dots, T\})$   
4:  $\epsilon \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$   
5: Take gradient descent step on  
    $\nabla_{\theta} \|\epsilon - \epsilon_{\theta}(\sqrt{\bar{\alpha}_t}\mathbf{x}_0 + \sqrt{1 - \bar{\alpha}_t}\epsilon, t)\|^2$   
6: until converged
```



- 정상 흉부 X-ray 영상으로 DDPM 신경망을 학습 (UNet 학습)
- 폐렴 환자 X-ray 영상을 DDPM에 입력하여 동일하게 Forward 시킨 후 Backward 과정을 통해 정상 흉부 X-ray 영상 생성

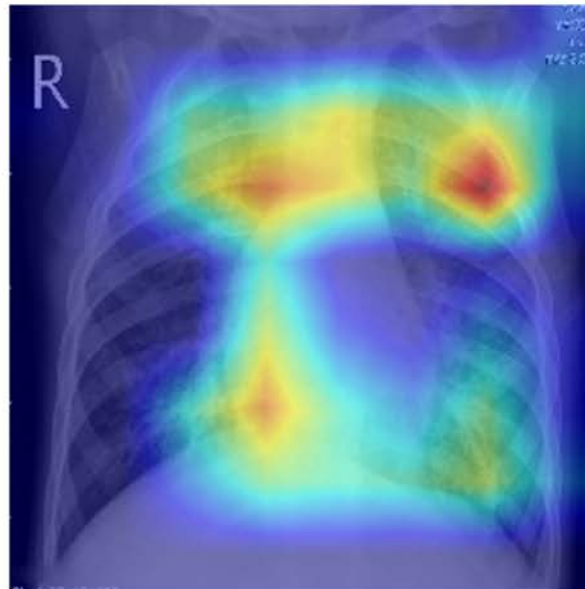


Class Activation Mapping (CAM)

- 기존 DDPM 을 사용할 경우 환자 X-ray 영상에 대한 왜곡이 심함
- 치료 경과 영상 생성시 동일환자가 아닌 별개 정상인 영상이 출력
- 인가하는 노이즈의 농도를 CAM을 결합, 비 균질적으로 인가하여 영상 생성



X_0



CAM



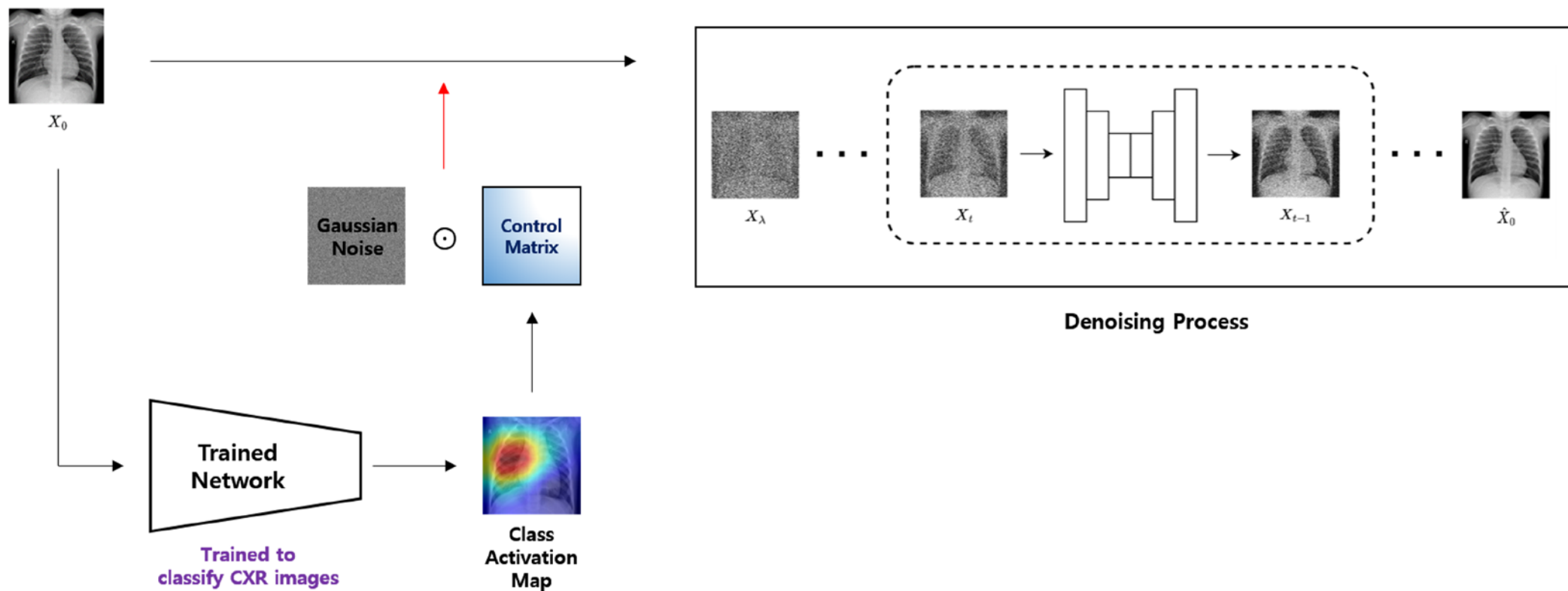
X_λ



ETRI 개발 치료경과 생성을 위한 DDPM 모델

- λ 값(치료 기간)을 조절하여 치료 경과에 따른 영상 생성

AnoDDPM + CAM





기존 방법과 ETRI 개발 DDPM + CAM 기반 흉부 X-ray 치료경과 비교

- $\lambda = 600$ 의 경우 (가장 가능한 노이즈를 많이 투입한 경우)

X_0 (input)



AnoDDPM



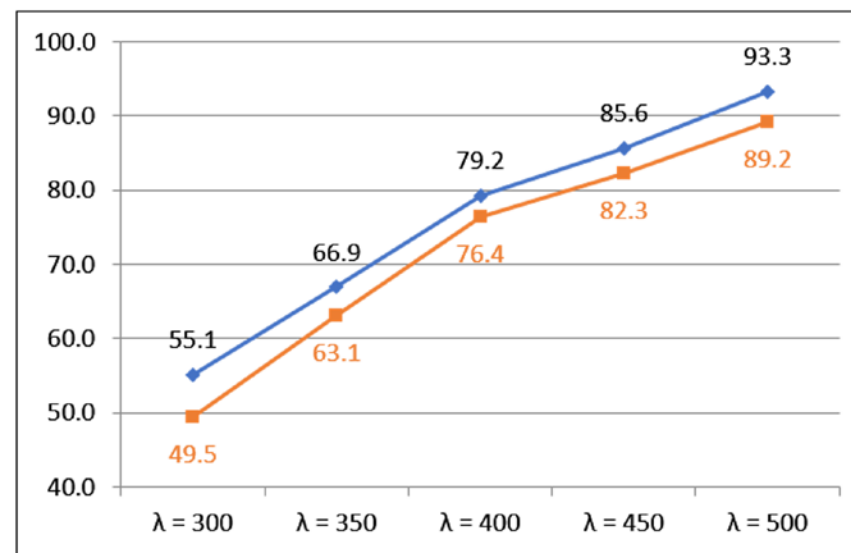
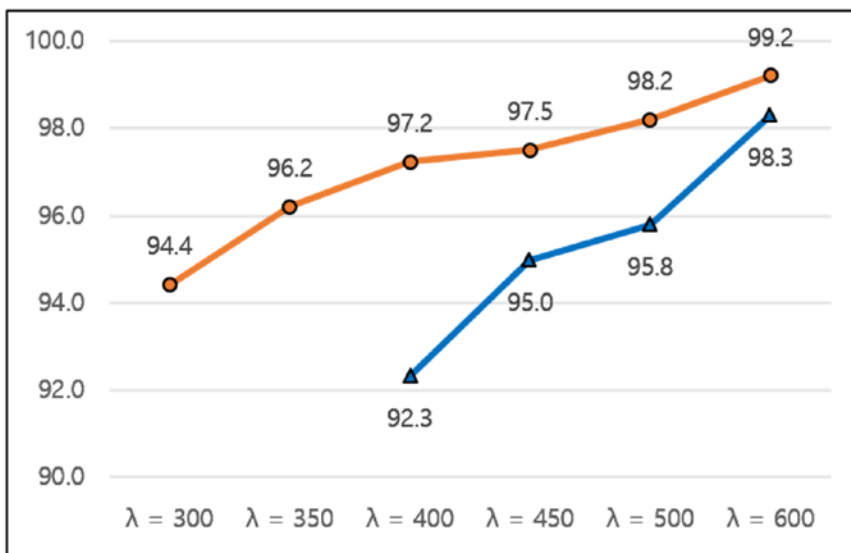
Ours





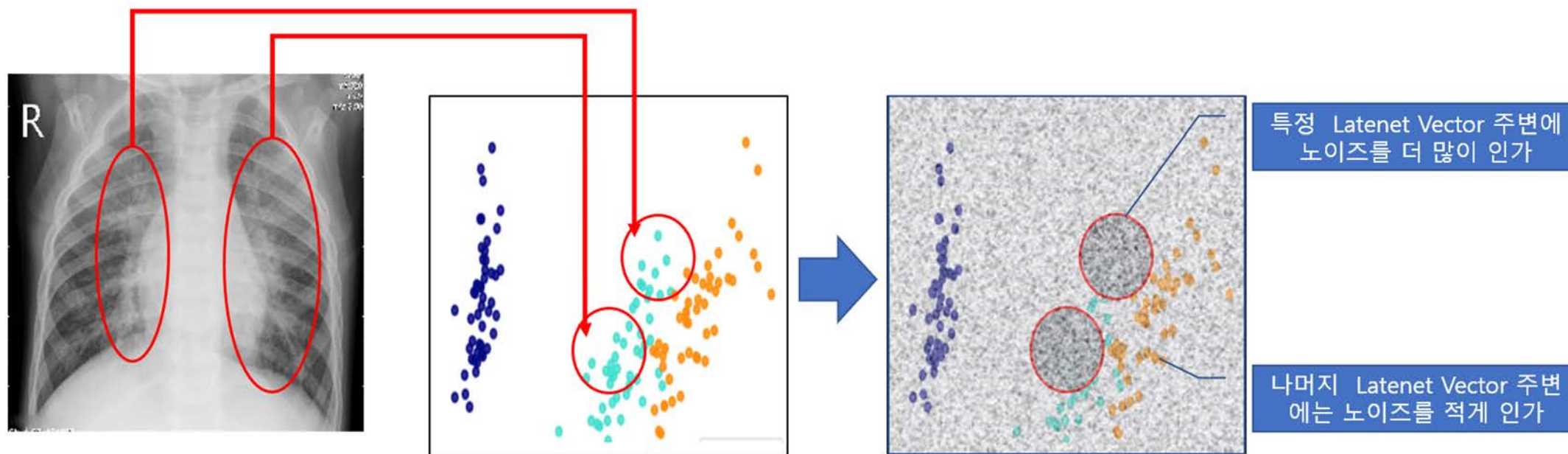
기존 방법과 ETRI 개발 DDPM + CAM 기반 흉부 X-ray 치료경과 비교

생성이미지 정상 클래스 판별률 (좌 : SNU data, 우 : Kaggle data)
(densenet121_2048fclayer 모델로 분류)



- 주황색 : 노이즈에 matrix 적용시 (Ours)
- 파란색 : 미적용 (기존 AnoDDPM)

- DDPM + CAM 이 아닌, Latent Diffusion + Latent CAM 개발
- 원본 환자 영상에 버다 일치되는 치료 경과 영상 생성 알고리즘 개발
- 고도화된 Nonlinear Optimization/Learning 알고리즘 개발

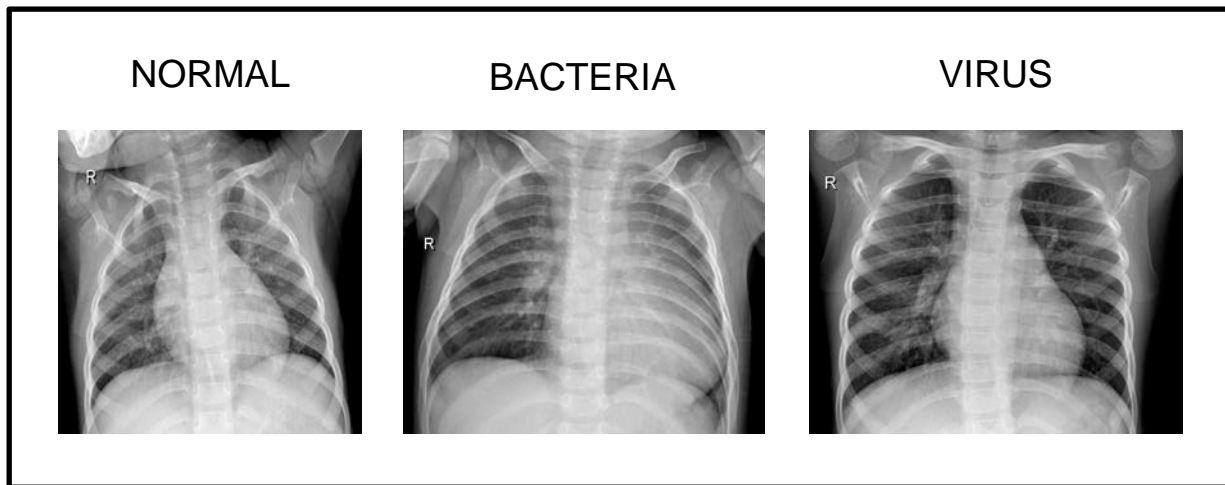


감사합니다.



Dataset (Kaggle Data)

Guangzhou Women and Children's Medical Center data (Kaggle)



데이터셋의 사용목적

- 딥러닝기반 폐렴 진단의 가능성을 확인
- 흉부 X-ray 사용한 딥러닝 진단의 우리만의 기준점 만들기

Dataset Configuration

set	normal	pneumonia	
		bacteria	virus
train	1349	2538	1345
test	234	242	148

- 1 ~ 5살 사이의 소아과 환자
- 2 categories (Pneumonia, Normal)
- 8-bit, Gray scale images
- 3000 x 3000 이하의 다양한 크기 이미지

Hyperparameters

1. Hyperparameter tuning with Grid search

Variables for Grid Search

hyper-parameter	variables
optimizer	SGD, NAG, Adam
learning rate	$0.25 \times \frac{1}{2^k}$, $k \in \{0, 1, 2, \dots, 11\}$
momentum	0.3, 0.9
epochs	100, 200, 300, ... , 1000
# of layers (ResNet)	18, 50, 152, 200

고정 변수

- 입력이미지는 256 x 256 크기, single channel (gray scale)로 변환하였음
- Augmentation : random rotation [-10, 10]
- Mini batch size : 64
- Pretrained weights 미사용 (randomly initialized weights 사용)

신경망에 따른 성능 비교

- VGG16
- ResNet
- InceptionV3
- DenseNet



Augmentation (with 'Albumentations')

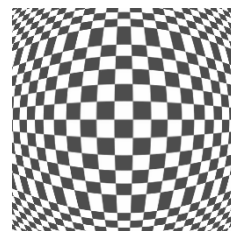
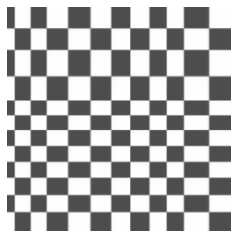
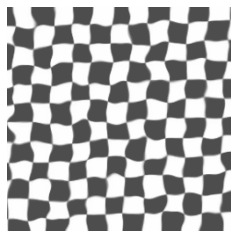
Non-rigid transforms

Original

Elastic deformation

Grid distortion

Optical distortion

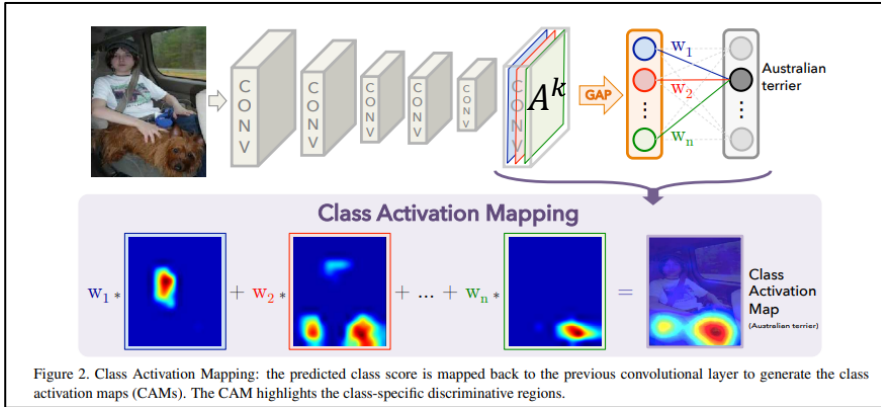


Random

- Rotate
- Scaling
- Horizontal flip



Summary of CAM and GRAD-CAM



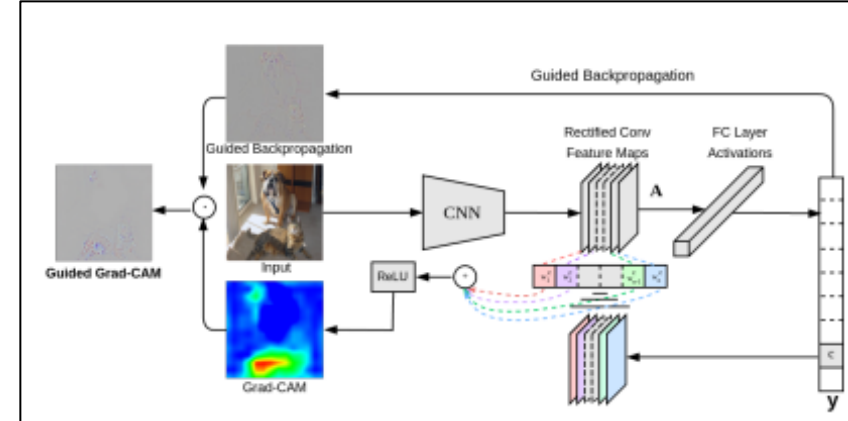
- Global Average Pooling(GAP) layer 에서 나온 feature map A^k 이 softmax 에 들어가는 구조가 있는 경우

$$y^c = \sum_k w_k^c \frac{1}{Z} \sum_i \sum_j A_{ij}^k$$

- GAP 에서 나와서 softmax 로 들어가는 weight(학습된), 마지막 conv layer 에서 나온 feature map 을 가지고 CAM 을 만들

$$L_{CAM}^c = \sum_k w_k^c A^k$$

- y^c : class c 에 대한 score , Z : feature map 의 모든 값들의 갯수



- 마지막 conv layer의 feature map 과 gradient information 사용
- 각 feature map A^k 에 대하여 class 에 관한 gradient

$$\frac{\partial y^c}{\partial A^k}$$

- 앞에서 말한 gradient 값들을 평균내서 the neuron importance weights α_k^c 를 구한다

$$\alpha_k^c = \frac{1}{Z} \sum_i \sum_j \frac{\partial y^c}{\partial A^k}$$

- 구한 α_k^c 로 feature map 들을 weighted sum 하고 ReLU 적용해서 CAM 만들

$$L_{Grad-CAM}^c = ReLU \left(\sum_k \alpha_k^c A^k \right)$$



Brief Introduction of Mathematical Analysis for DDPM



Reverse process $p(X_{t-1}|X_t)$ 를 근사할 $q(X_{t-1}|X_t, X_0)$ 를 구하는 방법 (Based on Intuitive SDE and Bayesian Analysis)

①

$$\alpha_t = 1 - \beta_t, \quad \bar{\alpha}_t = \prod_{i=1}^t \alpha_i$$

β 는 제어를 위한 파라미터

• $\{\beta_t \in (0,1)\}_{t=1}^T$

$q(X_t|X_{t-1}) := \mathcal{N}(X_t; \sqrt{1-\beta_t}X_{t-1}, \beta_t I)$ 에 **reparameterization trick** 적용

$$\Rightarrow X_t = \sqrt{1-\beta_t}X_{t-1} + \beta_t I$$

$$\Rightarrow X_t = \sqrt{\bar{\alpha}_t}X_{t-1} + (1 - \bar{\alpha}_t)I$$

$$\Rightarrow q(X_t|X_0) := \mathcal{N}(X_t; \sqrt{\bar{\alpha}_t}X_0, (1 - \bar{\alpha}_t)I)$$

②

이후 **베이즈 정리** 사용하여 다음과 같이 나타낼 수 있고

$$q(X_{t-1}|X_t, X_0) = \frac{q(X_t|X_{t-1}, X_0)q(X_{t-1}|X_0)}{q(X_t|X_0)}$$

KL divergence 로 두 분포간의 차이를 구하고 식을 정리해 나가면

노이즈를 예측함으로서 diffusion model 을 최적화 할 수 있다는 결론이 나온다.

(중간과정은 생략)

Additional Result

AnoDDPM / 사용이미지 = SNU CXR data (HE) / $\lambda = 400, 500, 600$

위: 원본 | 아래: 생성이미지

