

KSBi-BIML 2024

Bioinformatics & Machine Learning(BIML)
Workshop for Life and Medical Scientists

생명정보학 & 머신러닝 워크샵 (온라인)



Recent advances in AI for healthcare

예종철 _ KAIST



KSBI
KOREAN SOCIETY FOR
BIOINFORMATICS

| 한국생명정보학회



본 강의 자료는 한국생명정보학회가 주관하는 BIML 2024 워크샵 온라인 수업을 목적으로 제작된 것으로 해당 목적 이외의 다른 용도로 사용할 수 없음을 분명하게 알립니다.

이를 다른 사람과 공유하거나 복제, 배포, 전송할 수 없으며 만약 이러한 사항을 위반할 경우 발생하는 **모든 법적 책임은 전적으로 불법 행위자 본인에게 있음을 경고**합니다.

KSBI-BIML 2024

Bioinformatics & Machine Learning(BIML)

Workshop for Life and Medical Scientists

안녕하십니까?

한국생명정보학회가 개최하는 동계 교육 워크숍인 BIML-2024에 여러분을 초대합니다. 생명정보학 분야의 연구자들에게 최신 동향의 데이터 분석기술을 이론과 실습을 겸비해 전달하고자 도입한 전문 교육 프로그램인 BIML 워크숍은 2015년에 시작하여 올해로 벌써 10년 차를 맞이하게 되었습니다. BIML 워크숍은 국내 생명정보학 분야의 최초이자 최고 수준의 교육프로그램으로 크게 인공지능과 생명정보분석 두 개의 분야로 구성되어 있습니다. 올해 인공지능 분야에서는 최근 생명정보 분석에서도 응용이 확대되고 있는 다양한 인공지능 기반 자료모델링 기법들에 대한 현장 강의를 진행될 예정이며, 관련하여 심층학습을 이용한 단백질구조예측, 유전체분석, 신약개발에 대한 이론과 실습 강의를 함께 제공될 예정입니다. 또한 단일세포오믹스, 공간오믹스, 메타오믹스, 그리고 롱리드염기서열 자료 분석에 대한 현장 강의는 많은 연구자의 연구 수월성 확보에 큰 도움을 줄 것으로 기대하고 있습니다.

올해 BIML의 가장 큰 변화는 최근 연구 수요가 급증하고 있는 의료정보자료 분석에 대한 현장 강의를 추가하였다는 것입니다. 특히 의료정보자료 분석을 많이 수행하시는 의과학자 및 의료정보 연구자들께서 본 강좌를 통해 많은 도움을 받으실 수 있기를 기대하고 있습니다. 또한 다양한 생명정보학 분야에 대한 온라인 강좌 프로그램도 점차 증가하고 있는 생명정보 분석기술의 다양화에 발맞추기 위해 작년과 비교해 5강좌 이상을 신규로 추가했습니다. 올해는 무료 강좌 5개를 포함하여 35개 이상의 온라인 강좌가 개설되어 제공되며, 연구 주제에 따른 연관된 강좌 추천 및 강연료 할인 프로그램도 제공되며, 온라인을 통한 Q&A 세션도 마련될 예정입니다. BIML-2024는 국내 주요 연구 중심 대학의 전임 교원이자 각 분야 최고 전문가들의 강의로 구성되었기에 해당 분야의 기초부터 최신 연구 동향까지 포함하는 수준 높은 내용의 강의를 될 것이라 확신합니다.

BIML-2024을 준비하기까지 너무나 많은 수고를 해주신 운영위원회의 정성원, 우현구, 백대현, 김태민, 김준일, 김상우, 장혜식, 박종은 교수님과 KOBIC 이병욱 박사님께 커다란 감사를 드립니다. 마지막으로 부족한 시간에도 불구하고 강의 부탁을 흔쾌히 허락하시고 훌륭한 현장 강의와 온라인 강의를 준비하시는데 노고를 아끼지 않으신 모든 강사분들께 깊은 감사를 드립니다.

2024년 2월

한국생명정보학회장 이 인 석

Recent Advances in AI for Healthcare

2016년 딥러닝 기반 인공지능 기술이 의료 진단 영역에 소개되어 세상을 놀라게 한 이후 이미 많은 의료인공지능 기술이 발전하고 관련 스타트업이 빠르게 상장을 하고, 의료 현장 구석 구석에 인공지능 기술이 녹아들고 있다. 하지만 현존하는 AI 기술은 많은 데이터와 고비용의 레이블 데이터를 요구하고 있다,

본 강의에서는 이러한 문제를 해결하고 위해 최근에 제시되고 있는 다양한 AI 기술이 의료영역에 어떻게 사용되고 있는 지를 설명한다. 특히 Transformer, Vision Transformer, 및 강조학습 및 자기 지도 학습 기술 대해 원리를 설명하고, 이를 통하여 의료에 적용된 예를 소개한다. 또한 최근에 중요한 주제로 떠오르고 있는 원천 모델 (foundation model)에 대한 설명을 하며, 이것이 어떻게 의료현장의 AI를 혁신 시킬수 있는지 논의한다.

강의는 다음의 내용을 포함한다:

- Transformer, Vision Transformer
- Contrastive learning
- Self-Supervised learning
- Foundation model

* 참고강의교재:

Jong Chul Ye, "Geometry of Deep Learning: A Signal Processing Perspective", Springer, 2022

* 강의 난이도: 중급

* 강의: 예종철 교수 (KAIST AI 대학원)

Curriculum Vitae

Speaker Name: Jong Chul Ye, Ph.D.



► Personal Info

Name Jong Chul Ye
Title Professor
Affiliation KAIST AI

► Contact Information

Address KAIST, N5 Rm 2221, 291 Daehak-ro, Yuseong-gu,
Daejeon 34141, Republic of Korea
Email jong.ye@kaist.ac.kr
Phone Number 042-350-4320

Research Interest

Deep learning for Biomedical Imaging & Healthcare, Computer Vision, Generative Models, Diffusion Models

Educational Experience

1993 B.S. in Control and Instrumentation Engineering, Seoul National University, Korea
1995 M.S. in Control and Instrumentation Engineering, Seoul National University, Korea
1999 Ph.D. in Electrical and Computer Engineering, Purdue University, USA

Professional Experience

Jan 2022-Current Professor, Kim Jaechul Graduate School of AI, KAIST
Aug. 2004-Dec. 2021 Professor, Department of Bio and Brain Engineering, KAIST
March 2016 KAIST Endowed Chair Professor
2003-2004 Senior Researcher, X GE Global Research Center, New York
2001-2003 Senior Member Research Staff, Philips Research Center, NY

Selected Publications (5 maximum)

1. Kang, E., Min, J., & Ye, J. C. (2017). A deep convolutional neural network using directional wavelets for low-dose X-ray CT reconstruction. *Medical physics*, 44(10), e360-e375.
2. Oh, Yujin, Sangjoon Park, and Jong Chul Ye. "Deep learning COVID-19 features on CXR using limited training data sets." *IEEE transactions on medical imaging* 39.8 (2020): 2688-2700.
3. Ye, J. C., Han, Y., & Cha, E. (2018). Deep convolutional framelets: A general deep learning framework for inverse problems. *SIAM Journal on Imaging Sciences*, 11(2), 991-1048.
4. Hyungjin Chung, Byeongsu Sim, Dohoon Ryu, Jong Chul Ye, "Improving Diffusion Models for Inverse Problems using Manifold Constraints", Thirty-sixth Conference on Neural Information Processing Systems (NeurIPS), 2022.
5. Gwanghyun Kim, and Jong Chul Ye. "DiffusionCLIP: Text-guided image manipulation using diffusion models," *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2022.

Recent Advances in AI for Healthcare

Jong Chul Ye (예종철 교수)
Professor

Graduate School of AI
KAIST

Year 2016..

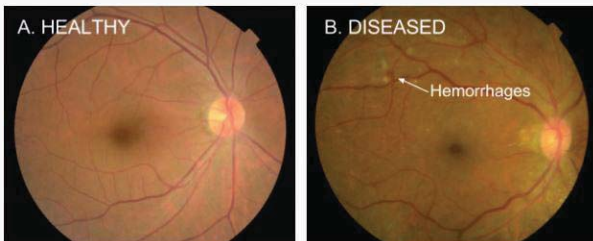
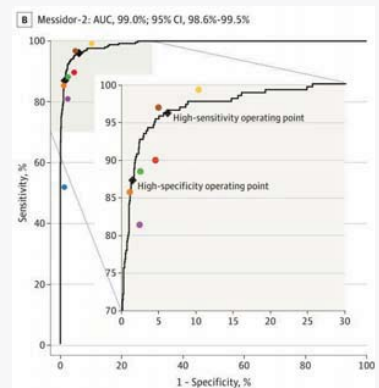


□ Deep Learning for Diabetic Retinopathy (By Google Brain)

JAMA | Original Investigation | INNOVATIONS IN HEALTH CARE DELIVERY

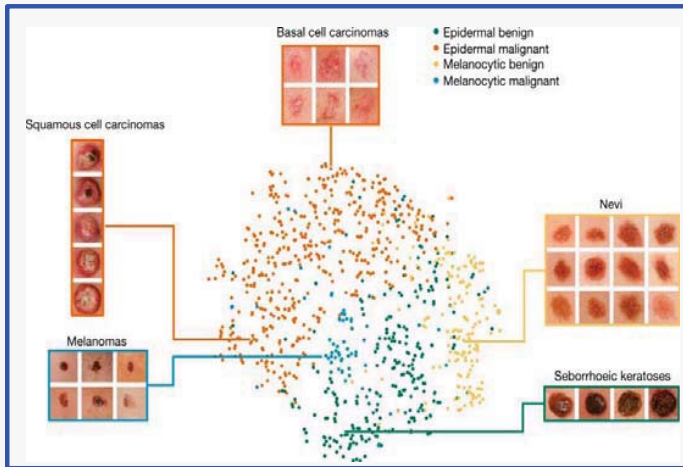
Development and Validation of a Deep Learning Algorithm for Detection of Diabetic Retinopathy in Retinal Fundus Photographs

Varun Gulshan, PhD; Lily Peng, MD, PhD; Marc Coram, PhD; Martin C. Stumpe, PhD; Derek Wu, BS; Arunachalam Narayanaswamy, PhD; Subhashini Venugopalan, MS; Kasumi Widner, MS; Tom Madams, MEng; Jorge Cuadros, OD, PhD; Ramasamy Kim, OD, DNB; Rajiv Raman, MS, DNB; Philip C. Nelson, BS; Jessica L. Mega, MD, MPH; Dale R. Webster, PhD



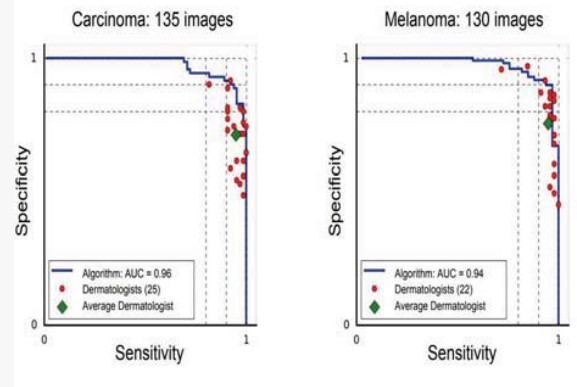
□ Skin Cancer Detection by Deep Learning

- Stanford Univ. Deep CNN(GoogleNet Inception v3)



Esteva et al. Nature, 2017

Test set: Dermatologist Comparison (376 images)



Low Dose CT Grand Challenge

2016년 8월

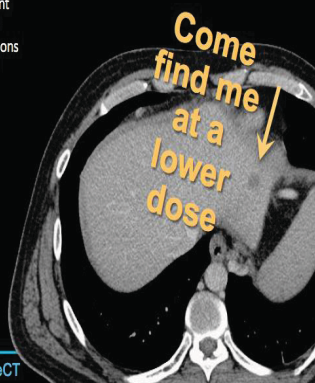
NIH National Institute of Biomedical Imaging and Bioengineering

AAEP AMERICAN ASSOCIATION OF PHYSICISTS IN MEDICINE

MAYO CLINIC

CT Clinical Innovation Center

- Radiologist-selected abdominal CT patient cases (10 training, 20 testing) with noise inserted to simulate lower dose acquisitions
- Projection data converted into an open format (user manual and reading tools provided)
- Apr 2016: Participants submit reconstructed images or denoised images to AAPM website
- Jun 2016: Images read by radiologists at the host site
- Aug 2016: Winners announced at AAPM Annual Meeting



www.aapm.org/GrandChallenge/LowDoseCT

□ The End of Radiology?



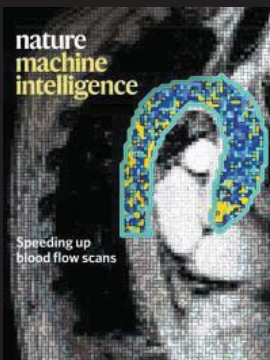
Jeffery Hinton 2016

“We should stop training radiologists now, it’s just completely obvious within five years deep learning is going to do better than radiologists.

‘If you work as a radiologist, you are like Wile E. Coyote in the cartoon: you’re already over the edge of the cliff, but you haven’t looked down.’



□ The Era of Augmented Intelligence



AI IN HEALTHCARE INNOVATION LANDSCAPE 2020

data root labs



© DataRoot Labs and datarootlabs.com, 2020. Feel free to share provided that full and clear credit is given to @datarootlabs with appropriate and specific direction to the original content.

□ AI role in healthcare

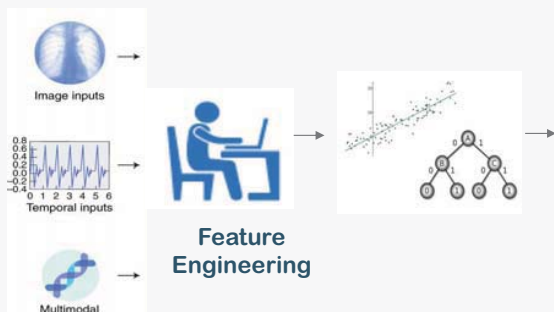
- Clinical assessment:
 - vitals, history, interview ← EMR/EHR analysis by AI
- Making a diagnosis
 - AI driven analysis of lab tests, imaging, bio-signals
- Treatment plan
- Prescribe medication
- Making a prognosis
- Making referrals for advanced treatment
- Surgical procedure

❑ Other AI's roles in healthcare

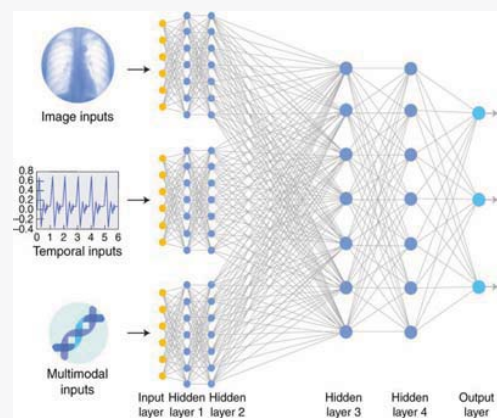
- Workflow improvements
 - Image enhancements
 - Segmentation
 - Registration
 - Quantification
 - Retrievals
- Drug discovery

❑ Classical vs Modern AI

Classical AI



Modern Deep Learning



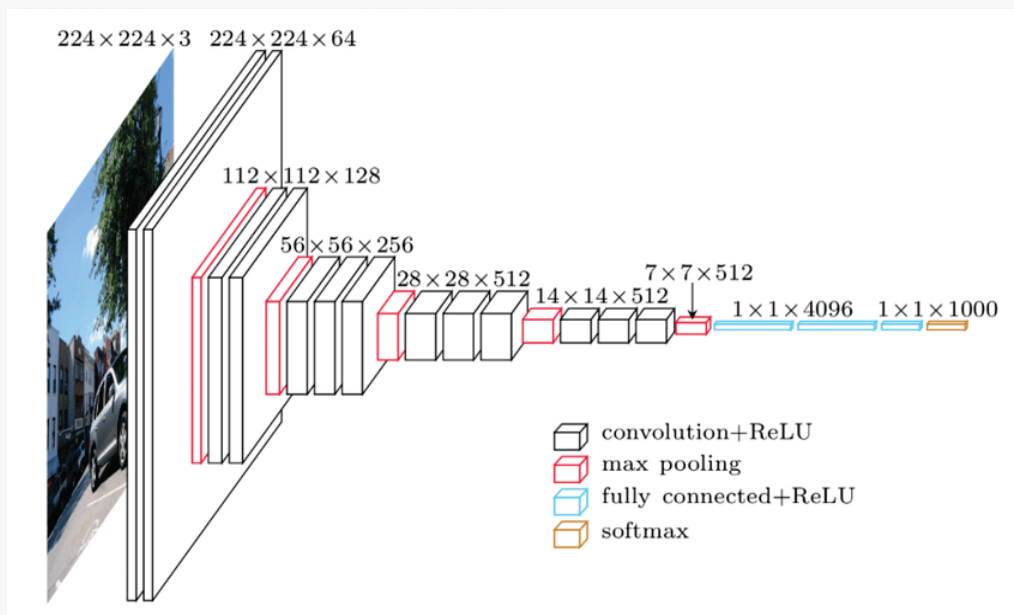
Esteva et al, Nature Medicine, (2019)

□ Technical Challenges in AI for Healthcare

- Limited data
 - Overfitting → Vision Transformer
 - Cost of labeling → self-supervised learning
 - No paired reference → generative models
 - Data privacy → federated learning
 - Multi-modal data → Vision-language pretraining

CNN, ViT

□ VGGNet : a CNN

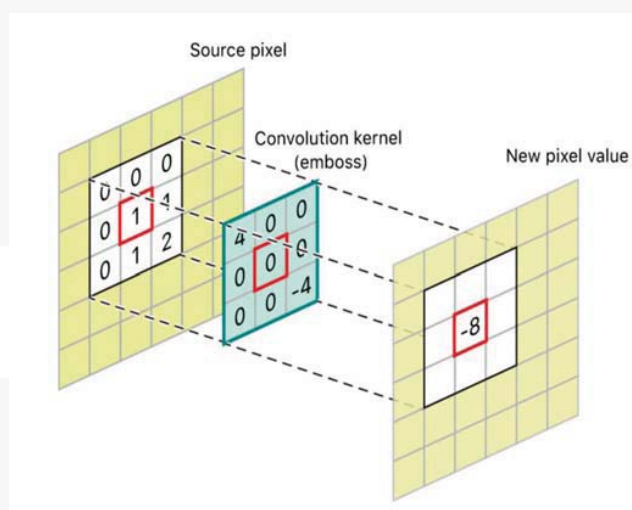


□ Convolution

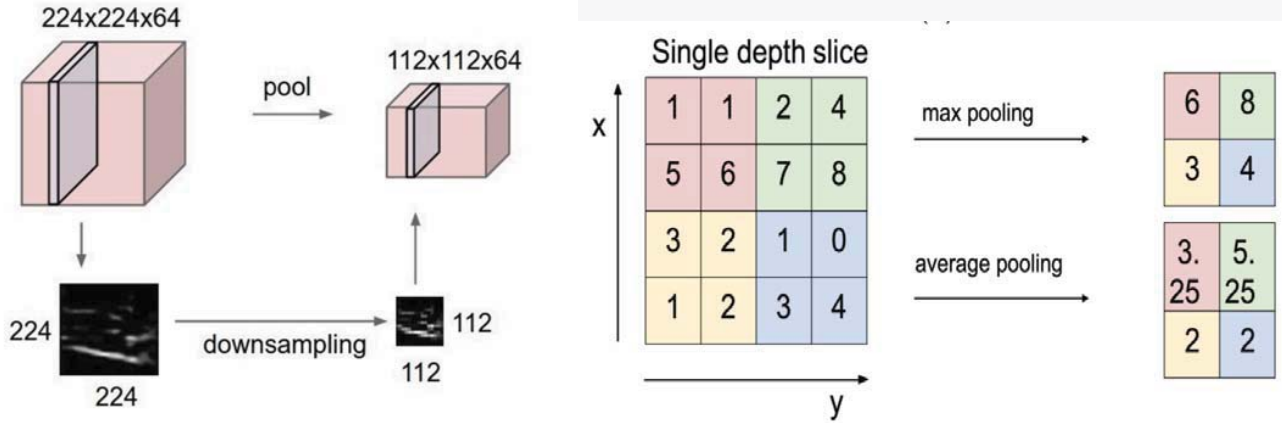
$$y = h * x$$

3x3 filter

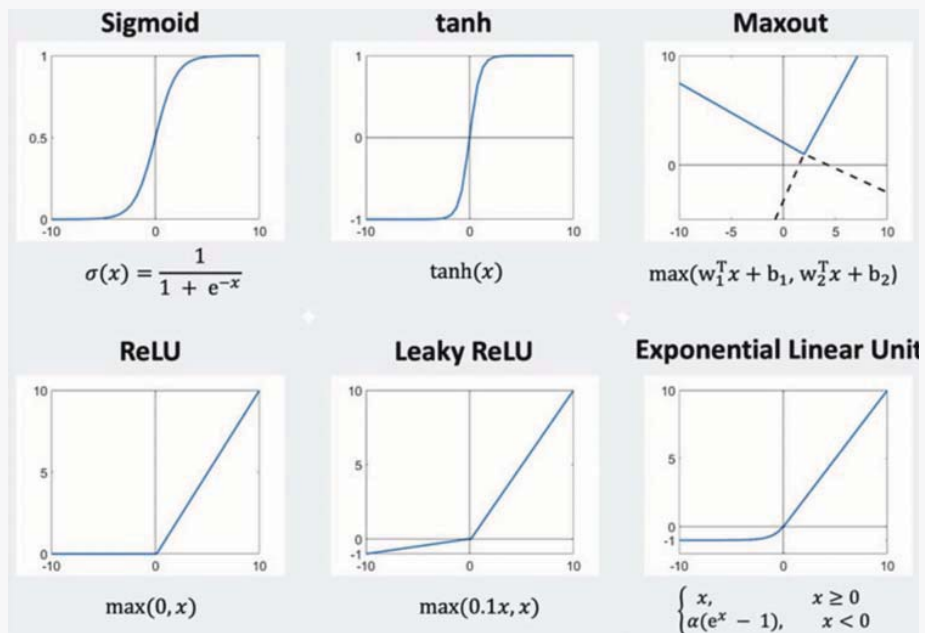
$$y[m, n] = \sum_{p, q=-1}^1 h[p, q] x[m - p, n - q]$$



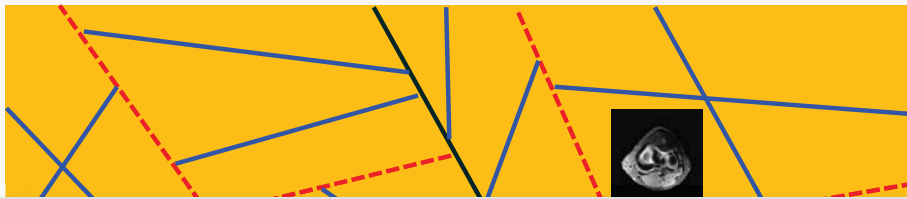
□ Pooling Layer



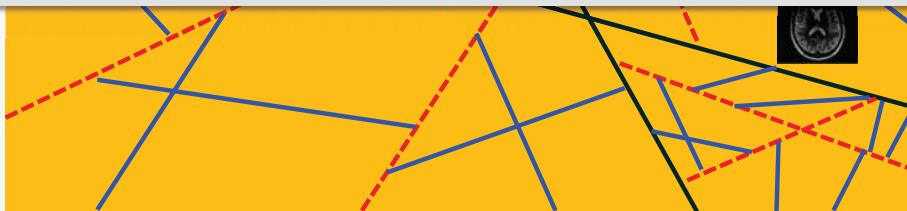
□ Nonlinearity



□ Nonlinearity is the Key for Learning!

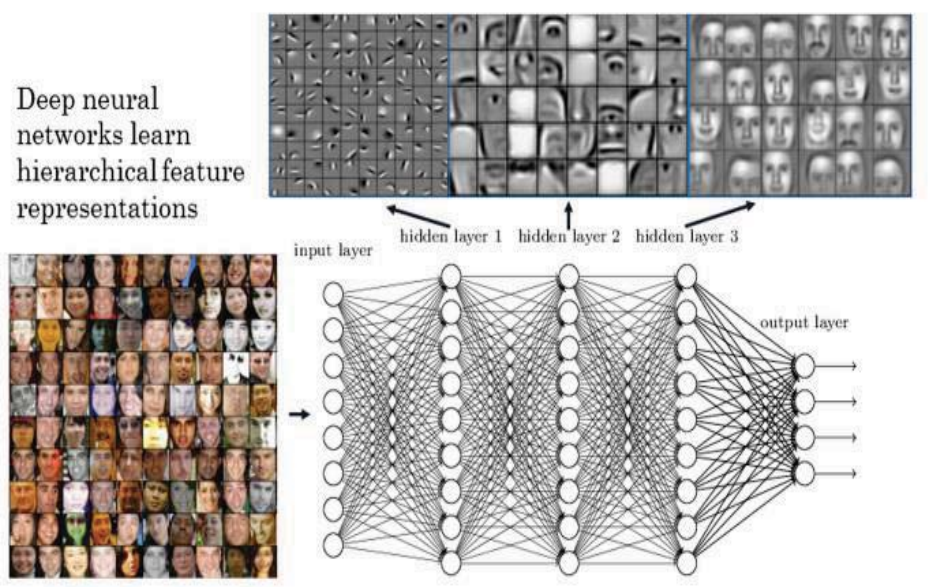


A CNN performs automatic assignment of **distinct linear representation** depending on input



□ Emergence of Hierarchical Features

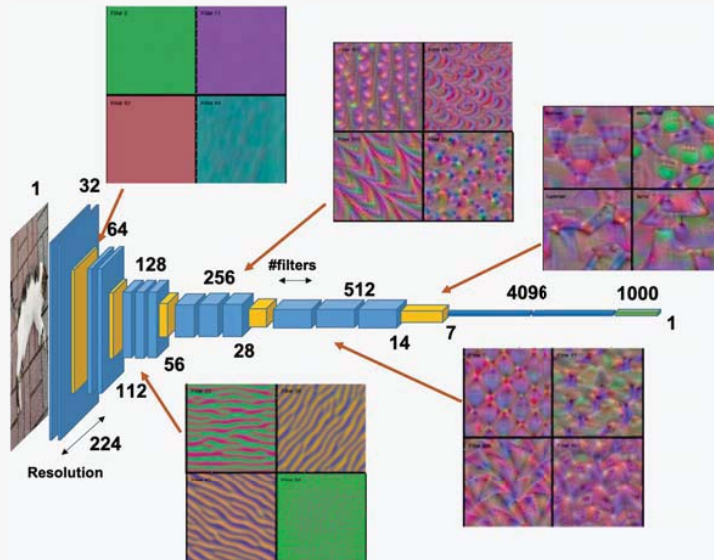
Deep neural networks learn hierarchical feature representations



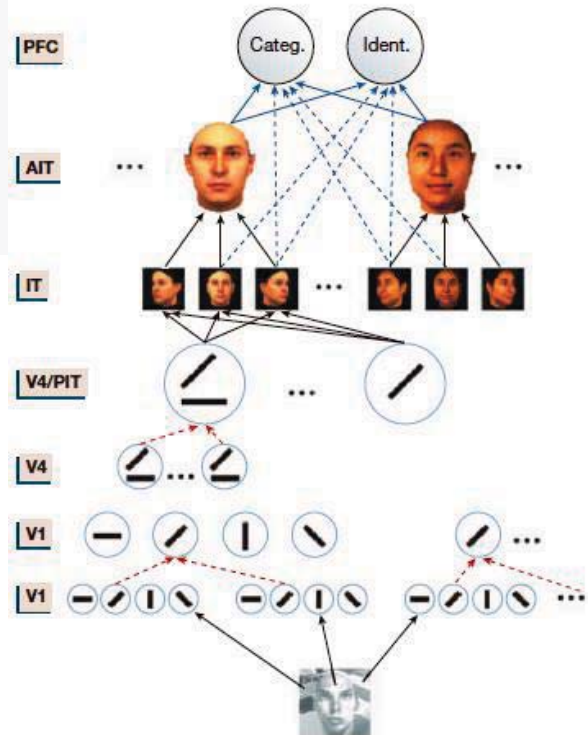
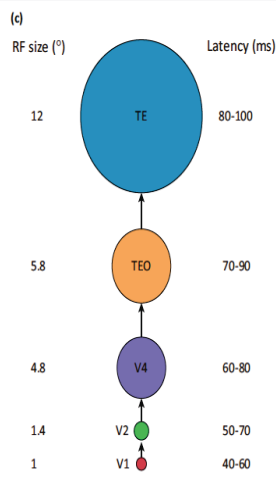
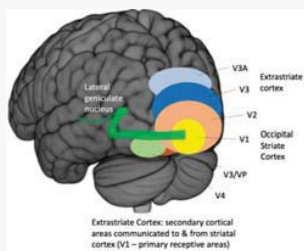
<http://klab.smp.northwestern.edu/wiki/images/4/43/NTM2.pdf>

20

□ Hierarchical Features in VGGNet

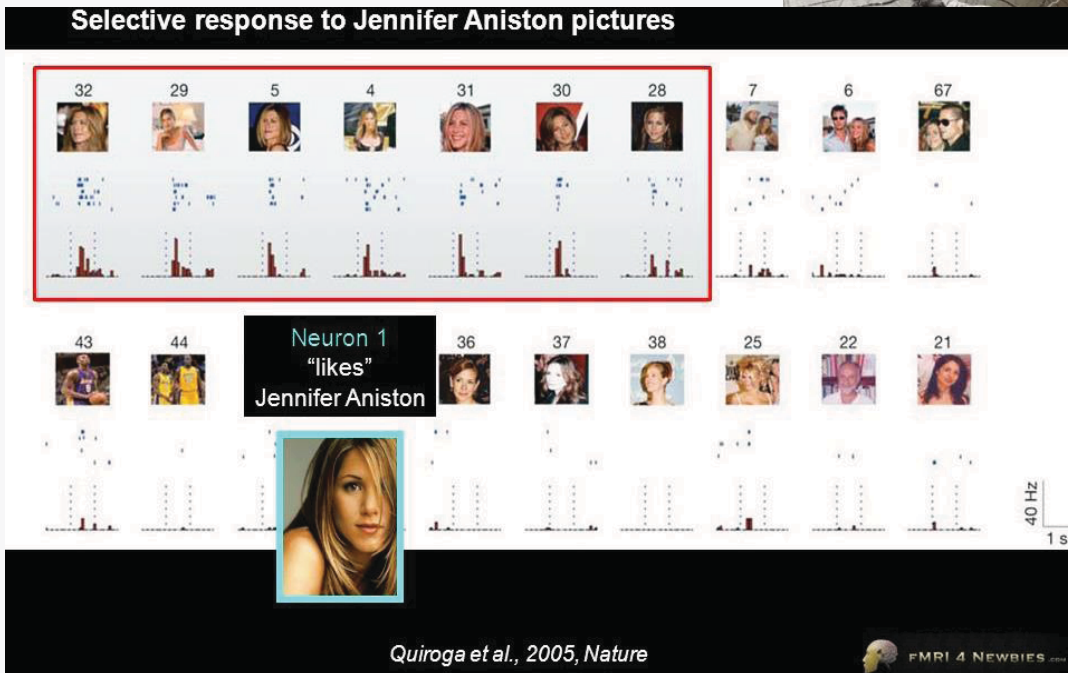


□ Information Processing in Brain



Poggio et al NATURE | VOL 431 | 14 OCTOBER 2004

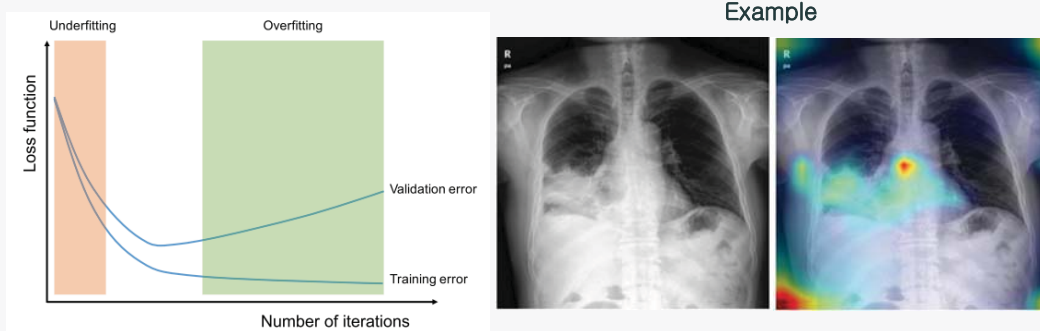
□ “The Jennifer Anniston Cell”



23

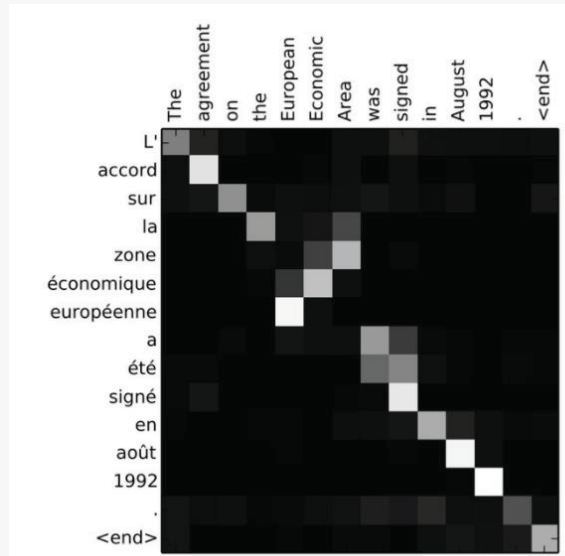
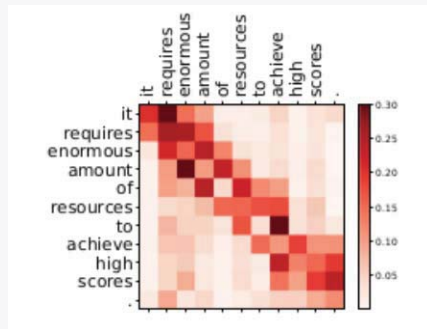
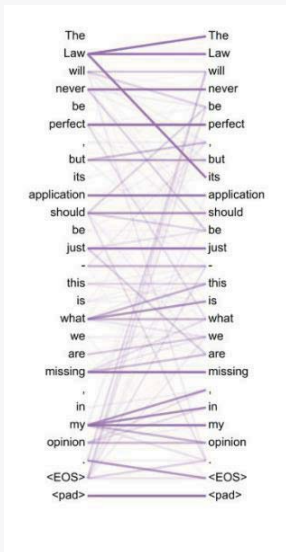
□ Limitation of CNN

- Overfitting: Especially critical in medical imaging

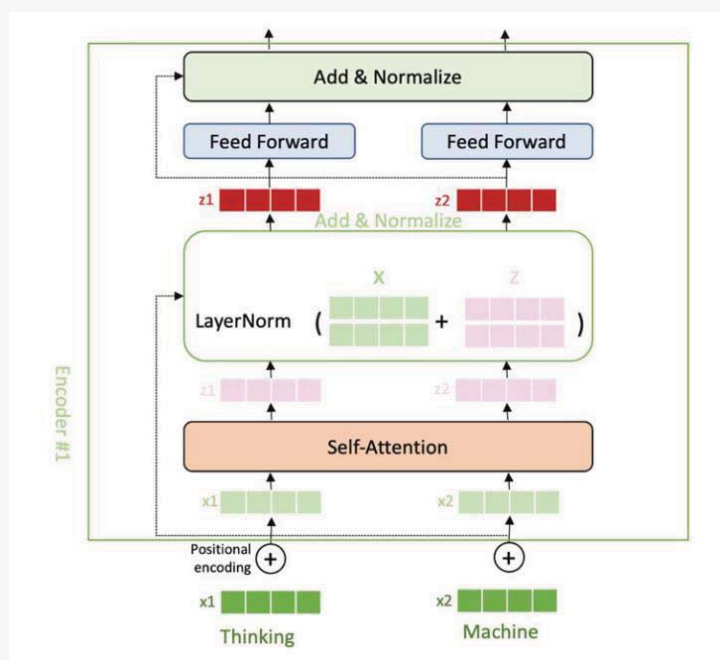


Yamashita et al. Insights imaging (2018)

Transformer: Attention is All You Need

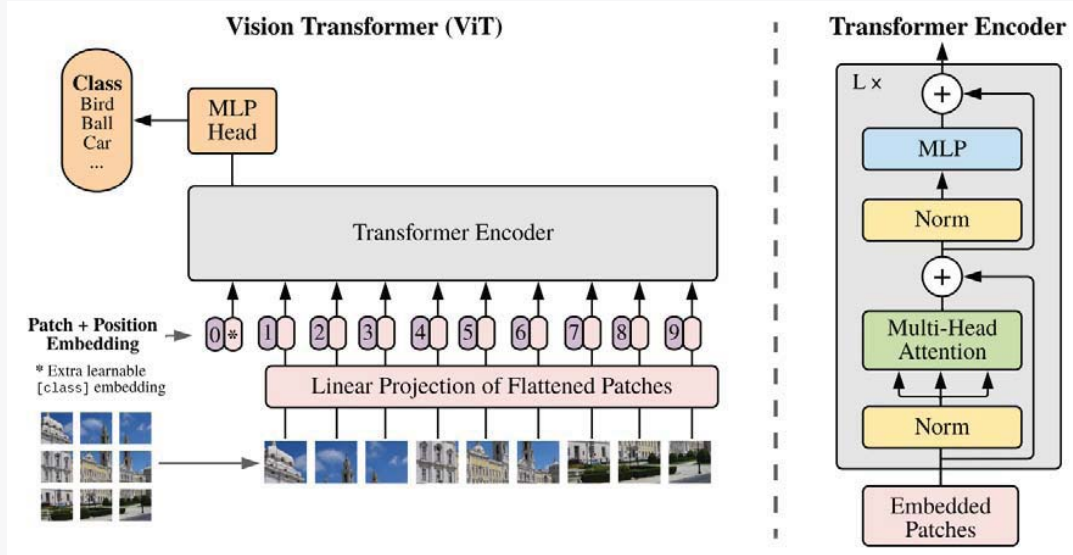


Transformer Architecture



□ ViT: Vision Transformer : Farewell to convolution

- First successful approach to introduce **pure Transformer** to vision.

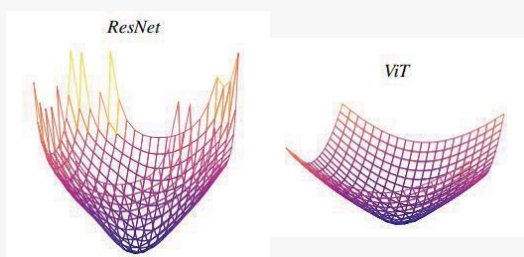


A Dosovitskiy, "an image worth 16 x 16 words for image recognition at scale"

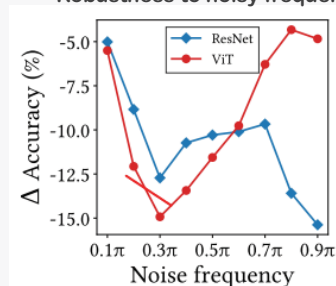
□ Why ViT works better than CNN?

- ViT can model long-range dependency between pixels.
- ViT has more flat loss landscape than CNN (less overfitting).
- ViT is less vulnerable to high frequency noise than CNN.
- ViT is more shape-biased than CNN, like human (**what we want!**).

Loss landscape visualization



Robustness to noisy frequency

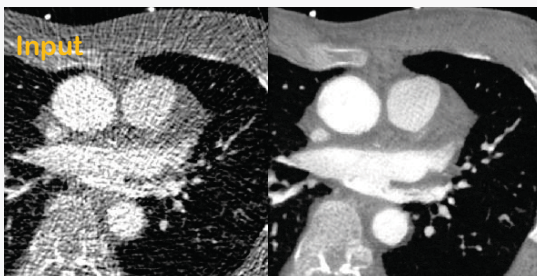
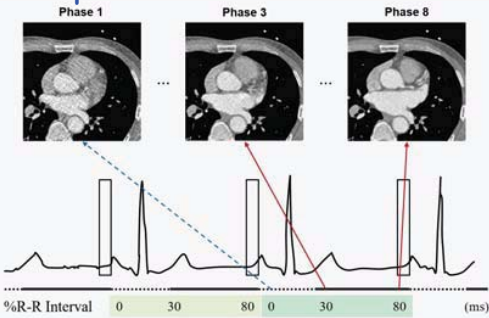


Park et al, ICLR (2022)

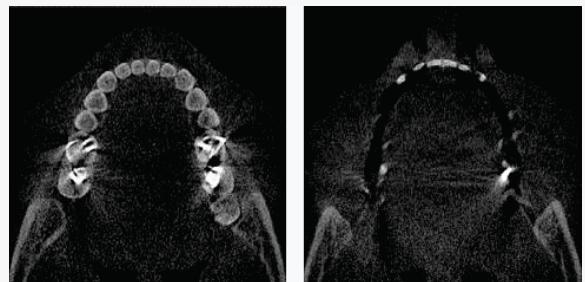
Self-Supervised Learning

□ Limitation of Supervised Learning in Medical AI

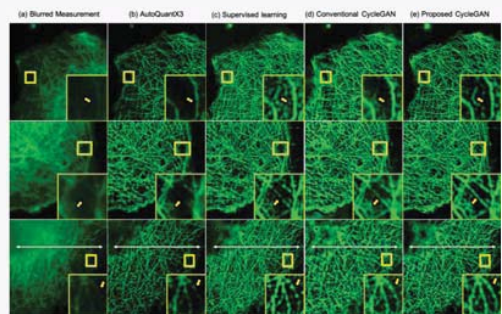
Multiphase Cardiac CT



Metal artifact removal



Blind deconvolution



□ Yann LeCun's Cake Analogy

■ "Pure" Reinforcement Learning (cherry)

- ▶ The machine predicts a scalar reward given once in a while.

- ▶ **A few bits for some samples**

■ Supervised Learning (icing)

- ▶ The machine predicts a category or a few numbers for each input
- ▶ Predicting human-supplied data
- ▶ **10→10,000 bits per sample**

■ Unsupervised/Predictive Learning (cake)

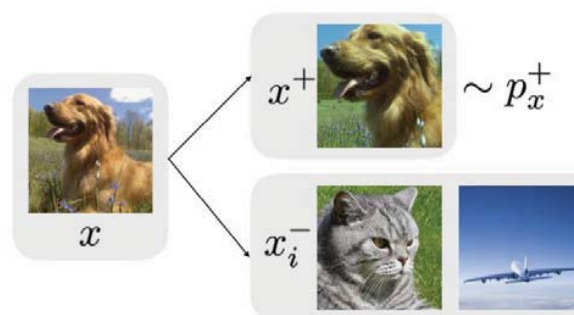
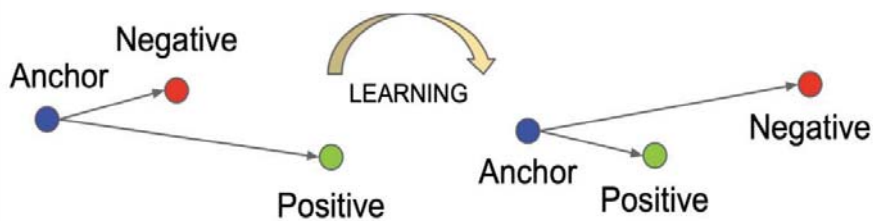
- ▶ The machine predicts any part of its input for any observed part.
- ▶ Predicts future frames in videos
- ▶ **Millions of bits per sample**



- (Yes, I know, this picture is slightly offensive to RL folks. But I'll make it up)

Slide courtesy of Yann LeCun's ICIP 2019 talk

□ Contrastive Learning



□ Contrastive Loss

$$l_{i,j} = -\log \frac{\exp(\text{sim}(\mathbf{z}_i, \mathbf{z}_j) / \tau)}{\sum_{k=1}^{2N} \mathbb{1}_{[k \neq i]} \exp(\text{sim}(\mathbf{z}_i, \mathbf{z}_k) / \tau)}$$

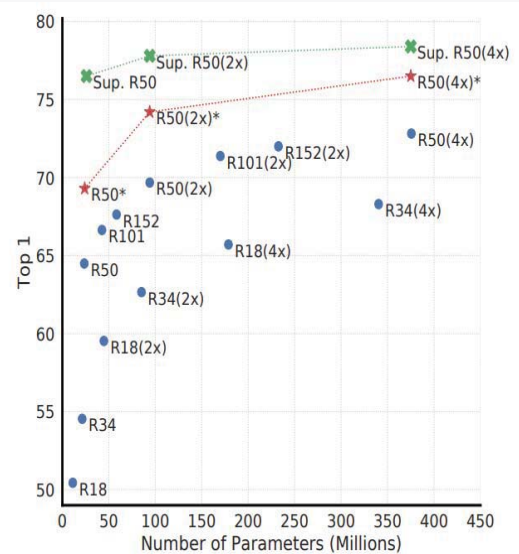
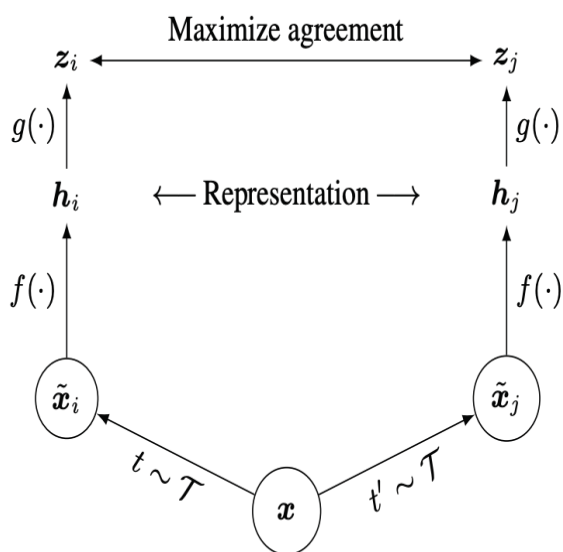
Anchor

positive samples

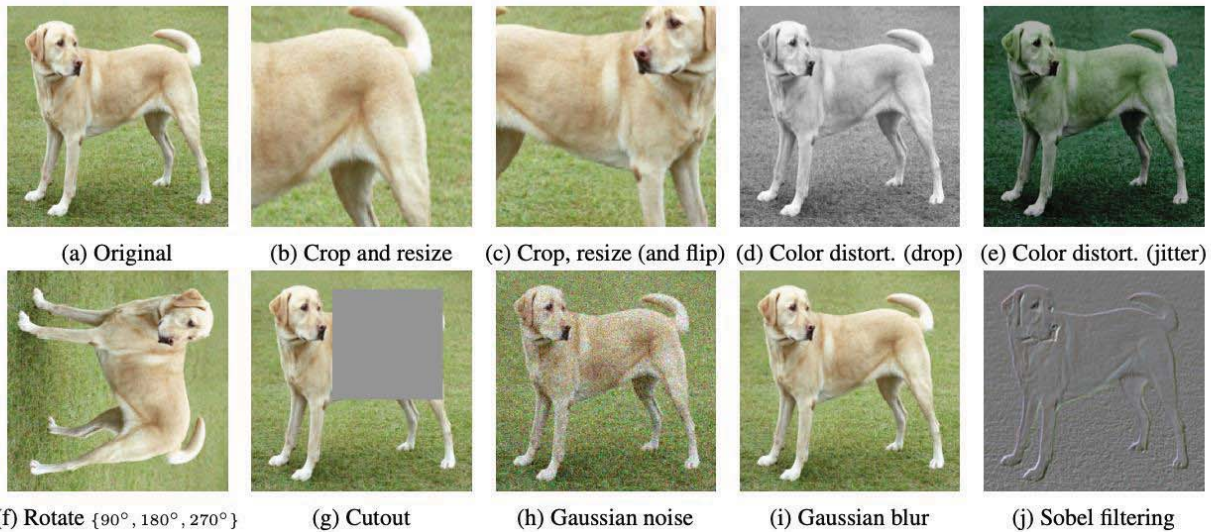
Negative samples

$$l(\mathbf{v}, \mathbf{v}^+, \mathbf{v}^-) = -\log \left[\frac{\exp(\mathbf{v} \cdot \mathbf{v}^+ / \tau)}{\exp(\mathbf{v} \cdot \mathbf{v}^+ / \tau) + \sum_{n=1}^N \exp(\mathbf{v} \cdot \mathbf{v}_n^- / \tau)} \right]$$

□ SimCLR (Chen et al, ICML 2020)



□ Data Augmentation



□ Self-supervised learning with distillation with no label (DINO)

Emerging Properties in Self-Supervised Vision Transformers

Mathilde Caron^{1,2} Hugo Touvron^{1,3} Ishan Misra¹ Hervé Jegou¹
Julien Mairal² Piotr Bojanowski¹ Armand Joulin¹

¹ Facebook AI Research ² Inria* ³ Sorbonne University



Figure 1: Self-attention from a Vision Transformer with 8×8 patches trained with no supervision. We look at the self-attention of the [CLS] token on the heads of the last layer. This token is not attached to any label nor supervision. These maps show that the model automatically learns class-specific features leading to unsupervised object segmentations.

□ Contrastive Prediction in DINO

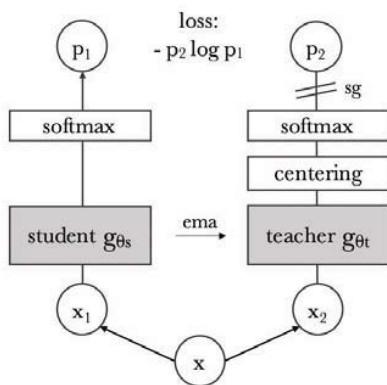
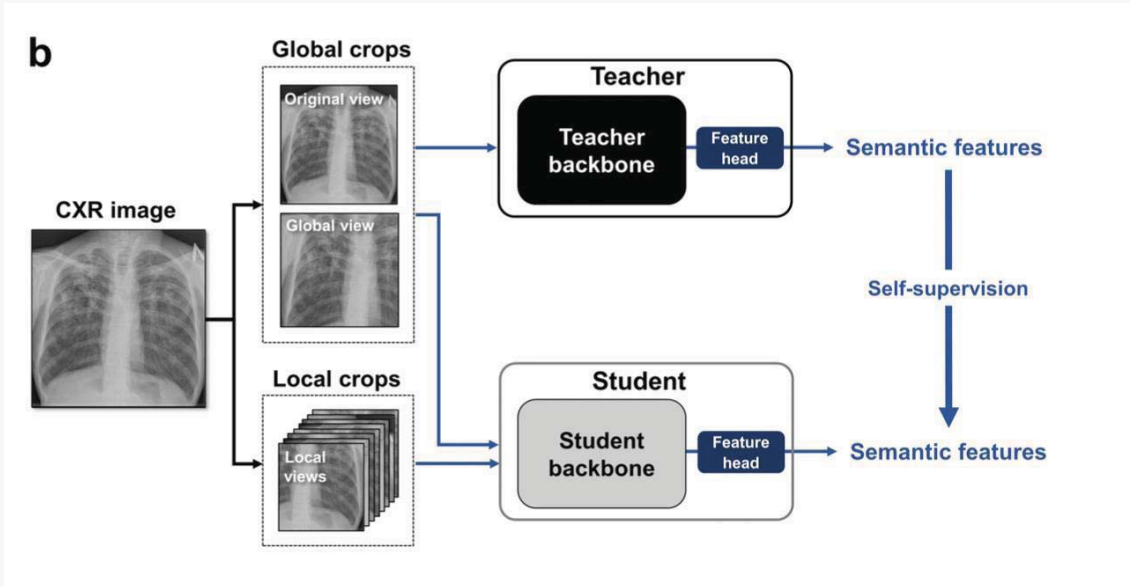
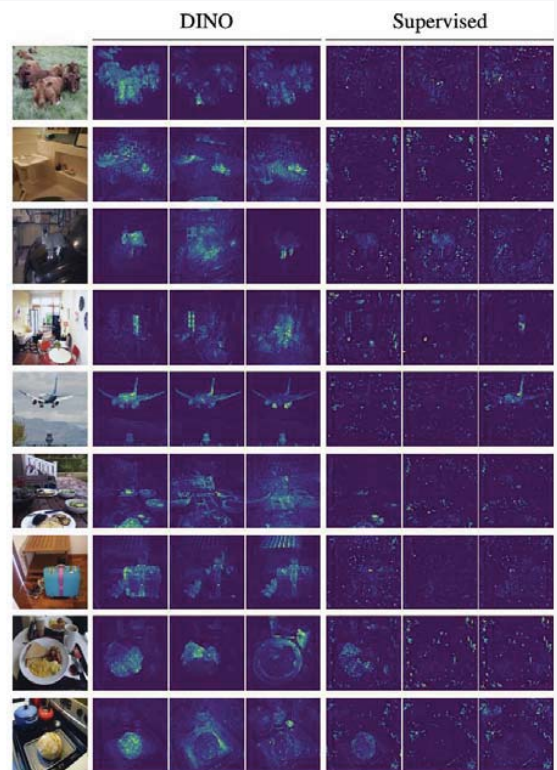
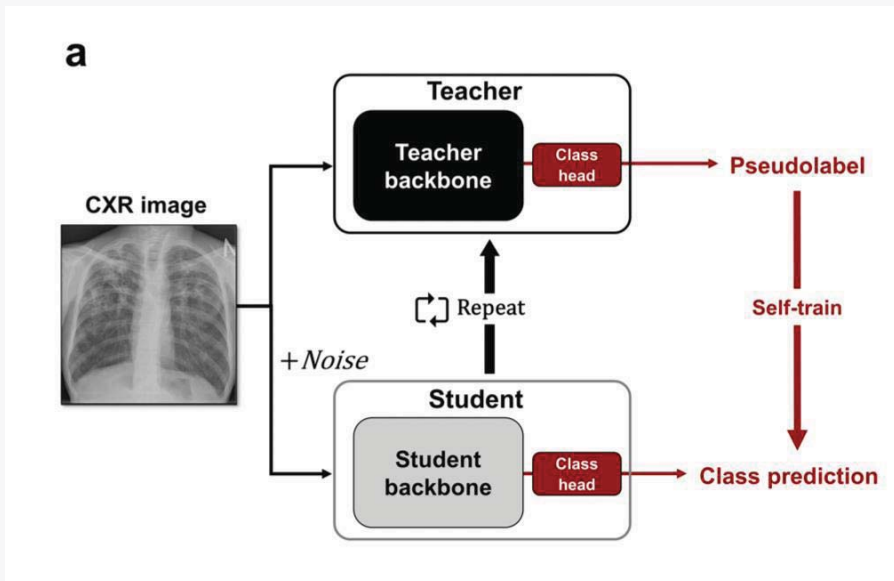


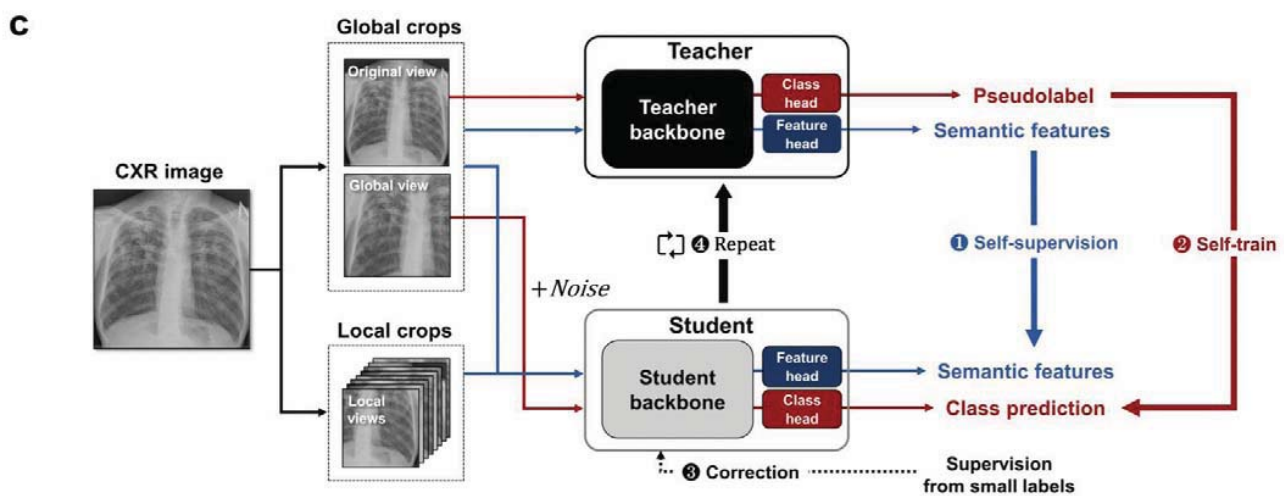
Figure 2: Self-distillation with no labels. We illustrate DINO in the case of one single pair of views (x_1, x_2) for simplicity. The model passes two different random transformations of an input image to the student and teacher networks. Both networks have the same architecture but different parameters. The output of the teacher network is centered with a mean computed over the batch. Each networks outputs a K dimensional feature that is normalized with a temperature softmax over the feature dimension. Their similarity is then measured with a cross-entropy loss. We apply a stop-gradient (sg) operator on the teacher to propagate gradients only through the student. The teacher parameters are updated with an exponential moving average (ema) of the student parameters.



□ Self-Training with Noisy Student

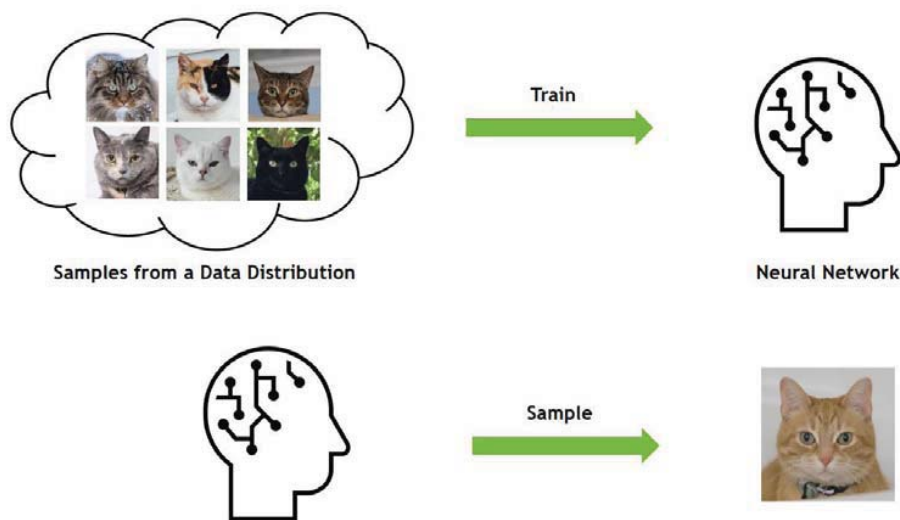


□ Distillation for self-supervised and self-train learning (DISTL): Park et al, Nature Comm 2022



Generative Models

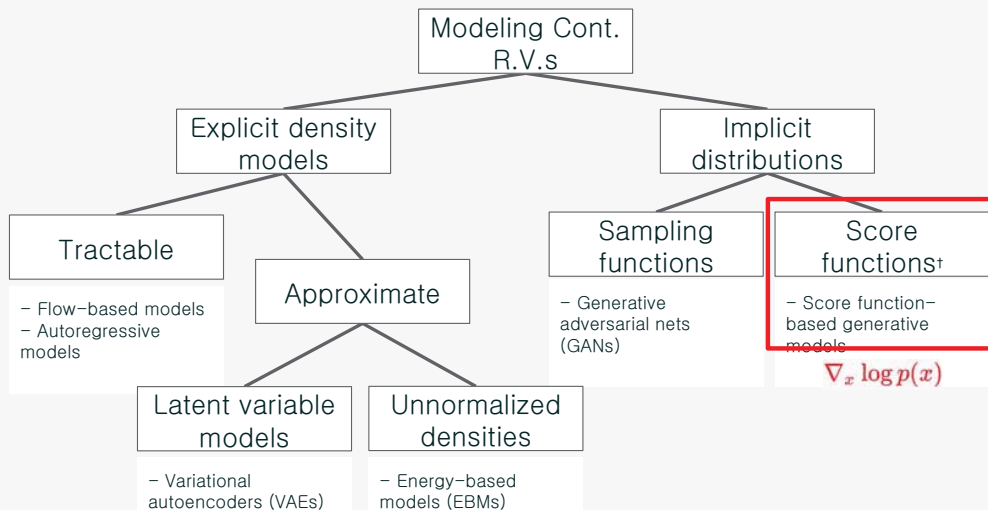
□ Deep Generative Learning



Slide from CVPR 2022 Diffusion Model Tutorials

□ Various Way of Modeling Continuous Variables

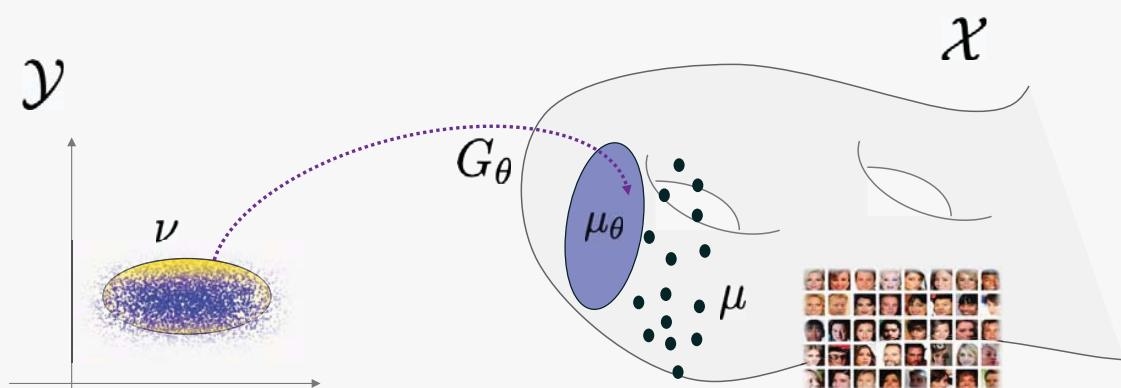
- Various ways to model continuous random variables
 - This taxonomic tree doesn't count on "training methods"



*copied and modified from I. Goodfellow,

2016. Here, we refer to score function as the log-density gradient wrt. input, instead of well-known definition in statistics

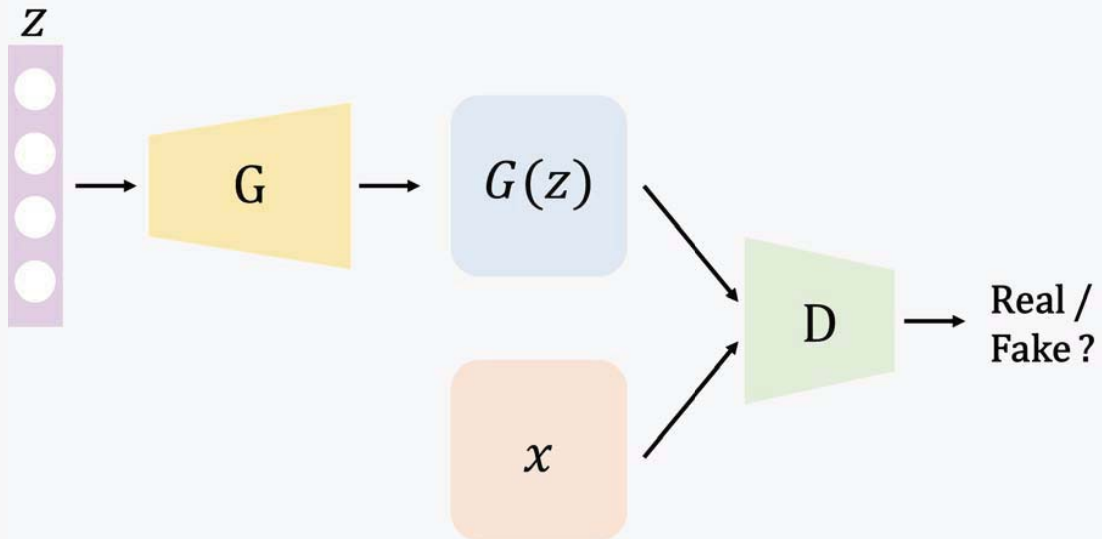
□ Generative Adversarial Nets (GAN)



$$\min_{\theta} \text{dist}(\mu_{\theta}, \mu)$$

$$\text{subject to } \mu_{\theta} = G_{\theta\#}\nu$$

□ Generative Adversarial Nets (GAN)



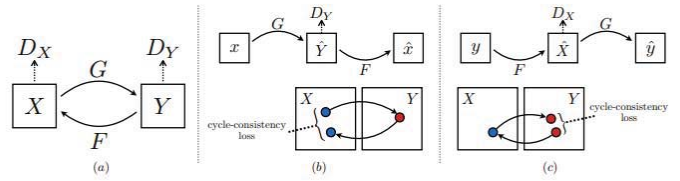
□ Generative Adversarial Nets (GAN)

Generative Adversarial Network (GAN)

<https://this-person-does-not-exist.com/en>

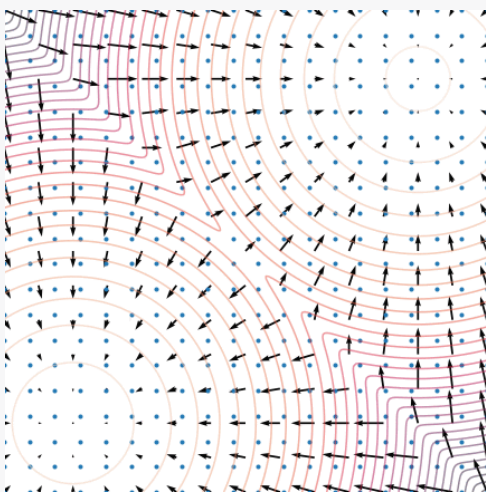


CycleGAN



J.-Y. Zhu, et al, CVPR, 2017

□ Diffusion-based Generative Models



- Once the score model is trained to optimality,
 - i.e. $s_\theta(\mathbf{x}) \approx \nabla_{\mathbf{x}} \log p(\mathbf{x})$
- Example: Use **Langevin dynamics** to draw samples

$$\mathbf{x}_{i+1} \leftarrow \mathbf{x}_i + \epsilon \nabla_{\mathbf{x}} \log p(\mathbf{x}) + \sqrt{2\epsilon} \mathbf{z}_i$$

$$i = 0, 1, \dots, K$$

<http://yang-song.github.io/blog/2021/score/>

□ Diffusion-based Generative Models



StyleGAN2-ADA
(Karras et al., 2020)



DDPM
(Ho et al., 2020)

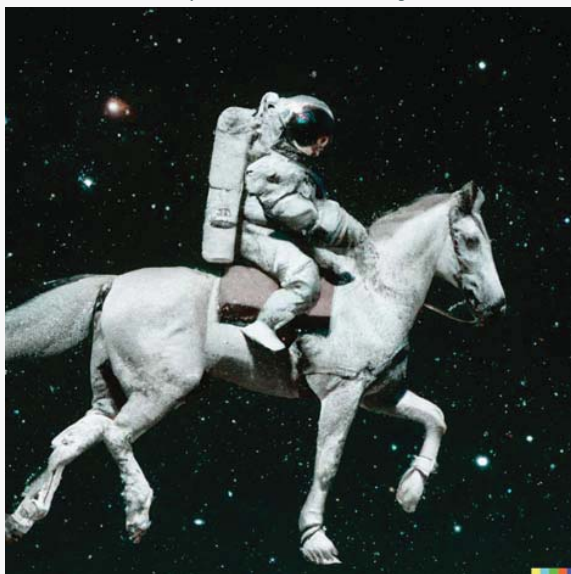


Reverse SDE
(Song et al., 2020)

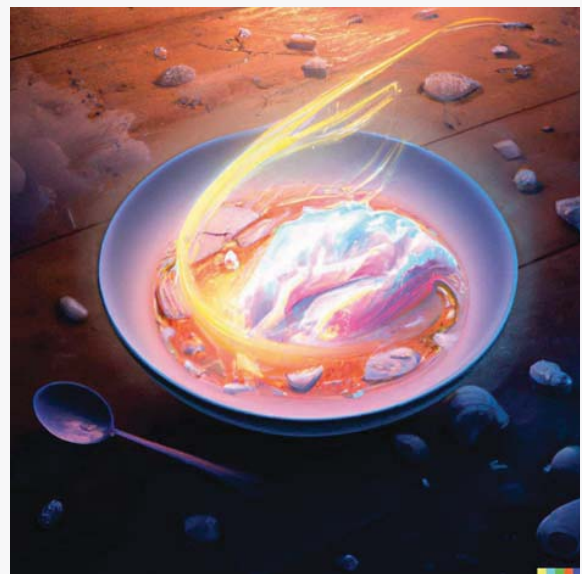
□ Text-Guided Image Generation

<https://openai.com/dall-e-2/>

“ An astronaut riding a horse
in a photorealistic style”

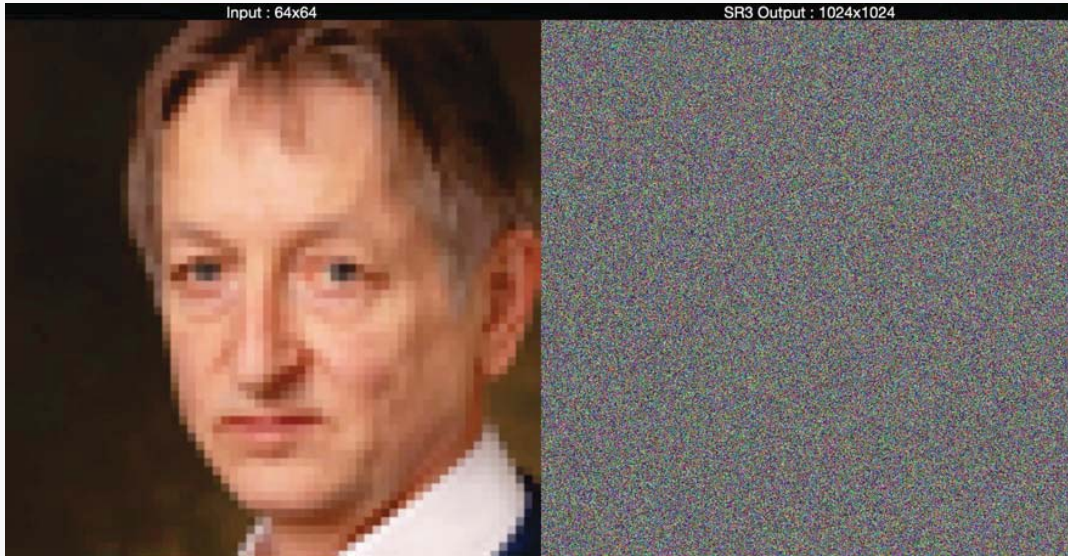


“ A bowl of soup that is a
portal to another dimension
as digital art”



□ Image Super-Resolution

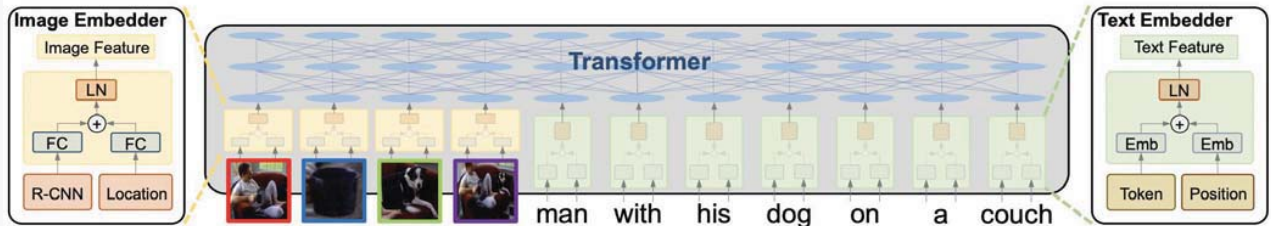
Conditional generation (SR)



<http://yang-song.github.io/blog/2021/score/>

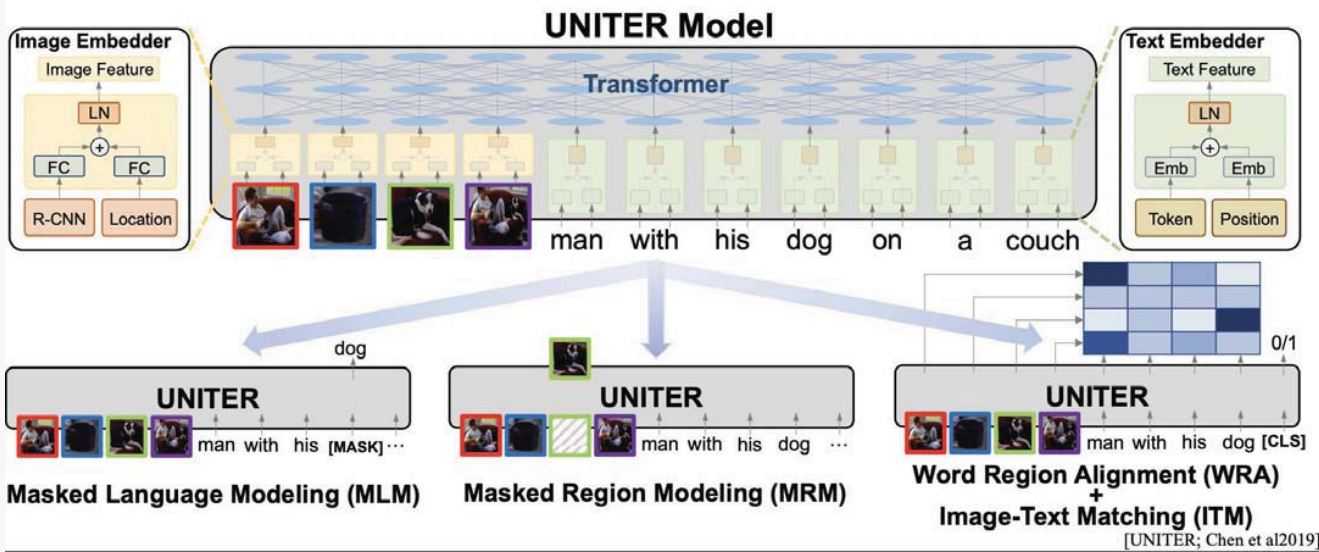
Vision-Language Pretraining

Single Stream Architecture: UNITER (Chen et al 2019)



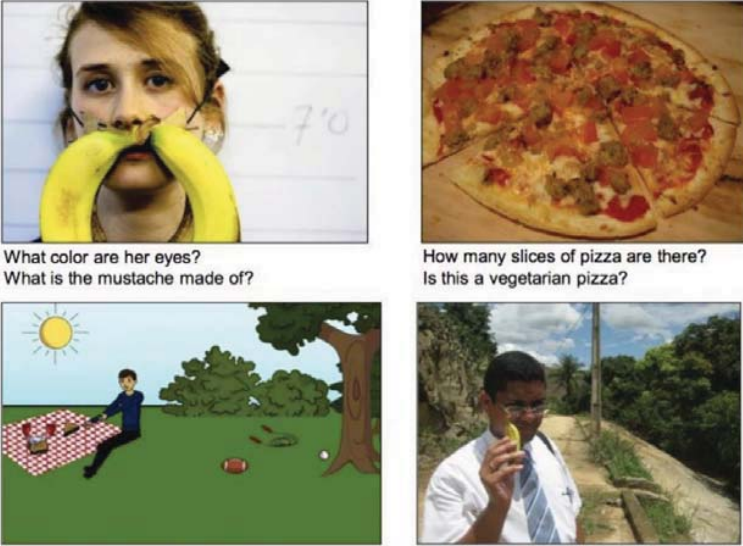
Slide courtesy from CVPR2022 tutorial

Pretraining



Slide courtesy from CVPR2022 tutorial

□ Downstream Tasks: Visual Question Answering



What color are her eyes?
What is the mustache made of?

How many slices of pizza are there?
Is this a vegetarian pizza?

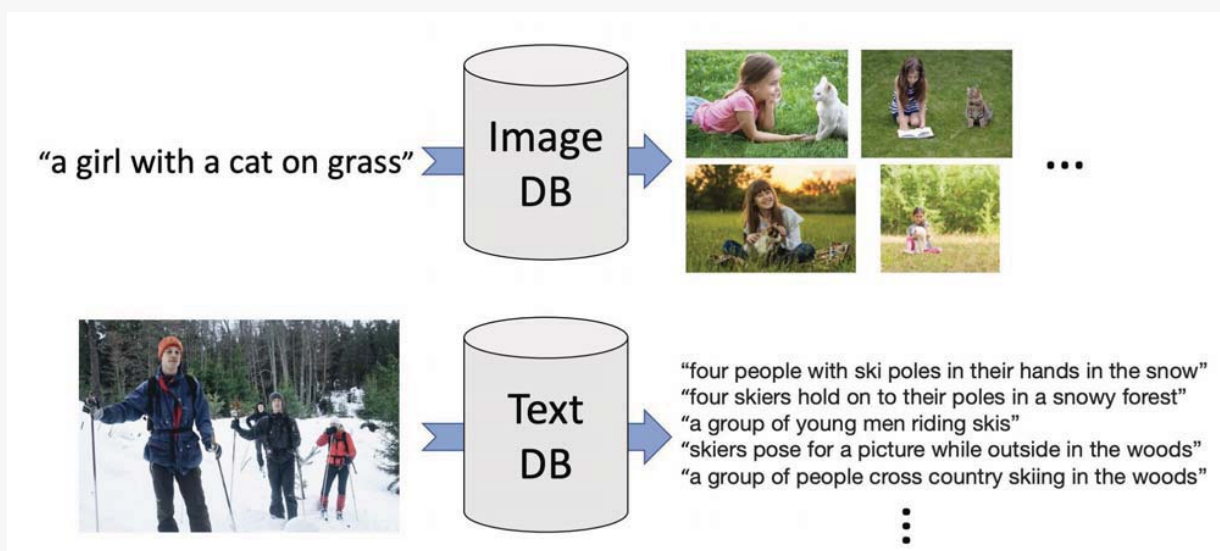
Is this person expecting company?
What is just under the tree?

Does it appear to be rainy?
Does this person have 20/20 vision?

[Antol et al., ICCV 2015]

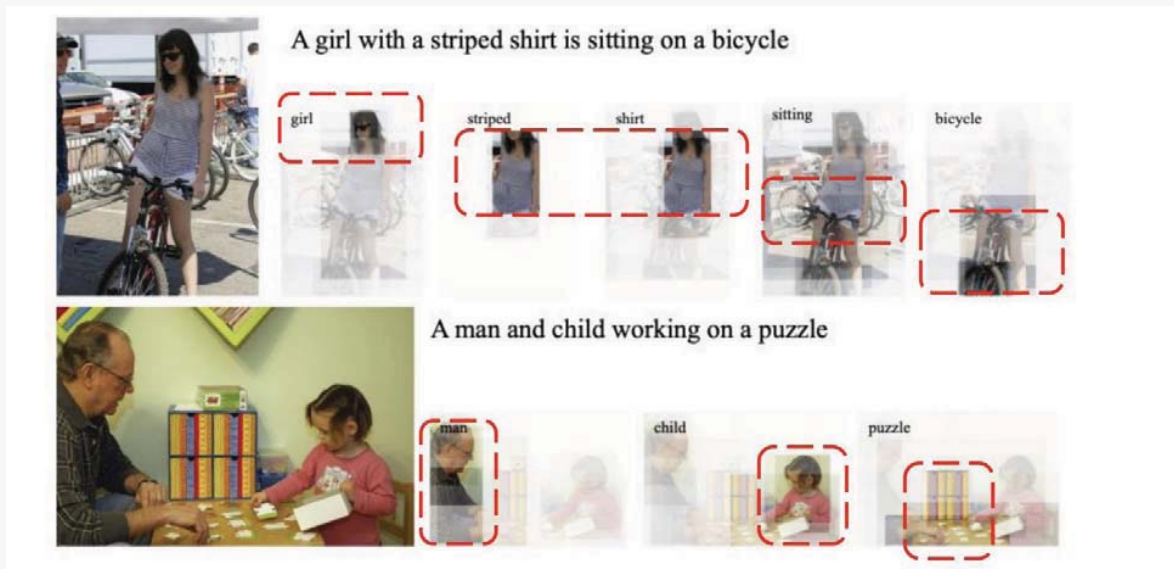
Slide courtesy from CVPR2022 tutorial

□ Downstream Tasks: Image-Text Retrieval



Slide courtesy from CVPR2022 tutorial

□ Downstream Tasks: Text-to-Image Attention



Slide courtesy from CVPR2022 tutorial

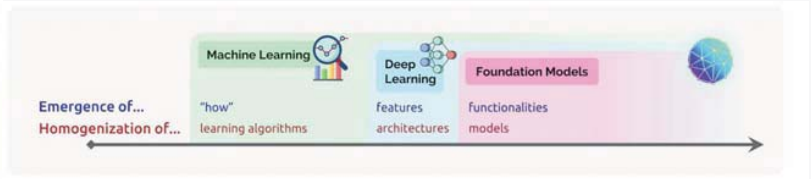
Foundation Models

□ Foundation Models: A Future AI?

On the Opportunities and Risks of Foundation Models

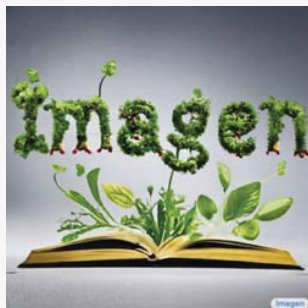
Rishi Bommasani* Drew A. Hudson Ehsan Adeli Russ Altman Simran Arora
 Sydney von Arx Michael S. Bernstein Jeannette Bohg Antoine Bosselut Emma Brunskill
 Erik Brynjolfsson Shyamal Buch Dallas Card Rodrigo Castellon Niladri Chatterji
 Annie Chen Kathleen Creel Jared Quincy Davis Dorottya Demszky Chris Donahue
 Moussa Doumbouya Esin Durmus Stefano Ermon John Etchemendy Kavin Ethayarajah
 Li Fei-Fei Chelsea Finn Trevor Gale Lauren Gillespie Karan Goel Noah Goodman
 Shelby Grossman Neel Guha Tatsunori Hashimoto Peter Henderson John Hewitt
 Daniel E. Ho Jenny Hong Kyle Hsu Jing Huang Thomas Icard Saahil Jain
 Dan Jurafsky Pratyusha Kalluri Siddharth Karamcheti Geoff Keeling Fereshte Khani
 Omar Khattab Pang Wei Koh Mark Krass Ranjay Krishna Robith Kudipudi
 Ananya Kumar Faisal Ladhak Mima Lee Tony Lee Jure Leskovec Isabelle Levent
 Xiang Lisa Li Xuechen Li Tengyu Ma Ali Malik Christopher D. Manning
 Suvir Mirchandani Eric Mitchell Zanele Mnyikwa Suraj Nair Avanika Narayan
 Deepak Narayanan Ben Newman Allen Nie Juan Carlos Niebles Hamed Nilforoshan
 Julian Nyarko Giray Ogut Laurel Orr Isabel Papadimitriou Joon Sung Park Chris Piech
 Eva Portelance Christopher Potts Aditi Raghunathan Rob Reich Hongyu Ren
 Frieda Rong Yusuf Roohani Camilo Ruiz Jack Ryan Christopher Ré Dorsa Sadigh
 Shiori Sagawa Keshav Santhanam Andy Shih Krishnan Srinivasan Alex Tamkin
 Rohan Taori Armin W. Thomas Florian Tramèr Rose E. Wang William Wang Bohan Wu
 Jiajun Wu Yuhuai Wu Sang Michael Xie Michihiro Yasunaga Jiaxuan You Matei Zaharia
 Michael Zhang Tianyi Zhang Xikun Zhang Yuhui Zhang Lucia Zheng Kaitlyn Zhou
 Percy Liang*¹

Center for Research on Foundation Models (CREM)
 Stanford Institute for Human-Centered Artificial Intelligence (HAI)
 Stanford University

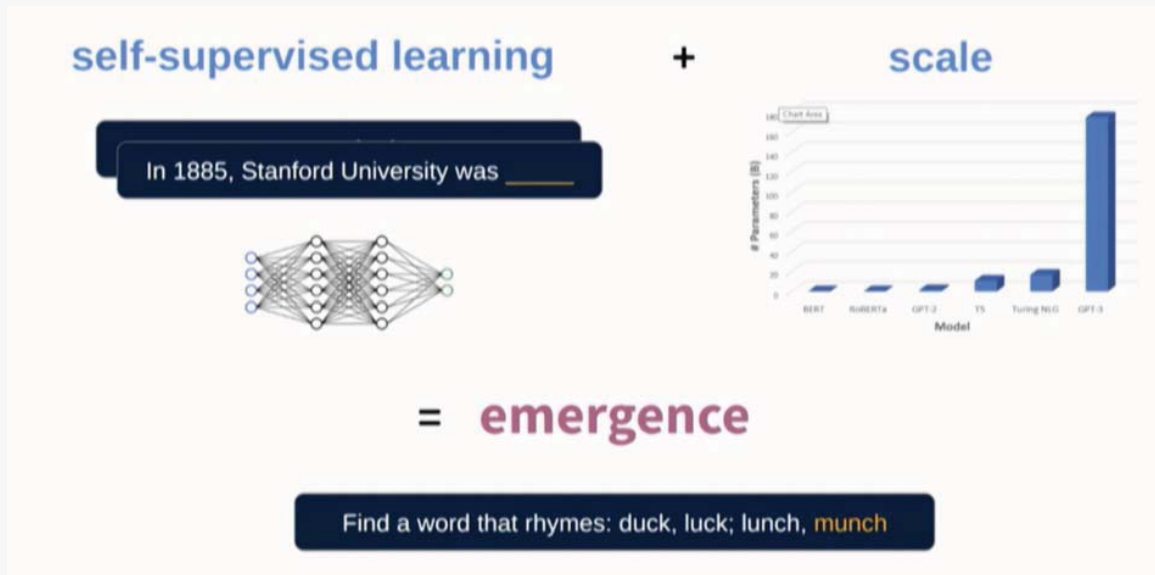


Bommasani, Rishi, et al. *arXiv:2108.07258* (2021).

□ Examples of Foundation Models

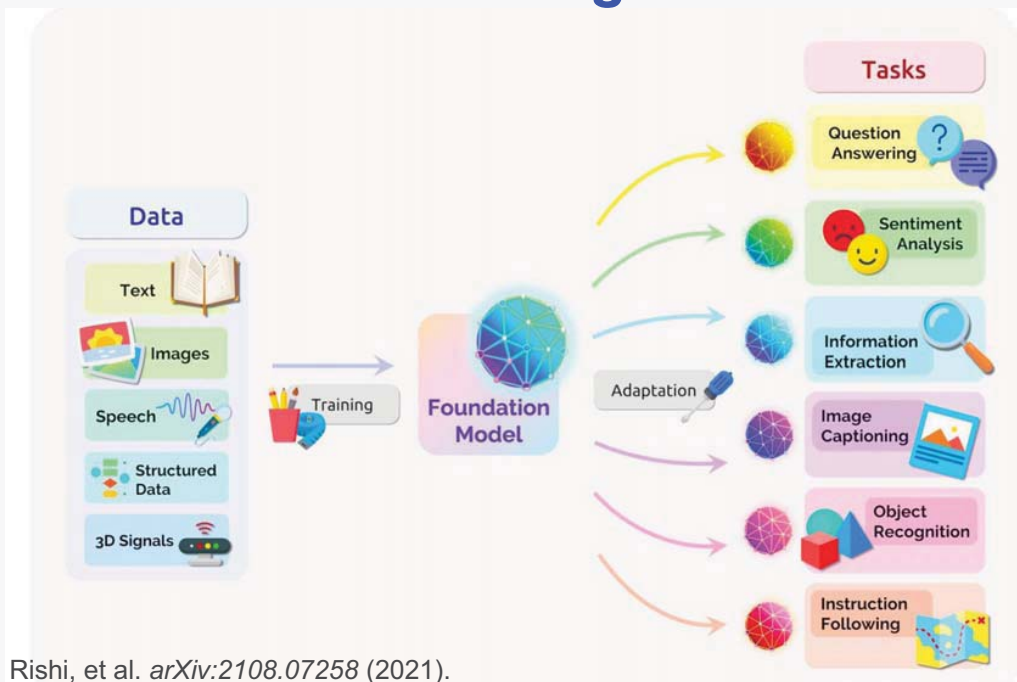


□ Foundation Models: Emergence



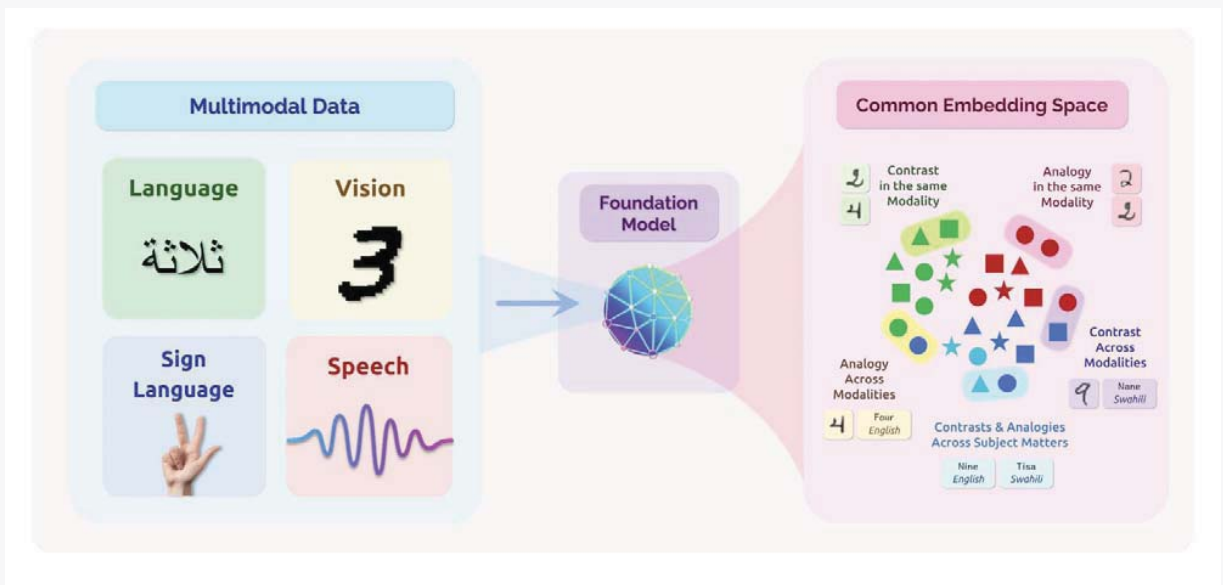
Slide courtesy of Percy Liang @ Foundation Model Workshop, 2021

□ Foundation Models: Homogenization



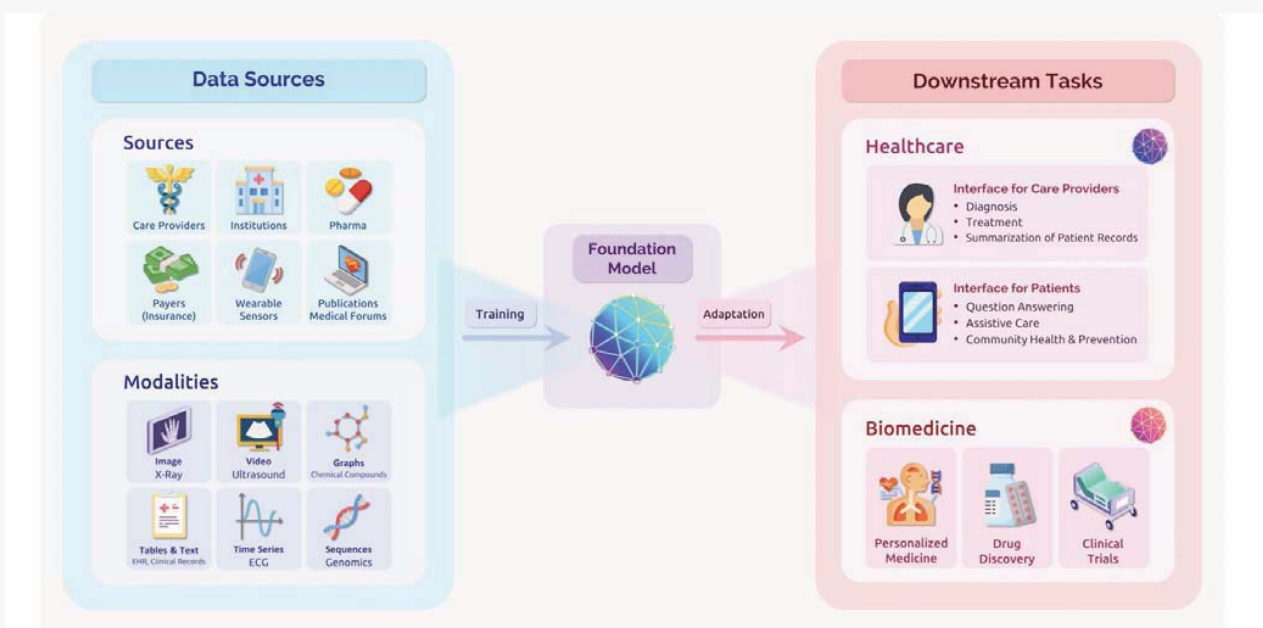
Bommasani, Rishi, et al. *arXiv:2108.07258* (2021).

□ Multimodal Embedding



Bommasani, Rishi, et al. *arXiv:2108.07258* (2021).

□ Foundation Models for Healthcare



Bommasani, Rishi, et al. *arXiv:2108.07258* (2021).

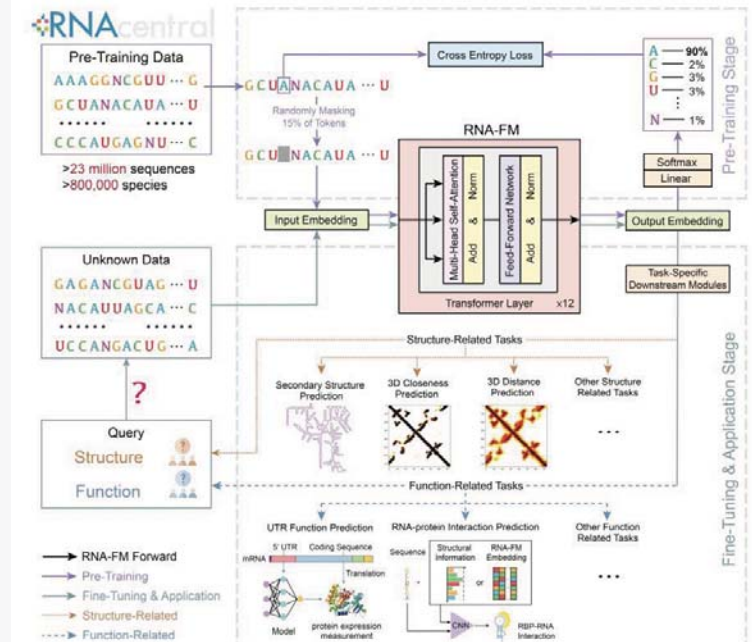
□ Foundation Models for Genomics

Interpretable RNA Foundation Model from Unannotated Data for Highly Accurate RNA Structure and Function Predictions

Jiayang Chen¹, Zhihang Hu¹, Siqi Sun^{*1,2,3}, Qingxiang Tan¹, Yixuan Wang^{1,4}, Qinze Yu^{1,5}, Licheng Zong¹, Liang Hong¹, Jin Xiao¹, Tao Shen^{1,9}, Irwin King¹, and Yu Li^{*1,6,7,8,10}

- ¹Department of Computer Science and Engineering, The Chinese University of Hong Kong, Hong Kong SAR, China
- ²Research Institute of Intelligent Complex Systems, Fudan University, Shanghai, China
- ³Shanghai AI Laboratory, Shanghai, China
- ⁴Harbin Institute of Technology, China
- ⁵University of Electronic Science and Technology of China, China
- ⁶Institute for Medical Engineering and Science, Massachusetts Institute of Technology, Cambridge, MA, USA
- ⁷Wyss Institute for Biologically Inspired Engineering, Harvard University, Boston, MA, USA
- ⁸Broad Institute of MIT and Harvard, Cambridge, MA, USA
- ⁹Zelixir Biotech, Shanghai, China
- ¹⁰The CUHK Shenzhen Research Institute, Hi-Tech Park, Nanshan, Shenzhen, China

Chen et al, doi: <https://doi.org/10.1101/2022.08.06.503062>



Federated Learning

□ Big Data and Data Privacy Issue

데이터 3법: 2020년 2월 4일 공포, 8월 5일 시행

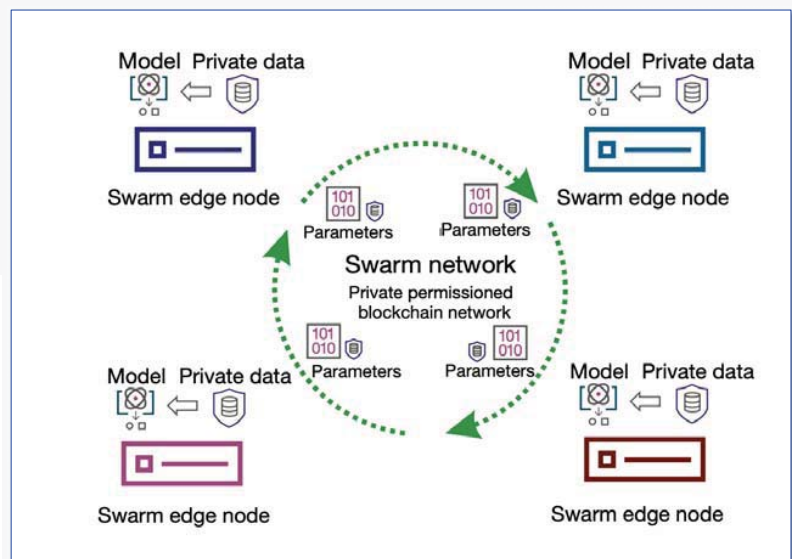
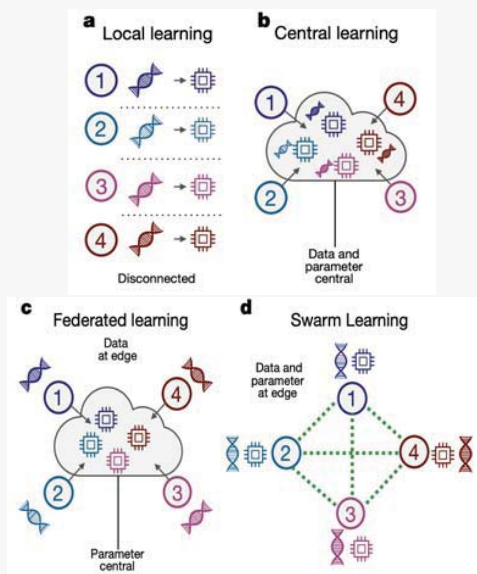
- 가명정보 이용: 데이터를 기반으로 하는 새로운 기술·제품·서비스의 개발 등 산업적 목적을 포함하는 과학적 연구, 통계작성, 공익적 기록보존 등의 목적
- 관련 법률의 유사·중복 규정은 「개인정보 보호법」으로 일원화

<표> 개인정보 외의 개념 확대

	개념	활용가능 범위
개인정보	개인을 알아볼 수 있는 정보. 다른 정보와 쉽게 결합해 알아볼 수 있는 정보	수집목적과 합리적으로 연관된 범위 내에서 정보주체 동의 없이 개인정보 추가 이용·제공 가능
가명정보	추가정보의 사용 없이는 특정 개인을 알아볼 수 없게 조치한 정보	다음 목적에 동의 없이 활용 가능 ① 통계작성 ② 과학적 연구 ③ 공익적 기록보존 등
익명정보	다른 정보를 사용하여도 더 이상 개인을 알아볼 수 없게 조치한 정보	개인정보가 아니기 때문에 제한 없이 자유롭게 활용

- Medical image, genome big data
- Even after pseudonymization, there is a potential of privacy infringement
- Challenges in the development of medical artificial intelligence

□ Federated Learning, Swarm Learning



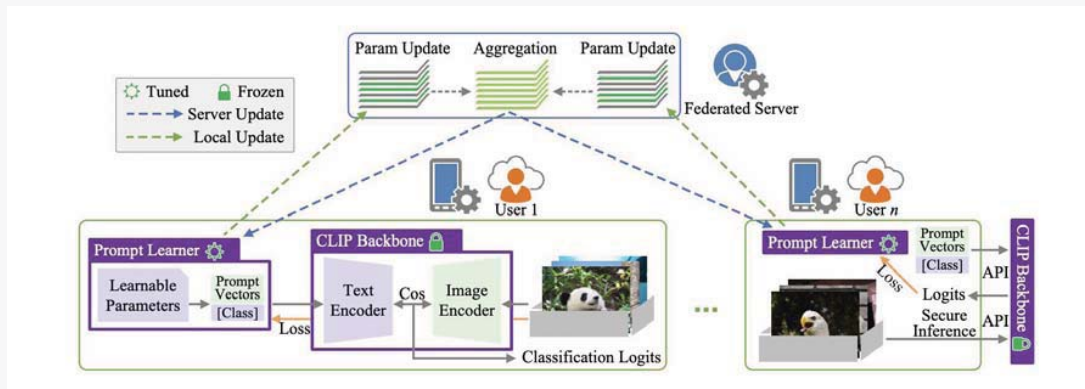
Warnat-Herresthal et al, *Nature* 594, 265–270 (2021)

□ FL in the age of Foundation Model

PROMPTFL: Let Federated Participants Cooperatively Learn Prompts Instead of Models — Federated Learning in Age of Foundation Model

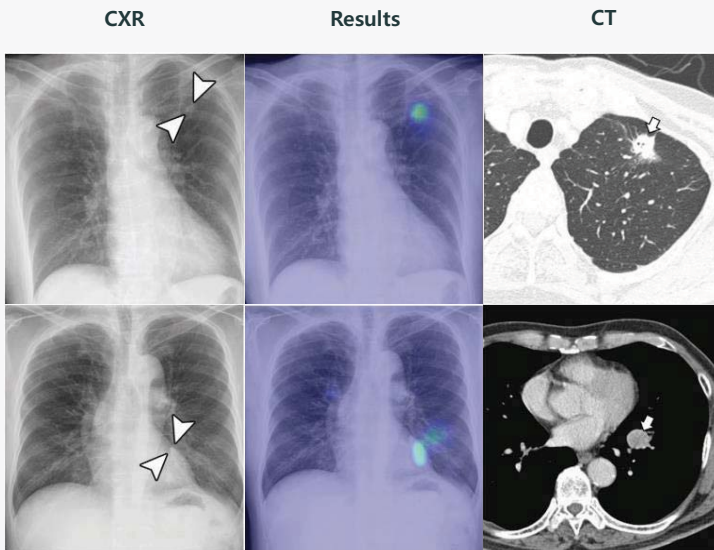
Tao Guo, Song Guo, Junxiao Wang, Wenchao Xu

Department of Computing, The Hong Kong Polytechnic University, Hong Kong, China



AI-driven Diagnosis

□ Lung Cancer Detection (Lunit)



- Development of a deep learning model consisting of a 25-layer convolutional neural network and 8 residual connections
- Learning with 34067 normal X-rays and 9225 malignant pulmonary nodule X-rays obtained at Seoul National University Hospital

Parameter	Radiograph Classification Performance				Nodule Detection Performance		
	AUROC	Sensitivity (%)	Specificity (%)	F1 Score (Precision, Recall)	JAFROC FOM	Sensitivity (%) ^a	Rate of FP Findings (%) ^b
Seoul National University Hospital	0.92	79.0 (94/119) [70.8, 85.4]	95 (59/62) [86.2, 98.9]	87.0 [94.9, 79.0]	0.885	69.9 (100/143) [62.0, 76.9]	0.34 [61/181]
Boramae Hospital	0.99	91.1 (112/123) [84.6, 95.1]	98 (58/59) [80.2, 100]	94.9 [99.1, 91.1]	0.924	82.0 (114/139) [74.7, 87.6]	0.30 [54/182]
National Cancer Center	0.94	71.2 (79/111) [62.1, 78.9]	100 (70/70) [93.8, 100]	83.2 [100, 71.2]	0.831	69.6 (80/115) [60.6, 77.3]	0.02 [3/181]
University of California San Francisco Medical Center	0.96	88 (78/89) [79.0, 93.1]	93 (56/60) [83.6, 97.8]	91.2 [95.1, 87.6]	0.880	75.0 (78/104) [65.8, 82.4]	0.25 [37/149]

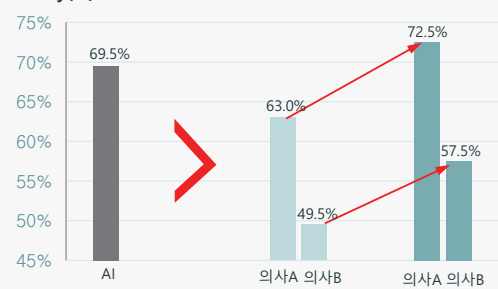
[1] Sunggyun Park et al. *Radiology*, 2018

[2] Lunit INSIGHT CXR (<https://www.lunit.io/ko/products/insight-cxr>), accessed on 2021.07.05

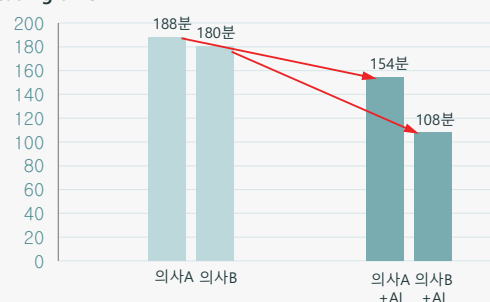
□ Bone Age (Vuno)

- test used to determine the normal growth of children and adolescents, and can be helpful in the timely treatment of musculoskeletal disorders.
- Bone age testing is increasing due to an increase in the number of patients with precocious puberty and a high interest in height growth.
- AI alone can give you more accurate results than experts.
- Use AI-generated results to improve physician reading speed and accuracy.

Accuracy(%)



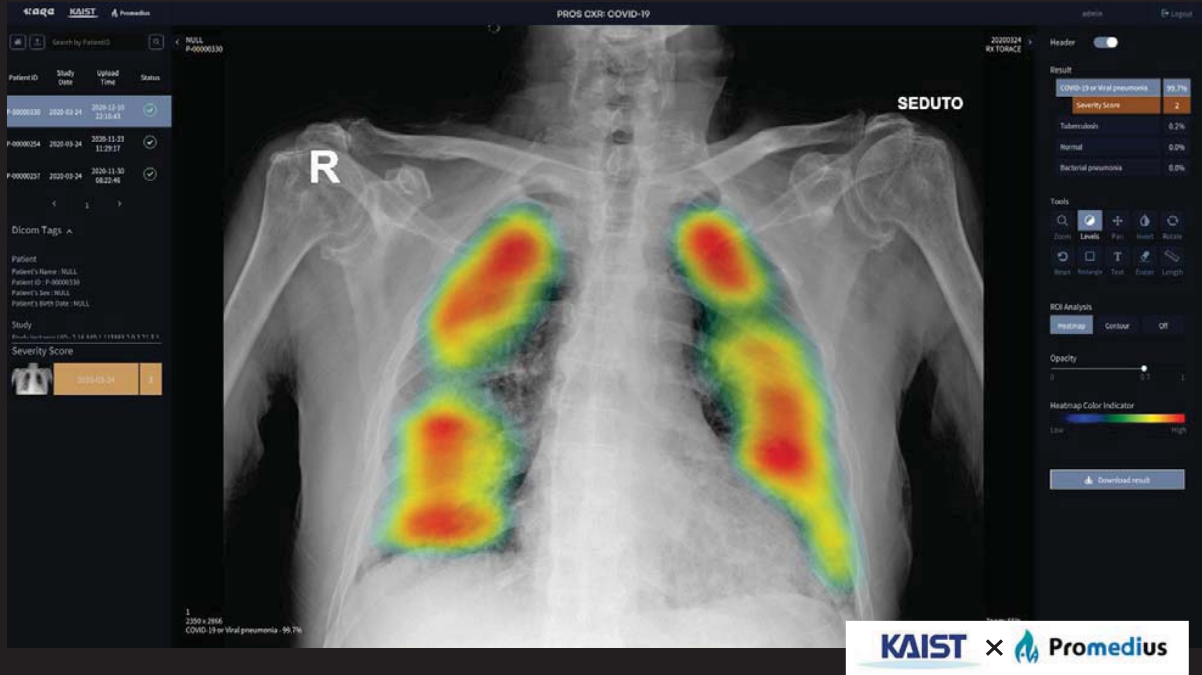
Reading time



[1] Kim JR et al., *AJR Am J Roentgenol*, 2017

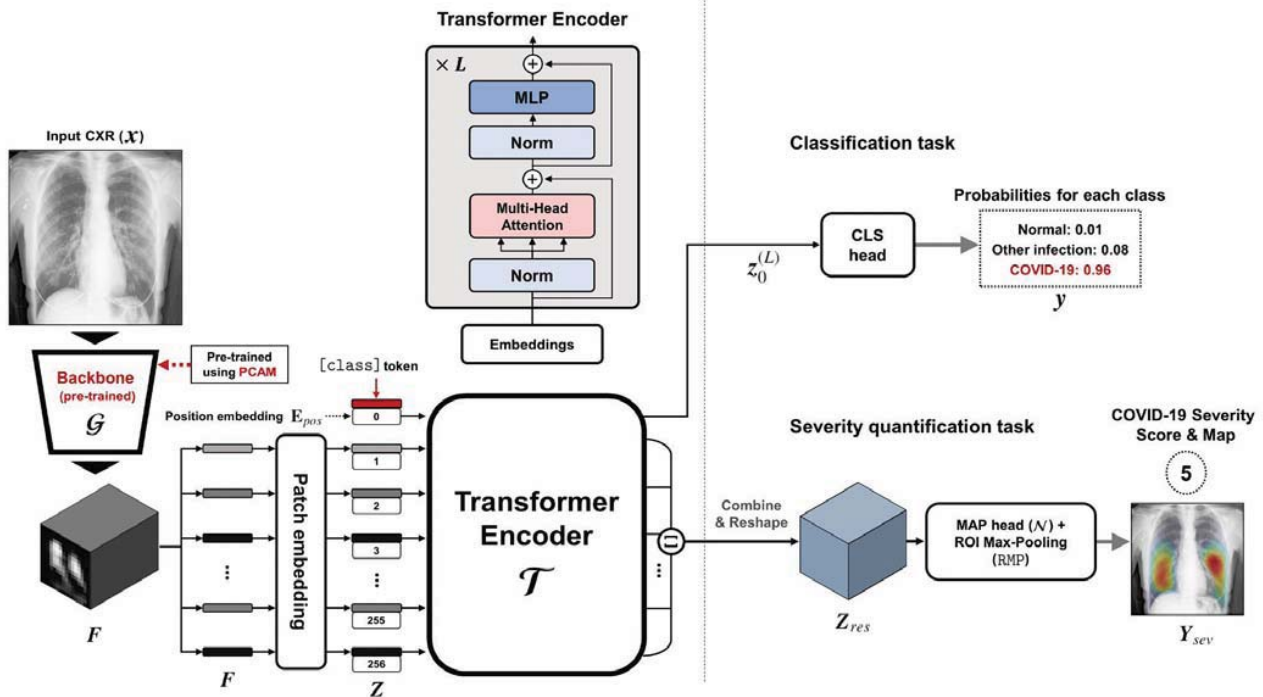
COVID-19 Detection by CXR

Park et al, MEDIA, 2021



(A) Shared layers between two tasks

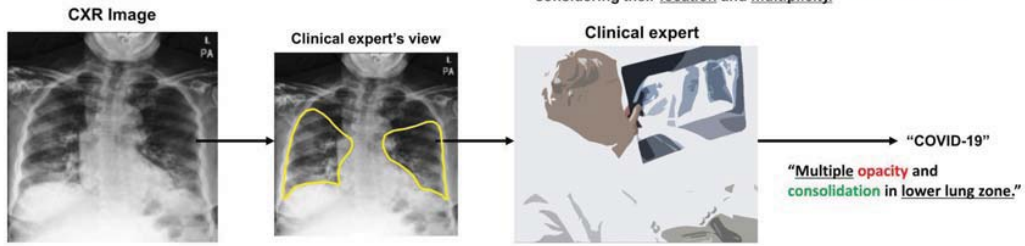
(B) Separated task-specific heads



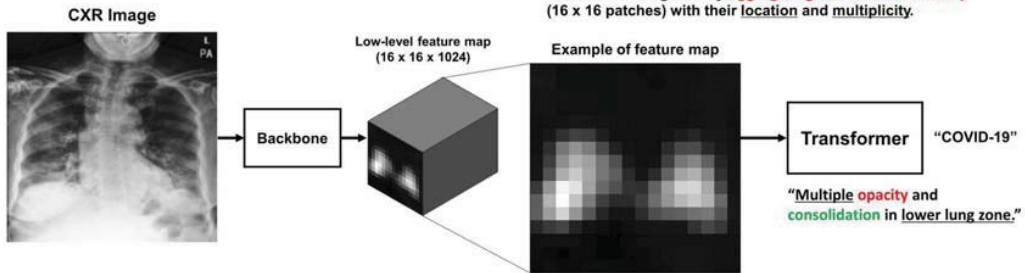
COVID-19 Detection by CXR

Park et al, MEDIA, 2021

Procedure by *Clinical expert*



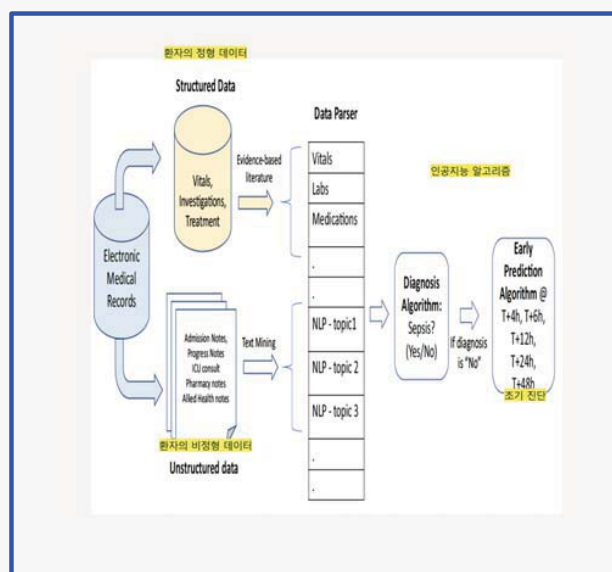
Procedure by *Our model*



AI-driven Prognosis and Prediction

□ Sepsis Prediction

- Severe infection by microorganisms, mortality: early 20-35, septic shock: 40-60%
- Electronic health record (temperature/pulse/respiratory rate/blood pressure/leukocyte count on blood test, etc.) and SOFA score (respiratory system/nervous system/circulatory system/liver/coagulation/kidney)
- Early diagnosis of sepsis Artificial intelligence (AI) technology Compared to medical experts, it was confirmed that the early diagnosis of sepsis was increased by 32% and the false-positive rate was lowered by 17%.



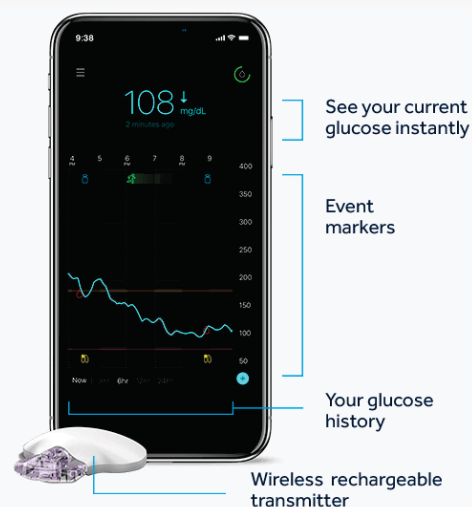
[1] KH Goh et al., *Nature Communications*, 2021

□ Blood Sugar Level Management

IBM, Meditronics

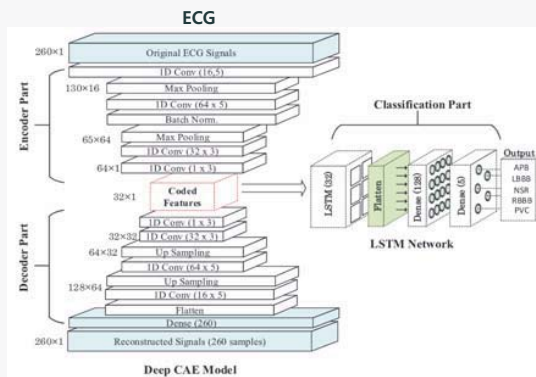
- Guardian connect CGM technology: Manage blood sugar in daily life by linking with mobile app
- Send blood glucose readings to mobile app every 5 minutes
- Predict a patient's hypoglycemia within 1-4 hours with 98.5% accuracy
- Bringing the patient's blood sugar levels to normal for an average of 39 more minutes per day.

Meditronic, IBM Watson



□ Arrhythmia Classification

- Predictive research using electrocardiogram (ECG) signals
- Combining Convolutional Neural Network (CNN) and Long Short-Term Memory Model (LSTM) to Classify Arrhythmic Disorders
- Most of the various types of arrhythmias show more than 99% accuracy.



Accuracy: 99.23%

Output Class	APB	LBBB	NSR	RBBB	PVC
APB	94.0% 359	0.0% 0	0.4% 45	0.2% 2	0.0% 0
LBBB	0.3% 1	99.8% 1181	0.0% 3	0.1% 1	0.1% 1
NSR	4.2% 16	0.2% 2	99.4% 11162	0.4% 4	1.4% 15
RBBB	1.0% 4	0.0% 0	0.0% 3	99.4% 1094	0.0% 0
PVC	0.5% 2	0.0% 0	0.1% 16	0.0% 0	98.6% 1093

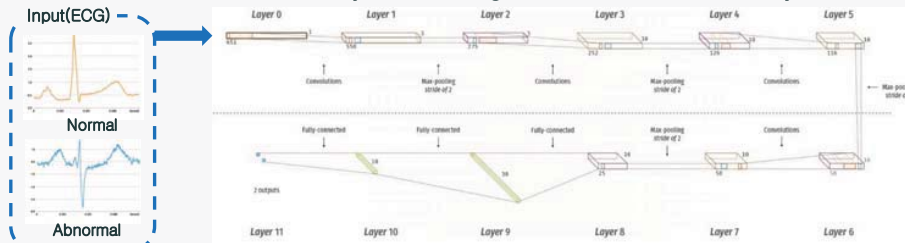
Target Class

부정맥 예측

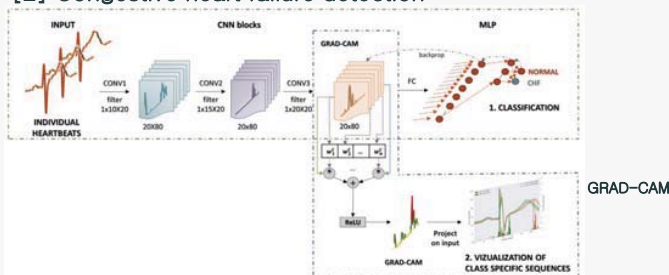
[1] Yildirim, Ozal, et al., *Computer methods and programs in biomedicine*, 2019

□ Detection of Heart Failure using ECG

- [1] Automated detection of arrhythmias using different intervals of tachycardia ECG segments



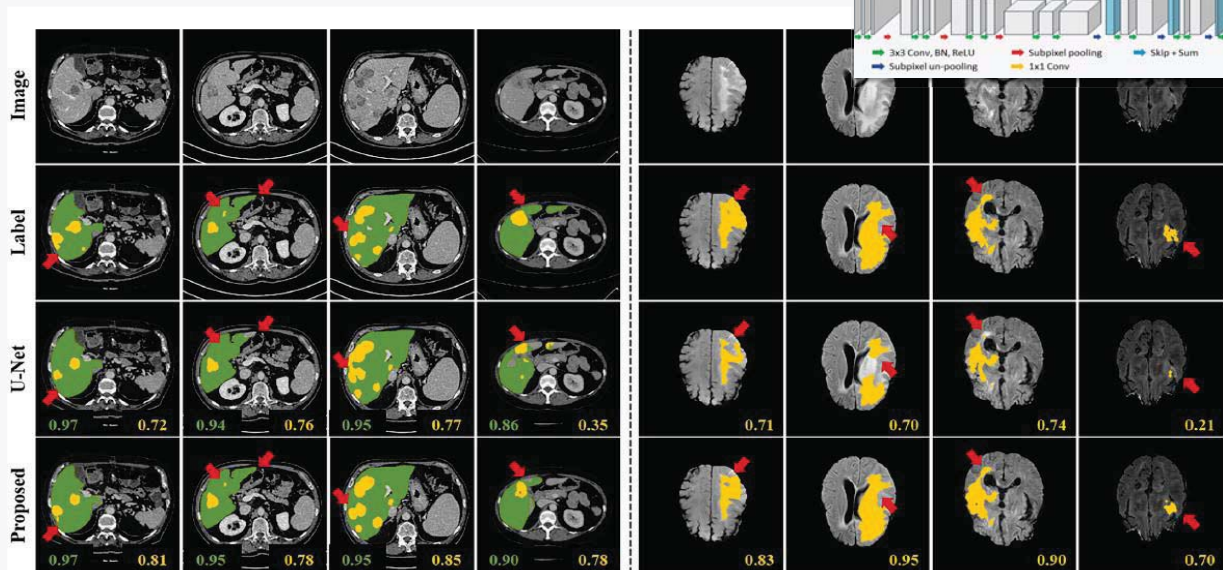
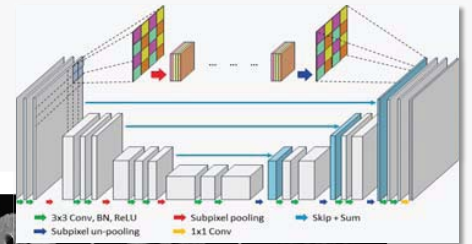
- [2] Congestive heart failure detection



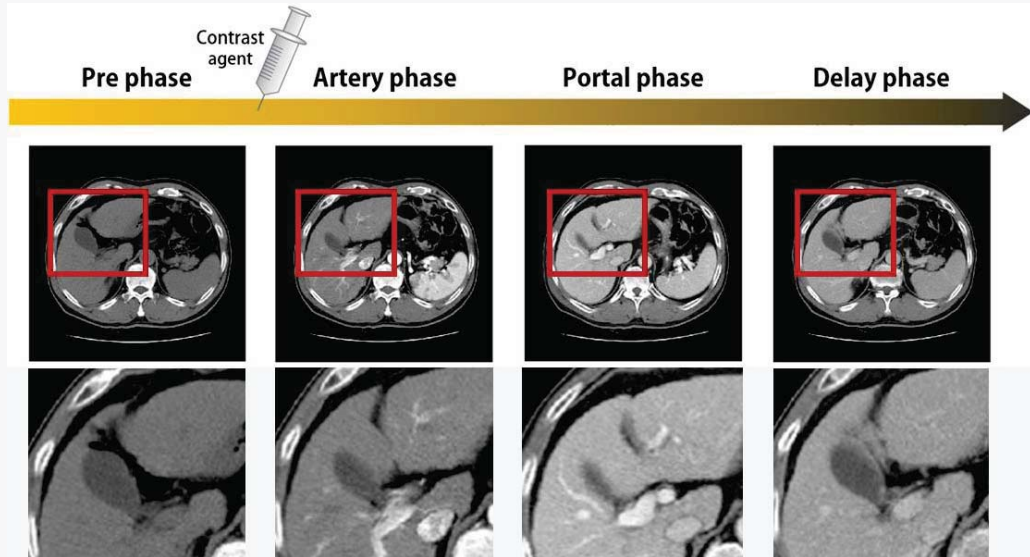
[1] Acharya, U. Rajendra, et al., *Information Sciences*, 2017
 [2] Porumb, M. et al., *Biomedical Signal Processing and Control*, 2020

Workflow

□ Segmentation



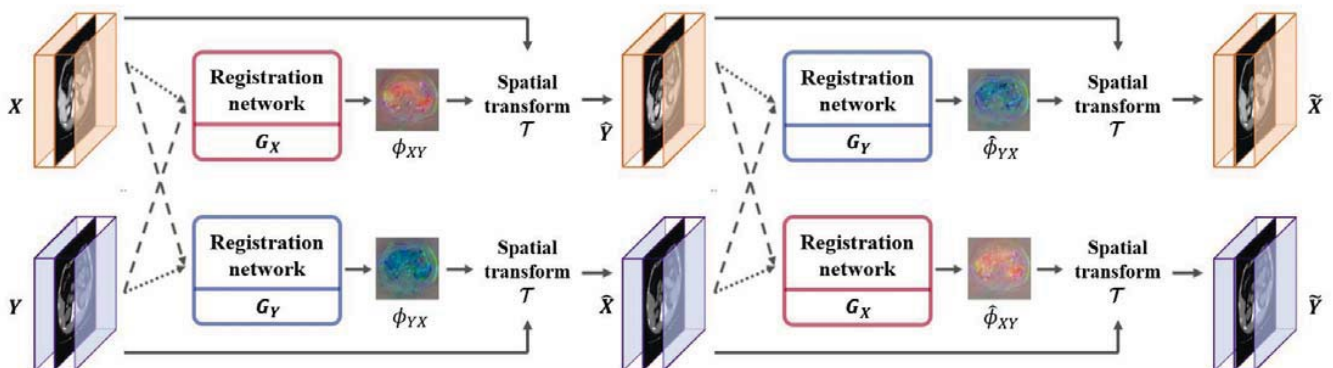
□ Image Registration



□ CycleMorph: Kim et al, MEDIA, 2021

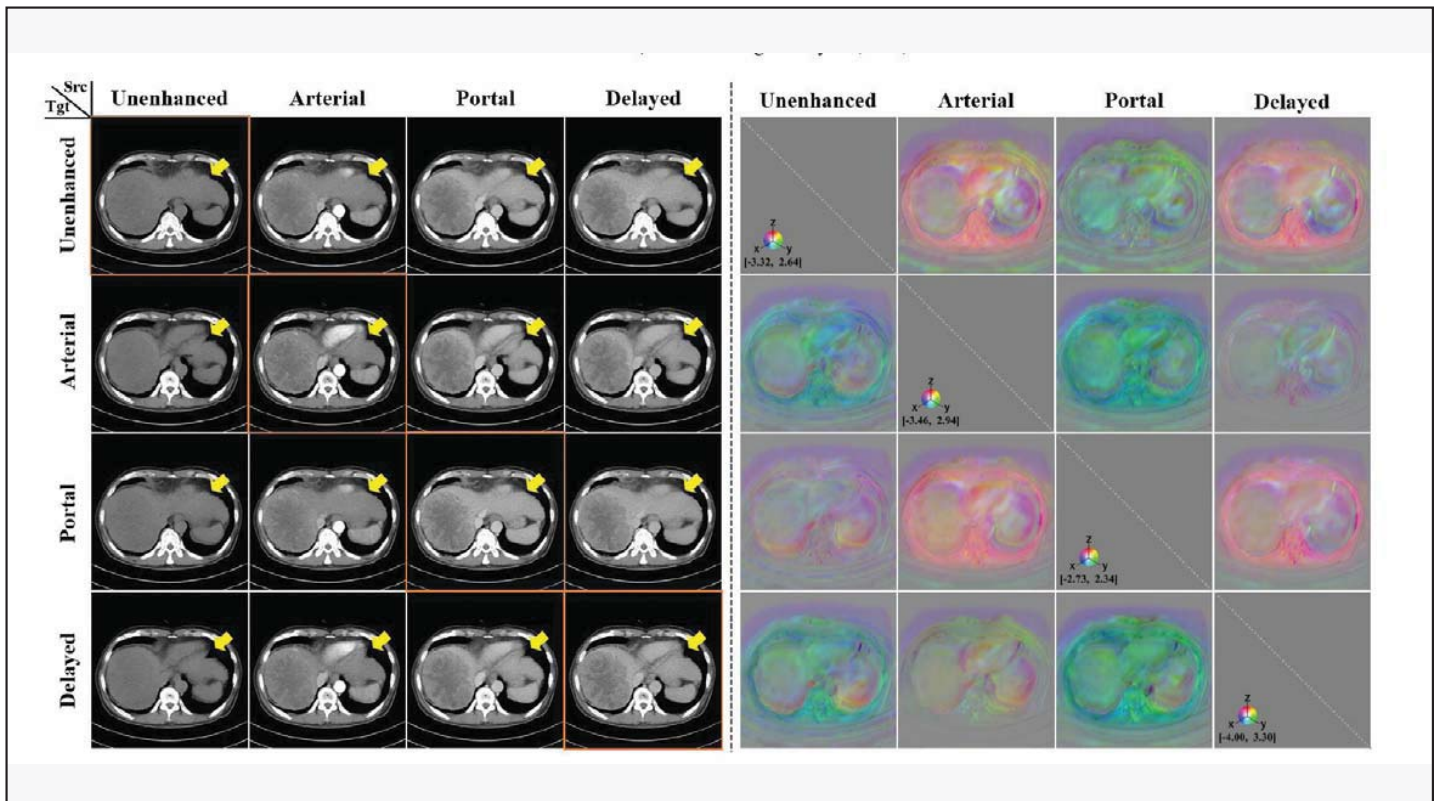
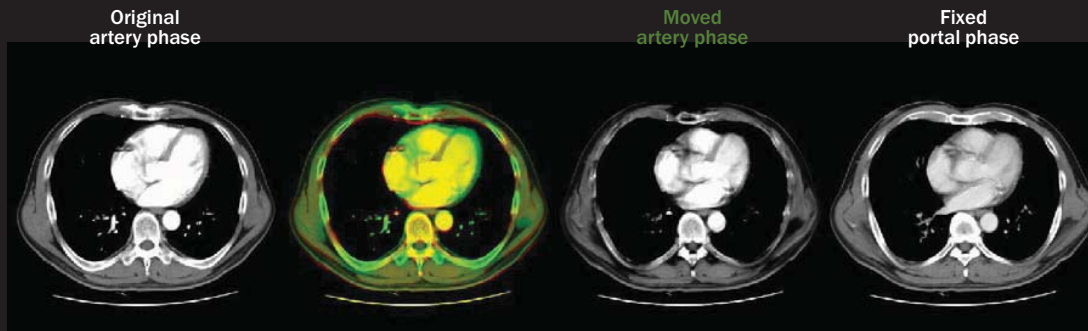
Boah Kim *et al.* / Medical Image Analysis (2021)

3



□ Multiphase Liver CT Registration

Artery phase → Portal phase



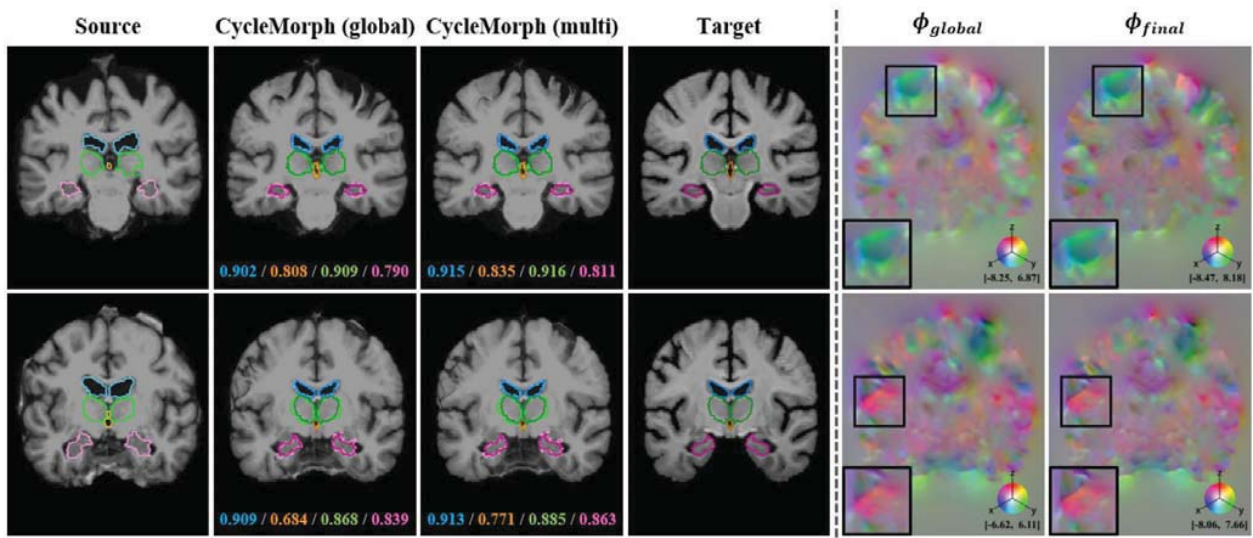


Image Enhancement

Low Dose CT Grand Challenge

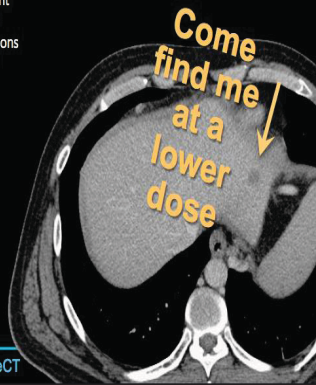
NIH National Institute of Biomedical Imaging and Bioengineering

AAEP AMERICAN ASSOCIATION of PHYSICISTS IN MEDICINE

MAYO CLINIC

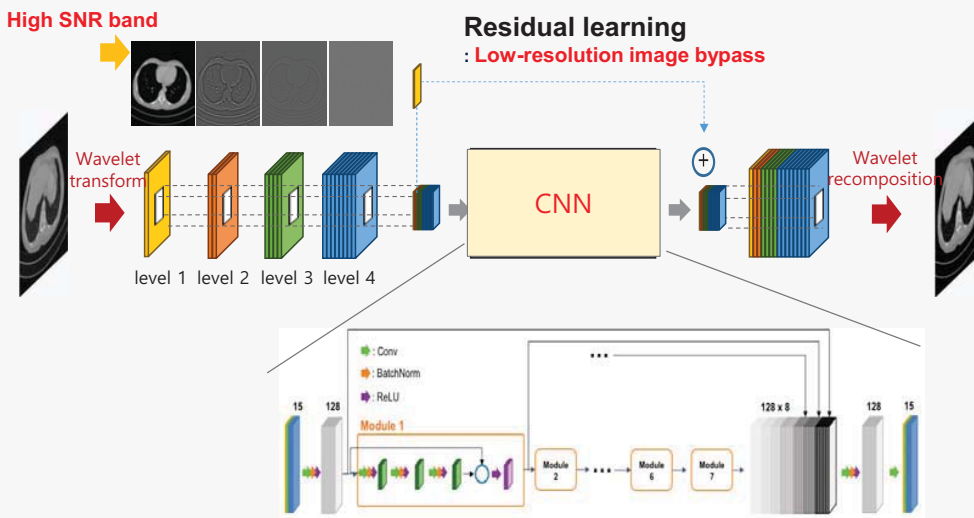
CT Clinical Innovation Center

- Radiologist-selected abdominal CT patient cases (10 training, 20 testing) with noise inserted to simulate lower dose acquisitions
- Projection data converted into an open format (user manual and reading tools provided)
- Apr 2016: Participants submit reconstructed images or denoised images to AAPM website
- Jun 2016: Images read by radiologists at the host site
- Aug 2016: Winners announced at AAPM Annual Meeting



www.aapm.org/GrandChallenge/LowDoseCT

□ AAPM-Net: first deep learning for low-dose CT

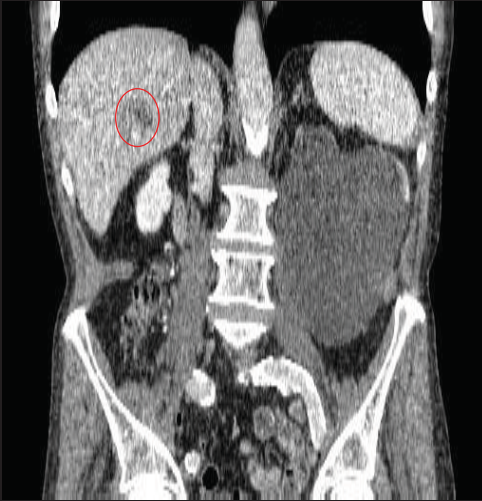


(Kang, et al, Medical Physics 44(10))

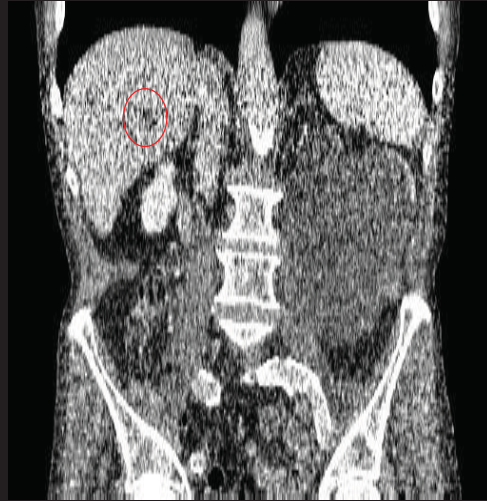
90

□ Low-dose CT: Kang et al, 2017

Full dose

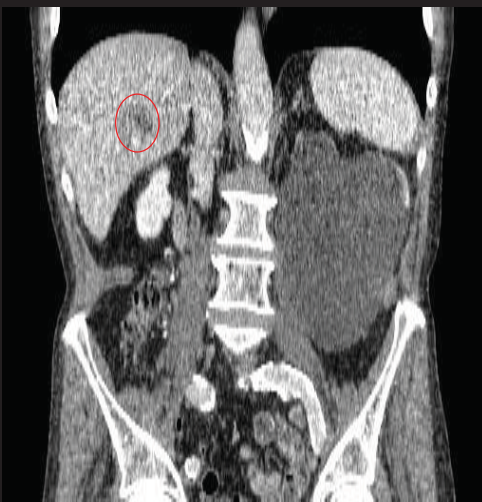


Quarter dose

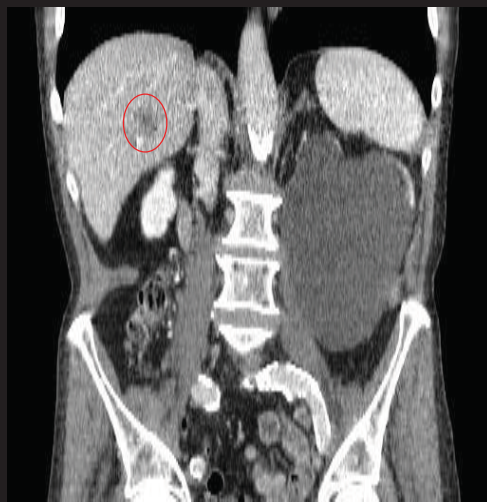


□ Low-dose CT: Kang et al, 2017

Full dose



Proposed



Canon

AiCE
Integrated Intelligence

See through the noise

World's 1st Deep Learning
Reconstruction for CT



GE Healthcare

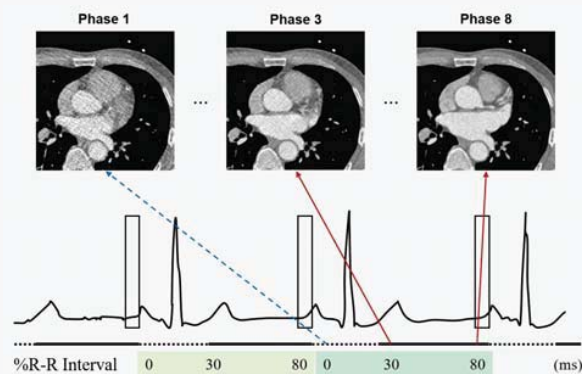
TrueFidelity

How the best see better.



□ Low-dose CT Denoising without Reference

- **Multiphase Cardiac CT denoising**
 - Phase 1, 2: low-dose, Phase 3 ~ 10: normal dose
 - Goal: dynamic changes of heart structure
 - **No reference available**

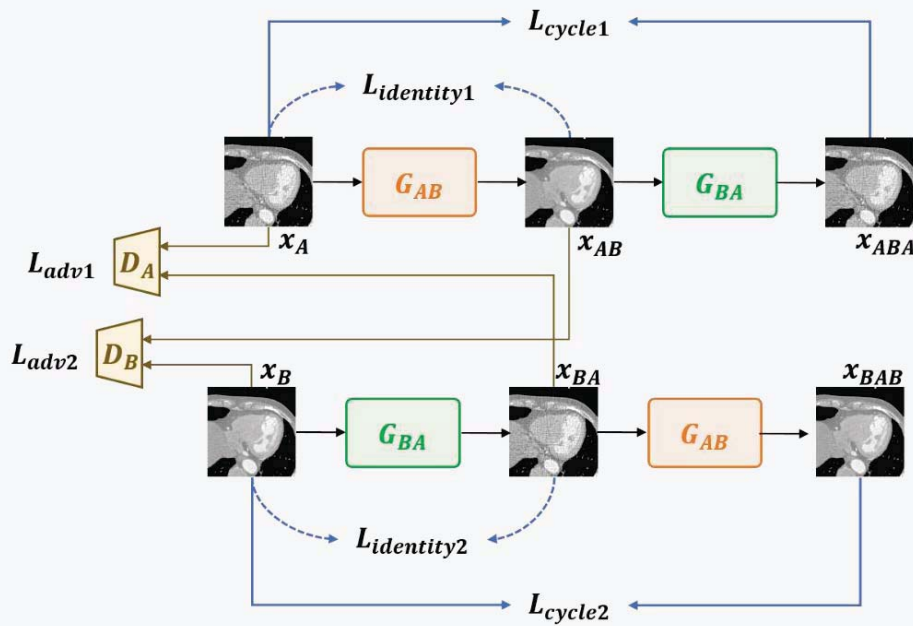


Kang et al, Medical Physics, 2018

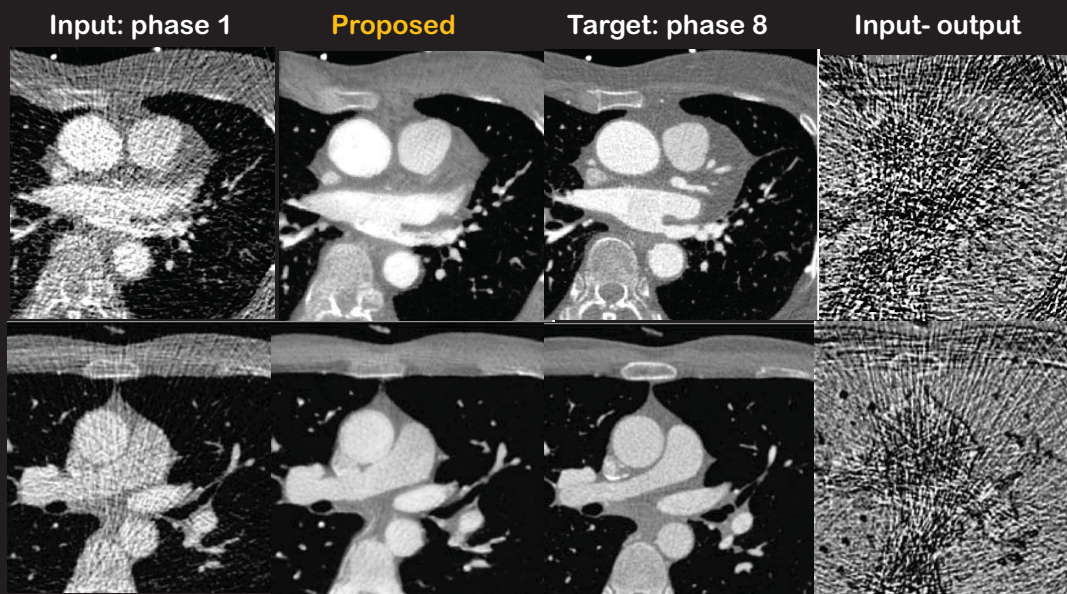
94

□ CycleGAN Denoising for Low-Dose CT

Kang et al, Medical Physics, 2018

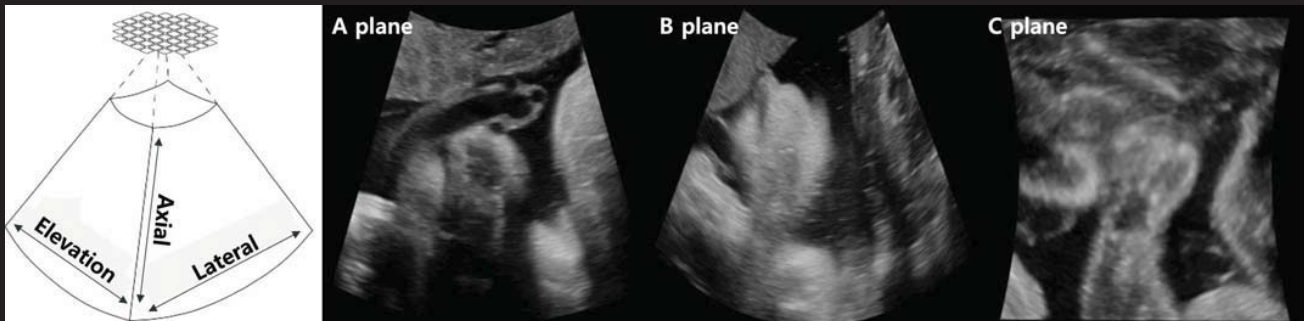


Lose dose (5%) → high dose



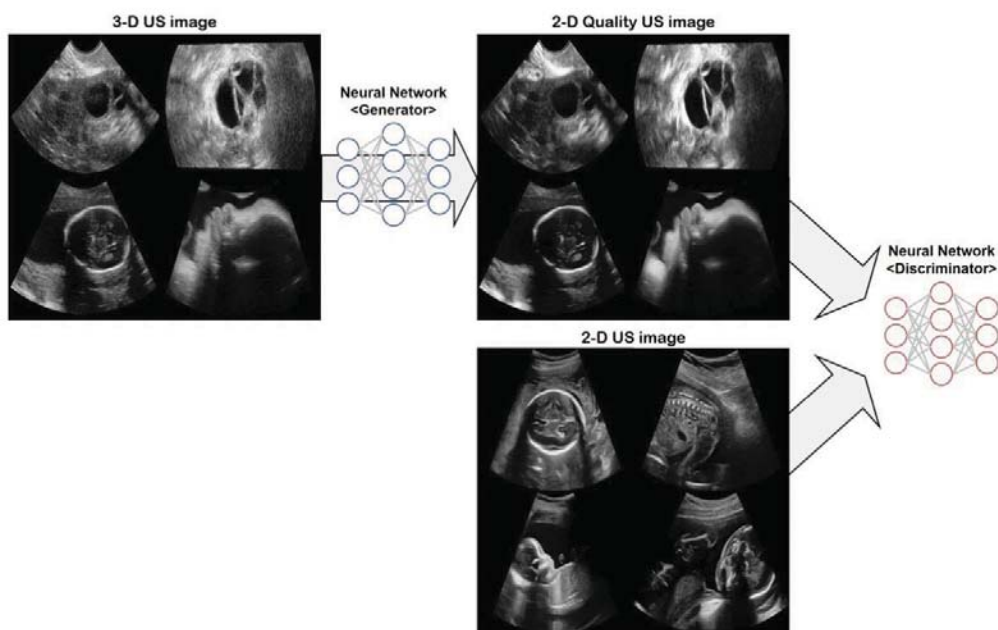
□ Ultrasound image

✓ To overcome the image quality degradation in 3D Ultrasound

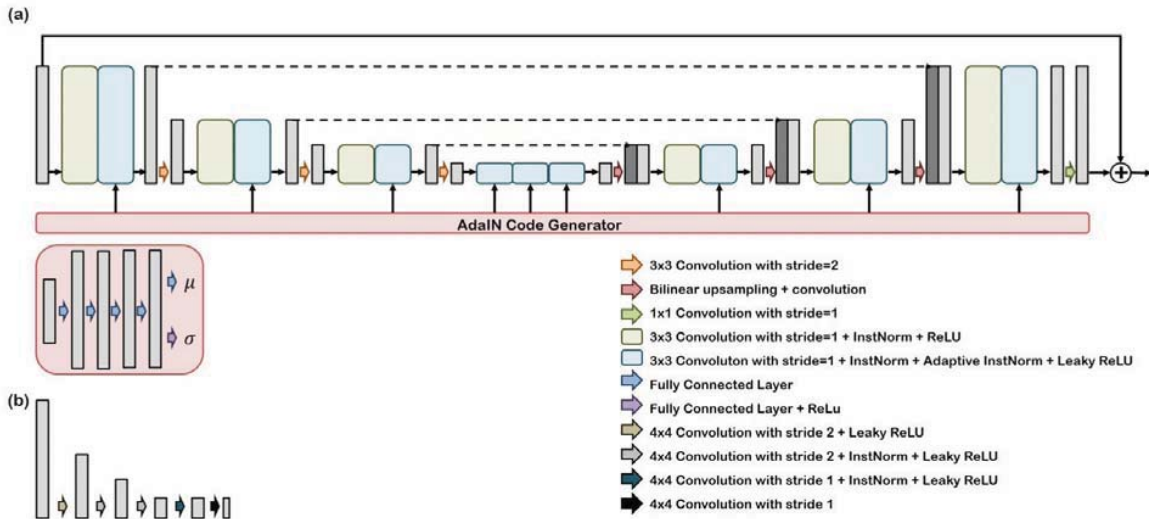


- Compared to 2D, 3D plane images have many artifacts
- In particular, B/C plane artifacts are sever.

□ Unsupervised Learning by CycleGAN



Switchable Architecture



GYN A-plane

GYN B-plane



$\alpha = 0.5 \sim 1.0$

$\alpha = 0.5 \sim 1.0$

OB A-plane



$\alpha = 0.5 \sim 1.0$

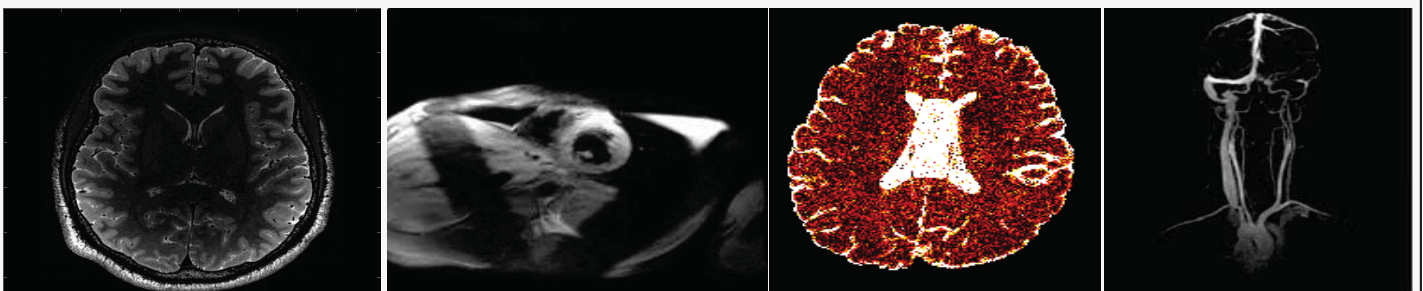
OB C-plane



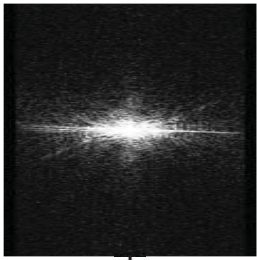
$\alpha = 0.5 \sim 1.0$

❑ Unmet Needs in MRI

- ❑ MR is an essential tool for diagnosis
- ❑ MR exam protocol : 30~60 min/patient
 - ✓ should increase the throughput of MR scanning
- ❑ Cardiac imaging, fMRI
 - ✓ Should improve temporal resolution
- ❑ Multiple contrast acquisition in a short time



Accelerated MRI



k-space, measurement space

\mathcal{F}^{-1}

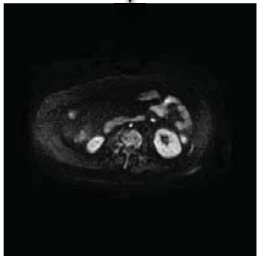
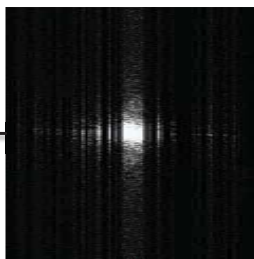
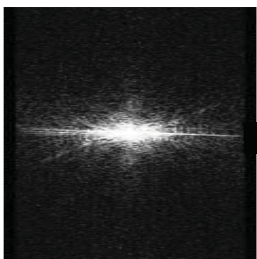


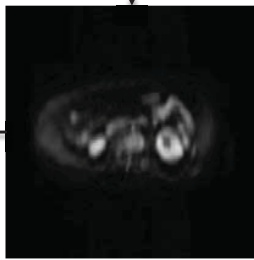
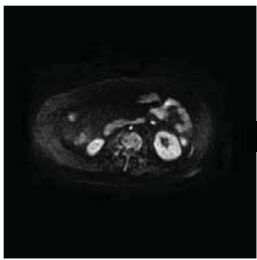
Image space

Accelerated MRI

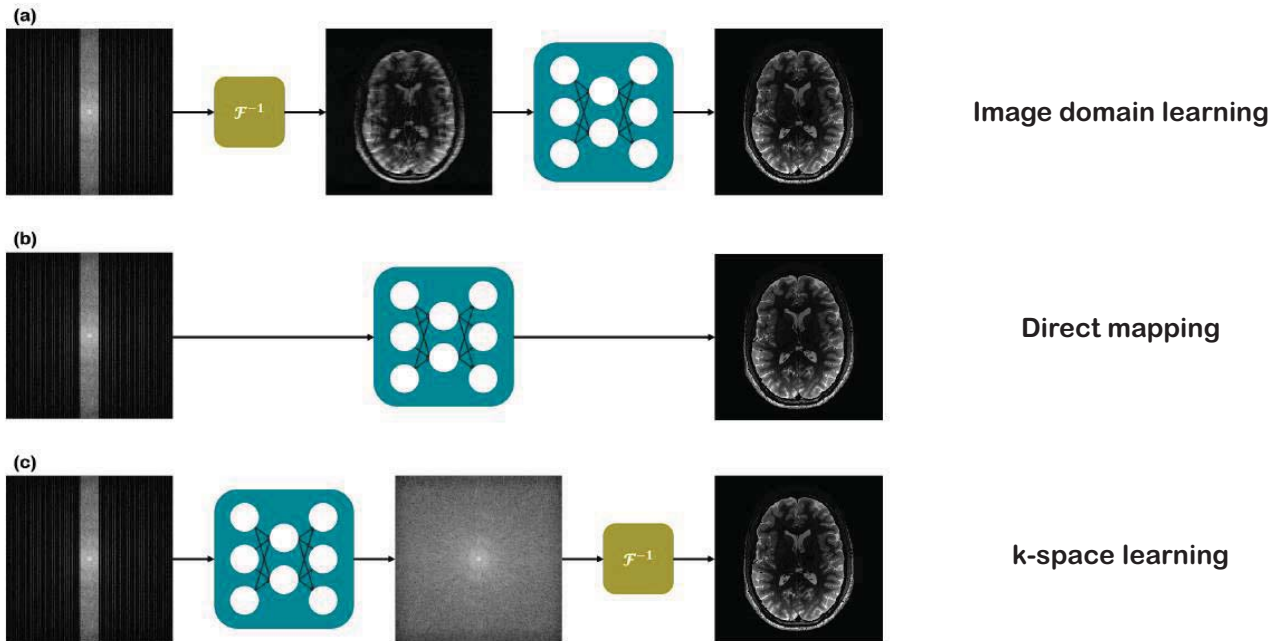


Reconstruction

\mathcal{F}^{-1}



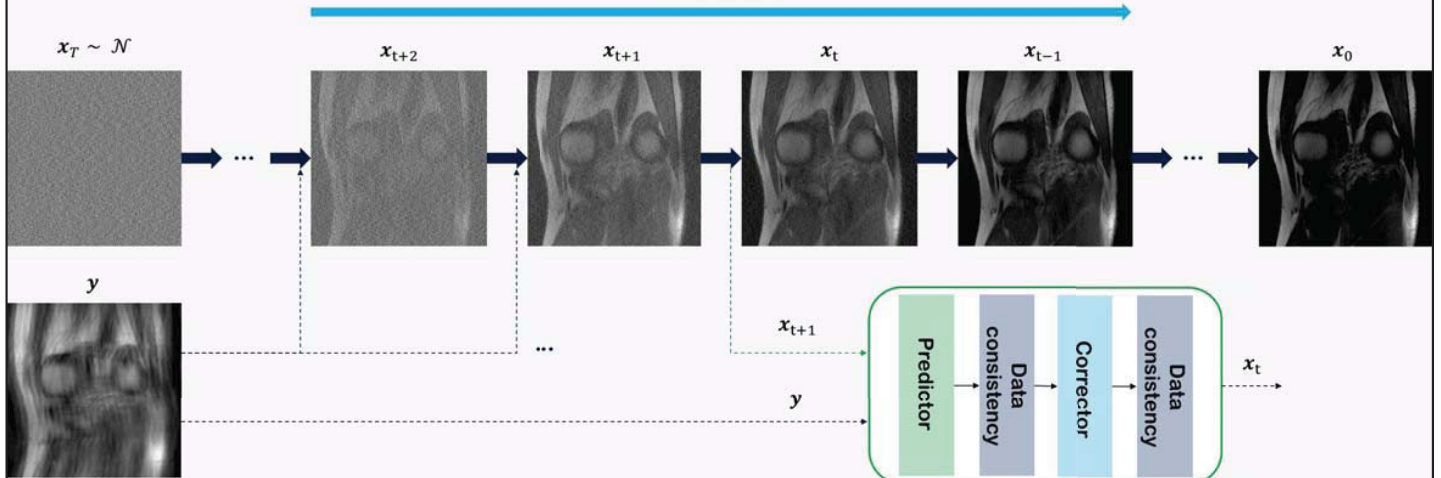
□ Deep Learning for Accelerated MRI



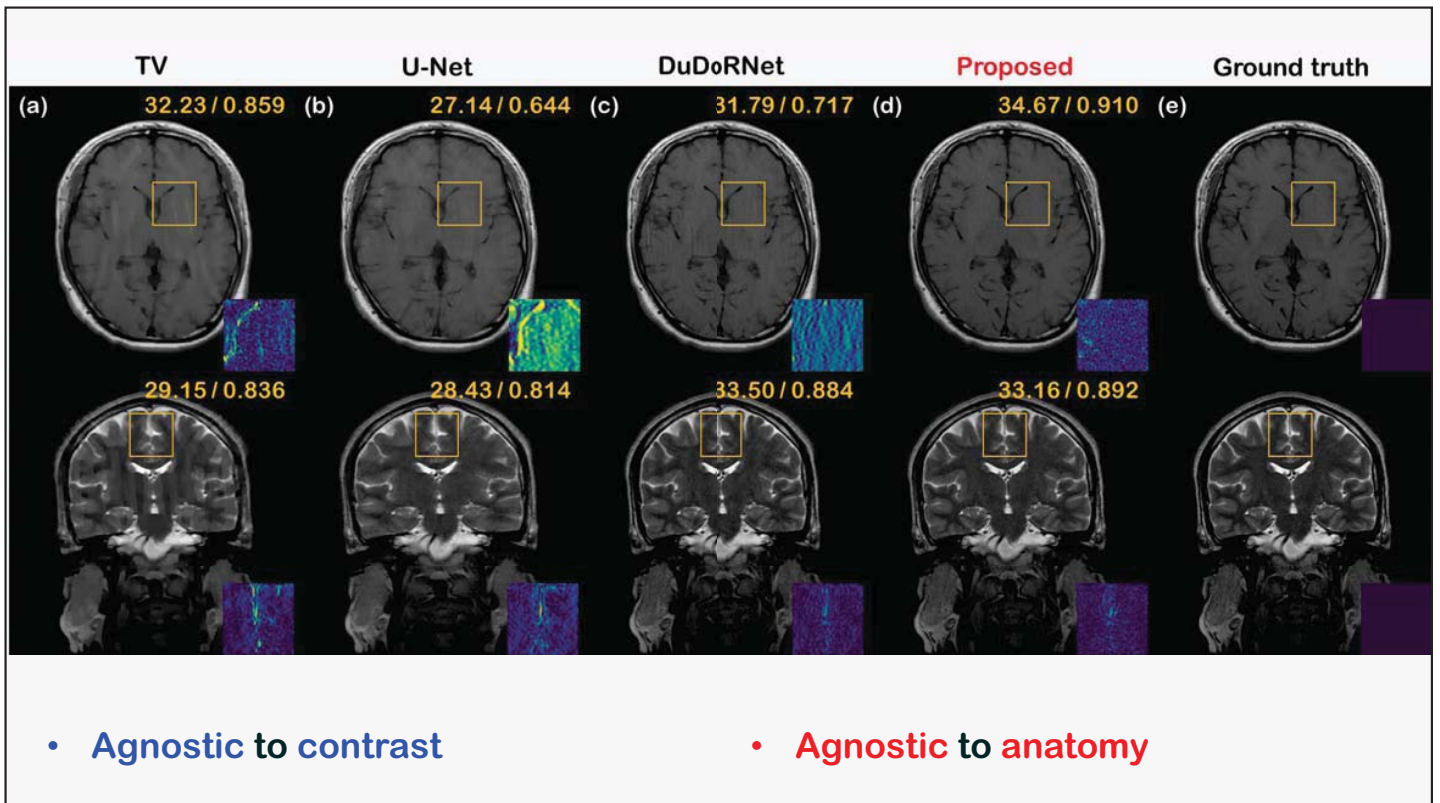
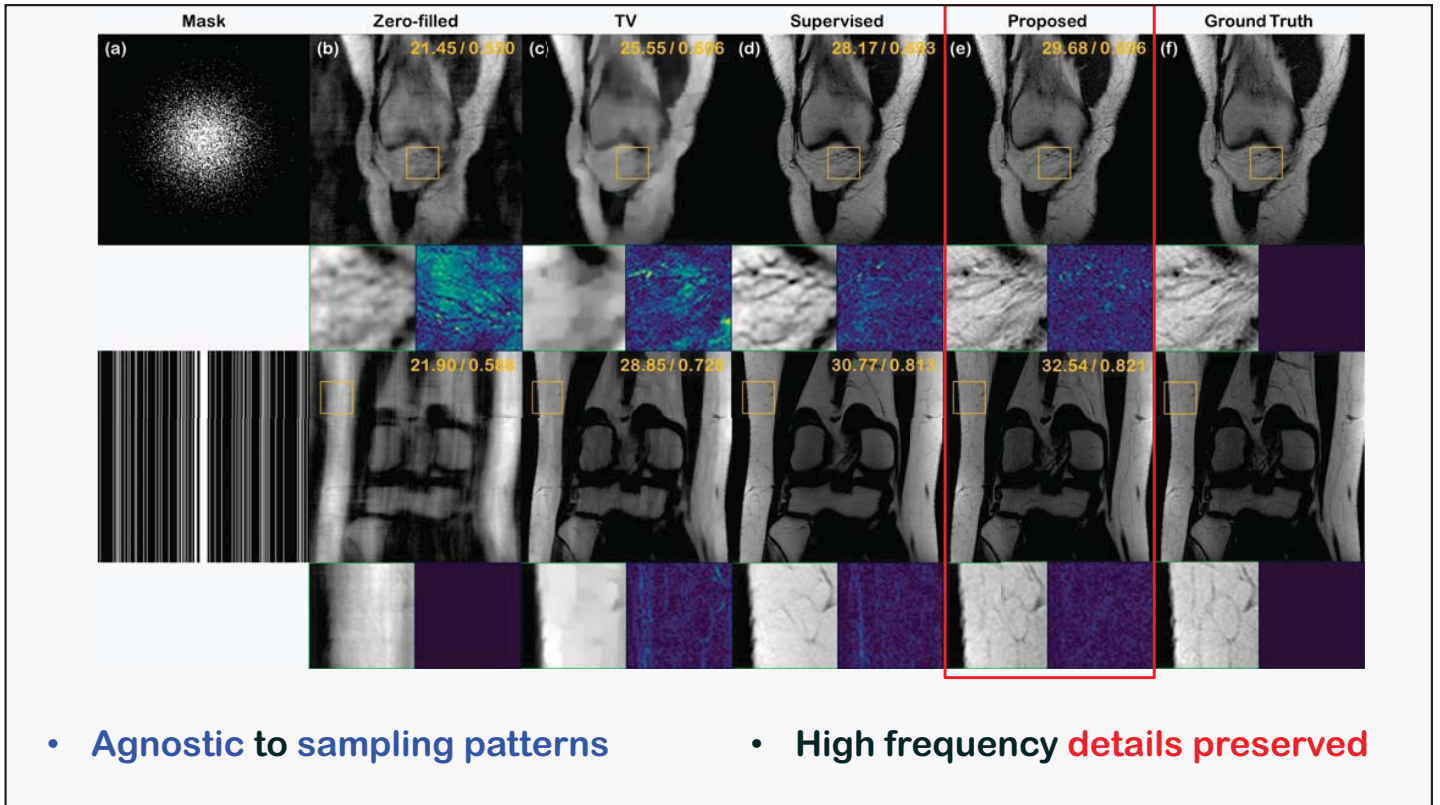
□ Diffusion models for accelerated MRI

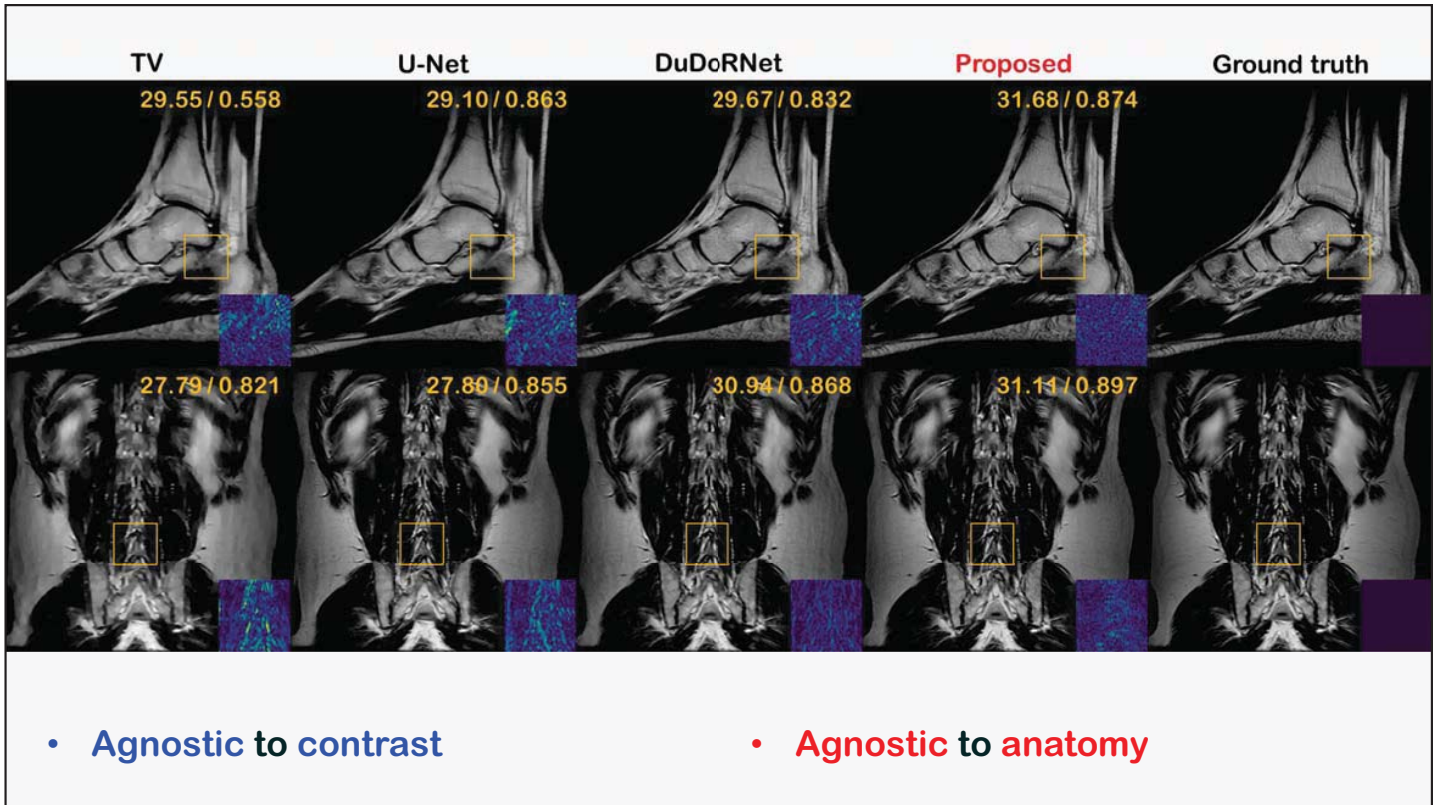
H. Chung, et al, MEDIA, 2022

Reverse SDE



Imposing **data consistency step** for every iteration





Oversight AI

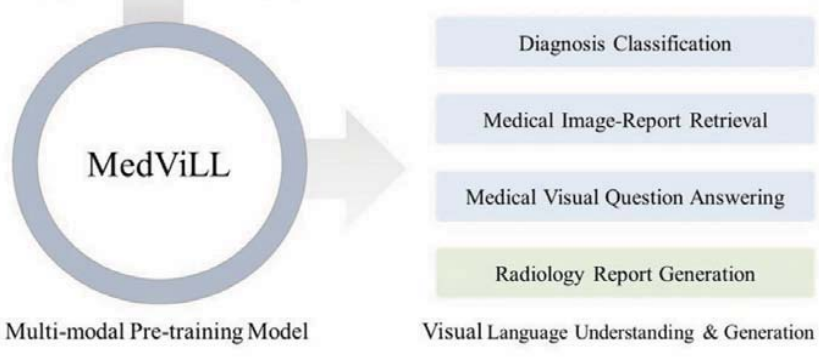
MedViLL: Single Stream VLP for CXR

Chest x-ray Image & Radiology Report



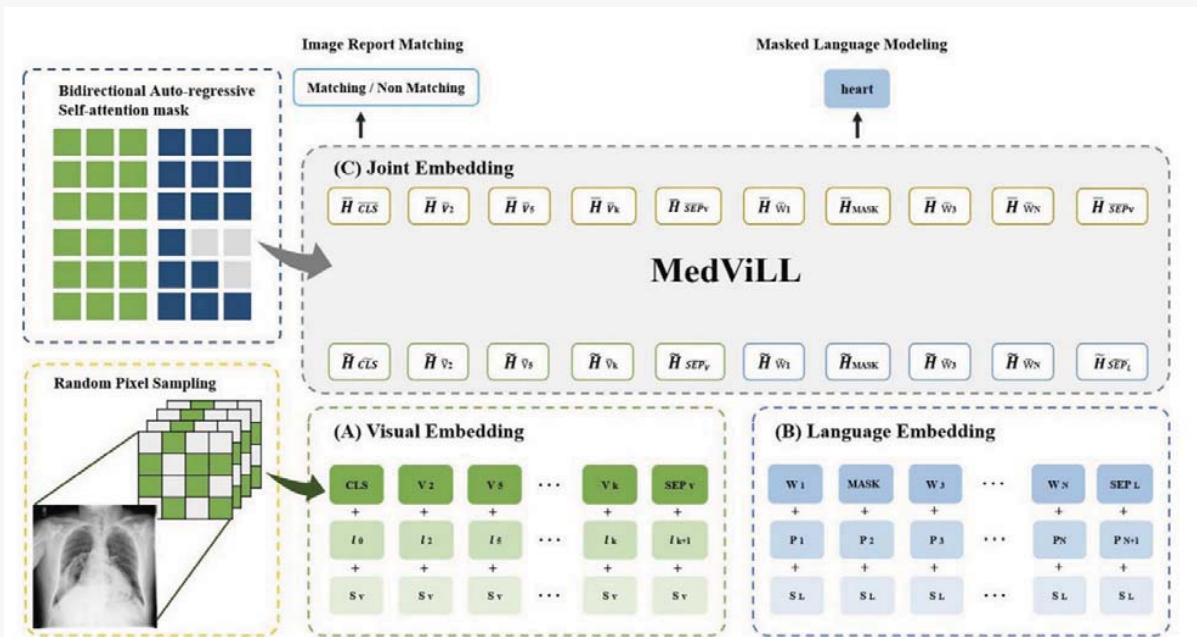
FINDINGS:
 Borderline heart size, similar. No pneumothorax. No effusion.
 Mildly increased pulmonary vascularity, more prominent.
 Segmental elevation left hemidiaphragm. Tortuous calcified aorta.
 Minimal basilar atelectasis. Probable scarring right costophrenic angle.

IMPRESSION:
 Minimal bibasilar atelectasis.

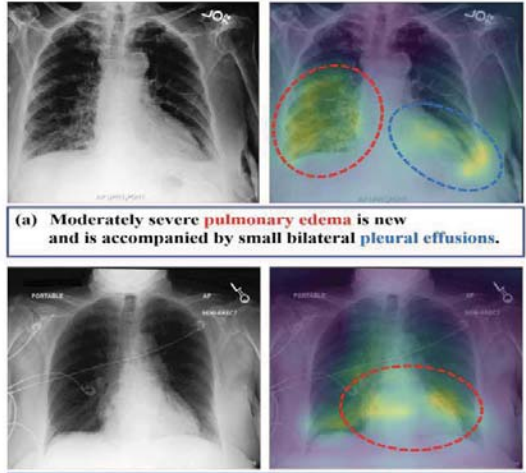


Moon, Jong Hak, et al.
 "Multi-modal understanding and generation for medical images and text via vision-language pre-training." *arXiv preprint arXiv:2105.11333* (2021).

MedViLL: Single Stream VLP for CXR



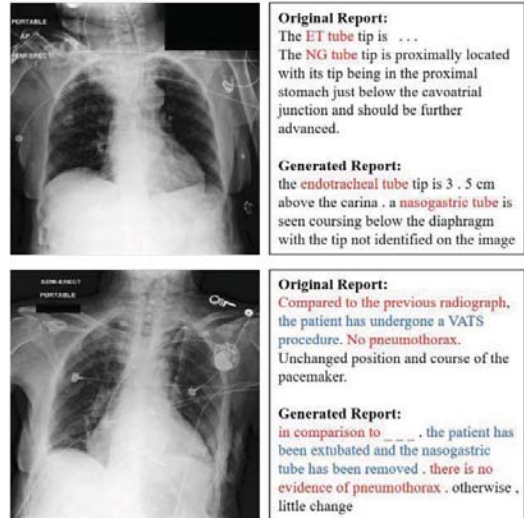
□ MedViLL: Single Stream VLP for CXR



(a) Moderately severe pulmonary edema is new and is accompanied by small bilateral pleural effusions.

(b) Heart size borderline enlarged. No pleural abnormality. Lungs clear. Normal pulmonary vasculature.

(a) Attention map visualization.



Original Report:
The ET tube tip is ...
The NG tube tip is proximally located with its tip being in the proximal stomach just below the cavoatrial junction and should be further advanced.

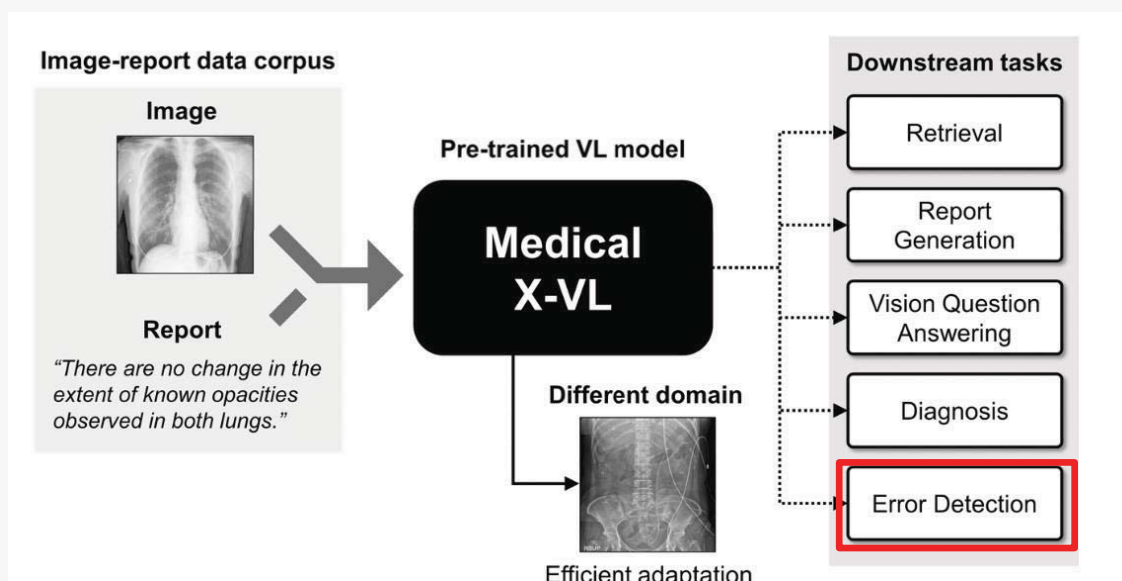
Generated Report:
the endotracheal tube tip is 3 . 5 cm above the carina . a nasogastric tube is seen coursing below the diaphragm with the tip not identified on the image

Original Report:
Compared to the previous radiograph, the patient has undergone a VATS procedure. No pneumothorax. Unchanged position and course of the pacemaker.

Generated Report:
in comparison to _ _ _ , the patient has been extubated and the nasogastric tube has been removed . there is no evidence of pneumothorax . otherwise , little change

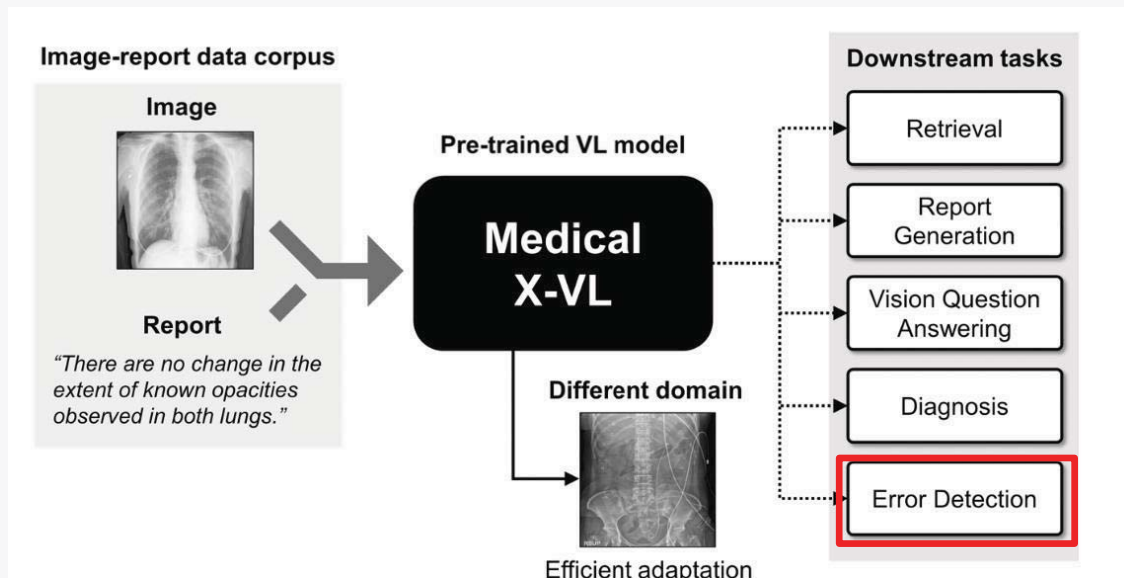
(b) Radiology report generation analysis.

□ Medical X-VL: Dual Stream VLP



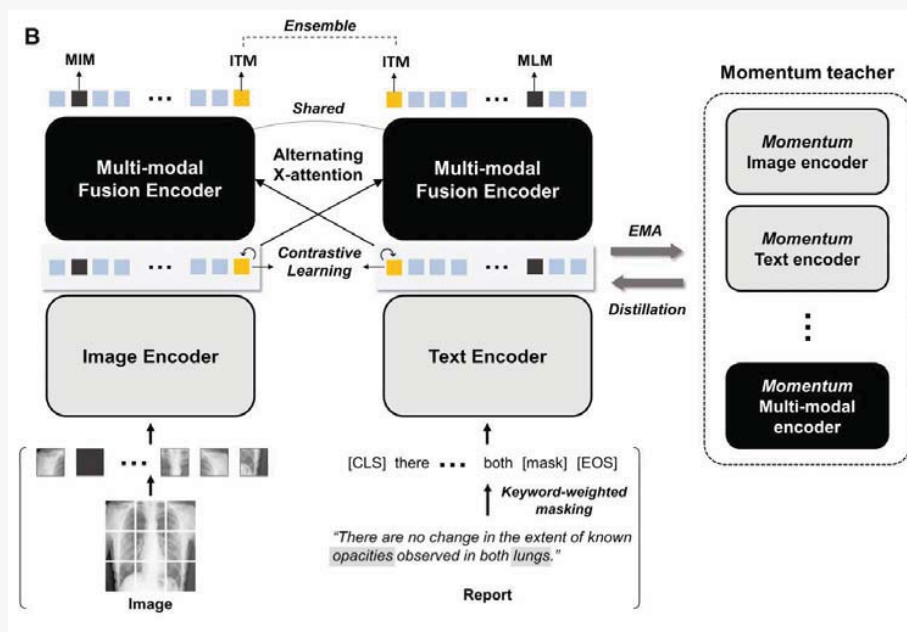
Medical X-VL: Dual Stream VLP

Park, et al. *arXiv preprint arXiv:2208.05140* (2022).



Medical X-VL: Dual Stream VLP

Park, et al. *arXiv preprint arXiv:2208.05140* (2022).



❑ Medical X-VL: Dual Stream VLP

Park, et al. *arXiv preprint arXiv:2208.05140* (2022).

A Right-left orientation confusion AUC = 0.759 ± 0.067



True: In comparison with the study of ___ the monitoring and support devices are unchanged. there is again substantial **enlargement** of the cardiac silhouette with pulmonary vascular congestion and bilateral pleural effusions more prominent on the **right**.

Wrong: In comparison with the study of ___ the monitoring and support devices are unchanged. there is again substantial enlargement of the cardiac silhouette with pulmonary vascular congestion and bilateral pleural effusions more prominent on the **left**.

Corrected: In comparison with the study of ___ the monitoring and support devices are unchanged. there is again substantial **widening** of the cardiac silhouette with pulmonary vascular congestion and bilateral pleural effusions more prominent on the **right**.

B Patient-report misregistration AUC = 0.981 ± 0.025

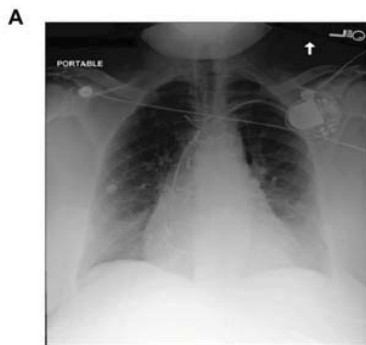


Wrong: The patient remains intubated with the ET tube tip is 7 cm above the carina. The NG tube tip passes below the diaphragm with its tip not clearly seen on the current examination. The right internal jugular line tip is at the level of mid SVC. The patient is in pulmonary edema, moderate, progressed since the prior study, associated bilateral pleural effusions.

Suggestion: Single portable view of the chest. Lower lung volumes seen on the current exam. Patchy region of opacity identified at the left lung base. Elsewhere, the lungs are clear. The cardiomeastinal silhouette is within normal limits. Tortuosity of the descending thoracic aorta is noted.

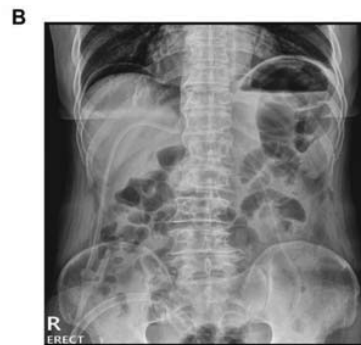
❑ Medical X-VL: Dual Stream VLP

Park, et al. *arXiv preprint arXiv:2208.05140* (2022).



Label: The patient is status **post median sternotomy**. Left-sided pacer device is seen with leads extending to the expected positions of the **right atrium and right ventricle**. The **cardiac silhouette is markedly enlarged**. Mediastinal contours are unremarkable. There may be **minimal central vascular engorgement** without overt pulmonary edema. No large pleural effusion is seen. There is no evidence of pneumothorax or focal consolidation. The lungs appear relatively hyperinflated.

Generated: Patient is status **post median sternotomy** CABG and left-sided dual-chamber pacemaker device with leads terminating in the **right atrium and right ventricle**. **Mild cardiomegaly** is unchanged. Mediastinal and hilar contours are similar. There is **mild pulmonary vascular congestion**. No focal consolidation pleural effusion or pneumothorax is present. No acute osseous abnormalities detected.



Label: Mild **post-op. ileus** and a small amount of pneumoperitoneum.

Generated: Mild **post-op. ileus** and a small amount of pneumoperitoneum.

Q&A