Detection and Diagnosis of Lesions in Medical Imaging with Ultra-lightweight Deep Learning Models

Kenji Suzuki, PhD

Professor & Founding Director of BMAI

Biomedical Artificial Intelligence Research Unit (BMAI)
Institute of Integrated Research
Institute of Science Tokyo*

*Tokyo Institute of Technology and Tokyo Medical and Dental University have been merged on October 1, 2024

Visiting Professor, School of Medicine, Kobe University
Visiting Professor, Graduate School of Dentistry, Tohoku University
Visiting Professor, St. Marianna University School of Medicine

Contents

- AI-aided Medical Image Diagnosis
  - “Small-data” deep learning
  - Efficiency of Small-data deep learning model
  - Small-data deep learning application to rare cancer
- AI/Deep-Learning Imaging
  - Bone suppression in chest radiographs
  - Radiation dose reduction in CT and Tomosynthesis

Computer-Aided Diagnosis (CAD)[1-5]

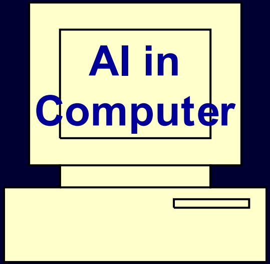
→ AI-aided Diagnosis "AI Doctor"

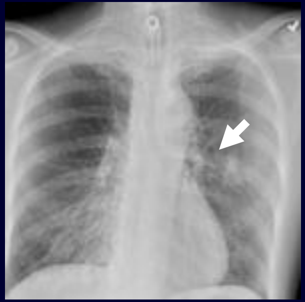
1) Doi K et al., Eur J Radiology (1999)
2) Giger ML & Suzuki K, *Biomed Info Tech* (2007)
3) Suzuki K, *Machine Learning in CAD* (2012)
4) Chang JZ et al., *Nature* (2016)
5) Chen Y & Suzuki K, *AI in Decision Support Systems* (2018)

Computer/AI-Aided Diagnosis/Detection (CAD)[1-5]

“AI Doctor”

Medical image





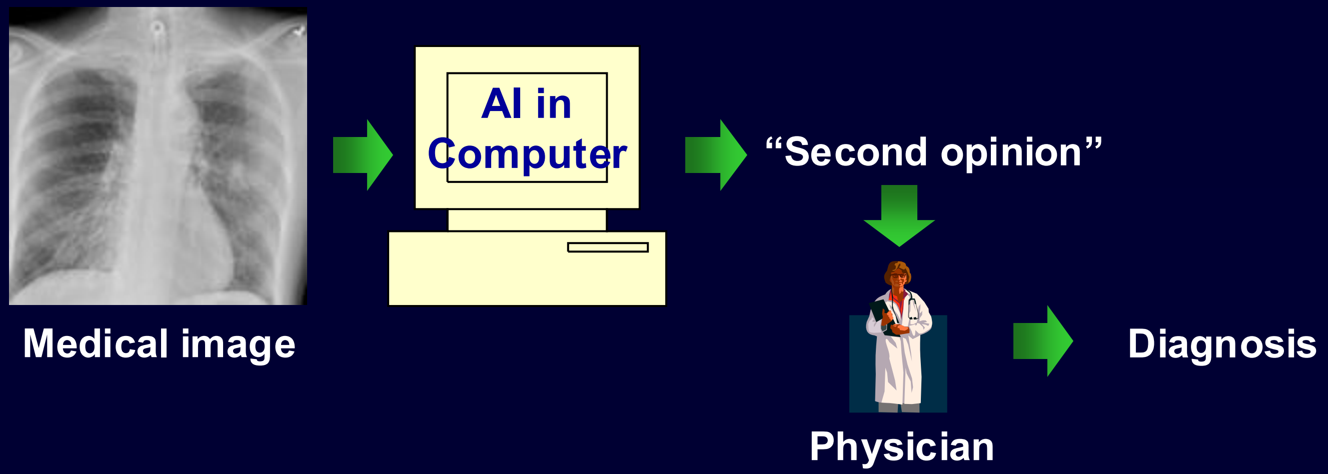
“Second opinion”
(e.g., “I found a pattern similar to a cancer”)

Diagnosis

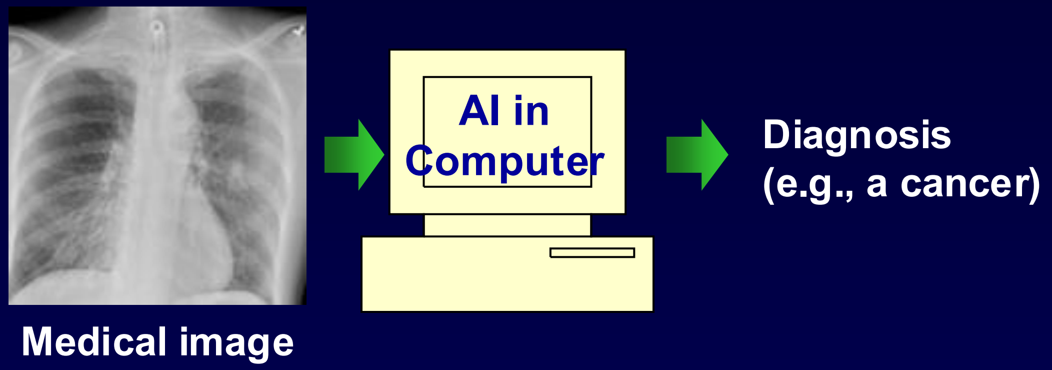
Physician

1) Doi K et al., Eur J Radiology (1999)
2) Giger ML & Suzuki K, *Biomed Info Tech* (2007)
3) Suzuki K, *Machine Learning in CAD* (2012)
4) Chang JZ et al., *Nature* (2016)
5) Chen Y & Suzuki K, *AI in Decision Support Systems* (2018)

Computer-Aided Diagnosis/Detection (CAD)



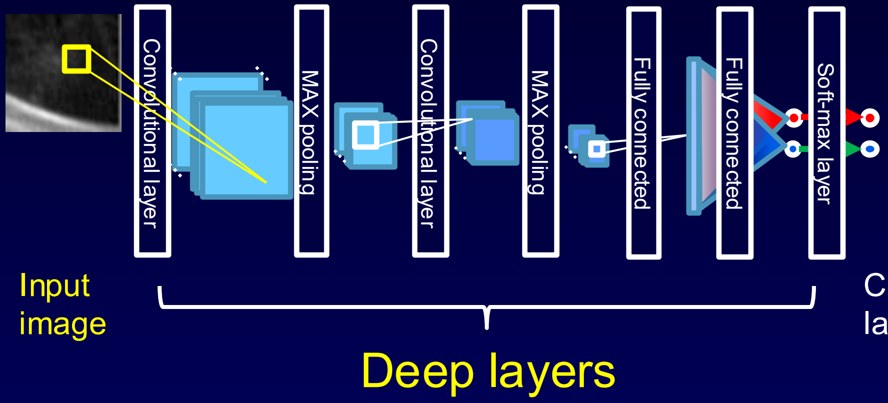
Automated Diagnosis

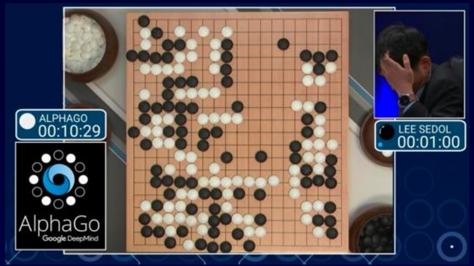


Machine Learning (ML) with Image Input: "Deep Learning"[1)]

- Deep learning: ML algorithms that attempt to model high-level representations of information processing in the brain by using deeply layered machine-learning architectures
- Input to the model is images, and the output is classes

Deep Convolution Neural Network

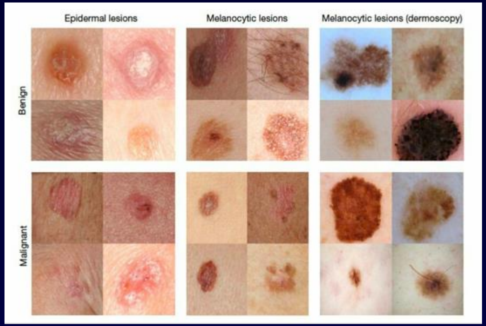




1) Y LeCun, et al. *Nature* (2015)

Deep Learning Application to Diagnosis of Disease on Medical Images

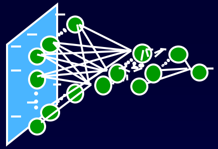
- Deep learning performance was equivalent to dermatologists' in diagnosis of skin cancer when it was trained with 130,000 medical + 1.28 M natural images (*Nature*, 2017）



Huge number of medical images

Problem

Deep Learning



Malignant

Benign

Dermatologists

Huge amount of work for annotations

✓ Deep learning training requires 10,000-100,000 cases per disease, which would take years to collect. Furthermore, it requires a huge amount of work for annotations of each of the images

Bottleneck of Deep Learning in Medicine

= Necessity of “Big Data"

(10,000 -100,000 cases)

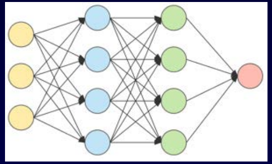
Method to Reduce the Big Data Issue

- Transfer Learning[1)]
  - Train a convolutional neural network (CNN) with 1.28M natural images; and then transfer the weights of the model and fine-tune it with 130,000 medical images

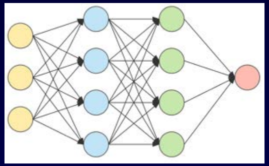
Nature (2017)

Transfer

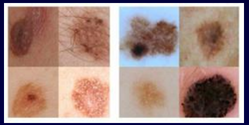
1.28M natural images



CNN



CNN

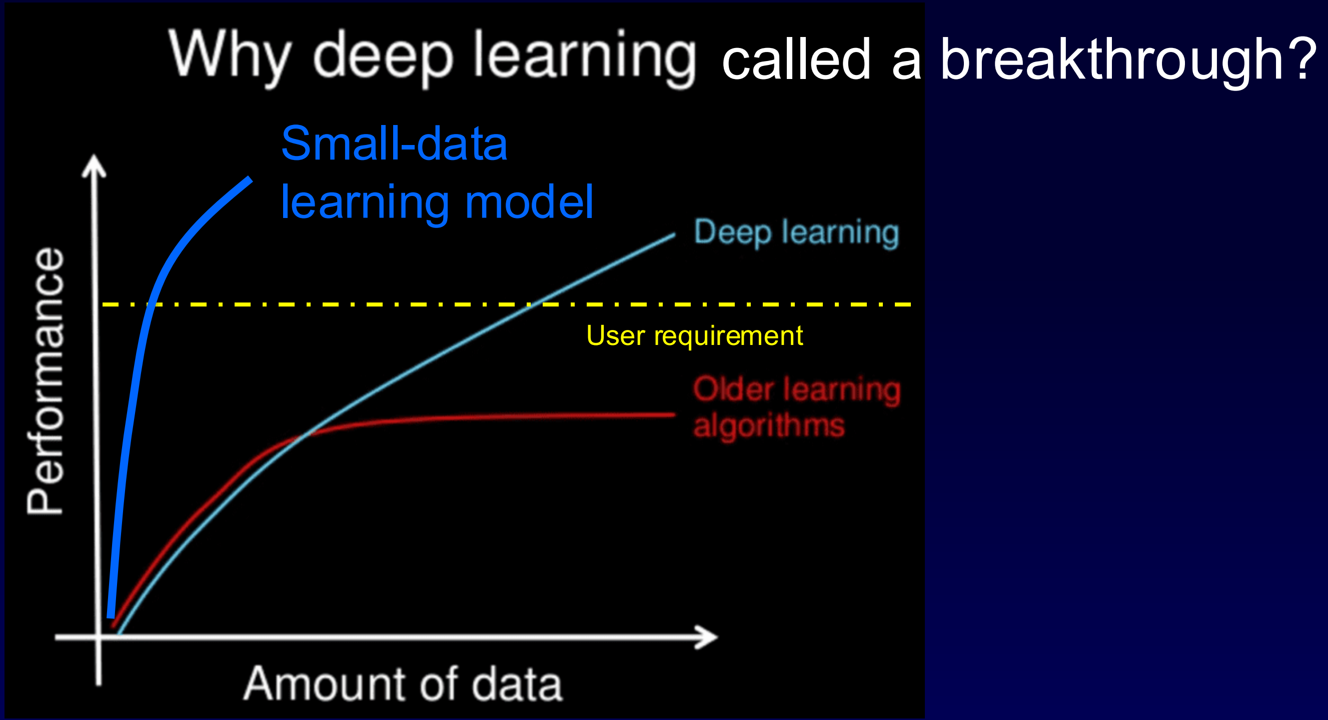


130K medical images

1) B Huynh et al. *J Med Imag* (2016)

“Small-Data Deep-Learning” Model[1)]

– Deep-learning model that can be trained with a small number of cases



Source: A. Ng (Stanford U), What Data Scientists Should Know about Deep Learning (slide 30), 2015; the small-data learning model & user requirement added by K Suzuki (Tokyo Tech)

1) K Suzuki, *Proc IEEE Big Data Services* (2022)

Is it possible to develop a deep-learning model that does not require 100,000 cases or transfer learning?

Small-Data MTANN:
Early "Deep Learning" Model

Noise reduction: Suzuki et al. ICSPAT (1996), Suzuki et al. Neural Proc Lett (2001), Suzuki et al. IEEE Trans Signal Proc (2002), Suzuki et al. IEICE Trans Info & Sys (2002), and Suzuki Neural Eng (2004)

Edge enhancement: Suzuki et al. IEEE Trans Pat Anal & Mach Intell (2003), Suzuki et al. IEEE Trans Med Imag (2004)

Lung CT CAD: Suzuki et al. *Med Phys* (2003), Arimura et al. *Acad Radiol* (2004), Li at al. *Radiology* (2005), Suzuki et al. *Acad Radiol* (2005), Suzuki et al. *IEEE Trans Med Imag* (2005), Suzuki et al. *Phys Med Biol* (2009)

CXR CAD: Suzuki et al. *Acad Radiol* (2005), Chen et al. *IEEE Trans Biomed Imag* (2013)

Bone separation (VDE) CXR: Suzuki et al. *IEEE Trans Med Imag* (2006), Oda et al. *AJR* (2009), Chen et al. *Med Phys* (2011), Chen et al. *IEEE Trans Med Imag* (2014), Chen et al. *Phys in Med & Biol* (2016)

CT colonography CAD: Suzuki et al. *Med Phys* (2006), (2008), (2010), Suzuki et al. *IEEE Trans Med Imag* (2010), Xu and Suzuki. *Med Phys* (2011)

Overview: Suzuki. *Int J Biomed Imag* (2012), Suzuki. *Quant Imag Med & Surg* (2012), Suzuki. *IEICE Trans Info & Sys* (2013), Suzuki. *J Med Imag & Info Sci* (2017), Suzuki. *Radiol Phys Tech* (2017), Suzuki. Med Imag Tech (2017)

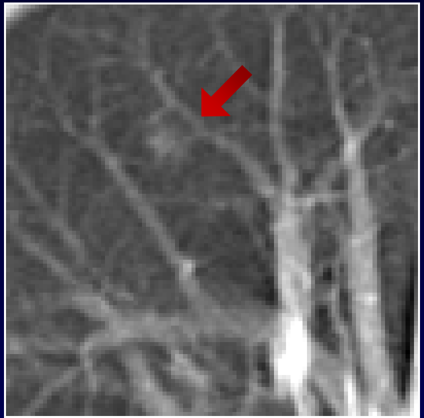
What is MTANN /émtæn/?

- Massive-training artificial neural network is
  - supervised image-processing / pattern-recognition technique based on machine learning (e.g., an artificial neural network)
  - One of earliest deep-learning models

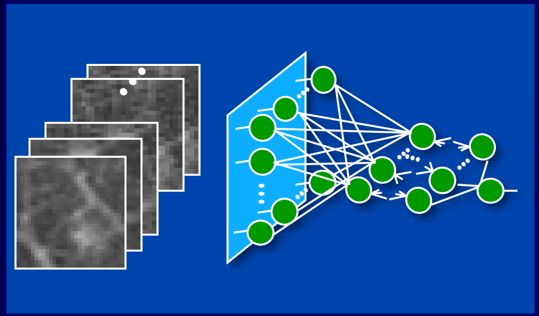
- **MTANN is award winning technology**
  - **Paul C Hodges Award from U of Chicago in 2002**
  - **Certificate of Merit Award at RSNA in 2003**
  - **Research Trainee Prize at RSNA in 2004**
  - **Young Investigator Award from Cancer Research Foundation in 2005**
  - **Certificate of Merit Award at RSNA in 2006**
  - **Certificate of Merit Award at RSNA in 2009**
  - **Best Paper Award, IEICE Journal in 2014**
  - **Most Cited Paper Award, EANM Springer-Nature in 2016**
  - **Most Citation Award, RPT Journal (Springer) in 2019**
  - **Award for Science & Technology, Ministry of Education, Culture, Sports, Science and Technology of Japan in 2021**

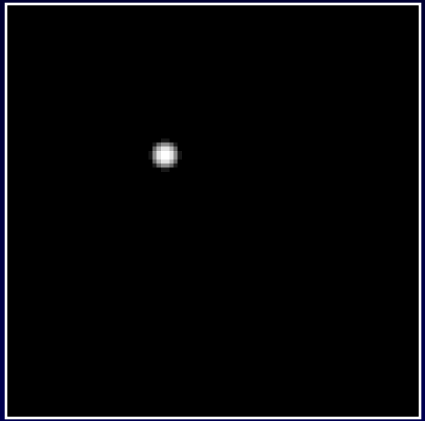
How does MTANN work?

- MTANN directly learns the relationship between input images and "teaching images." (c.f., other DL output is a class; e.g., cancer or not)
- MTANN enhances a specific pattern and suppresses other patterns in a medical image.



Input Image
(Arrow: Lung Cancer)

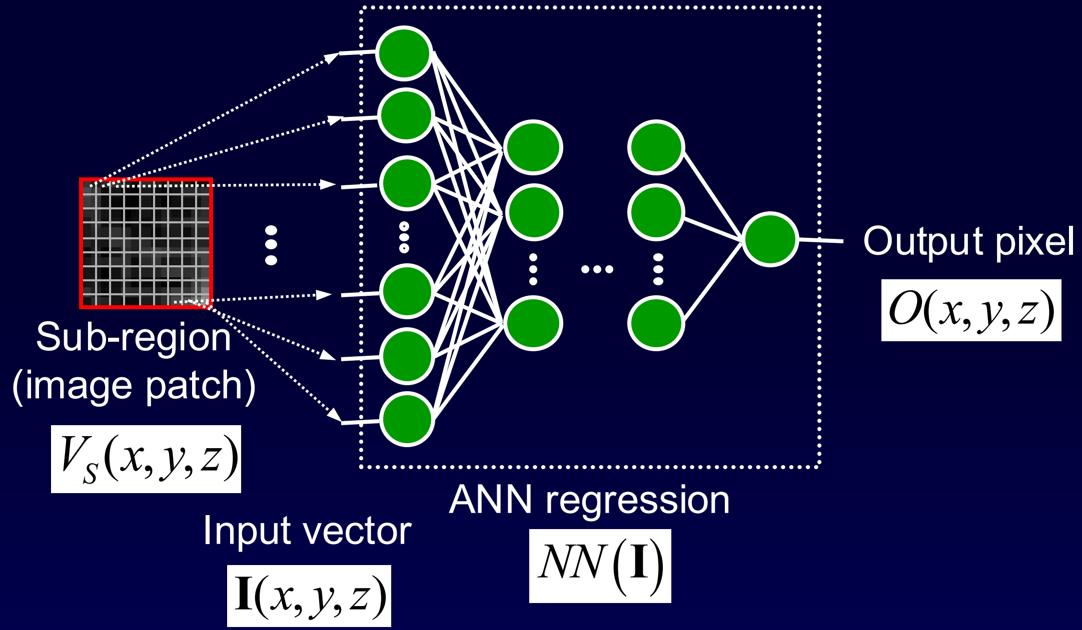




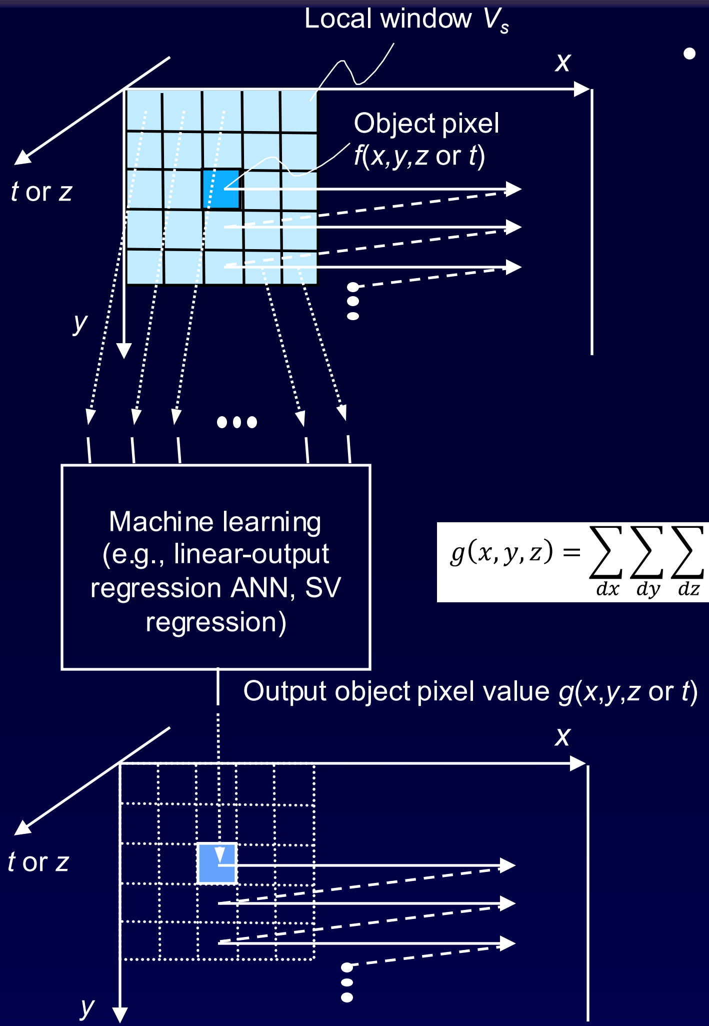
Teaching Image for
Enhancing Cancer

Architecture of MTANN
(Image-patch-based machine learning)

- Unlike deep CNN, pixel output (not category)



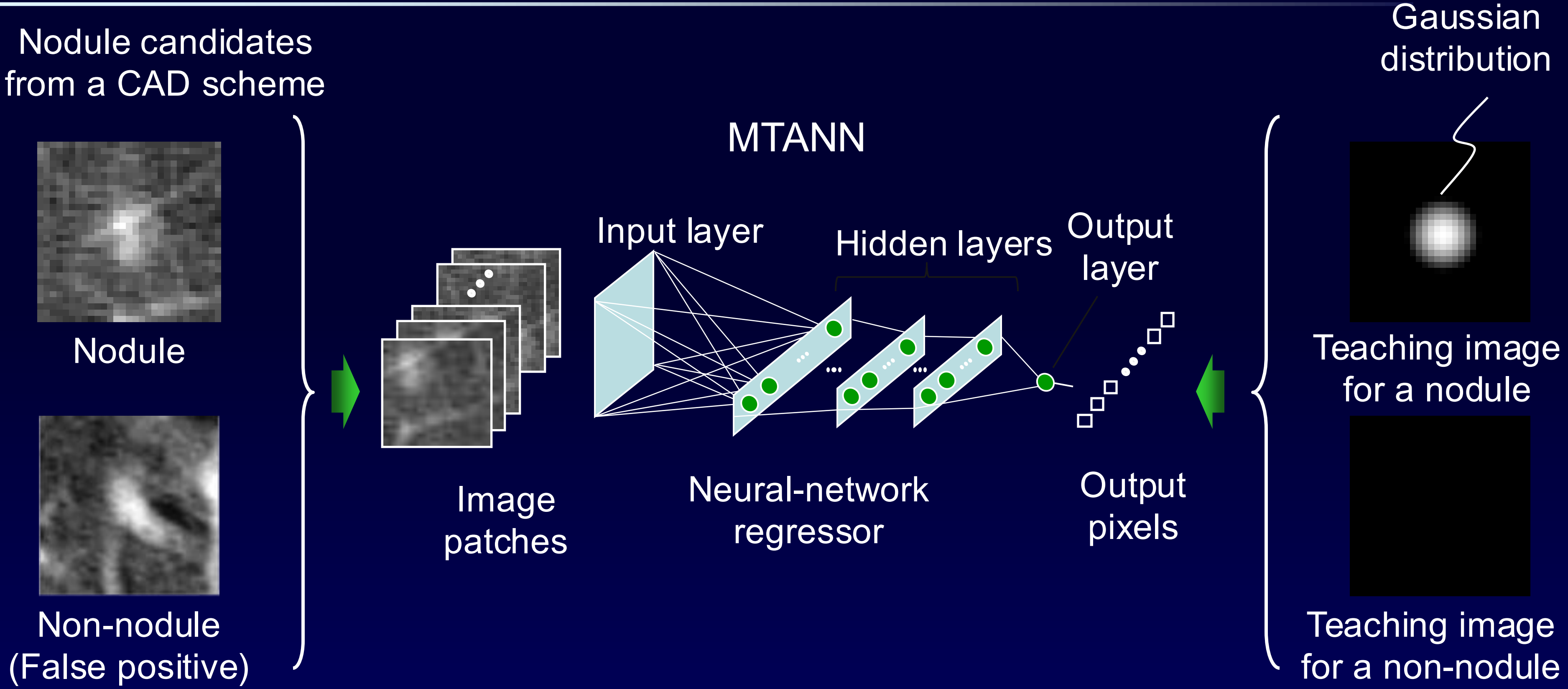
$$O(x, y, z) = NN\{\mathbf{I}(x-l, y-m, z-n) \mid l, m, n \in V_S\}$$



- Unlike convolutional neural network (CNN), convolution is done outside the network in the inference phase

$$g(x,y,z)=\sum_{dx}\sum_{dy}\sum_{dz}NN(dx,dy,dz)f(x+dx,y+dy,z+dz)$$

MTANN for Distinction between Nodules and Non-Nodules



# of training nodules and non-nodules: 10 and 90

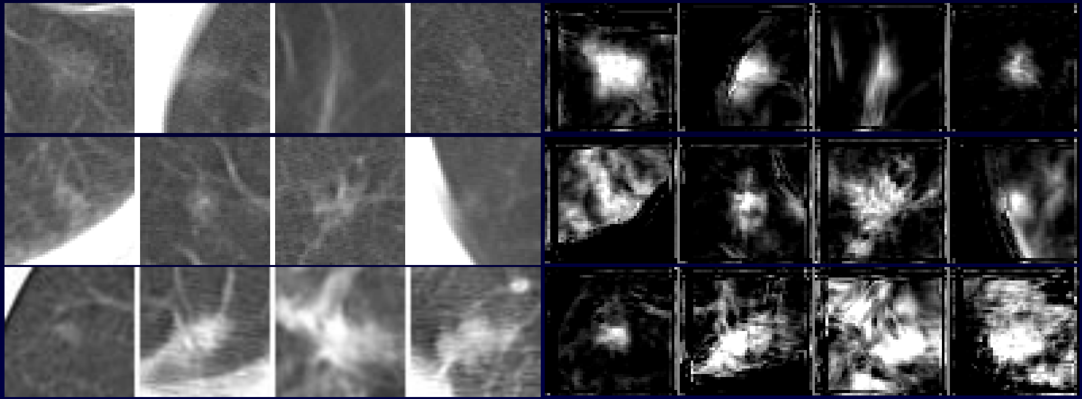
Output Images of MTANN for Non-Training Nodules

Nodule | Output image

Ground glass nodule

Part-solid nodule

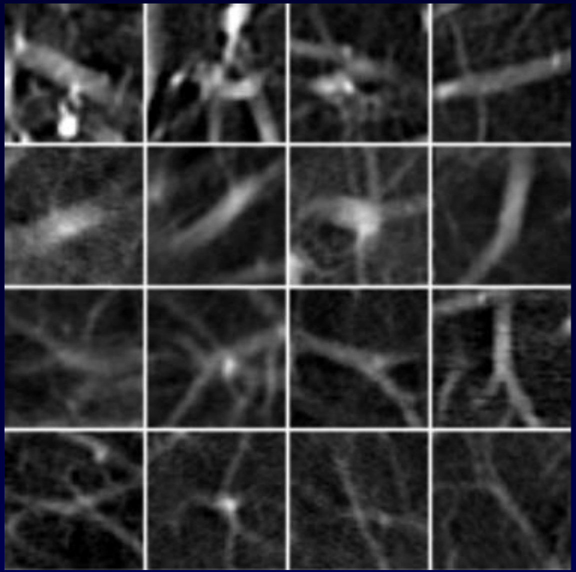
Solid nodule



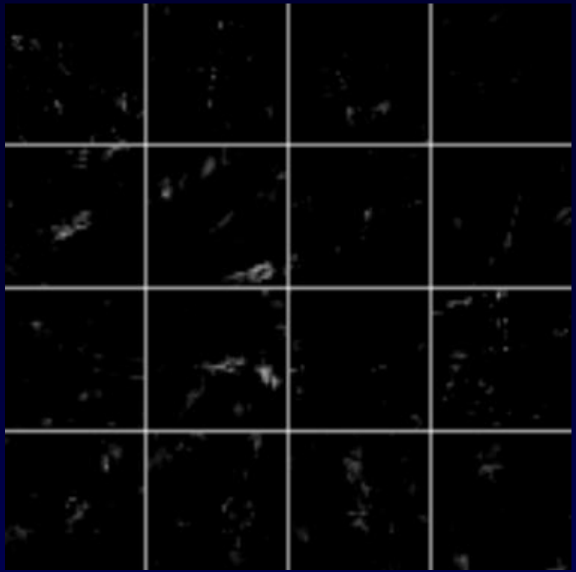
Suzuki K et al. *Med Phys* 2003

Output Images of MTANN for Various Non-Training Vessels

Different types of vessels



Output images



Output Images of Multi-MTANN for Various Types of Non-Nodules

Different types of non-nodules | **Output images**

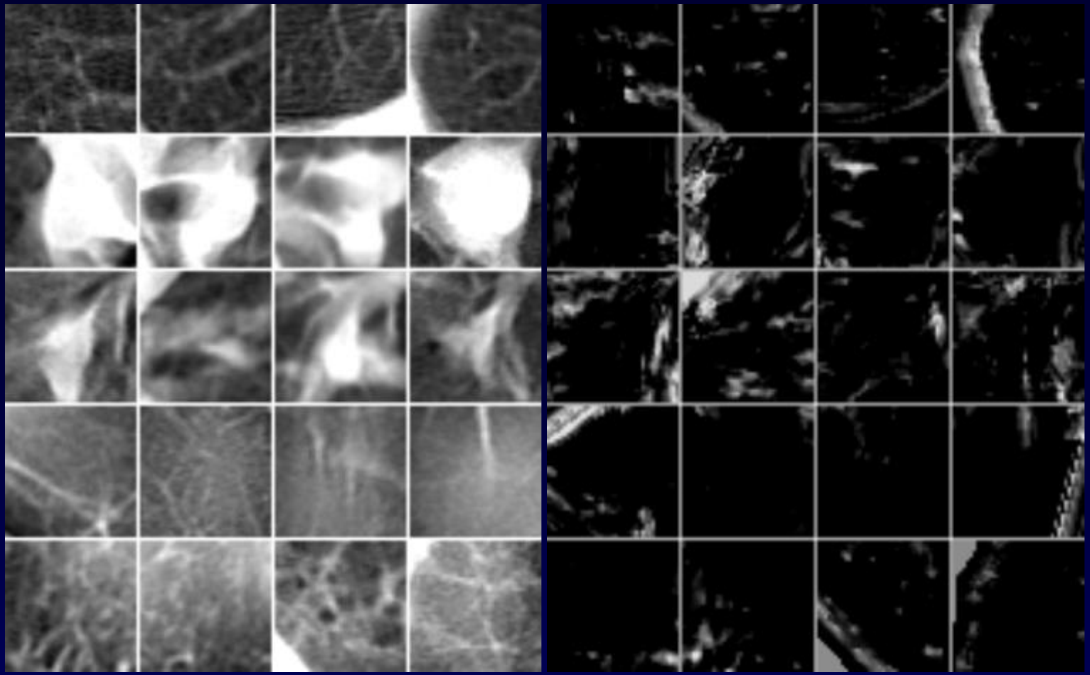
Periphera l vessels

Large vessels in the hilum

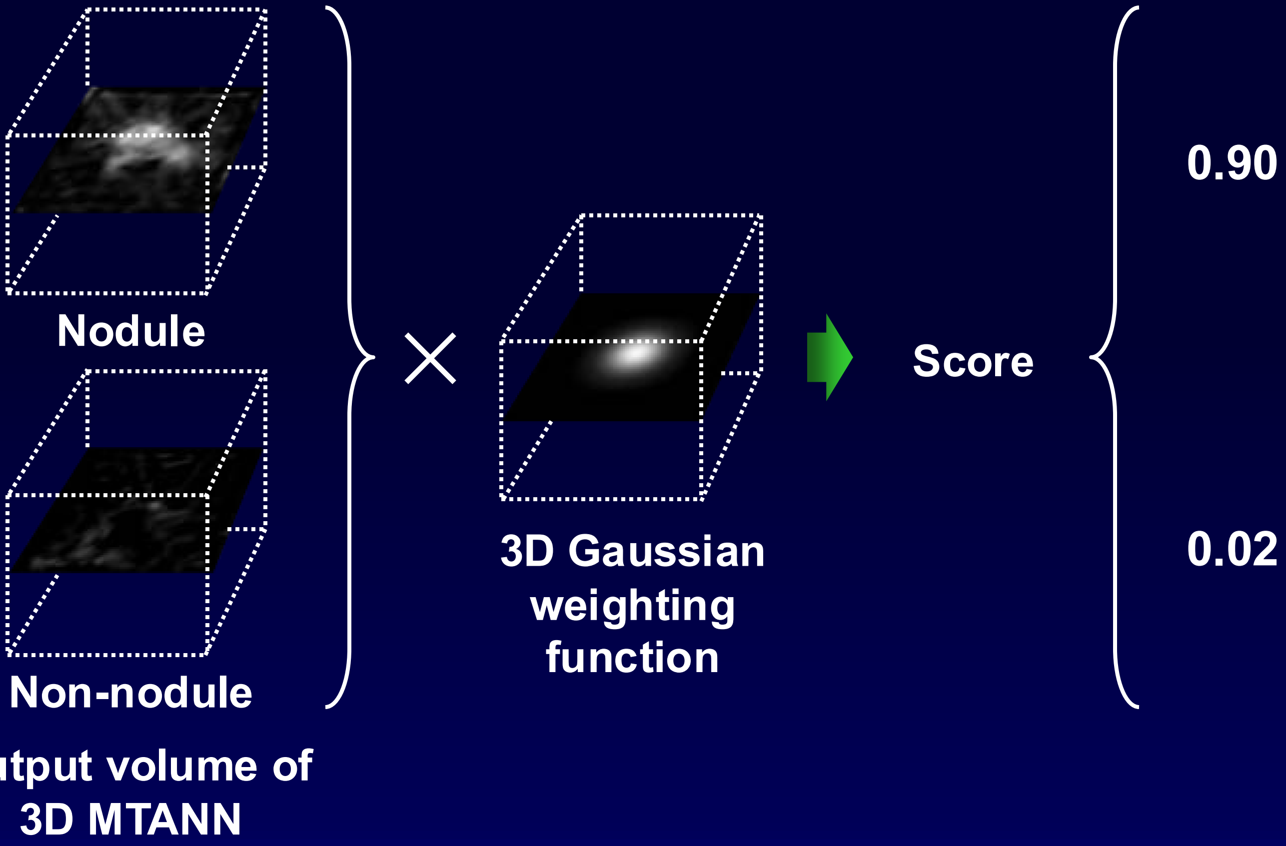
Vessels with some opacities

Soft-tissue opacities

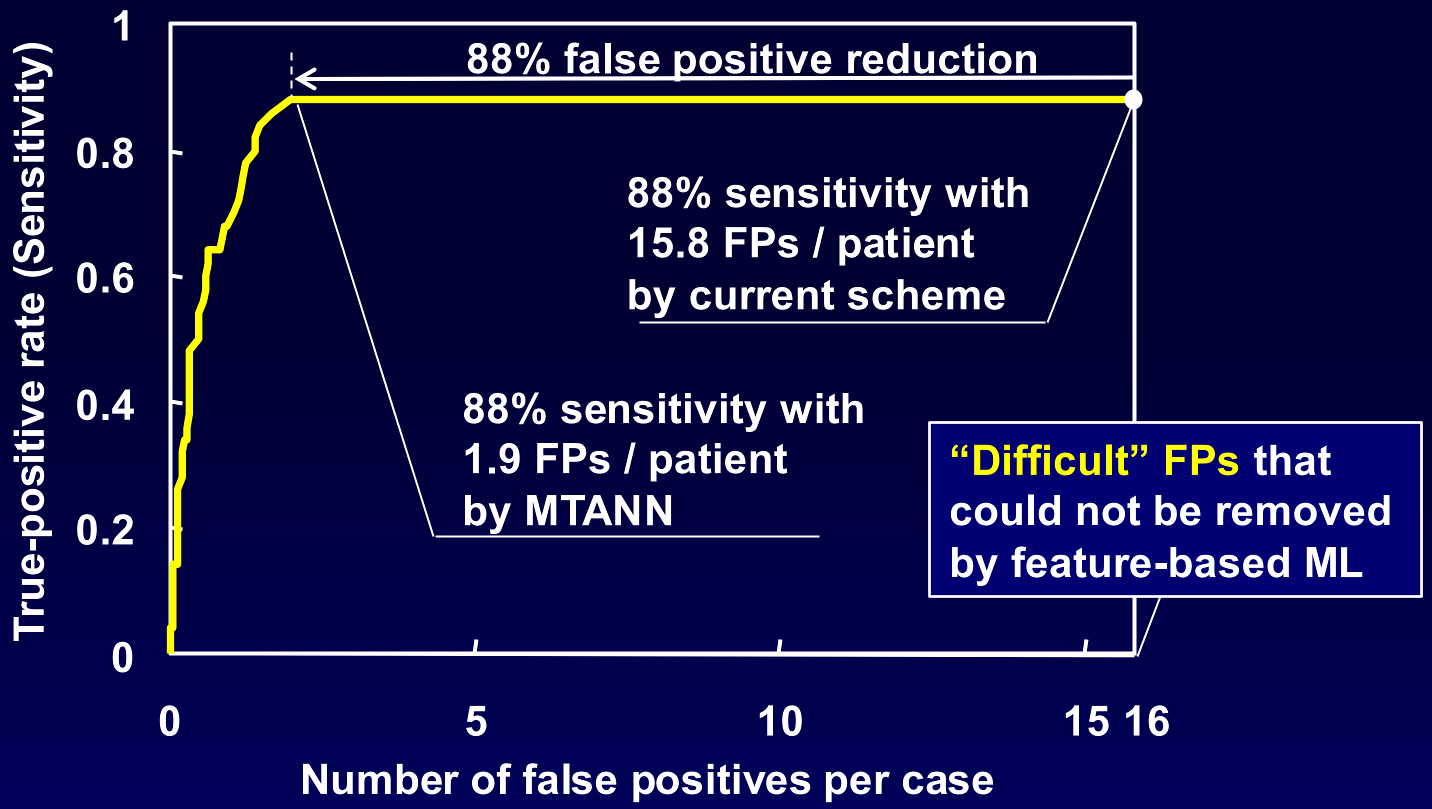
Abnormal opacities



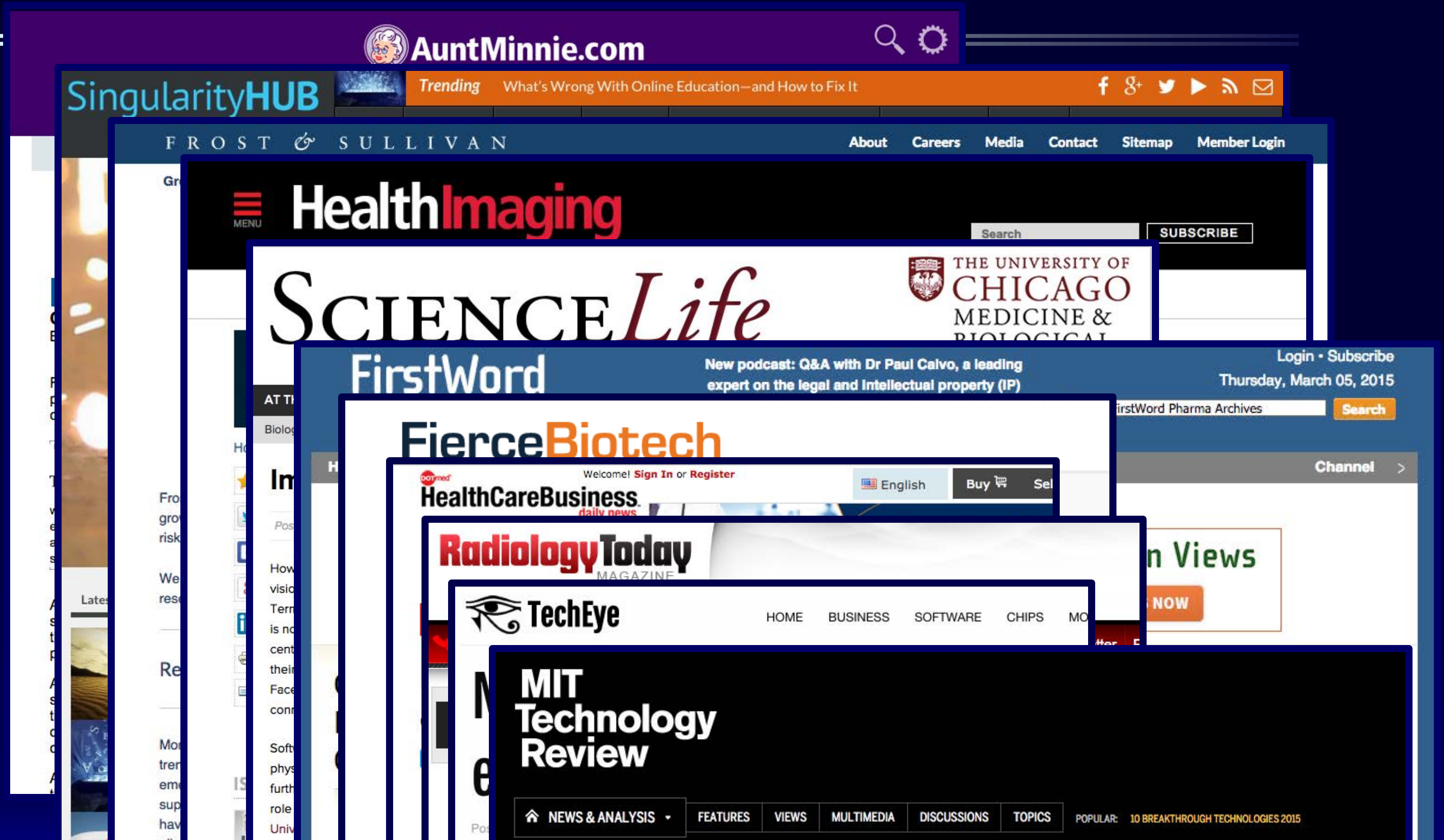
Scoring Layer for Distinction between Nodules and Non-nodules



Removal of False Positives by Use of MTANN



Press Coverage (selected from 49)

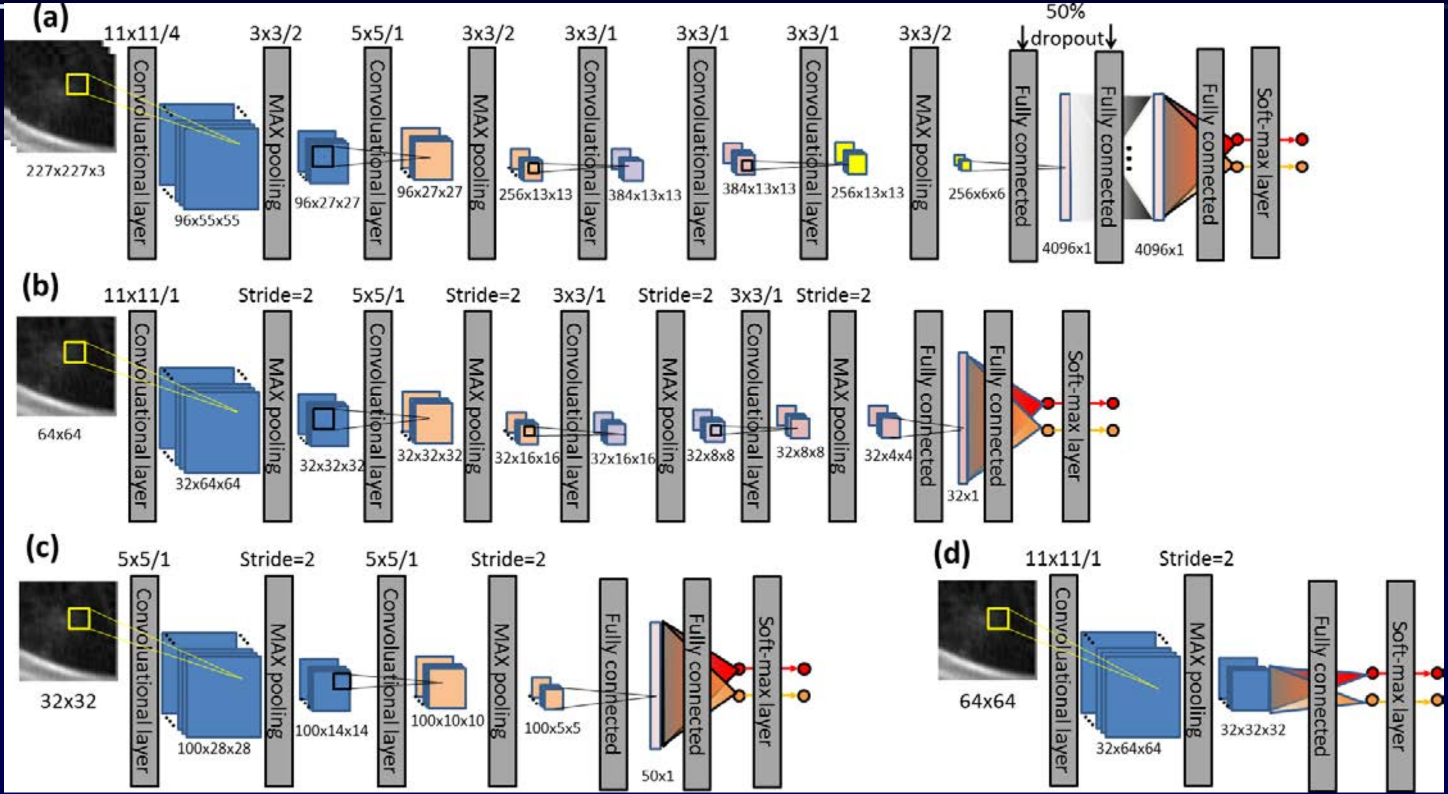


Comparing Two Classes of End-to-End Machine-Learning Models in Lung Nodule Detection and Classification:

MTANNs vs. CNNs[1)]

1) N Tajbakhsh & K Suzuki. *Pattern Recognition* (2016)

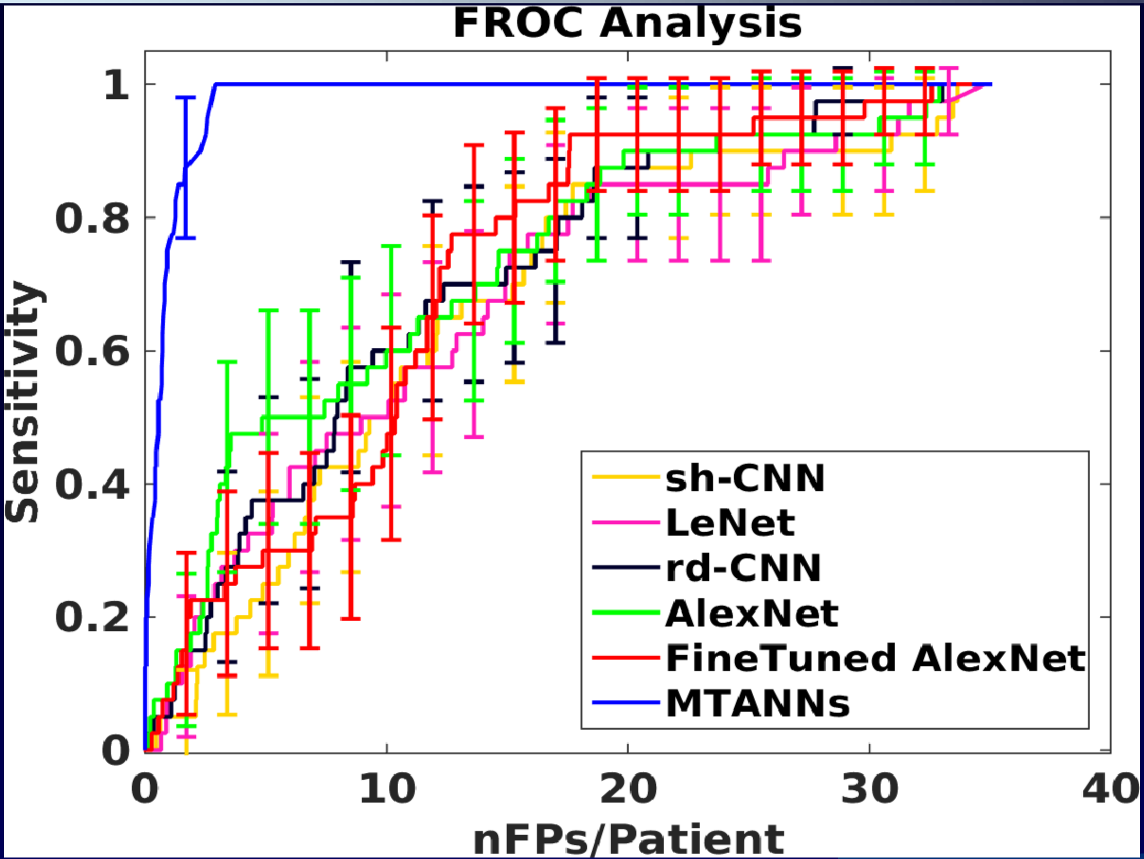
Reference Representative CNNs trained with 100 cases only



(a)Deep CNN (AlexNet)[1] (b) Relatively deep CNN (rd-CNN)
(c) LeNet[2] (d) Shallow CNN (s-CNN)

1) Alex Krizhevsky et al. NIPS (2012)
2) Y LeCun et al. Proc. IEEE (1998)

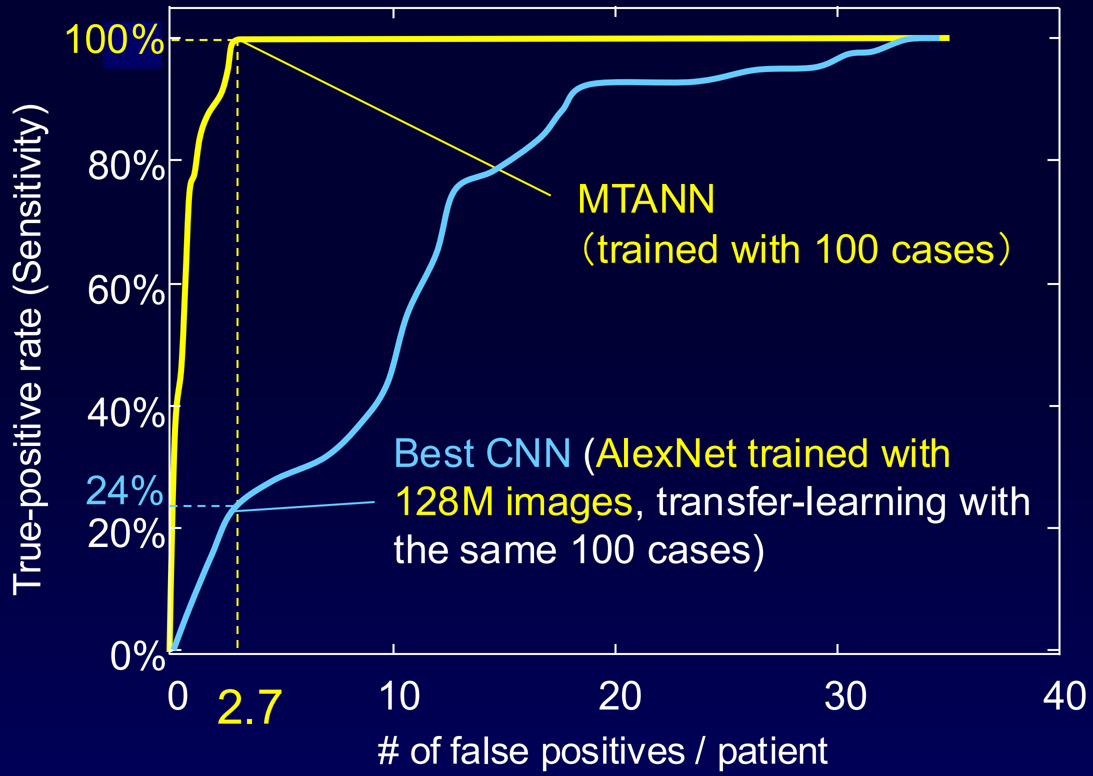
Comparison of MTANNs vs. CNNs: Lung Nodule Detection (Same # of 100 training cases)



N Tajbakhsh & K Suzuki. *Pattern Recognition* (2016)

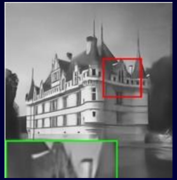
Comparison of MTANNs vs. CNNs with Small Data Learning

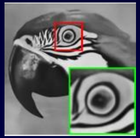
Lung Cancer Detection (training with 100 samples)



Advantages of MTANN Over Other Deep Learning Models

- Small required number of training samples
  - MTANN was trained with as small as 6 cases
- Low computational cost
  - half an hour to train, 1 sec. to execute on GPU
- Easy design of the architecture
  - easy to design the architecture and stable
- Stable in training
  - training is very stable, robust against parameter changes





	Required # of training samples	Training time	Performance
MTANN	10～100	< 10 min.	Higher
Other DL	5k～10k	a dozen hours to several days	Medium ～ High

Institute of SCIENCE TOKYO

Super-Efficient AI for Lung Nodule Classification in CT Based on Small-Data Massive-Training Artificial Neural Network (MTANN)

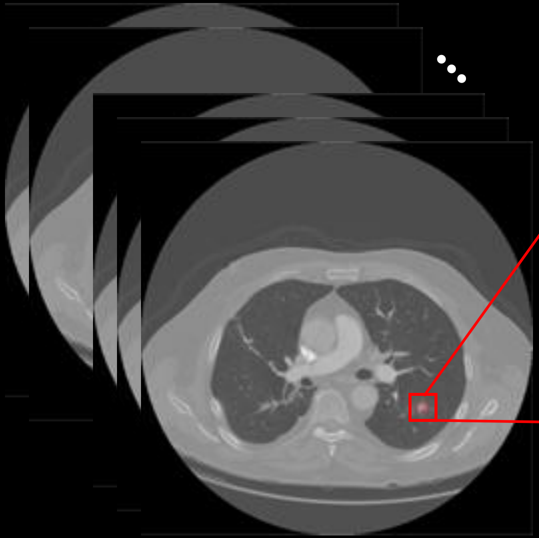
Shogo Kodera[1] B.E., Chavoshian Seyed Mohammad[1] B.E., Ze Jin[1], Ph.D., Takeyuki Watadani[2] M.D., Ph.D., Osamu Abe[2] M.D., Ph.D., and Kenji Suzuki[1], Ph.D.

1) BioMedical Artificial Intelligence (BMAI) Unit
Institute of Integrated Research (IIR)
Institute of Science Tokyo
2) Graduate School of Medicine and Faculty of Medicine
The University of Tokyo

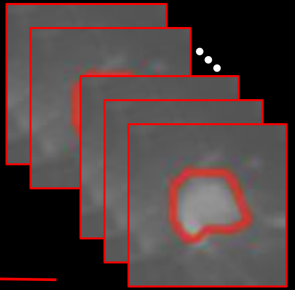
BMAI Biomedical Artificial Intelligence Research Unit

Kodera S, Chavoshian SM, Jin Z, Watadani T, Abe O, Suzuki K, RSNA, 2024

Workflow of Lung Nodule Classification



CT Image (Interpolated)
LIDC-IDRI database
512 x 512 x N



Volume of Interest (VOI) extracted from CT. Lung nodule located at the center of VOI (51 x 51 x 51)

Data preprocessing

DNN-based classification

Classify VOIs with trained deep neural network

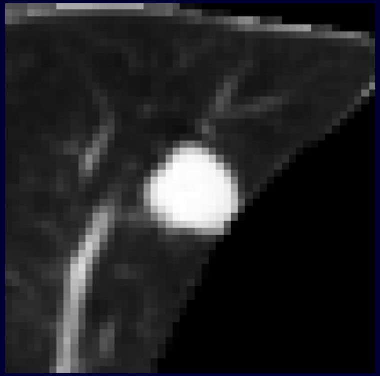
benign

malignant

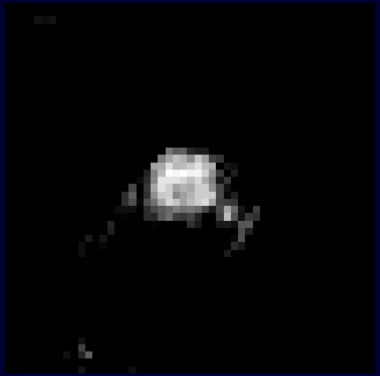
Results -Solid Cases-

- Small-data* MTANN emphasized a malignant nodule and suppressed a benign nodule.

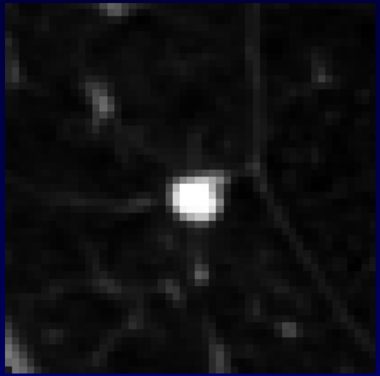
Malignant nodule



Output for malignant nodule



Benign nodule



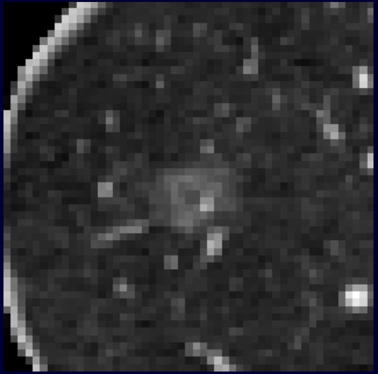
Output for benign nodule

*The MTANN was trained with only 38 malignant samples and 30 benign samples

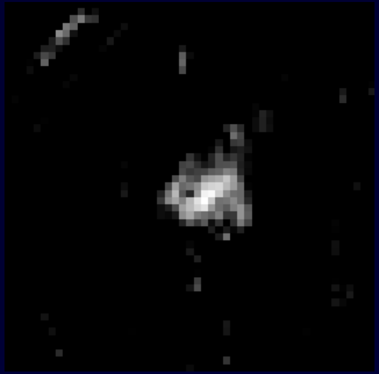
Results –Ground Glass Cases-

- Small-data* MTANN emphasized a malignant nodule and suppressed a benign nodule.

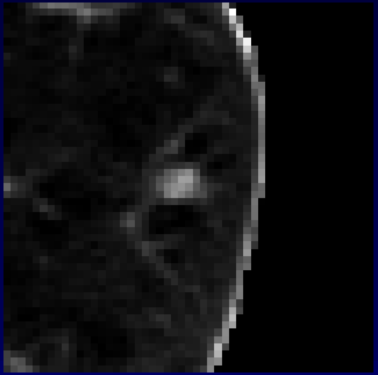
Malignant nodule



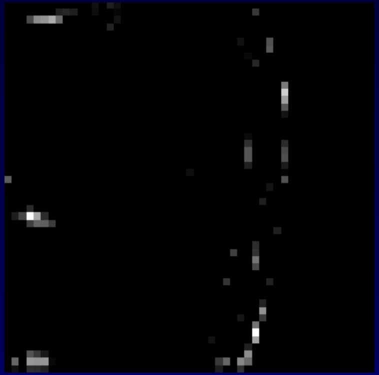
Output for malignant nodule



Benign nodule



Output for benign nodule



*The MTANN was trained with only 38 malignant samples and 30 benign samples

Performance & Efficiency in Lung Nodule Classification

Table 2. Comparison the performance (in terms of AUROC) and efficiency (in terms of the number of parameters and inference time) with the state-of-the-art models.

Models	Experiment Training cases: 68	Number of parameters	Inference time (per case)
Vision Transformer+	0.531 ± 0.065*	304M	23.0 ms
3D MTANN (ours)	**0.919 ± 0.004**	**0.017M**[++]	**8.15 ms**

Computational cost: 1/17800

NVIDIA

The experiments in this study were carried out using the NVIDIA RTX A4000 GPU.
+Transfer learning with 1.28M non-medical images in ImageNet 1K.
*Statistically significant ($P<0.05$) against our 3D MTANN.
[++]only 1/2100-1/18000 memory size is required.

Efficiency of Our MTANN Model in Lung Nodule Classification

Table 3. Comparison the calculating time between GPU server and CPU.

Time	NVIDIA GPU Server*	MacBook Air (CPU)**
Training time (68 cases)	4.5 sec.	50 sec.
Testing time (616 cases)	10 sec.	29.0 sec.
Inference time (/case)	**8.15 ms**	**46.5 ms**

*The experiments in this study were carried out using the NVIDIA RTX A4000 GPU.
**MacBook Air with M1 chip and 16GB memory.

Kodera S, Chavoshian SM, Jin Z, Watadani T, Abe O, Suzuki K, RSNA, 2024
received RSNA Cum Laude Award

Cum Laude Award at RSNA 2024

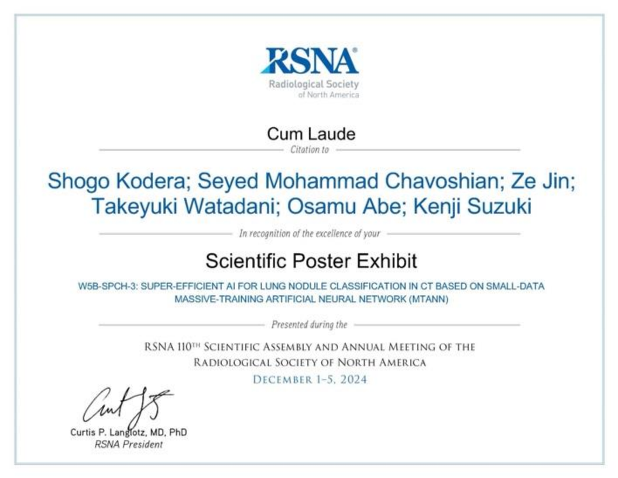
Two papers got awarded by RSNA 2024

Our two papers referenced below received the highest honor, Magna Cum Laude Award and the prestigious Cum Laude Award at RSNA 2024!

RSNA is a top conference in the clinical field of medical imaging with a very narrow acceptance rate of about 25% and attracts top radiologists and medical imaging researchers from all over the world.
At this year's RSNA, there were 1,312 poster presentations that passed through that narrow gate, 6 for Magna Cum Laude (0.46% chance of winning) and 19 for Cum Laude (1.45% chance of winning).
It is extremely rare and invaluable to medical imaging researchers like us to win the highest award of the conference, the Magna Cum Laude Award or the Cum Laude Award that is the second highest award at the prestigious RSNA conference.

Prof. Suzuki commented:
"I participated in and presenting at the RSNA in the past 24 years, but the Magna Cum Laude Award and Cum Laude Award were given to clinical research presented by radiologists, and I have merely seen that research by medical imaging researchers received these awards. Moreover, I don't recall that master students have won either of these awards; and thus, I am certain that those are great accomplishments. I would like to send my highest compliments to these students of mine."

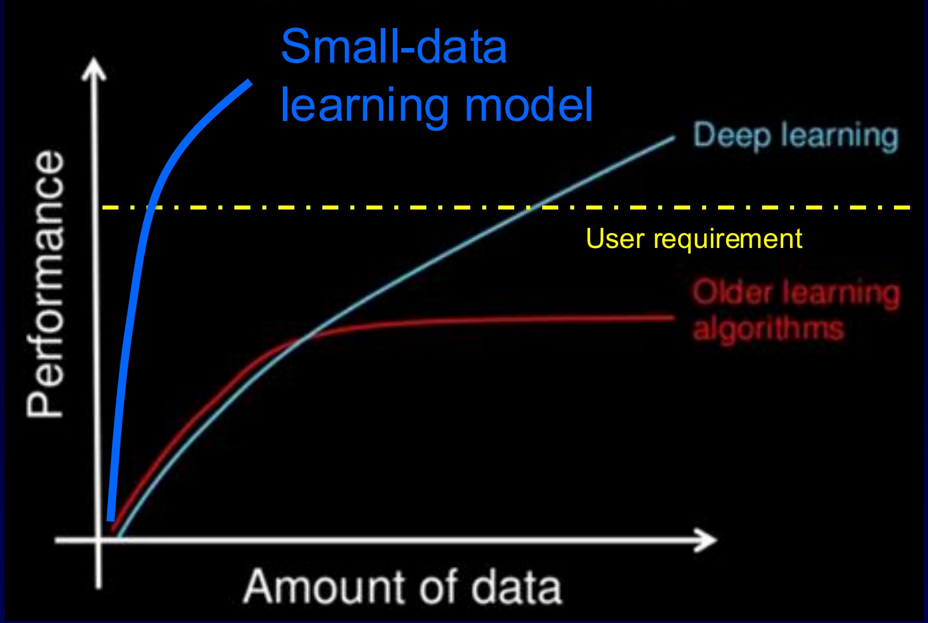
Magna Cum Laude	Qu T., Yang Y., Jin Z., and Suzuki K.: Annotation-free AI learning of lung nodule segmentation in CT using weakly-supervised massive-training artificial neural networks
Cum Laude	Kodera S., Chavoshian S. M., Jin Z., Watadani T., Abe O., and Suzuki K.: Super-efficient AI for lung nodule classification in CT based on small-data massive-training artificial neural network (MTANN)



"Small-Data Deep-Learning" Model[1)]

– Deep-learning model that can be trained with a small number of cases

Why deep learning called a breakthrough?

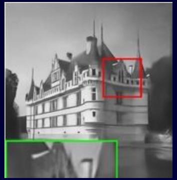


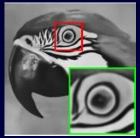
Source: A. Ng (Stanford U), What Data Scientists Should Know about Deep Learning (slide 30), 2015;
t he small-data learning model & user requirement added by K Suzuki (Tokyo Tech)

1) K Suzuki, *Proc IEEE Big Data Services* (2022)

Advantages of MTANN Over Other Deep Learning Models

- Small required number of training samples
  - MTANN was trained with as small as 6 cases
- Low computational cost
  - half an hour to train, 1 sec. to execute on GPU
- Easy design of the architecture
  - easy to design the architecture and stable
- Stable in training
  - training is very stable, robust against parameter changes

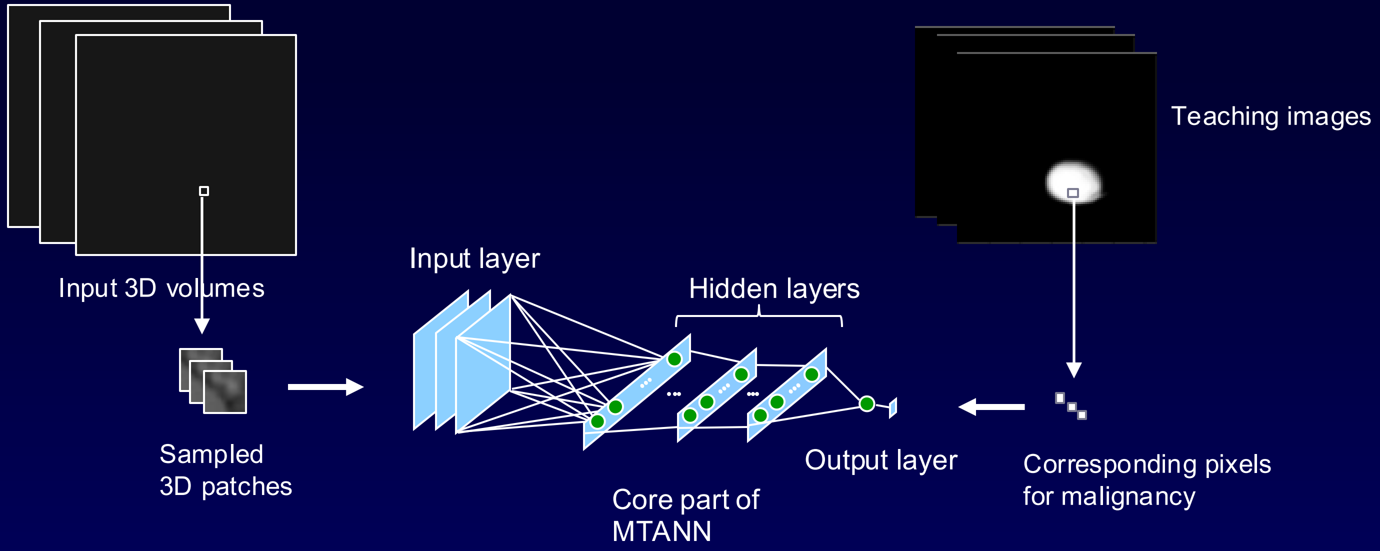




	Required # of training samples	Training time	Performance
MTANN	10～100	< 10 min.	Higher
Other DL	5k～10k	a dozen hours to several days	Medium ～ High

3D MTANN for Rare Cancer Diagnosis

- We trained a 3D MTANN to distinguish rare cancer (soft-tissue sarcoma) and benign tumors in femur T2w MRI
  - Training: 40 malignant tumors (soft-tissue sarcoma) + 40 benign tumors



Performance Comparison with State-of-the-art Transfer-learned Deep Learning Models in Distinction between Benign and Malignant Tumors

Model	AUC
AlexNet	0.59*
VGG-16	0.63*
ResNet-50	0.66*
MTANN	**0.78**

*p-value < 0.05 against MTANN
AUC: Area under the ROC curve

Y Yang,…, K Suzuki, *RSNA* (2023)

Virtual Deep-Learning/AI Imaging

1. Separation of Ribs from Soft Tissue in Chest Radiographs by Using MTANN
2. Radiation dose reduction in CT and mammography by Using MTANN

1-6) Suzuki et al. *IEEE Trans Med Imag* (IF:10.0) (2006), Oda et al. *AJR* (IF:4.0) (2009), Chen et al. *Med Phys* (IF:4.1) (2011), Chen et al. *IEEE Trans Med Imag* (IF:10.0) (2014), Chen et al. *Phys in Med & Biol* (IF:3.6) (2016), Zarshena et al. *Med Phys* (IF:4.1) (2019)

What Is Virtual Deep-Learning/AI Imaging?

- Definition*: Virtual deep-learning/AI imaging may be defined as imaging for enhancing/suppressing physical/semantic objects and/or materials by using deep learning regression
  - It may realize specific imaging virtually without certain physical equipment/materials/procedure

*K Suzuki, Tokyo Tech, Japan

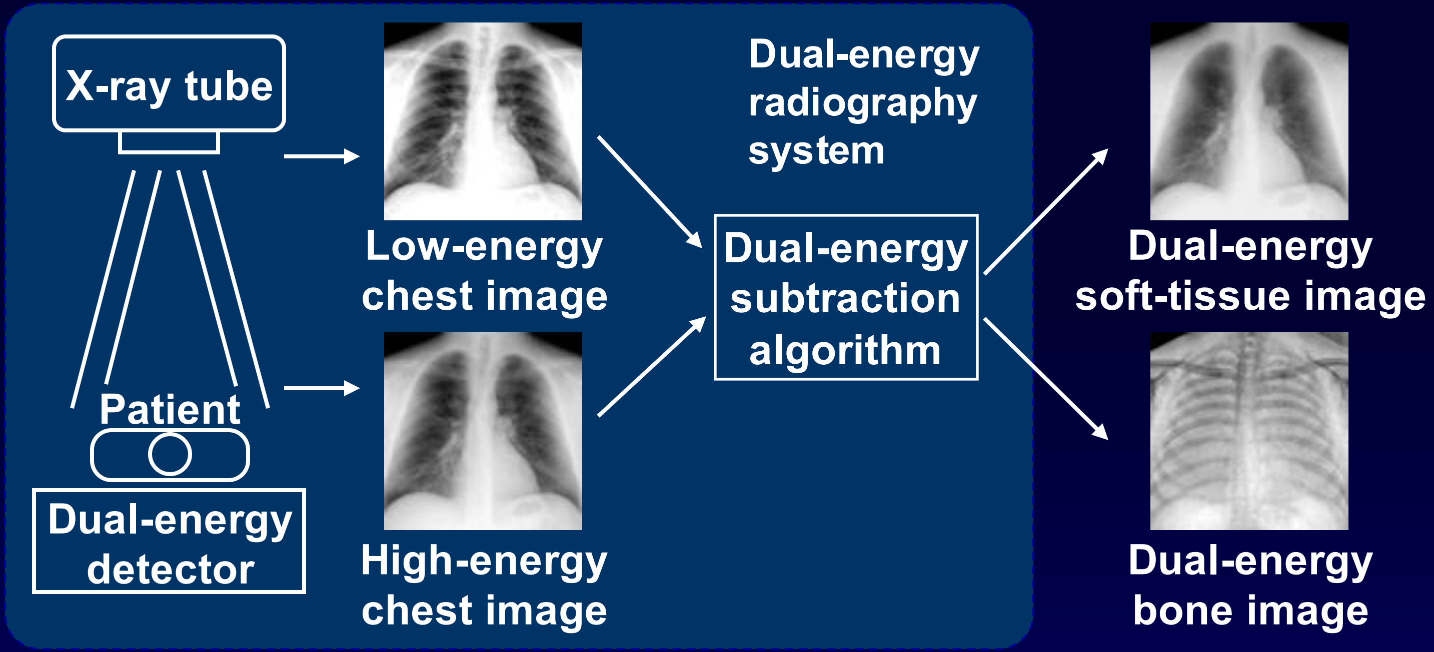
1-6) Suzuki et al. *IEEE Trans Med Imag* (IF:10.0) (2006), Oda et al. *AJR* (IF:4.0) (2009), Chen et al. *Med Phys* (IF:4.1) (2011), Chen et al. *IEEE Trans Med Imag* (IF:10.0) (2014), Chen et al. *Phys in Med & Biol* (IF:3.6) (2016), Zarshena et al. *Med Phys* (IF:4.1) (2019)

Motivation

- In one study[1)], more than 80% of the missed lung cancers by radiologists in CXR were partly obscured by overlying bones.

1) Austin et al. *Radiology* (1992)

Dual-Energy Subtraction Imaging



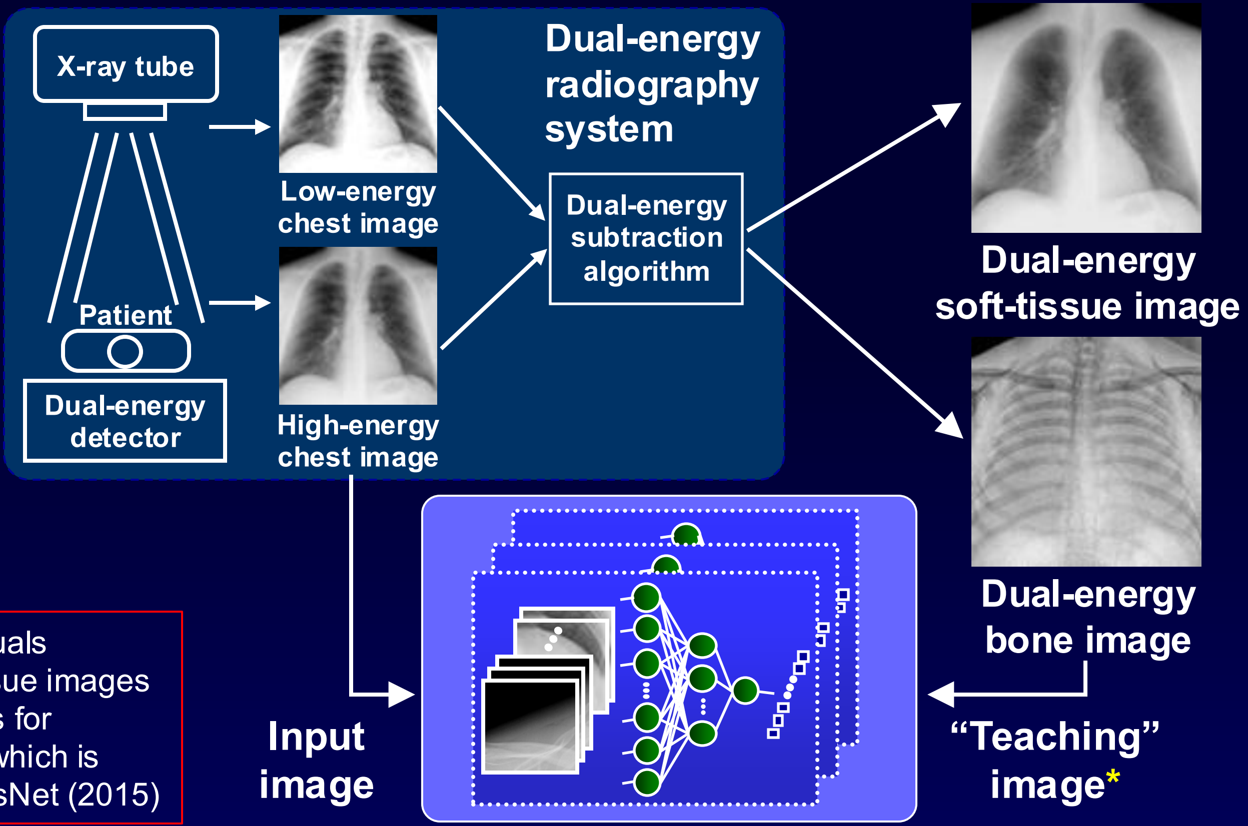
Kido S Et Al *Radiology* 1991, 1992

Limitation of Dual-Energy Imaging

- Despite great advantages, a limited number of hospitals use dual-energy radiography systems at present, probably because
  - Specialized equipment is required for obtaining dual-energy x-ray exposures,
  - Radiation dose may be greater than that for standard chest radiography,
  - Subtraction of high and low energy images causes an increased noise level.

Can we separate bones from soft tissue in CXR by software?

Basic Principle of "Virtual Dual-energy" Radiography: Training Step



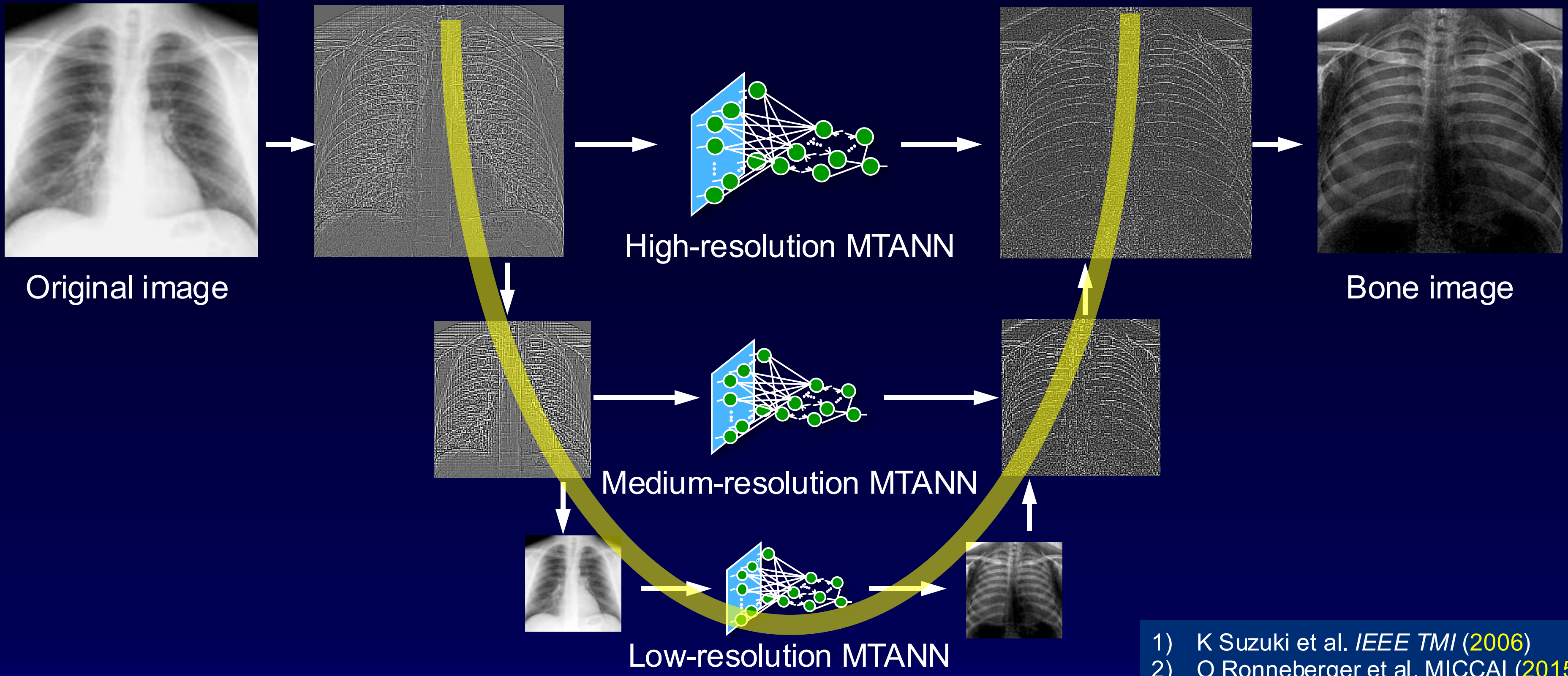
* Used the residuals between soft-tissue images and input images for training (2004), which is equivalent to ResNet (2015)

# of Training images: 4 pairs

Multi-resolution MTANN

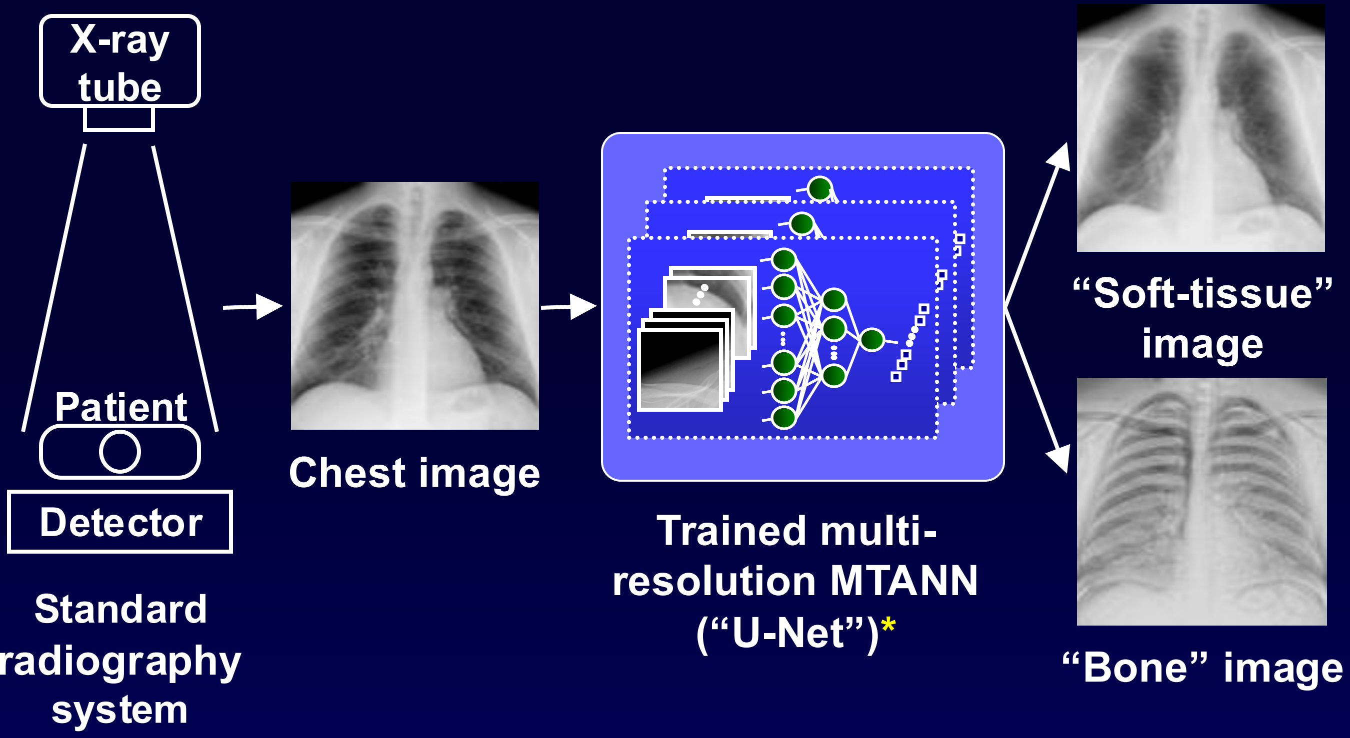
Multi-resolution MTANN[1)]

- Multi-resolution MTAN[1)] has a U-structure[2)] with decomposition (pooling) and reconstruction (up-sampling)



1) K Suzuki et al. *IEEE TMI* (2006)
2) O Ronneberger et al. MICCAI (2015)

Basic Principle of "Virtual Dual-energy" Radiography:
Application Step



* Multi-resolution MTANN[1] has U-Net[2]-like structure with decomposition (pooling) and reconstruction (up-sampling)

1) K Suzuki et al. *IEEE TMI* (2006)
2) O Ronneberger et al. MICCAI (2015)

Result of Separation of Ribs from Soft Tissue in CXR

Output image of MTANN

Suzuki K, Abe H, MacMahon H, Doi K *IEEE Trans Med Imag* 2006

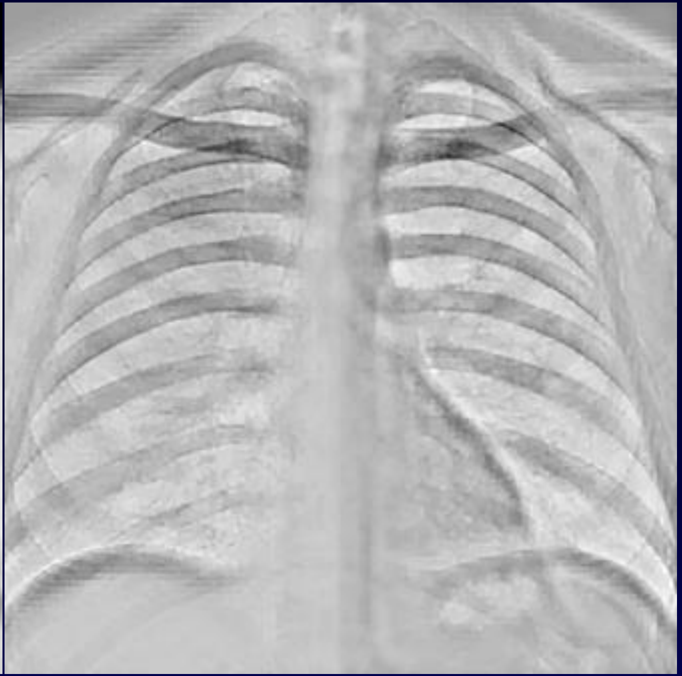
Rib Suppression by MTANN

Original chest image

MTANN soft-tissue image

Rib Separation by MTANN

Original chest image



MTANN bone image

Comparison with Dual-Energy Soft-Tissue Image

MTANN
soft-tissue image

"Gold-standard" dual-energy
soft-tissue image

Comparison of MTANN Soft-tissue Image with Dual-energy Soft-tissue Image[1)]

Original CXR

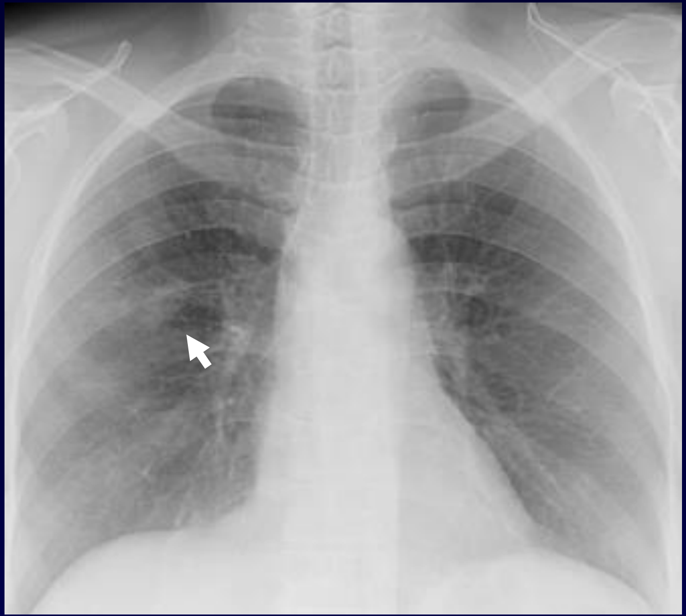
MTANN soft-tissue image

“Gold standard” dual-energy soft-tissue image

1) Suzuki K, et al. *IEEE Trans Med Imag* (2006)
2) Yang W, et al. *Med Imag Anal* (2017)

Results for Cancer Cases

Improved Conspicuity of Nodule with MTANN



Original chest image with nodule

Our MTANN soft-tissue image

Chen S, Suzuki K. *IEEE TMI* 2014

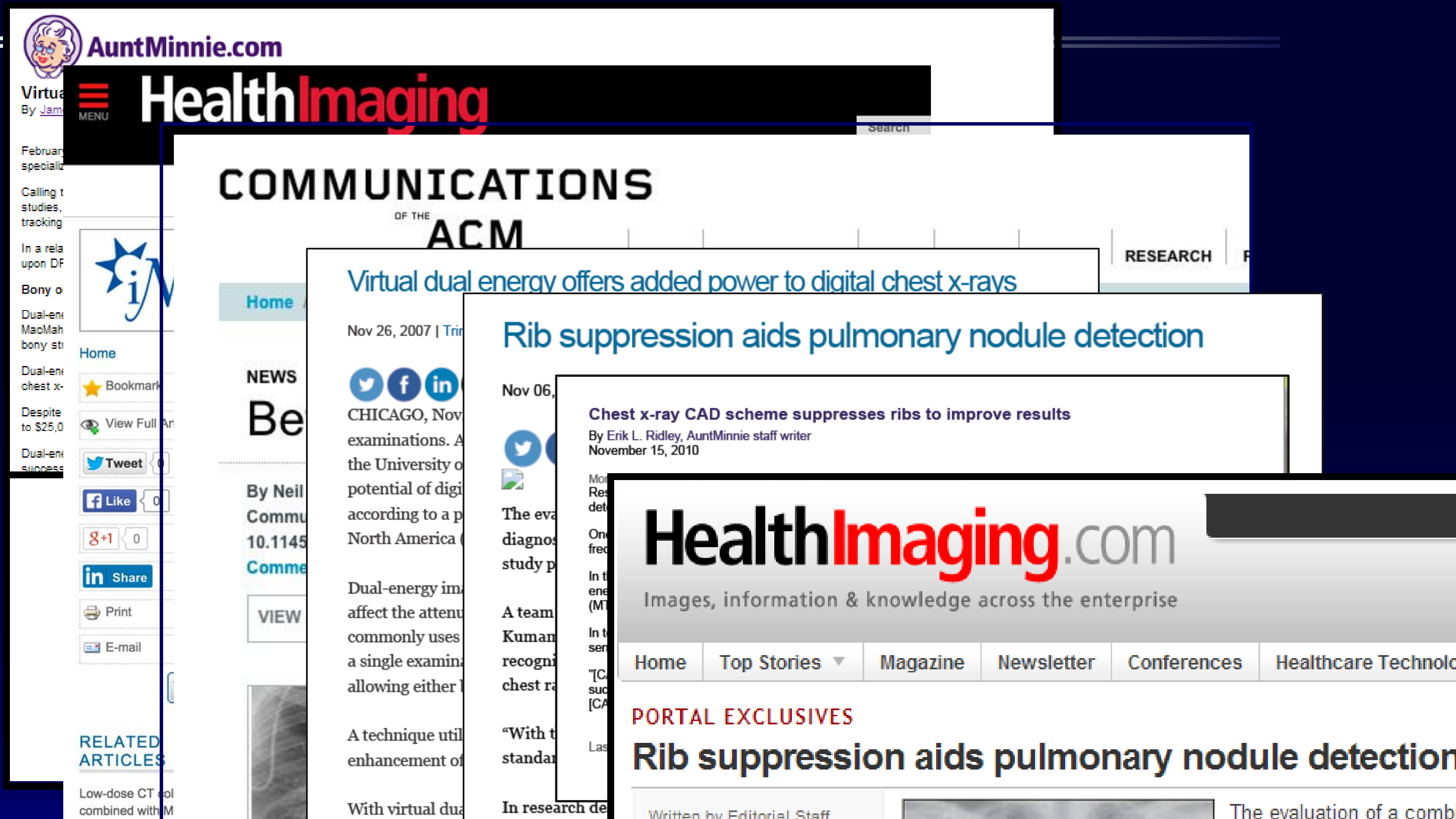
Comparison of Our MTANN "Virtual Dual-Energy" Imaging with Conventional Dual-Energy Imaging

- Our technique for separation of ribs from soft tissue requires:

  1. **No specialized equipment,**
  2. **No additional radiation dose to patients,**
  3. **But only software.**

- Our technique is applicable to any chest radiographs acquired with a standard radiography system ($3B global market).

Press Coverage



Radiation Dose Reduction in CT: Toward Sub-mSv CT Exams

in collaboration with Kazuo Awai[b)], MD, PhD, Wataru Fukumoto, MD, Toru Higaki, PhD

Department of Radiology
Hiroshima University, Japan

Tokyo Tech

1) K Suzuki et al. *RSNA* (2012)
2) K Suzuki, et al. LNCS-MLMI (2017)

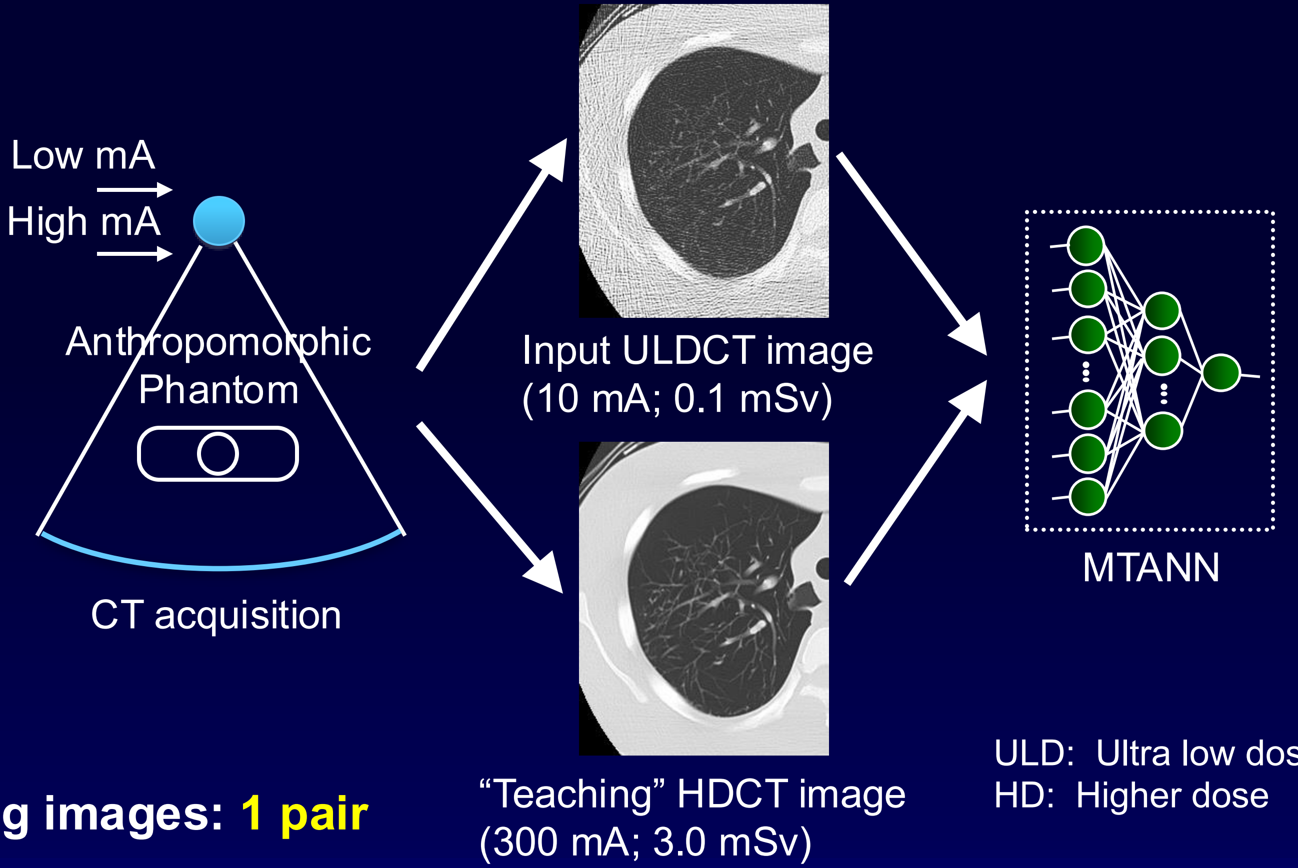
Radiation Dose Reduction in CT: Motivation

- CT scanners can expose patients to cumulative radiation doses which may elevate individuals' lifetime risk of developing cancer
- Studies[1-3] estimated
  - CT scans in the U.S. might be responsible for up to 1.5-2.0% of cancers
  - CT scans performed in the U.S. in 2007 alone would result in 29,000 new cancer cases in future years
  - CT scans of children each year would cause 4,870 future cancers in the U.S.

1) DJ Brenner et al. *N. Engl. J. Med.* (2007)
2) AB de González et al. *JAMA Intern. Med.* (2009)
3) DL Miglioretti et al. *JAMA Pediatrics* (2013)

Would you be happy if we could reduce the radiation dose in CT scans by 90%?

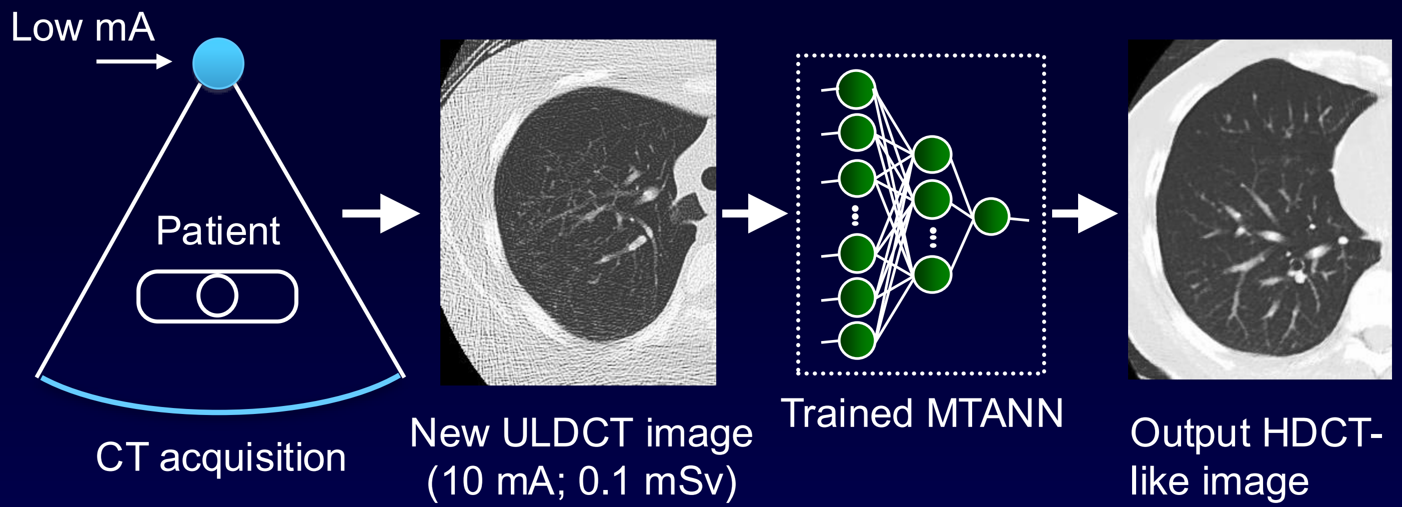
MTANN Deep Learning Reconstruction:
Training Phase



ULD: Ultra low dose
HD: Higher dose

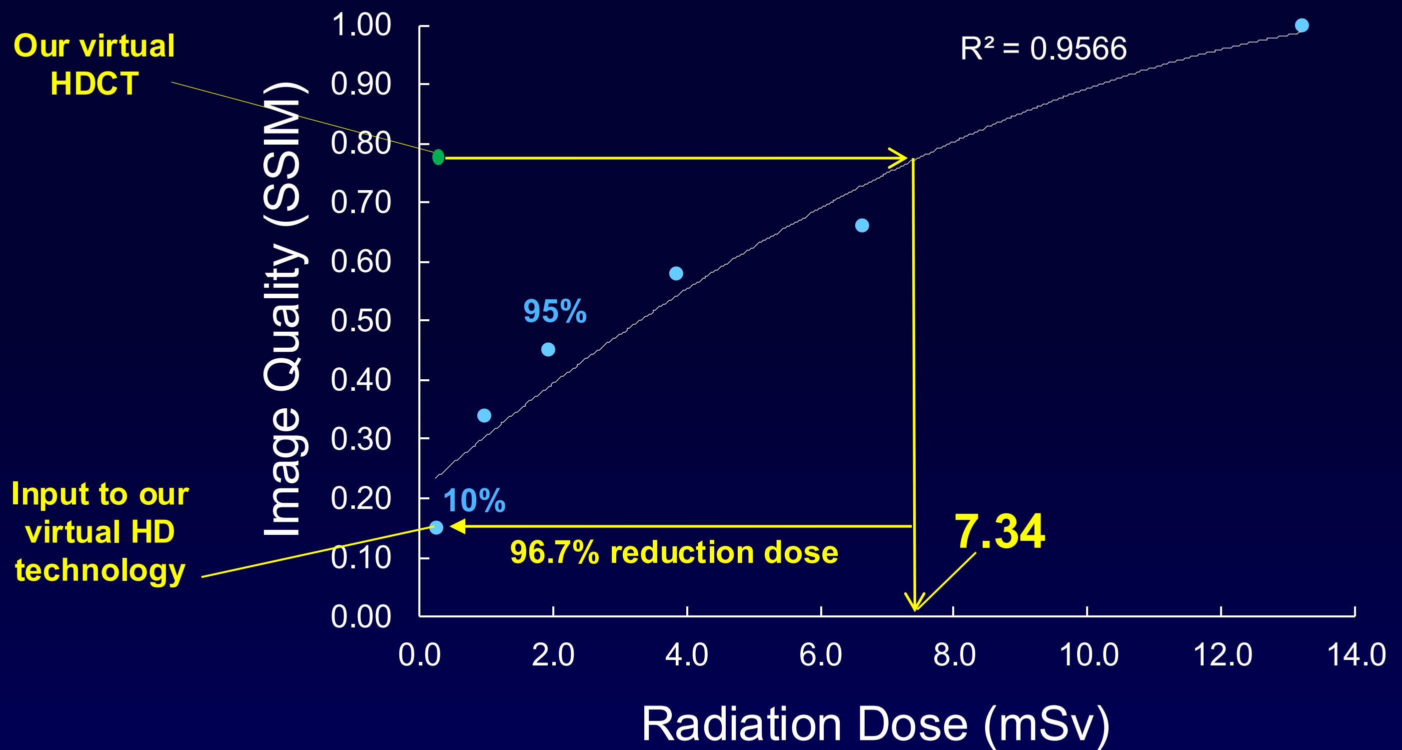
# of Training images: 1 pair

MTANN Deep Learning Reconstruction: Testing Phase



ULD: Ultra low dose
HD: Higher dose

Estimation for Equivalent Radiation Dose from Image Quality (SSIM)



SSIM = structural similarity index (Wang Z et al. *IEEE TIP*, 2004)

1) K Suzuki et al. *RSNA* (2012)
2) K Suzuki, et al. LNCS-MLMI (2017)

Conclusion

- Small-data deep learning requires small data to achieve high performance
  - MTANN small-data deep learning was able to achieve the state-of-the-art performance with only a dozen images to train
  - Small-data MTANN was super efficient for benign and malignant classification (47ms per case on MacBook Air for inference) in both speed and memory usage (1/2100-1/18000).
- Small-data deep learning would be useful for applications in the small-data domain

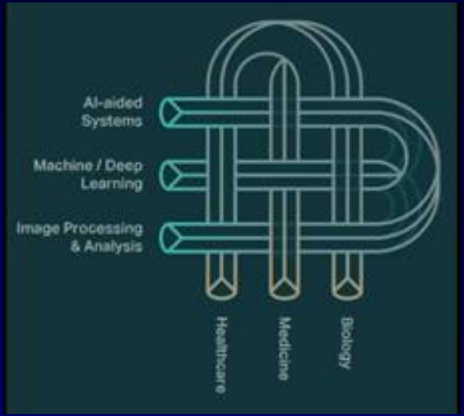
Acknowledgements

Institute of SCIENCE TOKYO

- I am grateful to all the members of BMAI and our collaborators

Thank You!

- This work was supported [in part] by JST GTIE GAP Fund Program, Grant Number GTIE2024_EX10, Japan.



BMAI Biomedical Artificial Intelligence Research Unit