Addressing Practical Challenges in Artificial Intelligence from the Medical Domain to General Tasks

理化学研究所
RIKEN Since 1917

医用機械知能チーム
谷　林　Lin Gu
研究員

Liberal Education Therapist

Finally publish my first paper on clinical psychology in 2021.

-2010

2010-13 **PhD in computer vision**
the Australian National University

2014-16 **Postdoctoral Fellowship**
A*STAR, Singapore

2016-19 **Project Researcher**, NII
Visiting Scholar, Kyoto University

2020- **Research Scientist**, RIKEN AIP
Special Researcher, The University of Tokyo

ANU CRAWFORD SCHOOL Crawford school of public policy. Students and friends are now working in various international organisations and national governments such as UN, WTO, etc.

From the Millennium Development Goals (MDGs) to Sustainable Development Goals (SDGs), regardless of technique details,

What AI could really contribute to human well-being?

-2010	
2010-13	**PhD in computer vision** the Australian National University
2014-16	**Postdoctoral Fellowship** A*STAR, Singapore
2016-19	**Project Researcher**, NII **Visiting Scholar**, Kyoto University
2020-	**Research Scientist**, RIKEN AIP **Special Researcher**, The University of Tokyo

自己紹介

Applying machine learning for neuroscience as well as medical tasks.

Has a wonderful life here!!

-2010	
2010-13	**PhD in computer vision** the Australian National University
2014-16	**Postdoctoral Fellowship** A*STAR BII, Singapore
2016-19	**Project Researcher**, NII **Visiting Scholar**, Kyoto University
2020-	**Research Scientist**, RIKEN AIP **Special Researcher**, The University of Tokyo

自己紹介

Participate ImPACT project and work with hardware (Canon, Hitachi, …) and doctors to design algorithms for early diagnose of breast cancer on photoacoustic imaging device.

Propose new design of camera.

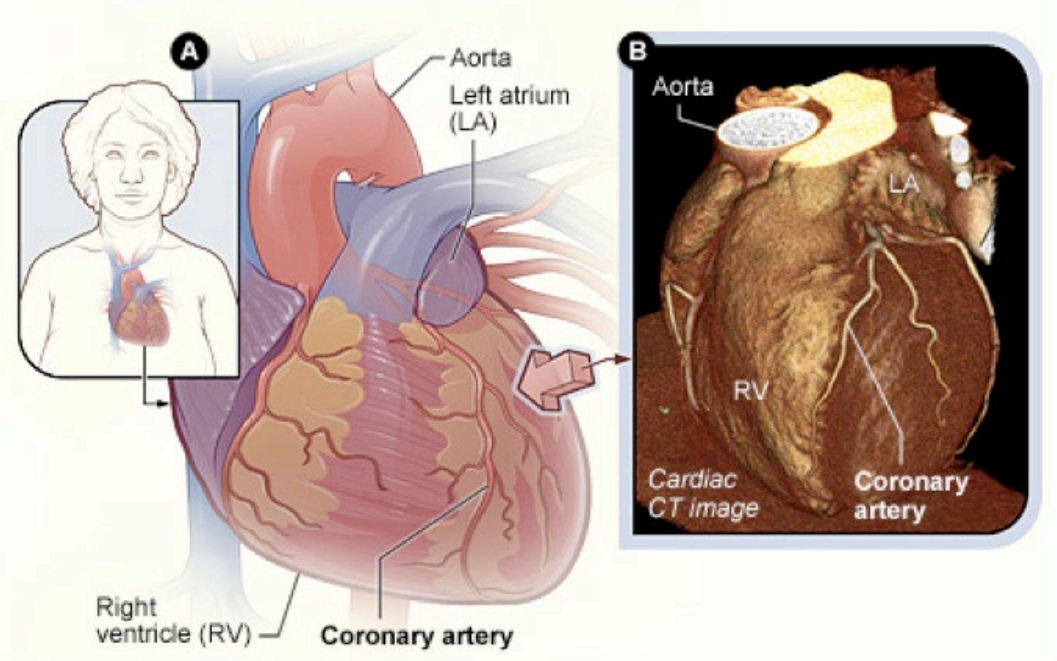
-2010	
2010-13	**PhD in computer vision** the Australian National University
2014-16	**Postdoctoral Fellowship** A*STAR, Singapore
2016-19	**Project Researcher**, NII **Visiting Scholar**, Kyoto University
2020-	**Research Scientist**, RIKEN AIP **Special Researcher**, The University of Tokyo

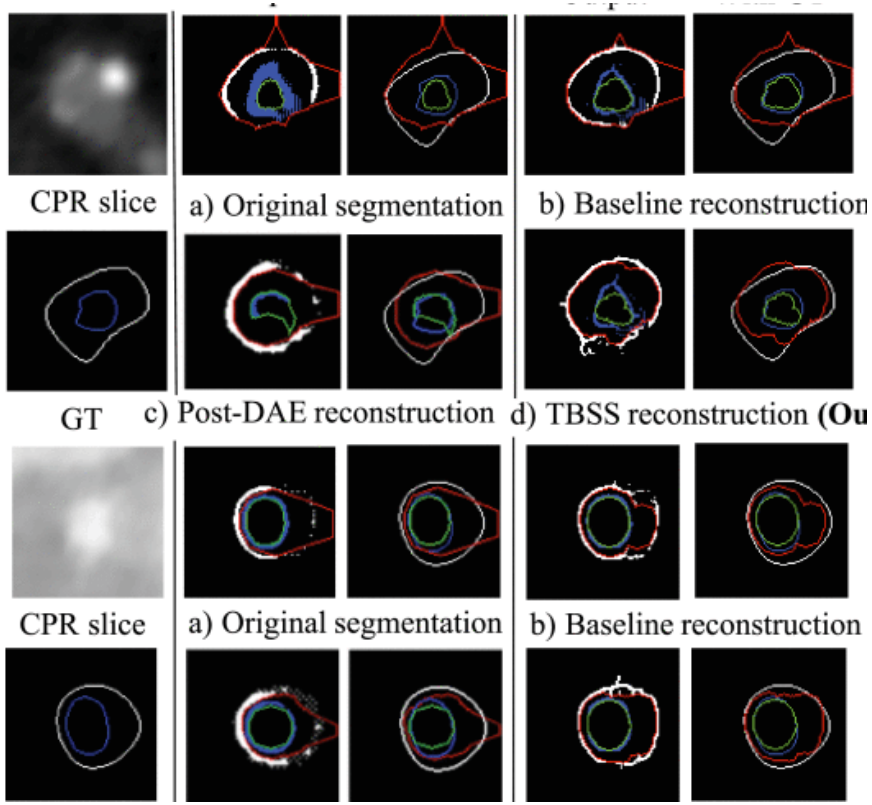
Joined RIKEN AIP under Prof. Harada Tatsuya's supervision.

Machine Intelligence for Medical Engineering Team.

-2010	
2010-13	**PhD in computer vision** the Australian National University
2014-16	**Postdoctoral Fellowship** A*STAR, Singapore
2016-19	**Project Researcher**, NII **Visiting Scholar**, Kyoto University
2020-	**Research Scientist**, RIKEN AIP **Special Researcher**, The University of Tokyo

Research ① Medical Research

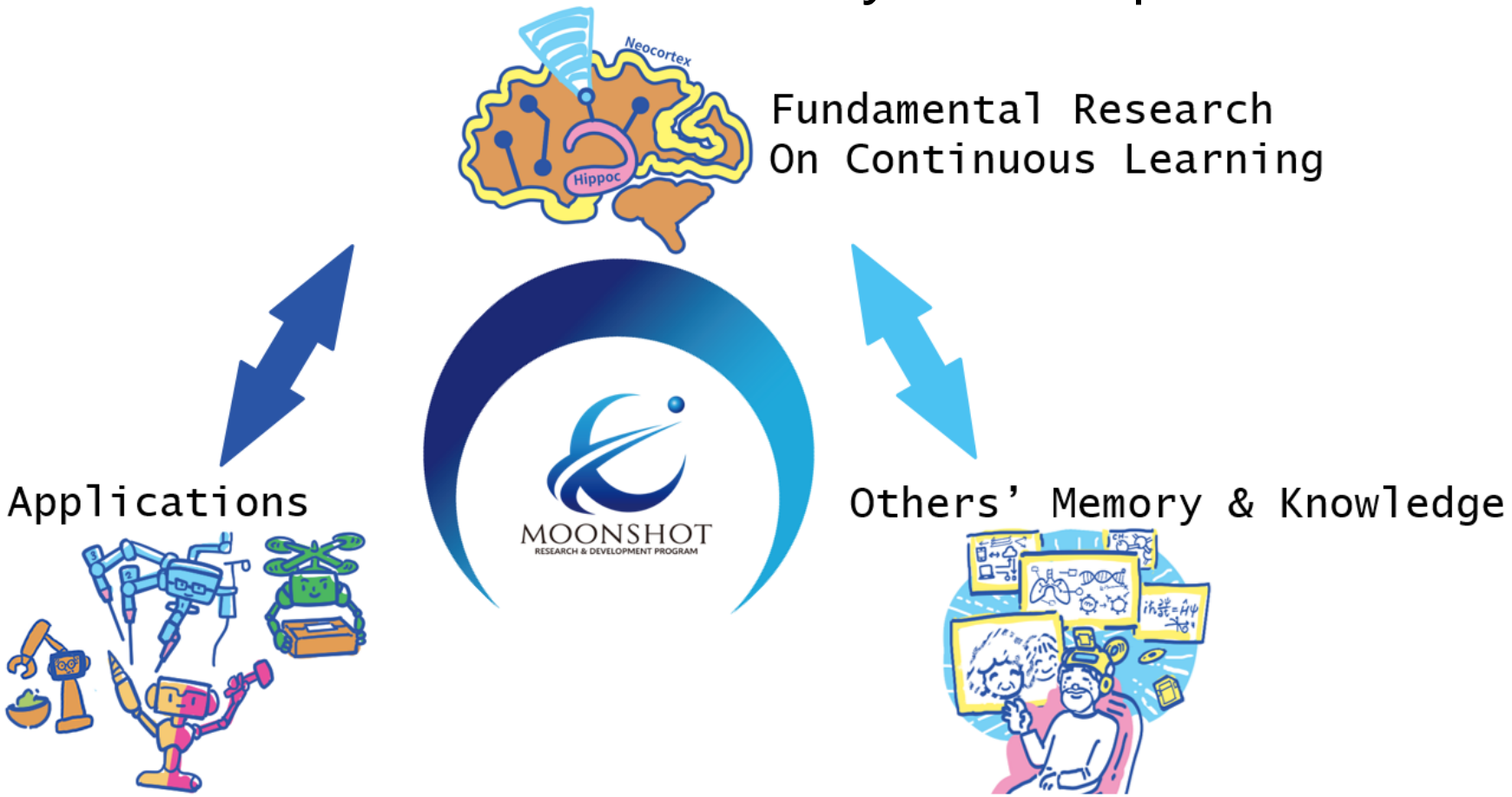




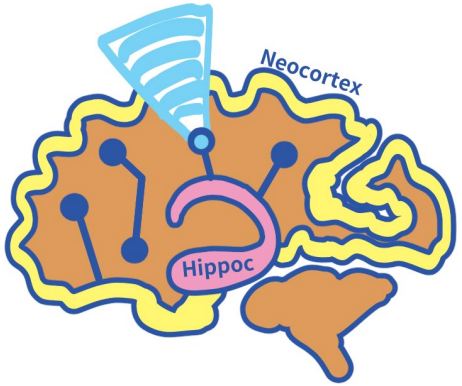
Clinical textbook **AI in Clinical Medicine** will be published in June.

Research ② Moonshot Project

- Moonshot project led by Cabinet Office of Japan.
- Human beings can be free from limitations of body, brain, space, and time by 2050.



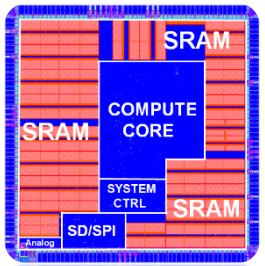
Design AI techniques for deploying robots for practical usage in Osaka Expo



Brain of Robot

Light weight AI architecture:
Towards an Effective Orthogonal Dictionary Convolution Strategy
AAAI 2022

High efficient AI chip inspired by neuroscience:
EtinyNet: Extremely Tiny Network for TinyML
AAAI 2022



Eye of Robot

Recognise under different scales:
Exploring Resolution and Degradation Clues as Self-supervised Signal for Low Quality Object Detection
ECCV 2022

See in the dark:
You Only Need 90K Parameters to Adapt Light
BMVC 2022

See under noise:
Memory-Efficient Deformable Convolution based Joint Denoising and Demosaicing for UHD Images
IEEE TNNLS

DnRCNN: Deep Recurrent Convolutional Neural Network for HSI Destriping
IEEE TNNLS

Heart of Robot

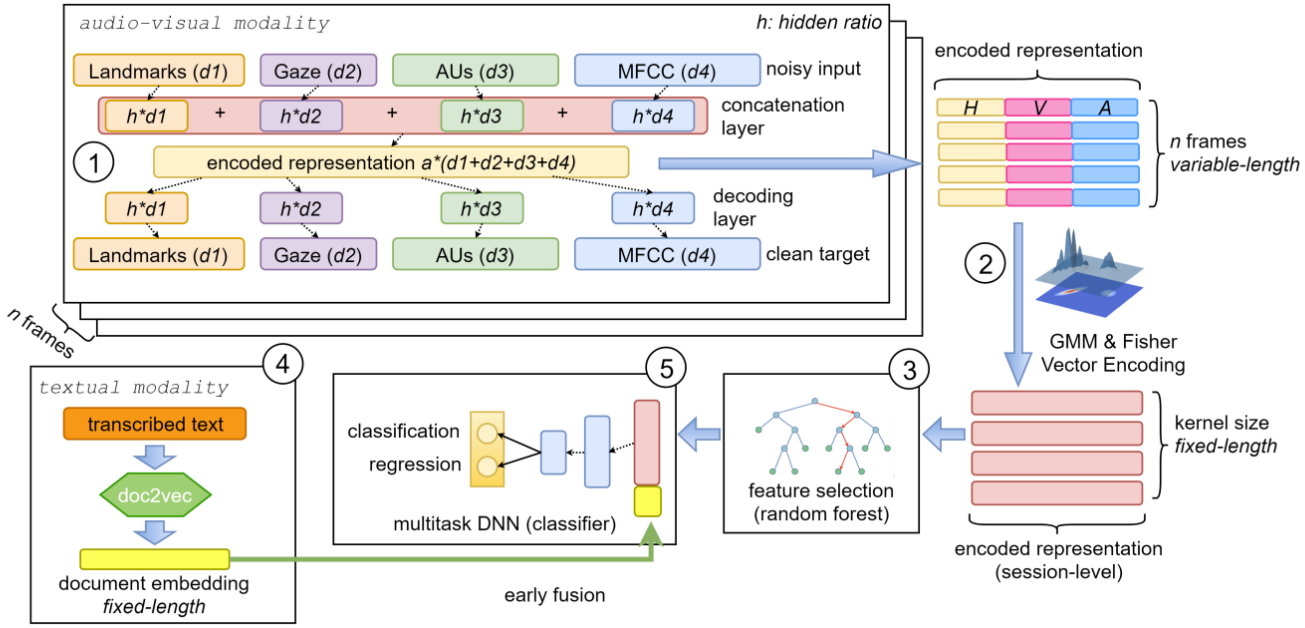
Protect Privacy and Enhance Fairness

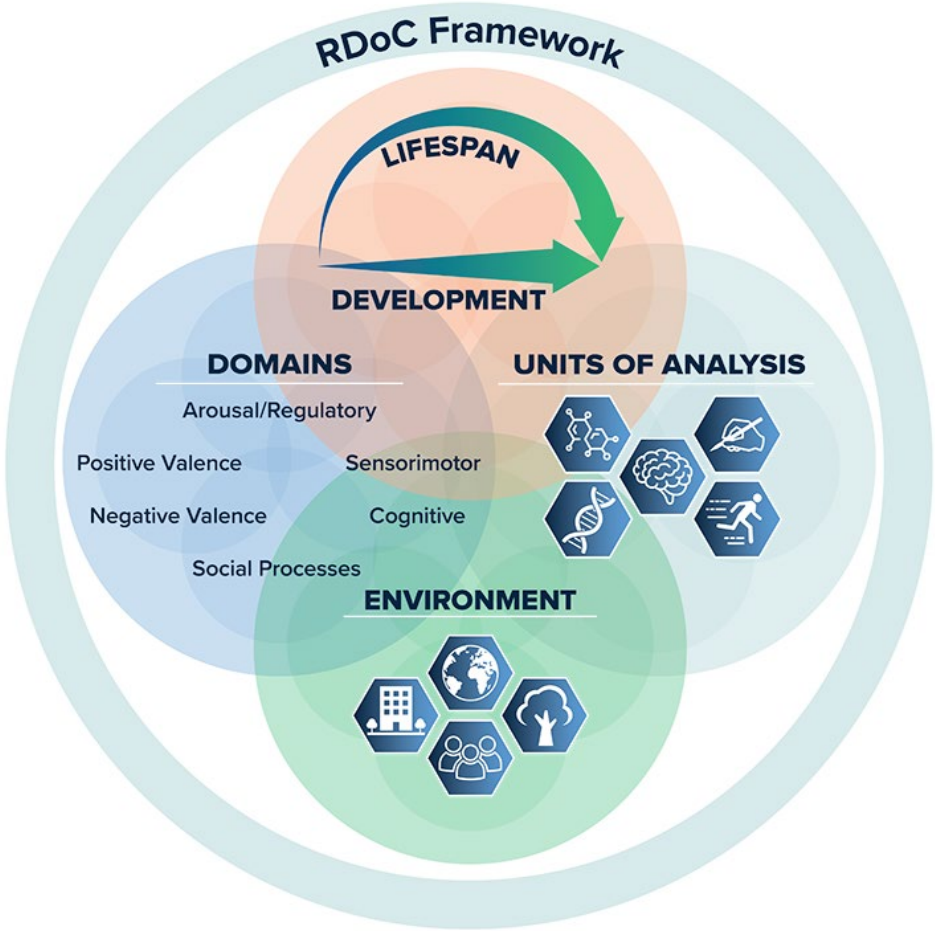
People taking photos that faces never share: Privacy Protection and Fairness Enhancement from Camera to User
AAAI 2023

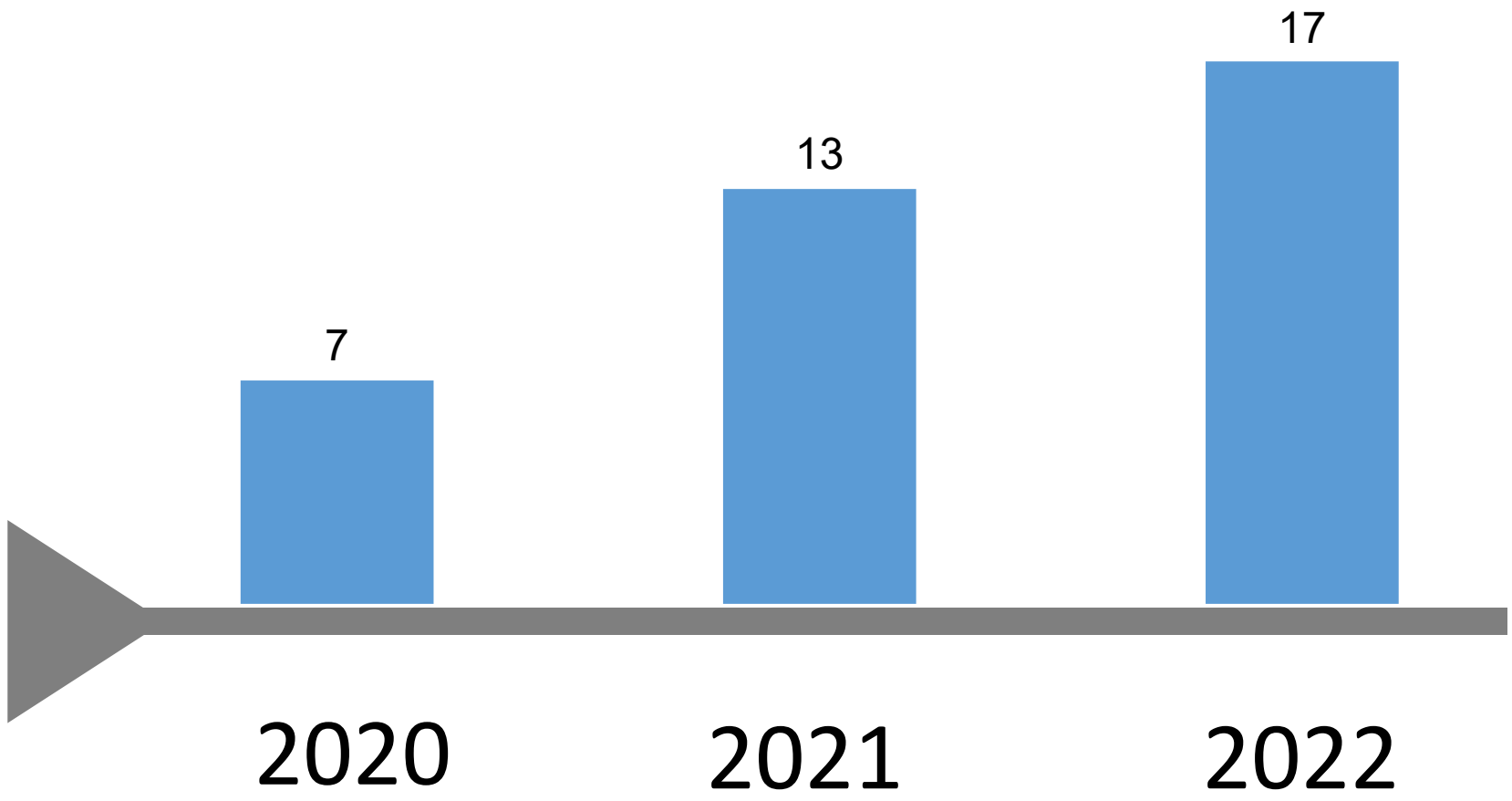
Improving Fairness in Image Classification via Sketching
Neurips 2022 Workshop

Research ③ Subtyping Schizophrenia with AI

- The joint grant between RIKEN and Zhejiang University Hospital
1 million USD







2022-2023

AAAI X 3

ECCV X 2

MICCAI X 2

CVPR

ICLR

Nature Methods (IF: 48)

Pattern Recognition (IF:7.1)

IEEE TNNLS (IF: 10.5)

Research Theme

- Word **Artificial** is introduced in English from *artificialis* in 15th century
- “made by man, contrived by human skill”
- There are much efforts on how to **make** Artificial Intelligence

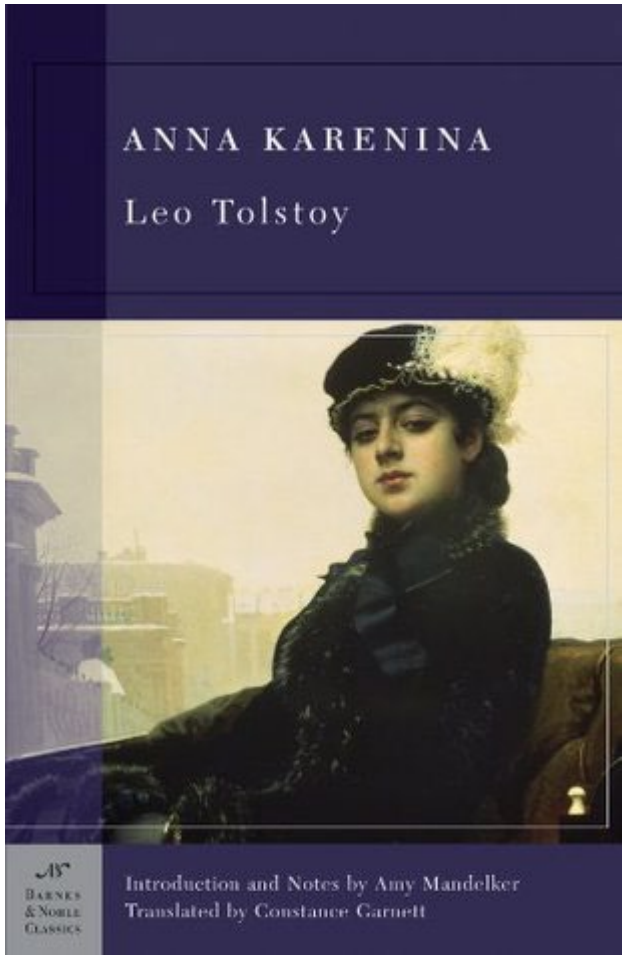
Research Theme

- Word **Artificial** is introduced in English from *artificialis* in 15th century
- “made by man, contrived by human skill”
- There are much efforts on how to **make** Artificial Intelligence

- How to make AI **for human**, in the real-world

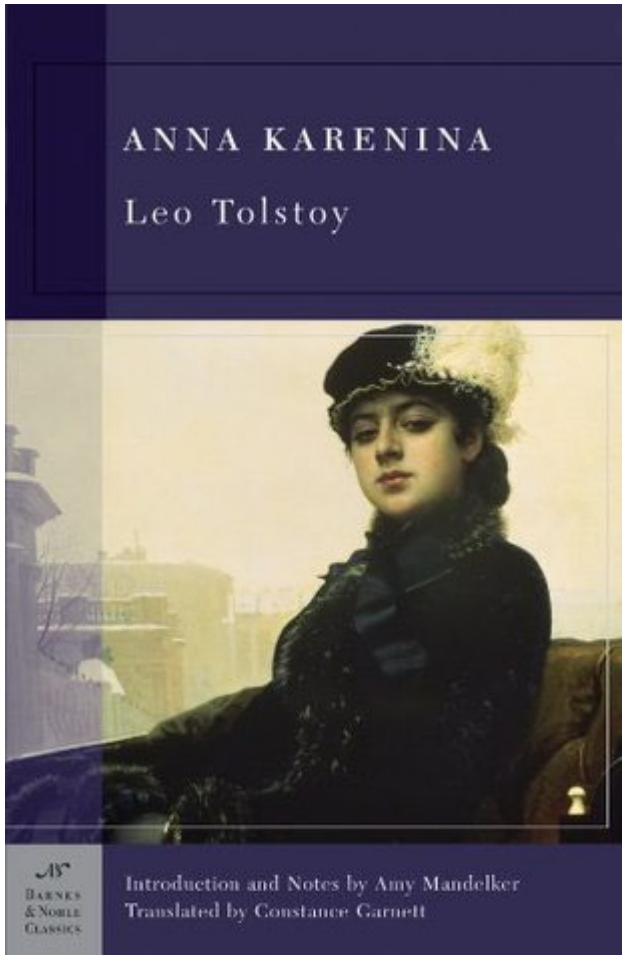
Bottlenecks

- All happy families are alike;
  - Existing computer vision and machine learning algorithms perform well on the ideal conditions and parameters.



Bottlenecks

- each unhappy family is unhappy in its own way.
  - Their accuracy and effectiveness can be limited or compromised in real-world scenarios that deviate from the ideal.

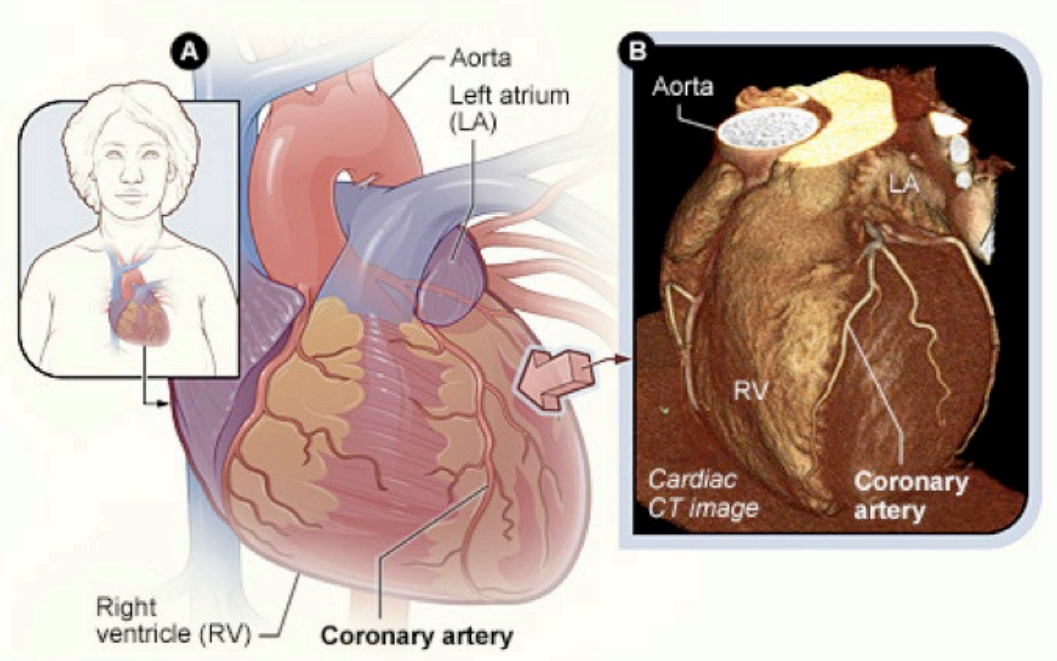


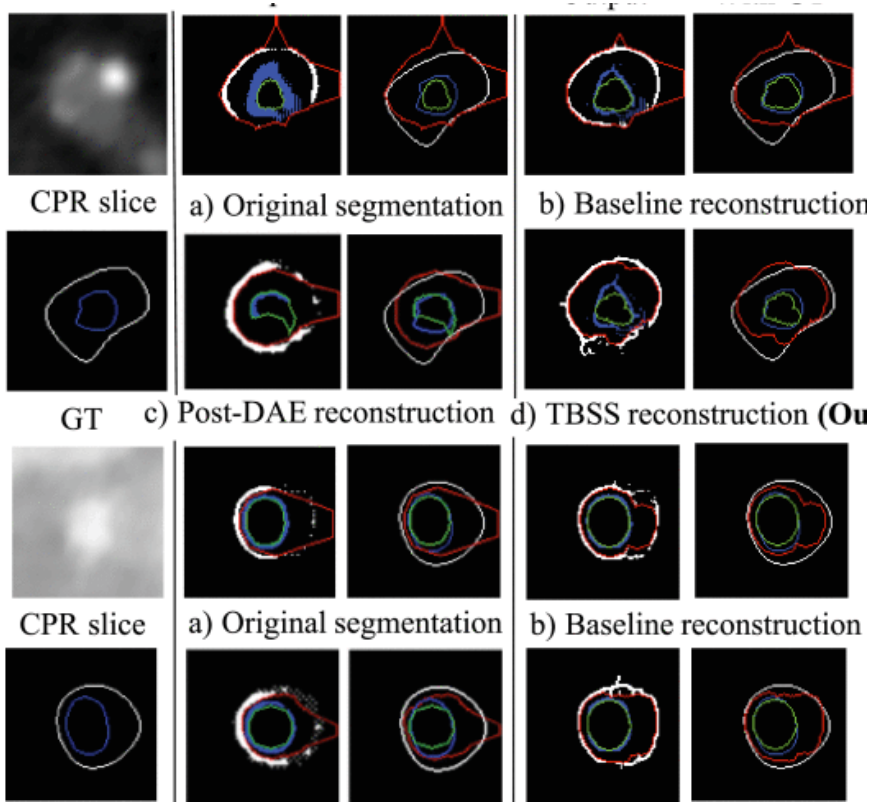
Research Theme

- The key for making AI for human is identifying the gap between ideal and real.

- Just like the title of our book to be released in June
- AI in Clinical Medicine: A **Practical Guide** for Healthcare Professionals

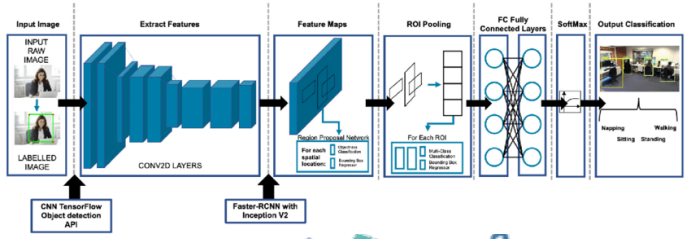
Research ① Medical Research

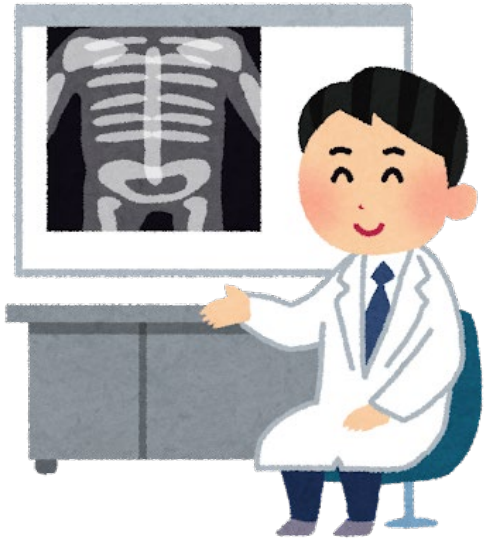




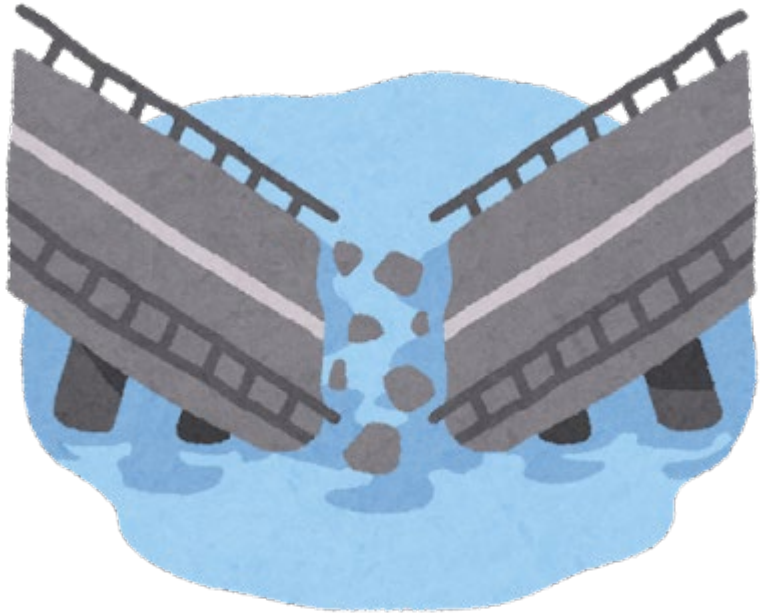
Personal Research Experience

Gap between engineering professionals and medical experts.





Tag(medical) ≠ Grountruth (engineering)



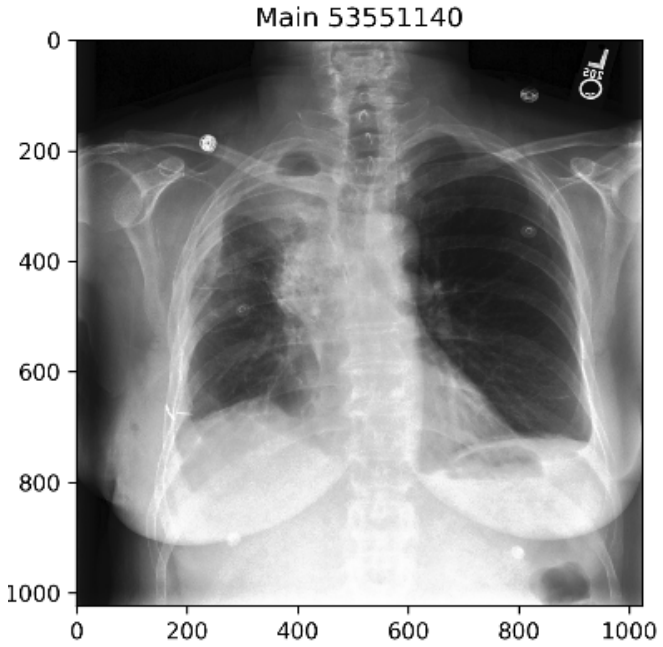
Medical:
Imperfect, uncertainty,
Ambiguous, real world

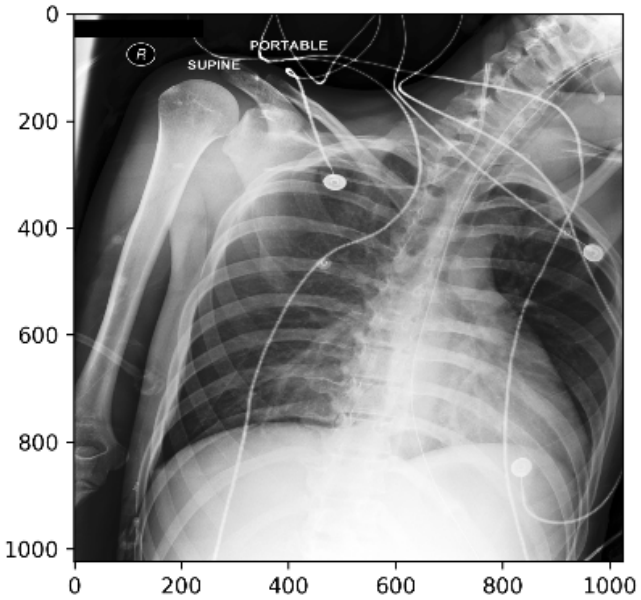
Engineering:
Well defined,
simple and elegant
abstract is beautiful

Medical Experts

To cure sometimes
to relieve often
to comfort always.

Medical world is full of uncertainty and imperfect.





Right upper opacity is **most likely** fluid in the pleural space. It has increased from prior study. There is **probably** atelectasis in the right middle lobe.
Note is also made of a **questionable** small right basilar pneumothorax

We are not gods.

We are only human.
(ἄνδρα θνητὸν)

It is impossible
(sometimes) to know
the ground-truth.

Medical Pipeline

Observation
Diagnose

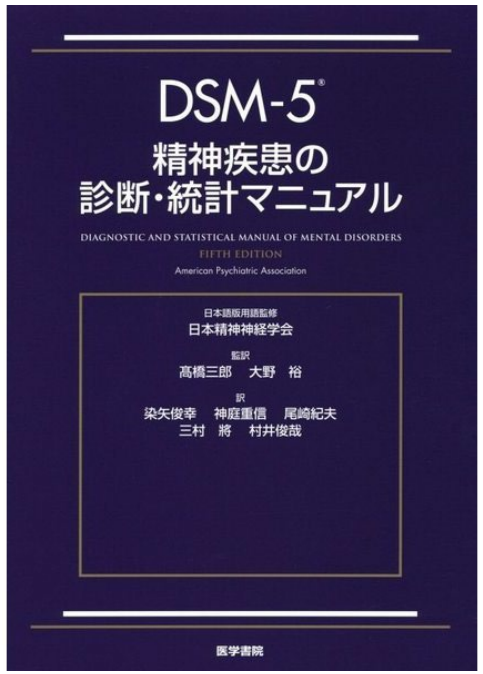
Take Action

Prey for the best

Observation and Diagnose

Observe the patient (data) and tag a label for the disease.

The reliability depends of the disease and field.



The symptoms used to define them often do not cluster neatly in the way that those of true syndromes would, and the statistical evidence for their existence is sometimes sparse. Nor, in most cases, **have the hoped-for biological markers turned up**—and to the extent that they have, **they have muddied the waters, rather than clarifying them.**

Take Action

Observe the patient (data) and tag a label for the disease.

What shall we do for a > 10% to ≤ 50% likelihood of malignancy?

A **BI-RADS 4** lesion under the breast imaging-reporting and data system refers to a suspicious abnormality.

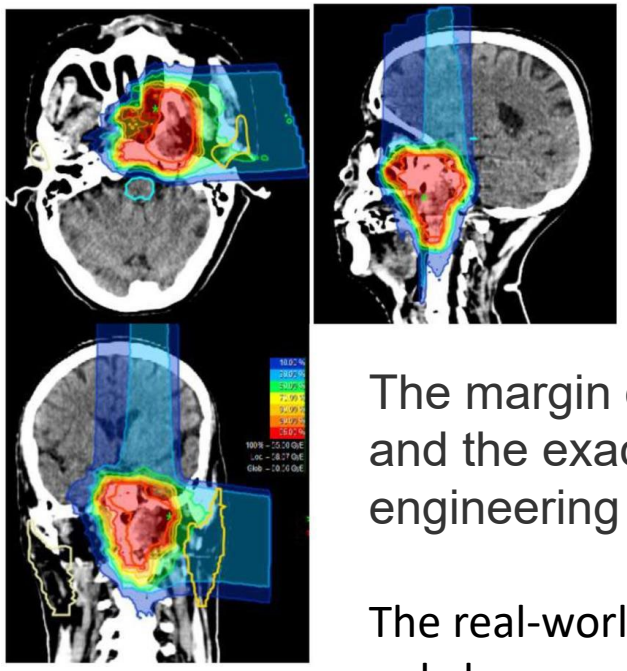
BI-RADS 4B: moderate level of suspicion for malignancy and has a > 10% to ≤ 50% likelihood of malignancy

Such probability is rooted in past population studies and is not individualized, still rendering uncertainty in individual clinical decision-making.

Pray for Best (Never Tell You)

After taking the procedure, doctors could only wait and hopes for the best.

The real-world result (survival rate e.g.) will only be seen after years.



The margin of CTVp1 and CTVp2 is quite arbitrary and the exact consequences is surprisingly (for engineering professionals) unclear.

The real-world patient outcome can be probabilistic and only be seen after years.

Gap Between Medical and Machine Learning

AI professionals will keeps on asking for grountruth label.

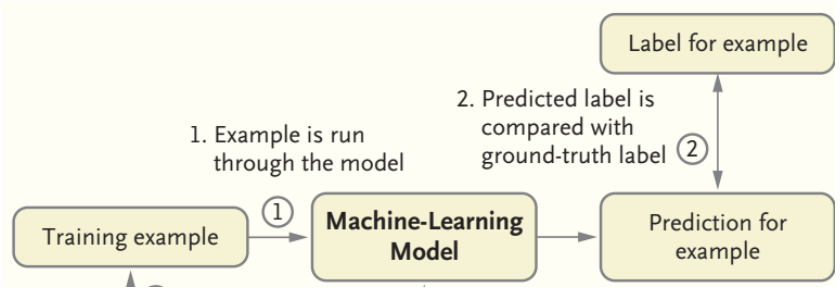
Medical to Machine Learning

Observation
Diagnose

Take Action

Prey for the best

Often unreliable disease label



Survival rate after 10 years

Gap Between Medical and Machine Learning

Sometimes diseases label are meaningless

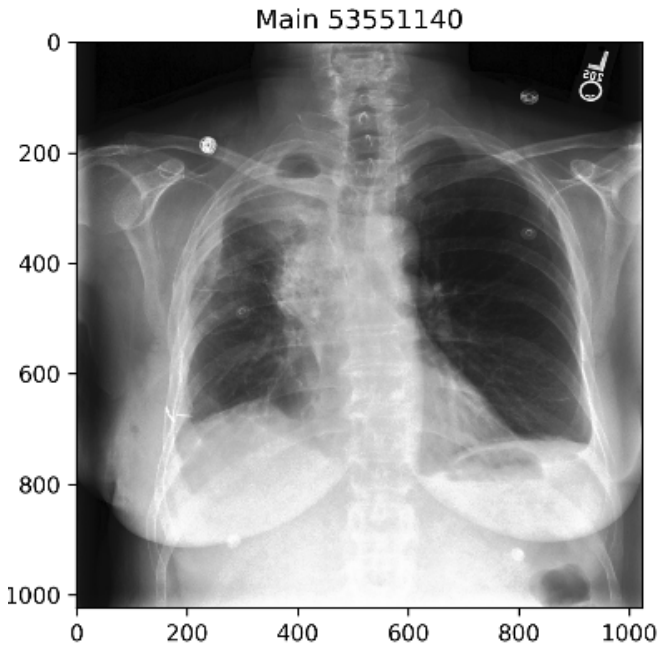
Sometimes diseases are not certain.

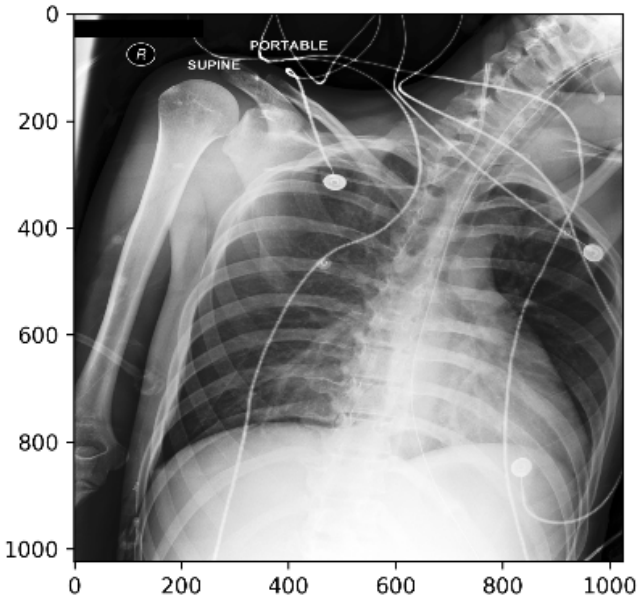
…..

Anyway, doctor will give a label and call it groundtruth.

Call for Collaboration

How to define this uncertainty in machine learning?

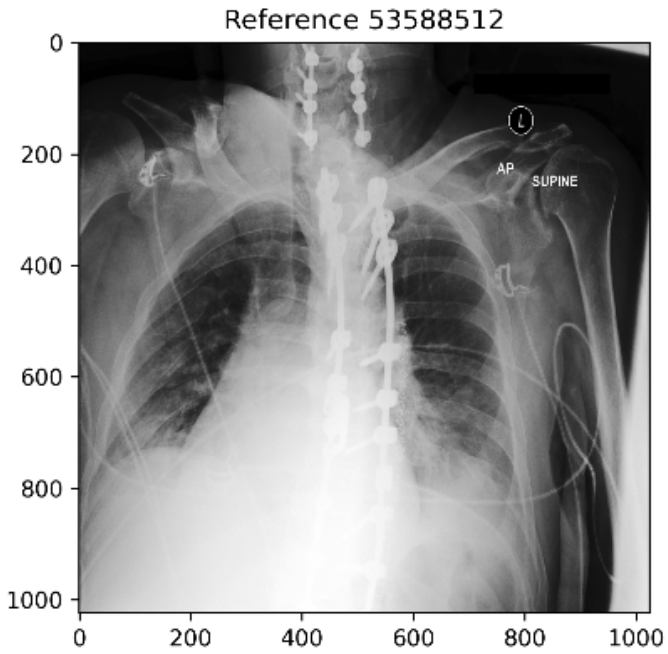




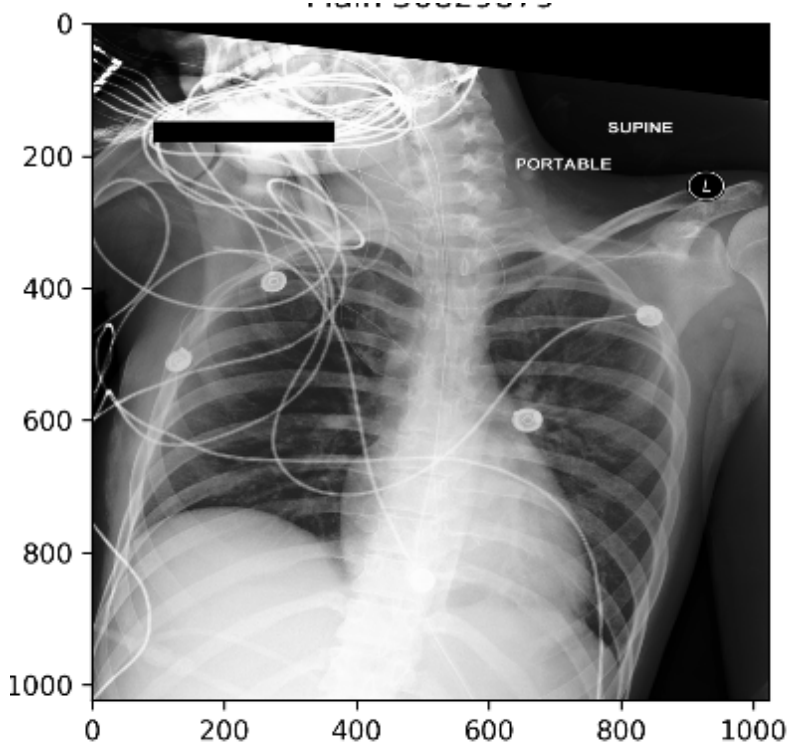
Right upper opacity is **most likely** fluid in the pleural space. It has increased from prior study. There is **probably** atelectasis in the right middle lobe.
Note is also made of a **questionable** small right basilar pneumothorax

Call for Collaboration

How to reflect these descriptions in causal relation?



Low lung volumes and patient rotation limits evaluation of *cardiomediastinal contours*. Bibasilar opacities likely reflect atelectasis at, and remain more severe on the left than the right.



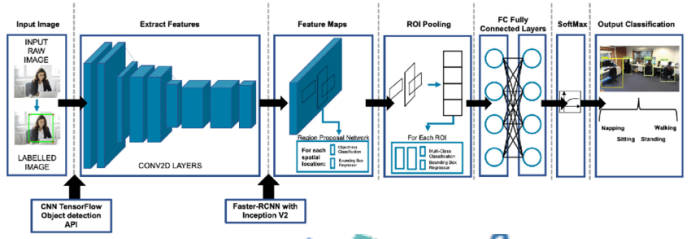
Again the left hemidiaphragm is not sharply seen, suggesting small pleural effusion and volume loss in the left lower lobe.

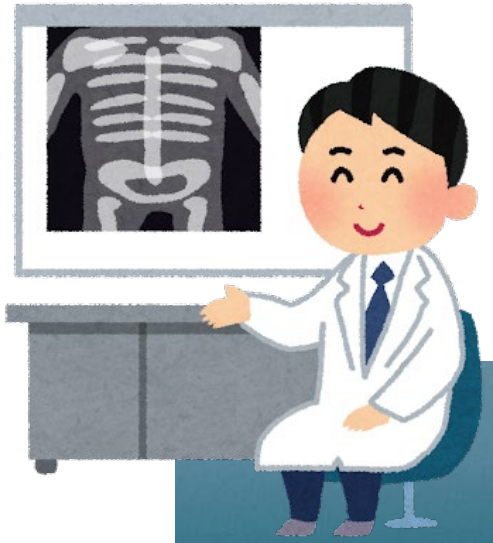
Calling for Collaboration

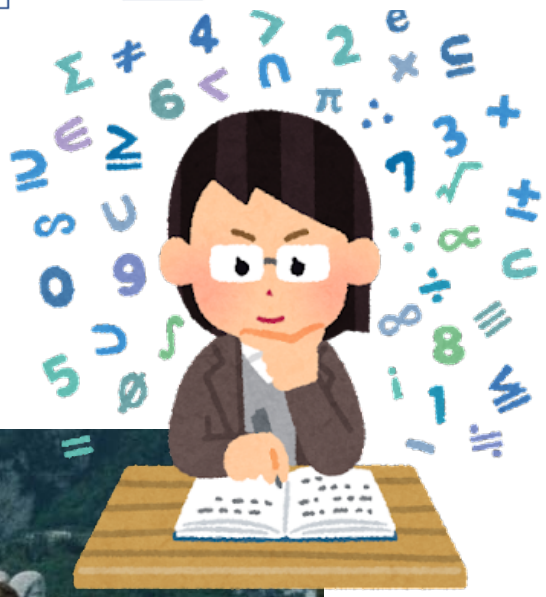
＼連携研究は大歓迎／

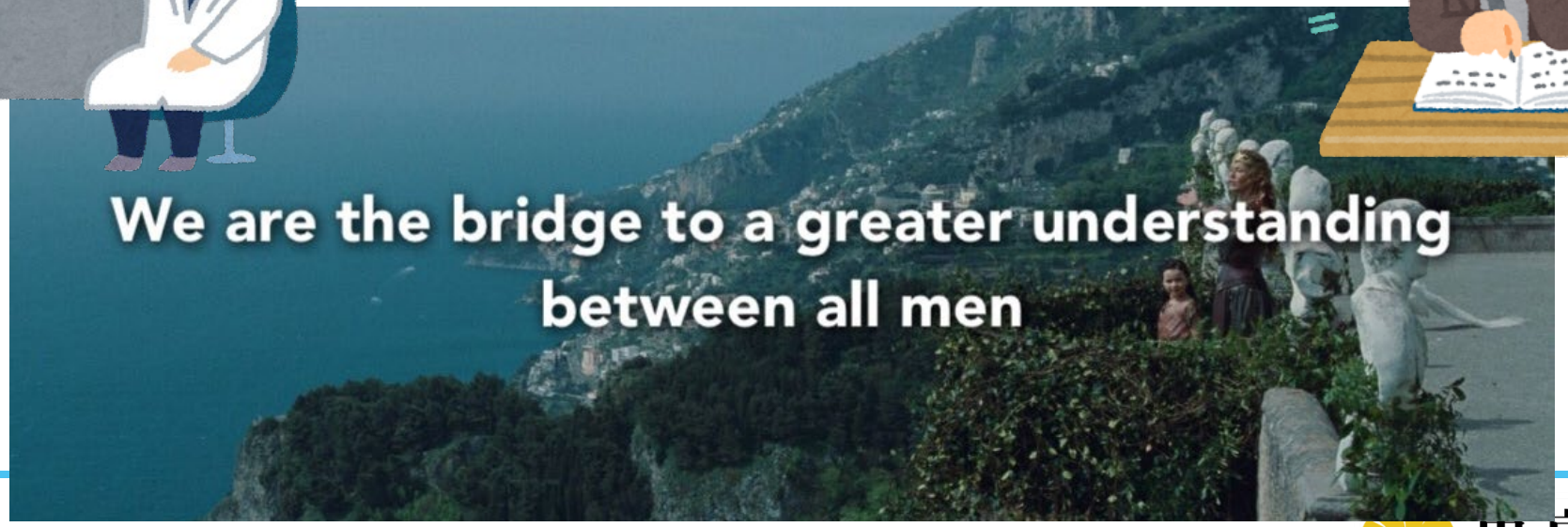
Our Mission

Deepen mutual understanding between two fields. Participate the real diagnose loop.









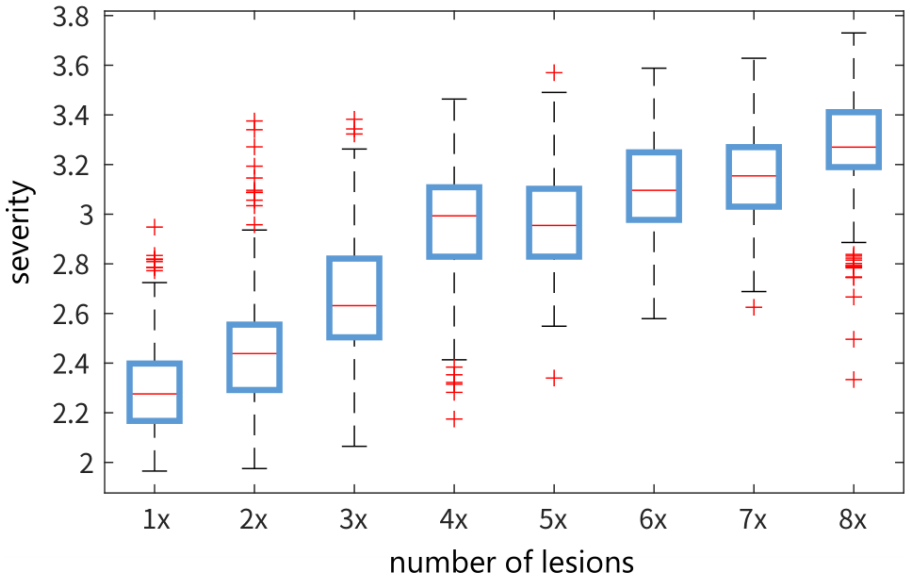
Our Solutions

- Explain how DL works from clinical aspect.
  - Pathological Evidence Exploration in Deep Retinal Image Diagnosis Niu et al. AAAI 2019
  - Explainable Diabetic Retinopathy Detection and Retinal Image Generation Nie et al. IEEE JBHI 2021

Koch's Postulates

- Double blind test on licensed ophthalmologists

- Quantity manipulation can result different severity diagnose



Gaze as External and Free Resource

Human gaze is a cost-efficient physiological data.

Goal-oriented gaze estimation for zero-shot learning
Liu et al. CVPR 2021

The selective attention mechanism helps the cognition system focus on task-relevant visual clues.

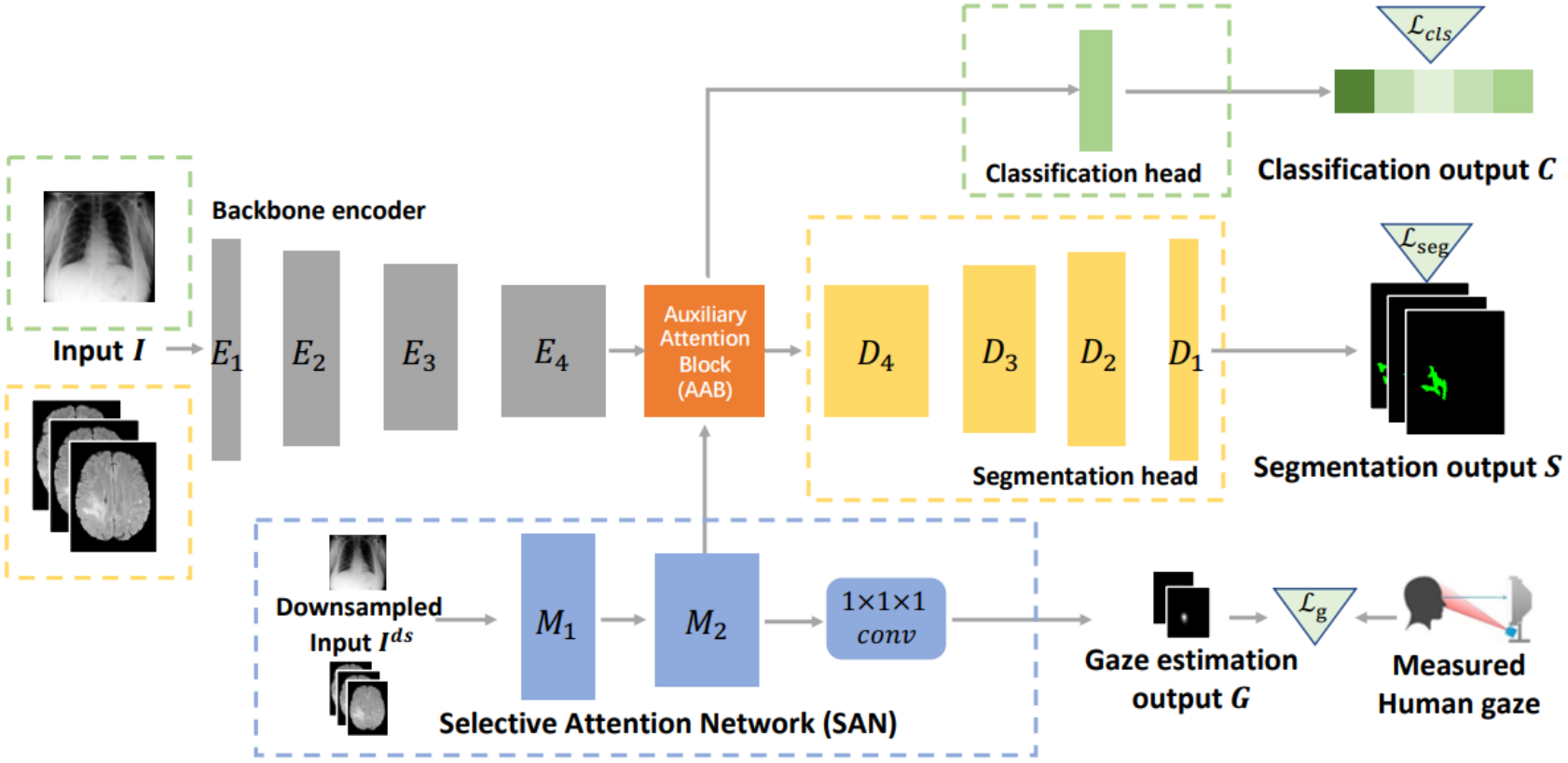
Where to Focus: Investigating Hierarchical Attention Relationship for Fine-Grained Visual Classification
Liu et al. ECCV 2022

Human beings (and thus network) can efficiently learn from a very limited number of training samples.

Leveraging Human Selective Attention for Medical Image Analysis with Limited Training Data
Huang et al. BMVC 2021

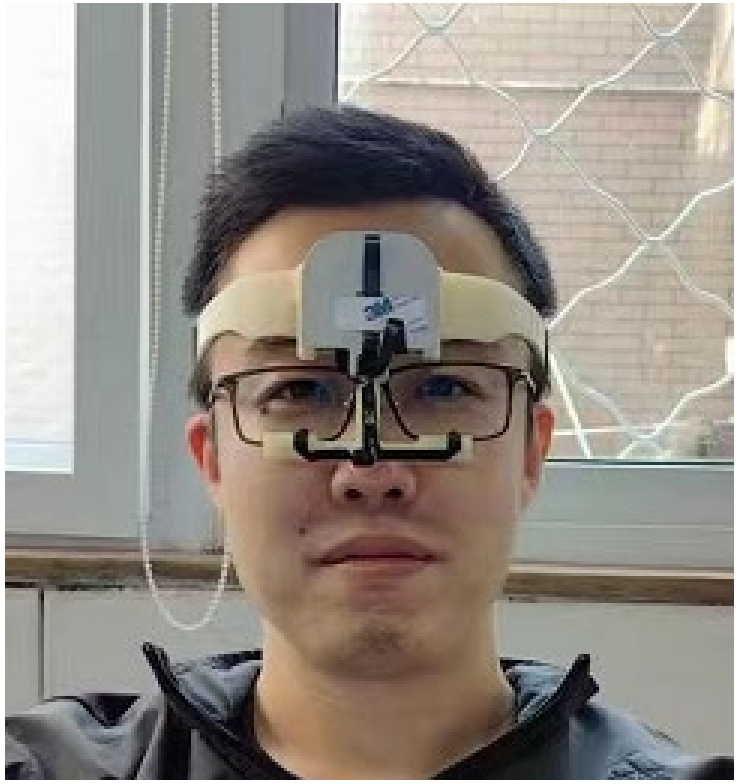
Gaze for 2D and 3D Medical Images

We propose a novel memory mechanism to simulate the cognition of radiologists when reading brain scan images. It could enhance arbitrary CNN based segmentation.



Gaze Tracking Hardware Platform

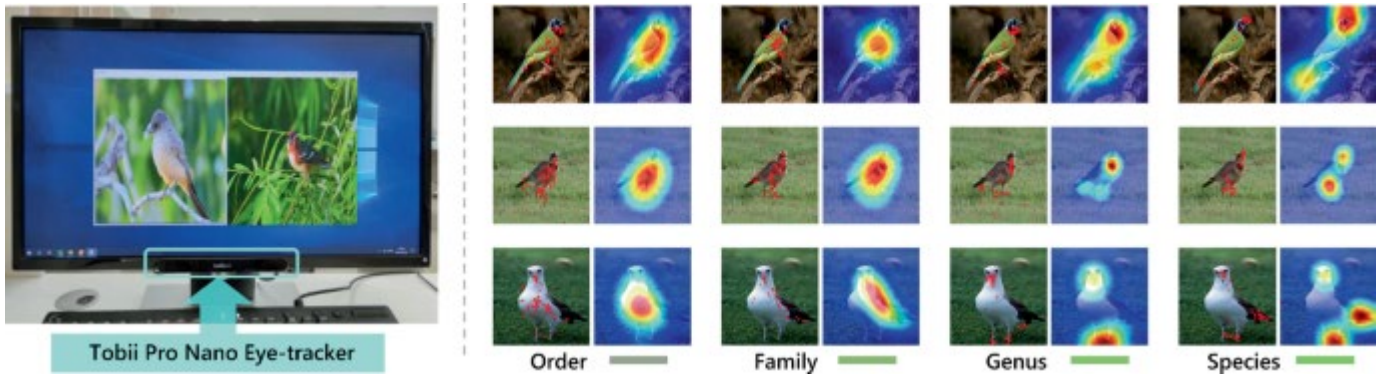
- Now we have developed our own gaze tracking hardware platform.
- More accurate and more energy efficient.
- Pave the way for other physiological signal collection.



Gaze for Hierarchical Classification

- Collect and release a human gaze dataset on the Caltech-UCSD birds (CUB) dataset at different category hierarchies.

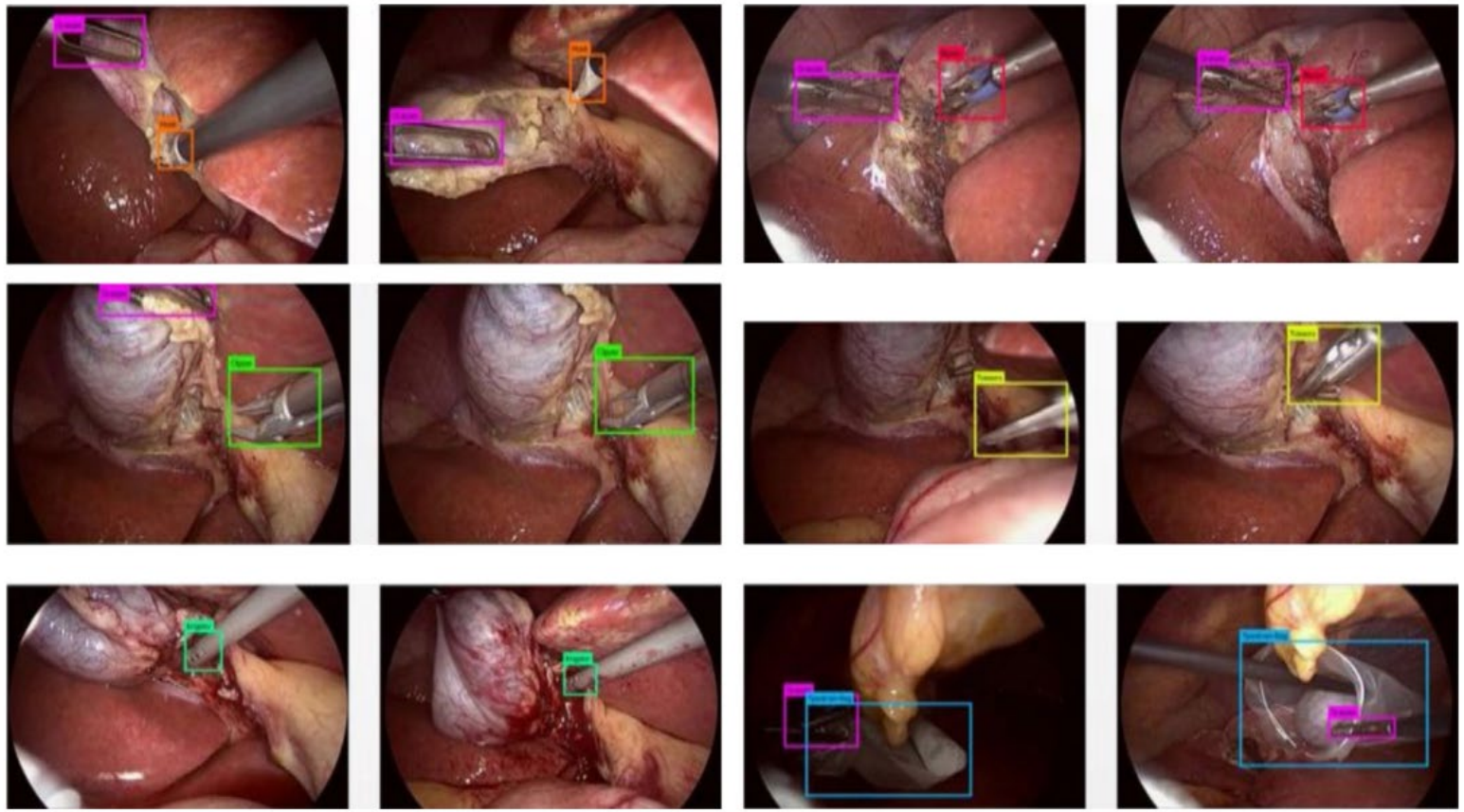
- Propose a cross-hierarchical region feature (CHRF) learning framework to mimic human attention behaviour at different granularities.



Where to Focus: Investigating Hierarchical Attention Relationship for Fine-Grained Visual Classification
Liu et al. ECCV 2022

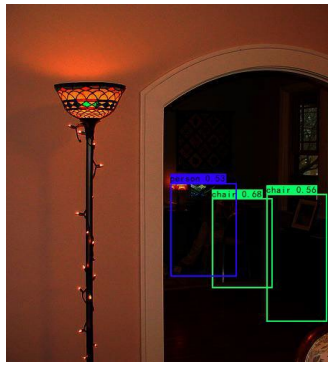
Surgery Assistance

Assist robotic surgery



Surgical Skill Assessment via Video Semantic Aggregation
Li et al.
MICCAI 2022

Gap in Vision



Low light

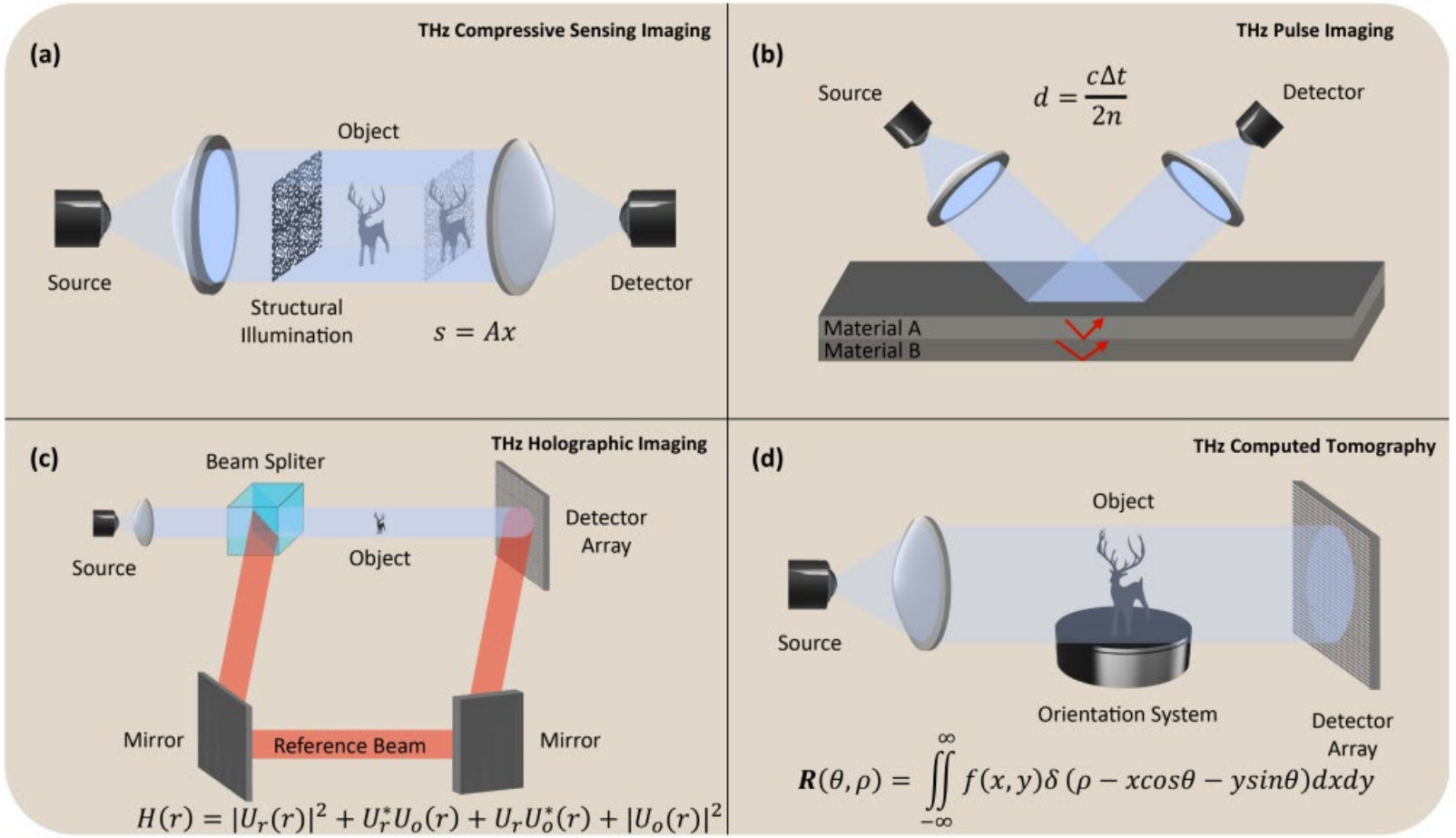
Over exposed

Hazy

Rainy

etc.

Terahertz Imaging



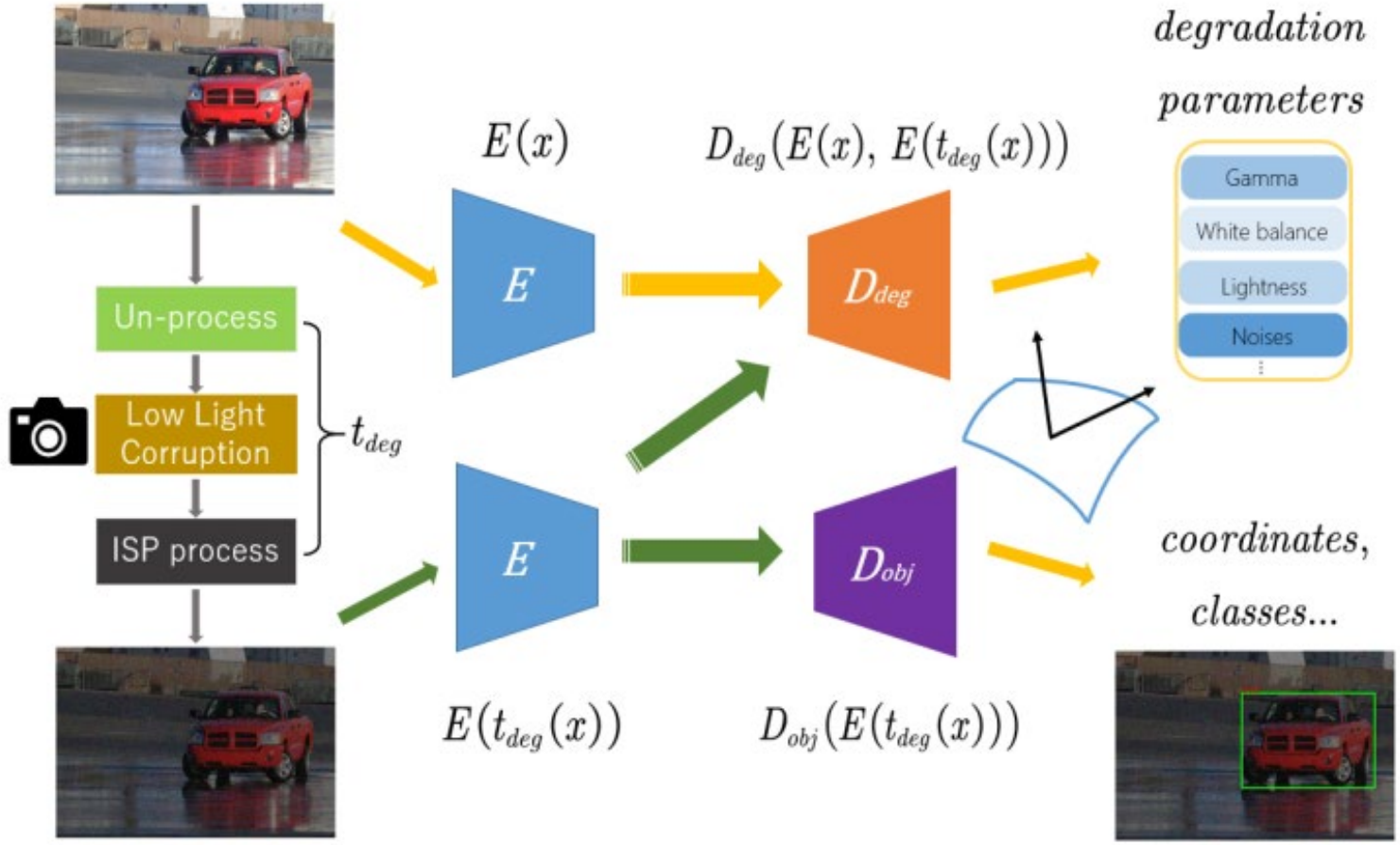
Physics-guided Terahertz Computational Imaging Hung et al.

Object Detection in Dark

Object detection in challenging conditions like dark place, rainy days.

Key for autonomous vehicle.

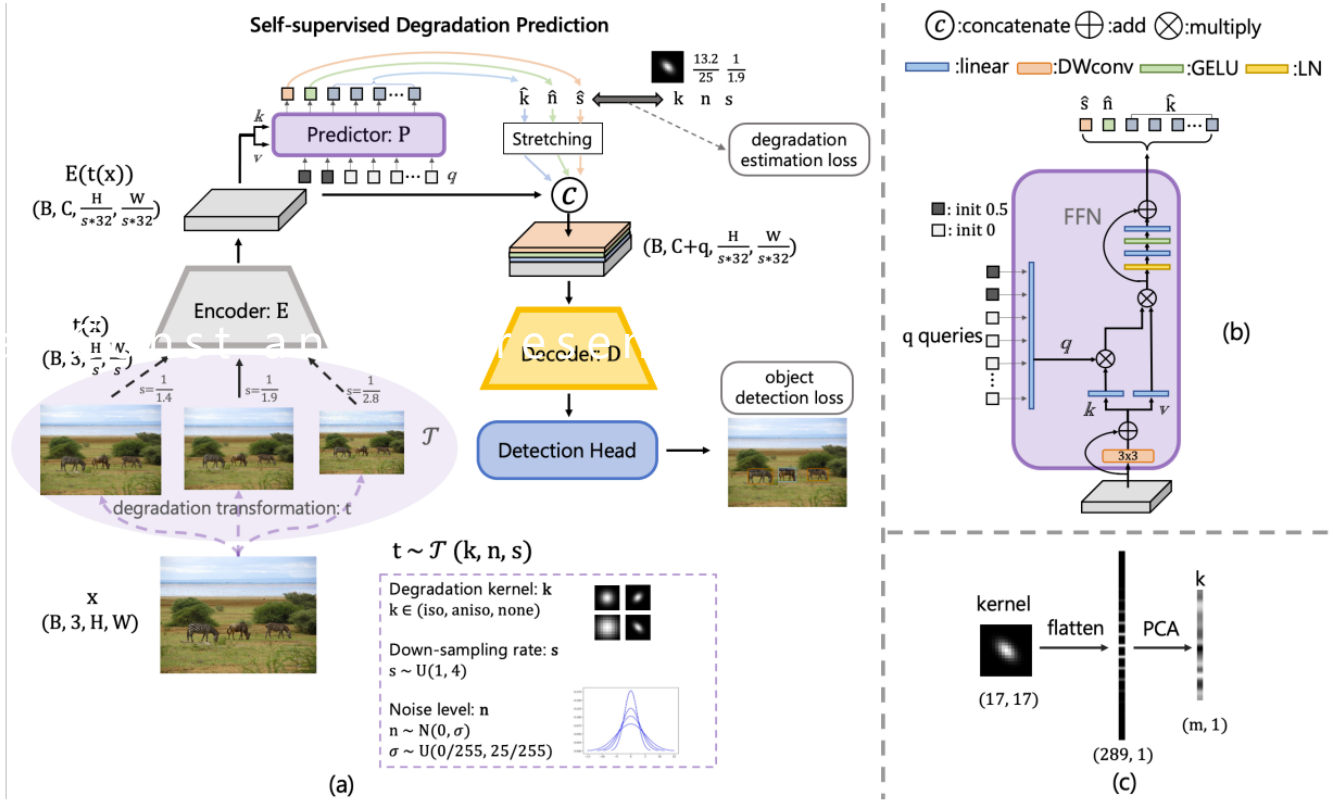
Perceptual Constant Representation



Perceptual Constant Representation

We have extended this model for scale invariant recognition.

Perceptual Constant Representation



Illumination Invariant

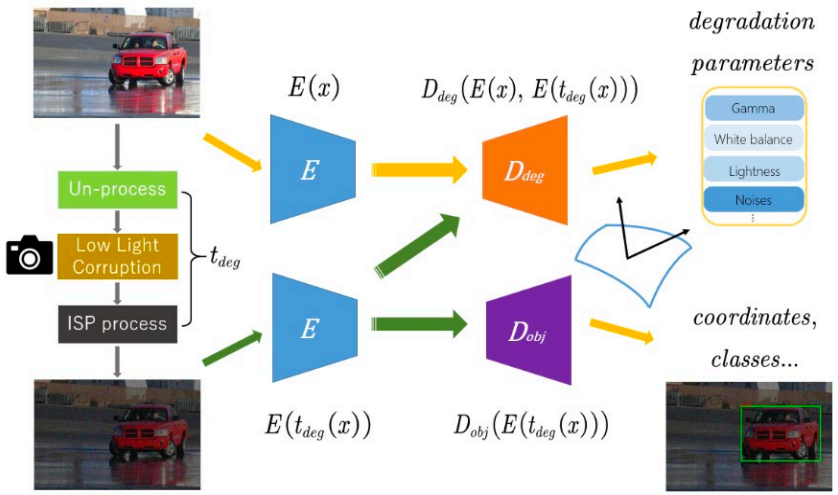
- From AET to MAET: Per aspera ad astra

AET (Auto-Encoding Transformation). Zhang *et.al.* 2019

Auto-Encoding Transformations Rather Than Data

Auto-Encoding Transformations and Detection Results

MAET Ours.



Encoder: $x \to E(x),\ t(x) \to E(t(x))$

Decoder: $D(E(x),\ E(t(x))) \to \hat{t}$

Encoder: $x \to E(x),\ t_{deg}(x) \to E(t_{deg}(x))$

Transformation Decoder : $D_{deg}[E(x),\ E(t_{deg}(x))] = \widehat{t_{deg}}$

Detection Decoder: $D_{obj}[E(t_{deg}(x))] \to detection\ feature$

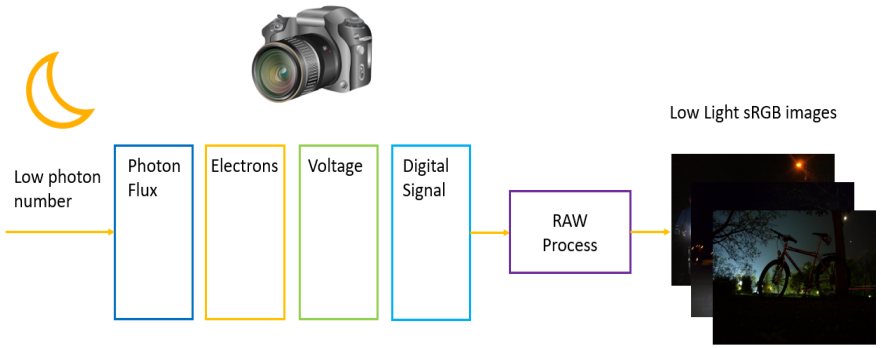
Multitask AET with Orthogonal Tangent Regularity for Dark Object Detection Cui et al. ICCV 2021

Practical Physics Model

- Photo **does not** directly reflect physical world.

In-Camera Low-Lit Degradation:

Light photons project through a lens on a capacitor cluster, each capacitor develops an electric charge corresponding to the lux of illumination of the environment.

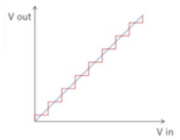


Low-Lit & Shot Noise & Read Noise:

Shot and read noises are common in a camera imaging system, thus, we model the noisy measurement on the sensor:

$$x_{noise} \sim N(\mu = kx, \sigma^2 = \delta_r^2 + \delta_s kx)$$
$$y_{noise} = kx + x_{noise},$$

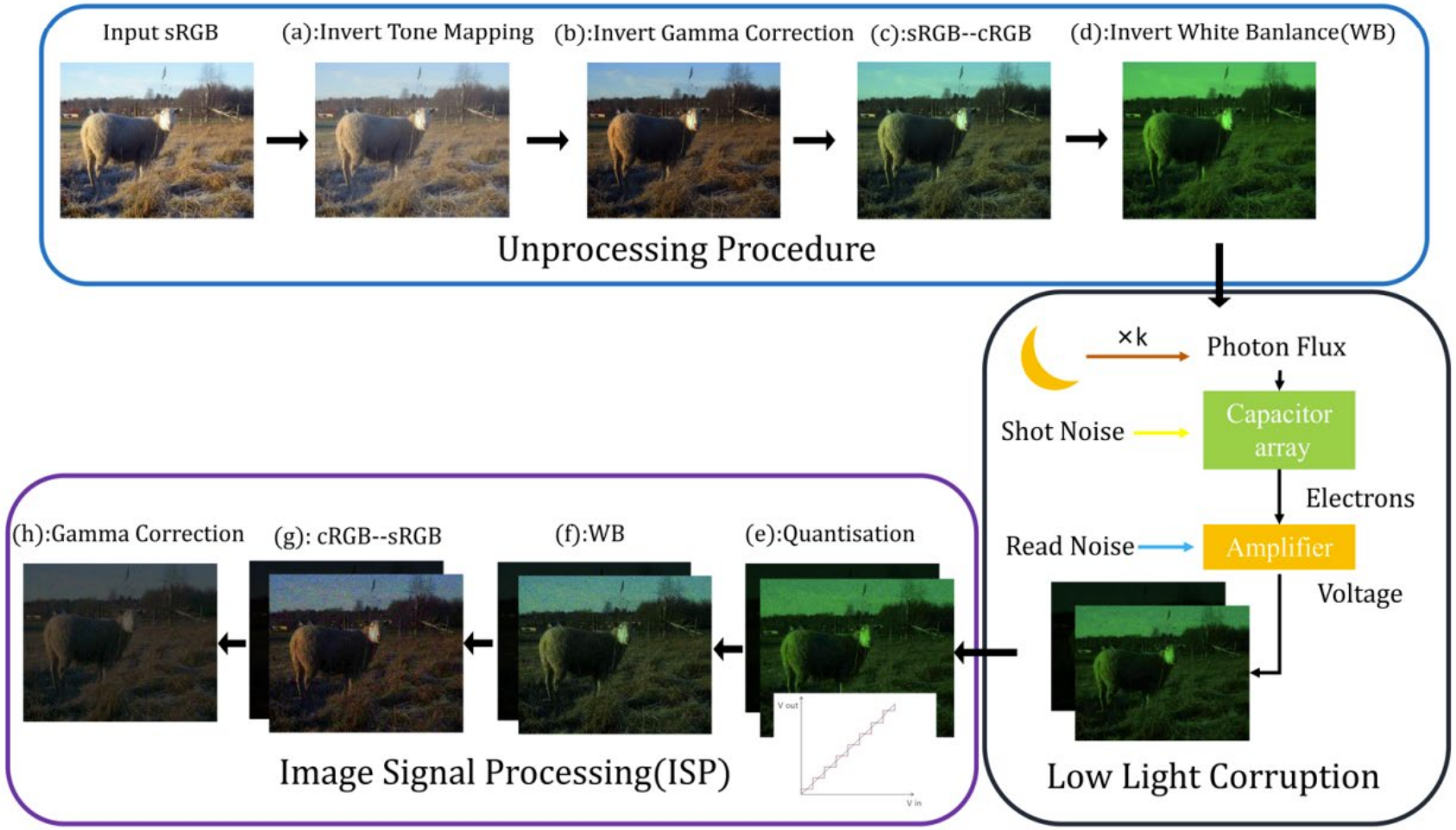
Quantisation Step (ADC):



$$x_{quan} \sim U(-\frac{1}{2B}, \frac{1}{2B})$$
$$y_{quan} = x + x_{quan}.$$

Practical Physics Model

- Photo does not directly reflect physical world.



- Enhance the real world dark image by integrating ISP in camera.

- ~90k parameters
- ~0.004s processing speed

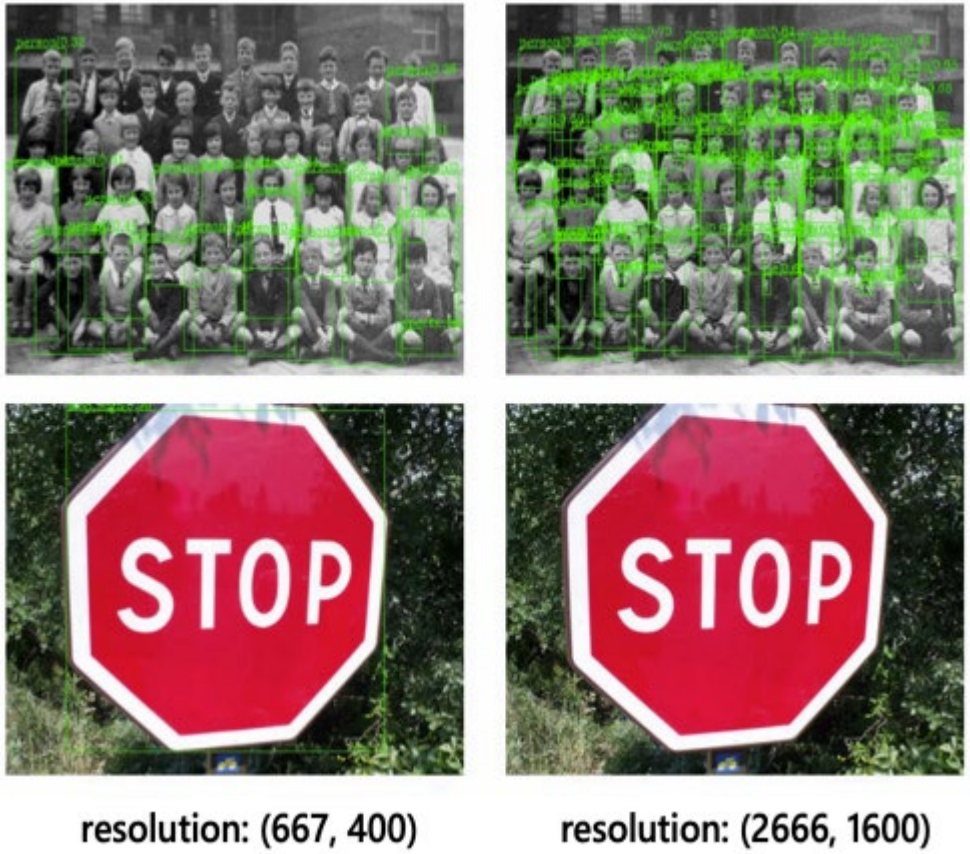
Nature lay hid in Night, apply our IAT, and all was light.

- State-of-The-Art (SOTA)

You Only Need 90K Parameters to Adapt Light: a Light Weight Transformer for Image Enhancement and Exposure Correction BMVC 2023

Scale Variance Bottleneck

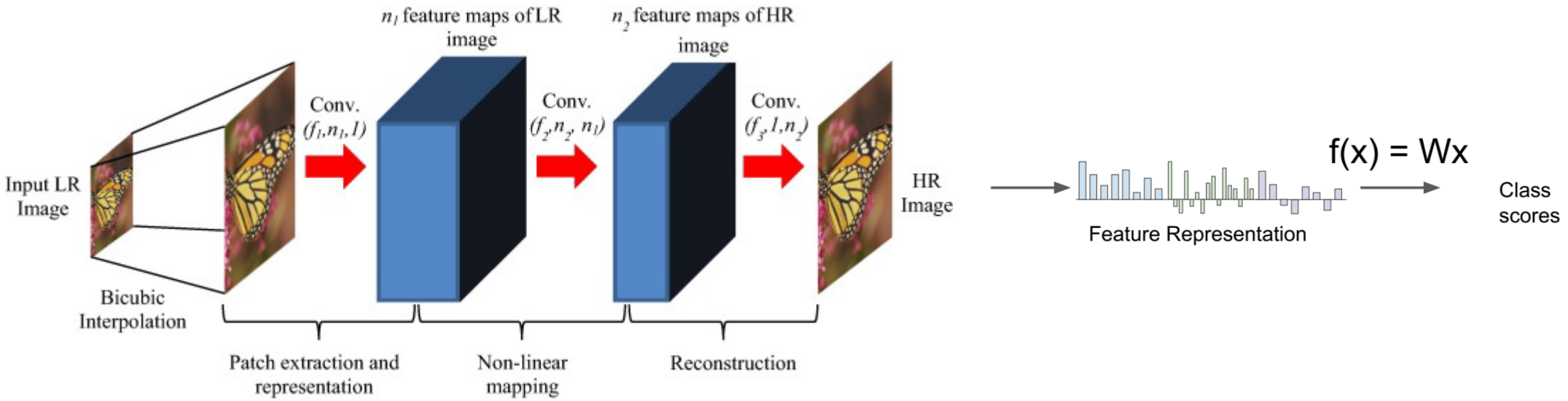
- Scale variance bottleneck haunts deep learning society.



Existing Solution: Super Resolution

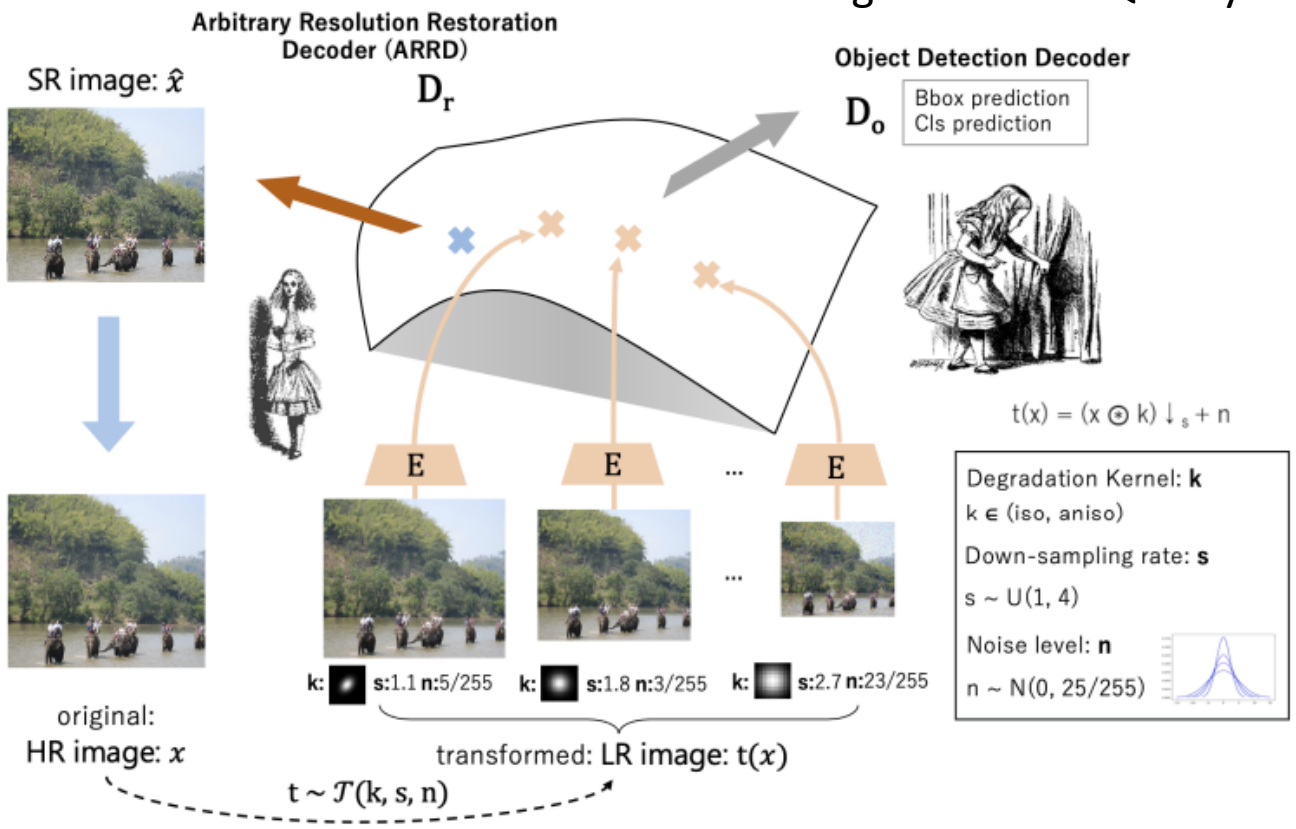
- Super resolution costs much resources, especially in inference stage.

- The optimal up sampling ratio is not known.



Our Solution: Self-supervision

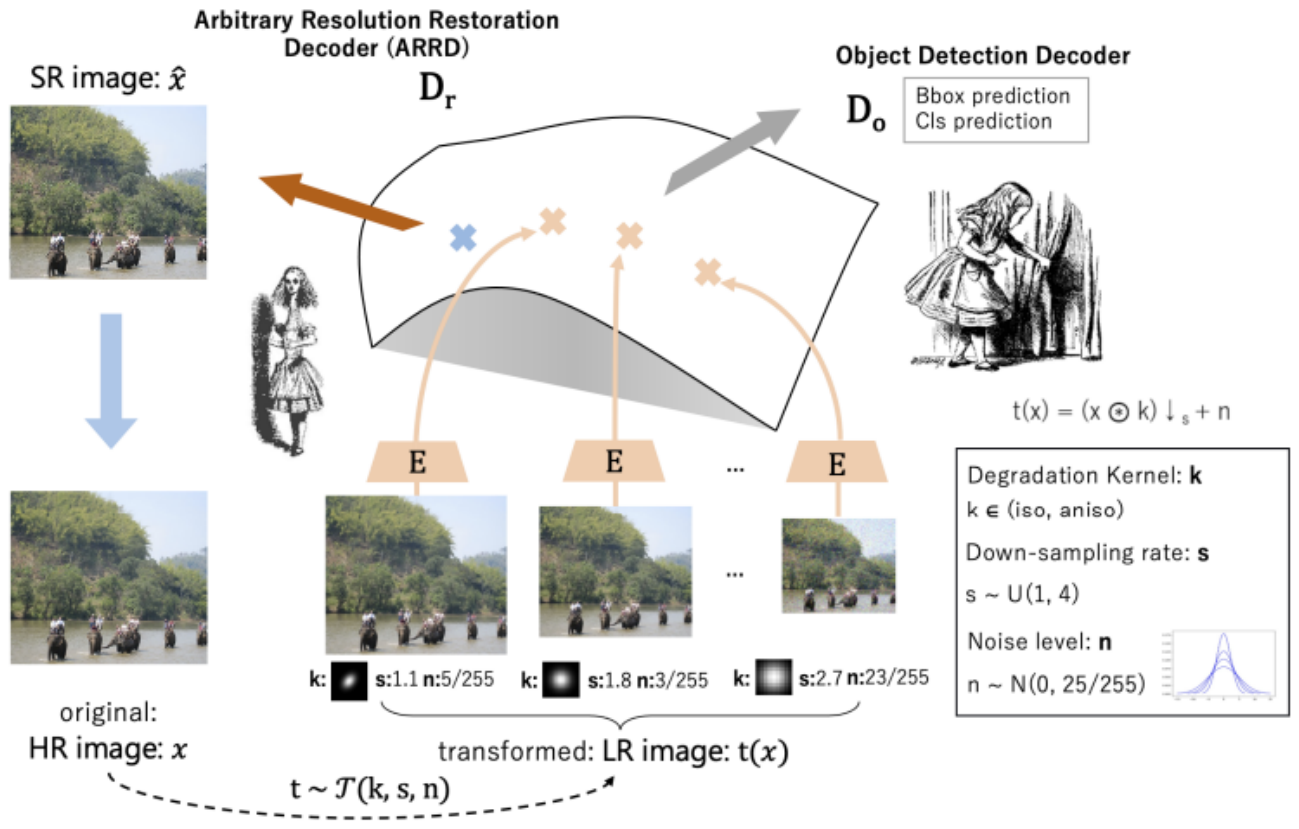
Exploring Resolution and Degradation Clues as Self-supervised Signal for Low Quality Object Detection ECCV 2022



Either being small enough to squeeze through the door (small resolution) or so big to shed a pool of tears (large resolution), Alice is what she is in the world.

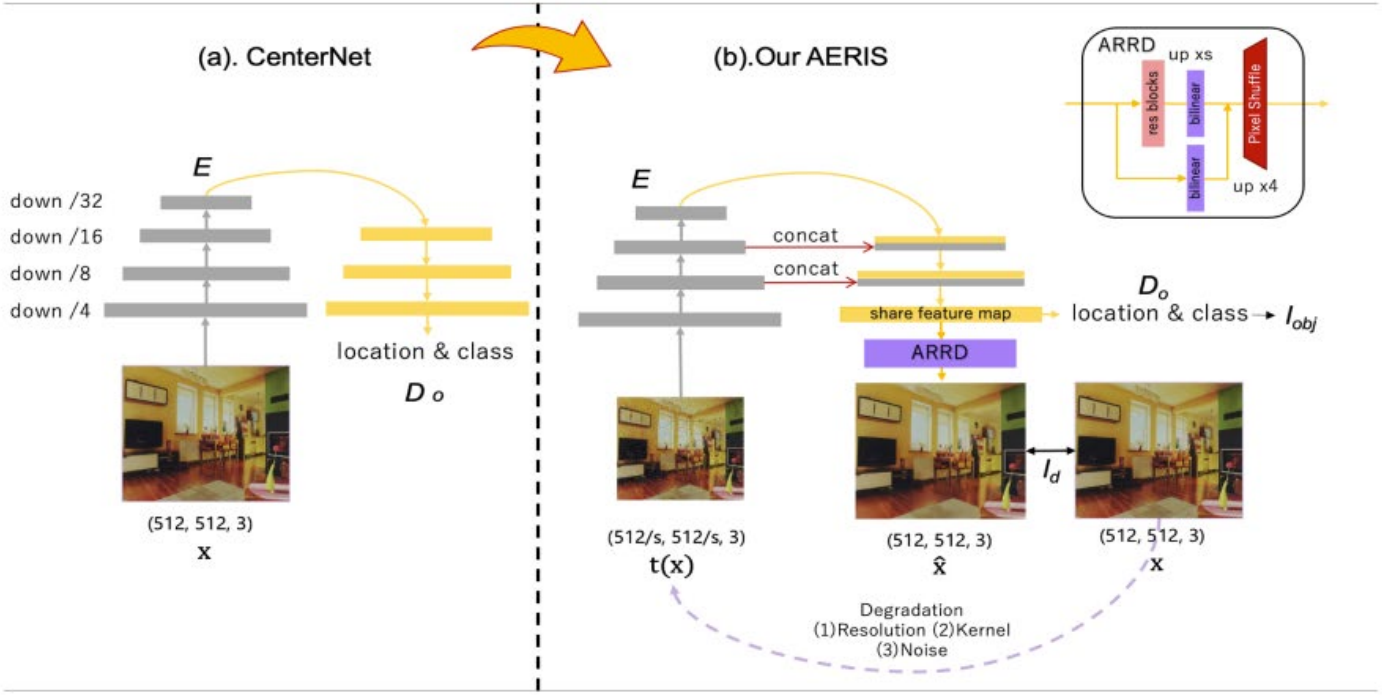
Resolution as Self-supervision Signal

- Solution: Resolution as self-supervised signal.
- Scale invariance manifold.



Resolution as Self-supervision Signal

- Work on main stream CNN and Transformer architecture
- Minimal modification during training.
- Inference stage: unchanged!!!



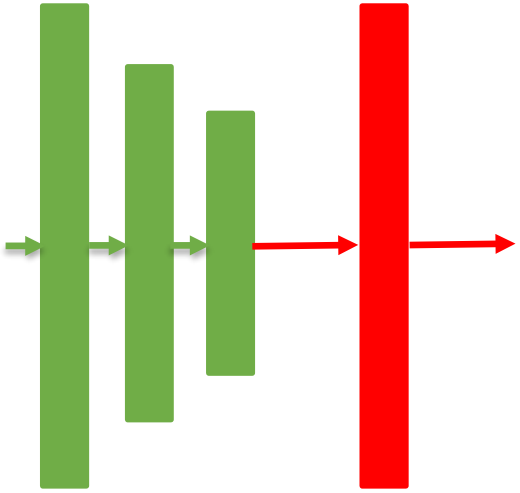
AERIS

Auto Encoding Resolution in Self-supervision (AERIS)

Exploring Resolution and Degradation Clues as Self-supervised Signal for Low Quality Object Detection ECCV 2022

3D Vision Under Radiation

Scene captured under radiation



3D Completion Network

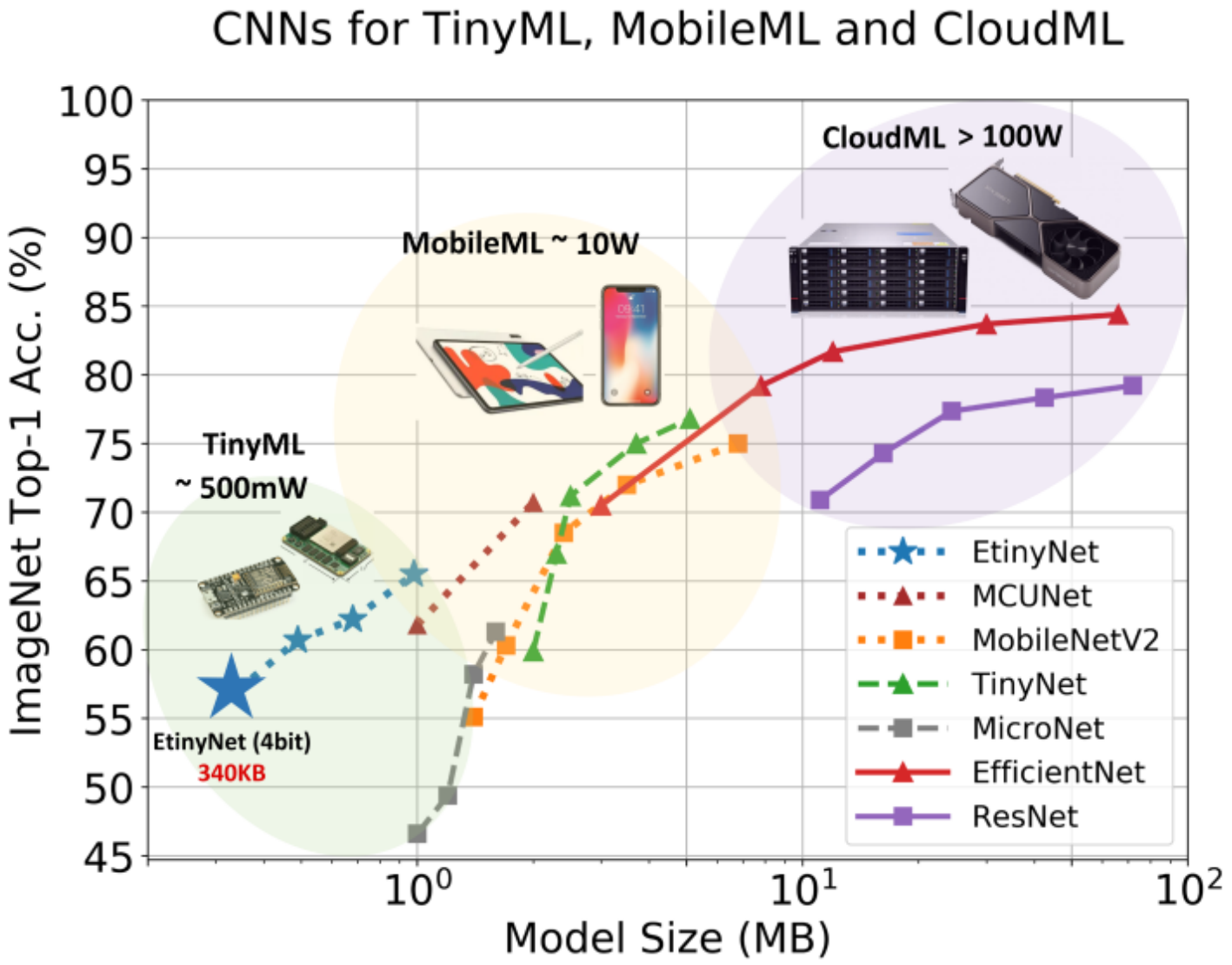
Our recovered radiation-free scan

2D knowledge distillation flow

Gap Among Human

Existing AI implementation depends on expensive GPU cards (~2000$) and reliable power supply (200W).

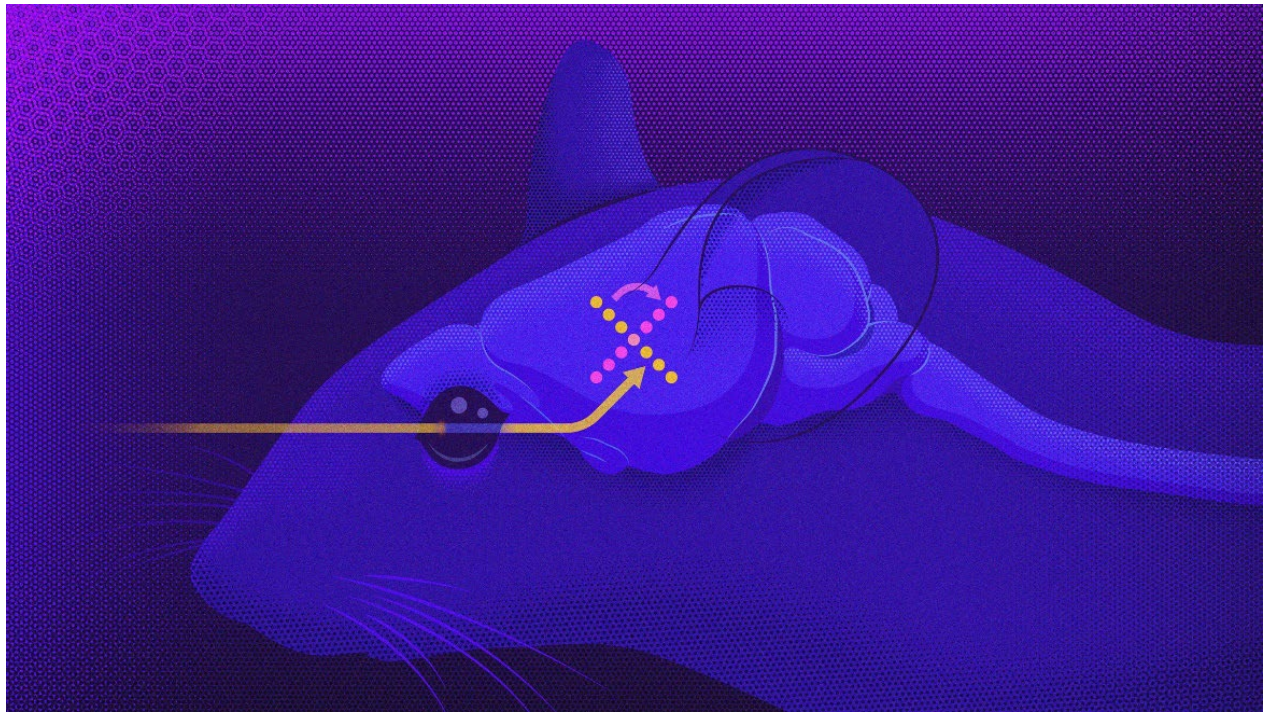
It is difficult to be implemented on robot and in resource poor countries.



Inspired by Neuronscience

Brain rotates sensory representations into orthogonal memory representations over time.

We simulate this into the CNN.



Rotational dynamics reduce interference between sensory and memory representations

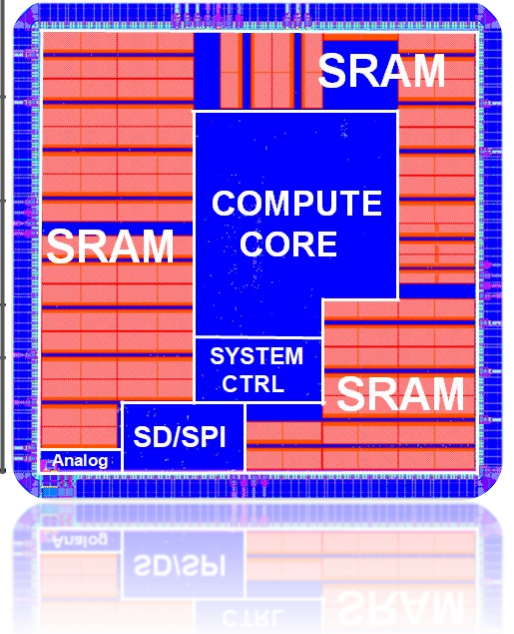
Libby et al. Nature Neuroscience April 05, 2021

Our designed FPGA, only using on-chip SRAM, achieves competitive throughput and reduction of power by 5.6 X.

Our solution achieves roughly10 × processing efficiency gains

17.1 Frames/s/mJ v.s.1.6 Frames/s/mJ

Component	AngelEye	SSA	Ours
Device	ZYNQ XCZ7020	ZYNQ XCZU7EV	Artix7 XC7A100T
On-chip RAM used (KB)	385	1483	576
DRAM used	yes	yes	**none**
ImageNet accuracy (%)	67.72	-	56.0
Model params. (MB)	138	3.5	0.34
Processing time (ms)	354	10.1	**9.7**
Power (W)	3.5	6.2	**0.62**
Efficiency (Frames/s/mJ)	0.0023	1.6	**17.1**

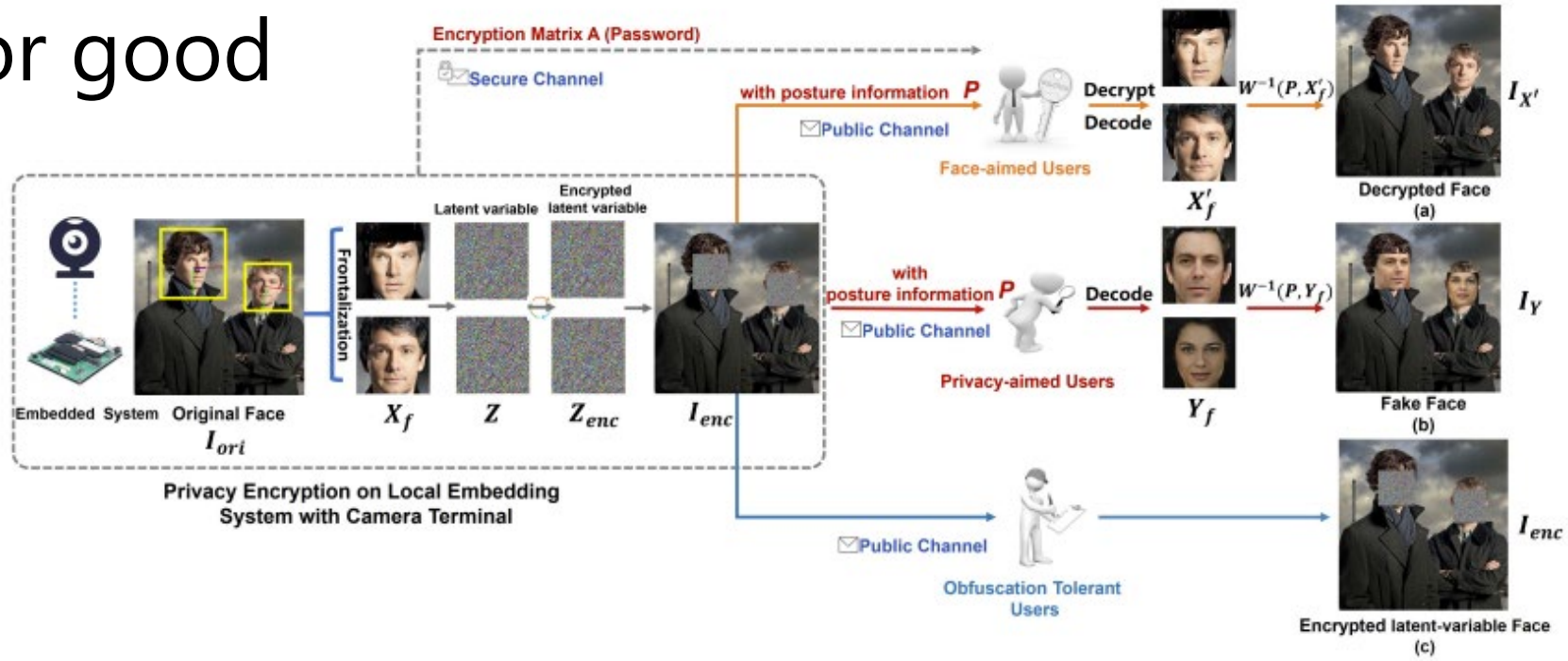


Gap in Heart

- Trustworthy Artificial Intelligence (AI)
- Transparency
- Privacy and Data governance
- Diversity, non-discrimination and fairness

Privacy and Fairness

- Propose systematic solution to protect the privacy and enhance fairness from the camera to the end users.
- Deepfake for good



People taking photos that faces never share: Privacy Protection and Fairness Enhancement from Camera to User AAAI 2023

- All of the codes are available online.
- Thank you for listening.
- lin.gu@riken.jp

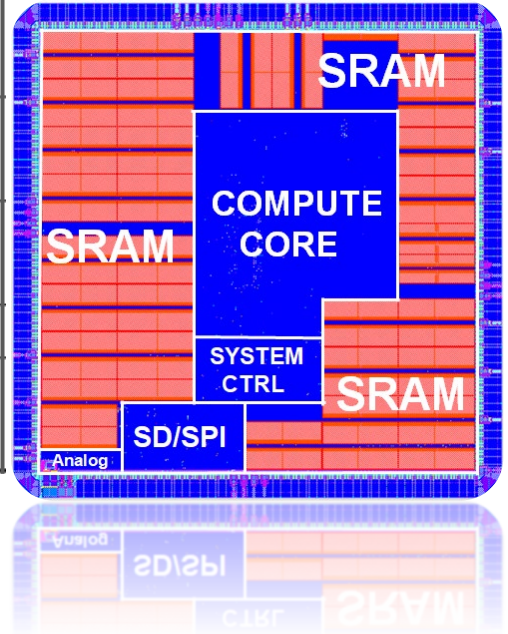
Semi-conduct Manufacturing

Our designed FPGA, only using on-chip SRAM, achieves competitive throughput and reduction of power by 5.6 X.

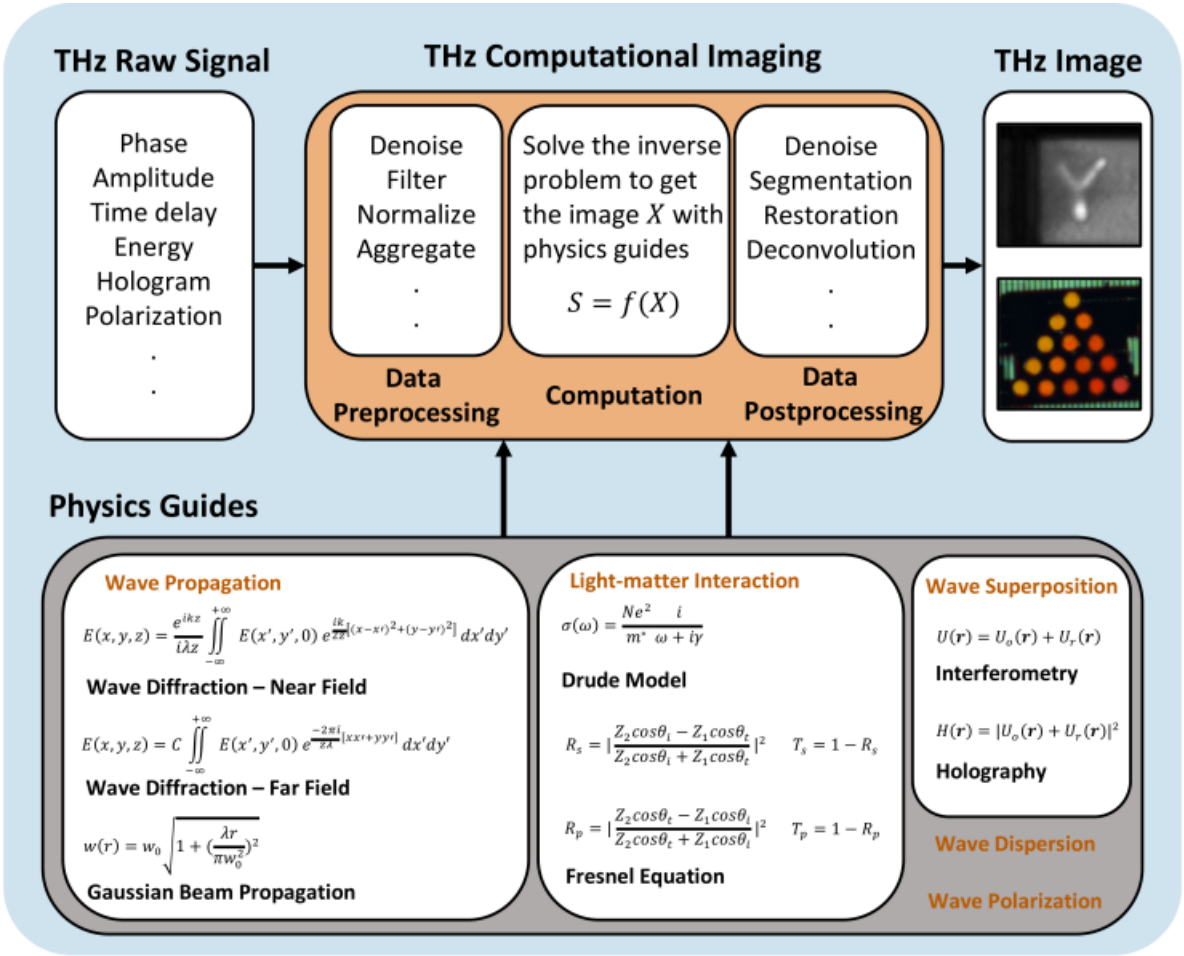
Our solution achieves roughly10 × processing efficiency gains

17.1 Frames/s/mJ v.s.1.6 Frames/s/mJ

Component	AngelEye	SSA	Ours
Device	ZYNQ XCZ7020	ZYNQ XCZU7EV	Artix7 XC7A100T
On-chip RAM used (KB)	385	1483	576
DRAM used	yes	yes	**none**
ImageNet accuracy (%)	67.72	-	56.0
Model params. (MB)	138	3.5	0.34
Processing time (ms)	354	10.1	**9.7**
Power (W)	3.5	6.2	**0.62**
Efficiency (Frames/s/mJ)	0.0023	1.6	**17.1**

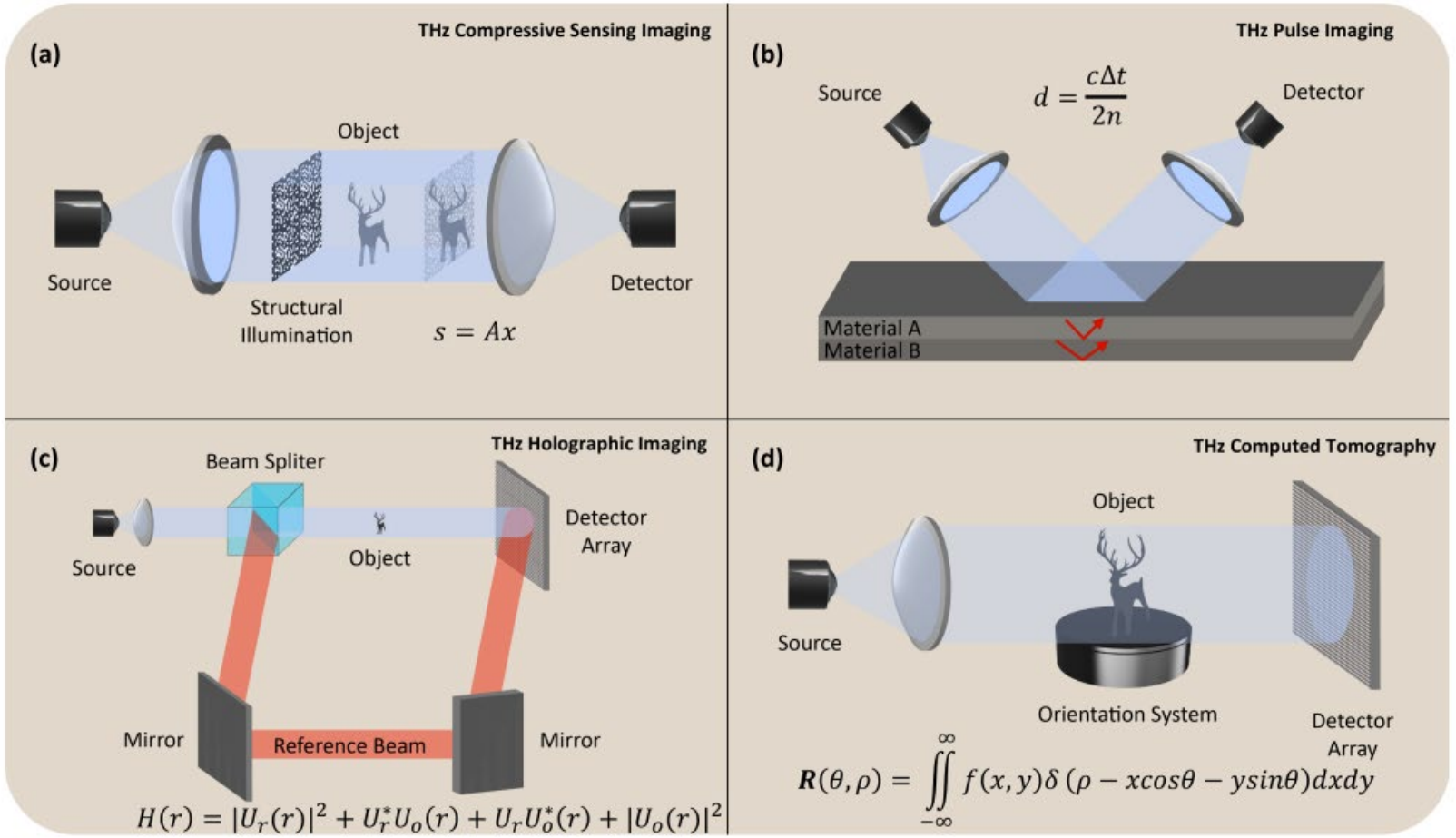


Terahertz Imaging



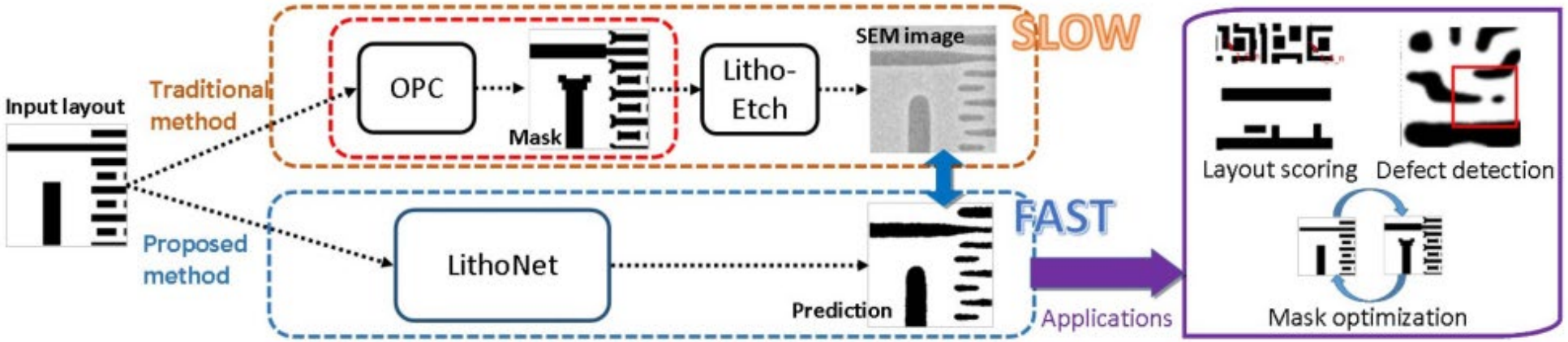
[Physics-guided Terahertz Computational Imaging et al.]

Terahertz Imaging



Physics-guided Terahertz Computational Imaging Hung et al.

Semi-conduct Manufacturing



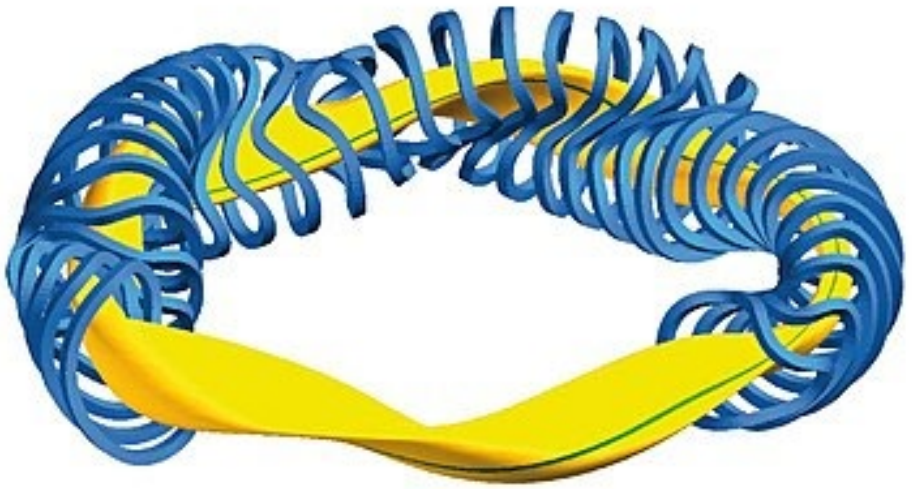
he OPC step, highlighted by the red dashed lines, suggests modifications of a layout mask so that the fabricated IC could have nearly the same shape as the original layout pattern. The proposed LithoNet and its applications are highlighted by purple contours.

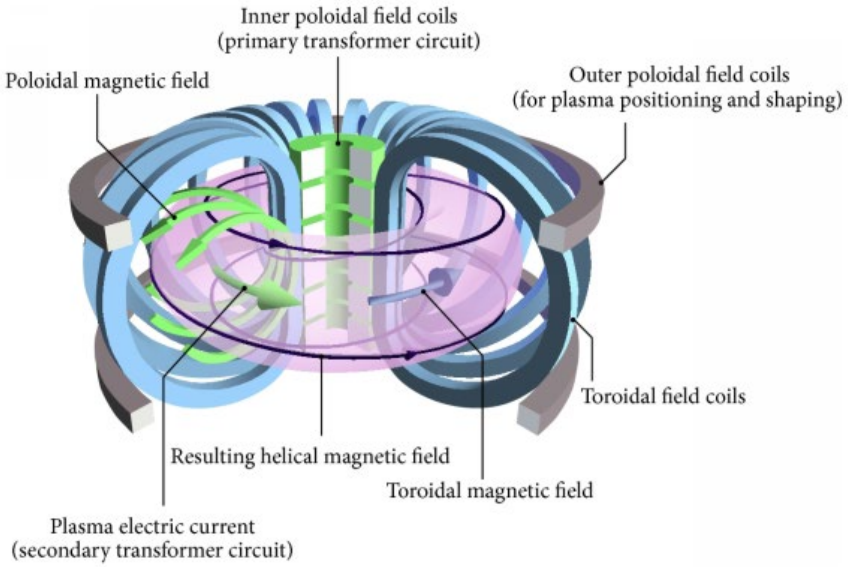
From IC Layout to Die Photograph: A CNN-Based Data-Driven Approach
Shao et al. IEEE TRANSACTIONS ON COMPUTER-AIDED DESIGN OF INTEGRATED CIRCUITS AND SYSTEMS,

Nuclear Fusion

- Collaborating with A*STAR and NTU in Singapore on designing parameters for nuclear fusion device including stellarator and Tokamak.

- Will visit Singapore in December to discuss the details.





Directly Correct Illumination

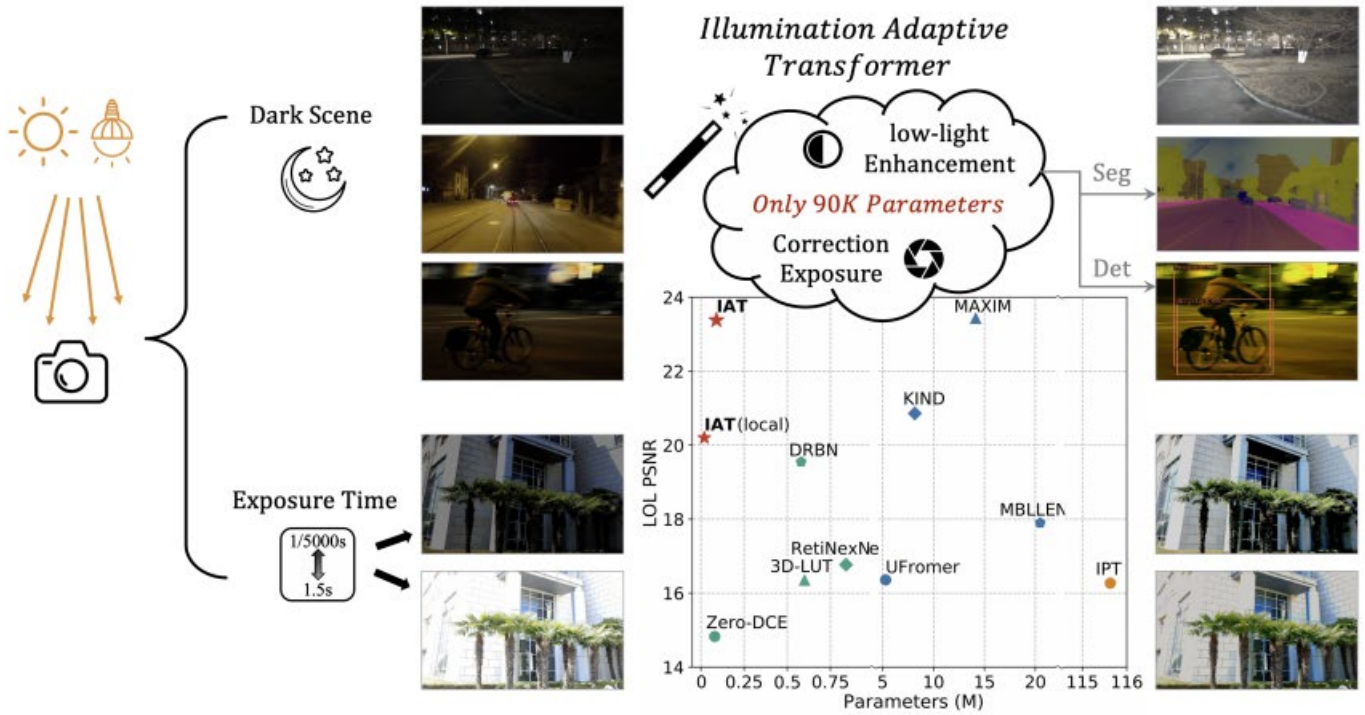
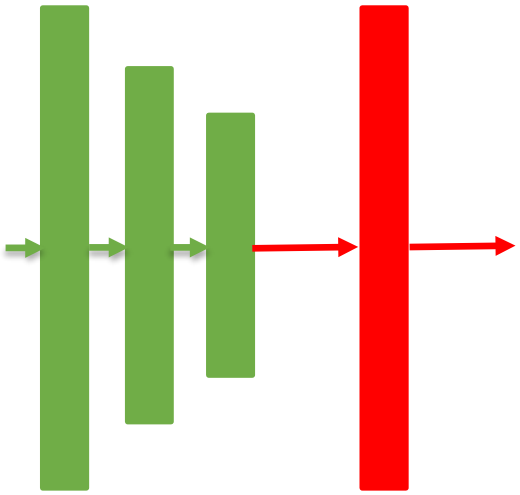


Figure 1: *Nature lay hid in Night*, apply our IAT, *and all was light*, middle figure shows our results compared with other SOTA methods on LOL-V1 dataset [64].

3D Vision Under Radiation

Scene captured under radiation

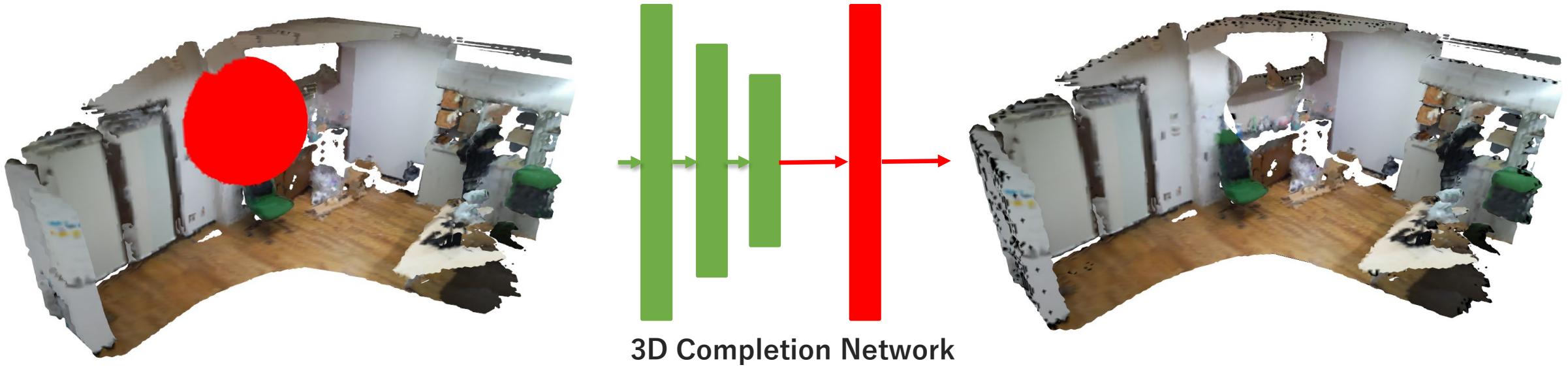


3D Completion Network

Our recovered radiation-free scan

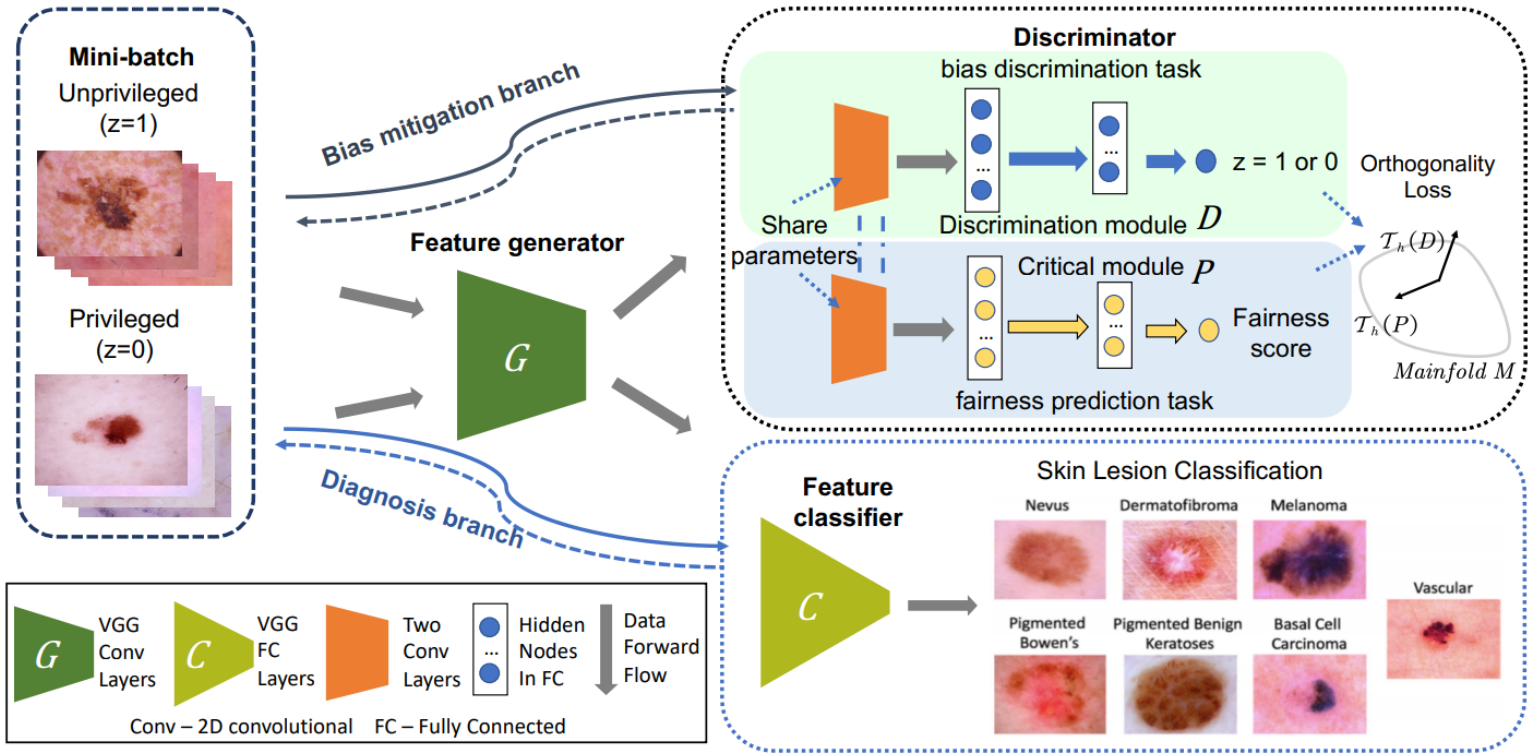
2D knowledge distillation flow

3D Vision Under Fire



Fairness Enhancement

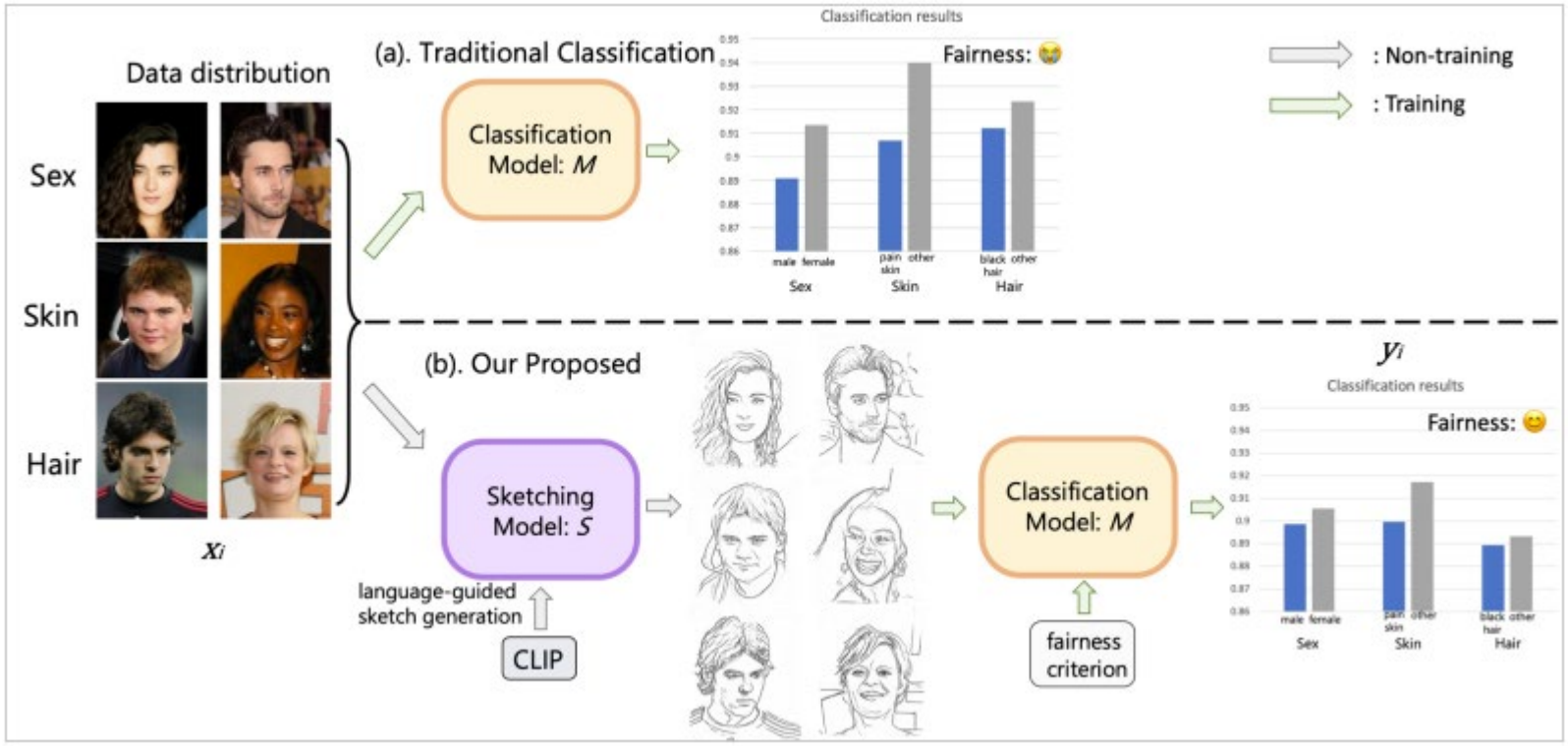
- Simultaneously mitigate and detect biases



Estimating and Improving Fairness with Adversarial Learning
Li X, et al. arXiv:2103.04243

Fairness Enhancement

- Explore the image-to-sketching that maintain useful semantic information for classification while filtering out the useless bias information.

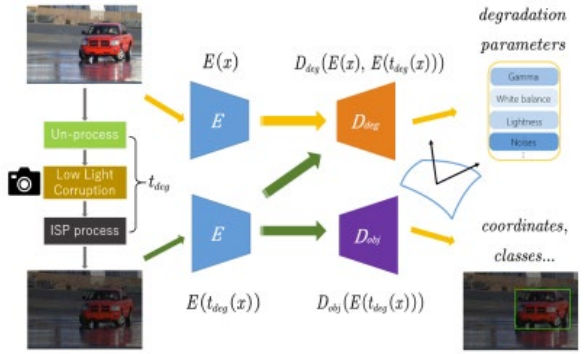


Improving Fairness in Image Classification via Sketching
Yao R, et al. Neurips 2023 workshp

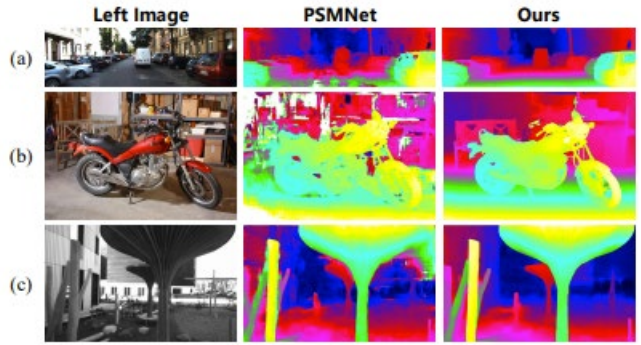
- Fairness and accountability are two essential pillars for trustworthy Artificial Intelligence (AI).

Dark

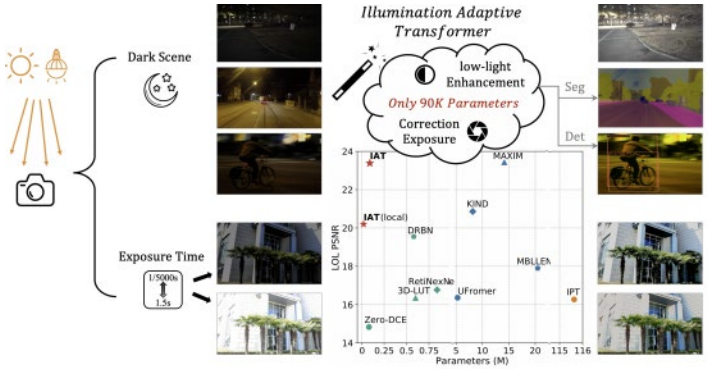
- Low light enhancement, super resolution, denosing, debluring, ….



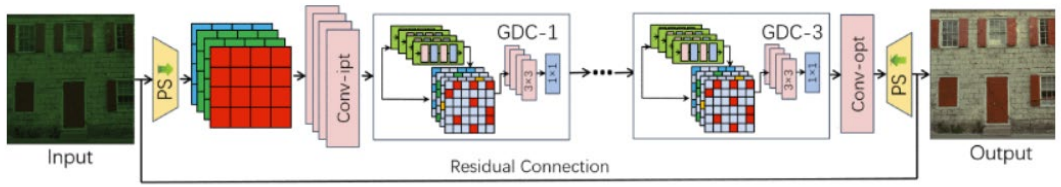
Multitask AET with Orthogonal Tangent Regularity for Dark Object Detection ICCV 2021



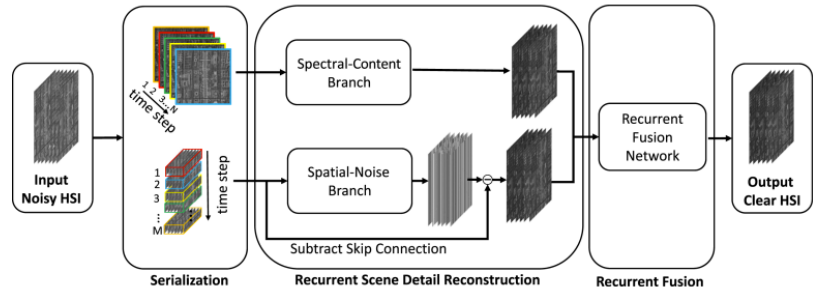
Revisiting Domain Generalized Stereo Matching Networks from a Feature Consistency Perspective CVPR 2022



You Only Need 90K Parameters to Adapt Light: A Light Weight Transformer for Image Enhancement and Exposure Correction BMVC 2022



Memory-Efficient Deformable Convolution Based Joint Denoising and Demosaicing for UHD Images TNNLS 2022



DnRCNN: Deep Recurrent Convolutional Neural Network for HSI Destriping TNNLS 2022