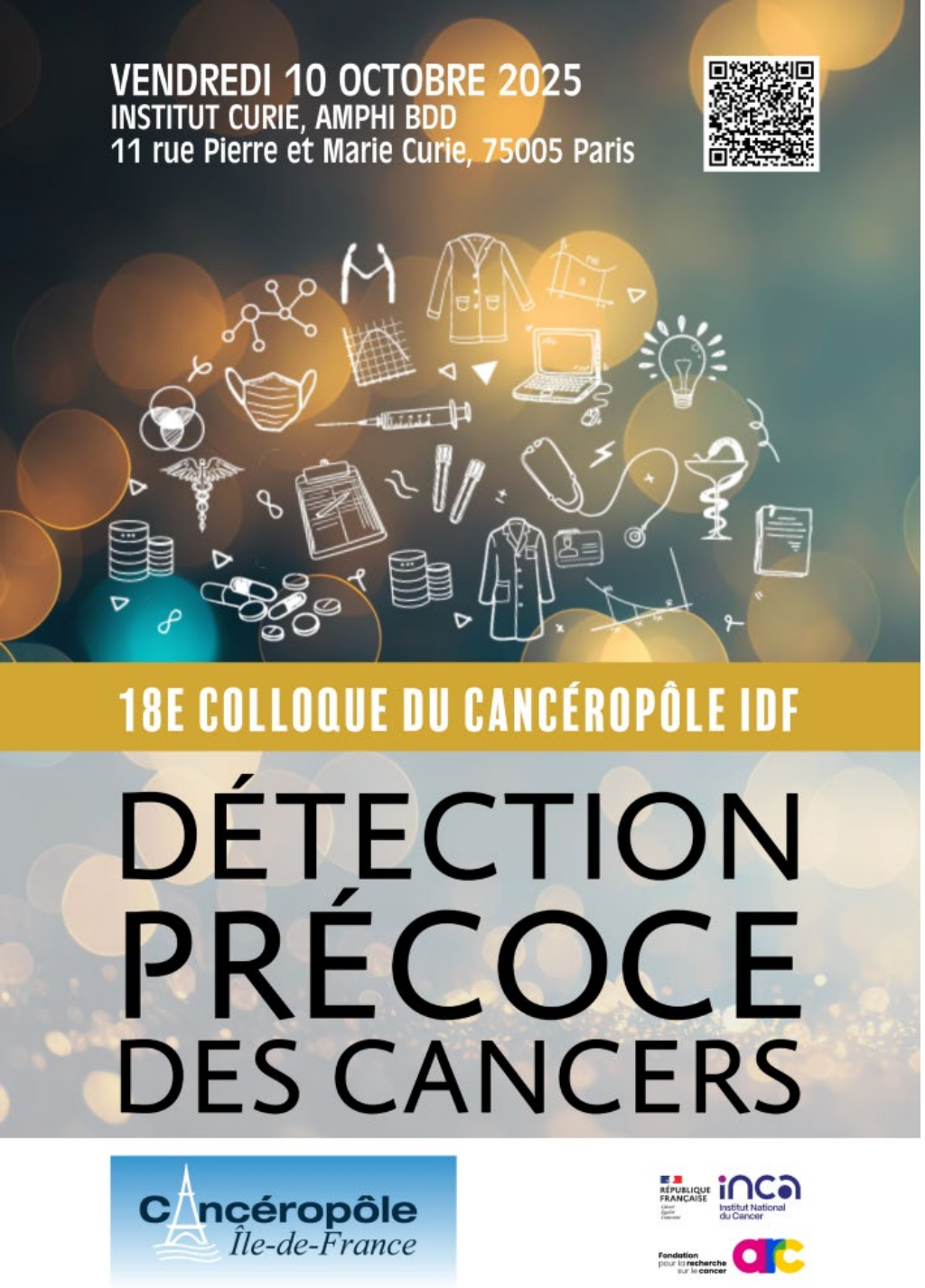


Deep Learning for Tumor Detection and Segmentation: Advances and Challenges

Wenxia Wu

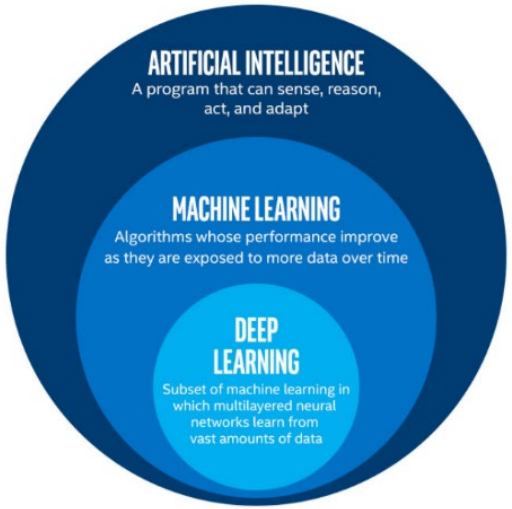
PhDc in Gustave Roussy Inserm U1030

Supervisor: Eric DEUTSCH

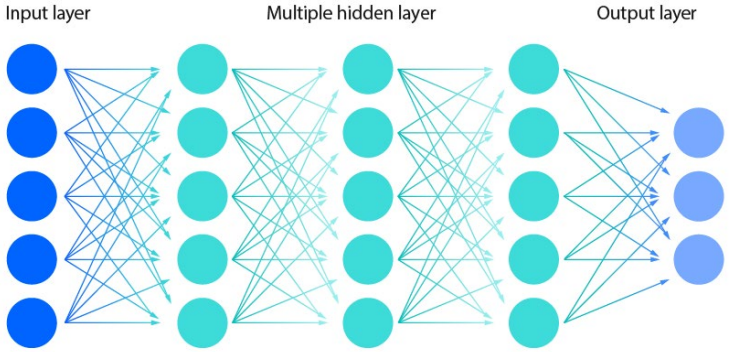
10 Oct. 2025

Deep Learning in the ML Landscape

Subfield of machine learning



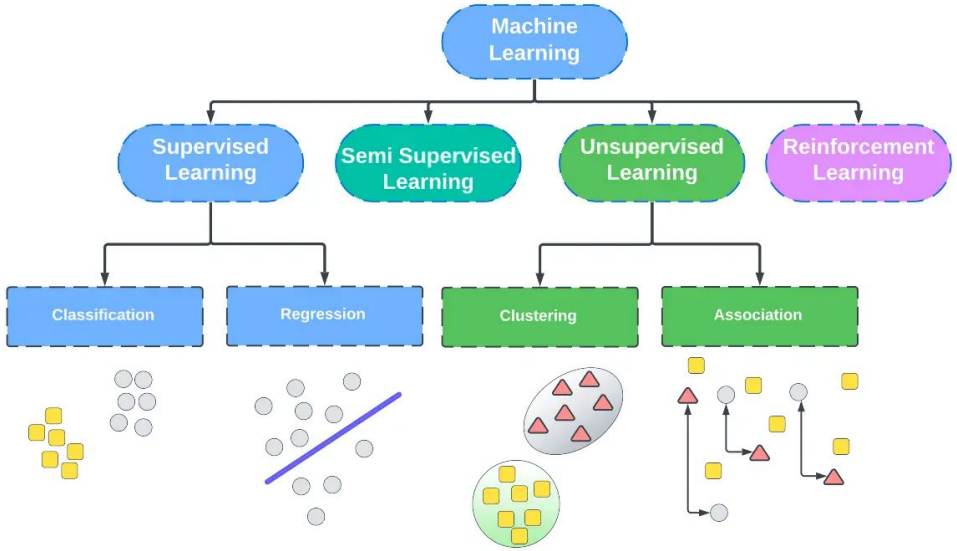
Multi-layer neural nets



Supervised

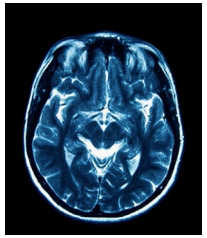
Unsupervised/self-supervised

Weak/semi-supervised



Deep Learning in the ML Landscape

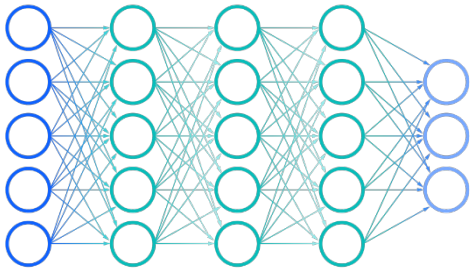
Machine Learning



Input

Feature Extraction

Shape
Texture
Intensity

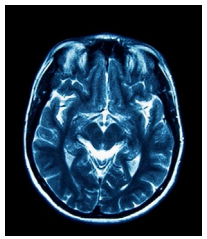


Classification

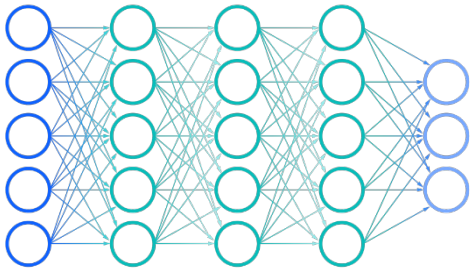
Normal?
Abnormal?

Output

Deep Learning

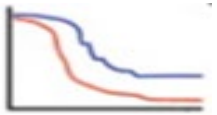


Input



Feature Extraction + Classification

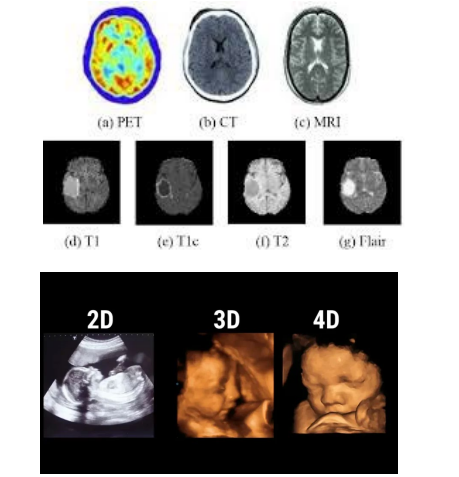
Normal?
Abnormal?

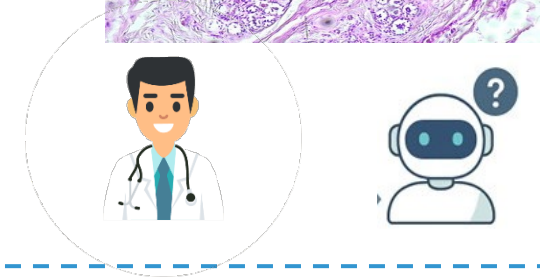


Output

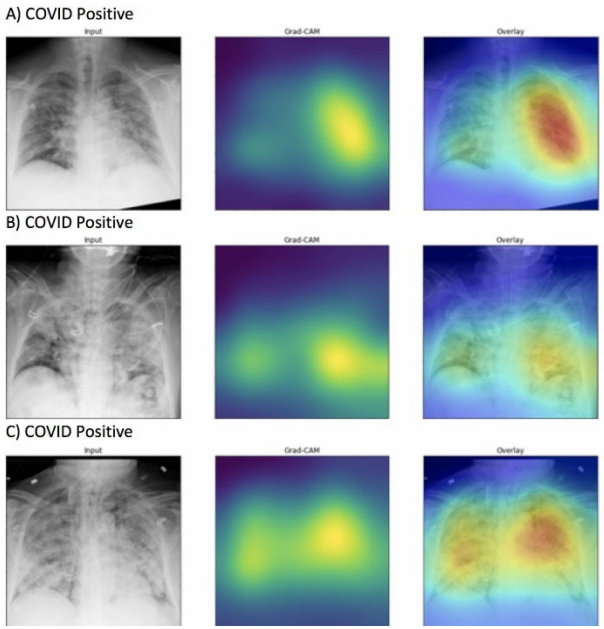
Clinical Value: Second Reader, Scale & Safety

Complex, large-scale clinical data

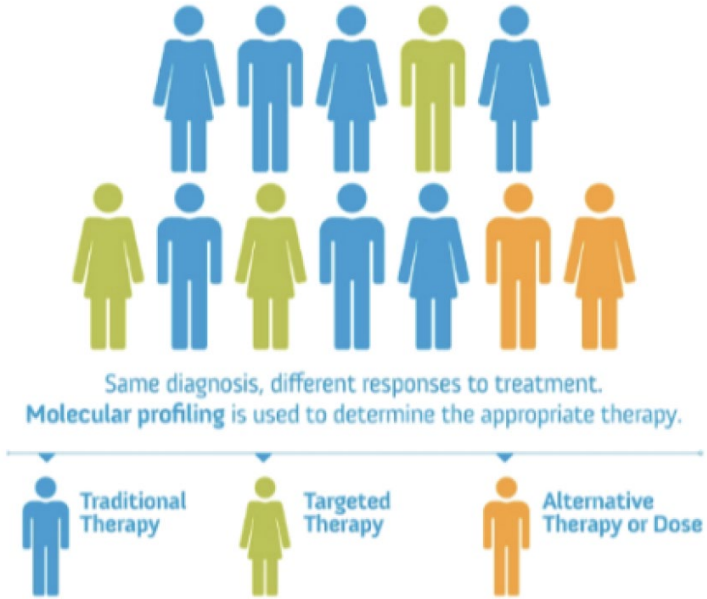




Novel biomarkers

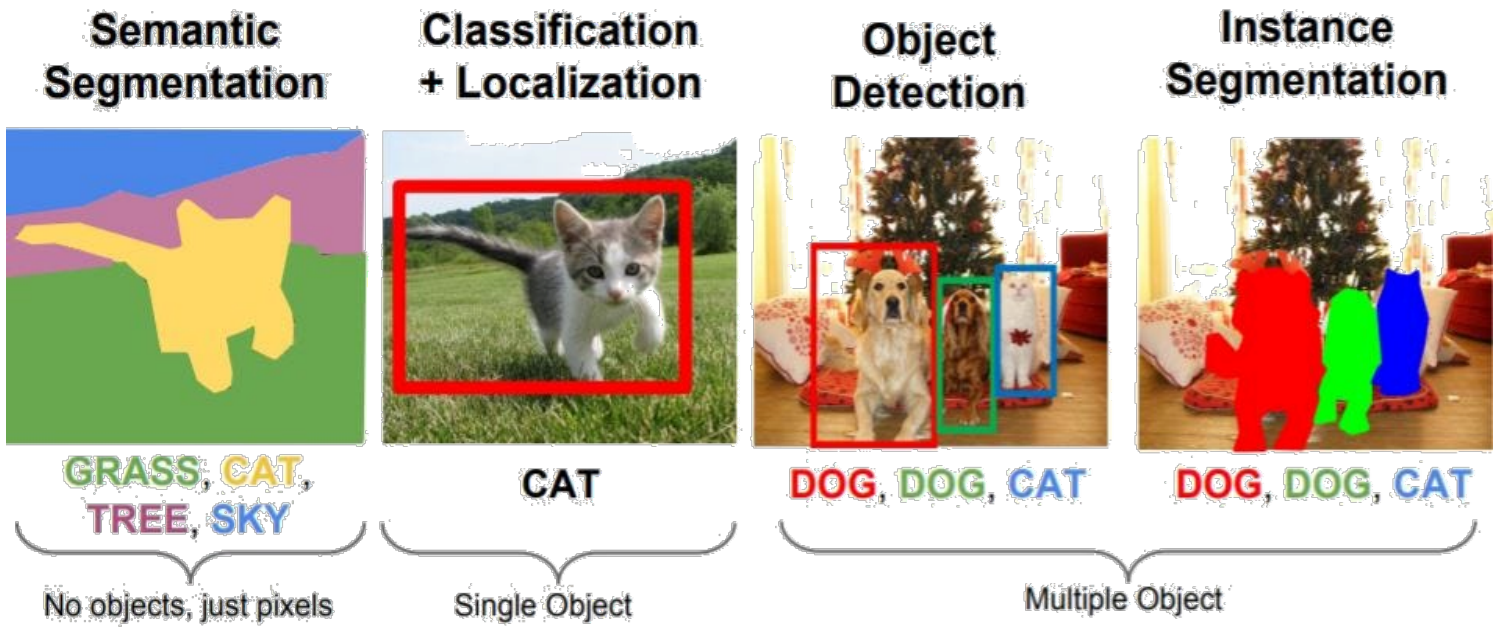


Personalized care

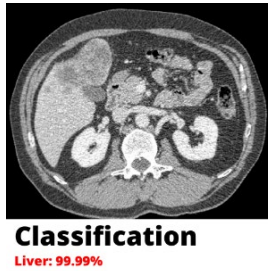


Detection & Segmentation

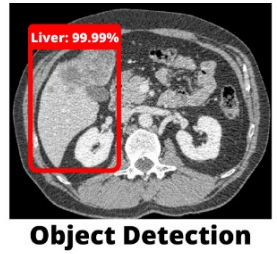
- Deep learning task on natural images



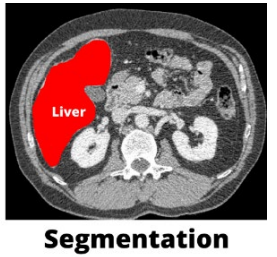
Is there a lesion?



Where is it?



What exactly is its boundary?



http://cs231n.Stanford.edu/slides/2017/cs231n_2017_lecture11.pdf

Detection & Segmentation

Comparison

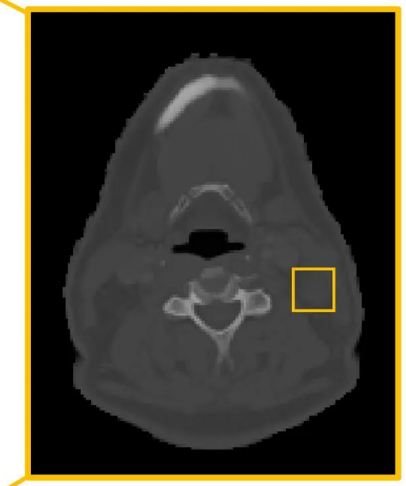
Uses bounding boxes to indicate locations

Identifies general position without boundaries

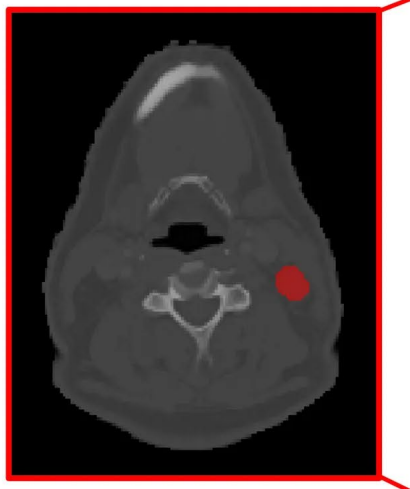
Less personalized

Treatment planning, response evaluation

Detection



Segmentation



Pixel-level analysis to outline boundary

High refinement with precise boundary outlining

Highly personalized

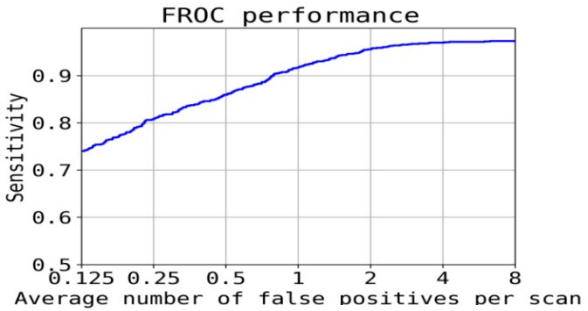
Supports automatic radiotherapy workflows, guiding decisions between involved field vs. elective lymph node irradiation.

Facilitates quantitative analysis over time, including volume analysis and biomarker research (radiomics)

Wu W, et al. Front Immunol 2025

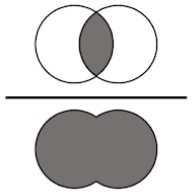
Evaluation

Detection:FROC

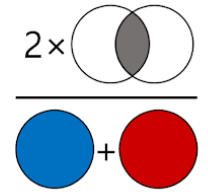


Segmentation:IoU/Dice

IoU



Dice Coefficient



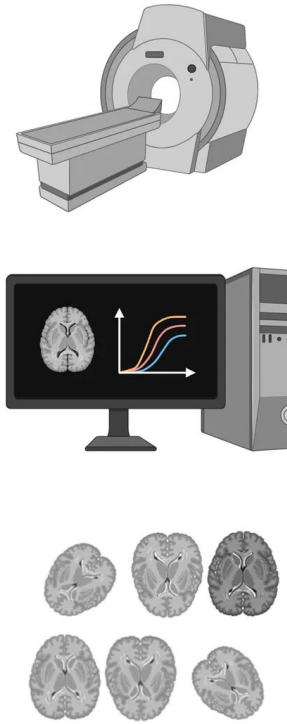
Detection & Segmentation

Data Preparation

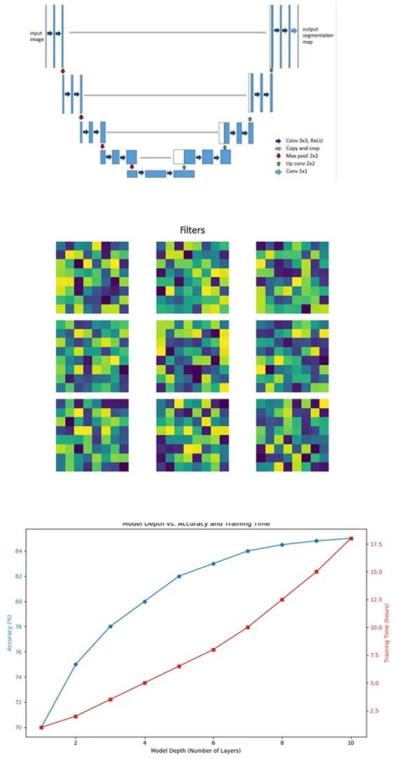
Network Architecture

Model Training

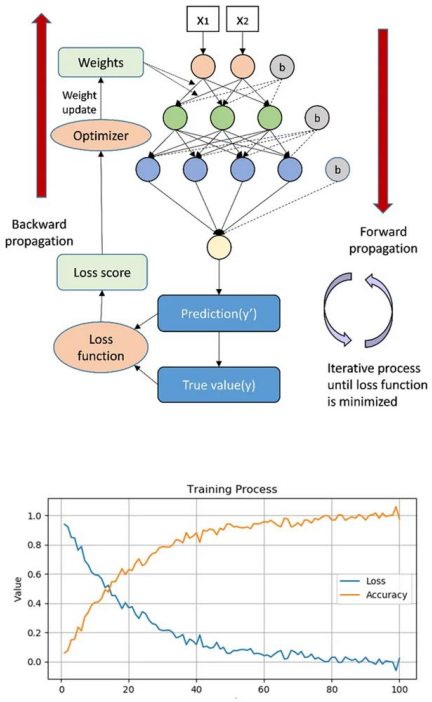
Performance Evaluation



(a)

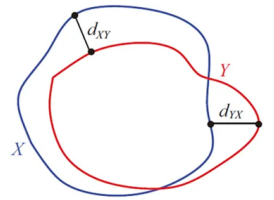


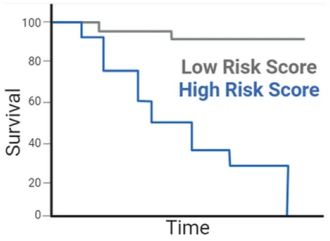
(b)



(c)

Real \ Predict	0	1
0	TN	FP
1	FN	TP





(d)

Deep learning medical image analysis pipeline

Wu W, et al. Front Immunol 2025

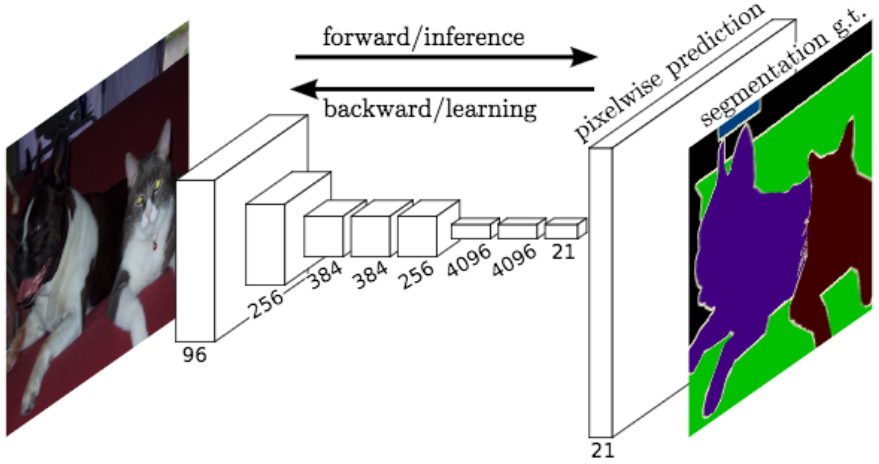
Segmentation: From FCN to U-Net to Transformers

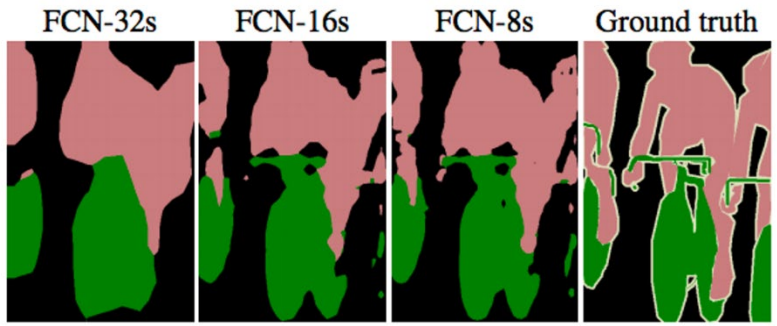
Fully convolutional networks for semantic **segmentation**

J Long, E Shelhamer, T Darrell - Proceedings of the IEEE ..., 2015 - openaccess.thecvf.com

... and transfer their learned representations by fine-tuning [3] to the **segmentation** task. We then ... Our **fully convolutional network** achieves stateof-the-art **segmentation** of PASCAL VOC (20...

☆ Save 99 Cite Cited by 58855 Related articles All 45 versions Web of Science: 3354 »



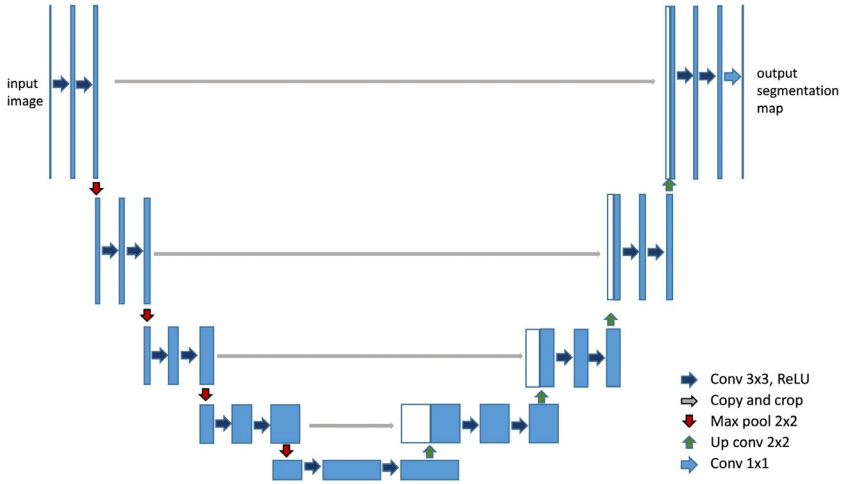


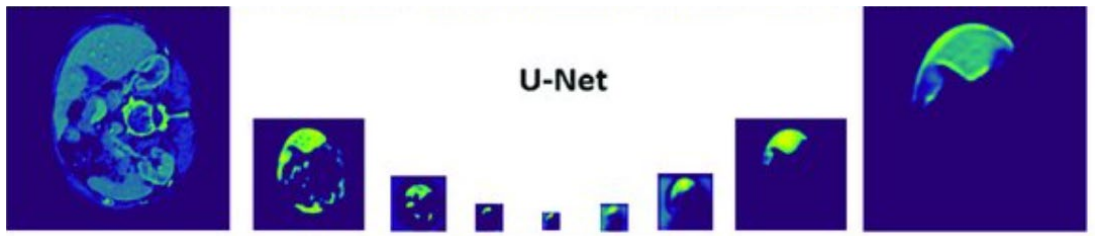
U-net: Convolutional networks for **biomedical image segmentation**

O Ronneberger, P Fischer, T Brox - ... Conference on Medical **image** ..., 2015 - Springer

... We demonstrate the application of the **u-net** to three different **segmentation** tasks. The first task is the **segmentation** of neuronal structures in electron microscopic recordings. An ...

☆ Save 99 Cite Cited by 120989 Related articles All 28 versions





Segmentation: From FCN to U-Net to Transformers

nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation

F Isensee, PF Jaeger, SAA Kohl, J Petersen... - Nature ..., 2021 - nature.com

... **nnU-Net**, we systematically tested the performance of common pipeline variations by systematically modifying some of **nnU-Net**'... against our default **nnU-Net** configuration, which ...

☆ Save 99 Cite Cited by 7188 Related articles All 9 versions Web of Science: 4986

nnu-net revisited: A call for rigorous validation in 3d medical image segmentation

F Isensee, T Wald, C Ulrich, M Baumgartner... - ... Conference on Medical ..., 2024 - Springer

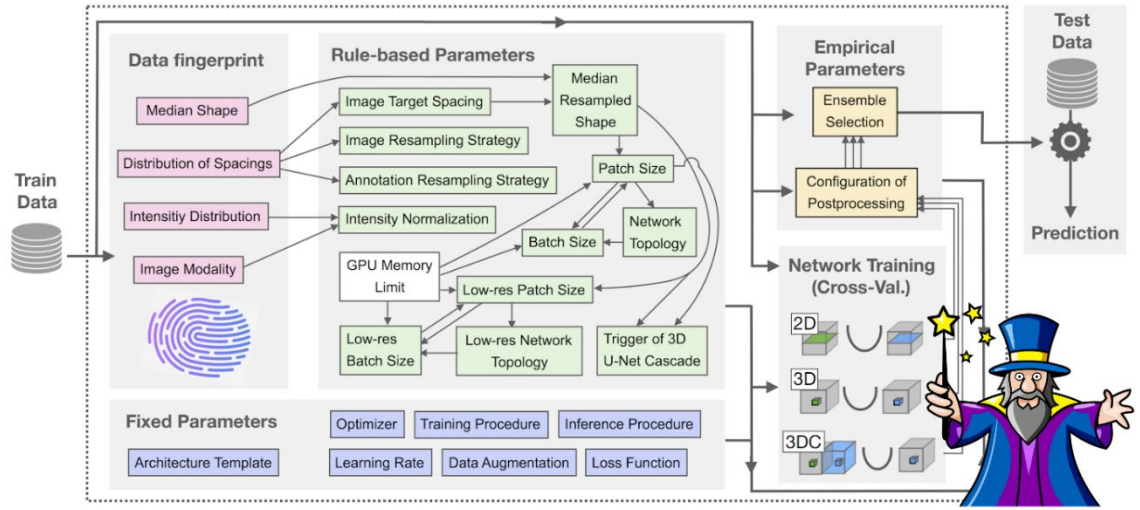
... P1), we introduce new **nnU-Net** ResEnc presets, which use **nnU-Net**'s existing automatic ... in the **nnU-Net** framework except SwinUNETR(V1+V2), which we integrate into the **nnU-Net** ...

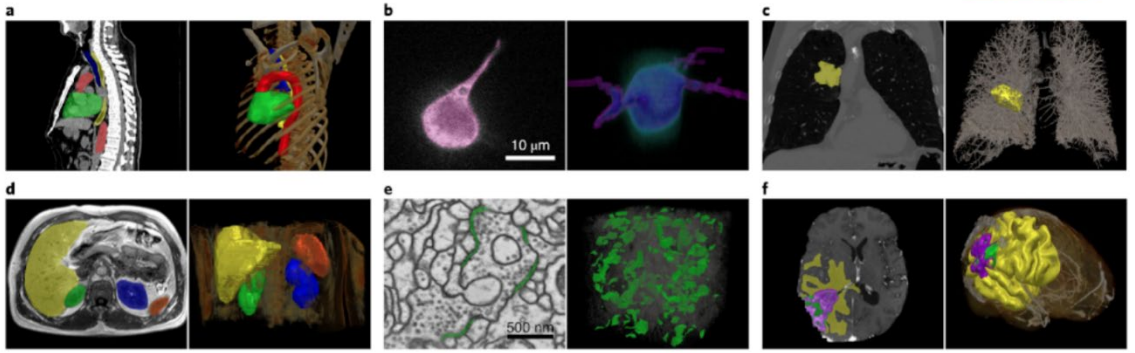
☆ Save 99 Cite Cited by 292 Related articles All 7 versions

MIC-DKFZ/**nnUNet**

MIC MEDICAL IMAGE COMPUTING

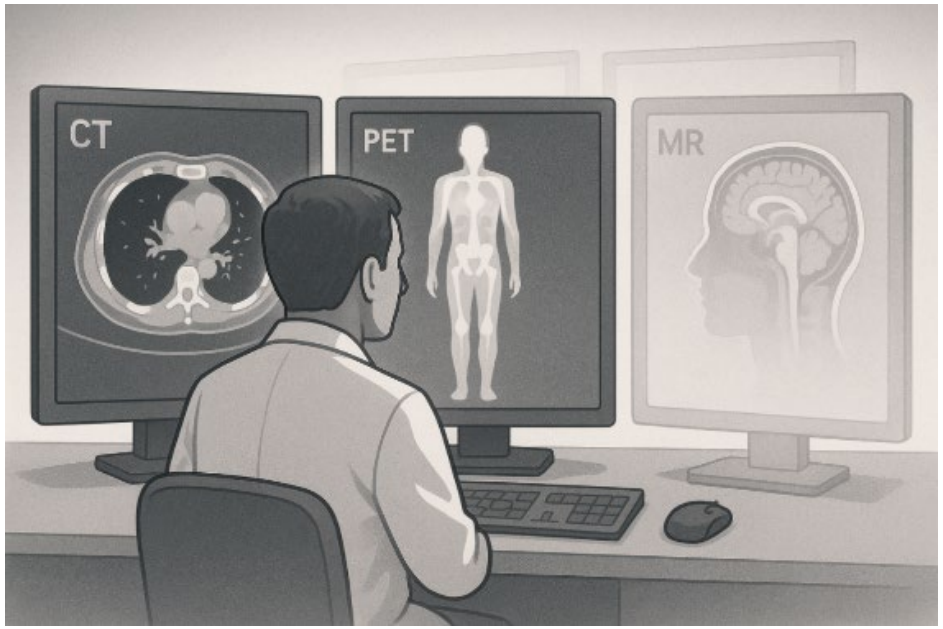
78 Contributors · 327 Issues · 290 Discussions · 7k Stars · 2k Forks



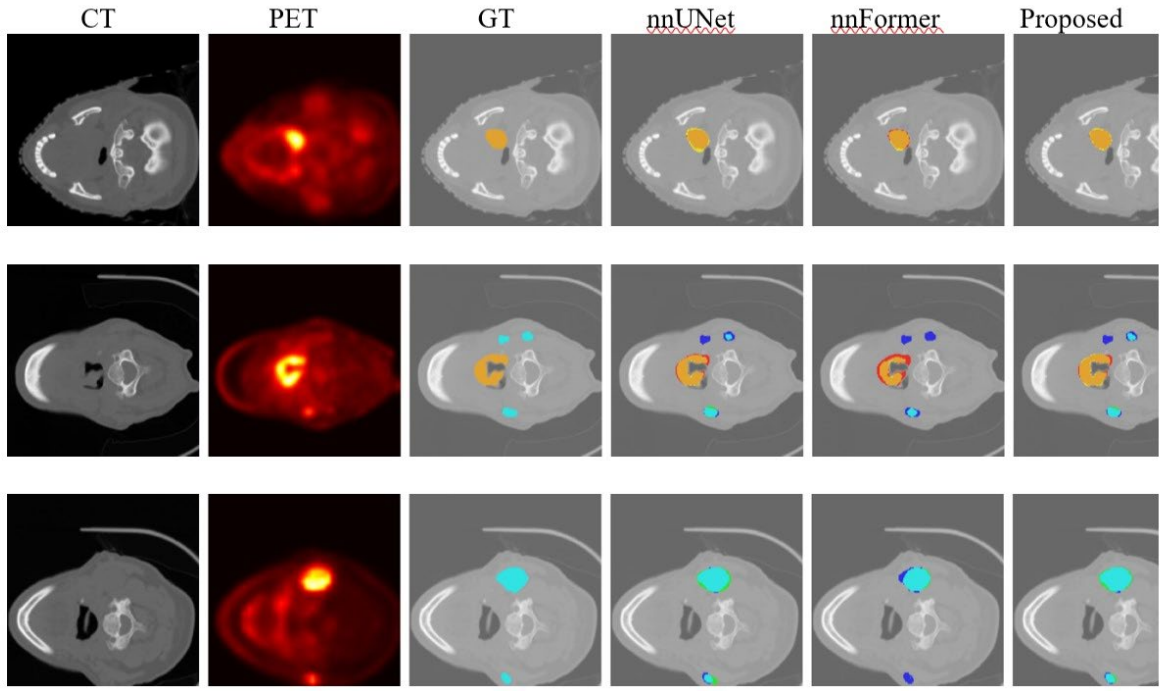


Segmentation: From FCN to U-Net to Transformers

Head and Neck Cancer GTV Segmentation with Limited PET



Missing Modality



Automated segmentation results. Columns show CT, PET, ground truth, nnU-Net, nnFormer, and the proposed model(Wu W, Sun R, in prep).

Segmentation: From FCN to U-Net to Transformers

[PDF] An image is worth 16x16 words: Transformers for image recognition at scale

A Dosovitskiy, L Beyer, A Kolesnikov, D Weissenborn, X Zhai, T Unterthiner, M Dehghani…

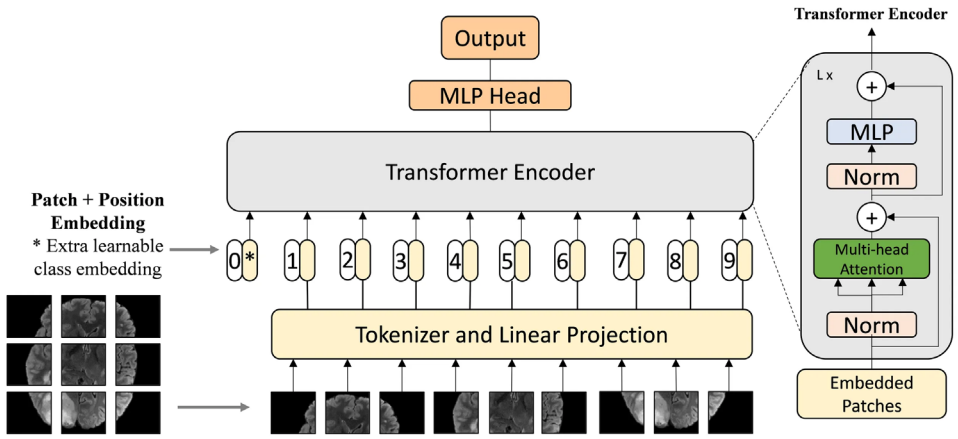
arXiv preprint arXiv:2010.11929, 2020 • arxiv.org

Abstract

While the Transformer architecture has become the de-facto standard for natural language processing tasks, its applications to computer vision remain limited. In vision, attention is either applied in conjunction with convolutional networks, or used to replace certain components of convolutional networks while keeping their overall structure in place. We show that this reliance on CNNs is not necessary and a pure transformer applied directly to sequences of image patches can perform very well on image classification tasks. When

SHOW MORE

☆ Save 99 Cite Cited by 73813 Related articles All 19 versions »

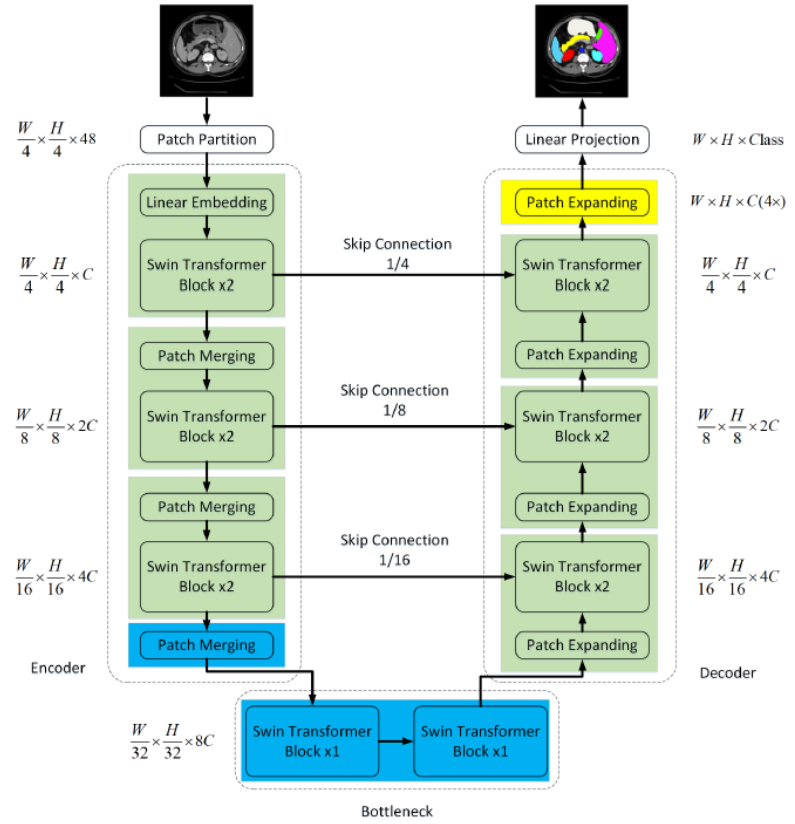


Swin-unet: Unet-like pure transformer for medical image segmentation

H Cao, Y Wang, J Chen, D Jiang, X Zhang… - European conference on …, 2022 - Springer

… **Swin** … **Swin** Transformer's [18] success, we propose **Swin-Unet** to leverage the power of Transformer for 2D medical image segmentation in this work. To our best knowledge, **Swin-Unet** …

☆ Save 99 Cite Cited by 5744 Related articles All 5 versions



Detection

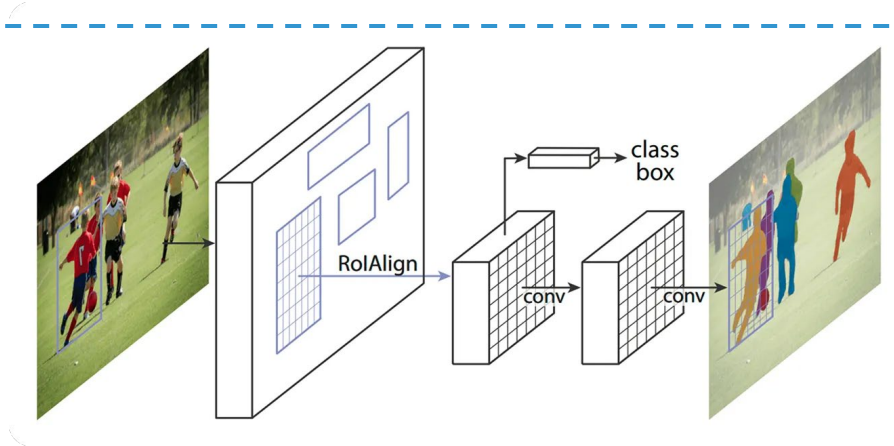
Mask r-cnn

K He, G Gkioxari, P Dollár... - Proceedings of the IEEE ..., 2017 - openaccess.thecvf.com

... **mask**, with minimal modification **Mask R-CNN** can be applied to detect instance-specific poses. Without tricks, **Mask R-CNN** ... **masks** on the top 100 detection boxes, **Mask R-CNN** adds a ...

☆ Save 99 Cite Cited by 45512 Related articles All 22 versions Web of Science: 10319 »

Two-stage: proposals → per-ROI class/box + mask; strong for precise instance delineation and overlapping targets.
Commonly applied on 2D slices; 3D variants exist but are heavier in memory/compute.



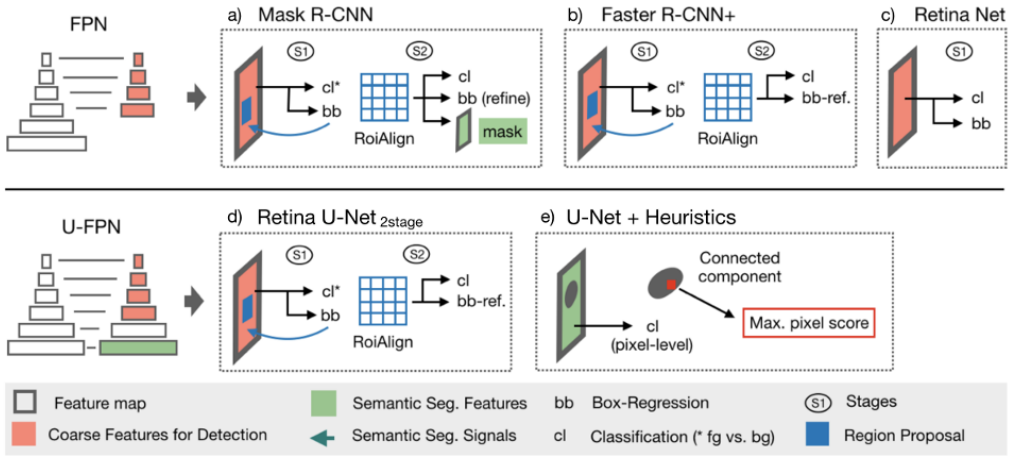
Retina U-Net: Embarrassingly simple exploitation of segmentation supervision for medical object detection

PF Jaeger, SAA Kohl, S Bickelhaupt... - ... learning for health ..., 2020 - proceedings.mlr.press

... **Retina U-Net**, a simple architecture, which naturally fuses the **Retina** Net one-stage detector with the **U-Net** ... Specifically, we propose **Retina U-Net**, a simple but effective method for ...

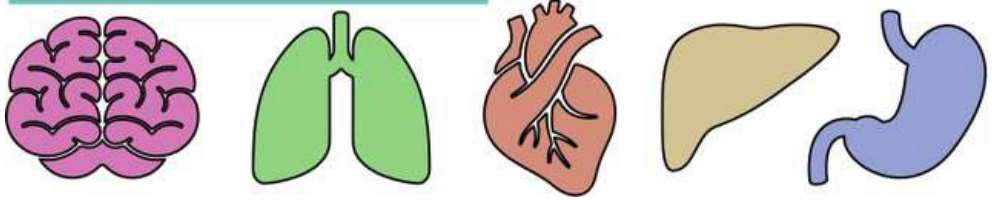
☆ Save 99 Cite Cited by 310 Related articles All 6 versions »

One-stage: RetinaNet-style anchors + U-Net decoder; uses focal loss and leverages segmentation labels to boost detection on small datasets.
Scales naturally to 2D/3D volumes; efficient when positives are rare and you need throughput.



Detection

nnDetection



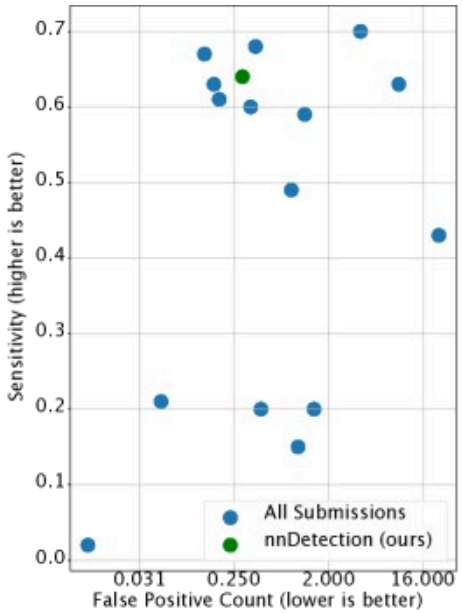
nnDetection: a self-configuring method for medical object detection

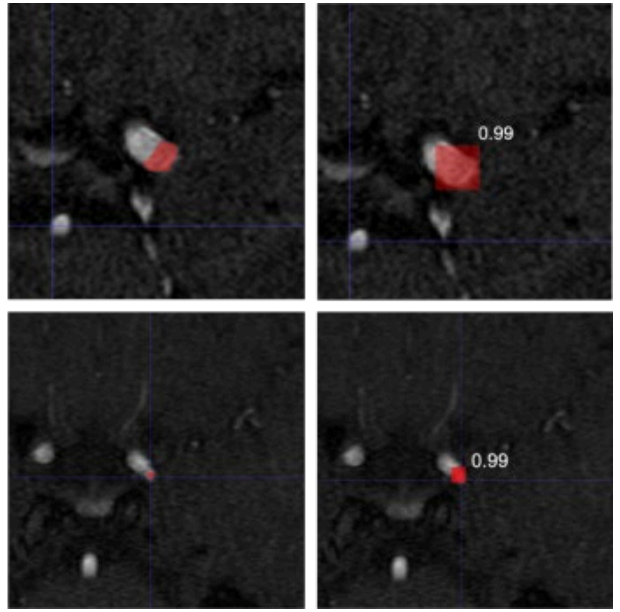
M Baumgartner, PF Jäger, F Isensee... - ... conference on medical ..., 2021 - Springer

... The resulting self-configuring method, **nnDetection**, adapts itself without any manual ...

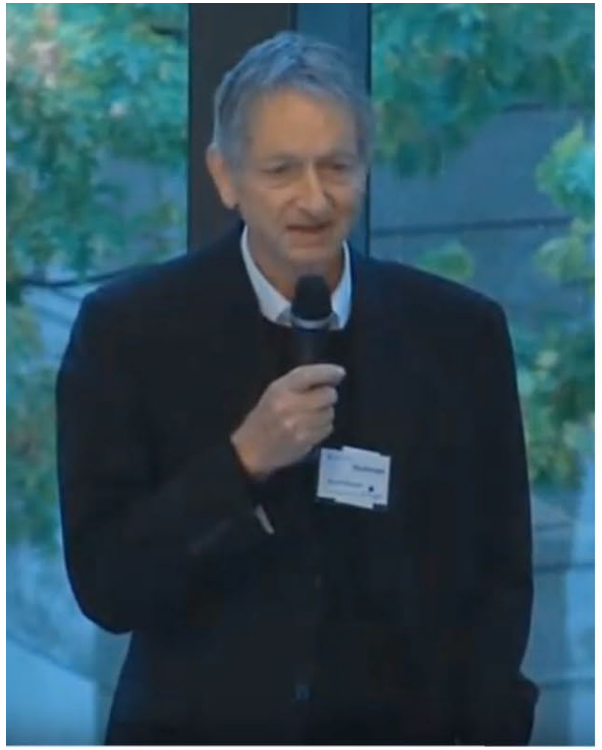
We demonstrate the effectiveness of **nnDetection** on two public benchmarks, ADAM and ...

☆ Save Cite Cited by 132 Related articles All 4 versions





Key Challenges in the Real World



Geoffrey Hinton

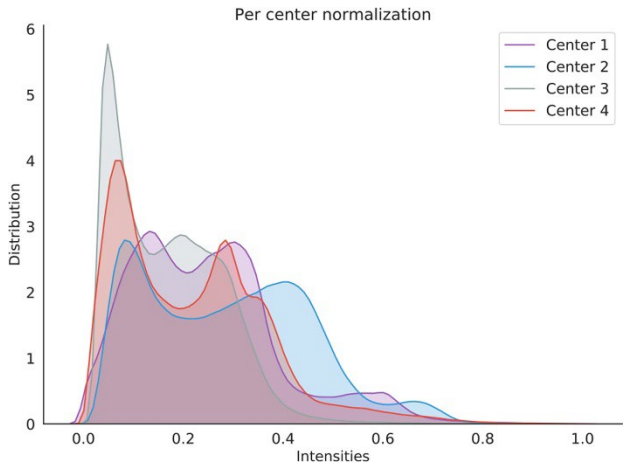
"I think that if you work as a radiologist, you are like Wile E. Coyote in the cartoon. You're already over the edge of the cliff, but you haven't yet looked down. There's no ground underneath. People should stop training radiologists now. It's just completely obvious that in five years deep learning is going to do better than radiologists."

Nov 24, 2016

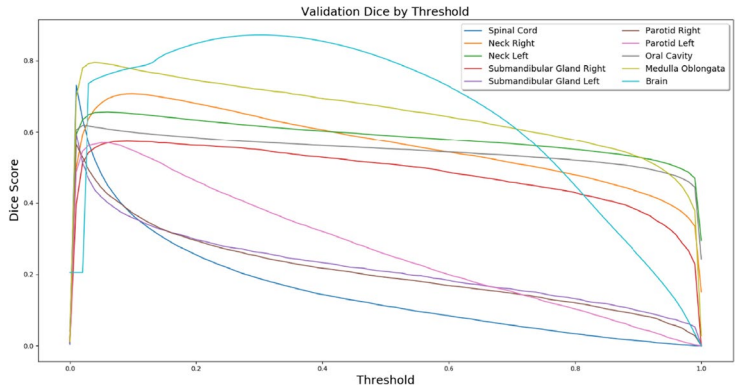
https://youtu.be/2HMPRXstSvQ?si=V_MLXfiG0N8qm67T

Key Challenges in the Real World

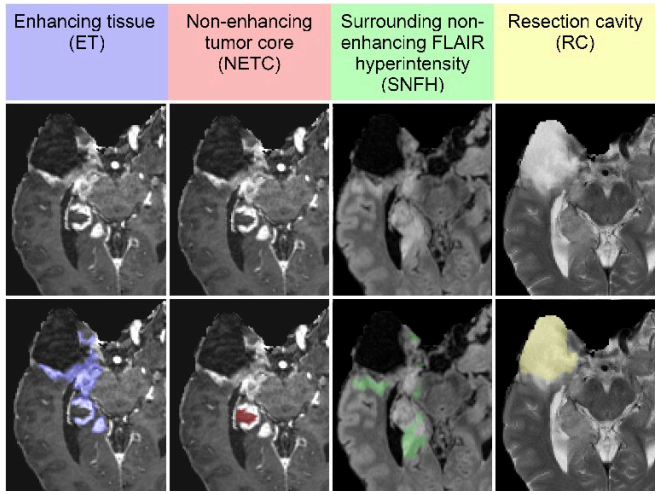
- Technical Challenges
  - Generalization



Cross-center / device / population shift



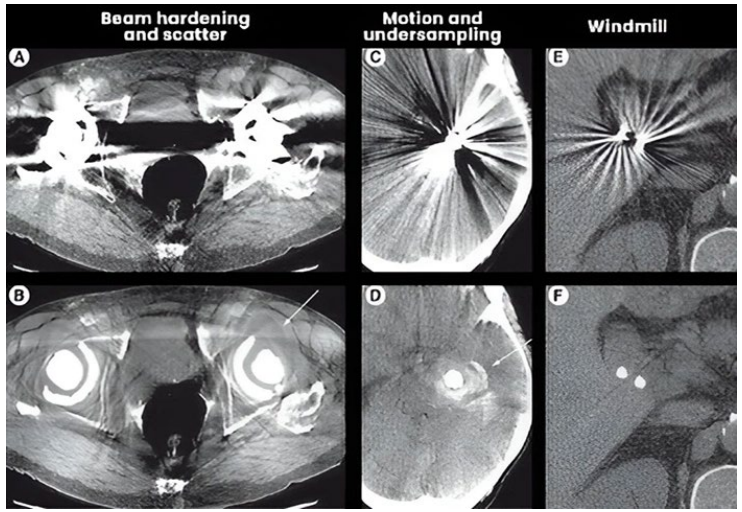
Sensitive to post-processing, decision thresholds, and protocols



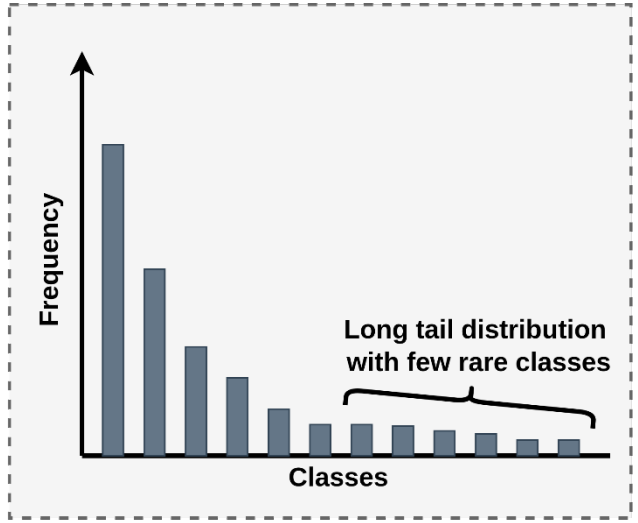
Post-treatment tumors are harder (e.g., BraTS 2024)

Key Challenges in the Real World

- Technical Challenges
  - Generalization
  - Reliability



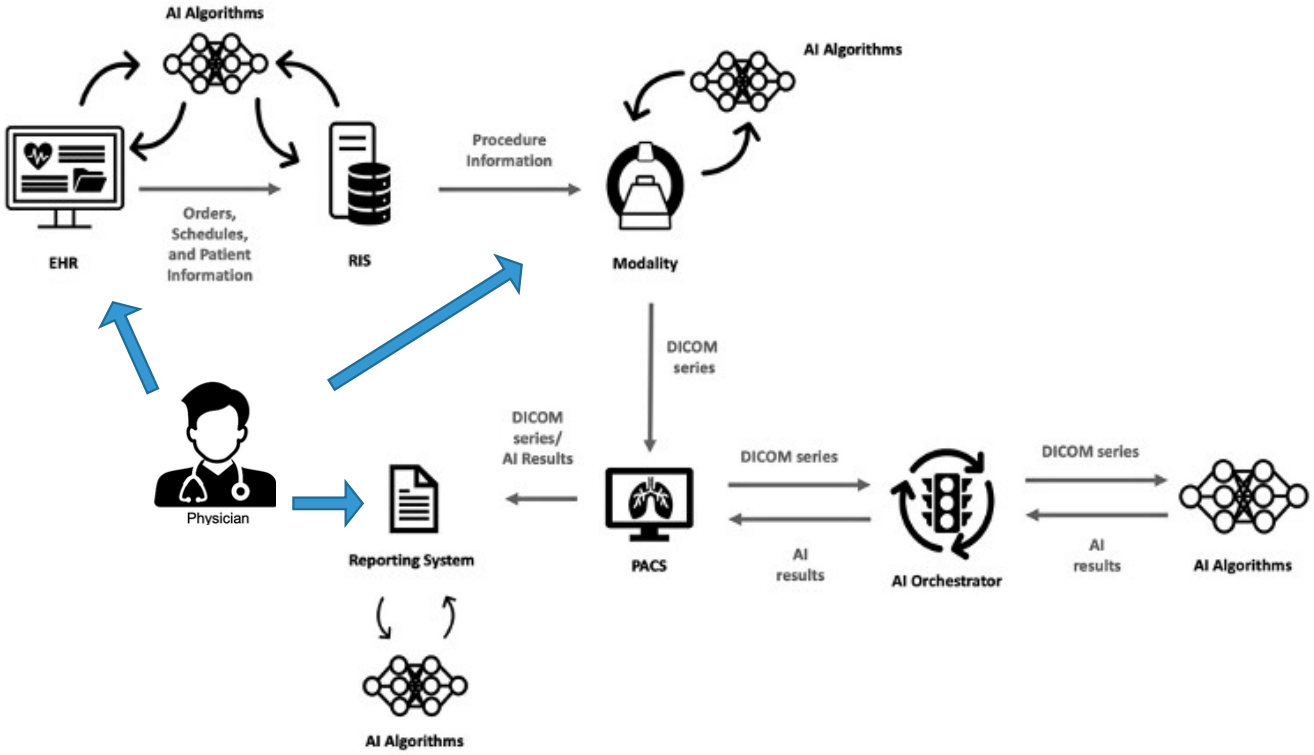
Aleatoric = data noise (blurry borders, artifacts, label disagreement)



Epistemic = model ignorance (rare/OOD cases)

Key Challenges in the Real World

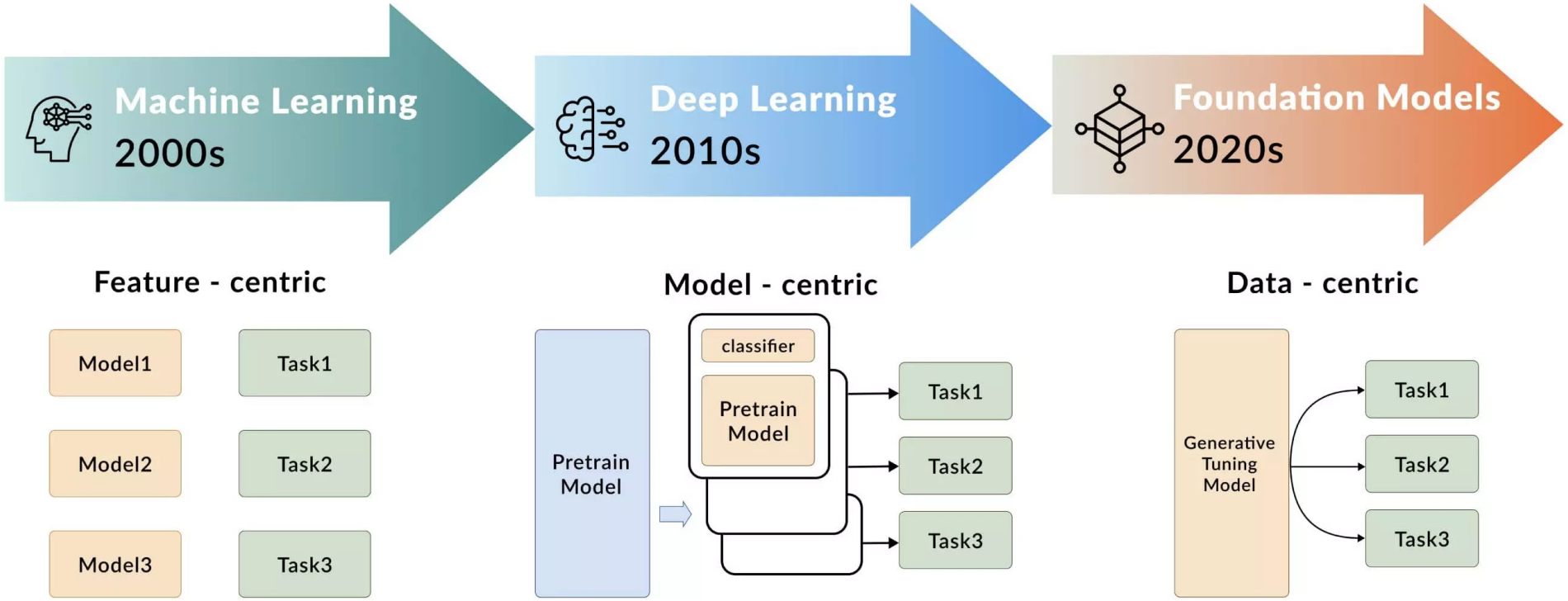
- Technical Challenges
  - Generalization
  - Reliability
  - Workflow & Interaction



Key Challenges in the Real World

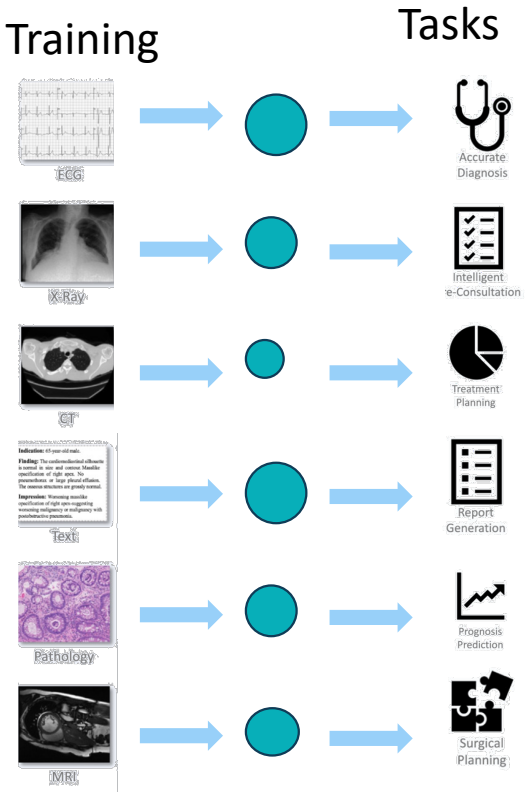
- Technical Challenges
  - Generalization
  - Reliability
  - Workflow & Interaction
- Regulatory Challenges
  - Privacy
  - Policy
  - Financial incentives
  - Equality

Foundation Models = Big Model × Big Data × Self-Supervision

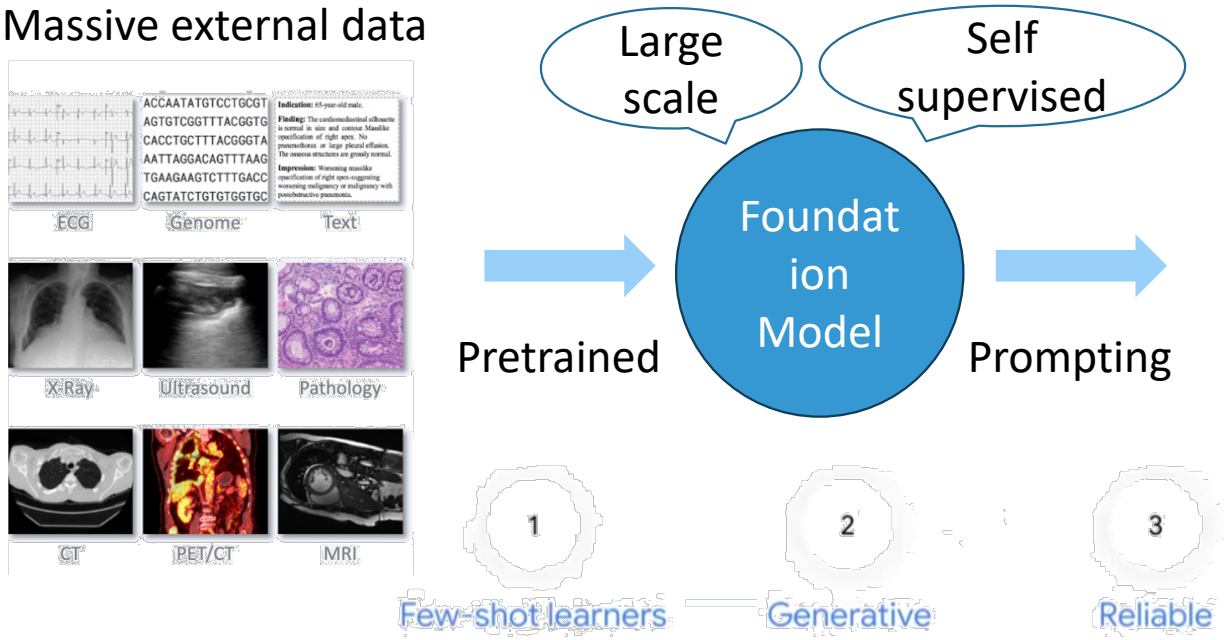


Foundation Models = Big Model × Big Data × Self-Supervision

Traditional AI models



Foundation models



Foundation Models = Big Model × Big Data × Self-Supervision

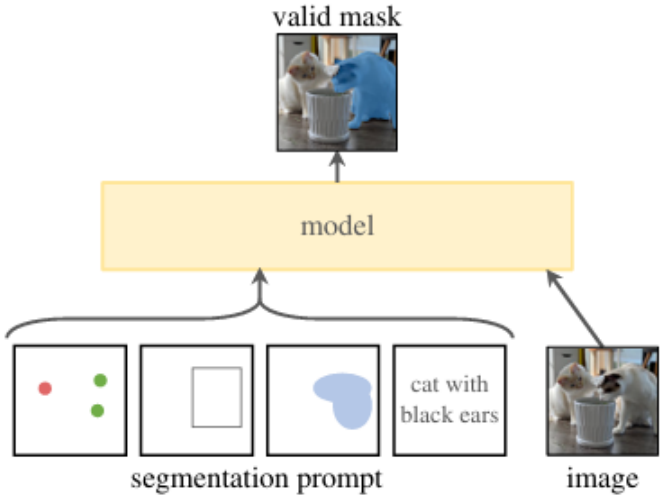
Segment anything

A Kirillov, E Mintun, N Ravi, H Mao... - Proceedings of the ...

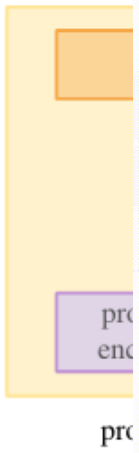
... **Segment Anything** 1B (SA-1B): Figure 1: We aim to build

... components: a promptable **segmentation** task, a **segmen**

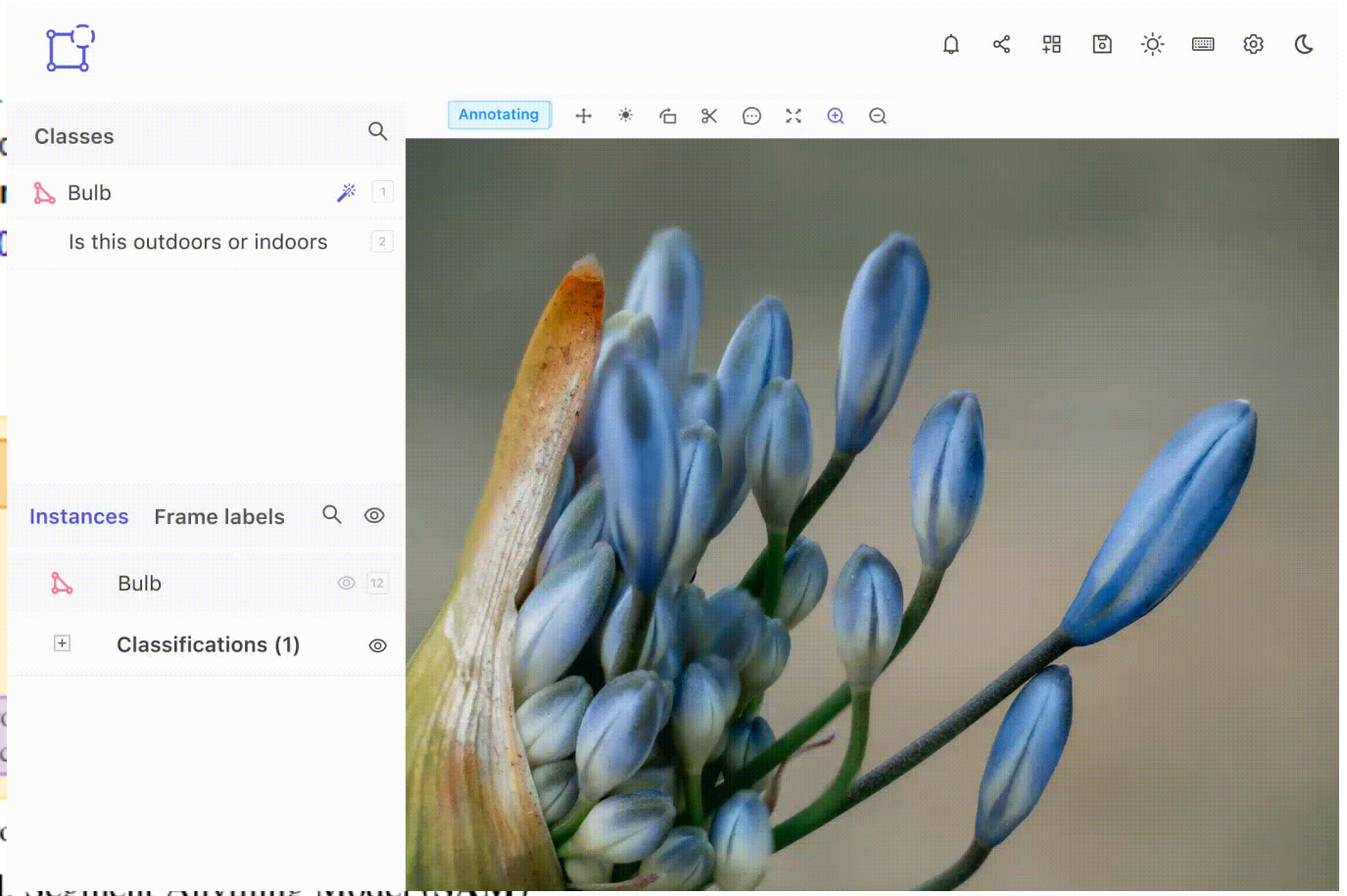
☆ Save 99 Cite Cited by 14166 Related articles All 10



(a) **Task**: promptable segmentation



(b) **Model**. Segment Anything Model (SAM)



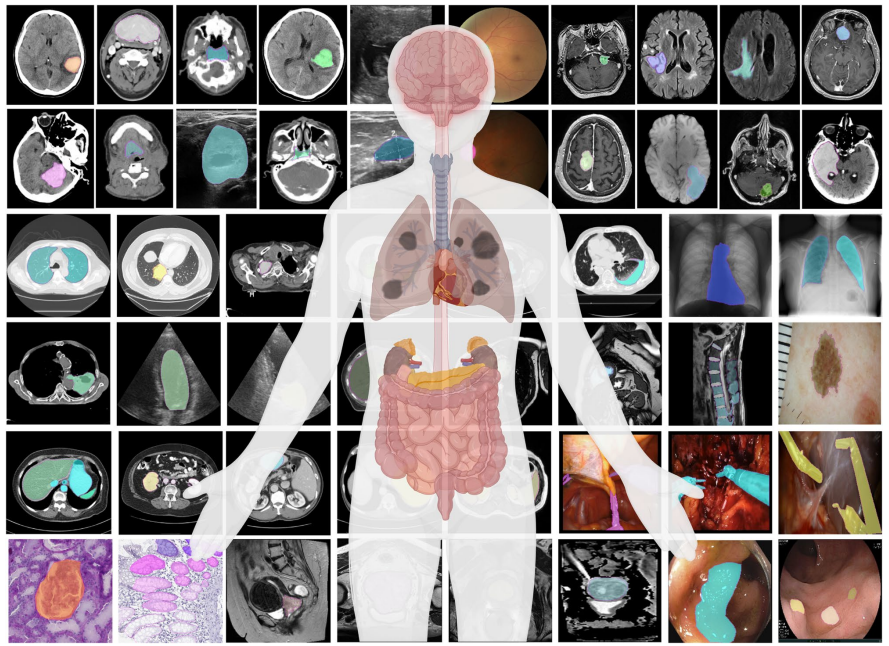
Foundation Models = Big Model × Big Data × Self-Supervision

nature communications

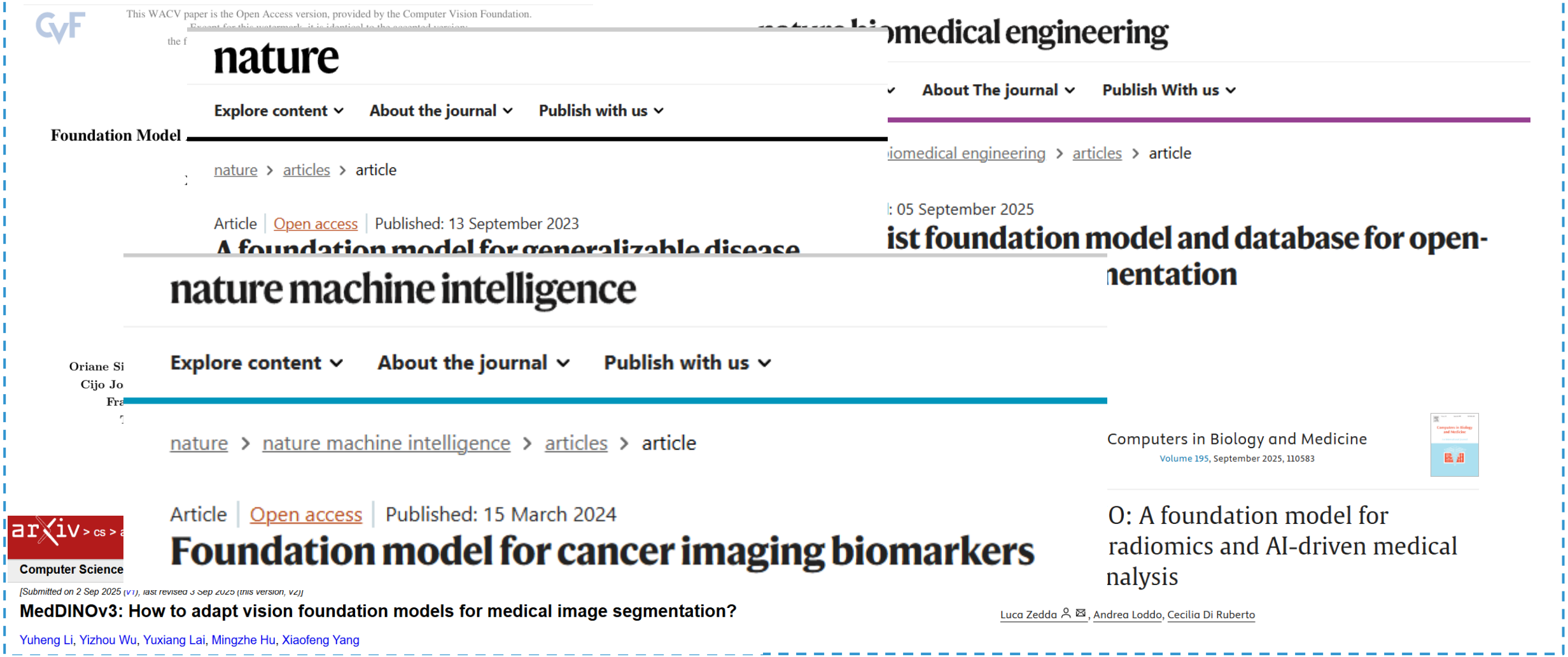
Article

https://doi.org/10.1038/s41467-024-44824-z

Segment anything in medical images



Foundation Models = Big Model × Big Data × Self-Supervision



This WACV paper is the Open Access version, provided by the Computer Vision Foundation.

Foundation Model

nature

Explore content · About the journal · Publish with us

nature > articles > article

Article | Open access | Published: 13 September 2023

A foundation model for generalizable disease

nature machine intelligence

Explore content · About the journal · Publish with us

nature > nature machine intelligence > articles > article

Article | Open access | Published: 15 March 2024

Foundation model for cancer imaging biomarkers

Oriane Si
Cijo Jo
Fra

arXiv > cs > a

Computer Science

[Submitted on 2 Sep 2025 (v1), last revised 3 Sep 2025 (this version, v2)]

MedDINOv3: How to adapt vision foundation models for medical image segmentation?

Yuheng Li, Yizhou Wu, Yuxiang Lai, Mingzhe Hu, Xiaofeng Yang

omedical engineering

About The journal · Publish With us

iomedical engineering > articles > article

: 05 September 2025

ist foundation model and database for open-

nentation

Computers in Biology and Medicine

Volume 195, September 2025, 110583

O: A foundation model for radiomics and AI-driven medical nalysis

Luca Zedda, Andrea Loddo, Cecilia Di Ruberto

Conclusion

- Deep Learning
  - A branch of machine learning using multi-layer neural networks to learn features end-to-end from data for classification, detection, and segmentation.
- Advantages
  - Handles complex, high-dimensional medical images
  - Enables new imaging biomarkers and more personalized care
- Challenges
  - Generalization, Reliability, Workflow & Interaction
- Future directions
  - Foundation models, Uncertainty quantification, Clinical integration
- The real breakthrough is making AI a *partner* to clinicians—calibrated, interpretable, robust, and woven into the workflow—not a replacement.

The team

Acknowledgments

French government grant by the National Research Agency integrated into the France 2030 # ANR-21-RHUS-0005.