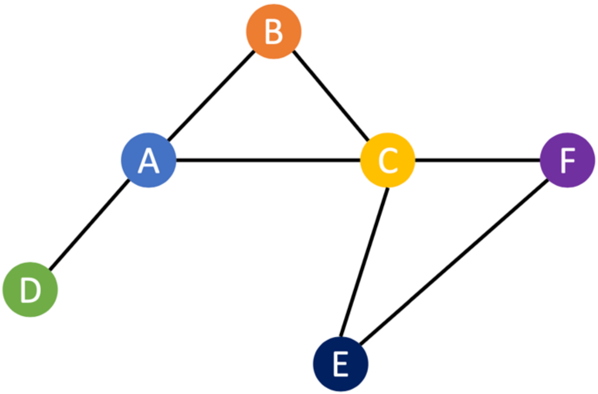
# CSC413/2516 Lecture 11:
# Additional architectures: GNNs, UNet, MedSAM

Bo Wang

# The missing piece

- Tabular data : Linear Models, MLP
- Sequence data (e.g., Language, speech): CNN, RNN, Transformer
- Imaging data : CNN, Vision Transformer
- What about ***graph*** data?

# What is a graph?




A graph is composed of

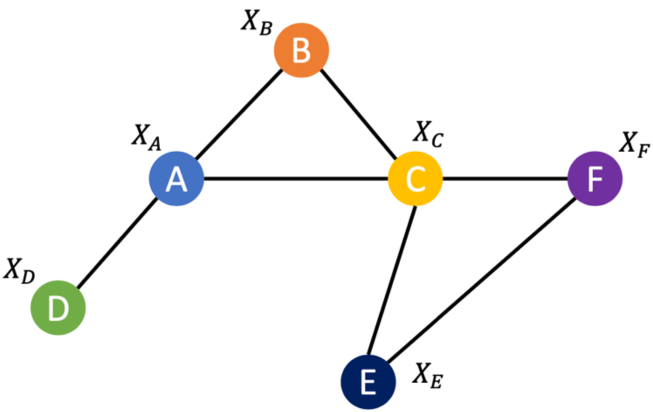
- **Nodes** (also called vertices)
- **Edges** connecting a pair of nodes

presented in an **adjacency matrix**

| | A | B | C | D | E | F |
|---|---|---|---|---|---|---|
| A | | 1 | 1 | 1 | | |
| B | 1 | | 1 | | | |
| C | 1 | 1 | | | 1 | 1 |
| D | 1 | | | | | |
| E | | | 1 | | | 1 |
| F | | | 1 | | 1 | |

Source: Minji Yoon, CMU

# What is a graph?




A graph is composed of

- **Nodes** (also called vertices)
- **Edges** connecting a pair of nodes

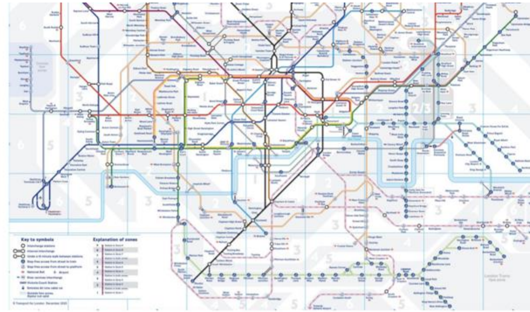
presented in an **adjacency matrix**

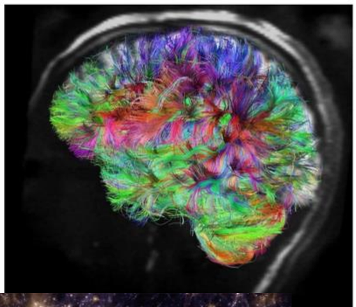
Nodes can have **feature vectors**

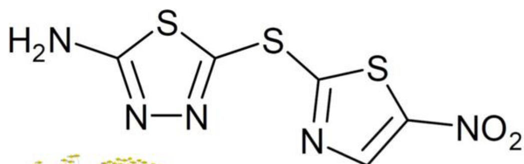
| | |
|---|---|
| A | $X_A$ |
| B | $X_B$ |
| C | $X_C$ |
| D | $X_D$ |
| E | $X_E$ |
| F | $X_F$ |

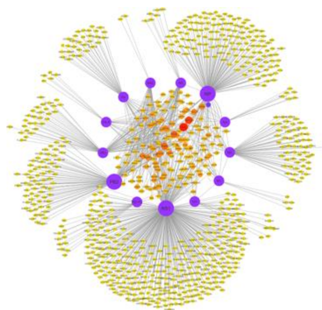
Source: Minji Yoon, CMU

# Graph is everywhere!

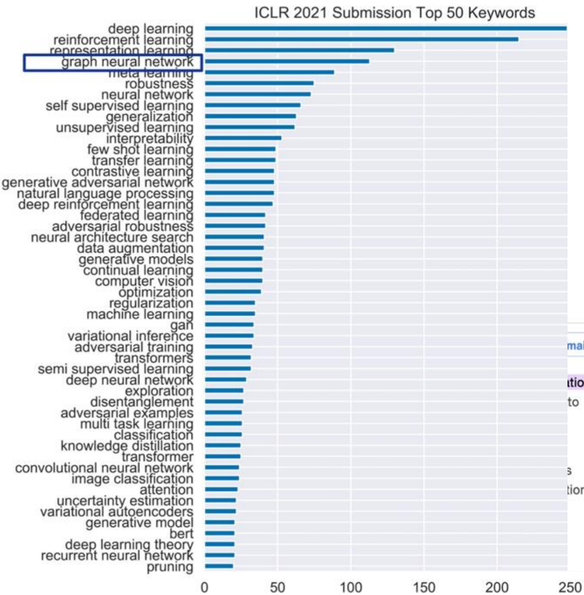






# Graph Neural Networks (GNN) is everywhere




npj | computational materials

**Explore content** ⌄ **About the journal** ⌄ **Publish with us** ⌄

nature > npj computational materials > articles > article

Article | Open Access | Published: 03 June 2021

## Benchmarking graph neural networks for materials chemistry

Victor Fung, Jiaxin Zhang, Eric Juarez & Bobby G. Sumpter

*npj Computational Materials* **7**, Article number: 84 (2021) | Cite this article

**7807** Accesses | **7** Citations | **41** Altmetric | Metrics

**nature** View all journals Search Login

**Explore content** ⌄ **About the journal** ⌄ **Publish with us** ⌄

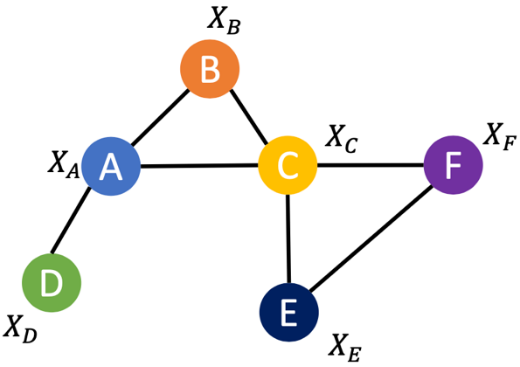
nature > articles > article

Article | Published: 09 June 2021

## A graph placement methodology for fast chip design

Azalia Mirhoseini, Anna Goldie, Mustafa Yazgan, Joe Wenjie Jiang, Ebrahim Songhori, Shen Wang, Young-Joon Lee, Eric Johnson, Omkar Pathak, Azade Nazi, Jiwoo Pak, Andy Tong, Kavya Srinivasa, William Hang, Emre Tuncer, Quoc V. Le, James Laudon, Richard Ho, Roger Carpenter & Jeff Dean

# What is GNN? – Problem Setup

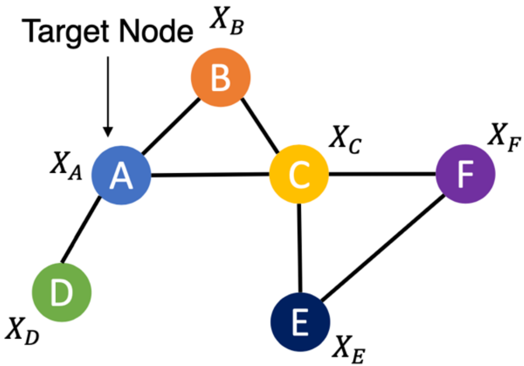



- **Given**
  - A graph
  - Node attributes
  - (part of nodes are labeled)
- **Find**
  - Node embeddings
- **Predict**
  - Labels for the remaining nodes

undirected unweighted graph

Source: Minji Yoon, CMU

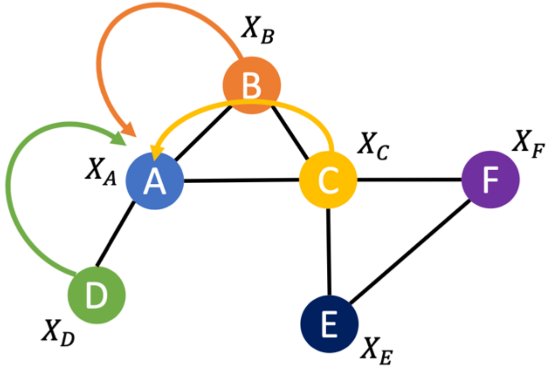
# What is GNN? – Problem Setup




**“Homophily: connected nodes are related/informative/similar”**

Source: Minji Yoon, CMU

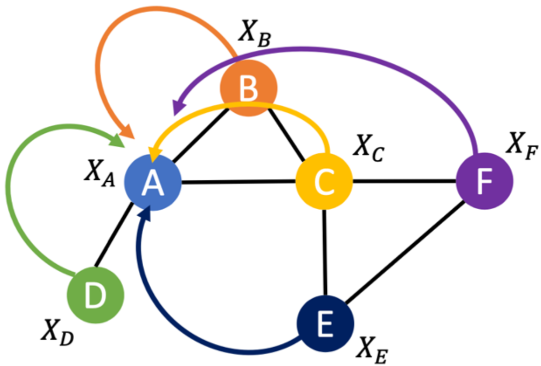
# What is GNN? – Problem Setup




**“Homophily: connected nodes are related/informative/similar”**

Source: Minji Yoon, CMU

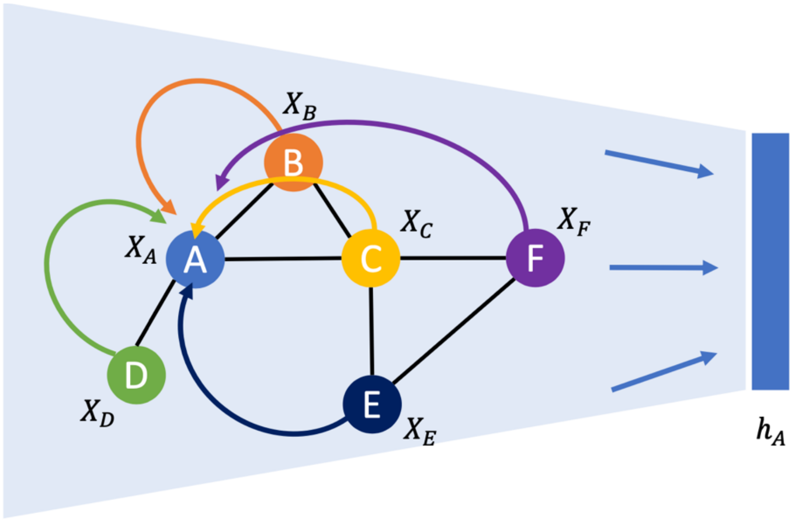
# What is GNN? – Problem Setup




**“Homophily: connected nodes are related/informative/similar”**

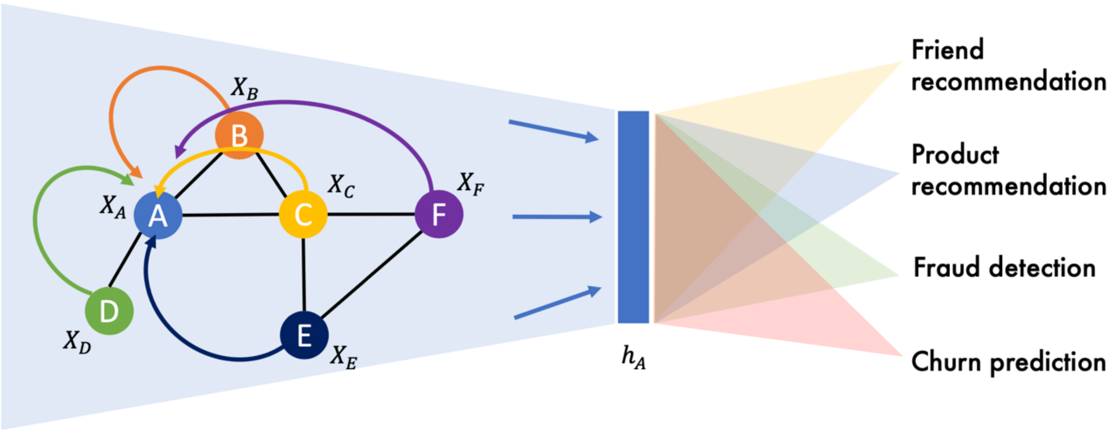
Source: Minji Yoon, CMU

# What is GNN? – Problem Setup



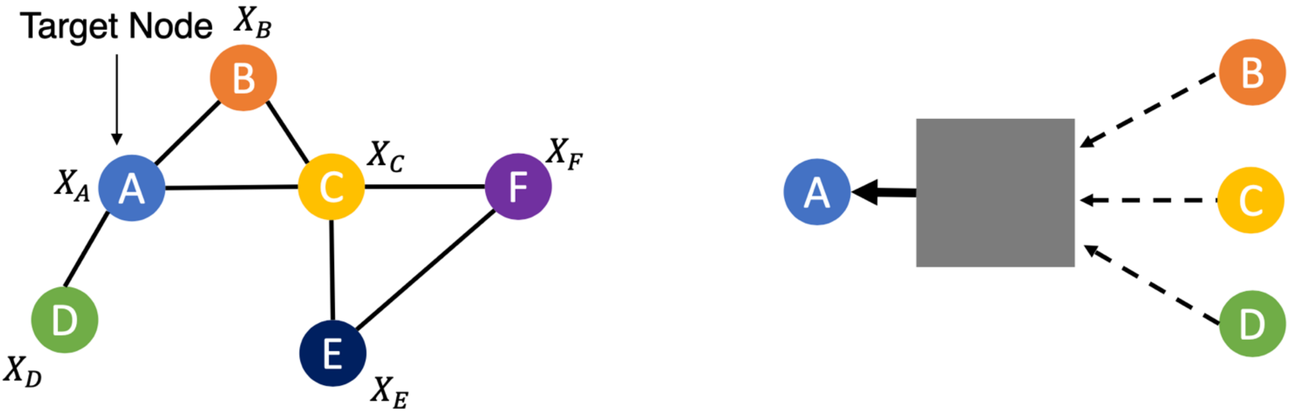

Source: Minji Yoon, CMU

# What is GNN? – Problem Setup



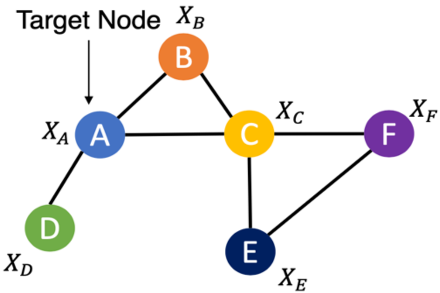

Source: Minji Yoon, CMU

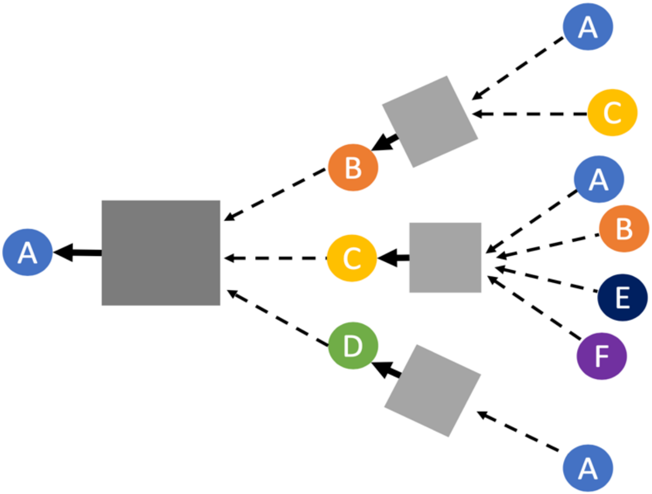
# What is GNN? – Forward propagation




Source: Minji Yoon, CMU

# What is GNN? – Forward propagation



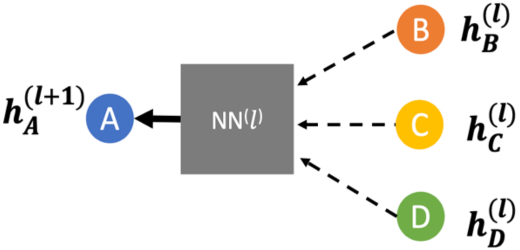

Source: Minji Yoon, CMU

# What is GNN? – Forward propagation

**1. Aggregate messages from neighbors**

$h_v^{(l)}$: node embedding of $v$ at $l$-th layer
$\mathcal{N}(v)$ : neighboring nodes of $v$
$\boldsymbol{f}^{(l)}$: aggregation function at $l$-th layer
$m_v^{(l)}$ : message vector of $v$ at $l$-th layer

$$
\begin{aligned}
m_A^{(l)} &= \boldsymbol{f}^{(l)}\left(h_A^{(l)}, \left\{h_u^{(l)} : u \in \mathcal{N}(A)\right\}\right) \\
&= \boldsymbol{f}^{(l)}\left(h_A^{(l)}, h_B^{(l)} h_C^{(l)} h_D^{(l)}\right)
\end{aligned}
$$




Neighbors of node A
$\mathcal{N}(A) = \{B, C, D\}$

Source: Minji Yoon, CMU

# What is GNN? – Forward propagation

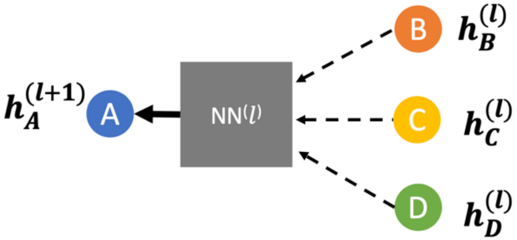
1. **Aggregate messages from neighbors**

$$m_A^{(l)} = \boldsymbol{f}^{(l)}\left(h_A^{(l)}, \left\{h_u^{(l)} : u \in \mathcal{N}(A)\right\}\right)$$
$$= \boldsymbol{f}^{(l)}\left(h_A^{(l)}, h_B^{(l)} h_C^{(l)} h_D^{(l)}\right)$$

2. **Transform** messages

$\boldsymbol{g}^{(l)}$: transformation function at $l$-th layer

$$h_A^{(l+1)} = \boldsymbol{g}^{(l)}(m_A^{(l)})$$




Neighbors of node A
$\mathcal{N}(A) = \{B, C, D\}$

Source: Minji Yoon, CMU

# What is GNN? – Forward propagation

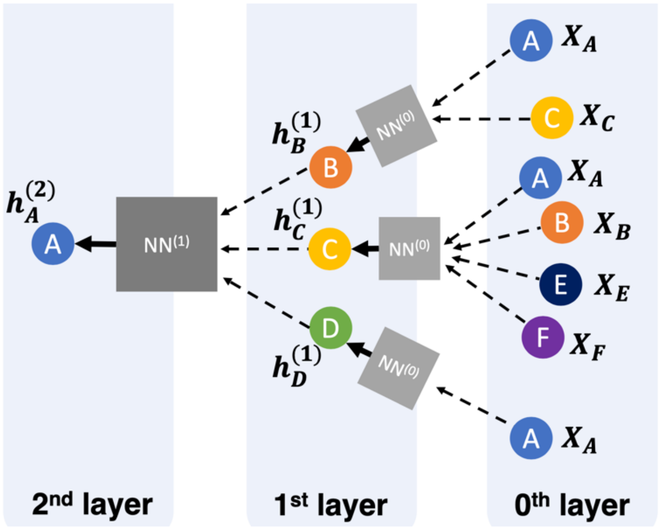
In each layer $l$ ,
for each target node $v$ :

**1. Aggregate messages**

$$m_v^{(l)} = \boldsymbol{f}^{(l)}\left(h_v^{(l)}, \left\{h_u^{(l)} : u \in \mathcal{N}(v)\right\}\right)$$

**2. Transform messages**

$$h_v^{(l+1)} = \boldsymbol{g}^{(l)}(m_v^{(l)})$$




Source: Minji Yoon, CMU

# What is GNN? – Forward propagation

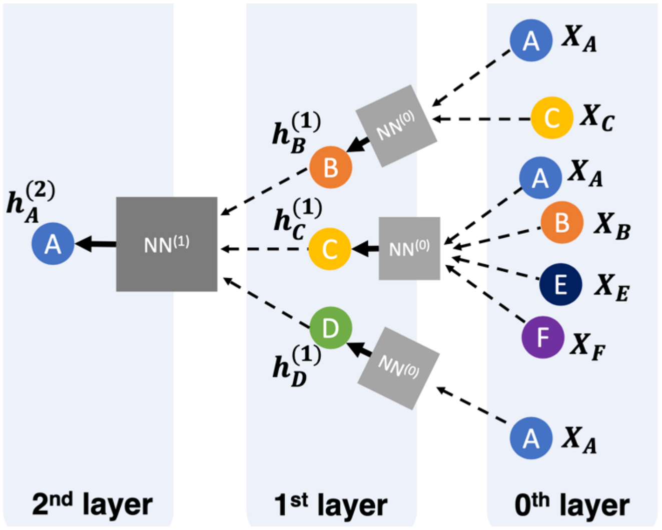
Graph Convolutional Networks[1]

**1. Aggregate messages**

$$m_v^{(l)} = \frac{1}{|\mathcal{N}(v) + 1|} \sum_{u \in \mathcal{N}(v) \cup \{v\}} h_u^{(l)}$$

**2. Transform messages**

$$h_v^{(l+1)} = \sigma(\boldsymbol{W}^{(l)} \circ m_v^{(l)})$$




[1] Kipf, Thomas N., et al. "Semi-supervised classification with graph convolutional networks."

Source: Minji Yoon, CMU

## What is GNN? – Forward propagation

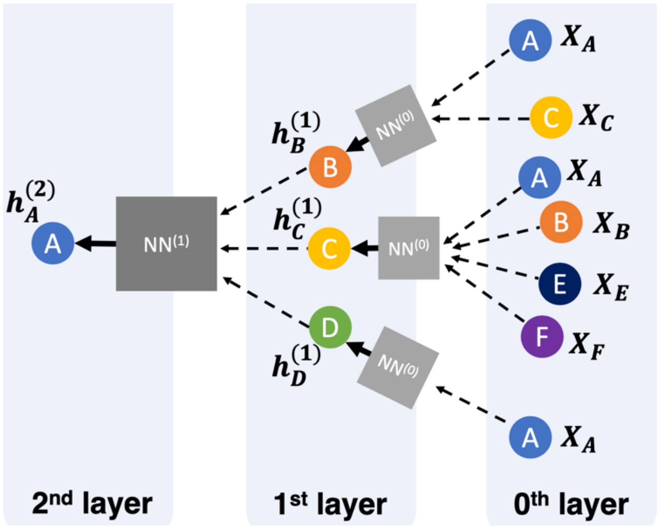
Graph Isomorphism Networks[2]

**1. Aggregate messages**

$$m_v^{(l)} = \sum_{u \in \mathcal{N}(v) \cup \{v\}} h_u^{(l)}$$

**2. Transform messages**

$$h_v^{(l+1)} = \sigma(\boldsymbol{W}^{(l)} \circ m_v^{(l)})$$




[2] Xu, Keyulu, et al. "How powerful are graph neural networks?."

Source: Minji Yoon, CMU

# What is GNN? – Forward propagation

Simplified GCN[3]

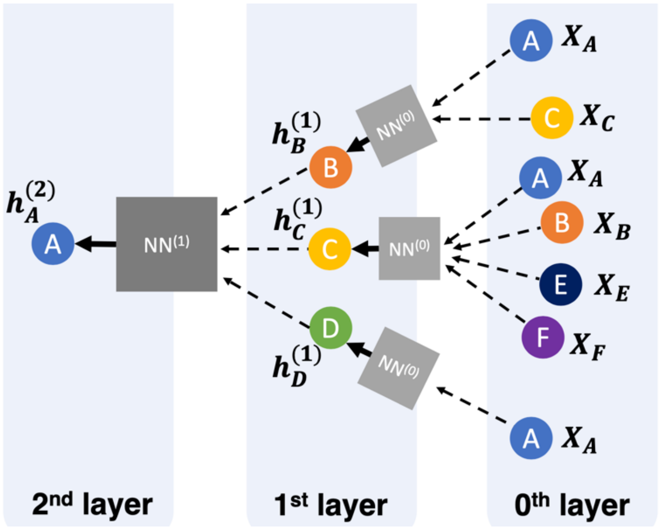
**1. Aggregate messages**

$$m_v^{(l)} = \frac{1}{|\mathcal{N}(v)+1|} \sum_{u \in \mathcal{N}(v) \cup \{v\}} h_u^{(l)}$$

**2. Transform messages**

$$h_v^{(l+1)} = \boldsymbol{W}^{(l)} \circ m_v^{(l)}$$

[3] Wu, Felix, et al. "Simplifying graph convolutional networks."




Source: Minji Yoon, CMU

# What is GNN? – Forward propagation

In each layer $l$ :

**Aggregate** over neighbors

$$m_v^{(l-1)} = \boxed{\boldsymbol{f}^{(l)}}\left(h_v^{(l-1)}, \left\{h_u^{(l-1)} : u \in \mathcal{N}(v)\right\}\right)$$

Core part of GNNs

**Transform** messages

$$h_v^{(l)} = \boxed{\boldsymbol{g}^{(l)}}\left(m_v^{(l-1)}\right)$$

1-layer MLP is commonly used

Source: Minji Yoon, CMU

# Graph Convolutional Network (GCN)

- GCN[1]
  - Average embeddings of neighboring nodes

$$m_v^{(l)} = \frac{1}{|\mathcal{N}(v) + 1|} \sum_{u \in \mathcal{N}(v) \cup \{v\}} h_u^{(l)}$$

$$+$$

$$h_v^{(l+1)} = \sigma(\boldsymbol{W}^{(l)} \circ m_v^{(l)})$$

$$h_v^{(l)} = \sigma(\boldsymbol{W}^{(l)} \circ (\frac{1}{|\mathcal{N}(v)+1|} \Sigma_{u \in \mathcal{N}(v) \cup \{v\}} h_u^{(l-1)}))$$

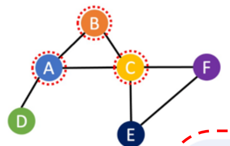
[1] Kipf, Thomas N., et al. "Semi-supervised classification with graph convolutional networks."

# Graph Convolutional Network (GCN)

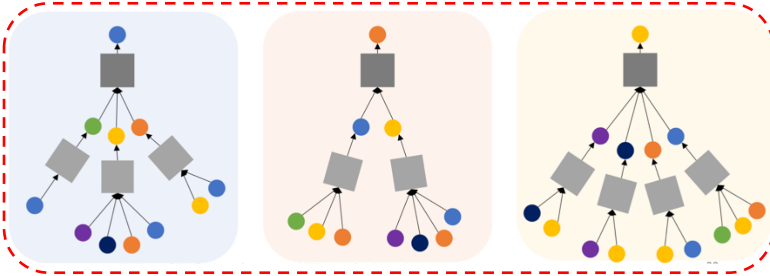
- GCN[1]

Can we use batch-mode?

$$h_v^{(l)} = \sigma(\boldsymbol{W}^{(l)} \circ (\frac{1}{|\mathcal{N}(v)+1|} \sum_{u \in \mathcal{N}(v) \cup \{v\}} h_u^{(l-1)}))$$




B= 3

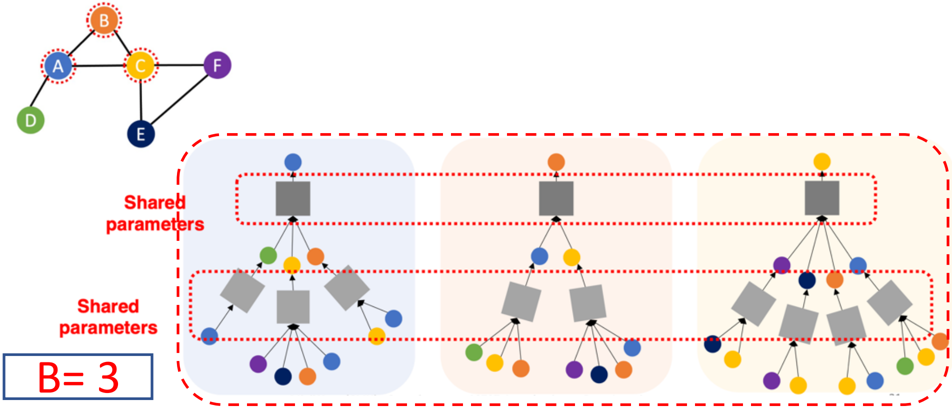


[1] Kipf, Thomas N., et al. "Semi-supervised classification with graph convolutional networks."

# Graph Convolutional Network (GCN)

- GCN[1]

$$h_v^{(l)} = \sigma(\boldsymbol{W}^{(l)} \circ (\frac{1}{|\mathcal{N}(v)+1|} \sum_{u \in \mathcal{N}(v) \cup \{v\}} h_u^{(l-1)}))$$




[1] Kipf, Thomas N., et al. "Semi-supervised classification with graph convolutional networks."

# Graph Convolutional Network (GCN)

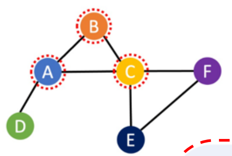
- GCN[1]

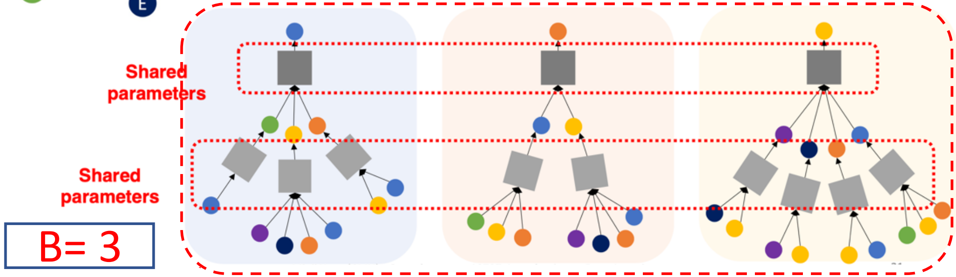
$$h_v^{(l)} = \sigma(\boldsymbol{W}^{(l)} \circ (\frac{1}{|\mathcal{N}(v)+1|}\sum\nolimits_{u\,\in\,\mathcal{N}(v)\cup\{v\}} h_u^{(l-1)}))$$

$$\boxed{\boldsymbol{H}^{(l)}} = \sigma(\boxed{(\widetilde{\boldsymbol{A}+\boldsymbol{I}})}\boldsymbol{H}^{(l-1)}\boldsymbol{W}^{(l)})$$

**Node embedding matrix**

(row-normalized) **Adjacency matrix**



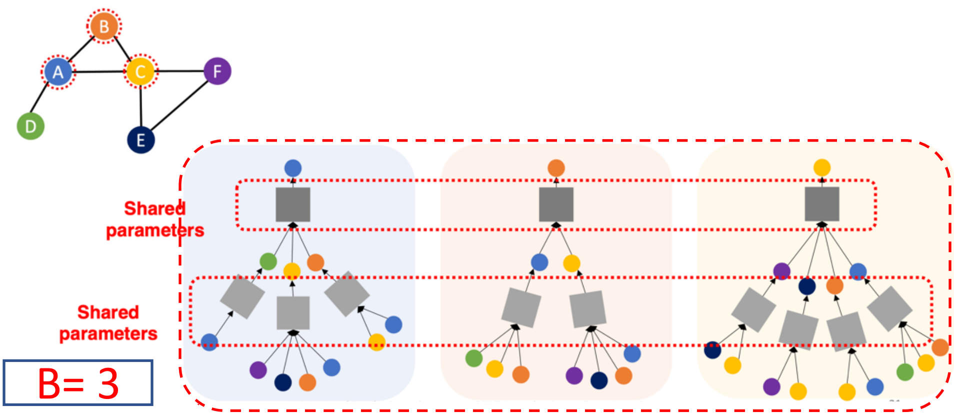

[1] Kipf, Thomas N., et al. "Semi-supervised classification with graph convolutional networks."

# Graph Convolutional Network (GCN)

- GCN[1]

$$h_v^{(l)} = \sigma(\boldsymbol{W}^{(l)} \circ (\frac{1}{|\mathcal{N}(v)+1|}\sum\nolimits_{u \in \mathcal{N}(v)\cup\{v\}} h_u^{(l-1)}))$$

$$\boldsymbol{H}^{(l)} = \sigma(\underset{\textbf{Fixed}}{\boxed{\widetilde{(\boldsymbol{A}+\boldsymbol{I})}}}\boldsymbol{H}^{(l-1)}\underset{\textbf{Trainable}}{\boxed{\boldsymbol{W}^{(l)}}})$$




[1] Kipf, Thomas N., et al. "Semi-supervised classification with graph convolutional networks."

# Graph Convolutional Network (GCN)

$$h_v^{(l)} = \sigma(\boldsymbol{W}^{(l)} \circ (\frac{1}{|\mathcal{N}(v)+1|}\sum_{u \in \mathcal{N}(v)\cup\{v\}} h_u^{(l-1)}))$$

$$\boldsymbol{H}^{(l)} = \sigma(\widetilde{(\boldsymbol{A}+\boldsymbol{I})}\boldsymbol{H}^{(l-1)}\boldsymbol{W}^{(l)})$$

**Node embedding matrix**

(row-normalized) **Adjacency matrix**

Assume A is the affinity matrix

Let $D$ be diagonal matrix where

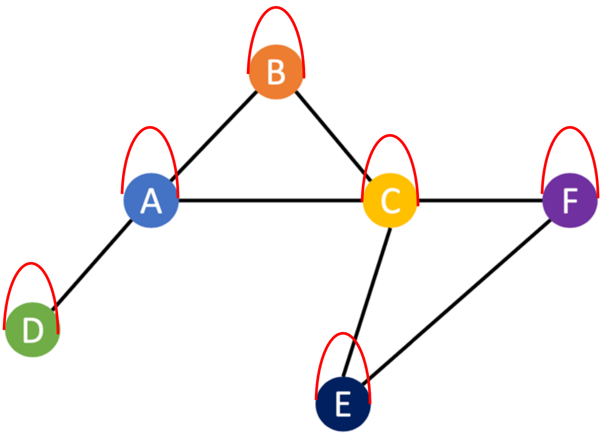
$D_{v,v} = \mathrm{Deg}(v) = |N(v)|$

- The inverse of $D$: $D^{-1}$ is also diagonal: $D_{v,v}^{-1} = 1/|N(v)|$

$$\boxed{(D+I)^{-1} * (A+I)} * H^{(l-1)}$$

# Graph Convolutional Network (GCN)

$$\widetilde{A+I} = (D+I)^{-1} * (A+I) \longrightarrow \tilde{A} = D^{-1} * A$$




A graph is composed of

- **Nodes** (also called vertices)
- **Edges** connecting a pair of nodes

presented in an **adjacency matrix**

| | A | B | C | D | E | F |
|---|---|---|---|---|---|---|
| A | 1 | 1 | 1 | 1 | | |
| B | 1 | 1 | 1 | | | |
| C | 1 | 1 | 1 | | 1 | 1 |
| D | 1 | | | 1 | | |
| E | | | 1 | | 1 | 1 |
| F | | | 1 | | 1 | 1 |

$$\boxed{\mathcal{L} = \mathbf{I} - (\mathbf{D})^{-1} * (\mathbf{A})}$$

Normalized Graph Laplacian

# Graph Convolutional Network (GCN) -- Summary

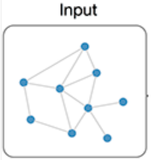
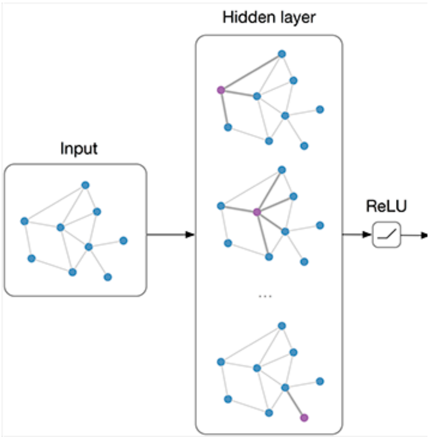



Image Credit: Defferrard et al. NIPS 2016

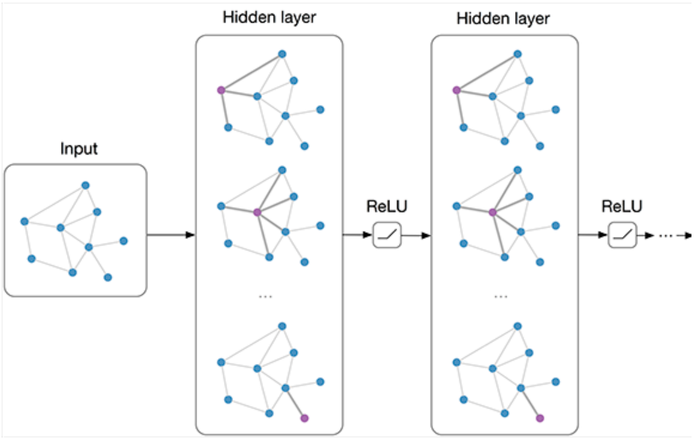
# Graph Convolutional Network (GCN) -- Summary




$$\mathrm{ReLU}\left(\hat{A}XW^{(0)}\right)$$

Image Credit: Defferrard et al. NIPS 2016

# Graph Convolutional Network (GCN) -- Summary



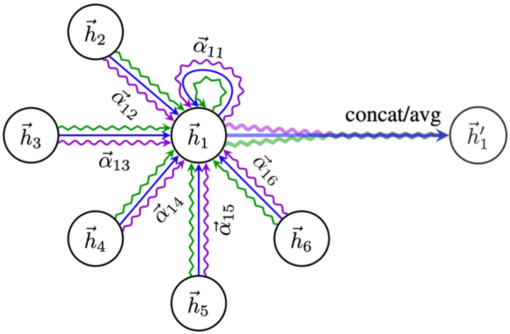

$$\hat{A}\,\mathrm{ReLU}\left(\hat{A} X W^{(0)}\right) W^{(1)}$$

Image Credit: Defferrard et al. NIPS 2016

# Graph Attention Network (optional)

- GAT[14]
  - Different weights to different nodes in a neighborhood
  - Multi-head attention

$$\alpha_{ij} = \frac{\exp\left(\text{LeakyReLU}\left(\vec{\mathbf{a}}^T[\mathbf{W}\vec{h}_i \| \mathbf{W}\vec{h}_j]\right)\right)}{\sum_{k \in \mathcal{N}_i} \exp\left(\text{LeakyReLU}\left(\vec{\mathbf{a}}^T[\mathbf{W}\vec{h}_i \| \mathbf{W}\vec{h}_k]\right)\right)}$$




[14] Petar Veličković., et al. "GRAPH ATTENTION NETWORKS."

## How to Train GNN?

- Semi-supervised learning
  - Input node features are given for all nodes in a graph
  - Only a subset of nodes have labels

$$\min_{\Theta} \mathcal{L}(\boldsymbol{y}, f(\boldsymbol{z}_v))$$

$\boldsymbol{y}$: node label

$\mathcal{L}$ could be L2 if $\boldsymbol{y}$ is real number, or cross entropy if $\boldsymbol{y}$ is categorical

Node embedding $\boldsymbol{z}_v$ is a function of input graph

## How to Train GNN?

### Unsupervised setting:

- No node label available
- **Use the graph structure as the supervision!**

- **"Similar" nodes have similar embeddings**

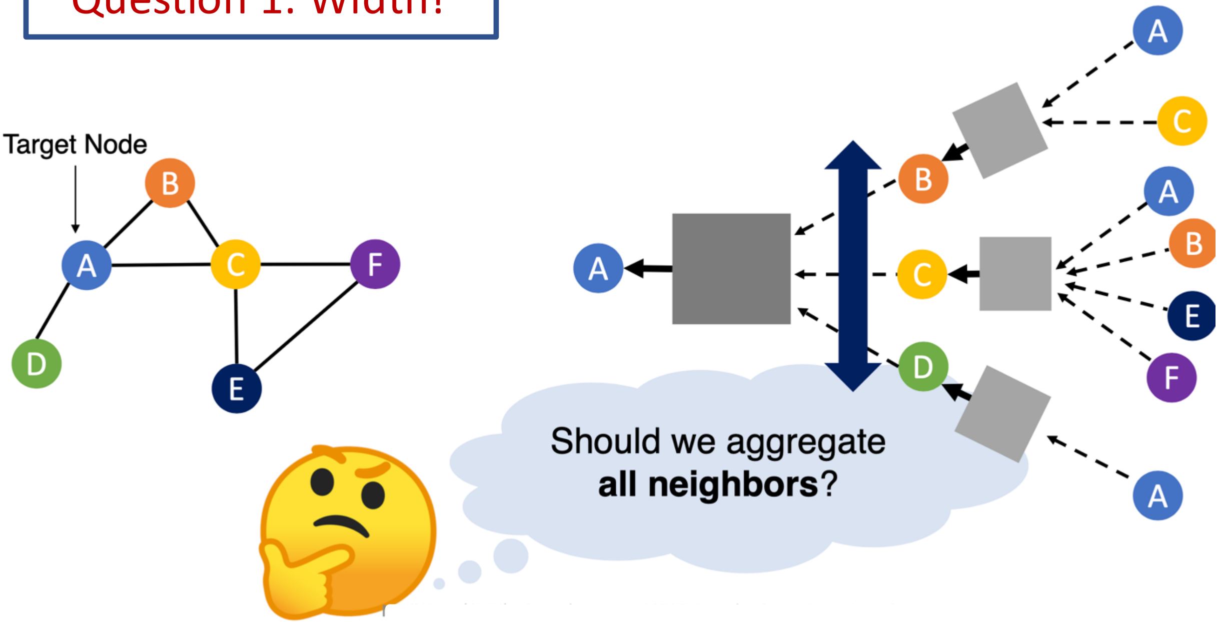
$$\mathcal{L} = \sum_{z_u, z_v} \mathrm{CE}(y_{u,v}, \mathrm{DEC}(z_u, z_v))$$

  - Where $y_{u,v} = 1$ when node $u$ and $v$ are **similar**
  - $\mathrm{CE}$ is the cross entropy
  - $\mathrm{DEC}$ is the decoder such as inner product

Still an active research topic!

# Two interesting questions about GNN

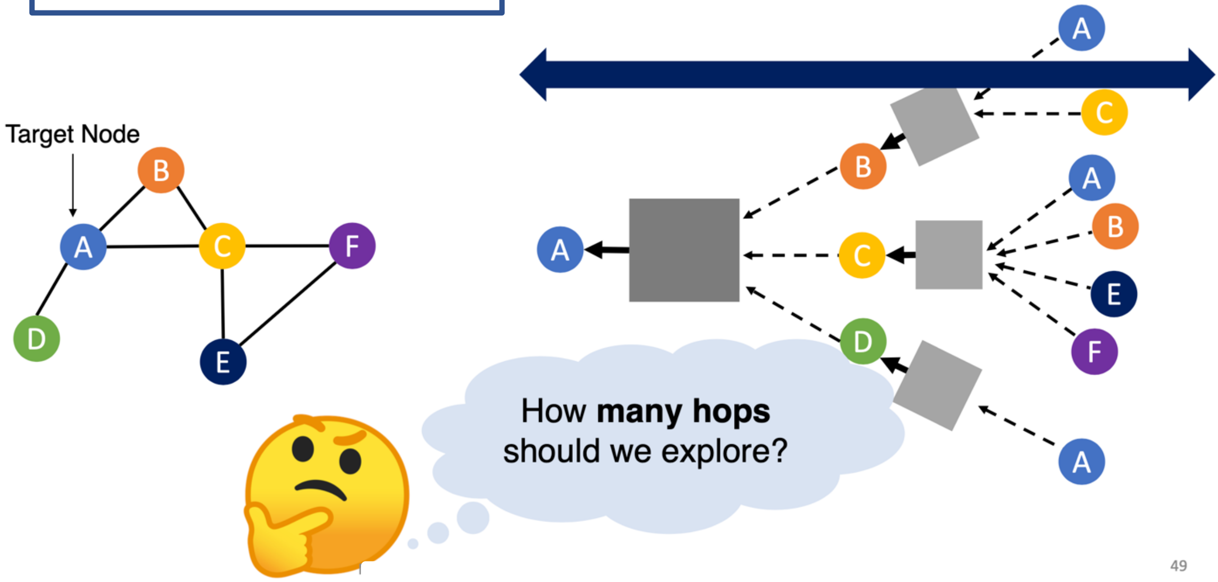
## Question 1: Width?




Source: Minji Yoon, CMU

# Two interesting questions about GNN

## Question 2: Depth?



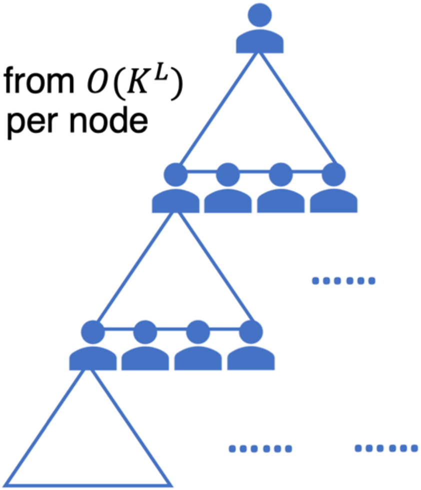



Source: Minji Yoon, CMU

# Two interesting questions about GNN

Question 1: Width?

- If we aggregate all neighbors, GNNs have scalability issues
- Neighbor explosion
  - In $L$ -layer GNNs, one node aggregates information from $O(K^L)$ nodes where $K$ is the average number of neighbors per node



Source: Minji Yoon, CMU

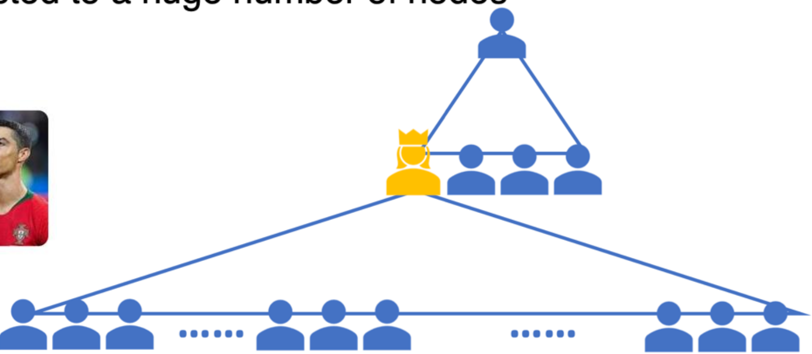
# Two interesting questions about GNN

Question 1: Width?

- If we aggregate all neighbors, GNNs have scalability issues
- Neighbor explosion
  - Hub nodes who are connected to a huge number of nodes

Cristiano Ronaldo

**Cristiano Ronaldo** is currently the most-followed individual on Facebook, with over 150 million followers.



Source: Minji Yoon, CMU

# Two interesting questions about GNN

Question 1: Width?

- Limit the neighborhood expansion by **sampling** a fixed number of neighbors


Source: Minji Yoon, CMU

# Two interesting questions about GNN

## Question 1: Width?

- Random sampling
  - Assign **same** sampling probabilities to all neighbors
  - *GraphSage*[4]
- Importance sampling
  - Assign **different** sampling probabilities to all neighbors
  - *FastGCN*[5], *LADIES*[6], *AS-GCN*[7], *GCN-BS*[8], *PASS*[9]

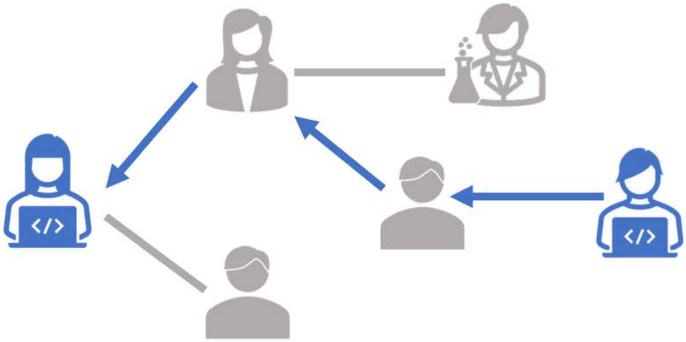
[4] Will Hamilton, et al. "Inductive representation learning on large graphs"
[5] Jie Chen, et al. "Fastgcn: fast learning with graph convolutional networks via importance sampling"
[6] Difan Zou, et al. "Layer-Dependent Importance Sampling for Training Deep and Large Graph Convolutional Networks"
[7] Wenbing Huang, et al. "Adaptive sampling towards fast graph representation learning"
[8] Ziqi Liu, et al. "Bandit Samplers for Training Graph Neural Networks"
[9] Minji Yoon, et al. "Performance-Adaptive Sampling Strategy Towards Fast and Accurate Graph Neural Networks"

Source: Minji Yoon, CMU

## Two interesting questions about GNN

**Question 2: Depth?**

- Informative neighbors could be indirectly connected with a target node



Source: Minji Yoon, CMU

# Two interesting questions about GNN

Question 2: Depth?

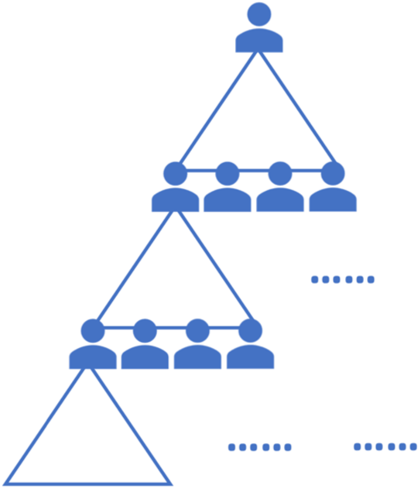
- 2-layer or 3-layer GNNs are commonly used in real worlds


Source: Minji Yoon, CMU

# Two interesting questions about GNN

Question 2: Depth?

- When we increase the depth $L$ more than this, GNNs face neighbor explosion $O(K^L)$
  - **Over-smoothing**



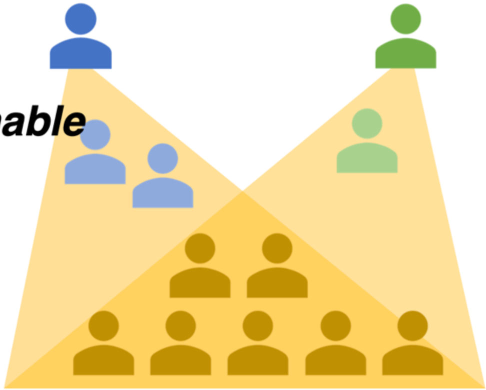

Source: Minji Yoon, CMU

# Two interesting questions about GNN

Question 2: Depth?

**Over-smoothing[10]**

- When GNNs become deep,
  nodes share many neighbors
- Node embeddings become ***indistinguishable***



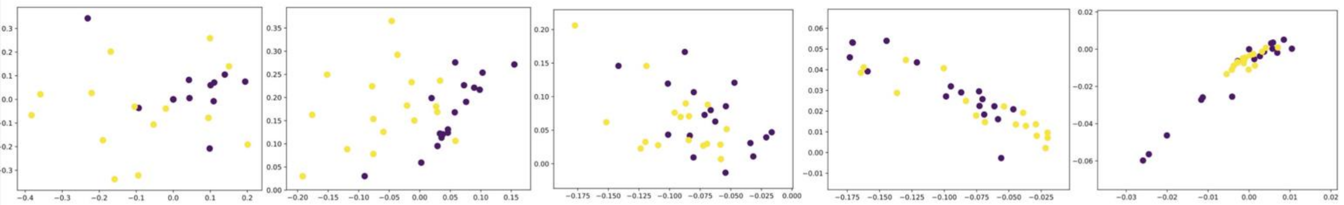
[10] Qimai Li, et al. "Deeper Insights into Graph Convolutional Networks for Semi-Supervised Learning"

Source: Minji Yoon, CMU

Question 2: Depth?

**Over-smoothing[10]**

- Node embeddings of Zachary's karate club network with GNNs



(a) 1-layer (b) 2-layer (c) 3-layer (d) 4-layer (e) 5-layer

[10] Qimai Li, et al. "Deeper Insights into Graph Convolutional Networks for Semi-Supervised Learning"

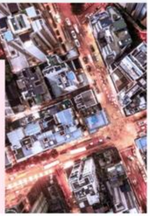
Source: Minji Yoon, CMU

# GNN Applications

DeepMind › Blog › Traffic prediction with advanced Graph Neural Networks

BLOG POST RESEARCH 03 SEP 2020

**Traffic prediction with advanced Graph Neural Networks**



**Food Discovery with Uber Eats: Using Graph Learning to Power Recommendations**

Ankit Jain, Isaac Liu, Ankur Sarda, and Piero Molino December 4, 2019

0

Pinterest Engineering
Aug 15, 2018 · 8 min read

**PinSage: A new graph convolutional neural network for web-scale recommender systems**

Ruining He | Pinterest engineer, Pinterest Labs

**Web image search gets better with graph neural networks**

ı to image search uses images returned by traditional search
des in a graph neural network through which similarity signals are
ıieving improved ranking in cross-modal retrieval.

ıral Network

'ER LABS Europe NAVER LABS FOLLOW

amazon | science

PUBLICATION

P-Companion: A principled framework for diversified complementary product recommendation

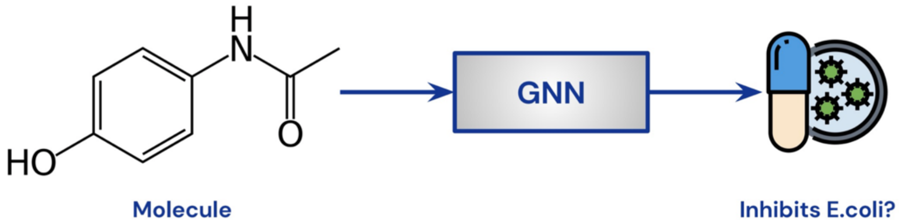
By Junheng Hao, Tong Zhao, Jin Li, Xin Luna Dong, Christos Faloutsos, Yizhou Sun, Wei Wang
2020



Source: Minji Yoon, CMU

# GNN applications

- Graph-level prediction: whether the molecule is a potent **drug**[29]
  - Execute on a large dataset of known candidate molecules
  - Select the ~ *top-100* candidates from the GNN model
  - Have chemists thoroughly investigate those

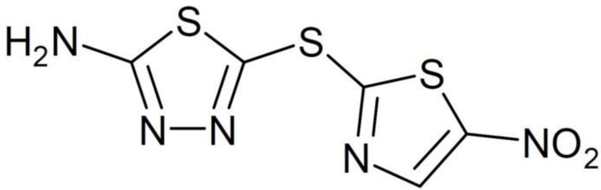



[29] Jonathan M.Stokes, et al. "A Deep Learning Approach to Antibiotic Discovery"

Source: Minji Yoon, CMU

# GNN applications

- Discover a previously overlooked compound that is a **highly potent** antibiotic[29]




Halicin

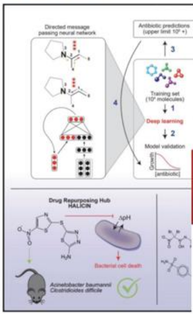
[29] Jonathan M.Stokes, et al. "A Deep Learning Approach to Antibiotic Discovery"

# GNN applications

**Cell**

**A Deep Learning Approach to Antibiotic Discovery**

Graphical Abstract



Authors

Jonathan M. Stokes, Kevin Yang, Kyle Swanson, ..., Tommi S. Jaakkola, Regina Barzilay, James J. Collins

Correspondence

regina@csail.mit.edu (R.B.), jimjc@mit.edu (J.J.C.)

**nature**

Subscribe

NEWS · 20 FEBRUARY 2020

## Powerful antibiotics discovered using AI

Machine learning spots molecules that work even against 'untreatable' strains of bacteria.

BBC Sign in News Sport Reel Worklife Travel Future

NEWS

Home Video World UK Business Tech Science Stories Entertainment & Arts

BBC WORKLIFE Our new guide for getting ahead

## Scientists discover powerful antibiotic using AI

21 February 2020 Share

FINANCIAL TIMES

COMPANIES TECH MARKETS GRAPHICS OPINION WORK & CAREERS LIFE & ARTS HOW TO SPEND IT

CORONAVIRUS **BUSINESS UPDATE**

Get 30 days' complimentary access to our Coronavirus Business Update newsletter

intelligence

obotics 'Death of the office' homeworking claims exaggerated Anti-social robots hari increase social distanc

Artificial intelligence + Add to myFT

## AI discovers antibiotics to treat drug-resistant diseases

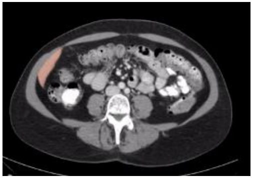
Machine learning uncovers potent new drug able to kill 35 powerful bacteria

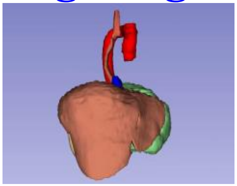
[29] Jonathan M.Stokes, et al. "A Deep Learning Approach to Antibiotic Discovery"

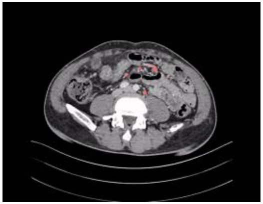
# After the break

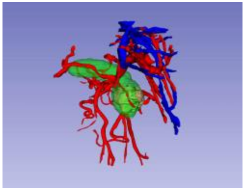
After the break: **U-Net and U-mamba, MedSAM**

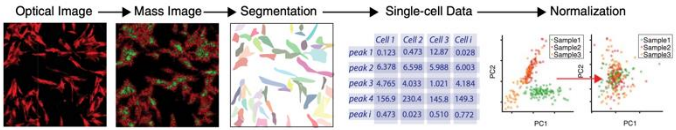
# Biomedical Image Segmentation: What and Why



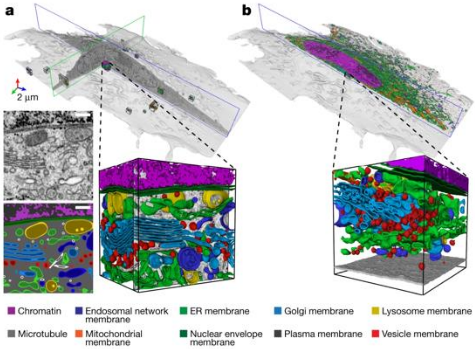









Capolupo, L et al. Science, 2022




Heinrich, L. et al. Nature, 2021

- Quantification of anatomical structures and disease progression
- Cancer microenvironment analysis

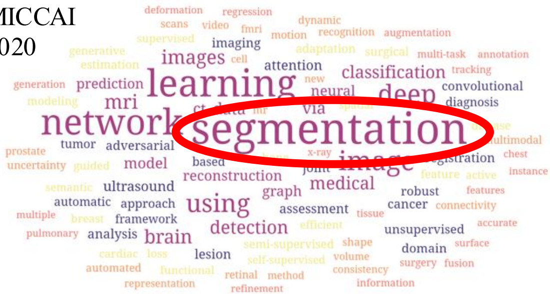
**Segmentation is the core technology towards precise biomedical image analysis!**

# Biomedical Image Segmentation

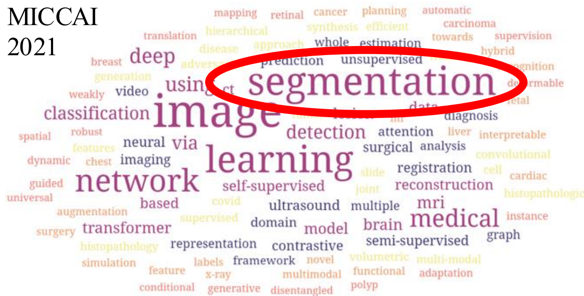
## Biomedical Image Segmentation is Still an Active Research Field!

**Word Cloud of paper titles in MICCAI 2020-2023**

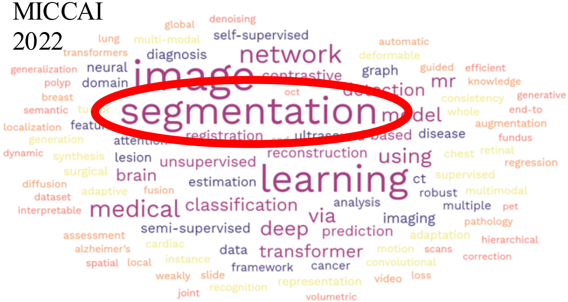
MICCAI 2020



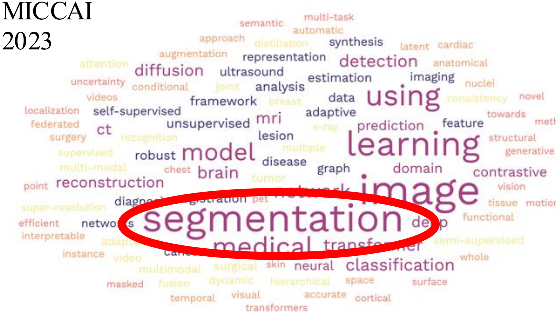

MICCAI 2021




MICCAI 2022




MICCAI 2023



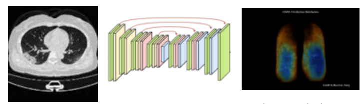

https://github.com/JunMa11/MICCAI-OpenSourcePapers

# Segmentation Paradigm Over the Past Half Century

| Heuristic methods | Optimization/model-based methods | DL-based tailored models | Foundation models |
|---|---|---|---|
| | 1980s | 2015-2022 | 2023 |
| Thresholding<br>Region Growing<br>… | Variational Models<br>(e.g., Snakes, GAC,…) | FCN, U-Net, DeepLab,<br>… | Segment Anything<br>(SAM) |

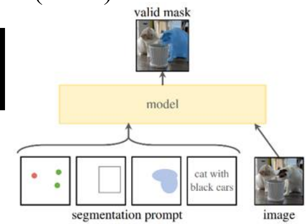
$$S^* = arg\min_S E(S)$$

Ma et al.
SIAM-IS,
2021



Ma et al., MedPhys,
2021

Customized Models




11M images & 256 GPUs

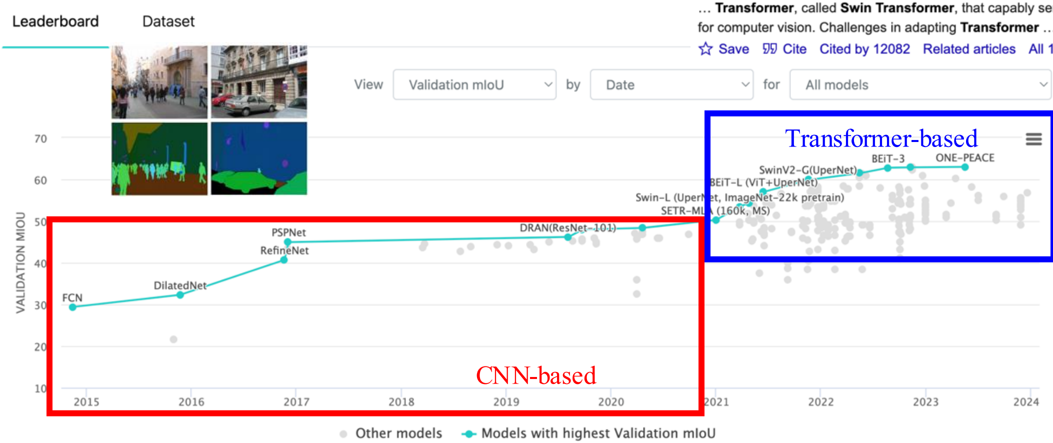
**More data**

**More powerful computing**

# Biomedical Image Segmentation

## What Are the SOTA Automatic Segmentation Networks?

Semantic Segmentation on ADE20K




**Swin transformer**: Hierarchical vision **transformer** using shifted windows
Z Liu, Y Lin, Y Cao, H Hu, Y Wei… - Proceedings of the …, 2021 - openaccess.thecvf.com
… **Transformer**, called **Swin Transformer**, that capably serves as a general-purpose backbone for computer vision. Challenges in adapting **Transformer** … a hierarchical **Transformer** whose …
Save Cite Cited by 12082 Related articles All 10 versions

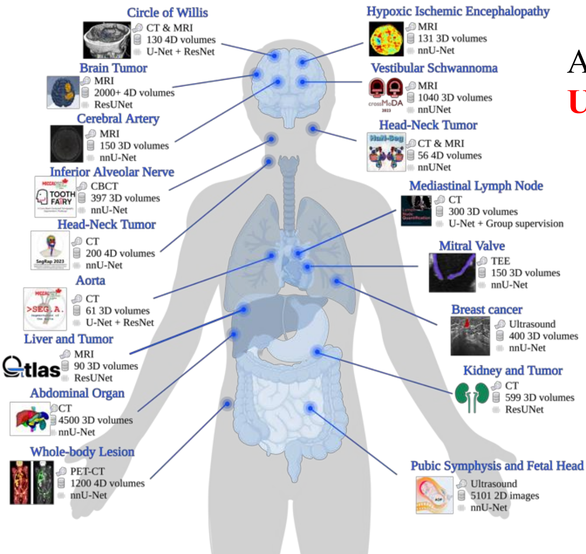
**Fully convolutional networks** for **semantic segmentation**
J Long, E Shelhamer, T Darrell - Proceedings of the IEEE …, 2015 - openaccess.thecvf.com
… for per-pixel tasks like **semantic** segmentation. We show that a **fully convolutional network** (FCN) trained end-to-end, pixels-to-pixels on **semantic** segmentation exceeds the state-of-the-…
Save Cite Cited by 44857 Related articles All 52 versions Web of Science: 2272

**Transformer-based networks are current SOTA on the natural image segmentation benchmark.**

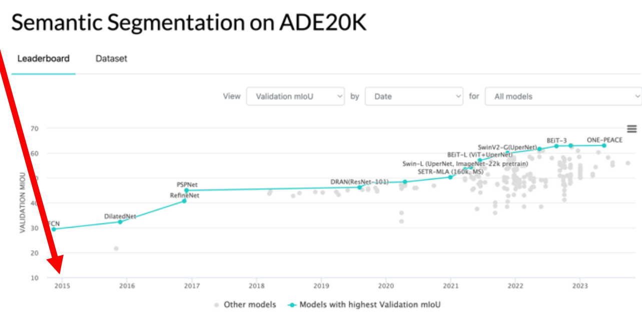
https://paperswithcode.com/sota/semantic-segmentation-on-ade20k

# Biomedical Image Segmentation

## What Are the SOTA Networks for Medical Image Segmentation?




All the winning algorithms are based on **U-Net** and its variants.



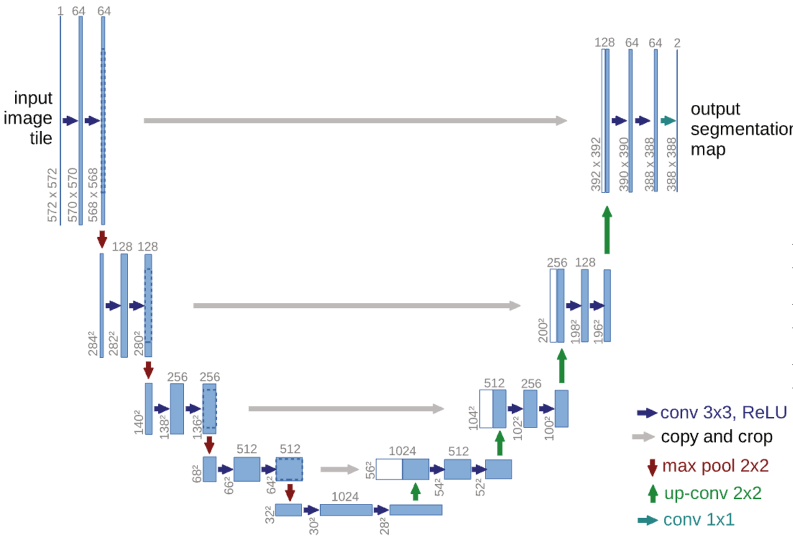

**U-net: Convolutional networks for biomedical image segmentation**
O Ronneberger, P Fischer, T Brox - ... image computing and computer ..., 2015 - Springer
... We demonstrate the application of the **u-net** to three different **segmentation** tasks. The first task is the **segmentation** of neuronal structures in electron microscopic recordings. An ...
☆ Save 99 Cite Cited by 96646 Related articles All 32 versions

A survey of 17 segmentation challenges in MICCAI 2023

https://github.com/JunMa11/SOTA-MedSeg

# Biomedical Image Segmentation

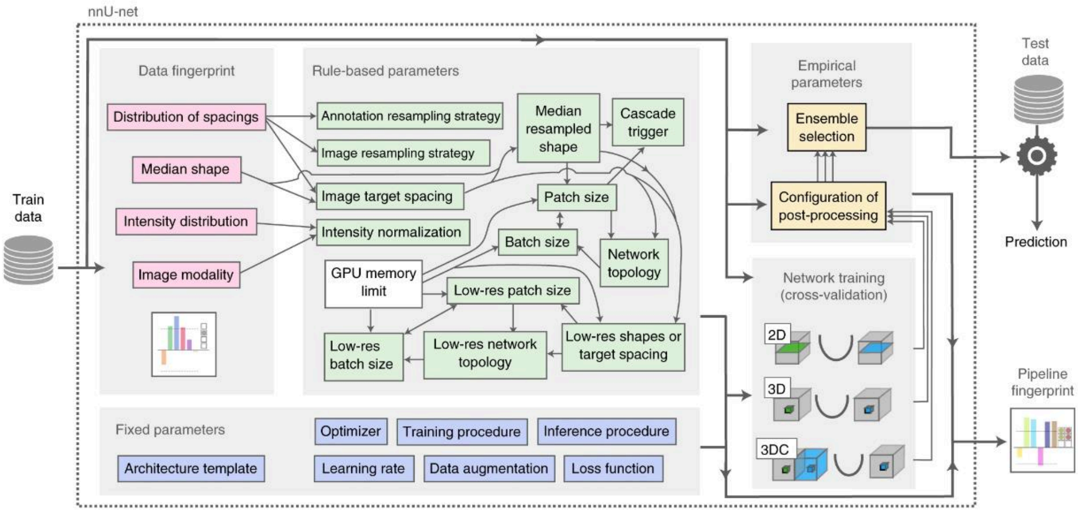
## What are the advantages of U-Net?




- Symmetry encoder-decoder design
- Skip connection to recover details
- Versatility across modalities

# Biomedical Image Segmentation

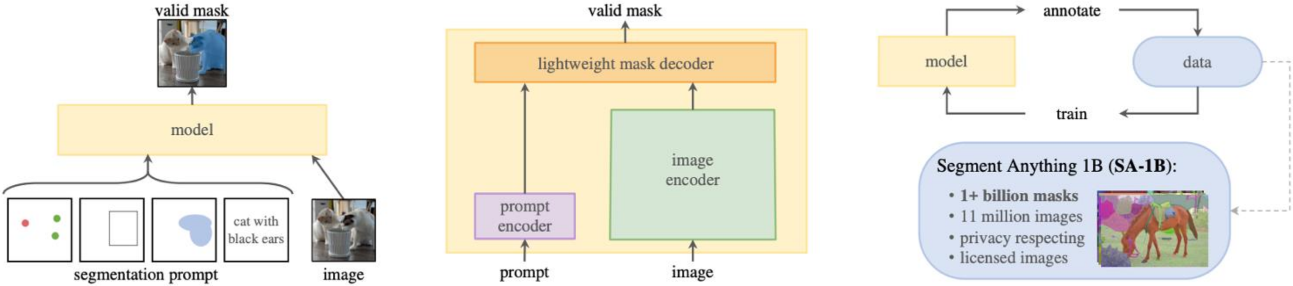
## nnU-Net: Automatically config U-Net



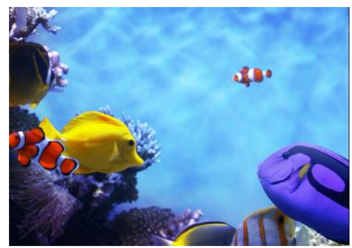

Isensee, Fabian, et al. "nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation." *Nature Methods* 18.2 (2021): 203-211.

# Biomedical Image Segmentation

## Segment Anything Model (SAM)



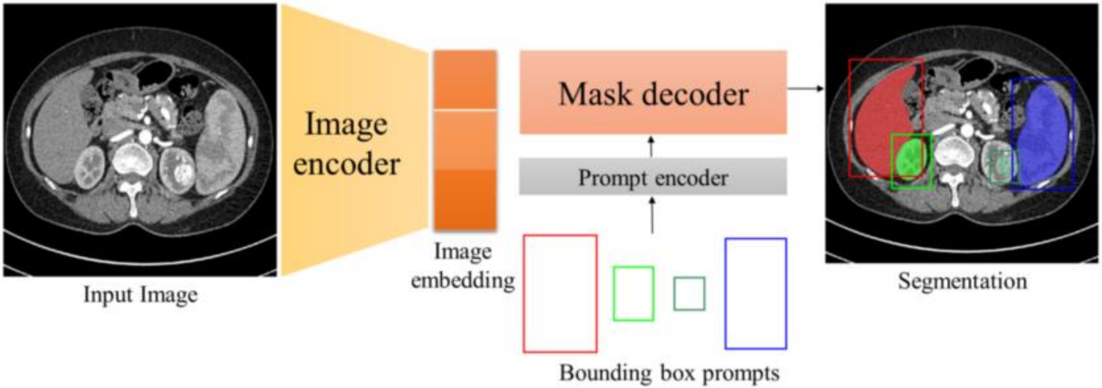

(a) **Task**: promptable segmentation (b) **Model**: Segment Anything Model (**SAM**) (c) **Data**: data engine (top) & dataset (bottom)



Kirillov, A., et al. "Segment anything." ICCV, 2023

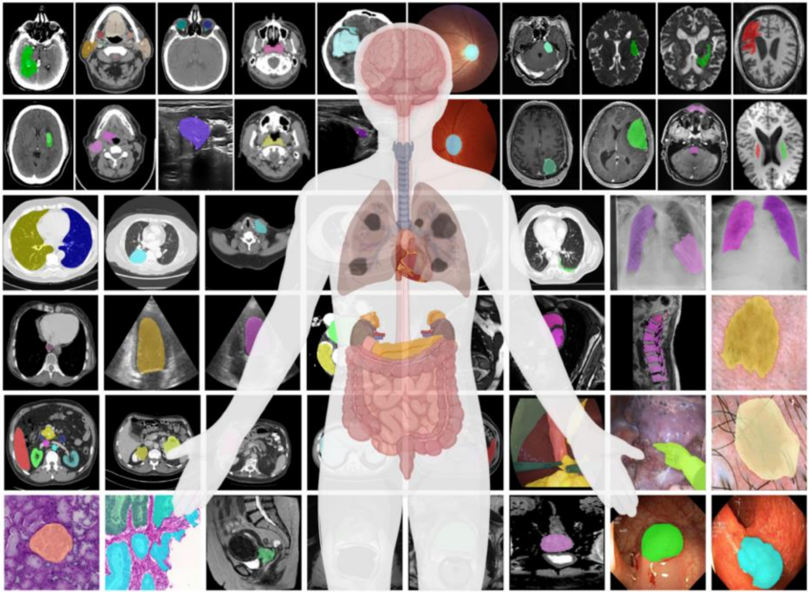
# MedSAM: Pipeline

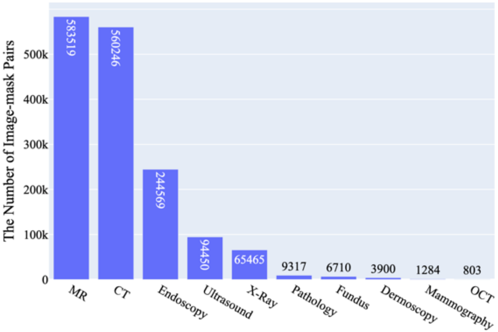
Fine-tune both image encoder and mask decoder



# Biomedical Image Segmentation

## MedSAM: 1M image-mask Pairs



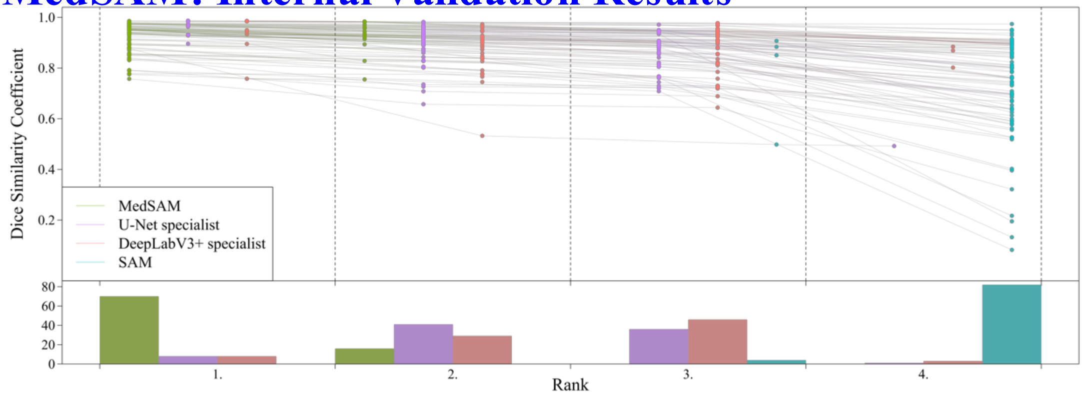


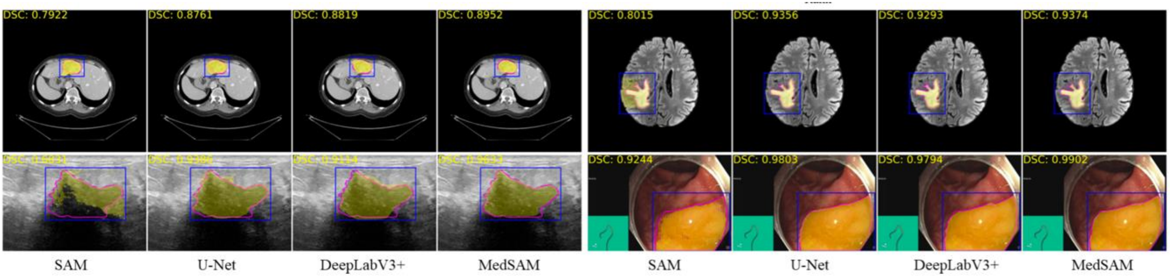

Experimental Settings

- 86 internal validation tasks
- 60 external validation tasks
- Compared to specialist U-Nets and DeepLabV3+ that are trained on each modality

# Biomedical Image Segmentation

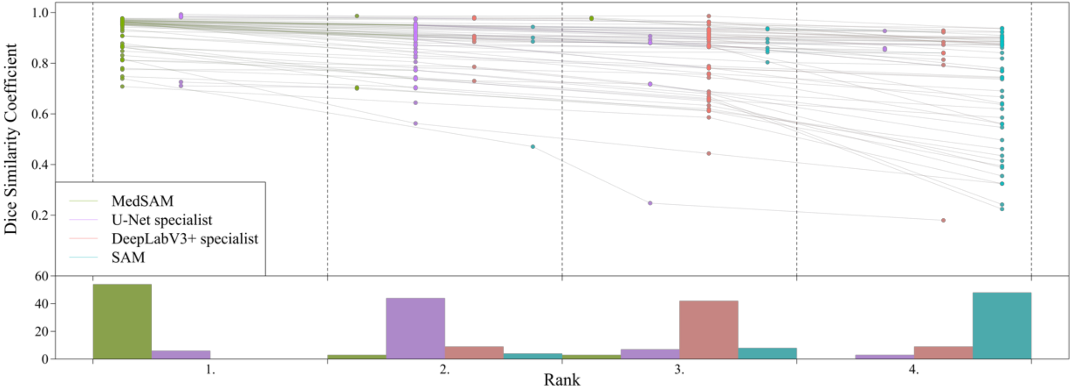
## MedSAM: Internal Validation Results

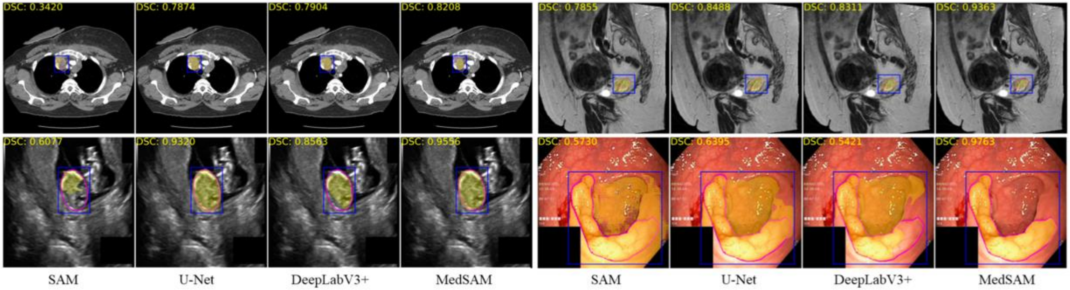


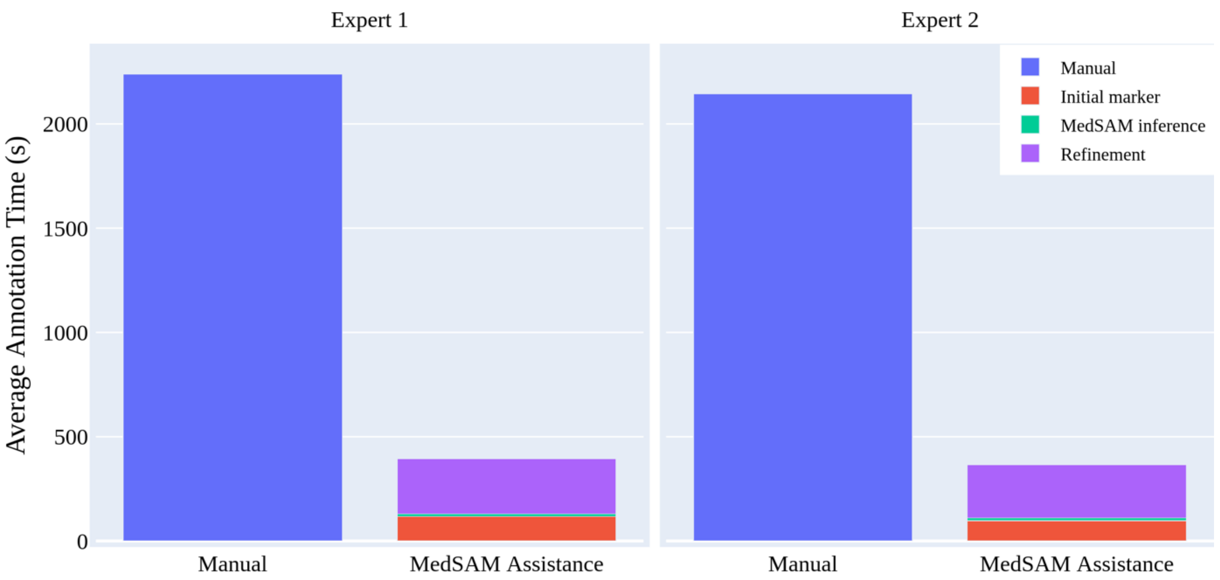
# Biomedical Image Segmentation

## MedSAM: External Validation Results



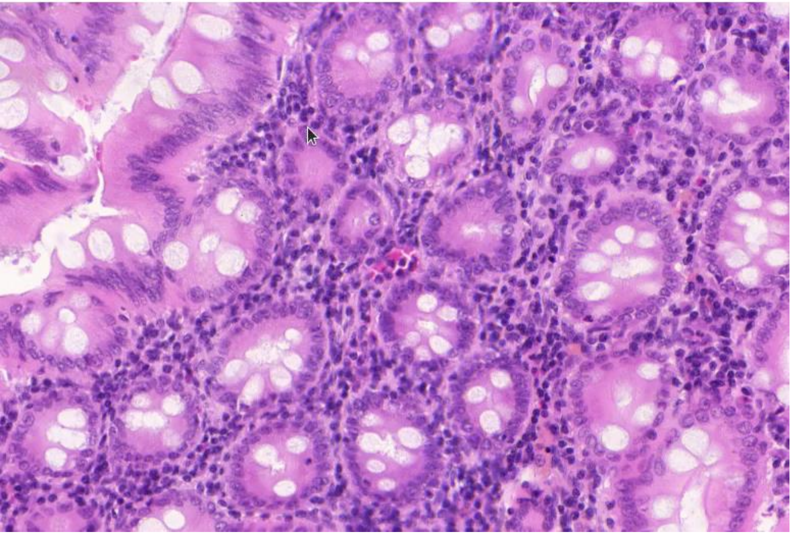



## Human Annotation Study



# Biomedical Image Segmentation

## MedSAM: Demo

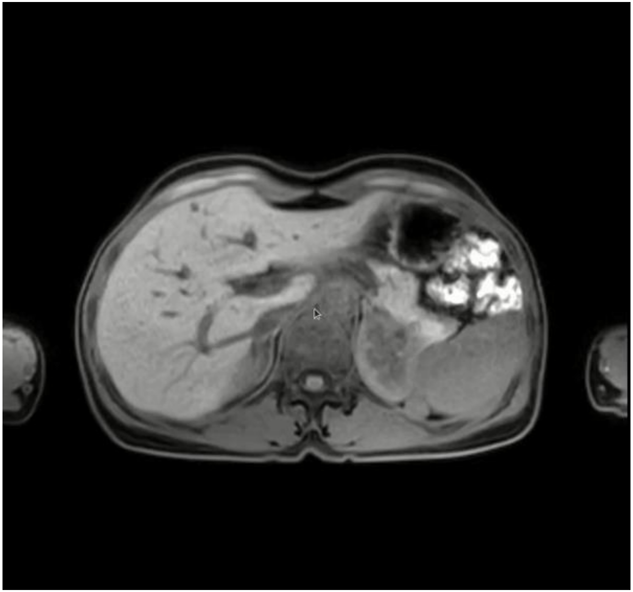


Gland Segmentation in Pathology Images

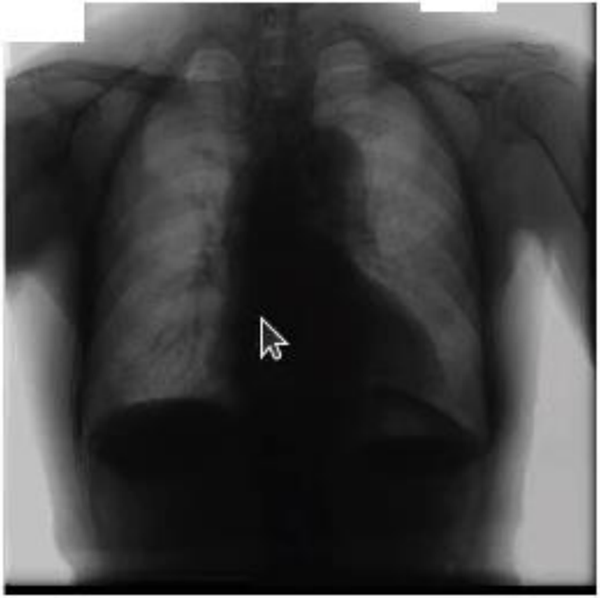
Liver and Tumor Segmentation in CT

# Biomedical Image Segmentation

## MedSAM: Demo

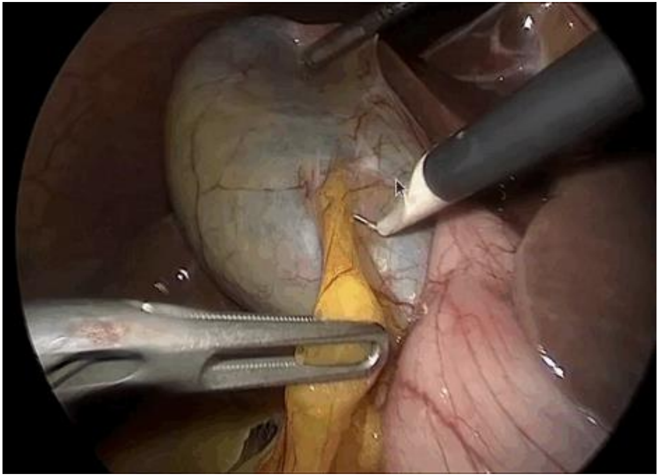


Abdominal Organ Segmentation in MR

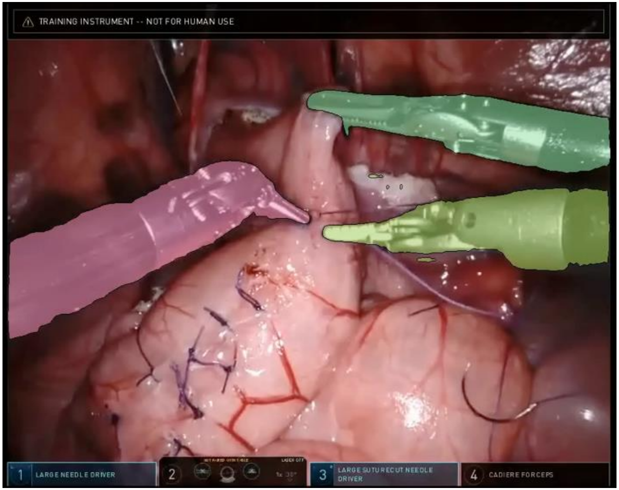


Lungs and Heart Segmentation in X-Ray

## MedSAM: Demo



Tissue and Instruments Segmentation in Endoscopy Image



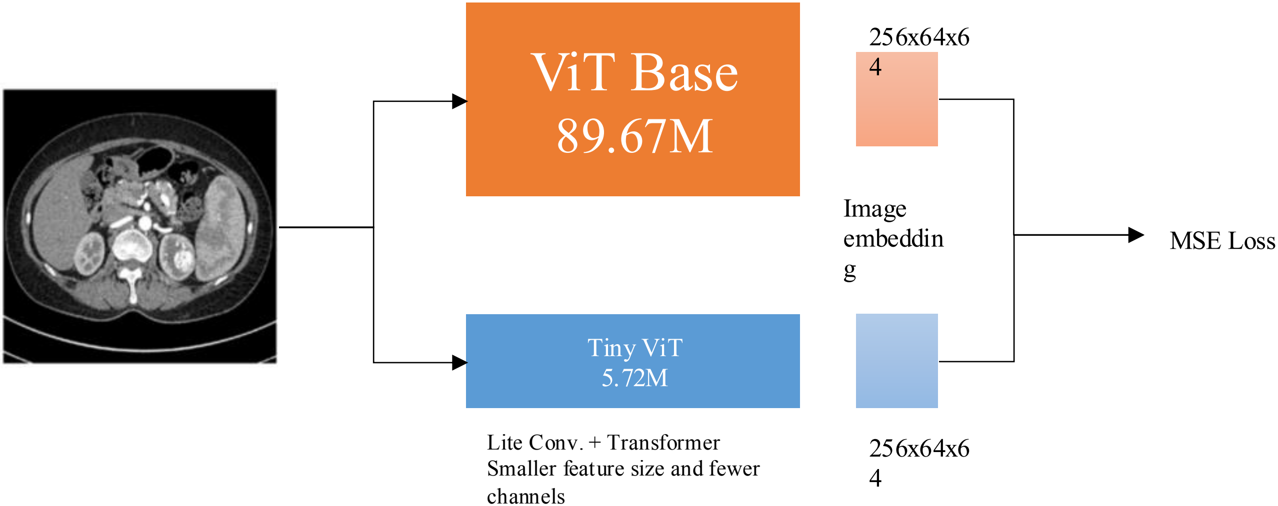

Instruments Segmentation and Tracking in Endoscopy Video

# Biomedical Image Segmentation

## How can we make the model accessible to medical professionals?

Answer: A Lightweight MedSAM (distillation and fine-tuning)

Stage 1. Distillation a small image encoder



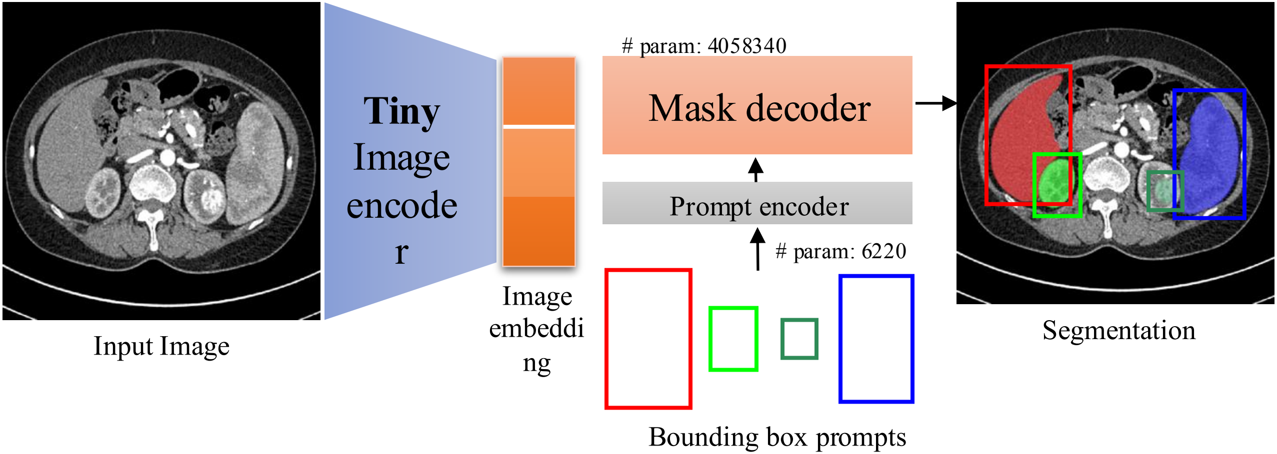


Wu, Kan, et al. "Tinyvit: Fast pretraining distillation for small vision transformers." ECCV, 2022.
Zhang, Chaoning, et al. "Faster Segment Anything: Towards Lightweight SAM for Mobile Applications." *arXiv preprint arXiv:2306.14289* (2023).
Zhao, Xu, et al. "Fast Segment Anything." arXiv preprint arXiv:2306.12156 (2023).

# Biomedical Image Segmentation

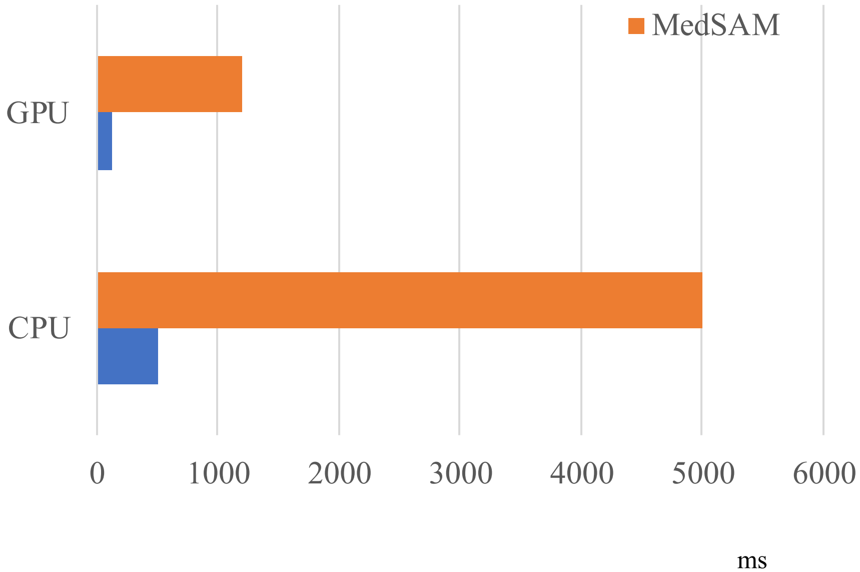
## Lite MedSAM: 10× Faster

Stage 2. Fine-tune the whole model

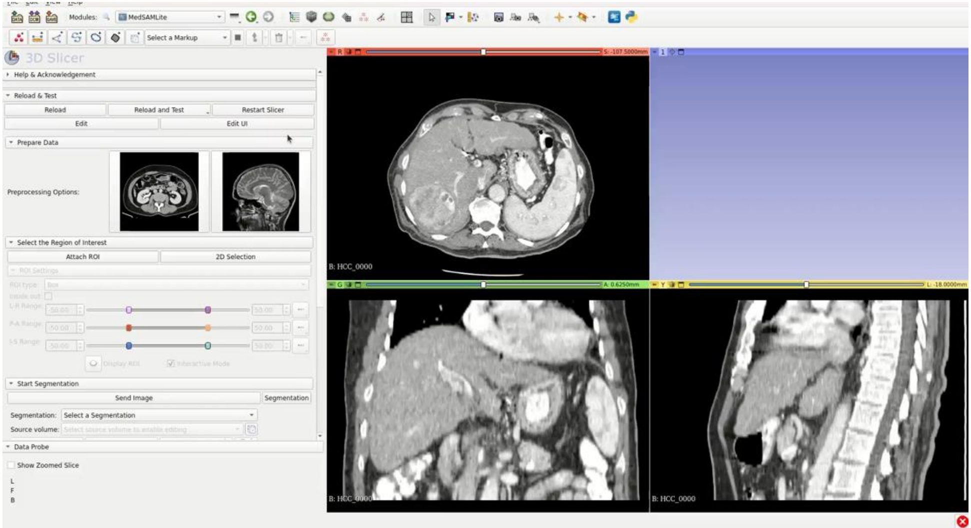


# Biomedical Image Segmentation

## Lite MedSAM: 10× Faster



# 3D Slicer Integration (Open-source Platform)




https://www.slicer.org /

https://github.com/bowang-lab/MedSAMSlicer

# Biomedical Image Segmentation

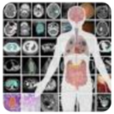
## MedSAM in Community

Google Scholar (~1000 citations in eight months)

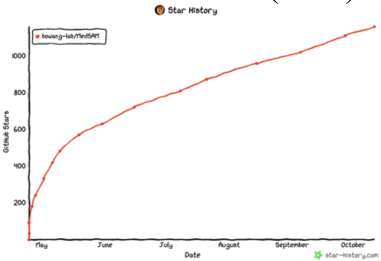
Nature
https://www.nature.com › ... › articles

**Segment anything in medical images - Nature**

by J Ma · 2024 · Cited by 975 — We introduce MedSAM, a deep learning-powered foundation model designed for the **segmentation** of a wide array of anatomical structures and lesions ...



GitHub Stars (2.2K)




MedSAM in HuggingFace

### Segment medical images with MedSAM

In this notebook, we're going to perform inference with MedSAM, a fine-tuned version of the SAM (segment-anything model) by Meta AI on the medical domain (thereby greatly improving its performance).

- Original repo
- Hugging Face docs.

https://github.com/NielsRogge/Transformers-Tutorials/blob/master/SAM/Run_inference_with_MedSAM_using_HuggingFace_Transformers.ipynb

MedSAM in napari

I integrated MedSam into napari FYI #36

Closed Karol-G opened this issue on May 5 · 2 comments

Karol-G commented on May 5

Hey,

I just wanted to let you know that I integrated MedSam already into my Napari SAM plugin: https://github.com/MIC-DKFZ/napari-sam

So you can check the mark on "3D slicer and napari support" on your todo list if you want ;)

Best,
Karol

https://github.com/MIC-DKFZ/napari-sam