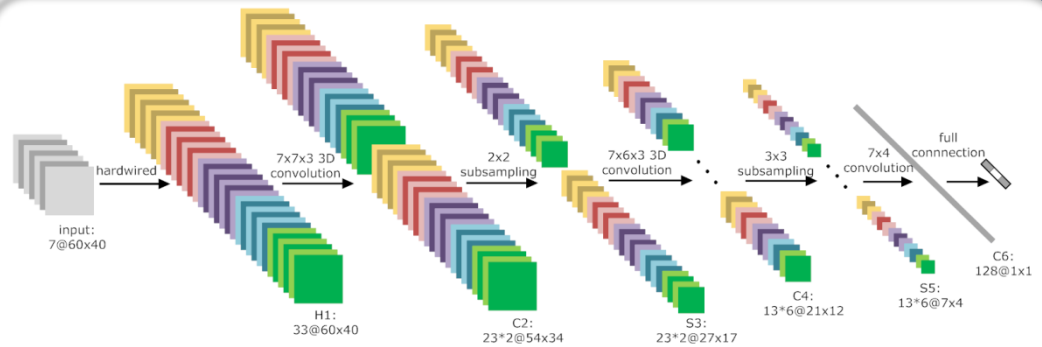
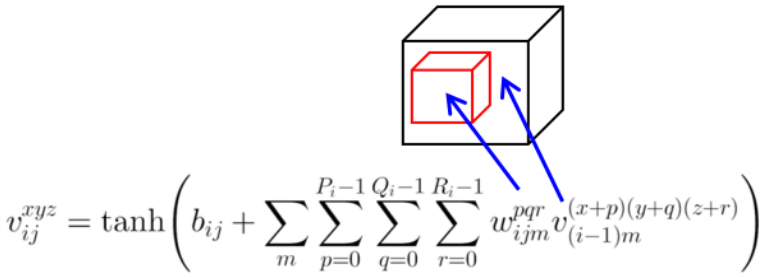
TEXAS A&M UNIVERSITY
Engineering

Image and Graph Deep Learning Methods for Cellular and Molecular Level Science Discoveries

Shuiwang Ji
Department of Computer Science & Engineering

Deep Learning in Pre-AlexNet Era (2009)





$$v_{ij}^{xyz} = \tanh\left(b_{ij} + \sum_{m}\sum_{p=0}^{P_i-1}\sum_{q=0}^{Q_i-1}\sum_{r=0}^{R_i-1} w_{ijm}^{pqr} v_{(i-1)m}^{(x+p)(y+q)(z+r)}\right)$$

Recognized examples of "CellToEar"

- Action recognition from real-world airport surveillance video
- One of the earliest work on deep learning for video analysis
- Winner of the 2009 TRECVID Video Surveillance Challenge

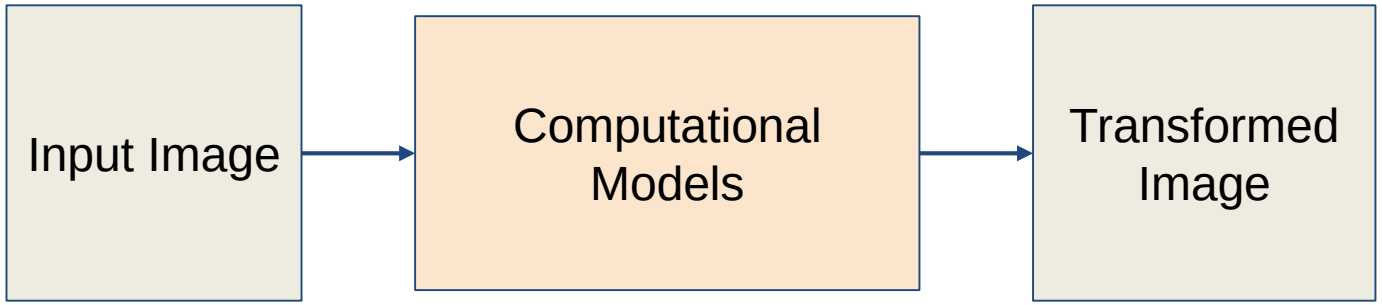
Ji *et al*., 3D Convolutional Neural Networks for Human Action Recognition. ICML'2010 and TPAMI'2013. 4,000+ citations

Today: Beyond Natural Image and Video Recognition

- More complex images in cellular and molecular imaging
- Image to image transformation problems
- Beyond grid-like data, such as graphs
- Use of graph techniques for molecular level analysis

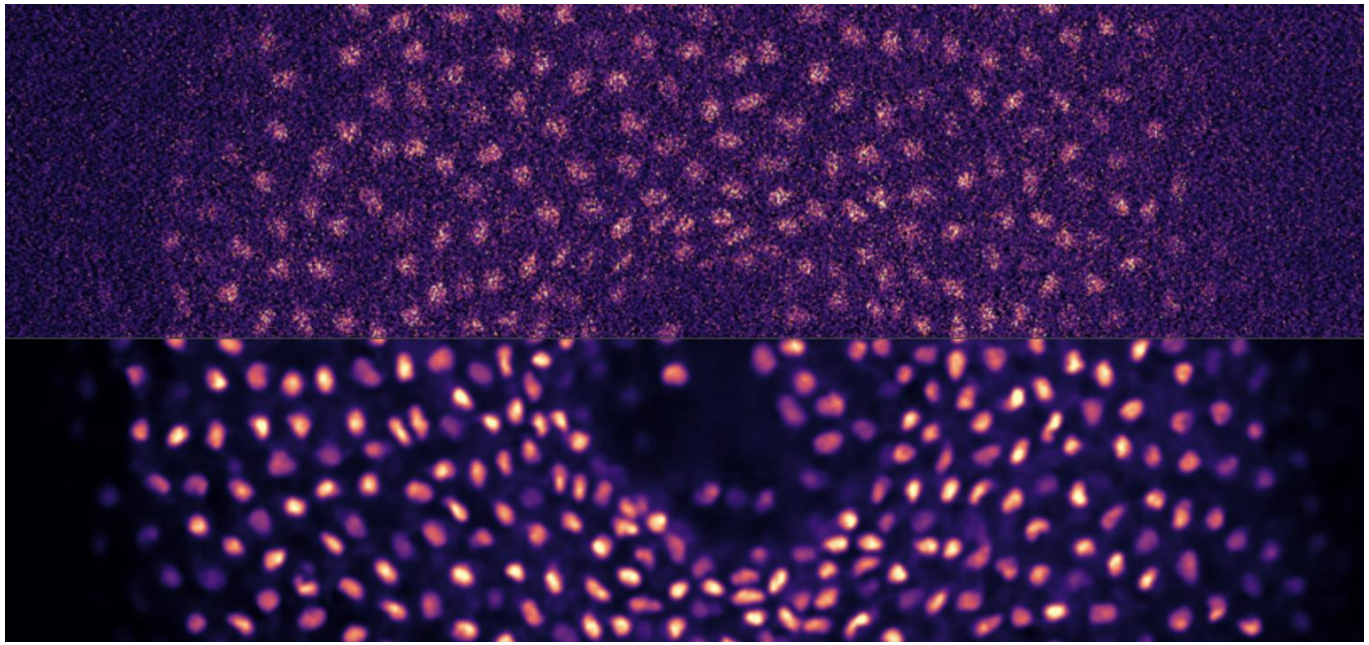
Biological Image Transformations

- Biological image transformations in general



Biological Image Transformations

- Image Denoising

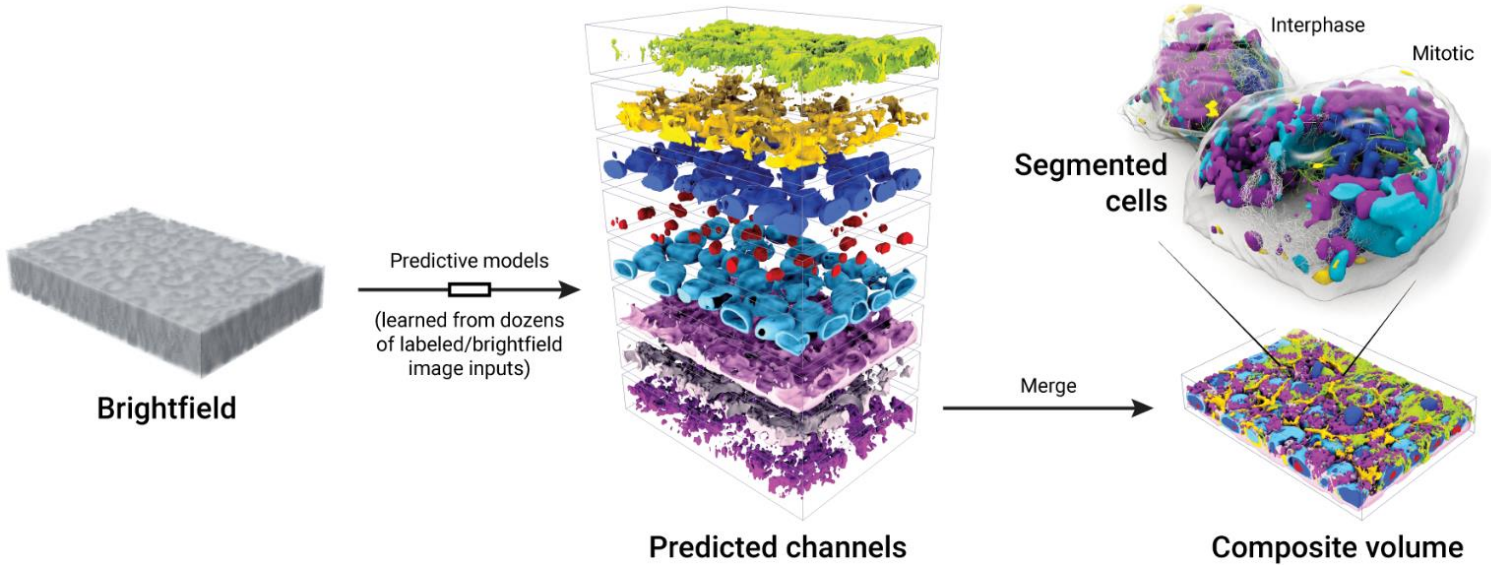


Weigert, Martin, et al. "Content-aware image restoration: pushing the limits of fluorescence microscopy." Nature methods 15.12 (2018): 1090.

Biological Image Transformations

TEXAS A&M UNIVERSITY
Engineering

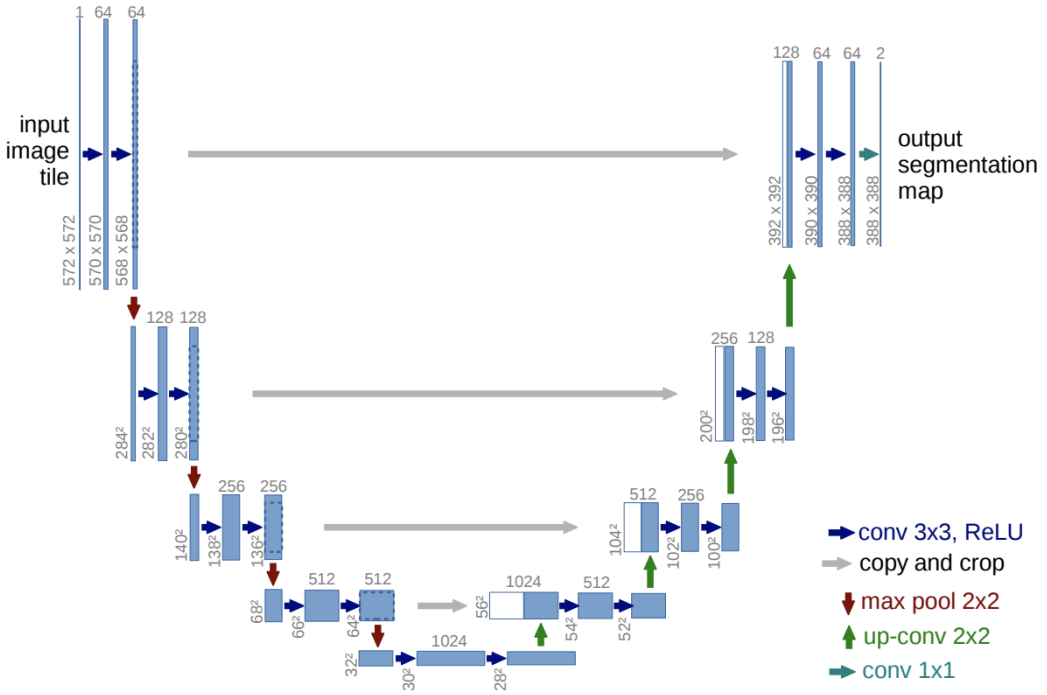
- Substructure Prediction & Segmentation



Ounkomol, Chawin, et al. "Label-free prediction of three-dimensional fluorescence images from transmitted-light microscopy." Nature methods 15.11 (2018): 917.

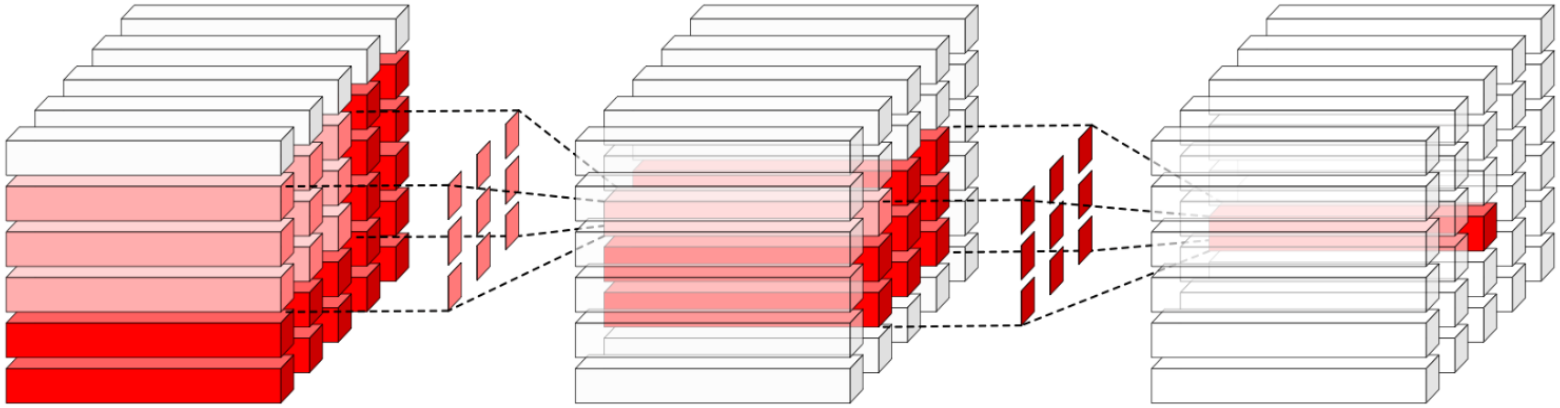
U-Net

- Current state-of-the-art and most popular network for bioimage transformations
- Architecture
  - Down-sampling path
  - Up-sampling path
  - Skip connections



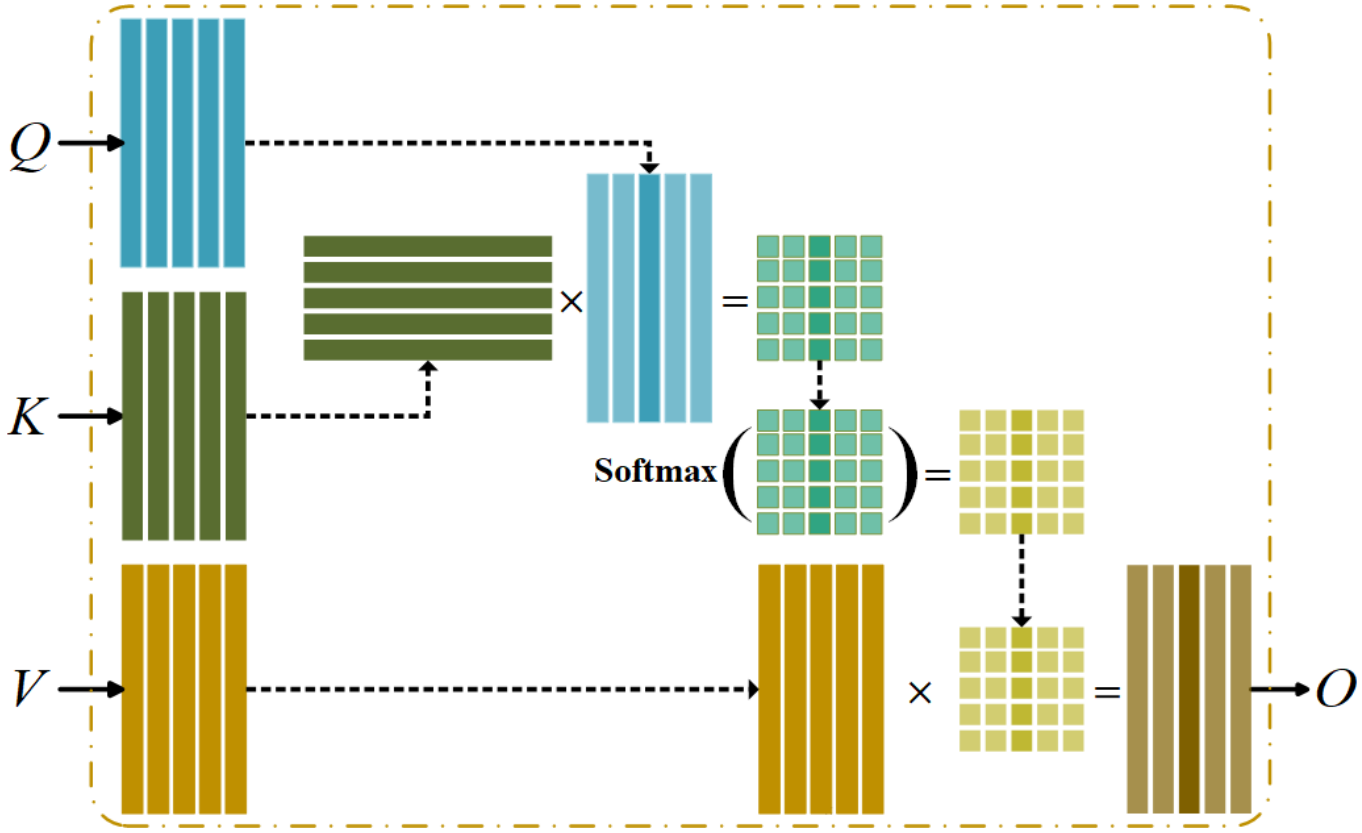
Ronneberger, Olaf, Philipp Fischer, and Thomas Brox. "U-net: Convolutional networks for biomedical image segmentation." International Conference on Medical image computing and computer-assisted intervention. Springer, Cham, 2015.

Local Operators in U-Net

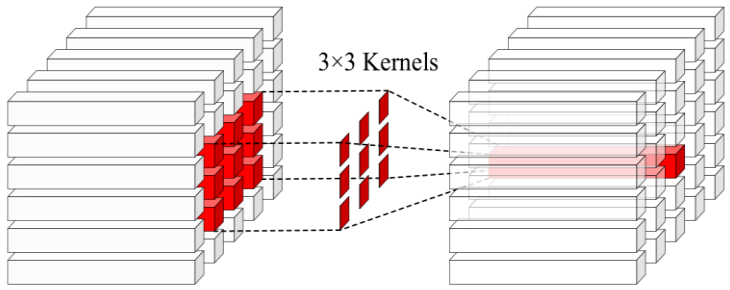


- Limited receptive field
- Focus more on local dependencies
- Fixed weights of kernels once trained

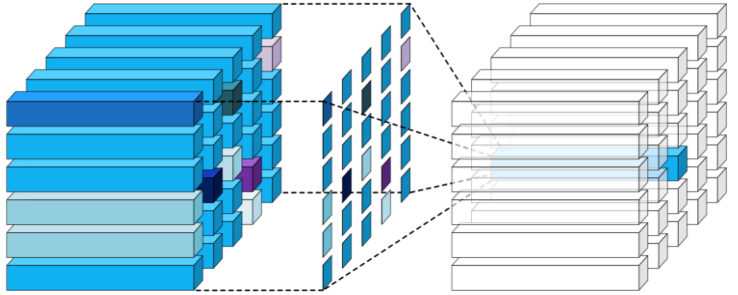
Attention Operator



Local v.s. Non-local



(a) Convolution



(b) Attention Operator

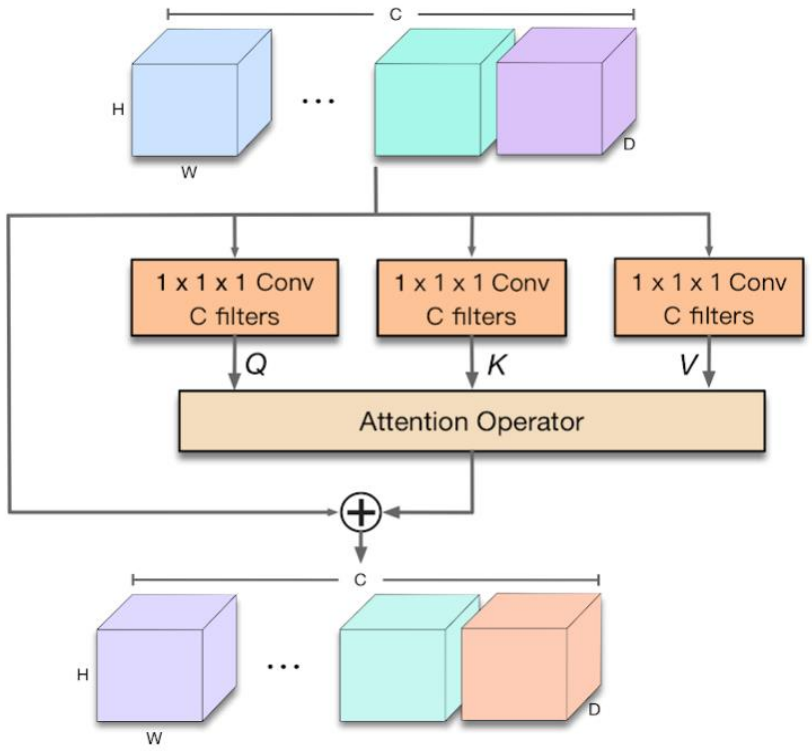
- Location relationship based
- Pixels only have access to their neighbour pixels in the previous layer
- Determined weights

- Relevance or similarity based
- Pixels have access to other pixels globally
- Input-dependent weights

Vaswani, Ashish, et al. "Attention is all you need." Advances in neural information processing systems. 2017.

Fundamentally different with fully connected layers

Global Voxel Transformation Operators (GVTOs)



Size-preserving GVTO

Input tensor $\mathscr{X} \in \mathbb{R}^{d \times h \times w \times c}$

1 x 1 x 1 Convs: $query\ (\mathscr{Q}), key\ (\mathscr{K}), value\ (\mathscr{V})$

Unfold: $\boldsymbol{Q}, \boldsymbol{K}, \boldsymbol{V} \in \mathbb{R}^{c \times dhw}$

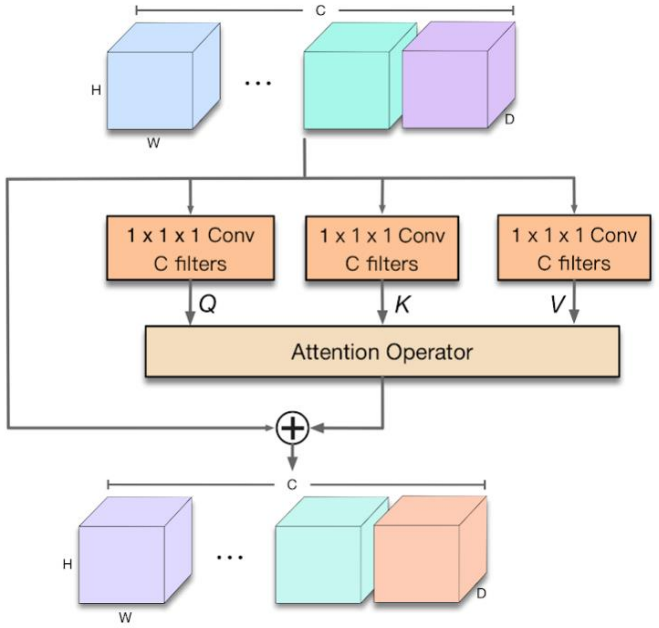
$$\boldsymbol{Y} = \boldsymbol{V} \cdot Normalize(\boldsymbol{K}^T \boldsymbol{Q}) \in \mathbb{R}^{c \times dhw}$$

Fold back: $\mathscr{Y} \in \mathbb{R}^{d \times h \times w \times c}$

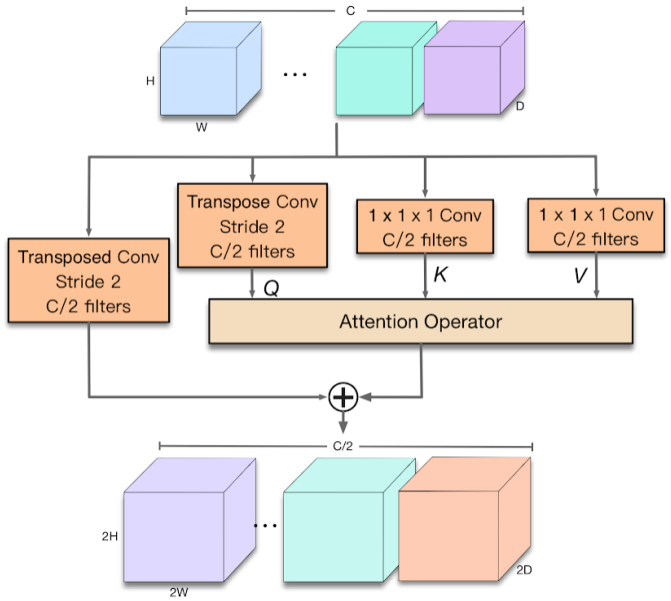
Output tensor $\mathscr{X} + \mathscr{Y}$

Global Voxel Transformation Operators (GVTOs)

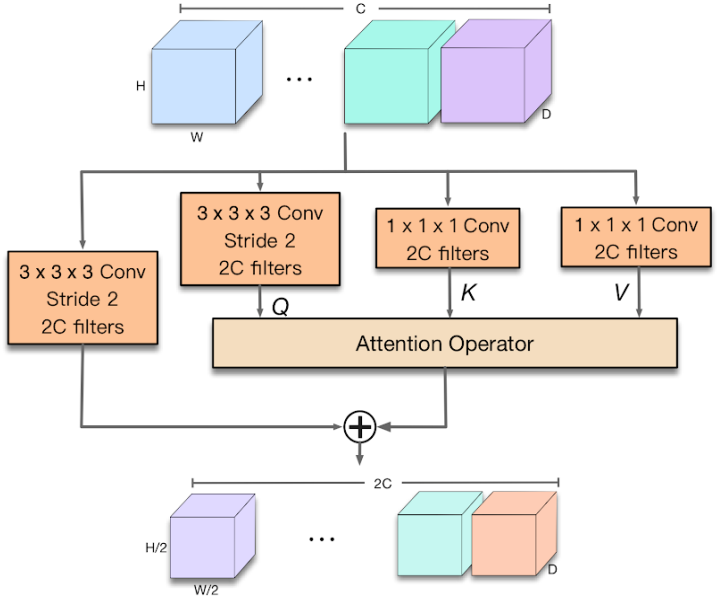
TEXAS A&M UNIVERSITY
Engineering



Size-preserving GVTO



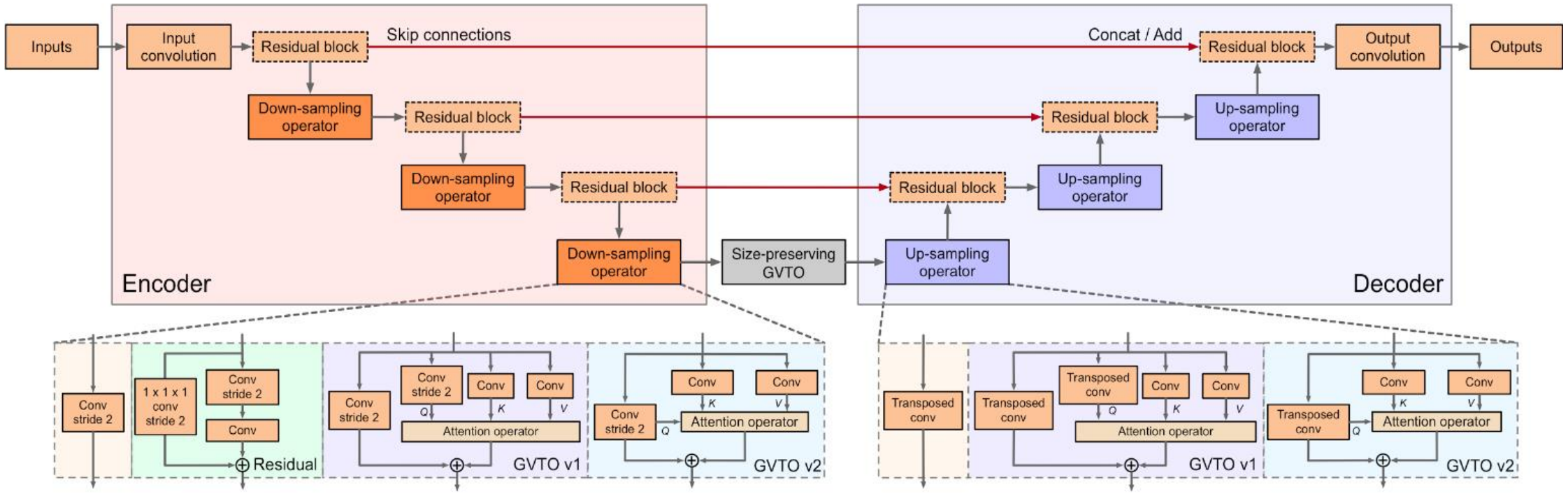
Up-sampling GVTO



Down-sampling GVTO

Global Voxel Transformer Networks for Augmented Microscopy. Nature Machine Intelligence, 2021

GVTNets



- Size-preserving GVTOs
- Apply Up-sampling / Down-sampling GVTOs optionally

Global Voxel Transformer Networks for Augmented Microscopy. Nature Machine Intelligence, 2021

Results: 3D Image Denoising

- CARE datasets

  Weigert, Martin, et al. "Content-aware image restoration: pushing the limits of fluorescence microscopy." Nature methods 15.12 (2018): 1090.

  - 3D Image Denoising

		SSIM			PSNR		
		Input	Baseline	Ours	Input	Baseline	Ours
Planaria	C0	0.226	0.771	**0.795**	22.22	31.57	**32.09**
	C1	0.183	0.740	**0.778**	21.73	30.15	**31.07**
	C2	0.156	0.644	**0.694**	21.44	28.13	**28.72**
Tribolium	C0	0.368	0.917	**0.921**	22.68	32.43	**32.57**
	C1	0.236	0.900	**0.907**	21.37	31.04	**31.21**
	C2	0.182	0.876	**0.885**	20.81	29.41	**29.55**

* ***SSIM*** *: The Structural Similarity Index, the higher the better.*
* ***PSNR*** *: Peak Signal-to-Noise Ratio, the higher the better.*
* ***Baseline****: Weigert, Martin, et al. "Content-aware image restoration: pushing the limits of fluorescence microscopy." Nature methods 15.12 (2018): 1090.*

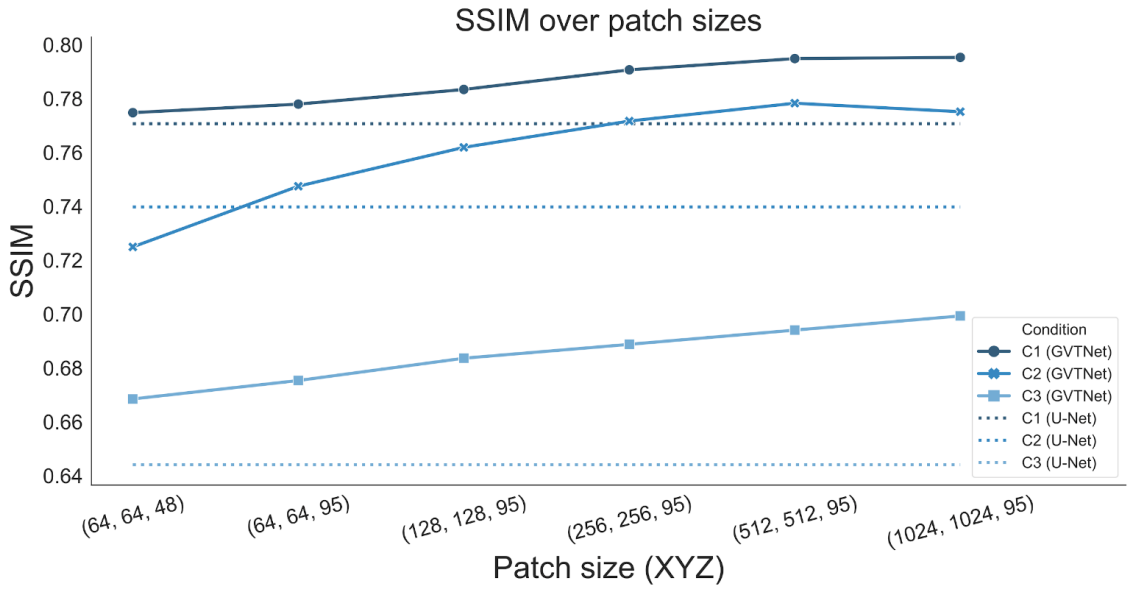
Results: 3D Image Denoising

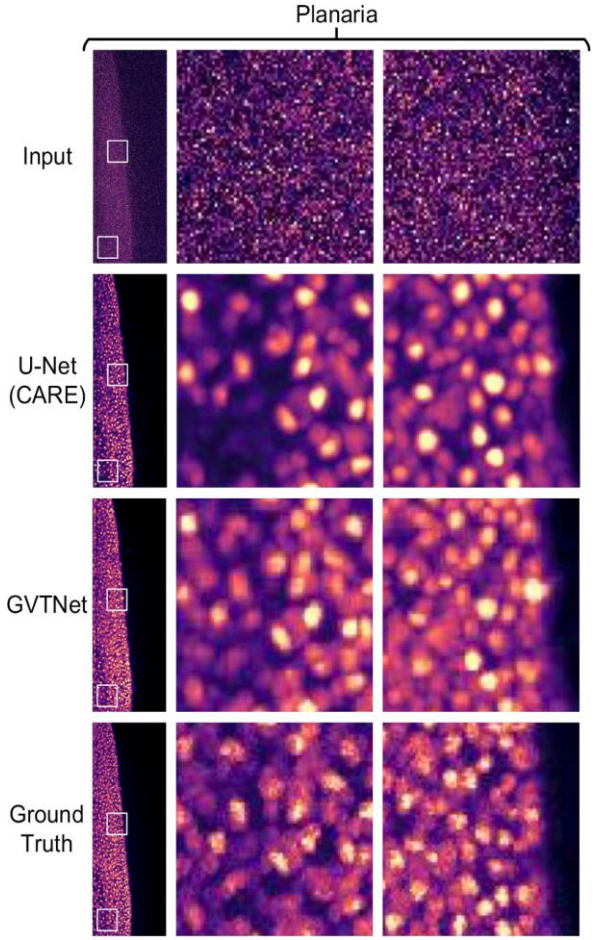
TEXAS A&M UNIVERSITY
Engineering

- CARE datasets

  Weigert, Martin, et al. "Content-aware image restoration: pushing the limits of fluorescence microscopy." Nature methods 15.12 (2018): 1090.

  - 3D Image Denoising (Planaria)





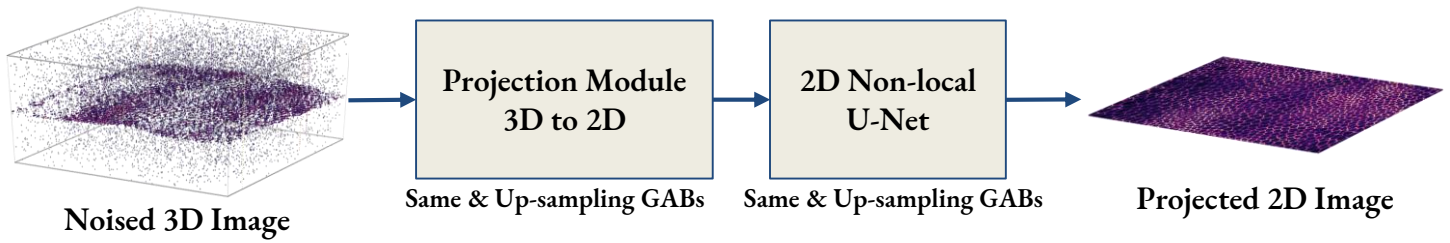
Results: 3D-2D Image Projection

TEXAS A&M UNIVERSITY
Engineering

- CARE datasets

  Weigert, Martin, et al. "Content-aware image restoration: pushing the limits of fluorescence microscopy." Nature methods 15.12 (2018): 1090.

  - 3D Image Projection (Flywing)

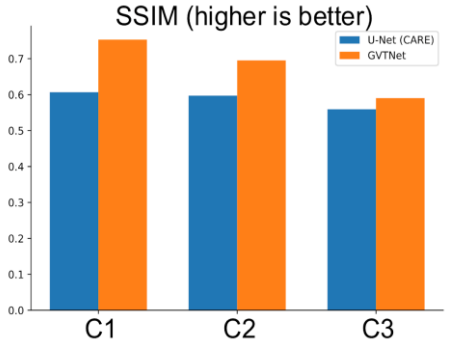


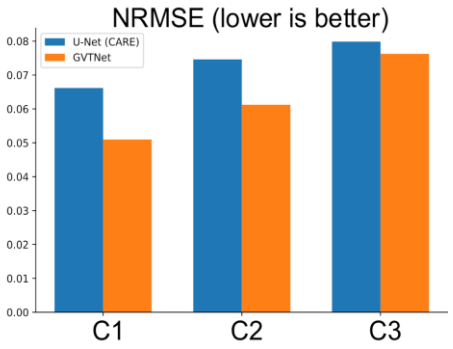
		SSIM			PSNR		
		Input	Baseline	Ours	Input	Baseline	Ours
Flywing	C0	0.190	0.607	**0.753**	18.38	23.66	**25.86**
	C1	0.080	0.597	**0.695**	17.24	22.55	**24.27**
	C2	0.024	0.559	**0.590**	16.63	21.96	**22.35**

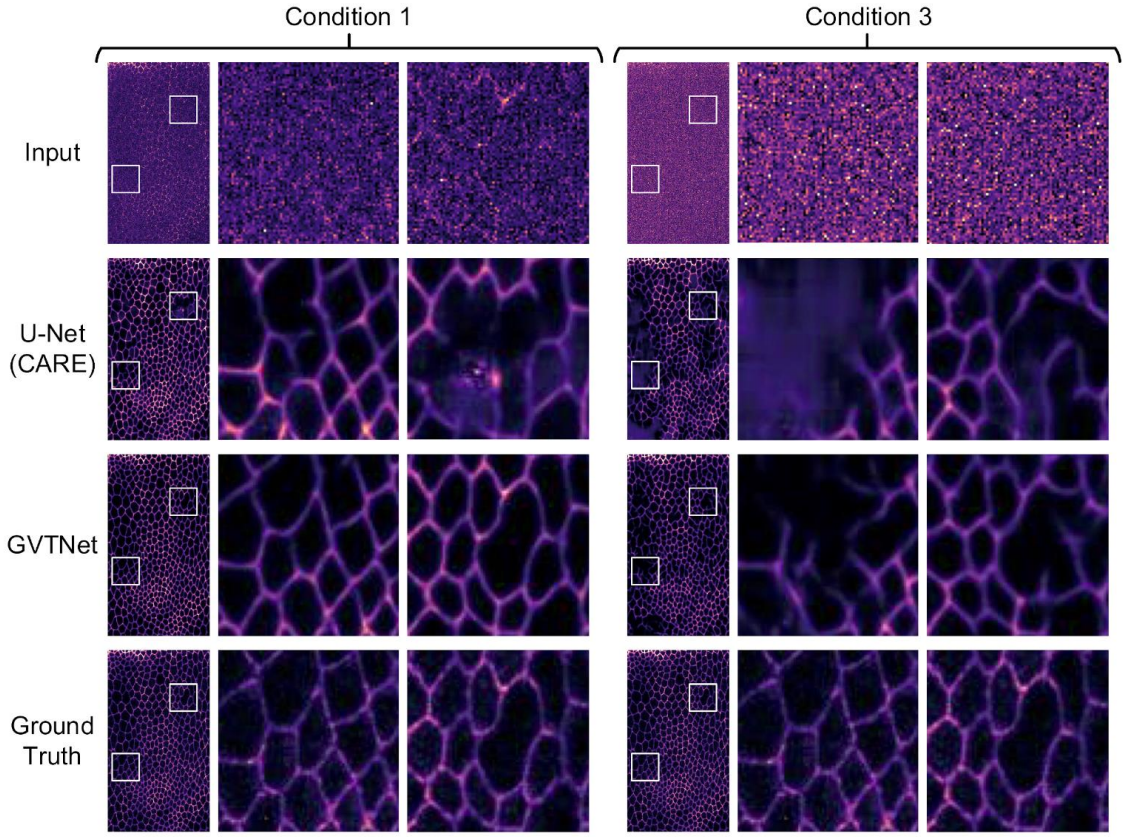
Results: 3D-2D Image Projection

TEXAS A&M UNIVERSITY
Engineering

- Projection (Flywing)







Results: Label-Free Prediction

TEXAS A&M UNIVERSITY
Engineering

- Label-free prediction

  Ounkomol, Chawin, et al. "Label-free prediction of three-dimensional fluorescence images from transmitted-light microscopy." Nature methods 15.11 (2018): 917.

	Pearson Correlation			
Dataset	Baseline (AVG)	Ours (AVG)	Difference	P-value
fibrillarin	0.876	**0.880**	0.005	0.021317
sec61_beta	0.722	**0.738**	0.015	0.029856
myosin_iib	0.481	**0.504**	0.023	0.002878
dna	0.626	**0.638**	0.012	0.004611
alpha_tubulin	0.800	**0.806**	0.007	0.051277
lamin_b1	0.843	**0.853**	0.009	4.04E-05
beta_actin	0.758	**0.765**	0.007	0.05905
dic_lamin_b1	0.645	**0.648**	0.003	0.01688
membrane_caax_63x	0.699	**0.711**	0.012	3.04E-06
zo1	0.460	**0.486**	0.026	0.023023
tom20	0.703	**0.719**	0.015	3.40E-07
st6gal1	0.200	**0.214**	0.014	0.003901

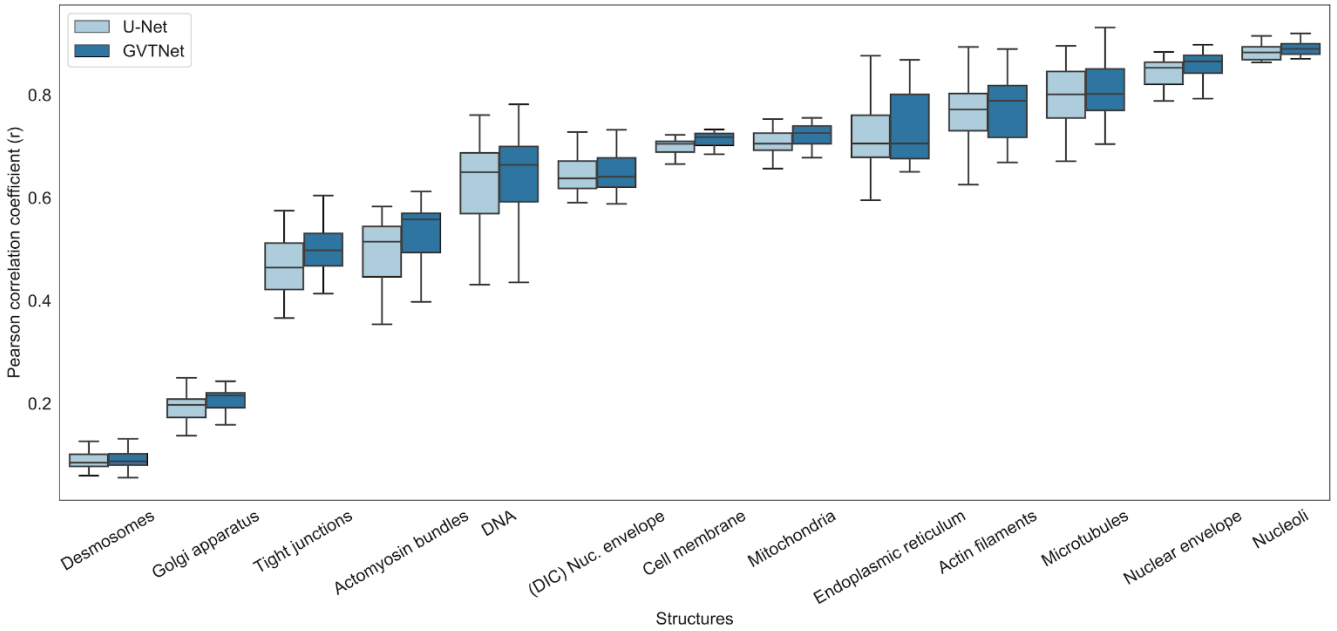
** Shown are **pearson correlations** averaged on 20 test samples per dataset*

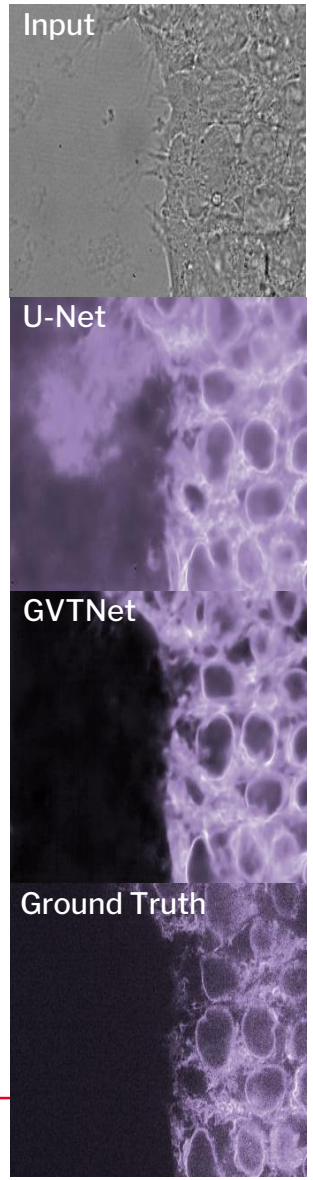
Results: Label-Free Prediction

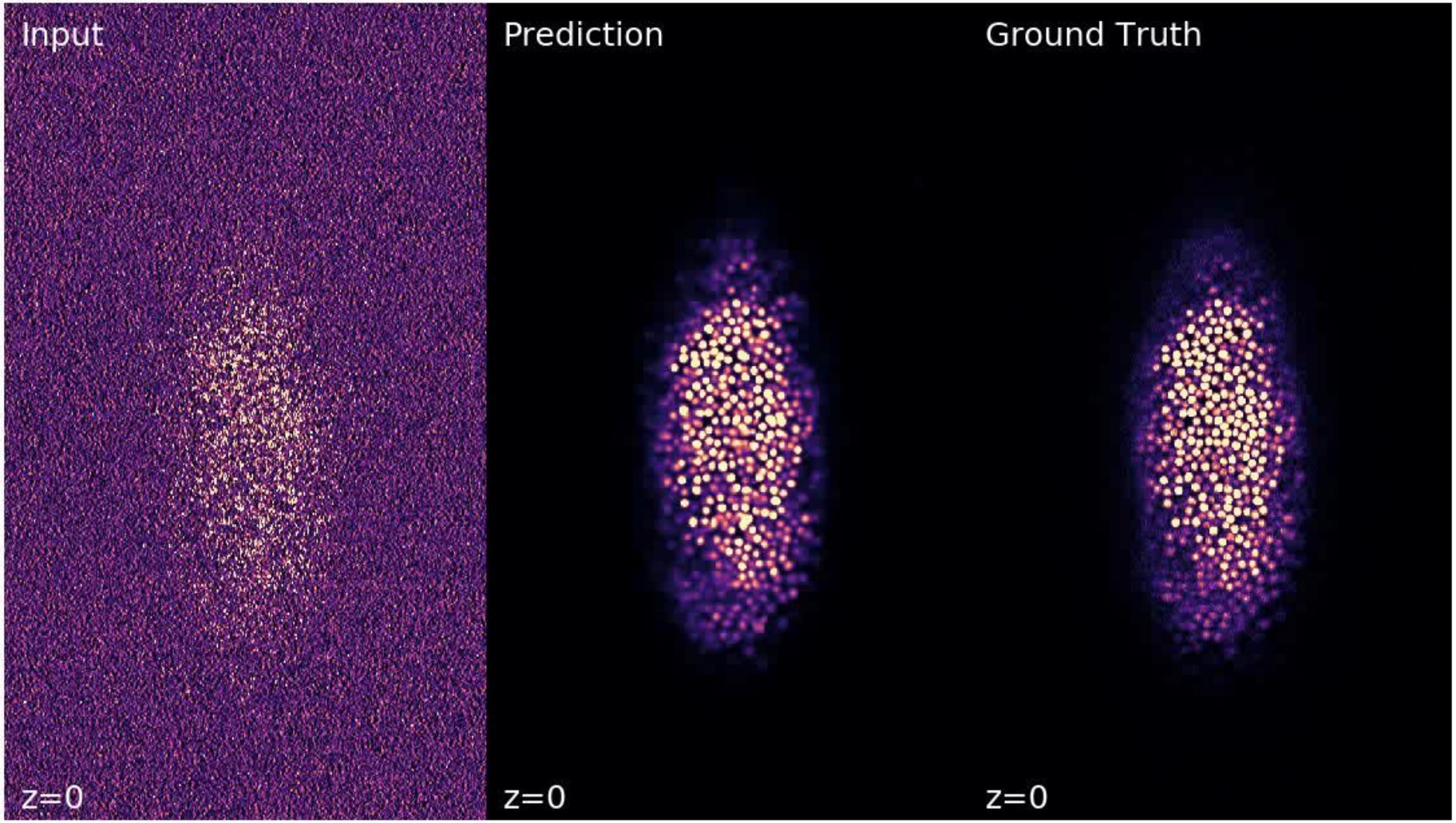
TEXAS A&M UNIVERSITY
Engineering

- Label-free prediction

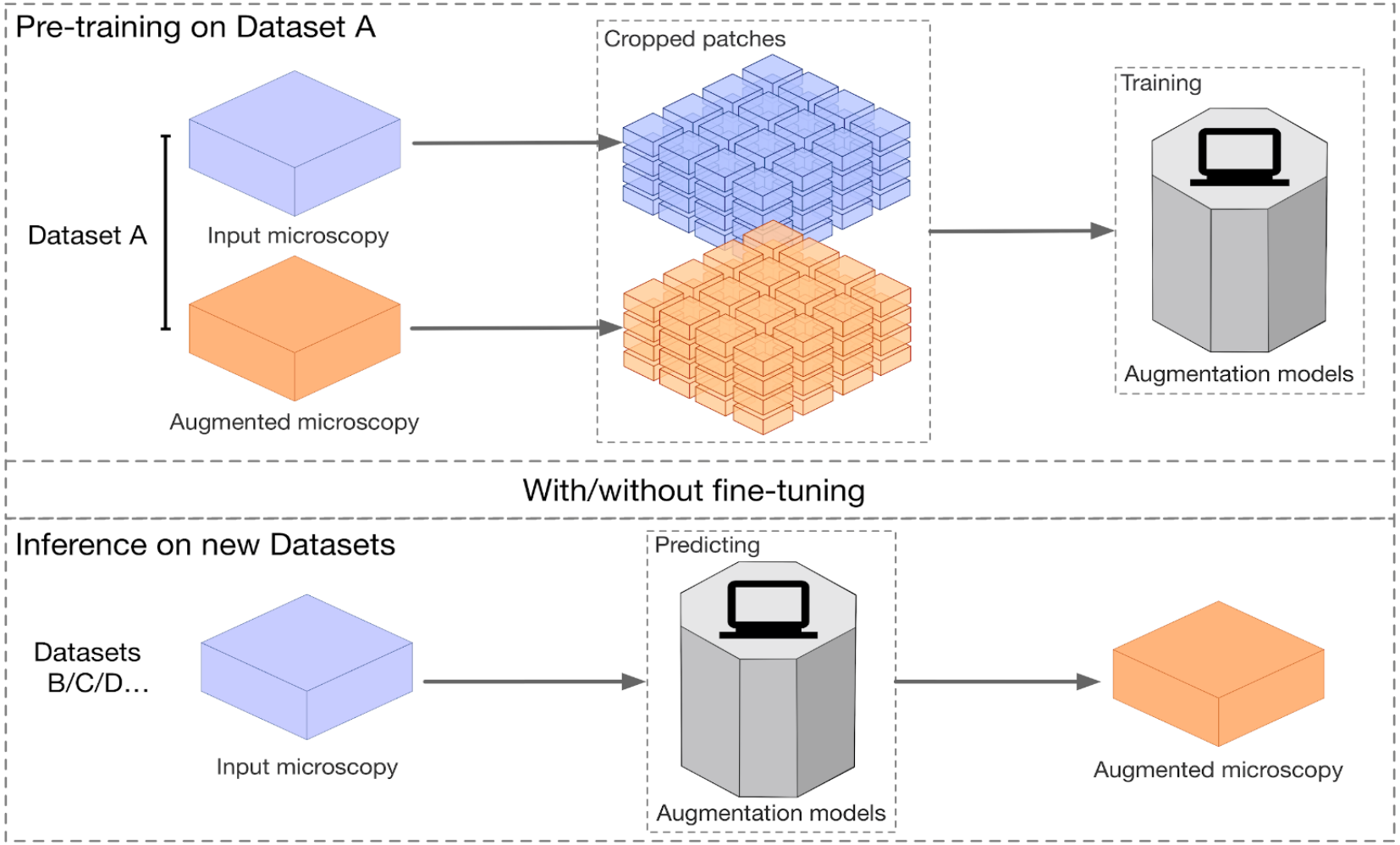
  Ounkomol, Chawin, et al. "Label-free prediction of three-dimensional fluorescence images from transmitted-light microscopy." Nature methods 15.11 (2018): 917.






Input
Prediction
Ground Truth
z=0
z=0
z=0

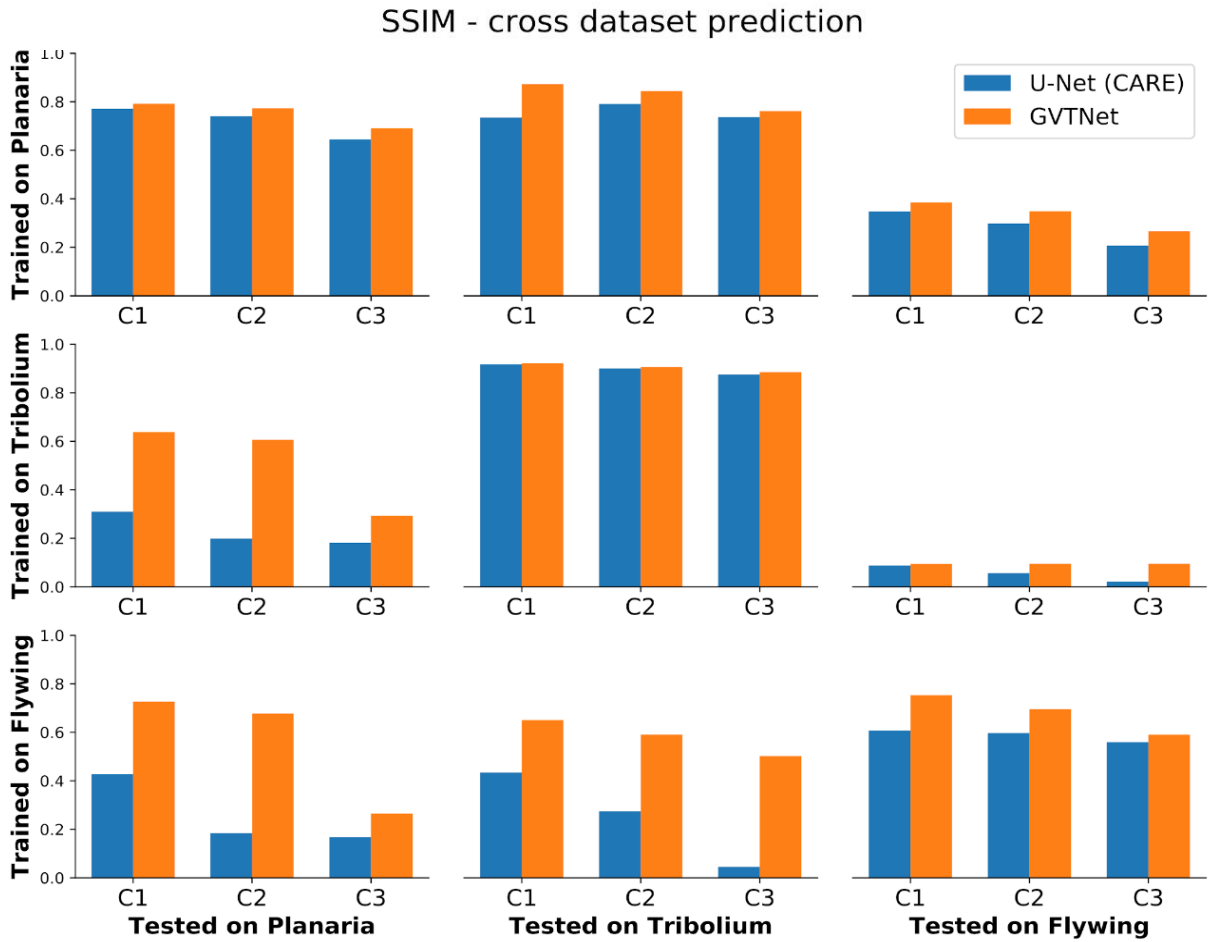
Transfer Learning



Results: Transfer Learning

TEXAS A&M UNIVERSITY
Engineering

- Transfer Learning - without fine-tuning



Self-Supervised Transformations

- Supervised image transformation requires matched pairs of training images, which might be hard to obtain
- Can we do unsupervised learning?
- We use self-supervised approaches
- Self-supervised learning is a type of unsupervised learning approach that predict one part of input from other parts

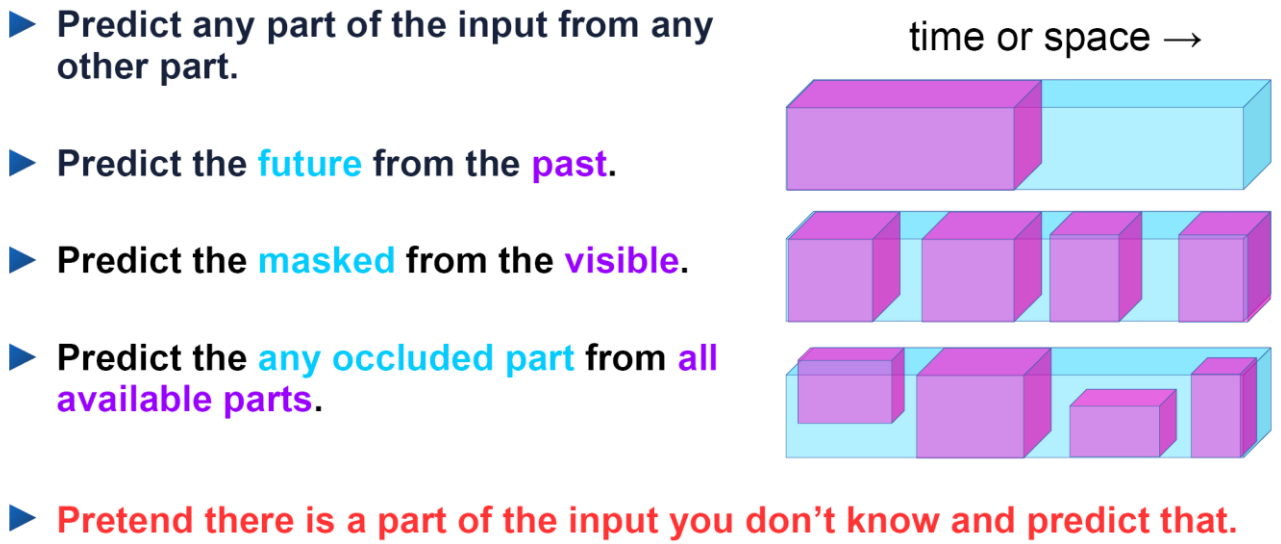


Figure: LeCun

Noise2Same

- We derive a **self-supervised upper bound** of the supervised loss

Theorem 1. *Consider a normalized noisy image $\boldsymbol{x} \in \mathbb{R}^m$ (obtained by subtracting the mean and dividing by the standard deviation) and its ground truth signal $\boldsymbol{y} \in \mathbb{R}^m$. Assume the noise is zero-mean and i.i.d among all the dimensions, and let J be a subset of m dimensions uniformly sampled from the image $\boldsymbol{x}$. For any $f : \mathbb{R}^m \to \mathbb{R}^m$, we have*

$$\mathbb{E}_{x,y} \left\| f(\boldsymbol{x}) - \boldsymbol{y} \right\|^2 + \left\| \boldsymbol{x} - \boldsymbol{y} \right\|^2 \leq \mathbb{E}_x \left\| f(\boldsymbol{x}) - \boldsymbol{x} \right\|^2 + 2m \, \mathbb{E}_J \left[\frac{\mathbb{E}_x \left\| f(\boldsymbol{x})_J - f(\boldsymbol{x}_{J^c})_J \right\|^2}{|J|} \right]^{1/2} \quad (6)$$

Supervised loss

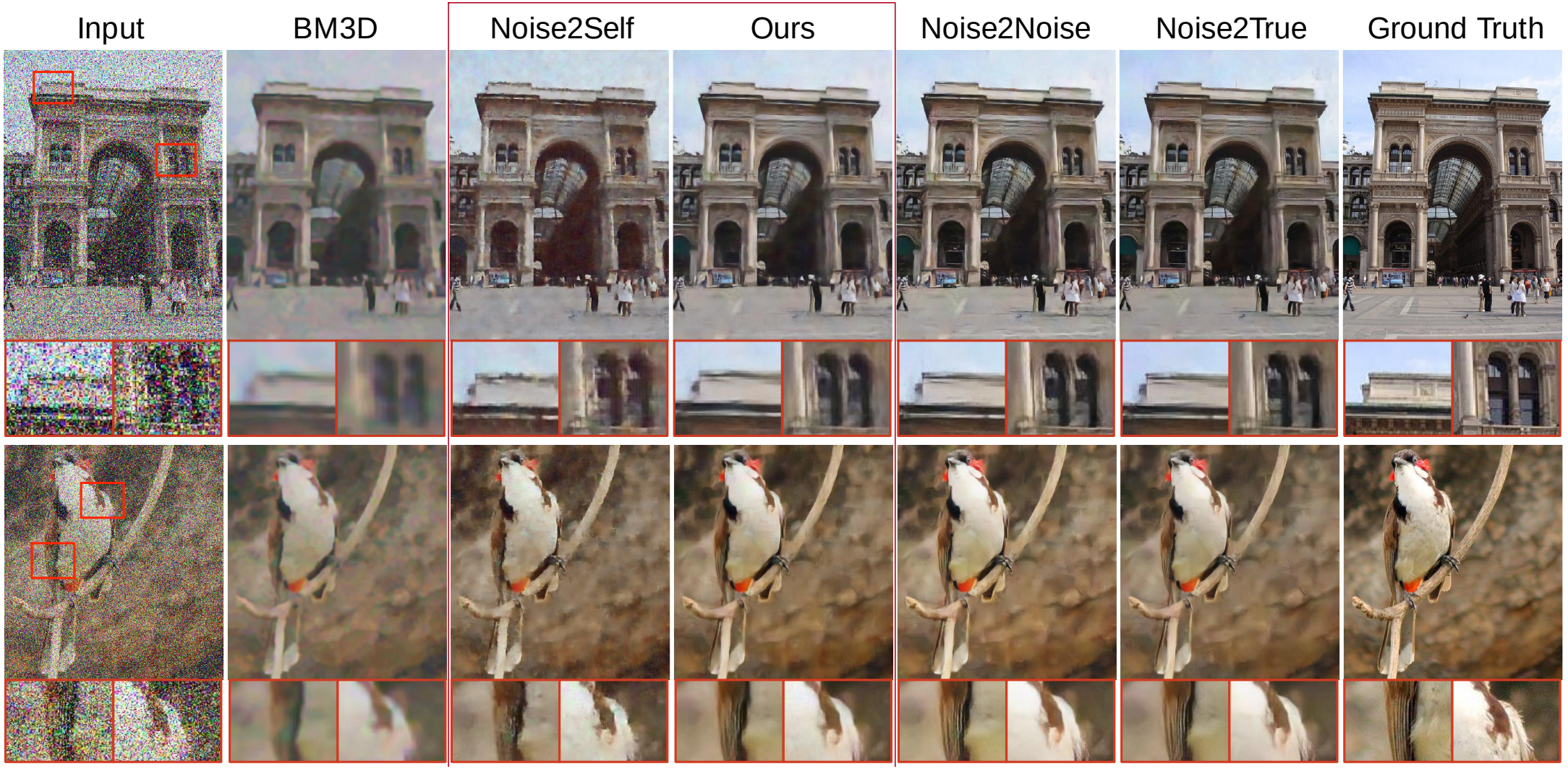
The self-supervised bound as the new loss

Yaochen Xie, Zhengyang Wang, and Shuiwang Ji: Noise2Same: Optimizing A Self-Supervised Bound for Image Denoising. Neural Information Processing Systems (NeurIPS), 2020

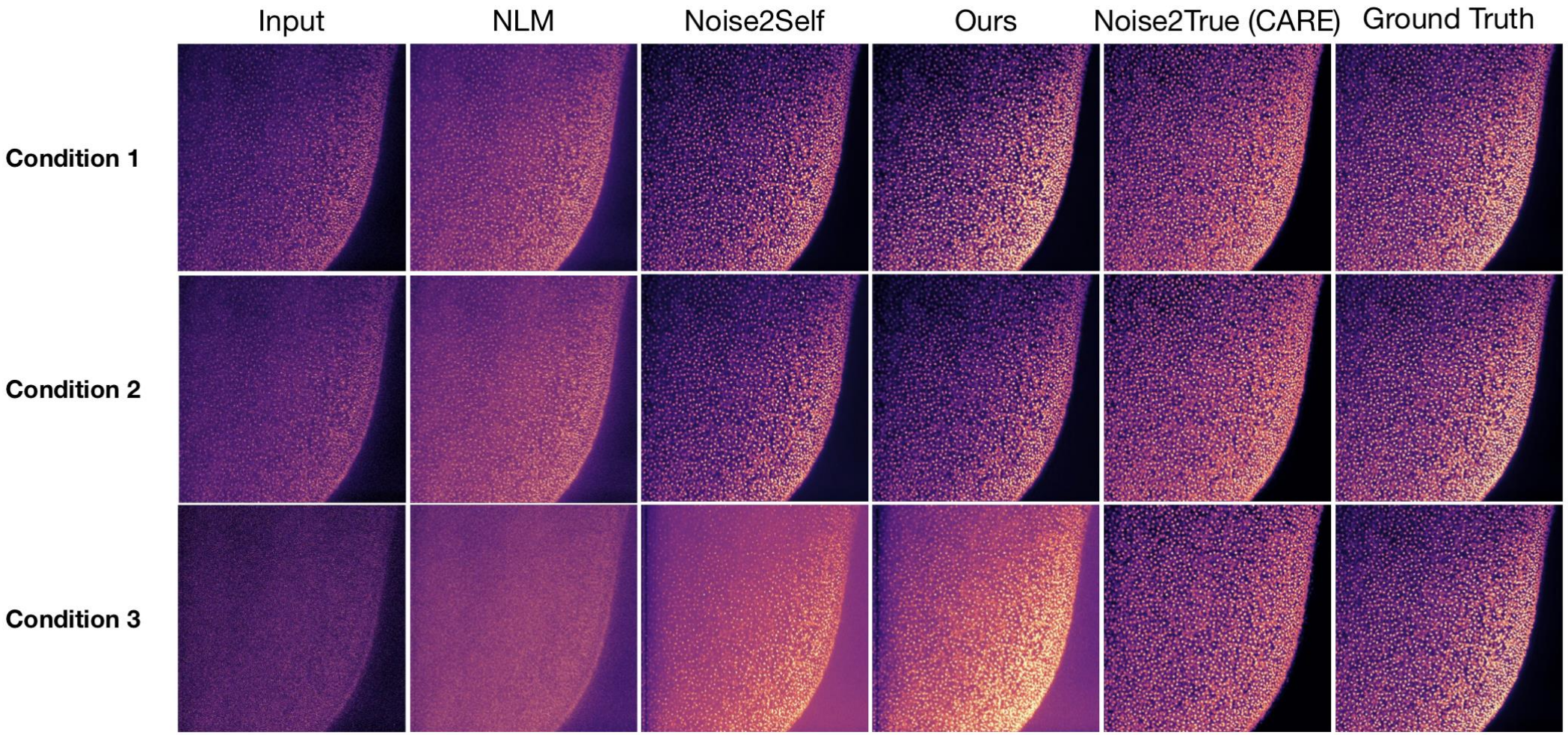
Results: Noise2Same

	Methods	Datasets			
		ImageNet	HànZì	Planaria	BSD68
Traditional	Input	9.69	6.45	21.52 / 21.09 / 20.82	20.19
	NLM [3]	18.04	8.41	25.80 / 24.03 / 21.62	22.73
	BM3D [5]	18.74	10.90	-	28.59
Supervised	Noise2True	23.39	15.66	31.57 / 30.15 / 28.13	29.06
	Noise2Noise [13]	23.27	14.30	-	28.86
Self-Supervised + noise model	Laine et al. [12]	-	-	-	28.84
Self-Supervised	Laine et al. [12]	20.89	10.70	-	27.15
	Noise2Void [10]	21.36	13.72	25.84 / 23.57 / 21.60	27.71
	Noise2Self-Noise [1]	20.38	13.94	27.58 / 24.83 / 21.83	26.98
	Noise2Self-Donut [1]	8.62	13.29	27.63 / 24.72 / 21.73	**28.20**
	Noise2Same	**22.26**	**14.38**	**29.48 / 26.93 / 22.41**	27.95

Results: Noise2Same



Results: Noise2Same

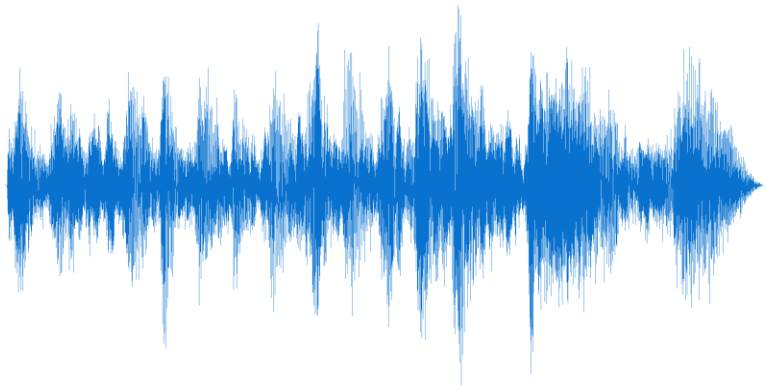


Deep Learning on Grid-Like Data

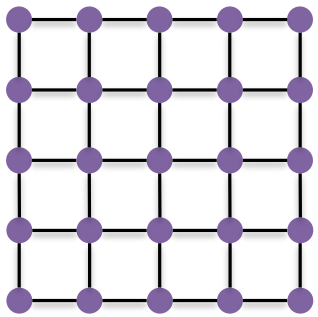
Image

"Hello world!"

Text



Audio

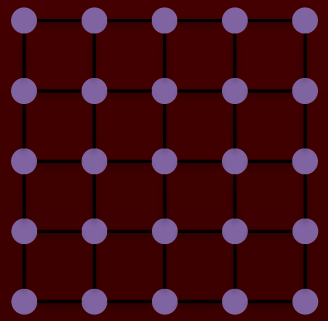


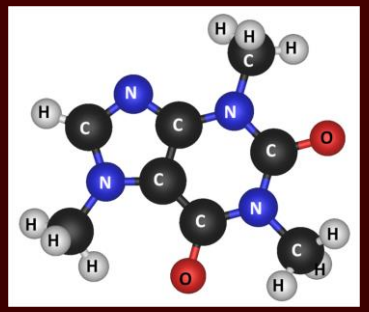
Grid-like Data!

Pictures from Google Images

TEXAS A&M UNIVERSITY
Engineering

From Images to Graphs



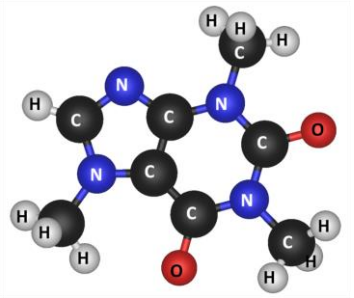


Ubiquitous Graph Data

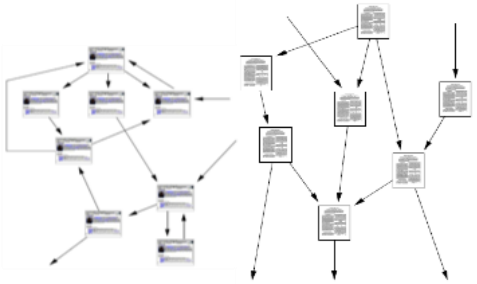
TEXAS A&M UNIVERSITY Engineering

Objects + Relationships = Graph

Social networks



Drug/Material molecules



Citation networks

Brain networks

.....

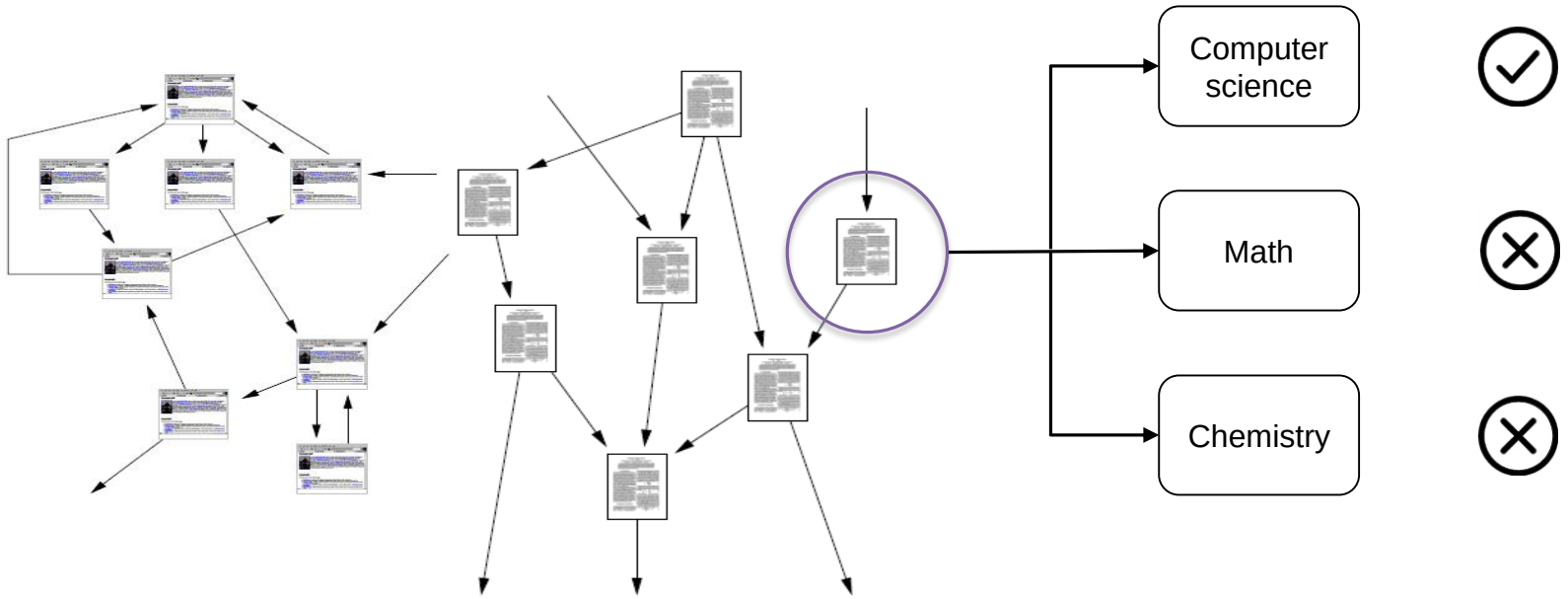
Photos from Google Images

Tasks on Graph Data

- Node classification/regression
- Graph classification/regression
- Link prediction
- Community detection
- …

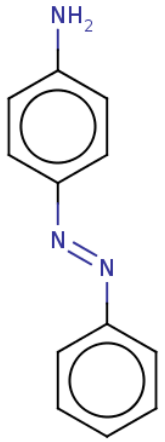
Tasks on Graph Data

- e.g., Node classification
  - Citation Networks

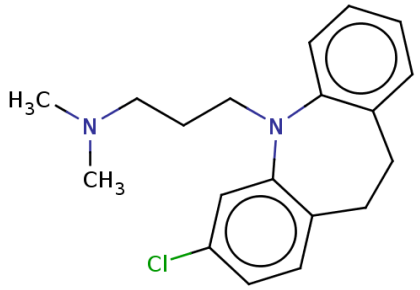


Tasks on Graph Data

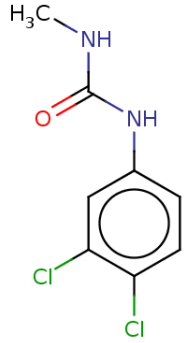
- e.g., Graph classification
  - Molecular property prediction



Toxic



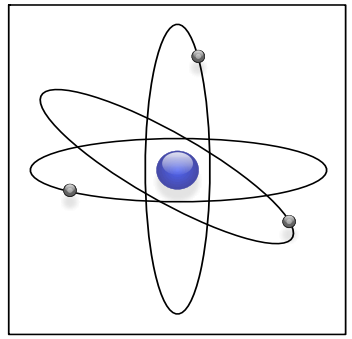
Non-toxic



?

Molecular Property Prediction

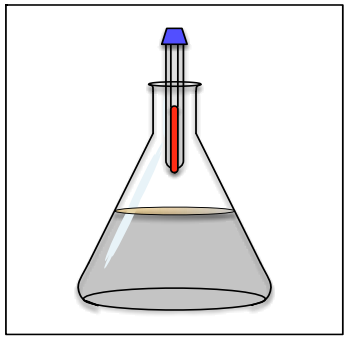
TEXAS A&M UNIVERSITY Engineering



Quantum Mechanics

- *QM8*
- *QM9*

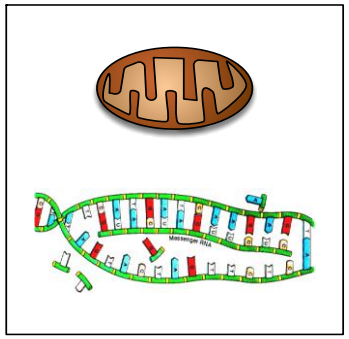
Geometric, energetic, electronic, thermodynamic properties



Physical Chemistry

- *ESOL*
- *Lipophilicity*
- *FreeSolv*

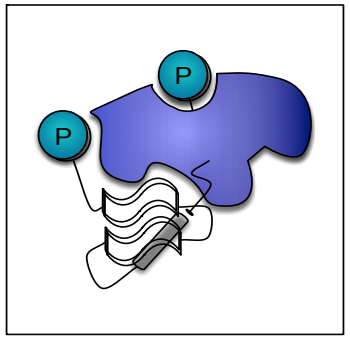
Water solubility, hydration free energy, lipophilicity



Biophysics

- *HIV*
- *BACE*
- *PCBA*
- *MUV*

Bioactivity, binding affinities



Physiology

- *BBBP*
- *Tox21*
- *ToxCast*
- *SIDER*
- *ClinTox*

Barrier permeability, toxicology, adverse reactions

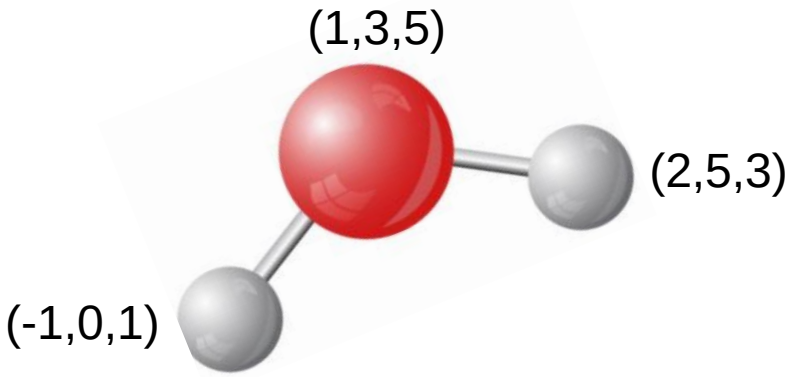
Wu, Zhenqin, et al. "MoleculeNet: a benchmark for molecular machine learning." Chemical science 9.2 (2018): 513-530.

3D Graphs

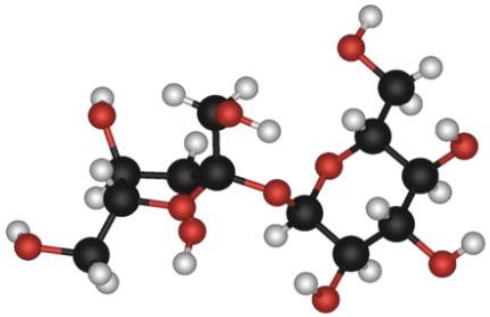
TEXAS A&M UNIVERSITY Engineering

3D graphs

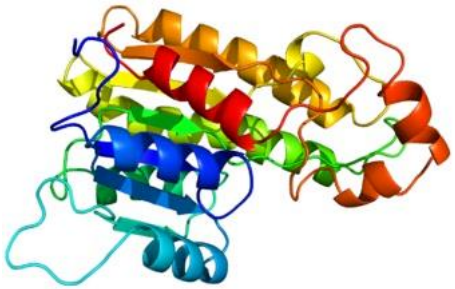
- Graph structure
- Node-level positional info



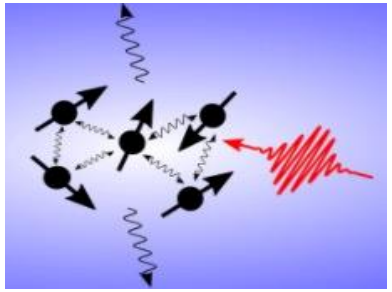
Important for a lot of real-world applications



Molecular modeling



Protein structural representation



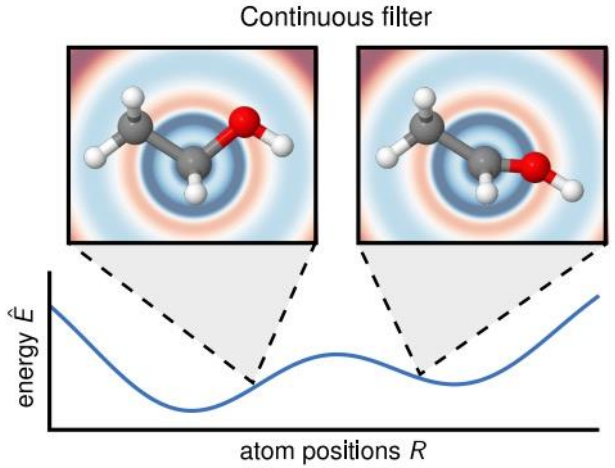
Quantum system simulation

Photos from Google Images

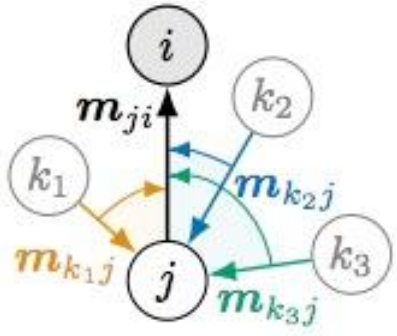
Research Obstacles

TEXAS A&M UNIVERSITY Engineering

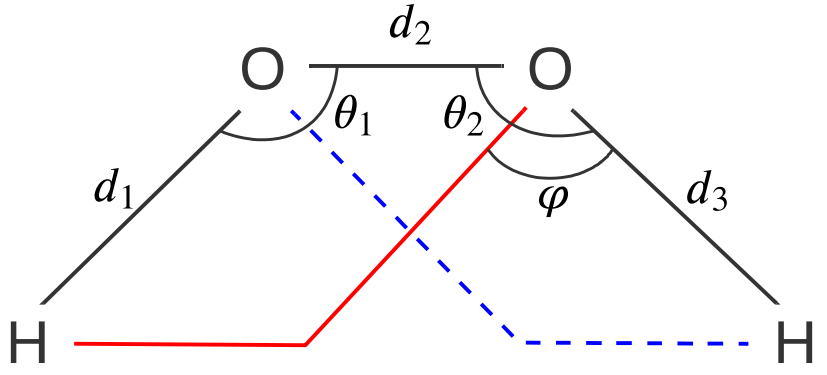
- No unified framework for 3D graphs
- Structural representation is not complete



SchNet: distance



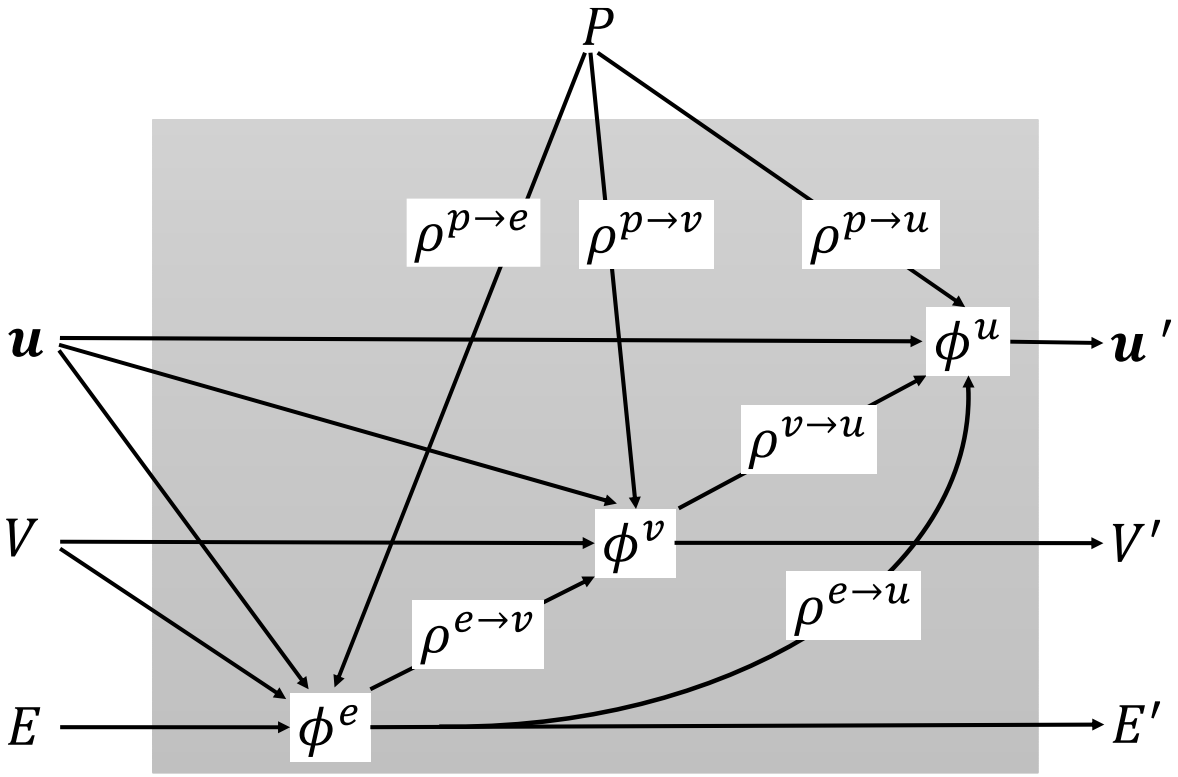
DimeNet: distance & angle



Only distance & angle is not enough!

Schütt, K. et al. Schnet: A continuous-filter convolutional neural network for modeling quantum interactions. In NIPS, pp. 991–1001, 2017.
Klicpera, J. et al. Directional message passing for molecular graphs. In ICLR, 2020.

A Unified Framework: 3DGN



$$\mathbf{e}'_k = \phi^e\left(\mathbf{e}_k, \mathbf{v}_{r_k}, \mathbf{v}_{s_k}, E_{s_k}, \mathbf{u}, \rho^{p\to e}\left(\{\mathbf{r}_h\}_{h=r_k\cup s_k\cup\mathcal{N}_{s_k}}\right)\right),$$
$$\mathbf{v}'_i = \phi^v\left(\mathbf{v}_i, \rho^{e\to v}\left(E_i\right), \mathbf{u}, \rho^{p\to v}\left(\{\mathbf{r}_h\}_{h=i\cup\mathcal{N}_i}\right)\right),$$
$$\mathbf{u}' = \phi^u\left(\rho^{e\to u}\left(E'\right), \rho^{v\to u}\left(V'\right), \mathbf{u}, \rho^{p\to u}\left(\{\mathbf{r}_h\}_{h=1:n}\right)\right).$$

- A generic and unifying framework for 3D graphs
- An interface for manipulating 3D graphs at different levels of granularity

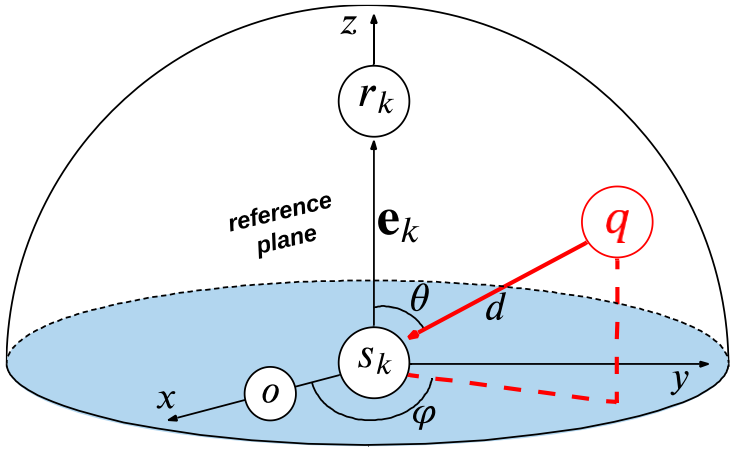
Liu, Yi, et al. "Spherical Message Passing for 3D Graph Networks." arXiv preprint arXiv:2102.05013 (2021).
Peter W. Battaglia et al.: Relational inductive biases, deep learning, and graph networks. https://arxiv.org/abs/1806.01261

A Novel MP Method: SMP

TEXAS A&M UNIVERSITY Engineering

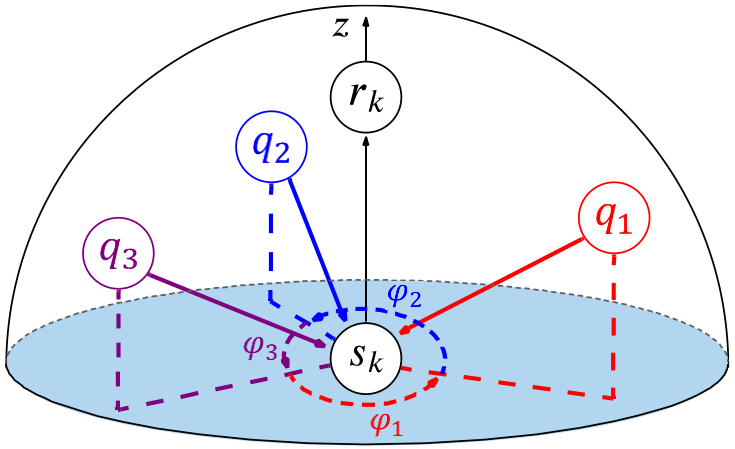
Spherical message passing

- ✓ An accurate and complete architecture for 3DGN
- ✓ Existing methods are special cases



(a)

SMP considers distance, angle, and torsion



(b)

Compute torsion

Liu, Yi, et al. "Spherical Message Passing for 3D Graph Networks." arXiv preprint arXiv:2102.05013 (2021).

Results: SphereNet

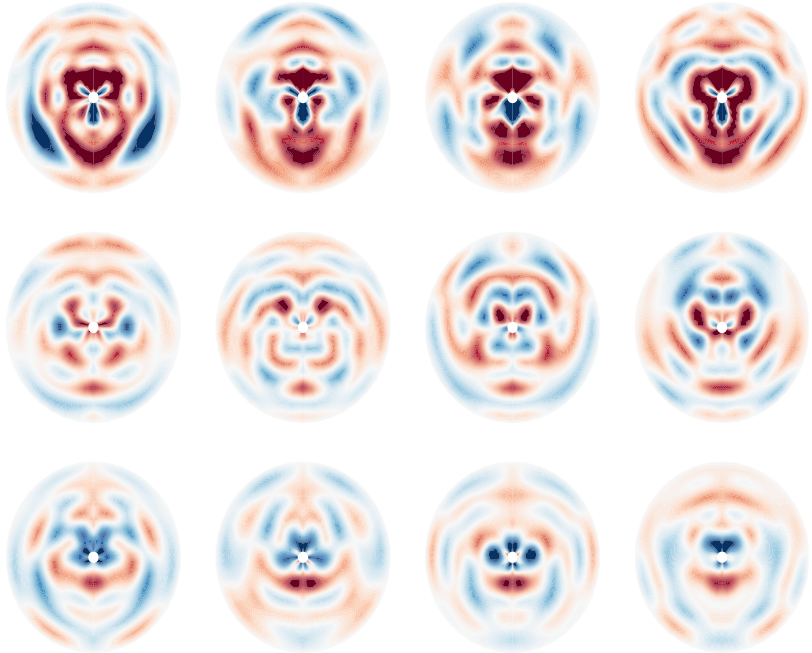
TEXAS A&M UNIVERSITY
Engineering

QM9

Property	Unit	PPGN	SchNet	PhysNet	Cormorant	MGCN	DimeNet	DimeNet++	SphereNet
μ	D	0.047	0.033	0.0529	0.13	0.0560	0.0286	0.0297	**0.0269**
α	$a_0{}^3$	0.131	0.235	0.0615	0.092	**0.0300**	0.0469	0.0435	0.0465
ϵ_{HOMO}	meV	40.3	41	32.9	36	42.1	27.8	24.6	**23.6**
ϵ_{LUMO}	meV	32.7	34	24.7	36	57.4	19.7	19.5	**18.9**
$\Delta\epsilon$	meV	60.0	63	42.5	60	64.2	34.8	32.6	**32.3**
$\langle R^2 \rangle$	$a_0{}^2$	0.592	**0.073**	0.765	0.673	0.110	0.331	0.331	0.292
ZPVE	meV	3.12	1.7	1.39	1.98	**1.12**	1.29	1.21	**1.12**
U_0	meV	36.8	14	8.15	28	12.9	8.02	6.32	**6.26**
U	meV	36.8	19	8.34	-	14.4	7.89	**6.28**	7.33
H	meV	36.3	14	8.42	-	14.6	8.11	6.53	**6.40**
G	meV	36.4	14	9.40	-	16.2	8.98	**7.56**	8.0
c_v	$\frac{\text{cal}}{\text{mol K}}$	0.055	0.033	0.0280	0.031	0.0380	0.0249	0.0230	**0.0215**
std. MAE	%	1.84	1.76	1.37	2.14	1.86	1.05	0.98	**0.94**

OC20

	Energy MAE [eV] ↓				EwT ↑			
Model	ID	OOD Ads	OOD Cat	OOD Both	ID	OOD Ads	OOD Cat	OOD Both
CGCNN	1.0479	1.0527	1.0232	0.9608	1.39%	1.38%	1.59%	1.57%
SchNet	1.0858	1.1044	1.0720	1.0391	1.34%	1.39%	1.42%	1.44%
DimeNet	1.0117	1.0734	0.9814	0.9767	1.45%	1.41%	1.53%	1.41%
DimeNet++	0.8819	0.9106	0.8357	0.8408	1.94%	1.69%	2.13%	1.84%
SphereNet	**0.8352**	**0.8723**	**0.7959**	**0.7952**	**1.96%**	**2.02%**	**2.19%**	**1.90%**

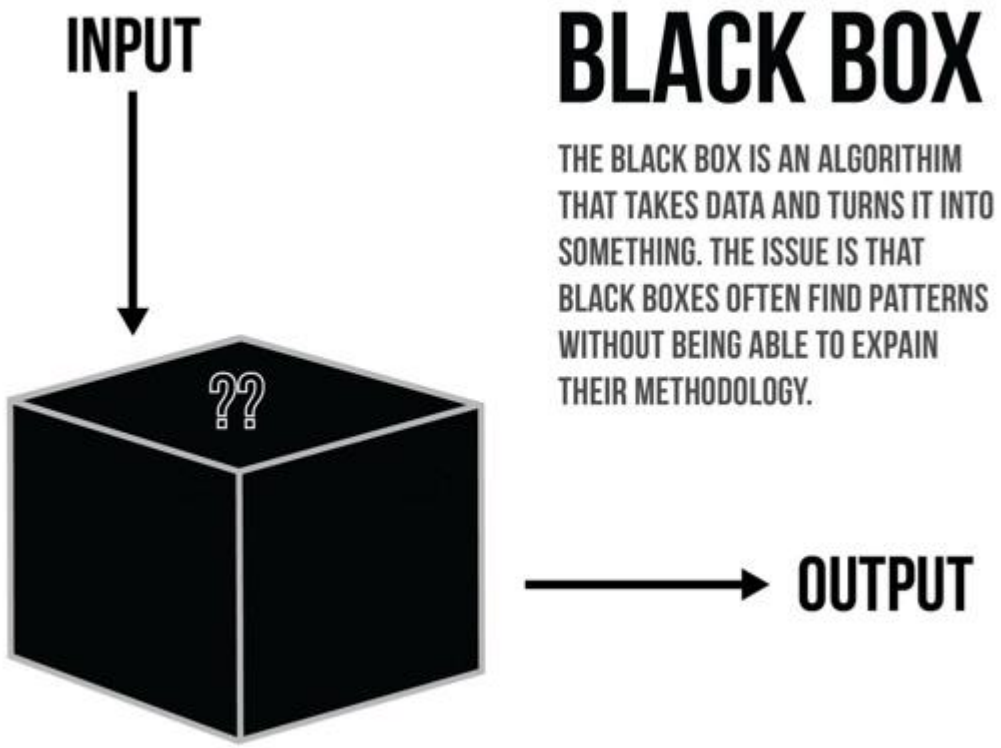


SphereNet filters

All the distance, angle, and torsion make contributions

Neural Nets are Black Boxes

- They cannot be fully trusted!
  - It prevents their use in critical applications pertaining to fairness, privacy, and safety
- We need explanation techniques
  - Explain the relationships between input graph and output predictions

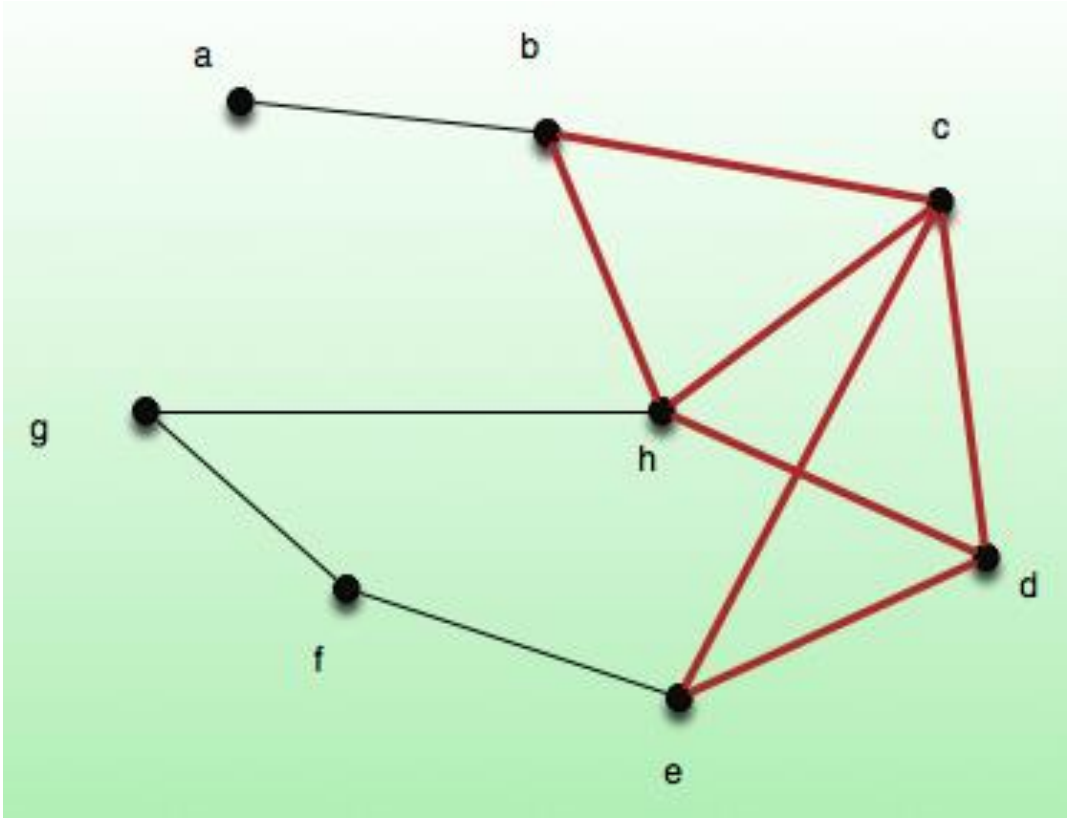


Reference: *https://towardsdatascience.com/guide-to-interpretable-machine-learning-d40e8a64b6cf*

Explaining GNNs

TEXAS A&M UNIVERSITY
Engineering

- Existing methods focus on nodes, edges, and features.
  - Less intuitive, not human-intelligible
- Subgraph-level explanations
  - Subgraphs are highly related to the functionalities of graphs
  - More human-intelligible
  - Example: network motifs, the simple building blocks of complex networks



Reference: *https://en.wikipedia.org/wiki/Dense_subgraph*

The Proposed SubgraphX

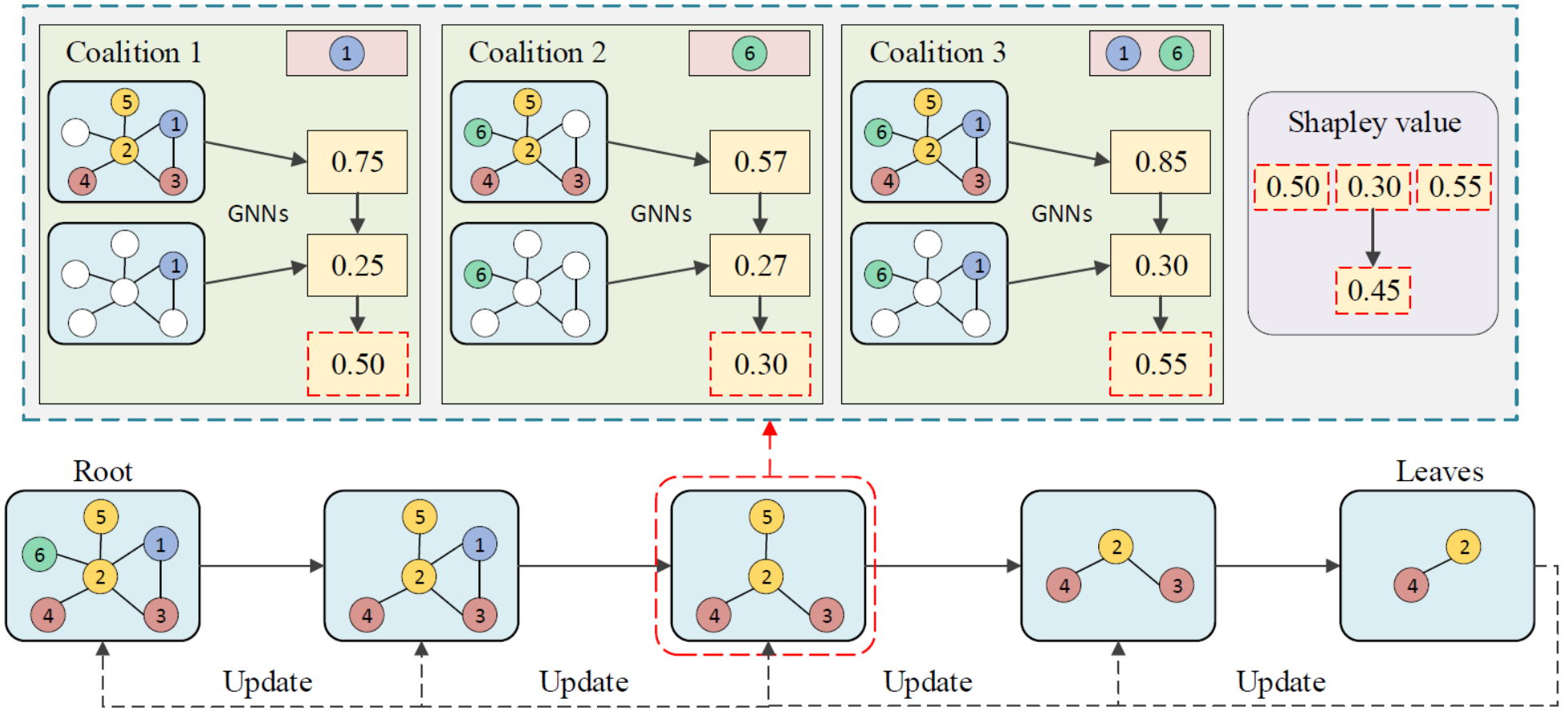
- Explain GNN predictions using subgraphs
  - Find the most important subgraph for a target prediction
- Explore subgraphs with the search algorithm
  - Monte Carlo Tree Search
  - Learn the trade-off between exploration and exploitation
- A Game-Theoretical Scoring Function
  - Shapley values to measure the importance of different subgraphs
  - Efficient approximation schemes to compute Shapley values by considering interactions within the information aggregation range

Hao Yuan….Shuiwang Ji: On Explainability of Graph Neural Networks via Subgraph Explorations. https://arxiv.org/abs/2102.05152

The Proposed SubgraphX

TEXAS A&M UNIVERSITY
Engineering

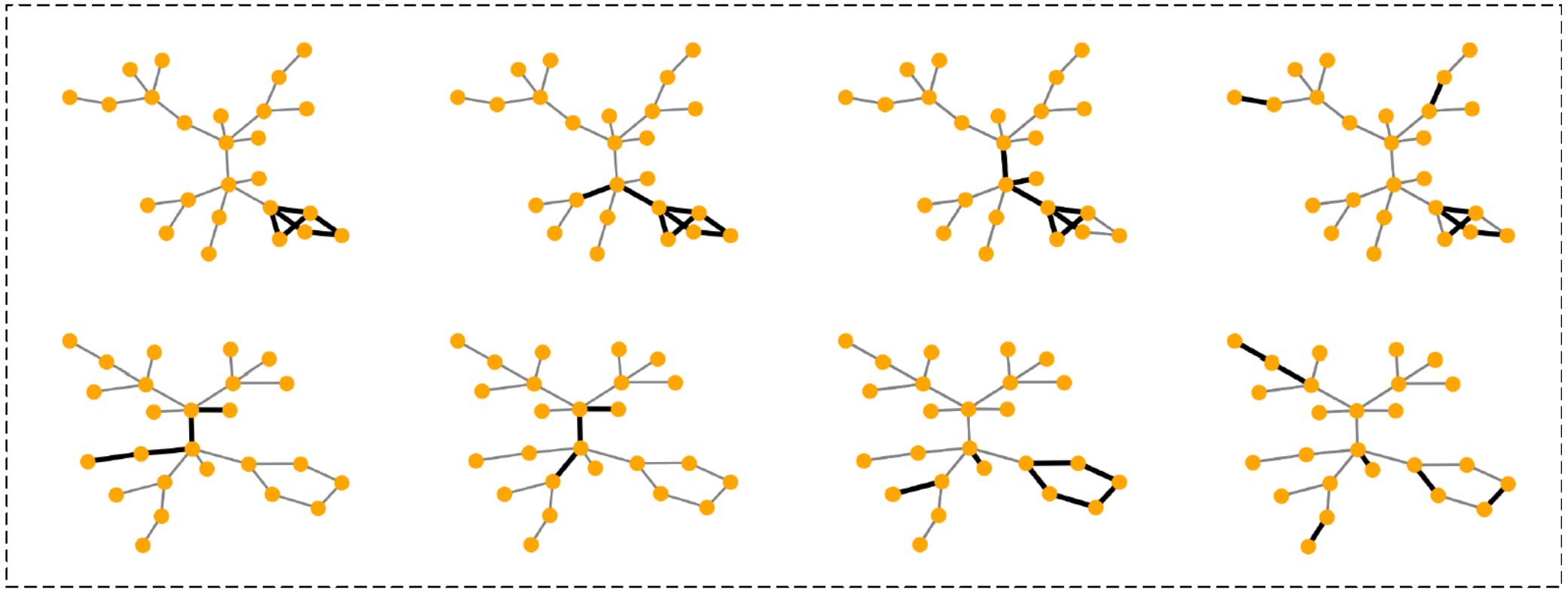
- Shapley values can capture the interactions in GNNs!



Results: SubgraphX

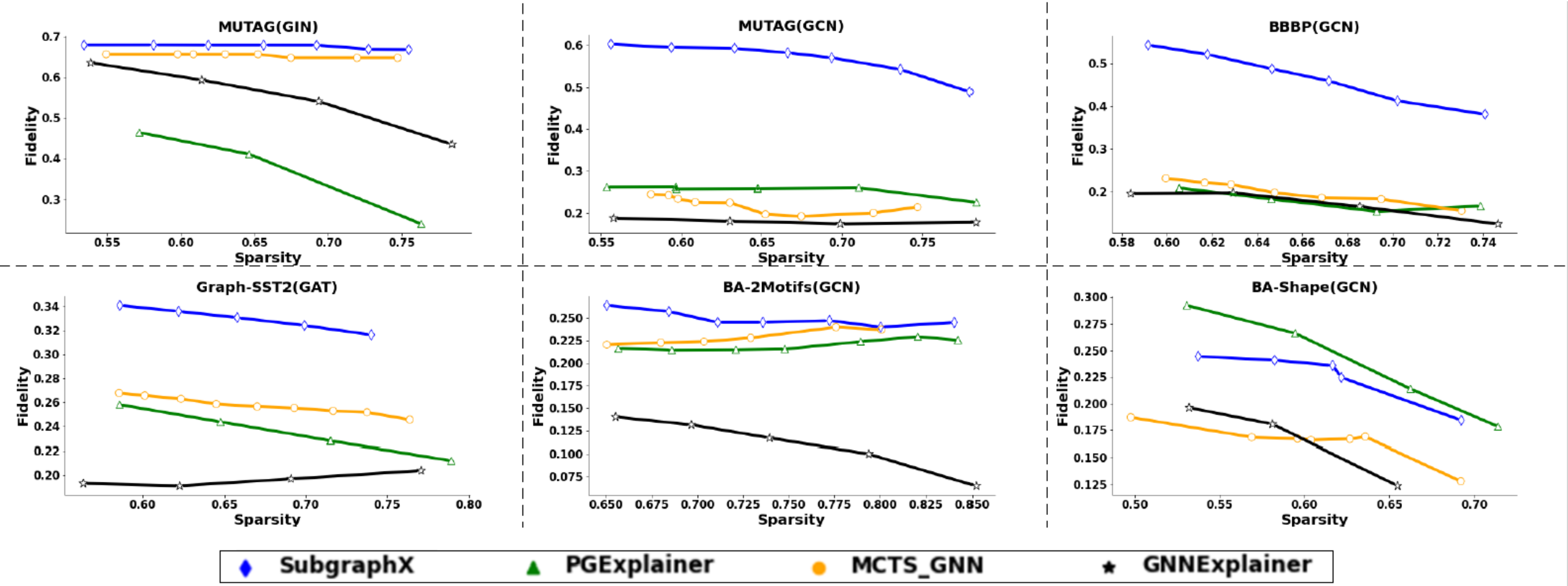
TEXAS A&M UNIVERSITY
Engineering

- Data are labeled by the motifs (house-like motif or cycle motif)
- First row: correct prediction; second row: incorrect prediction.



Results: SubgraphX

TEXAS A&M UNIVERSITY
Engineering



Fidelity: the change of prediction by removing important substructures (higher is better)

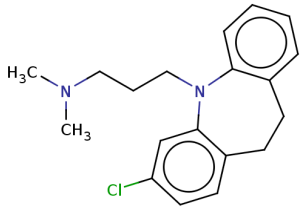
TEXAS A&M UNIVERSITY
Engineering

From Graph Prediction to Generation

Molecular Graph Generation

TEXAS A&M UNIVERSITY Engineering

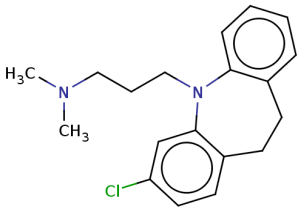
- A fundamental problem in drug discovery is to find novel molecules with desired properties.
  - Search in the chemical space with molecular property prediction.
    - Extremely expensive since chemical space is huge.



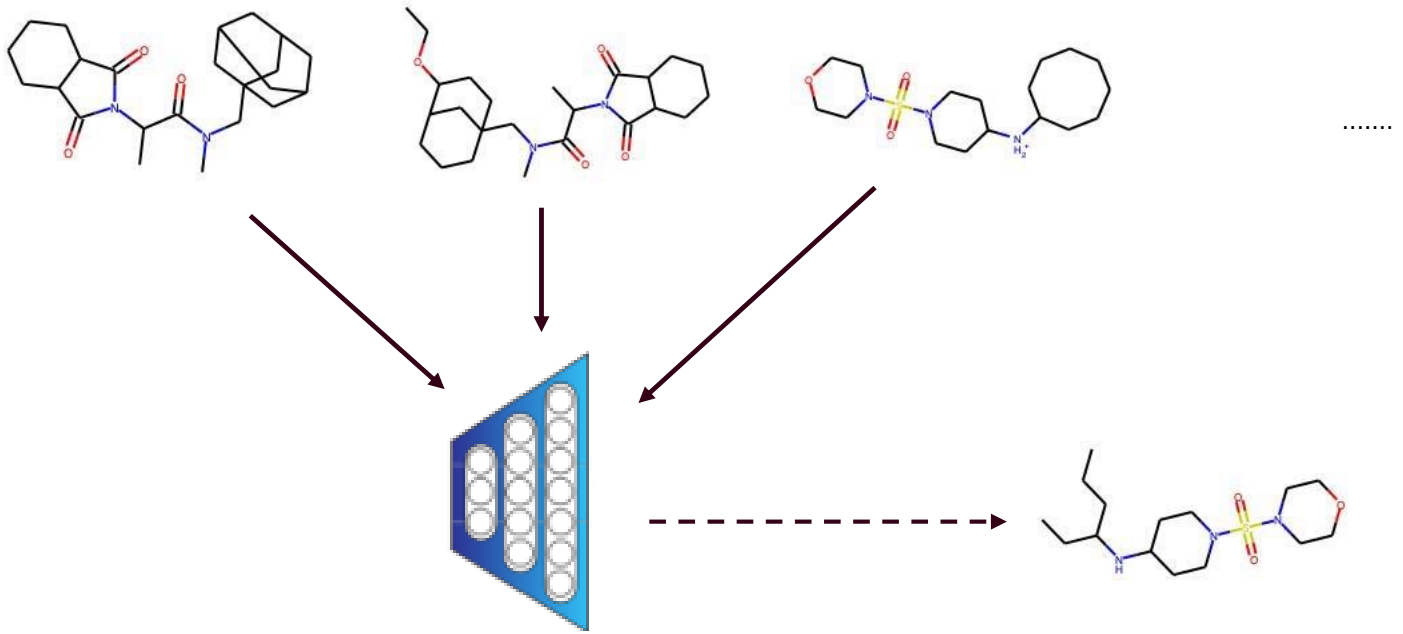
Properties (e.g., Toxicity)

  - Generate molecules with desired properties directly.
    - Circumvent the expensive search.
    - Molecule generation is still in the preliminary stage.

Properties (e.g., Druglikeness)

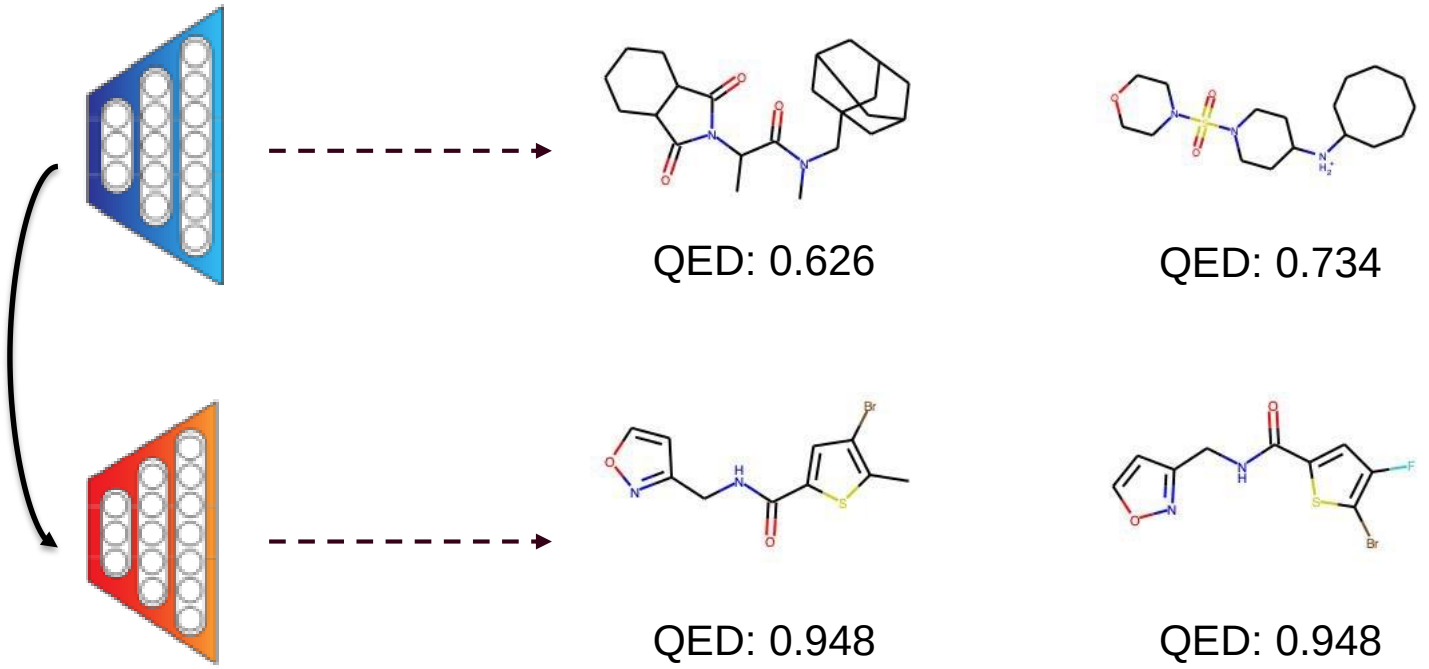


Molecular Graph Generation



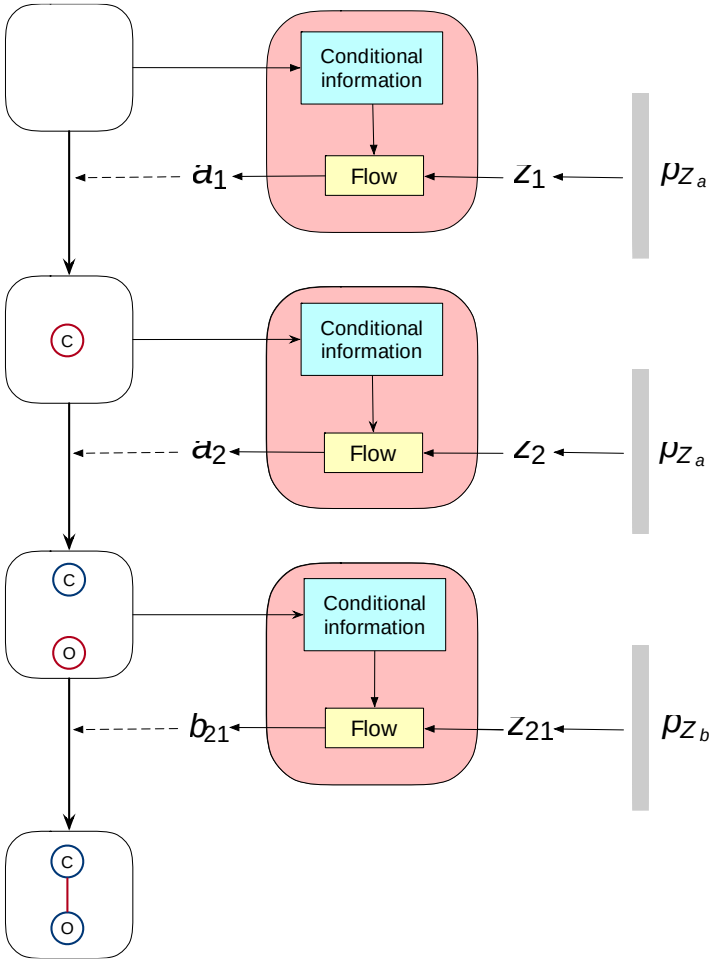
Given a molecule dataset, use it to train a molecule generator which can generate valid molecular graphs.

Property Optimization



Given a pretrained molecule generator, optimize it to generate molecules with high property scores.

A Discrete Flow Model



- Decompose the molecular graph into a sequence of discrete tokens $S_G = (a_1, a_2, b_{21}, a_3, \dots)$.
- Use the composition of multiple invertible modulo shift transforms to convert discrete latent variables z to the next discrete token x based on

$$x = (z + \mu) \bmod t,$$

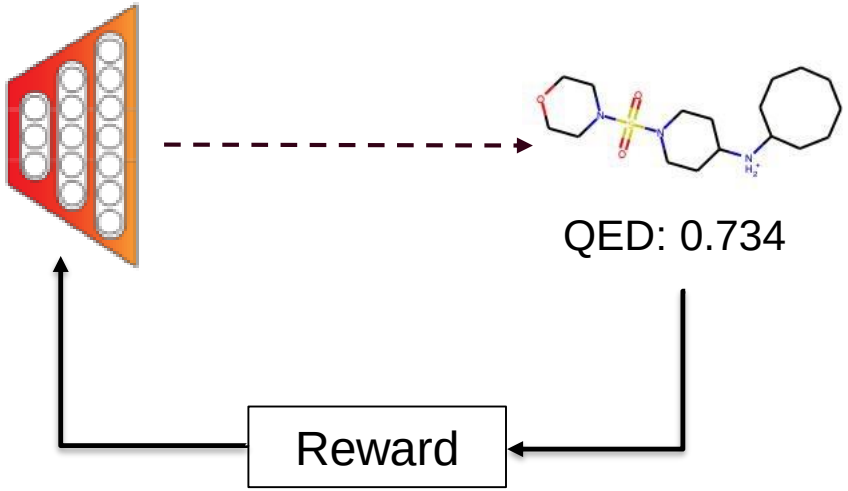
where t is the number of categories of x.

- Conditional information is extracted with Relational Graph Convolutional networks (R-GCN)* and incorporated into μ.

Youzhi Luo, Keqiang Yan, Shuiwang Ji: GraphDF: A Discrete Flow Model for Molecular Graph Generation. https://arxiv.org/abs/2102.01189

Optimize Property with Reinforcement Learning

TEXAS A&M UNIVERSITY Engineering



- Formulate the sequential generation of molecular graph as a Markov Decision Process,
  - state --- currently unfinished sub-graph,
  - action --- new node or edge addition,
  - reward --- property score of the finished molecular graph.
- Start from a model pretrained on a molecule dataset, then fine-tune the model with reinforcement learning by Proximal Policy Optimization (PPO)* algorithm.

Youzhi Luo, Keqiang Yan, Shuiwang Ji: GraphDF: A Discrete Flow Model for Molecular Graph Generation. https://arxiv.org/abs/2102.01189

Energy-Based Models

TEXAS A&M UNIVERSITY Engineering

- Energy-based models (EBMs)
  - $E_\theta(x): X \to R$ is the energy function, e.g., a deep neural network.
  - The probability distribution given by an EBM is:

$$p_\theta(x) = \frac{e^{-E_\theta(x)}}{Z_\theta}.$$

  - $Z_\theta = \int e^{-E_\theta(x)} dx$ for continuous data and $Z_\theta = \sum e^{-E_\theta(x)}$ for discrete data. Z_θ is usually intractable.

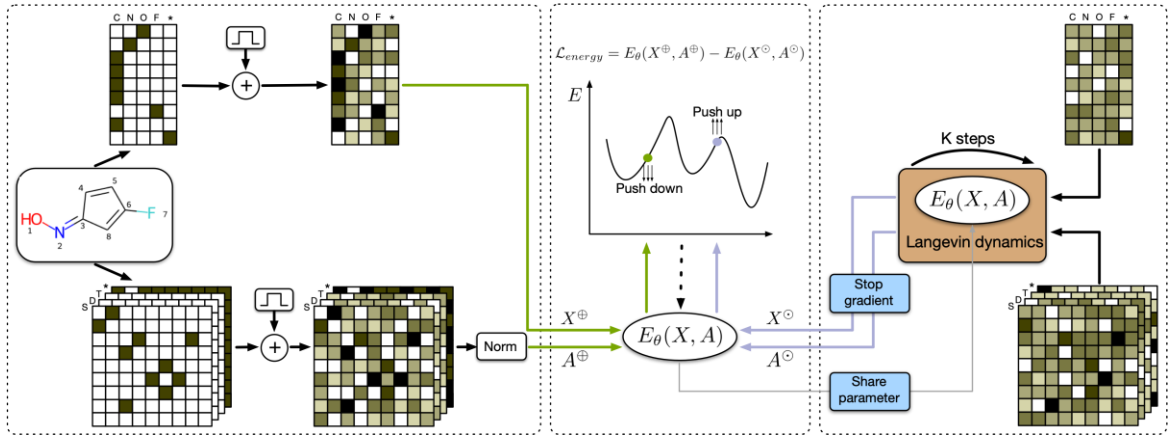
  - EBMs have been used as generative models in various domains, e.g., images and videos.

LeCun, Yann, et al. "A tutorial on energy-based learning." Predicting structured data 1.0, 2006.
Song, Yang, and Diederik P. Kingma. "How to Train Your Energy-Based Models." Preprint, 2021.
Du, Yilun, et al. "Improved Contrastive Divergence Training of Energy Based Models." Neural Information Processing Systems, 2019.

GraphEBM

TEXAS A&M UNIVERSITY
Engineering

- Energy function is parameterized as a permutation invariant graph neural network.
- Maximum likelihood training with MCMC (e.g., Langevin dynamics).



$$loss = E_\theta\left(X^\oplus, A^\oplus\right) - E_\theta\left(X^\odot, A^\odot\right)$$

- Goal-directed: push down energies with flexible degrees $f(y)$ based on property values y

$$loss = f(y)E_\theta\left(X^\oplus, A^\oplus\right) - E_\theta\left(X^\odot, A^\odot\right).$$

- Multi-objective: combine multiple trained energy functions associated with different properties

$$E_{\theta^*}(X, A) = E_{\theta_1^*}(X, A) + E_{\theta_2^*}(X, A).$$

- Generating with MCMC sampling.

Meng Liu, Keqiang Yan, Bora Oztekin, Shuiwang Ji: GraphEBM: Molecular Graph Generation with Energy-Based Models. https://arxiv.org/abs/2102.00546

Results: GraphDF

TEXAS A&M UNIVERSITY
Engineering

Method	Validity %	Validity w/o check %	Uniqueness %	Novelty %	Reconstruct %
JT-VAE	100	n/a	100	100	76.7
GCPN	100	20	99.97	100	n/a
MoFlow	100	81.76	**99.99**	100	100
GraphCNF	96.35	n/a	99.98	99.98	100
GraphAF	100	68	99.1	100	100
GraphDF (ours)	**100**	**89.03**	99.16	**100**	**100**

Table 1 Molecule generation performance on ZINC 250K dataset. GraphDF can model the underlying chemical rule more accurately and achieve much higher validity w/o check rate.

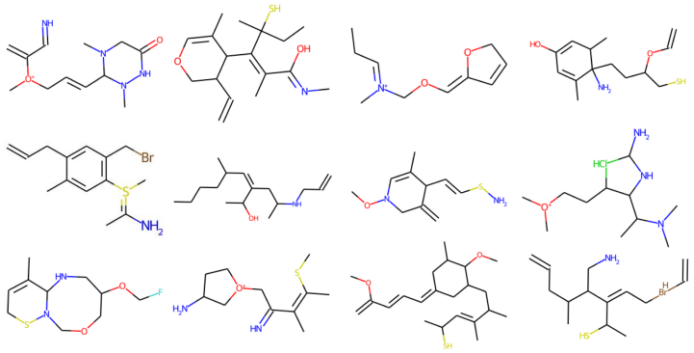
Method	Penalized logP			QED		
	1st	2nd	3rd	1st	2nd	3rd
JT-VAE	5.3	4.93	4.49	0.925	0.911	0.91
GCPN	7.98	7.85	7.8	0.948	0.947	0.946
MRNN	8.63	6.08	4.73	0.844	0.796	0.736
MoFlow	n/a	n/a	n/a	0.948	0.948	0.948
GraphAF	12.23	11.29	11.05	0.948	0.948	0.947
GraphDF (ours)	**13.70**	**13.18**	**13.16**	**0.948**	**0.948**	**0.948**

Table 2 Property optimization performance evaluated by top-3 property scores. GraphDF has stronger capacity to search molecules with high property scores.

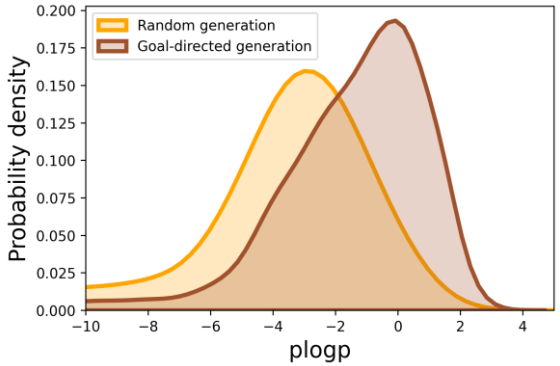
Results: GraphEBM

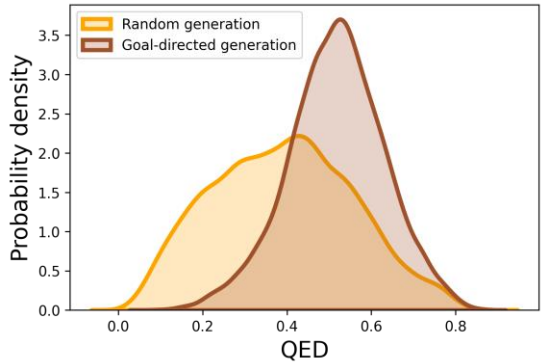
TEXAS A&M UNIVERSITY
Engineering

- Random generation
  - ➢ can generate non-trivial molecules

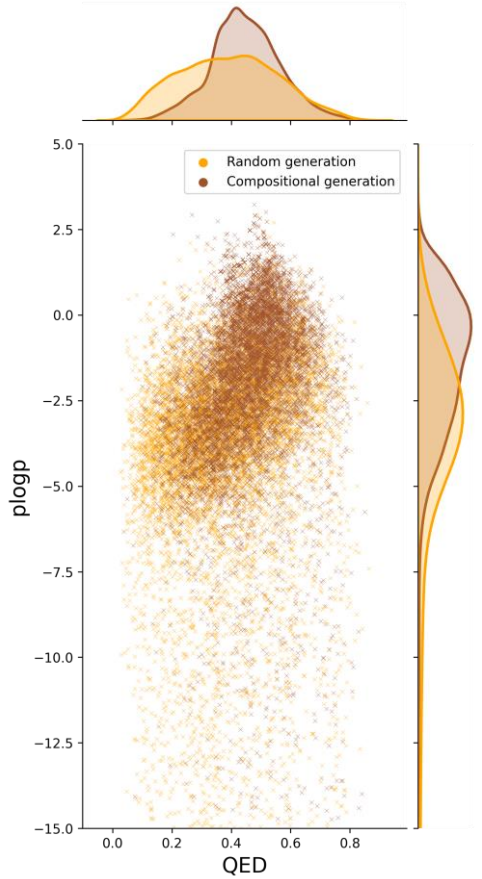


- Goal-directed generation
  - ➢ can improve the property of generated molecules





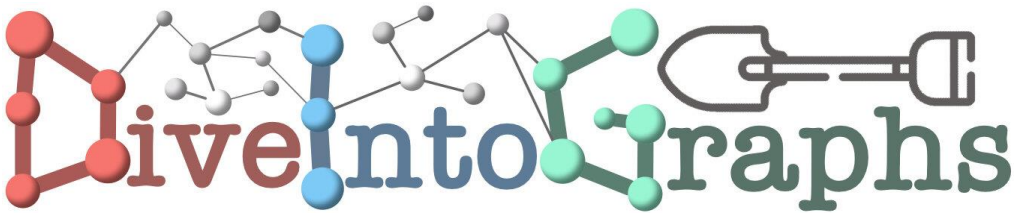
- Multi-objective generation
  - ➢ can improve multiple properties



DIG: Dive into Graphs

TEXAS A&M UNIVERSITY
Engineering

- A research-oriented testbed for graph neural networks
- Includes unified implementation of common methods, datasets, evaluations
- A turn-key solution to easily test new ideas and compare with baselines
- Currently, DIG includes four major research directions for graphs
  - Graph generation
  - Self-supervised learning on graphs
  - Interpretability of graph neural networks
  - 3D graphs



https://github.com/divelab/DIG (Coming soon!)

Conclusion and Outlook

TEXAS A&M UNIVERSITY
Engineering

- Concerted efforts to solve challenging real-world problems
- Develop new methods, and use them to solve problems in biology, neuroscience, chemistry, physics, etc.
- Scales range from cellular, molecular, to atomic in quantum world
- Interested in fundamental methodology, and fundamental science

Team and Acknowledgements

TEXAS A&M UNIVERSITY
Engineering

- Lei Cai (Scientist, Microsoft)
- Hongyang Gao (Assistant Professor, Iowa State)
- Zhengyang Wang (Scientist, Amazon)

- Meng Liu (Active PhD student)
- Yi Liu (Active PhD student)
- Youzhi Luo (Active PhD student)
- Bora Oztekin (Active Undergrad)
- Limei Wang (Active PhD student)
- Yaochen Xie (Active PhD student)
- Zhao Xu (Active PhD student)
- Hao Yuan (Active PhD student, joining Facebook)

TEXAS A&M UNIVERSITY
Engineering

Image and Graph Deep Learning Methods for Cellular and Molecular Level Science Discoveries

sji@tamu.edu
http://people.tamu.edu/~sji/
https://github.com/divelab/DIG/ coming in 3 weeks