ECCV

SparseSSP

3D Subcellular Structure Prediction from Sparse-View Transmitted Light Images

*Jintu Zheng, Yi Ding, Qizhe Liu, Yuehui Chen, Yi Cao, Ying Hu, and *Zenan Wang(Corresponding author)*

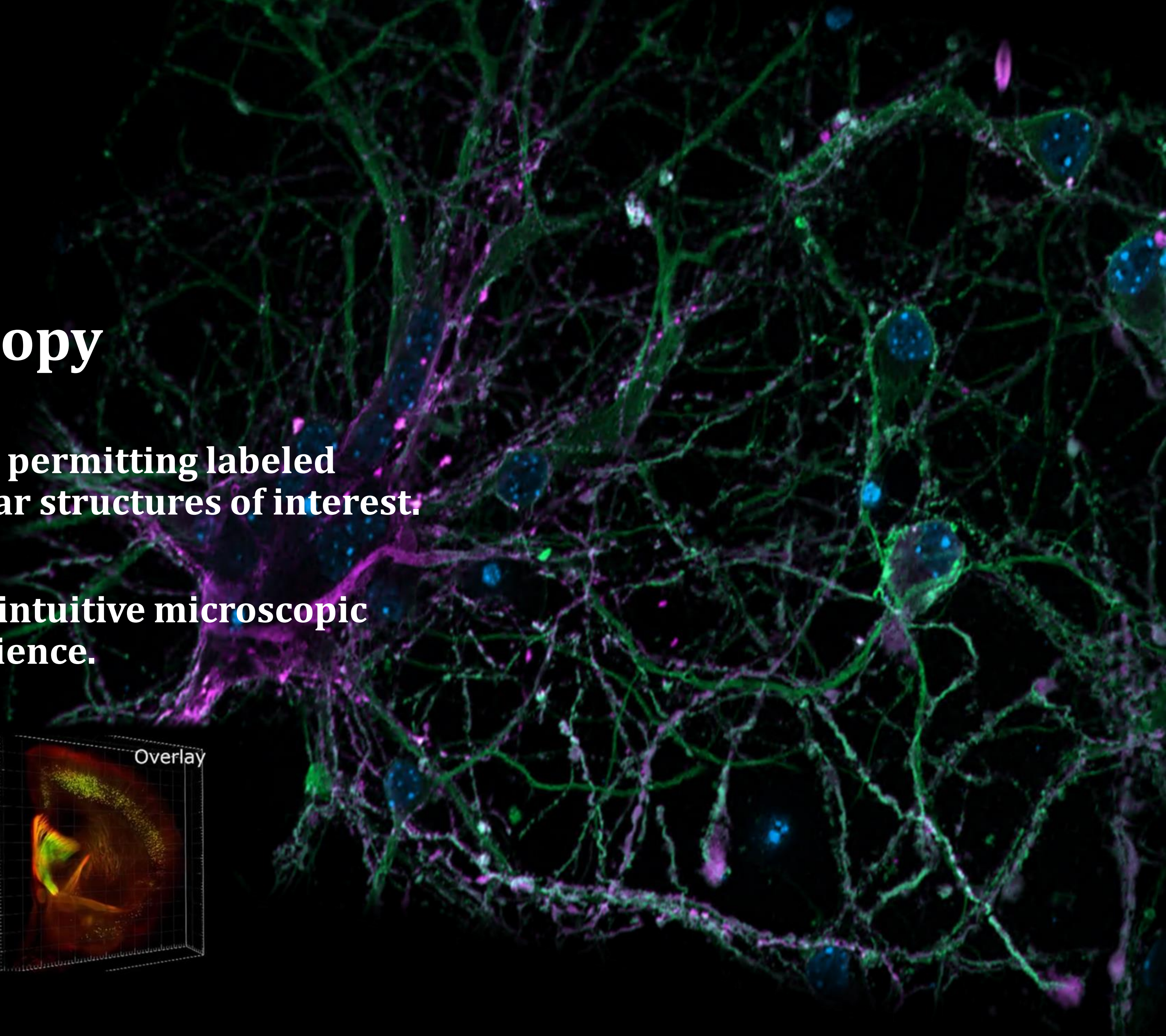
中国科学院大学
University of Chinese Academy of Sciences

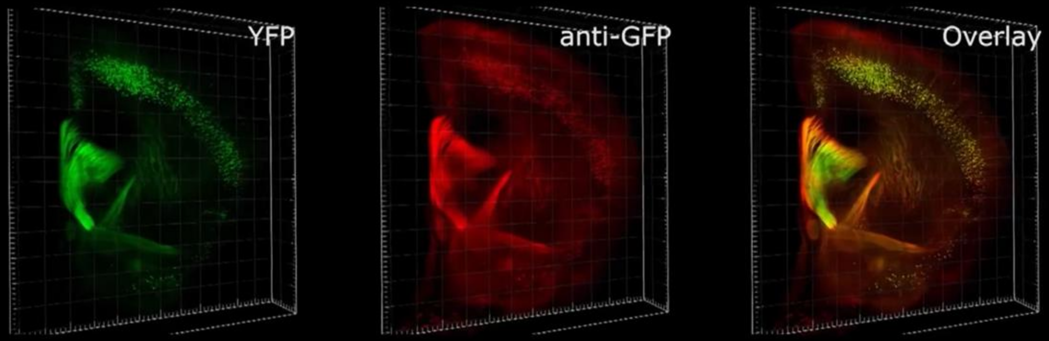
中国科学院深圳先进技术研究院
SHENZHEN INSTITUTE OF ADVANCED TECHNOLOGY
CHINESE ACADEMY OF SCIENCES

Fluorescence Microscopy

Revolutionized modern biology by permitting labeled imaging and quantifying subcellular structures of interest.

Current prevailing and inherently intuitive microscopic imaging mode in the field of life science.



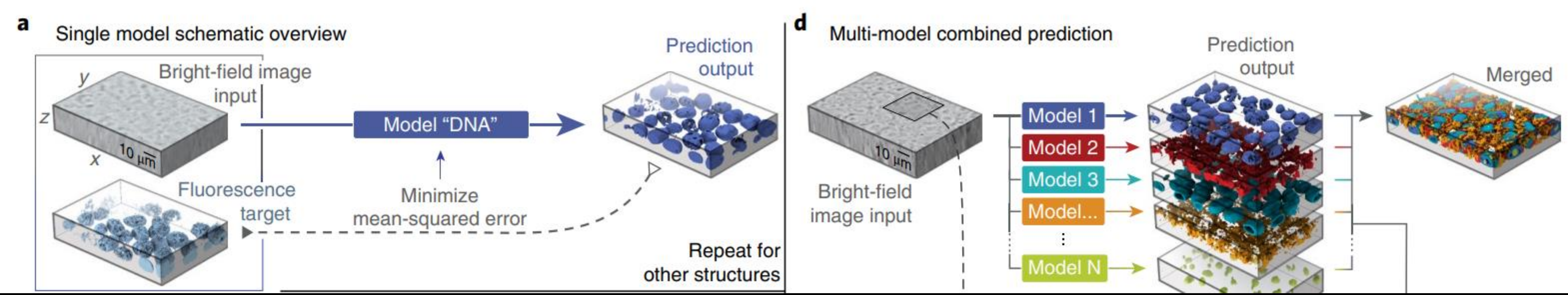


Subcellular Structure Prediction(SSP)

Fluorescence staining requires expensive and advanced instrumentation and time consuming preparation of materials.

Significant phototoxicity and photobleaching also damage the live cells.

An emerging technology, namely ***Subcellular Structure Prediction (SSP)***, enables direct prediction of 3D immunofluorescence (IF) from transmitted light (TL) images via 3D vision networks.



Cite from: Ounkomol, Chawin, et al. "Label-free prediction of three-dimensional fluorescence images from transmitted-light microscopy." Nature methods 15.11 (2018): 917-920.

Dense imaging process & Prolonged imaging time

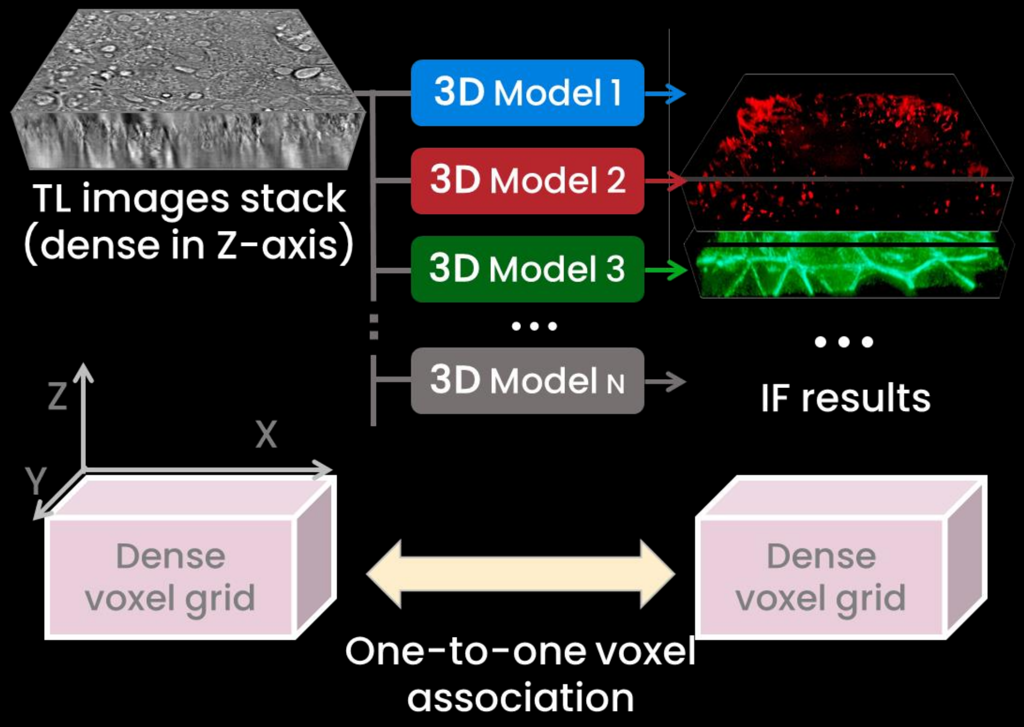
A motor is required to drive the lens to scan more layers on Z-axis for better data quality
In AllenCell collection, each subcellular type is imaged for up to ***2.5 hours*** on a Zeiss spinning disk microscope.

Prolonged imaging time is unfriendly to capturing the biodynamic process; the physiological motion, such as cell respiration, introduces the scanning position offset.

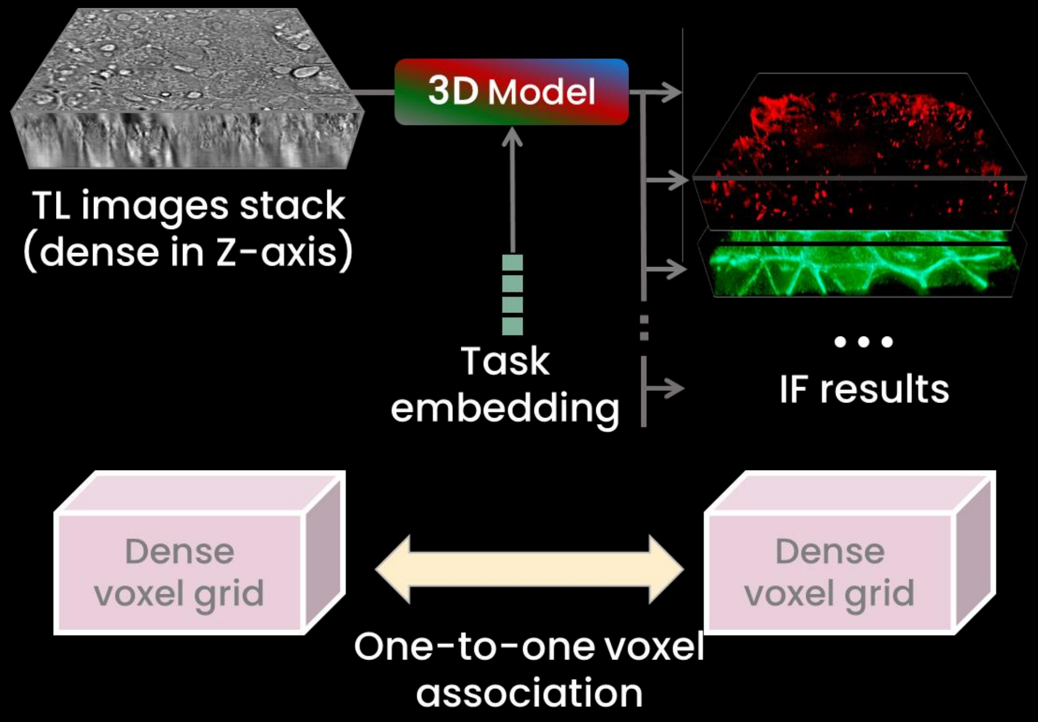
KEYWORDS

Fast, Live Cell Imaging, Low-Cost, Rapid biological dynamics visualization

Previous implementations



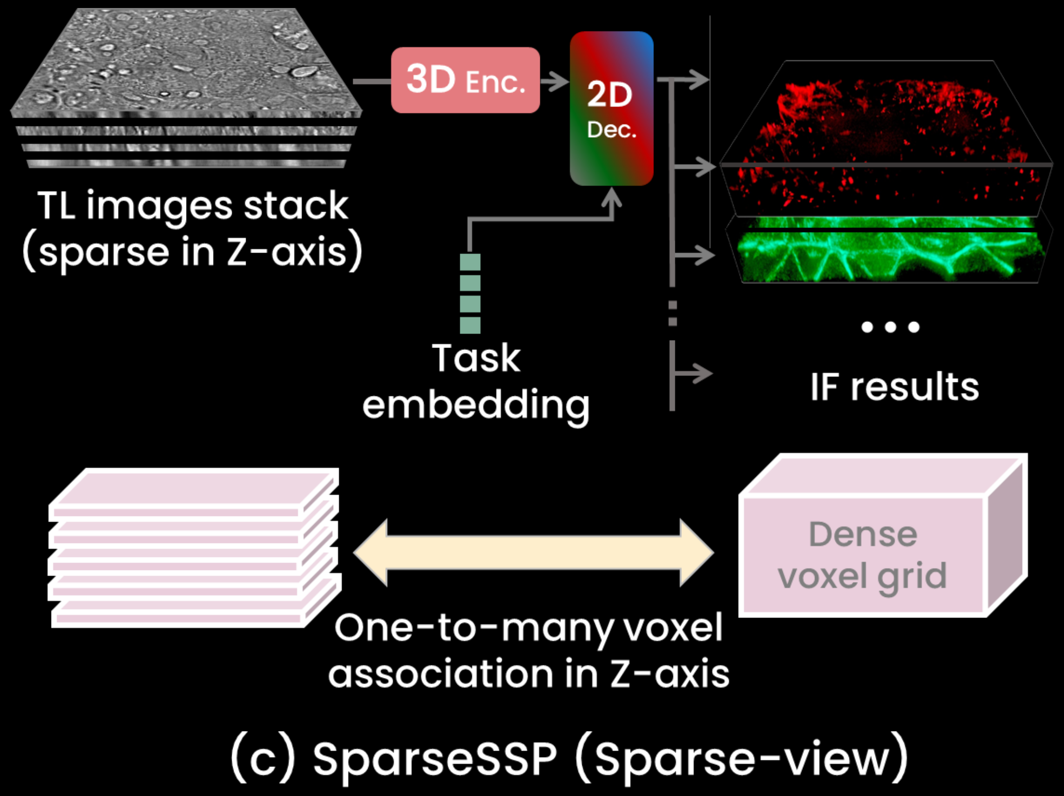
FNet (Nature Methods)



RepMode (CVPR 2023)

However, these one-to-one voxel learning approaches still require a long-time imaging process.

Sparse-View Techniques



Our proposal

Sparse-view techniques have emerged as a prominent research area in biological and medical garnering significant attention and interest.

For example, sparse-view techniques can reduce radiation dose in CT reconstruction with fewer projection times.

Similarly, less imaging time in SSP also reduces the phototoxicity of live cells.

Reduction of imaging times enables biologists to observe rapid biological dynamics in a cost-effective manner, facilitating better understanding of subcellular-level activities.

Insights of SparseSSP

Q. How to do sparse-view modeling?

A. Learning for one-to-many voxel regression

Presuppose a target voxel space implicit prior structural features are learned from the training data to assist in reconstructing the missing information

Q. How to learn better on sparse data?

A. Hybrid Topology Design

Fold sparse images along the Z axis onto the feature dimension, giving them the ability to regroup dense information during the channel transformation.

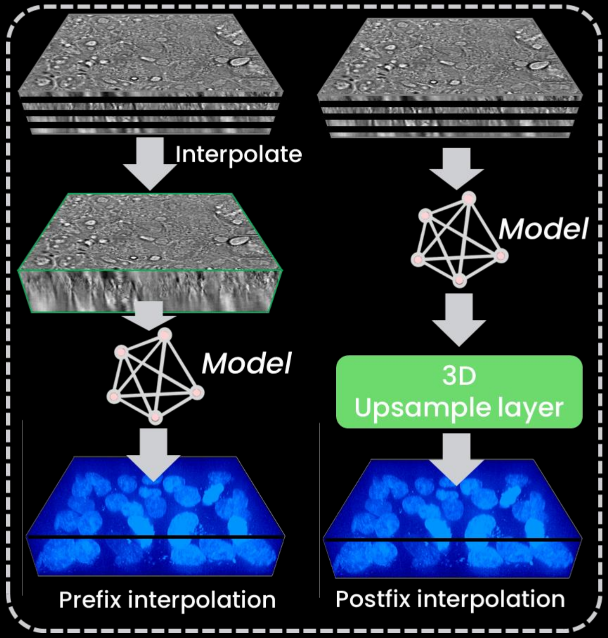
One-to-many Voxel Regression

Let A denotes the subsample operator, y denotes reconstructed images, b denotes sparse images.

Our goal is to learn the one-to-many mapping $f: b \rightarrow y$ by solving the $Ay = b$ problem.

But subsample operator A is not an invertible matrix, so there exists infinite solutions which indicates that this is an undetermined problem.

We can solve this problem by learning the follow mapping: $f: A'b \rightarrow y$, which A' denotes the reconstruction operation. To extract prior knowledge, deep learning method uses the training data $\{(b^{(k)}, y^{(k)})\}_{k=1}^{n}$ and fits the reconstruction process by learning the objective which is $\underset{f}{argmin}||(f(b^{(k)}) - y^{(k)})||_{l2}$.



Prefix strategy.

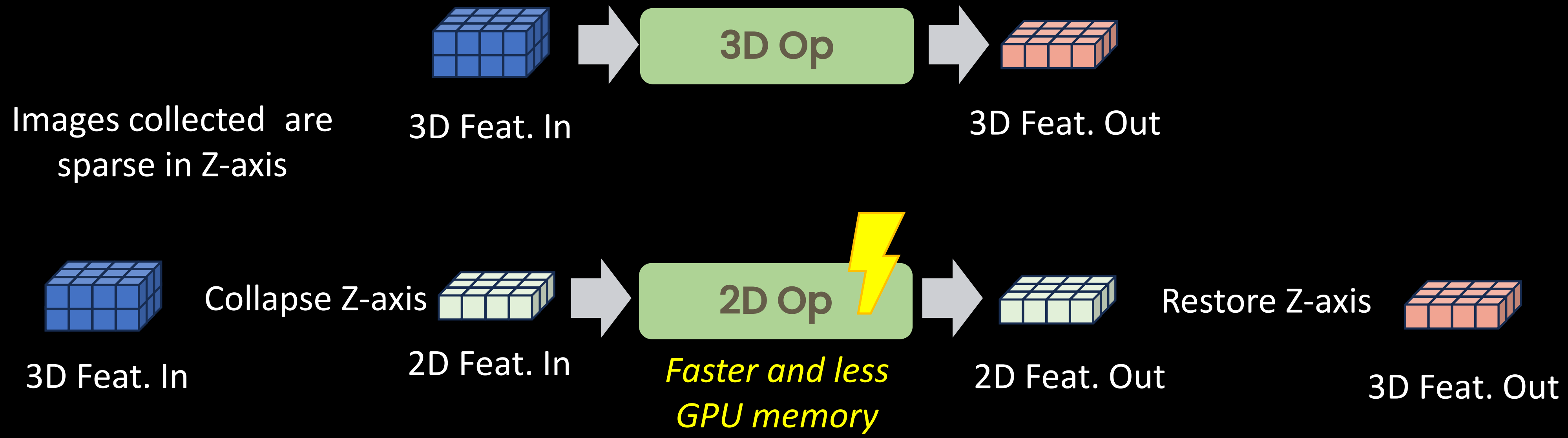
generates the pseudo voxel grid before the model input; in this strategy, the Z-axis information is implicitly restored through learning the fluorescence prediction.

Postfix strategy.

learns the restore procedure through an explicit upsampling layer, separates two processing purposes.

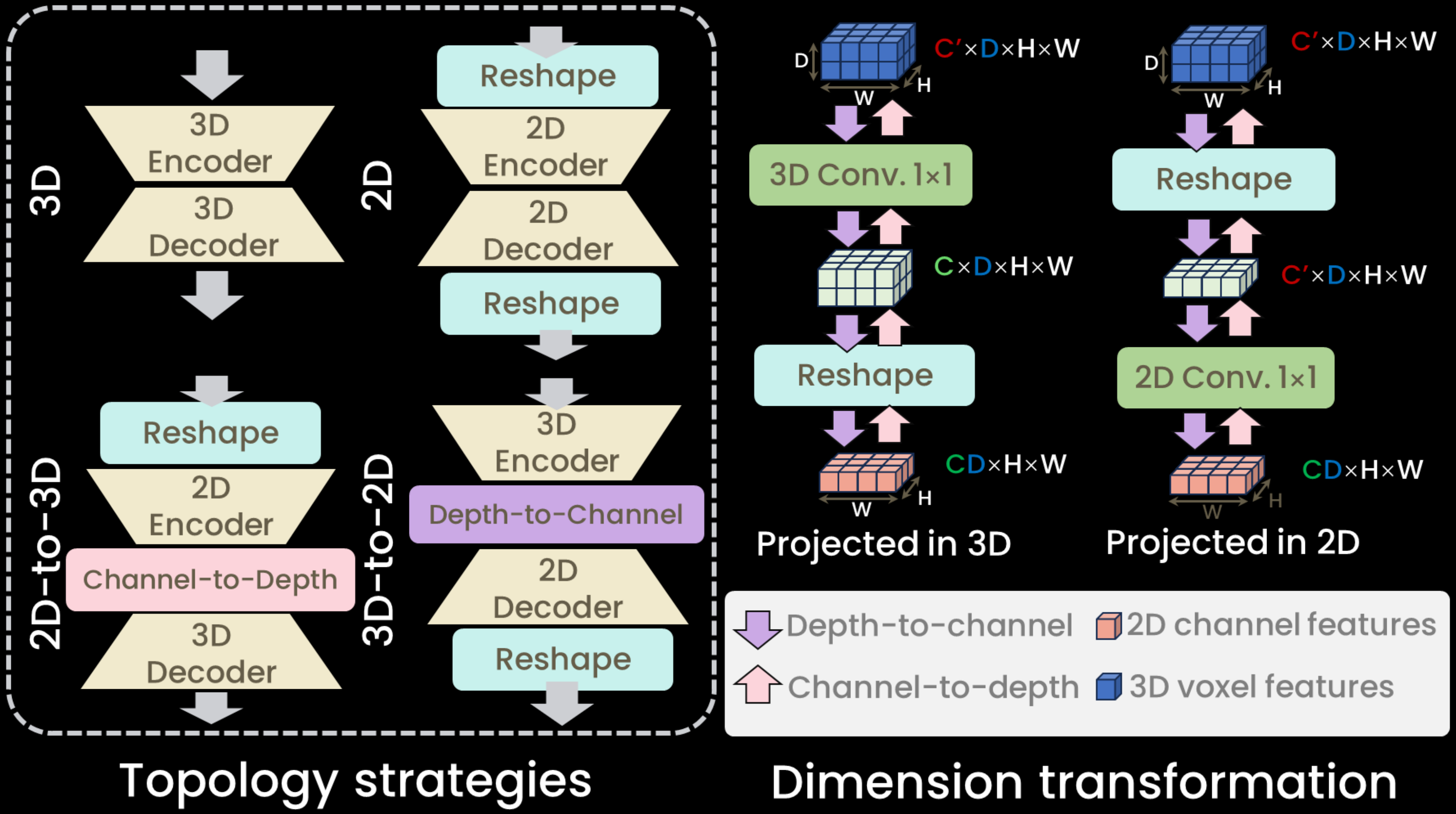
Hybrid Dimensions Topology

An Example of 3-to-2D

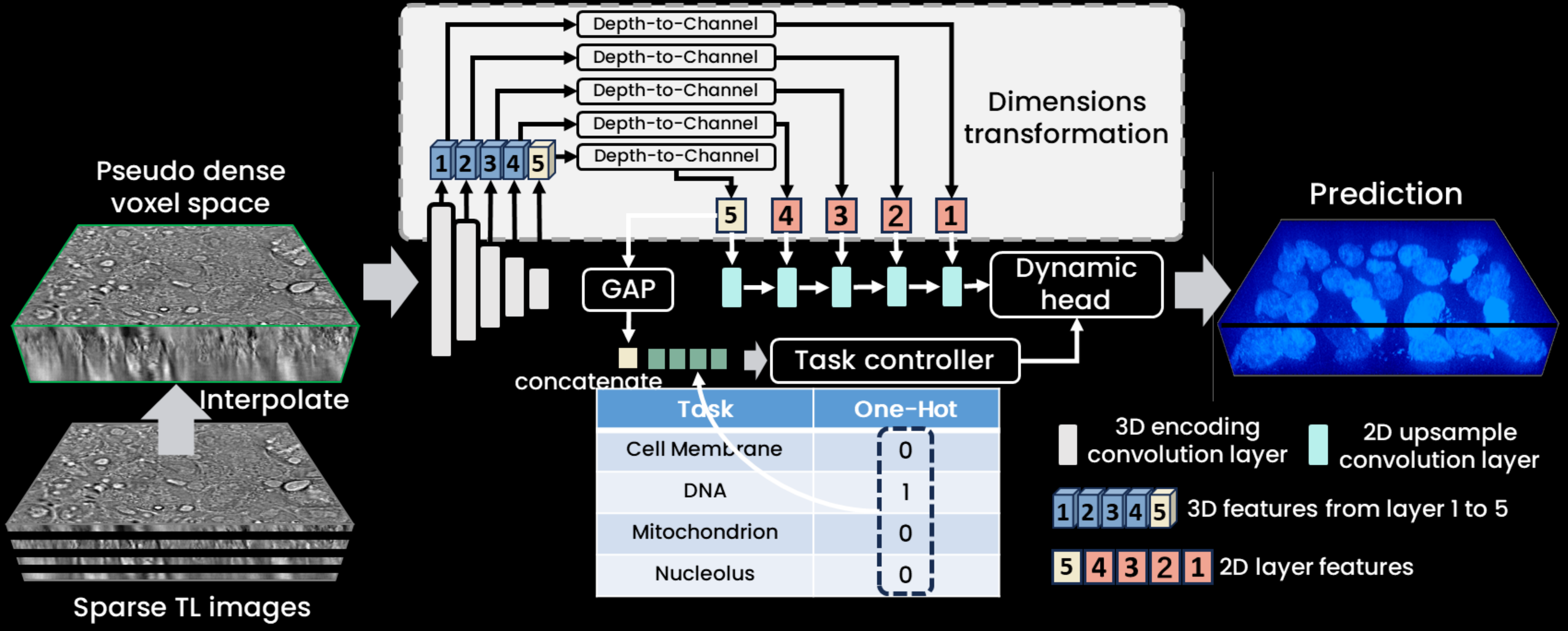


Learn the Z-axis information implicit reconstruction through collapse and reprojection
(It looks like an encode-to-decode procedure in depth-view)

Hybrid Dimensions Topology

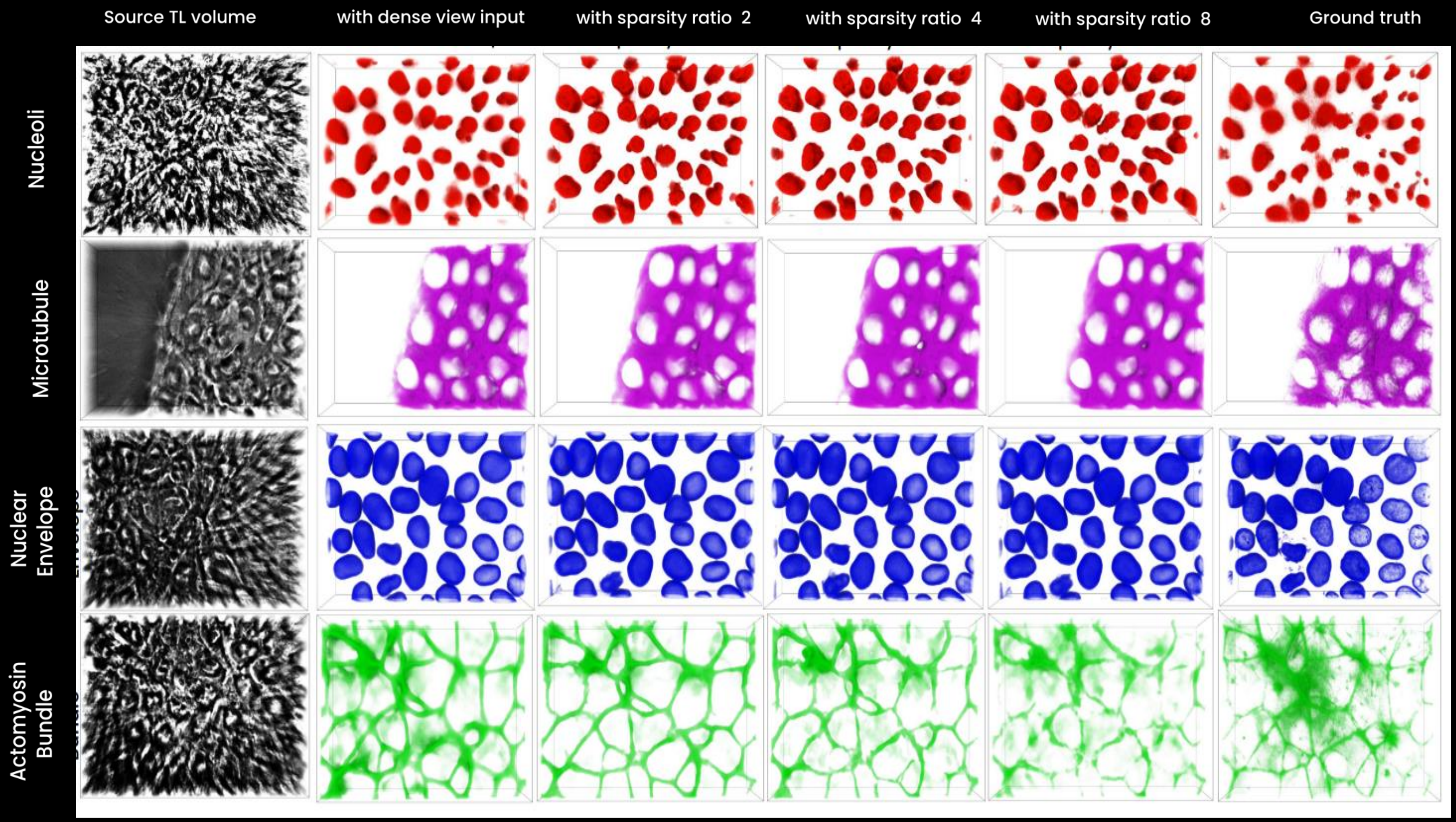


A 3-to-2D example apply on DoDNet*



*(CVPR 2021) Zhang J, Xie Y, Xia Y, et al. Dodnet: Learning to segment multi-organ and tumors from multiple partially labeled datasets[C]//Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2021: 1195-1204.

Visualization Results

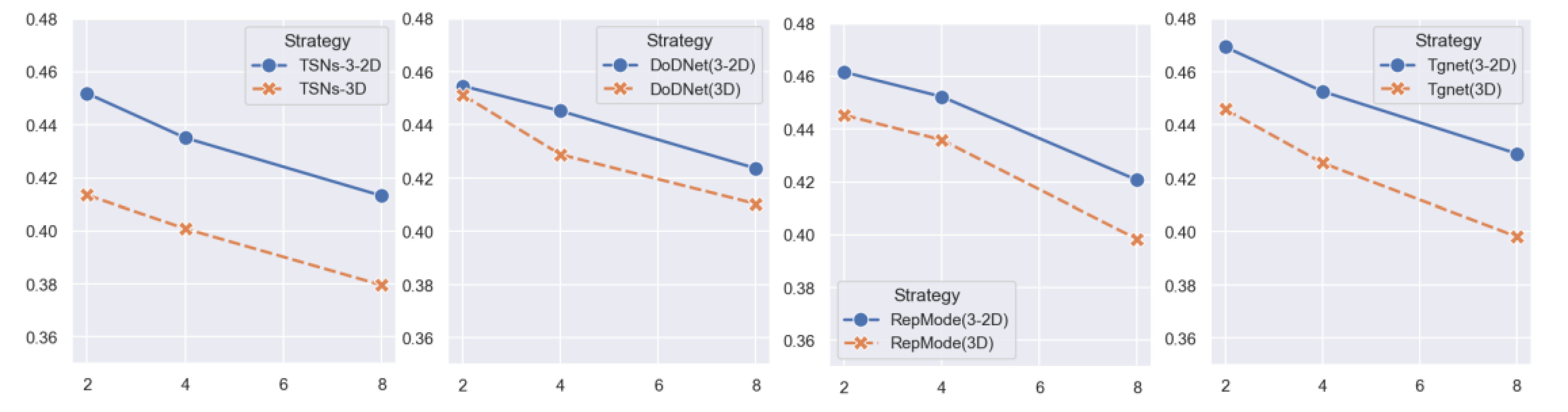


Comparisons of Strategy Combination

Approach	Interp.	Topology	r = 2			r = 4			r = 8		
			MSE	MAE	R^2	MSE	MAE	R^2	MSE	MAE	R^2
RepMode	None	2D	.6055	.4420	.3632	.6096	.4513	.3542	.6109	.4523	.3493
RepMode	None	3-to-2D	.5668	.4376	.4032	.5769	.4376	.3813	.6041	.4423	.3690
DoDNet	None	2D	.6103	.4509	.3524	.6238	.4623	.3243	.6298	.4602	.3103
DoDNet	None	3-to-2D	.5781	.4368	.3813	.5860	.4421	.3794	.6013	.4469	.3638
RepMode [30]	post	3D	.5455	.4234	.4193	.5539	.4360	.4137	.5886	.4424	.3736
RepMode	pre	3D	.5210	.4248	.4453	.5320	.4197	.4359	.5687	.4359	.3984
RepMode	post	2D	.6043	.4471	.3613	.6154	.4500	.3460	.6189	.4562	.3324
RepMode	pre	2D	.5704	.4532	.3984	.5812	.4383	.3857	.5896	.4484	.3849
RepMode	post	2-to-3D	.5458	.4236	.4250	.5523	.4312	.4153	.5896	.4447	.3790
RepMode	pre	2-to-3D	.5135	.4173	.4543	.5299	.4194	.4397	.5624	.4286	.3971
RepMode	post	3-to-2D	.5234	.4132	.4589	.5356	.4232	.4264	.5780	.4313	.3862
RepMode	pre	3-to-2D	**.5069**	.4140	**.4616**	**.5159**	.4136	**.4523**	.5468	.4222	.4207
DoDNet [28]	post	3D	.5343	.4241	.4243	.5541	.4313	.4117	.5768	.4381	.3861
DoDNet	pre	3D	.5173	.4257	.4512	.5392	.4318	.4288	.5572	.4316	.4103
DoDNet	post	2D	.6012	.4487	.3643	.6045	.4412	.3632	.6123	.4561	.3520
DoDNet	pre	2D	.5751	.4372	.3883	.5793	.4342	.3734	.5823	.4413	.3699
DoDNet	post	2-to-3D	.5486	.4329	.4232	.5554	.4382	.4143	.5774	.4367	.3903
DoDNet	pre	2-to-3D	.5244	.4185	.4463	.5367	.4150	.4221	.5535	.4234	.4145
DoDNet	post	3-to-2D	.5354	.4213	.4201	.5475	.4335	.4172	.5634	.4318	.4032
DoDNet	pre	3-to-2D	.5128	**.4118**	.4516	.5229	**.4133**	.4452	**.5440**	**.4209**	**.4236**

Combination of prefix interpolation and 3-to-2D strategies demonstrated significantly better performance than others.

Topology Strategies on Diverse Multi-task Methodologies



We compare 5 SOTA multi-task methodologies.
Trend of R2 value as sparsity ratio increased from 2 to 8.
Hybrid dimensions topology 3-to-2D (i.e., the blue lines in the figure) shows a slower decay and higher global scores than pure 3D topology (i.e., the orange lines).

Comparisons on Resource Consumption

		GPU Infer.		GPU Train.		Computation
Approach	Topology	time(s/iter)	Mem.(MiB)	Speed(iter/s)	Mem.(MiB)	MACs
RepMode	3D	4.47	9122	0.89	17843	66.29G
RepMode	2D	0.66	2472	3.31	3548	2.33G
RepMode	2-to-3D	1.86	3666	1.86	8984	30.11G
RepMode	3-to-2D	2.55	5428	1.29	15521	43.47G
DoDNet	3D	2.11	4710	2.64	16054	113.86G
DoDNet	2D	0.35	1982	4.70	2692	1.82G
DoDNet	2-to-3D	1.04	2897	3.31	4384	41.02G
DoDNet	3-to-2D	1.35	4124	4.53	14362	76.61G
TSNs	3D	1.09	6862	7.19	10458	55.70G
TSNs	2D	0.37	1720	31.12	2793	2.05G
TSNs	2-to-3D	0.86	2028	10.2	5729	13.94G
TSNs	3-to-2D	0.91	3554	8.91	7522	46.77G

Hybrid dimensions topologies demonstrate less resource consumption than pure 3D, especially in MACs.
The number of iterations in training is the number of the loss backward operations

THANK YOU

Paper link: https://arxiv.org/abs/2407.02159
Code link: https://github.com/JintuZheng/SparseSSP