RUTGERS
THE STATE UNIVERSITY OF NEW JERSEY

CARTA
CENTER FOR ACCELERATED REAL TIME ANALYTICS

Rutgers CARTA Site

Dimitris N. Metaxas

Distinguished Professor

Director Center for Computational Biomedicine, Imaging and Modeling (CBIM)

Director NSF-IUCRC CARTA Phase II

Email: dnm@cs.rutgers.edu

CARTA Phase II: carta.rutgers.edu

Multidisciplinary Team

- 17 PhD Students
- CBIM: 13 Faculty members
- Collaborations with PIs from major Universities, Hospitals and Companies

AI/ML Research

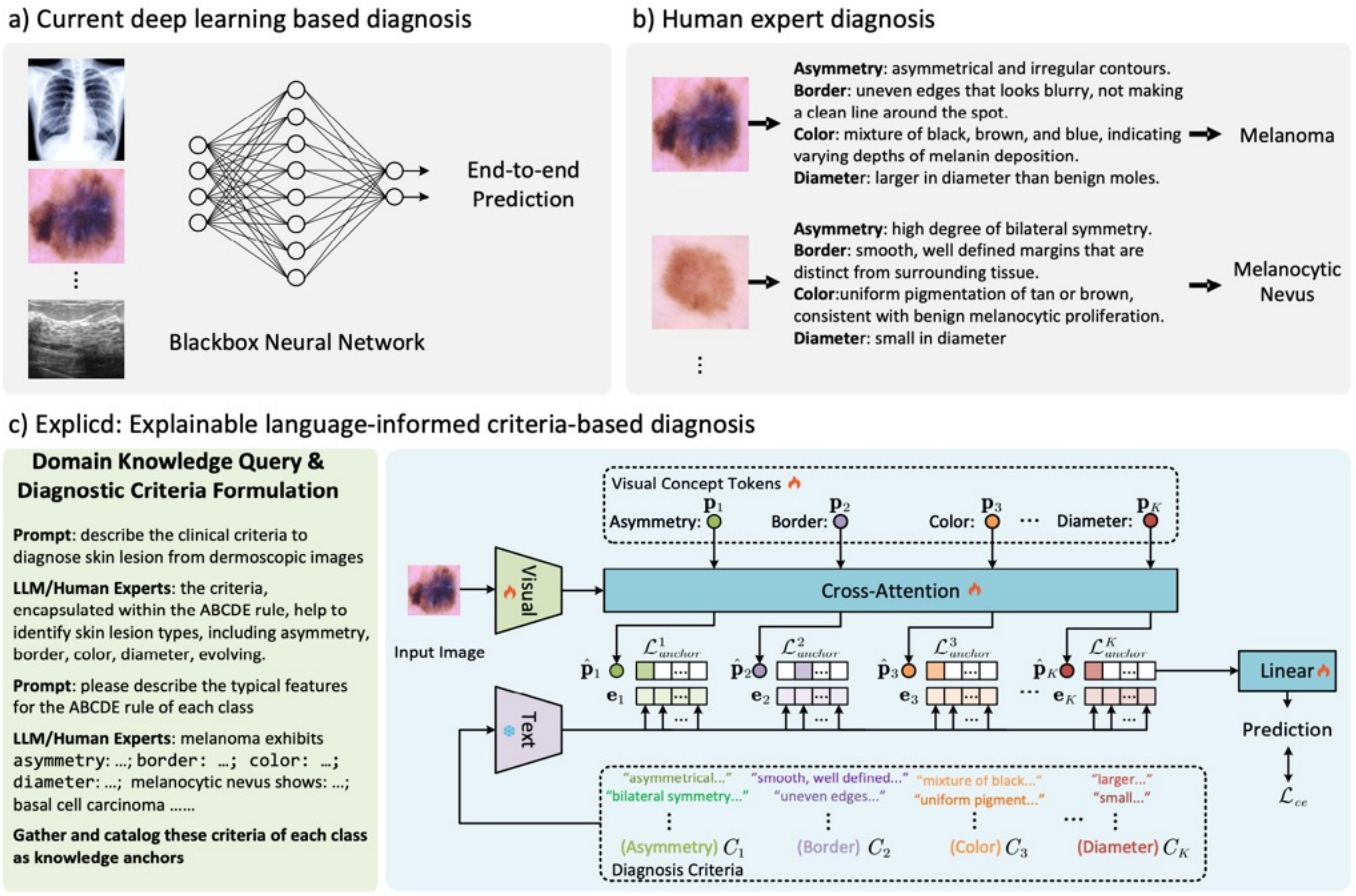
- Explainable AI
- Federated Learning
- Generative AI
- Large Language and Multimodal Foundation Models
- Multimodal Data Fusion
- Graph Neural Nets
- Semi and Unsupervised Learning
- Diffusion Methods
- Domain knowledge incorporation for improved inference
- Real time solutions

Application Domains

- Biomedical Applications
  - Cardiac, Cancer, Joints
  - K-space MRI Reconstruction
  - Histopathology
  - Spatial Biology and Multi-omics
  - Explainable Solutions
- Computer Vision
  - Segmentation, Registration, 3D object reconstruction, motion analytics
  - Explainable shape, motion analytics
  - Shape and motion generation with relationships

Biomedical Projects

Explainable Medical Inference
Aligning Medical Knowledge to Data



Histopathology Nuclei Segmentation Challenges

1.Data Annotation

- Laborious time
- Domain knowledge required

↓

Weakly supervised segmentation

2.Model Architecture

- Image details lost due to pooling

↓

Full resolution neural network

3.Loss Function

- Pixel-level cross entropy is limited
- There is a lack of spatial constraint

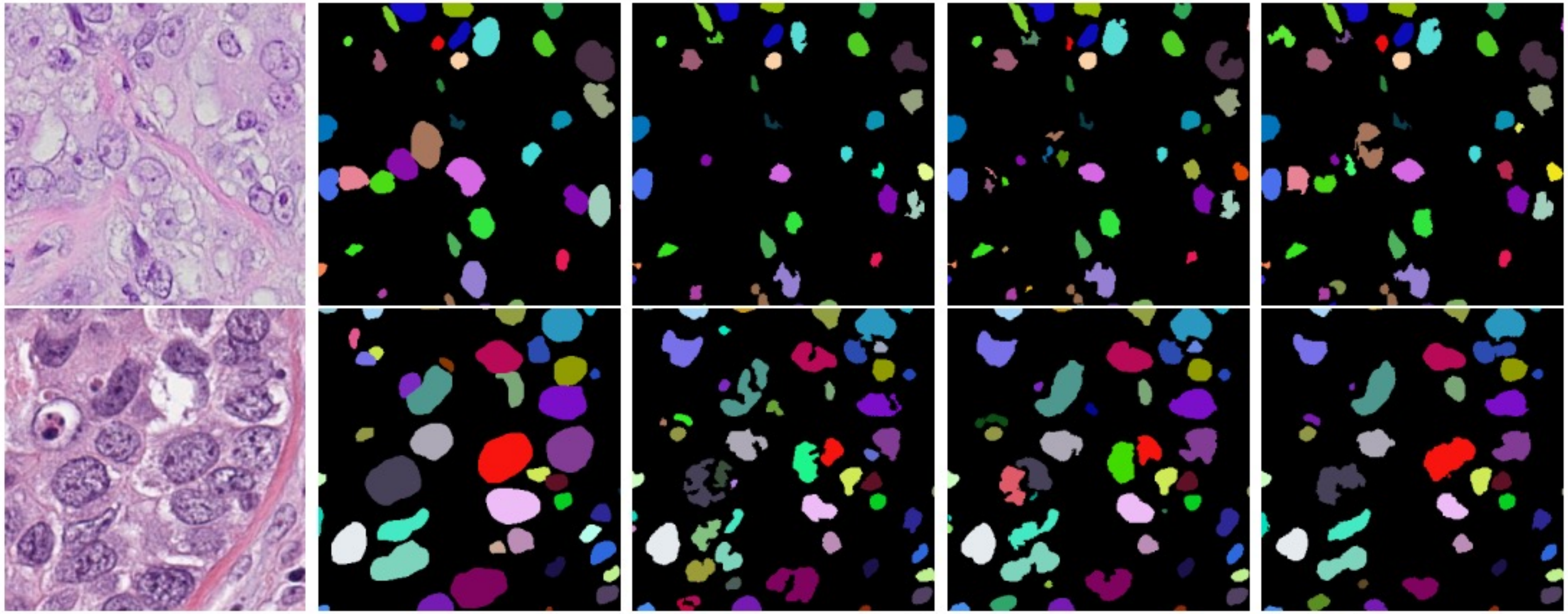
↓

Variance constrained loss

We proposed a series of novel deep-learning approaches to address the multiple challenges of nuclei segmentation

Weakly Supervised Nuclei Segmentation

- Extension: select 5% representative hard nuclei to annotate masks
  - Better than points and random selection



(a) image (b) GT (c) Points [73] (d) Mixed (rand) (e) Mixed (un)

Improving Nuclei Segmentation: FullNet & varCE loss

- More accurate boundaries help separate touching nuclei

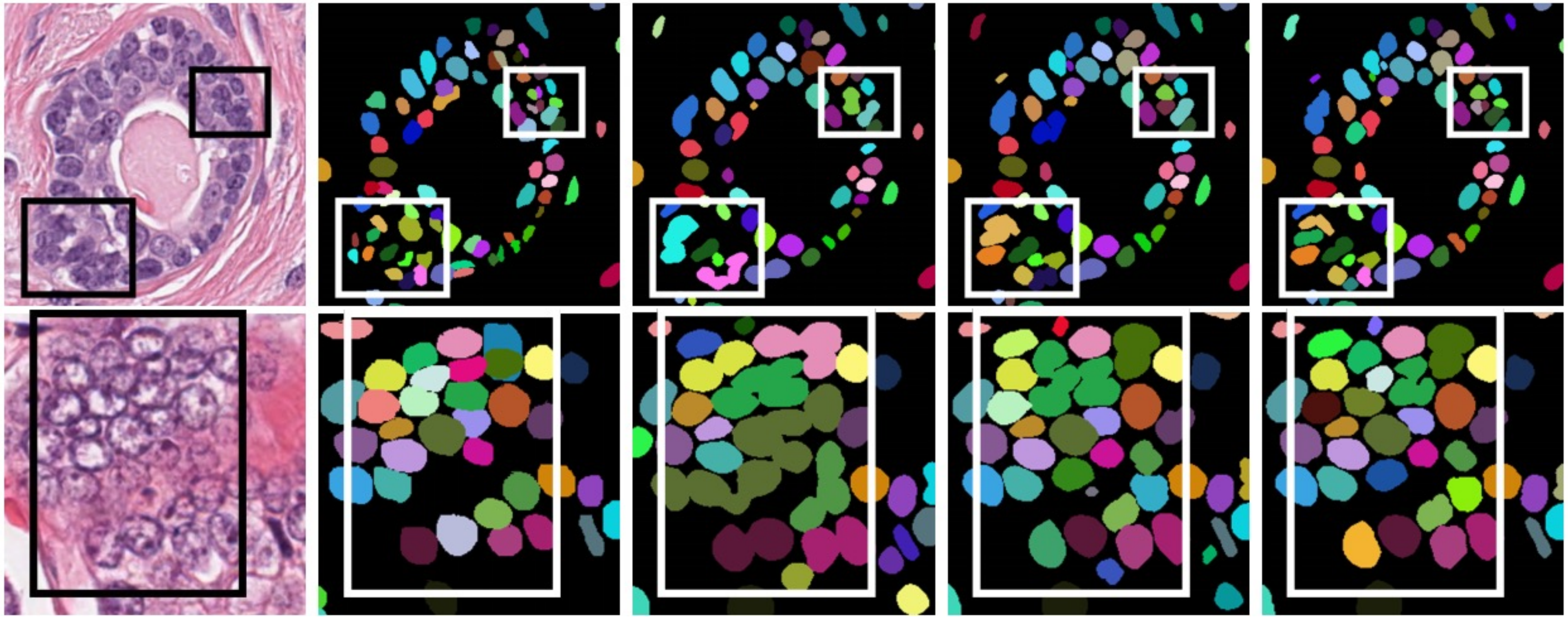


image groundtruth FCN-pooling FullNet FullNet & varCE

npj | precision oncology

Article | Open Access | Published: 23 September 2021

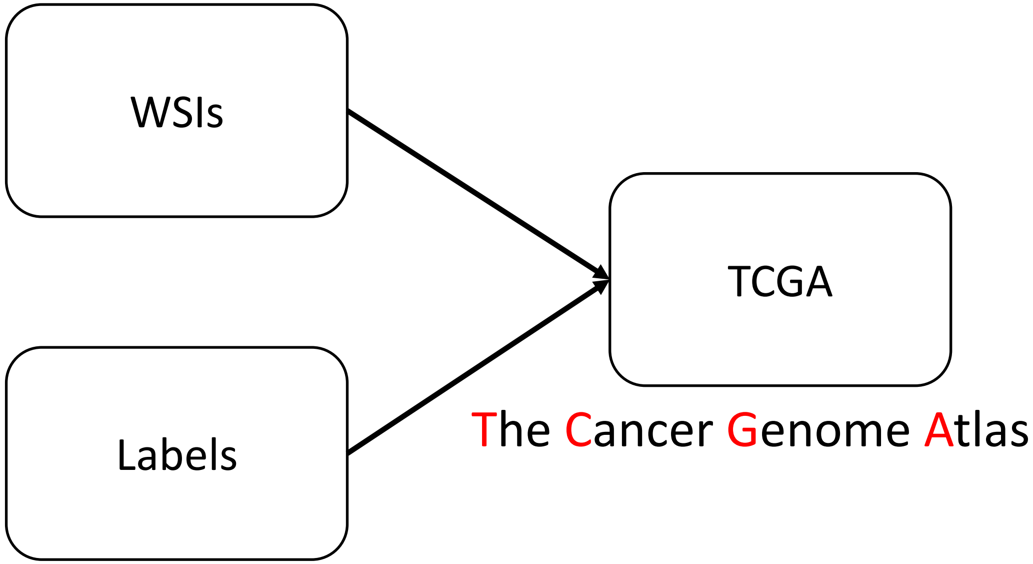
Genetic mutation and biological pathway prediction based on whole slide images in breast carcinoma using deep learning

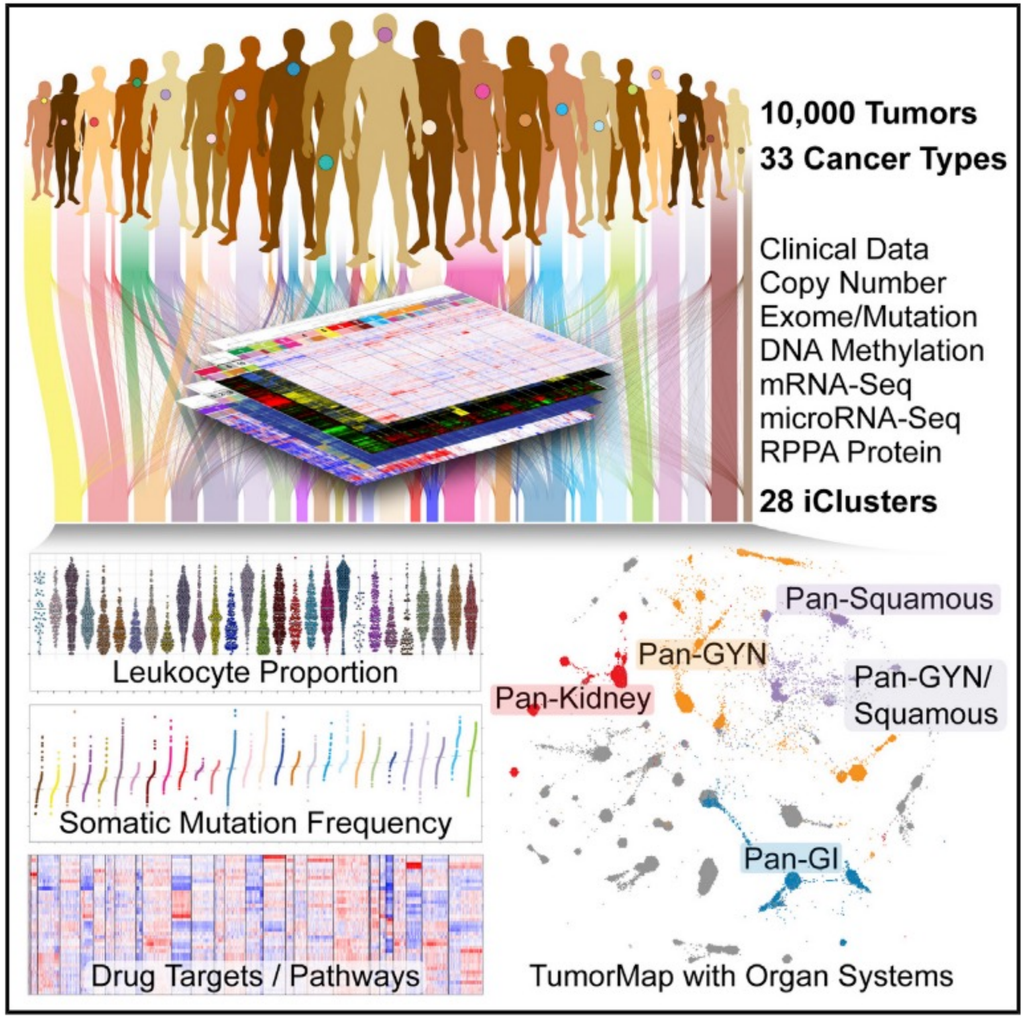
Hui Qu, Mu Zhou, Zhennan Yan, He Wang, Vinod K. Rustgi, Shaoting Zhang ✉, Olivier Gevaert ✉ & Dimitris N. Metaxas ✉

npj Precision Oncology **5**, Article number: 87 (2021) | Cite this article

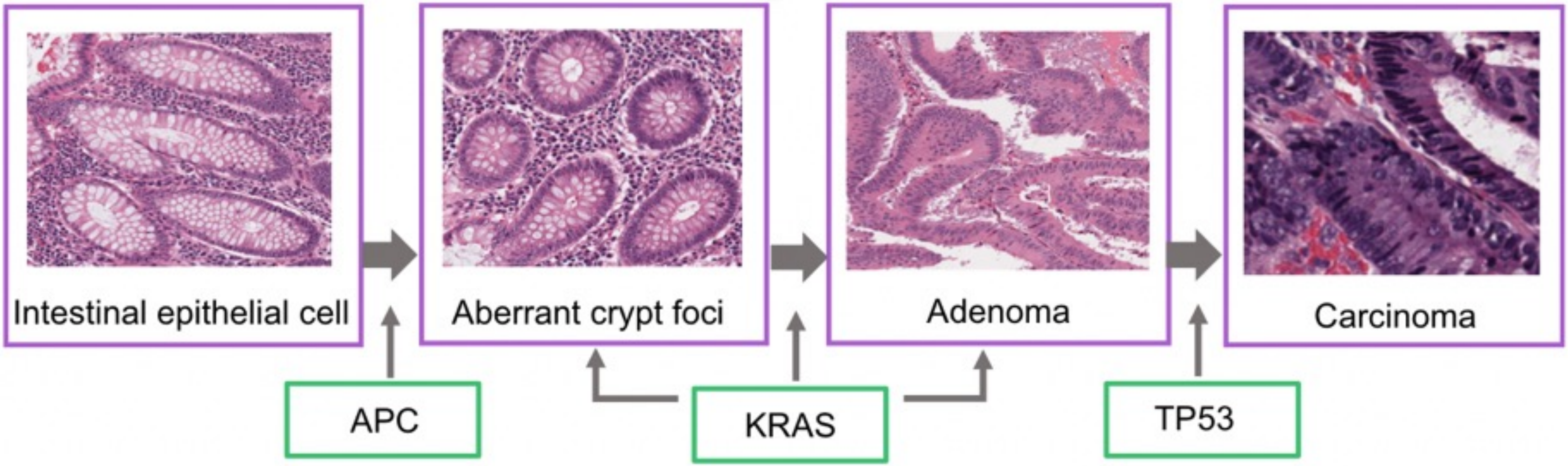
Gene Mutation and Pathway Prediction

- **Goal**: predict mutation and pathway activity from whole slide images (WSIs) directly





Colon Cancer Histopathology and Its Key Mutations

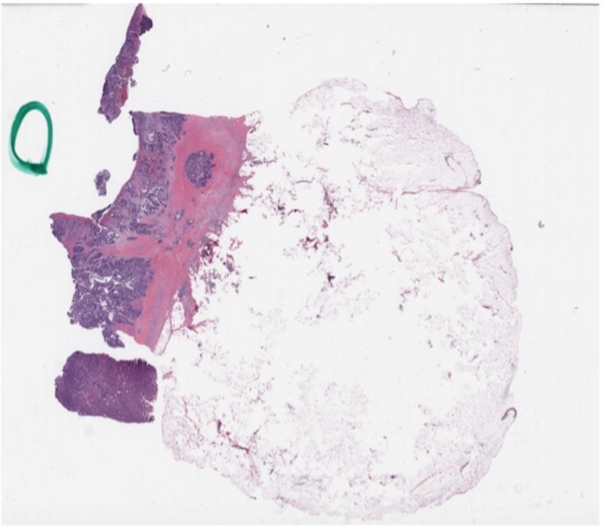


From Left to Right: Early to late Cancer Stage

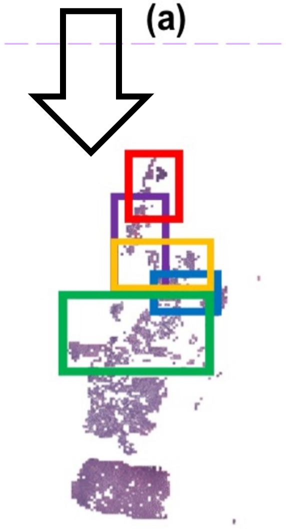
Goal:
We seek to build image-based graph and identify detectable evidence for cancer molecular outcomes with therapy implications

Kexin Ding, etc. MICCAI 2020.

Results Visualization

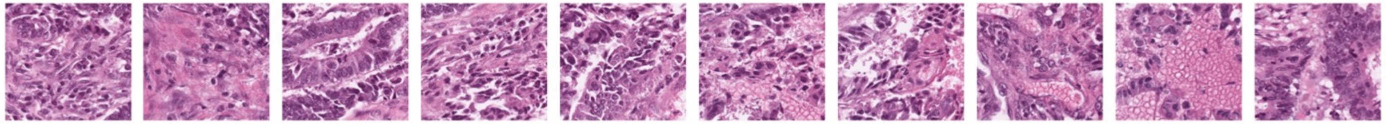


(a)

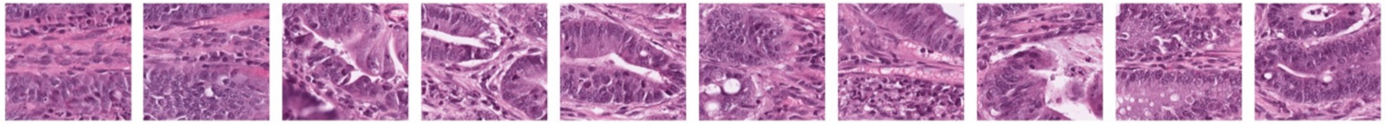


(b)

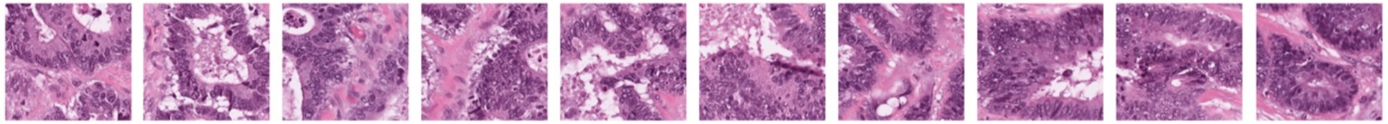
Region colors correspond to results from each subgraph model to the right.



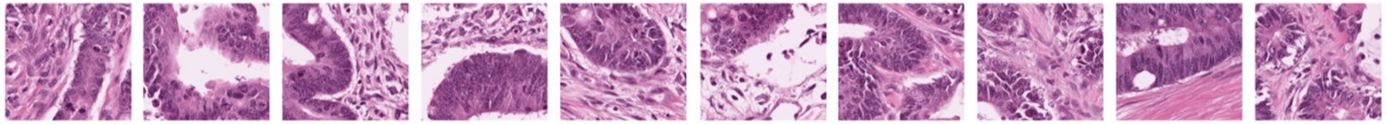
(c) Top-10 identified tiles from subgraph model 1 (within purple region)



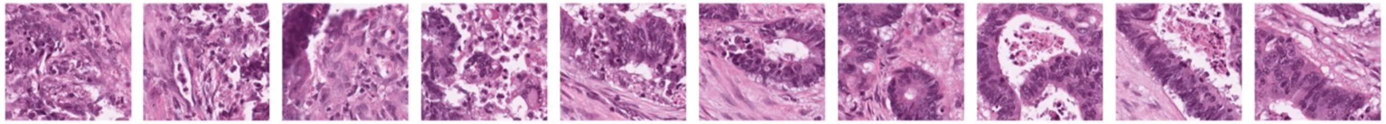
(d) Top-10 identified tiles from subgraph model 2 (within red region)



(e) Top-10 identified tiles from subgraph model 3 (within blue region)



(f) Top-10 identified tiles from subgraph model 4 (within green region)



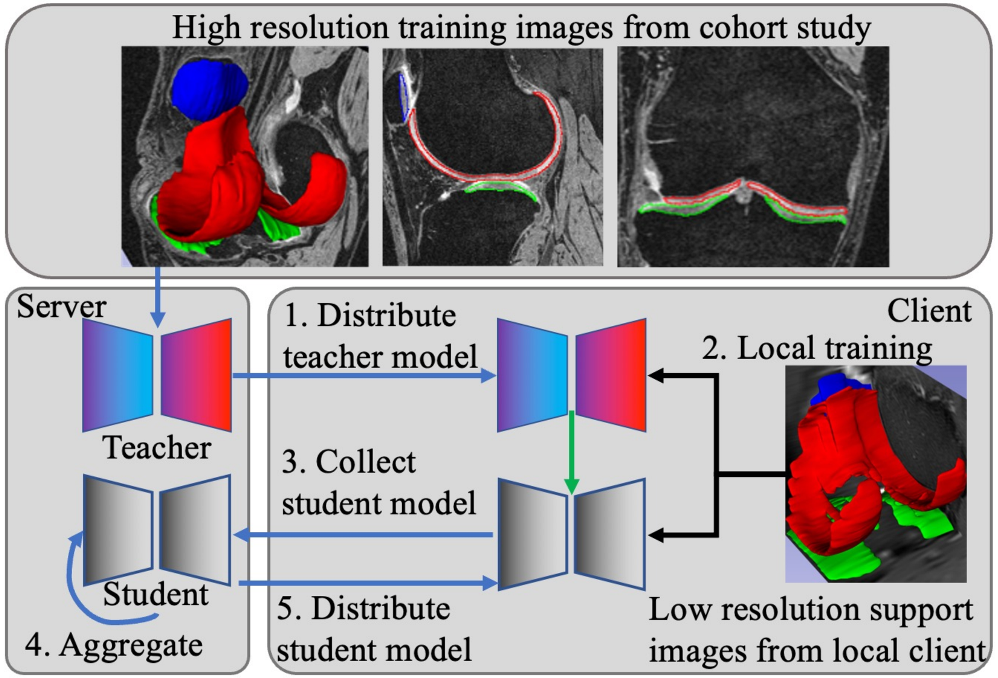
(g) Top-10 identified tiles from subgraph model 5 (within yellow region)

Statistics of subgraphs	Number of node	Node degree	Clustering coefficient	Closeness centrality	Betweenness centrality
	1000	915.4283	0.9400	0.9177	0.0001

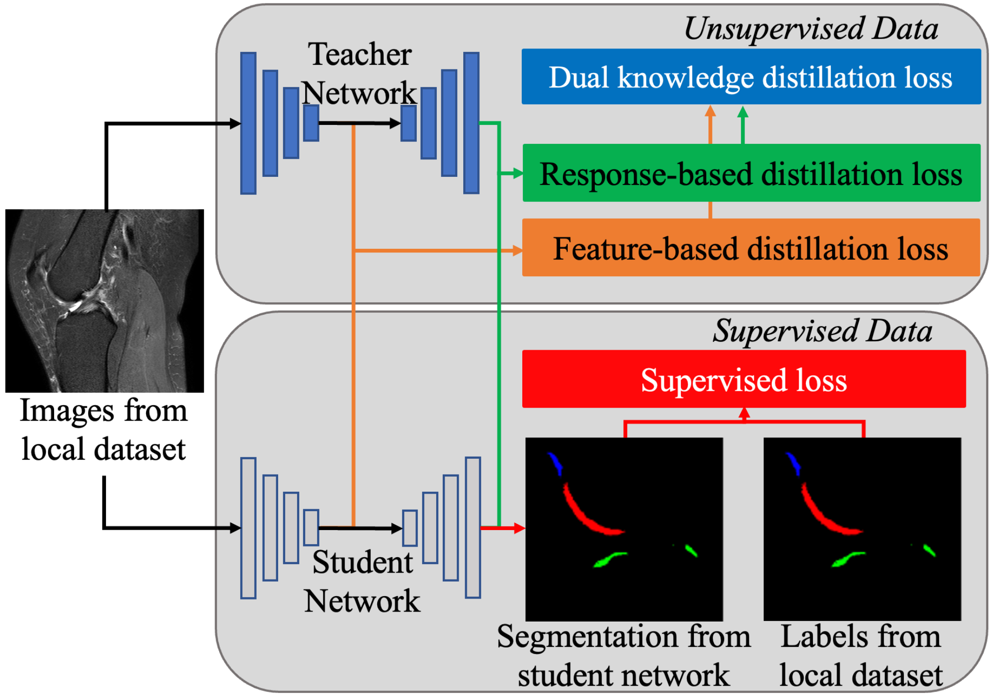
(h)

Federated Few-Shot Learning with Dual Knowledge Distillation on Medical Imaging

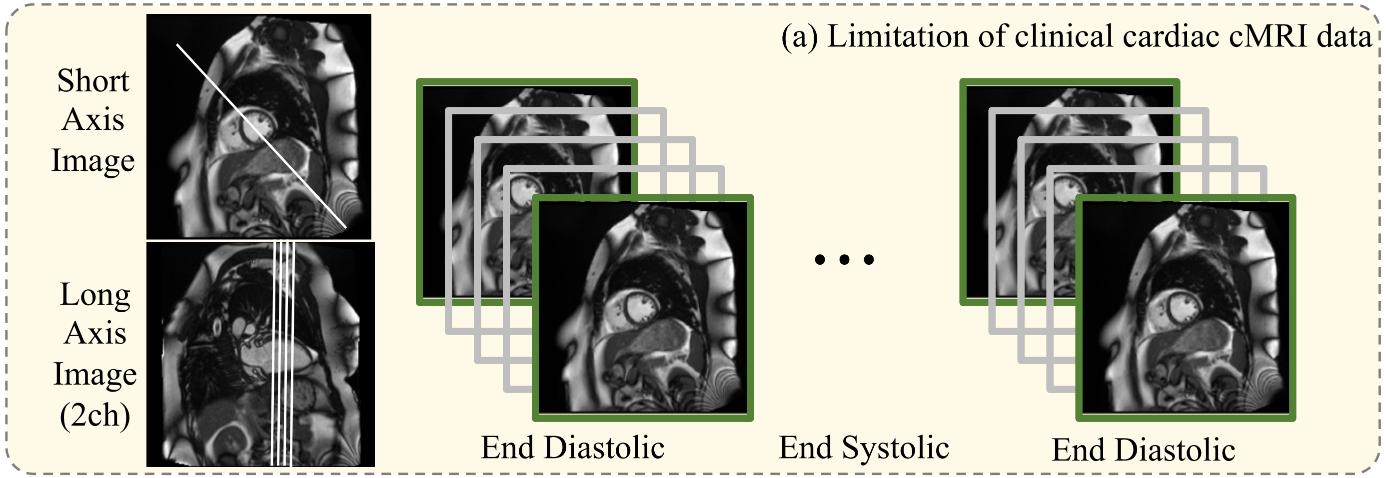
Motivation: **Sporadic** Distributed Data, **Limited** Annotations, **Heterogeneous** Data Distribution



Method: Distill both **feature-based** and **response-based** knowledge from teacher network

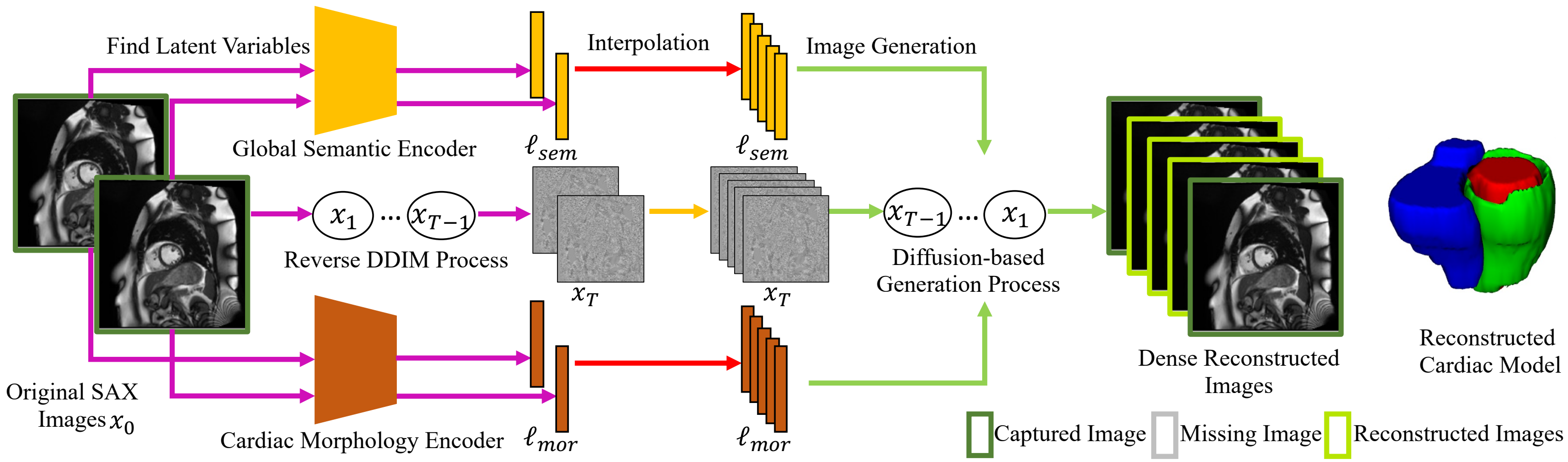


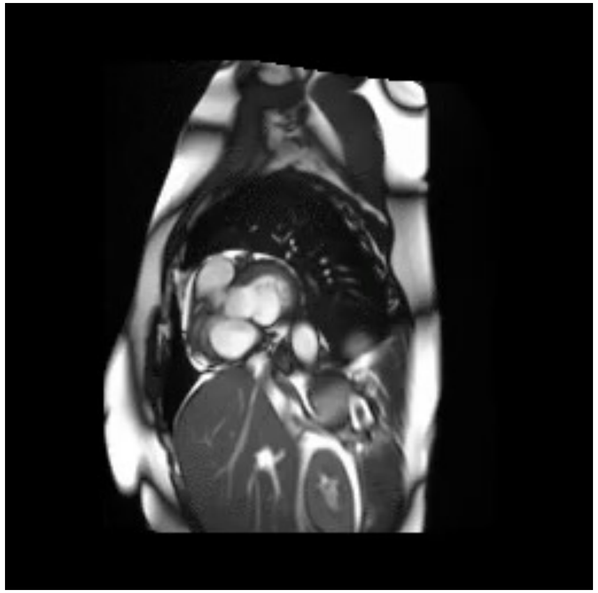
Morphology-Guided Diffusion Model for 3D Volume Reconstruction



Motivation: Improve **spatial** and **temporal resolution** of 4D MRI.

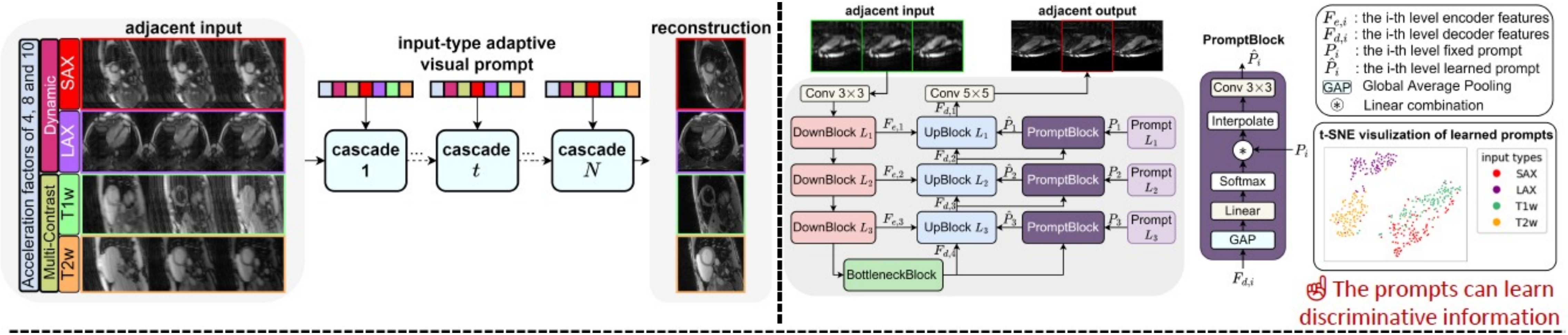
Method: Utilize **features** from **reconstruction** and **label** to condition on the diffusion process.



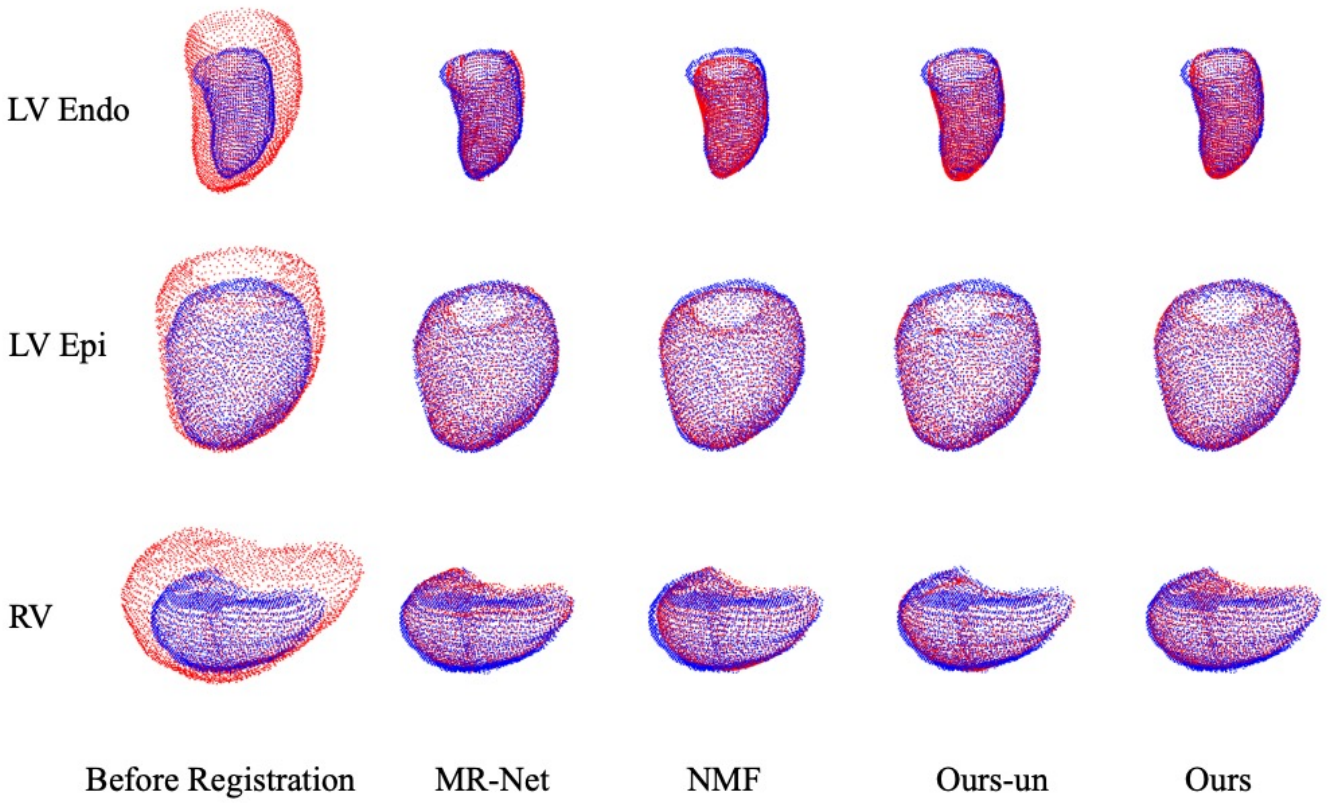


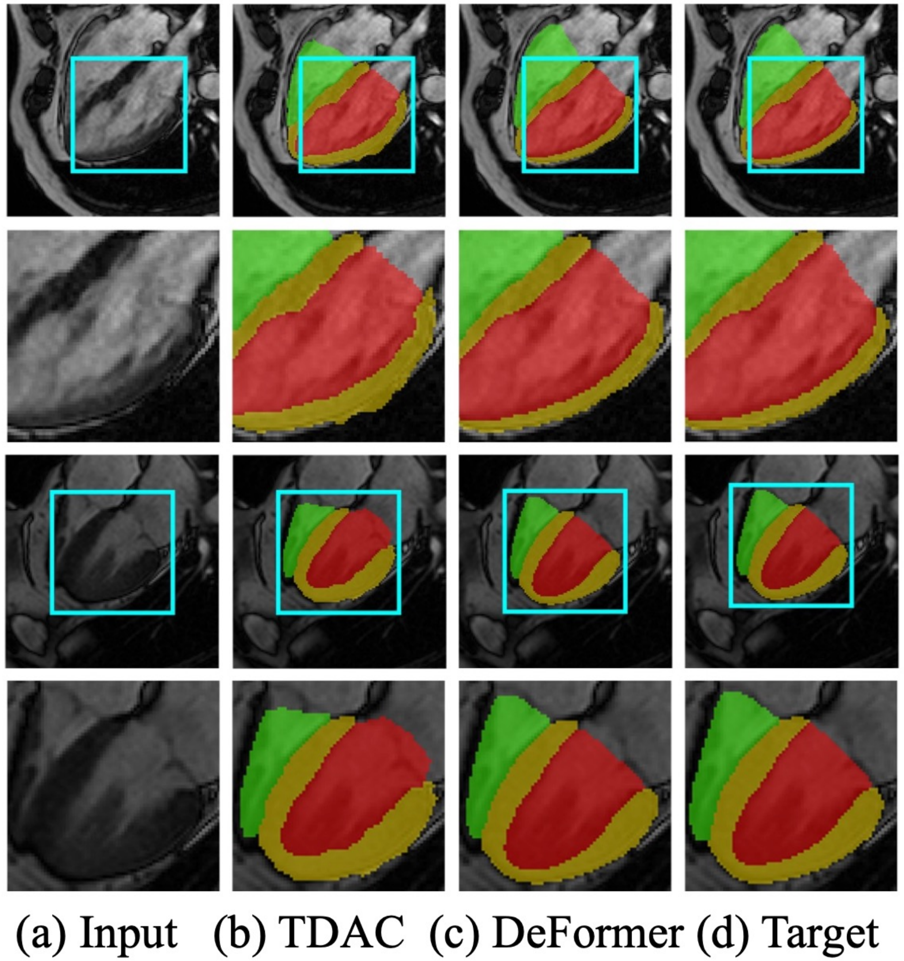
Promoting for All-In-One MRI reconstruction

Conventional CNN-based MRI reconstruction models often require training and deployment for each specific imaging scenario (imaging sequence, view, and device vendor), limiting their clinical application in the real world, we tackle this challenge via prompt-based learning for all-in-one MRI reconstruction.



Integrating Deep Learning with Physics-based Deformable Models





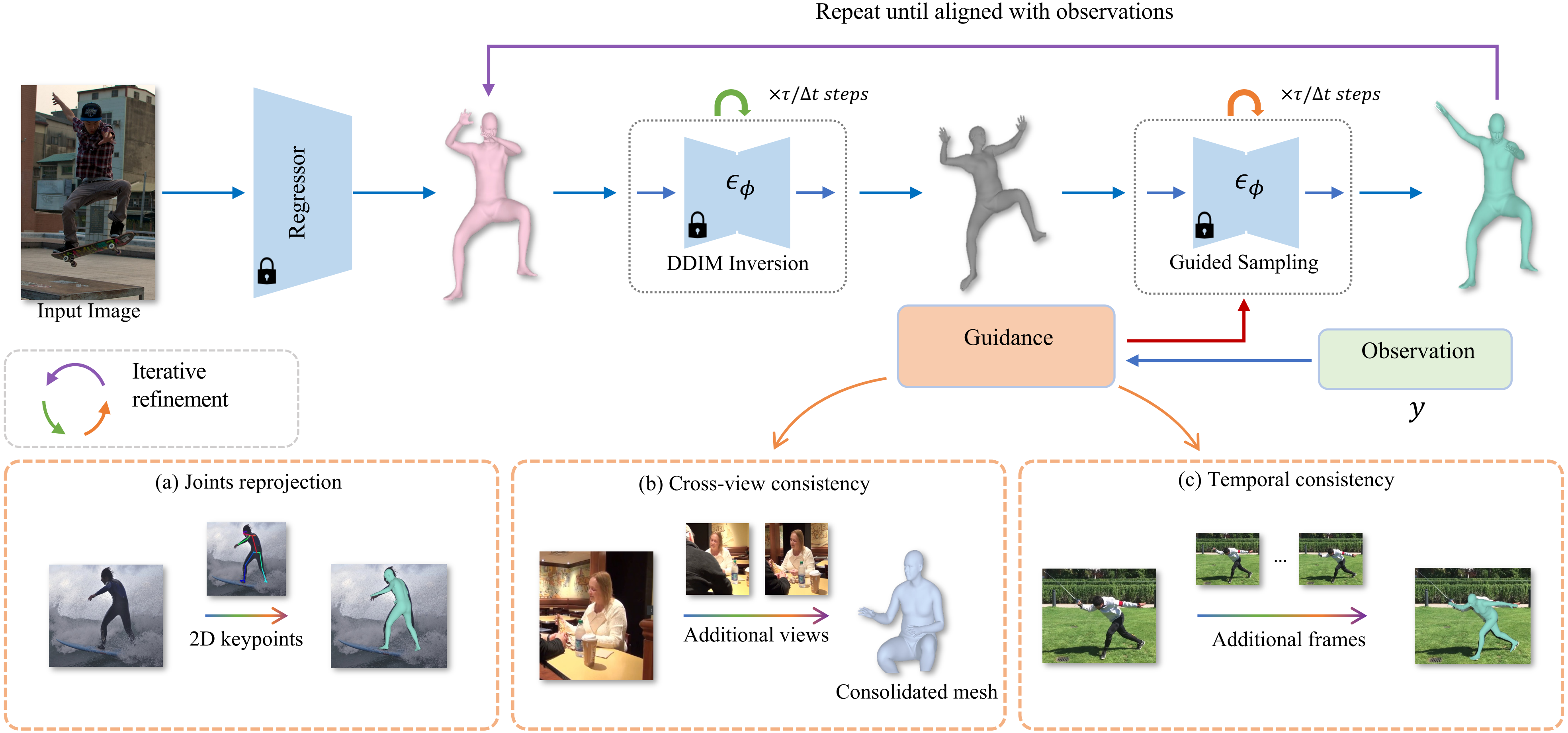
Shape Registration

Segmentation

Computer Vision
Generative/Multimodal AI
Explainable AI
Projects

Score-Guided Diffusion for 3D Human Recovery

Solving inverse problems for **3D human pose** and **shape** reconstruction with score guidance in the latent space of a diffusion model.



Social ODE: Multi-agent Trajectory Forecasting with Neural Ordinary Differential Equations

Method: Model an agent's trajectory spatial and temporal dimensions explicitly:

1. Model the temporal dimensions using Neural ODEs to learn continuous temporal dynamics
2. Model agent interactions using three variables: distance, agent interaction intensity, and agent aggressiveness.

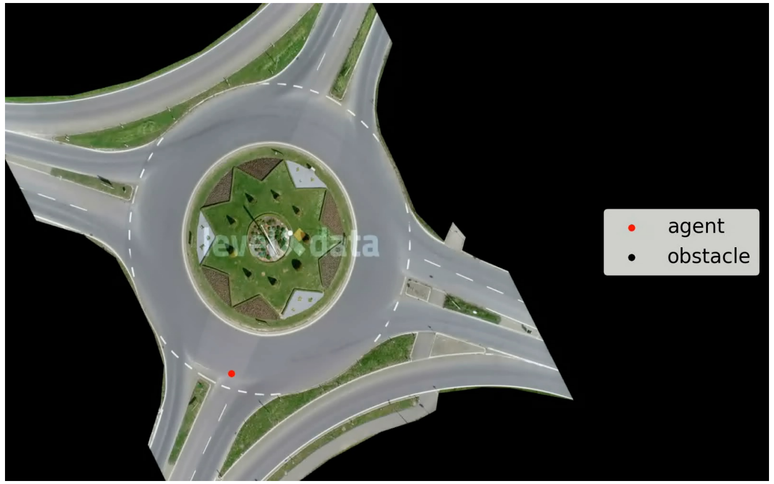
Trajectory forecasting:

Second-Order Graph ODEs for Multi-Agent Trajectory Forecasting

Method:

1. Incorporate distance and velocity information to model agent interactions by constructing dynamic interaction graphs in real-world space.
2. Model continuous temporal dynamics using second-order ODEs, following Newton’s Second Law.

Avoid obstacle



Near-miss forecasting

Diffusion Models for Sign Language Video Anonymization

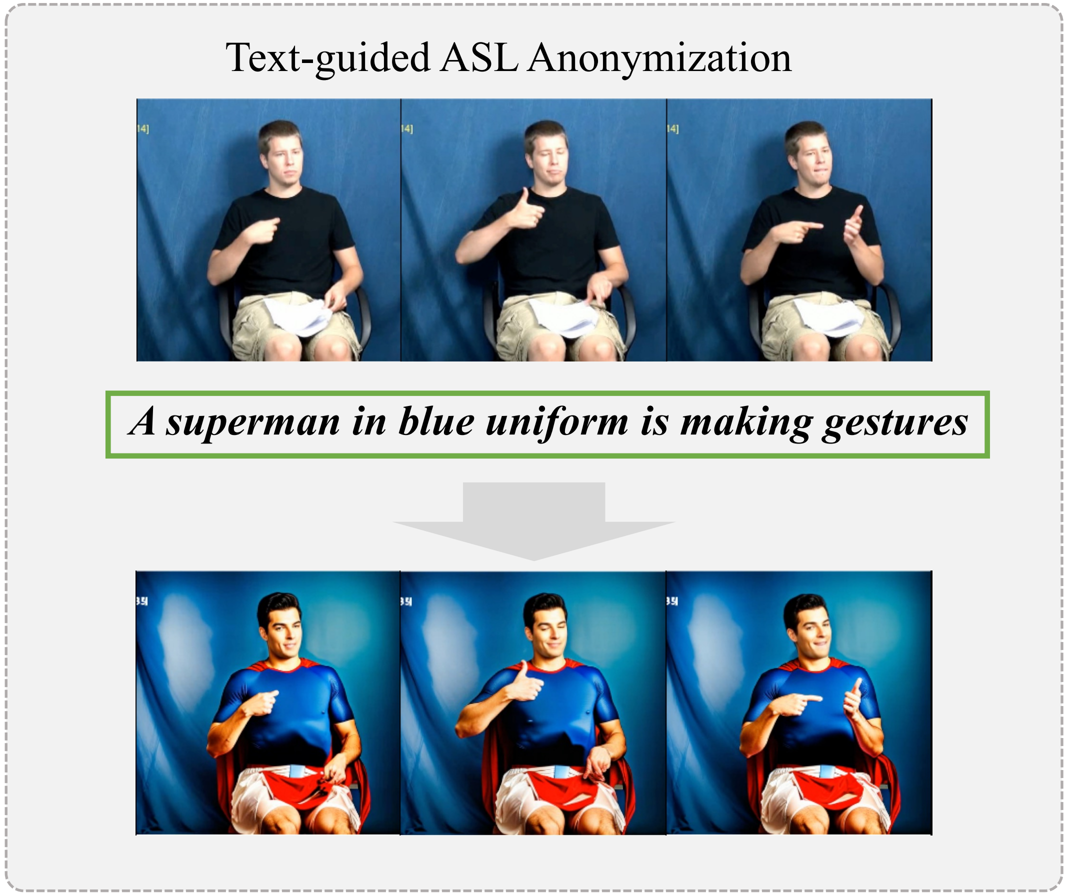
Diffusion Model, Text-to-video, Video Editing, Video Anonymization, ASL

Our research introduces DiffSLVA, a novel methodology that uses pre-trained large-scale diffusion models for text-guided sign language video anonymization.

(A) We incorporate ControlNet, which leverages low-level image features such as HED edges, to circumvent the need for pose estimation.

(B) Cross-Frame Attention Control and Optical Flow Based Guidance is applied for consistency in video editing.

(C) A specialized module based on motion estimation is developed to transfer linguistically essential facial expressions.



Sign Language Video Anonymization

Image Animation, Motion Transferring, Video Editing, Video Anonymization, ASL

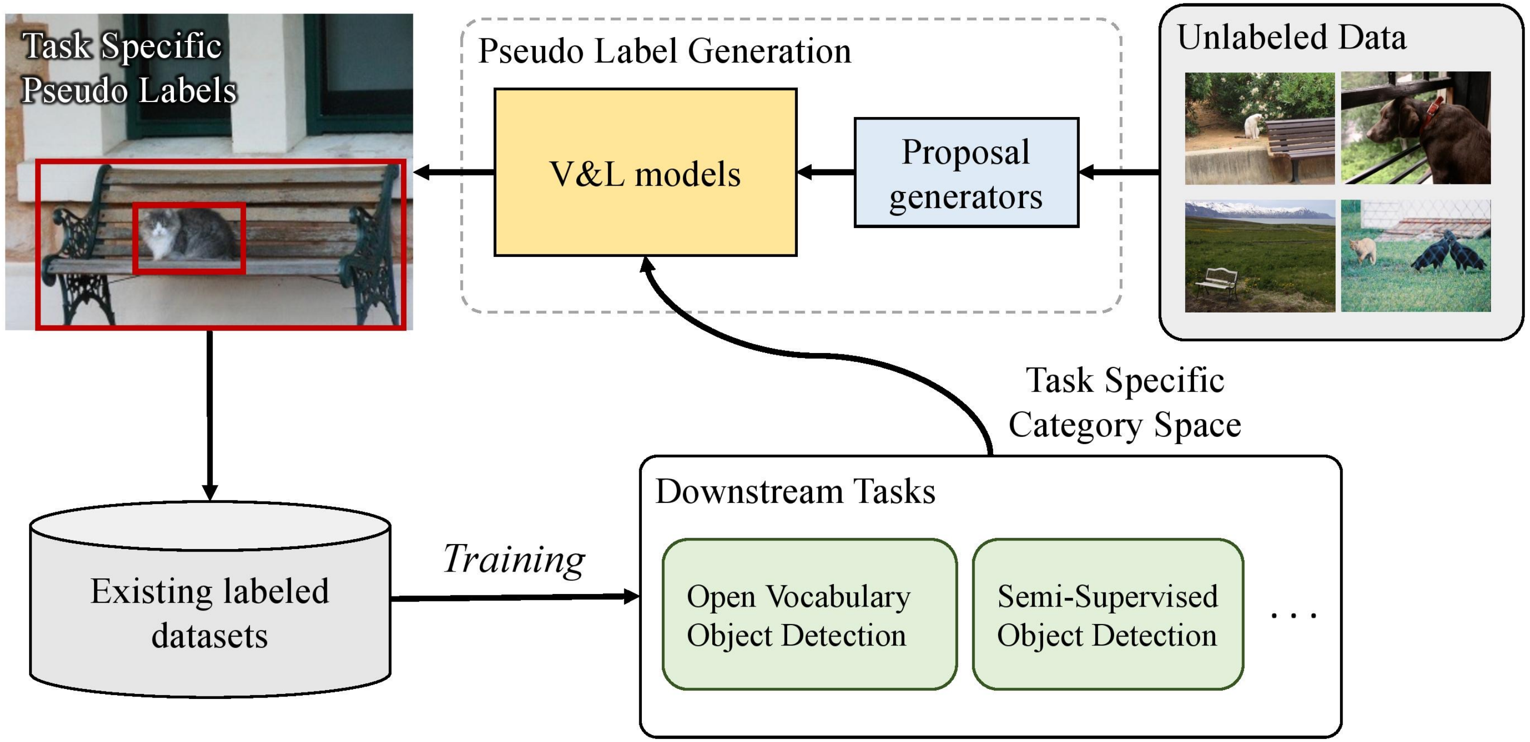
We propose to transfer a signer's identity to another signer based on the image animation model using a source frame and driving frames from ASL videos. Our contributions are:

(A) The Asymmetric Image Generator: An encoder-decoder structured network with a High-Resolution Generation **(HRG)** module for high quality and low computation cost image generation.

(B) Multiscale perceptual loss based on VGG-16 & **Hand & Face Focused Loss**: Computed between the high-resolution generated and driving frames, to improve face & hand generation.

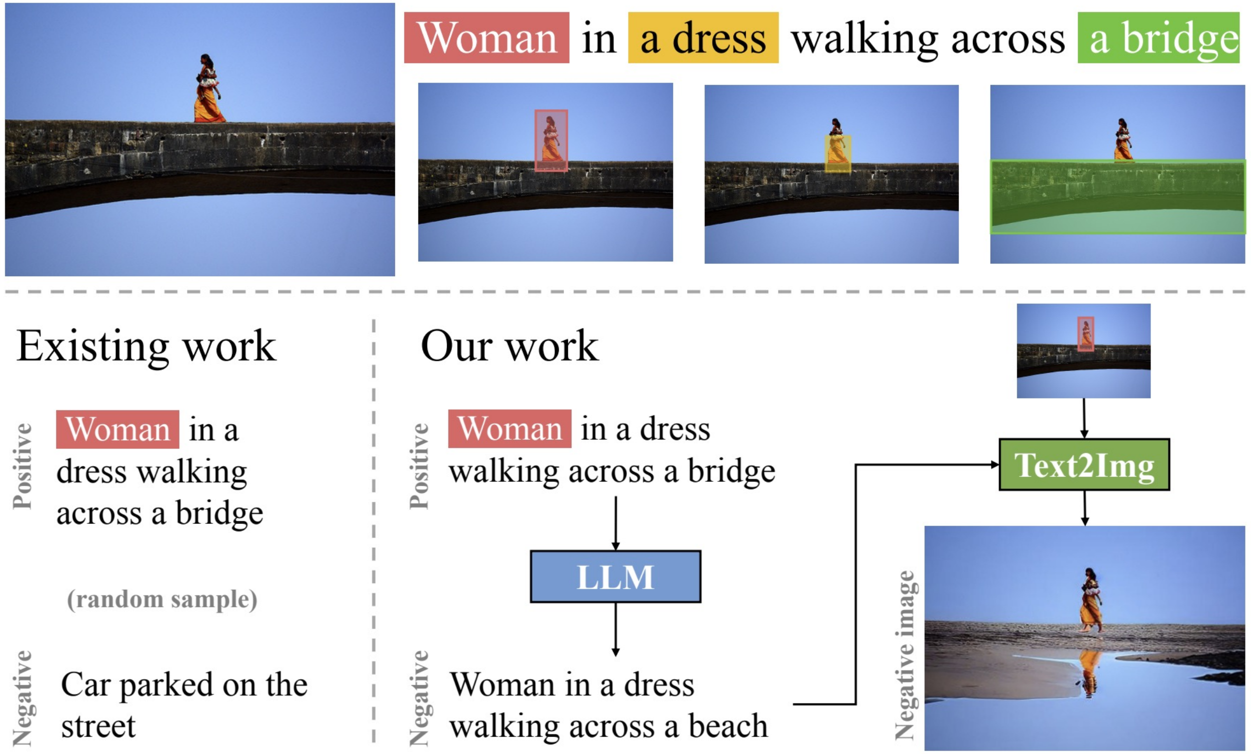
Exploiting Unlabeled Data with Vision and Language Models for Object Detection

Open vocabulary object detection, vision and language models, pseudo label generation



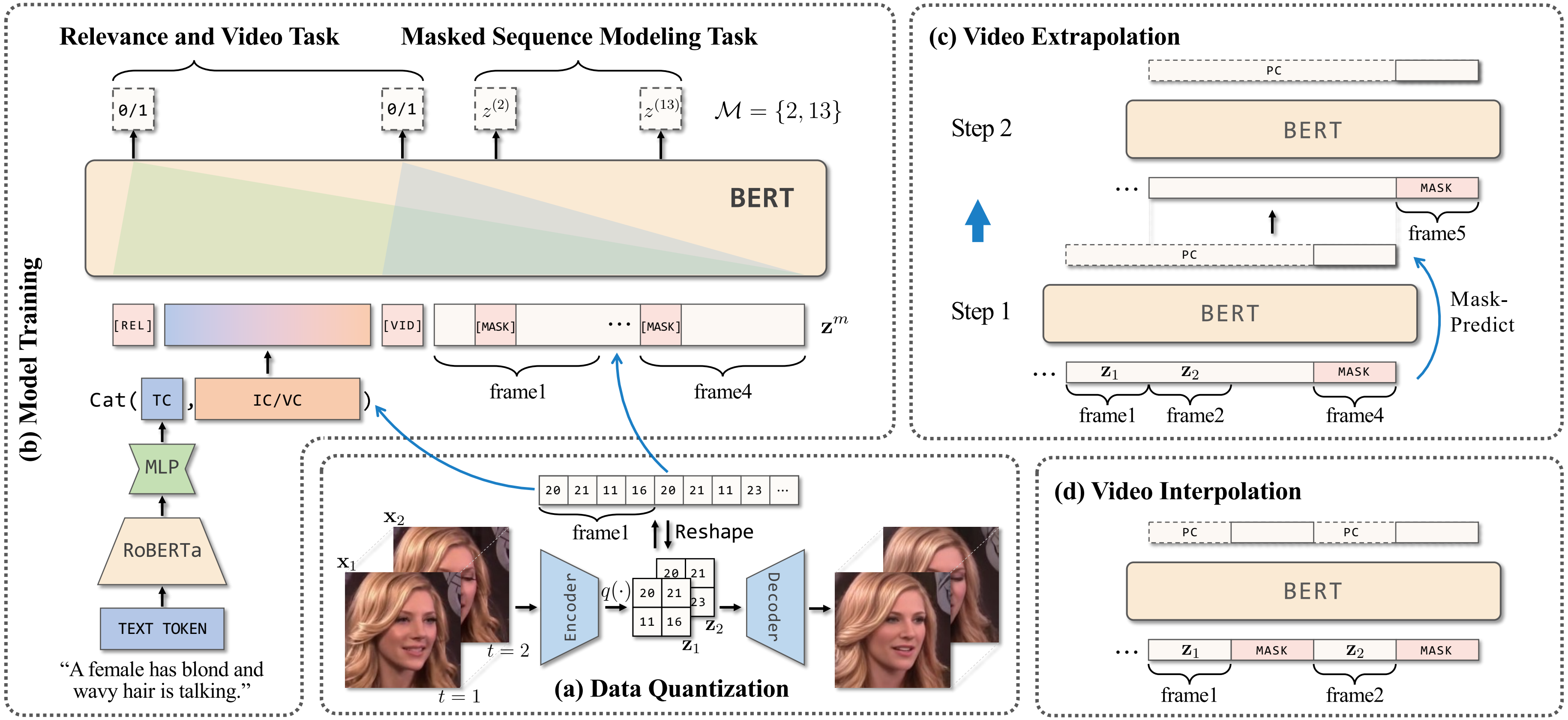
Generating Enhanced Negatives for Training Language-Based Object Detectors

Open vocabulary object detection, multi-modal learning, large language model, data augmentation



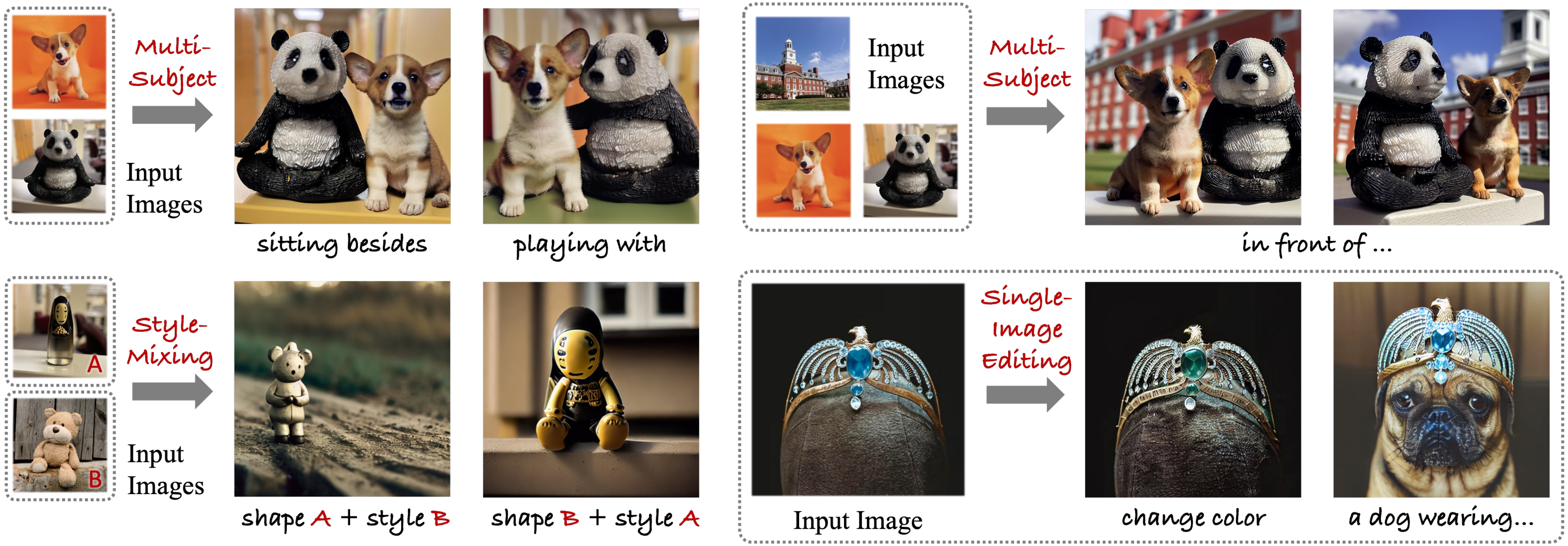
Show me what and tell me how

Video synthesis via multimodal conditioning



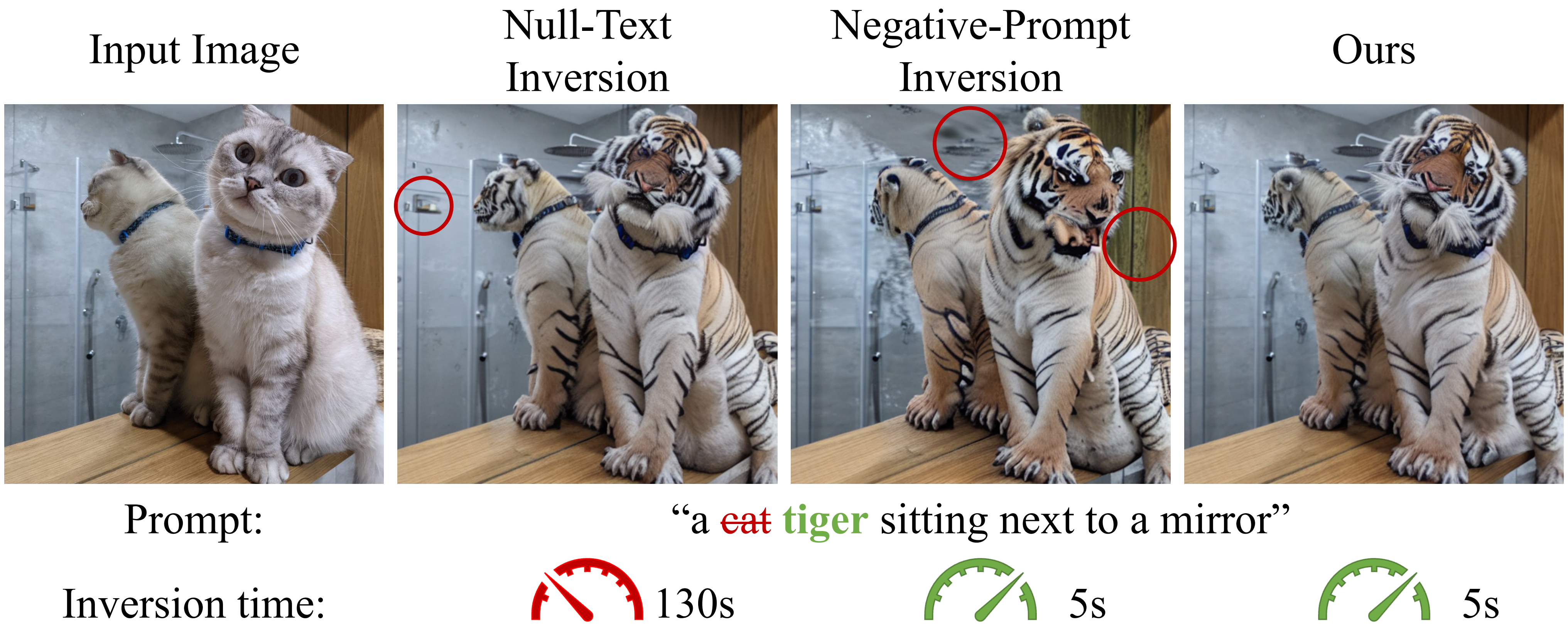
SVDiff: Compact Parameter Space for Diffusion Fine-Tuning

Parameter-efficient fine-tuning method for GenAI models (*e.g.* text-to-image diffusion)



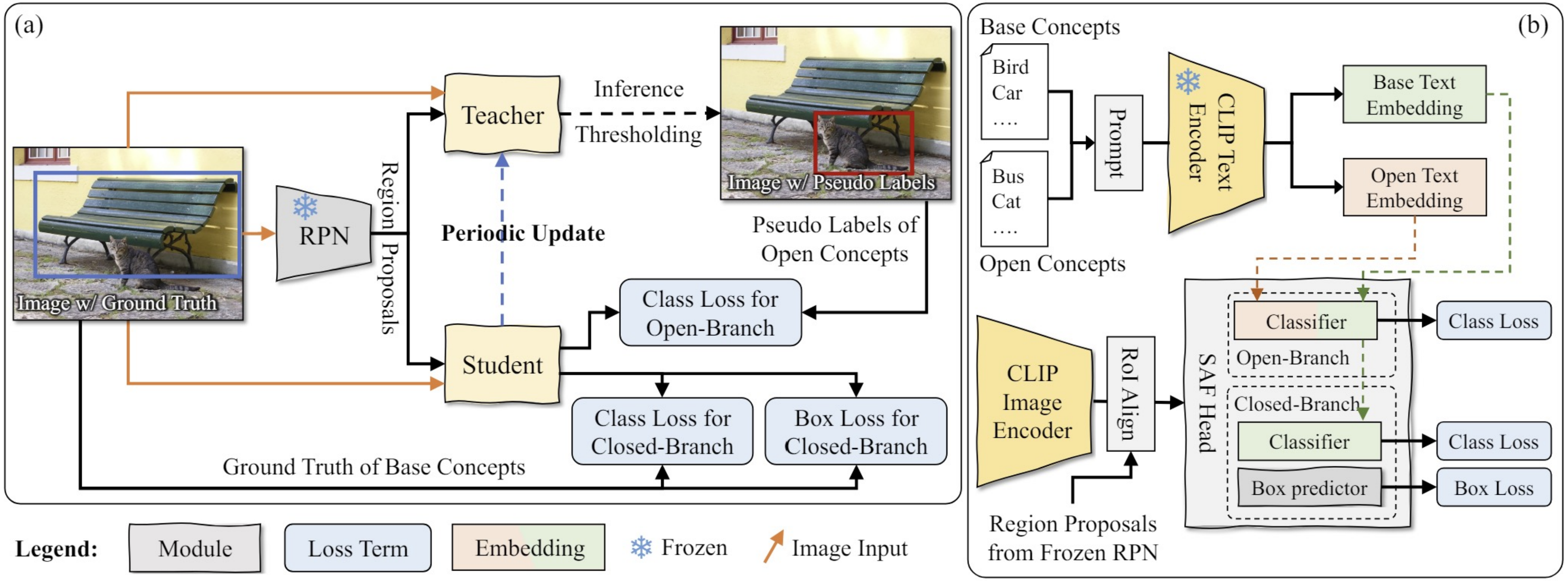
ProxEdit: Improving Tuning-Free Real Image Editing with Proximal Guidance

Tuning-free closed-form diffusion-based optimization for image editing



Self-Training for Open-Vocabulary Object Detection

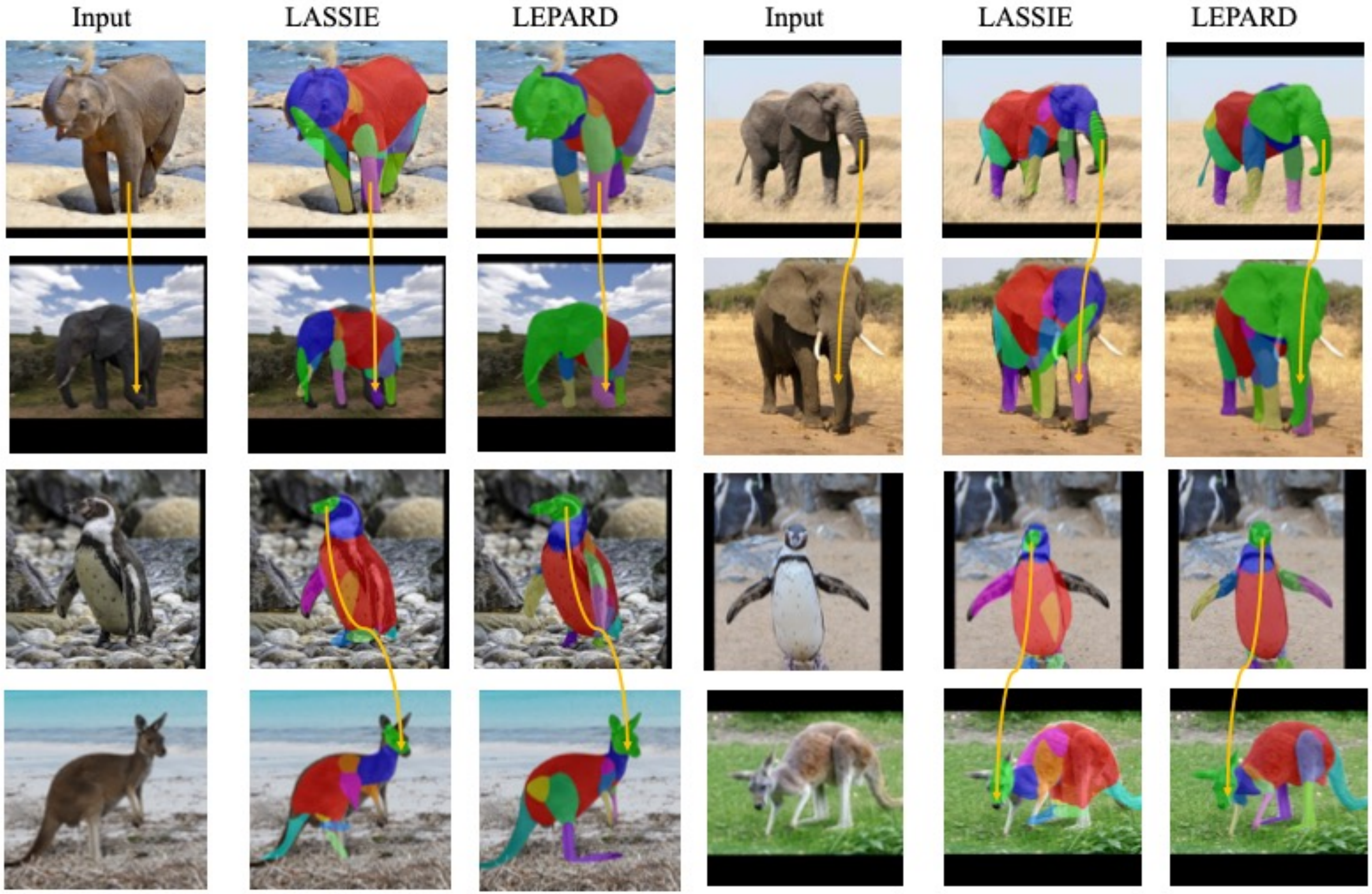
Open vocabulary object detection, vision and language models, self-training



LEPARD: Learning Explicit Part Discovery for 3D Articulated Shape Reconstruction (NeurIPS 2023)

- To generalize to scenarios where 3D annotations are not available, we
  - Illustrate the relationship of kinematics between 3D and 2D via projective geometry.
  - Project the primitives onto the image space and calculate the discrepancy between the projected primitives and 2D evidence.
  - Convert the image forces to their corresponding generalized forces that guide the deformation of the primitives.
  - Use deep features from DINO-ViT as supervision to train our model.

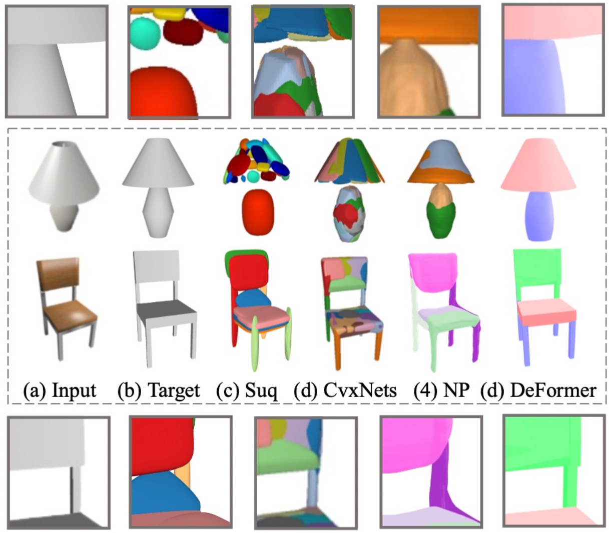
Results – Consistency Visualization



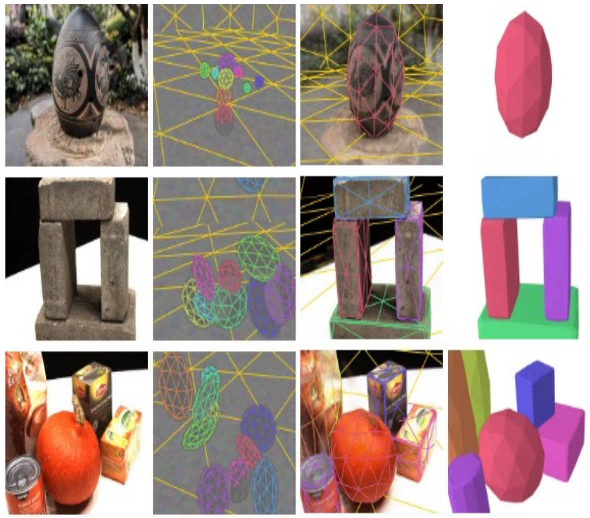
Deep Physics-based Deformable Models

Topic: 3D Scene Understanding and Object Shape Abstractions

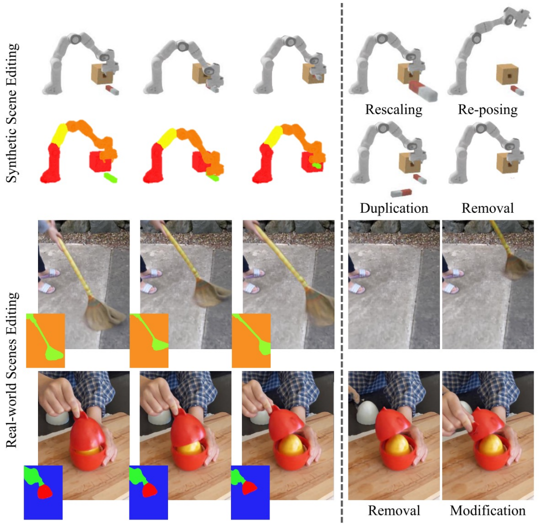
Applications:



Shape Reconstruction



Scene Understanding



Novel-view Synthesis/Editing

Thank You!

carta.rutgers.edu
dnm@cs.rutgers.edu