Final Project

Your task:

(i) **read** and understand a research paper,

(ii) **implement & reproduce** (a part of) the paper,

(iii) try a small **extension to the existing method**,

(iv) perform qualitative/quantitative **experimental evaluation**

Deliverables:

(a) 1-page proposal (Nov 19): read papers, choose topic, make detailed plan (will receive feedback)

(b) Presentation (Jan 6): brief introduction, clarify your contributions, results, visuals

(c) Report (Jan 13): 3-page description plus visuals and bibliography

Typically requires multiple weeks of intense work, therefore, starting early is recommended. Start now if you can. Wait for TAs announcement to claim your Google Cloud Credits.

Communicate with your project supervisor!

Final Project Topics

[A] Composed Image Retrieval

[B] 3D Human Motion Generation

[C] Generalizable Vision-Language Robotic Manipulation

[D] TokenCompose: Text-to-Image Diffusion with Token-Level Supervision

[E] Test-Time Training with Masked Autoencoders

[X] Your own chosen topic

Final Project Topics

[A] Composed Image **Retrieval**

[B] **3D Human** Motion **Generation**

[C] Generalizable Vision-Language **Robotic** Manipulation

[D] TokenCompose: **Text-to-Image Diffusion** with Token-Level Supervision

[E] Test-Time Training with **Masked Autoencoders**

[X] Your own chosen topic

Final Project Topics

[A] Composed Image Retrieval

[B] 3D Human Motion Generation

[C]

[D]

[E]

[X]
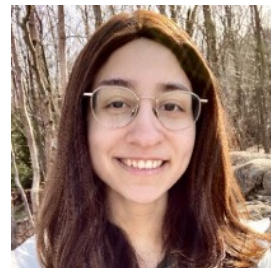
Lucas Ventura
IMAGINE

put this dog in a tropical back yard

CoVR: Learning composed video retrieval from web video captions.
Lucas Ventura, Antoine Yang, Cordelia Schmid, and Gul Varol. AAAI, 2024.

Final Project Topics

[A] Composed Image Retrieval

[B] 3D Human Motion Generation
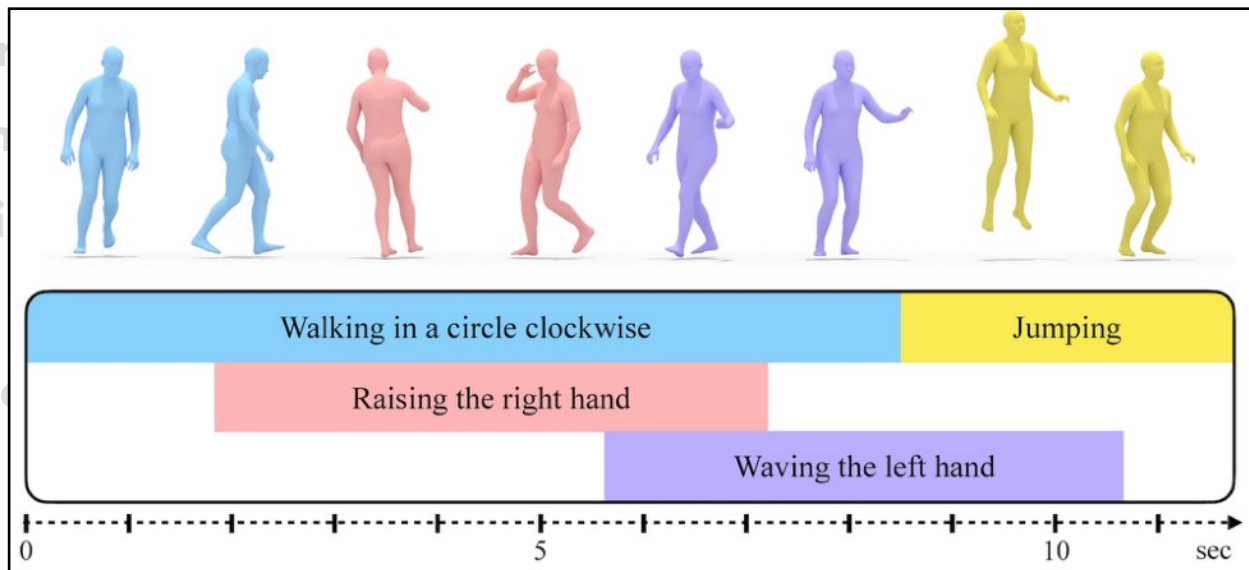
[C] Gener

[D] Token   Supervision

[E] Test-Ti

[X] Your c

Léore Bensabath
IMAGINE

A Cross-Dataset Study for Text-based 3D Human Motion Retrieval.

Léore Bensabath, Mathis Petrovich, and Gül Varol. CVPRW 2024.

Final Project Topics

[C] Generalizable Vision-Language Robotic Manipulation

Ricardo Garcia-Pinel
WILLOW
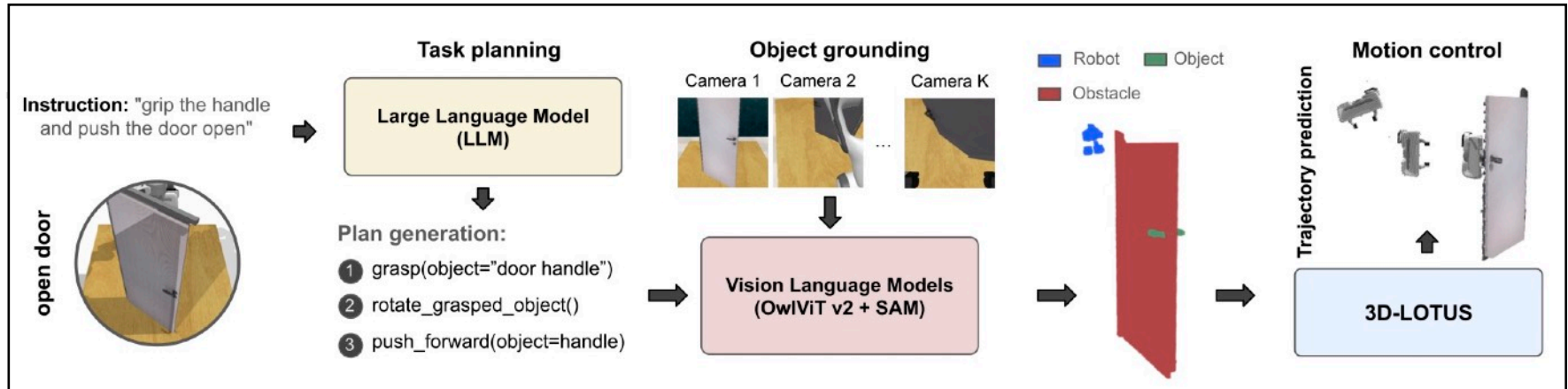
Towards Generalizable Vision-Language Robotic Manipulation: A Benchmark and LLM-guided 3D Policy.
Ricardo Garcia-Pinel, Shizhe Chen, and Cordelia Schmid. CoRL 2023.

Final Project Topics

[A] Composed Image Retrieval

[B] 3D Human Motion Generation
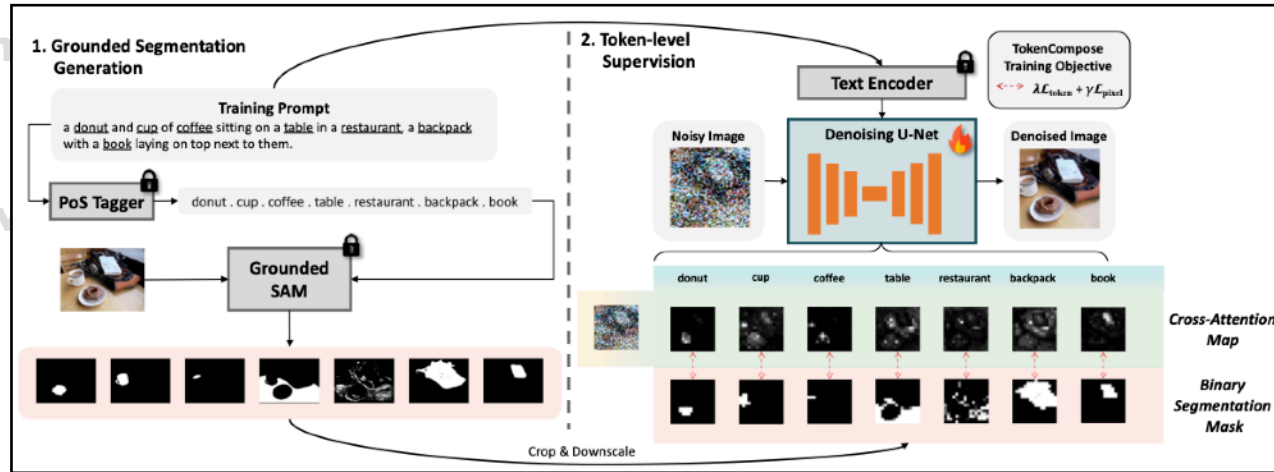
[C] Generalizable Vision-Language Robotic Manipulation

[D] TokenCompose: Text-to-Image Diffusion with Token-Level Supervision

[E] Test-Tim

[X] Your ov

Zeeshan Khan
WILLOW

Nicolas Dufour
IMAGINE

TokenCompose: Text to Image Diffusion with Token Level Supervision, Zirui Wang et al. CVPR 2024.

Final Project Topics

[A] Composed Image Retrieval

[B] 3D Human Motion Generation

[C] Generalizable Vision-Language Robotic Manipulation

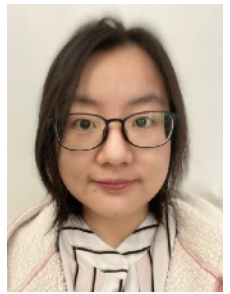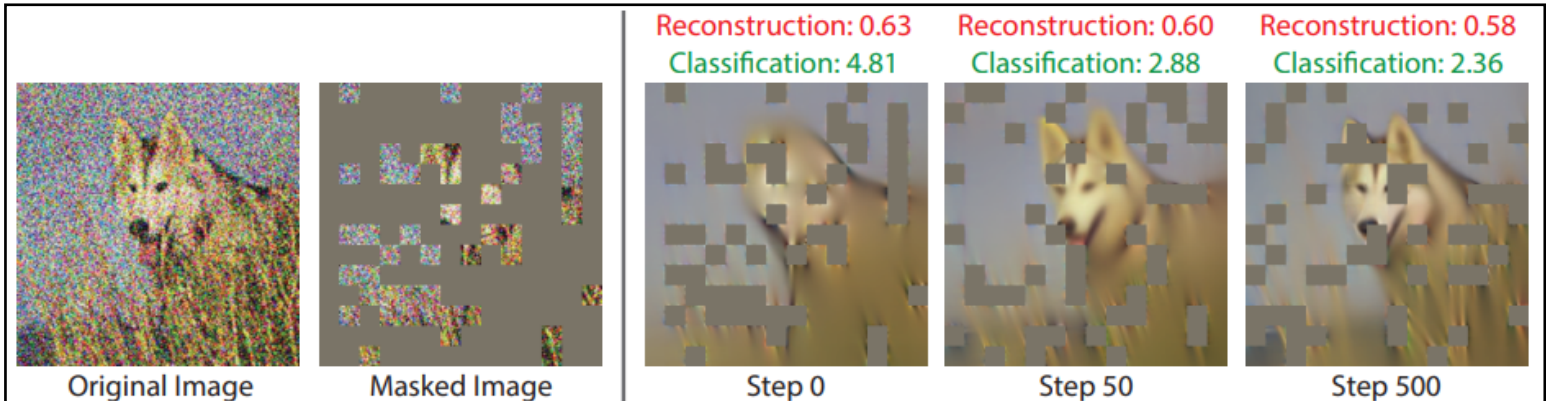[D] TokenCompose: Text-to-Image Diffusion with Token-Level Supervision

[E] Test-Time Training with Masked Autoencoders

[X] Your

Shizhe Chen
WILLOW

[Test-Time Training with Masked Autoencoders](). Gandelsman et al., NeurIPS 2022.

Final Project Topics

[A] Composed Image Retrieval

[B] 3D Human Motion Generation

[C] Generalizable Vision-Language Robotic Manipulation

[D] TokenCompose: Text-to-Image Diffusion with Token-Level Supervision

[E] Test-Time Training with Masked Autoencoders

[X] Your own chosen topic
- A paper discussed in the class, topics from previous years, your personal interests…
- At your own risk, i.e., supervision might be suboptimal.
- Validate with me beforehand by sending an email!

Final Project Topics

[A] Composed Image **Retrieval**

[B] **3D Human** Motion **Generation**

[C] Generalizable Vision-Language **Robotic** Manipulation

[D] TokenCompose: **Text-to-Image Diffusion** with Token-Level Supervision
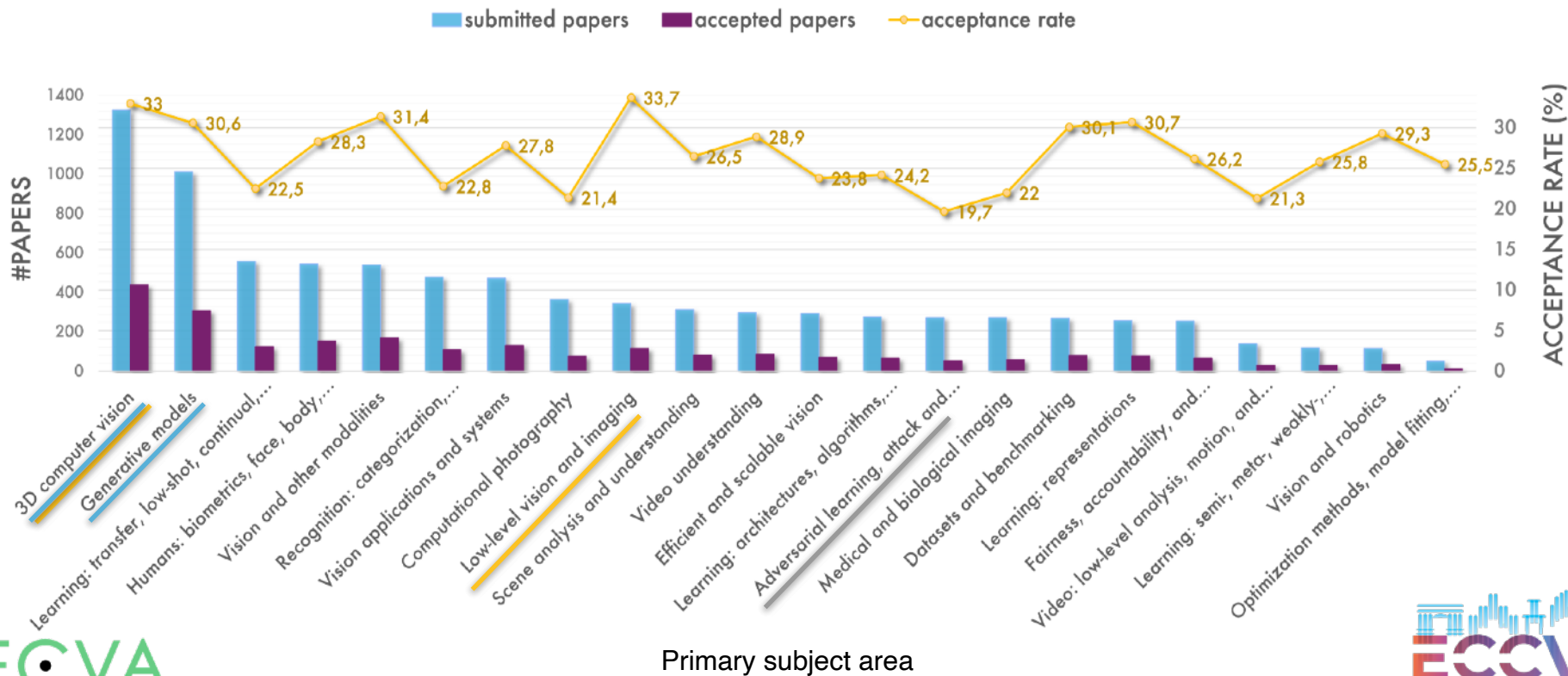
[E] Test-Time Training with **Masked Autoencoders**

[F] Personalization of CAD

[X] Your own chosen topic

Q?

ECCV 2024 Subject Areas Distribution

Source code adapted from Christopher Funk

2387 (100%)
Primary Categories
Selected