

Input Video



...



**VEGGIE: A Unified
Versatile Instructional
Video Generative Model**



+ [Concept Addition]



Instruction: Add a hat to the little girl.

🔍 [Concept Grounding]



Instruction: Highlight the woman in front of them with red masks.

? [Reasoning Segmentation]



Instruction: What can be used to contain food? Highlight with red masks.

🎨 [(Multimodal) Stylization]



Instruction:  Change to the reference image style.

- [Concept Removal]



Instruction: Remove the woman.

🌍 [Global Concept Editing]



Instruction: Change it to be in the forest.

📍 [Local Concept Editing]



Instruction: Change the bowl to pizza.