NSF REU 2026
Week 8

Tiffany Marcus

16 July 2026

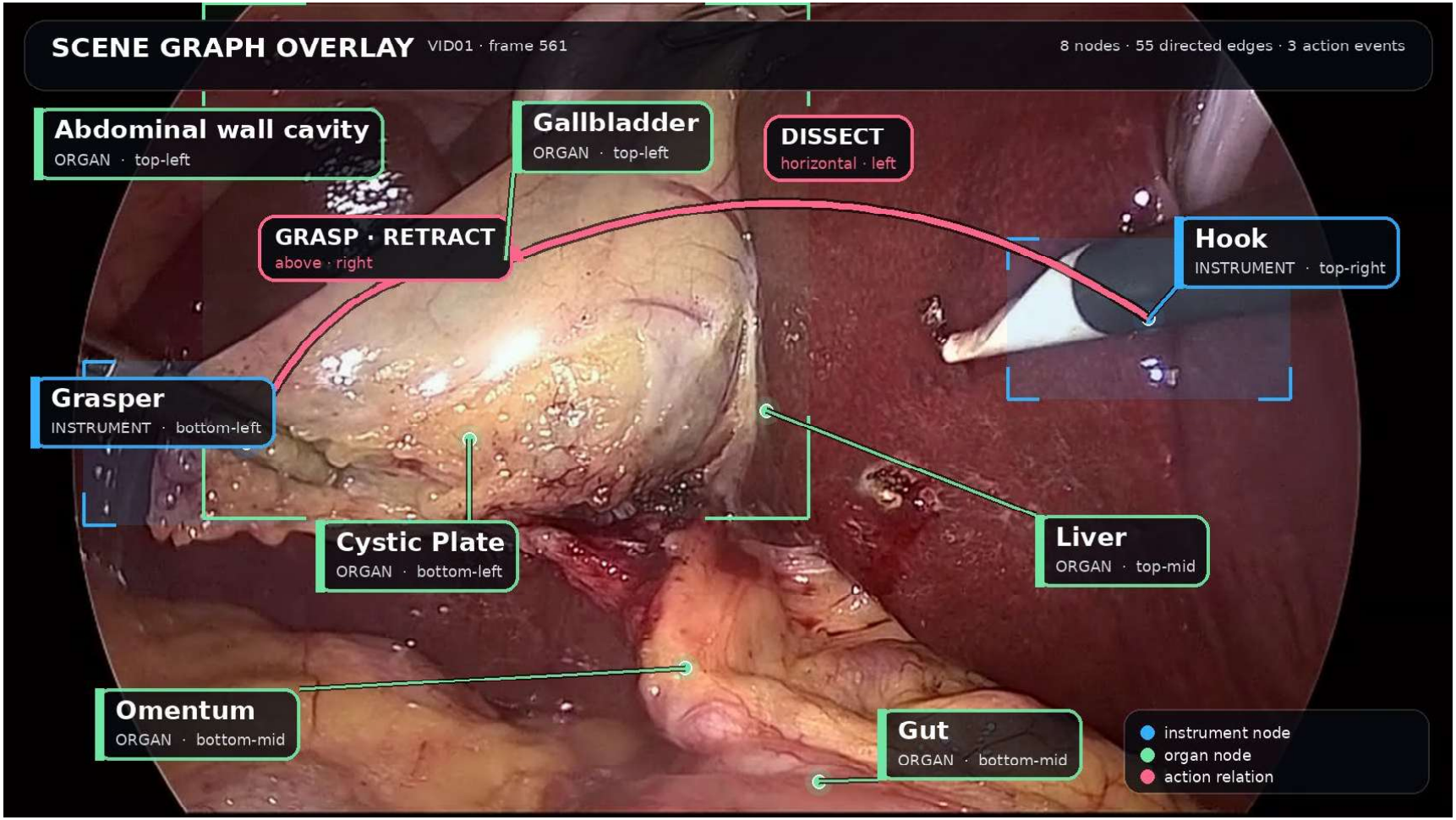
This week:

- Information about the dataset
- Generate simple scene graphs from ground truth, one for reach frame
- Train LLM-0 (train/test: 40/5) using LoRA
- Test VQA (ground truth from SSG-VQA)

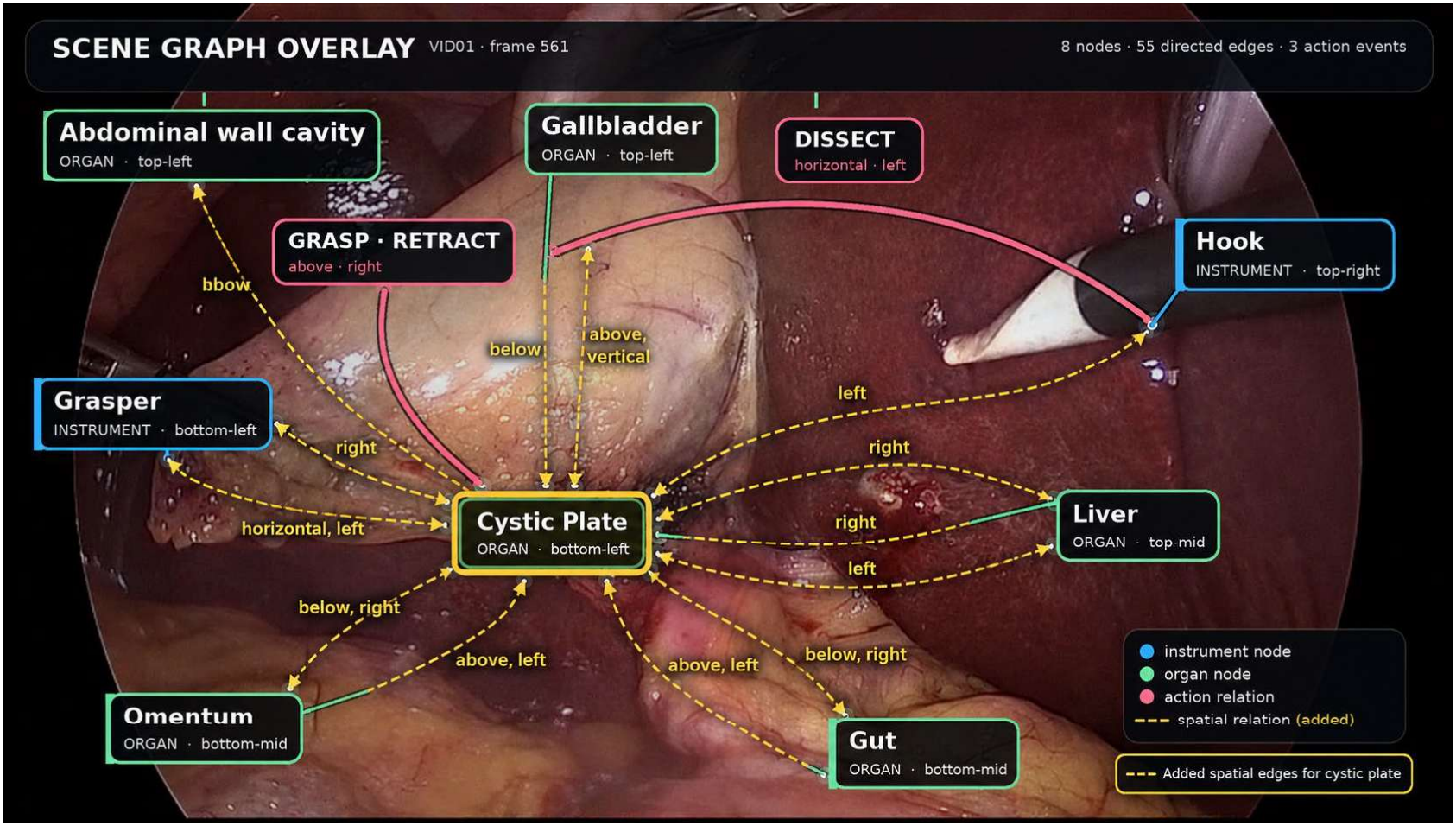
Dataset

- SSG-VQA builds scene graph using the triplets in CholecT45 (<instrument, verb, target>) and gets spatial data from image segmentation
- We combined the newly added annotations from SSG-VQA with the original triplets from CholecT45
- 45 videos, 1 subject per video
- 982 – 3,713 frames, average of ~1,200 frames/video, 1 frame/second

Visualized Scene Graph



Visualized Scene graph



Next Steps

- Week 8
  - Add more information about the custom dataset
  - Generate simple scene graphs from ground truth, one for reach frame
  - Train LLM-0 (train/test: 40/5) using LoRA
  - Test VQA (ground truth from SSG-VQA)

```
1%|| | 559/65341 [14:54<27:34:52,  1.53s/it]
```

- Week 9
  - From the existing ground truth, generate a template for adding temporal questions
  - Train LLM-t (train/test: 40/5)
  - Test VQA-t

- Week 10
  - Write report
  - Create poster
  - Presentation