Week 7 Progress Report

Generative Satellite Alignment for Cross-View Geolocalization

Student: Wyatt Petula **Mentors**: Soumik Ghosh, Qingyang Liu

-- Fine-Tuning GAReT from a Different Approach --

Terminology moving forward



Reference Image

WRIVA Challenge



F2F Patch

Frame-to-Frame,
“Small Aerial Gallery”

S2I Patch

Sequence-to-Image,
GAReT Reference Image



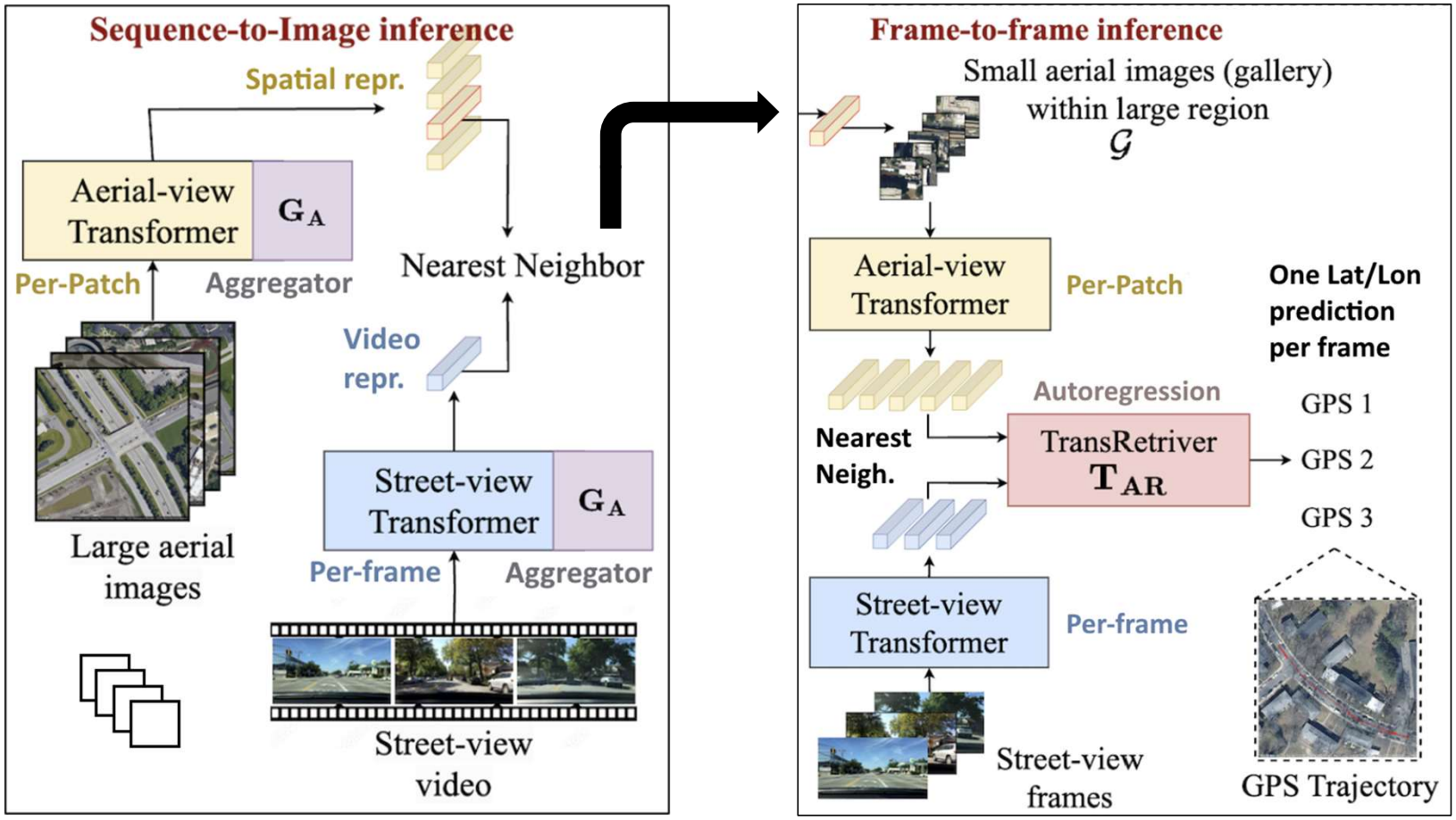
Query

Frames and Videos



https://www.magnific.com/free-photo/street-with-modern-buildings-green-trees_10352233.htm#fromView=keyword&page=1&position=0&uuid=67e4314b-0e43-4223-92dd-3aab1906185a&query=City+street+with+urban+landscape

Recap of GAReT's architecture



Modified from https://eccv.ecva.net/media/PosterPDFs/ECCV%202024/2461.png?t=1726216876.7051601. Credit to Pillai M., et. al.

Recap of my goal

WRIVA_Challenge_Dataset
|--Site_001
|----sequential_query_frames



|----reference_images



|--Site_002
|-- ...

← Reconcile →

GAReT (using GAMa)



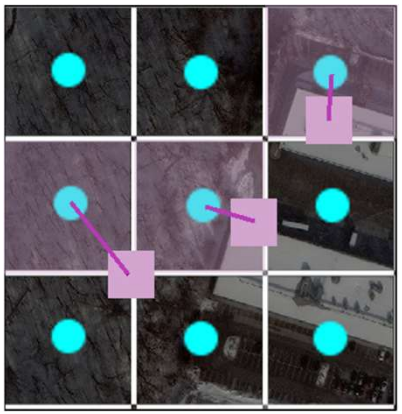
Query videos



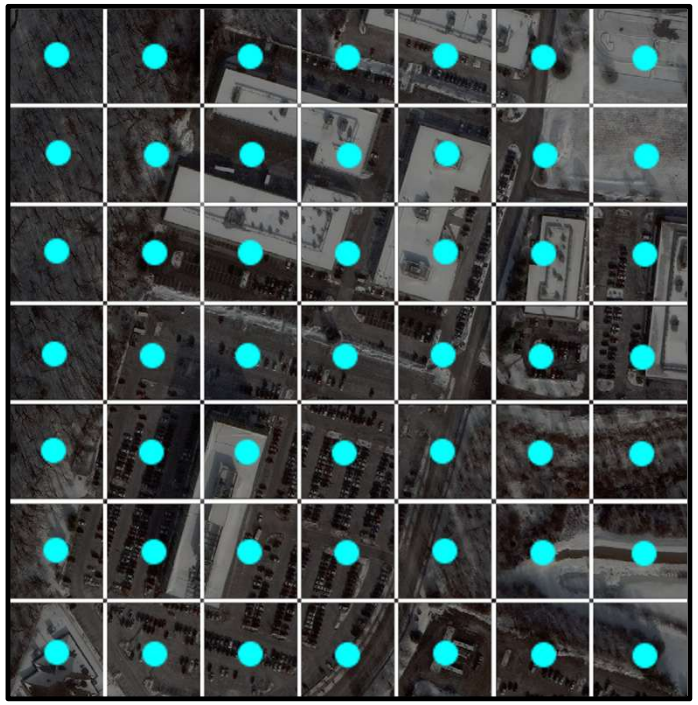
Reference images

My previous adaptation of GAReT created more problems than it solved

Garet's
F2F
Gallery



7x7 patches, 256 x 256 pixels
@ ~0.24 meters/pixel



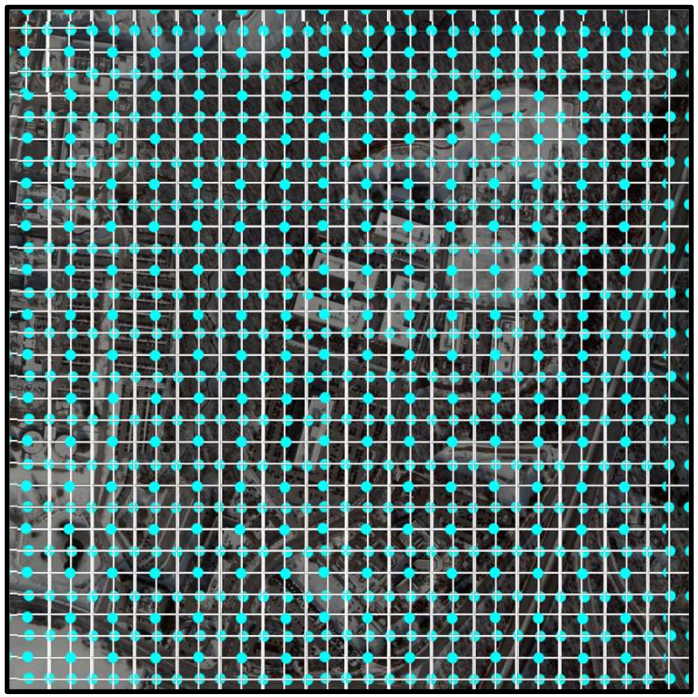
Problems:

- ***30-meter error***
- ***Different spatial resolution***

My previous adaptation of GAReT created more problems than it solved

16x16 patches, 128 x 128 pixels, 50% overlap
@ ~0.5 meters/pixel

Challenge's
F2F
Gallery



15-meter error!
But…

- No coarse retrieval

- Too many F2F patches

My current adaptation remedies those issues



WRIVA Challenge References:
2048 x 2048 pixels



S2I Patches:
512 x 512 pixels,
50% stride



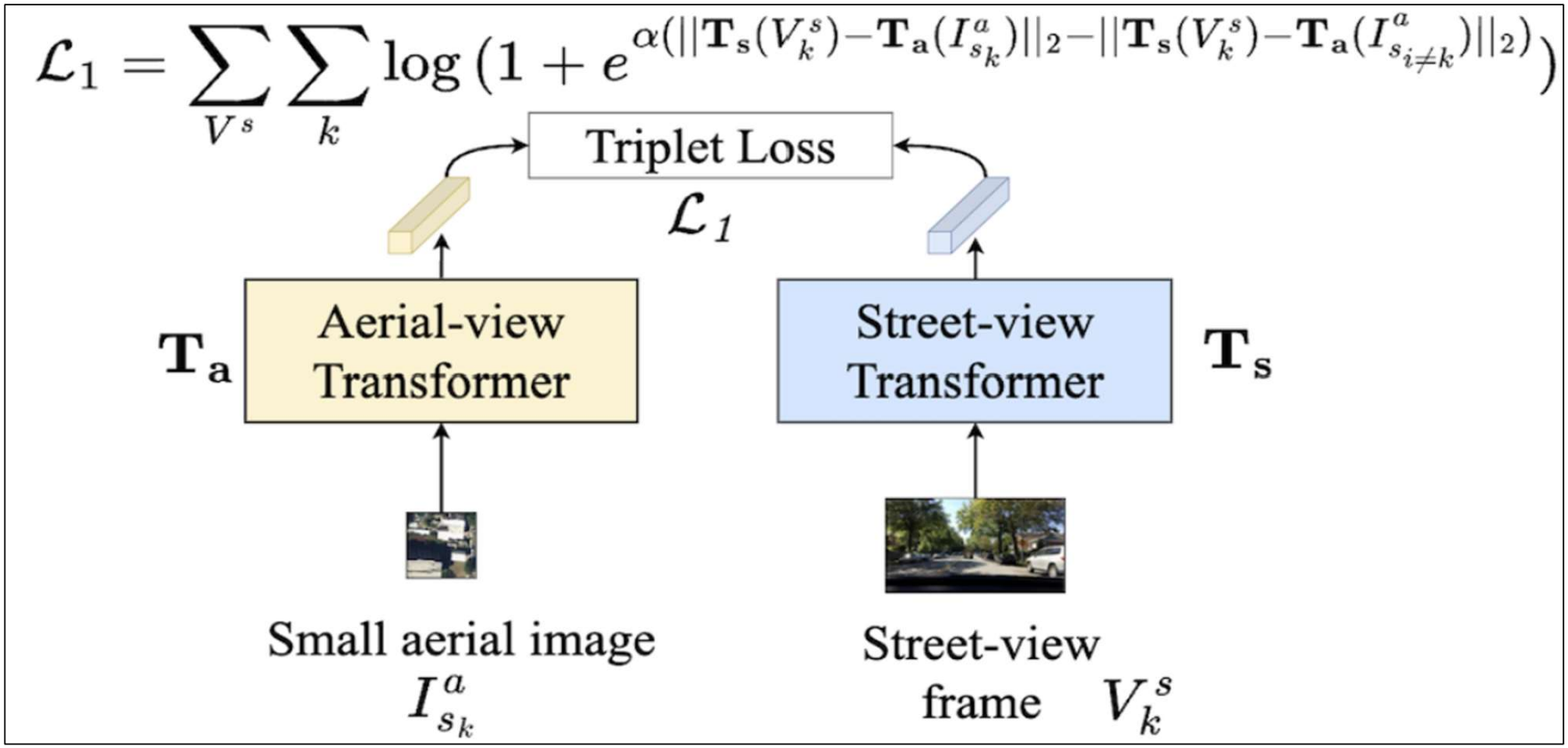
F2F Patches:
128 x 128 pixels,
50% stride

- 15-meter error

- Coarse Retrieval

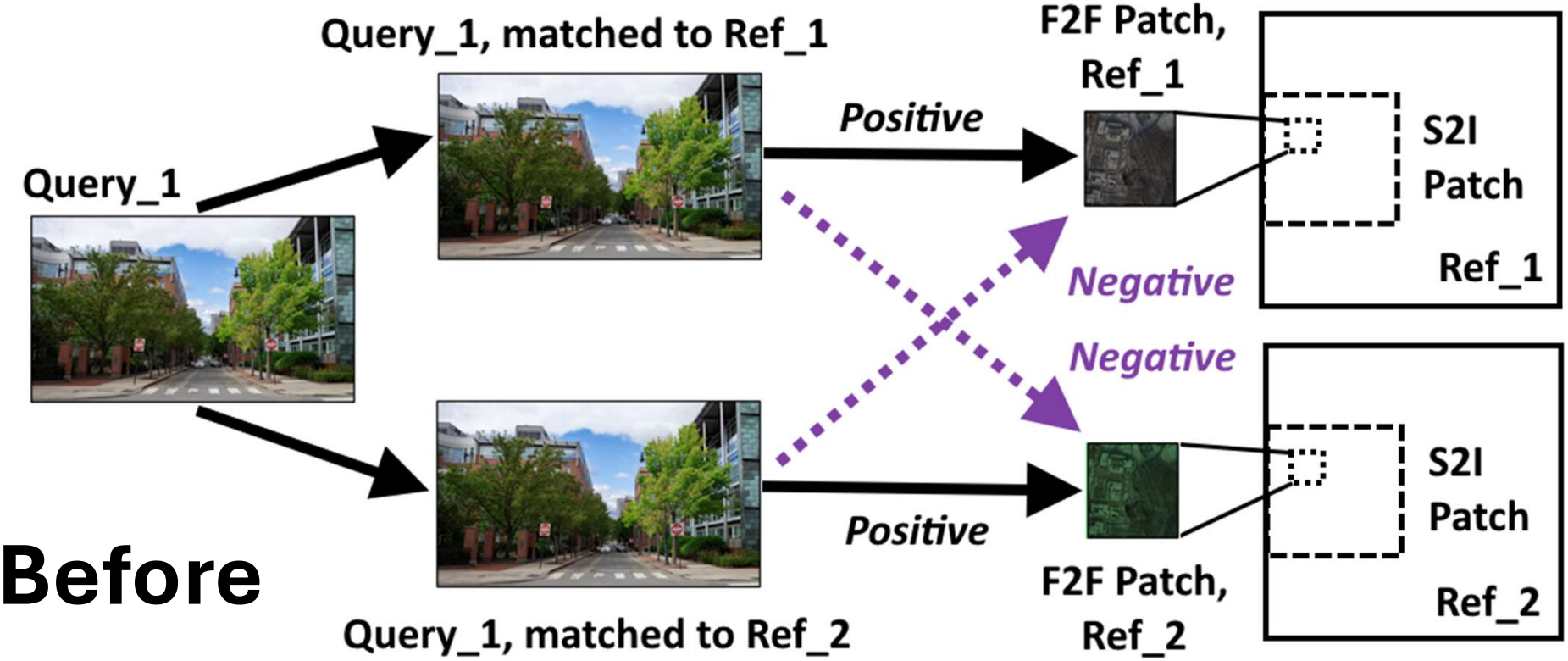
- Fewer F2F Patches

I am fine-tuning the two image encoders



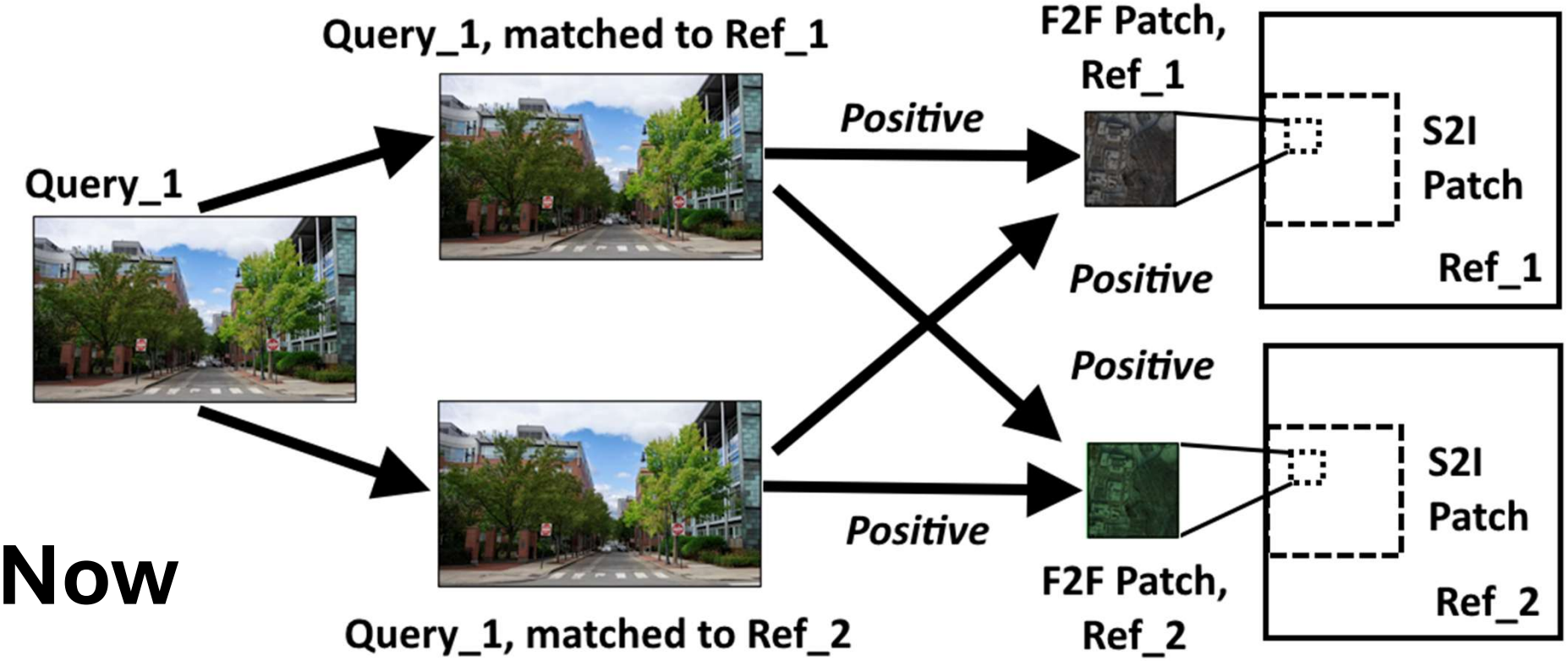
Taken from https://eccv.ecva.net/media/PosterPDFs/ECCV%202024/2461.png?t=1726216876.7051601. Credit to Pillai M., et. al.

I previously used single-positive triplet loss for fine-tuning the encoders



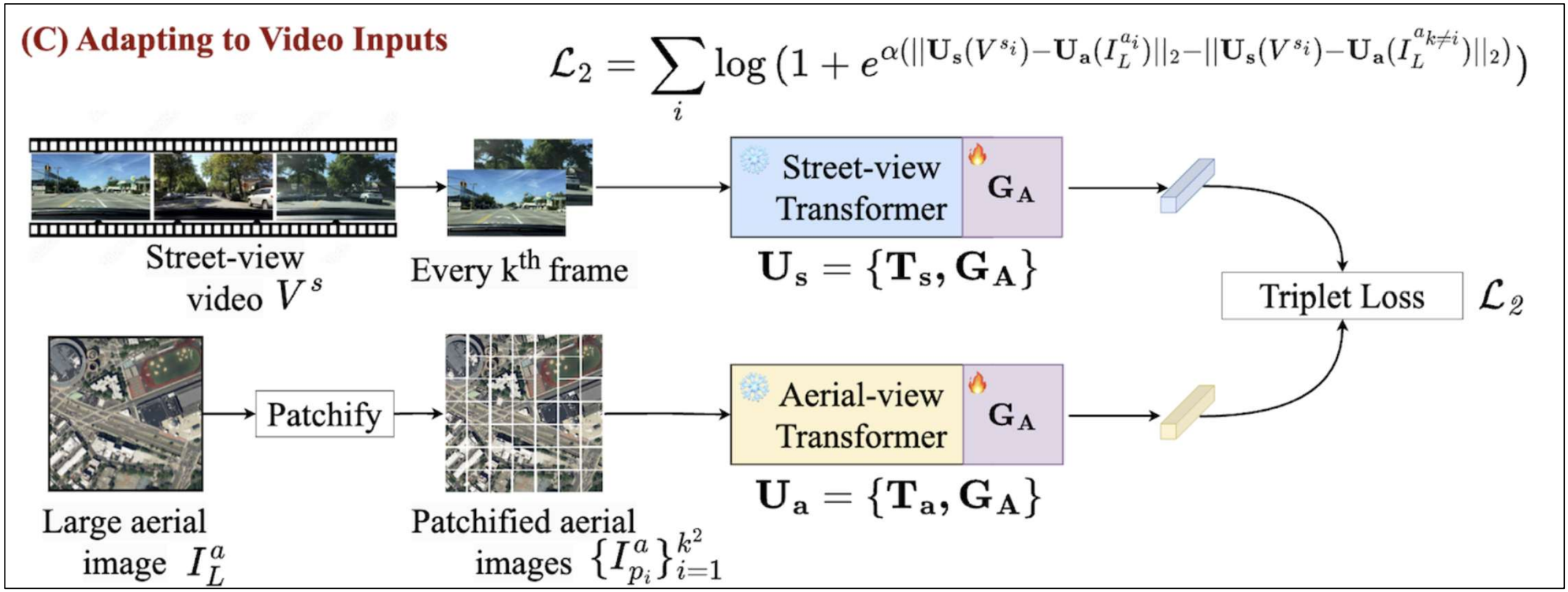
Before

Multi-positive triplet loss better aligns with the challenge



Now

I am also fine-tuning GeoAdapter using the new approach



Taken from https://eccv.ecva.net/media/PosterPDFs/ECCV%202024/2461.png?t=1726216876.7051601. Credit to Pillai M., et. al.

Goals for this/next week

- Finish fine-tuning S2I using my new approach
- If results are promising, fine-tune TransRetriever
- Swap out backbones for newer ones