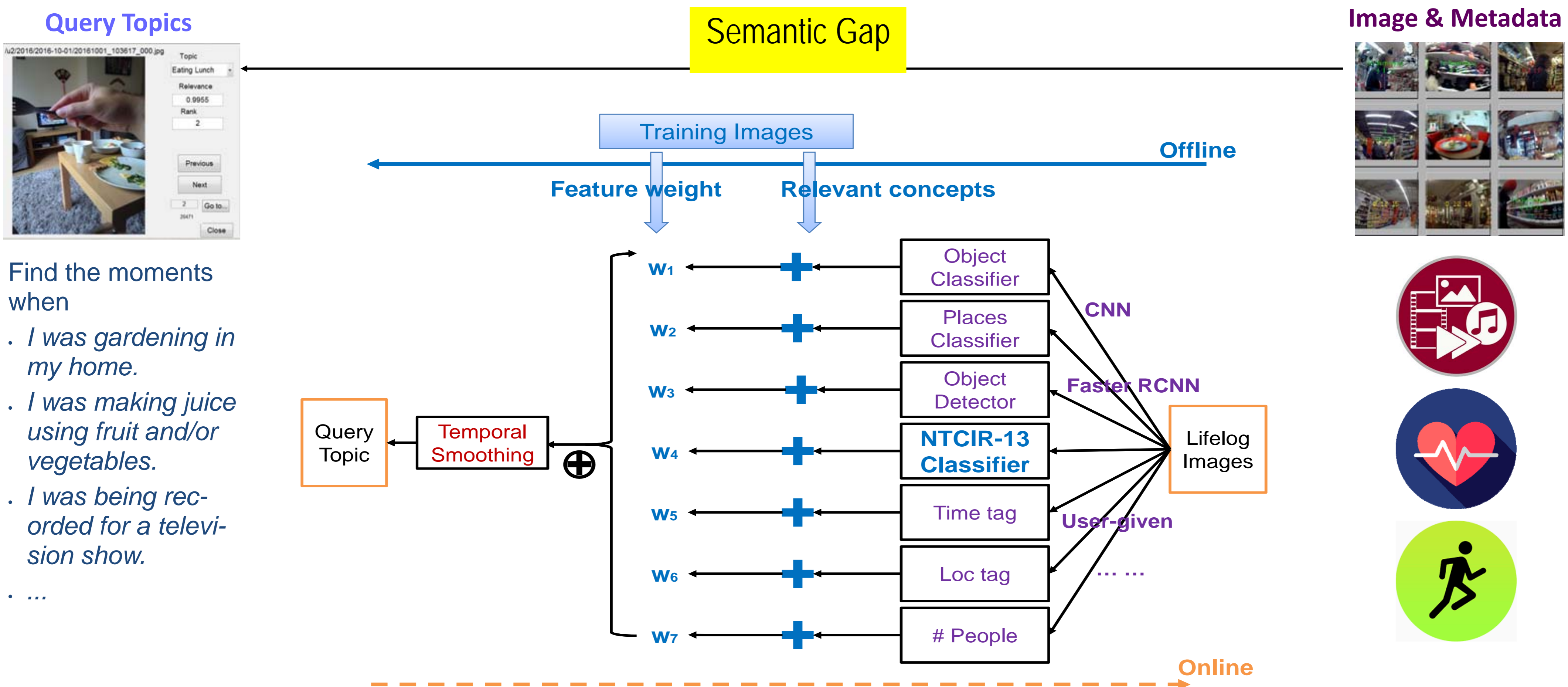


Semantic features extracted from deep learning is not enough for effective lifelog semantic access.

To bridge the semantic gap, we need

- **Relevant concepts:** *What are the basic semantics relevant to query topics?*
- **Feature weighting:** *Which feature contributes most to the query?*
- **Temporal smoothing:** *How to ensure temporal coherence?*
- **Post filtering:** *How to refine search using location and time?*



Find the moments when

- I was gardening in my home.
- I was making juice using fruit and/or vegetables.
- I was being recorded for a television show.
- ...

Basic Semantics

- Object: ResNet152-ImageNet1K
- Place: ResNet152-Place365
- Object detection: Faster R-CNN-MSCOCO (80)
- NTCIR classifier: VGG16-ImageNet1K (fine-tuned)

Feature Weighing

Conditional Random Field

$$E_{\theta}(s) = \lambda \sum_i \underbrace{\phi_u(s_i)}_{\text{unary}} + \sum_{ij} \underbrace{\phi_p(s_i, s_j)}_{\text{pairwise}}$$

Tuned to be adaptable to user and tasks.

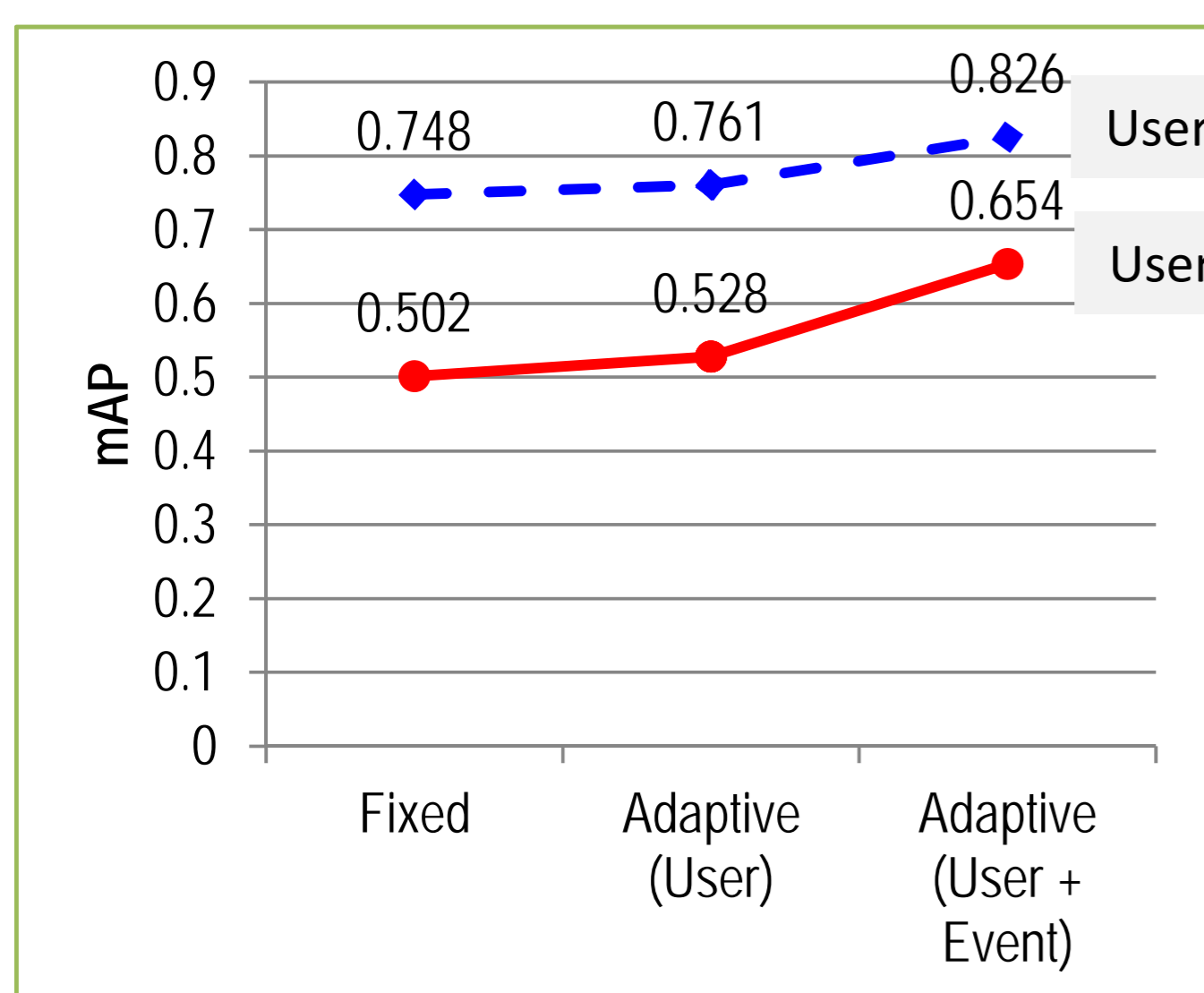
Temporal Smoothing

To ensure semantic coherence, a triangular window of size w is used, where w is adaptive to users.

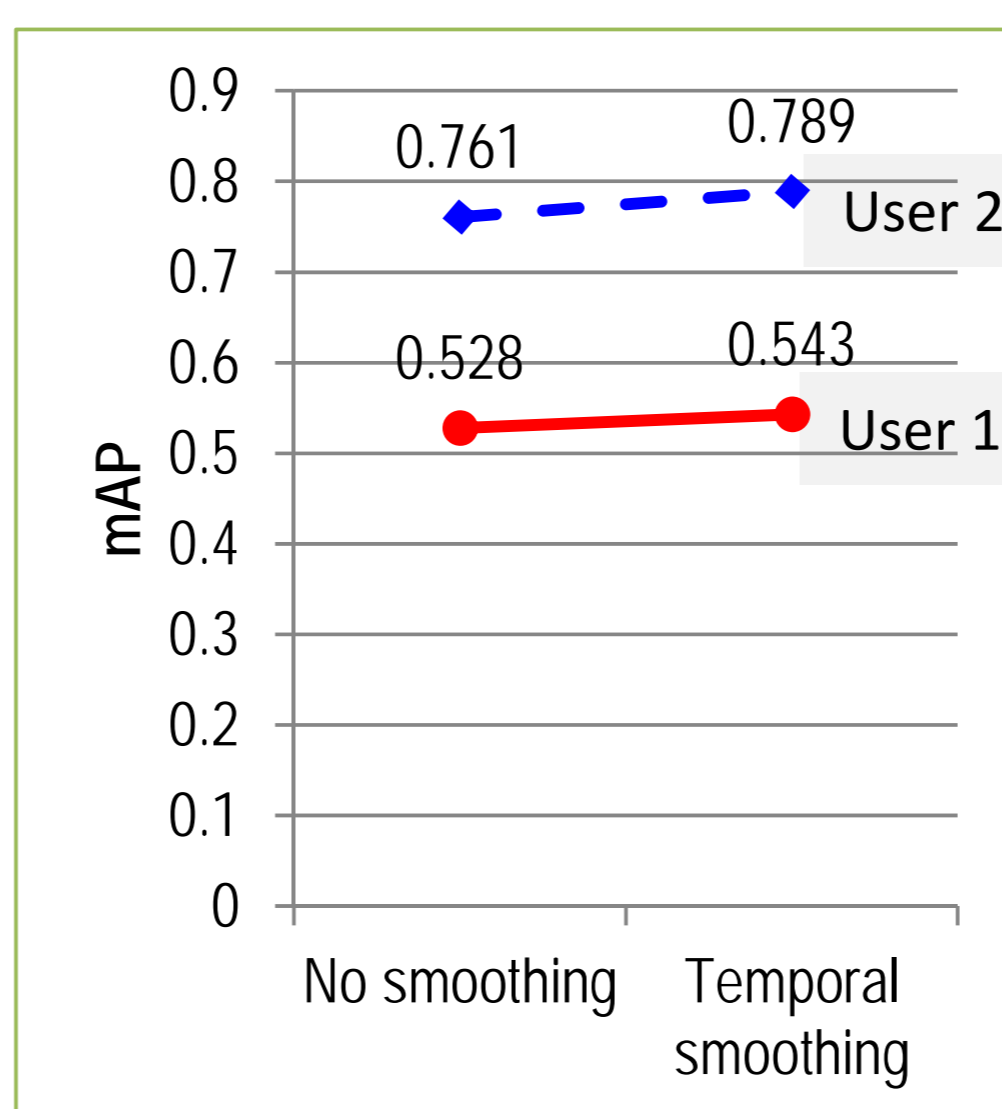
Post-Filtering

- Increase diversity of retrieved images.
- Use time and location (GPS) to filter images.
- Exclude images that are closer in time and location.

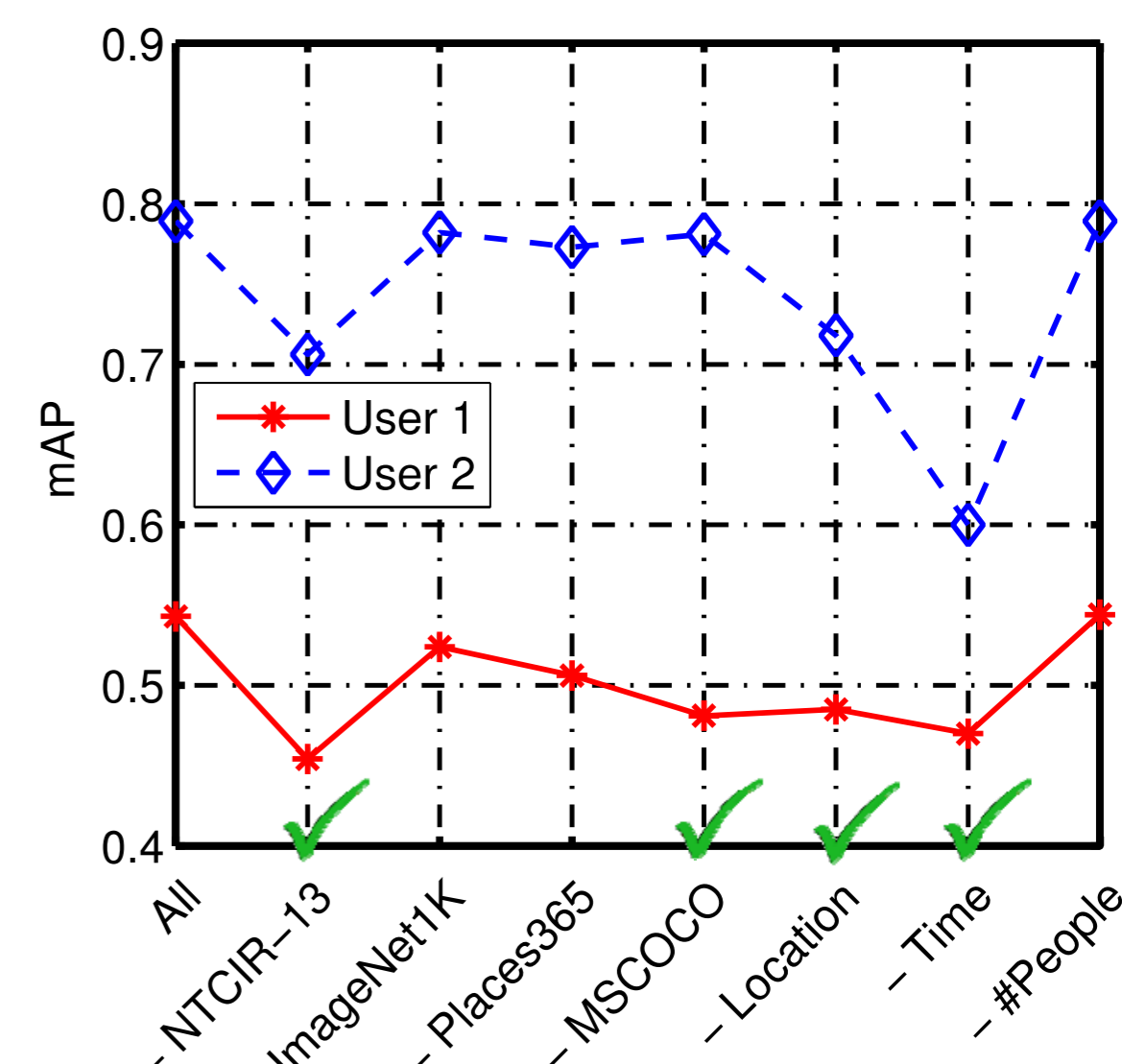
We achieved an average precision of 57.6% over 20 retrieval topics.



Effect of threshold for relevant concept searching



Effect of temporal smoothing



Feature importance