
National Institute of Meteorological Sciences


SkySearch: Satellite Video Search at Scale

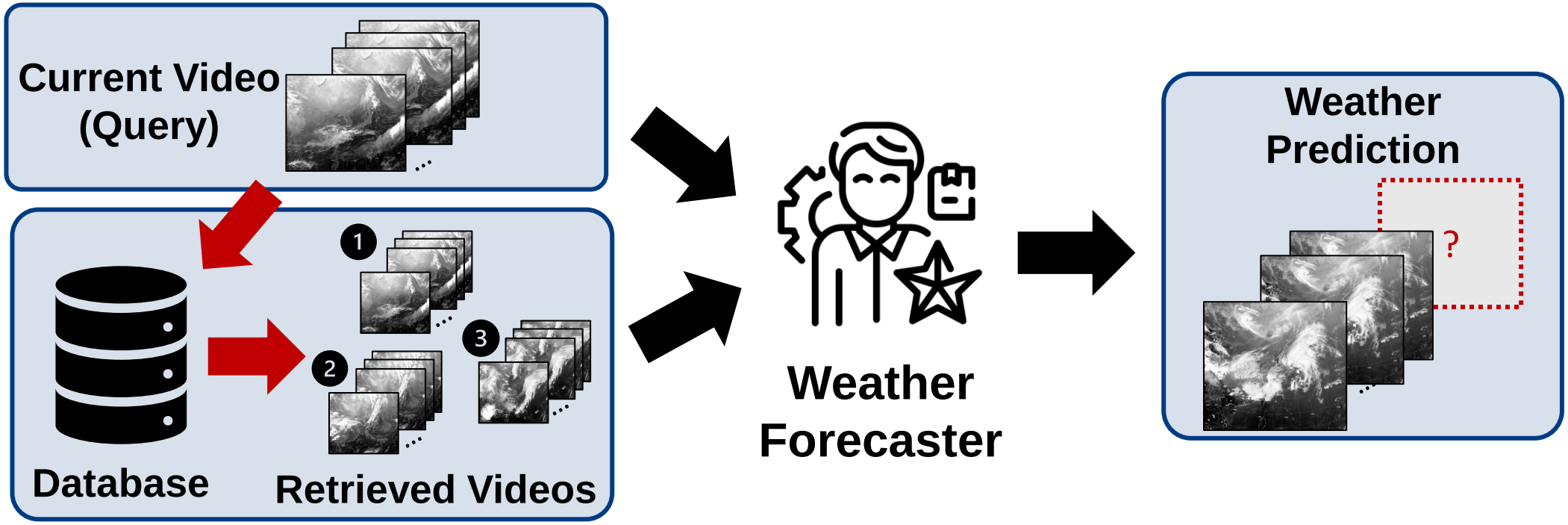
Minyoung Choe (KAIST), Geon Lee (KAIST), Changhun Han (Kookmin Univ.),
Suji Kim (Kookmin Univ.), Woong Hu (SI), Hyebeen Hwang (SI),
Geunseok Park (SI), Byeongyeon Kim (NIMS), Hyesook Lee (NIMS),
Ha-Myung Park (Kookmin Univ.), Kijung Shin (KAIST)

Motivations

- Weather patterns often **repeat**, making historical analysis essential.

Motivations

- Weather forecasters **refer to past satellite videos** to better understand and interpret current weather conditions.
- Our problem: Accurate and fast search of **similar satellite videos.**

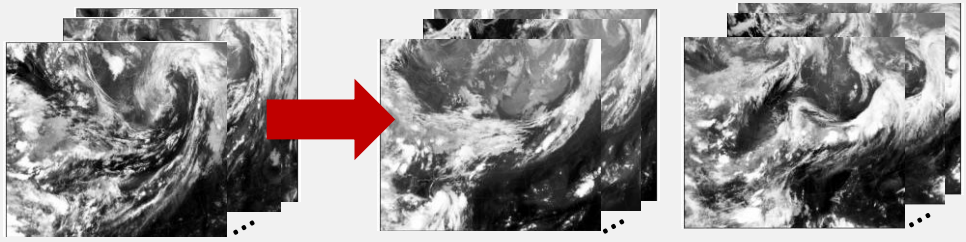


Challenges

- Satellite video retrieval presents **unique practical challenges**:
  1. **No ground truth for similarity**
  2. **Massive volume and resolution**

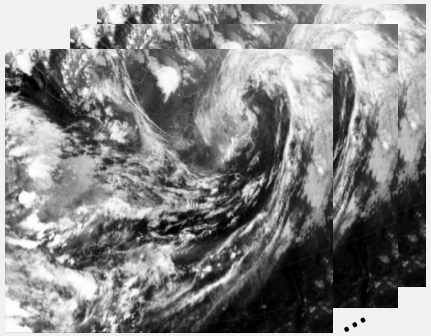
No Ground Truth for Similarity

"No label tells us which retrieved case is correct."

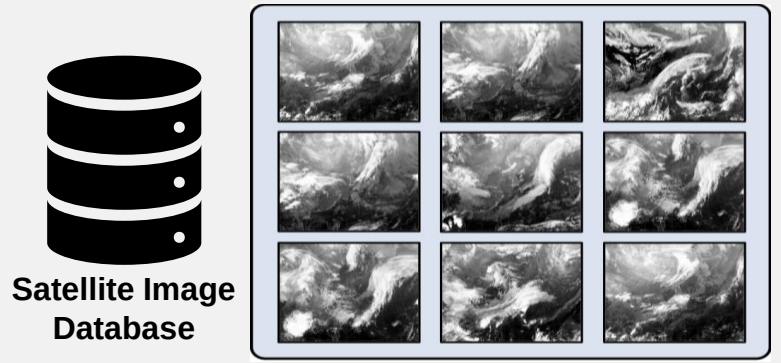


Query Video **Retrieved Cases**

Massive Volume and Resolution



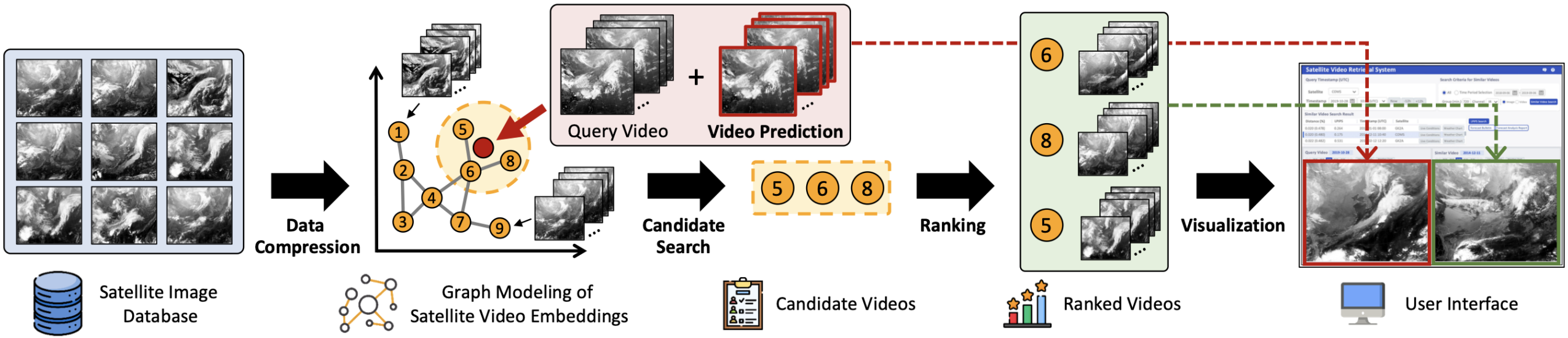
(600 ×748)×12 frames



Covers 12 years of data (2010–2021)

SkySearch: Satellite Video Search at Scale

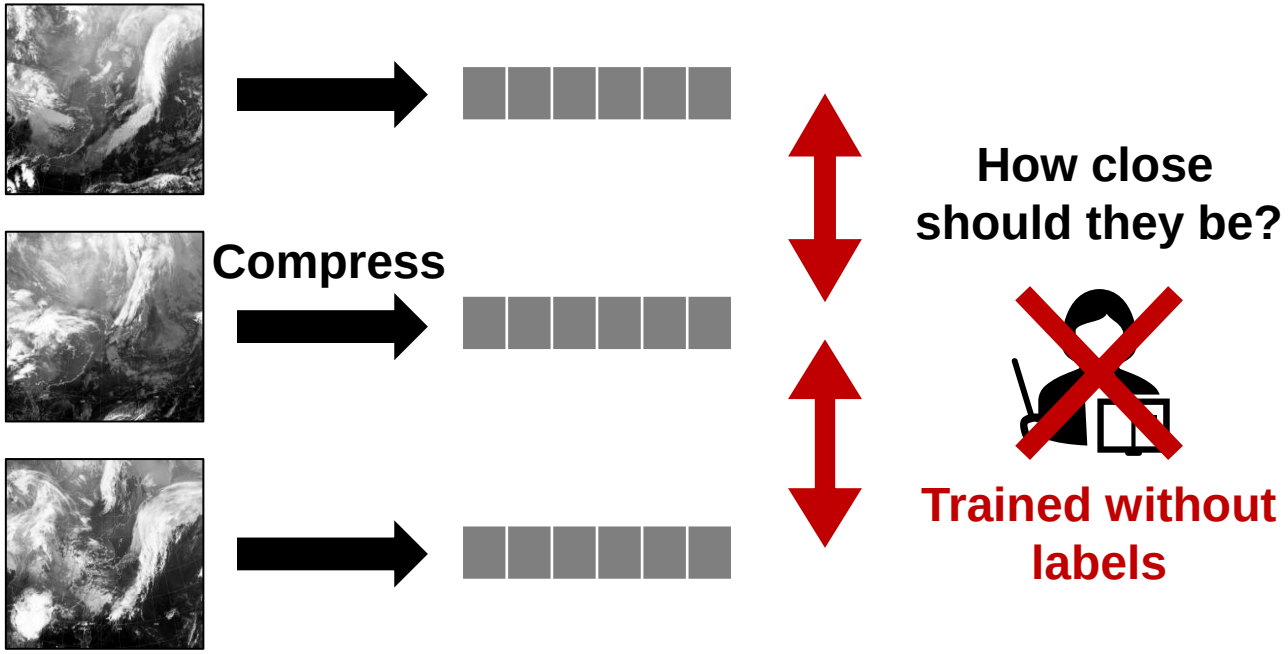
- An **unsupervised real-time satellite video search system** deployed at the Korea Meteorological Administration (KMA).



Key Components

- **SkySearch** has **three core** technical components,

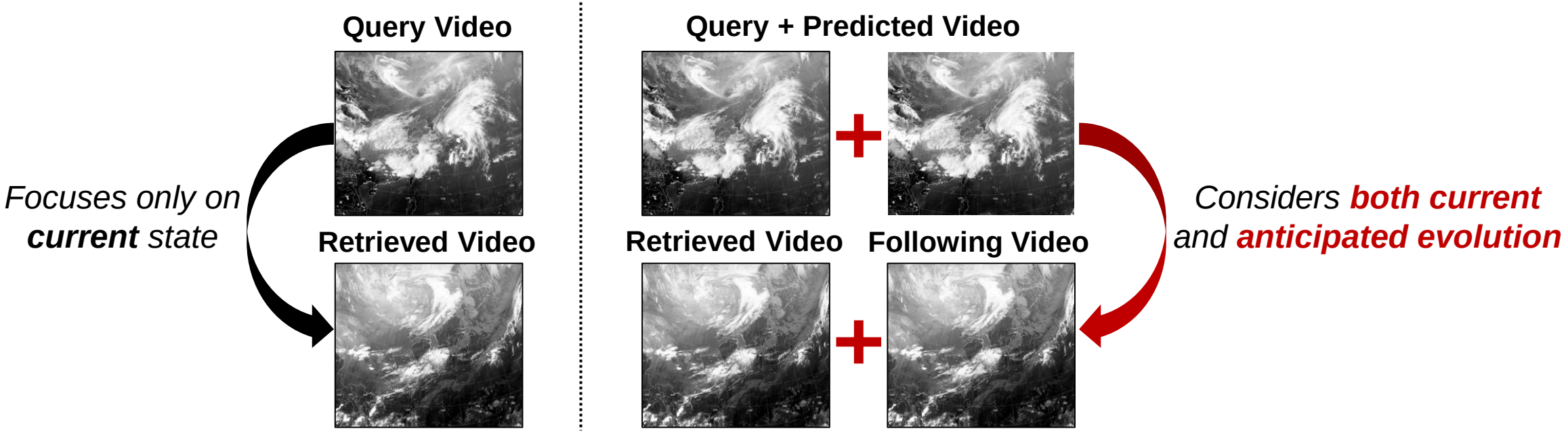
1. **Self-Supervised Video Compression:**
   Encode each video into **low-dimensional** vectors via **self-supervised learning**, capturing key spatiotemporal patterns **without labels**.



Key Components

- **SkySearch** has **three core** technical components,

2. **Prediction-Based Query Augmentation:**
   Predict future frames to **augment** the query, **enabling** retrieval aligned with **both current state and anticipated evolution**.

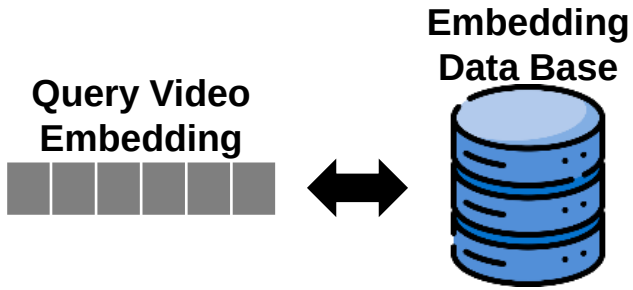


Key Components

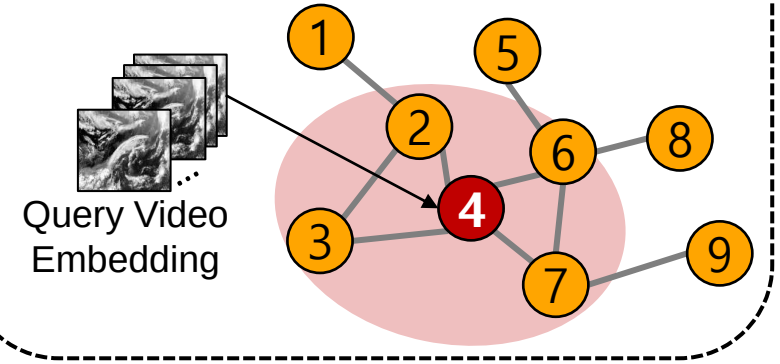
- **SkySearch** has **three core** technical components,

3. **Graph-Based Retrieval:**
   Traverse a **k-NN graph** in latent space to retrieve and rank similar past videos.

Naive Search:
Compare against entire database

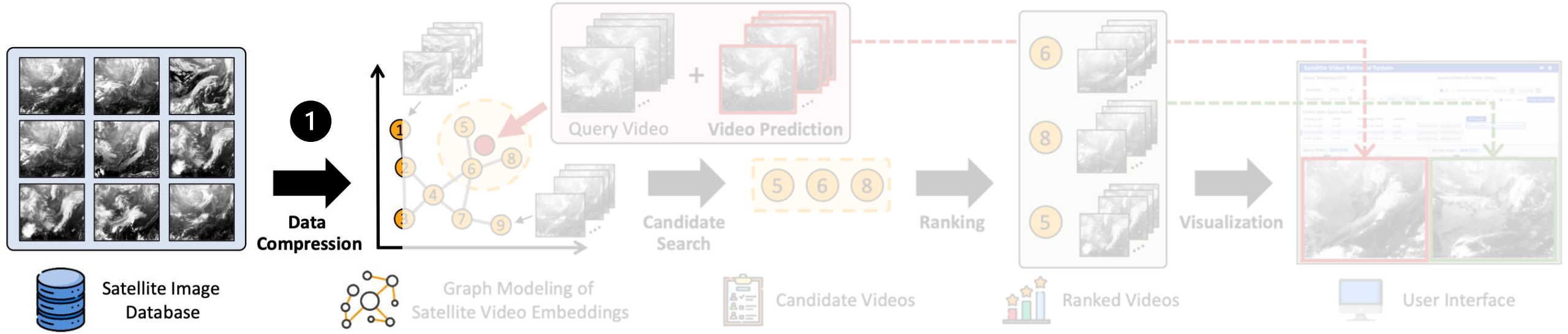


Graph-Based Retrieval:
Start from the query and explore its most similar neighbors



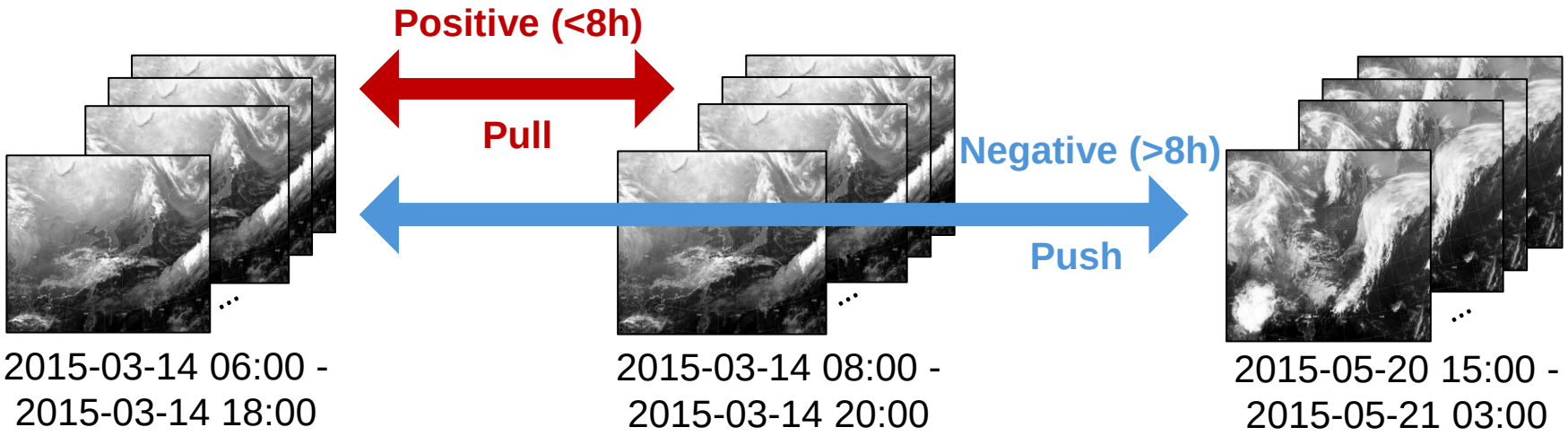
Step 1: Video Compression

- **Compression Objective:**
  - Compress satellite videos into **compact latent embeddings**.
  - Preserve **spatiotemporal similarity** between meteorologically similar videos.



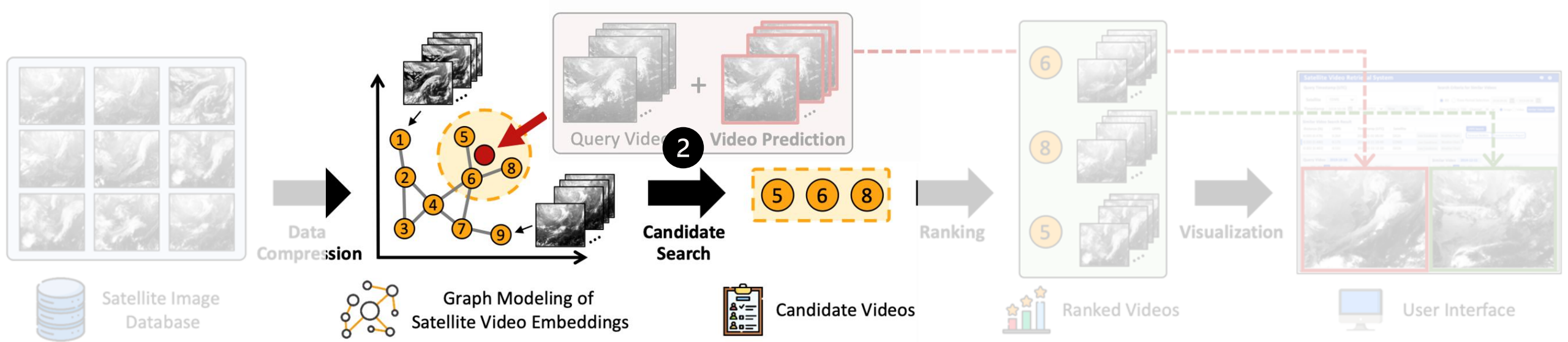
Step 1: Video Compression

- **Key Point:** Weather evolves **gradually** over time.
  → temporal proximity ≈ semantic similarity.
- **Contrastive Loss:** pull close positives, push away negatives.
  - **Positive Pairs:** videos within ±8 hours.
  - **Negative Pairs:** temporally distant videos.



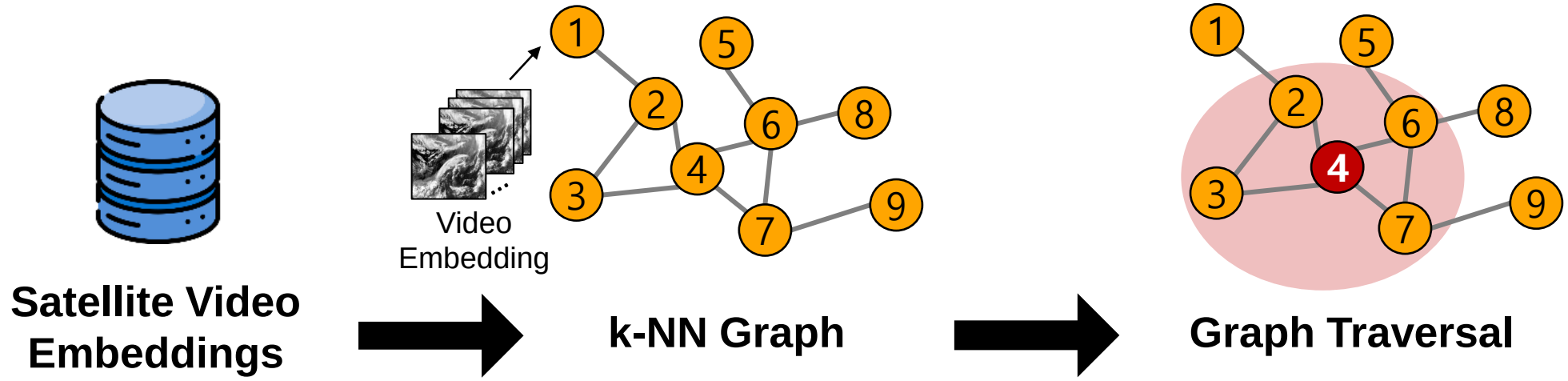
Step 2: Graph-Based Search

- **Goal: Efficiently** find videos most similar to a given query video.
- **Challenge**: Direct comparison with all videos is **too slow** for real-time use.
- **Solution**: Use **graph-based retrieval** to scale to large databases.



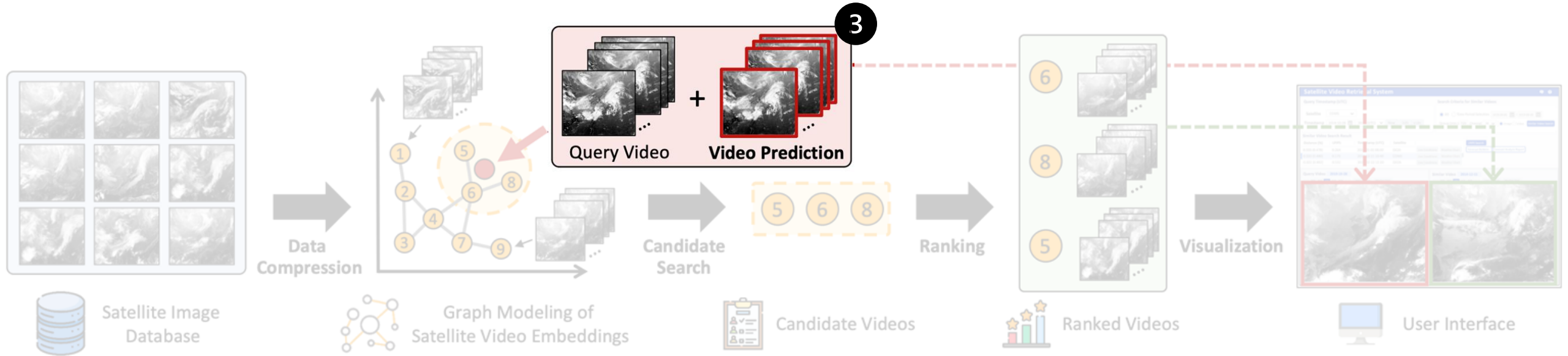
Step 2: Graph-Based Search

- **Candidate Search: Efficient and Time-Aware**
  - Build a **k-NN graph** from all video embeddings.
  - Start search from the query and **traverse nearby neighbors**.
  - **Restrict search** to videos from the **user-defined time window**.
  - Return a small set of highly similar candidates.



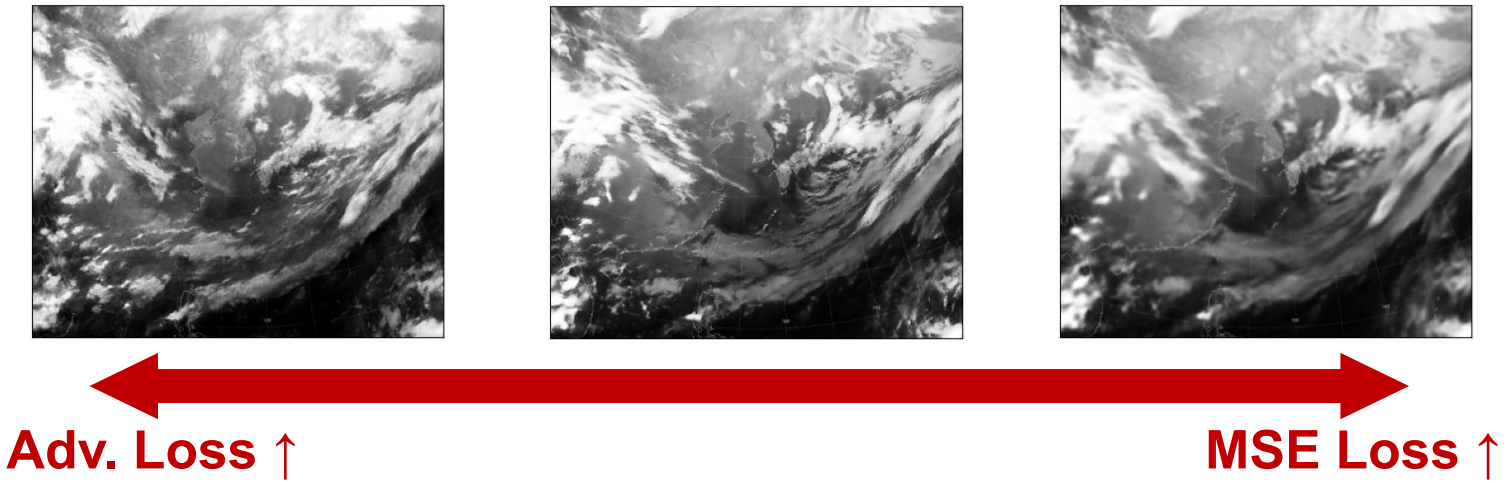
Step 3: Video Prediction

- **Goal:** Augment query videos with predicted future frames.
- **Purpose:** Retrieve past cases matching **both current and future conditions** to support forecast reasoning.



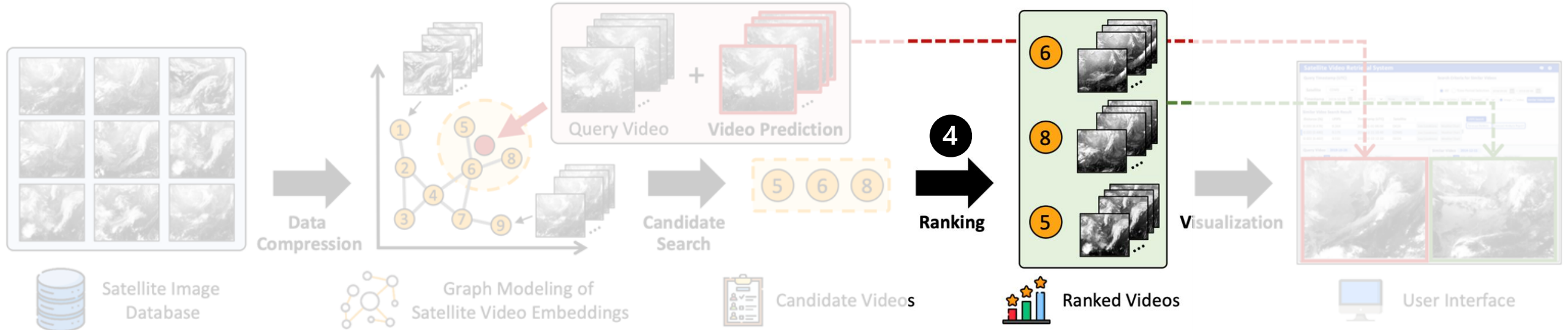
Step 3: Video Prediction

- **Backbone: SimVP**, a strong baseline for video prediction.
- **Challenge: Blurry** results on high-resolution satellite videos.
- **Solution:** Balance MSE and adversarial losses.
  - **MSE Loss:** Ensures pixel-level accuracy.
  - **Adversarial Loss:** Improves visual realism and sharpness.



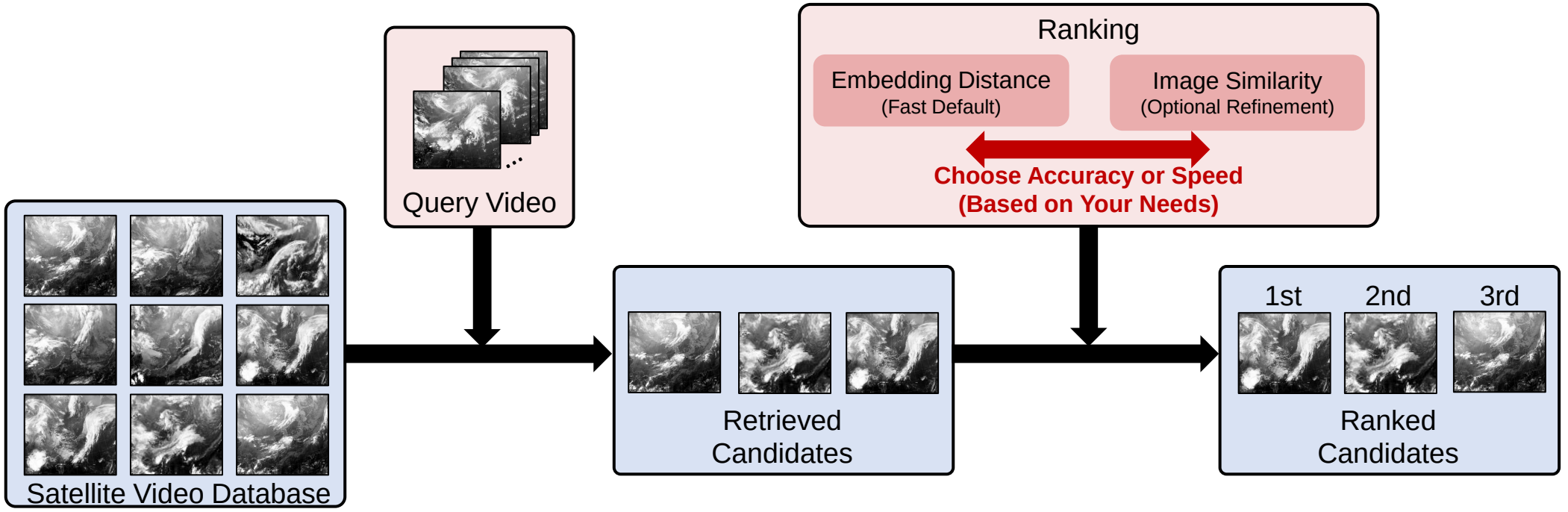
Step 4: Ranking

- **Goal**: Retrieved candidates are ranked by embedding distance or perceptual similarity to **prioritize** relevant results.



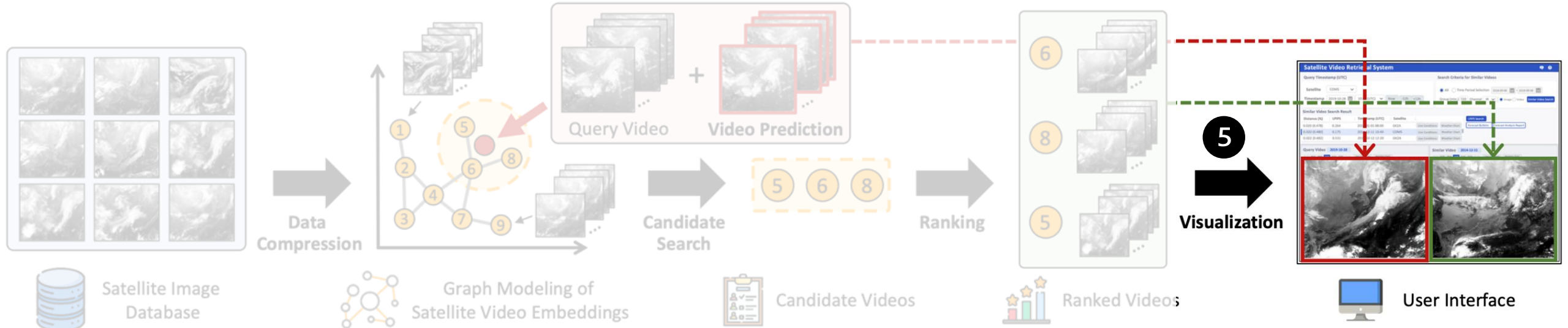
Step 4: Ranking

- **Default Ranking**: Use Euclidean distance between embeddings for speed.
- **Refined Ranking:** Optionally apply LPIPS, FSIM, or SSIM for better accuracy.
- **Trade-off:** Perceptual metrics improve relevance but increase computation time.



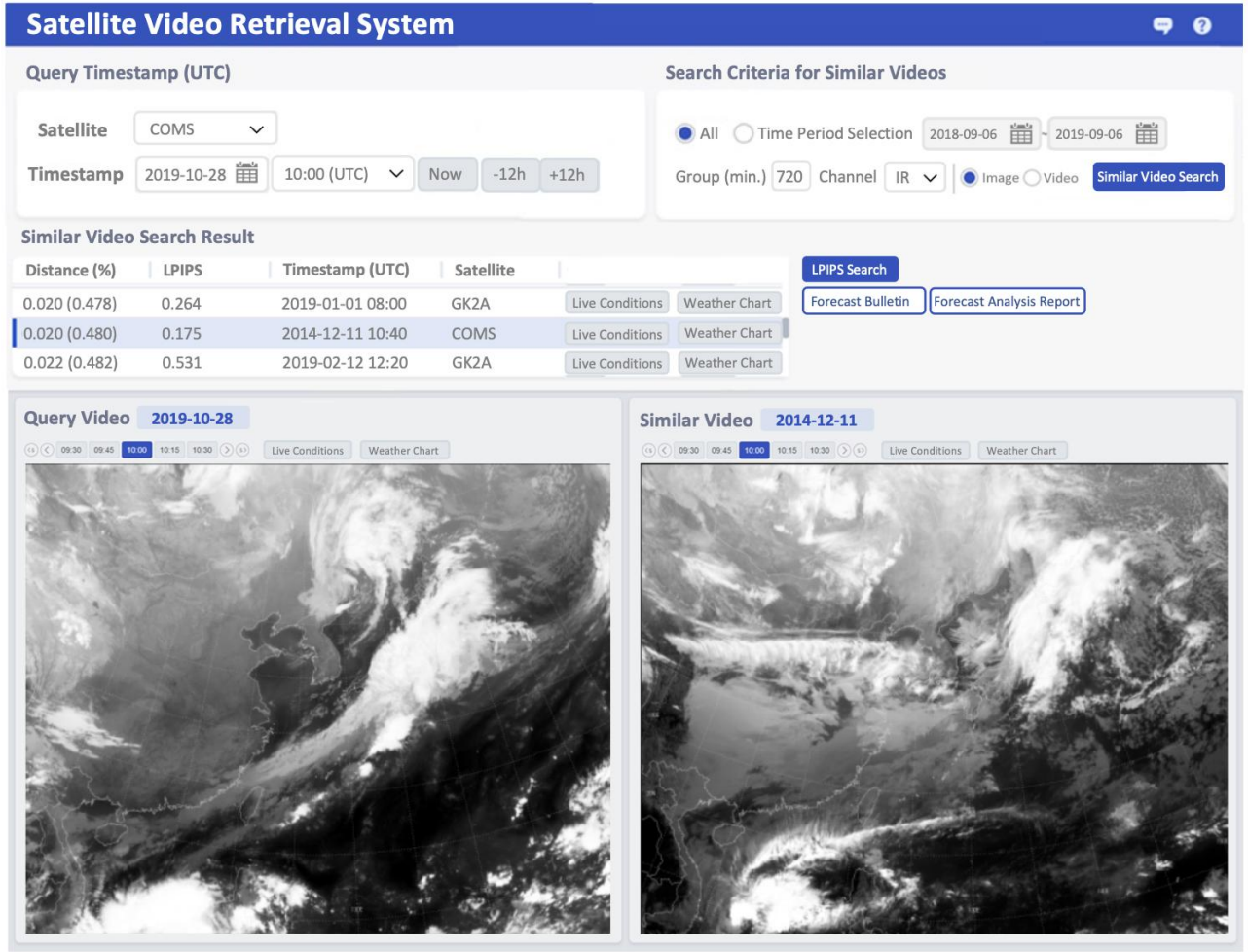
Step 5: Visualization

- **Goal**: The retrieved results are displayed through a **user interface** deployed at KMA.



Step 5: Visualization

- **Input:** satellite, channel, timestamp and the desired time range.
- **Output:** list of similar videos with date & similarity score.



Deployment at KMA

National Institute of Meteorological Sciences

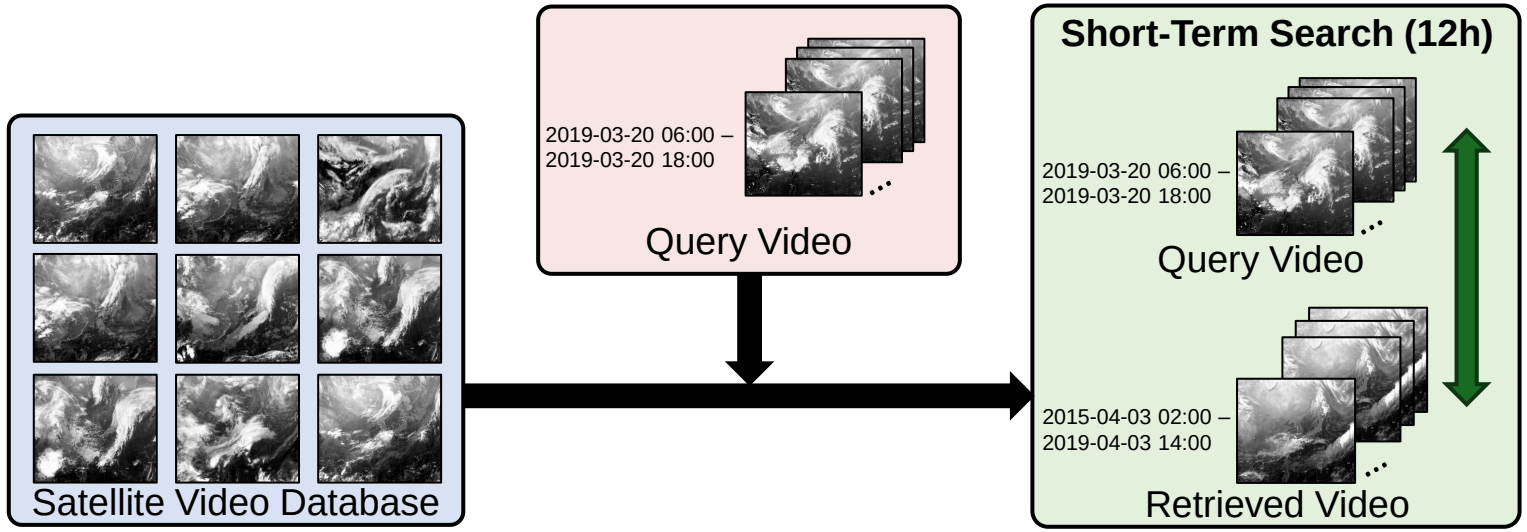
- **Satellite Sources**:
  - COMS (2010–2020): general-purpose satellite.
  - GK2A (2018–present): specialized for weather monitoring.
- **Supported Channels**:
  - IR (Infrared), SWIR (Shortwave IR), WV (Water Vapor).
- **Temporal Coverage**:
  - Dataset spans 12 years: Sep 2010 – Jun 2021.
  - Image intervals: 15 min (COMS), 10 min (GK2A).
- **Resolution & Scale**:
  - Original: 1300×1500, downsampled to 600×748.
  - Each video = 12 frames, totaling 5.4M pixels per video.

Evaluation of Final Outcomes

- **Evaluation Setup**
  - **Compare SkySearch with pre-trained and fine-tuned baselines**.
    - ResNet, EfficientNet, VideoMAE, TimeSformer
  - **Query Set:** 2,000 IR-channel COMS videos (2019–2020), 12h duration.
  - **Database:** 2014–2018 IR COMS videos.
  - **Metrics:** LPIPS, FSIM, SSIM over aligned frame pairs.

Evaluation of Final Outcomes

- **Short-Term Search Evaluation**
  - **Goal:** Retrieve visually similar past videos.
  - Compare retrieved 12h video with query 12h video.
  - Evaluates **current pattern matching** ability.



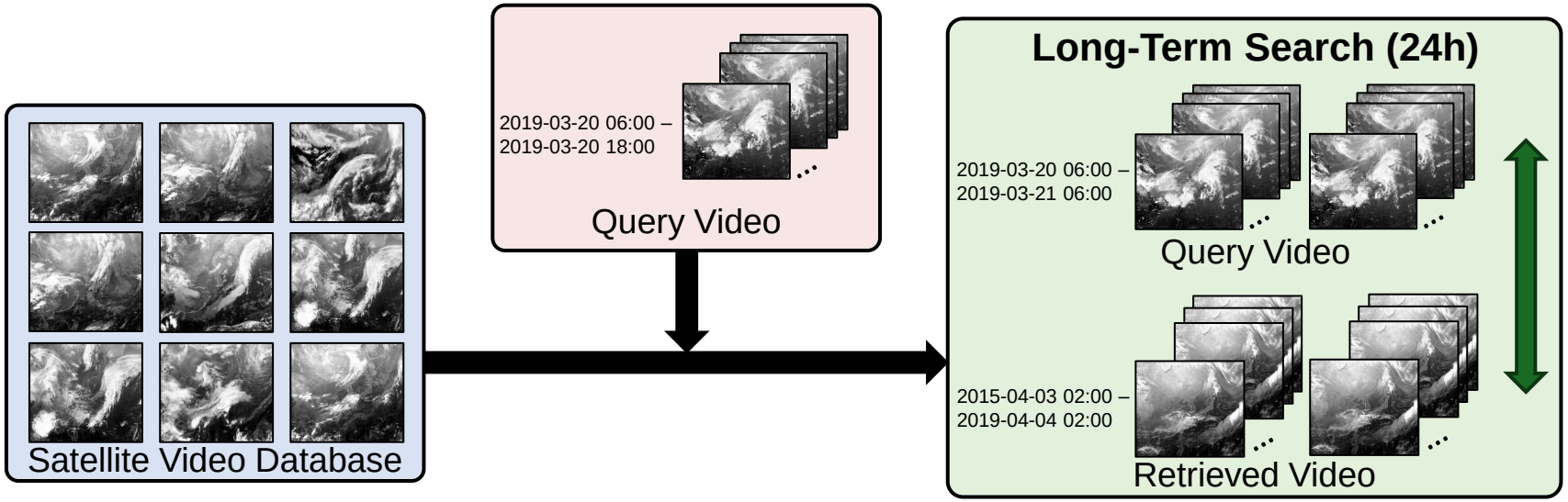
Evaluation of Final Outcomes

- **Short-Term Search Evaluation**
  - **SkySearch** achieves the **best** performance across all metrics.
  - The version without video prediction achieves the best results, demonstrating **strong representation quality**.

Model	Short-Term Search Accuracy (12 Hours)		
	LPIPS (↓)	FSIM (↑)	SSIM (↑)
Random	0.2406 ± 0.0364	0.3350 ± 0.0126	0.2249 ± 0.0609
ResNet (Pre-Trained)	0.2086 ± 0.0396	0.3465 ± 0.0159	0.2635 ± 0.0780
ResNet (Fine-Tuned w/ SimCLR)	0.2034 ± 0.0342	0.3487 ± 0.0149	0.2713 ± 0.0753
EfficientNet (Pre-Trained)	0.2087 ± 0.0376	0.3461 ± 0.0150	0.2633 ± 0.0751
EfficientNet (Fine-Tuned w/ SimCLR)	0.2039 ± 0.0365	0.3479 ± 0.0151	0.2703 ± 0.0767
VideoMAE (Pre-Trained)	0.2039 ± 0.0401	0.3486 ± 0.0164	0.2712 ± 0.0812
TimeSformer (Pre-Trained)	0.2090 ± 0.0395	0.3466 ± 0.0163	0.2627 ± 0.0782
SkySearch (w/o Video Prediction)	**0.1943 ± 0.0355**	**0.3502 ± 0.0150**	**0.2784 ± 0.0767**
SkySearch (w/ Video Prediction)	0.1960 ± 0.0356	0.3495 ± 0.0151	0.2775 ± 0.0759

Evaluation of Final Outcomes

- **Long-Term Search Evaluation**
  - **Goal:** Retrieve past videos that resemble **both current and future trends**.
  - **Tests whether retrieved video evolves similarly**.



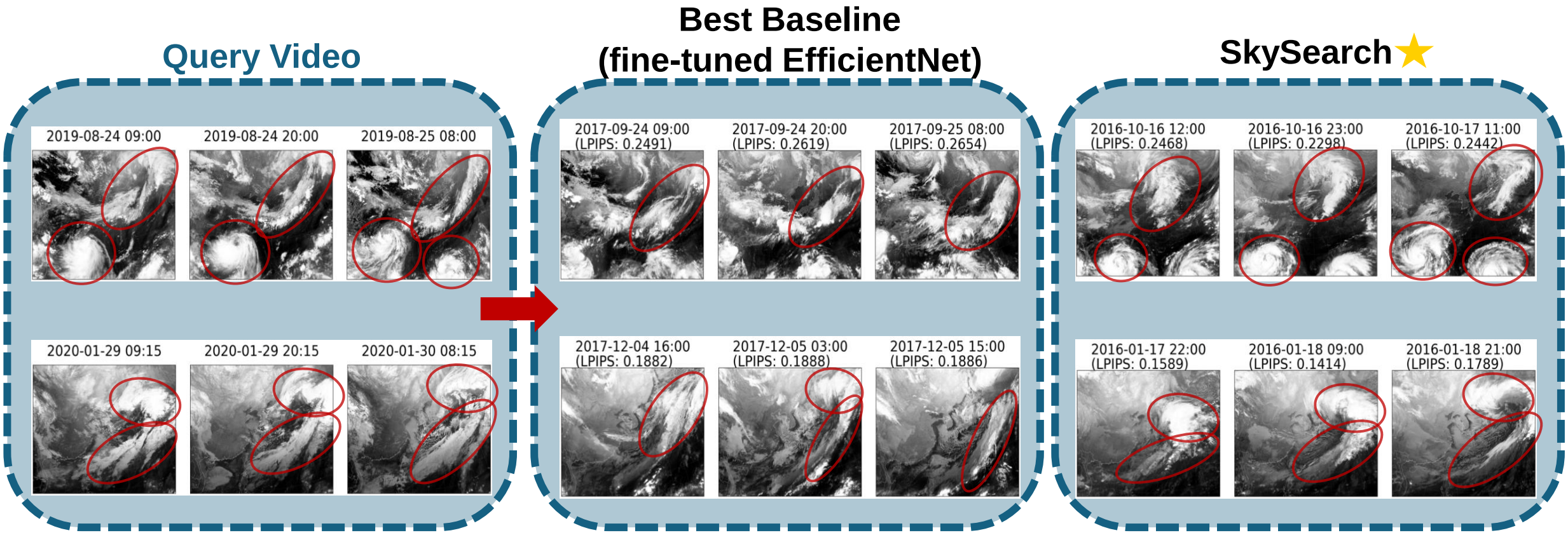
Evaluation of Final Outcomes

- **Long-Term Search Evaluation**
  - **SkySearch outperforms** all baselines by a clear margin.
  - The video prediction module enables alignment with future trends, boosting accuracy in long-term retrieval.

Model	Long-Term Search Accuracy (24 Hours)		
	LPIPS (↓)	FSIM (↑)	SSIM (↑)
Random	0.2405 ± 0.0356	0.3348 ± 0.0122	0.2248 ± 0.0595
ResNet (Pre-Trained)	0.2100 ± 0.0388	0.3459 ± 0.0154	0.2626 ± 0.0765
ResNet (Fine-Tuned w/ SimCLR)	0.2051 ± 0.0335	0.3478 ± 0.0144	0.2687 ± 0.0737
EfficientNet (Pre-Trained)	0.2103 ± 0.0367	0.3452 ± 0.0145	0.2619 ± 0.0739
EfficientNet (Fine-Tuned w/ SimCLR)	0.2047 ± 0.0357	0.3476 ± 0.0146	0.2690 ± 0.0751
VideoMAE (Pre-Trained)	0.2051 ± 0.0390	0.3480 ± 0.0159	0.2692 ± 0.0791
TimeSformer (Pre-Trained)	0.2102 ± 0.0388	0.3460 ± 0.0160	0.2615 ± 0.0763
SkySearch (w/o Video Prediction)	0.1974 ± 0.0349	0.3490 ± 0.0147	0.2751 ± 0.0750
SkySearch (w/ Video Prediction)	**0.1965 ± 0.0347**	**0.3494 ± 0.0146**	**0.2766 ± 0.0750**

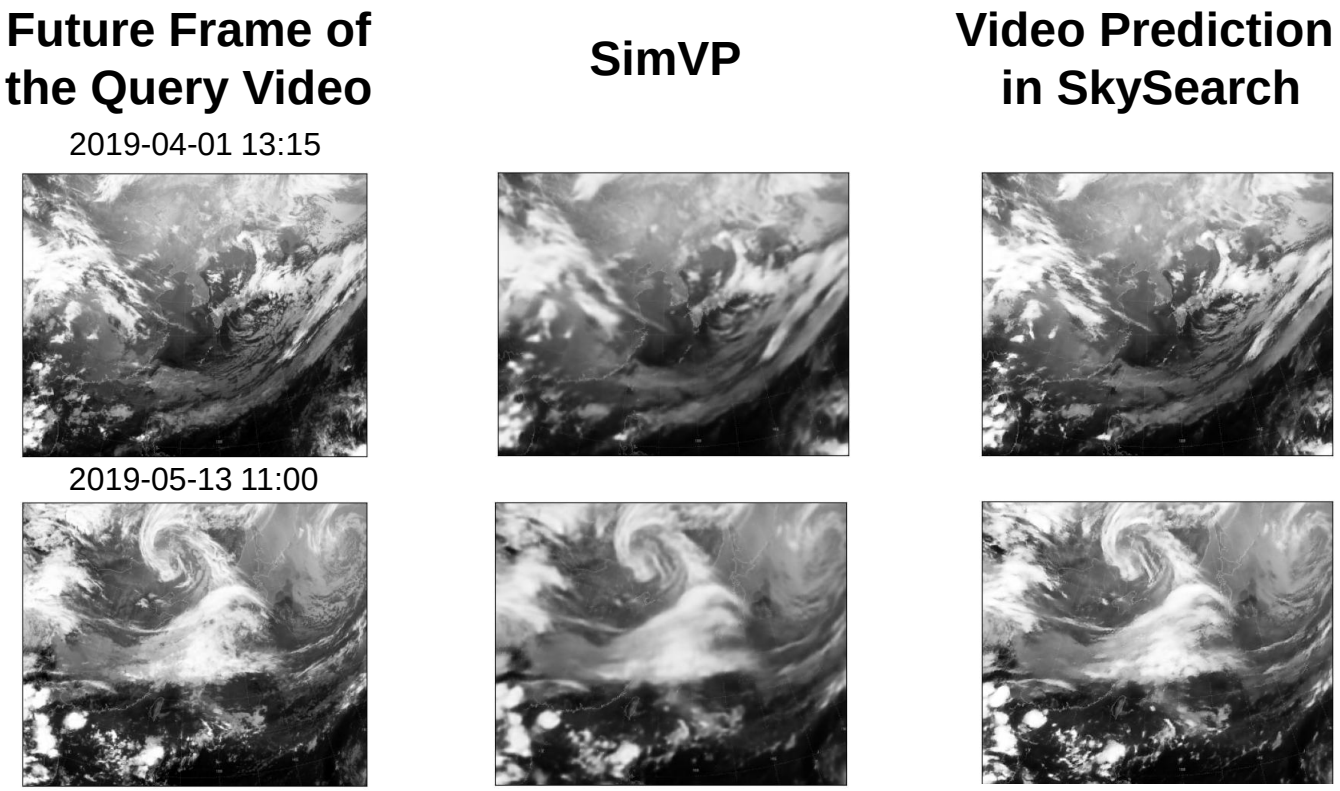
Evaluation of Final Outcomes

- SkySearch retrieves the **most visually coherent** past events.



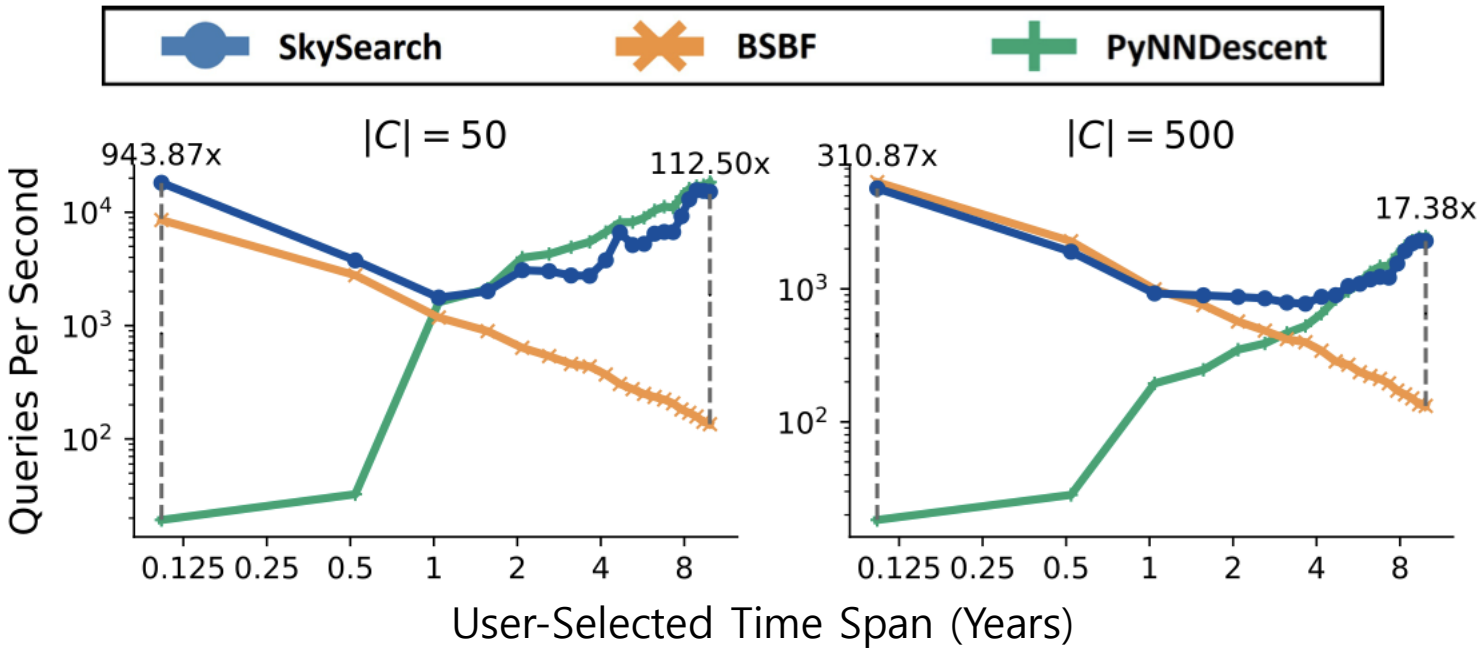
Evaluation of Components

- **Video prediction with adversarial loss** improves sharpness and realism.
- Achieves significantly **lower LPIPS** than SimVP: 0.2170 vs. 0.3907.



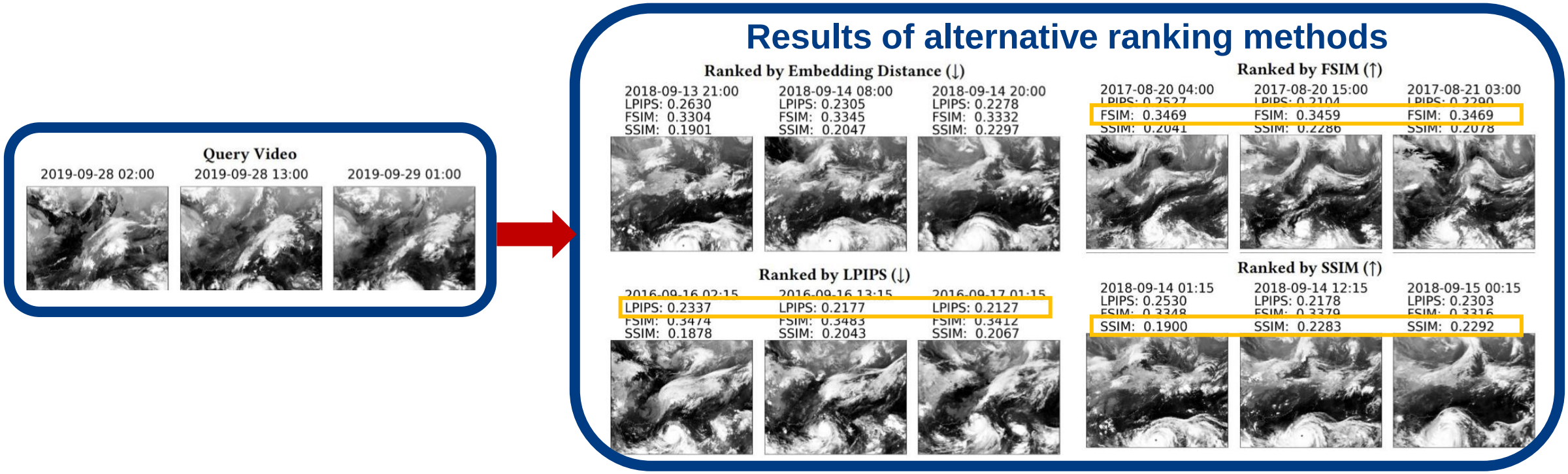
Evaluation of Components

- **Graph-based candidate search is fast and accurate**, even when users search across different time ranges.
- SkySearch achieves up to **943× faster** search than PyNNDescent, especially when the search is limited to recent years or specific seasons.



Evaluation of Components

- **SkySearch offers flexible ranking**: fast **embedding-based** by default, or **refined with LPIPS, FSIM, SSIM** for better visual alignment.
- **Users can trade off speed and accuracy** depending on their needs.



Conclusion

SkySearch, a large-scale satellite video retrieval system deployed at the Korea Meteorological Administration (KMA), offers the following advantages:

- **Efficient**: SkySearch processes queries from massive databases and provides results **within seconds**.
- **Accurate**: SkySearch retrieves videos that closely align with the input query video, **both numerically and visually**.
- **Unsupervised**: SkySearch employs **self-supervised learning** to address the lack of labeled data.

For reproducibility, we release our code and online appendix at https://github.com/geon0325/skysearch.