# OPTIMISING SEMANTIC CLOUD SEGMENTATION VIA SVOS FOR SOLAR FORECASTING

**Eduardo Sáez Martínez, Raphael Gut, Bijan Nouri, Yann Fabel, David Magiera, Niklas Blum and Luis F. Zarzalejo**

**EuroSun 2026**

**17/09/2026, Freiburg**

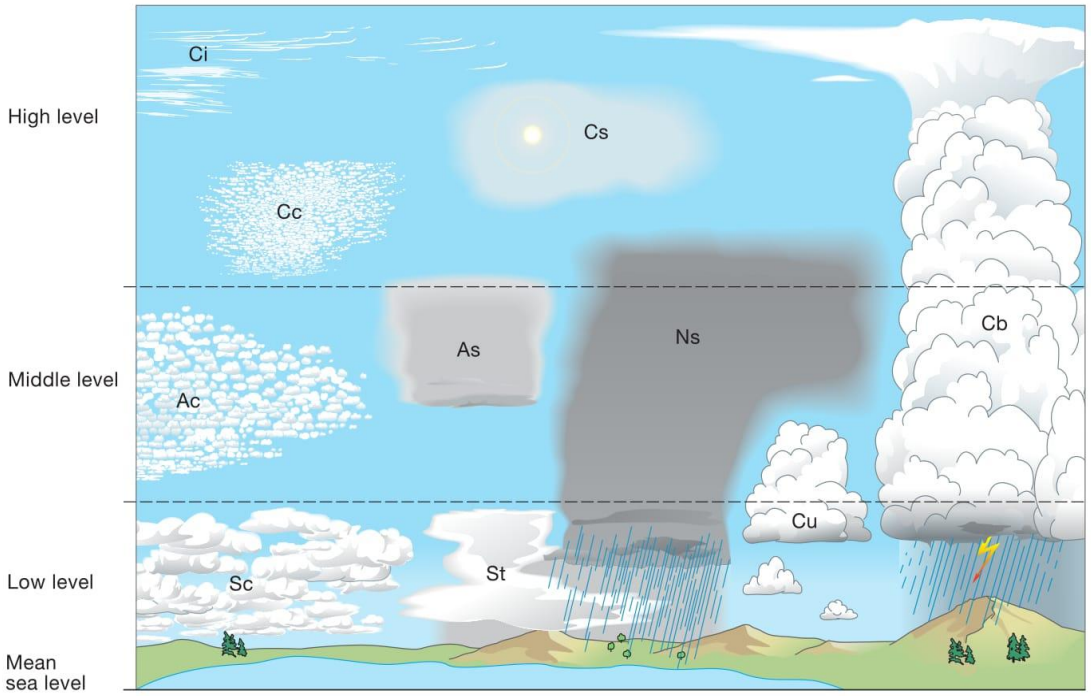
Ciemat

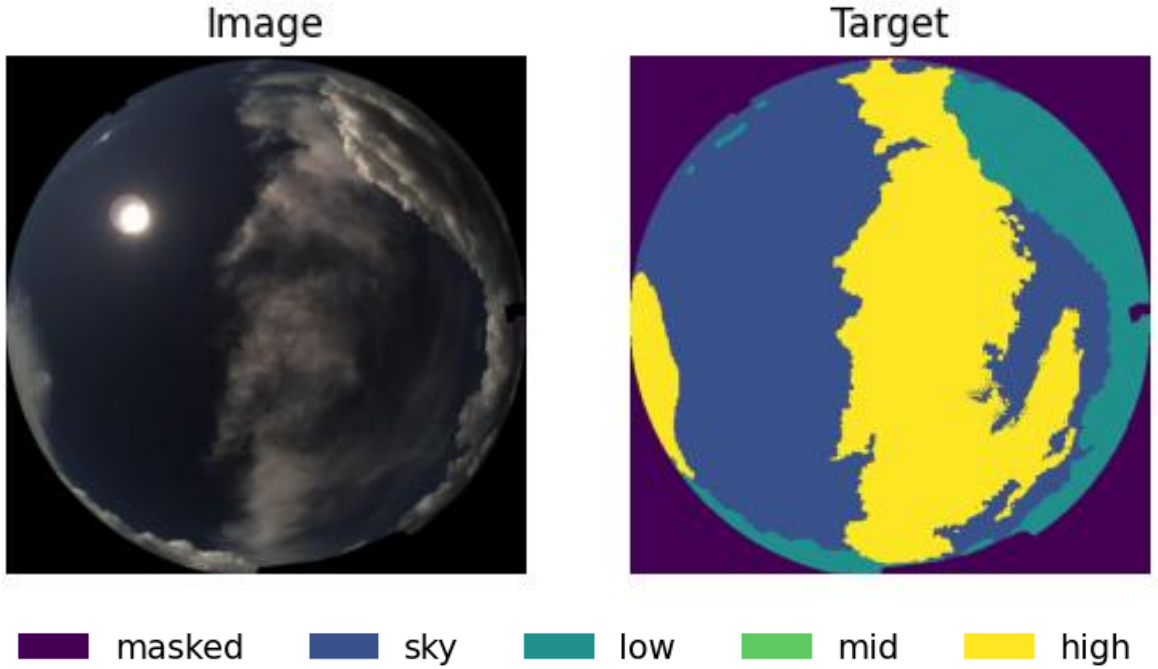
DLR

# Problem definition

**Motivation:** Cloud detection is essential for physics-based solar irradiance monitoring and forecasting, enabling reliable operation of solar power plants and electrical grids

**Goal:** Pixel-wise classification of clouds in all-sky images into three cloud (base) height layers defined by the WMO.



# Problem definition

**Approach:**

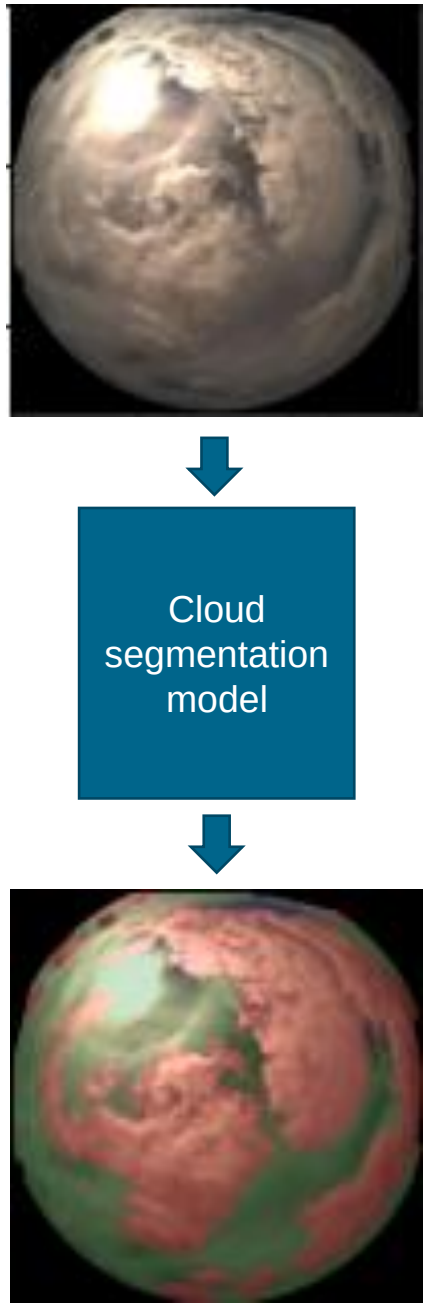
- Deep learning has emerged as the state-of-the-art approach

**Challenges:**

- Large amounts of high-quality pixel-level annotations are required
- Manually annotated ground-truth data are scarce
- Expert annotation is time-consuming

**Solution:**

- Automated expansion of manually annotated datasets.



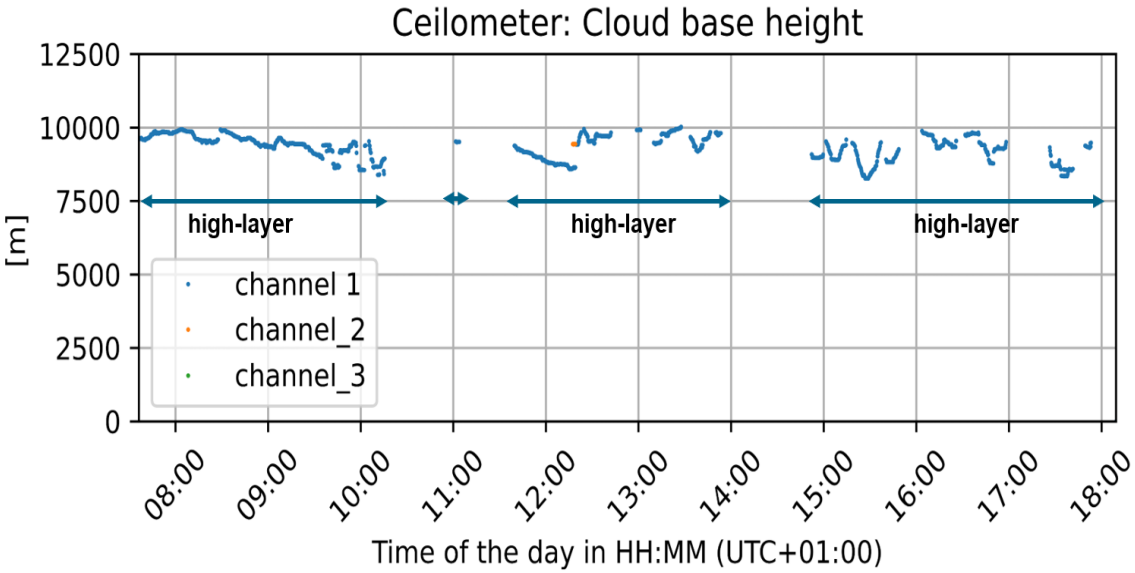
# Datasets

**Manually annotated dataset** [1]**:**

- 770 annotated images

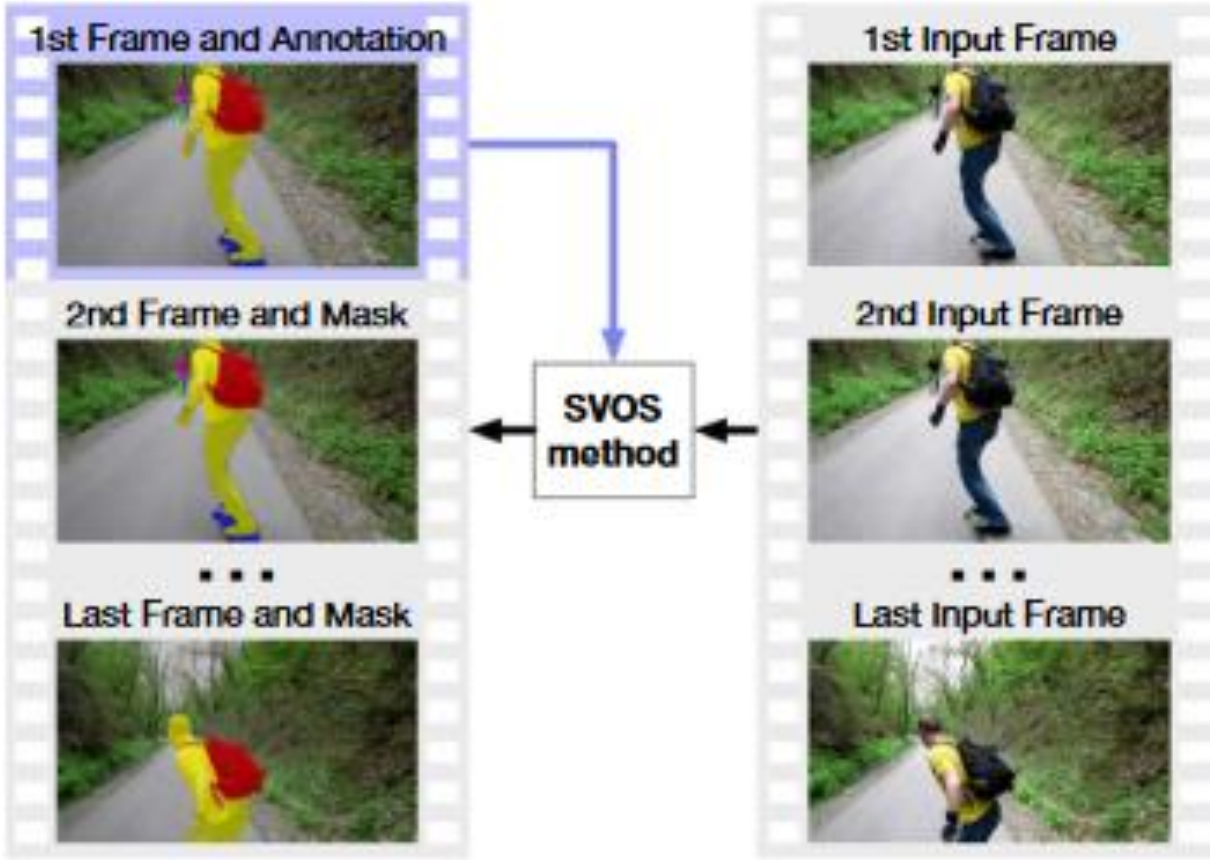
**Weakly annotated dataset** [2]**:**

- Ceilometer measures cloud height
- Measurements within certain thresholds represent single cloud layers
- Temporal stable single cloud layer conditions are identified
- Assign corresponding cloud layer as image-level weak label to binary cloud segmentations
- 47.000 segmented images



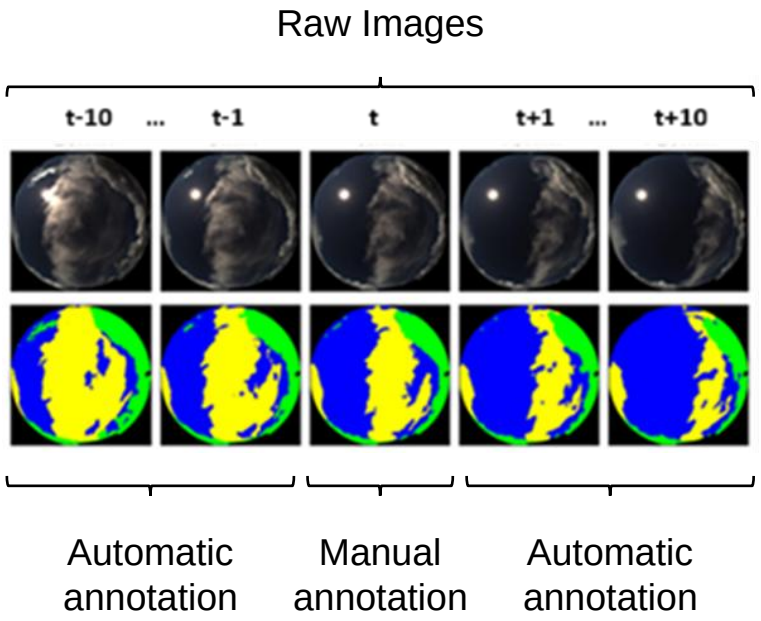
# Ground-truth data generation via Semi-supervised Video Object Segmentation: Overview

- VOS (Video Object Segmentation) is an AI technique that segments and tracks objects frame by frame in a video.
- In Semi-Supervised VOS (SVOS) an initial mask of the object to be segmented in the first frame is given.

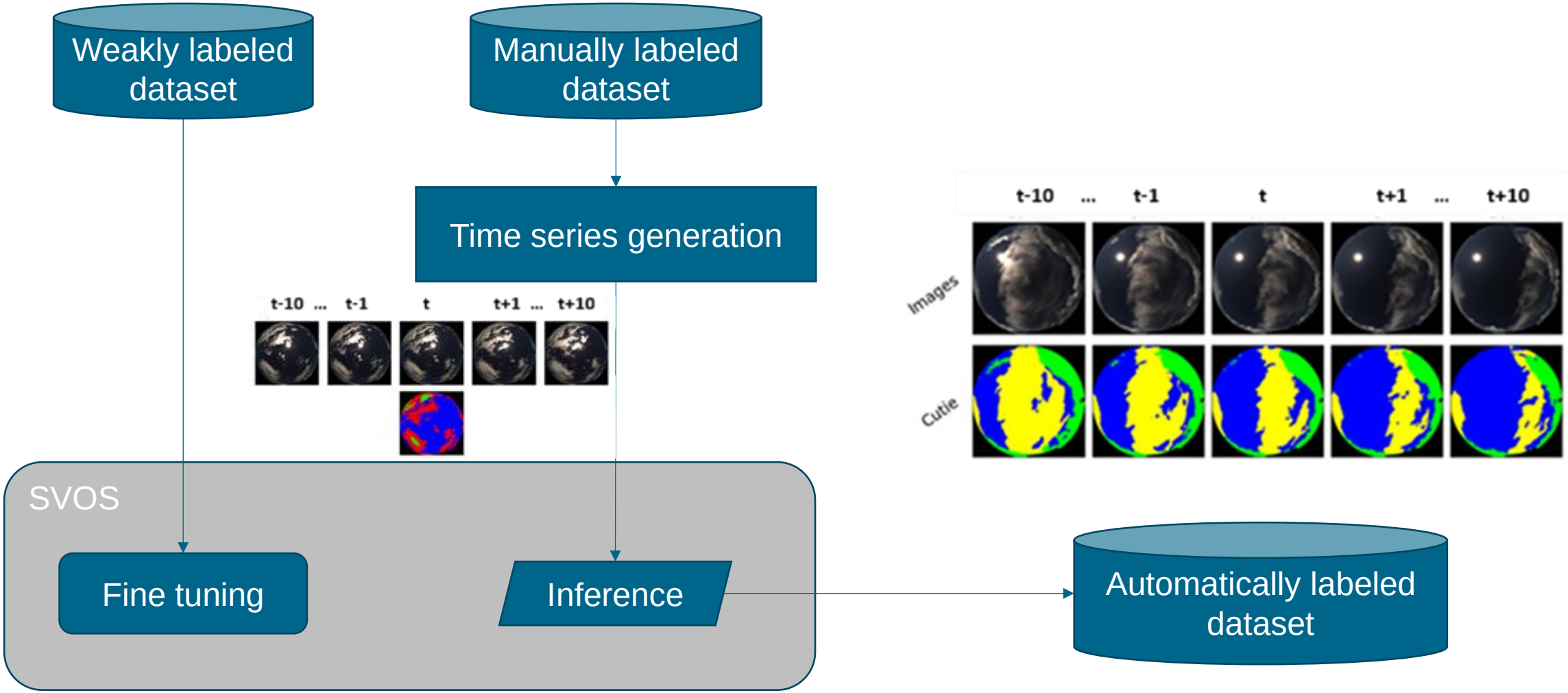


# Ground-truth data generation via Semi-supervised Video Object Segmentation: Strategy

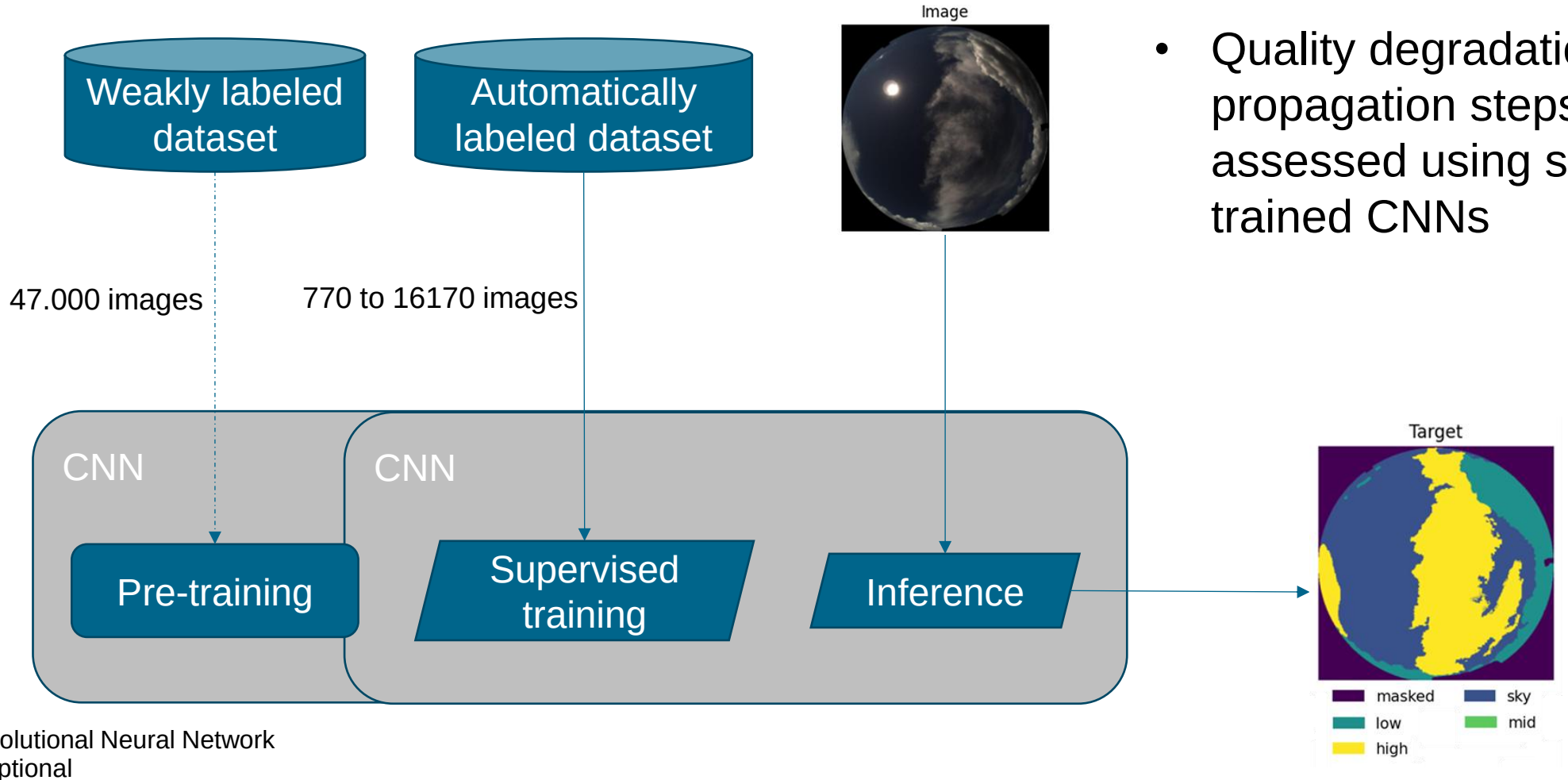
- Objectives of using SVOS:
  - Expand the existing manually annotated dataset
  - Study the degradation of quality in label propagation
- 4 SVOS models were analyzed
- Time series with a maximum of +/- 10 frames are evaluated



# Ground-truth data generation via Semi-supervised Video Object Segmentation: Strategy



# Ground-truth data generation via Semi-supervised Video Object Segmentation: Strategy



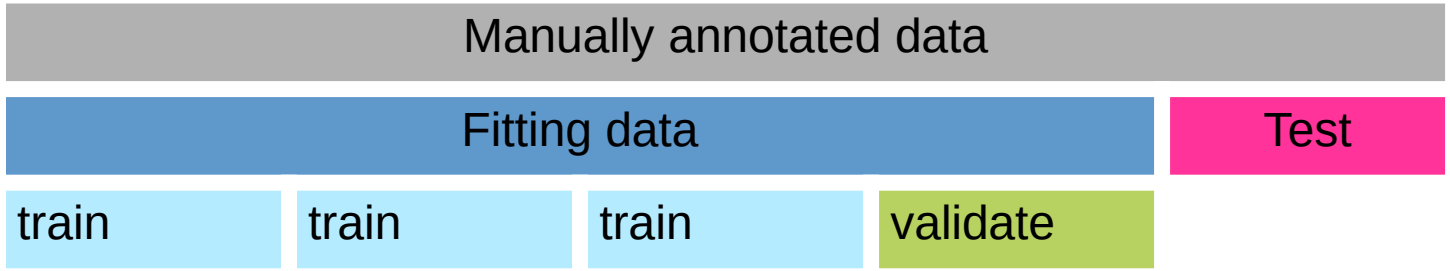

- Quality degradation across propagation steps assessed using separately trained CNNs

CNN: Convolutional Neural Network

Optional

# Experiment Results: Dataset split and Metric

- **Dataset split**:




- Train dataset increased with each propagation step.
- Test and validation dataset are manually annotated images only.

- **Metric**: Intersection over Union.

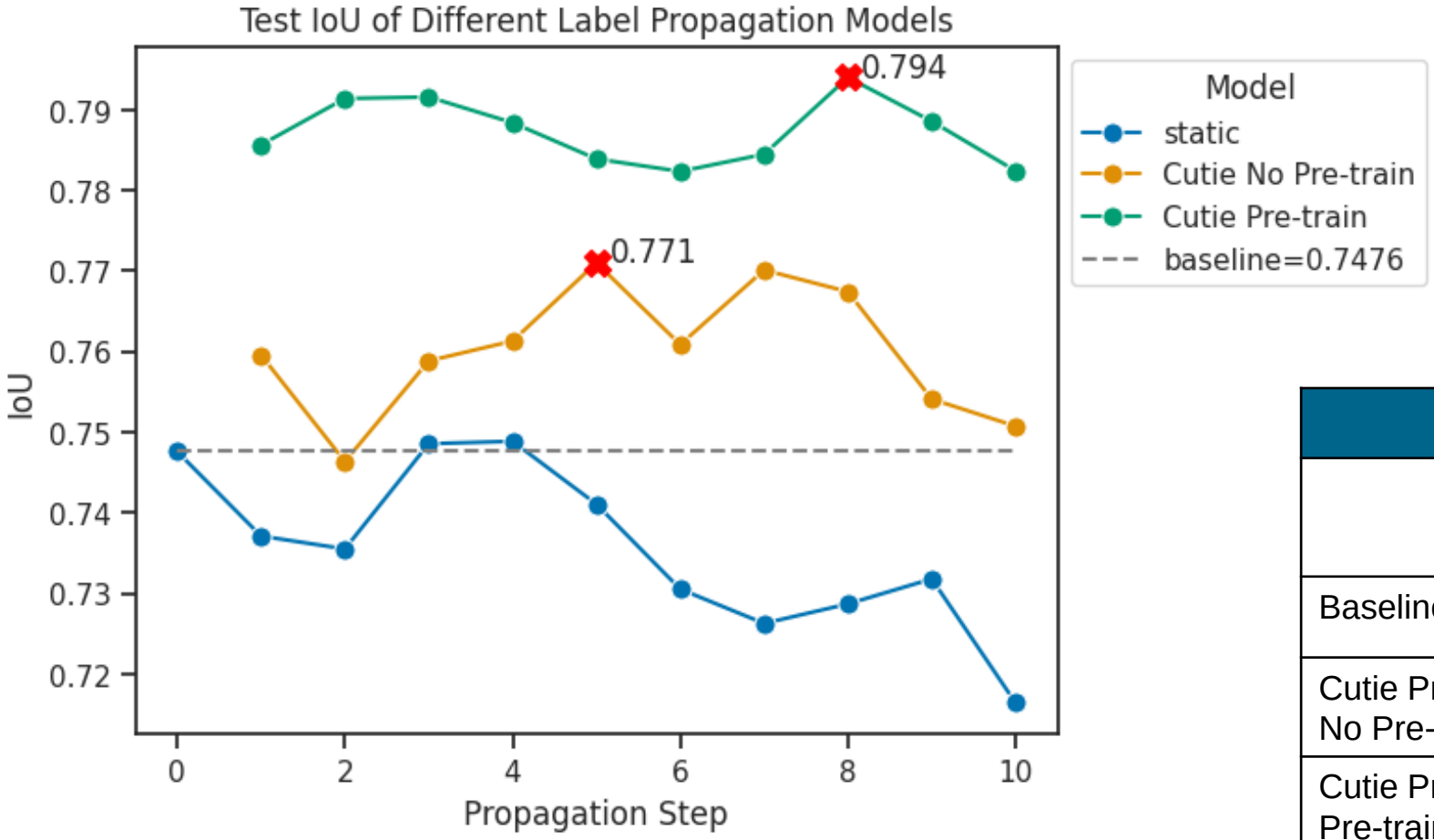



https://learnopencv.com/wp-content/uploads/2022/06/5-segmentation-iou.jpg

$$mean\ IoU = \frac{1}{|Classes|} \sum_{c \in Classes} \frac{Target_c \cap Prediction_c}{Target_c \cup Prediction_c}$$

$$= \frac{1}{|Classes|} \sum_{c \in Classes} \frac{TP_c}{TP_c \cup FP_c \cup FN_c}$$

# Experiment Results: Pre-train Cutie




Static: same cloud mask along propagation steps

| Test Results | | |
|---|---|---|
| | Test IoU | Test IoU over baseline [%] |
| Baseline | 0.746 | - |
| Cutie Prop. Steps 5 No Pre-train | **0.771** | **3.4** |
| Cutie Prop. Steps 8 Pre-train | **0.794** | **6.4** |

Pre-train: 4k iterations on the weakly labeled dataset (sun elevation above 10º).

Training: 4k iterations on the automatically labeled dataset.

# Summary, Conclusion and Outlook

## Summary:

- A strategy for automatically generating labeled datasets using SVOS has been proposed.
- Performance degradation for increasing mask propagation steps has been investigated.
- Pre-training has been integrated into the process

## Conclusion:

- The use of automatically labelled dataset with SVOS Cutie significantly outperform the baseline model.
- Pre-training with weakly-labeled data further improves performance.
- The combination of the SVOS (Cutie) with pre-training offers the best overall performance

## Outlook:

- Incorporate temporal information directly into model training
- Introduce confidence-weighted training for propagated labels
- Apply label relaxation to account for uncertainty at class boundaries

**Thank you for your attention!**
**Questions?**

yann.fabel@dlr.de
eduardo.saezmartinez@dlr.de

# References

[1] Y. Fabel, “Cloud Segmentation and Classification from All-Sky Images Using Deep Learning,” 2020.

[2] D. Magiera, “Semi-supervised Learning for Probabilistic Cloud Detection in Ground-based Imagery,” 2024.

[3] H. K. Cheng, S. W. Oh, B. Price, J.-Y. Lee, and A. Schwing, “Putting the Object Back into Video Object Segmentation,” 2023.