To mask or not to mask, that is the question.

**Background**
- **SLAM**: Simultaneous Localization and Mapping
- **Dynamic SLAM**: SLAM in dynamic environments. Track and match image features that are not on dynamic objects.

**Method: SLAM with Temporal Masking**

**New paradigm**: decide when to mask certain classes.

**Includes Temporal Masking Network trained with automatically annotated dataset.**

**Automatic dataset annotation**

1. Sample temporal masks uniformly (no masking decisions)
2. Benchmark all temporal masks
3. Aggregate temporal masks according to performance (e.g., our metric, USM)

Sample temporal masks uniformly among all possible temporal masks (space size: $2^n$) using a binary tree whose root-to-leaf iterations preserve sampling uniformity.

**Additional contributions**: USM Metric + ConsInv Dataset (150 sequences)

**Results**

To mask or not to mask, our network shall learn.

**Preventing drift in TUM RGB-D dataset**

**Preventing drift and excessive masking in ConsInv-Outdoors dataset**

**Conclusion**

With the proposed Temporal Masking paradigm, we overcome the current limits of Dynamic SLAM on real data, especially in difficult scenarios.