E²SAM: A Pipeline for Efficiently Extending SAM's Capability on Cross-Modality Data via Knowledge Inheritance

Sundingkai Su, Mengqiu Xu, Kaixin Chen, Ming Wu, Chuang Zhang

Contributions

- Proposing a pipeline to efficiently extending SAM's capability on cross-modality.
- No need to introduce additional data and any annotations since the model can be trained unsupervised.

Analyzing

To demonstrate the effectiveness of our proposed method, we design two sets of controlled experiments by comparing the input of different modalities and the advantages of inheriting the powerful capabilities of SAM. The quantitative results are as shown in tables, where 'Rand' indicates random initialization and 'SAM' means performing different operations based on SAM as the pre-trained model.

Conculusion

In this paper, in order to realize the capability expansion of SAM based on cross-modality data,

we propose a universal two-stage pipeline, which is the knowledge inheritance stage for inheriting SAM capability, and the fine-tuning stage for better downstream adaptation. An auxiliary branch including a Channel Selector and Merge Module is designed to separate different cross-modality to achieve feature alignment. It is worth mentioning that we do not need to lead into additional data and labels during the knowledge inheritance process, reducing the cost of collecting and annotating data. Through meticulous experiments and visualization results, it can be demonstrated that our method can efficiently inherit the ability of SAM on cross-modality data without compressing data.

References


References