

Forschungspraxis, Interdisciplinary Project

Scene graph generation models analysis using Visual Genome benchmark and self-recorded datasets

Scene graphs were first proposed [1] as a data structure that describes the object instances in a scene and the relationships between these objects. The nodes in the scene graph represent the detected target objects, whereas the edges denote the detected pairwise relationships. A complete scene graph can represent the detailed semantics of a dataset of scenes.

Scene graph generation (SGG) is a visual detection task for building structured scene graphs. In this work, we would like to compare the three traditional SGG models: Neural Motifs [2], Graph R-CNN [3], and IMP [4]. We will train and evaluate the models using the Visual Genome dataset and other commonly used datasets. In addition, our dataset will be annotated and then utilized.

[1]: J. Johnson, et al. "Image retrieval using scene graphs."

[2]: R. Zellers, et al. "Neural motifs: Scene graph parsing with global context."

[3]: J. Yang, et al. "Graph r-cnn for scene graph generation."

[4]: D. Xu, et al. "Scene graph generation by iterative message passing."

Prerequisites

- Good Programming Skills (Python, GUI design)
- Knowledge about Ubuntu/Linux/Pytorch
- Knowledge about Computer vision/Neural network
- Motivation to learn and conduct research

Contact

dong.yang@tum.de

Advisors

Dong Yang, Xiao Xu