Recent VQA Approaches

Wentao Mo[1]

[1]Department of Machine Intelligence
Peking University

2021 年 4 月 27 日

Outline

1. Visual Feature Extractor
2. Feature Fusion Methods

Visual Commonsense R-CNN

VinVL

MFH: Multimodal Factorized High-order Attention

MCAN: Deep Modular Co-Attention Networks

TRRNet: Tiered Relation Reasoning for Compositional VQA

OSCAR & UNITER: Transformer-like Fusion