

Training Scene Graph Generator with Synthetic Images

中山研究室 修士一年 増井 建斗 48156621

2015 年 12 月 18 日

1 概要

風景や物体などの一般的な環境を撮影した画像から、画像の内の情報を Scene Graph[1] と呼ばれるグラフ構造へ変換するモデルを教師あり学習する。Scene Graph は画像内の情報を、物体とその属性のペアによるグラフ構造で表現したものである。教師あり学習に必要な画像と Scene Graph は制作に多くの手間がかかり、学習に十分なデータ・セットが用意されていないため、コンピュータグラフィックスによって自動生成したデータ・セットを用いて学習を行う。

2 背景

家庭内で家事を行うロボットの意思決定や画像検索においては、画像内の物体認識に加えて、それら物体同士の位置関係などの2物体以上に渡る関係性も重要な情報となる。従来の画像認識では画像内の物体のラベル付けと位置検出が行われているが、物体同士の関係性を含む情報を抽出する試みには、多くの研究余地が残されている。この研究の最終的な目標は、画像から物体同士の関係を含む情報を抽出することで家事ロボットによる意思決定を補助することであるが、画像検索などの問題にも適用可能である。関連研究として Image Retrieval using Scene Graphs[1] があるが、彼らは検索対象の画像全てに Scene Graph があらかじめ作成されている前提で、自然言語による検索文と Scene Graph のマッチングを行い高精度の画像検索を実現した。データセットの作成にはクラウドソーシングが利用され、全て人力で Scene Graph が作成されている。この研究では画像からの Scene Graph を行うモデルを教師あり学習するが、彼らが用意したデータセットより多くの教師データを用いるため、コンピュータ・グラフィックスによる

データセットの自動生成を行うこととした。

3 問題設定

4 データセットの作成

5 提案手法

6 結論

参考文献

- [1] Justin Johnson, Ranjay Krishna, Michael Stark, Li-Jia Li, David Shamma, Michael Bernstein, and Li Fei-Fei. Image retrieval using scene graphs. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 3668–3678, 2015.