

ОБЗОР МЕТОДОВ ИЗВЛЕЧЕНИЯ ГРАФОВ ЗНАНИЙ ИЗ ИЗОБРАЖЕНИЙ

Дубовской А.А. (Национальный исследовательский университет ИТМО)

Научный руководитель – доцент, к.т.н. Баймуратов И.Р.

(Национальный исследовательский университет ИТМО)

Введение. Создание, управление и использование коллекций изображений требует большого количества трудозатрат. Эти задачи могут быть частично автоматизированы с помощью извлечения из коллекций изображений графов знаний. На данный момент уже предложен ряд подобных решений, однако не существует систематического обзора данной исследовательской области. Таким образом, в статье представлен обзор литературы по методам извлечения графов знаний из изображений, представленных различными архитектурами нейронных сетей, а также их применения для улучшений результатов классификации объектов.

Основная часть. Поиск источников осуществлялся в системе Google Scholar по ключевым словам: “knowledge graph extraction from images”, “ontology extraction from images”, “ontology extraction from videos”, “knowledge graph extraction from videos”. Результаты поиска отбирались по следующим критериям:

- 1) Статья была опубликована не ранее 2016 года.
- 2) Статья включает описание архитектуры решения, приведенные в статье эксперименты воспроизводимы.

В результате систематического анализа статей, подходящих под критерии, были сформированы две группы по целям исследований: 1) извлечение графа знаний из изображений, 2) применение графов знаний для улучшения точности классификации. Общим принципом, применяемым во всех найденных статьях, является использование сверточной нейронной сети. Для извлечения сущностей и их связей на изображениях авторы применяют такие архитектуры, как VGG16 [1][2], VGG19 [3], YOLO [4][5], Faster R-CNN [6], Mask R-CNN [7] и IntersectionResNetV2 [8][9]. Дальнейшую работу с графами знаний можно разделить на две группы: генерация графа знаний с нуля и дополнение уже существующего графа знаний. В первом случае применяются такие модели как HRE [1] и GSNN [6]. Во втором случае используются графы знаний WordNet [3], MMKG [2], ConceptNet [5], IKG [8] [9] и VEKG [4]. Результаты оцениваются различными метриками, наиболее часто используется точность, наилучший показатель был достигнут в решении [9] – 82,43%.

Выводы. Результатом является систематический обзор литературы на тему извлечения графов знаний из изображений, а также их применения для улучшения результатов классификации объектов.

Список использованных источников:

1. Mao J., Yao Y., Heinrich S., Hinz T., Weber C., Wermter S., Liu Z., Sun M. Bootstrapping Knowledge Graphs From Images and Text // *Frontiers in Neurorobotics*. – 2019. - № 13. – P. 93.
2. Liu Y., Li H., Garcia-Duran A., Niepert M., Onoro-Rubio D., Rosenblum D.S. MMKG: Multi-modal Knowledge Graphs // *The Semantic Web*. – 2019. – P. 459–474.
3. Mahon L., Giunchiglia E., Li B., Lukasiewicz T. Knowledge Graph Extraction from Videos // *19th IEEE International Conference on Machine Learning and Applications (ICMLA)*. – 2020. – P. 25–32.
4. Yadav P., Curry E. VEKG: Video Event Knowledge Graph to Represent Video Streams for Complex Event Pattern Matching // *First International Conference on Graph Computing (GC)*. – 2019. – P. 13–20.

5. Zhou Y., Sun Y., Honavar V. Improving Image Captioning by Leveraging Knowledge Graphs // IEEE Winter Conference on Applications of Computer Vision (WACV). – 2019. – P. 283–293.
6. Marino K., Salakhutdinov R., Gupta A. The More You Know: Using Knowledge Graphs for Image Classification. – 2016.
7. Fang W., Ma L., Love P.E.D., Luo H., Ding L., Zhou A. Knowledge graph for identifying hazards on construction sites: Integrating computer vision with ontology // Automation in Construction. – 2020. – № 119.
8. Menglong C., Detao J., Ting Z., Dehai Z., Cheng X., Zhibo C., Xiaoqiang X. Image Classification Based on Image Knowledge Graph and Semantics // IEEE 23rd International Conference on Computer Supported Cooperative Work in Design (CSCWD). – 2019. – P. 81–86.
9. Zhang D., Cui M., Yang Y., Yang P., Xie C., Liu D., Yu B., Chen Z. Knowledge Graph-Based Image Classification Refinement // IEEE Access. – 2019. – № 7. – P. 57678–57690.