



FashionGraph

understanding fashion data using scene graph generation

Shabnam Sadegharmaki¹; Marc A. Kastner²; Shin'ichi Satoh²
¹Technical University of Munich, ²National Institute of Informatics, Tokyo

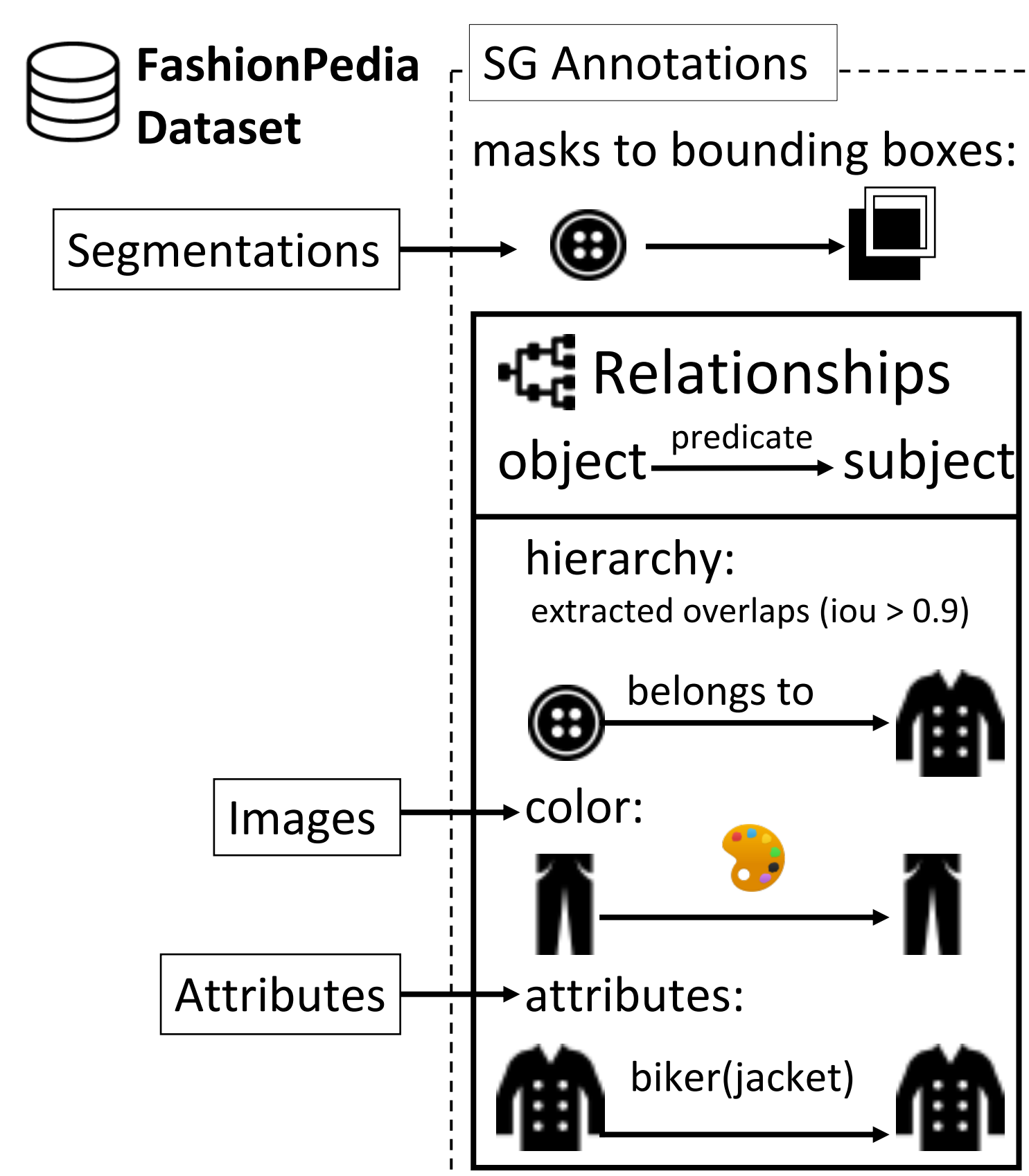


Motivation

- Brought the idea of scene graphs to fashion images.
 - ↳ helps in better understanding of fine-grained fashion data.
- A model to generate fashion scene graphs.
 - ↳ using object and relationship detection models.
 - ↳ we generated new annotations for this purpose.
- Integrated the attribute detection into the scene graph model.
- Highlighted the application of SG for fashion image retrieval.

Data Annotation

We provide relation detection annotations for Fashionpedia dataset[1].



To train a SG generator, we need the following annotations:

- Fine-grained segmentation: bounding boxes and object labels
- Relationships in the format of **object**, **subject**, **predicate**

Relationships for a fashion data include:

- Hierarchical. E.g. **pocket belongs to shirt**
- Attributes. E.g. **dress is A-line**
- Color. E.g. **Jacket is blue**

Architecture

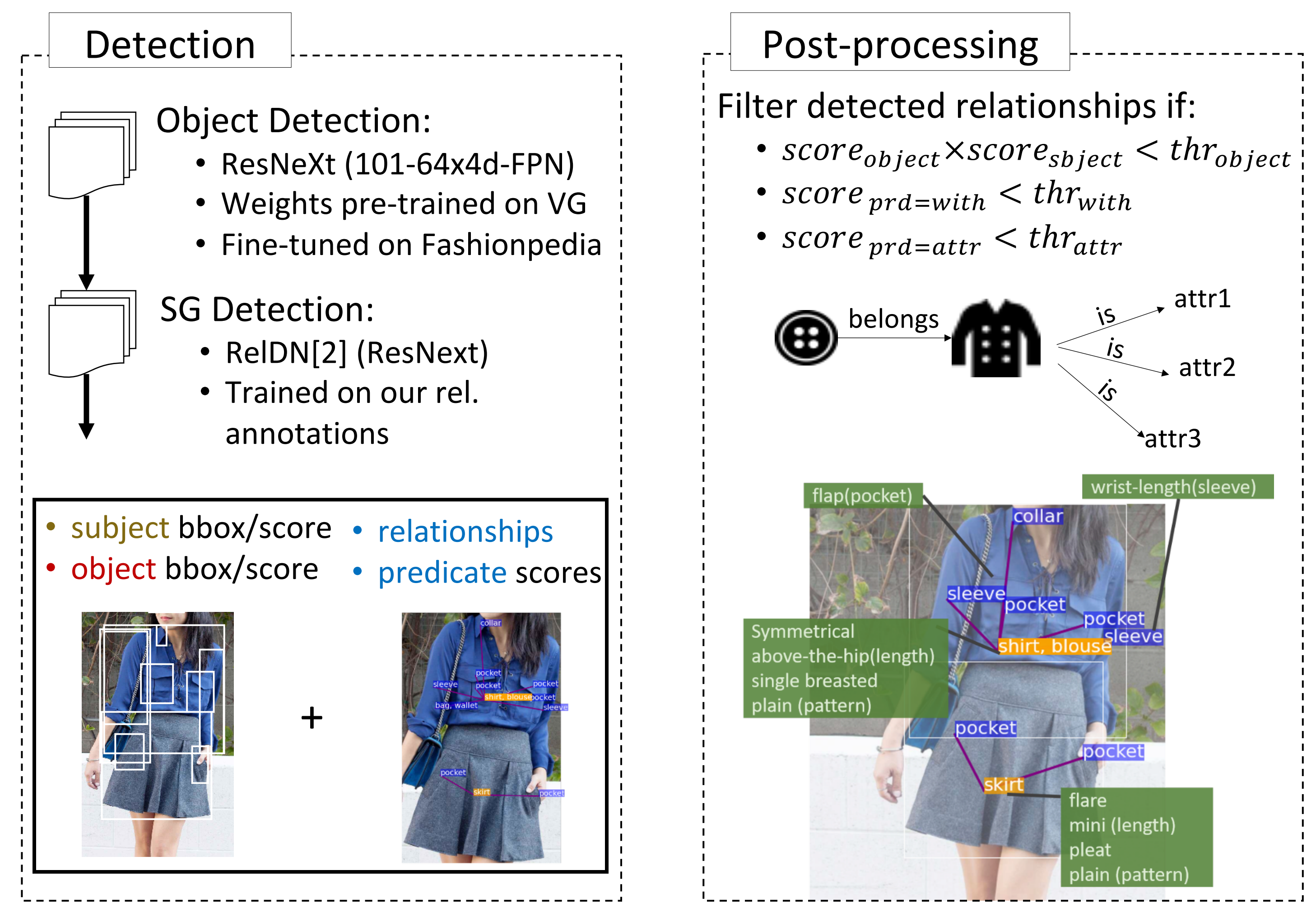
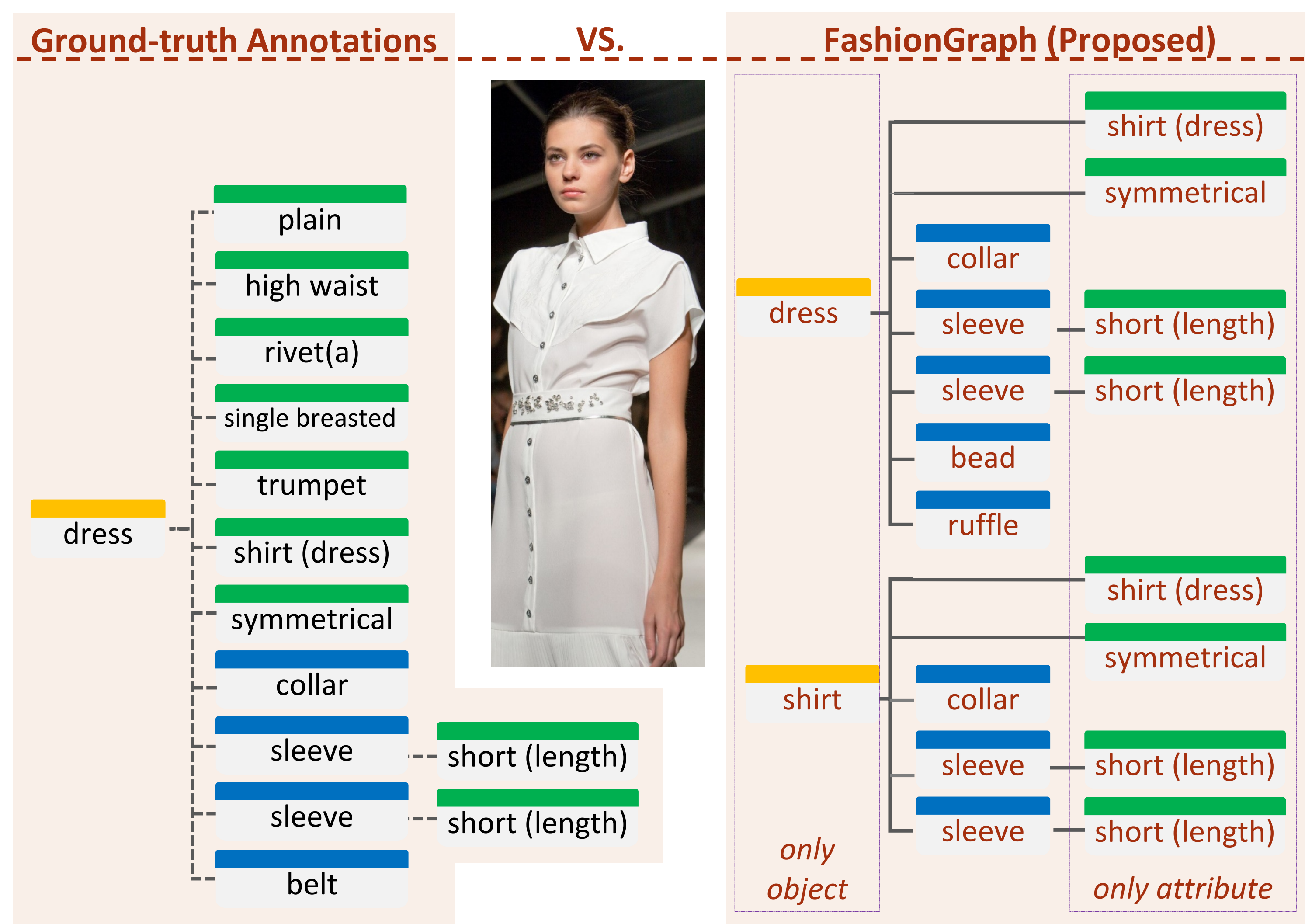


Table 1. ReIDN[2] on Visual Genome vs. FashionGraph on Fashionpedia

Method	Dataset	R@20	R@50	R@100
ReIDN[2]	Visual Genome	23	31	37
FashionGraph	Fashionpedia	18	22	24

Figure 1. Example of ground-truth annotations vs. predicted graph



Application: Image Retrieval

- To rank the images for a given query image, we represent the predicted scene graph by four matrices:

- Objects
- Hierarchical relationships
- Attributes
- Colors

- Then we calculate the cosine similarity of each type between the images and the query.

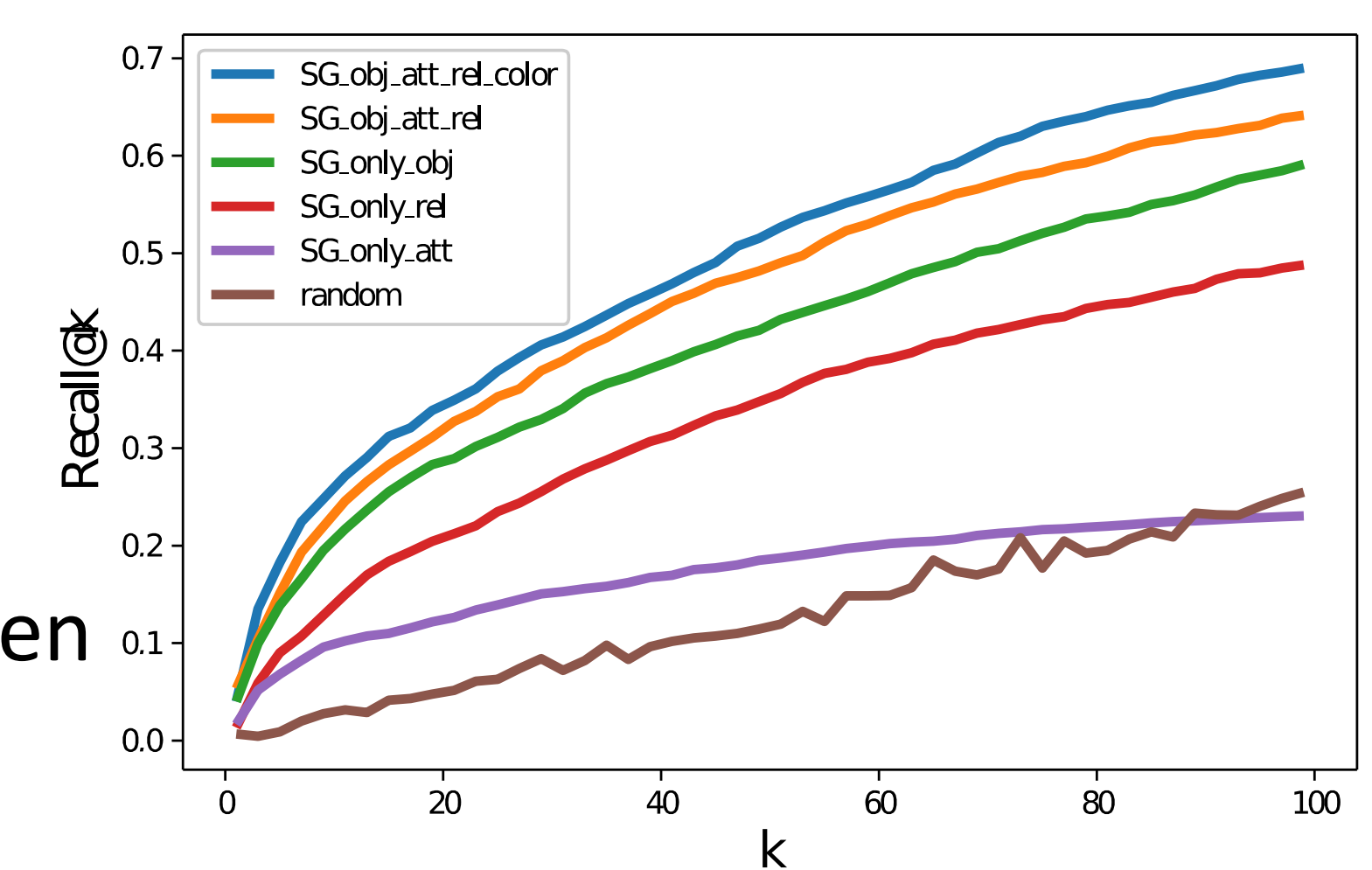


Figure 2. Qualitative evaluation on fashion image retrieval

Method	Query	Results
Ground-truth (Baseline)		
FashionGraph (Proposed)		

Contact

sadeghar@in.tum.de

Code

<https://github.com/shabnamsadegh/FashionGraph>

References

- Jia, Menglin, et al. "Fashionpedia: Ontology, Segmentation, and an Attribute Localization Dataset." arXiv preprint arXiv:2004.12276 (2020).
- Zhang, Ji, et al. "Graphical contrastive losses for scene graph parsing." Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2019.