# Notes of "Knowledge-Embedded Routing Network for Scene Graph Generation"

Jihang Li

January 17, 2020

## Contents

Contents 1

1 Introduction 1

4 Experiments 2

## Abstract

In this work, it is found that the statistical correlations between object pairs and their relationships can effectively regularize semantic space and make prediction less ambiguous.

## 1 Introduction

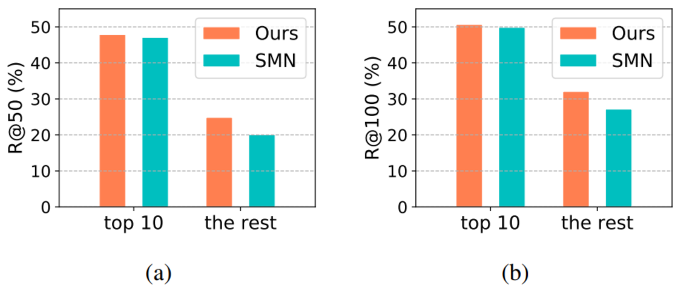



Figure 1: (a) Recall@50 and (b) Recall@100 of our proposed method and the SMN on the scene graph classification task on the Visual Genome dataset. Both models are trained on the whole training set and evaluated on the two subsets, respectively. Note that SMN is the previous best-performing method.

As shown in Figure 1, current best-performing method (*i.e.*, SMN) can achieve competitive performance if it has sufficient training samples, but its performance suffers from a severe drop otherwise.

Objects in visual scene commonly have strongly structured regularities (according to SMN).

A novel Knowledge-Embedded Routing Network (KERN) is introduced, which captures the interplay of target objects and their relationships under the explicit guidance of prior statistical knowledge and automatically mines contextual cues to facilitate scene graph generation.

Existing works utilize the recall@$K$ (short as R@$K$, according to VRD) as the evaluation metric.

***Original Texts***

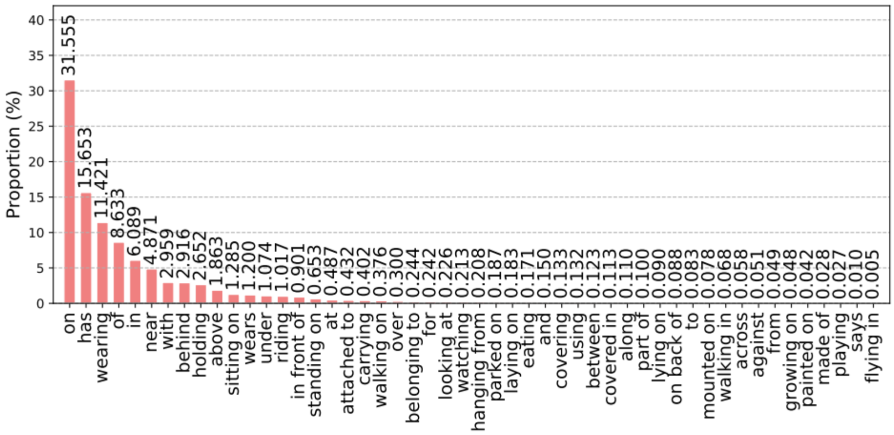
However, this metric is easily dominated by the performance of the relationships with a large proportion of samples. As the distribution of different relationships is severely uneven, if one method performs well on several most frequent relationships, it can achieve a high R@$K$ score. Thus, it can not well measure the performance of all relationships.

A mean R@$K$ is proposed as complimentary evaluation metric:

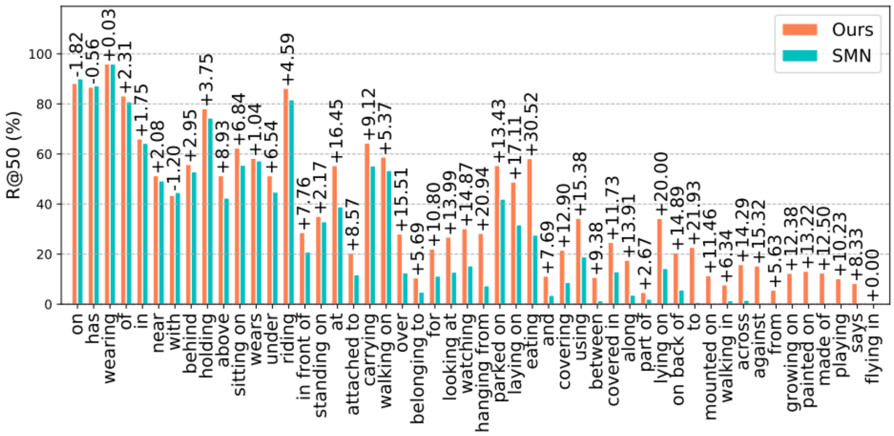
1. computes the R@$K$ for samples of each relationships
2. averages over all relationships to obtain mR@$K$

The KERN model improves the mR@50 and mR@100 from 15.4% and 20.6% to 19.8% and 26.2% on the scene graph classification task, with relative improvements of 28.6% and 27.2% respectively.

## 4 Experiments




(a)




(b)

Figure 2: (a) The distribution of different relationships on the VG dataset. The training and test splits share similar distribution. (b) The R@50 without constraint of our method and the SMN on the predicate classification task on the VG dataset.