Entity-Focused Dense Passage Retrieval for Outside-Knowledge Visual Question Answering

Jialin Wu and Raymond J. Mooney

Knowledge Retrieval for OK-VQA

Q: This plush toy was named after what US president?

The name **teddy bear** comes from former United States President **Theodore Roosevelt**, who was often referred to as “Teddy”

**EnFoRe Approach**

---

**Step 1: Entities Set Construction**

Q: This plush toy was named after what US president? A: Teddy Roosevelt

**Question-based Entities:**
Phrases in question: plush toy, president, Answer candidates: Roosevelt, ...
Sub-questions: Teddy bear, ...

**Image-based Entities:**
Azure Tagging: Teddy bear, ...
Clip Tagging: Teddy bear, ...
Phrases in captions: Infant, ...

---

**Step 2: Oracle entities detection**

Is the large yellow object a fruit or a vegetable?

- **Sparse Passage Retriever**
  - A banana is an elongated, edible fruit – botanically a berry – produced by several kinds of large herbaceous
  - 1/3 = 0.33

Is the large yellow object a fruit or a vegetable?

- **Bell Pepper**
  - Bell pepper The bell pepper is in different colors, including red, yellow, orange, ... they are commonly used as a vegetable ingredient
  - 1/1+1/4 = 1.25

---

**Step 3: Entity-Focused Dense Retrieval**

Q: This plush toy was named after what US president?

**Query**

plush toy

**Entities**

president
teddy bear (oracle)
infant

**Passage 1**

A teddy bear is a stuffed toy in the form of a bear ..., and named after President Theodore Roosevelt ..., the teddy bear ...

**VQA Results on OKVQA**

![VQA Score Chart]

<table>
<thead>
<tr>
<th>KAT-base</th>
<th>KAT-full</th>
<th>KAT-full + EnFoRe</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.49</td>
<td>0.552</td>
<td>0.552</td>
</tr>
<tr>
<td>0.513</td>
<td>0.544</td>
<td>0.544</td>
</tr>
<tr>
<td>0.537</td>
<td>0.522</td>
<td>0.522</td>
</tr>
</tbody>
</table>

**Why Entity-Focused Retrieval**

---

**Fitting in the VQA context**

**Q:** What holiday is this? A: Thanksgiving.

**Critical entity:** turkey

**Q:** Is the large yellow object a fruit or a vegetable? A: vegetable.

**Critical entity:** bell pepper

---

**Human Evaluation**

Judging whether KAT knowledge or EnFoRe knowledge contains the most supportive evidence to the answer?

Q: What is the main benefits of healthy properties these fruit contain? A: potassium

Naturally, the fruit of the Hassaku orange is a good source of vitamin C, folic acid, potassium and fiber.

Fruit vegetable is a fruit commonly referred to as a vegetable because they are savory (not sweet).