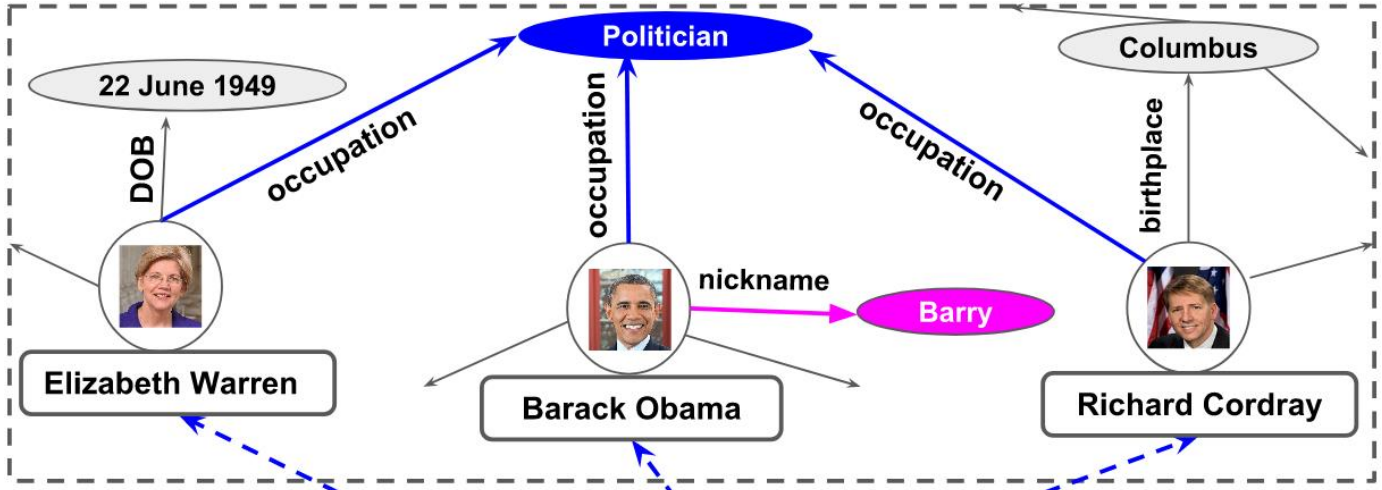
P7214

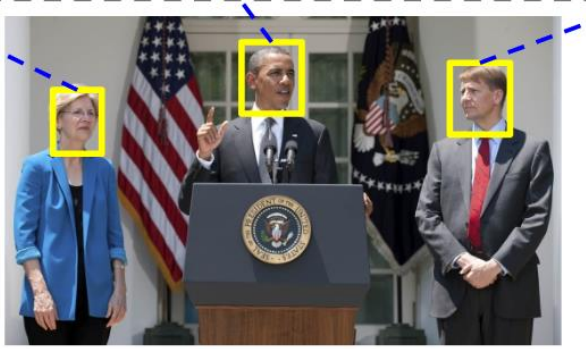
KVQA: Knowledge-aware Visual Question Answering

Sanket Shah[1*], Anand Mishra[2*], Naganand Yadati[2] and Partha Pratim Talukdar[2]

[1]IIIT Hyderbad, India [2]IISc Bangalore, India (*: Contributed equally)



A task which requires access to world Knowledge Graph



Conventional VQA

Q: How many people are there in the image?

A: 3

vs. KVQA (Ours)

Q: Do all the people in the image have a common occupation?

A: Yes

P7214

KVQA: Knowledge-aware Visual Question Answering

Sanket Shah[1*], Anand Mishra[2*], Naganand Yadati[2] and Partha Pratim Talukdar[2]

[1]IIIT Hyderbad, India [2]IISc Bangalore, India (*: Contributed equally)

Images containing proper nouns (named entities)

- Many datasets but **no bridge** between images and Knowledge Graphs
- **This work**
- A novel task and dataset: KVQA

KVQA - first dataset for the task of world knowledge-enabled VQA

- **24K images** covering **18K persons**
- Associated **Knowledge Graph facts**
- **183K Question-Answer pairs**
- State of the art is not adequate!

Available: http://malllabiisc.github.io/resources/kvqa/