Josiah Wang

Lucia Specia

Phrase Localization Without Paired Training Examples

Imperial College London

MultiMT

Newton Fund

Phrase Localization

the handles of the slide

How is this tackled?

Phrase Localization

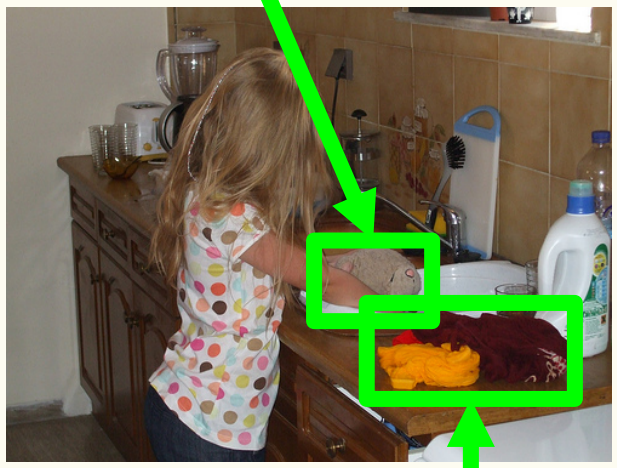
with strong supervision

Phrase Localization

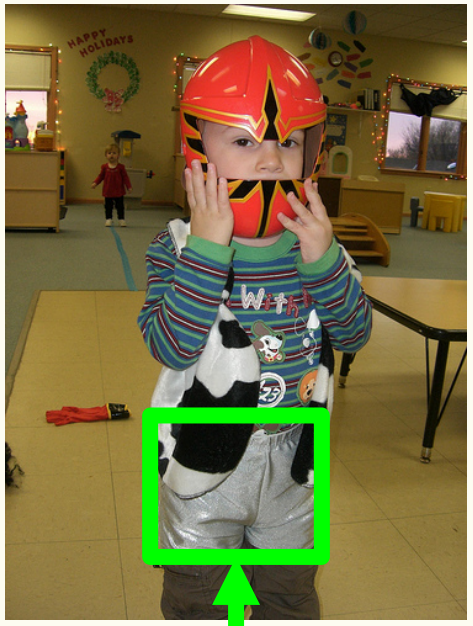
with strong supervision

the handles of the slide

a stuffed animal



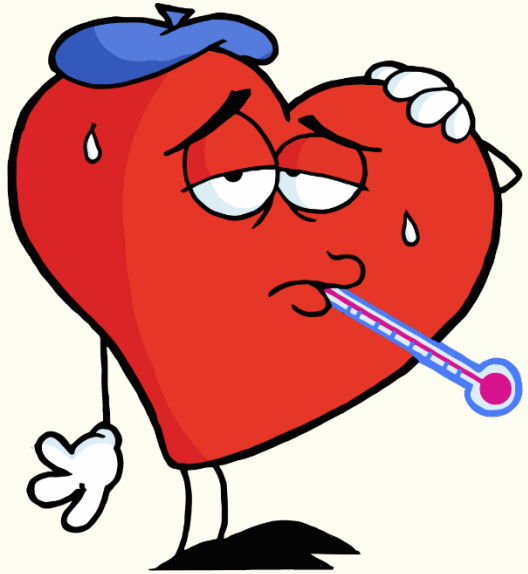
a few wet items



silver shorts

Phrase Localization

with weak supervision

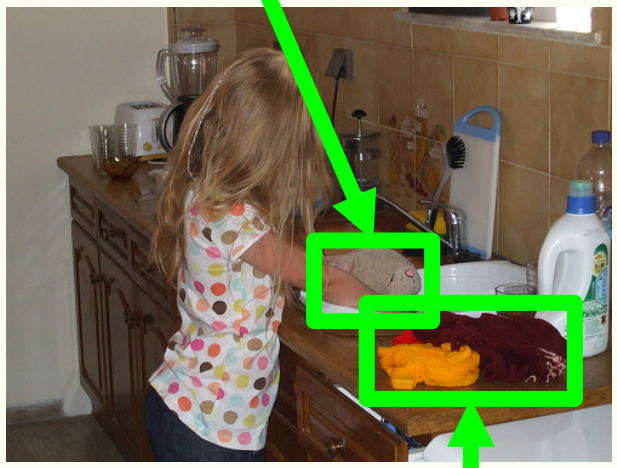


Phrase Localization

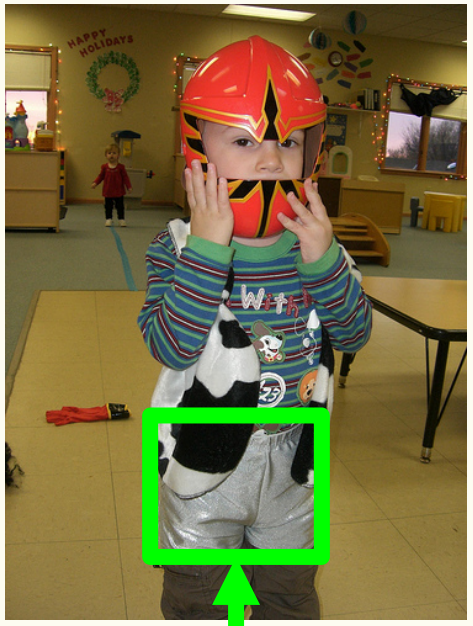
with weak supervision

the handles of the slide

a stuffed animal



a few wet items



silver shorts

Phrase Localization

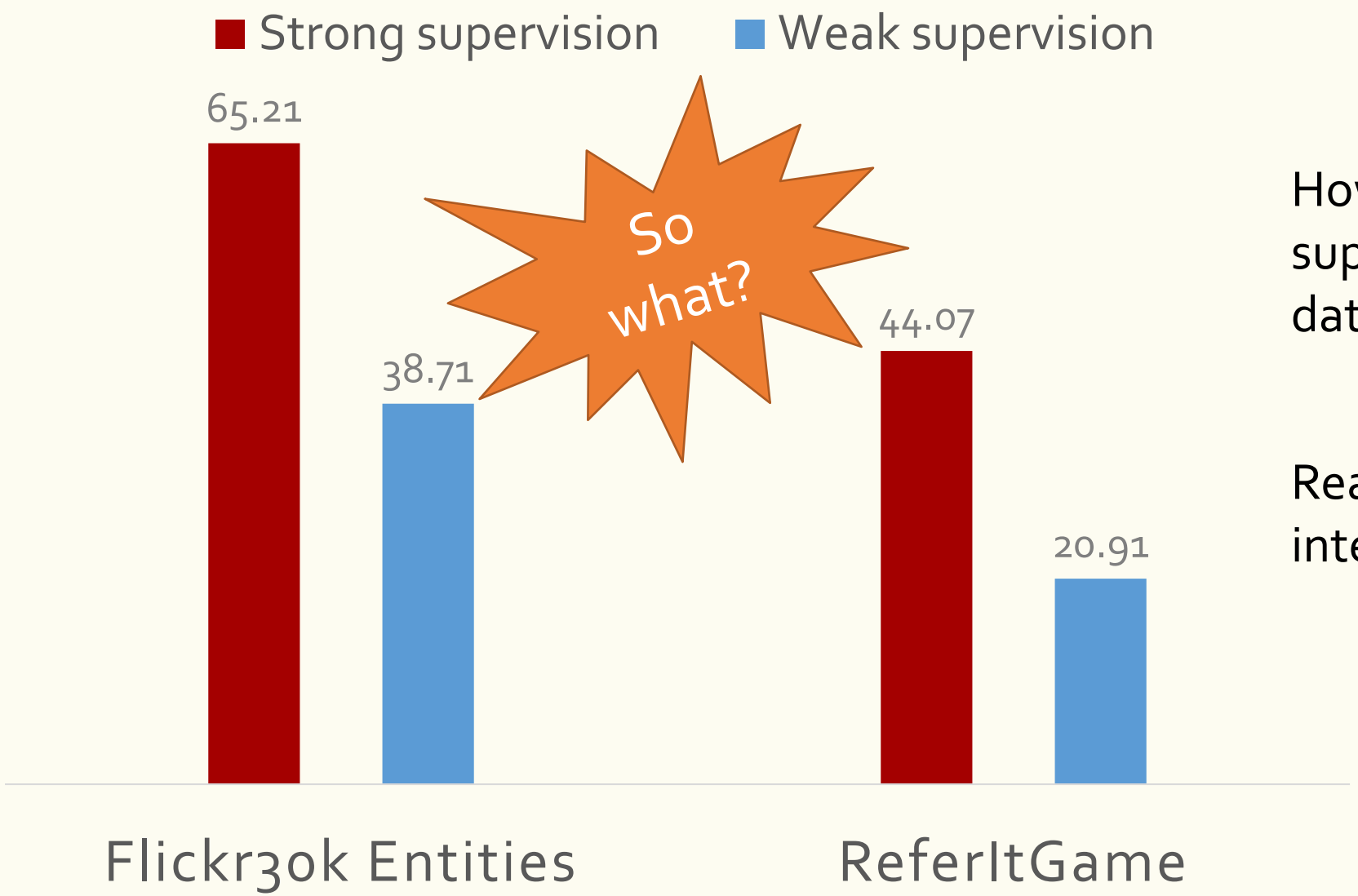
with weak supervision

the handles of the slide

a stuffed animal

a few wet items

silver shorts



How is the supervised data used?

Really intelligent?

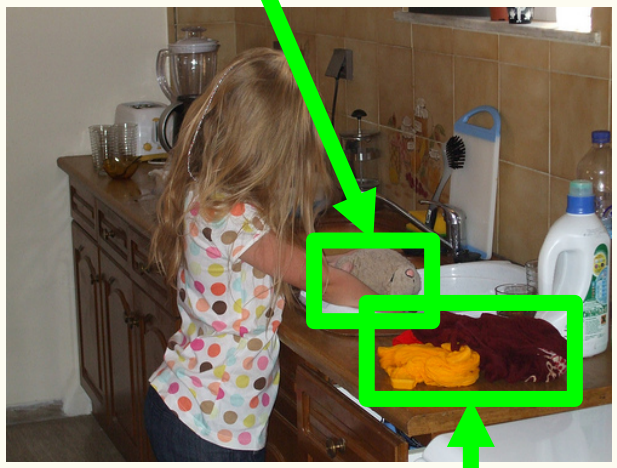
Memorize
all the phrases
with their
images!!!!

Phrase Localization

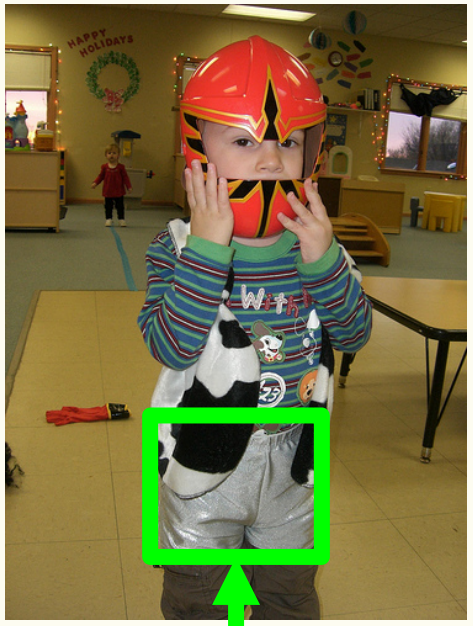
with NO paired data

the handles of the slide

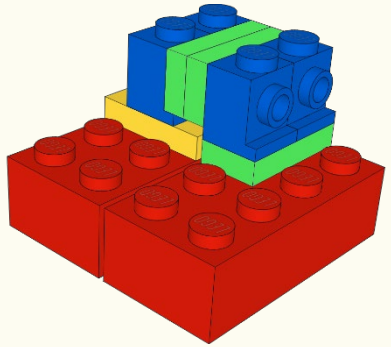
a stuffed animal



a few wet items



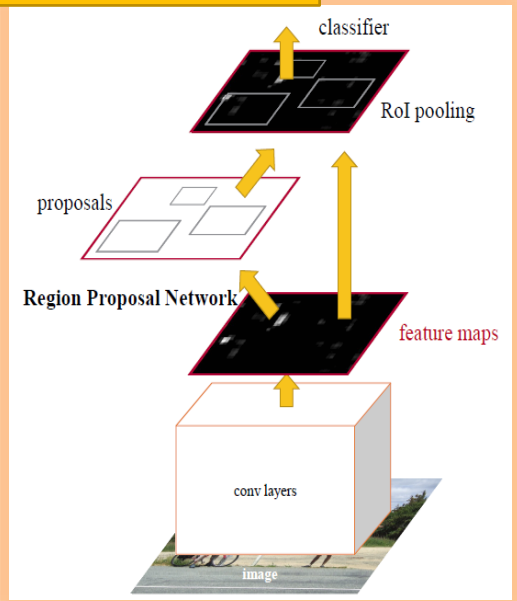
silver shorts



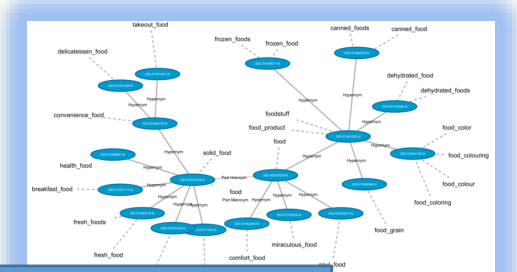
Phrase Localization

with NO paired data

Off-the-shelf object detectors

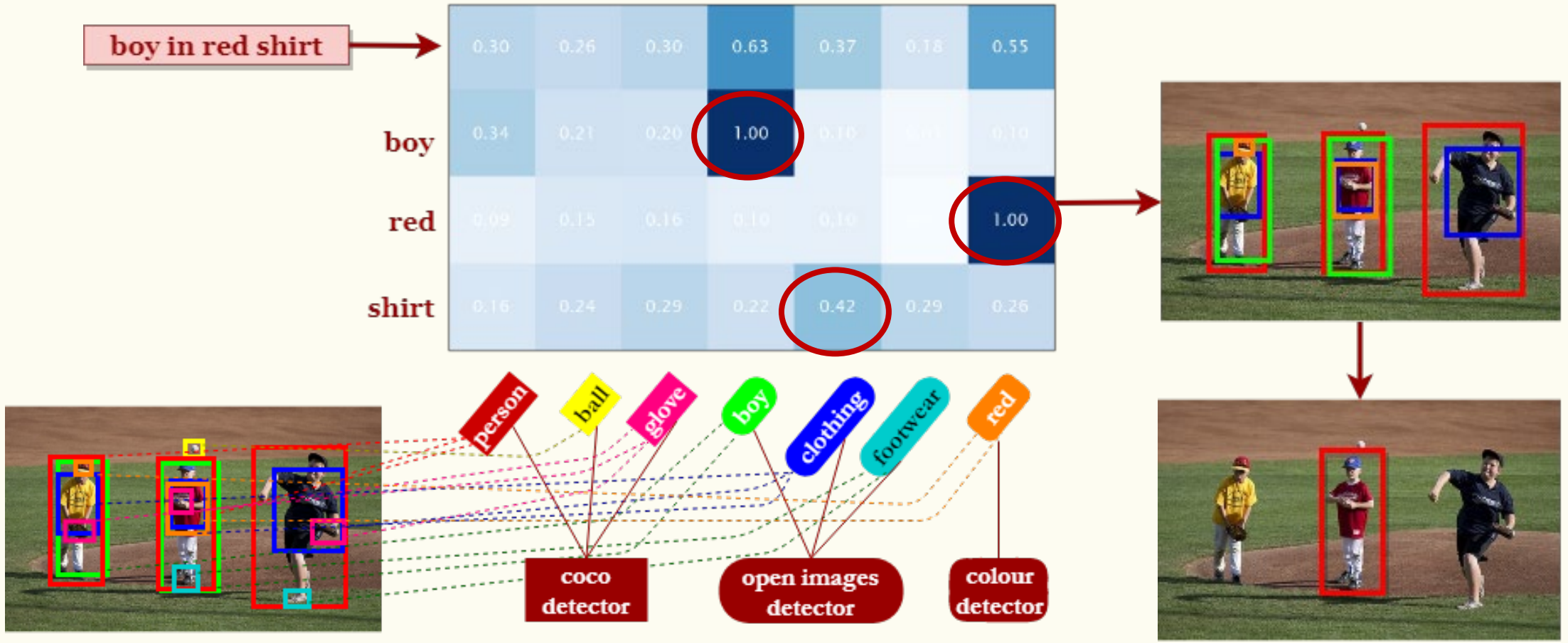


Text corpora/ Word embeddings



Knowledge bases

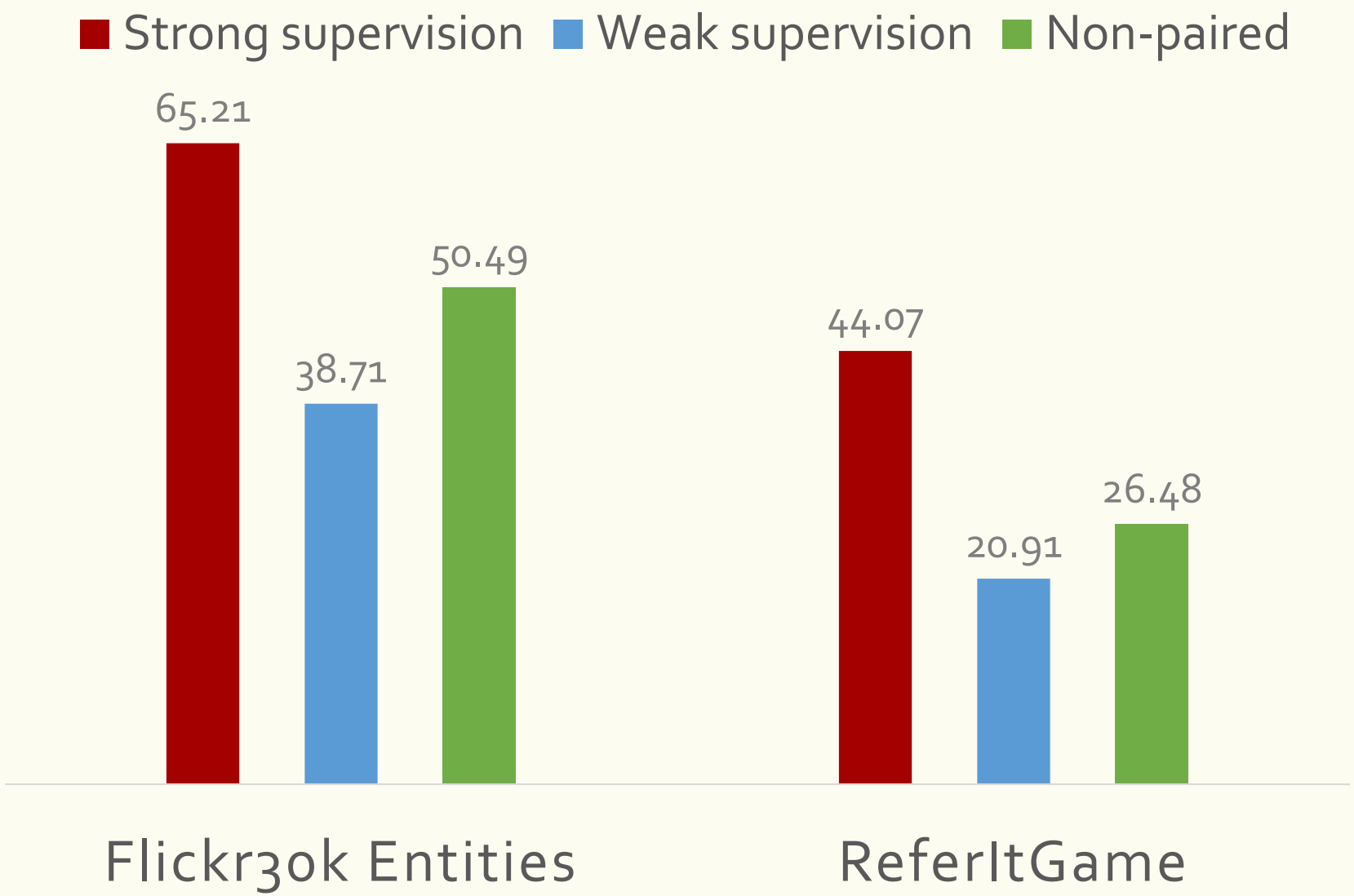
Generic images



1 Detect instances

2 Filter concepts

3 Localize


Strong supervision
Weak supervision
Non-paired
65.21
38.71
50.49
44.07
20.91
26.48
Flickr30k Entities
ReferItGame

Non-paired setting as a strong baseline

Simpler general methods *vs.* complex specialized methods

More effective use of paired data

…on top of what can already be achieved without.
Better generalization?

Poster #24

Josiah Wang

Lucia Specia

Phrase Localization Without Paired Training Examples

Imperial College London

MultiMT

Newton Fund