

**NANYANG
TECHNOLOGICAL
UNIVERSITY**
SINGAPORE

RIOSE
Rapid-Rich Object Search Lab
博云搜索实验室


Alibaba Group
阿里巴巴集团


2018
CVPR
SALT LAKE CITY • JUNE 18-22

Look, Imagine and Match: Improving Textual-Visual Cross-Modal Retrieval with Generative Models



Jiuxiang Gu



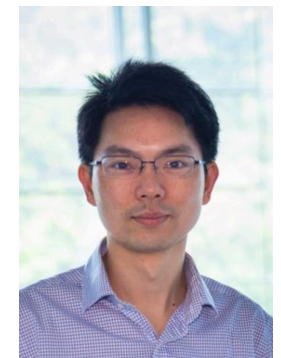
Jianfei Cai



Shafiq Joty



Li Niu



Gang Wang

Goal

Image-to-Text Retrieval



A young man doing a skateboard trick while others watch

A man doing a skate trick during a competition event with a audience

Guys on a course made for skate boarding

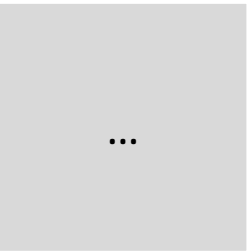
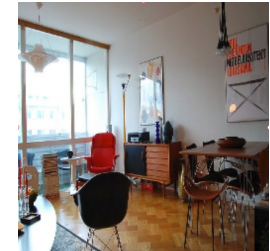
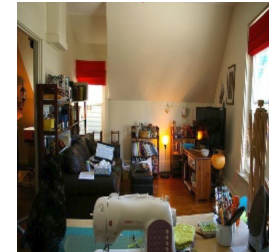
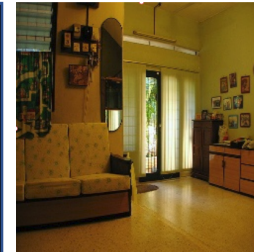
A group of people doing skateboarding tricks on a car

A boy riding on his skateboard at a skate park while other guys watch

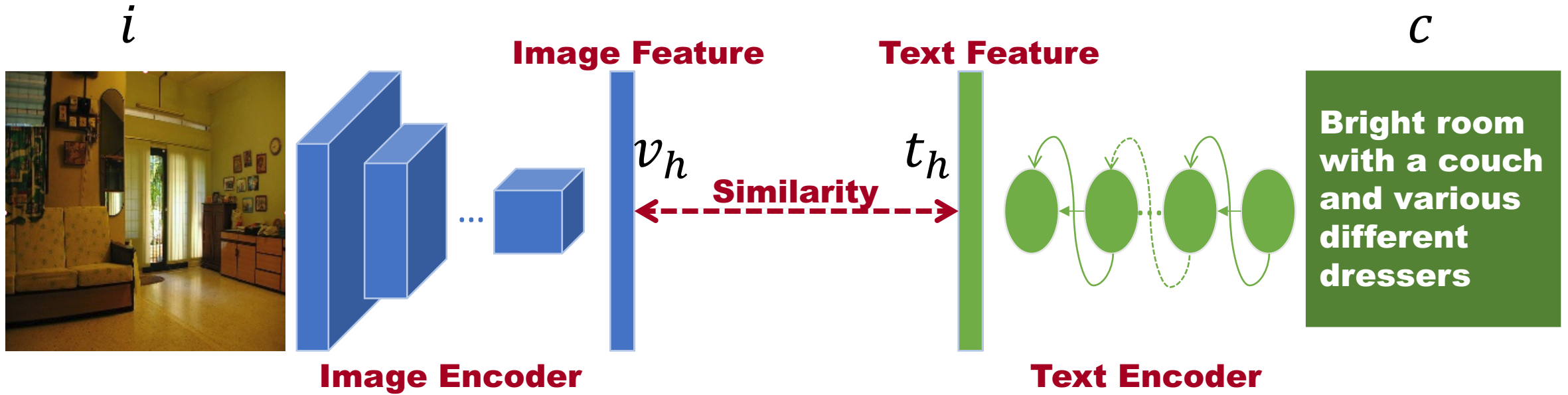
...

Text-to-Image Retrieval

Bright room with a couch and various different dressers



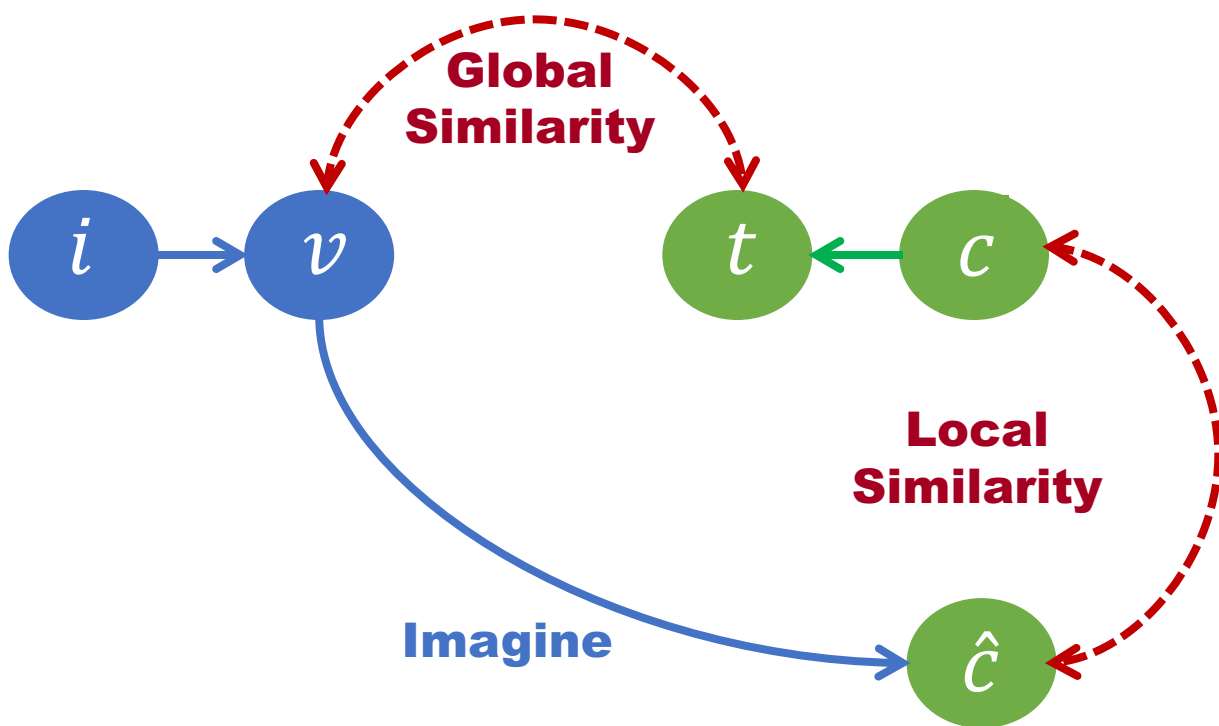
Classical Pipeline



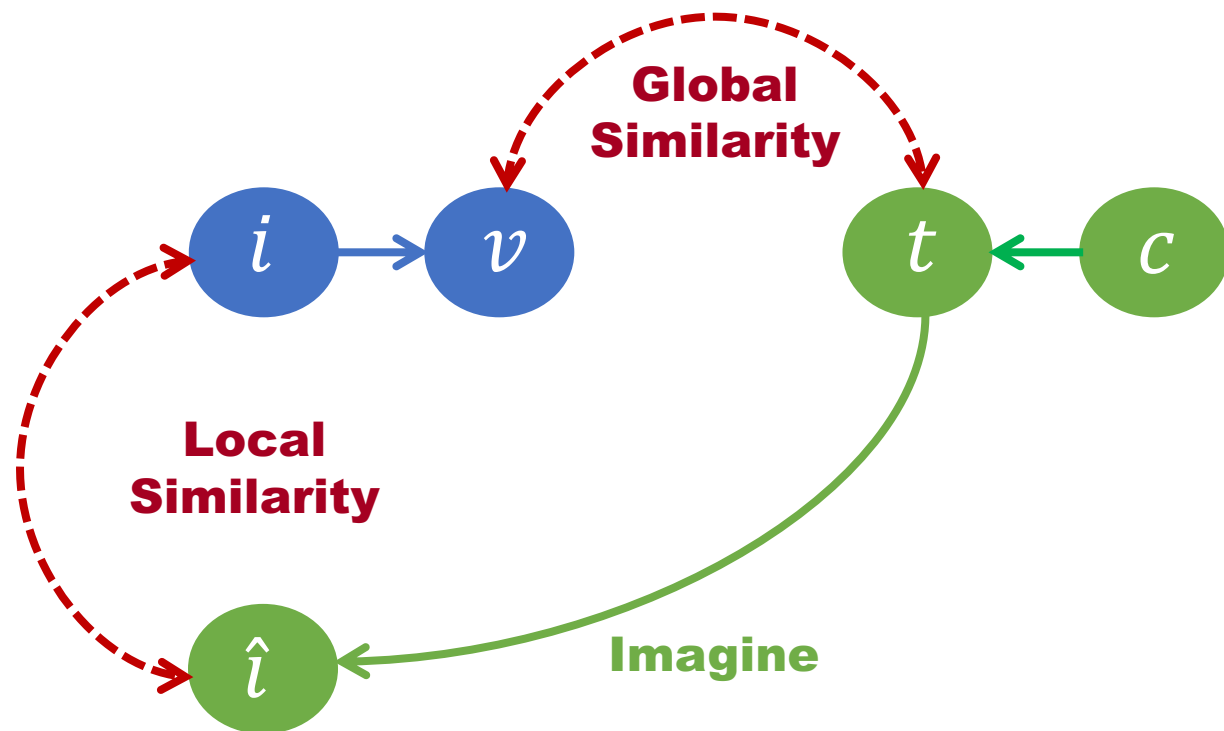
Motivation:

Look \rightarrow Imagine \rightarrow Match

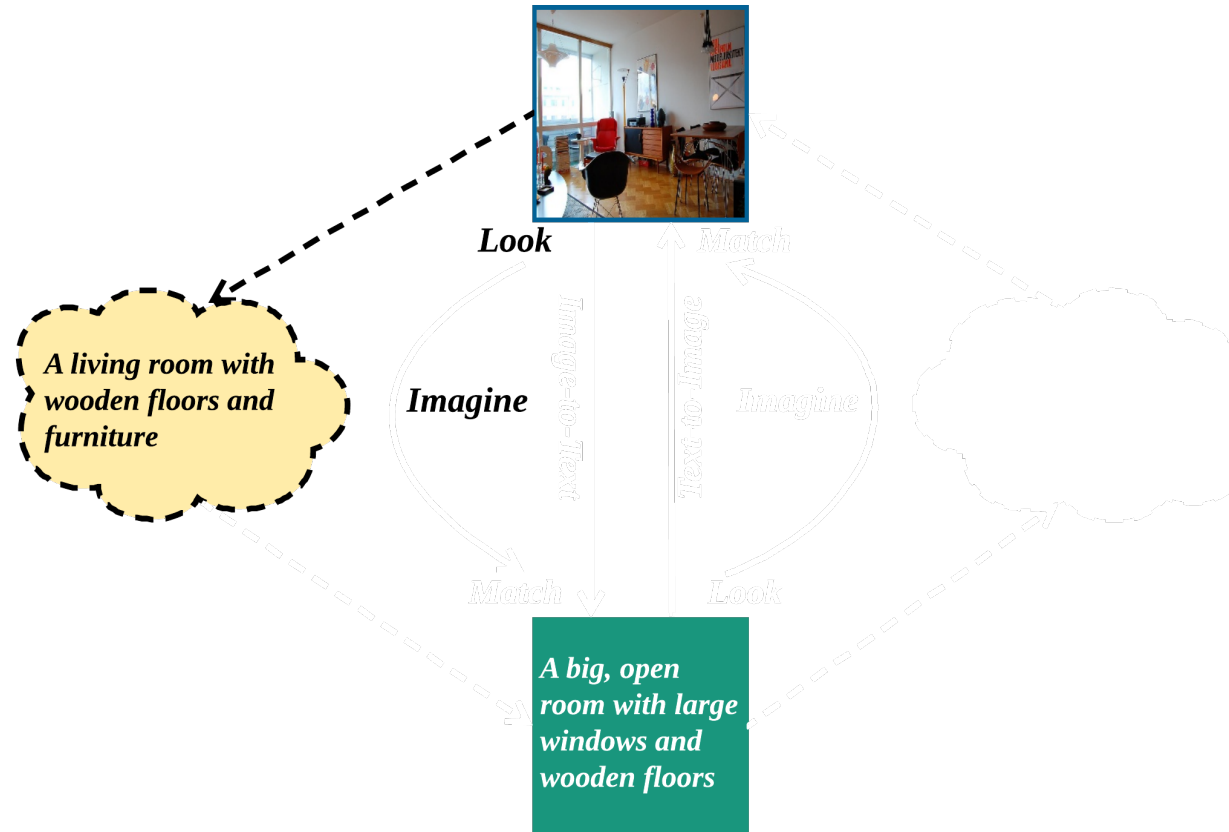
Image-to-Text Retrieval



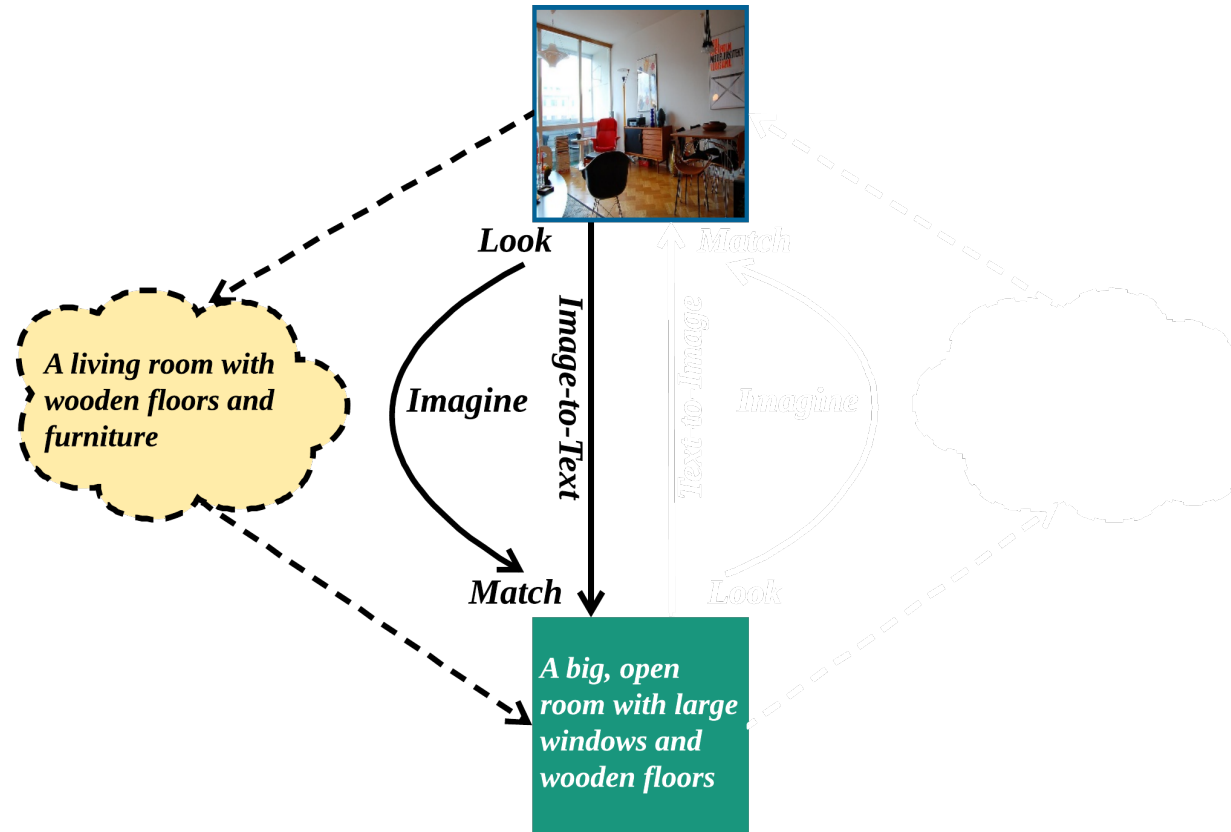
Text-to-Image Retrieval



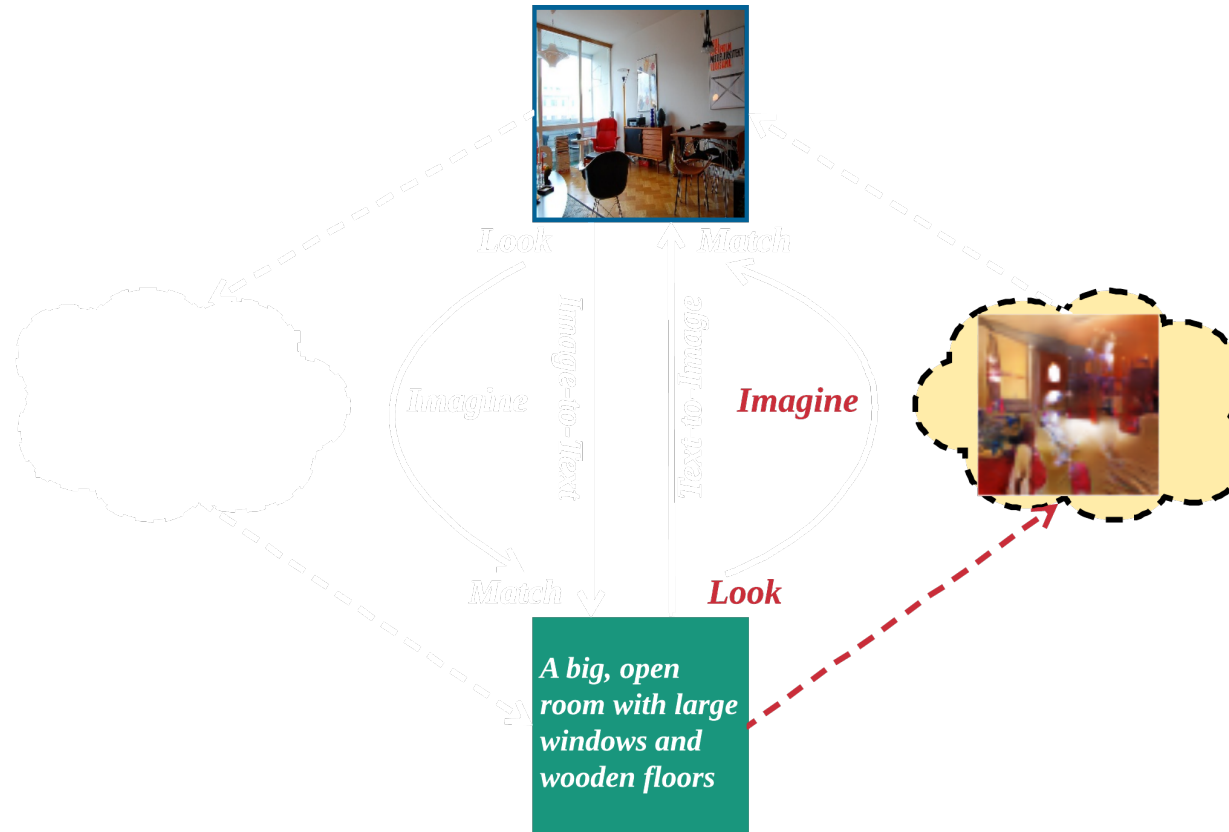
Look → Imagine



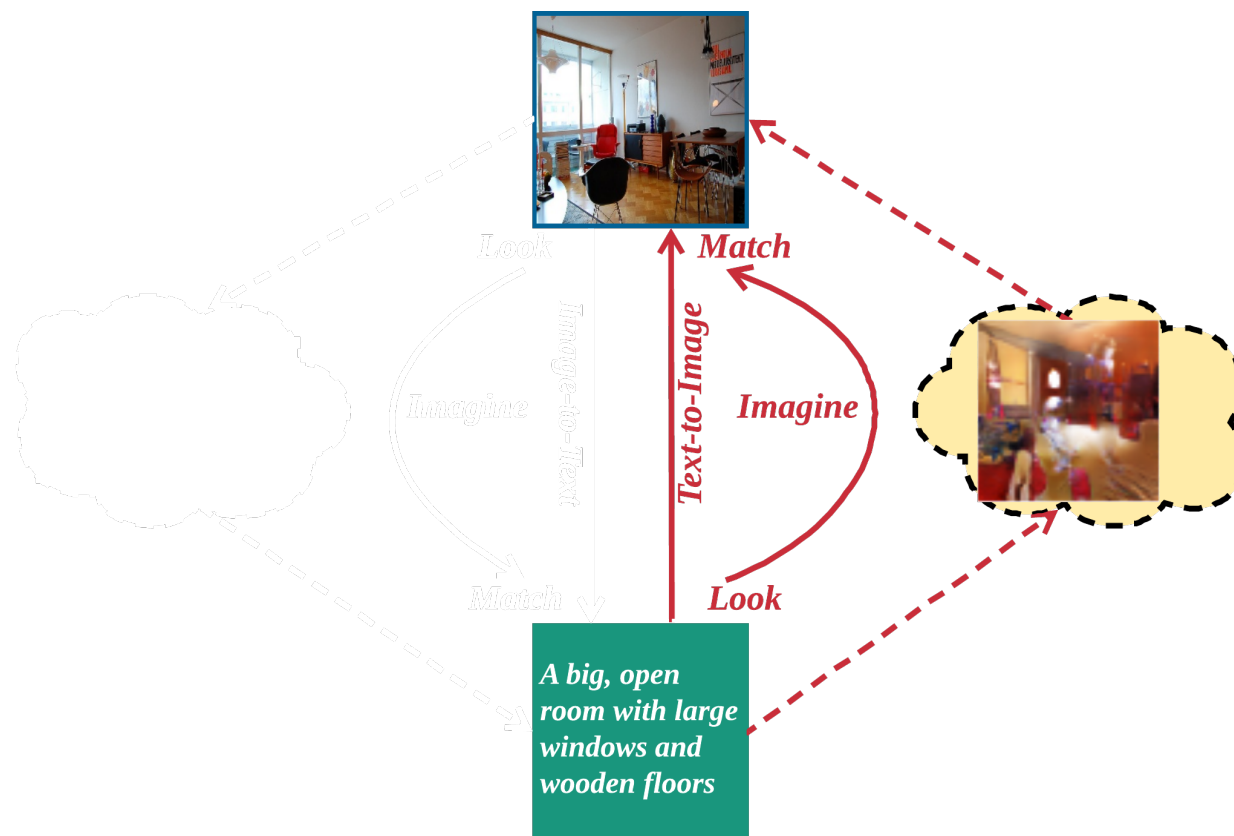
Match



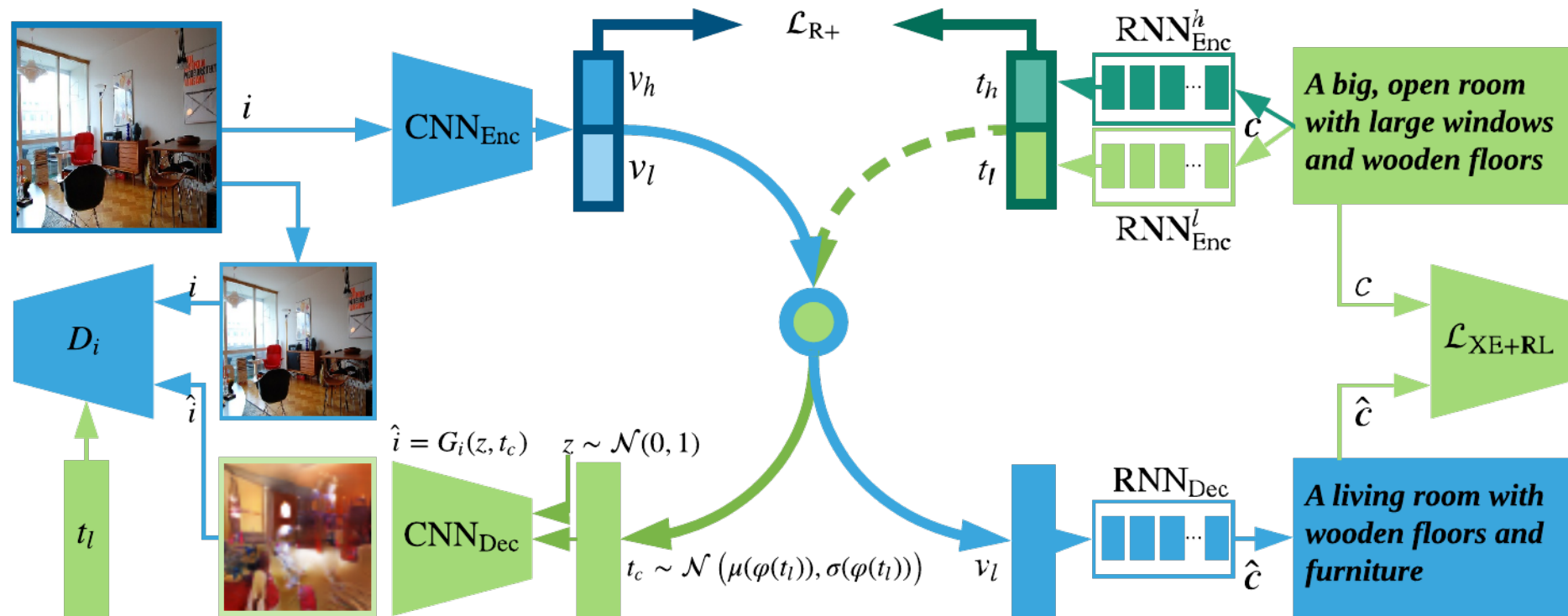
Look → Imagine



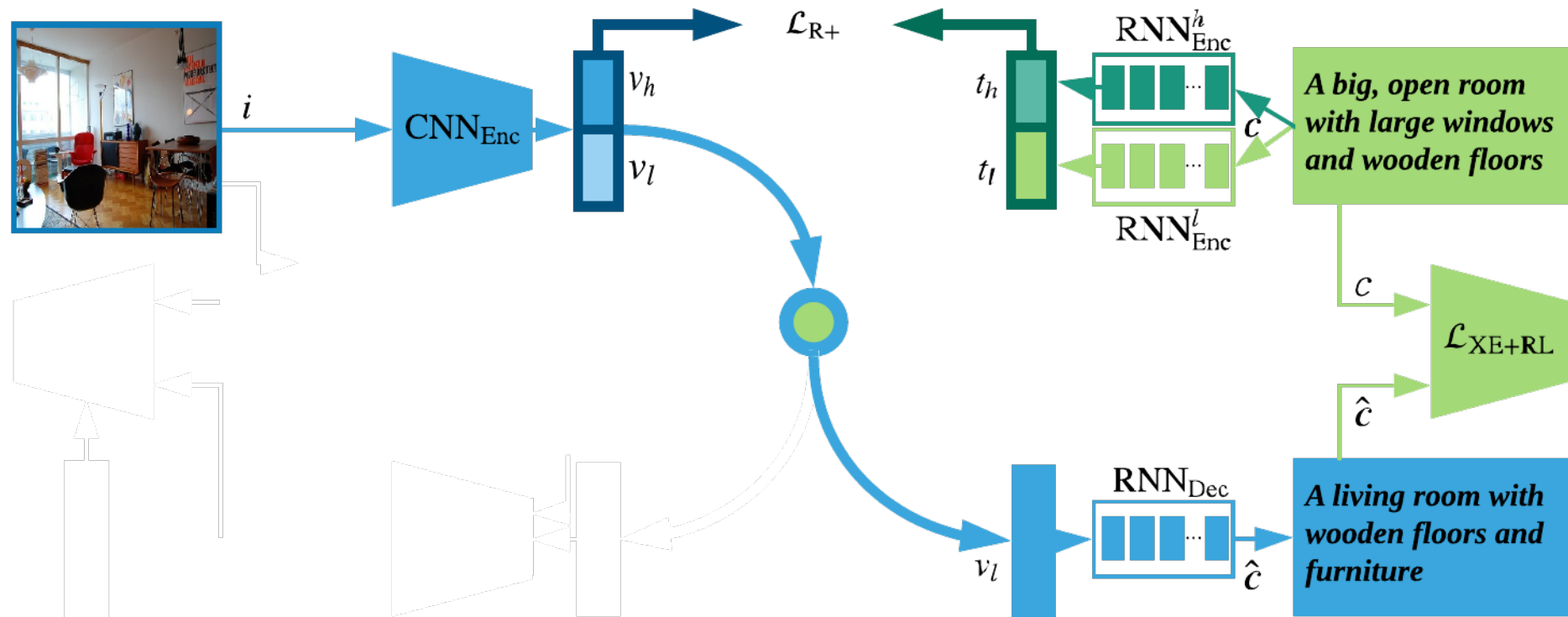
Match



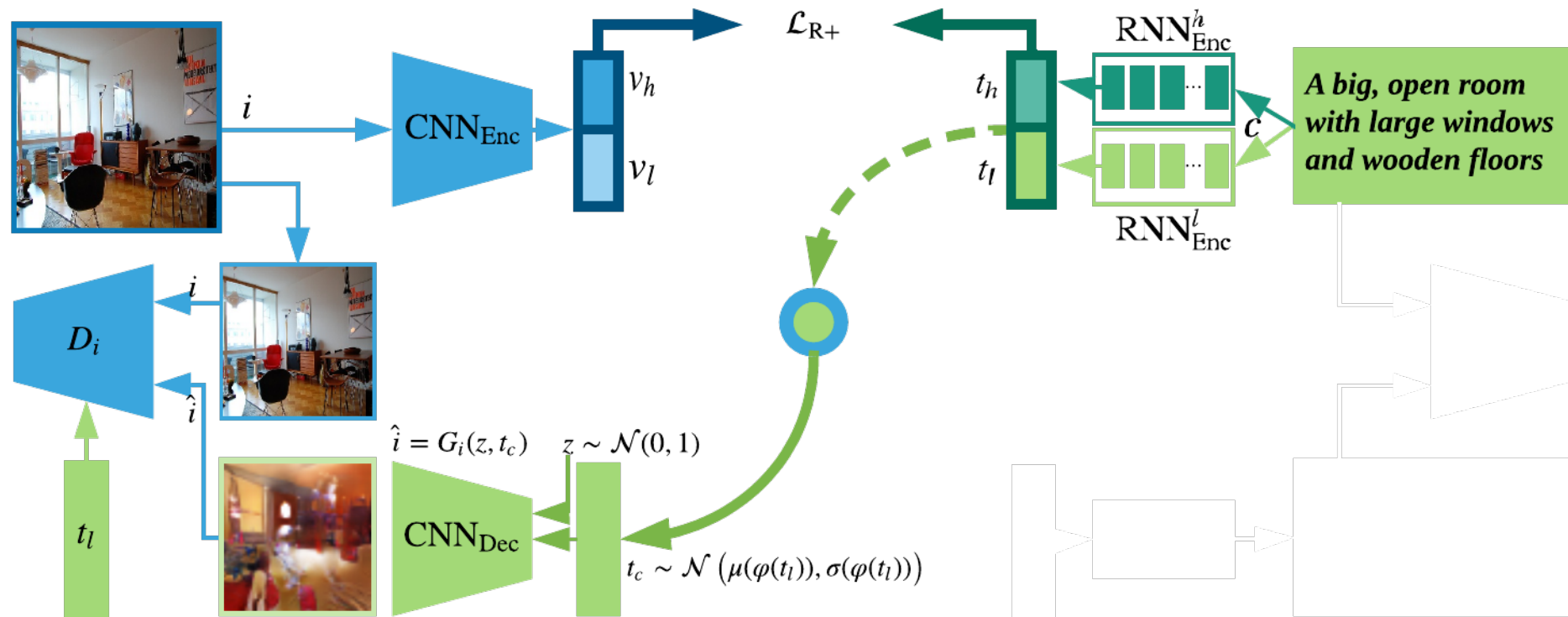
Proposed Approach



Cross-Modal Retrieval with Generative Learning



Cross-Modal Retrieval with Generative Learning

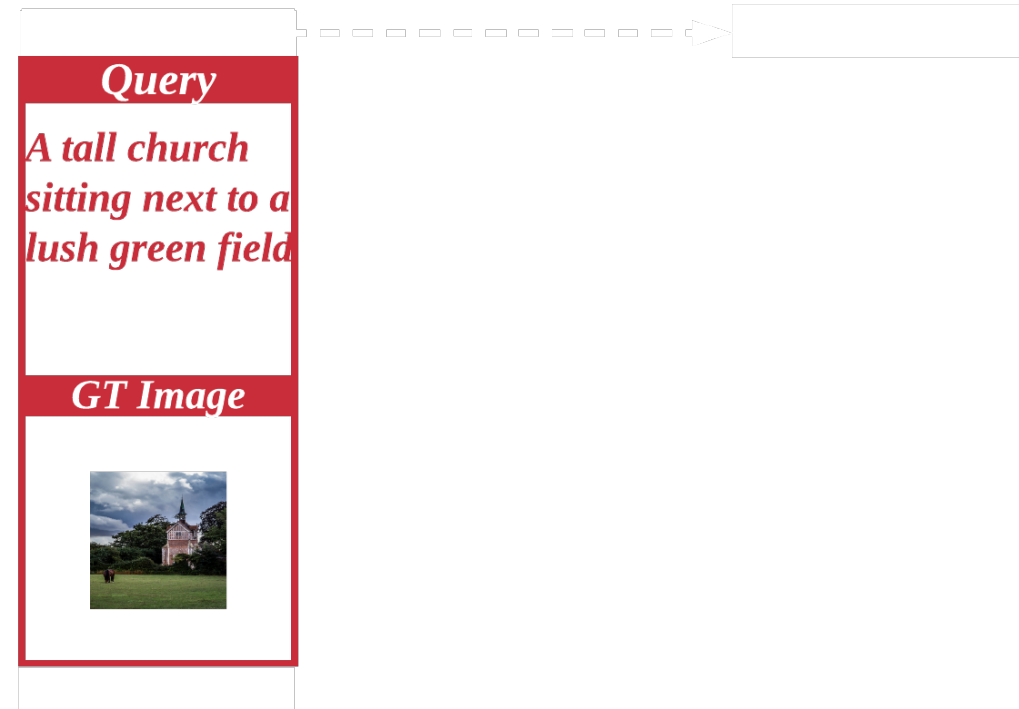


Results

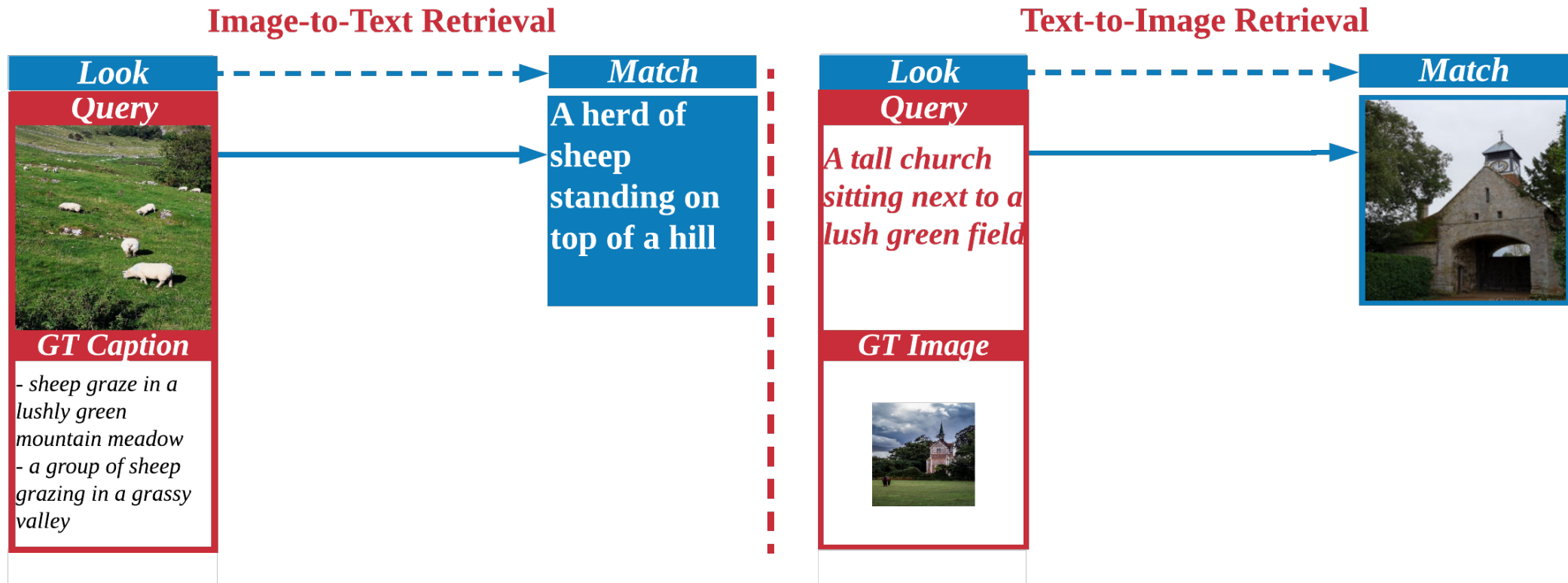
Image-to-Text Retrieval



Text-to-Image Retrieval

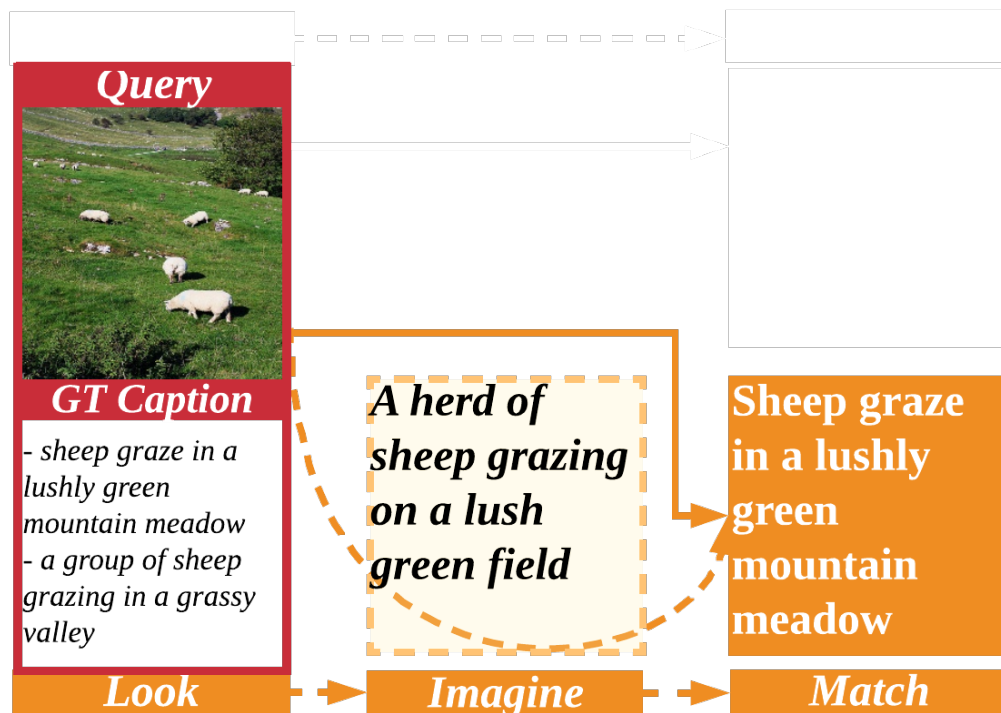


Results (Classical Pipeline)

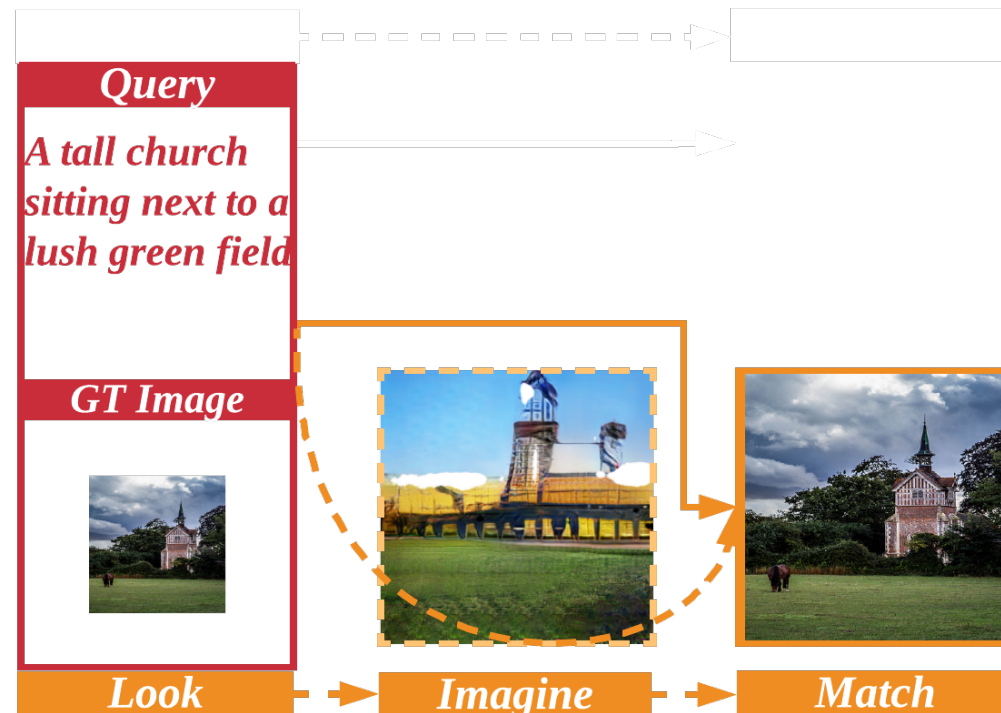


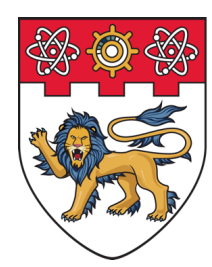
Results (Ours)

Image-to-Text Retrieval



Text-to-Image Retrieval





**NANYANG
TECHNOLOGICAL
UNIVERSITY**
SINGAPORE

RIOSE
Rapid-Rich Object Search Lab
博云搜索实验室


Alibaba Group
阿里巴巴集团


2018
CVPR
SALT LAKE CITY • JUNE 18-22

At the Poster:

- **Additional details**
- **Quantitative results**
- **Discussion**