

如何使用ViLT做视觉问答

by 唐云龙

- 连接远程服务器 11911607@10.20.1.155

密码: XXXXXXXXXX

(建议使用vscode)

```
(base) 11911607@nb1516:~$
```

- 进入虚拟环境project_21fall

```
conda activate project_21fall
```

```
(base) 11911607@nb1516:~$ conda activate project_21fall
```

- 进入文件夹ViLT

```
(project_21fall) 11911607@nb1516:~$ cd ViLT
```

- 执行以下指令:

```
python demo_vqa.py with num_gpus=1 load_path="vilt_vqa.ckpt" test_only=True
```

```
(project_21fall) 11911607@nb1516:~/ViLT$ python demo_vqa.py with num_gpus=1 load_path="vilt_vqa.ckpt" test_only=True
```

- 然后会有以下输出, 点击右下角Open in browser

```
(project_21fall) 11911607@nb1516:~/ViLT$ python demo_vqa.py with num_gpus=1 load_path="vilt_vqa.ckpt" test_only=True
WARNING - ViLT - No observers have been added to this run
INFO - ViLT - Running command 'main'
INFO - ViLT - Started
IMPORTANT: You are using gradio version 1.6.4, however version 2.4.2 is available, please upgrade.
-----
Running on local URL: http://0.0.0.0:8888/
To create a public link, set `share=True` in `launch()`.
Interface loading below...
<IPython.lib.display.IFrame at 0x7f38d1b09ca0>
```

📘 Your application running on port 8888 is available. [See all forwarded ports](#) ⚙️ ×

Open in Browser

Preview in Editor

就能进到这里面：

URL OF AN IMAGE.

QUESTION

IMAGE

ANSWER

CLEAR

SUBMIT

SCREENSHOT

GIF

FLAG

Examples

Run All

Load Previous
CTRL + ←

Load Next
CTRL + →

≡

■ ■

URL OF AN IMAGE.	QUESTION
https://s3.geograph....	What is the color of...
https://computing.ec...	What is the mustache...
https://computing.ec...	How many slices of p...
https://computing.ec...	Does it appear to be...



- 分别在 URL OF AN IMAGE 和 QUESTION 里输入一张图片的链接和针对该图片的提问，点击 SUBMIT，右边 ANSWER 里就会出答案（左下角是一些例子）


URL OF AN IMAGE.

https://computing.ece.vt.edu/~harsh/visualAttention/ProjectWebpage/Figures/vqa_2.png

QUESTION

How many slices of pizza are there?

IMAGE



ANSWER

8

Latency: 2.53s

CLEAR

SUBMIT

SCREENSHOT

GIF

FLAG

Examples

Run All

Load Previous
CTRL + ←

Load Next
CTRL + →

≡

■ ■

URL OF AN IMAGE.

QUESTION


URL OF AN IMAGE.

https://computing.ece.vt.edu/~harsh/visualAttention/ProjectWebpage/Figures/vqa_1.png

QUESTION

What is the mustache made of?

IMAGE



ANSWER

banana

Latency: 1.91s

CLEAR

SUBMIT

SCREENSHOT

GIF

FLAG

Examples

Run All

Load Previous
CTRL + ←

Load Next
CTRL + →

≡

■ ■

URL OF AN IMAGE.

QUESTION

<https://s3.geograph....>

What is the color of...

<https://computing.ec...>

What is the mustache...

接下来要做的就是，我们自己找一些图片，自己设计一些问题，然后让这个模型来回答，记录实验结果

目前已经有的点子：

1. 图灵测试：一个同学负责给出图片、提问题，两个同学要给出答案，两个同学其中一个的答案是自己给出的，另一个是采用模型输出的答案，然后让提问者试图区分人和机器的答案。
2. 歧义图片回答

