Dataset visualization



Reference caption: a plane on the tarmac

```
In [20]: # Load and show generated captions
# todo update the path to match your experiment
pred_captions_file = "outputs/eval_captioning/2023_05_09_00_02_51/pred_captions.
with open(pred_captions_file, "r", encoding="utf-8") as f:
    pred_captions = f.readlines()
assert len(pred_captions) == len(dataset)

def show_caption(n):
    print(f"Generated caption: {pred_captions[n]}")
In [21]: for i in range(10):
    show_datapoint(i)
```

show_caption(i)



Reference caption: a plane on the tarmac Generated caption: a plane and some people loading their luggage openers possible flight



Reference caption: two white and red trains parked next to each other Generated caption: a old train in the middle of some tracks meandscreen



Reference caption: a boat docked in a waterway next to a house Generated caption: boats on the water near houses - - this is a beautiful picture



Reference caption: a black and white photo of a train at a station Generated caption: a train at a train station at night establishment, monoch



Reference caption: a group of cyclists racing on a track Generated caption: cyclists competing at night on a track meander isbn 9780



Reference caption: two sheep laying in the grass

Generated caption: two sheep on a grassy field the herders kept keeping it cool



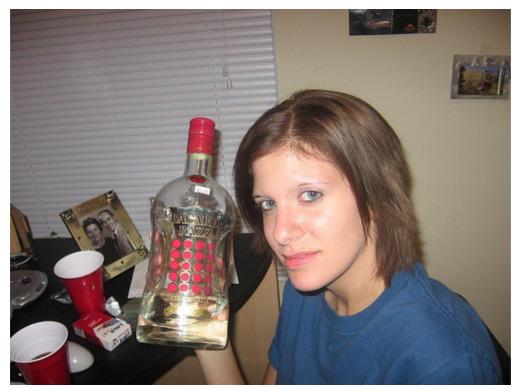
Reference caption: a computer is sitting on a desk next to a monitor Generated caption: a dell system of some kind - photo 2 an office desk



Reference caption: a man and a woman posing for a picture Generated caption: me and my friend so much - i wish i did his hair



Reference caption: a horse standing next to a green slide Generated caption: a child playing with a pony celestial mini slide with a pony



Reference caption: a woman holding up a bottle of vodka Generated caption: me holding a bottle of vodka with an empty bottle next to it

In []: