

Supplemental Material

1. More generated humorous captions on unseen images

As demonstrated in Figure 1, the humour generators that we have trained are capable of producing humorous captions for a diverse range of unseen images from OxfordTVG-HIC. However, the perceived humour intensity of the audience may differ depending on various factors such as culture, language, and personal experience, since humour is a complex cognitive phenomenon that requires high-level and subjective understanding.

Since humour generation is such an abstract and complicated task, our humour generators sometimes fail to induce the perception of humour even while the captions are somehow related to the input image as shown in Figure 3.

In addition to OxfordTVG-HIC, it is our intention to evaluate the efficacy of our trained humour generators on alternative image captioning datasets. As demonstrated in Figure 2, the humour generators are capable of generating humorous captions for images sourced from COCO [1]. From this, we can deduce that the humour generators exhibit both robustness and generalizability in the humour-oriented captioning task.

References

- [1] Xinlei Chen, Hao Fang, Tsung-Yi Lin, Ramakrishna Vedantam, Saurabh Gupta, Piotr Dollar, and C. Lawrence Zitnick. Microsoft coco captions: Data collection and evaluation server, 2015.



I hate shaving because it hurts so much.

The smell of my fart stinks more

I'm sitting here and playing with a nice guy called a stone

Second entrance to four-dimensional space.

Go back to school!



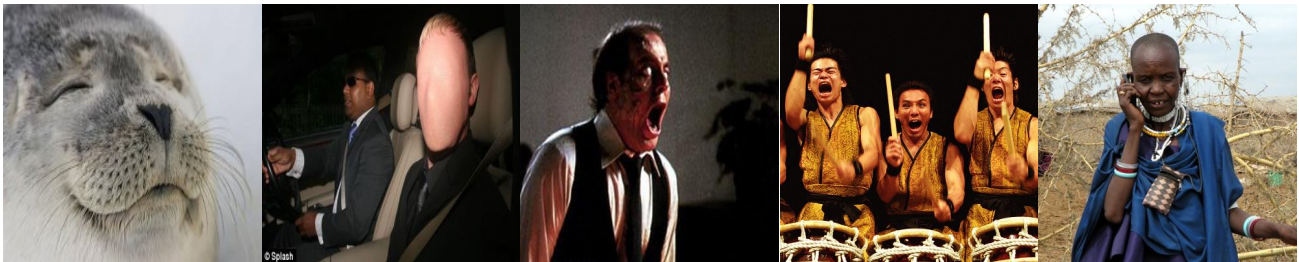
The expression on the face of the next patient.

I was up in the air, but I didn't fly.

I'm the head of the Strawberries team.

Your eyes are red... are you angry?

"You're fighting for the world."



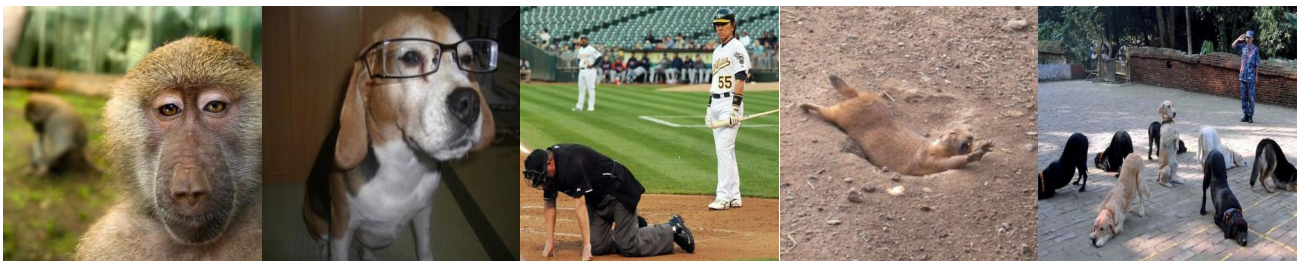
Awkward moment when you leave a store without buying

Watch my eyes.

When your boyfriend knows you are pregnant.

The front row of the roller coaster.

What are you talking about? What are YOU talking ABOUT?



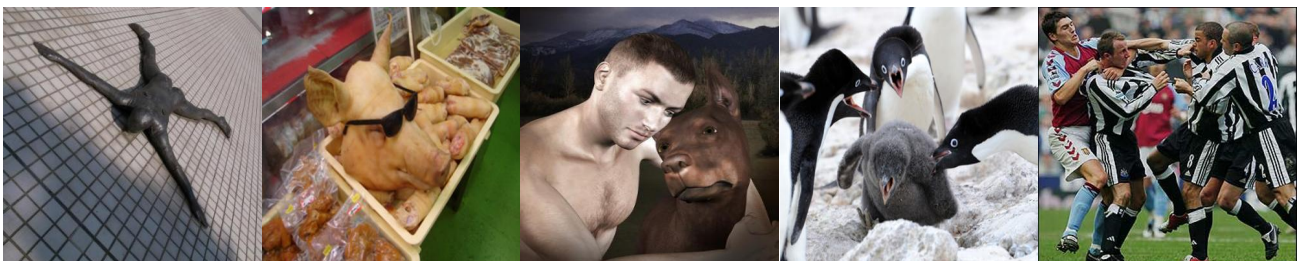
My husband opens the window and looks at his son in the park

I'm a dog, right?

The scene of a wild monkey being dragged away by a turtle.

I jumped into the pool, and there was no water.

"Hey, guys, let's get out of here!"



I've been frogs for a long time.

You've never been a pig before, have you?

I remember when your husband was a cow.

My girlfriend when I refused to confess I was surrounded by girls.

You're mine!

Figure 1: **Humorous captions generated on unseen images in the OxfordTVG-HIC dataset.** The humour generators have been able to generate funny captions given a wide variety of unseen images including paintings, animation, scenes of nature, animals, and human expression. Thus, the humour generators exhibit promising generalisability.

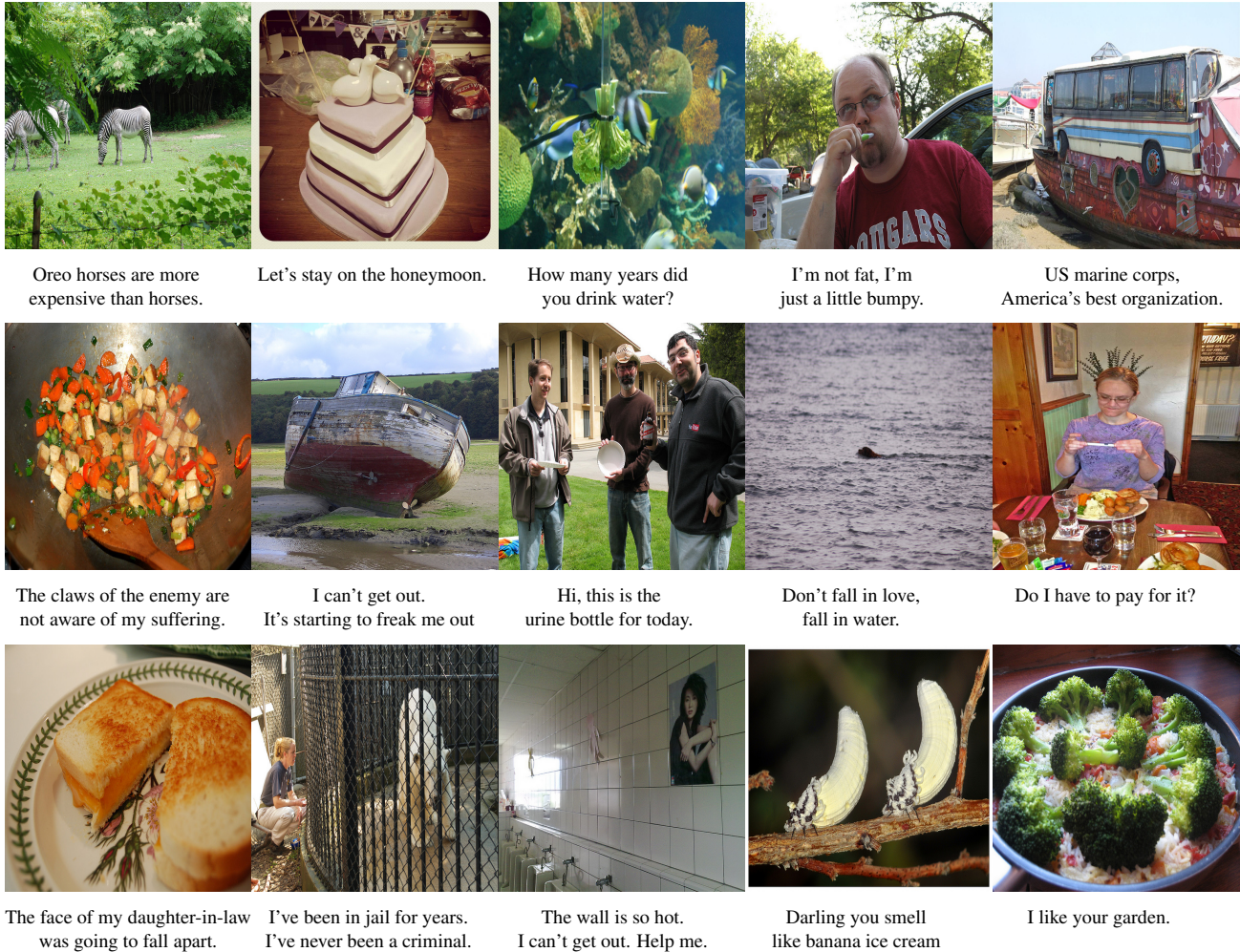


Figure 2: **Humorous captions generated on unseen images in COCO [1]**. The humour generators manage to generate funny captions of unseen images from other object-centric datasets such as COCO [1] as well, which further demonstrates the generalisability of our trained humour generators and the diversity of OxfordTVG-HIC in an implicit manner.

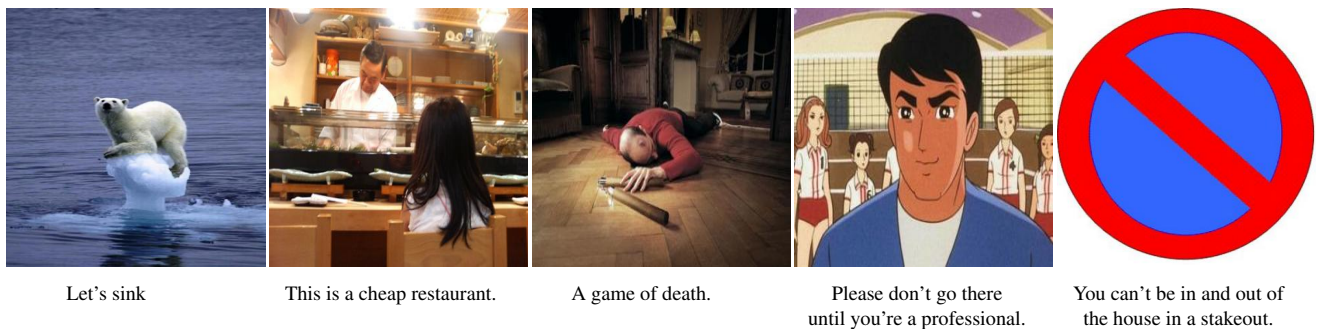


Figure 3: **Failed captions generated on unseen images in OxfordTVG-HIC**. While the generations of our humour generators are novel and an initial attempt towards a new direction: humour-oriented captioning, they still require significant improvement to match the sophistication and nuance of their human-made counterparts. Here, we illustrate how our neural speakers fail to elicit the perception of humour even if the generated captions are somewhat relevant to the input images.