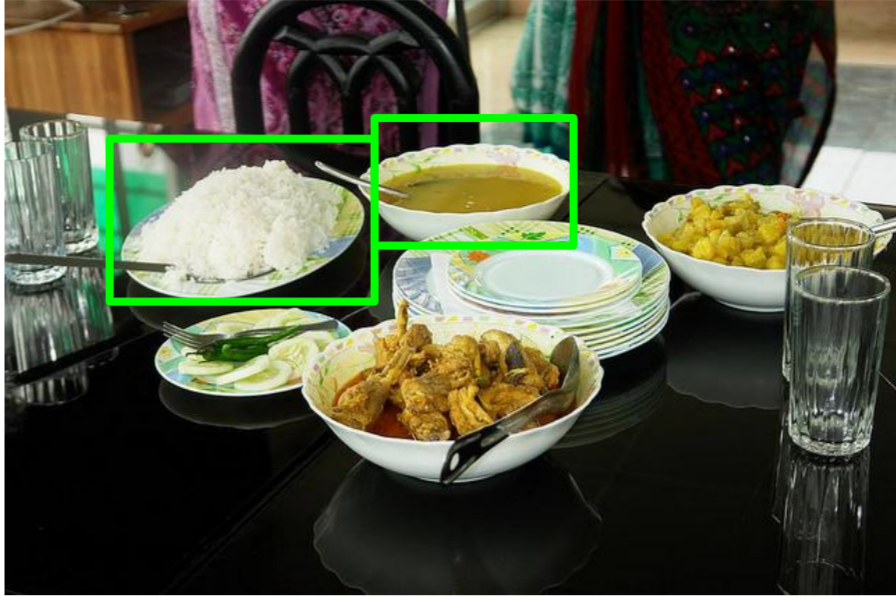


Referring Expression Comprehension



Book with three teddy bears on the cover.

Generalised Referring Expression Comprehension



Rice bowl and the yellow curry bowl

Grounded Visual Question Answering

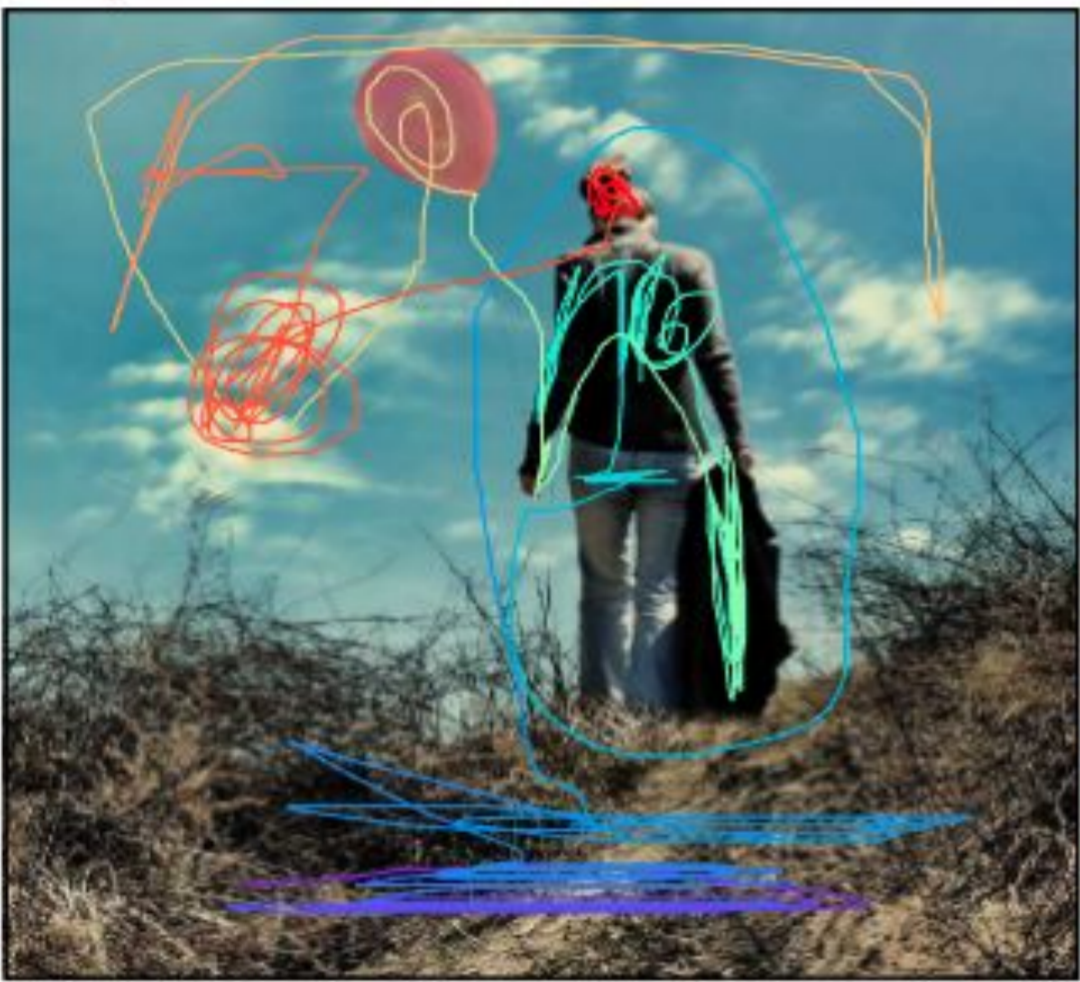


Which clock is a heart?
A, B, C, or D?



How many tablets in this box?

Grounded Captioning



In the front portion of the picture we can see a dried grass area with dried twigs. There is a woman standing wearing light blue jeans and ash colour long sleeve length shirt. This woman is holding a black jacket in her hand. On the other hand she is holding a balloon which is peach in colour. On the top of the picture we see a clear blue sky with clouds. The hair colour of the woman is brownish.



A snowman sits next to a campfire in the snow. He is wearing a hat scarf, and mittens. There are several pots nearby, likely containing a hot meal for the snowman.

GUI Grounding

