

Model Input (Prompt)

Identify the coordinate (row, column) of the **only one different** object:



Hint for readers:
Groundtruth:
(3,3)

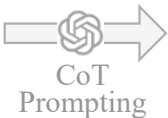


Model Output

ANSWER: (4,5).



Verbal (Concrete) Thinking



Model Output

1. **Observation:** The image consists of a grid of gift boxes with faces, resembling a cute character with a bow tie, but may be small differences.
2. **Labeling the Grid:** ... Row 1: (1,1) to (1,5); Row 2: (2,1) to (2,5); Row 3: ... (1,5).
3. **Features:** Each face has ..., and the facial expression changes slightly between faces.
We need to find the face that differs from the others.
4. **Examining the Grid:** ... It becomes clear that the face in ****row 3, column 5**** is different. It has a **facing different position.**
ANSWER: (3,5).



Different types of Visual Abstract

Contour Style



OR

PhotoSketch



OR

Binarization



...

Visual Abstract Thinking



Model Output

ANSWER: (3,3).

