

# FASHION VISUAL GROUNDING DICTIONARY

## 1. List entity

### - **Attribute:**

- + **Color:** "blue", "green", "black", "white", "yellow", "red", "pink", "purple", "flower", "colored", "polka dot"
- + **Sleeve length:** "short sleeve", "long sleeve"
- + **Length:** "short", "long"
- + **Collar:** "v-neck", "round neck"
- + **(Attribute)? + Accessory:** "black tie", "black bow"

### - **Position:** "left", "right", "middle", "top left", "top right", "top middle", "bottom left", "bottom right", "bottom middle",

### - **Category:** "shirt", "pant", "glass", "hat", "bag", "watch"

### - **Person:** "boy", "girl", "man", "woman"

## 2. List pattern generated from entity

### - **(Person)? 's + (Position)? + (Attribute)? + Category**

*Note: (Person)? = 'Person' or '' (just like regex)*

Example:

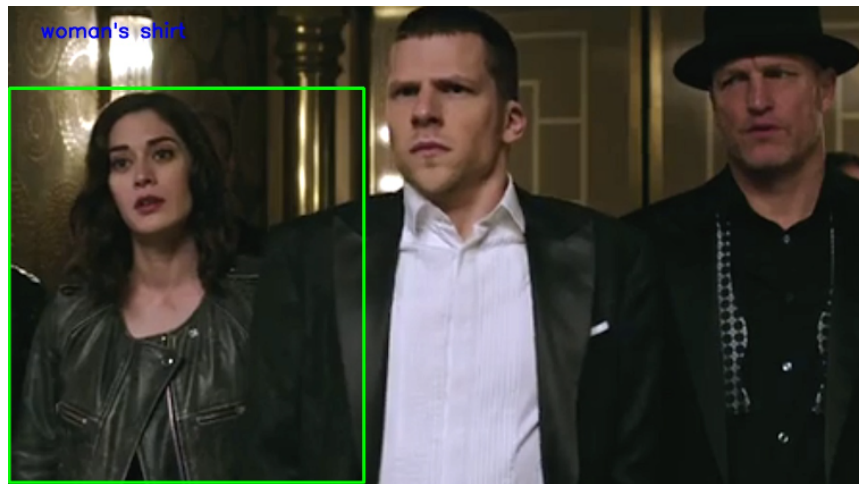
- + black glass



- + top right shirt



- + woman's shirt



+ man's hat



+ black bow shirt



+ girl's black bag



+ long sleeve



+ flower pant



Note: output của Fashion Visual Grounding API tập trung vào chủ thể người đang mặc trang phục đó hơn là chỉ tập trung local vào trang phục đó, đây cũng là input tương thích với input của model Fashion Visual Question Answering.

# FASHION VISUAL GROUNDING API

Link:

[https://drive.google.com/drive/folders/1EZ7zSm4ctinue3Pf1P6J4Ei0\\_8LzFIIX?usp=sharing](https://drive.google.com/drive/folders/1EZ7zSm4ctinue3Pf1P6J4Ei0_8LzFIIX?usp=sharing)

gồm có:

- File collection json để test trên postman (input: img base64 và text, output: start\_point và end\_point của bounding box)
- File python để show demo bounding box lấy từ API

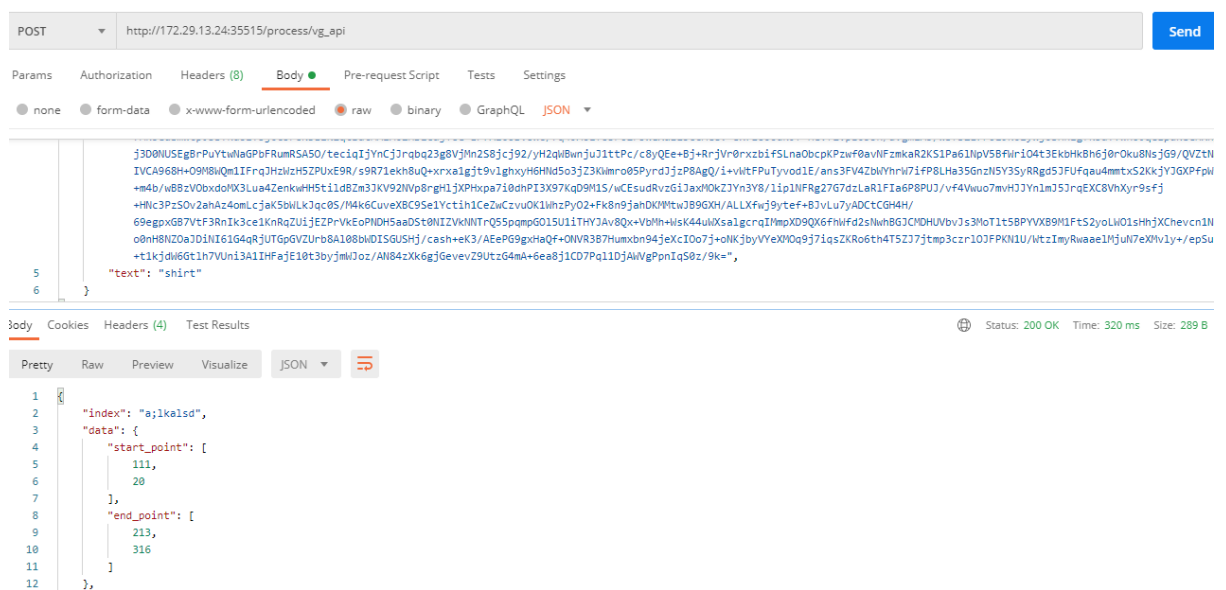
**GPU usage:** 2.045 GB

**Response time:** 0.3s/request

Example:

- API trên postman

Website để lấy base64 img online: <https://www.base64-image.de/>



- Test trên python: (requirement requests, cv2, pillow, numpy)

input: file\_path, text, output\_path

output: file demo.jpg lưu ở outputpath