



# Clothing Decision Benchmark **穿搭寶**

PRESENTED BY TEAM12 短髮容易長髮ㄉㄌ



 感情 · 2/22

## 男友穿搭太醜

隨便舉例一個：情人節穿搭 ... 除了讓我覺得很醜之外其實也會很失落

   103  75  51





# Clothing Decision Benchmark 穿搭寶

AI Dressing Recommendation System



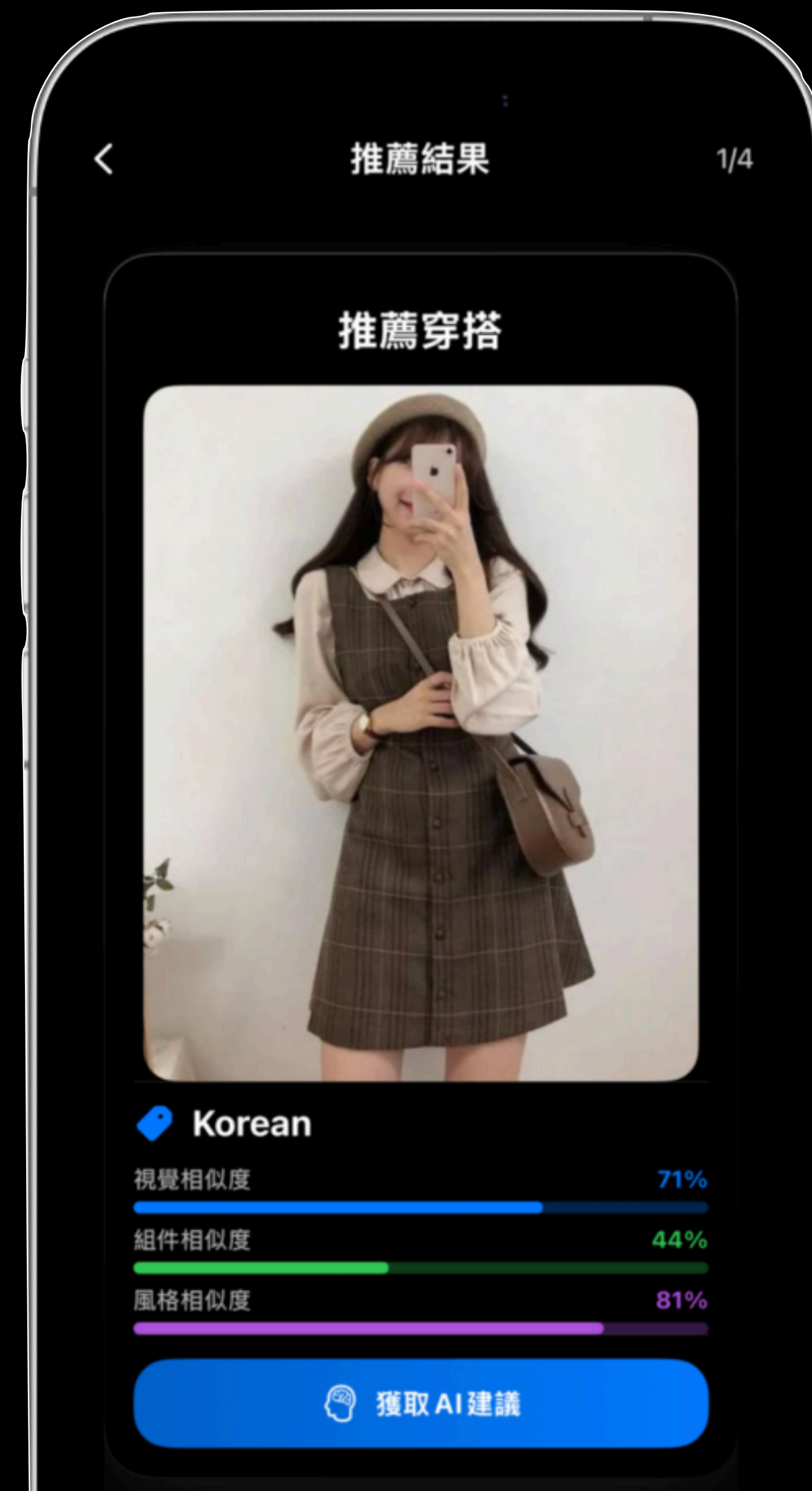
Not sure of your style?

**ML Model** can help.



Don't know how to dress?

**Our Model, your teacher.**



# Build a special you.

Personalized Recommendation System



How about a?

**Live Demo**

**How we did that ?**

# Define the problem

How we did that ?

**Recommend Good outfits  
that users might like**

**Generate Text-based Guide**

# Image Retrieval

Recommend Good outfits



Target

Which one is the most similar picture?

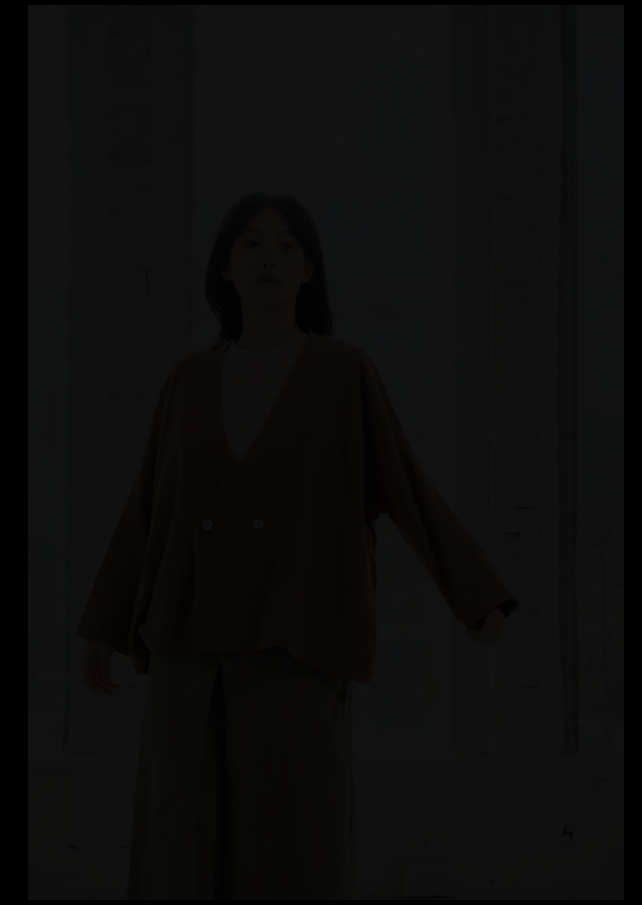
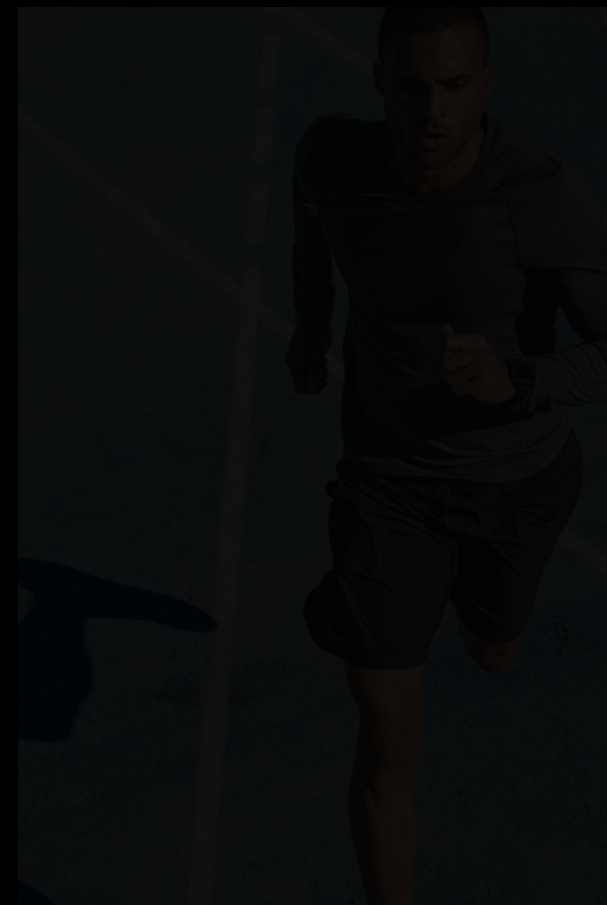


# Image Retrieval

Recommend Good outfits



Which one is the most similar picture?



# Image Retrieval

Recommend Good outfits

By simply tagging clothes?

Good idea ! But **too simple..**



Gray Coat

White Shirt

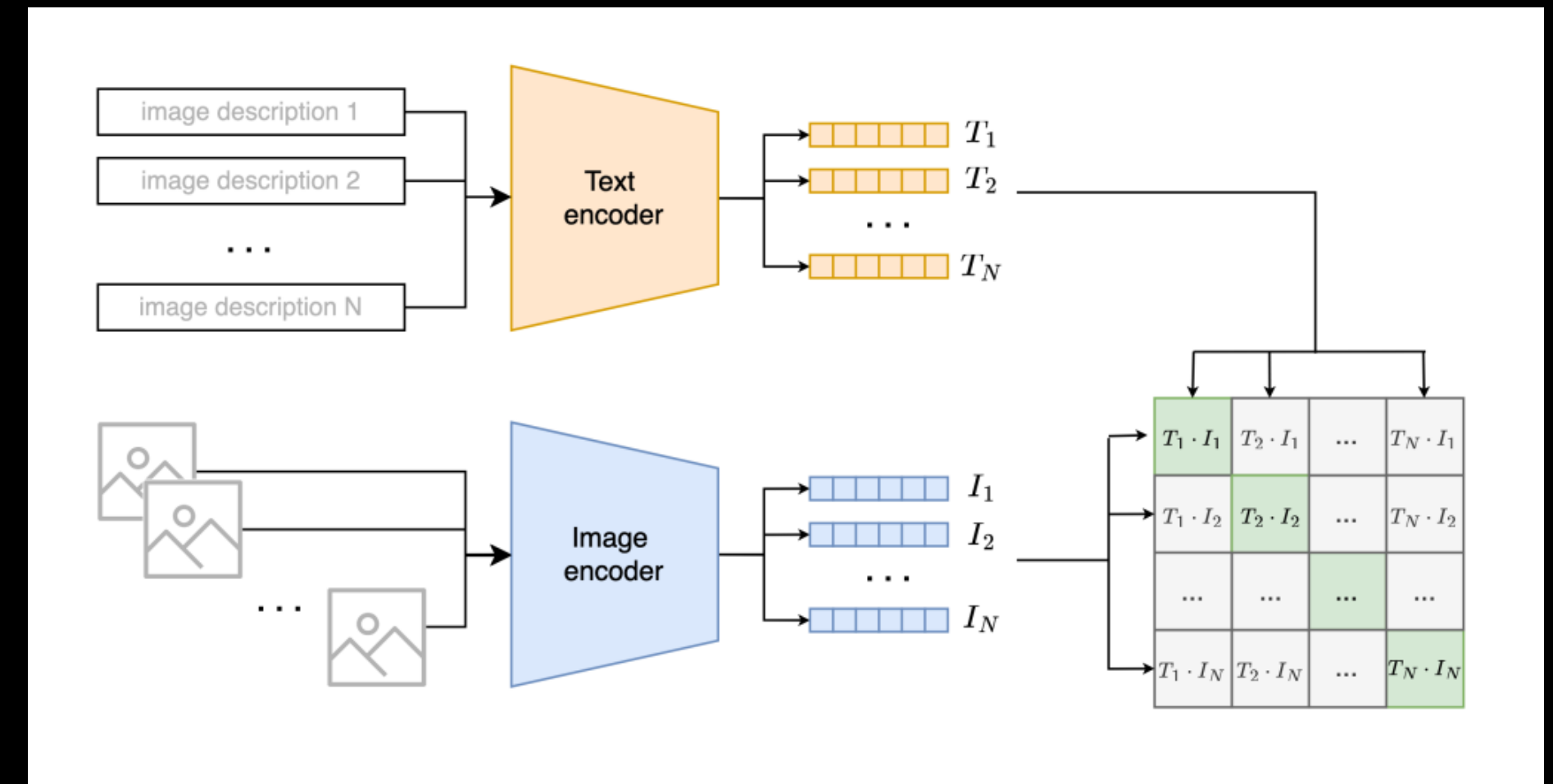
Black Pants

# Image Retrieval

Recommend Good outfits

Our method :

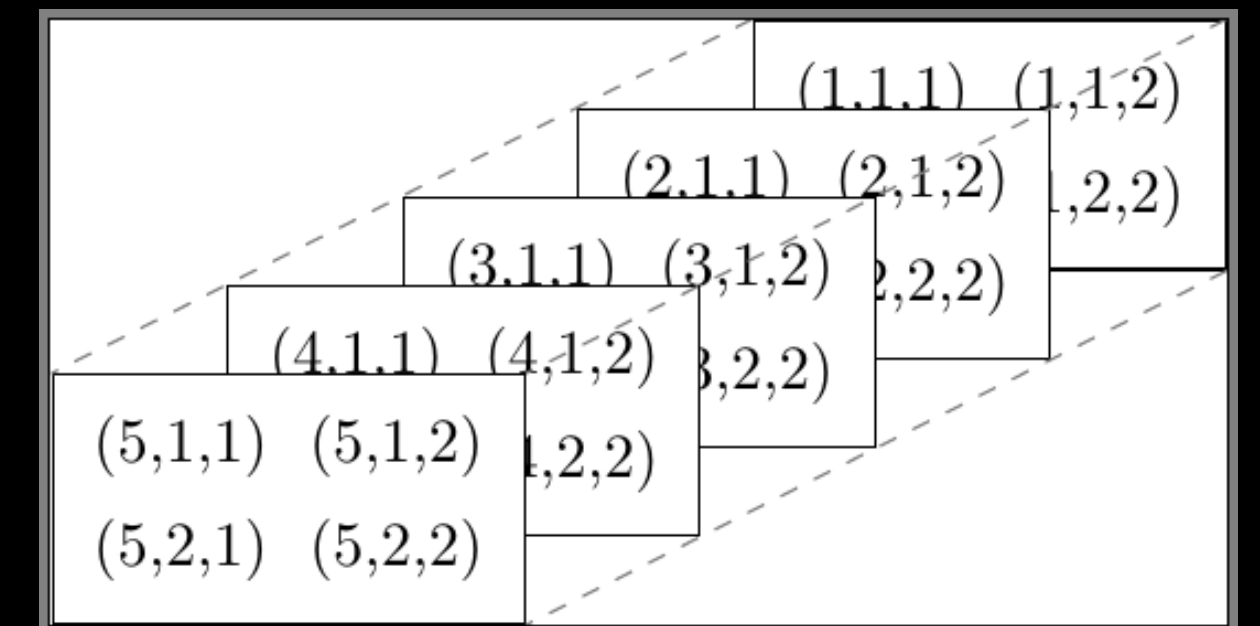
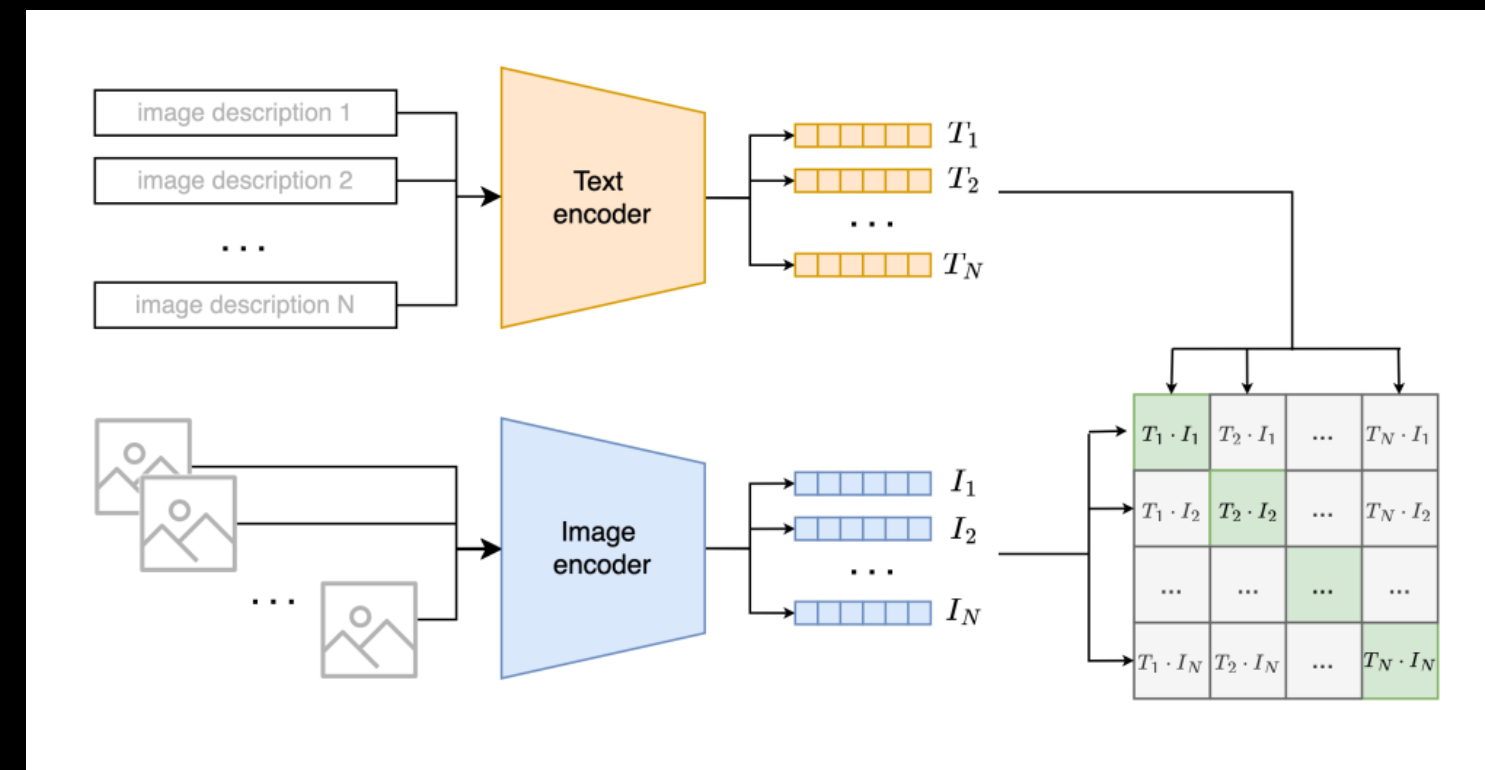
Using **FashionClip**  
To do Feature Extraction



# Image Retrieval

Recommend Good outfits

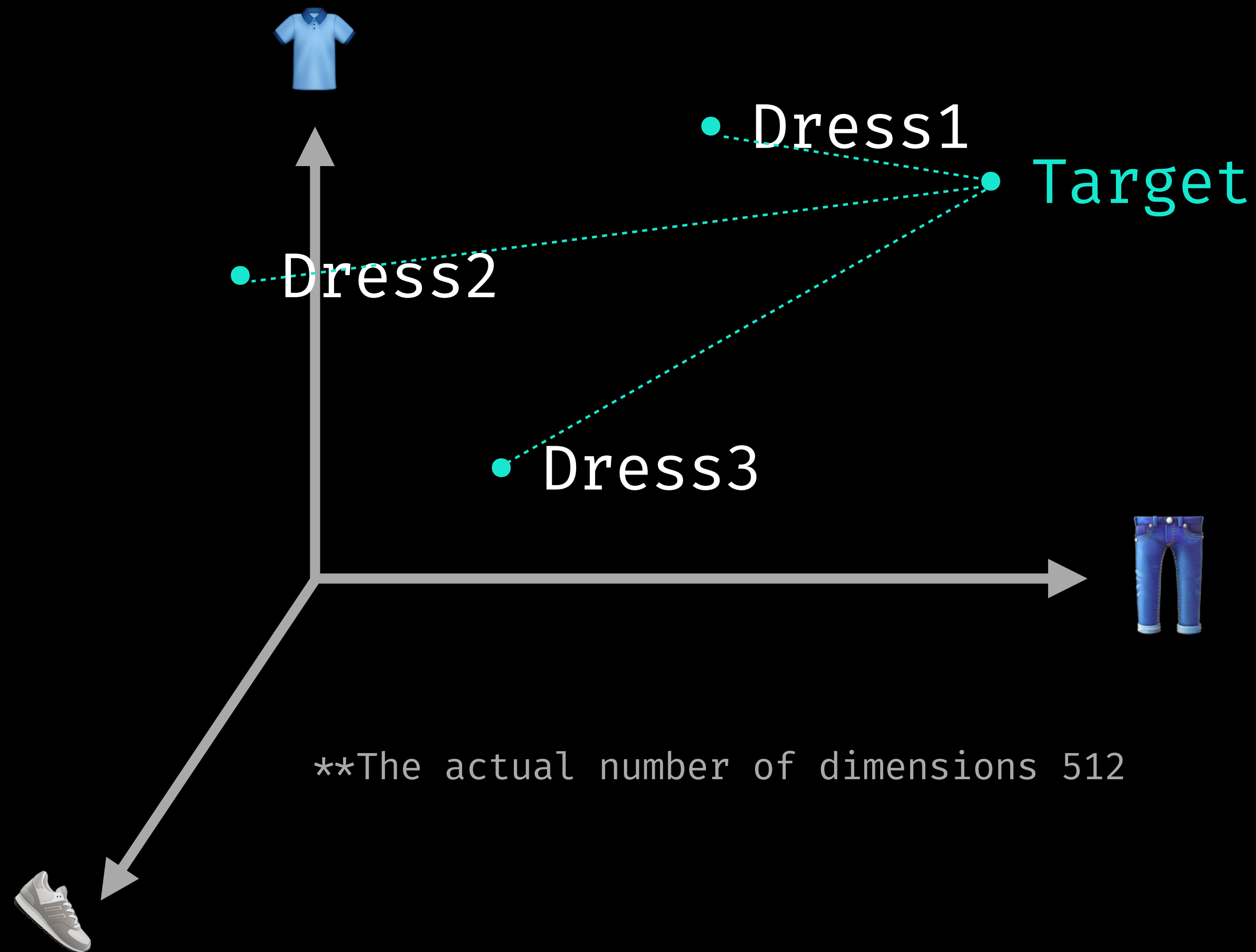
## Using FashionClip To do Feature Extraction



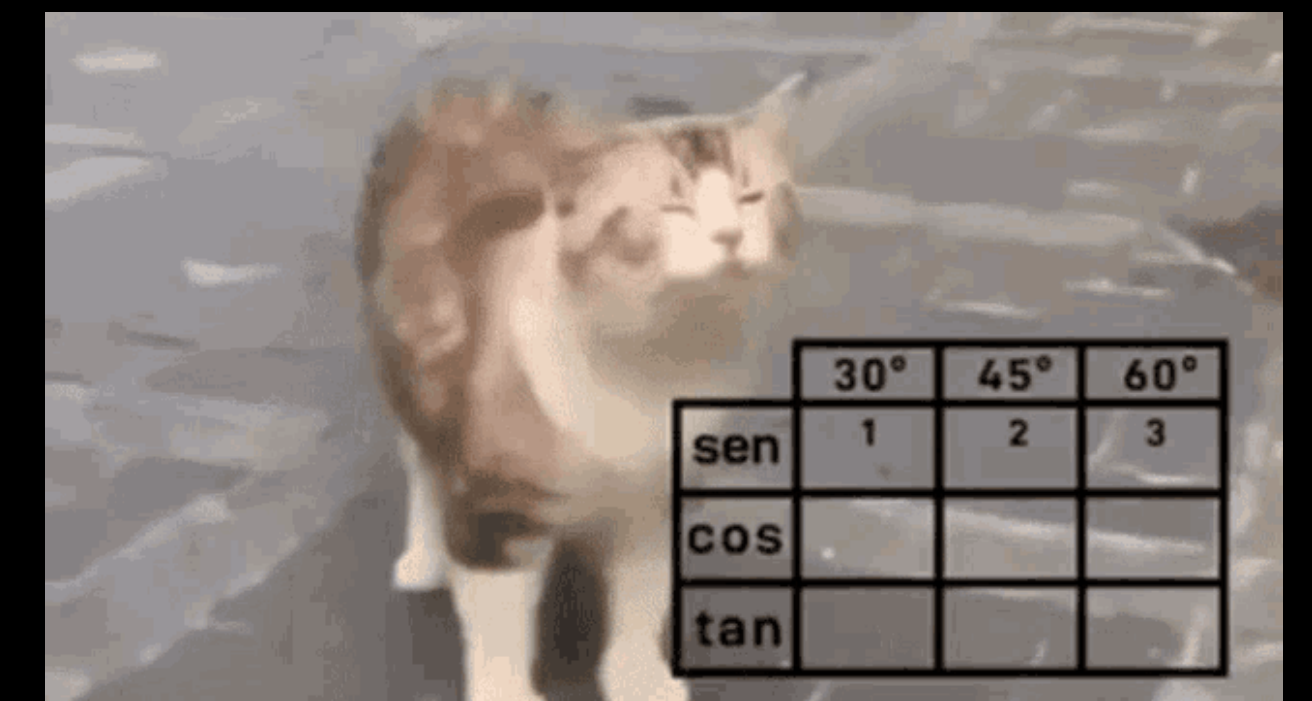
Features of the outfits photo

# Image Retrieval

Recommend Good outfits



**Dress1** is the closest one  
To the target

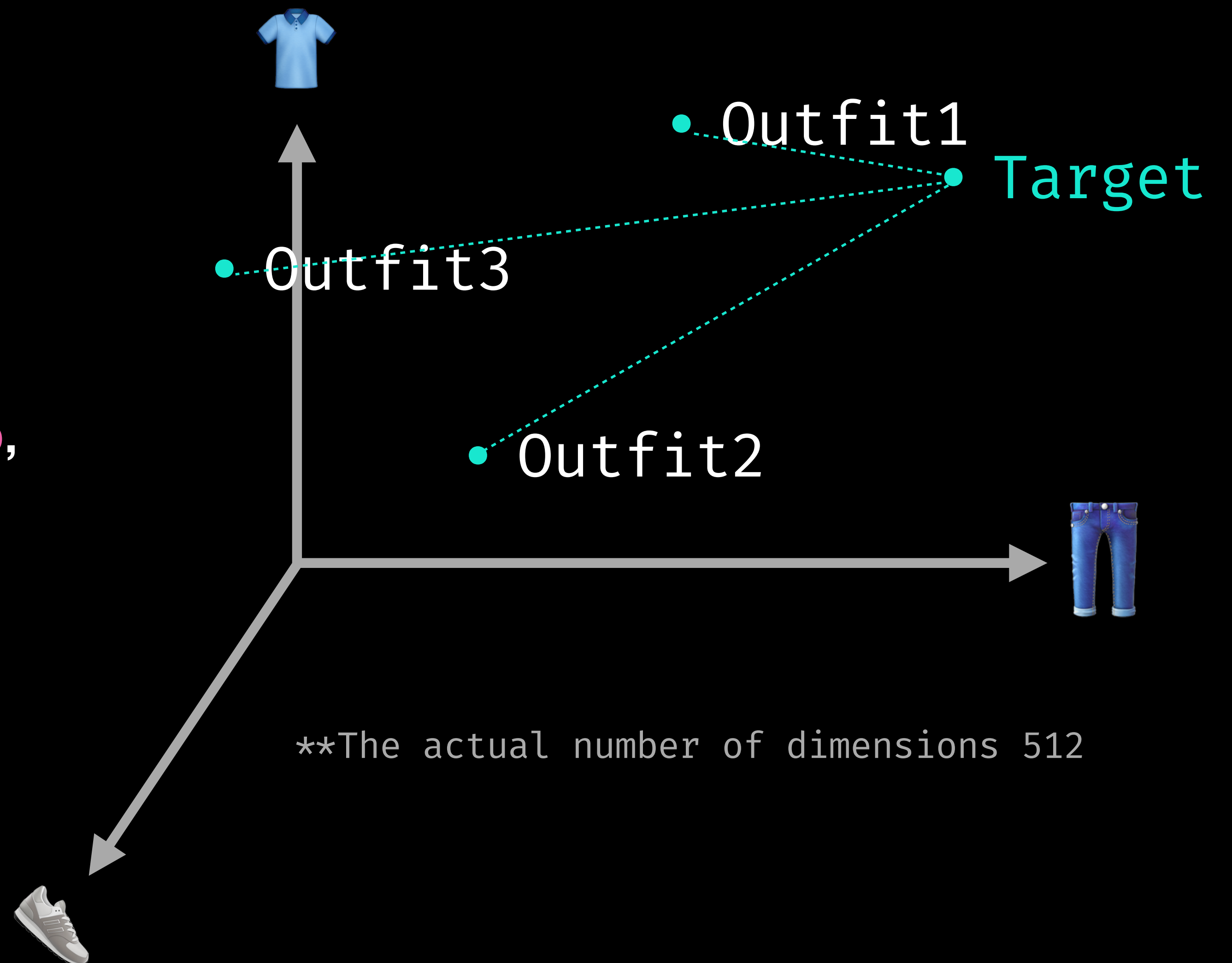


# Image Retrieval

Recommend Good outfits

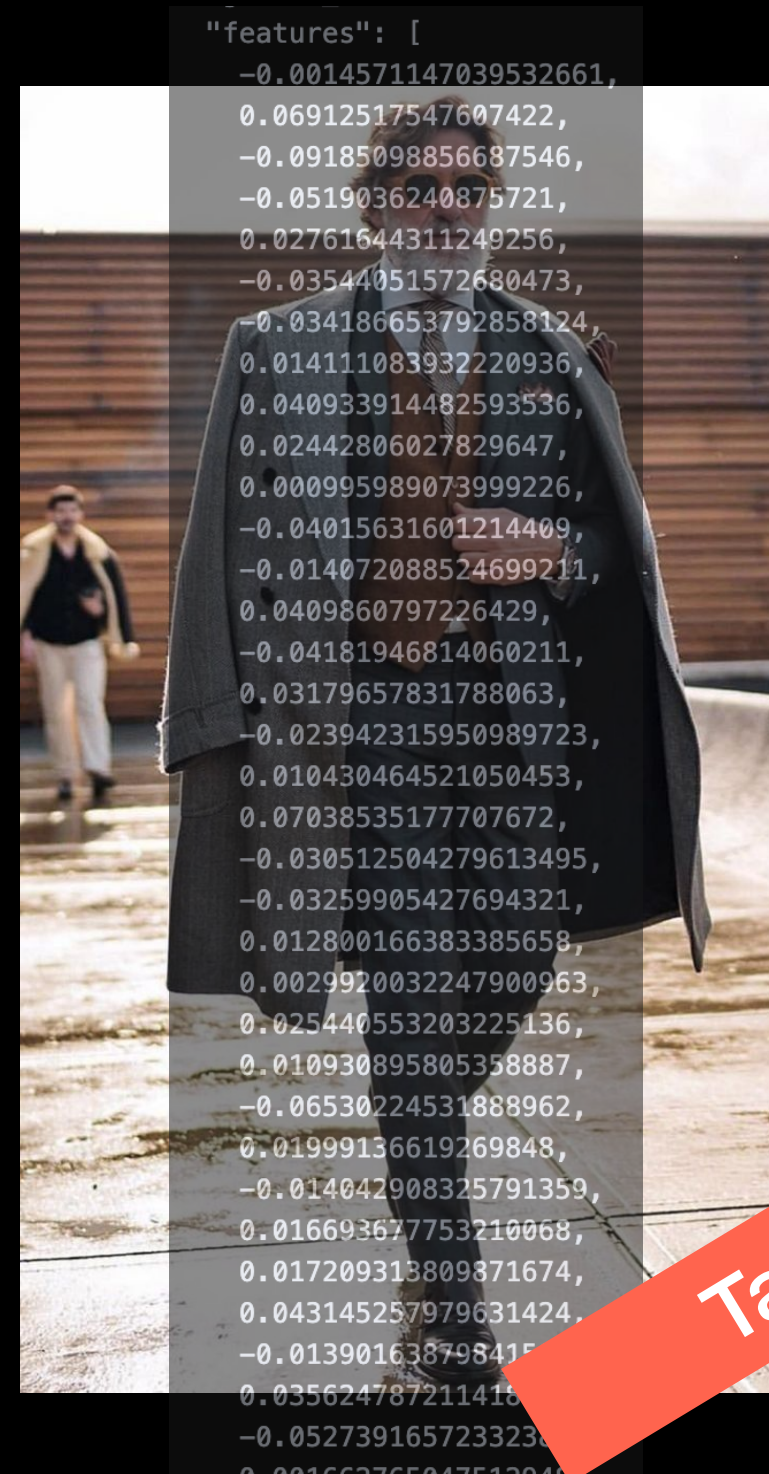
By **comparing the distance** of each dress in the 512-dimensional feature vector that extracted from **FashionClip**, we can know which photo is most similar to the target.

Note:  
We also use some techniques to optimize the performance of FashionClip.  
For example : Freezes FashionCLIP, adds mapping layer (512D→64D) to focus on fashion pairing semantics.



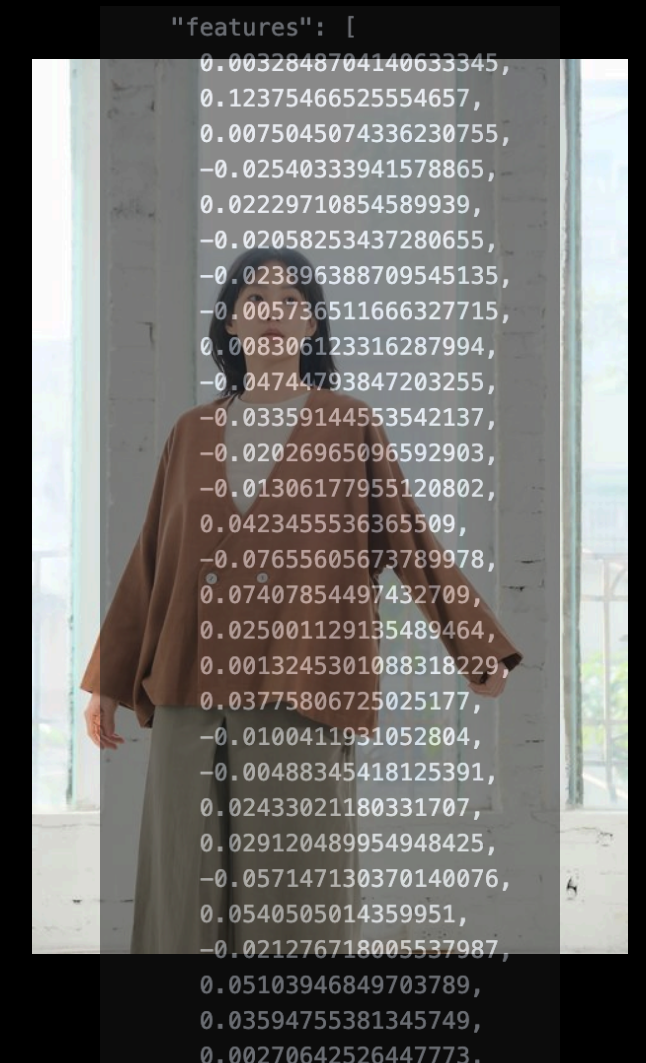
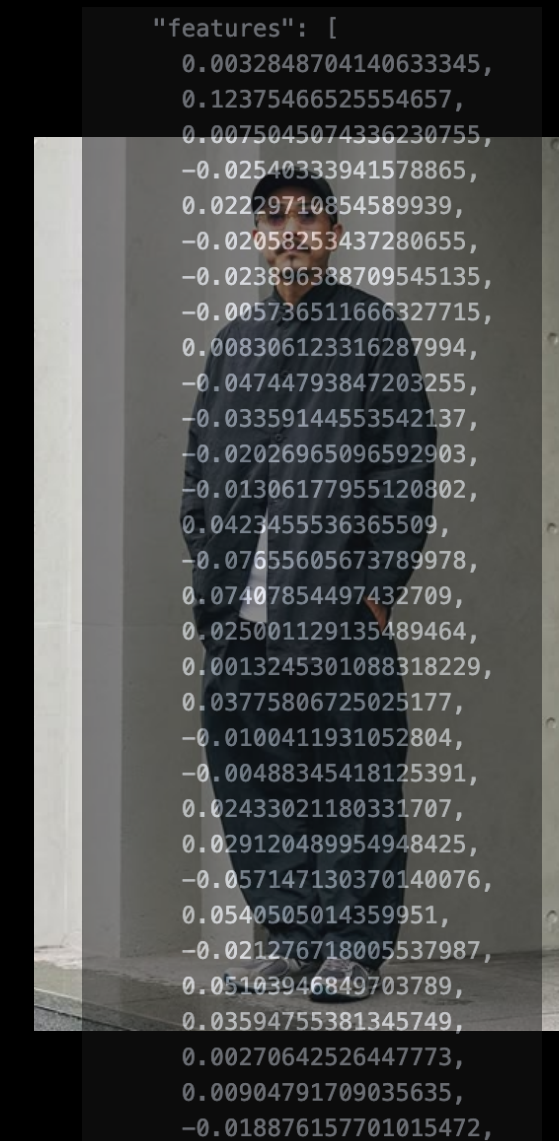
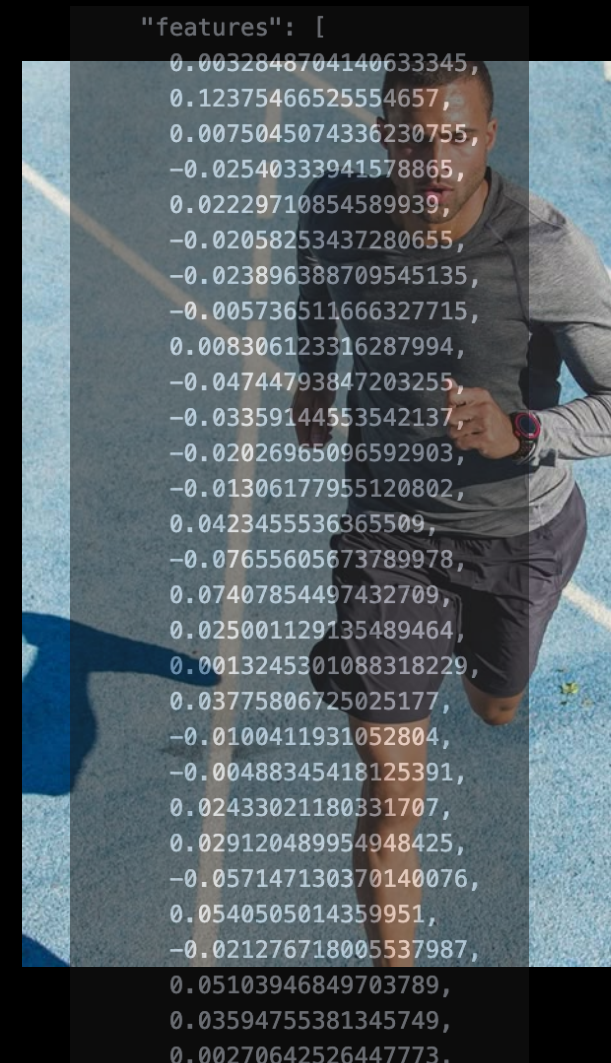
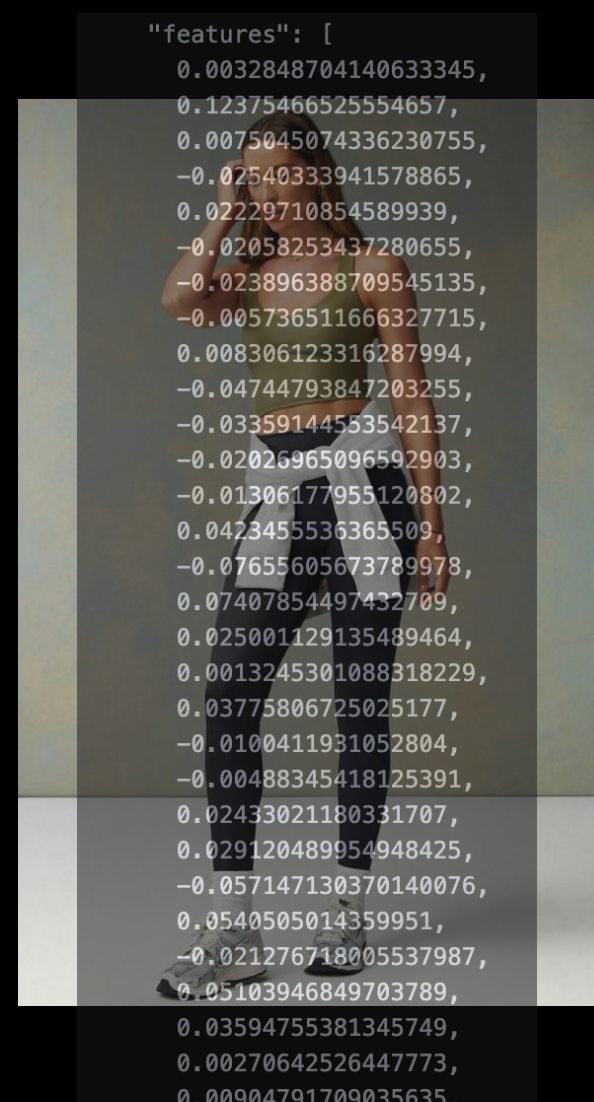
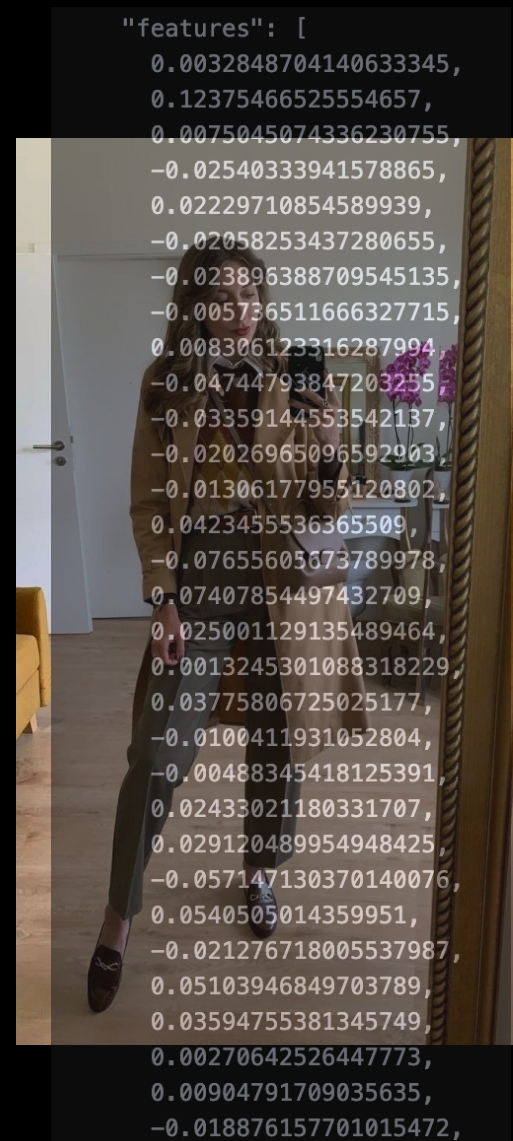
# Image Retrieval

## Recommend Good outfits



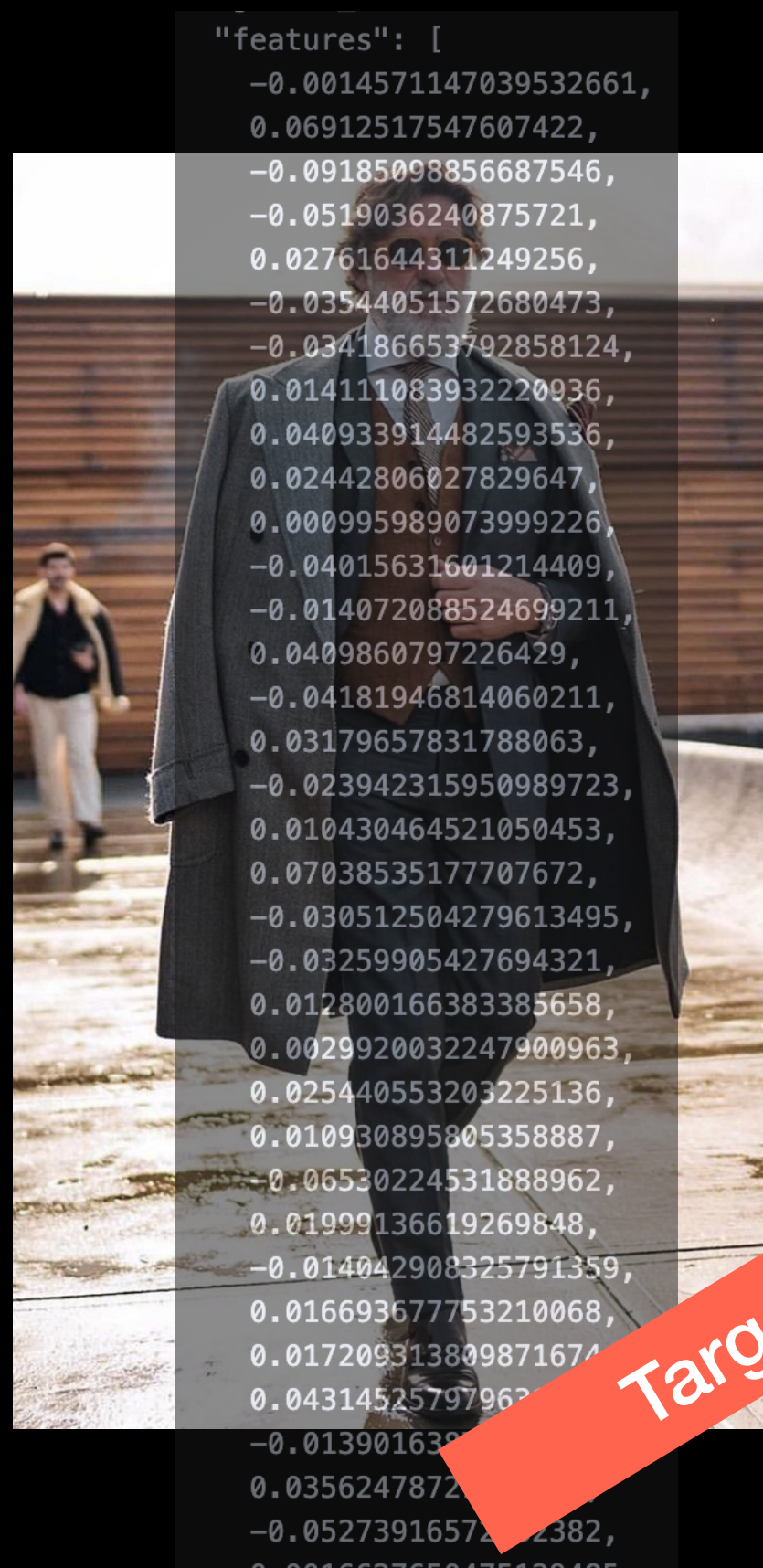
Target

### Which one is the most similar picture?

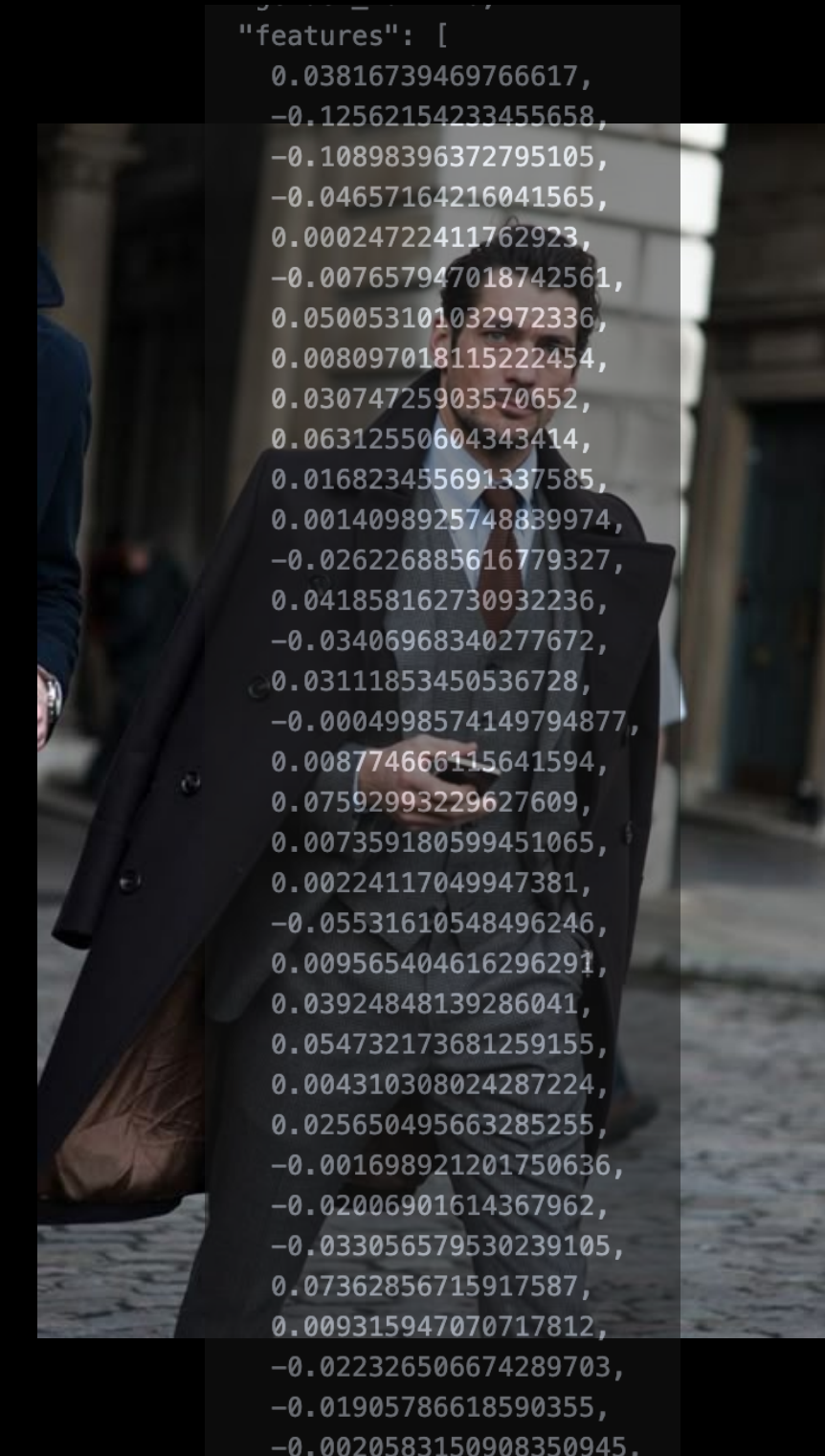


# Image Retrieval

Recommend Good outfits



Closest!



# Image Retrieval

Recommend Good outfits

Use "outfit style" to enhance the stability of recommendations



British



Athletics



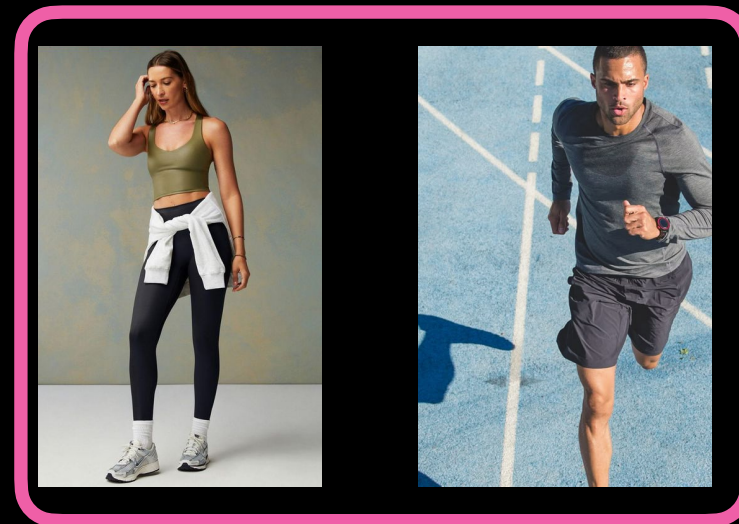
Japanese

# Image Retrieval

Recommend Good outfits



British



Athletics



Japanese



Artsy



Preppy



Casual



Goth



Kawaii



Kawaii



Street



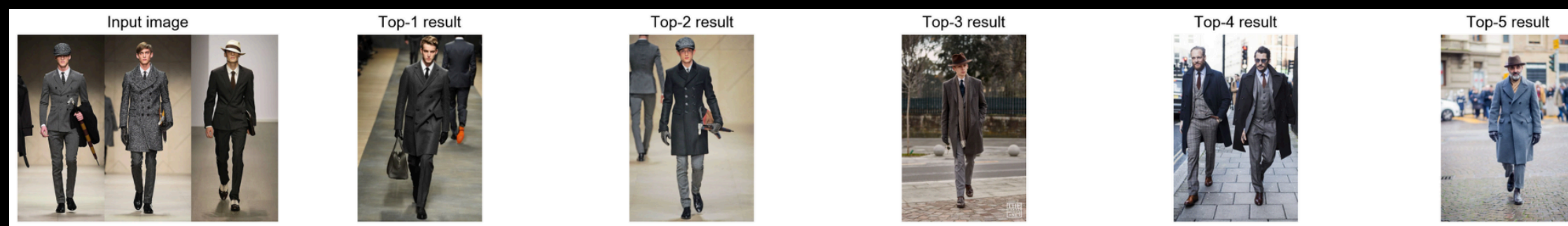
Vintage

Manually collect photos of 11 pre-categorized styles

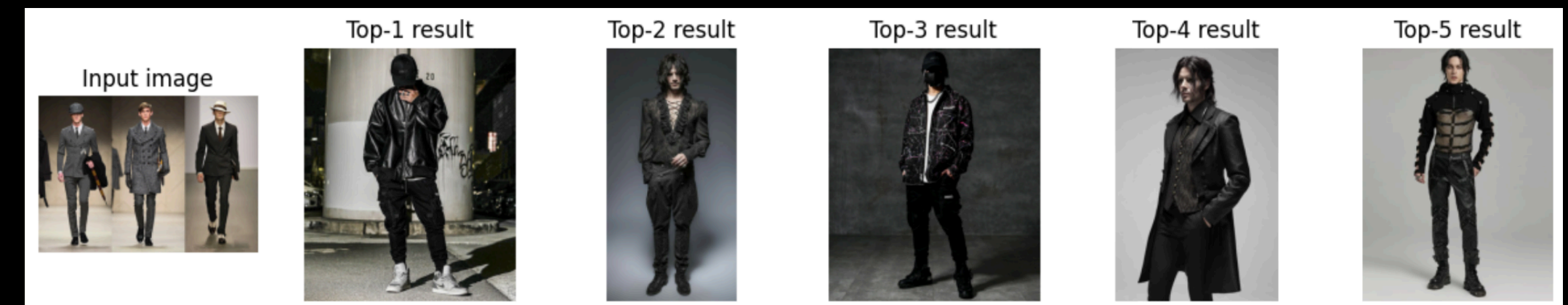
# Image Retrieval

Recommend Good outfits

Our Method: **84% ACCURACY**



MMFashion: **42% ACCURACY**



Optimized FashionCLIP

Style Pre-classified

open-mmlab/  
**mmfashion**



Open-source toolbox for visual fashion analysis  
based on PyTorch

5  
Contributors

12  
Used by

1k  
Stars

294  
Forks



# Image Retrieval

Recommend Good outfits



slasho\_k 2025-6-1

有人懂那種無力感嗎  
專題做到最後一刻發現阿里巴巴已經做過了



slasho\_k 2025-6-1

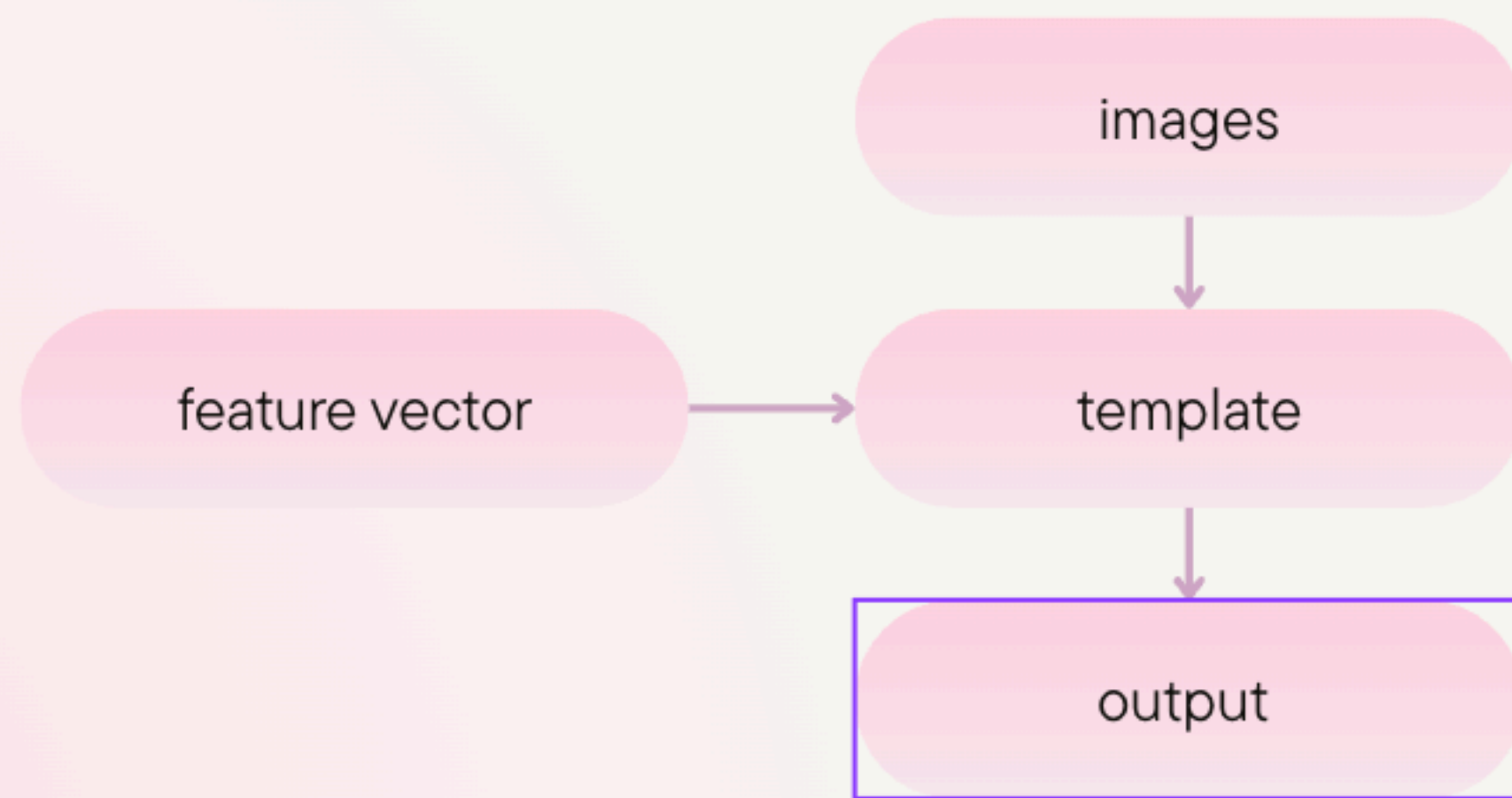
更：打贏阿里巴巴了  
🐮🍺



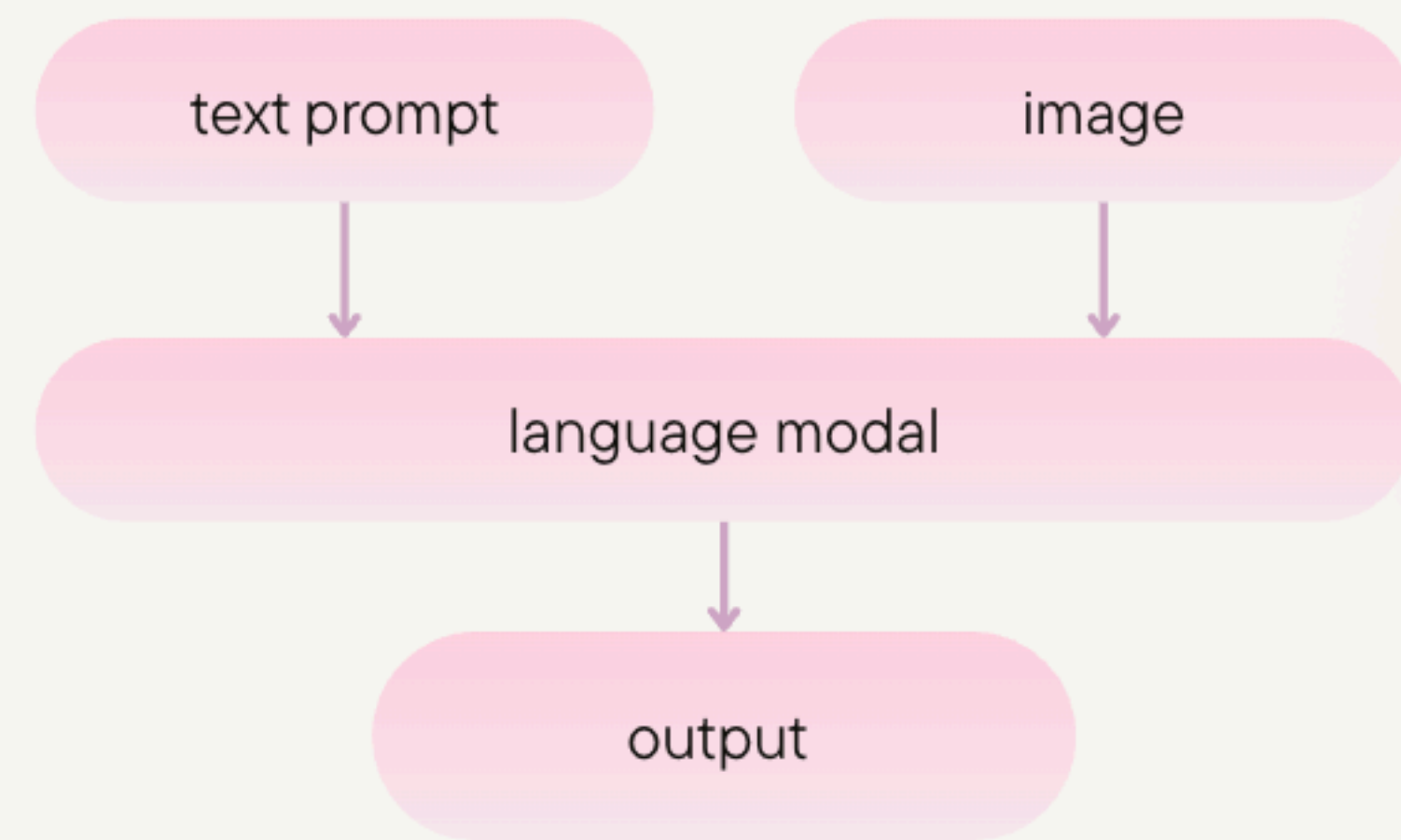
# Methods

## Generate Text-based Guide

### Rule-Based System



### Multimodal



# Results

## Generate Text-based Guide



### Rule-Based System (from FashionCLIP features)

```
- FashionCLIP: FashionCLIP 詳細特徵比較分析 (目標風格: STREET)
🔄 顏色: 您目前的blue clothing可以考慮調整為black clothing來更符合目標風格 (目標特徵中等: 0.64)
🔄 服裝類型: 建議將jacket替換為jeans以達到更好的效果 (目標特徵中等: 0.65)
✅ 風格: 很好! 您的casual style與目標完全匹配 (相似度: 0.98)
✅ 圖案: 很好! 您的solid color與目標完全匹配 (相似度: 1.00)
⚡ 配飾: 方向正確但可以更突出belt特徵
🔄 鞋類: 鞋類建議從casual shoes換成boots (目標特徵中等: 0.58)
✅ 材質: 很好! 您的denim material與目標完全匹配 (相似度: 1.00)
🔄 版型: 版型建議調整為oversized clothing而非loose fitting (目標特徵較弱: 0.48)
```

### Local Multimodal (LLaVA)

```
llava 穿搭分析:
1. 上衣改變
將目前休閒牛仔外套改為皮革外套, 以提升酷炫、街頭感。
搭配內搭換成帽T, 創造多層次且年輕的效果。

2. 下半身穿搭
牛仔褲保持寬鬆版型, 但可嘗試淺色系牛仔褲, 增加視覺對比感。

3. 配件轉變
取消黃色托特包, 換成運動感的棒球帽。
將鞋子從皮質感的休閒鞋轉換成更具街頭運動感的球鞋。

4. 顏色搭配
從深藍、黃、白色調轉變成更簡單的黑、白、淺藍色調, 保持整體簡約且明顯的街頭時尚感。
```

# Appendix: Personalized recommendation system

The biggest problem with our topic:

Everyone has a different definition of fashion classification/good-looking .

Our Solution:

- Personalized style customization:  
Choose your favorite dressing style
- Learn **users' preference weights** for different tag combinations through **sliding behavior**



3Q