

Pedestrains Detection

Soft Biometric

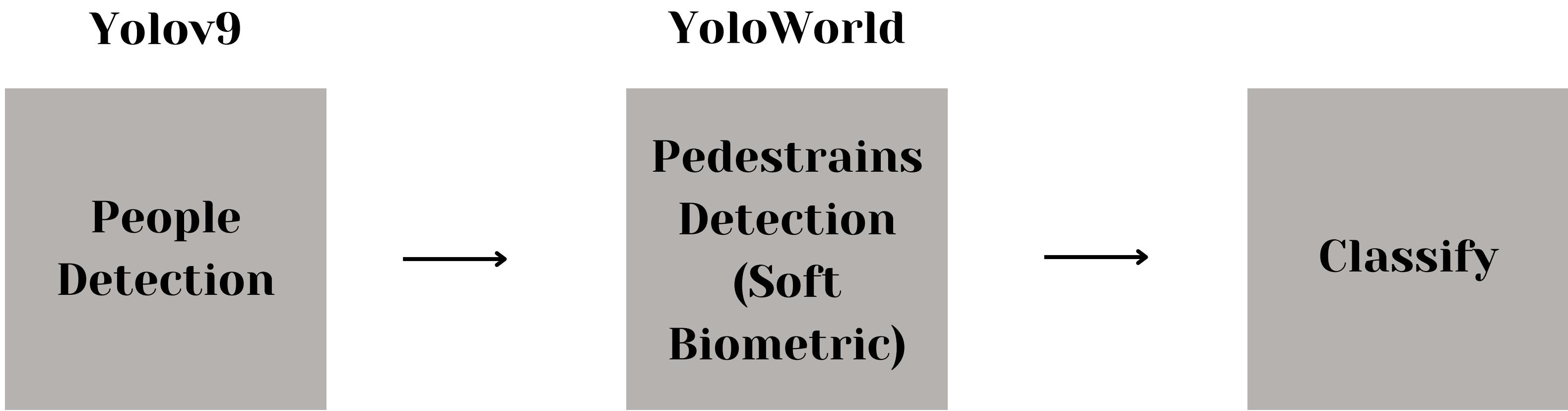
Pedestrains

Detection

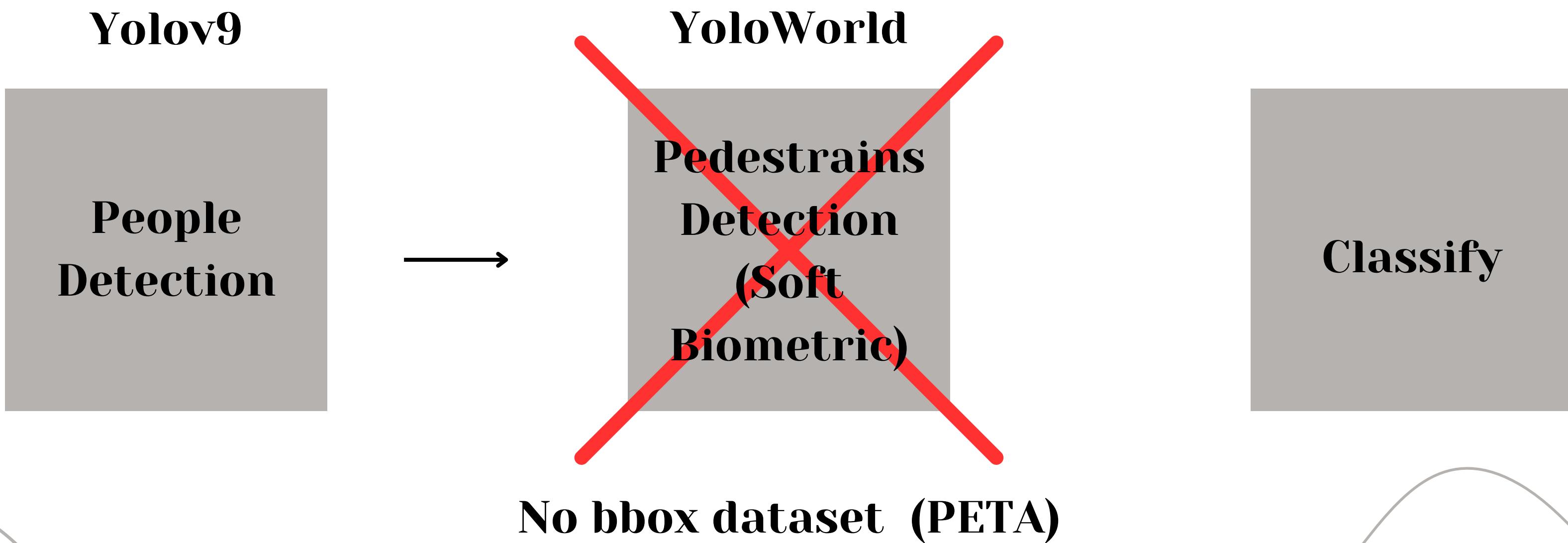
Goal

**To extract information from people in the
image/video as much as we can. Especially the
fashion (Cloths, Pants, ...)**

Ideal Plan



Ideal Plan



Ideal Plan

Yolov9



YoloWorld

~~Pedestrains
Detection
(Soft
Biometric)~~

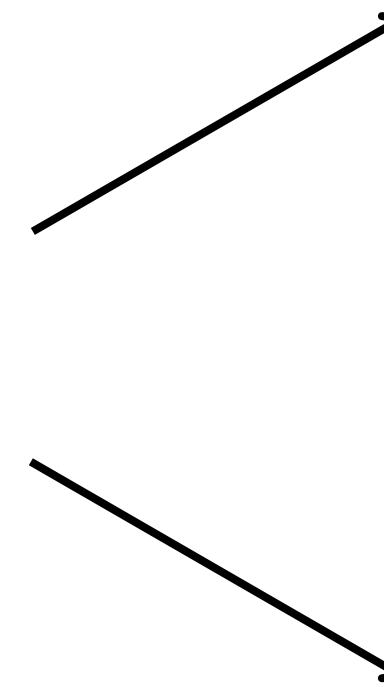
No bbox dataset (PETA)

~~Classify~~

Plan B

Yolov9

**People
Detection**



**Pedestrains
Classification**

?

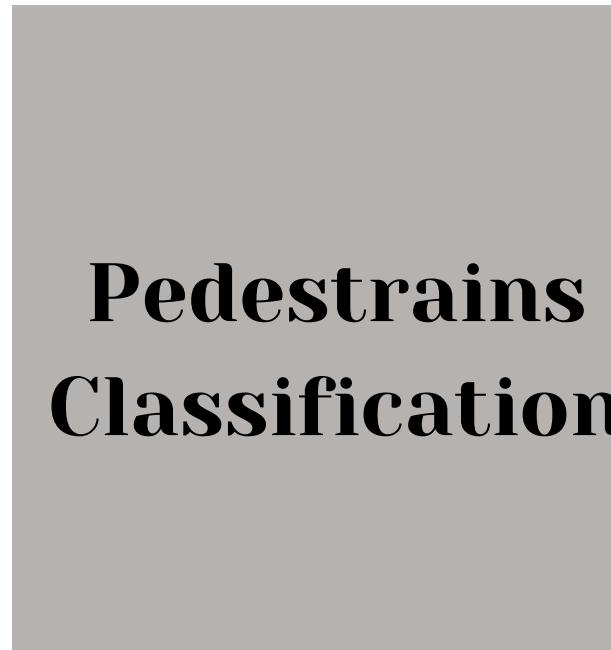
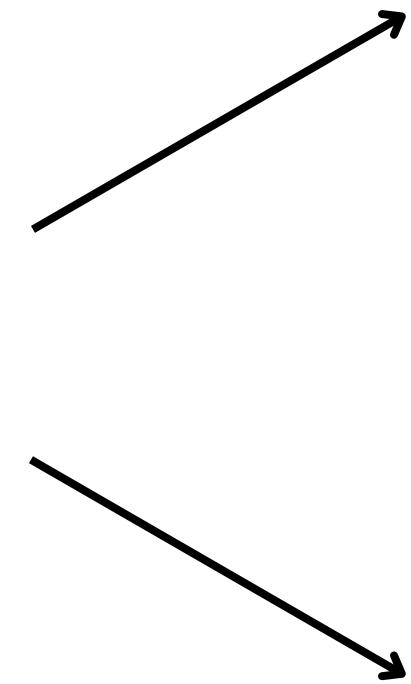
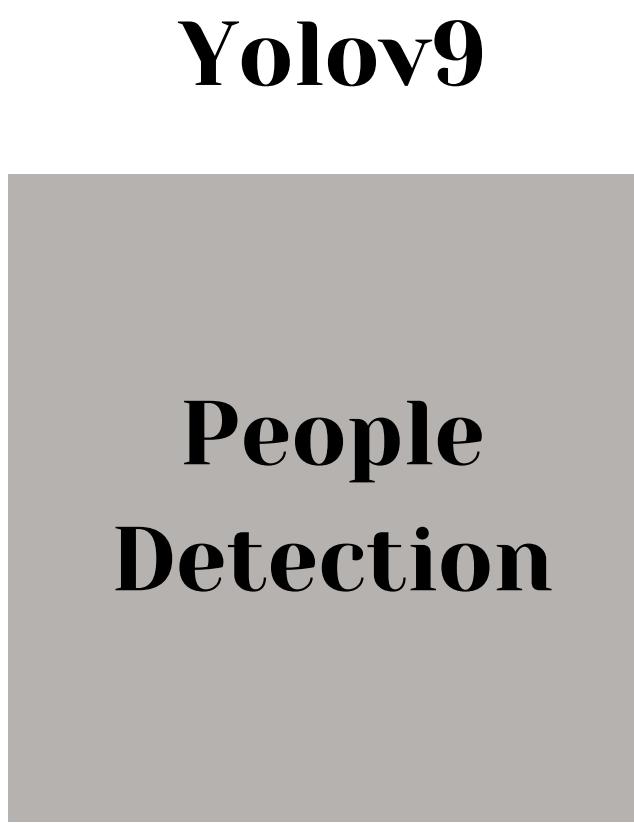
**Fashion
Detection**

Yolov9

**Others
Soft Biometric**

Cloths

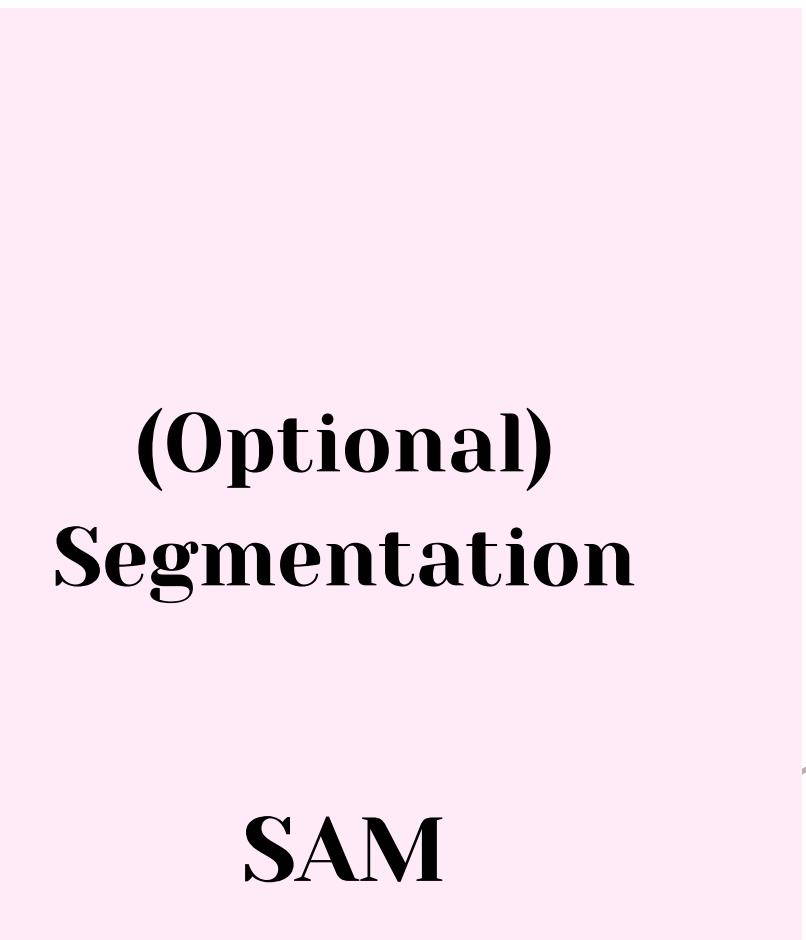
Plan B



?



Yolov9



Yolov9



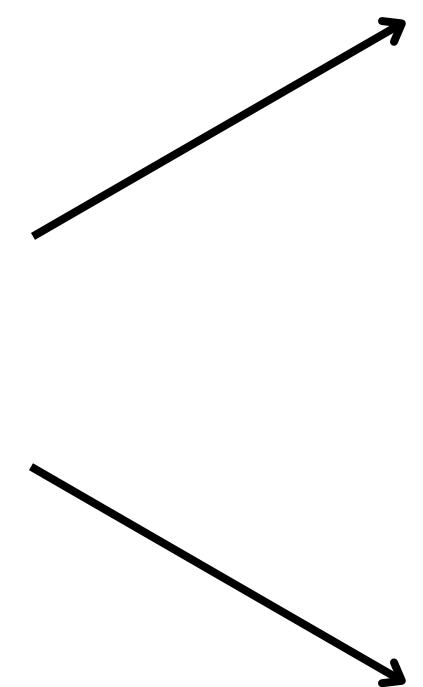
Plan B

Pedestrains
Classification

?

Fashion
Detection

Yolov9



**Why not use the instance
segmentation model
instead of using
detection +
segmentation?**

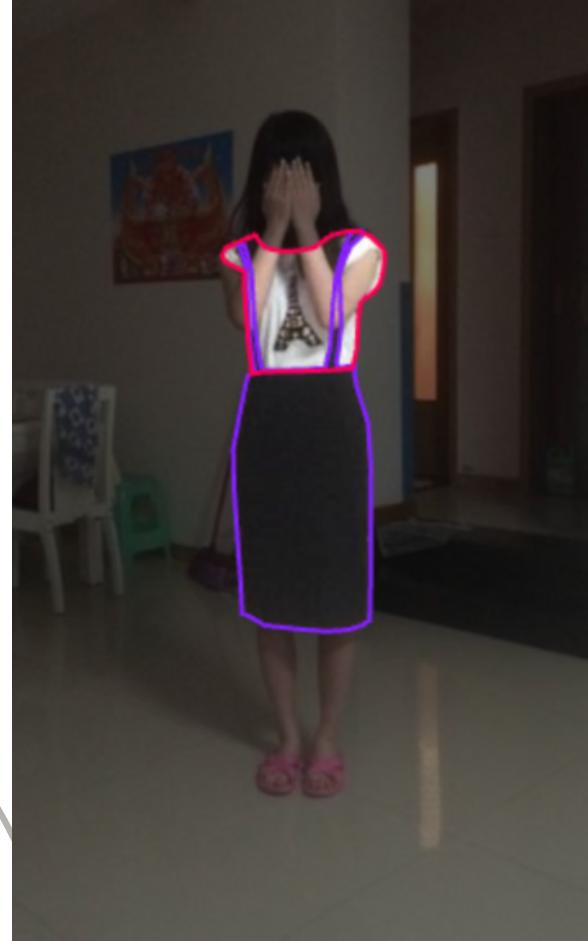


(Optional)
Segmentation

SAM

Fashion Detection

Yolov9



Deep Fashion2-m-11k

- 13 Attributes
- 6600 Images
- Instance Segmentation
- No color detection, Need to do some logic.
- From RoboFlow
- Original Deep Fasion2 has ~491,000 images





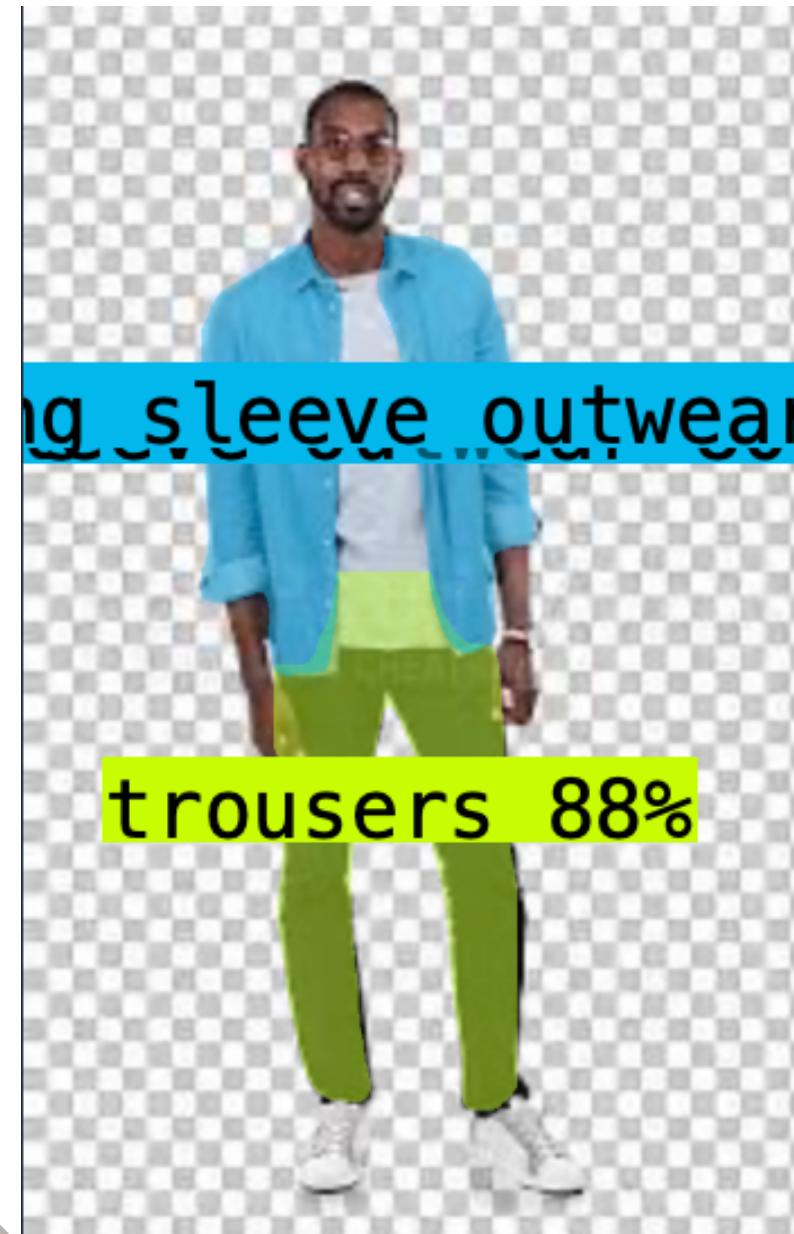
DeepFasion2

Attributes (13)

- **long sleeve dress**
- **long sleeve outwear**
- **long sleeve top**
- **short sleeve dress**
- **short sleeve outwear**
- **short sleeve top**
- **shorts**
- **skirt**
- **sling**
- **sling dress**
- **trousers**
- **vest**
- **vest dress**

Expected Output

Roboflow 2.1 Instance Segmentation



Expected Output

Roboflow 2.1 Instance Segmentation, **confidence = 0**



Actual Output

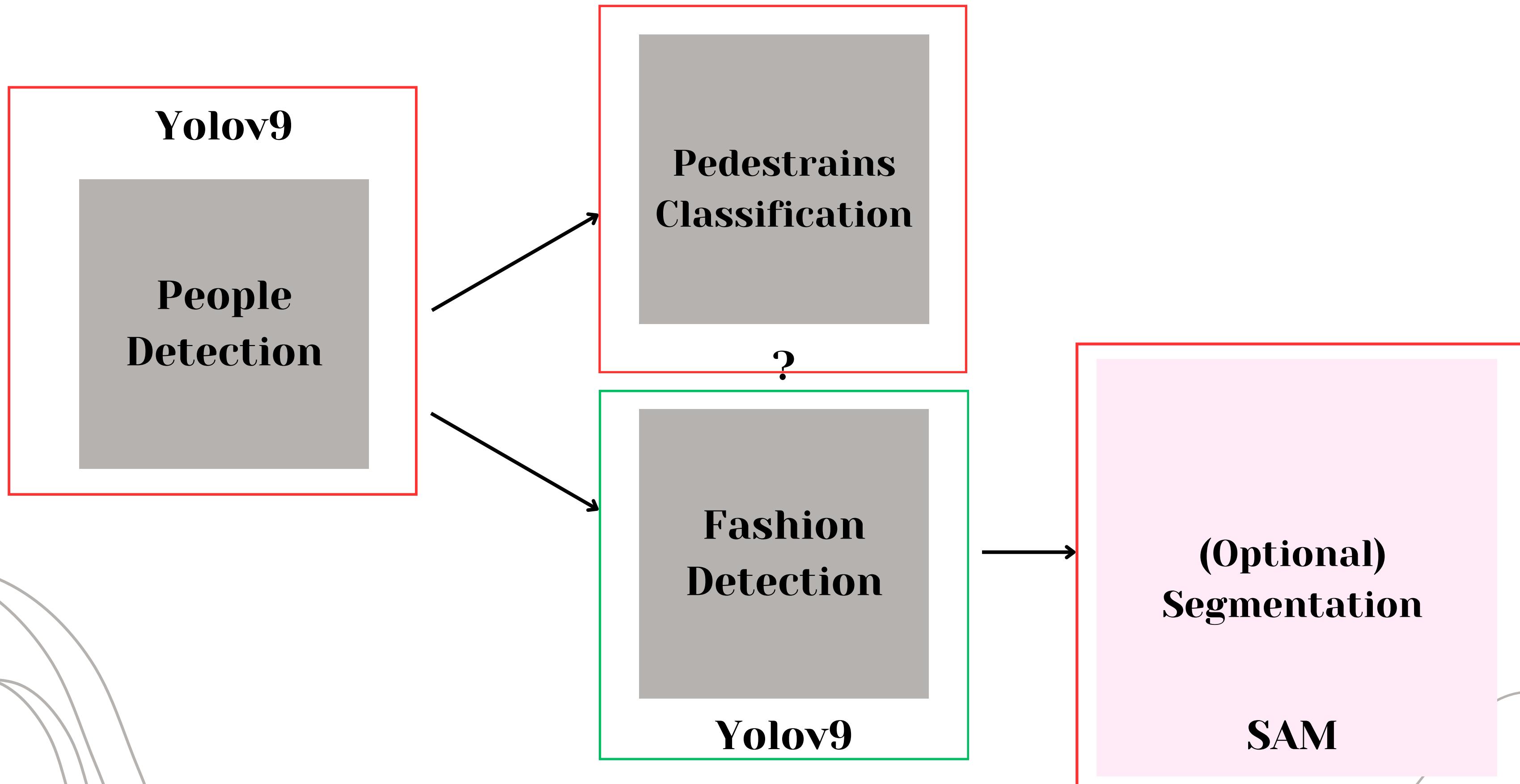
Yolov9 trained with 20 epochs



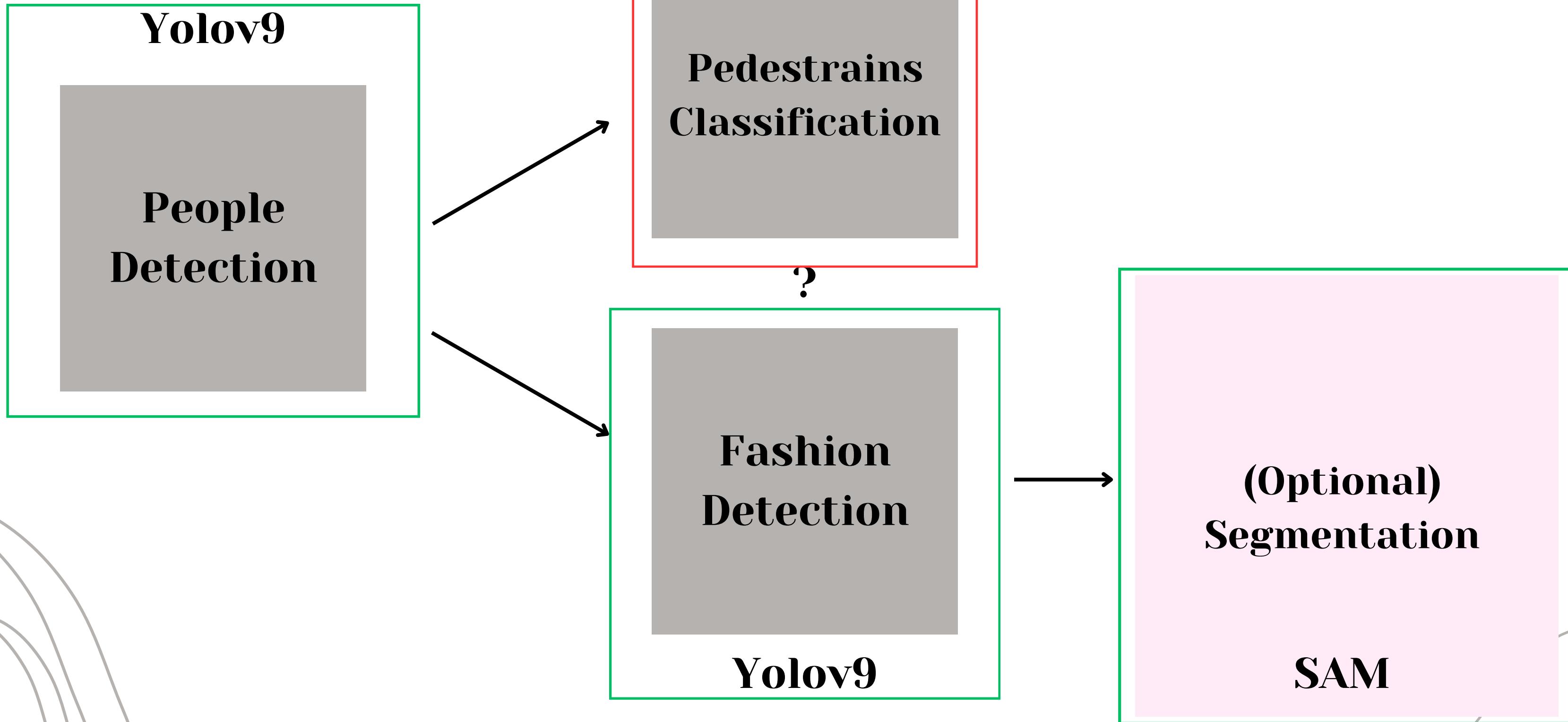


What's next?

What's next?



What's next?



People Detection

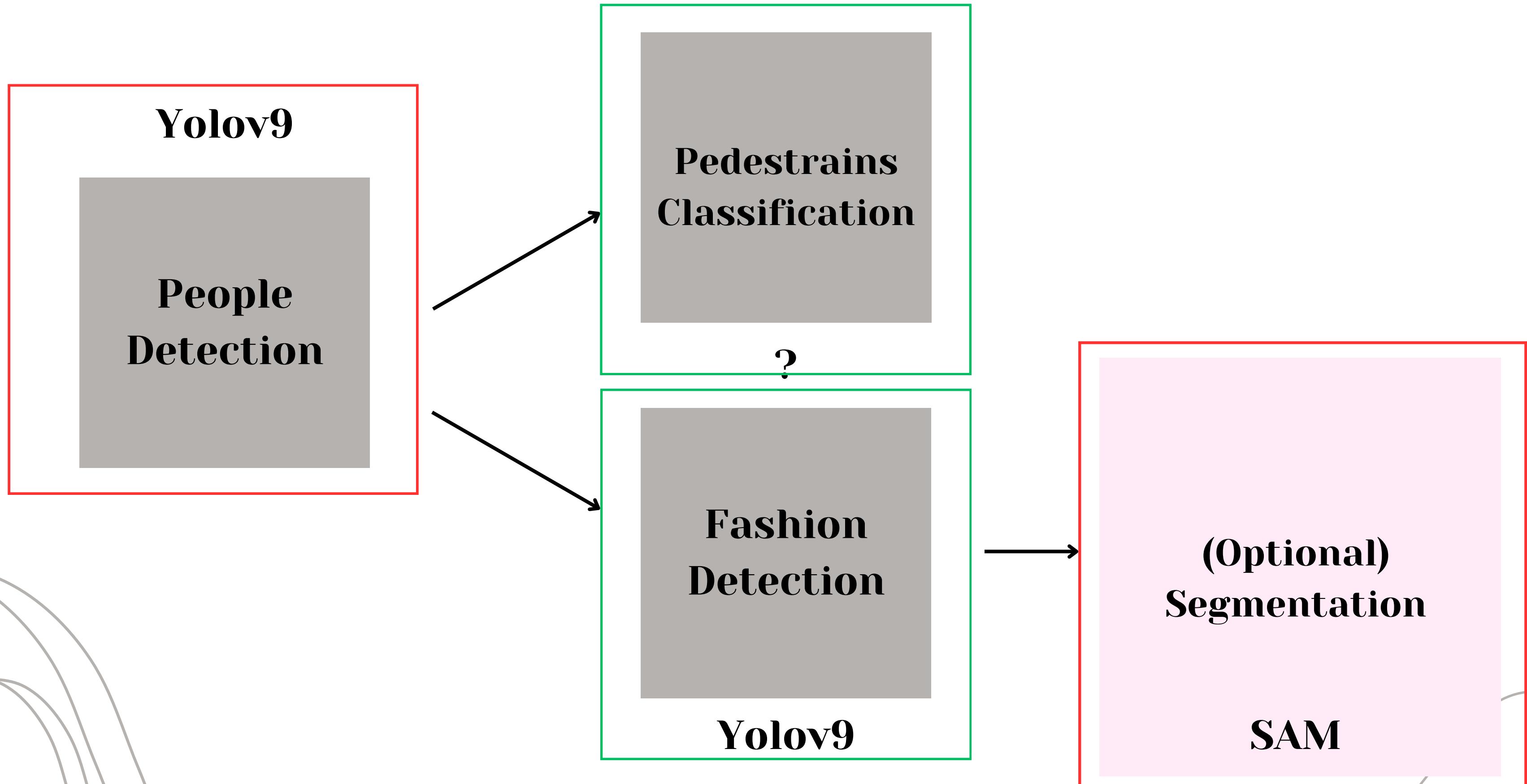
- In case we want high accuracy in detecting people.
- Might not need additional training, Yolov9 did well.

(Optional)
Segmentation

SAM

- In case the accuracy is so poor, we can do additional training on this model.

What's next?



Pedestrains Classification

UPAR (Unified Pedestrian Attribute Recognition and Person Retrieval)

- **Combination of PA100K, PETA, RAPv2, and Market1501**
- **Unifies 40 attributes across 12 categories**
- **Contains 224,737 images with annotations**
- **The paper used the CNN model to train on this dataset**

Project



Select Attributes



Upper Part



Lower Part



| long sleeve dress



| long sleeve top



| short sleeve dress



| sling dress



| vest dress



Physical Part



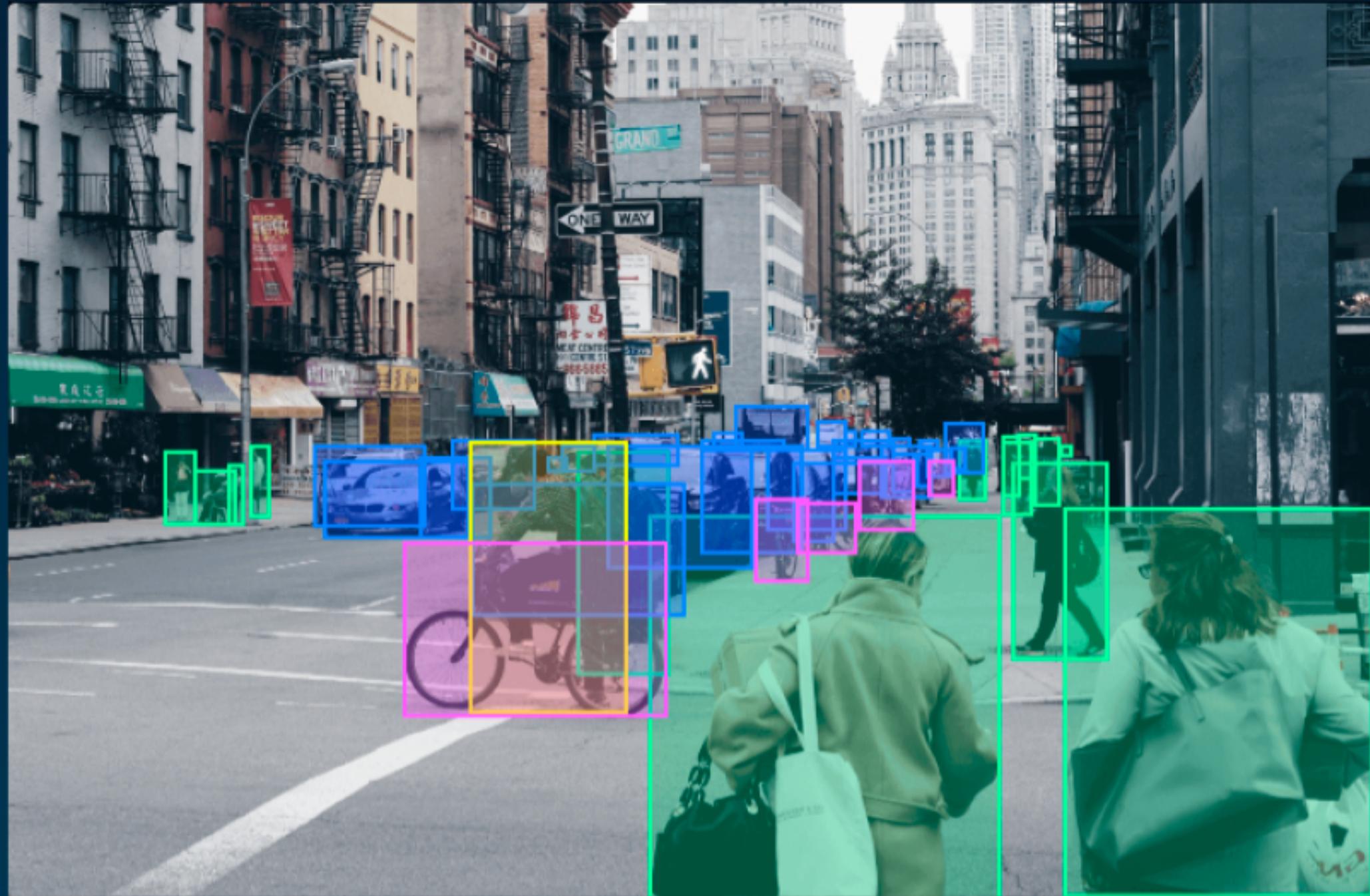
Location



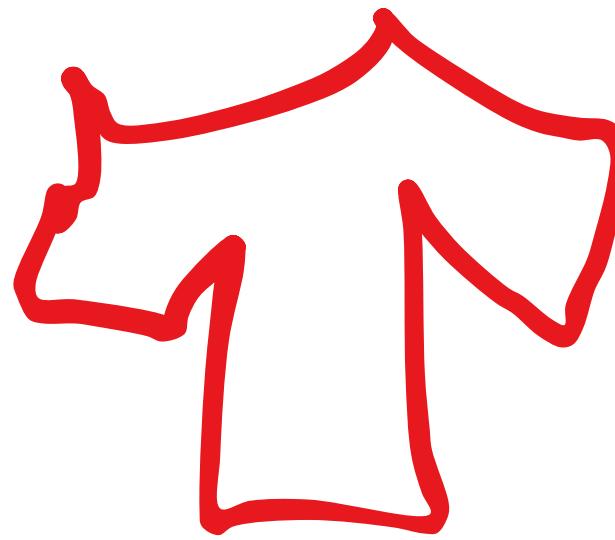
Selected 8/17

Finished 100%

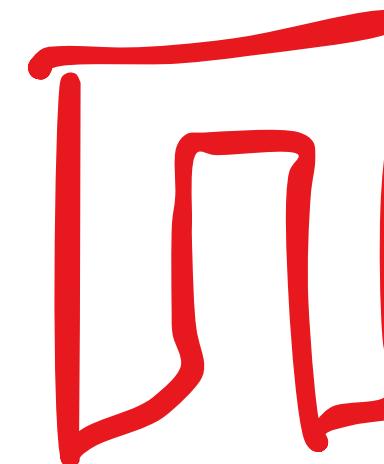
456 * 768 Pixel



Upload image



T-shirt
Red color



Trouser
Bluc
Confidence