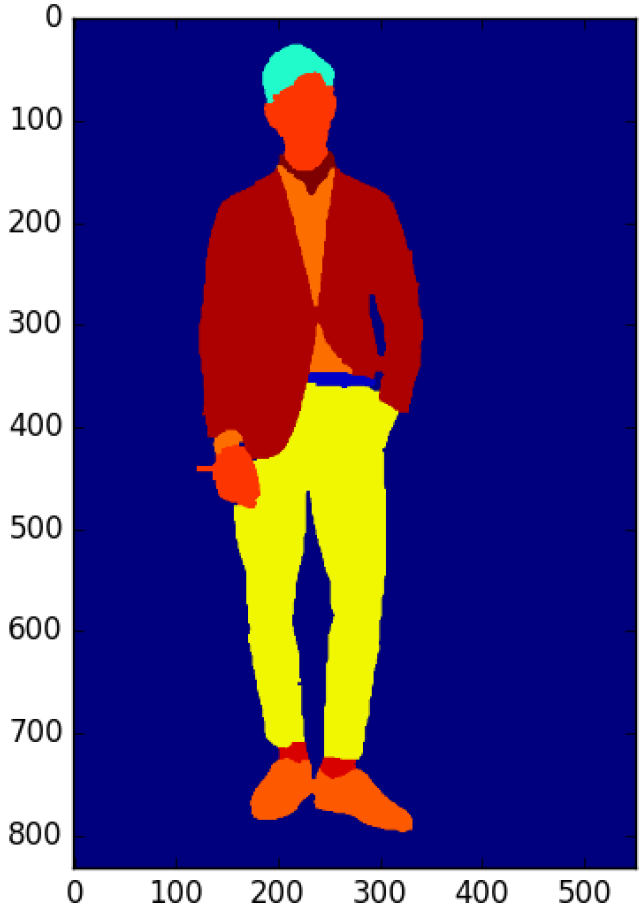
Clothing Parsing using MRF

Members: Daitao Xing
Jun Chen
Jin Huang
12/08/2016

Project: Clothing Parsing



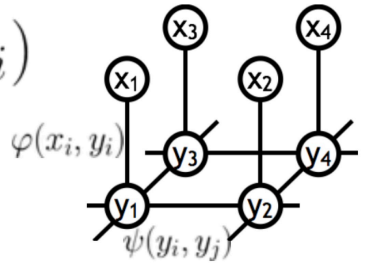
What are the difficulties?

- Background Influence
- Intro-class diversity
- Inter-class similarity
- What need to consider to better recognize?
  - Position
  - Color
  - Texture
  - Shape

Markov Random Field(MRF)

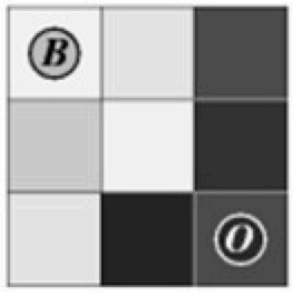
Energy Formulation

$$E(x,y) = \sum_i \varphi(x_i, y_i) + \sum_{i,j} \psi(y_i, y_i)$$



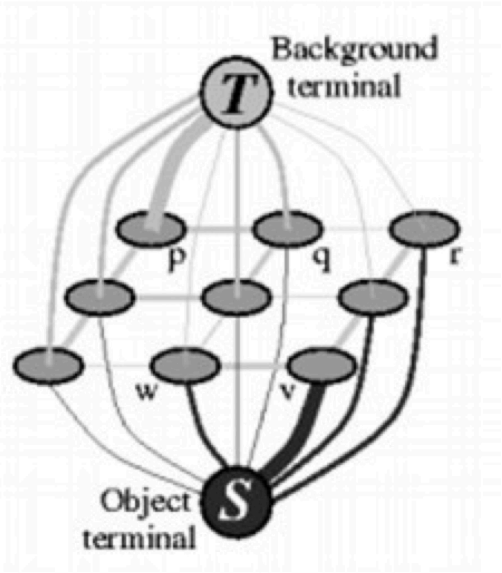
- Unary potentials φ
  - Local information about each pixel
  - e.g. how likely a pixel/patch belongs to a certain class

- Pairwise potentials ψ
  - Neighborhood information, enforces consistency
  - e.g. how different a pixel is from its neighbor in appearance

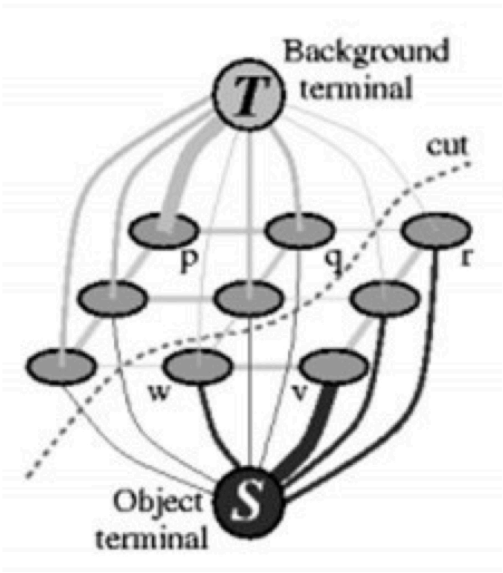


(a) Image with seeds.

(d) Segmentation results.



(b) Graph.



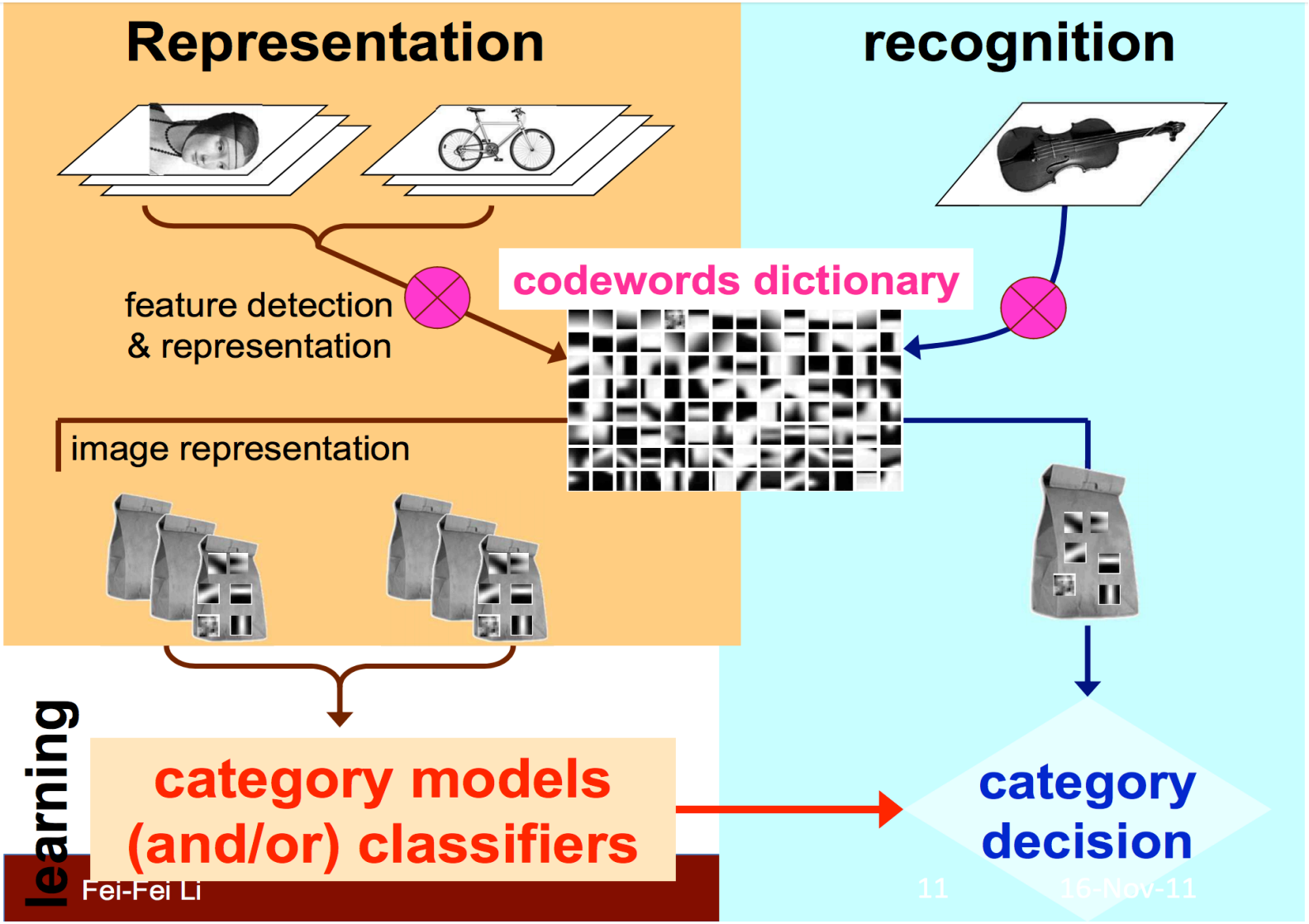
(c) Cut.

Build a good unary function(1/4)

$$\log \Pr[L|I] = \sum_{i \in U} \phi(l_i|I) + \lambda_1 \sum_{(i,j) \in V} \psi_1(l_i, l_j) + \lambda_2 \sum_{(i,j) \in V} \psi_2(l_i, l_j|I) - \log Z \quad (2)$$

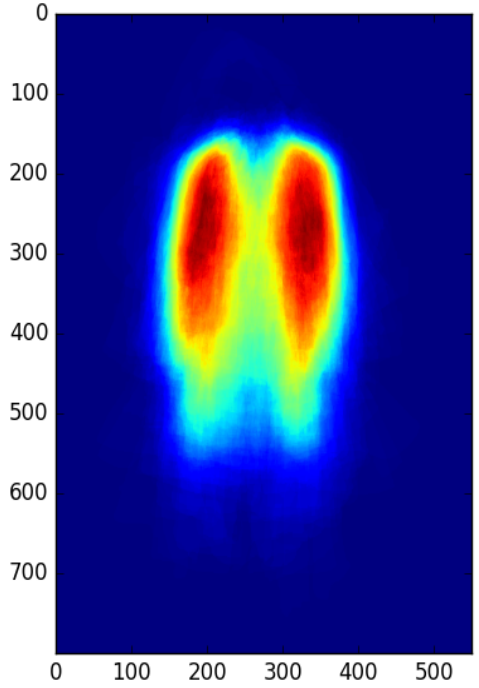
- Position Information: Human pose estimation
- Color Space : RGB and LAB
- Texture: Gabor Bank Filter
- BOW

Build a good unary function—Using BOW(2/4)

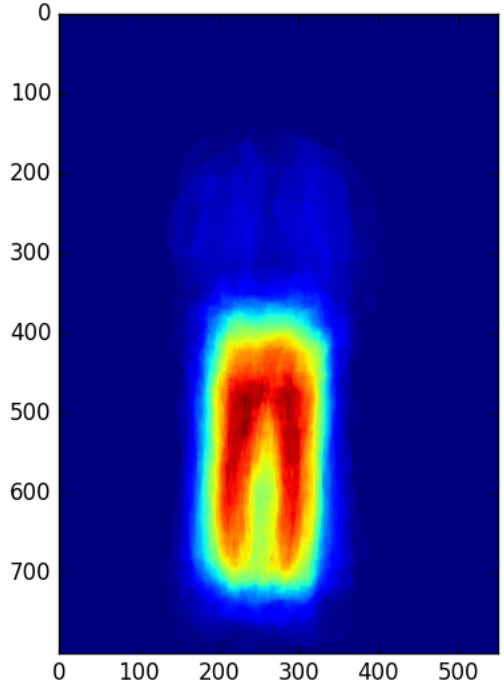


Feifei Li, Stanford

Build a good unary function(3/4)



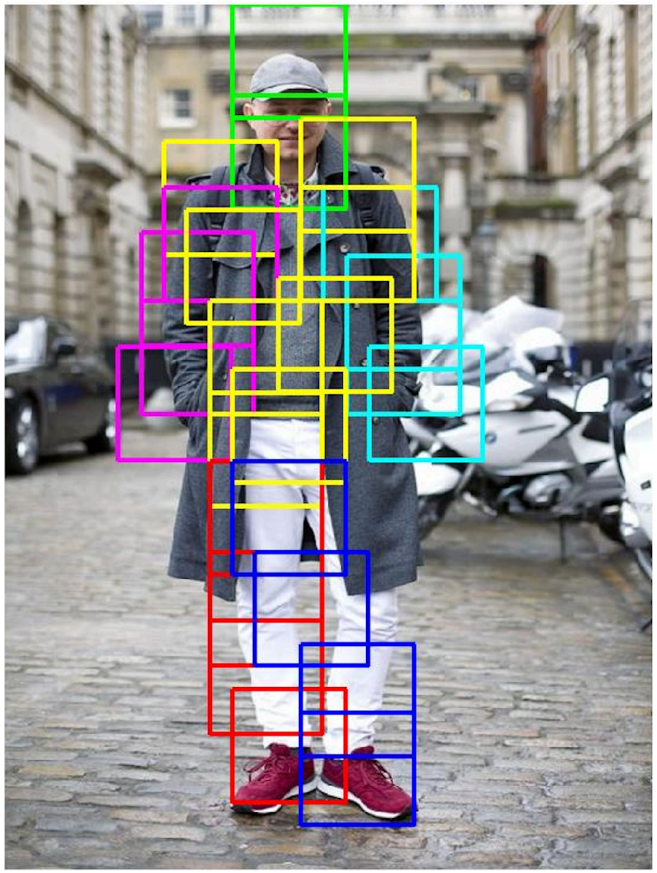
Coat



Jeans

Visualization of clothelets of images.

Build a good unary function(4/4)



The Pose detection

In our model, we select the following as features

(1) RGB color histogram,

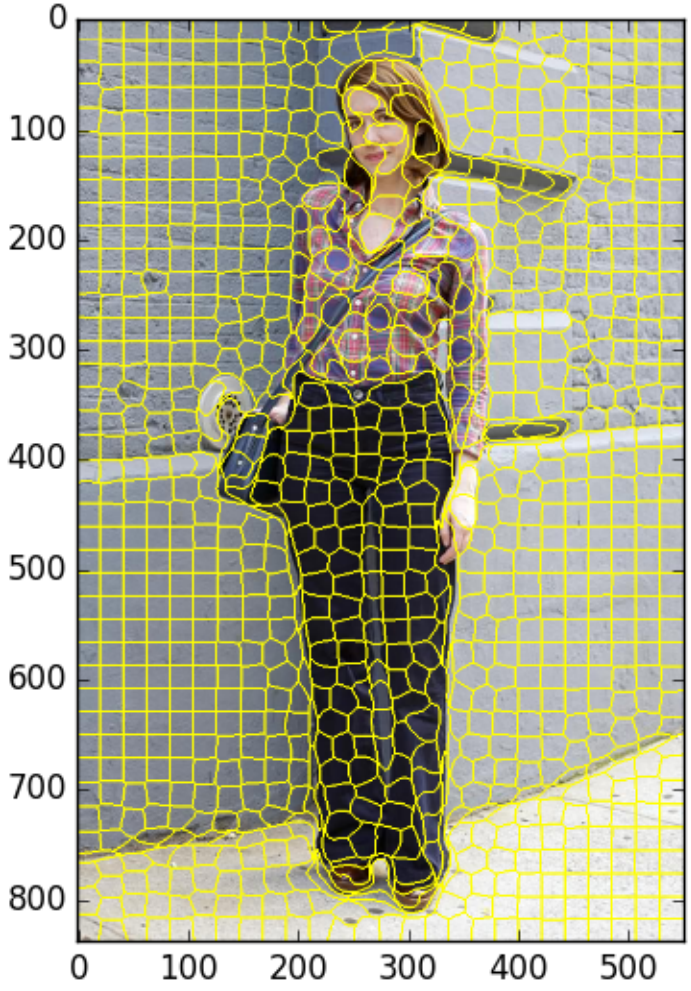
(2) LAB color space,

(3) normalized coordinates,

(4) pose descriptors,

(5) the combination of bag of visual words and SIFT features.

After extracting features, we use random forest to assign the labels to each superpixel.


0
100
200
300
400
500
600
700
800
0
100
200
300
400
500

Result (1/2)

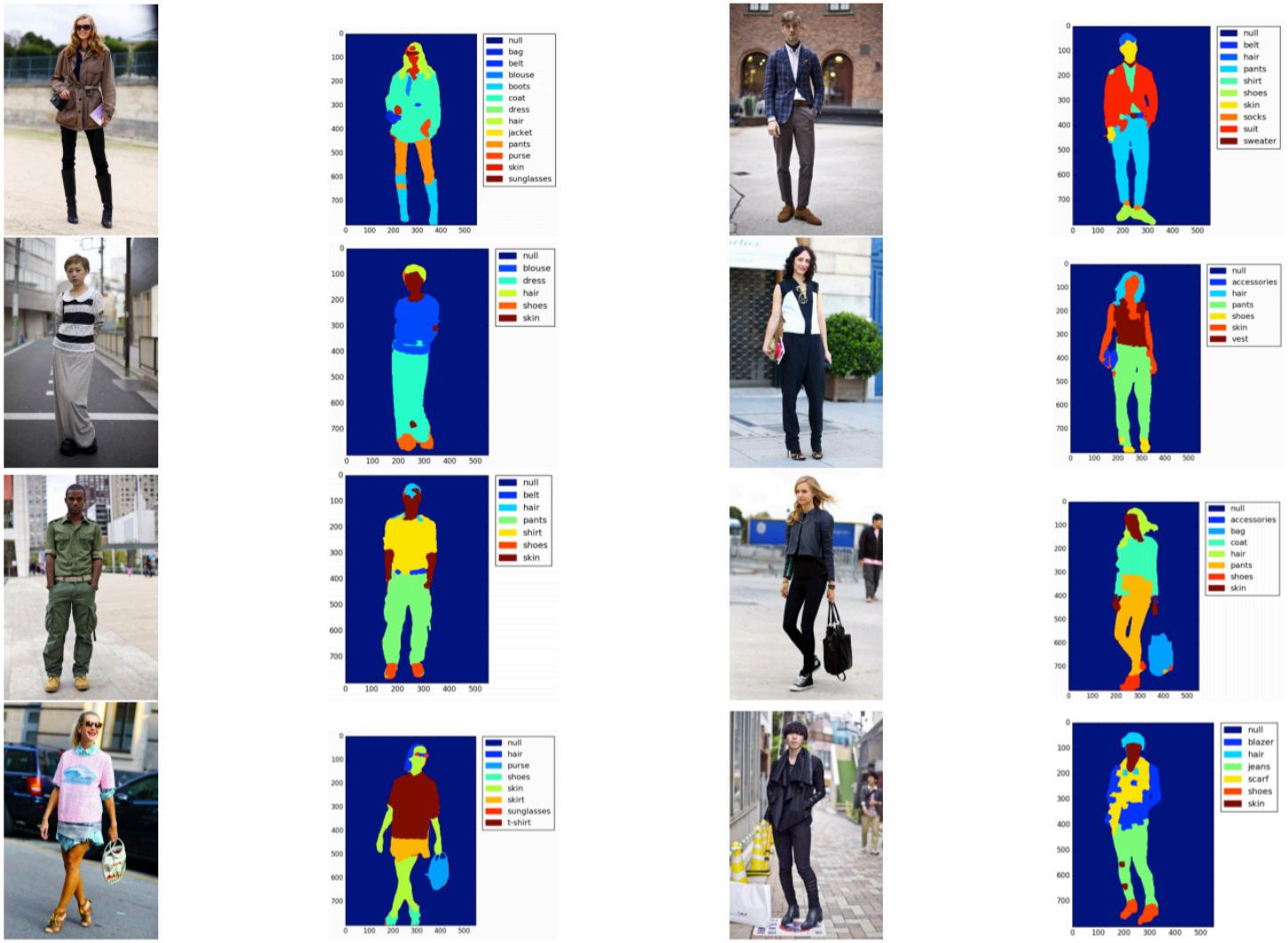


Fig. 2. Some good results.

Result (2/2)

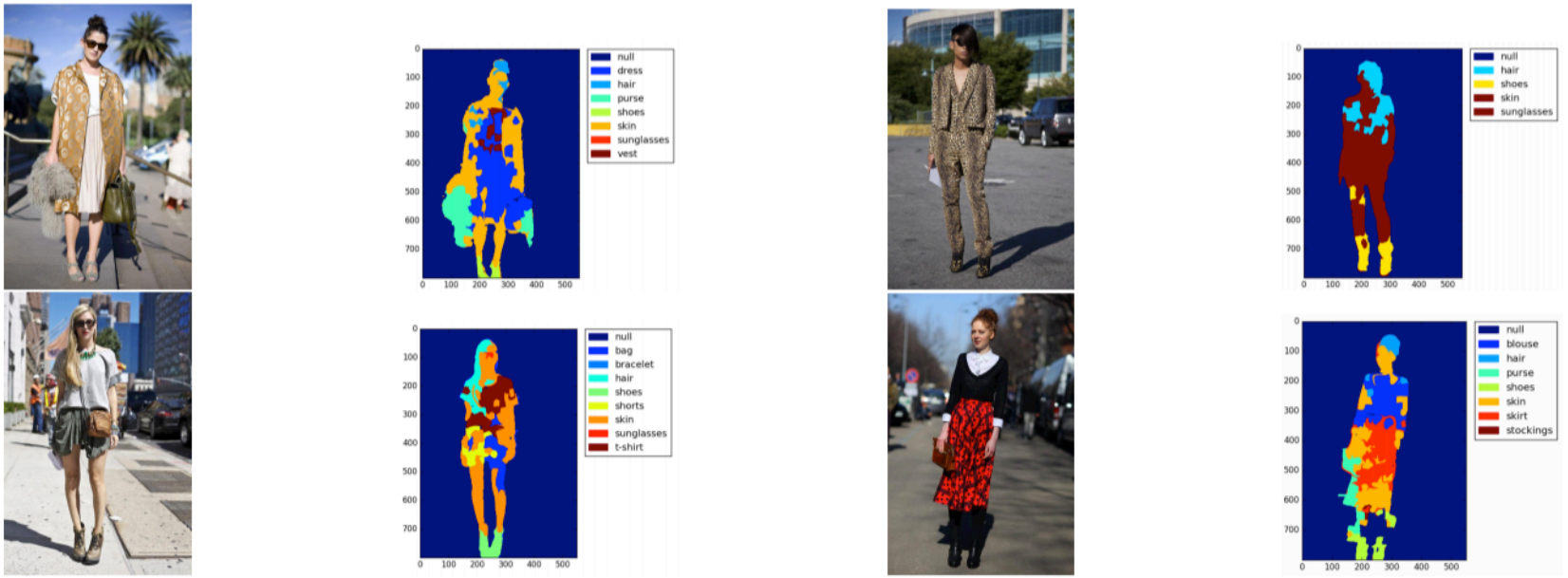


Fig. 3. Some unsuccessful results.

Future works

- We mainly have two targets in the next step:
- First, we plan to use a more general unary function, in order to get rid of the limitation of position. Thus to decline the dependence of pose detection.
- Secondly, we would like to improve the function to more complex environment and texture, aiming at increasing the accuracy of segmentation and labeling.

Thank You!