



# Deep Fashion Understanding

Ziwei Liu

Multimedia Lab, The Chinese University of Hong Kong

# Human-centric Analysis



# Human-centric Analysis



Person Re-identification



# Human-centric Analysis



Face Recognition

# Human-centric Analysis



Fashion Understanding

# Overall Pipeline



Clothes Detection



# Overall Pipeline

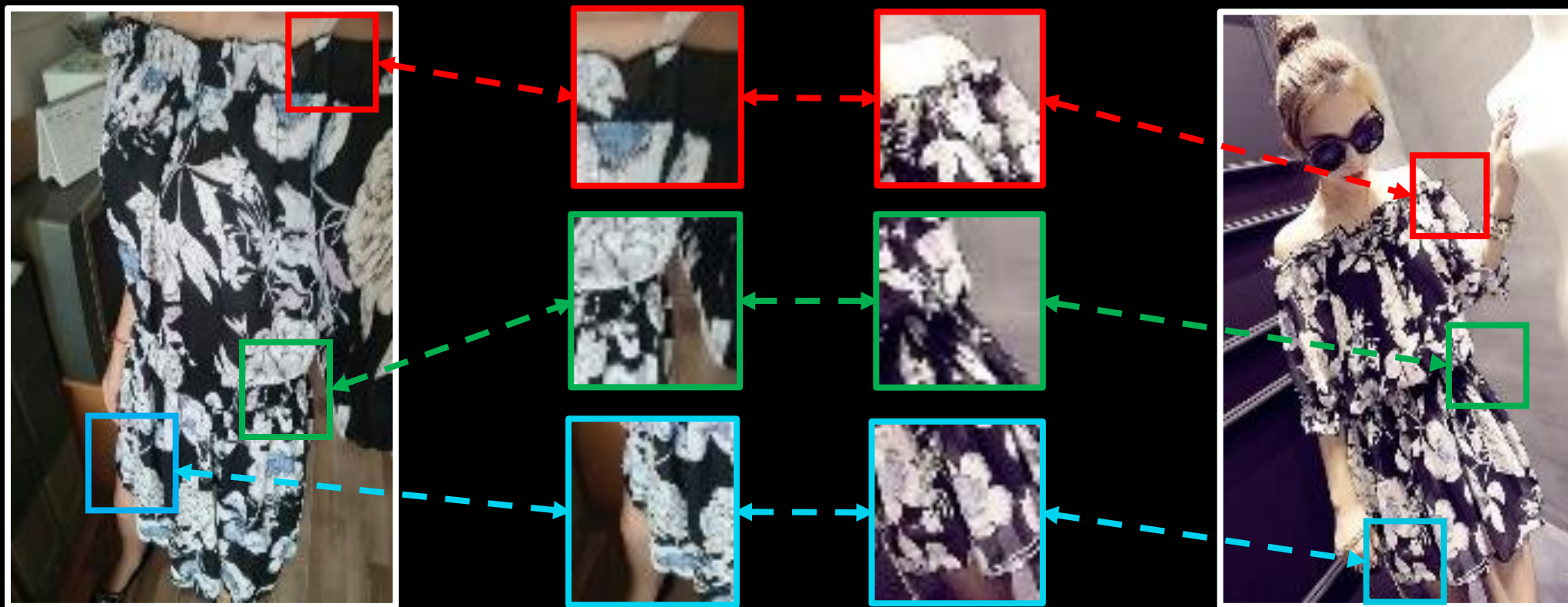


Clothes Detection



Clothes Alignment

# Overall Pipeline

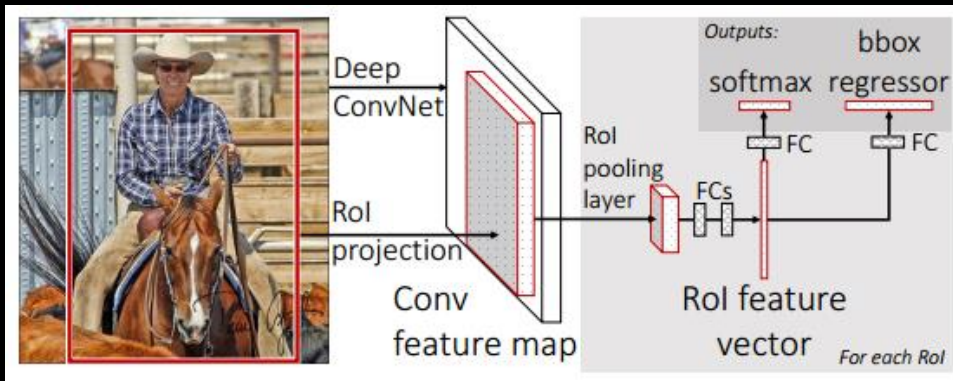


Clothes Recognition

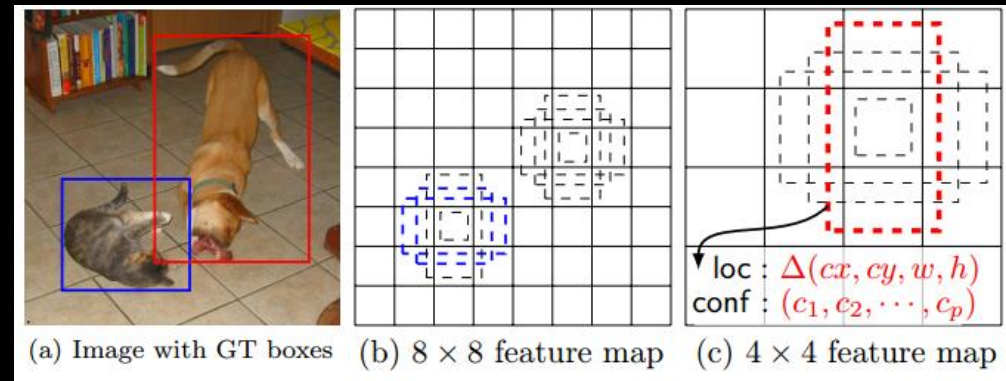


# Clothes Detection

A special class of general object detection



Fast R-CNN



SSD

Leverage domain knowledge

Divide & Conquer

Geometric Constraints

# Clothes Alignment

A set of fashion landmarks

Collars  
Cuffs  
Waistlines  
Hemlines  
...



(a.1)



(a.2)



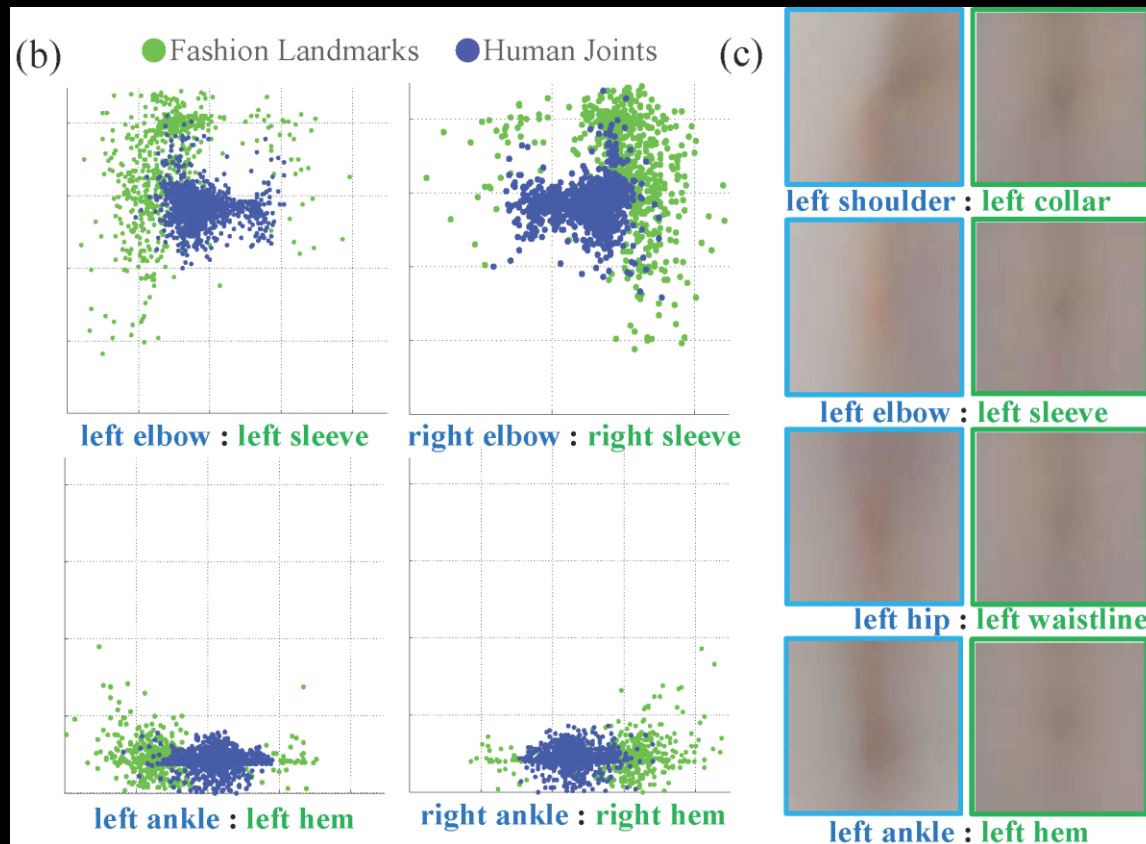
(a.3)



(a.4)

# Clothes Alignment

More challenging than human pose estimation



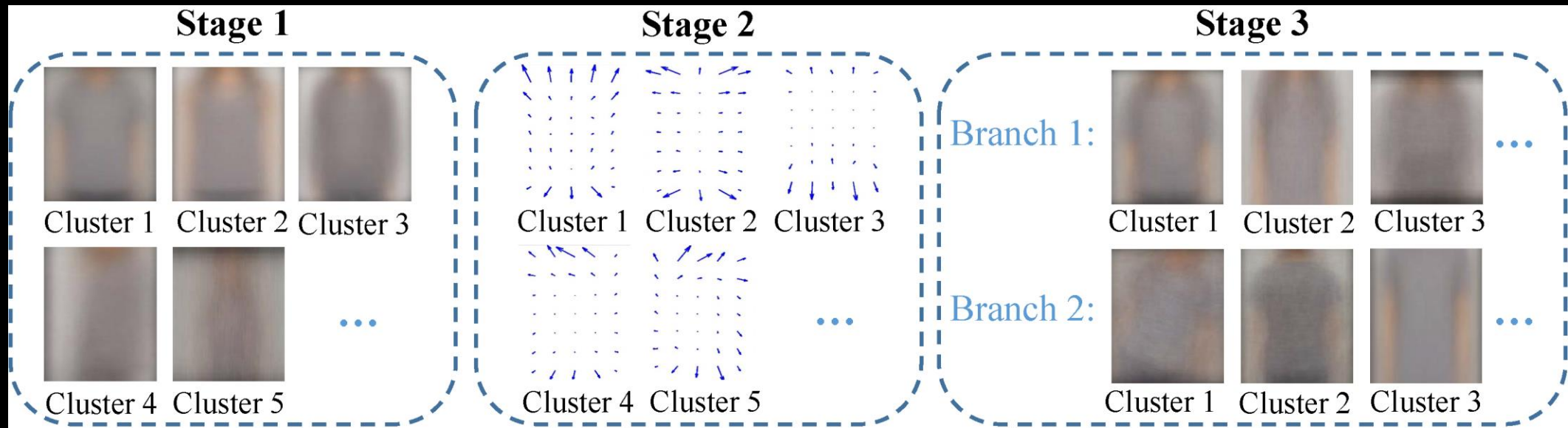
Geometry

Appearance



# Clothes Alignment

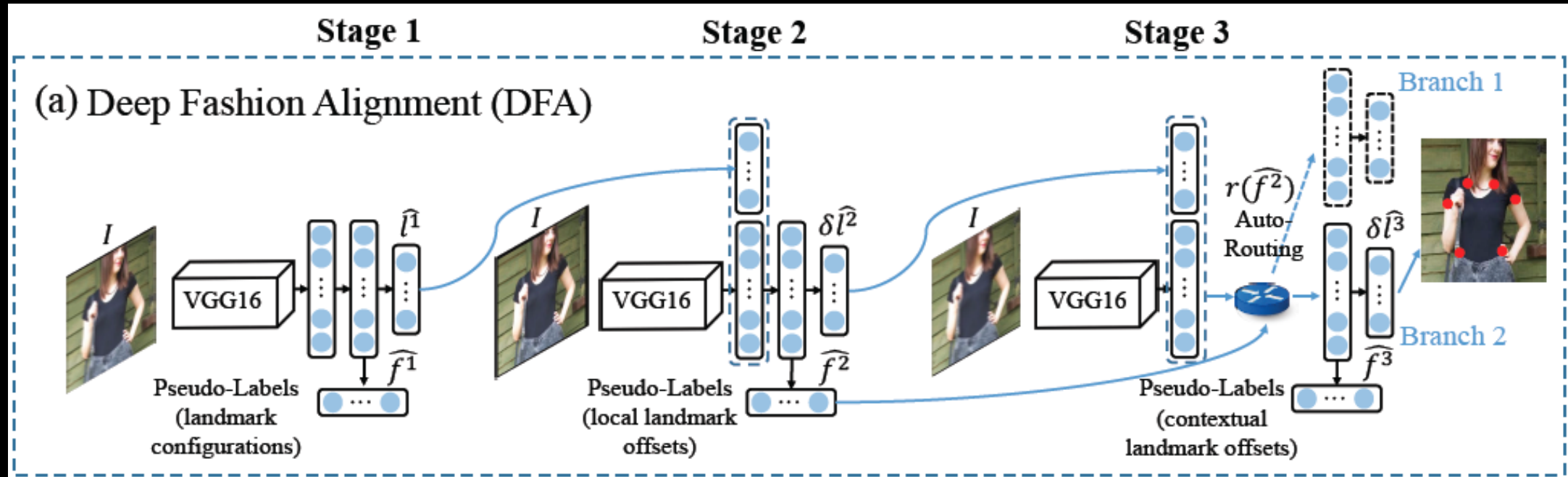
Reduce variations by pseudo-labels



Obtain codebook by k-means clustering in label space

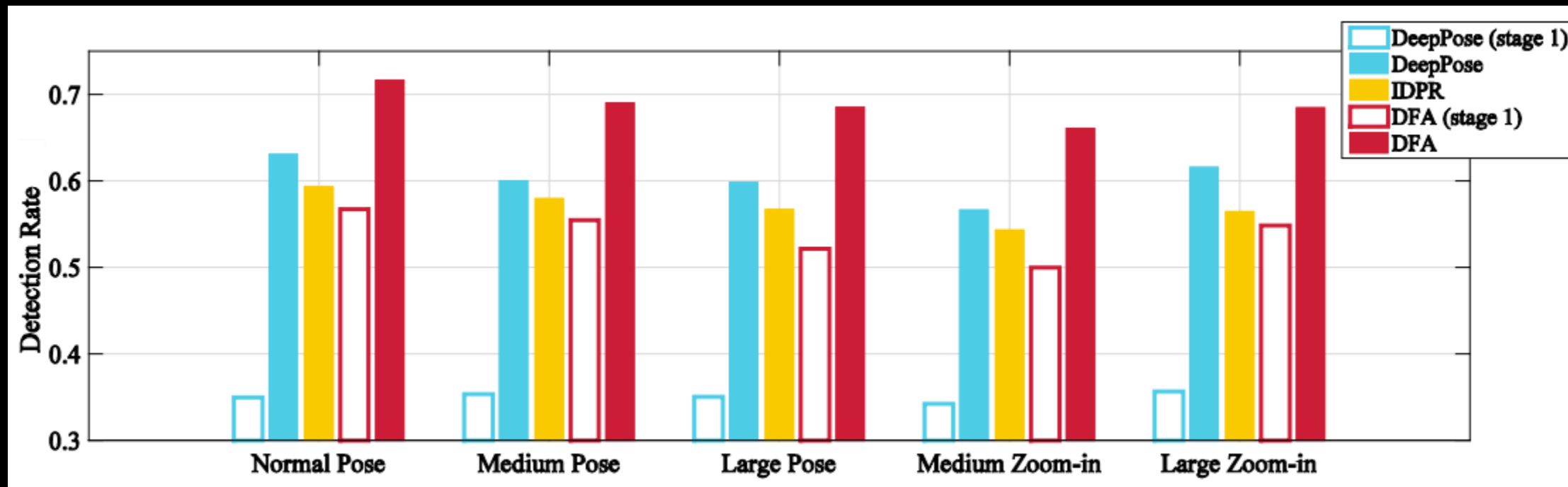
# Clothes Alignment

Reduce variations by pseudo-labels



# Clothes Alignment

























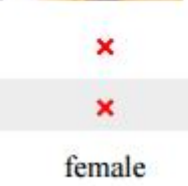
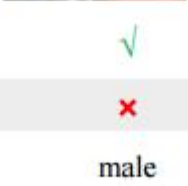
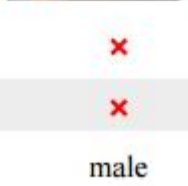

## Performance





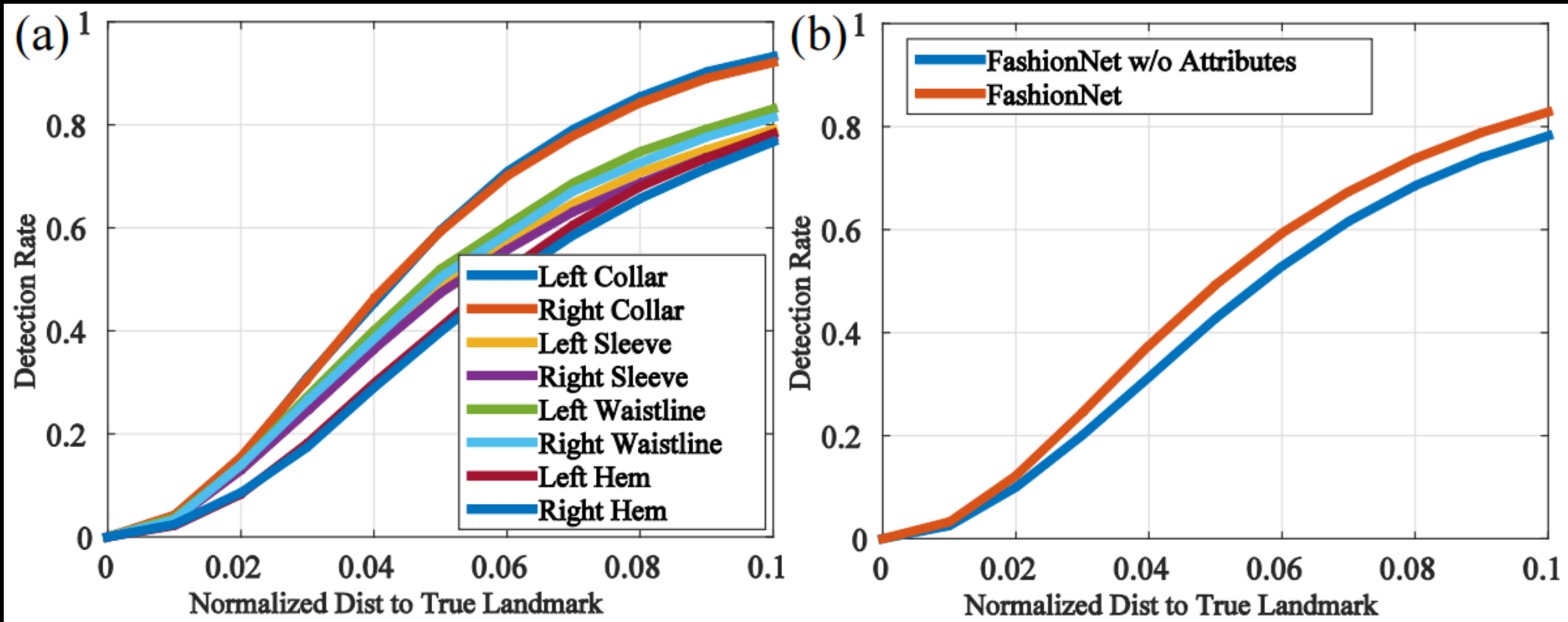
# Clothes Alignment

## Relationship to multi-task learning

Auxiliary Tasks	CNN							
	Cascaded CNN							
	TCDCN							
								
wearing glasses		×	×	✓	×	✓	×	×
smiling		×	✓	×	×	×	×	×
gender		female	male	female	female	male	male	female
pose		right profile	frontal	frontal	left	frontal	frontal	right profile

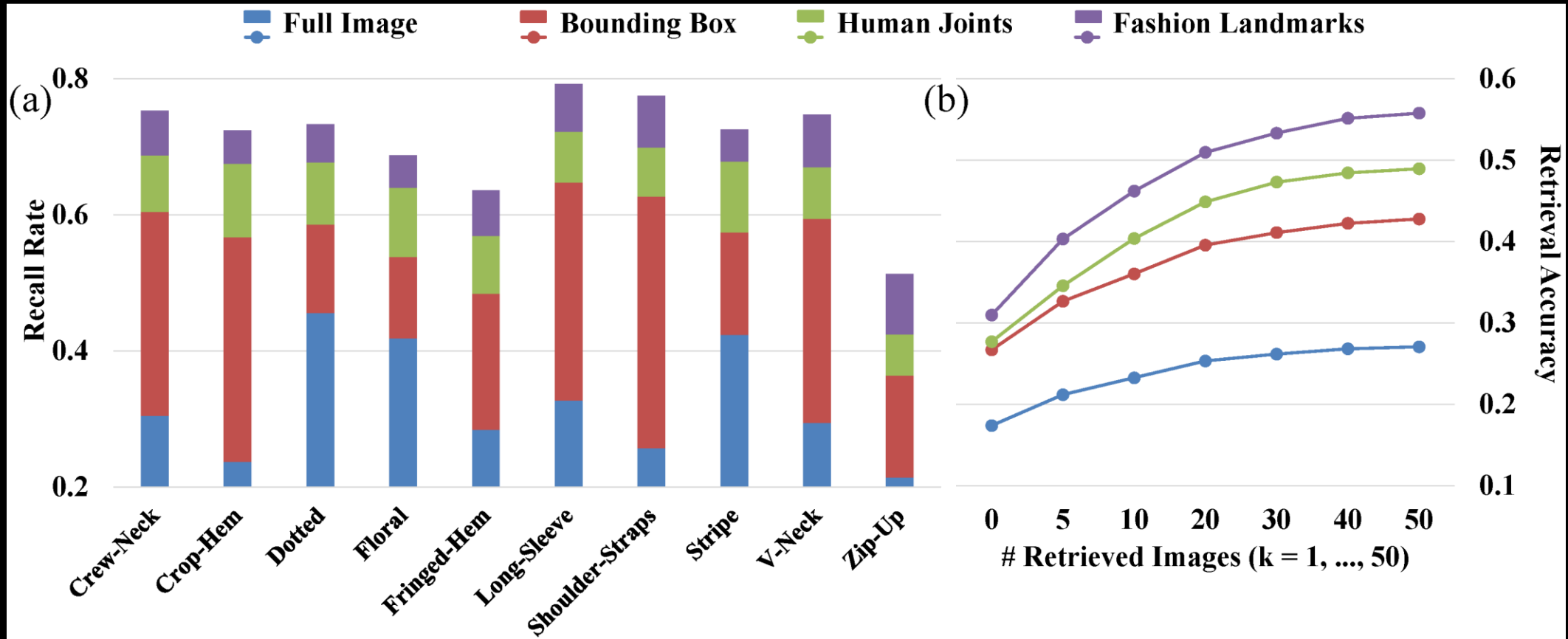
# Clothes Alignment

Relationship to multi-task learning



# Clothes Alignment

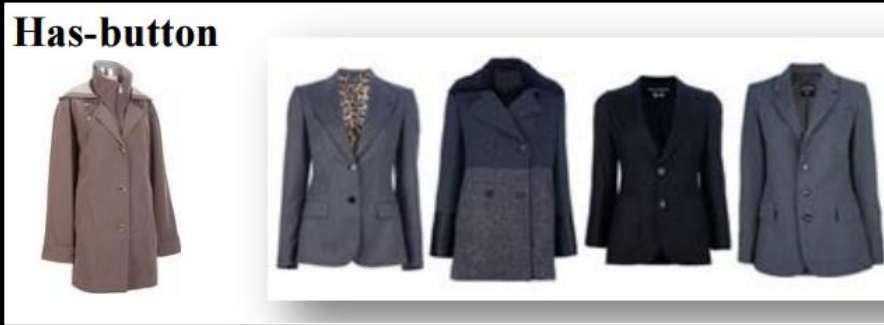
More effective representation



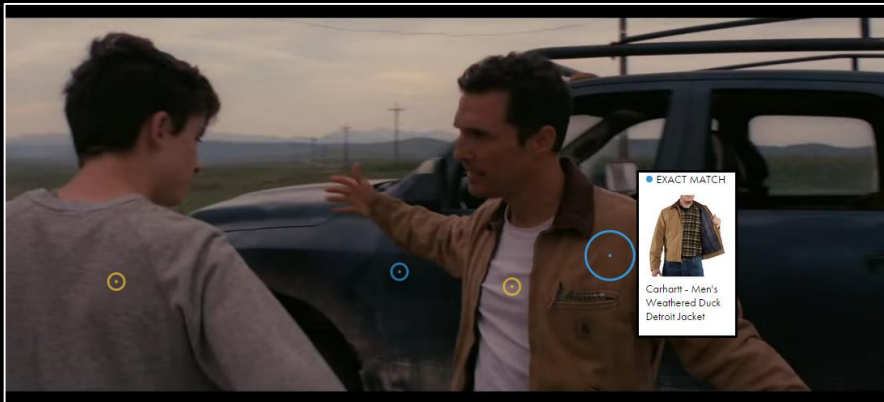


# Clothes Recognition

Think about the ultimate goal



Similar Style Retrieval



Cloth Spotting in Video



Street-to-shop



Fashion Assistant

# Clothes Recognition

The interplay between identities and attributes

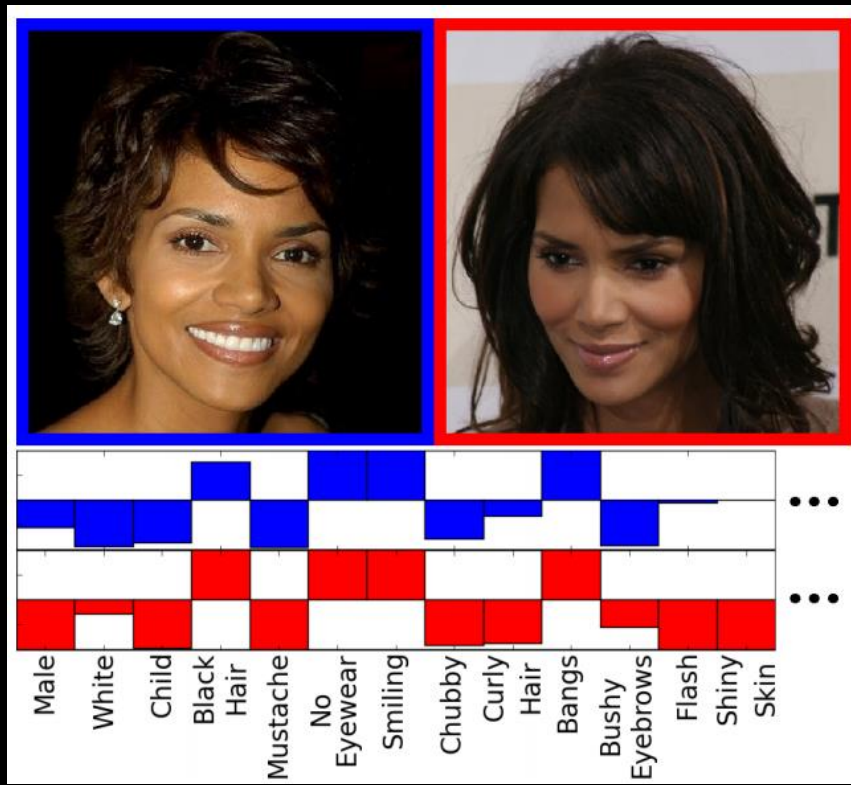


PID: 2000077658 (Forever21)

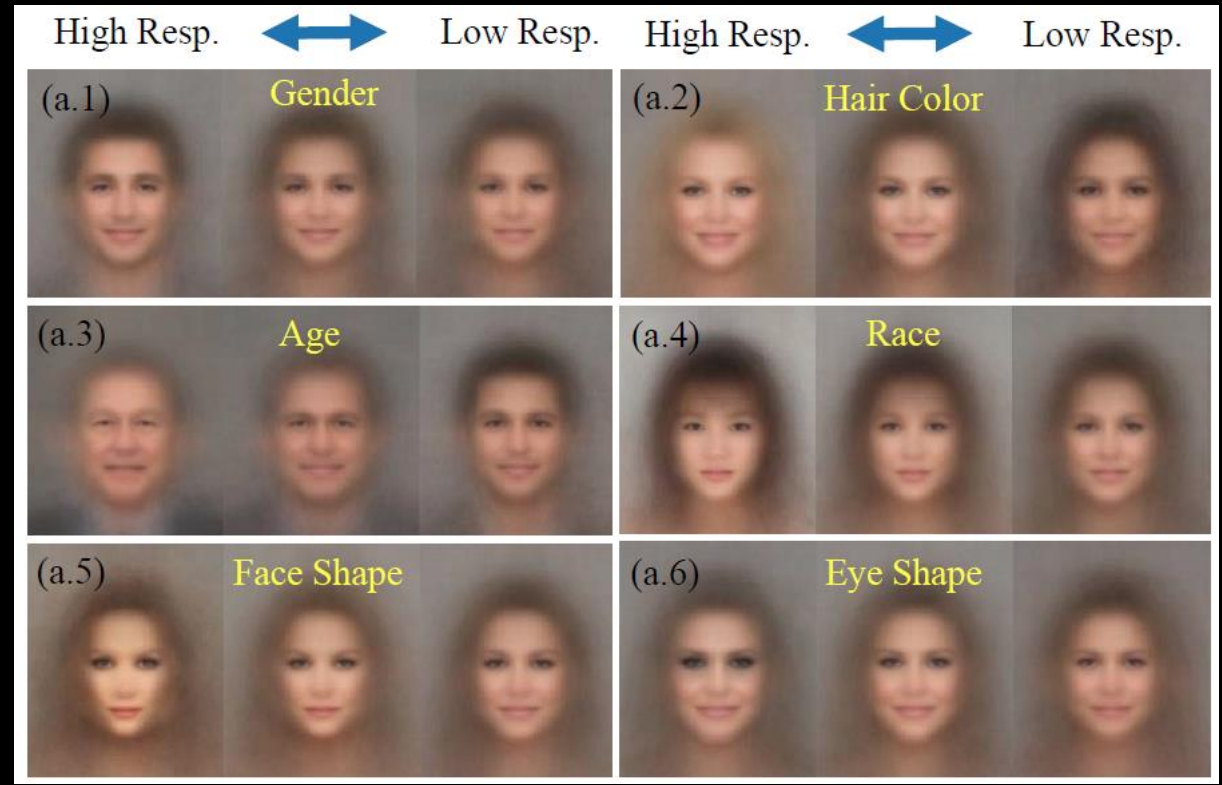
Ringer Tee (WOMEN)

# Clothes Recognition

The interplay between identities and attributes



Attributes facilitate identification.

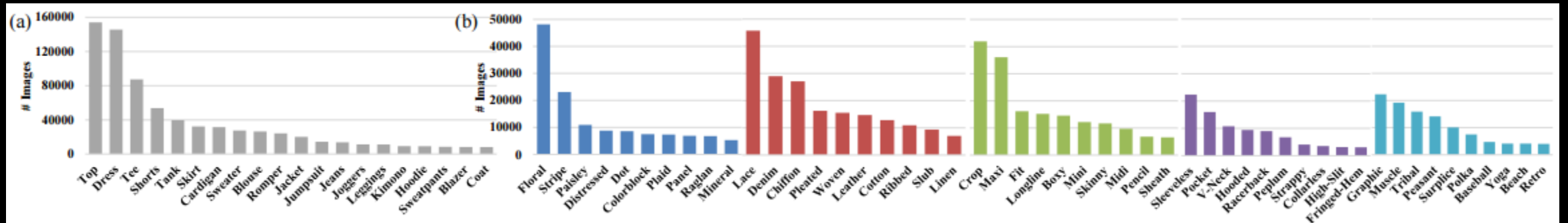


Identification discovers attributes.



# Clothes Recognition

Attributes are noisy and imbalanced

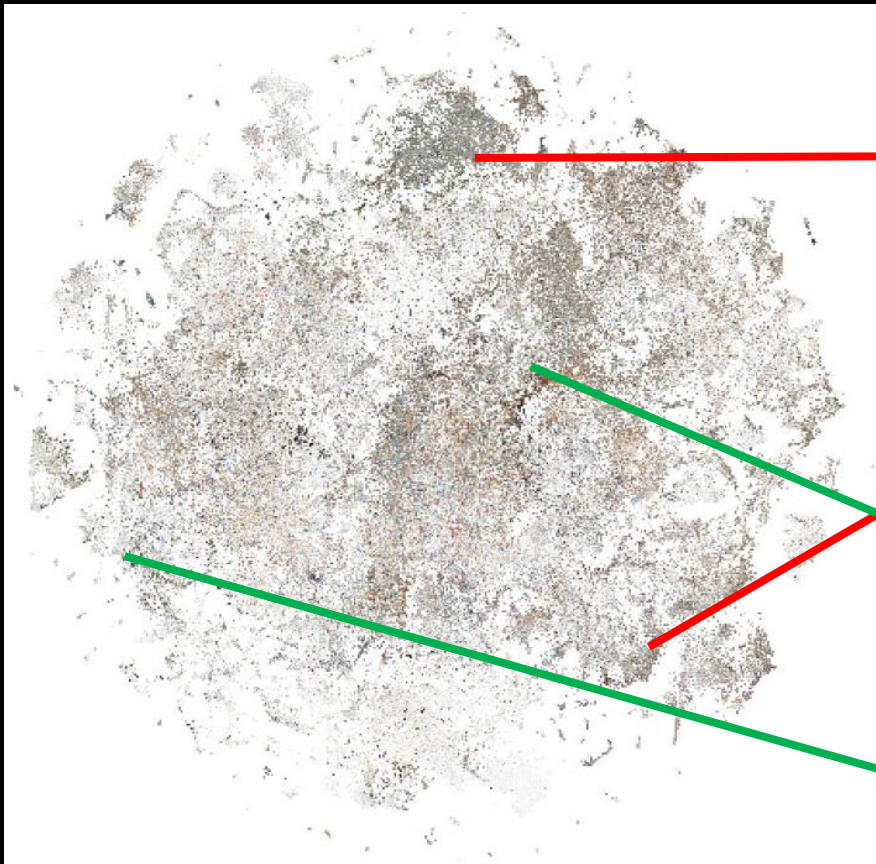


$$J = \sum_{i=1}^n \sum_{j=1}^{c_+} \sum_{k=1}^{c_-} \max(0, 1 - f_j(\mathbf{x}_i) + f_k(\mathbf{x}_i))$$

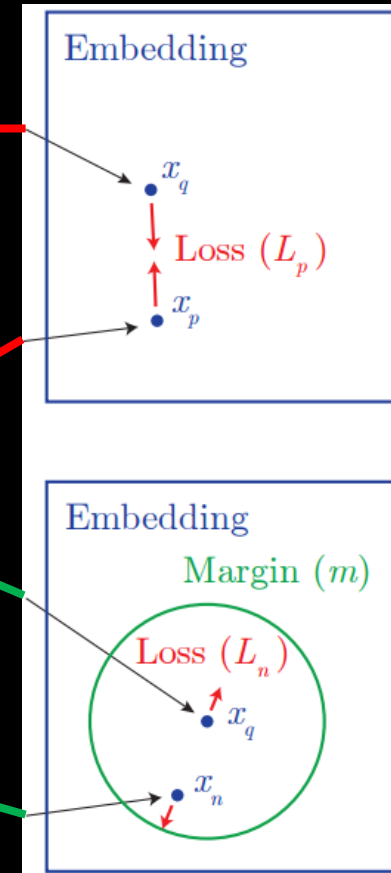
Multi-label Ranking Loss

# Clothes Recognition

The number of identities are huge



Millions of fashion identities



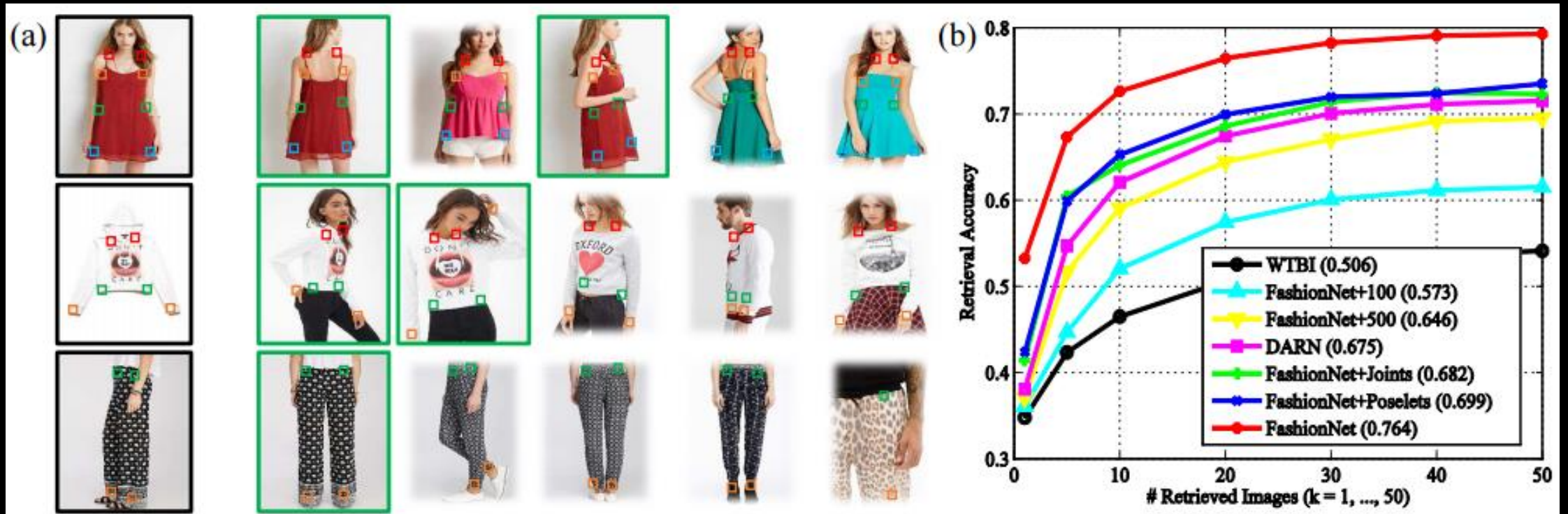
Positive Pair

Negative Pair

Hard Negative Mining

# Clothes Recognition

## In-shop Clothes Retrieval





# Clothes Recognition

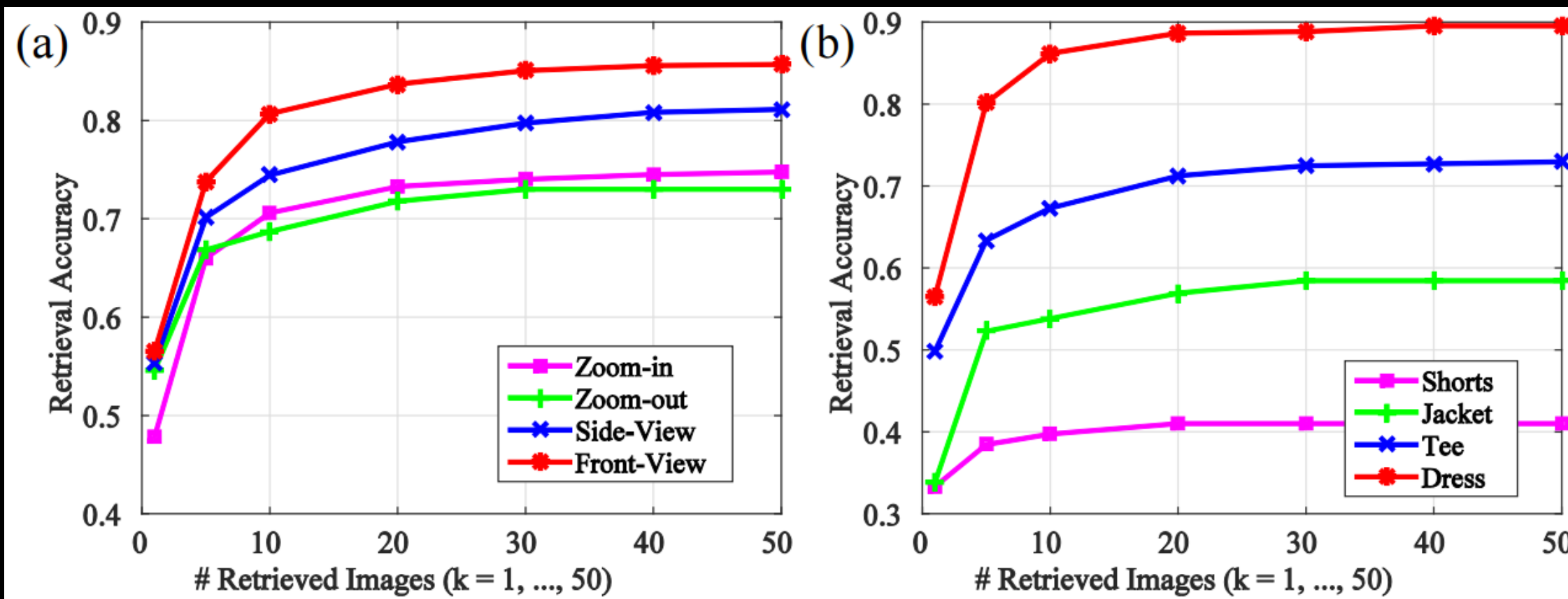
## Consumer-to-shop Clothes Retrieval





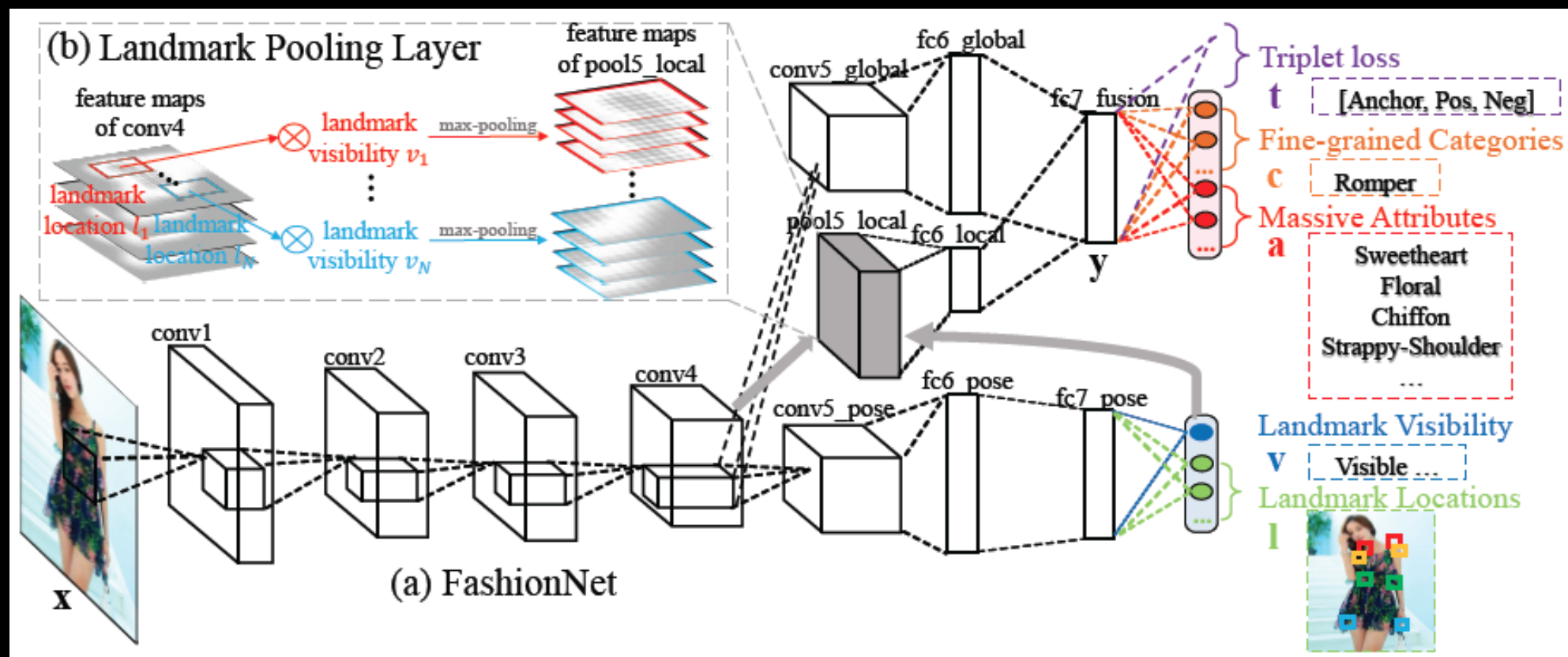
# Clothes Recognition

## Further Analysis



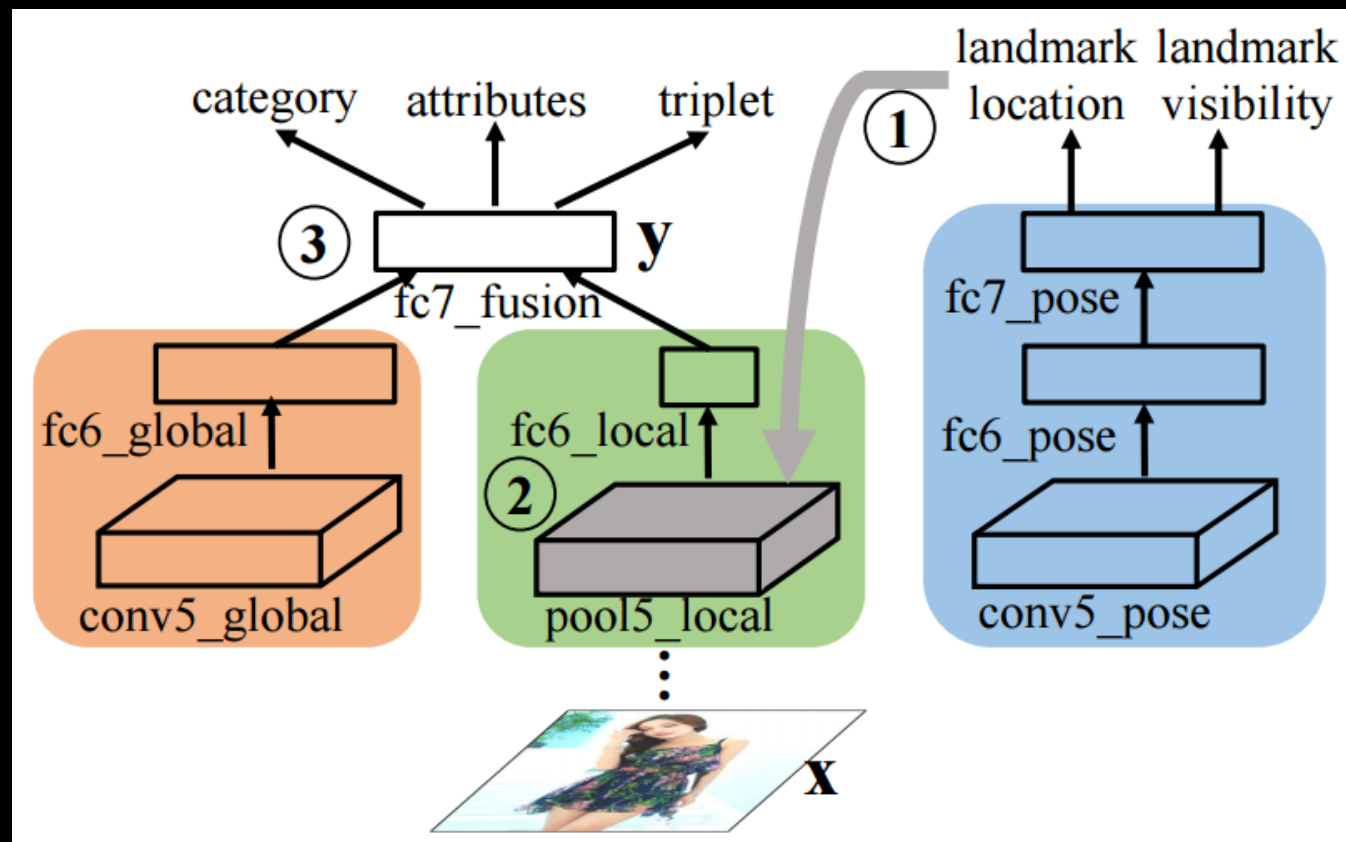
# FashionNet

## End-to-end System



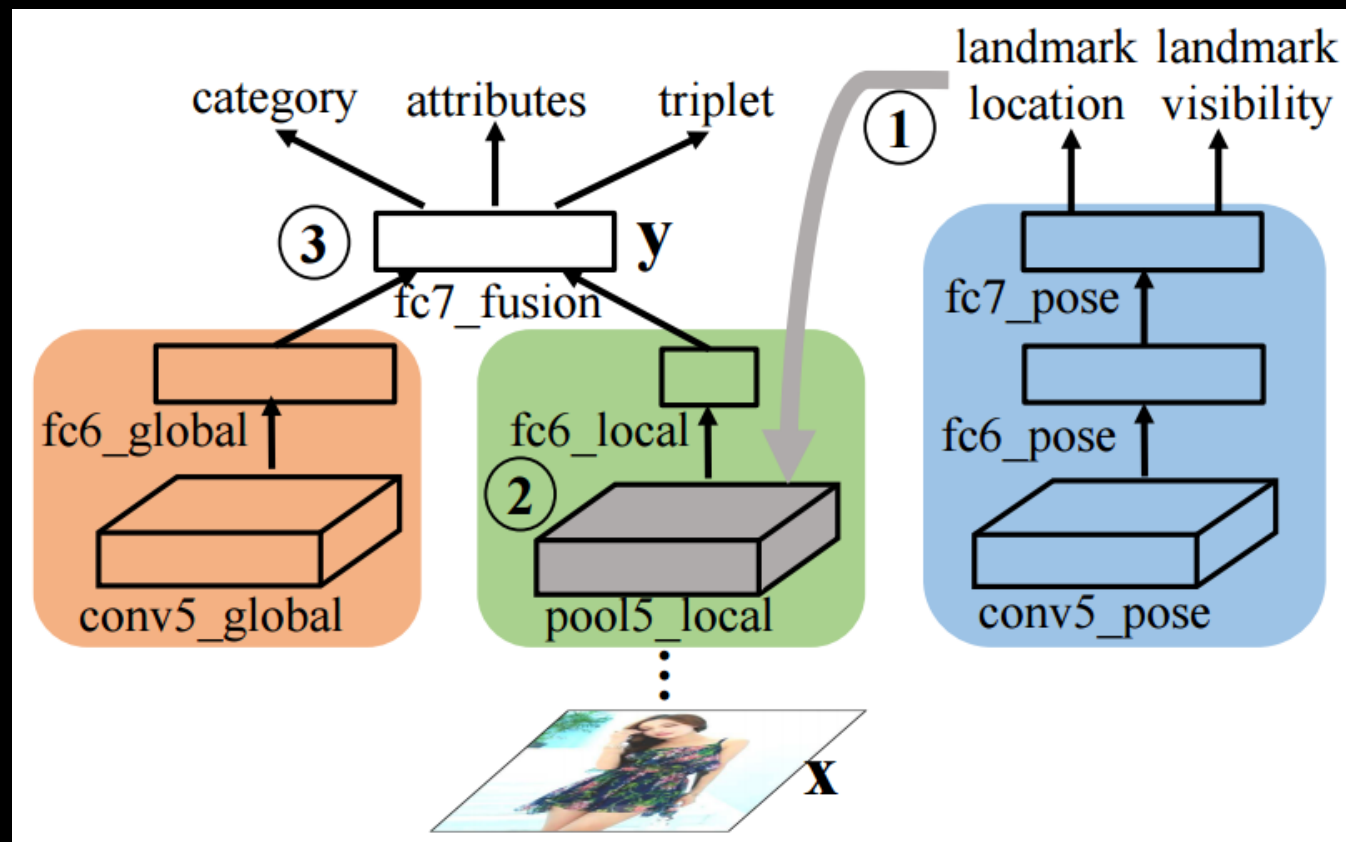
# FashionNet

## Forward Pass



# FashionNet

## Backward Pass





# Conclusions

- Large-scale Fashion Dataset DeepFashion
- Clothes Alignment by Fashion Landmarks
- End-to-end System with Heterogeneous Supervisions

# Future Work

- From “detection + alignment” to “parsing”
- From “pre-defined attributes” to “free-form descriptions”
- From “single clothes modeling” to “outfit understanding”

# Collaborators



Sijie Yan



Shi Qiu



Ping Luo



Xiaogang Wang



Xiaoou Tang

# Thanks!

DeepFashion Project by MMLAB, CUHK

Project Page: <http://mmlab.ie.cuhk.edu.hk/projects/DeepFashion.html>