DeepFashion: Powering Robust Clothes Recognition and Retrieval with Rich Annotations

Ziwei Liu[1]   Ping Luo[1]   Shi Qiu[2]   Xiaogang Wang[1]   Xiaoou Tang[1]

1. The Chinese University of Hong Kong   2. SenseTime Group Ltd.
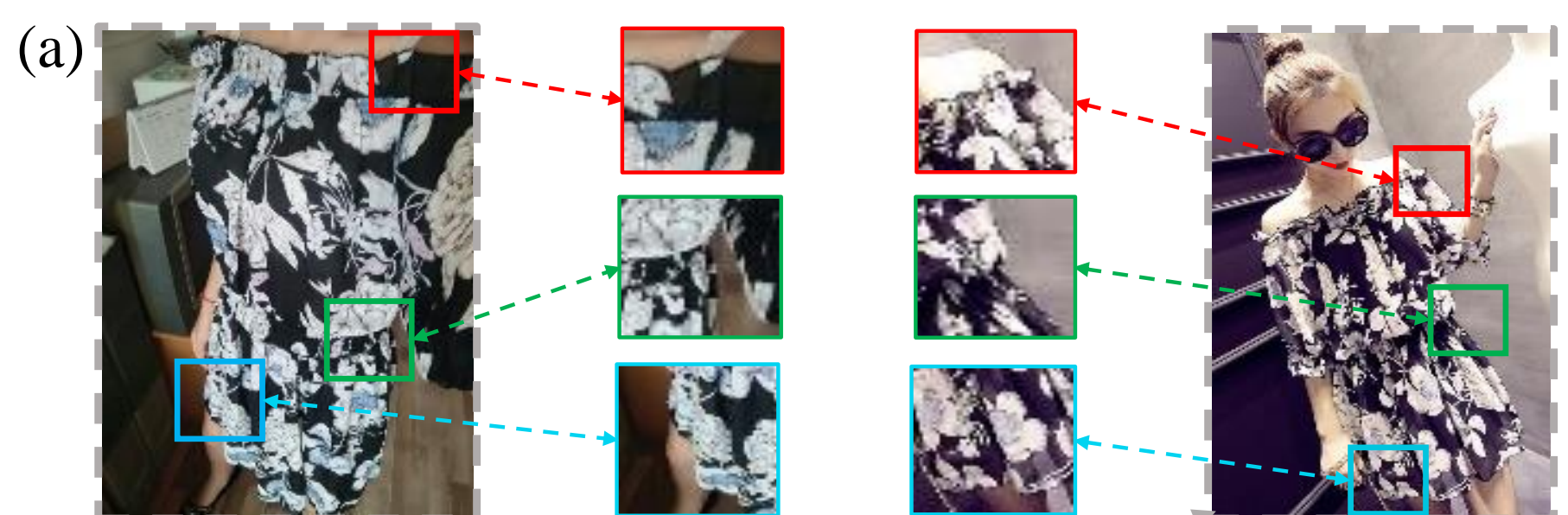
1. Motivation

Task：clothes recognition and retrieval

- **Landmarks** improve fine-grained recognition
- **Massive attributes** better partition feature space
- **Photo pairs** bridge the cross-domain gap

We provide

- Comprehensiveness: 50 categories, 1,000 attributes, 4~8 landmarks, 300K photo pairs
- Scale: over 800K real-life clothing images

	WTBI[1]	DARN[2]	**DeepFashion**
# image	78,958	182,780	**>800,000**
# attributes	11	179	**1050**
# pairs	39,479	91,390	**>300,000**
localization	bbox	N/A	**4~8 landmarks**

2. DeepFashion Dataset

Data Source

Search engines, online stores, user posts.

Quality Control

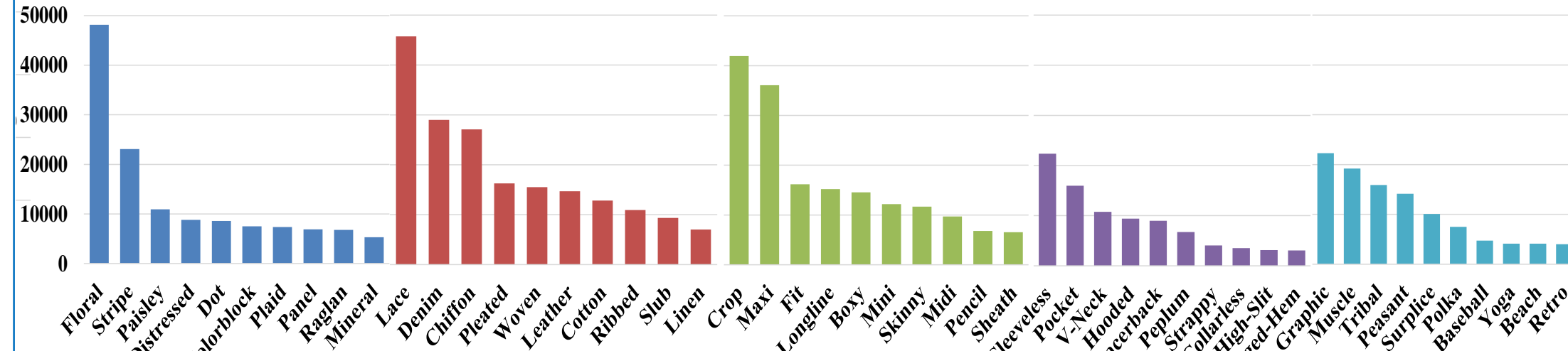Duplicate removal, fast screening, double checking

Annotation Assessment:

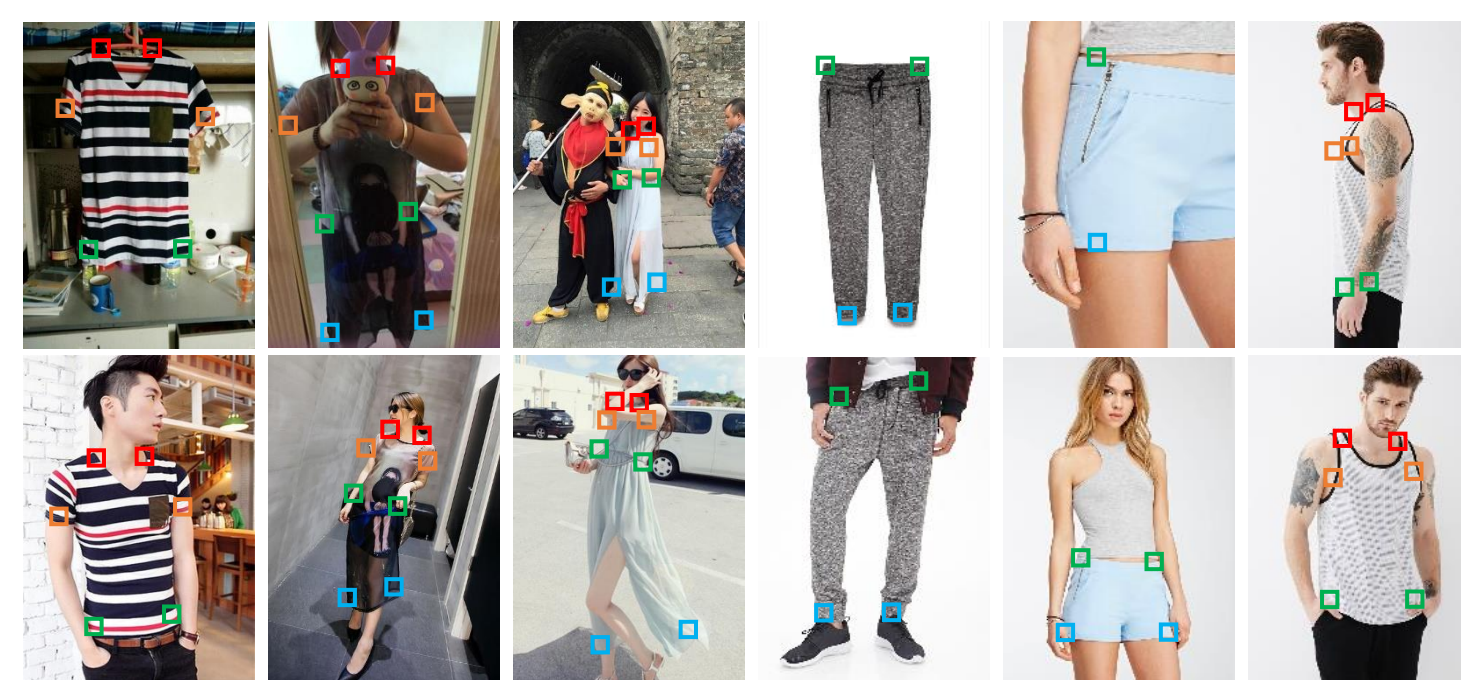| attribute	positive	negative
Label accuracies (%)	97.0	99.4

Sample Images

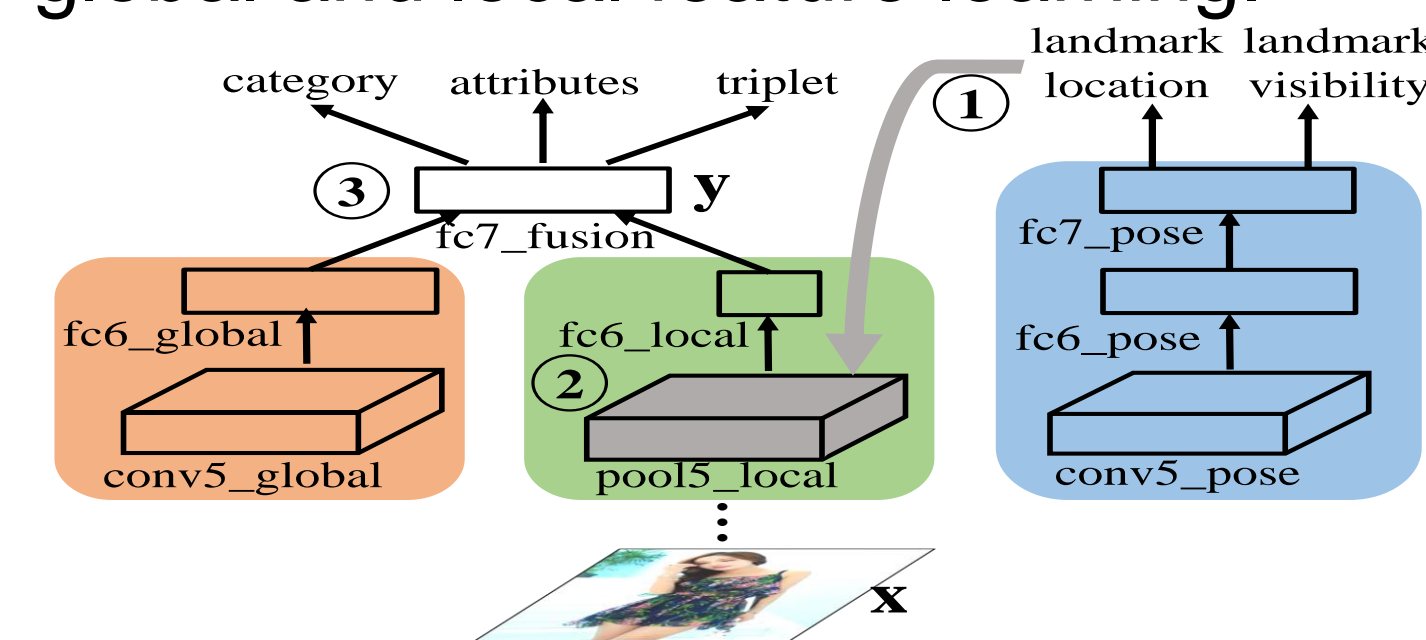### Attributes Statistics

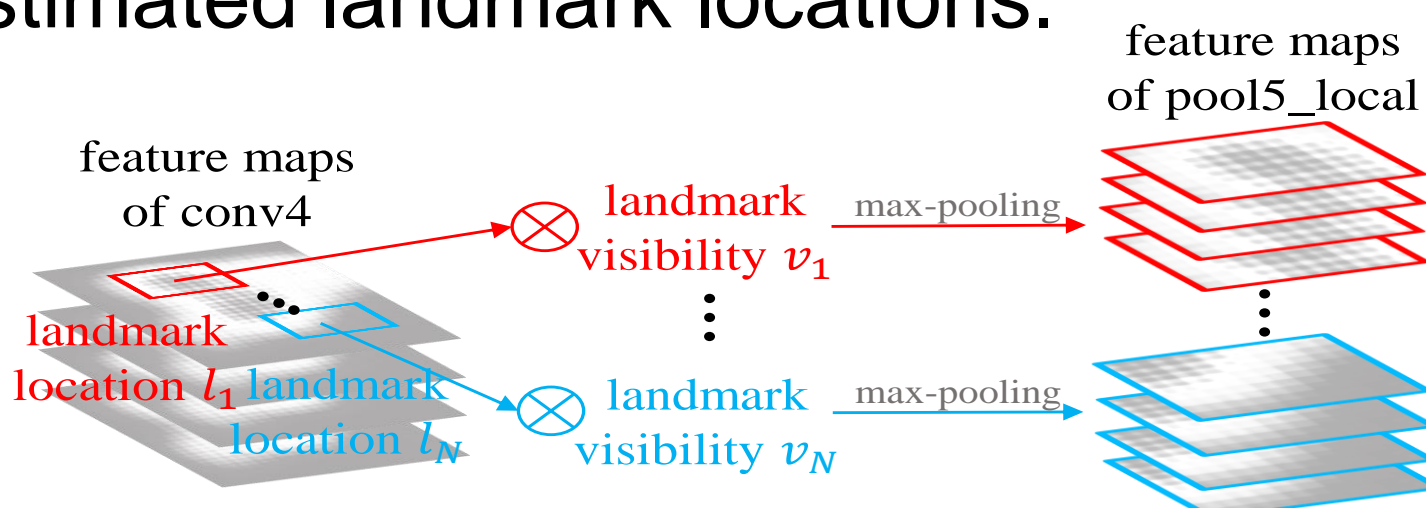### Landmarks and Pairs

3. FashionNet

Network Architecture

FashionNet jointly predicts landmarks and attributes to unify global and local feature learning.

Landmark Pooling Layer

Landmark pooling layer pools and gates features from estimated landmark locations.
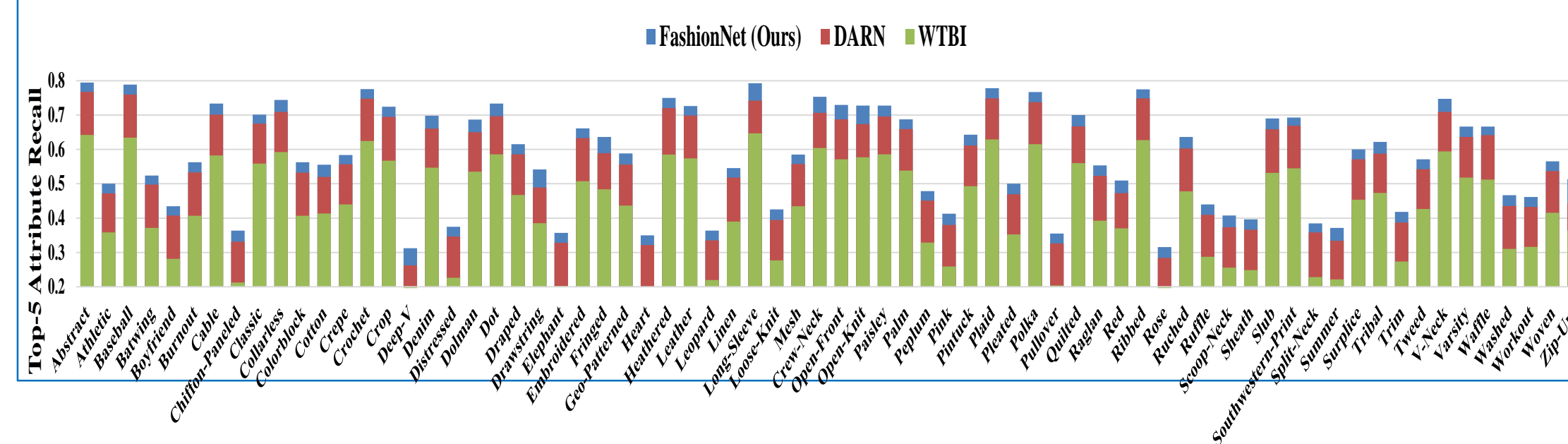

Multi-task Learning

Cross-entropy loss for attributes, Euclidean loss for landmarks, triplet loss for pairs.

4. Benchmarks

Category & Attribute Prediction

Metric: top-3 recall rate

	Categories (%)	Attributes (%)
WTBI[1]	43.73	27.46
DARN[2]	59.48	42.35
FashionNet	**82.58**	**45.52**

In-shop Clothes Retrieval

Metric: top-k retrieval accuracy

WTBI (0.506)
FashionNet+100 (0.573)
FashionNet+500 (0.646)
DARN (0.675)
FashionNet+Joints (0.682)
FashionNet+Poselets (0.699)
FashionNet (0.764)
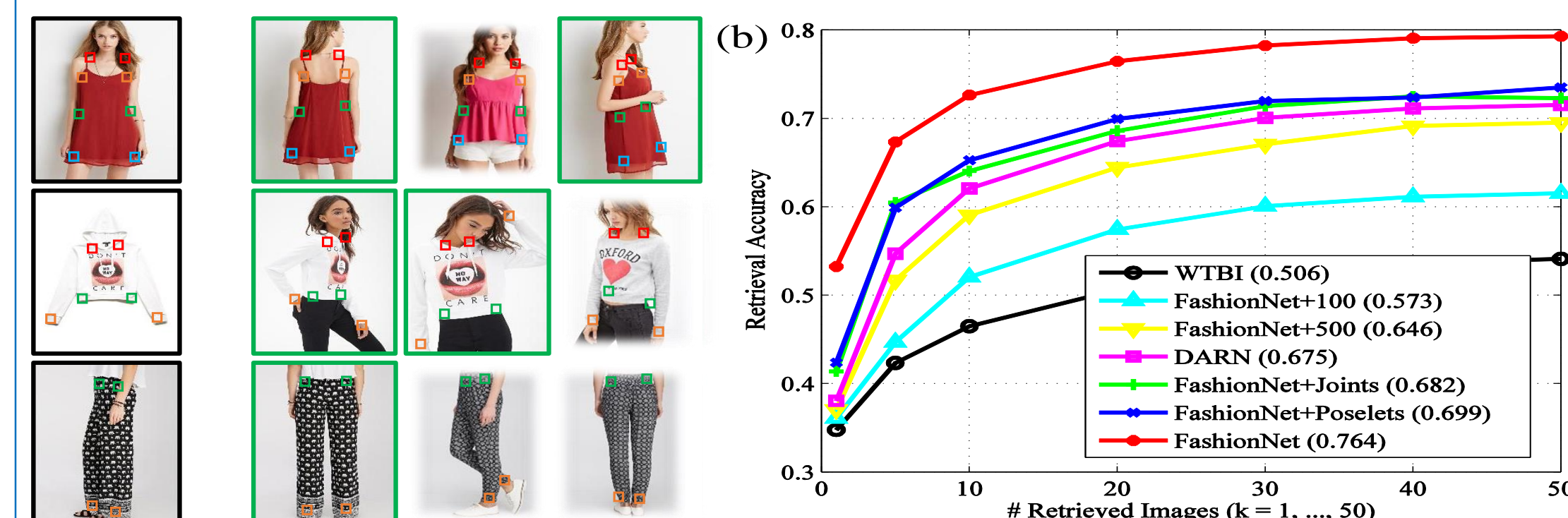
Consumer-to-shop Clothes Retrieval

Metric: top-k retrieval accuracy

WTBI (0.063)
FashionNet+100 (0.071)
FashionNet+500 (0.104)
DARN (0.111)
FashionNet+Joints (0.136)
FashionNet+Poselets (0.139)
FashionNet (0.188)
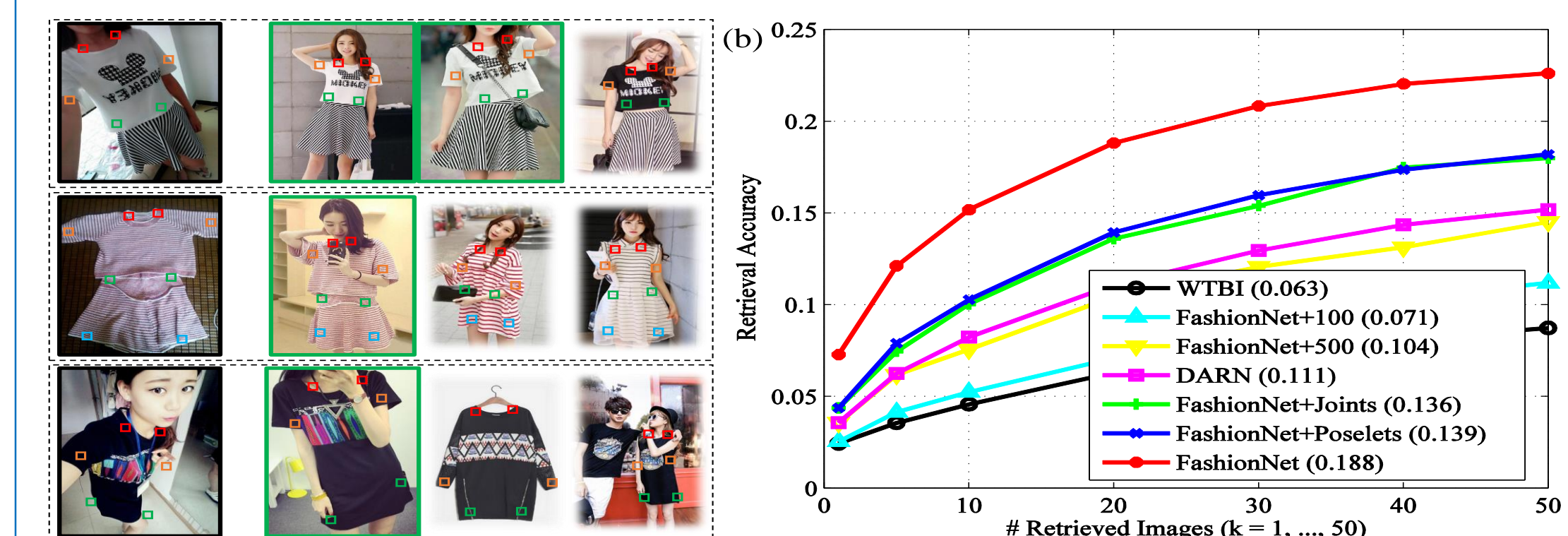
Further Analysis

How different variations affect performance?
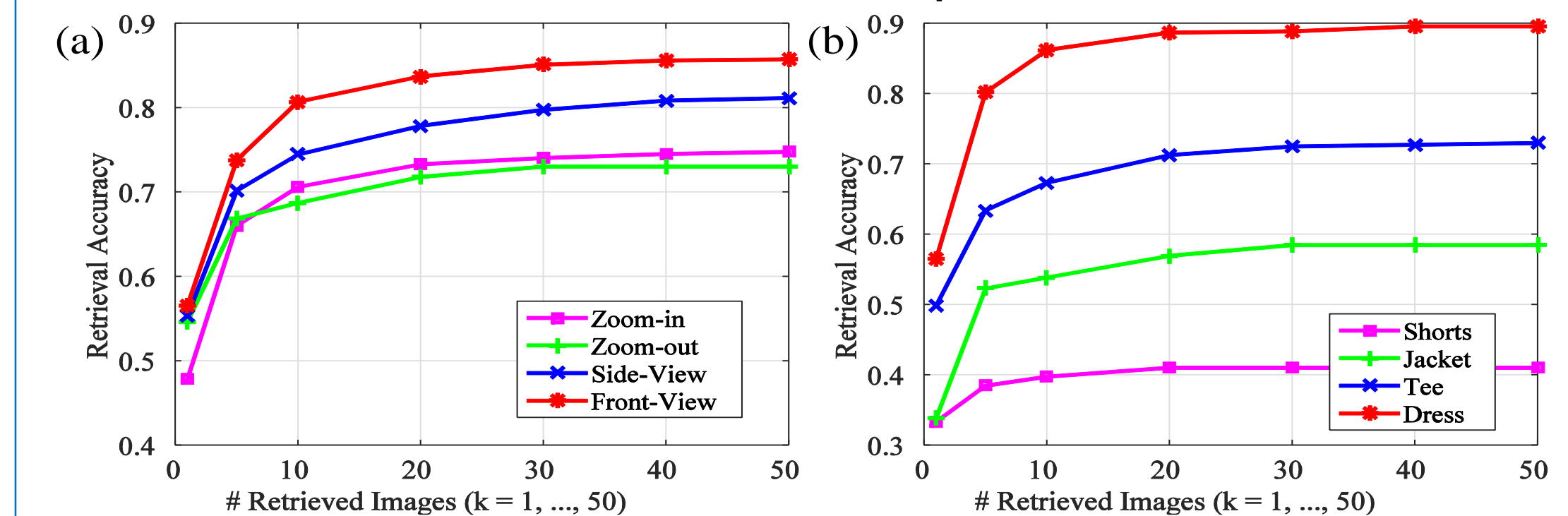
References

[1] M. H. Kiapour, et al. Where to buy it: Matching street clothing photos in online shops. In ICCV, 2015.
[2] J. Huang, et al. Cross-domain image retrieval with a dual attribute-aware ranking network. In ICCV, 2015