



Fashion is my profession



Robots can't understand fashion, they told us, but times have changed...

Nowadays, because of using the machine learning approaches, humans are in danger, in danger to be assessed as an outmoded.

To be protected everybody should know about state of the art approaches, current challenges and tasks for AI in the fashion domain and how robots can improve your look.

And remember, they are closer than you think.





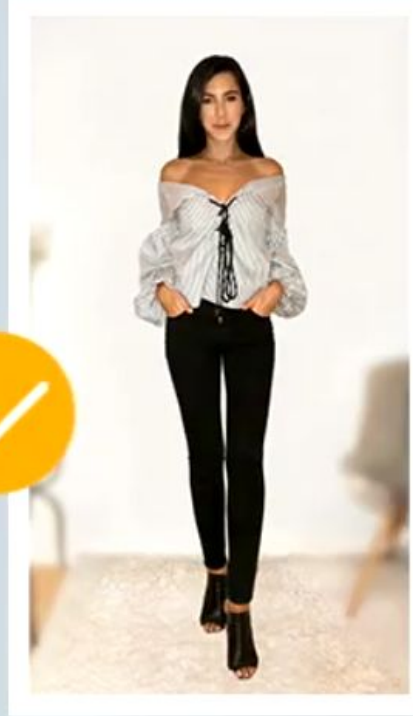
Echo Look | Hands-Free Camera and Style Assistant with Alexa—including Style Check to get a second opinion on your outfit



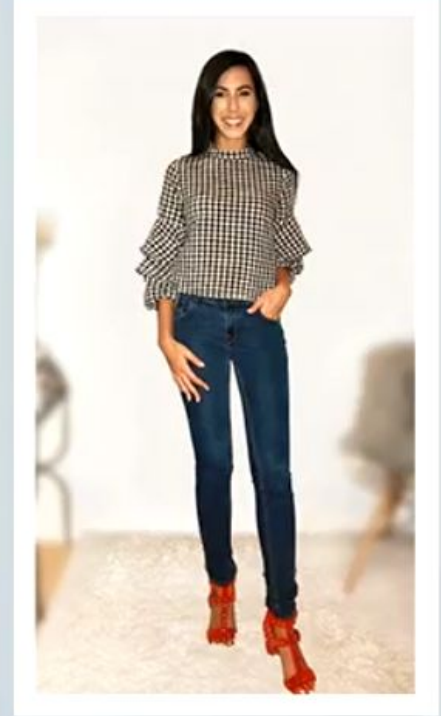
<https://www.amazon.com/Amazon-Echo-Look-Camera-Style-Assistant/dp/B0186JAEWK>

INTRODUCING STYLE CHECK

Style Check keeps your look on point using advanced machine learning algorithms and advice from fashion specialists. Submit two photos for a second opinion on which outfit looks better on you, and find out why your outfit was selected with feedback based on fit, color, styling, and current trends. Over time, these decisions get smarter through your feedback and input from our team of experienced fashion specialists.



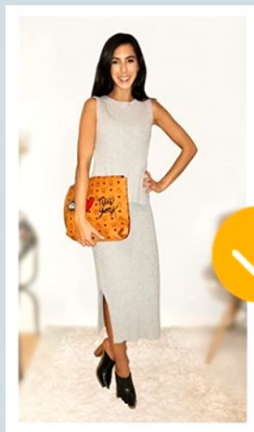
64%



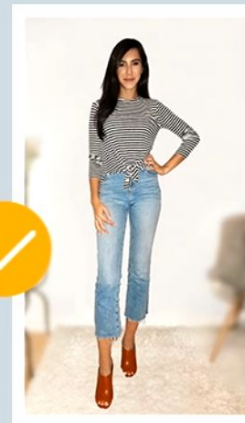
36%



45%



55%



72%



28%

AI meets Fashion:

- Semantic Segmentation | Clothing parsing;
- Object retrieval | Retrieval for e-commerce;
- Classification | Brand or attribute discovery;
- Recommendations and style detection;
- Image synthesis | Virtual Try-on;
- Caption synthesis | Friendly commenting;
- Pose estimation;

Semantic Segmentation | Clothing parsing

Fashionista Dataset

The directory with dataset itself **contains 685 annotated pictures** with good visibility of the full body and covering a variety of clothing items.



- null
- shoes
- shirt
- jeans
- hair
- skin



- null
- tights
- jacket
- dress
- hat
- heels
- hair
- skin

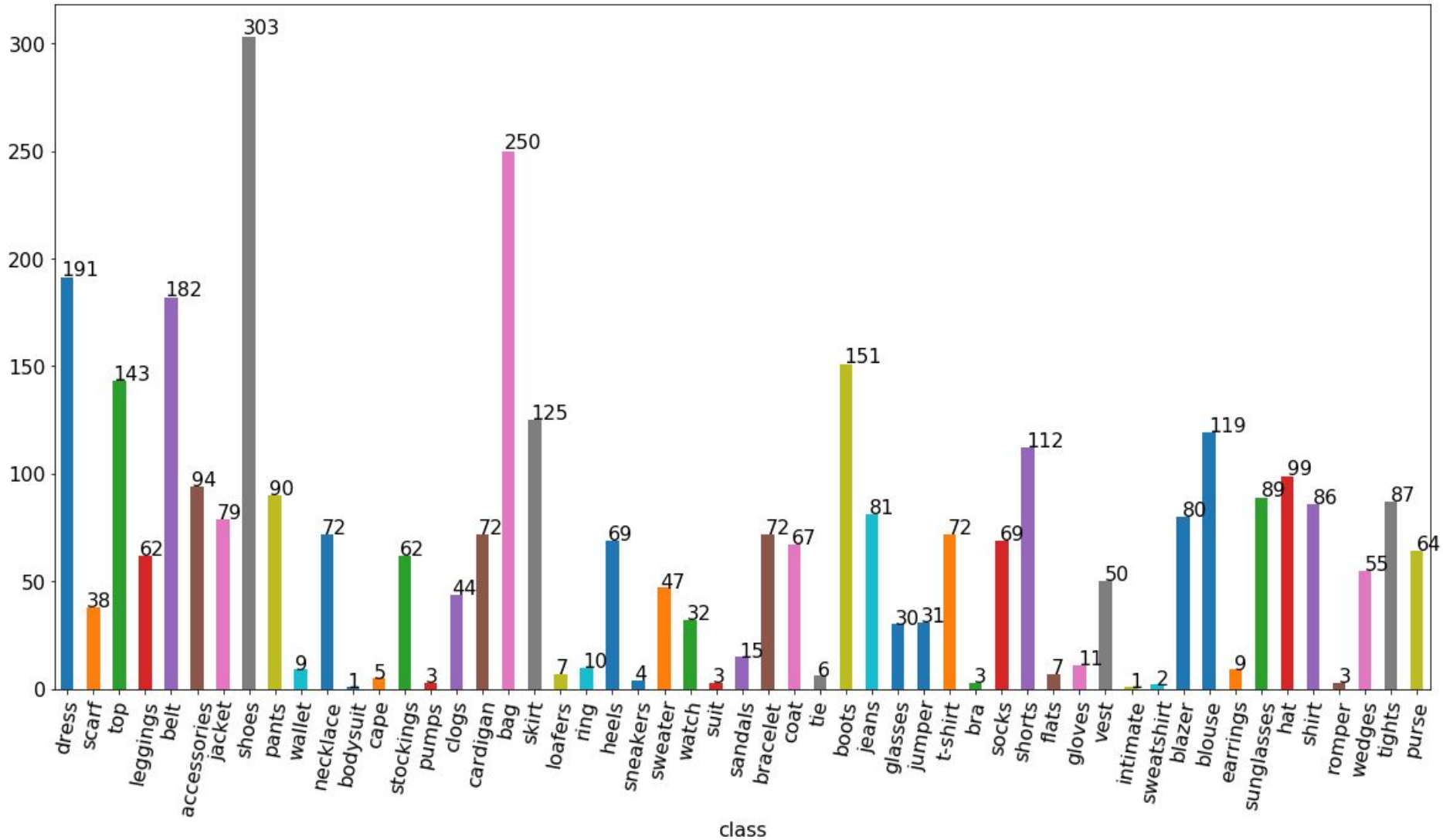


- null
- shorts
- blouse
- bracelet
- wedges
- hair
- skin



- null
- shoes
- top
- stockings
- hair
- skin

Fashionista Dataset



Total of 56 different possible clothing labels.

Fashionista Dataset

Demo from authors:

http://vision.is.tohoku.ac.jp/clothing_parsing

<http://clothingparsing.com/>

Input image



Final result



Colorful Fashion Parsing Data (cfpd dataset)

Fashion image dataset called Colorful-Fashion, in which all **2682 images** are labeled with pixel-level color-category labels. There are 23 categories in the dataset.

Images:



...

Color Labels:



...

Category Labels:



Cloting CoParsing (ccp dataset)



- null
- bag
- blouse
- cape
- hair
- loafers
- skin
- skirt
- stockings



- null
- bag
- boots
- coat
- hair
- skin
- skirt
- stockings
- sweater



- null
- bag
- blouse
- dress
- hair
- shoes
- skin



- null
- coat
- hair
- pants
- sandals
- skin
- t-shirt
- wallet



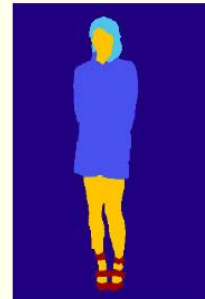
- null
- belt
- hair
- jacket
- jeans
- shirt
- shoes
- skin
- sunglasses
- tie



- null
- bag
- coat
- hair
- jeans
- shoes
- skin
- sunglasses
- t-shirt



- null
- accessories
- belt
- blouse
- hair
- hat
- purse
- shoes
- skin
- skirt



- null
- dress
- hair
- skin
- wedges

Totally 59 tags
with pixel-level
annotations.

Overall 1004 files

ATR dataset

7,700 images are included in the ATR dataset, 6,000 for training, 1,000 for testing and 700 for validation.

We use the Fashionista dataset after transforming the original labels to 18 categories.

We combine data from three small benchmark datasets:

Fashionista dataset containing 685 images

Colorful Fashion Parsing Data (CFPD) dataset containing 2,682 images

Daily Photos dataset containing 2,500 images.

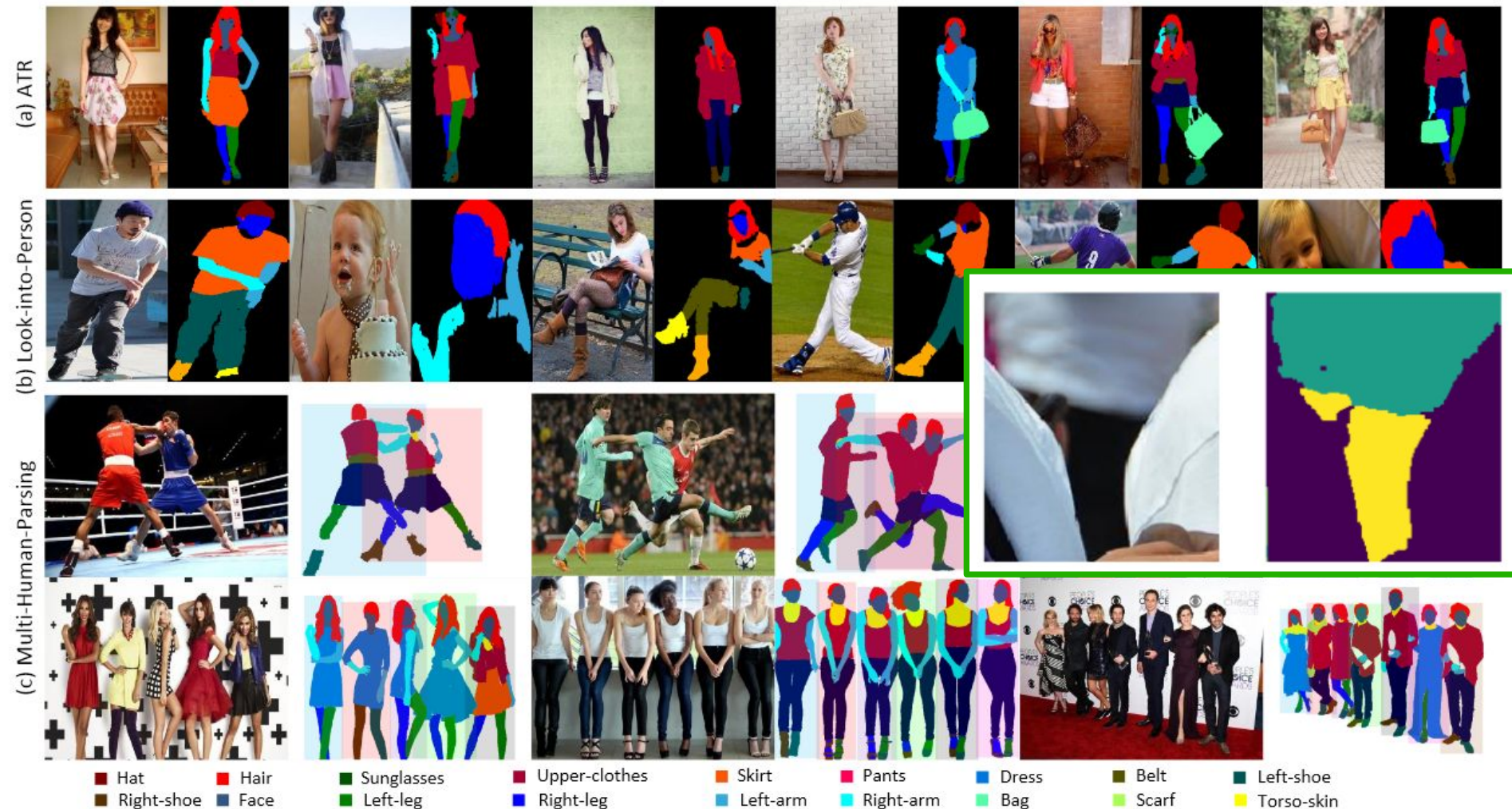
Chictopia10k

(could be a part of ATR)



Look into Person (LIP dataset)

In total, there are 50,462 images in the LIP dataset including **19,081 full-body images**, 13,672 upper-body images, 403 lower-body images, 3,386 head-missed images, 2,778 back-view images and 21,028 images with occlusions.



DeepFashion

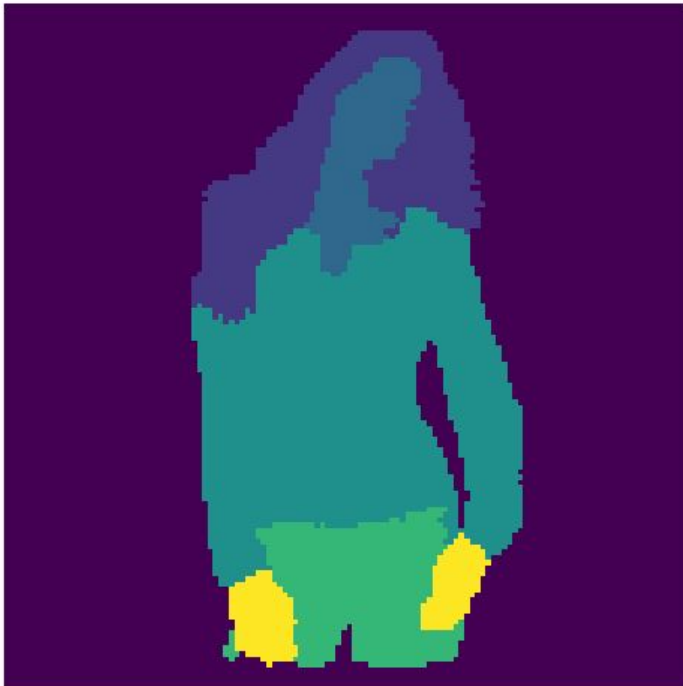
Largest clothing dataset to date, with over 800,000 diverse fashion images ranging from well-posed shop images to unconstrained consumer photos.



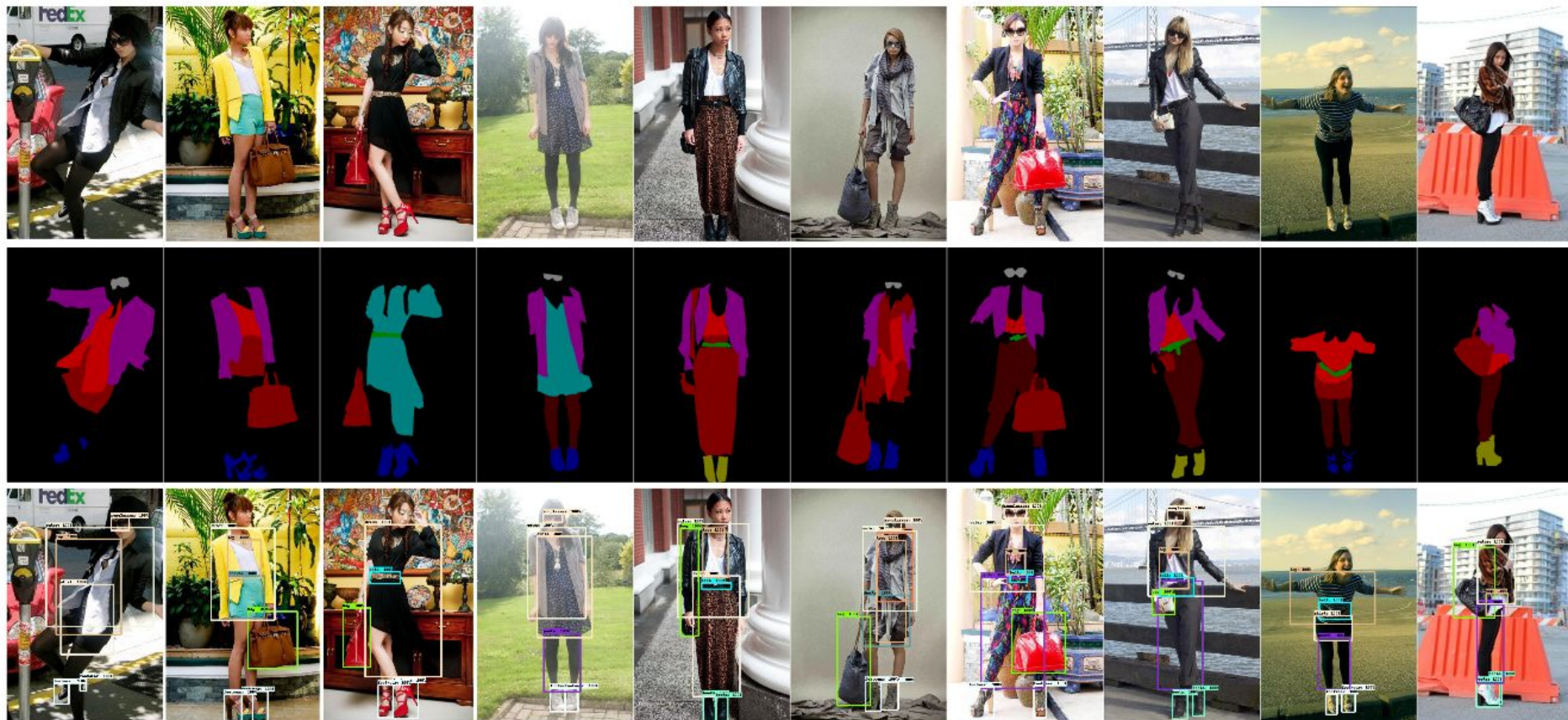
DeepFashion



Fashion Synthesis Benchmark facilitates the studies of generating new clothing images. It includes 78,979 images. Each image is associated with several sentences as captions and a **segmentation map**

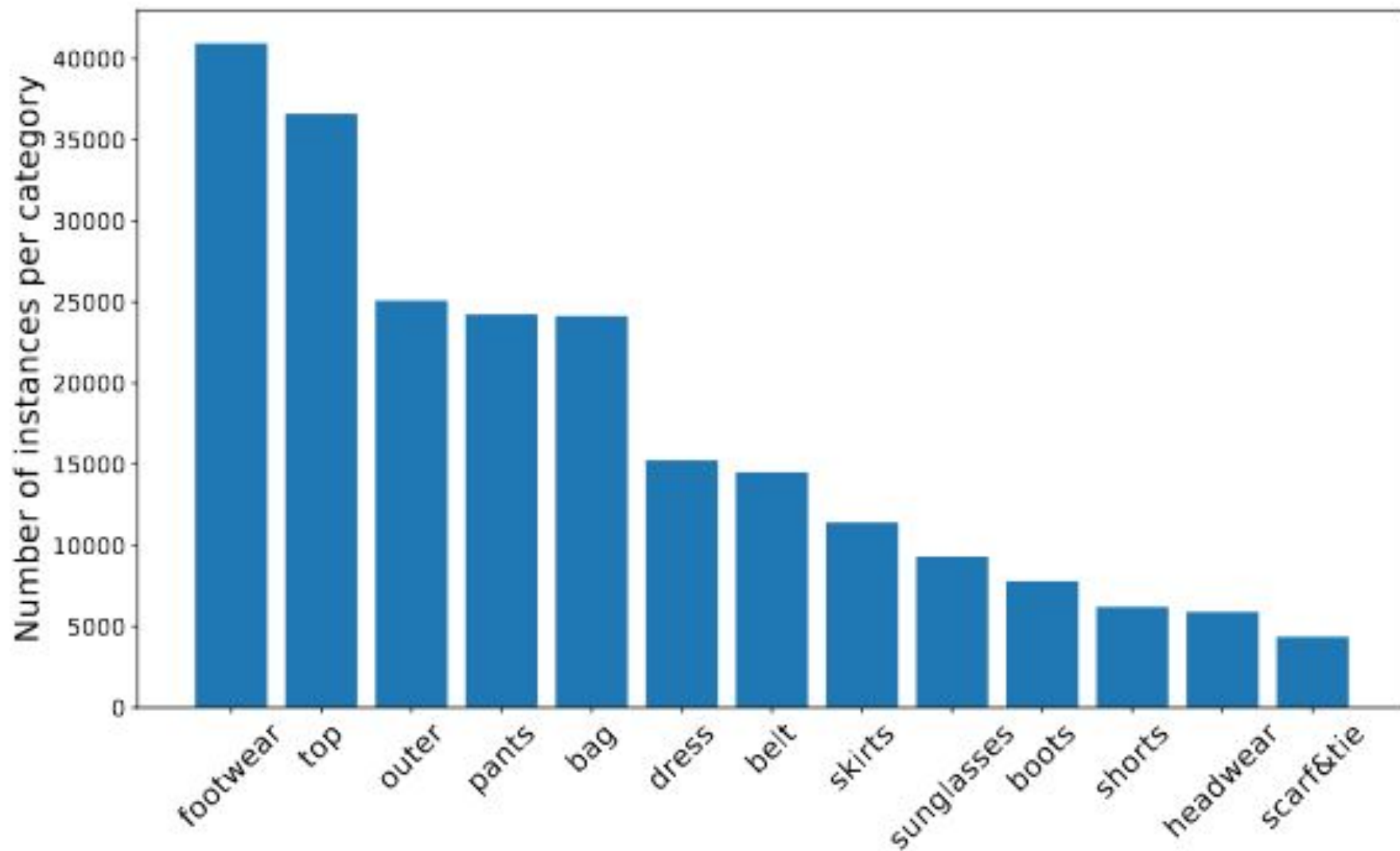


ModaNet

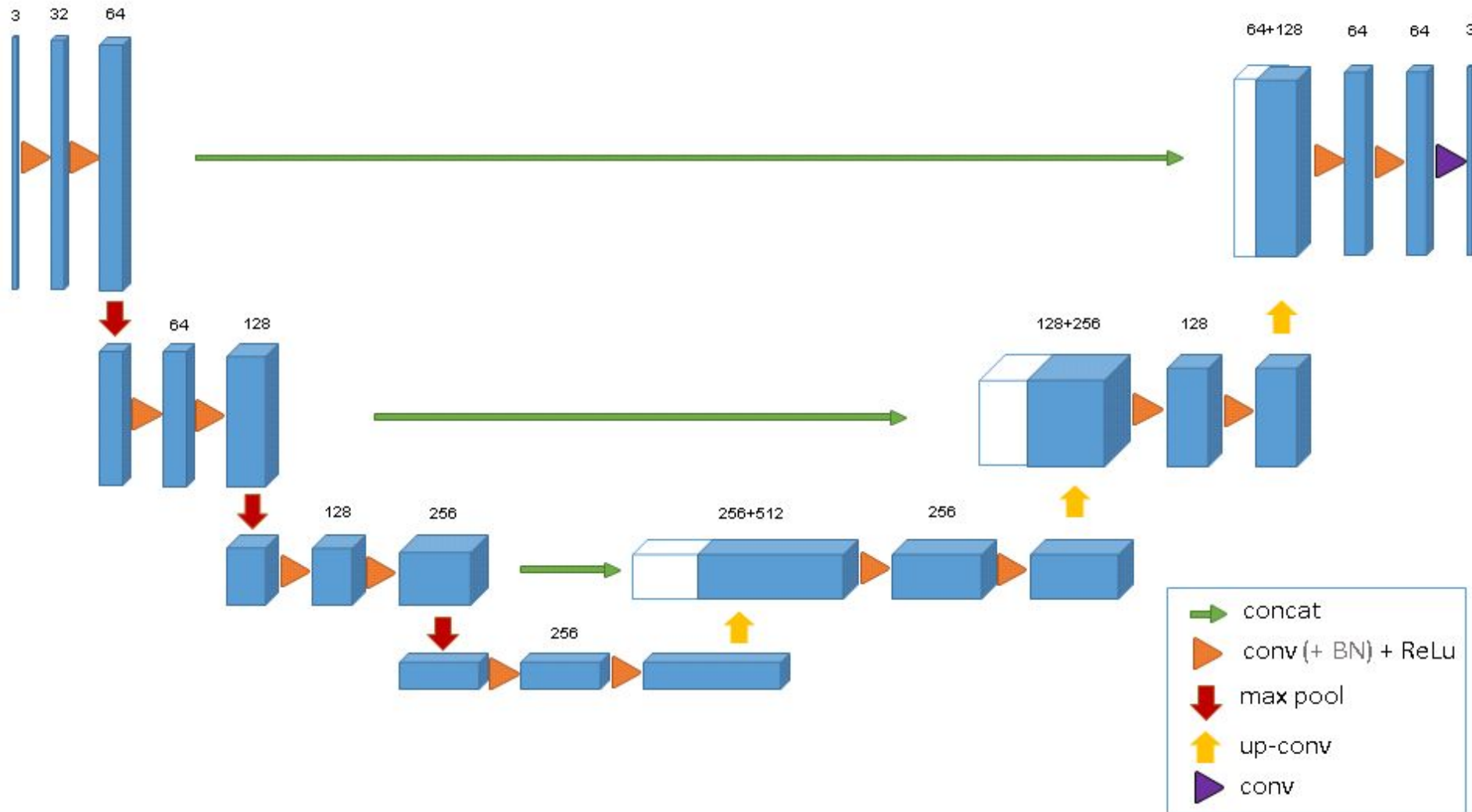


	DeepFashion [30]	CFPD [26]	CCP [42]	Fashionista [41]	HPW[20]	ModaNet
# of images	800,000	2,682	1,004	685	1,833	55,176
# of categories	50	19	56	53	11	13
Pixel annotation	×	✓	✓	✓	✓	✓
Bounding box	landmarks	✓*	✓*	✓*	✓*	✓
Polygon	×	×	×	×	×	✓

ModaNet

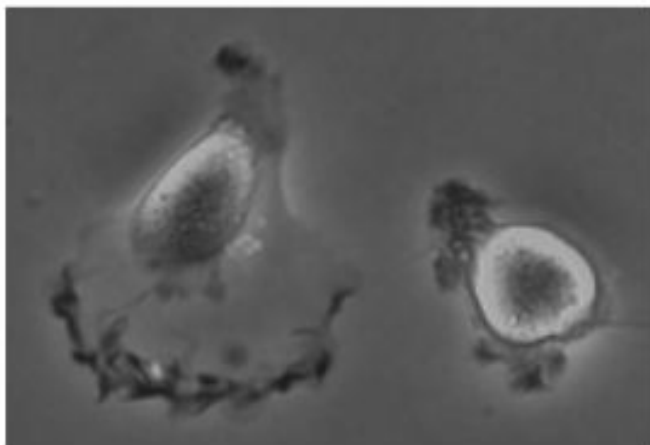


U-Net: Convolutional Networks for Biomedical Image Segmentation

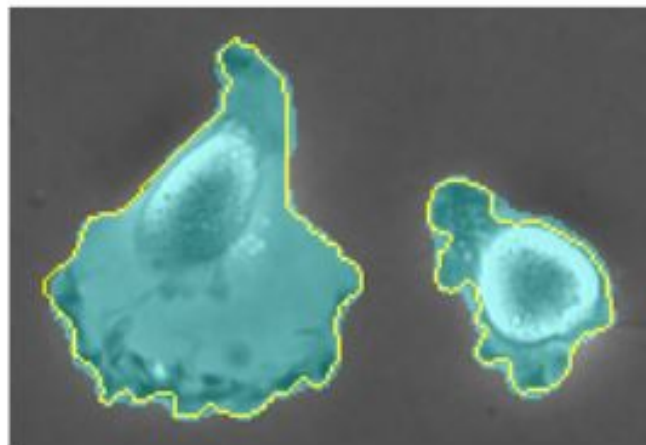


U-Net: Convolutional Networks for Biomedical Image Segmentation

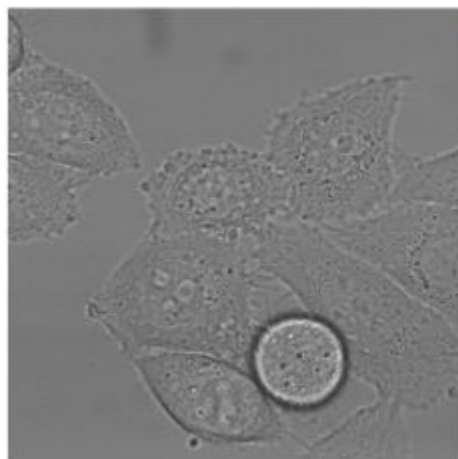
a



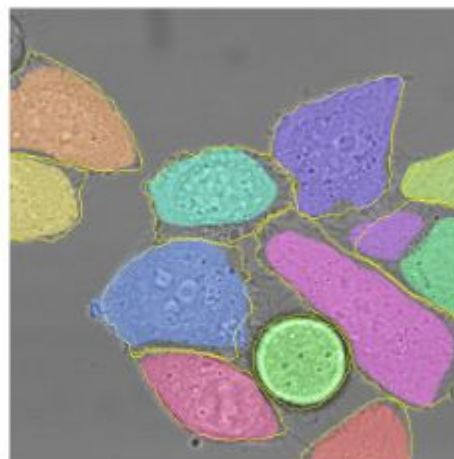
b



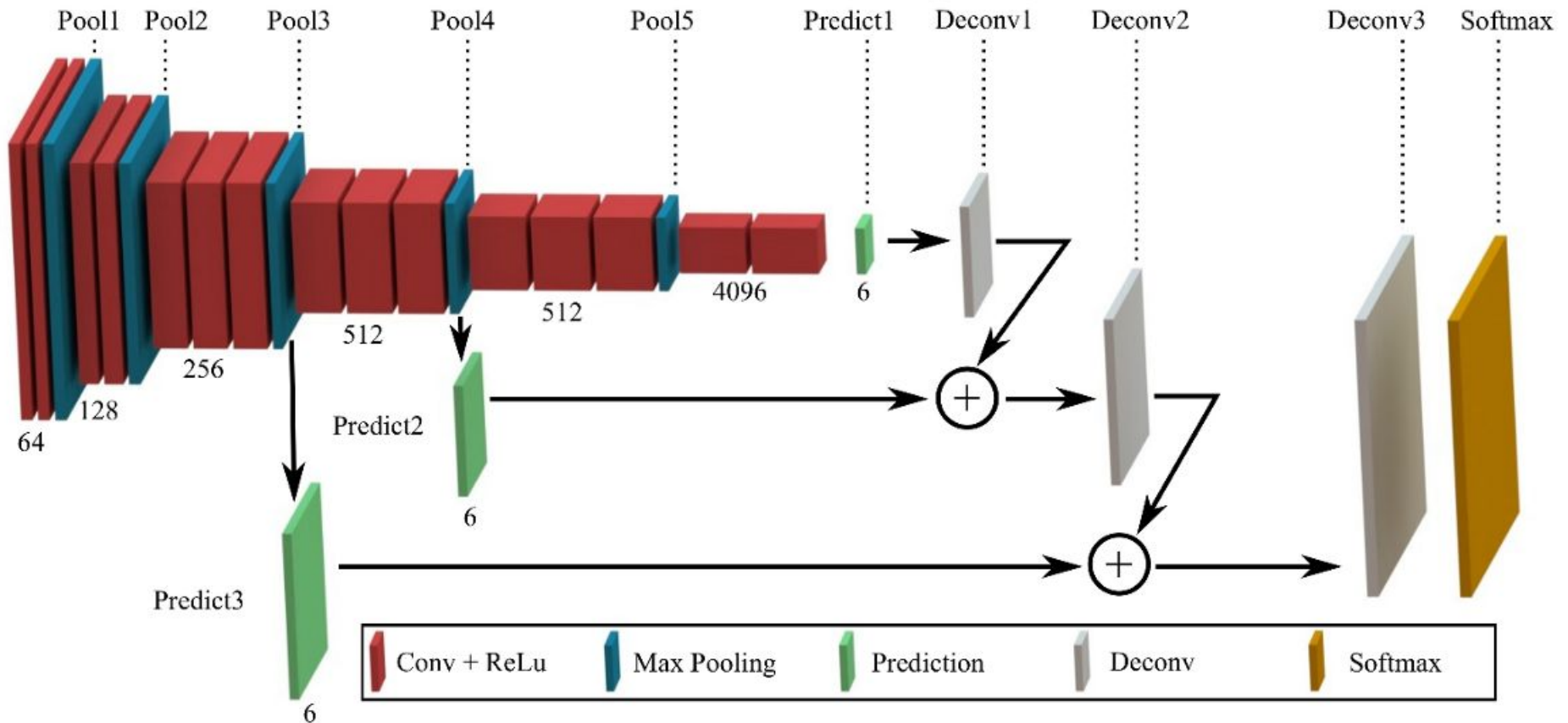
c



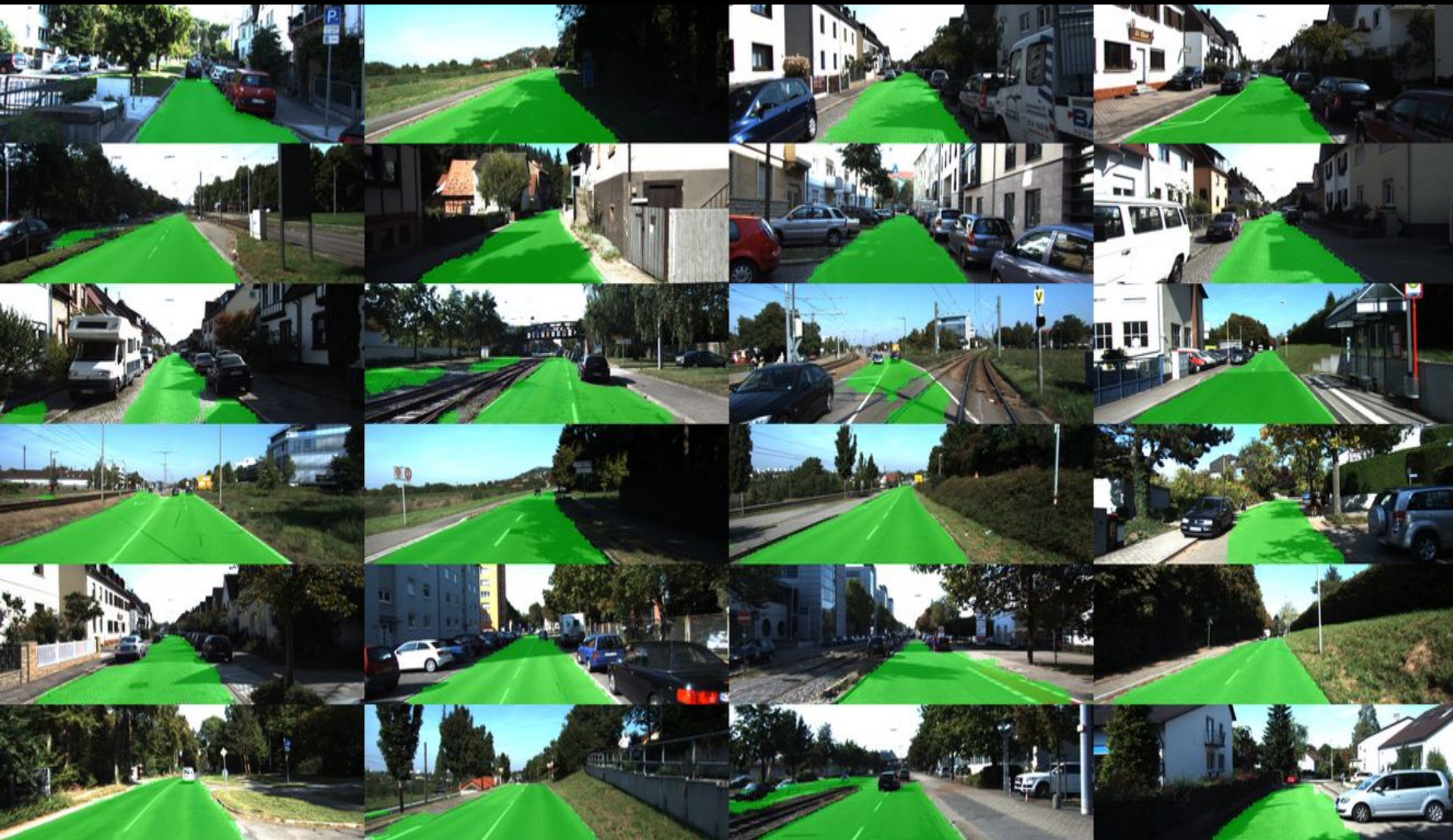
d



Fully Convolutional Networks for Semantic Segmentation



Fully Convolutional Networks for Semantic Segmentation



Looking at Outfit to Parse Clothing

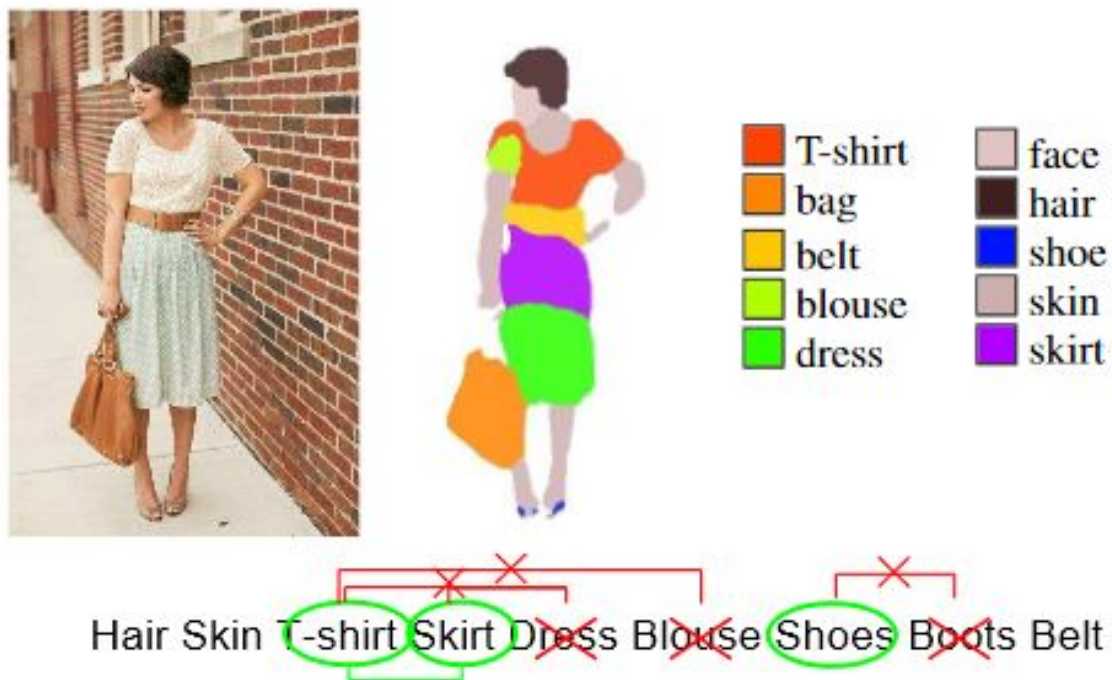


Figure 1: Combinatorial preference in clothing parsing: dress and skirt are in the exclusive relationship, yet independent pixel-wise prediction cannot encode such knowledge and results in mixture of patches (FCN-8s [22]). We propose the side-path outfit encoder and CRF alongside the segmentation pipeline to address the issue.

Looking at Outfit to Parse Clothing

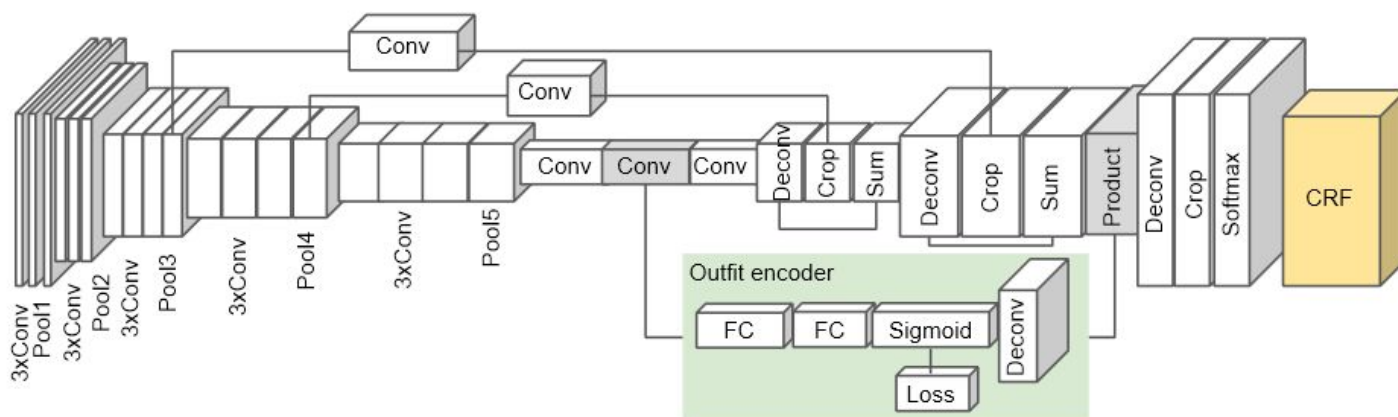
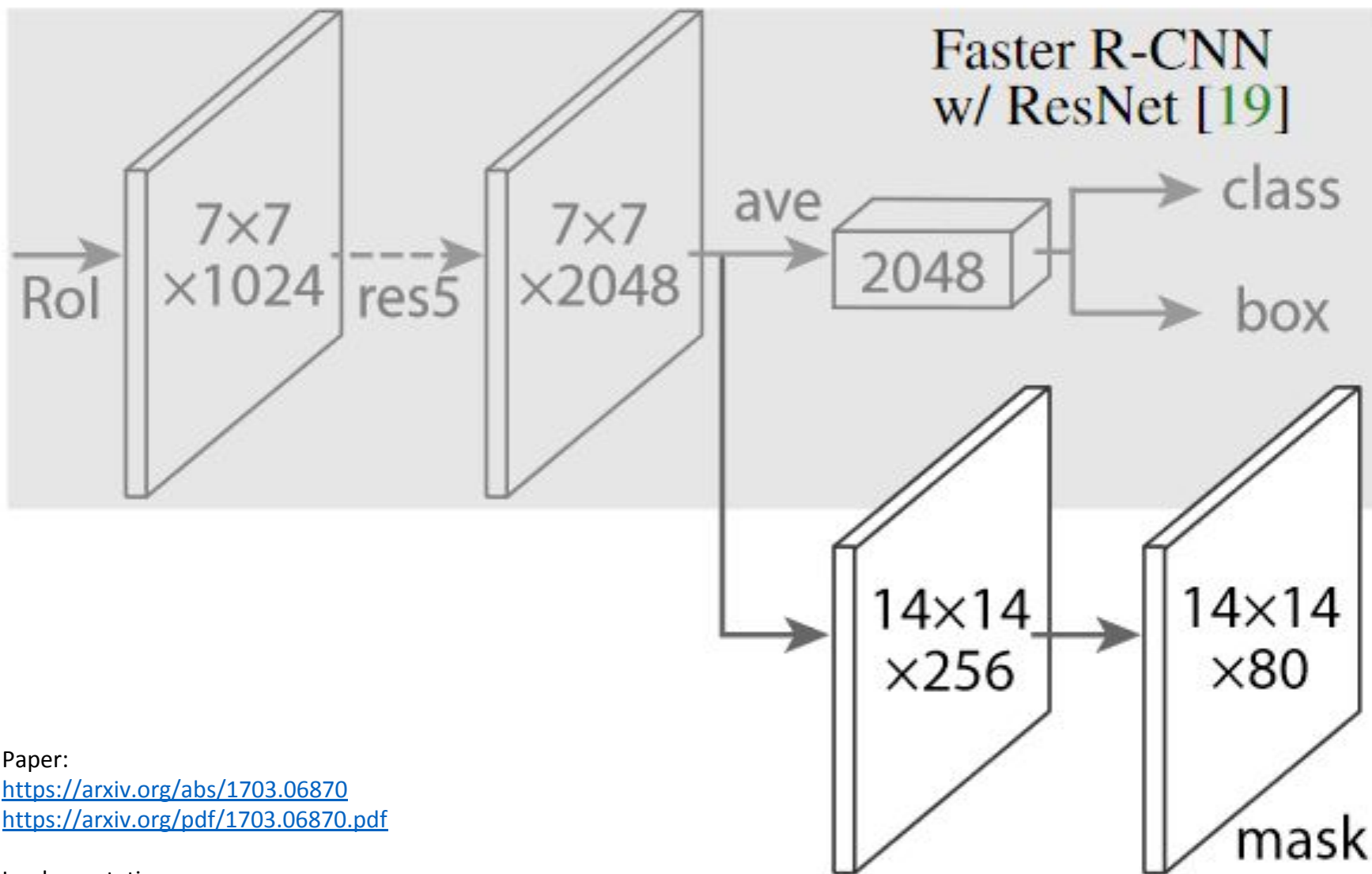


Figure 2: Our segmentation model based on FCN. We introduce 1) the outfit encoder to filter inappropriate clothing combination from segmentation, and 2) CRF to assign a visually consistent set of clothing labels.

Mask R-CNN



Paper:

<https://arxiv.org/abs/1703.06870>

<https://arxiv.org/pdf/1703.06870.pdf>

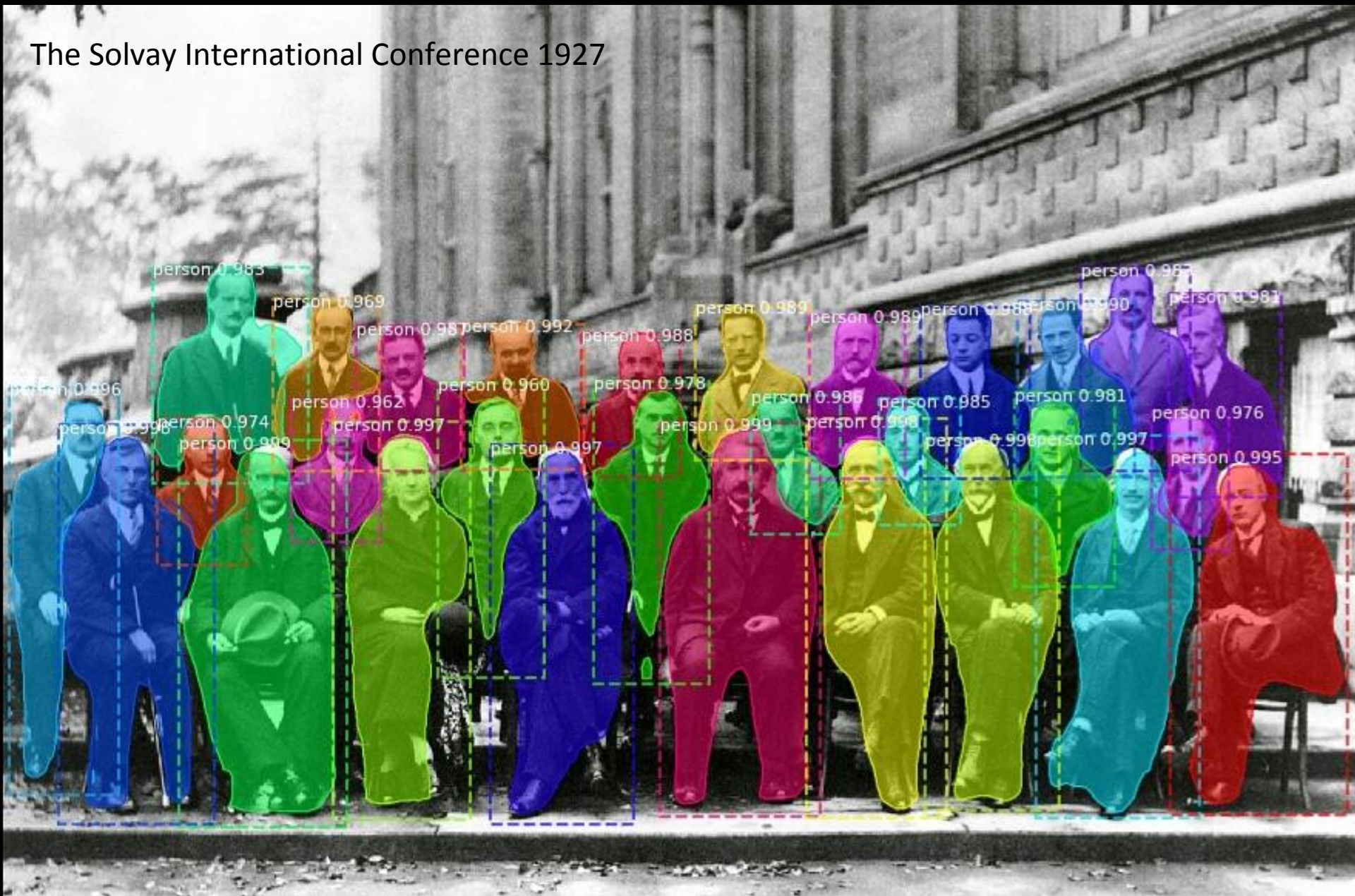
Implementations:

<https://github.com/facebookresearch/Detectron>

https://github.com/matterport/Mask_RCNN



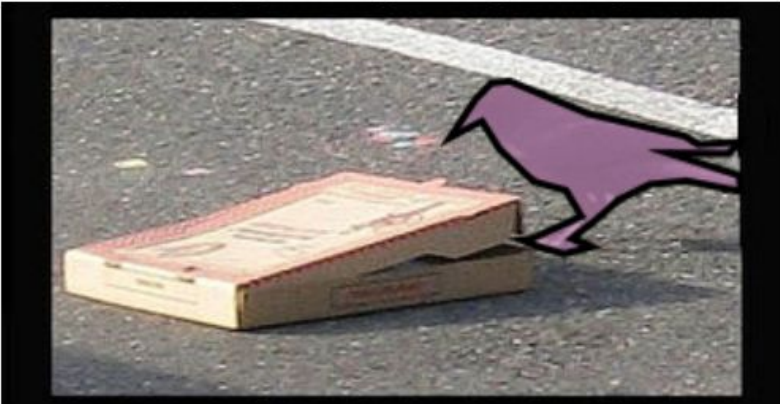
The Solvay International Conference 1927

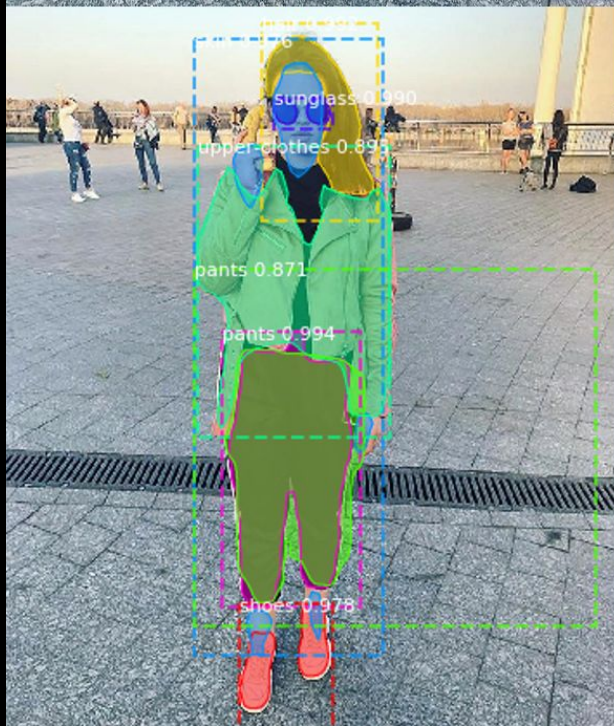
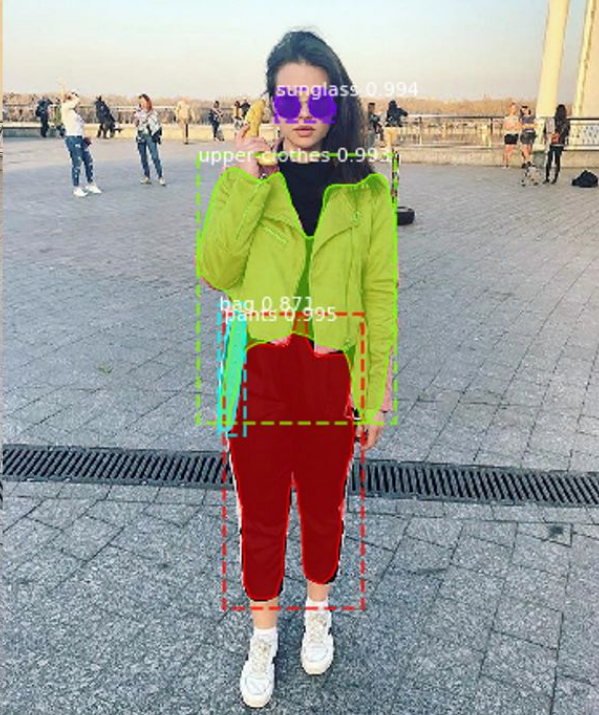


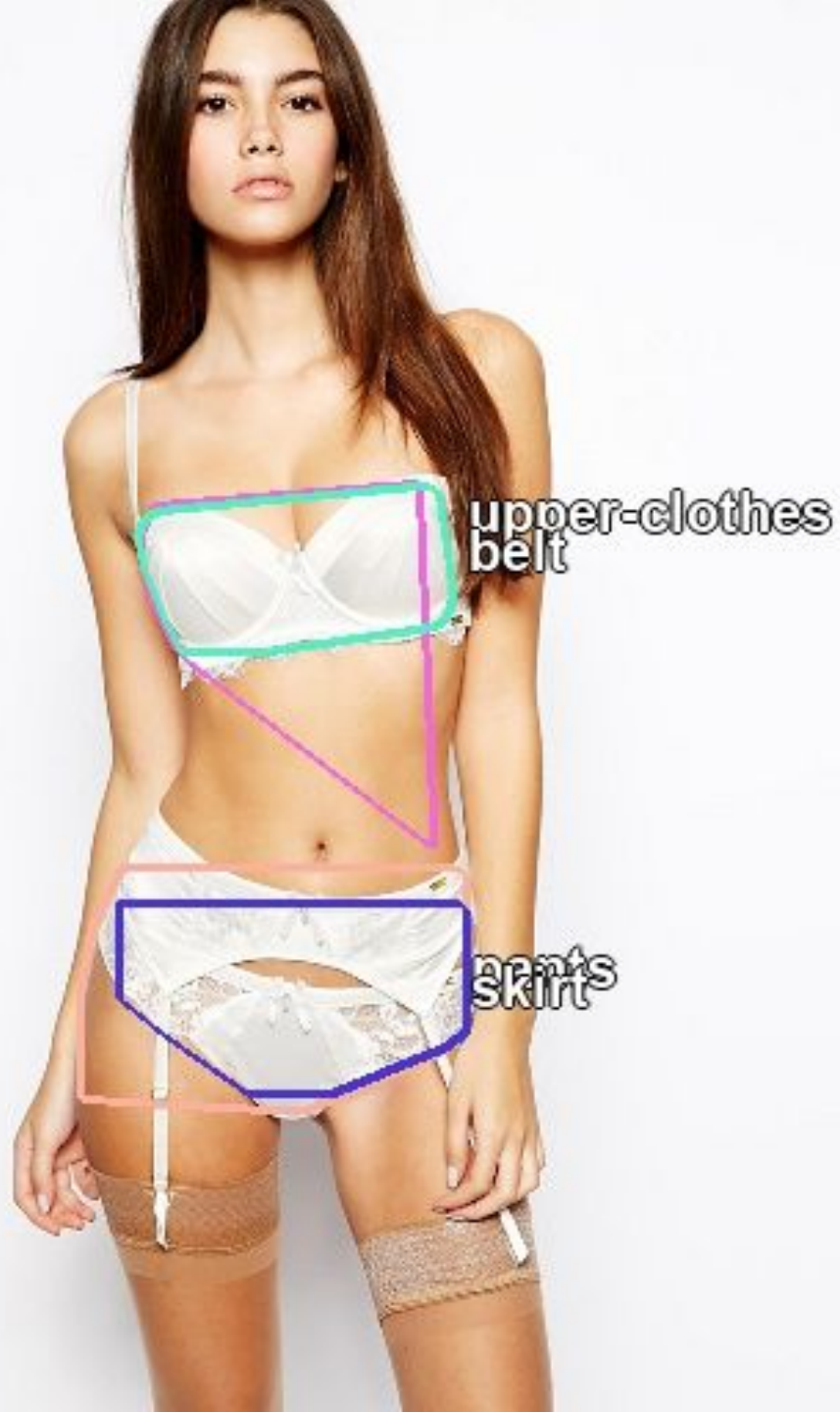
COCO 2017 train/val browser (123,287 images, 886,284 instances). Crowd labels not shown.



7 results

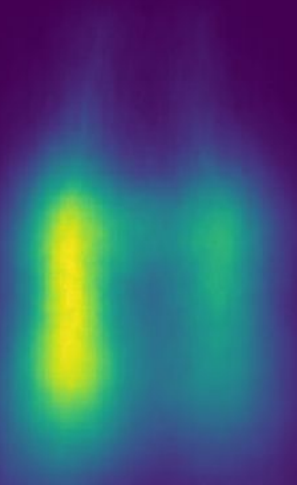
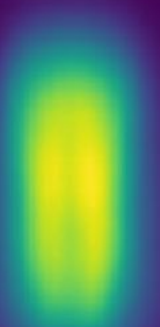
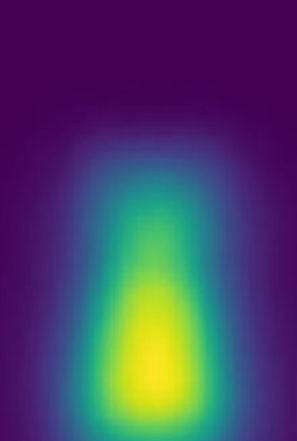
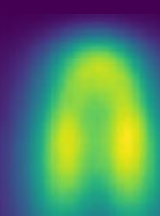
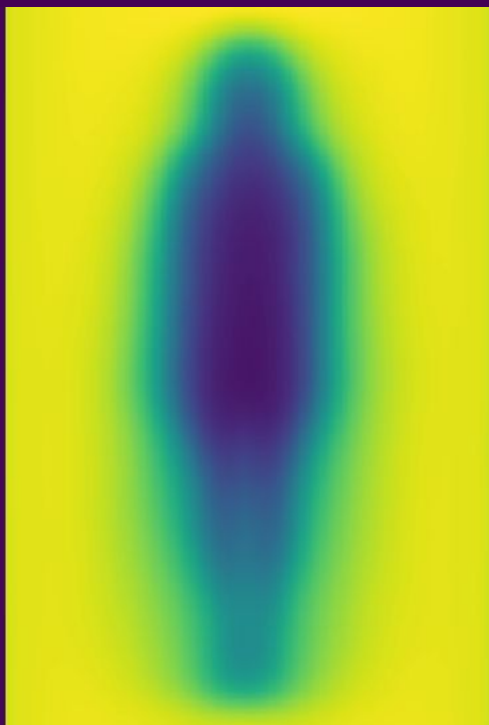






Generalization

There was no lingerie in the data set



Object retrieval | Retrieval for e-commerce

Street2Shop a.k.a. Where to Buy It dataset

The dataset contains **39,479 pairs** of exactly matching items worn in street photos and shown in shop images. Train subset contains **11 classes**.



Figure 2: Example *street outfit photos*, including large variations in pose, camera angle, composition and quality.



Figure 3: Example *shop photos*, displaying a wide range of apparel photography techniques.

Street2Shop a.k.a. Where to Buy It dataset



DeepFashion

ATTRIBUTE PREDICTION

IN-SHOP RETRIEVAL

CONSUMER-TO-SHOP RETRIEVAL

FASHION LANDMARK DETECTION

FASHION SYNTHESIS



- 7,982 number of clothing items;
- 52,712 number of in-shop clothes images, and ~200,000 cross-pose/scale pairs;

DeepFashion

ATTRIBUTE PREDICTION

IN-SHOP RETRIEVAL

CONSUMER-TO-SHOP RETRIEVAL

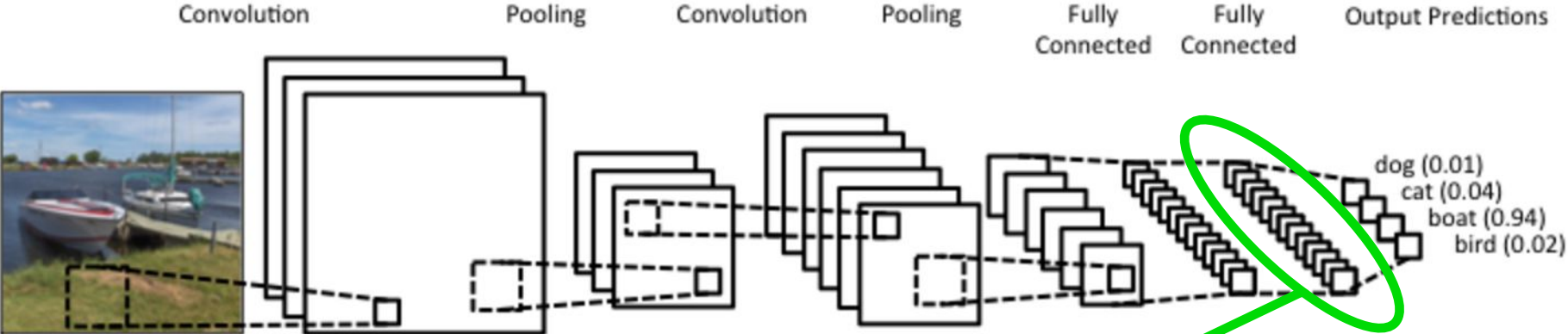
FASHION LANDMARK DETECTION

FASHION SYNTHESIS

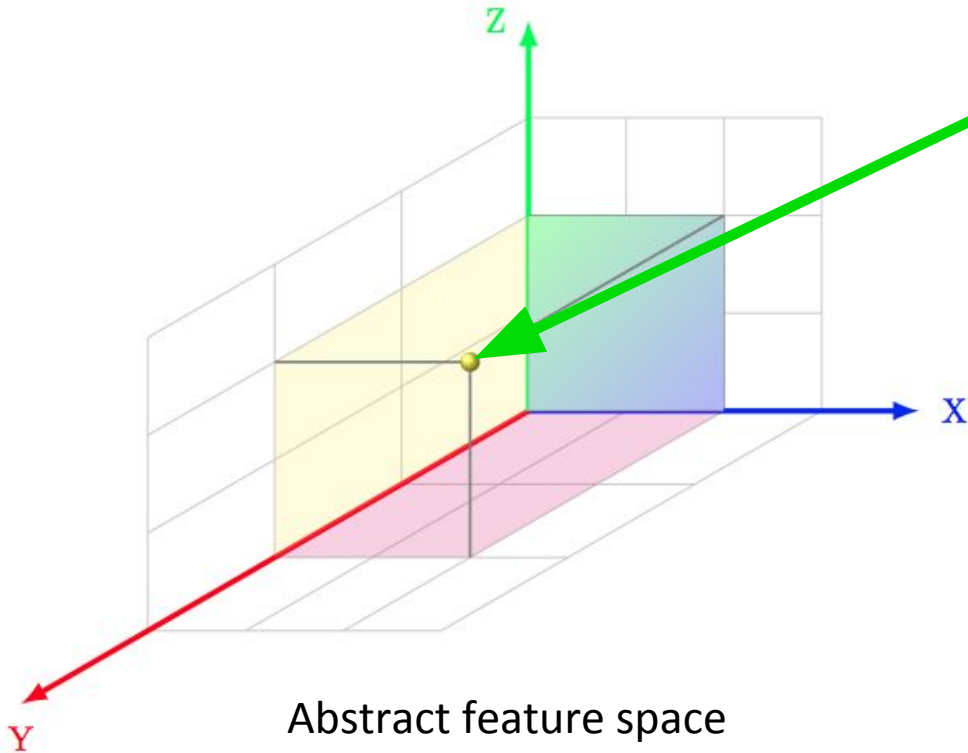


- 33,881 number of clothing items;
- 239,557 number of consumer/shop clothes images, and 195,540 cross-domain pairs;

Convolutional Neural Network

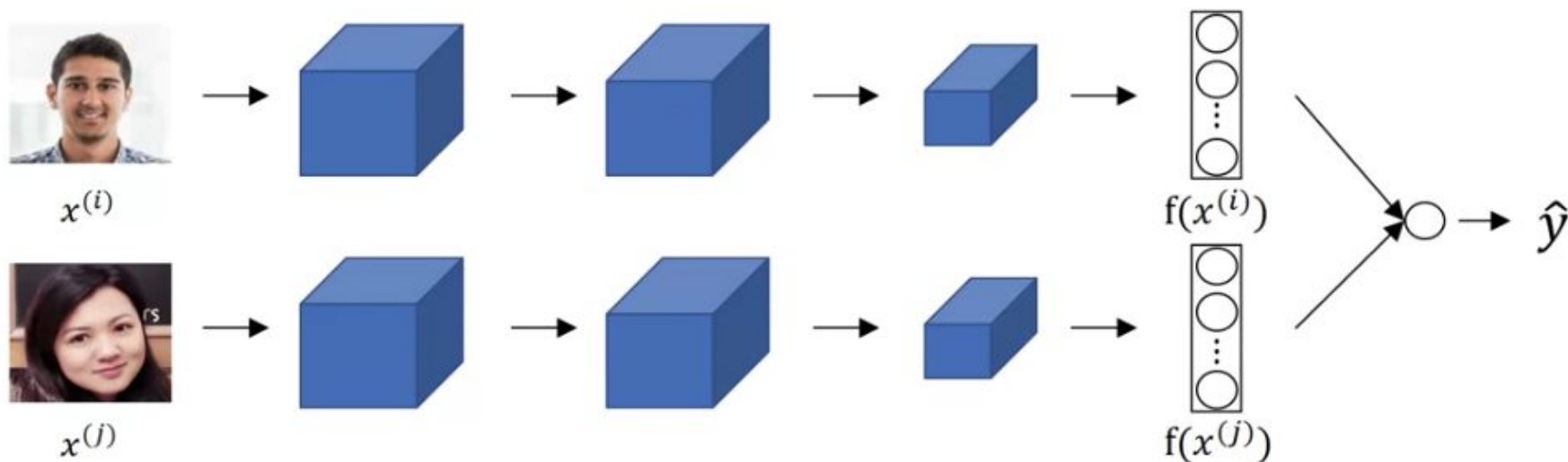


N-dimensional embedding vector



Siamese Network

Learning the similarity function



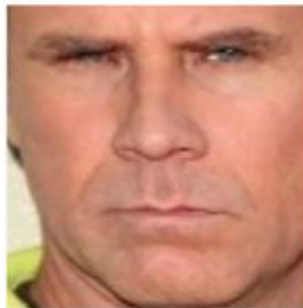
A single 'triplet' training step:

“metric learning”

Picture of
Chad Smith



Test picture of
Will Ferrell



Another picture of
Will Ferrell



↓
128 measurements
generated by neural net

↓
128 measurements
generated by neural net

↓
128 measurements
generated by neural net

↓
Compare results

↓
Tweak neural net slightly so that the
measurements for the two Will Ferrell
pictures are closer and the Chad Smith
measurements are further away

Triplet loss

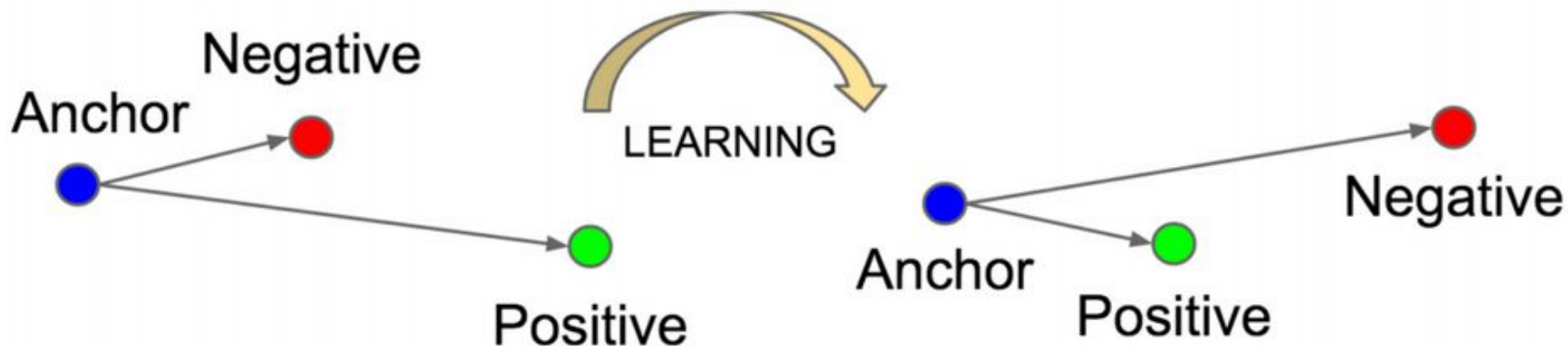


Figure 3. The **Triplet Loss** minimizes the distance between an *anchor* and a *positive*, both of which have the same identity, and maximizes the distance between the *anchor* and a *negative* of a different identity.

The loss that is being minimized is then $L =$

$$\sum_i^N \left[\|f(x_i^a) - f(x_i^p)\|_2^2 - \|f(x_i^a) - f(x_i^n)\|_2^2 + \alpha \right]_+$$

FaceNet: A Unified Embedding for Face Recognition and Clustering

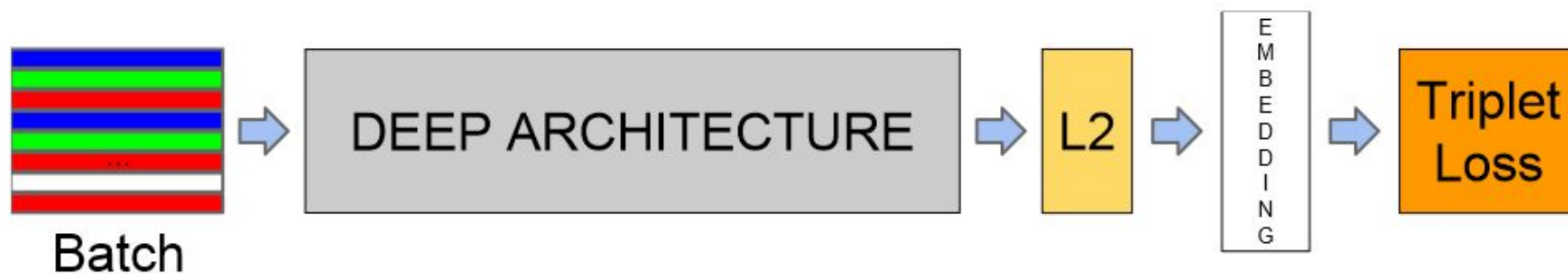
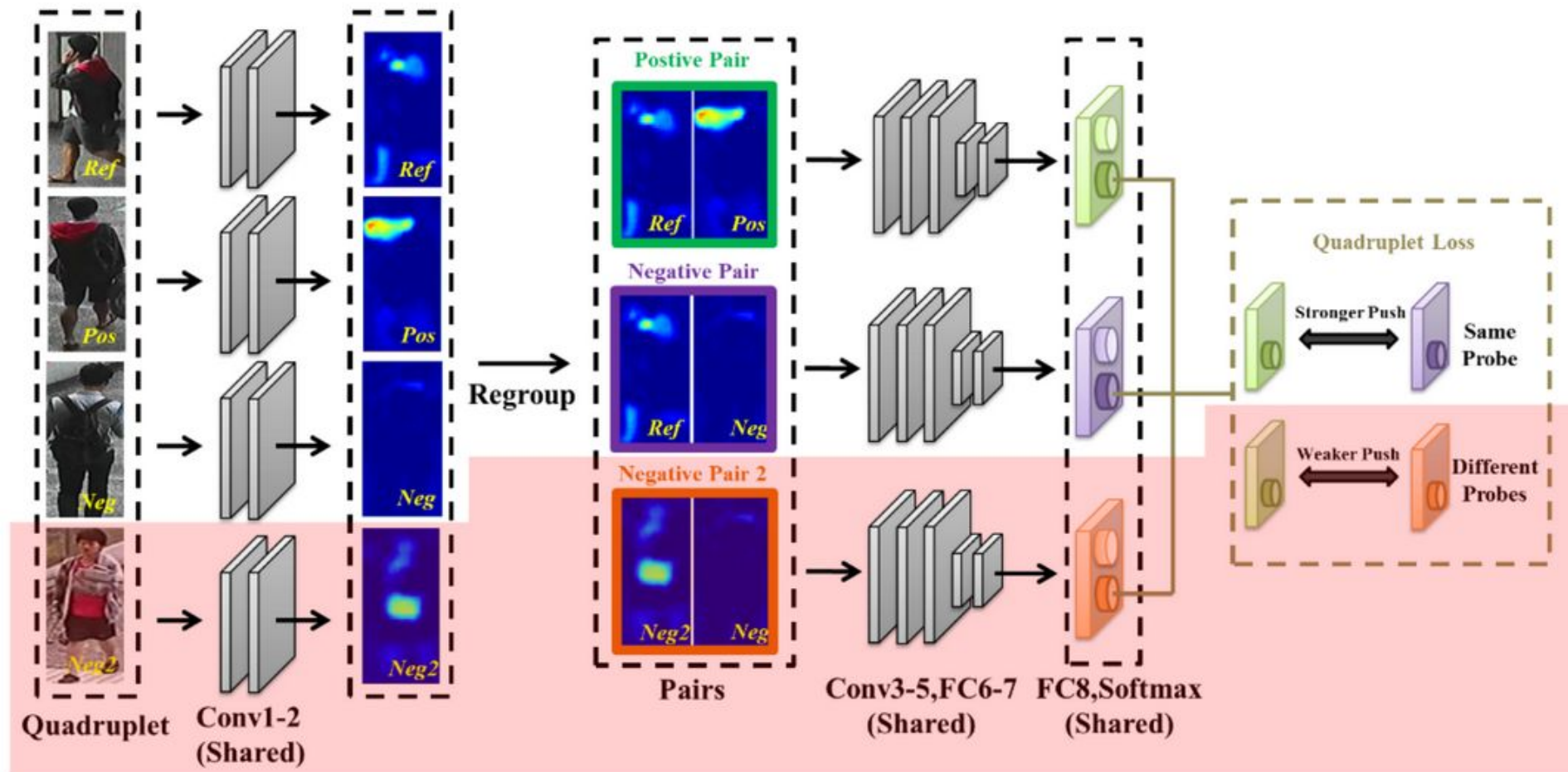


Figure 2. **Model structure.** Our network consists of a batch input layer and a deep CNN followed by L_2 normalization, which results in the face embedding. This is followed by the triplet loss during training.

Beyond triplet loss: a deep quadruplet network



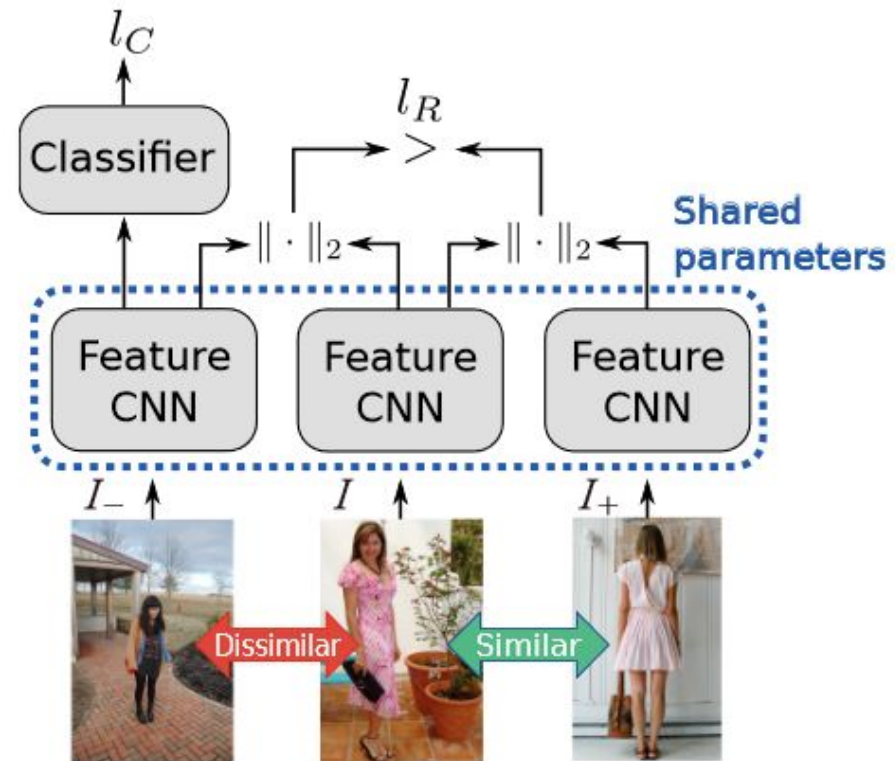
3. The framework of the proposed quadruplet deep network. The red shadow region indicates elements of the new constraint

Fashion Style in 128 Floats:

Joint Ranking and Classification using Weak Data for Feature Extraction

Table 1: Feature extraction network architecture. All convolutional layers have 1×1 padding and all layers besides the max pooling layer have a 1×1 stride, while the max pooling layers have a 4×4 stride.

	type	kernel size	output size	params
	convolution	3×3	384x256x64	1,792
	convolution	3×3	384x256x64	36,928
	dropout (25%)		384x256x64	
	max pooling	4×4	96x64x64	
	batch normalization		96x64x64	128
	convolution	3×3	96x64x128	73,856
	convolution	3×3	96x64x128	147,584
	dropout (25%)		96x64x128	
	max pooling	4×4	24x16x128	
	batch normalization		24x16x128	256
	convolution	3×3	24x16x256	295,168
	convolution	3×3	24x16x256	590,080
	dropout (25%)		24x16x256	
	max pooling	4×4	6x4x256	
	batch normalization		6x4x256	512
	convolution	3×3	6x4x128	32,896
	fully-connected		128	393,344
	TOTAL		128	1,572,544



Fashion Style in 128 Floats: Joint Ranking and Classification using Weak Data for Feature Extraction

I_-



I



I_+



I_-



I



I_+



Fashion144k dataset

Fashion144k (277,527 images)

Mean



As weak annotations, we use the “color” tags provided by the Fashion144k dataset which consist of color-garment tags such as “red-sweater” or “blue-boots”, the set of which has 3,180 unique tags

Fashion Style in 128 Floats: Joint Ranking and Classification using Weak Data for Feature Extraction

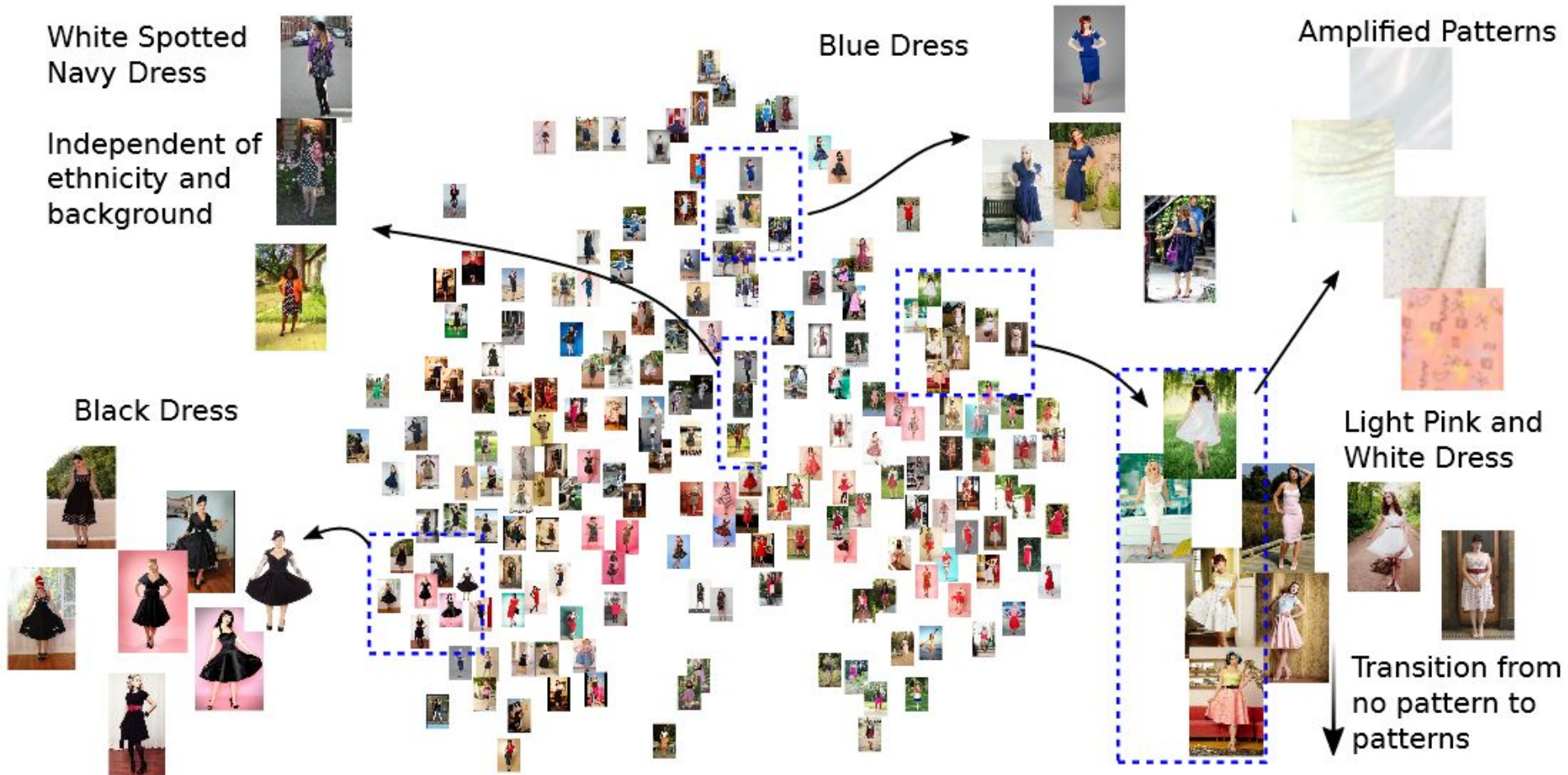
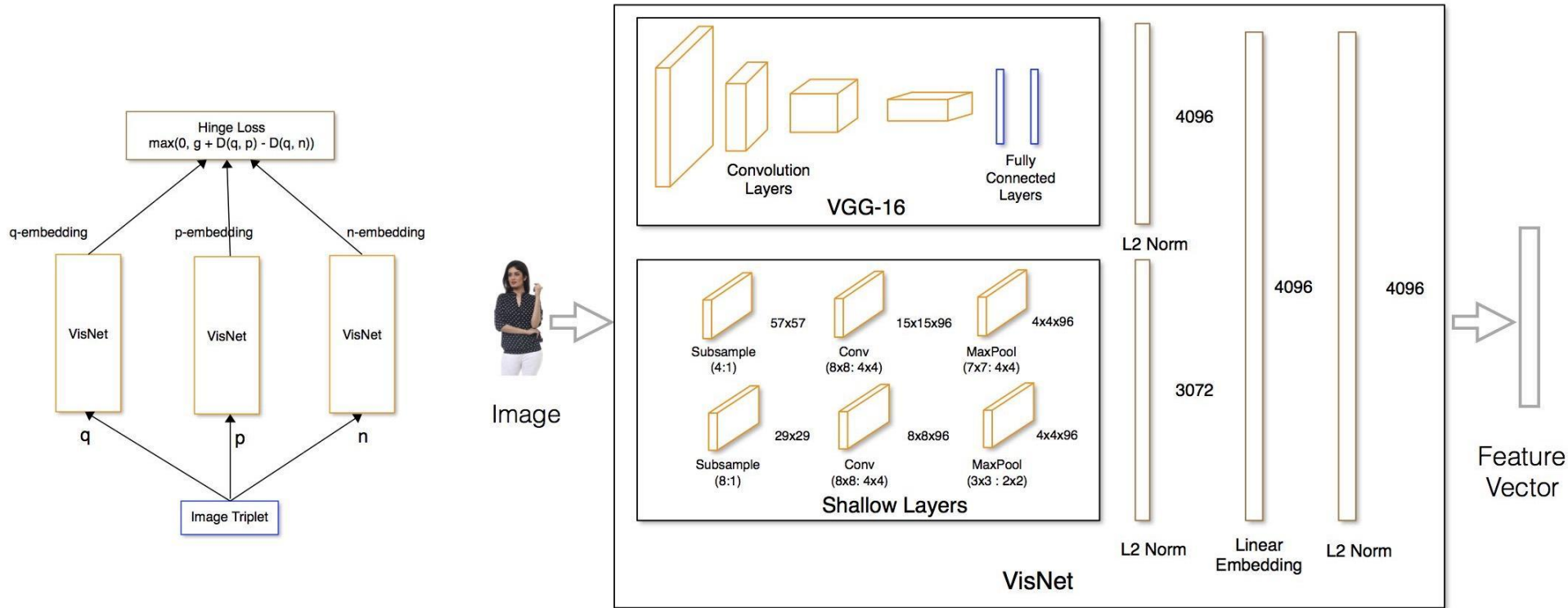


Figure 7: Visualization of the fashion style space of the Pinup class from the Hipster Wars [16] dataset using t-SNE [33].

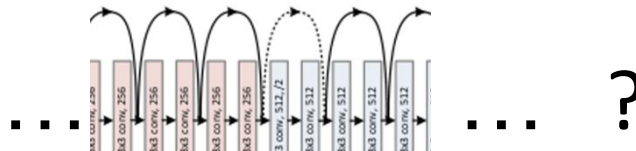
Deep Learning based Large Scale Visual Recommendation and Search for E-Commerce

We propose a unified Deep Convolutional Neural Network architecture, called VisNet



(a) Overall Deep Ranking Architecture

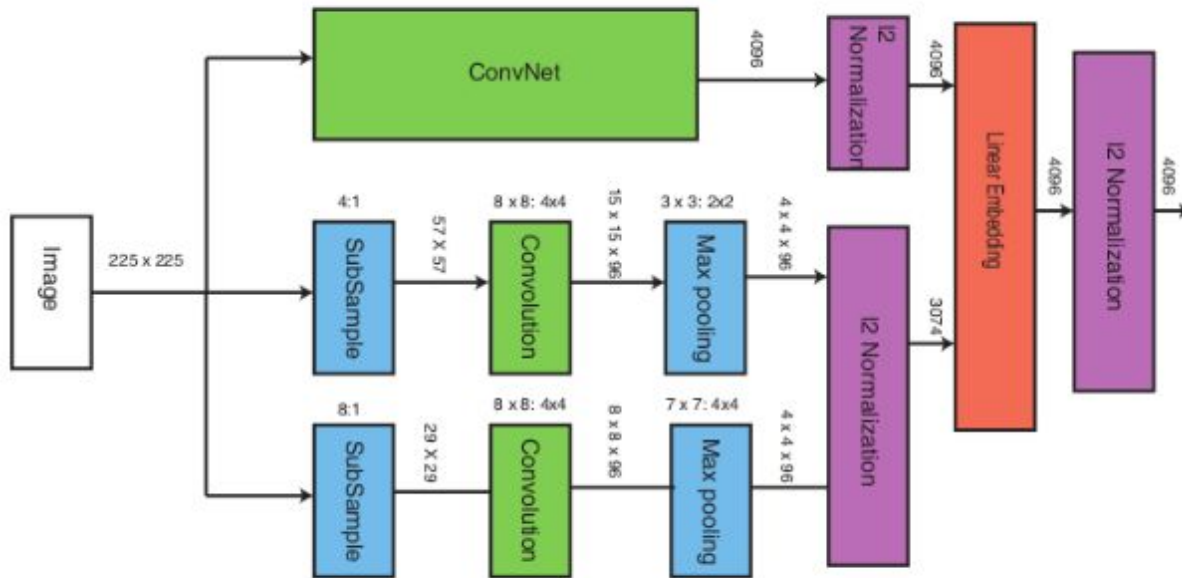
(b) VisNet Architecture



Fine-grained Image Similarity with Deep Ranking



Triplet Pairs



Multiscale Network Architecture

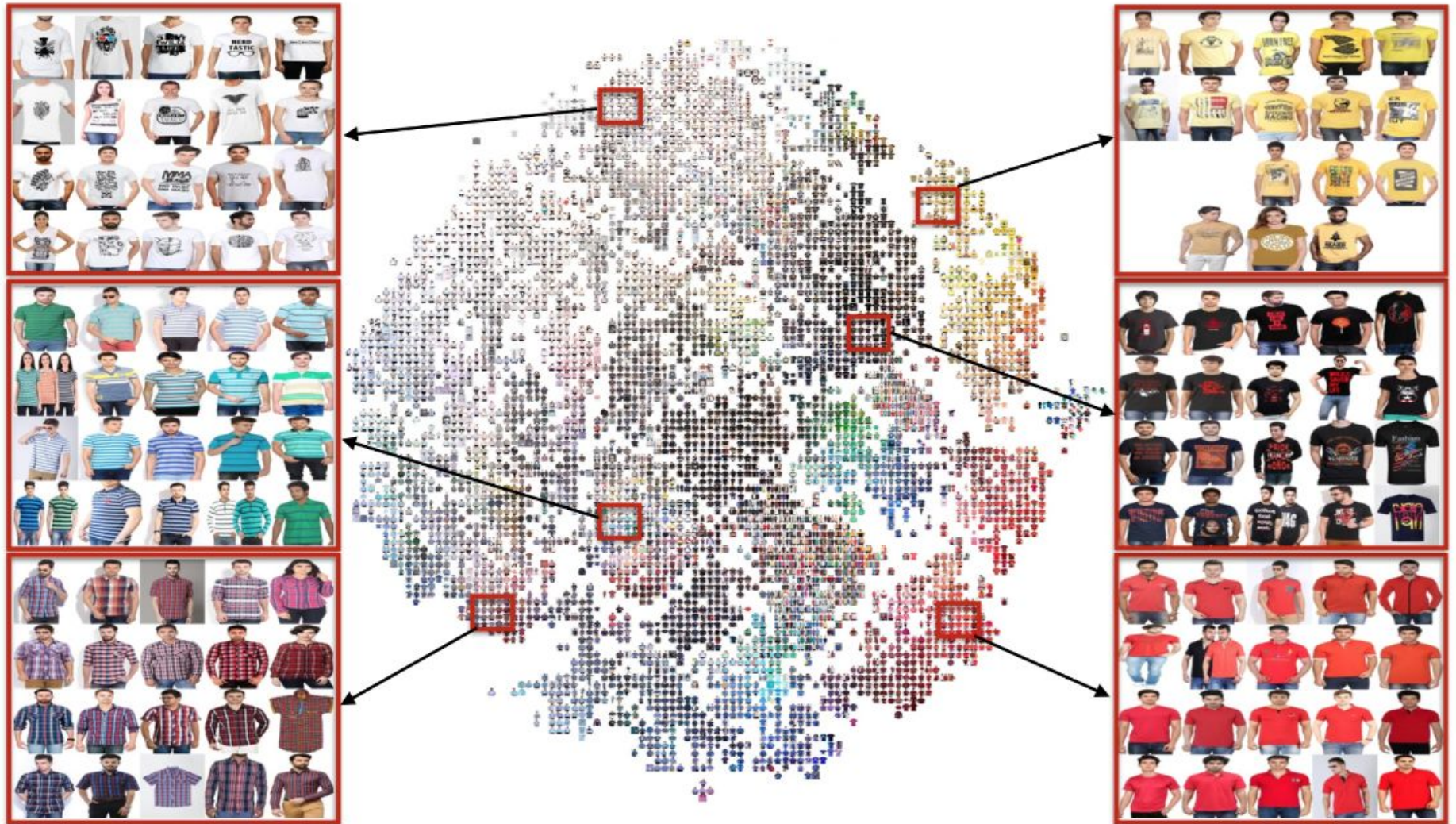


Figure 9: t-SNE Visualisation of final layer VisNet embeddings from catalog items

Clothing Retrieval with Visual Attention Model

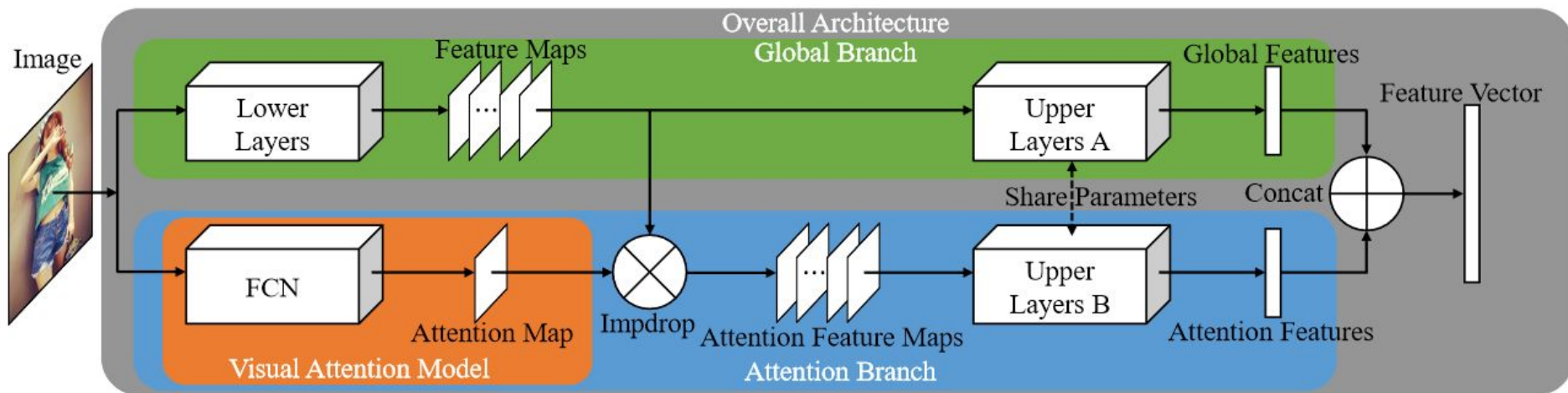
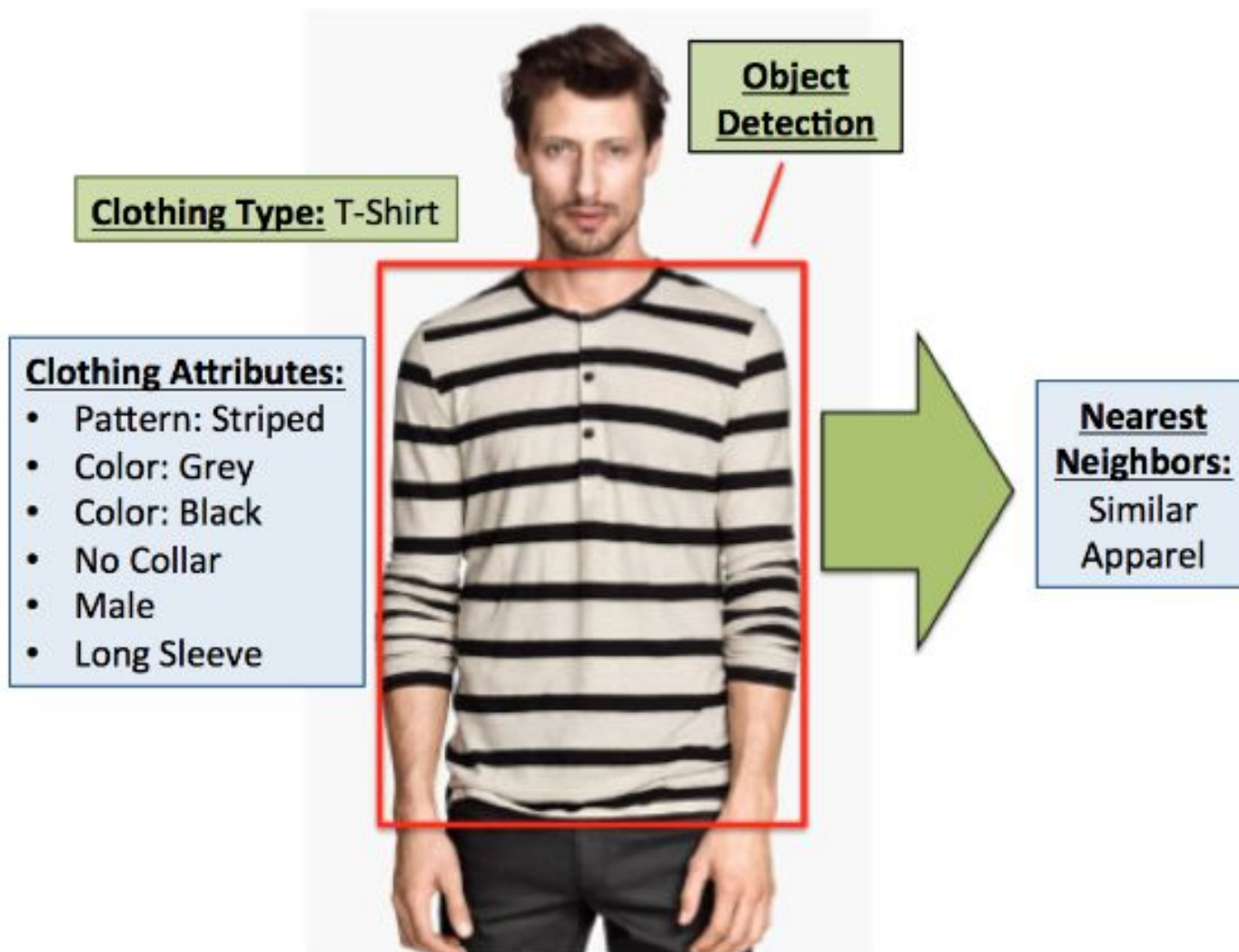
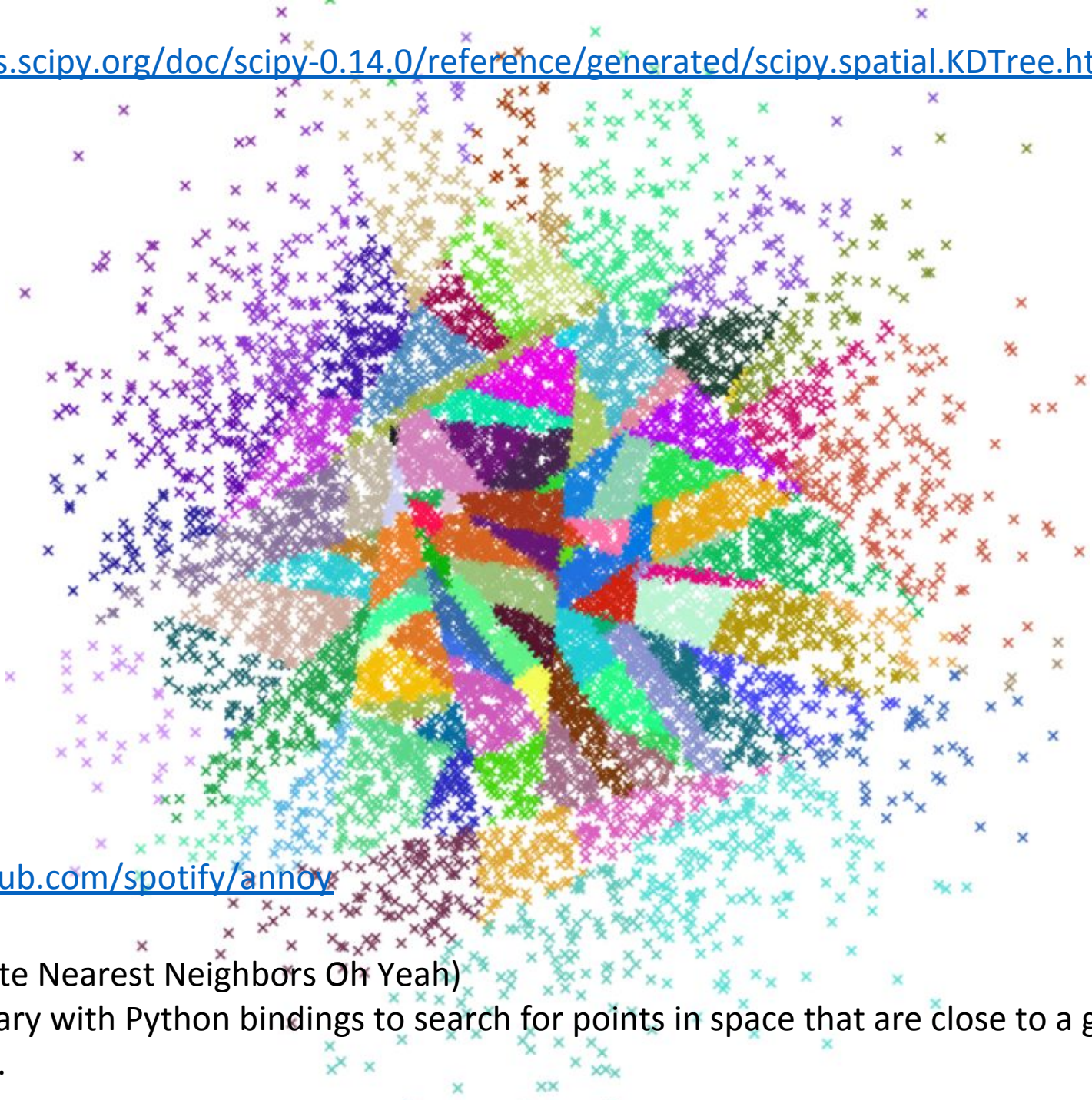


Fig. 1. The overall architecture of our model. The architecture consists of: 1) the **Global Branch** produces features maps from lower layers and outputs global features using the upper layers A; 2) the **Attention Branch** extracts the attention map with **VAM**, combines it with the feature maps into attention feature maps using **Imdrop** and outputs attention features through upper layers B.

Convolutional Neural Networks for Fashion Classification and Object Detection



<https://docs.scipy.org/doc/scipy-0.14.0/reference/generated/scipy.spatial.KDTree.html>



<https://github.com/spotify/annoy>

Annoy

(Approximate Nearest Neighbors Oh Yeah)

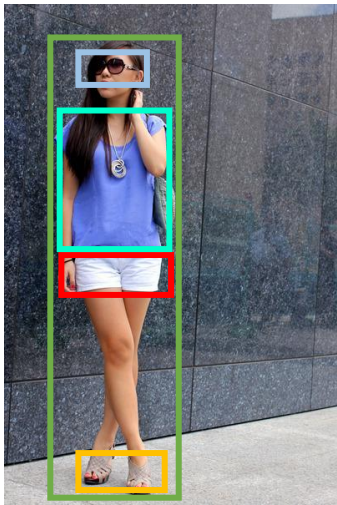
is a C++ library with Python bindings to search for points in space that are close to a given query point.



Original image



Semantic segmentation



Object detection



Background cleaning



Clothing attributes:

Color: blue

Type: top

Pattern: solid

Gender: female

Brand: brand name

Clothing attributes:

Color: white

Type: shorts

Pattern: solid

Gender: female

Brand: brand name



Steel blue bag

Indian Red sunglasses

Light Salmon belt

Gray footwear

Bisque outer

Gray pants



Lemon Chiffon bag

Pink dress

Rosy Brown top

Gray footwear

Steel Blue outer



Hard examples



 WeddingShoppelnc.com



Featured on
WEDDING
INSPIRATION
weddinginspiration.com

DONNA MORGAN



Featured on
WEDDING
INSPIRASI
weddinginspirasi.com

DONNA MORGAN

Áo dài





https://st.depositphotos.com/1561509/5121/1/950/depositphotos_51210809-stock-photo-ukrainian-cossack-in-national-dress.jpg

<http://soloma.in.ua/image/cache/data/Female/30-600x600.jpg>



T-SHIRT SHIRT £ 935



INDIGO





<https://www.harpersbazaar.com/uk/fashion/fashion-news/a13448923/moschino-is-selling-a-dry-cleaning-bag/>

<https://assets.vogue.com/>

Classification | Brand or attribute discovery

ATTRIBUTE PREDICTION

IN-SHOP RETRIEVAL

CONSUMER-TO-SHOP RETRIEVAL

FASHION LANDMARK DETECTION

FASHION SYNTHESIS



Image synthesis

Be Your Own Prada: Fashion Synthesis with Structural Coherence

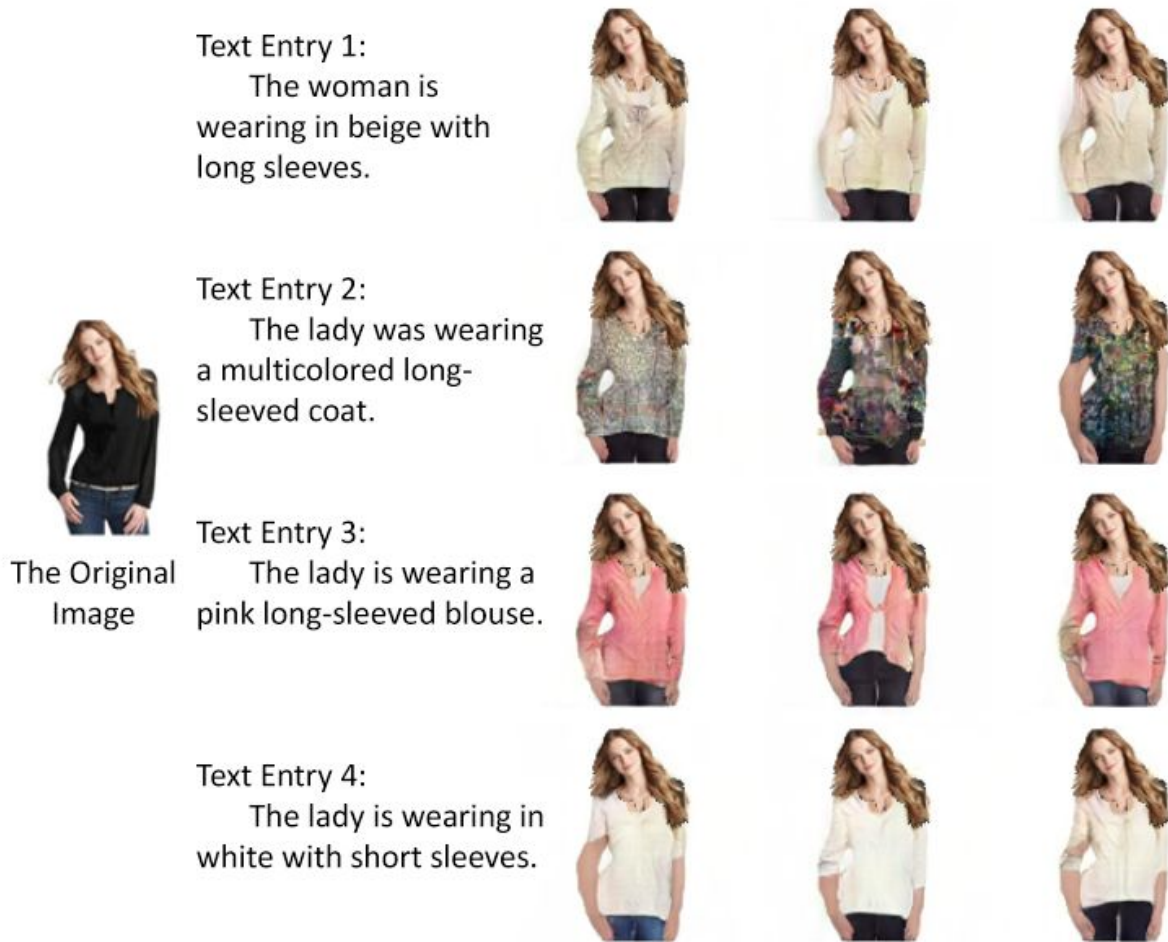


Figure 1. Given an original wearer's input photo (left) and different textual descriptions (second column), our model generates new outfits onto the photograph (right three columns) while preserving the pose and body shape of the wearer.

Be Your Own Prada: Fashion Synthesis with Structural Coherence

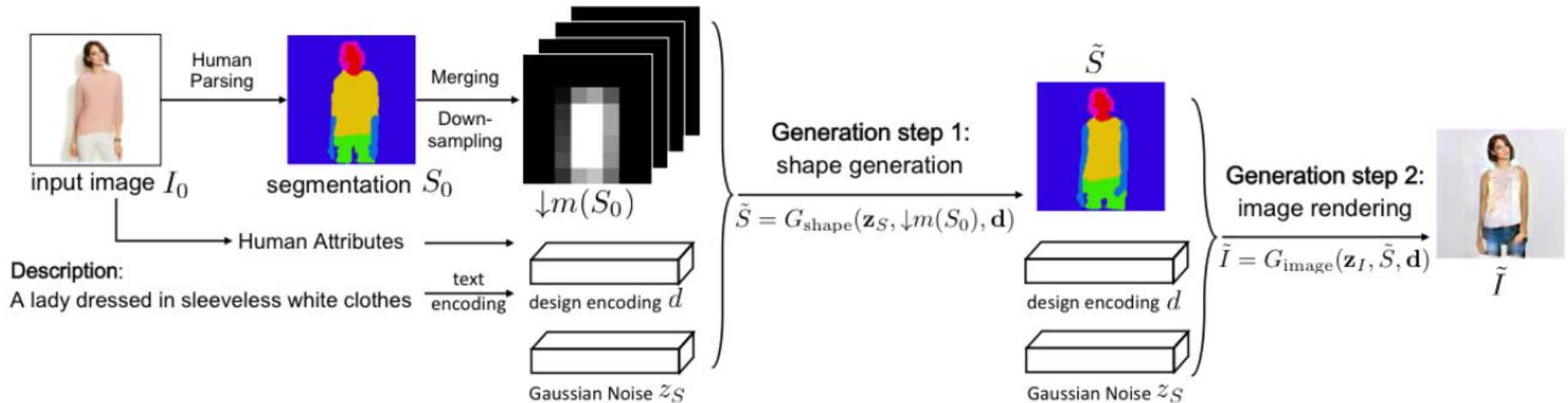


Figure 2. **Proposed framework.** Given an input photograph of a person and a sentence description of a new desired outfit, our model first generates a segmentation map \tilde{S} using the generator from the first GAN. We then render the new image with another GAN, with the guidance from the segmentation map generated in the previous step. At test time, we obtain the final rendered image with a forward pass through the two GAN networks.



VITON: An Image-based Virtual Try-on Network



VITON: An Image-based Virtual Try-on Network

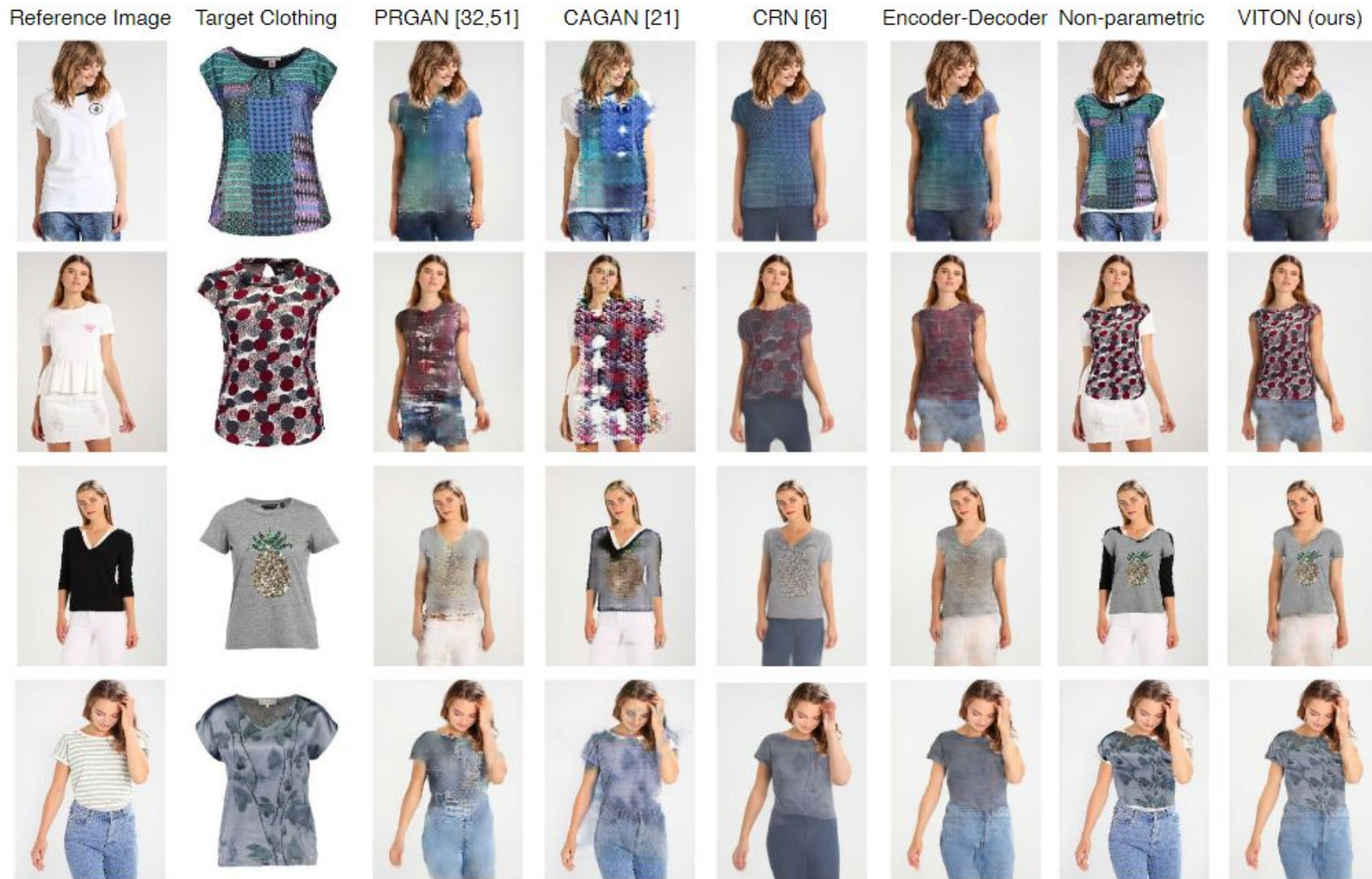


Figure 6: Qualitative comparisons of different methods. Our method effectively renders the target clothing on to a person.

VITON dataset







Dataset

The dataset is no longer publicly available due to copyright issues. For those who have already downloaded the dataset, please note that using or distributing it is illegal!

Caption synthesis

Netizen-Style Commenting on Fashion Photos: Dataset and Diversity Measures

WWW'2018, April 2018, Lyon, France

Methods						
	(a)	(b)	(c)	(d)	(e)	(f)
Human	Gorgeous! Love pants, sunglasses and everything else! ;)	amazing look! Love everything in it!Hyped!	amazing picture :-)	great, man!!	cool	Cute top!
CaptionBot	a couple of women standing next to a woman and she seems 😊.	a group of people standing next to a woman in a black wet suit and they seem 😊😊😊.	a man is jumping in the air on a skateboard.	a group of people standing around each other.	a little girl walking down the street and they seem 😊😊.	a woman standing next to a brick wall and she seems 😊.
NC	I love your dress!	I love your shoes <3	I love your hair and look	nice look	nice	Love your dress!
Attention	I love your look	I love your shoes	I love the shoes	I love your shoes	I love your look	I love your shoes
Our proposed NSC _{NC}	I love the combinations :)) My heart for today goes to you! :)	I love your shoes!!! HYPED! <3	I love your hair and the coat, unbelievably gorgeous. Hyped!	I love your style!! :D	I love your hair	Love the dress!



(a) **Human**: wow so stunning!! love this look, and u r so pretty ♥

(b) **CaptionBot**: A woman is laying on the grass.

(c) **NC**: love the skirt

(d) **Attention**: love the shoes

(e) **NSC**: Love this button up bustier! It's fabulous! This look is amazing.



(a) **Human**: Wow! Perfection!

(b) **CaptionBot**: A group of people standing around each other.

(c) **NC**: love the blazer!!!!

(d) **Attention**: love the shoes

(e) **NSC**: I love your outfits. This one isn't an exception <3



(a) **Human**: love ur ombre hair <3333 +1

(b) **CaptionBot**: A group of women standing next to a woman.

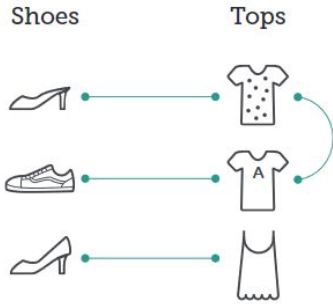
(c) **NC**: love the dress

(d) **Attention**: love the shoes

(e) **NSC**: I love the combinations :)) My heart for today goes to you! :)

Recommendations and style detection

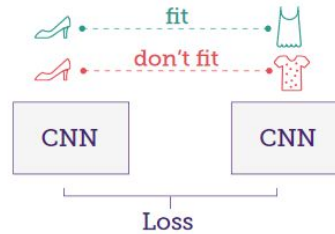
Step 1: Data collection



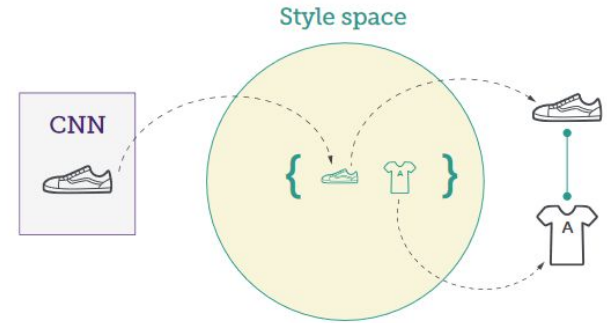
Step 2: Training data generation



Step 3: Siamese CNNs



Step 4: Recommendation



Learning Visual Clothing Style with Heterogeneous Dyadic Co-occurrences

"Styles and Substitutes" dataset

https://www.youtube.com/watch?v=ij-p_5oppl4&feature=youtu.be

<https://arxiv.org/pdf/1509.07473.pdf>

Visually-Aware Fashion Recommendation and Design with Generative Image Models



Visually-Aware Fashion Recommendation and Design with Generative Image Models



(a) Generated Images

(b) ℓ_1 Nearest Neighbors



The FashionGen Challenge

ELEMENT^{AI} SSENSE

Images above are AI-generated

<https://fashion-gen.com/>

FASHION

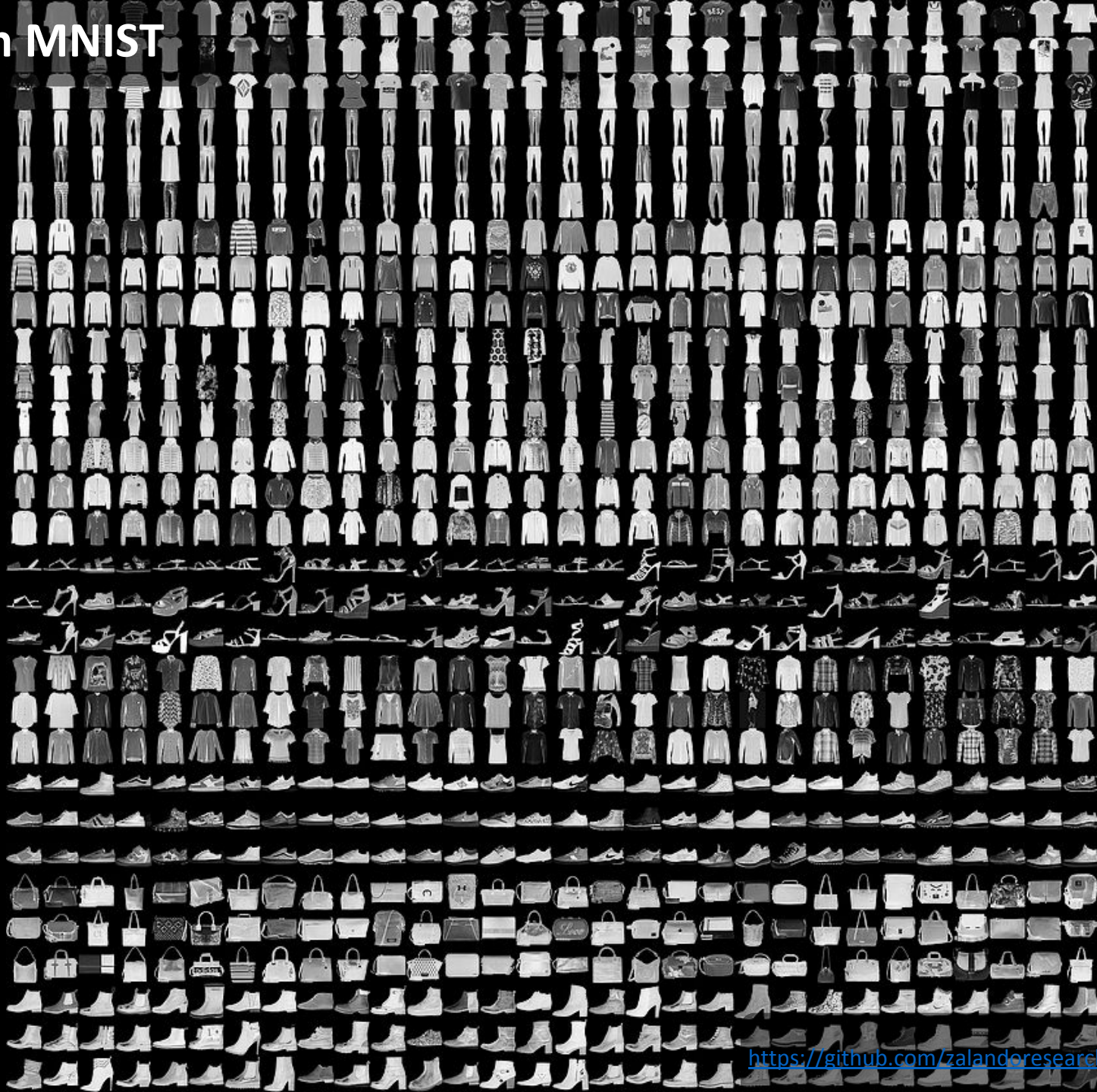
HONGKONG201

Attributes Recognition of Apparel

Apparel attributes are the basic knowledge of fashion field, which are large and complex. We constructed a hierarchical attributes tree as a structured classification target, to describe the cognitive process of apparel. Contestants are invited to design algorithms to recognize attributes of apparel images. This task might be widely applied for apparel image searching, navigating tagging, mix-and-match recommendation, etc.

https://tianchi.aliyun.com/markets/tianchi/FashionAleng?_lang=en_US#home

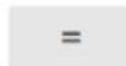
Fashion MNIST



Hipster Wars: Discovering Elements of Fashion Styles



Who's more **Hipster**?
(click on one)



(a) Snapshot of the game

Most

Least

Pinup



Goth



Hipster



Bohemian



Preppy



Hipster Wars: Discovering Elements of Fashion Styles

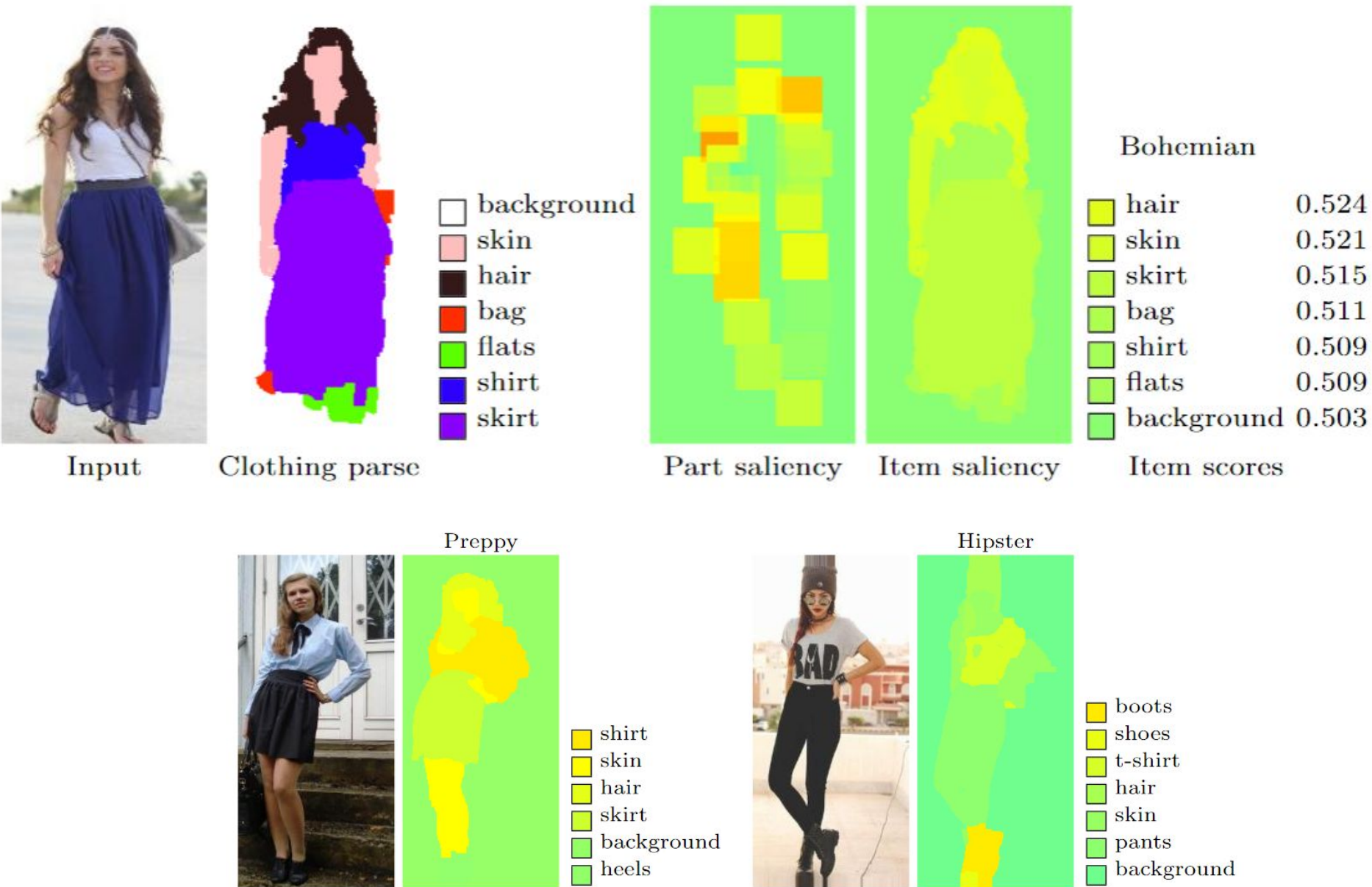
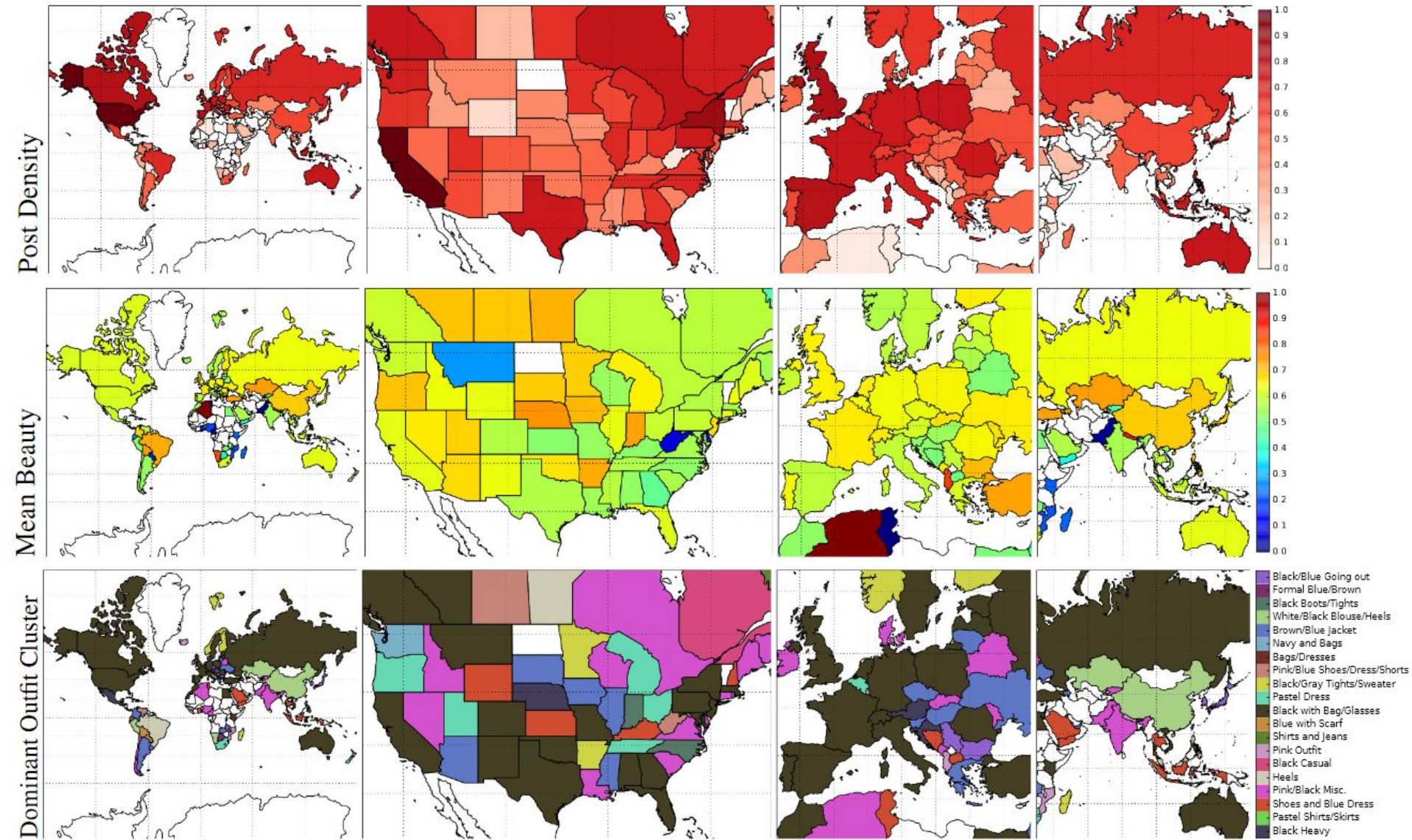


Fig. 10: Example predicted style indicators for individuals.

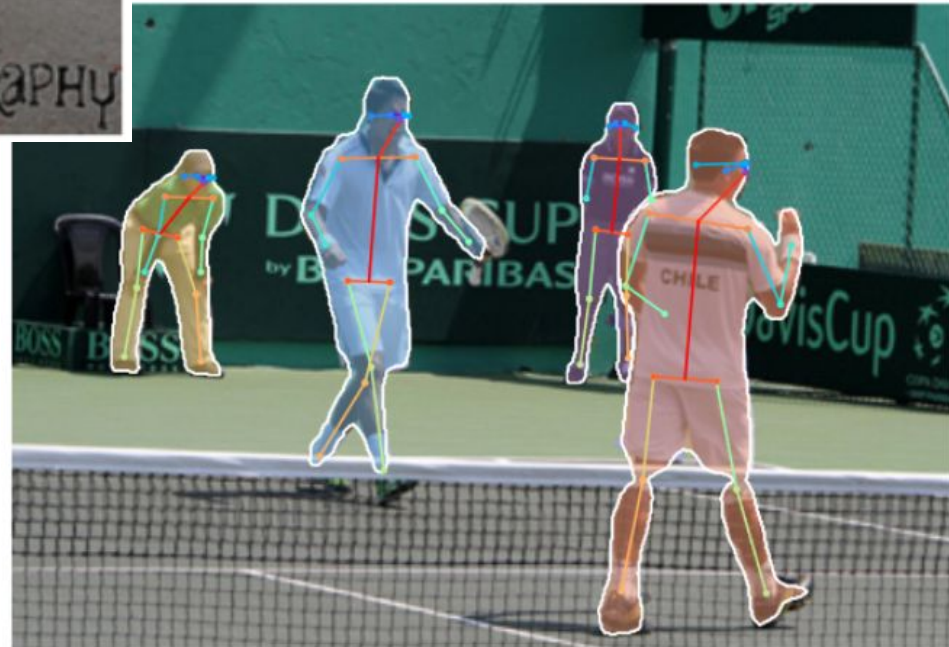
Neuroaesthetics in Fashion: Modeling the Perception of Fashionability



Pose estimation



Mask R-CNN for Human Pose Estimation





Social groups

[Murillo 12] [Kwak 13]



Occupation

[Song 11] [Shao 13]



Markable



Vasily



Fashwell



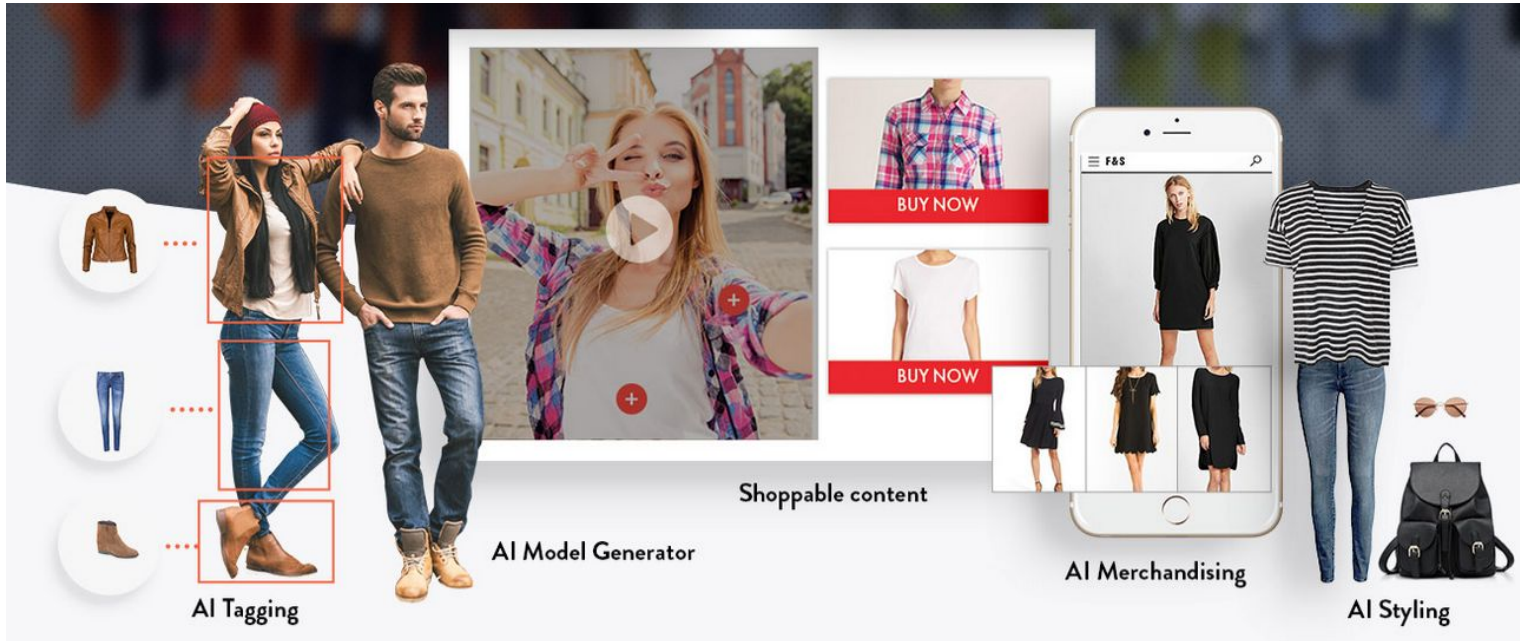
Shopagon



Wide Eyes Technologies

<http://demo.markable.ai/#/>

<https://shopagon.com/samplefeeds>



<https://vue.ai/>

<https://www.slideshare.net/KotaYamaguchi1/computer-vision-meets-fashion>



Only two of these images were taken by a camera. (Spoiler: Row 1, C and Row 2, D) (Vue.ai)

<https://qz.com/1090267/artificial-intelligence-can-now-show-you-how-those-pants-will-fit/>

Edit

- Same type
- Same color
- Same sleeve
- Same texture
- Same collar

Submit

衣服类型: 连衣裙, 样式: 乱花, 性别: 女



帛薇儿(Pokwai)2015原创设计春夏
欧美大牌一字领优雅长裙花朵真丝
雪纺连衣裙女 红色 S

JD.京东.COM



红叶精灵 2015年春装夏新款 气质
时尚中长款 韩版大裙摆修身花色连
衣裙 十号色 L

JD.京东.COM

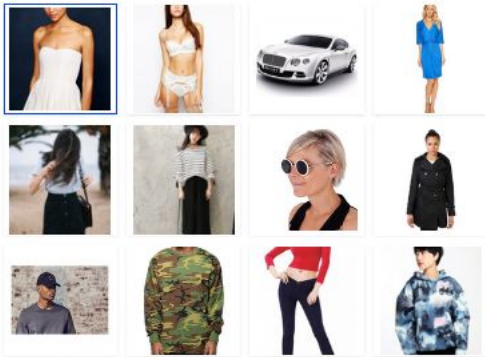


帛薇儿(Pokwai)2015原创设计春夏
欧美大牌一字领优雅长裙花朵真丝
雪纺连衣裙女 红色 S

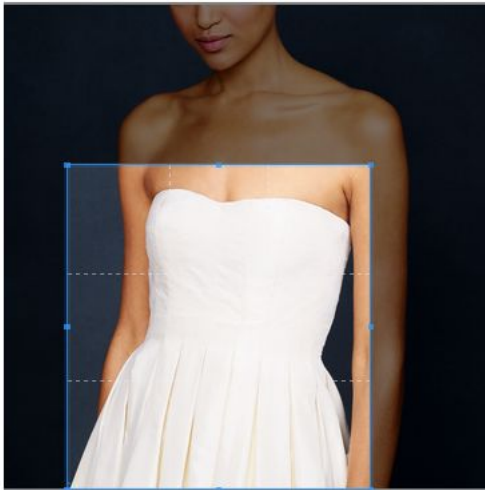
JD.京东.COM



唐狮 (TonLion) 女基本木耳边雪纺
无袖连衣裙 615221104445 钴蓝 M



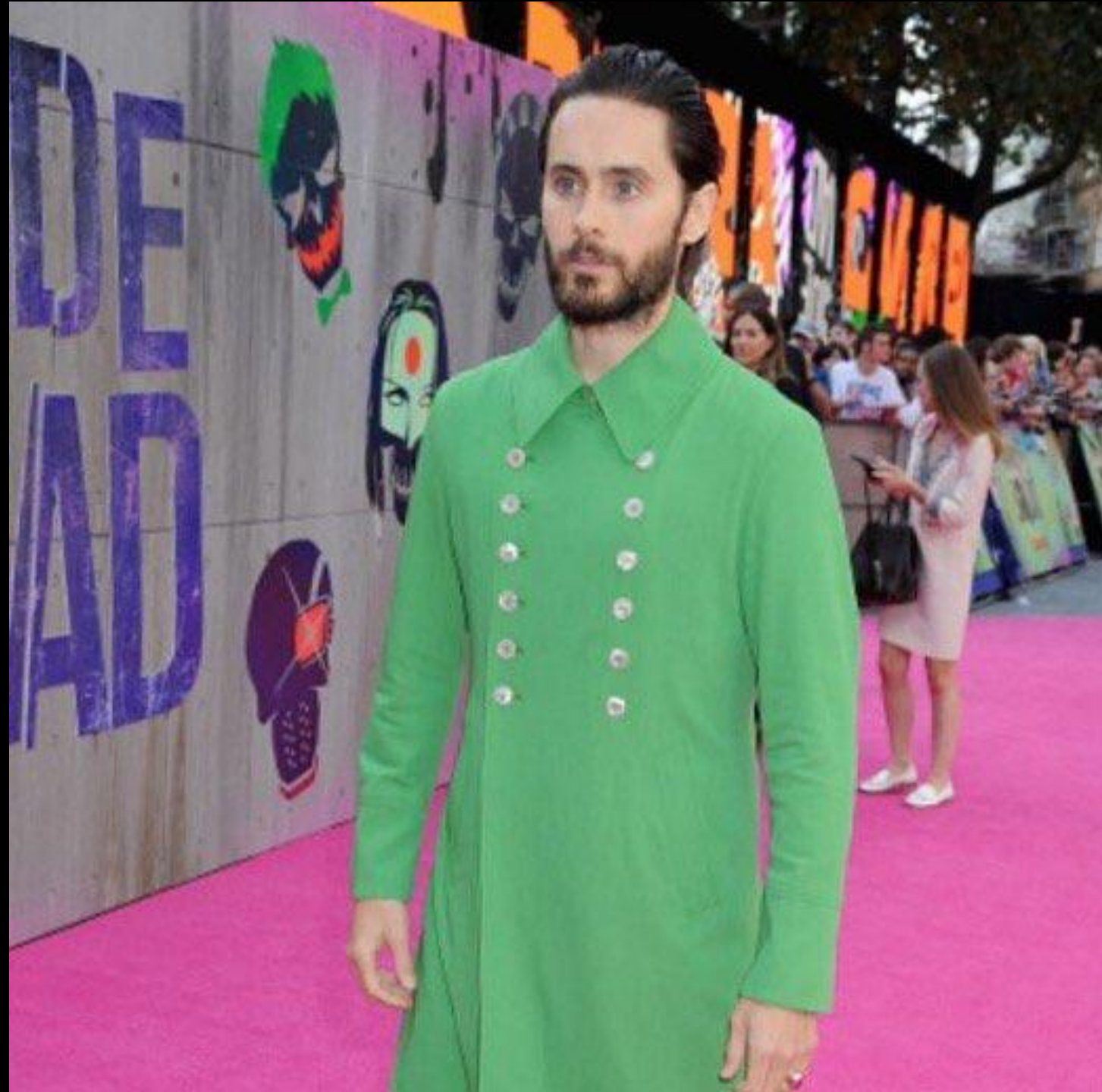
The region selected below will be used as the search query.





<https://twitter.com/broodybats/status/760906103659991040>





<https://www.facebook.com/neverdraw>



ROZETKA
интернет-супермаркет

Большой
выбор
платьев



Перейти →



ROZETKA
интернет-супермаркет



Выбрать →



ROZETKA
интернет-супермаркет



Выбрать →

Links & Used materials

<https://chatbotslife.com/small-u-net-for-vehicle-detection-9eec216f9fd6>

https://people.eecs.berkeley.edu/~jonlong/long_shelhamer_fcn.pdf

https://www.youtube.com/watch?v=ij-p_5oppl4&feature=youtu.be

<http://vision.is.tohoku.ac.jp/~kyamagu/research/chic-or-social/>

#



<https://www.slideshare.net/KotaYamaguchi1/computer-vision-meets-fashion>

<https://www.slideshare.net/MateuszOpala/6-pydata-warsaw-deep-learning-for-image-segmentation>

<https://medium.com/@fareeha/fashion-x-ml-6958cbfe7d4b>

#

Fashion x ML

A resource list for the latest papers and datasets in:

- [Style Recommendation / Trend Forecasting](#)
- [Clothing Recognition and Retrieval](#)
- [Feature Learning from clothing images](#)

<https://www.amazon.com/Amazon-Echo-Look-Camera-Style-Assistant/dp/B0186JAEWK>

<http://streetwearflair.com/vetements-total-fucking-darkness-jacket-black>

https://en.wikipedia.org/wiki/%C3%81o_d%C3%A0i

<https://arxiv.org/pdf/1711.02231.pdf>

<https://medium.com/retail-vuepoints/shoppable-video-the-next-big-digital-retail-opportunity-9d0036b6233c>

#

<https://medium.com/mlreview/how-to-apply-distance-metric-learning-for-street-to-shop-problem-d21247723d2a>

https://github.com/movchan74/street_to_shop_experiments

Demo: <http://vps389544.ovh.net:5555/>

#

<https://researchweb.iiit.ac.in/~praveen.krishnan/compreExam/slides/T4-DistanceMetricLearningBeyond0-1Loss.pdf>

<http://cs231n.stanford.edu/reports/2017/pdfs/108.pdf>

#

<https://github.com/ayushidalmia/awesome-fashion-ai>

