Learning about fashion from web photos

Kristen Grauman
Facebook AI Research
University of Texas at Austin

Fashion: what people wear

Fashion and culture

Postwar simplicity

Manufacturing fabrics

WWII austerity

Feminism

Anti-war

Women in the workplace

1920 1930 1940 1950 1970 1980

Vision + fashion: influence

Design & creativity

Business

amazon

ebay

Walmart

Etsy

YvesSaintLaurent

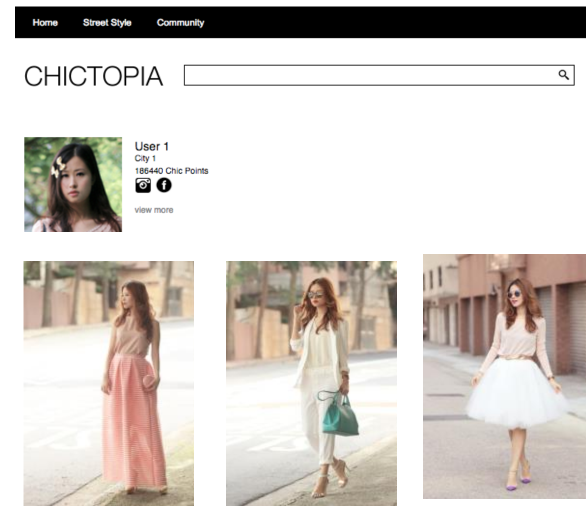
How we shop

Vision + fashion: problems

Fashion introduces new challenges for high-level vision:

Subtle distinctions

Composition and compatibility



Personalization and taste

Requires computational models for *style*

This talk: fashion & web photos

- Assembling fashionable outfits
  - What goes with what?
  - How to compose a wardrobe?
  - How could this outfit look better?

- Learning subtle attributes
  - How to distinguish slight differences?

How to learn visual compatibility?

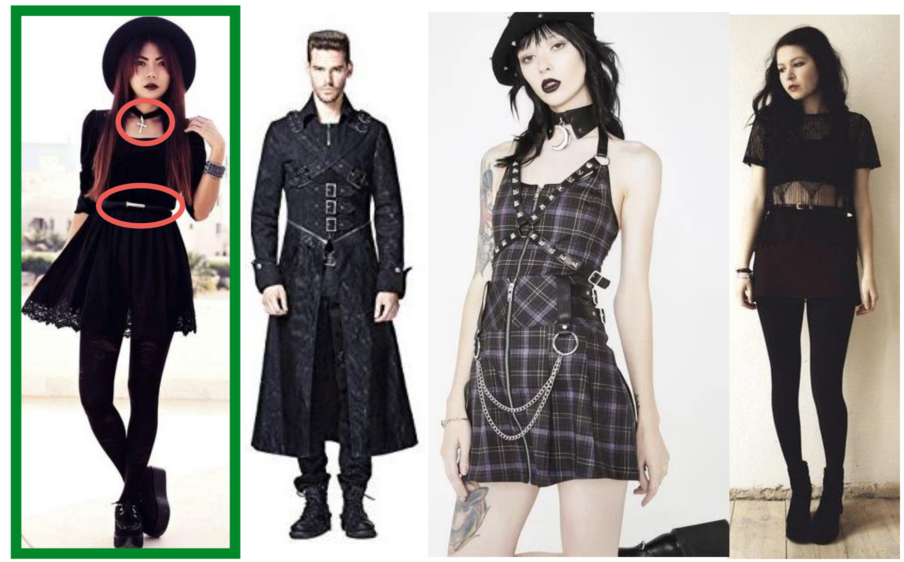
Co-purchase data
[McAuley 2015, Veit 2015, He 2016]

Manual curation
[Li 2017, Song 2017, Han 2017]

Supervised

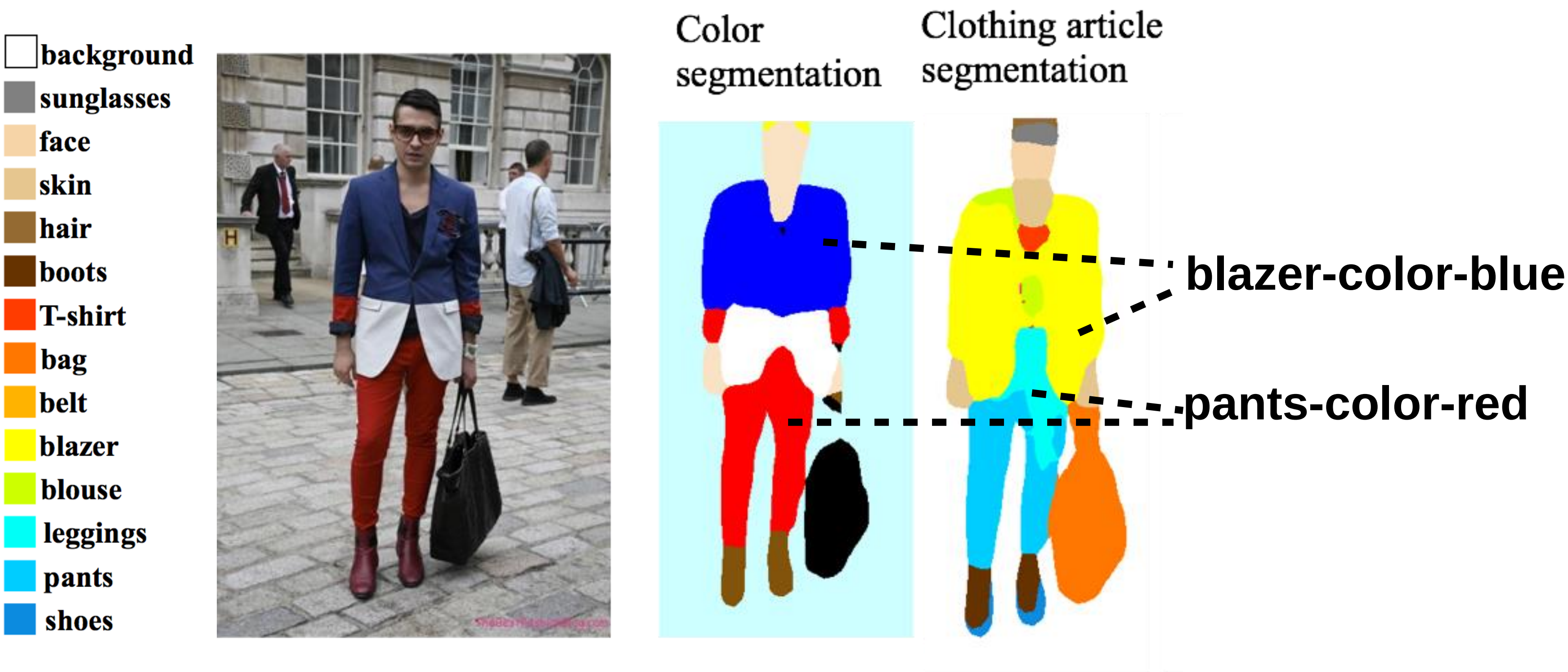
Our goal: Unlabeled in the wild photos

Attributes as style elements



Style: underlying compositions of elements.

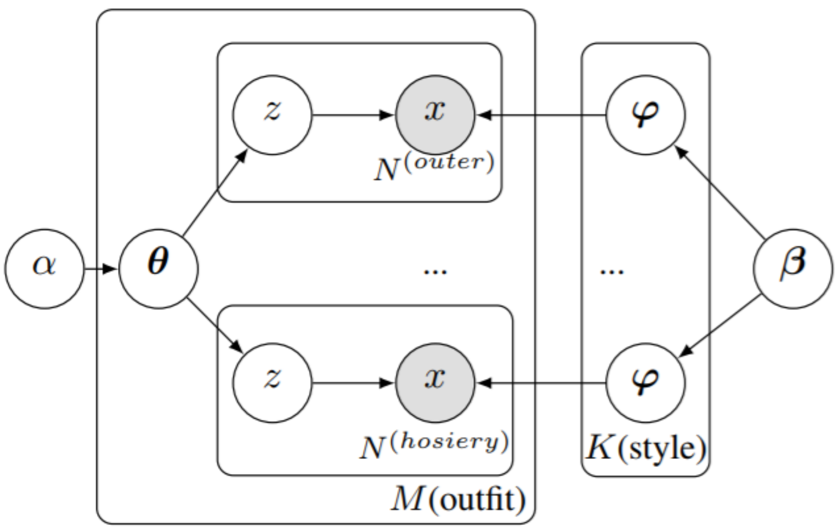
Attributes as style elements



- **Material, cut, pattern**
  - Fine-tune classification on ResNet50
- **Color, clothing article**:
  - Segmentation on DeepLab-DenseCRF

Learning styles from web photos

Unsupervised learning of a "visual style topics" with a polylingual topic model



outer layer
outer_color_orange
outer_color_white
outer_pattern_printed
outer_decoration_button
outer_sleeve_long
outer_length_short
outer_front_open
⋮
upper
shirt_color_white
shirt_pattern_plain
shirt_sleeve_short

An **outfit** is a mixture of (latent) **styles.**
A **style** is a distribution over **attributes.**

Mimno et al. "Polylingual topic models." EMNLP 2009.

Hsiao & Grauman, ICCV 2017

Example discovered styles (dresses)

sheath
knit
shift
sleeveless
bodycon
textured
stretch

chiffon
maxi
pleated
red
chiffon maxi
beaded
sleeveless

striped
stripe
knit
stripes
mini
midi
sleeve

denim
chambray
drawstring
classic
utility
button
wash

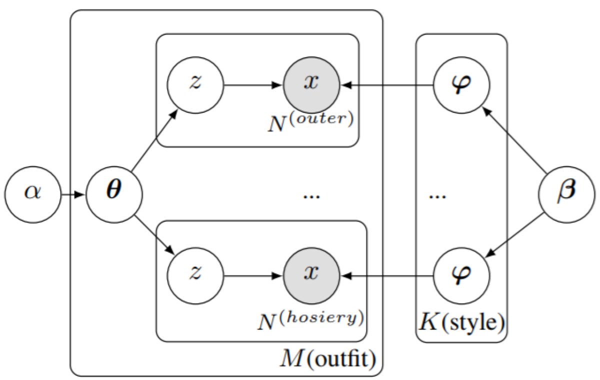
Styles we automatically discover in the **Amazon** dataset [McAuley et al. 2015]

Example discovered styles (full outfit)

Styles we automatically discover in the **HipsterWars** dataset [Kiapour et al]

Visual compatibility

Calculate compatibility of garments via likelihood under topic model



$$c(o_j) := p(o_j | \boldsymbol{\mu}, \boldsymbol{\Sigma}, \beta)$$

An **outfit** is a mixture of (latent) **styles.**
A **style** is a distribution over **attributes.**

Hsiao & Grauman, CVPR 2018

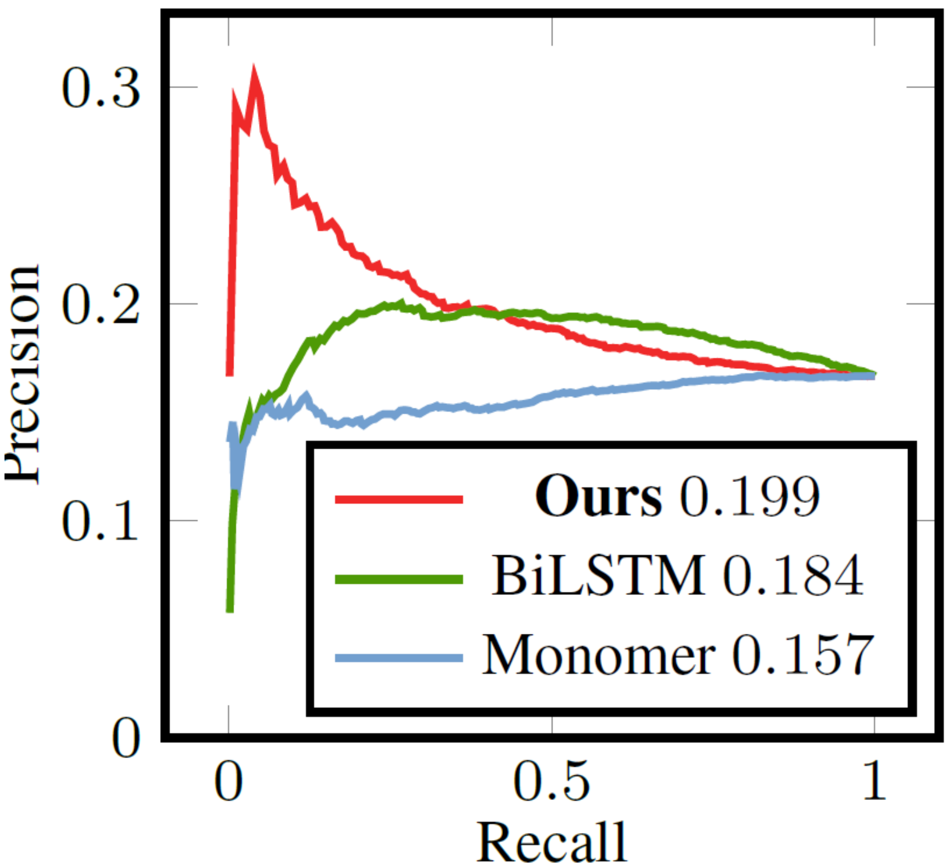
Visual compatibility results

Most compatible

Visual compatibility results

Least compatible

Visual compatibility results



BiLSTM [Han et al. 17]: unsupervised sequential model trained on Polyvore sets.

Monomer [He et al. 16]: supervised embedding trained on Amazon products co-purchase info.

Encouraging results for learning compatibility from unlabeled, full-body Web images

Hsiao & Grauman, CVPR 2018

Creating a "capsule" wardrobe

Goal: Select minimal set of pieces that mix and match well to create many viable outfits

Creating a “capsule” wardrobe

Pose as *subset selection* problem

set of garments = argmax **compatibility** + **versatility**

Hsiao & Grauman, CVPR 2018

Capsule via subset selection

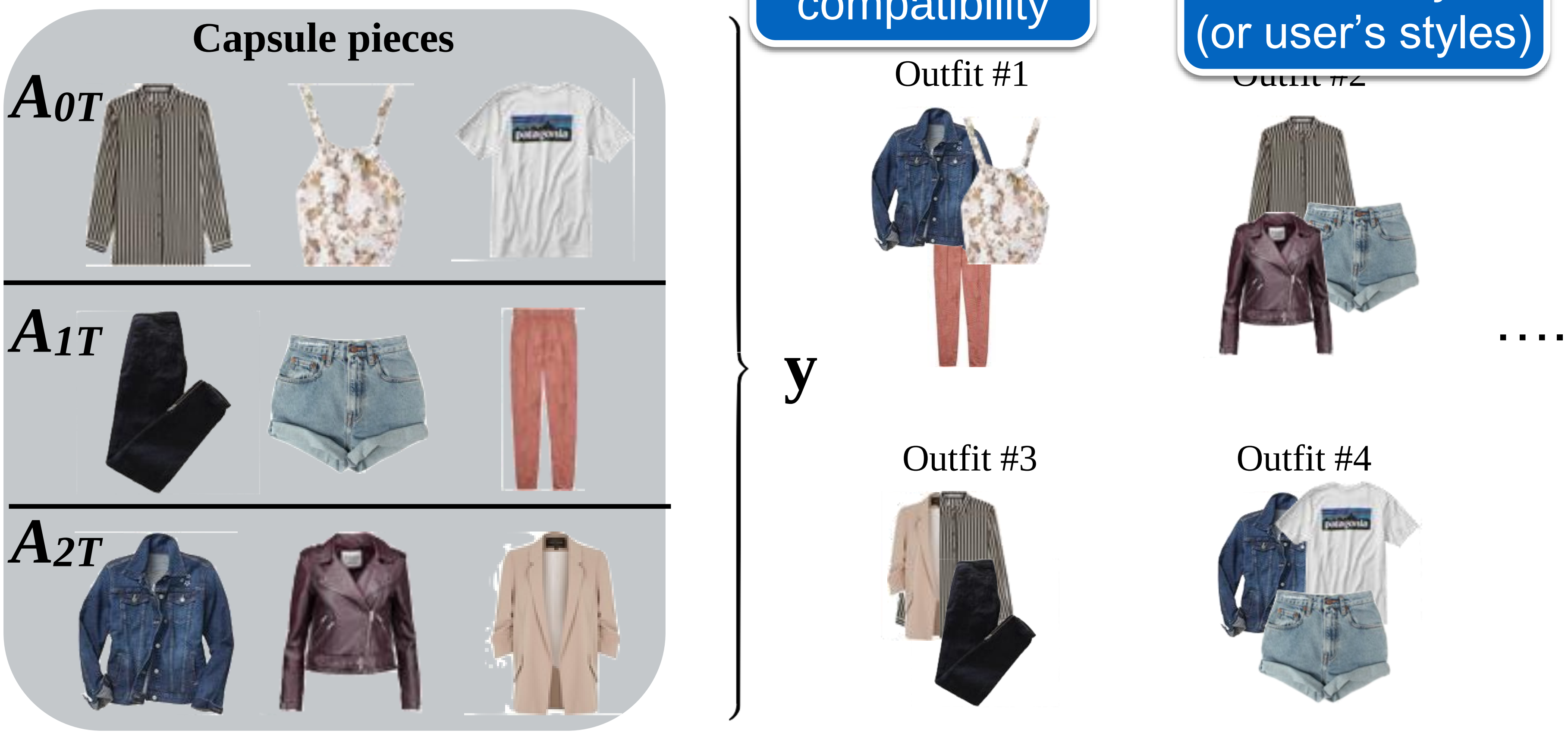
optimal set of composed outfits

…se as *subset selection* problem

…set of garments = argmax compat…ity + **ver…tility**

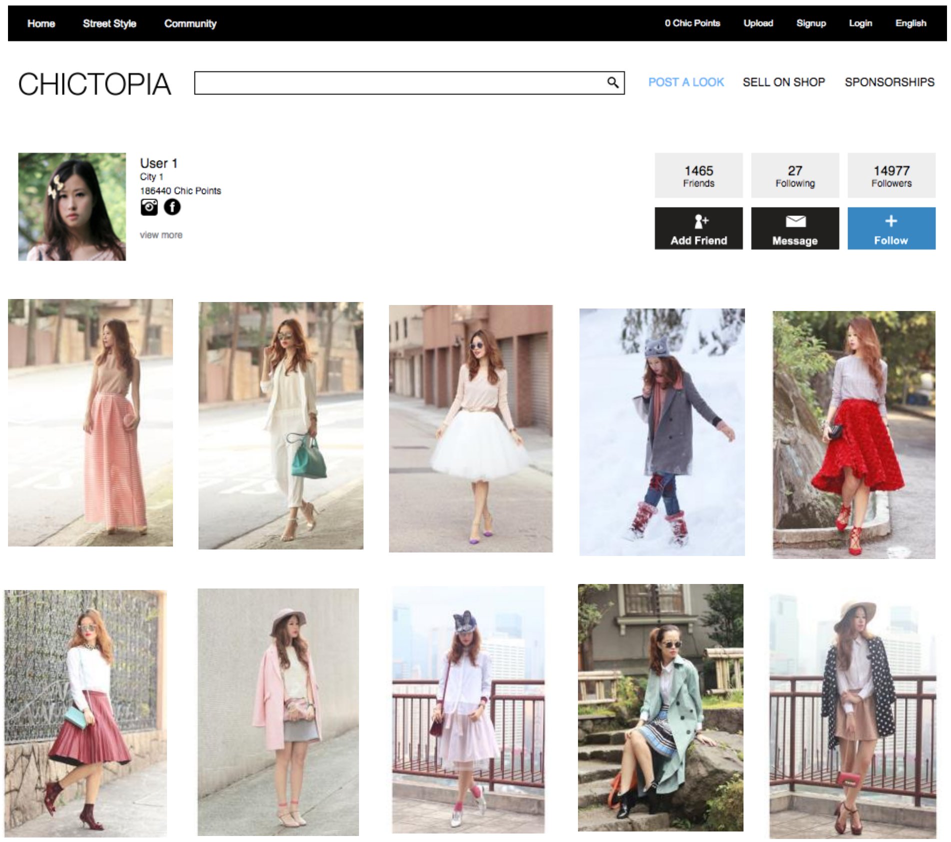
increase outfit compatibility

cover all styles (or user's styles)



Personalized capsule example

Discover user's style preferences from album

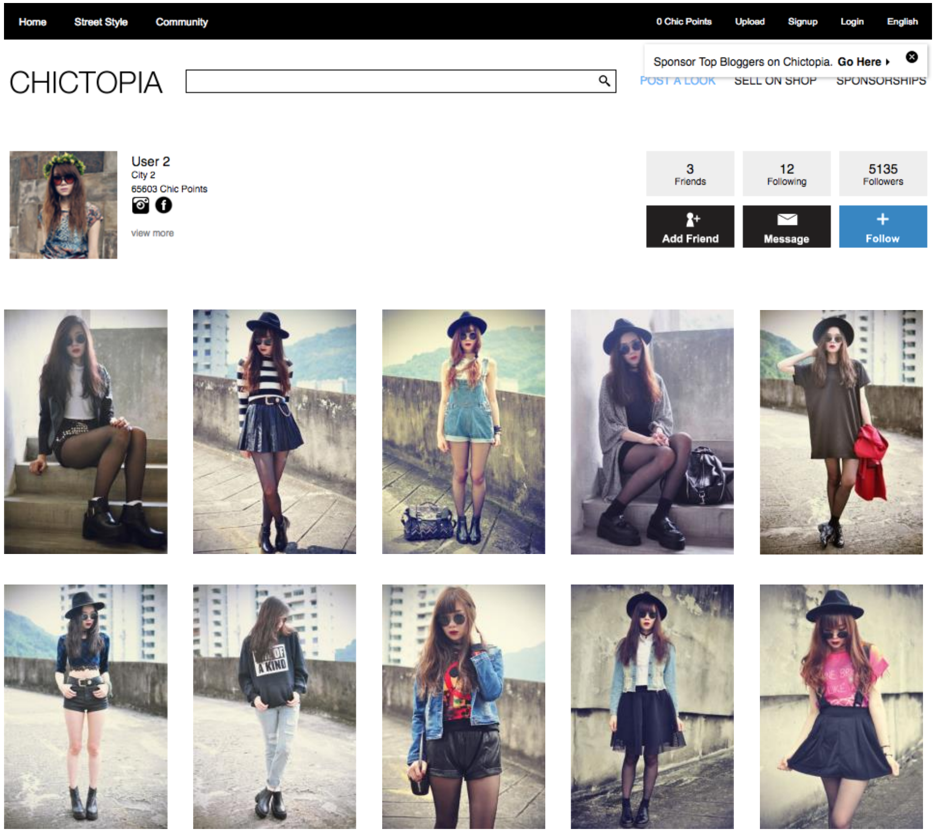


Personalized capsule

Hsiao & Grauman, CVPR 2018

Personalized capsule example

Discover user's style preferences from album



Personalized capsule

Hsiao & Grauman, CVPR 2018

This talk: fashion & web photos

- Assembling fashionable outfits
  - What goes with what?
  - How to compose a wardrobe?
  - How could this outfit look better?
- Learning subtle attributes
  - How to distinguish slight differences?

Idea: Minimal edits for outfit improvement

[Hsiao et al. Fashion++, arXiv 2019]

Fashion++

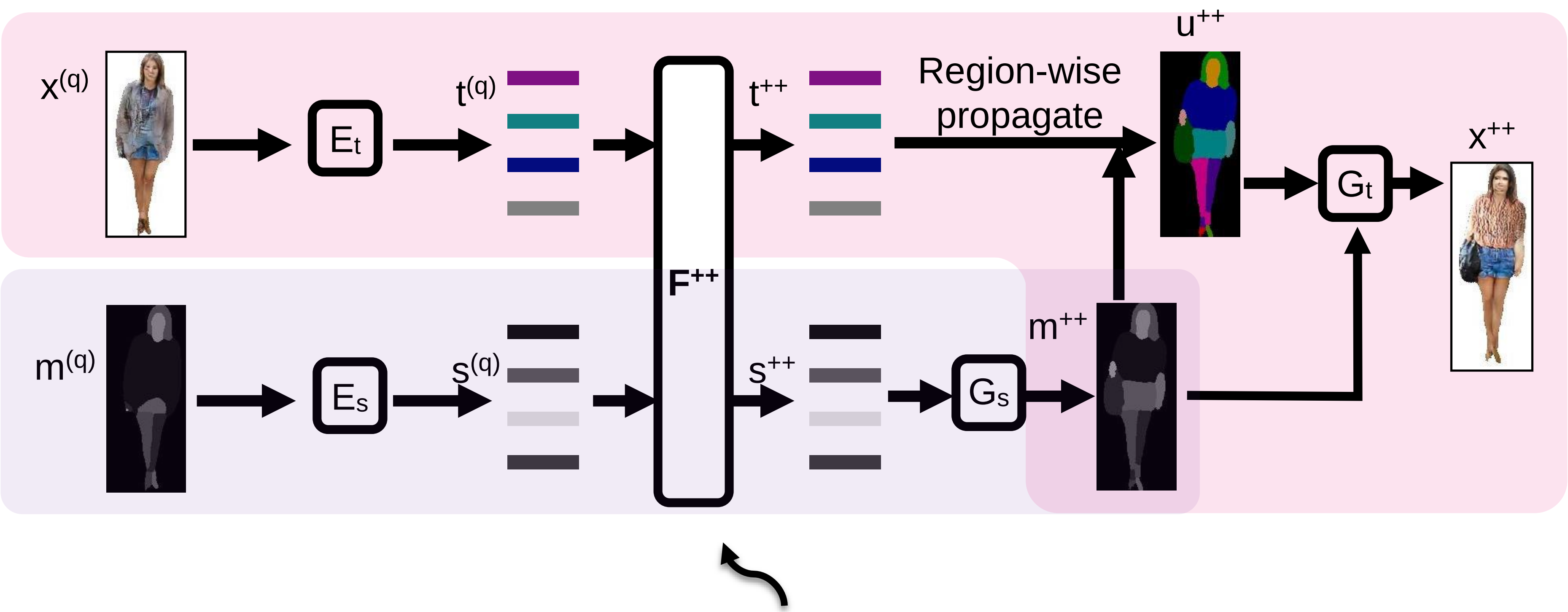
[Hsiao et al. Fashion++, arXiv 2019]

Fashion++

[Hsiao et al. Fashion++, arXiv 2019]

Our approach: Fashion++

Image generation pipeline factorizes shape and texture



Edit via activation maximization with discriminative fashionability model

[Hsiao et al. Fashion++, arXiv 2019]

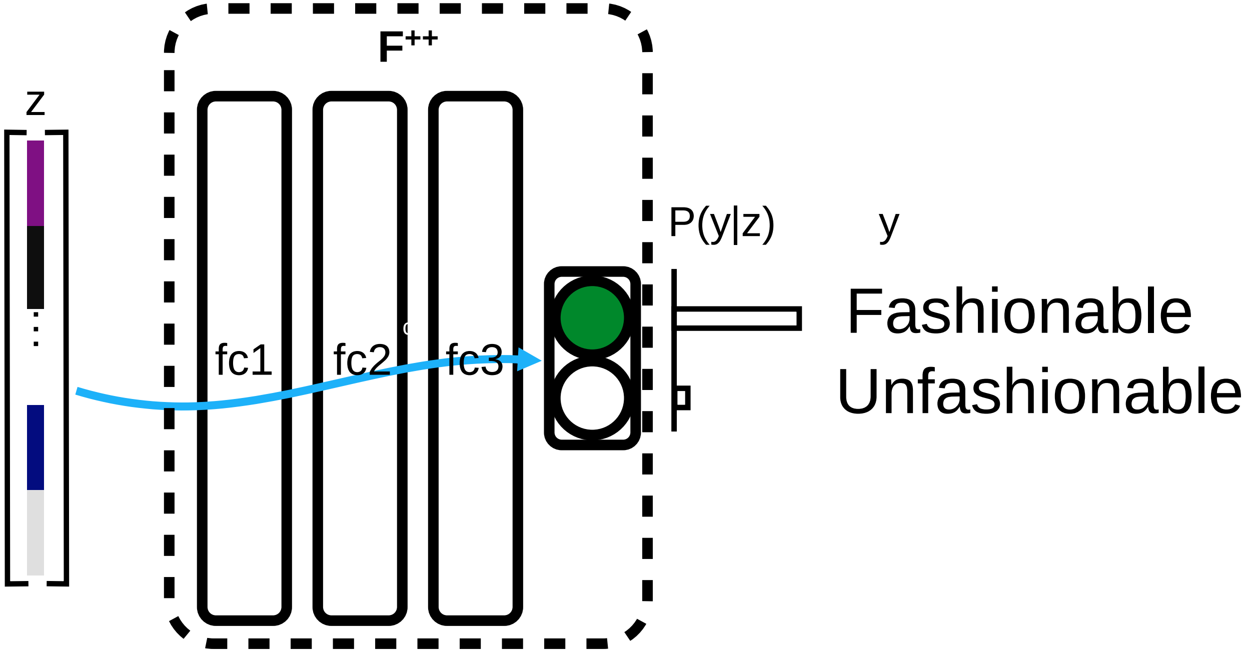
Learning discriminative fashionability model

Represent an outfit as concatenation of its latent codes

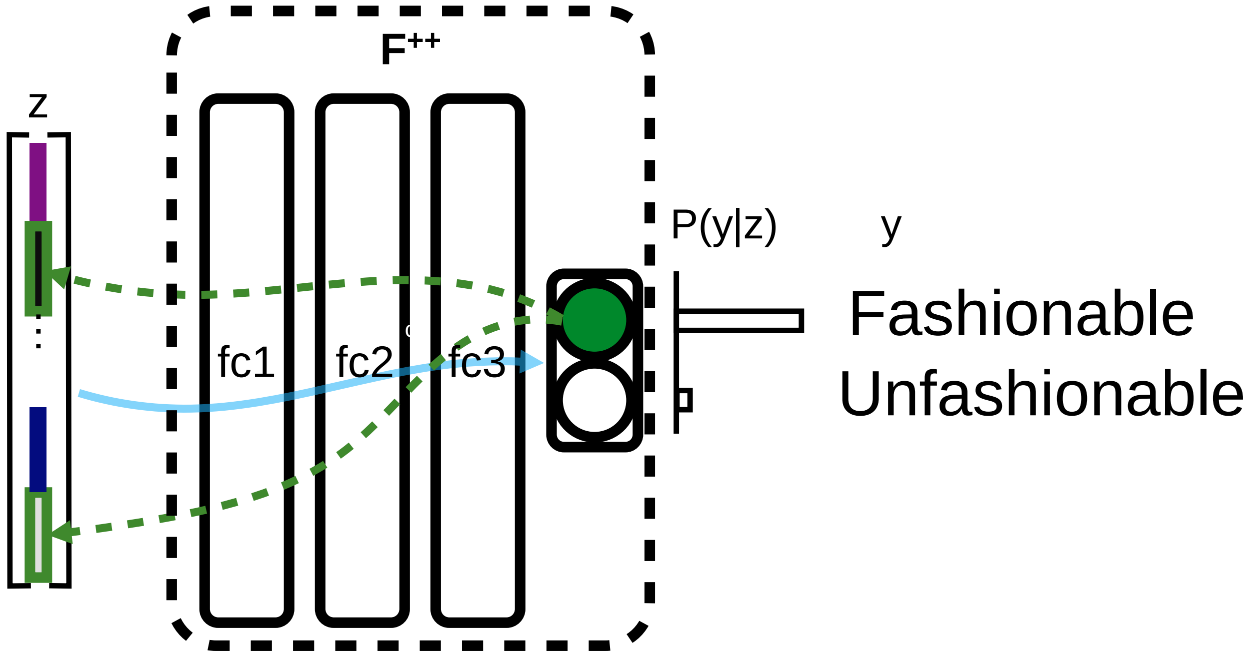
Learning discriminative fashionability model

Bootstrap web photos for "negatives" to learn fashionability

Computing an outfit edit



Computing an outfit edit



Propagate gradients to iteratively update the target feature (e.g., shape)

Fashion++ minimal edits

Visualization of **changed garment** in **new presentation.**





Swap to garment from inventory.





[Hsiao et al. Fashion++, arXiv 2019]

Fashion++ minimal edits

Change pants fit

Change pants length

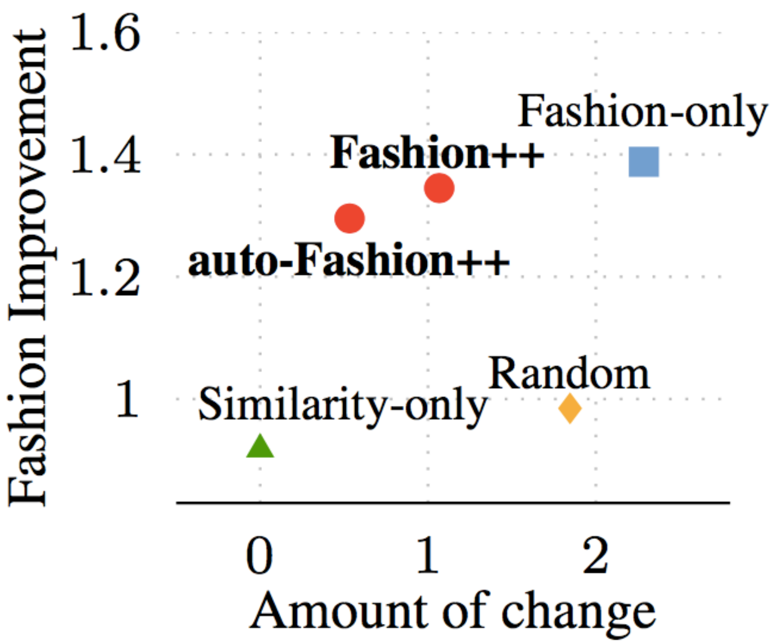
Tuck shirt in

Blouse shirt out

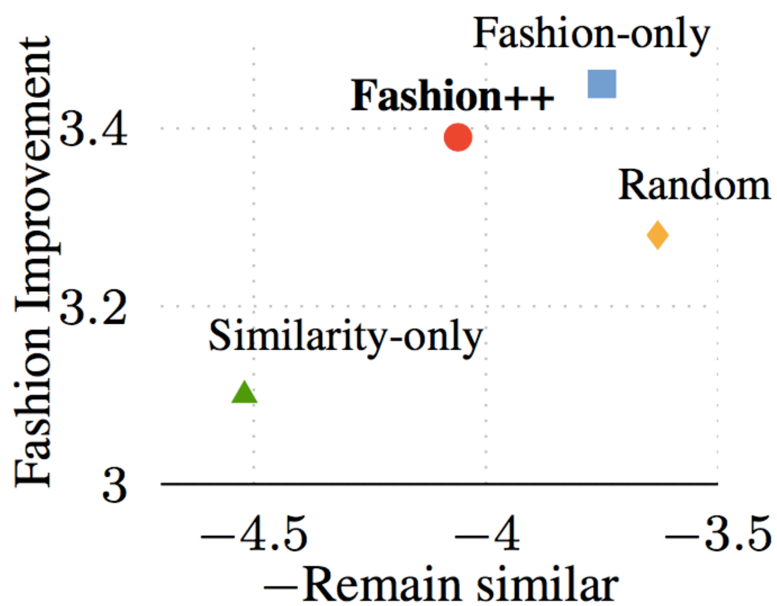
[Hsiao et al. Fashion++, arXiv 2019]

Fashion++ minimal edits

Automatic evaluation



Human subject study



Fashion++ balances best by improving fashionability while not changing too much.

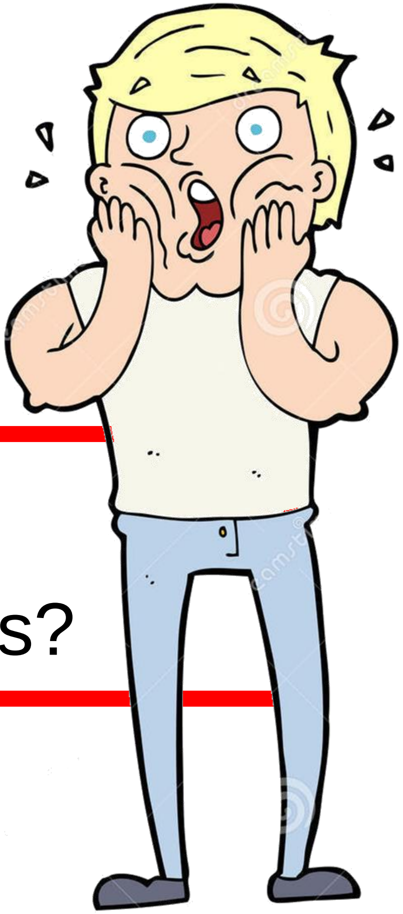
[Hsiao et al. Fashion++, arXiv 2019]

This talk: fashion & web photos

- Assembling fashionable outfits
  – What goes with what?
  – How to compose a wardrobe?
  – How could this outfit look better?

- Learning subtle attributes
  – How to distinguish slight differences?

The limits of web photos!



Fine-grained attribute comparisons

Which is *more sporty*?

vs.

Which is *more flirty*?

vs.

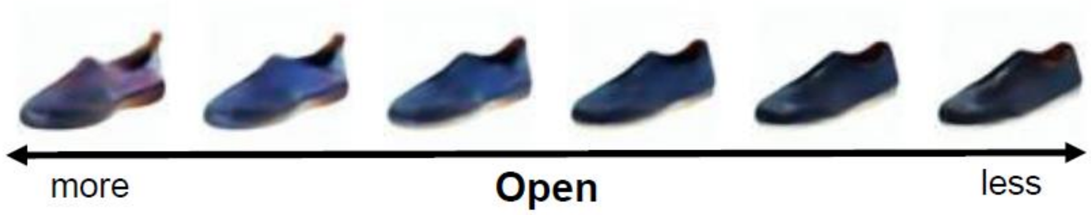
Goal: Subtle visual comparisons

Challenge: Curating training image pairs

Web photos and the streetlight effect

Idea: Semantic jitter

Overcome sparsity of available fine-grained image pairs with attribute-conditioned image generation



Images generated by Yan et al. 2016 Attribute2Image CVAE approach

Yu & Grauman, ICCV 2017

Idea: Semantic jitter

Overcome sparsity of available fine-grained image pairs with attribute-conditioned image generation

vs.

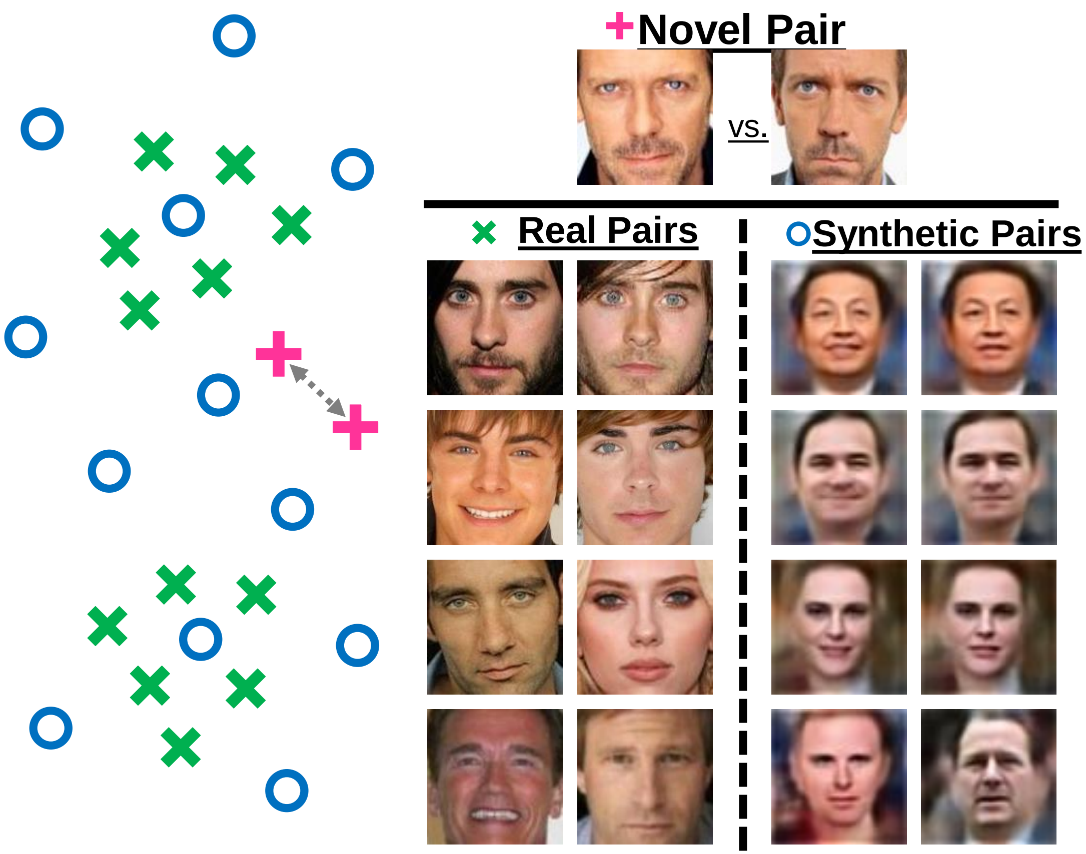
Status quo:
Low-level jitter

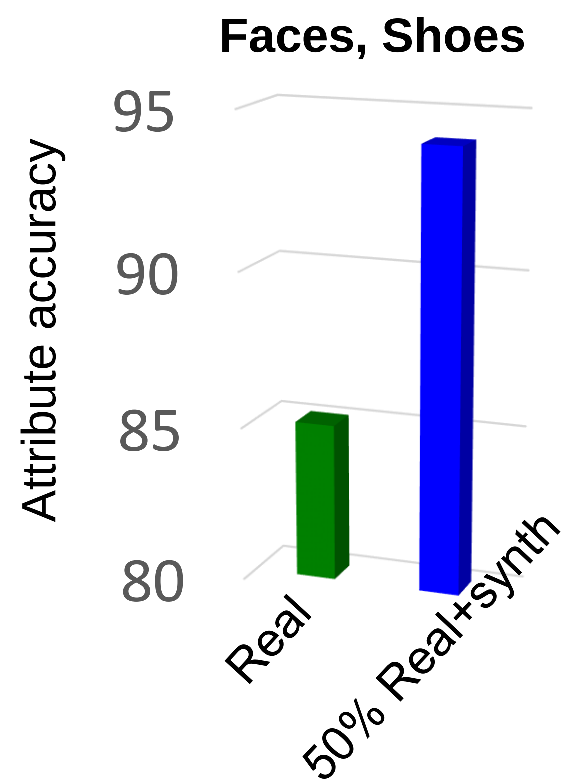
Our idea:
Semantic jitter

Yu & Grauman, ICCV 2017

Semantic jitter for attribute learning

Train rankers with both real and synthetic image pairs, test on real fine-grained pairs.



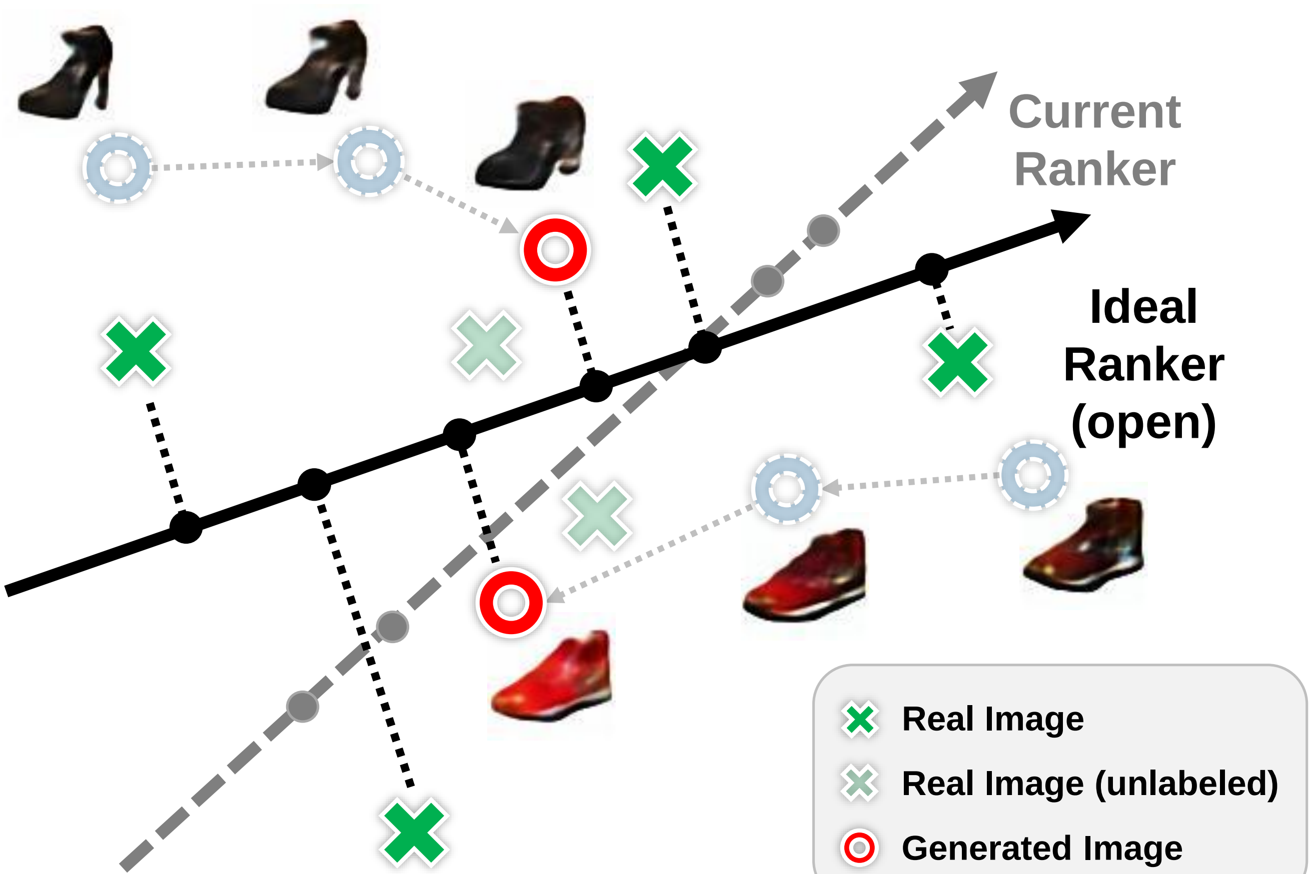


Ranking functions trained with deep spatial transformer ranking networks [Singh & Lee 2016] or Local RankSVM [Yu & Grauman 2014]

Yu & Grauman, ICCV 2017

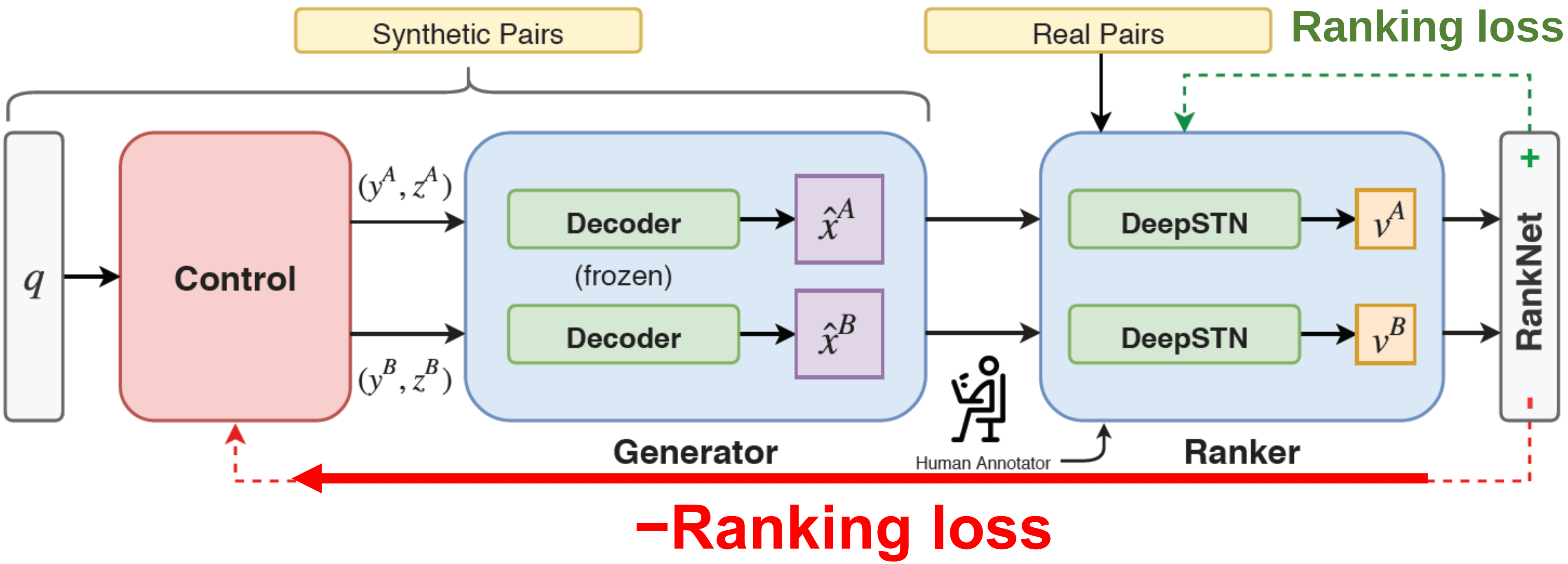
Idea: Active training image creation

Actively generate best image pairs for human labeling



Yu & Grauman, CVPR 2019

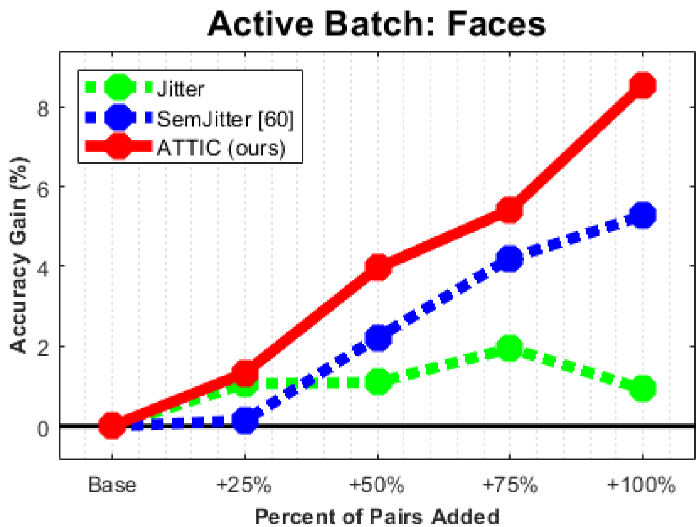
Active training image creation

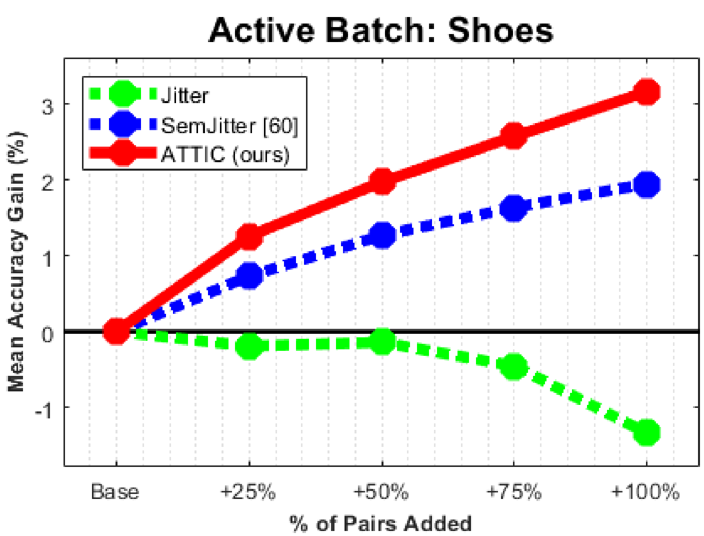


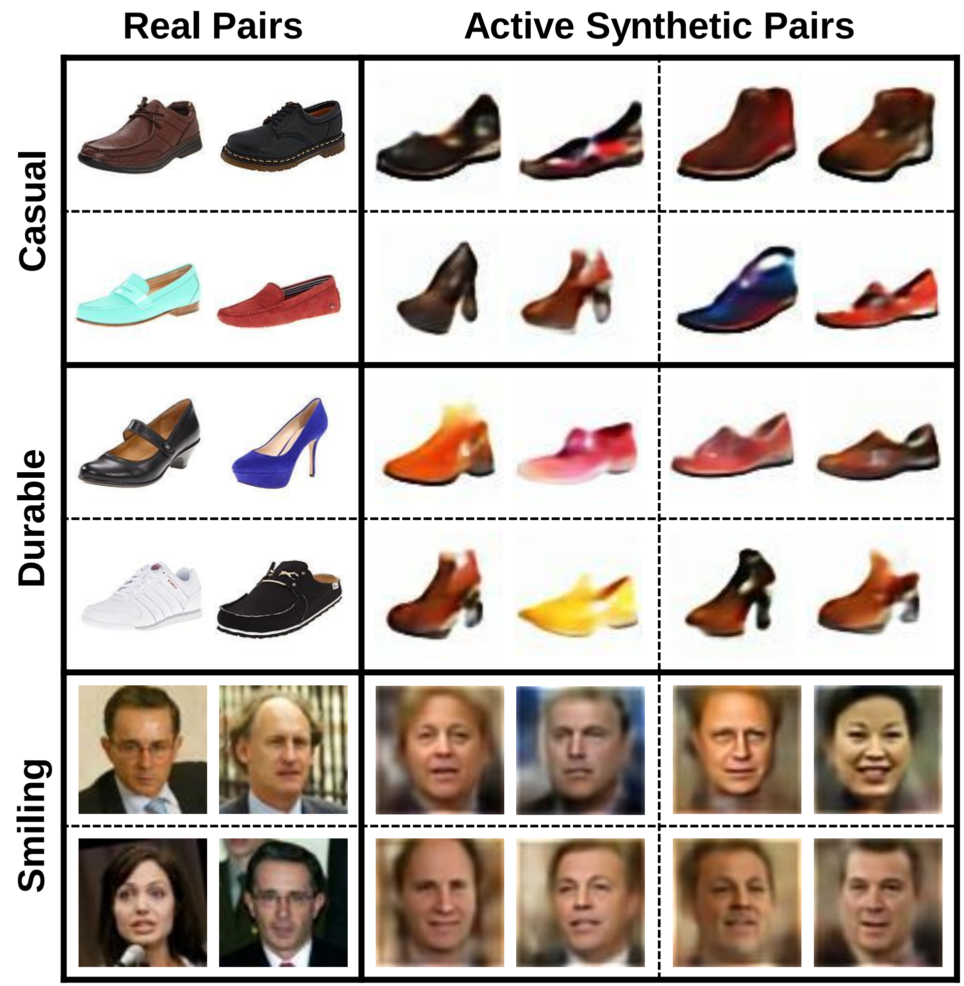
System "imagines" image pairs that would confuse current ranking model

Yu & Grauman, CVPR 2019

Active training image creation







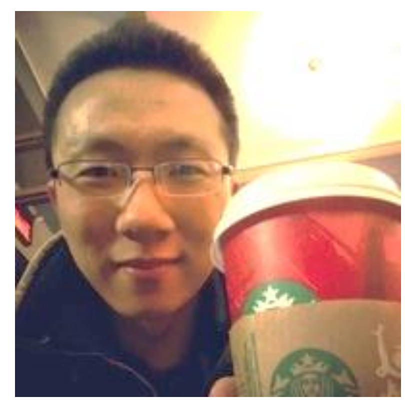
Actively curating synthetic training images → more accurate model

Summary

Kimberly Hsiao

- Learning styles and fashionability from web photos
- New ideas and methods for:
  - Style and compatibility
  - Capsule wardrobe creation
  - Incremental outfit improvement
  - Subtle visual comparisons

Aron Yu



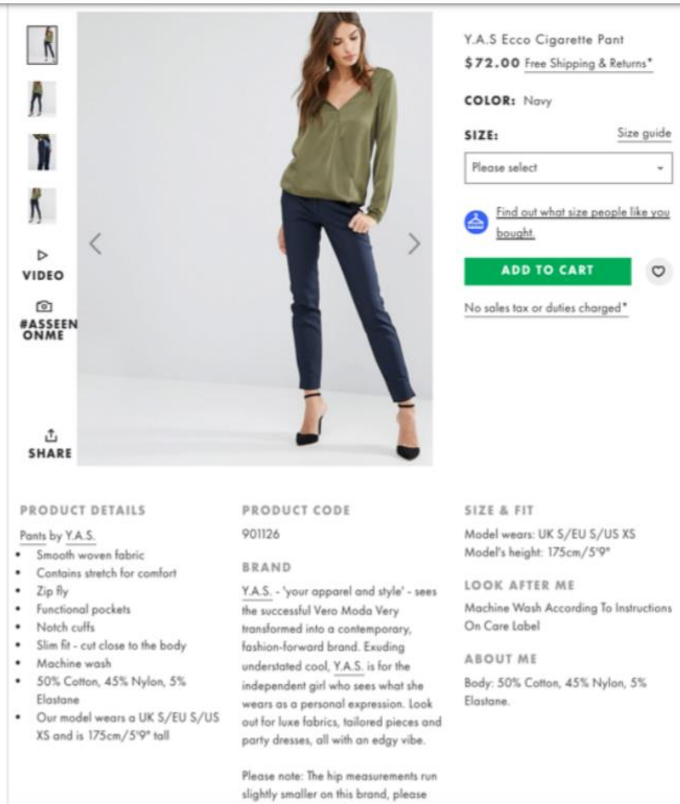
Kristen Grauman
FAIR & UT Austin

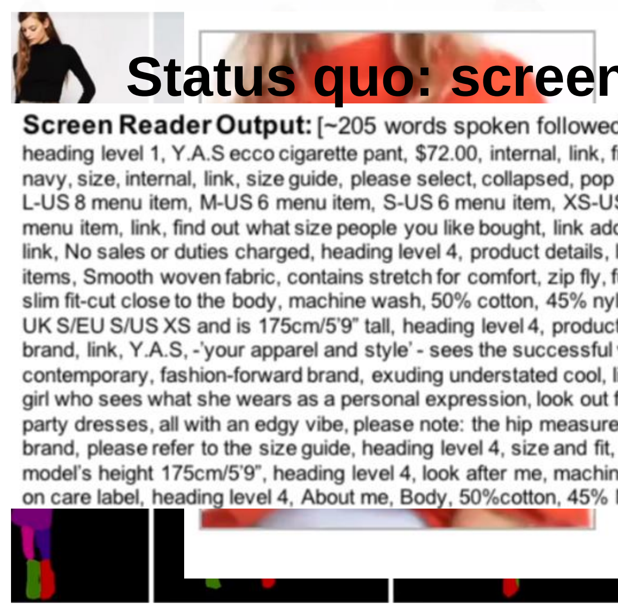
Papers

- **Creating Capsule Wardrobes from Fashion Images**. W-L. Hsiao and K. Grauman. In Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Salt Lake City, June 2018.
- **Thinking Outside the Pool: Active Training Image Creation for Relative Attributes.** A. Yu and K. Grauman. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Long Beach, CA, June 2019.
- **Semantic Jitter: Dense Supervision for Visual Comparisons via Synthetic Images**. A. Yu and K. Grauman. In Proceedings of the International Conference on Computer Vision (ICCV), Venice, Italy, Oct 2017.
- **Learning the Latent "Look": Unsupervised Discovery of a Style-Coherent Embedding from Fashion Images**. W-L. Hsiao and K. Grauman. In Proceedings of the International Conference on Computer Vision (ICCV), Venice, Italy, Oct 2017.
- **Fashion++: Minimal Edits for Outfit Improvement**. Wei-Lin Hsiao, Isay Katsman, Chao-Yuan Wu, Devi Parikh, Kristen Grauman. arXiv:1904.09261v1. April 2019.

BrowseWithMe

Make online shopping more accessible for visually impaired users





Status quo: screen reader

Screen Reader Output: [~205 words spoken followed by this product description]: heading level 1, Y.A.S ecco cigarette pant, $72.00, internal, link, free shipping and returns, color, navy, size, internal, link, size guide, please select, collapsed, pop up button, XL-US 10 menu item, L-US 8 menu item, M-US 6 menu item, S-US 6 menu item, XS-US 2 menu item, Please select, menu item, link, find out what size people you like bought, link add to cart, internal, link, 6987053, link, No sales or duties charged, heading level 4, product details, link Pants, by, link, Y.A.S, list 9 items, Smooth woven fabric, contains stretch for comfort, zip fly, functional pockets, notch cuffs, slim fit-cut close to the body, machine wash, 50% cotton, 45% nylon, 5% elastane, our model wears UK S/EU S/US XS and is 175cm/5'9" tall, heading level 4, product code, 901126, heading level 4, brand, link, Y.A.S, -'your apparel and style' - sees the successful very moda very transformed into a contemporary, fashion-forward brand, exuding understated cool, link, Y.A.S, is for the independent girl who sees what she wears as a personal expression, look out for luxe fabrics, tailored pieces and party dresses, all with an edgy vibe, please note: the hip measurements run slightly smaller on this brand, please refer to the size guide, heading level 4, size and fit, model wears UK S/EU S/US XS, model's height 175cm/5'9", heading level 4, look after me, machine wash according to instructions on care label, heading level 4, About me, Body, 50%cotton, 45% Nylon, 5% Elastane.

System: `cherry red`

https://www.youtube.com/watch?v=lUH5sWDSGuw&feature=youtu.be

Stangl et al. ASSETS 2018