StreetStyle: Exploring world-wide clothing styles from millions of photos

Kevin Matzen, Kavita Bala, and Noah Snavely
Experiments by: Angela Lin and Lakshay Narula
StreetStyle Dataset Statistics

Size:
- Total: 27084 entries
- 23908 Distinct Instagram Posts

Fields Associated With Each Image:
- City ID
- Month ID
- Upload Time
- Upload Location
- Bounding Box Coordinates
- Image Size
- Attribute Annotations
### Cities in the dataset

- Bangkok
- Beijing
- Berlin
- Bogotá
- Budapest
- Buenos Aires
- Cairo
- Delhi
- Dhaka
- Guangzhou
- Istanbul
- Jakarta
- Karachi
- Kiev
- Kolkata
- Lagos
- London
- Los Angeles
- Madrid
- Manila
- Mexico City
- Milan
- Moscow
- Mumbai
- New York City
- Osaka
- Paris
- Rio de Janeiro
- Rome
- SÃO Paulo
- Seoul
- Shanghai
- Sofia
- Sydney
- Tianjin
- Tokyo
- Austin
- Chicago
- Johannesburg
- Nairobi
- Seattle
- Singapore
- Toronto
- Vancouver

### Number of Samples in Each City

Median: 755 Samples
### Examples from the dataset

<table>
<thead>
<tr>
<th>Clothing Category</th>
<th>Shirt</th>
<th>Outerwear</th>
<th>T-shirt</th>
<th>Dress</th>
<th>Tank top</th>
<th>Suit</th>
<th>Sweater</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="shirt.png" alt="Image" /></td>
<td><img src="outerwear.png" alt="Image" /></td>
<td><img src="t-shirt.png" alt="Image" /></td>
<td><img src="dress.png" alt="Image" /></td>
<td><img src="tank-top.png" alt="Image" /></td>
<td><img src="suit.png" alt="Image" /></td>
<td><img src="sweater.png" alt="Image" /></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Neckline Shape</th>
<th>Round</th>
<th>Folded</th>
<th>V-shape</th>
<th>Sleeve Length</th>
<th>Long Sleeve</th>
<th>Short Sleeve</th>
<th>No Sleeve</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="round.png" alt="Image" /></td>
<td><img src="folded.png" alt="Image" /></td>
<td><img src="v-shape.png" alt="Image" /></td>
<td><img src="long-sleeve.png" alt="Image" /></td>
<td><img src="short-sleeve.png" alt="Image" /></td>
<td><img src="no-sleeve.png" alt="Image" /></td>
<td><img src="short-sleeve.png" alt="Image" /></td>
<td><img src="no-sleeve.png" alt="Image" /></td>
</tr>
</tbody>
</table>
Style In The Wild

- Multiple People & Occlusions
- Variation in Pose
- Varying Image Quality
- Cultural Holidays
- Variety of Sources
Co-occurrence Trends Using Conditional Probabilities

Style based

Functionality based

Reasonable Dependencies
Attributes Occur Outside of Bounding Boxes

Sleeve Length: Short Sleeve

Sleeve Length: Long Sleeve

Label: Dress Category

Label: Floral Clothing Pattern
The paper has the following theme:
1. Need large data to find trends
2. Large data tough to label
3. Label 27K images
4. Train a CNN to generate features, cluster to find trends

However…
1. We don’t have access to code
2. We don’t have access to the large dataset (>10 million images)
3. We do have access to the labeled 27K images

These set of experiments attempt to expose possible interesting trends in the dataset that could be explored!
Image Upload Trends

Slight increase in uploads on weekends!

We normalize all results by the average upload frequency to observe true trends.
Daily Trends

Higher fraction of glasses during day time might suggest about 3% presence of sunglasses!

Black color appears to be more popular at night.
Weekly Trends

Round Neckline for Weekends

Slight preference for round neckline on weekends!
Instagram uploads, population density, and median income are all positively correlated.
Instagram uploads and population density are positively correlated, but the median income is negatively correlated!
Black dominates the Manhattan region
White and multiple colors comparatively more trendy elsewhere
Black and white predominant in downtown region
Blue comparatively more popular elsewhere
T-shirts more popular elsewhere, as compared to Manhattan
Intra-city Trends (LA)

- T-shirts more popular elsewhere, as compared to downtown area
Thanks for listening!

Check out our code at:

https://github.com/angelaslin/streetstyle-experiment