Approach

Outfit for You (O4U) Dataset
- Labeling system is carefully designed
- 15,748 outfits with 5.25 incompatible physical labels on average

Fashion Convolutional Network
- Outfit Encoder: A convolutional layer to encode input outfit of size $142 	imes 85$ into $32$ feature maps of dimension $142 	imes 85$
- Multi-label Graph Convolutional Networks: Learn classifier based on label correlations
- Input space: Extracted attribute features from item images and physical label correlations represented by conditional probability.

Conditional Probability
- If ‘athletic’ is not compatible with an outfit, there is a high probability (0.709) that ‘round’ is also not compatible with this outfit.

Experiments

Quantitative Results
We compared FCN with four baselines to show its effectiveness.

<table>
<thead>
<tr>
<th>Method</th>
<th>top hours</th>
<th>top Mr.</th>
<th>Am [23]</th>
<th>FCN</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linear</td>
<td>15.47</td>
<td>63.90</td>
<td>66.47</td>
<td>76.14</td>
</tr>
<tr>
<td>SVM</td>
<td>63.41</td>
<td>63.00</td>
<td>71.96</td>
<td>70.00</td>
</tr>
<tr>
<td>ResNet [5]</td>
<td>10.05</td>
<td>31.31</td>
<td>33.17</td>
<td>70.22</td>
</tr>
<tr>
<td>Am [23]</td>
<td>9.48</td>
<td>30.20</td>
<td>31.69</td>
<td>69.68</td>
</tr>
<tr>
<td>FCN</td>
<td>15.39</td>
<td>65.53</td>
<td>70.18</td>
<td>83.48</td>
</tr>
<tr>
<td>mAP</td>
<td>76.29</td>
<td>70.92</td>
<td>77.62</td>
<td>76.35</td>
</tr>
</tbody>
</table>

Fashion Compatibility
- For example, if the user is wearing a pair of pants with high waist and an outfit with a skirt, a lower waist height will deteriorate the performance.
- Using too many kernels cannot dramatically improve performance, and it may lead to overfitting.

Qualitative Results
- Different customers with various appearances
- Defined physical information

Overview
Unlike previous research [4,10] focused on learning the fashion compatibility of an outfit, we introduce this task, namely fashion cognitive learning, which targets learning the compatibility relationships between outfits and personal physical information. A new outfit dataset with tremendous personal physical information and a new end-to-end framework called Fashion Convolutional Network is proposed to tackle this task. Through extensive experiments, our network outperforms several alternative methods with clear margins. The text in red is the wrong prediction. FCN precisely predicts all incompatible body shapes for the given outfit.

Dress Well via Fashion Cognitive Learning
Kaicheng Pang1,2, Xingxing Zou1,2, Waikeung Wong1,2
1Laboratory for Artificial Intelligence in Design, School of Fashion and Textiles, The Hong Kong Polytechnic University
2Laboratory for Artificial Intelligence in Design, School of Fashion & Textiles in Above 5-10cm

References

Ablation Study
1. Effect of filter region size. Only using one kind of convolutional filter size shows the worst performance. Using filters with a big region size has a negative effect on model performance. Using multiple filters with the same size achieves lower than FCN on the top-3 labels. The combination (1, 2, 4, 6) shows the best performance on CF1 and Top-3 metrics.

2. Effect of numbers of kernels for each filter. The performance achieves the best results when the number of kernels is 24. Using too few convolutional kernels will deteriorate performance significantly. Using too many kernels cannot dramatically improve performance, and it may cause a negative impact on recall metrics.

3. Effect of numbers of GCN layers. Deeper multi-layer GCNs degrade the performance on CF1 and Top-3 metrics. Using the number of kernels is 24. Using too few convolutional kernels will deteriorate performance significantly. Using too many kernels cannot dramatically improve performance, and it may cause a negative impact on recall metrics.