FANCY: Human-centered, Deep Learning-based Framework for Fashion Style Analysis

Youngseung Jeon
Ajou University
Suwon, Republic of Korea
jeonyoungs@ajou.ac.kr

Seungwan Jin
Ajou University
Suwon, Republic of Korea
jin6491@ajou.ac.kr

Kyungsik Han*
Ajou University
Suwon, Republic of Korea
kyungsikhan@ajou.ac.kr

ABSTRACT

Fashion style analysis is of the utmost importance for fashion professionals. However, it has an issue of having different style classification criteria that rely heavily on professionals' subjective experiences with no quantitative criteria. We present FANCY (Fashion Attributes detectioN for Clustering stYle), a human-centered, deep learning-based framework to support fashion professionals' analytic tasks using a computational method integrated with their insights. We work closely with fashion professionals in the whole study process to reflect their domain knowledge and experience as much as possible. We redefine fashion attributes, demonstrate a strong association with fashion attributes and styles, and develop a deep learning model that detects attributes in a given fashion image and reflects fashion professionals' insight. Based on attribute-annotated 302,772 runway fashion images, we developed 25 new fashion styles (FANCY dataset [1]). We summarize quantitative standards of the fashion style groups and present fashion trends based on time, location, and brand.

CCS CONCEPTS

• **Computing methodologies** → **Machine learning approaches**; • **Applied computing**; • **Human-centered computing**;

KEYWORDS

Fashion, Deep learning, Human-centered AI, Quantitative fashion trend analysis, Model application

ACM Reference Format:
Youngseung Jeon, Seungwan Jin, and Kyungsik Han. 2021. FANCY: Human-centered, Deep Learning-based Framework for Fashion Style Analysis. In *Proceedings of the Web Conference 2021 (WWW '21), April 19–23, 2021, Ljubljana, Slovenia.* ACM, New York, NY, USA, 12 pages. https://doi.org/10.1145/3442381.3449833

1 INTRODUCTION

The fashion industry is an important part of the economy. According to a fashion business research company [2], the global market for apparel is estimated at $3 trillion USD and 2% of the world's GDP. Due to the development and proliferation of the Internet and mobile technologies, fashion-related images are increasingly being shared online, and the use and dependence of online shopping is increasing [13, 14]. Online information (e.g., image, text) to describe fashion products plays an important role in the analysis of fast-changing fashions for fashion professionals as well as in consumer purchasing decisions. Recently, with machine and deep learning techniques, fashion image data are now utilized for modeling, which enables a quantitative and objective analysis of fashion style trend. Studies have investigated the detection of clothing categories and attributes [6, 29, 40], fashion style classification [23, 36, 39], fashion trend analysis [5, 35], fashion recommendation [12, 37, 43].

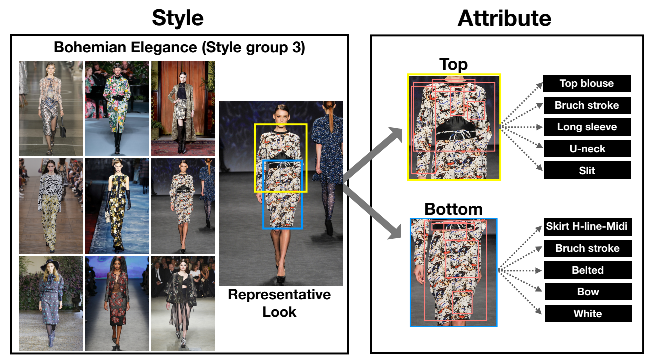

Figure 1: The hierarchical relationship between styles and attributes

Although fashion research with computing techniques gives such ample opportunities in many research domains, our interviews with fashion professionals and related literature review indicate that existing research has not sufficiently reflected fashion professionals' perspectives especially in data preparation. Fashion has two salient factors: *attributes* and *styles*. A visual attribute is naturally amenable to fashion tasks, since garments are often described by their materials, fit, and patterns (e.g., denim, polka-dotted, tight). A style is a higher-level concept of how clothing comes together (Figure 1). It refers to the overall characteristics coming out of the entire ensemble of clothes, which is a combination of attributes [1, 20]. Yet, according to the fashion professionals, prior research has the following issues: (1) Attributes defined by non-professionals (researchers) [2, 4] have an issue of fashion representativeness. (2) The number of attributes is not enough for representing fashion style [4, 34]. (3) Attributes and styles are considered as horizontal relationships [2, 29]. However, style is defined by a mix of attributes, meaning that the two concepts have a hierarchical relationship.

---
*Corresponding author
[1]The FANCY dataset is available at: https://github.com/youngseungjeon/FANCY
[2]https://bit.ly/2ulgHLU

WWW '21, April 19–23, 2021, Ljubljana, Slovenia

ACM ISBN 978-1-4503-8312-7/21/04.
https://doi.org/10.1145/3442381.3449833

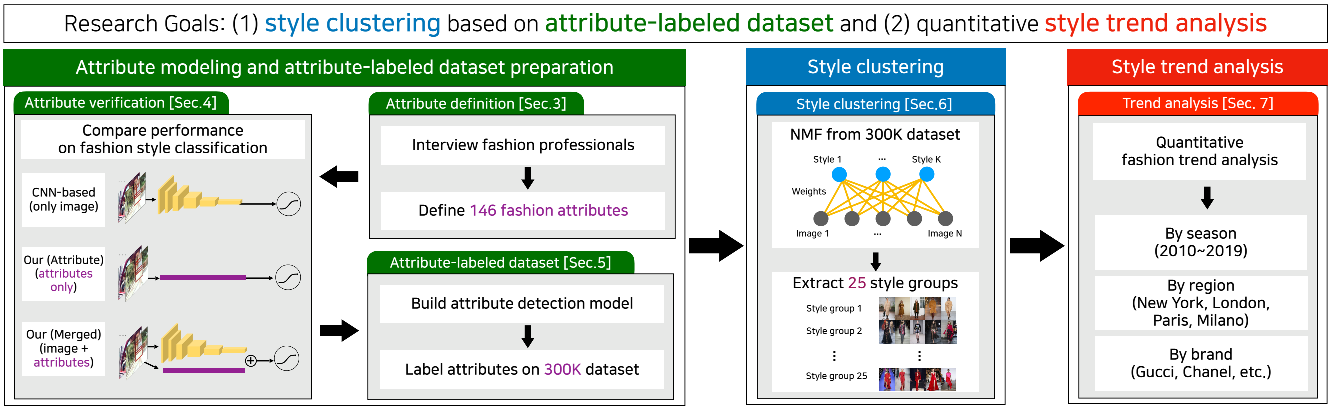

Figure 2: Study procedure, which consists of three phases: attribute dataset preparation, style clustering, and style trend analysis.

Data reliability and representation is of the utmost importance in big data analysis and modeling. Fashion style data should reflect the criteria applied in the fashion industry. Satisfying these conditions provides greater meaning for the results of analysis and improves utilization in the fashion industry. Thus, we aim to make use of the technology more relevant by *basing our models on interactions with fashion professionals* in order to (1) examine limitations on the fashion datasets used in prior studies, (2) identify core attributes that define fashion styles, and (3) investigate how to utilize computational model results. Achieving these goals is important because it not only allows to develop computational models that reflect fashion professionals' needs and are applicable to support fashion professionals' tasks but also offers stakeholders (e.g., researchers, practitioners) insights on model application and allows them to elicit ideas (e.g., feature identification, engineering) for model improvement and wider application in the context of fashion industry.

In this paper, we introduce FANCY (Fashion Attribute detection for Clustering stYle), a human-centered, deep learning-based framework that has the following components (Figure 2).

- *Attribute definition (Sec. 3)*: Through interviews with fashion professionals, we gained an understanding of current processes in fashion trend analysis, identify challenges and find methods for efficient and reliable decision-making. Through analyzing the process to make a fashion design (style), we defined fashion attributes as the factor to quantify fashion styles. We summarized six fashion design steps and defined the attributes used in each step (Figure 3).
- *Attribute verification (Sec. 4)*: We proved the meaningful relationship between our attributes and styles, comparing the performance of models to classify fashion style. Our model (attribute-merged) yields higher performance than the existing state-of-the-art models of style classification (Table 2), proving the effectiveness of attributes in characterizing styles.
- *Attribute-labeled dataset (Sec. 5)* : We developed an object detection model that identifies fashion attributes from a given image (45.0% F1 score) for labeling attributes on large-scale fashion data. We applied the model and prepared a dataset from 10 years of 302,772 runway fashion images (2010 SS-2019 FW) with attribute annotations.
- *Style clustering (Sec. 6)*: Based on the detected fashion attributes, we sorted 302,772 runway images into 25 style groups (Figure 6), using non-negative matrix factorization (NMF). The name of each style was decided upon by the fashion professionals, making the results of the analysis interpretable. We built the FANCY dataset.
- *Style trend analysis (Sec. 7)*: Using the frequency of a particular style in a certain year, we quantified style popularity. We analyzed style trends by year (Figures 7, 8), region (Figure 9), and brand (Figure 10).

Our research contributes to an application of fashion professionals' knowledge and insights to prepare a dataset that reflects fashion characteristics, and provides a demonstration of improved deep learning models for classifying attributes. An externalization of FANCY into a more detailed and reliable fashion trend analysis benefits fashion professionals.

2 RELATED WORK

2.1 Attribute and style in fashion

Attributes (e.g., dress, sleeve, stripe, red color) are used to classify types of clothing [2, 16, 29], describe clothing [4, 6], and retrieve products [18, 26, 32]. Relative attributes (i.e., the strength of an attribute in an image with respect to other images) [30] in fashion are used for classifying styles [24, 45], which are important to grasp a fashion trend. Different from attributes, style is a higher-level concept that can further represent a fashion trend. Bather [1] defined two concepts, attributes and styles, and established a hierarchical relationship between the two, which provides a view of fashion structurally. However, in prior studies reflecting computational approaches, such as computer vision, *the concepts of style and attribute were not considered at the same time.* In this paper, we defined essential attributes through interviews with fashion professionals and proved meaningful relationships between attributes and styles by developing deep learning models.

2.2 Detecting fashion attributes

Many studies on attribute detection have applied a Convolutional Neural Network (CNN)-based approach. Liu et al. [29] developed a CNN architecture based on VGG-16 to classify fashion categories and attributes. Tangseng et al. [40] tackled segmentation tasks with a fully-convolutional neural network (FCN) approach. Several studies employed Region-based Convolutional Neural Networks (R-CNN) models to detect a body or generate clothing proposals [6, 11, 18]. Unlike prior work, *we adopted state-of-the-art object detection models* such as Faster R-CNN [32] and RetinaNet [26], which are representative of a one-stage model and two-stage model, respectively. We demonstrated that these two models yield outstanding performance in object detection work.

2.3 Classifying fashion styles

Compared to attribute-based studies, relatively fewer studies have focused on fashion *style classification*. Early studies used supervised learning such as SVM [23], ResNet [39] for style classification (e.g., feminine, preppy, rock). For labeling garments for each style, researchers have collaborated with fashion professionals [39] and customers [36]. Veit et al. [41] employed weak supervised learning which uses a limited amount of labeled data to classify fashion styles. Unlike such prior studies, *we pursued an unsupervised approach* (e.g., NMF) for discovering fashion styles with attributes. This has the following advantages: (1) provides large-scale trend analysis, (2) allow for a broader style criteria covering all clothes, and (3) enables style classification for brand-new clothes in a fashion show.

2.4 Fashion attribute and style dataset

Table 1 summarizes fashion attribute and style datasets used in prior studies. To find the attributes of clothing, fashion-related image datasets were constructed in many studies [6, 29]. Attributes were extracted from images in each dataset in a similar way. First, attribute annotation was done through crowdsourcing and finalized by the researcher. The HipsterWars [41] and the FashionStyle14 dataset [39] were organized for the purpose of identifying clothing styles. The determination of style types in the HipsterWars dataset [41] was done by a researcher, and style labeling was done by people. The FashionStyle14 [39] dataset was handled by fashion professionals with respect to style selection and labeling. The Fashion144K [35] dataset summarizes attributes and styles to measure the fashionability of clothing. Unlike previous studies, *we annotated attributes based on professional knowledge, and labeled a style based on the combination of attributes. We then constructed a dataset in which both attributes and styles exist in a hierarchical relationship.*

3 ATTRIBUTE DEFINITION

3.1 Interviews with fashion professionals

We interviewed ten current fashion professionals (five fashion designers, two fashion magazine editors, three fashion business managers) to understand the work process of fashion analysis in a lab seminar room (September 16-30, 2019). Considering previous studies [3, 39, 46] (8-15 professionals), the number of professionals involved in our research is appropriate. The primary purpose of the interviews was to find availability of machine learning technology in professionals' current main work such as style trend analysis. In the interview, we discussed challenges of the style trend work, the limitation of previous fashion-oriented datasets and how to improve methods to address identified challenges. In general, a fashion style is defined as follows. A style is defined by fashion professionals, and the public (or normal users) evaluates or adopts it [10]. Thus, working with professionals for constructing the FANCY framework is appropriate, given the nature of the fashion industry.

Table 1: Overview of fashion-oriented datasets.

Dataset	*Number of Attributes*	*Number of Styles*	*Professional Participation*	*Number of Images*
ACWS [2]	15	No	No	145,718
DCSA [4]	26	No	No	1,856
Fashionpedia [21]	46	No	No	50,000
FashionGen [34]	48	No	No	293,000
DDAN [6]	67	No	No	341,021
DeepFashion [29]	1,000	No	No	800,000
HipsterWars [41]	No	5	No	1,893
FashionStyle14 [39]	No	14	Yes	13,126
Fashion 144K [35]	No	No	No	144,169
Our dataset (FANCY)	**146**	**25**	**Yes**	**302,772**

3.2 Results of interviews

3.2.1 Inconveniences in work processes with subjective criteria. We found that in the fashion industry professionals tend to rely on their field experience when classifying styles. While professionals had agreed-upon standards to some extent, there were no clear quantitative criteria for classifying fashion styles. While professionals show similar responses to fashion images (e.g., style name, style classification criteria), they often respond very differently due to the subjective nature of one's fashion insights.

3.2.2 Limitation of previous fashion datasets. We identified the following limitations of the previous fashion-oriented datasets: (1) the relationship between attributes and styles; in previous work, attributes and styles are considered as horizontal relationships [29, 41]. However, the professionals mentioned that attributes subordinate to styles, meaning the two concepts have a hierarchical relationship. (2) the number of attributes; professionals responded that the number of attributes is too small to represent fashion styles [2, 4, 34]. They mentioned at least 100 attributes are needed for usage in a fashion design process [3]. (3) data validation; attributes need to be accepted by professionals. For example, Liu et al. [29] crawled the brand name of online shops and defined 1,000 attributes. Although the number of attributes is large, there are 402 semantically similar attributes (e.g., "abstract geo" vs. "abstract geo print"), which results in ambiguity in the data [22]. The professionals mentioned that such ambiguities are caused by the absence of expertise.

3.2.3 Defining attributes in the fashion design process. We discussed with fashion professionals to define quantitative criteria for style trend analysis. We focused the core fashion factor, fashion attributes, which could be defined quantitatively. Furthermore, the professionals responded that fashion attributes should be used as quantitative criteria for clustering fashion style, considering fashion design process where designers try to make combination of fashion attributes to create new fashion style. We investigated the

[3]Professionals mentioned that defining the exact number of attributes does not mean much because core attributes vary by fashion domains

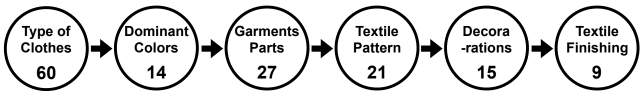

Figure 3: Number of attributes based on design steps. The appendix (Sec.A) shows 146 whole fashion attributes in detail.

process of designing clothes and derived a six-step process. The process consisted mainly of selecting the attributes required for the style from the attributes corresponding to each step.

We asked about the attributes that the fashion professionals use in practice in each of the six steps and then extracted 160 initial attributes. To avoid inaccurate annotations, we excluded 14 material-related attributes (e.g., wool, cotton, silk) that are even difficult for professionals to perfectly classify. Figure 3 illustrates the final 146 attributes extracted from six design steps.

4 ATTRIBUTE VERIFICATION

We defined a total of 146 fashion attributes that reflect fashion professional views, but have not been verified whether these attributes are useful in classifying styles. As the base data, we used the FashionStyle14 [39] dataset because it was labeled by fashion professionals. We additionally manually labeled our defined 146 fashion attributes to the FashionStyle14 dataset and tested whether the classification model with the attribute-added FashionStyle14 dataset performs better than the existing model with ResNet-50 presented in [39].

The verification was carried out by comparing the performance of the attribute-based logistic regression model with that of other CNN-based models, such as VGG-16, VGG-19, Xception, and Inception v3 except for ResNet-50.

4.1 Dataset for attribute verification

The FashionStyle14 dataset consists of 14 fashion style groups: conservative, dressy, ethnic, fairy, feminine, gal, girlish, kireimecasual, lolita, mode, natural, retro, rock, and street. These classes were chosen by fashion professionals who labeled the classes for each fashion image [39]. We finally used all 14 fashion styles and 12,190 images in the dataset, with the exception of 936 images with a 0 byte size.

4.2 Data annotation

4.2.1 Attribute extraction (excluding dominant colors). Since labeling methods corresponding to the dominant color attributes and the remaining attributes are different, we first labeled 132 attributes manually, except dominant color attributes, in the FashionStyle14 dataset. We labeled attributes in each image by marking bounding boxes in the attribute area. Because this required the consideration of the existence of all 132 attributes in a single image, the task was highly labor-intensive and required professional knowledge of fashion for accurate labeling. Thus, we worked with three fashion design students for a month (October 1-31, 2019). Finally, one professional with four years of experience in the fashion field double-checked the annotations to complete the data preparation for the analysis.

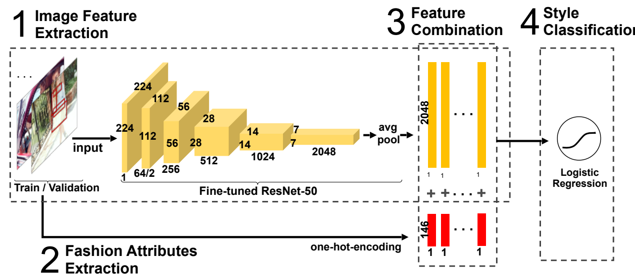

Figure 4: Style classification modeling with combined features.

4.2.2 Dominant color attribute extraction. We identified the importance of dominant color in style classification through the interviews with the professionals. Accordingly, the extraction and labeling of the dominant color was carried out for the images in the same FashionStyle14 dataset. Since the information we need is color information for the corresponding pixel value of the garment, we excluded unnecessary color information such as background and human skin through the following two steps (note that these colors are different from the dominant colors). First, to address the background, we built a semantic segmentation model using Mask R-CNN [17] which is pre-trained with COCO dataset [27]. Second, we used a color extraction algorithm [4] to extract the dominant color of the garment in the image. After completing the dominant color labeling work for all images, three professionals who participated in the interview evaluated the appropriateness of the results.

4.3 Experiments

4.3.1 Attributes verification. We first tested whether 146 attributes are reliable in detecting/identifying styles. We only used the fashion attributes as features to train a logistic regression model using scikit-learn [31] and conducted performance comparisons with CNN-based models tested by [39]. The attributes were applied to datasets using one-hot encoding. To evaluate the performance of the style classification algorithms, we used accuracy as a metric, since there is no class imbalance issue with the FashionStyle14 dataset and it is necessary for direct comparison with prior studies as they used accuracy [39]. We used 60% of the dataset for training, 5% for validation, 35% for testing as done in [39]. In Table 2, the performance of the attribute-based logistic regression model is 64%, slightly lower than that of ResNet-50, but higher than that of VGG-19, Xception, Inception v3, and VGG-16. This result highlights the importance of our proposed fashion attributes for improving the performance of 14 fashion style classification.

4.3.2 Modeling with feature combination. There is a difference in the learning process between ResNet-50 and attribute-based logistic regression. ResNet-50 did not learn professionals' views regarding the presence or absence of fashion attributes. Meanwhile, the attribute-based logistic regression model did not learn textures in an image. Therefore, in this paper, we propose a model that *combines the features extracted by ResNet-50 and the fashion attributes* as illustrated in Figure 4.

First, we used ResNet-50 that provides the highest performance as a feature extractor. After fine-tuning ResNet-50 pre-trained on

[4]https://github.com/algolia/color-extractor

Table 2: The performance comparison for style classification. Each type of models with Our model (Merged) and Our model (Attribute) indicate attribute and image feature-based logistic regression model and attribute-based logistic regression model respectively.

Model	*conserv.*	*dressy*	*ethnic*	*fairy*	*feminine*	*gal*	*girlish*	*kireime-casual*	*lolita*	*mode*	*natural*	*retro*	*rock*	*street*	***mean***
Our model (Merged)	**0.71**	**0.93**	**0.74**	**0.91**	**0.70**	**0.79**	**0.73**	0.64	**0.94**	**0.75**	**0.79**	**0.69**	**0.76**	**0.79**	**0.78**
ResNet-50	0.66	0.91	0.74	0.88	0.64	0.74	0.47	**0.66**	0.92	0.72	0.70	0.62	0.68	0.69	0.72
Our model (Attribute)	0.65	0.89	0.61	0.67	0.61	0.70	0.48	0.60	0.91	0.53	0.60	0.54	0.55	0.60	0.64
VGG19	0.54	0.79	0.57	0.81	0.43	0.50	0.26	0.54	0.80	0.62	0.56	0.42	0.53	0.60	0.58
Xception	0.44	0.79	0.63	0.84	0.45	0.50	0.33	0.54	0.80	0.61	0.56	0.44	0.52	0.53	0.58
Inception v3	0.37	0.73	0.54	0.78	0.41	0.39	0.27	0.45	0.78	0.55	0.44	0.35	0.47	0.46	0.51
VGG16	0.31	0.78	0.49	0.78	0.42	0.45	0.22	0.43	0.81	0.58	0.57	0.23	0.43	0.43	0.51

ImageNet [9] using the same dataset and hyperparameters of previous research [39], we extracted features for all images of the dataset from the average pooling layer. Next, we combined the image features extracted from the fined-tuned ResNet-50 with our 146 fashion attributes to form final features. The shape of the features is $(2048 + 146)\times 1$. The final dataset for our method consisted of 12,190 images and 2,195 features, with a training set (65%), validation set (5%) and test set (30%).

The accuracy of our model, which merged ResNet-50 and attributes, was 6% higher than that of ResNet-50 alone, which was used in the prior work [39] as shown in Table 2. This demonstrated that adding the features of the fashion attributes to those of ResNet-50 improves model performance.

4.3.3 Significance of fashion attributes. We further investigated the significance of the attributes through visualization. There were 1,126 false-positive images from the ResNet-50, and 463 of them were true-positive images from our model. The 463 images were utilized to indirectly identify the effects of the attributes in improving the performance of style classification.

We applied a novel explanation technology LIME that explains the predictions of any classifier [33] to the classification process of ResNet-50 for those 463 images. Figure 5 illustrates some examples. We can see that the superpixel areas (the bright areas inside the yellow border), which ResNet-50 considered to correctly classify the image, and the areas of attributes of the image marked with the red bounding boxes. We categorized the style prediction results of ResNet-50 into two cases as follows.

- Case 1 (Figure 4-a, b): The bright areas inside the yellow border are considered key features and utilized to classify the style correctly by ResNet-50. Most attributes of the style are contained within the considered key feature area. But since the area has textures which are confusing even to professionals to classify correctly, it is still challenging for ResNet-50 which uses a key feature as texture of image.
- Case 2 (Figure 4-c, d): A few attributes of the style are contained within the considered key feature area. The attributes outside the key feature area are not well-considered by ResNet-50, even though they are components that make up the style. The attributes included in the areas considered as a key feature are utilized to correctly classify styles, but may also be included in other styles. Due to these characteristics of style images, the ResNet-50 has difficulty classifying correctly.

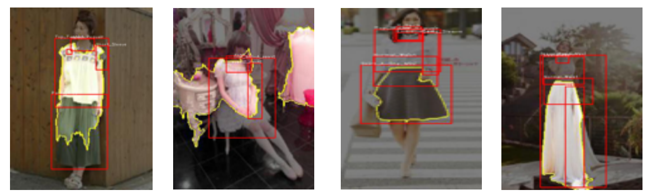

Figure 5: Comparisons between regions explaining predictions of a fine-tuned ResNet-50 presented as the bright areas inside the yellow border and attribute regions presented as red bounding boxes in an image. R, M, and G indicate ResNet-50, Merged (our model), and Ground-truth, respectively. These images show the cases that are hard to be identified by ResNet-50 only, but can be addressed by using attributes as additional features in modeling.

CNN-based ResNet-50 has these weaknesses because it relies on image textures, as in the two cases above. The result has shown possibility that attributes could create synergy with features of a CNN-based model that performs very well in style classification. These results show the adequacy of using the 146 attributes in style classification by presenting the missing parts by CNN (the yellow area) but covered by attributes (the red boxes). In this way, we demonstrated the appropriation of the use of 146 attributes, which also reflects fashion professionals' insights and suggestions.

5 ATTRIBUTE-LABELED DATASET

By quantitatively demonstrating the significant relationship between attributes and styles, we justify attributes as a feature for the new style grouping. Our next step is to label attributes to a large-scale runway data that we collected for fashion trend analysis. This work will be done most accurately by professionals, however is realistically challenging given that the number of runway images is about 300,000. Hence, we employed object detection algorithms and built an attribute detection model (ADM) that yields high performance for labeling over a large-scale dataset.

5.1 ADM Dataset

The dataset (25,470 images) used in ADM (attribute detection model) includes the images used in Sec.4.1 (12,190 images) and the ones crawled from Google Images (13,280 images). We used Google Images for additional data collection, because the number of images per an attribute was not evenly distributed. We crawled for images on Google Image using each of the 100 attributes which are relatively deficient (e.g., "V neck", "Top T-shirt Boxy"). After collection, we manually selected the images that placed clothes in the middle. As a result, we added 13,280 images, amassing a total of 25,470 images for the analysis (ADM dataset[5]).

The images used in Sec.4.1 (12,190 images) already had attribute labels, however Google Image crawling dataset needs attribute labeling. Thus we collaborated again with the same three fashion design students who had annotated attributes on the verification dataset (Sec.4.1). Using the BBox-Label-Tool, 132 attributes were multi-labeled for each image. This work was done over one month, and again one fashion professional double-checked the annotations.

5.2 Modeling

Labeling fashion attributes for 300K runway data requires a proper dataset as well as a sophisticated object detection algorithm with high performance. The fashion attribute dataset has two unique characteristics. The first is that many attributes are of a small size (e.g., beads, pockets). The second is that the positions of the attributes often overlap because the they are located in a limited space (e.g., on clothing).

Huang et al.[19] confirmed that when small objects are in confined spaces, there are differences in performance for various object detection algorithms. Typical object detection algorithm models such as SSD [28], R-FCN [8] and Faster R-CNN [32] were compared, among which Faster R-CNN performed best. The Facebook AI research team recently announced RetinaNet [26], which exceeded the performance of Faster R-CNN. In addition to small objects, overall performance was better [26]. Considering the corresponding performance of the algorithms, we decided to use Faster R-CNN and RetinaNet to find algorithms that fit AMD dataset.

5.3 Results

The indicator used to evaluate the performance of object detection models is mAP (mean Average Precision) [8, 19, 26, 28, 32]. However, we used F1-score, which is used as an evaluation indicator for multi-label classifications, because we require the presence or absence of a specific attribute in the fashion image, not location information.

ResNet50 and 101 were used as a backbone [26, 32] and compared in terms of performance for four cases. The RetinaNet model, using ResNet101 as a backbone, showed the best performance of 45.0% F1-score (Table 3). By comparing the performance of prior attribute classification models with that of our model, we verified the appropriateness of the attributes identified in our study that better reflect fashion professionals' needs and their tasks. We used these attributes to define styles through clustering, which will be explained in the next section. We summarized the problems of other studies as follow (Table 4).

- **[P1] Small number of classes**: Some studies defined the small number of fashion attributes as their classes [2, 4, 6]. However, it is difficult to represent fashion styles with the low number of fashion attributes.

[5]The ADM dataset is available at: https://github.com/youngseungjeon/FANCY

Table 3: Performance of four different attribute detection models using our attribute dataset. The RetinaNet with ResNet-101 model yielded the greatest performance, and we used this model for labeling attributes for 300K fashion image dataset.

Model	*Backbone*	*Precision*	*Recall*	*F-1 score*
Faster RCNN	ResNet-50	0.228	0.309	0.242
Faster RCNN	ResNet-101	0.324	0.461	0.365
RetinaNet	ResNet-50	0.459	0.426	0.426
RetinaNet	**ResNet-101**	**0.470**	**0.468**	**0.450**

Table 4: Model performance overview in prior research. Problem 1 (P1), P2, and P3 indicate small number of classes, no performance result, and evaluation metric issue, respectively.

Paper	*The number of attributes*	*Performance (average)*	*Metric to evaluate*	*Type of problem*
ACWS[2]	15	41.36%	Accuracy	P1, P3
DSCA[4]	26	N/A	Accuracy	P1, P2, P3
DDAN[6]	67	N/A	Accuracy	P1, P2, P3
DARN[18]	179	42.35% (top 3)	Accuracy	P2, P3
AttentiveNet[42]	1,000	51.53% (top 3)	Accuracy	P2, P3
Corbiere[7]	1,000	23.10% (top 3)	Accuracy	P2, P3
Deepfashion[29]	1,000	45.52% (top 3)	Accuracy	P2, P3
Our model	**132**	**45.00% (top 1)**	**F1 score**	

- **[P2] Evaluation criteria issue**: Model performance of attribute classification or not measured at all [4, 6]. Some research considered top-k attributes as evaluation criteria. However, wrong attributes can be included in k, which could also undermine the performance of clustering and its applicability to fashion analysis tasks.
- **[P3] Evaluation metrics issue**: F1-scores for both recall and precision were not used [18, 29]. Other studies used accuracy which may be deleterious when the number of samples varies by classes [2, 4, 6].

5.4 Large-scale attribute dataset

We crawled fashion images to build a large-scale attribute dataset. A total of 302,772 images from 8,121 fashion shows were collected from Vogue US [6] between 2010 and 2019. The fashion images covers 987 brand names ranging from mega couture (e.g., Gucci, Chanel) to high street brands (e.g., JCrew, Topshop) [16]. We labeled attributes in each image by using machine learning models. The dominant color model (Sec.4.2.2) was used to label dominant color attributes (14) and ADM (Sec.5.2) was used to label all other attributes (132). We had 302,772 images with 146 attributes.

6 STYLE CLUSTERING

6.1 Non-negative factorization (NMF)

We pose our style discovery problem in a non-negative factorization (NMF) [44]. It has benefits because each axis in the space derived by the NMF has a straightforward correspondence with each document cluster, and document clustering results can be directly derived without additional clustering operations. Regarding NMF, we

[6]https://bit.ly/2lV6NzQ

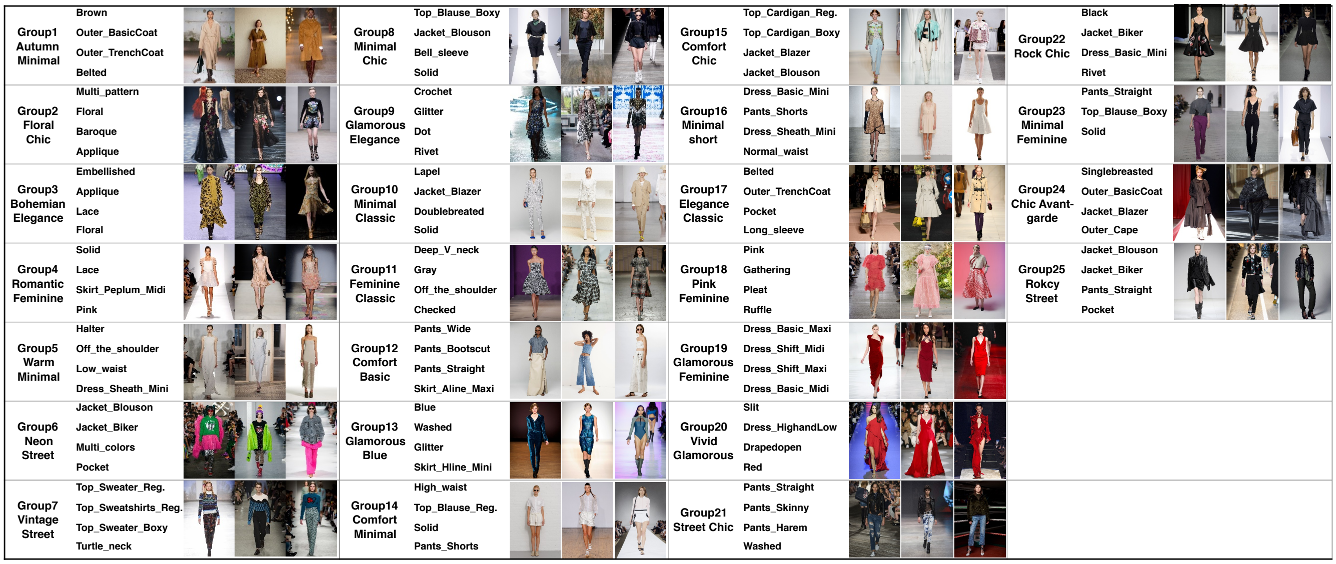

Figure 6: Representative images and attributes of the 25 style groups.

are aware that many deep learning algorithms for clustering have been introduced. However, such algorithms have disadvantages because they do not indicate how much an attribute impacts on the development of its assigned style group. Identifying attributes within a particular style is important for style configuration and analysis; Hence, we employed NMF for FANCY.

In NMF, the entire data V is divided into matrix parameters and expressed in times of matrix W (Weight Matrix), and matrix H (Feature Matrix) [25]. Data V, which consists of attribute i which is the number of attributes (146) and u which is the number of images (302,722) (Equation 1).

$$V \approx WH \;\; where \;\; V \in \mathbb{R}^{i \times u}, W \in \mathbb{R}^{i \times a}, H \in \mathbb{R}^{a \times u} \tag{1}$$

Data V is decomposed r (the number of styles) times via matrix decomposition into W consisting of attributes i and style a, and feature matrix consisting of a and image u (Equation 2).

$$\boldsymbol{V_{iu}} \approx (\boldsymbol{W H_{iu}}) = \sum_{\boldsymbol{a}=1}^{\boldsymbol{r}} \boldsymbol{W_{ia} H_{au}} \tag{2}$$

In the feature, for H, the style group is labeled according to the attributes of a particular image, and the specific attributes and their importance that make up that style can be seen in W.

In summary, the style is clustered according to the attributes of each image, and it is possible to see which attributes contributed to the creation of a particular style cluster and the extent to which particular attributes are important in clustering for a particular style through NMF. We are able to intuitively explain the fashion style we grouped and efficiently classify large volumes of data.

6.2 Clustering fashion styles

By conducting a third round of interviews with fashion professionals, we identified the appropriate number of the groups (clusters) and verified that each style group well represents a fashion style. Eight out of ten answered between 25 and 30 style groups would be the most appropriate, which is significantly different from the 14 style groups defined in FashionStyle14 [39]. The professionals named each style group based on what they commonly use in the fashion domain, which helps make the clustering results more reasonable and understandable.

Fashion professionals said that mixed style names (e.g., "Comfort Feminine," "Rocky Street") are used more than single style names (e.g. "Feminine," "Rock") that were defined and used in the FashionStyle14 dataset [39]. This is because designers participating in a runway presents their new styles by combining various attributes from different styles [15]. As a result, clothes on the runway are often very vague when attempting to define them with a single style name. We finally extracted 25 style groups using NMF results (Figure 6). Thus, we we believe our 25 styles reflect professionals' need and the variability of fashion style. Finally, we built the large-scale fashion style dataset (302,772 images with 146 attributes and 25 styles). We named this datset as FANCY dataset.

6.3 Style prediction model

We verify the hierarchical relationship between the attributes and styles in FANCY dataset by developing style prediction model based on attributes. We built a logistic regression model (80% training and 20% test FANCY datasets) with five-fold cross-validation. It yielded very high performance (91.2% F1 score), which means that the relationship is well established in our dataset. Furthermore, this demonstrates the possibility of using a classification model for new images in the future fashion collections.

7 STYLE TREND ANALYSIS

Fashion trends consider the appearance and construction of particular fashion styles that relate to a particular season and fashion trends change very rapidly [20]. For example, a popular style in a certain period of time may quickly fall in value after just one season [20]. Due to the rapid variation of fashion styles which is

Figure 7: Representative fashion images of each trend group.

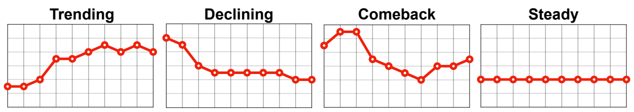

Figure 8: Life-cycle of fashion styles.

qualitative factor, trend analysis could be an appropriate case in our study.

The life cycle of a fashion product has four levels [38]: introduction, growth, maturity, and decline. According to the professionals, knowing the life-cycle of a certain style is key to generating profits. For these reasons, we conducted year-, region-, and brand-based trend analyses.

7.1 Year-based trend grouping

A fashion style trend can be defined as a change in popularity of a particular style over time [20]. We identified the frequency of the style shown in fashion runways over 10 years (2010-2019) and analyzed style trends.

7.1.1 The popularity of a style. The fashion trend index determines how many a style is shown in a given year. As the number of images shown on runways varies across years, we used the relative frequency of each style in a given year ($\boldsymbol{y}_t^s$). The number of images I of a particular style s in a given year t divided by the total number of images Q was used as the fashion style trend indicator (Equation 3).

$$\boldsymbol{y}_t^s = \frac{\boldsymbol{I}_t^s}{\boldsymbol{Q}_t} \tag{3}$$

7.1.2 Groups in fashion trends. We analyzed a temporal trajectory of 25 styles. The results were largely organized into four trend groups. In Figure 8, we visualized a style representing four groups: (1) Trending, (2) Declining, (3) Comeback, and (4) Steady.

7.1.3 Representative images. We arranged the 25 styles represented in each trend group and the corresponding representative images (Figure 7). In the Trending group, the strong features were in Autumnal Minimal, which unraveled a neutral color with minimal design, and the Bohemian Elegance style, which seeks a sense of etching in elegance. In comparison, the styles Neon Street and Minimal Short, centered on dark and vivid colors, had passed their peak and were becoming increasingly unpopular (Declining). Warm Minimal and Pink Feminine styles stopped their decline and began a comeback in the trending (Comeback). Chic Avant-garde and Rocky Street style were constantly on the runway, and less influenced by fads because they are reflective of niche groups (Steady).

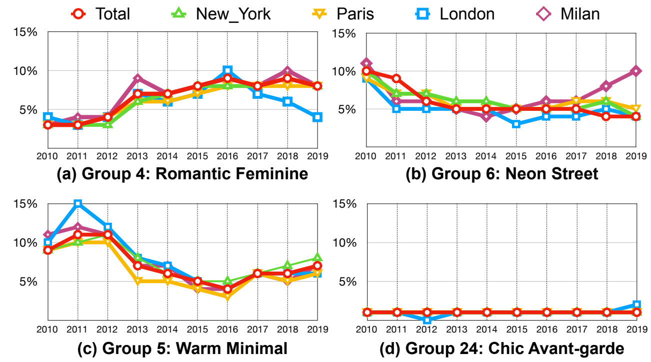

Figure 9: Trend flow changes by city (x-axis: year, y-axis: group ratio).

7.2 Trend analysis by region

New York, Paris, Milan and London are the main runways in the world. Vogue US has compiled runway region information for each brand. The number of images of the New York Fashion show accounts for 44.4% of the total (30% for the Paris collection, about 30% of the rest for Milan and London). With the dataset, we looked at the 25 fashion styles in each of the four cities based on the same method used in the previous section (Equation 3) where the Q is relative to the city in the city-based analysis.

- New York and Paris: Figure 9 illustrates a similar graph with no significant difference between New York, Paris and the entire cities. This means that New York and Paris not only share a similar style group distribution, but they best reflect/influence overall fashion trends around the world.
- London and Milan: Their main projection of styles reflected several differences, although most were similar to the entire set of data. In London, for example, the trend of Romantic Feminine (Group 4) in (Figure 9a) the Trending group was different from other cities. Three other cities showed rising styles, but London showed a downward curve that peaked and came down. For Milan, the trend of Neon Street (Group 6) was different from other cities. Three other cities presented falling styles, but Milan showed a steep upward curve (Figure 9b).

7.3 Trend analysis by brand

We analyzed the distribution of style groups for each brand across a region. Among brands, Chanel and Gucci were chosen for an analysis of style distribution (Figure 10), because of their popularity. Gucci has a high frequency in Groups 2 and 3, and was very high in Group 9. Group 9 is Glamorous Elegance, a style that shows splendid elegance with brilliant decorations such as beading. Unlike Gucci, which has colorful styles, in Chanel, a very high ratio was found in Group 4 (Romantic Feminine styles).

7.4 Evaluation from professionals

We have conducted a posteriori sample-based evaluation of style clustering and trend analysis results through additional interviews with five fashion professionals (three fashion designers, one fashion editor, and one fashion business manager). We randomly selected

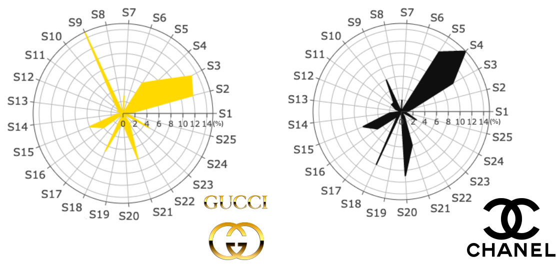

Figure 10: Example of the style group distribution by brand: Gucci, Chanel.

300 fashion looks and prepared the style clustering results (Figure 6), the trend analysis results for each look (Figure 9), and the brand analysis results (Figure 10). Each interview took approximately one hour and was held in the authors' laboratory. From the results, we identified three main points:

- New definition boundary of style: The professionals responded that they could broaden their limited style boundaries through the representative looks of style groups defined in quantitative ways. One fashion designer noted, *"In the representative photo of the style, there was a design that I hadn't usually thought of, but it was a recommendation in an understandable range, which expanded the range of the style I was thinking of"*. The other fashion designer noted *"The style I chose is a sexy style. Looking at the attribute information, such as the representative blazer and boxy sweater, which are far from the sexy style standard I thought of. It seemed necessary to distinguish the sexy style boundary that I already knew in greater detail. I was able to come up with a style group that I hadn't thought of"*.
- New insights in trend analysis: The professionals mentioned that a quantitative analysis of the frequency variation of a particular style provided them with new insights. One fashion designer noted, *"It is difficult to know 'when' the style trend increased and 'how many' of those styles have come up. Arranging the frequency of a particular style in chronological order will be very useful when fashion professionals decide to a long-term design strategy"*. In addition, one fashion editor noted, *"I can identify different trends in cities, which opened up the possibility for fashion professionals to understand the local trend"*.
- Reconfirmation of professionals' knowledge: The professionals responded that they were able to reconfirm information gained previously. The results of the brand analysis did well in presenting the brands' design characteristics. One fashion business manager noted, *"Quantitative viewing of the brand's style can provide advantages to our work. We can digitally identify the styles of their brands, and competitive brands, helping us strategically plan designs"*.

8 DISCUSSION AND CONCLUSION

In this work, we (1) identify limitations of prior work especially in data preparation, (2) redefine 146 attributes strongly associated with styles, (3) build the deep-learning based attribute detection model reflecting professionals' insight, (4) the large-scale fashion dataset presenting 25 styles and (5) showcase the applications of the model to fashion trend analysis that can be used by professionals.

Through meetings with fashion professionals, we gained an understanding of fashion analysis as a whole, the areas that needed improvement and how to solve problems that many fashion professionals are currently experiencing. The professionals shared the need for a quantitative style indicator and provided us with important insights on the design process for clothing as well as the attributes used at each design stage. Unlike prior studies that organized attributions from the mass data of the online market, *a deeper understanding of the stage in which the style was created and how professionals approached such trends* allowed us to prepare the dataset that has 146 important attributes and is used to generate 25 fashion styles. The datasets, the method of labeling and modeling, style generation, and trend analysis used in our study better reflect the fashion-related work practices and norms than those used in many prior studies. Furthermore, to the best of our knowledge, this is the first research study that combines and leverages user-centered and deep-learning methods in fashion research. Our study also gives researchers, practitioners, and designers insights on the application of machine/deep learning techniques to domains that have high variability and many subjective decision-makings are conducted (e.g., fashion, interior design, food).

We have sorted out limitations that we plan to address in future work. First, the attributes of the materials (e.g., cotton, silk) were not considered in this study. Although the detection of the material is still a significant challenge, as mentioned earlier, it is expected that we can define more style groups by developing models with more diverse attributes. Second, although our detection models yielded good performance, they can be further improved and more generalized. Having more data and diverse attributes will help improve performance for both style and attribute detection. We are now in the process of obtaining more data through a collaboration with a fashion company. Third, although our research has been framed based on close work with ten fashion professionals, their insights may not represent the perspectives of all professionals nor all the dynamics of the fashion industry. Thus, the attributes or styles derived from our study may not be applicable to some case. Last, the possibility that same looks can belong to multiple clusters. Based on the definition of clustering, one look can only belong to one cluster. The 25 style clusters identified in our paper were based on close collaboration with the 10 fashion professionals. However, there is flexibility regarding the number of clusters. If the number of style clusters is changed, it is possible that a particular look could belong to a totally different style cluster. Because 25 is not an absolute number applying to all fashion domains, there are possible changes in the relationship between style looks and clusters based on the number of clusters that can be defined by different fashion professionals.

Our proposed FANCY framework can go beyond fashion style modeling and trend analysis to develop a recommendation or decision-making support system. Such systems would allow professionals to analyze fashion images by attribute or style, recommend the most relevant images by various metrics (e.g., attribute, season, brand), and compare fashion images in different styles. Understanding effective ways of using inherent attributes of the image (attributes) and the derived latent space of those images (styles) is important because that will support professionals' decision-making and enhance user experience in content mining of multimedia Web data. This will also benefit many researchers in the perspective of proposing novel computational techniques that leverage characteristics of Web data.

ACKNOWLEDGMENTS

This research was supported by the MSIT (Ministry of Science and ICT), Korea, under the program (2020-0-01523) supervised by the IITP (Institute for Information & Communications Technology Planning & Evaluation) and the program (2020R1F1A1076129) by the NRF (National Research Foundation).

REFERENCES

[1] Roland Barthes et al. 1983. *The fashion system.* University of California Press Berkeley.

[2] Lukas Bossard, Matthias Dantone, Christian Leistner, Christian Wengert, Till Quack, and Luc Van Gool. 2012. Apparel classification with style. In *Asian conference on computer vision.* Springer, 321–335.

[3] Minsuk Chang, Léonore V Guillain, Hyeungshik Jung, Vivian M Hare, Juho Kim, and Maneesh Agrawala. 2018. Recipescape: An interactive tool for analyzing cooking instructions at scale. In *Proceedings of the Conference on Human Factors in Computing Systems.* 1–12.

[4] Huizhong Chen, Andrew Gallagher, and Bernd Girod. 2012. Describing clothing by semantic attributes. In *European conference on computer vision.* Springer, 609–623.

[5] KuanTing Chen, Kezhen Chen, Peizhong Cong, Winston H Hsu, and Jiebo Luo. 2015. Who are the devils wearing prada in new york city?. In *Proceedings of the 23rd ACM international conference on Multimedia.* ACM, 177–180.

[6] Qiang Chen, Junshi Huang, Rogerio Feris, Lisa M Brown, Jian Dong, and Shuicheng Yan. 2015. Deep domain adaptation for describing people based on fine-grained clothing attributes. In *Proceedings of the IEEE conference on computer vision and pattern recognition.* 5315–5324.

[7] Charles Corbiere, Hedi Ben-Younes, Alexandre Ramé, and Charles Ollion. 2017. Leveraging weakly annotated data for fashion image retrieval and label prediction. In *Proceedings of the International Conference on Computer Vision Workshops.* 2268–2274.

[8] Jifeng Dai, Yi Li, Kaiming He, and Jian Sun. 2016. R-fcn: Object detection via region-based fully convolutional networks. In *Advances in neural information processing systems.* 379–387.

[9] Jia Deng, Hao Su, Jonathan Krause, Sanjeev Satheesh, Sean Ma, Zhiheng Huang, Andrej Karpathy, Aditya Khosla, Michael Bernstein, Alexander C Berg, et al. 2014. Imagenet large scale visual recognition challenge. *arXiv preprint arXiv:1409.0575* (2014).

[10] Fiona Dieffenbacher. 2013. *Fashion thinking: Creative approaches to the design process.* Vol. 14. A&C Black.

[11] Qi Dong, Shaogang Gong, and Xiatian Zhu. 2017. Multi-task curriculum transfer deep learning of clothing attributes. In *IEEE Winter Conference on Applications of Computer Vision (WACV).* IEEE, 520–529.

[12] Xue Dong, Xuemeng Song, Fuli Feng, Peiguang Jing, Xin-Shun Xu, and Liqiang Nie. 2019. Personalized Capsule Wardrobe Creation with Garment and User Modeling. In *Proceedings of the 27th ACM International Conference on Multimedia.* 302–310.

[13] Ann Marie Fiore and Hyun-Jeong Jin. 2003. Influence of image interactivity on approach responses towards an online retailer. *Internet Research* 13, 1 (2003), 38–48.

[14] Ann Marie Fiore, Hyun-Jeong Jin, and Jihyun Kim. 2005. For fun and profit: Hedonic value from image interactivity and responses toward an online store. *Psychology & Marketing* 22, 8 (2005), 669–694.

[15] Patrizia Gazzola, Enrica Pavione, Roberta Pezzetti, and Daniele Grechi. 2020. Trends in the Fashion Industry. The Perception of Sustainability and Circular Economy: A Gender/Generation Quantitative Approach. *Sustainability* 12, 7 (2020), 2809.

[16] Yu-I Ha, Sejeong Kwon, Meeyoung Cha, and Jungseock Joo. 2017. Fashion conversation data on instagram. In *Eleventh International AAAI Conference on Web and Social Media.*

[17] Kaiming He, Georgia Gkioxari, Piotr Dollár, and Ross Girshick. 2017. Mask r-cnn. In *Proceedings of the IEEE international conference on computer vision.* 2961–2969.

[18] Junshi Huang, Rogerio S Feris, Qiang Chen, and Shuicheng Yan. 2015. Cross-domain image retrieval with a dual attribute-aware ranking network. In *Proceedings of the IEEE international conference on computer vision.* 1062–1070.

[19] Jonathan Huang, Vivek Rathod, Chen Sun, Menglong Zhu, Anoop Korattikara, Alireza Fathi, Ian Fischer, Zbigniew Wojna, Yang Song, Sergio Guadarrama, et al. 2017. Speed/accuracy trade-offs for modern convolutional object detectors. In *Proceedings of the IEEE conference on computer vision and pattern recognition.* 7310–7311.

[20] Tim Jackson. 2007. The process of trend development leading to a fashion season. In *Fashion Marketing.* Routledge, 192–211.

[21] Menglin Jia, Mengyun Shi, Mikhail Sirotenko, Yin Cui, Bharath Hariharan, Claire Cardie, and Serge Belongie. 2019. The fashionpedia ontology and fashion segmentation dataset. *Cornell University* (2019).

[22] Menglin Jia, Yichen Zhou, Mengyun Shi, and Bharath Hariharan. 2018. A Deep-Learning-Based Fashion Attributes Detection Model. *arXiv preprint arXiv:1810.10148* (2018).

[23] M Hadi Kiapour, Kota Yamaguchi, Alexander C Berg, and Tamara L Berg. 2014. Hipster wars: Discovering elements of fashion styles. In *European conference on computer vision.* Springer, 472–488.

[24] Adriana Kovashka, Devi Parikh, and Kristen Grauman. 2012. Whittlesearch: Image search with relative attribute feedback. In *2012 IEEE Conference on Computer Vision and Pattern Recognition.* IEEE, 2973–2980.

[25] Daniel D Lee and H Sebastian Seung. 1999. Learning the parts of objects by non-negative matrix factorization. *Nature* 401, 6755 (1999), 788.

[26] Tsung-Yi Lin, Priya Goyal, Ross Girshick, Kaiming He, and Piotr Dollár. 2017. Focal loss for dense object detection. In *Proceedings of the IEEE international conference on computer vision.* 2980–2988.

[27] Tsung-Yi Lin, Michael Maire, Serge Belongie, James Hays, Pietro Perona, Deva Ramanan, Piotr Dollár, and C Lawrence Zitnick. 2014. Microsoft coco: Common objects in context. In *European conference on computer vision.* Springer, 740–755.

[28] Wei Liu, Dragomir Anguelov, Dumitru Erhan, Christian Szegedy, Scott Reed, Cheng-Yang Fu, and Alexander C Berg. 2016. Ssd: Single shot multibox detector. In *European conference on computer vision.* Springer, 21–37.

[29] Ziwei Liu, Ping Luo, Shi Qiu, Xiaogang Wang, and Xiaoou Tang. 2016. Deepfashion: Powering robust clothes recognition and retrieval with rich annotations. In *Proceedings of the IEEE conference on computer vision and pattern recognition.* 1096–1104.

[30] Devi Parikh and Kristen Grauman. 2011. Relative attributes. In *2011 International Conference on Computer Vision.* IEEE, 503–510.

[31] Fabian Pedregosa, Gaël Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel, Peter Prettenhofer, Ron Weiss, Vincent Dubourg, et al. 2011. Scikit-learn: Machine learning in Python. *Journal of machine learning research* 12, Oct (2011), 2825–2830.

[32] Shaoqing Ren, Kaiming He, Ross Girshick, and Jian Sun. 2015. Faster r-cnn: Towards real-time object detection with region proposal networks. In *Advances in neural information processing systems.* 91–99.

[33] Marco Tulio Ribeiro, Sameer Singh, and Carlos Guestrin. 2016. Why should i trust you?: Explaining the predictions of any classifier. In *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining.* ACM, 1135–1144.

[34] Negar Rostamzadeh, Seyedarian Hosseini, Thomas Boquet, Wojciech Stokowiec, Ying Zhang, Christian Jauvin, and Chris Pal. 2018. Fashion-gen: The generative fashion dataset and challenge. *arXiv preprint arXiv:1806.08317* (2018).

[35] Edgar Simo-Serra, Sanja Fidler, Francesc Moreno-Noguer, and Raquel Urtasun. 2015. Neuroaesthetics in fashion: Modeling the perception of fashionability. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition.* 869–877.

[36] Edgar Simo-Serra and Hiroshi Ishikawa. 2016. Fashion style in 128 floats: Joint ranking and classification using weak data for feature extraction. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition.* 298–307.

[37] Xuemeng Song, Xianjing Han, Yunkai Li, Jingyuan Chen, Xin-Shun Xu, and Liqiang Nie. 2019. GP-BPR: Personalized Compatibility Modeling for Clothing Matching. In *Proceedings of the 27th ACM International Conference on Multimedia.* 320–328.

[38] George B Sproles. 1981. Analyzing fashion life cycles-Principles and perspectives. *Journal of Marketing* 45, 4 (1981), 116–124.

[39] Moeko Takagi, Edgar Simo-Serra, Satoshi Iizuka, and Hiroshi Ishikawa. 2017. What makes a style: Experimental analysis of fashion prediction. In *Proceedings of the IEEE International Conference on Computer Vision.* 2247–2253.

[40] Pongsate Tangseng, Zhipeng Wu, and Kota Yamaguchi. 2017. Looking at outfit to parse clothing. *arXiv preprint arXiv:1703.01386* (2017).

[41] Andreas Veit, Balazs Kovacs, Sean Bell, Julian McAuley, Kavita Bala, and Serge Belongie. 2015. Learning visual clothing style with heterogeneous dyadic co-occurrences. In *Proceedings of the IEEE International Conference on Computer Vision.* 4642–4650.

[42] Wenguan Wang, Yuanlu Xu, Jianbing Shen, and Song-Chun Zhu. 2018. Attentive fashion grammar network for fashion landmark detection and clothing category classification. In *Proceedings of the Conference on Computer Vision and Pattern Recognition.* 4271–4280.

[43] Xin Wang, Bo Wu, and Yueqi Zhong. 2019. Outfit Compatibility Prediction and Diagnosis with Multi-Layered Comparison Network. In *Proceedings of the 27th*

ACM International Conference on Multimedia. 329–337.
[44] Wei Xu, Xin Liu, and Yihong Gong. 2003. Document clustering based on non-negative matrix factorization. In *Proceedings of the international conference on Research and development in informaion retrieval*. ACM, 267–273.
[45] Aron Yu and Kristen Grauman. 2015. Just noticeable differences in visual attributes. In *Proceedings of the IEEE International Conference on Computer Vision*. 2416–2424.
[46] Xiong Zhang, Jonathan Engel, Sara Evensen, Yuliang Li, Çağatay Demiralp, and Wang-Chiew Tan. 2020. Teddy: A System for Interactive Review Analysis. In *Proceedings of the Conference on Human Factors in Computing Systems*. 1–13.

A WHOLE 146 ATTRIBUTES

Table 5: Overview of attributes that make up fashion styles.

Design Step	*Attribute*					
Type of clothes	Top_Blause_Boxy	Top_Blause_Reg.	Top_Cardigan_Boxy	Top_Cardigan_Reg.	Top_Crop_Boxy	Top_Crop_Reg.
	Top_Sweater_Boxy	Top_Sweater_Reg.	Top_Sweatshirts_Boxy	Top_Sweatshirts_Reg.	Top_T-shirt_Boxy	Top_T-shirt_Reg.
	Jacket_Biker	Jacket_Blazer	Jacket_Blouson	Pants_Bootscut	Pants_Harem	Pants_Shorts
	Pants_Skinny	Pants_Straight	Pants_Wide	Skirt_A-line_Maxi	Skirt_A-line_Midi	Skirt_A-line_Mini
	Skirt_H-line_Maxi	Skirt_H-line_Midi	Skirt_H-line_Mini	Skirt_HighandLow	Skirt_Mermaid_Maxi	Skirt_Mermaid_Midi
	Skirt_Mermaid_Mini	Skirt_Peplum_Maxi	Skirt_Peplum_Midi	Skirt_Peplum_Mini	Dress_A-line_Maxi	Dress_A-line_Midi
	Dress_A-line_Mini	Dress_Basic_Maxi	Dress_Basic_Midi	Dress_Basic_Mini	Dress_HighandLow	Dress_Mermaid_Maxi
	Dress_Mermaid_Midi	Dress_Mermaid_Mini	Dress_Peplum_Midi	Dress_Peplum_Mini	Dress_Peplum_Maxi	Dress_Sheath_Maxi
	Dress_Sheath_Midi	Dress_Sheath_Mini	Dress_Shift_Maxi	Dress_Shift_Midi	Dress_Shift_Mini	Outer_BasicCoat
	Outer_Cape	Outer_Parka	Outer_PeaCoat	Outer_Puffer	Outer_TrenchCoat	Jumpsuit
Dominant colors	Black	Blue	Brown	Gray	Green	Maroon
	Mustard	Orange	Pink	Purple	Red	White
	Yellow	Multi_colors				
Garment parts	U_neck	V_neck	Boat_neck	Collar	Cowl_neck	Deep_U_neck
	Deep_V_neck	Halter	Illusion_neckline	Keyhole	Off_the_shoulder	One_shoulder
	Squared_neck	Turtle_neck	Bell_sleeve	Cap_sleeve	Long_sleeve	Puff_sleeve
	Short_sleeve	Sleeveless	Double-breated	Draped-open	Single-breasted	Zip-up
	Normal_waist	High_waist	Low_waist			
Textile pattern	Argyle	Baroque	Brushstroke	Camouflage	Checked	Chevron
	Dot	Floral	Geometric	Grid	Hearts	Houndstooth
	Leopard	Lettering	Marble	Paisley	Stripe	Tie_dye
	Zebra	Multi-Pattern	Solid			
Decoration	Applique	Belted	Bow	Crochet	Embellished	Fringe
	Fur	Glitter	Hood	Lace	Lapel	Pocket
	Rivet	Ruffle	Tassel			
Textile finishing	Distressed	Frayed	Gathering	Pleat	Quilted	Slit
	Smocking	Tiered	Washed			