Modeling Fashion Influence from Photos

Ziad Al-Halah[1]

ziadlhlh@gmail.com

[1]The University of Texas at Austin

Kristen Grauman[1,2]

grauman@cs.utexas.edu

[2]Facebook AI Research

Abstract

The evolution of clothing styles and their migration across the world is intriguing, yet difficult to describe quantitatively. We propose to discover and quantify fashion influences from catalog and social media photos. We explore fashion influence along two channels: geolocation and fashion brands. We introduce an approach that detects which of these entities influence which other entities in terms of propagating their styles. We then leverage the discovered influence patterns to inform a novel forecasting model that predicts the future popularity of any given style within any given city or brand. To demonstrate our idea, we leverage public large-scale datasets of 7.7M Instagram photos from 44 major world cities (where styles are worn with variable frequency) as well as 41K Amazon product photos (where styles are purchased with variable frequency). Our model learns directly from the image data how styles move between locations and how certain brands affect each other's designs in a predictable way. The discovered influence relationships reveal how both cities and brands exert and receive fashion influence for an array of visual styles inferred from the images. Furthermore, the proposed forecasting model achieves state-of-the-art results for challenging style forecasting tasks. Our results indicate the advantage of grounding visual style evolution both spatially and temporally, and for the first time, they quantify the propagation of inter-brand and inter-city influences. Project page: https://www.cs.utexas.edu/~ziad/influence_from_photos.html

1. Introduction

'The influence of Paris, for instance, is now minimal. Yet a lot is written about Paris fashion.' —Geoffrey Beene

The clothes people wear are a function of personal factors like comfort, taste, and occasion—but also wider and more subtle influences from the world around them, like changing social norms, art, the political climate, celebrities and style icons, the weather, and the mood of the city in which they live. Fashion itself is an evolving phenomenon because of these changing influences. What gets worn continues to change, in ways fast, slow, and sometimes cyclical.

Pinpointing the influences in fashion, however, is nontrivial. To what extent did the runway styles in Paris last year affect what U.S. consumers wore this year? How much did the designs by J. Crew influence those created six months later by Everlane, and vice versa? How long does it take for certain trends favored in New York City to migrate to Austin, if they do at all? And how did the infamous cerulean sweater worn by the protagonist in the movie *The Devil Wears Prada* make its way into her closet?[1]

To quantitatively answer such questions would be valuable to both social science and the fashion industry, yet it remains challenging. Clothing sales records or social media "likes" offer some signal about how tastes are shifting, but they are indirect and do not reveal the sources of influence.

We contend that images are exactly the right data to answer such questions. Unlike vendors' non-visual meta-data or hype from haute couture designers, photos of what people are wearing or buying in their daily lives provide a unfiltered glimpse of current clothing styles "on the ground". Our idea is to discover fashion influence patterns in community photos (*e.g.* Instagram) and catalog photos (*e.g.* Amazon), and leverage those patterns to forecast future style trends. We explore fashion influence along two channels–cities and fashion brands—and forecast the popularity of a given style conditioned on the place in the world it will be worn or the brand from which it will be purchased. See Fig. 1. Specifically, we aim to discover *which cities influence which other cities* and *which brands influence which other brands* in terms of propagating their clothing styles, and *with what time delay*.

To this end, we introduce a unified approach to discover style influences from photos. First, we extract a vocabulary of visual styles from unlabeled, timestamped photos. We consider two sources of photos: geolocated social media photos from around the world, and catalog product photos from an online retailer. Each style is a mixture of detected

---

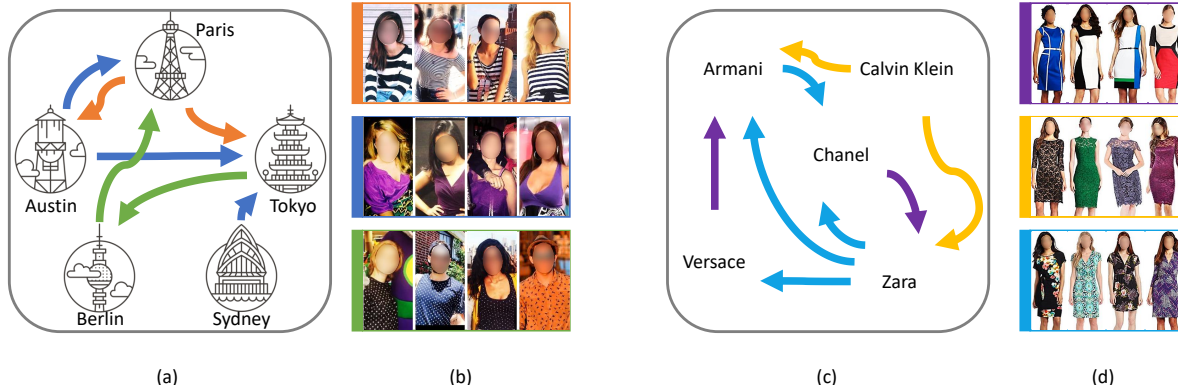[1]The Devil Wears Prada: Cerulean influence https://bit.ly/3dBAQ5W

Figure 1. Fashion styles propagate according to certain patterns of influence around the world. Our idea is to learn visual styles from large-scale (b) social media photos or (d) catalog product images, and then discover style influence relations among (a) major cities worldwide or (c) fashion brands. Our model leverages these relations to accurately forecast future fashion trends. City icons are attributed to https://bit.ly/2WDyFqB

visual attributes. For example, one style may capture short floral dresses in bright colors (Fig. 1d). Next, we record the past trajectories of each style's popularity, *i.e.* the frequency with which it is seen in the photos (or purchased from the catalog) over time.

To build a predictive model, we identify two key properties of an influencer—time precedence and novelty—and use a statistical measure that captures these properties to calculate the degree of influence between cities or brands as well as the influences among the various styles themselves. Next, we introduce a neural forecasting model that exploits the influence relationships discovered from photos to better anticipate future popular styles in any given location or brand. Finally, we propose a novel coherence loss to train our model to reconcile the local predictions with the global trend of a style for consistent forecasts. We demonstrate our approach on two large-scale datasets: GeoStyle [40], which is comprised of everyday photos of people with a wide coverage of geographic locations, and AmazonBrands [42], which is comprised of catalog photos of clothing. We gauge popularity for the visual styles using the frequency the various garments are worn or purchased, respectively.

Our work is the first to model and quantify fashion influence relations between cities and brands from real-world images. Our results shed light on the spatio-temporal migration of fashion trends across the world—revealing which cities are exerting and receiving more influence on others, which most affect global trends, which contribute to the prominence of a given style, and how a city's degree of influence has itself changed over time. Our findings hint at how computer vision can help democratize our understanding of fashion influence, sometimes challenging common perceptions about what parts of the world are driving fashion (consistent with designer Geoffrey Beene's quote above). Furthermore, our results examining influence between fashion brands exposes the latent effects that one line of clothing can have on another, as well as brands that

are relatively impervious to the influence of any others. In addition, we demonstrate that by incorporating influence, the proposed forecasting model yields state-of-the-art accuracy for predicting the future popularity of styles. Unlike prior work that learns trends with a monolithic worldwide model [2] or independent per city models [40], our influence-based predictions catch the temporal dependencies between when different cities or brands will see a style climb or dip, producing more accurate forecasts.

2. Related Work

Visual fashion analysis, with its challenging vision problems and direct impact on our social and financial life, presents an attractive domain for vision research. In recent years, many aspects of fashion have been addressed in the computer vision literature, ranging from learning fashion attributes [5, 6, 9, 10, 38], landmark detection [55, 57], cross-domain fashion retrieval [36, 26, 13, 34, 60, 31], body shape and size based fashion suggestions [43, 19, 23, 18], virtual try-on [54, 37, 11], clothing recommendation [35, 42, 59, 24], inferring social cues from people's clothes [49, 44, 32], outfit compatibility [33, 22, 56, 15], visual brand analysis [30, 14], and discovering fashion styles [29, 52, 21, 2]. Our work opens a new avenue for visual fashion understanding: modeling influence relations in fashion directly from images.

Statistics of styles Analyzing styles' popularity in the past gives a window on people's preferences in fashion. Prior work considers how the frequency of attributes (*e.g.* floral, neon) changed over time [53, 20], and how trends in (non-visual) clothing meta-data changed for the two cities Manila and Los Angeles [48]. Qualitative studies suggest how recommendation models can account for past temporal changes of fashion [17] or what cities exhibit strong style similarities [27]. However, all this prior work analyzes style popularity in an "after the fact" manner, and looks only

qualitatively at past changes in style trends. We propose to go beyond this historical perspective to forecast future changes in styles' popularity, and we provide supporting quantitative evaluation.

Trend forecasting Most previous methods do not leverage visual information, instead focusing on demand forecasting of fashion based on sales records of clothing items (*e.g.* [46, 28, 51, 3, 45]). Only limited prior work explores forecasting visual styles into the future [2, 40]. The FashionForward model [2] uses fashion styles learned from Amazon product images to train an exponential smoothing model for forecasting, treating the products' transaction history (purchases) as a proxy for style popularity. Similarly, the GeoStyle project [40] uses a seasonal forecasting model to predict changes in style trends per city. Both prior models assume that style trends in different cities are independent from one another and can be modeled monolithically [2] or in isolation [40]. In contrast, we introduce a novel model that accounts for influence patterns discovered across different cities (or brands). Our concept of fashion influence discovery is new, and our resulting forecasting model outperforms the state of the art.

Influence modeling To our knowledge, no previous work tackles influence modeling in the visual fashion domain. The closest study looks at the correlation among attributes popular in New York fashion shows and those attributes seen in street photos, as a surrogate for fashion shows' impact on people's clothing; however, no influence or forecasting model is developed [10]. Outside the fashion domain, models for influence are developed for connecting text in news articles [47], linking video subshots for summarization [39], or analyzing intellectual links between major AI conferences from their papers [8].

In a prior conference paper, we introduced the first influence model based on visual fashion trends that is grounded by the style forecasting task [1]. Our previous model captures influence relations between major cities around the world, and we showed that it discovers interesting influence patterns in fashion that go beyond simple correlations. Building on our earlier findings, this article expands both the model and experiments. In particular, we extend our model to go beyond cities' influences to capture how styles themselves influence each other. We propose a model that leverages both cities' and styles' influence relations, and we show it yields higher forecasting accuracy. Furthermore, we provide an in-depth analysis of the relations discovered by our model and how the measured influence correlates with fashion experts and public opinion on what is fashionable. Finally, we generalize our model to capture not only geographic influences but also influences among fashion *brands*. We demonstrate that our approach captures how major fashion brands influence each other with a new set of

experiments on another public dataset, AmazonBrands.

3. Fashion Influence Model

We propose an approach to model influence relations in fashion based on visual data. We consider two domains of influence: a) *location influence*, where we quantify influence relations between different locations in the world (*e.g.* cities) on what people wear in their everyday life; b) *brand influence*, where we capture the influence relations between fashion brands on what clothing items they sell. Both locations and brands can be seen as equivalent concepts for our influence modeling since both can be represented with a distribution over visual fashion styles. For brevity, we refer to a location or a brand as a *fashion unit* in the following.

Starting with images of fashion garments, 1) we learn a visual style model that captures the fine-grained properties common among the garments (Sec. 3.1); then 2) we construct style popularity trajectories by leveraging images' temporal and fashion unit meta information (Sec. 3.2); 3) we model the influence relations between different units (Sec. 3.3) and styles (Sec. 3.5). Finally, 4) we introduce a forecasting model that utilizes the learned influence relations together with a coherence loss (Sec. 3.4) for consistent and accurate predictions of future style trends (Sec. 3.6).

3.1. Visual Fashion Styles

Our model captures the fashion influence among different fashion units—*i.e.* either locations or brands. We begin by discovering a set of visual fashion *styles* from images of people's garments. Such images could be of people in everyday life and from around the world collected from photo-sharing social media platforms (*e.g.* a public dataset of Instagram images) or catalog images of garments from an online retailer platform (*e.g.* a public dataset of Amazon images).

Let $X = \{x_i\}^N$ be a set of clothing images. We first learn a semantic representation that captures the elementary fashion attributes like colors (*e.g.* cyan, green), patterns (*e.g.* stripes, dots), shape (*e.g.* vneck, sleeveless) and garment type (*e.g.* shirt, sunglasses). Given a fashion attribute model $f_a(\cdot)$ trained on a set of disjoint labeled images, we can then represent each image in X with $\mathbf{a}_i = f_a(x_i)$, where $\mathbf{a} \in \mathbb{R}^M$ is a vector of M visual attribute probabilities. We train a convolutional neural network (CNN) for $f_a(\cdot)$.

Next, we learn a set of fashion styles $S = \{S^k\}^K$ that capture distinctive attribute combinations. Hence, given an image of a new garment x_i, the style model $f_s(\cdot)$ can predict the probabilities of that garment to be from each of the learned styles $\mathbf{s}_i = f_s(\mathbf{a}_i)$. For example, the model $f_s(\cdot)$ can be realized using a Gaussian mixture model with K components (see Sec. 4.2 for details on f_s and f_a). The
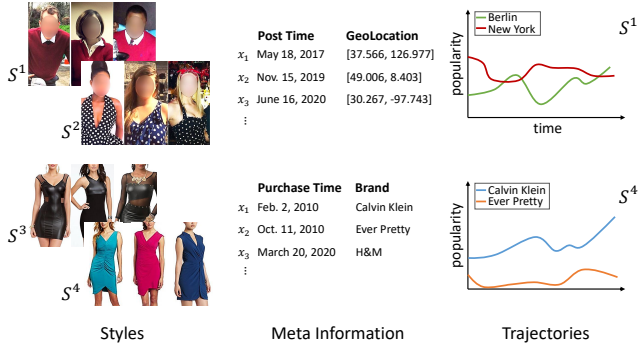
Figure 2. Style trajectories. First, we learn a set of fashion styles from images (left). The images have meta information in terms of either timestamps and geolocations from social media (center top), or purchase time and brand labels from commercial catalog data (center bottom). Then based on the images' meta information, we measure the popularity of a style at a given city or for a given brand and a time period (*e.g.* week) to build up its style popularity trajectory (right).

number of styles K is selected following [2, 40] to capture coherent visual appearance of a mid-level granularity.[2]

Fig. 2 (left) shows a set of fashion styles discovered by our style model from Instagram (S^1 and S^2) or Amazon product (S^3 and S^4) images. Each style is a distribution over the visual attributes, capturing a particular "look" that recurs within the data. For example, in Fig. 2, S^2 corresponds to a black and white spotted and sleeveless top whereas S^4 represents short dresses with uniform color and a v-neck.

3.2. Style Trajectories

We measure the popularity of a fashion style S^i in the context of a certain fashion unit U^j through the frequency of the style in the photos associated with that unit. Specifically, let Q be a set of transactions (*e.g.* social media posts, purchases) such that each $q \in Q$ is a tuple of an image of fashion garment or outfit x, a time stamp t, and a fashion unit id U. When we model city-city influences, the latter is the city; when we model brand-brand influences, it is the brand. We construct a temporal trajectory y^{ij} for each pair of style and unit (S^i, U^j):

$$y^{ij}_t = \frac{1}{|U^j_t|} \sum_{x_k \in U^j_t} p(S^i|x_k),\qquad(1)$$

where U^j_t is the set of images from fashion unit U^j in the time window t, $p(S^i|x_k)$ is the probability of style S^i given image x_k based on our style model $f_s(\cdot)$, and y^{ij}_t is the popularity of style S^i in unit U^j during time t. The time step t can have different temporal resolutions (*e.g.* weeks,

---

[2]Prior work [2] analyzes the impact of K on forecasting: as K increases the styles become less coherent and the noise in the trajectories increases since these styles have less supportive samples across time.

months) depending on the sought granularity of the trajectory. Finally, by getting all values for $t = 1, \ldots, T$ we build the trajectory y^{ij}.

Fig. 2 shows how these trajectories are constructed from social media posts and purchase logs via the large-scale datasets used in our experiments (detailed in Sec. 4.1). As depicted in the top panel, given the timestamps and geolocations of the photos, we quantize these properties to a meaningful temporal resolution (*e.g.* weeks) and locations (*e.g.* cities). We then measure the popularity of each style in a city as in Eq. 1, which captures the frequency with which people in city j are seen wearing the style i over time in the social media photos. Similarly, when modeling fashion brands, we obtain style trajectories representing the popularity of style i over time for brand j in terms of the worldwide purchase rate of garments shown in the brand's catalog photos (Fig. 2 (bottom)).

Next, we describe our influence model that analyzes these trajectories to discover the influence patterns among the various fashion units.

3.3. Fashion Unit Influence Modeling

We propose to ground fashion influence through style popularity forecasting. This enables us to quantitatively evaluate influence using learned computational models based on real world data.

We say fashion unit U^i influences unit U^j in a given fashion style S^n if our ability to accurately forecast the popularity of S^n in U^j significantly improves when taking into consideration the past popularity trend of S^n in U^i, in addition to its past popularity trend in U^j. In other words, past observations in $y^{ni}_{1...t}$ provide us with new insight on the future changes in $y^{nj}_{t+1...t+h}$ that are not available in $y^{nj}_{1...t}$.

Fig. 3 demonstrates our idea for the case where the fashion units are cities. We see that the style trend in *Milan* foreshadows the changes in *Paris* with a time lag of h. By taking into consideration this influence relation, we can accurately predict the future changes in *Paris* (purple curve).
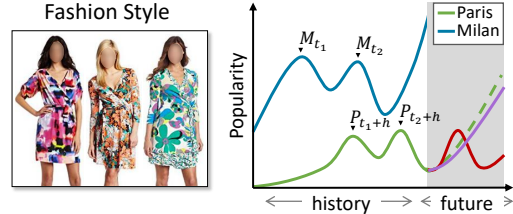

Figure 3. Influence modeling. We propose to ground influence among fashion units with trend forecasting. An influencer city (*e.g.* Milan) has unique cues to accurately predict the future changes of a style trend (purple curve) for the influenced city (*e.g.* Paris). In contrast, forecasting without regard to influence can falter in the presence of complex trends (red curve). We leverage this insight both for inter-city influences as well as inter-brand influences.

On the other hand, modeling the style popularity for *Paris* in isolation leads to significant forecasting error (red curve) since previous observations of style changes for this particular style in *Paris* (*i.e.* in the absence of the influencer) do not provide any cues for the sudden increase in the style's popularity.

We identify two main properties that an influencer U^i should possess: 1) time precedence, that is the influencer unit's changes happen before the observed impact on the influenced unit and 2) novelty, that is the influencer unit has novel past information not observed in the history of the influenced unit.

A naive approach to capture such relations is to use a multivariate model to learn to predict y_t^{nj} by feeding it all available information from the other units. However, this approach does not satisfy the second property for an influencer since it does not constrain the influencer to have novel information that is not present in the influenced entity. Instead, we capture our fashion influence relations using the Granger causality test [12]. The test determines that a time series y^1 Granger causes a time series y^2 if, while taking into account the past values of $y_{1...t}^2$, the past values of $y_{1...t}^1$ still have statistically significant impact on predicting y_{t+1}^2. The test proceeds by modeling y^2 with an autoregressor of order d:

$$y_t^2 = \phi_0 + \sum_{k=1}^{d} \phi_k y_{t-k}^2 + \sigma_t, \quad (2)$$

where σ_t is an error term and ϕ_k contains the regression coefficients. Then the autoregressor of y^2 is extended with lagged values of y^1 such that:

$$y_t^2 = \phi_0 + \sum_{k=1}^{d} \phi_k y_{t-k}^2 + \sum_{l=m}^{q} \psi_l y_{t-l}^1 + \sigma_t, \quad (3)$$

where the third term on the righthand side is a regressor on the lags from y^1. If these extended lags from y^1 do add significant explanatory power to y_t^2, *i.e.* the forecast accuracy of y^2 is significantly better ($p < 0.05$) according to a regression metric (mean squared error), then y^1 Granger causes y^2.

We estimate the influence relations across all units' trajectories for each fashion style S^i. In experiments we consider lags ranging from 1 to 8 temporal steps, meaning up to two months. In this way, we establish the influence relations among units and at which lag this influence occurs. Note that an influence relationship discovered for two cities or two brands indicates that their trends are linked in some predictable way. In particular, a Granger causality tie between two units does not mean that the influenced unit necessarily repeats the shape of the influencer's trajectory at a later time, only that the style trajectory for one unit can be more accurately forecasted by knowing the past trajectory of the other.
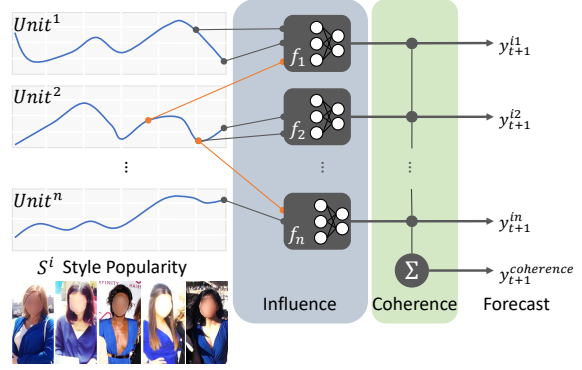

Figure 4. Influence coherent forecaster. Our model captures influence relations between fashion units for a given fashion style (orange connections) and uses them to predict future changes in the style popularity for each unit. Additionally, our model regularizes the forecasts to be coherent with the global trend of the style observed across all units.

3.4. Coherent Style Forecaster

After we estimate the influence relations across the fashion units, we build a forecaster for each trajectory y^{ij} such that:

$$\tilde{y}_{t+1}^{ij} = f_{unit}(L(y_t^{ij}), I(y_t^{ij})|\theta), \quad (4)$$

where $I(y_t^{ij})$ is the set of lags from the influencer of y^{ij} relative to time step t as determined in the previous section, and $L(y_t^{ij})$ are the lags from y^{ij}'s own style popularity trajectory. We model $f(\cdot)$ using a multilayer perceptron (MLP) and estimate the parameters θ by minimizing the mean squared error loss:

$$\mathcal{L}_{forecast} = \sum_{t} (y_{t+1}^{ij} - f(L(y_t^{ij}), I(y_t^{ij})|\theta))^2, \quad (5)$$

where y_{t+1}^{ij} is the ground truth value of y^{ij} at time $t+1$.

Our forecast model in its previous form does not impose any constraints on the forecasted values in relation to each other. However, while we are forecasting the style popularity for each individual unit given the influence from the others, the forecasted popularities $(y_t^{i1}, y_t^{i2} \ldots y_t^{in})$ are still for a common fashion style S^i that by itself exhibits a worldwide trend across all locations or brands.

We propose to reconcile the base forecasts produced at each location through a *coherence loss* that captures the global trend. For all forecasts $\hat{y}_{t+1}^{ij}$ for a fashion style S^i and across all units $U^j \in U$, we constrain the distribution mean of the predicted values to match the distribution mean of the ground truth values:

$$\mathcal{L}_{coherence} = \frac{1}{|U|} (\sum_{k} y_{t+1}^{ik} - \sum_{k} \hat{y}_{t+1}^{ik}). \quad (6)$$

The coherence loss, in addition to capturing the global trend of S^i, helps in combating noise at the trajectory level
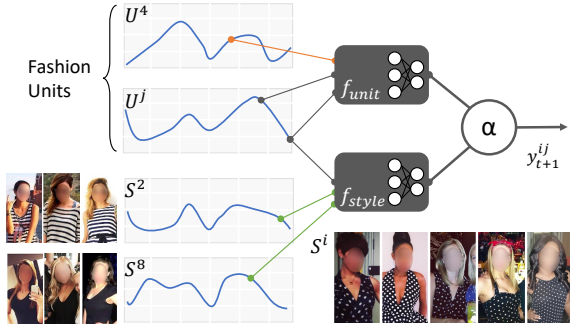
5

Figure 5. Our influence-based model forecasts a style popularity by jointly modeling two types of influence: unit (orange) and style (green) influences.

Style Trajectories   Influence Modeling α-Weighted Layer   Style Forecasts
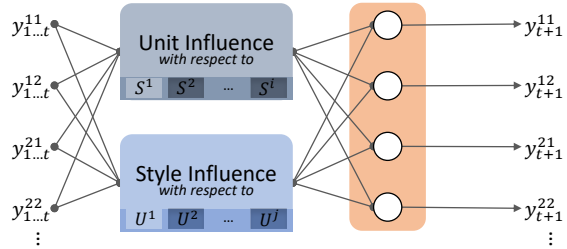
Figure 6. Model overview. For each style trajectory, our model captures the influence of other fashion units and styles on the input trajectories using the Unit and Style Influence modules. Then we use an α-weighted layer to combine the predictions of these modules and forecast the future trend of the styles. This figure unifies the model components depicted in Figures 4 and 5.

of each unit through regularizing the mean distribution of all forecasts. The model is trained with the combined forecast and coherence losses:

$$\mathcal{L} = \mathcal{L}_{forecast} + \lambda \mathcal{L}_{coherence}. \tag{7}$$

Fig. 4 illustrates our model. For a style S^i, we model its popularity trajectory for each fashion unit with a neural network of two layers and sigmoid non-linearity. The input of the network is defined by the lags from its own trajectory (shown in black) and any other influential lags from other units discovered by the previous step (Sec. 3.3), which are shown in orange. Furthermore, the output of all local forecasters is further regularized to be coherent with the overall observed trend of S^i using our coherence loss.

3.5. Style Influence Modeling

We have so far modeled how fashion units (cities or fashion brands) influence each other for a given fashion style. However, another type of influence to consider is how fashion styles impact each other's popularity. For example, a spike in popularity of solid purple v-necks could foreshadow a spike in v-necks with more complicated textures and colors, or an upward tick in more conservative styles may foreshadow a downward trend in less conservative styles. To capture these links within our model, we represent how a style S^i influences another style S^j in a unit U^n in a similar manner as in Sec. 3.3 by simply swapping styles and units.

3.6. Influence-based Forecasting

Our final model forecasts the popularity of a style based on both types of influence:

$$\hat{y}_{t+1}^{ij} = \alpha f_{style}(L(y_t^{ij}), I(y_t^{ij})|\theta_s) + \\ (1-\alpha)f_{unit}(L(y_t^{ij}), I(y_t^{ij})|\theta_u), \tag{8}$$

where f_{style} and f_{unit} are instances of our influence-based coherent forecaster (Eq. 4) that capture style and unit influence relations, respectively. Fig. 5 shows our joint model of

both types of influence relations. Each f_{style} and f_{unit} is trained for its respective loss in Eq. 7. The hyperparameter α is learned over the validation data to weight each type of influence contribution to the final forecast. We model theses influence relations using Eq. 8 for each of the trajectories in our data with respect to all fashion units and styles, as shown in the model overview in Fig. 6.

4. Evaluation

In the following experiments, we demonstrate our model's ability to forecast fashion styles' popularity changes by utilizing discovered influence relations. Furthermore, we analyze the influence patterns revealed by our model between major cities and brands, how they influence global fashion trends, and their influence dynamics trends through time.

4.1. Datasets

We evaluate our approach on two datasets:

GeoStyle [40] This dataset, collected by researchers at Cornell University, extends the StreetStyle dataset [41]. It is based on public Instagram photos showing people from 44 major cities from around the world. Since a photo may contain multiple persons and a cluttered background, the images are first preprocessed with a person detector to extract the regions of interest (see [41] for details). In total, the dataset has 7.7 million images that span a time period from July 2013 until May 2016.

AmazonBrands This dataset is derived from the data collected by researchers at the University of California [42] from the Amazon website which contains images of garments along with their transaction history (purchases). We use the *dresses* subset [2] which has $41K$ images. We consider a list of the most famous fashion brands and match

(a) GeoStyle   (b) AmazonBrands

Figure 7. Examples of the learned fashion styles from GeoStyle (a) and AmazonBrands (b) datasets. For each dataset, each row shows two example styles, with five sample images for each.

their names with the textual description associated with each sample in order to label each garment with its respective brand. Finally, we select the top 10 most frequent brands and their corresponding $4K$ images and $21K$ transactions to use in our experiments. The samples span a time period from March 2012 until July 2014. For this dataset, no person detection is needed since the catalog photos have white backgrounds with the fashion item centered. We use both datasets for research purposes only.

These two datasets provide us with different perspectives on people's preferences in fashion. While in GeoStyle we estimate a style's popularity based on its frequency in social media posts (*i.e.* what you wear is what you like), in AmazonBrands the popularity is estimated based on purchase frequency (*i.e.* what you buy is what you like).

However, both datasets, like any Internet photo dataset, have certain biases based on their respective data sources and sampling methods. These biases may affect the type of styles considered and their measured popularity. For example, images in GeoStyle are usually of young people in major cities since younger generations are more likely to upload photos to Instagram and from places with easy access to high Internet bandwidth [41]. On the other hand, while fashionability may be the main drive to select a given dress in AmazonBrands, other factors like price and brand loyalty could contribute to one's decision to buy as well. Nonetheless, the GeoStyle dataset is the largest public fashion dataset with the most temporal and geographic coverage, providing a unique glimpse on people's fashion preferences in daily life around the globe. Similarly, the AmazonBrands data gives us a unique insight on the influence dynamics among major fashion brands.

4.2. Fashion Styles

A style is a combination of certain attributes describing materials, colors, cut, and other factors, such as *V-neck, red,*

formal dress. Hence, we adopt an attribute-based representation of images for style learning.

For GeoStyle we use attribute predictions from [40] to represent each photo with $M = 46$ fashion attributes (*e.g.* colors, patterns and garment types). The attribute predictor is a multi-task CNN (GoogLeNet [50]) with separate heads to predict separate attributes. This yields an M-dimensional vector of attribute probabilities per image. Based on these, we learn $K = 50$ fashion styles using a M-dimensional Gaussian mixture model: each style is a mode in the data capturing a distribution of attributes. Then, we get the style probabilities of a garment by calculating the posterior probability of each mode in the mixture with respect to the inferred attributes' probabilities from the photo.

For AmazonBrands we train a ResNet-18 model [16] to predict $M = 1000$ fashion attributes learned from the Deep-Fashion dataset [38]. We learn fashion styles from these attributes using a non-negative tensor factorization (NMF) style model similar to [2]. That is, given a tensor representing the samples in the dataset by their attributes, we factorize it into two matrices using a set of learned latent variables. Each of these latent variables represents a fashion style. The NMF model uses a probabilistic formulation [25] and each discovered style captures a distribution of attributes in a similar fashion to the previous GMM model. We find that for the high-dimensional attribute representation in AmazonBrands, NMF produces more coherent and diverse visual styles compared to the GMM model, which seemed to collapse on a few attribute dimensions for this data. We learn $K = 20$ styles, which was sufficient to cover the diverse styles in this dataset since AmazonBrands is focused on one kind of garment—dresses.

Fig. 7 shows examples of the fashion styles learned by our style models from the GeoStyle and AmazonBrands datasets. We notice that in GeoStyle (Fig. 7a) some of the learned fashion styles may reflect a season-related type of

garment (*e.g.* the yellow jacket and scarf style), or a local traditional or cultural clothing (*e.g.* upper left row). However, many of the learned styles are common across different countries and cultures. Differently, the styles learned in AmazonBrands (Fig. 7b) do not suggest any clear seasonal variations. They capture a coherent visual representation that can be described with signature attributes like the slim-fit leopard dresses (upper right row) or the casual long-sleeved short dresses (lower left row). In both datasets, since each style is a mixture of attributes, it captures a recurring *configuration* of properties, rather than one isolated property (*e.g.* leopard-print plus tight fit; blazer paired with V-neck t-shirt). Thus styles offer a coarser representation than individual clothing items, and they emerge bottom-up from the data.

4.3. Style Trends Forecasting

We evaluate how well our model produces accurate forecasts by leveraging the influence patterns, and we compare it to several baselines and existing methods that model trajectories in isolation [2, 40].

4.3.1 Trajectories and data splits

For GeoStyle, we infer the popularity trajectory of each style in each city based on its frequency in the images. Additionally, to quantify the impact of possible seasonal yearly trends in fashion styles, we also consider forecasting the deseasonalized style trajectories. To do so, we subtract the yearly seasonal lag from the trajectories. The deseasonalized test is interesting because it requires methods to capture the more subtle visual trends not simply associated with the location's weather and annual events. For AmazonBrands, we build the temporal trajectory of each style and brand based on its transaction frequency.

For both datasets, we use a temporal resolution of weeks. We adopt the long-term forecasting data split from [40]. That is, we allocate the last 26 weeks of each trajectory for testing, the previous 4 for validation, and the rest for training.

4.3.2 Implementation details

We set $\lambda = 1$ for the coherence loss weight (see Eq. 7) and optimize our neural influence model using Adam for stochastic gradient descent with a learning rate of 10^{-2} and l_2 weight regularization of 10^{-8}. We select the best model and $\alpha \in [0, 1]$ from Eq. 8 that controls the contributions of unit and style influences based on the performance on a disjoint validation split.

We compare all models using the forecast accuracy captured with the mean absolute error $\mathrm{MAE} = \frac{1}{n}\sum_{t=1}^{n}|\hat{y}_t - y_t|$, which measures the absolute difference between the forecasted $\hat{y}_t$ and ground truth y_t values, and the mean absolute percentage error $\mathrm{MAPE} = \frac{1}{n}\sum_{t=1}^{n}|\frac{\hat{y}_t - y_t}{y_t}| \times 100$, which measures the forecast error scaled by the ground truth values, following prior work [2, 40]. The models are evaluated on forecasting a long time horizon of 26 temporal steps, meaning six months forward in time.

4.3.3 Baselines and existing methods

We arrange the baselines into three main groups:

Naive models: these models rely on basic statistical properties of the trajectory to produce a forecast. We consider five variants of these baselines:

- Gaussian: this model fits a Gaussian distribution based on the mean and standard deviation of the trajectory and forecasts by sampling from the distribution.

- Seasonal: this model forecasts the next step to be similar to the observed value one season before $y_{t+1} = y_{t-season}$. We set a yearly season of 52 weeks.

- Mean: it forecasts the next step to be equal to the mean observed values $y_{t+1} = \mathrm{mean}(y_1, \ldots, y_t)$.

- Last: it uses the value at the last temporal step to forecast the next $y_{t+1} = y_t$.

- Drift: it forecasts the next steps along the line that fits the first and last observations.

Per-Trajectory models: These models fit a separate parametric model trained on the history "lags" of each of the trajectories [7], without accounting for relationships between the units (locations or brands) or styles.

- AR: a standard autoregression model.

- ARIMA: a standard autoregressive integrated moving average model.

- FashionForward [2]: an exponential decay model which forecasts based on a learned weighted average of the historical values.

- GeoModel [40]: a parametric seasonal forecaster which models a trajectory with a sinusoidal model and drift components.

To our knowledge, the two existing methods [2, 40] represent the only prior approaches for style forecasting from images. Further, unlike our approach, all of these models consider the popularity trajectories of the styles in isolation, *i.e.* they do not take into consideration possible interactions among the units or styles.

Across-Trajectories models: The next set of baselines does model dependencies between the units or styles. To

Table 1. Forecast errors of fashion style trajectories on the GeoStyle and the AmazonBrands datasets.

| | GeoStyle | | | | AmazonBrands | |
| | Seasonal | | Deseasonalized | | | |
Model	MAE	MAPE	MAE	MAPE	MAE	MAPE
Naive						
Gaussian	0.1301	33.23	0.1222	26.08	0.1268	35.33
Seasonal	0.0925	22.64	0.1500	33.39	0.1329	36.48
Mean	0.0908	23.57	0.0847	18.97	0.0964	26.29
Last	0.0893	22.20	0.1053	23.08	0.1134	31.41
Drift	0.0956	23.65	0.1163	25.32	0.1207	33.53
Per-Trajectory Models						
AR	0.0846	21.88	0.0846	18.95	0.0934	25.61
ARIMA	0.0919	23.70	0.1033	22.70	0.1095	30.77
FashionForward [2]	0.0779	19.76	0.0848	18.94	0.0930	25.56
GeoModel [40]	0.0715	17.86	0.0916	20.31	0.1028	28.41
Across-Trajectories Models						
VAR - All Units	0.0771	19.25	0.0929	20.41	0.0955	26.12
VAR - All Styles	0.0752	19.15	0.0939	20.78	0.0949	26.00
Influence-based (ours)	**0.0688**	**17.13**	**0.0814**	**18.10**	**0.0914**	**24.86**

Table 2. Ablation study of our model on the GeoStyle deseasonalized style trajectories.

Model	MAE	MAPE
Ours (full model)	0.0814	18.10
Style Influence Only	0.0825	18.32
City Influence Only	0.0824	18.29
w/o Influence	0.0859	19.24
w/o Influence & Coherence	0.0942	20.62

represent this group, we use the VAR model [7], which fits a parametric model trained on the trajectories of a style across all units (VAR - All Units) or the trajectories from one unit and across all styles (VAR - All Styles). Such models assume a full and simultaneous interaction between all units or styles.

4.3.4 Forecasting results

Table 1 shows the performance of all models when forecasting the styles' future popularity. The proposed model outperforms all the naive, per-trajectory, and across-trajectories models, attaining the lowest forecast errors. This shows the value of discovering influence for the quantitative forecasting task. We notice that the styles' popularity trajectories in GeoStyle do have a strong seasonal component: seasonal models (like GeoModel [40] and Seasonal) do well compared to non-seasonal ones (like AR and Drift), but still underperform our approach. This ranking changes on the deseasonalized test of GeoStyle and on Amazon-Brands, where models like FashionForward [2] and AR outperform the seasonal ones. Our model outperforms all competitors on both types of trajectories, which demonstrates the benefits of accounting for influence. More-

over, we compare our model's predictions to the best "non-influence" per-trajectory or across-trajectories competitors on both datasets using a two-sided t-test with the null hypothesis that the expected forecast error is identical. We found that our model's improvements are statistically significant with $p < 0.05$.

The ablation studies in Table 2 show the impact of each component of our model. We evaluate our approach using only style-based (Sec. 3.5) or city-based (Sec. 3.3) influence modeling, as well as two additional versions—one without any influence modeling, *i.e.* it assumes a full interaction pattern among all cities, and a second version that is not trained for coherent forecasts (Sec. 3.4). We see that the joint modeling of the two types of influence leads to better performance. Furthermore, we notice a larger drop in accuracy when the model does not account for influence and coherence, which shows the importance of these concepts for accurate predictions of popularity changes in the future.

Next, we analyze the learned α from Eq. 8 for our influence models from both datasets. Interestingly, on GeoStyle and for the seasonal and deseasonalized splits, our model assigns equal importance for both city and style influence $\alpha = 0.5$; however, on AmazonBrands the model assigns higher weight for style influence $\alpha = 0.6$ compared to brand influence which is weighted with 0.4. As we will see later in influence relations analysis (Sec. 4.4) this can be attributed to the wide spread of the *monopoly* type of relations between brands and styles (see Fig. 11) which leads to a more limited direct pairwise interactions between brands compared to what we observe among cities. This is an interesting quantitative finding that agrees with an intuition that styles "spread" very organically between different parts of the world as people adopt new trends seen elsewhere,
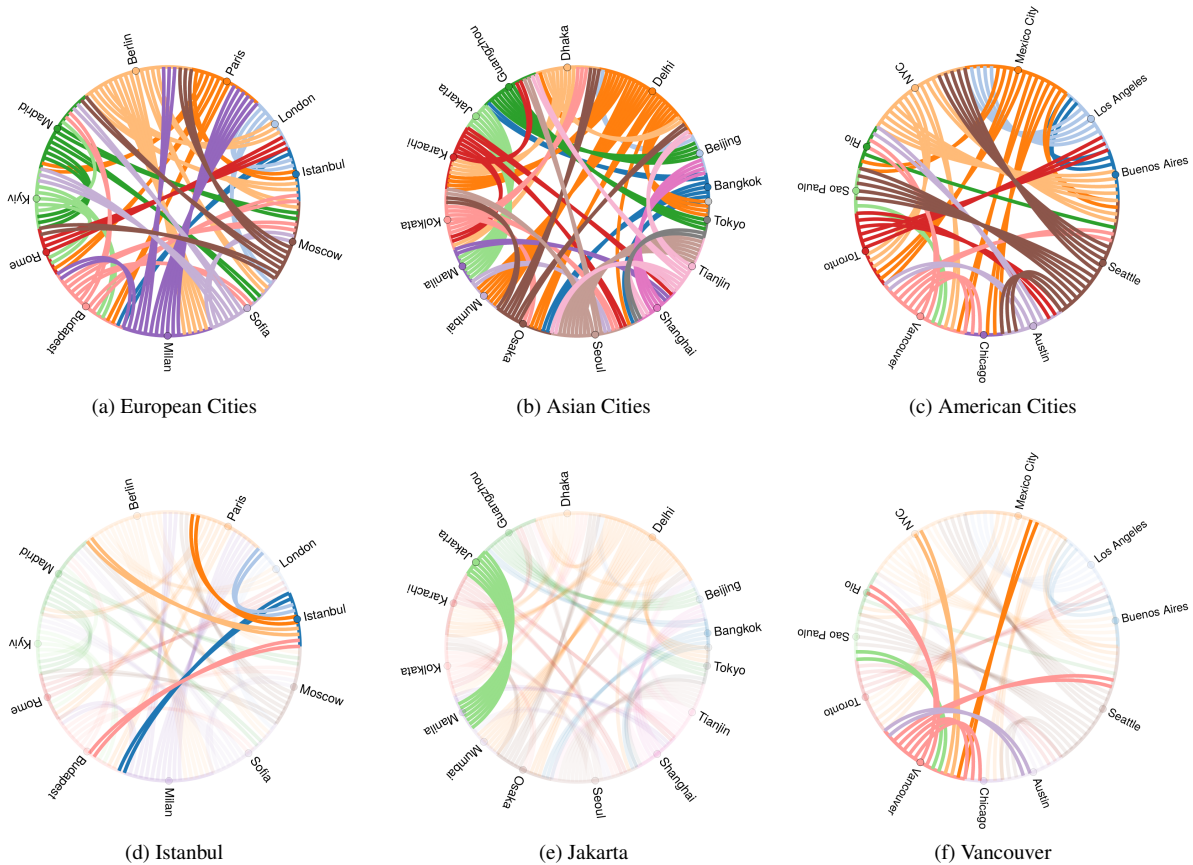
Figure 8. Style influence relations discovered by our model among European (a), Asian (b) and American (c) cities. The number of chords coming out of a node (*i.e.* a city) is relative to the influence weight of that city on the receiver. Chords are colored according to the source node color, *i.e.* the influencer. Our model discovers various types of influence relations from multi-city (*e.g.* Paris and New York City) and single-city (*e.g.* Jakarta) influencers to cities that are mainly influence receivers (*e.g.* Istanbul and Beijing) or influence focal points that exert and receive influence from multiple cities (*e.g.* Vancouver).

whereas distinct corporate brands may be more resistant to such intermingling.

4.4. Influence Relations

The results thus far confirm that our method's discovered influence patterns are meaningful, as seen by their positive quantitative impact on forecasting accuracy. Next, we analyze them qualitatively to understand more about what was learned. We consider influence interactions along two axes: 1) a local one that looks at pairwise influence relations among the cities and the brands; and 2) a global one which examines how cities and brands influence the world's fashion trends.

1) Unit → Unit influence For each visual style, our model estimates the influence relation between units and at which temporal lag, yielding a tensor $\mathbf{B} \in \mathbb{R}^{|U| \times |U| \times |S|}$ such that B_{ij}^k is the influence lag of U^i on U^j for style S^k. By averaging these relations across all visual styles, we get an estimate of the overall influence relation between all units

weighted by the temporal length, *i.e.* long term influencers are given more weight than instantaneous ones. We visualize the influence relations using a directed graph where each node represents a unit, and we create a weighted edge from unit U^i to U^j if U^i is found to be influencing U^j.

Fig. 8 shows an example of the influence pattern for fashion styles discovered by our model among major European (Fig. 8a), Asian (Fig. 8b), and American (Fig. 8c) cities, where the number of connections between two cities is relative to the weight of the influence relation. Our model discovers interesting patterns. For example, there are a few fashion hubs like *Paris* and *Berlin* which exert influence on multiple cities while at the same time being influenced by few (one or two) cities. *Paris* influences four cities in Europe while being influenced by *Milan* only (Fig. 8a). Other cities like *Vancouver* have exerted and received influence relations with multiple cities (Fig. 8f). Cities like *Jakarta* have a one-to-one influence relation with *Manila* (Fig. 8e). On the other end of the spectrum, we find cities like *Istanbul* (Fig. 8d) and *Beijing* (Fig. 8b) that mainly receive influence

(a) Fashion Brands   (b) Calvin Klein   (c) Jessica Howard

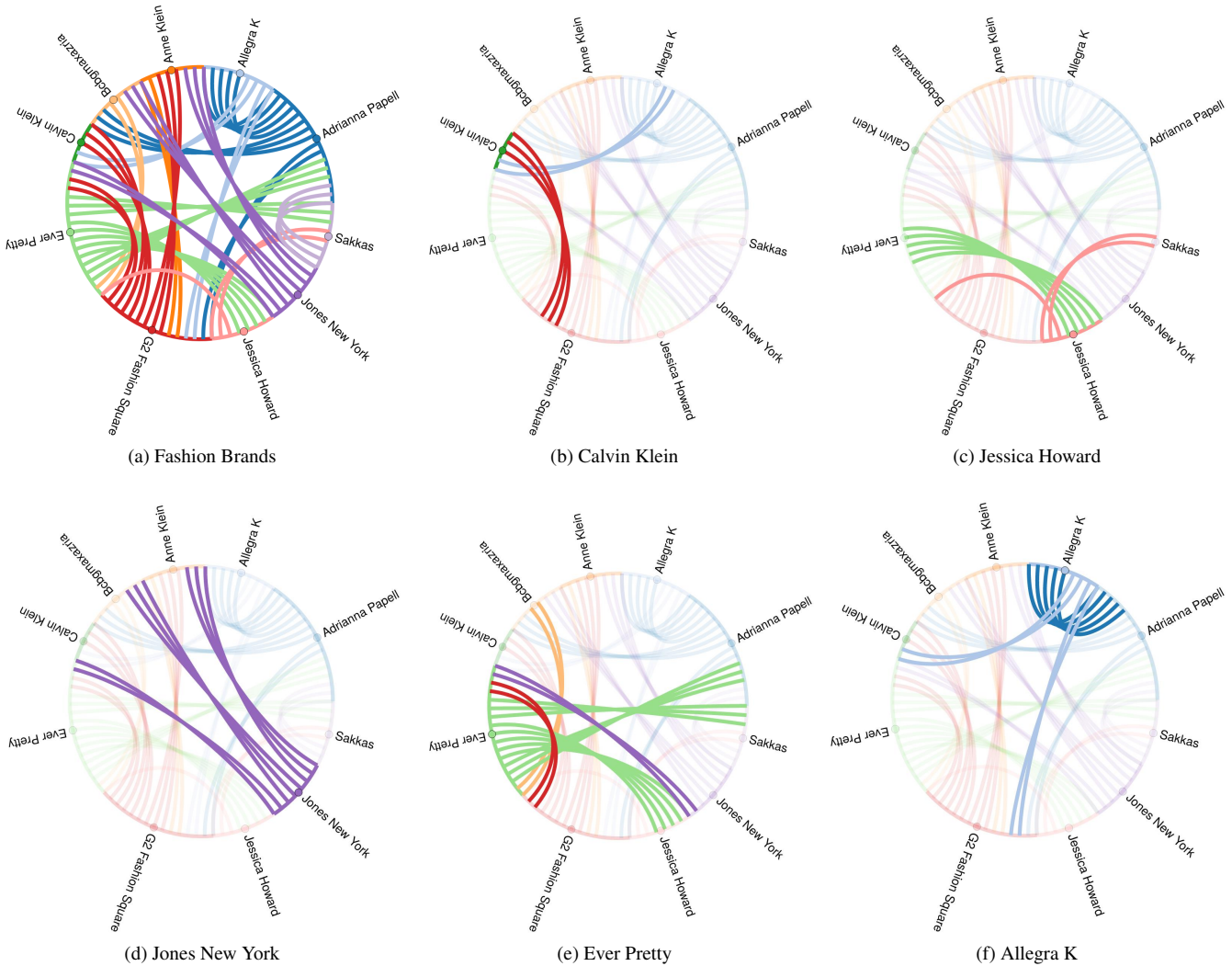(d) Jones New York   (e) Ever Pretty   (f) Allegra K

Figure 9. Style influence relations discovered by our model among fashion brands (a). The number of chords coming out of a node (*i.e.* a brand) is relative to the influence weight of that brand on the receiver. Chords are colored according to the source node color, *i.e.* the influencer. While some brands have limited influence interactions (*e.g.* Calvin Klein), others show tendency towards mainly receiving (*e.g.* Jessica Howard) or exerting (*e.g.* Jones New York) influence. Fashion brands like Ever Pretty have diverse and more balanced influence relations with the rest.

from multiple sources while influencing few.

Fig. 9 shows the influence relations discovered by our model among fashion brands from the AmazonBrand dataset. We see here similar patterns as those discovered for cities, where we can identify fashion hubs like *Jones New York* (Fig. 9d); influence receivers like *Calvin Klein* (Fig. 9b); and influence focal points like *Ever Pretty* (Fig. 9e).

2) Unit → Global influence Alternatively, we can analyze the influence relation between a unit and the global trend for a specific style. This helps us better understand who are the main influencers on the global stage for each of the styles. We capture this relation by modeling the interaction of a unit's popularity trajectory on the global one (*i.e.* the

observed trend of the collective popularity of the same style across the world).

Fig. 10 shows a set of Asian (Fig. 10a) and American (Fig. 10b) cities and their influence on global fashion styles. We see that for some of the fashion styles, like S^4, S^5 and S^7, a couple of cities maintain a monopoly of influence on them, whereas others, like S^3, S^2 and S^9, are influenced almost uniformly by multiple cities. Our influence model also reveals the influence strength (measured by the temporal lag) of these cities relative to their peers at the world stage. See for example the strong influence of *Seoul*, *Bangkok*, *Chicago* and *Mexico City* compared to the delicate one of *Manila*, *Jakarta* and *Sao Paulo*, as represented by the width displayed for their respective influence relation to the global
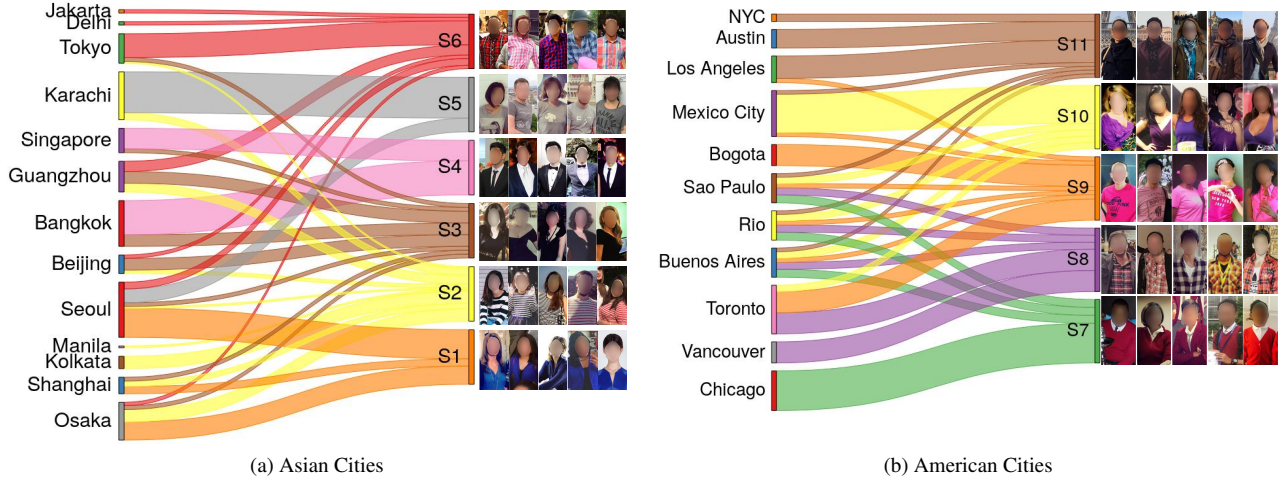
Figure 10. Discovered influence by our model of Asian (a) and American (b) cities on global trends of fashion styles learned from the GeoStyle dataset. The width of the connection is relative to the influence weight of that city in relation to other influencer of the same style.
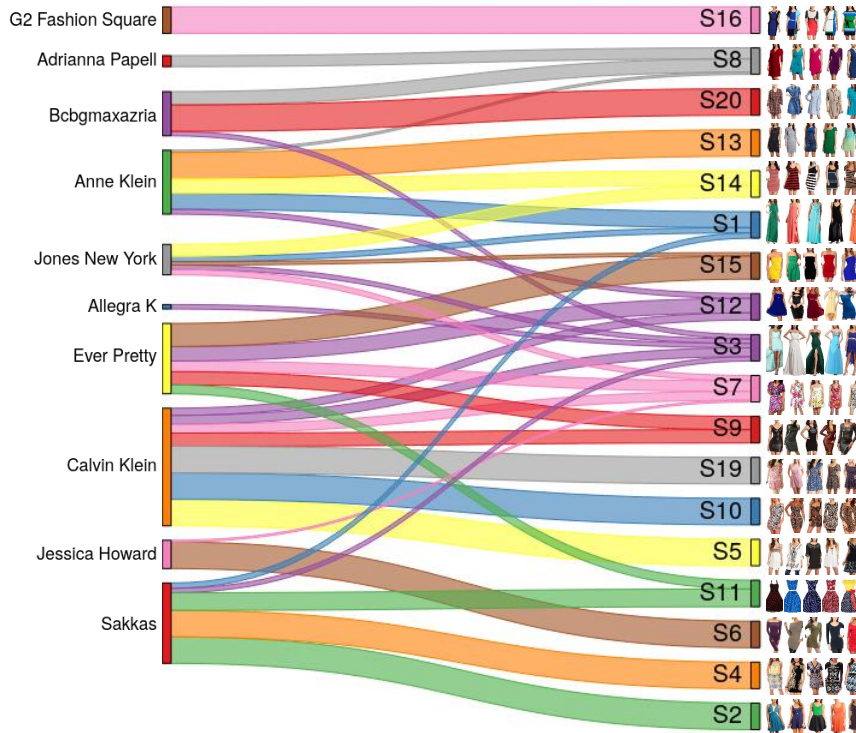

Figure 11. Discovered influence by our model of fashion brands on global trends of 20 dresses' fashion styles learned from the AmazonBrand dataset. The width of the connection is relative to the influence weight of that brand in relation to other influencers of the same style.

trends.

Fig. 11 show the fashion brands and their influence on global fashion styles for dresses. Interestingly, we see here that the influence monopoly pattern is more prominent than in the GeoStyle dataset. This is expected since fashion brands usually try to develop their own signature styles, hence they are more focused on a subset of styles where they can maintain an impact on their market share. More-

over, this view can provide us with unique information on the fashion industry. While a brand like *Calvin Klein* has little pairwise influence interaction with others (Fig. 9b), it maintains an influence monopoly at the global scale on a larger and diverse set of styles than the rest. We can also identify main brand competitors and on which styles. For example, *Ever Pretty* & *Calvin Klein* are the main competitors on style S^9, and *Jones New York* and *Anne Klein* on
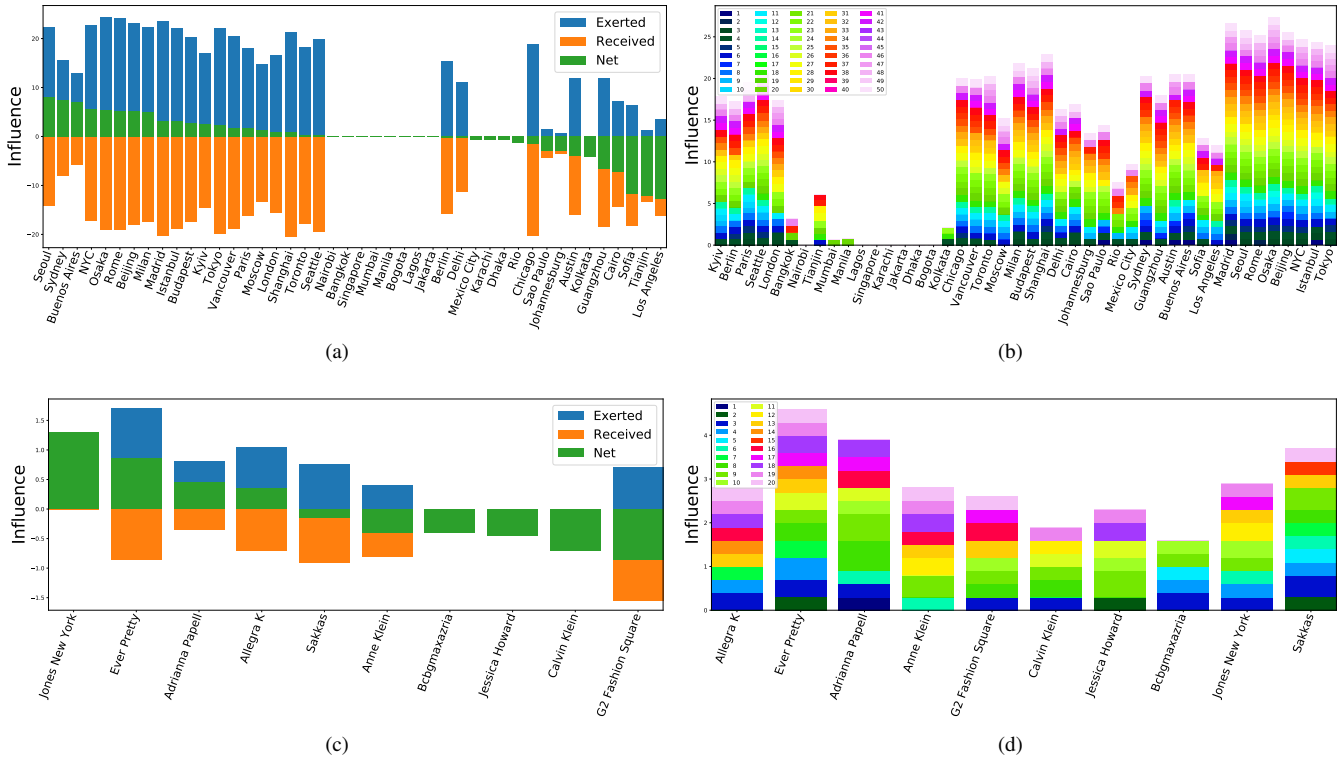
Figure 12. Left: Global ranking of major (a) cities and (c) brands according to their fashion influence on their peers. The fashion units are sorted by their *net* influence score (green). The fashion units with no bars indicate that they do not have influential relations with above average weight. Right: The exerted influence score is split into influence score per individual style for each (b) city and (d) brand (sorted by style influence similarity).

style S^{14}. Other styles like S^3 and S^7 exhibit a dispersed influence pattern with no clear main influencer.

4.5. Influence Ranking

We rank all cities in the GeoStyle dataset according to their accumulated influence power on their peers. That is, we assign an influence score for each city according to the sum of weighted influence relations *exerted* by that city on the rest. Similarly, we also calculate the sum of *received* influence as well as the difference in both as the *net* influence score.

Fig. 12a shows these three influence scores for all cities across the world, sorted by the *net* score. The ranking reveals that some cities, like *London* and *Seattle*, act like focal points for fashion styles, *i.e.* they receive and exert a high volume of influence simultaneously. Others, like *Seoul* and *Osaka*, have a high net influence, which could indicate having some unique fashion styles not influenced by external players. We see in Fig. 12c a similar pattern among the fashion brands ranking, however at a lower scale compared to what is observed among the cities. While the most influential cities may interact with up to 20 cities, an influential brand interacts with around 1.6 other brands on average.

Furthermore, breaking down the exerted influence score

for each city to per-style influence scores, we see in Fig. 12 (right) that we can identify and group influencers into "teams" based on their common set of styles where they exert their influence. For example, *Chicago*, *Vancouver*, and *Toronto* constitute a team since they seem to be influencing similar sets of visual styles (Fig. 12b). Likewise, *Jones New York* and *Sakkas*, and *Calvin Klein* and *Jessica Howard* share comparable style influence profiles (Fig. 12d).

Fig. 13 provides a global view of fashion influence on the world map. It shows the exerted influence score of each city represented by its circle size and color. We notice that most of the influential cities (among the 44 world cities present in the GeoStyle dataset) lay in the northern hemisphere and specifically in its upper part. Furthermore, we aggregate the influence score inferred by our model by country. We see that while some individual Asian and American countries have high fashion influence scores (*e.g.* *Japan*, *South Korea*, and *Canada*), most of the influential fashion cities and countries are located in *Europe* which leads in terms of global fashion influence according to these social media images.
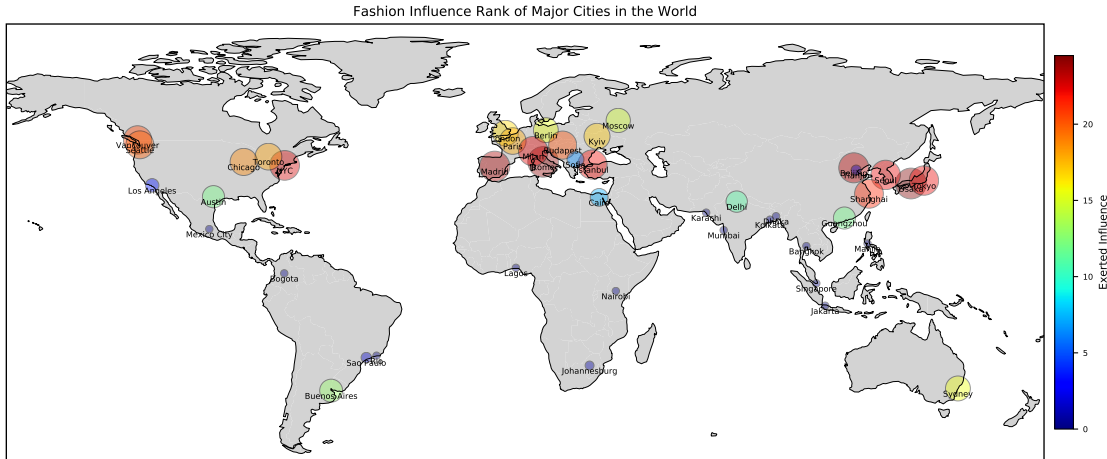
Figure 13. Fashion influence scores as inferred by our model from everyday images of people around the world at the city level.

4.6. Visual Influence vs. People's Perception of Fashionability

Our work is the first to quantify fashion influence relations directly from large-scale visual data. We stress that the trends visible in the photos are exactly what our model measures; there is no separate "ground truth" against which to score the influence measurements. However, to gain more insight into what visual fashion influence captures, we next explore how the patterns we have discovered align (or not) with existing metrics for related properties that are derived from means other than images. In particular, we analyze the alignment between a) how our model ranks cities by their influence and b) how fashion experts rank cities by their fashionability. "Fashionability" means the extent to which a unit (city) exhibits the popular clothing styles. This is a distinct property from influence, which means the extent to which a unit affects another unit's styles—here, in the Granger causality sense tied to forecastability. Nonetheless, they are linked concepts and their degree of alignment can help us understand how well the image-driven trends our model discovers agrees with the manually-made conclusions of fashion experts and enthusiasts.[3]

We consider two types of fashionability ranking:

1. *Expert ranking*: this ranking is based on fashion journalists' assessment and takes into account factors like how many fashion shows, luxury brands head quarters, and fashion schools are located in each city. For this type, we consider a ranking from Zalando, a fashion online retailer based in Europe, for the most elegant cities in the world [58]. While this ranking also considers 'urban factors' such as cleanliness and architecture, here we only consider the ranking produced based on

the 'fashion factors' like the ones mentioned earlier.

2. *Public ranking*: this ranking is based on people's perception on which cities are more fashionable. For this type, we consider a ranking generated by the Internet Fashion Map project [4] that ranks each city based on how often people discuss that city in social media (*e.g.* Twitter, Facebook) in the context of fashion using language-based statistics.

We use the Spearman rank correlation coefficient to compare our exerted and received influence-based ranking from the GeoStyle dataset to these expert and public fashionability rankings. Interestingly, we find that our *exerted* influence ranking is strongly aligned with public perception of fashionablity when compared to expert opinion, 0.87 to 0.31 respectively. This is consistent with the data source for our model: our model discovers influence relations based on images of people's clothing from everyday activities, which makes it better aligned with public opinions in social media. In contrast, the expert opinions are explicitly scoring cities based on high-end fashion shows and fashion from magazines. Moreover, we find that the *received* influence ranking has no correlation with fashionability, scoring only 0.05 for both expert and public ranking. This shows that for both the experts and the public, a fashion trend setter (*i.e.* with high exerted influence score) is perceived as more fashionable compared to an influence receiver.

4.7. Visual Influence Correlations with Cities' Properties

Next we analyze the correlations of influence relations discovered by our model with known real-world properties of the cities. While the influences discovered in the photos need not align with these properties—the data-driven influences stand on their own— it is nonetheless interesting to
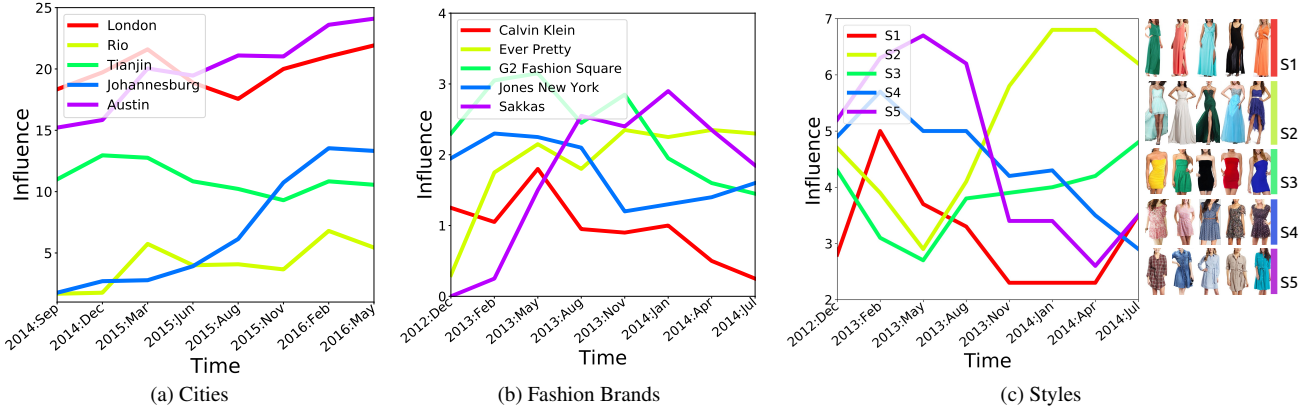
---

[3]We focus this analysis on cities, as we are not aware of similar resources that succinctly record the perceived fashionability of brands.

Figure 14. Dynamics analysis of exerted fashion influence at multiple time steps (with a three month interval) reveals the cities' (a), the brands' (b) and the styles' (c) temporal changes in influence strength.

Table 3. Correlations of the discovered influence patterns with meta information about the cities.

| Meta Info. | Fashion Influence | |
	Direction	World Rank
GDP	0.037	0.373
Temperature	- 0.319	- 0.616
Latitude	- 0.348	0.596
Population	0.038	- 0.193
Distance	- 0.165	n/a
Num. Samples	- 0.148	0.086

understand what external factors may contribute to fashion influence. Correlating against other properties simply helps unpack what the trends do or do not relate to.

We collect information about the annual gross domestic product (GDP), the geolocation, the population size, and the yearly average temperature for each of the cities. We calculate the correlations of these properties with the influence information discovered by our model at two levels: 1) influence world ranking (*i.e.* does a high influence rank correlate with the population size of the city? do cities with warm weather have a high influence score?), and 2) relation direction (*i.e.* does influence flow from high to low GDP cities? do cities influence those that are geographically close to them?).

World rank In Table 3 (second column) we correlate the discovered influence ranking of all the cities with the ranking derived from each of the meta properties using the Spearman coefficient. The correlation uncovers some curious cases. We see a weaker but positive correlation with GDP, *i.e.* a higher GDP could be a faint indicator of a higher fashion influence. There is an above average correlation between the city influence rank and its latitude; many of the influential fashion cities are on the northern hemisphere. This reinforces our previous observation from Sec. 4.5 that most influential cities are ofter of higher latitude. Finally, we ob-

serve a negative correlation with average temperature; influential fashion cities are often colder.

Relation direction Table 3 (first column) shows the correlation of the influence directions discovered by our model with differences in each of the meta properties between the influencer and the influenced city. Specifically, for each city and meta property P_i (*e.g.* GDP), first we measure the differences between that city and the rest in regards to P_i, then we correlate these differences with the influence exerted by that city. Interestingly, none show high correlation with fashion influence directions. The relation type cannot be reliably estimated based on the differences in GDP (*e.g.* high GDP cities do not always influence lower GDP ones), population (*e.g.* cities with high population do not necessarily influence others with lower population or vice versa), nor distance (*e.g.* influence does not correlate well with how far one city is from its influencer). A weak and negative correlation is found with temperature and latitude differences, showing that cities with similar temperature or at similar latitudes (*i.e.* similar seasonal variations) tend to influence each other slightly more. These results suggest our model discovers complex fashion influence relations that are hard to infer from generic properties of the constituent players.

As a sanity check, we also explore the correlation of the number of image samples collected from each city in the GeoStyle dataset with the two types of influence information (Table 3 last row). We find that there is no strong correlation between the learned influences and the number of images available in the data for each city (*i.e.* influential cities are not those with a higher number of samples in the dataset).

4.8. Influence Dynamics

Finally, we study the changes in the influence rank of fashion units and styles through time. We carry out our influence modeling based on the style trajectories of the var-

15

ious units and styles as before, but at multiple sequential time steps. Then we collect the overall influence score of each unit or style at each step.

Fig. 14a shows the change in the influence score for a subset of five cities spanning different continents. We notice that cities show various dynamic behaviors across time. While some cities like *London* and *Rio* maintain a steady influence score through time (at different levels), others like *Austin* and *Johannesburg* demonstrate a positive trend and are gaining more influence in the fashion domain over time but at varying speeds. Other cities like *Tianjin* exhibit a mild decline in their fashion influence. Similarly, the dynamic influence analysis of fashion brands in Fig. 14b reveals that some fashion brands are gaining increased influence through time like *Sakkas* and *Ever Pretty* while others like *G2 Fashion Square* and *Calvin Klein* are gradually losing some of their exerted influence.

Finally, in Fig. 14c we analyze the influence changes detected by our model for a subset of visual styles from the AmazonBrands dataset. Interestingly, we find that some of the styles like S^1 and S^2 show almost opposite trends in their influence, which may indicate a possible competition of these styles on a similar sector of customers that swing between these styles. We also see a general decline in influence for styles like S^4 and S^5, which could indicate a broad shift in the market away from similar fashion styles.

5. Conclusion and Future Work

We introduced a model to quantify influence of visual fashion trends, capturing the spatio-temporal propagation of styles around the world and between major fashion brands. Our approach integrates both fashion unit and style influence relations along with coherence regularizers to predict the future popularity of a style. Our influence-based model outperforms state-of-the-art in forecasting methods and our analysis of the discovered influences sheds light on intriguing patterns among cities and brands. Our findings suggest potential applications in social science and the fashion industry, where computer vision can unlock trends that are otherwise hard to capture.

We focused on this work on learning pairwise influence relations among various fashion units and styles; however, influence may exist at higher degrees of complexity. Discovering chains of influence among multiple units may provide us with a more comprehensive view of the influence cycle and propagation, and it could help in building more compact influence-based forecasters. Furthermore, modeling fashion influence at a finer granularity, such as individual people, could be an interesting future direction for our work. Influencers on social media platforms are highly visible and interesting to users and marketers alike. Quantifying the impact of these individual influencers on their social network and their visual style trends offers a challenging

scientific problem with impact on real-world applications.

References

[1] Ziad Al-Halah and Kristen Grauman. From Paris to Berlin: Discovering Fashion Style Influences Around the World. In *CVPR*, 2020.

[2] Ziad Al-Halah, Rainer Stiefelhagen, and Kristen Grauman. Fashion Forward: Forecasting Visual Style in Fashion. In *ICCV*, 2017.

[3] Logica Banica, Daniela Pirvu, and Alina Hagiu. Neural networks based forecasting for romanian clothing sector. In *Intelligent fashion forecasting systems: Models and applications*, pages 161–194. Springer, 2014.

[4] Thomas Eugen Bauknecht, Jule Schaefer, and Peter Bug. Internet fashion map: the web representation of certain fashion hot spots. Technical report, Reutlingen University, Reutlingen, Germany, 2017.

[5] Tamara L Berg, Alexander C Berg, and Jonathan Shih. Automatic attribute discovery and characterization from noisy web data. In *ECCV*, 2010.

[6] Lukas Bossard, Matthias Dantone, Christian Leistner, Christian Wengert, Till Quack, and Luc Van Gool. Apparel classification with style. In *ACCV*, 2012.

[7] George EP Box, Gwilym M Jenkins, Gregory C Reinsel, and Greta M Ljung. *Time series analysis: forecasting and control*. John Wiley & Sons, 2015.

[8] Chengyao Chen, Zhitao Wang, Wenjie Li, and Xu Sun. Modeling scientific influence for research trending topic prediction. In *AAAI*, 2018.

[9] Huizhong Chen, Andrew Gallagher, and Bernd Girod. Describing clothing by semantic attributes. In *ECCV*, 2012.

[10] Kuanting Chen, Kezhen Chen, Peizhong Cong, Winston H Hsu, and Jiebo Luo. Who are the Devils Wearing Prada in New York City? In *ACM Multimedia*, 2015.

[11] Haoye Dong, Xiaodan Liang, Xiaohui Shen, Bochao Wang, Hanjiang Lai, Jia Zhu, Zhiting Hu, and Jian Yin. Towards multi-pose guided virtual try-on network. In *ICCV*, 2019.

[12] Clive WJ Granger. Investigating causal relations by econometric models and cross-spectral methods. *Econometrica: Journal of the Econometric Society*, pages 424–438, 1969.

[13] M Hadi Kiapour, Xufeng Han, Svetlana Lazebnik, Alexander C Berg, and Tamara L Berg. Where to buy it: Matching street clothing photos in online shops. In *ICCV*, 2015.

[14] M Hadi Kiapour and Robinson Piramuthu. Brand > logo: Visual analysis of fashion brands. In *ECCV Workshops*, 2018.

[15] Xintong Han, Zuxuan Wu, Weilin Huang, Matthew R. Scott, and Larry S. Davis. Finet: Compatible and diverse fashion image inpainting. In *ICCV*, 2019.

[16] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *CVPR*, 2016.

[17] Ruining He and Julian McAuley. Ups and Downs: Modeling the Visual Evolution of Fashion Trends with One-Class Collaborative Filtering. In *WWW*, 2016.

[18] S. C. Hidayati, T. W. Goh, J. G. Chan, C. Hsu, J. See, W. Lai Kuan, K. Hua, Y. Tsao, and W. Cheng. Dress with style: Learning style from joint deep embedding of clothing styles and body shapes. *IEEE Transactions on Multimedia*, pages 1–1, 2020.

[19] Shintami Chusnul Hidayati, Cheng-Chun Hsu, Yu-Ting Chang, Kai-Lung Hua, Jianlong Fu, and Wen-Huang Cheng. What dress fits me best?: Fashion recommendation on the clothing style for personal body shape. In *ACM Multimedia*, 2018.

[20] Shintami C Hidayati, Kai-Lung Hua, Wen-Huang Cheng, and Shih-Wei Sun. What are the fashion trends in new york? In *ACM Multimedia*, 2014.

[21] Wei-Lin Hsiao and Kristen Grauman. Learning the latent look: Unsupervised discovery of a style-coherent embedding from fashion images. In *ICCV*, 2017.

[22] Wei-Lin Hsiao and Kristen Grauman. Creating capsule wardrobes from fashion images. In *CVPR*, 2018.

[23] Wei-Lin Hsiao and Kristen Grauman. Dressing for diverse body shapes. In *CVPR*, 2020.

[24] Wei-Lin Hsiao, Isay Katsman, Chao-Yuan Wu, Devi Parikh, and Kristen Grauman. Fashion++: Minimal edits for outfit improvement. *ICCV*, 2019.

[25] Changwei Hu, Piyush Rai, Changyou Chen, Matthew Harding, and Lawrence Carin. Scalable bayesian non-negative tensor factorization for massive count data. In *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*, pages 53–70. Springer, 2015.

[26] Junshi Huang, Rogerio S Feris, Qiang Chen, and Shuicheng Yan. Cross-domain image retrieval with a dual attribute-aware ranking network. In *ICCV*, 2015.

[27] Hirokatsu Kataoka, Yutaka Satoh, Kaori Abe, Munetaka Minoguchi, and Akio Nakamura. Ten-million-order human database for world-wide fashion culture analysis. In *CVPR Workshops*, 2019.

[28] Murat Kaya, Engin Yeşil, M Furkan Dodurka, and Sarven Sıradağ. Fuzzy forecast combining for apparel demand forecasting. In *Intelligent fashion forecasting systems: Models and applications*, pages 123–146. Springer, 2014.

[29] M Hadi Kiapour, Kota Yamaguchi, Alexander C Berg, and Tamara L Berg. Hipster wars: Discovering elements of fashion styles. In *ECCV*, 2014.

[30] Gunhee Kim and Eric Xing. Discovering pictorial brand associations from large-scale online image data. In *ICCV Workshops*, 2013.

[31] Zhanghui Kuang, Yiming Gao, Guanbin Li, Ping Luo, Yimin Chen, Liang Lin, and Wayne Zhang. Fashion retrieval via graph reasoning networks on a similarity pyramid. In *ICCV*, 2019.

[32] Iljung S Kwak, Ana Cristina Murillo, Peter N Belhumeur, David J Kriegman, and Serge J Belongie. From bikers to surfers: Visual recognition of urban tribes. In *BMVC*, 2013.

[33] Y. Li, L. Cao, J. Zhu, and J. Luo. Mining fashion outfit composition using an end-to-end deep learning approach on set data. *IEEE Transactions on Multimedia*, 19(8):1946–1955, 2017.

[34] X. Liang, L. Lin, W. Yang, P. Luo, J. Huang, and S. Yan. Clothes co-parsing via joint image segmentation and labeling with application to clothing retrieval. *IEEE Transactions on Multimedia*, 2016.

[35] Si Liu, Jiashi Feng, Zheng Song, Tianzhu Zhang, Hanqing Lu, Changsheng Xu, and Shuicheng Yan. Hi, magic closet, tell me what to wear! In *ACM Multimedia*, 2012.

[36] Si Liu, Zheng Song, Guangcan Liu, Changsheng Xu, Hanqing Lu, and Shuicheng Yan. Street-to-shop: Cross-scenario clothing retrieval via parts alignment and auxiliary set. In *CVPR*, 2012.

[37] Y. Liu, W. Chen, L. Liu, and M. S. Lew. Swapgan: A multi-stage generative approach for person-to-person fashion style transfer. *IEEE Transactions on Multimedia*, 21(9):2209–2222, 2019.

[38] Ziwei Liu, Ping Luo, Shi Qiu, Xiaogang Wang, and Xiaoou Tang. Deepfashion: Powering robust clothes recognition and retrieval with rich annotations. In *CVPR*, 2016.

[39] Zheng Lu and Kristen Grauman. Story-driven summarization for egocentric video. In *CVPR*, 2013.

[40] Utkarsh Mall, Kevin Matzen, Bharath Hariharan, Noah Snavely, and Kavita Bala. Geostyle: Discovering fashion trends and events. In *ICCV*, 2019.

[41] Kevin Matzen, Kavita Bala, and Noah Snavely. Streetstyle: Exploring world-wide clothing styles from millions of photos. *arXiv preprint arXiv:1706.01869*, 2017.

[42] Julian McAuley, Christopher Targett, Qinfeng Shi, and Anton Van Den Hengel. Image-based recommendations on styles and substitutes. In *ACM SIGIR Conference on Research and Development in Information Retrieval*, 2015.

[43] Rishabh Misra, Mengting Wan, and Julian McAuley. Decomposing fit semantics for product size recommendation in metric spaces. In *ACM Conference on Recommender Systems*, 2018.

[44] Ana C Murillo, Iljung S Kwak, Lubomir Bourdev, David Kriegman, and Serge Belongie. Urban tribes: Analyzing group photos from a social perspective. In *CVPR Workshops*, 2012.

[45] Shuyun Ren, Hau-Ling Chan, and Pratibha Ram. A comparative study on fashion demand forecasting models with multiple sources of uncertainty. *Annals of Operations Research*, 257(1-2):335–355, 2017.

[46] Shuyun Ren, Tsan-Ming Choi, and Na Liu. Fashion sales forecasting with a panel data-based particle-filter model. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 45(3):411–421, 2014.

[47] Dafna Shahaf and Carlos Guestrin. Connecting the dots between news articles. In *ACM SIGKDD*, 2010.

[48] Edgar Simo-Serra, Sanja Fidler, Francesc Moreno-Noguer, and Raquel Urtasun. Neuroaesthetics in Fashion: Modeling the Perception of Fashionability. In *CVPR*, 2015.

[49] Zheng Song, Meng Wang, Xian-sheng Hua, and Shuicheng Yan. Predicting occupation via human clothing and contexts. In *ICCV*, 2011.

[50] Christian Szegedy, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott Reed, Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke, and Andrew Rabinovich. Going deeper with convolutions. In *CVPR*, 2015.

[51] Sébastien Thomassey. Sales forecasting in apparel and fashion industry: A review. In *Intelligent fashion forecasting systems: Models and applications*, pages 9–27. Springer, 2014.

[52] Andreas Veit, Balazs Kovacs, Sean Bell, Julian McAuley, Kavita Bala, and Serge Belongie. Learning visual clothing style with heterogeneous dyadic co-occurrences. In *ICCV*, 2015.

[53] Sirion Vittayakorn, Kota Yamaguchi, Alexander C. Berg, and Tamara L. Berg. Runway to realway: Visual analysis of fashion. In *WACV*, 2015.

[54] Bochao Wang, Huabin Zheng, Xiaodan Liang, Yimin Chen, Liang Lin, and Meng Yang. Toward characteristic-preserving image-based virtual try-on network. In *ECCV*, 2018.

[55] Wenguan Wang, Yuanlu Xu, Jianbing Shen, and Song-Chun Zhu. Attentive fashion grammar network for fashion landmark detection and clothing category classification. In *CVPR*, 2018.

[56] Cong Yu, Yang Hu, Yan Chen, and Bing Zeng. Personalized fashion design. In *ICCV*, 2019.

[57] Weijiang Yu, Xiaodan Liang, Ke Gong, Chenhan Jiang, Nong Xiao, and Liang Lin. Layout-graph reasoning for fashion landmark detection. In *CVPR*, 2019.

[58] Zalando. The most elegant cities in the world. https://www.zalando.co.uk/worlds-most-elegant-cities/. Accessed April 7, 2020.

[59] X. Zhang, J. Jia, K. Gao, Y. Zhang, D. Zhang, J. Li, and Q. Tian. Trip outfits advisor: Location-oriented clothing recommendation. *IEEE Transactions on Multimedia*, 19(11):2533–2544, 2017.

[60] Bo Zhao, Jiashi Feng, Xiao Wu, and Shuicheng Yan. Memory-augmented attribute manipulation networks for interactive fashion search. In *CVPR*, 2017.