

# Towards Explanations for Visual Recommender Systems of Artistic Images

Vicente Dominguez  
PUC Chile  
Santiago, Chile  
vidominguez@uc.cl

Christoph Trattner  
University of Bergen  
Bergen, Norway  
christoph.trattner@uib.no

Pablo Messina  
PUC Chile  
Santiago, Chile  
pamessina@uc.cl

Denis Parra  
PUC Chile  
Santiago, Chile  
dparra@ing.puc.cl

## ABSTRACT

Explaining automatic recommendations is an active area of research since it has shown an important effect on users' acceptance over the items recommended. However, there is a lack of research in explaining content-based recommendations of images based on visual features. In this paper, we aim to fill this gap by testing three different interfaces (one baseline and two novel explanation interfaces) for artistic image recommendation. Our experiments with  $N=121$  users confirm that explanations of recommendations in the image domain are useful and increase user satisfaction, perception of explainability, relevance, and diversity. Furthermore, our experiments show that the results are also dependent on the underlying recommendation algorithm used. We tested the interfaces with two algorithms: Deep Neural Networks (DNN), with high accuracy but with difficult to explain features, and the more explainable method based on Attractiveness Visual Features (AVF). The better the accuracy performance –in our case the DNN method– the stronger the positive effect of the explainable interface. Notably, the explainable features of the AVF method increased the perception of explainability but did not increase the perception of trust, unlike DNN, which improved both dimensions. These results indicate that algorithms in conjunction with interfaces play a significant role in the perception of explainability and trust for image recommendation. We plan to further investigate the relationship between interface explainability and algorithmic performance in recommender systems.

## KEYWORDS

Recommender systems, Artwork Recommendation, Explainable Interfaces, Visual Features

### ACM Reference format:

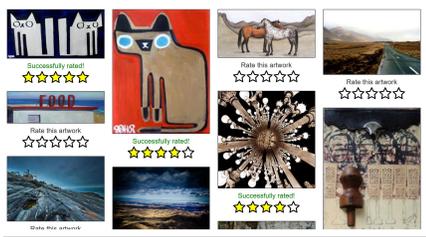
Vicente Dominguez, Pablo Messina, Christoph Trattner, and Denis Parra. 2018. Towards Explanations for Visual Recommender Systems of Artistic Images. In *Proceedings of IntRS Workshop, Vancouver, Canada, October 2018 (IntRS'18)*, 5 pages.

## 1 INTRODUCTION

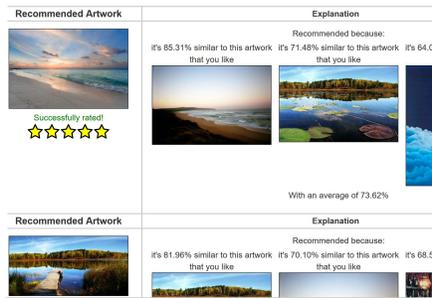
Online artwork recommendation has received little attention compared to other areas such as movies [1, 10], music [4, 16] or points-of-interest [25, 28, 29]. The first works in the area date from 2006-2007 such as the CHIP [2] project, which implemented traditional techniques such as content-based and collaborative filtering for artwork recommendation at the Rijksmuseum, and the *m4art* system by Van den Broek et al. [26], which used histograms of color to retrieve similar artworks where the input query was a painting image. More recently, deep neural networks (DNN) have been used for artwork recommendation and are the current state-of-the-art model [7, 12], which is rather expected considering that DNNs are the top performing models for obtaining visual features for several tasks, such as image classification [15], and scene identification [23]. However, no user study has been conducted to validate the performance of DNNs versus other visual features. This aspect is important since past works have shown that off-line results might not always replicate when tested with actual users [14, 17]. Moreover, we provide evidence of the important value of explanations in artwork recommender systems over several dimensions of user perception. Visual features obtained from DNNs are still difficult to explain to users, despite current efforts to understand them and explain them [20]. In contrast, features of visual attractiveness could be easily explained, based on color, brightness or contrast [21]. Explanations in recommender systems have been shown to have a significant effect on user satisfaction [24], and, to the best of our knowledge, no previous work has shown how to explain recommendations of images based on visual features. Hence, there is no study of the effect on users when explaining images recommended by a Visual Content-based Recommender (Hereinafter, VCBR).

*Objective.* In this paper, we research the effect of explaining artistic image suggestions. In particular, we conduct a user study on Amazon Mechanical Turk under three different interfaces and two different algorithms. The three interfaces are: i) no explanations, ii) explanations based on similar images, and iii) explanations based on visual features. Moreover, the two algorithms are: Deep Neural Networks (DNN) and Attractiveness Visual Features (AVF). In our study, we used images provided by the online store *UGallery* (<http://www.UGallery.com/>).

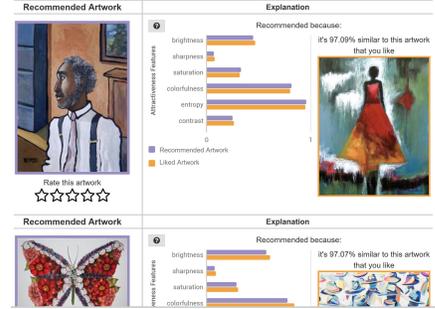
*Research Questions* To drive our research, the following two questions were defined:



**Figure 1: Interface 1: Baseline recommendation interface without explanations.**



**Figure 2: Interface 2: Explainable recommendation interface with textual explanations and top-3 similar images.**



**Figure 3: Interface 3: Explainable recommendation interface with features' bar chart and top-1 similar image.**

- **RQ1.** Given three different types of interfaces, one baseline interface without explanations and two with them, employing similar image explanations and a feature bar chart, which one is perceived as most useful?
- **RQ2.** Furthermore, based on the visual and content-based recommender algorithm chosen, are there observable differences in how the three interfaces are perceived?

## 2 RELATED WORK

Relevant related research is collated in two sub-sections: First, we review research on recommending artistic images to people. Second we summarize studies on explaining recommender systems. Both are important to our problem at hand. The final paragraph in this section highlights the differences to previous work and our contributions to the existing literature in the area.

**Recommendations of Artistic Images.** The works of Aroyo et al. [2] with the CHIP project and Semeraro et al. [22] with FIRST (Folksonomy-based Item Recommender syStem) made early contributions to this area using traditional techniques. More complex methods were implemented recently by Benouaret et al. [3], using context obtained through a mobile application, that makes a museum tour recommendation. Finally, the work of He et al. addresses digital artwork recommendations based on pre-trained deep neural visual features [12], and the work of Dominguez et al. [7] and Messina et al. [18] compared neural against traditional visual features. None of the aforementioned works performed a user study under explanation interfaces to generalize their results.

**Explaining Recommender Systems.** There are some related works on explanations for recommender systems [24]. Though a good amount of research has been published in the area, to the best of our knowledge, no previous research has conducted a user study to understand the effect of explaining recommendation of artwork images based on different visual features. The closest works in this aspect are researches oriented to automatically add caption to images [9, 19] or to explain image classifications [13], but they are not directly related to personalized recommender systems.

**Differences to Previous Research & Contributions.** Although we focus on artistic images, to the best of our knowledge this is the first work which studies the effect of explaining recommendations of images based on visual features. Our contributions are two-fold: i) we analyze and report the positive effect of explaining artistic recommendations especially for the VCBR based on neural

features, and ii) by a user study we validate off-line results stating the superiority of neural visual features compared to attractiveness visual features over several dimensions, such as users' perception of explainability, relevance, trust and general satisfaction.

## 3 METHODS

In the following section we describe in detail our study methods. First, we introduce the dataset chosen for the purpose of our study. Second we introduce the three different explainable visual interfaces implemented which we evaluate. Third the two algorithms chosen for our study are revealed. Finally, the user study procedure is explained.

### 3.1 Materials

For the purpose of our study we rely on a dataset provided by the online web store *UGallery*, which has been selling artwork for more than 10 years [27]. They support emergent artists by helping them sell their artwork online. For our research, *UGallery* provided us with an anonymized dataset of 1,371 users, 3,490 items and 2,846 purchases (transactions) of artistic artifacts, where all users have made at least one transaction. On average, each user bought 2-3 items over recent years .

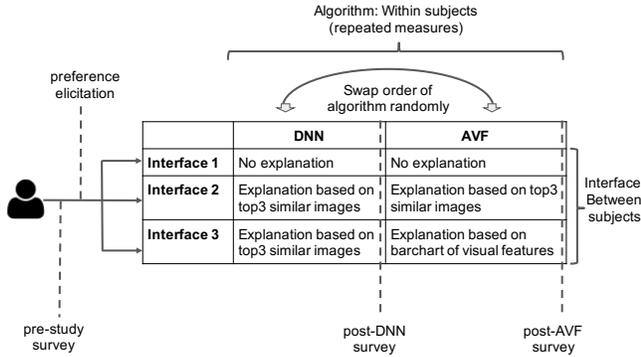
### 3.2 The Explainable Recommender Interfaces

In our study we explore the effect of explanations in visual content-based artwork recommender systems. As such, our study contains conditions depending on how recommendations are displayed: i) no explanations, as shown in Figure 1, ii) explanations given by text and based on the top-3 most similar images a user liked in the past, as shown in Figure 2, and iii) explanations employing a visual attractiveness bar chart and showing the most similar image of the user's item profile, as presented in Figure 3.

In all three cases the interfaces are vertically scrollable. While Interface 1 (baseline) is able to show 5 images in a row at the same time, interfaces 2 and 3 are capable of showing one recommended image at the same time in one row to the user.

### 3.3 Visual Recommendation Approaches

As mentioned earlier in this paper, we make use of two different content-based visual recommender approaches in our work. The reason for choosing content-based methods over collaborative filtering-based methods is grounded in the fact that once an item is sold via the *UGallery* store, it is not available anymore (every item



**Figure 4: Study procedure.** After the pre-study survey and the preference elicitation, users were assigned to one of three possible interfaces. In each interface they evaluated recommendations of two algorithms: DNN and AVF.

is unique) and hence traditional collaborative filtering approaches do not apply.

**DNN Visual Feature (DNN) Algorithm.** The first algorithmic approach we employed was based on image similarity, itself based on features extracted with a deep neural network. The output vector representing the image is usually called an image’s visual embedding. The visual embedding in our experiment was a vector of features obtained from an AlexNet, a convolutional deep neural network developed to classify images [15]. In particular, we use an AlexNet model pre-trained with the ImageNet dataset [6]. Using the pre-trained weights, for every image a vector of 4,096 dimensions was generated with the Caffe (<http://caffe.berkeleyvision.org/>) framework. We resized every image to a 227x227 image. This is the standard pre-processing needed to use the AlexNet.

**Attractiveness Visual Features (AVF) Algorithm.** The second content-based algorithmic recommender approach employed was a method based on visual attractiveness features. San Pedro and Siersdorfer in [21] proposed several explainable visual features that to a great extent, can capture the attractiveness of an image posted on Flickr. Following their procedure, for every image in our *UGallery* dataset we calculated: (a) average brightness, (b) saturation, (c) sharpness, (d) RMS-contrast, (e) colorfulness and (f) naturalness. In addition, we added (g) entropy, which is a good way to characterize and measure the texture of an image [11]. These metrics have also been used in another study [8], where we show how to nudge people with attractive images to take up more healthy recipe recommendations. To compute these features, we used the original size of the images and did not pre-process them.

Due space constrains, the details to calculate the features are described in the article by Messina et al. [18]

**Computing Recommendations.** We maximize the utility score that an item provides to a user. Given a user  $u$  who has consumed a set of artworks  $P_u$ , a constrained profile size  $K$ , and an arbitrary artwork  $i$  from the inventory, the score of this item  $i$  to be recommended to  $u$  is:

$$\text{score}(u, i)_X = \frac{\sum_{r=1}^{\min\{K, |P_u|\}} \max_{j \in P_u} \{ \text{sim}(V_i^X, V_j^X) \}}{\min\{K, |P_u|\}}, \quad (1)$$

**Table 1: Evaluation dimensions and statements asked in the post-study survey.** Users indicated their agreement with the statement on a scale from 0 to 100 (= totally agree).

| Dimension              | Statement   |
|------------------------|---|
| Explainable            | I understood why the art images were recommended to me.                         |
| Relevance              | The art images recommended matched my interests.                                |
| Diverse                | The art images recommended were diverse.  |
| Interface Satisfaction | Overall, I am satisfied with the recommender interface.                         |
| Use Again              | I would use this recommender system again for finding art images in the future. |
| Trust                  | I trusted the recommendations made.   |

where  $V_z^X$  is a feature vector of item  $z$  obtained with method  $X$ , where  $X$  can be either a pre-trained AlexNet (DNN) or attractiveness visual features (AVF).  $\max^{(r)}$  denotes the  $r$ -th maximum value, e.g., if  $r = 1$  it is the overall maximum, if  $r = 2$  it is the second maximum, and so on. We compute the average similarity of the top- $K$  most similar images because as shown in Messina et al. [18], for different users, the recommendations match better using smaller subsets of the entire user profile. Users do not always look to buy a painting similar to one they bought before, but they look for one that resembles a set of artworks that they liked.  $\text{sim}(V_i, V_j)$  denotes a similarity function between vectors  $V_i$  and  $V_j$ . In this particular case, the similarity function used was cosine similarity, expressed as:

$$\text{sim}(V_i, V_j) = \cos(V_i, V_j) = \frac{V_i \cdot V_j}{\|V_i\| \|V_j\|} \quad (2)$$

Both methods use the same formula to calculate the recommendations. The difference is in the origin of the visual features. For the DNN method, the features were extracted with the AlexNet [15], and in the case of AVF, the features were extracted based on San Pedro et al. [21].

### 3.4 User Study Procedure

To evaluate the performance of our explainable interfaces we conducted a user study in Amazon Mechanical Turk using a 3x2 mixed design: 3 interfaces (between-subjects) and 2 algorithms (within-subjects, DNN and AVF). The interface conditions were: *Interface 1*: interface without explanations, as in Figure 1; *Interface 2*: each item recommendation is explained based on the top 3 most similar images in the user profile, as in Figure 2; and *Interface 3*: only for AVF, based on a bar chart of visual features, as in Figure 3. Notice that in the condition Interface 3, for DNN we used the explanation based on top 3 most similar images, because the neural embedding of 4,096 dimensions has no *human-interpretable* features to show in a bar chart.

To compute the recommendations for each of the three interface conditions two recommender algorithms were chosen: one based on DNN visual features, and the other based on attractiveness visual features (AVF). The order in which the algorithms were presented was chosen at random to diminish the chance of a learning effect.

The full study procedure is shown in Figure 4. Participants accepted the study on Mechanical Turk (<https://www.mturk.com>) and were redirected to a web application. After accepting a consent form, they are redirected to the pre-study survey, which collects

**Table 2: Results of users' perception over several evaluation dimensions, defined in Table 1 . Scale 1-100 (higher is better), except for Average rating (scale 1-5). DNN: Deep Neural Network, and AVF: Attractiveness visual features. The symbol  $\uparrow^1$  indicates interface-wise significant difference (differences between interfaces using the same algorithms). The \* symbol denotes algorithm-wise statistical difference (comparing a dimension between algorithms, using the same interface).**

| Condition  | Evaluation Dimensions |                   |                   |      |                   |       |                        |      |           |      |       |      |                |      |
|--|-----------------------|-------------------|-------------------|------|-------------------|-------|------------------------|------|-----------|------|-------|------|----------------|------|
|  | Explainable           |                   | Relevance         |      | Diverse           |       | Interface Satisfaction |      | Use Again |      | Trust |      | Average Rating |      |
|  | DNN                   | AVF               | DNN               | AVF  | DNN               | AVF   | DNN                    | AVF  | DNN       | AVF  | DNN   | AVF  | DNN            | AVF  |
| <b>Interface 1</b><br>(No Explanations)                            | 66.2*                 | 51.4              | 69.0*             | 53.6 | 46.1              | 69.4* | 69.9                   | 62.1 | 65.8      | 59.7 | 69.3  | 63.7 | 3.55*          | 3.23 |
| <b>Interface 2</b><br>(DNN & AVF: Top-3 similar images)            | 83.5 $\uparrow^1$     | 74.0 $\uparrow^1$ | 80.0*             | 61.7 | 58.8              | 69.9* | 76.6*                  | 61.7 | 76.1*     | 65.9 | 75.9* | 62.7 | 3.67*          | 3.00 |
| <b>Interface 3</b><br>(DNN: Top-3 similar, AVF: feature bar chart) | 84.2 $\uparrow^1$     | 70.4 $\uparrow^1$ | 82.3 $\uparrow^1$ | 56.2 | 65.3 $\uparrow^1$ | 71.2  | 69.9*                  | 63.3 | 78.2*     | 58.7 | 77.7* | 55.4 | 3.90*          | 2.99 |

Stat. significance between interfaces by multiple t-tests, Bonferroni corr.  $\alpha_{bonf} = \alpha/n = 0.05/3 = 0.0017$ . Stat. significance between algorithms using pairwise t-test,  $\alpha = 0.05$ .

demographic data (age, gender) and a subject's previous knowledge of art based on the test by Chatterjee et al. [5].

Following this, they had to perform a preference elicitation task. In this step, the users had to "like" at least ten paintings, using a Pinterest-like interface. Next, they were randomly assigned to one interface condition. In each condition, they again provided feedback (rating with 1-5 scale to each image) to top ten recommendations of images with employing either the DNN or the AVF algorithm (also assigned at random as discussed before). Finally, the participants were asked to next answer a post-algorithm survey. The dimensions evaluated in the post-algorithm survey are the same for DNN and AVF algorithms, and they are shown in Table 1. This process is repeated for the second algorithm as well. Once the participants finished answering the second post study survey, they were redirected to the final view, where they received a survey code for later payment in Amazon Mechanical Turk.

## 4 RESULTS

The study was finished by in total 200 users out of which 121 were able to answer our validation questions successfully and hence were included in the results. In total, we had two validation questions set to check for attention of our study participants. Filtering out users not responding properly to these questions allowed us to include 41 users for the Interface 1 condition, 41 users for Interface 2 condition and 39 users for Interface 3 condition. In total, participants were paid an amount of 0.40 USD per study, which took them around 10 minutes to complete.

Our subjects reported to be between 18 to over 60 years old. Most of them were between 25 to 32 years old (36%) or 32 to 40 years old (29%). Females made up 55.4% . 12% just finished high school, 31% had a some college degree, 57% had a bachelor's, master's or Ph.D. degree. Only 8% reported some visual impairment. W.r.t. the subject understanding about art, 20% had null experience, 48% had attended 1 or 2 lessons, and 32% reported to have attended 3 or more at the high school level or above. 20% of our subjects also reported that they have almost never visited a museum or an art gallery; 36% do this once a year; and 44% do this once every 1 or 6 months.

**Differences between Interfaces.** Table 2 summarizes the results of the user study. First we compared interface performance and then we looked at the algorithmic performance. The explainable interfaces (Interface 2 and 3) significantly improved the perception of explainability compared to Interface 1 under both algorithms.

There is also a significant improvement over Interface 1 in terms of relevance and diversity, but this is only achieved by the DNN method when this is compared against the AVF method using the interface 3. Interestingly, this is the condition where the interface is more transparent, since it explains exactly what is used to recommend (brightness, saturation, sharpness, etc.). People report that they understand why the images are recommended (70.4), but since the relevance is rather insufficient (56.2), the perception of trust is reported as low (55.4).

**Differences in Algorithms.** With the only exception of the dimension *Diverse* where AVF was significantly better, DNN was perceived more positively than AVF at large. In interfaces 2 and 3, the DNN method was perceived significantly better in 5 dimensions (explainability, relevance, interface satisfaction, interest for eventual use, and trust), as well as higher average rating.

Overall, the results indicate that the explainable interface based on top 3 similar images works better than an interface without explanation. Moreover, this effect is enhanced by the accuracy of the algorithm, so even if the algorithm has no explainable features (DNN) it could induce more trust if the user perceives a larger predictive preference accuracy.

## 5 CONCLUSIONS & FUTURE WORK

In this paper, we have studied the effect of explaining recommendation of images employing three different recommender interfaces, as well as interactions with two different visual content-based recommendation algorithms: one with high predictive accuracy but with unexplainable features (DNN), and another with lower accuracy but with higher potential for explainable features (AVF).

The first result, which answers RQ1, shows that explaining the images recommended has a positive effect vs. no explanation. Moreover, the explanation based on top 3 similar images presents the best results, but we need to consider that the alternative method, explanations based on visual features, was only used with the AVF. This result is preliminary and opens a path of research in terms of new interfaces which could help to explain the features learned by a deep neural network of images.

Regarding RQ2, we see that the algorithm used plays an important role in conjunction with the interface. DNN is perceived better than AVF in most dimensions evaluated, showing that further research should focus on the interaction between algorithm and explainable interfaces. In the future we will expand this work to other datasets, beyond artistic images, to generalize our results.

## REFERENCES

- [1] Xavier Amatriain. 2013. Mining large streams of user data for personalized recommendations. *ACM SIGKDD Explorations Newsletter* 14, 2 (2013), 37–48.
- [2] LM Aroyo, Y Wang, R Brussee, Peter Gorgels, LW Rutledge, and N Stash. 2007. Personalized museum experience: The Rijksmuseum use case. In *Proceedings of Museums and the Web*.
- [3] Idir Benouaret and Dominique Lenne. 2015. Personalizing the Museum Experience through Context-Aware Recommendations. In *Systems, Man, and Cybernetics (SMC), 2015 IEEE International Conference on*. IEEE, 743–748.
- [4] Oscar Celma. 2010. Music recommendation. In *Music Recommendation and Discovery*. Springer, 43–85.
- [5] Anjan Chatterjee, Page Widick, Rebecca Sternschein, William Smith II, and Bianca Bromberger. 2010. The Assessment of Art Attributes. 28 (07 2010), 207–222.
- [6] Jia Deng, Wei Dong, Richard Socher, Li-Jia Li, Kai Li, and Li Fei-Fei. 2009. Imagenet: A large-scale hierarchical image database. In *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on*. IEEE, 248–255.
- [7] Vicente Dominguez, Pablo Messina, Denis Parra, Domingo Mery, Christoph Trattner, and Alvaro Soto. 2017. Comparing Neural and Attractiveness-based Visual Features for Artwork Recommendation. In *Proceedings of the Workshop on Deep Learning for Recommender Systems, co-located at RecSys 2017*. DOI: <http://dx.doi.org/10.1145/3125486.3125495> arXiv:arXiv:1706.07515
- [8] David Elsweller, Christoph Trattner, and Morgan Harvey. 2017. Exploiting food choice biases for healthier recipe recommendation. In *Proceedings of the 40th international acm sigir conference on research and development in information retrieval*. ACM, 575–584.
- [9] Hao Fang, Saurabh Gupta, Forrest Iandola, Rupesh Srivastava, Li Deng, Piotr Dollár, Jianfeng Gao, Xiaodong He, Margaret Mitchell, John Platt, and others. 2015. From captions to visual concepts and back. (2015).
- [10] Carlos A Gomez-Urbe and Neil Hunt. 2016. The netflix recommender system: Algorithms, business value, and innovation. *ACM Transactions on Management Information Systems (TMIS)* 6, 4 (2016), 13.
- [11] Rafael C Gonzalez, Steven L Eddins, and Richard E Woods. 2004. *Digital Image Publishing Using MATLAB*. Prentice Hall.
- [12] Ruining He, Chen Fang, Zhaowen Wang, and Julian McAuley. 2016. Vista: A Visually, Socially, and Temporally-aware Model for Artistic Recommendation. In *Proceedings of the 10th ACM Conference on Recommender Systems (RecSys '16)*. ACM, New York, NY, USA, 309–316. DOI: <http://dx.doi.org/10.1145/2959100.2959152>
- [13] Lisa Anne Hendricks, Zeynep Akata, Marcus Rohrbach, Jeff Donahue, Bernt Schiele, and Trevor Darrell. 2016. Generating visual explanations. In *European Conference on Computer Vision*. Springer, 3–19.
- [14] Joseph A Konstan and John Riedl. 2012. Recommender systems: from algorithms to user experience. *User Modeling and User-Adapted Interaction* 22, 1-2 (2012), 101–123.
- [15] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. 2012. Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems*. 1097–1105.
- [16] Pattie Maes and others. 1994. Agents that reduce work and information overload. *Commun. ACM* 37, 7 (1994), 30–40.
- [17] Sean M McNee, Nishikant Kapoor, and Joseph A Konstan. 2006. Don't look stupid: avoiding pitfalls when recommending research papers. In *Proceedings of the 2006 20th anniversary conference on Computer supported cooperative work*. ACM, 171–180.
- [18] Pablo Messina, Vicente Dominguez, Denis Parra, Christoph Trattner, and Alvaro Soto. 2018. Content-Based Artwork Recommendation: Integrating Painting Metadata with Neural and Manually-Engineered Visual Features. *User Modeling and User-Adapted Interaction* (2018). DOI: <http://dx.doi.org/10.1007/s11257-018-9206-9>
- [19] Margaret Mitchell, Xufeng Han, Jesse Dodge, Alyssa Mensch, Amit Goyal, Alex Berg, Kota Yamaguchi, Tamara Berg, Karl Stratos, and Hal Daumé, III. 2012. Midge: Generating Image Descriptions from Computer Vision Detections. In *Proceedings of the 13th Conference of the European Chapter of the Association for Computational Linguistics (EACL '12)*. Association for Computational Linguistics, Stroudsburg, PA, USA, 747–756. <http://dl.acm.org/citation.cfm?id=2380816.2380907>
- [20] Chris Olah, Alexander Mordvintsev, and Ludwig Schubert. 2017. Feature Visualization. *Distill* (2017). DOI: <http://dx.doi.org/10.23915/distill.00007> <https://distill.pub/2017/feature-visualization>.
- [21] Jose San Pedro and Stefan Siersdorfer. 2009. Ranking and Classifying Attractiveness of Photos in Folksonomies. In *Proceedings of the 18th International Conference on World Wide Web (WWW '09)*. ACM, New York, NY, USA, 771–780. DOI: <http://dx.doi.org/10.1145/1526709.1526813>
- [22] Giovanni Semeraro, Pasquale Lops, Marco De Gemmis, Cataldo Musto, and Fedelucio Narducci. 2012. A folksonomy-based recommender system for personalized access to digital artworks. *Journal on Computing and Cultural Heritage (JOCCH)* 5, 3 (2012), 11.
- [23] Ali Sharif Razavian, Hossein Azizpour, Josephine Sullivan, and Stefan Carlsson. 2014. CNN features off-the-shelf: an astounding baseline for recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*. 806–813.
- [24] Nava Tintarev and Judith Masthoff. 2015. Explaining recommendations: Design and evaluation. In *Recommender Systems Handbook*. Springer, 353–382.
- [25] Christoph Trattner, Alexander Oberegger, Lukas Eberhard, Denis Parra, Leandro Marinho, and others. 2016. Understanding the Impact of Weather for POI Recommendations. *Proceedings of RecTour Workshop, co-located at ACM RecSys* (2016).
- [26] Egon L van den Broek, Thijs Kok, Theo E Schouten, and Eduard Hoekamp. 2006. Multimedia for art retrieval (m4art). In *Multimedia Content Analysis, Management, and Retrieval 2006*, Vol. 6073. International Society for Optics and Photonics, 60730Z.
- [27] Deborah Weinswig. 2016. Art Market Cooling, But Online Sales Booming. <https://www.forbes.com/sites/deborahweinswig/2016/05/13/art-market-cooling-but-online-sales-booming/>. (2016). [Online; accessed 21-March-2017].
- [28] Mao Ye, Peifeng Yin, Wang-Chien Lee, and Dik-Lun Lee. 2011. Exploiting geographical influence for collaborative point-of-interest recommendation. In *Proceedings of the 34th international ACM SIGIR conference on Research and development in Information Retrieval*. ACM, 325–334.
- [29] Quan Yuan, Gao Cong, Zongyang Ma, Aixin Sun, and Nadia Magnenat Thalmann. 2013. Time-aware point-of-interest recommendation. In *Proceedings of the 36th international ACM SIGIR conference on Research and development in information retrieval*. ACM, 363–372.