IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 31 January 2024, accepted 3 February 2024, date of publication 8 February 2024, date of current version 14 February 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3363459

RESEARCH ARTICLE

Furniture Recommendations Based on User Propensity and Furniture Style Compatibility

MASAYOSHI TAKEDA[1], (Graduate Student Member, IEEE), KEIKO ONO[2], (Member, IEEE), AND AYUMU TAISHO[1]

[1]Master's Program in Information and Computer Science, Doshisha University, Kyoto 610-0394, Japan
[2]Department of Intelligent Information Engineering and Sciences, Doshisha University, Kyoto 610-0394, Japan

Corresponding author: Masayoshi Takeda (takeda.masayoshi@mikilab.doshisha.ac.jp)

This work was supported by JSPS KAKENHI under Grant 21K12097.

ABSTRACT As digital information becomes more voluminous and e-commerce becomes more widespread, there is a growing demand for item recommendations that match the users' sensibilities. However, learning users' propensities is a difficult problem, especially in the field of furniture, which requires the consideration of many factors, such as color and shape. In addition, pieces of furniture should not be recommended only as stand-alone items, but must also be considered in terms of their affinity with other pieces, making the compatibility of styles among them an important factor. However, a consumer's furniture style is an ambiguous concept that is difficult to define. To reduce this ambiguity, Siamese networks are often used to estimate style compatibility by adding various features that represent styles, but even when they make use of alternative features associated with images, they are difficult to represent accurately. This paper proposes a method for recommending multiple pieces of furniture by learning style compatibility properties with a high degree of accuracy, taking users' preferences and styles' compatibility into account. To this end, we engaged in two tasks: (1) extracting users' preferences and (2) improving the accuracy of style suitability estimation. For (1), we applied matrix factorization to identify users whose sensitivities were close to those of the users who will receive recommendations. For (2), we used the Siamese network we have already proposed, which can learn from multiple furniture images simultaneously. Specifically, we propose a one-to-many input ratio to maintain high performance even when the input is ambiguous. Validation experiments were conducted for each task, and the results showed that the performance was improved; the actual recommendation results also showed a high performance.

INDEX TERMS Furniture recommendation, matrix factorization, Siamese network, user propensity.

I. INTRODUCTION

The demand for item recommendation systems continues to increase due to the growing volume of digital data and the spread of e-commerce [1], [2]. Item recommendation systems must be able to identify user-preferred items from a large database, and collaborative filtering technologies are commonly used today [3], [4], [5]. However, learning users' preferences remains a challenging problem. Especially in the field of interior design, it is not easy to understand these preferences, because a variety of data, such as color, shape, size, material, and texture, must be considered in a composite manner [6], [7]. In addition, furniture needs to consider not only users' preferences, but also affinity with other pieces of furniture. In this case, one of the factors to consider is the compatibility among styles of furniture.

Singular value decomposition (SVD) and matrix factorization (MF) are also frequently studied as item recommendation methods that take into consideration users' preferences. These are based on collaborative filtering because they aim to recommend items preferred by users and require dimensionality compression of increasingly large amounts of data.

The associate editor coordinating the review of this manuscript and approving it for publication was Rongbo Zhu.

In particular, MF tends to achieve higher recommendation accuracy than collaborative filtering and SVD because it is a learning-efficient method that refers only to data evaluated by a large number of users.

Further, research focusing on furniture style compatibility uses the embedding of furniture images in a Euclidean space to classify the complex styles of each piece based on the outputs of multiple networks [8], [9]. In this field, the Siamese network, a type of deep metric learning, optimizes the Euclidean distance between the embedding of each furniture image, keeping furniture of similar styles close and those of dissimilar styles further away in Euclidean space. In this way, the Siamese network can not only estimate the similarity, but even their degree of similarity.

Due to the wide variety of furniture styles, their perceptions can vary greatly among individuals, which makes them difficult to define. Styles can also be considered ambiguous concepts, making it challenging to estimate style using a Siamese network. For example, some people may perceive a chair as modern, while others may perceive it as traditional. Therefore, style ambiguity needs to be mitigated to accurately estimate style compatibility among furniture pieces. Addressing this problem, Aggarwal et al. improved the compatibility among furniture styles by combining classification loss when a Siamese network is trained [8]. Weiss et al. also more accurately represented furniture styles by assigning multiple possible applicable style labels to specific pieces [10]. These existing studies have addressed style ambiguity by using complementary information when training the Siamese network; however, it is not easy to properly represent furniture styles with complementary information. With this in mind, we improved the ambiguous conventional Siamese network to create a model that evaluates each furniture piece using three or more images [11]. The reason why a multi-layered Siamese network, which takes multiple images as input in a one-to-many manner, is superior to conventional Siamese networks is that it can learn pieces' stylistic features more accurately. The multiple furniture images that constitute the "many" are all of the same style; therefore, during the inference phase, the multi-layered Siamese network is capable of extracting multiple pieces of furniture that have a higher degree of style compatibility and affinity with the "one" furniture image.

By contrast, the conventional Siamese model only recommends images based on their similarity, without considering users' preferences. In this study, we propose a recommendation method that takes into account users' preferences by combining the multi-layered Siamese model with MF, which can extract user preferences from a browsing history. This method enables furniture recommendations that consider users' preferences and additionally recommends multiple pieces of furniture that are compatible with that furniture. To verify the effectiveness of the proposed method, we trained the user's furniture ratings on the MF and compared the furniture recommendations with the conventional method using the trained MF and the Siamese model. Our results showed that MF training tended to improve the accuracy of individual sensitivity extraction, and the proposed method was able to recommend multiple pieces of compatible furniture compared to the conventional method. Note that users' recommendations are often evaluated using benchmark datasets, such as MovieLens; however, in this study, we estimate user preferences using actual e-commerce data from the Rakuten Ichiba dataset [12] and evaluate the performance of furniture images taken from the Born Furniture Styles Dataset [8]. The main contributions of this papers are as follows:

(1) The proposed method can extract users' preferences by applying MF taught from the ratings data in the Rakuten Ichiba dataset, which comprises actual e-commerce data.
(2) The proposed method, which uses a multi-layered Siamese network with increased input and MF to extract user preferences, is capable of recommending multiple furniture images that take into account users' preferences and style compatibility between pieces of furniture.

II. RELATED STUDIES

A. ITEM RECOMMENDATIONS CONSIDERING USERS' PREFERENCES

In item recommendation, it is necessary to recommend items that match users' preferences. Existing studies in this field rely on collaborative filtering and MF inference results applied on actual ratings data [14]. In recent years, MF that is especially skilled in handling a huge amount of evaluation data has often been used, and, in general, recommendations are made based on the inference results evaluated highly by the trained MF [13], [15]. In another MF-based method, He et al. improved performance by weighting items not rated by users according to their popularity, to account for item popularity bias [16]. Liang et al. improved performance by implementing a joint factorization between the user who evaluated the item and the item in question, rather than the conventional factorization between the user and the item, respectively [17].

In these studies, we found out that MF improves the accuracy of recommendations by considering the user's feelings. By contrast, there are other methods, apart from recommending items based on the inferred results of trained MF, that identify users with similar sentiments and recommend items that have been rated highly by those users. However, validating the effectiveness of such methods is not easy. Therefore, we aimed to recommend items that are rated highly by similar users, taking into account individual feelings, and conducted evaluation experiments to verify the efficacy of this approach.

B. COMPATIBILITY OF FURNITURE STYLES

Furniture style is an ambiguous concept, with varying degrees of style depending on the piece or pieces in question, such as a "slightly modern" or "very modern" chair. Thus, it is necessary to go beyond simple style-based classifications to quantitatively evaluate style compatibility among furniture pieces. Siamese networks are often used to achieve this goal. In this field, the ambiguity of the furniture does not facilitate accurate learning, so it needs to be mitigated. Therefore, we propose a method to train a Siamese network by adding additional information that represents the furniture style, instead of using only conventional image features.

In their study, Aggarwal et al. considered classification loss when training Siamese networks [8]. Specifically, they added a softmax layer to the subnetwork of a Siamese network to simultaneously learn image features and the classification loss associated with the class classification results. In so doing, they succeeded in improving the accuracy of evaluating the suitability of furniture styles. Weiss et al. gave multiple possible style labels to the furniture, which elucidated more ambiguous furniture styles [10]. Specifically, they asked 10 interior designers to assign a style to each piece of furniture to ensure accurate learning. In other works, Bell et al. proposed a method for learning the similarity between symbolic furniture images on a white background and actual room images, while Li et al. proposed learning the joint embedding of images and 3D models [18], [19].

In these studies, the features used when training the Siamese network were devised, so that the structure of the network was the conventional basic structure of comparing two images. However, even with the most creative use of features, the accurate representation of styles is difficult, and if even one of the two furniture images contains ambiguous images, accurate learning is not easy. Aware of these challenges, we proposed a multi-layered Siamese network that learns multiple furniture images in a one-to-many relationship [11]. Specifically, the model structure uses multiple furniture images in a one-to-many ratio wherein the furniture images that make up the "many" are all of the same style, which allows the model to learn style features more accurately. Furthermore, learning in a one-to-many relationship mitigates style ambiguity, bringing more conformable furniture closer together in the feature space and maintaining high accuracy even when ambiguous images are mixed in.

III. RELATED TECHNOLOGY

A. MATRIX FACTORIZATION

The implemented MF evaluates users' preferences with the following formula:

$$r_{ij} = v_i \cdot v_j + b_i + b_j \tag{1}$$

In equation (1), r_{ij} represents the evaluation value of user i for item j. v_i and v_j represent the feature vectors for the user and item, respectively, so $v_i \cdot v_j$ represents the dot product of the feature vectors for user i and item j. b_i and b_j represent the biases for the evaluation value of user i and item j, respectively, indicating the tendency of the user to assign high or low evaluation values and the tendency of the item to receive high or low evaluation values. Therefore, this MF considers not only the dot product of the feature vectors for users and items but also the biases that occur for each user and item, thus learning users' preferences and predicting recommended items.

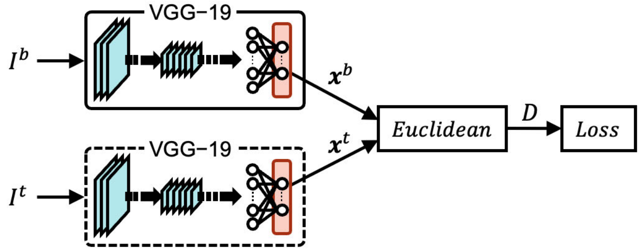

FIGURE 1. Structure of a conventional siamese model.

B. SIAMESE NETWORK

While Siamese networks are effective for simple image classification and clustering, their ability to measure quantitative similarities allows them to be used for estimating the compatibility of furniture styles [21], [22], [23], [24]. The configuration of the conventional Siamese model used in this study is shown in Fig.1. Since this study deals with images, the VGG-19 [25], a CNN that excels in image classification, was adopted as a subnet. Note that VGG-19 was pretrained on the ImageNet dataset [26]. This study assumes the use of datasets with multiple styles (e.g., modern, Asian) and categories (e.g., chair, table) and uses the Siamese network to learn and estimate the compatibility of styles. In the conventional Siamese model, two images are defined as the base image I^b and the target image I^t, and (I^b, I^t) are input to a VGG-19 network with shared weights. In this study, to estimate styles' compatibility regardless of furniture category, the furniture category C of the input (I^b, I^t) is defined as $C(I^b) \neq C(I^t)$. The 512-dimensional outputs $(\boldsymbol{x}^b, \boldsymbol{x}^t)$ from the VGG-19 network are embedded into the feature space, and the Euclidean distance D between $(\boldsymbol{x}^b, \boldsymbol{x}^t)$ is calculated. The Euclidean distance D is defined as follows:

$$D(I^b, I^t) = ||x^b - x^t||_2 \tag{2}$$

To achieve the optimal D, the conventional Siamese model uses the Contrastive Loss as the loss function.

IV. PROPOSED ITEM RECOMMENDATION METHODOLOGY

In this study, we propose a furniture recommendation method that takes into account users' preferences and the style compatibility between furniture items, as shown in Fig.2. To consider users' preferences, we used MF to learn preference information from each user's rating history. Additionally, to address styles' compatibility, we employ a multi-layered Siamese network, which increases the number of recommended furniture images and allows for the consideration of overall style compatibility. We apply MF learning and validation using the following procedure.

Step 1 Learn MF using the evaluation dataset (hereafter referred to as "original MF").

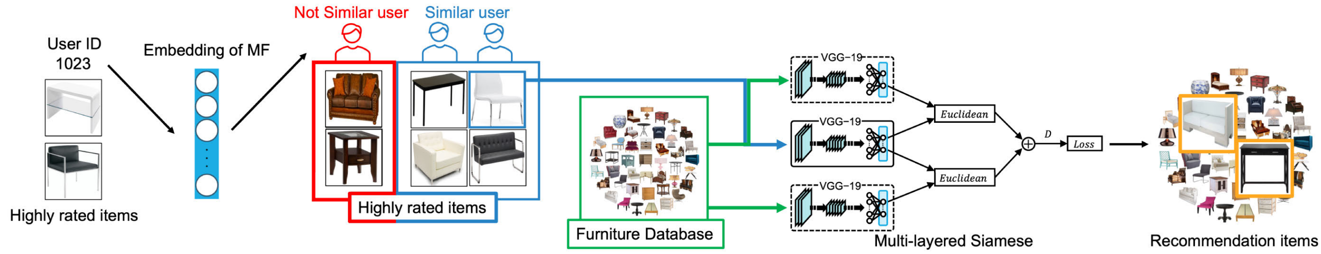

FIGURE 2. **Proposed furniture recommendation method.**

Step 2 Select one user randomly and extract similar users based on the original MF.
Step 3 Remove one piece of evaluation data from the training evaluation dataset and create a missing-value dataset.
Step 4 Learn MF using the missing-value dataset (hereafter referred to as "missing MF").
Step 5 Infer the rating of the removed evaluation data using the missing MF.
Step 6 Repeat Steps 2-5 50 times and investigate the differences between the actual ratings of similar users and the inferred rating of the missing MF.

The details and results of the experiment are described in the subsequent section.

Next, we provide a detailed explanation of the proposed recommendation methodology. First, the proposed method uses MF to learn evaluation data for furniture and to extract users who have similar sensibilities to the user receiving recommendations. Specifically, the embedding layer in MF, which has a shape of 1×10 based on user IDs, represents the distributed representation of each user. Therefore, by extracting the weight of the embedding layer and calculating the cosine similarity among all users, it can be said that the user with the highest cosine similarity is the most similar user. Then, since similar users are considered to have similar sensibilities to the user receiving recommendations, the proposed method recommends furniture that similar users have rated highly but which the recommendation-receiving user has not yet rated. We then input the furniture images to the multi-layered Siamese model that we previously constructed [11], and calculate the distance between the furniture images and all the furniture images in the database. In this case, the furniture image with the smallest output from the multi-layered Siamese model, that is, the furniture that is most compatible with the recommended furniture's style, is suggested as an additional recommendation. In this way, the proposed method can recommend not only furniture that similar users have rated highly but also multiple furniture items that are compatible with the recommendation-receiving user's furniture style.

V. EXPERIMENT WITH SIMILAR USER EXTRACTION USING MATRIX FACTORIZATION

In this experiment, we aimed to extract users' preferences using MF to make furniture recommendations that take into account individual sensibilities. Specifically, we identify similar users based on cosine similarity calculations from user features extracted from the MF's embedding layer. However, it was necessary to verify the effectiveness of recommending furniture that has been highly rated by similar users. Therefore, we propose a method to evaluate which is better at extracting user preferences, furniture that MF inferred to be highly rated by conventional MF, or furniture that is actually highly rated by similar users, and then verify the method using the procedure described above.

In Step 3, the missing value was created by deleting the rating data (e.g., userID 1, itemID 1, rating "5") of the recommendation-receiving user's item, which was rated "5" by both the recommendation-receiving user and similar users. By learning the missing MF using the above missing-value dataset, we could confirm the inference ability for the actually highly rated items when the training data were lost. Specifically, when the missing MF highly rated the missing value during inference, it indicated a high MF inference ability, whereas when the missing MF rated the missing value low, it indicated a low MF inference ability. Thus, recommending items rated highly rated by similar users is better. In this way, we could verify the effectiveness of recommending furniture that has been rated highly by similar users, compared to furniture with high MF inference values.

A. RATING DATA SET

In this experiment, we used the Rakuten dataset [12], from Rakuten Ichiba, which collected users' rating data for items to learn their preferences. These data refer to products sold between 2015 and 2019; we extracted only rating data for furniture items. In addition, for the MF training, we extracted only the rating data from users who had rated 20 or more furniture items, as users need to rate a certain number of furniture items for MF training. The rating values had six levels, ranging from 0 to 5, and the data consisted of three elements: userID, itemID, and rating. The newly created rating dataset taken from the above data shaping consisted of 7,644 users and 127,819 items, for a total of 225,095 ratings (e.g., rating "5" for itemID 1 by userID 1). The created rating dataset was used in a 4:1 ratio for training and validation during the MF training.

B. EXPERIMENT SETTINGS

When training of the original MF model, we adopted the stochastic gradient descent algorithm for optimization, with

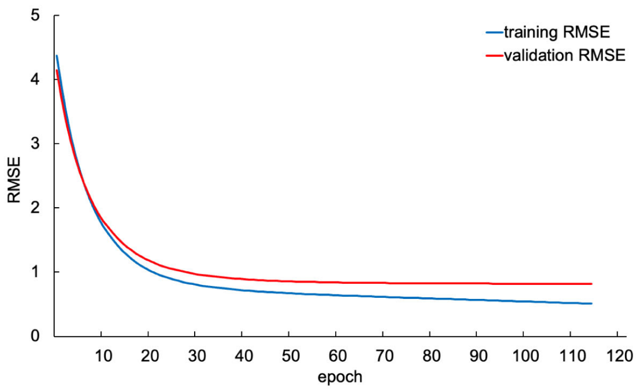

FIGURE 3. RMSE transition of original MF.

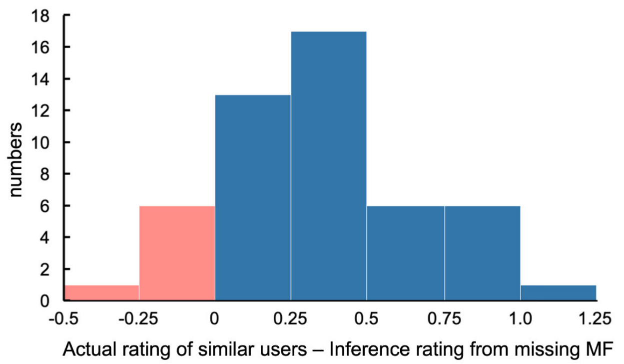

FIGURE 4. Histogram of the differences between the actual ratings by similar users and the predicted ratings of the missing MF.

a learning rate of 0.01. The rating dataset mentioned earlier was used as the training data, with a batch size of 50. The Root Mean Squared Error (RMSE) was used as the evaluation metric during training, and the training was terminated if there was no improvement in RMSE after five epochs due to early stopping. We performed a series of 10 training sessions, after which the model with the best performance was used for evaluation.

C. EXPERIMENT RESULTS

The learning results of the original MF are shown in Fig. 3. After 10 learning sessions, the highest accuracy of validation RMSE in a single learning session was, on average 0.822, and the best model achieved 0.821. Therefore, the original MF training has an error of about 0.8 in the inferred values in the rating range of 0 to 5. Compared to previous studies [27], [28], [29] that evaluated MF using RMSE in the same rating range, this level of accuracy is considered sufficient for creating a model.

Next, we conducted 50 trials and plotted a histogram, as shown in Fig. 4, to compare the differences between the actual ratings of similar users and the inferred ratings from the missing MF. In the histogram, the difference is calculated by subtracting the inferred ratings from the missing MF from the actual ratings of similar users. Positive values indicate that the actual ratings of similar users are better, while negative values indicate that the inferred ratings from the missing MF are better. As shown in Fig.4, there were 43 positive values and 7 negative oens. This means that, in 43 out of 50 cases, the actual ratings of similar users were higher than the ratings inferred from missing MF. Furthermore, compared to negative values, positive values were distributed in a larger range, from 0 to 1.25. Therefore, we determined that, when training with missing rating data for furniture that the user actually rated highly, it tends to be difficult to infer high ratings. From the above, we concluded that the actual ratings of similar users reflect the target user's preferences more accurately, compared to the ratings inferred ratings from missing MF.

VI. EXPERIMENT OF PROPOSED RECOMMENDATION METHOD

This experiment aims to verify the performance of a proposed recommendation method that recommends furniture items that have not only been rated highly by similar users, but also furniture items that match the user's preferred style. Specifically, in Section IV, we selected one image rated highly by similar users out of 50 test cases and used the multi-layered Siamese model to recommend a set of furniture images from the database that were the closest in distance. Furthermore, we validated the recommended furniture items using the proposed one-to-two Siamese model, which is a Siamese model that learns three furniture images in a one-to-two relationship, and proposed a one-to-three Siamese model, which learns four furniture images in a one-to-three relationship. We also conducted the same validation using only MF as a comparison, and used MF and a conventional Siamese model to recommend furniture items. Then, other items that were highly rated by similar user were also listed and compared with the furniture recommended by the proposed method. In this experiment, four users were randomly selected for recommendation, and four examples were shown as validation results.

A. EXPERIMENT SETTINGS

In this experiment, each Siamese model is validated using a test set of input images from a dataset. Specifically, we create a new $\boldsymbol{I}^t$ by replacing one of the images in $\boldsymbol{I}^t$ with I^b. This results in a set of 4,800 $\boldsymbol{I}^t$, where 2,400 sets consist of furniture images with the same style, and the remaining 2,400 sets contain furniture images with mixed styles. Additionally, a furniture image that was highly rated by a similar user is defined as the new I^b. Therefore, we input the new I^b and $\boldsymbol{I}^t$ into each Siamese model, extract the $\boldsymbol{I}^t$ that is closest to I^b in the feature space, and analyze the results.

B. EXPERIMENT RESULTS

Fig. 5 shows four examples of output images from each Siamese model and items that were highly rated by similar users of MF. From Verification example 1, in Fig. 5, we can see the furniture recommended based on the ratings by users

FIGURE 5. **Examples of output images for each siamese model and images from similar user on MF.**

similar to target user, and the furniture that the Siamese model estimated to have the highest style compatibility. The results show that using only the MF method results in one furniture image being recommended, while using each Siamese model increases the number of recommended images. In particular, the proposed one-to-two Siamese model recommends furniture with blue and white tones, both of which are of the ''beach'' style. By contrast, the proposed one-to-three Siamese model recommends furniture with blue and neutral tone tones, which are all of different styles but have a unified feel to the recommendation. In contrast, the conventional Siamese model recommends a lamp in white, which appears to be compatible with the style. However, since only one additional recommendation is made, the multi-layered Siamese model, which recommends more furniture images, is considered capable of recommending furniture that has higher overall style compatibility. Compared to other furniture highly rated by similar users using MF, the character was designed in the design of simplicity. Therefore, there was not as much style compatibility between the highly rated items as the furniture recommended by the proposed Siamse model.

Next, from Verification example 2 in Fig. 5, we can see that each Siamese model is recommending furniture with high similarity in color and shape, as in Validation Example 1. In addition, the proposed one-to-two Siamese model recommends all ''farmhouse'' style furniture, and the one-to-three Siamese model recommends all ''modern'' style furniture. However, we found that these pieces have style compatibility with furniture that similar users rated highly, thus indicating the possibility of recommending different style systems. In comparison to other furniture items highly rated by similar user using MF, leaf-patterned furniture was found among the modern furniture items. Therefore, there was not as much style compatibility between the highly rated items as the furniture recommended by the proposed Siamse model. In Verification exmaple 3 and Verification example 4, the Proposed Siamese model showed a higher fit than MF. It not only recommended more furniture than conventional Siamese, but also succeeded in taking affinity into account.

From the above, it was determined that the proposed recommendation method can recommend multiple furniture items of compatible styles, not just based on personal preferences. We also found that, while the conventional Siamese model can recommend furniture with style compatibility, the multi-layered Siamese model can not only increase the number of candidate furniture recommendations but also consider the overall affinity among furniture items. Therefore, proposed method using the multi-layered Siamese model is effective for practical use.

VII. CONCLUSION

In this study, we proposed a furniture recommendation method that takes into consideration the compatibility of styles among furniture pieces while considering users' preferences. With regard to these preferences, we achieved a more accurate extraction of sensibilities, compared to conventional methods, by extracting similar users' ratings based on an MF model trained with rating data for furniture. Additionally, we verified the effectiveness of using rating data from similar users compared to conventional MF-based inference values through a new evaluation method that omits rating data during training. Regarding the compatibility of styles, we used a Siamese model that evaluates the compatibility of three or more furniture images to improve the accuracy of compatibility estimation by increasing the number of images in $\boldsymbol{I}^t$ [11].

The proposed furniture recommendation method uses the above MF and the multi-layered Siamese model, and the actual recommendations were verified using the proposed method. As a result of the verification, the proposed method succeeded in recommending multiple pieces of furniture that are compatible with the user's preferred style, not only taking into account the individual sensitivity of the user with MF. Therefore, the proposed method is also considered be a high-performance method, from a practical standpoint, since it is capable of recommending multiple pieces of furniture that are compatible with one another and can even consider the overall affinity of multiple pieces of furniture. We think it would be best if we could quantitatively verify the compatibility and affinity (e.g., similarity) in this study. However, since actual compatibility differs from person to person, a data set that fits each user is necessary. This is not feasible in this experiment, so we conducted a visual verification.

ACKNOWLEDGMENT

In this article, they used the Rakuten Dataset (https://rit.rakuten.com/data_release/), provided by Rakuten Group Inc., via the National Institute of Informatics' IDR Dataset Service.

REFERENCES

[1] C. Mu, J. Zhao, G. Yang, J. Zhang, and Z. Yan, ''Towards practical visual search engine within elasticsearch,'' 2018, *arXiv:1806.08896*.

[2] S. Li and E. Karahanna, ''Online recommendation systems in a B2C e-commerce context: A review and future directions,'' *J. Assoc. Inf. Syst.*, vol. 16, no. 2, pp. 72–107, Feb. 2015.

[3] H. Hwangbo, Y. S. Kim, and K. J. Cha, ''Recommendation system development for fashion retail e-commerce,'' *Electron. Commerce Res. Appl.*, vol. 28, pp. 94–101, Mar. 2018.

[4] P. B. Thorat, R. M. Goudar, and S. Barve, ''Survey on collaborative filtering, content-based filtering and hybrid recommendation system,'' *Int. J. Comput. Appl.*, vol. 110, no. 4, pp. 31–36, Jan. 2015.

[5] N. Polatidis and C. K. Georgiadis, ''A multi-level collaborative filtering method that improves recommendations,'' *Exp. Syst. Appl.*, vol. 48, pp. 100–110, Apr. 2016.

[6] Z. Hu, Y. Wen, L. Liu, J. Jiang, R. Hong, M. Wang, and S. Yan, ''Visual classification of furniture styles,'' *ACM Trans. Intell. Syst. Technol.*, vol. 8, no. 5, pp. 1–20, Sep. 2017.

[7] S. Yoon, H. Oh, and J. Y. Cho, ''Understanding furniture design choices using a 3D virtual showroom,'' *J. Interior Design*, vol. 35, no. 3, pp. 33–50, May 2010.

[8] D. Aggarwal, E. Valiyev, F. Sener, and A. Yao, ''Learning style compatibility for furniture,'' 2018, *arXiv:1812.03570*.

[9] B. Liu, J. Zhang, X. Zhang, W. Zhang, C. Yu, and Y. Zhou, ''Furnishing your room by what you see: An end-to-end furniture set retrieval framework with rich annotated benchmark dataset,'' 2019, *arXiv:1911.09299*.

[10] T. Weiss, I. Yildiz, N. Agarwal, E. Ataer-Cansizoglu, and J. Choi, "Image-driven furniture style for interactive 3D scene modeling," *Comput. Graph. Forum*, vol. 39, no. 7, pp. 57–68, Nov. 2020.

[11] A. Taisho, K. Ono, E. Makihara, N. Ikushima, and S. Yamakawa, "Furniture style compatibility estimation by multi-branch deep Siamese network," *Math. Comput. Appl.*, vol. 27, no. 5, p. 76, Sep. 2022.

[12] Rakuten Group. (2014). *Rakuten Dataset, Informatics Research Data Repository, National Institute of Informatics. (Dataset)*. [Online]. Available: https://doi.org/10.32130/idr.2.0

[13] A. Pujahari and D. S. Sisodia, "Pair-wise preference relation based probabilistic matrix factorization for collaborative filtering in recommender system," *Knowl.-Based Syst.*, vol. 196, May 2020, Art. no. 105798.

[14] H. Zhang, I. Ganchev, N. S. Nikolov, Z. Ji, and M. O'Droma, "FeatureMF: An item feature enriched matrix factorization model for item recommendation," *IEEE Access*, vol. 9, pp. 65266–65276, 2021.

[15] H. Ma, H. Yang, M. R. Lyu, and I. King, "SoRec: Social recommendation using probabilistic matrix factorization," in *Proc. 17th ACM Conf. Inf. Knowl. Manage.*, California, NV, USA, 2008, pp. 931–940.

[16] X. He, H. Zhang, M-Y. Kan, and T.-S. Chua, "Fast matrix factorization for online recommendation with implicit feedback," in *Proc. 39th Int. ACM SIGIR Conf. Res. Develop. Inf. Retrieval*, Pisa, Italy, 2016, pp. 549–558.

[17] D. Liang, J. Altosaar, L. Charlin, and D. M. Blei, "Factorization meets the item embedding: Regularizing matrix factorization with item co-occurrence," in *Proc. 10th ACM Conf. Recommender Syst.*, 2016, pp. 59–66.

[18] S. Bell and K. Bala, "Learning visual similarity for product design with convolutional neural networks," *ACM Trans. Graph.*, vol. 34, no. 4, pp. 1–10, 2015.

[19] Y. Li, H. Su, C. R. Qi, N. Fish, D. Cohen-Or, and L. J. Guibas, "Joint embeddings of shapes and images via CNN image purification," *ACM Trans. Graph.*, vol. 34, no. 6, pp. 1–12, Nov. 2015.

[20] Y. Koren, R. Bell, and C. Volinsky, "Matrix factorization techniques for recommender systems," *Computer*, vol. 42, no. 8, pp. 30–37, Aug. 2009.

[21] A. Veit, B. Kovacs, S. Bell, J. McAuley, K. Bala, and S. Belongie, "Learning visual clothing style with heterogeneous dyadic co-occurrences," 2015, *arXiv:1509.07473*.

[22] L. F. Polanía and S. Gupte, "Learning fashion compatibility across apparel categories for outfit recommendation," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Taipei, Taiwan, Sep. 2019, pp. 4489–4493.

[23] H. Yuan, G. Liu, H. Li, and L. Wang, "Matching recommendations based on Siamese network and metric learning," in *Proc. 15th Int. Conf. Service Syst. Service Manage. (ICSSSM)*, Hangzhou, China, Jul. 2018, pp. 1–6.

[24] G. Gao, L. Liu, L. Wang, and Y. Zhang, "Fashion clothes matching scheme based on Siamese network and AutoEncoder," *Multimedia Syst.*, vol. 25, no. 6, pp. 593–602, May 2019.

[25] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2014, *arXiv:1409.1556*.

[26] J. Deng, W. Dong, R. Socher, L. J. Li, K. Li, and L. Fei-Fei, "ImageNet: A large-scale hierarchical image database," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Miami, FL, USA, Jun. 2009, pp. 248–255.

[27] W. Wang and Y. Lu, "Analysis of the mean absolute error (MAE) and the root mean square error (RMSE) in assessing rounding model," *IOP Conf. Ser., Mater. Sci. Eng.*, vol. 324, no. 1, 2018, Art. no. 012049.

[28] H. Byström, *Movie Recommendations from User Ratings*. Stanford, CA, USA: Stanford University, 2013.

[29] S. Sedhain, A. K. Menon, S. Sanner, and L. Xie, "AutoRec: Autoencoders meet collaborative filtering," in *Proc. 24th Int. Conf. World Wide Web*, 2015, pp. 111–112.

MASAYOSHI TAKEDA (Graduate Student Member, IEEE) received the B.E. degree in information and computer science from Doshisha University, Japan, in 2022, where he is currently pursuing the M.E. degree. His research interests include deep learning and recommendations that take into consideration users' preferences.

KEIKO ONO (Member, IEEE) received the B.S., M.S., and Ph.D. degrees in engineering from Doshisha University, Kyoto, Japan, in 2001, 2003, and 2007, respectively. In April 2009, she joined Doshisha University as an Assistant Professor with the Organization for Research Initiatives and Development. In April 2010, she joined Ryukoku University, Kyoto. Currently, she is an Associate Professor with the Department of Intelligent Information Engineering and Sciences, Doshisha University, in 2020. Her research interests include parallel computing, evolutionary optimization, and machine learning. She is a member of the Japanese Society for the Information Processing (IPSJ) and the Japanese Society for Evolutionary Computation (JPNSEC).

AYUMU TAISHO received the B.S. degree in information and computer science from Doshisha University, in 2021, and the M.S. degree from the Graduate School of Science and Engineering, Doshisha University, in 2023. His research interests include deep learning and recommendations that take into consideration the users' preferences.