# Intelligent Recommendation System Based on Image Processing

**Qi Lin***

International School, Beijing University of Posts and Telecommunications, Beijing 100876, China

* Corresponding author e-mail: linwel@bupt.edu.cn

**Abstract**. Recommendations for images are the core tasks of many website recommendation systems. This paper develops a relatively complete image recommendation system based on the relative data analysis of the hotel website. In this paper, the image processing methods such as computer vision in the field of deep learning are applied to the information processing classification of the image of the hotel website. The image recommendation system established in this paper is more targeted than the recommendation system of some previous websites. This paper evaluates the aesthetics of NIMA images of the photos of the houses, and uses the show and tell model to describe the image and text, and finally according to the star rating and user preferences. Use collaborative filtering algorithms for house listings and recommendations. The accuracy of the recommendation is greatly improved, and the quality of the service selected by the user can be effectively guaranteed.

## 1. Introduction

The amount of information is growing rapidly on various services, applications, and websites. The huge amount of information on the network poses a problem for the user's choice. The large amount of online information displayed on the website lacks objectivity. For example, many home-stay descriptions and pictures come from the landlord, and users may be confused and misled by false propaganda to choose risky services. Many online information is not clearly differentiated. Some ratings are similar. The website does not want to give the service a differentiated display. Instead, it wants more sales and more users, but it may affect the user's rights and interests. There is still a lot of reference to online information, and some of the introductory information does not describe the key issues, so that users are very upset when choosing the same service products.

In view of the above problems, this paper establishes an intelligent recommendation system based on deep learning. The system used the photo of the house uploaded by the landlord, built a photo dataset of the B&B, used NIMA's model on the ava dataset, and carried out the migration study on the photo dataset of the B&B to get the image aesthetic score. We use the show and tell network to implement image text descriptions, and we can use the text description results and the house description to calculate the similarity. Through the use of massive online data for analysis, we fully explore and use all aspects of valuable listing information, refer to the existing hotel star rating standards, the relevant regulations of the hotel industry and other systems, and build a comprehensive rating and rating system, presented to User information with uniform user standards and clear grades. Based on the existing information, we choose the appropriate collaborative filtering algorithm, build a recommendation system, and recommend personalized listings for users to make users more reasonable choices. This kind of intelligent analysis recommendation system that can be extended to other service areas will greatly improve the service experience, standardize the hidden dangers, protect the legitimate rights and interests of consumers, and enable people to fully enjoy the sharing of the mobile Internet era by standardizing and standardizing the local service industry such as B&B. The convenience brought by the economy.

The main innovations and contributions of this paper are:

1. Link the scoring system to the recommendation system, so that the information recommended to the user will not only traditionally depend on whether the user's preferences and item attributes match, but also whether the product quality level based on the data analysis meets the user's expectations;

2. Apply the information processing and the recommendation technology to the home-stay and other new daily-life service fields, and provide a reference to standardize the market order under the mode of sharing economy, which can promote the integration of the sharing economy and the mobile Internet.

## 2. Related Work

We have carried out a series of optimization improvements using the recommendation system proposed in [1].In Liu et al.'s paper [1], the recommendation system has room for improvement. This article uses a content-based recommendation algorithm to treat content-related keywords in collections of movies and users equally. Experiments show that there are many keywords with semantic duplication and low reference value, which affects the accuracy of recommendation results. To solve this problem, label clustering can be performed, and the weight of the keyword is considered. In addition, the content-based recommendation algorithm has the problem of over-specialization. Therefore, this paper chooses the collaborative filtering algorithm to make recommendation results more diverse and personalized.

In terms of image aesthetic evaluation, Guo et al. [2] proposed a method for establishing an image aesthetic database and evaluating image aesthetics based on web crawlers. First, they used the web crawler to capture the database image, and then extracted the aesthetic features of the image for aesthetic evaluation. The aesthetic features extracted by this method involve both global and regional features. Therefore, it was necessary to separate the subject and background regions of the image by using the power spectral slope before feature extraction. The experimental results show that this method can capture and analyze the network image well, and the evaluation method is in good agreement with human subjective perception and can accurately reflect the human aesthetic appreciation system. However, this method extracts features by artificial design. At present, the recognition performance of some specific categories is improved, but it is not good on other types of images under realistic conditions, and it is not robust against unknown data. Therefore, this paper chooses the NIMA by using the convolutional neural network to learn and extract the features of massive pictures. It is more robust against unknown data and has improved accuracy, which can better adapt to practical applications.

Considering the advantages and disadvantages of these existing systems, we propose our system to better perform data mining and recommend to users.

## 3. Method

### *3.1. Data Processing*

The image data we crawled is the real shot of the home uploaded by the host. What information can be obtained from these images is the key to the image processing of this system.

First of all, we explore from the perspective of image aesthetics. Through the work of Liu et al. [3], we understand that human aesthetic features mainly include the recognition and understanding of color and the main body of image, where color relates to the color distribution, color contrast, and color harmony of the image. In order to allow the computer to simulate the human brain to make an aesthetic judgment on the image, we use NIMA's deep learning method to evaluate the image of the home. NIMA's quality and aesthetic predictor stands on image classifier architecture, replacing the last layer of the baseline CNN with a fully-connected layer with 10 neurons followed by softmax (as shown in Fig. 1). Baseline CNN weights are initialized by training on the ImageNet dataset [4,5], and then an end-to-end training on quality assessment is performed [6]. In order to make the model more suitable for the

Journal of Physics: Conference Series **1449** (2020) 012131 doi:10.1088/1742-6596/1449/1/012131

B&B (Bed and Breakfast, homestay) field analyzed by this system, we have determined the weight of the sensitive feature set (including numerical features and location features) that can be used as the judging criteria for the image of the home. Then we build the dataset of images, and use NIMA's pre-trained model on the AVA dataset for migration learning with the convolutional neural network on this homestay images dataset. In order to reduce overfitting during training, we rescaled the input image to 256*256 and randomly extracted a clipping block of size 224*224 to increase the amount of data. The test results show that our model has good robustness, and the image aesthetic score of the final prediction evaluation can reasonably reflect the aesthetic quality of the image, which is a reference for the overall rating of the stay.

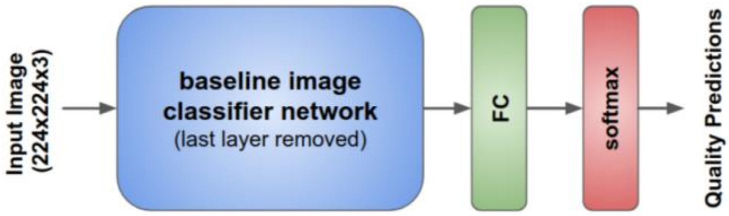


Fig. 1: Modified baseline image classifier network used in the framework.

For image data, we also want to mine the home information of environment, facilities, etc. from the images. This requires the computer not only to detect the object properties in the image, but also to understand the relationship between the objects, and finally use a reasonable language to express it. In order to achieve the mapping from the picture to the descriptive statement, we choose the encoder-decoder framework of the "show and tell" basic model to encode the picture into a fixed vector, and then decode the description of the picture from this vector. We collect a large number of photos shot inside the room and corresponding descriptions as our training set, and train in three parts. The first part is image embedding. We first encode the picture into the vector of 224*224*3, and then output the vector of 512 - ImageVec by the encoding function – CNN, which composed of 512 neurons fully connected layer followed by the google inception. The second part is word embedding. The m words in the corpus are used as m-dimensional input vectors, which are transformed into 512-dimensional output vectors by transformation matrix. The third part is decoding. This process is in a layer of Long Short-Term Memory (LSTM). Each step of the output is followed by a softmax classifier (as shown in Fig. 2). The number of classifiers is equal to the vocabulary size in the corpus [7], [8]. In order to make the trained model able to be associative -- it can also produce a correct description for the unknown picture, we let the model accept the predicted word as the next input with certain probability during the training phase, so that it can enter the state of prediction during training, and learn to correct the error itself. In the prediction phase, the word with the highest probability of output from the model is selected at each step of generating word by LSTM until it stops when the end mark is predicted. After the final caption result is processed by natural language processing (NLP), it is compared with the introduction of the corresponding stay, and the similarity between the two is calculated to help determine whether the description from the host is as objective and accurate as the fact.

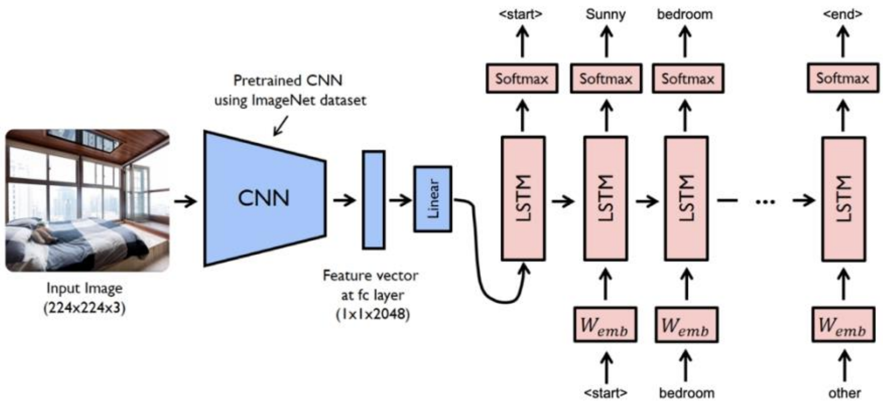


Fig. 2: Example of our Neural Image Caption (NIC) model

### *3.2. Scoring System*

The aesthetic score of images of the home can objectively reflect the environmental conditions accounting for 60% of the environment score. The similarity between the "Show and tell" results and the description written by the host can objectively reflect this point, accounting for 60% of the description score.

According to the rating of the stay more comprehensive, we divide the service quality of the hotel into five levels: five-star, four-star, three-star and two-star, which are corresponding to luxury, exquisite, comfortable and economical. 108 points and above is five-star; 96 to 107 points is four-star; 84 to 95 points is three-star and 72 to 83 points is two-star. The higher the star rating of the hotel, the better the reception facilities and the higher the service quality. Users can make reasonable choices according to the star rating and meet their own needs.

### *3.3. Recommendation System*

When users face a large number of similar listings displayed by the system, they often need a lot of browsing time to find a satisfactory one. How to improve information utilization, and help users quickly find their favorite listings by exploring long tails is an important problem that this paper needs to solve. It is the final step of this paper to make full use of the characteristics and grades of the stays obtained in all previous processes to personalize the recommendation based on the preference and the need of a user.

We analyzed the advantages and disadvantages of four types of recommendation algorithms: popularity-based, content-based, model-based, and collaborative filtering (CF). The popularity-based algorithm is simpler and suitable for new users, but it cannot implement personalized recommendation. The content-based algorithm uses the keywords in the text content as attributes, and there is an over-specialization problem, which makes the recommended content lose diversity. The regression prediction is a model-based algorithm that can fit accurate regression functions and calculate the weights corresponding to each eigenvalue. However, manual intervention is needed to combine and filter attributes, and mathematical models online need to be updated repeatedly to adapt to changes. The collaborative filtering algorithm is a commonly used algorithm for e-commerce websites, including user-based CF and item-based CF. The CF algorithm is to calculate the cosine between the preference vectors in the correlation matrix to represent the similarity. The closer the value is to 1, the more similar the two users or two items are [9]. For the problem of matrix sparsity, we solve it by decomposing a matrix of n*m into a matrix of n*k multiplied by a matrix of k*m by matrix factorization. On the whole, as long as we design a method of solving the cold start problem, the collaborative filtering recommendation algorithm can be perfectly applied to the homestay product recommendation of our system, making use of the information such as the stay label obtained by the information processing.

Since the data mining and processing of this system is mainly for housing, we choose the item-based CF for recommendations. For different users, we analyze their browsing history and reservation record.

According to the label, rating, grade and other information of the stay in the record, we calculate the similarity between all the stays. For the stay with high user rating, we find the most similar N stays and recommend them to the user. When a new user enters the system, in order to let the system know the user's preferences, we set the "initial question" in the user registration process, that is, show 5 different model homes in different levels with one or more labels among "well-equipped facilities", "clean environment", "convenience location", "good service", "accurate description", "good experience" (the example is shown in Fig. 3), and requires the user to select three of the favorite ones. We set a larger weight for the scores of the top three labels in the three listings, and sort the list for new users by the weighted scores from high to low, thus solving the cold startup problem of the recommendation system.

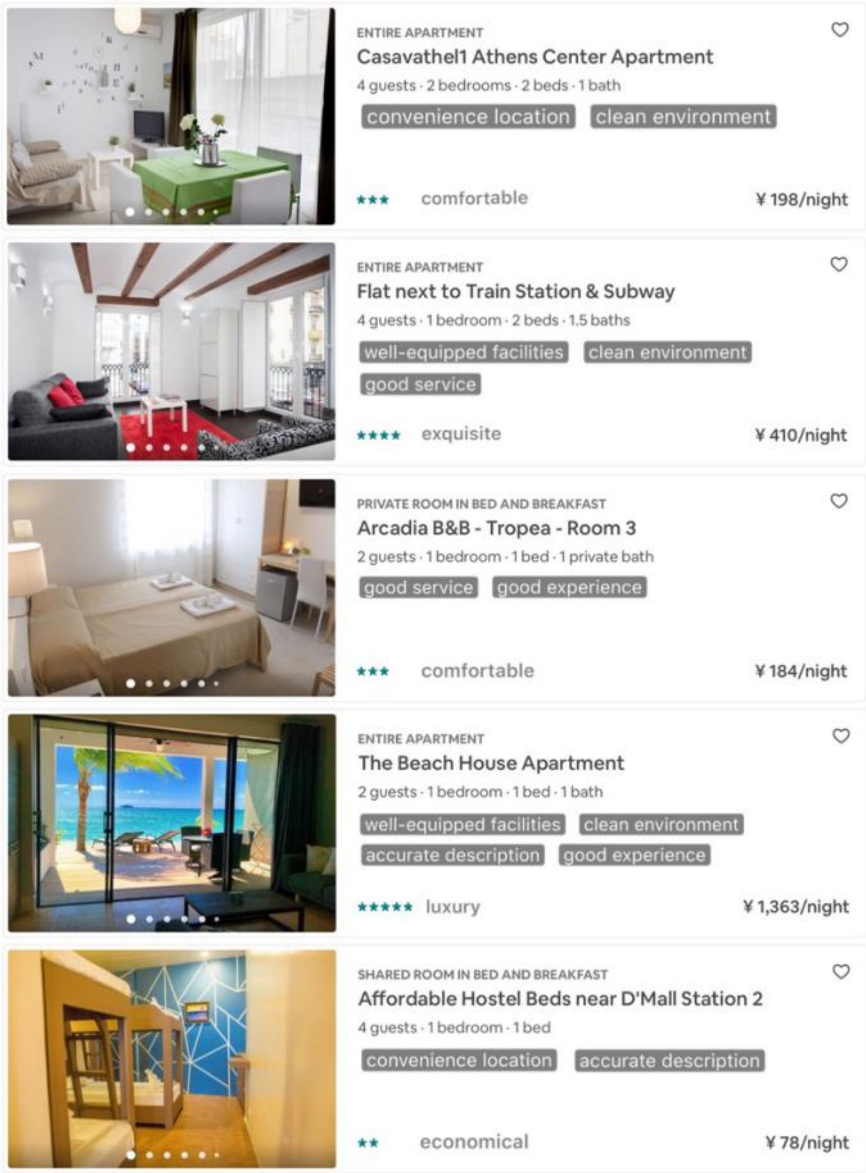


Fig. 3: Example of the "initial question" for new users (For the convenience of the reader, we translated the Chinese in the screenshot of interface here into English. )

To verify the effectiveness of our recommendation system, we conducted actual tests. We use flask to develop the website as a user interface and connect to the back-end database to present the full functionality of the system (the interface is shown in Fig. 4). Users can set the sorting rules of the list of stays according to individual needs (such as: level descending, price ascending, guess what you like, etc.) and filter conditions to browse the recommended listings. We randomly searched for 100 people of different ages and occupations to use our front-end system. Our improved recommendation system can be used to achieve personalized homestay recommendations that meet the individual needs and preferences of users.

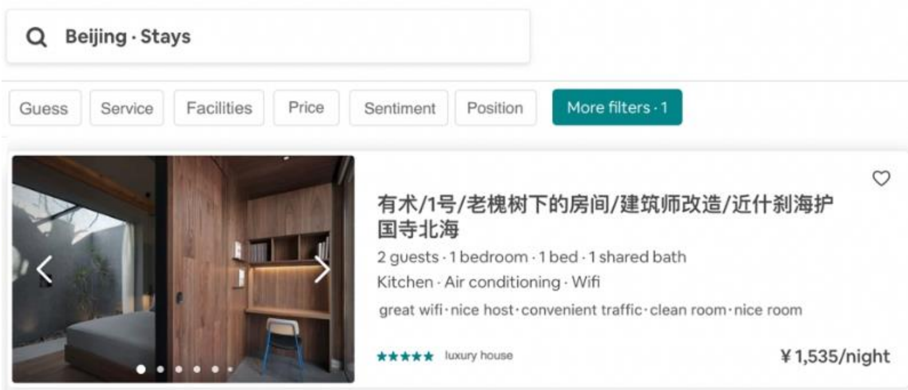


Fig. 4: Interface of our website

## 4. Experiments

### *4.1. Results of different methods of NIMA*

We use different basic backbones as feature extraction modules to extract valid information from images. We build a B&B dataset with about 10,000 images and use 70% of the images for training, 20% for verification, and 10% for testing. We use two 1080Ti GPUs for transfer learning based on NIMA's pre-trained model on the AVA dataset. We test the model after adjusted parameters using the verification set, and the test results are shown in Table 4.

Table1: Test results of different methods of NIMA

| Basic network | RMSE | Model size | Time |
|---|---|---|---|
| **Inception** [26] | 2.82 | 219.1M | 72.9 |
| **NASNet** [27] | 3.36 | 20M | 45.9 |
| **MobileNet** [28] | 3.03 | 13.2M | 22.45 |

Experiments show that the root mean square error (RMSE) of NasNet as the backbone is too large. The RMSE of Inception as the backbone is the smallest among the three, but the model takes up a lot of space and takes a long time. The MoblileNet architecture demonstrates the fastest speed with the smallest model in the test, due to the dense filter is replaced by a separable filter, so it is indeed an efficient deep CNN. So, we finally choose MobileNet as the backbone to get the results of our image quality.

## 5. Conclusion

Taking the website of the hotel as an example, this paper constructed an intelligent recommendation system by using image processing technology in deep learning. Compared with other existing recommendation systems, the establishment of this system was based on the mining of massive online data, so the consideration of similarity was more critical, and the recommended basis was more adequate. At the same time, it has opened up new application fields for natural language processing, image aesthetic evaluation and other technologies, making it more practical.

**References**

[1] Liu Hui, Li Fengyin, Qi Jiguo, Cui Wei, & Ge Rui. (2018). Design and Implementation of Film Recommendation System Based on Film Criticism Mining. Electronic Technology, 47(12), 89-92.

[2] Guo Lihua, Xiong Yangchao, Ma Jun, & Liu Yaneng. (2014). Image aesthetic quality database of web crawlers and its aesthetic evaluation system design. Systems Engineering Theory and Practice (S1), 228-233.

[3] Gao Han, Tang Jianglong, Liu Jiafeng, & Jin Ye. (2013). Image aesthetic evaluation based on image classification. Intelligent Computer and Application, 3(4), 39-41.

[4] Krizhevsky, A. , Sutskever, I. , & Hinton, G. . (2012). ImageNet Classification with Deep Convolutional Neural Networks. NIPS(Vol.25). Curran Associates Inc.

[5] QI, X. (2017). The Understanding of Convolutional Neuron Network Family. DEStech Transactions on Computer Science and Engineering, (cii).

[6] Talebi, H. , & Milanfar, P. . (2017). Nima: neural image assessment. IEEE Transactions on Image Processing, PP(99), 1-1.

[7] Luo Zhengqing, & Zheng Tao. (0). Collaborative filtering recommendation algorithm based on label clustering. The 13th (2018) China Management Annual Conference.

[8] Guo Lihua, Xiong Yangchao, Ma Jun, & Liu Yaneng. (2014). Image aesthetic quality database of web crawlers and its aesthetic evaluation system design. Systems Engineering Theory and Practice (S1), 228-233.

[9] Weiliang, K. (2013). Research on Key Issues of Collaborative Filtering Recommendation System [D]. Central China Normal University.