Digital Inheritance of POPF Based on Image Database and Identification Model

Guancan Yang, Huilin Zhang, Xiaomei Zhang, Yue Yu, Jie Yang

School of Information Resource Management, Renmin University of China, Beijing, China
yanggc@ruc.edu.cn, lhz@ruc.edu.cn, 2016202313@ruc.edu.cn,
2017202132@ruc.edu.cn, 2548264230@qq.com

Abstract. Pecking Opera Painted Faces (POPF) is a Chinese intangible cultural heritage full of aesthetic value. However, their cultural connotations become less known in modern society. It is therefor, necessary to make public easily understand unique cultural connotation and aesthetic value of POPF. In this project, we built a POPF image database based on the image classification model, with the aim to better showcase POPF. The images are classified into multiple categories according to the traditional connotation patterns. By using the MLP (Multilayer Perceptrons) algorithm classifier, the classification accuracy of model has approached 70%. Using those tools, the uncertainty of information about images of POPF can be reduced appreciably, and would benefit the innovation of derivatives around POPF.

Keywords: Pecking Opera Painted Faces (POPF), Intangible Cultural Heritage, Image Database, Identification Model.

1 Introduction

Peking Opera Painted Faces (POPF) is born of the art of Peking Opera, which is an essential type of intangible cultural heritage. They are painted by Peking Opera actors when they perform on the stage, used to describe the background and characteristic of unique roles with symbols. There are specific formats of facial painting in the aspect of color, type, and shape, so professional spectators would quickly tell the characteristic of a role. The earliest POPF were crude and straightforward, but with time, the designs became more elaborate and ornamental. However, the unique cultural connotation and aesthetic value of POPF are less known to the general public.

On the other hand, facial make-up is extremely stylized in the colors and patterns used; and no two painted faces are alike. There are many genres and types in POPF, which is a challenge to understand the POPF. For instance, there are more than 100 face-painting styles designed for Xiang Yu, the hero in Farewell My Concubine, and each one is very different.

In this case, there are problems with POPF's inheritance: One of the problems is its multi-genre inheritance, the other one is unclear cognition of connotation. At present,

POPF was inherited by part of people discretely. In such a way of inheritance, many images of POPF were lost. What is sadder is that the cultural connotations behind POPF are increasingly being forgotten. This oblivion makes the inheritance of POPF more difficult. To solves the above problems. Our research team designed an image database to store the information we collected. Besides, our team takes account of some connotation information missed, so an image classification model is needed to automatic supplement connotation labels. Eventually, we designed a website to make all the things accessible to the public.

2 Related Work

Early study about POPF concentrated on simply explaining what the color means in POPF. Recently, the focus of the study in POPF has been turned to deeply reveal connotation of POPF and automatically classify types of POPF based on algorithms. Nowadays, there are two main directions for the study of POPF. One is to use the algorithm to classify the POPF types, identify and decompose the various patterns of POPF. Zhang, Zhu, & Wang (2017) (Peng, Qing, & Zhiqiang, 2017) used the Support Vector Machine (SVM) algorithm as classifier on the basis of the SIFT (Scale Invariant Feature Transform) feature to improve the recognition speed and reduce the model size in the field of POPF. Wang, Kang, & Qin (2014) (Ding, Jinsheng, & Shengfeng, 2014) found that POPF is useful to express emotion and character of the roles by building a model of facial motion. Besides, the other direction is to analyze the connotation of POPF from a cultural perspective, and use examples to illustrate how to apply its connotation and thought to literary creation. Cheng &Kang (2010) (Di & Kang, 2010) introduced the Chinese zodiac character creation under FuGou view and studied them based on symbolic thinking.

3 Methodology

The aim of this project was to facilitate "the centralized preservation and inheritance of POPF". There are three key things that we prefer to highlight during the implementation of this project. Firstly, the project plants to provide a web service to the public, which would promote public engagement. Another point is that a specialized image database need built for supporting above web service. In order to avoid possible missing data, an image classification model is needed to automatic supplement information. Therefore, we use the KNIME software to build an identification model to supplement some missing data.

3.1 Data Collection & Database Designation

To construct the identification model for revealing the connotation of POPF, the data collection mainly contains two aspects. On the one hand, in order to ensure the accuracy and efficiency of the identification model, a total of 272 frontal images of POPF are

collected for creating a training dataset, which is consists of 16 different types. All these images come from an English language website of introduction to Peking Opera (http://www.paulnoll.com). Besides, about two hundred of images were also collected from various search engines to test the accuracy of the identification model. Moreover, related information about the images was collected simultaneously, such as their role names, repertoires, and so on. Then a typical image database using MySQL is created to save the images and their associated information.

3.2 Identification Model Construction

The identification model building of the study relied on a small dataset containing about 500 images, which contains a database of 272 images with labels. Moreover, we used the KNIME software to learn how to classify unknown images. The main steps are as follows:

Data collection and image processing: The image dataset consists of JPG and PNG format images, including frontal and side face types, and the size of each image is not consistent. All those 272 images with labels are imported and are divided into a training dataset and test dataset according to the proportion of 4:1. All imported images are converted into 100*100 gray-scale images; then the image pixels are converted to float and normalize the pixels between 0 and 1 in order to calculate the arithmetic values of image pixel values and calculate mean of image pixel values.

Role	Jiang Shang
Repertoire	*Opera "Wei River"*
Type of POPF	Entire face
Introduction	The leading character in Wei River, adapted from an episode in Canonization of the Gods. King Wen of Zhou, seeking capable men to help him establish his dynasty, met Jiang Shang on the banks of the Wei and made him his prime minister.

Table 1. Presentation of collected data (Typical example of POPF)

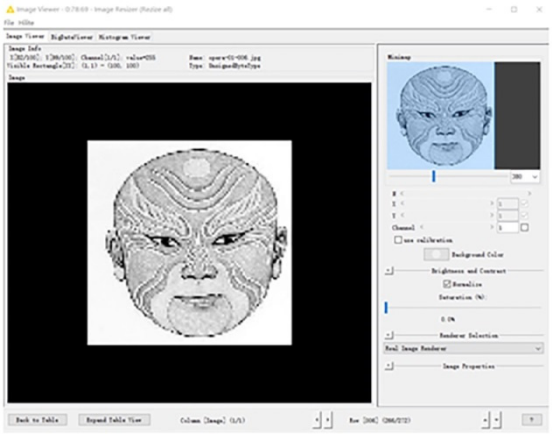

Figure 1.2 Image processing (Take Jiang Shang as an example)

Fig. 1. Data collection and image processing

Learning and evaluating the identification model: Firstly, use all the images of the learning set to learn the identification model and select image features in the *Feature* calculator and use the MLP (Multilayer Perceptrons) algorithm in the *Learning* node.

Secondly, the identification model will calculate the probability of belonging to one of these 16 categories, and the category with the highest value of the likelihood is the result of the identification model discriminating the image category.

In order to evaluate this model honestly, the study inserted the Partitioning node to divide the learning set into a train set and a test set, that is, randomly 80% of the pictures as the training set, and extracted the rest as the test set. It's found that the model was feasible by using images of the prediction set to evaluate the identification model, and its accuracy rate was about 70% after 10-fold cross-validation. The need of particular note is that only six types of POPF's sample are used to build identification model, which includes three-tile face, fragmented pattern face, cross-shaped pattern face, pictographic face, entire face, colored three-tile face. Because of only those amount of POPF's samples reach the baseline (at least 20 images) for training data, and the current identification model can not classify the remain 10 type of POPF's samples due to the shortage of training samples. Its accuracy rate was acceptable because the trained data set was small, and the data samples of each type of POPF were not uniform. In the future, the space for improving the accuracy of the model is enormous. In order to improve it, the following work needs to do: collect more high-quality images for different types to enrich the train set, adjust the calculated features, and so on.

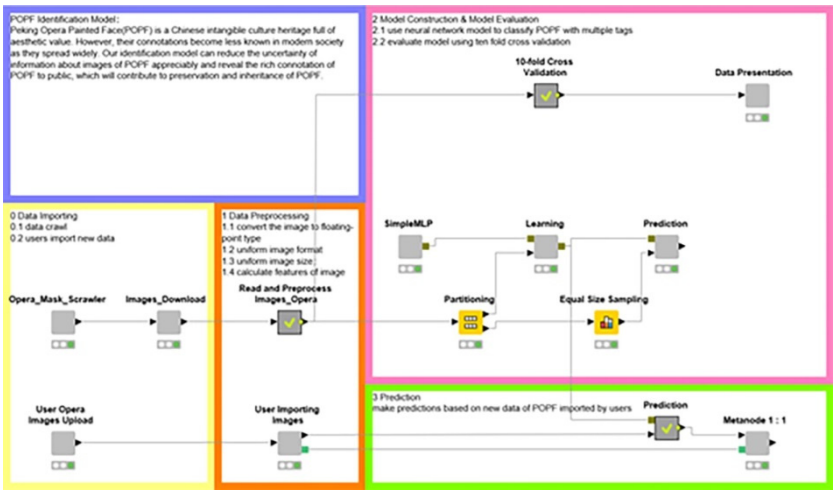

Fig. 2. The whole workflow

3.3 Web Designs of POPF

Function: The website has two main functions. On the one hand, it links to the image database, and all the information in the image database can be accessed on the Internet through it, which is dynamic. On the other hand, there is an upload port on the website. Users can upload images of POPF and then it will give them feedback by using the identification model. For those who are interested in POPF culture, the website is a

repository of knowledge. For designers, the website can be a repository of material. For POPF inheritors, the website is a communication platform.

Website designs: In general, the website has a classic look. The overall frames use scrolling design to provide users an immersive experience. The website consists of welcome, modules, an upload port and, about us.

The welcome page (shown in Figure 3.1 below) uses some magic pictures of POPF to ease the users into the learning environment. There are also some basic introductions about POPF.

Modules page (shown in Figure 3.2 below) shows the information in the image database.16 Icons represent 16 types of images of POPF, which includes the entire face (the first picture from the left in Figure 3.2 below), the three-tile face (the second picture from the left in Figure 3.2 below), the colored three-tile face (the third picture from the left in Figure 3.2 below), the six-tenths face (the fourth picture from the left in Figure 3.2 below) etc.

Through clicking on this icon, users can get this kind of representative images of POPF. For example, click on "the eunuch face", you can get Liu Jin's and Jia Gui's faces, whose characters in Peking Opera are eunuchs. Besides, there is some information about the POPF, including the introduction of the character, from which opera and the plot, which color is used and what it means, and facial structure. Users can easily access and understand representative images of POPF through above website element designs.

In the upload port (shown in Figure 3.3 below), there is a big "+" that allows users to upload their pictures. Then the identification model will compare this face with the features of 16 types of POPF, analyze its likelihood of belonging to different types, and then feedback the results with the highest probability to the website page. For example, if users upload Kong Xuan's face; the web will tell them it belongs to the goblin face because the model shows that it is the most likely to be a goblin face. After this, "yes" and "no" icons will appear on the web page. If users find problems with the results analyzed by the identification model, they can click "no" to feedback information, to help us improve our model and realize the optimization of web services based on machine learning.

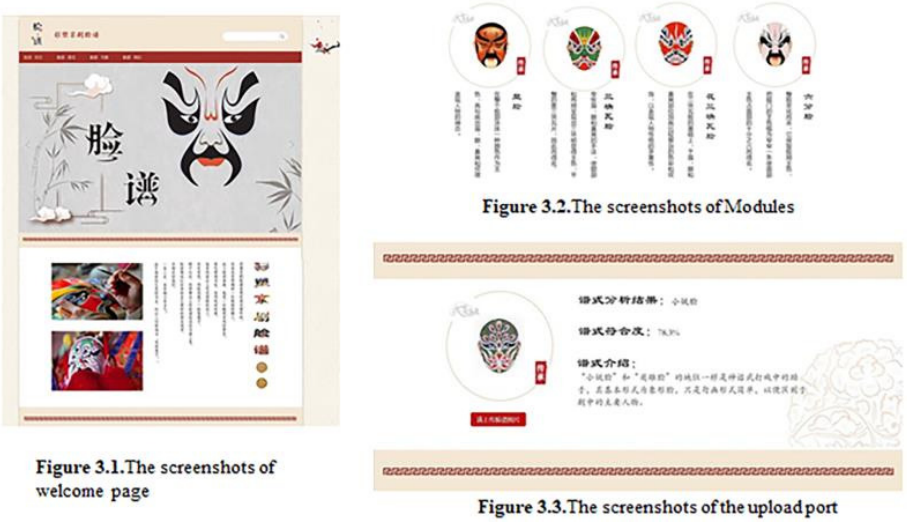

Figure 3.1.The screenshots of welcome page

Figure 3.2.The screenshots of Modules

Figure 3.3.The screenshots of the upload port

Fig. 3. The screenshots of the website

4 Conclusion

In response to the problems in the inheritance of POPF, including multi-genre inheritance and unclear cognition of connotation, the study collected data and built the image database, the identification model, and designing the website of POPF. Through the database, POPF and their related information were stored centrally and could be updated and expanded at any time. By the identification model, the public could identify the type of an unknown image to understand its relevant information. Through the website of POPF, the history, types and connotations of POPF were presented in an orderly manner in front of the world. However, although the image database stored pictures and their related information, it was small in scale. In addition, limited to the size of the data, the accuracy of the identification model was not high, and the content of the website was not sufficient. The future work is mainly to collect more high-quality images, improve the identification model and enrich the content of the website.

Acknowledgements.

This paper was supported by fund for building World-Class Universities (Disciplines) of Renmin University of China. In addition, part of data used in the paper has been from the website (http://www.paulnoll.com) and the owners of the website allow the use of the materials at no cost when not used for financial gain.

References

Di, C., & Kang, S. (2010). Study on the application of Peking Opera sign in the creation process of character animation under semiotics. *IEEE International Conference on Computer-aided Industrial Design & Conceptual Design*. IEEE.

Ding, W., Jinsheng, K., & Shengfeng, Q. (2014). The significance of information visualization based on the symbolic semantics of Peking Opera Painted Faces (POPF). *Twentieth International Conference on Automation & Computing, ICAC 2014*. Bedfordshire, UK.

Peng, Z., Qing, Z., & Zhiqiang, W. (2017). Research on classification and recognition of Peking opera facial images based on SIFT features and support vector machine. *Second International Conference on Electromechanical Control Technology and Transportation, ICECTT 2017*. Zhuhai, China.

Received: June 05, 2019
Reviewed: June 25, 2019
Finally Accepted: July 17, 2019