Art Portrait Design Based on Mobile Internet in Digital Media

Ran Zhang[1]

Abstract

With the improvement of people's living standards and art appreciation, more and more people are engaged in the design and collection of art portrait. At the same time, the development of new generation information technology, such as digital media and mobile internet, has also made art portrait design techniques more diverse and more rich artistic styles. However, facing a huge number of art portrait, the current management work is mainly carried out manually by professionals, with high people and financial costs. Therefore, it is of great significance to study the effective and correct design and management of art portrait and apply digital mobile technology to provide users with more accurate art portrait. For the demand of art portrait design, this paper takes print, oil painting, ink painting and watercolor painting as the research object, and proposes a deep learning model Mask R-CNN. Based on the multi-disciplinary fields such as deep neural network, digital media and art, this paper analyzes and studies an art portrait design system based on mobile internet. In the paper, the in-depth learning model Mask R-CNN is applied in the system. By comparing the accuracy of the training of Mask R-CNN and U-Net models, this paper analyzes the effectiveness of the two models in extracting art portrait features. In addition, Mask R-CNN model is applied to the art portrait design system based on mobile internet. The final experimental results show that the art portrait design system using Mask R-CNN model has higher prediction and detection accuracy and has better practicability for art portrait design and art communication.

Key Words: Art Portrait Design, Digital Media, Mask R-CNN Model, Mobile Internet.

I. INTRODUCTION

Digital media technology studies the theories, methods and technologies related to digital information processing and storage [1]. At present, the iterative updating of mobile internet and cloud information technology provides the greatest convenience for resource sharing and communication interaction in time and space, which is a powerful guarantee for the success of art communication. It can broaden the breadth of learning, promote the development of learners' high-level cognitive processing and high-level thinking. It helps them to achieve deep understanding, application, migration and recreation of art, and bring diversity and convenience to learners [2-3].

The establishment and development of mobile internet has played an effective and good supporting role for art [4]. Especially for art portrait, mobile internet breaks the limitations of traditional art portrait, enriches the communication mode of traditional art portrait, and can effectively help people to allocate time reasonably. It brings great convenience for the improvement of people's artistic ideas. We will encourage people to learn about art and culture and appreciate art portrait in their spare time, so as to promote artistic ideas more efficiently. We let more people understand and inherit Chinese art, and promote the development of the art system.

Art portrait plays an important role in the existing types of portraits. The use of lines increases the semantic information and visual information of art portrait. Artists express their spiritual feelings through artistic creation, which will also attract appreciation and artistic exchange of people who resonate with the artistic conception of paintings [5-6]. The rapid development of art portrait types comes from the development of network digital media technology. But with the increasing number of art portrait, how to effectively present and design massive art portrait is an urgent problem to be solved. On the one hand, early portrait design technology mainly relied on manual portrait annotation and feature extraction [7]. However, in the face of massive portrait data, the traditional manual feature extraction method may have some problems, such as annotation errors and insufficiently objective annotation, and it also costs much manpower and material resources. On the other hand, only a small number of professionals have the ability to judge the types of art. Most of them lack the professional knowledge and skills to design art portrait [8]. Sometimes it is difficult to avoid confusing works of different artistic styles in the design of art portrait. Moreover, compared with ordinary photographic

images, the painted art portrait has its inherent attribute characteristics in terms of emotional expression, shape, color and texture. Using traditional methods to extract features from the art portrait, the extracted features cannot effectively express the style feature information of the entire art portrait. The existing research work on the management of art portrait mainly focuses on the identification of true and false art portrait based on art portrait themes and expression techniques, and according to the creative style of artistic painters. However, there is less research on the design of art portrait based on mobile internet in digital media.

Art portrait design integrates a variety of digital media technologies. Based on the current powerful computer hardware system and other mobile internet platforms, it has the advantages of good complementarity, high recognition, reconstruction ability and high measurement dimension. In digital media, the general features and local details of art portrait is mostly learned and extracted by deep learning technology [9-10]. It improves the accuracy of analysis and presentation of art portrait, helps people have the basic ability to analyze different art portrait and declines the professional requirements of personnel who extract and sort art portrait. This is also of great significance to the subsequent research work of art portrait design. In addition, equipment security in digital media is also a factor to be considered [11].

This paper takes print, oil painting, ink painting and watercolor painting as the research object, and devises an art portrait design system based on mobile internet using the in-depth learning model Mask R-CNN. The main work and structure of this paper are as follows:

(1) First, it introduces the background and significance of art portrait research in digital media, as well as the research status of art portrait design. Then, this paper obtains four types of art portrait data, and conducts data preprocessing on art portrait. The data composition and characteristics of art portrait are preliminarily analyzed.

(2) The in-depth learning model Mask R-CNN and U-Net are described in detail. The experiment analyzes and compares the effects of the model through accuracy and other indicators. We have good advantages in applying the in-depth learning model Mask R-CNN to art portrait design system based on mobile internet.

(3) The design and implementation of an art portrait design system based on mobile internet is carried out. The design and implementation of the art portrait system are mainly combined with the model studied in this paper. It also introduces the framework and components of the system, as well as the functions that the system can realize. Finally, the operation effect of the art portrait design system is shown.

The rest of this paper is composed of four parts. The second part is the literature related to our work in this paper.

The third part introduces the characteristics of artistic data, deep learning model and design system in detail. The fourth part analyzes the proposed model and art portrait design system through experiments. Finally, we summarize the main research contents and conclusions of this paper.

II. LITERATURE REVIEW

The integration of mobile internet platform and art system is the only way under the current trend of "Internet plus art" [12], and is also one of the effective ways to improve the proportion of high-quality content of mobile internet platform. The integration of the two is conducive to the dissemination of the mainstream cultural content and artistic aesthetic concepts in line with the youth groups. In the process of artistic communication, reasonable and effective use of internet and mobile internet platform resources can effectively expand the coverage of young audiences [13]. It forms a dual scientific and art communication mode that conforms to the laws of modern art communication from content selection to audience acceptance, from subjective evaluation to objective evaluation. This has changed the previously relatively closed communication mode of excellent content, and highlighted the value of high-quality content in art. Continuous high-quality content input will gradually improve the cultural ecology, aesthetic ecology and mainstream value guidance ecology of the mobile internet platform, and build a composite mobile internet platform with the ability to spread mainstream culture and multiple subcultures.

Art portrait can be divided into fine brushwork painting and freehand brushwork painting according to the expression methods of painting [14]. Fine brushwork in art portrait pays attention to drawing the outline of things first, the handwriting is clear, the strength of writing is emphasized, and then the color is filled. The overall picture is fine and rigorous, and the line outline is clear, meticulous and even. Freehand brushwork is different from meticulous brushwork. It uses bold and bold brushwork, pays attention to the form of god, and highlights the artist's subjective emotional expression of things. Jiang et al. [15] used low-level features such as color feature information, autocorrelation texture features, and edge size histograms to classify traditional Chinese fine brushwork and freehand brushwork. However, the extracted low-level features were common features of portraits, and could not represent the unique attributes of fine brushwork and freehand brushwork. The traditional portrait was composed of the main body of the painting, the signature seal and the inscription. Bao2 et al. [16] located the inscription position according to the overall layout characteristics of the art portrait, as well as the color and structure of the inscription part, and extracted the in-

scription information in the painting.

The advantages of deep learning in such tasks as feature selection and portrait analysis provide new choices for art. The research on the stylization of portrait art by using machine learning convolutional neural network has attracted extensive attention and has also been applied to a certain extent. Lexey Moiseenkoy created Prisma portrait stylized application software based on deep learning convolutional neural network. But this software was limited by specific kinds of filters with different artistic styles. Sergey Morugin created Ostagram portrait stylization software based on the deep dream algorithm of deep learning. It is no longer limited to the given filter template, and can identify the content of any two pictures, and transfer the artistic style of one picture to the other. Gatys et al. [17] took the lead in proposing to use the deep learning convolution neural network method to extract the texture features of art portrait style, so as to realize the stylization of art portrait in graphic design. However, due to the relatively complex algorithm of deep learning neural network, it requires a large amount of memory resources and a long running time. In order to bring users a good user experience, it is necessary to run the processing on the background server at the same time to meet the real-time requirements of portrait stylization processing speed, which makes the application of deep learning algorithm in graphic design portrait stylization subject to certain restrictions on the mobile terminal [18]. Recently, some scholars have proposed a large number of art portrait multi style methods based on generation confrontation network [19]. Another researcher has also proposed an algorithm to transfer the artistic style of multiple portraits to a common portrait [20]. The improvement of these algorithms has also improved the speed of stylization of art portrait. Pathak et al. [21] combined the coding and decoding architecture and the generation countermeasure network technology to complete the picture information through the judgment of the prediction map. In the case of a large area of information defect, the picture can also be filled, and the completed picture is semantically consistent with the original picture. Bertalmio et al. [22] adopted the idea of partial differentiation to complete and repair the portrait, and used the information outside the area to be completed to repair inward along the contour.

However, there is little research on the methods of feature extraction for various types of artistic style portrait in the above research. So based on the characteristics of the bottom layer of art portrait, it is not good to distinguish the style characteristics among various types of art portrait [23-24]. Facing a huge number of art portrait, feature extraction of art portrait by hand is subjective and requires a lot of manpower and material resources [25-26]. According to the style characteristics of art portrait such as print, ink painting

and oil painting, this paper uses a depth convolution neural network method that can extract the style and details of portrait features. The Mask R-CNN model used in this paper is an integrated learning method. According to the comparison between the method prediction results and U-Net prediction results, the method is superior to U-Net in the accuracy of the features extracted from art portrait and network performance. The art portrait design system based on mobile internet according to this model can effectively analyze the portrait, and has a high practical significance for the design and dissemination improvement of art portrait.

III. ART PORTRAIT DESIGN BASED ON MOBILE INTERNET

The development of digital media technology led to the discovery of portraits of different artistic styles by the general public [27]. A large number of mixed artistic creation styles of different artistic creation technologies come from the new creativity generated by artists from different styles of art portrait. In the era of full development of mobile terminal applications, the integration of users with mobile terminals and networks is more in-depth. Massive consumption data and behavior data are generated by mobile terminal users and network users. The larger data scale provides more optional features for the design of art portrait. By establishing an accurate analysis model of art portrait, better traffic conversion rate can be achieved.

Mobile internet technology can enable more people to learn and appreciate art with the help of mobile terminal devices (such as smartphones and tablets) and wireless networks, providing great convenience for online art communication. However, as an art form that is constantly being innovated and improved by people, the variety of styles and the richness of content increase the complexity of the characteristics of art portrait [28]. Therefore, this paper applies a Mask R-CNN depth learning method to feature extraction, and establishes an art portrait design system based on mobile internet.

3.1. Composition and Characteristics of Art Data

As the starting point of all training, it is necessary to organize and construct data sets. In terms of images, they should all be composed of art portrait. The current general image description data set mainly focuses on photos, which has a certain deviation from the goal of this article. In addition, in terms of data form, it is required that the form of data must be the coexistence of image and description to achieve one-to-one correspondence. In terms of the selection of source datasets, three aspects are focused on: artistry, content and consistency.

Artistry is not only a requirement for portraits, but also a requirement for images. This is the most important thing to support the establishment of the art portrait description data set, and it is also the fundamental difference from the ordinary data set. The artistic focus is on the expression of lines, strokes and colors [29]. In the collected album, many paintings are attached to the side of the writer's biography in an auxiliary form. Although such a portrait is also meaningful, it is difficult to complete the mapping of the author's life from the portrait itself, which requires a lot of other knowledge. In fact, it is observed that there is no artistic description composed solely of portrait content. In most cases, it is a combination of content description, author's description and era background. Consistency is a requirement for the overall data set. It is expected that the image, style and trend performance described by the data set will be consistent to protect the stability of data distribution. Stable portrait style also helps to reduce the portrait feature space and improve the accuracy of the model. To sum up, it is a synchronous requirement of quality and quantity.

Based on the results of multiple studies, we used web crawler technology to download art portrait from Dayi [30], Artlib World Art Appreciation Library website [31] and Baidu Search Pictures. Because you use keywords to search and download art portrait, the portrait obtained may be related to the keywords, but the artistic style does not necessarily belong to this feature category. Some portraits are too small in size or low in definition. In order to make the data more accurate and representative, the size of the objects that do not match the artistic style is less than 125×125 pixels, line texture and other seriously blurred art portrait was cleaned and screened, and 6,700 oil paintings, 7,724 ink paintings, 5,690 prints and 8,389 watercolor paintings were finally obtained.

In this paper, all the collected samples are carefully screened and data cleaned to make their presentation more reasonable and avoid noise interference caused by irrelevant images, so as to obtain accurately labeled art portrait data sets. At the same time of annotation, necessary pre-processing is performed on the image information, therefore, the existing art portrait data is enhanced in this paper. Data enhancement mainly refers to the process of adding image data by using specific methods to create deformed images that belong to the same category as the original image. In order to obtain multiple data images of this artistic style 299×299 pixels to extract high resolution art portrait with rich style information. This is because the style information of each art portrait is evenly distributed. In this way, the details of the portrait will be displayed more fully, so as to minimize the loss of local details and increase the image data. The specific algorithm is as follows:

To a portrait $X \in DS^{W \times H \times D}$, Where W represents the width of the portrait, H represents the height of the portrait, and D represents the portrait channel. Use 299 to round down the length and width values of the portrait to get the number of S_h segments with length of 299 in the upward direction of the rectangle, and the number of S_w segments with width of 299 in the wide direction. Finally, S art portrait with the size of $DS^{299 \times 299 \times D}$ is obtained from one image.

$$S_h = \left\lfloor \frac{h}{299} \right\rfloor. \tag{1}$$

$$S_w = \left\lfloor \frac{w}{299} \right\rfloor. \tag{2}$$

$$S = S_h \times S_w. \tag{3}$$

A slight change in the horizontal, for its overall artistic style, art portrait of different vertical degrees or sizes has little influence. So, this paper not only enhances the data of the portrait sample database, but also enhances the data of the training set during the training network. In the training process, let the training set rotate randomly within the range of 0° to 25°, and translate 0.02 times the length and width of the portrait in horizontal and vertical directions respectively.

3.2. Deep Learning Model

3.2.1. U-Net

Portrait design in digital media is to extract the required information from the image to be tested by computer and identify it. At present, the portrait recognition technology is becoming more and more mature, and has been widely used in all aspects of real life, which has important practical significance. Visual system is a process of multi-layer transmission from concrete to abstract. Low level features combine to form high level features. From low level to high level, features become more and more abstract, which is more and more able to express the original semantics of objects. In the portrait, pixels are the lowest level features, and objects themselves are the highest-level semantics. The higher the level of abstraction, the more accurate the results of brain judgment, and the fewer doubts. Deep learning simulates the visual center of the human brain. By building a multi-layer network, the original input signals are continuously feature extracted until the features available for the classifier are abstracted. The final output layer of the system has only a small amount of key information [32].

U-Net model is a network model proposed by Ronneberger et al. [33] for medical image segmentation in 2015. This method won several first prizes in the ISBI cell tracking competition in 2015. The U-Net network model is proposed based on the most basic deep full convolution network model. Fig. 1 shows the U-Net network structure [34].
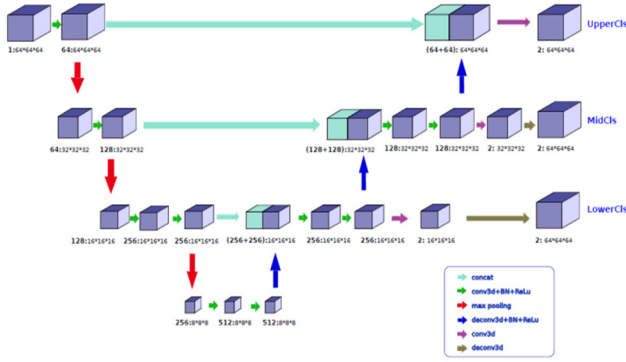
Fig. 1. U-Net network structure.

The network structure is symmetrical from left to right, forming a structure similar to the letter "U". The left side is a down sampled coding network. Convolution, pooling, and activation functions are used in the coding structure to extract image features. On the right is the up sampled decoding network. The decoding network restores images through repeated up sampling, convolution, and activation functions. The blue arrow in Fig. 1 indicates that the size of convolution kernel is 3×3 convolution operation and ReLU activation function. The red arrow represents maximum pooling, and the size is 2×2. After each down sampling, the number of filters in the convolution doubles. The green arrow represents that the convolution kernel size is 2×2, the feature channel will be reduced by half. The jump connection is represented by the gray arrow in the figure, which transmits the features extracted from the encoder structure to the decoder for feature fusion. The yellow arrow represents that the convolution kernel is 1×1. The U-Net model only needs a few times to make the network converge, and at the same time, the training data set used by the network is relatively small.

3.2.2. Mask R-CNN

In the development process of machine learning, single task network structure has become commonplace. Now more promising is the integrated and complex multi task network model, and Mask R-CNN is a typical representative. The Mask R-CNN paper was published by the team of He Kaiming in 2017, and obtained the best paper of ICCV 2017, which is one of the important achievements in the field of machine learning computer vision [35].

The network structure of Mask R-CNN consists of two parts. One part is used by Backbone to extract features, and the other part is used by Head to classify, box regression and mask prediction for each ROI.So two architectures are proposed to generate corresponding masks, namely, the left and right are faster R-CNN/ResNet and R-CNN/FPN respectively, as shown in Fig. 2 [36]:

For the building on the left, Backbone uses the pre trained ResNet, which is the fourth last layer of ResNet. The
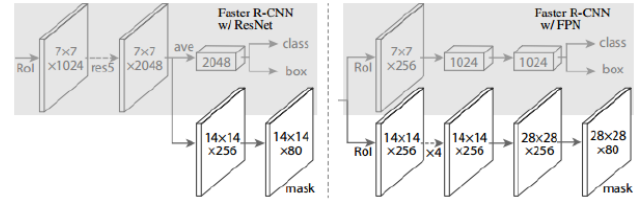

Fig. 2. Mask R-CNN two architectures.

ROI entered first gets 7×7×1,024 ROI feature, and then upgrade it to 2,048 channels, and then divide it into two subfields. The upper and lower branches are respectively responsible for classifying regression and generating corresponding masks. Due to the previous multiple convolutions and pooling, the corresponding resolution is reduced. The mask branch starts to use deconvolution to improve the resolution. At the same time, the number of channels is reduced to 14×14×256, finally output 14×14×80 mask template.

On the right the backbone used by the architecture is the FPN network. By inputting a single scale image, you can finally get the corresponding feature pyramid. It has been proved that the network can improve the detection accuracy to a certain extent, and many current methods have used it. Because the FPN network already contains res5, which can be used more efficiently, fewer filters are used here. The architecture is also divided into two branches. Although it has the same function as the former, the classification branch and mask branch are very different. FPN network uses fewer filters in classification, and may obtain a lot of useful information about features of different scales. In the mask subfield, the manoeuvre operation is performed for many times. First, the ROI is changed to 14×14×256 feature, perform the same operation four times, then perform the deconvolution operation, and finally output 28×28×80. This architecture outputs a larger mask than the former, and more detailed masks can be obtained. The loss function of each ROI of Mask R-CNN is as follows:

$$L = L_{cls} + L_{box} + L_{mask}. \qquad (4)$$

L_{cls} and L_{box} are the same as those defined in Faster R-CNN. For each ROI, mask subfield has $K×m×m$ dimension output, and its size for K is $m×m$ mask, each mask has K categories. Mask R-CNN uses a per pixel sigmoid, and defines L_{mask} as the average binary cross entropy loss.

3.3. Design System Based on Mobile Internet

The art portrait design system based on mobile internet consists of two main parts: server and client. The wireless mobile internet system follows the regulations of IETF IMPP on instant messaging and online status exchange, and expands some additional services for it. In addition, it also

provides seamless connection between different wireless devices and wired devices, as well as interoperability with existing major devices. Between the clients and the servers, it uses TCP/IP based protocols for secure communication [37]. The system adopts the B/S (Browser/Server) structure, and realizes the feature extraction and design of art portrait through the three-tier architecture of mobile layer, logical layer and data layer. The specific structure of the system is shown in Fig. 3.

The mobile layer is the interaction interface between mobile internet users and the system, and also the main business processing interface of the system. It mainly includes: data upload, result query, user error correction and data management. This paper presents the design results of art portrait to mobile internet users in a visual way. Mobile internet users can upload art portrait through the mobile layer, query the recognition results of art portrait, design the recognition resultsAt the same time, they can also view the help information and data management in the mobile layer, realizing the design of art portrait.

The logic layer is the main processing module of the function and business logic of the art portrait recognition and design system. Its main functions include art portrait processing and recognition. The art portrait processing receives the art portrait information uploaded by mobile internet users. After image enhancement, drying and morphological processing, it extracts the art portrait style features, stores the style feature information in the database, and sends the preliminarily processed portrait to the deep learning model for recognition after standardized processing. And return the identification result to the Web server.

The data layer is located on the server and database side. The data layer stores user data, deep learning model data

Fig. 3. Overall structure of art portrait design system based on mobile internet.

and art portrait processing data, providing important data support for art portrait recognition. Processing data of art portrait includes photos stored in the process of art portrait design, information data of art portrait and photos uploaded by users. The administrator can regularly send the designed art portrait to the model for training to further improve the recognition accuracy of the model.

Under the background of big data and mobile internet, art portrait technology has evolved rapidly, and its effectiveness, accuracy and real-time performance have been significantly improved, which has also made the technology fully applied in various fields of society. The system abstracts each functional module into an independent component, and realizes the impact of each module on the system through configuration files. The logic and code separation between the core components not only solves the problem of code confusion, but also provides convenience for subsequent system maintenance.

IV. EXPERIMENT AND RESULT ANALYSIS

4.1. Model Evaluation Index

In order to evaluate the feature extraction effect of art portrait after online training, the following three evaluation indicators are selected: specificity, $F1$ value and accuracy. S stands for specificity, A stands for accuracy, TP is the real sample, which stands for the number of positive classes successfully predicted to be positive, FP is the false positive sample, which stands for the number of negative classes incorrectly predicted to be positive, FN is the false negative sample, which stands for the number of positive classes incorrectly predicted to be negative, and TN stands for the number of negative classes successfully predicted to be negative. The corresponding expression is as follows:

$$S = \frac{TN}{TN+FP}. \tag{5}$$

$$F1 = \frac{2TP}{2TP+FP+FN}. \tag{6}$$

$$A = \frac{TP+TN}{TP+TN+FP+FN} \times 100\%. \tag{7}$$

In order to solve the problem of gradient disappearance in the training process of the model, and to use time for stable training of the model, this paper selects the Binary Cross Entropy (BCE) loss function as the target loss function to test during network training, and updates and optimizes the network. The binary cross entropy loss function is defined as follows:

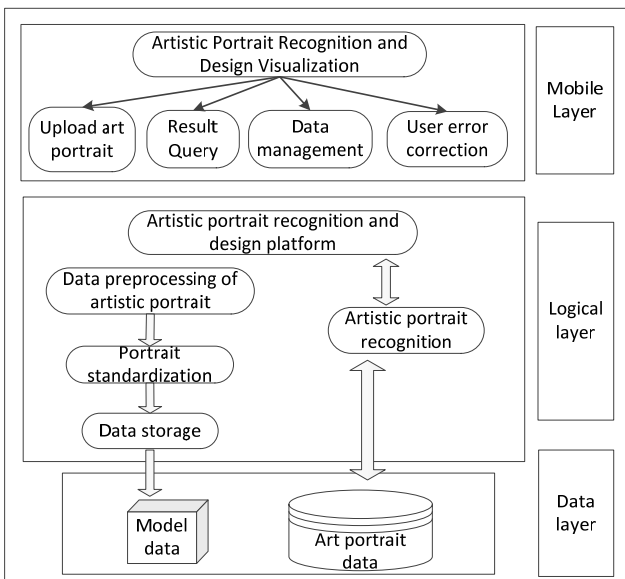$$Loss = -\sum_{i=1}^{N}\sum_{j=1}^{L}[(g_{ij}logp_{ij}) + (1-g_{ij})(1-logp_{ij})]. \tag{8}$$

In the formula, g_{ij} represents the real category manually marked by the portrait, and p_{ij} represents the predicted result value after model feature extraction.

4.2. System Experiment Results

In order to compare the feature extraction performance of the two models proposed in this paper, the depth learning model U-Net and Mask R-CNN were compared. On the data sets of the above four kinds of art portrait (oil painting, ink painting, print, and watercolor painting), the specificity, $F1$ value, and accuracy changes on the data sets were obtained. The experimental results recorded with MATLAB are shown in Fig. 4−Fig. 6, respectively.

From the comparison of deep learning models in Figure 4 that the super robust performance of the Mask R-CNN model is good during the training process, and the specificity of the Mask R-CNN model is higher than that of the U-Net model in the four types of art portrait, which means that the prediction results of the Mask R-CNN model are close to the actual situation.

From the comparison results of the deep learning models in Fig. 5 that the Mask R-CNN model has a good performance on the $F1$ value in the test data set. $F1$ in the training set is at a higher position, and the feature extraction results of the art portrait are more accurate.

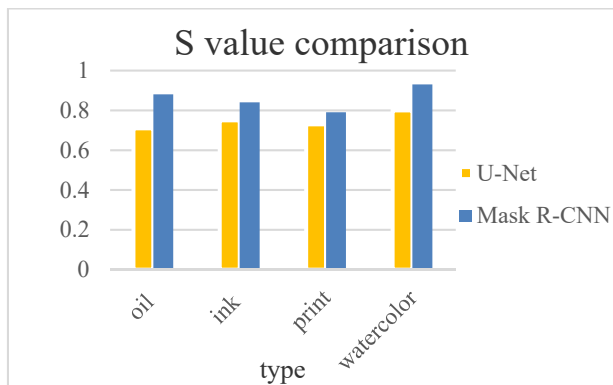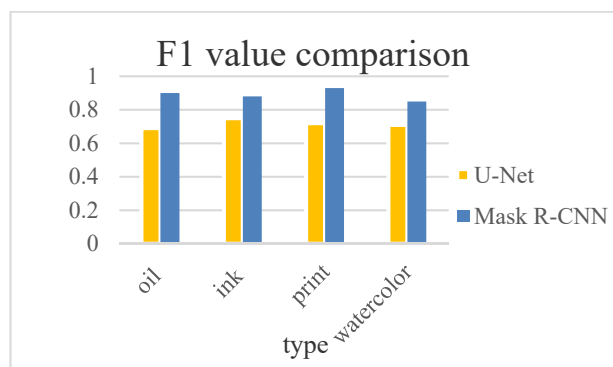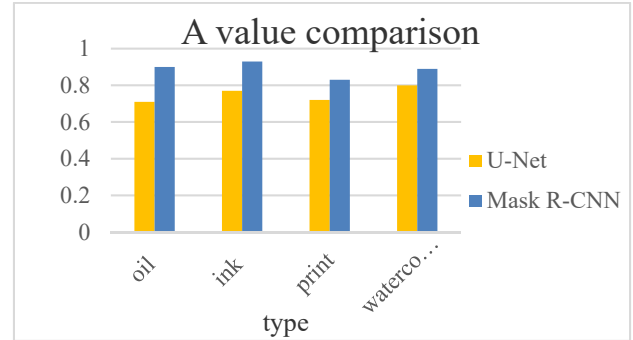Compared with U-Net, the Mask R-CNN model proves

Fig. 6. Change of accuracy rate.

its superiority. In Fig. 6, we can see that the accuracy of the Mask R-CNN model is at a higher level in the training set. In oil painting feature extraction, Mask R-CNN model has an accuracy rate of 26.76% higher than U-Net. Therefore, Mask R-CNN model has advantages in accuracy rate, strong feature extraction ability, and can reasonably design art portrait.

In addition, the experiment tested the art portrait design system based on mobile internet. Fig. 7 shows the loss value when applying Mask R-CNN model and U-Net model for dataset training under different epochs.

From the comparison of the model application in Fig. 7 that under the mobile connected art portrait system, the loss rate of the Mask R-CNN model decreases steadily with the increase of the sample set, and the data enhancement has a promoting effect on the loss rate. Compared with the U-Net model, it can speed up the convergence speed while shortening the training time.

The detection effect of U-Net and Mask R-CNN models on art portraits is shown in Fig. 8. From the figure, it can be seen that the right side has achieved better effect on the recognition of art portraits, can easily achieve the detection of multiple objects, and can use boxes to frame the positions of different types of objects in the figure. In the comparison figure, the left side is the detection effect of U-Net model on art portraits, and the right side is the detection effect of Mask R-CNN model on art portraits. The boxes with differ-

Fig. 4. Specific changes.

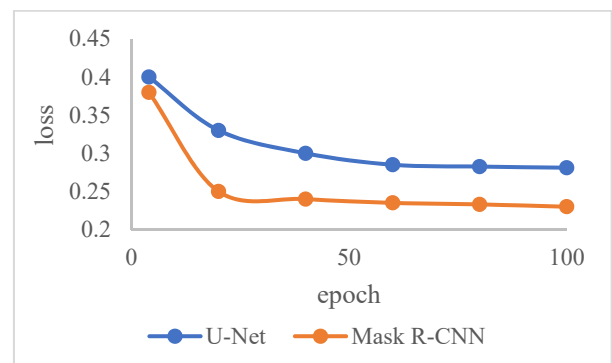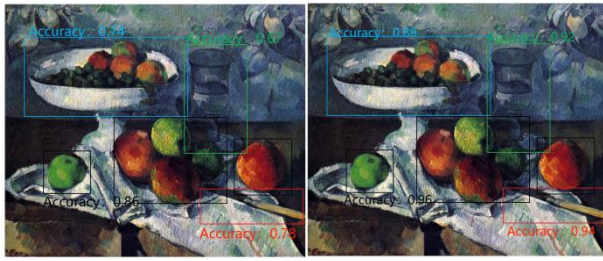Fig. 5. *F*1 value change.

Fig. 7. Loss value of different epochs.

(a) Oil

(b) Watercolor

(c) Ink painting and print

Fig. 8. The accuracy of different epochs.

ent colors in the figure represent different portrait objects, and the accuracy of the objects is indicated in the boxes.

Through experiments, it is found that the U-Net and Mask R-CNN models perform well on the whole, but the accuracy is different. In the recognition of watercolor painting and ink painting, the U-Net model has error detection and does not correctly recognize the object type, while the Mask R-CNN model correctly recognizes the art portrait. Later, the Mask R-CNN model can be further studied and used in practical applications.

To sum up, the Mask R-CNN model is superior to U-Net in specificity, $F1$ value and accuracy, and has verified that the Mask R-CNN model has good generalization ability in the art portrait design system based on mobile internet. The Mask R-CNN model extracts the feature information of art portrait more clearly and accurately, which has practical significance for the design of art portrait.

V. CONCLUSION

With the increasing variety and quantity of art portrait, the traditional design method of art portrait can no longer meet the needs of efficient management of art portrait. So the design of art portrait requires higher professional knowledge and skills of personnel. With the development of deep learning, convolutional neural network has been widely used in the task of art portrait design because of its good feature extraction ability for images. This paper first outlines the background and significance of art portrait design in the current research, and the main research object is the art portrait in digital media. In order to design an art portrait design system based on mobile internet, we should not only consider from the algorithm level, but also help improve the performance of the model from other aspects. Computational ability, feature extraction algorithm and experimental data support the progress of technology together. These three aspects are indispensable and depend on each other. Therefore, the text launches experiments in relevant research fields from the model level, and summarizes the experimental results. In order to deeply explore the relationship between the art portrait data and accurately design the art portrait, this paper has built an art portrait design system based on mobile internet. First, the data of the art portrait has been preprocessed. Second, the two in-depth learning models, Mask R-CNN and U-Net, have been compared and analyzed. The data set has been used for comparative experiments, and the accuracy, $F1$ value, specificity and other indicators have been obtained through method training. Each index of Mask R-CNN model is superior to U-Net, which proves the superiority of Mask R-CNN and the rationality of design. It is applied in the design system. The loss value shows that Mask R-CNN model can better analysis of detailed features of the art portrait, and can realize the design of art portrait.

REFERENCES

[1] M. Chen, "Design of animation professional interactive training platform based on digital media information flow fusion algorithm," in *2022 6th International Conference on Intelligent Computing and Control Systems (ICICCS), IEEE*, May 2022, pp.1805-1808.

[2] R. Chen "The design and application of college english-aided teaching system based on web," *Mobile Information Systems*, vol. 2022, 2022.

[3] Y. Zeng and M. Luo. "Data sharing and online political education based on edge computing network optimization," *Mobile Information Systems*, vol. 2022, 2022.

[4] S. Hou and J. Ahn "Design and empirical study of an online education platform based on B2B2C, focusing on the perspective of art education," *KSII Transactions on Internet and Information Systems (TIIS)*, vol. 16, no. 2, pp. 726-741, 2022.

[5] Y. Chen, "Development and implementation strategy of students' concept innovation in Chinese painting teaching based on big data," *Journal of Physics: Conference Series*, vol. 1992, no. 4, p. 042063, 2021.

[6] F. Chen, "Analysis of the characteristics of art intangible cultural heritage in cross-cultural communication," *Art and Design Review*, vol. 10, no. 3, pp. 389-396, 2022.

[7] N. van Noord, "A survey of computational methods for iconic image analysis," *Digital Scholarship in the Humanities*, vol. 37, no. 4, pp. 1316-1338, 2022.

[8] V. Shunkov, O. Shevtsova, V. Koval, T. Grygorenko, L. Yefymenko, and Y. Smolianko, et al., "Prospective directions of using multimedia technologies in the training of future specialists," 2022.

[9] T. T. Nguyen, C. M. Nguyen, D. T. Nguyen, D. T. Nguyen, T. Huynh-The, and S. Nahavandi, et al., "Deep learning for deepfakes creation and detection," arXiv preprint arXiv:1909.11573, vol. 223, p. 103525, 2022.

[10] I. Santos, L. Castro, N. Rodriguez-Fernandez, A. Torrente-Patino, and A. Carballal, "Artificial neural networks and deep learning in the visual arts: A review," *Neural Computing and Applications*, vol. 33, no. 1, pp. 121-157, 2021.

[11] J. Miao, Z. Wang, X. Miao, and L. Xing, "A secure and efficient lightweight vehicle group Authentication protocol in 5G networks," *Wireless Communications and Mobile Computing 2021,* pp. 1-12, 2021.

[12] H. Zhou, *Characteristics of User Experience in Art E-Commerce: Case "buybuy Art",* 2018.

[13] X. Kong, X. Liu, B. Jedari, M. Li, L. Wan, and F. Xia, "Mobile crowdsourcing in smart cities: Technologies, applications, and future challenges," *IEEE Internet of Things Journal,* vol. 6, no. 5, pp. 8095-8113, 2019.

[14] J. Sheng, C. Song, J. Wang, and Y. Han, "Convolutional neural network style transfer towards Chinese paintings," *IEEE Access*, vol. 7, pp. 163719-163728, 2019.

[15] W. Jiang, Z. Wang, J. S. Jin, Y. Han, and M. Sun, "DCT–CNN-based classification method for the Gongbi and Xieyi techniques of Chinese ink-wash paintings," *Neurocomputing*, vol. 330, pp. 280-286, 2019.

[16] H. Bao, Y. Liang, H. Z. Liu, and D. Xu, "A novel algorithm for extraction of the scripts part in traditional Chinese painting images," in *2010 2nd International Conference on Software Technology and Engineering*, IEEE, 2010, vol. 2, pp. V2-26-V2-30.

[17] L. Gatys, A. S. Ecker, and M. Bethge, "Texture synthesis using convolutional neural networks," *Advances in Neural Information Processing Systems*, vol. 28, 2015.

[18] Z. Tian, "Dynamic visual communication image framing of graphic design in a virtual reality environment," *IEEE Access*, vol. 8, pp. 211091-211103, 2020.

[19] M. Z. Alom, T. M. Taha, C. Yakopcic, S. Westberg, P. Sidike, and M. S. Nasrin, et al., "A state-of-the-art survey on deep learning theory and architectures,"

[20] A. Elgammal, "AI is blurring the definition of artist: Advanced algorithms are using machine learning to create art autonomously," *American Scientist*, vol. 107, no. 1, pp. 18-22, 2019.

[21] B. Ma, X. An, and N. Sun, "Face image inpainting algorithm via progressive generation network," in *2020 IEEE 5th International Conference on Signal and Image Processing (ICSIP)*, IEEE, 2020, pp. 175-179.

[22] W. Huang, Y. Li, Z. Xu, and C. Huang, "Image inpainting by reducing edge blur and error accumulation," in *International Conference on Computer Graphics, Artificial Intelligence, and Data Processing (ICCAID 2021)*. SPIE, 2022, vol. 12168, pp. 18-24.

[23] G. Castellano and G. Vessio, "A deep learning approach to clustering visual arts," *International Journal of Computer Vision*, vol 130, no. 11, pp. 2590-2605, 2022.

[24] F. Milani and P. Fraternali, "A dataset and a convolutional model for iconography classification in paintings," *Journal on Computing and Cultural Heritage (JOCCH)*, vol. 14, no. 4, pp. 1-18, 2021.

[25] C. Zhang and Y. Lu, "Study on artificial intelligence: The state of the art and future prospects," *Journal of Industrial Information Integration*, vol. 23, p. 100224, 2021.

[26] Y. Gong, "Application of virtual reality teaching method and artificial intelligence technology in digital media art creation," *Ecological Informatics*, vol. 63, p. 101304, 2021.

[27] Y. Ruan, "The cultural value validity of digital media art based on deep learning network model," *Advances in Multimedia*, vol. 2022, 2022.

[28] K. K. Fan and T. T. Feng "Sustainable development strategy of Chinese animation industry," *Sustainability*, vol. 13, no. 13, p. 7235, 2021.

[29] H. Kim, H. Y. Jhoo, E. Park, and S. Yoo, "Tag2pix: Line art colorization using text tag with secat and changing loss," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2019, pp. 9056-9065.

[30] C. Celik and H. S. Bilge, "Content based image retrieval with sparse representations and local feature descriptors: A comparative study," *Pattern Recognition*, vol. 68, pp. 1-13, 2017.

[31] Y. Lecun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436-444, 2015.

[32] J. Zhang, Z. Yin, P. Chen, and S. Nichele, "Emotion recognition using multi-modal data and machine learning techniques: A tutorial and review," *Information Fusion*, vol. 59, pp. 103-126, 2020.

[33] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmenta-

tion," *International Conference on Medical Image Computing and Computer-Assisted Intervention.* Cham, pp. 234-241, 2015.

[34] G. Zeng, X. Yang, J. Li, L. Yu, P. A. Heng, and G. Zheng, "3D U-net with multi-level deep supervision: Fully automatic segmentation of proximal femur in 3D MR images," *International Workshop on Machine Learning in Medical Imaging*, Cham: Springer, 2017, pp. 274-282.

[35] Z. Hu, W. Fang, T. Gou, W. Wu, J. Hu, and S. Zhou, et al., "A novel method based on a Mask R-CNN model for processing dPCR images," *Analytical Methods*, vol. 11, no. 27, pp. 3410-3418, 2019.

[36] Y. Liu, X. Yao, Z. Gu, Z. Zhou, X. Liu, and X. Chen, et al., "Study of the automatic recognition of landslides by using InSAR images and the improved mask R-CNN model in the Eastern Tibet Plateau," *Remote Sensing*, vol. 14, no. 14, p. 3362, 2022.

[37] J. Miao, Z. Wang, X. Ning, N. Xiao, W. Cai, and R. Liu, "Practical and secure multifactor authentication protocol for autonomous vehicles in 5G," *Software: Practice and Experience,* 2022.

AUTHOR

Ran Zhang received B.S. degree from the China Academy of Art in 2006 and M.S. degree from the Henan Normal University in 2013. He is currently a senior lecture at Henan Institute of Technology. His research interests include visual communication, digital media, mobile Internet, etc.