Tech Science Press

Research on the Identification of Hand-Painted and Machine-Printed Thangka Using CBIR

Chunhua Pan[1,*], Yi Cao[2] and Jinglong Ren[3]

[1]key Laboratory of Artificial Intelligence Application Technology State Ethnic Affairs Commission, Qinghai Minzu University, Xining, 810007, China
[2]Department of Electrical and Computer Engineering, University of Windsor, Windsor, ON, N9B 3P4, Canada
[3]Qinghai Qianxun Information Technology Co. LTD, Xining, Qinghai, 810007, China
*Corresponding Author: Chunhua Pan. Email: chunhuapan123@163.com

Abstract: Thangka is a unique painting art form in Tibetan culture. As Thangka was awarded as the first batch of national intangible cultural heritage, it has been brought into focus. Unfortunately, illegal merchants sell fake Thangkas at high prices for profit. Therefore, identifying hand-painted Thangkas from machine-printed fake Thangkas is important for protecting national intangible cultural heritage. The paper uses Content-Based Image Retrieval (CBIR) techniques to analyze the color, shape, texture, and other characteristics of hand-painted and machine- printed Thangka images, in order to identify Thangkas. Based on the database collected and established by this project team, we use Local Binary Pattern (LBP) texture analysis combined with the color histogram of Hue,aturation, Value (HSV) space, scale invariance, K-Means clustering, perceptual Difference Hash (DHASH) and other algorithms to extract the color lines and texture features of Thangka images, in order to identify hand-painted and machine-printed Thangkas. Three algorithms, LBP algorithm and HSV algorithm and DHASH algorithm, are compared, and the experimental results show that the color histogram algorithm based on HSV space is efficient. This algorithm can be applied broadly to retrieve and identify hand-painted Thangkas and help protect this precious intangible cultural heritage.

Keywords: CBIR; hand-painted; machine-printed Thangka; LBP; HSV

1 Introduction

In Qinghai, many ethnic groups live in compact communities on the 720,000 square kilometers of the vast and mysterious Qinghai-Tibet Plateau. More than 20 ethnic groups, including Han, Tibetan, and Hui, have lived and multiplied here for generations. The people have created their history and culture, formed and maintained their unique, colorful customs and art forms in their production and daily life for a long time.

Thangka is a special artistic way of painting in Tibetan traditional culture. In 2006, Thangka was approved by The State Council and the Ministry of Culture as the first batch of national intangible cultural heritage under protection [1], attracting more and more art lovers and collectors' attention. At the

same time, a lot of printed Thangkas appeared on the market, and some businessmen sold printed Thangkas as hand-painted Thangkas at a high price. Since currently we don't have many effective techniques to identify the authenticity and quality of Thangkas [2], this ancient and exquisite craftsmanship is facing a huge challenging with the market full of faked Thangkas.

In the 21st century, benefiting from the Second Industry Civilization, we need to think about how to protect the traditional arts and crafts to make them inherit and spread. We should find out a right ascension path of progress for exquisite national crafts, which can go out of their own nation, move towards the world, and make the national culture furtherly carry forward.

2 The Cultural Value Differences between Hand-Painted Thangka and Printed Thangka

The theme of Thangka involves the history, politics, culture, and social life of the Tibetan people, with distinct ethnic characteristics, strong religious color, and unique artistic style. It depicts the world of the Holy Buddha with bright colors. Thangka pigments are traditionally made of precious mineral gems such as gold, silver, pearl, agate, coral, turquoise, malachite, and cinnabar, as well as plant pigments such as saffron, rhubarb, and indigo, which also show their sanctity. These natural raw materials make the color of Thangka paintings bright and dazzling even after hundreds of years. Therefore, Thangka is regarded as the treasure of Chinese national painting art, the "encyclopedia" of Tibetan people, and the precious intangible cultural heritage of Chinese folk art [3].

The drawing of traditional Thangka has strict requirements and extremely complex procedures, which must be carried out by the rituals in the scriptures and the requirements of the master. These procedures include pre-drawing ceremony, canvas making, composition crafting, dyeing, marking and shaping, laying gold and tracing silver, opening eyes, sewing, and mounting light, etc [4,5]. It usually takes half a year or more than a year to draw a Thangka painting by hand.

The mystery of Thangka comes from the special rules while drawing. The drawing of Thangka has high requirements for the artist, including purifying oneself before drawing, chanting sutras, and thinking. The completed Thangka must be illuminated and blessed by lamas. Therefore, every hand-painted Thangka embodies the blood and sweat of the painter and the protection of the gods. Therefore, the hand-painted Thangka is regarded as a living Buddha. However, the machine-printed fake Thangkas lose all the doctrinal connotation and the meaning of life vitality with only electricity and machines.

In this paper, we will use three classification retrieval methods: LBP texture based on color HSV histogram, scale invariance based on K-Means clustering method, CBIR based on the perceptual hash value. With some experiments, the retrieval efficiency of each algorithm is compared to find out a suitable algorithm for identifying Thangka images.

3 Algorithms to Distinguish Printed Thangka from Hand-Painted Thangka

3.1 Color Histogram Algorithm Based on Texture LBP and HSV

The canvas of a hand-painted Thangka is usually made of specially made cotton, hemp, or silk materials. The cotton is formed by manual grinding with bovine glue and lime. When paint on the cloth, we need to use large white powder as coatings to polish repeatedly. In this way, the canvas can reflect the variety of pigments depicting the design more evenly. At the same time, the polished canvas is insect-proof and corrosion resistant.

The printed fake Thangka usually use a custom-made plastic cloth. Most of the pigments used are polymeric pigments synthesized by human beings.

The pigments are very bright in color and have a large adhesion force, which can be applied to multiple intermediate carriers. The machine has great versatility in all kinds of painting techniques, but there are some deficiencies in cotton painting. It dries fast, has water resistance. The dry pigment film feels tough and stiff. It has no flexibility, cracking, and even slag. We can see differences between these two canvases under the magnifying glass. We use the same equipment and sample in green to avoid color difference. The result of the printed Thangka is shown in Fig. 1, and the hand-painted Thangka in Fig. 2. We can see that there are obvious small dots and regular textures on the printed Thangka sample, but none on the hand-painted Thangka sample.

Figure 1: Printed Thangka under the maximum magnification of Honor 8

Figure 2: Hand-Painted Thangka under the maximum magnification of Honor 8

It can be seen from the effect drawing that there are obvious small dots and regular textures on the printed Thangka, not the hand-painted Thangka.

3.1.1 LBP (Local Binary Pattern)

LBP uses the characteristics of the local texture of the image to process the image [6]. In specific processing, the image is first processed with grayscale, and then 3 * 3 = 9 pixels are selected to take the gray value of the central pixel as the target, and the gray value of the adjacent 8 pixels is compared with the central target. If the gray value of the adjacent pixel is greater than that of the target, the pixel value of the adjacent point is set to 1; otherwise, it is set to 0.

These adjacent 8 pixels to the central target form a binary number of one byte. We use the binary of this byte to represent the LBP value of the central target, as shown in Fig. 3. The model of this algorithm is expressed by formula (1).

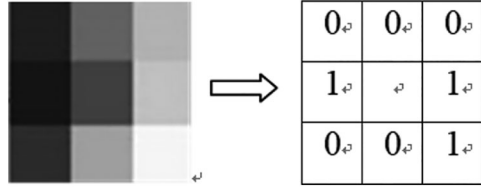$$LBP(x_c, y_c) = \sum_{p=0}^{p-1} 2^p s(i_p - i_c) \tag{1}$$

Figure 3: An example of central target LBP

In formula (1), (x_c, y_c) represents the central element, whose gray value is i_c, i_p represents the value of adjacent pixels. For sign function $s(x)$, if $x \geq 0$, then $s(x) = 1$, otherwise $s(x)$ is 0 [7].

In this way, the LBP texture feature is extracted for each pixel of the whole Thangka image. In order to distinguish hand-painted Thangka from printed Thangka, we need to measure image similarity with similarity calculation formula (2) based on image histogram, and Maximum Likelihood Estimation formula (3).

$$D(S, M) = \sum_i \min(S_i, M_i) \tag{2}$$

$$L(S, M) = -\sum_i S_i \log M_i \tag{3}$$

In this paper, a color histogram is used for similarity comparison. The color histogram will be introduced in detail in Section 3.1.2 with LBP and similarity calculation of hand-painted and printed Thangkas, the image is unified to obtain the color histogram with LBP local texture features.

3.1.2 HSV Color Histogram

Besides texture features, the color features of Thangka images are also very intuitive. Therefore, color features are also one of the basic visual features of Thangka images. In this paper, an HSV color histogram is used to extract color features of hand-painted Thangka and machine-printed Thangka. HSV color space is a color model for visual perception [8,9]. The standard simulation model is represented by an inverted cone. The long axis represents brightness, the distance from the axis represents saturation, and the angle around the long axis represents hue. HSV standard simulation model is designed, as shown in Fig. 4.

In this paper, HSV feature extraction of the Thangka image was carried out according to the following process, as shown in Fig. 5.

3.2 Scale Invariance with SFIT and K-Means

3.2.1 Scale Invariance Algorithm: SFIT

Scale-Invariant Feature Transform (SIFT) is an image processing algorithm, which is used to detect and describe local features in images. It seeks the limit point on the spatial scale of distributed action and selects its specific action position, scale, and invariant rotation. This algorithm was published by David Lowe in 1999 [10,11].
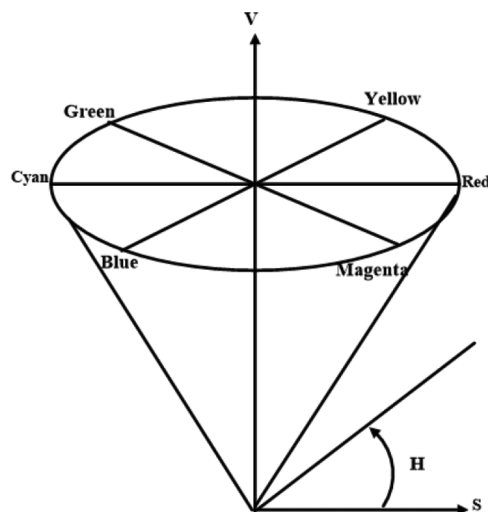
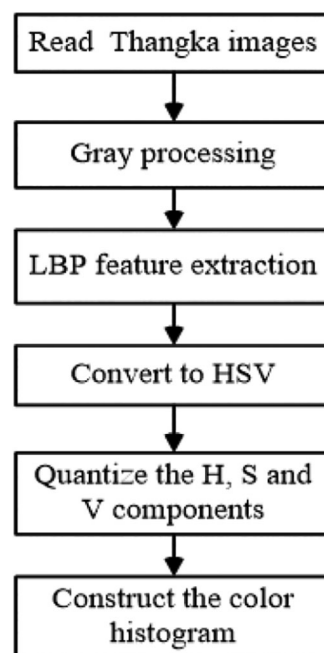**Figure 4:** Standard simulation model of HSV

Figure 5: Flow chart of HSV color feature extraction

SIFT algorithm steps are as follows:

(1) Limit value of test scale distribution action space: search scale distribution action space. A gauss differential mathematical function is used to identify potential concerns about scale invariance.

(2) Positioning key points: Location and scale are determined by a simulation standard simulation model for all candidate-specific action positions.

(3) Determine the direction: one or more distribution directions are allocated to the positions of all points, which is the partial one-sided direction of the image gradient distribution. Each subsequent manipulation and management of image data information is transformed relative to the distribution

direction, scale, and specific action position of the core points, thus providing invariance for the above transformation.

(4) Describe key points: Measure the local gradient of the image at the selected scale in the neighborhood around each key point. These gradients are transformed into a representation that allows for relatively large local shape deformation and illumination changes [12,13].

3.2.2 K-Means Method

K-Means is a clustering algorithm based on unsupervised learning. Aiming at the established standard sample set, k-means distinguishes the standard sample set into K-clusters according to the distance between standard samples. The points inside the cluster should be linked as closely as possible, and the distance between clusters should be as large as possible [14,15]. Suppose the partitioned cluster is $(C_1, C_2, \ldots C_i)$, so our goal is to minimize the squared error E, calculated by the formula (4). In (4), μ_i is the mean vector of cluster C_i, expressed as the formula (5).

$$E = \sum_{i=1}^{k} \sum_{x \in C_i} \|x - \mu_i\|_2^2 \tag{4}$$

$$\mu_i = \frac{1}{|C_i|} \sum_{x \in C_i} x \tag{5}$$

3.3 DHASH Method Based on Image Perception

The perceptual hash algorithm is the general name of a kind of hash algorithm. Its function is to generate the "fingerprint" string of each image and compare the fingerprint information of different images to judge the similarity of images [16–18]. The closer you get to the image, the more similar it becomes. Perceptual hashing algorithm includes aHash (mean hashing), pHash (perceptual hashing), and dHash (difference hashing). AHash is faster but less accurate; pHash, by contrast, is more accurate but slower; dHash takes advantage of two strengths to achieve high accuracy and speed. After the hash value is obtained, the hamming distance is used to quantify the similarity of the two images. The larger the hamming distance is, the smaller the image similarity is. Conversely, the smaller the Hamming distance is, the greater the image similarity is. In this paper, we will apply DHSH to extract fingerprint strings of hand-painted Thangka and printed Thangka.

The steps are as follows:

(1) Reduce the size: In order to remove the details of the image, we will retain only the basic information of the image, such as structure, light, and shade. The Thangka image is reduced to 8*8 with 64 pixels in total. In this way, the image differences brought by different sizes and proportions are eliminated.

(2) Color simplification: The image reduced in the first step is converted to a 64-level grayscale, that is, each pixel contains only 64 colors;

(3) Calculate the average value: Calculate the gray average value of 64 pixels.

(4) Compare the grayscale of pixels: Compare the grayscale of each pixel with the average value. If it is greater than or equal to the average value, it is marked as 1. If it is less than the average value, it is marked as 0.

(5) Calculate the hash value: Combine the comparison results of the previous step together to form a 64-bit integer, which is the fingerprint of this Thangka image. Once you have a fingerprint, you can compare different Thangka images to see how many of the 64 bits are different. Theoretically, this is equivalent to the Hamming distance [19]. If the different data bits do not exceed 5, it indicates that the two images are very similar. When it's greater than 10, it means there are two different images.

4 The Simulation Experiment

In this paper, the experimental setting is CPU Intel-I5 quad-core, 16 G memory, and 512 G solid state disk. "Python3.6 + Opencv3.4.2" is configured in the Ubuntu system environment to complete the simulation experiment of the above algorithms. The experiment steps is shown in the following flow chart (Fig. 6).
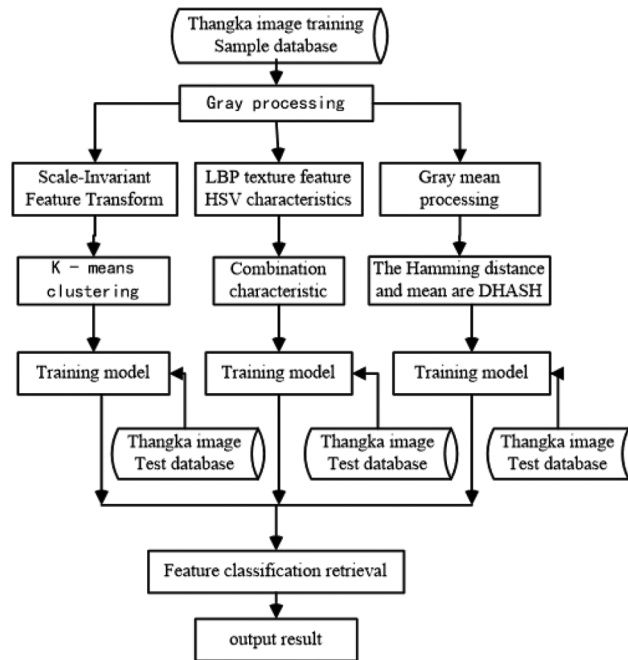

Figure 6: Flow chart of system experiment

4.1 The Experimental Results by Color Histogram HSV and LBP Texture

Firstly, LBP is used to extract texture features for the canvas of hand-painted Thangka and machine-printed Thangka. The texture analysis results of printed Thangka are shown in Fig. 7, and that of the hand-painted Thangka is shown in Fig. 8.

Figure 7: Results of printing Thangka texture points

From the experimental results, the texture distribution of printed Thangka is very uniform, while the texture distribution of hand-painted Thangka has no obvious rule.

Figure 8: Analysis results of hand-painted Thangka texture

HSV color histogram can help extract color features of hand-painted Thangka and machine-printed Thangka. The color histogram of printed Thangka images is shown in Fig. 9, and that of hand-painted Thangka images is shown in Fig. 10.
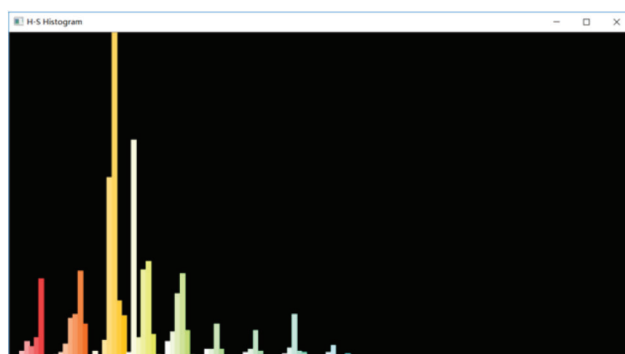

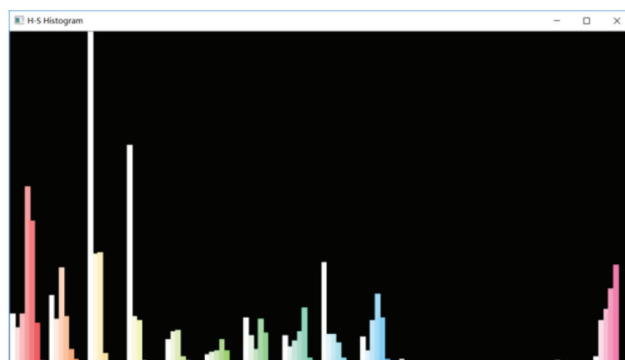**Figure 9:** Color histogram of printed Thangka images

Figure 10: Color histogram of hand-painted Thangka image

From the results, the color gamut of machine-printed Thangka images is not bright enough, while the color gamut of hand-painted Thangka images is richer, and their yellow color is relatively brighter.

4.2 The Retrieval Results by SIFT

The feature points extracted by SIFT algorithm are generally not fixed values [20], and the dimensions are unified 128 dimensions [21]. In Fig. 10, we can obtain the search result of machine-drawn Thangka with

invariant SIFT scale. And Fig. 11 is the search result of hand-drawn Thangka with invariant SIFT scale. In the picture, "search" is the image to be searched, and "require" is the image with similar features found. You can easily see the results of the search in Figs. 11 and 12.

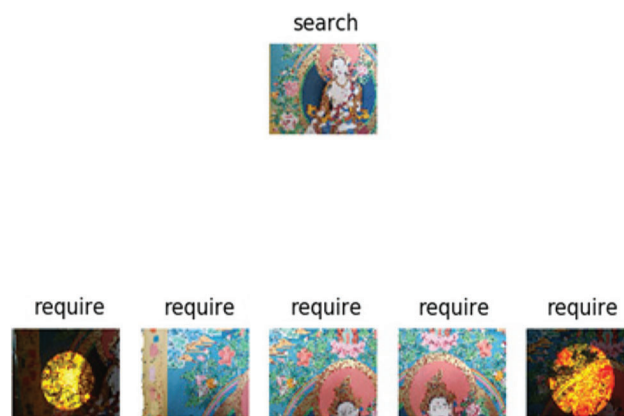**Figure 11:** Search results of printed Thangka with invariant SIFT

Figure 12: Search results of hand-painted Thangka with invariant SIFT

4.3 The Retrieval Results by DHASH

The retrieval steps are as follows:

(1) Load the image and calculate the difference of the hash value.

(2) Extract filename from path and update database.

(3) Add the file name to the value list using Hash as the dictionary key.

(4) The fingerprint of the Thangka image is extracted from the database and the corresponding hash database is established.

(5) Open the database, and then load the picture, calculate the fingerprint of the picture, find all the pictures with the same fingerprint. If any images have the same hash value, the graph images are iterated over and displayed. The retrieval results are shown in Fig. 13.
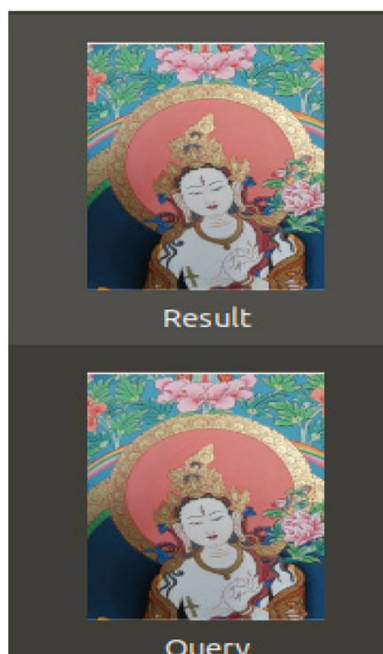
Figure 13: Retrieval results of image-aware DHASH algorithm

In total, there are 83 photos drawn in the image data set with 3968 * 2976 size. For the photo labeled 2, the searching time in this paper is 0.00023651123046875 seconds. For search, we use the VP-tree binary Tree method. It can be seen from the results that this method is effective and efficient.

5 Conclusion

This paper uses three methods to retrieve Thangka photos. The first method can retrieve two Thangkas with similarity from the image library based on LBP texture and HSV for multiple different regions. Basically, it can recognize images with relatively accurate results, and the recognition rate of the HSV color histogram for different regions is better. This method is of practical significance to the retrieval and recognition of hand-painted Thangka and machine-painted Thangka.

The second method uses SIFT feature points to mark pixels and forms feature clusters of Thangka images through K-Means clustering, which can effectively identify images. Compared with color histograms, this method is more suitable for identifying specific objects.

The third method applies the image perceptual hash algorithm to calculate the difference between two adjacent pixels in the Thangka image. Due to the property of the hash algorithm, even a very small difference of the input can result in big changes of the output. However, in our use case, we expect that similar inputs have similar outputs, so satisfied results cannot be achieved. Experiments show that this method has a large error in Thangka image recognition, but perceptual hash is a worthy field of image search because of its fast search time and a single hash value.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

[1] Y. L. Qu, "Research on the construction of Digital resources of Thangka archives from the perspective of Co-construction and Sharing," *Natural and Cultural Heritage Research*, vol. 4, no. 12, pp. 1–5, 2019.

[2] Q. Y. Wang and H. B. Dong, "Book retrieval method based on QR code and CBIR technology," *Journal on Artificial Intelligence*, vol. 1, no. 2, pp. 101–110, 2019.

[3] C. Cheng and D. Lin, "Based on compressed sensing of orthogonal matching pursuit algorithm image recovery," *Journal of Internet of Things*, vol. 2, no. 1, pp. 37–45, 2020.

[4] W. Sun, G. Z. Dai, X. R. Zhang, X. Z. He and X. Chen, "TBE-Net: A three-branch embedding network with part-aware ability and feature complementary learning for vehicle re-identification," *IEEE Transactions on Intelligent Transportation Systems*, pp. 1–13, 2021. https://doi.org/10.1109/TITS.2021.3130403.

[5] P. Xu and J. Zhang, "An expected patch log likelihood denoising method based on internal and external image similarity," *Journal of Internet of Things*, vol. 2, no. 1, pp. 13–21, 2020.

[6] X. R. Zhang, X. Sun, X. M. Sun, W. Sun and S. K. Jha, "Robust reversible audio watermarking scheme for telemedicine and privacy protection," *Computers, Materials & Continua*, vol. 71, no. 2, pp. 3035–3050, 2022.

[7] K. T. Jia, "Brilliant tibetan thangka handicraft: A case study of thangka feature film resource bank constructed by xizang library," *Tibetan Art Research*, vol. 31, no. 2, pp. 83–86, 2017.

[8] Y. S. Sun, X. J. Li and K. Zhao, "Surface state recognition based on improved LBP and HSV color histogram," *Bulletin of Surveying and Mapping*, vol. 65, no. 2, pp. 29–36, 2020.

[9] L. C. Jiang, G. Q. Shen and G. X. Zhang, "Image retrieval algorithm based on HSV block color histogram," *Mechanical & Electrical Engineering*, vol. 38, no. 11, pp. 54–58, 2009.

[10] C. Cheng and D. Lin, "Image reconstruction based on compressed sensing measurement matrix optimization method," *Journal of Internet of Things*, vol. 2, no. 1, pp. 47–54, 2020.

[11] X. R. Zhang, W. F. Zhang, W. Sun, X. M. Sun and S. K. Jha, "A robust 3-D medical watermarking based on wavelet transform for data protection," *Computer Systems Science & Engineering*, vol. 41, no. 3, pp. 1043–1056, 2022.

[12] H. Zhang, H. B. Zhang, Q. Li and X. M. Niu, "Image perception hashing algorithm based on human vision system," *Acta Electronica Sinica*, vol. 36, no. 12, pp. 30–34, 2008.

[13] X. R. Zhang, X. Sun, X. M. Sun, W. Sun and S. K. Jha, "Robust reversible audio watermarking scheme for telemedicine and privacy protection," *Computers, Materials & Continua*, vol. 71, no. 2, pp. 3035–3050, 2022.

[14] J. Xu and W. Chen, "Convolutional neural network-based identity recognition using ECG at different water temperatures during bathing," *Computers, Materials & Continua*, vol. 71, no. 1, pp. 1807–1819, 2022.

[15] K. M. Ding and C. Q. Zhu, "A perceptual hash algorithm for remote sensing image integrity authentication," *Journal of Southeast University (Natural Science Edition)*, vol. 44, no. 4, pp. 723–727, 2014.

[16] J. Zhang, Z. Wang, Y. Zheng and G. Zhang, "Design of network cascade structure for image super-resolution," *Journal of New Media*, vol. 3, no. 1, pp. 29–39, 2021.

[17] H. L. Lu, "Research on Image Retrieval Technology Based on Perceptual Hash," Wuhan Hubei: M. S. Hubei University of Technology, 2020.

[18] T. Zhang, Z. Zhang, W. Jia, X. He and J. Yang, "Generating cartoon images from face photos with cycle-consistent adversarial networks," *Computers, Materials & Continua*, vol. 69, no. 2, pp. 2733–2747, 2021.

[19] Y. Sun, "An image processing method of Tibetan murals based on wavelet compression sensing," *Laboratory Research and Exploration*, vol. 35, no. 5, pp. 138–140, 2016.

[20] H. He, Z. Zhao, W. Luo and J. Zhang, "Community detection in aviation network based on k-means and complex network," *Computer Systems Science and Engineering*, vol. 39, no. 2, pp. 251–264, 2021.

[21] W. Sun, L. Dai, X. R. Zhang, P. S. Chang and X. Z. He, "RSOD: Real-time small object detection algorithm in UAV-based traffic monitoring," *Applied Intelligence*, vol. 92, no. 6, pp. 1–16, 2021.