Folk art classification using support vector machine

Malay Bhatt, Apurva Mehta
Department of Computer Engineering, Faculty of Technology, Dharmsinh Desai University, Nadiad, India

Article Info

Article history:

Received Feb 8, 2024
Revised May 1, 2024
Accepted May 12, 2024

Keywords:

Classification
Data augmentation
Folk art
Generalized co-occurrence matrix
Local binary pattern
Support vector machine

ABSTRACT

Tremendous amounts of effort have been carried out every year by the governments of all the countries to preserve art and culture. Art in the form of paintings, artifacts, music, dance, and cuisines of every country has the utmost importance. The study of Tribal arts provides deep insight into our history and acts as a milestone in the roadmap of our future. This paper focuses on three popular folk arts namely: Gond, Manjusha, and Warli. 300 images of each artwork have been collected from various online repositories. To generate a robust system, data augmentation is applied which results in 7510 images. A feature vector based on a generalized co-occurrence matrix, local binary pattern, HSV histogram, and canny edge detector is constructed and classification is performed using a linear support vector machine. 10-fold cross-validation produces 99.8% accuracy.

Corresponding Author:

Malay Bhatt
Department of Computer Engineering, Faculty of Technology, Dharmsinh Desai University
Nadiad, India
Email: malaybhatt202@yahoo.com

1. INTRODUCTION

Cultural heritage includes tangible and intangible culture along with natural heritage. Cultural heritage helps maintain identity and cultural diversity which is essential for today's globalized world [1]. Governments are sincerely making efforts for the preservation and digitization of different cultural heritage [2]. Tribal art is one of the important cultural heritages. It has its own relevance based on the spatio-temporal continuum. It includes arts in the form of paintings, artifacts, music, and dance. The study of Tribal arts provides deep insight into our history and acts as a stepping stone for our future.

In the beginning, folk art was born to appease Gods/Goddesses, to articulate joy and happiness, to celebrate festivals, and to satisfy the psychological requirements of human beings. Over the period of time, folk arts have greatly influenced, attracted, and uplifted people all over the world. The artisans have been acclaimed worldwide, and folk art become a prominent source of earning for the artists.

India is well-known for its art and architecture. The history of India in terms of art is very immense. The government of India is making all-around efforts to bring folk art and folk artists into the global market. Some of the prominent Indian folk arts are Warli, Khovar, Tanjore/Thanjavur, Kawad, Madhubani, Pattachitra, Saura, Gond, Bhil, and Manjusha. The sample art forms for all these prominent folk arts are shown in Figure 1. Figure 1(a) shows a painting of Warli folk art. Figure 1(b) represents a painting made using Khovar folk art. Figure 1(c) shows a wall painting of Tanjore/ Thanjavur folk art. Figure 1(d) depicts an artifact decorated with Kawad folk art designs. Figure 1(e) shows a Madhubani-style painting over a clothing material. Figure 1(f) shows a painting of a dancer using Pattachitra folk art. A decorated cloth piece is shown in Figure 1(g) showing Saura folk art. Figure 1(h) shows a canvas painting of Gond folk art. Figure 1(i) presents Bhil folk art painting. Figure 1(j) demonstrates a painting of Manjusha folk art over a clothing material.

The 'Gond' art belongs to the Madhya Pradesh State of India and was developed by the Gond Community. These paintings are rich in details, lines, colors, mystery, and humor. These paintings are also drawn on paper, canvas, and cloths. In these paintings, lines, dots, and dashes are important features. The 'Wali' art is an asset of Maharashtra State; it is made up of basic geometrical shapes like squares, circles, and triangles. Scenes of daily life in ancient India such as hunting, festivals, fishing, farming, dancing, and others are portrayed in the paintings.

'Manjusha' Art originated in the state capital, Champa, which is currently located in Bhagalpur, Bihar. Pink, yellow, and green colors are mainly used in Manjusha's painting. These colors have esoteric and symbolic meanings. The pink and yellow colors signify excitement and exuberance, while green is a symbol of gloom and growth. Borders in Manjusha art include designs of Belpatra, Lehariya, Triangle, Mokha, and a series of Snakes [3], [4].

Figure 1. Prominent Indian folk arts: (a) Warli, (b) Khovar, (c) Tanjore/ Thanjavur, (d) Kawad, (e) Madhubani, (f) Pattachitra, (g) Saura, (h) Gond, (i) Bhil, and (j) Manjusha

These arts in the form of paintings are used by interior designers to decorate the living room and other commercial and non-commercial premises. Other than paintings, this artwork is carried out to decorate walls; and printed on clothes. The people of India are fond of wearing clothes and having any of these artworks. We can also see the presence of these artworks on the cover page of diaries and we can witness it in a decorative wall clock.

Today's globalized market is supported through e-commerce websites and mobile applications. The folk arts are sold across the globe through these mediums. E-commerce sites are selling hand-crafted pen stands painted with different folk art having varying costs as shown in Figure 2. A novice buyer cannot differentiate between these artworks. The high costs of these pen stand over others puzzle the buyer. A buyer has a lack of in-depth knowledge of the intricacies of folk art. These different art forms share similarities and sometimes the difference is not perceivable by everyone. In this situation, an automated classification system can help distinguish between different folk art.

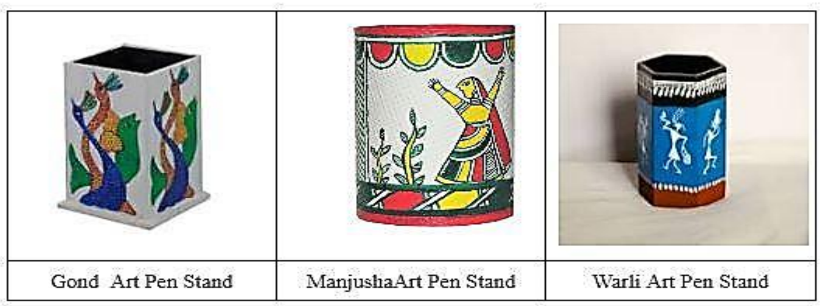

Figure 2. Pen stands with art work

The main objective of the paper is to promote different folk arts of India; to investigate the use of technology especially through e-commerce sites and mobile applications; to prepare an image dataset of the fork artwork images as no such image dataset is in existence for Indian folk art; to evaluate the performance of existing feature descriptor, which comes under traditional machine learning techniques, for the classification of different folk art images. Such work can assist in creating a digital archival system. That segregates provided images (i.e. museums' catalog, and artists' catalog) into respective categories based on classification algorithms.

This paper is divided into five sections. Section 2 covers the Literature Survey. The image data set for folk art images is prepared and it is discussed in section 3. The Feature descriptor proposed by the author is described in section 4. The result analysis is covered in section 5.

2. LITERATURE SURVEY

The recent development in the field of Computer Vision has made tremendous progress in various image processing tasks. These bring significant changes to the lives of people. Such applications are in the healthcare sector [5], remote sensing [6], defense [7], agriculture [8], music [9], and art [10]. Folk art classification is one such area that can empower the buyer to recognize the correct art form. These folk arts are drawn on different objects. The images of such objects are useful for making automatic recognition possible through computer vision and machine learning.

Visual descriptors of the images can be extracted from the local binary pattern (LBP) [11]. These descriptors act as features during modeling. Features based on LBP are used in the work [12] for the classification of brain tumors. LBP is utilized to find distance-based and angle-based features from the images of tumor samples. LBP-based feature descriptors are used for facial recognition. The work of [13] uses Orthogonal Difference-LBP for facial recognition. This approach represents each pixel of an image as 3 OD-LBP transformed images. These images are used for histogram construction for feature extraction. The LBP-based features are also used for facial recognition for the e-attendance system [14]. Texture features are extracted using a co-occurrence matrix. A co-occurrence matrix of an image can reflect the comprehensive information about the direction, adjacent interval, and variation range of the image [15]. It is used for Alzheimer's disease diagnosis [16] and Lung cancer prediction [17]. HSV colour space is more enriched than the RGB colour space for representing colours. In image processing tasks where image data has a range of colours, a feature vector based on an HSV histogram provides key colour details. HSV colour histogram is used in Content-based image retrieval systems [18], [19], recognizing induced emotions [20], and soil classification [21]. The canny edge detector is essential in detecting structural information from the images. It is more powerful than thresholding for detecting different types of contours from an image. It has been used for text recognition [22], Tuberculosis detection [23], and automatic flow structure detection [24].

Cultural heritage value mining has become essential in today's digitized world. The works on folk arts classification can help in this direction. Work is proposed for folk handicraft image identification. The work uses the ALOI database. The probability distribution of image feature vectors is obtained using Bayesian and Gaussian methods. The k-means algorithm is used for evaluating the accuracy of image extraction [25]. An interesting work attempts cross-media retrieval [26]. The work explores the emotional correlation for music and image data retrieval. The modeling is performed using a Differential Evolutionary-Support Vector Machine. This work reports motivating results in comparison to works related to semantics for cross-media retrieval.

It is observed that no work is reported on the Classification of Indian folk arts. The rich Indian folk arts are elaborate and essential for the identity of the community. This work plans to prepare a representative dataset of 3 Indian folk arts: Gond, Manjusha, and Warli. A study of the dataset will be performed to find a set of suitable feature extraction methodologies. Later, modeling is to be performed to automate the classification process.

3. DATA SET

We have manually created a dataset of images that incorporates three artworks: Manjusha, Warli, and Gond. The dataset consists of 300 images for each artwork and these images are downloaded from various websites. In total, there are 900 images. To make the system robust we have applied Data Augmentation methods [27]. We have performed: Flipping, Zooming, and Rotation operations. Details of Augmentation are given in Table 1. After augmentation, the augmented data set consists of approximately 2500 images for each artwork and there are 7510 images in total.

Table 1. Data augmentation

Property	Description
Flipping	1. X-Direction 2. Y-Direction
Zooming	1. 1.5x 2. 2x
Rotation (Clockwise)	1. 30 Degree 2. 90 Degree 3. 120 Degree 4. 180 Degree
Rotation (CounterClockwise)	30 Degree 90 Degree 120 Degree 180 Degree

4. PROPOSED APPROACH

This section is taken verbatim from the author's paper [28] to increase the readability. Figure 3 provides a flow chart of a content-based image retrieval system. An image query is the image file that is given as input to the system. The features of the input are calculated. A query of the extracted features is then generated and is compared with all the other features of the image files present in the database. Based on similarity measures, the system retrieves the required image files from the database and presents them in the form of the result.

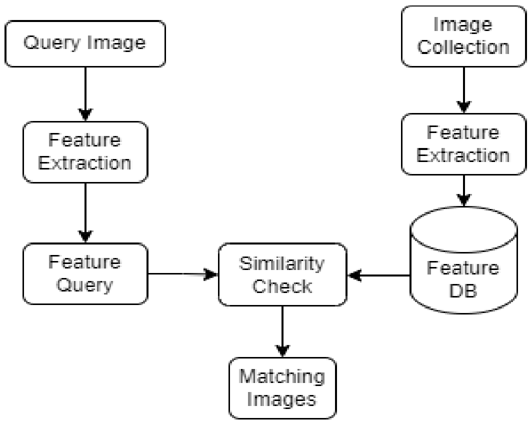

Figure 3. Content-based image retrieval system

4.1. Pre-processing

A set of 7510 input images is re-sized into 128x128 Resulted images are converted from RGB color space into hue, saturation, and value (HSV) color space [29], [30]. The layers of the human retina sense the light through rod cells and cone cells [31]. The gray levels are perceived by rod cells at low levels of illumination while at higher levels of illumination cone cells are also excited. The human perceives the color the same as the HSV color space. RGB color representation is different and not as per human perception. Hue indicates the pure color; S indicates the percentage of white added in the pure color and V represents intensity. The HSV color space can be represented as a hexacone [32]. When saturation is zero, we get only shades of gray from black to white by increasing the intensity. Incident light is composed of many spectral components but causes loss of color information when saturation is low even though illumination is very high. By changing the saturation from 0 to 1, perceived color changes from shades of gray to pure color under the given hue and intensity. It is known that HSV color space has more discriminating power as compared to RGB color space.

Generalized co-occurrence matrix properties with different distances and directions are also computed and it results in 20 additional features. H, S, and V planes are also extracted from the input image. The Computation of the Centre Symmetric Local Binary Pattern with 16 bins and Histogram with 16 bins are carried out on each plane. Generalized co-occurrence matrix properties are also extracted from each plane. Figure 4 covers several techniques which were merged together to generate the feature vector. Figure 5 shows a block diagram of preprocessing and the feature vector generation process.

Figure 4. Feature vector

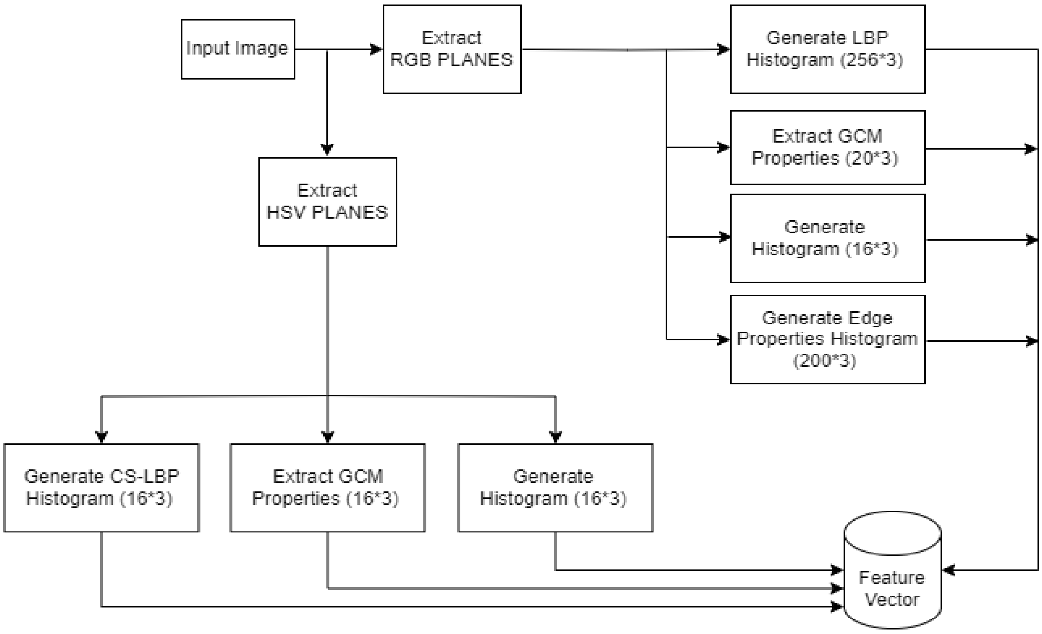

Figure 5. Pre-processing and feature vector generation

4.2. Generalized co-occurrence matrix

The Generalized CO-Occurrence Matrix is useful for extracting the texture of the image. It is represented as a 4-tuple (i, j, d, Ɵ) [33]. Here, í' and 'j' represent grey levels, and d is the distance between pixels p1 and p2. Gray Levels of p1 and p2 are i and j respectively. Ɵ is the angle between pixels p1 and p2.

Table 2 shows generalized co-occurrence matrices (GCM) of size 128x128 are calculated for inter-pixel distances 8, 16, 32, and 64 in a horizontal direction for H, S, and V planes. (3 planes x 4 inter- pixel distance). Working of the Generalized Co-Occurrence Matrix for a 5x5 matrix with 4 distinct values and x-direction distance '1' is described in Figure 6. Figure 6(a) describes the working of GCM for a sample image of size 5x5 with 4 gray levels while Figure 6(b) contains calculated GCM in the horizontal direction with inter-pixel distance '1'. There are four main properties of GCM namely contrast, correlation, energy, and homogeneity. These properties are described in Table 3.

Table 2. Generalized co-occurrence matrix used as features

Number of gray levels	Distance	Direction
128	8	Horizontal
128	16	Horizontal
128	32	Horizontal
128	64	Horizontal
128	64	Vertical

1	1	2	3	4
2	3	4	4	1
3	4	1	1	2
4	1	2	2	3
1	2	3	3	4

(a)

	1	2	3	4
1	2	4	0	0
2	0	1	3	0
3	0	0	1	4
4	3	0	0	1

(b)

Figure 6. Co-Occurrence Matrix (a) a sample image (b) the GCM in horizontal direction with inter-pixel distance '1'

Table 3. Generalized co-occurrence matrix properties

Property	Description
Contrast	It measures the intensity contrast between a pixel and its neighbor over the whole image.
Correlation	It indicates how correlated a pixel is to its neighbor over the whole image.
Energy	It represents the sum of squared elements in the GLCM. It is also known as uniformity.
Homogeneity	It measures the closeness of the distribution of elements in the GLCM to the GLCM diagonal.

Contrast: $\sum_{i,j}|i-j|^2 p(i,j)$ (1)

Correlation: $\sum_{i,j}\frac{(i-\mu_i)(j-\mu_j)p(i,j)}{\sigma_i\sigma_j}$ (2)

Energy: $\sum_{i,j} p(i,j)^2$ (3)

Homogeneity: $\sum_{i,j}\frac{p(i,j)}{1+|i-j|}$ (4)

p (i, j) represents count at position (i, j) in GLCM, μ denotes mean and σ indicates the standard deviation in the above equations. Here, small, medium, and large distance values are considered to capture the span of the shape in the horizontal and vertical directions.

4.3. LBP and CS-LBP

The local binary pattern effectively captures texture information from the local neighborhood. Figure 7 explains the working of the local binary pattern and the centre-symmetric local binary pattern.

$$LBP(x,y) = \sum_{i=0}^{n-1} s(n_c - n_i)2^i$$
$$S(x) = 1 \; if \; x \geq 0 \; otherwise \; 0 \quad (5)$$

Here, nc indicates the gray level of the center pixel of the 8-neighborhood, ni indicates ith pixel of the neighborhood. The signs of the differences in a neighborhood are interpreted as N-bit binary numbers resulting in 2N distinct values in the binary pattern. The LBP features are robust against illumination changes, they are very fast to compute, do not require many parameters to be set, and have high discriminative power [34]. In CS-LBP, center symmetric pairs of pixels are compared. LBP produces 256 distinct binary patterns, whereas CS-LBP generates 16 distinct binary patterns. The robustness of flat image regions is obtained by thresholding the gray level differences with a small value of T. In our proposed system, a histogram of CS-LBP is generated for all 3 planes of the HSV image resulting in 48 (16*3) while the histogram of LBP is obtained for all 3 planes of the RGB image resulting in 768 (256*3) features.

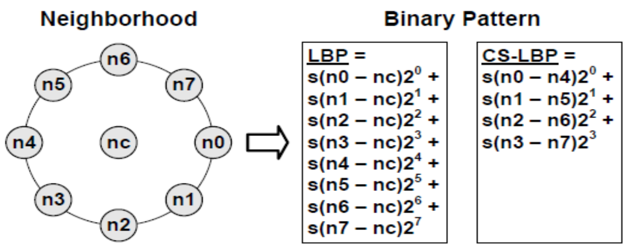

Figure 7. Local binary pattern and centre symmetric local binary pattern [33]

4.4. Edge histogram

Color information is obtained through histograms, Area information is added to the feature vector using a generalized co-occurrence matrix using different distances and directions, and texture information is achieved using LBP and CS-LBP histograms. To add the structural (behavior at the edge points) information in the feature descriptor, a canny edge detector is used with a threshold of 0.2 so that the most prominent edges are preserved. Canny edge detector consists of smoothing, finding gradients, non-maxima suppression, double thresholding, and edge tracking by hysteresis [35]. For each detected edge point, a 5x5 neighborhood is considered and the mean and the standard deviation are calculated. The unique values obtained from these statistical properties vary for every image because the detected edge points are not fixed. It is observed that unique values are in the range of 2,000-10,000. Two Histograms with bin size 100 are generated for mean and standard deviation.

4.5. Fitness function

Here, we have adopted the classification accuracy calculated by a linear SVM classifier on the training set as well as the testing set. The overall fitness 'Er' is the average of the tenfold cross-validation accuracy. In our case, the value of n is 10. Accuracy (i) represents the accuracy of fold 'i' by the SVM. The fitness function is defined as:

$$E_r = (1 - (\sum \frac{(SVM[accuracy(i)])}{n})) * \ 100\% \quad (6)$$

5. RESULTS AND DISCUSSION

We have evaluated our proposed method using 64-bit MATLAB 2022a, 16GB of RAM running on the Windows 11 OS with an i7 processor. We adopted 10-fold cross-validation for which the total dataset is divided randomly into 10 equal-sized parts and performed ten repetitions of training the SVM on 9/10 of the set and testing on the remaining 1/10 [36]. We have achieved 99.8% accuracy during 10-fold cross-validation. We have also performed 5-fold cross validation which contains 80% training images and 20% testing images. We have achieved 99.7% accuracy during 10-fold cross-validation. The confusion matrix for one iteration of 5-fold cross-validation is shown in Table 4. Precision, Recall, and F1-Score are given in Table 5.

$$\text{Precision } P = \frac{tp}{tp + fp} \quad (7)$$

$$\text{Recall } R = \frac{tp}{tp+fn} \quad (8)$$

$$\text{F1-score} = \frac{2PR}{P+R} \quad (9)$$

Table 4. Confusion matrix

Class	Gond	Manjusha	Warli
Gond	430	0	1
Manjusha	0	364	0
Warli	1	1	705

Table 5. Precision, Recall and F1-Score

Class	Precision	Recall	F1-Score
Gond	99.7%	99.7%	99.7%
Manjusha	99.7%	100%	99.8%
Warli	99.8%	99.7%	99.7%

The details of misclassified images are shown in Figure 8. Three images are misclassified. Figure 8(a), which belongs to class 'Gond', is misclassified as 'Warli'. Figures 8(b) and 8(c) belong to the class 'Warli'. Figure 8(b) is misclassified as 'Gond' while the last image is misclassified as 'Manjusha'. In Figure 8(a), as we zoom in, even though it contains lots of small *' and straight lines, it creates an illusion of geometric shapes such as triangles and circles. That may have triggered the misclassification as 'Warli'. In Figure 8(b) the Warli characters are in the center while the prominent area has no details of the Warli art form. The colors are used heavily for the demonstration of different festivals in Gond paintings. The same is present in Figure 8(b) leading to the misclassification as Gond instead of Warli. Figure 8(c) has shades of green color in the majority of the image area. It is a highlighting feature of the Manjusha art form. This could be the reason for the misclassification.

(a) (b) (c)

Figure 8. Misclassified images: (a) Gond, (b) Warli, and (c) Warli

6. CONCLUSION

Content-based image classification has generated potential applications in different areas like agriculture, arts, surveillance, and many more. The artworks are essential in the holistic representation of the country's tradition. The image dataset of 3 prominent Indian folk arts Gond, Manjusha, and Warli is considered. The feature vector is generated using histograms, local binary patterns, a generalized co-occurrence matrix, and a canny-edge detector. For classification, a linear support-vector machine is used. The proposed work reports an average accuracy of 99.8% based on 10-fold cross-validation. F1-score for Gond, Manjusha, and Warli are 99.7%, 99.8%, and 99.7% respectively.

REFERENCES

[1] A. Bhattacharya, "Art for life: intangible cultural heritage as a tool for entrepreneurial development in India," 2021, pp. 65–94.

[2] S. Kolay, "Cultural heritage preservation of traditional indian art through virtual new-media," *Procedia - Social and Behavioral Sciences*, vol. 225, pp. 309–320, Jul. 2016, doi: 10.1016/j.sbspro.2016.06.030.

[3] Holidify, [Online]. Available: https://www.holidify.com/pages/tribal-art-of-india-4497.html

[4] Manjusha art, [Online]. Available: https://umsas.org.in/manjusha-art/

[5] T. Dhar, N. Dey, S. Borra, and R. S. Sherratt, "Challenges of deep learning in medical image analysis—improving explainability and trust," *IEEE Transactions on Technology and Society*, vol. 4, no. 1, pp. 68–75, Mar. 2023, doi: 10.1109/tts.2023.3234203.

[6] X. Cheng, Y. Sun, W. Zhang, Y. Wang, X. Cao, and Y. Wang, "Application of deep learning in multitemporal remote sensing image classification," *Remote Sensing*, vol. 15, no. 15, p. 3859, Aug. 2023, doi: 10.3390/rs15153859.

[7] C. Shi, Y. Liu, M. Zhao, C. M. Pun, and Q. Miao, "Attack-invariant attention feature for adversarial defense in hyperspectral image classification," *Pattern Recognition*, vol. 145, p. 109955, Jan. 2024, doi: 10.1016/j.patcog.2023.109955.

[8] Z. Luo, W. Yang, Y. Yuan, R. Gou, and X. Li, "Semantic segmentation of agricultural images: A survey," *Information Processing in Agriculture*, Feb. 2023, doi: 10.1016/j.inpa.2023.02.001.

[9] M. Alfaro-Contreras, J. M. Iñesta, and J. Calvo-Zaragoza, "Optical music recognition for homophonic scores with neural networks and synthetic music generation," *International Journal of Multimedia Information Retrieval*, vol. 12, no. 1, p. 12, Jun. 2023, doi: 10.1007/s13735-023-00278-5.

[10] G. Xiaomo and S. Yongqing, "Genre and subject matter characteristics of Zuoquan folk songs from the perspective of intangible cultural heritage," *ournal of Namibian Studies: History Politics Culture*, vol. 33, pp. 3691–3704, 2023, doi: 10.59670/jns.v33i.2488.

[11] J. Tang, Q. Su, B. Su, S. Fong, W. Cao, and X. Gong, "Parallel ensemble learning of convolutional neural networks and local binary patterns for face recognition," *Computer Methods and Programs in Biomedicine*, vol. 197, p. 105622, Dec. 2020, doi: 10.1016/j.cmpb.2020.105622.

[12] K. Kaplan, Y. Kaya, M. Kuncan, and H. M. Ertunç, "Brain tumor classification using modified local binary patterns (LBP) feature extraction methods," *Medical Hypotheses*, vol. 139, p. 109696, Jun. 2020, doi: 10.1016/j.mehy.2020.109696.

[13] S. Karanwal and M. Diwakar, "OD-LBP: Orthogonal difference-local binary pattern for Face Recognition," *Digital Signal Processing: A Review Journal*, vol. 110, p. 102948, Mar. 2021, doi: 10.1016/j.dsp.2020.102948.

[14] B. Tej Chinimilli, A. Anjali, A. Kotturi, V. Reddy Kaipu, and J. Varma Mandapati, "Face recognition based attendance system using haar cascade and local binary pattern histogram algorithm," in *Proceedings of the 4th International Conference on Trends in Electronics and Informatics, ICOEI 2020*, Jun. 2020, pp. 701–704, doi: 10.1109/ICOEI48184.2020.9143046.

[15] S. Gao, "Gray level co-occurrence matrix and extreme learning machine for Alzheimer's disease diagnosis," *International Journal of Cognitive Computing in Engineering*, vol. 2, pp. 116–129, Jun. 2021, doi: 10.1016/j.ijcce.2021.08.002.

[16] Y. Kim, S. Lhatoo, G. Q. Zhang, L. Chen, and X. Jiang, "Temporal phenotyping for transitional disease progress: An application to epilepsy and Alzheimer's disease," *Journal of Biomedical Informatics*, vol. 107, p. 103462, Jul. 2020, doi: 10.1016/j.jbi.2020.103462.

[17] L. Hussain *et al.*, "Lung cancer prediction using robust machine learning and image enhancement methods on extracted gray-level co-occurrence matrix features," *Applied Sciences (Switzerland)*, vol. 12, no. 13, p. 6517, Jun. 2022, doi: 10.3390/app12136517.

[18] A. Nazir, R. Ashraf, T. Hamdani, and N. Ali, "Content based image retrieval system by using HSV color histogram, discrete wavelet transform and edge histogram descriptor," in *2018 International Conference on Computing, Mathematics and Engineering Technologies: Invent, Innovate and Integrate for Socioeconomic Development, iCoMET 2018 - Proceedings*, Mar. 2018, vol. 2018-January, pp. 1–6, doi: 10.1109/ICOMET.2018.8346343.

[19] N. Danapur, S. A. A. Dizaj, and V. Rostami, "An efficient image retrieval based on an integration of HSV, RLBP, and CENTRIST features using ensemble classifier learning," *Multimedia Tools and Applications*, vol. 79, no. 33–34, pp. 24463–24486, Sep. 2020, doi: 10.1007/s11042-020-09109-9.

[20] S. A. Mohseni, H. R. Wu, J. A. Thom, and A. Bab-Hadiashar, "Recognizing induced emotions with only one feature: A novel color histogram-based system," *IEEE Access*, vol. 8, pp. 37173–37190, 2020, doi: 10.1109/ACCESS.2020.2975174.

[21] D. N. Kiran Pandiri, R. Murugan, and T. Goel, "Smart soil image classification system using lightweight convolutional neural network," *Expert Systems with Applications*, vol. 238, p. 122185, Mar. 2024, doi: 10.1016/j.eswa.2023.122185.

[22] F. Wu, C. Zhu, J. Xu, M. W. Bhatt, and A. Sharma, "Research on image text recognition based on canny edge detection algorithm and k-means algorithm," *International Journal of System Assurance Engineering and Management*, vol. 13, no. S1, pp. 72–80, Mar. 2022, doi: 10.1007/s13198-021-01262-0.

[23] S. K. T. Hwa, A. Bade, M. H. A. Hijazi, and M. S. Jeffree, "Tuberculosis detection using deep learning and contrast-enhanced canny edge detected X-Ray images," *IAES International Journal of Artificial Intelligence*, vol. 9, no. 4, pp. 713–720, Dec. 2020, doi: 10.11591/ijai.v9.i4.pp713-720.

[24] I. A. Znamenskaya and I. A. Doroshchenko, "Edge detection and machine learning for automatic flow structures detection and tracking on schlieren and shadowgraph images," *Journal of Flow Visualization and Image Processing*, vol. 28, no. 4, pp. 1–26, 2021, doi: 10.1615/jflowvisimageproc.2021037690.

[25] X. Yu and Y. Dai, "Research on folk handicraft image recognition based on neural networks and visual saliency," *Mathematical Problems in Engineering*, vol. 2022, pp. 1–7, May 2022, doi: 10.1155/2022/6290814.

[26] B. Xing *et al.*, "Emotion-driven Chinese folk music-image retrieval based on DE-SVM," *Neurocomputing*, vol. 148, pp. 619–627, Jan. 2015, doi: 10.1016/j.neucom.2014.08.007.

[27] K. Maharana, S. Mondal, and B. Nemade, "A review: data pre-processing and data augmentation techniques," *Global Transitions Proceedings*, vol. 3, no. 1, pp. 91–99, Jun. 2022, doi: 10.1016/j.gltp.2022.04.020.

[28] M. S. Bhatt and T. P. Patalia, "Indian monuments classification using support vector machine," *International Journal of Electrical and Computer Engineering*, vol. 7, no. 4, pp. 1952–1963, Aug. 2017, doi: 10.11591/ijece.v7i4.pp1952-1963.

[29] S. M.Waikar and K. B. K. K.B.Khanchandani, "Determination of image features for content-based image retrieval using interactive genetic algorithm," *International Journal of Computer Applications*, vol. 89, no. 17, pp. 13–17, 2014, doi: 10.5120/15722-4583.

[30] H. FathyAtlam, G. Attiya, and N. El-Fishawy, "Comparative study on CBIR based on color feature," *International Journal of Computer Applications*, vol. 78, no. 16, pp. 9–15, 2013, doi: 10.5120/13605-1387.

[31] R. C. Gonzalez, "Digital image processing, 4th Edition," *Pearson Education iIdia*, 2009. .

[32] Y. An, M. Riaz, and J. Park, "CBIR based on adaptive segmentation of HSV color space," in *UKSim2010 - UKSim 12th International Conference on Computer Modelling and Simulation*, 2010, pp. 248–251, doi: 10.1109/UKSIM.2010.53.

[33] M. Heikkilä, M. Pietikäinen, and C. Schmid, "Description of interest regions with center-symmetric local binary patterns," 2006, pp. 58–69.

[34] K. Meena, A. Suruliandi, and R. Rose, "Face recognition based on local derivative ternary pattern," *IETE Journal of Research*, vol. 60, no. 1, pp. 20–32, Jan. 2014, doi: 10.1080/03772063.2014.890811.

[35] B. Syam, J. S. R. Victor, and Y. S. Rao, "Efficient similarity measure via genetic algorithm for content based medical image retrieval with extensive features," in *Proceedings - 2013 IEEE International Multi Conference on Automation, Computing, Control, Communication and Compressed Sensing, iMac4s 2013*, Mar. 2013, pp. 704–711, doi: 10.1109/iMac4s.2013.6526499.

[36] J. Han, M. Kamber, and D. Mining, "Concepts and techniques," *Morgan Kaufmann*, pp. 94104–3205, 2006.

BIOGRAPHIES OF AUTHORS

Prof. Malay Bhatt was born at Bhuj in Gujarat, India on 20th February, 1981. Prof. Bhatt completed Bachelor of Engineering with specialization in Computer Engineering from C.U. Shah College of Engineering & Technology, Surendranagar, Gujarat, India in the year 2002. Prof. Bhatt completed Masters in Engineering with computer engineering as specialization from Dharmsinh Desai University, Nadiad, Gujarat, India in the year 2004. Prof. Bhatt has completed Ph.D. in Computer Engineering from Rai University, Ahmedabad, Gujarat, in the year 2017. The author's major field of study covers Image Processing and Multimedia Information Retrieval. He has been working as Associate Professor in the Computer Engineering Department at Dharmsinh Desai University, Nadiad, Gujarat, India since 2008. He delivers expert talks and serves as reviewer in international conferences and journals. He has also presented papers in international Journals, IEEE Conferences, National Journals and National Conferences. He can be contacted at email: malaybhatt202@yahoo.com.

Prof. Apurva Mehta was born at Bhavnagar in Gujarat, India on 10th June, 1990. He completed B.E., M.Tech. and Ph.D. with a specialization in Computer Engineering from Dharmsinh Desai University, Nadiad, Gujarat, India in the years 2011, 2014, and 2022 respectively. The author's major field of study covers computational biology, machine learning and image processing. He has been working as Assistant Professor in the Computer Engineering Department at Dharmsinh Desai University, Nadiad, Gujarat, India since 2011. He delivers expert talks and serves as reviewer in international conferences and journals. He has also presented papers in International Journals, IEEE Conferences. He can be contacted at email: apurvamehta2015@gmail.com.