Sustainable Marine Structures
https://journals.nasspublishing.com/index.php/sms

ARTICLE

Integrated Application Research on Marine Image Recognition Models

Chih-Chen Kao [1] ***, Yu-Fen Peng*** [2] ***, Bo-Wen Wu*** [3*]

[1] *School of Intelligent Manufacturing, Shanghai Zhongqiao Vocational and Technical University, Shanghai 201514, China*

[2] *Physical Education Office, Yuanpei University of Medical Technology, Hsinchu 30015, Taiwan*

[3] *Department of Optometry, College of Medical Technology and Nursing, Yuanpei University of Medical Technology, Hsinchu 30015, Taiwan*

ABSTRACT

Marine environments present significant challenges for image processing due to factors such as low light intensity, suspended particles, and varying degrees of water turbidity. These conditions severely degrade the clarity and quality of captured marine images, making accurate image recognition difficult. The problem is further compounded by the limited availability of high-quality, labeled training samples, which restricts the effectiveness of conventional recognition algorithms. Existing techniques in both academic and industrial settings—such as Principal Component Analysis (PCA), Neural Networks, and Wavelet Transforms—typically involve converting color images to grayscale prior to feature extraction. While this simplifies processing, it also results in the loss of essential color information, which is often critical for distinguishing features in marine imagery. To address these issues, this paper proposes a novel approach that preserves and utilizes the full color information of marine images during processing and recognition. The method combines color image representation with Hu's invariant moments to extract stable and rotation-invariant features. These features are then input into a Back Propagation Neural Network (BPNN), which is trained to recognize and classify various marine targets. The integration of color-based feature extraction with BPNN significantly improves recognition performance, particularly under complex environmental conditions. Experimental results show that the proposed system achieves a recognition accuracy exceeding 98%, demonstrating its effectiveness and potential for practical applications in marine exploration, environmental monitoring, and underwater robotics.

Keywords: Marine Image Color Preprocessing; Pattern Recognition; BPNN; Invariant Moments

*CORRESPONDING AUTHOR:
Bo-Wen Wu, Department of Optometry, College of Medical Technology and Nursing, Yuanpei University of Medical Technology, Hsinchu 30015, Taiwan; Email: wubwen@mail.ypu.edu.tw

ARTICLE INFO
Received: 28 March 2025 | Revised: 22 April 2025 | Accepted: 29 April 2025 | Published Online: 6 May 2025
DOI: https://doi.org/10.36956/sms.v7i2.1915

CITATION
Kao, C.-C., Peng, Y.-F., Wu, B.-W., 2025. Integrated Application Research on Marine Image Recognition Models. Sustainable Marine Structures. 7(2): 38–44. DOI: https://doi.org/10.36956/sms.v7i2.1915

1. Introduction

In the field of image recognition, most research focuses on the processing of grayscale images or utilizes commonly employed classifiers for example Naive Bayes Classifier and Support Vector Machine [1, 2]. There is a scarcity of studies that retain color images while integrating BPNN technology [3–7], let alone the design and development of a comprehensive image recognition model to assist in marine image recognition. Chapter 1 is an introduction, which outlines the structure and research direction of this paper. Chapter 2 details the methods for image preprocessing and color retention. Chapter 3 employs Hu's invariant moments for feature extraction. Chapter 4 discusses the experimental results of recognizing marine images using Back Propagation Neural Network (BPNN). Finally, Chapter 5 presents the conclusions drawn from the study.

2. Image Conversion

2.1. Image Capture

The purpose of image preprocessing is to effectively separate targets from blurred images, preparing them for further recognition. The main steps involved in the image preprocessing are illustrated in **Figure 1**.

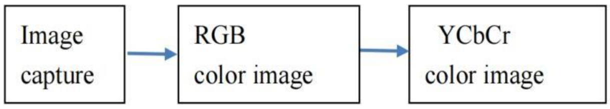

Figure 1. Steps in Image Preprocessing.

Marine images are captured using marine cameras. Due to insufficient lighting in the marine environment, the images tend to be blurred. The image acquisition system consists of marine imaging cameras, personal computers, image capture cards, image processing software, and other components.

2.2. YCbCr Color Model

An image in the RGB color model consists of three independent image planes, each corresponding to one of the three primary colors: red, green, and blue. When these three image planes are transmitted to an RGB display, they combine to form a full-color image.

Color information plays a crucial role in image processing. However, to effectively reduce data size and computational complexity, color images are often converted to grayscale. This conversion inevitably results in the loss of significant color information. Therefore, this study focuses on transforming RGB color space into the YCbCr color space, which is less sensitive to luminance variations. This choice is motivated by the fact that many studies on image and color detection are conducted in this color model, as shown in Equation (1) [3,4].

$$f(x,y) = f_y(x,y)i + f_{cb}(x,y)j + f_{cr}(x,y)k \tag{1}$$

The YCbCr color model is a type of full-color representation widely used in continuous image processing for video and digital video formats, such as JPEG and MPEG compression standards. The method proposed in this paper also operates within this color space. In this model, luminance information is represented by a single component, Y, while color information is stored in two chrominance components, Cb and Cr. The Cb component represents the difference between the blue channel and a reference value, whereas the Cr component represents the difference between the red channel and a reference value. The conversion formula from RGB toYCbCr is given in Equation (2).

A key advantage of using the YCbCr color model is that the transformation between RGB and YCbCr is linear, fast, and reversible. If the YCbCr color model needs to be converted back to the RGB color model, the inverse transformation can be performed using Equation (3).

$$\begin{gathered} Y = 0.299R + 0.587G + 0.114B \\ Cb = -0.1687R - 0.3313G + 0.5B + 128 \\ Cr = 0.5R - 0.4187G - 0.0813B + 128 \end{gathered} \tag{2}$$

$$\begin{gathered} R = Y + 1.402(Cr - 128) \\ G = Y - 0.3441(Cb - 128) - 0.714(Cr - 128) \\ B = Y + 1.772(Cb - 128) \end{gathered} \tag{3}$$

This paper uses color images and the commonly used YCbCr model to extract feature values, which is different from traditional image processing methods (such as grayscale images, edge enhancement, noise elimination, and binarization).

3. Invariant Moment Feature Extraction

Hu's proposed that these invariant moments possess properties of translation, rotation, and scale invariance [8, 9]. If $f(x,y)$ is a piecewise continuous function with nonzero values only within a finite region of the plane, then invariant moments of all orders uniquely exist. Mathematically, an infinite sequence of moments can uniquely represent $\{\mu_{p,q},\ p + q = 0, 1,2 \ldots\}$ using the $f(x,y)$, as shown in Equation (4).

$$f(x,y) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} exp[-2j\pi(ux+vy)] \cdot \left[\sum_{p=0}^{\infty}\sum_{q=0}^{\infty} m_{p,q}\frac{(2j\pi)^{p+q}}{p!\,q!}\right] dudv \tag{4}$$

The central moment μ_{pq} can be expressed by m_{pq}. For example, μ_{00} as shown in Equation (5). The physical meanings of each central moment are shown in Table 1.

In this study, the method of integrating color information with Hu's invariant moments considers the transformed image's color function as the vector part of a three color components, denoted as $f(x, y) = f_y(x, y)i + f_{cb}(x, y)j + f_{cr}(x, y)k$.

$$\mu_{00} = \sum_{x=1}^{M}\sum_{y=1}^{N}(x-x_c)^0(y-y_c)^0 f(x,y) = m_{00}\,f(x,y) \tag{5}$$

Table 1. Physical interpretation of central moments.

μ_{pq}	Physical Meaning	>0	<0
μ_{20}	Represents the image's horizontal spread		
μ_{02}	Represents the image's vertical spread		
μ_{11}	Represents the image's tilt	Tilted to the left	Tilted to the right
μ_{30}	Represents the centroid shift in the horizontal direction	Centroid shifts left	Centroid shifts downward
μ_{03}	Represents the centroid shift in the vertical direction	Centroid shifts upward	Upper part expands more than the lower part
μ_{21}	Represents the balance of horizontal expansion	Lower part expands more than the upper part	
μ_{12}	Represents the balance of vertical expansion	Right side expands more than the left side	Left side expands more than the right side

The central moment μ_{pq} reflects the distribution of pixel values relative to the pixel centroid of the image. Based on the mathematical formulation of μ_{pq} central moments are invariant to translation and scale transformations. To ensure that invariant moments also possess rotational invariance, Hu's proposed seven invariant moments defined using normalized central geometric moments. These seven moments exhibit translation invariance, scale invariance, and rotation invariance, as shown in Equation (6).

$$\begin{gathered}
\phi_1 = \mu_{20} + \mu_{02} \\
\phi_2 = (\mu_{20} + \mu_{02})^2 + 4{\mu_{11}}^2 \\
\phi_3 = (\mu_{30} - 3\mu_{12})^2 + (\mu_{03} - 3\mu_{21})^2 \\
\phi_4 = (\mu_{30} - \mu_{12})^2 + (\mu_{03} - \mu_{21})^2 \\
\phi_5 = (\mu_{30} - 3\mu_{12})(\mu_{30} + \mu_{12})\phi_x + (\mu_{03} - 3\mu_{21})(\mu_{03} + \mu_{21})\phi_y \\
\phi_6 = (\mu_{20} - \mu_{02})[(\mu_{30} + \mu_{12})^2 - (\mu_{03} + \mu_{21})^2] + 4\mu_{11}(\mu_{30} + \mu_{12})(\mu_{03} + \mu_{21}) \\
\phi_7 = (3\mu_{21} - \mu_{03})(\mu_{30} + \mu_{12})\phi_x + (\mu_{30} - 3\mu_{21})(\mu_{03} + \mu_{21})\phi_y
\end{gathered} \tag{6}$$

The formulas for ϕ_x and ϕ_y are shown below.

$$\phi_x = (\mu_{30} - \mu_{12})^2 - 3(\mu_{03} + 3\mu_{21})^2, \phi_y = (\mu_{03} - \mu_{21})^2 - 3(\mu_{30} + \mu_{12})^2 \quad (7)$$

4. Back Propagation Neural Network (BPNN)

The images used in the experiment were taken by divers in the waters of the Penghu Islands (Taiwan) and are indeed used for training and testing recognition. These images were taken by divers in the waters of the Penghu Islands (Taiwan). The images were selected to include various shapes of animals and plants and take into account high, medium and low frequency components. This study utilizes 96 images from 12 common types of marine imagery, focusing on the recognition of blurred images in low-resolution marine environments, as shown in **Figure 2**.

Because marine images are difficult to collect and the number of images is relatively small, this study increased the hidden units and number of operations of the neural network. After applying Backpropagation Neural Network (BPNN) recognition, with a sufficiently large number of iterations (50,000), the average recognition accuracy exceeds 98% [10–14]. The training and recognition results are presented in **Figure 3** and **Tables 2–5.**

Figure 2. Marine Images (12 Categories).

Table 2. Neural Network Parameter Settings.

Name	Parameter Settings
Input Units	7 units (Φ_1, Φ_2, Φ_3, Φ_4, Φ_5, Φ_6, Φ_7)
Hidden Units	10 units
Output Units	12 units (corresponding to 12 image categories)
Number of Iterations	50000
Learning Rate	0.5
Momentum Term	0.5
Number of Training and Testing Samples	48 training samples, 48 testing and validation samples

Table 3. Normalized Invariant Moments of Marine Images (12 Categories).

Sample\Invariants	Φ_1	Φ_2	Φ_3	Φ_4	Φ_5	Φ_6	Φ_7
1	0.1188	0.3944	0.4396	0.4247	0.9219	0.6346	0.8423
2	0.1138	0.3620	0.4557	0.4619	0.9599	0.6673	0.9297
3	0.1173	0.3907	0.4596	0.4482	0.9169	0.6454	0.8883
4	0.1123	0.3082	0.4183	0.4253	0.8476	0.5807	0.8270
5	0.1189	0.3171	0.4306	0.4770	0.9092	0.6366	0.8937
6	0.1188	0.2970	0.4018	0.4142	0.8126	0.5919	0.8322
7	0.1153	0.3655	0.5000	0.4657	0.9523	0.6831	0.9816

Table 3. *Cont.*

Sample\Invariants	Φ_1	Φ_2	Φ_3	Φ_4	Φ_5	Φ_6	Φ_7
8	0.1194	0.3809	0.4559	0.4527	0.9108	0.6432	0.9208
9	0.1201	0.3310	0.4329	0.4727	0.9248	0.6620	0.9515
10	0.1191	0.3679	0.4729	0.5188	0.9929	0.7029	1.0000
11	0.1199	0.3955	0.4674	0.4957	0.9868	0.7167	0.9472
12	0.1175	0.4222	0.4437	0.4429	0.8698	0.6779	0.8821

Table 4. Weights and Biases After Neural Network Training.

Weights						Biases						
Weighting = [Input][Hide] = [7][10]	11.38	−102.55	−47.84	9.2	−2.65	19.37	21.16					
	7.74	−3.5	−8.37	−17.25	−40.23	20.61	37.35					
	2.9	−49.95	−11.09	56.16	−14.74	−38.77	43.32					
	9.74	8.92	−42.54	−40.95	−25.42	45.05	5.78					
	5.03	−20.81	3.2	21.74	−20.91	−17.77	38.88					
	12.22	10.3	14.03	−21.32	−21.22	−68.58	−10.74					
	−2.37	−39.53	32.22	−15.03	−1.75	−61.9	24.66					
	9.55	−15.63	−10.89	−8.6	−39.3	9.38	33.3					
	−9.57	18.25	20.87	−80.56	−2.02	32.92	−1.51					
	9.04	−0.59	−56.88	21.17	18.86	16.71	−33.4					
Weighting = [Hide][output] = [10][12]	−10	−4	−14.3	−0.08	−5.02	31.22	3.14	−5.22	2.84	−4.6		
	15.9	−7.66	−3.65	1.67	−5.33	−16.2	8.54	−6.89	16.77	10.7		
	−43	−10.9	23.18	−7.28	−16.0	1.63	10.1	−12.1	9.49	−0.3		
	5.93	−7.29	−3.86	−11.8	0.36	24.17	19.6	−4.8	−5.41	−10		
	22.4	−9.58	−3.04	−15.3	−1.46	−6.42	15.4	−1.65	−23.6	6.38		
	20.3	4.08	−0.67	15.6	−3.78	4.98	−0.9	6.17	0.31	5.88		
	−11	0.23	−0.28	−12.8	−4.72	−17.4	8.25	−1.97	14.86	−15		
	−25	6.15	35.03	−3.32	3.14	10.09	7.86	7.28	−16	−2.7		
	22.2	11.06	−8.04	9.59	−0.34	−19.7	−6.1	12.48	0.68	2.23		
	−1.3	−3.52	14.31	−13.5	7.71	−31.4	−4.0	−1.94	−17.1	−4.3		
	−19	2.28	4.02	1.00	3.54	−36.5	−23	−3.63	−6.52	5.53		
	−11	2.46	−17.1	7.46	−8.09	−9.43	−14	−2.47	11.93	−4.3		
Bias hide	−24	−0.87	0.32	−21.2	9.27	−71.7	−27	−12.4	−4.28	−18		
Bias output	17	21.6	15	21.6	24	42	0	31	23.4	6.1	−8.4	2.5

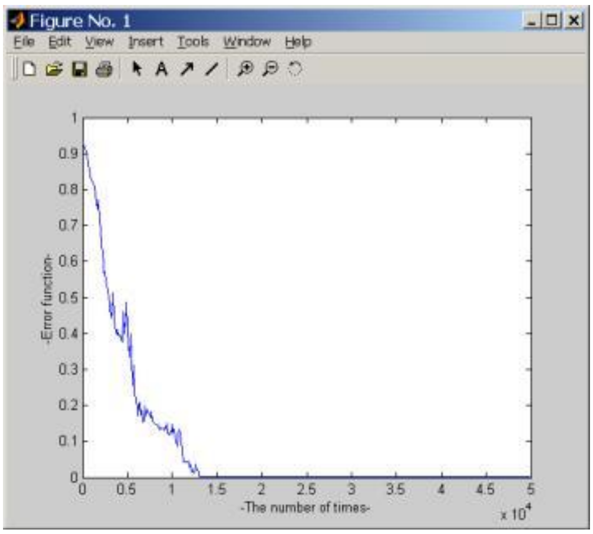

Figure 3. Neural Network Training Curve.

Table 5. Neural Network Recognition Results.

Rec. rate	1	2	3	4	5	6	7	8	9	10	11	12	Er. fnc.
1	1.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00007
2	0.00	**0.99**	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00007
3	0.00	0.00	**0.99**	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00026
4	0.00	0.00	0.00	**0.99**	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00007
5	0.00	0.00	0.00	0.00	**0.99**	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00009
6	0.00	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00005
7	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.01	0.00	0.00	0.00	0.00	0.00006
8	0.00	0.00	0.01	0.00	0.00	0.00	0.00	**0.98**	0.00	0.00	0.00	0.00	0.00018
9	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00005
10	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	1.00	0.00	0.00	0.00005
11	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00	0.00002
12	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00001

The bold fonts in the table are 12 output units, and the correct recognition rate of each unit is over 98%.

5. Conclusions

This paper adopts an approach for preserving color image processing, combined with Hu's invariant moments for feature extraction, which includes rotational and non-integer scaling invariance. These features are applied to achieve rotation and scale invariance in object recognition. The Backpropagation Neural Network (BPNN) is currently one of the most efficient algorithms in the field of image processing, achieving an overall recognition accuracy exceeding 98%. In the future, if we can improve the image pre-processing technology and collect more training samples, Convolutional Neural Network (CNN) deep learning techniques could be applied to train recognition models, enabling the broader application of the research results.

Author Contributions

The author individual contributions are described as follows: "Conceptualization, C.K. and B.W.; methodology, B.W.; software, Y.P.; validation, C.K., B.W. and Y.P.; formal analysis, B.W.; investigation, B.W.; resources, C.K.; data curation, Y.P.; writing—original draft preparation, B.W.; writing—review and editing, B.W.; visualization, B.W.; supervision, C.K.; project administration, C.K.; funding acquisition, C.K. All authors have read and agreed to the published version of the manuscript.

Funding

This work received no external funding.

Institutional Review Board Statement

This paper does not involve research on humans or animals and does not require ethical review and approval.

Informed Consent Statement

This paper does not involve research on humans, so we chose to exclude this statement.

Data Availability Statement

This paper has no additional data links.

Acknowledgments

This paper did not receive any additional administrative and technical support or donation.

Conflicts of Interest

No conflict of interest is declared for this paper.

References

[1] Koutroumbas, K., Theodoridis, S., 2010. Pattern Recognition, 4th Edition. Academic Press: New York, NY, USA. pp. 13-86.

[2] Everitt, B., Krzanowski, W.J.,1988. Principles of Multivariate Analysis. Oxford University Press: Oxford, UK. pp. 332-364.

[3] Wu, B.W., Huang, C.C., Lin, W.C., 2023. Development of Neural Network Recognition Model for Optical Illusion Images. Journal of Physics: Conference Series. 2468, 012123. DOI: https://doi.org/10.1088/1742-6596/2468/1/012123

[4] Wu, B.W., Fang, Y.C., Wen, C.C., et al., 2022. A Study of Artificial Neural Network Technology Applied to Image Recognition for Underwater Images. IEEE ACCESS. 10, 13844–13851. DOI: https://doi.org/10.1109/ACCESS.2022.3144742

[5] Wu, B.W., Fang, Y.C., 2021. Application of blurred circular 3D images on the human vision model. Microsystem Technologies. 27, 1099–1105.

[6] Wu, B.W., Fang, Y.C., 2015. Applications of neural networks in human shape visual perception. Journal of the Optical Society of America A. 32(12), 2338–2345.

[7] Wu, B.W., Fang, Y.C., 2007. Neural Network Application to Thermal Image Recognition of Low Resolution Objects. Journal of Optics A- Pure and Applied Optics. 9(2), 134–144.

[8] Hu, M.K., 1962. Visual pattern recognition by moment invariants. IRE Transactions on Information Theory. 8(2), 179–187.

[9] Fine, N.J., Wilf, H.S., 1965. Uniqueness theorems for periodic functions. American Mathematical Society. 16(1), 109–114.

[10] Backes, A.R., Junior, J.S., 2017. LBP maps for improving fractal based texture classification. Neuro computing. 266, 1–7.

[11] Abuzneid, M.A., Mahmood, A., 2018. Enhanced human face recognition using LBPH descriptor multi-KNN and back-propagation neural network. IEEE Access. 6, 20641–20651.

[12] Traore, A.B., 2018. Deep convolution neural network for image recognition. Ecological Informatics. 48, 257–268.

[13] Dheir, M., Mettleq, A., 2019. Classifying nuts types using convolutional neural network. International Journal of Academic Information Systems Research. 3, 12–18.

[14] Kuo, C.-C.J., Zhang, M., Li, S.Y., et al., 2019. Interpretable convolutional neural networks via feedforward design. Journal of Visual Communication and Image Representation. 60, 346–359.