IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received December 19, 2021, accepted January 7, 2022, date of publication January 25, 2022, date of current version February 7, 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3144742

A Study of Artificial Neural Network Technology Applied to Image Recognition for Underwater Images

BO-WEN WU[1], YI-CHIN FANG[2], CHAN-CHUAN WEN[3], CHAO-HSIEN CHEN[4], HSIAO-YI LEE[5], AND SHUN-HSYUNG CHANG[6]

[1]Department of Optometry, Yuanpei University of Medical Technology, Hsinchu 30015, Taiwan
[2]Department of Mechatronics Engineering, National Kaohsiung University of Science and Technology, Kaohsiung 82444, Taiwan
[3]Department of Shipping Technology, National Kaohsiung University of Science and Technology, Kaohsiung 82444, Taiwan
[4]Department of Mechanical Engineering, National Kaohsiung University of Science and Technology, Kaohsiung 82444, Taiwan
[5]Department of Electrical Engineering, National Kaohsiung University of Science and Technology, Kaohsiung 82444, Taiwan
[6]Department of Microelectronics Engineering, National Kaohsiung University of Science and Technology, Kaohsiung 82444, Taiwan

Corresponding author: Yi-Chin Fang (yfang@nkust.edu.tw)

ABSTRACT In this study, the researchers developed holographic image software for the Polaris, a nongovernmental Taiwanese oceanographic research vessel. It is a survey vessel that was codeveloped through an industry–academia collaboration between National Kaohsiung University of Science and Technology and Dragon Prince Hydro-Survey Enterprise Co. With a weight of 260 tons, length of 36.98 m, and width of 6.80 m, the vessel can travel at a speed of 11 knots. It has undergone underwater rescue and exploration operations and is therefore fairly experienced in such operations. When performing underwater exploration missions, survey vessels are often faced with interferences caused by factors such as current velocity; water temperature, refraction, and spectral conditions; climate; ocean current; presence of algae; and light reflection from schools of fish. Therefore, instantaneous image analysis is imperative for marine exploration. In accordance with the instantaneous recognition needs of the Polaris, the researchers developed artificial-neural-network-based recognition software for rapidly recognizing the category of a detected underwater object. Recognition of shapes in low-resolution underwater images was improved using a neural network resulting in an average recognition rate of 95%. Analysis of variance also indicated that the neural network yielded a significantly higher recognition rate than did manual recognition.

INDEX TERMS Image processing, moment invariants, neural network, image recognition.

I. INTRODUCTION

The Polaris is the first nongovernmental oceanographic research and survey vessel of Taiwan. The vessel was codeveloped through an industry–academia collaboration between National Kaohsiung University of Science and Technology (the Nanzih Campus) and Dragon Prince Hydro-Survey Enterprise Co. It was constructed by Shing Sheng Fa Boat Building Co., Ltd. and has a weight of 260 tons, length of 36.98 m, and width of 6.80 m. With a top speed of 11 knots, the vessel has been stationed at the Innovation Incubation Center of National Kaohsiung University of Science and Technology since September 2008, and it is usually parked at the jetty of the Cijin Campus. The vessel has undergone numerous underwater rescue and exploration operations. For example, after the mysterious disappearance of Malaysia Airlines flight MH370 on March 8, 2014, various parties were involved in the searching operation. The Indian Ocean has an average depth of approximately 3900 m, with the deepest point being approximately 8000 m. The search operation was difficult and required advanced equipment. At the time, the Polaris was invited to assist with the salvation operation in Australia through its underwater detection and deep tow system; this indicated that the system's superior quality is recognized internationally. The internal configuration of the Polaris is illustrated in Figure 1.(a)–(c).

The associate editor coordinating the review of this manuscript and approving it for publication was Donato Impedovo.

The major equipment and devices on the Polaris are an underwater remotely operated vehicle, differential satellite positioning system, single-beam echo sounder, multibeam echo sounder, side-scan sonar, subbottom profiler,

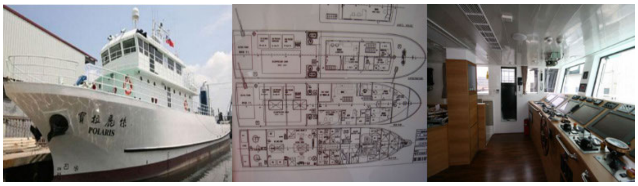


FIGURE 1. **(a)The Polaris (b)Cabin configuration diagram (c)Data processing center.**

marine magnetometer, underwater positioning system, sediment corer, and penetrometer.

With regard to the underwater imaging operations of the Polaris, the focus is not on capturing beautiful and magnificent underwater images with high contrast, which is the primary aim for underwater photographers. The numerous disaster relief and emergency salvation operations that the Polaris has been involved in indicate that in terms of underwater imaging, enhancement of the underwater object recognition rate is more essential. The recognition rate is defined as the probability of correctly identifying underwater objects (e.g., is it a school of fish, a torpedo, a sunken ship, or a reef rock; what is the probability?) from captured images. Therefore, to ensure that the Polaris recognizes objects instantaneously and accurately, the focus of the current study was to enhance its recognition rate.

Some of the most popular artificial neural network technologies are the back propagation neural network (BPNN), support vector machine (SVM), and self-organizing linear output (SOLO). Both the BPNN and SVM can be regarded as nonlinear regression analysis techniques. However, they have different mechanisms for deciding the form and parameter of the regression equation. The SOLO is a linear regression analysis technique in which nonlinear problems are described using the concept of piecewise linearity. After considering the rescue and detection missions that the Polaris is involved in, the researchers decided to use the BPNN technique in developing a system to enhance the instantaneous recognition rate for underwater images; this system was developed on the theoretical basis that object shape affects the manual recognition rate [2]–[21], [24], [25].

II. IMAGE PRE-PROCESSING

A. IMAGE FILTERING AND REMOVE NOISE

After underwater images were obtained, they were subjected to preprocessing, which involved the use of a low-pass filter to remove noise (or high-frequency noise) from the shape images. The method used in the current study was the average masking method, which involves summing all the grayscale values in the mask, obtaining an average value, and writing the value in the corresponding pixel point. This is therefore known as the moving average filter, and noise can be removed using this method. In the low-pass filter, the average value is written in the corresponding pixel at the middle of the mask. Therefore, only the value of one pixel can be handled at a time; the adjacent pixel is processed during the next mask movement. Therefore, the grayscale value of a pixel may be or may not be the same as the grayscale value of its surrounding pixels.

1/9	1/9	1/9
1/9	1/9	1/9
1/9	1/9	1/9

For a high-pass filter, high-frequency sharpening of character images is performed. The sum of weights of the mask processed using a high-pass filter is 1. The purpose of high-pass filter processing is to highlight the edge portions of an image as well as the finer and complex areas of the image. The image may become clearer after the processing. However, the added clarity comes with the accompanying problem of increased high-frequency noise, causing additional uncertainty in terms of analysis and processing. Therefore, a suitable sharpening value must be identified.

-1	-1	-1
-1	9	-1
-1	-1	-1

B. BINARIZATION

Binarization of an image involves using a threshold gray level to segment an image into two parts. After removing high-frequency noise using the low-pass filter (or performing high-frequency sharpening of the image using the high-pass filter). In order to focus on the segmentation of the image that must be recognized. An image can be segmented into the background and objects. In the case of the image to be identified in this study, the target to be recognized is the object, whereas the background is the environmental image on the iron plane. Therefore, when identifying the object, the background image must be filtered out first, just leaving the objects in the image. To ease the segmentation process, the gray level threshold value between the background and object to be identified was selected in accordance with the study published by J.J. Thomson Physics Lab, University of Reading; for gray level ≥ 50, the value was set to 255; for gray level < 50, the value was set to 0.

C. EDGE DETECTION

After the static images were subjected to image morphology operations, the researchers performed edge detection on the objects to be identified to ease the subsequent extraction of eigenvalues. The Sobel method was used for this purpose. In the Sobel method, the gradient value of each pixel along the x and y directions is calculated on the basis of the Sobel masks shown in the following figure; Gx and Gy represent the mask computation values obtained in the x and y directions, respectively. Therefore, the calculation formula for the radial vector of each pixel is $G(x,y) = \sqrt{}(Gx+Gy)$.

-1	-2	-1		-1	0	1
0	0	0		-2	0	2
1	2	1		-1	0	1

To rapidly search for the characteristic eigendata in the image, we made use of the concept of image interconnection, which involves framing the region with the maximum area (grayscale value not equivalent to 0) in the image. The following figure is a mask that makes use of the interconnection concept; e is taken as the center (i.e., the grayscale value of e is not equivalent to 0). As long as the grayscale value of point d, b, f, or h is not equivalent to 0, the point is considered to belong to the same region as e. The region with the maximum area in the image is framed using the concept of image interconnection.

a	d	g
b	e	h
c	f	i

The design process for the image processing program (C++ Builder) was as follows:

- Load the image (self-define the pathway and the format).
- Place the pixels of the image in the array of the memory cell.
- Extract each pixel for computation or transformation by using various filter methods.
- Once the computations are complete, paste the obtained results back into the original memory cells.

III. ARTIFICIAL NEURAL NETWORK (BPNN) RECOGNITION

The learning process of the algorithm consists of forward and backward information propagation (Figure 2(a), Figure 2(b), Figure 2(c)). In the process of forward propagation, weights are calculated at the hidden layer through the input layer. After transformation through the activation function, the network output value is calculated and propagated toward the output layer. The neurons in each layer only affect the status of the neurons in the next layer. If the target output values cannot be obtained at the output layer, backpropagation is activated; the error signal is backpropagated along the original connection channel. The weight and bias values of neurons in each level are modified continually until an error value within the tolerable error range is acquired.

A. RECOGNITION MODEL

Recognition systems comprise components such as underwater cameras, personal computers, and graphic cards. To simulate the external appearance and structures of common underwater objects, images of 12 basic shapes were captured using underwater cameras. Showing a low-resolution environment, the blurred images were used as the identification targets (Figure 3).

The neural network model was used to replace the human eyes and brain in recognizing the 12 target shapes. The recognition program was constructed using the C++ programming language, and complete target images were obtained using image preprocessing methods (e.g., filtering, binarization, and edge detection). Finally, the invariant moment was adopted as the main characteristic for recognition and input unit for the neural network.

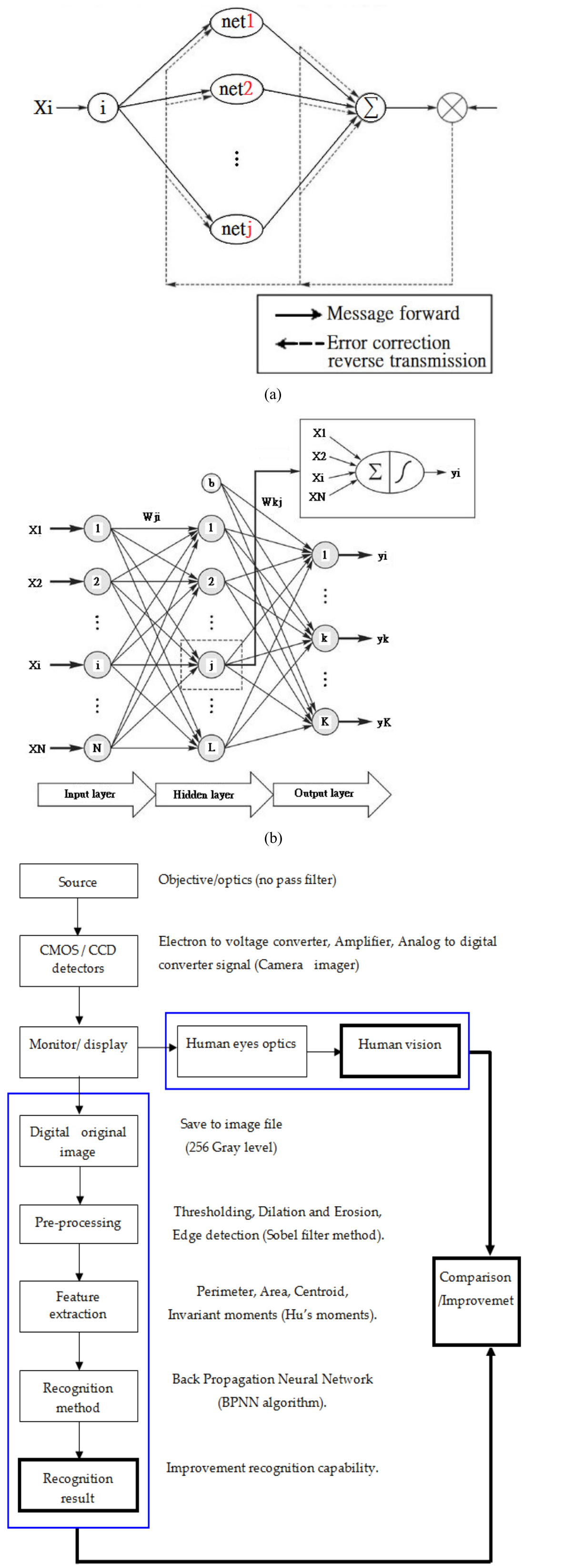


FIGURE 2. (a) Transfer direction of the BPNN. (b) Illustration of the BPNN. (c) Image recognition processes.

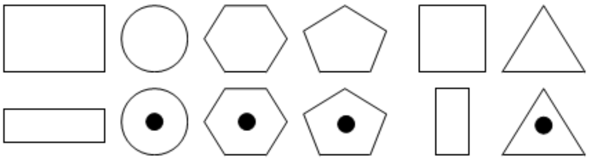


FIGURE 3. Basic shapes (12 types).

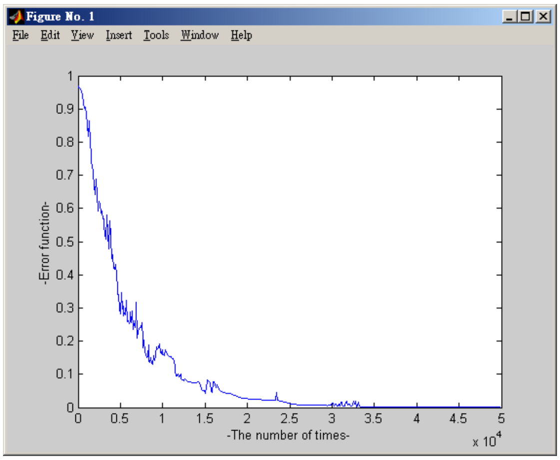


FIGURE 4. Neural network training curve.

The learning process of the algorithm consisted of forward and backward information propagation (Fig. 2). In the process of forward propagation, weights are calculated at the hidden layer through the input layer. After transformation through the activation function, the network output value is calculated and propagated toward the output layer. The neurons in each layer only affect the status of the neurons in the next layer. If the target output values cannot be obtained at the output layer, backpropagation is activated; the error signal is back propagated along the original connection channel. The weight and bias values of neurons in each level are modified continually until an error value within the tolerable error range is acquired. In the algorithm, the input value of the jth neuron of layer n is the nonlinear function of the output value of neurons in layer n-1:

$$y_i^n = f(net_j^n), \quad net_j^n = \sum_i w_{ji}^n y_i^{n-1} - b_j^n$$

$$E = (1/2)\sum_k (d_k - y_k)^2, \quad w_{ji} = -\rho \frac{\partial E}{\partial w_{ji}} \tag{1}$$

where y is the output value of layer n, serving as the input value for layer 1; f is the activation function; net_j^n is the accumulated weight of the output value for layer n – 1; w_{ji}^n is the connection weight between the jth neuron of layer n and ith neuron of layer n-1; b_j^n is the bias value of the j_{th} neuron in layer n; E is the error function; d_k is the target output value of the kth neuron; y_k is the network output value of the kth neuron of the output layer; and ρ is the learning rate, the value of which decides the modification magnitude of the method of steepest descent.

Because the algorithm is a type of supervised learning method, the purpose of learning is to reduce the differences between the network output values and target output values.

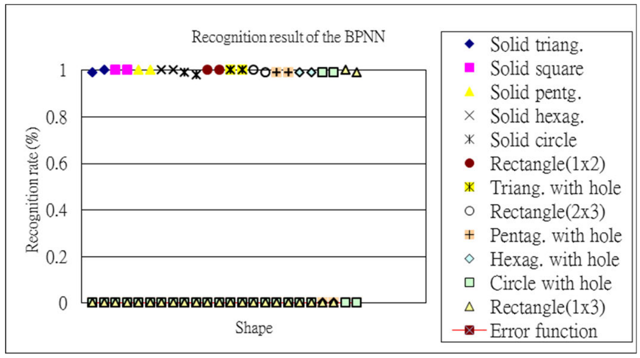


FIGURE 5. Neural network recognition results.

TABLE 1. Parameter settings of the neural network.

Input unit	7 units (Φ1,Φ2,Φ3,Φ4,Φ5,Φ6,Φ7)
Hidden unit	10 units
Output unit	12 units (12 shapes)
Number of computations	50000 times
Learning rate	0.5
Inertia item	0.5
Number of training and testing samples	72 training samples plus 48 training and

TABLE 2. Normalized invariant moments of the shape images (12 types).

Sample\ Invariants	Φ1	Φ2	Φ3	Φ4	Φ5	Φ6	Φ7
Solid triangle	0.1115	0.3757	0.4791	0.4534	0.9290	0.6451	0.9320
Solid square	0.1118	0.3470	0.4633	0.4551	0.9123	0.6289	0.9385
Solid pentagon	0.1104	0.3707	0.4209	0.4391	0.8755	0.6320	0.8750
Solid hexagon	0.1114	0.3482	0.4519	0.5162	0.9546	0.6932	0.9591
Solid circle	0.1113	0.3514	0.4519	0.4606	0.9438	0.6425	0.9266
Rectangle(2x3)	0.1109	0.3224	0.4323	0.4679	0.9285	0.6525	0.9275
Triangle with hole	0.1094	0.3448	0.4061	0.4096	0.8346	0.6124	0.8063
Rectangle(1x2)	0.1120	0.3329	0.4757	0.4756	0.9582	0.6579	0.9528
Pentagon with hole	0.1102	0.3358	0.4350	0.4675	0.9454	0.6380	0.9194
Hexagon with hole	0.1108	0.3581	0.4277	0.4349	0.8731	0.6382	0.8646
Circle with hole	0.1115	0.3361	0.4461	0.4681	0.9439	0.6435	0.9296
Rectangle(1x3)	0.1098	0.3194	0.4749	0.4866	1.0000	0.6482	0.9679

Therefore, the learning process of the network is minimization of the error function E. We employed the method of steepest descent to search for the best solution of E, or in other words, the smallest sum of the squared errors. The network slightly adjusted the weight each time a training data point is input. The magnitude of adjustment and error function were proportional to the sensitivity of the connection weight values, as shown in the following equation. With recurring calculation and training process, the minimum value of E was identified when a state of convergence was achieved to obtain the optimal recognition results.

B. EIGENVALUE SELETION

Next, the effective eigenvalues were identified and used as the inputs for the algorithm. We trained the neural network such that it could recognize the object to be identified based on this eigenvalue, and the recognition results acquired using human vision and the algorithm under various conditions could be compared. In this study, the researchers used Hu's moments

TABLE 3. **Weights and threshold values after completion of neural network training.**

1.68	−8.72	−13.83	−8.74	−6.92	−34.19	8.72
7.33	110.91	92.12	−62.50	−16.41	−71.13	33.42
−0.59	131.01	49.88	−75.79	−20.71	−13.83	−39.06
−7.40	−1.72	107.04	11.41	7.31	3.92	36.38
2.30	−22.12	−9.55	−5.56	−7.59	−21.23	0.11
0.87	8.04	−22.32	−9.86	19.07	−32.31	−25.94
10.14	−73.09	66.95	−29.12	7.88	44.66	−89.73
0.63	−2.12	−1.07	−5.80	7.72	−15.69	2.45
−5.27	22.21	−79.61	12.62	60.66	−49.64	−23.00
−4.94	49.62	−54.48	−0.68	75.36	−67.15	−11.52

Weighting: [Hide][Output] =[10][12]

−11.40	15.19	17.52	11.51	−12.47	−4.36	−9.74	−3.31	−11.98	−3.13
6.19	26.06	−6.95	−16.59	5.60	−7.21	−10.87	3.76	−25.64	−21.05
2.53	16.50	20.42	−44.57	−2.03	−3.18	−18.20	−1.19	2.39	−3.31
−14.40	−13.41	−18.85	10.18	−10.86	−12.54	−16.50	−7.29	−11.04	−13.92
−7.74	44.42	−1.13	−19.65	−15.18	−3.91	−27.51	0.91	−7.29	15.99
1.79	−6.68	−18.07	−54.88	7.43	−5.38	−3.39	−1.85	−7.61	−9.09
5.13	0.87	6.55	−19.73	5.61	8.42	12.97	−0.94	4.34	3.02
−2.84	15.46	−25.71	18.55	−4.47	−8.09	−20.28	−0.99	−19.69	−16.99
6.57	−7.35	−2.26	−40.04	5.68	10.62	−25.38	1.52	16.85	14.73
−6.70	−4.68	12.53	−25.58	−6.92	−16.24	14.63	−3.79	−18.08	−17.86
−0.82	7.34	−16.76	−6.86	9.22	−0.68	30.17	0.86	−5.84	−32.22
−5.34	−13.40	−8.12	−1.57	−17.23	0.67	0.17	14.65	−1.41	17.68

Bias hide

−30.89	17.68	−28.70	92.57	−32.02	−34.36	−49.70	−4.28	−16.06	8.10

Bias output

23.70	14.68	25.34	−11.83	36.58	−7.93	30.13	11.18	13.32	−1.96	4.89	15.23

TABLE 4. **ANOVA results for human recognition of various easily confused shapes.**

One-way ANOVA

Recognition method	Shape type	Sum	Average	Variation
Triangle	5	3.31	0.551667	0.025857
Hexagon	5	4.15	0.691667	0.049457
Square	5	4.75	0.791667	0.037577
Rectangle(1x2)	5	5.03	0.838333	0.025977
Rectangle(1x3)	5	5.29	0.881667	0.007777
Rectangle(2x3)	5	5.42	0.903333	0.002627

Variation source	SS(Sum of squares)	Degree of freedom	SS/Degree of freedom	F-value	P-value	Threshold value
Between group	0.533681	4	0.106736	4.290324	0.004613	2.533555
Within group	0.74635	24	0.024878			

TABLE 5. **ANOVA analysis of human and BPNN recognition results.**

Recognition method	Shape type	Sum	Average	Variation
Human eye	5	4.65	0.775	0.01779
BPNN	5	5.97	0.995	0.00003

Variation source	SS(Sum of squares)	Degree of freedom	SS/ Degree of freedom	F-value
Between group	0.1452	1	0.1452	16.2963
Within group	0.0891	8	0.00891	
	P-value	0.002373	Threshold value	4.964603

TABLE 6. **Parameter settings of the neural network.**

Input unit	7 units (Φ1,Φ2,Φ3,Φ4,Φ5,Φ6,Φ7)
Hidden unit	8 units
Output unit	8 units (8 types of figure)
Number of computations	50000 times
Learning rate	0.5
Inertia item	0.5
Number of training and testing samples	32 training samples plus 32 testing and verification samples

theory [1] in which the absolute value of the image's moment is employed as the basis of object recognition (Table 2).

Many of the geometrical characteristics (e.g., size, location, direction, and shape) of a planar surface are related

TABLE 7. **Normalized invariant moments of the underwater images (eight types).**

Sample\ Invariants	Φ1	Φ2	Φ3	Φ4	Φ5	Φ6	Φ7
1	0.1049	0.2563	0.3973	0.3960	0.7885	0.5293	0.7834
2	0.1102	0.2674	0.4921	0.4194	0.8863	0.5546	0.8822
3	0.1075	0.2363	0.4111	0.4047	0.8301	0.5393	0.7918
4	0.1106	0.2699	0.4257	0.3852	0.8103	0.5206	0.7915
5	0.1093	0.2650	0.4345	0.4363	0.8715	0.5950	0.8512
6	0.1116	0.2473	0.4667	0.4861	0.9821	0.6107	0.9886
7	0.1083	0.2411	0.4223	0.4252	0.8595	0.5478	0.8563
8	0.1159	0.2779	0.4625	0.4590	1.0000	0.6372	0.9099

TABLE 8. **Weights and bias after completion of neural network training.**

−12.47	0.54	−0.62		−12.22	−41.15	−25.95	61.24
8.91	5.05	7.34		−14.88	11.55	−9.98	35.66
−0.50	6.34	9.46		6.27	9.40	7.23	36.60
−8.08	18.12	−23.59		16.65	−5.27	13.18	−1.84
−7.17	3.24	−4.46		5.16	−6.60	2.82	−24.34
19.98	−7.92	12.48		16.61	14.28	−21.22	−4.25
7.94	19.68	26.39		15.81	−6.16	14.84	3.10
10.35	−9.43	7.62		4.28	7.38	−3.86	−15.14

−1.91	0.73	−0.32	0.32	−8.84	−17.95	12.29	−11.05
13.60	−7.46	1.58	21.42	23.80	−2.25	−15.67	3.37
−2.99	−23.10	7.55	−8.95	−7.35	−24.33	23.27	−12.00
3.48	−5.03	3.10	−7.03	2.75	−10.33	1.54	10.81
18.16	−10.89	−26.05	29.08	−5.11	2.53	−5.41	−5.61
−0.49	−14.93	−33.49	−4.07	−5.19	2.87	−9.31	−6.22
−14.82	−43.63	17.65	−22.73	−12.32	4.39	−18.43	−12.98
−5.33	18.87	−1.57	−21.59	−0.46	−13.40	−16.04	14.08

Bias hide

3.48	14.65	13.26	4.32	−0.09	7.78	30.67	4.91

Bias output

−1.80	14.99	−3.35	1.89	−2.26	−3.55	−8.77	−0.51

to the parameter moment. The (p+q) moment of the two-dimensional and binary figure b in the (m,n) region is defined as follows:

$$m_{pq} = \sum_{m=0}^{m-1}\sum_{n=0}^{n-1} m^p n^q b(m,n) \quad p, q = 0, 1, 2, \ldots,$$

$$\mu_{pq} = \sum_{m=0}^{m-1}\sum_{n=0}^{n-1} (m-\bar{x})^p (n-\bar{y})^q b(m,n)$$

$$\bar{x} = \frac{m_{10}}{m_{00}}, \quad \bar{y} = \frac{m_{01}}{m_{00}}$$

m_{00}: the pixel is equal to 1 of total

m_{10} : the moment of the ′x′ direction

m_{01} : the moment of the ′y′ direction

$$\eta_{pq} = \frac{\mu_{p,q}}{\mu_{00}^{(p+q+2)/2}}$$

(Normalize of the central moment) (2)

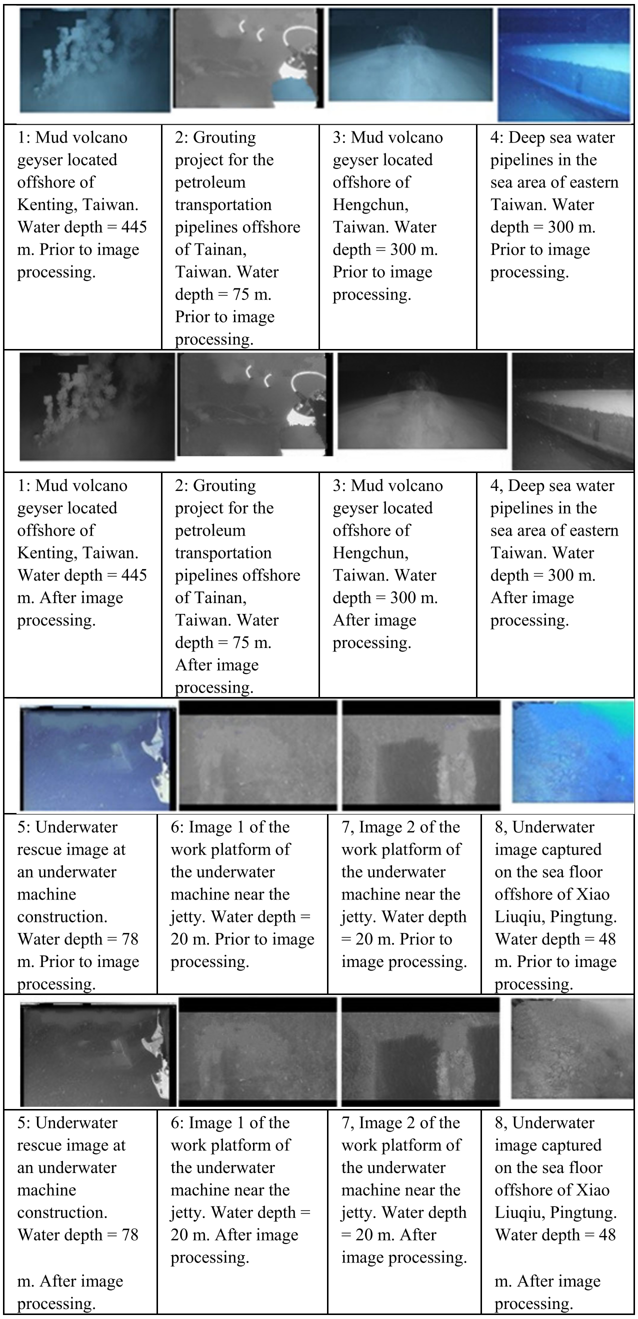


FIGURE 6. Comparison of eight original underwater images with the corresponding grayscale images produced through image preprocessing.

The normalized second and third moments can be combined and used in the derivation of seven moment invariants Φi (i = 1,2,…,7), which can be used to describe the geometrical characteristics related to the planar surface (e.g., shape, size, location, and direction). First moments are related to shape, second moments to the level of expansion of curves, and third moments to a curve's symmetry. The group of invariant moments does not change under the influence of translation, rotation, or size change during image processing [1].

$$\phi_1 = \eta_{20} + \eta_{02}$$

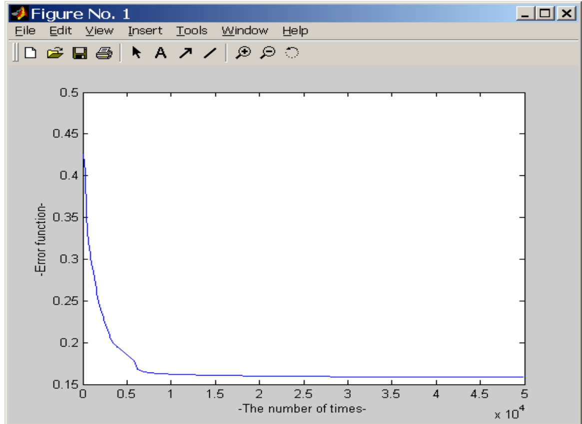


FIGURE 7. Neural network training curve.

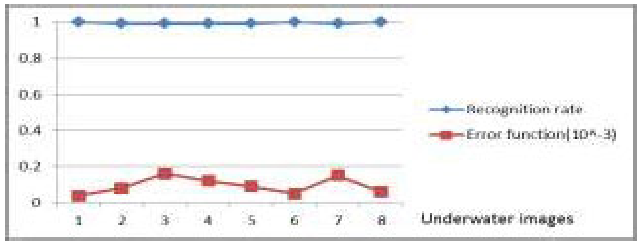


FIGURE 8. Results of neural network recognition.

TABLE 9. Results of recognition by the neural network.

Recognition rate / Underwater images	1	2	3	4	5	6	7	8	Error function
1	**1.0**	0.0	0.0	0.0	0.0	0.0	0.0	0.0	$0.04x10^{-3}$
2	0.0	**0.9**	0.0	0.0	0.0	0.0	0.0	0.0	$0.08x10^{-3}$
3	0.0	0.0	**0.9**	0.0	0.0	0.0	0.0	0.0	$0.16\ x10^{-3}$
4	0.0	0.0	0.0	**0.9**	0.0	0.0	0.0	0.0	$0.12\ x10^{-3}$
5	0.0	0.0	0.0	0.0	**0.9**	0.0	0.0	0.0	$0.09\ x10^{-3}$
6	0.0	0.0	0.0	0.0	0.0	**1.0**	0.0	0.0	$0.05\ x10^{-3}$
7	0.0	0.0	0.0	0.0	0.0	0.0	**0.9**	0.0	$0.15\ x10^{-3}$
8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	**1.0**	$0.06\ x10^{-3}$
3.48	14.65	13.26	4.32	−0.09	7.78	30.67	4.91		
Bias output									
−1.80	14.99	−3.35	1.89	−2.26	−3.55	−8.77	−0.51		

$$\phi_2 = (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2$$
$$\phi_3 = (\eta_{30} - 3\eta_{12})^2 + (\eta_{03} - 3\eta_{21})^2$$
$$\phi_4 = (\eta_{30} + \eta_{12})^2 + (\eta_{03} + \eta_{21})^2$$
$$\begin{aligned}\phi_5 = {} & (3\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 \\ & - 3(\eta_{21} + \eta_{03})^2] \\ & +(3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03}) \times [3(\eta_{30} + \eta_{12})^2 \\ & -(\eta_{21} + \eta_{03})^2]\end{aligned}$$
$$\begin{aligned}\phi_6 = {} & (\eta_{20} - \eta_{02})[(\eta_{30} + \eta_{12})^2 \\ & -(\eta_{21} + \eta_{03})^2] + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03})\end{aligned}$$
$$\begin{aligned}\phi_7 = {} & (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] \\ & +(3\eta_{12} - \eta_{30})(\eta_{21} + \eta_{03}) \times [3(\eta_{30} + \eta_{12})^2 \\ & - (\eta_{21} + \eta_{03})^2]\end{aligned} \quad (3)$$

C. RESULTS OF THERMAL IMAGE RECOGNITION

Use Thermal-CAM to take images of objects in the process of cooling. Among the 120 thermal images, 72 were used for neural network training and 48 for testing and verification. Table 1 presents the parameter settings of the neural network; Table 2 details the normalized invariant moments of 12 types of shape image; and Table 3 presents the weights and threshold values after completion of neural network training. Figure 4 shows the neural network training curve, and the neural network recognition results are displayed in Figure 5.

IV. VERIFICATION OF THE RECOGNITION RESULTS USING ANALYSIS OF ANALYSIS OF VARIANCE (ANOVA)

The ANOVA analysis results for human recognition of the shapes are shown in Table 4. Both the threshold value and P value were < 0.05, indicating that shape recognition differs slightly from human to human, and the difference is slightly significant.

The ANOVA results for recognition performed by the neural network are shown in Table 5. The F value was much higher than the threshold value, and the P value was much smaller than 0.05; this indicated significantly superior results for neural network recognition than for human recognition.

V. ANALYSIS OF THE UNDERWATER IMAGE RECOGNITION RESULTS

Based on the eight types of underwater image captured by the Polaris (Fig. 1), a more blurred image was produced for each underwater image by using the Gaussian blurring method (number of pixels = 5); this resulted in 16 raw underwater images. Next, each underwater image was rotated and sampled at 0°, 90°, 180°, and 270°, yielding 64 raw underwater images. Subsequently, low-resolution object identification based on blurred images was conducted for images captured in an underwater environment. A comparison of the original underwater images (eight types) with the grayscale images produced through image preprocessing is shown in Figure 6. After low-resolution underwater image recognition was improved using the neural network, an average recognition rate of 95% or higher could be achieved. The results are shown in Tables 6-9. The neural network training curve and identification results are displayed in Figure 7 and Figure 8.

VI. CONCLUSION

Regardless of the type of underwater detection vehicle be it an oceanographic research vessel or unmanned underwater vehicle such as a remotely operated vehicle, autonomous underwater vehicle, or autonomous underwater glider the ability to obtain underwater images, instantaneously identify the objects in them, and achieve a high discrimination rate is critical for underwater surveys.

Through image pre-processing, moment invariant features, and thermal image simulation of underwater images, after training the BPNN to perform the recognition model, we successfully solved the problem of underwater image recognition difficulties. In this study, recognition of shapes in low-resolution underwater images was optimized using a combination of neural network and human vision characteristic techniques. After the recognition of shapes in low-resolution underwater images had been improved using a neural network (with 50000 iterations), an average recognition rate of 95% was achieved. ANOVA indicated that compared with human recognition, recognition using the neural network resulted in a significantly higher recognition rate.

Many of the above causes will distort and decrease the qualities of the original images and many important information of futures of the original images will lose. We list 8 main types of these and they are not only limited to these 8 types. So we develop the empirical novel algorithm-BPNN to remedy this problem.

We know a lot of famous network, such as AlexNet, VGG, Inception and ResNet, they are based on empirical intuition to succeed, after success, people try to build the complete theory. Neural network is a very creative, vigorous and vital research field, any novel technological and theoretical idea would be possible to bring new resolution. We publish our research on this solid example maybe could intrigue many scientists to apply our novel idea to many different research field, so we regard it as a valuable work. And of course, in the future, when we have enough large data set, we will compare it with CNN and to see if any novel work we can do.

CNN is one of the effective methods for image recognition, but when the amount of effective data is not large enough, the CNN model will have an overfitting problem, that is, it cannot grasp the features. The CNN training process uses samples to slowly adjust the model, and the samples are too few. In the case of CNN, this will cause CNN to only recognize specific samples. For other pictures that have not been seen, CNN may not be able to recognize because of the lack of training features. To train the CNN model to be practical, it needs considerable The number of samples is limited by the underwater environment, so it is impossible to obtain a large number of samples, so this article is for the study of limited samples. Because the number of samples is too low, CNN is not suitable. When the amount of data is large enough, The CNN method will be used as a topic for future research.

REFERENCES

[1] A. McAuley, A. Coker, and K. Saruhan, "Effect of noise in moment invariant neural network aircraft classification," in *Proc. IEEE Nat. Aerosp. Electron. Conf.*, Jan. 1999, pp. 743–749.

[2] Y.-C. Fang and B.-W. Wu, "Neural network application for thermal image recognition of low-resolution objects," *J. Opt. A, Pure Appl. Opt.*, vol. 9, no. 2, pp. 134–144, Feb. 2007.

[3] B. W. Wu and Y. C. Fang, "Prediction of the thermal imaging minimum resolvable (circle) temperature difference with neural network application," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 30, no. 12, pp. 2218–2228, Dec. 2008.

[4] L.-S. Chang and B.-W. Wu, "A new application of human visual simulated images in optometry services," *J. Opt. Soc. Korea*, vol. 17, no. 4, pp. 328–335, Aug. 2013.

[5] B. W. Wu and C. Lin-Song, "New application of human eye model in the verification of spot diagram with refractive diopter," *Optik*, vol. 125, pp. 1784–1788, Oct. 2014.

[6] B. W. Wu and Y. C. Fang, "Applications of neural networks in human shape visual perception," *J. Opt. Soc. Amer. A, Opt. Image Sci.*, vol. 32, pp. 2338–2345, Oct. 2015.

[7] D. Sedra and G. Huang, K. Weinberger, and Y. Z. Liu, "Deep networks with stochastic depth," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, 2016, pp. 646–661.

[8] A. R. Backes and J. J. de Mesquita Sá Junior, "LBP maps for improving fractal based texture classification," *Neurocomputing*, vol. 266, pp. 1–7, Nov. 2017.

[9] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proc. CVPR*, 2017, pp. 4700–4708.

[10] A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, and H. Adam, "MobileNets: Efficient convolutional neural networks for mobile vision applications," 2017, *arXiv:1704.04861*.

[11] C. Chuan-Yu and W. Wei-Chun, "Integration of CNN and faster R-CNN for tire bubble defects detection," in *Proc. Int. Conf. Broadband, Wireless Comput. Commun. Appl.*, 2018, pp. 285–294.

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

[13] G. Ciocca and P. Napoletano, "CNN-based features for retrieval and classification of food images," *Comput. Vis. Image Understand.*, vol. 1, pp. 176–177, Oct. 2018.

[14] P. Deeprasertkul and W. Praikan, "An application of numbers and characters recognition and classification on radar images using for flood monitoring," in *Proc. 3rd Int. Conf. Comput. Commun. Syst. (ICCCS)*, Apr. 2018, pp. 222–226.

[15] X. Xu, J. Zhou, and H. Zhang, "Screen-rendered text images recognition using a deep residual network based segmentation-free method," in *Proc. Int. Conf. Pattern Recognit.*, Beijing China, 2018, pp. 2741–2746.

[16] H.-Q. Bong, Q.-B. Truong, H.-C. Nguyen, and M.-T. Nguyen, "Vision-based inspection system for leather surface defect detection and classification," in *Proc. 5th NAFOSTED Conf. Inf. Comput. Sci. (NICS)*, Nov. 2018, Art. no. 18374617.

[17] B. B. Traore, "Deep convolution neural network for image recognition," *Ecol. Inf.*, vol. 48, pp. 257–268, Jun. 2018.

[18] I. M. Dheir, A. Mettleq, and A. Elsharif, "Classifying nuts types using convolutional neural network," *Int. J. Acad. Inform. Syst. Res.*, vol. 3, pp. 12–18, Oct. 2019.

[19] M. W. Akram, G. Li, and Y. Jin, "CNN based automatic detection of photovoltaic cell defects in electroluminescence images," *Energy*, vol. 189, Oct. 2019, Art. no. 116319.

[20] C.-C.-J. Kuo, M. Zhang, S. Li, J. Duan, and Y. Chen, "Interpretable convolutional neural networks via feedforward design," *J. Vis. Commun. Image Represent.*, vol. 60, pp. 346–359, Apr. 2019.

[21] P. P. Banik and R. Saha, "An automatic nucleus segmentation and CNN model based classification method of white blood cell," *Expert Syst. Appl.*, vol. 149, Oct. 2020, Art. no. 113211.

[22] R. Gonzalez and R. Wood, *Digital Image Processing*. Upper Saddle River, NJ, USA: Prentice-Hall, 2006.

[23] J.-S. Lee, "Digital image enhancement and noise filtering by use of local statistics," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-2, no. 2, pp. 165–168, Mar. 1980.

[24] L. Chen, Z. Liu, L. Tong, Z. Jiang, S. Wang, J. Dong, and H. Zhou, "Underwater object detection using invert multi-class Adaboost with deep learning," in *Proc. Int. Joint Conf. Neural Netw. (IJCNN)*, 2020, pp. 1–8.

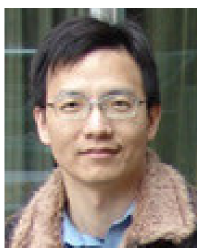
[25] L. Chen, Z. Jiang, L. Tong, Z. Liu, A. Zhao, Q. Zhang, and H. Zhou, "Perceptual underwater image enhancement with deep learning and physical priors," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 31, no. 8, pp. 3078–3092, Oct. 2021.

BO-WEN WU was born in Taiwan. He is currently an Associate Professor with the Department of Optometry, Yuanpei University of Medical Technology. His research interests include image processing, artificial visual recognition, optical design, human visual model, 3D stereo vision, and optometry.

YI-CHIN FANG is currently a Professor at the National Kaohsiung University of Science and Technology. His research interests include optics engineering and design involved in the development of image, non-image systems, and infrared systems. He was a Featured Editor of *Applied Optics*, OSA, IEEE, *Applied Sciences*, and *Materials*.

CHAN-CHUAN WEN is currently the Deputy Professor with the National Kaohsiung University of Science and Technology. He has engaged in both underwater surveying and integrated navigation system design for more than 30 years. He is also the Chairman of the Kaohsiung Marine Engineering Association that is the largest ocean engineering in Taiwan and now has paid a major attribution to the Taiwanese offshore wind farm development.

CHAO-HSIEN CHEN received the Ph.D. degree in electro-optical engineering from the NCTU. In 1999, he joined the Industrial Technology Research Institute, where he worked on optical design and measurement. In 2005, he joined the Mechanic's Department, National Kaohsiung University of Science and Technology, where he is currently a Professor. His research interests include optical aberration theorems, optical measurement, lens design, and illumination systems.



HSIAO-YI LEE is currently a Professor at the Optoelectronic System Design Laboratory, Electrical Engineering Department, National Kaohsiung University of Science and Technology. He has published more than 70 peer-reviewed journal articles and held more than 70 patents. His major research interests include optics design, lighting design, and optoelectronic systems.

SHUN-HSYUNG CHANG is currently a Professor with the Department of Microelectronics Engineering, National Kaohsiung University of Science and Technology. He has published over 250 scientific-technical publications, among which are 15 monographs. His research interests include electrical engineering, electronics and communication engineering, signal processing and array processing, underwater and acoustic communication, underwater acoustics, and Kalman filtering and detection. He is also the Editor-in-Chief of the *Journal of Ocean and Underwater Technology*.