Webology (ISSN: 1735-188X)
Volume 18, Number 6, 2021

A Study Of Image Identification For Underwater Images Using ANN Technique

[1]Dr. A. Divya , [2]Dr. M. Sathya

Assistant Professor Department of Computer Science Valluvar College of Science and Management Karur.

Assistant Professor Department of Computer Science Valluvar College of Science and Management Karur.

Abstract

Research on a Indian oceanographic research vessel, Polaris includes developing holographic image software for the ship. With a weight of 260 tonnes, a length of 36.98 metres, and a breadth of 6.80 metres, this ship can reach 11 knots of speed. For this reason, it has performed a number of underwater rescue and exploratory operations. A wide range of aspects, includes water temperature, water deflection, and phantom circumstances, as well as the weather, oceanic current, algae, and light replication from schools of fish, could obstruct underwater research efforts. For marine exploration, immediate picture analysis is essential. Scientists built Artificial Neural Network(ANN)-based credit software to quickly identify an underwater object's category to meet Polaris's need for immediate identification. A neural network was used to improve shape recognition in low-resolution underwater photos, resulting in a 95% recognition rate on average. Using a neural network instead of manual recognition proved to be more accurate.

Keywords: Image processing, Neural network, moment invariants, underwater images, image recognition.

I. Introduction

A non-government oceanographic research vessel, the Polaris is the first of its kind for India[1]. Underwater rescue and exploring missions have taken place on the vessel multiple times. The deepest point in the Indian Ocean is around 8000 metres below sea level, with an average depth of 3900 metres[2], [3]. An extensive search necessitated the use of specialised tools and training. With its international reputation for quality, the Polaris was invited to aid Australia in its rescue attempts at the time[4].

Among the Polaris' many tools and gadgets are an underground water wirelessly functioned vehicle, a variance satellite aligning structure, echo sounders (both solitary and multibeam), side scan sonars, a sub-bottom profiler, a marine magnetometer, a positioning system for the seafloor, sediment corers, and penetrometers[5]–[7]. The Polaris also has a sub-bottom profiler and a penetrometer. There's little emphasis on obtaining stunning underwater photographs with strong contrast, which is the major goal of underwater photographers when they use the Polaris[8]. It is clear from the Polaris's involvement in multiple disaster relief and emergency rescue operations that underwater object detection is more important than other aspects of underwater imaging[9]. According to this definition, a system's recognition rate is the percentage of collected photos that properly identify some sort of underwater item. The current study's goal was to increase the Polaris' identification rate in order to ensure that items are recognised quickly and properly[10], [11]. Regression equations are formed and parameterized in different ways for each method. Piecewise linearity is used to characterise nonlinear situations in SOLO, a linear regression analysis technique[12]. Following an examination of the Polaris's rescue and detection missions, the researchers chose to employ the Back propagation neural network techniques in the construction of a structure to increase the sudden identification rate for submerged images[13].

Image Pre-Processing

Elimination of Noise and Image Filtration

Using low-pass filter to reduce high-frequency sound in the shape images after they were taken underwater was the first step in pre-processing. All the grayscale values of a mask were added together, resulting in mean value and the values was written to the image at the pixel point where a mask was applied[14]. The moving average filter can be used to reduce noise. The average value in a pixel at the centre of the mask is written in the middle of the mask using low-pass filters. Each time the mask is adjusted, the next pixel is processed. This method allows for processing of only one pixel at a time[15]. A pixel's grayscale value can differ from the surrounding pixels depending on lighting conditions.

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

Using a high-pass filter, character images are sharpened at a high frequency. The mask's weights are equal to 1 when processed through high-pass filter. Goal of high-pass filter proceeding to emphasise image's edges as well as its finer and more intricate regions[16]. After the processing, the image may become clearer. Increased high-frequency noise has the unintended side effect of increasing analysis and processing uncertainty. Identifying an appropriate sharpness value is therefore necessary.

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

Binarization

A threshold grey level is used to divide an image into two pieces during binarization. After the low-pass filter has been used to remove the high-frequency noise. It is necessary to distinguish the image's division in order to focus on it. The background and objects in an image can be separated in Photoshop. Objects to be identified in this investigation include a target and a background image that appears on an iron plane[17]. Consequently, the background image must be eliminated before identifying an object. If the object to be detected has a grey level above 50, the threshold value should be set to 255; if the level is below 50, it should be set to zero.

Edge Detection

To expedite the extraction of eigenvalues from images, researchers used edge detection after performing image morphological procedures on the objects to be detected. In order to accomplish this, the Sobel method was applied. This approach uses Sobel masks to compute each pixel's gradient value in both the x axis and y axis; The facade computational values in the x axis and y axis are represented by the variables Gx and Gy, respectively. This means that the radial vector computation formula for each pixel must be the same.

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

Image interconnection was used to quickly locate the image's characteristic eigendata using the notion of using the biggest possible space as a frame in the image. Interconnection is used in the design of the following mask, which features e as the focal point. Points d, b, f, and h are regarded to be in the same region as e as long as their grayscale values are not equal to zero[18]. The notion of image connectivity is used to frame the portion of the image that has the greatest surface area.

a	d	g
b	e	h
c	f	i

Following steps were taken in the creation of the image processing programme (CCC Builder):

- Upload the picture (self-define the path and the arrangement).
- Memory cells can hold the image's pixels in an array.
- Each pixel can be computed or transformed using a variety of different methods.
- When you're done computing, go back and paste your results into the initial memory cells.

ANN recognition

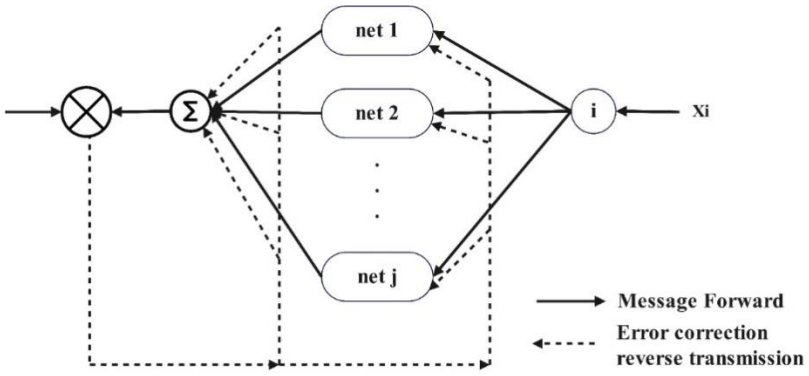
Figures 1(a), 1(b), and 1(c) show the forward and backward propagation of information throughout the algorithm's learning phase. During the forward proliferation operation,

weightage are computed at the concealed layer. The activation function transforms the network's output before sending it to the output layer. There are no connections between the neurons in any one sheet[19]. Backpropagation is launched if the output layer fails to collect the desired output values, and the fault signal is returned via the initial connecting channel.

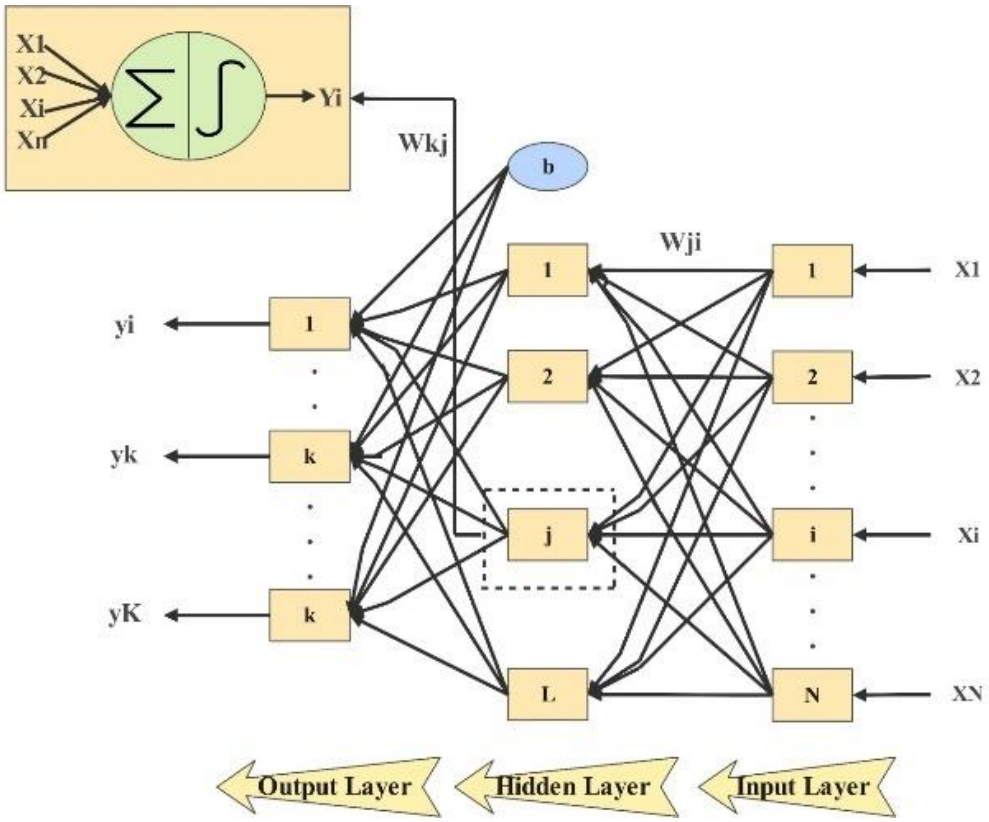
Recognition Model

personal computers, graphics cards and Underwater cameras are all part of a recognition system. Undersea cameras captured images of 12 basic forms in order to approximate the presence and structure of underwater things[20]. The blurred pictures served as identification targets because they depicted a low-resolution environment (Fig 2).

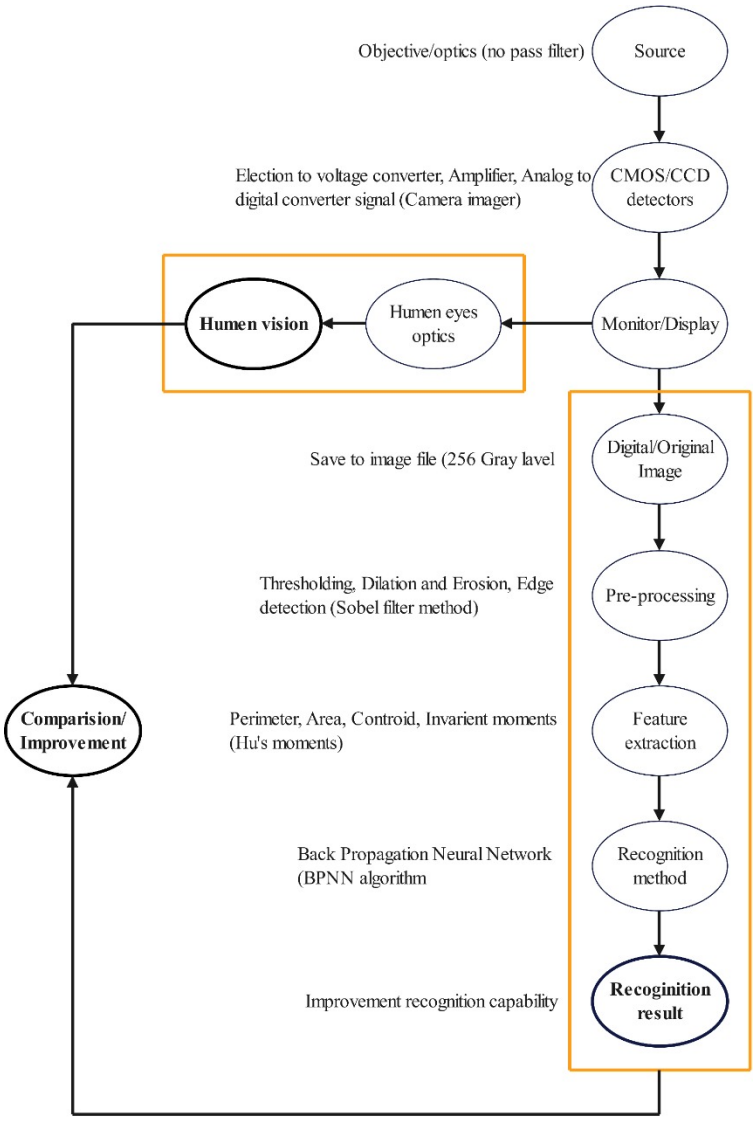
For the 12 target objects, neural network model was used as a surrogate for the human eye and brain. The recognition application's target images were created using image pre-processing and the CCC programming language (e.g., filtering, binarization, and edge detection). A neural network's basic characteristic and input unit for recognition was thus an invariant instant.



(a)


X1
X2
Xi
Xn
Yi
Wkj
b
Wji
1
2
j
L
1
2
i
N
X1
X2
Xi
XN
yi
yk
yK
1
k
k
Output Layer
Hidden Layer
Input Layer

(b)



(c)

Figure 1. (a)The BPNN's transfer. (b) BPNN illustration. (c) Image-recognition systems.

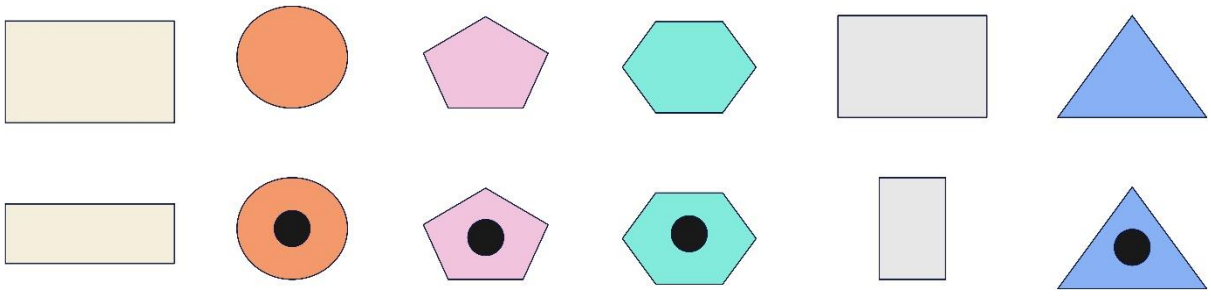


Figure 2. Elementary shapes (12 types).

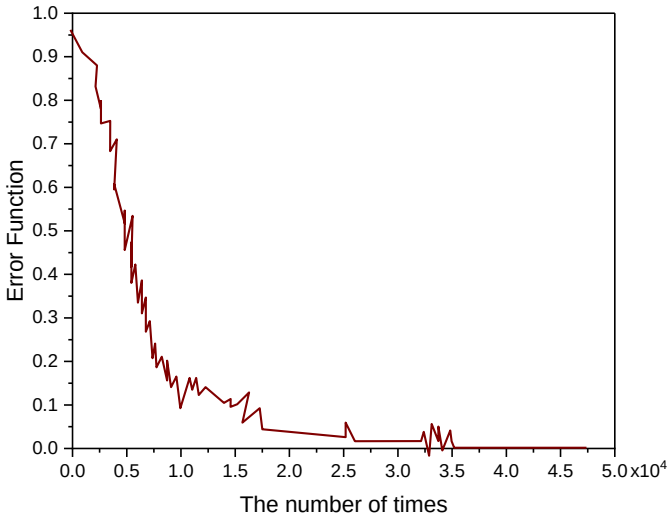


Figure 3. Neural network practicing curve.

The algorithm's training approach included both frontward and rear ward data propagation (Figure. 2). It is during the forward circulation phase that the hidden layer is responsible for determining weights at the input layer[21]. The outcome range of the network is evaluated and delivered to the outlet layer for additional processing when the activation function has completed its change. Only the neurons above it in the layering process have any effect on the state of a neuron. When an error value falls in the acceptable limits, the weights and biases of neurons at each level are adjusted accordingly. The inletrate of the jth nerve cell in layer n is determined using the outlet of neurons in layer n-1.

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

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

The primary objective of the algorithm's learning process is to narrow the discrepancy among the network's output and the desired results.

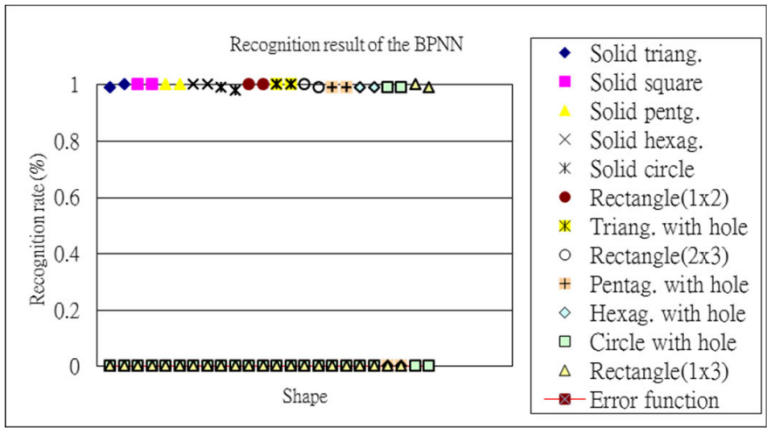


Figure 4. Results of recognition by neural networks .

Table 1. Neural network arrangement.

Input units	7
Hidden units	10
Output units	12
No of computations in times	50,000
Rate of Learning	0.5
Element at Inertia	0.5
Number of testing samples and trainings	72 Samples+ 48 Trainings

Table 2. Moments of invariance of the shape images normalised (12 types).

Sample	**1Φ**	**Φ2**	**Φ3**	**Φ4**	**Φ5**	**Φ6**	**Φ7**
Triangle	0.1084	0.3457	0.4062	0.4097	0.8366	0.6125	0.8073
Square	0.1119	0.3472	0.4634	0.4552	0.9124	0.6288	0.9386
Pentagon	0.1105	0.3708	0.4208	0.4392	0.8756	0.6234	0.8740
Rectangle(1x2)	0.1120	0.3329	0.4758	0.4756	0.9582	0.6679	0.9538
Circle	0.1113	0.3516	0.4528	0.4608	0.9439	0.6426	0.9278
Rectangle(2x3)	0.1108	0.3126	0.4138	0.4679	0.9874	0.6534	0.9341
Solid triangle	0.1115	0.3758	0.4791	0.4534	0.9290	0.6451	0.9322
Solid hexagon	0.1114	0.3482	0.4518	0.5162	0.9547	0.6942	0.9591
Pentagon with hole	0.1103	0.3358	0.4360	0.4676	0.9655	0.6380	0.9184
Hexagon with hole	0.1108	0.3591	0.4277	0.4359	0.8732	0.6383	0.8646
Circle with hole	0.1116	0.3361	0.4561	0.4681	0.9439	0.6435	0.9297

Rectangle(1×3)	0.1098	0.3184	0.4749	0.4876	1.0000	0.6583	0.9689

This means that the network's learning goal is to reduce error function E to the minimum possible value. We utilised the steepest descent approach to treasure the finest solution to E in order to get the lowest possible sum of square errors. The weight of the network is adjusted slightly each time a new piece of training data is added[22]. Error function magnitude and sensitivity of connection weight values are depicted in the figure using this equation. The lowest value of E needed to produce optimal recognition results was discovered using a recurrent computation and training process.

Eigen Value Selection

Essential eigenvalues were determined and employed as input for the procedure following this step. A neural network was trained using this eigenvalue to identify the target item, and the outcomes achieved using humanoid vision and the algorithms in different scenarios were related. A theory put forth by [23]states that the total value of an image's moment determines its quality (Table 2). In many cases, the parameters of a planar surface may be traced back to the constraint moment.

Moment invariants can be utilised to explain the geometric properties of the plane surface by combining the normalised second and third moments (e.g., figure, size, position, and path). Shape, expansion, and symmetry all play a role in the first, second, and third base moments of a curvature. During image processing, collection of invariant moments is not affected by translation, rotation, or size change.

Table 3. Result of neural network training in terms of final weights and threshold.

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**]						

-10.18	14.97	16.96	10.97	-11.84	-3.97	-10.08	-4.12	-12.14	-4.07
5.97	25.96	-7.19	-15.89	6.12	-7.12	−1088	3.77	−25.65	−21.06
2.53	1650	20.42	−44.57	−203	−3.18	−1820	−1.19	239	−331
−14.40	−13.41	−18.85	10.18	−10.86	−12.54	−1650	−7.29	−11.04	−13.92
−7.74	44.42	−1.13	−19.65	−15.18	−391	−2751	0.91	−729	1599
1.79	−6.68	−18.07	−54.88	7.43	−538	−339	−1.85	−7.61	−9.09
5.13	0.87	6.55	−19.73	5.61	8.42	1297	−0.94	434	3.02
−2.84	15.46	−25.71	18.55	−4.47	−8.09	−2028	−0.99	−19.69	−16.99
6.57	−7.35	−2.26	−40.04	5.68	10.62	−2538	1.52	1685	14.73
−6.70	−4.68	12.53	−25.58	−6.92	−1624	14.63	−3.79	−1808	−17.86
−0.82	734	−16.76	−6.86	922	−0.68	30.17	0.86	−5.84	−3222
−534	−13.40	−8.12	−1.57	−17.23	0.67	0.17	14.65	−1.41	17.68
Bias hide									
−308	17.68	−28.7	92.57	−32.0	−34.3	−49.7	−4.28	−16.0	
9		0		2	6	0		6	
Bias Output									
23.70	14.68	25.34	−11.83	36.58	−7.93	30.13	11.18	13.32	−1.96

Table 4. Analysis of Variance results for human acknowledgment of different simply jumbled shapes.

Recognition technique	Shape type	Summation	Mean	Difference
Triangle	5	331	0551667	0025857
Hexagon	5	4.15	0691667	0049457
Square	5	475	0791667	0037577
Rectangle (1X2)	5	503	0838333	0025977
Rectangle (lX3)	5	529	0.881667	0007177
Rectangle (2X3)	5	542	090333	0002627

Table 5. Analysis of Variance of humanoid and BPNN identification outcomes.

Recognition technique	Shape type	Quantity	Mean	Difference
Human eye	5	4.66	0.785	0.01879
Back Propagation Neural Network	5	5.98	0.996	0.00004
Difference source	SS	DOF	SS	F-value

Between	0.1452	1	0.1452	16.2963
group	0.0891	8	0.0891	
Within group	P-value		value of Threshold	
	0.002412		5.4216	

Table 6. Invariants of the Normalized underwater pictures.

Sample	**Φ**						
	1	**2**	**3**	**4**	**5**	**6**	**7**
1	0.1048	0.2564	0.3974	0.3961	0.7886	0.5294	0.7835
2	0.1103	0.2675	0.4922	0.4195	0.8864	0.5547	0.8823
3	0.1076	0.2364	0.4112	0.4048	0.8302	0.5496	0.7919
4	0.1107	0.2689	0.4258	0.3853	0.8104	0.5306	0.7916
5	0.1094	0.2651	0.4386	0.4364	0.8716	0.5951	0.8513
6	0.1117	0.2474	0.4668	0.4862	0.9822	0.6108	0.9987
7	0.1073	0.2412	0.4214	0.4253	0.8596	0.5479	0.8564
8	0.1169	0.2789	0.4626	0.4591	0.9914	0.6373	0.9098

Table 7. A post-trained neural network's weights and bias

−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
13.60	−7.46	1.58	21.42	23.80	−2.25	−15.67
−2.99	−23.10	7.55	−8.95	−7.35	−24.33	23.27
3.48	−5.03	3.10	−7.03	2.75	−10.33	1.54
18.16	−10.89	−26.05	29.08	−5.11	2.53	−5.41

−0.49	−14.93	−33.49	−4.07	−5.19	2.87	−9.31
−14.82	−43.63	17.65	−22.73	−12.32	4.39	−18.43
−5.33	18.87	−1.57	−21.59	−0.46	−13.40	−16.04
Bias hide						
3.54	15.42	14.62	5.18	-0.10	7.94	31.24
Bias output						
-1.94	15.12	-3.42	1.96	-2.34	-3.69	-8.61

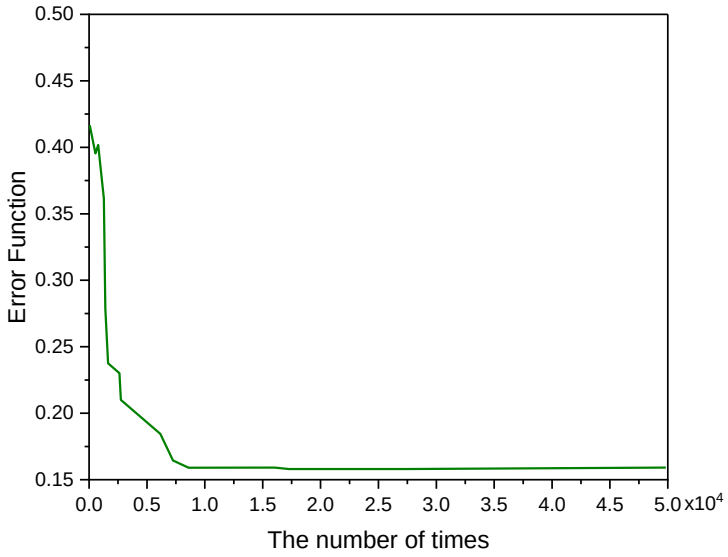


Figure 5. Neural network training curve.

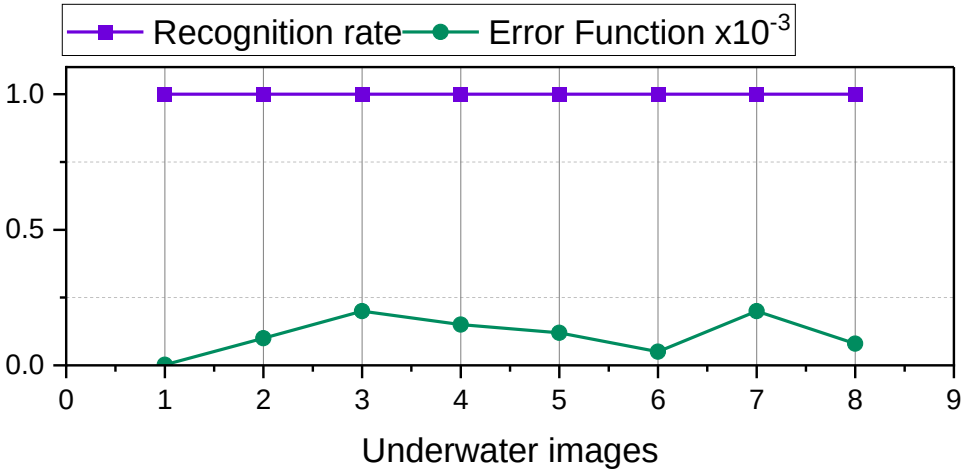


Figure 6. Neural network recognition Results.

Table 9. neural network recognition results.

Recognition rate	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**Error**
1	1.0	0	0	0	0	0	0	0	$0.05x10^{-3}$
2	0	0.9	0	0	0	0	0	0	$0.08\ x10^{-3}$
3	0	0	0.9	0	0	0	0	0	$0.17\ x10^{-3}$
4	0	0	0	0.9	0	0	0	0	$0.13\ x10^{-3}$
5	0	0	0	0	0.9	0	0	0	$0.10\ x10^{-3}$
6	0	0	0	0	0	1.0	0	0	$0.06\ x10^{-3}$
7	0	0	0	0	0	0	0.9	0	$0.16\ x10^{-3}$
8	0	0	0	0	0	0	0	1.0	$0.07\ x10^{-3}$
Bias output									
-1.80	14.99	-3.35	1.89	-2.26	-3.55	-8.77			-0.51

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

$$\phi_2 = (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2$$

$$\phi_3 = (\eta_{30} - 3\eta_{12})^2 + (\eta_{03} - 3\eta_{21})^2$$

$$\phi_4 = (\eta_{30} + \eta_{12})^2 + (\eta_{03} + \eta_{21})^2$$

$$\begin{aligned}\phi_5 = {} & (3\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 \\ & -3(\eta_{21} + \eta_{03})^2] \\ & +(3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03}) \times [3(\eta_{30} + \eta_{12})^2 \\ & -(\eta_{21} + \eta_{03})^2]\end{aligned}$$

$$\begin{aligned}\phi_6 = {} & (\eta_{20} - \eta_{02})[(\eta_{30} + \eta_{12})^2 \\ & -(\eta_{21} + \eta_{03})^2] + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03})\end{aligned}$$

$$\begin{aligned}\phi_7 = {} & (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] \\ & +(3\eta_{12} - \eta_{30})(\eta_{21} + \eta_{03}) \times [3(\eta_{30} + \eta_{12})^2 \\ & -(\eta_{21} + \eta_{03})^2]\end{aligned}$$

Thermal image Recognition Results

Photos of cooled things can be taken with Thermal-CAM. A total of 72 of the 120 thermal images were utilised to train neural networks, while the lasting 48 were employed to test and validate the networks' abilities[24]. After training neural networks, the weights and threshold values for 12 types of form images are provided in Tables 1 and 2, respectively, as indicated in Table 2. Neural network training outcomes are shown in Fig 3, whereas training outcomes are shown in Fig 4.

Authentication of the Recognition Results via ANOVA

Table 4 displays the results of the ANOVA test for shape identification by humans. We found a small but statistically significant difference between humans in terms of shape recognition threshold and P value. Table 5 summarises the ANOVA outcomes for neural network recognition. The F value above the threshold greatly and the P value was significantly less than 0.05, showing that neural network appreciation outperformed human oidac knowledgment significantly.

Underwater Image Recognition Analysis.

Sixteen raw underwater images were taken by the Polaris utilising the Gaussian blurring method (no of pixels D 5) to create eight different sorts of underwater images. This was followed by rotating and sampling each underwater image at 0 to produce 64 raw underwater photographs[25]. After that, photos taken underwater were subjected to low-resolution object recognition using blurred images. Image pre-processing results in a grayscale version of the original underwater photos (8 types). The neural network was used to improve low-resolution underwater image recognition, 90, 180, and 270, resultant in an mean significant rate of 95% or more. The findings are summarised in Tables 6 to 8. The ANN training curvature and documentation outcomes are depicted in Figures 5&6.

Conclusion

Unmanned underwater vehicles, such as remotely functioned submersibles or self-guided underwater gliders, must be able to quickly identify objects in images taken under water, and must have a high discrimination rate. Shape recognition in less-resolution underwater photographs can be enhanced by utilizing a amalgamation of neural network and humanoid vision characteristics. Low-resolution underwater photographs were able to be recognised at 95 % accuracy after shape recognition was improved by 50000 neural network iterations. An ANOVA shows that the neural network's recognition rate is much higher than human recognition.

The quality of the original photographs will be distorted and degraded as a result of many of the aforementioned causes, and valuable future information will be lost as a result. In addition to the eight basic categories listed above, there are many others. To address this issue, we created the empirical novel algorithm BPNN. People have tried to create a full theory after a series of successes with well-known networks such as Alex Net and Res Net that were founded on empirical perception. In the field of neural network research, any new technological or theoretical idea could lead to a breakthrough. We believe that publishing our findings on such a solid example would pique the interest of other scientists and lead them to consider applying our ground breaking concept across a wide range of disciplines. Of course, we'll compare our data to that of CNN in the future to see if we can come up with anything new.

Using CNN for image identification can be useful, but it suffers from an overfitting problem if the quantity of active data is not great sufficient. The CNN training procedure relies on a small number of data to gradually update the model. In CNN's example, this means that CNN will only recognise a few samples. The lack of training features means that CNN may be unable to recognise images that have yet to be presented to it. In order to make the CNN model useful, it requires a significant amount of For this subject, we only have a small number of samples to

work with because they are constrained by the underwater environment. To utilise CNN, the quantity of samples is insufficient. If enough data is gathered, the CNN method will be studied in depth in the future.

REFERENCES

[1] R. C. Gonzalez and R. E. Woods, "No Title," Digit. Image Process., 1992.

[2] A. McAuley, A. Coker, and K. Saruhan, "Effect of noise in moment invariant neural network aircraft classification," in IEEE Proceedings of the National Aerospace and Electronics Conference, 1991, vol. 2, pp. 743–749.

[3] 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, 2007, doi: 10.1088/1464-4258/9/2/003.

[4] Y.-C. Fang and B.-W. Wu, "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, 2008, doi: 10.1109/TPAMI.2007.70839.

[5] 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, 2013, doi: 10.3807/JOSK.2013.17.4.328.

[6] B.-W. Wu and L.-S. Chang, "New application of human eye model in the verification of spot diagram with refractive diopter," Optik (Stuttg)., vol. 125, no. 6, pp. 1784–1788, 2014, doi: 10.1016/j.ijleo.2013.09.053.

[7] B.-W. Wu, Y.-C. Fang, and D. P.-C. Lin, "Applications of neural networks in human shape visual perception," J. Opt. Soc. Am. A Opt. Image Sci. Vis., vol. 32, no. 12, pp. 2338–2345, 2015, doi: 10.1364/JOSAA.32.002338.

[8] G. Huang, Y. Sun, Z. Liu, D. Sedra, and K. Q. Weinberger, "Deep networks with stochastic depth," Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), vol. 9908 LNCS. pp. 646–661, 2016, doi: 10.1007/978-3-319-46493-0_39.

[9] A. R. Backes and J. J. de Mesquita Sá Junior, "LBP maps for improving fractal based texture classification," Neurocomputing, vol. 266, pp. 1–7, 2017, doi: 10.1016/j.neucom.2017.05.020.

[10] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017, 2017, vol. 2017-Janua, pp. 2261–2269, doi: 10.1109/CVPR.2017.243.

[11] A. G. Howard et al., "Mobilenets: Efficient convolutional neural networks for mobile vision applications," CoRR, 2017.

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

[13] 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, doi: 10.1109/ACCESS.2018.2825310.

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

[15] P. Deeprasertkul and W. Praikan, "An Application of Numbers and Characters Recognition and Classification on Radar Images Using for Flood Monitoring," in 2018 3rd International Conference on Computer and Communication Systems, ICCCS 2018, 2018, pp. 301–305, doi: 10.1109/CCOMS.2018.8463346.

[16] X. Xu, J. Zhou, and H. Zhang, "Screen-rendered text images recognition using a deep residual network based segmentation-free method," in Proceedings - International Conference on Pattern Recognition, 2018, vol. 2018-Augus, pp. 2741–2746, doi: 10.1109/ICPR.2018.8545678.

[17] 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 NICS 2018 - Proceedings of 2018 5th NAFOSTED Conference on Information and Computer Science, 2019, pp. 300–304, doi: 10.1109/NICS.2018.8606836.

[18] B. B. Traore, B. Kamsu-Foguem, and F. Tangara, "Deep convolution neural network for image recognition," Ecol. Inform., vol. 48, pp. 257–268, 2018, doi: 10.1016/j.ecoinf.2018.10.002.

[19] I. M. Dheir, A. S. A. Mettleq, A. A. Elsharif, and S. S. Abu-Naser, "Classifying nuts types using convolutional neural network," Int. J. Acad. Inf. Syst. Res., vol. 3, no. 12, pp. 12–18, 2020.

[20] M. W. Akram et al., "CNN based automatic detection of photovoltaic cell defects in electroluminescence images," Energy, vol. 189, 2019, doi: 10.1016/j.energy.2019.116319.

[21] 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, 2019, doi: 10.1016/j.jvcir.2019.03.010.

[22] P. P. Banik, R. Saha, and K.-D. Kim, "An Automatic Nucleus Segmentation and CNN Model based Classification Method of White Blood Cell," Expert Syst. Appl., vol. 149, 2020, doi: 10.1016/j.eswa.2020.113211.

[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, 1980, doi: 10.1109/TPAMI.1980.4766994.

[24] L. Chen et al., "Underwater object detection using Invert Multi-Class Adaboost with deep learning," 2020, doi: 10.1109/IJCNN48605.2020.9207506.

[25] L. Chen et al., "Perceptual Underwater Image Enhancement with Deep Learning and Physical Priors," IEEE Trans. Circuits Syst. Video Technol., vol. 31, no. 8, pp. 3078–3092, 2021, doi: 10.1109/TCSVT.2020.3035108.