Dorothy Rajendran[1], Thankappan Sasilatha[1], Suvakeen Amala Doss Hebciba Mary[2], Susai Santhammal Rajendran[2,3], Caslav Lacnjevac[4], Gurmeet Singh[5]*

[1] AMET University, Department of EEE, Chennai, India, [2]Department of Chemistry, St. Antony's College of Arts and Sciences for Women Thamaraipady, Tamil Nadu, India,[3] Pondicherry Uniersity, Puducherry, India [4]University of Belgrade, Faculty of Agriculture, Belgrade, Serbia, [5]Pondicherry University, Puducherry, India.

Scientific paper
ISSN 0351-9465, E-ISSN 2466-2585
https://doi.org/10.5937/zasmat2201005R

Zastita Materijala 63 (1)
5 - 14 (2022)

Deep learning based underwater metal object detection using input image data and corrosion protection of mild steel used in underwater study - A case study

Part A - Deep learning based underwater metal object detection using input image data

ABSTRACT

Due to the importance of underwater exploration in the development and utilization of deep-sea resources, underwater autonomous operation is more and more important to avoid the dangerous high-pressure deep-sea environment. For underwater autonomous operation, the intelligent computer vision is the most important technology. In an underwater environment, weak illumination and low-quality image enhancement, as a pre-processing procedure, is necessary for underwater vision. In this paper, introduced the Deep learning based Underwater Metal object detection using input Image data by using several step to improve the model performance. In this experimentation we are using TURBID dataset 100 images to validate the performance. And also we compare the performance result by given the input images in different validation level. In first input image is initially preprocessed and that images is given to the KFCM-Segmentation. The segmented images are given to the DWT Extraction to extract the features from those images. And finally the Convolution Neural Network (CNN) is used to classify the images to detect the objects. Also this proposed model attained the classification accuracy of 98.83%. This method is much suitable for detect the objects in underwater robotically. Metallic parts of machines of ships or aero planes may submerge in sea water. They may undergo corrosion when they come in contact with sea water which contains 3.5% sodium chloride. This is most commonly responsible for the corrosive nature of the sea water. The robots made of materials such as mild steel may also undergo corrosion when they come in contact with sea water, while is search. If a paint coating is given, it will control the corrosion of these proposed materials. Hence this work is undertaken. Mild steel is coated with Asian guard red paint. Corrosion resistance of mild in3.5% sodium chloride solution is measured before coating and after coating by electrochemical studies such as such as polarization study and AC impedance spectra . The corrosion inhibition efficiency offered by red paint to mild steel in 3.5% sodium chloride is 99.98%.

Keywords: *input Image data**,** Convolutional Neural Network (CNN), Fuzzy c-means clustering and TURBID dataset polarization study, AC impedance spectra, sea water.*

1. INTRODUCTION

This article is divided into two parts. Part A deals with "Deep learning based Underwater Metal object detection using input Image data". Part B deals with "Corrosion protection of mild steel used in underwater study- A case study"

The underwater environment is one of the most challenging conditions for object detection. The signal received by any sensor can be significantly absorbed and distorted by the water medium [1]. This significantly degrades the performance of object detection methods, leading to high false positive and false negative ratios. Moreover, in underwater environments, it is quite difficult to deploy and control sensors [2]. Many state-of-the-art devices and technologies are not suited to underwater environment operation. In general, sonar and cameras are two typical sensors widely used for underwater object detection [3-6]. Sonar sensors are sensitive to geometrical structure

*Corresponding author: Dorothy Rajendran
E-mail: rdorothyraj.dgl@gmail.com
Paper received: 30. 07. 2021.
Paper corrected: 20. 08. 2021.
Paper accepted: 25.08. 2021.
Paper is available on the website: www.idk.org.rs/journal

information and can provide information of underwater scenes even in low- and zero-visibility environments. However, the data acquired by sonar can only present the difference of the distance over the scanning points. Other factors such as visual features are missed by this type of sensor. As a result, sonar-based systems are feasible for top-down tasks, such as hydrographic surveying and charting [7], shipwreck searching [8], and marine geological surveys [9]. In contrast to sonar, cameras can provide more types of visual information at high spatial and temporal resolutions. Prominent objects can be identified by the various visual features such as color, intensity [10], texture, and contours [11]. Recently developed binocular or multi-ocular underwater systems can generate non-scale depth maps [12-14]. Hence, in addition to these top-down tasks, underwater vision systems possess a better ability to handle down-top tasks where we have few prior knowledge of the current underwater scenes, such as marine ecology monitoring [15] and underwater entertainment [16]. However, vision-based underwater object detection methods have not yet provided satisfactory results, although, in contrast, sonar has more opportunities to succeed in underwater object detection.

The drawback of underwater vision systems lies in their instability for underwater object detection. Underwater images acquired by cameras suffer from intensity degeneration, color distortion, and haze effects [17]. In order to make the underwater images clear and distinguishable, several underwater image enhancement or restoration methods have been introduced into object detection models as a pre-processor before feature extraction [18]. However, if these image pre-processors do not adapt to the underwater optical environment, many new noise sources and false colors will be mistaken for the objects themselves. The increasing demand for vision-based applications enhances the importance of camera-based object detection methods in underwater scenes.

The monocular camera system may be a better option for underwater environments provided it is sufficiently robust to the underwater conditions. In order to reach this goal and improve the performance of the monocular vision system for underwater object detection, light transmission information is introduced as a novel cue to identify underwater objects in the region of interest (ROI). This transmission information is combined with the color and intensity features to detect the ROI, which is then filtered and segmented to produce the results of underwater object detection.

2. RELATED WORKS

Zhu et al. [19] proposed an underwater object detection method based on the discriminative regional feature integration method. In this method, three features, including regional contrast, regional property, and regional background descriptors, are jointly used to establish a comprehensive saliency map for underwater images. Li et al. [20] proposed a region contrast-based method by using the image segmentation method as the pre-processor. The region segmentation method may benefit from the removal of noisy data points but will cause false detections in the high-intensity regions. As a result, the region segmentation-based method can detect all salient regions with high-intensity; however, most regions are not consistent with the ground-truth. Chang et al. [21] introduced a significant amount of polarization into light at scattering angles near 90 degrees: This light can then be distinguished from light scattered by an object that remains almost completely unpolarised. Results were obtained from a Monte Carlo simulation and from a small-scale experiment, in which an object was immersed in a cell filled with polystyrene latex spheres suspended in water.

Krizhevsky et al. [22] applied CNN method to deal with classification problem winning the champion of ILSVRC (ImageNet Large Scale Visual Recognition Challenge), which reduce the top 5 error rate to 15.3%, from then on deep CNN has been widely applied. Perez et al. [23] proposed an underwater image enhancement method based on deep learning, which constructed a training data set consisting of groups of degraded underwater images and restored underwater images. The model between degraded underwater images and restored underwater images was obtained from a large number of training sets by deep learning method, which is used to enhance the underwater image quality. Yang et al. [24] proposed a method of detecting underwater weak target based on Gabor transform, which is processed on laser underwater complicated non-stationary signal to turn it to become an approximate stationary signal, and then the triple correlation is computed with Gabor transform coefficient and it can eliminate random interference and extrude target signal's correlation.

3. PROPOSED SYSTEM

Figure 1 shows the Proposed System Block diagram.

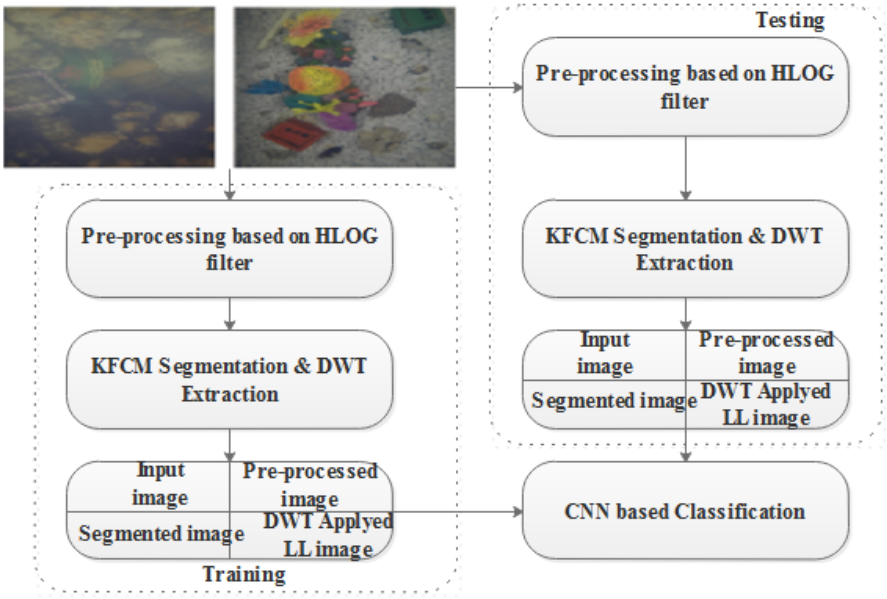

Figure 1. Proposed system block diagram

Slika 1. Predloženi sistemski blok dijagram

Pre-processing

The Initial step of the pre-Processing is Rgb2gray colour conversion. With the help of mat lab inbuilt command we have did this process. The second step of the pre-processing is block conversion for this process input image is converted into 256*256 matrixes. From that matrix the image is converted in to 64*64 sub blocks over all the 256*256 matrix size of image consist of 16 blocks. The block converted images are given to the hybrid- LOG filter. HLOG is a combination of LOG and Gaussian filter. Generally, the Laplacian filters are derivative filters and generally used in finding the areas of edges of images. As this derivative filter is very sensitive to noise, the research work used Gaussian filters for smoothening the image. In LOG filter, the Gaussian filtering is performed before the Laplacian filtering. After these two processes, again the Gaussian filtering is applied to smoothen the images. The output from the LOG filter is given as an input to the Gaussian filter. The filtering of LOG is shown in Eq. (9), and the Gaussian filtering of LOG is shown in Eq. (1).

$$\nabla^2 G(x,y) = \frac{x^2+y^2-2\sigma^2}{\pi\sigma^4} exp\left(\frac{-x^2+y^2}{2\sigma^2}\right) \quad (1)$$

$$G(x,y;\sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} exp\left(\frac{-x^2+y^2}{2\sigma^2}\right) \quad (2)$$

where the Laplacian operator is ∇^2, standard deviation is σ, and x and y are coordinates of the images. The standard deviation plays a major role in the filtering behavior of the HLOG filter. The regions of rapid intensity change are highlighted by Laplacian. The LOG filter is an orientation-independent operator and the scale is provided by σ. The corners, curves, and locations have varied intensity function, which is in non-linear manner breakdown by using the LOG filter. A set of images with various levels of smoothness is achieved by applying Gaussian filter of different scales of standard deviation to an image. In Gaussian filter, zero crossings of the second-order derivatives are found to detect the edges from the images. The purpose of the hybrid LOG filter in the proposed method is to remove the noise from the restored images because the images from the database may be corrupted by noises from the environment. Thus, here, the hybrid LOG filter is used to enhance the contrast and brightness of the restored image. The hybrid LOG filter provides higher significance to the pixels near the edge. Effective smoothening and noise removal is obtained by using the combination of LOG and Gaussian filtering in the proposed method [25].

KFCM-Segmentation

One of the important aspects of our Total Generalized Variation Fuzzy c-means clustering (TGVFCMS) method is robustness to reduce noise and its ability to edge-preserving. Particularly, the desirable properties of TGV regularization up to a certain order of differentiation make it of a useful tool to measure image characteristics, such as noise removing and sharp edge preserving. To eliminate the undesired noise and artifacts from the FCM-based methods, in this study the TGV regularization was adapted in the smoothing term of our TGVFCMS. Let us first define the framework of TGV as follows:

$$TGV_a^k(u) = sup\left\{\int_\Omega udiv^k vdx |\ v \epsilon \square_c^k(\Omega, Sym^k(\square^d)), ||div^l v||_\square \leq a_l\right\} \quad (3)$$

where $l = 0,1,\dots,k-1$, and $k \in \mathbb{N}$ indicates an order of TGV, and $a = (a_0, a_1, \dots, a_{k-1})$ denotes the positive weight to TGV. $Sym^k(\mathbb{R}^d)$ represents the space of symmetric k- tensors. For each component $\eta \epsilon M_{k-1}$, the l-divergence of the symmetric k tensor field is given by

$$(div^l v)_\eta = \sum_{r\epsilon M_1} \frac{l!}{\gamma!} \frac{\partial^l v_{\eta+\gamma}}{\partial x^\gamma} \quad (4)$$

where $M_k$is the multi-index of order k

$$M_k = \{\eta \epsilon \mathbb{N}^d | \sum_{i=1}^d \eta_i = k\} \quad (5)$$

The ∞- norm for symmetric k-vector field is given as

$$||v||_\infty = \underset{x\epsilon\Omega}{sup}\left\{\left(\sum_{\eta\epsilon M_k} \frac{k!}{\eta!} v_\eta(x)^2\right)\right\} \quad (6)$$

The first-order gradient and the high-order gradient achieved in (3) are both constrained to be sparse, which can obviously reduce the staircase artifacts.

Here, we take into account the second order TGV, i.e.

$$TGV_a^2(u) = sup\left\{\int_\Omega udiv^2 vdx\ |v \epsilon \mathbb{C}_c^2(\Omega, S^{d\times d}), \left||v|\right|_\infty \leq a_0, ||div\ v||_\infty \leq a_1\right\} \quad (7)$$

where $\mathbb{C}_c^2(\Omega, S^{d\times d})$ denotes the vector space of compactly supported under the set of symmetric matrices $S^{d\times d}$. Particularly, the respective definitions for the divergence and norms can be calculated as follows:

$$(div\ v)_i = \sum_{j=1}^d \frac{\partial v_{ij}}{\partial x_j}, (div^2\ v)_i = \sum_{j=1}^d \frac{\partial^2 v_{ii}}{\partial x_i^2} + 2\sum_{i<j} \frac{\partial v_{ij}}{\partial x_j \partial x_j} \quad (8)$$

and

$$||v||_\infty = \underset{x\ \in\ \Omega}{sup}(\sum_{i=1}^d |v_{ii}(x)|^2 + 2\sum_{i<j} |v_{ij}(x)|^2)^{1/2}$$

$$||div\ v||_\infty = \underset{x\ \in\ \Omega}{sup}\left\{\sum_{i=1}^d |\sum_j^d \frac{\partial^2 v_{ij}}{\partial x_j^2}(x)|^2\right\}^{1/2} \quad (9)$$

where minimum solution is taken over all vector fields on Ω and $\varepsilon(v) = (\nabla v + \nabla u\, T)/2$ indicates the symmetrized derivative. Here, the definition of (9) shows that $\nabla^2 u$ *contributes less than* $\nabla u = v$ *to smooth regions. In the edg*e neighbors, $\nabla^2 u$ is locally 'larger' than ∇u in these regions, and minimization could work well with $v = 0$. Thus, provides a way to achieve the balance between the first and second derivative (via the ratio of positive weights α_0 andα_1. For practical purposes, the two weights α_0 and α_1 are tuned to 0.1 and 0.15, respectively. Through the definition of second-order TGV, the proposed TGVFCMS can yield results that are more robust to noise and detail-preserving [26].

DWT Extraction

To compute the wavelet coefficient of the input images, the DWT is used and it consider a square shaped function. The DWT technique has better compression energy and proper reconstruction with short support filters, low-computation, and no redundancy. The DWT follows the fuzzy de-noising procedure, which provides shift capable sub-bands and better directional selectivity with less redundancy. In the multi-resolution process, the real texture of the image is computed by zooming in and zooming out process. Usually, the images are decomposed into a number of sub-image at various resolution serving the low and high frequency information. The DWT property helps to extract the information of the texture from the images. The square integral function $f(u)$ and wavelet transform w is represented as the inner product$f$, and $\psi(u)$ is original valued function. The wavelet function is given in the Eq. (10)

$$w[f(s,\tau)] = (f, \psi_{s,t}^k) = \int_\infty^\infty f(u)\psi_{s,t}^k(u)du \quad (10)$$

$$Where, \psi_{s,t}^k(u) = \left(\frac{1}{\sqrt{s\psi_{s,t}^k}}/s\right) denotes \quad (11)$$

Wave family, $s\ \epsilon\ z$ is scale, τ is translation and $k \in \{h, v, d\}$ is orientation parameters. The orientation parameters h, v and d denote vertical, horizontal and diagonal direction respectively. The dyadic wavelet decomposition achieved during $s = 2^j$ and $\tau = 2^j, n, j, n\varepsilon z$. the dyadic wavelet decomposition is a scalable sample of DWT; it follows a geometric sequence of ratio 2. The following wavelet decomposition employs dyadic wavelets that are implemented using perfect reconstruction filter banks. By using wavelet function $\psi(u)$ and the scaling function $\varphi(u)$, which showed in Eq. (10) and (11). The wavelet atoms describe by scaling and three mother atomsψ^h, $\psi^v$and ψ^d. This mother atoms computed as the tensor products of 1-dimensional $\psi(u)$ and $\varphi(u)$,, that is denoted in Eq. (13) and (14)

$$\psi_{j,n}^{k}(u) = \frac{1}{\sqrt{2^j}}\psi^{k}\left(\frac{u-2^j.n}{2^j}\right) \quad (12)$$

$$\varphi_{j,n}^{k}(u) = \frac{1}{\sqrt{2^j}}\psi^{k}\left(\frac{u-2^j.n}{2^j}\right) \quad (13)$$

$$\varphi(u) = \varphi(u_1)\varphi(u_2), \psi^{k}(u) \quad (14)$$

$$\varphi^{v}(u) = \varphi(u_1)\psi(u_2), \psi^{d}(u) \quad (15)$$

Two dimensional DWT is implemented by combination of down samplers and digital filter banks. The digital filter bank made up of a low-pass filter and high-pass filter. The number of the bank is grouped as per the desired functions in the wavelet configuration structure. Next, the rows and columns of the underwater image are separately undergone through the 1-dimensional wavelet transform in order to create 2-dimensional wavelet coefficient. The original images $A_{2^{j+1}}f$ at resolution 2^{j+1} decompose into 4-subband images in the frequency domain. Three sub-band images such as $D_{2^i}^{h}f, D_{2^i}^{h}f$ and $D_{2^i}^{h}f$ are the original images in resolution in vertical, horizontal and diagonal between the 4-subband images. The fourth image is an approximation image, $A_{2^i}f$ found at coarse resolution, so the entire under water image $A_{2^{j+1}}f$ is denoted in the Eq. (16).

$$A_{2^{j+1}}f = D_{2^i}^{h}f, +D_{2^i}^{v}f + D_{2^i}^{d}f + A_{2^i}f \quad (16)$$

The decomposed sub-images are the 2-dimensional orthogonal wavelet. The results of the wavelet decomposition of an image is 4-orthogonal sub-bands such as Low-Low (LL) band, Low-High (LH) band, High-Low (HL) band and High-High (HH) band, which is represented as $D_{2^i}^{h}f, +D_{2^i}^{v}f + D_{2^i}^{d}f and A_{2^i}f$ respectively.

CNN classification

Convolutional Neural Network (CNN) is a multiple layer feed forward neural network, which includes different types of layers, which are convolution layer, ReLU layer, pooling layers, and fully connected output layers. CNN is designed to recognize features in images such as edges and shapes.

Convolutional Layer

The layer which comes first in CNN architecture is always a Convolutional Layer. Typically an input layer to a CNN accepts MxNx1. Here MxN is the two-dimensional size of image with Single layers. CNN uses a filter with particular parameters which are having the same depth as the input image, and the filter is convolved with the image. The filter represents a curve orshape to which the input image is convolved with. The shape that resembles the curve in the input image that is represented by the filter ends up in higher values as a result of convolution. Convolution operation can be represented by Eq.(17).

$$s(t) = (x^{*}w)(t) \quad (17)$$

Pooling layer

The pooling layer is used to reduce the data size. Pooling involves arranging the matrix data in different segments and replacing the whole segment with a single value hence reducing the metrics data dimension. Some of the popular pooling functions are Maxpooling and Average-pooling, in which the segmented matrices are replaced by the maximum or average of all the values in the current segment as shown in Fig. 2.

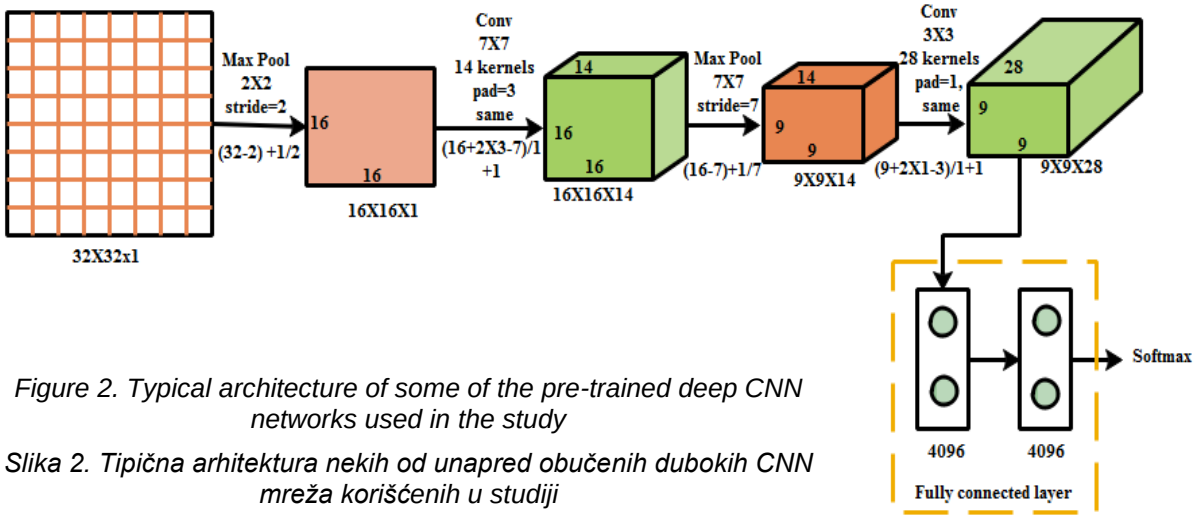

Figure 2. Typical architecture of some of the pre-trained deep CNN networks used in the study

Slika 2. Tipična arhitektura nekih od unapred obučenih dubokih CNN mreža korišćenih u studiji

Fully Connected layer

To fit the network layer architecture, dimensionality of layers is altered in a fully connected layer. A fully connected layer is a function operation that is between m and n each dimension of input and output are connected to each other. A

fully connected layer connects all the activations from the previous layer to the next layer of the network, just as it is usually in a conventional artificial neural network.

Softmax layer

The Softmax function translates input from earlier layers into a probability for the classes that sum to one. Thus this layer plays a critical role in output as the predicted output is the class which has maximum probability for the given input data. Several deep neural networks that have been used to classify imagesare available. There are two types of deep networks used in this study firstly, linear networks which as the name suggests have simple linear architecture although these networks are pre-trained to classify other images we can modify them to suit our classification problem using transfer learning by tweaking required parameters.

For all the networks have been kept training hyperparameters for constant. We have divided data in multiple epochs witch are allowed to go maximum 25. Mini-batch size is the number that represents the number of samples after which the internal parameters of the model are updated. Mini-batch size for training in our experiment was kept 7, and the Initial learning rate for each training was kept 0.0001.

4. RESULT AND DISCUSSION

The proposed system is experimented using MATLAB (version 2018a) with 3.0 GHz Intel i3 processor, 1TB hard disc and 8 GB RAM. For determining the effectiveness of the proposed system is compared with the existing systems on the publically available TURBID dataset. In the proposed system, in TURBID dataset 100 images.

TURBID Dataset description

The TURBID dataset was an initial dataset proposal for the algorithms evaluation procedure that will be present in the next section. The dataset consist of three different high quality printed real scenes previously photographed at the Bahamas. These images was called here as Photo1, Photo2 and Photo3. These scenes contains structures of the underwater floor and some human made objects. The pictures were re-photographed inside a 1000 litres tank made of plastic, illuminated by two 30 watts fluorescent light strips. As the image capture device we used a static Go Pro Hero3 Black Edition with 12 mega pixels (3000x4000) of resolution. Some of the sample images are defined in the figure 3.

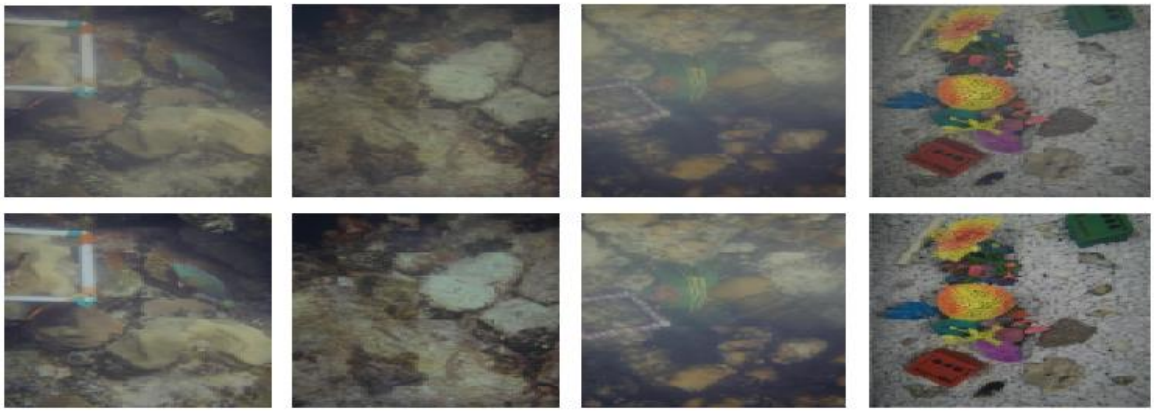

Figure 3. Sample Dataset Images
Slika 3. Uzorci slika skupa podataka

Evaluation Metrics

The challenge evaluation metrics is used for evaluating the both segmentation and classification performance of our method. For the segmentation, the evaluation criteria include sensitivity (SE), specificity (SP), accuracy (AC), Recall (R) and Precision (P). The performance criteria are defined is as:

$$SE = \frac{tp}{tp + fn}$$

$$AC = \frac{tp + tn}{tp + fp + tn + fn}$$

$$SP = \frac{tn}{tn + fp}$$

where tp, tn, fp and fn denote the number of a true positive, true negative, false positive and false negative. As for the classification, there are four evaluation criteria, including sensitivity (SE), specificity (SP) and accuracy (AC).

Performances analysis

The proposed system performance has validated in various is discussed in the below section.

Table 1. Comparision of performance analysis by using various input image data

Tabela 1. Poređenje analize performansi korišćenjem različitih ulaznih slikovnih podataka

S.no	Method	SE (%)	SP (%)
1	Only Input image	49.58	52.71
2	Only Preprocessed image	55.83	66.43
3	Only segmented image	60.00	62.00
4	Only DWT image	64.00	60.00
5	Input and pre-processed images	78.36	87.44
6	Segmented and DWT images	87.692	94.658
7	Input and Segmented images	89.04	90.35
8	Segmented and DWT images	97.58	96.89
9	pre-processed and Segmented images	89.231	95.54
10	pre-processed and DWT images	97.65	97.56
11	Input, pre-processed and Segmented images	90.6156	89.3844
12	Input, pre-processed and DWT images	95.2976	94.7024
13	Input, DWT and Segmented images	98.61	98.58
14	pre-processed, Segmented and DWT images	100	99.73
15	Input, pre-processed, Segmented and DWT images	100	100

Table 2. Comparision of classification accuracy by using various input image data

Tabela 2. Poređenje tačnosti klasifikacije korišćenjem različitih ulaznih slikovnih podataka

S.no	Method	AC (%)
1	Only Input image	55.87
2	Only Preprocessed image	59.64
3	Only segmented image	61.11
4	Only DWT image	60.64
5	Input and pre-processed images	65.38
6	Segmented and DWT images	68.23
7	Input and Segmented images	66.47
8	Segmented and DWT images	70.54
9	pre-processed and Segmented images	85.44
10	pre-processed and DWT images	80.30
11	Input, pre-processed and Segmented images	88.02
12	Input, pre-processed and DWT images	94.89
13	Input, DWT and Segmented images	96.77
14	pre-processed, Segmented and DWT images	96.38
15	Input, pre-processed, Segmented and DWT images	98.83

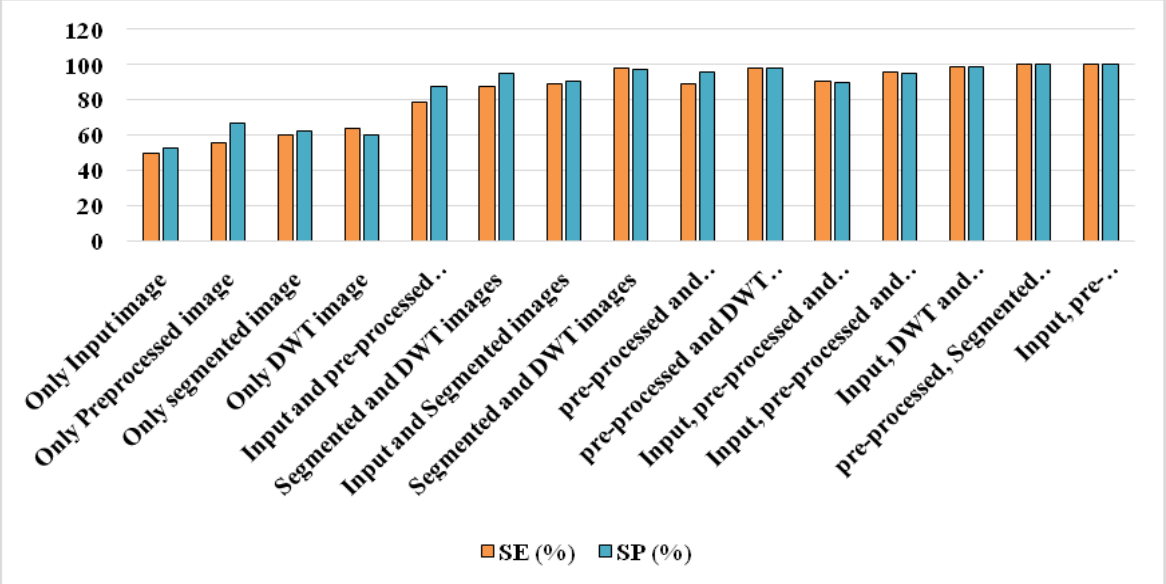

Figure 4. Graphical representation of different input image data performance

Slika 4. Grafički prikaz različitih performansi ulaznih slikovnih podataka

In above table 1 and figure 4 signifies the complete performance analysis of different kinds of input data is given to achieve the better performance analysis in different stages. In this validation, first we preferred to analysis the classification performance by giving only image data to achieve the SE value of 49.58% and SP of 52.71%. And next we gives the only preprocessed image by achieved the SE value of 55.83%, it is slightly better than only input image. Then gives the only segmented images to classifier to achieve 60% of SE and 62% of SP value. Then gives only DWT images to achieve the SE value of 64%. And then gives the input and preprocessed image. 78.36% of SE. after that Segmented and DWT images is given to classifier to achieve the SE value of 97.58%, it is better than the previous inputs results. Then Input, pre-processed and segmented images is given to the classifier to achieve the SE value of 90.61% of SE value and 89.38% of SP value. Then pre-processed, Segmented and DWT images is given to the classifier to achieve the 100% of SE value and 99.73% of SP value. Then finally Input, pre-processed, Segmented and DWT images is given to the classifier to attain the SE value of 100% and SP value of 100%. In this input image comparisons, finally the Input, pre-processed, Segmented and DWT images is attained better results than other input images.

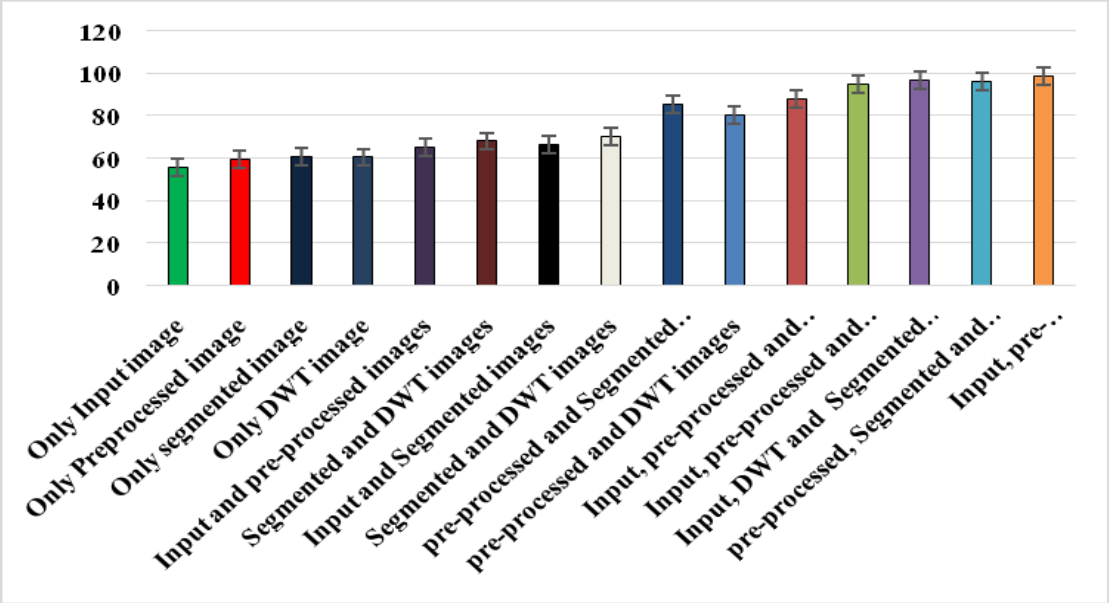

Figure 5. Graphical representation of classification accuracy

Slika 5. Grafički prikaz tačnosti klasifikacije

In above table 2 and figure 5 signifies the classification accuracy performance of different kinds of input data in different stages. In this validation, first we desired to analysis the classification accuracy performance by giving only image data to achieve the accuracy of 55.87%. And next we gives the only preprocessed image by achieved the accuracy value of 59.64%, it is slightly better than only input image. Then gives the only segmented images to classifier to achieve 61.11% of accuracy value. Then gives only DWT images to achieve the accuracy value of 60.64%. And then gives the input and preprocessed image. 65.68% of accuracy. After that Segmented and DWT images is given to classifier to achieve the accuracy of 68.23%, it is better than the previous inputs results. Then Input, pre-processed and segmented images is given to the classifier to achieve the accuracy value of 88.02%. Then pre-processed, Segmented and DWT images is given to the classifier to achieve the 96.38% of accuracy value. Then finally Input, pre-processed, Segmented and DWT images is given to the classifier to attain the 98.83% of accuracy. In this input image comparisons, finally the Input, pre-processed, Segmented and DWT images is attained better classification results than other input images.

5. CONCLUSION

Considering the underwater vision characteristics, some new image processing procedures are proposed to deal with the low contrast and the weakly illuminated problems. A deep CNN method is proposed to achieve the detection and classification of marine organisms, which is commonly recognized as the fastest object detection method.

The effectiveness and capability of the proposed method are obviously verified by the qualitative and quantitative evaluation results. The proposed method is suitable for our underwater

robot to detect the objects, which is not better than the typical methods for the other dataset. And dropout layers and other technologies are not significant in this model; the reconstruction of the network by using a more complicated algorithm would be more effective.

6. REFERENCES

[1] X.Xue, D.Pan, X.Zhang, B.Luo, J.Chen, H.Guo (2015) Faraday anomalous dispersion optical filter at 133 Cs weak 459 nm transition, Photonics Res., 3, 275–278.

[2] W.Liu, Z.Xu, L.Yang (2015) SIMO detection schemes for underwater optical wireless communication under turbulence. Photonics Res., 3, 48–53.

[3] C.Spampinato, Y.Chen-Burger, G.Nadarajan, R.Fisher (2008) Detecting, Tracking and Counting Fish in Low Quality Unconstrained Underwater Videos, VISAPP p.514–519.

[4] G.Foresti, S.Gentili (2000) A vision based system for object detection in underwater images, Int. J. Pattern Recognit. Artif. Intell., 14, 167–188.

[5] H.Cho, J.Gu, H.Joe, A.Asada, S.Yu (2015) Acoustic beam profile-based rapid underwater object detection for an imaging sonar, J. Mar. Sci. Technol., 20, 180–197.

[6] I.Masmitja, S.Gomariz, J.Del Rio, B.Kieft, T.O'Reilly (2016) Range-only underwater target localization: Path characterization, In Proceedings of the OCEANS 2016 MTS/IEEE Monterey, Monterey, CA, USA, p. 1–7.

[7] C.Gostnell, J.Yoos (2005) Efficacy of an interferometric sonar for hydrographic surveying: Do interferometers warrant an in-depth examination, Hydrogr. J., 118, 17–22.

[8] R.Ballard, L.Stager, D.Master, D.Yoerger, D.Mindell, L.Whitcomb, H.Singh, D.Piechota (2002) Iron age shipwrecks in deep water off Ashkelon, Israel. Am. J. Archaeol., 106, 151–168.

[9] D.Piper, P.Cochonat, M.Morrison (1999) The sequence of events around the epicentre of the 1929 Grand Banks earthquake: Initiation of debris flows and turbidity current inferred from sidescan sonar, Sedimentology, 46, 79–97.

[10] A.Caffaz, A.Caiti, G.Casalino, A.Turetta (2010) The hybrid glider/AUV Folaga, IEEE Robot. Autom. Mag., 17, 31–44.

[11] A.Ortiz, M.Simó, G.Oliver (2002) A vision system for an underwater cable tracker. Mach. Vis. Appl., 13, 129–140.

[12] F.Bruno, G.Bianco, M.Muzzupappa, S.Barone, A.Razionale (2011) Experimentation of structured light and stereo vision for underwater 3D reconstruction, ISPRS J. Photogram. Remote Sens., 66, 508–518.

[13] M.Johnson-Roberson, O.Pizarro, S.Williams, I.Mahon (2010) Generation and visualization of large-scale three-dimensional reconstructions from underwater robotic surveys, J. Field Robot., 27, 21–51.

[14] S.Negahdaripour, H.Madjidi (2003) Stereovision imaging on submersible platforms for 3-D mapping of benthic habitats and sea-floor structures, IEEE J. Ocean. Eng., 28, 625–650.

[15] B.Brown, R.Dunne, M.Goodson, A.Douglas (2000) Marine ecology: Bleaching patterns in reef corals, Nature, 404, 142–146.

[16] I.Holjevac (2003) A vision of tourism and the hotel industry in the 21st century, Int. J. Hosp. Manag., 22, 129–134.

[17] R.Smith, K.Baker (1981) Optical properties of the clearest natural waters (200–800 nm). Appl. Opt., 20, 177–184.

[18] R.Schettini, S.Corchs (2010) Underwater image processing: State of the art of restoration and image enhancement methods. EURASIP J. Adv. Signal Process, 10, 746–752.

[19] Y.Zhu, L.Chang, J.Dai, H.Zheng, B.Zheng (2016) Automatic object detection and segmentation from underwater images via saliency-based region merging, In Proceedings of the OCEANS, Shanghai, China, 10–13 April 2016.

[20] L.Li, R.M.Eustice, M.Johnson-Roberson (2015) High-level visual features for underwater place recognition, In Proceedings of the IEEE International Conference on Robotics and Automation (ICRA), Seattle, WA, USA.

[21] P.C.Chang, J.Flitton, K.Hopcraft, E.Jakeman, D.Jordan, J.Walker (2003) Improving visibility depth in passive underwater imaging by use of polarization, Applied Optics, 42(15), 2794–2803.

[22] A.Krizhevsky, I.Sutskever, G.E.Hinton (2017) ImageNet classification with deep convolutional neural networks, Communications of the ACM, 60(6), 84–90.

[23] J.Perez, A.C.Attanasio, N.Nechyporenko, P.J.Sanz (2017) A deep learning approach for underwater image enhancement, In International Work-Conference on the Interplay Between Natural and Artificial Computation, Springer, Cham, p.183–192.

[24] S.Yang, F.Peng (2009) Laser underwater target detection based on Gabor transform, 4th International Conference on Computer Science & Education, Nanning, China, p.95–97

[25] D.Basha, D.Khalandar, T.Venkateswarlu (2019) Linear Regression Supporting Vector Machine and Hybrid LOG Filter-Based Image Restoration, Journal of Intelligent Systems, 29(1), 1480-1495.

[26] X.Zhang, P.Weijun, Zh.Wu, J.Chen, Yi.Mao, R.Wu (2020) Robust Image Segmentation Using Fuzzy C-Means Clustering With Spatial Information Based on Total Generalized Variation, IEEE Access, 8, 95681-95697.

IZVOD

DETEKTOVANJE PODVODNIH METALNIH OBJEKATA POMOĆU VEŠTAČKE INTELIGENCIJE I ZAŠTITA OD KOROZIJE PREDMETA OD MEKOG ČELIKA KORIŠĆENIH U PODVODNOJ STUDIJI - STUDIJA SLUČAJA

Deo A - detektovanje podvodnih metalnih objekata pomoću veštačke inteligencije

Zbog važnosti podvodnog istraživanja u razvoju i korištenju dubokomorskih resursa, podvodni autonomni rad je sve važniji kako bi se izbeglo opasno dubokomorsko okruženje pod visokim pritiskom. Za podvodni autonomni rad, inteligentni računarski vid je najvažnija tehnologija. U podvodnom okruženju, slabo osvetljenje i nekvalitetno poboljšanje slike, kao postupak prethodne obrade, neophodni su za podvodni vid. U ovom radu predstavljeno je otkrivanje podvodnih metalnih objekata zasnovanog na veštačkoj inteligenciji pomoću ulaznih podataka o slici koristeći nekoliko koraka za poboljšanje performansi modela. U ovom eksperimentu koristi se TURBID skup podataka od 100 slika za proveru performansi. Takođe, upoređuje se rezultat performansi prema datim ulaznim slikama na različitim nivoima validacije. U prvom slučaju, ulazna slika se prethodno obrađuje i te slike se daju u KFCM-segmentaciji. Segmentirane slike se daju DVT ekstrakciji da izdvoje karakteristike iz tih slika. I na kraju, Convolution Neural Netvork (CNN) se koristi za klasifikaciju slika radi otkrivanja objekata. Takođe, ovaj predloženi model dostigao je tačnost klasifikacije od 98,83%. Ova metoda je veoma pogodna za robotsko otkrivanje objekata u morskim dubinama. Metalni delovi mašina brodova ili aviona mogu potonuti u morsku vodu. Mogu doći do korozije u kontaktu sa morskom vodom koja sadrži 3,5% natrijum hlorida. Ovo je najčešće odgovorno za korozivnu prirodu morske vode. Roboti napravljeni od materijala kao što je meki čelik, takođe, mogu pretrpeti koroziju kada dođu u kontakt sa morskom vodom, dok je u toku pretraga. Ako se nanese premaz boje, on će kontrolisati koroziju ovih predloženih materijala. Zbog toga se ovaj posao preduzima. Meki čelik premazan je azijskom zaštitnom crvenom bojom. Otpornost na koroziju blagog 3,5% rastvora natrijum hlorida meri se pre nanošenja i nakon nanošenja elektrohemijskim studijama, kao što su polarizacione studije i spektri impedanse naizmenične struje. Efikasnost sprečavanja korozije koju crvena boja nudi mekom čeliku u 3,5% natrijum hloridu je 99,98%.

Ključne reči: *ulazni podaci o slici, Konvoluciona neuronska mreža (CNN), klasterisanje sa srednjim vrednostima, Studija polarizacije TURBID skupa podataka, Spektar impedanse naizmenične struje, morska voda.*

Naučni rad
Rad primljen: 30. 07. 2021.
Rad korigovan: 20. 08. 2021.
Rad prihvaćen: 25. 08. 2021.
Rad je dostupan na sajtu: www.idk.org.rs/casopis