Segmentation of *Enhalus acoroides* seagrass from underwater images using the Mask R-CNN method

S A Pamungkas, I Jaya* and M Iqbal

Department of Marine Science and Technology, Faculty of Fisheries and Marine Sciences, IPB University, Dramaga, Bogor 16680, West Java, Indonesia

*E-mail: indrajaya@apps.ipb.ac.id

Abstract. Seagrass is a Spermatophyta plant that has many roles, including as a primary producer in the food chain in the waters. Monitoring of seagrass meadows and conditions needs to be done in order to achieve a healthy marine ecosystem. The steps taken in monitoring seagrass are by detecting and segmenting it. The purpose of the study is to implement and get information about the performance of the Mask R-CNN algorithm in detecting and segmenting the *Enhalus acoroides*. The dataset consists of 500 *Enhalus acoroides* images that had gone through a color correction and labelling process. The training process was performed with the configuration of 0.001 learning rate, batch size of 4 and some image augmentation was used to avoid overfitting. The optimum weight value was obtained after conducting the learning process with 100 epochs. A confusion matrix was used to evaluate detection performance, and linear regression was used to evaluate the segmentation produced by the model. The model evaluation results showed an accuracy value of 0.9246, a precision value of 0.9507, a recall value of 0.9712 and a correlation coefficient value of 0.8771. The value indicates that the model can detect and segment the seagrass *Enhalus acoroides* well and accurately.

Keywords: mask R-CNN, object detection, seagrass, segmentation

1. Introduction

Seagrass is a Spermatophyta plant that can live in the marine environment by fully adapting from low salinity levels to high salinity levels. Seagrass acts as a primary producer in the food chain in the waters, habitat for marine biota, protection of coastal areas, maintaining natural resources [1], producing organic carbon through absorption and decay [2], also seagrass can support local economic income for people in coastal areas [3]. Therefore, monitoring seagrass meadows and seagrass health conditions needs to be done to achieve a healthy marine ecosystem.

The step to detect and monitor the seagrass can be through various types of images, including spectral images-based satellites [4], acoustic images [5], underwater video images [6], and underwater digital images [7]. Researchers of marine ecology use underwater digital images for real-time monitoring of the distribution, health, and percent of seagrass coverage [8]. The underwater image data is processed and analyzed with computer vision technology. The use of computer vision in the field of marine ecology now is widely done because of its ability to process images to obtain information quickly and accurately, making it possible for researchers to monitor individuals or seagrass populations on a spatial and temporal scale [9].

Detection and analysis of images through computer vision becomes more accurate with the implementation of deep learning models. Deep learning has better feature extraction capabilities for object detection and classification compared to traditional machine learning methods such as random forest, support vector machine, maximum likelihood classifier, etc. In recent years, deep learning has been successfully applied to classification and detection in some sea objects such as fish [10-11],...
plankton [12-14], and coral [15-16] using underwater image data. However, so far, the implementation of deep learning for seagrass detection and segmentation is still little done. This prompted researchers to implement the Mask R-CNN (region-based convolutional neural network) algorithm to detect and segment seagrass species of *Enhalus acoroides* from underwater image data.

The purpose of the study is to implement and get information about the performance of the Mask R-CNN in detecting and segmenting the seagrass *Enhalus acoroides*.

2. **Methodology**

The proposed approach for *Enhalus acoroides* detection and segmentation is performed in five steps: data collection, color correction, data labelling, training, and finally, evaluation of the model.

2.1. **Data collection**

Seagrass image data was obtained from Mahesa Glagah Agung Satria research in 2019 on the Beralas Pasir Island, Bintan. Based on the research, an underwater video recording for 19 minutes was obtained with a resolution of 640x480 pixels and a framerate of 30 fps [17]. The underwater video is then extracted to get an image on each of its frames. The extraction process produces 20,823 images for the data selection process. The selected data must represent a variety of backgrounds and feature objects so that the learning outcomes of the Mask R-CNN algorithm become better. The results of the data selection were then divided by a ratio of 80%:10%:10% as data training, data validation, and data tests.

2.2. **Color correction**

The use of computer vision requires a clear image. Getting a clear picture in the field is very difficult, especially when influenced by lighting, wavelength, depth, and noise that occurs during the recording of underwater objects [18]. Therefore, the dataset used for training will go through the white balance process first for color correction and improving image quality. White balance is a process for calibration of white color or a process that aims to eliminate colors that are not real. Color correction with the white balance method is done by inserting the image and changed into 16-bit mode on the Adobe photoshop layer. The purpose of converting images to 16-bit mode is to minimize information loss. The next step to correct the color is detecting and arranging three color dots (black dots, gray dots, and white dots) on the image. After the white balance process is completed, the next step is the evaluation of image quality. The image can be said to have good quality if the image looks natural, the color tone doesn't look too white or too dark, does not look greenish or bluish. Images that already have good quality are then saved in .jpg format.

2.3. **Data labeling**

Training and validation data are processed with Labelme software (MIT, USA). Labelme is an open-source software developed and published by MIT to create annotations and labels on data [19]. The process of labelling training and validation data is done by digitizing seagrass using polygons tools. The data is divided into two classes, the 0th class is the background, and the 1st class is the *Enhalus acoroides*. The training data and validation data that have been labelled are then stored in JSON. The data labelling display is illustrated in Figure 1.
2.4. Training

The training process was performed with 450 images from 500 dataset images produced. The dataset used only as many as 500 images. This is because the images generated from video extraction have very low quality, so only 500 images are obtained with better quality for the training process. The training process should use large amounts of image data (more than 1000 datasets). However, because the number of datasets is small, transfer learning methods are used. The transfer learning technique is suitable to use in the convolutional neural network (CNN) for image classification with datasets of less than 1000 images. This method will quickly transfer the learning weight to a new task with a small number of training data [20].

Datasets for the training process are divided by an 80%:10%:10% ratio commonly used in machine learning. This division is based on the Pareto distribution principle, where in most cases, 80% of the effect comes from 20% of causes [21]. Training Mask R-CNN uses reference code derived from the Matterport repository (MIT, USA). The code can be accessed on the https://github.com/matterport/Mask_RCNN page. Several changes were made to the file train.py on the configuration and load of the dataset to adjust to the Enhalus acoroides dataset. Backbone Resnet101 is used for training because the accuracy of the Resnet101 is better than other models.

Preprocessing was done by resizing the image with a square mode that has a minimum scale of 400 pixels and a maximum scale of 512 pixels. The resizing image needs to be done to reduce the memory load on the GPU. In the training process, mini-batch gradient descent is used as an optimizer. The mini-batch gradient descent will randomly select multiple samples from the training data to calculate gradient descent and update the model parameter value. A batch size of 4 is used for optimization in the training process. Batch size selection of 4 due to the size of GPU memory is not too large. Batch sizes that are too large will slow down the training process, and it is difficult to reach the local minimum [22].

Coco weight was used in this study as a pre-trained weight model for transfer learning. The optimum weight value in this study was obtained after conducting the learning process with 100 epochs and 25 steps per epoch. The learning rate used is 0.001, with a total number of steps taken during the training process is 2500 steps to complete the training. Image augmentation such as random horizontal flip, random vertical flip, rotation, scaling, multiply, and Gaussian blur was used to multiply datasets to avoid overfitting.

2.5. Model evaluation

Validation produces images of seagrass that have been detected and segmented as seagrass species of Enhalus acoroides. Validation results are used to calculate accuracy, precision, and recall models using a confusion matrix. Accuracy describes how accurate the model is in classifying it correctly. Precision is the level of accuracy of the data requested by the results of predictions generated by the model. Recall or sensitivity is the ratio of True Positive predictions compared to the overall positive data. The mathematical expression of accuracy, precision, and recall are [23]:

---

Figure 1. Enhalus acoroides data labelling with Labelme.
accuracy = \frac{TP+TN}{TP+FP+FN+TN} \tag{1} \\
precision = \frac{TP}{TP+FP} \tag{2} \\
recall = \frac{TP}{TP+FN} \tag{3}

Where TP is True Positive, the amount of positive data is classified correctly by the system. FP is False Positive, which is the amount of positive data but classified incorrectly by the system. FN is a False Negative in which an object or ground truth bounding box is not detected. TN is True Negative is the amount of negative data predicted correctly by the model.

The seagrass coverage is calculated to evaluate the segmentation produced by the model against the shape of the actual object. The area of seagrass coverage was calculated using Adobe Photoshop software, with pen tools and histogram tools. The data used is the image test data before segmentation and the image test data after segmentation with Mask R-CNN. The pixel value of seagrass is obtained from the digitization process using pen tools, and then the digitization result region is displayed with the histogram. After all the data has finished passing through the digitization process, the pixel value then passes through a calculating process of seagrass coverage. The value of the seagrass coverage is obtained from the comparison of pixel values on seagrass objects with the overall pixel value of the image. Here is a formula for calculating the percentage of seagrass coverage:

\text{coverage} = \frac{n_i}{N_i} \times 100\% \tag{4}

Information:

\(n_i\) = Number of seagrass pixels in the i-figure
\(N_i\) = Total number of pixels in the i-figure

The final step to evaluating the segmentation generated by the model is to look for a correlation between the percent coverage before segmentation and the percent coverage after segmentation. The correlation value is obtained by using linear regression. Regression analysis is a technique for estimating the relationship among variables. The main focus of regression is to analyze the relationship between a dependent variable and one independent variable also formulates the linear equation between dependent and independent variables [24]. For this reason, we chose a regression method to obtain significant results. Based on the shape or direction of the relationship, the correlation coefficient value is expressed (-1 ≤ KK ≤ +1). If the correlation coefficient is positive, where the closer the value of the correlation coefficient to +1, the stronger the positive correlation between percent coverage before segmentation and percent coverage after segmentation by Mask R-CNN.

3. Result and discussion

3.1. Dataset

This research has produced a seagrass (Enhalus acoroides) dataset consisting of 500 data with a 640x480 pixels resolution. The dataset was divided with a ratio of 80%:10%:10% for training, validation, and testing data. Training data is the majority of datasets that will be used for the learning process. Training data will be seen by the model during the training process and will be used to learn the parameters of the model or the value of the connection weight between the nodes of the neural network [25]. Validation data is also used in the training process to validate the model performance at the end of each training iteration process. Test data is data that is not related and does not go through the training process. Test data is used to test the performance of the model. Figure 2 shows an example of a dataset that has been generated.
Figure 2. Seagrass datasets (a) before white balance, (b) after white balance.

3.2. Training result

The training process in this study uses transfer learning methods because the dataset used is relatively small. Transfer learning is suitable in the Convolutional Neural Network (CNN) for classification with datasets of less than 1000 images [26]. CNN's algorithm can recognize complex non-linear patterns on big data, so it takes a lot of time for the training process [27]. Transfer learning is a technique that can overcome this problem. This method will quickly transfer information to a new task with a small number of training data. Network customization with transfer learning is much faster and easier than building and training a new network from scratch. This is because transfer learning used a pre-trained weight model and then adjusted the architecture to include training parameters as needed. The pre-trainer model used was the coco weight from ms coco to take trained tissue on millions of images and train it for new objects classification. In this case, the new objects are seagrass (*Enhalus acoroides*). The training process on the seagrass (*Enhalus acoroides*) dataset is taken approximately 2 hours 40 minutes with a length per epoch of 59 seconds.

Figure 3. Loss of R-CNN Mask training process.
The error value (loss) is a measure of the error generated by the model. Loss values imply how well or poorly a model is designed and optimized in each iteration. Optimization is done by updating the vector weight value in each iteration until it reaches an optimal or minimum value [28]. A loss value that closes to 0 and decreases indicates that the model has better classification and high accuracy.

Figure 3 shows loss movements in the Mask R-CNN training process from the seagrass (*Enhalus acoroides*) dataset. Train loss is the value of loss in training data during the learning process. Train loss continuously decreased from 2.5314 to 0.6012 at the end of the training process. Train loss that is decreasing and approaching 0 indicates that the model is good at classifying objects in training data. Val loss is the value of a loss on validation data during the learning process. Val loss also continuously decreased from 2.2543 to 1.3765 at the end of the training process. The validation loss value moves down but not too close to 0 because the model has a wide variety. The high variation also led the model to be less consistent in predicting new data that had never been seen. Complex machine learning algorithms tend to have higher variations because they are more prone to overfitting certain data [29]. Train loss and validation loss values fluctuate in the 1st epoch to the 40th epoch then begin to stabilize until the end of the training process. Fluctuations can be minimized by improving model performance through regularization [30].

Getting the best hyperparameter can help models in generalizing better [31]. This suggests that efforts to get a good hyperparameter from the model are still required by any type of model to help improve classification accuracy. The best weight value in the 56th epoch was selected as the weight for the validation test. Weight value selection is based on the difference in value between training loss, and validation loss that is not too large, with the difference is not too large is expected to produce good accuracy in object classification.

### 3.3. Validation test
Validation tests were performed on 50 image test data that have never been through the training process. The validation test produced 50 seagrass images that contained bounding box classifier, category, confidence score, and masking in the form of maps colors. The bounding box describes the spatial location of the detected object. The bounding box is predicted to be true positive if the bounding box overlaps more than 50% with the ground truth box [32]. The category displayed is a description of the detected object. This study only displays the *Enhalus acoroides* category. The confidence score is a value that indicates the possibility in the anchor box, and there is an object detected. This value is obtained from the prediction in the RoIPool layer [33]. Instance segmentation is applied to produce masking in every object at the pixel level [34]. This means that each object with the same class on one image will be segmented as a different entity. This technique is particularly useful considering that datasets have multiple objects on a single image. The validation test results are represented in Figure 4.
A validation test is done to measure the model performance in detecting seagrass (Enhalus acoroides). Validation results obtained three classes to measure the performance of the model. The classes are seagrass detected correctly, seagrass detected wrong, and seagrass undetected. Seagrass detected correctly entered the True Positive (TP) parameter, indicated by Figure 4(a), with the correct bounding box, category, confidence score, and masking characteristics. Seagrass detected incorrectly entered into the False Positive (FP) parameter, indicated by Figure 4(b) with other objects detected as seagrass. Seagrass detected incorrectly has a bounding box, category, confidence score, and masking on other objects than seagrass. Seagrass not detected entered into the False Negative (FN) parameter, indicated by Figure 4(c), characterized by seagrass objects that do not have a bounding box, category, confidence score, and masking.

3.4. Model performance evaluation

Evaluation of model performance is done by looking for accuracy, precision, and recall values from the resulting model. Precision is the ability of an object detector to identify objects that are targeted detection [8]. Recall or sensitivity is the value used to measure the detection rate produced by the model [35].

| True label | Enhalus acoroides | Background |
|------------|-------------------|------------|
| Enhalus acoroides | 270 | 8 |
| Background | 14 | - |

Based on validation tests, the total Region of Interest (RoI) is 292 objects. 292 RoI is then grouped into three parameters: True Positive (TP), False Positive (FP), and False Negative (FN). True Positive is the sum of true detection values where the intersection of union (IoU) of bounding box
seagrass ≥ threshold value [8]. The threshold value used in this study was 0.5. False Positive (FP) is the number of misdetection seagrass where the intersection of union (IoU) of bounding box seagrass ≤ threshold value. The third parameter is False Negative, where the ground truth bounding box on the object of *Enhalus acoroides* seagrass is undetectable. Table 1 is a comparison of the results of classification by the model presented in the confusion matrix. True positive (TP) parameters or the number of seagrasses detected correctly is 270 objects, False positive (FP) parameters or the number of incorrect detection of seagrass produced is 14 objects, and false negative (FN) parameters or the number of undetected seagrasses produced is 8 objects. The True Negative (TN) parameter or the number of backgrounds detected correctly is not detected and displayed by the model. Therefore, the True Negative (TN) parameter is not calculated. The results of the confusion matrix are then used to calculate the accuracy, precision, and recall values of the model.

| Evaluation parameter | Value  |
|----------------------|--------|
| Accuracy             | 0.9246 |
| Precision            | 0.9507 |
| Recall               | 0.9712 |

Table 2 shows the results of accuracy, precision, and recall calculations from the model in detecting seagrass (*Enhalus acoroides*). The model produces an accuracy value of 0.9246. A high accuracy value means that the model is accurate in classifying seagrass correctly. The precision value is 0.9507, and the recall value is 0.9712. A large precision value indicates that the classifier produced by the model is accurate in detecting required objects. A large recall value indicates a very well-recognized class, the majority of all positive samples detected by the model [36]. Based on these results, it can be known that the model is able to detect the seagrass (*Enhalus acoroides*) well.

### 3.5. Seagrass coverage

The seagrass coverage is calculated to evaluate the segmentation produced by the model against the actual shape of the seagrass. The seagrass coverage is produced in the form of a percentage of cover on each data test image and segmentation image. The results of the seagrass coverage are represented in Figure 8 below.

![Figure 5](image-url) Figure 5. Percentage of seagrass coverage in test data image and segmentation image.

Based on Figure 5, the percentage of seagrass coverage in the test data image before going through the segmentation process ranges from 2.64% - 28.48%. The seagrass coverage on the
segmentation image has values ranging from 2.05% - 19.45%. Based on the calculation of the seagrass coverage, it is known that the percentage value of the seagrass coverage in the model segmentation image tends to be close to the actual area of the data test. However, the wide percentage of the seagrass coverage of segmentation produced by Mask R-CNN has a smaller value than segmentation by a human. This is because the results of segmentation by Mask R-CNN are not perfect according to the shape of the object. The results of Mask R-CNN segmentation are always smaller than segmentation by humans due to poor input image quality. So, at the end of the leaf or a small part of the seagrass model tends to be wrong or cannot segment perfectly.

![Figure 6. Correlation of seagrass coverage in the data test image and segmentation image.](image)

Evaluation of segmentation results produced by the model is done by comparing the coverage on the data test image and the segmentation results image. Based on Figure 6, the determination coefficient result by the model is 0.7693, and the correlation coefficient is 0.8771. Large correlation values indicate a strong correlation between the coverage in the data test image and the segmentation image. Based on a strong correlation and seagrass coverage segmentation results that tend to approach the area of the data test, the resulting model can be said to be good in segmenting seagrass objects.

4. Conclusion
The study successfully applied the Mask R-CNN (region convolutional neural network) algorithm model to detect and segment *Enhalus acoroides* seagrass. The results of this study showed an accuracy value of 0.9246, a precision value of 0.9507, a recall value of 0.9712 and a correlation coefficient value of 0.8771. These values indicate that the model is accurate in detecting seagrass correctly.

Acknowledgments
This work is partly supported by IPB Agromaritim Research Grant No. 7824/IT3.L1/PT.01.03/p/b/2021 to Prof. Indra Jaya.

References
[1] Waycott M, Mcmahon K, Mellors J, Calladine A and Kleine AD 2004 *A Guide to Tropical Seagrasses of the Indo-West Pacific* (Townsville: James Cook University Press) p 72
[2] Macreadie P I, Baird M E, Trevathan-Tackett S M, Larkum A W D and Ralph P J 2014 Quantifying and modelling the carbon sequestration capacity of seagrass meadows—a critical assessment *Mar. Pollut. Bull.* **83** 430-39
[3] Hossain M S, Bujang J S, Zakaria M H and Hashim M 2015 Application of Landsat images to seagrass areal cover change analysis for Lawas, Terengganu and Kelantan of Malaysia Cont. Shelf Res. 110 124–48

[4] Roelfsema C M, Phinn S R, Udy N and Maxwell P 2009 An integrated field and remote sensing approach for mapping seagrass cover, moreton bay, Australia J. Spat. Sci. 54 45–62

[5] Vasilijevic A, Miskovic N, Vukic Z and Mandic F 2014 Monitoring of seagrass by lightweight AUV: A Posidonia oceanica case study surrounding Murter island of Croatia 2014 22nd Mediterranean Conf. on Control and Automation (Palermo: IEEE) pp 758–63

[6] Norris J G, Wylie-Echeverria S, Mumford T, Bailey A and Turner T 1997 Estimating basal area coverage of subtidal seagrass beds using underwater videography Aquatic Bot. 58 269–87

[7] Gonzalez-Cid Y, Burguera A, Bonin-Font F and Matamoros A 2017 Machine learning and deep learning strategies to identify Posidonia meadows in underwater images Ocean 2017 – Aberdeen (Aberdeen : IEEE) 1–5

[8] Moniruzzaman M, Islam S M S, Lavery P and Bennamoun M 2019 Faster R-CNN based deep learning for seagrass detection from underwater digital images Digit. Image Comput.: Tech. Appl. DICTA pp 1–7

[9] LeCun Y, Bengio Y and Hinton G 2015 Deep learning Nature 521 436–44

[10] Li X, Shang M, Qin H, and Chen L 2016 Fast accurate fish detection and recognition of underwater images with Fast R-CNN Ocean 2015 - MTS (Washington : IEEE) pp 1-5

[11] Villon S, Chaumont M, Subsol G, Villéger S, Claverie T and Mouillot D 2016 Coral reef fish detection and recognition in underwater videos by supervised machine learning: Comparison between deep learning and HOG+SVM methods Int. Conf. on Advanced Concepts for Intelligent Vision System vol 10016 (Lecce : Springer) pp 160–71

[12] Py O, Hong H and Zhongzhi S 2016 Plankton classification with deep convolutional neural networks Conf. on Information Technology, Networking, Electronic and Automation Control (Chongqing : IEEE) pp 132–6

[13] Lee H, Park M and Kim J 2016 Plankton classification on imbalanced large scale database via convolutional neural networks with transfer learning Int. Conf. on Image Processing (Phoenix : IEEE) pp 3713-7

[14] Dai J, Wang R, Zheng H, Ji G and Qiao X 2016 ZooplanktontoNet: Deep convolutional network for zooplankton classification Ocean 2016 (Shanghai : IEEE ) pp 1-6

[15] Mahmood A, Bennamoun M, An S, Sohle F, Boussaid F, Hovey R, Kendrick G and Fisher R B 2016 Coral classification with hybrid feature representations Proc. Int. Conf. on Image Process. (Edinburgh : IEEE ) pp 519–23

[16] Elawady M 2015 Sparse Coral Classification Using Deep Convolutional Neural Networks MSc Thesis University of Girona, Girona

[17] Satria M G A 2019 Rancang bangun dan uji kinerja unmanned surface vehicle (usv) untuk pemetaan kualitas perairan dan vegetasi bawah air Msi Thesis IPB University, Bogor

[18] Prabowo D P and Pramunendar R A 2019 Implementasi metode CLAHE menggunakan parameter distribusi untuk meningkatkan kualitas citra objek bawah air Pros. SNST Fakultas Teknik 1 264-9

[19] Russell B C, Torralba A, Murphy K P and Freeman W T 2008 LabelMe: a database and web-based tool for image annotation Int. J. Comp. Vision 77 157-73

[20] Darmanto H 2019 Pengenalan spesies ikan berdasarkan kontur otolith menggunakan convolutional neural network Joined Journal: Journal of Informatics Education 2 41–59

[21] Newman M E J 2006 Power laws, Pareto distributions and Zipf’s law Contemp. Phys. 46 323–51

[22] Wu D, Yuan Y, Huang J and Tan Y 2020 Optimize TSK Fuzzy Systems for Regression Problems: Minibatch Gradient Descent with Regularization, DropRule, and AdaBound (MBGD-RDA) Trans. Fuzzy Syst. vol 28 (Wuhan: IEEE) pp 1003–15

[23] Vanderbeck S and Bockhorst J 2011 A machine learning approach to identifying section in legal briefs In: Visa S, Inoue A, Ralescu A, editor Proc. of Twenty-Second Midwest
Artificial Intelligence and Cognitive Science Conference 2011 (Wisconsin: Omnipress) pp 16-22

[24] Uyanik G K and Guler N 2013 A Study on Multiple Linear Regression Analysis Int. Conf. on Fourth New Horizons in Education vol 106 (Sakarya: ELSEVIER) pp 234-40

[25] Nugroho M A and Sebatubun M M 2020 Klasifikasi Varietas Kopi Berdasarkan Green Bean Coffee Menggunakan Metode Machine Learning J. Inf. Syst. Manag. 1 1–5

[26] Darmanto H 2019 Introduction to Fish Species Based on Otolith Conture Using the Convolutional Neural Network Joined J. 2 41–59

[27] Fan C, Xiao F and Zhao Y 2017 A short-term building cooling load prediction method using deep learning algorithms J. Appl. Energy 195 222–33

[28] Mahmood A, Bennamoun M, An S, Sohel F, Boussaid F, Hovey R, Kendrick G and Fisher RB 2016 Coral classification with hybrid feature representations Proc. Int. Conf. on Image Process (Edinburgh : IEEE) pp 519–23

[29] Doroudi S 2020 The Bias-Variance Tradeoff: How Data Science Can Inform Educational Debates AERA Open 6(4) 1-18

[30] Yamada Y, Iwamura M, Akiba T and Kise K 2019 Shakedrop regularization for deep residual learning IEEE Access 7 186126–36

[31] Ramdan A, Zilvan V, Suryawati E, Pardede HF and Rahadi VP 2020 Tea clone classification using deep CNN with residual and densely connections J. Tek. Sist. Komputer 8 289–96

[32] Hariharan B, Arbeláez P, Girshick R and Malik J 2014 Simultaneous detection and segmentation European Conf. on Computer Vision vol 7 (Switzerland : Springer) pp 297–312.

[33] He K, Gkioxari G, Dollár P and Girshick R 2017 Mask R-CNN Int. Conf. on Computer Vision vol 42 (Venice : IEEE) pp 386–97

[34] Arnab A and Torr PH 2017 Pixelwise instance segmentation with a dynamically instantiated network Conf. on Computer Vision and Pattern Recognition (Honolulu : IEEE) pp 441-50

[35] Lubin A 2018 Cyber law and espionage law as communicating vessels 10th Int. Conf. on Cyber Conflict vol 220 (Tallinn: NATO CCD COE Publications)

[36] Schlegl T, Waldstein S M, Bogunovic H, Endstraber F, Sadeghipour A, Philip A M, Podkowinski D, Gerendas B S, Langs G and Schmidt-Erfurth U 2018 Fully automated detection and quantification of macular fluid in OCT using deep learning J. Ophthalmology 125 549–58