

# Enhanced Localization and Classification of Coral Reef Structures and Compositions

Kirill Bogomasov, Philipp Grawe, and Stefan Conrad

Heinrich Heine University, Universitätsstraße 1, 40225 Düsseldorf, Germany  
<http://dbs.cs.uni-duesseldorf.de>  
{bogomasov,grawe,stefan.conrad}@hhu.de

**Abstract.** The automatic annotation of coral images is important for researching the underwater ecosystem, which is the focus of the ImageCLEFcoral task. We participated by refining our approaches from the last years challenge for localization and classification of corals within images of sea floor. Underwater images bear multiple difficulties which we tackle with applying image enhancement algorithms. To locate and classify the corals we applied multiple deep learning approaches and also revisioned our two-staged algorithm. The results show that deep learning approaches are the most convincing. Still, the localization of corals is the most challenging part for us, but we managed to increase our models performance significantly.

**Keywords:** Image Segmentation · Image Classification · Object Localization

## 1 Introduction

Monitoring coral reefs and their health is an important component to understand effects of the climate change on maritime life [8]. Experts annotate underwater images, who not only have to deal with the complex morphology of the corals but also the large number of pictures. Computer vision based localization and classification of corals seems to be a reasonable solution. Unlike typical datasets for object detection tasks, underwater images hold more problems regarding the image quality and thus need very specific features and preprocessing.

In this paper we present the improvements of our approaches which are based on the last years ImageCLEFcoral [3][5] submission. Additionally we used another deep learning approach, namely RetinaNet [13], since this seemed to be the most promising. The classical machine learning is revisioned, but still not compatible with deep learning approaches regarding its performance. Lastly we implemented a popular suggestion to increase the image quality by preprocessing the images with algorithms made for underwater photography [7][1].

Overall we increased the performance of our approaches and can provide more insights, which we present in the following sections.

---

Copyright © 2020 for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0). CLEF 2020, 22-25 September 2020, Thessaloniki, Greece.

## 2 Related Work

The research of last year coral task can be divided into classical feature engineering and deep learning approaches. Caridade and Marçal [4] used random forest classification, based on a selected feature set, consisting of color and texture features to localize and classify the substrate types. Jaisakthi et al. [10] used a faster R-CNN to solve this task. Another solution proposal presented by Steffens et al. [22] is based on a DCNN architecture. Our approach differs from the mentioned research. Considering the different properties, distributions and sizes of the corals, we rely on a combination of both categories of image processing. The good results substantiate our approach and make it one of the most promising so far.

## 3 Data

For the purpose of the task [9][6], a training dataset with 440 images and 12077 annotated substrates, which are labeled with one of 13 substrate types, is provided. An additional dataset contains 400 raw images, that is used for testing the predictions while no further information about the images is given to the participants.

Table 1: **Substrate types** with their relative frequency in the training set.

Class label	Relative frequency
c_algae_macro_or_leaves	0.00761463
c_fire_coral_millepora	0.00157259
c_hard_coral_boulder	0.13590465
c_hard_coral_branching	0.09774872
c_hard_coral_encrusting	0.07829829
c_hard_coral_foliose	0.01464989
c_hard_coral_mushroom	0.01845721
c_hard_coral_submassive	0.01638802
c_hard_coral_table	0.00173812
c_soft_coral	0.46871379
c_soft_coral_gorgonian	0.0074491
c_sponge	0.13996027
c_sponge_barrel	0.01150472

The substrate representatives have a highly imbalanced distribution as shown in table 1. A random split of the data can lead to an even more disadvantageous class distribution, since it can exclude the representatives of the rare classes in one of the sets or increase the impact of frequent classes on the set. We give an example with a split into train and validation subsets with a ratio of 80 : 20 which distribution can be seen in table 2. To prevent the tendency of high imbalance, we propose the following procedure similar to [22].

May  $P(D)$  be the relative distribution of classes of the complete image dataset  $D$ . Since both splits of the dataset should have the same distribution of classes,  $P(D)$  is our target distribution. May  $A$  and  $B$ , with  $A \cap B = \emptyset$ , be the two sets after splitting  $D$  with the relative distributions  $P(A)$  and  $P(B)$ . The key idea is to swap images between the two initial random splits  $A$  and  $B$  to make  $P(A)$ ,  $P(B)$  and  $P(D)$  as similar as possible. Let  $a, b$  with  $a \in A$  and  $b \in B$  be the two images to swap between  $A$  and  $B$ . We define  $A^* = (A \setminus \{a\}) \cup \{b\}$  and respectively  $B^* = (B \setminus \{b\}) \cup \{a\}$ . If the similarity between  $P(A^*)$  and  $P(D)$  is smaller than the similarity between  $P(A)$  and  $P(D)$ , we swap the items, so that  $A = A^*$  and  $B = B^*$ . Since the similarity between  $P(B^*)$  and  $P(D)$  decreases when the similarity between  $P(A^*)$  and  $P(D)$  decreases, this approach works w.l.o.g. The loop over the images is running until there are no more swaps, i.e. no swap increases the similarity. To measure this similarity we use Jensen-Shannon divergence [11]. Since this procedure only converges to a local, but not global optimum, not every random split ends up in the same balanced split but consequently the optimal split is not found every time. Further research is needed to evaluate this approach, in regards of optimizations and metrics. The results of the balancing algorithm are shown in the table 3.

The Jensen-Shannon divergence between the train and validation set before swapping is 0.040313, whereas after swapping its divergence is 0.0061. This is an improvement by a factor of almost 8.

Table 2: **Substrate distribution** before balancing.

Class label	Relative frequency train	Relative frequency valid
c_algae_macro_or_leaves	0.00799747	0.005827505
c_fire_coral_millepora	0.00126276	0.00271950272
c_hard_coral_boulder	0.12722298	0.167443667
c_hard_coral_branching	0.0967063	0.101787102
c_hard_coral_encrusting	0.07965906	0.0730380730
c_hard_coral_foliose	0.01515311	0.012810513
c_hard_coral_mushroom	0.01894139	0.0167055167
c_hard_coral_submassive	0.01746817	0.01243201
c_hard_coral_table	0.0021046	0.000388500389
c_soft_coral	0.47606019	0.442113442
c_soft_coral_gorgonian	0.00683995	0.00971250971
c_sponge	0.13869304	0.144910645
c_sponge_barrel	0.01189098	0.101010101

## 4 Approaches

The "Coral reef image annotation and localisation task" can be divided into two tasks. Segmentation of various coral objects from images of sea ground and classification of those with their specific type of one of the 13 known types of

Table 3: **Substrate distribution** after balancing.

Class label	Relative frequency train	Relative frequency valid
c_algae_macro_or_leaves	0.00758972	0.00735809
c_fire_coral_millepora	0.00140952	0.00210231
c_hard_coral_boulder	0.13574759	0.13594954
c_hard_coral_branching	0.09779898	0.09775753
c_hard_coral_encrusting	0.07828255	0.07813595
c_hard_coral_foliose	0.01474574	0.0143658
c_hard_coral_mushroom	0.01843218	0.01857043
c_hard_coral_submassive	0.01637211	0.01646811
c_hard_coral_table	0.00173479	0.00175193
c_soft_coral	0.46882793	0.4688157
c_soft_coral_gorgonian	0.0074813	0.00735809
c_sponge	0.14008457	0.13980378
c_sponge_barrel	0.01149301	0.01156272

substrates.

Due to multiple difficulties that underwater images bear, strategies to enhance the image quality which should help to find better features are used. We applied two of the state-of-the-art deep learning approaches and additionally combined these with an improvement of our own development [3]. Those approaches are presented in the following subsections.

#### 4.1 Image Enhancement

Underwater images inherit problems, like the attenuation of light or the suspension of particles reflecting the light. Those conditions distort colors and visibility, which affect the performance of machine learning algorithms. Therefore we applied and evaluated multiple enhancement algorithms, that are specialized on underwater images.

Ancuti et al. [1] use Fusion [24] and work without knowledge of a physical model of the lighting conditions. Two derivations of the image, improving the white balance and the contrast, are fused together using different weight measures to restore the image. The authors show that they retrieve more features using SIFT [15] through applying their image enhancement. Ghani and Isa [7] use Rayleigh-stretching, as well as stretching using the HSV color model, to first correct the contrast and then correct the color. Further on the processings are referred to as Fusion and RD. Since we do not have access to already corrected images of the instant dataset, we use three evaluation measures that predict how human would perceive the image, based on learned examples. Namely the measures are BRISQUE [16], NIQE [17] and PIQUE [23]. For all of them smaller values mean better image quality.

The evaluation [18] of the dataset of 2019 showed an improvement in image quality using the two enhancement algorithms, as seen in table 4 as well as in the subsections 4.4 and 5. Example images are shown in figure 1.

Table 4: **Image enhancement** algorithms evaluated on ImageCLEF Coral dataset 2019.

Algorithm	BRISQUE	NIQE	PIQUE
<i>None</i>	25.98	3.61	26.34
Fusion	22.66	3.43	30.99
RD	<b>20.92</b>	<b>2.94</b>	<b>25.06</b>

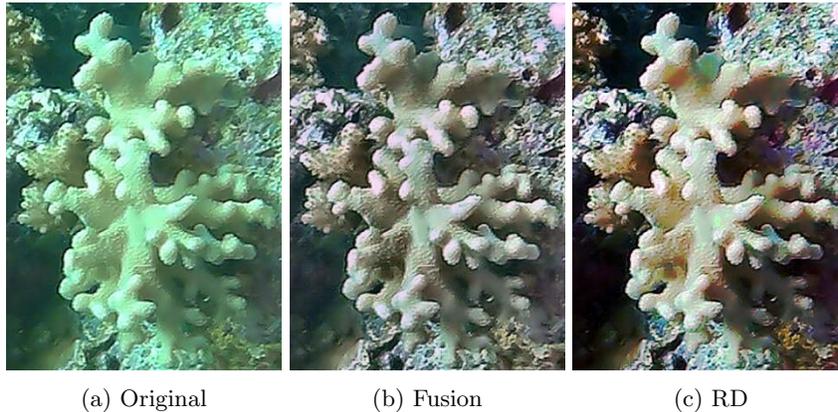


Fig. 1: Comparison of the image enhancements.

## 4.2 Yolo - Improvement

Neural networks are still state of the art in segmentation and classification tasks. Last year we used Yolov3 [20] as our main neural network approach. Especially areas with a particularly large denseness of smaller corals were challenging. The reason for this is on one hand Yolo's native ROI restraints which sets a natural limit on the regions considered within a certain area. On the other hand we were limited by the input data size with the largest resolution we could use with GPU of  $608 \times 608$  pixels. Regarding the original resolution of images of  $4032 \times 3024$ , we preserve a scaling factor of at least of 5, i.e. each pixel represents an area of more than 25 pixels in the original image. Considering that the smallest corals

consists of  $12 \times 12$  pixels, the information loss is significant. We split the images into overlapping subimages of  $608 \times 608$  pixels and trained the network on unscaled input. Consequently we got a large amount of images which also meant that the training time of our network was almost one month. Unfortunately, there was a mistake in the training data, therefore a revision of the results of the last years challenge was just recently announced. This left us with not enough time to retrain our working setup.

### 4.3 RetinaNET

For our second neural network approach we chose RetinaNet. The network showed impressive results on the COCO dataset and outperformed Yolov2 with a 17% higher  $AP_{0.5}$  value. RetinaNet is well suited, since it is able to produce more predictions and is capable to work with less balanced data. At its core, the architecture consists of the following components: a feature pyramid network [12] (based on Resnet), a regressor for bounding box prediction and a classifier. Basically, it is a one-stage detector. The particular advantage of RetinaNet is the focal loss [14]. In case of end-to-end object detection, background predictions oftentimes dominate. The optimizer rates the prediction as correct and the loss of the positive background prediction forms the complete return loss. This mostly leads to an optimizer return value of zero for the background areas in case of cross entropy and thus reduce the loss. Focal loss weights the positive samples higher and ensures that the network performs better on unbalanced data.

Right suited anchor boxes are the key to quality of object detection for any architecture that works with a "regions of interest". If the anchors are not properly prepared, the network has in many cases no chance of finding particularly small, neither large objects. In our dataset, we experience a wide variety of sizes of corals. Starting from a box size of  $18 \times 9$  to a size of  $3966 \times 2662$  pixels, the standard deviation of the areas is 629 assuming a square size. That means that irregular or peripherally sized objects present a special challenge. To tackle this problem, we chose a solution that was originally used on medical data [25]. In our opinion, the potential for improvement can be easily transferred to coral context, since tumors and nodules as smaller objects are comparable to coral objects of small size.

### 4.4 Own Developments

Besides the deep learning we also increased the performance of our classic feature based approaches. We evaluated the use of principal component analysis (PCA) [19] to select the best features, which increased the performance slightly. Apart from the features the choice of the classifier is important. Last year we used k-NN, which is depended from the parameter  $k$ . To overcome the search for the right parameter we evaluated the use of naïve bayes [21] for locating and classifying substrates.

When classifying the coral areas and non-coral areas, the features along with the approach are the same as in [3], which is illustrated in figure 2. A problem

that can be seen with k-NN is the low precision. This is due to labeling non-coral areas as coral areas, because most often water gets falsely classified. There are multiple ways to evaluate, as well as multiple things to evaluate. Beside the pure evaluation of the coral and non-coral tiles, we also evaluate the bounding boxes that enclose those coral tiles. In the end the evaluation of the found bounding boxes is more significant, but the pure evaluation of the coral and non-coral tiles helps with the assessment of the performance of finding bounding boxes in the generated black and white images (see figure 2b - 2d). Likewise this creates the opportunity to compare different image enhancement algorithms. All results are discussed in the following.

Table 5 holds the evaluation of the coral/non-coral grid that is compared with a grid representing the ground truth. The results show that the naïve bayes classifier has increased the accuracy, as well as the precision but greatly decreased the recall. PCA on the other hand did not have a big impact.

It could be shown that image enhancement increases the performance, even if just slightly. Table 6 showed that connected components works much better with the naïve bayes classifier than k-NN. The naïve bayes classifier has an insignificantly increased, whereas k-NN has significantly worse performance. We believe that this is caused by k-NN having too many false-positives which results in big boxes, that cover rather more than less area. Another indication for this assumption is the decrease of precision but increase of recall (compare figure 2b).

Table 5: **Evaluation** of the coral/non-coral classification, based on the tiles.

Approach	Image enhancement								
	<i>None</i>			Fusion			RD		
	Acc	Prec	Rec	Acc	Prec	Rec	Acc	Prec	Rec
k-NN	0.617	0.4496	0.5283	0.5823	0.4066	0.4813	0.6146	0.4355	0.4275
k-NN with PCA	0.6146	0.4471	<b>0.5343</b>	0.5858	0.4075	0.4637	0.6171	0.4371	0.4134
Bayes with PCA	0.6488	0.4545	0.1317	0.6570	0.3550	0.0031	<b>0.6575</b>	<b>0.4857</b>	0.0192

Table 6: **Evaluation** of the coral/non-coral classification, based on the bounding boxes.

Approach	Image enhancement								
	<i>None</i>			Fusion			RD		
	Acc	Prec	Rec	Acc	Prec	Rec	Acc	Prec	Rec
k-NN	0.4234	0.3920	0.8755	0.3919	0.4386	0.8543	0.4812	0.3973	0.8302
k-NN with PCA	0.4088	0.4088	<b>0.8927</b>	0.4097	0.4345	0.83106	0.4879	0.3888	0.8186
Bayes with PCA	0.6518	0.4833	0.2558	0.6571	0.3373	0.0025	<b>0.6606</b>	<b>0.5948</b>	0.0254

We also used the same two classifiers for classifying the bounding boxes. While using the same set of features, k-NN outperformed naïve bayes by far. Image enhancement shows its contribution again with tripling the accuracy of the naïve bayes classier, as seen in table 7. This substantiate that the enhancement proposed by Ghani and Isa [7] leads to increased performance with our dataset.

Table 7: **Evaluation** of substrate classification. The given values represent the accuracy.

Approach	Image enhancement		
	<i>None</i>	Fusion	RD
k-NN	0.4332	<b>0.4550</b>	0.4374
Bayes	0.1261	0.3328	0.3672

## 5 Evaluation of the Submitted runs

The evaluation was processed on test data that consists of images from four different geographical regions than the training set:

- same location
- similar location
- geographically distinct but ecologically connected
- geographically and ecologically distinct

In total the test dataset has 400 images, made by 100 images per subset. The results are examined in more detail below, while the interesting and informative values are discussed in the text. For the complete list of results, the reader is referred to the task overview working note [6]. Each submitted result was produced by one of our approaches, all of which were trained on the full training set.

The classification of the substrate types based on the classic features alone fails largely with an  $MAP_0$  value of around 27.4. and an  $MAP_{0.5}$  of 1. Although we improved the  $MAP_{0.5}$  value compared to last year by 300 percent, the results are still not really useful due to the low absolute values. This applies to both experiments, for the classification by means of k-NN based on the chosen features boosted by PCA and to statistical label assignment as well.

Our neuronal Network based approaches show a significantly better performance.

Since we had a limitation in the number of submissions, we chose none linear composition of available options. Our pool of options consisted alongside to classical features of RetinaNet and Yolov3, which both were trained on the unpreprocessed data and also on enhanced images by RD and Fusion. In addition, we worked with a variation of threshold  $\tau \in \{0.001, 0.1, 0.2, 0.5\}$ , which limits the MAP.

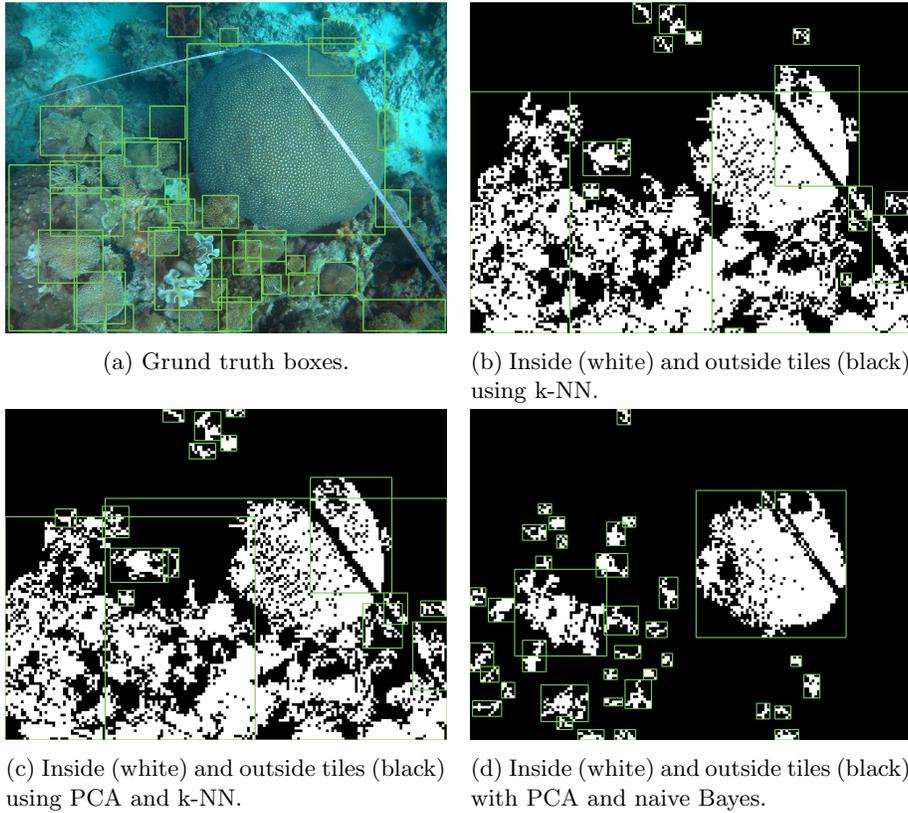


Fig. 2: Visualized process of localization corals with our approach. The raw picture is taken from the ImageCLEFcoral dataset [6].

We achieve our best result with an ensemble of RetinaNet and Yolov3. Whereby the predictions of RetinaNet were extended by the predictions of Yolov3. Both Networks were trained on RD-preprocessed images, with a  $\tau$  of 0.1. The combination of both systems results in a  $\text{MAP}_{0.5}$  of 39.2 % and  $\text{MAP}_0$  of 80.6 %. It is noticeable that we predicted very few bounding boxes for a  $\text{MAP}_{0.5}$  both through Yolov3 and through RetinaNet, additionally the found predictions were far from being present in all of the images.

In case of reduction of the accepted overlap, we get significantly more bounding boxes which also leads to a higher chance of hitting the right coral within the test images on cost of our overall accuracy. However, if we increase the accuracy, the  $\text{MAP}$  value drops. The following are significant examples: RetinaNet ( $\tau = 0.01$ ) combined with Yolov3 ( $\tau = 0.01$ ) has an  $\text{MAP}_{0.5}$  of 0.303 and an  $\text{MAP}_0$  of 0.727, but only an overall accuracy of 7 %. In contrast to it RetinaNet alone produces significantly fewer boxes with ( $\tau = 0.2$ ), but achieves the best overall

accuracy of 14.2 % with an  $\text{MAP}_{0.5}$  of only 30.3 and a  $\text{MAP}_0$  of 66.3 % while the accuracy is 10 % higher than in our best run.

A possible explanation is probably that RetinaNet has not finished training in 35 epochs. The large amount of predictions with a low overlap value could probably be boosted by non-maximum suppression. However, we were not aware of this problem until the results were published.

When provided with such a variety of possibilities, it can not be clearly determined which enhancement variant is the best choice. Certainly enhanced images improve the predictions, while we suspect that RD is superior to Fusion for coral images.

### 5.1 Transferability of the results within the test data subsets

One major question is the transferability of the results within the test dataset, considering that results, that form the average measure, vary widely. For our best run the  $\text{MAP}_{0.5}$  for *same location* increases by 6 percent to 45.7 % and by 1 % to 81.5 % the  $\text{MAP}_0$ . For *similar location*, however, it drops to 28.3 % for  $\text{MAP}_{0.5}$ , but reaches a value of 86.4 % for  $\text{MAP}_0$  and thus has the highest score among all presented submissions by all participants. Our best approach performs very well on *geographically similar* data. The  $\text{MAP}_{0.5}$  is 42.6 % and the  $\text{MAP}_0$  81.2 %. Overall, this part of the test dataset seems to be less complex, which is shown by the fact that the performance of almost all of our approaches have an increased performance on it.

It is also worth mentioning that we perform significantly worse on *geographically distinct* data with just an  $\text{MAP}_{0.5}$  0.125 and an  $\text{MAP}_0$  0.362. This tendency is also evident in the other approaches we used. The data seems to differ significantly, we either generalize less or the classes that are harder to recognize are more present. Some additional, yet unknown substrate types may be more dominant in geographically and ecologically distinct rocky reef and lead to distorted results. A further investigation of the dataset is required.

## 6 Conclusion

Overall, our approaches show significantly better results than last year. A comparison of our best approaches between the two years shows that we have improved the  $\text{MAP}_{0.5}$  by 13.4 %. The classification of the tiles into in- and outside boxes could also be improved by a Bayesian classifier, but is still far from being accurate.

Image enhancement techniques on the last years data were confirmed by the evaluation of the current test dataset, which leads us to the conclusion that the correction of blurry images in terms of contrast and sharpness is necessary. The RD algorithm makes the greatest contribution to improving the quality of the images. Less noteworthy results are made with the classical feature engineering approach. A deeper examination of the features and their information value is needed.

It can be assumed that the fair partitioning of images according to our balancing strategy, also has significantly contributed to the improved results. Deep learning strategies generalize quite well and are superior when using for this task. Especially the performance of RetinaNet, since it is not only better on the coco dataset than the state of the art. So far, the complexity of the images can hardly be handled by a single approach. We still see the most potential in an ensemble of several architectures. The combination of advantages of different approaches is the key to a stable solution. Since the amount of data has grown, while remaining relatively small, we still cannot exclude the potential of classic machine learning. Usually neural networks show a much better performance with an increased amount of data. For this reason, the images should be split into overlapping images and thus increase the number of training samples.

For future approaches, we recommend the usage of more specific features that are suitable for corals, such as used in [2]. However, we would rather rely on an ensemble of neural networks of the gold standard, which we would reduce in depth to shorten the training time and increase in terms of input resolution to decrease the information loss.

## References

1. Ancuti, C., Ancuti, C.O., Haber, T., Bekaert, P.: Enhancing underwater images and videos by fusion. In: 2012 IEEE Conference on Computer Vision and Pattern Recognition. pp. 81–88. IEEE (2012)
2. Ani Brown Mary, N., Dharma, D.: Coral reef image classification employing improved ldp for feature extraction. *J. Vis. Comun. Image Represent.* **49**(C), 225–242 (Nov 2017). <https://doi.org/10.1016/j.jvcir.2017.09.008>, <https://doi.org/10.1016/j.jvcir.2017.09.008>
3. Bogomasov, K., Grawe, P., Conrad, S.: A two-staged approach for localization and classification of coral reef structures and compositions. In: Working Notes of CLEF 2019 - Conference and Labs of the Evaluation Forum, Lugano, Switzerland, September 9-12, 2019 (2019), [http://ceur-ws.org/Vol-2380/paper\\_106.pdf](http://ceur-ws.org/Vol-2380/paper_106.pdf)
4. Caridade, C.M.R., Marçal, A.R.S.: Automatic classification of coral images using color and textures. In: CLEF (2019)
5. Chamberlain, J., Campello, A., Wright, J.P., Clift, L.G., Clark, A., García Seco de Herrera, A.: Overview of ImageCLEFcoral 2019 task. In: CLEF2019 Working Notes. CEUR Workshop Proceedings, CEUR-WS.org (2019)
6. Chamberlain, J., Campello, A., Wright, J.P., Clift, L.G., Clark, A., García Seco de Herrera, A.: Overview of the ImageCLEFcoral 2020 task: Automated coral reef image annotation. In: CLEF2020 Working Notes. CEUR Workshop Proceedings, CEUR-WS.org (2020)
7. Ghani, A.S.A., Isa, N.A.M.: Underwater image quality enhancement through composition of dual-intensity images and rayleigh-stretching. *SpringerPlus* **3**(1), 757 (2014)
8. Hoegh-Guldberg, O., Mumby, P.J., Hooten, A.J., Steneck, R.S., Greenfield, P., Gomez, E., Harvell, C.D., Sale, P.F., Edwards, A.J., Caldeira, K., et al.: Coral reefs under rapid climate change and ocean acidification. *science* **318**(5857), 1737–1742 (2007)

9. Ionescu, B., Müller, H., Péteri, R., Abacha, A.B., Datla, V., Hasan, S.A., Demner-Fushman, D., Kozlovski, S., Liauchuk, V., Cid, Y.D., Kovalev, V., Pelka, O., Friedrich, C.M., de Herrera, A.G.S., Ninh, V.T., Le, T.K., Zhou, L., Piras, L., Riegler, M., Halvorsen, P., Tran, M.T., Lux, M., Gurrin, C., Dang-Nguyen, D.T., Chamberlain, J., Clark, A., Campello, A., Fichou, D., Berari, R., Brie, P., Dogariu, M., Ștefan, L.D., Constantin, M.G.: Overview of the imageclef 2020: Multimedia retrieval in medical, lifelogging, nature, and internet applications. In: *Experimental IR Meets Multilinguality, Multimodality, and Interaction. Proceedings of the 11th International Conference of the CLEF Association (CLEF 2020)*, vol. 12260. LNCS Lecture Notes in Computer Science, Springer, Thessaloniki, Greece (September 22-25 2020)
10. Jaisakthi, S.M., Mirunalini, P., Aravindan, C.: Coral reef annotation and localization using faster r-cnn. In: *CLEF (2019)*
11. Lin, J.: Divergence measures based on the shannon entropy. *IEEE Transactions on Information theory* **37**(1), 145–151 (1991)
12. Lin, T., Dollár, P., Girshick, R.B., He, K., Hariharan, B., Belongie, S.J.: Feature pyramid networks for object detection. *CoRR* **abs/1612.03144** (2016), <http://arxiv.org/abs/1612.03144>
13. Lin, T.Y., Goyal, P., Girshick, R., He, K., Dollár, P.: Focal loss for dense object detection. In: *Proceedings of the IEEE international conference on computer vision*. pp. 2980–2988 (2017)
14. Lin, T.Y., Goyal, P., Girshick, R., He, K., Dollar, P.: Focal loss for dense object detection. *2017 IEEE International Conference on Computer Vision (ICCV) (Oct 2017)*. <https://doi.org/10.1109/iccv.2017.324>, <http://dx.doi.org/10.1109/ICCV.2017.324>
15. Lowe, D.G.: Distinctive image features from scale-invariant keypoints. *International journal of computer vision* **60**(2), 91–110 (2004)
16. Mittal, A., Moorthy, A.K., Bovik, A.C.: No-reference image quality assessment in the spatial domain. *IEEE Transactions on image processing* **21**(12), 4695–4708 (2012)
17. Mittal, A., Soundararajan, R., Bovik, A.C.: Making a “completely blind” image quality analyzer. *IEEE Signal processing letters* **20**(3), 209–212 (2012)
18. Nguyen, H.C.: Comparison of methods for underwater image improvement (2020)
19. Pearson, K.: Liii. on lines and planes of closest fit to systems of points in space. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* **2**(11), 559–572 (1901)
20. Redmon, J., Farhadi, A.: Yolov3: An incremental improvement. *arXiv preprint arXiv:1804.02767* (2018)
21. Rish, I., et al.: An empirical study of the naive bayes classifier. In: *IJCAI 2001 workshop on empirical methods in artificial intelligence*. vol. 3, pp. 41–46 (2001)
22. Steffens, A., de A. Campello, A.C., Ravenscroft, J., Clark, A., Hagra, H.: Deep segmentation: using deep convolutional networks for coral reef pixel-wise parsing. In: *CLEF (2019)*
23. Venkatanath, N., Praneeth, D., Bh, M.C., Channappayya, S.S., Medasani, S.S.: Blind image quality evaluation using perception based features. In: *2015 Twenty First National Conference on Communications (NCC)*. pp. 1–6. IEEE (2015)
24. Xydeas, C., Petrovic, V.: Objective image fusion performance measure. *Electronics letters* **36**(4), 308–309 (2000)
25. Zlocha, M., Dou, Q., Glocker, B.: Improving retinanet for ct lesion detection with dense masks from weak recist labels. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. pp. 402–410. Springer (2019)