

SCIENTIFIC REPORTS



OPEN

Compact-Morphology-based poly-metallic Nodule Delineation

Timm Schoening¹, Daniel O. B. Jones² & Jens Greinert¹

Poly-metallic nodules are a marine resource considered for deep sea mining. Assessing nodule abundance is of interest for mining companies and to monitor potential environmental impact. Optical seafloor imaging allows quantifying poly-metallic nodule abundance at spatial scales from centimetres to square kilometres. Towed cameras and diving robots acquire high-resolution imagery that allow detecting individual nodules and measure their sizes. Spatial abundance statistics can be computed from these size measurements, providing e.g. seafloor coverage in percent and the nodule size distribution. Detecting nodules requires segmentation of nodule pixels from pixels showing sediment background. Semi-supervised pattern recognition has been proposed to automate this task. Existing nodule segmentation algorithms employ machine learning that trains a classifier to segment the nodules in a high-dimensional feature space. Here, a rapid nodule segmentation algorithm is presented. It omits computation-intensive feature-based classification and employs image processing only. It exploits a nodule compactness heuristic to delineate individual nodules. Complex machine learning methods are avoided to keep the algorithm simple and fast. The algorithm has successfully been applied to different image datasets. These data sets were acquired by different cameras, camera platforms and in varying illumination conditions. Their successful analysis shows the broad applicability of the proposed method.

Poly-metallic nodules (PMNs - also referred to as manganese nodules) are a marine mineral resource. They contain relevant concentrations of Copper, Nickel and Cobalt¹. Most PMN research and resource assessment focusses on the deep abyssal plains in the eastern Pacific Ocean. Economically interesting areas were found within the *Clarion Clipperton Zone* (CCZ) where several countries are exploring PMN claims. These are license areas provided to the countries through the International Seabed Authority.

International research projects are investigating aspects of nodule mining (*Ecological Aspects of Deep-Sea Mining*: <https://jpio-miningimpact.geomar.de>; *Managing Impacts of Deep-Sea Resource Exploitation*: <https://www.eu-midas.net/>). Current foci of these research projects lie on environmental aspects: e.g. epifauna associated with PMNs, mining-induced habitat disturbances and the resilience of the ecosystem to cope with this anthropogenic pressure^{2,3}.

Quantitative measurements of PMN abundance are required for various applications. Individual nodule sizes (in cm²) provide basic measurements that can be aggregated to create descriptive spatial statistics. Examples are PMN density (nodules/m²) and seafloor coverage (in %) that data can related to environmental parameters, like faunal densities or chemical gradients.

Quantitative nodule information is traditionally provided by two methods. First, by physical sea floor sampling, e.g. using *box cores*. Second, by hydro-acoustic sensors, e.g. *Multi-beam Echo Sounder* (MBES) or *Side Scan Sonar* (SSS). Physical sampling by box cores provides precise size measurements of individual PMNs at minimum areal coverage. MBES and SSS provide high areal coverage but minimum quantitative resolution.

It is possible to correlate hydro-acoustic data and physical sampling for specific PMN abundances, but nodules need to be frequent and homogeneously distributed. Box cores can provide misleading data in heterogeneously populated areas and might even provide empty measurements despite an actual PMN occurrence in sparsely populated areas.

Optical imaging provides a bridge technology between these two methods. It features higher quantitative resolution than hydro-acoustics and higher areal coverage than physical sampling (see Fig. 1)⁴. It can be used to measure local PMN heterogeneity as well as regional abundance.

Semantics have to be assigned to pixels to determine nodule quantities in optical imagery. The objects of interest, i.e. PMNs, have to be segmented from the sediment background making this semantic segmentation is a

¹GEOMAR Helmholtz Centre for Ocean Research, Kiel, Germany. ²National Oceanography Centre, Southampton, UK. Correspondence and requests for materials should be addressed to T.S. (email: tschoening@geomar.de)

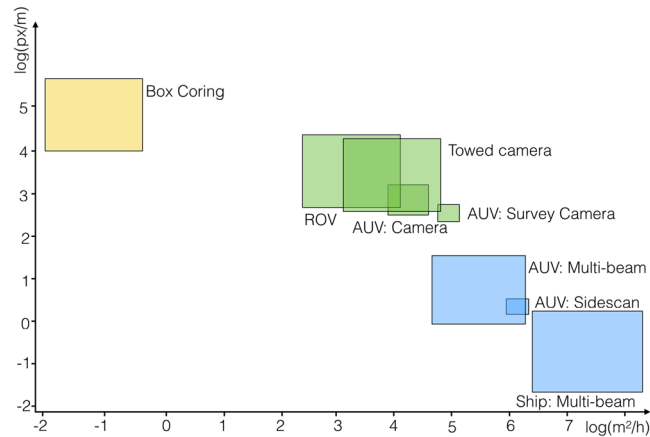


Figure 1. Comparison of mapping and sampling technology by their resolution (in $[\log_{10}(\text{px/m})]$) and aerial coverage (in $[\log_{10}(\text{m}^2/\text{h})]$). Physical sampling with box cores (yellow box) provides maximal resolution at minimal aerial coverage while hydro-acoustic (Multi-beam Echo Sounder (MBES) and Side Scan Sonar (SSS) - blue boxes) methods provide minimum resolution at maximum aerial coverage. Optical imaging (green boxes) provides a bridge technology that enables seafloor measurements at intermediate resolution and aerial coverage. Hydro-acoustic and optical methods are separated by platform (Ship, AUV, ROV, etc.) and sensor (MBES, SSS, Camera (e.g. Survey Camera⁵), etc.).

binary task. Each pixel of an image needs to be assigned to either the PMN class ω_0 or the sediment background class ω_1 .

After the segmentation, individual PMNs need to be delineated. This delineation provides nodule sizes in cm^2 - like physical sampling. By aggregating those size measurements imaging provides spatial nodule abundance statistics for large seafloor areas - like hydro-acoustic mapping. Examples of spatial statistics are size distributions, seafloor percent coverage and the number of nodules per area.

Traditionally, manual image annotation is conducted to add semantic information to images and specific tools for manual annotation of underwater imagery have been developed (e.g. *Squidle* or *BIIGLE*^{6,7}). In manual annotation, human experts inspect images and perform two steps: detection and classification of objects of interest. Detections are quantified by geometrical markers placed on top of the images (e.g. rectangles, polygons). Classifications are quantified by category names (e.g. “nodule”).

Classification is easy for the binary nodule case (ω_0 or ω_1), detection however is complicated. Nodules are embedded within the sediment and a gradual transition is frequent between the two classes. Delineating nodules manually can not be achieved in a pixel-perfect way and early methods use percent coverage to measure nodule abundance in images^{8,9}. These coverage measurements have to be seen as subjective *inspections* rather than objective *annotations*. Also, all manual image interpretation - inspection and annotation - is prone to bias by human factors like fatigue. Several studies showed that manual underwater image annotation is an error-prone task^{10–12} as it is in other image analysis domains¹³.

Pattern Recognition (PR) has been proposed to reduce human observer bias. PR-based image analysis methods are usually developed for imagery acquired in air. Their application to in-water imagery is complicated by the underwater environment due to physical factors like scattering, wavelength-dependent light absorption and inhomogeneous illumination. Biological factors add turbidity, marine snow and biofouling of image acquisition gear as further challenges. Some underwater imagery can be pre-processed to make in-air PR methods applicable¹⁴.

Successful underwater application of PR has been performed for specific use cases. Fish were classified by a deep neural network¹⁵ and by manually tuning shape features¹⁶. Nematode biomass was assessed by manually tuning intensity thresholds¹⁷ and scallop abundance has been quantified using Adaboost¹⁸. General-purpose megafauna detection has been initiated using rich feature representations and Support Vector Machines¹¹.

Efforts to automate the PMN segmentation have emerged as well but because of the aforementioned challenges of underwater imaging, traditional segmentation methods like *Region growing*¹⁹, *Pyramid linking*²⁰, *Normalized-cuts*²¹, *Mean-shift*²² and *Otsu thresholding* are not directly applicable²³. Otsu thresholding for example heuristically tunes a threshold value based on the colour frequencies in an image but does not take the spatial distribution of colours into account.

Successful PMN segmentation methods are characterised by varying manual annotation efforts and intense computational runtimes, which is higher than the image acquisition time. An early approach to assess PMNs provides seafloor percent coverage only⁹. This approach requires few manual annotations as input for an unsupervised pixel clustering. A more sophisticated method implements manual annotation of clusters in a feature space²⁴. Parts of this ‘pixel clustering’ method have been sped up using high-performance computing techniques²⁵. Other parts of this method were fully automated at increased computational cost based on a compactness criterion²³.

All methods mentioned above were tuned for specific image data sets. Those data sets are characterised by a particular illumination pattern, prevalence of backscatter and camera view angle. One method was successfully

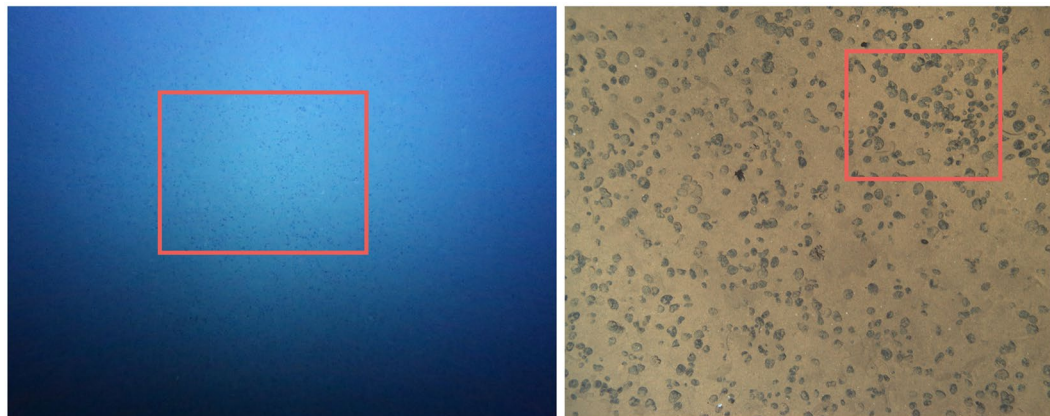


Figure 2. Two example images from $I^{(1)}$ (left) and $I^{(2)}$ (right). Both were acquired by an autonomous underwater vehicle (AUV). The left image was acquired by the *Deep Survey Camera* on board the GEOMAR AUV *Abyss*⁵. *Abyss* flew in an altitude of ca. 8.9 m above the seafloor. Image in $I^{(1)}$ on average show a seafloor area of ca. 180 m². The image on the right was acquired by NOCS' AUV *Autosub6000*. It flew in an altitude of ca. 2.6 m. Images in $I^{(2)}$ show a seafloor area of ca. 1.4 m². Red boxes mark sections that are shown as zoom-ins in following figures.

linked to hydro-acoustic measurements and contributed to PMN abundance assessment on a spatial scale larger than 1,000 m²²⁶.

Here, the *Compact Morphology-based Nodule Delineation* (CoMoNoD) method is presented. It combines the advantages of the above-mentioned methods but overcomes some of their individual shortcomings. *CoMoNoD* competes by neglecting time-consuming feature computation and machine learning. It requires no manual image annotation, applies data-driven parameter tuning and is computationally optimised. *CoMoNoD* maximises the between-class contrast by a compactness criterion, similar to the criterion proposed in²³. Comparable to the Otsu method, a colour threshold is determined. In *CoMoNoD* the threshold is based on spatial colour distributions within the image. The algorithm is implemented in a GPU-optimised way to allow for rapid image processing. Images acquired at 1 Hz can be processed in realtime on one computer. This provides PMN abundance assessments shortly after image acquisition and thus helps scientists to e.g. pin-pointing subsequent sampling locations during research cruises based on the determined spatial nodule statistics.

Material

CoMoNoD was applied to two diverse image sets, $I^{(1)}$ and $I^{(2)}$, showing poly-metallic nodules. Both image sets were acquired in the deep sea of the Pacific Ocean. The images show a vertical view down onto the seafloor (see Fig. 2).

Image set $I^{(1)}$ ($|I^{(1)}| = 34,200$) was acquired by GEOMAR using the *Deep Survey Camera*⁵ on board AUV *Abyss*. Images were acquired in the DISCOL experimental area of the Peru Basin (station SO242-1_083_AUV10²⁷). The imaged area was chosen to reinvestigate a benthic disturbance experiment conducted in 1989²⁸. $I^{(1)}$ is a subset of a much larger AUV image collection acquired during the *Ecological Impacts of Deep Sea Mining* cruises SO239²⁹ and SO242-1²⁷. Image and meta data are available through OSIS (<https://portal.geomar.de>) and the web-based image annotation software DIAS (<https://dias.geomar.de>). The images in $I^{(1)}$ were acquired from an average altitude of 7.5 m above the sea floor. A FishEye lens was used to capture a 90° field of view. FishEye un-distortion was applied to create rectified images of 4096 (3072) pixel width (height). Resolution in the rectified images is ca. 1 px/cm. Those rectified images were then analysed with *CoMoNoD*.

Image set $I^{(2)}$ ($|I^{(2)}| = 88,630$) was acquired by the National Oceanography Centre Southampton (NOCS) in an *Area of Particular Environmental Interest* (APEI No. 6) in the CCZ, using AUV *Autosub6000*³⁰. The images in $I^{(2)}$ were acquired from an average altitude of 3 m. They feature a footprint of ca. 1.8 m² and a resolution of ca. 16 px/cm.

The segmentation of nodules and sediment is considered a binary task for $I^{(1)}$ and $I^{(2)}$. As less than 3% of the image pixels show other objects like fauna, these objects are neglected in the analysis. Both image sets come with meta data for each image including latitude and longitude for geo-referencing as well as altitude to compute image footprints in m².

Method

The proposed *CoMoNoD* method consists of two phases: 1) contrast maximisation (see Figs 3 and 4) and 2) nodule delineation (see Fig. 5). The core heuristic - assuming that PMNs are elliptical, mostly convex objects - is exploited in both phases. All image processing is conducted using the OpenCV C++ library and GPU acceleration is used where applicable.

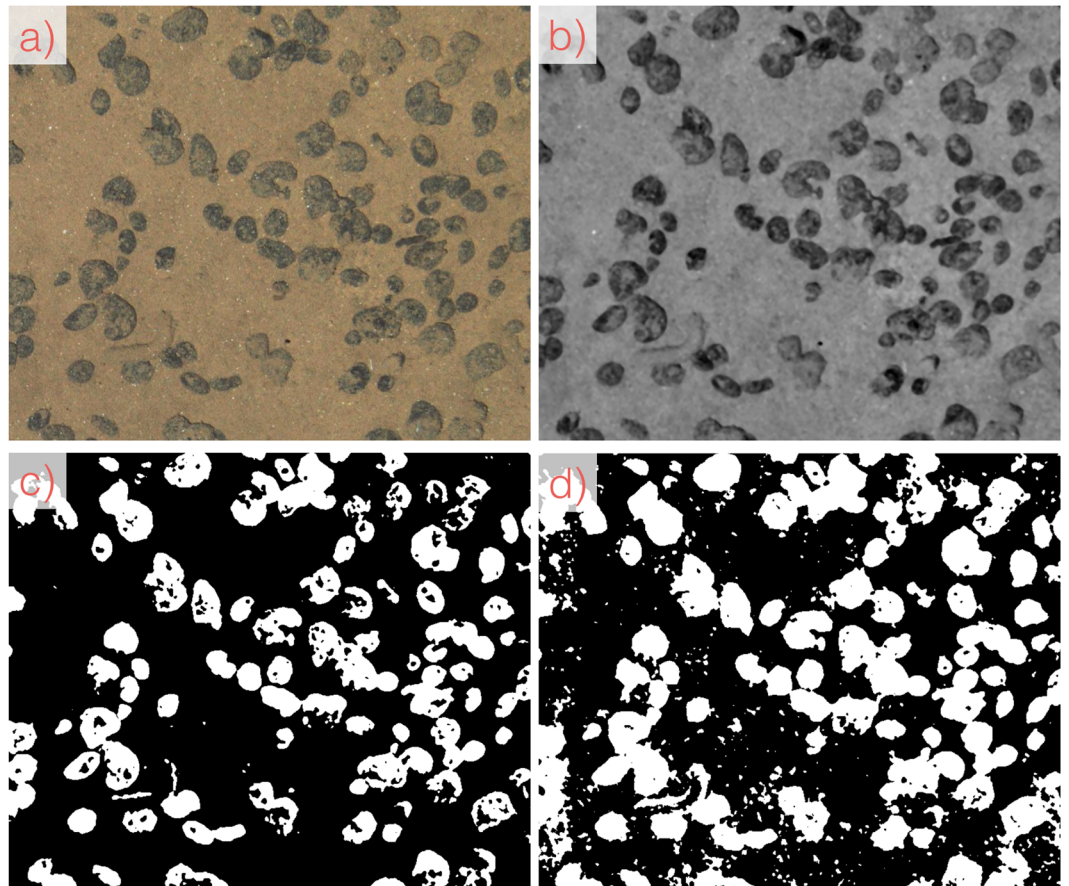


Figure 3. Workflow of image processing steps for contrast maximisation in the first phase of *CoMoNoD*. A zoom-in of the entire image is shown (see Fig. 2). (a) Shows the input image from $I^{(2)}$ (ca. 1/10th of source image). (b) Shows the colour corrected version after applying the *fSpice* method. (c) Shows the contrast-maximised binary image B_i after applying the heuristically tuned threshold t_i . (d) Shows a noisy result created by a threshold t_i that was chosen too large (by erroneously setting θ_γ too high).

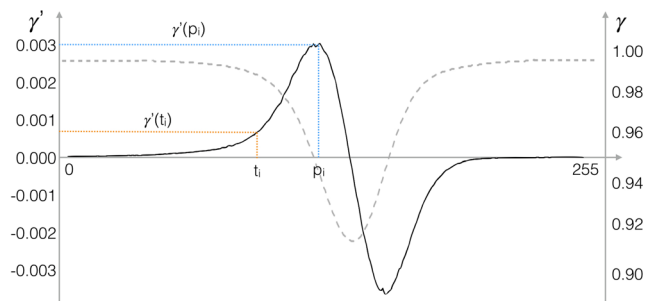


Figure 4. Plot of the compactness curve $\gamma(\cdot)$ (dashed, grey curve) and the first derivative $\gamma'(\cdot)$ (solid, black curve). The curves represent the compactness in one image from $I^{(1)}$. To select the intensity threshold t_i , first the peak p_i is determined in $\gamma'(\cdot)$. Then, t_i is chosen as the highest value $b < p_i$ for which $\gamma'(b) < \gamma'(p_i) * \theta_\gamma$.

Contrast maximisation. Each image I_i of an image set I is first filtered with a 3×3 px median filter to remove sensor noise and small artefacts. Artefacts can be caused by floating particles like marine snow or shell fragments on the seafloor. The images are then scaled with a scale factor $f_i = A^*/A_i$ where A_i is the footprint of image I_i in m^2 . A^* is the median image footprint in the image set I . Cubic interpolation is used while scaling the images. This is the most accurate interpolation method in the GPU-accelerated part of OpenCV. Scaling leads to a uniform px/cm ratio within the scaled image set.

Scaled images are smoothed by a Gaussian filter of 3×3 px size. Afterwards, they are colour corrected using the *fSpice* algorithm (see Figs 3b) and 6(b))¹¹. This data-driven colour correction method removes an illumination

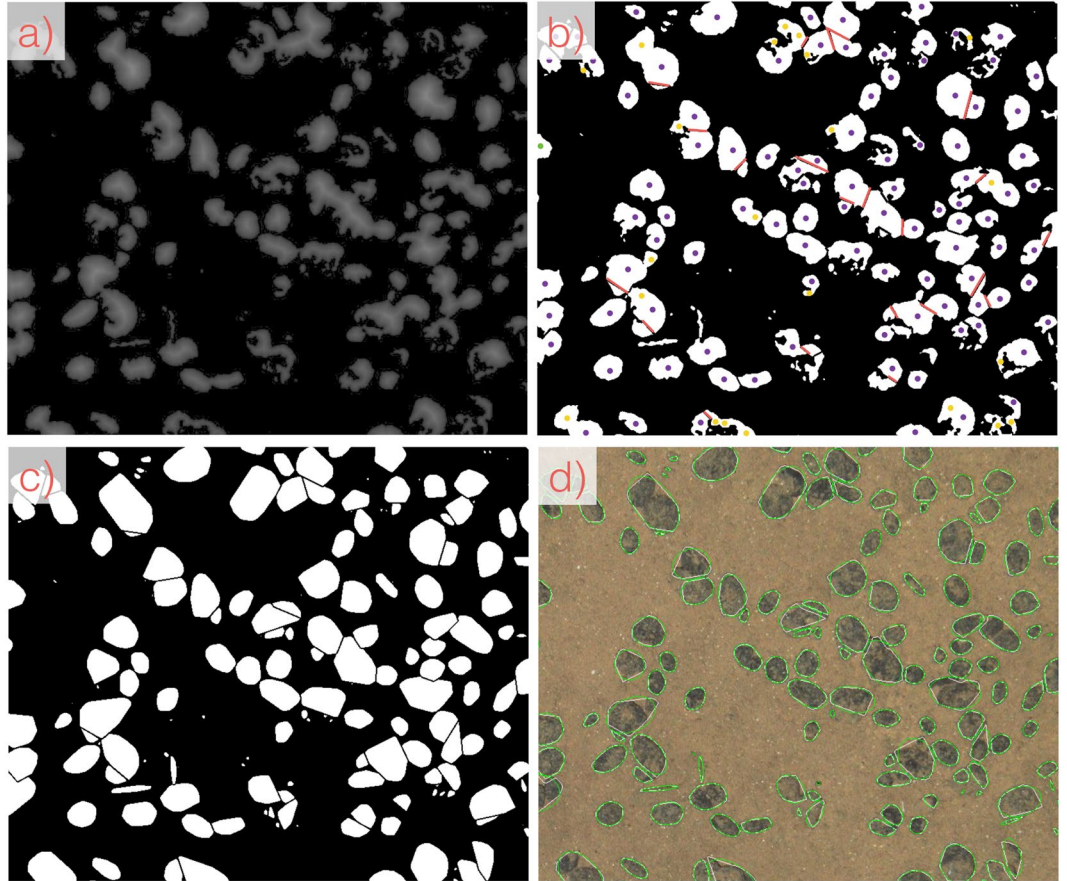


Figure 5. Workflow of the nodule delineation steps in the second phase of *CoMoNoD*. Again the zoom-in section is shown (see Fig. 2). **(a)** Distance image D_i computed from B_i (see Fig. 3c). **(b)** Peaks within the distance image constitute nodule candidate centroids (yellow markers). These are filtered to determine the regional maximum within a neighbourhood (purple markers - all overlaying further yellow markers). Shape bottlenecks between adjacent nodules are used to separate connected nodule candidates (red lines). **(c)** Each nodule candidate blob is delineated by its convex hull. **(d)** Convex hulls from c) are fit by an ellipsoid and shown as an overlay on top of the original image (green nodule delineations). The size of these delineations can be measured in cm^2 and provides the basis for statistical assessments of PMN abundance.

come from individual images and equalises the colour histograms of all images within an image set. The algorithm was developed to reconstruct the natural appearance of the seafloor. Here, *fSpice* is applied in a way which increases the contrast between nodules and the sediment. The colour corrected images are then converted to 8 bit grey-scale.

Each grey-scale image is transformed to a binary image B_i to maximise the contrast. An image-specific threshold t_i is used for binarisation ($t_i \in [0..255]$, see Fig. 3c). To determine t_i , first a grey level co-occurrence matrix $G^{(i)}$ is computed considering the four Moore-neighbouring pixels in 1 px distance. At this point the compactness heuristic is exploited: few colour co-occurrences of pixels below t_i (likely PMNs - i.e. ω_0) and above t_i (likely background - i.e. ω_1) are targeted. Hence for each t_i , a compactness value $\gamma(t_i)$ is computed:

$$\gamma(t_i) = \frac{\sum_{a=0}^t \sum_{b=0}^t G_{a,b}^{(i)} + \sum_{a=t+1}^{255} \sum_{b=t+1}^{255} G_{a,b}^{(i)}}{\sum_{a=0}^{255} \sum_{b=0}^{255} G_{a,b}^{(i)}} \in [0..1] \quad (1)$$

This $\gamma(t_i)$ is maximised for $\gamma(0) = \gamma(255) = 1$.

Next, the first derivative of $\gamma(t_i)$ is computed and its peak position p_i determined:

$$p_i = \underset{a}{\operatorname{argmax}} \gamma'(a). \quad (2)$$

The threshold t_i is selected as:

$$t_i = \underset{b < p_i}{\operatorname{argmax}} \gamma'(b) \quad (3)$$

with:

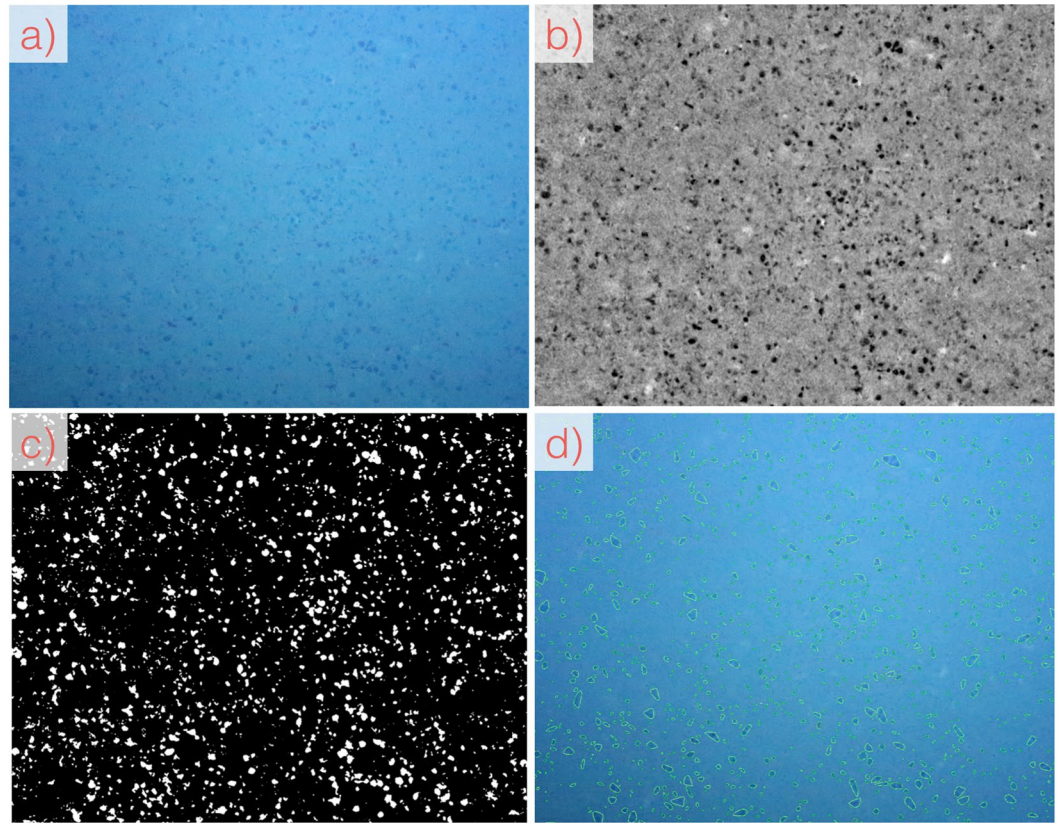


Figure 6. Processing workflow for $I^{(1)}$. Shown are the zoom-ins to the original images (see Fig. 2). (a) Source image, (b) *fSpice* result, (c) binary image, (d) nodule delineation.

$$\gamma'(b) < \gamma'(p_i) \cdot \theta_\gamma \quad (4)$$

where θ_γ is one of the *CoMoNoD* input parameters.

The concept for selecting the threshold t_i like this is as follows. The compactness $\gamma(p_i)$ decreases when the binary image B_i becomes noisy. This is the case when background pixels are erroneously added to the PMN class ω_0 (see Figs 3d) and 4). It will happen when t_i is chosen too high. The maximum compactness change $\gamma'(p_i)$ thus serves as the upper limit for the threshold value t_i .

Nodule delineation. After contrast maximisation, the binary image B_i is subject to multiple blob detection, splitting and fusion steps. These steps delineate individual nodules in the second phase of *CoMoNoD* (see Fig. 5).

First, the distance image D_i is computed, which has the same size as B_i . Each pixel value in D_i encodes the shortest distance to a pixel in B_i assigned to class ω_1 (see Fig. 5a)).

Local maxima are determined in D_i . Maxima are pixels that exceed their Moore-neighbouring pixels. These local maxima constitute the initial nodule candidate centroids. Neighbouring local maxima are filtered and only the highest peak is retained within a $5 * \theta_r$ px radius (see Fig. 5b)). θ_r is another input parameter for *CoMoNoD*. It represents the minimum nodule radius in pixels and depends on image resolution.

Each pixel in B_i , which is set to ω_0 , needs to be assigned to one of these peaks. Here it is again assumed that PMNs are compact objects of elliptical shape. The pixels set to ω_0 form connected pixel clusters or *blobs*.

Bottlenecks in the blob contours are evaluated to separate adjacent PMNs in the binary image (see Fig. 5b) - red lines). Each blob is iteratively split up to virtual blobs to find the optimal separation of the peaks within the blob. The iterative splitting is discontinued when each peak is contained in its own blob (see Fig. 5b)).

All pixel blobs are fused with their largest neighbour when they are smaller than $\pi * \theta_r^2$ pixels. To be fused the neighbour has to be closer than $2 * \theta_r$ pixels distance. Fusion avoids over-segmentation. Small blobs are discarded if no such neighbour exists.

Nodule candidate blobs are delineated by their convex hull to account for gaps between PMNs. Gaps can be caused by sediment coverage or epifauna (see Fig. 5c)). Each convex hull is finally fit with an ellipsoid (see Fig. 5d)). The size of these ellipsoids provides individual nodule sizes in cm^2 from which descriptive nodule statistics can be computed. The main axes of the ellipse also provide measurements of two of the nodule axes.

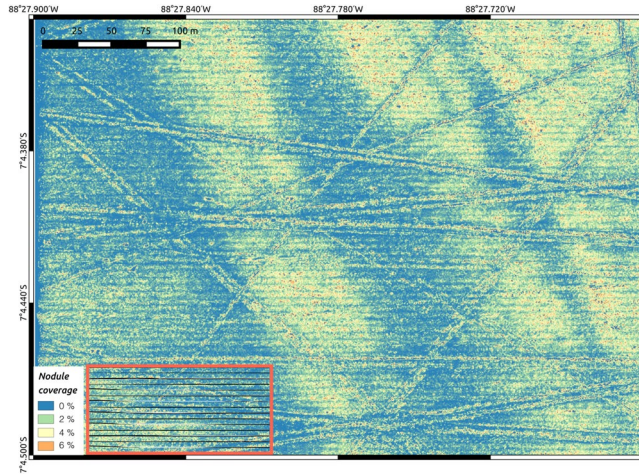


Figure 7. Nodule abundance map showing nodule coverage in percent. 34,200 images contributed to this map. It was computed in about 12 h. The map represents an area of ca. $500 \times 400 \text{ m}^2$ size. It lies in the DISCOL experimental area in the Peru Basin, Pacific Ocean. Narrow linear structures, criss-crossing the area, are anthropogenic plough-marks (8 m wide) from a mining simulation conducted in 1989²⁸. The larger scale (ca. 100 m wide), east-west facing wave pattern correlates with the seafloor micro-bathymetry. The striped horizontal pattern follows the AUV dive trajectory (see black lines in the inset in the lower left corner). Eighty horizontal AUV lines were conducted at a spacing of 5 m. The horizontal pattern is likely caused as an artefact of the max-pooling value selection for overlapping images.

Nodule statistics. The number of PMNs in an image is denoted by N_i . Each PMN nodule $n_{i,j}$, $j = 0, \dots, N_i$ in an image I_i is described by a size value $s_{i,j}$ in cm^2 . Several PMN abundance statistics can be computed from these $s_{i,j}$. The most straightforward are the number of PMNs per square meter ($\Phi_i^N = N_i/A_i \text{ m}^{-2}$) and the percentage coverage of the seafloor with PMNs ($\Phi_i^c = \sum_{j=0}^{N_i} s_{i,j}/A_i$). To represent the nodule size distribution in an image, *CoMoNoD* additionally provides seven characteristic nodule size values:

$$\Phi_{i,\alpha}^s, \alpha \in [0.01, 0.10, 0.25, 0.5, 0.75, 0.90, 0.99] \quad (5)$$

These are computed by first sorting the $s_{i,j}$ by increasing value. Then:

$$\Phi_{i,\alpha}^s = s_{i,k_\alpha} \quad (6)$$

with:

$$k_\alpha = \underset{k}{\operatorname{argmax}} \sum_{l=0}^{k-1} s_{i,l} < \alpha \cdot \Phi_i^c \cdot A_i \quad (7)$$

The interpretation of $\Phi_{i,\alpha}^s$ is that a fraction of α of Φ_i^c in image I_i is owing to nodules smaller than $\Phi_{i,\alpha}^s$ (e.g. $\Phi_{i,0.5}^s$: half of the nodule coverage in I_i is created by nodules smaller than $\Phi_{i,0.5}^s \text{ cm}^2$). The $\Phi_{i,\alpha}^s$ are thus values along the accumulation curve of the nodule size distribution.

Common geological particle size descriptors can be computed from the $\Phi_{i,\alpha}^s$ ³¹. An example are the sorting value: $0.5 \cdot (\Phi_{i,0.75}^s - \Phi_{i,0.25}^s)$ and the skewness: $\Phi_{i,0.75}^s + \Phi_{i,0.25}^s - 2 \cdot \Phi_{i,0.5}^s$ as defined by Trask³².

Results

Figure 6 shows an example of the delineation result for $I^{(1)}$ (in d). It also shows intermediate results of the *fSpice* pre-processing and the binarization step. An example delineation is shown in Fig. 5d) for $I^{(2)}$.

By applying *CoMoNoD* to all 34,200 images in $I^{(1)}$, nodule abundance was assessed within a contiguous seafloor area of ca. $500 \times 400 \text{ m}^2$. Nodule abundance was quantised to a $0.25 \times 0.25 \text{ m}^2$ resolution grid to render the map (see Fig. 7). Therefore, each image was subdivided into $25 \times 25 \text{ cm}^2$ tiles. *CoMoNoD*'s descriptive statistics were then computed for each of these grid tiles.

The average nodule number per square meter for the entire area is $\Phi^N = 2.7 \text{ m}^2$, the average seafloor coverage is $\Phi^c = 1.7\%$ and the median nodule size is $\Phi_{0.5}^s = 48.6 \text{ cm}^2$ (corresponding to a radius of 3.9 cm). Figure 8 shows the nodule size distribution and the coverage distribution.

More than 2.6 million nodules were delineated in the surveyed area. They cover a total area of 8.8 thousand square meters. The nodule volume can be estimated by assuming that the nodules are discoidal rotational ellipsoids. They would form a contiguous manganese nodule cube of 6.6 meter edge length. Assuming a metal composition as reported in the literature (Cobalt: 0.26%, Nickel: 1.2%, Copper: 1%)³³, contained metal amounts can be deduced. The nodule cube would have an estimated worth of ca. 1 million EUR at current market prices (Cobalt:

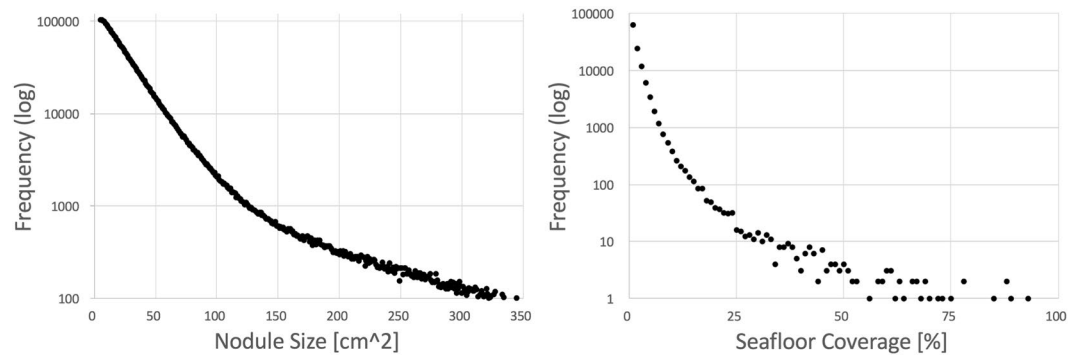


Figure 8. Nodule size distribution (left) and coverage distribution (right) for $I^{(1)}$.

Method	Training [min / MPix]	Training for $I^{(1)}$ [h]	Application [s / MPix]	Application to $I^{(1)}$ [h]
<i>CoMoNoD</i>	—	—	0.1	19
<i>PCCA</i> ²⁴	4.7	3.9	1.9	233
<i>Rapid PCCA</i> ²⁵	4.7	3.9	0.2	25
<i>ES4C</i> ²³	10.3	8.6	1.9	233

Table 1. Runtimes for different nodule detection algorithms. *CoMoNoD* operates without a training phase and has the fastest application time per image. *PCCA* stands for *Pixel-classification by cluster annotation*. This method has been improved by a software company to create *Rapid PCCA*. *ES4C* stands for *Evolutionary tuned Segmentation using Cluster Co-occurrence and a Convexity Criterion*. It first uses the same PR methods as *PCCA* in the training phase. An additional optimisation step is conducted in the training phase of *ES4C*. It allows omitting the manual cluster annotation required in *PCCA*. Training of *PCCA*, *Rapid PCCA* and *ES4C* is conducted on a data subset corresponding to ca. 5 images.

69 kEUR/t, Nickel: 12 kEUR/t, Copper: 7.5 kEUR/t). These values are rough estimates, they do not consider the mining costs and completely neglect the value of the destroyed habitat and fauna.

Coverage values provided by *CoMoNoD* are comparable to results created by two other methods (*PCCA*²⁴ and *ES4C*²³) showing an agreement in the coverage estimates of ca. 92%.

Images in $I^{(2)}$ were acquired along linear transect lines and no contiguous map can be rendered. The nodule statistics for all 88,630 images are $\Phi^N = 111.3 \text{ m}^{-2}$, $\Phi^c = 6.2\%$ and $\Phi_{0.5}^s = 10.4 \text{ cm}^2$ (corresponding to a radius of 1.8 cm).

Further nodule statistics are being computed for the remainder of the image set which $I^{(1)}$ belongs to (cruises SO239 and SO242-1 - see above). Results will be presented in a follow-up publication.

Runtime. *CoMoNoD* was implemented and tested on a high-end desktop computer. It was equipped with a GeForce GTX 980 Ti GPU, Intel Xeon E5-1650 CPU and 64GB RAM. An average runtime of ca. 0.1 s/MPix was observed. Computational complexity of the contrast maximisation is linear in the pixel size. Complexity of the nodule delineation steps is quadratic in the number of nodule segments. Thus images showing more nodules will lead to a longer runtime.

Images in $I^{(1)}$ are ca. 20 MPix in size and the runtime for the entire data set was about 19 hours. Images in $I^{(2)}$ are ca. 5 MPix in size, resulting in a runtime for the entire data set of about 12 hours. Table 1 shows average runtimes for image set $I^{(1)}$.

Source Code. The source code for *CoMoNoD* is available from the GEOMAR OpenSource Git repository (<https://git.geomar.de/open-source/comonod>). Code in this repository is maintained to include future algorithmic updates. A snapshot of the code used for this publication has been archived in Pangaea³⁴. Additional software required to run the algorithm is the OpenCV image processing library (available through <http://opencv.org/>) and the *fSpice* algorithm (<https://git.geomar.de/open-source/oceanvcv>).

Discussion

So far, assessing PMN abundance variations was not possible over multiple spatial scales in near realtime. Measurements performed after physical sampling provide local abundance measurements. In case of box corers the sampled area represents 0.25 m² of the seafloor. Patterns on scales larger than 0.5 m are missed. Ship-based hydro-acoustic data enables assessing nodule patterns at scales of several hundred square kilometres. Its beam-footprint of ca. 50 × 50 m² oversees patterns on smaller scales.

Optical imaging provides a tool to capture PMN abundance variations over several scales. Automated nodule delineation allows to rapidly extract quantitative data from those images. Applied as a joined methodology,

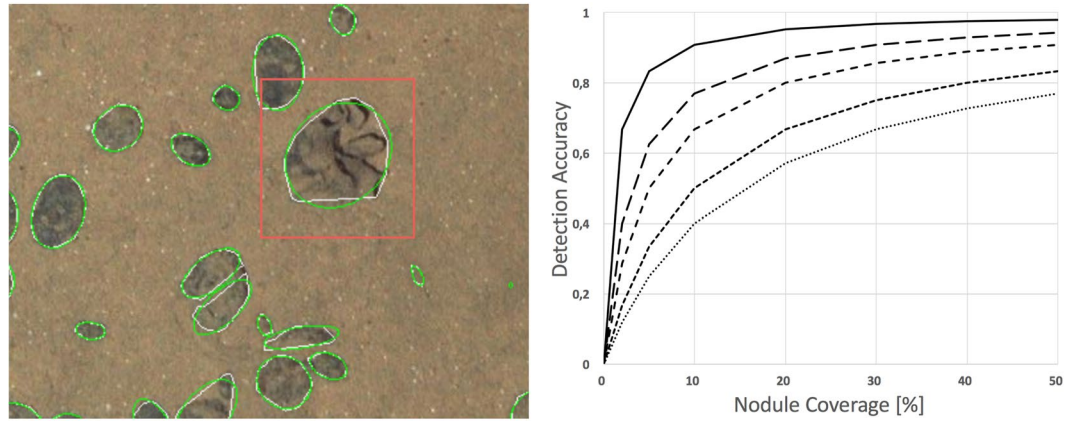


Figure 9. Zoom-in on a *Xenophyophore* that has been misclassified as poly-metallic nodule (left). Detection accuracy per observed nodule coverage (in %) for different fauna coverage values (right). The curves represent different pixel coverage by fauna (solid: 1%, long-dash: 3%, medium-dash: 5%, short-dash: 10%, dotted: 15%). *CoMoNoD* should be accompanied by other detection methods when dark fauna represents more than 10% of the pixels. Bright fauna will be assigned to the sediment class and will not skew the nodule statistics. For $I^{(1)}$ and $I^{(2)}$ dark fauna contributes to less than 3% of the pixels.

imaging and automated analysis can assess PMN abundance and abundance variations from centimetre to kilometre scale.

CoMoNoD is currently applied to image data sets of several hundred thousand photos acquired across the Pacific Ocean. This study will provide nodule abundance statistics for local, regional and ocean basin scale.

Statistically assessing the natural heterogeneity requires to select a sampling size for the seafloor area *quants*. In physical sampling this is restricted to one specific size (e.g. 0.25 m² for box cores). For hydro-acoustics the minimum sampling size depends on the beam opening angle and the distance to the seafloor (meter scale for AUV-based data, 40–100 m for ship-based data). For image-based methods, this area *quant* can be selected. Its size will affect the derived descriptive statistics. When choosing a very small area *quant* (e.g. 10 × 10 cm²) it will be rare to detect nodules. When a nodule does occur within the area, the coverage can reach 100%. Similarly, local nodule abundance heterogeneity might be occluded when the area is chosen too large (e.g. 100 × 100 m²). The summative character of the Φ_i will then average out those variations. A study on assessing nodule abundance heterogeneity on various scales is in preparation and will provide suggestions for sampling size selection.

Abundance maps like Fig. 7 provide a subjective impression of the quality of the nodule delineation algorithm. Ground truth information on nodule abundance would be needed to verify the detection accuracy. As manual annotations are lacking for nodule image and as the quality of manual annotations is disputable, robust verification is currently impossible. Ground truth data from physical sampling could be linked with our image analysis method, yet the natural abundance heterogeneity and the limited accuracy of underwater navigation currently prohibit such an assessment. No data set exists where an image has been taken prior physical sampling to assure a one-to-one comparison of image-derived and ground truth data. Video material of TV-guided box cores is of too low quality to be analysed quantitatively.

The contrast maximisation step provides good binary images in all cases. Sediment coverage creates holes in the third example. The nodule delineation succeeds for the first two examples. Percent coverage could be measured in the latter two only.

To overcome these shortcomings, an imaging survey should be conducted prior to the physical seafloor sampling. This would allow assessment of the visual baseline of the undisturbed seafloor. The physical sampling afterwards would provide precise ground truth measurements of individual nodules. Accuracy of underwater navigation is far from the required accuracy of few tens of centimetres for direct comparison. To overcome ambiguities sampling sites need to be distant enough from each other (>50 m). This will allow identification of each sampling location at the seafloor based on underwater navigational data. A second imaging survey over the same area would then determine the exact positions of the sampling impact sites using visual navigation.

Frequently occurring fauna can become a challenge for *CoMoNoD*. Under such conditions, the assumption that these objects do not contribute to the nodule class does not hold any more. An example of a misinterpretation is shown in Fig. 9. There, a *Xenophyophore* has been erroneously detected as a PMN, contributing an additive error to the coverage and nodule count measures and likely affects Φ_i^s as well. Interestingly it has been shown, that *CoMoNoD* can also be used to detect specific fauna by selecting images with outliers for $\Phi_{i,0.99}^s$. These outliers can be an indicator for large *Holothurians*. However, most complex objects are characterised by more complex visual features. Those objects require pattern recognition systems that are more sophisticated than *CoMoNoD*. Deep sea fauna can be colourful, complex-shaped or textured. Even nodules can present rich morphologies and colour features when imaged at high resolution. In those cases *CoMoNoD* can be challenged as well (see below).

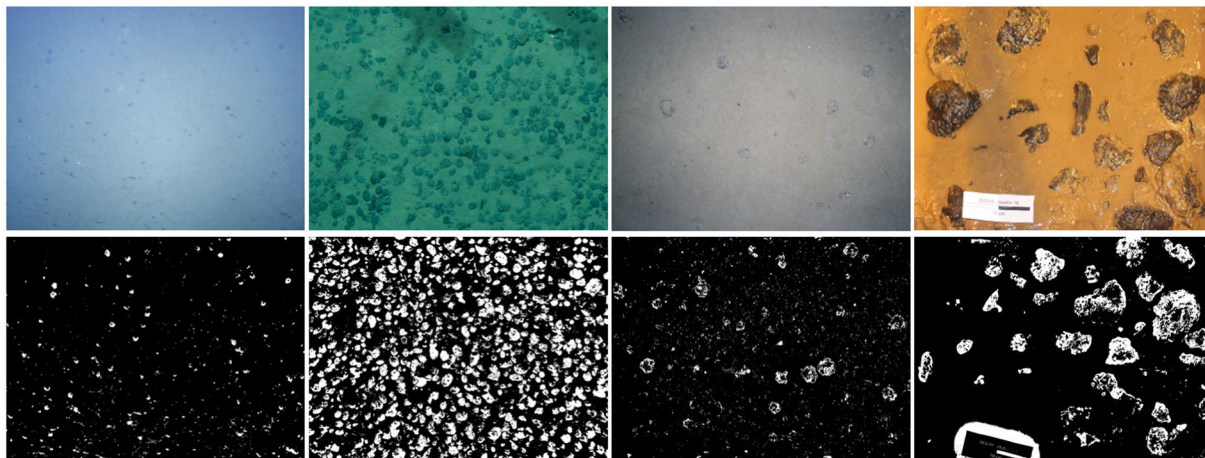


Figure 10. Examples showing *CoMoNoD* results for other imagery. From left to right: image acquired by AUV *Abyss* from lower altitude (resolution 320 Pix/m); image acquired by a towed camera by the Federal Institute for Geosciences and Natural Resources (1890 Pix/m); image acquired by a towed camera by the Alfred-Wegener Institute Helmholtz-Center for Polar and Ocean Research (1900 Pix/m)³⁵; image acquired on deck by an SLR camera (8492 Pix/m).

As binary segmentation and delineating convex objects are also topics elsewhere, *CoMoNoD* could be applied in those contexts as well. Potential fields are microscopy (e.g. for blood cells) and remote sensing (e.g. for permafrost thaw craters).

CoMoNoD was applied to further nodule image data sets (see Fig. 10).

These were obtained by different camera systems mounted on different platforms. Illumination was different in terms of light sources and direction. The data sets are mostly *in-situ* imagery like $I^{(1)}$ and $I^{(2)}$. One data set shows *ex-situ* images of box core samplings acquired on deck of a research vessel. For all of these data sets, appropriate settings for θ_γ provided binary images with maximised contrast between nodules and the sediment background.

For the two image data sets with the highest pixel resolution, and thus highest pixel number per nodule, no appropriate nodule delineation could be achieved. This is likely because the high resolution makes the internal heterogeneity of the nodule itself visible and detectable by *CoMoNoD*. In such cases, *CoMoNoD* could be extended by a fallback system using surrounding pixel's colour information (as described above). This would derive nodule outlines more robustly when the initial delineation does not provide satisfying results. Further studies are needed to determine the limits of the delineation process.

CoMoNoD is governed by the parameter θ_γ . An intelligent parameter guessing system could be implemented that would again use the nodule convexity-heuristic. This system would compute delineations for various settings of θ_γ and θ_r in a brute-force manner. The most promising settings could then be determined by picking the one that produces mostly convex pixel blobs. It would also be possible to analyse multiple images with one parameter setting and assess the resulting nodule statistics as a quality criterion for the chosen parameters.

In the current implementation t_i is estimated for each image. An improved *CoMoNoD* could use the estimate of the previous image to reduce the parameter search space for the current image. When images are acquired in quick succession (e.g. at 1 Hz) it can be assumed that they are rather similar. The nodule distribution, image illumination and seafloor distance should remain comparable and hence t_i should be similar.

The second phase of *CoMoNoD* merges pixel blobs where nodules were cut into multiple segments. It also breaks up pixel blobs where nodules are connected to each other in B_i . These steps can fail when large epifauna occurs on a nodule or when a sediment cover occludes the visibility on the nodules. Both steps are steered by the compactness heuristic and θ_r . Using aerial feature descriptors, further information about pixel neighbourhoods could be included in the nodule fusion and decision process. This additional information would come at an additional computational cost. Further intelligence could be added to *CoMoNoD* by making the initial median filter adaptive, i.e. using the median only when the intensity change exceeds a threshold. These improvements were purposely neglected for *CoMoNoD* to maintain its rapid data analysis capability.

For new users of the algorithm it is easy to explore the influence of different settings for θ_γ and θ_r . A graphical user interface with two sliders would enable users to rapidly create nodule delineations for visual inspection. No feature computation, data normalisation or model training is necessary. *CoMoNoD* can directly be applied to single images. When promising settings have been determined, these can then be applied to the remainder of the data set in an automated way.

Conclusion

The proposed *CoMoNoD* algorithm segments poly-metallic nodules from the sediment background by delineating individual nodules and extracting quantitative abundance data from optical seafloor imagery. It links to the traditional methods for nodule abundance assessment that use physical sampling and hydro-acoustic mapping.

CoMoNoD is governed by only few simple parameters compared to other pattern-recognition-based nodule detection methods. Annotation is not required. Its rapid execution time enables quantitative mapping of nodule distributions during research cruises. CoMoNoD quantifies nodule abundance and abundance variations from square-centimetre to square-kilometre scale.

References

- Kuhn, T., Węgorzewski, A., Rühlemann, C. & Vink, A. Composition, formation, and occurrence of polymetallic nodules. In *Deep-Sea Mining*, 23–63 (Springer, 2017).
- Purser, A. *et al.* Association of deep-sea incirrate octopods with manganese crusts and nodule fields in the Pacific Ocean. *Current Biology* **26**, R1268–R1269 (2016).
- Vanreusel, A., Hilario, A., Ribeiro, P. A., Menot, L. & Arbizu, P. M. Threatened by mining, polymetallic nodules are required to preserve abyssal epifauna. *Scientific Reports* **6** (2016).
- Durden, J. *et al.* Perspectives in visual imaging for marine biology and ecology: from acquisition to understanding. *Oceanography and Marine Biology: An Annual Review* **54**, 1–72 (2016).
- Kwasnitschka, T. *et al.* DeepSurveyCam - A Deep Ocean Optical Mapping System. *Sensors* **16**, 164 (2016).
- Ontrup, J., Ehnert, N., Bergmann, M. & Nattkemper, T. W. BIIGLE-Web 2.0 enabled labelling and exploring of images from the Arctic deep-sea observatory HAUSGARTEN. In *OCEANS 2009-EUROPE*, 1–7 (IEEE, 2009).
- Langenkämper, D., Zurowietz, M., Schoening, T. & Nattkemper, T. W. BIIGLE 2.0-Browsing and Annotating Large Marine Image Collections. *Frontiers in Marine Science* **4**, 83 (2017).
- Sharma, R. Quantitative estimation of seafloor features from photographs and their application to nodule mining. *Marine Georesources & Geotechnology* **11**, 311–331 (1993).
- Schoening, T. & Nattkemper, T. W. Seabed classification using a novel bag-of-prototypes feature representation. In *Proceedings of the ICPR 2014* (2014).
- Culverhouse, P. F., Williams, R., Reguera, B., Herry, V. & González-Gil, S. Do experts make mistakes? A comparison of human and machine identification of dinoflagellates. *Marine Ecology Progress Series* **247**, 17–25 (2003).
- Schoening, T. *et al.* Semi-automated image analysis for the assessment of megafaunal densities at the Arctic deep-sea observatory HAUSGARTEN. *PLoS one* **7**, e38179 (2012).
- Durden, J. M. *et al.* Comparison of image annotation data generated by multiple investigators for benthic ecology. *Marine Ecology Progress Series* (2016).
- Duncan, J. S. & Ayache, N. Medical image analysis: Progress over two decades and the challenges ahead. *IEEE transactions on pattern analysis and machine intelligence* **22**, 85–106 (2000).
- Hoth, J. & Kowalczyk, W. Colour reconstruction of underwater images. *Underwater Technology* **34**, 54–61 (2017).
- Siddiqui, S. A. *et al.* Automatic fish species classification in underwater videos: exploiting pre-trained deep neural network models to compensate for limited labelled data. *ICES Journal of Marine Science* (2017).
- Navarro, A. *et al.* Imafish_ml: A fully-automated image analysis software for assessing fish morphometric traits on gilthead seabream (*Sparus aurata* L.), meagre (*Argyrosomus regius*) and red porgy (*Pagrus pagrus*). *Computers and Electronics in Agriculture* **121**, 66–73 (2016).
- Mazurkiewicz, M., Górska, B., Jankowska, E. & Włodarska-Kowalczyk, M. Assessment of nematode biomass in marine sediments: A semi-automated image analysis method. *Limnology and Oceanography: Methods* **14**, 816–827 (2016).
- Dawkins, M., Stewart, C., Gallager, S. & York, A. Automatic scallop detection in benthic environments. In *Applications of Computer Vision (WACV), 2013 IEEE Workshop on*, 160–167 (IEEE, 2013).
- Gonzalez, R. C., Woods, R. E. & Eddins, S. L. *Digital image processing using MATLAB*, vol. 2 (Gatesmark Publishing Knoxville, 2009).
- Antonisse, H. J. Image segmentation in pyramids. *Computer Graphics and Image Processing* **19**, 367–383 (1982).
- Shi, J. & Malik, J. Normalized cuts and image segmentation. *Pattern Analysis and Machine Intelligence, IEEE Transactions on* **22**, 888–905 (2000).
- Comaniciu, D. & Meer, P. Mean shift: A robust approach toward feature space analysis. *Pattern Analysis and Machine Intelligence, IEEE Transactions on* **24**, 603–619 (2002).
- Schoening, T., Kuhn, T., Jones, D. O., Simon-Lledo, E. & Nattkemper, T. W. Fully automated image segmentation for benthic resource assessment of poly-metallic nodules. *Methods in Oceanography* (2016).
- Schoening, T. *Automated detection in benthic images for megafauna classification and marine resource exploration* (2015).
- Schoening, T., Langenkämper, D., Steinbrink, B., Brün, D. & Nattkemper, T. W. Rapid image processing and classification in underwater exploration using advanced high performance computing. In *IEEE OCEANS* (2015).
- Kuhn, T., Rühlemann, C. & Wiedicke-Hombach, M. Developing a strategy for the exploration of vast seafloor areas for prospective manganese nodule fields. In *Proc. of the 41st Conference of the Underwater Mining Institute (UMI)*, **9** (2012).
- Greinert, J. RV SONNE Cruise Report SO242-1: JPI OCEANS Ecological Aspects of Deep-Sea Mining, DISCOL Revisited, Guayaquil-Guayaquil (Equador), 28.07.–25.08. 2015. *GEOMAR Report* **26** (2015).
- Bluhm, H. Re-establishment of an abyssal megabenthic community after experimental physical disturbance of the seafloor. *Deep Sea Research Part II: Topical Studies in Oceanography* **48**, 3841–3868 (2001).
- Martnez Arbizu, P. & Haeckel, M. RV SONNE Fahrtbericht/Cruise Report SO239: EcoResponse Assessing the Ecology, Connectivity and Resilience of Polymetallic Nodule Field Systems, Balboa (Panama)–Manzanillo (Mexico), 11.03.–30.04. 2015. *GEOMAR Report* **25** (2015).
- Jones, D. *et al.* RRS James Cook Cruise JC120 15 Apr–19 May 2015. Manzanillo to Manzanillo, Mexico. Managing Impacts of Deep-sea resource exploitation (MIDAS): Clarion-Clipperton Zone North Eastern Area of Particular Environmental Interest. *Cruise Reports* (2015).
- Füchtbauer, H. *Sedimente und Sedimentgesteine* (Schweizerbart'sche Verlagsbuchhandlung, 2009).
- Trask, P. D., Hammar, H. E. & Wu, Z. *Origin and environment of source sediments of petroleum* (Gulf Publishing Company, 1932).
- Halbach, P., Friedrich, G. & Von Stackelberg, U. *The Manganese Nodule Belt of the Pacific Ocean*, vol. 254 (1988).
- Schoening, T. Source code for the Compact Morphology-based Nodule Delineation (CoMoNoD) algorithm (2017).
- Purser, A. *et al.* Photographs of incirrate octopods observed during SONNE cruise SO242/2 to the DISCOL Experimental Area, Peru Basin (PANGAEA, 2016).

Acknowledgements

We thank the crews of SO239, SO242-1 and JC120 as well as the AUV teams of NOCS and GEOMAR for their support in acquiring the images used in this article. We thank the data management team at GEOMAR for their support in managing the large image sets and making them available online. This manuscript was much improved by the comments of two anonymous reviewers.

Author Contributions

T.S. developed and implemented the *CoMoNoD* algorithm. D.J. and J.G. provided image data sets and evaluated algorithmic results. All authors contributed to the article.

Additional Information

Competing Interests: The authors declare that they have no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2017