Supplementary Material

SM1. Post-Processing of Images for Automated Classification

Imagery was collected without artificial light and using a fisheye lens to maximise light capture, therefore each image needed to be processed prior annotation in order to balance colour and to minimise the non-linear distortion introduced by the fisheye lens (Figure S1). Initially, colour balance and lenses distortion correction were manually applied on the raw images using Photoshop (Adobe Systems, California, USA). However, in order to optimize the manual post-processing time of thousands of images, more recent images from the Indian Ocean and Pacific Ocean were post-processed using compressed images (jpeg format) and an automatic batch processing in Photoshop and ImageMagick, the latter an open-source software for image processing (www.imagemagick.org). In view of this, the performance of the automated image annotation on images without colour balance was contrasted against images colour balanced using manual post-processing (on raw images) and the automatic batch processing (on jpeg images). For this evaluation, the error metric described in the main text (Materials and Methods) was applied to the images from following regions: the Maldives and the Great Barrier Reef (Figures S2 and S3). We found that the colour balance applied regardless the type of processing (manual vs automatic) had an important beneficial effect on the performance of the automated image annotation as errors were reduced for critical labels in both regions (e.g., Algae labels; Figures S2 and S3). Importantly, no major differences in the performance of the automated annotations were observed between manual and automated adjustments for colour balance.

Once adjusted for colour and lens distortion, images were cropped to 1 m^2 in order to standardise the observation size (Figures S1d and S1e). Since altitude from the camera to the reef substrate can vary do to reef topography, the size of the image will invariably change. To control for this variability and standardise the sampling size, images were cropped to 1 m^2 using the altitude data logged by the Doppler transponder[1].

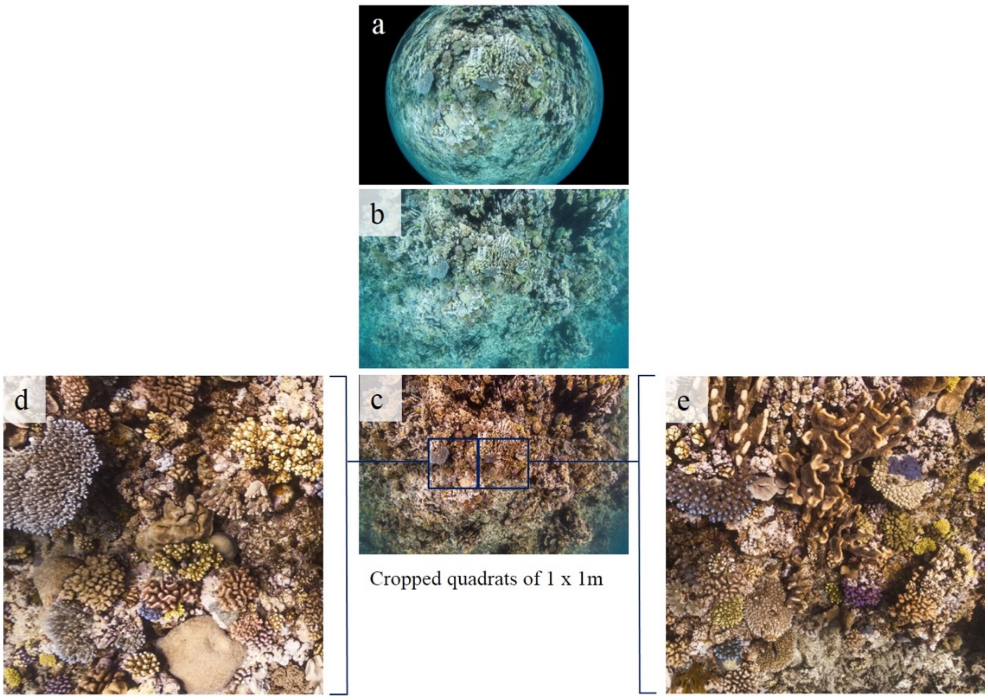

Figure S1. Post-processing of images for automated classification: (a) Raw image. (b) Lenses distortion corrected image. (c) Colour balanced image. (d and e) Cropped quadrat of 1 x 1 m.

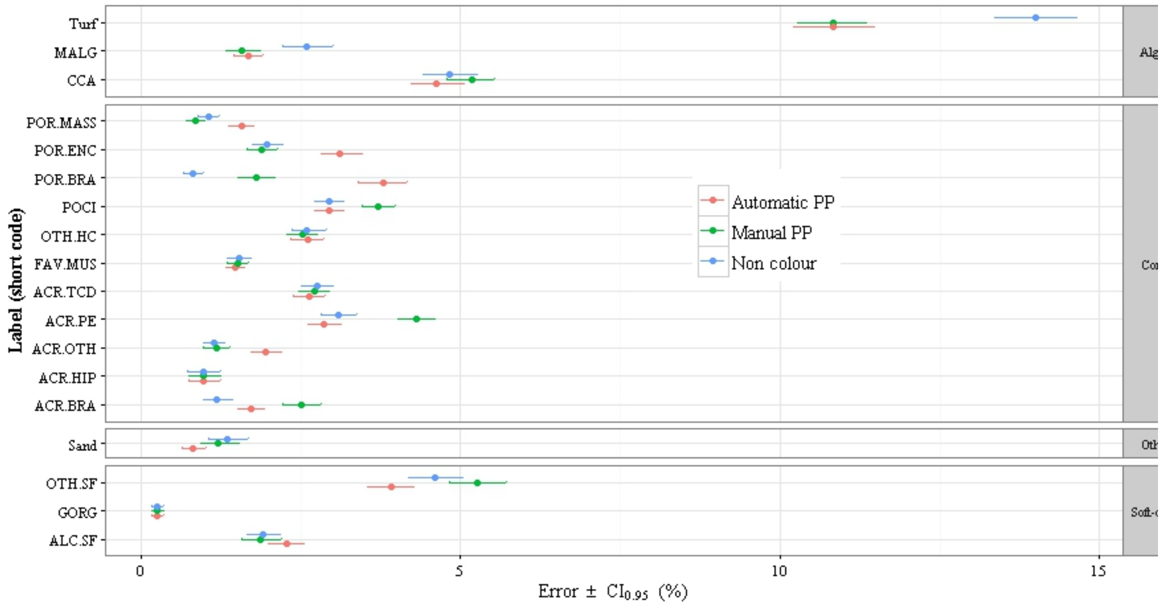

Figure S2. Comparison between post-processing methods of precision errors of automated image annotation for each label within the Maldives, Central Indian Ocean. Automatic Post-processing=Automatic PP; Manual Post-processing=Manual PP; Original colour = Non-colour. Error bars represent the 95% confidence interval.

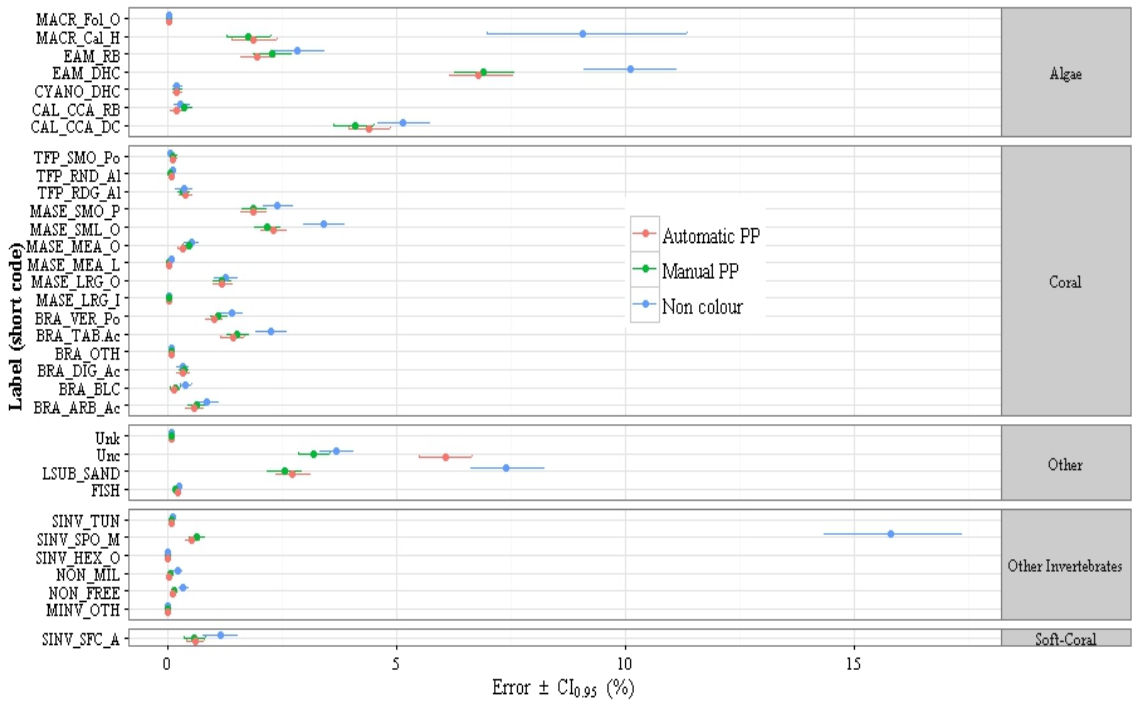

Figure S3. Comparison between post-processing methods of precision errors of automated image annotation for each of label within the Great Barrier Reef. Automatic Post-processing=Automatic PP; Manual Post-processing=Manual PP; Original colour = Non-colour. Error bars represent the 95% confidence interval.

SM2. Deep Learning Network design

In this study, we used the VGG-D 16 network architecture, hereafter denoted VGG. The VGG net is a convolutional neural network (CNN) with over 138 million parameters organised in 16 layers which directly maps a 224 x 224 RGB input image, x, to a score vector $s \in \Delta^k \subset R^k$, representing the posterior probabilities that the input belongs to each of a set of K classes. Δ^k here represents the K simplex:

$$\sum_{k=1}^{K} s(k) = 1, s(k) > 0 \tag{1}$$

$$f^{(\theta)}: N^{224x224x3} \rightarrow R^k \tag{2}$$

where θ is the network parameterisation. These network parameters define the weights of each connection in the neural network. Classification is given by predicting the class pertaining to the highest score:

$$y^{CNN}(x) = argmax f^{\theta}(x) = argmax s \tag{3}$$

The net is trained using back-propagation to minimize classification loss on a giving set of images and labels [2]. We will refer to a “sample", as a pair of image x and target label y, and a “training-data" as a set of samples used to learn the parameters of the network. Given a single sample (x, y), the loss is given by the cross-entropy:

$$E = -log(s_y) \tag{4}$$

where $s \in \Delta^k$, as previously, is given by $f^{\theta}(x)$. The cross-entropy loss is intuitive: to minimize the loss, the entry of the score vector which corresponds to the correct class, y should be as large as possible. In practice, optimization is performed across batches of m samples $(x_1, y_1), \ldots, (x_m, y_m)$, so that m images are simultaneously forward-propagated through the network, and the loss calculated as:

$$E = \frac{-1}{m} \sum_{i=1}^{m} -log\left(s_{y_i}\right) \tag{5}$$

where s_{y_i} represents the score given to ground truth class y_i of image x_i.

Each region has a unique taxonomic composition represented by the label set (SM3). All networks, however, where initialised with ImageNet dataset and then fine-tuned with the manually annotated training images for each country or region (Table 1, main text). Similarly, parameters of learning rate and image scale in the network were independently optimised for each network. A total of 5,255 images, randomly selected from the project's open-access repository (https://espace.library.uq.edu.au/view/UQ:734799), were manually annotated for training of the deep learning networks. The manual annotation was conducted on CoralNet, an online platform designed for image analysis (www.coralnet.ucsd.edu).

Network Fine-Tuning

Traditionally, deep learning algorithms for computer vision require large datasets for training [3]. The most popular benchmark dataset, ImageNet, contains one million images from one thousand categories. But in practice, training dataset are typically smaller and training a neural network's weights from scratch, using smaller datasets and starting from random initialized parameters, would largely over-fit the training set [3].

One approach to get around the aforementioned problem is to first pre-train a deep net on a large-scale dataset, like ImageNet. Then, given a new dataset, the network can be initialised with these pre-trained weights when training on our new task [3]. This process is commonly called *fine-tuning* [3] and, there are a number of variations of fine-tuning. For example, the initial neural network can be used only as a *feature extractor*. That means that every or some layers, prior to the output layer, are frozen and the network simply learn a new output layer. Another approach is to update all of the network's weights for the new task, and that's the approach used in this study.

In this study, a VGG network [4], using 16 weight layers and pre-trained on ImageNet after 370K iterations (74 epochs), was initialised and fine-tuned using the training data. To fine-tune this network, the last fully-connected layer, containing one thousand classes from ImageNet, was replaced by a new one that outputs the desired number of classes (Table 1, main text). Pre-trained weights were initialised using a Gaussian sampling, where weights were drawn from a normal distribution, and adjusted by backpropagation, during the training process, as described above. This fine-tuning exercise ran for 40 K iterations, where the final classification was contrasted against a validation set of images, an independent 20% subset from the original set of images. During each iteration, all training images are forward-propagated to the network to calculate the cross-entropy loss and the weights are adjusted by back-propagation (eq. 5). Overall, the intention of the training exercise is an optimisation problem aimed at finding the best combination of model parameters or weights that minimise the cross-entropy loss. In other terms, the loss function can be defined as the goodness-of-fit of the model to the validation data and by optimising the minimal loss the network best represents the classification data. While less granular in the definition of goodness of fit, the overall accuracy of the classification also defines the fit of the model to the validation data and proportionally increases as the cross-entropy loss decreases (Figure S4). Finding a sample minima for the network loss and the maxima of the accuracy, ensure that the classification outputs from the best network configuration is not achieved by chance.

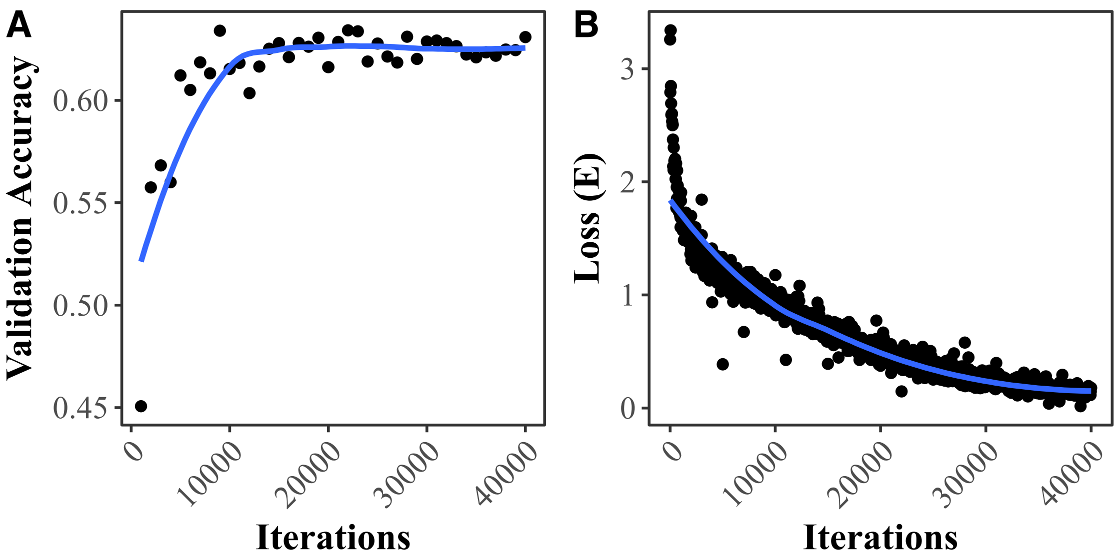

Figure S4. Fine-tuning of deep learning networks. Fine-tuning iterations are aimed at maximising accuracy (A) and minimising cross –entropy loss (B) of network classification of validation samples (subset of images and annotations). This figures is taken an example from the fine-tuning results during the training of the deep learning network for the region Eastern Australia.

For each network, two more hyper-parameters were also calibrated: the learning rate and the representation of the receptive field. Learning rate is a hyper-parameter that control the degree or steps at which the network weights are adjusted with respect the loss gradient (E, eq. 4). Receptive field is the region in the input space that a particular CNN's feature is looking at and activated by (e.g., filter size). Typically, higher resolution in the receptive fields will lead to increases in the accuracy of the classification, at a cost in the computational time. However, the receptive fields in VGG are fixed, and while this, admittedly, should be optimised to maximise the classification performance, the chosen architecture does not allow altering the receptive field resolution. As discussed in the main text, the accuracy of smaller and patchy organisms (e.g., turf algae) in the benthic imagery tend to be penalised by the size of the receptive field while larger receptive field may increase the accuracy of larger organisms (e.g., hard corals). Ideally, a multi-scale architecture would be better suited for this problem, but further development work is required. One initial step towards solve this problem, implemented here, is to up-scale the image size prior cropping the patches that will serve as training samples for the network (see below classification of random point annotations). Increasing the size of the images (total number of pixels) by a factor, reduces the he proportional representation of the cropped patches within the image by the same factor. Thus, balancing the per-class classification accuracy between larger objects (e.g., hard corals) and smaller or patchy organisms (e.g., turf algae, Figure S5). Details of the model parametrisation for each region is provided below (Table S1).

To optimise the learning rate and representation of the receptive field in each network, a suite of experiments was designed to train the same network multiple times using different values of learning rates and image size, simultaneously (i.e., grid search). Learning rate was evaluated using a range of values starting from 0.01, as the initial rate described for the pre-trained VGG network [4], and decreasing by a factor of 10 (e.g., 0.01, 0.001, 0.0001). The image size was adjusted by two different methods: upscaling by a scaling factor (Scale) or by defining the pixel/cm ratio (Ratio). The Scale method increased the total number of pixels in an image by a multiplying factor (e.g., 1.0, 2.0, 3.0), while Ratio set the image spatial resolution (e.g., 11, 22, 33). Both methods yielded the same results only varying the input variable (scaling factor or desired spatial resolution) and the resulting image is resized by bilinear interpolation. The experiment produced a number of networks, using the same number of interactions as a described above, that were contrasted against the absolute error per label to select the best learning rate and image size for each country/region.

Table S1. Configuration parameters for each network after fin-tuning weights and calibrating the learning rate and receptive field. ID = Network identifier, Region and country define the location where images were collected, Training images is the total number of images collected, training samples is the total number of cropped patches from 100 random points per image used for training the network, validation samples is the random subset of samples (20%) used for fine-tuning the network weights, *K* is the total number of classes used to replace the final layer of the pre-trained network, method and factor describe the approach to explore the amount of information contained in the receptive field (see text), Learn. rate is the final learning rate selected and S.I. is the iteration number at which the network achieved the loss minimum.

ID	Region	Country	Training images	Training samples	Validation samples	*K*	Method	Factor	Learn. rate	S.I
1	Central Pacific Ocean	Hawaii	501	40,080	10,020	21	scale	1	0.001	40,000
2	Central Indian Ocean	Chagos Archipelago	359	28,720	7,180	33	scale	2	0.001	40,000
3		Maldives	1,171	93,680	23,420		scale	2	0.0001	30,000
4	Western Atlantic	All	449	35,920	8,980	38	scale	1	0.001	37,000
5	Eastern Australia	Australia	1,234	98,720	24,680	22	scale	2	0.001	40,000
6	South East Asia	Indonesia and Philippines	752	60,160	15,040	35	scale	2	0.0001	30,000
7		Timor Leste	551	44,080	11,020		ratio	30	0.001	30,000
8		Solomon Islands	439	35,120	8,780		scale	1	0.0001	40,000
9		Taiwan	350	28,000	7,000		scale	1	0.001	40,000

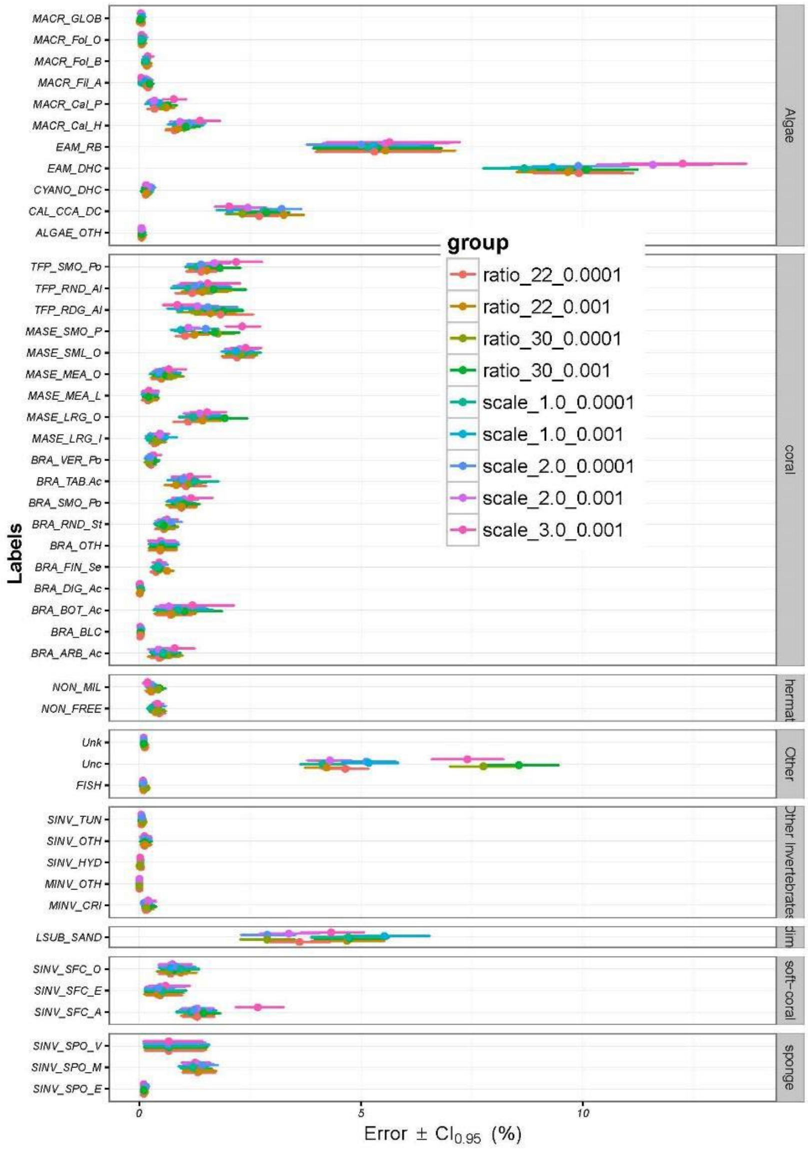

Figure S5. Absolute Error for the abundance estimation of benthic classes (Labels) by different net configurations. Network configurations correspond to the parametrization of learning rate and image size (proportional representation of the receptive field, see main text). This figure is an example of the experiments designed to find the best learning rate and image size that represents the labels with the Maldives dataset. A group of networks were produce with different combinations of parameters, labelled as "method_factor_lerning rate" in the legend of this figure. Colour dots represent the mean error while the error bars represent the 95% confidence interval for each group of experiments. In this figure, EAM_DHC (Turf Algae), Unc (Unclear substrate), LSUB_SAND (Sand) and SINV_SFC_A (Soft Corals, family Alcyoniidae) where the labels where the image size and learning rate had the most significant impact.

Classification of Random Point Annotations

In this study, the relative cover of benthic groups or labels was estimated using 50 points per image. Each cropped patch was then classified independently and then the relative abundance for each of the benthic classifications was the ratio between the numbers of patches classified for a given class by the total number of patches evaluated in an image (Figure S6).

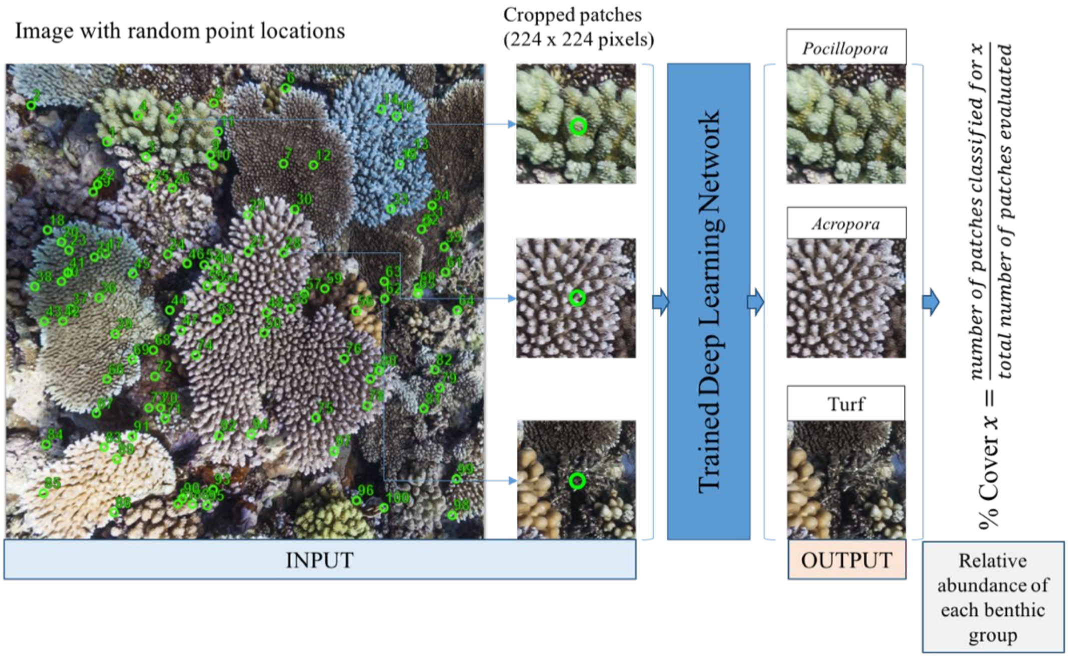

Figure S6. Diagram to visualise the approach to automatically estimate the relative abundance of benthic groups using a sample image. From random point locations, an area cropped around each point (patch) was automatically classified and the relative abundance was calculated by the proportional number of classes.

Data Augmentation

As afore mentioned, deep learning algorithms often require large training datasets. While fine-tunning a pre-trained network is the preferred method to deal with small datasets, data augmentation can help. Data augmentation is a technique to synthetically increase the dataset by manipulating the images, in our case the patches. Minor alterations to the existing dataset, such as flips or translations or rotations will increase the number of images and the network will identify each manipulated image as distinct sample.

Furthermore, in the real-world scenario, a dataset of images may be taken in a limited set of conditions. But, the images expected to be classified may exist in a variety of conditions, such as different orientation, location, scale, brightness etc. A convolutional neural network that can robustly classify objects, even if it's placed in different orientations, is said to have the property called invariance. More specifically, a CNN can be invariant to translation, viewpoint, size or illumination (or a combination of the above). Here we trained the neural networks with additional synthetically modified data by rotating in 90 degrees each patch image, doubling the size of the training set.

SM3. Label-Set Designed for Automated Classification of Coral Reef Benthos in Each Region

Table S2. Label set defined for automatically identifying coral reef benthos per region. Group refers to major functional group of benthic classes. “Name” represents the final classes for the classification process, descried in the column “description”. “Attribute” refers to the taxonomical or morphological attribute that defines each name.

Group	Name	Attribute	Description
Region: Western Atlantic			
Algae	Cyanobacteria	Morphology	Filamentous cyanobacteria
	Epilithic Algal Matrix		Multi-specific assemblages of filamentous algae and CCA smothering reef surface 1 cm or less in height
	Macroalgae		Upright macroalgae > 1 cm in height (all genera and species)
Hard corals	*A. cervicornis*	Species	*Acropora cervicornis*
	A. palmata		*Acropora palmata*
	C. natans		*Colpophyllia natans*
	D. labyrinthiformis		*Diploria labyrinthiformis*
	E. fastigiata		*Eusmilia fastigiata*
	M. cavernosa		*Montastraea cavernosa*
	Madracis spp	Genera	*Madracis sp*
	Meandrina spp		*Meandrina sp*
	Meandroid Faviidae and Mussidae	Family and Morphology	Massive meandroid corals, including: *Dendrogyra, Isophyllia, Manicina, Mycetophyllia, Musa, Scolymia*
	Orbicella spp	Genera	*Orbicella* complex: *O. Annularis, O. faveolata* and *O. franksi*
	P. astreoides	Species	*Porites astreoides*
	P. porites		*Porites porites*
	Plates/encrusting Agariciidae	Family and Morphology	Agariciidae. Plates and encrusting corals of the genera *Agaricia, Undaria* and *Helioseris*
	Pseudodiploria spp	Genera	*Pseudodiploria strigosa* and *P. clivosa*
	S. siderea	Species	*Siderastrea siderea*
	Sub-massive hard coral	Morphology	Sub-massive corals, including: *Dichocoenia, Favia, Solenastrea* and *Stephanocoenia*
Other	Other	Broad functional groups	Other organisms, materials and substrate nor reported in other labels. Includes: fish, trash, transect hardware
	Sand		Sand. Unconsolidated reef sediment
	Seagrass		Seagrass
	Sediment other		Terrigenous sediments
Other Invertebrates	*A. fistularis*	Species	*Aplysina fistularis*
	C. vaginalis		*Callyspongia vaginalis*
	Cliona spp	Genera	*Cliona viridis* complex
	Encrusting sponge	Morphology	Less than 5 cm height (for whole individual); e.g. *S. ruetzleri, Haliclona sp, Monanchora sp, Chondrilla sp, Clathria sp, C. varians*
	Erect Sponge		Upright narrow tubes or branches e.g. *A. compressa*. >1 cm height; height >>basal area
	Ircinia spp	Genera	*Ircinia sp*
	Massive sponge	Morphology	Large basal area to body size and height > 5 cm, compact shape, irregular shapes. Include: chain-tubes spherical castle; e.g. *E. ferox* and *Verongula sp.*
	Millepora spp	Genera	*Millepora sp*
	Mobile invertebrates	Broad functional groups	Mobile invertebrates: e.g., Echinoderms, Lobsters
	Other sponge		Other sponges
	Rope sponge	Morphology	Height >> basal area, spread along the substrate. e.g. *S. aura*
	Sessile invertebrates	Broad functional groups	Other sesile invertebrates. Includes: bryozoa, molluscs, ascidians, soft hexacorrallia and hydroids
	Tube sponge	Morphology	Tube species: *A. archeri, A. tubulata, C. plicifera, N. digitalis* and unknown vase and small barrel group (tube opening ~ height; basal area < opening)
	X. muta	Species	*Xestospongia muta*
Soft Corals	Other soft corals	Broad functional groups	Other soft corals
	Sea fans, plumes and whips		Sea fans, whips and plumes
Region: Australia			
Algae	Crustose Coralline Algae	Morphology	Crustose coralline algae
	Epilithic Algal Matrix	Morphology	Multi-specific assemblages of filamentous algae and CCA smothering reef surface 1 cm or less in height
	Macroalgae	Morphology	Upright macroalgae > 1 cm in height (all genera and species)
Hard Corals	Branching Acroporidae	Family and Morphology	Family Acroporidae, branching morphology (excluding hispidose morphology)
	Branching Poritidae		Family Poritidae, branching morphology e.g. *Porites cylindrica*
	Encrusting Acroporidae		Family Acroporidae, plate and encrusting morphologies
	Encrusting Poritidae		Family Poritidae, encrusting morphologies e.g. *Porites lichen*
	Hispidose Acroporidae		Family Acroporidae, hispidose morphology

	Massive Poritidae		Family Poritidae, massive and sub-massive morphologies, e.g. *Porites lobata; P. lutea*
	Meandroid Faviidae and Mussidae		Families Faviidae and Mussidae, massive and meandroid morphologies
	Other Acroporidae		Other corals from the family Acroporidae (e.g. *Isopora spp*)
	Other hard corals	Broad functional groups	Other hard corals
	Pocilloporidae	Family	Family Pocilloporidae, all genera and species
	Table Acroporidae	Family and Morphology	Family Acroporidae, table, corymbose and digitate morphologies
Other	Other	Broad functional groups	Other organisms, materials and substrate nor reported in other labels. Includes: fish, trash, transect hardware
	Sand		Sand. Unconsolidated reef sediment
Other Invertebrates	Mobile invertebrates	Morphology	Mobile invertebrates: e.g., echinoderms, lobsters
	Sessile invertebrates	Broad functional groups	Other sesile invertebrates. Includes: bryozoa, molluscs, ascidians, soft hexacorrallia and hydroids
Soft Coral	Digitate Alcyoniidae	Family and Morphology	Family Alcyoniidae, digitate morphology. Common genera *Lobophytum, Sarcophyton* and *Sinularia*
	Other soft corals	Broad functional groups	Other soft corals
	Sea fans, plumes and whips		Sea fans, whips and plumes
Region: Central Indian Ocean			
Algae	Crustose Coralline Algae	Morphology	Crustose coralline algae
	Cyanobacteria		Filamentous cyanobacteria
	Epilithic Algal Matrix		Multi-specific assemblages of filamentous algae and CCA smothering reef surface 1 cm or less in height
	Macroalgae		Upright macroalgae > 1 cm in height (all genera and species)
Hard corals	Branching Acroporidae	Family and Morphology	Family Acroporidae; branching morphology (including hispidose morphology)
	Branching hard corals		Non-Acropora branching genera including: *Seriatopora, Anacropora, Echinopora, Montipora, Tubastrea* (excluding *Pocillopora* and *Stylophora*)
	Branching Poritidae		Family Poritidae, branching morphology e.g. P. cylindrica
	Encrusting Poritidae		Family Poritidae, encrusting morphologies e.g. P. lichen
	Foliose corals	Morphology	Thin Foliose and Plate colonies (excluding genera Acropora and Porites): e.g. Echinophyllia, Turbinaria, Montipora, Echinopora
	Lobophyllia spp	Genera	*Lobophyllia sp*
	Massive Agariciidae/ Coscinaraeidae	Family and Morphology	Massive, Submassive and Encrusting colonies with small or invisible polyps (including columnar forms). E.g., *Pavona; Psammocora, Coscinaraea, Gardineroseris*
	Massive Other		Massive, Submassive and Encrusting colonies with small or invisible polyps (including columnar forms). E.g., *Pavona; Psammocora, Coscinaraea, Gardineroseris*
	Massive Poritidae		Family Poritidae, massive and sub-massive morphologies, e.g. *P. lobata; P. lutea*
	Meandroid Faviidae/Mussidae		Massive Submassive Encrusting colonies with meandering ridges and valleys resembling brain. Includes: *Platygyra, Leptoria, Goniastrea*
	Other Acroporidae		Other corals from the family Acroporidae (e.g. *Isopora spp*)
	Other hard corals	Broad functional groups	Other hard corals
	Pocilloporidae	Family	Family Pocilloporidae; all genera (excluding Stylophora) and species
	Stylophora spp	Genera	Branching *Stylophora spp*
	Table Acroporidae	Family and Morphology	Family Acroporidae, table and corymbose morphologies
Other	Other	Broad functional groups	Other organisms, materials and substrate nor reported in other labels. Includes: fish, trash, transect hardware
	Sand		Sand. Unconsolidated reef sediment
	Seagrass		Seagrass
Other Invertebrates	*A. planci (COTS)*	Species	Crown of Thorns Sea Star. *Acanthaster planci*
	Cliona spp	Genera	*Cliona viridis* complex
	Massive sponge	Morphology	Massive or encrusting sponges
	Millepora spp	Genera	*Millepora sp*
	Mobile invertebrates	Morphology	Mobile invertebrates: e.g., Echinoderms, Lobsters
	Other sponge	Broad functional groups	Other sponges
	Sessile invertebrates	Broad functional groups	Other sesile invertebrates. Includes: bryozoa, molluscs, ascidians, soft hexacorrallia and hydroids
Soft Coral	Digitate Alcyoniidae	Family and Morphology	Family Alcyoniidae, digitate morphology. Common genera Lobophytum, Sarcophyton and Sinularia
	Other soft corals	Broad functional groups	Other soft corals
	Sea fans, plumes and whips		Sea fans, whips and plumes
Region: Central Pacific Ocean			

Algae	Crustose Coralline Algae	Morphology	Crustose coralline algae
	Cyanobacteria		Filamentous cyanobacteria
	Epilithic Algal Matrix		Multi-specific assemblages of filamentous algae and CCA smothering reef surface 1 cm or less in height
	Macroalgae		Upright macroalgae > 1 cm in height (all genera and species)
Hard corals	Massive Poritidae	Family and Morphology	Family Poritidae, massive and sub-massive morphologies, e.g. *P. lobata; P. lutea*
	Montipora spp	Genera	*Montipora sp*
	Other hard corals	Broad functional groups	Other hard corals
	P. compressa	Species	*Porites compressa*
	Pocilloporidae	Family	Family Pocilloporidae, all genera and species
	Porites rus/monticulosa	Species	*Porites rus* and *P. monticulosa*
	Porites spp	Genera	*Porites spp*. Includes: nodular and encrusting morphologies
Other	Other	Broad functional groups	Other organisms, materials and substrate nor reported in other labels. Includes: fish, trash, transect hardware
	Sand		Sand. Unconsolidated reef sediment
	Sediment other		Terrigenous sediments
Other Invertebrates	Mobile invertebrates	Morphology	Mobile invertebrates: e.g., echinoderms, lobsters
	Sessile invertebrates	Broad functional groups	Other sesile invertebrates. Includes: bryozoa, molluscs, ascidians, soft hexacorallia and hydroids
Region: South-East Asia			
Algae	Crustose Coralline Algae	Morphology	Crustose coralline algae
	Cyanobacteria		Filamentous cyanobacteria
	Epilithic Algal Matrix		Multi-specific assemblages of filamentous algae and CCA smothering reef surface 1 cm or less in height
	Macroalgae		Upright macroalgae > 1 cm in height (all genera and species)
Hard corals	Branching Acroporidae	Family and Morphology	Family Acroporidae; branching morphology (including hispidose type branching)
	Branching hard corals		Non-Acropora branching genera including: *Seriatopora, Anacropora, Echinopora, Montipora, Tubastrea* (excluding *Pocillopora* and *Stylophora*)
	Branching Poritidae		Family Poritidae, branching morphology e.g. *Porites cylindrica*
	Encrusting Poritidae		Family Poritidae, encrusting morphologies e.g. *Porites lichen*
	Foliose corals	Morphology	Thin Foliose and Plate colonies (excluding genera *Acropora* and *Porites*): e.g. *Echinophyllia, Turbinaria, Montipora, Echinopora*
	Lobophyllia spp	Genera	*Lobophyllia sp*
	Massive Agariciidae/ Coscinaraeidae	Family and Morphology	Family Agariciidae and Coscinaraeidae. Massive, submassive and encrusting morphologies with small or invisible polyps (including columnar forms). E.g., *Pavona; Psammocora, Coscinaraea, Gardineroseris*
	Massive Other		Massive, Submassive and Encrusting colonies with small or invisible polyps (including columnar forms). E.g., Pavona; Psammocora, Coscinaraea, Gardineroseris
	Massive Poritidae		Family Poritidae, massive and sub-massive morphologies, e.g. *Porites lobata; P. lutea*
	Meandroid Faviidae/Mussidae		Massive Submassive Encrusting colonies with meandering ridges and valleys resembling brain. Includes: *Platygyra, Leptoria, Goniastrea*
	Other Acroporidae		Other corals from the family Acroporidae (e.g. *Isopora spp*)
	Other hard corals	Broad functional groups	Other hard corals
	Pocilloporidae	Family	Family Pocilloporidae; all genera (excluding Stylophora) and species
	Stylophora spp	Genera	Branching Stylophora spp
	Table Acroporidae	Family and Morphology	Family Acroporidae, table, corymbose and digitate morphologies
Other	Other	Broad functional	Other organisms, materials and substrate nor reported in other labels. Includes: fish, trash, transect hardware
	Sand		Sand. Unconsolidated reef sediment
	Sediment other		Terrigenous sediments
Other Invertebrates	*A. planci* (COTS)	Species	Crown of Thorns Sea Star. *Acanthaster planci*
	Massive sponge	Morphology	Massive or encrusting sponges
	Millepora spp	Genera	*Millepora sp*
	Mobile invertebrates	Morphology	Mobile invertebrates: e.g., echinoderms, lobsters
	Other sponge	Broad functional groups	Other sponges
	Rope sponge	Morphology	Height >> basal area, spread along the substrate

	Sessile invertebrates	Broad functional groups	Other sesile invertebrates. Includes: bryozoa, molluscs, ascidians, soft hexacorrallia and hydroids
	Tube sponge	Morphology	Tube species: tube opening ~ height, basal area < opening
Soft Coral	Digitate Alcyoniidae	Family and Morphology	Family Alcyoniidae, digitate morphology. Common genera *Lobophytum, Sarcophyton* and *Sinularia*
	Other soft corals	Broad functional	Other soft corals
	Sea fans, plumes and whips		Sea fans, whips and plumes

SM4. Test Transects

To locate the 30m units (test transects) within the 2km-transects, we used images (for all surveyed years in each transect) and applied hierarchical clustering with Ward's method [5], as implemented in the "hclust" function of R [6]. The Ward's method creates clusters that minimize within-cluster variance in the distance metric, in our case, we use the physical distance among points (images) and cut the cluster at 30m to identify the aggregation groups. From these units, a number test transects per region (Table 1, main text) were selected at random and the benthic composition of the images comprised within these transects was averaged and contrasted between the two methods evaluated in this study: manual vs automated annotation (observer vs machine).

In order to assess changes in community composition over time, in a replicable and predictable fashion, we created a buffer polygon around the 30m units and extracted the images for each year that were contained within this polygon. For each polygon, coral cover was estimated by manual and automated methods, and the absolute change in cover was compared. Buffer polygons of 30 m length and 20m width were designed around the centroid of each cluster unit using the package "rgeos" from R [6]. The buffer polygons defined the sub-transect region and the benthic coverage for each year was extracted by averaging the cover across all images contained within each polygon for a given year.

As discussed in the manuscript, images were grouped within transects for three main reasons: 1) consistency in the definition of sample unit for coral reef monitoring; 2) evaluating the ability of automated methods in detecting change over time, and; 3) compatibility of observations with existing monitoring data to evaluate continuity in coral reef monitoring data. However, we do recognised that a reader may be interested on the error of automated estimations within an image and not aggregated within transects. When summarising the errors for benthic estimation within images (Figure S7) shows a slight increment in the values (~ 1 - 2%), indicating noisier estimations at the image level. Such increase in error can be attributed in part to the sampling unit size and to the heterogeneity in quality properties across images that may affect the classifier (e.g., luminescence, light reflectance, composition, topographic complexity). As discussed in the manuscript, image quality can affect the accuracy of the classifier resulting in either over- or under- estimation of benthic groups. Some of this variability in the performance of the classifier can then be cancelled or reduced by considering a larger sampling unit size.

Figure S7. Absolute error (|E|) for automated estimation of benthic abundance within an image. Errors are aggregated by functional groups along the y axis and, regions along the x axis. Solid and error bars represent the mean and the 95% Confidence Intervals of the absolute error, respectively. Please note, while this figure resembles Figure 4 in the manuscript, the absolute errors represented here are averaged among all images within a region, without aggregating by test transects.

SM5. Monitoring Program Data

Table S3. Summary of locations from monitoring programs in Hawaii, Bermuda and Australia used to compare the capacity of automated image analysis to ensure data continuity. These locations were chosen because their proximity in time and space to XL-CSS survey sites (within a radius of 2km).

Region	Program	Sub-region	Reef	site	survey years	Latitude	Longitude	Depth
Central Pacific	NCRMP	Main Hawaiian Islands	Hawaii	HAW-49	2016	19.93	-155.89	14
Central Pacific	NCRMP	Main Hawaiian Islands	Lanai	LAN-12	2016	20.74	-156.92	15
Central Pacific	CRAMP	Main Hawaiian Islands	Maui	MaKah07m	2015	20.94	156.69	7
Central Pacific	CRAMP	Main Hawaiian Islands	Maui	MaOlo07m	2015	20.81	156.61	7
Australia	LTMP	Cairns	ST CRISPIN REEF	1	2012, 2014, 2016	-16.07	145.85	8
Australia	LTMP	Cairns	ST CRISPIN REEF	2	2012, 2014, 2016	-16.07	145.85	8
Australia	LTMP	Cairns	ST CRISPIN REEF	3	2012, 2014, 2016	-16.08	145.85	8
Australia	LTMP	Cairns	OPAL (2)	1	2012, 2014, 2016	-16.20	145.91	8
Australia	LTMP	Cairns	OPAL (2)	2	2012, 2014, 2016	-16.21	145.91	8
Australia	LTMP	Capricorn Bunkers	ONE TREE REEF	3	2012, 2014, 2016	-23.49	152.10	8
Australia	LTMP	Lizard Island	YONGE REEF	3	2012, 2014, 2016	-14.58	145.62	8
Australia	LTMP	Townsville	MYRMIDON REEF	1	2012, 2014, 2016	-18.26	147.38	8
Australia	LTMP	Townsville	MYRMIDON REEF	2	2012, 2014, 2016	-18.25	147.39	8
Australia	LTMP	Townsville	MYRMIDON REEF	3	2012, 2014, 2016	-18.25	147.39	8
Australia	LTMP	Townsville	KNIFE REEF	1	2012, 2014, 2016	-18.57	147.58	8
Australia	LTMP	Townsville	KNIFE REEF	2	2012, 2014, 2016	-18.57	147.58	8
Australia	LTMP	Townsville	KNIFE REEF	3	2012, 2014, 2016	-18.57	147.58	8
Australia	LTMP	Townsville	CHICKEN REEF	2	2012, 2014, 2016	-18.66	147.72	8
Australia	LTMP	Townsville	CHICKEN REEF	3	2012, 2014, 2016	-18.66	147.72	8
Australia	LTMP	Townsville	DAVIES REEF	1	2012, 2014, 2016	-18.81	147.67	8

Australia	LTMP	Townsville	DAVIES REEF	2	2012, 2014, 2016	-18.81	147.67	8
Western Atlantic	BREAM	North Bermuda	Devils	LTEM 02	2015	32.43	-64.87	10
Western Atlantic	BREAM	Southeast Bermuda	Sonesta	LTEM 26	2015	32.24	-64.83	10
Western Atlantic	BREAM	Southwest Bermuda	Chaddock	LTEM 29	2015	32.25	-64.95	10
Western Atlantic	BREAM	Southwest Bermuda	Chub Head	LTEM 33	2015	32.29	-65.00	10
Western Atlantic	BREAM	West Bermuda	WBC	LTEM 40	2015	32.37	-64.92	10

SM6. Cost-Benefit Analysis

Given the validation results from the evaluation of performance (see main text), the cost and benefit of implementing deep learning automated image annotation in coral reef monitoring were explored using a case-based example.

From our experience, an expert observer tends to annotate approximately 50 images in an eight-hour day (~ 6 img.hr-1), using 50 random point annotations per image. Based on the casual hour rates at the University of Queensland, we estimated that the cost of annotating as single image by an expert observer would be US $5.41.

Typically, it would take about five working days invested from a machine learning expert to train, fine-tune and optimise network architecture and hyper-parameters, validate the outputs and produce the final data (e.g., benthic cover estimations for the entire dataset). Considering the average cost of US $1 per computation hour (AWS EC2 P2 instance, including storage costs) and that about 14 days of computing time will be required to produce the data (including the steps aforementioned), the computing costs for training and producing the data would be equal to US $336.

If we consider that an average of 600 images is needed to train and optimise the deep learning network, at the rate aforementioned, it would cost about US $3,247 to train the machine using manually annotated images. Based on that deep learning networks can estimate the benthic coverage of about 12,000 images per hour (200 fold more efficient than manual annotations), the individual cost of annotating a single image using machine learning would be equal to US $0.07 (Figure 7a and 7b, main text).

Importantly, the above-mentioned costs will carry an on-going cost for the manual labour from a machine learning expert required to re-train, optimise and produce new data. Every time a machine will need to be reconfigured, for example for classifying a new region, an additional cost of five days of expert labour and 14 days of computing time will be added to the total cost (about US $1740). If a machine is already trained and optimised for a given dataset, and no further developments in the architecture or framework are required, new images from the same region will only require the expert labour and computing time to run the classifier on the new images. This would be about two days from a machine learning expert and a day of AWS usage per every 25k images required to be classified (about US $600 for as much as 25k images to process).

References:

1. González-Rivero, M.; Beijbom, O.; Rodriguez-Ramirez, A.; Holtrop, T.; González-Marrero, Y.; Ganase, A.; Roelfsema, C.; Phinn, S.; Hoegh-Guldberg, O. Scaling up Ecological Measurements of Coral Reefs Using Semi-Automated Field Image Collection and Analysis. *Remote Sens.* **2016**, *8*, 30, doi:https://doi.org/10.3390/rs8010030.
2. Lee, J.; Weger, R.C.; Sengupta, S.K.; Welch, R.M. A neural network approach to cloud classification. *IEEE Transactions on Geoscience and Remote Sensing* **1990**, *28*, 846–855, doi:10.1109/36.58972.
3. Yosinski, J.; Clune, J.; Bengio, Y.; Lipson, H. How transferable are features in deep neural networks? Available online: https://arxiv.org/abs/1411.1792 (accessed on 1 December 2019).
4. Simonyan, K.; Zisserman, A. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556* **2014**. Available online: https://arxiv.org/abs/1409.1556 (accessed on 1 December 2019).
5. Ward Jr, J.H. Hierarchical grouping to optimize an objective function. *Journal of the American statistical association* **1963**, *58*, 236–244.
6. R Core Team. *R: A language and environment for statistical computing.*, R Foundation for Statistical Computing: Vienna, Austria, 2013. Available online: https://repo.bppt.go.id/cran/web/packages/dplR/vignettes/intro-dplR.pdf (accessed on 1 December 2019).