A fully automated method for monitoring the intertidal topography using video monitoring systems

Antoine Soloy, Imen Turki, N. Lecoq, Ángel David Gutiérrez Barceló, Stéphane Costa, Benoît B. Laignel, Benjamin Bazin, Yves Soufflet, Loïc Le Louargant, Olivier Maquaire

A fully automated method for monitoring the intertidal topography using Video Monitoring Systems

AUTHORS:

Antoine Soloy[a], Imen Turki[a], Nicolas Lecoq[a], Ángel David Gutiérrez Barceló[d], Stéphane Costa[b], Benoit Laignel[a], Benjamin Bazin[a], Yves Soufflet[e], Loïc Le Louargant[c], Olivier Maquaire[b]

[a] Normandie Univ, UNIROUEN, UNICAEN, CNRS, M2C, 76000 Rouen, France

[b] Normandie Univ, UNICAEN, CNRS, LETG-Caen, 14000 Caen, France

[c] Département de la Seine-Maritime, Service Ouvrages Littoral et Seine, Hotel du Département, Quai Jean Moulin, 76000 Rouen

[d] SandS Corp, S.L., Centro de Negocios Finca Pontanía (La Albericia), Calle Río Danubio 1, Planta 1, Oficina 16, 39012 - Santander, Spain, www.sands.es

[e] Waves'n See, Meteo France, Incubateur Green Tech Verte, Bâtiment D'Alembert, 42 Avenue Gaspard Coriolis, 31057 Toulouse Cedex, www.wavesnsee.com

Corresponding author : Antoine Soloy, antoine.soloy@univ-rouen.fr

ABSTRACT:

Coastal systems are extremely dynamic environments exposed to many hazards, making accurate and regular monitoring a major challenge, particularly in the context of global change and sea level rise. In this frame of reference, high-frequency, high-resolution coastal Video Monitoring Systems (VMS) have been installed on three megatidal (tidal amplitude > 9 m) sites of Normandy (France) including a sandy beach at Villers-sur-Mer, a pebble beach at Etretat and a composite beach at Hautot-sur-Mer. This article proposes the use of Mask R-CNN to process images acquired at these sites and perform the automatic segmentation of the visible bodies of water in order to extract the waterline. The extracted waterlines are associated with a measured water level, which makes it possible to reconstruct the topography of the beaches at the scale of the tidal cycle. After training the neural network on manually labeled data, the segmentation by Mask R-CNN is very efficient by achieving a satisfactory segmentation on 69.87% of the images of Villers-sur-Mer, on 67.11% at Hautot-sur-Mer, and on 97.33% at Etretat. Once the waterlines have been extracted and georeferenced, the reproduction of the beaches' morphology is satisfactory (averaged vertical RMSE = 28 cm). These results confirm that segmentation by Mask R-CNN is a particularly powerful tool that allows efficient and low-cost monitoring of the evolution of beach morphology,

particularly in response to marine conditions. Its capabilities to detect and segment bodies of water while not being affected by the various sources of noise make it a remarkably effective tool for coastal science applications.

KEYWORDS:

Intertidal Topography, Shoreline Detection, Coastal Video Monitoring, Mask R-CNN, Coastal Morphodynamics, Pebble Beach

1. INTRODUCTION[1]

Coastal environments are continuously subjected to the natural processes of weathering, marine erosion and flooding (Nicholls et al., 2007). The impacts of these processes and events vary depending on the geometric structure of the coastlines, their characteristics, and their relative exposure to the impacts of waves and tides.

In the context of climate change and the growing environmental strain caused by human activity, coastal communities are increasingly vulnerable to environmental hazards. Additionally, these communities are facing an intensification of natural hazards including shoreline change induced by coastal erosion, and changes in nearshore topography (Jongejan et al., 2016; Le Cozannet et al., 2019; Ranasinghe et al., 2012; Wainwright et al., 2015).

In this frame of reference, beaches provide protection against wave action by dissipating their energy when their sediment is mobilized, especially during storm event. These systems are subject to multiple maritime forcing whose relative influence determines the morphological response of the beach. Although the dynamics of these systems are complex, they are not impossible to predict (Davidson et al., 2017; Esmail et al., 2019; Hanson, 1989; Jara et al., 2015; Montaño et al., 2020; Yates et al., 2009). It is therefore important to monitor the geomorphology of coastal systems in order to acquire capabilities of prediction.

Among the different methods used for monitoring the coastal morphology, Video Monitoring Systems (VMS) allow the morphology of a site to be measured thanks to both variations of the water level and the position of the waterline, with a spatial resolution of a few centimeters, and a temporal resolution of one day or less.

With this in mind, the Normandy coastline has been equipped with VMS on 3 study sites in 2018 and 2019: Etretat, Hautot-sur-Mer and Villers-sur-Mer. The Normandy coasts are distinguishable among others by three main characteristics: (1) the presence of pebble ridges, (2) the mega-tidal range (> 9 m), and (3) the presence of wide sandy intertidal area (> 200 m cross-shore). With their high-frequency measurement capabilities and low cost, VMS are ideally suited to monitor these environments in consideration of their tidal and morphological specificities.

However, extracting the position of the coastline from VMS images remains a long and often tedious task. Historically, it was first necessary to manually delineate the waterline (Holman et al., 1991), then it was made easier with the development of several detection

methods (Plant et al., 2007). A large part of these methods is inherited from the SLIM (ShoreLine Intensity Maximum) model of Plant and Holman (1997), based on the presence of a wave breaking zone with a visible high light intensity near the waterline. Examples include the PIC (Pixel Intensity Clustering) or the method of Aarninkhof et al., (2003) also known as BIM (Intertidal Beach Mapper), the ANN (Artificial Neural Network) model of (Kingston, 2003), and the CCD (Color Channel Divergence) model of (Turner et al., 2001).

Although these methods have greatly improved data processing time, their application is limited by variability in data quality and requires extensive expertise (Osorio et al., 2012). More recently, new approaches have been developed in order to address these constraints and move closer to a fully automatic detection. For instance, Osorio et al. (2012) developed the Physical and Statistical Detection Model (PSDM) based on edge detection algorithms. Another example is the SDM (Shoreline Detection Model) of Valentini et al. (2017) which proposes to perform the semantic segmentation of the image pixels in order to determine the land/water interface by processing the histogram of RGB channels. Nevertheless, most of these algorithms remain strongly constrained by the multiple environmental variabilities (luminosity, contrast, rain, fog, presence of users, boats, dirty lenses, etc.) and therefore require the action of a human operator to fine-tune the parameters before being applied to a new site.

To overcome such limitations, the present research proposes the use of deep learning techniques to be applied to a large dataset of remote sensing images from coastal VMS for mapping nearshore intertidal topography. Many advances have been made in the field of image classification, notably with the rise of the use of convolutional neural networks (CNN) as reported in various academic articles (Chen et al., 2020, 2019; He et al., 2017; Huang et al., 2019; Kirillov et al., 2020; Wang et al., 2020).

More specifically, if trained using enough data, the algorithm called Mask R-CNN (He et al., 2017) can detect and delineate concepts as abstract as coastline delimited bodies of water on images under various conditions, unaffected by the presence of objects and people. During this study, Mask R-CNN was used for mapping the nearshore topography in three megatidal coastal areas characterized by different morphological properties with the aim to propose a fully automated method for monitoring coastal morphology using VMS.

This paper is structured as follows: following this introduction, section 2 describes the 3 study sites, the dataset and presents the new methodological approach used to determine the nearshore intertidal topography. The methodology results are validated and discussed in section 3. Finally, concluding remarks and further works related to coastal applications are suggested in section 4.

2. Datasets and Methodological Approach

2.1. Study sites and monitoring systems

The monitored sites are located in France, on the Normand side of the English Channel (Figure 1). These coastal locations are facing some of the largest macrotidal ranges in the world, with an average range of 9.5 m observed on the study sites, and a record of

15 m observed at Mont Saint Michel (Bonnefille, 1968; Chabert D'Hières and Le Provost, 1978; SHOM, 1953), further west. As 4 m is the threshold above which tidal ranges are classified as macrotidal in the Davies (1964) classification, greater ranges are poorly represented by this scale. Because of this, Levoy et al. (2000) proposed to add the megatidal category for amplitudes greater than 8 m into this classification, thus better representing the extreme ranges observed in Normandy. According to this classification, the study sites' shorelines, with more than 9 m amplitude, fall into this category.

The Normand seaboards are also heavily man-influenced, with the presence of engineering structures (dikes, groins, breakwaters, ports...), and numerous users and usages (tourism, fishing, navigation...). Therefore, it is useful and relatively cost-efficient to install VMS on existing buildings in order to survey how these structures and users influence the shoreline dynamics.

The three study sites covered in this paper have been chosen to represent the wide variability of coastal systems present in Normandy. They are located in Villers-sur-Mer, Etretat, and Hautot-sur-Mer (Figure 1). Tidal dimensions of each site are given in Table 1.

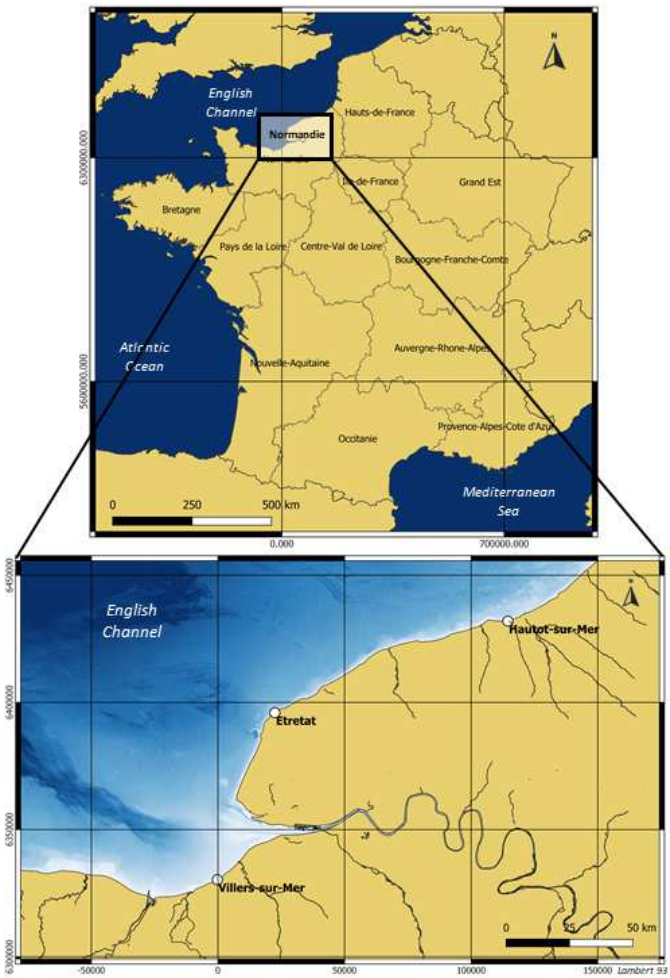

Figure 1 - Location map of Villers-sur-Mer, Etretat and Hautot-sur-Mer, in Normandy, France

Table 1 – Tidal and intertidal dimensions of the Normandy beaches of Villers-sur-Mer, Etretat and Hautot-sur-Mer

Sites	Villers-sur-Mer	Etretat	Hautot-sur-Mer
Average tidal range	5.11 m	6.08 m	6.79 m
Max tidal range	9.60 m	9.13 m	9.86 m
Min tidal range	0.16 m	3.20 m	2.96 m
Lowest Water Level (LWL) (datum: IGN69)	-3.81 m	-4.80 m	-4.45 m
Beach width at LWL	310 m	150 m	210 m

2.2. Methodological approach

The methodology for building intertidal topography from VMS is a multi-step process that was described by Aarninkhof et al. (2003). It was applied to the newly installed coastal VMS in Normandy, which is original in that it implements the use of the Mask R-CNN as the waterline classification model. As presented in Figure 3, the method includes two branches.

The first branch represents the series of steps that are necessary to implement a model able to detect bodies of water, in the event that there is no existing model already trained leading to satisfying results. Training the neural network requires manually labeled image data, as explained in section 2.2.3 in more detail. This series of steps can be repeated to increase the detection performances of a previously trained neural network, using new labeled data.

The second branch is the operational stage. Once the model is trained and shows satisfying results, the neural network can delineate waterlines without further human labeling, allowing the methodology described in section 2.2.4 to efficiently produce intertidal Digital Elevation Models (DEMs).

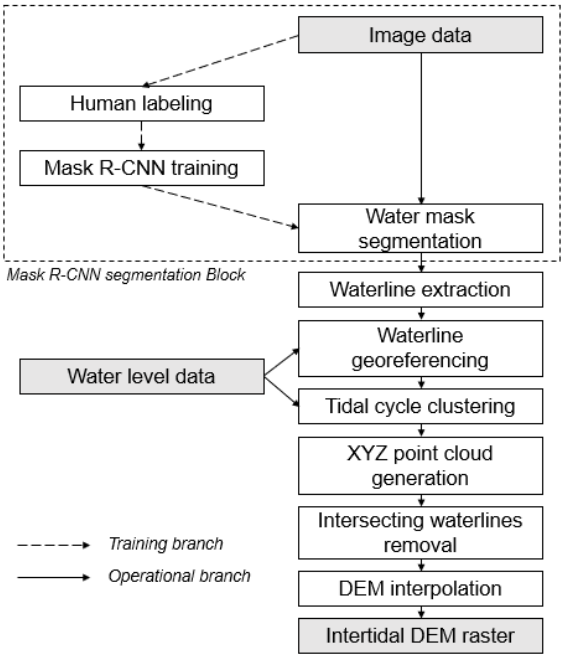

Figure 2 - Algorithm for the intertidal topography reconstruction process: First an image is analyzed by the Mask R-CNN model to be segmented, either for training or operational purpose. Then the shoreline is extracted and stacked with other shorelines belonging to the same tidal cycle. Intersecting waterlines are removed before interpolating the point cloud.

2.2.1. Data acquisition

All three VMS are composed of three cameras with different fields of view, each of which records 6 images per hour during daylight. This produces a total of 18 images per hour per study site, covering the entire beach (Figure 4). These so-called "timex" images are recorded with an exposure time of 10 minutes. Without considering possible losses likely due to malfunctions and bad imaging conditions (dirty lenses, foggy weather, rain drops, etc.), the expected data set for one year includes about 78000 images, with a resolution of 1936x1216 px at Hautot-sur-Mer, 3264x1856 px at Villers-sur-Mer and 3840x2160 px at Etretat.

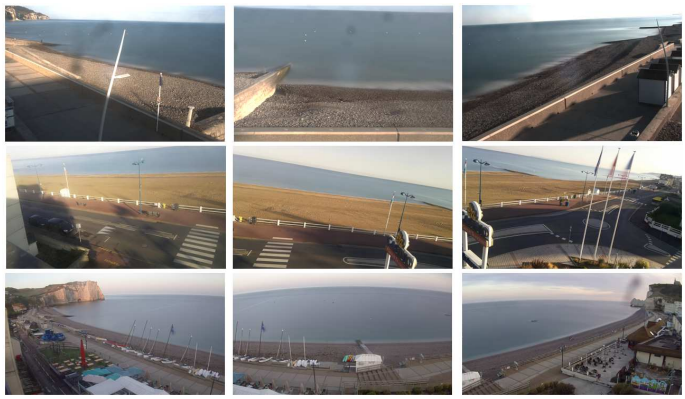

Figure 3 - Example of timex panoramas at Hautot-sur-Mer (top), Villers-sur-Mer (center) and Etretat (bottom)

120 *2.2.2. Shoreline definition and manual delineation*

121 Besides Dolan's et al. (1980) definition stating that the shoreline represents the interface between land and sea, there is currently no
122 real consensus on a more accurate definition for a shoreline. However, Mask R-CNN's output fully relies on the way that the human
123 operator providing the model's training data will understand this definition and eventually delineate the shoreline accordingly by hand.
124 Clearly defining the object of our interest is therefore an important step in this methodology.

125 In previous video-based monitoring studies, the shoreline has often been defined as a visible component of the wave breaking zone
126 that can be identified thanks to the SLIM (ShoreLine Intensity Maximum) introduced by Plant and Holman (1997). As the objective is
127 to measure the intertidal topography using the waterline as an indicator associated with a previously known water level, these
128 indicators remain reliable. Nevertheless, they strongly depend on the presence of a foamy wave breaking zone, which is likely to
129 constrain the shoreline detection capabilities to wavy conditions, while calmer and waveless ones wouldn’t be processable despite
130 their likely better accuracy.

131 In this study, it was decided to define the shoreline as the line that a human operator is able to detect as the separation between land
132 and water on timex images, regardless to the presence of wave breaking conditions. This definition is closer to the approaches adopted
133 by Aarninkhof et al. (2003), Kingston (2003), Osorio et al. (2012) and Turner et al. (2001), considering the shoreline as the interface
134 between a wet and a dry domains, but remains more open to interpretation as it allows the operator to include or exclude elements to
135 either domains depending on the context (e.g. a floating boat can be considered wet, an object occulting the shoreline can be ignored).
136 With this in mind, an operator could manually delineate the body of water that represents the sea on a timex image. The polygon thus
137 digitized will be limited by the sides of the image, the horizon and the defined shoreline (Figure 6). In order to avoid the ambiguity
138 implied by the multiple meanings of the word “shoreline”, the detected lines will be called “waterlines” in the rest of this paper.

139 *2.2.3. Mask-RCNN description, parametrization and training*

140 The first objective of the present study is to detect the waterline on shoreline images. In order to carry out this work in an
141 automatable way that would be adaptable to the many sources of noise specific to coastal VMS data, a deep learning algorithm was
142 trained using a bank of previously labeled images.

143 The algorithm that was chosen to achieve this task is called Mask Regional Convolutional Neural Network (Mask R-CNN). This
144 meta-algorithm developed by He et al. (2017) combines the proven object detection algorithms Faster R-CNN (Ren et al., 2016), that
145 allows the identification of the nature and position of objects by bounding boxes, and Fully Convolutional Network (FCN) (Long et
146 al., 2015), allowing semantic segmentation, i.e. classification at the pixel scale.

147 The Mask R-CNN working process can be broken down into two stages (Figure 5). The first stage scans the image and generates
148 anchors (areas likely to contain an object). To do this, a backbone consisting of a Residual Neural Network 101 (ResNet101, a
149 convolutional neural network architecture of 101 layers deepness) and a Feature Pyramid Network (FPN, a succession of features

maps with an increasing level of abstraction) first extracts the spatial information. Next, a Region Proposal Network (RPN) is responsible for generating random areas of interest, and then classifies their relevance according to the output of the backbone.

The second stage classifies the proposals from the first stage and generates bounding boxes and masks in parallel. This is made possible by a Regions Of Interest (ROI) classifier that performs object detection (nature and bounding box) on the anchors, in order to extract the ROI. These are then homogenized by ROI pooling at a scale of 28 by 28 pixels, in order to perform a semantic segmentation on each of them and to obtain a mask. Finally, the mask is resized to the initial scale, and repositioned in the image.

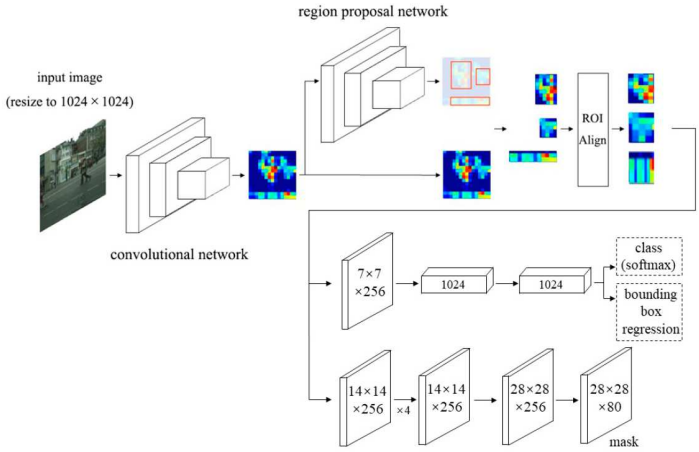

Figure 4 - The Pipeline of Mask R-CNN (Zhang et al., 2020)

The neural network was trained on a total of 1062 manually labelled images from the study sites' VMS. The images cover as wide a set of situations as possible, including sunny weather, storms, rain, fog, backlight sun, sunrise and sunset, presence of people and objects on the beach (tourists, boats), in the water and in front of the waterline, presence of insects on the camera lens, and so on. A total of 95 epochs each including 1000 steps with various augmentation operations (rotations, crop, addition of noise) were necessary to obtain a satisfactory detection quality.

The algorithm learning performance is evaluated during learning steps by comparing the detected masks to the manually digitized ones. This is done using a set of metrics including class loss, mask loss and boundary box loss, each of which is calculated separately for both the training and the validation datasets. The evolution of these metrics helps to prevent the model from overfitting. The detection is considered satisfactory when the algorithm plateaus at a minimum value of loss.

The influence of the human labeling bias on the algorithm's results is considered negligible as the mask output provided by Mask R-CNN will mandatorily be 28 x 28 pixels in size, as constrained by the neural network's architecture, before being upscaled to the image original size. Therefore, an aliasing effect can sometimes be visible (Figure 6). As a consequence, the maximal resolution of detectable morphological structures on a specific object is equal to 1/28 of the total size of the object. On shoreline images for instance, some structures called cusps sometimes appear as a series of oscillations of the waterline along the beach. In this case, only cusps with a wavelength appearing larger than 3/28 of the body of water's width in pixels (often equivalent to the width of the image)

will be detectable. For the same reasons, the pixel uncertainty range of the detected mask edge will be relative to the object itself, equaling 1/28 of the detected object's bounding box dimensions. Thus, the metric accuracy is not consistent as it is always relative to the detected object size, which changes from one image to another, and is expected to decrease with the distance to the camera, as pixels cover wider surfaces. In addition, it is interesting to note that as the body of water's apparent size may depend on the state of the tide – especially in images from cameras pointing towards the alongshore direction - the detection accuracy is expected to increase at low tide as the body of water becomes smaller, as compared to high tide.

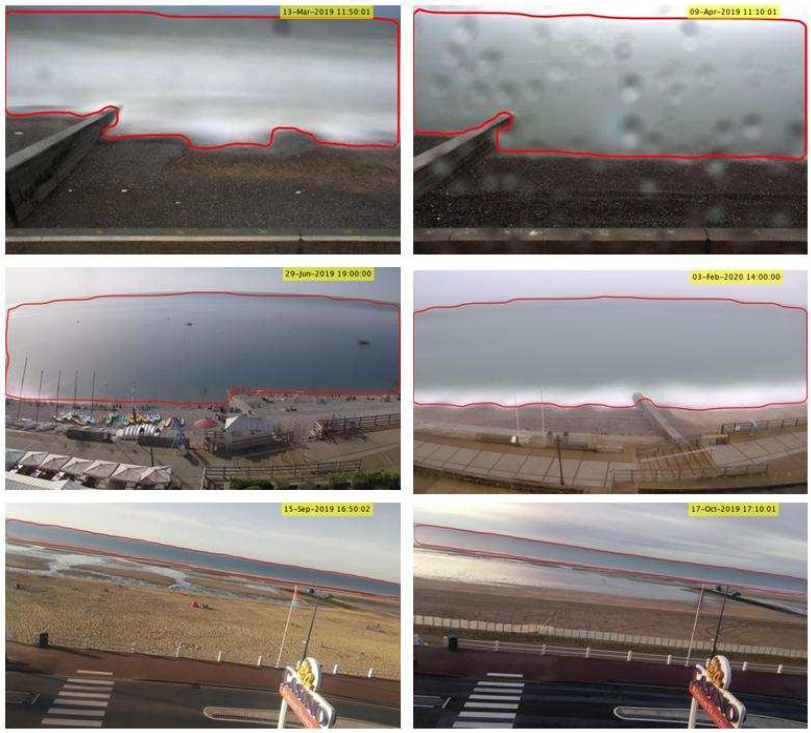

Figure 5 - Example of body of water detection using Mask R-CNN at Hautot-sur-Mer (top), Etretat (center) and Villers-sur-Mer (bottom), during various filming conditions

As shown on Figure 6 the algorithm performs well in most cases even when conditions are not optimal such as during rain, storms, presence of users, presence of channels and ponds at low tide, etc. However, these conditions are also likely to confuse the model (Figure 9) and therefore can represent a limit to detection. The strongest advantage of the Mask R-CNN lies in the generalization and abstraction capabilities inherent in convolutional neural networks, which makes it possible to train a model only once, and then use it on various types of sites, orientations and conditions. Moreover, it is always possible to improve it by completing or even complexifying the training dataset in order to reduce the number of errors and open the detection capabilities to new object classes.

2.2.4. Intertidal topography reconstruction

The first step of the operational branch is the Mask R-CNN segmentation. Each timex image is subjected to detection by the trained neural network, which will delineate the edges of the visible body of water.

Then, the waterline is extracted by a ROI mask that also excludes the groins from the detected polygon (Figure 7). and is associated to its respective water level. This water level allows georeferencing, i.e. the transformation of image coordinates into geographical XYZ coordinates by projecting the waterline on the plane of the water surface.

Figure 6 - Example detected waterlines on the 3 cameras of the Hautot-sur-Mer's VMS

Once georeferenced, waterlines belonging to the same half semi-diurnal tidal cycle (i.e over about 6h12) are clustered together, thus creating XYZ intertidal point clouds (Figure 8a). This allows the identification of the potentially miss delineated waterlines which are likely to cross one other waterline or more. To do so, the number of intersections is calculated for each waterline, and the waterline that crosses the highest number of different other ones is removed. This operation is iteratively performed until the total number of intersections goes down to zero.

Finally, the point cloud is converted into a raster with normalized coordinates (Figure 8b) using a 2D linear interpolation. The raster is then cropped to the surface covered by the waterlines in order to remove the irrelevant extrapolated information, thus allowing the tide-to-tide comparison of the beach intertidal morphology.

a b

Figure 7 - XYZ point cloud (a) formed by the georeferenced waterlines detected using Mask R-CNN during a complete tidal cycle at Etretat and (b) interpolated DEM. Dotted lines show the surface covered by each camera of the VMS.

Although the Mask R-CNN algorithm works in the vast majority of cases (Table 2) it sometimes misidentifies the waterline, usually due to factors of confusion such as low light, presence of fog, presence of obstructions on the camera lens (water droplets or insects), or unclear waterline (too smooth transition from dry to wet sand) (Figure 9). In order to limit detection errors, it is possible to

filter out ambiguous data prior to detection. The filtering applied in this study uses an established threshold for both brightness and blur metrics to separate the good quality images from the darker and blurrier ones. For each camera, these metrics are calculated over an appropriate section of the image, in this case the sea, and objects with visible and contrasted edges respectively. Brightness is measured by adding the averaged red, green and blue channels of the image, and the blurriness index has been outlined by Crété-Roffet et al. (2007). After calculating both metrics on the whole dataset, thresholds are empirically selected from their respective histogram for each camera in order to determine the combination that allows the best filtering possible.

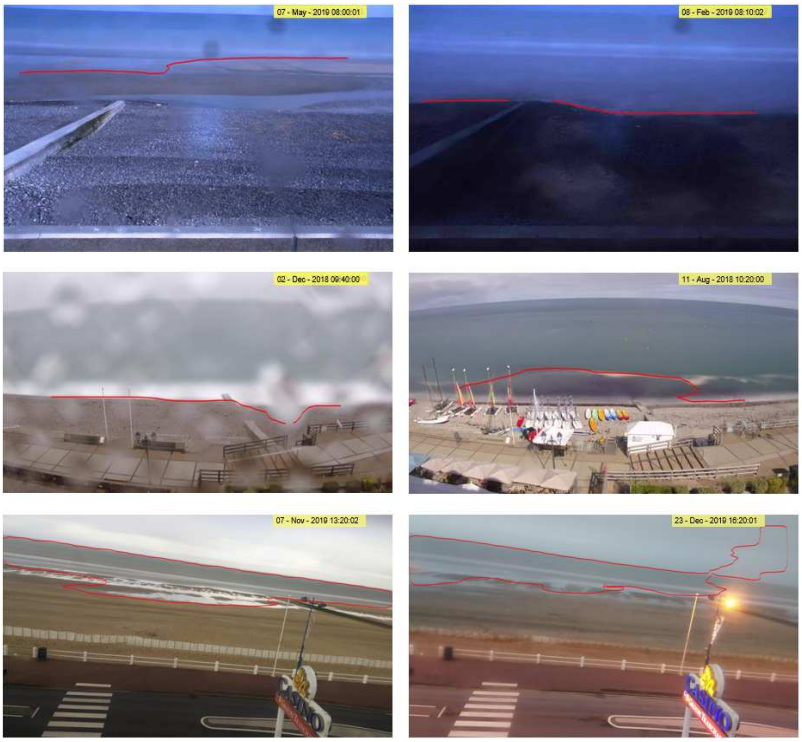

Figure 8 - Example of unsatisfactory detections using Mask R-CNN at Hautot-sur-Mer (top), Etretat (center) and Villers-sur-Mer (bottom)

In addition, it is important to mention that on images showing alongshore perspectives, the water mask resolution remains 28 x 28 pixels although the mask is then being scaled up to the body of water's actual size. Therefore, because of both the limited accuracy of 1/28 of the body of water's size and the smoothing effect due to the mask resizing process, closed angles may appear rounder, and impose a horizontal offset that can become significant for long distances to the camera (e.g. the extremity of the beach on Figure 10) . Consequently, care should be given to the selection of the ROI's size being used for extracting the waterline to a reasonable extent.

Figure 9 - Example of body of water detected by Mask R-CNN on a lateral perspective.

3. RESULTS AND DISCUSSIONS

3.1. Ground truth validation

3.1.1. Detection validation

For the purpose of estimating Mask R-CNN's capabilities to delineate the waterline, the methodology was applied to the data acquired by the central camera of each site, covering periods of 9, 20 and 11 months, between 2018 and 2020 at Villers-sur-Mer, Etretat and Hautot-sur-Mer respectively (Table 2).

After stacking the extracted waterlines on their respective images, each detection was individually and manually analyzed by a human operator in order to empirically classify the detected waterlines' quality as satisfactory or unsatisfactory. It was then possible to calculate a satisfactory detection performance value by comparing the number of images classified as satisfactory to the number of images leftover after the parametric filtering.

At Etretat, the detection by Mask R-CNN followed by parametric filtering is successful in 97.33% of the data, i.e. the manual operation will have classified only 2.67% of the detections as unsatisfactory. With 69.87% and 67.11% success respectively, Villers-sur-Mer and Hautot-sur-Mer also show very satisfactory performances, although slightly lower. This difference between Etretat, Villers-sur-Mer and Hautot-sur-Mer beaches can be explained by the texture contrasts, the slope, and the morphological complexity of the beach. At Villers-sur-Mer, the association between the gentle slope and the megatidal ranges makes the interface between the dry and the wet sand move very fast which tends to blur the intertidal area whose sand already smooths the texture of, and thus makes the identification of a clear shoreline particularly difficult on timex images, even for a human eye. In addition, the complex topography sandy dissipative beaches (bars, ponds, channel networks…) is hard to catch for the mask resolution that can't store more complex structures than 1/28 of its own size. The issue is even stronger at Hautot-sur-Mer and seem to be making the waterlines over the site's sandy area even more difficult to detect for unidentified reasons. However, the steep slope and contrasted texture of the pebble ridges makes the waterline clearly identifiable in almost all cases at Etretat and Hautot-sur-Mer, which compensates the poorer sampling

248 capabilities over the sandy area for the last site in the presented detection performance value.

249 It is important to mention that these misdetections are in the vast majority of the cases possible to automatically identify and
250 remove during the intersecting waterline removal part of this methodology (Figure 3).

251 **Table 2 – Mask R-CNN detection validation results**

Sites	Villers-sur-Mer	Etretat	Hautot-sur-Mer
Image dataset start	11/03/2019	28/06/2018	11/12/2019
Image dataset end	31/12/2019	13/02/2020	04/01/2020
Total number of images	23579	11720	27532
Images after parametric filtering	17373	10000	18490
Images manually classified as satisfactory	12138	9733	12409
Detection performance	69.87%	97.33%	67.11%

252

253 *3.1.2. DEM Validation*

254 For each study site, a validation geodesy campaign was carried out using a differential Global Navigation Satellite System
255 (dGNSS) in order to measure the ground truth topography. Measurements have been carried out with a non-gridded uniform spread of
256 the sampling points over the video recorded area, covering the beach from the highest point of its top part to the low tide waterline,
257 with number of points ranging from 100 to 150 per site in total, measured on both slope breaks and centers.

258 The intertidal topography was reconstructed at the geodesy campaign dates using the methodology described 2.2.4 and then
259 compared to the dGNSS data for validation. Comparisons are evaluated using the Root Mean Squared Error (RMSE) calculated
260 between the dGNSS points' elevation and the DEM's elevation at the same XY locations, thus corresponding to an average vertical
261 offset. As some of the dGNSS points were out of the DEM coverage, these points were excluded from the calculation, resulting in the
262 counts visible in Table 3.

263 Table 3 shows the vertical RMSE values calculated for all sites, with 22 cm at Etretat, 29 cm at Hautot-sur-Mer and 33 cm at
264 Villers-sur-Mer, which correspond to ranges found by previous studies (Uunk et al., 2010). R^2 are also provided with values ranging
265 from 0.93 to 0.99, which confirms method's ability to provide reliable beach morphology estimations on all sites.

Table 3 - DEM validation results

	Villers-sur-Mer	Etretat	Hautot-sur-Mer
DEM date	31/12/2019	10/01/2020	14/02/2020
GNSS campaign date	08/01/2020	10/01/2020	14/02/2020
GNSS points count	59	48	54
Vertical RMSE (m)	**0.33**	**0.22**	**0.29**
R^2	0.95	0.99	0.93

In addition, the evolution of the vertical offset with the omnidirectional distance to the camera was analyzed using a Mann Kendall statistical test. Results presented in Table 4 show that although the error tends to increase with the distance in all cases, the only found significant p-value at the 0.05 threshold is seen at Hautot-sur-Mer. For this site, the trend remains under 2.4 vertical mm / per horizontal meters. However, on this site the low tide shoreline can go as far as 210 m in the cross-shore direction (Table 1), therefore the vertical error can be increased by an order of 50 cm at this distance and is thus important to consider. Similar calculations would show an increase of about 20 cm and 4 cm at Villers-sur-Mer and Etretat respectively, for low tide vertical uncertainties. This error increasement is due to a combination between the lower resolution for longer distance to the camera and the Mask R-CNN intrinsic uncertainty. Nevertheless, these values should be considered as orders of magnitude as the vertical error doesn't properly fit to a linear trend, as showed by the low R^2 values in Table 4.

Table 4 - Sensitivity analysis results of the vertical error to the distance from the camera

	Villers-sur-Mer	Etretat	Hautot-sur-Mer
RSE/Dist R^2	0.04	0.01	0.06
RSE/Dist Mann Kendall Tau	0.05	0.02	0.20
RSE/Dist Mann Kendall p-value	0.58	0.82	0.04
RSE/Dist Slope	6.35E-04	3.21E-04	2.39E-03

(m/m)

3.2. Discussions

The task of extracting the shoreline position from camera images to reproduce the intertidal bathymetry has been realized with an increasing efficiency over the last decades. In most cases, authors' methodology implied the classification of pixels based on their individual intensities in order to to identify either the swash zone (Plant and Holman, 1997) or the wet and dry pixels (Aarninkhof et al., 2003; Kingston, 2003; Turner et al., 2001; Valentini et al., 2017).

The SLIM model (ShoreLine Intensity Maximum) of Plant and Holman (1997) locates the waterline along pre-defined cross-shore transects by thresholding pixel intensities in order to parametrically extract the pixels that are considered belonging to the swash zone. The edge of swash zone is a good proxy of the shoreline, but its inconsistent presence makes the technique unreliable in a lot of cases, such as dissipative beaches, intertidal zones with complex morphologies (bars, channels…), and waveless weathers.

The CDD (Color Channel Divergence) model of Turner et al. (2001) and the PIC model (Pixel Intensity Clustering) developed by Aarninkhof et al. (2003) were both created to overcome some of the issues of the SLIM model. CDD uses the expected difference between sand and water colors to find their interface point along pre-defined cross-shore transects. PIC identifies wet and dry pixels within a pre-defined ROI on images. This classification is possible using two discriminator functions that consider the hue, saturation, and grayscale values applied on each pixel. The classification of each individual pixel is then performed by thresholding the function's results. Contrary to SLIM, these two models don't rely on the presence of a swash zone, although they require a fine calibration to the site to which it is applied, as well as a significant contrast between the two domains of interest.

Both the ANN (Artificial Neural Network) and SDM (Shoreline Detection Model) models from Kingston (2003) and Valentini et al. (2017) respectively are a semantic segmentation algorithms that produce a classification of each pixel in an image as water and sand, the interface of which is the waterline. Models require to be trained on manually labeled images and the misclassified pixels need to be manually filtered.

While these methods remain efficient, they require case-specific calibration processes, rely on the concomitance of a wide range of conditions (presence of a single swash zone, absence of users/obstructions, specific difference in color between wet and dry pixels) and require a significant amount of manual work in order to be successfully deployed.

The present study overcame some of these difficulties thanks to the use of instance segmentation, in particular Mask R-CNN, with the aim of building an automatically working algorithm. Mask R-CNN is a deep learning algorithm that performs instance segmentation on RGB images, which has been used for a wide variety of applications such as measuring snow depth (Kopp et al., 2019), counting the number of cows present in a farm corridor (Qiao et al., 2019), mapping the grain size of pebbles (Soloy et al., 2020), and determining the shape of molds present on the walls of a tunnel (Zhao et al., 2020).

Mask R-CNN allowed more characteristics to be additively taken into account when identifying a body of water than parametric solutions would (textures, shapes, presence/absence of a swash zone, variabilities in brightness, colors, weather, human activity, sediment types…) thus increasing the range of satisfactory identifications while significantly lightening the need for calibration and thresholding steps. In addition, using a modular algorithm that is not specific to the coastal field such as Mask R-CNN and training the model including multiple sources data tends to optimize the generalization capabilities of this methodology, thus allowing it to be applied on new sites with limited parametrization work.

This method has proven to be reliable when detecting contours of bodies of water on shoreline images at Villers-sur-Mer, Etretat and Hautot-sur-Mer, with a detection performance rate of 69.87%, 97.33% and 67.11% respectively, and considering an accuracy of 1/28 of the detected body of water's bounding box size in pixels. For comparison, Plant et al. (2007) recorded detection performance rates ranging from 15% to 50% for SLIM, from 0 to 91% for CDD, from 24 to 78% for PIC and from 32 to 92% for ANN while comparing different methods applied to four different study sites. Rates obtained with Mask R-CNN therefore surpasses the other detection models with higher average detection rates and less difference from site to site. The exceptional rates are made possible thanks to the adaptability of Mask R-CNN and its ability to detect abstract concepts, such as a body of water, unaffected by the many sources of image variability (e.g. weather conditions, waterline obstruction, etc.), which are difficult to take into account using a parametric detection model.

Extracting and georeferencing multiple waterlines along tidal cycles allows the reconstruction of good intertidal topographies at Villers-sur-Mer, Hautot-sur-Mer and Etretat, with RMSE values of 33 cm, 29 cm and 22 cm respectively. It is important to mention that at Villers-sur-Mer the RMSE value was calculated using a DEM from a week before the geodesy campaign due to a malfunction of the VMS during this period. The value is therefore likely to be overestimated. With on average 10 cm for SLIM, 20 for CDD, and 20 to 34 cm for PIC (Plant et al., 2007; Uunk et al., 2010), vertical biased calculated for this methodology remain of the same order of magnitude as other intertidal DEM building strategies from video monitoring techniques.

However, the methodology remains sensitive to the distance of the waterline and faces difficulties to correctly identify the proper waterline in the context of a megatidal and gently sloped beach. Additionally, the neural network has troubles classifying large, complex geometries, including the presence of ponds and channels due to sand bars, which therefore represent a limit to its use. Nevertheless, misidentified waterlines can be parametrically detected and cleared out. It is therefore possible to use this new methodology for monitoring the daily intertidal topography in a fully automated way.

More generally, the model's resilience and abstraction abilities could make Mask R-CNN a powerful asset to the coastal science field, as it already is for other remote sensing applications (Maxwell et al., 2020; Nie et al., 2018; Soloy et al., 2020; Yang et al., 2020; Zhao et al., 2018). Moreover, coastal VMS are relatively efficient, high resolution and low priced in comparison to aerial and satellite technologies, especially when associated with Mask R-CNN for measuring multi-scale variables.

The Mask R-CNN method has been successfully used for other scientific and societal applications (Kopp et al., 2019; Qiao et al., 2019; Soloy et al., 2020; Zhao et al., 2020), and is now proven to be useful for coastal science applications despite its current limitations. Recently, major progress have been made in the instance segmentation field and new tools are now available with better capabilities such as BlendMask (Chen et al., 2020), TensorMask (Chen et al., 2019), Mask Scoring R-CNN (Huang et al., 2019), PointRend (Kirillov et al., 2020), SOLOv2 (Wang et al., 2020), and even more. One can expect that these tools, freely accessible for the most part, will become a standard in further coastal applications.

4. Conclusion

Monitoring the evolution of coastline morphology is a key challenge in the context of global change. Using coastal VMS addresses this problem with a high frequency, high resolution and low price that enables the surveillance of the waterline's position. The multiscale changes of the waterline's position are very important and depend on the physical characteristics of sediment, the topography and the local hydrodynamic conditions of waves and tides. The large variability of these influential factors complexifies the assessment of the shoreline and the study of its evolution at different timescales, from storm events to seasonal and interannual variations, with the aim of quantifying the long-term coastal erosion due to climate change. In this frame of reference, coastal VMS have been installed to survey the morphodynamics of 3 megatidal coastal systems of Normandy (France) including a sandy beach at Villers-sur-Mer, a pebble beach at Etretat, and a composite beach at Hautot-sur-Mer.

Extracting the waterline from the images provided by coastal VMS can be a complicated and time-consuming task because of the many sources of noise specific to this type of data (e.g. weather, sun angle, brightness, sea state variabilities, presence of obstructing objects, people etc.). In recent years, the development of convolutional neural network methods, including Mask R-CNN to process instance segmentations, made it possible to automatically extract detailed information from images and identify complex and abstract concepts in a wide variety of contexts.

The methodology presented in this research proposes the use of Mask R-CNN as a tool to classify bodies of water, with the aim of extracting the waterline from a large dataset of VMS images. This approach provides a robust technique to automatically identify the waterline of the Normandy beaches on RGB images provided by VMS, as Mask R-CNN is able to perform segmentation over 67% to 97% of the provided datasets. The waterlines can then be georeferenced with low uncertainty estimates, with vertical RMSEs of 33 cm, 29 cm and 22 cm at Villers-sur-Mer, Hautot-sur-Mer and Etretat respectively. Some limitations remain on sandy and gently sloped areas such as Villers-sur-Mer and Hautot-sur-Mer due to the very specific conditions of the Normandy sandy beaches (complex morphology, low slope, very wide intertidal zone, and unclear water-land delimitation). These conditions tend to increase the cumulative uncertainty of both segmentation and georeferencing as they make it a complicated task to identify the waterline, even for the human eye. Despite these limitations, this methodology was successfully and satisfactory deployed at the three monitored

study sites of this research: Etretat, Villers-sur-Mer and Hautot-sur-Mer.

In addition to being freely available, Mask R-CNN applications can be extended to a wide spectrum of scientific questions due to its adaptability and its abstraction abilities and can therefore be a very interesting tool for coastal science. The recent major advances made in the field of instance segmentation could extensively open the possibilities for the coastal science studies.

ACKNOWLEDGMENT

The authors wish to thank the cities of Villers-sur-Mer, Etretat and Hautot-sur-Mer, as well as the Seine-Maritime Departmental Council and the Normandy Region for supporting this study.

Additionally, the authors would like to thank Jacuelyn Steele who helped and verified the English writing of this work.

REFERENCES

Aarninkhof, S.G.J., Turner, I.L., Dronkers, T.D.T., Caljouw, M., Nipius, L., 2003. A video-based technique for mapping intertidal beach bathymetry. Coast. Eng. 49, 275–289. https://doi.org/10.1016/S0378-3839(03)00064-4

Bonnefille, R., 1968. Contribution théorique et expérimentale à l'étude du régime des marées. Université de Grenoble.

Chabert D'Hières, G., Le Provost, C., 1978. Atlas des composantes harmoniques de la marée dans la Manche. Ann. Hydrogr. 6.

Chen, H., Sun, K., Tian, Z., Shen, C., Huang, Y., Yan, Y., 2020. BlendMask: Top-down meets bottom-up for instance segmentation, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 8573–8581.

Chen, X., Girshick, R., He, K., Dollár, P., 2019. Tensormask: A foundation for dense object segmentation, in: Proceedings of the IEEE International Conference on Computer Vision. pp. 2061–2069.

Crété-Roffet, F., Dolmiere, T., Ladret, P., Nicolas, M., 2007. The Blur Effect: Perception and Estimation with a New No-Reference Perceptual Blur Metric, in: SPIE Electronic Imaging Symposium Conf Human Vision and Electronic Imaging. San Jose, United States, p. EI 6492--16.

Davidson, M.A., Turner, I.L., Splinter, K.D., Harley, M.D., 2017. Annual prediction of shoreline erosion and subsequent recovery. Coast. Eng. 130, 14–25. https://doi.org/10.1016/j.coastaleng.2017.09.008

Davies, J., 1964. A morphogenic approach to world shorelines. Zeitschrift fur Geomorphol. 8, 127–142.

Dolan, R., Hayden, B.P., May, P., May, S., 1980. The reliability of shoreline change measurements from aerial photographs. Shore and beach, 48, 22–29.

Esmail, M., Mahmod, W.E., Fath, H., 2019. Assessment and prediction of shoreline change using multi-temporal satellite images and statistics: Case study of Damietta coast, Egypt. Appl. Ocean Res. 82, 274–282. https://doi.org/10.1016/j.apor.2018.11.009

Hanson, H., 1989. GENESIS - a generalized shoreline change numerical model. J. Coast. Res. 5, 1–27.

He, K., Gkioxari, G., Dollar, P., Girshick, R., 2017. Mask R-CNN. Proc. IEEE Int. Conf. Comput. Vis. 2017-Octob, 2980–2988. https://doi.org/10.1109/ICCV.2017.322

Holman, R.A., Lippmann, T.C., O'Neill, P. V., Hathaway, K., 1991. Video estimation of subaerial beach profiles. Mar. Geol. 97, 225–231. https://doi.org/10.1016/0025-3227(91)90028-3

Huang, Z., Huang, L., Gong, Y., Huang, C., Wang, X., 2019. Mask scoring r-cnn, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 6409–6418.

Jara, M.S., González, M., Medina, R., 2015. Shoreline evolution model from a dynamic equilibrium beach profile. Coast. Eng. 99, 1–14. https://doi.org/10.1016/j.coastaleng.2015.02.006

Jongejan, R., Ranasinghe, R., Wainwright, D., Callaghan, D.P., Reyns, J., 2016. Drawing the line on coastline recession risk. Ocean Coast. Manag. 122, 87–94. https://doi.org/10.1016/j.ocecoaman.2016.01.006

Kingston, K.S., 2003. Applications of complex adaptive systems approaches to coastal systems. University of Plymouth.

Kirillov, A., Wu, Y., He, K., Girshick, R., 2020. Pointrend: Image segmentation as rendering, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 9799–9808.

Kopp, M., Tuo, Y., Disse, M., 2019. Fully automated snow depth measurements from time-lapse images applying a convolutional neural network. Sci. Total Environ. 697, 134213. https://doi.org/10.1016/j.scitotenv.2019.134213

Le Cozannet, G., Bulteau, T., Castelle, B., Ranasinghe, R., Wöppelmann, G., Rohmer, J., Bernon, N., Idier, D., Louisor, J., Salas-y-Mélia, D., 2019. Quantifying uncertainties of sandy shoreline change projections as sea level rises. Sci. Rep. 9, 1–11. https://doi.org/10.1038/s41598-018-37017-4

Levoy, F., Anthony, E.J., Monfort, O., Larsonneur, C., 2000. The morphodynamics of megatidal beaches in Normandy, France. Mar. Geol. 171, 39–59. https://doi.org/10.1016/S0025-3227(00)00110-9

Long, J., Shelhamer, E., Darrell, T., 2015. Fully Convolutional Networks for Semantic Segmentation. 2015 IEEE Conf. Comput. Vis. Pattern Recognit. 10. https://doi.org/10.1109/CVPR.2015.7298965

Maxwell, A.E., Pourmohammadi, P., Poyner, J.D., 2020. Mapping the topographic features of mining-related valley fills using mask R-CNN deep learning and digital elevation data. Remote Sens. https://doi.org/10.3390/rs12030547

Montaño, J., Coco, G., Antolínez, J.A.A., Beuzen, T., Bryan, K.R., Cagigal, L., Castelle, B., Davidson, M.A., Goldstein, E.B., Ibaceta, R., Idier, D., Ludka, B.C., Masoud-Ansari, S., Méndez, F.J., Murray, A.B., Plant, N.G., Ratliff, K.M., Robinet, A., Rueda, A., Sénéchal, N., Simmons, J.A., Splinter, K.D., Stephens, S., Townend, I., Vitousek, S., Vos, K., 2020. Blind testing of shoreline evolution models. Sci. Rep. 10, 1–10. https://doi.org/10.1038/s41598-020-59018-y

Nicholls, R.J., Wong, P.P., Burkett, V.R., Codignotto, J.O., Hay, J.E., McLean, R.F., Ragoonaden, S., Woodroffe, C.D., 2007. Coastal systems and low-lying areas, Cambridge. ed, Climate Change 2007: Impacts, Adaptation and Vulnerability. Contribution ofWorking Group II to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge, UK.

Nie, S., Jiang, Z., Zhang, H., Cai, B., Yao, Y., 2018. Inshore ship detection based on mask r-cnn. Int. Geosci. Remote Sens. Symp. 2018-July, 693–696. https://doi.org/10.1109/IGARSS.2018.8519123

Osorio, A.F., Medina, R., Gonzalez, M., 2012. An algorithm for the measurement of shoreline and intertidal beach profiles using video imagery: PSDM. Comput. Geosci. 46, 196–207. https://doi.org/10.1016/j.cageo.2011.12.008

Plant, N.G., Aarninkhof, S.G.J., Turner, I.L., Kingston, K.S., 2007. The performance of shoreline detection models applied to video imagery. J. Coast. Res. 23, 658–670. https://doi.org/10.2112/1551-5036(2007)23[658:TPOSDM]2.0.CO;2

Plant, N.G., Holman, R.A., 1997. Intertidal beach profile estimation using video images. Mar. Geol. 140, 1–24. https://doi.org/10.1016/S0025-3227(97)00019-4

Qiao, Y., Truman, M., Sukkarieh, S., 2019. Cattle segmentation and contour extraction based on Mask R-CNN for precision livestock farming. Comput. Electron. Agric. 165, 104958. https://doi.org/10.1016/j.compag.2019.104958

Ranasinghe, R., Callaghan, D., Stive, M.J.F., 2012. Estimating coastal recession due to sea level rise: Beyond the Bruun rule. Clim. Change 110, 561–574. https://doi.org/10.1007/s10584-011-0107-8

Ren, S., He, K., Girshick, R., Sun, J., 2016. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. arXiv1506.01497 [cs].

SHOM, 1953. Les courants de marée dans la mer de la Manche et sur les côtes françaises de l'Atlantique. Service Hydrographique et Océanographique de la Marine, Paris.

Soloy, A., Turki, I., Fournier, M., Costa, S., Peuziat, B., Lecoq, N., 2020. A deep learning-based method for quantifying and mapping the grain size on pebble beaches. Remote Sens. https://doi.org/10.3390/rs12213659

Turner, I., Leyden, V., Symonds, G., McGRATH, J., Jackson, A., Jancar, T., Aarninkhof, S.G.J., Elshoff, I.E., 2001. Comparison of observed and predicted coastline

changes at the gold coast artificial (surfing) reef, Sydney, Australia, in: Proceedings of the International Conference on Coastal Engineering.

Uunk, L., Wijnberg, K.M., Morelissen, R., 2010. Automated mapping of the intertidal beach bathymetry from video images. Coast. Eng. 57, 461–469. https://doi.org/10.1016/j.coastaleng.2009.12.002

Valentini, N., Saponieri, A., Molfetta, M.G., Damiani, L., 2017. New algorithms for shoreline monitoring from coastal video systems. Earth Sci. Informatics 10, 495–506. https://doi.org/10.1007/s12145-017-0302-x

Wainwright, D.J., Roshanka, R., Callaghan, D.P., Woodroffe, C.D., Jongejan, R., Dougherty, A.J., Rogers, K., Cowell, P.J., 2015. Moving from deterministic towards probabilistic coastal hazard and risk assessment : development of a modelling framework and application to. Coast. Eng. 92–99.

Wang, X., Zhang, R., Kong, T., Li, L., Shen, C., 2020. SOLOv2: Dynamic, Faster and Stronger. arXiv Prepr. arXiv2003.10152.

Yang, F., Feng, T., Xu, G., Chen, Y., 2020. R-CNN, Applied method for water-body segmentation based on mask. J. Appl. Remote Sens. 14, 014502. https://doi.org/https://doi.org/10.1117/1.JRS.14.014502

Yates, M.L., Guza, R.T., O'Reilly, W.C., 2009. Equilibrium shoreline response: Observations and modeling. J. Geophys. Res. Ocean. 114, 1–16. https://doi.org/10.1029/2009JC005359

Zhang, Y., Chu, J., Leng, L., Miao, J., 2020. Mask-Refined R-CNN : A Network for Refining. Sensors 20. https://doi.org/https://doi.org/10.3390/s20041010

Zhao, K., Kang, J., Jung, J., Sohn, G., Street, K., Drive, M., York, N., Mb, O.N., 2018. Building Extraction from Satellite Images Using Mask R-CNN with Building Boundary Regularization. CVPR Work. 247–251.

Zhao, S., Zhang, D.M., Huang, H.W., 2020. Deep learning–based image instance segmentation for moisture marks of shield tunnel lining. Tunn. Undergr. Sp. Technol. 95, 103156. https://doi.org/10.1016/j.tust.2019.103156