Enhancing the predictive performance of remote sensing for ecological variables of tidal flats using encoded features from a deep learning model

Logambal Madhuanand, Catharina J. M. Philippart, Jiong Wang, Wiebe Nijland, Steven M. de Jong, Allert I. Bijleveld & Elisabeth A. Addink

View supplementary material

Published online: 03 Jan 2023.

Submit your article to this journal

Article views: 254

View related articles

View Crossmark data

Taylor & Francis
Taylor & Francis Group

Enhancing the predictive performance of remote sensing for ecological variables of tidal flats using encoded features from a deep learning model

Logambal Madhuanand [a], Catharina J. M. Philippart [a,b], Jiong Wang[a,c], Wiebe Nijland [a], Steven M. de Jong [a], Allert I. Bijleveld [b] and Elisabeth A. Addink [a]

aDepartment of Physical Geography, Utrecht University, Utrecht, Netherlands; bDepartment of Coastal Systems, NIOZ Royal Netherlands Institute for Sea Research, Texel, Netherlands; cGeo-Information Science and Earth Observation (ITC), University of Twente, Enschede, Netherlands

ABSTRACT

Tidal flats are among the ecologically richest areas of the world where sediment composition (e.g. median grain size and silt content) and the macrozoobenthic presence play an important role in the health of the ecosystem. Regular monitoring of environmental and ecological variables is essential for sustainable management of the area. While monitoring based on field sampling is very time-consuming, the predictive performance of these variables using satellite images is low due to the spectral homogeneity over these regions. We tested a novel approach that uses features from a variational autoencoder (VAE) model to enhance the predictive performance of remote sensing images for environmental and ecological variables of tidal flats. The model was trained using the Sentinel-2 spectral bands to reproduce the input images, and during this process, the VAE model represents important information on the tidal flats within its layer structure. The information in the layers of the trained model was extracted to form features with identical spatial coverage to the spectral bands. The features and the spectral bands together form the input to random forest models to predict field observations of the sediment characteristics such as median grain size and silt content, as well as the macrozoobenthic biomass and species richness. The maximum prediction accuracy of feature-based maps was close to 62% for the sediment characteristics and 37% for benthic fauna indices. The encoded features improved the prediction accuracy of the random forest regressor model by 15% points on average in comparison to using just the spectral bands. Our method enhances the predictive performance of remote sensing, in particular the spatiotemporal dynamics in median grain size and silt content of the sediment thereby contributing to better-informed management of coastal ecosystems.

1 Introduction

Tidal flats are marine ecosystems characterized by high biodiversity and productivity across the world. They host a wide variety of benthic invertebrates and provide essential foraging areas for waterbirds (Beukema 1976; Bakker et al. 2021; Piersma et al. 1993). Globally, tidal flats are under pressure due to coastal developmental activities such as harbor creation and shoreline stabilization, sea level rise, and erosion (Murray et al. 2019). Effective management of biodiversity, addressing anthropogenic pressures, and climate change impacts require sustained efforts to monitor the ecosystem (Miloslavich et al. 2018). Changes in water levels, sediment composition, and benthic community composition and biomass can have a significant impact on the presence of estuarine birds that forage on tidal

flats (Piersma et al. 1993; Zwarts and Wanink 1993; Bijleveld et al. 2016; Park et al. 2014). Thus, regular monitoring of environmental and ecological variables is essential for conserving the tidal flat ecosystem.

Optical remote sensing is an important tool for monitoring the dynamics of tidal flats due to its wide coverage and synoptic view, even in inaccessible areas. Several studies have demonstrated the use of satellite images for mapping ecological indicators in tidal systems, with or without extensive field datasets (Adolph et al. 2018). Bartholdy and Folving (1986) used a multi-temporal classification approach to map the sediment characteristics of the Danish Wadden Sea with Landsat images. Images from Landsat 5 TM, Landsat 7 ETM+, Landsat 8 OLI or Airborne Thematic Mapper (ATM) were utilized to

map median grain size (60–250 μm) with approaches like maximum likelihood, multiple regression, and principal component analysis (PCA). These produced good results but only for a limited range of grain sizes (Yates et al. 1993; Rainey et al. 2000; Ryu et al. 2004). Van der Wal, Herman, and Ysebaert (2004) explored the potential of synthetic aperture radar (SAR) for habitat mapping and found correlations between sediment characteristics and macrofauna density. This was followed by research by Van der Wal et al. (2008) that evaluated different response models to predict the spatial dynamics of benthic species using hyperspectral images, altimetry information, and field data of sediment grain sizes. Ryu, Kuk Choi, and Kyung Lee (2014) proposed an integrated approach with high spatial, temporal resolution, multispectral, hyperspectral, and SAR data to monitor topography, sediment, and bio facies of tidal flats in Korea. Hedley et al. (2018) discussed the use of image-derived attributes (bathymetry, slope, reflectance from water surface and bottom) that can contribute to the discriminatory ability and repeatability while mapping benthic features from Sentinel-2 images. They concluded that Sentinel-2 provided a new level of monitoring information (in terms of spatial, spectral, and temporal resolution) of these features compared to previously available platforms. Most of the studies discussed above concluded that information required to study the tidal flats was not presented in optical imagery due to its lack of outspoken spectral variations. The low spectral contrast on tidal flats presents a substantial challenge to optical remote sensing applications in this environment.

Machine learning models that automatically learn patterns within the data have gained popularity in ecology due to the advancements in data-driven modeling (Willcock et al. 2018) and their efficiency in describing complex relationships often with nonlinear data (Olden, Lawler, and Leroy Poff 2008). Lee et al. (2013) used a simple neural network to evaluate the mapping of macrobenthos habitat through a supervised approach using eight control factors acquired with field survey. They showed an average accuracy of 70% between the probability maps and species locations. However, acquiring a spatial database of all the control factors on a large scale would be practically not feasible and hence more advanced models are required for large-scale monitoring. Convolutional neural network (CNN), a subset of

deep neural networks (DNN), has outperformed several machine learning algorithms and statistical models, and are capable of performing complicated tasks such as image classification and segmentation in both supervised and unsupervised manners (Sehgal 2012; Bhandare et al. 2016; Zhang, Wang, and Liu 2019; Madhuanand, Nex, and Ying Yang 2021; Guo et al. 2018). They are composed of an encoder/decoder architecture where the information from input images is compressed through encoder layers and then gets up-sampled by decoder layer to create output layers, which match the input extent (Goodfellow, Bengio, and Courville 2016). Deep learning is a rapidly evolving area that could be applied to improve many ecological applications (Rammer and Seidl 2019). Most studies that use deep learning models, follow a straightforward train-test approach where the model is trained to predict a particular phenomenon. However, the potential of the learned representations from the hidden layers of CNNs has been barely explored.

Zeiler and Fergus (2014) visualized the encoded features from CNNs and revealed that these hidden features, far from random, showed many intuitively desirable properties such as compositionality, increasing invariance (object remain unchanged after certain operations), and class discrimination. Nguyen, Yosinski, and Clune (2016) added that the CNNs learn the global structure of images and general patterns found within them from these features. A particular type of framework that learns efficient coding from the input data are autoencoders where CNN is used for encoding and decoding parts (Baldi 2012).

Autoencoders (Figure 1) are unsupervised networks that are trained to reconstruct the input with minimal distortion (Baldi 2012). The model learns to reconstruct the input images by minimizing the difference between the original and reconstructed image in the manner of loss-term optimization. This process of reconstructing the output image to be identical to the input image, enables the network to learn discriminative features that contain accurate spatial and textural information describing the input image. These features could then act as additional information for remote sensing tasks.

This study aimed to determine the added value of the encoded features of Sentinel-2 images extracted from a deep learning autoencoder model in the
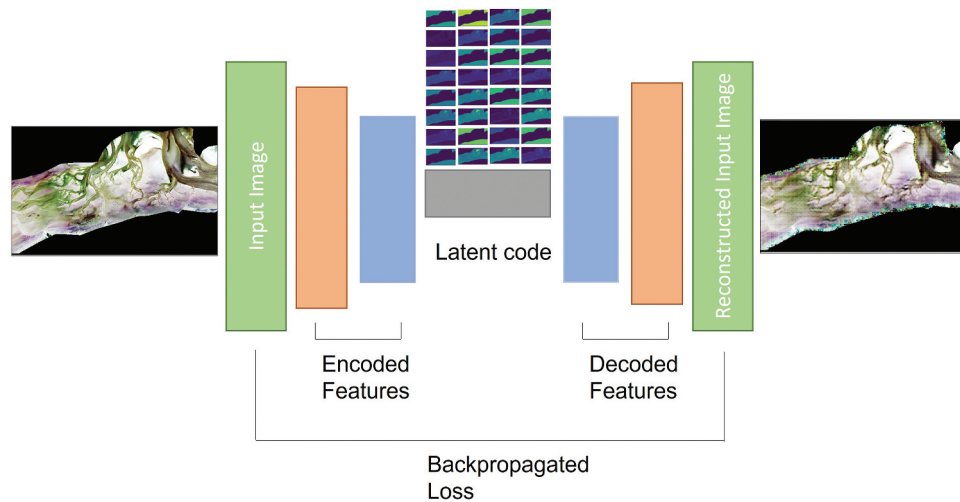
Figure 1. Autoencoder model – with input and reconstructed input as output, tidal flats of Pinkegat site, Dutch Wadden Sea.

prediction of environmental and ecological variables of the tidal flats in comparison to spectral bands. We hypothesized that context and textural information extracted from the autoencoder model captures the spatial expression of differing environmental conditions which are not well represented in local spectral information and thus elevates mapping performance. We used random forest to predict sediment characteristics: median grain size and silt content, and macrozoobenthic biomass and species richness.

2 Materials & methods

2.1 Study area

The Wadden Sea is the largest tidal-flat system in the world and extends along the coasts of the Netherlands, Germany and Denmark (Kloepper, Strempel, and Bostelmann 2017). It is a protected site under Natura 2000 and has been designated as a UNESCO World Heritage site and has been given Ramsar site status. It is a dynamic tidal system of high ecological significance providing wide-ranging ecosystem services like shrimp fisheries, mussel culture, and tourism (Beukema 1976; Boere and Piersma 2012; Levin et al. 2001; Marencic 2009; Philippart et al. 2007).

The study site contains two adjacent tidal basins in the Dutch Wadden Sea, Pinkegat (53°25N, 5°46E) and Zoutkamperlaag (53°26N, 6°15E), with a surface area of 170 km^2 and 150 km^2 respectively (Figure 2). The

diversity of macrozoobenthic species (Drent et al. 2017) and the availability of long term detailed field data (Compton et al. 2013) makes this area an ideal study site.

The intertidal zone spans almost 8 km along the north-south axis and 40 km along the east-west axis, covering a total area of about 320 km^2. The sediment composition shows a gradient from muddy to sandy from south to north. The tides in the area are semi-diurnal with a mean tidal range of 2.3 m with summer tides 15 cm higher than winter tides (Hollebrandse 2005). The tide along with the input of fresh water from the discharge sluices (in the southern part of the study site) has a strong effect on the nutrient supply and the organisms present (Compton et al. 2013). The region encompasses a wide variety of environmental gradients with many transitional zones, providing habitats for various ecosystems.

2.2 Data

2.2.1 Satellite images
Freely available Sentinel-2 (L1C) satellite imagery was used in this study. Sentinel-2 imagery comes with 13 spectral bands (visible to short-wave infra-red) varying spatial (10 m and 60 m) ("User Guides - Sentinel-2 MSI 2015). Sentinel-2 captures the study site between 11:30 and 12:30 CET. All images that are cloud free and have exposed tidal flats were selected and downloaded through USGS Earth Explorer for the period 2018 to 2020. The study
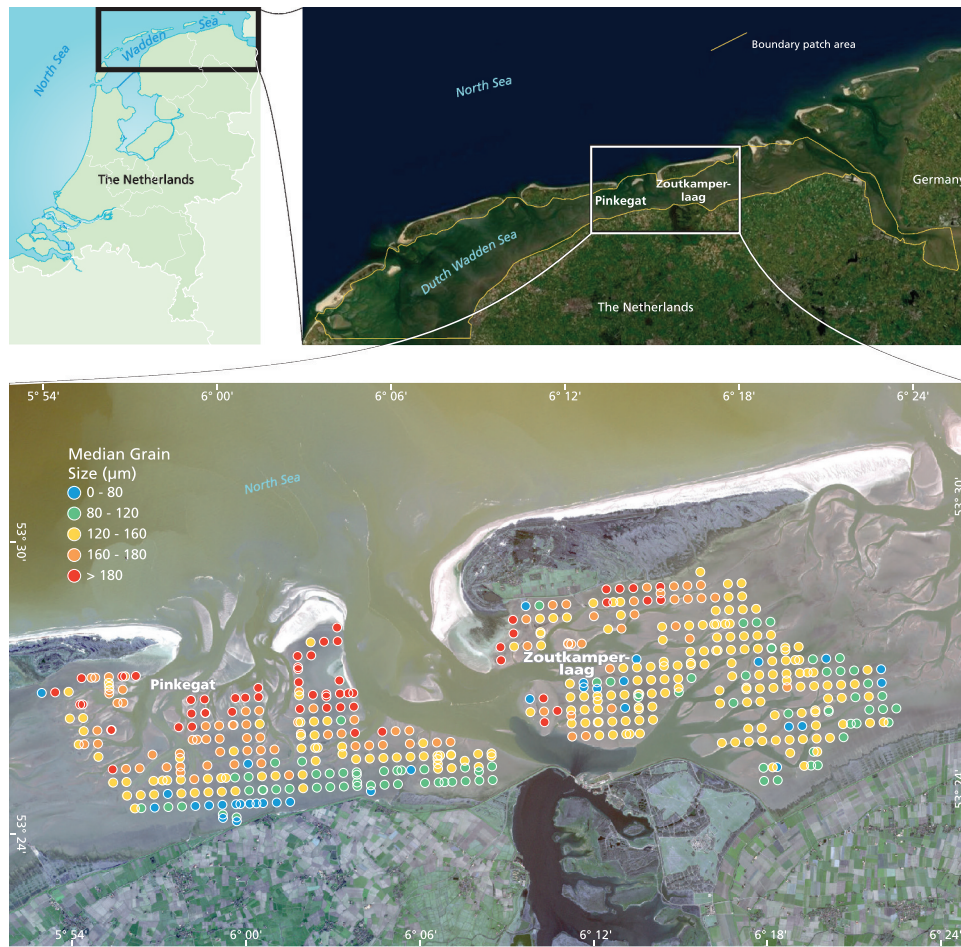
Figure 2. (Top) Study sites, Pinkegat and Zoutkamperlaag in the Dutch Wadden Sea, the Netherlands. The inset shows the Netherlands, and the yellow line indicates the outline of regions that were used to create image patches for the training dataset. (Bottom) Median grain size from Synoptic Intertidal Benthic Survey (SIBES) data for 2018. Source background image: Sentinel-2, 2018.

sites, Pinkegat and Zoutkamperlaag, were covered by two Sentinel-2 tiles (L1C-T31UGV and L1C-T31UFV). The tides with respect to mean sea level at the time of image acquisition were obtained from the Dutch Ministry of Infrastructure and Water Management (Rijkswaterstaat n.d.) (Table 1).

2.2.2 Field data
Since 2008, the Royal Netherlands Institute of Sea Research (NIOZ) has been conducting extensive field sampling through the Synoptic Intertidal Benthic Survey (SIBES) program, a long-term ecological monitoring program to study the intertidal macrofauna of the Dutch Wadden Sea (Bijleveld et al. 2012). The data is collected with an inter-sampling distance of 500 m corresponding to approximately 4,700 sampling locations covering the entire Dutch Wadden Sea. In addition to the grid points sampling, 20%

random points are included to allow estimating autocorrelation parameters (Bijleveld et al. 2012). Within both tidal basins, 200 or more samples are taken each year (Table 2). For a detailed description about the techniques used by SIBES to collect the data, see (Compton et al. 2013).

In our study, we focused on two sediment characteristics: median grain size and silt content and two macrozoobenthic variables: biomass and species richness. Based on the field data for the period from 2018 to 2020, the median grain size in the study region ranged between 20 μm and 231 μm with a mean of 145 μm (Figure 2, Table 2). The percentage of silt content by volume ranged between 1.5% and 82% with a mean of 13%. The total biomass (ash-free dry mass (AFDM)) for all species ranged from 0.05 to 550 gAFDM/m^2 with a mean of 30 gAFDM/m^2 and the species richness (number of benthic species per

Table 1. Satellite images with the corresponding tide levels, *
represents the images selected to be used for prediction.

Date of image acquisition	Tide level from NAP datum (in m)
Pinkegat	
2018/05/26	−0.21
2018/11/17*	−0.73
2019/02/15*	−0.75
2019/04/01	−0.33
2019/04/16	−0.44
2020/04/05*	−0.26
2020/04/20	0.15
Zoutkamperlaag	
2018/01/08	−0.46
2018/05/08	−0.14
2018/05/23	−0.84
2018/06/07	−0.34
2018/08/06*	−0.79
2019/02/15	−0.90
2019/02/27	−0.99
2019/08/24	−0.10
2019/08/26*	−0.94
2020/04/05*	−0.41

sample) varied between 1 and 21 species with a mean of 8. For convenience, we have only chosen the stations with macrozoobenthos presence for calculation of biomass and species richness.

2.3 Method

In this study, we trained a CNN model with Sentinel-2 images of the Dutch Wadden Sea to learn the representative information. The encoded features were tested for their ability to enhance the predictability of random forest models for median grain size, silt content, biomass, and species richness for the spectrally poor tidal regions in the Dutch Wadden Sea. The random forest models were then used to create full coverage maps of the four variables. The steps involved in the process include data preparation, training the CNN model, and feature extraction followed by the prediction of the variables (Figure 3).

2.3.1 Data preparation and processing

The input images for training the model were prepared from the Sentinel-2 tiles that cover the tidal flats of Pinkegat and Zoutkamperlaag for the years 2018, 2019, and 2020. The Sentinel-2 images were stacked to form a single image with four bands (B, G, R, and NIR) of 10 m resolution. The selection of bands was chosen after several experiments as described in section 2.5.1. Only areas with exposed intertidal flats were used in this study, and pixels outside this area (on land or permanent sea) were

discarded (Figure 2). The images were further normalized as per the recommendation of the Sentinel-2 data reference guide ("User Guides - Sentinel-2 MSI 2015). We created patches sized 64×64 pixels (640 m $\times$ 640 m) with an overlap of 10% between the patches and repeated this for all the images. In total, we extracted over 38,000 patches from which 80% were used for training (31,154 patches) and 20% were used for validation (7,788 patches).

2.3.2 Autoencoder model

Unlike supervised deep learning models, the autoencoder model does not use separate ground truth data for learning. A regularized version of the autoencoder model is the variational autoencoder (VAE) model (Kingma and Welling 2014), which learns the statistical distribution of the input data to make the generative process plausible using the latent code parameters. This is achieved through two loss functions (backpropagated loss, Figure A1), regularization, and reconstruction loss (Kingma and Welling 2014) as given in Equation (1).

$$L = \lambda_1 L_m + \lambda_2 L_{kl} \qquad (1)$$

where λ_1, λ_2 represent the weights between the two loss terms used. L_m, represents the mean squared error loss (MSE), L_{kl} represents the Kullback – Leibler (KL) divergence loss (Joyce 2011). The regularization loss is KL divergence loss, which is efficient for learning the underlying parameters of the distribution of the training data in VAE models (Equation (2)).

Table 2. Details of field sampling campaigns along with descriptive statistics of the four variables and the number of points collected at each site over different years.

| Region	Field sampling dates	Median grain size (µm)			Silt content (%)			Biomass (gAFDM/m²)			Species richness (no. of species/sample)			No. of points
		Mean	Std	Min-Max	Mean	Std	Min-Max	Mean	Std	Min-Max	Mean	Std	Min-Max	
Pinkegat	2018/08/01 – 2018/08/07	147.5	37.8	42-231	13.3	13.35	1.5-60	37.8	51.9	0.02-314.2	9.07	3.9	1-21	220
	2019/08/07 – 2019/08/31	144.4	39.9	21-222	14.5	15.09	1.5-78.1	27.47	55.66	0.005-684.1	7.54	3.24	1-16	224
	2020/06/21 – 2020/06/22	146.9	44.75	20.6-222.6	13.03	15.38	1.55-82	17.84	25.84	0.02-280	8.6	4.08	1-19	235
Zoutkamperlaag	2018/08/02 – 2018/08/07	140.0	30.65	32.4-211	15.73	13.26	1.81-65.5	30.79	49.57	0.014-482	9.18	3.56	2-18	235
	2019/08/08 – 2019/08/11	137.92	32.58	26.2-219.8	16.8	1.8	1.25-73.5	20.26	42.62	0.011-297.4	7.58	3.07	1-18	238
	2020/06/23 – 2020/06/24	143.4	29.6	44.5-217.4	14.78	11.1	1.76-57.9	24.9	56.8	0.103-546.5	8.36	3.15	1-17	236
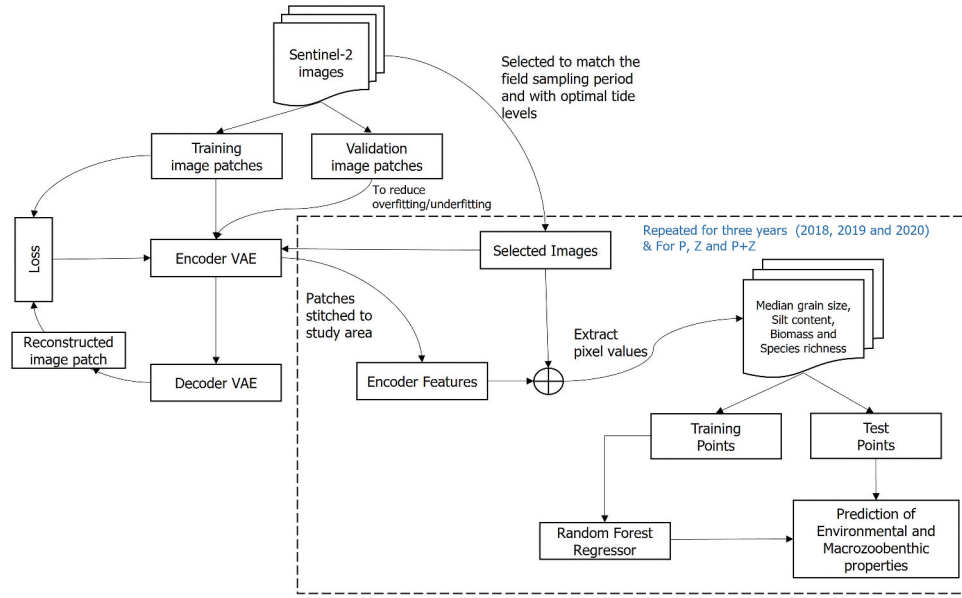
Figure 3. Stepwise procedure for data processing, feature extraction, and prediction of the environmental and macrozoobenthic properties. The dotted line delineates the processes which were repeated for different scenarios separately: Pinkegat (P), Zoutkamperlaag (Z), and combined sites (P+Z).

$$L_{kl} = \sum_{i=1}^{n} \sigma_i^2 + \mu_i^2 - \log \sigma_i - 1 \qquad (2)$$

where σ_i and μ_i represent the variance and mean of the latent variables, n represents the total number of pixels in the image patch. The reconstruction loss is calculated from the MSE that optimizes the model performance by comparing the pixel differences of the input and the reconstructed image to improve the reconstruction quality.

2.3.2.1 Implementation

The architecture of the VAE model includes three important components, i.e. encoder, bottle neck and decoder. The model takes the image patches as input. These patches pass through the encoder with multiple skip connections that skips certain layers to extract both low-level and high-level features from

the ResNet-50 (Residual Networks) (Kaiming et al. 2016) block structures (Figure A3). This was followed by a series of up-sampling layers with rectified linear unit (ReLU) activation (Agarap 2018) to introduce non-linearity (Goodfellow, Bengio, and Courville 2016) in each layer. The output from the last layer passes through the sigmoid activation function and an output image with the same dimensions as that of the input patch is formed. The loss terms calculated based on input and output images were backpropagated through the model to enable the learning process. The weights for the two loss terms were tuned and the respective weights were determined after experiments (Table 3).

We experimented with three different network architectures for the encoder layers in our VAE model. These were ResNet-50, DenseNet-121

Table 3. Hyperparameters used in experiments to tune the deep learning model along with the selected hyperparameter values applied for the further parts of the study.

Hyperparameters	Experiments	Optimal Values
Encoder architecture	ResNet-50, DenseNet-121, MobileNet	ResNet-50
Batch size	8, 16, 20, 32	20
Learning rate	10^{-3}, 10^{-4}, 10^{-5}, 10^{-6}	10^{-4}
Patch size	32×32, 64×64, 128×128	64×64
Epochs	500, 1000	500
Spectral band	B+G+R+NIR	B+G+R+NIR
	B+G+R+NIR+SWIR	
	B+G+R+NIR+RedEdge+SWIR	
Weights between losses λ_1, λ_2	1, 0.1, 0.001, 0.0001	1, 0.001
Optimizer	Adam	
GPU	Nvidia Quadro P4000 of 8GB memory	

(Gao et al. 2017) and MobileNet (Howard et al. 2017). All models in this work were implemented in PyTorch (Paszke et al. 2017) and were fine-tuned to achieve optimal hyperparameters (Table 3).

The ResNet-50 model converged fastest (17 h) with minimal loss value for both training and validation datasets (Figure A1 & A2). Also, to study the impact of various combinations of spectral bands, we experimented by stacking different combination of bands (Table 3) and found that there was no appreciable improvement in the reconstruction process with the use of more than four bands. In addition, the four bands B, G, R and NIR are commonly available in most optical remote sensing platforms with different spatial specifications (Planet images, UAV images) and limiting to just the four bands will facilitate the use of our model with other platforms. Based on our experiments, we used the ResNet-50 VAE model with four spectral bands of patch size 64×64 that was trained for 500 epochs with a learning rate of 10^{-4} and a batch size of 20 for the feature extraction task (Table 3).

2.3.3 Feature extraction

The various blocks in the encoder (Figure A3) contain an increasing number of features of smaller dimensions when moving from the input image toward the bottleneck structure. After training the model, the information content of each block optimally describes the input image. To retrieve the features, one block would be up-sampled through bilinear interpolation to 64×64 pixels. The first layer with 64 features of dimension 32×32 pixels each was selected here for the feature extraction process for two reasons. First, the deeper layers have more features, which increases the computational demand for the predictive model. Secondly, the smaller dimensions of the features in the deeper layers may introduce more uncertainties when up-sampled to match the input patch size. The up-sampling process amplified the border effect in the individual patches and to overcome this, we used a 30% overlap while dividing the image into patches for testing.

2.3.4 Random forest regression

We used a random forest regression model (Breiman 2001) for the prediction of the four selected variables. A random forest regression model fits many decision trees (ensemble learning) over sub-samples of data and of variables and predicts by taking the average of the output of all trees. The random forest model was chosen because of its ability to reliably fit models with many input variables even when nonlinear relations or collinearity are present in the data. The field samples for both Pinkegat and Zoutkamperlaag for the four variables were split into training and testing data points for each year. For silt content and biomass, a common logarithmic transformation was applied, as the values were skewed and of a very large range. To determine the added value of the features extracted from the VAE model, we predicted the four variables in two rounds. In the first round, we trained the random forest model with the four spectral bands and in the second round we trained the model with the four spectral bands, encoded features from VAE and the field data. In both rounds, predictions were made for the two study sites combined and individually, for each of the three years (Figure 3). Also, for the full images of the study site, predictions of the four variables were made to explore if the predicted values and spatial patterns were plausible.

The random forest regressor was tuned for hyperparameters using cross-validation technique (Jung and Jianhua 2015) with 90% data for training and 10% for validation. For the model, the number of trees was set to 800, the minimum number of data points at each node was set to ten and the maximum number of features was taken as the square root of number of input variables. We implemented a k-fold cross-validation strategy that optimizes the number of observations when fitting the trees and can also lead to a more stable estimation of the model results. Here, we used 10-fold which was repeated three times to further reduce the variance in the model performance. All 30 folds were averaged to estimate the model accuracy. The accuracy of the point predictions was evaluated using R^2 (coefficient of determination) value (Chicco, Warrens, and Jurman 2021). In addition to R^2, we also estimated the relative importance of the features through the Gini coefficient of importance, which is used as a general indicator of feature relevance (Menze et al. 2009).

3 Results

3.1 Feature extraction

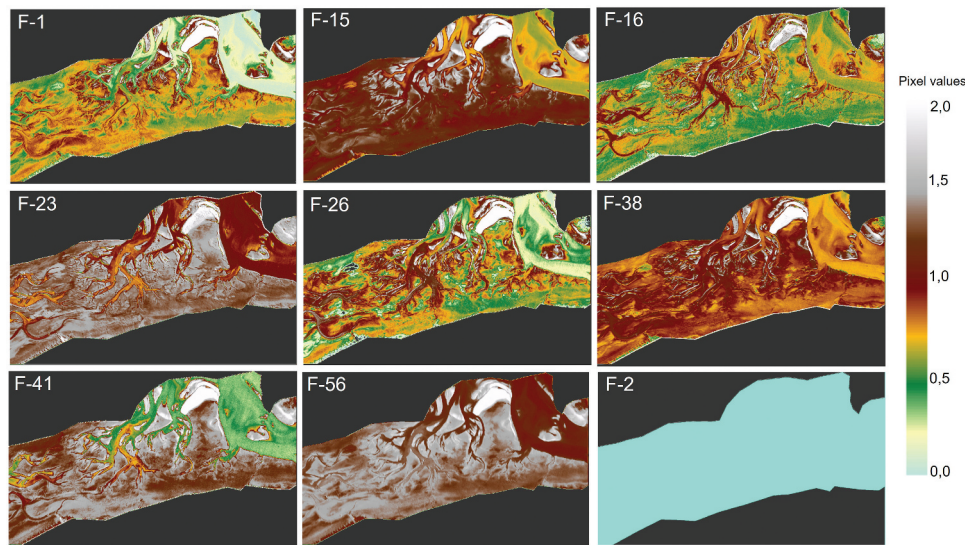The VAE model with ResNet-50 architecture converged after 500 epochs for the training dataset with

Figure 4. Examples of the encoded features for Pinkegat, derived from the Sentinel-2 image captured on 17 November 2018. The feature numbers are specified in each feature image. The background is shown gray for all the features to differentiate it from the tidal basin.

32,000 patches. For the first 100 epochs, a steep decrease in loss values was observed, followed by a gradual decrease after 250 epochs, and ending with saturation after 500 epochs. The VAE model was tested with selected images (Table 1) from 2018, 2019 and 2020 for the Pinkegat and Zoutkamperlaag regions. The average Root Mean Square Error between the reconstructed and original test image was maintained around 0.016.

For each of the test image patches, we extracted the features from the first layer of the encoder. The mean value of the 64 features ranges from 0.11 to 0.47 with the standard deviation ranging from 0.15 to 0.45. Around 30 features, e.g. Feature (F)-2 (Figure 4), were entirely made up of zeros in the tidal flat regions. The correlation values between the non-zero features ranged between −0.08 (F-18 & F-22) and 0.9 (F-15 & F-56) while the correlation values between the four spectral bands ranged from 0.998 to 0.999. By visualizing all non-zero features for a single patch, we observed that there was spatial variation within the patches that created clustering patterns with the spectral information (Figure 5).

3.2 Prediction model

The spectral bands of Sentinel-2 images (B, G, R, and NIR) served as input together with and without the 64 encoded features to the random forest regressor model.

The predictions, which included autoencoder features outperformed those with only the spectral bands for every variable, area, and year (Figure 6; Table 4). The median grain size predictions ranged from 25.5% to 61.7%; silt content predictions ranged from 32% to 59.5%; biomass predictions ranged from 12.5% to 37%; and species richness from 9.4% to 34.1% (Figure 6, Table 4). These values are all well within realistic ranges for the respective variables. The average increase in R^2 values was 14% points for median grain size, silt content, and biomass and for species richness the increase in R^2 was 18% points. These values imply that the features contributed consistently well for all four variables considered. For the combined dataset, R^2 increased by 15% points on average. The contribution of the encoded features was stronger at Zoutkamperlaag than at Pinkegat with an average increase of 18 and 13%, respectively. The improvement in model performance when including features was consistent for the 3 years. We tested with an average increase in R^2 values of 14% points for 2018, 16% points for 2019, and 15% points for 2020.

The absolute performance of the model (with spectral bands & extracted features) showed considerable variation between the sites and for the variables tested (Figure 7 & 8; Table 5A & B). We tested
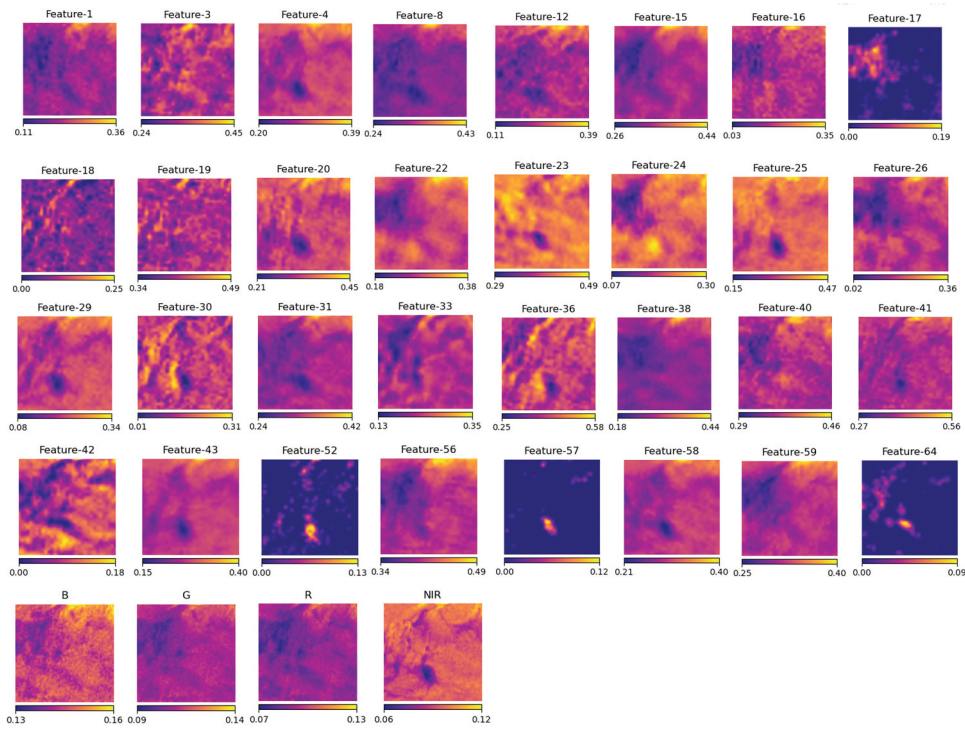
Figure 5. Examples of the encoded features for a patch from Pinkegat, derived from the Sentinel-2 image captured on 17 November 2017 2018. The feature numbers are specified in each feature image.
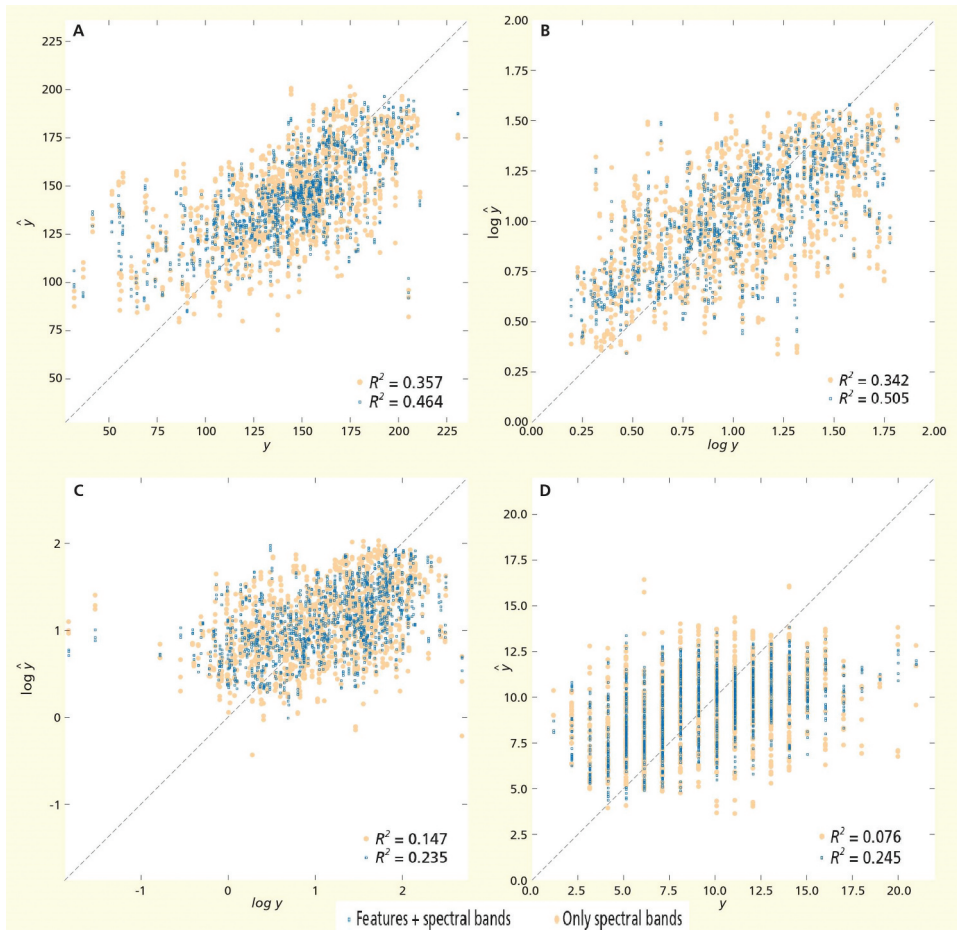

Figure 6. Predictions of the four variables for Pinkegat and Zoutkamperlaag combined, 2018, by using only the spectral bands and using the combination of spectral bands and the VAE features. $\hat{y}$ is the predicted and y is the observed value. A) Median grain size (in μm), B) Silt content (in %), C) Biomass (in g AFDM/m^2), and D) Species richness (no of species).

Table 4. Predictions (R^2) over the test dataset for all variables over different years and spatial constellations. The highest improvements for each variable is made bold. The training and testing are done over individual sites and then on combined site.

| Region | Year | Median grain size | | | Silt content | | | Biomass | | | Species richness | | |
		Spectral band (A)	Features + Spectral band (B)	B-A (%)	Spectral band (A)	Features + Spectral band (B)	B-A (%)	Spectral band (A)	Features + Spectral band (B)	B-A (%)	Spectral band (A)	Features + Spectral band (B)	B-A (%)
Pinkegat (P)	2018	0.496	0.617	12.1	0.412	0.535	12.3	0.191	0.334	14.3	0.108	0.275	16.7
	2019	0.526	0.612	8.6	0.475	0.597	12.2	0.077	0.235	15.8	0.215	0.316	10.1
	2020	0.301	0.478	17.7	0.374	0.468	9.40	0.202	0.370	16.8	0.258	0.341	8.30
Zoutkamperlaag (Z)	2018	0.136	0.315	17.9	0.253	0.486	**23.3**	0.084	0.086	0.20	0.008	0.170	16.2
	2019	0.248	0.366	11.8	0.319	0.474	15.5	−0.119	0.134	**25.3**	−0.271	0.094	**36.5**
	2020	0.126	0.255	12.9	0.201	0.321	12.0	−0.015	0.125	14.0	−0.033	0.156	18.9
Combined (P + Z)	2018	0.357	0.464	10.7	0.342	0.505	16.3	0.147	0.235	8.8	0.0761	0.245	16.9
	2019	0.392	0.473	8.1	0.441	0.528	8.7	0.055	0.212	15.7	−0.004	0.230	23.4
	2020	0.143	0.385	**24.2**	0.276	0.421	14.5	0.137	0.277	14.0	0.123	0.281	15.8

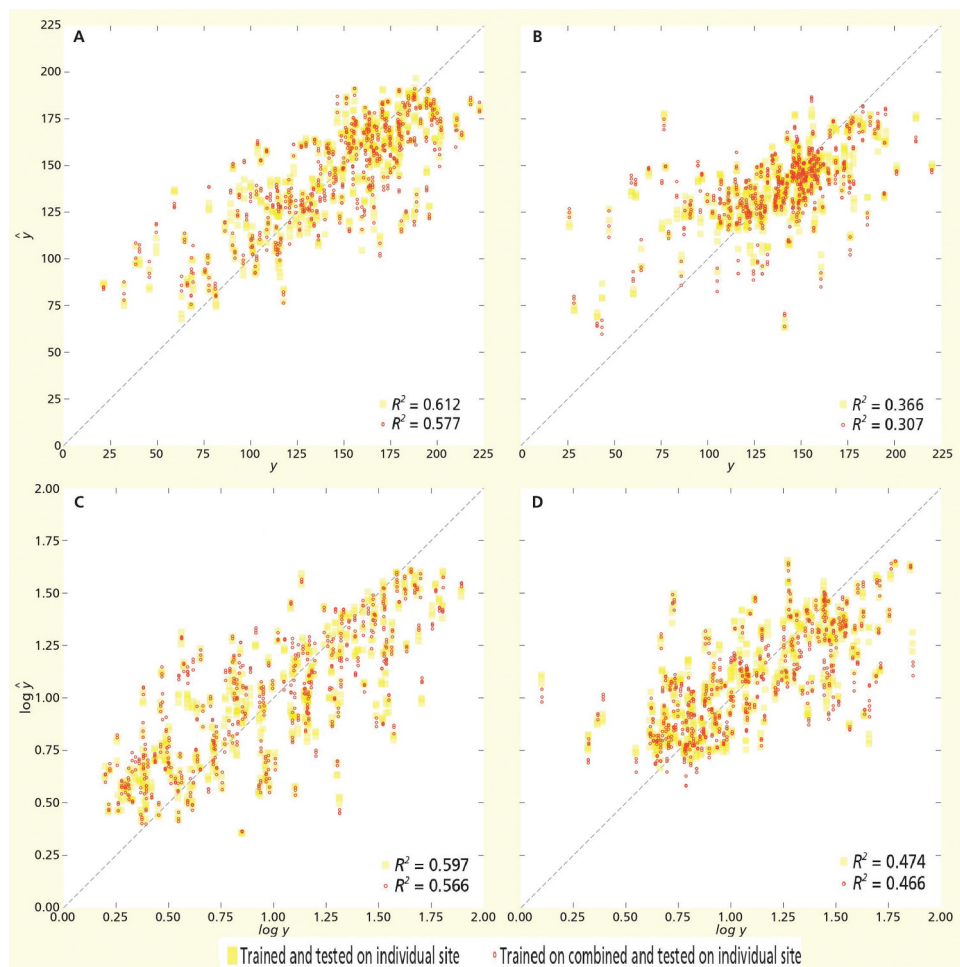A- Only Spectral bands B- With Encoded Features and Spectral bands (B-A)-% point change

Figure 7. Prediction results for 2019, A) Median grain size for the model trained and tested on Pinkegat vs trained on combined and tested on Pinkegat, B) Median grain size predictions for trained and tested on Zoutkamperlaag vs trained on combined, C) Silt content predictions for Pinkegat and combined, and D) Silt content predictions for Zoutkamperlaag and combined. $\hat{y}$ is the predicted and y is the observed value.

these scenarios to understand how well the model with encoded features could generalize over different sites and years. While the improvement in performance with features is the main focus of our study, the R^2 values that we achieve from our model are still comparable with some of the studies in the
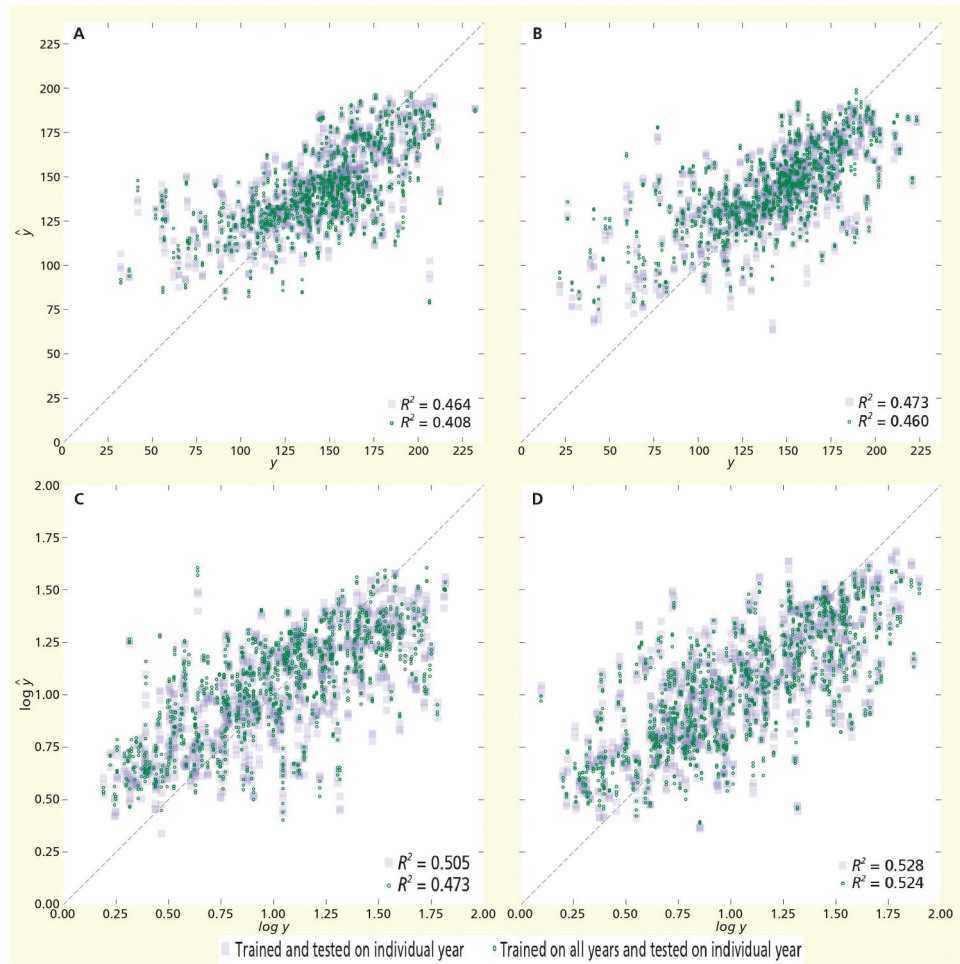
Figure 8. A) Combined site, median grain size predictions for 2018 and all years together, B) Combined site, median grain size predictions for 2019 and all years together, C) Combined site, silt content predictions for 2018 and all years together, and D) Combined site, silt content predictions for 2019 and all years together. ŷ is the predicted and y is the observed value.

literature discussed in *section 4.2*. The R^2 values for median grain size and silt content are consistently higher than for biomass and species richness for the combined dataset and for Pinkegat for all 3 years. For Zoutkamperlaag, all R^2 values were lower than for Pinkegat and the combined data. The R^2 values for silt content were consistently higher than those for the other three variables. For certain cases of species richness, without the use of encoded features the model even showed negative R^2 values, showing the model's inability to fit for that data.

The model's ability to generalize was explored by training and testing over various spatial and temporal scenarios (Table 5). For median grain size and silt content, the prediction values were better when trained and tested over the same dataset. But for biomass and species richness, there was an increase in the coefficient of determination values when trained over the

combined dataset and tested on an individual dataset. Similarly, for temporal combinations, we could observe that the model performed better when it was trained and tested on data from the same year. An exception is found for 2020, the biomass and species richness performed well when all years were combined than just using a single year.

For every tidal flat pixel in Pinkegat and Zoutkamperlaag, the four variables were predicted for 2019 (Figure 9) using the random forest model trained with the combined dataset from 2019. The descriptive statistics from the spatial maps for the four variables were compared to that of the field sampled points over the same year (Table 6). The mean values of the predictions all lie within the one standard deviation range from the mean values of the field observations. Standard deviations are much lower for the predictions, and the median values are closer to the mean values. For

Table 5. Predictions (R^2) for all variables over different years and regions to test the model's ability to generalize. A) Predictions when the model is trained and tested on individual sites compared with trained on combined site and tested on individual sites. B) Predictions done by training on a single year and testing over different years. Also, trained on all years combined and tested over different years for all variables.

(A)

Region	Year	Median grain size		Silt content		Biomass		Species richness	
		C	D	C	D	C	D	C	D
Pinkegat (P)	2018	0.617	0.580	0.535	0.515	0.334	0.369	0.275	0.306
or	2019	0.612	0.577	0.597	0.566	0.235	0.254	0.316	0.335
Combined (P + Z)	2020	0.478	0.470	0.468	0.462	0.370	0.358	0.341	0.359
Zoutkamperlaag (Z)	2018	0.315	0.292	0.486	0.482	0.086	0.110	0.170	0.171
or	2019	0.366	0.307	0.474	0.466	0.134	0.134	0.094	0.112
Combined (P + Z)	2020	0.255	0.227	0.321	0.340	0.125	0.121	0.156	0.173

(B)

Region	Year	Median grain size			Silt content			Biomass			Species richness		
		2018	2019	2020	2018	2019	2020	2018	2019	2020	2018	2019	2020
Combined (P + Z)	2018	0.464	0.216	0.086	0.505	0.264	0.153	0.235	0.112	0.139	0.245	−0.043	0.094
	2019	−0.250	0.473	−0.023	−0.209	0.528	0.049	−0.260	0.212	−0.020	−0.380	0.230	0.061
	2020	−0.001	0.085	0.385	−0.178	−0.090	0.421	−0.228	−0.724	0.277	−0.177	0.149	0.281
Combined (P + Z)	All years	0.413	0.460	0.356	0.473	0.522	0.416	0.256	0.221	0.295	0.251	0.203	0.297

C- Trained and Tested on individual site (Pinkegat-P, Zoutkamperlaag-Z) D- Trained on combined and tested with individual site (P/Z)
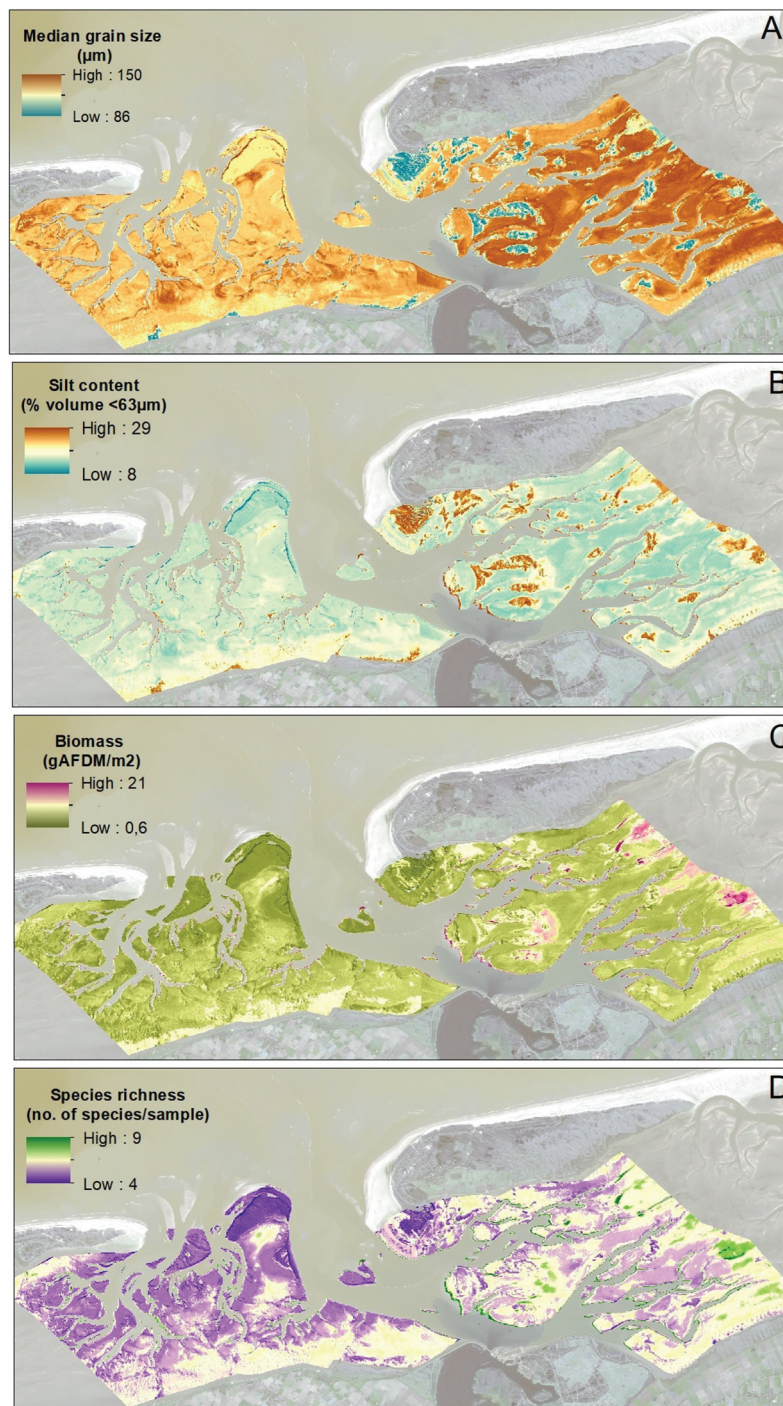
Figure 9. Spatial mapping of Pinkegat and Zoutkamperlaag for 2019 using the combined 2019 model. A) Median grain size, B) Silt content, C) Biomass, and D) Species richness.

Table 6. Descriptive statistics of the four variables for the field observation and the predicted values for 2019.

|	Field points					Predicted points				
Variables	Mean	Std	Median	Min-Max	No. of points	Mean	Std	Median	Min-Max	No. of points
Median grain size (µm)	140	36.8	147.5	21-222	460	123	4.32	123.8	86-150	1534889
Silt content (%)	15.6	14.4	9.9	1.3–78.1	447	16.3	1.26	15.9	7.6–29.5	1534889
Biomass (gAFDM/m2)	23.7	49.5	10.3	0.01–290	462	2.6	0.62	2.42	0.6–21.07	1534889
Species richness (no. of species/sample)	7	3.15	7	1-18	462	6	0.4	6	4-9	1534889
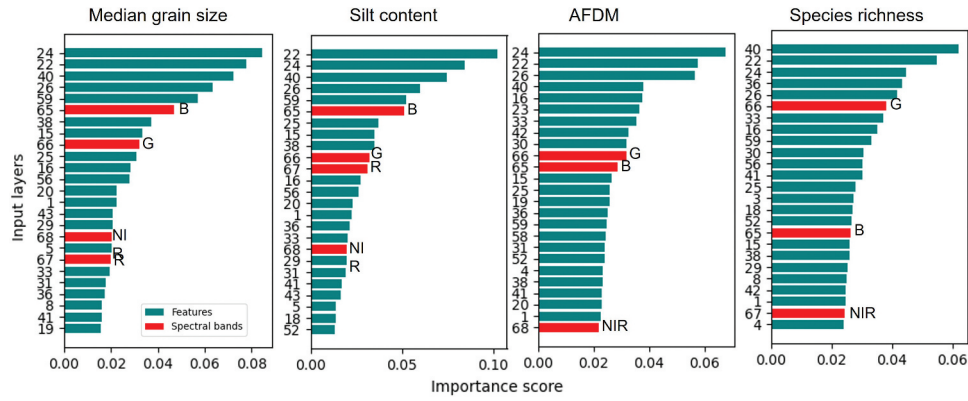
Figure 10. Feature importance scores for all variables for Pinkegat and Zoutkamperlaag in 2018 (top 18). X axis represents the importance score, Y axis represents the input layers including features (1–64) and spectral bands (65-B, 66-G, 67-R, 68-NIR).

biomass, the mean value is much smaller (ca. 10%) for the predictions than for the field observations.

3.3 Feature importance in random forest model

The most important predictors in the random forest models were autoencoder features for all the four target variables. Of the original bands, Green had the highest importance in each model, ranking between second and fifteenth place (Figure 10, Figure A4 & A5). Among the encoded features, features F-1, F-4, F-24, F-26 and F-38 (Figure 4) repeatedly occurred at the top positions of feature importance. For median grain size and silt content (Figure A3), two or three features contributed more with values up to 0.08 and then showed a steep reduction in plot. While for biomass and species richness more than three or four features contributed equally showing a vertical plot toward the bottom. The top 18 most important variables (among 36 non-zero input layers) for median grain size and silt content included three spectral bands, while for biomass and species richness only one spectral band was included.

4 Discussion

4.1 Feature extraction

In this paper, we proposed using a deep learning VAE model to improve information extraction from Sentinel images representing tidal flats. The spectral signatures show little contrast, and the VAE model created features that represented the spectral and spatial patterns in the image. We hypothesized that

the autoencoder features capture the spatial expression of differing environmental conditions and thereby improve mapping performance. When testing this on images over different years and different sites, the VAE model showed consistent features which indicates its ability use pre-trained VAE model when new images of the study area are acquired.

Several studies have tried to visualize the information contained inside these hidden layers within a deep learning network to improve the model performance and to understand what a model learns during training (Kahng et al. 2019; Junjie, Zhang, and Okatani 2019; Simonyan, Vedaldi, and Zisserman 2014). Yosinski et al. (2014) did experiments with classification models and discussed that the features in the initial layers of CNNs were formed with filters that tend to learn specific frequency content in any direction or identify regions that differ in spectral properties from the surroundings. This is also evident in the features that we extracted which showed varying patterns (Figure 4) indicating that these filters identified specific information. We observed features that are characterized by directional filters (Figure 4, F-15, F-41) and different color blobs targeting pixels of particular spectral values.

There are also features where the tidal flat regions are entirely populated with zeros (Figure 4, F-2). During training, we observed that these features showed very small negative values in the tidal flat regions after the convolution layer, which became null when passed through the ReLU activation function. This is probably due to the lack of representative information in the input image that is supposed to be derived by some filters. Due to the null values, these

features did not contribute to the prediction process. Mahendran and Vedaldi (2015) elaborated in their findings that the hidden layers retain an invariant and abstract notion of the image content. This could also be observed in our extracted features where the images still retain the basic spatial structure after passing through the convolutional layer.

The extracted features from our VAE model showed information that represents the spatial texture of the input tidal flats. It showed varied patterns and differences in the kind of information derived from a particular feature. In spectrally flat regions like tidal flats, deserts, and snow-covered mountains, these kinds of features extracted from VAE model can provide valuable additional information in terms of texture which complement the spectral bands in remote sensing images. The information contained in these features can be used as input for secondary processes such as classification, segmentation, and object detection. Nugroho et al. (2020) discussed the important information that is retained by the VAE model when it compresses the input. They also suggested that the features from the latent vector can provide suitable information and can substitute the original images.

In our study, we did not substitute the spectral bands but used the features together with the input spectral bands. We also tested and evaluated the case where we use only the features and remove spectral bands totally. For all the cases, the features showed better predictions than the spectral bands. The predictions accuracy further increased while combining both. The features provided complementary patterns and spatial arrangement that significantly enhanced the prediction results. This supports our hypothesis that spatial context and texture information captured the autoencoder has notable added value to spectral information when related to environmental and ecological variables

4.2 Prediction models

We observed that the encoded features are effective in enhancing the model capability. Additionally, the importance score clearly showed that the encoded features contributed substantially more to the model performance than the spectral bands (Figure 10). The features in top ranks showed different combinations of spatial characteristics, delineating,

and characterizing tidal flats, water channels, and the sand bar region (Figure 4).

While the overall performance of the prediction model improved with the use of extracted features, there is a substantial variation in predictive capabilities between environmental and ecological variables. Particularly, median grain size and silt content are predicted well. Biomass and species richness predictions are not to a level where the model is capable enough for operational applications, but our prediction accuracies are in line with existing remote sensing studies and need more data/images for better predictions (Etter and Frederick Grassle 1992; Puls et al. 2012; Van der Wal et al. 2008). Compton et al. (2013) suggested that mapping and monitoring these variables require the use of hyperspectral information along with other factors like hydrodynamics, land surface patterns, and field sampling. Also, Van der Wal and Herman (2007) showed that a combination of microwave and optical remote sensing images is required to map sediment characteristics with reliable results. Our study aimed to find a methodology that is less complex, more affordable, and widely applicable over different regions with freely available remote sensing data. We achieved similar accuracy to that of Van der Wal and Herman (2007) for the environmental variables, median grain size, and silt content with just the help of features from the deep learning model, which is much easier to acquire than microwave images.

For the two ecological variables, biomass and species richness, the inclusion of features did not improve the prediction accuracies as strongly as those of the environmental variables. Puls et al. (2012) obtained a prediction result (Cohen's kappa) of 0.25 ± 0.077 for macrozoobenthic communities' occurrence through four field collected sediment parameters and 0.44 ± 0.044 through five hydrodynamic parameters. The macrozoobenthos properties were related with NDVI based on ground reflectance spectra by Van der Wal et al. (2008) to achieve the R^2 of 0.40 for total biomass and 0.43 for species richness. Our model achieved a prediction score of 0.37 for some cases of biomass with the use of spectral bands and the extracted features of just multispectral images.

As far as the validity of using spectral bands for predicting species richness is concerned, most studies only try to predict the biomass. To progress research further, we wanted to test whether features have

some information content that is useful to predict species richness. Other studies have worked on developing response models using species richness variables based on ground spectra (Van der Wal et al. 2008). Since species richness is useful and pertinent information for the ecological community, we evaluated prediction capacity by using the spectral bands. We also observed variations in the results between the two sites when tested individually. While the pattern of increasing R^2 values with the use of the encoded features was observed at both sites, the overall prediction ability for Pinkegat was better than for Zoutkamperlaag. The model was unable to generalize well for the ecological variables at Zoutkamperlaag as can be seen from the insignificant R^2 values from Table 4.

The difference in performance in prediction results between Pinkegat and Zoutkamperlaag was also observed when the prediction was done using just the four spectral bands (Table 4). For median grain size, the difference in performance can be explained by the fact that the field data over Pinkegat is widely distributed, whereas in Zoutkamperlaag the data is more narrowly distributed (Figure 2). Some of the variation in predictions between different years may be caused by the time gap between the image acquisition, field sample collection, tide level at the time of image capture, and the level of exposure of the tidal flats.

The prediction model uses image features and field data that are matched as close in time as possible given tides and image availability, even though the feature extraction is done on a generic model. The generalizing capability of the prediction model is less than the VAE (Table 5) and requires the training dataset to have some points representative of the testing dataset. Predictions showed better R^2 values when the training data included samples from the same site or year as testing data, but the two ecological variables, biomass and species richness, had improved prediction accuracy when training data was increased beyond the test data either spatially or temporally (Table 4). These results indicate that while our prediction model requires more data from the same time and area of interest, they can still be improved by including adjacent areas or periods which is enabled more easily by a pre-trained general autoencoder model.

From the random forest models, a map of the tidal flats was created for each of the four variables (Figure 9). The model predicted values did not violate the constraints of the compositional data, and they fell within a realistic range of values (Table 6). It was challenging for the model to encompass all the variability in the data, as the standard deviations of the predictions are much smaller than those of the field data (Table 6). In particular, the model found it difficult to predict the values that were more than one standard deviation away from the mean on the field data distribution.

The spatial distribution of the four variables showed the patterns as expected. Median grain size and silt content showed inverse patterns (Cao et al. 2016), and sediments were generally less coarse and contained more silt in the less dynamic parts of this tidal basin, such as the tidal divides perpendicular to the islands and areas close to the shores of the islands and the mainland (Alonso et al. 2021). The maps of macrozoobenthic biomass and richness reflected more or less those of sediment composition, with higher total biomass and more species in areas with finer sediments with high silt contents (Compton et al. 2013; Drent et al. 2017).

To summarize, although the model's prediction performance could improve (e.g. the predicted mean value for biomass is still relatively low), the predicted values generally fall within the expected ranges and the spatial patterns of the variables were captured well.

5 Conclusion

This paper presents a novel approach showing that encoded features derived from a deep learning model enhance the predictive performance for coastal environmental and ecological variables when used along with the original spectral remote sensing data. We used a Variational Autoencoder Model trained with Sentinel-2 satellite images. The encoded features improved the prediction performance for the scenarios considered consistently by an average of 15% points in comparison to spectral bands. Prediction capability was higher for environmental variables than for ecological variables. The maps created with the random forest models revealed realistic values and spatial patterns for the four variables.

The VAE features clearly show differing spatial expressions of image information, and their contribution to model performance was also evident from the

importance score where the top positions were mostly occupied by the features. The enhancement of the spectral information with the features extracted from the deep learning model to predict environmental and ecological variables is shown to be an efficient way to elevate mapping and monitoring of tidal systems as it does not require additional data collection. Our results suggest that features extracted from VAE models may also improve remote-sensing-based mapping in other spectrally poor regions like deserts and snow-covered mountains.

ORCID

Logambal Madhuanand 🔟 http://orcid.org/0000-0002-1534-5234

Catharina J. M. Philippart 🔟 http://orcid.org/0000-0001-7238-9113
Wiebe Nijland 🔟 http://orcid.org/0000-0002-2665-0947
Steven M. de Jong 🔟 http://orcid.org/0000-0002-1586-9601
Allert I. Bijleveld 🔟 http://orcid.org/0000-0002-3159-8944
Elisabeth A. Addink 🔟 http://orcid.org/0000-0002-0919-6498

References

Adolph, W., H. Farke, S. Lehner, and M. Ehlers. 2018. "Remote Sensing Intertidal Flats with TerraSar-X. A SAR Perspective of the Structural Elements of a Tidal Basin for Monitoring the Wadden Sea." *Remote Sensing* 10 (7): 1085. doi:10.3390/rs10071085.

Agarap, A. F. 2018. "Deep Learning Using Rectified Linear Units (ReLu)." *ArXiv*, no. 1: 2–8. http://arxiv.org/abs/1803.08375.

Alonso, A. C., D. S. V. M. van Maren, E. P. L. Elias, S. J. Holthuijsen, and Z. B. Wang. 2021. "The Contribution of Sand and Mud to Infilling of Tidal Basins in Response to a Closure Dam." *Marine geology* 439 (February): 106544. doi:10.1016/j.margeo.2021.106544.

Bakker, W., B. J. Ens, A. Dokter, H. -J. van der Kolk, K. Rappoldt, M. van de Pol, K. Troost, et al. 2021. "Connecting Foraging and Roosting Areas Reveals How Food Stocks Explain Shorebird Numbers." *Estuarine, Coastal and Shelf Science* 259 (June): 107458. doi:10.1016/j.ecss.2021.107458.

Baldi, P. 2012. "Autoencoders, Unsupervised Learning, and Deep Architectures." *ICML Unsupervised and Transfer Learning* 37–50. doi:10.1561/2200000006.

Bartholdy, J., and S. Folving. 1986. "Sediment Classification and Surface Type Mapping in the Danish Wadden Sea by Remote Sensing." *Netherlands Journal of Sea Research* 20 (4): 337–345. doi:10.1016/0077-7579(86)90001-3.

Beukema, J. J. 1976. "Biomass and Species Richness of the Macro-Benthic Animals Living on the Tidal Flats of the Dutch Wadden Sea." *Netherlands Journal of Sea Research* 10 (2): 236–261. doi:10.1016/0077-7579(76)90017-X.

Bhandare, A., M. Bhide, P. Gokhale, and R. Chandavarkar. 2016. "Applications of Convolutional Neural Networks." *International Journal of Computer Science and Information Technologies* 7 (5): 2206–2215. http://ijcsit.com/docs/Volume7/vol7issue5/ijcsit20160705014.pdf.

Bijleveld, A. I., R. B. MacCurdy, Y. -C. Chi Chan, E. Penning, R. M. Gabrielson, J. Cluderay, E. L. Spaulding, et al. 2016. "Understanding Spatial Distributions: Negative

Density-Dependence in Prey Causes Predators to Trade-Off Prey Quantity with Quality." *Proceedings of the Royal Society B: Biological Sciences* 283 (1828): 20151557. doi: 10.1098/rspb.2015.1557.

Bijleveld, A. I., J. A. van Gils, J. van der Meer, A. Dekinga, C. Kraan, H. W. van der Veer, and T. Piersma. 2012. "Designing a Benthic Monitoring Programme with Multiple Conflicting Objectives." *Methods in Ecology and Evolution* 3 (3): 526–536. doi:10.1111/j.2041-210X.2012.00192.x.

Boere, G. C., and T. Piersma. 2012. "Flyway Protection and the Predicament of Our Migrant Birds: A Critical Look at International Conservation Policies and the Dutch Wadden Sea." *Ocean & coastal management* 68: 157–168. doi:10.1016/j.ocecoaman.2012.05.019.

Breiman, L. 2001. "Random Forests." *Machine Learning* 45 (1): 5–32. doi:10.1007/978-3-030-62008-0_35.

Cao, Z., G. Liu, H. Zhan, L. Chaozheng, Y. You, C. Yang, and H. Jiang. 2016. "Pore Structure Characterization of Chang-7 Tight Sandstone Using MICP Combined with N 2 GA Techniques and Its Geological Control Factors." *Nature*, Vol. 6. Nature Publishing Group. 10.1038/srep36919

Chicco, D., M. J. Warrens, and G. Jurman. 2021. "The Coefficient of Determination R-Squared is More Informative Than SMAPE, MAE, MAPE, MSE and RMSE in Regression Analysis Evaluation." *PeerJ Computer Science* 7: e623. doi:10.7717/PEERJ-CS.623.

Compton, T. J., S. Holthuijsen, A. Koolhaas, A. Dekinga, J. ten Horn, J. Smith, Y. Galama, et al. 2013. "Distinctly Variable Mudscapes: Distribution Gradients of Intertidal Macrofauna Across the Dutch Wadden Sea." *Journal of Sea Research* 82: 103–116. doi:10.1016/j.seares.2013.02.002.

Drent, J., R. Bijkerk, M. Herlyn, M. Grotjahn, J. Voß, M. -C. Carausu, and D. W. Thieltges. 2017. "Wadden Sea Quality Status Report - Alien Species." *Wadden Sea Quality Status Report*. qsr.waddensea-worldheritage.org/reports/alien-species.

Etter, R. J., and J. Frederick Grassle. 1992. "Patterns of Species Diversity in the Deep Sea as a Function of Sediment Particle Size Diversity." *Nature* 360 (6404): 576–578. doi:10.1038/360576a0.

Gao, H., Z. Liu, L. Van Der Maaten, and K. Q. Weinberger. 2017. "Densely Connected Convolutional Networks." In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2261–2269. doi: 10.1109/CVPR.2017.243.

Goodfellow, I., Y. Bengio, and A. Courville. 2016. *Deep Learning*. MIT Press. www.deeplearning.org.

Guo, W., W. Yang, H. Zhang, and G. Hua. 2018. "Geospatial Object Detection in High Resolution Satellite Images Based on Multi-Scale Convolutional Neural Network." *Remote Sensing* 10 (1): 131. doi:10.3390/rs10010131.

Hedley, J. D., C. Roelfsema, V. Brando, C. Giardino, T. Kutser, S. Phinn, P. J. Mumby, O. Barrilero, J. Laporte, and B. Koetz. 2018. "Coral Reef Applications of Sentinel-2: Coverage, Characteristics, Bathymetry and Benthic Mapping with Comparison to Landsat 8." *Remote Sensing of Environment* 216 (April): 598–614. doi:10.1016/j.rse.2018.07.014.

Hollebrandse, F. A. P. 2005. "Temporal Development of the Tidal Range in the Southern North Sea." Delft University of Technology. https://www.semanticscholar.org/paper/Temporal-development-of-the-tidal-range-in-the-Sea-Hollebrandse/dfb2742f3009b87125d8e9df7d421a9162e9bcfe.

Howard, A. G., M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, and H. Adam. 2017. "MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications." *Arxiv*. http://arxiv.org/abs/1704.04861.

Joyce, J. M. 2011. "Kullback-Leibler Divergence." In *International Encyclopedia of Statistical Science*, edited by M. Lovric, 720–722. Berlin: Springer. doi:10.1007/978-3-642-04898-2_327.

Jung, Y., and J. Jianhua. 2015. "A K -Fold Averaging Cross-Validation Procedure." *Journal of Nonparametric Statistics* 27 (2): 167–179. doi:10.1080/10485252.2015.1010532.

Junjie, H., Y. Zhang, and T. Okatani. 2019. Visualization of Convolutional Neural Networks for Monocular Depth Estimation, 3869–3878. http://arxiv.org/abs/1904.03380.

Kahng, M., N. Thorat, D. Horng Polo Chau, F. B. Viégas, and M. Wattenberg. 2019. "GAN Lab: Understanding Complex Deep Generative Models Using Interactive Visual Experimentation." *IEEE Transactions on Visualization and Computer Graphics* 25 (1): 310–320. doi:10.1109/TVCG.2018.2864500.

Kaiming, H., X. Zhang, S. Ren, and J. Sun. 2016. "Deep Residual Learning for Image Recognition." *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 770–778. 10.1109/CVPR.2016.90.

Kingma, D. P., and M. Welling. 2014. "Auto-Encoding Variational Bayes." *2nd International Conference on Learning Representations, ICLR 2014 - Conference Track Proceedings*, Banff, AB, Canada, April 14-16.

Kloepper, S., R. Strempel, A. Bostelmann et al. 2017. *CWSS(2017) Introduction In: Wadden Sea Quality Status Report 2017*. Common Wadden Sea Secretariat, Wilhelmshaven, Germany: Wadden Sea Ecosystem. https://qsr.waddensea-worldheritage.org/reports/introduction

Lee, S., I. Park, B. J. Joo Koo, J. -H. Hyung Ryu, J. -K. Kuk Choi, and H. J. Jun Woo. 2013. "Macrobenthos Habitat Potential Mapping Using GIS-Based Artificial Neural Network Models." *Marine pollution bulletin* 67 (1–2): 177–186. doi:10.1016/j.marpolbul.2012.10.023.

Levin, L. A., D. F. Boesch, A. Covich, C. Dahm, C. Erséus, K. C. Ewel, R. T. Kneib, et al. 2001. "The Function of Marine Critical Transition Zones and the Importance of Sediment Biodiversity." *Ecosystems* 4 (5): 430–451. doi:10.1007/s10021-001-0021-4.

Madhuanand, L., F. Nex, and M. Ying Yang. 2021. "Self-supervised monocular depth estimation from oblique UAV videos." *ISPRS Journal of Photogrammetry and Remote Sensing* 176: 1–14. doi:10.1016/j.isprsjprs.2021.03.024.

Mahendran, A., and A. Vedaldi. 2015. "Understanding Deep Image Representations by Inverting Them." *Proceedings of the IEEE Computer Society Conference on Computer Vision and*

Pattern Recognition June 07-12: 5188–5196. doi: 10.1109/CVPR.2015.7299155.

Marencic, H. 2009. "Quality Status Report 2009. Wadden Sea Ecosystem No. 25." http://www.waddensea-secretariat.org/sites/default/files/downloads/qsr-2009.pdf.

Menze, B. H., B. M. Kelm, R. Masuch, U. Himmelreich, P. Bachert, W. Petrich, and F. A. Hamprecht. 2009. "A Comparison of Random Forest and Its Gini Importance with Standard Chemometric Methods for the Feature Selection and Classification of Spectral Data." *BMC Bioinformatics* 10 (1): 1–16. doi:10.1186/1471-2105-10-213.

Miloslavich, P., N. J. Bax, S. E. Simmons, E. Klein, W. Appeltans, O. Aburto-Oropeza, M. Andersen Garcia, et al. 2018. "Essential Ocean Variables for Global Sustained Observations of Biodiversity and Ecosystem Changes." *Global Change Biology* 24 (6): 2416–2433. doi:10.1111/gcb.14108.

Murray, N. J., S. R. Phinn, M. DeWitt, R. Ferrari, R. Johnston, M. B. Lyons, N. Clinton, D. Thau, and R. A. Fuller. 2019. "The Global Distribution and Trajectory of Tidal Flats." *Nature*, Vol. 565. Springer US. doi: 10.1038/s41586-018-0805-8.

Nguyen, A., J. Yosinski, and J. Clune. 2016. "Multifaceted Feature Visualization: Uncovering the Different Types of Features Learned by Each Neuron in Deep Neural Networks." http://arxiv.org/abs/1602.03616.

Nugroho, H., M. Susanty, A. Irawan, M. Koyimatu, and A. Yunita. 2020. "Fully Convolutional Variational Autoencoder for Feature Extraction of Fire Detection System." *Journal of Computer Science and Information* 13 (1): 9. doi:10.21609/jiki.v13i1.761.

Olden, J. D., J. J. Lawler, and N. Leroy Poff. 2008. "Machine Learning Methods Without Tears: A Primer for Ecologists." *The Quarterly Review of Biology* 83 (2): 171–193. doi:10.1086/587826.

Park, W., Y. Kyung Lee, D. Jae Kwon, and J. Sun Won. 2014. "Optical Remote Sensing for Long-Term Changes of Surface Sediments on Tidal Flats: Preliminary Results in the West Coast of Korea." *International Geoscience and Remote Sensing Symposium (IGARSS)*, 4335–4338. doi:10.1109/IGARSS.2014.6947449.

Paszke, A., G. Chanan, Z. Lin, S. Gross, E. Yang, L. Antiga, and Z. Devito. 2017. "Automatic Differentiation in PyTorch. NIPS 2017 Workshop Autodiff Submission. Dec 4-9, Long Beach Convention Center, Long Beach." Nips 1–4.

Philippart, C. J. M., J. J. Beukema, G. C. Cadée, R. Dekker, P. W. Goedhart, J. M. Van Iperen, M. F. Leopold, and P. M. J. Herman. 2007. "Impacts of Nutrient Reduction on Coastal Communities." *Ecosystems* 10 (1): 95–118. doi:10.1007/s10021-006-9006-7.

Piersma, T., R. Hoekstra, A. Dekinga, A. Koolhaas, P. Wolf, P. Battley, and P. Wiersma. 1993. "Scale and Intensity of Intertidal Habitat Use by Knots Calidris Canutus in the Western Wadden Sea in Relation to Food, Friends and Foes." *Netherlands Journal of Sea Research* 31 (4): 331–357. doi:10.1016/0077-7579(93)90052-T.

Puls, W., K. H. van Bernem, D. Eppel, H. Kapitza, A. Pleskachevsky, R. Riethmüller, and B. Vaessen. 2012.

"Prediction of Benthic Community Structure from Environmental Variables in a Soft-Sediment Tidal Basin (North Sea)." *Helgoland Marine Research* 66 (3): 345–361. doi:10.1007/s10152-011-0275-y.

Rainey, M. P., A. N. Tyler, R. G. Bryant, D. J. Gilvear, and P. McDonald. 2000. "The Influence of Surface and Interstitial Moisture on the Spectral Characteristics of Intertidal Sediments: Implications for Airborne Image Acquisition and Processing." *International Journal of Remote Sensing* 21 (16): 3025–3038. doi:10.1080/01431160050144938.

Rammer, W., and R. Seidl. 2019. "Harnessing Deep Learning in Ecology: An Example Predicting Bark Beetle Outbreaks." *Frontiers in plant science* 10. doi:10.3389/fpls.2019.01327.

Rijkswaterstaat. n.d. "Getij | Rijkswaterstaat."|"Getij | Rijkswaterstaat." Accessed November 1, 2021. https://www.rijkswaterstaat.nl/water/waterdata-en-waterberichtgeving/waterdata/getij.

Ryu, J. -H., Y. -H. Ho Na, J. -S. Sun Won, and R. Doerffer. 2004. "A Critical Grain Size for Landsat ETM+ Investigations into Intertidal Sediments: A Case Study of the Gomso Tidal Flats, Korea." *Estuarine, Coastal and Shelf Science* 60 (3): 491–502. doi:10.1016/j.ecss.2004.02.009.

Ryu, J. -H., J. -K. Kuk Choi, and Y. -K. Kyung Lee. 2014. "Potential of Remote Sensing in Management of Tidal Flats: A Case Study of Thematic Mapping in the Korean Tidal Flats." *Ocean & coastal management* 102: 458–470. doi:10.1016/j.ocecoaman.2014.03.003.

Sehgal, S. 2012. "Remotely Sensed LANDSAT Image Classification Using Neural Network Approaches." *Computer Science* 2 (5): 43–46. https://www.semanticscholar.org/paper/Remotely-Sensed-LANDSAT-Image-Classification-Using-Sehgal/ff66711f110389327c7aa3b34b81c664866d8d25.

Simonyan, K., A. Vedaldi, and A. Zisserman. 2014. "Deep Inside Convolutional Networks: Visualising Image Classification Models and Saliency Maps." *2nd International Conference on Learning Representations, ICLR 2014 - Workshop Track Proceedings*, April 14-16, Banff, Canada, 1–8.

"User Guides - Sentinel-2 MSI." 2015. ESA-Sentinel 2. https://sentinel.esa.int/web/sentinel/user-guides/sentinel-2-msi.

Van der Wal, D., and P. M. J. Herman. 2007. "Regression-Based Synergy of Optical, Shortwave Infrared and Microwave Remote Sensing for Monitoring the Grain-Size of Intertidal Sediments." *Remote Sensing of Environment* 111 (1): 89–106. doi:10.1016/j.rse.2007.03.019.

Van der Wal, D., P. M. J. Herman, R. M. Forster, T. Ysebaert, F. Rossi, E. Knaeps, Y. M. G. Plancke, and S. J. Ides. 2008. "Distribution and Dynamics of Intertidal Macrobenthos Predicted from Remote Sensing: Response to Microphytobenthos and Environment." *Marine Ecology Progress Series* 367 (June 2014): 57–72. doi:10.3354/meps07535.

Van der Wal, D., P. Herman, and T. Ysebaert. 2004. "Space-Borne Synthetic Aperture Radar of Intertidal Flat Surfaces as a Basis for Predicting Benthic Macrofauna Distribution." *EARSeL EProceedings* 3 (1): 69–80.

Willcock, S., J. Martínez-López, D. A. P. Hooftman, K. J. Bagstad, S. Balbi, A. Marzo, C. Prato, et al. 2018. "Machine Learning for Ecosystem Services." *Ecosystem Services* 33: 165–174. doi:10.1016/j.ecoser.2018.04.004.

Yates, M. G., A. R. Jones, S. McGrorty, and J. D. Goss-Custard. 1993. "The Use of Satellite Imagery to Determine the Distribution of Intertidal Surface Sediments of the Wash, England." *Estuarine, Coastal and Shelf Science* 36 (4): 333–344. doi: 10.1006/ecss.1993.1020.

Yosinski, J., J. Clune, Y. Bengio, and H. Lipson. 2014. "How Transferable are Features in Deep Neural Networks?" *Advances in Neural Information Processing Systems* 4 (Jan): 3320–3328.

Zeiler, M. D., and R. Fergus. 2014. "Visualizing and Understanding Convolutional Networks." In *ECCV*, 8689 LNCS:818–833. doi:10.1007/978-3-319-10590-1_53.

Zhang, G., M. Wang, and K. Liu. 2019. "Forest Fire Susceptibility Modeling Using a Convolutional Neural Network for Yunnan Province of China." *International Journal of Disaster Risk Science* 10 (3): 386–403. doi:10.1007/s13753-019-00233-1.

Zwarts, L., and J. H. Wanink. 1993. "How the Food Supply Harvestable by Waders in the Wadden Sea Depends on the Variation in Energy Density, Body Weight, Biomass, Burying Depth and Behaviour of Tidal-Flat Invertebrates." *Netherlands Journal of Sea Research* 31 (4): 441–476. doi:10.1016/0077-7579(93)90059-2.