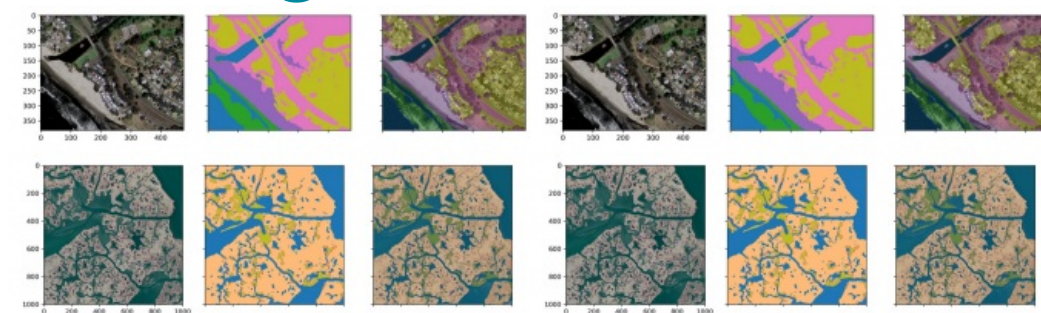# New dataset provides insights into coastal change and management

A new paper published in '*Scientific Data*' presents the 'CoastTrain' dataset, a collection of orthomosaic and satellite images of coastal environments along with corresponding labels. The dataset includes 1.2 billion labelled pixels, representing over 3.6 million hectares of diverse coastal environments, and was created using a human-in-the-loop tool designed for rapid and reproducible Earth surface image segmentation.

The authors argue that the CoastTrain dataset will enable the use of machine learning models for the mapping and monitoring of coastal change. Machine learning models that carry out supervised (i.e. human-guided) pixel-based classification, or image segmentation, have transformative applications in the spatio-temporal mapping of dynamic environments, including transient coastal landforms, sediments, habitats, waterbodies and water flows. However, these models require large and well-documented training and testing datasets consisting of labelled imagery.

Understanding Coastline Dynamics

Coastlines are highly spatially variable, coupled human-natural systems that comprise a nested hierarchy of component landforms, ecosystems and human interventions, each interacting over a range of space and time scales. Understanding and predicting coastline dynamics necessitates frequent observation from imaging sensors on remote sensing platforms.

The availability of imagery from Earth observation platforms in coastal areas has enabled models of physical processes in the coastal zone to focus on coastal change measured in decades to centuries and tens to hundreds of kilometres. Remotely sensed photography has been used to monitor coastal ecosystems and hazards, such as hurricanes, flooding and cliff erosion for almost a century. In some areas, aerial photos of the coast predate extensive modification of coastal morphology and ecosystems by humans.

This new dataset, CoastTrain, will be a valuable tool for coastal scientists, remote sensing experts and machine learning practitioners for understanding and predicting coastline dynamics, which is important for coastal management and conservation. The dataset will be made publicly available and will be an important resource for researchers in the field.

[The paper can be found here](#)

The CoastTrain dataset, a collection of orthomosaic and satellite images of coastal environments along with corresponding labels.