

Pixlab-Kmask

Instructions for use

[Pixlab repository link](#) [Kmask code link](#)

The Pixlab repository hosts a variety of ML applications for lithologic characterization including an Mask R-CNN instance segmentation model, a fully integrated application to generate training data, and a semantic segmentation model (Rodriguez, 2025). K-mask is a platform for automatic characterization and modal analysis of carbonate sample images. K-mask utilizes stochastic vector quantization (clustering) algorithms to extract associations in input data. To access the Kmask API use the script:

```
from Kmask import KMask
```

where KMask is the object contained in the file Kmask.py. Generating a model instance of K-mask requires one parameter and can be done with the following code:

```
model = KMask(im)
```

where im is an array of paths to images (.png, .tif, or .tiff). Making predictions using Kmask requires specifying a ground truth as well as a 'recipe' for the vector design process (e.g. selecting principal components or generating band compositions). The full code for making a prediction may look something like:

```
from Kmask import KMask
k = 5
model = KMask(im)
# If no ground truth data is present this may also be 'None' (This does not affect the
predictions made by K-mask)
ground_truth = {the path to the ground truth annotations -- also an array like im}
transformation = []
model.predict(5, gt, transformation)
plt.imshow(model.predicted_labels)
```

Here I specify that the model should look for 5 facies, then I instantiate a KMask object, define the ground truth, define the transformations, then run a prediction. Here, the transformation is empty -- I have specified that I don't want the model to design an input vector, and instead just use the regular image data. I may choose to define:

```
transformation = [(model.PCA, [[0, 1, 2]], {})]
```

Here, I am specifying that I want to select the principal components 0, 1, and 2 as my input data. I may also choose to define:

```
transformation = [(model.PCA, [[0, 1, 2]], {}, model.build_filter_predictor, [k], {})]
```

Here, I am additionally including texture filter response in my input vector. I may also choose to define:

```
transformation = [(model.generate_image_composition, [[0, 7]], {my_function})]
```

Here, I define the input vector to be a composition of multispectral bands, where `my_function` is a user-defined function which takes two arguments (corresponding to band 0 and 7). For example we can utilize the Normalized Difference Vegetation Index if we define my function to be the NDVI function:

```
def my_function(a, b):  
    return(b-a)/(b+a)
```

Included in this package are the data used in Rodriguez 2025, "Automatic Characterization of Marine Carbonate Microfacies Across the OAE". Data are easily accessed via `Kmask`:

```
dataset = model.sampledatapaths
```

For ease of use, I've included the following script to generate an example prediction in google colab, utilizing a multispectral image as input, and taking PCA 2 as the data vector. Pycharm or Jupyter Lab are the preferred IDE for running `Kmask`.

```
from google.colab import drive  
!pip install colorama # The only package in Kmask not included in colab  
drive.mount('/content/drive')  
import os  
os.chdir("drive/My Drive/Kmask") # cd into the relevant dir  
from Kmask import KMask #Import the module  
k = 5 # Define k  
model = KMask(None) # Initialize an empty model instance  
dataset = model.sampledatapaths # grab the dataset  
# Initialize a model instance with data  
model = KMask(dataset["65R-01W-15-17"]["partial"])  
# Dont pass a ground truth annotation -- bc this isn't colab-friendly for some reason?  
# does work on Jupyter  
ground_truth = None  
transformation = [(model.PCA, [[1]], {})] # 1 bc arrays are 0-indexed  
model.predict(k, ground_truth, transformation)  
plt.imshow(model.predicted_labels)
```

Additional Notes

Much of the code in `Kmask` is uncommented/poorly organized. This base version is a temporary build for the submission of "Automatic Characterization of Marine Carbonate Microfacies Across the OAE." `Kmask` is described here as a part of the larger [Pixlab](#) repository, though I have not yet merged the two in order to maintain the integrity of my JP submission -- the code described here is my 'submitted' version. Future iterations of `Kmask` will appear in `Pixlab` and will be updated periodically.

For this 'submitted' version, please direct pull-requests and inquiries to cr5357@princeton.edu