# Python pseudo code outlining the image analysis process

# Required Data:
# irSpectra - 2D data array with shape (number of spectra, number of points in each spectrum)
# irImage - 2D IR image with shape (x,y) where x*y = number of spectra

import numpy as np
from scipy import ndimage
import pymorph

# median filter to remove noise (3x3 pixel filter in this case)
fltImage = ndimage.filters.median_filter(irImage, size=3)

# set background pixels to zero (0.03 used as minimum signal value here)
fltImage[fltImage<0.03] = 0

# determine regional maxima locations (fltImage must be converted to bytes for pymorph package, not shown here)

rmax = pymorph.regmax(fltImage)

# label each regional maximum
seeds, nr_nuclei = ndimage.label(rmax)

# use the regional maxima as seeds for the watershed
labels = pymorph.cwatershed(-fltImage, seeds)

# calculate mean spectra per labelled cell
meanSpectra = np.zeros([labels.max()+1, irSpectra.shape[1]])

for i in range(0, labels.max()+1):
    # determine which pixels are in the cell and which spectra they correspond to
    inCell = labels.reshape(irSpectra.shape[0]) == i
    # average all the spectra in the cell
    meanSpectra[i, :]= irSpectra[inCell, :].mean(axis=0)

# meanSpectra now contains all the average spectra for each of the isolated cells
# meanSpectra[i, :] is the average spectrum corresponding to the image labels == i