Appendix 2: Quantitative and Qualitative analysis of FLocK features

This appendix includes 1) description of a FLocK GUI, which provide a quantitative and interactive approach to let the user check the FLocK configuration interactively; 2) quantitative analysis on the effect of some feature values with respect to feature space bandwidth (h_s and h_m).

A. FLocK GUI

In the code package [34], we provided a GUI for user to test the FLocK construction by customizing the parameter setting (see Fig. S1). Also, we provide a subset of FLocK feature values, which include *'FLocK-portion of intersected FLocK', 'FLocK-abs. number of intersected FLocK', 'FLocK-portion of highly intersected FLocK:T=0.3', 'FLocK-mean(size of non-1-2cell FLocK)', 'FLocK-min(nuclei no. in FLocK)', 'FLocK-max(nuclei no. in FLocK)', 'FLocK-mean(var of nuclear feat1 to the centroid)', 'FLocK-mean(var of nuclear feat2 to the centroid)'*.
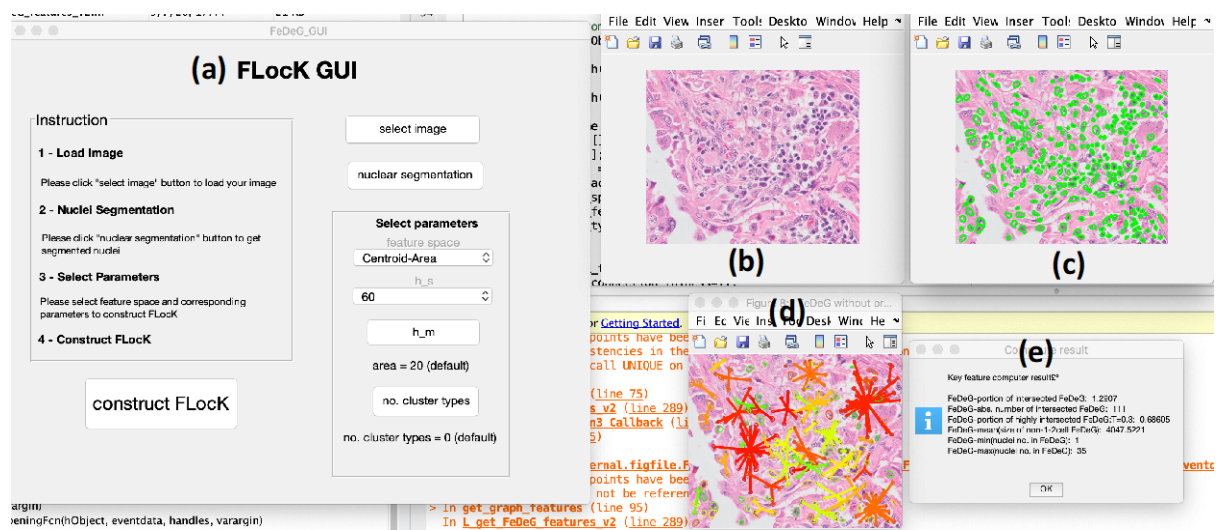

Figure S1: FLocK GUI. User can set up parameters for FLocK construction (left panel (a)) for a given image (b). These parameters include feature space, bandwidth for the space (h_s) and morphology (h_m), and number of cluster types. Once the parameters are set, the FLocK result will be generated and displayed (bottom-right panels (d) and (e)).

B. Qualitative analysis of the effect of FLocK parameters on typical FLocK features

We performed a qualitative analysis on the effect of parameters for representative FlocK feature values with respect to feature space bandwidth (h_s and h_m). The example image we used is shown in Fig. S1(b). Here we used feature space 'centroid-area-mean intensity' for the analysis. We evaluated three typical FLocK features: 1) *portion of intersected FLocKs*, 2) *mean(size of non-102 cell FLocKs),* and 3) *mean(var of nuclear feat1 to the centroid),* which reflects the interaction between FLocKs, size of FLocKs, and disorder of the nuclear area within a FLocK, respectively. Fig.S2 illustrates these feature values with respect to the space bandwidth h_s and the bandwidths for nuclear morphology, i.e., area and mean intensity. As one may expect, with increase in space bandwidth h_s, the larger of the FLocKs, therefore the larger of feature values for *portion of intersected FLocKs* and *mean of the FlocKs size.* However, the value of the feature *mean(var of nuclear feat1 to the centroid)* does not appear to be substantially affected by the space bandwidth h_s. Bandwidths of the nuclear morphology, i.e., area and mean intensity correlate with the feature value, since the higher the bandwidth of nuclear morphology, the greater the number of nuclei included in that particular FLocK. This in turn results in higher variance of nuclear feature 1 (i.e., area) within a FLocK.

One could use our released code at Github for executing a more quantitative analysis of different FLocK features. (https://github.com/hacylu/FLocK/quantitative_analysis_3d.m).
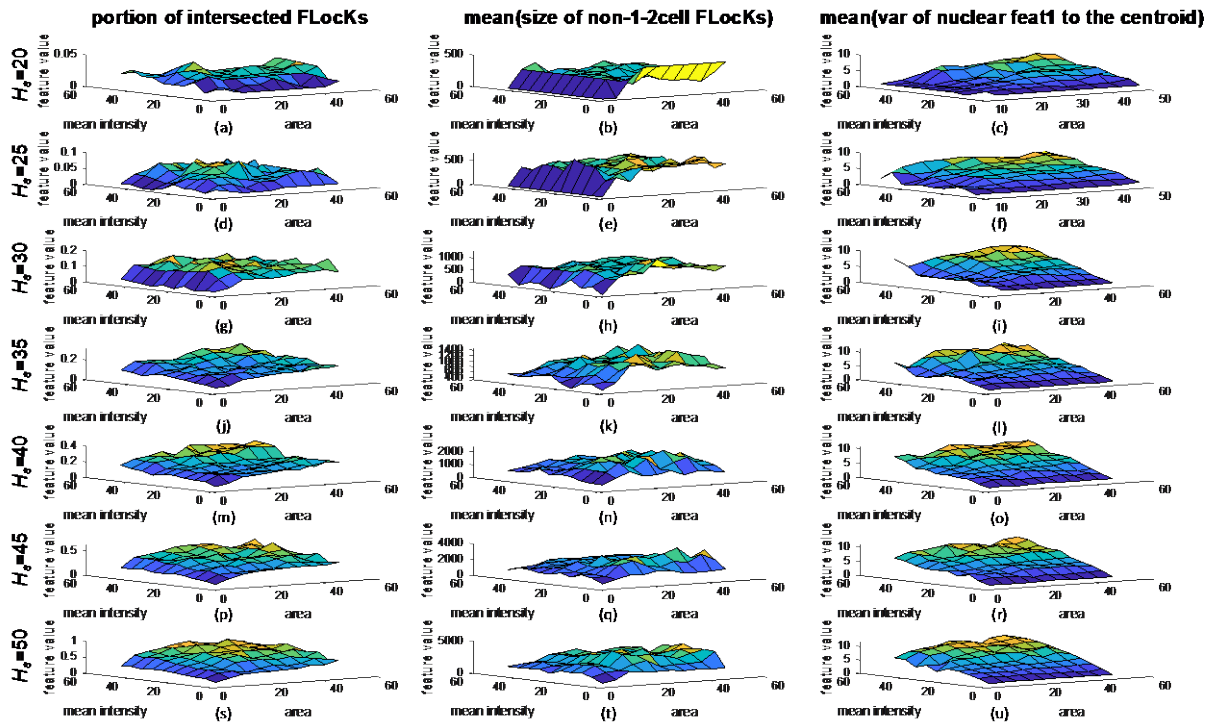

Figure S2: Three fLocK feature values with respect to bandwidths for space (h_s) and nuclear morphology (h_m). In this case, we chose feature space 'centroid-area-mean intensity' for the analysis. This figure shows three different FLocK features: *portion of intersected FLocKs, mean(size of non-102 cell FLocKs),* **and** *mean(var of nuclear feat1 to the centroid)***). The corresponding feature values are shown in (a)-(u), in which each row represents one space bandwidth h_s, from h_s=20 to h_s=50. The bandwidths for nuclear morphology, i.e., area and mean intensity, are the x and y axes for the mesh plot.**