HuBMAP Image Quality Control Metrics

Robert F. Murphy, Haoran Chen, Ted Zhang
Computational Biology Department
School of Computer Science
Carnegie Mellon University
HIVE Tools Component
Version 1.5 – September 8, 2021

SPRM saves quality control metrics in a JSON file for each processed image. In the list below, the underlined headings are the top-level tags in the JSON file. The headings in bold are the second-level tags. There may be one or more levels below that. Note that the order of the tags in the file is not necessarily the order below.

Image Information

Number of Channels

Image Quality Metrics not requiring image segmentation

Signal To Noise Otsu (per channel)

The Otsu method is used to choose an intensity threshold for each channel and the ratio of the average intensities above and below the threshold is calculated for each channel.

Signal to Noise Z-Score (per channel)

The ratio of the mean intensity to the standard deviation of intensity is calculated for each channel.

Total Intensity (per channel)

The sum of intensities is calculated for each channel.

Image Quality Metrics Requiring Background Segmentation

Fraction of Pixels in Image Background

See definition of background under Segmentation Quality

1/(AvgCVBackground+1)

The average coefficient of variation (CV) of the pixels in the background areas across all channels is calculated as

$$\mathrm{ACVB} = \frac{1}{n_h} \sum_{i=1}^{n_h} \frac{\sigma_{ib}}{\mu_i}$$

where n_h is the total number of image channels, σ_{ib} is the standard deviation of pixel intensities of i^{th} channel in the background, μ_i is the mean of all pixel intensities in i^{th} channel.

The reciprocal of ACVB+1 is used as the metric, so that higher values indicate better segmentation and its upper bound is 1.

FractionOfFirstPCBackground

As an alternative measure of similarity of background pixels,

$$\mathrm{FPCB} = \frac{\lambda_{b1}}{tr(\Sigma_b)}$$

Is calculated, where λ_{b1} is the variance explained by the first principle component across channels for the background pixels, Σ_b is the covariance matrices from PCA analysis on the background pixels across all channels after z-score standardization on each channel, $tr()$ is trace calculation on a matrix.

Image Quality Metrics that require cell segmentation

Number of Cells

Fraction of Image Occupied by Cells

Silhouette Scores From Clustering

The metrics below can be used to determine if the image is within the range of the expected number of clusters/cell types.

“Mean-All”

The silhouette scores for clustering into from 2 to 9 clusters.

“Number of Clusters with Max Score”

“Max Silhouette Score”

Channel Statistics

There is an entry for each channel for the two metrics below. These can be used either to identify outlier images or to assess whether values typical of a given channel for a given tissue have been obtained.

“Average per Cell Ratios”

‘Cell / Background’

Average intensity per cell divided by the average intensity of the background pixels

‘Nuclear / Cell’

Average intensity in the nucleus divided by the average intensity in the whole cell

Segmentation Evaluation Metrics

Matched Cell

The principle is that a nuclear mask should be entirely within the corresponding cell membrane mask. All cells that have corresponding nuclei are counted as matched. For

cells that have multiple corresponding nuclei, the one with the smallest fraction of mismatch pixels is kept. A matched cell mask is obtained by removing all unmatched cells. A matched nuclear mask is obtained by removing all unmatched nuclei and pixels outside the corresponding cell membrane for each nucleus. Some of the following metrics use this mask or the number of matched cells.

Each image was separated into regions consisting primarily of foreground and background. To do this, mean thresholding was applied to the nuclear, cytoplasmic, and cell membrane channels respectively, followed by performing two rounds of morphological closing with disk kernel of radius 1 and 10 respectively on three thresholded images combined. Each round consists of a foreground closing to bridge the small and large gaps between cells within the tissue and a background closing on the inverted image to reunite the overly scattered background. An area closing was subsequently applied and those areas that have less than 5000 pixels were considered to be the foreground. To correct the resulting rounded boundaries, a morphological geodesic active contour was applied on the inverted image with current background areas as seeds, followed by an area closing to remove the small dots in the background. If no information is given for which channels contain markers for nuclei, cytoplasm, and cell membrane, the same procedure was applied but thresholding was applied to all channels separately in the first step and the results from all channels combined by voting: a pixel is classified as foreground if it is foreground in at least 50% of the channels. In either case, the final output is a binary mask with 1 as foreground and 0 as background.

Given the binary foreground-background separation, the following two coverage metrics were designed with the general assumption that an accurate segmentation mask will cover most tissue areas (i.e., the foreground) but few background areas, along with other metrics with various assumptions.

"**FractionOfForegroundOccupiedByCells**":

$$\mathrm{FFC} = \frac{a_{cf}}{a_f}$$

where a_{cf} is the area of a cell mask in the foreground and a_f is the area of the foreground

"**1-FractionOfBackgroundOccupiedByCells**":

$$\mathrm{FBC} = \frac{a_{cb}}{a_b}$$

where a_{cb} is the area of a cell mask in the background and a_f is the area of the background.

1 minus FBC is used as the metric, so that larger values indicate better segmentation.

"**FractionOfCellMaskInForeground**":

$$\mathrm{FCF} = \frac{a_{cf}}{a_c}$$

where a_{cf} is the area of a cell mask in the foreground and a_c is the area of the cell mask

“**FractionOfMatchedCellsAndNuclei**": ,

The fraction of matched cells and nuclei is calculated, based on the assumption that in well-segmented images segmented cells and nuclei should have a one-to-one correspondence relationship.

$$\mathrm{FMCN} = \frac{n_m}{n_{cmi} + n_{nmi} + n_m}$$

where n_m is the number of matched cells and nuclei, n_{cmi} is the number of mismatched cells, n_{nmi} is the number of mismatched nuclei.

"**NumberOfCellsPer100SquareMicrons**":

The principle is that cell density is a measure of the quality of both the image and the segmentation method. The pixel size in square microns varies in different

image modalities. To calculate the density, pixel sizes in the X and Y dimensions are directly obtained from OME-TIFF metadata. The metric is defined as

$$\mathrm{NC} = \frac{n_c}{s_p * n_p} * 100$$

where n_c is the number of cells in the cell mask, s_p is the size of one pixel in squared microns, n_p is the total number of pixels in the corresponding image.

"**1/(AvgCVForegroundOutsideCells+1)**":

The principle is that the pixels outside of the cells but still within the image foreground should be similar in protein composition (i.e., should consist of extracellular matrix of similar composition). The coefficient of variation (CV) of the foreground pixels outside the cells for each channel was calculated and then the average CV across all channels was taken.

$$\mathrm{ACVF} = \frac{1}{n_h} \sum_{i=1}^{n_h} \frac{\sigma_{if}}{\mu_i}$$

where n_h is the total number of image channels, σ_{if} is the standard deviation of pixel intensities of the i^{th} channel in the foreground outside the cells, μ_i is the mean of all pixel intensities in the i^{th} channel.

The reciprocal of ACVF+1 is used as the metric, so that higher values indicate better segmentation and its upper bound is 1.

"**FractionOfFirstPCForegroundOutsideCells**":

Principal Component Analysis (PCA) on the matrix of all foreground outside-the-cell pixel intensities across all channels after z-score standardization on each channel was applied and the fraction of variance explained by the first principal component was calculated. A higher value of the fraction stands for a more conserved relationship across channels in the foreground outside the cell areas.

$$\mathrm{FPCF} = \frac{\lambda_{f1}}{tr(\Sigma_f)}$$

where λ_{f1} is the variance explained by the first principle component across channels for the foreground pixels outside the cells, Σ_f is the covariance matrices from PCA analysis on the foreground pixels outside the cells across all channels, $tr()$ is trace calculation on a matrix.

"1/(ln(StandardDeviationOfCellSize)+1)":

This is based on assumption that properly segmented images should have similarly sized cells.

$$\mathrm{CSSD} = \sqrt{\frac{\sum_i (s_i - \mu_s)}{n_c}}$$

where s_i is the size of i^{th} cell, μ_s is the mean size of all cells, n_c is the total number of cells.

The reciprocal of ln(CSSD)+1 is used as the metric, so that higher values indicate better segmentation and its upper bound is 1. The natural logarithm is taken to close the large difference of CSSD between methods.

Cell Not Including Nucleus (cell membrane plus cytoplasm)

Nucleus (including nuclear membrane)

To evaluate the segmentation performance on each cellular component, two cellular component masks were derived from the original matched whole-cell and nuclear masks. The Cell Not Including Nucleus mask was calculated by removing the nuclear mask from the cell mask. Each pixel in a given cell thus belongs to only one cellular compartment mask. For each cellular compartment mask, the following 3 metrics were calculated.

"1/(AvgOfWeightedAvgCVMeanCellIntensitiesOver1~10NumberOfClusters+1)":

The principles are that (a) cell types should be roughly similar in composition across all channels, (b) cell types can be approximated by cell clusters, (c) the

average CV across all clusters for a given number of clusters is a proxy for similarity in composition within cell types, and (d) averaging the average CV across all clusters across different numbers of the cluster is also a proxy for similarity in composition within cell types. Cell types were defined using KMeans clustering performed on the mean cell intensities across all channels after applying z-score standardization for each channel. Since the number of cell types in an arbitrary image is unknown, K, the number of clusters, was varied from 1 to 10. Note that the KMeans clustering was applied on the cell mask to define the cell type, which generated cell type labels of each number of clusters used in the following metric calculation for each cellular component mask.

For the clusters from each K value, the average CV for each cluster across channels was calculated followed by an average weighted by the cluster size. The final metric was calculated as the average of the weighted average coefficient of variation across all clusters over 1 to 10 clusters.

$$\mathrm{ACVC} = \frac{1}{K_{max} - K_{min} + 1} \sum_{K=K_{min}}^{K_{max}} \left(\frac{1}{n_c} \sum_{k=1}^{K} (CV)_k \cdot n_k\right)$$

$$(CV)_k = \frac{1}{n_h} \sum_{i=1}^{n_h} \frac{\sigma_{ki}}{\mu_{ki}}$$

where K_{min} is the smallest number of clusters, K_{max} is the largest number of clusters, K is the current number of clusters, n_c is the total number of cells, CV_k is the coefficient of variation of mean cell intensities in k^{th} cluster, σ_{ki} and μ_{ki} is the standard deviation and the mean of mean cell intensities of k^{th} cluster in i^{th} channel respectively, n_k is the number of cells in k^{th} cluster.

The reciprocal of ACVC+1 is used as the metric, so that higher values indicate better segmentation and its upper bound is 1.

"**AvgOfWeightedAvgFractionOfFirstPCMeanCellIntensitiesOver1~10NumberOfClusters**": ,

A similar measure was derived using principal component analysis. The first two principles are the same as above but (c) the fraction of variance accounted for by the first principal component of the cells in each cluster is a proxy for similarity in composition of each cell type, and (d) averaging this fraction over different numbers of clusters is also a proxy for similarity of each cell type, We, therefore, applied PCA on the mean cell intensities across z-score standardized channels of each number of cluster and calculated the average fraction of variance accounted for by the first principal component across all numbers of clusters. The output vector was averaged over different K values to get the final metric.

$$\mathrm{FPCC} = \frac{1}{K_{max} - K_{min} + 1} \sum_{K=K_{min}}^{K_{max}} (\frac{1}{n_c} \sum_{k=1}^{K} \frac{\lambda_{k1}}{tr(\Sigma_k)} \cdot n_k)$$

where K_{min} is the smallest number of clusters, K_{max} is the largest number of clusters, K is the current number of clusters of choice, n_c is the total number of cells, λ_{k1} is the variance explained by the first principle component in k^{th} cluster, Σ_k the covariance matrices from PCA on mean cell intensities across all channels in k^{th} cluster, n_k is the number of cells in the current cluster.

"**AvgSilhouetteOver2~10NumberOfClusters**": ,

$$\mathrm{AS} = \frac{1}{K_{max} - K_{min} + 1} \sum_{K=K_{min}}^{K_{max}} (\frac{1}{n_c} \sum_{i=1}^{n_c} \frac{b(i) - a(i)}{\max\{a(i), b(i)\}})$$

$$a(i) = \frac{1}{|K_i| - 1} \sum_{j \in K_i, i \neq j} d(i, j)$$

$$b(i) = \min_{l \neq i} \frac{1}{|K_l|} \sum_{j \in K_l} d(i, j)$$

where K_{min} is the smallest number of clusters, K_{max} is the largest number of clusters, K is the current number of clusters of choice, $a(i)$ is the average distance between cell i and all the other cells j in the cluster K_i to which cell i belongs, $b(i)$ is the minimum average distance from cell i to all cells j in all clusters K_l to which i does not belong, n_c is the total number of cells.

QualityScore

The method above yields 14 metrics for one segmentation method on an individual image. A PCA model was trained using these metrics for 603 images from 4 imaging modalities (CODEX, Cell DIVE, MIBI, IMC) after segmentation by 11 segmentation methods (Deepcell 0.9.0 with cell membrane, Deepcell 0.9.0 with cytoplasm, Cellpose 0.6.1, DeepCell 0.8.0 with cell membrane, DeepCell 0.8.0 with cytoplasm, Cellpose 0.0.3.1, CellProfiler, AICS classic, Cellsegm, CellX, Voronoi). The first principal component accounted for most of the variation across these images. The second principal component which accounted for the second most of the variation also represented the overall coverage of a segmentation mask on the image. We, therefore, defined the overall quality score as the variation-weighted sum of the top 2 principal components. For each image analyzed by SPRM, the vector of the 14 individual metrics is transformed using the learned PCA model to yield this overall quality score.