Electronic Supplementary Material (ESI) for Integrative Biology.

Title:

Morphology-based Prediction of Cancer Cell Migration Using Artificial Neural Network and Random Decision Forest

Authors and affiliations:

Zhixiong Zhang[†1], Lili Chen[†1], Brock Humphries[2], Riley Brien[1], Max S. Wicha[3, 4], Kathryn E. Luker[2], Gary D. Luker[2, 3, 6], Yu-Chih Chen[*1, 3, 4], and Euisik Yoon[*1, 6]

[1] Department of Electrical Engineering and Computer Science, University of Michigan, 1301 Beal Avenue, Ann Arbor, MI 48109-2122;

[2] Center for Molecular Imaging, Department of Radiology, University of Michigan, 109 Zina Pitcher Place, Ann Arbor, MI 48109-2200, USA;

[3] Comprehensive Cancer Center, University of Michigan, 1500 E. Medical Center Drive, Ann Arbor, MI 48109, USA;

[4] Forbes Institute for Cancer Discovery, University of Michigan, 2800 Plymouth Rd., Ann Arbor, MI 48109, USA;

[5] Department of Microbiology and Immunology, University of Michigan, 109 Zina Pitcher Place, Ann Arbor, MI 48109-2200, USA;

[6] Department of Biomedical Engineering, University of Michigan, 2200 Bonisteel, Blvd. Ann Arbor, MI 48109-2099, USA

[†] These authors contributed equally to this work.

* Corresponding author

Yu-Chih Chen
1301 Beal Avenue, Ann Arbor, MI 48109-2122, USA
Tel: 734-272-7113; E-mail: yuchchen@umich.edu.

Euisik Yoon
1301 Beal Avenue, Ann Arbor, MI 48109-2122, USA
Tel: 734-615-4469; E-mail: esyoon@umich.edu.

Supplementary Figures

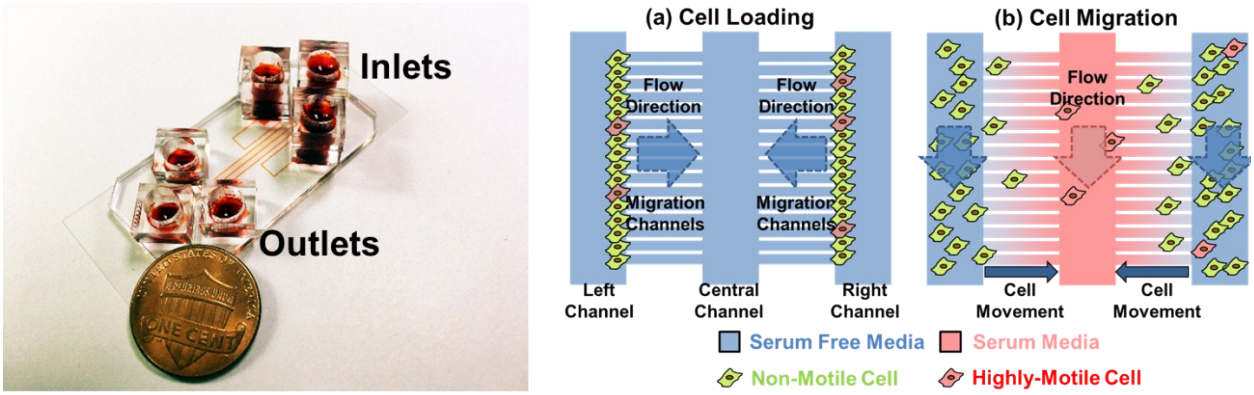

Figure S1. Microfluidic migration chip overview. (a) Photograph of a migration device with a one-cent coin, depicting the 3 inlets and 3 outlets reservoirs with red dye flowing through each device. The serpentine loading channels can be seen running vertically, with the horizontal migration channels running perpendicularly between them. (b) Cells were loaded along both left side and right side of the vertical cell loading channels and migrate through the horizontal migration channels due to chemoattraction.

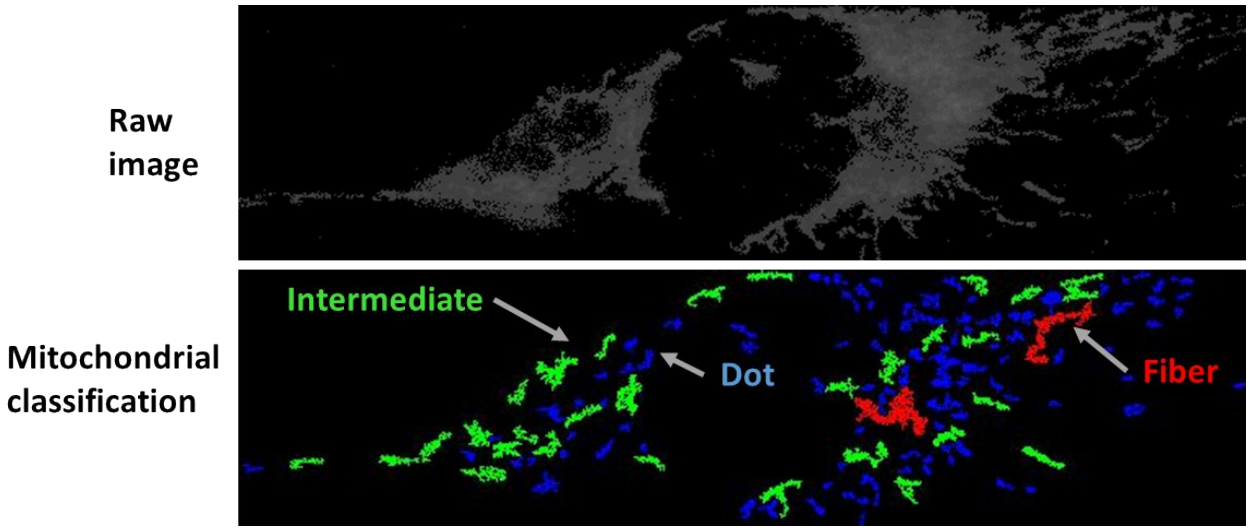

Figure S2. Sample microscopy image of out-of-focus mitochondria and classification. (a) Fluorescent image of mitochondria. (b) Mitochondrial classification of the out-of-focus image, showing low fiber ratio.

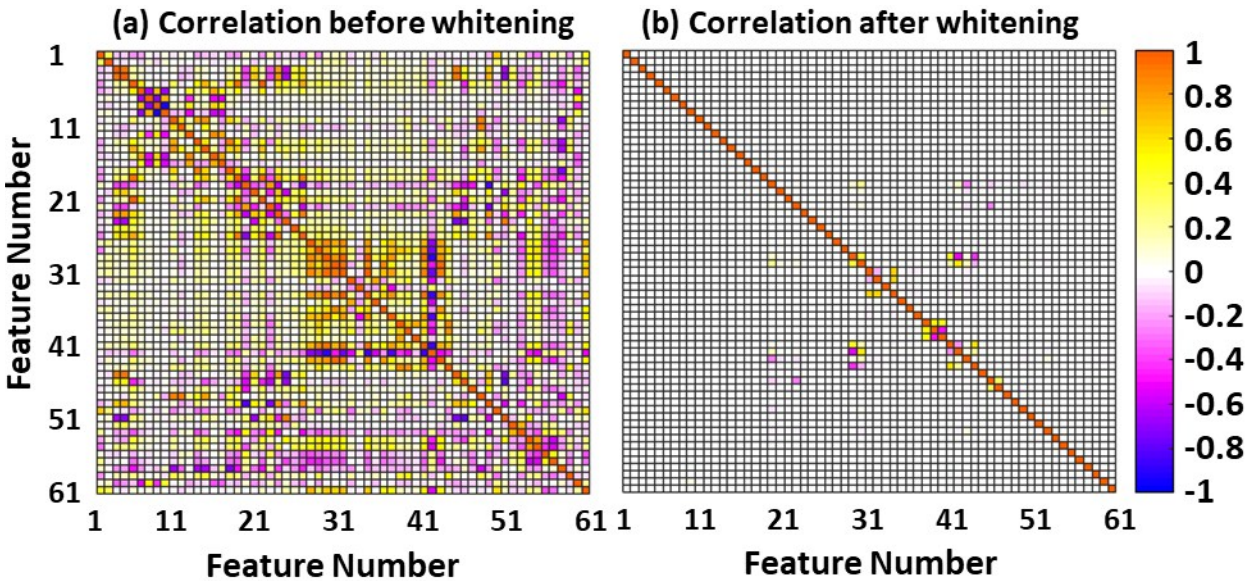

Figure S3. Morphological feature correlation heatmap (a) before whitening transformation, (b) after whitening transformation

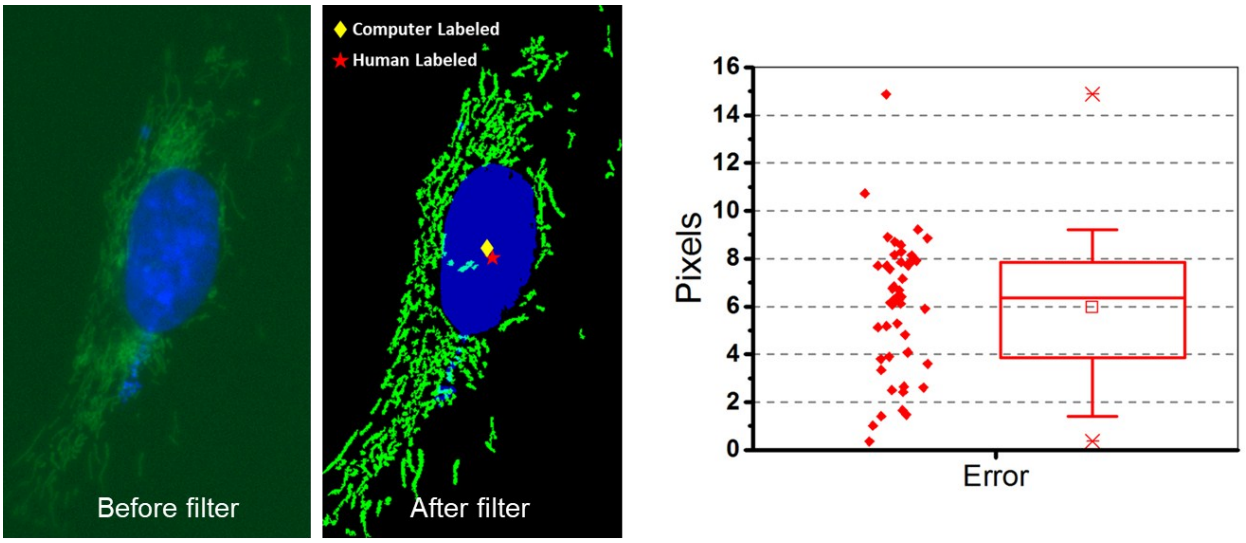

Figure S4. Comparison between human labeled nucleus center and computer labeled nucleus center. (a) Overlaid images of mitochondria (green) and DAPI-stained cell nucleus (blue). (b) Nucleus was labelled by human (shown as yellow diamond after applying thresholding filter of raw mitochondria, while computer calculated the center of mass (shown in red star) based on DAPI staining image. (b) scattering plot showing the distance error between human labeled nucleus center and computer labeled nucleus center, with an average around 6 pixels.

Supplementary information: Cellular morphology features

1. CenterShift

CenterShift is defined as the logarithm of the ratio between the distance of front frontier to nucleus center and the distance of rear frontier to nucleus center. Front frontier and rear frontier are defined based on cell moving direction.

2. FiberUpDownRatio
   FiberUpDownRatio is defined as the logarithm of the ratio between total fiber area of the front half of a cell and total fiber area of the rear half of a cell. Front half and rear half of a cell are determined by the nucleus of a cell and cell moving direction.
3. AspectRatioGreen
   AspectRatioGreen is defined as the ratio of length and width of a single cell under FITC channel.
4. AspectRatioRed
   AspectRatioRed is defined as the ratio of length and width of a single cell under TRITC channel.
5. RedTotalArea
   RedTotalArea is defined as the total cell area under TRITC channel.
6. GreenTotalArea
   GreenTotalArea is defined as the total mitochondria area under FITC channel.
7. GreenRedAreaRatio
   GreenRedAreaRatio is defined as the ratio between GreenTotalArea and RedTotalArea.
8. FiberRatio
   FiberRatio is defined as the ratio of fiber areas over all mitochondria areas.
9. DotRatio
   DotRatio is defined as the ratio of dot areas over all mitochondria areas.
10. Prctile99
    Prctile99 is defined as the value of 99^{th} percentile pixel brightness for a cell under TRITC channel.
11. Prctile50
    Prctile50 is defined as the value of 50^{th} percentile pixel brightness for a cell under TRITC channel.
12. GreenPrctile50
    GreenPrctile50 is defined as the value of 50^{th} percentile pixel brightness for a cell under FITC channel.
13. Prctile99Prctile50Ratio
    Prctile99Prctile50Ratio is defined as the ratio between Prctile99 and Prctile50.
14. TotalIntensityRed
    TotalIntensityRed is defined as the sum of intensity of a cell under TRITC channel.
15. TotalIntensityGreen
    TotalIntensityGreen is defined as the sum of intensity of a cell under FITC channel.
16. TotalIntensityRedGreenRatio
    TotalIntensityRedGreenRatio is defined as the ratio between TotalIntensityGreen and TotalIntensityRed.
17. CenterThickness
    CenterThickness is defined as the ratio between the average intensity over the nucleus region (15 pixels*15 pixels around the nucleus center) and Prctile50.
18. ConvexArea
    ConvexArea is defined as the value obtained from “ConvexArea” of “regionprops” in *MATLAB* for cell area under TRITC channel.
19. Solidity
    Solidity is defined as the value obtained from “Solidity” of “regionprops” in *MATLAB* for cell area under TRITC channel.
20. Eccentricity

Eccentricity is defined as the value obtained from "Eccentricity" of "regionprops" in *MATLAB* for cell area under TRITC channel.

21. Equivdiameter
    Equivdiameter is defined as the value obtained from "Equivdiameter" of "regionprops" in *MATLAB* for cell area under TRITC channel.
22. Extent
    Extent is defined as the value obtained from "Extent" of "regionprops" in *MATLAB* for cell area under TRITC channel.
23. MajorAxis
    MajorAxis is defined as the value obtained from "MajorAxisLength" of "regionprops" in *MATLAB* for cell area under TRITC channel.
24. MinorAxis
    MinorAxis is defined as the value obtained from "MinorAxisLength" of "regionprops" in *MATLAB* for cell area under TRITC channel.
25. Orientation
    Orientation is defined as the absolute value obtained from "Orientation" of "regionprops" in *MATLAB* for cell area under TRITC channel.
26. Perimeter
    Perimeter is defined as the value obtained from "Perimeter" of "regionprops" in *MATLAB* for cell area under TRITC channel.
27. MaxWidth
    MaxWidth is defined as the logarithm of the ratio between the max width of top 30% of the front half of a cell and the max width of back 30% of the rear half of a cell. Front half and rear half of a cell are determined by the nucleus of a cell and cell moving direction. Max width is defined as the distance between the leftmost point and the rightmost point for a row of a cell.
28. MaxWidthSum
    MaxWidthSum is defined as the logarithm of the ratio between the max width sum of top 30% of the front half of a cell and the max width sum of back 30% of the rear half of a cell. Front half and rear half of a cell are determined by the nucleus of a cell and cell moving direction. Max width sum is defined as the sum of pixels between the leftmost point and the rightmost point for a row of a cell.
29. TotalAreaRatio
    TotalAreaRatio is defined as the logarithm of the ratio between the area of top 30% of the front half of a cell and the area of back 30% of the rear half of a cell. Front half and rear half of a cell are determined by the nucleus of a cell and cell moving direction.
30. TotalPerimeterRatio
    TotalPerimeterRatio is defined as the logarithm of the ratio between the perimeter of top 30% of the front half of a cell and the perimeter of back 30% of the rear half of a cell. Front half and rear half of a cell are determined by the nucleus of a cell and cell moving direction.
31. TopDownConvexAreaRatio
    TopDownConvexAreaRatio is defined as the logarithm of the ratio between the convex area of top 30% of the front half of a cell and the convex area of back 30% of the rear half of a cell. Front half and rear half of a cell are determined by the nucleus of a cell and cell moving direction. Convex area is obtained from "ConvexArea" of "regionprops" in *MATLAB*.
32. TopDownSolidityRatio
    TopDownSolidityRatio is defined as the logarithm of the ratio between the solidity of top 30% of the front half of a cell and the solidity of back 30% of the rear half of a cell. Front half and rear half

of a cell are determined by the nucleus of a cell and cell moving direction. Solidity is obtained from "Solidity" of "regionprops" in *MATLAB*.

33. TopDownEccentricityRatio
    TopDownEccentricityRatio is defined as the logarithm of the ratio between the eccentricity of top 30% of the front half of a cell and the eccentricity of back 30% of the rear half of a cell. Front half and rear half of a cell are determined by the nucleus of a cell and cell moving direction. Eccentricity is obtained from "Eccentricity" of "regionprops" in *MATLAB*.
34. TopDownEquivdiameterRatio
    TopDownEquivdiameterRatio is defined as the logarithm of the ratio between the equivdiameter of top 30% of the front half of a cell and the equivdiameter of back 30% of the rear half of a cell. Front half and rear half of a cell are determined by the nucleus of a cell and cell moving direction. Equivdiameter is obtained from "Equivdiameter" of "regionprops" in *MATLAB*.
35. TopDownExtentRatio
    TopDownExtentRatio is defined as the logarithm of the ratio between the extent of top 30% of the front half of a cell and the extent of back 30% of the rear half of a cell. Front half and rear half of a cell are determined by the nucleus of a cell and cell moving direction. Extent is obtained from "Extent" of "regionprops" in *MATLAB*.
36. TopDownMajorAxisRatio
    TopDownMajorAxisRatio is defined as the logarithm of the ratio between the major axis length of top 30% of the front half of a cell and the major axis length of back 30% of the rear half of a cell. Front half and rear half of a cell are determined by the nucleus of a cell and cell moving direction. Major axis length is obtained from "MajorAxisLength" of "regionprops" in *MATLAB*.
37. TopDownMinorAxisRatio
    TopDownMinorAxisRatio is defined as the logarithm of the ratio between the minor axis length of top 30% of the front half of a cell and the minor axis length of back 30% of the rear half of a cell. Front half and rear half of a cell are determined by the nucleus of a cell and cell moving direction. Minor axis length is obtained from "MinorAxisLength" of "regionprops" in *MATLAB*.
38. TopDownPrctile99
    TopDownPrctile99 is defined as the logarithm of the ratio between Prctile99 of top 30% of the front half of a cell and Prctile99 of back 30% of the rear half of a cell. Front half and rear half of a cell are determined by the nucleus of a cell and cell moving direction.
39. TopDownPrctile50
    TopDownPrctile50 is defined as the logarithm of the ratio between Prctile50 of top 30% of the front half of a cell and Prctile50 of back 30% of the rear half of a cell. Front half and rear half of a cell are determined by the nucleus of a cell and cell moving direction.
40. TopDownPrctile99Prctile50Ratio
    TopDownPrctile99Prctile50Ratio is defined as the logarithm of the ratio between Prctile99Prctile50Ratio of top 30% of the front half of a cell and Prctile99Prctile50Ratio of back 30% of the rear half of a cell. Front half and rear half of a cell are determined by the nucleus of a cell and cell moving direction.
41. TopDownTotalIntensityRatio
    TopDownTotalIntensityRatio is defined as the logarithm of the ratio between the total intensity of top 30% of the front half of a cell and the total intensity of back 30% of the rear half of a cell. Front half and rear half of a cell are determined by the nucleus of a cell and cell moving direction.
42. TopDownPerimeterAreaRatio
    TopDownPerimeterAreaRatio is defined as the logarithm of the ratio between the perimeter area ratio of top 30% of the front half of a cell and the perimeter area ratio of back 30% of the rear half

of a cell. Front half and rear half of a cell are determined by the nucleus of a cell and cell moving direction.

43. TopDownPerimeterBoundingboxRatio
    TopDownPerimeterBoundingboxRatio is defined as the logarithm of the ratio between the perimeter boundingbox ratio of top 30% of the front half of a cell and the perimeter boundingbox ratio of back 30% of the rear half of a cell. Front half and rear half of a cell are determined by the nucleus of a cell and cell moving direction. Boundingbox is defined as the minimum rectangle which contained the target area of a cell.
44. TopDownMeanIntensityRatio
    TopDownMeanIntensityRatio is defined as the logarithm of the ratio between the mean intensity of top 30% of the front half of a cell and the mean intensity of back 30% of the rear half of a cell. Front half and rear half of a cell are determined by the nucleus of a cell and cell moving direction.
45. MajorMinorAxisRatio
    MajorMinorAxisRatio is defined as the ratio between MajorAxis and MinorAxis.
46. PerimeterAreaRatio
    PerimeterAreaRatio is defined as the ratio between the perimeter of a cell under TRITC channel and the area of a cell under TRITC channel.
47. PerimeterBoundingboxRatio
    PerimeterBoundingboxRatio is defined as the ratio between the perimeter of a cell under TRITC channel and the perimeter of the bounding box for that cell. Bounding box is defined as the minimum rectangle which contained the target cell.
48. MeanIntensity
    MeanIntensity is defined as the ratio between total intensity of a cell and total area of a cell under TRITC channel.
49. PerimeterSquareAreaRatio
    PerimeterSquareAreaRatio is defined as the ratio between square of the perimeter of a cell and the area of that cell.
50. TotalPerimeterHalfRatio
    TotalPerimeterHalfRatio is defined as the ratio of the perimeter of the front half of a cell and the perimeter of the rear half of a cell. Front half and rear half of a cell are determined by the nucleus of a cell and cell moving direction.
51. AverageWidth
    AverageWidth is defined as the average width of a cell under TRITC channel. Width of a cell is defined as the distance between the leftmost pixel and rightmost pixel in a row.
52. LeftRightBound
    LeftRightBound is defined as the maximum of lengths of a column 10 pixels to the left/right bound. Length of a column is defined as the distance between the highest pixel and the lowest pixel in a column.
53. LeftRightBoundRatio
    LeftRightBoundRatio is defined as the ratio between LeftRightBound and the length of a cell under TRITC channel.
54. HeadAverageWidth
    HeadAverageWidth is defined as the average width of front one-sixth of a cell. Front one-sixth of a cell is determined by cell moving direction. Width of a cell is defined as the distance between the leftmost pixel and rightmost pixel in a row.
55. HeadAverageWidthRatio
    HeadAverageWidthRatio is defined as the ratio between HeadAverageWidth and AverageWidth.

56. CenterRed
    CenterRed is defined as the ratio between the width of the center of a cell and AverageWidth. Center of a cell is determined by the nucleus. Width of a cell is defined as the distance between the leftmost pixel and rightmost pixel in a row.
57. RedShift
    RedShift is defined as the logarithm of the ratio between the distance from the row of widest width of a cell to the front row of a cell and the distance from the row of widest width of a cell to the rear row of a cell. Front row and rear row of a cell are determined by cell moving direction.
58. RedFoot
    RedFoot is defined as the ratio of the average intensity of the front foot area and the average intensity of the nucleus area under TRITC channel. The front foot area is defined as the front 1-25 pixels near the contour of a cell in each column. The nucleus area is defined as a 15 pixels*15 pixels square centered with the nucleus.
59. CenterMito
    CenterMito is defined as the total mitochondria area in nucleus center region under FITC channel. Nucleus center region is defined as a 41 pixels*41 pixels square centered with the nucleus.
60. RedGreenDist
    RedGreenDist is defined as the ratio between the distance from the front of a cell under TRITC channel to the front of mitochondria under FITC channel and the length of a cell.
61. TotalAreaHalfRatio
    TotalAreaHalfRatio is defined as the ratio of the area of the front half of a cell and the area of the rear half of a cell. Front half and rear half of a cell are determined by the nucleus of a cell and cell moving direction.