Machine Learning Classification of Cancer Cells Migration in 3D Multi-cue Microenvironments[1]

K. Esmaeili Pourfarhangi[1] and B. Gligorijevic[1,2]

1. Bioengineering Department, College of Engineering, Temple University, Philadelphia, PA, USA
2. Cancer Biology Program, Fox Chase Cancer Center, PA, USA

{Kamyar.esmaeili, Bojana.gligorijevic} @temple.edu

Metastasis is the leading cause of deaths among cancer patients. Metastatic dissemination, during which cancer cells from the primary tumor reach the secondary organ, consists of a cascade of events, starting with cancer cell invasion and migration through the surrounding tissue [1-3]. During cancer cells invasion, tumor cell movement is directed by multiple guiding cues present in the tissue at different concentrations [1,4]. Guiding cues can be biophysical, such as aligned collagen fibers inducing contact guidance or biochemical, such as gradients of chemoattractants inducing chemotaxis. Following individual cues will lead to directed cell migration, where the cell velocity and cell persistence will be determined by the concentration of the cue and the presence of the cell receptors that are responsible for communication with the particular cue. However, in complex microenvironments, where multiple cues are present at different levels, distances and orientations, cues can compete with or synergize each other to guide the cell migration. In such conditions, analyzing the cell migration parameters (velocity, persistence, directionality etc.) is a challenging task.

In this work, we cultured breast cancer cells in a 3D microchip (ibidi Chemotaxis μ-slides) containing collagen fibers. Fibers were aligned using magnetic beads [5]. In addition, a chemoattractant gradient was introduced, with direction perpendicular to the fiber alignment. This set up exposed cancer cells simultaneously to two cues, contact guidance cue, from aligned collagen fibers, and a competing chemotaxis cue, from the chemoattractant gradient. Cell migration in the microchip was monitored over 48 hours, using fluorescent microscopy. Images were then segmented, and cells tracked over time; cell tracking data was then used as input for a nonlinear support vector machine (SVM) classifier in MATLAB Classification Learner App [6].

Training of the SVM classifier was done by utilizing data from single-cue conditions, where either the fiber alignment or the chemoattractant gradient were present. 20 predictors were extracted from each cell's migration trajectory. Predictors are average cell persistence values (net distance migrated divided by total distance migrated) with 20 various frame intervals ranging from 2 to 21 frames:

$$P_i = \frac{\Sigma_{j=1:i:f-i} \frac{\sqrt{\left(x_{j+i}-x_j\right)^2+\left(y_{j+i}-y_j\right)^2}}{\Sigma_{n=j:1:j+i}\sqrt{(x_{n+1}-x_n)^2+(y_{n+1}-y_n)^2}}}{\left|\frac{f}{i}\right|}. \quad (1)$$

In (1), f represents the number of data points in each cell's trajectory, x and y are the coordinates of the cells in the cartesian system and the subscripts of x and y refer to the frame number in which

[1] Research reported in this work was supported by the National Cancer Institute of the National Institute of Health under award number K99CA172360. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute of Health.

the coordinates of a cell are being read. For example, if a cell is tracked for 200 frames, f is equal to 200 and x_{122} refers to the horizontal coordinate of the cell at frame number 122. i refers to the frame interval which varies from 2 to 21. The value of i defines the number of frames that would be skipped in order to calculate the net distance migrated by a cell.

Once the SVM was trained, it was tested on data extracted from the bi-cue conditions. In bi-cue microenvironments, machine learning approach successfully classified migrating cells into the contact-guidance and chemotaxis classes, with an accuracy of >90%. Interestingly, classic comparison of the average cell velocity or average migration persistence between the two classes of cells is not able to distinguish any difference between them, showing the importance of the automated, single cell-based approach to classification of the cell migration trajectories.

Results of this work highlights the possibility of using machine learning techniques to analyze migratory subpopulations of cancer cells and generate predictive models of metastasis. While our current setup is a bi-cue one, both the microchip as well as the SVM approaches are generalizable and can be used as multi-cue and multi-class methods. Analyzing cancer cell behavior in controlled microenvironments with increasing complexity will help in understanding the microenvironment role in guiding trajectories of the cancer cells that successfully metastasize. In turn, this can guide development of predictive models of metastasizing cells and preventive therapies against metastasis.

REFERENCES

[1] K. Esmaeili Pourfarhangi, E. Cardenas De La Hoz, A. R. Cohen, and B. Gligorijevic, "Contact guidance is cell cycle-dependent," *APL bioengineering*, vol. 2, no. 3, pp. 031904, 2018.

[2] K. E. Pourfarhangi, A. Bergman, and B. Gligorijevic, "ECM cross-linking regulates invadopodia dynamics," *Biophysical journal*, vol. 114, no. 6, pp. 1455-1466, 2018.

[3] K.E. Pourfarhangi, A. Bergman, and B. Gligorijevic, "Integrating live-cell fluorescent microscopy and signal processing to discover the relationship of invadopodia digging cycles with extracellular matrix crosslinking ratio," *Signal Processing in Medicine and Biology Symposium (SPMB)*, pp. 1-2, 2016.

[4] M. J. Oudin, and V. M. Weaver, "Physical and chemical gradients in the tumor microenvironment regulate tumor cell invasion, migration, and metastasis," *Cold Spring Harbor symposia on quantitative biology*, pp. 030817, 2016.

[5] K. Esmaeili Pourfarhangi, E. Cardenas De La Hoz, A. R. Cohen, and B. Gligorijevic, "A methodology to investigate the relationship between cancer cells cell-cycle phase and their migratory behaviors," *Optical Tomography and Spectroscopy*, pp. JTu3A-13, 2018.

[6] B. Bayarmagnai, L. Perrin, K. E. Pourfarhangi, and B. Gligorijevic, "Intravital Imaging of Tumor Cell Motility in the Tumor Microenvironment Context," in *Cell Migration*, A. Gautreau, Ed. New York: Humana Press, 2018, pp. 175-193.

Machine Learning Classification of Cancer Cells Migration in 3D Multi-cue Microenvironments

Kamyar Esmaeili Pourfarhangi[1], Bojana Gligorijevic[1,2]

[1]Bioengineering Department, College of Engineering, Temple University, Philadelphia, PA, USA

[2]Cancer Biology Program, Fox Chase Cancer Center, Philadelphia, PA, USA

Background

- Cancer invasion is one of the early steps of metastasis, during which cancer cells disseminate through the tumor associated tissue.
- Disseminating tumor cells are exposed to various guiding cues (e.g. aligned collagen fibers and gradients of chemoattractant molecules) stimulating them to migrate in a directed fashion.
- In complex microenvironments, where multiple cues are present at different levels, distances and orientations, cues can compete with or synergize each other to guide the cell migration.
- It is shown that cell migration parameters (velocity, persistence, directionality etc.) in response to different cues are different. However, in *in vivo* set ups where multiple cues affect cell migration simultaneously, it is impossible to find out which cue is affecting cell migration.

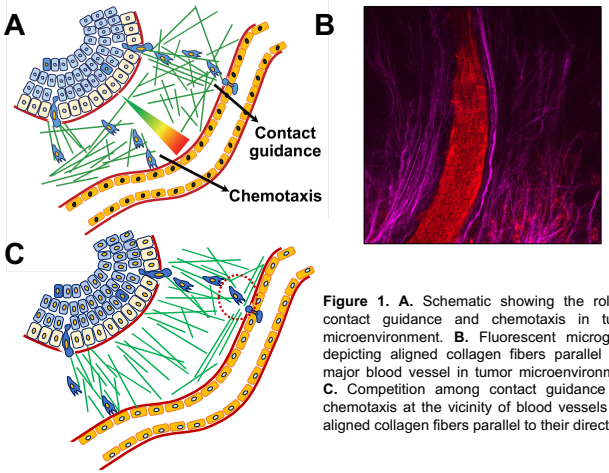


Figure 1. A. Schematic showing the role of contact guidance and chemotaxis in tumor microenvironment. **B.** Fluorescent micrograph depicting aligned collagen fibers parallel to a major blood vessel in tumor microenvironment. **C.** Competition among contact guidance and chemotaxis at the vicinity of blood vessels with aligned collagen fibers parallel to their direction.

- In this work, we embark on assessing the possibility of using machine learning classification algorithms on cell migration data captured from cells exposed to a bi-cue environment to indicate what guiding cue is the cells attracted toward.

Methodology & Results

1. 3D model of chemotaxis

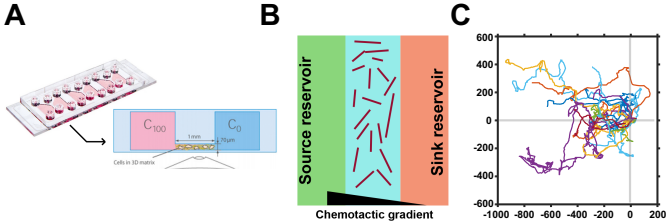


Figure 2. A. ibidi chemotaxis microslides utilized to create a linear fetal bovine serum (FBS) gradient across a 3D collagen gel in which HS-578T breast carcinoma cells are cultured. **B.** Schematic showing the direction of gradient forming across the collagen gel. **C.** Cell migration trajectories in the system. Cells are attracted towards the FBS source.

2. 3D model of contact guidance

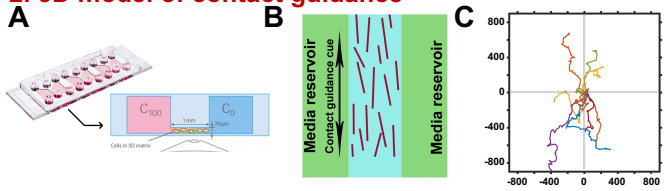
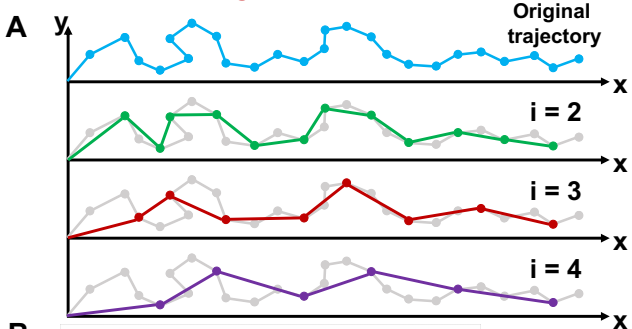


Figure 3. A. ibidi microslides utilized to create collagen alignment via incorporation of paramagnetic microbeads within the 3D collagen gel in which HS-578T breast carcinoma cells are cultured. **B.** Schematic showing the direction of fiber alignment in the collagen gel. **C.** Cell migration trajectories in the system. Cells depict contact guidance along the direction of fibers.

3. Extracting migration features from cell trajectories recorded from single-cue conditions



$$P_i = \frac{\sum_{j=1:i:f-i} \sqrt{(x_{j+i}-x_j)^2+(y_{j+i}-y_j)^2}}{\sum_{n=j:1:j+i} \sqrt{(x_{n+1}-x_n)^2+(y_{n+1}-y_n)^2}}, \quad i = 2{:}21$$

Figure 4. A. Feature extraction from individual cell's trajectories. **B.** Mathematical formula of feature calculation from the trajectories. The formula calculates average persistence of the cells at various sampling frequencies (i).

4. Using MATLAB machine learning classifier to find the best model distinguishing highest amount of differences between the features of the two groups

Average Error:	Tree (Fine)	Tree (Coarse)	Linear Discriminant	Logistic Regression	Linear SVM	Cubic SVM	Medium Gaussian SVM	Fine KNN	Coarse KNN
CG	13 ± 2.1	13 ± 1.9	17 ± 3.2	11 ± 1.2	10 ± 1.6	13 ± 1.9	7 ± 1.1	16 ± 2.1	38 ± 4.5
Ch	12 ± 2.4	12 ± 2.3	24 ± 2.9	25 ± 1.9	12 ± 1.8	12 ± 2.5	11 ± 1.5	19 ± 2.8	27 ± 2.3

Medium Gaussian SVM Predicted class

True class	CG	Ch
CG	93	11
Ch	7	89

5. 3D model of contact guidance-chemotaxis

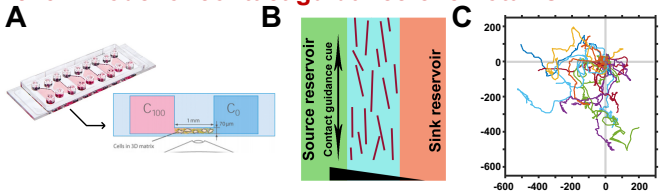


Figure 5. A. ibidi microslides utilized to create both FBS gradient and collagen alignment via at various directions within the 3D collagen gel in which HS-578T breast carcinoma cells are cultured. **B.** Schematic showing the direction of fiber alignment and FBS gradient. **C.** Cell migration trajectories in the system. Cells depict both contact guidance and chemotaxis.

6. Using the Medium Gaussian SVM to classify the cell trajectories recorded from the bi-cue conditions

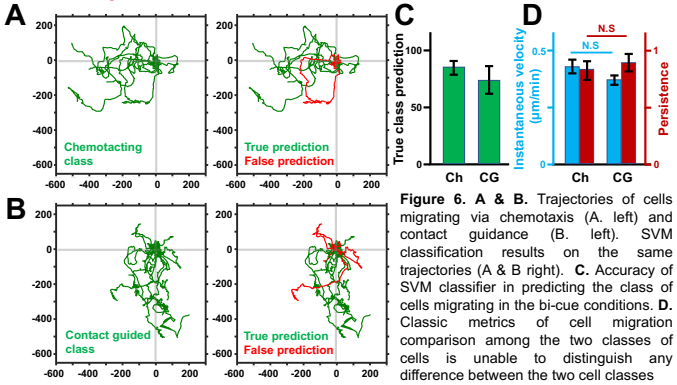


Figure 6. A & B. Trajectories of cells migrating via chemotaxis (A. left) and contact guidance (B. left). SVM classification results on the same trajectories (A & B right). **C.** Accuracy of SVM classifier in predicting the class of cells migrating in the bi-cue conditions. **D.** Classic metrics of cell migration comparison among the two classes of cells is unable to distinguish any difference between the two cell classes

Conclusion

- While classic metrics of cell migration do not recognize any difference among cells in bi-cue environments, a Medium Gaussian SVM classification algorithm can classify the cells into correct classes with more than 75 % accuracy.
- Such methods can be further extended and leveraged in analyzing in vivo data to recognize the specific guiding cues affecting cell motility

References

[1] K. Esmaeili Pourfarhangi, E. Cardenas De La Hoz, A. R. Cohen, and B. Gligorijevic, *APL bioengineering*, vol. 2, no. 3, pp. 031904, 2018.

[2] K. E. Pourfarhangi, A. Bergman, and B. Gligorijevic, *Biophysical journal*, vol. 114, no. 6, pp. 1455-1466, 2018.

[3] K.E. Pourfarhangi, A. Bergman, and B. Gligorijevic *Signal Processing in Medicine and Biology Symposium (SPMB)*, pp. 1-2, 2016.

[4] M. J. Oudin, and V. M. Weaver, *Cold Spring Harbor symposia on quantitative biology*, pp. 030817, 2016.

[5] K. Esmaeili Pourfarhangi, E. Cardenas De La Hoz, A. R. Cohen, and B. Gligorijevic, *Optical Tomography and Spectroscopy*, pp. JTu3A-13, 2018.

[6] B. Bayarmagnai, L. Perrin, K. E. Pourfarhangi, and B. Gligorijevic, in *Cell Migration*, A. Gautreau, Ed. New York: Humana Press, 2018, pp. 175-193.