Morphological clustering of cancer cell images

Ilmari Ahonen[1], Jaakko Nevalainen[1] and Matthias Nees[2]

[1]University of Turku, Finland, Department of Mathematics and Statistics

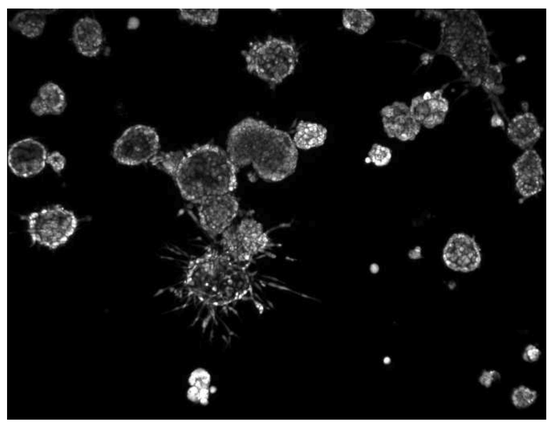
[2]VTT Technical Research Centre of Finland

Introduction

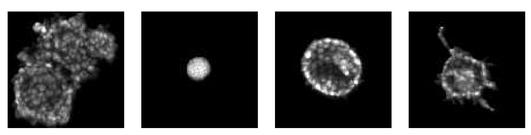
Cancer cells are grown in a fully 3-dimensional environment where they can move freely and form multicellular **spheroids**.

Multiple treatments can be tested and the effects are seen in the appearance of the formed spheroids.

We classify cell structures based on their appearance and thus create useful data for treatment effect assessment.

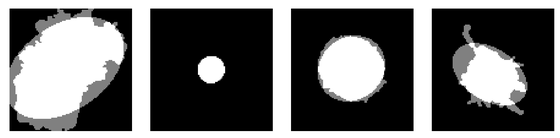


A 672×512 image taken from PC-3 cancer cells after 10 days in three-dimensional culture

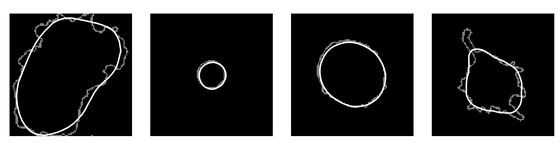


Four example cell structures

Shape features



The overall regularity of the shape can be assessed with an ellipse fit.



Principal curve fit captures roughness in smaller scale.

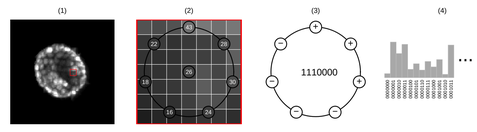
In addition to the general size of the objects, a lot of information relating to the behaviour of cells can be measured from the outlines of the structures

Rough and smooth surfaces, regular and irregular shapes and appendages are all local biologically relevant and interpretable features that can be linked to disease characteristics.

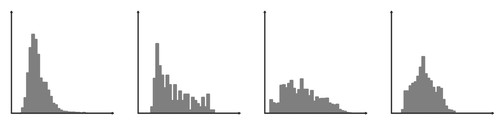
Texture Features

The appearance of the structure surface is in some cases an extremely informative feature for researches but is rarely addressed in existing cancer cell image analysis applications.

Texture operators such as central moments and LBPs typically produce a histogram for each image, whose bin densities are used as features.



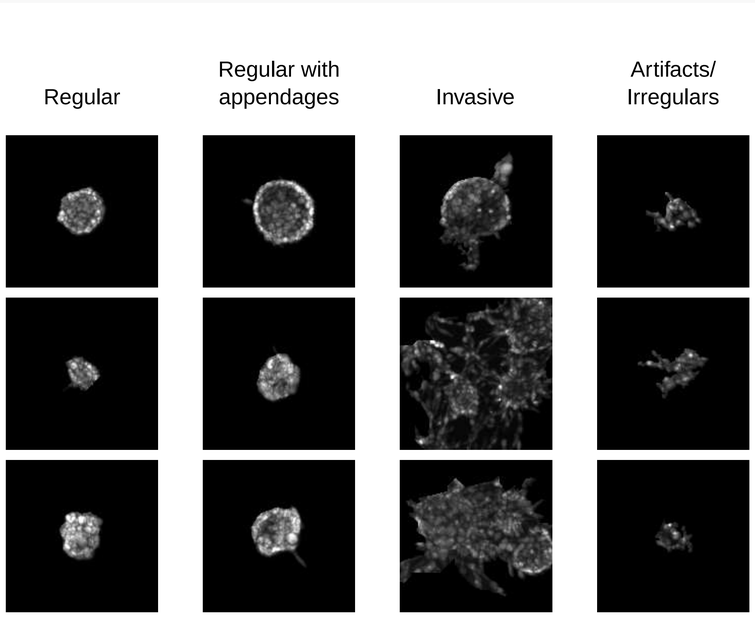
Local binary patterns (LBP) map images into a distribution of binary codes.



Any operators can be calculated from smaller windows within images. Here we have plotted the distributions of standard deviations calculated using 10×10 windows from our four example cell structures.

Clustering

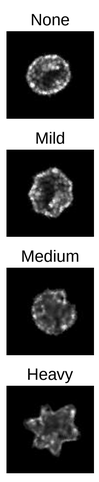
Cell structures are clustered based on the obtained shape and texture features. Before clustering, the dimension of the features is reduced by extracting a set of first principal components. Obtained cluster labels can then be used as response variables in subsequent treatment effect analyses

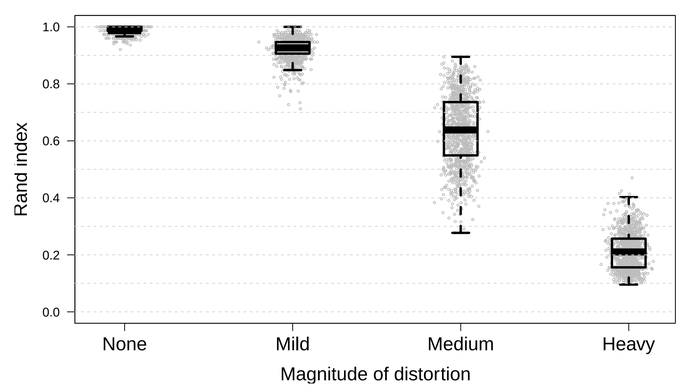


Examples of clusters that can be obtained

Simulation results

The robustness of our proposed analysis process was tested using an artificial dataset simulated from our four example cell structures with varying degree of distortions.





Four clusters were extracted from the simulated dataset and Rand index was used to measure the agreement between the cluster labels and the original structures.

Summary

- Structure phenotype can be summarized by a set of numerical features.
- Structures can be accurately grouped using a combination of dimension reduction and clustering
- Cluster labels can be used for evaluating and ranking treatments
- The computations are fairly fast and automated relative to the large mass of data

email: **ilmari.ahonen@utu.fi**, website: **www.tilastotieteenkeskus.fi**