Scale space classification of ovarian cancer transcriptomics data

Jeroen de Ridder^{1,*}, Theo Knijnenburg^{2,\$}, Ilya Shmulevich³, Lodewyk Wessels^{1,2}, Marcel Reinders¹

*j.deridder@tudelft.nl, ^{\$}t.knijnenburg@nki.nl

¹Delft Bioinformatics Lab, Faculty of Electrical Engineering, Mathematics and Computer Science, Delft University of Technology, Delft, The Netherlands. ²Bioinformatics and Statistics, Division of Molecular Biology, The Netherlands Cancer Institute, Amsterdam, The Netherlands, ³Institute for Systems Biology, Seattle, United States of America

The notion of scale plays an important role in how high-throughput biomolecular measurements are analyzed and interpreted. At one level of functional scale, for instance, all genes involved in the cell cycle may be considered as a functional unit, whereas a more fine grained scale would capture more specific cell-cycle related events. Capturing components that manifest themselves as functional units across different scales may reveal key biological mechanisms.

We propose to improve classification of gene expression data by exploiting relevant structure in the data at multiple scales. The method is based on concepts of scale space analysis in machine vision, used for scale independent detection of features in images.

This multiscale framework was applied to a large gene expression dataset of patients diagnosed with ovarian cancer, measured as part of the Cancer Genome Atlas Project (TCGA). Our goal was to build a classifier that accurately discriminates between patients with local recurrence and distant recurrence (metastasis) of the tumor after the primary tumor was surgically removed.





25NN Leave-one-out error per scale and per TF







Delft University of Technology