

Automated classification of circulating tumor cells and the impact of inter-observer variability on classifier training and performance.

Svensson C-M, Hübler R, Figge MT (2015) Automated classification of circulating tumor cells and the impact of inter-observer variability on classifier training and performance. *Journal of Immunology Research* 2015(Article ID 573165), 9.

[Details](#)



Abstract

Application of personalized medicine requires integration of different data to determine each patient's unique clinical constitution. The automated analysis of medical data is a growing field where different machine learning techniques are used to minimize the time consuming task of manual analysis. The evaluation, and often training, of automated classifiers requires manually labelled data as ground truth. In many cases such labelling is not perfect, either because of the data being ambiguous even for a trained expert or because of mistakes. Here we investigated the inter-observer variability of image data comprising fluorescently stained circulating tumor cells and its effect on the performance of two automated classifiers, a random forest and a support vector machine. We found that uncertainty in annotation between observers limited the performance of the automated classifiers, especially when it was included in the test set on which classifier

performance was measured. The random forest classifier turned out to be resilient to uncertainty in the training data while the support vector machine's performance is highly dependent on the amount of uncertainty in the training data. We finally introduced the consensus data set as a possible solution for evaluation of automated classifiers that minimizes the penalty of inter-observer variability.

Beteiligte Forschungseinheiten

[Angewandte Systembiologie](#) [Marc Thilo Figge](#) [Mehr erfahren](#)

Leibniz-HKI-Autor*innen



Marc Thilo Figge

[Details](#)



Ron Hübler

[Details](#)



Carl-Magnus Svensson

[Details](#)

Identifier

doi: 10.1155/2015/573165

PMID: 26504857