

# Cell Classifier for Biological Research

Studiengang: BSc in Informatik | Vertiefung: Computer Perception and Virtual Reality

Betreuer: Prof. Dr. Bernhard Anrig

Experte: Dr. René Bach

Industriepartner: Institut für Zellbiologie, Universität Bern, Bern

Biologen an der Universität Bern erforschen das Verhalten von einzelnen Zellen im Zusammenhang mit der Entstehung von Krebs. Ein Teil der beobachteten Zellen kann die Auswertung der Experimente jedoch verfälschen. In dieser Thesis wurden Verfahren zur Aussortierung dieser Zellen untersucht.

## Ausgangslage und Zielsetzung

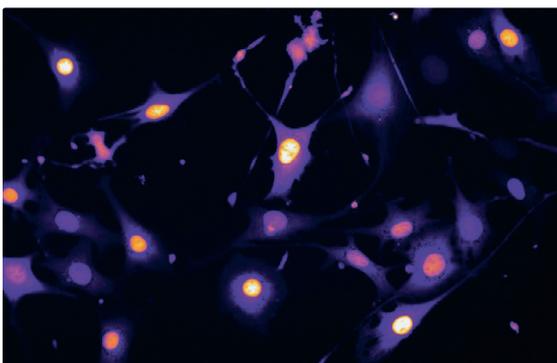
Biologen sind heute in der Lage, menschliche Zellen im Labor einzeln zu untersuchen und so Unterschiede zwischen den Zellen zu erforschen. Interessanterweise reagieren die Zellen oft unterschiedlich auf äussere Einflüsse, sogar wenn sie genetisch identisch sind. Dies ist in der Krebsforschung ein topaktuelles Thema: oft überleben einzelne Krebszellen nach einer medizinischen Behandlung und können dann neue Tumore bilden. Für diese Untersuchungen filmen die Forscher hunderte von Zellen unter dem Mikroskop. Daraus entstehen Zeitreihen, welche unterschiedliche Merkmale der Zellen über die Zeit wiedergeben (z. B. die Flächen der Zellen). Ein Teil dieser Zeitreihen kann die darauffolgende Auswertung jedoch verfälschen, z. B. weil sich Zellen während der Aufnahme teilen und dadurch falsche Messwerte entstehen.

Biologen investieren viel Zeit, um solche unbrauchbaren Zellen von Hand aus ihren Datensätzen auszusortieren. Das Ziel der Thesis war es, Verfahren zu implementieren, um die Aussortierung von unbrauchbaren Zellen zu automatisieren.

## Vorgehen

Es wurden drei unterschiedliche Ansätze getestet, um unbrauchbare Zellen auszusortieren:

- Hauptkomponentenanalyse der Zeitreihen,
- ein «sliding window»-Verfahren und
- maschinelles Lernen.



Menschliche Zellen unter dem Mikroskop

Als Beispiel sei hier das «sliding window»-Verfahren erläutert. Dabei wird ein «Fenster» über eine Zeitreihe gelegt und darüber verschoben (siehe Abbildung unten). An jeder Position wird der letzte Wert im Fenster mit allen Werten im Fenster verglichen um zu bestimmen, ob es ein Ausreisser ist (rote Linie in der Abbildung). Wenn die Anzahl Ausreisser in einer Zeitreihe einen Schwellwert übersteigt, wird die zugehörige Zelle als «nicht brauchbar» klassifiziert.

## Resultate

Die drei getesteten Verfahren haben eine ähnliche Effizienz. Ein Grossteil der nicht brauchbaren Zellen konnte aussortiert werden – jedoch nicht alle: eine kleine Teilmenge davon konnte nicht von den brauchbaren Zellen unterschieden werden. Beim «sliding window»-Verfahren beispielsweise wurde in etwa jede zehnte unbrauchbare Zelle nicht als solche erkannt.

Insgesamt konnte eine Genauigkeit der Klassifizierung von 85–90% erreicht werden, was ein guter Wert ist. Wenn diese Resultate mit den Zeitreihen von weiteren Experimenten wiederholt werden können, so eignen sich diese Verfahren für den produktiven Einsatz in der Forschung.



Steven René Cardini  
s.cardini@gmail.com



Das «sliding window»-Verfahren schematisch dargestellt