Cell Classifier

Introduction to object classification using graphic descriptors

Évariste DALLER

Université Caen Normandie

mars 2016

A Cell Classifier

Purpose

Given a set of pictures representing cells, we want to define classes (categories), based on observation or predefined, in which we put presented cells.

Example

For instance, we would like to count and describe different kind of cells in a blood sample image :

- Erythrocytes (Red blood cells)
- Lymphocytes
- Monocytes
- ...

A Cell Classifier

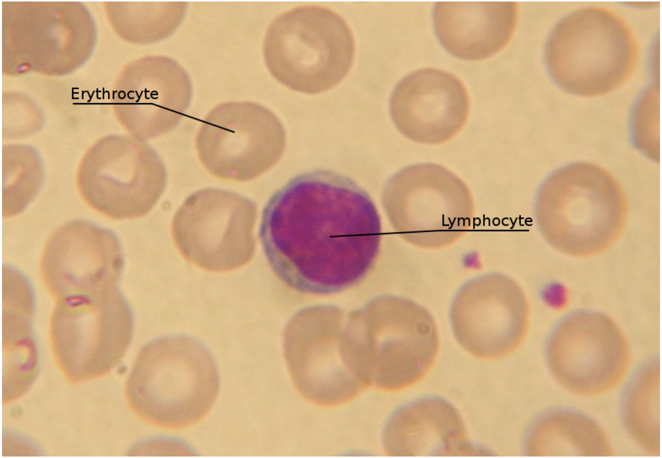

Figure: Kinds of blood cells

Classification : How it works

Descriptors

To allow a computer to do so, we translate the **qualitative** description of the objects into a **quantitative** one, using a vector of numbers, in order for the machine to make computations.

Detecting classes

The basic idea is to highlight gatherings of examples to be seen as classes.

Classifying

Once we have those groups, we compute the probability for a new example to be in one or another of these.

Classification : How it works

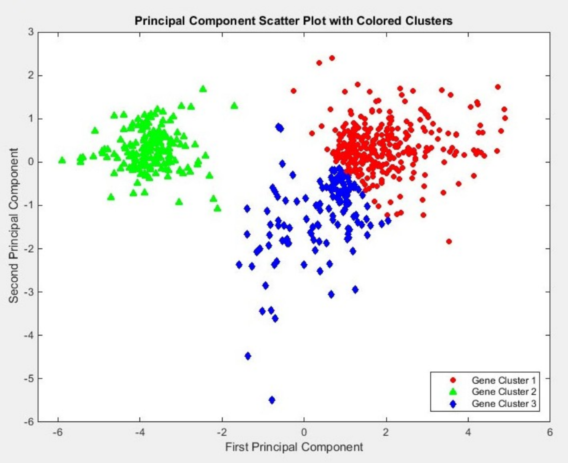

Figure: Example of a classification

Classification : Automatic Valuation

Defining classes

One way to define categories automatically from data (**data clustering**) is to do some statistical tests on the examples database in order to express clouds of points.

Examples of clustering methods

- *k*-means
- Expectation-maximisation algorithm (EM)
- hierarchical clustering

Classification : Supervised Learning

First step : constitute training set

We use a consistent-sized set of data (images here), as a training set. An oracle (expert, human) label cells according to different pre-defined classes.

Second step : classify

We use data from the training set to compute probability for a new example to be in one or another class.

Descriptors

While describing a sample set of images, we consider a relevant model in which descriptors will be expressed.

What are descriptors

In a given model, a descriptor is a vector describing a qualitative consideration on data. It can be obtained by :

- computations on the signal
- extraction of informations
- computation on extracted informations

A Global Approach : Texture

Global Approach

In a global approach, we define descriptors directly from the signal. These methods refer to **signal processing**.

Texture refer to the variations of the signal (here a 2D signal) :

- contrast variation
- granularity

Texture Descriptors

By the use of different signal processing methods, we extract a set of numbers (a vector) describing variations of the signal

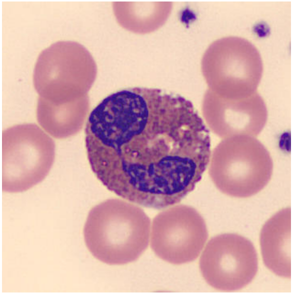

Figure: Texture differences

A Local Approach : Contours

Local Approach

Descriptors based on local approaches are computed from local information already extracted.

Examples

- Contours and shape
- Colours

In the project, we took sides with the use of contours descriptors, in order to classify various-shaped cells.

Fourier Descriptors

After extraction of cell-contours, we want to translate them into sets of numeric values.

Fourier Transform

Let $S(t)$ be a 1D periodic signal (continuous function). We can express S as a sum (over infinity) of trigonometric functions.

$$S(t) = \sum_{k=1}^{\infty} A_k cos(\omega_k t + \phi_k)$$

Fourier Descriptors

A contour can be expressed as a periodic function :

A Contour from an angle

Given a point C_t on the contour, let α_t be the orientation variation of the unitary vector $\vec{v}_C$ adjacent to the contour, between t and $t - \epsilon$.

Contour

We define then a contour as :

$$C(t) = \alpha_t \textit{ with } 0 \leq t \leq 2\pi$$

Fourier Descriptors

Our contour is now a periodic function, and as we saw before, we can express this function with a Fourier Transform.

Harmonics

Let k_M be a number of harmonics we want to use.
$\forall k \leq k_M$, we settle ω_k and ϕ_k as :

- $\omega_k = 2k\pi$
- $\phi_k = 0$

An approximation of the contour is :

$$C(t) = \sum_{k=1}^{k_M} A_k cos(2k\pi.t)$$

Fourier Descriptors

Fourier Descriptors

$\forall k \leq k_M$, we call A_k the k^{th} harmonic.
A Fourier descriptor representing a contour is a k_M-dimensions vector containing all A_k

Classification Examples : Pathology

We can classify ill cells of the sickle-cell disease (SCD) according to their shapes :

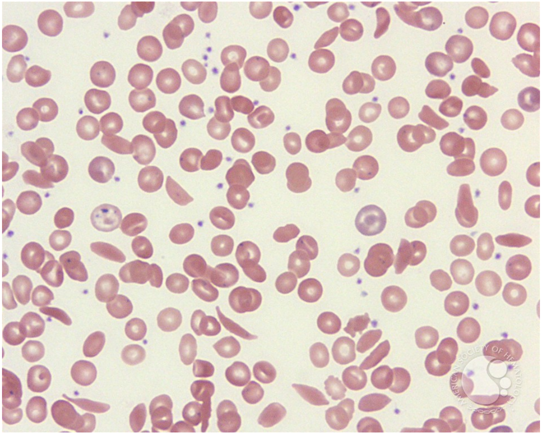

Figure: Sickle cells and normal blood cells

Classification Examples : Bacteriology

Bacteria have various recognizable shapes, often sticks-like shapes :

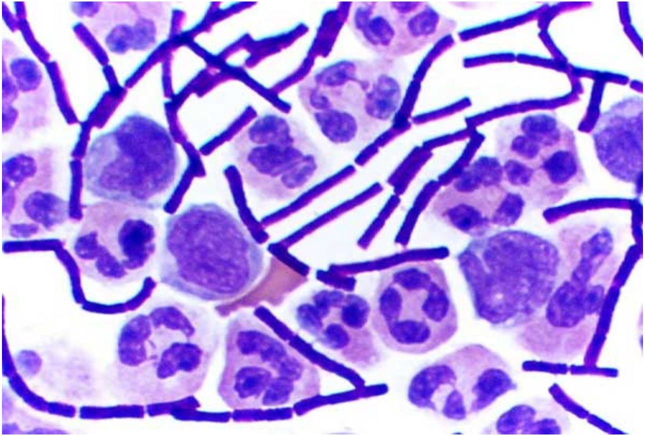

Figure: Bacillus anthracis in cerebrospinal fluid