# Fast and Robust Ellipse Detection

This document summarizes a sub-task from a project I worked on at Oneirix Labs. The objective was to extract the geometrical parameters of blood cells from microscopic images by fitting ellipses to noisy, edge-detected binary images. While the code and results are subject to an NDA and cannot be disclosed, I will outline the problem formulation, core idea, and the modifications I introduced to improve computational performance.

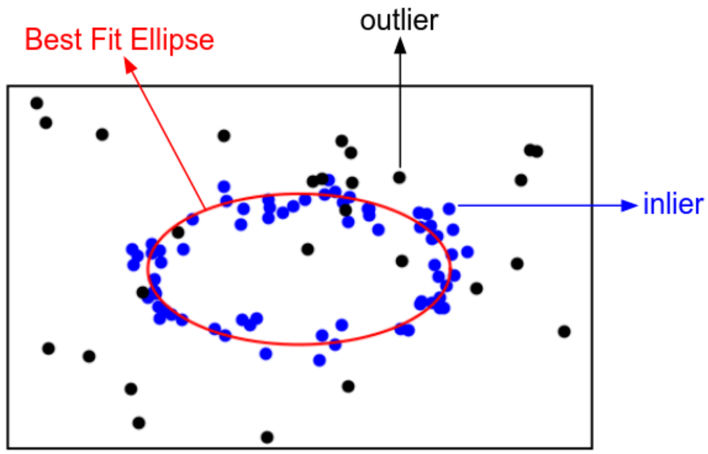


Figure: Simulated example illustrating RANSAC-based ellipse fitting

**I) Problem:**

1. **Input**: Noisy, edge-detected binary microscopic images containing multiple blood cells.
2. **Task**: Rapidly, accurately identify ellipse-shaped cells and fit ellipses to their boundaries.
3. **Output**: Geometric parameters of each fitted ellipse (e.g., major/minor axes, orientation, centroid).

**II) Mathematical Formulation:**

The problem involved detecting ellipses in a set of 2D points $(x, y)$ which may contain significant noise and outliers. An ellipse is represented by the general conic equation:

$$Ax^2 + Bxy + Cy^2 + Dx + Ey + F = 0;\ b^2 - 4ac < 0$$

Given a set of points, the goal was to find the parameters $(A, B, C, D, E, F)$ that best fit these points, while also being robust to noise. For instance, given $N$ points, we need to minimize:

$$\Sigma\,(Ax_i^2 + Bx_iy_i + Cy_i^2 + Dx_i + Ey_i + F)^2$$

This is formulated as a least squares problem, and the non-trivial solution is obtained using Singular Value Decomposition.

**III) Challenge:**

1. **Large number of points**: Increasing $N$ directly impacts computation time.
2. **Noise and outliers**: Standard least squares is sensitive to outliers

**IV) Existing Solutions:**

Using tools from standard libraries like OpenCV:

1. **Hough Transform**
2. **Active Contours**

Both methods operated directly on binary images without requiring a mathematical formulation, but were highly sensitive to outliers and performed poorly in terms of speed and convergence.

**V) Proposed Solution:**

1. **RANSAC (Random Sample Consensus)**: Iteratively select minimal subsets of 5 data points (sufficient to uniquely define a general ellipse), fit an ellipse model to these points, and evaluate the number of inliers to identify the best-fitting model.
2. **Vectorization**: Eliminate ALL explicit loops by operating on entire NumPy arrays

**VI) Programming Stack:**

1. **NumPy**: Numerical computing and efficient array operations.
2. **OpenCV**: Image processing and binary image handling.

**VII) Results**: The detection time for a single ellipse decreased by **~2-5×**. While this improvement may appear modest for a single detection, it translated to a considerable performance gain when processing multiple ellipses across numerous microscopic slides.

**VIII) Reference:** The core idea was adapted from two literature sources, while I developed a custom implementation from the ground up with a focus on vectorization and loop reduction to improve computational efficiency.