Computer Vision Selection Camp 2025

Background: A karyotype is a pictorial representation of an individual's, population's, or species' chromosomes during the metaphase stage of cell division. Abnormalities in chromosome number or structure such as aneuploidies, translocations, deletions, inversions aid in the diagnosis of various genetic disorders and cancers. Traditional methods of chromosome analysis are labour-intensive as they require a technician to manually sort through multiple images taken by pathology cameras to take a higher-resolution image and then assign chromosomal identities to the high-resolution images. In the current paradigm, a commercial system takes images of a cell in metaphase that looks like this:
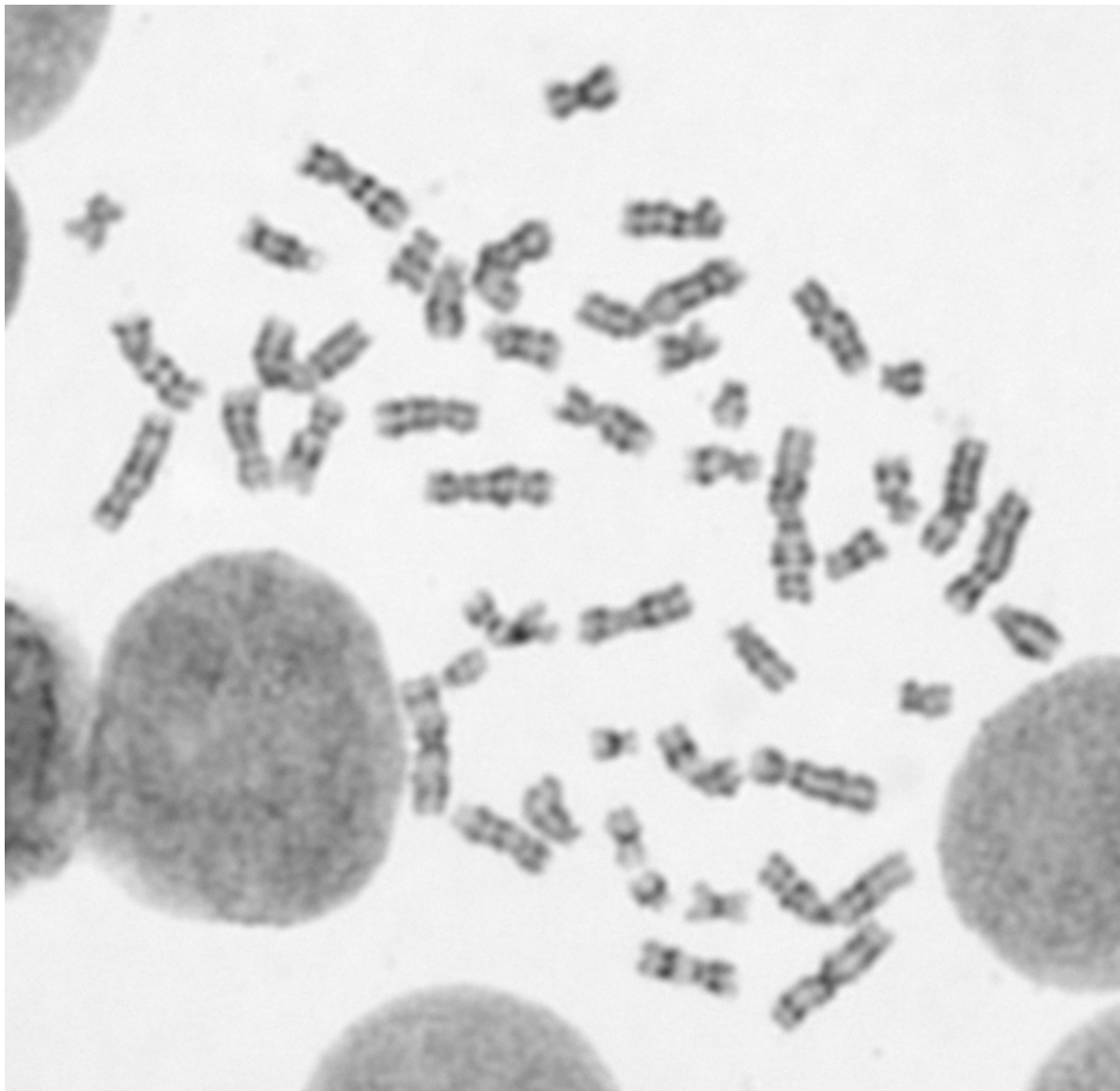
Figure 1. A typical metaphase image

Notice the big blobs - these are nuclei (nucleus) and not informative. We want to separate out (segment) the worm like structures - chromosomes and then assign them an identity. Human chromosomes occur as 23 pairs - pair of chromosomes 1 to 22 and a pair of sex chromosomes XX or XY.
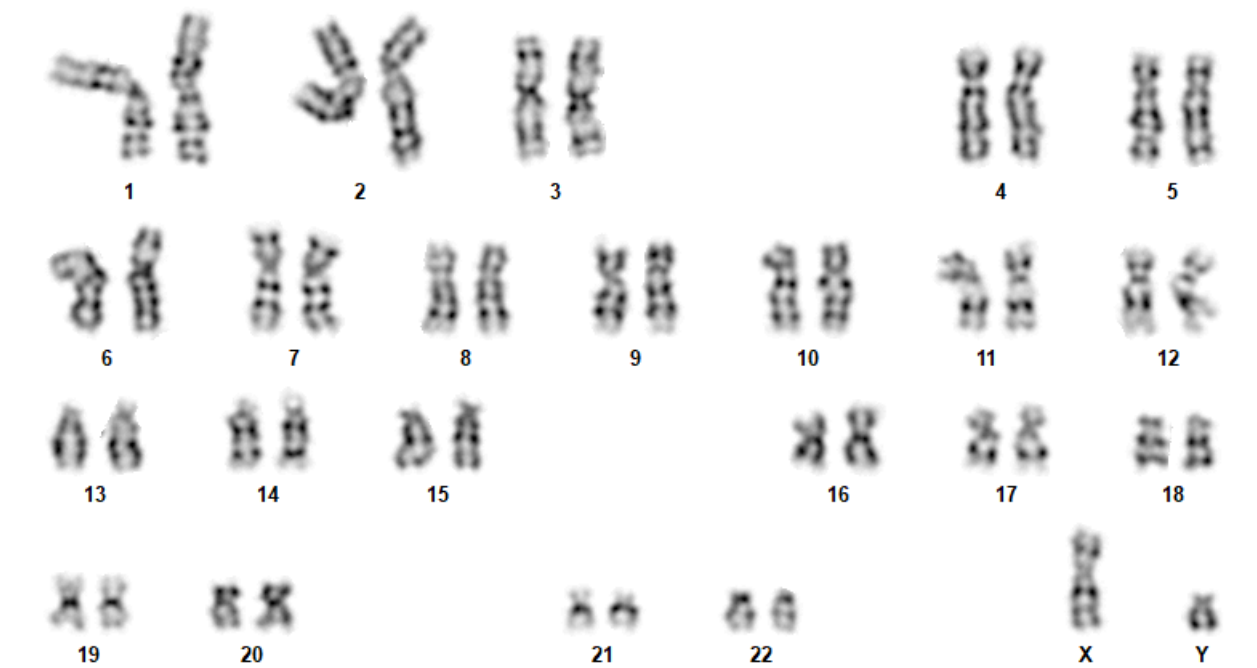

Figure 2. Karyogram - Chromosomes with identities assigned to each pair of chromosomes from the metaphase image in Figure 1

Currently a human needs to assign the identity to each chromosome manually - this is possible because each chromosome has its own banding pattern. Your goal is to train a classifier that can start from Figure 1 and give you identities of individual pairs of chromosomes as in Figure 2.

Dataset and goal: The dataset contains 5000 "24 chromosome class annotations" and 2000 "single chromosome annotations". Each directory has an image and the associated annotation available in xml format.

Single chromosomal annotations are stored in the "single_chromosomes_object" files. This data set only provides users with a distinction between chromosomes and backgrounds.

Chromosomes are not classified in this dataset --> this can be used to train a classifier to separate chromosomes from non-chromosomes.

"24_chromsomes_object" has images along with individual annotation for each chromsome (1-23, X, Y).

Goal: Train a classifer to output the chromosome identity for a given metaphase image. You should plot the auPRC curve. Use the list of images in train.txt to train and in test.txt to evaluate your model.

Data credits: Tseng et al. 2023 https://pmc.ncbi.nlm.nih.gov/articles/PMC9950090/ , **KCDH IIT Bombay**

Data: 📷 2025_Karyogram_CV_Camp

We are not looking for a perfect solution, even if decent efforts are put, its highly appreciated!

Submission Form (Fill with LDAP)- https://forms.gle/vahZsRJeimhnzwGp9
Submission Deadline - 18th October, 11:59 pm