Automated Chromosome Classification for Karyotype Construction Using Deep Learning

Laura G. Speggiorin[1,2], Thayne W. Kowalski[2,3,4], Claudio R. Jung[1], Mariana Recamonde-Mendoza[1,2]

[1]Institute of Informatics – UFRGS, Porto Alegre, RS, Brazil

[2]Bioinformatics Core – HCPA, Porto Alegre, RS, Brazil

[3]PPGBM, Genetics Department – UFRGS, Porto Alegre, RS, Brazil

[4]Medical Genetics Service – HCPA Porto Alegre, RS, Brazil

lgspeggiorin@inf.ufrgs.br,tkowalski@hcpa.edu.br,

crjung@inf.ufrgs.br, mrmendoza@inf.ufrgs.br

Abstract. *This project aims to automate human karyotype generation to optimize cytogenetic analysis, seeking to improve accuracy and reduce manual workload. Using a public dataset, the YOLOv10 model was employed for chromosome detection, demonstrating a mean average performance of 99.4%. For segmentation, the SAM model was used as a proof of concept, proving very successful in most cases for generating masks, although it faces challenges with overlapping chromosomes. In conclusion, the detection achieved excellent performance, and segmentation is promising with room for improvement.*

1. Introduction

A karyotype is the complete set of an individual's chromosomes that describes their number and morphology under the microscope. In diploid organisms such as humans, chromosomes occur in pairs, totaling 46 chromosomes: 22 pairs of autosomes and one pair of sex chromosomes (typically XX for females and XY for males). The term karyotype also refers to a lab-produced image of chromosomes isolated from a cell and arranged in numerical order, used to detect abnormalities in number, type, shape, or banding. This exam helps to determine fetal sex or to diagnose chromosomal disorders, such as Trisomy 21 in Down syndrome.

To produce a karyotype, the sample is prepared, photographed under a light microscope, and the chromosomes are isolated, classified, and organized. Classification relies on visual analysis of size, banding pattern, and centromere position [Erwinsyah et al. 2017]. This technique is very time-consuming and prone to human error, and the current technology used to aid in the process is not efficient. Most automated methods still require manual preprocessing or focus only on classifying pre-cropped chromosomes [Remani Sathyan et al. 2022].

Based on the need to optimize the visual analysis process, this project proposes a pipeline for the classification and arrangement of chromosomes in a karyotype, aiming to improve accuracy, reduce manual workload, and accelerate cytogenetic analysis. More

specifically, the goal of this work is to train a detection model that, based on photomicrography, automatically locates and classifies each chromosome. The segmentation step to isolate each detected chromosome was tested as a proof of concept, but was not explored in depth in the scope of this work.

2. Materials and Methods

2.1. Data

A public Kaggle dataset, published by [Tseng 2022], was used for the experiment[1]. It contains grayscale microscopy images in various resolutions, each with an XML file providing bounding box annotations and class labels for each chromosome. The class distribution is mostly balanced, with most classes having two instances per image. The X chromosome appears in $3/4$ of the cases and the Y chromosome in $1/4$. The reason is that the samples are well divided into male and female (with a few samples with more or fewer sex chromosomes).

The dataset was pre-divided into Train (75%) and Test (25%), with the Train set further split into Train (85%) and Validation (15%). Although 5000 samples were listed, six Train samples were missing. In order to train the model, all images and annotations were organized into a YOLO-friendly directory without modifying the original files.

Table 1. Number of samples per partition after dataset division.

Dataset	% of total	num of samples
Train	$63,74\%$	3183
Validation	$11,23\%$	561
Test	$25,03\%$	1250

2.2. Architecture

For the detection task, YOLO (You Only Look Once) version 10 [Wang et al. 2024] was chosen. The pre-trained large model was imported for fine-tuning on the collected data, with 364 layers and 25,888,688 parameters. YOLO is a one-stage detector, and in version 10, it is more efficient for real-time applications. Moreover, its large version shows better results than the former versions. Thus, it was chosen for consistently achieving state-of-the-art results.

For fine-tuning, the training was executed on an A100 GPU. Since the largest image dimensions were 1024 by 1360 pixels, and the model requires a fixed image size, the chosen value for resizing the images was 1024 in both height and width, the highest resolution manageable by the RAM, with a batch size of 16. The model was trained for 100 epochs.

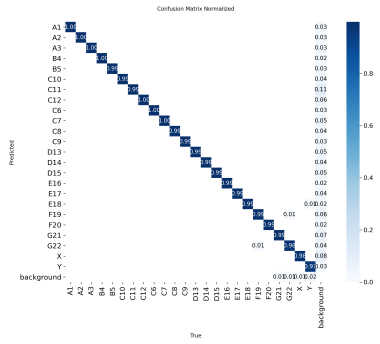
To validate the idea of using a two-step process to form a karyotype, SAM (Segment Anything) [Kirillov et al. 2023] was employed for the segmentation task. SAM has two main advantages relevant for this project: First, it is said to have zero-shot generalization to unfamiliar objects and images, without the need for additional training, and since

[1] `https://www.kaggle.com/datasets/aliabedimadiseh/chromosome-image-dataset-karyotype`

there was no segmentation mask to use as ground truth, there was no easy way to train a model. The second advantage is the fact that SAM can take prompts in the form of text, points, or, interestingly, bounding boxes, which is exactly the detector's output. Since we cannot train a segmentation model and chromosomes are not typical objects in which the system was trained, we can use the predicted boxes as prompts in the segmentation task. For this work, the vision transformer used for the image encoder was trained on high-resolution images.

3. Experimental Results

The detection model performed surprisingly well. As can be seen in the confusion matrix in Figure 1a, all classes present very high performance, the lowest belonging to the sex chromosomes, probably due to having fewer samples to train with. However, the lowest accuracy is 97% for the Y class. To further illustrate this, Figure 1b shows some performance metrics for each class, and for the whole Test dataset. In particular, the overall performance stays high even in higher confidence thresholds.



(a) Normalized confusion matrix.

Class	Images	Instances	Box(P	R	mAP50	mAP50-95)
all	1250	57484	0.991	0.99	0.994	0.828
A1	1250	2499	0.996	0.998	0.995	0.905
A2	1250	2500	0.995	0.995	0.995	0.902
A3	1250	2500	0.995	0.995	0.995	0.89
B4	1250	2501	0.996	0.994	0.995	0.889
B5	1250	2500	0.994	0.993	0.995	0.884
C10	1250	2499	0.992	0.992	0.994	0.854
C11	1250	2499	0.988	0.991	0.994	0.847
C12	1250	2499	0.994	0.995	0.995	0.856
C6	1250	2500	0.993	0.995	0.995	0.878
C7	1250	2501	0.992	0.995	0.995	0.866
C8	1250	2500	0.994	0.989	0.994	0.853
C9	1250	2499	0.994	0.993	0.995	0.854
D13	1250	2500	0.992	0.99	0.995	0.818
D14	1250	2500	0.993	0.991	0.995	0.82
D15	1250	2499	0.99	0.991	0.995	0.817
E16	1250	2497	0.994	0.992	0.994	0.805
E17	1250	2497	0.991	0.991	0.995	0.809
E18	1250	2503	0.996	0.99	0.995	0.783
F19	1250	2498	0.988	0.986	0.993	0.767
F20	1250	2500	0.992	0.987	0.994	0.76
G21	1250	2501	0.991	0.991	0.995	0.731
G22	1250	2495	0.985	0.978	0.993	0.729
X	1248	1835	0.985	0.977	0.991	0.843
Y	651	662	0.976	0.965	0.983	0.723

(b) Performance metrics in the Test set.

Figure 1. Detection evaluation in the Test dataset as a normalized confusion matrix (a) and a table of metrics (b). in Figure 1b, in order, the columns show for each class its name, number of images with it, total number of instances, the bounding box precision, recall, mean Average Precision (mAP) at a threshold of 50%, and mean mAP at varying thresholds, from 50% to 95%.

For the segmentation experiment, the best approach was to run the full image with each bounding box as a prompt, creating an individual mask for each run, for each chromosome. Since there was no ground truth mask available for a quantitative evaluation, a qualitative visual analysis was employed, including specialists in the field. Some examples can be seen in Figure 2. This approach is very successful in most cases, yielding a result similar to Figure 2b, except in eventual cases where chromosomes cross, as can be seen in Figure 2c.

4. Discussion

Although the current work presented excellent results, it still has its limitations. It was trained and tested on a single dataset of images that might not be representative of all real use cases. Also, the approach was not tested in different models, so no claim can

(a) (b) (c)

Figure 2. Original image (a), segmentation mask of a successful (b) and a non-successful (c) example.

be made regarding this being the best approach. The issue of overlapping chromosomes was not deemed as important by the specialists, as it represented cases where part of a chromosome is hidden so another karyotype from a different cell would be needed for that chromosome anyways.

5. Conclusion

This work aims to automate part of the process of generating a human karyotype. The detection model presented great performance, both in locating and in classifying the chromosomes, indicating the visual differences are notable enough to be learned. The segmentation model showed some promise with the prompt, although there is space for improvement, maybe through fine-tuning or a different model/approach. Another issue to be approached is the touching and overlapping chromosomes, which certainly make the segmentation issue more difficult.

Acknowledgments

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001, and by grants from the Fundação de Amparo à Pesquisa do Estado do Rio Grande do Sul - FAPERGS [22/2551-0000390-7 (Project CIARS)] and Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) [308075/2021-8; 440242/2024-0].

References

Erwinsyah, R., Riandi, and Nurjhani, M. (2017). Relevance of human chromosome analysis activities against mutation concept in genetics course. *IOP Conf. Ser. Mater. Sci. Eng.*, 180:012285.

Kirillov, A., Mintun, E., Ravi, N., Mao, H., Rolland, C., Gustafson, L., Xiao, T., Whitehead, S., Berg, A. C., Lo, W.-Y., Dollár, P., and Girshick, R. (2023). Segment anything.

Remani Sathyan, R., Chandrasekhara Menon, G., S, H., Thampi, R., and Duraisamy, J. H. (2022). Traditional and deep-based techniques for end-to-end automated karyotyping: A review. *Expert Systems*, 39(3):e12799.

Tseng, C.-H. L.-E. K.-J. (2022). CIL:54816, homo sapiens linnaeus, 1758, epithelial cell.

Wang, A., Chen, H., Liu, L., Chen, K., Lin, Z., Han, J., and Ding, G. (2024). YOLOv10: Real-time end-to-end object detection.