1. Provide your team background and organization description (if applicable).

I am a data scientist with a strong foundation in software engineering and extensive industry experience. I am passionate about solving challenging problems on platforms like Trustii and have achieved the rank of Competition Grandmaster on Kaggle - becoming the 4th French data scientist to reach this milestone.

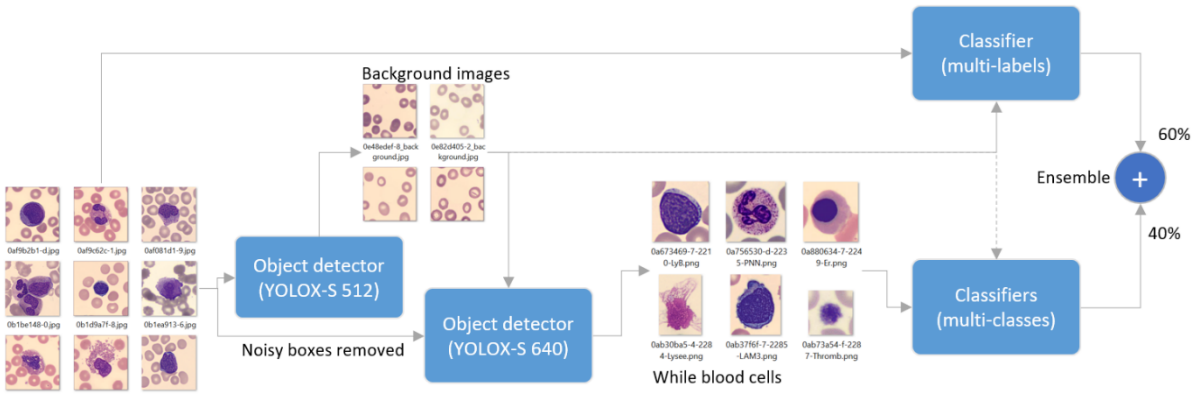
2. Explain why you participated in the Cytologia challenge.

I first discovered the Trustii platform in 2023 while participating in the DigiLut challenge. Later, I came across the Cytologia challenge, supported by France 2030, and decided to tackle it during my spare time over the Christmas break. With no prior expertise in hematology, I approached this competition as I would a typical computer vision problem.

3. Describe how you built your winning model and elaborate on the technical and modeling choices you made.

My approach leverages a two-stage modeling pipeline to effectively detect and classify white blood cells. Stage 1 is a universal White Blood Cell Detection. This stage employs an object detection model to identify white blood cells, irrespective of their underlying class. The goal is to robustly locate and distinguish white blood cells from other elements in the image. Stage 2 is classification once white blood cells are detected in the first stage, a second-stage model classifies them into specific categories (PNN, LY, MO, EO, BA …). Alternatively, this stage can directly learn class labels when working with full images in the absence of explicit white cell detections. This modular pipeline ensures accurate localization and precise classification, optimizing performance for varying data scenarios and improving the adaptability of the solution.

I have created a PDF report that details the technical pipeline and modeling choices: https://github.com/MPWARE-TEAM/Cytologia/blob/main/documentation/overview.pdf



4. What GPU/CPU/RAM resources you used to build your model

Desktop 64GB RAM, IntelCorei9, 10 CPUs, SSD Disk.
1x NVidia RTX3090/24GB-VRAM GPU (local)
1x NVidia A6000/48GB-VRAM GPU (hosted)

5. Do you have any positive feedback or improvement opportunities for the Trustii.io platform?

Trustii.io hosts interesting challenges. Cytologia duration was a bit short (6 weeks) to investigate more solutions / pipelines / ideas.