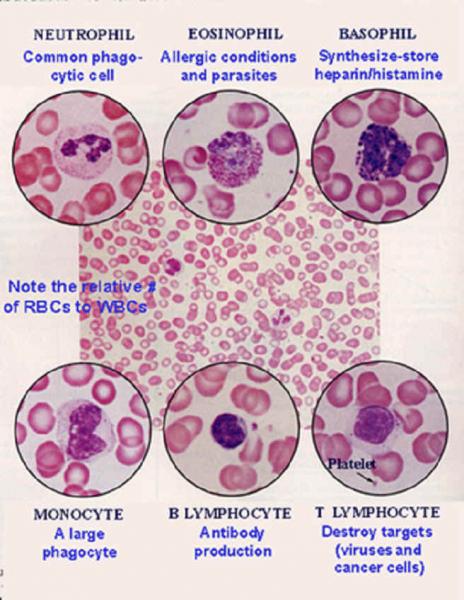
**Blood cells characterization**

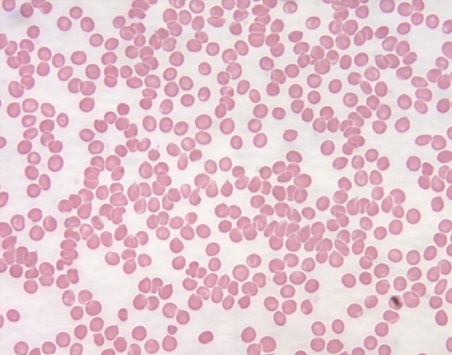
In the biomedical field, accurate and precise blood evaluation plays a crucial rule. Nowadays millions of blood test are analyzed in order to control and improve mankind overall health, and therefore fast and high-throughput process are required. One example is complete blood count (CBC), which quantifies the different cell components present in your blood. CBC is used to evaluate a person’s and detect a wide range of disorders such as leukemia, anemia, malaria, etc. In the past, blood sample analyses were performed manually, leading to a time-consuming and human error-prone process. In the last years, computer-aided systems have been incorporated in the workflow accelerating and facilitating the process. However, totally automated systems are not yet implemented and validated. Therefore, artificial intelligence could play a vital role in improving and optimizing this kind of tests, in part also thanks to the increasing number of datasets (both unlabeled and labelled) provided in the literature. A wide variety of image processing and machine learning techniques are good candidates to tackle the different hurdles present in the workflow.

Focusing on CBC, there are three main components that can be found in all the blood test performed, which are listed below:

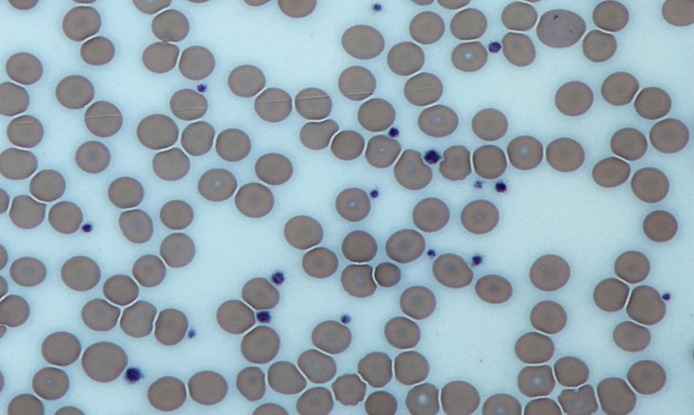
* **White blood cells (WBC):** also called leukocytes, are the cells of the immune system that are involved in protecting the body against both infectious disease and foreign invaders. They flow through your bloodstream to fight viruses, bacteria, and other foreign invaders that threaten your health. Both high and low amounts of this type of cells could indicate important health disorders such as infection, AIDS or leukemia, among others. There are several types of WBC such as monocytes, neutrophils, basophils or eosinophils, each one playing a different role during immune response and triggered by different scenarios.



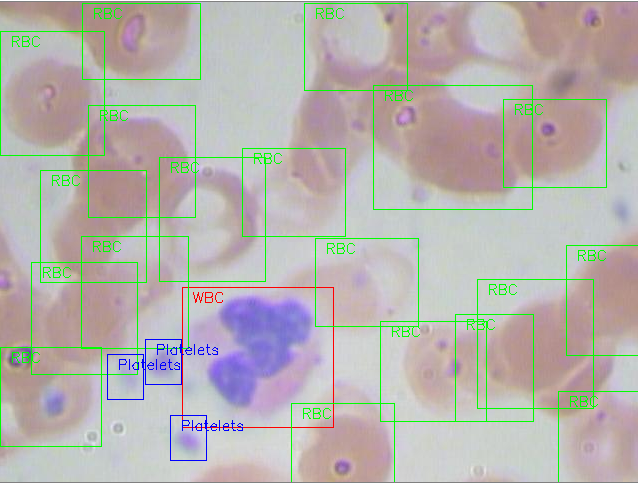
* **Red blood cells (RBC):** are the most common type of blood cell and the vertebrate's principal means of gas exchange in the body tissues. They have a concave shape and no nucleus in order to optimize the oxygen carrying phenomena. Low numbers of RBC could be caused by strong anemias or potential heart disorders.



* **Platelets:** are tiny blood cells that help your body form clots to stop bleeding. If one of your blood vessels gets damaged, it sends out signals to the platelets. The platelets then rush to the site of damage. they form a plug (clot) to fix the damage. Typically, a low platelet count is the result of a medical condition, such as leukemia, or certain medications. In the picture below, platelets are reprsented by the small blue dots.



Bearing in mind the information exposed until now, a labelled data set of 364 images are provided. Each image may contain WBC, RC and platelets in different ratios and positions. Apart from the images, the dataset also includes xml documents containing the coordinates of each cell and its type. The first goal would be creating and algorithm that is able to find with certain precision the different cell components of the image.



In addition, as discussed before, is not only useful to count WBC but also to categorize or classify the different subtypes present in the blood. The dataset also includes labels corresponding to each WBC, but it is not present on the XML documents. Labels could be attached not only to the WBC delimited region but also to the full image (in case only one WBC is present on it). Multiclass classification or binary classification ( polynuclear versus mononuclear) are good cases to be considered Both algorithms could be developed in order to compare the different approaches.

Images and XML files could be uploaded in a certain server, whereas the associated labels could be uploaded to a SQL database in order to retrieve them. Associated labels and locations calculated after the algorithm could also be uploaded to the SQL or maybe create a separate xml (or similar format) with all the predicted information