UCLouvain

University Year **2020-2021**

Using Artificial Intelligence to identify image biomarkers of mature White Blood Cells from peripheral blood smears

Written by **Carvalho Gonçalves José Miguel**
NOMA: **8124-14-00**

Supervised by **John Lee**, a professor at UCLouvain.

Master Thesis
Faculty of Medecine
UCLouvain

Abstract

This thesis introduces a novel workflow whose aim is to identify biomarkers from medical images. The proposed approach is made up of two-steps. In the first step, we automate the cell analysis and in the second step, we try to identify biomarkers among several radiomics (or textural) features and more traditional morphologic (or geometric) features.

Even though such approach can be conceptually applied to a large variety of problems, we chose mature White blood cells (WBC) from peripheral blood smears, mainly because the issue has been largely neglected by the AI community *(see section 2.2)* despite the clinical and educational relevance *(see section 6)*.

Despite the analysis of more than 100 features and almost 55.000 images, the proposed workflow did not identify new image biomarkers of WBC. Nevertheless, it allowed us to verify commonly used cell characteristics values and establish new gold-standard norms for WBC. These results can, however, be criticized because of the data selection bias.

Yet, the major issue of our work is the feature selection process (step 2) and hence, one should improve it before implementing our workflow to more ambitious tasks such as the diagnosis of malaria and hematological diseases – of particular usefulness for undeveloped countries – and for educational purposes by, for instance, helping biomedical students during their histology courses.

Key Terms: Deep learning, Artificial Intelligence, biomarkers, radiomics, medical imaging, peripheral blood smear.

Acknowledgments

I would want to thank Prof. John Lee for his many insights and help with this long project, as well as the *Miro laboratory* who gracefully let me use their GPUs. I would also like to thank several CHL contributors, including Dr. Niel for his guidance, as well as Dr. Vincent and Dr. François, who let me collect data while giving me precious advice.

Abbreviations

WBC	White Blood Cells	**AI**	Artificial Intelligence
SD	Standard Deviation	**ML**	Machine Learning
NI	Normality Interval	**DL**	Deep Learning
CI	Confidence Interval	**ANN**	Artificial Neural Networks
GLCM	Gray Level Co-occurrence Matrix	**PCA**	Principal Component Analysis
GLSZM	Gray Level Size Zone	**PC**	Principal Component
GLRLM	Gray Level Run Length Matrix	**ROI**	Region of Interest
GLDM	Gray Level Dependence Matrix		

Table of contents

1. Introduction

1.1. Motivation

Since the earlier days of my medical curriculum, I had been astonished by the lack of gold-standard norms that could help me, and my colleagues, to discriminate cells from each other. To make things worse, the available criteria such as size and nuclei-to-cell ratio could even be sometimes misleading. At that time, it seemed that such expertise – the ability to easily classify cells and tissues - could only be acquired by seeing many instances - by practicing.

Solving such ambitious task is not easy as one needs not only to correctly identify the most relevant characteristics but also needs a tool that can automatically analyse hundreds of thousands of cells to quantify their respective range of values.

This paper attempts to address those problems with a novel workflow that uses a two-step approach. In the first step we automate cell analysis with deep learning and in the second step, we try to identify **biomarkers*** among several **radiomics*** (or textural) **features** and more traditional morphologic (or geometric) features.

The first step - cell analysis – uses two independent **artificial neural networks***. The first classifies cells (the cell's label), and the second segments cells – that is, it gives us the localization of pixels that belong to the cytoplasm and/or nuclei.

Even though such workflow can be used on a large variety of imaging modalities and pathological samples, I chose to restrict it to mature white blood cells from peripheral blood smears for several reasons: images are easily retrievable from high-through modern microscopes on a large scale, image analysis is reasonably simple and achievable with the current state-of-the-art deep learning algorithms, but mainly because the likely findings are original since cytology has been largely neglected by AI community *(see section 2.2)* despite the clinical and educational relevance. Besides, it allows me to gain further insight into deep learning technology and biomarkers development - sine qua non for more complex and critical clinical problems *(see section 6).*

1.2. Artificial Intelligence

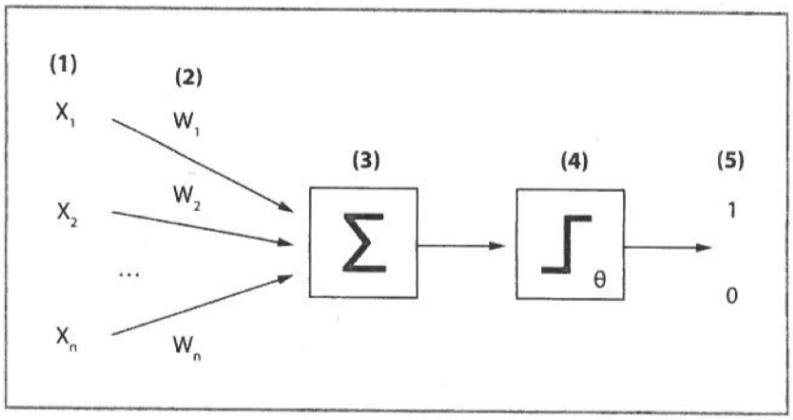
Artificial Intelligence* (AI), and more specifically **Deep Learning*** (DL) a subfield of **Machine learning*** (ML) is ubiquitous - from Google to the small start-up company, the industry quickly adopted it for its ability to **analyse large amounts of data** in several domains such as speech analysis [1], natural language processing [2] and computer vision [3]. However, their massive use is rather new which is mostly due to recent advances in computing power, algorithms efficiency, and the availability of massive datasets.

Although this technology promises to revolutionize medical practice in the coming years [4], it remains alien for the vast majority of medical professionals. Nonetheless, their way of working is rather familiar. Deep learning methods operate like the human brain [5] - information is processed by several layers of transformation[6] into abstract representations which enables him to automatically find out the needed features to classify an object (e.g., wheels, wing mirrors, and windshield for a car). This process, which needs the analysis of a large number of examples to work, is rather different from other more traditional technics that use manual hand-crafted features that are labour-intensive and human dependant. In other words, these methods automatically find a non-linear function between inputs and outputs - as opposed to traditional ones where one must provide it beforehand.

1.3. Standard artificial Neural Networks (ANN)

Deep learning is based on very deep neural networks, which act, as already mentioned, like the ones found in brains. Artificial Neural Networks (**ANN**) are composed of a network of distributed units called **neurons** whose behaviour is based on **Hebbian theory** [7]. The latter stipulates that a synaptic connection is created or reinforced if only two or more neurons are simultaneously activated.

Table 1: Analogy between a biological neuron and an artificial neuron

Biological Neuron	Artificial Neuron
The neuron receives **inputs** from upstream neurons via synapses located on their dendrites. These inputs are then processed in their soma and sent to downstream neurons via the synapses located on their axon. Each synapse has a specific **synaptic efficacy** which is regulated by learning.	 The neuron receives inputs X_n (0 or 1) that are processed in sequential order. First, it computes a weighted sum (whose **weights** are determined after a learning phase) to get an **activation value**, which is then non-linearly converted to value **M** using a **transfer function** (e.g., *sigmoid* and *ReLu* function). Finally, it thresholds the value **M** (e.g. If M > threshold, then give 1. Otherwise, give 0) to get the final output **Y**. We can also add an extra value to the weighted sum called **bias** (not shown here). The weights and biases are called the **parameters** of the ANN.

Training an ANN is to find the parameters (weights and biases) that allow the ANN to give the correct output for every set of inputs $\mathbf{X_n}$. In supervised learning, where we provide labelled data, we train the ANN with inputs $\mathbf{X_n}$ and labels **Z**, and minimize the dissimilarity between the desired outputs **Z** and the predicted outputs **Y**. That dissimilarity is mathematically quantified with a loss function such as *cross-entropy* and minimized with gradient descent algorithm - whose gradients are obtained with backpropagation [8].

1.4. Deep learning architectures

Neural networks design or architecture - that is, how neurons are interconnected and communicate - is an active area of research [9] since it is a major component of ANN performance. Modern ANN are highly complex, but all can be seen as a stack of layers (input, hidden, and output layers) where information propagates through them. If information flows unidirectionally from the input to the output layer, we called them ***feedward networks***. If, however, information flows bidirectionally, we called them ***recurrent networks***.

In this thesis, we only used *feedforward networks* since recurrent networks are mostly used for the unrelated task of sequence processing which includes translation, speech recognition, and DNA analysis.

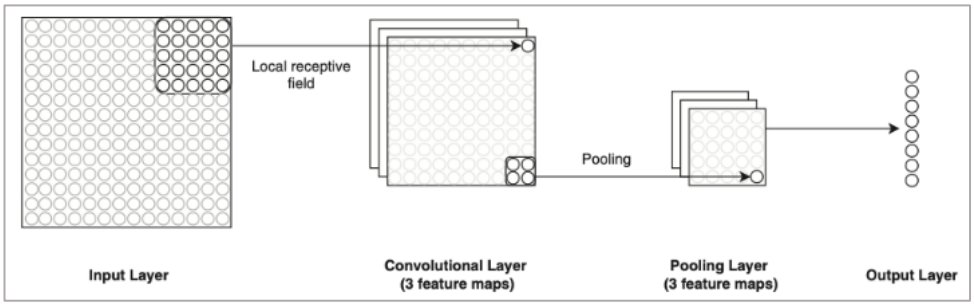
There are many *feedforward networks*, but the most important one in computer vision is probably **convolutional neural networks** (ConvNet) that had been widely adopted [9] by the AI community. Their key characteristic – responsible for their state-of-the-art results - is **convolutional layers**. Those layers act as a filter that detects relevant patterns which are invariant to the location such as edges and motifs. To do so, it sequentially scans the whole input image with **convolutions operations*** to output what we call an Activation map (or feature map). The latter is then (usually) further simplified with **Max pooling layers*** by reducing the input's dimension with **Max pooling operation*** for better generalization capabilities. Finally, information completes its course in the traditional **fully connected layer*** where each neuron is connected to all neurons of the previous layer.

As with simple and complex cells of the visual cortex, this specific organisation allows the convolutional neural network to detect increasingly more abstract and complex patterns.

Table 2: Deep Learning architectures used in this paper

Convolution neural networks (ConvNet) [4]

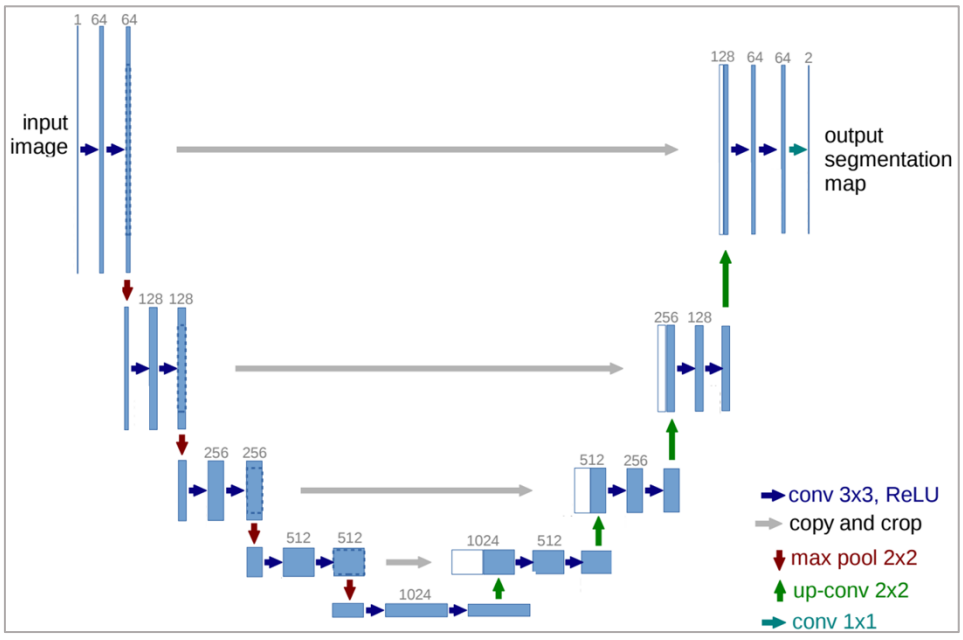
(for image classification)



The input image is sequentially divided into small regions to be analysed, called local **receptive fields**. Each receptive field is connected to a single neuron in the following **convolutional layer** which acts as a filter that detects patterns. This process - the analysis of the whole image by a single filter - results in a **feature map** (or Activation Map) which means that one needs several of them to detect different patterns. The convolution layer is then further condensed by the **pooling layer** for better generalization. Finally, the fully connected layer allows the classification of the input image into one of the output categories.

U-Net [10]

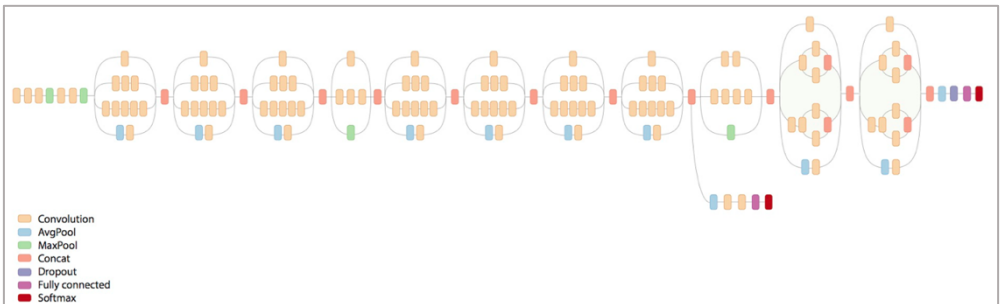
(for semantic segmentation)



U-Net is made up of two interconnected paths - a **contracting path** and an **expansive path**. First, the input image is sequentially downsampled with convolutions, ReLU, and pooling operations which enable the network to compute feature maps (blue boxes) containing relevant patterns. The number of these feature maps progressively increases during this step. Secondly, these feature maps are upsampled to the original size with (up-)convolution operations.

InceptionV3 [11, 12]

(for image classification)



InceptionV3 is made up of a stack of building blocks called **inception modules** that contain a variable number of convolution and pooling operations. This enables each module to compute *multiple different transformations over the same input map in parallel*.

1.5. Deep Learning Workflow

There are three main types of training for deep learning[13]. **Supervised learning*** where the algorithm is trained with labelled data (we assign a label for each object). **Unsupervised learning** where it is trained with unlabelled objects and without any guidance. Finally, **reinforcement learning** where it learns by trial and error. Each method is used for very specific tasks that are, respectively: Regression and classification (like in this thesis), association and clustering, and reward-based problems. In this section, we present the typical workflow for supervised learning:

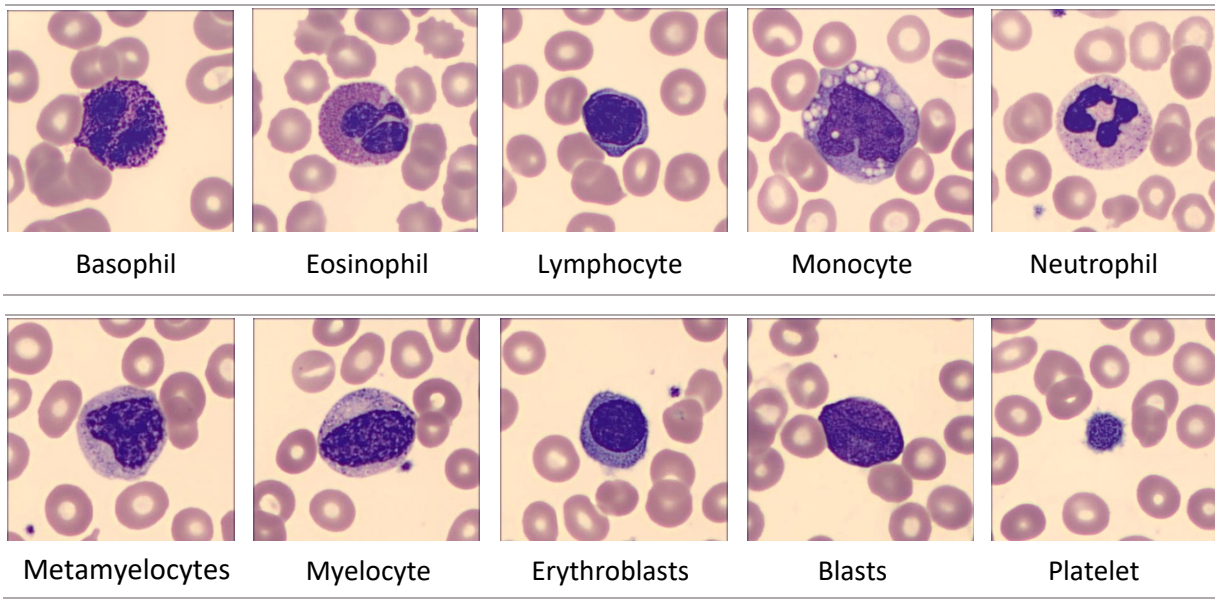
1) Getting the data and split it into three subsets. A **training set** whose purpose is to find the neural network's parameters (weights and biases). A **validation set** whose aim is to find the best **hyperparameters*** by trial and error. Finally, the **test set** to evaluate the neural network's performance.
2) Define the most suited neural network architecture for the task. Nowadays, there're so complex that it is common practice to use an open-source implementation that has been pre-trained on a different dataset. This technic, called **Transfer learning***, allows the learning process to be much faster because filters that had been trained to detected patterns such as edges and motifs in the previous dataset are also meaningful and useful in our dataset.
3) Define the **evaluation metrics*** to assess neural network performance.
4) Train the neural network while comparing the metrics between the training and validation set. This allows us to know whether the model is **underfitting*** (poor training metric and validation metric), or **overfitting*** (good training metric but poor validation metric) - which means that our model has respectively, poor performance, or that it's memorizing the raw examples instead of learning the needed features to correctly classify them.
5) Fine-tune the model architecture and/or hyperparameters to improve the ANN performance. For instance, one could increase the model complexity if it's underfitting, or add regularization, and/or increase the dataset size with **Data Augmentation*** if it's overfitting.
6) Evaluate the neural network in the testing set. If **generalization*** is good enough, deploy it.

1.6. Medical background

Blood tests are one of the most important diagnostic tools for doctors since they can be used to screen and/or confirm conditions such as inflammation, organ dysfunction, and serious diseases like leukaemia. Nowadays, blood analysis is an automatic process done by high-throughput digital microscopes that stain the patient's **peripheral blood smear*** (usually with MGG) and deliver high-resolution images while giving the differential blood count.

The blood is made up of plasma (60%) and three types of cells (40%) - white blood cells, red blood cells and, platelets. Their function is respectively, human defence, oxygen transportation, and coagulation [14]. The White blood cells are further subdivided into five **mature** cells: basophils (0.5-1%), eosinophils (1-4%), lymphocytes (20-40%), monocytes (2-8%), and neutrophils (55-70%). Some pathological samples can also contain **immature** cells such as (meta-pro-)myelocytes, erythroblasts, and blasts.

Figure 1: Cells found in a peripheral blood smear (from sick and non-sick patients)



1.7. Medical Image Analysis

Medical imaging covers a large variety of imaging modalities whose aim is to aid clinicians to diagnose and treat patients. These modalities encompass tomographic images - X-ray, computed tomography (CT), magnetic resonance imaging (MRI), positron emission tomography (PET) - but also histology and cytology images. Unsurprisingly, they produce a still-growing volume of data, which makes medical imaging particularly well-fitted for deep learning.

Image analysis is a broad term that includes several tasks: **Image classification*** where an image is labelled (e.g., benign vs malign). **Image segmentation*** where we label each pixel (e.g., cytoplasm, nuclei, background). **Instance segmentation** where we identify individual objects with their associated pixels – which allows them, in contrast to segmentation, to separate two identical objects like two nuclei that overlap. Finally, there's also the new-growing field of "**Radiomics**" [6].

Radiomics* - defined as the extraction and analysis of quantitative features obtained from the conversion of regions of interest (ROI) of grayscale images to data [15] - is based on the assumption that images are more than just pictures, but data [16]. As such, **textural features** such as the spatial distribution of signal intensities and pixel interrelationships of a given region of interest (ROI), may contain relevant clinical information and disease attributes that the human eye can't see [17]. Consequently, it may be used to identify new **biomarkers*** that could assist clinicians in diagnosis, prognosis, therapy choice, and disease monitoring.

Even though radiomics has been mostly applied in the field of tomographic imaging and oncology [18], it can be used in other imaging modalities and diseases. Like the other "omics" (genomics, metabolomics, and proteomics), radiomics seems to be a necessary step to push precision medicine - whose promise is the right treatment for the right patient at the right time - even further. [19]

It is also worth noting that the terms “radiomics features” and “texture features” are interchangeable.

1.8. Overview

The rest of the thesis is organized as follows: In **Section 2**, we review the related works and present our contributions. In **Section 3** we describe our workflow and the steps needed to identify biomarkers: Image collection, image preparation for the classification and segmentation task, the artificial neural networks building, training, and validation, the post-processing step, the image analysis, and finally, the feature extraction and selection. The results and discussion are given in **Section 4**, which contains the artificial neural networks evaluation, and the image analysis results and discussion. Finally, we conclude and give some perspectives in **Section 6**.

Words with the symbol “ ***** “ are defined in section 3.2 (Glossary).

2. Related works & contributions

2.1. Related works

Deep learning has already been widely used in biomedical images. May this be image segmentation, lesion detection (mainly used on oncology for cancer detection) and classification, but also to make predictions like a disease's prognosis. Some algorithms have even achieved human-level performance on some very specific tasks such as breast lesion detection in mammograms [20], and diabetic retinopathy [21] [22]. **Table 3** provides several high profile-works related to the analysis of biomedical images with deep learning.

Nonetheless, the majority of conducted studies are mostly designed to tackle topics of clinical interest – particularly in radiology [23] and to a lesser extent, digital pathology (also called histopathology) and cytology (as this paper).

The same observation applies to biomarkers identification, whose main goal is to assist clinicians. Unsurprisingly, the usage of radiomics, a not-so-recent method employed on radiology for phenotype quantification and thus, biomarkers identification, has been increasingly applied with deep learning – especially in oncology [24]. However, this association - deep learning and radiomics - has only been used in radiology for clinical issues. Some of these works can be found in **Table 4** whereas **Table 5** provides some papers that use a radiomics approach with more traditional machine learning technics.

This strong emphasis given to clinical issues, but also the scarcity of cytological datasets, explains why cytology has been largely neglected by the deep learning and the radiomics community. Thus, one should not be surprised by the current cytology literature that still implements traditional machine learning techniques and hand-crafted features [25]. Likewise, the current popular automatable cell analysis tools such as Fiji [26], supersegger [27], and OMERO [28] do not feature either, deep learning methods. The only and recent exceptions are CellProfiler [29] and ImageJ [30]. It should also be noted that even though some commercial blood smears systems like HemaCAM, CellaVision® DM96 and DM1200, use artificial neural networks to classify cells, they are still unable to segment them [31].

Table 3: Overview of papers using deep learning technics on biomedical images

Target task	Title	Tissue/ Organ
Radiology		
• Segmentation of retinal vessels in digital fundus images	• Artificial Intelligence in Ophthalmology: A Meta-Analysis of Deep Learning Models for Retinal Vessels Segmentation [32]	• Eye
• Detection of diabetic retinopathy in retinal fundus	• Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs [22]	• Eye
• Classification of masses in mammography	• Representation learning for mammography mass lesion classification with convolutional neural networks [33]	• Breast
• Classification of skin lesions	• Dermatologist-level classification of skin cancer with deep neural networks. [34]	• Skin
• Predict bone mineral density (BMD) and lung percentage of emphysema from CT scans	• Deep learning for biomarker regression: application to osteoporosis and emphysema on chest CT scans. [35]	• Bone and lung
• Predict age using MRI brain images	• Predicting brain age with deep learning from raw imaging data results in a reliable and heritable biomarker. [36]	• Brain
• Predict Alzheimer's disease from FDG-PET	• A DL model to predict a diagnosis of Alzheimer's disease by using 18F-FDG PET of the brain. [37]	• Brain
• Predict BPCO staging and outcome from chest CT	• Disease staging and prognosis in smokers using deep learning in chest computed tomography. [38]	• Lung
Histopathology		
• Predict mutations of prostate cancer	• H&E-stained whole slide image deep learning predicts SPOP mutation state in prostate cancer. [39]	• Prostate
• Predict mutations of lung cancer	• Classification and mutation prediction from non-small cell lung cancer histopathology images using deep learning. [40]	• Lung
• Predict mutations of gastric and colorectal cancer	• Deep learning can predict microsatellite instability directly from histology in gastrointestinal cancer. [41]	• Gastric and colorectal
• Classification of coeliac disease, non-specific duodenitis, and normal tissue	• Automated detection of celiac disease on duodenal biopsy slides: a deep learning approach. [42]	• Duodenum
• Detection of gastric cancer	• Clinically applicable histopathological diagnosis system for gastric cancer detection using deep learning. [43]	• Stomach
• Predict mutations of lung cancer	• Application of automated image analysis reduces the workload of manual screening of sentinel lymph node biopsies in breast cancer Histopathology. [44]	• Lymph nodes
• Detection of breast cancer in sentinel lymph nodes	• Diagnostic assessment of deep learning algorithms for the detection of lymph node metastases in women with breast cancer. [45]	• Breast
• Classification of breast cancer	• Using deep convolutional neural networks to identify and classify tumor-associated stroma in diagnostic breast biopsies. [46]	• Breast
• Mitosis detection in breast cancer	• Mitosis detection in breast cancer histology images with deep neural networks. [47]	• Skin
• Predict TIL scoring in melanoma	• An open-source automated tumor infiltrating lymphocyte algorithm for prognosis in melanoma. [48]	• Skin
• Classification of ovarian cancer	• Automatic classification of ovarian cancer types from cytological images using deep convolutional neural networks. [49]	• Ovary
• Nuclei segmentation	• Deep adversarial training for multi-organ nuclei segmentation in histopathology images. [50]	• Multi-organ
Cytology		
• Classification of Papanicolaou smear images	• Automated classification of Papanicolaou smear images to detect cervical dysplasia. [51]	• Cervix
	• DeepPap: deep convolutional networks for cervical cell classification. [52]	
• Segmentation of Papanicolaou smear images	• Accurate segmentation of cervical cytoplasm and nuclei based on multiscale convolutional network and graph partitioning. [53]	• Cervix
• Detection of thyroid cancer	• Artificial intelligence in cytopathology: a neural network to identify papillary carcinoma on thyroid fine-needle aspiration cytology smears. [54]	• Thyroid
• Detection of urothelial cancer	• Artificial neural network in the diagnosis of urothelial cell carcinoma in urine cytology. [55]	• Urine
• Classification of breast lesions	• Artificial neural network in breast lesions from fine-needle aspiration cytology smear. [56]	• Breast
• Detection of pancreatic cancer	• Computer-assisted cytologic diagnosis in pancreatic FNA: an application of neural networks to image analysis. [57]	• Pancreas
• Detection of mesothelioma	• Detection of malignant mesothelioma using nuclear structure of mesothelial cells in effusion cytology specimens. [58]	• Pleural effusion
• Classification of blood cells	• Improving blood cells classification in peripheral blood smears using enhanced incremental training. [59]	• Peripheral blood smears
	• Analyzing Microscopic Images of Peripheral Blood Smear Using Deep Learning. [60]	
	• Recognition of peripheral blood cell images using convolutional neural networks. [61]	

Table 4: Overview of papers using deep learning segmentation algorithms and radiomics approach on radiology

Target task	Title	Tissue/Organ
• **Extraction of radiomics features** of cervical cancer from MRI	• Deep learning for fully automated tumor segmentation and extraction of magnetic resonance radiomics features in cervical cancer. [62]	• Cervix
• **Extraction of radiomics features** of breast masses from breast CT.	• Deep learning-based segmentation of breast masses in dedicated breast CT imaging: Radiomic feature stability between radiologists and artificial intelligence. [63]	• Breast
• **Classification** of sacral tumors from CT.	• Machine and Deep Learning Based Radiomics Models for Preoperative Prediction of Benign and Malignant Sacral Tumors. [64]	• Sacrum
• **Prediction** of preoperative risk of endometrial cancer from MRI.	• Endometrial carcinoma: MR imaging-based texture model for preoperative risk stratification—a preliminary analysis. [65]	• Endometrium

Table 5: Overview of papers using manual segmentation and radiomics approach on digital pathology

Target task	Title	Tissue/Organ
• **Prediction** of lung or head-and-neck cancer prognosis.	• Decoding tumor phenotype by non-invasive imaging using a quantitative radiomics approach. [66]	• Lung and head-and-neck
• **Classification** of Central Nervous System tumors	• Shedding Light on the 2016 World Health Organization Classification of Tumors of the Central Nervous System in the Era of Radiomics and Radiogenomics. [67]	• Brain and spinal cord
• **Classification** of lung cancer	• Exploratory study to identify radiomics classifiers for lung cancer histology, Front. [68]	• Lung

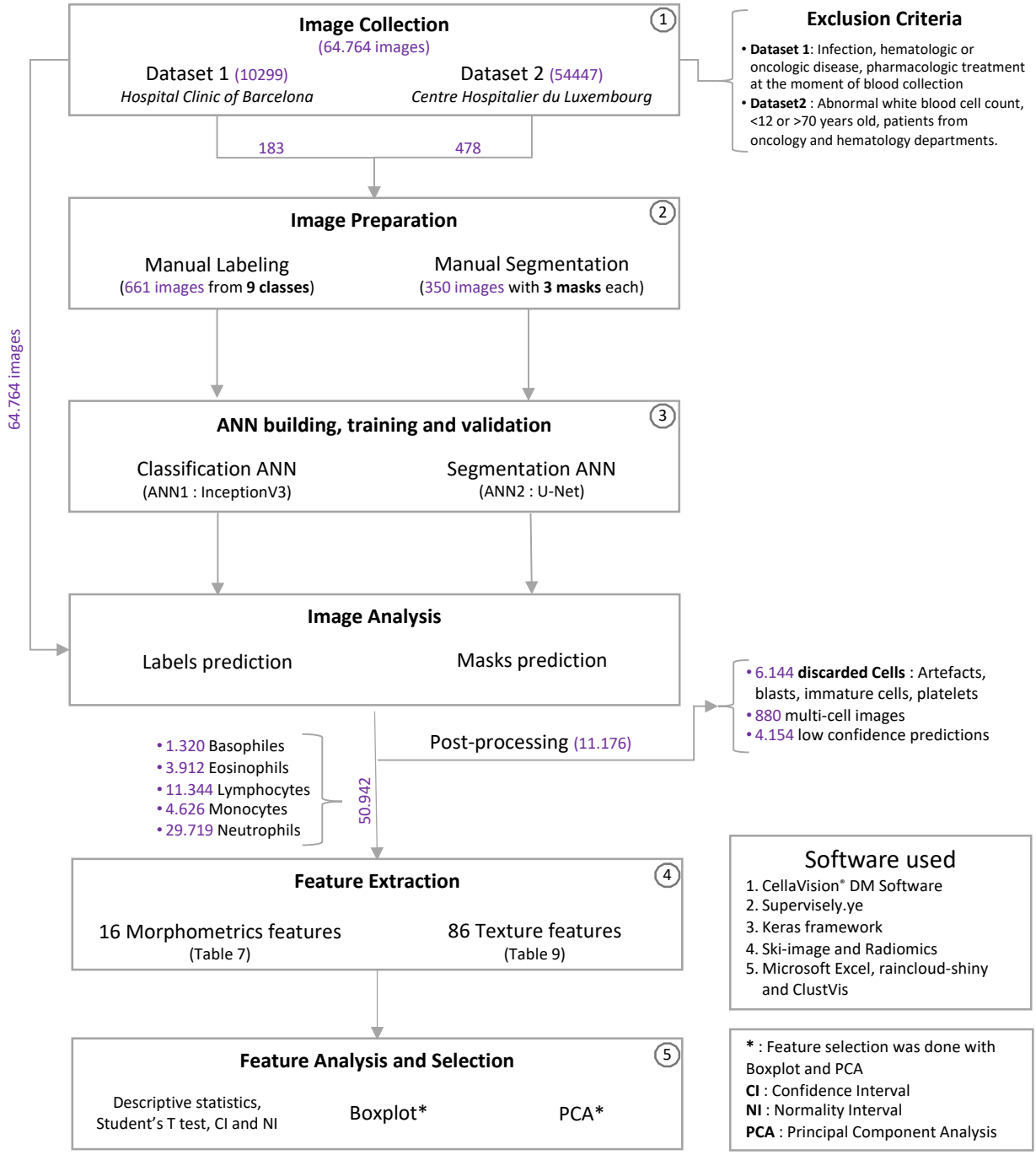
2.2. Contributions

After reading the above-mentioned sections, one realizes that this thesis makes several scientific contributions by being the first work to:

- Use a radiomics approach in peripheral blood smears.
- Use deep learning in such a massive biomedical dataset (more than sixty hundred images) of cells.
- Create and quantify novel morphometric features such as border irregularity and nuclei eccentricity.
- Create the first publicly available image repository of classified AND segmented white blood cells from peripheral blood smears which can be found at https://www.dropbox.com/sh/9s5yfw3p8mymyt8/AAAKSDnTTC_xh6n27xex9-cxa?dl=0. Hopefully, it will contribute to reducing the current gap between cytology and other biomedical image modalities.
- Verify and corroborate, automatically, several commonly taught cell characteristics in biomedical courses like the diameter, nuclei-to-cytoplasm, etc.

3. Materials and methods

3.1. Flowchart of thesis



3.2. Glossary

Computer Science [9, 6, 69]			
Artificial Intelligence	A branch of applied computer science wherein computer algorithms are trained to perform tasks typically associated with human intelligence		
Machine learning	A branch of artificial intelligence where algorithms learn (e.g., how to make predictions) through data without explicit programming.	**Deep Learning**	A branch of machine learning that is capable of learning abstract representations from data. The term Deep Learning and artificial neural networks are often used interchangeably.
Generalization	The capability of the model to perform well on unseen data.	**Deep learning architecture**	How neurons are interconnected and communicate.
Overfitting	The inability of a model to perform well on unseen data despite good performance on the training data.	**Deep learning framework**	Software that is used to implement deep learning models. E.g., Keras, Tensorflow, Torch….
Underfitting	The inability of a model to perform well on the training data and unseen data.	**Artificial neural networks (ANN)**	Computing system inspired by biological neural networks (e.g, neurons, synapses, and synaptic efficacy), and used to analyse and process information.
Supervised learning	Type of learning where the model is trained with labelled data (we assign a label for each object).	**Transfer learning**	A technic where one trains an already pre-trained deep learning model in a new dataset. This method allows the model to converge much faster.
Boolean image	Black and white image that only contains two-pixel values, 0 and 1 - as opposed to classical B&W images whose pixel intensity range is 0 to 255.	**Data Augmentation**	A technic where we apply random transformations to the images used to train the deep learning model. This technic artificially augments our original dataset which allows it to be trained with less data while being more robust.
Classification	A task where an entire image is categorized (or labelled) into a class such as "people", "animals".	**Hyperparameter**	A parameter whose value is set before training. **E.g.,** learning rate, number of layers.
Segmentation	A task where we identify parts of the image belonging to a specific class. Conceptually, it's a pixel classification task.	**Convolutional layer**	A specific layer used to detect and extract specific features (e.g., edges and motifs) from an input image by using learned filters and convolution operation. It enables the model to preserve the spatial relationship between pixels and to be invariant to object translation.
Mask	An image that only contains the desired class (e.g., nuclei).	**Convolution operation**	Mathematical operation where a specific Matrix **Y** (or Filter whose size is called *receptive field*) slides over all regions of the Matrix **X** (or Input) to compute Matrix **Z** (or Output).
Bayesian error	The best optimal theoretical error for any classifier. It's usually similar to human-level performance.	**Max pooling layer**	A specific layer used to remove irrelevant information from the previous layer by shrinking it with convolution and pooling operations.
Evaluation metrics	Mathematical formulas used to assess the model performance. **E.g**: Precision, recall, intersection over union, ….	**Max Pooling operation**	Mathematical operation where we retrieve the maximum value of a fixed-shape window from the input. We repeat that procedure until the entire input is analysed to get the Output.
		Fully connected layer	A layer where neurons are connected to all the neurons of previous layers.

Image Analysis [70]			
Biomarkers	A characteristic that is objectively measured and evaluated as an **indicator** of normal biological processes (e.g., cells, tissues), pathogenic processes, or responses to a therapeutic intervention.	**Peripheral blood smear**	A sample of blood smeared on a microscope slide (after staining) and used to inspect blood cells.
Radiomics	A method that includes the extraction and analysis of quantitative features obtained from the conversion of regions of interest (ROI) of grayscale images to data.	**White blood cells**	Cells of the immune system that mainly serve to protect the human body against infections by fighting germs such as virus, bacteria, and other foreign invaders.

3.3. Image collection

We acquired a total of 64 746 images from two datasets collected from Cellavision® analysers DM9600 and DM1200. The latter is a popular high-throughput microscope used in laboratories and hospitals that outputs standardized uni-cell images that can be easily exported in JPG format. As opposed to the more manual intense time-consuming preparation method of taking pictures from individual cells from a traditional microscope, this method allowed us to create a much larger dataset.

Data set 1: 10 299 samples from a publicly available dataset captured with DM9600 device of Core Laboratory at the Hospital Clinic of Barcelona. The exclusion criteria include individuals with infection, hematologic or oncologic disease and patients free of any pharmacologic treatment at the moment of blood collection. [71]

Data set 2: **54 447 samples** from a total of 195 patients, captured with DM1200 device at the Centre Hospitalier du Luxembourg from 12/11/2020 to 22/01/2021. The exclusion criteria include patients below 12 and above 70 years old, abnormal white blood count, and patients hospitalized on the oncology and haematology department. This dataset, whose images are completely anonymised, has been made public to be used, hopefully, for the development of new algorithms.

3.4. Image preparation

The image preparation was carried out in two stages. First, we classified images into 9 classes that were later verified by experts. Secondly, we segmented each image into 3 classes using the annotation platform Supervisely.ly [72].

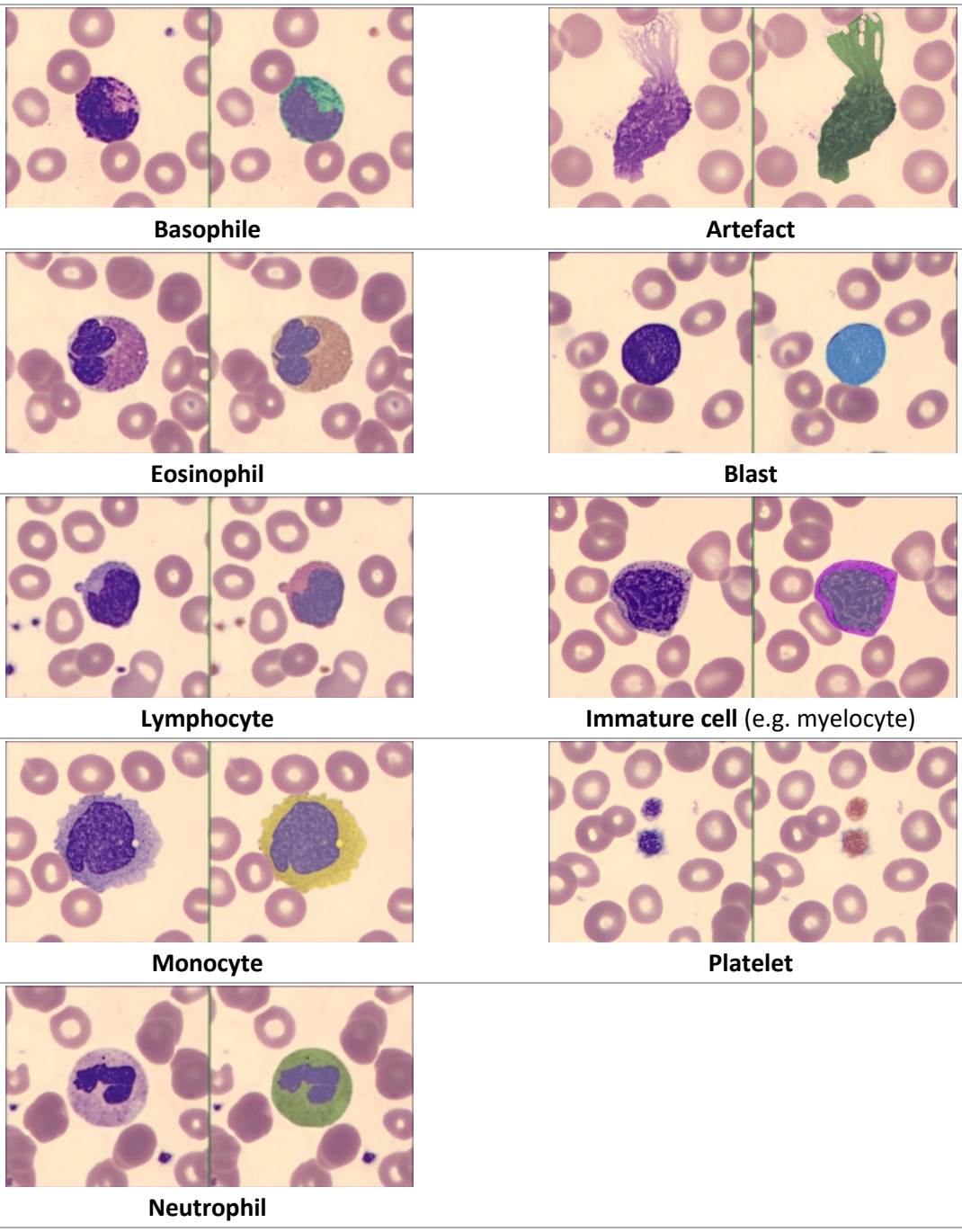
Classification Task (1er stage): A total of 661 – of which 183 from dataset 1 and 478 from dataset 2 - were manually classified into 9 classes (see **Figure 2**). Only 5 classes were used for the feature extraction process. Nonetheless, the remaining 4 classes (or discarded cells) were still classified since it allows our model to perform better in the classes of interest. The number of samples per class is indicated in brackets.

- **White blood cells** (5): Basophils (70), eosinophils (70), lymphocytes (87), monocytes (101), neutrophils (70).
- **Discarded cells** (4): Artefacts (destroyed or busted cells) (47), blasts (83), immature cells (promyelocytes, myelocytes, metamyelocytes, and erythroblasts) (73), and platelets (60).

Segmentation task (2e stage): A total of 398 cells, chosen from the previous 661 classified classes, were manually segmented into two classes - cytoplasm and nuclei. A third class, the background, was computed since it corresponds to the pixels that do not belong to either cytoplasm or the nuclei.

- **Masks** (3): Cytoplasm, nuclei, background.

Figure 2: Samples of segmentation masks for each cell



3.5. Artificial Neural Networks building, training, and validation

We built and trained three publicly available pre-trained neural networks using the high-level framework *Keras* [73]. The **ANN1** to classify cells (e.g., Eosinophils, Lymphocytes, …), the **ANN2** to segment cells and/or their nuclei, and finally, the **ANN3** to classify the few images containing several cells. More technical information can be found in **Table 6**.

We used two independent neural networks for the classification and segmentation task – instead of one single neural network – because this approach proved to be, after empirical experimentation, to be much more efficient. Not only it requires less training data, but it has also the added benefit of being more flexible for future works (e.g., add an extra class).

The **ANN3**, however, is only required for the post-processing step. Since our ANN1 and feature scripts were built for single-cell images, we eliminated multi-cell images. This model was trained with 148 manually classified samples of which 80 single-cell images, and 68 samples multi-cell images. These samples were collected from ANN2 results – in other words, we used the outputted masks from ANN2 as the training data for the ANN3. Thus, we can seamlessly eliminate multi-cell images during the image analysis process; instead of doing it at the very end. Additionally, the training process is easier since the images are **Boolean*** instead of RGB.

Table 6: Artificial Neural Networks (ANN) used in this thesis

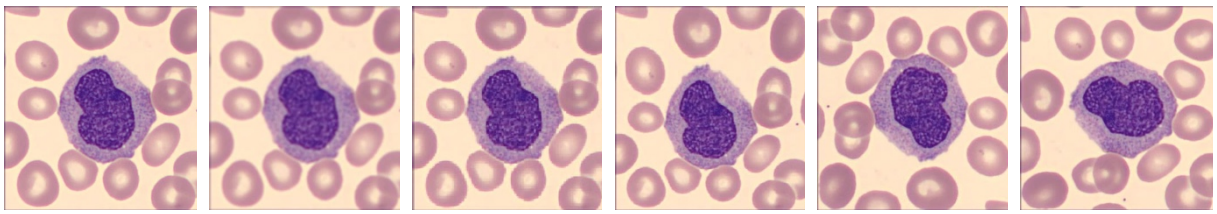
	Task	Architecture	Comments
ANN1	Classification of cells	InceptionV3	This architecture was empirically chosen after testing several publicly available networks. More details can be found in **Table 12** in section 4.1.
ANN2	Segmentation of cells and/or nuclei	U-Net with backbone EfficientNetb4	This architecture was chosen for its state-of-the-art results in medical images and small datasets.
ANN3	Classification of single-cell and multi-cell images	InceptionV3	

Furthermore, we used two popular technics to increase the performance of the neural networks.

Transfer learning* by using pre-trained models (InceptionV3 and U-Net) in *ImageNet* [3] - an image database with currently more than 14 million images classified between 20.000 categories.

Data Augmentation* by using five image transformations - Blur, Downscale, GridDistortion, Flip, Transpose - from the library *Albumentations* [74] and *ImageDataAugmentor* [75]. We choose these specific variations since they mimic the physiological cell variability found among individuals, and/or the alterations induced by the coloration method and/or the blood test method.

Figure 3: Image transformations used for Data Augmentation



Original Image **Blur** **Downscale** **GridDistortion** **Flip** **Transpose**

As for the evaluation metrics implemented to **train and validate the model**, we used standard metrics such as precision, recall, and accuracy. Additionally, we also used the dice coefficient and IoU for the segmentation task.

Table 7: Evaluation metrics - including formulas - [6] used for classification and segmentation

IoU (Intersection over Union; = Jaccard index)	
Measure from 0 to 1 (no overlap, and perfect overlap, respectively), which quantifies how closely two objects match. E.g., the predicted mask, and the real mask.	$\frac{Area\ of\ Intersection}{Area\ of\ union} = \frac{\lvert P \cap GT \rvert}{\lvert P \rvert + \lvert GT \rvert - \lvert P \cap GT \rvert}$
Precision (or positive predictive value)	
Fraction of true positive predictions.	$\frac{TP}{TP + FP}$
Recall (or sensitivity)	
Fraction of positives examples detected by the model.	$\frac{TP}{TP + TN}$
Dice coefficient (= F1 score)	
A variant of IoU that also quantifies how closely two objects match	$\frac{2\ x\ Area\ of\ Intersection}{Total\ number\ of\ pixels\ in\ both\ images} = \frac{2\ x\ \lvert P \cap GT \rvert}{\lvert P \rvert + \lvert GT \rvert}$
Accuracy	
Fraction of predictions that our model got right.	$\frac{Number\ of\ correct\ predictions}{Total\ number\ of\ predictions} = \frac{(TP + TN)}{(TP + TN + FP + FN)}$

True Positive (TP) represents the correctly classified predicted positives, **False Positive** (FP) represents the misclassified predicted positives, **True Negative** (TN) represents the correctly classified predicted negatives, and **False Negative** (FN) represents the misclassified predicted negatives.
"P" represents the **Predictions** given by the model, and "**GT"** the **Ground Truth**

The remaining neural networks training details like the selected hyperparameters and data-augmentation parameters can be found in the Appendices.

3.6. Image analysis

The collected data was naturally analysed with ANN1 and ANN2, whose aim is respectively, to label each image into one of the nine categories, and to segment each image into three masks.

3.7. Post-processing

As previously mentioned, we eliminated three categories of images. **Low confidence** predictions that we defined as ANN1 predictions with probability confidence < 95%. Thus, we avoid the few misclassifications of ANN1 and improve the precision of our results. **Discarded cells** that include facts, blasts, immature granulocytes, and platelets – since they're not the subject of this thesis. Finally**, multi-cell images** using ANN3 because, as already stated in section 3.5, the ANN1 and feature scripts were built for single-cell images.

Moreover, we also removed very small objects and holes from the final masks, as they're most likely small artefacts from ANN2.

3.8. Feature extraction

We computed a total of 102 features of which 16 **morphometric features** and 86 **texture features**. These were extracted from either the nuclei and/or cell (cytoplasm AND nuclei), with *Ski-image* and *PyRadiomics* libraries [76]. It should be noted that these libraries computed the features from original images. The predicted masks from ANN2 were only used to retrieve the coordinates of the nuclei and cytoplasm. Furthermore, we rectified their results according to the scale of 1px = 0,101 um. The latter was confirmed by the Cellavision®company.

Morphometric features (also called morphological features) [77]**:** We computed 9 different morphometric features, for a total of 16 features since 7 features were computed twice (cell and nuclei).

Table 8: Morphometric features

Cell and/or **nuclei features**		**Cell features**	
Perimeter	Number of pixels around the boundary of the region of interest (ROI)	**Nuclei-to-cell ratio**	Nuclei area divided by the cytoplasm area.
Area	Number of pixels in the region of interest	**Nuclei eccentricity**	Euclidean distance between central centroids of "cell u" and "nuclei v", divided by the cell equivalent diameter. The value 0 means a perfectly centered nucleus. *The central centroid is defined as the Geometric center of the ROI, given by the arithmetic mean position of all the points.*
Equivalent diameter	Diameter of a circle with the same area as the ROI		
Major axis length	Length in pixels of the largest axis of the ROI-enclosing ellipsoid		
Minor axis length	Length in pixels of the smallest axis of the ROI-enclosing ellipsoid.		
Perimeter-to-area ratio	Perimeter divided by the area.		
Border irregularity	Perimeter divided by the perimeter of a perfect circle with the same area (defined as $2\,x\,\pi/\sqrt{Area/\pi}$). The value 1 means a perfect round boundary.		

Texture Features: We computed a total of 84 features for each cell. These features are grouped into 5 methods - First-order statistics, Gray Level Co-occurrence Matrix (**GLCM**), Gray Level Size Zone (**GLSZM**), Gray Level Run Length Matrix (**GLRLM**), and Gray Level Dependence Matrix (**GLDM**). For simplicity purposes, we only provide in this section a summary of the key concepts in **Table 9**, as well as the list of these features in **Table 10**. More detailed information, including formulas, can be found in the Appendices and the *Imaging Biomarker Standardization Initiative (IBSI)* [78].

Table 9: Key concepts behind texture characterization

Entropy	A measure of the average amount of information required to encode the image values. It's also a measurement of pixel randomness.	**Variance**	A measure of the spread of the distribution of values in the ROI.
Mean	Average gray-level intensity within the ROI.	**Uniformity**	A measure of the homogeneity of the image area.
Energy	A measure of the magnitude of the pixels within the ROI. It's a measurement of pixel randomness.	**Autocorrelation**	A measure of the magnitude of the fineness and coarseness of texture.
Skewness	A measure of the asymmetry of the distribution of values in the ROI	**Emphasis**	A Measure of the distribution of a specific zone or run (for a given size and/or gray-level intensity). A greater value means that the latter is present in great proportion.
Kurtosis	A Measure of the 'peakedness' of the distribution of values in the ROI.		

Table 10: Texture features ordered by method (with their definition) [79]

GLDM (12)	GLRLM (16)	GLSZM (16)	GLCM (22)	First Order statistics (18)
They quantify gray level dependencies in an image, which are defined as the number of connected pixels within distance δ that are dependent on the center pixel	They quantify gray level runs, which are defined as the length in number of pixels, of consecutive pixels that have the same gray level value.	They quantify gray level zones in an image, which are defined as the number of connected voxels that share the same gray level intensity.	They quantify how often pairs of pixels with specific values and in a specified spatial relationship occur in an image. Put differently, it measures how often different combinations of pixel brightness values (grey levels) occur in an image.	They describe the distribution of voxel intensities within the ROI.
• Small Dependence Emphasis (SDE) • Large Dependence Emphasis (LDE) • Gray Level Non-Uniformity (GLN) • Dependence Non-Uniformity (DN) • Dependence Non-Uniformity Normalized (DNN) • Gray Level Variance (GLV) • Dependence Variance (DV) • Dependence Entropy (DE) • Low Gray Level Emphasis (LGLE) • High Gray Level Emphasis (HGLE) • Large Dependence Low Gray Level Emphasis (LDLGLE) • Large Dependence High Gray Level Emphasis (LDHGLE)	• Short Run Emphasis (SRE) • Long Run Emphasis (LRE) • Gray Level Non-Uniformity (GLN) • Gray Level Non-Uniformity Normalized (GLNN) • Run Length Non-Uniformity (RLN) • Run Length Non-Uniformity Normalized (RLNN) • Run Percentage (RP) • Gray Level Variance (GLV) • Run Variance (RV) • Run Entropy (RE) • Low Gray Level Run Emphasis (LGLRE) • High Gray Level Run Emphasis (HGLRE) • Short Run Low Gray Level Emphasis (SRLGLE) • Short Run High Gray Level Emphasis (SRHGLE) • Long Run Low Gray Level Emphasis (LRLGLE) • Long Run High Gray Level Emphasis (LRHGLE)	• Small Area Emphasis (SAE) • Large Area Emphasis (LAE) • Gray Level Non-Uniformity (GLN) • Gray Level Non-Uniformity Normalized (GLNN) • Size-Zone Non-Uniformity (SZN) • Size-Zone Non-Uniformity Normalized (SZNN) • Zone Percentage (ZP) • Gray Level Variance (GLV) • Zone Variance (ZV) • Zone Entropy (ZE) • Low Gray Level Zone Emphasis (LGLZE) • High Gray Level Zone Emphasis (HGLZE) • Small Area Low Gray Level Emphasis (SALGLE) • Small Area High Gray Level Emphasis (SAHGLE) • Large Area Low Gray Level Emphasis (LALGLE) • Large Area High Gray Level Emphasis	• Autocorrelation • Joint Average • Cluster Prominence • Cluster Shade • Cluster Tendency • Contrast • Correlation • Difference Average • Difference Entropy • Difference Variance • Joint Energy (or Angular Second Moment) • Joint Entropy • Informational Measure of Correlation (IMC) 1 • Informational Measure of Correlation (IMC) 2 • Inverse Difference Moment (IDM) • Inverse Difference Moment Normalized (IDMN) • Inverse Difference (ID) • Inverse Difference Normalized (IDN) • Inverse Variance • Maximum Probability (or Joint Maximum) • Sum Entropy • Sum of Squares (or Joint Variance)	• Energy • TotalEnergy • Entropy • Minimum • 10th percentile • 90th percentile • Maximum • Mean • Median • Interquartile range • Range • Root Mean Squared (RMS) • Mean Absolute Deviation (MAD) • Robust Mean Absolute Deviation (rMAD) • Skewness • Kurtosis • Variance • Uniformity

3.9. Feature analysis and selection

Data was summarized with Microsoft Excel and assessed by analysis of the mean and standard deviation (SD). The feature significance was established with the Student's t-test between each pair of cells, for a total of 10 tests. Results are considered statistically different if the p-value is <0.05. In addition, we computed the confidence interval (CI) as well as what we call "Normality interval" (NI), which we defined as the interval between **mean - 2SD** and **mean + 2SD**.

The most relevant features for cell classification – that is, the ones with minimum redundancy and maximum relevance to the cell of interest – were **selected** with two methods that use the same key assumption: The most relevant features for classification are the ones with the greatest inter-cell and smallest intra-cell variability.

First, we graphically summarized – with the web tool *raincloud-shiny* [80] [81] - morphometric features with **boxplot's** to visually select the ones respecting the above conditions.

Secondly, we applied **Principal Component Analysis** (PCA) - with the web tool *Clustvis* [82] - to morphometric and texture features to retrieve the ones with the greatest variability. As a reminder [83], PCA is an unsupervised learning method similar to clustering that reduces data dimensionality by transforming original dimensions (e.g., Area, diameter, perimeter, …) into fewer new but more abstract dimensions called Principal Components (PC). Each PC explains a precise amount of data variance or variability (e.g., If PC1 is 87%, it means that PC1 explains 87% of data variability) and is the result of a linear and "weighted" combination of original features. Therefore, features with the greatest ponderation (or weight) in the PC are also the ones with the greatest variability. In other words, PCA allows the identification of the key features behind the variance of data.

4. Results and Discussion

4.1. Artificial neural networks performance

Quantitative evaluation: Our three neural networks achieved a state-of-art accuracy of 99.33%, 100%, and 98.96% for cell classification, cell segmentation, and multi-cell images, respectively. More details can be found in **Table 11** which summarizes the performance – using several evaluation metrics – of the three neural networks after 120 epochs.

As already stated in section 3.5, we chose the architecture *InceptionV3* for cell classification since it proved to be the best one for our task. The details can be found in **Table 12**.

Table 11: Performance (using evaluation metrics) of the three artificial neural networks

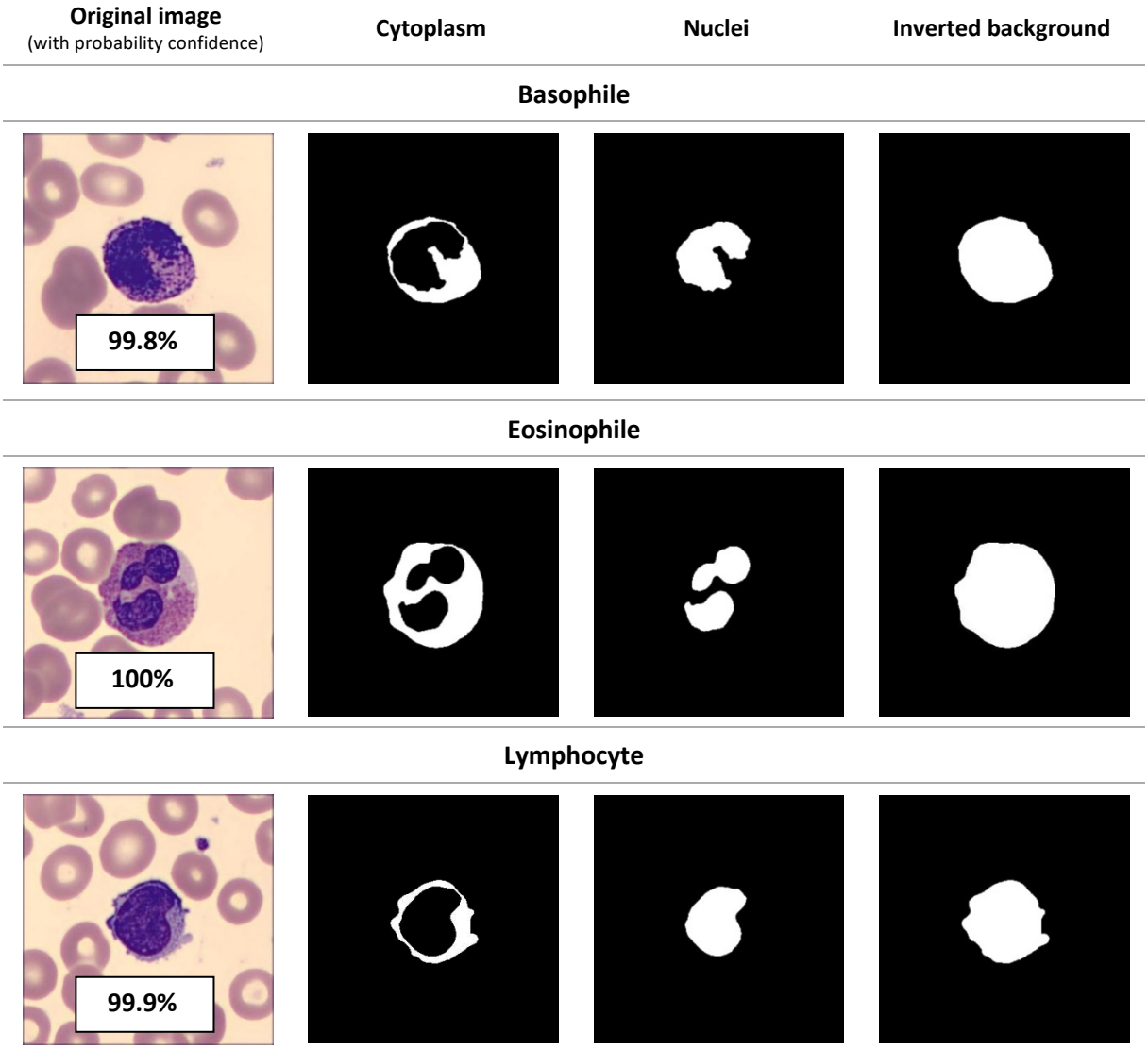
Neural Network	Precision	Recall	Accuracy	mean-IoU	Dice (F1-Score)
ANN1 (Classification of cells)	99.33	99.33	**99.33**		
ANN2 (Segmentation of cells and/or nuclei)	98.96	98.96	98.96	98.31	**99.32**
ANN3 (Classification of single-cell and multi-cell images)	100.0	100.0	**100.0**		

Table 12: Performance (using evaluation metrics) of the several tested cell classification ANNs

ANN for cell classification	Precision	Recall	Categorical Accuracy
VGG19	84.64	83.50	83.84
InceptionV3	99.33	99.33	**99.33**
EfficientNetB4	94.26	93.94	94.28
EfficientNetB7	94.22	93.27	93.94
Xception	92.48	89.87	91.15

Qualitative evaluation: We also manually checked some random results from ANN1, ANN2, and ANN3, to rule out any systematic error during the quantitative evaluation process. As one can notice in **Figure 4**, not only do those samples validate our quantitative evaluation, but they also reveal the admirable precision of our models. This finding is even corroborated by **Figure 5** which shows some random *low confidence predictions* whose abnormal phenotype makes it difficult to classify even for the human eye.

Figure 4: Random results from ANN1 (Original Image with probability), and ANN2 (Masks: Cytoplasm, Nuclei, and Inverted background)



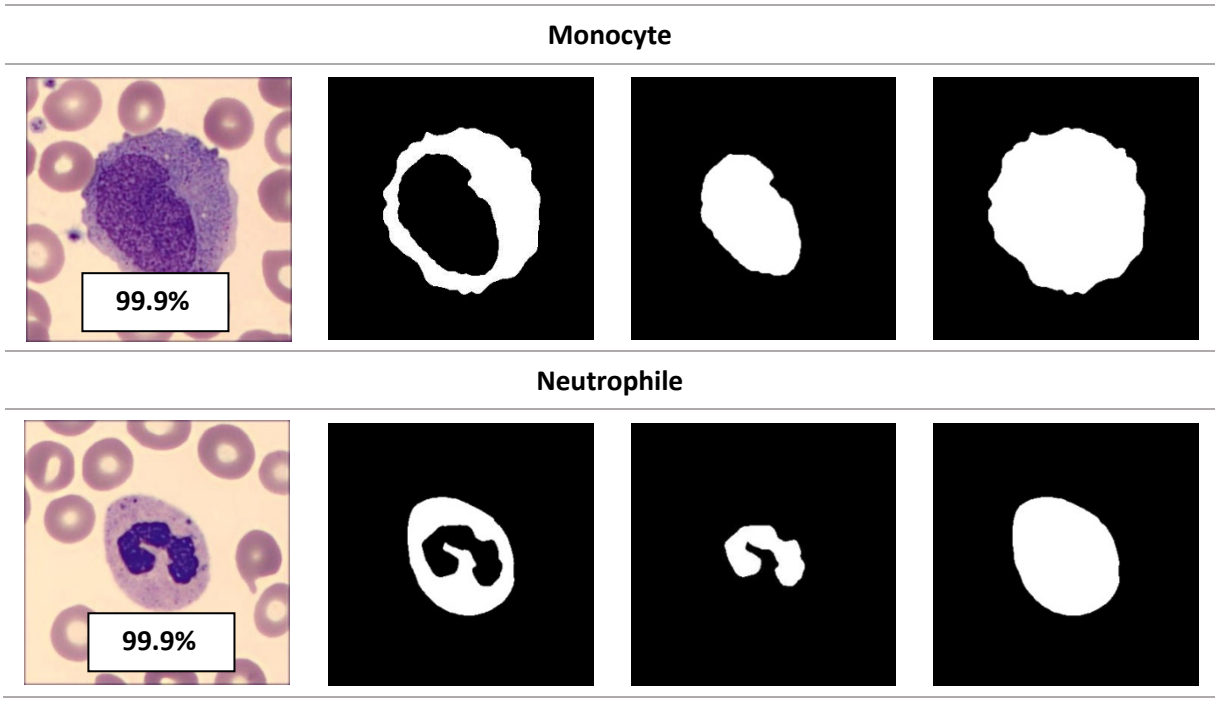
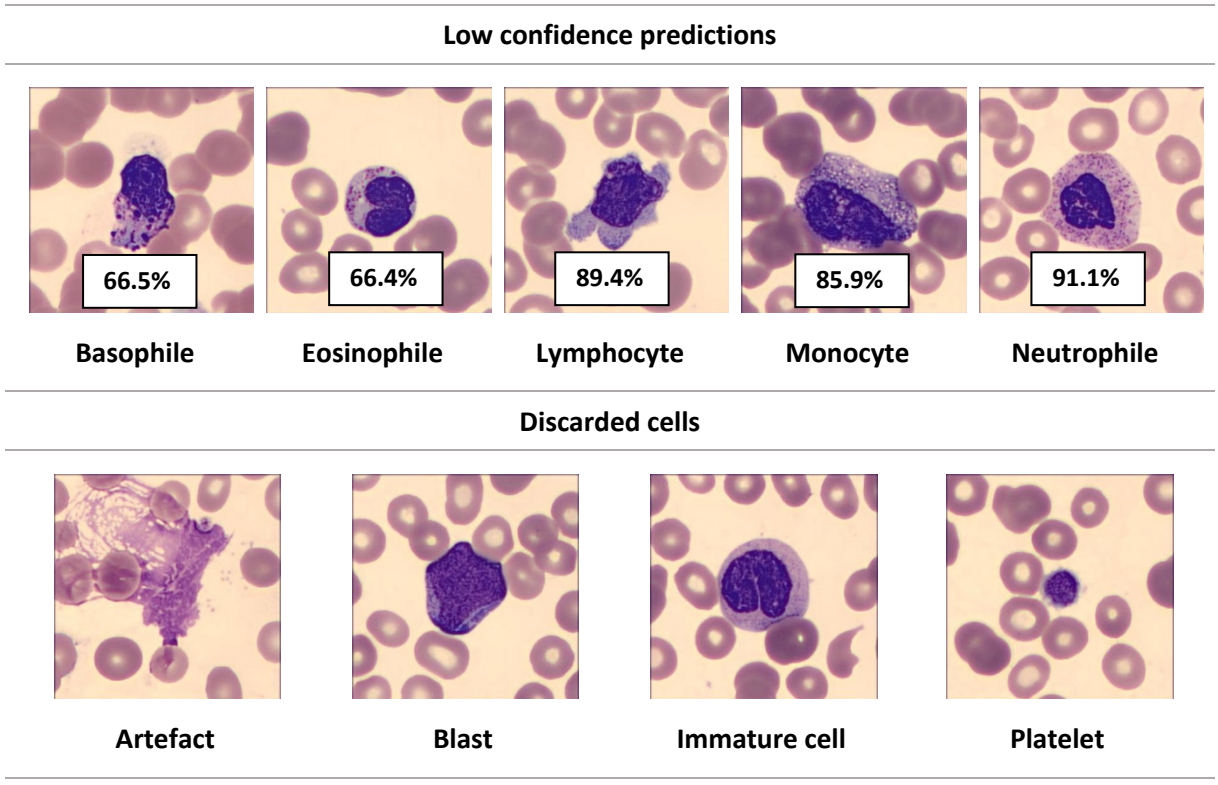
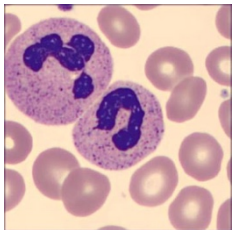
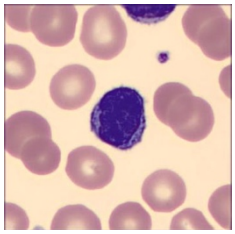
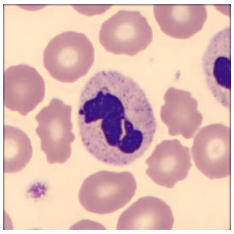


Figure 5: Random samples of the post-processing stage - Low confidence predictions, discarded cells (artefacts, blasts, immature granulocytes, and platelets), and multi-cell images.



Multi-cell images		
		
Sample 1	**Sample 2**	**Sample 3**

4.2. Feature analysis and selection

From the total of 62.118 collected images, we analysed 50.942 cells, of which 1.320 basophils, 3.912 Eosinophils, 11.344 Lymphocytes, 4.626 Monocytes, and 29.719 Neutrophiles. As a reminder, we computed 102 features for each sample, of which 16 morphological and 86 texture features. For the sake of clarity, the p-values which are almost always statistically significant ($p < 0.05$), confidence values, and the normality interval of the texture features, are presented in the Appendices.

We eliminated 11.176 samples during the post-processing step, of which 6.144 discarded cells, 4.152 low confidence predictions and 880 multi-cell images.

The collected results - including images, as well as raw and processed data – are also available at https://www.dropbox.com/sh/9s5yfw3p8mymyt8/AAAKSDnTTC_xh6n27xex9-cxa?dl=0.

4.2.1. Morphometric features

As expected, the variability between morphometric (or morphological) features of cells is consistent with what is currently taught in the biomedical field. Monocytes and their nuclei are the largest cells since they have the largest **area, diameter, perimeter,** and **minor axis**, whereas Lymphocytes are the smallest while having the biggest **nuclei-to-cell ratio**. The remaining cells - basophils, eosinophils and neutrophils - have a fairly similar size.

Furthermore, the **border irregularity of cells** and **cells perimeter-to-area ratio** are also similar among classes, as opposed to their nuclei: Neutrophils have the most irregular nuclei while having the biggest perimeter-to-area, whereas Lymphocytes have the most spherical and central nuclei since it has the smallest **nuclei eccentricity** value.

More interestingly, the inner variability of features within each class – also called the intra-cell variability - that can be depicted as the interquartile range and/or the *Normality Interval*, seems to be mostly high for the **cell's area** and **perimeter**, as well as for the **nuclei border irregularity**, **nuclei perimeter-to-area ratio**, **nuclear-to-cytoplasm ratio,** and the **nuclei eccentricity**. In contrast, the **nuclei area**, **cell irregularity**, and **cell perimeter-to-area ratio** are rather similar within each class.

Unfortunately, our results corroborate what histological experts state during histology class – that one cannot classify a white blood cell with one single morphometric feature. This statement is validated by the following two findings.

First, when inspecting the **boxplots**, one can realize that there is no feature with significant inter-cell variability while having low intra-cell variability. For example, when analysing the *Normality Interval*, we can easily see that the lower value of the major class (e.g., Neutrophil) is always smaller than the larger value of the minor class (e.g., Lymphocytes) – there is always an overlap. Unsurprisingly to biomedical readers, the feature that seems to be least affected by this finding seems to be the **nuclei-to-cytoplasm ratio**. In contrast, the **nuclei characteristics** (e.g., nuclei area, diameter, perimeter, eccentricity, etc) seem to be the less useful ones for white blood cell classification.

It should be noted that one could use the interquartile range to overcome the overlap problem, but the latter only contains fifty percent of values. Thus, we cannot use them to identify biomarkers and/or for cell classification since their accuracy is too coarse. It can however be used to gain some precious insight about feature variability.

Secondly, when reviewing PCA results, one can notice that data variance cannot be significantly explained by a single feature - only by several of them. For instance, PC1 which explains 51% of data variability, is equally reliant on 10 features from the total of 16 features – each of which has approximately an absolute value of 0,30. The remaining 6 features are basically the dominant features of PC2, PC3, PC4, PC5, PC6 that account for 19%, 9%, 7%, 6%, and 4% of data variance, respectively. In other words, all features are more or less equally important. There's no specific feature that stands out.

Additionally, when examining the PCA score plot (**Figure 8**), one realizes that the variability between WBC – which can be roughly estimated by the area and diameter of the prediction ellipses – is surprisingly homogenous. Besides, there's also a noticeable overlap between them (but less pronounced between Monocytes and Lymphocytes, which is probably due to their size difference) which means that there is no clustering. Thus, despite the complexity of our PCs, the latter are unsuitable for WBC classification.

In short, and despite the successful data dimensionality reduction of PCA - from 16 variables to roughly 6 variables that explain 96% of data variance – it was **unable to identify new image morphological biomarkers** of mature white blood cells.

For the sake of simplicity, we only described in this section (without going into too much detail) the most relevant findings. The details, including values of the mean and their standard deviation, the *Normality Interval,* the boxplots of the data, PCA Loadings and PCA score plot, are shown in **table 13**, **table 14**, **Figure 6** and **7**, **Table 15**, and **Figure 8**, respectively.

Hopefully, this data can and will be used as new gold-standard norms for future textbooks and histology courses.

Table 13: Heatmap of the mean and Standard deviation (SD) of the morphometric features

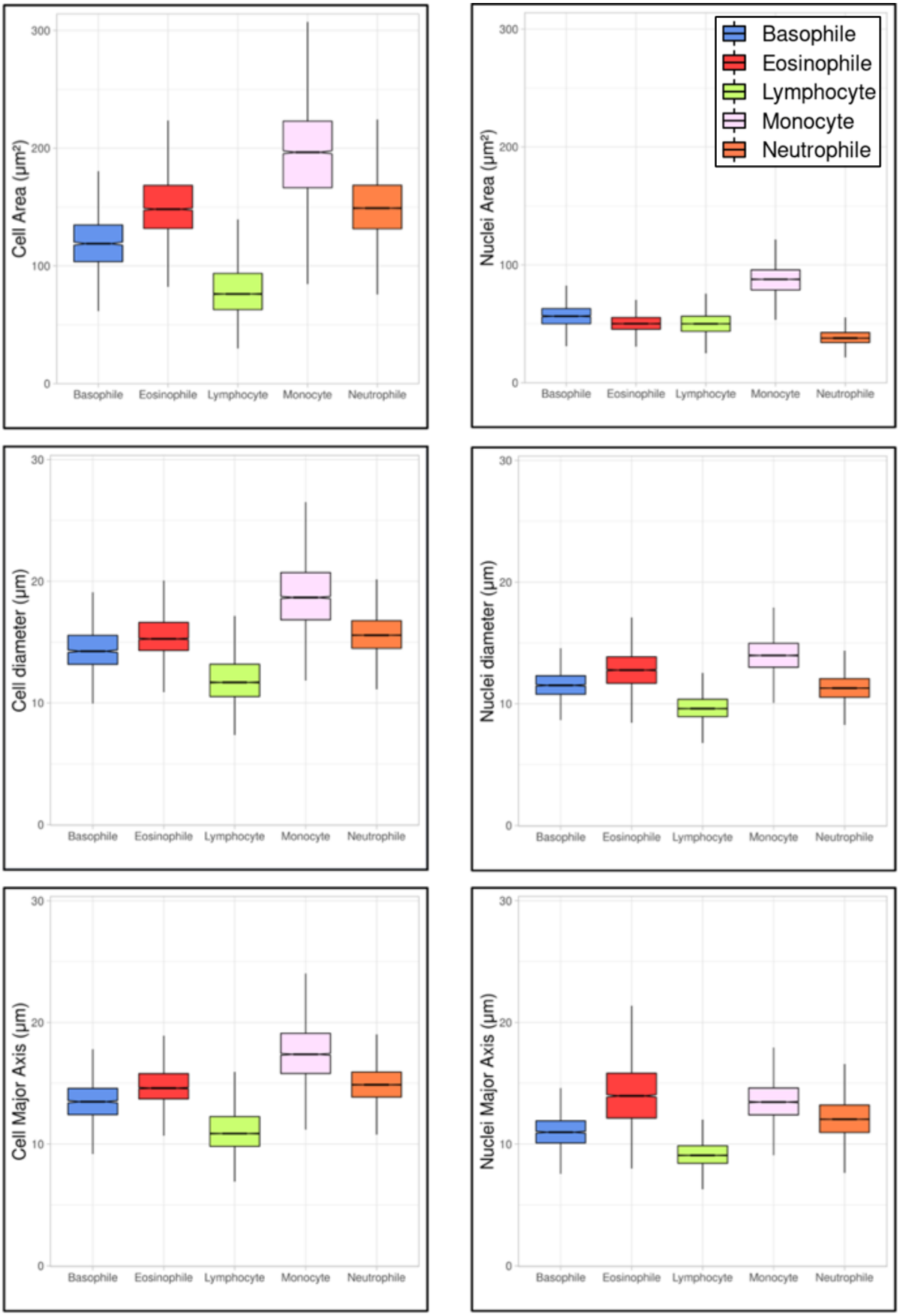
		MEAN					STANDARD DEVIATION				
		BASOPHILE	EOSINOPHILE	LYMPHOCYTE	MONOCYTE	NEUTROPHILE	BASOPHILE	EOSINOPHILE	LYMPHOCYTE	MONOCYTE	NEUTROPHILE
MORPHOLOGY	Nuclei area	56,61	50,66	49,97	87,30	38,96	10,85	7,92	9,72	13,14	6,92
	Nuclei perimeter	178,81	188,29	124,94	195,26	192,55	36,28	31,67	16,80	32,62	32,29
	Nuclei diameter	11,57	12,83	9,71	14,02	11,36	1,24	1,66	1,17	1,57	1,19
	Nuclei major axis	11,06	14,07	9,21	13,55	12,14	1,42	2,63	1,21	1,72	1,71
	Nuclei minor axis	8,12	7,31	7,46	9,59	7,78	1,07	1,28	0,83	1,14	1,41
	Nuclei perimeter-to-area ratio	3,37	3,92	2,72	2,35	5,19	0,92	0,84	1,25	0,51	1,10
	Cell area	119,82	151,01	80,10	196,55	150,78	23,08	27,83	23,51	42,23	27,88
	Cell perimeter	192,98	212,18	158,25	253,41	211,08	31,17	31,81	30,67	40,36	26,44
	Cell diameter	14,53	15,59	12,09	18,95	15,74	2,27	1,93	2,26	3,03	1,85
	Cell major axis	13,61	14,84	11,22	17,70	14,99	1,76	1,69	2,00	2,77	1,69
	Cell minor axis	11,89	13,59	9,65	15,03	13,31	1,26	1,25	1,45	1,68	1,36
	Cell perimeter-to-area ratio	1,70	1,48	2,12	1,36	1,46	0,26	0,18	0,35	0,17	0,17
	Nuclei-to-cell ratio	0,48	0,34	0,65	0,45	0,26	0,08	0,04	0,11	0,07	0,04
	Cell border irregularity	4,99	4,88	5,03	5,12	4,87	0,60	0,48	0,53	0,47	0,36
	Nuclei border irregularity	6,76	7,51	5,03	5,92	8,83	1,35	1,21	0,58	0,94	1,53
	Nuclei eccentricity	0,066	0,093	0,059	0,073	0,061	0,05	0,05	0,05	0,05	0,04

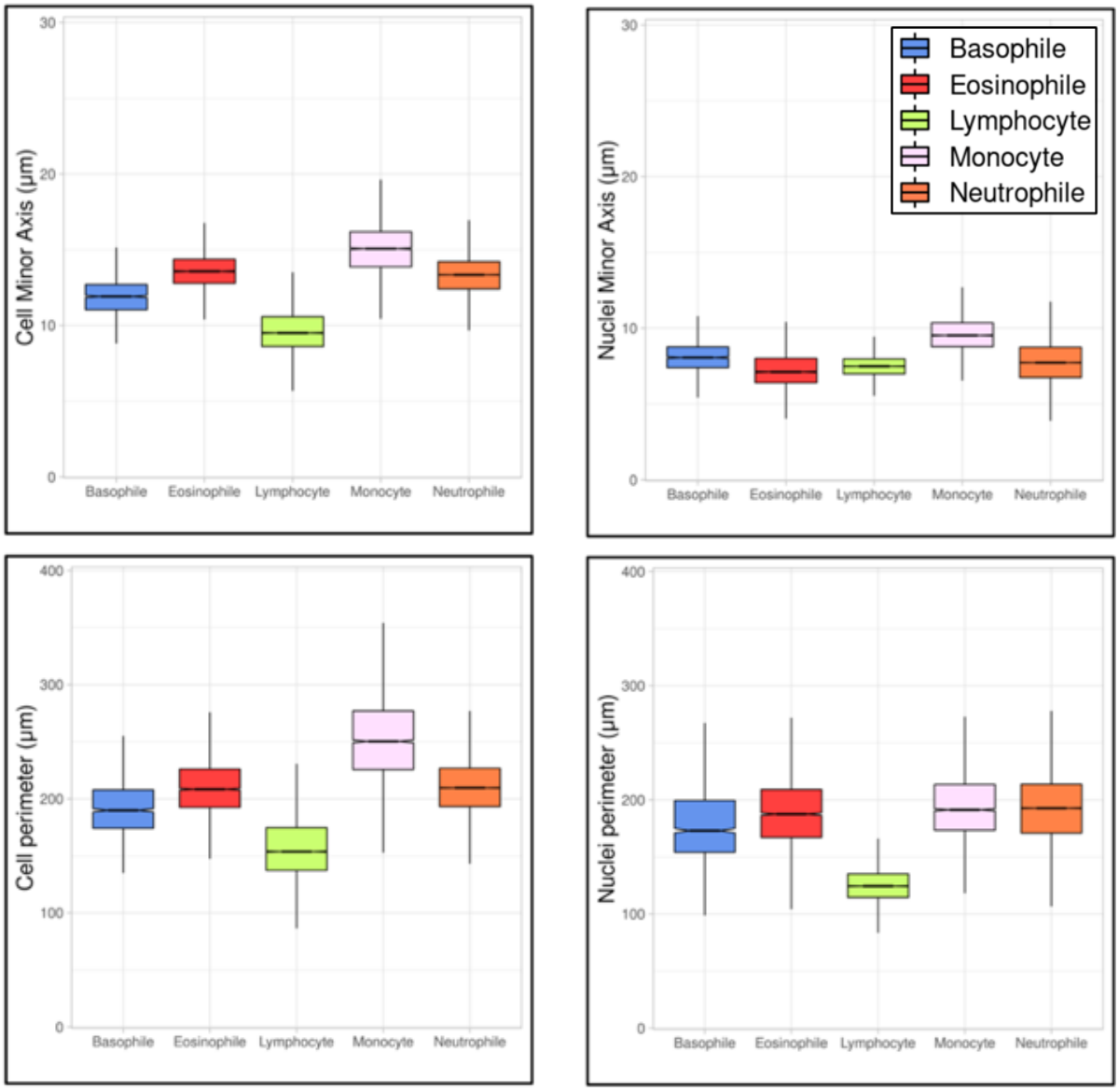
Description: *An heatmap was applied for each row for better visualization. The green cell stands for the largest value and the red cell for the smallest value.*

Table 14: Normality Interval of the morphometric features

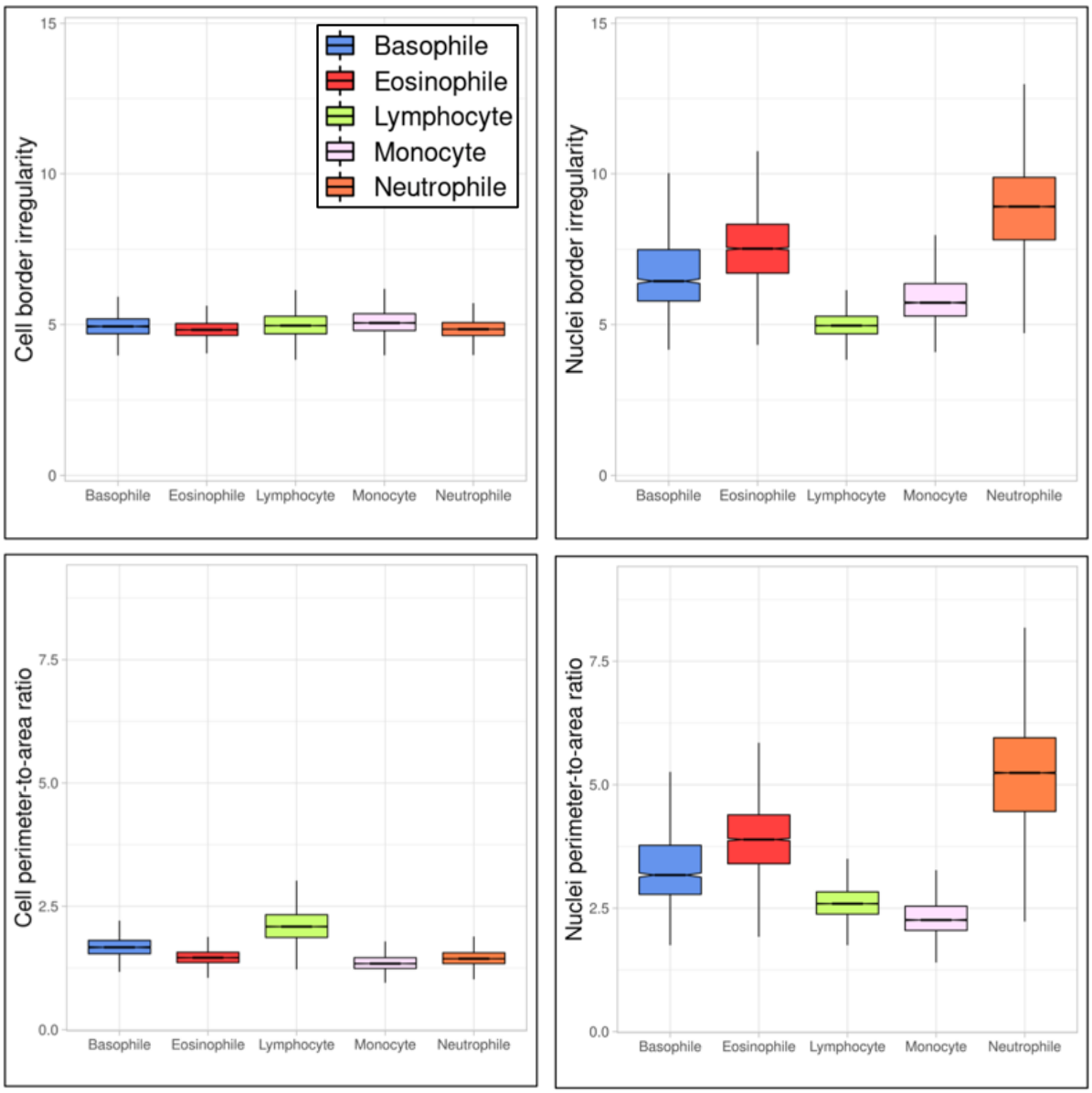
		MEAN ± 2SD = "Normality Interval"									
		BASOPHILE		EOSINOPHILE		LYMPHOCYTE		MONOCYTE		NEUTROPHILE	
MORPHOLOGY	Nuclei area	34.91	78.30	34.82	66.50	30.54	69.40	61.02	113.58	25.12	52.79
	Nuclei perimeter	106.24	251.38	124.95	251.63	91.33	158.55	130.02	260.49	127.97	257.13
	Nuclei diameter	9.09	14.05	9.50	16.15	7.37	12.04	10.88	17.16	8.98	13.74
	Nuclei major axis	8.22	13.91	8.81	19.33	6.79	11.62	10.10	17.00	8.72	15.56
	Nuclei minor axis	5.98	10.27	4.75	9.88	5.80	9.13	7.32	11.86	4.95	10.61
	Nuclei perimeter-to-area ratio	1.53	5.22	2.25	5.60	0.21	5.23	1.34	3.37	2.99	7.40
	Cell area	73.67	165.97	95.36	206.66	33.09	127.11	112.09	281.01	95.02	206.54
	Cell perimeter	130.64	255.32	148.55	275.80	96.91	219.60	172.69	334.12	158.20	263.97
	Cell diameter	9.98	19.07	11.72	19.45	7.57	16.62	12.88	25.02	12.04	19.44
	Cell major axis	10.08	17.13	11.47	18.22	7.22	15.22	12.15	23.24	11.61	18.37
	Cell minor axis	9.36	14.41	11.08	16.09	6.76	12.54	11.67	18.39	10.59	16.04
	Cell perimeter-to-area ratio	1.18	2.21	1.11	1.84	1.43	2.82	1.01	1.71	1.12	1.80
	Nuclei-to-cell ratio	0.32	0.64	0.25	0.43	0.42	0.87	0.32	0.59	0.18	0.34
	Cell border irregularity	3.80	6.19	3.92	5.85	3.98	6.08	4.18	6.06	4.14	5.59
	Nuclei border irregularity	4.06	9.47	5.09	9.92	3.88	6.18	4.03	7.81	5.76	11.89
	Nuclei eccentricity	-0.04	0.17	-0.02	0.20	-0.04	0.15	-0.02	0.17	-0.01	0.14

Figure 6: Morphometric features of cells and their nuclei: Area, diameter, major axis, minor axis, perimeter, border irregularity, perimeter-to-area



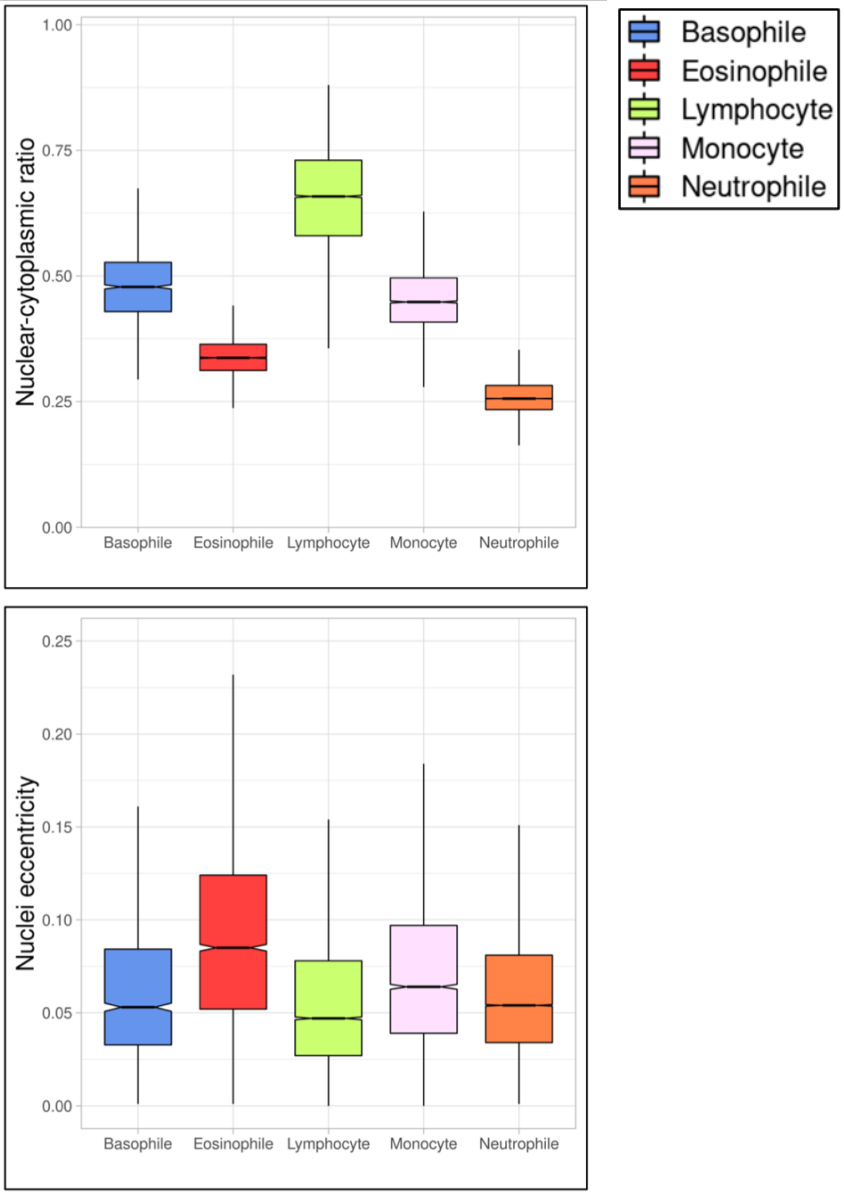


Description: *The summary of the data is shown as a boxplot, with the box indicating the IQR, the whiskers showing the range of values that are within 1.5*IQR and a horizontal line indicating the median. The notches represent for each median the 95%. The nuclei features are always on the right side of the figure.*



Description: *The summary of the data is shown as a boxplot, with the box indicating the IQR, the whiskers showing the range of values that are within 1.5*IQR and a horizontal line indicating the median. The notches represent for each median the 95%. The nuclei features are always on the right side of the figure.*

Figure 7: Morphometric features of cells or their nuclei - Nuclear-cytoplasmic ratio, and nuclei eccentricity



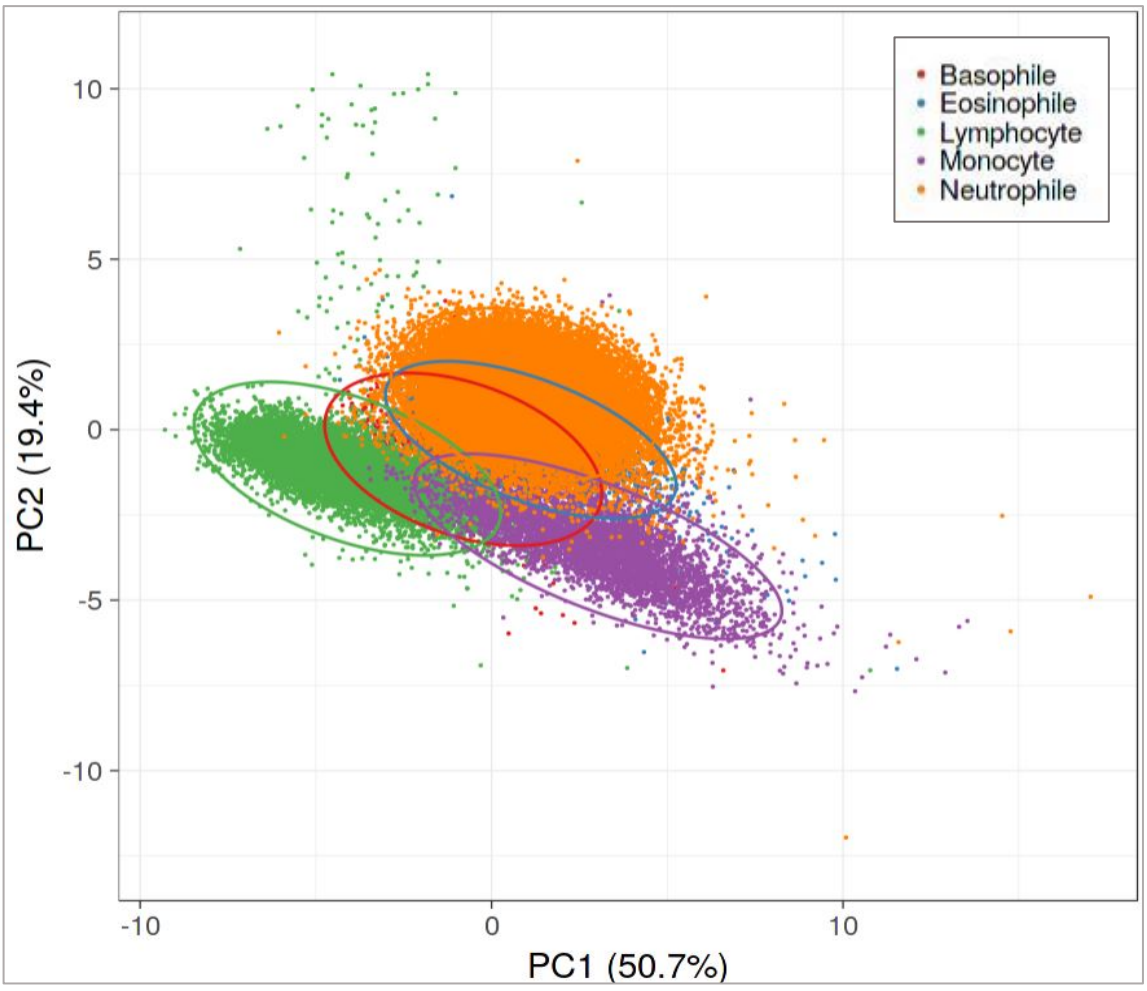
Description: *The summary of the data is shown as a boxplot, with the box indicating the IQR, the whiskers showing the range of values that are within 1.5*IQR and a horizontal line indicating the median. The notches represent for each median the 95%.*

Table 15: PCA Loadings of morphometric features

	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8
Nuclei area	-0,10	0,51	-0,18	-0,15	-0,01	0,03	-0,03	0,42
Nuclei perimeter	-0,29	-0,16	-0,01	-0,34	-0,15	0,23	-0,07	0,44
Nuclei diameter	-0,29	0,20	-0,11	0,08	-0,06	0,49	-0,05	-0,05
Nuclei major axis	-0,28	0,03	-0,09	0,18	-0,08	0,61	0,07	-0,42
Nuclei minor axis	-0,15	0,13	-0,09	-0,80	-0,13	-0,19	0,07	-0,51
Nuclei perimeter-to-area ratio	-0,12	-0,50	0,17	-0,11	-0,09	0,05	-0,08	0,00
Cell area	-0,34	0,08	-0,05	0,07	0,07	-0,18	0,09	0,09
Cell perimeter	-0,31	0,15	0,28	0,08	-0,04	-0,15	0,20	0,05
Cell diameter	-0,31	0,15	0,17	0,09	0,05	-0,19	-0,46	-0,05
Cell major axis	-0,32	0,11	0,12	0,10	0,06	-0,18	-0,51	-0,08
Cell minor axis	-0,32	0,03	-0,13	0,07	0,06	-0,16	0,61	0,14
Cell perimeter-to-area ratio	0,30	0,10	0,36	-0,09	-0,17	0,19	-0,03	0,02
Nuclei-to-cell ratio	0,26	0,33	-0,07	-0,21	-0,06	0,21	-0,13	0,24
Cell border irregularity	-0,02	0,23	0,74	0,01	-0,24	0,04	0,24	-0,01
Nuclei border irregularity	-0,20	-0,42	0,10	-0,20	-0,11	0,12	-0,07	0,31
Nuclei eccentricity	0,01	0,00	-0,27	0,23	-0,91	-0,22	-0,05	-0,01
Individual Explained Variance	0,51	0,19	0,09	0,07	0,06	0,04	0,02	0,01
Cumulative Explained Variance	0,51	0,70	0,79	0,85	0,91	0,96	0,97	0,98

Description: *An heatmap was applied for each column for better visualization. The green cell stands for the largest value and the red cell for the smallest value.*

Figure 8: Principal component analysis (PCA) score plot of morphometric features



Description: *Unit variance scaling is applied to rows; SVD with imputation is used to calculate principal components. X and Y axis show PC 1 and PC 2 that explain 50,7% and 19,4% of the total variance, respectively. Prediction ellipses are such that with probability of 0.95, a new observation from the same group will fall inside the ellipse. N = 50942 data points.*

4.2.2.Texture features

Even though the texture features and their values are abstract, one can easily retrieve useful information and gain some broad intuitions if one compares values between the five WBC using a heatmap. Based on this premise, and **Table 8** which give as the key intuitions behind radiomics concepts, one can notice that basophils hold the most **complex textures**, followed by Lymphocytes, Eosinophils and Monocytes (whose texture complexity is similar) and finally Neutrophiles that have the simplest texture. The concept of complexity was captured by Entropy, Uniformity, Autocorrelation and Variance.

Unfortunately, this finding is basically the only fact retained despite a large number of variables since PCA analysis was unable – as with morphometric features – to identify key features. To support such claim, one can review PCA results and see that data can only be explained by a substantial number of features, and not, by a single feature. For instance, PC1 which explains 56.7% of data variability, depends roughly on 60 features from the total of 86 features – and where each has approximately an absolute value between 0,10 and 0,14. Even worse, there's not a specific method among the five used (e.g., First Order, GLCM, etc) that stands out since the 60 features are equally distributed among them. The remaining 26 features are basically the dominant features of PC2, PC3, PC4, PC5 and PC6 which account for 10%, 8%, 6%, 4%, and 3% of data variability, respectively. These 6 PC explain a total of 88% data variance. The remaining 12% are distributed between a dozen of PC that accounts for 1-2% each.

Moreover, when reviewing the PCA score plot (**Figure 9**), one can notice that there is an overlap among classes. The sole exception is Basophils and the "trio" Eosinophils, Monocytes and Neutrophiles. This finding means that PCA, as with morphometric features, is unsuitable for WBC classification despite the higher and more complex PCs. Nevertheless, PCA allowed us to appreciate the magnitude of cell variability: Lymphocytes have by far the largest data variability as expected, followed by Basophils, Neutrophiles, Monocytes and Eosinophils.

Simply put, we were again **unable to identify new image texture biomarkers** of mature white blood cells. The mean values with their standard deviation, PCA loadings and PCA plot score, are provided in **Table 16**, **Table 17** and **Figure 9**, respectively.

Table 16: Heatmap mean and Standard deviation (SD) of texture features

		MEAN					STANDARD DEVIATION				
		BASOPHILE	EOSINOPHILE	LYMPHOCYTE	MONOCYTE	NEUTROPHILE	BASOPHILE	EOSINOPHILE	LYMPHOCYTE	MONOCYTE	NEUTROPHILE
First Order	10Percentile	24.80	82.25	57.51	105.98	137.17	17.98	13.79	30.86	22.20	16.38
	90Percentile	155.35	160.14	165.73	181.47	190.33	23.61	19.44	19.39	19.40	11.04
	Energy	49013761.62	120783370.97	34644557.73	198902283.62	251136622.17	29829176.19	47067700.99	40269012.57	102179268.86	75795810.25
	Entropy	2.84	2.31	2.65	2.30	1.94	0.21	0.17	0.29	0.26	0.21
	InterquartileRange	76.53	35.79	57.83	36.85	25.07	17.56	9.11	21.94	11.60	6.97
	Kurtosis	2.66	4.28	3.03	4.03	7.51	0.82	1.12	1.29	1.22	2.66
	Maximum	232.37	238.57	220.81	235.36	230.19	11.17	10.16	12.49	10.38	6.24
	MeanAbsoluteDeviation	41.40	24.11	33.58	23.51	17.86	6.70	4.00	9.09	5.54	3.27
	Mean	87.47	116.97	110.89	143.31	164.28	23.83	14.13	25.38	18.15	12.25
	Median	83.90	112.30	110.09	143.40	167.21	30.94	14.01	30.79	19.00	13.00
	Minimum	0.55	21.72	10.42	21.74	21.62	1.64	11.92	10.07	13.21	12.68
	Range	231.82	216.85	210.40	213.62	208.57	11.35	13.95	14.22	15.82	12.78
	RobustMeanAbsoluteDeviation	31.45	15.59	24.25	15.73	10.79	6.62	3.67	8.53	4.68	2.86
	RootMeanSquared	101.30	121.37	119.30	146.78	166.30	21.63	13.80	22.66	17.33	11.95
	Skewness	0.34	0.83	0.03	-0.20	-1.40	0.56	0.31	0.60	0.53	0.66
	TotalEnergy	49013761.62	120783370.97	34644557.73	198902283.62	251136622.17	29829176.19	47067700.99	40269012.57	102179268.86	75795810.25
	Uniformity	0.16	0.26	0.19	0.26	0.34	0.03	0.04	0.06	0.05	0.06
	Variance	2510.74	1039.30	1803.64	977.77	660.31	669.99	279.68	816.34	417.65	213.25
GLCM	Autocorrelation	19.57	23.94	25.62	35.65	46.24	8.61	5.55	10.85	9.09	9.00
	JointAverage	3.93	4.74	4.75	5.82	6.71	1.01	0.56	1.07	0.78	0.68
	ClusterProminence	496.26	122.13	257.32	112.74	65.72	255.24	71.51	251.65	107.24	75.30
	ClusterShade	13.68	9.11	2.28	0.07	-5.80	26.04	6.46	20.68	7.41	6.03
	ClusterTendency	13.71	5.28	8.71	5.24	3.07	4.33	1.81	4.53	2.64	1.29
	Contrast	1.31	0.48	1.22	0.39	0.32	0.39	0.09	0.86	0.14	0.07
	Correlation	0.82	0.82	0.76	0.85	0.81	0.04	0.04	0.08	0.05	0.03
	DifferenceAverage	0.77	0.40	0.70	0.34	0.27	0.14	0.05	0.32	0.10	0.06
	DifferenceEntropy	1.58	1.09	1.45	0.99	0.90	0.17	0.08	0.33	0.14	0.10
	DifferenceVariance	0.68	0.31	0.60	0.26	0.24	0.17	0.04	0.30	0.07	0.04
	JointEnergy	0.06	0.14	0.10	0.16	0.25	0.03	0.03	0.07	0.06	0.07
	JointEntropy	4.69	3.54	4.24	3.38	2.78	0.40	0.29	0.72	0.47	0.37
	Imc1	-0.30	-0.38	-0.31	-0.44	-0.42	0.05	0.04	0.09	0.07	0.04
	Imc2	0.89	0.89	0.86	0.91	0.87	0.03	0.03	0.05	0.04	0.03
	Idm	0.67	0.81	0.70	0.84	0.87	0.05	0.02	0.11	0.04	0.03
	Idmn	0.99	1.00	0.99	1.00	1.00	0.00	0.00	0.01	0.00	0.00
	Id	0.69	0.81	0.72	0.84	0.87	0.04	0.02	0.10	0.04	0.03
	Idn	0.93	0.96	0.94	0.97	0.97	0.01	0.00	0.03	0.01	0.01
	InverseVariance	0.44	0.34	0.39	0.30	0.24	0.04	0.03	0.08	0.06	0.05
	MaximumProbability	0.16	0.28	0.20	0.29	0.41	0.07	0.06	0.11	0.09	0.09
	SumEntropy	3.64	3.01	3.32	2.95	2.45	0.26	0.23	0.38	0.35	0.30
	SumSquares	3.75	1.44	2.48	1.41	0.85	1.14	0.46	1.28	0.68	0.33
GLRLM	GrayLevelNonUniformity	435,89	739,46	203,56	714,24	745,80	145,28	177,16	115,56	204,44	166,33
	GrayLevelNonUniformityNormalized	0,15	0,20	0,17	0,20	0,23	0,02	0,02	0,03	0,03	0,02
	GrayLevelVariance	4,02	2,30	3,29	2,17	2,01	0,83	0,38	1,10	0,55	0,40
	HighGrayLevelRunEmphasis	21,70	27,27	27,14	34,87	44,04	6,76	5,39	7,66	7,65	7,85
	LongRunEmphasis	3,65	9,55	5,24	13,69	19,94	1,30	2,29	6,43	7,64	7,99
	LongRunHighGrayLevelEmphasis	73,79	226,67	178,89	537,34	966,57	59,72	78,77	329,84	377,43	473,49
	LongRunLowGrayLevelEmphasis	1,21	0,53	0,57	0,44	0,46	0,97	0,20	3,17	0,21	0,24
	LowGrayLevelRunEmphasis	0,20	0,06	0,10	0,05	0,03	0,08	0,02	0,06	0,02	0,01
	RunEntropy	4,16	4,60	3,96	4,70	4,72	0,18	0,15	0,41	0,27	0,17
	RunLengthNonUniformity	1621,38	1253,96	635,97	1143,62	964,44	477,22	356,35	158,90	275,99	220,21
	RunLengthNonUniformityNormalized	0,56	0,34	0,60	0,33	0,30	0,06	0,04	0,15	0,07	0,04
	RunPercentage	0,67	0,46	0,68	0,42	0,36	0,06	0,04	0,16	0,08	0,05
	RunVariance	1,33	4,53	2,56	7,38	11,75	0,79	1,45	4,75	5,38	5,91
	ShortRunEmphasis	0,77	0,60	0,79	0,59	0,56	0,04	0,04	0,10	0,06	0,04
	ShortRunHighGrayLevelEmphasis	17,08	17,62	21,37	20,02	23,63	4,60	3,49	4,60	3,94	4,14
	ShortRunLowGrayLevelEmphasis	0,13	0,04	0,08	0,03	0,02	0,05	0,01	0,04	0,01	0,01
GLSZM	GrayLevelNonUniformity	155,87	119,47	68,06	110,91	98,20	45,28	24,89	15,65	21,13	18,08
	GrayLevelNonUniformityNormalized	0,14	0,15	0,14	0,15	0,15	0,01	0,01	0,01	0,01	0,01
	GrayLevelVariance	4,12	3,67	3,84	3,55	3,79	0,48	0,38	0,64	0,41	0,42
	HighGrayLevelZoneEmphasis	23,63	32,35	27,95	32,81	35,73	4,10	5,35	4,77	5,96	5,75
	LargeAreaEmphasis	221,07	6294,24	1189,22	9706,53	15518,55	623,67	4378,58	4141,27	12130,04	13941,21
	LargeAreaHighGrayLevelEmphasis	4376,13	130232,23	50359,38	371249,25	752614,65	12824,95	97944,36	244106,81	505235,25	735153,60
	LargeAreaLowGrayLevelEmphasis	81,80	339,79	126,76	281,07	340,63	186,08	270,38	1275,10	342,42	338,56
	LowGrayLevelZoneEmphasis	0,14	0,06	0,10	0,06	0,06	0,03	0,02	0,03	0,02	0,02
	SizeZoneNonUniformity	354,66	221,43	190,05	200,38	189,40	98,32	48,94	31,56	42,92	36,58
	SizeZoneNonUniformityNormalized	0,32	0,27	0,41	0,27	0,28	0,05	0,05	0,09	0,04	0,04
	SmallAreaEmphasis	0,59	0,54	0,66	0,53	0,55	0,05	0,05	0,07	0,04	0,04
	SmallAreaHighGrayLevelEmphasis	14,65	18,48	18,80	16,99	17,38	2,42	3,74	3,11	3,30	3,23
	SmallAreaLowGrayLevelEmphasis	0,07	0,03	0,06	0,04	0,04	0,01	0,01	0,02	0,02	0,01
	ZoneEntropy	5,34	5,61	4,79	5,55	5,43	0,27	0,27	0,36	0,24	0,22
	ZonePercentage	0,26	0,11	0,34	0,10	0,08	0,07	0,02	0,17	0,04	0,02
	ZoneVariance	202,64	6193,64	1165,64	9548,52	15328,35	609,10	4352,04	4107,53	12042,42	13882,39
GLDM	DependenceEntropy	5,49	5,25	5,09	5,17	4,66	0,20	0,19	0,35	0,34	0,32
	DependenceNonUniformity	684,29	980,82	319,38	1180,72	1342,21	217,02	271,95	203,37	534,74	457,82
	DependenceNonUniformityNormalized	0,16	0,12	0,19	0,13	0,15	0,02	0,00	0,08	0,02	0,02
	DependenceVariance	4,16	5,25	4,07	5,43	5,88	0,83	0,45	2,00	0,56	0,49
	GrayLevelNonUniformity	724,60	2026,59	452,19	2253,32	3064,85	317,49	528,00	476,08	1006,00	922,70
	GrayLevelVariance	4,11	1,75	2,97	1,65	1,14	1,08	0,45	1,31	0,67	0,34
	HighGrayLevelEmphasis	21,14	25,38	27,21	36,00	46,24	8,16	5,55	9,49	8,68	8,73
	LargeDependenceEmphasis	17,76	33,94	18,14	37,28	43,25	4,43	3,73	11,40	7,43	4,93
	LargeDependenceHighGrayLevelEmphasis	348,98	796,07	554,93	1408,20	2071,62	254,04	197,99	600,94	525,92	510,39
	LargeDependenceLowGrayLevelEmphasis	6,64	1,92	1,89	1,27	1,01	4,71	0,64	3,13	0,47	0,31
	LowGrayLevelEmphasis	0,24	0,06	0,11	0,04	0,03	0,12	0,02	0,08	0,02	0,01
	SmallDependenceEmphasis	0,23	0,11	0,30	0,10	0,08	0,06	0,02	0,15	0,03	0,02
	LowGrayLevelEmphasis	5,55	3,38	8,37	3,36	3,26	1,88	0,88	4,06	1,16	0,75
	SmallDependenceEmphasis	0,03	0,01	0,03	0,01	0,00	0,01	0,00	0,02	0,00	0,00

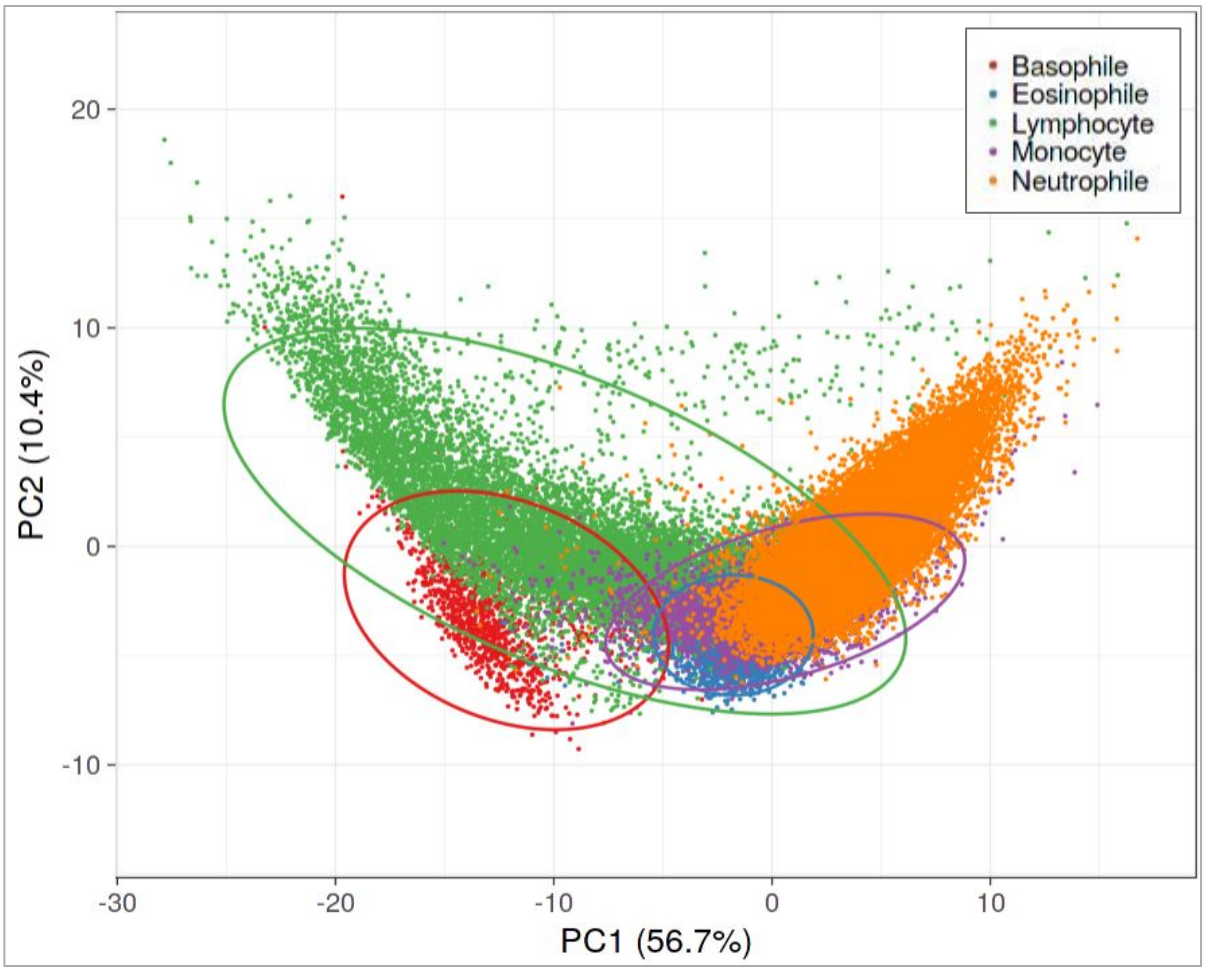
***Description**: An heatmap was applied for each row for better visualization. The green cell stands for the largest value and the red cell for the smallest value.*

Table 17: PCA Loadings of texture features

		PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8
First Order	10Percentile	0,14	-0,06	0,01	0,04	0,01	0,08	0,02	-0,10
	90Percentile	0,08	-0,13	-0,15	0,00	-0,02	0,27	-0,04	-0,24
	Energy	0,12	-0,04	-0,06	-0,09	0,16	0,15	0,01	-0,04
	Entropy	-0,13	0,05	-0,12	0,00	-0,02	0,06	-0,07	0,05
	InterquartileRange	-0,13	-0,02	-0,10	-0,09	0,00	0,09	-0,06	-0,05
	Kurtosis	0,11	-0,15	0,06	-0,05	0,03	-0,11	0,16	0,01
	Maximum	0,05	0,05	-0,19	-0,05	0,09	0,05	-0,02	-0,14
	MeanAbsoluteDeviation	-0,13	-0,02	-0,11	-0,08	-0,03	0,05	-0,04	-0,04
	Mean	0,13	-0,09	-0,05	0,05	-0,02	0,16	0,02	-0,16
	Median	0,12	-0,09	-0,06	0,06	-0,04	0,16	0,04	-0,16
	Minimum	0,06	0,08	0,15	0,03	-0,02	0,30	-0,07	-0,27
	Range	-0,02	-0,03	-0,28	-0,06	0,09	-0,24	0,05	0,14
	RobustMeanAbsoluteDeviation	-0,13	-0,02	-0,10	-0,09	-0,01	0,08	-0,06	-0,05
	RootMeanSquared	0,12	-0,10	-0,07	0,03	-0,02	0,17	0,02	-0,18
	Skewness	-0,10	0,18	0,01	0,00	0,04	-0,04	-0,17	0,09
	TotalEnergy	0,12	-0,04	-0,06	-0,09	0,16	0,15	0,01	-0,04
	Uniformity	0,13	-0,09	0,11	-0,03	0,02	-0,09	0,09	-0,04
	Variance	-0,13	-0,06	-0,12	-0,11	-0,03	0,04	0,00	-0,05
GLCM	Autocorrelation	0,11	-0,15	-0,15	0,05	-0,01	-0,01	0,02	-0,07
	JointAverage	0,12	-0,13	-0,14	0,07	-0,01	0,02	0,01	-0,06
	ClusterProminence	-0,09	-0,07	-0,15	-0,17	-0,07	0,02	0,06	-0,11
	ClusterShade	-0,06	0,09	0,08	-0,09	0,13	-0,10	-0,11	0,07
	ClusterTendency	-0,12	-0,03	-0,14	-0,13	-0,02	0,06	-0,02	-0,07
	Contrast	-0,12	-0,12	0,00	-0,01	0,11	0,07	-0,01	-0,04
	Correlation	0,06	0,14	-0,21	-0,16	-0,20	0,05	-0,08	0,08
	DifferenceAverage	-0,13	-0,08	0,00	0,01	0,12	0,05	-0,01	-0,02
	DifferenceEntropy	-0,14	-0,04	-0,02	0,02	0,09	0,01	0,02	-0,03
	DifferenceVariance	-0,13	-0,11	-0,01	-0,02	0,08	0,03	0,02	-0,05
	JointEnergy	0,12	-0,11	0,09	-0,06	-0,05	-0,08	0,08	-0,04
	JointEntropy	-0,14	0,03	-0,08	0,02	0,07	0,06	-0,06	0,03
	Imc1	-0,11	-0,04	0,11	0,13	0,20	-0,10	0,10	-0,06
	Imc2	0,01	0,15	-0,24	-0,12	-0,20	0,15	-0,16	0,10
	Idm	0,14	0,05	0,01	-0,03	-0,12	-0,04	0,01	0,01
	Idmn	0,13	0,11	-0,04	0,00	-0,10	-0,09	0,01	0,04
	Id	0,14	0,04	0,01	-0,03	-0,12	-0,03	0,02	0,01
	Idn	0,14	0,06	-0,02	-0,02	-0,11	-0,07	0,02	0,02
	InverseVariance	-0,13	0,08	-0,03	0,08	0,12	-0,01	-0,03	0,03
	MaximumProbability	0,12	-0,08	0,09	-0,06	-0,06	-0,10	0,07	-0,04
	SumEntropy	-0,13	0,08	-0,12	0,01	0,03	0,08	-0,08	0,04
	SumSquares	-0,12	-0,04	-0,12	-0,12	0,00	0,06	-0,02	-0,07
GLRLM	GrayLevelNonUniformity	0,11	0,09	-0,05	-0,06	0,25	0,06	0,06	-0,04
	GrayLevelNonUniformityNormalized	0,12	-0,01	0,13	0,04	0,09	-0,04	0,07	-0,09
	GrayLevelVariance	-0,11	-0,12	-0,14	-0,14	-0,10	0,01	0,06	-0,01
	HighGrayLevelRunEmphasis	0,11	-0,15	-0,16	0,06	0,03	-0,05	-0,02	-0,06
	LongRunEmphasis	0,12	-0,11	0,06	-0,16	-0,01	-0,01	-0,06	0,08
	LongRunHighGrayLevelEmphasis	0,11	-0,16	-0,01	-0,10	0,01	-0,02	0,00	0,11
	LongRunLowGrayLevelEmphasis	0,00	0,02	0,06	-0,22	-0,06	-0,24	-0,28	-0,37
	LowGrayLevelRunEmphasis	-0,12	0,01	0,07	-0,19	-0,01	-0,07	0,11	-0,05
	RunEntropy	0,13	0,06	-0,09	-0,12	-0,05	0,05	-0,02	0,09
	RunLengthNonUniformity	0,03	0,19	-0,17	-0,08	0,27	-0,01	0,20	-0,08
	RunLengthNonUniformityNormalized	-0,13	-0,09	0,02	0,05	0,01	-0,01	-0,01	-0,02
	RunPercentage	-0,14	-0,03	0,00	0,07	0,05	0,01	-0,02	-0,01
	RunVariance	0,11	-0,12	0,07	-0,16	-0,02	-0,03	-0,07	0,07
	ShortRunEmphasis	-0,13	-0,07	0,01	0,06	-0,02	-0,04	0,01	-0,02
	ShortRunHighGrayLevelEmphasis	0,04	-0,21	-0,20	0,15	0,01	-0,12	-0,06	-0,11
	ShortRunLowGrayLevelEmphasis	-0,12	-0,03	0,07	-0,14	-0,01	-0,02	0,11	-0,02
GLSZM	GrayLevelNonUniformity	0,04	0,18	-0,12	-0,08	0,31	-0,06	0,24	-0,11
	GrayLevelNonUniformityNormalized	0,04	0,13	0,17	0,10	0,18	0,03	-0,13	-0,23
	GrayLevelVariance	-0,02	-0,17	-0,15	-0,16	-0,15	-0,04	0,14	0,20
	HighGrayLevelZoneEmphasis	0,08	-0,10	-0,19	0,13	0,10	-0,19	-0,16	-0,02
	LargeAreaEmphasis	0,09	-0,10	0,08	-0,18	0,16	0,05	-0,16	0,22
	LargeAreaHighGrayLevelEmphasis	0,09	-0,14	0,04	-0,16	0,17	0,02	-0,12	0,23
	LargeAreaLowGrayLevelEmphasis	0,03	0,01	0,09	-0,24	0,02	-0,17	-0,36	-0,24
	LowGrayLevelZoneEmphasis	-0,10	0,01	0,14	-0,19	-0,09	0,13	0,22	0,00
	SizeZoneNonUniformity	-0,04	0,03	-0,14	-0,10	0,19	-0,18	0,41	-0,07
	SizeZoneNonUniformityNormalized	-0,10	-0,20	0,05	0,02	-0,09	-0,05	0,06	-0,01
	SmallAreaEmphasis	-0,10	-0,19	0,05	0,03	-0,10	-0,08	0,08	-0,01
	SmallAreaHighGrayLevelEmphasis	-0,02	-0,16	-0,16	0,16	0,06	-0,30	-0,16	0,01
	SmallAreaLowGrayLevelEmphasis	-0,08	-0,06	0,16	-0,15	-0,14	0,20	0,21	-0,02
	ZoneEntropy	0,09	0,19	-0,12	-0,05	0,11	0,06	-0,03	0,03
	ZonePercentage	-0,13	-0,12	0,03	0,04	0,05	0,04	-0,04	0,00
	ZoneVariance	0,09	-0,10	0,08	-0,18	0,16	0,05	-0,16	0,22
GLDM	DependenceEntropy	-0,08	0,22	-0,17	0,05	-0,07	0,00	-0,03	0,04
	DependenceNonUniformity	0,11	-0,02	-0,02	-0,18	0,19	0,09	-0,03	0,11
	DependenceNonUniformityNormalized	-0,07	-0,24	0,07	-0,07	0,13	0,12	-0,12	0,05
	DependenceVariance	0,11	0,05	-0,02	-0,01	-0,23	-0,15	0,13	-0,07
	GrayLevelNonUniformity	0,13	-0,04	0,02	-0,12	0,18	0,04	0,02	0,04
	GrayLevelVariance	-0,13	-0,06	-0,12	-0,11	-0,03	0,04	0,00	-0,05
	HighGrayLevelEmphasis	0,11	-0,15	-0,16	0,05	0,01	-0,02	0,01	-0,07
	LargeDependenceEmphasis	0,14	0,00	0,01	-0,08	-0,05	0,00	0,01	0,01
	LargeDependenceHighGrayLevelEmphasis	0,13	-0,11	-0,08	-0,02	-0,01	-0,01	0,04	-0,01
	LargeDependenceLowGrayLevelEmphasis	-0,05	0,08	0,06	-0,29	-0,04	-0,25	-0,05	-0,26
	LowGrayLevelEmphasis	-0,11	0,01	0,06	-0,21	0,01	-0,13	0,05	-0,12
	SmallDependenceEmphasis	-0,13	-0,12	0,03	0,04	0,05	0,04	-0,04	-0,01
	LowGrayLevelEmphasis	-0,12	-0,16	-0,03	0,08	0,05	0,01	-0,08	-0,03
	SmallDependenceEmphasis	-0,13	-0,11	0,06	-0,06	0,04	0,05	0,04	-0,02
	Individual Explained Variance	0,57	0,10	0,08	0,06	0,04	0,03	0,02	0,02
	Cumulative Explained Variance	0,57	0,67	0,75	0,81	0,85	0,88	0,90	0,92

Description: *An heatmap was applied for each column for better visualization. The green cell stands for the largest value and the red cell for the smallest value.*

Figure 9: Principal component analysis (PCA) score plot of texture features



Description: *Unit variance scaling is applied to rows; SVD with imputation is used to calculate principal components. X and Y axis show PC 1 and PC 2 that explain 56,7% and 10,4% of the total variance, respectively. Prediction ellipses are such that with probability of 0.95, a new observation from the same group will fall inside the ellipse. N = 50942 data points.*

5. Problems and recommendations for future works

In this section, we enumerate a series of issues that limit the scope and validity of our results, but also of the proposed workflow.

Data selection bias: The primary aim of this thesis was to identify biomarkers of mature White Blood Cells. However, for logistic reasons, data came from hospitals. Thus, it is reasonable to think that some patients were sick and that several white blood cells were indeed pathologic. To partially overcome such hurdles, we eliminated patients from oncological and haematological departments, but also patients with an abnormal white blood cell count. Even though such workaround is imperfect, since oncological-haematological patients can be hospitalized in other departments, and also because patients can have abnormal cells with normal WBC count, it will hopefully reduce the problem to such extent that it will not impact significantly the validity of our results.

Furthermore, we eliminated during the post-processing stage multi-cell images and low confident predictions. The main purpose of this approach was to avoid any misclassifications and consequently, improve the precision of the results. This can however create a systematic bias in our results by eliminating the more atypical cells and decrease results exactitude since we don't analyse the full spectrum of normality. To mitigate this problem, one could, first of all, increase our training data to obtain more confident predictions which would decrease the number of low confident predictions. Secondly, one could also alter the deep learning algorithms and radiomics scripts to take multi-cell images.

Bayesian Error: Some images, especially the subtypes of immature granulocytes, are difficult to classify even for experts with years of practice. Consequently, it is reasonable to assume that some samples were mislabelled and that our deep learning algorithms could be further improved by correcting them. Even though such a problem is not of critical significance for our thesis since our evaluation metrics were excellent, and the more troublesome classes (immature granulocytes) were not analysed, it cannot be forgotten for future and more complex works. One conceivable solution would be the usage of immunofluorescence staining markers to classify the cell labels with 100% certainty. Such proposal, which is not new [84], would allow the algorithms to outperform humans, but would also require a more labour-intensive pre-processing step.

Black box problem: One of the most challenging barriers to the deployment of such algorithms in the medical field, especially after the GDPR [85], is our inability to fully understand and explain the predictions. However, the implications of the black box problem are so profound to deep learning applications, that solutions will hopefully emerge in the coming years since it is a very active research area [86]. For example, the *Grad algorithm* [87] where one visualizes the pieces of the image that are important for the classification task.

Besides, we should not forget that experts classify cells primarily on intuition, which depends primarily on their clinical experience. In other words, the human mind is also a black box and yet, is still used to acquire knowledge. The same applies to deep learning. For instance, model interpretability can be used to identify new hallmarks of diseases [88].

Flexibility and adaptability: The nature of our deep learning algorithms implies that the acquired knowledge is fixed and cannot be altered - except if we re-train it. This can be particularly troublesome if we need to slightly adapt the algorithm (e.g., change the image resolution, the type of staining, add an extra class…) and discourage the industry to deploy them. One possible solution would be *Active learning* where the artificial neural network is continuously learning by receiving constant feedback from experts. However, this proposal is unlikely to thrive because of the profound implications of mislabelled medical data.

Generalization problem: Even though the performance of our deep learning algorithms seems impressive, it is only valid for our very particular dataset as deep learning is very data-specific. For example, a drop of performance of 20% was seen in some works when applying the algorithm on an independent data set [89]. To mitigate this problem and develop more robust and generalizable algorithms, one could simply train them in different datasets, at the expense of a more burdensome manual labelling. One could also standardize the pre-processing steps needed for blood analysis. The main goal would be to fully capture the broad biological and morphological variability of cells, as well as the technical variability inherent to cytopathology.

Translatability: We identify three main reasons that could discourage experts from the usage of our approach in research and clinical setting. First, the lack of gold-standard data for proper validation of the results. Second, the difficult interpretability of the texture features. Finally, the fact that data is usually collected retrospectively.

It is, therefore, necessary to overcome these 3 key points to allow such workflow, and especially the usage of radiomics features, to be widely accepted and deployed by the research and clinical community.

Also, one needs a regulatory framework to avoid the risk of non-reproducible laboratory-specific methods [90]. The AI algorithms need to output consistent results - sine qua non towards their clinical usage.

Feature selection: Of all the above-mentioned issues, feature selection is undoubtedly the most tangible problem currently affecting our thesis.

As a reminder, a biomarker has minimum intra-cell variability and maximum inter-cell variability, and our main selection method – PCA – should be able to detect those properties which is not the case. PCA is primarily designed for dimensionality reduction by maximising variance and not necessarily for feature selection (by finding clusters), which means that key features for PCA - the ones with most variation - are not necessarily the most relevant ones to classify a cell. Even though both go often hand in hand, they can sometimes diverge. For instance, features with high intra-cell variability which we do not want can be sometimes of key importance for PCA.

To overcome this problem, one could use other popular machine learning technics [91] such as Linear Discriminant Analysis (LDA), Recursive Feature Elimination, Random Forest algorithm and Lasso Regression – just to mention a few.

6. Conclusions and Perspectives

Despite the analysis of more than 100 features – including the most well-known ones by the biomedical community such as the nuclei-to-cytoplasm ratio, area and diameter – and almost 55.000 images, our proposed workflow that uses a two-step approach, was unable to identify new image biomarkers of mature white blood cells. Nevertheless, it allowed us to corroborate what histological experts state during histologic classes – that one can only classify a white blood cell with several features as there is not a distinct feature that stands out.

More importantly, our deep learning algorithms (step 1) proved to be notably accurate which allowed us to verify values of some well-known cell features. Not only that, but they also enabled us to establish some new gold-standard norms to be used by the biomedical community – provided that one is aware of their bias. In particular, the **data selection bias** since data came from hospitals and, some undesired images were not analysed.

Yet, the major issue of our work is the **feature selection process** (step 2) – sine qua non for biomarker identification. Hence, future works should focus on addressing it if one wants to implement our two-step approach to other imaging modalities and/or tasks.

For instance, one could use it to **diagnose** diseases by analysing patient's blood smears, but also to stratify patients based on **prognostic** biomarkers and phenotype. Such a tool is especially of high value for undeveloped countries where access to the **diagnosis of malaria** [92] and **haematological diseases** [93] is still a major problem due to the lack of resources. That is why one could release a smartphone application to be used by non-medical personnel – instead of implementing it directly on the few available high-tech microscopes. This application would be able to seamlessly analyse photos taken from traditional microscopes using a smartphone and should be sensitive enough to answer the following question: Should the patient be sent to a hospital?

Moreover, one could also use our approach for **educational purposes**. For instance, to identify and segment a specific cell type on histo(patho)logy slides on student's request and/or giving him feedback if (s)he is unsure about a specific cell. For example, it is well known by students that macrophages are difficult to find in gastrointestinal tract slides due to their scarcity. The software add-on could also give him/her details about a particularly atypical cell - that is, their characteristics and how they compare to cells of the same type. This would hopefully help him/her understand why (s)he had so much trouble when classifying that cell. Besides, such software add-on could be easily implemented on Cytomine [94], an open-source application used for the analysis of histological slides and used by students at UCLouvain.

Supervisor

The scientific supervisor of this publication is Professor John Lee.

Funding

The author states that this work has not received any funding.

Datasets, results and code availability

Link: https://www.dropbox.com/sh/9s5yfw3p8mymyt8/AAAKSDnTTC_xh6n27xex9-cxa?dl=0

Ethics approval and consent to participate.

Institutional review board approval was not required since no study was performed, and the data is completely anonymous.

Conflict of Interest:

The author declares that he has no competing of interests.

References

1. Hinton, G. *et al.* Deep Neural Networks for Acoustic Modeling in Speech Recognition: The Shared Views of Four Research Groups. *IEEE Signal Process. Mag.* **29**, 82–97 (2012).
2. Hirschberg, J. & Manning, C. D. Advances in natural language processing. *Science* **349**, 261–266 (2015).
3. Russakovsky, O. *et al.* ImageNet Large Scale Visual Recognition Challenge. *Int. J. Comput. Vis.* **115**, 211–252 (2015).
4. Esteva, A. *et al.* A guide to deep learning in healthcare. *Nat. Med.* **25**, 24–29 (2019).
5. Wan, J. *et al.* Deep Learning for Content-Based Image Retrieval: A Comprehensive Study. in *Proceedings of the 22nd ACM international conference on Multimedia* 157–166 (Association for Computing Machinery, 2014). doi:10.1145/2647868.2654948.
6. Anwar, S. M. *et al.* Medical Image Analysis using Convolutional Neural Networks: A Review. *J. Med. Syst.* **42**, 226 (2018).
7. Paulsen, O. & Sejnowski, T. J. Natural patterns of activity and long-term synaptic plasticity. *Curr. Opin. Neurobiol.* **10**, 172–180 (2000).
8. Bengio, Y. Learning Deep Architectures for AI. *Foundations* **2**, 1–55 (2009).
9. LeCun, Y., Bengio, Y. & Hinton, G. Deep learning. *Nature* **521**, 436–444 (2015).
10. Ronneberger, O., Fischer, P. & Brox, T. U-Net: Convolutional Networks for Biomedical Image Segmentation. *ArXiv150504597 Cs* (2015).
11. Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J. & Wojna, Z. Rethinking the Inception Architecture for Computer Vision. *ArXiv151200567 Cs* (2015).
12. *Inception v3 - Building Machine Learning Projects with TensorFlow*. (2016).
13. A guide to the types of machine learning algorithms. https://www.sas.com/en_gb/insights/articles/analytics/machine-learning-algorithms.html.
14. Ganong, W. F. Immunity, Infection & Inflammation. in *Review of medical physiology (21 ed.).* 518 (McGraw-Hill/Appleton & Lange).
15. Castellano, G., Bonilha, L., Li, L. M. & Cendes, F. Texture analysis of medical images. *Clin. Radiol.* **59**, 1061–1069 (2004).
16. Gillies, R. J., Kinahan, P. E. & Hricak, H. Radiomics: Images Are More than Pictures, They Are Data. *Radiology* **278**, 563–577 (2016).
17. Cook, G. J. R. *et al.* Radiomics in PET: principles and applications. *Clin. Transl. Imaging* **2**, 269–276 (2014).
18. Gatenby, R. A., Grove, O. & Gillies, R. J. Quantitative imaging in cancer evolution and ecology. *Radiology* **269**, 8–15 (2013).
19. van Timmeren, J. E., Cester, D., Tanadini-Lang, S., Alkadhi, H. & Baessler, B. Radiomics in medical imaging—"how-to" guide and critical reflection. *Insights Imaging* **11**, 91 (2020).
20. Kooi, T. *et al.* Large scale deep learning for computer aided detection of mammographic lesions. *Med. Image Anal.* **35**, 303–312 (2017).
21. Poplin, R. *et al.* Prediction of cardiovascular risk factors from retinal fundus photographs via deep learning. *Nat. Biomed. Eng.* **2**, 158–164 (2018).
22. Gulshan, V. *et al.* Development and Validation of a Deep Learning Algorithm for Detection of Diabetic Retinopathy in Retinal Fundus Photographs. *JAMA* **316**, 2402–2410 (2016).
23. Hosny, A., Parmar, C., Quackenbush, J., Schwartz, L. H. & Aerts, H. J. W. L. Artificial intelligence in radiology. *Nat. Rev. Cancer* **18**, 500–510 (2018).
24. Liu, Z. *et al.* The Applications of Radiomics in Precision Diagnosis and Treatment of Oncology: Opportunities and Challenges. *Theranostics* **9**, 1303–1322 (2019).
25. Deka, L. *et al.* Nuclear morphometry and texture analysis on cytological smears of thyroid neoplasms: a study of 50 cases. *Malays. J. Pathol.* **39**, 33–37 (2017).
26. Schindelin, J. *et al.* Fiji: an open-source platform for biological-image analysis. *Nat. Methods* **9**, 676–682 (2012).

27. Stylianidou, S., Brennan, C., Nissen, S. B., Kuwada, N. J. & Wiggins, P. A. SuperSegger: robust image segmentation, analysis and lineage tracking of bacterial cells. *Mol. Microbiol.* **102**, 690–700 (2016).
28. Allan, C. *et al.* OMERO: flexible, model-driven data management for experimental biology. *Nat. Methods* **9**, 245–253 (2012).
29. McQuin, C. *et al.* CellProfiler 3.0: Next-generation image processing for biology. *PLoS Biol.* **16**, e2005970 (2018).
30. Rueden, C. T. *et al.* ImageJ2: ImageJ for the next generation of scientific image data. *BMC Bioinformatics* **18**, 529 (2017).
31. Kratz, A. *et al.* Digital morphology analyzers in hematology: ICSH review and recommendations. *Int. J. Lab. Hematol.* **41**, 437–447 (2019).
32. Islam, M. M., Poly, T. N., Walther, B. A., Yang, H. C. & Li, Y.-C. J. Artificial Intelligence in Ophthalmology: A Meta-Analysis of Deep Learning Models for Retinal Vessels Segmentation. *J. Clin. Med.* **9**, (2020).
33. Arevalo, J., González, F. A., Ramos-Pollán, R., Oliveira, J. L. & Guevara Lopez, M. A. Representation learning for mammography mass lesion classification with convolutional neural networks. *Comput. Methods Programs Biomed.* **127**, 248–257 (2016).
34. Esteva, A. *et al.* Dermatologist-level classification of skin cancer with deep neural networks. *Nature* **542**, 115–118 (2017).
35. González, G., Washko, G. R. & Estépar, R. S. J. Deep learning for biomarker regression: application to osteoporosis and emphysema on chest CT scans. *Proc. SPIE-- Int. Soc. Opt. Eng.* **10574**, (2018).
36. Cole, J. H. *et al.* Predicting brain age with deep learning from raw imaging data results in a reliable and heritable biomarker. *NeuroImage* **163**, 115–124 (2017).
37. Ding, Y. *et al.* A Deep Learning Model to Predict a Diagnosis of Alzheimer Disease by Using 18F-FDG PET of the Brain. *Radiology* **290**, 456–464 (2019).
38. González, G. *et al.* Disease Staging and Prognosis in Smokers Using Deep Learning in Chest Computed Tomography. *Am. J. Respir. Crit. Care Med.* **197**, 193–203 (2018).
39. Schaumberg, A. J., Rubin, M. A. & Fuchs, T. J. H&E-stained Whole Slide Image Deep Learning Predicts SPOP Mutation State in Prostate Cancer. *bioRxiv* 064279 (2018) doi:10.1101/064279.
40. Coudray, N. *et al.* Classification and mutation prediction from non-small cell lung cancer histopathology images using deep learning. *Nat. Med.* **24**, 1559–1567 (2018).
41. Kather, J. N. *et al.* Deep learning can predict microsatellite instability directly from histology in gastrointestinal cancer. *Nat. Med.* **25**, 1054–1056 (2019).
42. Wei, J. W. *et al.* Automated Detection of Celiac Disease on Duodenal Biopsy Slides: A Deep Learning Approach. *J. Pathol. Inform.* **10**, (2019).
43. Song, Z. *et al.* Clinically applicable histopathological diagnosis system for gastric cancer detection using deep learning. *bioRxiv* 2020.01.30.927749 (2020) doi:10.1101/2020.01.30.927749.
44. Holten-Rossing, H. *et al.* Application of automated image analysis reduces the workload of manual screening of sentinel lymph node biopsies in breast cancer. *Histopathology* **71**, 866–873 (2017).
45. Ehteshami Bejnordi, B. *et al.* Diagnostic Assessment of Deep Learning Algorithms for Detection of Lymph Node Metastases in Women With Breast Cancer. *JAMA* **318**, 2199–2210 (2017).
46. Ehteshami Bejnordi, B. *et al.* Using deep convolutional neural networks to identify and classify tumor-associated stroma in diagnostic breast biopsies. *Mod. Pathol. Off. J. U. S. Can. Acad. Pathol. Inc* **31**, 1502–1512 (2018).
47. Cireşan, D. C., Giusti, A., Gambardella, L. M. & Schmidhuber, J. Mitosis detection in breast cancer histology images with deep neural networks. *Med. Image Comput. Comput.-Assist. Interv. MICCAI Int. Conf. Med. Image Comput. Comput.-Assist. Interv.* **16**, 411–418 (2013).

48. Acs, B. *et al.* An open source automated tumor infiltrating lymphocyte algorithm for prognosis in melanoma. *Nat. Commun.* **10**, 5440 (2019).
49. Wu, M., Yan, C., Liu, H. & Liu, Q. Automatic classification of ovarian cancer types from cytological images using deep convolutional neural networks. *Biosci. Rep.* **38**, (2018).
50. Mahmood, F. *et al.* Deep Adversarial Training for Multi-Organ Nuclei Segmentation in Histopathology Images. *ArXiv181000236 Cs* (2018).
51. Bora, K., Chowdhury, M., Mahanta, L. B., Kundu, M. K. & Das, A. K. Automated classification of Pap smear images to detect cervical dysplasia. *Comput. Methods Programs Biomed.* **138**, 31–47 (2017).
52. Zhang, L. *et al.* DeepPap: Deep Convolutional Networks for Cervical Cell Classification. *IEEE J. Biomed. Health Inform.* **21**, 1633–1643 (2017).
53. Song, Y. *et al.* Accurate Segmentation of Cervical Cytoplasm and Nuclei Based on Multiscale Convolutional Network and Graph Partitioning. *IEEE Trans. Biomed. Eng.* **62**, 2421–2433 (2015).
54. Sanyal, P., Mukherjee, T., Barui, S., Das, A. & Gangopadhyay, P. Artificial Intelligence in Cytopathology: A Neural Network to Identify Papillary Carcinoma on Thyroid Fine-Needle Aspiration Cytology Smears. *J. Pathol. Inform.* **9**, 43 (2018).
55. Muralidaran, C., Dey, P., Nijhawan, R. & Kakkar, N. Artificial neural network in diagnosis of urothelial cell carcinoma in urine cytology. *Diagn. Cytopathol.* **43**, 443–449 (2015).
56. Subbaiah, R. M., Dey, P. & Nijhawan, R. Artificial neural network in breast lesions from fine-needle aspiration cytology smear. *Diagn. Cytopathol.* **42**, 218–224 (2014).
57. Momeni-Boroujeni, A., Yousefi, E. & Somma, J. Computer-assisted cytologic diagnosis in pancreatic FNA: An application of neural networks to image analysis. *Cancer Cytopathol.* **125**, 926–933 (2017).
58. Tosun, A. B., Yergiyev, O., Kolouri, S., Silverman, J. F. & Rohde, G. K. Detection of malignant mesothelioma using nuclear structure of mesothelial cells in effusion cytology specimens. *Cytom. Part J. Int. Soc. Anal. Cytol.* **87**, 326–333 (2015).
59. Al-qudah, R. & Suen, C. Y. Improving blood cells classification in peripheral blood smears using enhanced incremental training. *Comput. Biol. Med.* **131**, 104265 (2021).
60. Mundhra, D., Cheluvaraju, B., Rampure, J. & Rai Dastidar, T. Analyzing Microscopic Images of Peripheral Blood Smear Using Deep Learning. in *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support* (eds. Cardoso, M. J. et al.) 178–185 (Springer International Publishing, 2017). doi:10.1007/978-3-319-67558-9_21.
61. Acevedo, A., Alférez, S., Merino, A., Puigví, L. & Rodellar, J. Recognition of peripheral blood cell images using convolutional neural networks. *Comput. Methods Programs Biomed.* **180**, 105020 (2019).
62. Lin, Y.-C. *et al.* Deep learning for fully automated tumor segmentation and extraction of magnetic resonance radiomics features in cervical cancer. *Eur. Radiol.* **30**, 1297–1305 (2020).
63. Caballo, M., Pangallo, D. R., Mann, R. M. & Sechopoulos, I. Deep learning-based segmentation of breast masses in dedicated breast CT imaging: Radiomic feature stability between radiologists and artificial intelligence. *Comput. Biol. Med.* **118**, 103629 (2020).
64. Yin, P. *et al.* Machine and Deep Learning Based Radiomics Models for Preoperative Prediction of Benign and Malignant Sacral Tumors. *Front. Oncol.* **10**, (2020).
65. Ueno, Y. *et al.* Endometrial Carcinoma: MR Imaging-based Texture Model for Preoperative Risk Stratification-A Preliminary Analysis. *Radiology* **284**, 748–757 (2017).
66. Aerts, H. J. W. L. *et al.* Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. *Nat. Commun.* **5**, 4006 (2014).
67. Colen, R. R., Hassan, I., Elshafeey, N. & Zinn, P. O. Shedding Light on the 2016 World Health Organization Classification of Tumors of the Central Nervous System in the Era of Radiomics and Radiogenomics. *Magn. Reson. Imaging Clin. N. Am.* **24**, 741–749 (2016).
68. Wu, W. *et al.* Exploratory Study to Identify Radiomics Classifiers for Lung Cancer Histology. *Front. Oncol.* **6**, 71 (2016).

69. He, J. *et al.* The practical implementation of artificial intelligence technologies in medicine. *Nat. Med.* **25**, 30–36 (2019).
70. Abramson, R. G. *et al.* Methods and challenges in quantitative imaging biomarker development. *Acad. Radiol.* **22**, 25–32 (2015).
71. Acevedo, A. *et al.* A dataset of microscopic peripheral blood cell images for development of automatic recognition systems. *Data Brief* **30**, 105474 (2020).
72. *supervisely/supervisely*. (Supervisely, 2021).
73. Chollet, F. & others. *Keras*. (2015).
74. Buslaev, A. *et al.* Albumentations: Fast and Flexible Image Augmentations. *Information* **11**, 125 (2020).
75. Tukiainen, M. *ImageDataAugmentor*. *GitHub repository* (GitHub, 2019).
76. Griethuysen, J. J. M. van *et al.* Computational Radiomics System to Decode the Radiographic Phenotype. *Cancer Res.* **77**, e104–e107 (2017).
77. Mohammed, E., Mohamed, M., Far, B. & Naugler, C. Peripheral blood smear image analysis: A comprehensive review. *J. Pathol. Inform.* **5**, 9 (2014).
78. Zwanenburg, A., Leger, S., Vallières, M. & Löck, S. Image biomarker standardisation initiative. *Radiology* **295**, 328–338 (2020).
79. Radiomic Features — pyradiomics v3.0.1.post3+g0c53d1d documentation. https://pyradiomics.readthedocs.io/en/latest/features.html.
80. Allen, M. *et al.* Raincloud plots: a multi-platform tool for robust data visualization. *Wellcome Open Res.* **4**, 63 (2021).
81. Kirstie Whitaker *et al. RainCloudPlots/RainCloudPlots: WellcomeOpenResearch*. (Zenodo, 2019). doi:10.5281/zenodo.3368186.
82. Metsalu, T. & Vilo, J. ClustVis: a web tool for visualizing clustering of multivariate data using Principal Component Analysis and heatmap. *Nucleic Acids Res.* **43**, W566–W570 (2015).
83. Lever, J., Krzywinski, M. & Altman, N. Principal component analysis. *Nat. Methods* **14**, 641–642 (2017).
84. Buggenthin, F. *et al.* Prospective identification of hematopoietic lineage choice by deep learning. *Nat. Methods* **14**, 403–406 (2017).
85. General Data Protection Regulation (GDPR) – Official Legal Text. *General Data Protection Regulation (GDPR)* https://gdpr-info.eu/.
86. Samek, W., Wiegand, T. & Müller, K.-R. Explainable Artificial Intelligence: Understanding, Visualizing and Interpreting Deep Learning Models. *ArXiv170808296 Cs Stat* (2017).
87. Selvaraju, R. R. *et al.* Grad-CAM: Visual Explanations from Deep Networks via Gradient-based Localization. *Int. J. Comput. Vis.* **128**, 336–359 (2020).
88. Mobadersany, P. *et al.* Predicting cancer outcomes from histology and genomics using convolutional networks. *Proc. Natl. Acad. Sci.* **115**, E2970–E2979 (2018).
89. Campanella, G. *et al.* Clinical-grade computational pathology using weakly supervised deep learning on whole slide images. *Nat. Med.* **25**, 1301–1309 (2019).
90. Health, C. for D. and R. Artificial Intelligence and Machine Learning in Software as a Medical Device. *FDA* (2021).
91. *Feature Selection for Data and Pattern Recognition*. (Springer-Verlag, 2015). doi:10.1007/978-3-662-45620-0.
92. Onile-ere, O., Openibo, J. & Olasehinde, G. Malaria Diagnosis: Current Approaches and Future Prospects. in (2016).
93. Molyneux, E., Scanlan, T., Chagaluka, G. & Renner, L. Haematological cancers in African children: progress and challenges. *Br. J. Haematol.* **177**, 971–978 (2017).
94. Marée, R. *et al.* Collaborative analysis of multi-gigapixel imaging data using Cytomine. *Bioinforma. Oxf. Engl.* **32**, 1395–1401 (2016).

Appendices

Supplementary Data

Table 18: Student's Test of morphometric and texture features

		STUDENT TEST									
		Basophile with...				**Eosinophil** with...			**Lymphocyte** with...		**Monocyte** with...
		Eosinophil	Lymphocyte	Monocyte	Neutrophil	Lymphocyte	Monocyte	Neutrophil	Monocyte	Neutrophil	Neutrophil
MORPHOLOGY	Nuclei area	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	Nuclei perimeter	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	Nuclei diameter	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	Nuclei major axis	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	Nuclei minor axis	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	Nuclei perimeter-to-area ratio	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	Cell area	0,00	0,00	0,00	0,00	0,00	0,00	0,64	0,00	0,00	0,00
	Cell perimeter	0,00	0,00	0,00	0,00	0,00	0,00	0,04	0,00	0,00	0,00
	Cell diameter	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	Cell major axis	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	Cell minor axis	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	Cell perimeter-to-area ratio	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	Nuclei-to-cell ratio	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	Cell border irregularity	0,00	0,04	0,00	0,00	0,00	0,00	0,04	0,00	0,00	0,00
	Nuclei border irregularity	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	Nuclei eccentricity	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
First Order	10Percentile	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	90Percentile	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	Energy	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	Entropy	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	InterquartileRange	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	Kurtosis	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	Maximum	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	MeanAbsoluteDeviation	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	Mean	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	Median	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	Minimum	0,00	0,00	0,00	0,00	0,00	0,95	0,63	0,00	0,00	0,57
	Range	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	RobustMeanAbsoluteDeviation	0,00	0,00	0,00	0,00	0,00	0,13	0,00	0,00	0,00	0,00
	RootMeanSquared	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	Skewness	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	TotalEnergy	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	Uniformity	0,00	0,00	0,00	0,00	0,00	0,54	0,00	0,00	0,00	0,00
	Variance	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
GLCM	Autocorrelation	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	JointAverage	0,00	0,00	0,00	0,00	0,18	0,00	0,00	0,00	0,00	0,00
	ClusterProminence	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	ClusterShade	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	ClusterTendency	0,00	0,00	0,00	0,00	0,00	0,37	0,00	0,00	0,00	0,00
	Contrast	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	Correlation	0,05	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	DifferenceAverage	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	DifferenceEntropy	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	DifferenceVariance	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	JointEnergy	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	JointEntropy	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	Imc1	0,00	0,01	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	Imc2	0,15	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	Idm	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	Idmn	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	Id	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	Idn	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	InverseVariance	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	MaximumProbability	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	SumEntropy	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	SumSquares	0,00	0,00	0,00	0,00	0,00	0,01	0,00	0,00	0,00	0,00
GLRLM	GrayLevelNonUniformity	0,00	0,00	0,00	0,00	0,00	0,00	0,03	0,00	0,00	0,00
	GrayLevelNonUniformityNormalized	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	GrayLevelVariance	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	HighGrayLevelRunEmphasis	0,00	0,00	0,00	0,00	0,24	0,00	0,00	0,00	0,00	0,00
	LongRunEmphasis	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	LongRunHighGrayLevelEmphasis	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	LongRunLowGrayLevelEmphasis	0,00	0,00	0,00	0,00	0,22	0,00	0,00	0,00	0,00	0,00
	LowGrayLevelRunEmphasis	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	RunEntropy	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	RunLengthNonUniformity	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	RunLengthNonUniformityNormalized	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	RunPercentage	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	RunVariance	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	ShortRunEmphasis	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	ShortRunHighGrayLevelEmphasis	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	ShortRunLowGrayLevelEmphasis	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
GLSZM	GrayLevelNonUniformity	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	GrayLevelNonUniformityNormalized	0,00	0,00	0,00	0,00	0,00	0,00	0,21	0,00	0,00	0,00
	GrayLevelVariance	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	HighGrayLevelZoneEmphasis	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	LargeAreaEmphasis	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	LargeAreaHighGrayLevelEmphasis	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	LargeAreaLowGrayLevelEmphasis	0,00	0,00	0,00	0,00	0,00	0,00	0,86	0,00	0,00	0,00
	LowGrayLevelZoneEmphasis	0,00	0,00	0,00	0,00	0,00	0,83	0,00	0,00	0,00	0,00
	SizeZoneNonUniformity	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,08	0,00
	SizeZoneNonUniformityNormalized	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	SmallAreaEmphasis	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	SmallAreaHighGrayLevelEmphasis	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	SmallAreaLowGrayLevelEmphasis	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	ZoneEntropy	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	ZonePercentage	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	ZoneVariance	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
GLDM	DependenceEntropy	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	DependenceNonUniformity	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	DependenceNonUniformityNormalized	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	DependenceVariance	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	GrayLevelNonUniformity	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	GrayLevelVariance	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	HighGrayLevelEmphasis	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	LargeDependenceEmphasis	0,00	0,02	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	LargeDependenceHighGrayLevelEmphasis	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	LargeDependenceLowGrayLevelEmphasis	0,00	0,00	0,00	0,00	0,43	0,00	0,00	0,00	0,00	0,00
	LowGrayLevelEmphasis	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	SmallDependenceEmphasis	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	LowGrayLevelEmphasis	0,00	0,00	0,00	0,00	0,00	0,57	0,00	0,00	0,00	0,00
	SmallDependenceEmphasis	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00

***Description**: A conditional formatting was used. Values below 0.05 (statistically significant) are highlighted in green whereas values ≥ 0.05 (statistically non-significant) are highlighted in red.*

Table 19: Normality Interval of morphometric and texture features

		MEAN ± 2SD = "Normality Interval"									
		BASOPHILE		EOSINOPHILE		LYMPHOCYTE		MONOCYTE		NEUTROPHILE	
MORPHOLOGY	Nuclei area	34,91	78,30	34,82	66,50	30,54	69,40	61,02	113,58	25,12	52,79
	Nuclei perimeter	106,24	251,38	124,95	251,63	91,33	158,55	130,02	260,49	127,97	257,13
	Nuclei diameter	9,09	14,05	9,50	16,15	7,37	12,04	10,88	17,16	8,98	13,74
	Nuclei major axis	8,22	13,91	8,81	19,33	6,79	11,62	10,10	17,00	8,72	15,56
	Nuclei minor axis	5,98	10,27	4,75	9,88	5,80	9,13	7,32	11,86	4,95	10,61
	Nuclei perimeter-to-area ratio	1,53	5,22	2,25	5,60	0,21	5,23	1,34	3,37	2,99	7,40
	Cell area	73,67	165,97	95,36	206,66	33,09	127,11	112,09	281,01	95,02	206,54
	Cell perimeter	130,64	255,32	148,55	275,80	96,91	219,60	172,69	334,12	158,20	263,97
	Cell diameter	9,98	19,07	11,72	19,45	7,57	16,62	12,88	25,02	12,04	19,44
	Cell major axis	10,08	17,13	11,47	18,22	7,22	15,22	12,15	23,24	11,61	18,37
	Cell minor axis	9,36	14,41	11,08	16,09	6,76	12,54	11,67	18,39	10,59	16,04
	Cell perimeter-to-area ratio	1,18	2,21	1,11	1,84	1,43	2,82	1,01	1,71	1,12	1,80
	Nuclei-to-cell ratio	0,32	0,64	0,25	0,43	0,42	0,87	0,32	0,59	0,18	0,34
	Cell border irregularity	3,80	6,19	3,92	5,85	3,98	6,08	4,18	6,06	4,14	5,59
	Nuclei border irregularity	4,06	9,47	5,09	9,92	3,88	6,18	4,03	7,81	5,76	11,89
	Nuclei eccentricity	-0,04	0,17	-0,02	0,20	-0,04	0,15	-0,02	0,17	-0,01	0,14
First Order	10Percentile	-11,16	60,77	54,68	109,82	-4,21	119,23	61,58	150,38	104,42	169,93
	90Percentile	108,13	202,57	121,25	199,02	126,95	204,51	142,66	220,28	168,26	212,40
	Energy	-10644590,75	108672113,99	26647969,00	214918772,94	-45893467,41	115182582,87	-5456254,11	403260821,34	99545001,66	402728242,68
	Entropy	2,42	3,25	1,97	2,65	2,08	3,22	1,78	2,81	1,52	2,35
	InterquartileRange	41,41	111,65	17,58	54,00	13,95	101,72	13,66	60,04	11,12	39,01
	Kurtosis	1,02	4,30	2,04	6,53	0,45	5,62	1,59	6,47	2,19	12,84
	Maximum	210,03	254,72	218,25	258,90	195,83	245,80	214,60	256,11	217,71	242,68
	MeanAbsoluteDeviation	28,01	54,80	16,12	32,10	15,41	51,76	12,44	34,59	11,32	24,40
	Mean	39,81	135,12	88,72	145,22	60,14	161,65	107,02	179,61	139,78	188,77
	Median	22,01	145,79	84,28	140,32	48,50	171,67	105,40	181,40	141,20	193,21
	Minimum	-2,73	3,83	-2,11	45,55	-9,73	30,57	-4,68	48,16	-3,74	46,98
	Range	209,12	254,52	188,96	244,75	181,95	238,84	181,98	245,25	183,01	234,13
	RobustMeanAbsoluteDeviation	18,22	44,68	8,24	22,94	7,19	41,32	6,36	25,09	5,07	16,50
	RootMeanSquared	58,04	144,55	93,77	148,96	73,98	164,61	112,13	181,44	142,39	190,20
	Skewness	-0,78	1,45	0,20	1,46	-1,16	1,22	-1,26	0,87	-2,72	-0,09
	TotalEnergy	-10644590,75	108672113,99	26647969,00	214918772,94	-45893467,41	115182582,87	-5456254,11	403260821,34	99545001,66	402728242,68
	Uniformity	0,09	0,23	0,18	0,33	0,08	0,30	0,15	0,36	0,22	0,46
	Variance	1170,76	3850,71	479,95	1598,66	170,96	3436,33	142,47	1813,06	233,81	1086,82
GLCM	Autocorrelation	2,36	36,79	12,84	35,04	3,92	47,33	17,46	53,83	28,24	64,23
	JointAverage	1,91	5,95	3,63	5,85	2,61	6,90	4,26	7,37	5,36	8,07
	ClusterProminence	-14,21	1006,73	-20,90	265,16	-245,99	760,62	-101,73	327,21	-84,88	216,33
	ClusterShade	-38,40	65,76	-3,81	22,03	-39,08	43,63	-14,75	14,89	-17,85	6,26
	ClusterTendency	5,05	22,37	1,66	8,90	-0,35	17,78	-0,04	10,51	0,50	5,65
	Contrast	0,53	2,08	0,30	0,66	-0,51	2,94	0,10	0,68	0,18	0,46
	Correlation	0,74	0,90	0,74	0,91	0,59	0,93	0,75	0,94	0,74	0,87
	DifferenceAverage	0,49	1,05	0,29	0,51	0,06	1,34	0,15	0,53	0,16	0,39
	DifferenceEntropy	1,25	1,92	0,94	1,25	0,79	2,10	0,70	1,28	0,71	1,09
	DifferenceVariance	0,33	1,02	0,22	0,40	0,00	1,21	0,13	0,40	0,16	0,31
	JointEnergy	0,00	0,12	0,08	0,21	-0,04	0,23	0,05	0,28	0,11	0,39
	JointEntropy	3,88	5,50	2,96	4,12	2,81	5,68	2,44	4,33	2,04	3,52
	Imc1	-0,40	-0,20	-0,46	-0,29	-0,49	-0,12	-0,58	-0,31	-0,51	-0,33
	Imc2	0,82	0,96	0,83	0,95	0,76	0,96	0,84	0,98	0,81	0,94
	Idm	0,57	0,77	0,76	0,86	0,48	0,92	0,75	0,92	0,81	0,92
	Idmn	0,98	0,99	0,99	1,00	0,97	1,00	0,99	1,00	0,99	1,00
	Id	0,61	0,77	0,77	0,86	0,53	0,91	0,76	0,92	0,82	0,92
	Idn	0,91	0,95	0,95	0,97	0,88	0,99	0,95	0,98	0,96	0,98
	InverseVariance	0,36	0,51	0,28	0,41	0,24	0,55	0,17	0,43	0,15	0,34
	MaximumProbability	0,01	0,30	0,16	0,40	-0,02	0,42	0,12	0,47	0,22	0,60
	SumEntropy	3,11	4,17	2,56	3,47	2,55	4,09	2,25	3,66	1,85	3,06
	SumSquares	1,48	6,03	0,52	2,36	-0,08	5,04	0,05	2,77	0,18	1,51
GLRLM	GrayLevelNonUniformity	145,33	726,46	385,14	1093,78	-27,56	434,67	305,36	1123,11	413,13	1078,47
	GrayLevelNonUniformityNormalized	0,11	0,19	0,17	0,24	0,11	0,22	0,15	0,26	0,19	0,28
	GrayLevelVariance	2,37	5,67	1,54	3,07	1,09	5,48	1,07	3,27	1,20	2,81
	HighGrayLevelRunEmphasis	8,19	35,21	16,50	38,04	11,83	42,46	19,57	50,16	28,34	59,74
	LongRunEmphasis	1,06	6,24	4,98	14,12	-7,62	18,10	-1,58	28,96	3,96	35,92
	LongRunHighGrayLevelEmphasis	-45,65	193,23	69,12	384,21	-480,79	838,56	-217,52	1292,20	19,58	1913,56
	LongRunLowGrayLevelEmphasis	-0,73	3,16	0,12	0,94	-5,77	6,90	0,02	0,85	-0,03	0,95
	LowGrayLevelRunEmphasis	0,03	0,36	0,02	0,10	-0,01	0,22	0,01	0,08	0,02	0,05
	RunEntropy	3,81	4,52	4,31	4,90	3,14	4,77	4,16	5,25	4,37	5,06
	RunLengthNonUniformity	666,93	2575,82	541,27	1966,65	318,17	953,77	591,63	1695,60	524,03	1404,86
	RunLengthNonUniformityNormalized	0,43	0,69	0,27	0,42	0,30	0,91	0,20	0,47	0,23	0,38
	RunPercentage	0,54	0,79	0,37	0,54	0,37	0,99	0,26	0,59	0,26	0,46
	RunVariance	-0,26	2,91	1,63	7,43	-6,94	12,06	-3,37	18,14	-0,08	23,57
	ShortRunEmphasis	0,69	0,86	0,53	0,67	0,60	0,99	0,46	0,72	0,49	0,64
	ShortRunHighGrayLevelEmphasis	7,88	26,27	10,64	24,61	12,17	30,57	12,14	27,90	15,36	31,90
	ShortRunLowGrayLevelEmphasis	0,04	0,22	0,01	0,06	-0,01	0,17	0,00	0,06	0,01	0,04
GLSZM	GrayLevelNonUniformity	65,32	246,43	69,69	169,25	36,76	99,36	68,64	153,18	62,03	134,37
	GrayLevelNonUniformityNormalized	0,12	0,16	0,13	0,16	0,12	0,16	0,13	0,17	0,13	0,17
	GrayLevelVariance	3,16	5,08	2,91	4,43	2,56	5,13	2,72	4,38	2,94	4,63
	HighGrayLevelZoneEmphasis	15,44	31,83	21,65	43,04	18,40	37,50	20,88	44,74	24,23	47,24
	LargeAreaEmphasis	-1026,27	1468,42	-2462,93	15051,41	-7093,32	9471,76	-14553,55	33966,61	-12363,86	43400,96
	LargeAreaHighGrayLevelEmphasis	-21273,76	30026,02	-65656,49	326120,96	-437854,24	538573,00	-639221,26	1381719,75	-717692,54	2222921,84
	LargeAreaLowGrayLevelEmphasis	-290,36	453,96	-200,96	880,55	-2423,43	2676,95	-403,77	965,91	-336,50	1017,76
	LowGrayLevelZoneEmphasis	0,08	0,20	0,02	0,11	0,03	0,16	0,02	0,11	0,02	0,10
	SizeZoneNonUniformity	158,02	551,30	123,56	319,30	126,92	253,17	114,54	286,23	116,23	262,57
	SizeZoneNonUniformityNormalized	0,22	0,43	0,18	0,36	0,24	0,58	0,20	0,34	0,21	0,36
	SmallAreaEmphasis	0,48	0,69	0,44	0,63	0,52	0,80	0,46	0,61	0,47	0,62
	SmallAreaHighGrayLevelEmphasis	9,80	19,49	11,00	25,95	12,59	25,02	10,39	23,60	10,92	23,84
	SmallAreaLowGrayLevelEmphasis	0,04	0,10	0,01	0,06	0,02	0,10	0,01	0,07	0,01	0,07
	ZoneEntropy	4,79	5,89	5,08	6,14	4,06	5,52	5,07	6,02	4,99	5,87
	ZonePercentage	0,13	0,40	0,06	0,15	0,00	0,69	0,01	0,18	0,04	0,12
	ZoneVariance	-1015,55	1420,84	-2510,44	14897,73	-7049,42	9380,70	-14536,32	33633,36	-12436,43	43093,14
GLDM	DependenceEntropy	5,09	5,90	4,87	5,64	4,39	5,78	4,48	5,85	4,01	5,31
	DependenceNonUniformity	250,25	1118,32	436,91	1524,72	-87,37	726,12	111,24	2250,20	426,56	2257,85
	DependenceNonUniformityNormalized	0,11	0,20	0,11	0,13	0,03	0,36	0,10	0,17	0,10	0,20
	DependenceVariance	2,49	5,82	4,35	6,16	0,07	8,06	4,32	6,55	4,90	6,87
	GrayLevelNonUniformity	89,62	1359,59	970,59	3082,59	-499,96	1404,34	241,32	4265,33	1219,45	4910,25
	GrayLevelVariance	1,96	6,27	0,85	2,64	0,35	5,60	0,30	2,99	0,45	1,82
	HighGrayLevelEmphasis	4,81	37,47	14,27	36,48	8,23	46,18	18,64	53,36	28,78	63,70
	LargeDependenceEmphasis	8,91	26,61	26,48	41,39	-4,65	40,93	22,42	52,15	33,40	53,10
	LargeDependenceHighGrayLevelEmphasis	-159,09	857,05	400,10	1192,04	-646,94	1756,81	356,37	2460,03	1050,83	3092,40
	LargeDependenceLowGrayLevelEmphasis	-2,78	16,06	0,63	3,20	-4,37	8,15	0,33	2,21	0,38	1,64
	LowGrayLevelEmphasis	0,00	0,48	0,02	0,10	-0,05	0,27	0,00	0,08	0,01	0,04
	SmallDependenceEmphasis	0,12	0,35	0,07	0,15	0,01	0,60	0,03	0,17	0,05	0,12
	LowGrayLevelEmphasis	1,79	9,31	1,61	5,14	0,26	16,48	1,04	5,68	1,76	4,75
	SmallDependenceEmphasis	0,01	0,06	0,00	0,01	-0,01	0,08	0,00	0,01	0,00	0,01

Description: *Normality Interval is defined as the Interval between Mean-2SD and Mean+2SD*

Table 20: Confidence Interval of morphometric and texture features

		MEAN					Confidence Interval				
MORPHOLOGY	Nuclei area	56,61	50,66	49,97	87,30	38,96	0,586	0,248	0,179	0,378	0,079
	Nuclei perimeter	178,81	188,29	124,94	195,26	192,55	1,959	0,993	0,309	0,937	0,367
	Nuclei diameter	11,57	12,83	9,71	14,02	11,36	0,067	0,052	0,022	0,045	0,014
	Nuclei major axis	11,06	14,07	9,21	13,55	12,14	0,077	0,082	0,022	0,050	0,019
	Nuclei minor axis	8,12	7,31	7,46	9,59	7,78	0,058	0,040	0,015	0,033	0,016
	Nuclei perimeter-to-area ratio	3,37	3,92	2,72	2,35	5,19	0,050	0,026	0,023	0,015	0,013
	Cell area	119,82	151,01	80,10	196,55	150,78	1,246	0,872	0,433	1,213	0,317
	Cell perimeter	192,98	212,18	158,25	253,41	211,08	1,683	0,997	0,564	1,160	0,301
	Cell diameter	14,53	15,59	12,09	18,95	15,74	0,123	0,061	0,042	0,087	0,021
	Cell major axis	13,61	14,84	11,22	17,70	14,99	0,095	0,053	0,037	0,080	0,019
	Cell minor axis	11,89	13,59	9,65	15,03	13,31	0,068	0,039	0,027	0,048	0,016
	Cell perimeter-to-area ratio	1,70	1,48	2,12	1,36	1,46	0,014	0,006	0,006	0,005	0,002
	Nuclei-to-cell ratio	0,48	0,34	0,65	0,45	0,26	0,004	0,001	0,002	0,002	0,000
	Cell border irregularity	4,99	4,88	5,03	5,12	4,87	0,032	0,015	0,010	0,013	0,004
	Nuclei border irregularity	6,76	7,51	5,03	5,92	8,83	0,073	0,038	0,011	0,027	0,017
	Nuclei eccentricity	0,066	0,093	0,059	0,073	0,061	0,003	0,002	0,001	0,001	0,000
First Order	10Percentile	24,80	82,25	57,51	105,98	137,17	0,971	0,432	0,568	0,638	0,186
	90Percentile	155,35	160,14	165,73	181,47	190,33	1,275	0,609	0,357	0,558	0,125
	Energy	49013761,62	120783370,97	34644557,73	198902283,62	251136622,17	1610649,665	1475387,427	741109,374	2935738,260	861905,513
	Entropy	2,84	2,31	2,65	2,30	1,94	0,011	0,005	0,005	0,007	0,002
	InterquartileRange	76,53	35,79	57,83	36,85	25,07	0,948	0,285	0,404	0,333	0,079
	Kurtosis	2,66	4,28	3,03	4,03	7,51	0,044	0,035	0,024	0,035	0,030
	Maximum	232,37	238,57	220,81	235,36	230,19	0,603	0,319	0,230	0,298	0,071
	MeanAbsoluteDeviation	41,40	24,11	33,58	23,51	17,86	0,362	0,125	0,167	0,159	0,037
	Mean	87,47	116,97	110,89	143,31	164,28	1,287	0,443	0,467	0,521	0,139
	Median	83,90	112,30	110,09	143,40	167,21	1,671	0,439	0,567	0,546	0,148
	Minimum	0,55	21,72	10,42	21,74	21,62	0,089	0,374	0,185	0,379	0,144
	Range	231,82	216,85	210,40	213,62	208,57	0,613	0,437	0,262	0,454	0,145
	RobustMeanAbsoluteDeviation	31,45	15,59	24,25	15,73	10,79	0,357	0,115	0,157	0,134	0,032
	RootMeanSquared	101,30	121,37	119,30	146,78	166,30	1,168	0,432	0,417	0,498	0,136
	Skewness	0,34	0,83	0,03	-0,20	-1,40	0,030	0,010	0,011	0,015	0,007
	TotalEnergy	49013761,62	120783370,97	34644557,73	198902283,62	251136622,17	1610649,665	1475387,427	741109,374	2935738,260	861905,513
	Uniformity	0,16	0,26	0,19	0,26	0,34	0,002	0,001	0,001	0,002	0,001
	Variance	2510,74	1039,30	1803,64	977,77	660,31	36,177	8,767	15,024	12,000	2,425
GLCM	Autocorrelation	19,57	23,94	25,62	35,65	46,24	0,465	0,174	0,200	0,261	0,102
	JointAverage	3,93	4,74	4,75	5,82	6,71	0,055	0,017	0,020	0,022	0,008
	ClusterProminence	496,26	122,13	257,32	112,74	65,72	13,782	2,242	4,631	3,081	0,856
	ClusterShade	13,68	9,11	2,28	0,07	-5,80	1,406	0,203	0,381	0,213	0,069
	ClusterTendency	13,71	5,28	8,71	5,24	3,07	0,234	0,057	0,083	0,076	0,015
	Contrast	1,31	0,48	1,22	0,39	0,32	0,021	0,003	0,016	0,004	0,001
	Correlation	0,82	0,82	0,76	0,85	0,81	0,002	0,001	0,002	0,001	0,000
	DifferenceAverage	0,77	0,40	0,70	0,34	0,27	0,008	0,002	0,006	0,003	0,001
	DifferenceEntropy	1,58	1,09	1,45	0,99	0,90	0,009	0,002	0,006	0,004	0,001
	DifferenceVariance	0,68	0,31	0,60	0,26	0,24	0,009	0,001	0,006	0,002	0,000
	JointEnergy	0,06	0,14	0,10	0,16	0,25	0,002	0,001	0,001	0,002	0,001
	JointEntropy	4,69	3,54	4,24	3,38	2,78	0,022	0,009	0,013	0,014	0,004
	Imc1	-0,30	-0,38	-0,31	-0,44	-0,42	0,003	0,001	0,002	0,002	0,000
	Imc2	0,89	0,89	0,86	0,91	0,87	0,002	0,001	0,001	0,001	0,000
	Idm	0,67	0,81	0,70	0,84	0,87	0,003	0,001	0,002	0,001	0,000
	Idmn	0,99	1,00	0,99	1,00	1,00	0,000	0,000	0,000	0,000	0,000
	Id	0,69	0,81	0,72	0,84	0,87	0,002	0,001	0,002	0,001	0,000
	Idn	0,93	0,96	0,94	0,97	0,97	0,001	0,000	0,001	0,000	0,000
	InverseVariance	0,44	0,34	0,39	0,30	0,24	0,002	0,001	0,001	0,002	0,001
	MaximumProbability	0,16	0,28	0,20	0,29	0,41	0,004	0,002	0,002	0,003	0,001
	SumEntropy	3,64	3,01	3,32	2,95	2,45	0,014	0,007	0,007	0,010	0,003
	SumSquares	3,75	1,44	2,48	1,41	0,85	0,061	0,014	0,024	0,020	0,004
GLRLM	GrayLevelNonUniformity	435,89	739,46	203,56	714,24	745,80	7,845	5,553	2,127	5,874	1,891
	GrayLevelNonUniformityNormalized	0,15	0,20	0,17	0,20	0,23	0,001	0,001	0,001	0,001	0,000
	GrayLevelVariance	4,02	2,30	3,29	2,17	2,01	0,045	0,012	0,020	0,016	0,005
	HighGrayLevelRunEmphasis	21,70	27,27	27,14	34,87	44,04	0,365	0,169	0,141	0,220	0,089
	LongRunEmphasis	3,65	9,55	5,24	13,69	19,94	0,070	0,072	0,118	0,219	0,091
	LongRunHighGrayLevelEmphasis	73,79	226,67	178,89	537,34	966,57	3,225	2,469	6,070	10,844	5,384
	LongRunLowGrayLevelEmphasis	1,21	0,53	0,57	0,44	0,46	0,053	0,006	0,058	0,006	0,003
	LowGrayLevelRunEmphasis	0,20	0,06	0,10	0,05	0,03	0,004	0,001	0,001	0,001	0,000
	RunEntropy	4,16	4,60	3,96	4,70	4,72	0,009	0,005	0,007	0,008	0,002
	RunLengthNonUniformity	1621,38	1253,96	635,97	1143,62	964,44	25,768	11,170	2,924	7,930	2,504
	RunLengthNonUniformityNormalized	0,56	0,34	0,60	0,33	0,30	0,003	0,001	0,003	0,002	0,000
	RunPercentage	0,67	0,46	0,68	0,42	0,36	0,003	0,001	0,003	0,002	0,001
	RunVariance	1,33	4,53	2,56	7,38	11,75	0,043	0,045	0,087	0,155	0,067
	ShortRunEmphasis	0,77	0,60	0,79	0,59	0,56	0,002	0,001	0,002	0,002	0,000
	ShortRunHighGrayLevelEmphasis	17,08	17,62	21,37	20,02	23,63	0,248	0,109	0,085	0,113	0,047
	ShortRunLowGrayLevelEmphasis	0,13	0,04	0,08	0,03	0,02	0,003	0,000	0,001	0,000	0,000
GLSZM	GrayLevelNonUniformity	155,87	119,47	68,06	110,91	98,20	2,445	0,780	0,288	0,607	0,206
	GrayLevelNonUniformityNormalized	0,14	0,15	0,14	0,15	0,15	0,001	0,000	0,000	0,000	0,000
	GrayLevelVariance	4,12	3,67	3,84	3,55	3,79	0,026	0,012	0,012	0,012	0,005
	HighGrayLevelZoneEmphasis	23,63	32,35	27,95	32,81	35,73	0,221	0,168	0,088	0,171	0,065
	LargeAreaEmphasis	221,07	6294,24	1189,22	9706,53	15518,55	33,676	137,251	76,216	348,511	158,531
	LargeAreaHighGrayLevelEmphasis	4376,13	130232,23	50359,38	371249,25	752614,65	692,493	3070,171	4492,532	14516,041	8359,736
	LargeAreaLowGrayLevelEmphasis	81,80	339,79	126,76	281,07	340,63	10,048	8,475	23,467	9,838	3,850
	LowGrayLevelZoneEmphasis	0,14	0,06	0,10	0,06	0,06	0,002	0,001	0,001	0,001	0,000
	SizeZoneNonUniformity	354,66	221,43	190,05	200,38	189,40	5,309	1,534	0,581	1,233	0,416
	SizeZoneNonUniformityNormalized	0,32	0,27	0,41	0,27	0,28	0,003	0,001	0,002	0,001	0,000
	SmallAreaEmphasis	0,59	0,54	0,66	0,53	0,55	0,003	0,002	0,001	0,001	0,000
	SmallAreaHighGrayLevelEmphasis	14,65	18,48	18,80	16,99	17,38	0,131	0,117	0,057	0,095	0,037
	SmallAreaLowGrayLevelEmphasis	0,07	0,03	0,06	0,04	0,04	0,001	0,000	0,000	0,000	0,000
	ZoneEntropy	5,34	5,61	4,79	5,55	5,43	0,015	0,008	0,007	0,007	0,003
	ZonePercentage	0,26	0,11	0,34	0,10	0,08	0,004	0,001	0,003	0,001	0,000
	ZoneVariance	202,64	6193,64	1165,64	9548,52	15328,35	32,889	136,419	75,595	345,994	157,862
GLDM	DependenceEntropy	5,49	5,25	5,09	5,17	4,66	0,011	0,006	0,006	0,010	0,004
	DependenceNonUniformity	684,29	980,82	319,38	1180,72	1342,21	11,718	8,525	3,743	15,364	5,206
	DependenceNonUniformityNormalized	0,16	0,12	0,19	0,13	0,15	0,001	0,000	0,002	0,001	0,000
	DependenceVariance	4,16	5,25	4,07	5,43	5,88	0,045	0,014	0,037	0,016	0,006
	GrayLevelNonUniformity	724,60	2026,59	452,19	2253,32	3064,85	17,143	16,551	8,762	28,904	10,492
	GrayLevelVariance	4,11	1,75	2,97	1,65	1,14	0,058	0,014	0,024	0,019	0,004
	HighGrayLevelEmphasis	21,14	25,38	27,21	36,00	46,24	0,441	0,174	0,175	0,249	0,099
	LargeDependenceEmphasis	17,76	33,94	18,14	37,28	43,25	0,239	0,117	0,210	0,213	0,056
	LargeDependenceHighGrayLevelEmphasis	348,98	796,07	554,93	1408,20	2071,62	13,717	6,206	11,060	15,110	5,804
	LargeDependenceLowGrayLevelEmphasis	6,64	1,92	1,89	1,27	1,01	0,254	0,020	0,058	0,013	0,004
	LowGrayLevelEmphasis	0,24	0,06	0,11	0,04	0,03	0,006	0,001	0,002	0,001	0,000
	SmallDependenceEmphasis	0,23	0,11	0,30	0,10	0,08	0,003	0,001	0,003	0,001	0,000
	LowGrayLevelEmphasis	5,55	3,38	8,37	3,36	3,26	0,10	0,03	0,07	0,03	0,01
	SmallDependenceEmphasis	0,03	0,01	0,03	0,01	0,00	0,00	0,00	0,00	0,00	0,00

***Description**: An heatmap was applied for each row (of the mean) for better visualization. The green cell stands for the largest value, and the red cell for the smallest value.*

Supplementary Methods

Training Details

The optimization method **Adam** was used in all training, and all networks have been trained for 120 **iterations**. The detailed data augmentation from Albumentations were: **Blur** (blur_limit=(3, 5), p=0.3), **Downscale**(scale_min=0.7, scale_max=0.9, p=0.3), **GridDistortion**(num_steps=2, distort_limit=(-0.2, 0.2), value=(0, 0, 0), p=0.3), **Flip**(p=0.5), **Transpose**(p=0.5)

For the classification ANN we used a **learning rate** $\alpha = 0.01$, **batch size** of 32, a split between **training and validation** of 80%/20%, and **image size** was (256, 256,3). The **loss function** was "categorical cross entropy" and we apply L1-L2 **regularization** (l1 = 1000, l2 = 1000)

For the segmentation ANN we used a **learning rate** $\alpha = 0.001$, **batch size** of 4, a split between **training and validation** of 90%/10%, and **image size** was (352, 352,3). The **loss function** was "Dice loss" and we did not apply regularization.

Software Used

All code was written in **Python 3.7.9** and used the following two frameworks: **Tensorflow** (1.14.0) an end-to-end open-source platform for machine learning, and **Keras** (2.3.1, a high-level neural networks API which as built-in GPU support. Moreover, we used several libraries, including: **pyradiomics** 3.0.1**, sci-kit image** 0.18.1, **scipy** 1.6.0, **matplotlib** 3.3.3, **segmentation-models** 1.0.1, **albumentations** 0.5.2, **ImageDataAugmentor**, **pandas** 1.2.1, **SimpleITK** 2.0.2, **six** 1.15.0, and **numpy** 1.19.5.

The data management and data analysis were done with **Microsoft Excel** for Mac 2021 (16.46), and two web-tools: **raincloud-shiny** for the boxplot's graphs, and **clustVis** for the PCA plots and loadings.

The neural networks training was done on 4 remote GPUs from the MIRO laboratory: 1 GTX Titan, 1 GTX Titan XP and 2 GTX Titan X.

Radiomic Features Definitions

First Order statistics (19 features): They describe the distribution of voxel intensities within the ROI.

- **Energy and TotalEnergy:** It measures the magnitude of voxel values in an image and consequently, of the pixel randomness.
- **Entropy**: It measures the average amount of information required to encode the image values,and is also a randomness measurement.
- **Minimum**: The smallest gray level intensity within the ROI.
- **10th percentile**: The 10th percentile gray level intensity within the ROI.
- **90th percentile**: The 90th percentile gray level intensity within the ROI.
- **Maximum**: The maximum gray level intensity within the ROI.
- **Mean**: The Average gray level intensity within the ROI.
- **Median**: The median gray level intensity within the ROI.
- **Interquartile range**: The difference between the 25th and 75th percentile of the image array.
- **Range**: The range of gray values in the ROI.
- **Root Mean Squared (RMS)**: It measures the magnitude of the image values.
- **Standard Deviation, Mean Absolute Deviation (MAD) and Robust Mean Absolute Deviation (rMAD)**: It measures the amount of variation or dispersion from the Mean Value.
- **Skewness**: It measures the asymmetry of the distribution of values about the Mean value.
- **Kurtosis**: It is a measure of the 'peakedness' of the distribution of values in the image ROI. A higher kurtosis implies that the mass of the distribution is concentrated towards the tail(s) rather than towards the mean. A lower kurtosis implies the reverse: that the mass of the distribution is concentrated towards a spike near the Mean value.
- **Variance**: It measures the spread of the distribution about the mean.

- **Uniformity**: It measures the homogeneity of the image array, where a greater uniformity implies a greater homogeneity or a smaller range of discrete intensity values.

GLCM (24): Gray Level Co-occurrence Matrix, describe the texture of an image by calculating how often pairs of pixels with specific values and in a specified spatial relationship occur in an image. Put differently, it measures how often different combinations of pixel brightness values (grey levels) occur in an image.

- **Autocorrelation**: Measure of the magnitude of the fineness and coarseness of texture.
- **Joint Average**: Mean gray level intensity of a given distribution.
- **Cluster Prominence**: Measure of the skewness and asymmetry whereas a higher value implies more asymmetry about the mean while a lower value indicates a peak near the mean value and less variation about the mean.
- **Cluster Shade**: Measure of the skewness and uniformity, whereas a a higher cluster shade implies greater asymmetry about the mean.
- **Cluster Tendency**: Measure of groupings of voxels with similar gray-level values.
- **Contrast**: Measure of the local intensity variation. A larger value correlates with a greater disparity in intensity values among neighboring voxels.
- **Correlation**: Value between 0 (uncorrelated) and 1 (perfectly correlated) showing the linear dependency of gray level values to their respective pixels.
- **Difference Average**: Measure of the relationship between occurrences of pairs with similar intensity values and occurrences of pairs with differing intensity values.
- **Difference Entropy**: Measure of the randomness/variability in neighbourhood intensity value differences.
- **Difference Variance**: Measure of heterogeneity that places higher weights on differing intensity level pairs that deviate more from the mean.
- **Energy**: Measure of homogeneous patterns in the image. A greater Energy implies that there are more instances of intensity value pairs in the image that neighbour each other at higher frequencies.
- **Joint Energy (or Angular Second Moment)**: Measure of homogeneous patterns in the image. A greater Energy implies that there are more instances of intensity value pairs in the image that neighbor each other at higher frequencies.
- **Joint entropy**: Measure of the randomness/variability in neighbourhood intensity values.
- **Informational Measure of Correlation (IMC) 1 and 2**: Measures the complexity of the texture
- **Inverse Difference Moment (IDM), Inverse Difference Moment Normalized (IDMN), Inverse Difference (ID), and Inverse Difference Normalized (IDN)**: Measures of the local homogeneity of an image, where higher values correspond to more uniform gray levels.
- **Maximal Correlation Coefficient (MCC)**: Measure of complexity of the texture.
- Inverse Variance
- **Maximum Probability (or Joint Maximum)**: Most predominant pair of neighboring intensity values.
- **Sum Average**: Measure of the relationship between occurrences of pairs with lower intensity values and occurrences of pairs with higher intensity values.
- **Sum Entropy**: Sum of neighborhood intensity value differences.
- **Sum of Squares (or Joint Variance)**: Measure in the distribution of neighbouring intensity level pairs about the mean intensity level.

GLSZM (16): It quantifies gray level zones in an image. A gray level zone is defined as the number of connected voxels that share the same gray level intensity.

- **Small Area Emphasis (SAE)**: Measure of the distribution of small size zones, with a greater value indicative of more smaller size zones and more fine textures.

- **Large Area Emphasis (LAE):** measure of the distribution of large area size zones, with a greater value indicative of more larger size zones and more coarse textures.
- **Gray Level Non-Uniformity (GLN):** measures the variability of gray-level intensity values in the image, with a lower value indicating more homogeneity in intensity values.
- **Gray Level Non-Uniformity Normalized (GLNN): M**easures the variability of gray-level intensity values in the image, with a lower value indicating a greater similarity in intensity values. This is the normalized version of the GLN formula.
- **Size-Zone Non-Uniformity (SZN):** Measures the variability of size zone volumes in the image, with a lower value indicating more homogeneity in size zone volumes.
- **Size-Zone Non-Uniformity Normalized (SZNN):** measures the variability of size zone volumes throughout the image, with a lower value indicating more homogeneity among zone size volumes in the image. This is the normalized version of the SZN formula.
- **Zone Percentage (ZP):** measures the coarseness of the texture by taking the ratio of number of zones and number of voxels in the ROI.
- **Gray Level Variance (GLV):** GLV measures the variance in gray level intensities for the zones.
- **Zone Variance (ZV):** Measures the variance in zone size volumes for the zones.
- **Zone Entropy (ZE):** Measures the uncertainty/randomness in the distribution of zone sizes and gray levels. A higher value indicates more heterogeneity in the texture patterns.
- **Low Gray Level Zone Emphasis (LGLZE):** measures the distribution of lower gray-level size zones, with a higher value indicating a greater proportion of lower gray-level values and size zones in the image.
- **High Gray Level Zone Emphasis (HGLZE):** Measures the distribution of the higher gray-level values, with a higher value indicating a greater proportion of higher gray-level values and size zones in the image.
- **Small Area Low Gray Level Emphasis (SALGLE):** Measures the proportion in the image of the joint distribution of smaller size zones with lower gray-level values.
- **Small Area High Gray Level Emphasis (SAHGLE):** Measures the proportion in the image of the joint distribution of smaller size zones with higher gray-level values.
- **Large Area Low Gray Level Emphasis (LALGLE):** Measures the proportion in the image of the joint distribution of larger size zones with lower gray-level values.
- **Large Area High Gray Level Emphasis (LAHGLE): M**easures the proportion in the image of the joint distribution of larger size zones with higher gray-level values.

GLRLM (16): It quantifies gray level runs, which are defined as the length in number of pixels, of consecutive pixels that have the same gray level value.

- **Short Run Emphasis (SRE):** Measure of the distribution of short run lengths, with a greater value indicative of shorter run lengths and more fine textural texture
- **Long Run Emphasis (LRE)**: Measure of the distribution of long run lengths, with a greater value indicative of longer run lengths and more coarse structural textures.
- **Gray Level Non-Uniformity (GLN):** Measures the similarity of gray-level intensity values in the image, where a lower GLN value correlates with a greater similarity in intensity values.
- **Gray Level Non-Uniformity Normalized (GLNN)**: Measures the similarity of gray-level intensity values in the image, where a lower GLNN value correlates with a greater similarity in intensity values. This is the normalized version of the GLN formula.
- **Run Length Non-Uniformity (RLN)**: Measures the similarity of run lengths throughout the image, with a lower value indicating more homogeneity among run lengths in the image.
- **Run Length Non-Uniformity Normalized (RLNN)**: Measures the similarity of run lengths throughout the image, with a lower value indicating more homogeneity among run lengths in the image. This is the normalized version of the RLN formula.
- **Run Percentage (RP)**: Measures the coarseness of the texture by taking the ratio of number of runs and number of voxels in the ROI.

- **Gray Level Variance (GLV)** : Measures the variance in gray level intensity for the runs.
- **Run Variance (RV)**: Measure of the variance in runs for the run lengths.
- **Run Entropy (RE)** : Measures the uncertainty/randomness in the distribution of run lengths and gray levels. A higher value indicates more heterogeneity in the texture patterns.
- **Low Gray Level Run Emphasis (LGLRE)**: Measures the distribution of low gray-level values, with a higher value indicating a greater concentration of low gray-level values in the image.
- **High Gray Level Run Emphasis (HGLRE)**: Measures the distribution of the higher gray-level values, with a higher value indicating a greater concentration of high gray-level values in the image.
- **Short Run Low Gray Level Emphasis (SRLGLE):** Measures the joint distribution of shorter run lengths with lower gray-level values.
- **Short Run High Gray Level Emphasis (SRHGLE)**: Measures the joint distribution of shorter run lengths with higher gray-level values.
- **Long Run Low Gray Level Emphasis (LRLGLE)**: Measures the joint distribution of long run lengths with lower gray-level values.
- **Long Run High Gray Level Emphasis (LRHGLE)**: Measures the joint distribution of long run lengths with higher gray-level values.

GLDM (12): It quantifies gray level dependencies in an image. A gray level dependency is defined as the number of connected voxels within distance δ that are dependent on the center voxel.

- **Small Dependence Emphasis (SDE)**: Measure of the distribution of small dependencies, with a greater value indicative of smaller dependence and less homogeneous textures.
- **Large Dependence Emphasis (LDE)**: Measure of the distribution of large dependencies, with a greater value indicative of larger dependence and more homogeneous textures.
- **Gray Level Non-Uniformity (GLN)**: Measures the similarity of gray-level intensity values in the image, where a lower GLN value correlates with a greater similarity in intensity values.
- **Dependence Non-Uniformity (DN)**: Measures the similarity of dependence throughout the image, with a lower value indicating more homogeneity among dependencies in the image.
- **Dependence Non-Uniformity Normalized (DNN)**: Measures the similarity of dependence throughout the image, with a lower value indicating more homogeneity among dependencies in the image. This is the normalized version of the DLN formula.
- **Gray Level Variance (GLV)**: Measures the variance in grey level in the image.
- **Dependence Variance (DV)**: Measures the variance in dependence size in the image.
- **Dependence Entropy (DE)**
- **Low Gray Level Emphasis (LGLE)**: Measures the distribution of low gray-level values, with a higher value indicating a greater concentration of low gray-level values in the image.
- **High Gray Level Emphasis (HGLE)**: Measures the distribution of the higher gray-level values, with a higher value indicating a greater concentration of high gray-level values in the image.
- **Small Dependence Low Gray Level Emphasis (SDLGLE)**: Measures the joint distribution of small dependence with lower gray-level values.
- **Small Dependence High Gray Level Emphasis (SDHGLE)**: Measures the joint distribution of small dependence with higher gray-level values.
- **Large Dependence Low Gray Level Emphasis (LDLGLE)**: Measures the joint distribution of large dependence with lower gray-level values.
- **Large Dependence High Gray Level Emphasis (LDHGLE):** Measures the joint distribution of large dependence with higher gray-level values.

More detailed information, including formulas, can be found in the paper "Imaging Biomarker Standardization Initiative (IBSI)" [78], and at https://pyradiomics.readthedocs.io/en/latest/features.html.