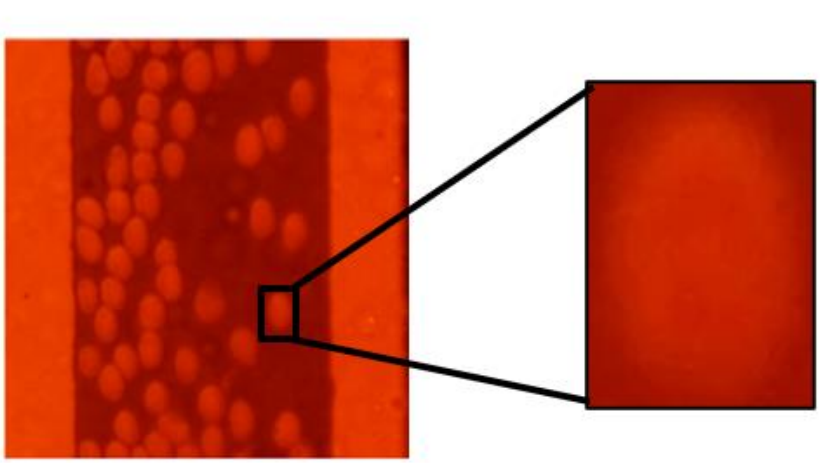
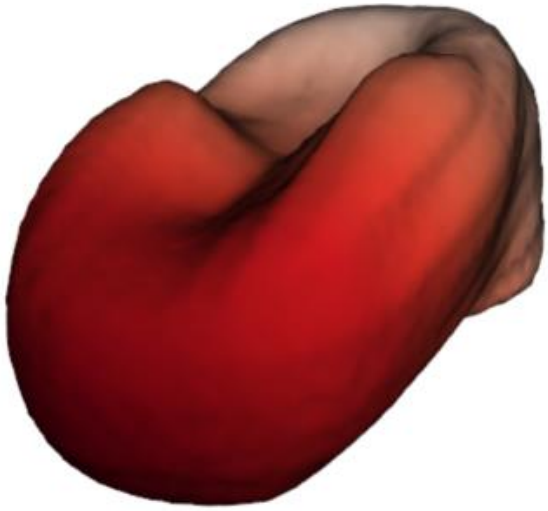




Bachelor Thesis

3D Blood Cell Prediction by 2D Microscopic Images

Degree:	B.Sc. Industrial Engineering
Author:	Dominique Peytrignet
Supervisor:	Prof. Dr. Cédric Bessire
Expert:	Matt Stark
Date:	28.06.2023

Bern University of Applied Sciences

Technik und Informatik
Wirtschaftsingenieurswesen

Abstract

The following Bachelor Thesis covers the topic on how the shape of an object can be estimated based on 2D information, i.e. images, which is a difficult problem for machines. This work treats this problem in the specific case of reconstructing and estimating the shape and volume of red blood cells, based on single microscopic images. Therefore the goal of this thesis was to present a method on how this task can be fulfilled and evaluate its accuracy. The SHAPR-Framework provides a open-source package which can generate 3D cell predictions based on single cell images. Over 10'000 microscopic images, containing single red blood cells from a patient, have been used to make predictions. The results show that these predictions create realistic cell shapes and accurate volume estimations, which align with the true values from the blood sample. It is recommended to further evaluate the accuracy of the results with further blood samples from additional patients. With the presented SHAPR-Framework, the diagnosis of home patients can be further enhanced.

Table of contents

1 Introduction

1.1 Initial Situation and relevance

The human blood, more specifically the red blood cells, plays an important role for the human health as it supplies all the organs with enough oxygen. Analysing the blood and its red blood cells can therefore yield important insights to the current health situation of a person. To analyse the red blood cell and measure important metrics such as shape, volume, haemoglobin concentration, there exists costly equipment which are used in the field of haematology, the study of blood, namely haematology analysers or confocal microscopes, which are further described in chapter 2.

To reduce the costs of utilizing such equipments, a cost- and time-saving automized alternative would have to be developed, which can analyse these metrics of a RBC based on a single image.

An internal research group from the BFH, located at the Swiss Institute for Translational and Entrepreneurial Medicine (sitem) in Bern, is doing research in the field of haematology. Being able to measure the shape and the dimensions of blood cells to measure certain important parameters, can help gain important information to indicate the current health status of a person.

With such an image-based approach certain blood disorder, such as sickle cell anemia, could be detected earlier and countermeasures could be taken. Also it could help standardize the process of analysing red blood cells and reduce the potential for human error in the diagnosis.

1.2 Sitem-insel

Sitem, also known as the "Swiss Institute for Translational and Entrepreneurial Medicine", is a research and innovation center located in Bern, Switzerland. It was established in 2010 as a joint venture between the University of Bern, the Inselspital University Hospital, and the Canton of Bern.

It aims to bridge the gap between academic research and the commercialization of innovative healthcare technologies. Its focus is on translating scientific discoveries into practical applications that can benefit patients.

The research activities span a wide range of fields, including biotechnology, medical devices, digital health, and personalized medicine. They provide support to researchers and entrepreneurs at every stage of the innovation process, from idea generation to commercialization. [1]

1.3 Project objectives

The goal of this Bachelor Thesis is to present a method on how the shape and the volume of a red blood cell can be estimated, based on a single 2D image, provided by the microscope at the sitem in Bern, Switzerland.

To achieve this goal, following questions have been formulated to be examined:

- What are current methods to create a 3D model out of images?
- How accurate is the method applied to the microscopic images?

1.4 Intended / expected results, deliverables and recommendations

The results of this Bachelor Thesis will be a Python script which applies the selected method to predict the shape of the RBC images, provided by the sitem microscope. The created results should then be visualized and important metrics can then be read out. At the end, a quantitative comparison should be presented, which shows the difference between the results of the image absorption method, explained further below, and the selected method.

2 Red blood cells

2.1 Properties / Function

Red blood cells (RBCs), also known as erythrocytes, are the most abundant cells in human blood, accounting for about 40 to 45% of its volume. That is also why the human blood appears mostly red, when they contain oxygen.

RBCs are made in the bone marrow and their primary function is to carry oxygen from the lungs to the rest of the body and to bring carbon dioxide back to the lungs to be exhaled. To achieve this task they are flowing through arteries, veins and capillaries, which the latter often only has a diameter of 8 to 10 µm. The protein which is responsible for transporting the oxygen and the carbon dioxide is called haemoglobin (Hb). It contains also iron to be able to take the carbon dioxide. [2]

2.2 Diseases

There are various diseases related to the human blood. Probably the most popular blood disease is anaemia, which is a condition where the number of RBC in the blood or the concentration of haemoglobin, within the cell, is lower than normal.

The lack of RBC or haemoglobin results that the organs of the human bloody receive not enough oxygen, which can lead to fatigue, weakness, dizziness or shortness of breath. According to the World Health Organization (WHO) roughly 42% of children less than 5 years old and 40% of pregnant women worldwide are anaemic. [3]

2.3 Structure

In a normal healthy state where the RBC is not flowing, the RBC has a distinctive biconcave disc shape, which means it is flattened in the middle and thicker at the edge, resembling much a typical donut without a hole. The cell of an average healthy human is about 7-8µm in diameter and 2µm in thickness at the thickest point. The biconcave shape provides a large surface area-to-volume ratio, which allows for efficient gas exchange of oxygen and carbon dioxide between the cell and its surrounding environment. The shape also allows the cell to be flexible, which enables it to travel through small blood vessels and capillaries, which are often smaller than the cell itself. The RBC is essentially a membrane-bound sac filled with haemoglobin. [4]

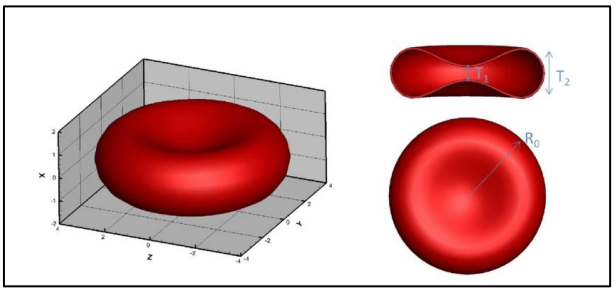


Figure 1: The typical shape of a Red blood cell, discocytes, resembles a biconcave disc shape with a flattened middle part and thick edges, or in simpler terms, a donut with no hole. The high surface-area to volume ratio allow the cell to exchange oxygen faster and deform itself better to flow through narrow capillaries. [4]

2.4 Other Shapes

Deformation due to flow speed

As the RBCs are flowing through capillaries, which can be smaller in diameter, 3-5ym, than the cell itself, the RBCs can change their shape by stretching and deforming to fit through them. The cells then can look like croissants or are deformed as shown in the figure below. [5]

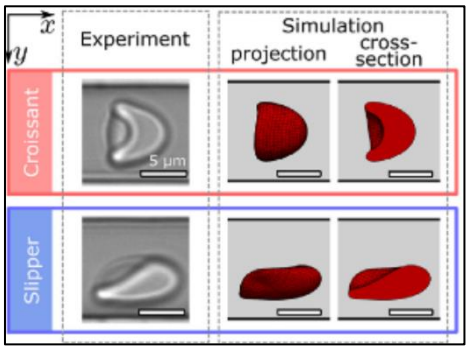


Figure 2: The shape of a red blood cell can change due to flow speed and look like croissant or slippers. [5]

Deformation due to diseases

The RBCs can also take other shapes, regardless of the flowing speed. One example is the spherocyte, where the RBC takes a spherical shape, where there is a lack of central pallor compared to a healthy disc shaped RBC, see figure below. Due to this shape, the RBC have a lower surface area to volume ratio and the deformability is limited, which leads to that the cells die much earlier and don't provide the body with enough oxygen. Patients having these types of RBCs in their body are diagnosed with the Spherocytosis illness, a different type of anaemia. [6]

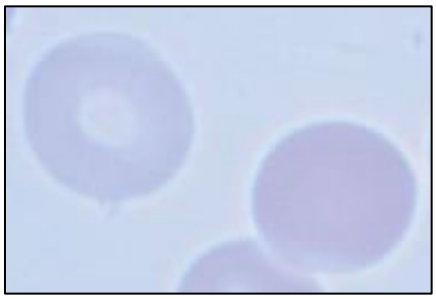


Figure 3: Left : Healthy disc shaped RBC (discocyte) with pallor in the middle due to the reduced thickness. Right: spherical shaped RBC (spherocyte) with lack of pallor due to the spherical shape. [6]

Another disease is the sickle cell, in which the shape of the RBC looks like a sickle, shown in the figure below. This is because the person with the disease has inherited abnormal haemoglobin, which causes the cell to be hard and sticky, and therefore can't serve the body with enough oxygen because they are blocking the blood flow. Normally a healthy RBC lives up to 120 days, but the sickle cell only about 10 to 20 days, which leads to anaemia. [7]

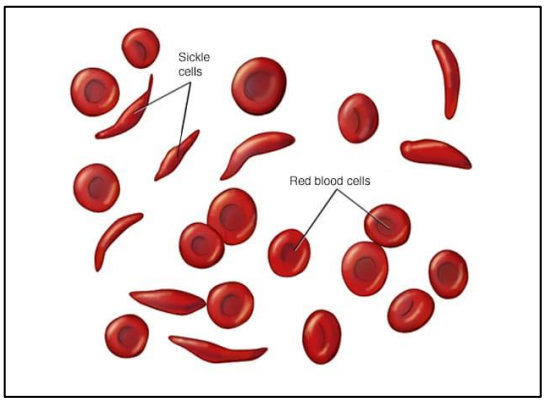


Figure 4: Sickle cell next to normal disc shaped RBC. They get their name because of their shape, which resembles the shape of sickles. mayoclinic.org

2.5 Important data about blood

To identify the diseases mentioned above, blood tests are carried out either with traditional blood tests, described in the next chapter, or with point of care devices. These blood tests try to measure important data about the human blood to identify whether the patient has a disease or is at risk of it.

In the following, metrics are described which are of importance for determining the health of a person based on their RBCs.

RBC Count
The red blood cell count is the number of RBCs in the human blood and is assessed with a complete blood count test (CBC). For male there are normally 4.7 to 6.1 million red blood cells per microliter of blood. For women it is 4.2 million to 5.4 million. A higher RBC count can make the blood thicker and cause blood clots. [8]

Mean Corpuscular Volume (MCV)
The mean corpuscular volume is the mean volume size of the RBCs and is denoted usually in femtoliters (fL). For males, normal volumes are 89.8 to 93.6 fL and for females 90-92.5 fL. Having a too low MCV value can cause iron deficiency and leads to microcytic anaemia. [9]

Red Blood Cell Distribution width (RDW)
The RBC Distribution width is assed as part of the complete blood count test, to gain information about the variation in the size of the RBC sample. The RDW is calculated by determining first the standard deviation of the RBC sizes and expressing it as a percentage of the mean cell volume (MCV). A normal reference range for the RDW is 12 to 15% [10]. Higher RDW indicates greater variability in RBC size while a lower RDW suggests more uniform cells. [11]

Haemoglobin Count
The Haemoglobin count measures how much haemoglobin is in the human blood. For males, normal amounts are 13.8 to 17.2 grams per deciliter (g/dL) and for females 12.1 to 15.1 g/dL. A low Haemoglobin count means that the body doesn't get enough oxygen and can lead to feeling fatigue. [12]

2.6 Methods to measure human blood properties

Haematology Analysers

To measure the volume of RBC and the amount of haemoglobin in it, clinical haematology analysers are utilised. These analysers work by using a combination of optical and electrical methods to measure various parameters of blood cells, such as their size, shape, and count.

When a blood sample is loaded into the analyser, it is mixed with a reagent that lyses the red blood cells, releasing the haemoglobin and other intracellular components. The analszer then uses optical methods to measure the absorbance of the released haemoglobin, which can be used to determine the total haemoglobin concentration in the sample.

The analyser also uses electrical impedance to determine the number and size of different types of blood cells, such as red blood cells, white blood cells, and platelets. As the blood cells pass through a small aperture in the analyser, they cause changes in the electrical impedance, which can be used to determine the cell size and count. [13] [14]

Figure 5: Example of a haematology analyser. [14]

Confocal Microscope

A method to measure single cell properties is the confocal microscope. It is a type of microscope that uses a laser or other light source to illuminate a cell probe for example. The light is focused onto the sample by a lens, and the reflected light is then collected by the same lens and directed to a detector.

However, unlike a traditional microscope, a confocal microscope uses a pinhole aperture to block out-of-focus light. This allows for better resolution and contrast in the image, as only light from a very small focal plane is detected.

The microscope can then scan the sample in the z-axis (up and down) to create a series of 2D images, which can be combined to create a 3D image of the sample. The stacking of these 2D images, is time consuming, as it can take up to several minutes in complex images, and therefore not suitable for high throughput applications, where the quick diagnoses have to be made for the patient.

Overall, a confocal microscope provides high-resolution images of a sample by eliminating out-of-focus light and allowing for the capture of only in-focus light. [15]

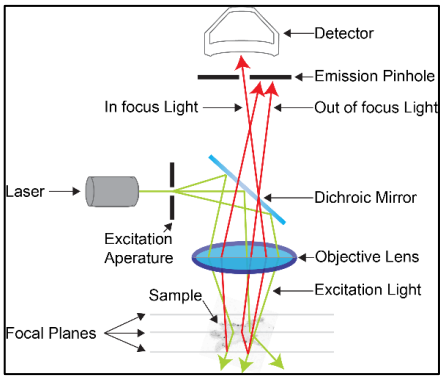


Figure 6: Function principal of a confocal microscope. Note that only the “in focus light” gets to the detector.[15]

Scanning Electron Microscopy
Another method to get information about the texture, composition or the inner structure of a single is the scanning electron microscope (SEM).

The SEM uses, instead of light, a focused, with the help of an electron gun, beam of high-energy electrons to generate a variety of signals when hitting the surface of a cell sample. This electric beam is held in vacuum, for better controlling. Detectors can then collect these ejected signals, such as electrons and x-rays, to convert them to the final image. Because the SEM uses electrons and vacuum conditions to generate images, the samples have to go through some preparation for the scanning. For example, all the water has to be removed because it would boil in the vacuum. Another preparation is that non-metallic samples have to be coated with a thin layer of conductive material. The fact that the sample is in an vacuum, also means that the RBC is no longer alive.

An advantage of the SEM over other typical light microscopes, is the fact that it has a larger depth of field which allows it to have more than one sample be in focus at a time. Another advantage is that the resolution is much higher and better adjustable compared to other microscopes, thanks to the fact that is operated with electromagnets rather than lenses.
[16]

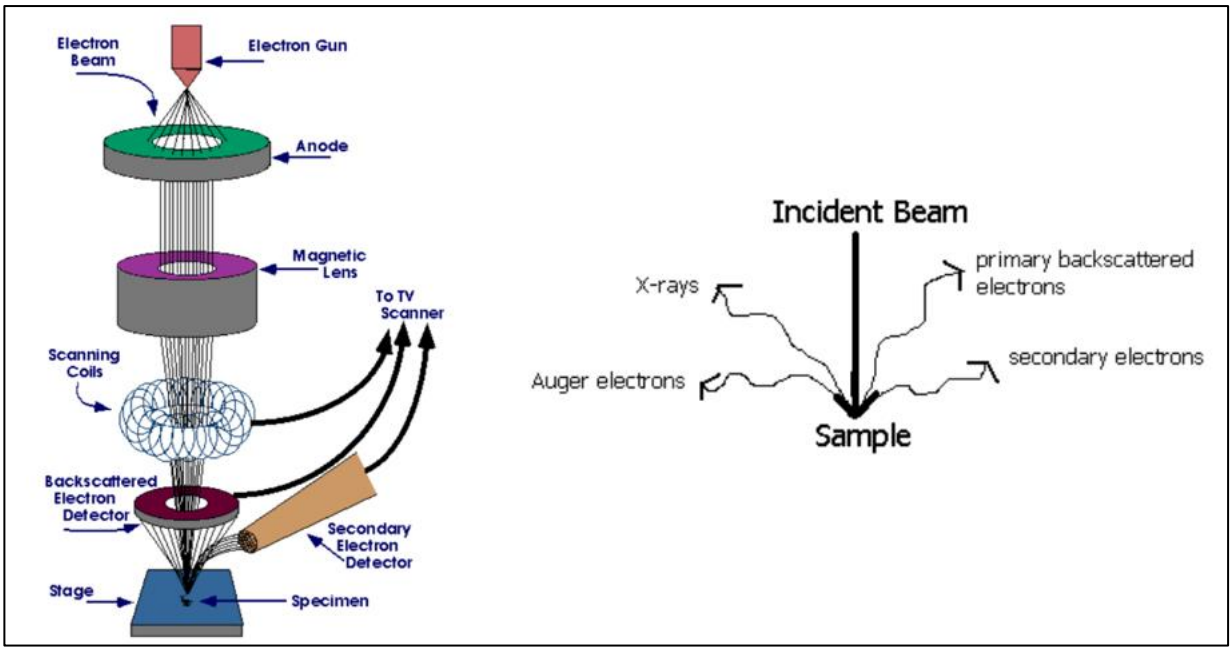


Figure 7:Left : Functional principle of a SEM. The microscope consists of electron gun that creates a beam which is focused on a specimen (sample). As soon as the beam hits the sample it will reflect from its surface and create multiple signals such as x-rays or auger électrons (Right image).The reflected signals are detected by multiple detectors to produce an image. [16]

3 State of the art: Methods for volume extraction of objects in 2D images

The traditional hematology instruments, explained in the previous chapter, are not suitable for high throughput applications, either because of their time consumption or inability to create 3D models of the RBC. That is why in this chapter, alternative solutions have been searched to provide a solution on how 3D models can be created based on 2D images.

During a pre-study (Industry Project 2) for this Bachelor Thesis, some methods to estimate the volumes of objects were presented and evaluated, which are also included in this chapter together with other methods.

3.1 Machine Learning

Cobo et. al. presented a method to use a deep-learning algorithm, namely a convolutional neural network, to be able to estimate the red wine volume filled in wine glasses. [17]

They created a database with various images, which show filled wine glasses in different angles and distances to the camera. They have also used different glasses. To each of the image, the filled in volume of the wine is attached, which range from 50ml to 300ml.

Figure 8: Sample images of dataset. The dataset consists of images with different types of wine glasses with different amounts of wine in them. The images were taken in different locations and varying distance and view to the camera. [17]

By training the model they were able to estimate the volume with no mean absolute error (MAE) higher than 10%. The following images show the MAEs for each amount of wine filled in the glasses.

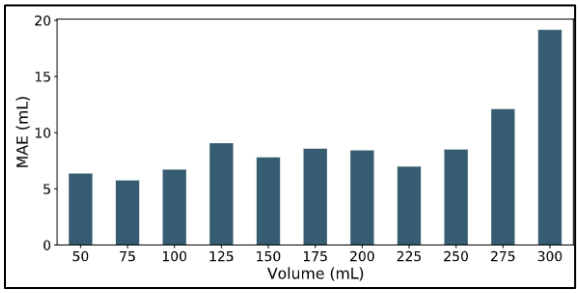


Figure 9: The Mean Absolue Error (mL) of the model with different wine volumes. An image which has a wine glass filled with 125mL of wine was predicted with a MAE of 10mL. [17]

If applied to the RBC image, this method would only be able to estimate the volume of the cell and not gather more information about its shape. Therefore no real 3D visualization could be made with this method.

3.2 Molecular Depth Estimation Network

An extension of the method, shown by *Cobo et. al.* [17] to predict the wine volume, is the usage of molecular depth estimation networks, presented by *Graikos* et al.[18], to estimate food volume from a single input image.

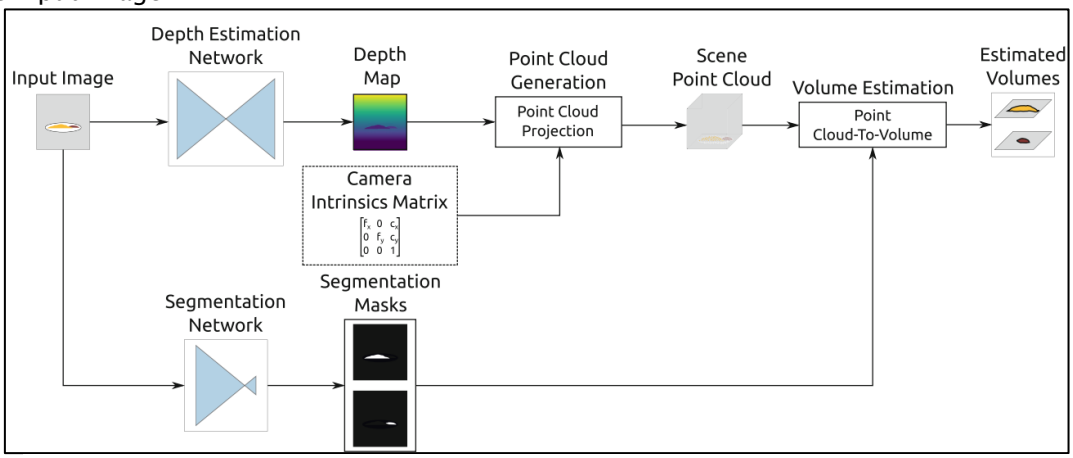


Figure 10: The food volume estimation system can be separated in three distinct parts: the depth estimation network, the segmentation network and the point cloud-to-volume algorithm. [18]

The authors propose a framework that combines a depth prediction network with a food segmentation network to estimate the volume of individual food items in the image.
The depth-estimation network is trained on image sequences, 3 consecutive frames of a video, which the authors have collected from a database, which consists of over 50 hours of food video material. With the help of the captured frames, in which each of those frames the camera has another position relative to the food, it is possible, together with the intrinsic parameters of the camera, to estimate the depth map of the image scene, without having a reference object.

Together with the food segmentation, which was also trained with a neural network, they are able to project each pixel to a corresponding point in a 3D space (Point cloud representation).

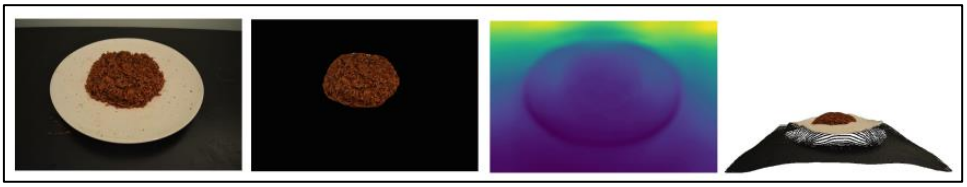


Figure 11: The system stages of estimating the volume of a food object. From left to right: Input Image, Segmentation mask, Depth estimation map, Point cloud generation. [18]

This method of estimating the volume would match very good with the case of the RBCs, because it does not only estimates the volume, it also takes the topographical information into consideration, which allows for reconstructing the 3D model in form of a point cloud.

Problems that could occur when applying this method is the fact that, to have accurate performance of the model, enough prelabelled training data of the RBCs has to be available. Further in the research paper, in each frame of the videos, the food is shown from different angles from the camera. In case of this Bachelor Thesis , it would not be the camera that is changing the position, rather it is the RBC. And because the cells are moving really fast through the channel, it would be challenging being able to track them across multiple frames.

3.3 Pixel2Mesh

The paper "Pixel2Mesh: Generating 3D Mesh Models from Single RGB Images" by *Wang et. al.* proposes a novel approach for generating 3D mesh models from 2D RGB images. The method is based on a deep neural network that takes an image as input and produces a 3D mesh as output.

The network consists of two main components: an encoder and a decoder. The encoder extracts features from the input image, while the decoder generates the 3D mesh. To generate the mesh, the decoder first predicts a set of 3D points and their corresponding 2D projections. It then constructs a triangular mesh by connecting nearby 3D points based on their proximity and the coherence of their 2D projections. The resulting mesh is optimized using a loss function that encourages the mesh to be both geometrically accurate and visually plausible.

To train the network, the authors use a dataset of RGB-D images that provide both 2D color information and depth information. They use the depth information to generate ground truth 3D meshes for each image, which they then use to train the network.

The authors evaluate their method on several benchmark datasets and show that it outperforms several state-of-the-art methods in terms of both geometric accuracy and visual quality. They also demonstrate the effectiveness of their approach in various applications, including object reconstruction and shape completion. [19] One problem, when utilizing such a network, could be that there are no ground-truth information about the dimensions of the objects, i.e. whether they are in millimeter or meter etc.

Figure 12: Created meshes applied to real-world images from online product datasets and the internet. [19]

4 Neural-Network based shape prediction (SHAPR)

The problem with the methods presented in the previous chapter, is that the existing methods for reconstructing the 3D shape of an object, based on a single image, can't be applied to the RBC, because of various reasons: First, the object properties of a RBC are fundamentely different from real-world images in terms of color, contrast and object orientation.

Second, unlike the shapes of cars or furniture that might vary because of differing photographic viewpoints, the shapes of single cells are similar but never the same, and it is often not feasible to image the same cell from different angles in high throughput microscopy.

Finally, existing computer vision algorithms have been trained on tens of thousands of photographs where synthetic 3D models are available. [20] [21] For red blood cells there are much less available.

4.1 SHAPR Framework

Waibel et al. present with the SHAPR open-source framework a solution on how the shape of single RBCs can be predicted based on a single 2D image.

They address the challenges of costly and limited imaging techniques such as confocal and electron imaging by leveraging a machine learning algorithm to predict the 3D shape of cells from 2D images. The framework consists of two main components: an encoder, that extracts features from 2D images, and a decoder, which acts as a shape prediction module, that uses the extracted features to predict the 3D shape of the cell. Both the encoder and decoder are integrated in a convolutional neural network.

To train and validate the SHAPR framework, they used a recently published dataset containing 3D red blood cell shapes taken from confocal microscopy images, where each cell was reconstructed from 64 confocal images which are stacked up in the z-direction, resulting in a 64x64x64 format. By using the 2D image, that intersects the RBC at the center-slice and its corresponding binary segmentation, SHAPR was trained in a supervised manner, as shown in the figure below. To increase the prediction accuracy, a discriminator model, an additional neural-network, was trained to differentiate between true and predicted 3D cell shapes (Adversarial training).
[22]

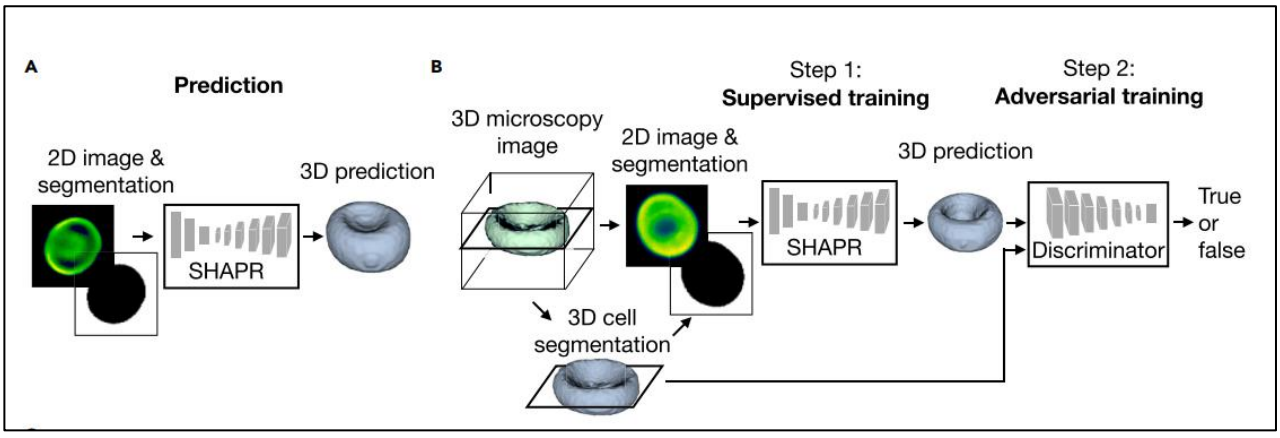


Figure 13: A: The SHAPR framework needs an image of the RBC and its corresponding binary segmentation to generate a synthesized 3D model of the RBC. B: In a first step the network is trained in a supervised manner by using a publicly available dataset containing 3D confocal microscopy images from RBCs with the corresponding 2D image and segmentation. To improve the model parameters, a discriminator is trained which tries to differentiate between true and predicted 3D cell shapes. [22]

4.2 Convolutional Neural Networks

As SHAPR and some of the methods presented in the previous chapter make use of the term convolutional neural networks, this section aims to briefly describe what they are, how they work and what role they play for the SHAPR Framework.

Deep Learning is a subfield of machine learning, which usage has grown rapidly in recent years [23], and the techniques, that are used in this field, try to mimic the way how the human brain learns. One of these techniques is the convolutional neural network (CNN), which can be used to detect tumors in medical images, recognize object in images, or classify if a bone is fractured in an X-ray image. [24]

For the SHAPR Framework, a CNN is used to extract import features from the input image, also referred to as the encoder, which are then used to predict new 3D shapes, also referred to as decoder. So the encoder and decoder are integrated in a single CNN.

The structure of a CNN involves multiple building blocks, such as convolutional layers and pooling layers, batch normalization layers or dropout layers.

The encoder consists of three convolutional blocks, each with two convolutional layers which are each separated with a batch normalization layer, and pooling layers. At the end of every block a dropout layer is included.

The decoder, on the other hand, has seven blocks. These blocks not only consists of normal convolution layers, instead of pooling layers which are used to decrease the feature map resolution, it has transpose layers, which are used to increase the resolution of the feature map, leading finally to the desired 64x64x64 shape. They, in simple terms, take the feature map and try to undo the convolution form the encoder, effectively increasing the feature map size and dimension. [25]

Convolutional Layer

The convolution layers are responsible for extracting the features, by applying a combination of linear and non-linear operations. Convoluting over an image, also known as a grid of numbers or a tensor, involves performing a linear operation for feature extraction. This operation utilizes a small array of numbers, called a kernel, which is element-wise multiplied with the input tensor at each pixel. The resulting products are then summed to obtain the output value in the corresponding position, creating the feature map, as shown in the figure below.

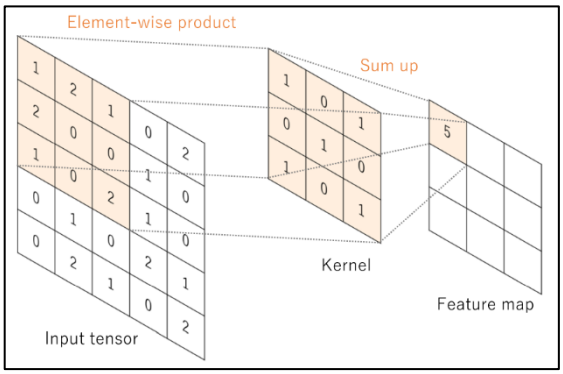


Figure 14: The kernel is convolving across the input tensor (image). The element-wise product between the input tensor and the kernel is calculated at every location and stored in a feature map. [25]

The output from the linear convolution operation is then passed through a non-linear operation function, namely the rectified linear unit (ReLu), which equals to f(x) =max(0,x). This function returns the bigger value between x and zero. If x is bigger than zero, it will return x, if the number is lower than zero, the function will zero as an output. The purpose of this operation is to introduce non-linearity to the network, enabling it to model complex relationships and capture non-linear patterns in the data.

Pooling Layer

Pooling layers is a typical down sampling operation, used to reduce the feature map dimension. SHAPR utilizes the max pooling operation for the encoder and the average pooling layer for the decoder. The max pooling extracts the biggest value from a patch and the average pooling calculates the average value. [26]

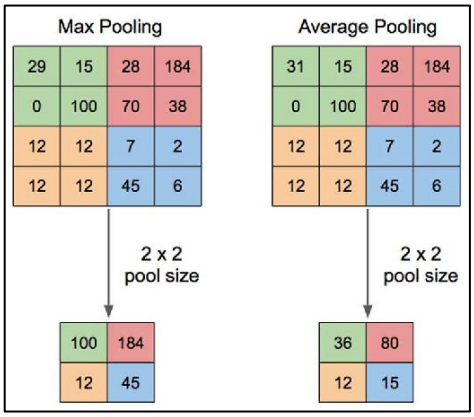


Figure 15: Illustration of Max Pooling and Average Pooling. In max pooling the 2x2 pool/filter is convolving over an image and finds the biggest value from the patch and stores it in a new output. In average pooling the average from the patch is stored as an output. [26]

Batch Normalization Layer

The batch normalization layer, as the name itself already says, normalized the data, preventing that the weights, the elements in the kernels (filters), of the network get too large, which could results in feature maps with pixels that are spread across a wide range. It also increases the efficiency of the networks, allowing higher learning rates .

Dropout Layer

Dropout layers are regularization technique that helps prevent overfitting, which occurs when the network becomes too specialized in learning form the training data and fails to generalize to new, unseen data. The dropout layer randomly and temporarily "drops out" a fraction of the connections between the units in the previous layer and the next layer. This means that during each training iteration, a different subset of connections is deactivated. As a result, the network is forced to learn more robust and generalizable representations by relying on the remaining connections.
[27]

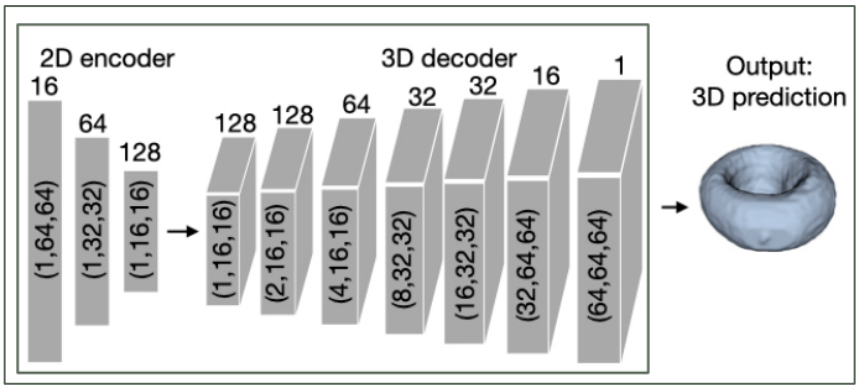


Figure 16: SHAPR Neural-Network with encoder that extract the features of the 2D input and reducing the feature map with multiple convolutional layers. The decoder upsamples the feature map with the help of transpose layers, leading to the 3D shape prediction. The process of the input data being processed through the network is called forward propagation. Backward progragation is used during the training where the errors are propagated back from the last layer to the first layer and constantly updating the model parameters, also called weights, i.e. kernels, which are then stored as pretrained weights.[22]

4.3 Generative Adversaria Networks

Generative Adversarial Networks (GANs) can generate samples that duplicate the distribution of the training dataset. They do this by training two neural-networks, usually CNNs, together, namely the generator, the SHAPR encoder and decoder used for prediction, and the discriminator. The generator tries to predict and generate new synthesized images which are indistinguishable from real-world data to mislead the discriminator into thinking they are real inputs. On the other hand, the discriminator is a classifier that discriminates whether a given input looks like real data from the dataset or an artificially synthesized data, i.e. the prediction. As both networks continuously optimize themselves to improve their result and learn from their own as well as each other's weaknesses, the networks become stronger during the training process. The reason that it is called "Adversarial" is because the two networks are always competing against each other up until the point that the discriminator fails to distinguish real data from fake synthesized data.
[28]

GANs can be used for several applications in the field of computer vision, machine learning or photoshop. In the case of SHAPR, the two CNNs build up the GAN, namely the predictor, consisting of an encoder and decoder, shown in figure 16, that predicts new shapes, and the discriminator which aims to classify if the 3D input is an prediction or ground truth data, as shown in the figure below.

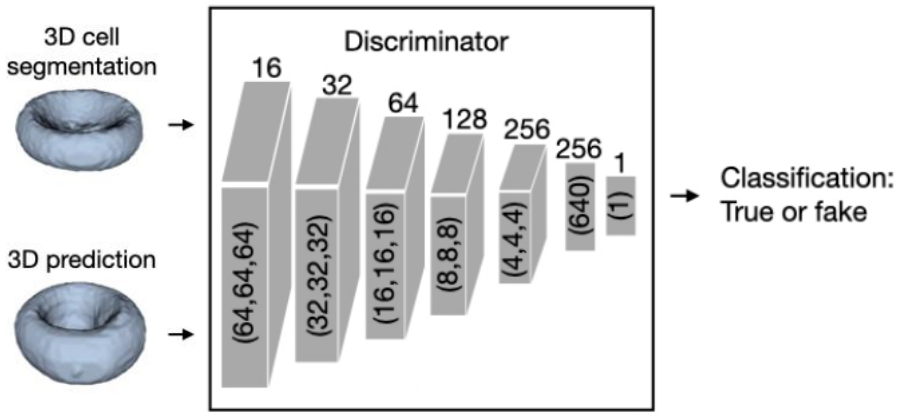


Figure 17: The discriminator is a classifier, consisting of multiple dense layers, that tries to predict whether a 3D input is a prediction or ground truth data, i.e. an actual 3D cell segmentation from a confocal microscope .[22]

In the following some examples are presented where these GANs can be used elsewhere.

Image to Image Translation

GANs are able to generate new samples of images that match that of a given training set. With style transfer methods it is possible to change the visual style of images. CycleGAN is able to alter the style of a given image and still preserve the original information, so that only the colours, i.e. the ambience, are changed. [29]

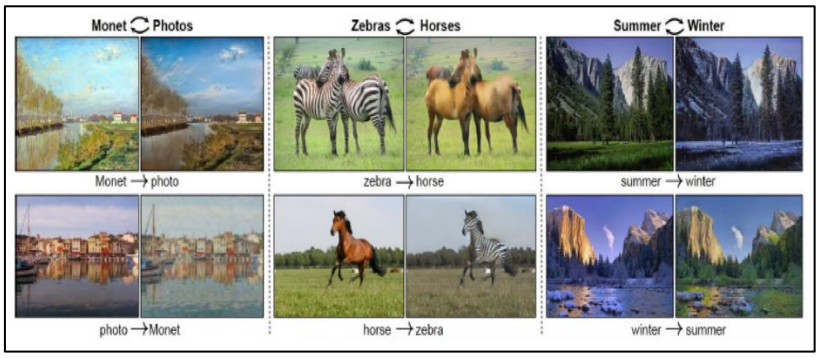


Figure 18: Examples of image translations performed by the CycleGan model. In each example left is the input image and right the result. [29]

Text to Image Translation

Reed et. al. used a GAN architecture to synthesize images from text descriptions. With a given text caption of a bird, trained GANs can generate several samples of images that match the description. [30]

Figure 19 : With a given text caption of a bird, a trained Gan can generate several samples of images.

Image Inpainting

Image inpainting is the process of reconstructing missing parts of an image so that observers are unable to tell that these regions have undergone restoration. Yu et al. present a GAN approach which can utilize surrounding image feature as references during network training to make better predictions. These applications can be used in photo editing, image-based rendering and computational photography. [31]

Figure 20: Example of inpainting result. Missing regions are shown in white. The left is the input image and right is the output of the generative adversarial network. [31]

5 Methods

5.1 Microscope setup for red blood cell images

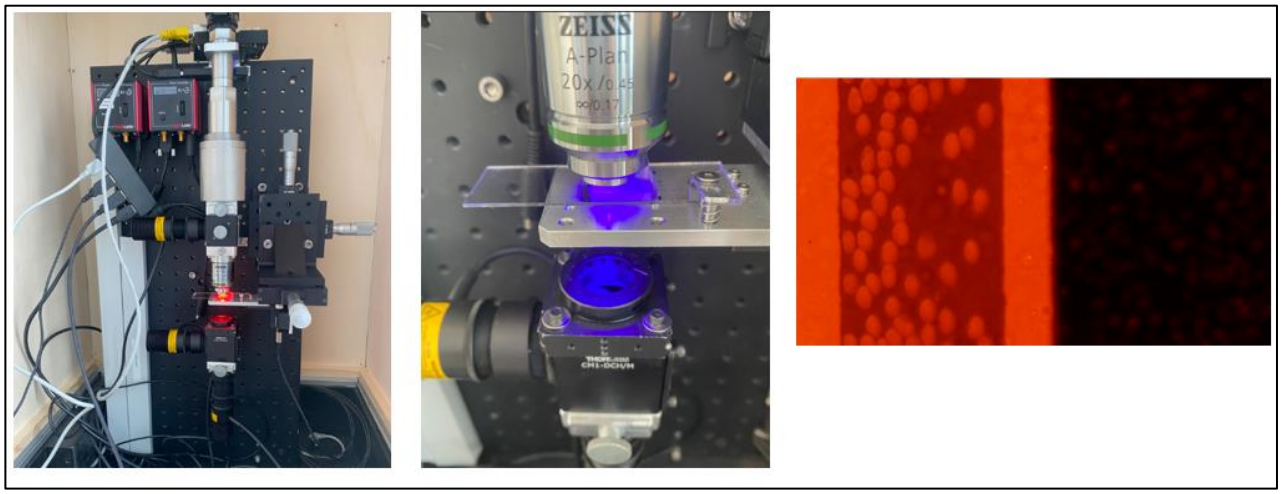


Figure 21: Left: The microscope installation at the sitem in Bern with two LEDs. Middle: Close-up view of the focal-point of the microscope with blue LED light, the focal point is where the micron-sized channel is and the blood is flowing through. Right: Frame from the Raspberry Pi Camera filming the RBCs flowing through the channel, which has a width of 50µm and height of 5µm. Source: sitem-insel, Bern

At the Swiss Institute for Translational and Entrepreneurial Medicine (sitem) in Bern there is a microscope, shown in the figure above, that is used for research purposes in the field of haematology and medicine.

To measure the RBC, micron-sized channels have been used which let the cells flow through the channel thanks to the capillary effect created when a drop of blood is put in the entrance. The channels were made with polydimethylsiloxane (PDMS) molding. The pattern were the blood is flowing through was developed with photolithography in a clean room at the BFH in Burgdorf. The mixture was spin-coated on a silicon waver before the UV exposure, to achieve a channel height of 5µm.

5.2 Absorption imaging

With this setup, it is possible for the researchers at sitem to create a height and haemoglobin mapping based on the Beer-Lambert law of light absorption and the procedure proposed by *Schonbrun et al.* [32]

The setup at sitem is very similar to that presented in the paper by *Paul et al.* [33] With the help of dye exclusion absorption imaging, the height of the RBC can be measured, given the known channel height. When the dye solution aborbs light of a certain wavelength, the presence of a cell in the optical path will increase the light transmission by removing the dye particle from the optical path. For this work a red absorption dye has been used.

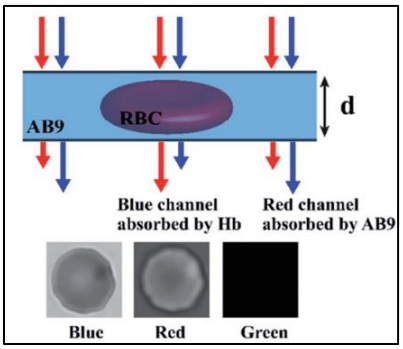


Figure 22: Graphical representation of the light absorption of AB9 and RBC. The intensity of the blue light decreases while passing through an RBC. The intensity of the red light, on the other hand, decreases while passing through the AB9 solution. Three channels are shown that are extracted from a typical color image. As there is no green light in the source, the green channel is completely dark. [33]

With this method a height map can be created which shows the thickness of the RBC at every pixel, as shown in the figure below. Like this the volume of the cell can be calculated, altough no 3D model.

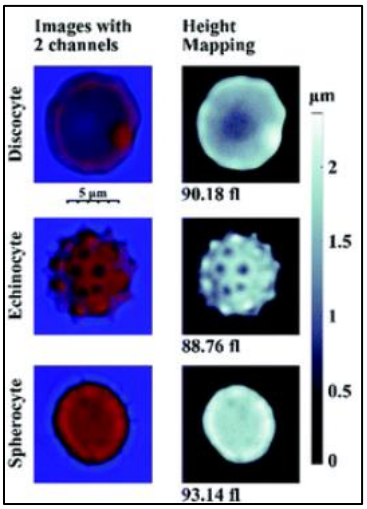


Figure 23: Height mapping of different RBC types: Discocyte, Echinocyte and Spherocyte. The maximum height in the height mapping is 2.8 µm. The scale in the figure is applicable for all three RBCs. [33]

5.3 Literature research

In chapter 3 a literature research has been conducted based on the pre-study from the Industry Project 2. Multiple methods have been presented on how the volume of an object can be extracted based on 2D images. Based on this research, the method presented by *Waibel et al*. with their created open source package, SHAPR, has been chosen to be the most optimistic method to be predict 3D shapes of RBCs.

5.4 Applying the trained models to the images

Dataset

The same dataset was used for training as stated in the research paper by *Waibel et. al.* The dataset consists of 825 publicly available 3D images of RBC of size (64,64,64) voxels, each assigned to different classes: SDE shapes, cell cluster, multilobate, keratocyte, knizocyte and acanthocyte. In the SDE class the cell shape of spherocytes, stomatocyte, discocyte and echinocyte are represented. The other shape classes occur in samples from patient with blood disorders. The number of cells in each classes were: 602 for SDE shapes(93 spherocytes, 41 stomatocytes, 176 discocytes, and 292 echinocytes), 69 for cell clusters, 12 multilobates, 31 keratocytes, 23 knizocytes, and 88 acanthocytes.

The 2D images were extracted from the central slide of each 3D image and the segmentation was created with thresholding. [34]

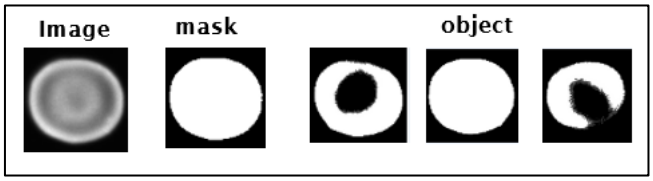


Figure 24: The dataset consists of 2 "features" in the training set, the greyscale image which is stored in the folder "image", the segmentation stored in “mask” . The ground truth 3D data can be seen as label data, and are stored in the folder “obj”. In the figure an example of a discocyte cell can be seen and its 3 slices of the total 64 stacked up images, (slice 26, 36, 42 out of 64). The segmentation image, stored in the “mask” folder is the center slice of the 64 images. [34]

Training

The model was split up in 20% for the test set, 20% for the validation during training and the remaining 60% were used to optimize the model weights during training.

The model was trained on 5 folds, while the folds were randomly split. To optimize the parameters and the accuracy of the model, 30 epochs were used during training.

These hyper-parameters correspond with the ones used in the SHAPR scientific paper presented previously, so that the generated models are similar to the “original” from the paper.

The training gave as an out 10 models with pretrained weight parameters as hdf5 files: 5 models which were trained in an adversarial manner (GANs) and 5 which were trained without the discriminator. To create and compare the results with the absorption imaging method, only the GANs (Nr. 0-4) were used for the comparison, as they were only created when they performed better in the training than the previous models.

Important to mention is that the models were trained by the supervisor because the computational capacity was not available from the author.

Pre-Processing RBC images

For this project, multiple images are generated by sitem, which show single RBCs captured from the video sequence of the blood flowing through the micron-channel. In total 11'668 images were captured. The blood samples are taken from a single patient at the Insel hospital in Bern. Together with the blood sample comes the corresponding data of each single cell, including the width and height in pixel terms and a multiple of the true volume.

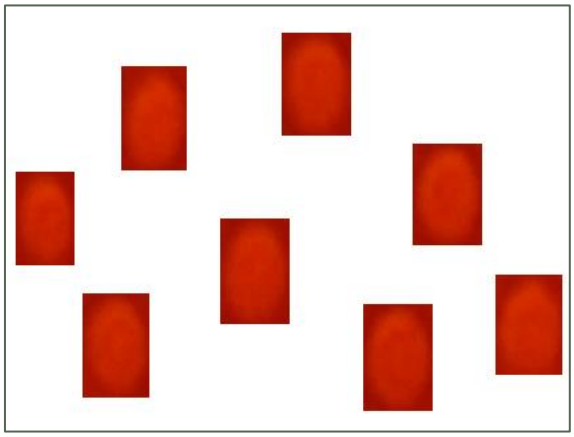


Figure 25 : Some samples from the total of 11'668 single RBC images.

Before the model can be fed with the images, they first have to be pre-processed, so that they meet the requirements of the SHAPR framework and look similar to the images used during training, see figure below.

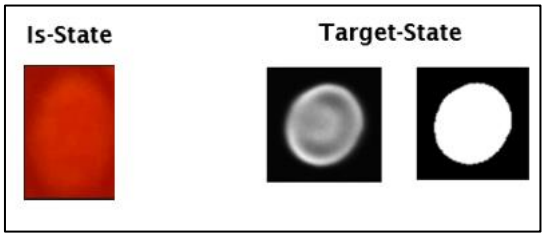


Figure 26: Left: Image of cropped out RBC from video sequence, provided by sitem. Right: Greyscale image and binary segmentation of single discocyte RBC used for training the SHAPR model. They are stored as tif-files and have size of 64x64 pixels. The current images should be put in greyscale and a binary segmentation has to pe performed. At last the image file should have the same size and type as the ones used for training.

Prediction and visualization

As a next step the images are used to make predictions about the shape and the RBC is then visualized as a 3D model. Together with the prediction the cell volume can be read out in terms of μm^3, respectively fL . For the visualization either external software will be used or a python script will be written.

5.5 Compare results

After the predictions were made, the results can then be quantitatively compared with the height map data created from the image absorption method, provided by sitem. The data involves the mean cell volume and the standard deviation of the volume distribution from the blood sample, which can then be compared with the predictions.

6 Results

6.1 Pre-processing RBC images

Because the images provided by sitem, containing the single RBCs, have different sizes and usually aren't square-shaped, they have to be pre-processed. The pre-processing involves only extracting the red channel of the image, because the video of the microscope was taken with red LED lights and absorption dye. For this purpose two python scripts have been created, in form of a jupyter notebooks, one which extracts the red channel and one which resizes the image so that it fits the requirements of 64x64pixel size.

This image resizing was achieved by first determining the aspect ratio of each image, which is the ratio between the width and height of the image and is calculated by dividing them. Because the ratio of a square is 1 (width divided by the height), the aspect ratio of the RBC image is either higher or lower. When it is higher, the width is the limiting factor of which the image will be resized to the desired 64pixels and when it is lower, the height is the limiting factor.

In the images provided, almost every image had an lower aspect ratio than 1, because they are taller than wider and therefore the image-scaling-factor was calculated so that image had the desired height of 64pixels. The width of the image was resized with the same image-scaling-factor, to keep the aspect of the RBC.

To make the image square-shaped a grey background canvas was added, which has a width of 64pixels and takes the background color of the grayscaled images. The color of the canvas has to be the same as the background color of these images, because in the next step they have to be thresholded to create binary segmentations.

The binary thresholding of the images labels every pixel in a greyscale image above a certain threshold as white and the rest as black. If the intensity of the canvas color is too high or differs too much from the original background, the segmentation can't be performed accurately and the prediction may produce bad results. [35]

The whole process is visualized, with an example, in the figure below.

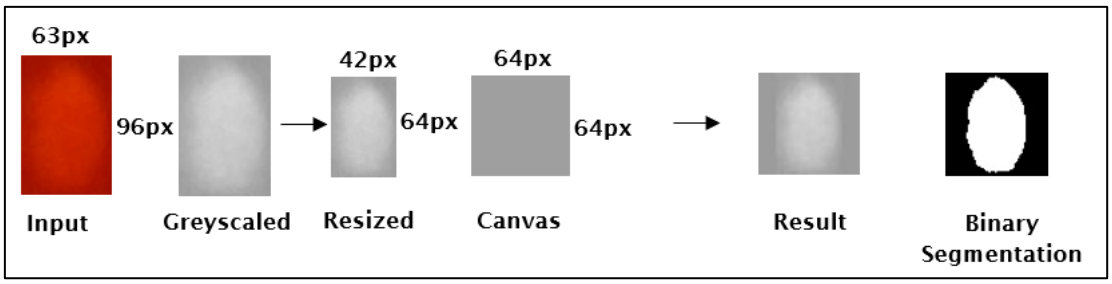


Figure 27: In a first step, the red channel of the input image has to be extracted, to create the grayscaled image. As a next step, this image has to be resized so that the height equals to 64pixels, without losing the aspect ratio from the width and the height. Because of the resizing, the width of the RBC image is smaller than 64pixels. To solve this, a canvas image is created with the desired 64x64 pixels dimensions and the same color value as the background from the greyscale image. The resized image is then put in the center of the canvas creating the final result. As a last step the resulted image is then thresholded at a specific pixel value to create a binary segmentation which is required by the model. Both these images are stored in separate folders, see Jupyter Notebooks. Note that the metric-to-pixel ratio of the setup, 75nm/pixel, described in the next chapters, is increased according to the image-scaling-factor.

6.2 Results from the image absorption method

11'668 images of single RBCs, from a blood sample of a single patient, were provided by the microscope at sitem. The images were automatically cropped out from a video sequence filming the RBCs flowing through the micron-channel. Information about the position in the video sequence, dimensions and cell volume of the RBC were stored in an excel file. Note that not all RBCs from the sample could be captured as images.

The Mean Cell Volume (MCV) of the whole blood sample equals to **91µm³** or **fL** and the Red Cell Distribution Width (RDW) is **12.7%** which equals to a standard deviation of **11.56fL.**

The values for the cell volume, don't represent the real volume of a RBC in terms of femto-liters, rather it is a multiple of it and represents a pixel value. To get the real metric, the values have to be multiplied with the absorption- and scaling-coefficient.

The absorption coefficient determines how much light the RBC can absorb as it passes through it. A high absorption coefficient means that the cell absorbs more light, making the cell appear darker on the image. On the other hand, a low absorption coefficient indicates that the cell absorbs less light, making it appear brighter. With this coefficient it is possible to distinguish between thick and thin parts of the RBC, as thicker regions let less light through than thinner regions. For this blood sample it wasn't possible to determine this coefficient for every single RBC image, because it depends on multiple factors which couldn't be obtained yet, such as blood quality, wavelength of the light and the plasma in which the RBC is flowing in.

The scaling coefficient determines in every provided image, how long 1 pixel is in the real world. In this work the metric would be µm³. In the setup of the microscope at sitem, the pixel to metric ratio is set to 75.5nm per pixel respectively 0.0755µm per pixel.

The two coefficient can be multiplied together, resulting in a combined coefficient to get the real world cell volume in femto-liters [fL], as shown in the figure below.

$$\text{vol [fL]} = C_{\text{scaling}} \times C_{\text{absorption}} \times CV$$
$$\text{vol [fL]} = C_{\text{combined}} \times CV$$

Figure 28 : The cell volume (CV) is a multiple of the real world value in femto liters [fL], represented in a pixel value. To get the real world dimension this value has to be multiplied by the scaling- and absorption coeffiecient, which both can be combined in one single coefficient.

To compare the results of the SHAPR prediction method with the absorption imaging results, a normal distribution curve was created with all the provided images, showing the mean value and the standard deviation of the RBC images.

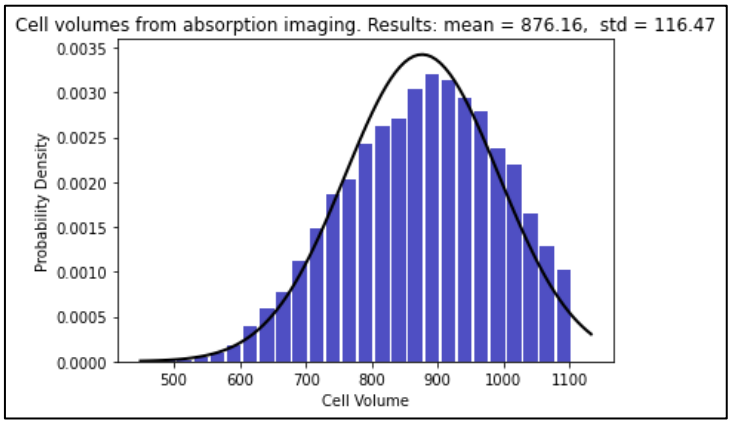


Figure 29: Normal Distribution histogram of the cell volume calculated with the absorption imaging method. The data consists of 11'668 values and the mean is 876 and the standard deviation is 116. Note that these values don't represent the real world cell volume in term of fL, rather they are a multiple of it, shown in pixel values.

6.3 Results from SHAPR prediction

Because of simplicity reasons, only the results of the last generated model is presented here, as it is usually also the best performing. The results from the other models can be found in the appendix.

6.3.1 Examples GAN 4

The predicted cell don't resemble much of that of a typical biconcave disk shape, as presented in figure 1. However, it can be seen that the model tries to implement some sort of concavity in the prediction as can be seen especially in cell C,E and F, in the figure below. The cells look like initial spherocytes with swell up spherical round shape that were stretched in the flow direction of the capillaries.

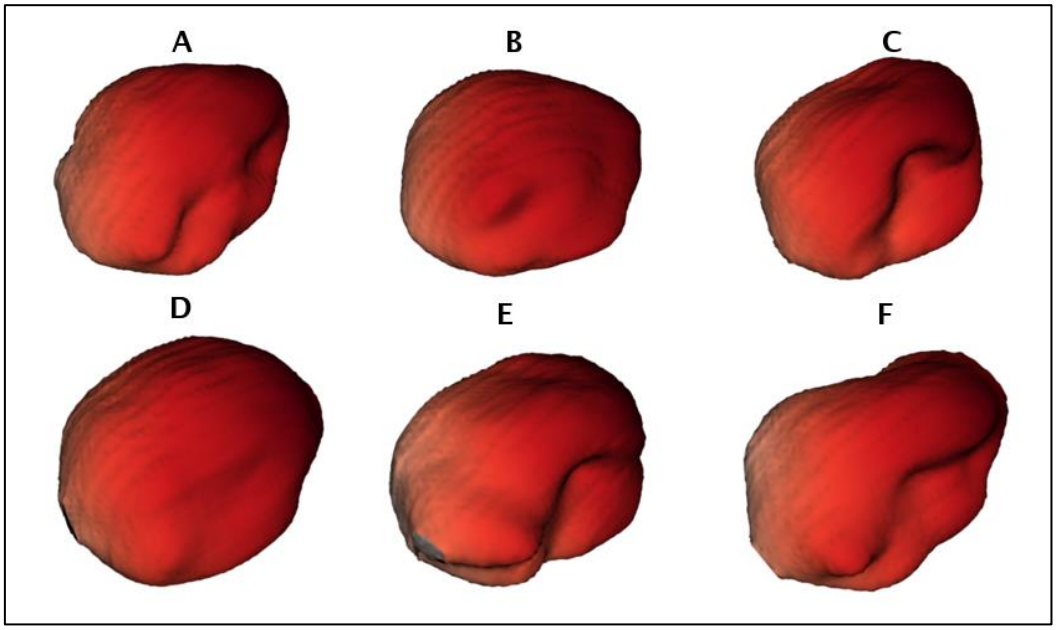


Figure 30 : 3D Volumina examples of predicted RBC shapes with the fifth GAN model « GAN4 ». The results show that the cell have some sort of concavity in their shape, however they don't resemble much that of a typical discocyte.

As the occurrence of patients having spherical shaped RBC and the related disease called spherocytosis is very low in Europe (1 out of 5000 people) and mostly appear in children, the assumption can be made that this blood sample doesn't contain spherocytes or any other disease. [36]

Because of that, the existing dataset, in which the examples above were generated, was altered so that only healthy discocytes were represented and new prediction could be made. This resulted in a dataset with 176 different discocytes, were 60% were used to optimize the model weights, 20% for testing and 20% for validation.

The training of the dataset resulted in 9 models, 5 non-adversarial models and 4 GANs, where the last model is the GAN 3 (GAN0-GAN3). For simplicity reasons, only the results from this model are presented in this chapter. Also for this training set, only the GANs were evaluated, as they would be only created if they performed better during the training than the non-adversarial models. The remaining results from the models can be found in the appendix

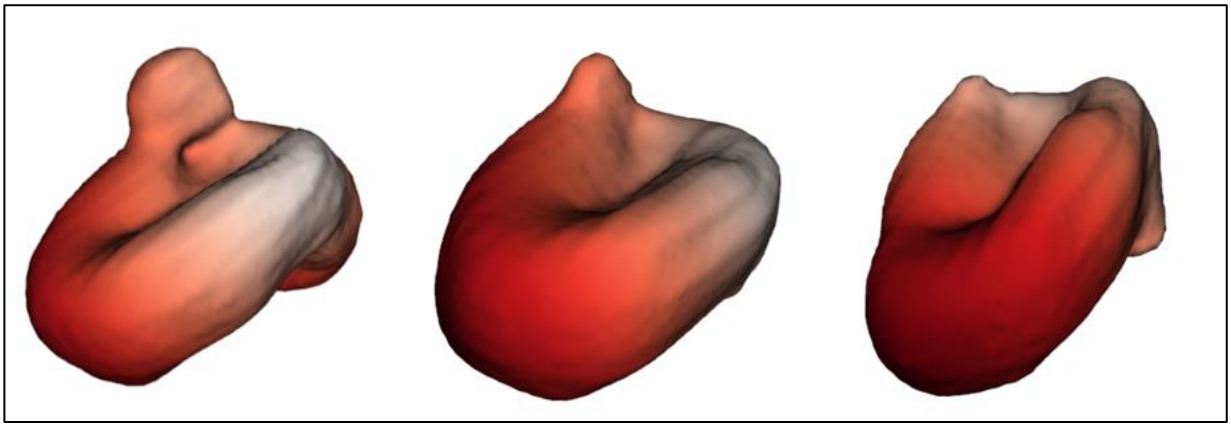


Figure 31 : 3D Volumina examples of predicted RBC shapes with the fourth GAN model « GAN3 », which was trained on a training set which only included discocyte images. Note that there are sometimes little spikes at the edge of the RBC.

The predictions from the discocyte-only training set show that the cells take more of a biconcave disc shape with a flattened middle and thicker edges, than the prediction with all cells included in the training set.
However, in some cases, the model generates little spikes at the edge of the RBC, which are sometimes even fully separated from the whole cell, as shown in the figure below.

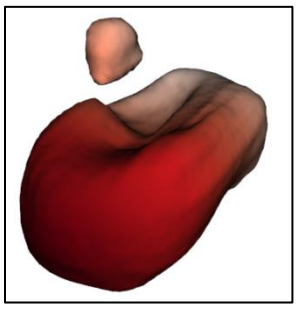


Figure 32: Predicted RBC with spike, that is seperated from the actual cell.

A reason for these little spikes could be, that the pre-processing, where the grayscale and the binary segmentations are created, was not precise enough, or other errors might have an influence. These little spikes, which are separated from the cell, could negatively influence the volume calculation, as they are still taking into account for it.

6.3.2 Volume Calculation

Each of the predictions consist of 64 stacked up grayscale images, where each image represents a slice of the whole RBC. The 3D meshes generated in figure 29 can also be represented as matrix with the shape of 64x64x64 pixels, where every pixel can take values of 0 to 255, setting the pixel intensity.

Two methods have been used to calculate the volume. In both methods the scaling coefficient mentioned in the previous chapter has to be adjusted, because the input images were resized, so that it meets the 64x64 pixel requirements, as described in chapter 6.1. To calculate the new pixel to metric ratio, following formula has been applied:

$$new\ pixel\ to\ metric\ ratio\ = \left(\frac{pixel\ to\ metric\ ratio}{scaling factor * 1000}\right)^3 \quad (1)$$

where the new pixel to metric ratio is in μm^3 or fL and the scaling factor is calculated as following:

$$scaling\ factor = \frac{desired\ image\ height}{input\ image\ height} \quad (2)$$

where the desired image height is the 64 pixel height, which is required for the model to make predictions, and the input image height is the actual pixel height of the input image.
The initial pixel to metric ratio of 75nm per pixel has to be devided by the scaling factor, because the resized image is now smaller, and one pixel represents now more nm in the real world. The division by 1000 is to convert the units to μm^3. The quotient of pixel to metric ratio and scaling factor has to be put to the power of three, to put the new pixel to metric ratio in μm^3, respectively in femtoliters, making every pixels to voxels.

For the first method, in each image slice of every prediction, it is evaluated if the pixels are above a certain pixel threshold or not. The program then counts all the pixels above the threshold and multiplies this number with the newly calculated pixel to metric ratio in equation 1, resulting in the volume in femtoliters. The threshold was manually fixed, by experimenting with different values. The value of 175 resulted that not too much of the predicted cell was cut off from the initial cell prediction, but also not too much was included. Lower threshold would result in higher volumes as more pixels would be counted and vice versa as shown in the figure below.

The second volume calculation method, involves using the code provided by *Waibel et. al*. at their GitHub Repository [37], where they first normalize the data in the RBC, putting every pixel in relation to the highest pixel value in the matrix and then apply the otsu-thresholding method, from the scikit-image library [38]. This method creates a binary matrix of the RBC, consisting of only ones and zeros. To calculate the volume, the script then only needs to count the ones in the matrix and multiples the sum with the newly calculated pixel to metric ratio in equation 1, resulting in the volume in femtoliters.

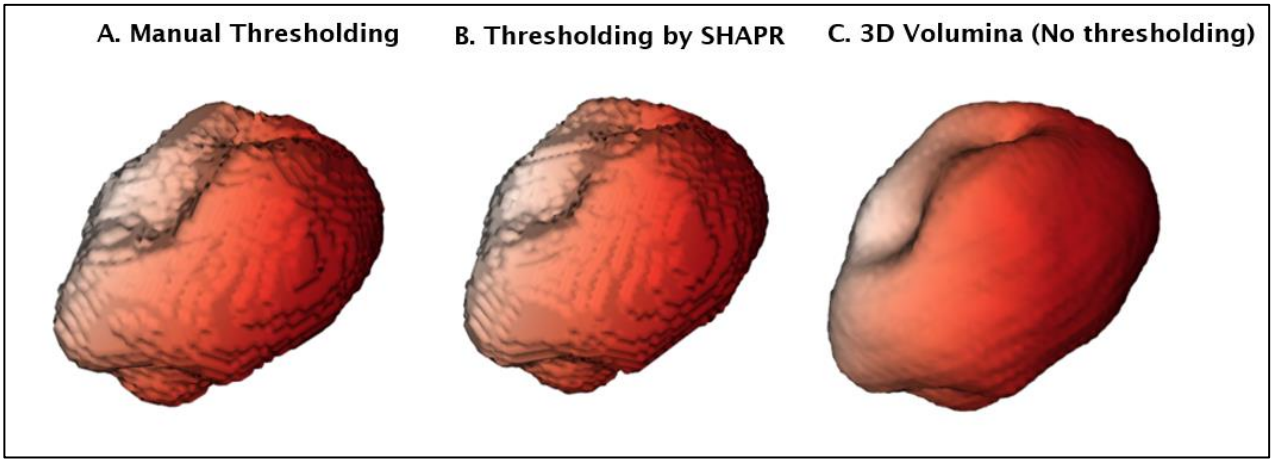


Figure 33 : The 3D visualizations of a single RBC when applying different thresholding methods to calculate the volume. A : Manually thresholding the predicted cell at a threshold of 175 pixel intensity. Resulting in a cell volume of 65fL. B : Thresholding by SHAPR resulting to a volume of 73fL. C : 3D Visualization with all the pixels included. Note that the surface of model C is much smoother than the other two. The assumption is that this is because of the pixels which have lower intesity, to create a smoothed surface.

6.3.3 Histograms training set with all cells

To visualize the calculated volumes with the different methods, two histograms have been created, which show the mean cell volume (MCV) and the standard deviation of the distribution. The results for the fifth model (GAN4), which was trained with all RBC types, show that the difference between the two methods is not very large.

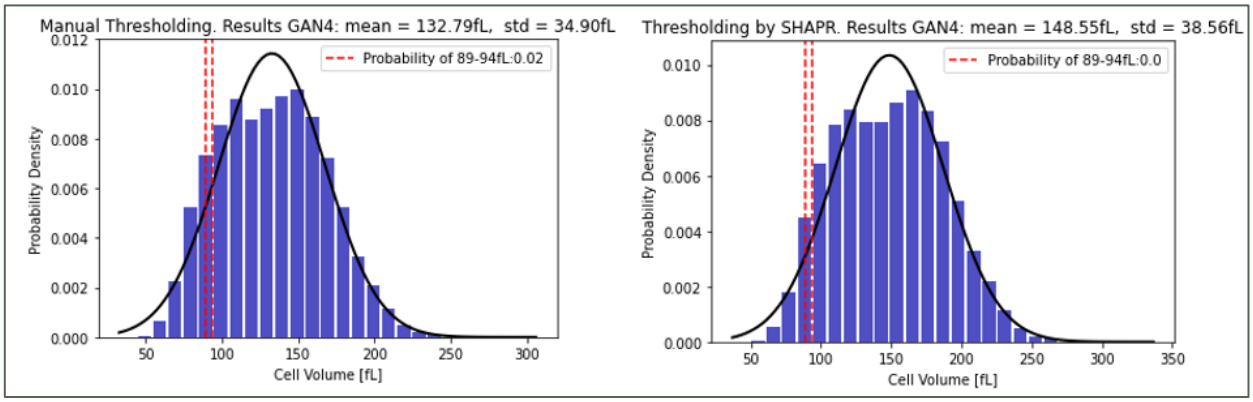


Figure 34: Normal distribution histograms of the prediction where all cells were included in the training set. The histograms include the 11668 single RBC images, showing the mean cell volume and the standard deviation. The values are in fL.

The histograms also show that the mean value is larger than the MCV of 91fL, given by the absorption analysis and the standard deviation is more than double of 11.56fL, described in chapter 6.2.

The probability density function (pdf) doesn't provide probabilities on discrete values, rather it calculates it by integrating over a specific section, in this case over 89 to 94fL, which are normal MCVs for a healthy male. The histograms show the probability, that the prediction lies between 89 and 94fL, is approximately 2%, for the manual thresholding method. For the volume calculation method by SHAPR the probability is too small to be able to calculate it, i.e. smaller than 1%.

To compare the volumes for each RBC, the results from the predictions and the absorption analysis histogram from figure 29 have been normalized, utilizing the formula below:

$$z_i = \frac{x_i - \min(x)}{\max(x) - \min(x)} \quad (3)$$

where z_i is the i^{th} normalized value, x^i the i^{th} value, min(x) the minimum value and max(x) the maximum value in the dataset [39]. This equation puts every volume in relation to the maximum volume in the data, so that numbers from zero to one are created.

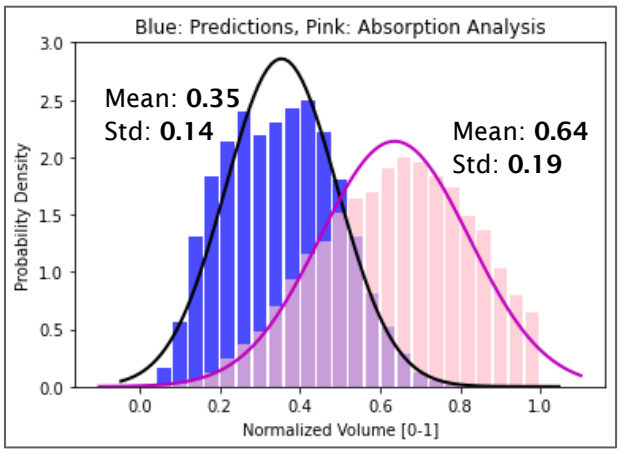


Figure 35 : Overlapped histograms of the normalized prediction data from the manual thresholding method (blue), where all cell types were included in the trainingset, and the normalized absorption analysis data (pink).The mean values are significantly different from each other.

The results show, that the mean values of the predictions are significantly different from the absorption analysis provided by the sitem microscope, although the standard deviations are similar. The reason for the large shift of both means, could be that both initial histogram, with no normalization applied where differntley shaped. The predictions where more shifted to the left, and the absorption results were shifted to the right.

6.3.4 Histograms training set with discocytes only

Also two histograms have been created with the results of the fourth model (GAN3), which was trained with only discocytes in the training set. The histogram also shows here that the difference between the two methods is not very large.

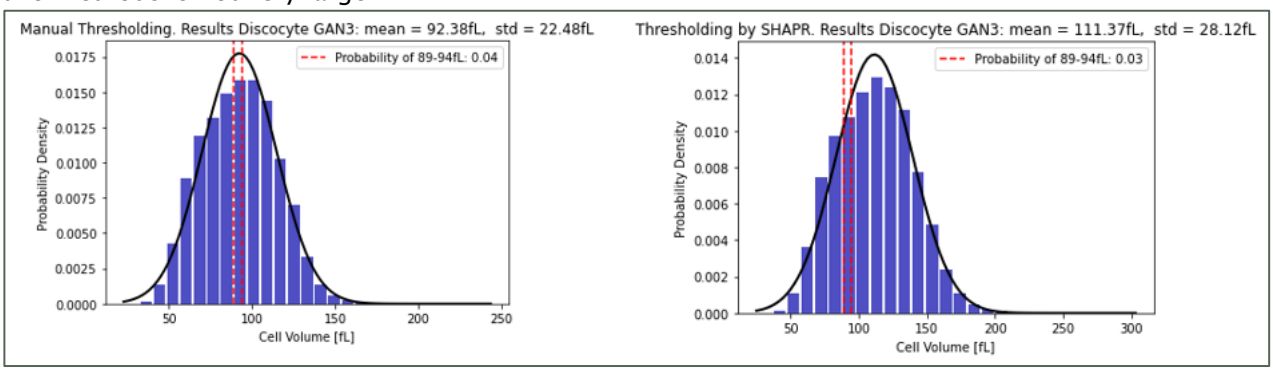


Figure 36 : Normal distribution histograms of the prediction where only discocytes were included in the training set. The histograms include the 11'668 single RBC images, showing the mean cell volume and the standard deviation. The values are in fL.

The predicted volumes of the RBC come very close to the 91fL MCV of the blood sample, especially the manual thresholding method. On the other hand the standard deviation is nearly the double of 11.56fL. The histograms show the probability, that the prediction lies between 89 and 94fL, is approximately 4%, for the manual thresholding method, and 3% for the volume calculation by SHAPR.

To compare the volumes for each RBC, the results from the predictions and the absorption analysis histogram from figure 29 have been normalized, utilizing the equation 3.

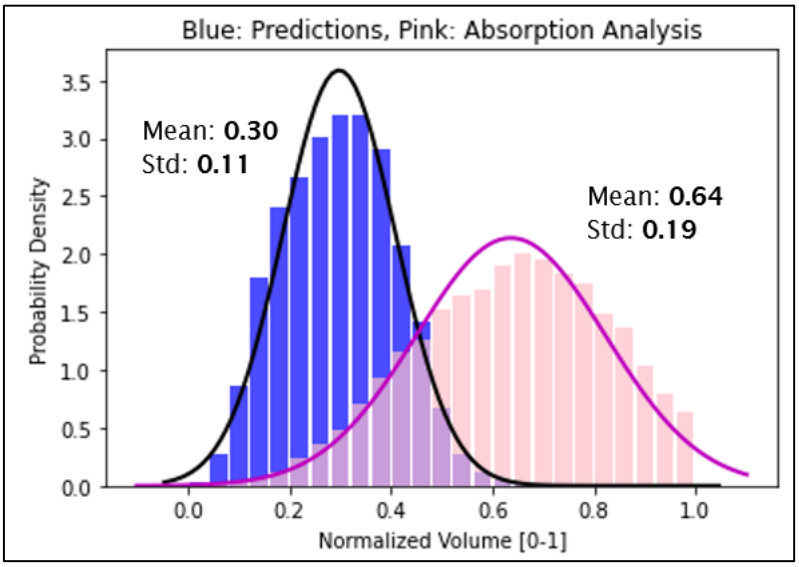


Figure 37: Overlapped histograms of the normalized prediction data from the manual thresholding method (blue), where only discocytes were included in the training set, and the normalized absorption analysis data (pink).The mean values are significantly different from each other.

The results also show here , that the predictions are significantly different from the absorption analysis provided by sitem. Also the standard deviation is a bit more off.

6.4 Limiting and causing factors of results

Nature of Generative Adversarial Networks

As the discriminator tries to distinguish the 3D RBC shapes between prediction and real data, the predictor, i.e. the generator, tries to fool the discriminator by generating shapes that look similar to the training set. This way, the assumption is there that the model is overfitted and can't generalize to new data, that are different from the training set.

To see if this assumption is true, the predictions with the sitem RBC images are compared with the predictions which were generated during the training and validation of the model, using the same images from the public dataset.

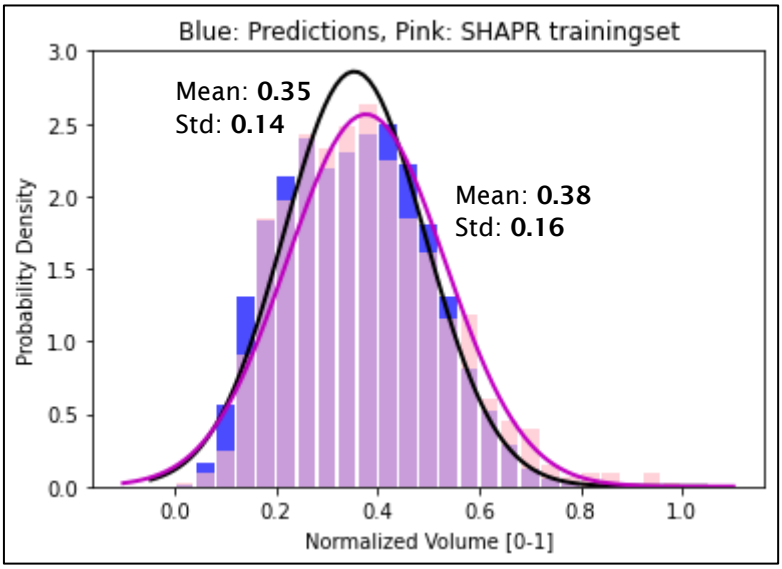


Figure 38: Overlapped histograms of the normalized prediction data from the manual thresholding method (blue) with the GAN4 model, where all cell types were included in the training set, and the normalized data from the predictions which were generated during the training and validation of the SHAPR models (pink).The mean values are very similar.

The results show that the predictions with the sitem images are very similar to the predictions from the training and validation of the models, done by *Waibel et. al.* in their published paper. Therefore it can be said that the model can't generalize to the new images provided by sitem.

Image Quality

The figure below shows the difference of the image quality between the images which were used for training and the images used for prediction. The line plots on the right side show the pixel values along the white line of the grayscale image. It can be seen that the training image has much more contrast in where the RBC is located. On the line plots the pallor in the middle of the cell is also better visible for the training images. An other difference is the fact that the training images are taken with an confocal microscope, which doesn't let in the light which is out of focus. This results that everything around the RBC is black.

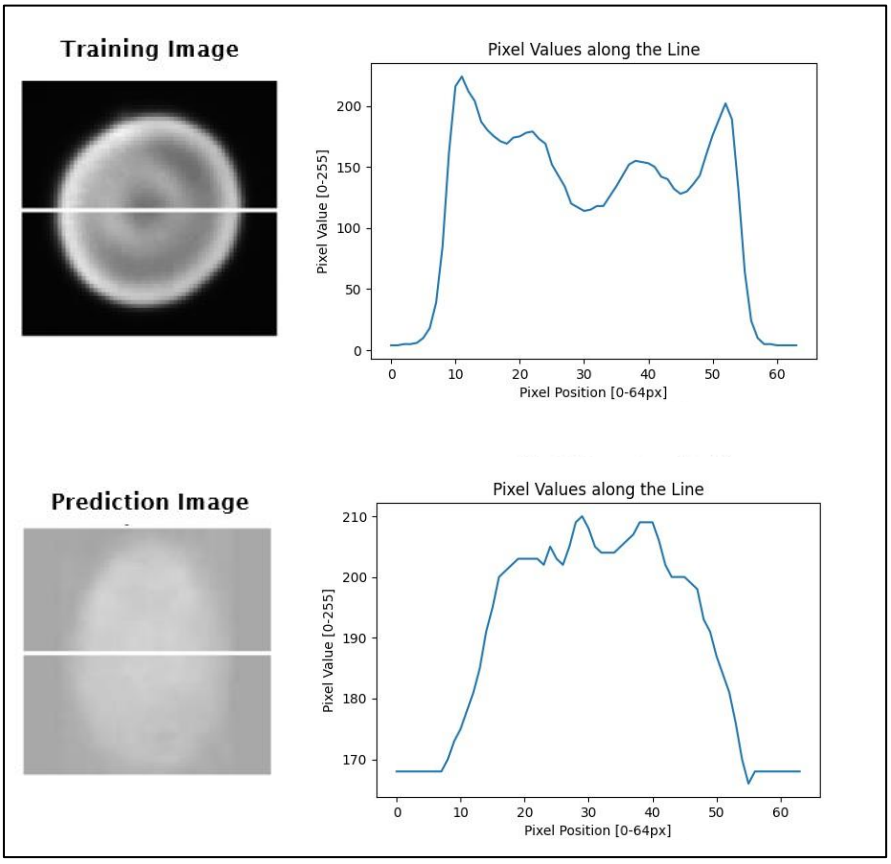


Figure 39 : Top : Example of an greyscale image with an discocyte, which was used for training the models together with the pixel values across the white line. Bottom : Example of an greyscaled image which was used to perform the predictions, which resulted to the 3D volumes, shown in figure 30 and figure 32. It can be seen that the typical pallor, which is caused due to the flattened middle of a biconcave discocyte, is better visible on the training image.

Other Factors
A factor which may be considered, is the fact that the images, used for the training, are always the center slice of the whole RBC [22]. This means that the models assumes, when the grayscale image and the binary segmentation are given as an input, that those images are the center slice of the whole cell. This is not true, because the images provided by sitem and used for the predictions are taken with a conventional microscope and only show the top-down view of the RBC.

For typical discocytes it may not play a crucial role, because usually the center slice is also mostly the widest part of the RBC, resulting that the segmentation equals the center slice. But for other RBC types, it can be possible that the RBC has a shape similar to a mushroom where the center slice is much smaller than the top-view image from the microscope. Another possibility is that the RBC takes an unusual position, where the cell is rotated away from the microscope camera, which lead to that the RBC is depicted wider or on the images wider than it actually is assumes that the RBC is wider or shorter than it actually is.

6.5 Created Python Scripts

For this Bachelor Thesis several scripts have been created for each different purpose and task. The files can be accessed via the public GitHub repository, which has further explanation of the folder structure and how the files should be utilized:
https://github.com/DominiquePeytrignet/BachelorThesis_DPeytrignet

The notebook "*BTHE_Extracting_Red_Channel*" can be used when the images, containing a single RBC, isn't put into grayscale mode. The script then takes the images, extracts only the red channel and stores them as grayscale images.

With the help of the notebook "*BTHE_Altering_Images*" it is possible to alter an image so that it meets the requirements for the SHAPR Framework and perform predictions.

To predict the shape of a RBC the notebook "*BTHE_Applying_SHAPR*" can be used. It loads the SHAPR framework and the pretrained models to predict the shapes.

To calculate the volume of the predicted shapes, two methods were presented, as described previously. For each method, one notebook was written: "*BTHE_Volume_Calculation_bySHAPR*" and "*BTHE_Volume_Calculation_manualThresh*". Both these script calculate the volume in femto liters and store the results in an excel file.

The histograms shown previously to compare the results, were created with the scripts "*Normalization_Comparison*", "*histogram_absorption_data*" and "*histogram_prediction_data*".

To have a 360° view of the predicted RBC shape, the script "BTHE_3D_Visualization" can be used.

Finally the python script "BTHE_Image_quality" is written to create the line plots, which show the pixel values the white horizontal line that crosses the image.

7 Discussion

7.1 Answering the research questions

- What are current methods to create a 3D model out of images?
There are multiple solutions for the problem of reconstructing three-dimensional objects based from two-dimensional image information.

Pixel2Mesh by *Wang et. al.* creates, with the help of deep neural networks, 3D triangular meshes from 2D RGB-D images, for example from chairs, tables or planes.

Graikos et. al. utilize a molecular depth estimation network to segment food portions on a plate and the volume, i.e. the calories, of it. This networks, coulnd't be taken into consideration for the RBC images, because no training data was available for the depth estimation network, and it required multiple frames of the object of interest.

Cobo et al. trained a neural network to estimate the amount of wine that is in a glass, based on a single image. This networks also failed, for the application of the RBC images, because it couldn't generate 3D models.

All these methods can be used to either predict the shape or estimating the volume of an object. They were trained on datasets with over 10'000 models from the object of desire, hence with ground true data. For RBC there currently don't exist that much data.

For estimating the volume of a single RBC there exists the quantitative absorption imaging method, presented in chapter 5, and deeper explained by *Paul et. al.* in their paper [33].

For this Bachelor Thesis the SHAPR Framework, explained in chapter 4, was applied to the microscopic images provided by sitem, to predict the shape and volume of the RBCs.

-How accurate is the method applied to the microscopic images?

The results show that the predictions from the model, which was trained purely on discocytes, performed better than the model which was trained on all RBC types, because the calculated mean cell volume of 92.38fL, is closer to the real mean cell volume of 91fL, both with the manual thresholding calculation method and the method used by SHAPR. Also the probability that the prediction lies between 89 and 94fL is higher, but still small, given the fact that the patient was healthy. Further, the predicted shapes resemble much more to a typical discocyte, although it is not determinable if the cell in the image actually looks like this.

Also when considering the results from all models, which can be found in the appendix, the predictions of the mean volumes from the models, which were trained solely on discocytes, range from 61fL to 92fL. Compared to a mean volume range from 83fL to 211fL, from the models, which were trained on all RBC types, which is a way better performance.

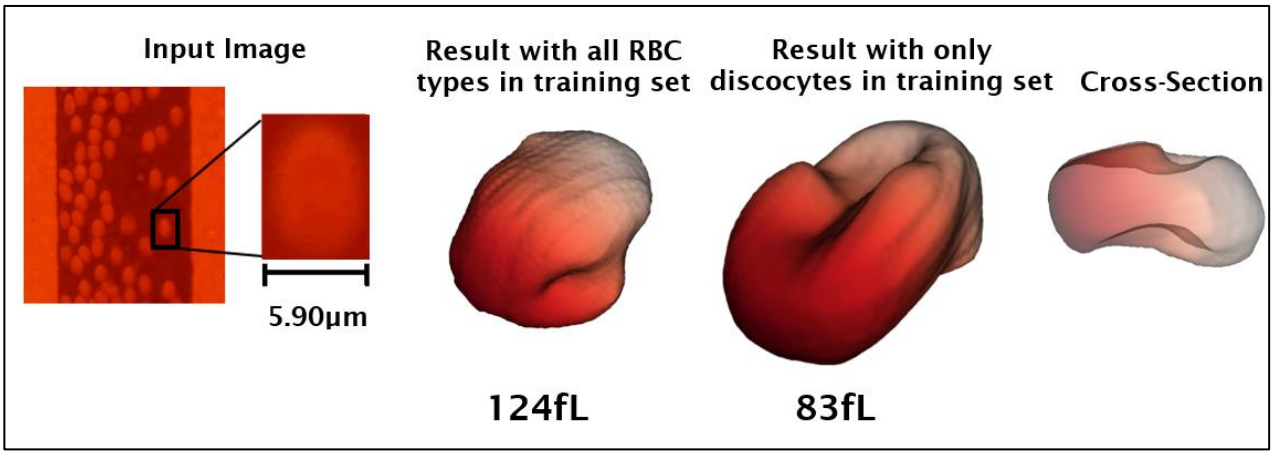


Figure 40: 2 3D voluminas from the same input image. The first shape, with a volume of 124fL was predicted by a model which was trained on a dataset, containing all RBC types. The second shape, with a volume of 83fL, was predicted by a model which was trained on a dataset, containing only discocytes. Not only is the predicted volume in the second shape closer to the real mean volume, but also the shape is accurate to a discocyte, with thick edges and a flattend middle, as seen In the cross-section.

7.2 Retrospective of goals

The goal for this Bachelor thesis was to present a method that can estimate, i.e. predict the shape and the volume of a red blood cell, based on a single 2D image. As a result a Python script should be created, which applies the presented method to predict the shape of the RBC based on the images. These results should then be visualized and important metrics can then be read out. Also a comparison between the image absorption method and the selected method should be made.

These goals were achieved, as in chapter 3 various methods were presented for estimating the volume of an object and generating a 3D Model from a single 2D image. However, these approaches were solely tested on regular daily objects such as chairs, tables or food and big datasets were available for training the neural networks. Additionally, RBC images are fundamentally different from these objects and there aren't much datasets available with the ground truth data of the cells. Therefore the SHAPR framework, developed by *Waibel et. al.* was presented in chapter 4 to be applied on the RBC images, provided by sitem, to predict their shape and volume.

In chapter 6, with the help of histograms, the predicted volumes were compared with the results from the absorption imaging method from the single images and the mean cell volume of the whole blood sample.

For the visualization a Python script was created which allows the user to have a 360° view of the whole predicted RBC shape, see GitHub Repository.

7.3 Potential

Additional Metrics

The SHAPR Framework provides further functions on how other metrics can be read out, such as the roughness, convexity or surface of the RBC. They were not included in this work, because the main focus was on the volume of the RBC, and there were no ground truth data of these metrics to validate the predictions. Including such metrics, however, could yield additional important insights.

Additionally, SHAPR provides a Python script, which can classify the predictions on the different RBC types, which were used for the training. With such a tool, diagnosis of diseases and illnesses could be improved significantly.

Diffusion Models

During the progress of this Bachelor Theses and further research, the paper from Waibel et. al. [40] was discovered, where they present a diffusion model, DISPR (Diffusion based shape prediction), which fulfills the same task as SHAPR, predicting RBC shapes based on single 2D images. Diffusion models are a type of generative models different to GANs. *Nichol and Dhariwal* even say in their paper [41], that diffusion models produce better image synthesis than GANs. So it would be interesting to see, whether the performance of the shape predictions can be optimized with this alternative approach.

Digital Twin

With the presented method a first step may have been taken to create some sort of digital twin of a single red blood cell. What with this work was created, is a single snapshot of the RBC flowing through a microfluidic chip channel. However, if one would want to combine the shape predictor with a highspeed camera, constantly tracking the RBC flowing through the channels, this would result in a digital twin, which constantly shows the shape of the RBC together with important metrics. This would be one step closer to the vision of a digital human twin from Benedikt von Thuengen, which enables early detection of diseases. [42]

"A crucial step is to move towards detecting diseases as early as possible – by building a digital human twin. This needs a wide range of sensors to capture a variety of data points such as digital (sound, image, etc.), biophysical (blood, saliva, etc.) and behavioral (emotional state, language, etc.). These data points then need to be aggregated into a dynamic digital human twin that can monitor your health and alert the healthcare system as soon as something changes." – Benedikt von Thuengen

7.4 Critical Points

It is important to mention that the threshold value, which is used to determine to what intensity (0 to 255) threshold the pixel should be included in the volume calculation or not, was determined manually, by testing multiple values. Therefore it is possible that a certain bias can be included in the volume calculation so that the results meets the expectation of the user. The value of 175 was defined as a good value from the author, as it is also the pixel value which was used to create the binary segmentation from the grayscale image. To create higher volumes the threshold could simply be lowered, or increased for higher volumes, although the author of this work strongly advices not to do this.

Another point, which is important to mention, is that the author assumes the pixel to metric ratio, i.e. how much one pixel equals in real world dimensions, stays the same, both for the generated 64 images and the resized grayscale and binary segmentation images. For the input images one pixel equals to 75.5nm. This ratio is increased, due to the fact that images were scaled down to 64x64 images and therefore one pixels equals more nanometers in the real world. A problem now might be that, because the model generates prediction that resemble the RBCs from the training set, the ratio is changed to the one used for taking the confocal images. This pixel to metric ratio is unknown for the author and therefore no comparison can be made between this possibility and the results from this Bachelor Thesis.

As a last point, it has to be taken into consideration, that the results which are shown in this Bachelor Thesis only come from one single patient. That means, to further validate the accuracy of the presented method, the same measurements would have to be carried out on additional patients.

8 Conclusion , recommandations, future works

In this Bachelor Thesis the SHAPR-Framework was presented, as a time and cost saving alternative, to reconstruct 3D shapes of RBCs and estimate their volumes. To evaluate the accuracy of the presented method, the shape and volume of over 10'000 single RBC have been predicted, based on 2D microscopic images, and compared with the real values provided by sitem. For the prediction two training sets were created, one which contains all RBC types, including unhealthy, and one which contains only discocytes.

The results show that for the training set, which only includes discocytes, the predicted volumes come very close to the real mean volume of 91fL, provided by sitem. The probability that the prediction lies between 89fL and 94fL is 4%, which is still small, when considering that the patient was a healthy male. Also the shapes resemble perfectly a typical disked shape discocyte. However in some predictions there are little spikes, which are fully separated from the actual cell. An algorithm which detects these spikes in the shape, which are not attached to the bigger cell, could easily recalculate the correct volume.

The author recommends therefore, to implement such an algorithm and verify if the predicted volume is significantly impacted or not. Furthermore, it is recommended to also evaluate the shortly presented DISPR-Framework, which tries to achieve the same task as SHAPR, but with a different, presumably better approach, namely a diffusion model. These two frameworks should therefore be compared on their accuracy and see if it is possible to increase the probability of a prediction to be between 89fL and 94fL.

Because of the fact that the results in this work are generated with solely one single patient, it is also recommended to perform the same measurements with additional blood samples from different patients and see if the results are still accurate.

As a last recommendation, to remove the overfitting of the model, it should be evaluated whether the training set should be expanded with additional ground truth data. For this, ground truth data is needed and therefore a confocal microscope, which can create stacks of 2D images to generate the 3D volume.

For future works, the presented method, on how the 3D shape of a RBC can be reconstructed from 2D images, could be integrated in to a point-of-care device, for home diagnoses of patients. One such device is currently developed at the Institute for Human Centered Engineering (HUCE), which allows nursing staff to perform blood tests at the patient's home and stay mobile.

Figure 41: Prototype of point-of-care device, currently developed at the HUCE. Source: BFH

9 Contents of figures

10 Bibliography

[1] "sitem-insel." https://sitem-insel.ch/de/ (accessed Apr. 13, 2023).
[2] "Capillaries: Function, Anatomy, Related Conditions," *Cleveland Clinic*. https://my.clevelandclinic.org/health/body/21988-capillaries (accessed Apr. 12, 2023).
[3] "Anaemia." https://www.who.int/health-topics/anaemia (accessed Apr. 12, 2023).
[4] J. F. Hoffman, "Biconcave shape of human red-blood-cell ghosts relies on density differences between the rim and dimple of the ghost's plasma membrane," *Proc. Natl. Acad. Sci.*, vol. 113, no. 51, pp. 14847–14851, Dec. 2016, doi: 10.1073/pnas.1615452113.
[5] "Red blood cell shape transitions and dynamics in time-dependent capillary flows | Elsevier Enhanced Reader." https://reader.elsevier.com/reader/sd/pii/S000634952103900X?token=87F64F02ED2DF9DABFC0C1576931262DB40CF78B6251EB5DD881BCACEF82C6BED1DE95D658665FFC6E444B99A768A03C&originRegion=eu-west-1&originCreation=20230412092402 (accessed Apr. 12, 2023).
[6] M. To and V. Villatoro, "Spherocytes", Accessed: May 17, 2023. [Online]. Available: https://openeducationalberta.ca/mlsci/chapter/abnormal-rbc-morphology-spherocyte/
[7] "Sickle Cell Disease," Nov. 19, 2019. https://www.hopkinsmedicine.org/health/conditions-and-diseases/sickle-cell-disease (accessed Apr. 13, 2023).
[8] "High Red Blood Cell Count: Symptoms, Meaning, Causes," *Cleveland Clinic*. https://my.clevelandclinic.org/health/symptoms/17810-high-red-blood-cell-count (accessed Apr. 14, 2023).
[9] J. Y. Lee *et al.*, "Age-related changes in mean corpuscular volumes in patients without anaemia: An analysis of large-volume data from a single institute," *J. Cell. Mol. Med.*, vol. 26, no. 12, pp. 3548–3556, 2022, doi: 10.1111/jcmm.17397.
[10] "RDW Blood Test: What It Is, Procedure & Results," *Cleveland Clinic*. https://my.clevelandclinic.org/health/diagnostics/22980-rdw-blood-test (accessed May 26, 2023).
[11] N. Li, H. Zhou, and Q. Tang, "Red Blood Cell Distribution Width: A Novel Predictive Indicator for Cardiovascular and Cerebrovascular Diseases," *Dis. Markers*, vol. 2017, p. 7089493, 2017, doi: 10.1155/2017/7089493.
[12] "Hemoglobin Information | Mount Sinai - New York," *Mount Sinai Health System*. https://www.mountsinai.org/health-library/tests/hemoglobin (accessed Apr. 14, 2023).
[13] G. Di Caprio, C. Stokes, J. M. Higgins, and E. Schonbrun, "Single-cell measurement of red blood cell oxygen affinity," *Proc. Natl. Acad. Sci.*, vol. 112, no. 32, pp. 9984–9989, Aug. 2015, doi: 10.1073/pnas.1509252112.
[14] "Hematology Analyzers—From Complete Blood Counts to Cell Morphology | Labcompare.com." https://www.labcompare.com/10-Featured-Articles/162042-Hematology-Analyzers-From-Complete-Blood-Counts-to-Cell-Morphology/ (accessed Apr. 13, 2023).
[15] "Confocal Microscopes - Institute for Molecular Bioscience - University of Queensland." https://imb.uq.edu.au/facilities/microscopy/hardware-software/confocal-microscopes (accessed Apr. 21, 2023).
[16] "Scanning Electron Microscope - Radiological and Environmental Management - Purdue University." https://www.purdue.edu/ehps/rem/laboratory/equipment%20safety/Research%20Equipment/sem.html (accessed May 24, 2023).
[17] "Artificial intelligence to estimate wine volume from single-view images | Elsevier Enhanced Reader." https://reader.elsevier.com/reader/sd/pii/S240584402201845X?token=23012F6EABC3FDC65B414A479FB632ADA2DD1F5517A3FDCE59FD67E5AAF6CC2B3882BDE55EA6C28C761A76BA8AE0EDE4&originRegion=eu-west-1&originCreation=20230319164442 (accessed Mar. 19, 2023).
[18] "Graikos et al. - 2020 - Single Image-Based Food Volume Estimation Using Mo.pdf." Accessed: Mar. 19, 2023. [Online]. Available: https://protein-h2020.eu/wp-content/uploads/2020/09/Single-Image-Based-Food-Volume-Estimation-Using-Monocular-Depth-Prediction-Networks.pdf
[19] N. Wang, Y. Zhang, Z. Li, Y. Fu, W. Liu, and Y.-G. Jiang, "Pixel2Mesh: Generating 3D Mesh Models from Single RGB Images," in *Computer Vision – ECCV 2018*, V. Ferrari, M. Hebert, C. Sminchisescu, and Y. Weiss, Eds., in Lecture Notes in Computer Science, vol. 11215. Cham: Springer International Publishing, 2018, pp. 55–71. doi: 10.1007/978-3-030-01252-6_4.
[20] A. X. Chang *et al.*, "ShapeNet: An Information-Rich 3D Model Repository." arXiv, Dec. 09, 2015. Accessed: Apr. 19, 2023. [Online]. Available: http://arxiv.org/abs/1512.03012

[21] X. Sun *et al.*, "Pix3D: Dataset and Methods for Single-Image 3D Shape Modeling," in *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Jun. 2018, pp. 2974–2983. doi: 10.1109/CVPR.2018.00314.
[22] D. J. E. Waibel, N. Kiermeyer, S. Atwell, A. Sadafi, M. Meier, and C. Marr, "SHAPR predicts 3D cell shapes from 2D microscopic images," 2021.
[23] I. H. Sarker, "Machine Learning: Algorithms, Real-World Applications and Research Directions," *SN Comput. Sci.*, vol. 2, no. 3, p. 160, Mar. 2021, doi: 10.1007/s42979-021-00592-x.
[24] M. Taye, "Theoretical Understanding of Convolutional Neural Network: Concepts, Architectures, Applications, Future Directions," *Computation*, vol. 11, p. 52, Mar. 2023, doi: 10.3390/computation11030052.
[25] A. Kumar, "Transposed Convolution vs Convolution Layer: Examples," *Data Analytics*, Mar. 30, 2023. https://vitalflux.com/transposed-convolution-vs-convolution-layer-examples/ (accessed Jun. 09, 2023).
[26] M. Yani, S. Irawan, and C. Setianingsih, "Application of Transfer Learning Using Convolutional Neural Network Method for Early Detection of Terry's Nail," *J. Phys. Conf. Ser.*, vol. 1201, p. 012052, May 2019, doi: 10.1088/1742-6596/1201/1/012052.
[27] R. Yamashita, M. Nishio, R. K. G. Do, and K. Togashi, "Convolutional neural networks: an overview and application in radiology," *Insights Imaging*, vol. 9, no. 4, Art. no. 4, Aug. 2018, doi: 10.1007/s13244-018-0639-9.
[28] P. Salehi, A. Chalechale, and M. Taghizadeh, "Generative Adversarial Networks (GANs): An Overview of Theoretical Model, Evaluation Metrics, and Recent Developments".
[29] H. Alqahtani, M. Kavakli, and Dr. G. Kumar Ahuja, "Applications of Generative Adversarial Networks (GANs): An Updated Review," *Arch. Comput. Methods Eng.*, vol. 28, Dec. 2019, doi: 10.1007/s11831-019-09388-y.
[30] S. Reed, Z. Akata, X. Yan, L. Logeswaran, B. Schiele, and H. Lee, "Generative Adversarial Text to Image Synthesis." arXiv, Jun. 05, 2016. Accessed: May 01, 2023. [Online]. Available: http://arxiv.org/abs/1605.05396
[31] J. Yu, Z. Lin, J. Yang, X. Shen, X. Lu, and T. S. Huang, "Generative Image Inpainting with Contextual Attention." arXiv, Mar. 21, 2018. Accessed: May 01, 2023. [Online]. Available: http://arxiv.org/abs/1801.07892
[32] E. Schonbrun, R. Malka, G. Di Caprio, D. Schaak, and J. M. Higgins, "Quantitative Absorption Cytometry for Measuring Red Blood Cell Hemoglobin Mass and Volume," *Cytom. Part J. Int. Soc. Anal. Cytol.*, vol. 85, no. 4, pp. 332–338, Apr. 2014, doi: 10.1002/cyto.a.22450.
[33] "Paul et al. - 2020 - Quantitative absorption imaging of red blood cells.pdf." Accessed: Apr. 07, 2023. [Online]. Available: https://pubs.rsc.org/en/content/articlepdf/2020/ra/d0ra05421f
[34] D. Waibel, N. Kiermeyer, S. Atwell, B. Rieck, M. Meier, and C. Marr, "Datasets for 3D shape reconstruction from 2D microscopy images." Zenodo, Sep. 01, 2022. doi: 10.5281/zenodo.7031924.
[35] "OpenCV: Image Thresholding." https://docs.opencv.org/4.x/d7/d4d/tutorial_py_thresholding.html (accessed May 10, 2023).
[36] "Hereditary Spherocytosis: What It Is, Symptoms, Causes & Treatment," *Cleveland Clinic*. https://my.clevelandclinic.org/health/diseases/23058-hereditary-spherocytosis (accessed May 26, 2023).
[37] "SHAPR." Marr Lab, May 16, 2023. Accessed: May 31, 2023. [Online]. Available: https://github.com/marrlab/SHAPR_torch
[38] "Thresholding — skimage v0.20.0 docs." https://scikit-image.org/docs/stable/auto_examples/segmentation/plot_thresholding.html (accessed May 31, 2023).
[39] Zach, "How to Normalize Data Between 0 and 1," *Statology*, Apr. 26, 2021. https://www.statology.org/normalize-data-between-0-and-1/ (accessed Jun. 01, 2023).
[40] D. J. E. Waibel, E. Röell, B. Rieck, R. Giryes, and C. Marr, "A Diffusion Model Predicts 3D Shapes from 2D Microscopy Images." arXiv, Mar. 14, 2023. Accessed: Jun. 13, 2023. [Online]. Available: http://arxiv.org/abs/2208.14125
[41] P. Dhariwal and A. Nichol, "Diffusion Models Beat GANs on Image Synthesis." arXiv, Jun. 01, 2021. Accessed: Jun. 13, 2023. [Online]. Available: http://arxiv.org/abs/2105.05233
[42] "The Dawn of a New Era of Digital Twin Diagnostics," *Diagnostics from Technology Networks*. http://www.technologynetworks.com/diagnostics/articles/the-dawn-of-a-new-era-of-digital-twin-diagnostics-360148 (accessed Jun. 13, 2023).

11 Declaration of authorship

I hereby certify that I composed this work completely unaided, and without the use of any other sources or resources other than those specified in the bibliography. All text sections not of my authorship are cited as quotations and accompanied by an exact reference to their origin.

Bern, 28.06.23

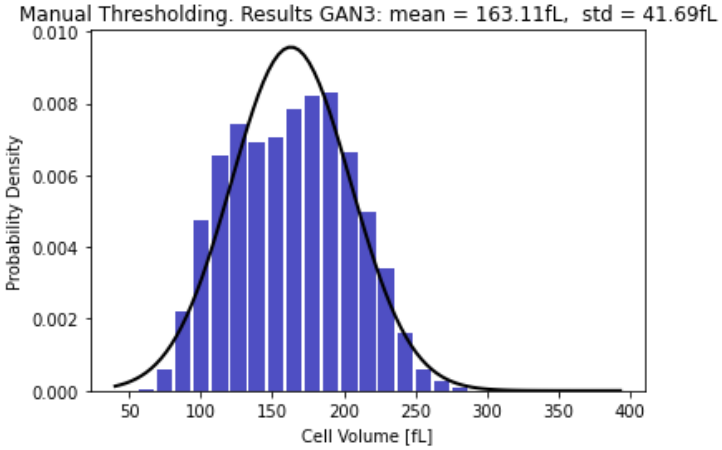
_____
City, Date Dominique Peytrignet

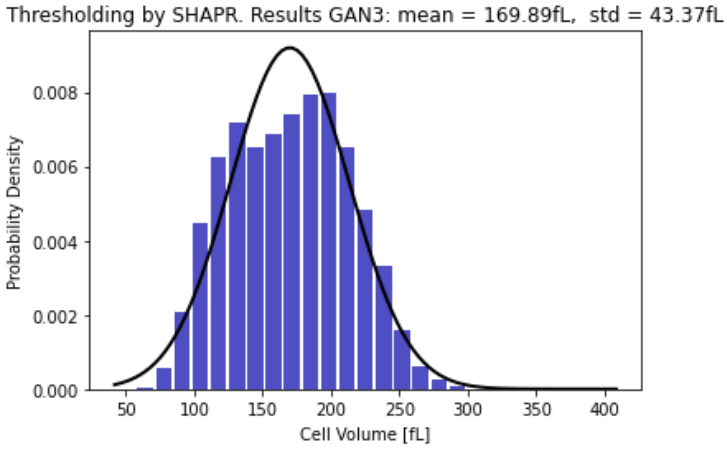
12 Version control

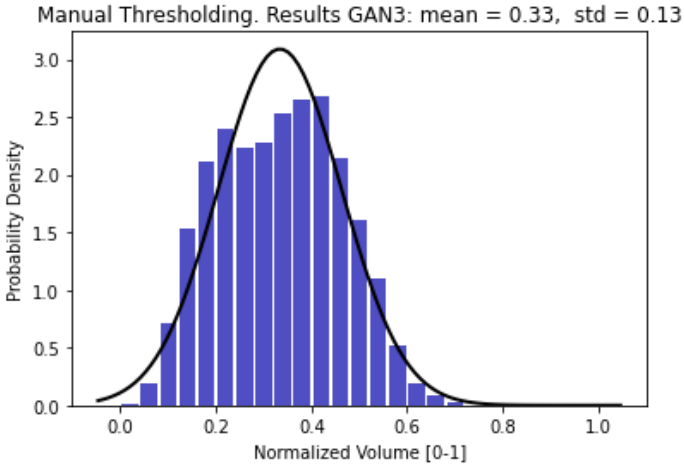
Version	Date	Description	Author
0.1	03.04.2022	Document created	Dominique Peytrignet
0.2	28.06.2023	Document finished	Dominique Peytrignet

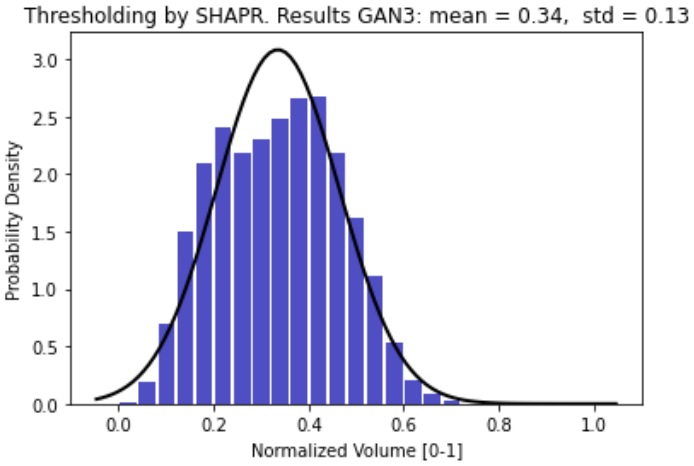
13 Appendix

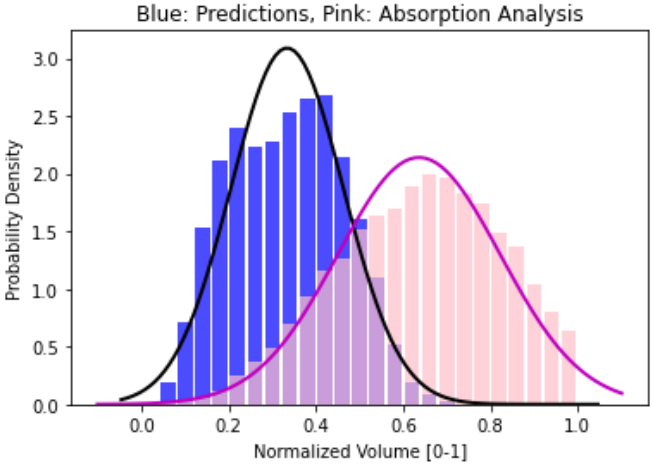
13.1 Results GAN3 All RBC Types

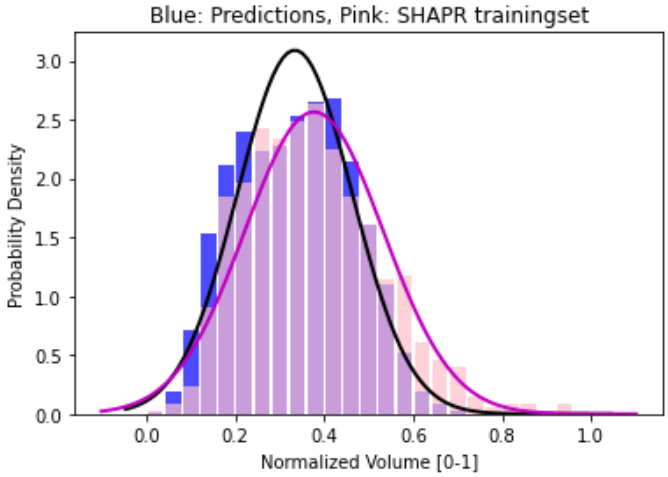




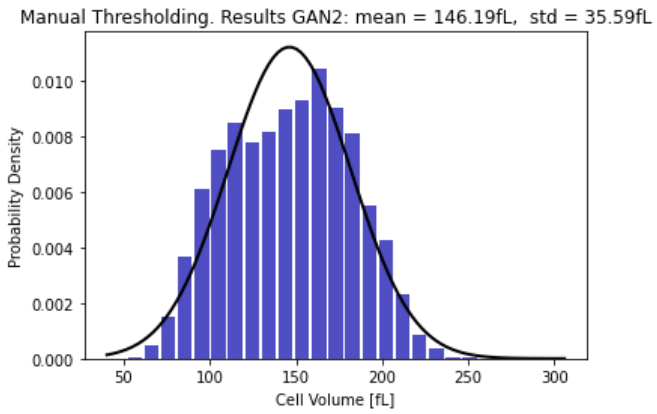


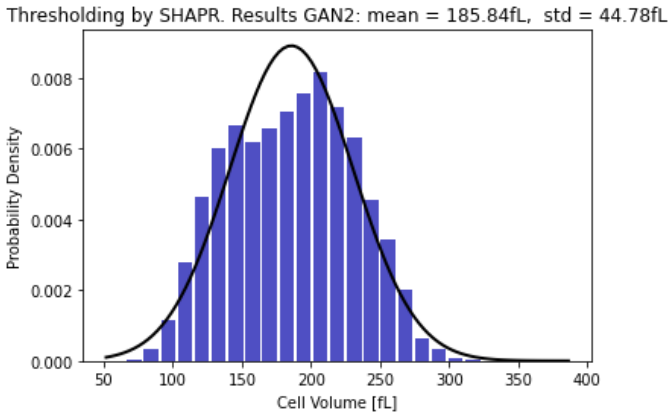


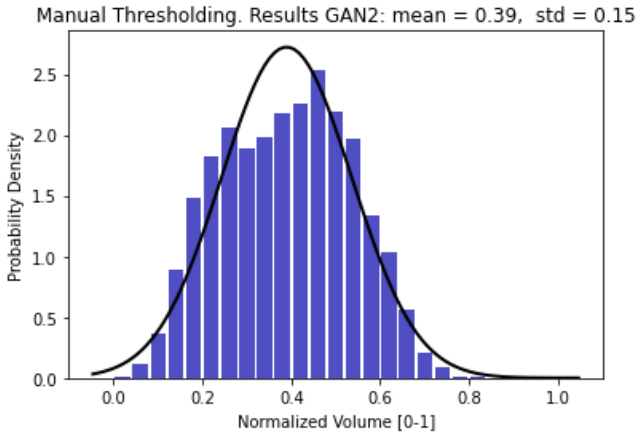


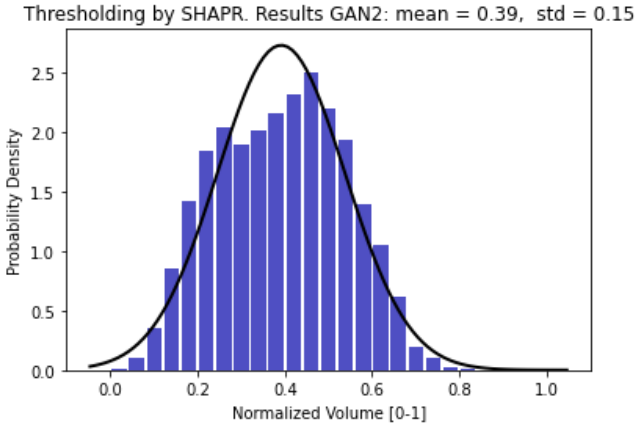


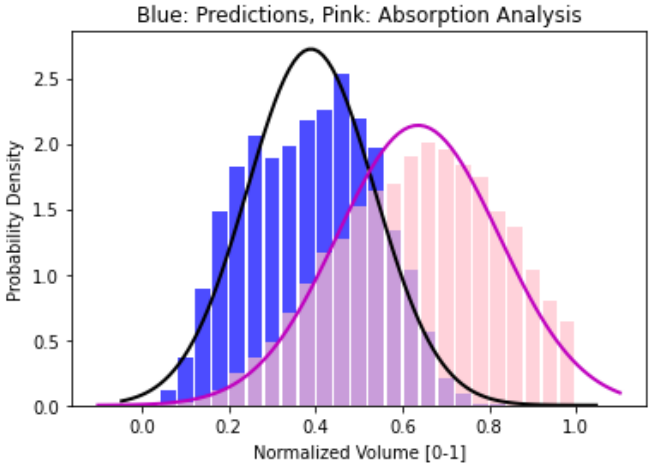
13.2 Results GAN2 All RBC Types

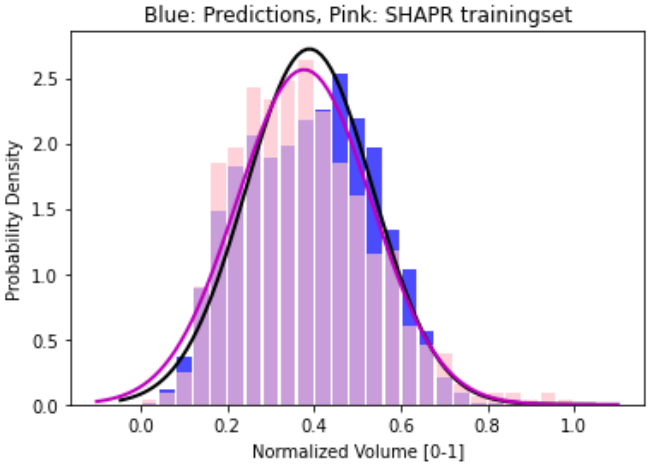




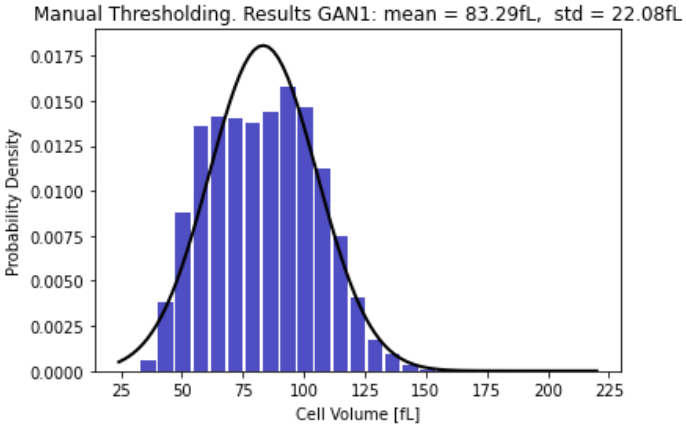


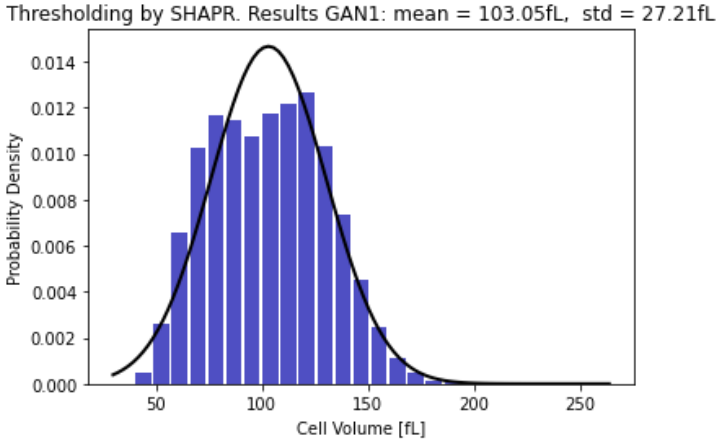


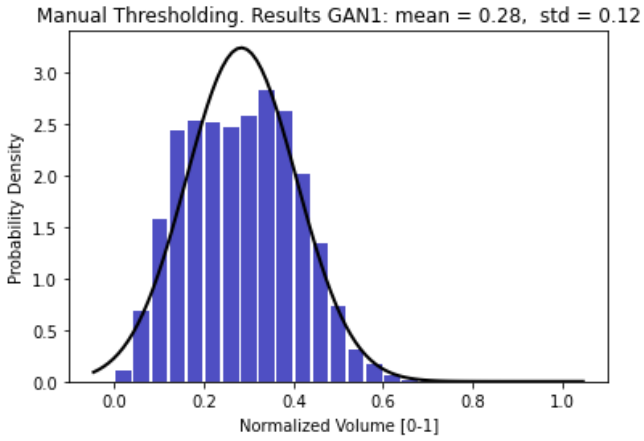


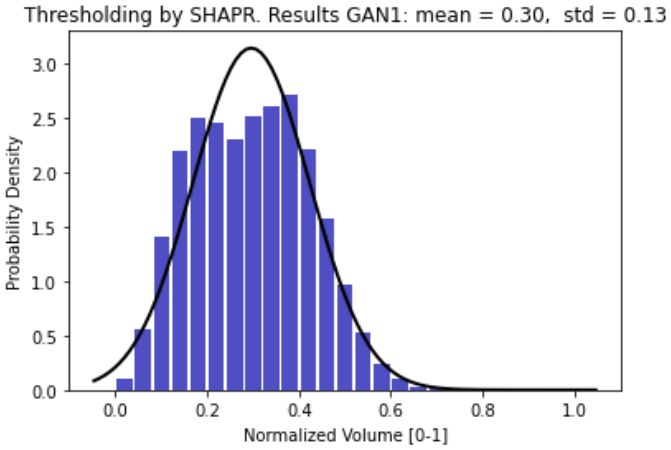


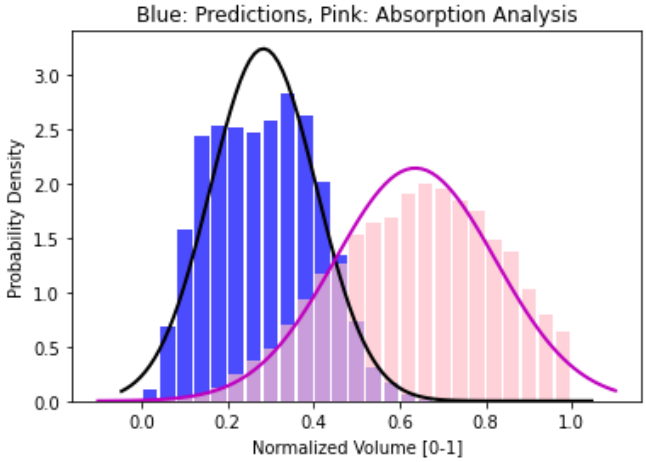
13.3 Results GAN1 ALL RBC Types

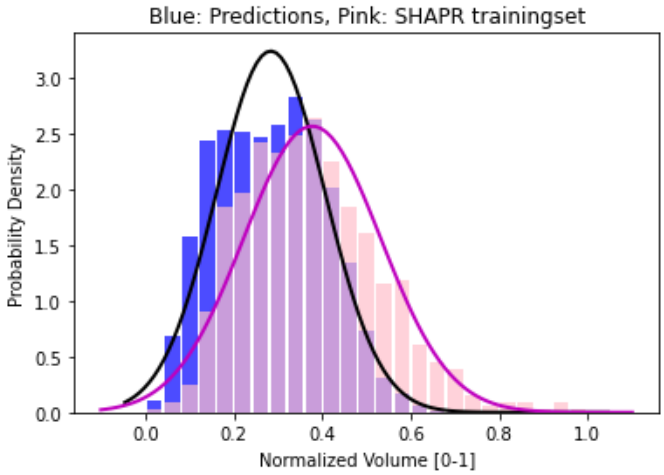




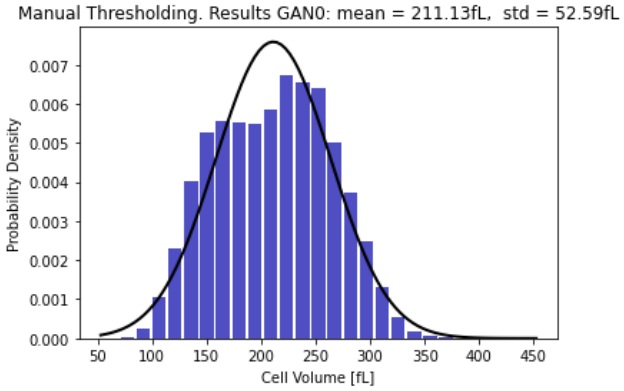


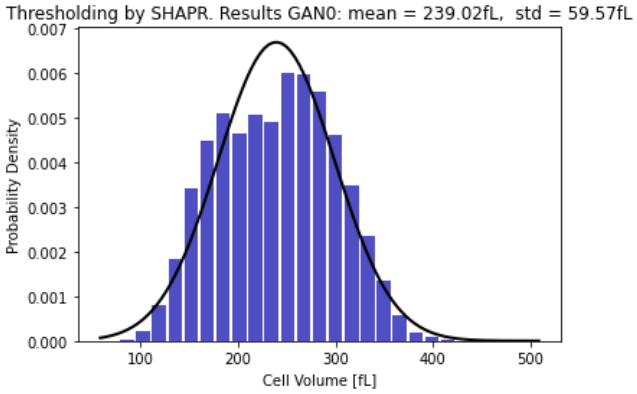


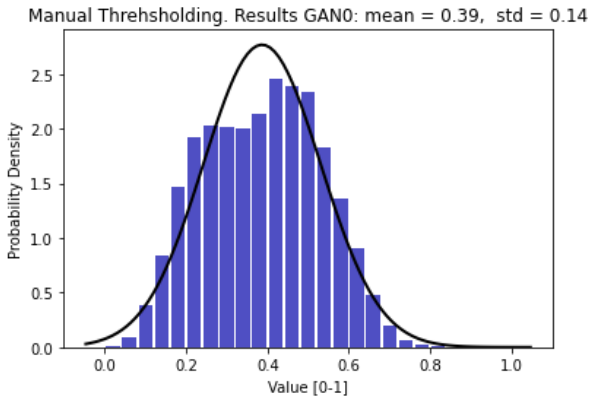


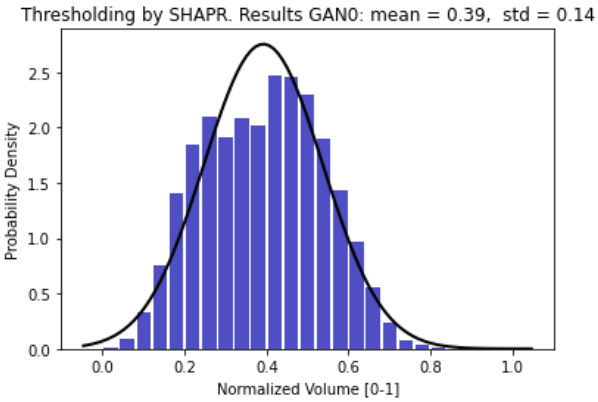


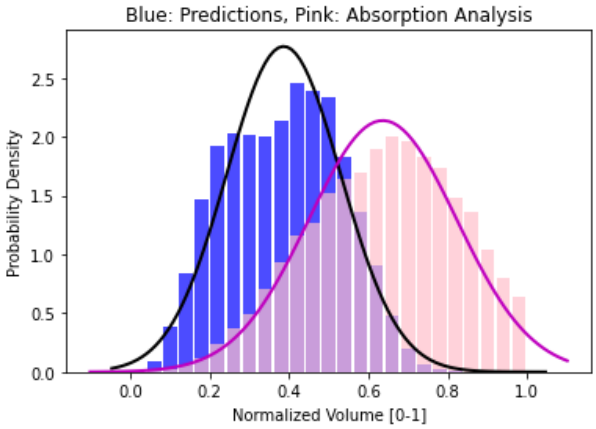
13.4 Results GAN0 ALL RBC Types

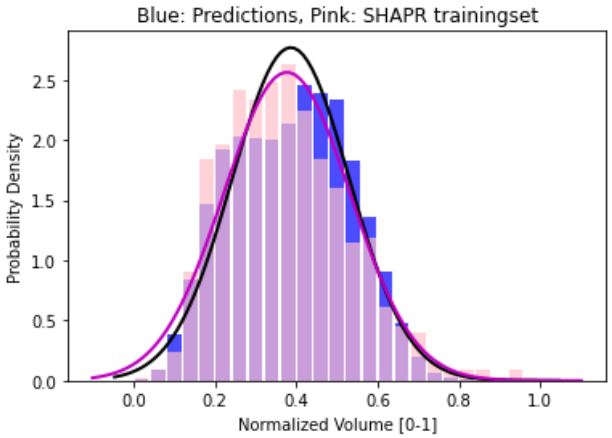






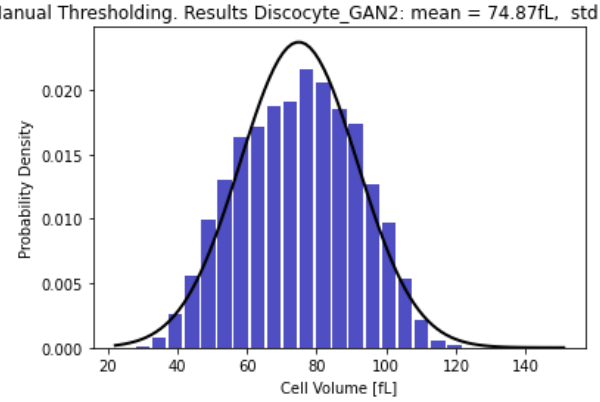




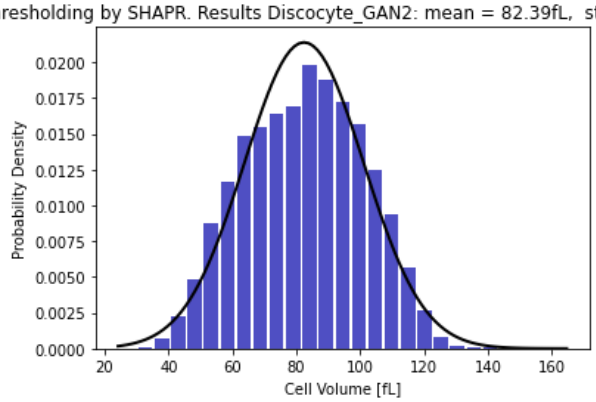


13.5 Results Discocyte_GAN2 Discocyte only training set

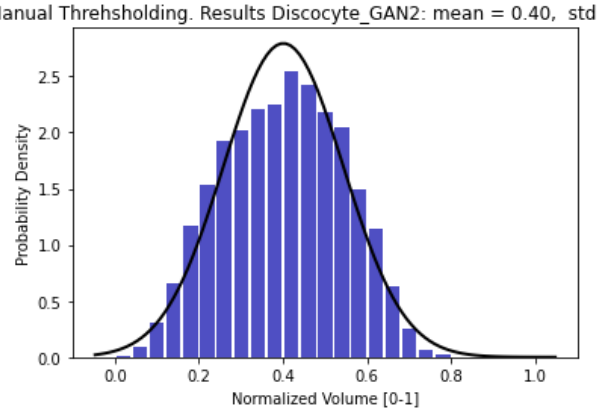
Manual Thresholding. Results Discocyte_GAN2: mean = 74.87fL, std = 16.81fL



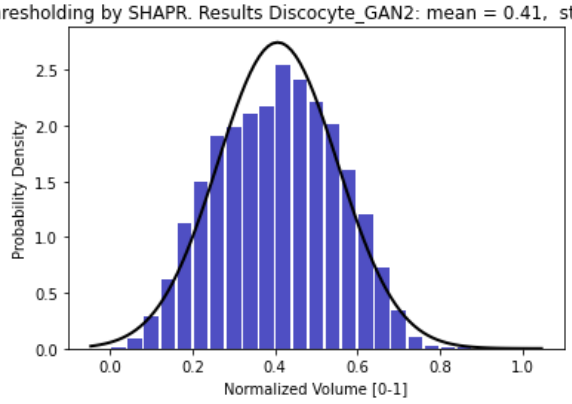
Thresholding by SHAPR. Results Discocyte_GAN2: mean = 82.39fL, std = 18.65fL



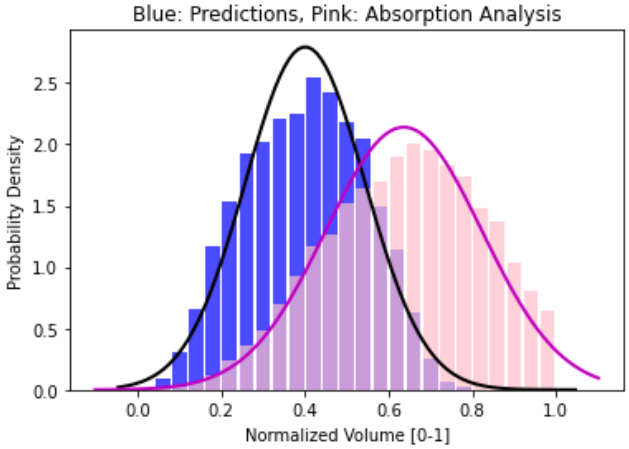
Manual Threhsholding. Results Discocyte_GAN2: mean = 0.40, std = 0.14



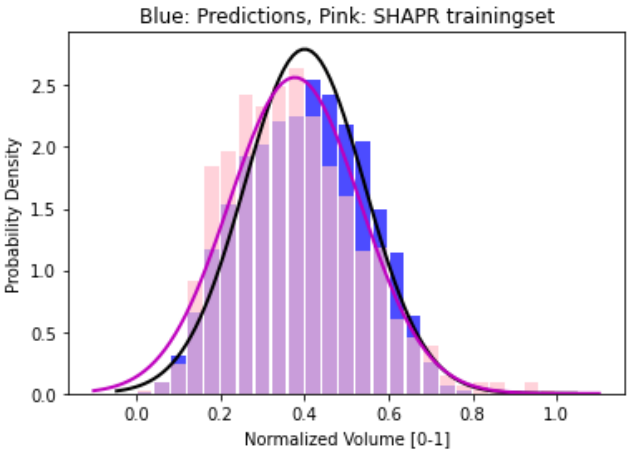
Thresholding by SHAPR. Results Discocyte_GAN2: mean = 0.41, std = 0.15



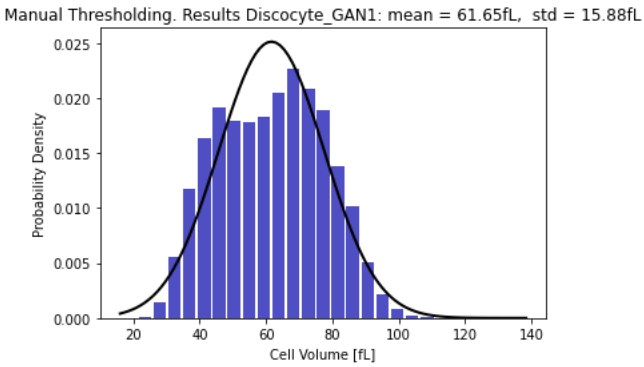
Blue: Predictions, Pink: Absorption Analysis

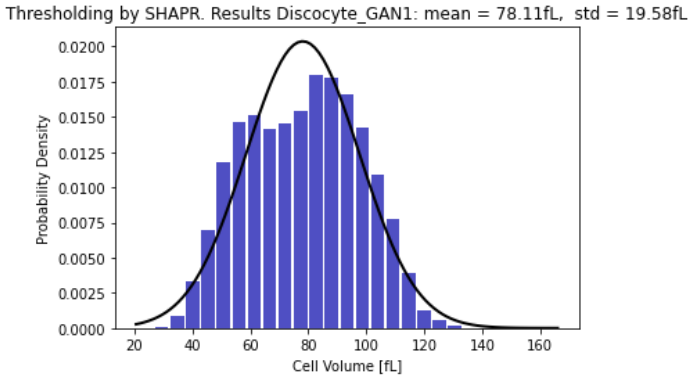


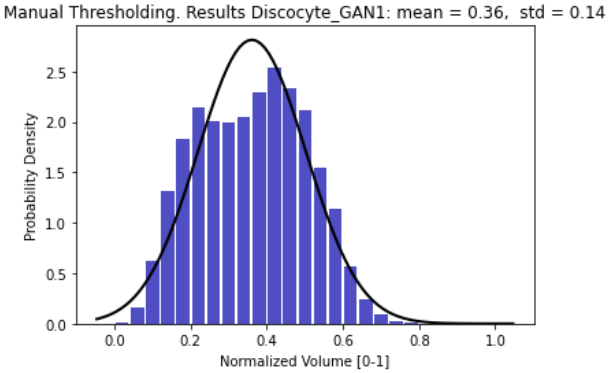
Blue: Predictions, Pink: SHAPR trainingset

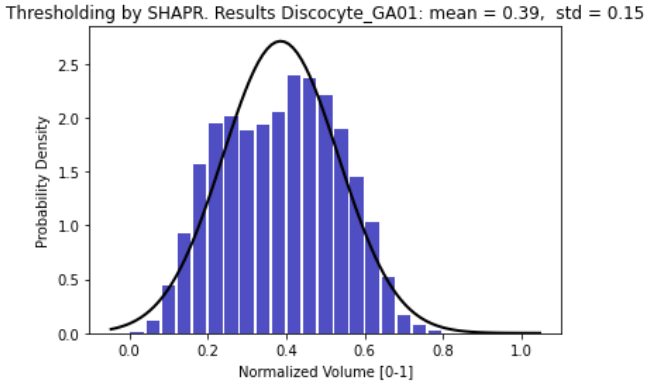


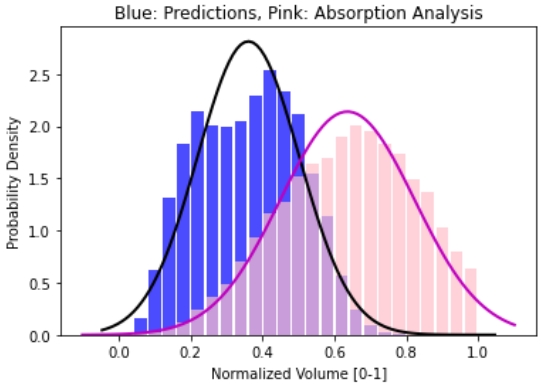
13.6 Results Discocyte_GAN1 Discocyte only training set

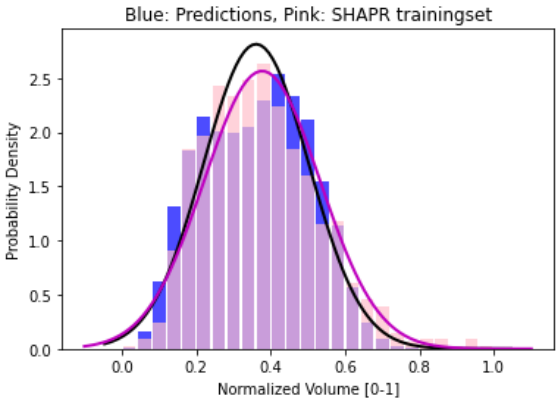












13.7 Results Discocyte_GAN0 Discocyte only training set