



Lehrstuhl für Medizinische Biotechnologie  
Technische Fakultät  
Friedrich-Alexander-Universität Erlangen-Nürnberg

# Semi-Supervised Z-stack segmentation using Random Forest Classifier

Mini Thesis  
*In the program Advanced Optical Technologies*

*Submitted by*  
**Sergei Dobrovolskii**  
Matrikelnummer: 22745251

Erlangen, 01.04.2022

*Supervisors:*  
Dr. Sebastian Schürmann



## Zusammenfassung

Die Biologie ist eine quantitative Wissenschaft, die eine beträchtliche Datenmenge benötigt, um Hypothesen zu überprüfen. Bilder von Zellen und Geweben müssen dazu verarbeitet, analysiert und quantifiziert werden, um wichtige Erkenntnisse zu gewinnen.

Für die Zwecke dieser Arbeit ist es notwendig, Immunzellen in dreidimensionale Stapel zu segmentieren, aber es ist eine Herausforderung, da die Zellsegmentierung mit herkömmlichen Methoden schwierig durchzuführen ist. Hier wird die Segmentierung von Zellen mit einem maschinellen Lernverfahren wie Random Forest angegangen.

Es kann die Produktivität steigern und Benutzer unterstützen, indem es das gesamte Bild mit nur einer kleinen Menge an Eingabedaten beschriftet.

In dieser Arbeit wird ein Random-Forest-Klassifikator als Werkzeug zur halb-automatischen Markierung untersucht und auf Z-Stapel von Dickdarmgewebe angewendet, die mit einem Multiphotonenmikroskop aufgenommen wurden.

## Abstract

Biology is a quantitative science, that requires a significant amount of data to test a hypothesis. Images of cells and tissues are a great source of data, but must be processed, analyzed, and quantified to get important insights.

It is necessary to segment immune cells in three-dimensional stacks for the purposes of this thesis, but it is challenging since cell segmentation is difficult to do using conventional methods. Here, the segmentation of cells will be addressed using a machine learning method like random forest.

It can increase productivity and assist users, by labeling the whole image, given only a small amount of input data.

In this thesis, a Random Forest classifier as a tool for semi-automated labeling will be investigated and applied to z-stacks of colon tissue acquired by a multiphoton microscope.



# Table of Content

<b>1</b>	<b>Introduction</b>	<b>8</b>
<b>2</b>	<b>State of the Art</b>	<b>10</b>
2.1	Imaging in Medical Diagnostics .....	10
2.1.1	Imaging in medicine .....	10
2.2	Motivation.....	11
2.2.1	Human immune cells in colon tissue.....	11
2.2.2	Multiphoton Microscope.....	12
2.3	Aims of Image Analysis .....	13
2.3.1	Machine learning in image processing .....	14
2.4	Machine learning algorithms.....	15
2.4.1	Decision Tree and Random Forest .....	15
2.4.2	Feature extraction .....	17
2.5	Evaluation metric.....	18
<b>3</b>	<b>Methods</b>	<b>20</b>
3.1	Data labeling and cleaning.....	20
3.2	Weka segmentation plugin .....	22
3.3	Intersection over Union calculation .....	23
<b>4</b>	<b>Results</b>	<b>24</b>
<b>5</b>	<b>Discussion</b>	<b>27</b>
5.1	Classifier feature selection.....	27
<b>6</b>	<b>Conclusion</b>	<b>31</b>
<b>7</b>	<b>References</b>	<b>32</b>



# **Abbreviation**

MPM	Multiphoton Microscope
CNN	Convolutional Neural Network
H&E	Hematoxylin and Eosin
RF	Random Forest
DTC	Decision Tree Classifier

# 1 Introduction

In the frame of this project, labeling for further investigation of immune cells in human colon tissue is required. Each individual cell must be located, identified, and its features analyzed. This provides objective and statistically reliable evidence to support a hypothesis. To achieve it, however, each cell must first be segmented. Even though segmenting images is a tedious operation, doing it with three-dimensional data is far more challenging.

The studies of immune infiltrate in the infected tissues play a role to research the disease progress, which would be beneficial for future treatment or at least in the diagnostics [1]. Immune cells in a tissue are acquired using the multiphoton microscope, which creates a three-dimensional image of the tissue. Each channel of this image corresponds to a different fluorophore, that binds to specific proteins, that helps to visualize different parts of the specimen like immune cells, collagen matrix, epithelium cells, etc. That helps to distinguish between them. For further research, the quantification of the immune cells' number, their total intensity, and shape will help to unveil different types of immune cells and their statistics over the disease progress. Acquiring instances of the immune cells and quantifying them can be addressed as a segmentation problem, that can be solved using an automated machine learning approach.

Getting insights from volumetric biological data is a sophisticated task, that depends upon the modality that was used to acquire images, sample preparation, and software that will quantify the raw data and give significant results, that can be interpreted. Labeling for segmentation tasks is much more difficult than any other classification task. It requires from user to outline the pixel regions carefully for each image. Especially for images with cells, all cells must be labeled to avoid algorithm classifying them as non-target ones, which in the end will lead to low prediction performance or even false predictions.

But this problem becomes even more complicated when the number of data dimensions is increased by 1. The three-dimensional data is difficult to display. Scientists are used to working with planar data, therefore with 3-dimensional data, only a single image slice is visible at a time during labeling. This escalates the complexity and number of annotations, and the time investment from the

user further. Speeding up the labeling process by any means will save time and money during the data labeling. Hence assistance for labeling would be a helping hand to reduce the time for labeling.

For the segmentation task, one of the popular classical machine learning algorithms is called Random Forest classifier (RF). It is widely used in different fields and has proven to be efficient and time-competitive nowadays [2]. It exploits the statistical benefit of ensembling to pursue a competitive per-pixel classification, that results in good precision and smooth object boundaries. This algorithm is supervised and requires a set of labeled annotations, but in comparison to deep learning approach requires 1000x fewer examples which is beneficial for the task. RF uses for pixel-wise classification of the features generated by a set of image filters, selected by the user. The increased number of features leads to better precision but slower computation, hence a clever feature set selection is required. In this Thesis classical machine learning approach using Random Forest will be tested, optimized, explained, and utilized for human immune cells segmentation.

## 2 State of the Art

### 2.1 Imaging in Medical Diagnostics

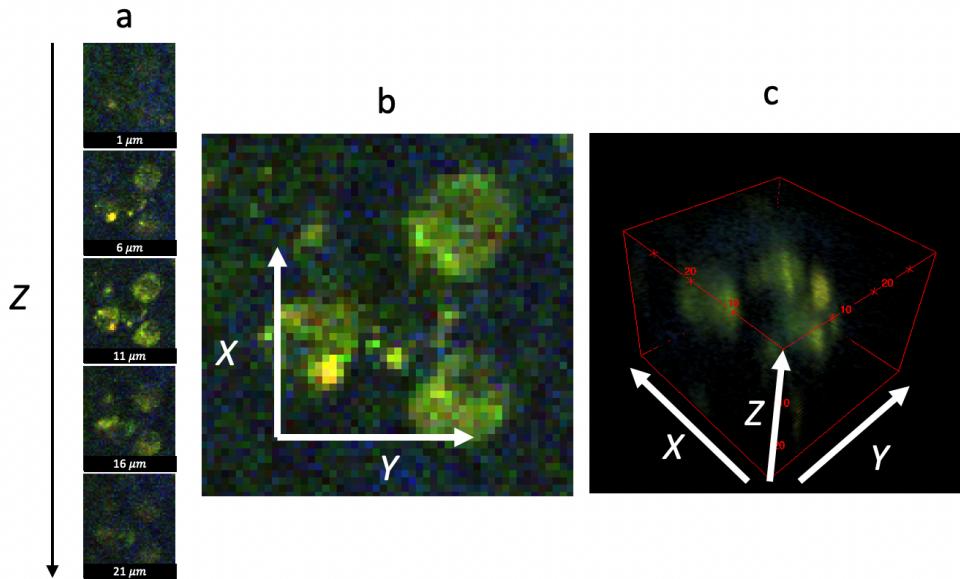
In this section, the imaging modalities in diagnostics, and the type of acquired data will be explained. This is important to understand the underlying principle of what volumetric data is and how image data is acquired in different modalities.

#### 2.1.1 Imaging in medicine

The human body is a highly complex system with a plethora of organs, consisting of tissues, that are formed by billions of cells that group together. These cells, interact with each other and support life in a body. Like any complex system, this one tends to fail too – any disease can make life complicated or impossible. The first step of curing any disease is diagnostics. There are multiple imaging modalities in medicine these days: X-ray, Ultrasound, Computed tomography, Optical coherence tomography, Magnetic resonance imaging, and Microscopy.

Microscopy is used for visualizing single cells and tissue. Oftentimes prior to image acquisition, tissue samples must be fixed (with a substance like formaldehyde) and stained. For brightfield imaging, the most common stain is Hematoxylin and Eosin. For fluorescence microscopy the most common stains are DAPI, Hoechst, FITC, TRITC. Those are used for making tissue or cell structures visible and isolated from other structures. Different microscopic modalities can deliver fluorescent volumetric data such as Multiphoton Microscopy or Confocal laser Microscopy.

Cells are usually nonplanar and tend to live in three dimensions. The examination under the coverslip is not accurate enough and can hide some of the cellular behavior. To receive more data from the tissue it is required to record volumetric images also known as Z-stacks or stacks. This is a three-dimensional image with multiple color channels.



**Figure 1.** Representation of volumetric data. A: The montage of images acquired at different depths Z. B: Zoomed in a single image. C: 3D representation of a stack.

Stack allows seeing the cells in volume, which is oftentimes necessary, for example, direct observation of skin cells together with immune cells lying under the skin layer [1]. Volumetric data is more difficult to work with, due to the high memory consumption and visualization tradeoffs that we must do to observe the data.

## 2.2 Motivation

In biological research, any experiment consist of multiple steps like hypothesis, subject preparation, acquisition and analysis. In this section subject of experiment, acquisition modality, and general data analysis concepts will be explained.

### 2.2.1 Human immune cells in colon tissue

The immune cells are present mostly everywhere, but if the inflammation occurs, then the immune cells will accumulate in that area, hence their concentration and composition will change. The types of cells in the infected area and their relative concentration can tell a lot to the doctor about the inflammation process.

A high concentration of immune cells is concrete evidence of inflammation in human colon tissue. But not only the presence of immune cells is a source of information about the inflammatory process, but also the cell types, their concentration, and location. Once the fluorescence signal is created, it is important not to mix it with the signals, that arise from different objects, that are out of interest. Hence the exact immune cell location is important for the measurement of the immune infiltrate.

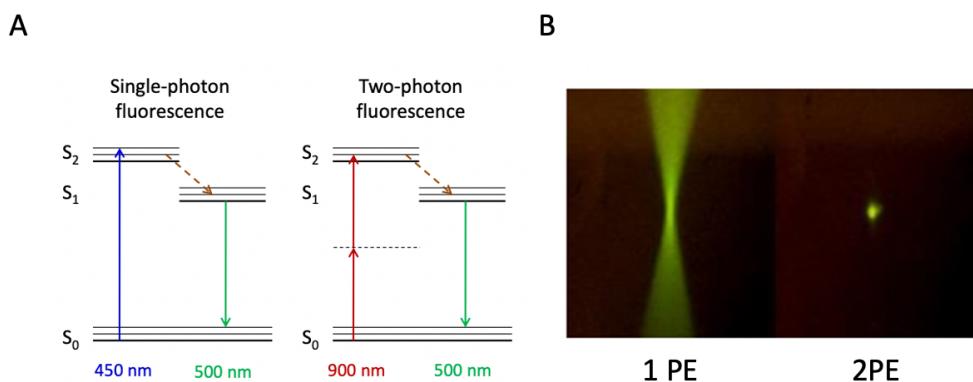
To detect the presence, quality, and quantity of immune cells in tissue imaging techniques like Multiphoton Microscopy might be very useful.

### 2.2.2 Multiphoton Microscope

Multi-photon microscopy (MPM) is a powerful tool that allows three-dimensional mapping of samples that have a measurable nonlinear optical response such as second harmonic generation (SHG), third-harmonic generation (THG), or fluorescence induced by multiphoton absorption. MPM provides a way to observe stained tissue and single cells with high resolution, in 3D. It reduces the scattering from non-focal planes, by excitation only at the focal plane.

Fluorescence is the process when the electron of the fluorophore absorbs the excitation photon and settles on a higher energetic level of the molecule. Then it relaxes down to the ground state, with the emission of a light photon of a different wavelength (**Figure 2A**). Two-photon excitation works the same way, but instead of a single excitation photon, it requires two incident photons each with half the energy of the required one. For this effect to happen, these photons must hit the same atom simultaneously. To achieve this seldom event the density of photons must be high.

The principle of this modality differs from the fluorescent microscope. In a fluorescence microscope, the excitation response of the fluorophore is linear, more excitation light - more fluorescence response. But with the two-photon excitation, it is different (**Figure 2B**). Excitation is visible only at the focal point – the place where the photon density will be the highest.



**Figure 2 A:** Multiphoton fluorescence energy diagram comparison of a single-photon (1 PE) fluorescence and two-photon (2 PE) fluorescence. In two-photon fluorescence, 2 photons of the energy twice lower than required are exciting electrons together. This is achieved by higher energy density which results in a higher probability of excitation events happening. **B:** The comparison of the excited volumes – 1 PE has a lot of exciting molecules out of the focal plane which results in worse image quality, compared to 2 PE.

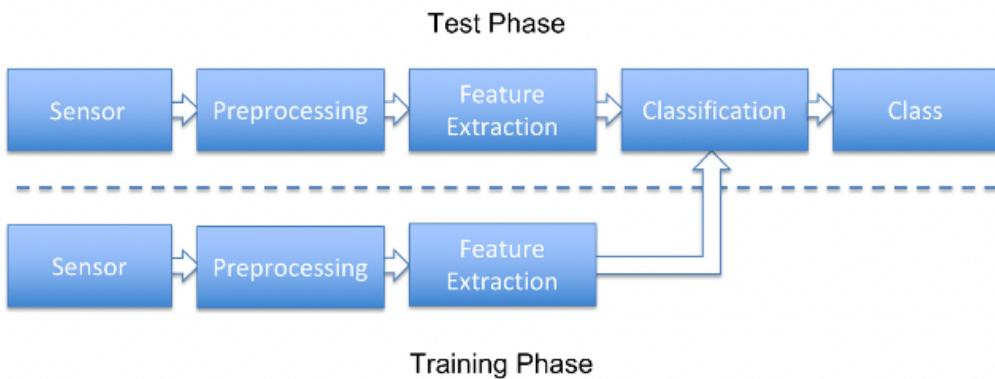
MPM has found important applications in nonlinear materials characterization, biological research, and diagnosing medical conditions. The data in this thesis was acquired using this modality.

### 2.3 Aims of Image Analysis

For the experiment conduction, the data is acquired by any given sensor (photo-multiplier tube in our case). This raw information provides a good view on the processes, structure, intensity, but to reveal a full potential and extract more information from them some analysis is required. To prove the hypothesis this information must be cleaned, distilled, and processed. All these tasks are faced by image analysis – the field of science and the mathematical toolset. Image analysis involves processing images into fundamental components to extract important information. It may involve tasks such as finding shapes, detecting edges, removing noise, counting objects, texture analysis, etc. achieved by means of mathematical operations, or dynamically, using machine learning approaches. In this Thesis, only the segmentation procedure for human immune cells in three-dimensional data will be explained.

### 2.3.1 Machine learning in image processing

Machine learning is a set of sophisticated mathematical operations performed on data to receive an expressive result from it. These algorithms are conventionally generalized in form of a pattern recognition pipeline (**Figure 3**). At first, the data is acquired using any type of electronic device such as a camera, microphone, or microscope. Then this data is stored and preprocessed. In preprocessing step data is filtered, enhanced, and prepared for further steps. Next goes a feature extraction, which extracts representative features from data using a set of mathematical operations. These features can represent a simplified version of original data or can create new data, that allows for algorithms to find intra-data correspondences. Those features could be obtained by applying convolutional filters like Gaussian, Laplacian, Gradient etc. At a later stage, these features are used for the *so-called* learning or training step. During training Algorithm will try to find the best possible split to classify the data with minimal error, based on extracted features.



**Figure 3** Typical pattern recognition pipeline [3]. It is divided into two parts: The test phase and the Training phase. Sensor, preprocessing, and feature extraction steps are common for both phases.

These algorithms in image processing are conventionally divided into two groups: Classical and Deep Learning approaches. Differences between those groups can be compared using the pattern recognition pipeline (**Figure 3**):

- Classical machine learning approaches follow this pipeline. They require a small amount of data and small computational costs. Lack of generalization and precision.
- Deep Learning approaches bypass part of this pipeline by combining feature extraction with classification. There are no predefined feature

extraction procedures in neural networks, and these procedures are estimated, during the training process. They require a hundreds of human-labeled data, high computational costs, and time. They are good at generalization and most of the state-of-the-art methods now are using neural networks.

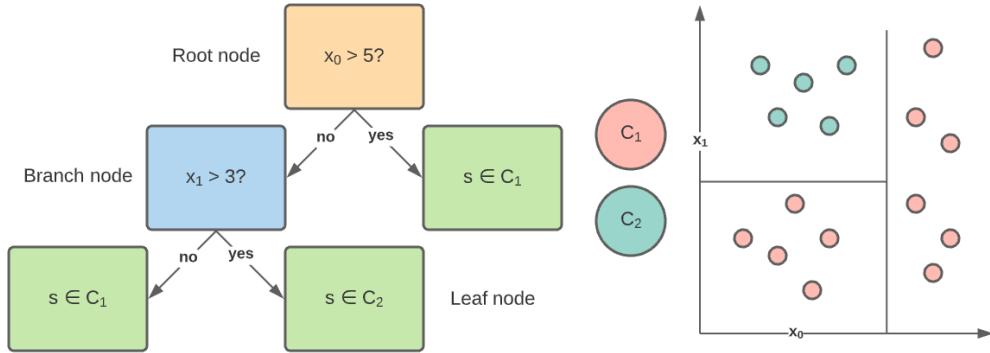
Deep Learning approaches require a gigantic amount of human-labeled data like hundreds or even thousands of images. Due to that, it is always preferable to find a way to minimize the labeling time from humans because it is tedious and expensive. For that reason, it is important to assist the human operator with labeling tasks, by iteratively annotate full image automatically via machine learning tool, and refine annotation by providing small manual adjustments. One of the promising methods is to use classical machine learning approaches like Random Forests for easier data labeling.

## 2.4 Machine learning algorithms

With the use of machine learning (ML), which is a form of artificial intelligence (AI), software programs may anticipate outcomes more accurately without having to be explicitly instructed to do so. In order to forecast new output values, machine learning algorithms using previous data as input. There are concepts like Decision Tree and Random Forest that allow to map several values from one domain to another by classifying them with simple rules, based on few annotated examples. Those can be used for image processing for automated pixel-wise classification, which is also known as segmentation task.

### 2.4.1 Decision Tree and Random Forest

Decision Trees are a non-parametric supervised learning method used for classification and regression. The goal is to create a model that predicts the value of a target variable by learning simple decision rules inferred from the data features. The selection process can be described as a sequence of binary selections corresponding to the traversal of a tree structure. One limitation of decision trees is that the division of input space is based on hard splits in which only one model is responsible for making predictions for any given value of the input variables.



**Figure 4** Classification Tree example. On the left is the example of the classification tree: orange is the root node, blue is the branch node, and green is the leaf node (decision node). On the right is the sample space division based on the classification tree from the left. Red samples correspond to class 1 and blue to class 2. Image by author.

The tree model consists of nodes and each node will ask a true-false question about one of the features ( $x_0, x_1$ ) (Figure 4). And in response to this question, the data is split into two subsets. These subsets then become the input to two child nodes that are added to the tree. And the goal of the question is to unmix the labels as proceeding down to leaf nodes. Or in other words, to produce the purest possible distribution of the labels at each node. In other words a Decision Tree is dividing a multi-dimensional space of input parameters in small domains, where the label distribution is the purest, using a set of generated questions. With unseen data, Decision tree spans input parameters in these domains, based on “learned” set of questions to each domain. That is called classification.

The quantification of a split uncertainty at a single node can be achieved using a metric called Gini impurity:

$$G = \sum_{i=1}^C p(i) * (1 - p(i)) \quad (2-1)$$

$C$  – number of classes,  $p(i)$  – probability of correctly classifying class  $i$ . And we can quantify how much a question reduces that uncertainty using a concept called information gain. Given that question, we'll recursively build the tree on each of the new nodes. We'll continue dividing the data until a limit called a tree height will occur, at which data is no longer divided and a class is assigned

based on the majority of samples present in the last subset of points. This height is set up manually to prevent a tree from overfitting.

A single classification tree is a powerful algorithm that lacks generalization and is prone to overfitting. Ensembling of multiple classification trees is a common strategy to achieve higher generalization and accuracy.

Random forest is a classical machine learning method for data classification. It is based on the ensembling of multiple Decision Trees and “decides” based on the majority voting of all decision trees. Each tree is trained on a random subset of data, which leads to a random set of uncorrelated trees. This algorithm is used for the classification of pixels for the annotation of digital stacks.

#### 2.4.2 Feature extraction

Why do images' pixel values can not be directly used with Random Forest for segmentation purposes? Intensity values on their own are a weak data representation, they do not represent the neighborhood regions, also they cannot be used with RF, because it will only divide the image dataset based on intensity threshold, which is ineffective. Spatial information is much more valuable and tends to generalize better. How to blend the information from a neighborhood region and extract important information about it? It can be done using a mathematical operation, called convolution.

Convolution is a mathematical operation, performed on two functions, that produces a third one, that expresses how the shape of one function will modify the shape of another one:

$$(f * g)(t) := \int_{-\infty}^{+\infty} f(\tau)g(t - \tau)d\tau \quad (2-2)$$

Once used for a discrete domain, it must be reformulated for image stack applications like:

$$h[x, y, z] = \sum_{i,j,k} f[i, j, k]I[x + i, y + j, z + k] \quad (2-3)$$

Here  $x, y, z$  are the pixel coordinates in the image stack, and  $i, j, k$  are the pixel-wise iterators. From these two definitions, it is defined, that a single pixel intensity is a function of the pixel's neighborhood and a new function, called the kernel.

In image processing different types of functions or kernels are convolved with an image to obtain different spatial effects:

1. Box kernel – spatial linear filter, each pixel in the resulting image has a value averaged of its neighborhood pixels.  

$$f_{box} = \frac{1}{w^2} \sum_{i,j,k} I(x+i, y+j, z+k),$$
where  $w$  is a size of a box kernel. This kernel blurs the image.
2. Gaussian kernel –  $G(x, y, z, \sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2+y^2+z^2}{2\sigma^2}}$ , where  $\sigma$  is a standard deviation. Blur the image with an effect of a circular aperture – the kernel has radial symmetry.
3. Sobel operator – used in image processing for edge detection. Composed of two operations: Finding the spatial x and y image derivatives  $G_x, G_y$  and combining their results with  $G = \sqrt{G_x^2 + G_y^2}$
4. Difference of Gaussians – a kernel composed as a difference of two different Gaussian kernels  $f_{DoG} = G_{\sigma_1} - G_{\sigma_2}$  is a close approximation of  $\nabla^2 G(x)$  and is used as a replacement for – blurring and edge detection procedure, due to decreased computational cost.
5. Median filter – nonlinear filter, that assigns the median intensity value of the neighborhood pixel area to a given pixel. Allows minimizing the presence of high-frequency noise.

All these kernels with different parameters are used to create image features, necessary for preserving the neighborhood correspondence and extraction of important features from it, like edges, textures, and shapes.

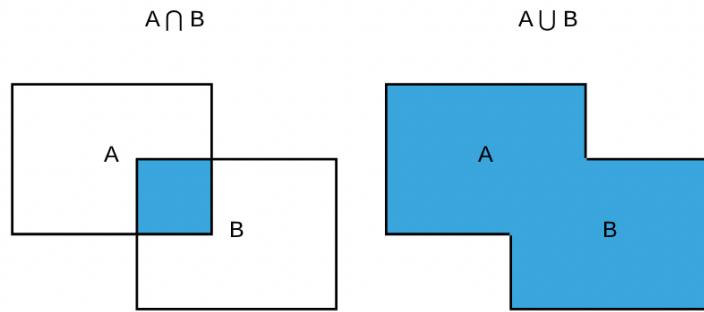
## 2.5 Evaluation metric

For a representative comparison of the segmentation algorithms, the metric called Intersection over Union (IoU) is used. Its calculation is simple and explainable. The IoU spans between 0 and 1, where 0 means no overlap of the mask area with the ground truth area, and 1 means the complete overlap.

$$IoU = \frac{\text{Area of intersection}}{\text{Area of union}} = \frac{A \cap B}{A \cup B}$$

Where A and B mean the predicted and ground truth areas. If the predicted area contains a complete ground truth area, but the predicted one is a bit larger, the misclassified pixels will result in a smaller IoU than 1. This metric

scales for any given size, and data type, which made it an industry standard for machine learning and particularly segmentation tasks evaluation.

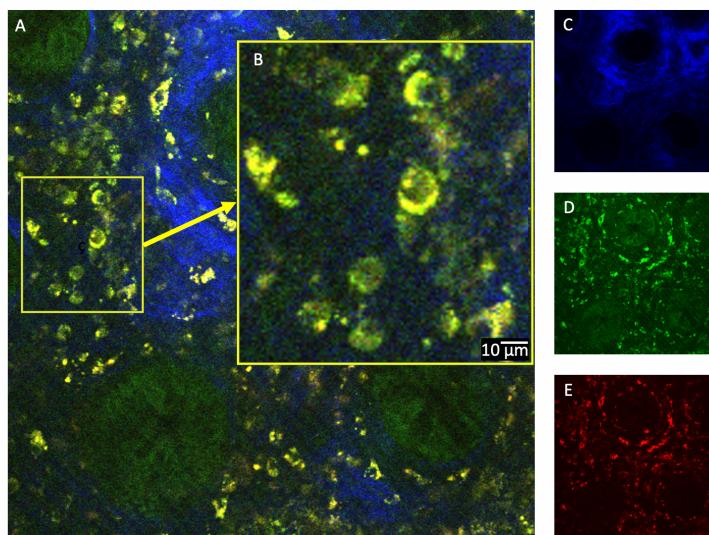


**Figure 5.** Intersection over Union (IoU) visual representation of Intersection and Union operations over areas.

## 3 Methods

In this section, the exact image processing pipeline for immune cell segmentation in three dimensional stacks will be explained.

The 3d cell segmentation will be solved as an example on one of the stacks, acquired at our lab in the Institute of Medical Biotechnology FAU Erlangen. This is the stack recorded from the patient's colon tissue. It is 400 µm x 400 µm x 100 µm stack with three channels: Collagen autofluorescence (blue), NADH (green), FAD (red). This sample has a visible presence of immune cells (yellow), which can be seen as the presence of the immune cells (**Figure 6**).



**Figure 6** A: Image of the human colon tissue taken from a stack, acquired with the multiphoton microscope. B: Magnified area. C: Collagen autofluorescence channel. D: Green channel (NADH fluorescence). E: Red channel (FAD fluorescence)

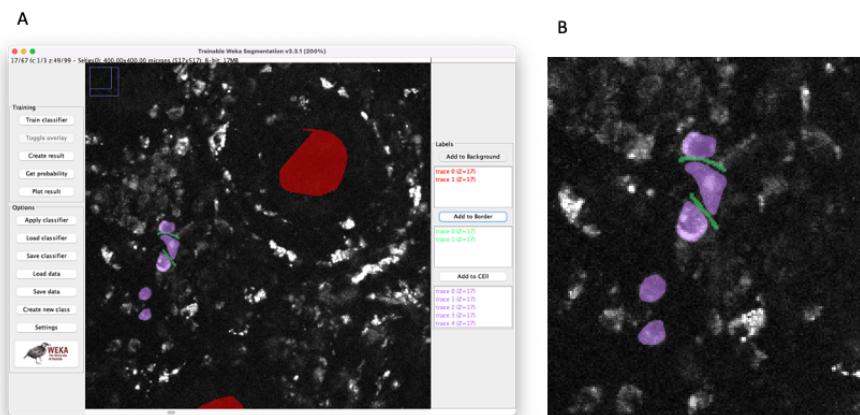
One of the typical human immune cell types is Neutrophil. Its size lies in the range of 6 to 14 µm, which can be seen in **Figure 6** B. The segmentation of these cells in semi-supervised fashion is the aim of this work. But first, it is necessary to simplify the data and eliminate the unnecessary parts.

### 3.1 Data labeling and cleaning

For correct data processing, we must select only information, that contributes to result, to avoid unnecessary computations and to make the algorithm work faster.

The stack is a three-channel image. Each channel represents a signal from a certain fluorophore. The yellow color of the immune cells is a combination of green and red channels. The green and blue channels indicate the colon crypts and the collagen matrix, which regions are not interesting in the scope of the research **Figure 6 C, D**. Therefore, to highlight the immune cells, it is enough to use only the red channel **Figure 6 E**. Hence crypts and collagen matrix signals will be suppressed. Next, the segmentation plugin will be used for the labeling of data.

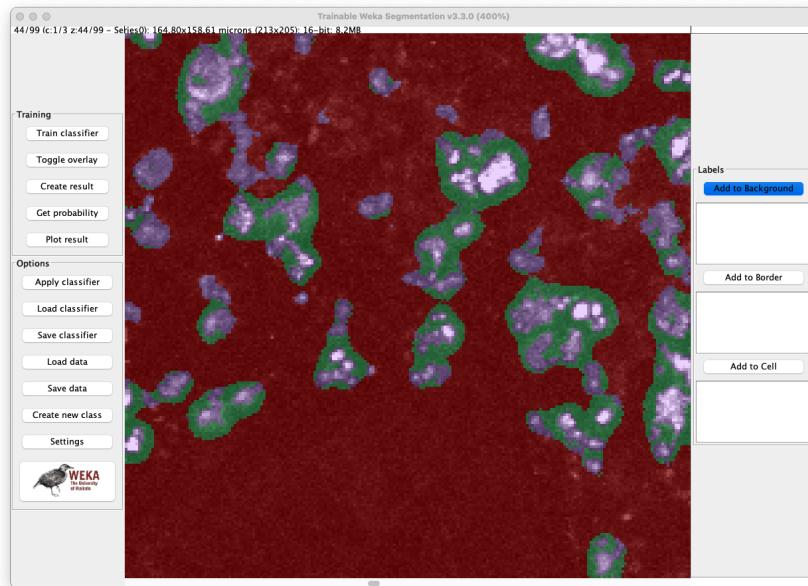
The number of classes for segmentation is three: Background, Border, Cell. The background class shows the signal, which is assumed to be a background signal. Border class defines the extracellular boundaries between cells. This class is aimed to improve the accuracy of background classification and separate cells from each other. The labeling procedure is easy for ImageJ experienced users – it requires the usage of default labeling tools from ImageJ like pen, polygon, and rectangle selections. Once the region is selected, the pixelated area appears to be of the class color. Then on different depths, the labeling also must be performed. Because the features, for segmentation, are three dimensional, it is necessary to label the cell boundaries in the Z direction too. They appear to be dim, but still must be considered for more accurate prediction.



**Figure 7 A:** Weka segmentation 3D plugin window. Colored segments are the human input for the training procedure. All of them are listed on the right part of the window. **B:** Enlarged labeling example of cells (purple) and borders (green) – the separation of the cells is required to be labeled.

### 3.2 Weka segmentation plugin

Open-source FIJI [4] (based on ImageJ) has a built-in plugin for human-aided image segmentation, using random forests. This plugin is called Weka Segmentation [5]. It utilizes the random forest approach together with the feature extraction described in the previous chapters. It can be accessed from the toolbar Plugins menu -> Segmentation -> Trainable Weka segmentation 3D.



**Figure 8** Example image of the Weka Segmentation plugin. The red segment is the background, green is the cell neighborhood, and purple is the cell.

Usage of this plugin is simple, but some preparations are required for the given task. Setup requires:

1. Select number of classes: it is necessary to segment 3 classes – Background, Cells, Cell borders.
2. Select a set of the feature types: these features will be used for classification by RF. In 5.1 it is explained, which set of features to select for the best performance.
3. Draw a polygon around individual cell, add it to the Cell class; Draw a polygon in area where no immune cells are present and add it to the Background class; Draw line in area where there is a small gap between cells, add it to Border class;

4. Repeat step “3” 5 times for different depths and press “Train Classifier”
5. Observe segmentation result and add corrections to it by repeating step “3” to “5”
6. If result looks good enough, press “Create result” to generate a mask. This mask can be used further for cell counting or segmentation.

### 3.3 Intersection over Union calculation

When performance evaluation is needed, a metric called Intersection over Union (IoU) can be calculated using the script below:

```
import numpy as np
import matplotlib.pyplot as plt
from skimage import io, filters, morphology
import os

def load_mask(path: str):
    '''Load and process the mask'''
    mask = io.imread(path)
    mask = mask.astype(np.bool)
    return mask

def compute_iou(mask: np.bool, GT: np.bool):
    '''Computes IoU'''
    intersection = np.logical_and(mask, GT)
    union = np.logical_or(mask, GT)
    return np.sum(intersection) / np.sum(union)
```

It utilizes 3 important python libraries:

1. NumPy [6] – open-source library for python, that enables multidimensional linear algebra operations. It was created in 2005, building on the early work of the Numeric and Numarray libraries. It is proven to be faster and easier to use than traditional python lists.
2. Matplotlib [7] – the python library for creating static and animated plots and visualizations.

3. Scikit-image (skimage) [8] – python open-source library for image processing. Contains a collection of classical algorithms for image transformation segmentation, data reading, and writing.

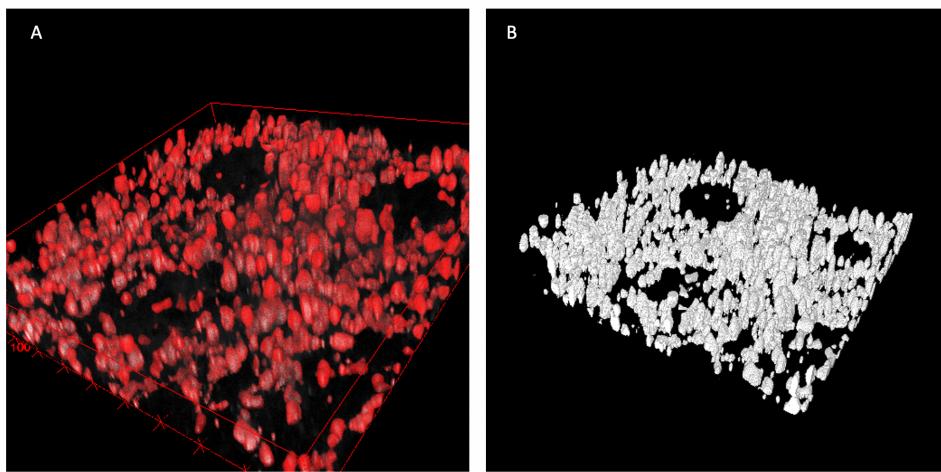
The first method “load\_mask” takes as an input a path to a file as a string, that should be loaded. By using the “imread” method of the module “io” of skimage, this image stack is loaded and saved to the “mask” variable. Then mask is converted to boolean type, to be processed later for IoU calculation.

The second method “compute\_iou” takes as an input 2 NumPy boolean arrays – mask that an algorithm produces and the ground truth (GT) that will be compared to the mask. Then separate Boolean operations of intersection and union are performed with these 2 arrays. Then in the end we divide the sum of all entries considered to be intersected to calculate the intersection area and divided by the sum of all union pixels. This produces the IoU result that is calculated later.

## 4 Results

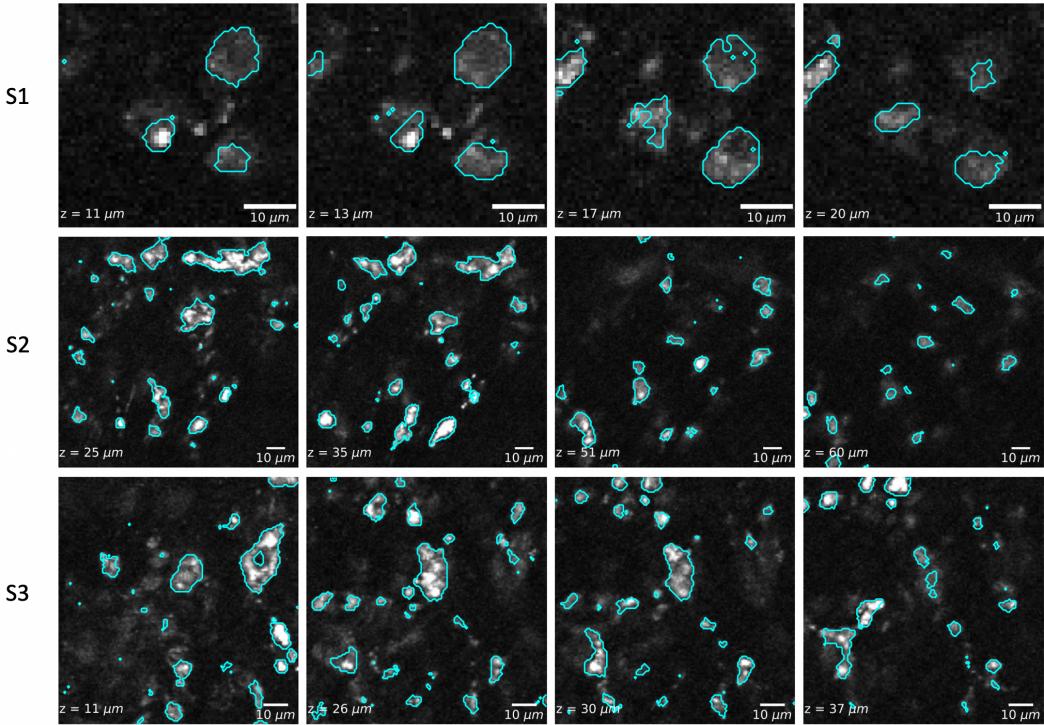
In this section, the question of cell segmentation using Random Forest will be answered.

One of the stacks from was completely segmented, using the Random Forest approach (**Figure 9**). This 3d representation shows how well the RF algorithm performed, when it comes to immune cell segmentation. The red – white image **Figure 9A** shows the overlay of intensity channel with mask, and B shows just mask.



**Figure 9.** Examples of immune cells segmentation using Random Forest classifier. A: overlay of mask and real data. Red volume refers to cells and Dark to the original image. B: The segmentation results only

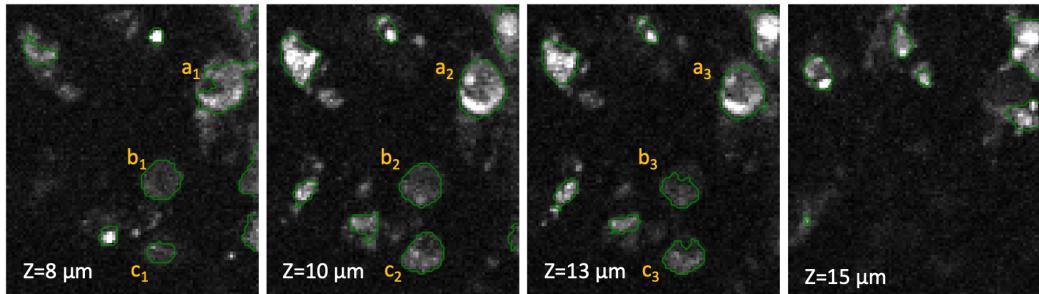
On the image next figure (**Figure 10**) can be observed 3 image regions (S1, S2, S3). For each of these regions, images from different depths were taken and overlaid with the predicted mask (cyan outline). It shows How algorithm works. For S1,  $z=11\mu\text{m}$  the cell outlines were indicated properly. But for S3,  $z=11\mu\text{m}$  can be seen a false segmentation of cell nucleus – it was excluded from segmentation.



**Figure 10:** Visualization of the algorithm performance. 3 regions S1, S2, and S3 were used to acquire 4 image samples from different depths. The mask overlay is indicated with a cyan contour.

The segmentation results can be observed on one of the stacks (**Figure 11**). Here the images were taken from different z coordinates to show how the algorithm can connect regions in three dimensions. It can be seen three labeled cells. The cells' outlines are clear for  $a_{2,3}$ ,  $b_{1,2}$  and  $c_{1,2}$ , the segmentation is clear and defect-free. The contour line goes exactly on the cells' membrane. But this is not always the case. Sometimes the algorithm can fail when segmenting the cells with a black spot on, which is caused by the nucleus (**Figure 11** a<sub>1</sub>, b<sub>3</sub>, c<sub>3</sub>). That is a common issue with this dataset. The nuclei do not produce the fluorescence signal, because it is not stained, and the only signal, that we can get for segmentation produces a signal from intercellular volume, except a nucleus. The algorithm struggles to differentiate between separation of extracellular space and nuclei space.

Current segmentation lacks cell separation (**Figure 9**). Often, when performing the labeling it is difficult to distinguish between separate cells. Having easily distinguishable anchors like separately stained cell nuclei might improve accuracy and help to quantify cells.



**Figure 11:** The images, taken from the same z-stack ( $68\mu\text{m} \times 78\mu\text{m}$ ) at different z coordinates with the segmentation contour applied with green color. With the same non-capital letters the same region with an x/y coordinate is noted.

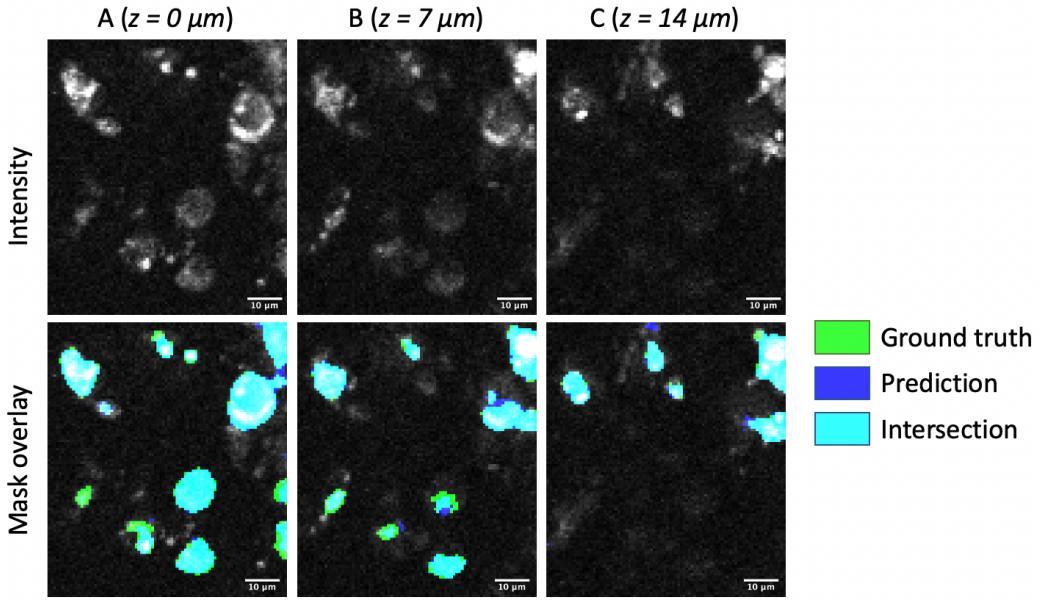
## 5 Discussion

### 5.1 Classifier feature selection

Random forest is an effective approach for semantic segmentation, for three-dimensional images. More filters are used for classification the more stable and reliable results can be acquired. Generalization gets higher (the model does not overfit) and precision rises. On the other hand, the memory consumption rises too together with the computation time required for feature computation. To increase processing time main goal is to find the golden ratio between computation time and model precision. For smaller stacks, more different features can be computed in comparison to larger stacks with better precision and speed. For larger stacks, it is important to select the most adequate feature set as a precision trade-off.

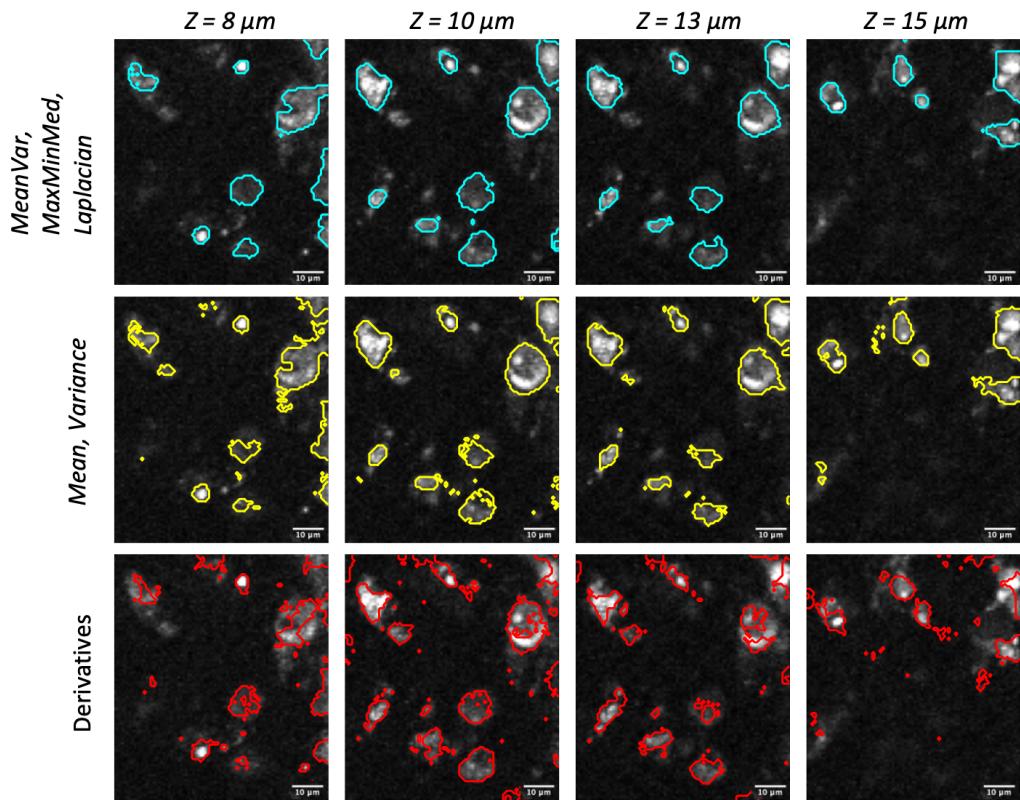
The performance comparison will be performed using the IoU metric (Intersection over Union). It goes to 0 when there is a small overlap between the ground truth area and the predicted area.

The segmentation result can be seen in **Figure 9**. It was slightly filtered after classification, using a median filter of size 1 and cancellation of all blobs with a size  $< 10$  pixels.

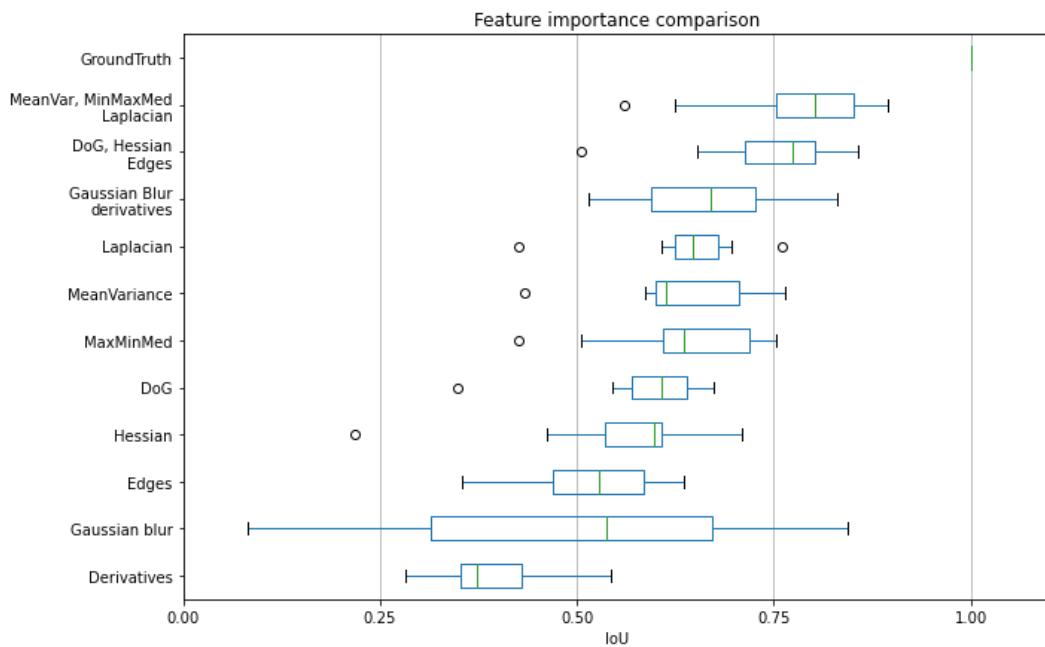


**Figure 12.** Segmentation results are taken from different z depths. The grayscale channel is the cell fluorescence, green is the ground truth mask, and blue is the prediction of the model, made by using the best filter set: Mean, Variance, Min, Max, Median, Laplacian. Cyan is the intersection of the ground truth and the model prediction.

As mentioned in section 2.4.2 features are important, but heavy to compute and store in memory. Therefore, finding the right set of features is crucial. The influence of the feature selection on the classification accuracy will be shown in **Figure 13** and **Figure 14**. The training data was labeled and then a classifier was trained only using a single feature set and then the IoU (Intersection over Union) was calculated and compared to the ground truth ( $IoU = 1.0$  **Figure 14A**). It can be observed that the Mean-Variance feature set alone will perform better than the rest ( $IoU = 0.68$ ). Then comes the Max-Min-Median set comparable to the previous one with  $IoU = 0.66$ . The most underperforming ones are the Derivatives and Gaussian Blur sets with  $IoU = 0.39$  and  $0.51$  units. Any single feature set, used for training, didn't allow to fit the model well – the model under fitted. Hence the amount of training data must be extended by combining feature sets during the training **Figure 14B**. Even a combination of two worst feature sets together outperforms a single best set ( $IoU_{single\ best} = 0.68$ ;  $IoU_{double\ worst} = 0.72$ ). A combination of better performing sets will give even a better result (up to  $IoU = 0.82$  and more). But with increased accuracy comes the downside – increased computational time and memory consumption.



**Figure 13:** Visual comparison of three different feature sets performance. Cyan outline means the mask, acquired with the Mean, Variance, Max, Min, Med, Laplacian filters – the best feature set; Yellow outline is for Mean, Variance only; Red is for Derivatives filter – it induces the highest noise level in segmentation. It is clear, how much the segmentation performance drops with the decrease in the number of filters.



**Figure 14.** Feature importance comparison to ground truth in RF classification using intersection over union metric. From this plot, the importance of the features for this dataset can be rated. Feature combination comparison – a combination of multiple features performs better than a single feature.

## 6 Conclusion

For semi-supervised segmentation tasks like cell segmentation in volumetric data, Random Forest is applicable. It allows with few annotations predict the segmentation of three-dimensional stack. It can help a scientist to perform tedious task faster. It does not require specific knowledge and can be learned by doing. It can work with different settings that can bring better performance with slower computation, or vice versa. But how different is it in comparison to the deep learning approach? The Deep Learning approach will require hundreds of completely hand-labeled images, a powerful workstation, and hours of training. The preliminary result can be observed only after a day. But segmentation of unseen data can be done relatively fast, with good generalization and accuracy. With the Random Forest approach, you can get the stack-wise result in less than a minute, depending on the stack size and number of features. The amount of data required can be equal to 3-4 labeled cells on 1 slice! The rest generalizes itself. But with simplicity comes underperformance – Random Forest lacks generalization for different images. The capacity of this classifier is limited, and features are not optimized for the application and are selected mostly intuitively. Deep Learning approach on the other hand optimizes a huge number of feature extractors to be most efficient for selected applications and hence performs better.

Random Forest can be used as a labeling aid for Deep Learning dataset creation. Human manual labeling of 3d microscopic images is a tedious task. It requires a great amount of time and money to be invested. This routine can be simplified using machine learning algorithms. Human aided software might label data at a higher speed than conventional methods. And it requires less training time in comparison to deep-learning methods. Common open-source software like FIJI provides users with a convenient tool to segment 3d images and classify the respective pixels.

In future work, the segmentation masks will be used to identify single cells and isolate their volume from the rest for further investigation, such as their size, shape, intensity, and relative concentration in the infiltrate. This work might give a different perspective on the disease progress and understanding of how the body's defenses work overtime in human colon tissue.

## 7 References

- [1] C. K. A. S. M. V. T. P. B. C. F. K. J. C. A. C. K. J. Z. P. B. S. M. S. S. O. F. M. J. W. Kristina Scheibe, "Inhibiting Interleukin 36 Receptor Signaling Reduces Fibrosis in Mice With Chronic Intestinal Inflammation," *Gastroenterology*, vol. 156, no. 4, pp. A1-A22, e1-e18, 827-1224, 2019.
- [2] B. K. a. T. Q. Nguyen, "Random Forest With Learned Representations for Semantic Segmentation," *IEEE Transactions on Image Processing*, vol. 28, no. 7, pp. 3542-3555, 2019.
- [3] C. S. T. L. C. R. Andreas Maier, "A gentle introduction to deep learning in medical image processing," *Zeitschrift für Medizinische Physik*, vol. 29, no. 2, pp. 86-101, 2019.
- [4] J. A.-C. I. F. E. e. a. Schindelin, "Fiji: an open-source platform for biological-image analysis," *Nature Methods*, vol. 9, p. 676–682, 2012.
- [5] M. Hall, "The WEKA data mining software: an update Share on," *ACM SIGKDD Explorations Newsletter*, vol. 11, no. 1, pp. 10-18, 2009.
- [6] C. R. H. a. K. J. M. a. S. J. v. d. W. a. R. G. a. P. V. a. D. C. a. E. W. a. J. T. a. S. B. a. N. J. Smit, "Array programming with NumPy," *Nature*, vol. 585, no. 7825, pp. 357-362, 2020.
- [7] J. D. Hunter, "Matplotlib: A 2D graphics environment," *Computing in Science & Engineering*, vol. 9, no. 3, pp. 90-95, 2007.
- [8] J. L. a. N.-I. J. a. B. F. a. W. J. D. a. Y. N. a. G. E. a. Y. T. a. t. s.-i. c. van der Walt Stefan and Schönberger, "scikit-image: image processing in Python," *PeerJ*, vol. 2, p. e453, 2014.
- [9] K. A. J. J. R. a. R. Z. Andrew H. Fischer, Hematoxylin and Eosin Staining of Tissue and Cell Sections, Cold Spring Harbor Laboratory Press, 2008.

## **8 Appendix**

# Curriculum vitae

## Personal Details

Name	<b>Sergei Dobrovolskii</b>
Date and place of birth	03.02.1996 in Moscow
Address	Koldestraße 10, 91052 Erlangen
Email	dobrik4ever@gmail.com
Telefon	1516 3610896

## Education

10/2019 – present	Master program Advanced Optical Technologies, Friedrich Alexander Universität Erlangen-Nürnberg
09/2017 – 07/2018	<i>Bachelor of Science (B.Sc.)</i> Bachelor Thesis: "Phase microscope based on multi-step phase calculation method"
09/2014 – 05/2018	Bachelor program Optical engineering, MIREA – Russian Technological University
09/2002 – 07/2014	School №875, Moscow

## Professional Career

08/2020-present	Interherence GmbH – Optical engineer
10/2019 – 07/2020	JSC "AMTEO M" - Software Developer. Developing the automated histopathology scanning brightfield microscope.

**10/2018 – 10/2019**

JSC "AMTEO M":

- Chief Engineer. Developing the automated histopathology scanning brightfield microscope.
- Field service engineer. Repairment of medical devices.
- Crisis manager. B2B communication, factory processes repairment.

**03/2016 – 08/2018**

State budget vocational educational institution of Moscow "Polytechnic College No 47 named after VG Fedorov" – teacher (subjects: microcontrollers, electronics, robotics, programming)

# **Erklärung**

Ich versichere, dass ich diese Arbeit selbständig verfasst und keine anderen als die angegeben Quellen und Hilfsmittel verwendet habe. Die Arbeit hat in dieser oder ähnlicher Form noch keiner anderen Prüfungsbehörde vorgelegen.

# **Declaration**

I confirm that I have written this thesis without any external help and not using sources other than those I have listed in the thesis. I confirm also that this thesis or a similar version of it has not been submitted to any other examination board and has not been previously accepted as part of a exam for a qualification.

Erlangen, den 01.04.2022

(sign here)

Sergei Dobrovolskii