Lehrstuhl für Medizinische Biotechnologie

Technische Fakultät

Friedrich-Alexander-Universität Erlangen-Nürnberg

Design and evaluation of a Deep Learning approach to quantify synthetic volumetric autofluorescence data of immune cell infiltrate

Master’s Thesis

*In the program Advanced Optical Technologies  
to acquire the academic degree*

*Master of Science (M.Sc.)*

*Submitted by*

Sergei Dobrovolskii

Matrikelnummer: 22745251

Erlangen, date 19.09.2022

*Supervisors:*

Dr. Sebastian Schürmann

Dr. Daniel F. Gilbert

Prof. Dr. Dr. Oliver Friedrich

Zusammenfassung

**Design und Evaluierung eines Deep-Learning-Ansatzes zur Quantifizierung synthetischer volumetrischer Autofluoreszenzdaten von Immunzellinfiltraten**

Da die Biologie eine quantitative Disziplin ist, sind viele Beweise erforderlich, um eine Hypothese zu stützen. Bilder von Zellen und Geweben sind eine hervorragende Datenquelle, aber um aussagekräftige Erkenntnisse zu erhalten, müssen sie verarbeitet, analysiert und quantifiziert werden. Deep Neural Networks sind eine der Techniken, die im Datenanalyseprozess verwendet werden. Diese Netzwerke können bei Aufgaben wie der Bildsegmentierung und -klassifizierung bemerkenswerte Leistungen erbringen, benötigen dafür jedoch viele Trainingsdaten. Wenn Bilder dreidimensional sind, wird die Kennzeichnung von Trainingsdaten erheblich schwieriger. Um dieses Problem anzugehen, ist ein Simulator für synthetische Daten erforderlich, um dieses Problem zu lösen. Es kann unbegrenzt kommentierte Daten für neuronale Netze erstellen, um deren Leistung zu testen. Diese Arbeit entwickelt einen Simulationsrahmen, vergleicht seine Ergebnisse mit tatsächlichen Stapeln, die mit einem Multiphotonenmikroskop aufgenommen wurden, und trainiert ein tiefes Faltungsnetzwerk, das diese künstlichen Daten verwendet, um Immunzellen zu zählen und zu klassifizieren.

Abstract **Design and evaluation of a Deep Learning approach to quantify synthetic volumetric autofluorescence data of immune cell infiltrate**

As biology is a quantitative discipline, it requires a lot of evidence to support a hypothesis. Images of cells and tissues are an excellent source of data, but to get meaningful insights, they must be processed, analyzed, and quantified. Deep Neural Networks are one of the techniques used in the data analysis process. These networks may perform remarkably in tasks like picture segmentation and classification, but they need a lot of training data to do so. When images are three-dimensional, labeling training data becomes considerably more difficult. To address this issue, a synthetic data simulator is necessary to solve this problem. It can create unlimited annotated data for neural networks to test their performance. This thesis develops a simulation framework, compares its results to actual stacks acquired with a multiphoton microscope, and trains a deep convolutional network using this artificial data to count and classify immune cells.

Table of Content

[1 Introduction 8](#_Toc111840969)

[2 State of the Art 10](#_Toc111840970)

[3 Methods 11](#_Toc111840971)

[4 Results 12](#_Toc111840972)

[5 Discussion 13](#_Toc111840973)

[6 Conclusion 14](#_Toc111840974)

[7 References 15](#_Toc111840975)

[8 Appendix 16](#_Toc111840976)

Abbreviations

|  |  |
| --- | --- |
| MPM | Multiphoton Microscope |
| CNN | Convolutional Neural Network |
| H&E | Hematoxylin and Eosin |
| RF | Random Forest |
| DTC | Decision Tree Classifier |
| FAD | Flavin Adenine Dinucleotide |
|  |  |
|  |  |

# Introduction

In contemporary drug, therapy, or disease research, researchers not only need to develop a method, but they also need to be able to see what is going on, whether it be with a patient or with a test subject. First comes the hypotheses, then it must be tested via experiment, and this experiment must be somehow observed. The observation of experiment can be done via imaging modalities, such as conventional microscope, cell-counter, computed tomography or by naked eye. For investigation of small cells usually microscopy is a good modality. There is a plethora of different microscopic modalities, and in the scope of this project, a multiphoton microscope was used. It can establish three-dimensional image acquisition and utilize fluorescence for morphological sample understanding.

Acquired images themselves are useful at a first glance. The images can be examined by knowledgeable pathologists or biologists, and then a conclusion can be drawn from those observations. Because biology is a quantitative field of study, the significance of statistical analysis is something that must not be overlooked. To gain a deeper understanding of it, analysis is required to find more insights. Images need to be processed and analyzed before they can provide a better understanding of the experiment.

The problem of image processing and analysis can be solved using a variety of mathematical and software techniques. These techniques can either be traditional, such as thresholding, watershed, or region growing algorithms, or they can be addressed to more sophisticated techniques, such as decision tree, random forest, and support vector machines. All these methods are still in use today, but recent developments in deep learning have made it possible to perform image processing that is both more general and more concise.

The way these networks are working is not the same as for the traditional approaches. For Neural Networks no particular algorithm is being developed to perform image processing and classification, but network is being “trained” on datasets that consist of original images and the output that is required to get from the network given that image. The output is usually created manually, by hand, and this time-consuming process is named labeling. Network “learns” mapping between input and output, and hence can be used on unseen data to do the same. This makes it possible to process a great variety of images, which makes it superior to any of the traditional methods.

These networks may demonstrate remarkable performance in tasks such as image segmentation and classification, but they require a substantial amount of training data to do so. When images contain a third dimension, labeling the training data becomes considerably more difficult. A data simulation tool is required to find a solution to this issue. It is able to generate an unlimited amount of annotated data for use in testing the performance of neural networks. This simulator must generate images that resemble real ones and a random assortment based on parameter settings. The advantage of using a simulator instead of manually annotated data is data with flawless annotations and unlimited dataset capacity. Any experiment of image annotation and augmentation can be done in a matter of minutes, but not weeks.

In this dissertation, a simulation framework for FAD cytoplasm fluorescence in 3D is developed, its results are compared to actual stacks acquired with a multiphoton microscope, and a deep convolutional network is trained to count and classify immune cells using the simulated data.

# State of the Art

# Methods

# Results

# Discussion

# Conclusion

# References

**There are no sources in the current document.**

# Appendix

Curriculum vitae

|  |  |
| --- | --- |
| Personal Details |  |
| Name | **Sergei Dobrovolskii** |
| Date and place of birth | 03.02.1996 Moscow |
| Address | Koldestraße 10, 91052 Erlangen |
| Email | dobrik4ever@gmail.com |
| Phone number | +49 1516 3610896 |
| Education |  |
| 10/2019 – present | Master program Advanced Optical Technologies,  Friedrich Alexander Universität Erlangen-Nürnberg |
| 09/2017 – 07/2018 | *Bachelor of Science (B.Sc.)*  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 19.09.2022

(sign here)

Sergei Dobrovolskii

Acknowledgement