Digital Holographic Microscopy for High-throughput Analysis of Tumour Cells

TOMI PITKÄAHO M.Sc.

A thesis submitted for the degree of Doctor of Philosophy

Department of Computer Science

Maynooth University – National University of Ireland Maynooth

Maynooth, Co. Kildare, Ireland

October 2019

Supervisor: Prof. Thomas J. Naughton

Head of Department: Dr. Joseph Timoney

Contents

List of Figures

List of Tables

Isälle ja Viiville

Declaration

I hereby declare that this thesis is my own work and has not been submitted in any form for another award at any university or other institute of tertiary education. Information derived from the published and unpublished work of others has been acknowledged in the text and a list of references is given.

TOMI PITKÄAHO M.Sc.

October 2019

Acknowledgements

I would like to offer my sincerest gratitude to my supervisor Professor Thomas J. Naughton for his support, guidance, and patience. Prof. Naughton demanded high-quality research throughout this project, for which I am grateful and will not forget.

I would like to thank Dr. Aki Manninen and his research group from the University of Oulu, Finland where I visited several times during the project. I can not thank enough for Dr. Manninen's patience in explaining difficult biological concepts and his willingness to sacrifice his time for this project. Without him, his research group and our collaboration this project would not have been possible.

Thanks to Professor Ken Duffy, Hamilton Institute, Maynooth University, for helping with the statistical analysis of patient data.

In Maynooth, I would like to thank the Department of Computer Science for providing facilities that helped to accomplish this project. I would like to express my thanks to the postgraduates and staff. I wish to thank Taina, Lingfei, James and Hao, to name few.

I would like to acknowledge the financial contributions of Irish Research Council, Science Foundation Ireland, Maud Kuistila memorial foundation and Kerttu Saalasti Foundation.

I want to thank my parents for their continuous support and believing in me over the years of my studies.

To my wife, Tiina – thank you for all your support and encouragement. Finally to our lovely daughter, Viivi – daddy will be more home from now on.

Abstract

Cell imaging for cancer diagnosis typically requires invasive procedures to make cells visible, and in time-lapse imaging, light levels can be harmful for cells. These routines can affect the behaviour of cells and introduce unwanted artefacts in the captured images.

Digital holographic microscopy (DHM) enables stain-free single-shot imaging of living cells, uses low light intensities, and requires low data storage capacity. Due to advantages over many other methods, DHM is a strong candidate for high-throughput analyses. Use of DHM for imaging complex three-dimensional organoids has not been reported before.

The major limitation preventing widespread-use of DHM is the lack of analysis algorithms and limitations with those that are available. There are no algorithms to detect the presence of objects encoded in digital holograms. Existing algorithms used to find the in-focus depths of objects encoded in digital holograms require multiple reconstructions that increase processing time making them unsuitable for high-throughput analyses. There are no algorithms that could be used to segment digital hologram reconstructions of biological objects to multiple distinctive regions. To the best of our knowledge, using digital hologram reconstructions for classification of cysts has not been reported before.

This thesis introduces novel approaches for efficient analyses of holograms of cell lines and real patient samples. By analysing and interpreting different features of organoids, cancer-specific signatures are identified. In this thesis, a number of novel contributions are reported. A model-based object presence detection approach exploiting information extracted from a CNN is reported. CNNs are trained to find in-focus depths of organoids encoded in digital holograms without any numerical propagation. Reconstructions from

digital holograms of organoids are segmented to multiple discrete regions using a CNN, allowing novel quantitative analyses. Different classifiers using either extracted feature vectors or phase reconstructions are trained to discriminate healthy and tumorigenic organoids. A large-scale experiment is conducted for finding a CNN model with sufficient classification accuracy and minimum number of learning parameters; hand-crafted features are added to these shallow networks to improve the classification accuracy. Organoids derived from tissue samples of thirteen prostate cancer patients are shown to introduce additional challenges. Based on the existing data, there is an indication that prostate cancer is unique for each patient thus complicating detection of cancer.

List of publications

This list of publications contains all the publications published or submitted during my studies. Publications that are part of this thesis are in bold.

Journal publications

I. **Tomi Pitkäaho, Kai Zhang, Aki Manninen, and Thomas J. Naughton, "Quantitative live cell imaging and three-dimensional analysis of epithelial organoids with digital holographic microscopy," Nature Methods (x), x–x 2018. (in preparation)**

II. **Tomi Pitkäaho, Kai Zhang, Aki Manninen, and Thomas J. Naughton, "Distinguishing background fingerprints of digital holograms for convolutional neural network," Optics Letters (x), x–x 2019. (in preparation)**

III. **Tomi Pitkäaho, Aki Manninen, and Thomas J. Naughton, "Classification of digital holograms of oncogenic multicellular samples," Biomedical Optics Express x, x–x, 2019. (in preparation)**

IV. **Tomi Pitkäaho, Aki Manninen, and Thomas J. Naughton, "Detection of an object in the field-of-view of digital holograms," Optics Letters x, x–x, 2019. (in preparation)**

V. **Tomi Pitkäaho, Aki Manninen, and Thomas J. Naughton, "Digital hologram reconstruction segmentation of microscopic objects using a convolutional neural network," Optics Letters x, x–x, 2019. (in preparation)**

VI. **Tomi Pitkäaho, Aki Manninen, and Thomas J. Naughton, "Focus prediction in digital holographic microscopy using deep convolutional neural networks," Applied Optics 58, 5, A202–A208, 2019.**

VII. Tomi Pitkäaho, Ville Pitkäkangas, Mikko Niemelä, Sudheesh K. Rajput, Naveen K. Nischal and Thomas J. Naughton, "Space-variant video compression and processing in digital holographic microscopy sensor networks with application to potable water monitoring," Applied Optics, 57, 22, E190–E198, 2018.

VIII. T. Qaiser, A. Mukherjee, PB C. Reddy and S. D. Munugoti and V. Tallam, T. Pitkäaho, T Lehtimäki, T. Naughton, M Berseth and A. Pedraza, R. Mukundan, M. Smith, A. Bhalerao, E. Rodner, M. Simon, J. Denzler, C-H. Huang, G. Bueno, D. Snead, I. O. Ellis, M. Ilyas and N. Rajpoot, "HER2 challenge contest: a detailed assessment of automated HER2 scoring algorithms in whole slide images of breast cancer tissues," Histopathology, 72, 2, 227-238, 2017.

Conference publications

IX. **Tomi Pitkäaho, Aki Manninen and Thomas J. Naughton, "Digital hologram reconstruction segmentation using a convolutional neural network," Digital Holography & 3-D Imaging, Th3A.1, Bordeaux, France, 2019.**

X. **Tomi Pitkäaho, Aki Manninen and Thomas J. Naughton, "Detection of an object in the field of view of a digital hologram with an heuristic algorithm parameterized using a convolutional neural network," Digital Holography & 3-D Imaging, Th3A.3, Bordeaux, France, 2019.**

XI. **Tomi Pitkäaho, Aki Manninen and Thomas J. Naughton, "Temporal deep learning classification of digital hologram reconstructions of multicellular sample," Biophotonics Congress: Biomedical Optics, JW3A-14, Hollywood, Florida, USA, 2018.**

XII. **Tomi Pitkäaho, Aki Manninen and Thomas J. Naughton, "Principal component analysis on time-lapse captured digital holograms of cell clusters," Biophotonics Congress: Biomedical Optics, JW3A-68, Hollywood, Florida, USA, 2018.**

XIII. **Tomi Pitkäaho, Aki Manninen and Thomas J. Naughton, "Detecting the Presence of a Transparent Object in Off-Axis Digital Holograms," Digital Holography & 3-D Imaging, JM4A-42, Orlando, Florida, USA, 2018.**

XIV. **Tomi Pitkäaho, Aki Manninen and Thomas J. Naughton, "Classification of Digital Holograms with Deep Learning and Hand-Crafted Features," Digital Holography & 3-D Imaging, DW2F-3, Orlando, Florida, USA, 2018.**

XV. Tomi Pitkäaho, Ville Pitkäkangas, Mikko Niemelä, Sudheesh K. Rajput, Naveen K. Nishchal, and Thomas J. Naughton, "Digital holographic microscopy in remote potable water monitoring," Digital Holography & 3-D Imaging, Jeju-do, South Korea, 2017.

XVI. **Tomi Pitkäaho, Aki Manninen and Thomas J. Naughton, "Performance of autofocus capability of deep convolutional neural networks in digital holographic microscopy," Digital Holography & 3-D Imaging, Jeju-do, South Korea, 2017.**

XVII. **Tomi Pitkäaho, Aki Manninen and Thomas J. Naughton, "Focus classification in digital holographic microscopy using deep convolutional neural networks," Advances in Microscopic Imaging, Vol. 10414 of SPIE Proceedings (Optical Society of America), (E. Beaurepaire, ed.), 104140K, Munich, Germany, 2017.**

XVIII. Tomi Pitkäaho, Aki Manninen and Thomas J. Naughton, "Monitoring MDCK cell vesicles by digital holographic microscopy and image processing," 2016 Imaging and Applied Optics Congress, Optical Society of America, JW4A.21, Heidelberg, Germany, 2016.

XIX. **Tomi Pitkäaho, Aki Manninen and Thomas J. Naughton, "Digital Holographic Microscopy in Morphometric Analysis of Three-dimensional Cell Cultures," 2015 European Conference on Lasers and Electro-Optics - European Quantum Electronics Conference, Optical Society of America, CLEO_ECBO_1_3, Munich, Germany, 2015.**

1 Introduction

1.1 Background and motivation

Cancer is and will remain one of the biggest global challenges concerning human health and is one of the leading causes of death. For instance in Europe every year, nearly four million Europeans are diagnosed with cancer [1]. Early detection and diagnosis of cancer and subtypes thereof are key factors for effective treatment. At present, cancer detection and assessments of at-risk patients are mainly conducted by blood and tumour tissue tests, urine tests, X-ray scans, CT scans, and endoscopy. The clinical diagnosis and the treatment plan is largely based on a subjective expert evaluation. For example, a pathologist estimates aggressiveness of a cancer based on histological evaluation of a tissue sample and the treatment plan is largely derived from this. This is a typical example of a traditional medical treatment, in which less consideration is given to the differences between individuals. As the number of different drugs and treatments increase, it is of great importance to find the best possible treatment plan for an individual patient. This fact is partially emphasized by a transition from traditional medical treatment to personalized precision medicine [2–5]. This approach allows clinicians to more accurately predict which treatment for a particular disease is the best possible for each individual. Reducing the number of unnecessary treatments and improving drug responses via tailored treatment plans can drastically reduce health care costs and significantly improve the well-being of patients living with cancer.

In response to this challenge, one possibility is to analyse tumour cells [6] and their response to different drugs. A tumour cell in the context of this thesis is defined to be a cell

that has abnormal cell growth and abnormal cell division properties. Over time, these cells typically form abnormal structures differentiating them from their healthy counterparts. Based on the quantitative analysis of a bank of tumour cells, each exposed to a different drug, the best possible drug for an individual could be chosen. Unfortunately, there is no standard test available that works at the cellular level, observing a cell and its response to different drugs. A natural way to observe a cell is through imaging. For this kind of imaging to be meaningful and also reliable, we define the system to enable:

- high-throughput (thousands of images per day),
- quantitative measurements,
- non-invasive measurements (low light levels),
- freedom from tagging,
- long term observation, and
- freely moving cells in a sample.

Conventional microscopes require tagging (labelling) the sample, which can be a highly invasive procedure that affects the behaviour of the cells. Typically, tagged cells are imaged with confocal microscopes that are based on scanning, which can limit the size and movement of cells inside a sample. Also, they require a high amount of light that may rapidly impair cell viability and morphology [7, 8]. Electron microscopes require the object to be fixed and sputtered with metal, which kills the sample. Phase-contrast techniques such as Zernike, Normarsky, and dark-field microscopies only increase the visibility of the internals of the cell and cannot be used to infer any quantitative three-dimensional (3D) information about the cell. X-ray microscopes can achieve nanometre resolution, but again they will kill a living sample.

Digital holographic microscopy (DHM) [9] and other interferometry based microscopy techniques, in principle, are the only *ex vivo* high-resolution technologies capable of capturing and thus following dynamics of a living sample completely non-invasively and without introducing artefacts and ambiguities in the measurements. DHM enables the capture of quantitative transmitted phase of an object over long periods of time, does not

require any labelling, uses small light intensities (approximately 1000 times less than in a confocal microscope), allows free movement of a sample within the field of view, and has true potential in high-throughput imaging. The suitability of DHM for high-throughput imaging is supported by the fact that the imaging is not based on any form of scanning; a single hologram capture time is limited only by the integration time of the imaging sensor. Due to this property, another advantage of DHM is that it produces a small amount of data compared with other 3D microscopy techniques.

In light of these properties and advantages, DHM as a technology is a perfect candidate for high-throughput imaging and analysis. However, a severe lack of algorithms exists. For this technology to really be usable for clinicians, biologists, and other medical experts, there need to be algorithms that can be used to analyse holographic data. This thesis decreases this gap by implementing novel solutions that can be used together with data captured by sophisticated hardware to do analysis for a range of purposes.

1.2 Problem statement

The problem tackled in this thesis is how to build a high-throughput three-dimensional analysis system for cysts and organoids.

The most pressing need for any solution to this problem is reliable processing and analysis software components. There are no available computational methods for cancer detection that use the morphology of cancer cells. This thesis provides a new dimension to cancer cell detection and characterisation with a rapid and robust *ex vivo* microscopical assessment of tumour cell morphology. Research methods include imaging, observation, and analysis conducted at the cellular level by use of DHM. Crucial solutions required for a high-throughput imaging system are identified, defined, implemented, and tested.

In addition to the general problem statement of what is needed to build a high-throughput analysis system comprising the optical imaging system and processing and analysis algorithms, we are interested in:

- Can we extract information/features from digital holograms of biological samples that can be used for high accuracy two-class classification (healthy and oncogenic)?

- If so, which and how many features are needed for reliable classification?
- Can we improve two-class classification with deep learning?
- If so, how deep do convolutional neural network models need to be to solve the problem?
- If we need convolutional neural networks, how much data is needed?

1.3 Scope and contribution of the thesis

This thesis forms a first stage for exploring the possibilities of DHM for high-throughput analysis in cancer research. The technology enables capturing the phase information of transparent objects that can potentially be used in identification and morphological characterisation of tumourigenic cells. By analysing and interpreting morphological features of living samples, that manifest themselves in the quantitative optical density, cancer-specific signatures will be identified and will inform our software development.

The aim of this thesis is to design a method for high-throughput screening (allowing to analyse thousands of images per day) and increased knowledge of cell-with-cancer specific features and structural changes through cell morphology and physiological quantitative measurements. In addition, possibilities to use sophisticated machine learning algorithms will be explored. The main focus is on high-throughput imaging, cancer cell detection and characterisation, which may display various specific features like distorted cell division. The ultimate goal of the thesis is to form a high-throughput analysis method and design reliable classifiers of morphological phase signatures which can be used to detect cancer.

This thesis investigates possibilities for combining image processing, image analysis, and DHM. Novel image processing approaches, that can be used for processing and feature extraction of cells, are introduced. These extracted features are further used with machine learning algorithms in order to discriminate between healthy and tumourigenic cells, and moreover, solutions presented in this thesis can be used to quantify the tumourigenic potential of cells. Our convolutional neural network (CNN) design decisions are summarised in Table 1.1.

Table 1.1 Table of CNN design decisions.

Decision	Description
Size normalisation	Input to a CNN is scaled to a proper size as designed by the CNN. This allows magnification independent CNN training.
Wavelength normalisation	Inputs to a CNN are normalised to allow wavelength independent CNN training.
Off-the-shelf architectures	Where possible, off-the-shelf CNN architectures such as AlexNet are used.
Minimal pre-processing	Minimum amount of numerical pre-processing giving a sufficient result is applied to holograms before feeding to a CNN.

Size normalisation - digital holographic microscopes can be used with different microscope objectives that change the magnification of objects. As the CNNs are not scale invariant, due to nature of the convolution operation, this changes the perceived size. Therefore, to be able to do inference using a trained network with digital holograms captured with different microscope objectives, it is important to scale the objects to the same size as the samples used to train the network.

Wavelength normalisation - digital holographic microscopes operate at various different wavelengths. The number of phase wraps in phase reconstructions are dependent on the used wavelength, i.e. a wrapped phase reconstruction of a digital hologram of an object captured with a short wavelength displays more phase wraps than the same object captured with a longer wavelength. This information is transferred to the unwrapped phase reconstruction and phase values would be higher for the unwrapped phase reconstruction captured using the shorter wavelength. Therefore, to be able to do inference using a trained network with digital holograms captured with a different wavelength, it is important to normalise unwrapped input values to the same range that they were while training the network.

Off-the-shelf architectures - well-known, widely tested and accepted, and well performing CNN models are preferred whenever possible. These existing models have been proven to work with images of real-world objects. Minimal changes are applied to these models so that they can work on holographic data.

Minimal pre-processing - before feeding a hologram input to the network, minimal pre-processing is applied. For example, inputs to the CNNs that were trained to find the

in-focus depth were hologram-plane amplitude. Minimal pre-processing will increase the throughput of the system.

Throughout the thesis, two main themes are present: simplification and paradigms. Simplification in the context of this thesis is defined as: "Design and implementation of solutions and model architectures in a simplified manner using basic elements whenever possible; entirely without reducing the efficiency". Simplification is shown with different design choices made in the thesis: CNN models are successively simplified, training data is reduced, and the number of features used for training is reduced.

Paradigms in the context of this thesis is defined as: "Model- and learning-based approaches are used in collaboration to improve the performance of the other." This theme is used both ways; model-based approaches are used to improve learning-based approaches and vice versa.

1.4 Organisation of the thesis

Figure 1.1 shows identified components of an automated high-throughput imaging system. These three components need to be addressed for a fully functional system. Each of these components can be split to smaller problems that need to be answered (Fig. 1.2). This thesis provides solutions to most of the identified problems, but not to all (see Fig. 1.2).

The imaging method relies on holography and its principles. The algorithms, in turn, rely strongly on different machine learning algorithms. Chapter 2 contains a short introduction to both of these concepts. In addition, this chapter contains a literature review of machine learning approaches used in conventional microscopy and quantitative phase imaging (including DHM). At the end of the chapter, biological aspects of the thesis are introduced.

Chapter 3 describes optical hardware that is needed to implement an automated high-throughput imaging system. In addition, it defines the digital holographic system that was used for the experiments. The software forms the core of this thesis. Chapter 4 contains first set of machine learning algorithms used for classification. Chapter 5 contains a novel solution used to detect if an object is in the field of view captured in a digital hologram.

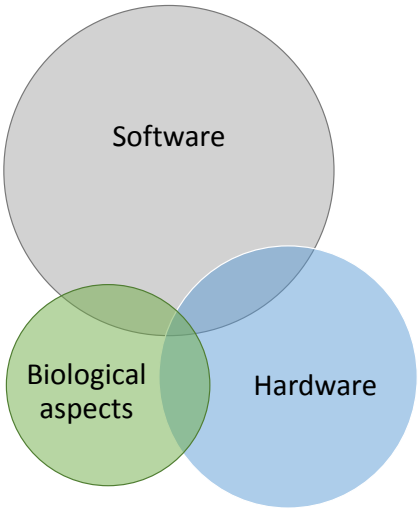

Figure 1.1 The main components of an automated high-throughput imaging system. Such a system is a collection of interlocked components that work in collaboration. Different sizes represent their importance in this thesis. The emphasis is on software while biological aspects and hardware, have important roles in a high-throughput analysis system. Hardware consists of optical, mechanical, and computer hardware (realisations of trained classifiers).

Chapter 6 describes a novel approach used for finding the in-focus depth of objects encoded in digital holograms. Chapter 7 introduces another novel approach that can be used to segment objects encoded in digital holograms.

Chapter 8 introduces a shallow convolutional neural network architecture that is used to classify holograms of oncogenic and healthy living samples. The suitability of the system for time-lapse imaging and real patient samples is demonstrated in Chapter 9.

Chapter 10 summarises this thesis with sections for discussion and future work.

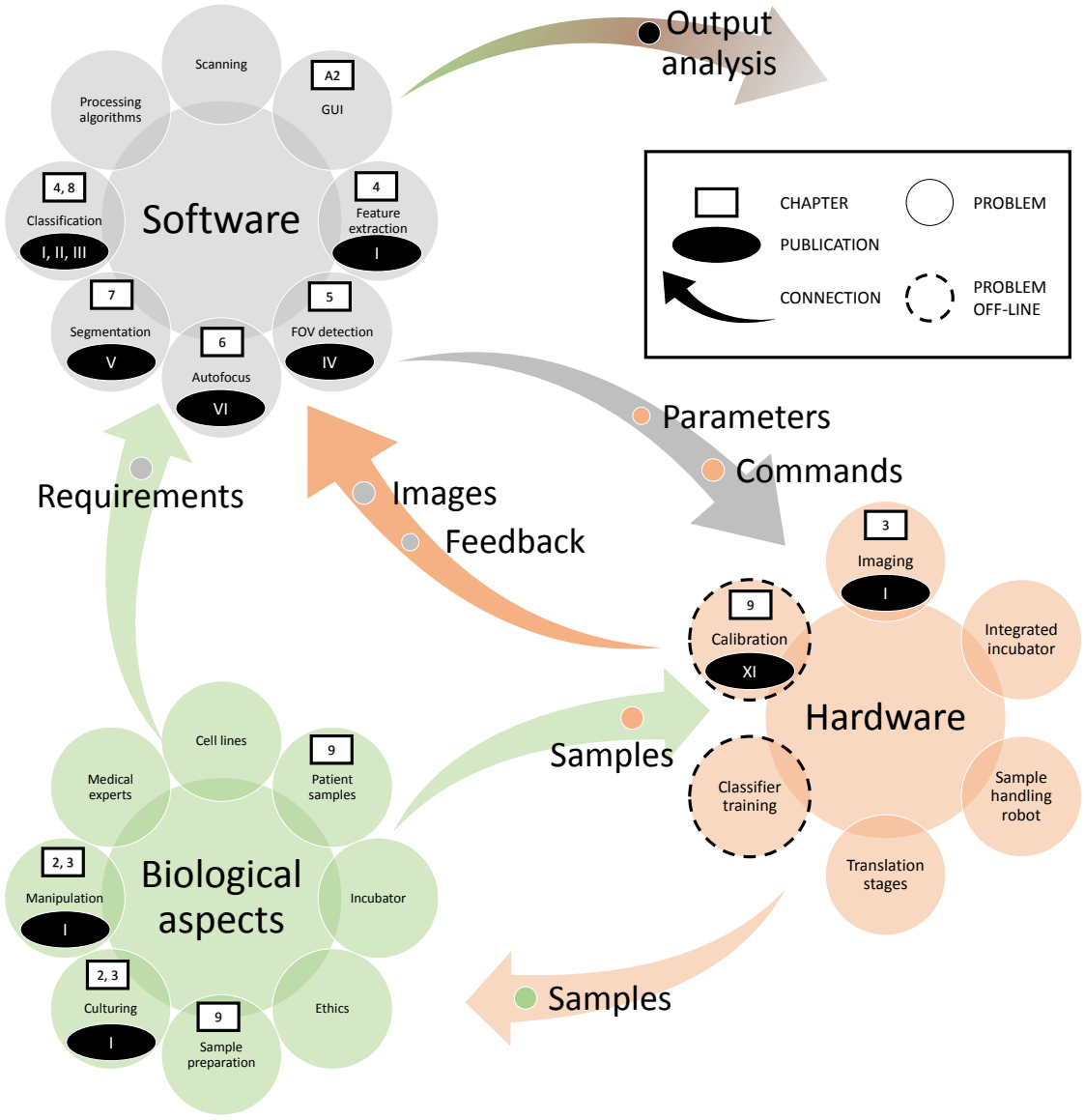

Figure 1.2 Automated high-throughput imaging system. Three main components contain smaller problems that need to be solved for a fully functional automated high-throughput system. The figure displays chapter numbers where chapters contain solutions to particular problems. In addition, publications are marked. Most of the system works online, however, the hardware calibration and classifier training can be accomplished off-line. The problems with either a chapter number or a publication number have been solved in this thesis.

2 Background

This chapter gives a short introduction to holography and machine learning. The section on holography follows the chronological development of the science and technology that took approximately 50 years. The last section of this chapter is about biological aspects related to this thesis.

2.1 Introduction

The background of holography is in electron microscopy, a field that Dennis Gabor worked on. Gabor [10] was working on improving the resolution of the electron microscope when he invented holography, work for which he was awarded the Nobel prize more than two decades later. From his early work on conventional analogue holography, holography has evolved into digital holography and digital holographic microscopy. Nevertheless, the basic principle remains the same – this technology enables the capture of a full wavefront.

One definition of machine learning given by Tom Mitchell is: "Well posed learning problem: a computer program is said to learn from experience E with respect to some task T and some performance measure P, if its performance on T, as measured by P, improves with experience E" [11]. Machine learning itself has a rather long history starting from the work of Karl Pearson [12], Ronald A. Fisher [13], and Arthur Samuel [14]. Frank Rosenblatt [15] designed the perceptron, the first image recognition machine in 1957. Ten years later in 1967, the nearest neighbour algorithm was published [16]. From these early discoveries, machine learning has evolved new approaches such as support vector machines, random forests, convolutional neural networks, and boosting, amongst others.

Interestingly, holography has a history as an enabling technology for artificial neural networks (ANNs) [17], and ANNs have been applied before in the fields of digital holography [18, 19], digital holographic microscopy [20, 21] and more generally in quantitative phase imaging [22]. These are just examples, as many other machine learning algorithms and approaches have been applied in the field of digital holography. These algorithms are important aspects of the modern digital holography and are necessary for efficient hologram processing and analysis.

2.2 Holography

2.2.1 Introduction to holography

Gabor's idea was to have two interfering waves and record the formed interference pattern on a photographic medium [10]. The first one of the waves, the object wave, is altered by the object under investigation and the second wave, the reference wave, goes through the volume unaltered (Fig. 2.1). The recorded hologram contains both amplitude and phase information of the wavefront. Mathematically, a hologram, I, which is an interference pattern formed by the reference wave, U_R, and the object wave, U_O, can be expressed as

$$I = |U_O + U_R|^2 = U_O U_R^* + |U_R|^2 + |U_O|^2 + U_O^* U_R \tag{2.1}$$

where $*$ is the complex conjugate, $|U_R|^2$ and $|U_O|^2$ denote the dc terms, and $U_O U_R^*$ and $U_O^* U_R$ are the conjugate and real images, respectively.

Later the recording process can be inverted, and a reconstruction obtained from the hologram by illuminating the hologram with an optical imitation of the original reference wave (Fig. 2.2).

The two most important properties of a conventional reconstruction (a consequence of recording the quantitative phase) are as follows:

- The viewer can naturally focus at any depth of the scene.
- The viewer can view the scene from different perspectives (limited by the hologram size and the illumination angle).

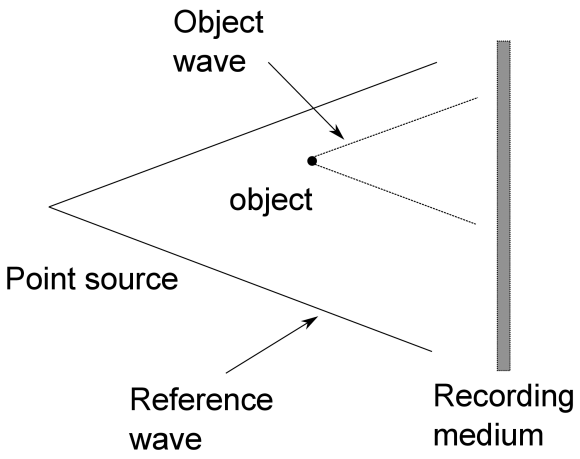

Figure 2.1 Recording a hologram as explained by Gabor [10]. The hologram is formed from the object and the reference waves interfering. See text for more explanation.

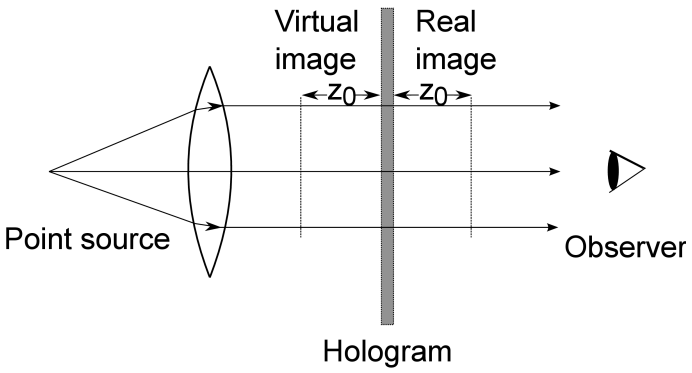

Figure 2.2 Reconstructing from a Gabor hologram. The hologram is reconstructed by illuminating the hologram with the original reference wave. An observer is able to view two identical replicas of the reconstructed object at equal but opposite distances, z_0, from the hologram.

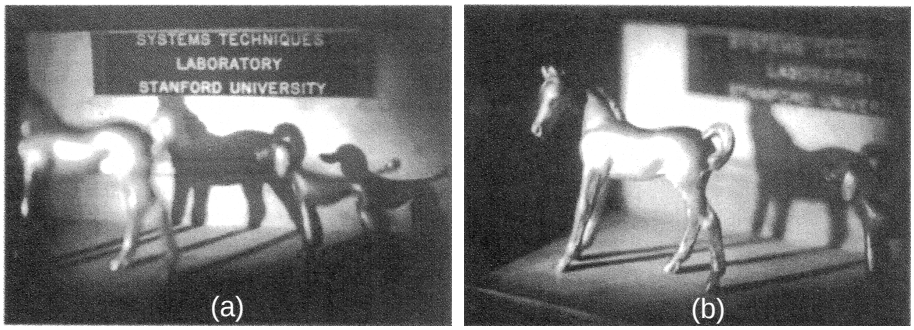

Figure 2.3 Special properties of a hologram. This figure shows two photographs captured of a reconstructed hologram. (a) background in focus, viewed from left, (b) front in focus, viewed from right. Figure adapted from [23].

Fig. 2.3 shows hologram reconstructions which demonstrate both of these properties.

The original recording approach, in-line holography, as invented and described by Gabor has one major challenge; the reconstructed wavefront contains two exact copies of the object, twin images (a virtual and a real image), at the same but opposite distances from the hologram plane (Fig. 2.2). This is due to the fact that the virtual image is in the opposite phase to the real image. When viewing the reconstruction, the in-focus real image is partially occluded by the out-of-focus twin. To overcome this problem, when the first commercial helium-neon lasers became available, Leith and Upatnieks [24] invented off-axis holography in the early 1960s. The off-axis approach introduces an angle between the reference and the object wave (Fig. 2.4). This method enables spatial separation of the virtual and the real image in the reconstruction (Fig. 2.5). In addition, the zero-order, which is located at the centre of the reconstruction in an in-line hologram, can be spatially separated.

Fourier-transform holography was invented by Stroke and Falconer [25]. The difference with the original idea by Gabor is that a reference point-source is placed at the object plane (Fig. 2.6). After capturing a hologram with this approach, a reconstruction from the hologram can be accomplished by taking a Fourier transform of the hologram. Both twins appear in-focus at the object plane.

2.2.2 Digital holography

The first digital holograms were computer generated and made by Brown and Lohman in 1966 [26]. They printed holograms on a computer line printer and reconstructed them op-

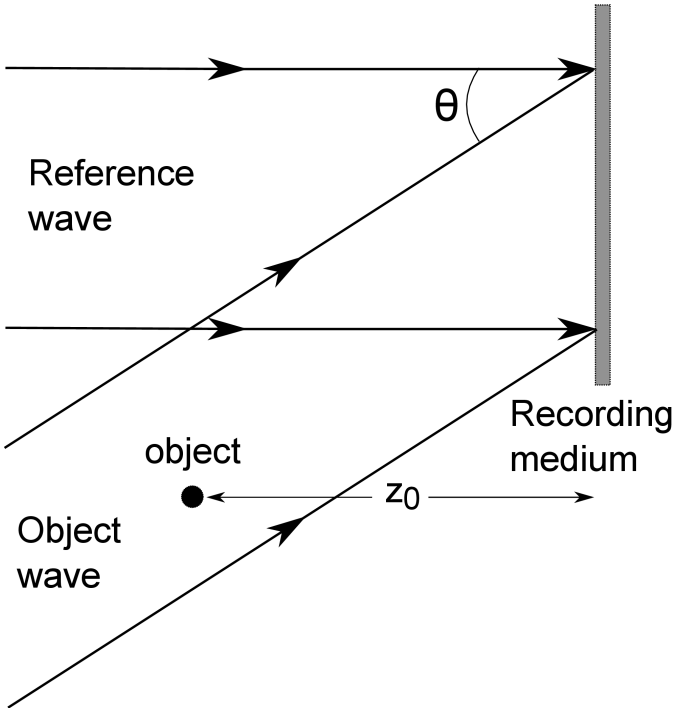

Figure 2.4 Off-axis setup. The reference and the object waves are separated by an angle allowing spatial separation of different hologram terms.

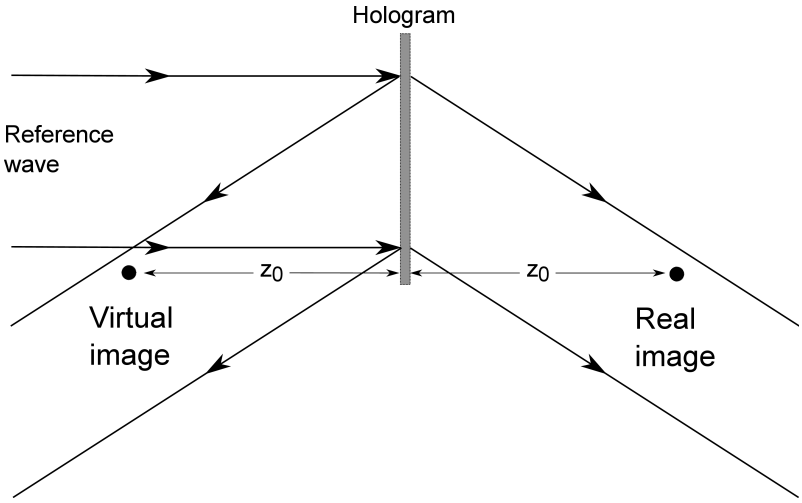

Figure 2.5 Off-axis reconstruction. The process to reconstruct from a hologram is similar to Gabor's. However, due to the separation angle of two waves in capturing, different hologram terms are spatially separated allowing to view the reconstructed object free of disturbance of the out-of-focus twin.

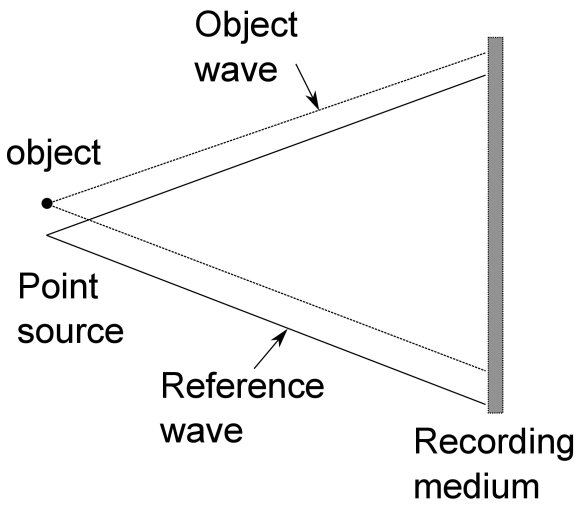

Figure 2.6 Recording a Fourier hologram. A point source is located on the same plane with the object.

tically by using coherent laser illumination. Shortly after this, digitally captured holograms were produced by Goodman and Lawrence [27]. They captured Fourier holograms by using a vidicon with the front lens removed. The reconstruction (numerical two-dimensional Fourier transform) was performed by a PDP 6 computer. Later Goodman introduced an efficient numerical reconstruction approach that can be used on Fresnel holograms [28]. A digital hologram, $I(x,y)$ can be reconstructed at any depth z as

$$U(x,y;z) = \frac{-\mathrm{i}}{\lambda z}\exp(\mathrm{i}kz)\, I(x,y) \otimes \exp\left(\mathrm{i}\pi\frac{x^2+y^2}{\lambda z}\right), \tag{2.2}$$

where λ is the wavelength of the light, $\otimes$ denotes a convolution operation and $k = 2\pi/\lambda$. Eq. 2.2 can be solved numerically with the angular spectrum method by using discrete Fourier transforms (DFTs) as

$$U(x,y;z) = \mathscr{F}^{-1}\left\{\mathscr{F}\left[I(x,y)\right]\mathscr{F}\left[h(x,y;z)\right]\right\} \tag{2.3}$$

where $\mathscr{F}$ and $\mathscr{F}^{-1}$ denote the DFT and the inverse DFT, respectively; x and y denote the spatial discrete coordinates; and $h(x,y;z)$ is the diffraction kernel defined as

$$\begin{aligned} h(x,y;z) &= \frac{-1}{2\pi}\frac{\partial}{\partial z}\frac{\exp(jkr)}{r} \\ &= \frac{-z}{2\pi}\left(jk-\frac{1}{r}\right)\frac{\exp(jkr)}{r^2} \end{aligned} \tag{2.4}$$

where r is

$$r = (x^2 + y^2 + z^2)^{\frac{1}{2}} \tag{2.5}$$

The DFT of the $h(x,y;z)$ can be replaced with [23]

$$\begin{aligned} \mathscr{F}\{h(x,y;z)\} &= F(f_x, f_y; z) \\ &= \exp\left\{-jkz[1-(\lambda f_x)^2 - (\lambda f_y)^2]^{\frac{1}{2}}\right\} \end{aligned} \tag{2.6}$$

where f_x and f_y are the spatial frequencies. From the complex-valued reconstruction, $U(x,y;z)$, an intensity reconstruction, $O(x,y;z)$, can be extracted as its absolute square

$$O(x,y;z) = \{\mathrm{Re}[U(x,y;z)]\}^2 + \{\mathrm{Im}[U(x,y;z)]\}^2 \tag{2.7}$$

and phase reconstruction can be extracted as

$$\Phi(x,y;z) = \text{arc tan}\left\{\frac{\mathrm{Im}[U(x,y;z)]}{\mathrm{Re}[U(x,y;z)]},\right\} \tag{2.8}$$

where the arc tan function is implemented computationally using atan2.

Yaroslavskii and Merzlyakov introduced the theory of digital holography in 1977 [29], however, it was the technological evolution in the '90s that brought digital holography to wider use. Digital recording became more efficient due to the evolution of imaging sensors to their current state. The first recordings on a charge-coupled device (CCD) was reported by Schnars and Jüptner [30]. Pixels in such sensors became small enough, and sensor sizes were large enough for a sufficient sampling for a hologram. Also, frame rates were high enough to take holographic video recordings of moving objects. In addition, the increased calculation power of computers enabled numerical reconstructions from a digital hologram to be performed in a reasonable amount of time.

One important discovery that provided a solution to "the twin problem" of in-line holograms was the development of phase shifting digital holography (PSDH) by Yamaguchi and Zhang in 1997 [31]. In PSDH one captures multiple holograms of the same object each with a different phase. The phase is typically altered (shifted) by using a piezoelectric mirror or another precise controller to change the length of one of the paths along the

optical axis of the system in a holographic capture setup; usually the reference path is altered. Let us consider one example: four equally spaced phase shifts. The complex-valued hologram $U(x,y)$ is computed from a sequence of four interferograms $I(x,y;0)$, $I(x,y;\frac{\pi}{2})$, $I(x,y;\pi)$, and $I(x,y;\frac{3\pi}{2})$, each differing from the previous by a phase shift of $\pi/2$, according to [31, 32] with reference R and object O waves as

$$\begin{aligned} U(x,y) &= \frac{1}{4U_{R_0}^*}\left\{[I(x,y;0)-I(x,y;\pi)]+\mathrm{i}[I(x,y;\frac{\pi}{2})-I(x,y;\frac{3\pi}{2})]\right\} \\ &= \frac{1}{4U_{R_0}^*}\Big[\,(|U_{R_0}|^2+|U_O|^2+U_{R_0}U_O^*+U_{R_0}^*U_O-|U_{R_\pi}|^2-|U_O|^2 \\ &\quad -U_{R_\pi}U_O^*-U_{R_\pi}^*U_O)+\mathrm{i}(|U_{R_{\frac{\pi}{2}}}|^2+|U_O|^2+U_{R_{\frac{\pi}{2}}}U_O^*+U_{R_{\frac{\pi}{2}}}^*U_O \\ &\quad -|U_{R_{\frac{3\pi}{2}}}|^2-|U_O|^2-U_{R_{\frac{3\pi}{2}}}U_O^*-U_{R_{\frac{3\pi}{2}}}^*U_O)\Big] \\ &= \frac{1}{4U_{R_0}^*}\Big[\,(U_{R_0}U_O^*+U_{R_0}^*U_O-U_{R_\pi}U_O^*-U_{R_\pi}^*U_O) \\ &\quad +\mathrm{i}(U_{R_{\frac{\pi}{2}}}U_O^*+U_{R_{\frac{\pi}{2}}}^*U_O-U_{R_{\frac{3\pi}{2}}}O^*-U_{R_{\frac{3\pi}{2}}}^*U_O)\Big] \end{aligned} \tag{2.9}$$

which, using $\mathrm{Re}(A)=\frac{A+A^*}{2}$ and the complex number property $(AB)^*=A^*B^*$, can be expressed as

$$U(x,y)=\frac{1}{4RU_0^*}\left\{[2\mathrm{Re}(U_{R_0}^*U_O)-2\mathrm{Re}(U_{R_\pi}^*U_O)]+\mathrm{i}[2\mathrm{Re}(U_{R_{\frac{\pi}{2}}}^*U_O)-2\mathrm{Re}(U_{R_{\frac{3\pi}{2}}}^*U_O)]\right\}. \tag{2.10}$$

By using the properties $U_{R_\pi}=-U_{R_0}$ and $U_{R_{\frac{3\pi}{2}}}=-U_{R_{\frac{\pi}{2}}}$ equation (2.10) can be rewritten as

$$U(x,y)=\frac{1}{4U_{R_0}^*}\left[4\mathrm{Re}(U_{R_0}^*U_O)+\mathrm{i}4\mathrm{Re}(U_{R_{\frac{\pi}{2}}}U_O^*)\right]. \tag{2.11}$$

With the complex number property of $U_{R_{\frac{\pi}{2}}}U_O^*=\mathrm{Im}(U_{R_0}^*U_O)+\mathrm{i}\mathrm{Re}(U_{R_0}U_O^*)$ this can be rewritten as

$$U(x,y)=\frac{1}{U_{R_0}^*}\left[\mathrm{Re}(U_{R_0}^*U_O)+\mathrm{i}\mathrm{Im}(U_{R_0}^*U_O)\right]=U_O\,, \tag{2.12}$$

which is the full complex-valued description of the object wave.

This shows how PSDH gives a representation of the full complex-valued Fresnel field in the plane of the camera, free of the dc and the second twin. The benefit of this approach over an off-axis holography approach is that the full spatial size of the imaging sensor can

be used to sample a hologram. As the object wave is assumed to be identical between each interferogram, some downsides to the approach are that it is sensitive to vibrations and fast moving objects pose a challenge for imaging. PSHD is not used in this thesis.

Digital holograms have clear benefits and disadvantages compared with traditional glass-plate holograms. No chemical processing is required after the capture. Digital holograms can be transmitted over networks, and they can be numerically processed (e.g. noise removal, compression). Although pixel sizes are relatively small and imaging sensors are physically large, there is still a vast difference in the optical quality of digital and conventional holograms. While digital camera pixel sizes are a few micrometres, grains in a holographic film are hundreds of nanometres in length. This means that conventional holograms have a higher sampling density and therefore a wider viewing angle. Also, imaging sensors are smaller in size than conventional holographic recording mediums and therefore are limited to imaging relatively small objects. Still, digital holograms, if captured at a large-enough distance, contain the same special properties as conventional holograms. They can be viewed from multiple perspectives, by reconstructing a different portion of the hologram. In addition, any plane of a volume can be viewed (reconstructed) numerically.

In addition to numerical reconstruction, digital holograms can be reconstructed optically by using spatial light modulators [33, 34]. This is out of scope for this thesis, and will not be covered here.

2.2.3 Digital holographic microscopy

Haddad et al. introduced a Fourier-transform holographic microscope in 1992 [35]. This system used a drop of glycerol on top of a microscope cover slide to enable the formation of Fourier holograms. An improvement to this was made by Boyer et al. in 1996 when they replaced the glycerol drop with a reflecting metal sphere [36]. Zhang and Yamaguchi proposed to use phase-shifting digital holography in microscopy [37]. These approaches allow the use of the reconstructed intensity to observe microscopic samples.

The first proposal to use the reconstructed phase was by Schnars in 1994 who proposed to use hologram interferometry with digitally recorded holograms [38].

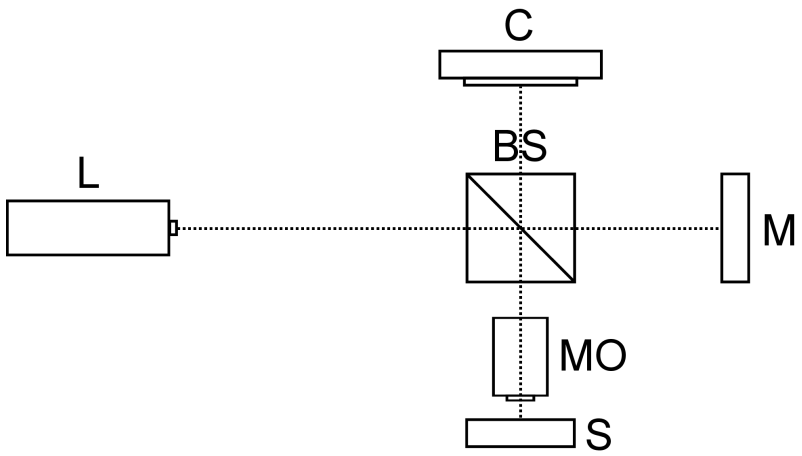

Figure 2.7 Principle of Michelson interferometry for reflecting objects. Collimated light from a laser (L) is split by a beamsplitter (BS) into an object wave and a reference wave. The reference wave is reflected back to the beamsplitter by a reference mirror (M). The object wave goes through a microscope objective (MO) to the sample (S) and is reflected back from the sample through the microscope objective. The two beams are combined at the beamsplitter and the formed interference pattern is recorded by the camera (C).

In 1999, Cuche et al. introduced an application of off-axis digital holography that could use the reconstructed phase of a single hologram for profile measurements of reflective objects [9]. Later, this was extended to microscopy (digital holographic microscopy, DHM) and also to transparent objects, enabling thickness measurements of transparent samples (if refractive indices of the sample and medium are known) based on the reconstructed phase from a single hologram [39]. In this type of DHM, the magnification is realised by using a microscope objective. A typical optical set-up for capturing reflective objects with this approach is a Michelson interferometer (Fig. 2.7,) and a Mach-Zehnder interferometer (Fig. 2.8) can be used for transmissive objects.

Digital in-line holographic microscopy (DIHM) follows the original idea by Gabor and was introduced by Kreuzer et al. [40] in 1999 (Fig. 2.9). In DIHM, a coherent light source is focused onto a pinhole producing a spherical wave that goes through a volume containing a sparse collection of objects. Light going through objects forms an object wave and the light going through the volume without interacting with the objects forms a reference wave. The light emerging from a pinhole contains a natural magnification (a natural free-space propagation) as the beam expands as a function of distance. Objects positioned closer to the pinhole therefore have higher magnification (and the system resolution is also better for these objects) [41]. The greatest benefit of DIHM, if compared to systems using imaging lenses, is its simplicity; a microscope can be assembled by using just three components:

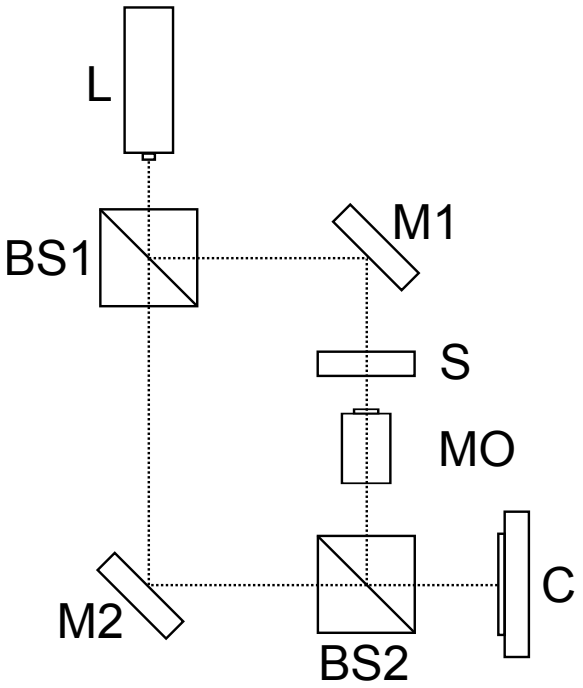

Figure 2.8 Principle of Mach-Zehnder interferometry for transparent objects. Collimated light from a laser (L) is split by a beamsplitter (BS1) to an object wave and a reference wave. The object wave is steered by a mirror (M1) through the sample (S) and is magnified by a microscope objective (MO). The reference wave is steered by a mirror (M2) to a beamsplitter (BS2) where the two beams are combined and the formed interference pattern is recorded by the camera (C).

a light source, a pinhole, and an imaging sensor. Typically a microscope objective is used to focus light onto a pinhole, however, the objective is not used for imaging. This type of setup is referred to as lensless. The downside of this approach is that for the reconstructed phase to be meaningful, relatively heavy numerical processing is required. This is due to the fact that the phase of the out-of-focus twin corrupts or disturbs the in-focus phase. However, with weakly scattering objects the phase is usable even without intense processing [42].

In-line holographic microscopy can also be realised by using a microscope objective for imaging. In this approach, all the components are in-line and the reference and object waves are formed as described in the original idea by Gabor. An imaging lens (microscope objective) is placed in front of the imaging sensor, after the object/volume (Fig. 2.10).

Because of the near proximity of the sensor to the object in DHM, DHM does not have a conventional holography property of viewing the object from different perspectives. The most modern approach to DHM is tomographic DHM [43–48] where the object is rotated or the angle of illumination is varied, to enable the capture of multiple holograms

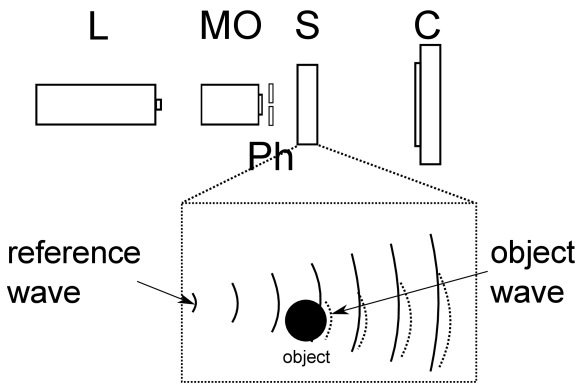

Figure 2.9 Gabor interferometer with a free space propagation. Collimated light from a laser (L) is focused by a microscope objective (MO) onto a pinhole (Ph) forming a spherical wave that goes through a sample (S). The formed interference pattern is captured by a camera (C). The inset shows how the object wave and the reference wave are formed.

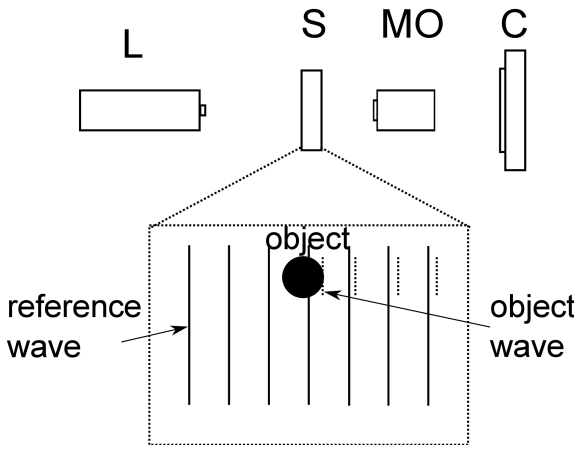

Figure 2.10 Gabor interferometer with a microscope objective magnification. Collimated light from a laser (L) is directed through a sample. The formed interference pattern is magnified by the microscope objective (MO) and captured by a camera (C). The inset shows how the object wave is formed by the reference wave.

at different object orientations. By processing captured holograms, a full 3D refractive index map can be produced. The greatest benefit of this approach is that the object can be viewed from different angles revealing its internal structures. This approach has been used successfully to increase the imaging resolution of a DHM system beyond the optical diffraction limit [49].

2.2.4 Digital holographic image processing

Digital holograms can be modified with digital processing, a process that can be called digital holographic image processing. Image processing includes noise removal, contrast enhancement, edge detection, segmentation, and compression among other operations. The output of image processing is another image. In digital holography the standard processing of a digital hologram before propagation is:

1. dc terms and conjugate image removal
2. aberration removal

The out-of-focus twin (conjugate image) and dc terms must be removed from a digital hologram reconstruction for the reconstructed phase to be usable. Kreis and Jüptner [50] showed that the dc can be removed by subtracting the average intensity from a hologram. Subtracting a simulated (computer-generated) reference wave intensity from an off-axis hologram in order to suppress dc term was proposed by Chen et al. [51]. Zhang et al. [52] used Sobel differential gradient algorithm to remove the dc. These procedures, of course, do not remove the conjugate image.

In an off-axis setup the dc term and the conjugate image removal operation can be realised by taking a Fourier transform of a digital hologram, removing the frequencies representing the dc and one of the twins, centring frequencies that belong to the other twin, and taking an inverse Fourier transform as proposed by Cuche et al. [53]. This is possible as different terms are spatially separated in the Fourier domain. An automated method to remove these unwanted terms was proposed by He et al. [54]. They used iterative thresholding in the Fourier plane to find three blobs belonging to different terms.

A microscope objective used for magnifying an object introduces phase aberrations in the reconstructed phase. These aberrations can be removed numerically by carefully

adjusting reconstruction parameters as proposed by Cuche et al. [39]. Ferraro et al. [55] proposed three methods to remove phase curvature: extrapolating extracted information of an object free region, estimating a correcting wavefront, and double-exposure. Colomb et al. [56] proposed a method where known flat regions of a phase reconstruction are manually selected and extracted as a reference for aberration-free sections. Information in these regions combined with polynomial fitting was used to remove aberrations. Zuo et al. [57] proposed a method where aberrations could be removed automatically without prior knowledge of the scene by using principal component analysis. Goldstein and Creath [58] proposed to use the square of the phase gradient map to determine background pixels and a best-fit plane to remove tilt present in a quantitative phase image. Quite recently, Liu et al. [59] proposed to use optimisation on Zernike polynomials in order to create a phase correction mask. Nguyen et al. [60] proposed to use Zernike polynomial fitting as well, but they trained a CNN to detect the background of a scene.

After a hologram is pre-processed and propagated, the subsequent steps usually are

1. phase unwrapping, and
2. segmentation.

The reconstructed phase is wrapped (modulo 2π) due to the arctangent function used in the phase extraction. To compensate for this, and to allow relatively thick objects to be quantitatively imaged, phase unwrapping needs to be applied. Over the years, various different methods for phase unwrapping have been proposed [61–71].

Before any analysis can be accomplished, typically one segments the object(s) from the background. Segmentation is an operation where an image can be divided into regions based on some attribute [72, 73]. In digital holography, many different methods have been introduced and applied. Active contours [74], block-wise variance [75, 76], adaptive thresholding with border following [77], adaptive thresholding with level set [52], and watershed segmentation [78] have been reported. In the field of particle holography, segmentation using thresholding [79, 80], thresholding based on combining knowledge of amplitude minimum along the longitudinal axis for each (x, y) position on the transverse plane with a particular signal to noise ratio was reported in [81], and using information of

the conjugate image by propagating between two twins was reported in [82]. Our CNN based segmentation approach is described in Chapter 7.

2.2.5 Digital holographic image analysis

Image analysis involves extracting meaningful information from an image. These tasks include extracting morphological information or other features from objects present in an image. Typically, the output from this process is a set of features, for example continuous numbers or other abstractions that are useful for later analysis. Different analysis algorithms such as

1. finding in-focus depth of objects encoded in digital holograms,
2. detection,
3. classification, and
4. tracking,

have been reported in the field of digital holography.

DHM overcomes the problem present in conventional optical microscopes of a shallow depth-of-field, permitting one to reconstruct at different in-focus planes of a volume. Despite this powerful property of digital holography, an object of interest is usually in-focus at few depths, possibly only one, as a single reconstruction layer still carries a shallow depth-of-field. Many methods have been proposed to solve auto-focusing, the problem of finding an in-focus plane. Our deep learning based solution is described in Chapter 6.

The detection and classification of objects is one of the active research areas in digital holography [83–88, 80, 89]. Correlation was used to classify protein crystals [83]. Morphological feature vectors of biological samples were extracted using Gabor wavelets and classification was realised by graph matching in [84]. This work was extended by adding standard statistical methods in [85, 86]. Principal component analysis with statistical hypothesis testing for biological microorganisms in general was reported in [87]. Malaria detection by using the cross-correlation of control and test samples was proposed in [88]. Morphological feature extraction and texture analysis for detecting and classifying para-

sites in drinking water was performed in [80]. Stem cell clustering based on the entropy estimation was reported in [89].

Another active research area that is closely related to auto-focusing is object tracking, as an object to be tracked typically needs to be in-focus. When tracking, a holographic video is recorded and tracking is applied to each frame. A comprehensive review of holographic 3D particle tracking was published by Memmolo et al. [90]. Reported holographic tracking approaches are based on a maximum a posteriori (MAP) tracker for the tracking of biological microorganisms [91], particle tracking velocimetry (PTV) algorithm for tracking red blood cells (RBC) in a microtube [92], living cell tracking (RBC and fibrosarcoma cells) by applying logarithmically weighted and bandpass-filtered power spectra [93], 3D particle-tracking algorithm that was used in the tracking of swimming micro-organisms [94]. Other similar applications using the PTV algorithm was reported in [95], and hidden Markov models in [96]. Two-dimensional cross-correlation of two successive frames was used in tracking cell mobility [97]. Summing differences of two consecutive holograms of multiple holograms (video) and thresholding was reported in [98]. Tamura coefficient and intrinsic holographic shearing between two successive frames were used in particle tracking in [99].

Other recent reported applications of DHM are particle characterisation [100], mitosis detection [101, 102], cell death detection [103] and classification [104], quantifying RBC volume [105], helical trajectories of sperm [106], cell life cycle characterisation [107], cyto-toxicity screening [108], morphometry of living cells in methanol [109], microscopic fiber characterisation [110], cell phenotype screening [111], monitoring wound healing [112], RBC morphometry [113–115], and measurement of cardiomyocytes dynamics [116].

2.3 Machine learning in the context of this thesis

2.3.1 Introduction to machine learning

Machine learning and learning algorithms can be divided into two major paradigms: supervised and unsupervised. In supervised learning, the training data is a set of input vectors along with their corresponding targets (labels), while in unsupervised learning the data does not contain corresponding targets [117]. In this thesis, multilayer perceptrons,

convolutional neural networks, random forest classifiers, and support vector machines are used for different applications and they represent supervised learning. k-means clustering and principal component analysis used in this thesis, in turn, are examples of unsupervised learning. Another paradigm, although out of the scope of this thesis, falling between these two is semi-supervised learning where typically some of the data is labelled.

2.3.2 Convolutional neural networks

Convolutional neural networks (CNNs) were pioneered by LeCun et al. [118]. A CNN is an artificial neural network where some of the layers perform convolution operations on their multi-dimensional input. The weights of convolution kernels are learned during training. Four main ideas behind convolutional neural networks are local connections, shared weights, pooling, and the use of many layers [119].

Each convolution operation produces a feature map where each feature in the map shares the same kernel, and different feature maps in a layer use different kernels. Pooling layers merge semantically similar features down to one reducing the dimensionality of feature maps. Pooling is based on small patches of a feature map for which an operation (for example maximum or averaging) is calculated and stored as a new value of a sub-sampled feature map. Pooling improves the shift invariance by increasing the size of the receptive field [119]. This should not be confused with shift invariance defined in optics textbooks as pointed out by Barbastathis et al. [22].

Currently, one of the promoting factors and a benchmark that pushes CNN progress further are different challenges and competitions such as the ImageNet Large Scale Visual Recognition Challenge [120]. Many of the modern and well-known architectures and discoveries related to CNNs have been introduced within this context [121–125]. Krizhevsky et al. [121] were the first to successfully use CNNs with large size images. They showed that using rectified linear unit (ReLU) activation decreases learning times significantly compared with other activation functions. In addition, they used dropout as a regularisation method to prevent overfitting [126]. Dropout has been successfully replaced in many applications by batch normalisation [127]. Szegedy et al. [122] introduced GoogLeNet; the most important part of this architecture was its Inception module that concatenates outputs

of multiple convolution layers. Simonyan and Zisserman [123] used a CNN architecture with a fixed 3×3 pixel kernel in its convolution layers. He et al. [124] used residual learning in which outputs of lower layers are combined (via elementwise addition operations) with outputs of higher layers. Zagoruyko and Komodakis [125] added more width to a residual network to improve network's performance and reduce the training time.

CNNs are not restricted to conventional classification and regression tasks. CNNs have been used successfully for image segmentation using image patches [128], and later using fully convolutional neural networks [129]. The encoder-decoder architectures [130, 131] allow to maintain spatial information of features. Networks using dilated convolutions [132–134], and a combination of encoder-decoder and dilated convolutions [135], allow image segmentation without pooling layers.

One important image analysis and computer vision challenge is object detection where objects belonging to different classes are spatially identified in images. CNNs have been used to tackle this problem [136–140].

An important aspect that needs to be mentioned here is overfitting. The easiest ways to avoid it is to use regularisation and have a large enough training dataset. As it is not always possible to obtain such a large dataset, the existing dataset can be extended through data augmentation [141]. Data augmentation can be based on affine transforms, such as translation, scaling, reflection, rotation, shear mapping, and elastic deformations. In addition, data can be cropped at different spatial locations. There is no single theory or rule on how much data is necessary. This should be experimentally tested on each application and signs of overfitting need to be monitored by comparing training and validation loss values.

2.3.3 Other machine learning algorithms used in the thesis

Support vector machines are supervised learning algorithms that map the input vectors into a high dimensional feature space through some non-linear mapping chosen a priori [142, 143]. A decision surface is constructed in this space. The datapoints that are located closest to the decision surface are used to measure a margin and are called support vectors. The basic idea is to find support vectors that separate two classes with an optimal margin. If the

data is linearly non-separable, it is possible to use different kernel functions (such as the radial basis function (RBF) kernel [144]) that compute a dot product in a higher dimensional space that possibly is linearly separable. As the algorithm was originally designed for a binary (two class) problem, different approaches to do multi-class classification have been proposed [145]. In a study by Pal [145] the one-vs-one approach was found to provide the highest accuracy and the lowest computational cost compared to the other five investigated approaches.

k-means clustering is an unsupervised learning clustering algorithm [146]. *k*-means clustering divides data to k clusters. This algorithm is iterative and contains two steps. After the centre of each cluster is initialised at some data point:

1. data points find the shortest Euclidean distance to a cluster centre and are assigned to that cluster, and

2. a new mean is calculated based on the formed clusters.

Steps 1 and 2 are repeated until convergence is achieved i.e. the means do not change.

Principal component analysis (PCA) [12] is a linear data dimensionality reduction technique aiming to find a set of orthogonal vectors in a dataset that can then be used to describe the data in a lower dimensional space. PCA is an important tool, especially as a preprocessing step for some machine learning tasks where the data is high dimensional.

Random forests is an example of ensemble learning using a collection of decision trees and is less prone to overfitting than an individual decision tree [147]. A decision tree contains nodes for features that are linked (rules, represented as branches) down to conclusions (e.g. classes) in the leaf nodes. The idea of combining the results of multiple classifiers or predictors that are trained with randomly picked portions (with replacement) of the trained data is called bootstrap aggregating (bagging) and was introduced by Breiman [148]. If using random forests for classification, a class is determined by a majority vote; the class that gets most votes from the trees is the classified class. In a regression task, the predicted value is the mean of all the outputs of the trees in the forest.

2.4 Microscopy and machine learning

2.4.1 Conventional microscopy and machine learning

As most of the microscopes nowadays are equipped with digital cameras, digital image processing and analysis are easily applied to the captured data. A wide variety of machine learning approaches have been applied to different fields of microscopy. A common reason to use machine learning within microscopy is to extract features of the captured image data [149], instead of using the whole image data as an input e.g. to a classifier.

A relatively new imaging modality, whole slide imaging, that is used as part of digital pathology has strongly influenced diagnostic pathology [150, 151]. Typically, in digital pathology, a pathologist prepares tissue samples that are stained and scanned by using whole slide imaging, which produces billions of pixels of data depending on the resolution of sensor used in the imaging. Different machine learning approaches within whole image sliding have been reported [152–163]. Beck at al. [152] used a logistic regression classifier with features extracted from hematoxylin and eosin (H&E) stained histological images of breast cancer tissue microarrays. Linder et al. [153] used a SVM together with extracted features to identify tumour epithelium and stroma. Cosatto et al. [154] used multiple-instance learning (MIL) to classify H&E stained gastric tissue sections. In MIL, the input is a set of multiple vectors (bag) with a single label [164]. Cireşan et al. [165] used a CNN to detect mitosis in breast cancer histology images. Cruz-Roa et al. [155] used CNNs to detect basal cell carcinoma in H&E stained histopathology images. Rexhepaj et al. [156] used a SVM to discriminate melanoma from non-melanoma cells. Wang et al. [157] combined handcrafted features with a CNN in order to detect mitosis in breast cancer pathology images. Lewis et al. [158] applied a SVM on extracted features of H&E images in order to classify carcinoma. Arevalo et al. [159] used an unsupervised learning approach to classify H&E stained basal cell carcinoma cells. Xu et al. [160] applied a CNN to segmenting and classifying epithelial and stromal regions of images of H&E stained samples. Yu et al. [161] predicted prognosis of lung cancer patients by extracting features of H&E stained histopathology images and using seven different classifiers. Yamamoto et al. [162] used a SVM classifier with morphometric features extracted from p63 and H&E

stained myoepithelial cells of breast tumours. Cruz-Roa et al. [163] used CNNs to detect invasive breast cancer regions of whole slide histopathology images.

SVMs have been used successfully in other fields of microscopy and in different microscopy applications [166–168]. Neumann et al. [166] used a SVM to classify nuclei in time-lapse RNAi screening. Osowski et al. [167] reported an application of automated detection of blood cells in bone marrow smear samples by combining a genetic algorithm and SVMs. Han et al. [168] used Haematoxylin and H&E -stained fibroblast samples and a SVM to detect cell nuclei.

Usage of conventional artificial neural networks in microscopy have been reported in [169–173]. Stotzka et al. [169] did prostate grading by using extracted features of microscope images of lesions as an input to a neural network. Binder et al. [170] used epiluminescence microscopy and reported an automated detection of pigmented skin lesions by using an artificial neural network. Boland and Murphy [171] used an artificial neural network to recognise different parts of HeLa cells imaged with a fluorescence microscope. Jurrus et al. [172] used multiple ANNs to detect and finally to segment neuron membranes by using electron microscopy images. A general approach for the classification and segmentation of medical images by using ANN was given by Sharma et al. [173].

As deep learning has become the most popular machine learning approach to image-based applications, CNNs have been used in various different applications [174–185]. Cireşan et al. [174] used a CNN to segment neuronal membranes of electron microscopy images. Prentašić et al. [175] applied CNNs to segment foveal microvasculature of optical coherence tomography (OCT) angiography images. Rezaeilouyeh et al. [176] used a shearlet transform on H&E-stained histopathology images and extracted the magnitude and phase of shearlet coefficients which were then fed to a CNN in order to classify cancerous and normal cells. Kraus et al. [177] combined CNNs and MIL to segment and classify microscopic images of fluorophore labelled cells. Eulenberg et al. [178] used imaging flow cytometry (brightfield and darkfield) and a CNN to classify (seven class classification) cell cycle phase of Jurkat cells. Dürr and Sick [179] performed cell phenotype classification by using an existing image database of fluorophore labelled cells and a CNN. Gopakumar et al. [180] applied deep belief network and CNN on microfluidics-

based imaging flow cytometry to classify leukemia cells. They used a pre-trained CNN to extract features of images and on the output of the last fully connected layer they applied PCA to reduce output vector dimensionality from 4096 to 20. These 20 element vectors were then used to train the actual classifiers (SVM, feed-forward network, naïve Bayes, and k-nearest neighbour). Kraus et al. [181] captured fluorophore labelled samples with a confocal microscope and used a CNN to classify different compartments of yeast cells. Abdolmanafi et al. [182] used OCT and the combination of a CNN as a feature extractor and three different classifiers (CNN, SVM, and Random forest) to identify Kawasaki disease from tissue samples. Karri [183] et al. used OCT to image a retina and used a CNN to identify retinal pathologies. Rivenson et al. [184] used a CNN in order to improve the resolution of an optical microscope. Later Rivenson et al. [185] used a CNN to virtually stain imaged auto-fluorescing tissue samples.

Approaches that are based on other machine learning algorithms or various different algorithms have been reported in [186–192]. Farjam et al. [186] reported automated Gleason grading (five class classification) on pathological images of the prostate gland by using tree-structured classification. Boucheron et al. [187] used multispectral imaging on H&E-stained histopathology samples and six different classifiers to classify captured images as benign or malignant. Timelapse RNAi screening was used by Jones et al. [188] who applied Gentle AdaBoost [193] on extracted features to classify different cell phenotypes. Theriault et al. [189] captured phase-contrast images of fibroblast cells and classify their morphological state by using AdaBoost. Koydemir et al. [190] used fluorescent microscopy and bagging to classify particles in a water sample in order to detect cysts of a Giardia lamblia parasite. A comprehensive review of machine learning in phenotypic profiling was written by Grys et al. [191]. Lannin et al. [192] reported an application of circulating tumour cell identification by using k-nearest neighbour, SVM, random forest, and Bayesian classifiers.

2.4.2 Quantitative phase imaging and machine learning

Despite the fact that DHM is a relatively new imaging technology, different machine learning approaches have been reported in the literature [84, 85, 87, 88, 20, 21, 194–209].

These approaches are solving traditional image interpretation problems. In addition, some of the approaches utilise machine learning to solve new quantitative phase imaging related problems such as poor resolution, shallow depth-of-field, and propagation. Javidi et al. [84, 85] used DHM and extracted features that were used with rigid graph matching to identify microorganisms. Moon and Javidi [87] showed that PCA can be used to classify digital holograms of biological microorganisms. Anand et al. [88] used DHM to capture malaria-infected and healthy red blood cells (RBCs) and correlation to detect the malaria-infected cells. Conventional artificial neural networks have been used successfully in holographic tomography [20], and particle characterisation with a simulated in-line DHM [21]. Feizi et al. [194] classified both stained and unstained cells, and quantified cell viability and concentration by using in-line digital holographic microscopy and an SVM. Park et al. [195] detected malaria by extracting features of quantitative phase images of erythrocytes together with linear discriminant, logistic regression, and k-nearest neighbour classifiers. Roitshtain et al. [196] used human cell lines, both healthy and tumourigenic, to identify cancer cells by using an interferometric phase microscopy. They extracted spatial morphological and textural features of unwrapped phase reconstructions, applied PCA to the features and used a SVM as a classifier. Jo et al. [197] captured holograms of anthrax spores and four different Bacillus species in order to identify anthrax spores by using a CNN. Rivenson et al. [198] designed a CNN that was able to learn propagation from digital holograms and recover dc and twin free amplitude and phase reconstructions. Wu et al. [199] extended the depth of field of digital hologram reconstructions with the use of a CNN. Göröcs et al. [200] used in-line holographic microscopy and deep learning to analyse water samples. Wu et al. [201] used in-line holographic microscopy and deep learning to monitor and classify bio-aerosols. Ren et al. [202] trained a CNN to detect the in-focus plane of digital holograms. Hannel et al. [203] used a CNN to detect and localise particles in a holographic video. Zhang et al. [204] applied CNNs to extract in-focus unwrapped phase of digital off-axis holograms. Another in-focus depth detection of digital holograms utilising a CNN was proposed by Lee et al. [205]. Nguyen et al. [206] proposed to use a CNN in reference beam free optical diffraction tomography where a CNN learns the inverse scattering. Luo et al. [207] were able to improve the resolution of digital

holograms by using low-resolution holograms (smaller quantity than in conventional pixel super-resolution approaches) and deep learning. The same application was reported by Byeon et al. [210]. Trujillo and Garcia-Sucerquia [208] showed that a CNN can be used to detect and count phase objects present in digital holograms. Ren et al. [209] trained a CNN to perform hologram reconstruction of amplitude and phase objects (at a macroscopic level).

Besides DHM, other quantitative phase imaging technologies have been used with machine learning algorithms. Nguyen et al. [211] reported an application of automated prostate grading by using a spatial light interference microscope (SLIM) for imaging. They combined random forests (segmentation), the bag-of-words model and a SVM to do the classification. Chen et al. [212] utilised time-stretch quantitative phase imaging to image cells, and extracted features from quantitative phase images that were then input to a multilayer artificial neural network in order to discriminate tumourigenic cells from normal, and classify algal lipid content. Guo et al. [213] used time-stretch quantitative phase imaging and a SVM in order to detect microalgal lipids. Li et al. [214] showed that a CNN can learn quantitative phase of objects imaged behind a scattering media. Sinha et al. [215] showed that a CNN can learn quantitative phase from intensity diffraction images. This work was later improved by applying spectral modulation to the training inputs [216] and extended to biological objects [217]. Zheng et al. [218] used quantitative differential interference contrast (TI-DIC) imaging and deep learning to identify cancerous lung cells.

Based on this literature review, machine learning algorithms have been used with different microscopy techniques for more than two decades. Quite recently, deep learning has started to have more and more impact in different research fields, including the field of DHM. Deep learning has been used within many interesting applications that extend use cases for machine learning.

2.4.3 Commercial software tools

Several companies manufacture digital holographic microscopes and software tools that can be used to process and reconstruct holographic data. This section contains a short description of features that different software products are advertised to have.

OsOne Image Analysis Software from Ovizio Imaging Systems NV/SA (Brussels, Belgium) has implementations of machine learning algorithms for analysing digital holograms. Machine learning is used to detect dead cells, activated cells, viral infection, and beads [219]. In addition, this software allows users to create their own machine learning models for classifying samples. In the available public material, no details of the machine learning algorithms are given.

Koala Acquisition & Analysis from Lyncée Tec SA (Lausanne, Switzerland) is capable of performing operations such as hologram stitching [220], sample roughness quantification [221], and aberration removal [56]. None of the algorithms utilize machine learning.

The HoloMonitor App Suite from Phase Holographic Imaging PHI AB (Lund, Sweden) has algorithms for time lapse kinetic cell proliferation, quality control, and kinetic cell motility [222]. None of the algorithms utilize machine learning.

4-Deep Inc. (Halifax, Canada) manufactures in-line digital holographic microscopes and software. Octopus [223] is able to perform particle counting. Swordfish [224] from the same company can be used to detect particles based on size, intensity and shape. The third software product from the same company is called Stingray [225]; it has functionality for morphology analysis and contains an unknown deep neural network classifier that users can train and use with their own data.

Based on all of the publicly available information on different manufactures' websites and advertising brochures, to the best of our knowledge, none of the commercial software can be used for performing tasks needed in this thesis. These products do not have functionality to identify the presence of objects in the field-of-view, they cannot be used for finding in-focus depth in real-time, and they cannot be used to segment and classify complex three-dimensional cysts. All the algorithms presented in this thesis are therefore considered novel.

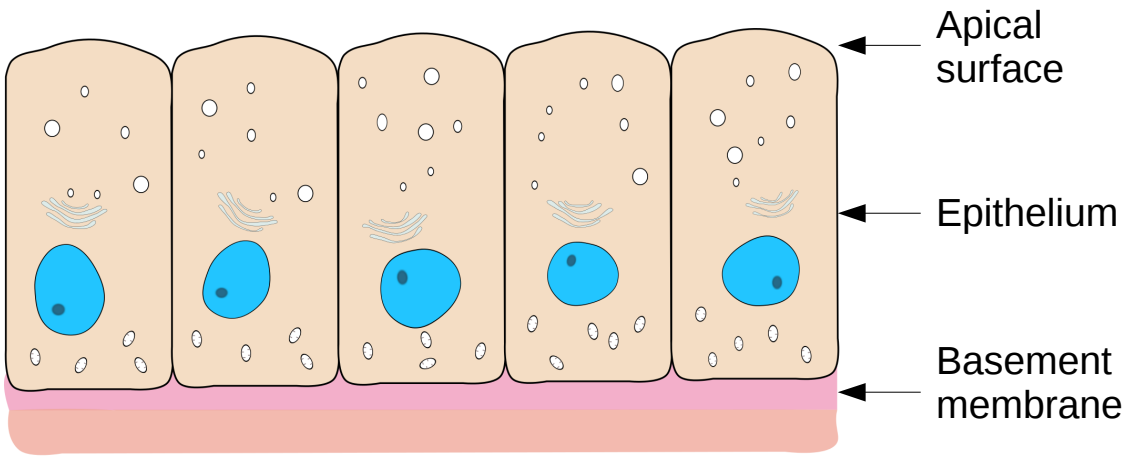

Figure 2.11 Epithelium. Epithelial cells forming tight cell-to-cell connections form the epithelium that is built on the basement membrane. The free surface facing the opposite direction of the basement membrane is an apical surface. (Author's own image).

2.5 Biological aspects of the work reported in this thesis

2.5.1 Epithelial cells

Epithelial cells form an epithelial tissue, called an epithelium, that is one of four major tissue types in animals. The epithelium coats surfaces of an organism, line organs, ducts or secretory acini and is therefore usually an interface between the organism and the environment, or between an organ and a fluid space [226]. Functions of epithelial cells are [227]:

- to protect tissues and organs,
- to maintain homoeostasis of an organism,
- the secretion of hormones and other products, and
- to register sensations.

The epithelial cells of most organs can be grown in tissue cultures. These cells and models derived from them have been used actively in studying the regulation of cell proliferation and differentiation [226]. The basic structure of epithelium is shown in Fig. 2.11.

Three-dimensional (3D) cell culture systems provide cells with *in vivo*-like microenvironmental growth conditions and have become a standard model for diverse cell biology and drug discovery applications. One of the best-characterised examples of 3D cell cultures

are Madin-Darby canine kidney (MDCK) epithelial cells embedded in 3D collagen or Matrigel™ substrates, where the normal cells proliferate to form spherical multicellular cysts with a hollow interior, called a lumen [228]. The main components of a cyst are epithelial cells themselves, an apical surface (walls of a lumen when grown in 3D medium), and a basement membrane (Fig. 2.12). These 3D cultures are particularly useful when detecting growth properties of abnormal cells, such as oncogenic cancer cells that are poorly differentiated and so fail to form structurally organised polarised cysts [229, 230].

As a model for abnormal cells, we used MDCK cells that express an oncogenic G12V-mutant of K-Ras proto-oncogene ($KRas^{V12}$) [230]. $KRas^{V12}$ is a strong oncogene that is frequently found in several different types of solid tumours including lung, pancreatic, and colorectal cancers with poor prognosis [231]. A human prostate epithelial cell line, RWPE-1, was the other cell line used in the experiments. This cell line underwent the same processing and culturing methods as the MDCK line. Lentiviral transgenesis was used to overexpress $KRas^{V12}$ in otherwise healthy cells. In lentiviral transgenesis, lentiviruses replace the chromosome of target cells with their genome [232].

DHM measures a projection along the optical axis (the direction that light propagates from the sample to the camera). Therefore, if a hollow lumen is present in a spherical cyst, the quantitative phase values should be lower at the centre of the projection and higher between the basolateral membrane and apical surface (the edge region consisting of a monolayer of polarized cells). Fig. 2.13 shows same cyst captured with a confocal microscope (63x oil immersion microscope objective) and DHM (40x microscope objective). It can be observed that the hollow lumen (confirmed with Fig. 2.13b) is displayed as a "crater" -shape in the hologram reconstruction.

2.5.2 Prostate

The prostate is located between the bladder and the penis in front of the rectum (Fig. 2.14). The urethra goes from the bladder through the centre of the prostate to the penis allowing urine to exit the body. The function of the prostate is to secrete a fluid during an ejaculation. The prostate is composed of epithelial and stromal cells [233]. The epithelial cells form glands with hollow lumens, and the stroma surrounds these glands [233] (Fig. 2.15).

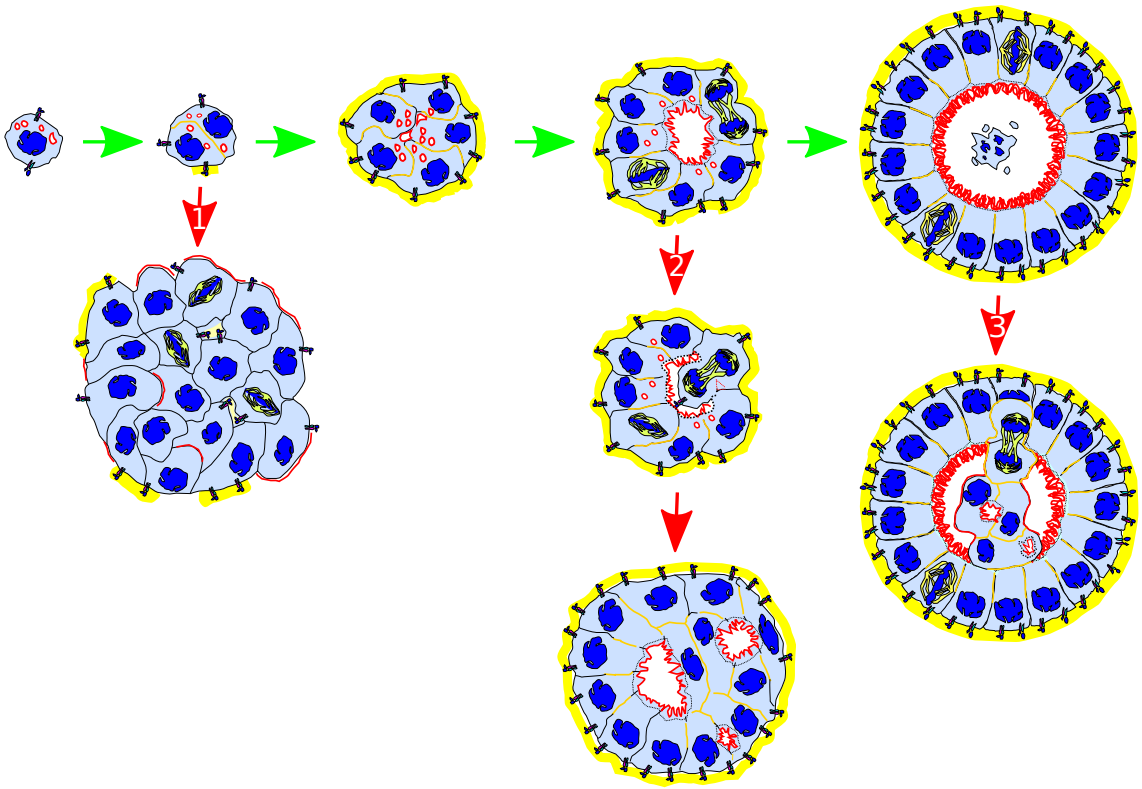

Figure 2.12 Cyst formation. When epithelial cells are embedded in a 3D medium, cells form cell-cell contacts and form their basolateral membrane (yellow). With a healthy cyst (green arrows) the apical lumen (red) is formed when cells are correctly polarised. In the case of failure, the formation of a lumen can be disturbed, and cysts can grow as an unpolarised cell mass (red arrow 1) or partially polarised structures (multiple lumens) (red arrow 2). In case of resistance to anoikis (programmed cell death), that should take place when a cell is trapped inside a lumen, cells can continue to proliferate inside the lumen, blocking the lumen (red arrow 3). Figure inspired by [228].

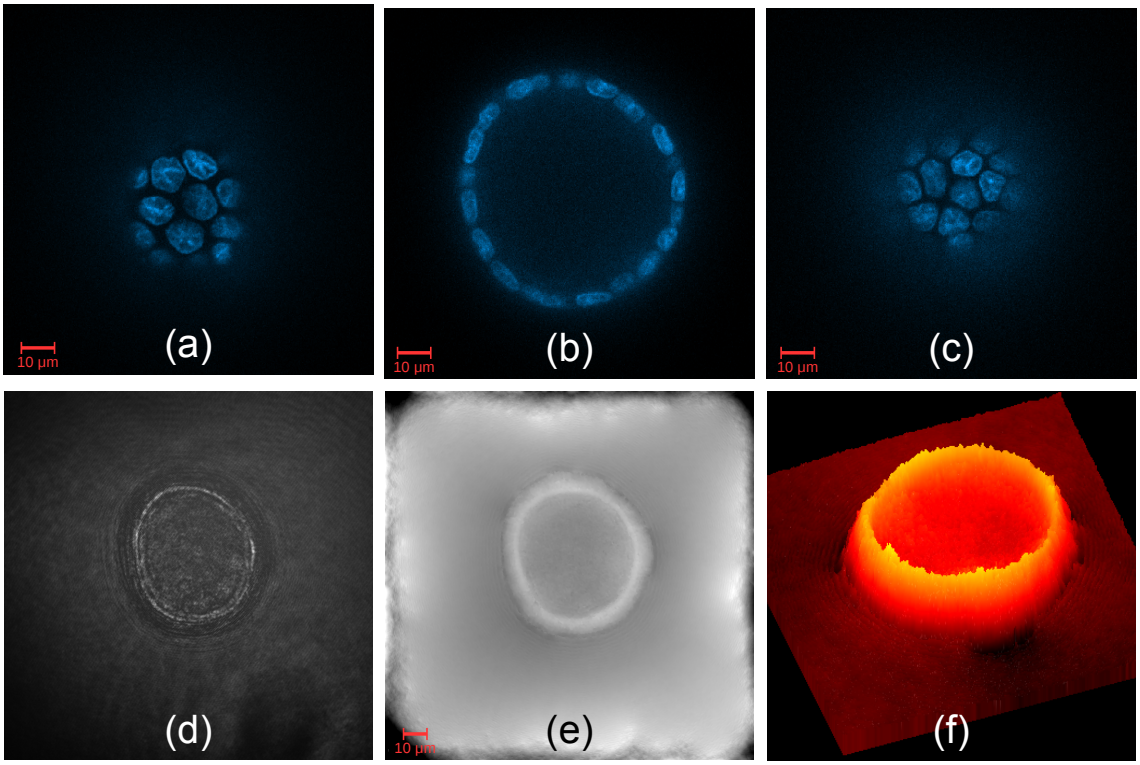

Figure 2.13 Confocal and DHM of MDCK cyst. The same MDCK cyst imaged with a spinning disc confocal microscope and DHM. Confocal images at different sections of a cyst (a)-(c), at (a) 6μm, (b) 34.8μm, and (c) 63.6μm, (d) digital hologram, (e) quantitative phase reconstruction at -4.8 cm from the hologram plane, (f) 3D mesh representation of (e).

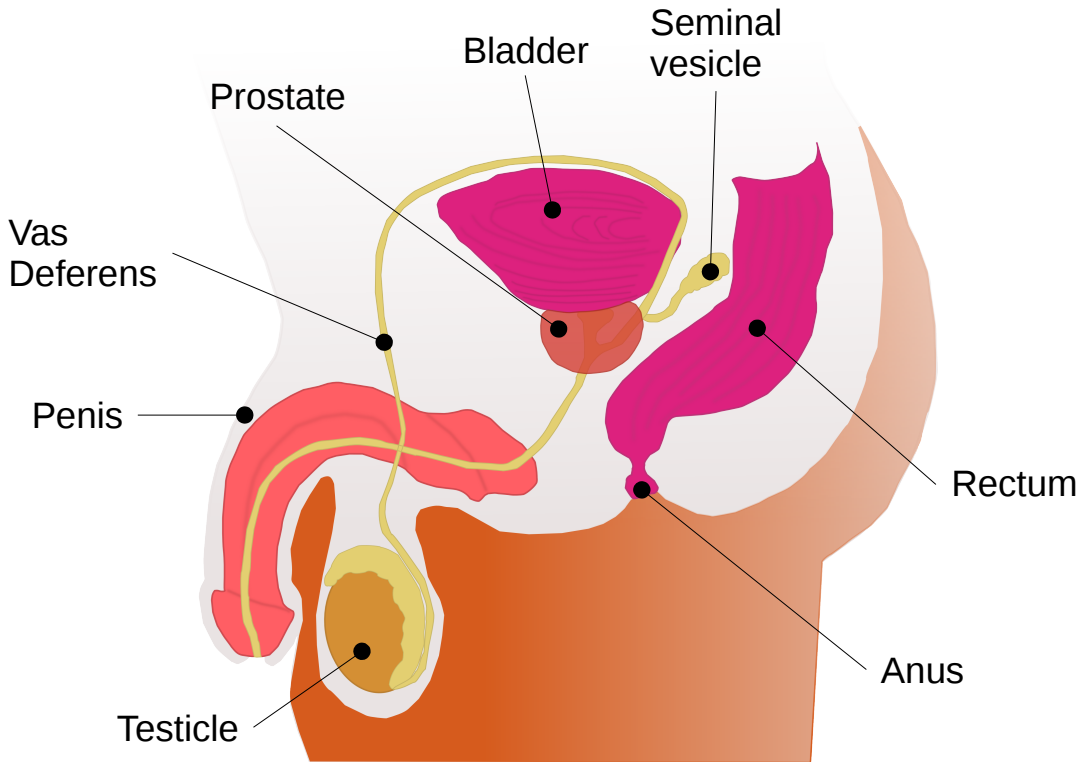

Figure 2.14 The male reproductive system showing the location of the prostate. (Author's own image.)

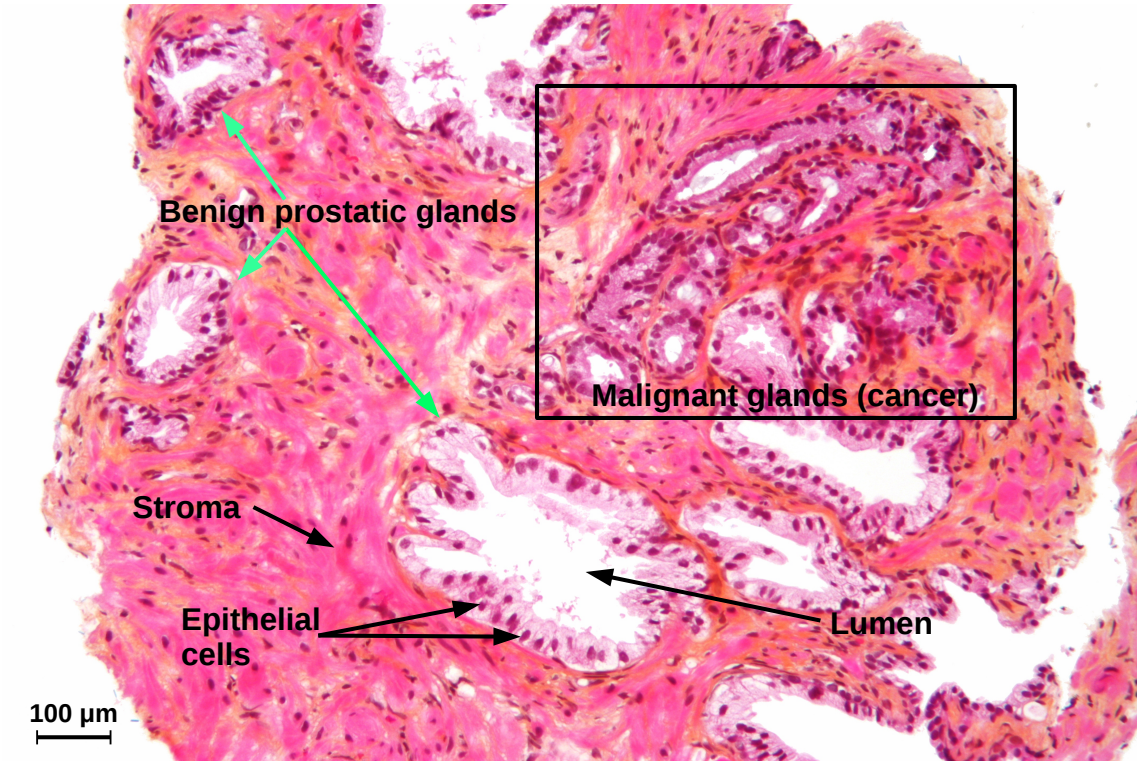

Figure 2.15 Hematoxylin phloxine saffron (HPS) stained prostate tissue sample obtained through biopsy, showing a thin cross section of a prostate. Green arrows point out benign prostatic glands that are formed of hollow lumens (hollow tubular structures inside prostatic glands) surrounded by polarised epithelial cells and stromal cells. The region inside the rectangle contains cancerous tissue with malignant prostatic glands. The scale bar on the bottom left is an estimated scale. [The original figure: Copyright ©2011 Michael Bonert (https://commons.wikimedia.org/wiki/User:Nephron).] License: CC BY-SA 3.0.

Prostate cancer is the second most common cancer in men after lung cancer [234]. In 2018, globally 1.2 million new cases were diagnosed [235]. There is no standard test for prostate cancer screening. A prostate specific antigen (PSA) is generated by the prostate, and the quantity of this can be measured from blood of a patient. High levels of PSA may indicate prostate cancer, however, only 25 % of men with elevated PSA levels have prostate cancer [236]. A digital rectal examination conducted by a medical expert can be performed together with the PSA test, and is routinely performed if PSA levels are high or if a patient has prostate related symptoms. The final diagnoses and evaluation (Gleason score) of prostate cancer is performed by a pathologist from tissue samples obtained through biopsy. Radical treatments (prostatectomy or radiotherapy) do not necessarily improve life expectancy when compared to active-monitoring [237]. Therefore, it would be important to be able to estimate cancer aggressiveness accurately and treat actively only men with a lowered life-expectancy or lowered life quality. If this was possible, not only financial benefits could be achieved, but also patients suitable for active-monitoring would not need to suffer from unnecessary treatment side-effects such as erectile dysfunction [238], urinary problems [239], and infections.

3 Equipment and method

This chapter discusses challenges present in current imaging systems and proposes a system based on DHM that can be used in high-throughput analysis of cysts. Imaging results obtained by DHM are shown together with data captured with a confocal microscope.

3.1 Introduction

Multicellular 3D structures are often relatively large and thus pose immense challenges for imaging and image-based analyses [240, 241]. Typically, confocal microscopy or selective plane illumination microscopy (SPIM) [242] is used to analyse these structures, each of which relies on fluorescent marker molecules that must be introduced into the cells (Fig. 3.1). Moreover, and particularly in the case of confocal microscopy, high-intensity laser excitation of the samples is required to capture high-quality data, potentially resulting in photobleaching of the fluorophores, phototoxicity, and production of oxygen radicals [7, 8]. In addition, these techniques capture multiple optical slices that are aggregated to generate a 3D dataset that requires significant storage capacity, in particular for time-lapse imaging applications. Further, imaging live and fast-moving specimens can be challenging for these scanning-based techniques, and selecting and analysing features in the scanned 3D data is largely a laborious manual process. Despite these challenges, the field of imaging and image-based analysis of 3D cysts is active, and in particular high-throughput analysis is a new topic that is seeing growing interest. State-of-the-art methods [243, 244] however, by using only 2D imaging techniques, are not exploiting the full 3D information available for analysis in 3D cysts. As explained, existing 3D

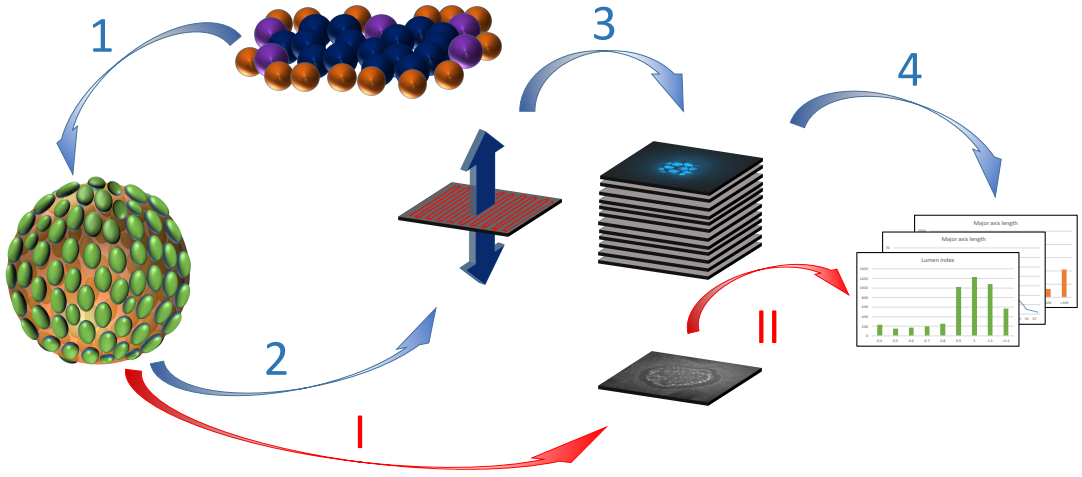

Figure 3.1 Conventional cyst imaging and analysis compared to DHM. Conventional approach (blue arrows) starts with introducing fluorescence molecules into a sample (1), the sample is imaged using raster scanning (2) producing confocal stacks (3), that are reconstructed and analysed (4); DHM approach (red arrows) is a fluorescence-free and completely non-scanning imaging modality producing a small amount of data compared to many other techniques (I) that can be efficiently processed and analysed (II).

imaging techniques have various disadvantages that mean they are unavailable as options for the specific needs of the high-throughput 3D cyst analysis community. This motivates the pressing need for an automatic system that can image and analyse tens or hundreds of thousands of cysts, removing bias in the selection of the cysts for measurement, thus increasing the reliability and tenability of any resulting scientific findings.

Quantitative phase imaging, such as DHM allows one to quantify the optical density of a living specimen without fluorophores (Fig. 3.1) [37, 39], for example to obtain noninvasive dry mass [245]. DHM can be characterized as a non-scanning imaging technique that uses low-intensity light and produces comparatively small amounts of data (a modest data storage requirement) compared to other 3D microscopy techniques. Without data loss, DHM captures the full wavefront of laser light passing through a specimen, allowing a single camera frame to be used to reproduce brightness and phase measurements at multiple focus depths through a 3D sample. Although the technology behind DHM is known and understood, and rudimentary hologram processing algorithms such as segmentation have been applied to digital holograms, robust and reliable automated hologram analysis algorithms for many fields, including organoid research, are not found in the literature.

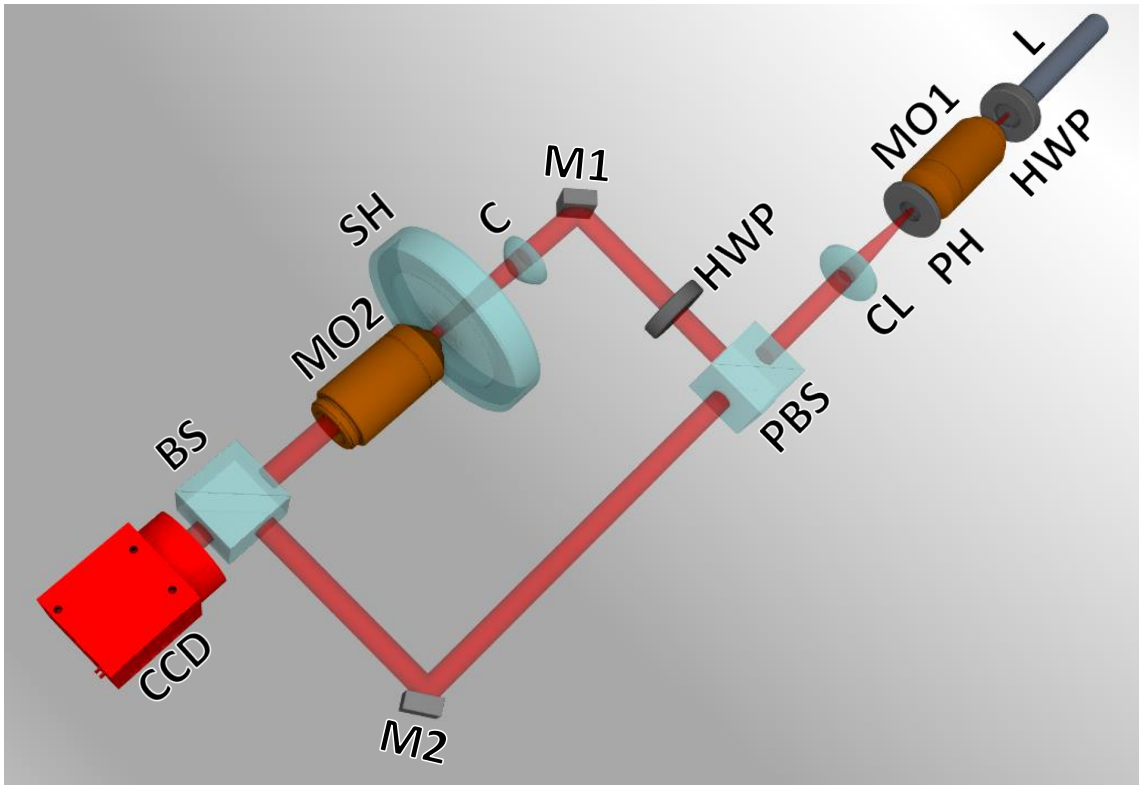

Figure 3.2 DHM setup. The optical architecture based on an off-axis Mach-Zehnder interferometer: the light from the laser (L) goes through a half-wave plate (HWP) and is focused by the microscope objective (MO1) onto a pinhole (PH). The resulting spherical wave is collimated by the collimating lens (CL) and split into two mutually-coherent beams, called the reference and object beams, by the polarizing beam splitter (PBS). The object beam goes through another half-wave plate (HWP) matching the polarization of the reference light and illuminates the sample (SH) through the condensing lens (C). The second microscope objective (MO2) magnifies the object beam and is combined by the beam splitter (BS). The digital camera (CCD) samples the hologram formed by the interference of the two beams. M1 and M2 are mirrors used for beam steering. A small angle is introduced between the directions of the reference and object beams enabling more convenient decoding of the digital holograms.

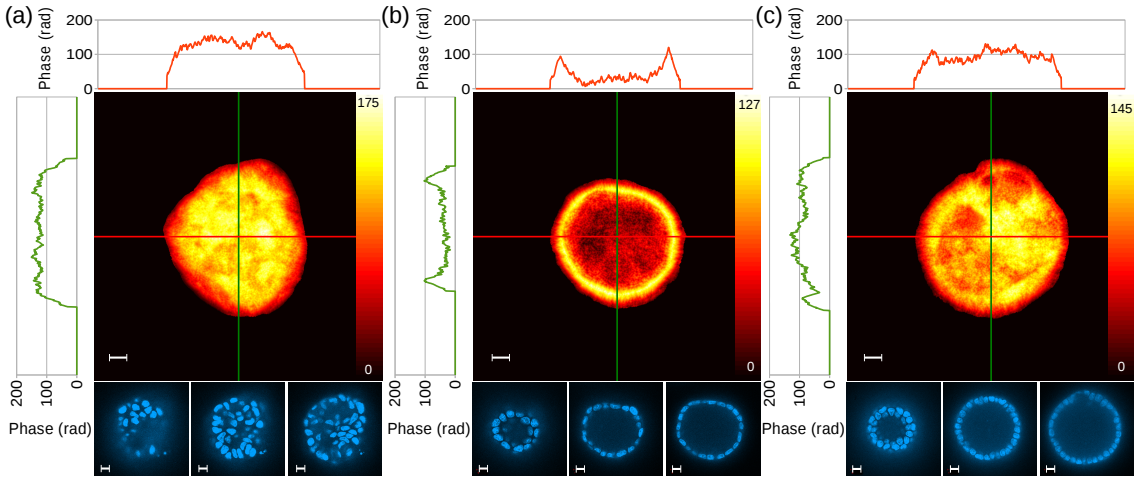

Figure 3.3 DHM analysis and confocal data. Reconstructed phase images with vertical (green) and horizontal (red) lines show positions of the extracted phase profiles plotted above and to the left of the images. Below each phase image, three confocal sections of a z-stack of the same cyst are shown. Confocal data was captured with a Zeiss spinning disc confocal microscope, and DHM data with a commercial digital holographic microscope (Lyncée Tec SA, Lausanne, Switzerland). Scale bars are 10 µm. (a) segmented quantitative phase reconstruction of a KRas^{V12} transformed MDCK cyst reconstructed at 5.3 cm from the hologram plane, lumen index 0.01; (b) segmented quantitative phase reconstruction of a wild-type MDCK cyst reconstructed at 3.6 cm from the hologram plane, lumen index 0.83; (c) segmented quantitative phase reconstruction of a KRas^{V12} transformed MDCK cyst reconstructed at 5.8 cm from the hologram plane, lumen index 0.19.

We couple DHM with a novel image analysis algorithm for living 3D cell cultures. To validate the novel imaging and analysis modality, wild-type MDCK cells were embedded into a Matrigel^{TM} substrate where they formed hollow polarized cysts within 4-to-5 days in culture. The total number of wild-type and KRas^{V12} transformed cysts analysed was 281 and 217, respectively. A subset of the cysts (20 and 24, wild-type and KRas^{V12}-transformed cysts, respectively) were imaged with both a digital holographic and a spinning disc confocal microscope. After being imaged with the DHM, the cell nuclei were stained with DAPI to facilitate confocal imaging. The majority of the KRas^{V12}-transformed MDCK cells grew as solid clusters without lumen as evidenced by multiple DAPI-stained nuclei inside the cell clusters (Fig. 3.3a). In these cases, the DHM quantitative phase images also unambiguously showed the absence of lumen in KRas^{V12}-transformed cysts (the edge region could not be distinguished from the central region). For most wild-type MDCK cysts, DHM clearly reveals higher quantitative phase values between the basolateral and luminal membranes (Fig. 3.3).

While both confocal and DHM techniques, in general, were able to resolve whether the cysts contained lumen or not, there is evidence that DHM can be more discriminating. This claim is justified by the observation that in some cases, K-Ras^{V12}-MDCK cysts reconstructed from confocal slices appeared to have central lumens that were devoid of DAPI-stained nuclei whereas the DHM quantitative phase image of the same cyst did not detect any lumen. It was found that in these cases DNA had been digested resulting in a loss of DAPI-staining although cell debris was still present (Fig. 3.3). The ability of DHM to detect any biological material, including that which would be unlabelled in a confocal fluorescence setup, could therefore be beneficial for more reliable imaging and analysis in some cell clusters. In terms of storage requirements, the DHM data consists of a single image (digital hologram) whereas the confocal data consists of multiple individual optical slice images. Importantly, storage requirements also have an implication for processing speeds: whereas complex reconstruction and volumetric analysis of the confocal 3D stack is required for the quantitative measurement of confocal images, the single DHM phase image can be easily converted to a non-biased quantitative measure that determines a "lumen index".

To verify the feasibility of applying the proposed approach to other cell types, human prostate (RWPE) samples were prepared (the number of wild-type and KRas^{V12} cysts in the analysis was 85 and 114, respectively), imaged with DHM, and analysed. The most distinguishing feature for RWPE samples was related to lumen. This is surprising as this cell type does not form clear hollow lumens like MDCK cysts do, it was expected that other features would have a greater significance. No individual extracted feature alone can be used to differentiate wild-type and transformed samples with a high degree of accuracy. Instead, multiple extracted features should be used with machine learning algorithms to classify different cyst phenotypes. This is a feasible solution for applications where one needs to compare metamorphosis to an expected phenotype, for example, after some sample treatment.

3.2 Digital holographic microscope

In the imaging experiments a commercial T1000 series digital holographic microscope (Lyncé Tec SA, Lausanne, Switzerland) was used. The microscope comprised a 660 nm laser source and a 1024×1024 pixel CCD camera with 6.45 µm square pixels. All the imaging experiments were performed by using a 40x microscope objective with 0.7 numerical aperture (Leica HCX PL Fluotar). The optical setup was based on an off-axis Mach-Zehnder interferometer (Fig.3.2). Manual three-axes translation stages were used to manually find cysts in a petri dish and holograms were captured close to an in-focus plane.

The microscope had a built-in automated calibration procedure that allowed it to compensate for an optical path length difference caused by the sample. This calibration was performed at the beginning of each imaging experiment to compensate for differences in medium thickness between different samples.

After verification of the suitability of the proposed system, the system was automated to enable the system to be used for high-throughput imaging/analysis. Motorized actuators (SGDC10-25, OptoSigma Corp., Santa Ana, California, USA) were installed on two orthogonal axes translation stages besides a top-stage environmental chamber (model Uno-T-H-CO2, OKOLAB S.R.l, Ottaviano, Italy). Although holograms shown and analysed here are captured with a 40x objective, a 10x microscope objective (Leica HCX PL Fluotar 10x/ NA 0.3) is sufficient for imaging and extracting quantitative features related to morphology and phase, enabling capture with a larger field of view (Fig. 3.4). Table 3.1 shows a comparison of extracted features with these two microscope objectives. It can be observed that 10x is indeed suitable for high-resolution imaging of cysts.

3.3 Confocal microscope

The same cysts imaged by DHM were imaged by using a Zeiss Cell Observer microscope equipped with a Yokogawa CSU-X1 spinning disc confocal unit. Imaging was performed by using a 63x alpha Plan-Apochromat oil immersion objective with a numerical aperture of 1.46. The cysts were searched for manually by using a motorized three-axis translation stage and confocal stacks were captured from the top to the bottom part of a cyst. Start and

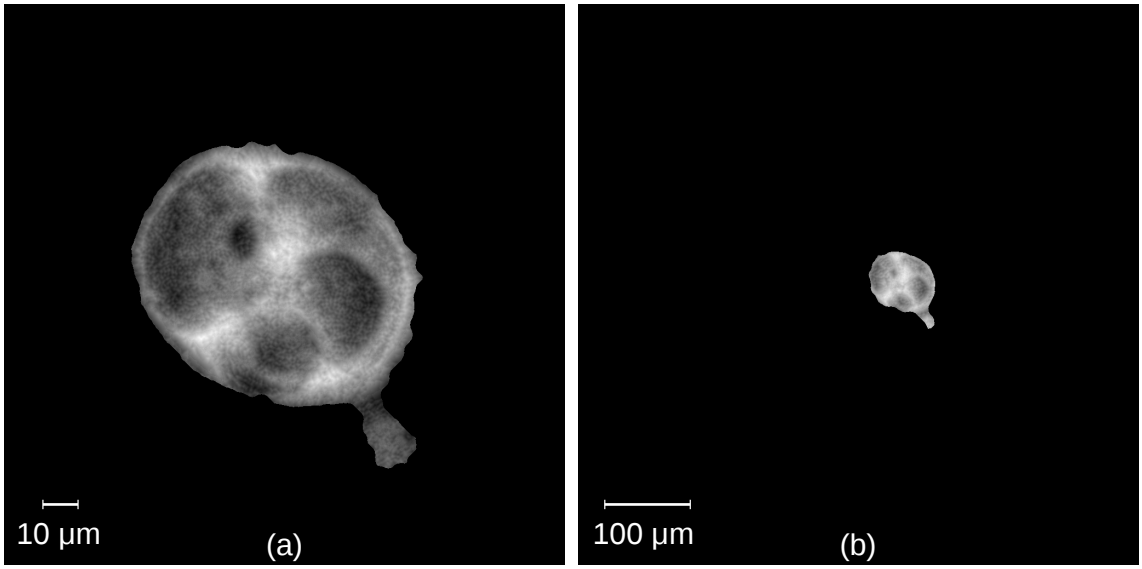

Figure 3.4 40x vs 10x objective. A cyst captured with two different microscope objectives. (a) 40x microscope objective, (b) 10x microscope objective.

Table 3.1 40x and 10x comparison. Comparison of extracted feature values of holograms captured with two different microscope objectives. Features were extracted using methods described in Chapters 4 and 7. Units are in micrometres where applicable.

Feature	40x	10x	Absolute difference
Major axis length	90.31	90.47	0.16 (0.18 %)
Minor axis length	67.96	65.96	2.00 (2.95 %)
Eccentricity	0.66	0.68	0.02 (3.94 %)
Area	4605.44	4450.43	155.01 (3.37 %)
Roundness	0.72	0.69	0.03 (3.72 %)
Perimeter	306.60	291.17	15.43 (5.03 %)
Convex area	5118.01	4927.94	190.07 (3.71 %)
Equivalent diameter	76.58	75.28	1.30 (1.70 %)
Extent	0.63	0.65	0.02 (2.39 %)
Minimum phase	7.37	39.24	31.87 (432.46 %)
Maximum phase	171.71	117.57	54.14 (31.53 %)
Orientation	-0.75	-0.76	0.01 (1.30 %)
Solidity	0.90	0.90	0.00 (0.00 %)
Skewness	0.42	-0.08	0.52 (119.92 %)
Kurtosis	-0.36	-0.60	0.24 (67.38 %)
Roughness	1.08	1.04	0.04 (3.43 %)
Standard deviation	24.86	11.24	13.62 (54.78 %)
Density	0.003	0.01	0.008 (269.35 %)
Average	82.94	31.75	51.19 (61.73 %)
Lumen index	0.72	0.76	0.04 (5.56 %)

end points were selected manually, and optical sections were gathered in 600 nm steps to fulfil the sampling theorem.

The captured data was processed with ZEISS ZEN microscope software that allows controlling a microscope, processing, and analysing images.

3.4 Sample preparation

150 µl of Matrigel^{TM} was pipetted onto a 35 mm high glass bottom u-dish (Ibidi 81158) and allowed to solidify at 37 °C for 20 minutes. Subconfluent MDCK cells were treated with trypsin, resuspended in complete growth medium (MEM-Glutamax; 41090-093, Life Technologies) supplied with 5 % fetal bovine serum (FBS, Life Technologies) and 1 % penicillin/streptomycin (15140122, Gibco) and cells were counted. Ten thousand cells were resuspended into 1.5 ml of complete medium containing 2 % Matrigel^{TM} and the mixture was gently pipetted onto thin layer Matrigel^{TM}-coated U-dish. Culture medium was exchanged every two days with fresh culture medium containing 2 % Matrigel^{TM}.

3.5 Conclusions

The full system is idealised in Fig. 3.5. The microscope operates online and can be accessed with client software through the internet. This enables a user to monitor the progress of imaging and analysis giving access to data on demand. An initialization step requiring user interaction with the system at the beginning of the imaging contains a sample preparation and placing the sample on the sample holder. After this initialization, the whole process of imaging and analysis is fully automated. With a state-of-the-art translation stage and 10x microscope objective a full 36 mm dish (with a 21 mm observation area, the largest square fitting a 21 mm circle diameter is 14.85 mm, resulting in 527 holograms) can be imaged in 271 seconds (with 500 ms settling time, 1 ms shutter time, and 13 ms translation time).

This chapter contained a description of hardware that is needed for an automated high-throughput imaging. In addition, the DHM hardware used for experiments in this thesis was defined. A lack of processing and analysis algorithms was identified. In the following chapters, we give novel tools that provide this missing piece of knowledge. An

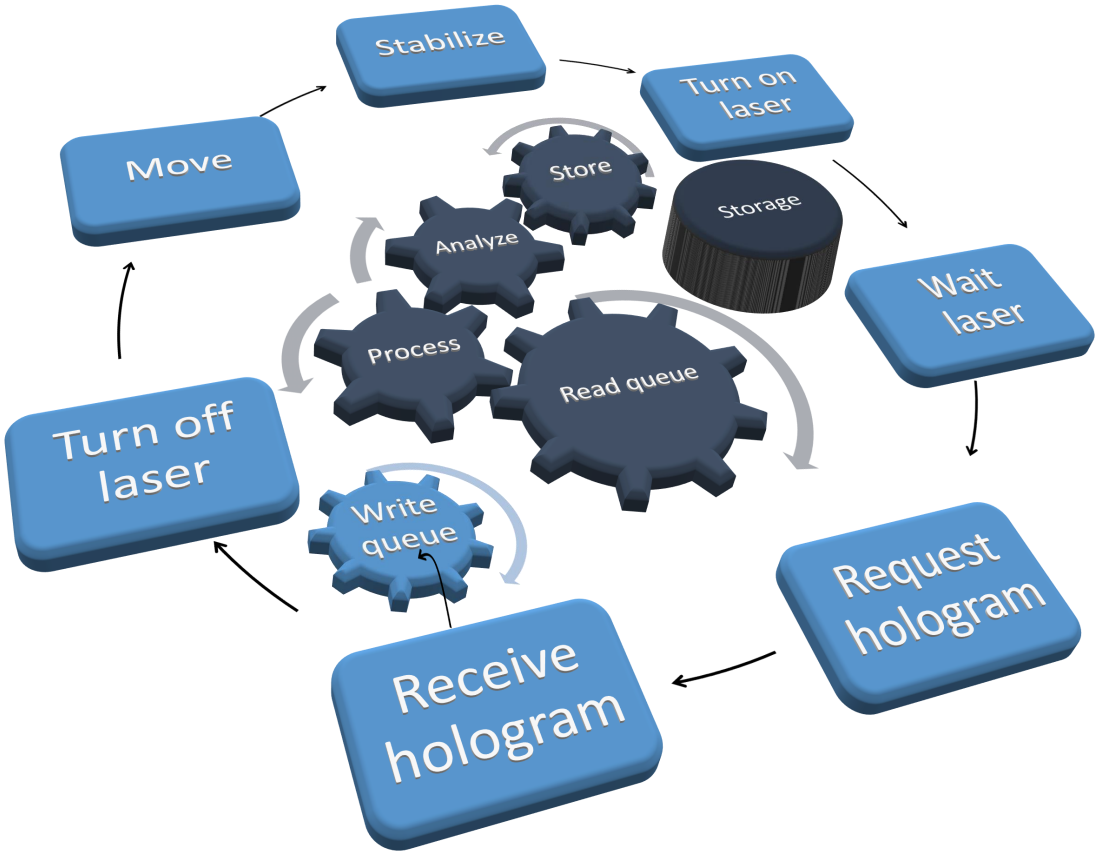

Figure 3.5 Asynchronous capturing, processing and analysing. Two different processes are run parallel. One of the processes is in charge of capturing data and translating the sample while the other process reads the captured data, processes and analyses the data, and stores values in a database.

automated solution to determine if a field of view contains an object is described in detail in Chapter 5. Finding in-focus depth of objects encoded in digital holograms is given in Chapter 6. A novel segmentation approach of holograms is introduced in Chapter 7. Further analysis of holograms is given in Chapters 4 and 8. Suitability of the system for time-lapse imaging and real patient samples is proven in Chapter 9.

4 Classification with extracted features

In this chapter, extracted quantitative features are used with three different classifiers to classify cysts as normal or oncogenic. In addition, a CNN is trained to perform the same classification task. An approach to use multiple different classifiers was chosen in order to investigate if some of the classifiers perform better than others. The object was to show that at least one configuration of deep learning network outperforms an exhaustive battery of less sophisticated techniques. It is not the intention in the thesis to explain for all deep learning design decisions and all hyperparameters, which are the most crucial. While transfer learning has been proved to be a viable technique, and there are concepts such as “reasonable default values" for certain hyperparameters, in general, if one has a novel dataset, one cannot predict the optimal properties and parameter values for a given deep learning network in advance. For the results in this thesis, we cannot claim that they will hold with different cell lines. All solutions presented in this chapter were implemented using the general-purpose Python programming language, and the scikit-image [246] and scikit-learn [247] libraries. The experiments related to CNNs were implemented using the PyTorch [248] deep learning framework.

4.1 Introduction

An automatic process that can correctly classify cysts as normal or oncogenic is a useful tool when there is a need to distinguish between the two. One potential application is in drug research where different drugs are applied to tumourigenic samples, and the drug

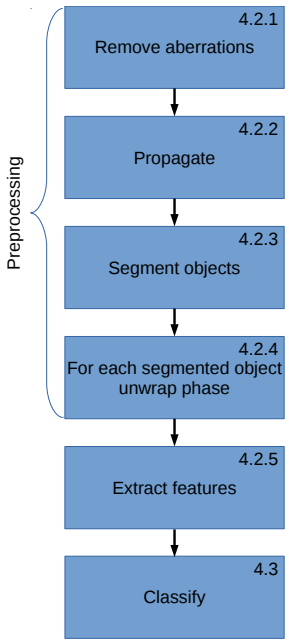

Figure 4.1 Processing and analysing steps. Numbers inside each box refer to a section in this chapter.

responses need to be monitored. This monitoring can be performed by classifying each cyst and tracking the number of each class over time.

Before any classification can be realised, the data for a classifier needs to be pre-processed into a suitable format. In this chapter, data for a classifier is extracted from a digital hologram phase reconstruction. For the corresponding classification using CNNs, segmented phase reconstructions are used as input. For data extraction, the phase reconstruction needs to be segmented from the background. In the literature, most of the segmentation methods for digital holographic microscopy exploit intensity reconstructions [74, 81, 92, 100, 94, 80]. As the cysts have a relatively large diameter, the in-focus plane is disturbed by the diffraction of the out-of-focus planes. This fact complicates the use of the reconstructed intensity.

4.2 Preprocessing and feature extraction

The overall process flow is shown in Fig. 4.1. Briefly, phase aberrations caused by the optical set-up and unwanted hologram terms are removed. A hologram is propagated to the in-focus plane and a phase reconstruction is extracted, objects are segmented, each object is unwrapped and quantitative features are extracted for further analysis.

4.2.1 Aberration removal

Holograms were preprocessed by removing the zero order, a twin term, and phase aberrations using a method introduced by Zuo et al. [57] (Fig. 4.2). This state-of-the-art technique for phase aberration correction, at the time, was chosen as it does not require prior knowledge of the imaged sample. A comparison of phase extraction methods was beyond the scope of the thesis. In addition, in justification of the use of this method, the problems of how to choose optimal phase extraction and phase aberration methods were not problems tackled in this thesis. Unfortunately, we did not have access to any other competing phase metrology technology in our laboratory (although we did compare it with confocal microscopy in Chapter 3). This meant that we had no ground truth phase data for our samples. If we had, we could have performed very interesting analyses such as how differences, if any, between reconstructions and ground truth affect the performance of the classification algorithms, including experimental upper bounds on the accuracy of our algorithms with perfect data. Not having ground truth phase data can be problematic if one has a self-built phase capture system and one wishes to argue for the general applicability of one's image processing algorithms. However, in our case, we use an off-the-shelf commercial DHM and well-known cell lines, so capturing identical phase data to ours can be achieved in principal in any laboratory. It therefore becomes less important whether the perfectly accurate phase is measured or not, as long as the classification accuracy is high (i.e. the transfer function of the repeatable capturing system could be enhancing classification accuracy) and the capturing system and capturing conditions can be easily replicated.

4.2.2 Propagation

The holograms were reconstructed along the optical axis at the centre region of the cyst by using the Fresnel approximation (Eq. 2.2) where the reconstruction depth was chosen automatically (our CNN based automated method for finding an in-focus plane is described in detail in Chapter 6)

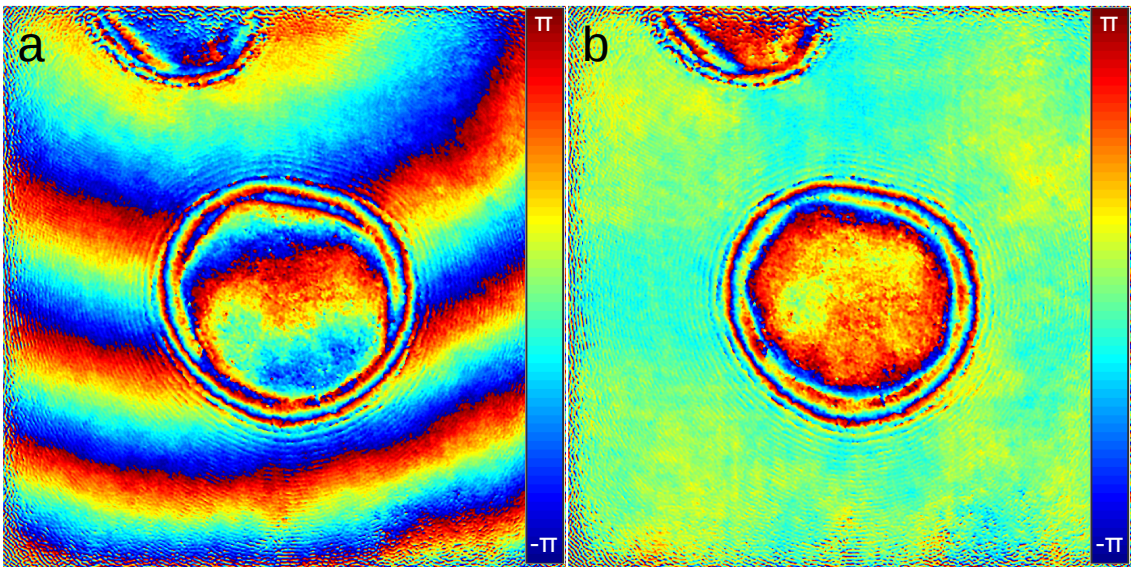

Figure 4.2 Hologram processing. (a) wrapped phase reconstruction at 3.5 cm from the hologram plane containing aberrations caused by the microscope objective, (b) aberration free wrapped phase.

4.2.3 Segmentation

Segmentation evolved during process of this thesis. Our convolutional neural network approach is described in detail in Chapter 7.

4.2.4 Phase unwrapping

Phase unwrapping was realised by using the method introduced by Ghiglia and Romero [62].

4.2.5 Feature extraction

The lumen condition can be analysed accurately by using the segmentation masks. For this we have derived a novel metric, r, that defines relation between a theoretically empty lumen and a measured lumen. Each segmentation mask $S(x,y)$ originally contains three distinct values (0: background, 1: cells between apical and basolateral membranes, 2: and hollow lumen). By binarising $S(x,y)$ (combining regions 1 and 2, and leaving value 0 for the background), and applying an image erosion operation to the resulting mask with a 50 pixel diameter isotropic structuring element, we obtain a binary mask $S'(x,y)$ that

segments the theoretical lumen only. By calculating total number of pixels of $S'(x,y)$ as

$$A' = \sum_{x=1}^{M} \sum_{y=1}^{N} S'(x,y) \tag{4.1}$$

we are able calculate lumen index, r, the ratio between the measured lumen, L, and theoretical area

$$r = \frac{L}{A'} \tag{4.2}$$

where L is

$$L = \sum_{x=1}^{M} \sum_{y=1}^{N} S''(x,y) \tag{4.3}$$

where

$$S''(x,y) = \begin{cases} 1, & \text{if } S(x,y) = 2. \\ 0, & \text{otherwise} \end{cases} \tag{4.4}$$

The r value defines how much of a lumen is hollow. In addition, as the lumen is segmented with such a high accuracy a multi-lumen phenotype can be detected; if the region inside basement membrane contains more than one connected region, the cyst is multi-lumen.

In addition, other features were extracted. Some of the features are the same as those used in a paper by Härmä et al. [249]. This shows that DHM can be used to extract the same quantitative features that are available in confocal images. These features can be divided into two groups: morphology and phase related as shown in Tables 4.1 and 4.2.

4.3 Classifiers

The lumen index can be used to quantify the lumen region; however, it is not sufficient alone on its own to discriminate different phenotypes within different cell types. The maximum phase can be regarded as the most descriptive single feature with MDCK samples, but still, it cannot be used alone for a high accuracy classification as will be shown later in this chapter. Indeed, the classification accuracy can be improved by additional features. For this, the extracted features (n=19) of MDCK data were used with a multilayer perceptron (MLP), a random forest classifier (RF), a support vector machine (SVM), and a CNN.

Table 4.1 Extracted morphological features and their derivations.

Feature	Equation
Area (A)	$\sum_{x=1}^{W}\sum_{y=1}^{H} S(x,y)$
Convex area (C)	number of pixels of convex hull image
Eccentricity (ε)	$\sqrt{1-\frac{\gamma^2}{\beta^2}}$
Equivalent diameter	The diameter of a circle with the same area as the region
Extent (E)	A/lc, where l and c are the height and width of a region, respectively
Major axis length (β)	The length of the major axis of the ellipse that has the same normalised second central moments as S(x,y)
Minor axis length (γ)	The length of the minor axis of the ellipse that has the same normalised second central moments as S(x,y)
Perimeter (P)	length of a line through the centres of border pixels of $S(x,y)$ using a 4-connectivity
Roughness (R)	$\frac{P}{P'}$
Convex image perimeter (P')	length of a line through the centres of border of a binary convex hull image which has the same size as bounding box of S(x,y) using a 4-connectivity
Roundness (O)	$\frac{4\pi A}{(\pi\beta)^2}$
Solidity (T)	$\frac{A}{C}$

Table 4.2 Extracted phase features and their derivations.

Feature	Equation
Average (α)	$\frac{1}{WH}\sum_{x=1}^{W}\sum_{y=1}^{H}\phi(x,y)S(x,y)$
Density (D)	$\frac{s^2}{\alpha}$
Kurtosis (K)	$WH\frac{\sum_{x=1}^{W}\sum_{y=1}^{H}[\phi(x,y)S(x,y)-\alpha]^4}{\{\sum_{x=1}^{W}\sum_{y=1}^{H}[\phi(x,y)S(x,y)-\alpha]^2\}^2}$
Lumen index (r)	see text for explanation
Maximum phase	$\max[\phi(x,y)S(x,y)]$
Minimum phase	$\min[\phi(x,y)S(x,y)]$
Skewness (B)	$\frac{\sum_{x=1}^{W}\sum_{y=1}^{H}[\phi(x,y)S(x,y)-\alpha]^3}{s^3}$
Standard deviation (s)	$\sqrt{\frac{1}{WH}\sum_{x=1}^{W}\sum_{y=1}^{H}[\phi(x,y)S(x,y)-\alpha]^2}$

For the MLP, RF and SVM, in total 12 independent imaging experiments were realized for MDCK and RWPE cysts; 6 for each cell line over a six-year period. In total 479 MDCK cysts were used for this experiment (266 wild type and 213 $KRas^{V12}$ samples, imaging experiments 0-5 in Appendix A). In addition, wild type and $KRas^{V12}$ RWPE samples were prepared and imaged holographically (containing 85 wild type and 114 $KRas^{V12}$ cysts, imaging experiments 6, 8, and 10-13 in Appendix A). The reason for having results for different classifiers is to show that all the different classifiers display the same trend in classification accuracy. By adding more features, but not necessarily all, classification accuracy can be increased. The data was split into training (90 %, 431) and testing (10 %, 48) sets. Features were standardized to have a zero-mean and standard deviation of 1. The data splitting was performed randomly for each run; the total number of runs being 1000. For each of the classifiers, multiple training runs and tests were realised by using a different number of features. Features were added in the order described in Tables 4.3 and 4.4 (page 60). This approach aligns with the simplification theme in this thesis.

For the CNN the same data (segmented phase reconstructions) was used, except 10 % of data was extracted for validation, leaving 80 % for the training data.

4.3.1 Multilayer perceptron

Extracted features (n=19) were used with a multilayer perceptron that had a single hidden layer with 3 neurons. ReLu was used as an activation function, the regularisation parameter was 0.001, the maximum number of iterations was 5000, and the tolerance for the optimisation was 1×10^{-6}. The tolerance for optimisation was used to determine if learning was improving. When the loss score does not improve by at least this value, convergence has been reached and training is finished. The optimiser was the limited-memory BFGS [250], which was found to result in faster convergence than the stochastic gradient descent.

4.3.2 Support vector machine

The support vector machine classifier was trained with the radial basis function kernel as that was found to provide better results than the linear kernel, the penalty parameter was 1, the kernel coefficient (γ) was determined by

$$\gamma = \frac{1}{n} \tag{4.5}$$

where n is the number of features, the shrinking heuristic was used, and the stopping criterion was set to 0.001.

4.3.3 Random forest

The random forest classifier contained 15 trees, and used the Gini impurity for the information gain. The maximum number of features for the best split was set to the number of features (1-11), and the minimum number of samples to split an internal node was set to 2. The nodes were expanded until all the leaves were pure or until all leaves contained less than the minimum number of samples, where the minimum number of samples required to be at a leaf node was set to 1. An unlimited number of leaf nodes were allowed, and the early stopping tree growth threshold was set to 10^{-7}.

4.3.4 Convolutional neural networks

Although conventional classifiers perform well with extracted data, as shown later in this chapter, we wanted to see if classification performance can be improved with more complex models. For this reason we used deep convolutional neural networks. In addition, conventional neural networks perform well with the data captured at the same time than the training data. However, if these classifiers are tested with the data captured at a different time, accuracy decreases or classifiers fail completely (this is shown later in Chapter 8, Fig. 8.2).

Two different deep convolutional neural networks were used (see Fig. 4.3). The first was based on AlexNet [121], which won the classification and localization tasks in the Large Scale Visual Recognition Challenge 2012. AlexNet has 5 convolution layers, 3 fully-connected layers, and uses convolutional filters up to 11×11 pixels in size. The second, VGG16, was based on a variant of the VGG [123] architecture that was successful in the classification task in the same challenge in 2014. It has 13 convolution layers, 3 fully-connected layers, and learns smaller features with its 3×3 pixel convolutional filters.

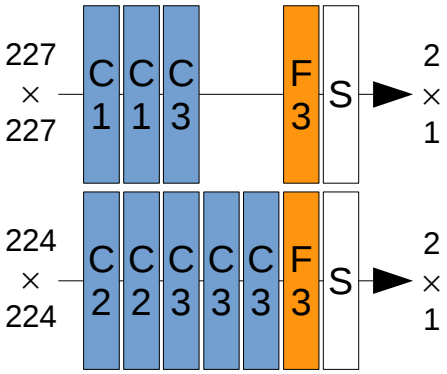

Figure 4.3 Network architectures based on AlexNet (above) and VGG16 (below): C, convolution block; F, fully-connected block; input size, 227×227 or 224×224 pixels depending on the network. Numbers show amount of layers in each block. Each convolution block is followed by a maxpooling layer (with kernel size of 3 and stride of 2 in AlexNet, and with kernel size of 2 and stride of 2 in VGG16).

Both of these network architectures are well-known and highly accepted models for a wide range of real-world image-based applications. In this chapter, we show that these standardised models can be used as-is with holographic data from biological samples, and that it is not necessary to invent a custom architecture for each new application. This aligns with the simplification theme in this thesis. The only modifications made were as follows. Both networks used batch normalisation. In addition, as our data contained a single channel, these network configurations were modified to take a single channel input instead of RGB input, while the rest of the network architectures were kept the same.

Cysts were segmented (the approach used is not described until Chapter 7). Square regions around the cysts in the reconstructed phase images were segmented. Each segmented cyst was scaled to an appropriate size as required by the network architecture and was randomly augmented during run-time by a random combination of rotations by 90 degrees, horizontal mirroring, and vertical mirroring.

4.4 Experimental results

Figure 4.4 shows example MDCK holograms and phase reconstructions. From Fig. 4.5 it can be observed that wild type MDCK holograms display higher maximum phase values than $KRas^{V12}$ transformed holograms. To display the importance of this feature, the maximum phase values were extracted and principal component analysis was performed on features other than maximum phase. A plot of the maximum phase and the 1st principal

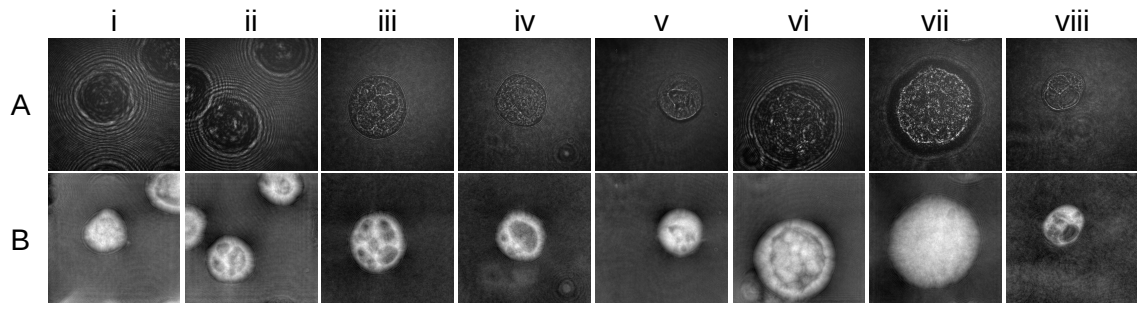

Figure 4.4 Example holograms and phase reconstructions of MDCK samples. A: holograms, B: in-focus unwrapped phase reconstructions of A.

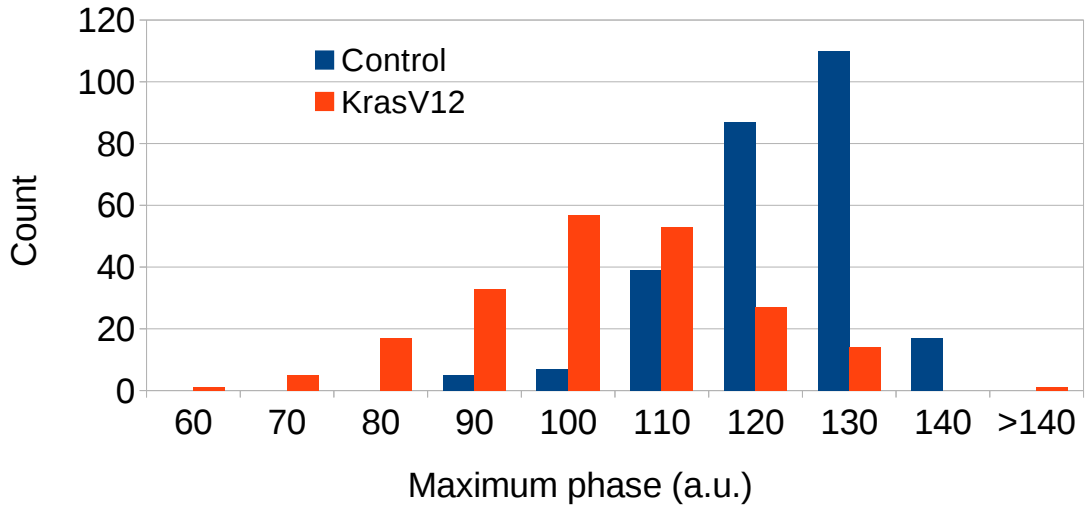

Figure 4.5 MDCK maximum phase histogram. Feature extraction on MDCK samples using DHM and quantitative phase analysis revealed that maximum phase (shown here) was the most representative individual feature for MDCK cysts from the set of features examined.

component on vertical and horizontal, respectively, is shown in Fig. 4.6. It can be observed that samples of two phenotypes are partially overlapping. The importance of maximum phase as a feature is demonstrated by running an analysis of variance (ANOVA) between two different sample populations on all extracted features. The results of ANOVA are shown in Table 4.3 where features are ordered in descending order by the F-values obtained from ANOVA. It can be observed that the maximum phase value is the most important feature. Nearly all of the extracted features within these populations are statistically significant (p-value < 0.05).

By running ANOVA on the RWPE samples, it is observed that lumen index is the most important single feature. As with the MDCK samples, nearly all extracted features are statistically significant between wild type and $KRas^{V12}$ samples (p-value < 0.05). These

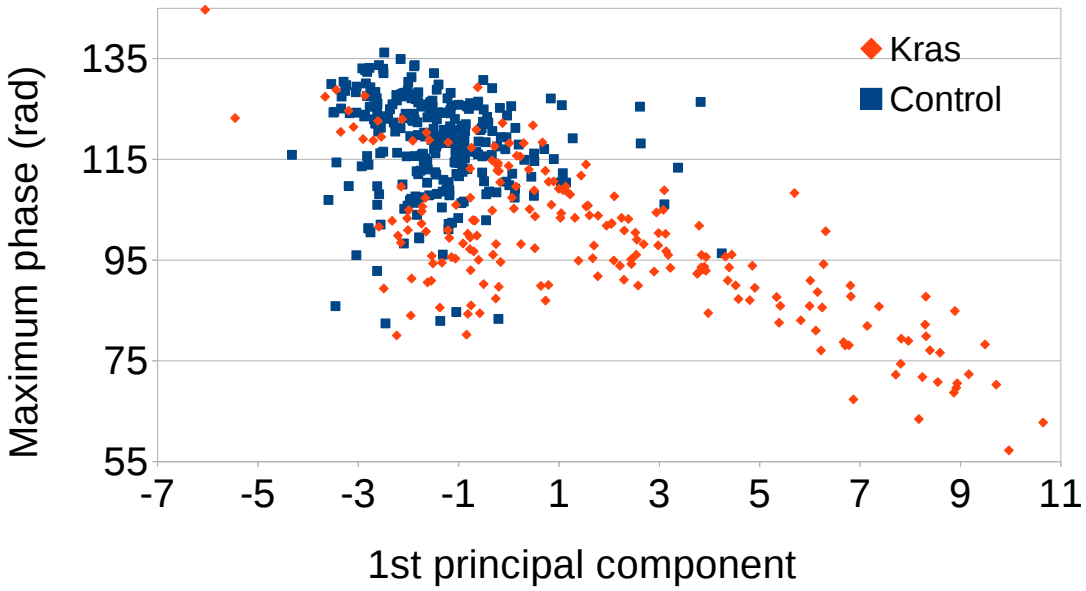

Figure 4.6 MDCK maximum phase and 1st principal component. Showing maximum phase values on vertical axis and 1st principal component of other features on horizontal axis.

Table 4.3 Features and their corresponding F and p values from ANOVA on MDCK samples. Features below the dashed horizontal line have p value > 0.05.

Feature	F-value	p-value
Maximum phase	304.3	6.2×10^{-53}
Skewness	204.6	8.4×10^{-39}
Kurtosis	172.0	1.0×10^{-33}
Minor axis length	139.0	2.8×10^{-28}
Perimeter	134.0	1.9×10^{-27}
Convex area	133.8	2.1×10^{-27}
Area	133.0	2.83×10^{-27}
Equivalent diameter	130.7	7.1×10^{-27}
Lumen index	120.6	3.9×10^{-25}
Standard deviation	117.9	1.1×10^{-24}
Major axis length	111.0	1.9×10^{-23}
Minimum phase	83.5	1.9×10^{-18}
Density	70.3	5.8×10^{-16}
Roughness	68.1	1.5×10^{-15}
Average	24.2	1.2×10^{-6}
Solidity	19.1	1.5×10^{-5}
Extent	15.5	9.7×10^{-5}
Eccentricity	1.7	0.19
Roundness	0.9	0.34

Table 4.4 Features and their corresponding F and p values from ANOVA on RWPE samples. Features below the dashed horizontal line have p value > 0.05.

Feature	F-value	p-value
Lumen index	165.5	6.8×10^{-28}
Average	56.6	1.8×10^{-12}
Minimum phase	49.6	3.0×10^{-11}
Density	43.2	4.4×10^{-10}
Minor axis length	39.3	2.3×10^{-9}
Equivalent diameter	35.6	1.1×10^{-8}
Area	24.8	1.4×10^{-6}
Convex area	22.6	3.8×10^{-6}
Major axis length	22.2	4.7×10^{-6}
Perimeter	18.5	2.6×10^{-5}
Kurtosis	18.5	2.7×10^{-5}
Standard deviation	10.6	1.3×10^{-3}
Skewness	10.1	1.7×10^{-3}
Extent	5.3	0.02
Roughness	5.1	0.03
Solidity	3.4	0.07
Eccentricity	2.2	0.14
Roundness	1.8	0.18
Maximum phase	0.1	0.71

results are shown in Table 4.4. A histogram showing lumen indices of wild type and $KRas^{V12}$ is depicted in Fig. 4.7. It can be observed that the two populations differ greatly on this extracted feature. $KRas^{V12}$ displays low lumen index values and is positively skewed; most of the values are below 0.1. The lumen index values of wild type cysts have a larger variance, the distribution is more symmetrical, and values spread more evenly. The range of the two distributions is approximately equal, which is not surprising as with these two phenotypes it is typical to have $KRas^{V12}$ cysts that look wild type and vice versa. By extracting lumen indices from the feature vectors and performing principal component analysis on other features, it can be observed that the two phenotypes can almost perfectly be linearly separated using lumen index and the 1st principal component (Fig. 4.8).

A normalised cross-correlation between features is shown in Fig. 4.9. It can be observed that size related features with all sample types have strong correlations, as expected. Phase related features do not show the same trend. By looking at MDCK control samples, there are strong correlations between average, lumen index, and skewness; in addition, density is correlated with size-related features and maximum phase with standard deviation. With $KRas^{V12}$ samples, the lumen index is correlating with skewness, maximum phase with

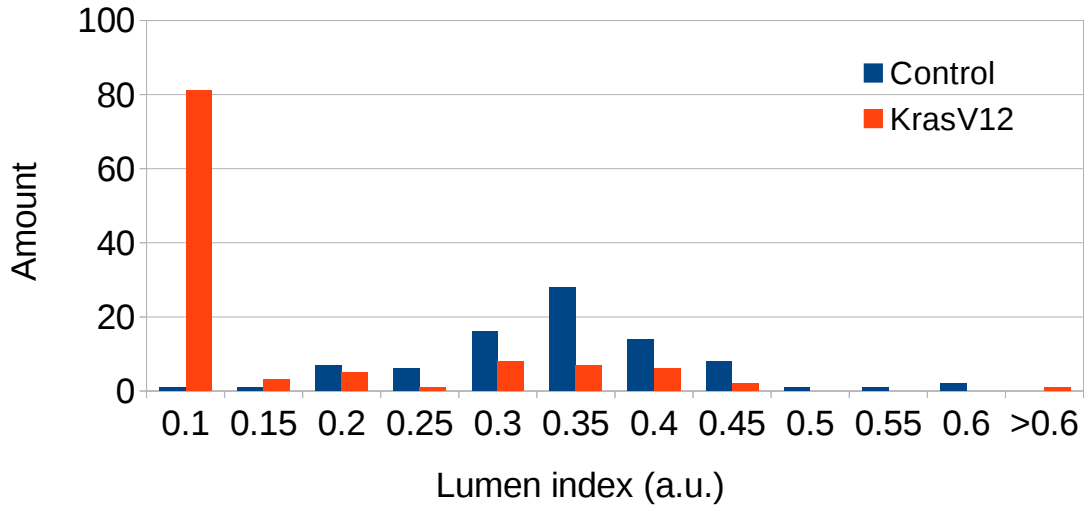

Figure 4.7 Lumen index on RWPE samples. Feature extraction on RWPE samples using DHM and quantitative phase analysis revealed that lumen index (shown here) was the most representative individual feature for RWPE cysts from the set of features examined.

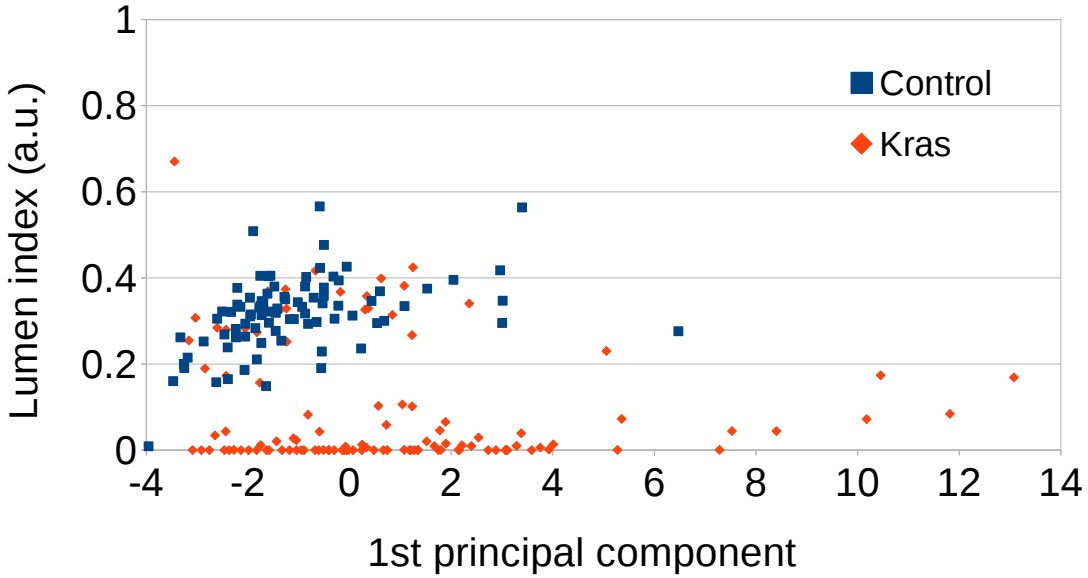

Figure 4.8 RWPE lumen index and 1st principal component, shows lumen index on the vertical axis and the 1st principal component of the other features on horizontal axis.

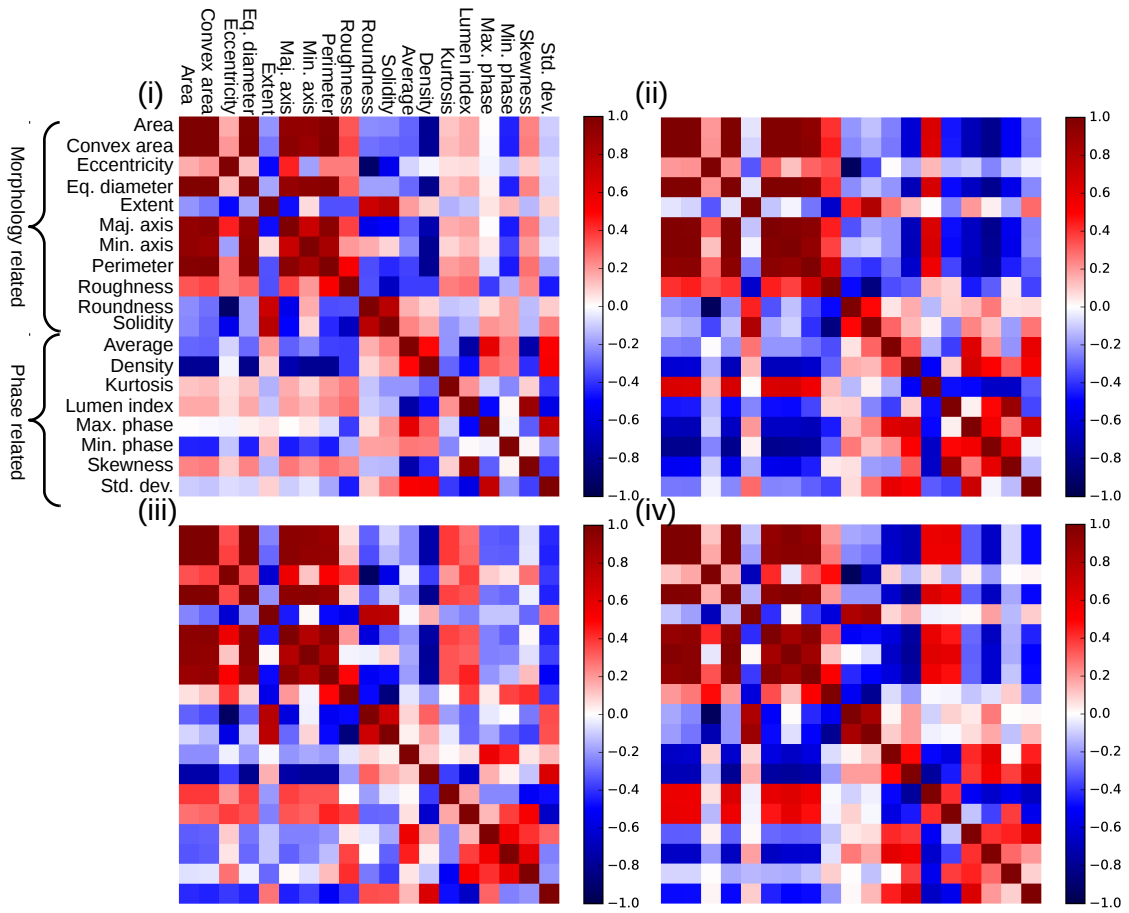

Figure 4.9 Normalised cross-correlation between extracted quantitative features. (i) MDCK wild-type, (ii) MDCK KRas^{V12}, (iii) RWPE wild-type, (iv) RWPE KRas^{V12}. See text for more explanation.

size related features and standard deviation, and minimum phase with some of the size related features.

With RWPE samples, strong correlations between phase related features are almost non-existent with few exceptions. Density is correlated with size-related features for both control and KRas^{V12}. In addition, with KRas^{V12} sample, density is correlated with kurtosis.

4.4.1 Multilayer perceptron

Results of using a different number of features are shown in Fig. 4.10. With MDCK samples, by using only the maximum phase value, the median classification accuracy was 81.3 %. By using a full set of features the median accuracy with the test data was 87.5 %. It can be observed that the multilayer perceptron provides the best median classification accuracy with only the two top features; the median accuracy being 89.6 %. With additional

features, there is no improvement or decrease in median classification accuracy. The lowest standard deviation of 4.3 was obtained with two features. With the RWPE samples, by using only the lumen index, the median classification accuracy was 80.0 %. The best median accuracy and lowest standard deviation was obtained by using six features (Fig. 4.11). With six features, the median classification accuracy was 85.0 % and the standard deviation was 7.7.

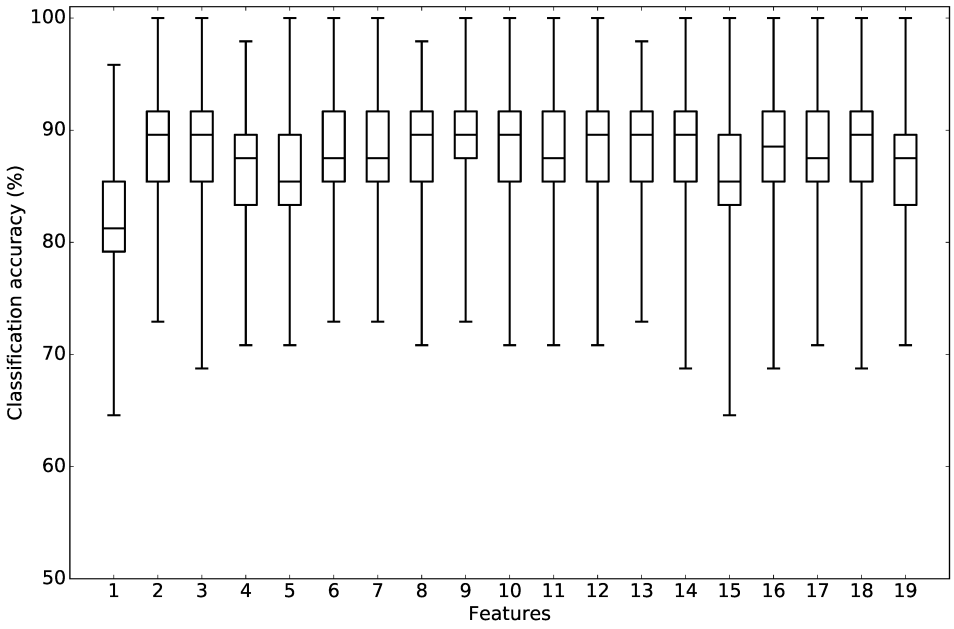

Figure 4.10 Multilayer perceptron classification accuracy by using different number of features with MDCK samples. Showing minimum, maximum, median, 25th, and 75th percentiles.

4.4.2 Support vector machine

With MDCK samples, and using only the maximum phase, the median classification accuracy was 83.3 %. Results of using a different number of features are shown in Fig. 4.12. It can be observed that the same median accuracy of 91.7 % can be achieved by using any number of features from 14 to 19. Of these, the one with 18 features has the lowest variation with a standard deviation of 3.8.

With the RWPE samples, 85.0 % median accuracy could be achieved with any number of features (Fig. 4.13). The smallest standard deviation was obtained by using 12 features.

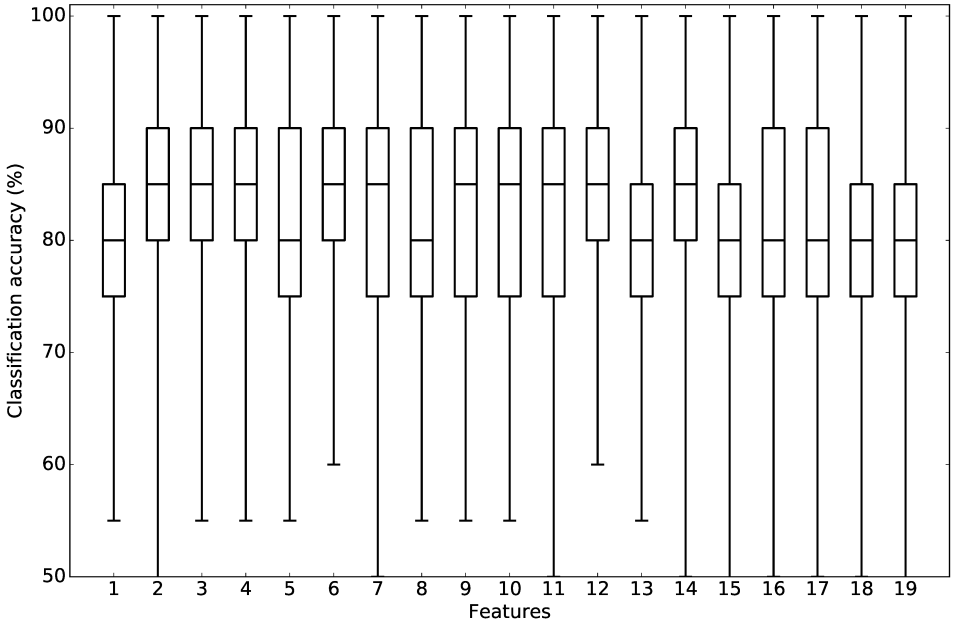

Figure 4.11 Multilayer perceptron classification accuracy by using different number of features with RWPE samples. Showing minimum, maximum, median, 25th, and 75th percentiles.

4.4.3 Random forest

With MDCK samples, and using only the maximum phase value, the median classification accuracy was 72.9 %. Results of using a different number of features are shown in Fig. 4.14. It can be observed that the same median accuracy of 91.7 % can be achieved by using any number of features from 14 to 19. Of these, the one with 16 features has the lowest variation with a standard deviation of 4.0.

With the RWPE samples and using only the lumen index, the median classification accuracy was 75.0 %. A 85 % median classification accuracy could be achieved with any number of features from three to 19 (Fig. 4.15). The smallest standard deviation of 7.1 was obtained by using 8 features.

4.4.4 Convolutional neural networks

Figure 4.16 shows training and validation loss curves together with validation accuracy over epochs for two typical training runs. There are no obvious signs of overfitting and

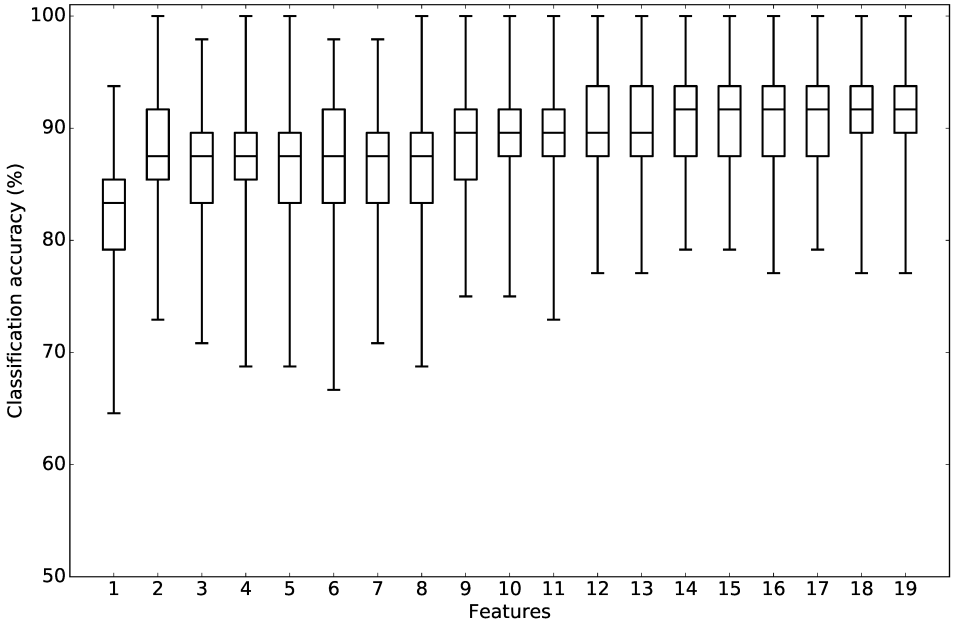

Figure 4.12 Support vector machine classification accuracy by using different number of features using MDCK samples. Showing minimum, maximum, median, 25th, and 75th percentiles.

the validation loss follows the training loss. The average classification accuracies over ten runs of AlexNet and VGG16 were 88.0 % and 81.4 %, respectively.

4.4.5 Results summary

For this particular application, and with the data used, the support vector machine and random forest give the best classification accuracy, with a best median accuracy of 91.7 %. Figure 4.17 shows receiver operating characteristic (ROC) curves and area under curve (AUC) values for MLPs, SVMs, RFs, and CNNs for one example run. It can be observed that all classifiers perform very well in their task. The AUC values of all classifiers are greater than 0.9 denoting an excellent classifier capability.

4.5 Conclusions

It was shown that extracted data from a segmented phase reconstruction can be used to classify cysts as tumourigenic or healthy. A novel metric to define a condition of the lumen was introduced. Classification was performed by using conventional classifiers.

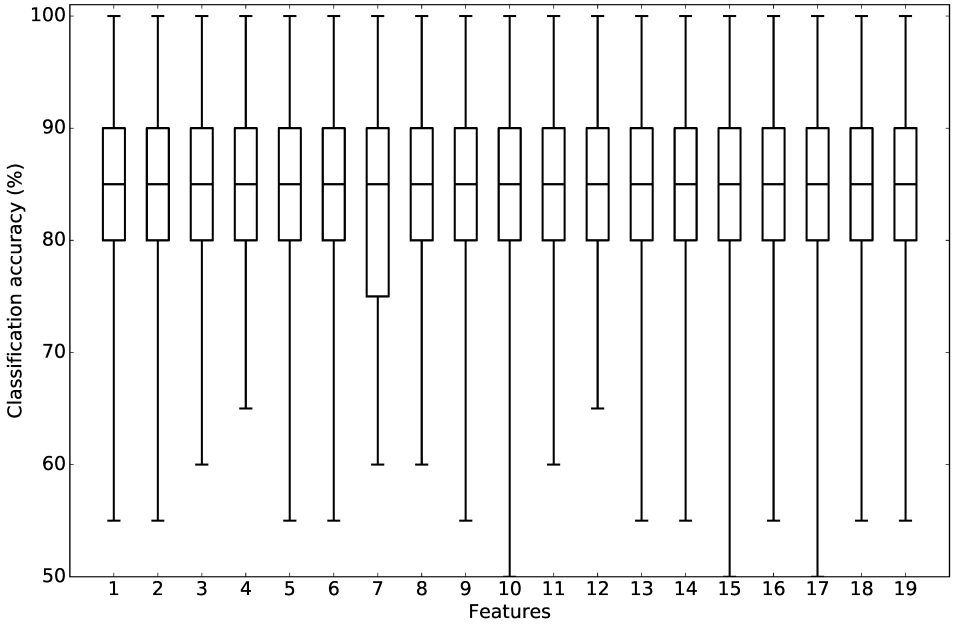

Figure 4.13 Support vector machine classification accuracy by using different number of features using RWPE samples. Showing minimum, maximum, median, 25th, and 75th percentiles.

In addition, two different CNNs were trained to conduct the same classification task. In total, classification was executed on five different classifiers and the performance of them all, based on their ROC curves, was excellent. It needs to be noted that test performance was measured only on data that was captured at the same time as the training data. If the conventional classifiers were tested on data captured on a different day, they would fail. This is shown later in Chapter 8.

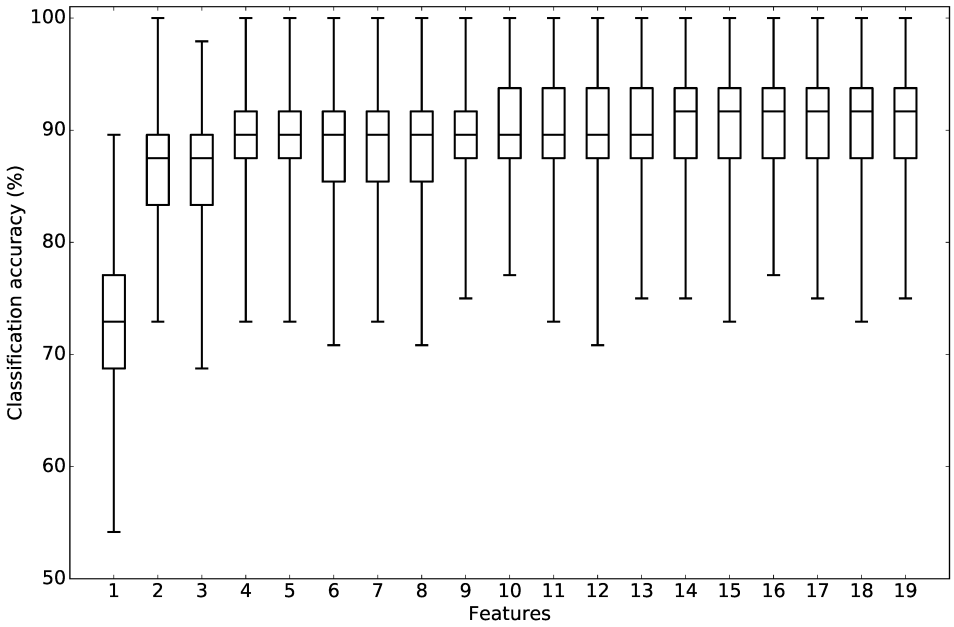

Figure 4.14 Random forest classification accuracy for different numbers of features on MDCK samples. Showing minimum, maximum, median, 25th, and 75th percentiles.

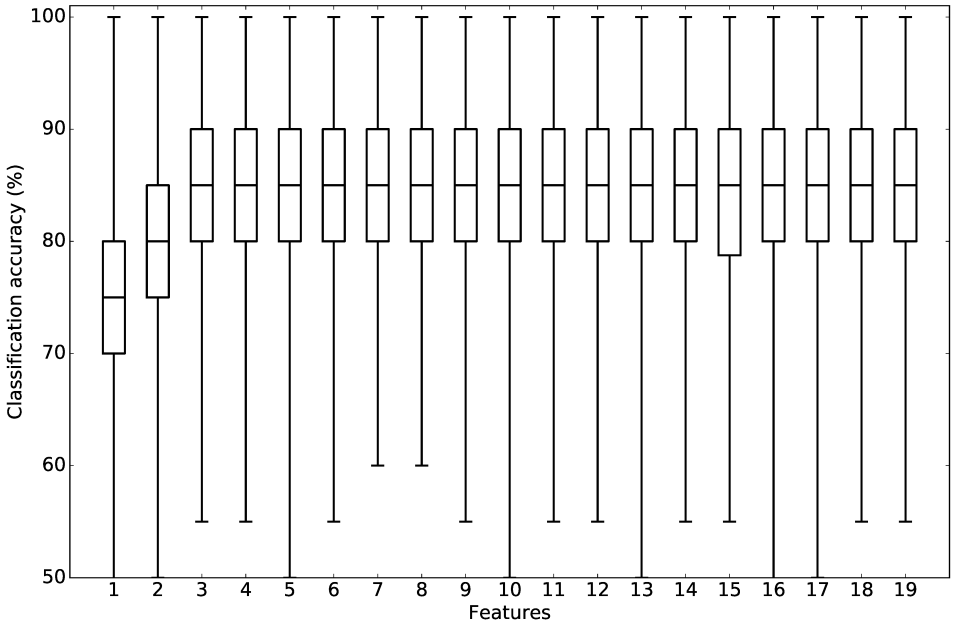

Figure 4.15 Random forest classification accuracy by using different number of features using RWPE samples. Showing minimum, maximum, median, 25th, and 75th percentiles.

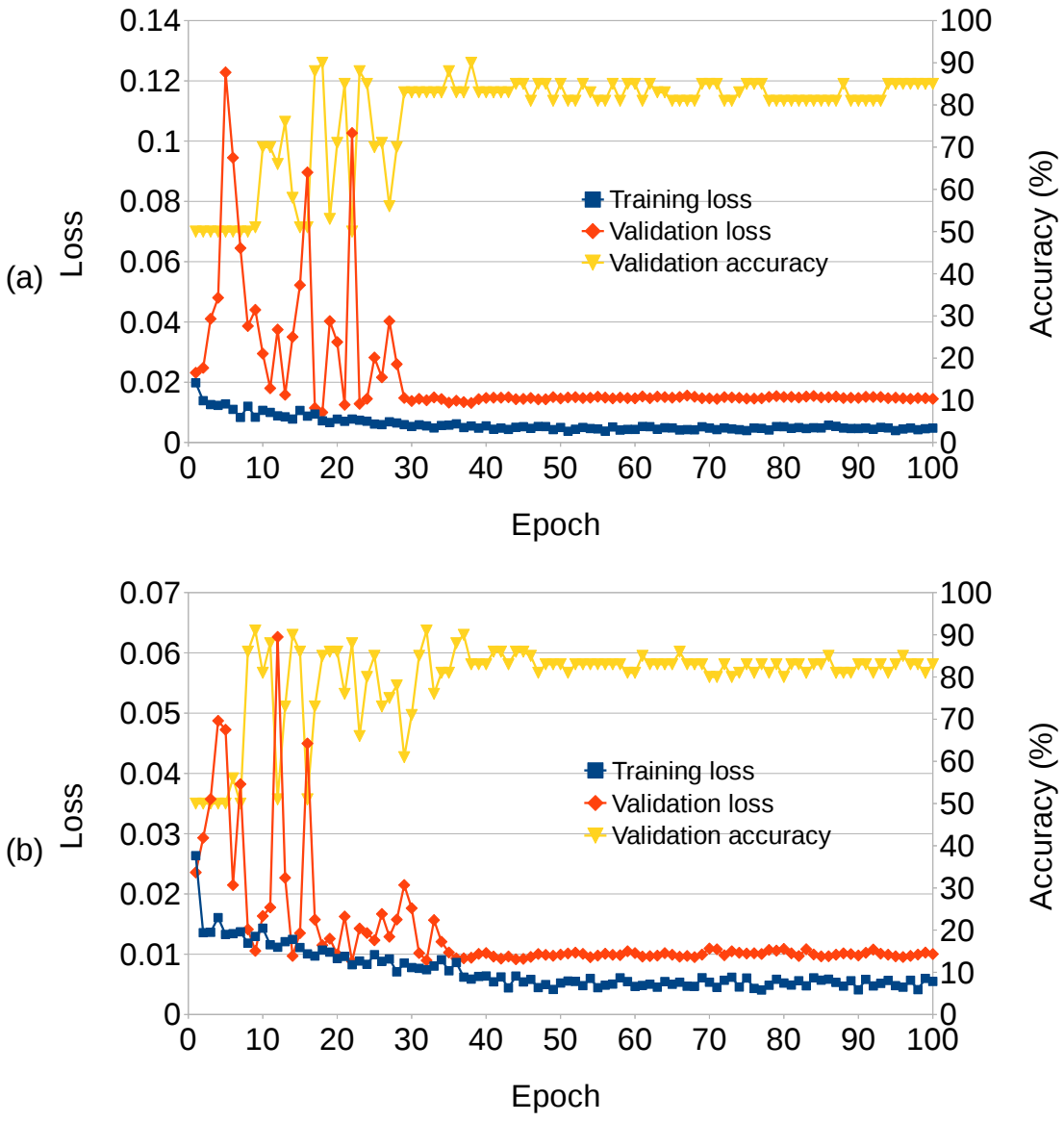

Figure 4.16 Training loss, validation loss, and accuracy curves for (a) AlexNet, (b) VGG16.

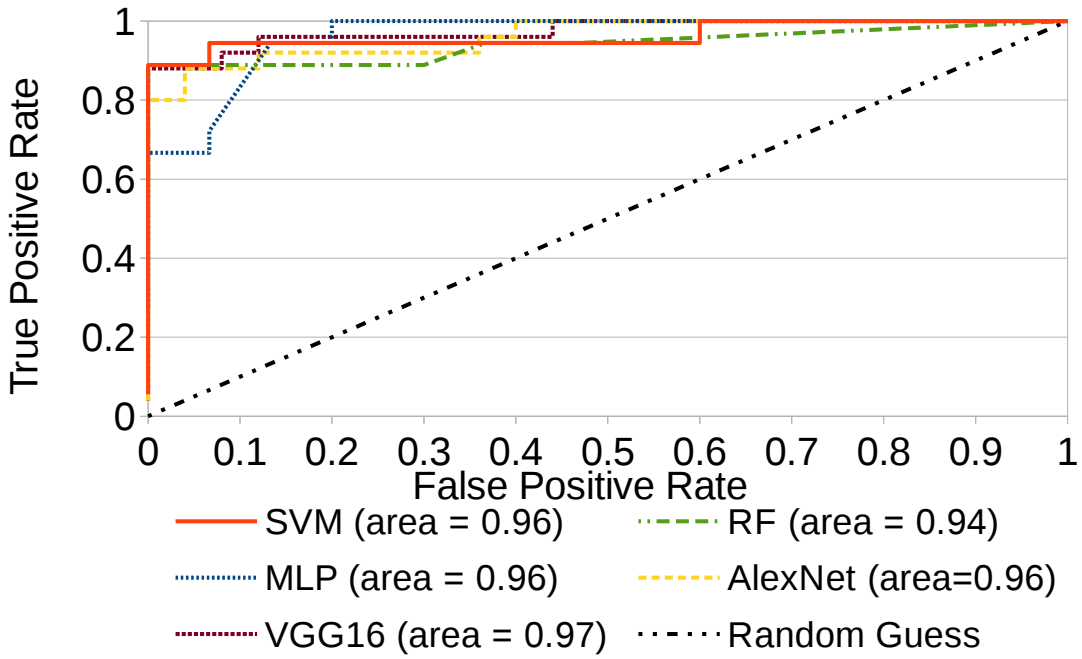

Figure 4.17 ROC curves for MLP, SVM, RF, AlexNet and VGG16. See text for more explanation.

5 Object detection in the field of view

This chapter introduces a novel approach that can be used for detecting if an object is in the field of view. This approach is parametrised based on analysis of deep learning results.

5.1 Introduction

Several approaches, such as [251, 252], have been proposed to determine the in-focus-depth of a transparent object. However, each of these methods assume that the hologram contains an object in its field of view (FOV). Furthermore, well-known techniques to segment a transparent object from a hologram reconstruction volume (one approach that as a side-effect can detect the presence or absence of an object) requires multiple reconstruction steps per hologram. Although straightforward to solve for reflective objects, efficiently detecting the presence or absence of a transparent object of interest in a noisy medium can be challenging.

When DHM is used for time-lapse imaging, particle field imaging, or any other application where a large number of holograms are captured, many holograms can be recorded without an object in their FOV. Subsequently processing these holograms as if they contained objects can be wasteful. To overcome this inefficiency, we propose a solution that can be used to detect the presence, or alternatively absence, of a transparent object in the FOV before any numerical processing on a captured digital hologram is performed, including for example searching for the in-focus plane.

5.2 Object detection

We propose a histogram-based approach to quantify gross features in the Fourier domain of digital holograms. Given a $M \times N$ hologram I we Fourier transform it and extract the amplitude to give us I'. The metric to determine low frequency content of a hologram is calculated as

$$\alpha = \frac{1}{mn} \sum_{x=0}^{m-1} \sum_{y=0}^{n-1} \Omega(x,y)\,, \tag{5.1}$$

where Ω is an $m \times n$ ($m << M$ and $n << N$) binary mask containing ones in a particular region of I'. The size and shape of the region was determined experimentally as

$$\Omega(x,y) = \begin{cases} 1, & \text{if } I'(x+u, y+v) > \tau \\ 0, & \text{otherwise}\,, \end{cases} \tag{5.2}$$

where $x \in [0, m-1]$, $y \in [0, n-1]$, τ is some threshold, and u and v are empirically chosen regions defined as

$$u = \frac{M}{2}, \quad v = \frac{5}{16}N\,. \tag{5.3}$$

I' is the amplitude of the two-dimensional Fourier transform of a $M \times N$ pixel hologram. An example amplitude of a hologram containing an object in the FOV showing the experimentally chosen region used for analysis is shown in Fig. 5.1.

Because the α metric requires a hyperparameter τ that might need to be tuned to the application at hand, the metric was also compared, on the basis of accuracy and computational efficiency, with an approach employing a well-known deep learning convolutional neural network as a binary classifier, trained on Fourier-domain amplitude inputs.

The CNN contained five convolution layers, four max pooling layers and two fully-connected layers (Fig. 5.2). The kernel size for all convolution layers was 3×3 pixels, except for the first layer that had 9×9 pixel kernels. The stride of the convolutions was always one pixel. Each convolution layer was followed by a batch normalisation layer, and then an activation function. The parametric rectified linear unit (PReLu) [253] was used for the activation function. The pooling layers had 2×2 pixel kernels with a 2-pixel stride, except for the last pooling layer that had 14×14 pixel kernels with a 13-pixel stride. These

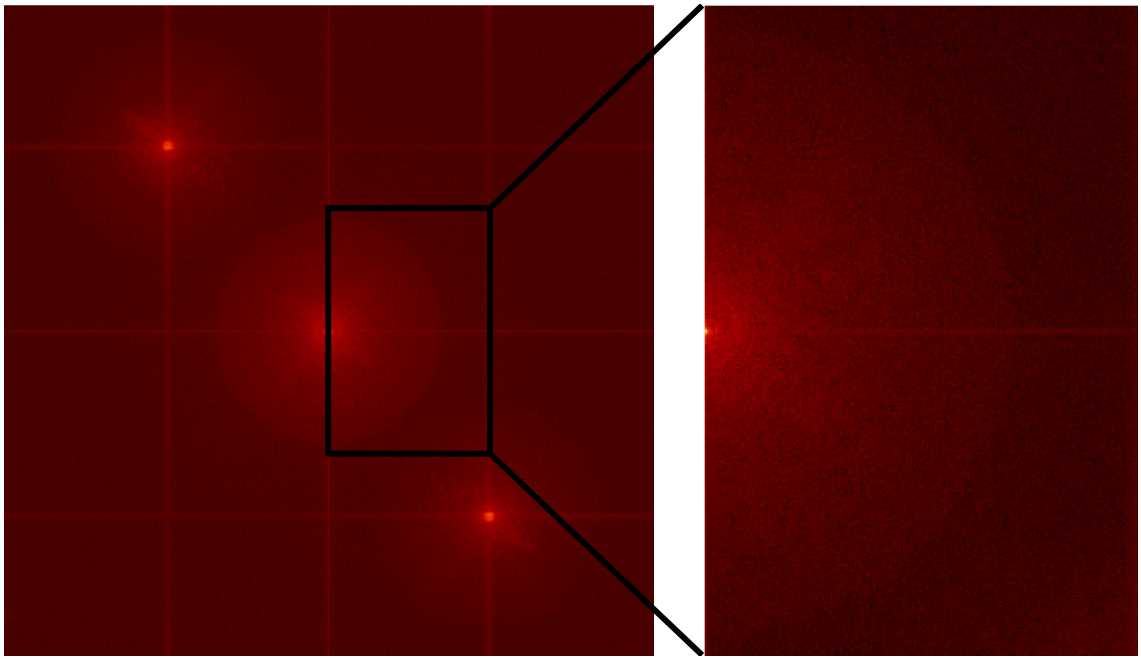

Figure 5.1 Fourier transform of a hologram with an object in the FOV. Black lines show borders of an empirically selected region. Zoomed-in version of this region is shown on right.

parameters for the last pooling layer were used to reduce the data dimensions before the fully-connected layers.

As data for the CNN, 290 holograms with, and 290 holograms without, an object (MDCK cell line) in the FOV were used (example holograms are shown in Fig. 5.3) (from the imaging experiment 23 in Appendix A). The data was split into training (80 %, 464), validation (10 %, 58), and testing (10 %, 58). Each subset had a similar number of examples of each class. The input for the CNN was the amplitude of the two-dimensional Fourier transform of a hologram. 1024×1024 pixel Fourier amplitude images were down-sampled (with interpolation) using the scikit-image library [246] to 224×224 pixels to allow for greater batch sizes.

PyTorch [248], a deep learning framework, was used. Training was executed on two Nvidia GTX 1080 graphics cards with a batch size of 50. The network was trained for 100 epochs with Adam [254] as the optimiser and categorical cross entropy loss. The initial learning rate was set to 1×10^{-4} and was reduced by a factor of 10 if the validation loss failed to improve for ten consecutive epochs.

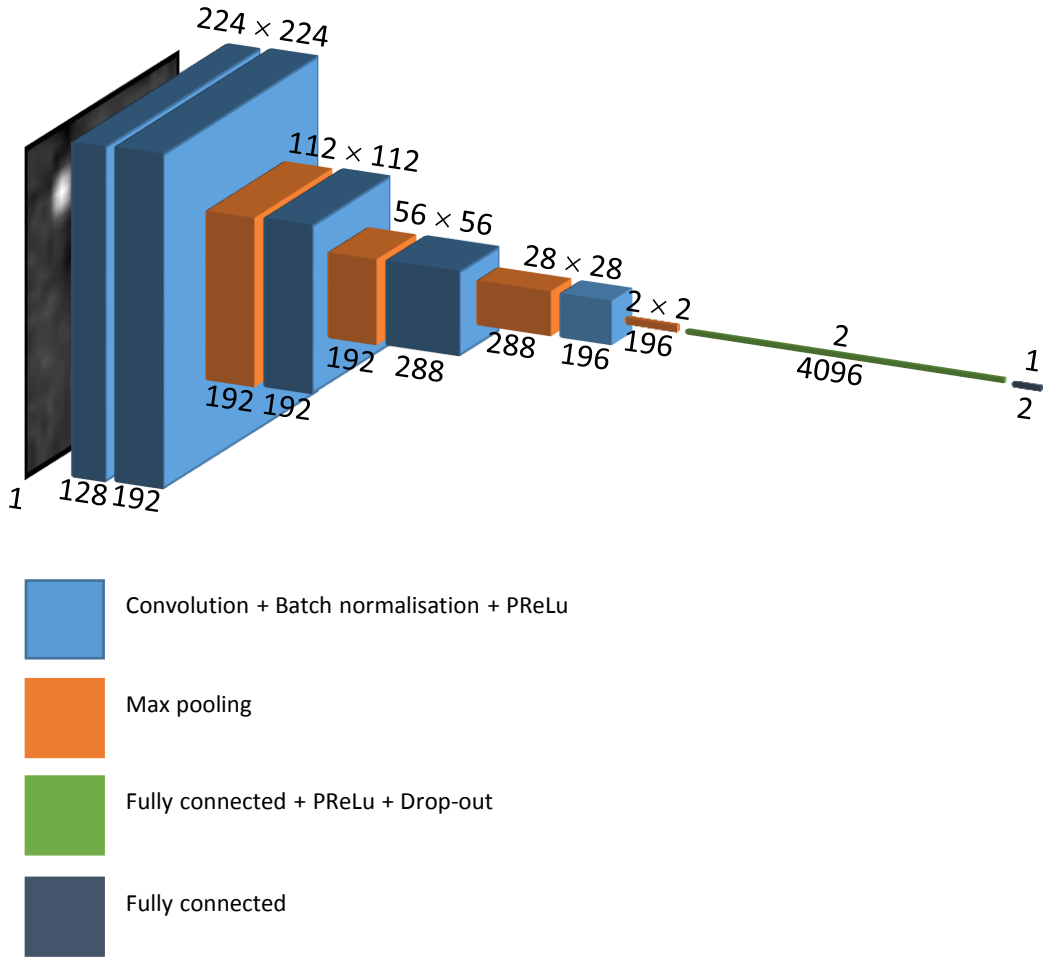

Figure 5.2 CNN architecture showing the different layers, and their sizes. Below each layer the number of output channels is shown. Numbers above denote spatial dimensions. See text for more explanation.

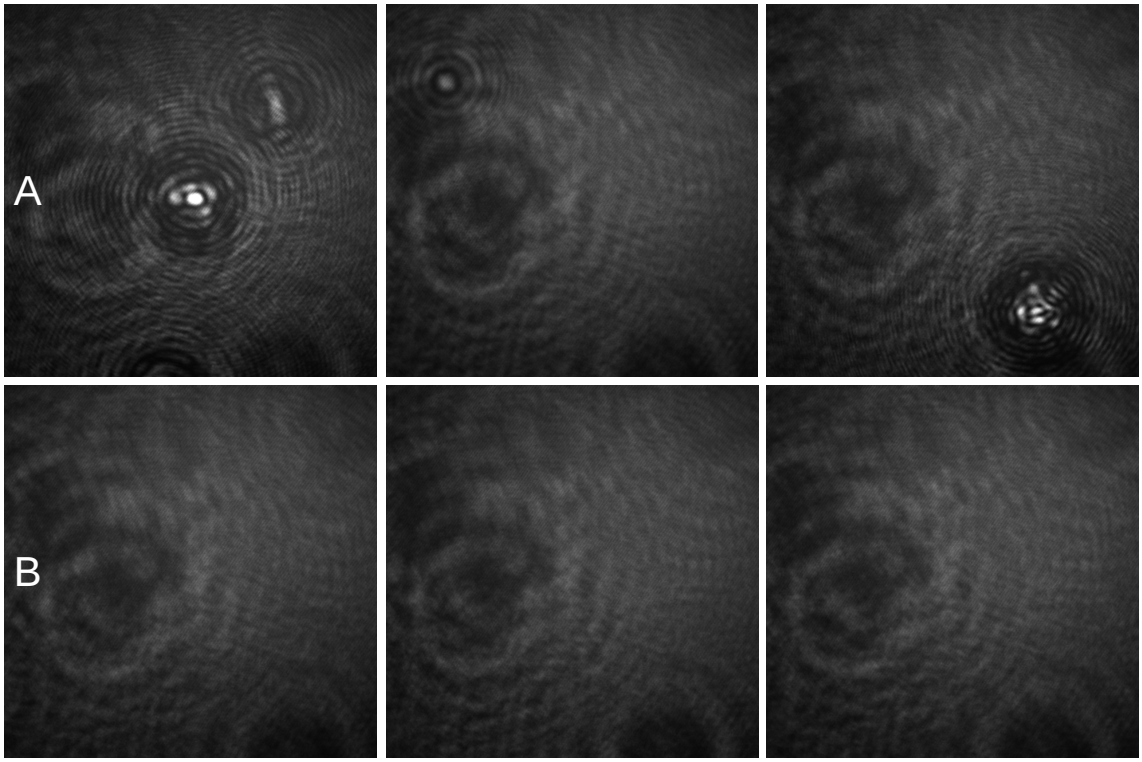

Figure 5.3 Example holograms. Holograms in row A contains objects in the FOV whereas holograms in row B are without objects in the FOV.

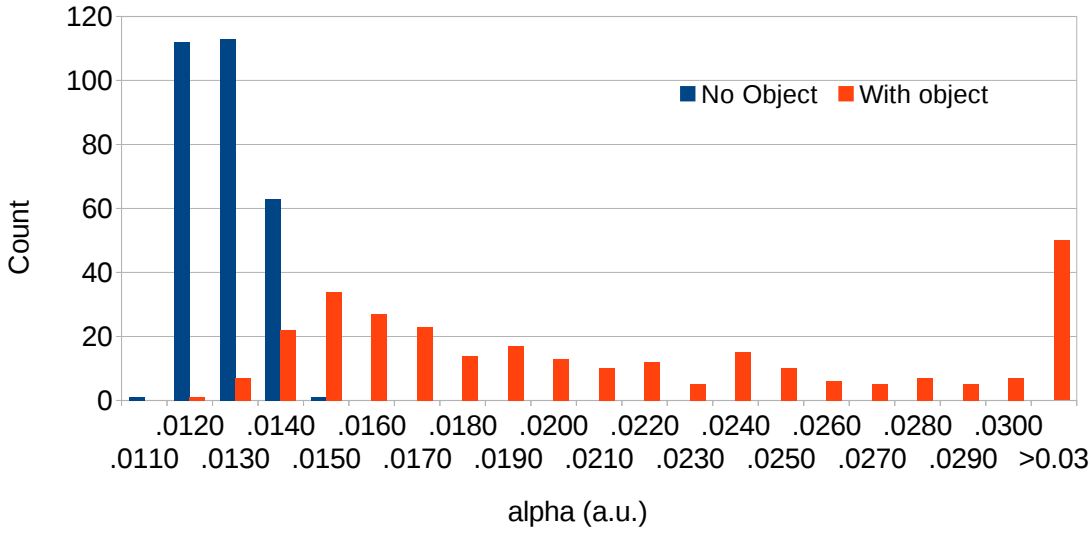

Figure 5.4 α results with an empirically set Fourier region. See text for more explanation.

Table 5.1 FOV confusion matrix, heuristic approach. Showing predicted and actual classes.

		Actual	
		Yes	No
Predicted	Yes	289	30
	No	1	260

5.3 Experimental results

A good value for τ was empirically found to be $max(I')/1000$. Results with this value are shown in Fig. 5.4. By setting the α threshold to 0.014, we can respectively identify 99.7 % and 89.7 % of holograms that do not and do have an object in the FOV (Table 5.1). The average accuracy with this threshold is 94.7 %. Loss values and validation accuracy for the CNN approach are shown in Fig. 5.5. The average CNN testing accuracy over three runs was 97.2 %. These values show that the CNN approach gives a better accuracy. However, it raised new questions:

- Why is the CNN so much better?
- Can we improve our heuristic approach by changing the Ω region?

If the performance of a heuristic approach is sufficient, clear advantages are:

- it will be faster than a CNN approach,
- it can potentially be used with several input types (different cell types),
- if assumed that the same threshold can be used, no training is needed, and
- it has more relaxed constraints (twin terms can be located at any position in the Fourier domain).

To answer these newly raised questions, we investigated which features/regions the CNN considers the most important, and if it was possible to construct our mask based on this analysis. This approach is in alignment with our paradigms theme. We ran gradient-weighted class activation mapping (Grad-CAM) [255] on several Fourier amplitude input images to discover what a CNN considers to be the most distinguishing features. Grad-CAM can be used to produce visual explanations of the functionality of a CNN. It uses gradient information flowing into the last convolutional layer of the CNN to understand the

significance of each neuron for a particular decision (in our case, the presence or absence of an object in the field of view). Regions of these features were used to construct a new region Ω', hence the trained CNN model parametrised our histogram based heuristic approach.

Fig. 5.6 shows example Grad-CAM outputs for seven example inputs and an average Grad-CAM image from 100 inputs. It can be observed that with an object in the FOV, the most important region is around the dc region at the centre of the Fourier amplitude. Without an object in the FOV, the most important features are found around different hologram terms. Based on this, a mask $S(x,y)$ was constructed as follows:

$$S(x,y) = \begin{cases} 1, & \text{if } A(x,y) \geq \tau \text{ and } x \geq \frac{M}{2} \\ 0, & \text{otherwise ,} \end{cases} \tag{5.4}$$

where $A(x,y)$ is an average image (Fig. 5.6, row A, last column) of 100 Grad-CAM outputs with an object in the field of view, and τ is an iteratively obtained threshold. An initial value for τ is given by Otsu's method [256], and is increased iteratively by steps of one until $S(x,y)$ contains only one blob. Fig. 5.7 shows the created mask.

$\Omega'(x,y)$ was obtained as

$$\Omega'(x,y) = \begin{cases} 1, & \text{if } S(x,y)H'(x,y) > \tau' \\ 0, & \text{otherwise ,} \end{cases} \tag{5.5}$$

where I' is the amplitude of two-dimensional Fourier transform of a $M \times N$ pixel hologram, τ' is a threshold defined as

$$\tau' = \frac{k}{1000} \tag{5.6}$$

where $k = \max[S(x,y)I'(x,y)]$.

Figure 5.8 shows α values for all 580 test holograms (290 with and 290 without an object in the FOV). It can be seen that holograms without an object in the FOV show significantly lower α values.

Figure 5.9 shows a histogram comparison of intensity, Tamura coefficient, and α from our proposed method. For the intensity calculation, the hologram plane intensities were

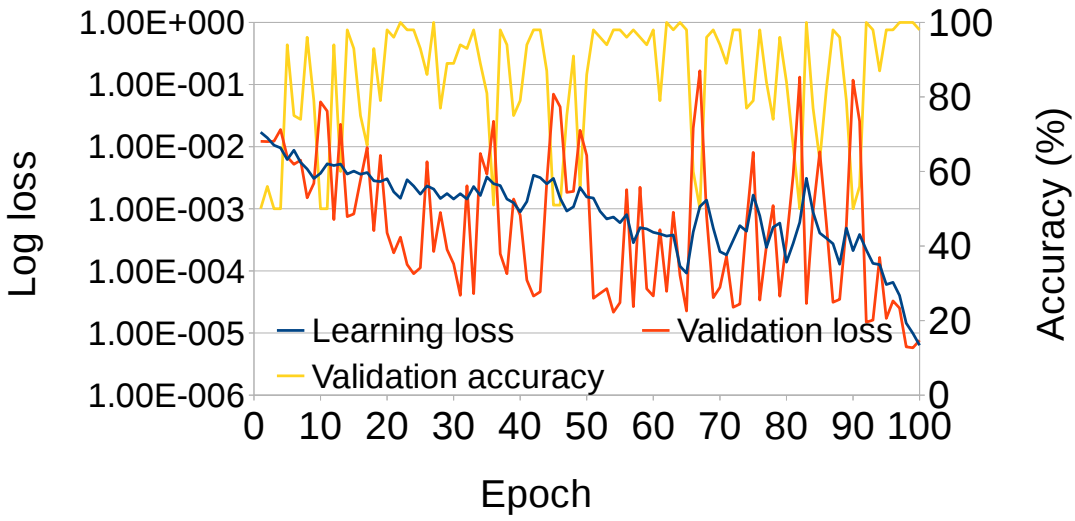

Figure 5.5 Training loss, validation loss and validation accuracy. Although the validation loss fluctuates there are no signs of overfitting.

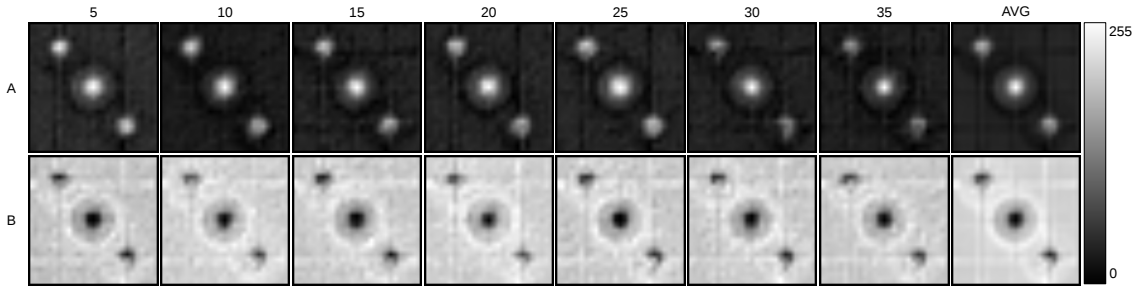

Figure 5.6 Grad-CAM results. Showing high values in regions the trained CNN model considers to contain most important features. Images in row A are with an object in the FOV and ones in row B are without an object in the FOV. Numbers above show each file index and the last column is the average of 100 image inputs.

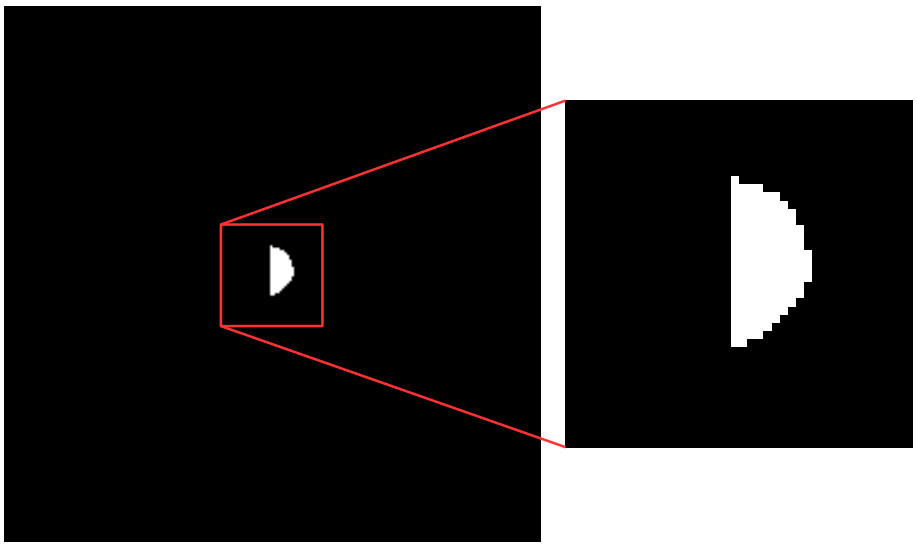

Figure 5.7 Segmentation mask $S(x,y)$. This mask is used to segment the region of Fourier transformed holograms that is used for α metric calculation. See text for more explanation.

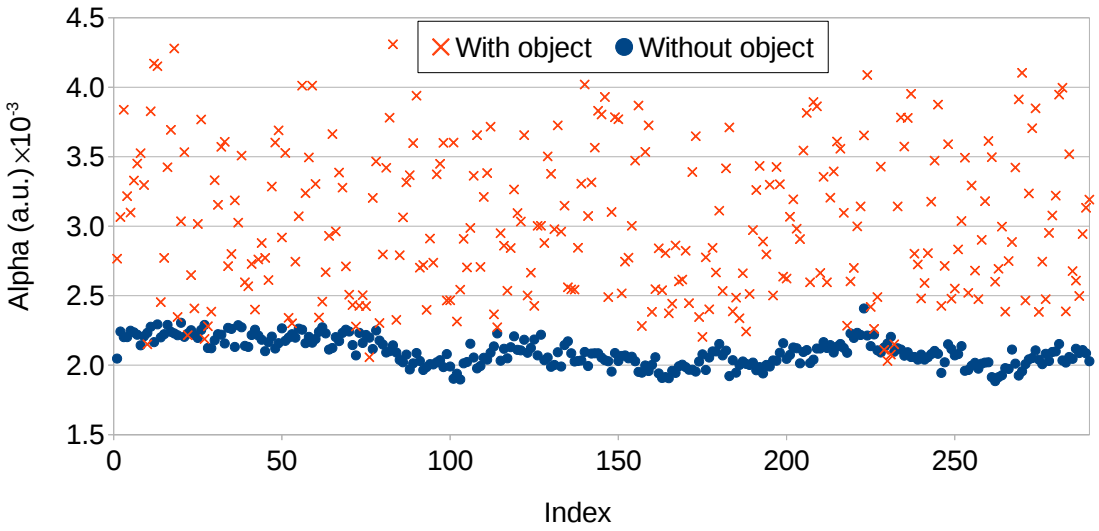

Figure 5.8 α result with 580 test holograms. Showing α values with and without objects in the FOV. It can be observed that holograms without an object in the FOV show, in general, significantly lower α value.

Table 5.2 FOV confusion matrix with the proposed method. Showing predicted and actual classes.

		Actual	
		Yes	No
Predicted	Yes	287	16
	No	3	274

summed. The Tamura coefficient was calculated as described in [251]. It can be observed from these results that the proposed method can be used to efficiently detect the presence or absence of an object with high reliability (see Fig. 5.9c). Intensity is not suitable to detect the presence of an object as can be deduced from Fig. 5.9a. The Tamura coefficient works to some extent (Fig. 5.9b). By setting a threshold to 2.3×10^{-3} (values below that indicate there is no object in the FOV), we can respectively identify 99.0 % and 94.5 % of holograms that do not and do have an object in the FOV (Table 5.2). The average accuracy with this threshold is 96.7 % which is comparable to the 97.2 % classification accuracy of the CNN. This is an obvious improvement to the empirically chosen region that had an average accuracy of 94.7 %.

When benchmarked, a single α value calculation takes 5.4 ms (averaged over 580 holograms) on a PC with 16 GB DDR3 memory and i5-4300U processor. The corresponding runtime for the CNN model (averaged over 58 test holograms) is 70.8 ms, including rescaling the input and converting it to a tensor.

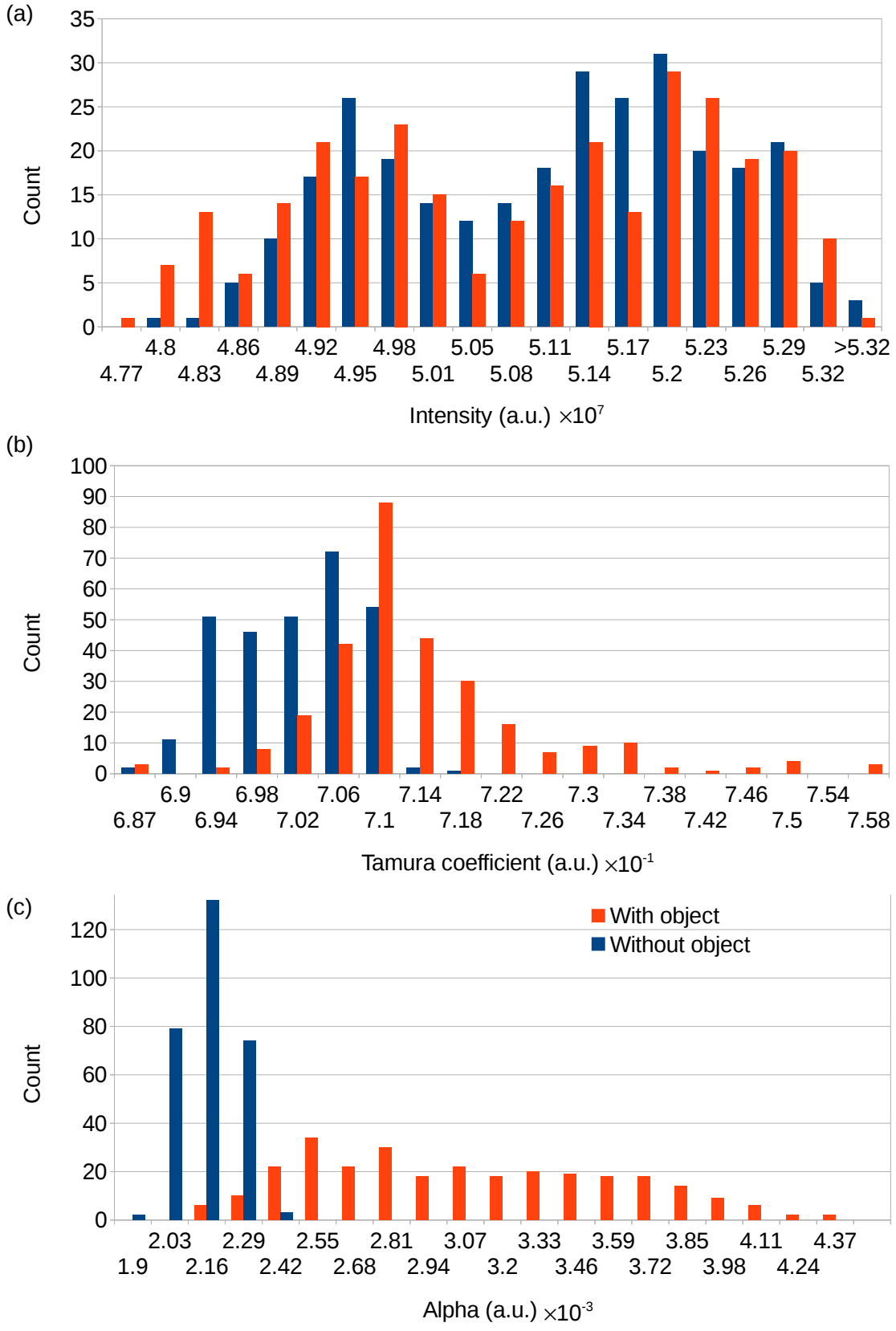

Figure 5.9 Comparison between intensity, Tamura coefficient and the proposed method. See text fore more explanation.

5.4 Conclusions

We have proposed a novel method to parametrise a heuristic approach based on the information obtained from a CNN model. It was shown that the proposed metric is able to identify holograms, with or alternatively without, an object in the FOV with high accuracy. The accuracy is at the same level as the CNN. As a heuristic approach, the proposed approach is fast and therefore facilitates a high data throughput. As such, it is well suited for high-throughput analysis, where only the holograms containing an object(s) should be processed further. The proposed approach was tested only with digital holograms of cysts, however in theory it should work with other types of transparent and semitransparent objects.

6 Deep learning for autofocusing

This chapter contains an application to find the in-plane of an object enclosed in a digital hologram.

6.1 Introduction

Deep learning [119] is a technique for solving hitherto open problems in image analysis, and other fields, that is starting to have an impact in the field of biomedical optics, for example OCT [175, 182, 183] and other forms of microscopy [174, 157, 176, 180, 184].

Image-based applications of deep learning [120] are characterised by neural networks with at least eight hidden layers, at least tens of thousands of images, at least hundreds of images per class, at least millions of learned parameters, and training times of at least weeks if run on a single-processor personal computer. This type of network has been used successfully in various different visual object recognition and object detection applications [121, 123]. In this chapter, we introduce an application that finds an in-focus depth of an object encoded in a digital hologram by using off-the shelf architectures.

6.2 Autofocusing

A digital hologram is an efficient encoding of a diffraction volume. It would be desirable for digital holography researchers to be able to edit the hologram directly, in order to produce some semantic change in the diffraction volume (such as 3D segmentation), or even more simply, analyse the hologram directly so we may construct an understanding of the 3D scene. Unfortunately, in the general case, this has eluded researchers in digital holography.

Researchers are limited to sampling the reconstruction volume (i.e. using numerical propagation to reconstruct from the digital hologram at a plurality of depths) before they can understand the encoded 3D scene. A handful of notable exceptions exist, such as the landmark papers by Vikram and Billet [257] and Onural and Özgen [258], and subsequently others over the past decade [259–264, 21] whose work allows one to determine the size and position of individual particles based on an analysis of the hologram directly. However, these approaches are limited to the special case of idealised spherical particles. Here, we consider significantly more complicated multi-cellular partially-transparent objects.

One open problem in DHM is how to efficiently determine the appropriate in-focus plane for an arbitrary semi-transparent microscopic object encoded in a digital hologram. If the sample contains more complicated objects than idealised spherical particles [258], researchers must perform a plurality of numerical propagation steps, which is not efficient. In this chapter, we demonstrate that it is possible to design a deep convolutional neural network to predict the in-focus distance of a living cell cluster from only the digital hologram plane amplitude. With deep learning, we propose that DHM researchers now have a tool at their disposal that is a major step towards removing the need to perform any propagation steps in order to determine the in-focus distance.

DHM overcomes the shallow depth-of-field problem in optical microscopes, by allowing one to change the in-focus plane after hologram capture. However, since each in-focus plane has a narrow depth of field, the object of interest is in focus only for a small range of reconstruction depths. The problem of determining the most appropriate in-focus depth is essential for applications such as auto-focusing, extended focus imaging, and 3D object segmentation/recognition.

The critical importance of this problem to digital holography researchers is evidenced by the regularity of newly proposed focus metrics to apply to Eqs. 2.7 and 2.8 [265–268, 76, 269, 251, 270–272, 252], among others. We highlight some of them as follows.

Gillespie and King [265] were the first to automate in-focus plane detection of digital holograms. They applied entropy as a focus metric. Ferraro et al. [266] detected phase changes and used this information as a focus metric. Liebling and Unser [267] used a wavelet approach to measure image energy as a focus metric. A reconstruction in

a stack that contains high levels of image energy in as few coefficients as possible is considered to be the in-focus plane. Dubois et al. [268] discovered that integrated amplitude reconstruction is at a minimum for pure amplitude objects and at a maximum for pure phase objects. McElhinney [76] et al. used the maximum of the grey level variance of an amplitude reconstruction stack to find in-focus objects in a scene. Langehanenberg et al. [269] compared power spectra, grey level variance, edge detection, and Laplacian filtering of pure phase objects and concluded that power spectra and edge detection are equally well suited for pure phase objects. Memmolo et al. [251] showed that Tamura coefficient estimation on pure phase objects is at the minimum in a reconstruction stack for the in-focus plane. Dohet-Eraly et al. [270] used multi-wavelength DHM and developed a focus metric that works on Fourier domain phase. They showed that their metric yielded a minimum at the in-focus plane of a reconstruction stack. He et al. [271] calculated the cosine score on an amplitude reconstruction stack, and by finding the minimum value they were able to discover the in-focus plane. An approach of Ren et al. [272] by using structure tensors and their eigenvalues was shown to work with overlapping objects. By finding maximum peak(s) in their focus metric, applied on a reconstruction stack, they discovered in-focus plane(s). Lyu et al. [252] calculated the sum of subtracted neighbouring reconstructions in a stack, and by finding the maximum of amplitude objects and the minimum for phase objects, they were able to find the in-focus plane.

Despite a great number of publications tackling this challenging problem, they each have the same drawback, namely, that a stack of reconstructed images must be computed, and the focus metric must be applied to each reconstruction. This time-consuming drawback is compounded by the fact that the whole procedure must be applied to each new hologram. There is a trade-off between computational search time and accuracy (by sampling the diffraction volume at too low a resolution longitudinally one may miss the in-focus plane and sampling the volume with a fine resolution leads to increased computation time).

The greatest benefit of the deep learning method outlined here is that after training, the in-focus depth can be obtained from the hologram plane intensity directly, and in constant time, without any numerical propagation. Ren et al. [202] reported a deep learning based

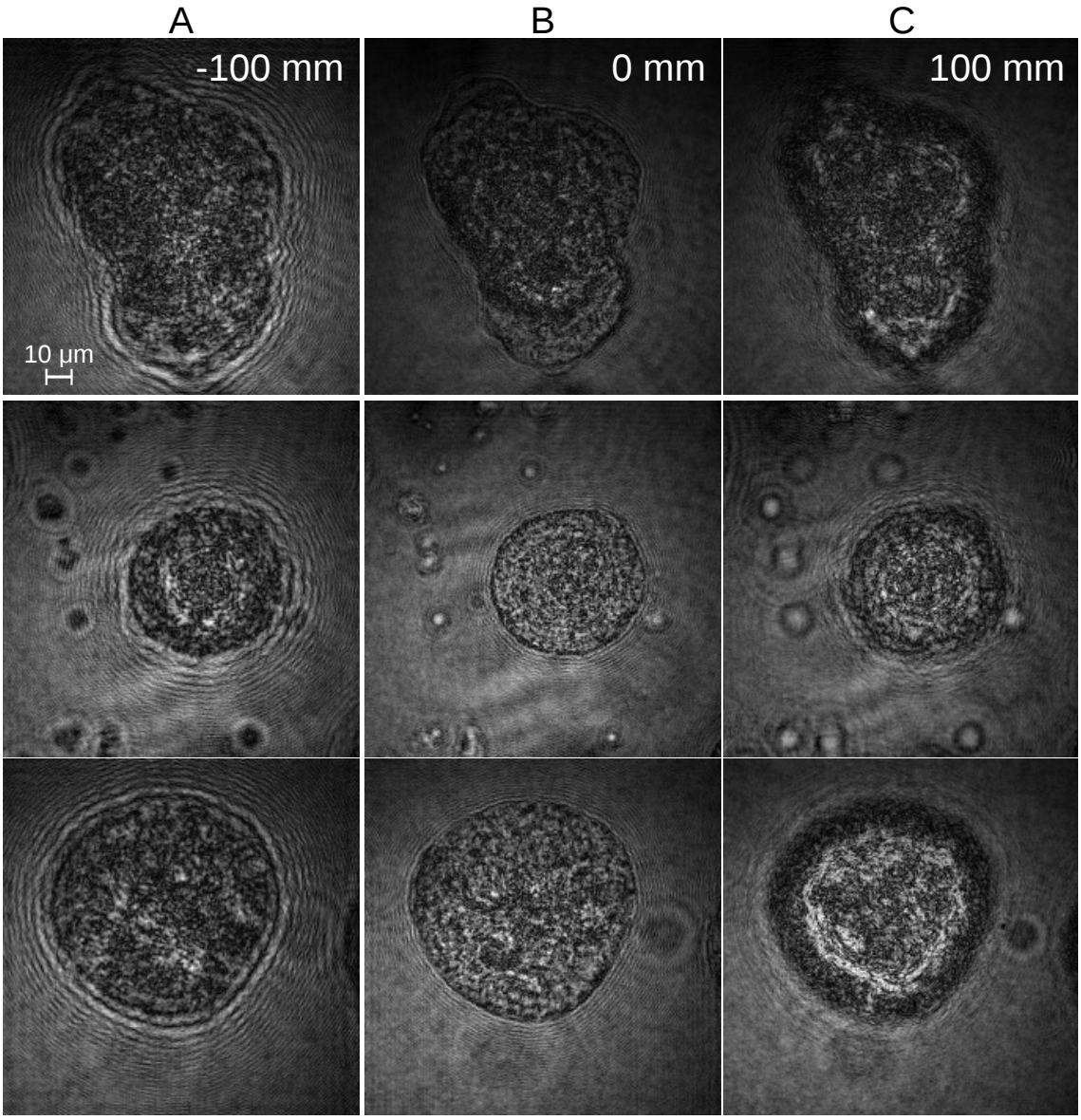

Figure 6.1 Example training images. Each row shows amplitude reconstructions from one hologram (at the in-focus plane, and at the distances ± 100 mm from the in-focus plane).

approach for both amplitude and phase objects. An interesting deep learning approach where the network learns to propagate a hologram of a transparent object to an in-focus plane and perform a phase unwrapping operation was reported by Rivenson et al. [198]. In contrast to these papers, the phase objects reported here are living biological 3D multicellular samples with all of the noise and inhomogeneities of living organisms.

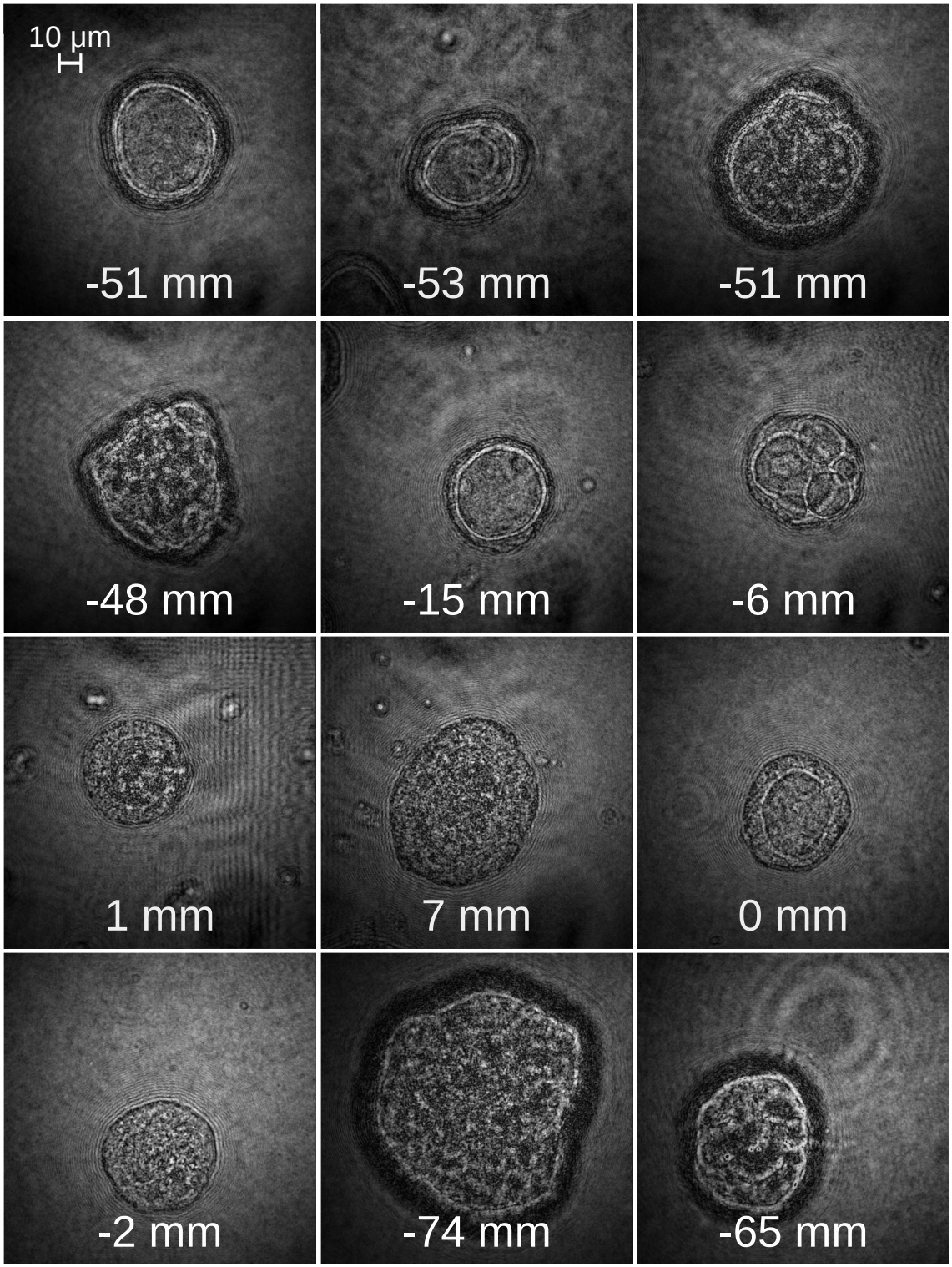

Figure 6.2 Amplitudes of the twelve dc- and twin-free holograms used for testing (with the ground truth in-focus distance shown).

6.2.1 Focusing with a regression model

Two different deep convolutional neural networks as described in Chapter 4 (see Fig. 4.3) were used. The regression model did not have the softmax layer and it returned a single real number denoting the in-focus depth.

A set of 494 holograms of semi-transparent MDCK cysts were used for this experiment (from imaging experiments 0-5 in Appendix A). Each hologram was preprocessed automatically by removing the zero order and twin terms according to [53]. To obtain the ground truth data, the in-focus depth for each hologram in the training, validation, and test datasets was determined manually. For this, holograms were reconstructed (by using Eq. 2.2) at multiple depths and the particular depth z that visually displayed the sharpest outer border was chosen as the in-focus depth.

The holograms were used to generate a database of images as follows. An amplitude reconstruction was obtained from each hologram at each of 21 depths distributed equally over the range ± 100 mm centred on the in-focus plane (see examples in Fig. 6.1). (We follow a convention [266, 269] of ignoring the effect of the microscope objective in our reconstruction distances.) Through different combinations of rescaling and cropping, each reconstruction was used to generate six similar but distinct 227×227 pixel rescaled and cropped images. Each such image was augmented with 90°, 180°, and 270° rotations, and each resulting image was further augmented through horizontal mirroring. This formed a database of 497 952 images. From this database, all augmented images from the twelve hand-picked holograms shown in Fig. 6.2 (comprising 12 096 images, 2.4 % of the set) were set aside as test data. The remaining images were partitioned randomly into training (87.8 %, 437 271) and validation (9.8 %, 48 585) data. Finally, a mean image (calculated from the training data only) was subtracted from each image.

6.2.2 Focusing with a classification model

Although finding the in-focus depth can be considered clearly as a regression task, treating this application as a classification task provides a few interesting properties that can be useful. When used for classification, a CNN returns confidence values (probabilities) for an input to be an example of a certain class. This type of network "forces" its input to be

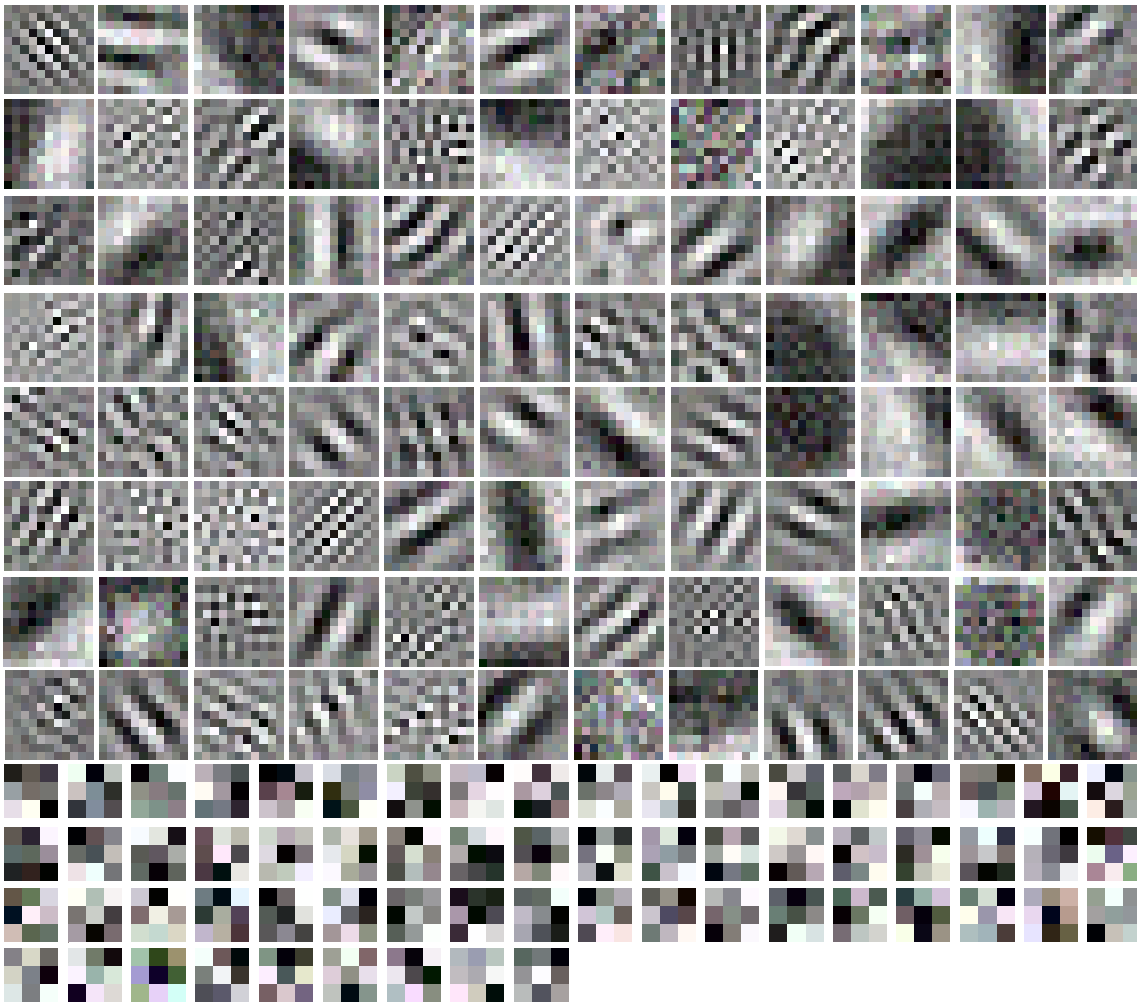

Figure 6.3 The learned filters from the first convolution layer: 96 11×11 pixel filters from AlexNet (top), 63 3×3 pixel filters from VGG16 (bottom).

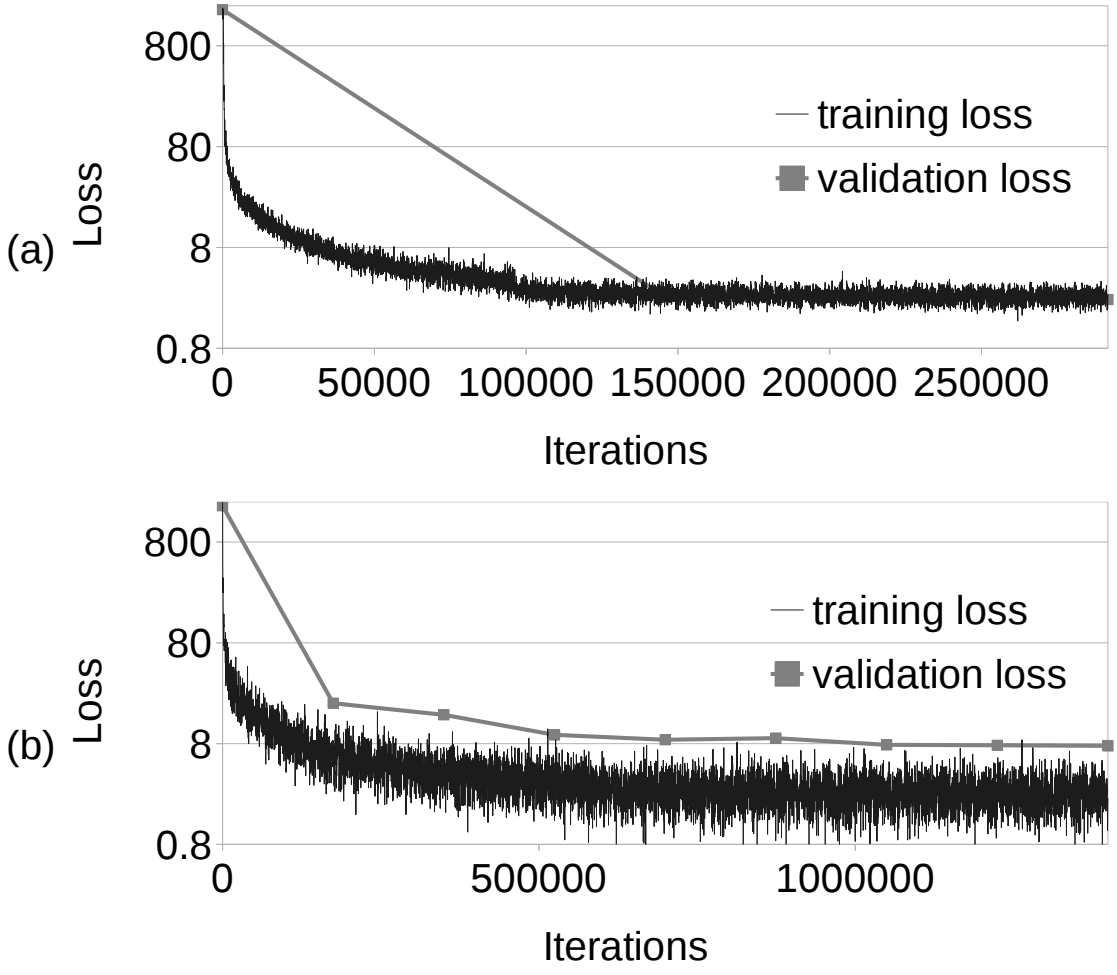

Figure 6.4 Loss functions (square mean error, logarithmic scale) for (a) AlexNet, (b) VGG16. Possibly VGG16 would have benefited from a greater minibatch size (as evidenced by the fluctuation in the loss function).

one of the pre-determined classes; the input cannot belong to any other class than the ones that have been decided at the beginning before any training takes place. In addition, there is no need for extrapolation, the allowed range is determined already. It was decided that the networks should learn 21 classes (labelled '-10', ..., '10'), representing reconstruction distances (in mm) in a range $\pm\rho$ from the in-focus plane, with quantization $\rho/10$.

6.3 Experimental results

6.3.1 Regression

The training used Nvidia's Deep Learning GPU Training System (DIGITS) software with two Nvidia GTX 1080 graphics cards. Learning rates were fixed at 1×10^{-5} and 1×10^{-6} for AlexNet and VGG, respectively. AlexNet was trained for 100 epochs (16 hours) and VGG16 for 80 epochs (210 hours) with a stochastic gradient descent solver. The loss function was mean square error and each network returned a single value for the predicted in-focus depth. The minibatch size was set to 150 and 25 for AlexNet and VGG16, respectively. The learned filters from each of the first convolution layers are shown in the Fig. 6.3, allowing one to infer the basic features that each network learned to extract from an image for analysis in subsequent layers. When the training was finished, the training loss was 2.73 and 4.47 for AlexNet and VGG16, respectively (Fig. 6.4).

Testing was performed on a separate computer to demonstrate the portability of deep learning. The trained models (with sizes 227.4 MB and 662.9 MB for AlexNet and VGG16, respectively) were imported into the Caffe [273] deep learning framework using the general-purpose Python programming language. The run times (mean of 200 holograms) were 247 ms (AlexNet) and 680 ms (VGG16) on a PC with 16 GB DDR3 memory and i5-4300U processor without a discrete graphics card. For comparison, on the same computer, and using a hologram of the same 1024×1024 dimensions, a single Tamura coefficient calculation (mean of 200) requires 932 ms (aberration removal 380 ms, reconstruction 318 ms, phase unwrapping 231 ms, Tamura coefficient calculation 3 ms).

Testing was performed by using holograms that were not used in training or validation, as explained. Of the 12 096 test images, 100 % and 99.9 % were within an acceptable error

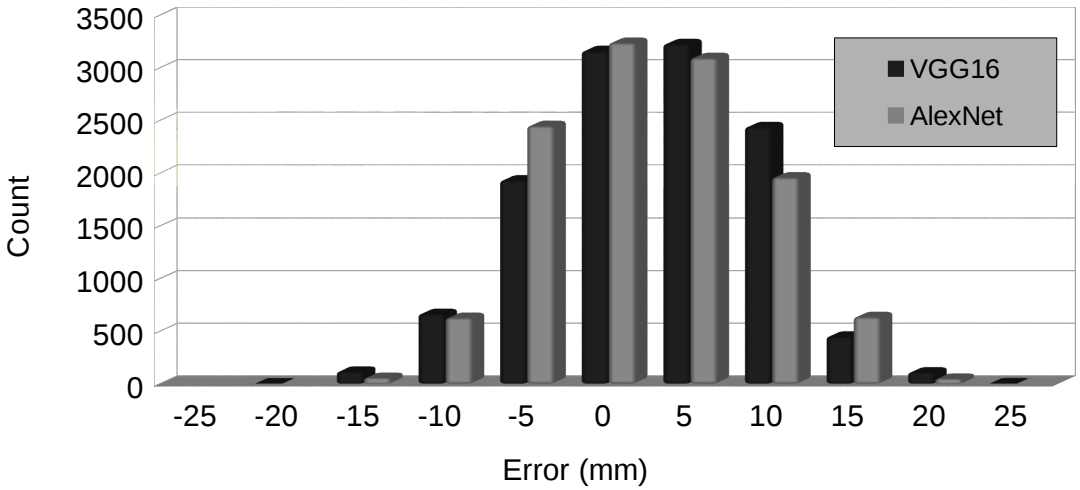

Figure 6.5 Absolute estimation errors with the testing data. In 88.7 % and 88.9 % of test cases, for AlexNet and VGG, respectively, the absolute depth estimation error is $\leq$10 mm; in 100 % and 99.9 % of test cases, respectively, the absolute error is $\leq$20 mm.

Table 6.1 Regression test results using the 12 holograms from Fig. 6.2, showing regression result from each network: G, ground truth; A, AlexNet; AE, absolute error; V, VGG16.

G (mm)	A (mm)	AE (mm)	V (mm)	AE (mm)
-51	-46.0	4.9	-52.5	1.5
-53	-52.4	0.6	-55.3	2.3
-51	-49.8	1.2	-57.7	6.7
-48	-41.1	6.9	-42.9	5.1
-15	-15.2	0.2	-15.5	0.5
-6	-5.1	0.9	-13.4	7.4
1	1.4	0.4	-4.3	5.3
7	0.0	7	5.1	1.9
0	-4.4	4.4	-4.7	4.7
-2	-1.0	1.0	2.7	4.7
-74	-83.0	9.0	-82.1	8.1
-65	-60.1	4.9	-67.8	2.8

margin of 20 mm for AlexNet and VGG16, respectively (see Figs. 6.5 and 6.6). The rms errors were 6.37 mm and 6.49 mm for AlexNet and VGG16, respectively. The error margin of 20 mm was determined experimentally; there was no significant visual difference in amplitude reconstruction quality within this range. All required morphological information about cell clusters can be extracted within this range with good accuracy. This is supported by extracting a single morphological feature, size, with test holograms (an example shown in Fig. 6.7).

To examine how the networks responded to holograms that may have an in-focus distance not a multiple of the 10 mm discretisation used in training, the holograms from

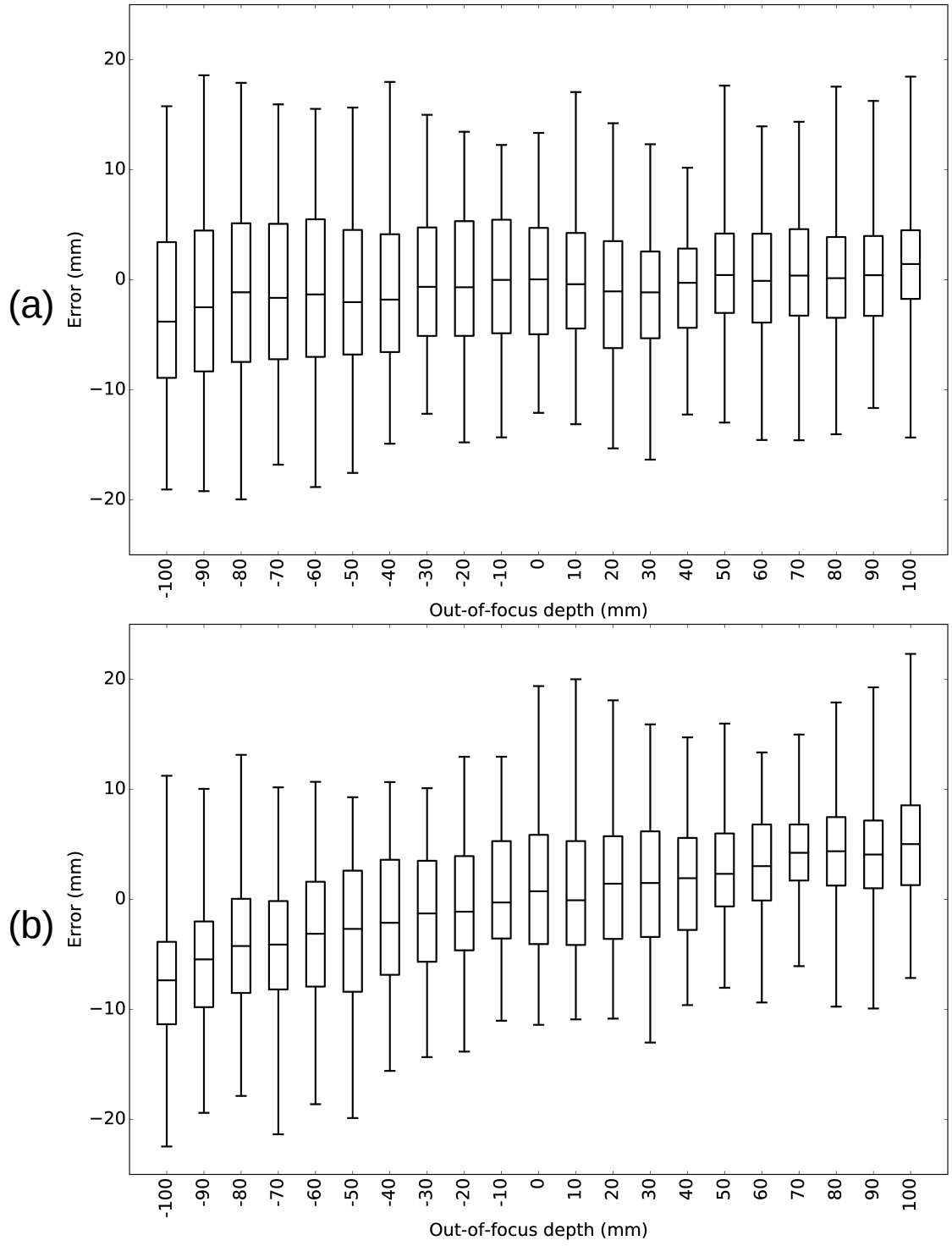

Figure 6.6 Box-plots. (a) AlexNet and (b) VGG16 showing distributions of in-focus depth prediction errors for different out-of-focus depths.

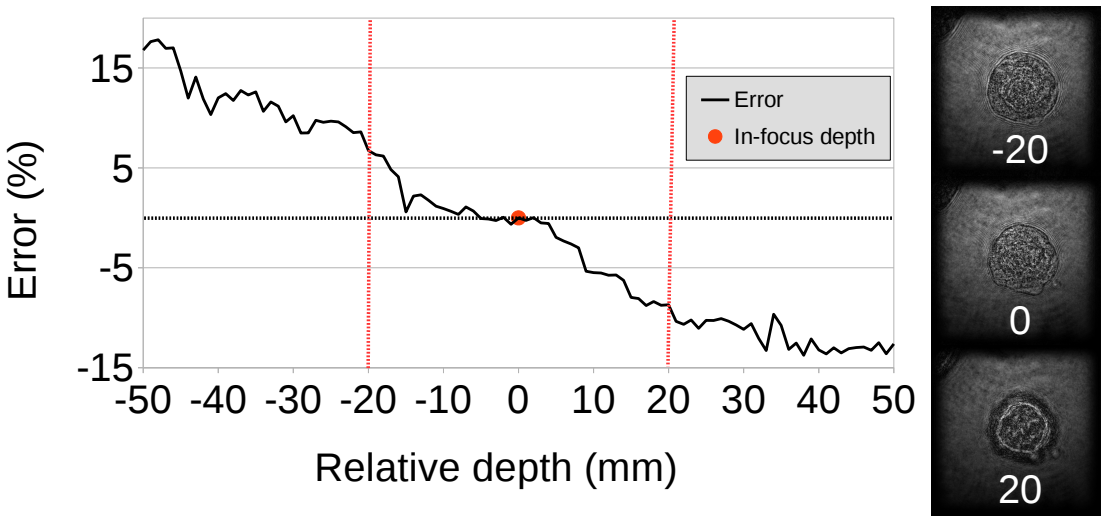

Figure 6.7 Area calculated at multiple depths on one of the test holograms. In addition, showing reconstructions at -20 mm, 0 mm, and 20 mm from the in-focus plane. The dashed horizontal line shows the 0 % error. The red dashed vertical lines show ± 20 mm range.

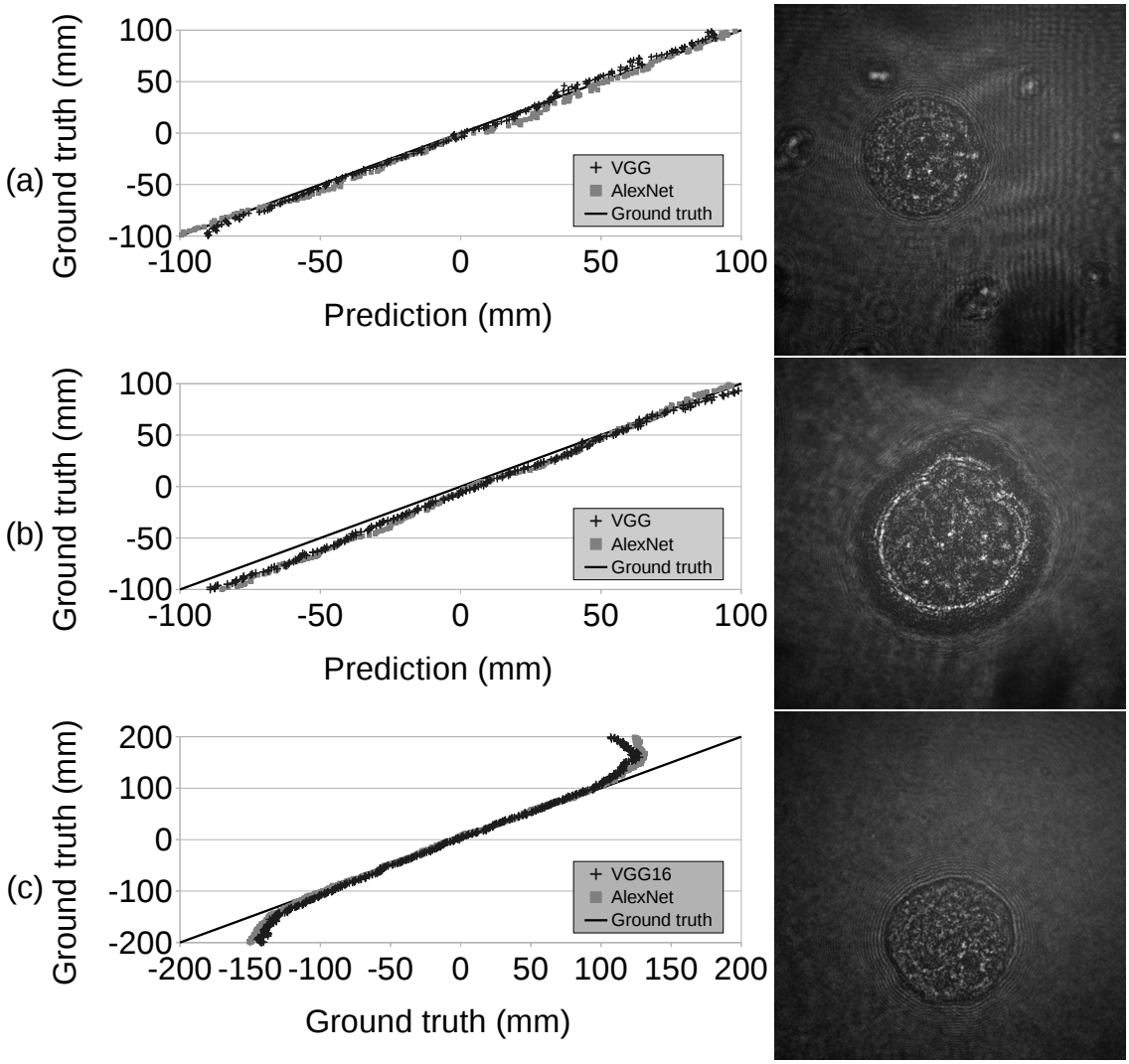

Figure 6.8 Fine (1 mm) depth-resolution test results with different size cell clusters. (a linear result indicates perfect estimation). (a) and (b) are two holograms reconstructed ± 100 mm around the in-focus plane, (c) a hologram reconstructed ± 200 mm around the in-focus plane.

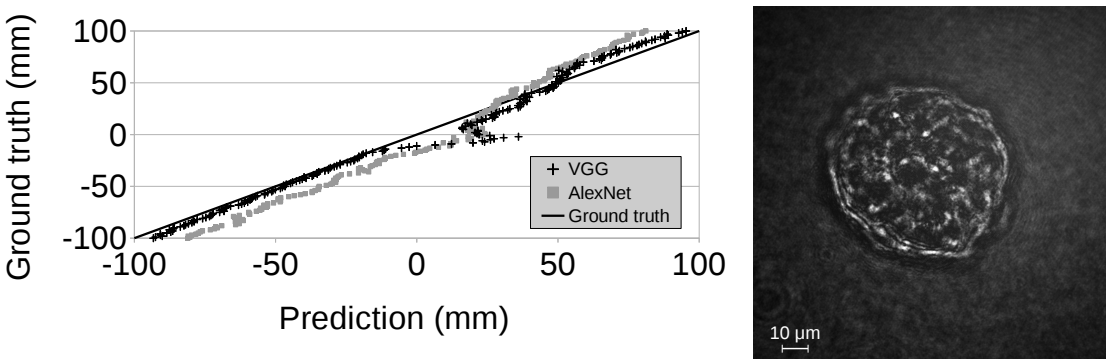

Figure 6.9 Fine (1 mm) depth-resolution test results with a human cell line sample.

Fig. 6.2 were used directly (see Table 6.1). The mean absolute error over the 12 holograms was 3.62 mm and 4.25 mm for AlexNet and VGG16, respectively.

Systematic testing was then performed with the holograms from Fig. 6.2 over the range $\pm$100 mm centred on the in-focus depth, but this time with a finer depth resolution of 1 mm. For a system to generalise well outside the discrete set of 21 in-focus depths with which it was trained, the shape of the scatter plot should be linear. Both networks generalised well with each test hologram and typical examples are shown in Figs. 6.8(a) and 6.8(b).

To test the extrapolation capability of the models beyond depths not used in the training, the test holograms were reconstructed at distances in the range $\pm$200 mm from the manually estimated in-focus plane, with steps of 1 mm (Fig. 6.8c). It can be observed that beyond depths that were not used in the training, the error increases. This is expected as neural networks, in general, are not good in extrapolation. As the volume size is known before the training, the system should be trained within the full depth of the volume so that there is no need for extrapolation during prediction.

To test the generalisation capability and push the networks past their designed capabilities, the networks were tested with a human cell line captured with the same DHM hardware. An example with 1 mm reconstruction steps is shown in Fig. 6.9. It can be observed that, in general, the CNN is able to determine in-focus depths correctly. However, the result displays greater error in the middle region. Typically, with new samples that differ greatly from a training set of a trained model, one needs to retrain a model with new data. However, transfer learning [274] can be used to speed up the training process.

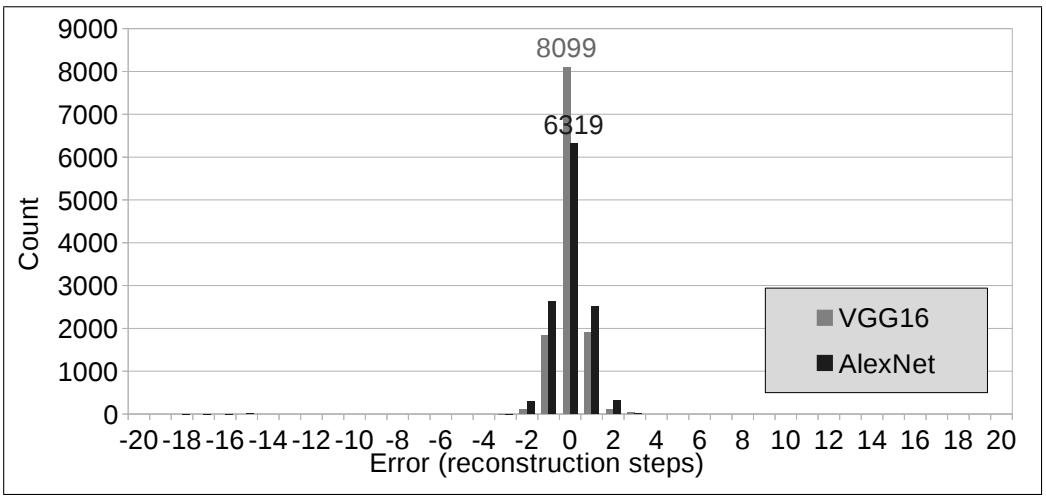

Figure 6.10 In 52.2 % and 66.9 % of test cases the correct depth class was predicted. In 94.8 % and 97.9 % of test cases a depth within one depth class of the ground truth depth was predicted.

6.3.2 Classification

For classification networks, AlexNet was trained for 60 epochs (13 hours) and VGG16 for 40 epochs (108 hours) with a stochastic gradient descent solver. The minibatch sizes were set to 100 and 40 for AlexNet and VGG16, respectively. When the training was finished, the training loss was 2.497×10^{-2} and 4.955×10^{-3} for AlexNet and VGG16, respectively.

Of the 12 096 test images, 94.8 % and 97.9 % were classified within one class of the ground truth depth for AlexNet and VGG16, respectively (see Fig. 6.10). Although the depth classes are completely unrelated as far as each network is concerned, it is a remarkable indicator of robustness that where each model incorrectly classifies an input, it invariably choses a neighbouring depth class instead.

To examine how the networks responded to holograms that may have an in-focus distance not a multiple of the 10 mm discretisation used in training, the holograms from Fig. 6.2 were used directly (see Table 6.2). Based on probability, the two top predictions for each network typically straddle the correct answer. Both networks typically classify with high confidence holograms with an in-focus distance close to a multiple of 10 mm. For holograms with in-focus depths halfway between two classes, AlexNet can partition its confidence appropriately between two classes, for example, the holograms with ground

Table 6.2 Classification test results using the 12 holograms from Fig. 6.2, showing the two top predictions from each network as ranked by their confidence values: G, ground truth; A, AlexNet; C, respective confidence values; V, VGG16.

G (mm)	A (mm,mm)	C (%,%)	V (mm,mm)	C (%,%)
-51	-40,-50	60,40	-40,-50	95,5
-53	-50,-60	99,1	-50,-60	96,4
-51	-60,-50	100,0	-60,-70	100,0
-48	-40,-30	99,1	-40,-50	100,0
-15	-10,-20	100,0	-10,-20	94,6
-6	**-10,0**	**77,23**	0,-10	99,1
1	10,0	100,0	0,10	100,0
7	0,10	99,1	10,0	100,0
0	0,-10	100,0	0,-10	100,0
-2	0, 10	99,1	0,10	100,0
-74	-90,-100	100,0	-80,-70	99,1
-65	**-70,-60**	**58,42**	-60,-50	100,0

truth values of -6 mm and -65 mm (highlighted in bold). The mean absolute error over the 12 holograms was 6.58 mm and 4.92 mm for AlexNet and VGG16, respectively.

Systematic testing was then performed with the holograms from Fig. 6.2 over the range $\pm$100 mm centred on the in-focus depth, but this time with a finer depth resolution of 1 mm. For a system to generalise well outside the discrete set of 21 in-focus depth classes with which it was trained, the shape of the scatter plot should exactly be a staircase with a linear trend. Both networks generalised well with each test hologram, and a typical example is shown in Fig. 6.11(a).

To push the networks past their designed capabilities, the networks were tested (over the same depth range and resolution) with a human cell line sample captured with the same DHM hardware. Both networks performed surprisingly well on this sample. However classification with AlexNet, while largely monotonic, is no longer linear [see Fig. 6.11(b)].

6.4 Needed amount of data

Typically, in real-world image classification or regression problems, the number of images is hundreds or thousands, and the magnitude of classes is in the same range. As we restricted our depth finding problem to a specific range, we were interested if it is possible to reduce the amount of data, while retaining the same performance (clasffication accuracy). To investigate this, we conducted an experiment where the number of images for each

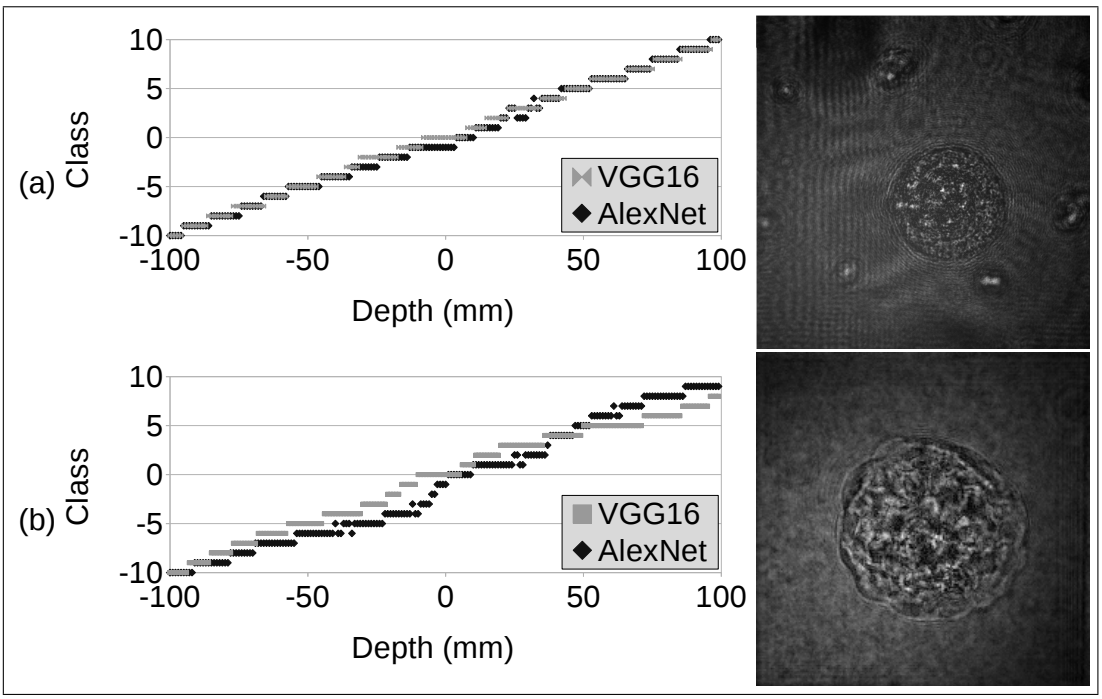

Figure 6.11 Fine (1 mm) depth-resolution test results with classification models (a linear staircase indicates perfect classification): (a) same DHM hardware and same sample type as in training, (b) same DHM hardware but different cell line.

depth was reduced systematically. This experiment is in alignment with out simplification theme. We did not augment all the images as before, but instead each input was augmented during run-time by a randomly chosen combination of rotation of 90 degrees, horizontal mirroring, and vertical mirroring. Data was reduced by steps of 10 % from 100 to 10 and then by steps of 1 % from 10 to 5. For each training, the starting point was the previous dataset and a proportion of data according to the reduction plan was used for training. For each reduction step, the training was repeated three times. The used network was based on VGG16 architecture. All networks were trained with two Nvidia GTX 1080 graphics cards using the PyTorch [248] deep learning framework.

Figure 6.12 shows the mean absolute error with different numbers of input images. It can be observed that reducing the amount of data by 60 % does not affect the accuracy; the performance remains the same. After a 60 % data reduction, the accuracy decreases, however, it is remarkable that even with 90% data reduction, the mean absolute error is only 5.8 mm.

Figure 6.13 shows focus prediction error using different amounts of training data. It can be observed that the data reduction flattens the data distribution, thus increasing the

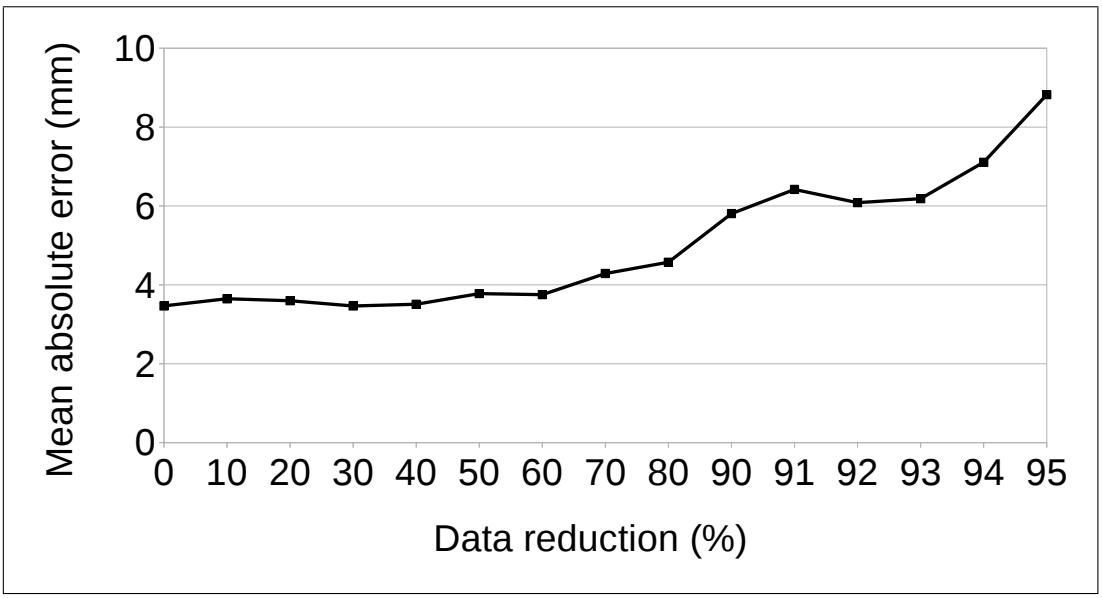

Figure 6.12 Data reduction mean absolute error. Showing mean absolute error by using VGG16 networks with different amounts of training data. The data reduction strategy used is explained in the text.

standard deviation and kurtosis. The standard deviation without data reduction was 4.0 mm, and it remains approximately the same up to 60 % data reduction (4.5 mm).

From these results it can be concluded that increasing amount of data does not always improve performance, at least if the data is as homogeneous as ours. It needs to be noted that less data is also favourable due to a faster training time.

6.5 Conclusions

It was shown that a CNN can learn the propagation distance from a hologram amplitude image to its in-focus plane. At the time of publication [275], this was one of the first applications of deep learning to digital holographic microscopy in the literature, where we show that a deep artificial neural network can be designed to learn the appropriate in-focus depth of an arbitrary MDCK cell cluster encoded in a digital hologram. Its greatest benefit is that the in-focus depth can be obtained from the hologram plane intensity only, and in constant time, without any numerical propagation. It generalises well to in-focus depths differing from those in its training set.

It was shown that the same problem can be solved using either regression or classification. Both of the approaches have pros and cons. In addition, a data reduction experiment

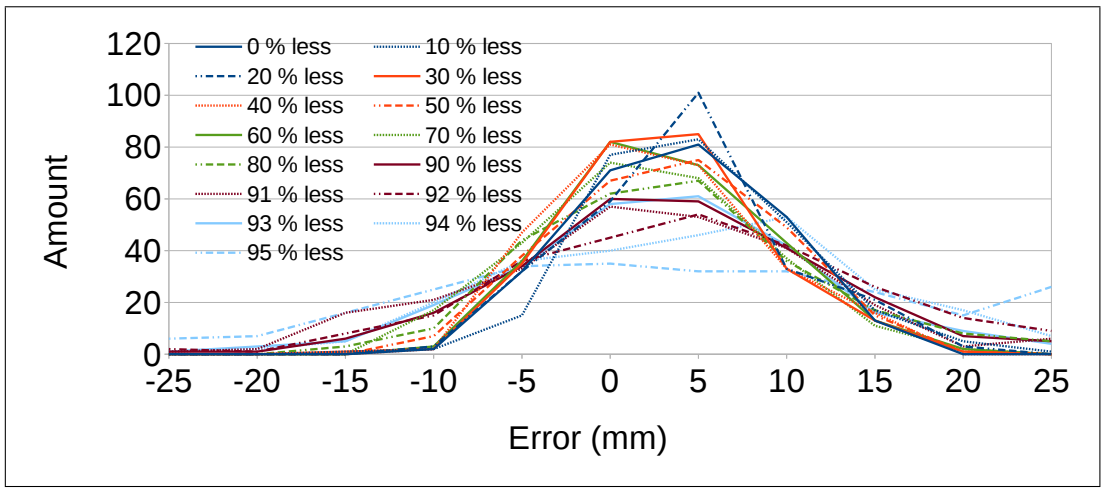

Figure 6.13 Data reduction error. Showing error range with networks trained with different amount of training data.

was reported. It was shown that data can be reduced greatly and still keep the same performance.

7 Deep learning for hologram segmentation

This chapter describes segmentation algorithm that can be used to extract quantitative morphological information from digital holograms. High accuracy segmentation allows precise quantification of lumenal space.

7.1 Introduction

Image segmentation is an image processing operation where an image can be divided into regions based on some attribute [73]. In the problem of segmentation, each pixel of an image must be given a single label, where pixels sharing one of a set of mutually-exclusive properties are given the same label. This process is useful later in the analysis stage as the obtained segmentation mask can be used to extract meaningful information such as morphology-related details (e.g. size). Conventional segmentation approaches include methods such as thresholding [276, 256] and edge detection [277]. More sophisticated methods include statistical approaches such as active contours [278], *k*-means clustering [279], and different region growth approaches [280].

Conventional neural networks have been used successfully in image segmentation [281–285]. After these early applications, quite recently CNNs have been used for the same task. Initially, segmentation using CNNs was performed on patches of an image [128]. The main reason for using image patches was the existence of fully-connected layers that required a fixed input size. Fully convolutional networks (FCNs) were introduced by Shelhamer et al. [129]. This approach replaced fully connected layers with convolutional layers, allowing dynamic input sizes. In addition, processing the whole image in one

iteration was possible — there was no need to divide the image into patches. In the network, the spatial dimensions of the data are usually reduced using pooling layers, which allow a greater receptive field (larger field of view) in the downstream layers of the network. This has a drawback of losing some spatial information about a feature, which can be undesirable in segmentation tasks. To overcome this problem, the encoder-decoder architecture was invented [130]. This architecture contains skip connections from earlier layers to later layers allowing the network to retain spatial information of features. One of the most common encoder-decoder architectures is the family of implementations based on U-Net [131]. Another approach that does not use pooling layers is based on dilated (atrous) convolutions [132–134]. The most recent development is a combination of encoder-decoder and dilated convolutions [135].

7.2 Segmentation

In the field of digital holographic microscopy, Yi et al. [286] were the first to use FCNs for segmentation. They segmented phase reconstruction images of red blood cells. Nguyen et al. [60] used FCNs for the segmentation of objects as part of their aberration removal. To the best of our knowledge, these are the only two papers that report the use of FCN for segmentation within the field of digital holographic microscopy. The two main differences that we have that those authors did not have are

1. we have complex 3D objects with extensive volumes (above 1000 picolitres), and
2. we are interested in segmenting with heterogeneous regions (multi-label segmentation).

As the lumen may contain multiple hollow 3D regions, if one can identify the basolateral and apical membranes accurately it is possible to make accurate measurements of the lumen. For this purpose, we have trained a CNN to segment objects into three regions:

0. background,
1. cells between apical and basolateral membrane (or non-hollow 3D region), and
2. hollow lumen

Pixels in our segmentation mask $S(x,y)$ therefore have one of three distinct values, one for each region, from $\{0,1,2\}$.

7.3 Combining two worlds: ground-truth masks

The performance of a CNN is highly dependent on the quality of training data: the better the data, the better the result. Because of this, we drew the ground-truth masks semi-manually. The extracted quantitative phase of the complex-valued reconstruction is in the range of $[-\pi,\pi]$ due to the arctangent function used in the phase extraction (Fig. 7.1b). Typically, the next step before any further processing and analysis would be to apply an unwrapping procedure on the wrapped phase image. Then, the unwrapped phase would be segmented using an algorithm such as [287, 78]. We are proposing an alternative way to use the wrapped phase in the process of making the segmentation mask as the extracted wrapped phase may have remaining aberrations (such as a tilt), and the phase unwrapping algorithms themselves may introduce errors in the data. In addition, the phase difference of the object and the surrounding medium can be too small in the unwrapped phase image for accurate differentiation between the object and the background.

Our initial segmentation mask is obtained through a morphological gradient operation as

$$G(x,y;z) = \phi(x,y;z) \oplus b(x,y) - \phi(x,y;z) \ominus b(x,y) \tag{7.1}$$

where $\phi(x,y;z)$ is the wrapped phase image, x,y define the reconstruction plane and z is the reconstruction distance, $\oplus$ and $\ominus$ are greyscale dilation and erosion operations, respectively, and b is a flat structuring element. The morphological gradient (Fig. 7.1c) is calculated for each pixel of ϕ by using an over-blocking window, where the size of the window is determined by the size and shape of b. $G(x,y;z)$ is thresholded to produce a binary mask $M(x,y;z)$ with a manually chosen fixed threshold τ as

$$M(x,y;z) = \begin{cases} 1, & \text{if } G(x,y;z) \geq \tau, \\ 0, & \text{otherwise} \end{cases} \tag{7.2}$$

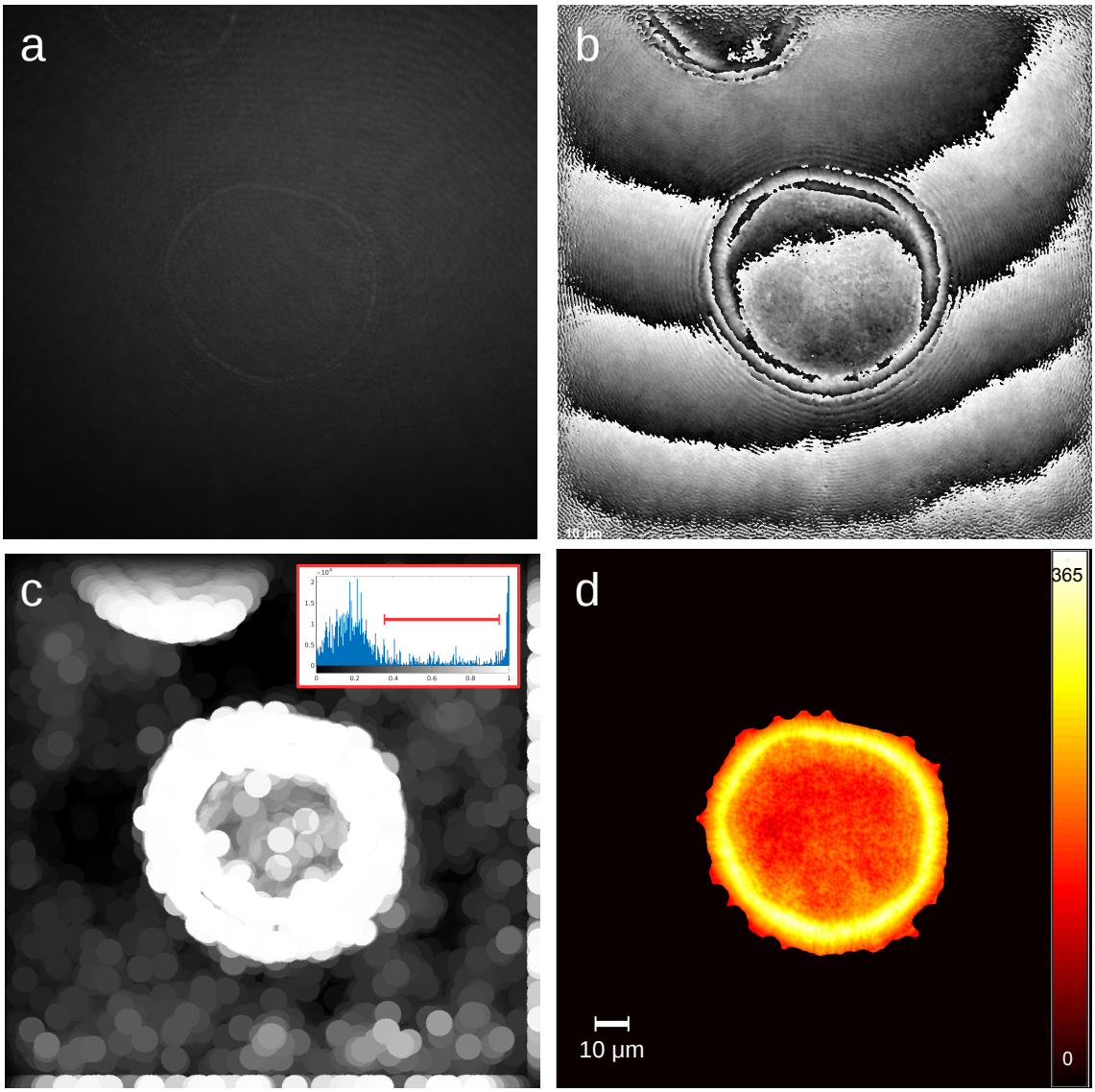

Figure 7.1 Segmentation. (a) digital hologram of MDCK cyst, (b) wrapped phase reconstruction at 3.5 cm from the hologram plane, (c) morphological gradient map that is obtained by using a disk-shaped structuring element with a 25 pixel radius, inset shows a histogram of the morphological gradient map illustrating how effortlessly the thresholding can be realized, (d) segmented and unwrapped quantitative phase map.

τ is 0.6 in all of our experiments. Three successive binary erosion operations are applied on the mask; first, with 15-pixel radius disk-shaped structuring element, second with five-pixel diamond-shaped structuring element, and third with two-pixel diamond shape structuring element. These operations are required because the resulting mask from the morphological gradient operation is larger than the object. A morphological binary open operation (followed by a binary fill operation) is applied to M. Regions smaller than a predefined threshold (20 000 pixels in our experiments) and regions connected to the phase image borders are removed. Unconnected regions in $M(x,y;z)$ are labelled using one label function of the measure module of scikit-image [246] library, to produce a label mask $M'(x,y;z)$, permitting each cyst in a scene to be segmented and unwrapped independently. Labelling is performed by using 2-connectivity. A bounding box is determined for each labelled region $R(x,y;z)$ in $M'(x,y;z)$, and the corresponding region of the $R(x,y;z)$ is cropped and binarised as

$$R'(x,y;z) = \begin{cases} 1, & \text{if } R(x,y;z) > 0 \\ 0, & \text{otherwise.} \end{cases} \tag{7.3}$$

Phase unwrapping was performed using the method of Ghiglia and Romero [62]. An example of segmented unwrapped phase reconstruction is shown in Fig. 7.1d. These initial binary masks were manually corrected using GNU Image Manipulation Program (GIMP) [288] using the unwrapped phase as a background layer. GIMP is a free and open source software application for such tasks as photo retouching, image composition and image authoring. In addition, a third region was manually defined for the lumen. With this approach, each ground-truth mask contained two to three distinct values depending on whether there was any lumen present. Example masks are shown in Fig 7.2. Obtaining initial segmentation masks as described using a model-based approach for use in deep learning is in alignment with our paradigms theme.

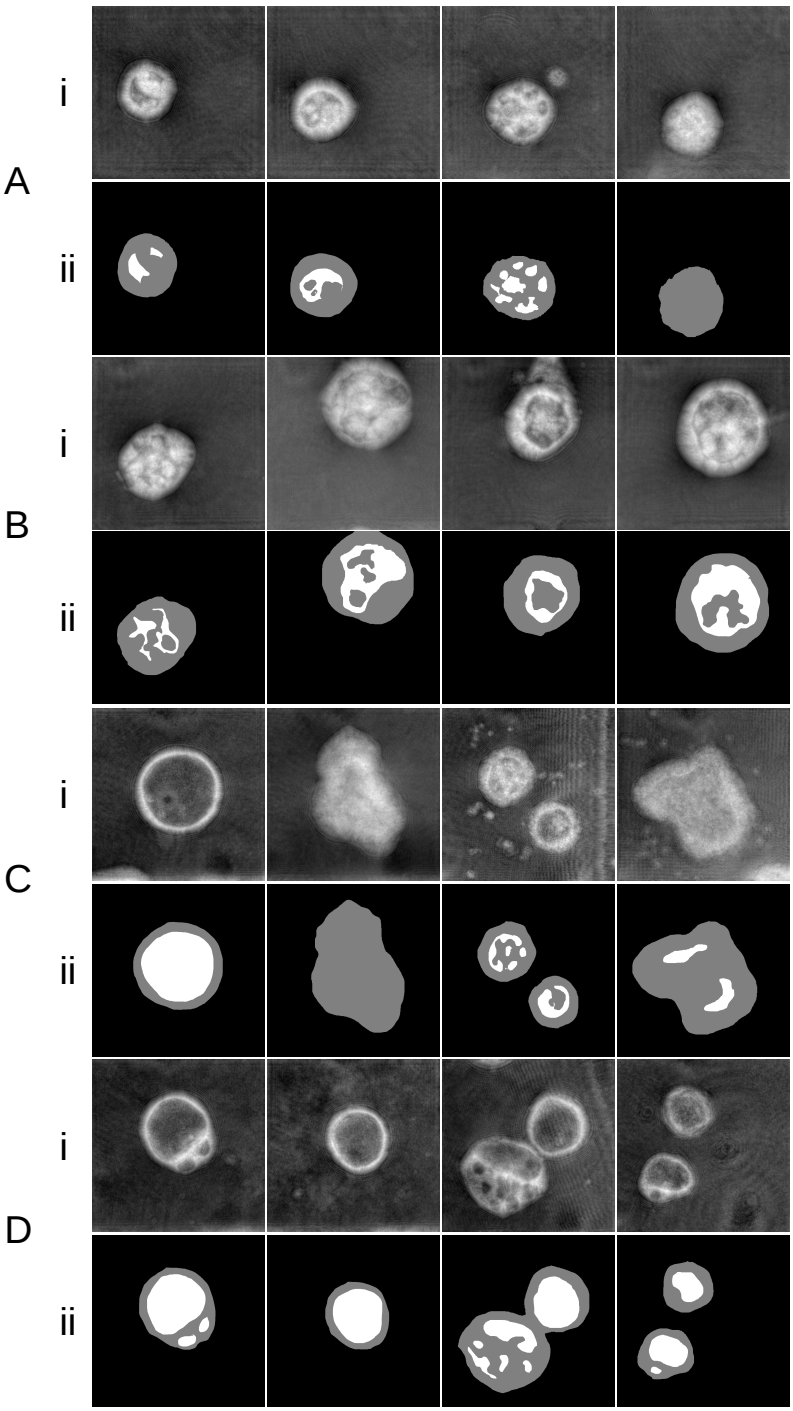

Figure 7.2 Example MDCK phase reconstructions with ground truth masks. Example phase reconstructions (i) from imaging experiments along with the corresponding ground-truth masks (ii). A and D are of healthy samples, and B and C show $KRas^{V12}$ samples. In the masks black is for background, grey is for basolateral membrane (or non-hollow 3D region), and white is for hollow lumen.

7.4 Experimental results

The ground-truth mask making process was repeated for 200 cyst holograms that were captured at eight distinct times (Appendix A, experiments 1-5, 24 and 30). From two of these imaging experiments (experiment 24 and 30), 50 holograms of healthy cysts and 50 holograms of $KRas^{V12}$ cysts were randomly picked. The remaining 100 were randomly picked from the remaining experiments. The reason for picking data as described was the different number of holograms captured in each imaging experiment. In total, half of the holograms (100) contained $KRas^{V12}$ transformed samples, and the remaining contained healthy samples (100). The network architecture was based on U-Net [131] with a full 1024×1024 pixel resolution input (Fig. 7.3). This architecture has been used successfully before with DHM [60, 199, 289] which is the reason for choosing this architecture.

The network was trained with three different input types: a hologram, an out-of-focus phase, and an in-focus phase; each subsequent type requiring more pre-processing for the network than the previous type. If the network was able to distinguish cyst and lumen using a real-valued hologram, it would allow for the smallest amount of pre-processing and therefore a faster throughput. Convolution kernels were 3×3 pixels with one-pixel stride, except for the last convolution that had a 1×1 pixel kernel. All max pooling layers had 2×2 pixel kernels with two-pixel stride making the output side lengths half of the input side lengths. Transposed convolutions had 2×2 pixel kernels with a stride of two making the output side lengths two times larger than the input side lengths. The network was trained with two Nvidia GTX 1080 graphics cards using the PyTorch [248] deep learning framework with a batch size of three. The network was trained for 100 epochs with the Adam [254] optimiser, weight decay of 1×10^{-4}, and two-dimensional cross entropy loss. The initial learning rate was set to 1×10^{-4}, and reduced by a factor of 10 every 25 epochs.

Figure 7.4 shows loss values during training with different inputs. With hologram and in-focus phase inputs, the loss is smaller than with out-of-focus phase inputs, however, in general, the loss values are small with all three different input types. Each network was trained once.

Figure 7.5 shows results for multiple different inputs, both healthy and tumourigenic. From these results, it is clear that the in-focus phase provides the best accuracy. These

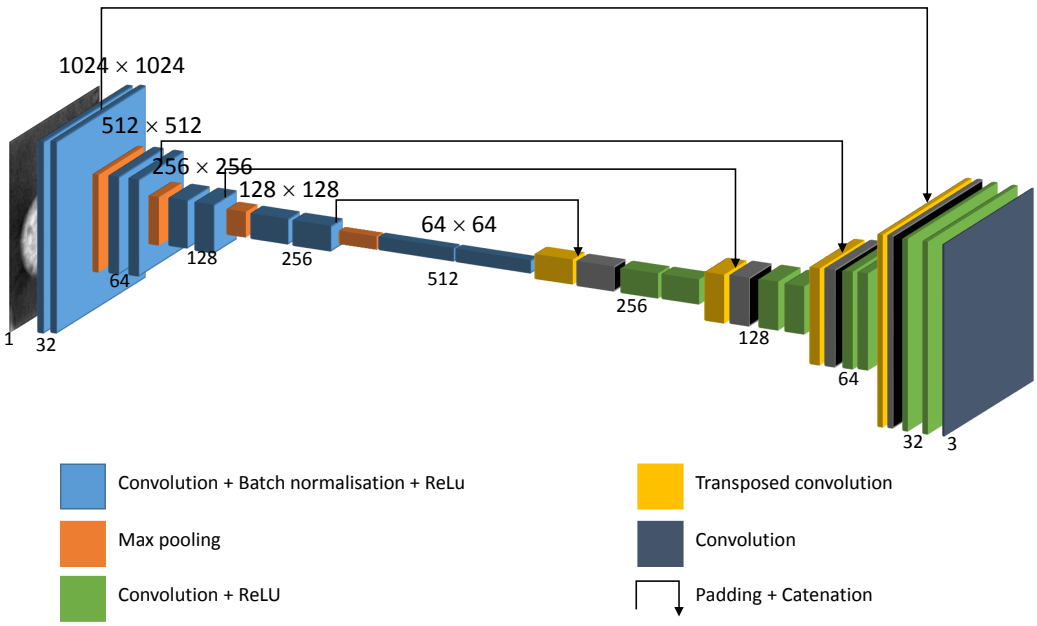

Figure 7.3 U-Net network structure, showing data flow through the network. Numbers above show iamge dimensions and numbers below show depths of layers. The input to the network was a greyscale image with 1024×1024 pixels. The output was 3×1024×1024 – containing different label probabilities at each of three channels.

3D objects are so complex that the lumen boundary cannot be determined directly from hologram nor from the out-of-focus phase images for our particular network architecture and training set.

To quantify segmentation accuracy for analysis, lumen indices r were calculated using different CNNs (r is defined in Section 4.2.5). Table 7.1 shows r values for test holograms shown in Fig. 4.4. Figure 7.6 shows absolute r value error with different network inputs. The average absolute errors were 0.158, 0.170, and 0.040 for hologram, out-of-focus phase, and in-focus phase input CNNs, respectively. This demonstrates that the in-focus phase images provide the smallest error. The segmentation CNN with in-focus phase as input is more robust than others, successfully segmenting all 11 samples, whereas the network with out-of-focus phase input successfully segmented only seven.

7.5 Conclusions

In this chapter it was shown that a CNN can perform an efficient segmentation of 3D cell clusters, and their lumens, encoded in digital holograms. A high accuracy segmentation is an important initial step in the processing and analysis pipeline for many biomedical

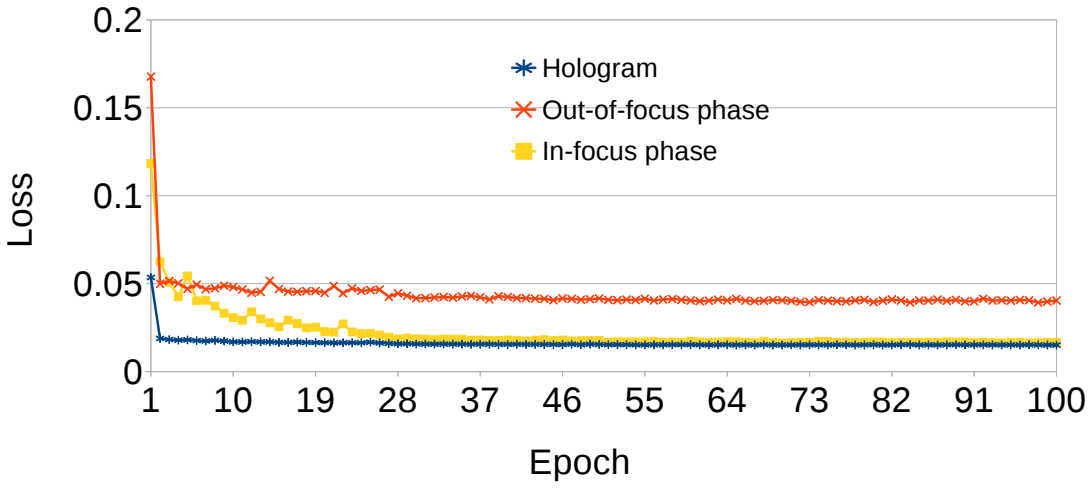

Figure 7.4 U-Net training loss values with different training inputs.

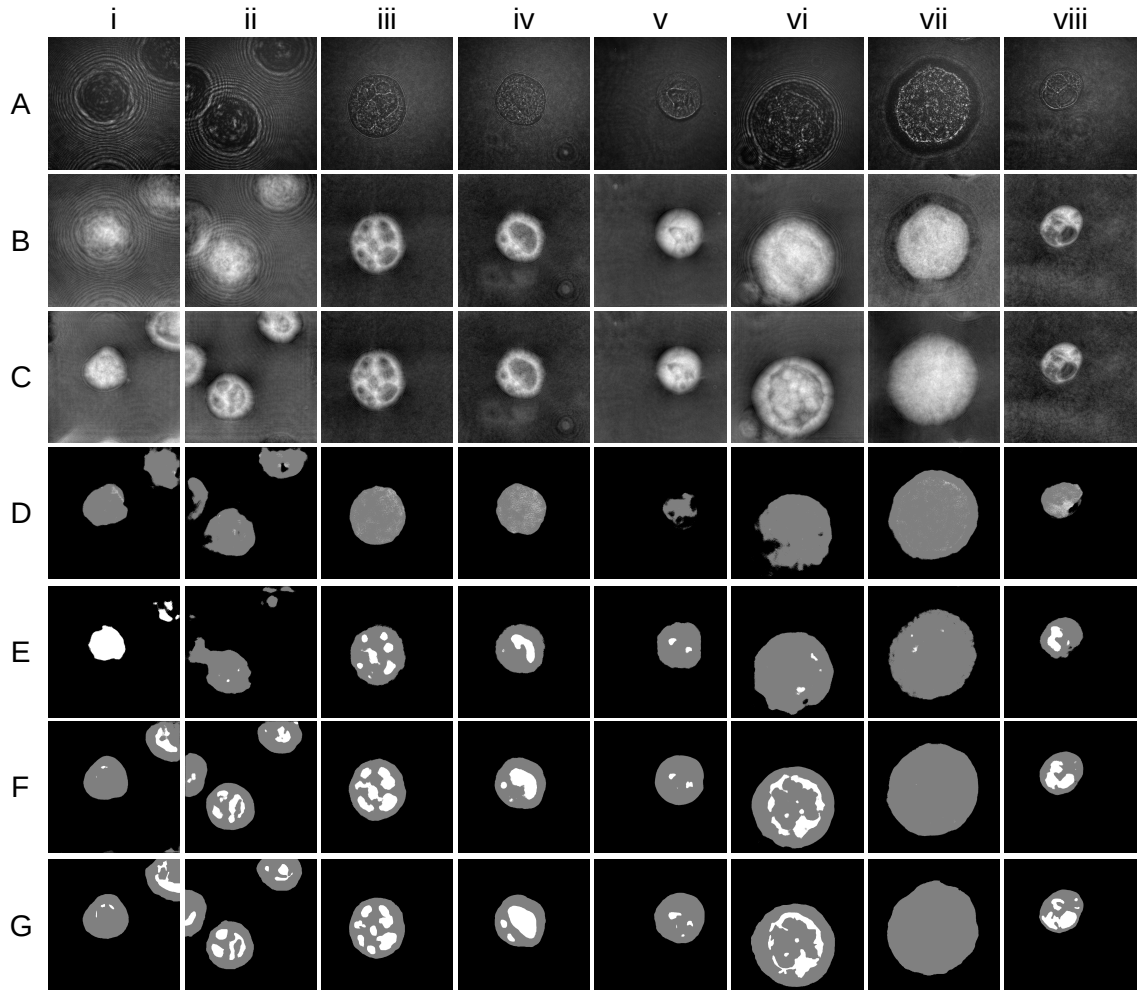

Figure 7.5 Segmentation results. A, hologram; B, out-of-focus phase; C, in-focus phase; D, mask with hologram input; E, mask with out-of-phase input; F, mask with in-focus input; G, ground truth mask. Columns i through iv are healthy cysts, and columns v through viii are oncogenic inputs.

Table 7.1 Lumen analysis. Showing *r* values of ground truth masks and masks obtained from different CNNs for example cysts shown in Fig 7.2. If a scene contained multiple cysts, cysts are indexed (e.g. H1.1), and indexes are increased from top to down; a cyst located at the top part of the input gets the lowest index. GT: ground-truth, H: hologram, OP: out-of-focus phase, IP: in-focus phase. NA in results means that this cyst with the input type could not be segmented.

Cyst	GT	H	OP	IP
H1.1	0.370	0.015	NA	0.361
H1.2	0.029	0.065	NA	0.014
H2.1	0.208	0.088	NA	0.235
H2.2	0.219	0.00	NA	0.104
H2.3	0.358	0.021	0.011	0.334
H3	0.396	0.196	0.240	0.457
H4	0.476	0.492	0.237	0.382
H5	0.070	0.000	0.059	0.048
H6	0.228	0.001	0.021	0.246
H7	0.000	0.050	0.007	0.000
H8	0.544	0.659	0.324	0.489

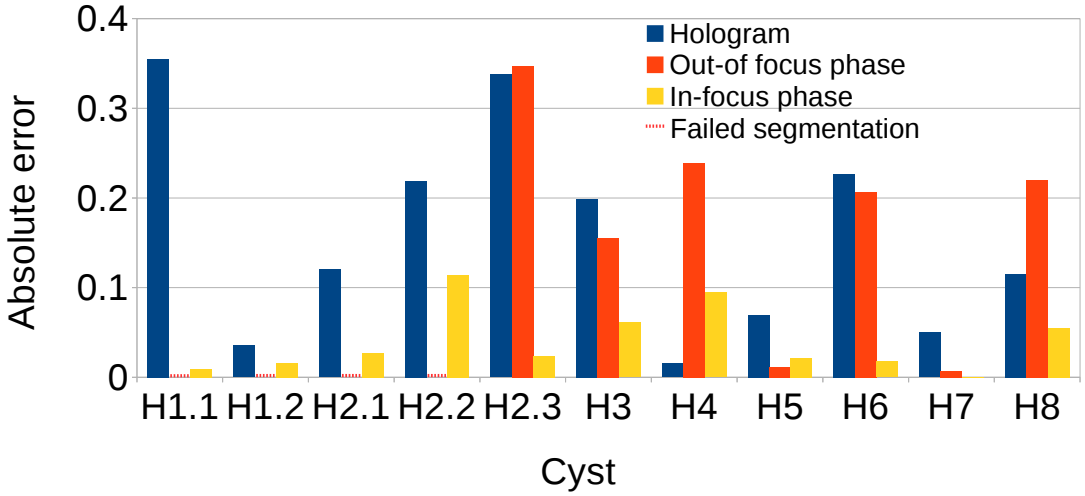

Figure 7.6 Lumen index *r* value absolute error, showing *r* value absolute errors of eight cyst holograms (Fig 7.2), containing 11 cysts in total.

applications. The best performance was provided using in-focus phase reconstructions. For the U-Net architecture and our training set, the lumens in these 3D cell clusters are too complex to be analysed from the digital holograms or from the out-of-focus phase reconstructions directly.

8 Shallow CNN architectures

This chapter contains a deep learning based approach to classify cysts as either healthy or oncogenic. An alternative to existing deep CNN architectures for two-class classification is described and results with various different shallow architectures is reported. In addition, this chapter investigates if a-priori knowledge of samples can be integrated into a neural network for our particular application domain.

8.1 Introduction

Our problem is binary classification; we want to classify our cysts as either normal or oncogenic. There does not exist a state-of-the-art CNN architecture specific to binary problems as all widely used and well-known architectures are invented for (usually harder) multi-class problems. Many of them are introduced as part of different challenge competitions, perhaps the most well-known is the ImageNet Large Scale Visual Recognition Challenge [120]. Although these architectures and model designs work well with binary problems, executing a single inference with these models can be time-consuming due their deep architectures. To tackle this problem, we investigated the possibility to simplify the network architecture for binary-classification. This approach is in alignment with our simplification theme.

In addition, we compare the performance of shallow networks with and without additional extracted quantitative hand-crafted features. Traditionally, a CNN contains a feature extraction part that, during training, automatically learns to extract distinguishing features from the data that can be used with a fully connected part of the network to

perform its task (classification or regression). It is also possible to combine CNN features with hand-crafted features that originate from elsewhere. In several published papers hand-crafted features are combined with CNN features and another classifier is trained with a composition of these features [290–293]. Jin et al. [290] used a combination of the locality-constrained linear coding (LLC) method and CNN features as an input to a SVM classifier. Li et al. [291] trained a support vector machine by concatenating the output of a first or second fully connected layer of a fine-tuned CNN with handcrafted features (root-SIFT, vectorised raw-pixel values and multi-resolution local patterns). Lee et al. [292] used random forest classifier with a concatenation of CNN features and hand-crafted features of shape and texture. Nguyen et al. [293] concatenated CNN features and a multi-level local binary patterns as an input to a SVM classifier. Other approaches either feed in hand-crafted features to a network [294] or use hand-crafted features as an additional input to a network [295]. Hosseini et al. [294] used Gabor filter responses as inputs to a CNN. Wu et al. [295] combined colour and texture histograms with features extracted by feature extractor part of a CNN based on the AlexNet architecture.

8.2 Shallow networks

8.2.1 Building blocks

We define the concept of building blocks where each building block is a combination of multiple layers. We enumerated our architectures using a binary encoding to ensure that all small building block combinations in our scheme are included. We keep our building blocks as simple as possible where the two blocks are:

- one convolution layer followed by max pooling (denoted with the symbol ‘0’) (Fig. 8.1a)
- two convolution layers followed by max pooling (denoted with the symbol ‘1’) (Fig. 8.1b)

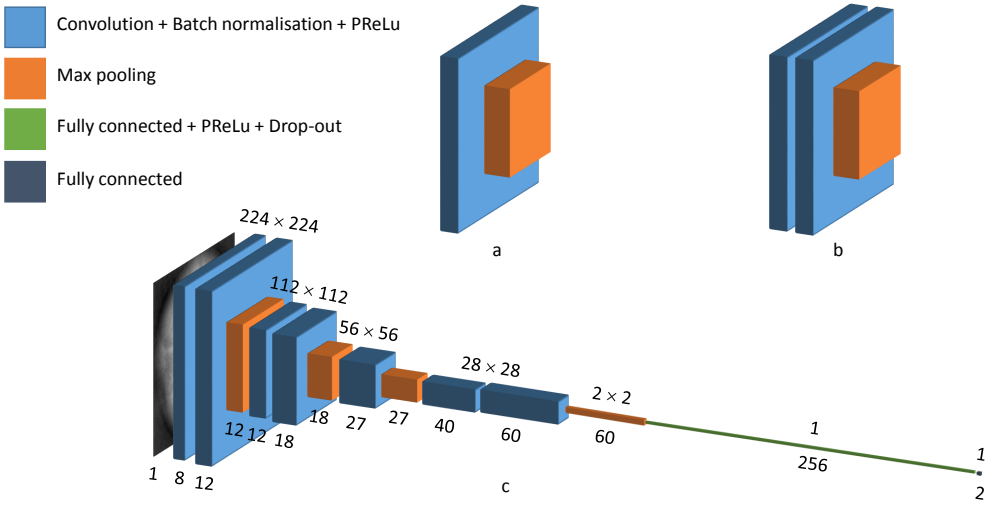

Figure 8.1 Building blocks. A: '0', B: '1', C: An example architecture with '1101' configuration. Numbers above layers show the spatial size of their output and numbers below show the number of output channels/features.

Tables 8.1 and 8.2 show tested building block combinations along with the number of learnable parameters. For reference, the AlexNet architecture contains 58.3 million learnable parameters.

8.2.2 Procedure to combine building blocks

In building block '0' there is one convolution layer followed by a maxpooling layer (see Fig. 8.1). In building block '1' there are two convolution layers followed by a maxpooling layer. In both cases, each convolution layer is followed by a batch normalisation layer that is followed by an PReLu activation layer. The first fully connected layer is followed by an activation layer and a drop-out layer with a 50 % drop-out ratio. The second (and last) fully connected layer outputs two features that are fed to a logarithmic softmax layer.

8.2.3 Network parameters

Each convolution layer has 1-pixel stride and a 3×3 pixel kernel size except the first convolution layer that has 9×9 pixel kernel size. Max pooling layers have 2×2 pixel kernel and 2-pixel stride except for the last max pooling layer. The last max pooling layer outputs a o×2×2 feature vector to the fully connected layer, where o is dependent on the depth of

Table 8.1 Shallow CNN architectures starting with a '0' building block, showing total number of learning parameters.

Index	Model	Parameters
1	0	9636
2	00	10237
3	01	15234
4	000	15234
5	001	23377
6	010	23377
7	011	37049
8	0000	23377
9	0001	37049
10	0010	37049
11	0011	60202
12	0100	37049
13	0101	60202
14	0110	60202
15	0111	102463
16	00000	37049
17	00001	60202
18	00010	60202
19	00011	102463
20	00100	60202
21	00101	102463
22	00110	102463
23	00111	182054
24	01000	60202
25	01001	102463
26	01010	102463
27	01011	182054
28	01100	102463
29	01101	182054
30	01110	102463
31	01111	337890

Table 8.2 Shallow CNN architectures starting with a '1' building block showing total number of learning parameters.

Index	Model	Parameters
32	1	14 633
33	10	15 966
34	11	24 109
35	100	24 109
36	101	37 781
37	110	37 781
38	111	60 934
39	1000	37 781
40	1001	60 934
41	1010	60 934
42	1011	103 195
43	1100	60 934
44	1101	103 195
45	1110	103 195
46	1111	182 786
47	10000	60 934
48	10001	103 195
49	10010	103 195
50	10011	182 786
51	10100	103 195
52	10101	182 786
53	10110	182 786
54	10111	338 622
55	11000	103 195
56	11001	182 786
57	11010	182 786
58	11011	338 622
59	11100	182 786
60	11101	338 622
61	11110	338 622
62	11111	653 267

the network as

$$o = \begin{cases} \lfloor a(g^{c-2}) \rfloor, & \text{if } c > 2, \\ \lfloor ag \rfloor, & \text{if } c = 2, \\ a, & \text{otherwise,} \end{cases} \tag{8.1}$$

where a is the number of output channels from the first convolution layer, g is the growth factor, and c is the total number of convolution layers. The growth factor, g, determines how rapidly the number of features (width of the network) is increased. The growth factor affects the third and subsequent blocks only. This layer outputs the same number of channels as the first convolution block. We set g to 1.5 and a to 4 in our experiments. Besides these, we keep the number of output features from the first fully connected layer, b, fixed at 256.

The networks were trained for 100 epochs with the Adam [254] optimiser, weight decay of 1×10^{-3} and cross entropy loss. An adaptive learning rate was used with an initial learning rate of 10^{-4} that was reduced by a factor of 10 if the validation loss had not decreased for ten consecutive epochs.

8.2.4 Justifying network parameters *a* and *b*

To justify these choices we trained the '1000' and '1001' architecture models with different values of a and b. The average accuracies obtained by using these parameter values were compared with AlexNet and a multilayer perceptron. The multilayer perceptron was trained with extracted features as explained in Chapter 4, and all classifiers used the same training and testing data. Three CNNs and 1000 multilayer perceptrons were trained. The averaged results of these networks are shown in Fig. 8.2. This figure shows classifier average accuracies on testing data captured on the same day as the training data, data captured at a different time than the training data, and averages of these two testing datasets weighted by the number of test samples (imaging experiments 24, 30, and 33 in Appendix A). It can be observed that all classifiers performed well on the data that was captured on the same day as the training data. However, when testing on the data that was captured at a different time (imaging experiments 0-6 in Appendix A), differences became more obvious. The multilayer perceptron failed on this dataset. With $a = \{2,4\}$

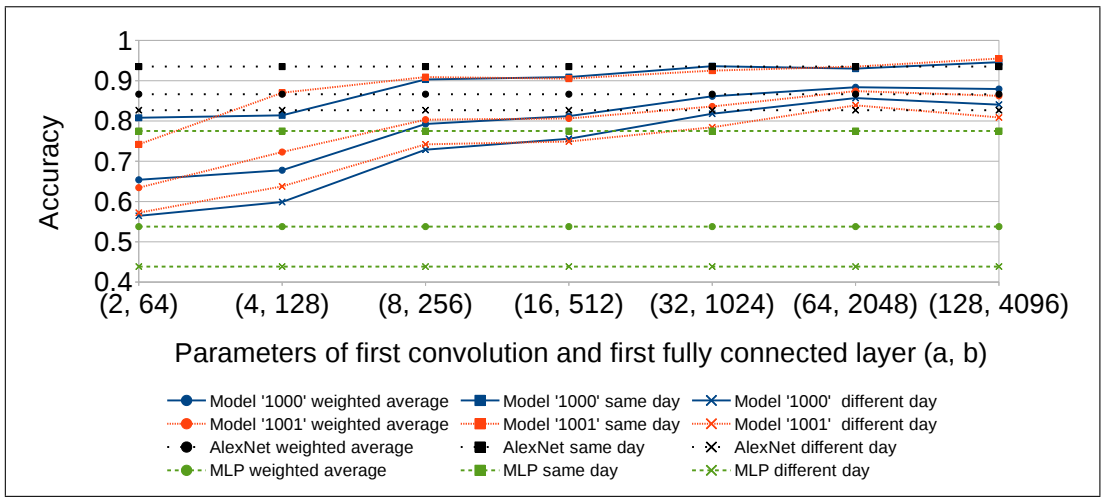

Figure 8.2 Choice of parameters. Showing classification accuracy of two shallow CNN models ('1000' and '1001'), AlexNet, and multilayer perceptron (MLP) with different testing data and their weighted averages. See text for more explanation.

and $b = \{64, 128\}$, the shallow models failed to classify this data. By looking at weighted averages, the shallow models with $a = 8$ and $b = 256$ were near to the performance of AlexNet so using these parameters was regarded as appropriate minima.

8.2.5 Claim and systematic search

The research question we are interested in is: given our architectural design restrictions, what is the minimum number of learning parameters that can compete with AlexNet for our particular problem?

Definition 8.2.1. We define an **architecture** to be a sequence of layers created using the building blocks defined in Section 8.2.1 using the procedure defined in Section 8.2.2, and parameters defined in Section 8.2.3.

Definition 8.2.2. We define an **application** to be classification of in-focus segmented phase reconstructions from digital holograms of healthy and oncogenic MDCK cysts. The cysts are cultured and imaged as described in Chapter 3. Classes are restricted to two, healthy and oncogenic, and the nature of classification is therefore binary (two-class classification).

Definition 8.2.3. We define the **reference architecture** to be AlexNet as described in Section 4.3.4. We define **performance metric** to be weighted average accuracy.

Claim 8.2.1. *For the class of architectures described in Definition 8.2.1 and for the application defined in Definition 8.2.2 no more than* 100 000 *learning parameters are required to solve the problem within 10 % of the performance of the reference architecture as defined in Definition 8.2.3.*

Systematic search. In order to validate this claim, we just need to show the existence of one instance of the architecture with sufficient performance for our binary-classification problem. We find this instance by performing a systematic search over the architectures.

The scope of this result is very limited. We hold the input domain and output domain fixed (for our specific dataset and binary classification problem), we restrict ourselves to a narrow architectural class (those CNNs that can be defined using our building blocks), and we keep various parameters fixed (such as learning rate behaviour, batch size, weight decay, dropout ratio). In this very restricted scenario, we then perform an exhaustive search with all combinations of building blocks and investigate which combinations of building blocks provides sufficient performance for our binary-classification problem.

Although this will demonstrate conclusively the existence of an architecture instance with no more than 100 000 learning parameters that solves the problem, this is just characterising an upper bound on the number of learning parameters required, and we are not claiming that no smaller network exists. Similarly, for a different architecture, more learning parameters may be required.

In order to go further and prove that this upper bound would be optimal, we would need to exhaustively search each possible test set of our particular dataset, exhaustively search all the parameters of each neural network in our collection of architectures (batch size, weight decay, dropout ratio), search all possible learning strategies (learning rate behaviour, choice of activation function), check all possible randomly-generated network weight initialisation values, and so on.

8.3 Experimental results and systematic search

A total of 2741 (2245 control and 496 $KRas^{V12}$) MDCK cysts were used in the classification experiment. The data was split into training, validation and, testing data as follows. First, from both phenotypes, 100 holograms were chosen for test data. From the remaining

holograms, 400 holograms from both phenotypes were randomly chosen for training. For validation data, 200 and 180 objects were picked from control and oncogenic phenotypes, respectively. We made sure that none of the holograms used for training or validation overlapped with test holograms. In addition, we supplemented the test dataset with an additional test dataset to make the classification problem harder (called additional test dataset in the remainder of this chapter). It contained 199 control and 146 oncogenic MDCK samples that were captured at a different time from the training data. All networks were trained with two Nvidia GTX 1080 graphics cards using the PyTorch [248] deep learning framework with a batch size of 30.

Each input was randomly augmented during run-time by a randomly generated combination of rotation by 90 degrees, horizontal mirroring, and vertical mirroring. Rescaling to 224×224 pixels enables magnification independency allowing the same network model to be used with data captured using different microscope objectives. Each unwrapped phase input was normalised to be between 0 and 1 to allow wavelength independency as the number of phase wraps is dependent on the used wavelength. Each network was trained three times.

Result of the systematic search. We demonstrate the existence of an architecture instance satisfying Claim 8.2.1 by examining the results of the exhaustive search of architectures as described in Definition 8.2.1 in increasing order of the number of learning parameters. From Fig. 8.3 it can be seen that model '00010' has accuracy (78.5 %) within 10 % of reference architecture (86.7 %) as defined in Definition 8.2.3. Model '00010' has 60 202 learning parameters. This demonstrates that no more than 100 000 parameters are needed for the application defined in Definition 8.2.2.

8.4 Combining two worlds: additional hand-crafted features

In the literature, defined at the beginning of this section, where hand-crafted features were used some way or another with CNNs, the magnitude of extracted features were typically the same as CNN extracted features (in the thousands). As we have a-priori morphological

and phase related quantitative features for cysts (defined in Chapter 4), we wanted to investigate if this information could be used to improve the classification accuracy. In total, we had 19 extracted features which is a small number compared to the typical number of extracted convolutional features. Our aim was to develop an end-to-end-trainable network where we integrate our additional data into the network, keeping the feature extraction part of our CNN intact. To accomplish this, we modified the last fully connected layer to output 19 values that were concatenated with our 19 a-priori features. The magnitudes were made equal so that the hand-crafted features do not vanish during the training and have the same significance as the CNN extracted features. We introduced an additional fully connected layer that took these 38 values as input and returned as output two values that were fed to a logarithmic softmax layer. The network was able to learn weights also for a-priori features. This procedure is illustrated in Fig. 8.4. Adding hand-crafted features increases the total number of model parameters by 4523. This approach is in alignment with our paradigms.

We trained all 62 shallow neural networks with the hand-crafted features. Figure 8.5 shows the mean classification accuracy of the test and additional test datasets with different building block combinations with and without a-priori hand-crafted features. It can be observed that adding a-priori hand-crafted features generally improves classification accuracy with few exceptions. Accuracy improvements are more pronounced with shallower models. With deeper models, containing more parameters, accuracy improvements are small or non-existent. The average accuracy improvement was 6.0 % with a standard deviation of 4.6.

8.5 Visualisation

A visualisation of the most representative image from CNN's point of view (as proposed by Simonyan et al. [296]) is shown in Fig. 8.6. In this visualisation, the starting image contains only zeros and the loss is minimised for a specific class keeping the trained weights of the model fixed at each iteration. At each iteration, we perform a forward pass, correct the classification, and perform a backward pass, updating the image. The center region of Fig. 8.6b shows more constant values than Fig. 8.6a, which might be an indication of a blocked lumen. However, these results are not easily interpreted.

A different visualisation and understanding of what features are important for a CNN can be obtained by using Grad-CAM [255]. Example results of this are shown in Fig. 8.7 and Fig. 8.8 for five normal and five $KRas^{V12}$ samples, respectively. These figures contain results of using and not using hand-crafted features. The most important features for normal cysts were found at edges and membrane borders (Fig. 8.7). The most important features for $KRas^{V12}$ cysts were larger homogeneous regions, at least without hand-crafted features (Fig. 8.8). With hand-crafted features, edges display the highest values.

Figure 8.11 shows individual weights for convolutional and hand-crafted features after the concatenation layer for the trained '00010' model. Figure 8.12 shows absolute differences of weights from Fig. 8.11. A greater difference means a more important feature. Based on these results, hand-crafted features are less important than convolutional features. The most important single hand-crafted feature for this model and with this data is roughness. Of the hand-crafted features, area was the least important discriminative features.

8.6 Data and importance of segmentation

The inputs to the networks were in-focus segmented unwrapped phase images. For this, we defined a square region around the object using our heuristic segmentation approach (presented in Chapter 7). Each segmented phase image was rescaled to 227×227 pixels as required by the network. This process is illustrated in Fig. 8.13. This approach kept some of the background in the inputs. In order to investigate how important the background is and if the cysts in the field of view should be segmented from the background alone, all the values within a circular region with a diameter of the bounding box side length were set to 0. The image was cropped relative to the bounding box with parameters +25 pixels to -45 pixels (with 5 pixel steps). Then the cropped region was rescaled back up to the original image size. This permitted us to preserve different amounts of background in input images (example inputs shown in Fig. 8.14). For this experiment, the AlexNet [121] architecture as defined in Section 4.3.4 was used.

In total, 980 MDCK cyst holograms were used in this experiment (500 wild-type and 480 oncogenic) and this was partitioned between training (400), validation (380), and

testing (200) (example phase reconstructions are shown in Fig. 8.15). In addition to the MDCK cell line, a human prostate (RWPE-1) epithelial cell line was utilised from 13 independent imaging experiments (example phase reconstructions shown in Fig. 8.16). In total 818 RWPE holograms were used and this data was split into training (423), validation (180), and testing (215).

Figure 8.17 shows classification accuracies with different background proportions. It can be observed that, for both cell lines, the classification accuracy remained high even with a small fraction of the background. With the MDCK cell line, accuracies varied from 93 % (+25 pixels) to 71 % (-45 pixels). With the RWPE cell line, accuracies varied from 81.4 % (+ 20 pixels) to 74.9 % (-40 pixels). Given that all accuracies are higher than 70 %, it can be verified that the networks were not failing with any of the inputs, and remarkably the networks were able to learn distinguishing features from small proportions of the background alone.

8.7 Conclusions

In this chapter, it was shown that a CNN can be used to reliably classify normal and oncogenic cysts. The phase data needs to be normalised for wavelength independency and magnification independency can be obtained by rescaling the data. The results of a large scale comparison of different CNN architectures for a binary classification task were given. It was shown that by adding a-priori information to a CNN, it is possible to improve the classification accuracy of shallow models. Shallow CNNs with hand-crafted features (in our experiments) have higher classification accuracy than the same models without hand-crafted features, and may be more robust to the noise present in real biological samples. CNNs containing more learning parameters are able to learn discriminative features that are equally or more important than hand-crafted features, making the hand-crafted features less useful.

In addition, an experiment on the importance of background removal was performed. Based on these results, it can be concluded that samples prepared and cultured at different times may leave visible artefacts on the background that can be learned by a CNN. Also, the imaging system needs to be examined carefully so that the quality of the captured

holograms remains constant between different captures. Dissimilar lighting conditions or other artefacts introduced into the system can fabricate a system-state-specific "fingerprint" that is transferred to captured data. Liquid solutions (Matrigel^{TM}, pH buffers etc.) inside the petri dish need to be carefully adjusted and constant amounts should be used between different captures. Finally, the environmental conditions (CO_2 levels, humidity, temperature) should be stable and constant between the captures. One solution to this problem is an accurate segmentation of objects-of-interest where the background cannot interfere with the classification.

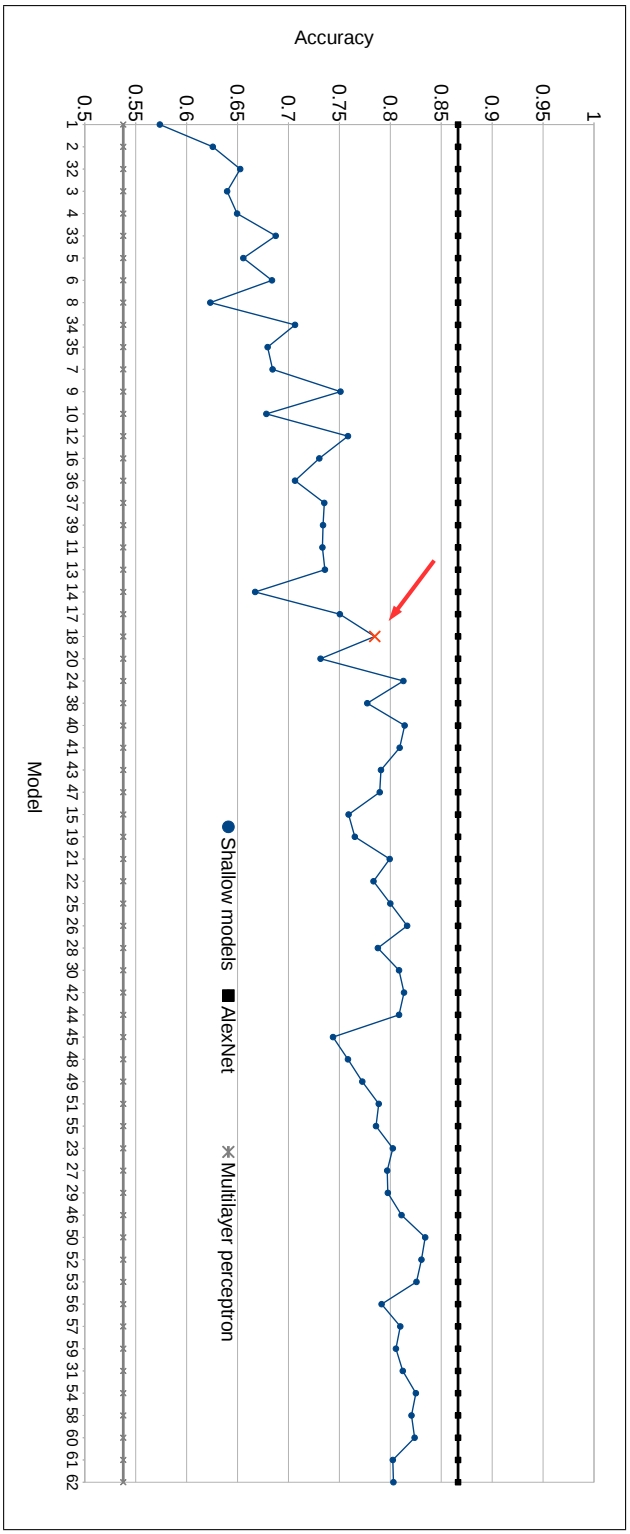

Figure 8.3 Accuracy of different architectures, showing accuracies of different shallow models. In addition, AlexNet and multilayer perceptron accuracies are shown. The model number on the horizontal axis refers to the indices in Tables 8.1 and 8.2. The arrow indicate the accuracy of model '00010'.

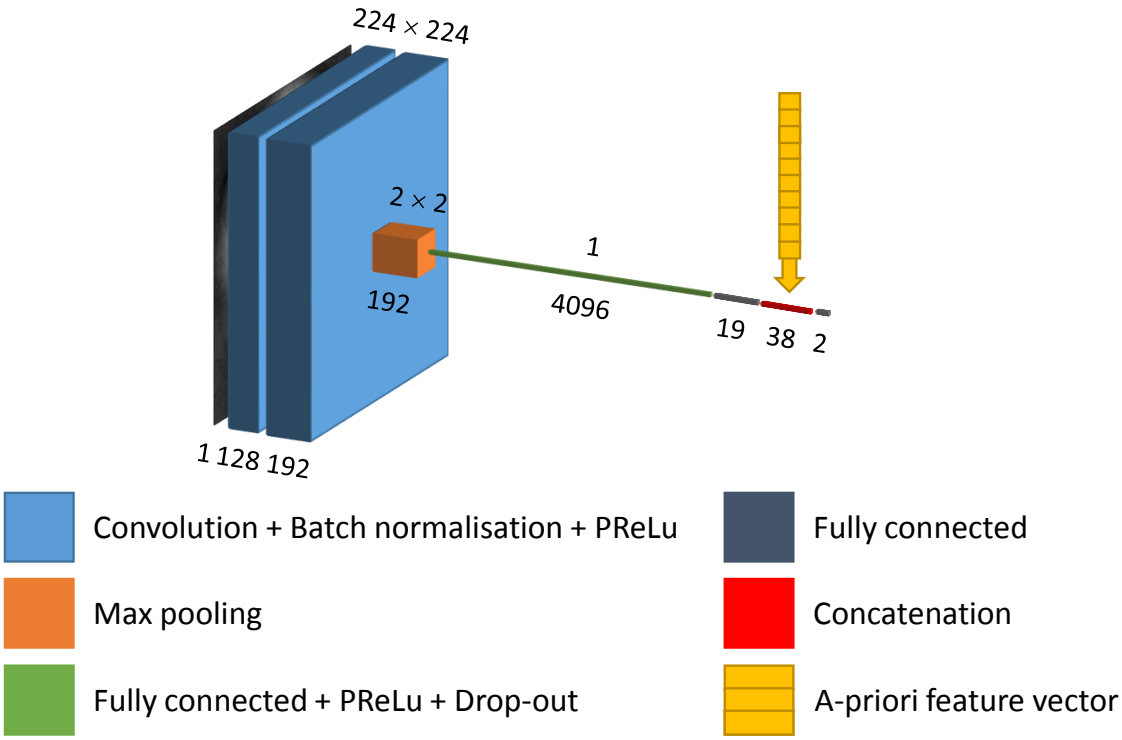

Figure 8.4 A-priori feature addition. See section 8.4 for more information.

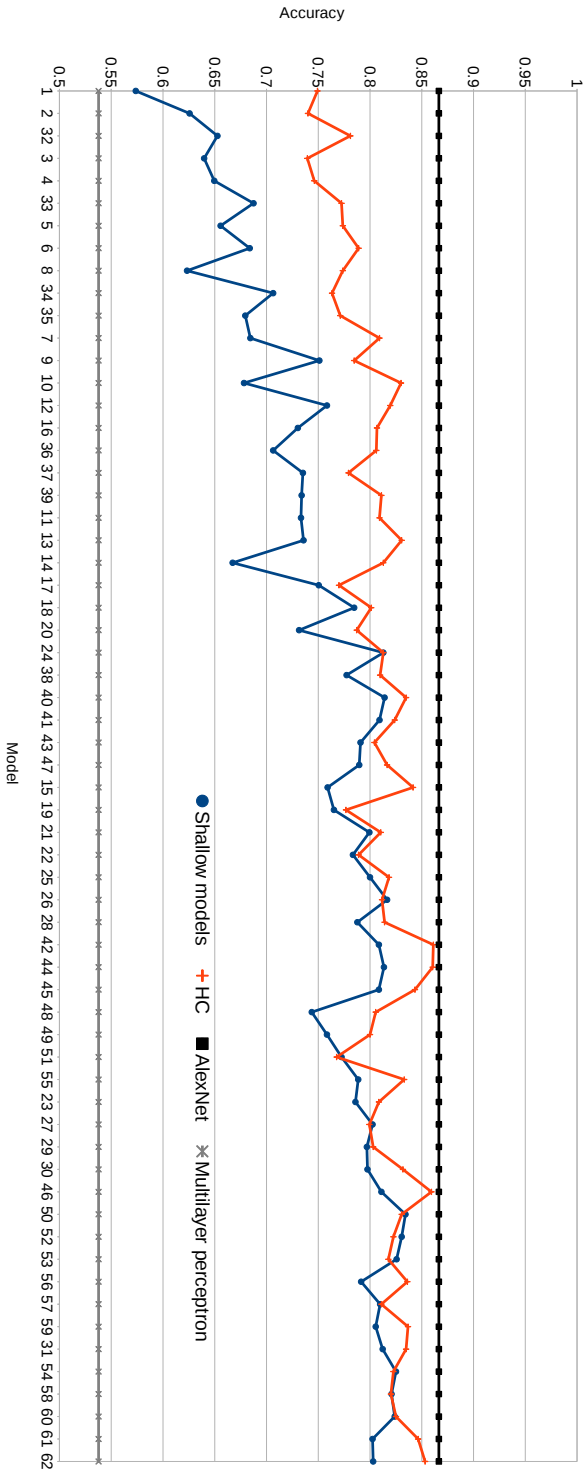

Figure 8.5 Classification accuracies of different shallow models with and without hand-crafted features. In addition, AlexNet and multilayer perceptron accuracies are shown. The model number on the horizontal axis refers to the indices in Tables 8.1 and 8.2.

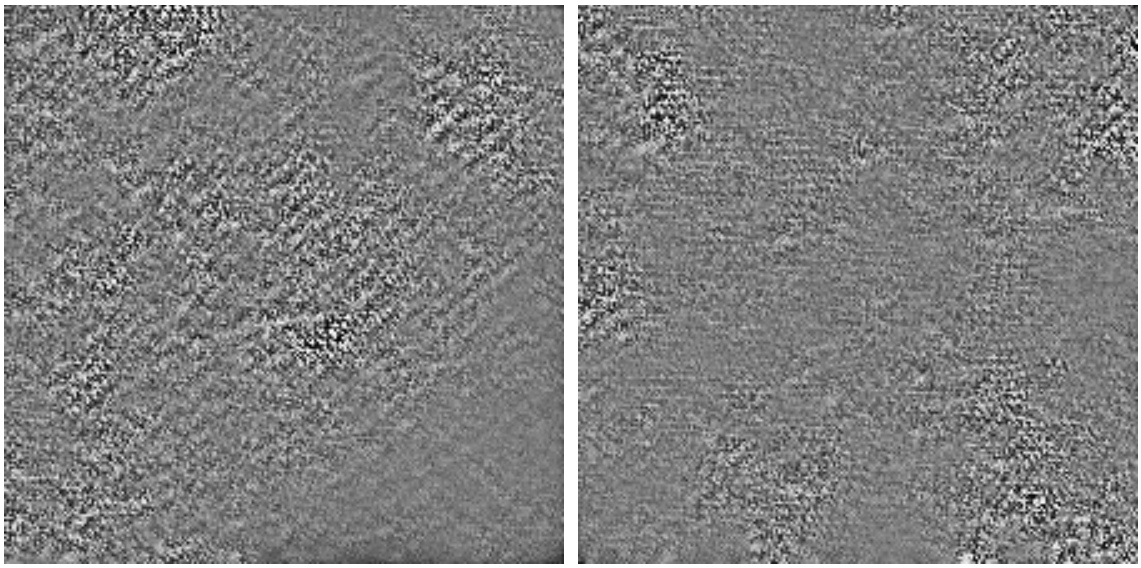

Figure 8.6 Class specific image creation. Model '00010' was used to create class specific images over 100 iterations. Visualisation of (a) normal and (b) $KRas^{V12}$ cysts. See text for more explanation.

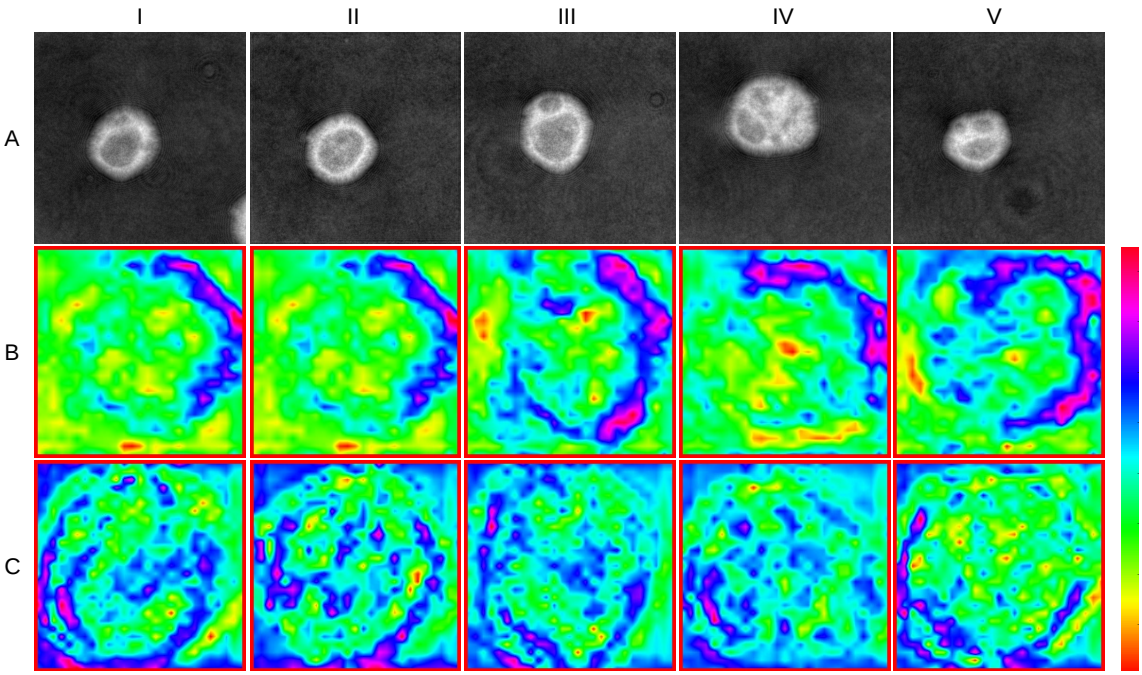

Figure 8.7 Grad-CAM for five normal MDCK cysts I-V. Row A: unwrapped phase reconstruction, row B: Grad-CAM on segmented phase reconstruction without hand-crafted features, row C: Grad-CAM on segmented phase reconstruction with hand-crafted features. CNN probabilities for class 0: BI: 0.853, BII: 0.981, BIII: 0.934, BIV: 0.845, BV: 0.965, CI: 0.645, CII: 0.859, CIII: 0.883, CIV: 0.807, CV: 0.911. Model '00010' was used to obtain these results. Colourbar: values are normalised to be from 0 to 255 for visualisation.

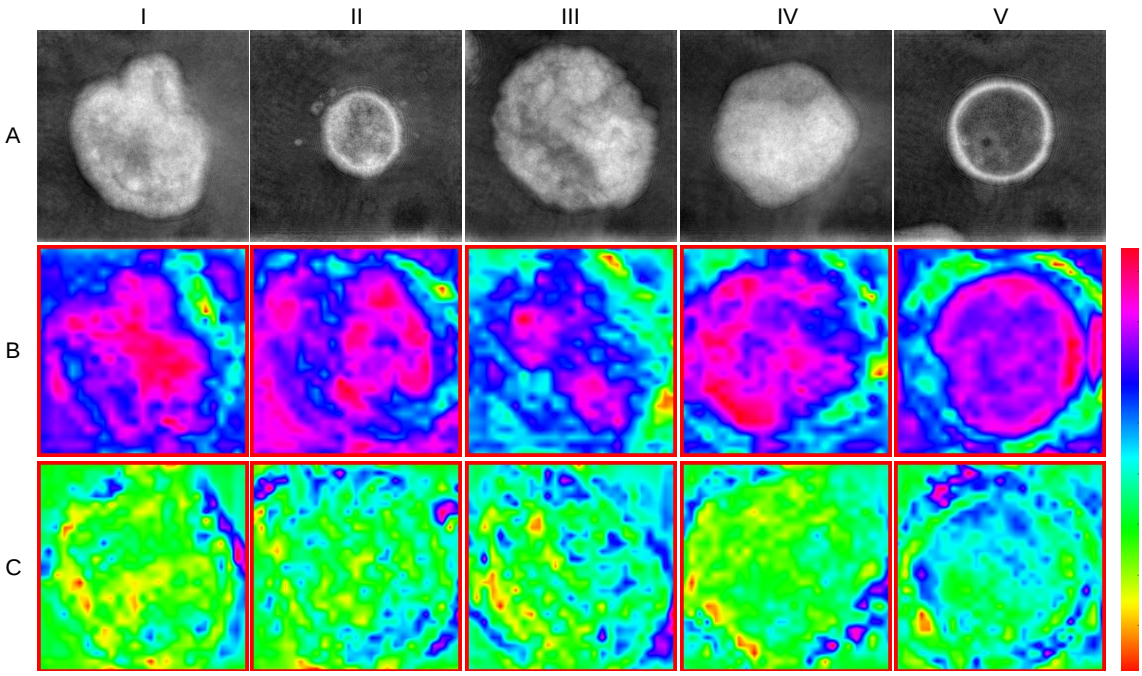

Figure 8.8 Grad-CAM for five KRas^{V12} MDCK cysts I-V. Row A: unwrapped phase reconstruction, row B: Grad-CAM on segmented phase reconstruction without hand-crafted features, row C: Grad-CAM on segmented phase reconstruction with hand-crafted features. CNN probabilities for class: 1 BI: 0.983, BII: 0.642, BIII: 0.999, BIV: 0.987, BV: 0.976, CI: 0.999, CII: 0.969, CIII: 0.999, CIV: 0.983, CV: 0.979. Model '00010' was used to obtain these results. Colourbar: values are normalised to be from 0 to 255 for visualisation.

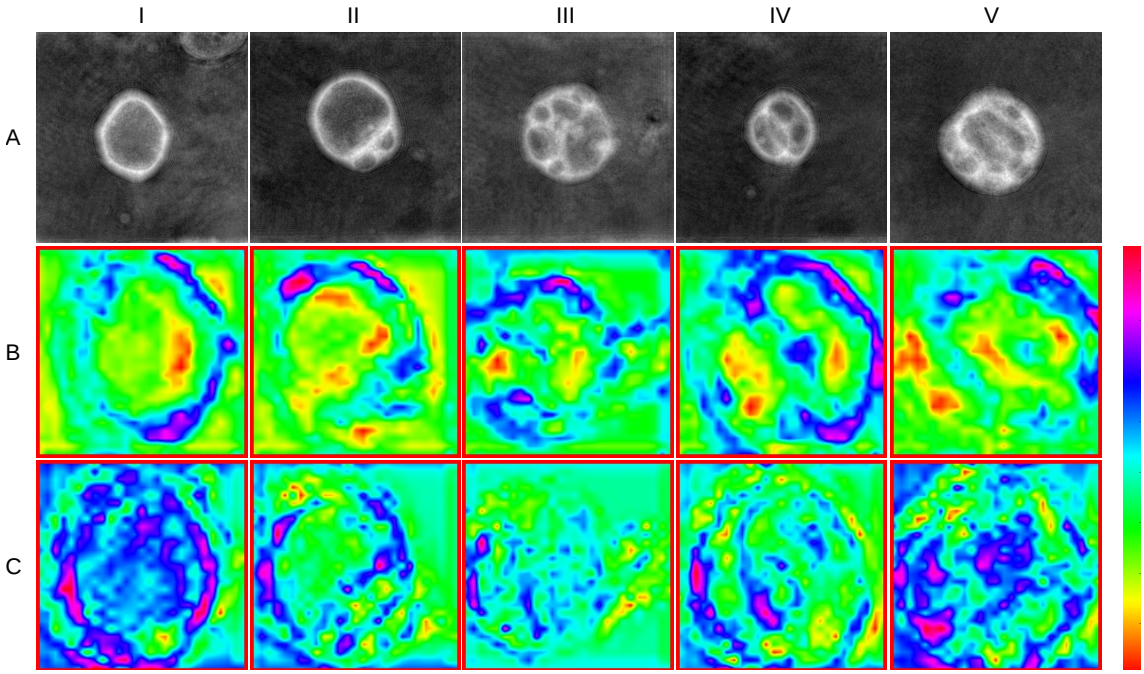

Figure 8.9 Grad-CAM incorrect classification for five normal MDCK cysts I-V. Row A: unwrapped phase reconstruction, row B: Grad-CAM on segmented phase reconstruction without hand-crafted features, row C: Grad-CAM on segmented phase reconstruction with hand-crafted features. CNN probabilities for class 0: BI: 0.304, BII: 0.102, BIII: 0.001, BIV: 0.033, BV: 0.044, CI: 0.072, CII: 0.045, CIII: 0.154, CIV: 0.069, CV: 0.001. Model '00010' was used to obtain these results. Colourbar: values are normalised to be from 0 to 255 for visualisation.

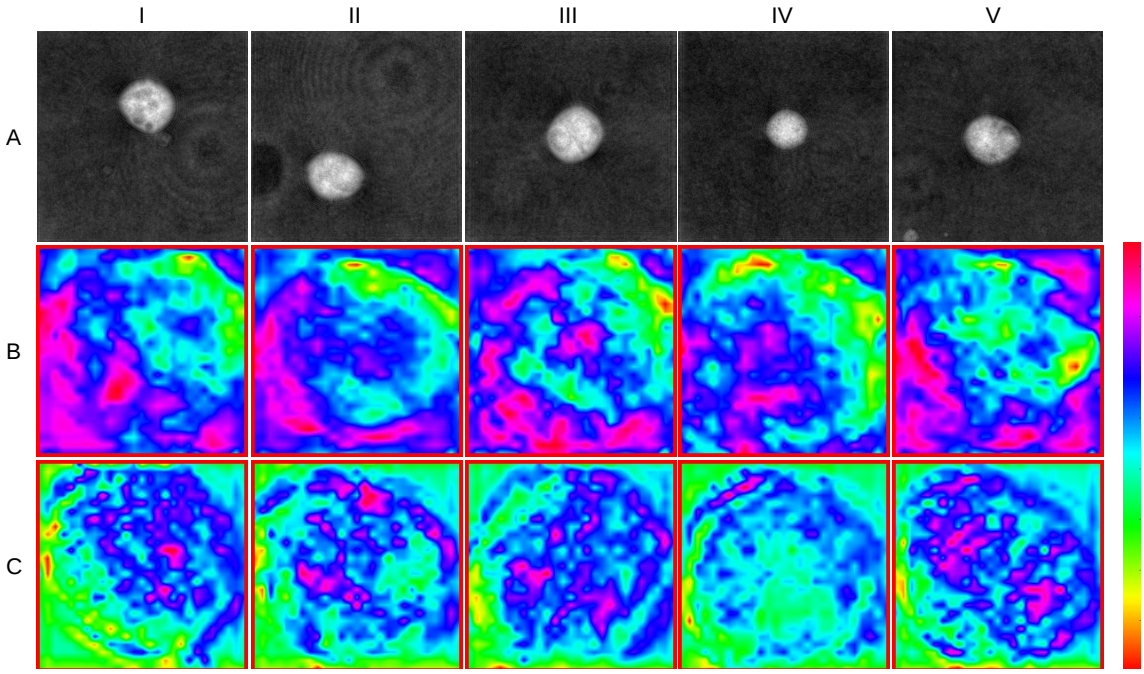

Figure 8.10 Grad-CAM incorrect classification for five $KRas^{V12}$ MDCK cysts I-V. Row A: unwrapped phase reconstruction, row B: Grad-CAM on segmented phase reconstruction without hand-crafted features, row C: Grad-CAM on segmented phase reconstruction with hand-crafted features. CNN probabilities for class 1: BI: 0.095, BII: 0.155, BIII: 0.243, BIV: 0.036, BV: 0.039, CI: 0.24, CII: 0.079, CIII: 0.042, CIV: 0.041, CV: 0.065. Model '00010' was used to obtain these results. Values are normalised to be from 0 to 255 for visualisation.

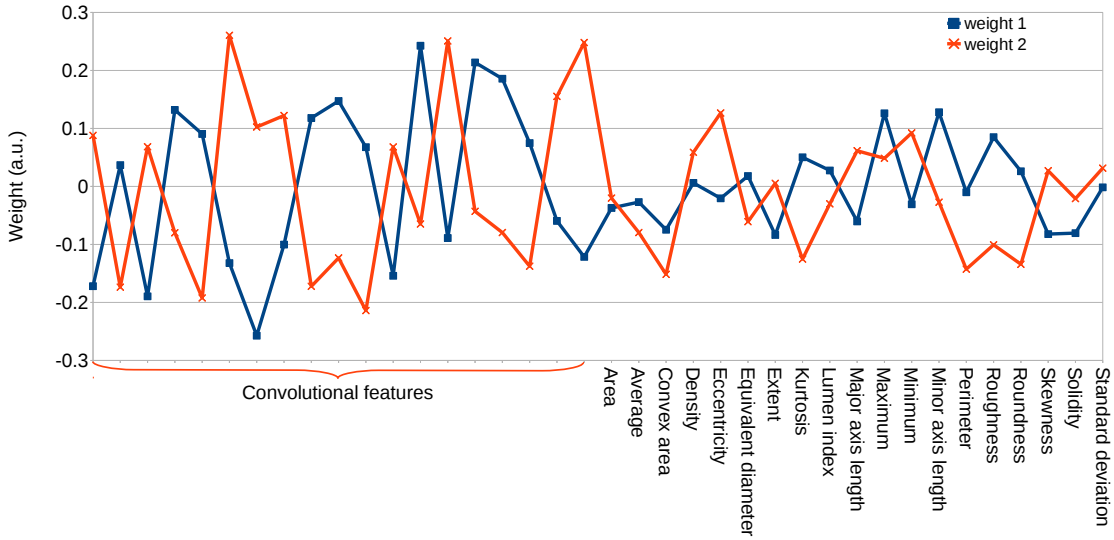

Figure 8.11 Concatenated weights plotted. Showing actual learned weight values of the concatenation layer of the '00010' architecture model.

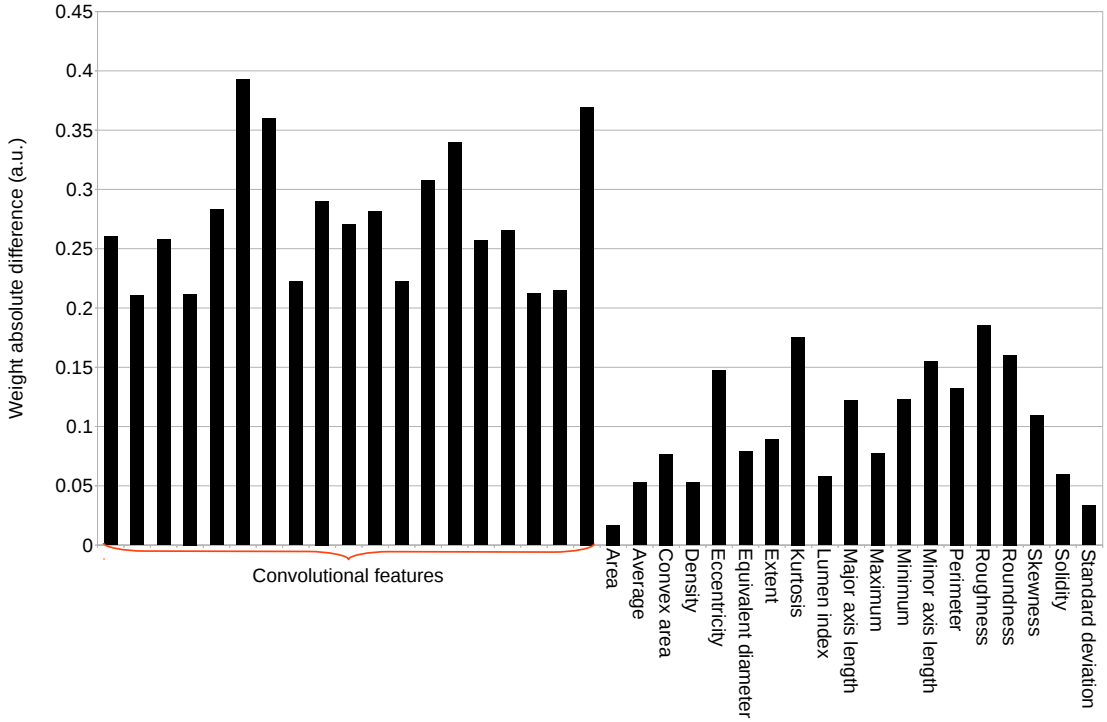

Figure 8.12 Absolute difference of weights (from Fig. 8.11) of the '00010' model.

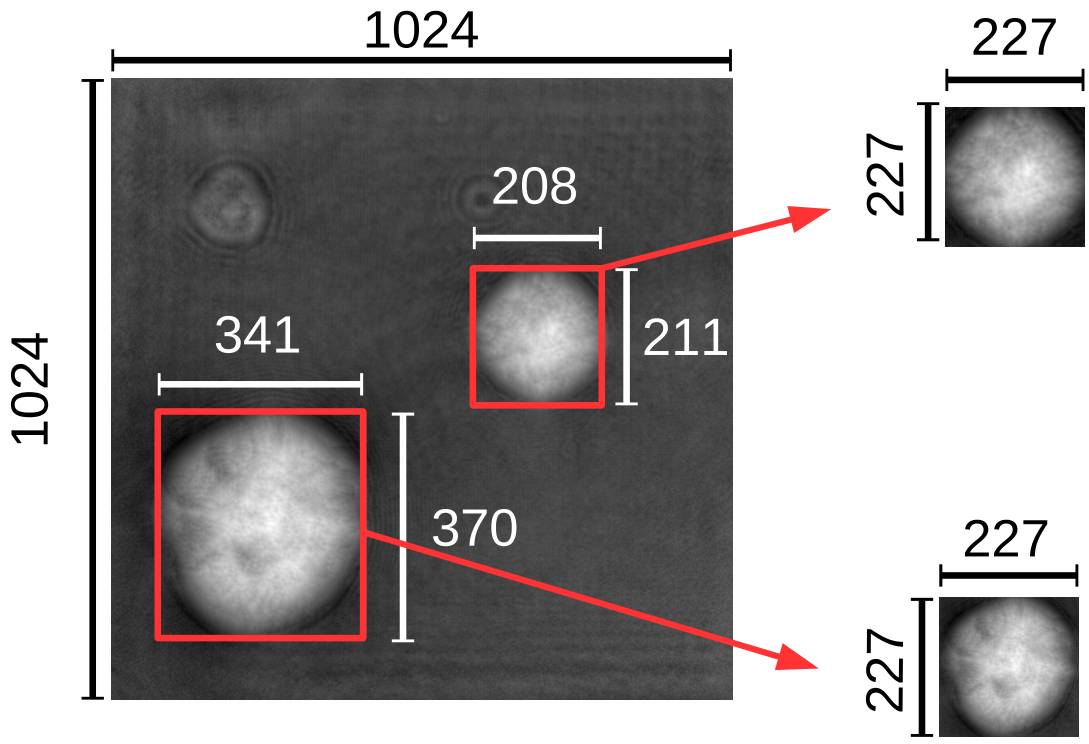

Figure 8.13 Phase image preprocessing. Cysts in 1024×1024 unwrapped phase images were segmented and rescaled to 227×227 pixels. This approach allows the use of different magnifications during imaging.

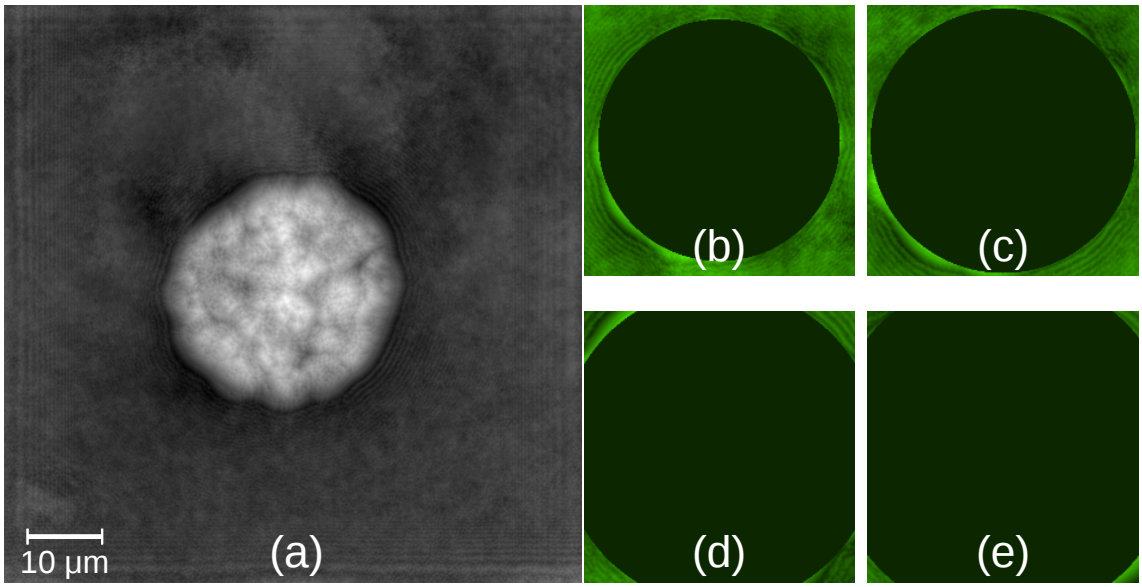

Figure 8.14 RWPE cell cluster with different amounts of background preserved. (a) original phase reconstruction, (b) - (e) example inputs with different portions of background (bounding box side length average) (b) +25 pixels, (c) +5 pixels, (d)-30 pixels, (e) -45 pixels. (b-e shown with a false colour-map for visualisation)

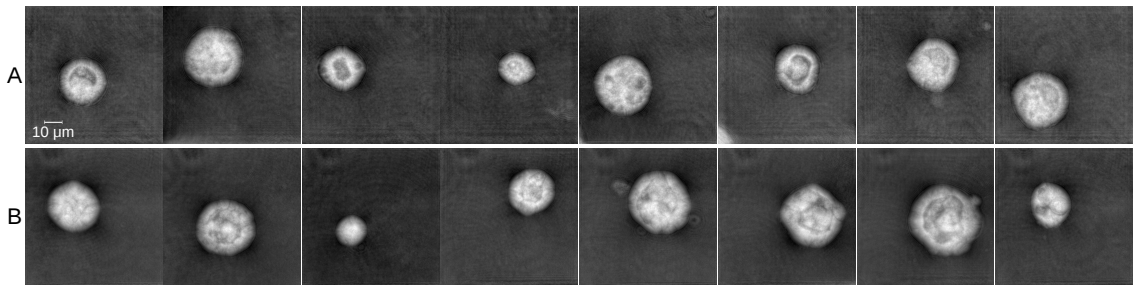

Figure 8.15 Example MDCK phase reconstructions. Row A: wild-type, row B: oncogenic.

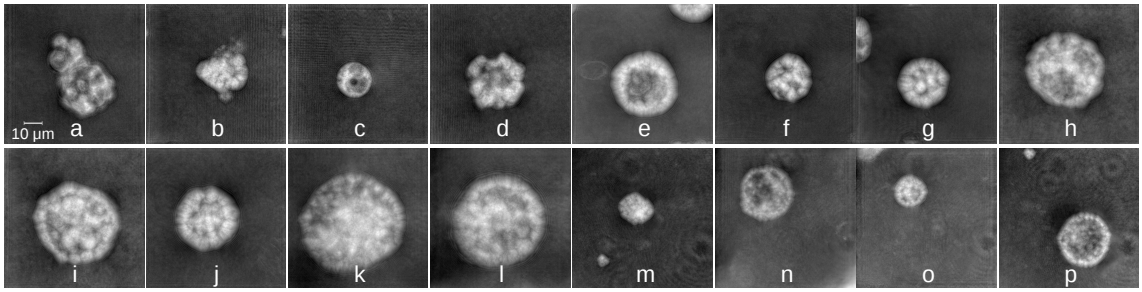

Figure 8.16 Example RWPE phase reconstructions. Example RWPE cell clusters from each imaging experiment. a-l are from experiments 1-12, and m-p are from experiment 13.

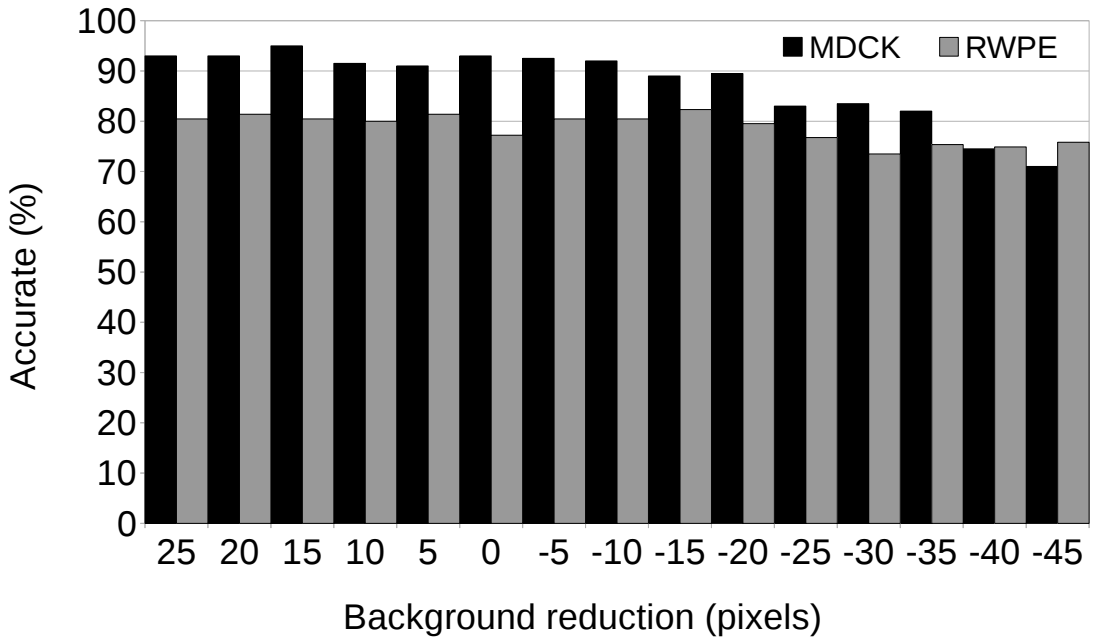

Figure 8.17 Accuracies with MDCK and RWPE samples, showing classification accuracies with different amounts of background in the inputs.

9 Time-lapse and patient samples

9.1 Introduction

Microscope modifications (integrated incubator, motorised translation stages, and related software) allow time-lapse imaging and monitoring morphological changes of cysts over long periods of time.

University Oulu Hospital (UOH) and Biocenter Oulu (BCO) have conducted collaborative research where patients diagnosed with prostate cancer went through a prostatectomy, providing tissue samples for research purposes. All participating patients were volunteers and gave written consent for the research. Tissue samples were prepared for further research purposes, and imaged holographically. Perhaps the greatest challenge with the patient derived data is that we do not have real control samples, i.e. we do not have samples of healthy tissue.

In this chapter, we illustrate that gathering time-lapse data of living cysts is possible. In addition, we show that the system can be used as-is for real patient samples. We use k-means clustering and different classifiers with the captured holograms.

9.2 Time-lapse data

An example time-lapse sequence of MDCK cysts is shown in Fig. 9.1. Figure. 9.2 shows change rates of different extracted features (see Chapter 4 for details of extracted features). Cysts were binary segmented using the heuristic used to create the ground truth mask with which to train the segmentation CNN from Chapter 7. The CNN was not used

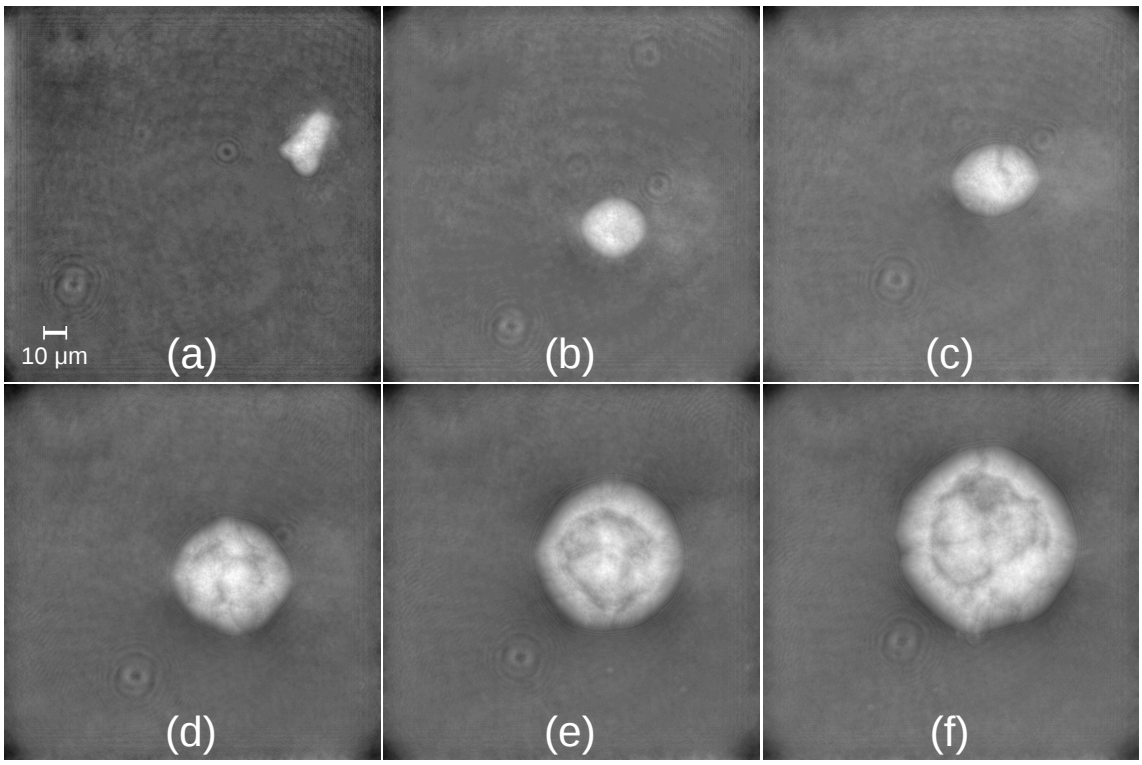

Figure 9.1 Example phase sequence from time-lapse imaging at different imaging times. (a) 0 h, (b) 24 h, (c) 48 h, (d) 72 h, (e) 96 h, and (f) 120 h.

for segmentation in this experiment (it was developed later chronologically). For these experiments we also use an alternative (earlier developed) lumen index compared to that defined in Section 4.2.5.

For each mask $R'(x,y;z)$ (defined in Section 7.3) where the ground-truth masks were applied we constricted the mask towards the centre of the mass by approximately 10 μm (50 pixels) and obtained a new mask $L(x,y;z)$ that covered only the lumen. By subtracting $L(x,y;z)$ from $R'(x,y;z)$ we obtain another mask, $B(x,y;z)$, for the cells between the basolateral and apical membranes. We calculated a heuristic lumen index, l, as

$$l = \frac{\sum_{x=0}^{M-1}\sum_{y=0}^{N-1} L(x,y;z)S(x,y;z)}{A\bar{B}} \tag{9.1}$$

where $\bar{B}$ is the average of maxima along the y axis of $B(x,y;z)S(x,y;z)$, and M and N are width and height of $L(x,y;z)$, respectively. $S(x,y;z)$ is defined as

$$S(x,y;z) = \phi'(x,y;z)R'(x,y;z) \tag{9.2}$$

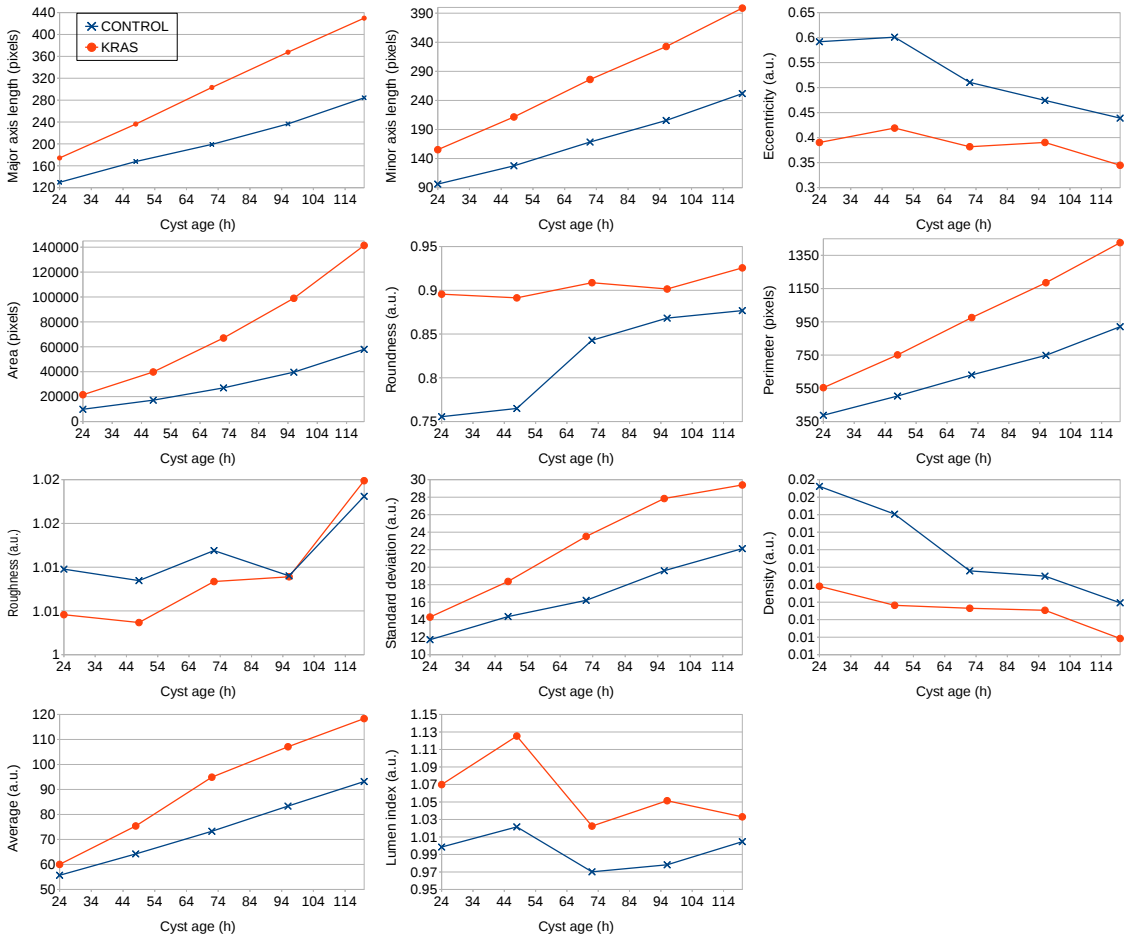

Figure 9.2 MDCK change rates of different extracted features. Values are averages of 20 healthy and KRas^{V12} MDCK cysts.

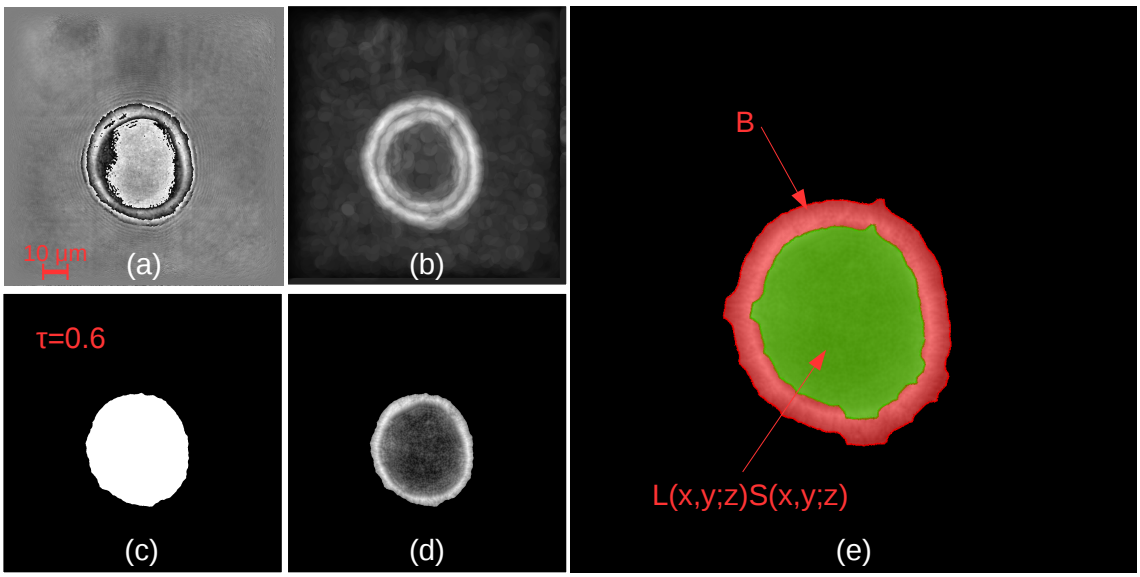

Figure 9.3 Cyst segmentation procedure. (a) aberration free wrapped phase $\phi(x,y;z)$, (b) morphological gradient image $G(x,y;z)$ calculated from (a), (c) binary mask $M(x,y;z)$ obtained through processing (b), (d) segmented unwrapped phase reconstruction, (e) illustration of different regions of a cyst: B (red) region of cells between the basolateral and apical membranes and lumen (green).

where $\phi'(x,y;z)$ is a part of the wrapped phase image $\phi(x,y;z)$ with same coordinates as the bounding box for the segmentation mask. A is defined as

$$A = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} L(x,y;z). \tag{9.3}$$

9.2.1 Principal component analysis and *k*-means clustering

For PCA-k-means combination, we used time-lapse data captured with our automated system. The hologram capture was realised over 144 hours (6 days) where discrete (x,y) positions were captured every 24 hours. This was performed for both healthy and KRas^{V12} samples. For the actual analysis and experiment, data from days 2 to 6 were used. 20 sequences from both healthy and KRas^{V12} cysts were used for analysis.

Principal component analysis can be used to reduce dimensions of high dimensional data. For this, we are interested if PCA together with k-means clustering can be used to split cysts into tumourigenic and healthy cysts. Each cyst was segmented from the background and quantitative phase and morphology related features (Table. 9.1) were extracted as well as the rate of change of each feature. Features were standardised so that

Table 9.1 Morphology and phase related features.

Index	Name	Type
1	Major axis length	Morphology
2	Minor axis length	Morphology
3	Eccentricity	Morphology
4	Area	Morphology
5	Roundness	Morphology
6	Perimeter	Morphology
7	Roughness	Morphology
8	Standard deviation	Phase
9	Density	Phase
10	Average	Phase
11	Lumen index (I)	Phase

Table 9.2 Component loadings without change rates. Index refers to Table 9.1.

Hours	24		48		72		96		120	
Index/PC	1	2	1	2	1	2	1	2	1	2
1	0.41	0.15	0.37	0.19	0.39	0.18	0.39	0.08	0.39	0.07
2	0.42	-0.19	0.40	-0.06	0.40	-0.01	0.39	-0.09	0.39	-0.06
3	-0.13	0.43	-0.25	0.38	-0.17	0.63	-0.11	0.61	-0.11	0.60
4	0.44	-0.07	0.40	0.02	0.40	0.08	0.39	0.00	0.38	0.04
5	0.09	-0.47	0.23	-0.46	0.17	-0.63	0.10	-0.64	0.11	-0.61
6	0.44	0.01	0.39	0.08	0.40	0.11	0.40	0.01	0.39	0.05
7	-0.03	0.27	-0.10	0.40	0.00	0.16	0.06	0.20	0.20	0.41
8	0.33	0.30	0.30	0.34	0.37	0.01	0.37	0.15	0.35	-0.07
9	-0.01	0.44	-0.17	0.40	-0.17	-0.14	-0.19	0.27	-0.25	-0.04
10	0.26	0.40	0.25	0.40	0.37	0.22	0.36	0.23	0.36	0.18
11	0.27	-0.16	0.29	0.07	0.15	-0.25	0.23	0.13	0.14	-0.20

the data had zero mean and a standard deviation of one. *k*-means clustering was run on each day's data ten times with different centroid seeds to avoid local minima. The maximum number of iterations was 300 for a single run. The seed initialisation was performed with the *k*-means++ initialisation scheme [297].

Table 9.2 and Table 9.3 show component loadings without and with change rates, respectively. It can be observed that there are no strong correlations between features and principal components and most of the features are therefore contributing equally to principal components. Based on this result, lumen index, l, is the only feature that does not show component loading values greater than 0.3 with any dataset without change rate. With the data using the change rate, component loadings are even more balanced.

Figures 9.4 and 9.5 show principal component analysis of different age cysts without and with change rates, respectively. These figures display also *k*-means clustering result.

Table 9.3 Component loadings with change rates. Index refers to Table 9.1. CR=change rate of corresponding feature.

Hours	24		48		72		96		120	
Index/PC	1	2	1	2	1	2	1	2	1	2
1	0.29	0.11	0.31	0.01	-0.29	0.05	0.31	0.00	0.32	-0.03
CR	0.29	0.11	0.21	0.24	-0.27	-0.19	0.28	0.10	0.25	0.16
2	0.29	-0.14	0.32	-0.11	-0.29	0.15	0.31	-0.11	0.31	-0.10
CR	0.29	-0.14	0.26	-0.22	-0.23	0.21	0.26	-0.19	0.24	-0.17
3	-0.09	0.30	-0.17	0.21	0.08	-0.36	-0.07	0.39	-0.04	0.37
CR	-0.09	0.30	0.02	0.29	-0.12	-0.42	0.03	0.37	0.01	0.42
4	0.31	-0.05	0.33	-0.07	-0.29	0.10	0.31	-0.06	0.32	-0.06
CR	0.31	-0.05	0.33	-0.04	-0.29	0.06	0.31	-0.06	0.29	-0.07
5	0.07	-0.33	0.15	-0.26	-0.08	0.36	0.06	-0.42	0.05	-0.38
CR	0.07	-0.33	-0.03	-0.34	0.16	0.41	-0.04	-0.39	-0.02	-0.42
6	0.31	0.01	0.33	-0.04	-0.29	0.11	0.31	-0.05	0.32	-0.05
CR	0.31	0.01	0.29	0.10	-0.28	-0.04	0.29	-0.02	0.29	0.03
7	-0.02	0.19	-0.05	0.24	0.03	0.16	0.06	0.15	0.18	0.15
CR	-0.02	0.19	-0.04	0.14	-0.02	0.07	0.08	0.10	0.18	0.17
8	0.24	0.21	0.26	0.17	-0.26	0.17	0.30	0.05	0.27	-0.04
CR	0.24	0.21	0.10	0.33	-0.23	-0.14	0.19	0.25	0.06	0.25
9	-0.01	0.31	-0.11	0.32	0.16	0.13	-0.12	0.20	-0.20	0.10
CR	-0.01	0.31	0.02	0.32	-0.14	-0.18	0.02	0.32	-0.01	0.20
10	0.18	0.28	0.23	0.18	-0.27	0.04	0.29	0.09	0.30	0.08
CR	0.18	0.28	0.08	0.31	-0.24	-0.21	0.14	0.23	0.17	0.31
11	0.19	0.11	0.23	0.03	-0.08	0.28	0.17	0.05	0.09	-0.03
CR	0.19	-0.11	0.14	0.08	0.11	0.06	0.05	0.10	-0.04	0.13

Table 9.4 shows classification accuracies that are based on confusion matrices shown in Figures 9.4 and 9.5. From these results, it can be seen that classification accuracies vary between different age cysts and accuracy does not necessarily improve with older cysts as could be expected. Adding the change rates to the features improves classification accuracy with some cases and decreases accuracy sometimes.

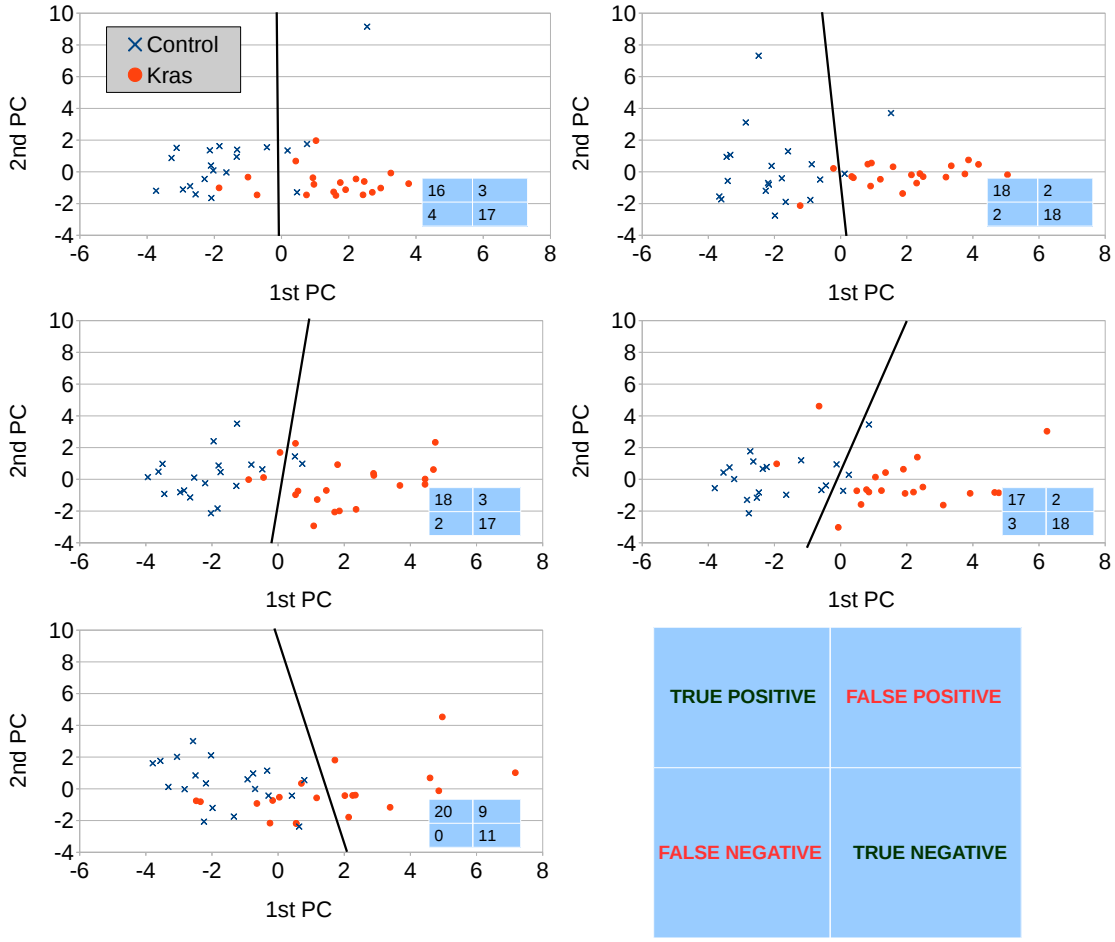

Figure 9.4 PCA and clustering without change rate. Principal component analysis at different imaging intervals showing first and second principal components on the 11 features in Table 9.1. (a) 24 h – 48 h, (b) 48 h – 72 h, (c) 72 h – 96 h, (d) 96 h – 120 h, (e) 120 h – 144 h. The black line shows the border between clusters.

9.3 Patient data

University Oulu Hospital (UOH) and Biocenter Oulu (BCO) have conducted a collaborative research where patients diagnosed with a prostate cancer went through a prostatectomy, providing tissue samples for research purposes. All participating patients were volunteers and gave written consent for the research.

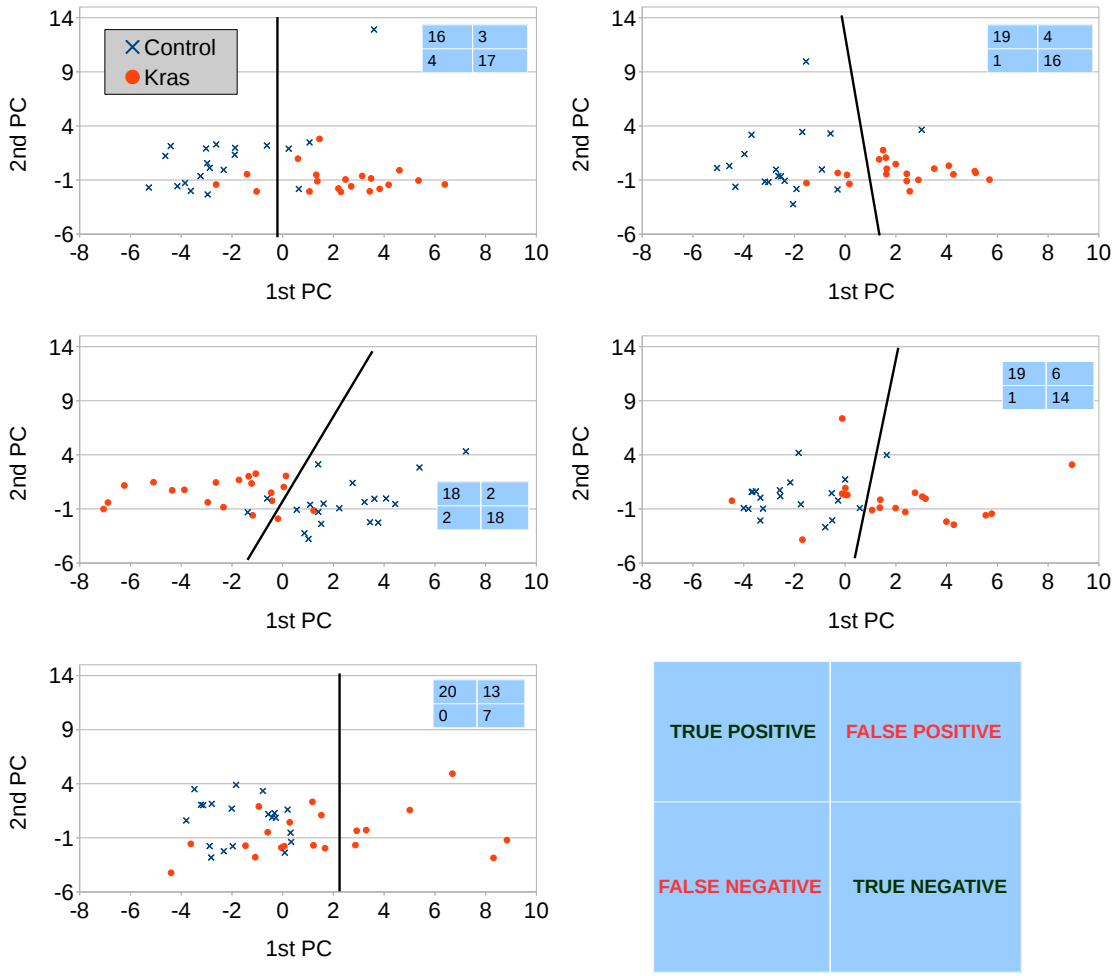

Figure 9.5 PCA and clustering with change rate. Same as Fig. 9.4 except 22 features used for PCA (11 features in Table 9.1 and their change rates).

Surgically removed prostate glands were analysed by a pathologist who performed a histological evaluation of an adjacent tissue slice of the slice going to be used later in research. The histological evaluation contained an estimation of possible amount of cancer tissue. In total 13 patients of 37 were used in this study. These patients were selected by researchers of BCO based on the histological evaluation containing tissue samples of benign hyperplasia (3 patients) and different proportions of cancerous tissue ('LOW' to 'HIGH'). For three of the patients histological evaluation was not performed ('UNKNOWN'). Patients are grouped based on proportion of cancerous tissue as shown in Table 9.5.

Table 9.6 shows grouping of selected 13 patients and number of organoids used in the study.

Table 9.4 Classification accuracy with *k*-means clustering.

Cyst age (h)	Accuracy (%)	Change rate included
24-48	82.5	
48-72	90.0	
72-96	87.5	
96-120	87.5	
120-144	77.5	
24-48	82.5	x
48-72	87.5	x
72-96	90.0	x
96-120	82.5	x
120-144	67.5	x

Table 9.5 Patient encoding. See text for more explanation.

Proportion of cancer cells (%)	Label
0	'BENIGN'
1 - 20	'LOW'
21 - 100	'HIGH'
Not determined	'UNKNOWN'

Table 9.6 Patient derived tissue samples showing group label based on evaluation of a pathologist and total number of organoids.

Patient ID	Group label	Number of organoids
Z0	'BENIGN'	313
L3	'BENIGN'	45
F0	'BENIGN'	295
Y8	'LOW'	196
D7	'LOW'	280
M8	'LOW'	288
B6	'LOW'	270
L9	'HIGH'	250
W9	'HIGH'	88
W5	'HIGH'	300
E2	'UNKNOWN'	176
T5	'UNKNOWN'	213
A7	'UNKNOWN'	307

9.3.1 From tissue samples to organoids

The following three paragraphs are largely taken from [298].

Tissue samples were transported from UOH to BCO where they were processed further. The taken tissue samples were prepared immediately for cell culturing. Tissue samples were sliced in small parts and cells were isolated through collagen treatment that was followed by a short-term trypsin treatment. Treated samples were preserved in freezing solution containing 90 % fetal bovine serum (FBS, Life Technologies) and 10 % dimethyl sulfoxide (DMSO) and kept in liquid nitrogen for later experiments. Cells in these samples were not homogeneous, instead each sample contained mixture of different cell types.

Cells were thawed on ice and then suspended into starting medium. Cells were centrifuged for 5 minutes, and a pellet was suspended into pre-warmed complete medium (MEM-Glutamax; 41090-093, Life Technologies). The cell suspension was mixed with Matrigel^{TM} in 1:4 ratio and divided into four wells in 24-well plate. Samples were incubated in +37 °C for 2 minutes. After that the plates were turned upside down and the Matrigel^{TM} was allowed to solidify in the same temperature for 15 minutes. A medium with rho-associated protein kinase inhibitor (ROCK inhibitor Y-27632) and nicotinamide was added on top of it. The medium was replaced every second day. During the first week both Y-27632 and nicotinamide were added to the medium and afterwards only nicotinamide. 35 mm glass-bottom μ-Dish (Ibidi GmbH, Gräfelfing, Germany) was blocked with phosphate buffered saline (PBS) containing 1 % bovine serum albumin (BSA) in +37 °C for an hour. After the blocking, the μ-Dish was washed with cold PBS and cooled down to +4 °C. A thin layer of Matrigel^{TM} was spread on top of the μ-Dish and the dish was placed back to +4 °C for 5 minutes and then transferred into +37 °C for 30 minutes to solidify.

Meanwhile organoids that were grown in Matrigel^{TM} drops were washed with PBS and cell recovery solution was added covering the blobs. Cells were incubated in a cold room on shaker for 30 minutes. The melted solution was transferred into Eppendorf tube and organoids were spun down 10 seconds by using a table top centrifuge. The supernatant was removed, and pellet was suspended into TrypLE. This was followed by 30 minutes incubation in +37 °C with 5 % CO_2. Organoids were dissolved by pipetting up and down several times during incubation. The TrypLE was inactivated by adding equal amount

of soybean trypsin inhibitor solution and pipetting up and down at least six times. This was followed by 5 minutes incubation at room temperature. The cells were spun down, supernatant was removed, and pellet was resuspended into culture medium (Complete Advanced DMEM/F12 medium with Y-27632 and nicotinamide). Equal amounts of cell suspension and 2 % Matrigel^{TM} were mixed and placed on top of the solidified 100 % Matrigel^{TM}. After the cells were attached to the bottom, Complete Advanced DMEM/F12 medium containing 2 % Matrigel^{TM} was added on top of the cells. Medium was replaced every second day. After three days, complete medium was added without Matrigel^{TM} and Y-27632. The patient-derived organoids were grown for one week in overlay before imaging. For these experiments, holographic imaging was performed manually — i.e. organoids in the dish were searched manually.

Cell sorting was not performed due to limited amount of cells and the fact that cell sorting is harmful for cells. Because the cell sorting was not performed, cultured samples contained not only epithelial prostate cells but also stromal cells. This fact can complicate the analysis as cultures with these different cell types can form complex structures [299] posing additional challenges.

9.3.2 Clustering

Clustering applied on patient data includes various different problems that were identified during the experiments. These problems together with approaches used to solve problems are shown in Table 9.7. Definitions of the problems are as follows.

Definition 9.3.1. Clustering consistency problem. Clustering of each patient is performed based on different set features. The problem exists only if patient samples are clustered one patient at a time based on extracted features of this very same patient.

Definition 9.3.2. Cluster relevance problem. Obtained clusters do not have relevance to samples being benign or cancerous. The problem exists if cluster centres having no relevance to disease are used for prediction.

Definition 9.3.3. Cluster indexing problem. Meaning of assigned cluster indexes is different for different patients. The problem is applicable only if patient samples are clustered one patient at the time.

Table 9.7 Clustering problems. Showing problems related to k-means clustering and approaches to solve the problems. All problems and solution approaches are presented in this chapter. x denotes existence of a problem, ✓ denotes that the problem is solved by an approach, N/A denotes that the problem is not applicable, and * denotes future research and this approach is not tested.

Approach /Problem	**Clustering each patient to two clusters**	**Group patient data and cluster to two clusters**	**Fix two cluster centres based on 'BENIGN' and 'HIGH' patients**	**Apply PCA, cluster each patient to two clusters, apply 2nd order clustering**	**Include real control data***
Clustering consistency problem	x	✓	✓	✓	✓
Cluster relevance problem	x	x	✓	x	✓
Cluster indexing problem	x	N/A	N/A	✓	✓
Index mapping problem	x	x	x	x	✓

Definition 9.3.4. Index mapping problem. A mapping from cluster centre indexes to cancer and benign can not be performed.

The last identified problem can not be tested within this thesis due to the fact that real control data is not available.

k-means clustering was realised by using scikit-learn machine learning library with Python [247]. The initialisation method was 'k-means++' with 20 random initial cluster center positions (k-means was run 20 times with different initial cluster positions), maximum number of iterations was set to 1000, and tolerance for stopping criteria was set to 10^{-4}.

Clustering was executed with multiple different inputs by using data of samples of single patients and combinations of patient samples. Each in-focus reconstruction was segmented using the CNN model that was trained with the MDCK cell line (Chapter 7). Fig. 9.6 shows example holograms, reconstructions and segmentation masks. It is remarkable that the trained CNN model works so well with the patient data. A larger set of

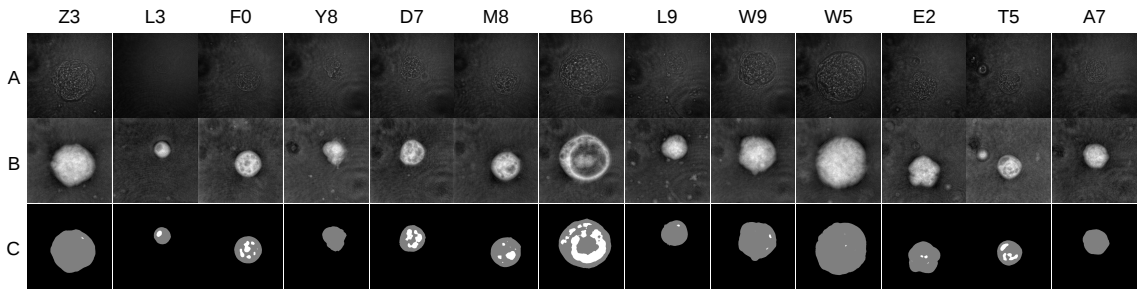

Figure 9.6 Example patient data with segmentation masks showing example in-focus phase reconstructions for each patient. Row A: holograms, row B: in-focus unwrapped phase reconstructions, row C: segmentation mask. Patient IDs shown above each column.

patient data examples is shown in Fig. 9.7. For clustering, features of cysts were extracted (Chapter 4) and extracted feature vectors were used as an input to the *k*-means algorithm.

As many of the (adjacent slice) tissue samples contained both cancer and healthy tissue, it was assumed at the beginning that this is true also for cultured samples. The first clustering experiment was performed by using samples of single patients and organoids of each patient were clustered to two clusters. All the identified problems exist with this approach (Table.. 9.7). The most severe problem here is the 'Clustering consistency problem'; when each patient is independently clustered using only features data specific to this patient, each patient is clustered into two clusters based on different features. Results of this is shown in Fig. 9.8. It can be observed that splits between clusters do not match pathological evaluation for majority of patients. For benign hyperplasia patients, data is split between two clusters. The same trend can be seen with the patients who received 'HIGH' pathological cancer evaluation.

To explore the functionality of *k*-means deeper and gain more understanding of the data, clustering was repeated by setting number of initial clusters to one and repeating clustering 50 times for each patient. Fig. 9.9 shows a boxplot where percentage of samples in smaller clusters are shown. It can be observed that with majority of patients variance between different runs is large. For an easier interpretation the following should be noted:

- patient 'Z0': maximum, median and 75th percentile are the same (35.14)
- patient 'Y8': minimum and 25th percentile are the same (30.1)
- patient 'D7': 75th percentile and maximum are the same (42.86)

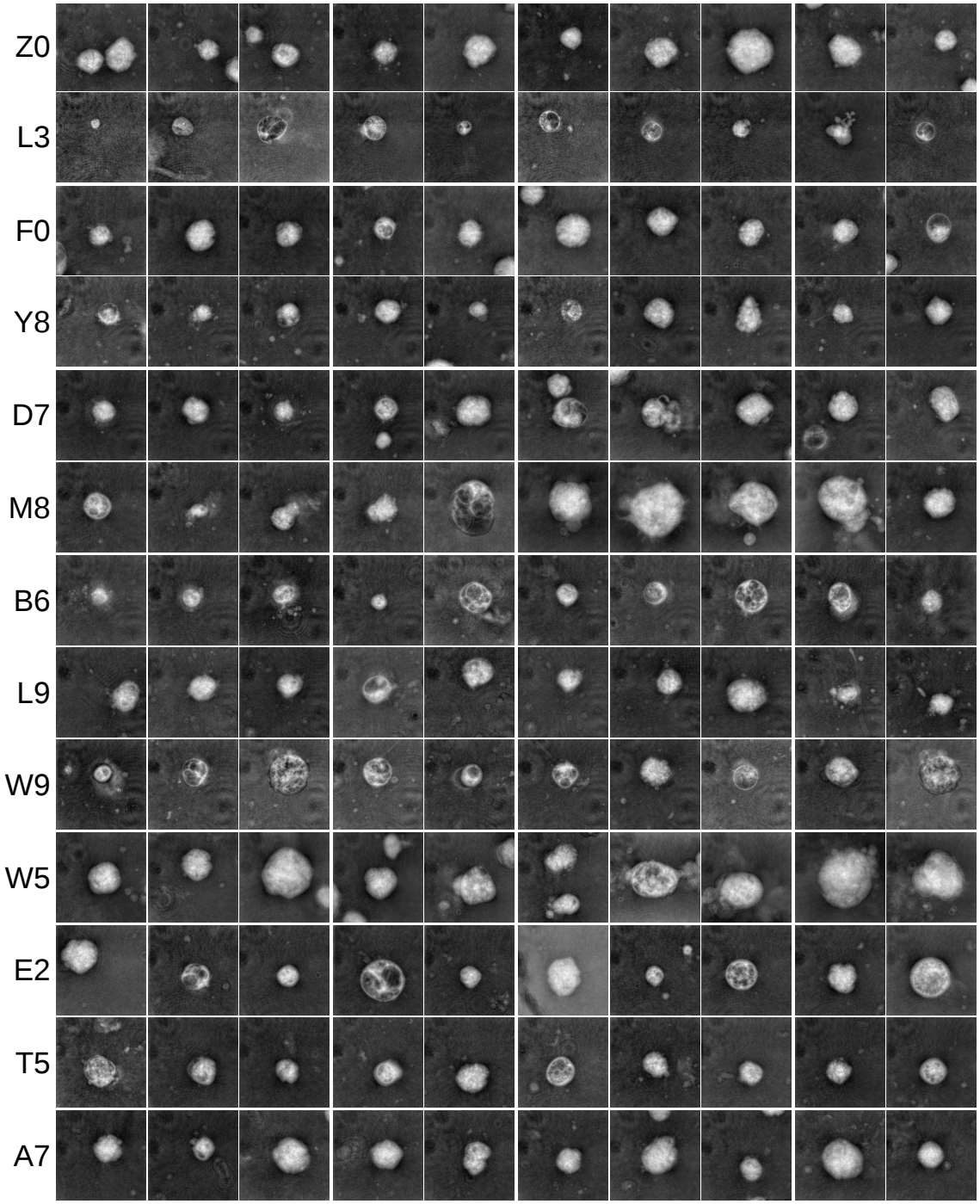

Figure 9.7 Example patient data. Showing example in-focus phase reconstructions for each patient. Patient IDs shown on beginning of each row.

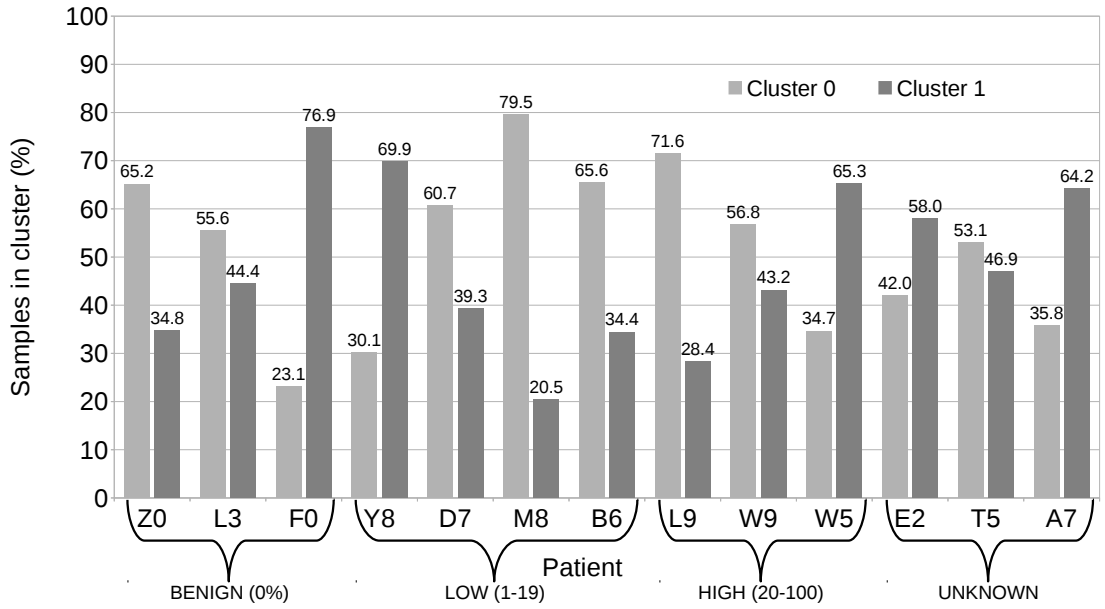

Figure 9.8 Single patient clustering. Each patient is clustered independently into two clusters by using *k*-means clustering.

- patient ‘M8’: 25th percentile, median, 75th percentile and maximum are the same (24.65)
- patient ‘W5’: minimum, 25th percentile, and median are the same (34.67)
- patient ‘A7’: median, 75th percentile and maximum are the same (36.48)

Our approach to the ‘Clustering consistency problem’ was to conduct the next experiment by using multiple patients simultaneously. All the data of all the patients was combined and clustered to two clusters. The result of this is shown in Fig. 9.10. Starting from patient ‘Y8’, a linear increase in one of the clusters can be expected, however, such a trend is not observed. Two patients, ‘W5’ (‘HIGH’ evaluation) and ‘A7’ (‘UNKNOWN’ evaluation) have completely different distributions between two clusters than other patients. The same experiment was repeated with all benign hyperplasia patients (‘Z0’, ‘L3’, and ‘F0’) and patient ‘W5’ with ‘HIGH’ pathological cancer evaluation. This was performed to see if distributions of the number of samples in two clusters are different. The result of this is shown in Fig. 9.11. It can be seen that for benign hyperplasia patients, the large majority of data is in cluster 0. The split for patient ‘W5’ is 58/42 for clusters 0 and 1, respectively. This is completely different from benign hyperplasia patients, however not as obvious when using all the patients where majority of data was in opposite cluster if

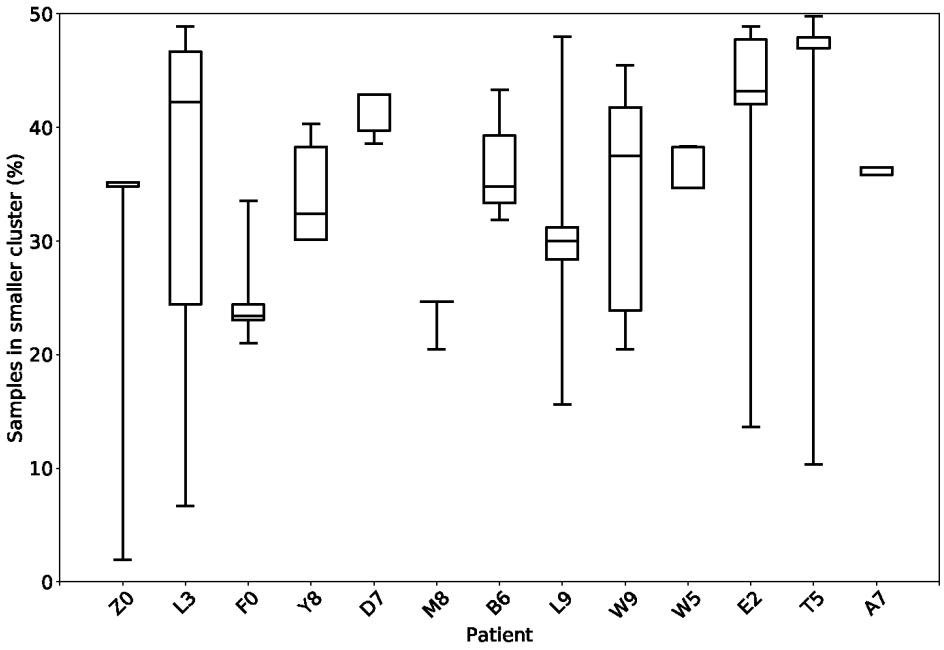

Figure 9.9 Boxplot of single patient clustering. Each patient is clustered independently to two clusters by using *k*-means clustering. Clustering is repeated 50 times with random initial cluster centers. Showing proportion of samples in a smaller cluster.

compared to benign patients. Although this approach bypasses the 'Clustering consistency problem' other problems still exists.

Our approach to the 'Cluster relevance problem' is to assume that data (Fig. 9.11) in Cluster 0 is related to benign cases and data in cluster 1 is related to cancer. In accordance with the assumption, *k*-means was "trained" by using benign hyperplasia patients and patient 'W5'. The resulting cluster centres were used to predict clusters for other patients not used in the "training". The result of this is shown in Fig. 9.12. A large majority of data

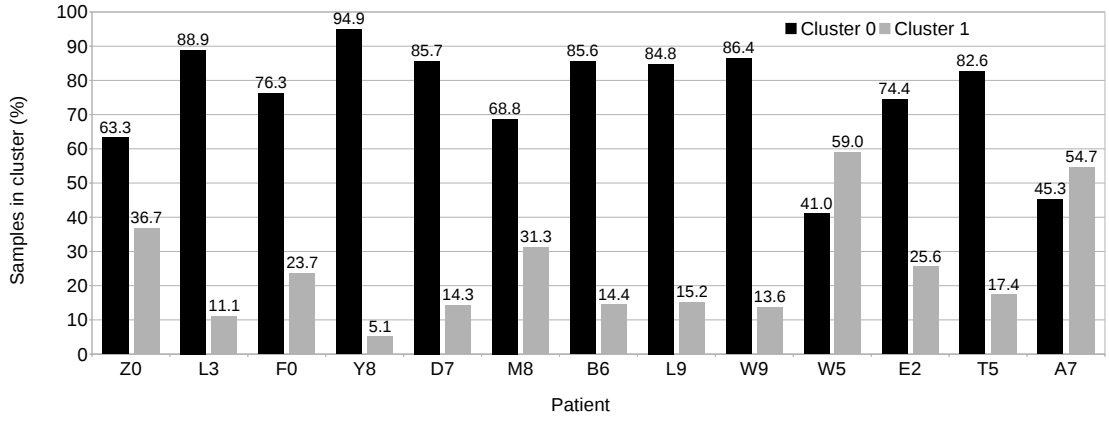

Figure 9.10 Data of all patients grouped together and clustered to two clusters.

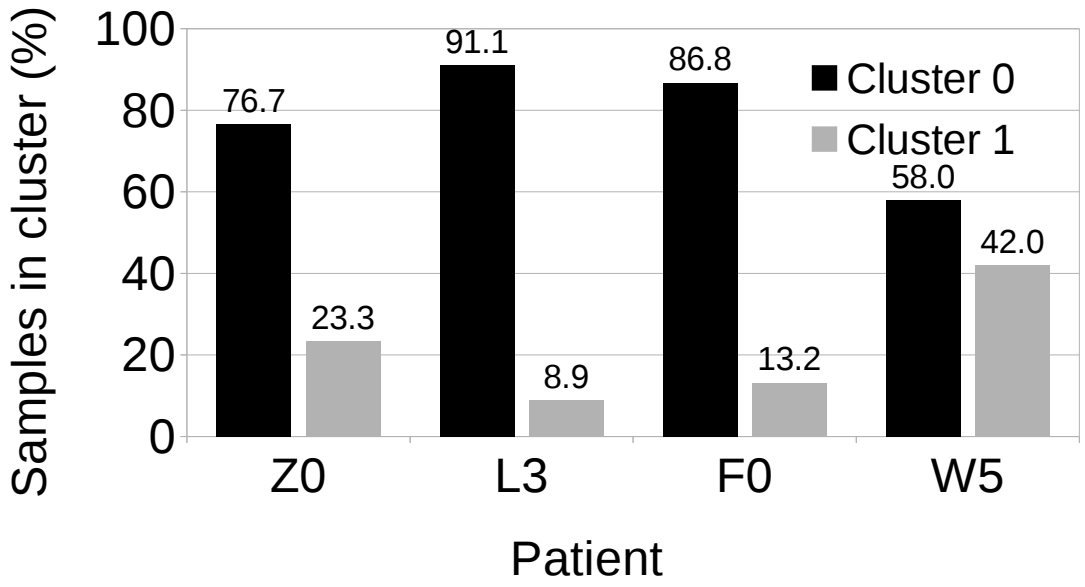

Figure 9.11 Benign hyperplasia patients and one 'HIGH' pathological cancer evaluation patient ('W5'). Clustered to two clusters.

is in cluster 0. There is no increase in number of samples in cluster 1 (cancer cluster) as could be expected based on cancer evaluation.

To investigate why *k*-means is not working (i.e. cancer proportions are not in agreement with the evaluation of a pathologist) in accordance with our tentative assumption, we approach the 'Cluster indexing problem' by using principal component analysis to reduce feature dimensionality, allowing for easier visual inspection of the data. With the human prostate cell line (RWPE), the lumen index was the most representative single feature (Chapter 4) between wild-type and oncogenic samples. To study if this is true also for patient samples, fig.9.13 shows the lumen indices of all patient organoids. 68 % of the organoids have a lumen index below 0.1. By calculating a single principal component on other features besides lumen index, plotting fitted data on x and lumen index on y axis, it can be observed that the data is largely overlapping as shown in Fig. 9.14. These two features alone do not form visually observable clusters.

A single principal component explains only 39.4 % of the original data variance (Fig. 9.15). By increasing number of components to three, 77.9 % of the original data variance is explained. Fig. 9.16 shows PCA fitted data by using all the patient data and three principal components. It can be observed that the data is spread, mixed and largely overlapping. However, possibly at some locations the data is concentrated into clusters that

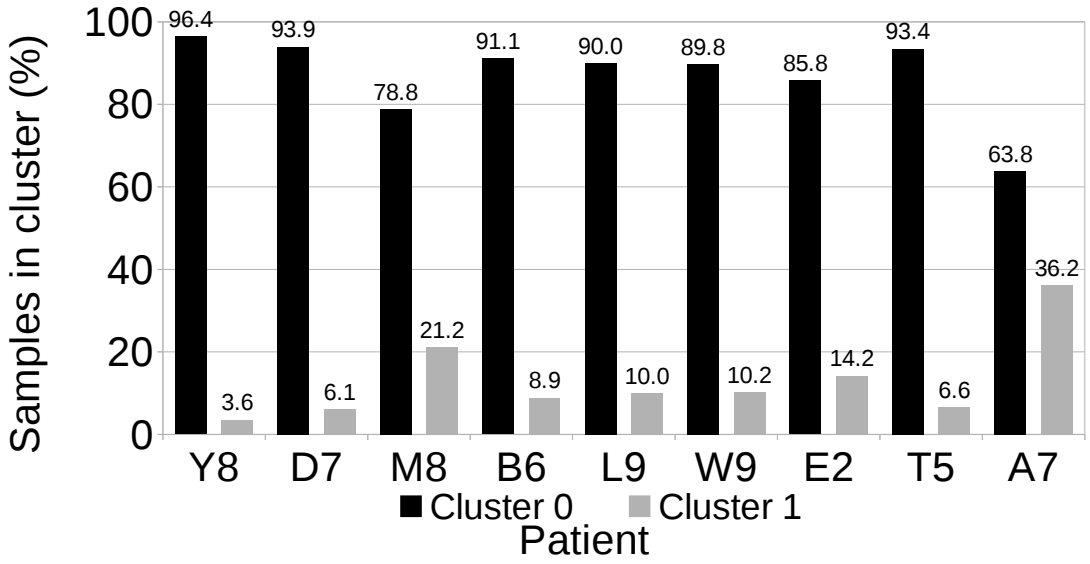

Figure 9.12 Patient data clustering by *k*-means with fixed cluster centres.

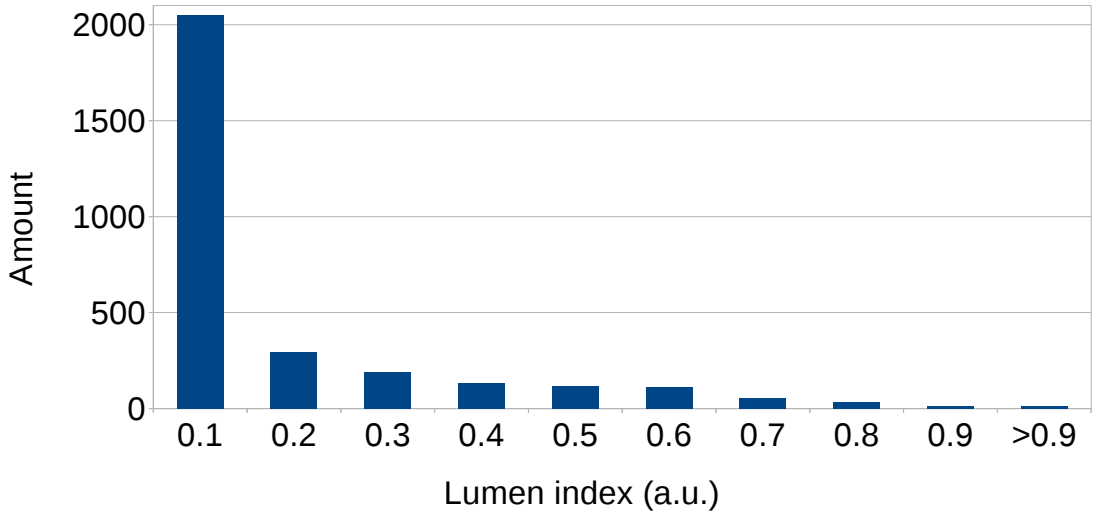

Figure 9.13 Lumen indices of all patient organoids. 68.1 % of 3013 organoids have lumen index below 0.1.

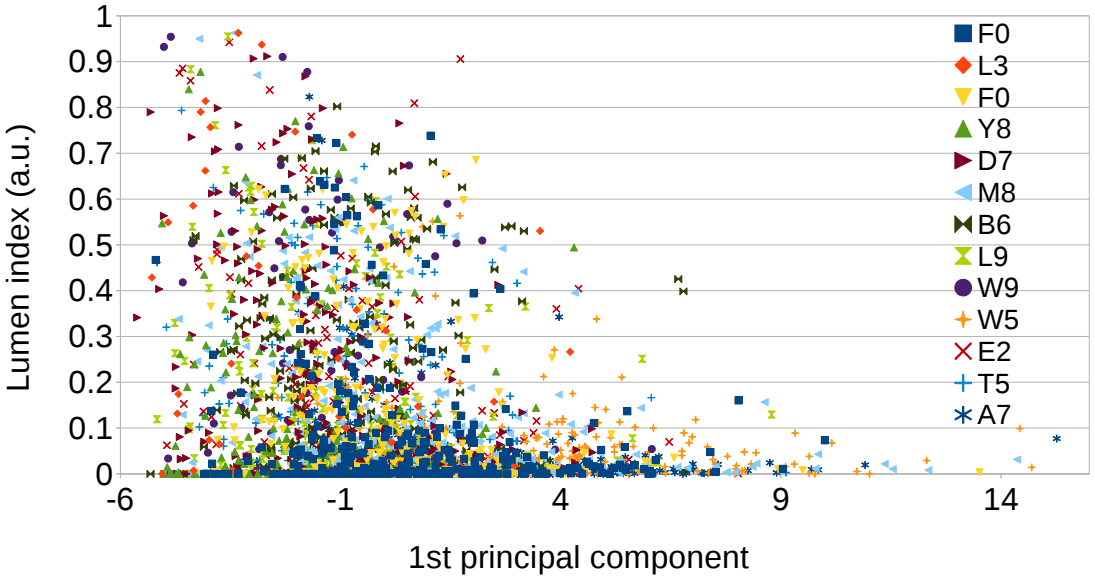

Figure 9.14 Lumen indexes and PCA fitted data of all patient organoids.

do not display clear borders. Table 9.8 shows component loadings of features for different principal components. It can be seen that depending on the principal component, some features are more important than others. For example, for the 1st principal component, axis lengths, area, perimeter, convex area, equivalent diameter, minimum phase, and maximum phase are more important than other features (absolute value >0.75). For the other two components, the loadings are more balanced.

Applying *k*-means on reduced data by using the same settings as the previous experiment verifies the previous observation. Fig 9.17 shows two cluster centers for each patient by using fitted data. It can be observed that the region highlighted with the red oval contains one cluster center for 12 out of 13 patients. Data points falling in this cluster all have something in common, however it is not directly related to a cancer evaluation performed by a pathologist. The same can be observed in the Fig. 9.18 where cluster centers are clustered by using *k*-means algorithm (2nd order clustering). The data is clustered into four clusters and most of the data is located in the cluster that had 12 out of 13 patients.

The majority of patient samples are in the cluster 0, it contains 12 out of 13 patients as shown in Fig. 9.19. Figures 9.20-9.23 show example organoids from four different clusters shown in Fig. 9.18. All 13 patients had organoids in all four clusters.

Table 9.8 Patient data component loadings. Showing component loadings of different features by using three principal components. Absolute values ≥ 0.75 are highlighted with red colour.

Feature	PC1	PC2	PC3
Major axis length	0.93	0.30	0.06
Minor axis length	0.96	0.05	-0.25
Area	0.96	0.17	-0.14
Eccentricity	-0.01	0.49	0.68
Average	0.42	-0.64	0.37
Roundness	0.01	-0.59	-0.68
Perimeter	0.88	0.43	0.01
Convex area	0.95	0.21	-0.11
Equivalent diameter	0.98	0.15	-0.12
Extent	0.02	-0.64	-0.49
Min phase	-0.81	-0.04	0.02
Max phase	0.61	-0.52	0.33
Solidity	0.12	-0.76	-0.46
Skewness	-0.43	0.32	-0.06
Kurtosis	-0.05	0.56	-0.52
Roughness	-0.10	0.75	0.22
Standard deviation	0.67	-0.56	0.35
Density	-0.41	-0.64	0.43
Lumen index	-0.33	0.59	-0.51

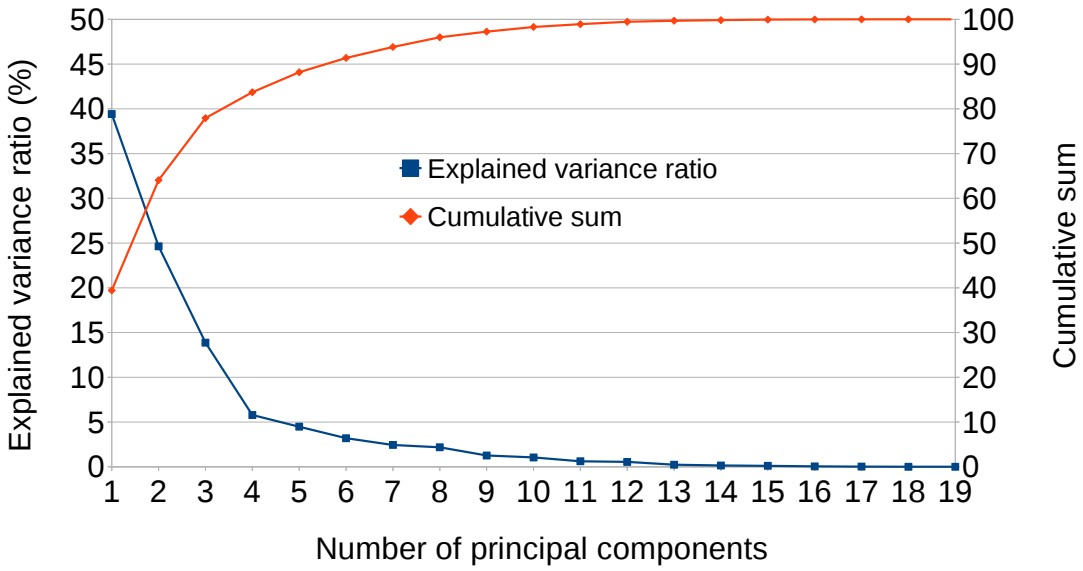

Figure 9.15 Explained variance ratio. Showing how much of the original data variance is explained by each principal component.

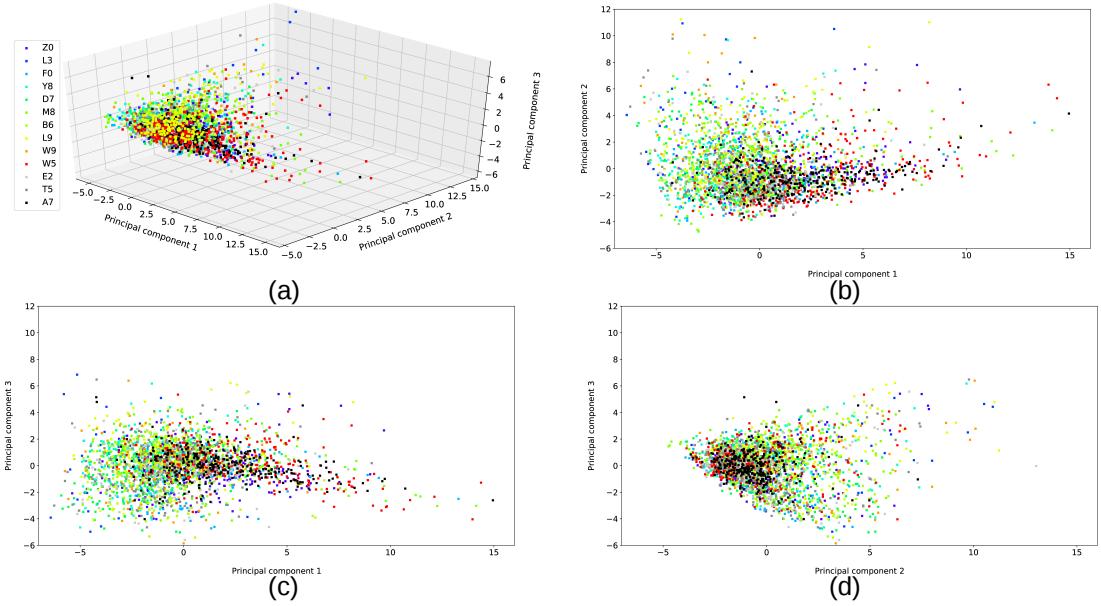

Figure 9.16 Fitted patient data with three principal components. (a) 3D plot showing three axes, (b)-(d) showing two out of three axes, one at the time. Data for each patient has a unique colour.

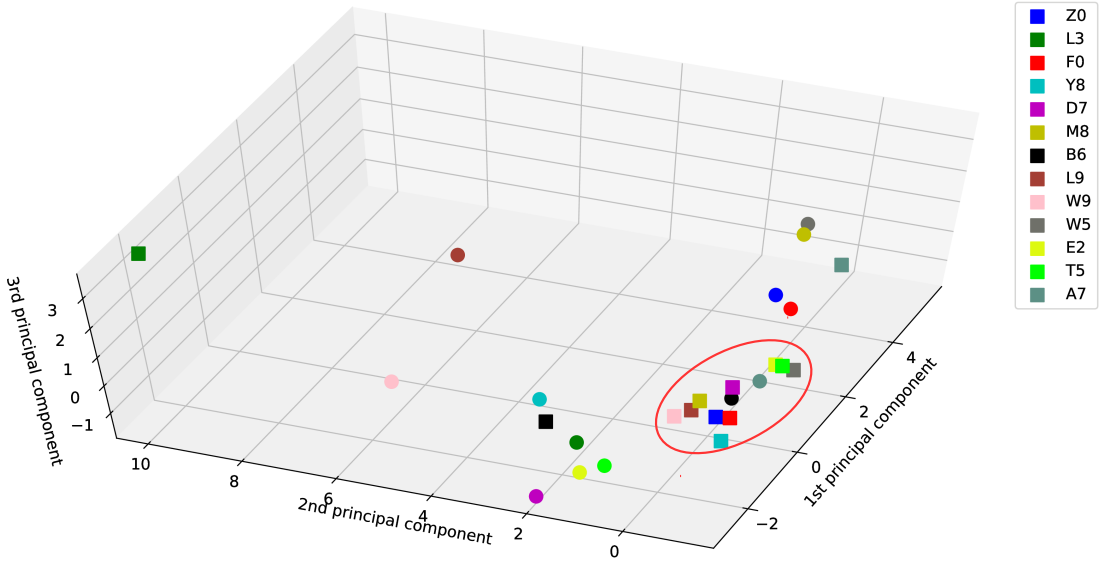

Figure 9.17 Clustering cluster centers. Each pair formed of a square and a circle with the same colour are two cluster centers of a single patient. The red oval contains one of the cluster centers that had 12 out of 13 patients.

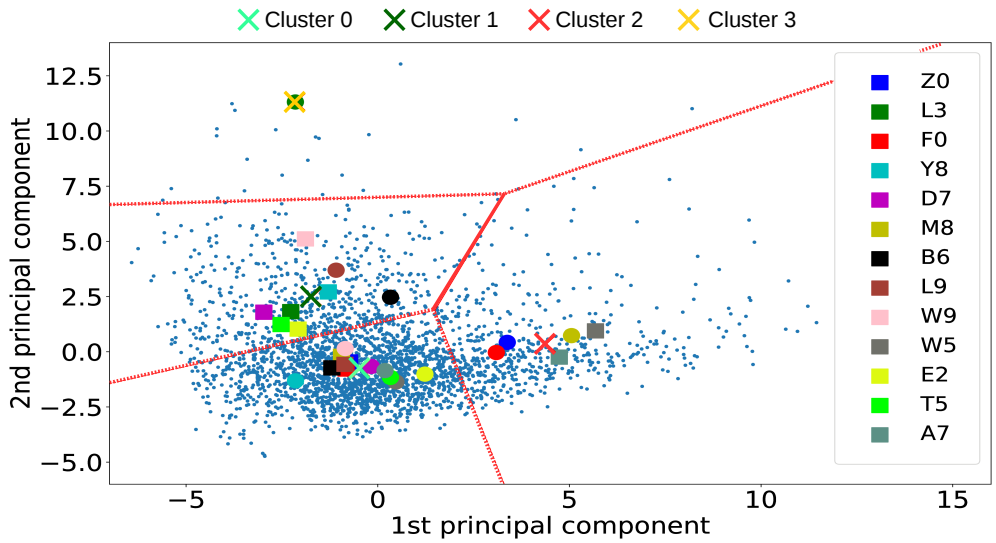

Figure 9.18 Voronoi diagram. Showing two cluster centers for each patient with circles and squares. Crosses show the location of cluster centers of clusterered cluster centers that were used to define Voronoi cells. Light blue dots show fitted data by using two principal components of the full dataset.

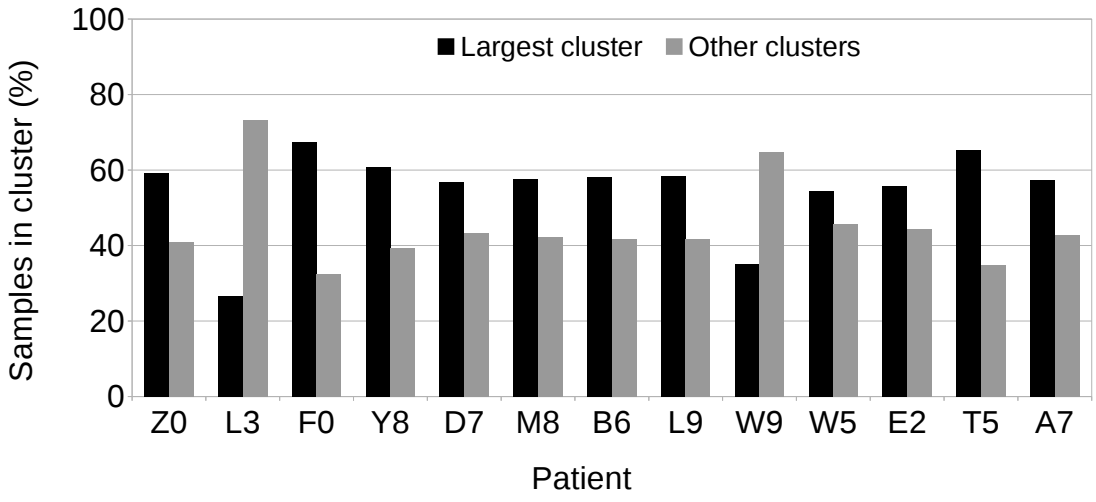

Figure 9.19 Patient data clustered. Showing proportion of samples in the cluster 0 that had cluster centers for 12 out of 13 patients (Largest cluster) and other three clusters 1-3 summed (Other clusters).

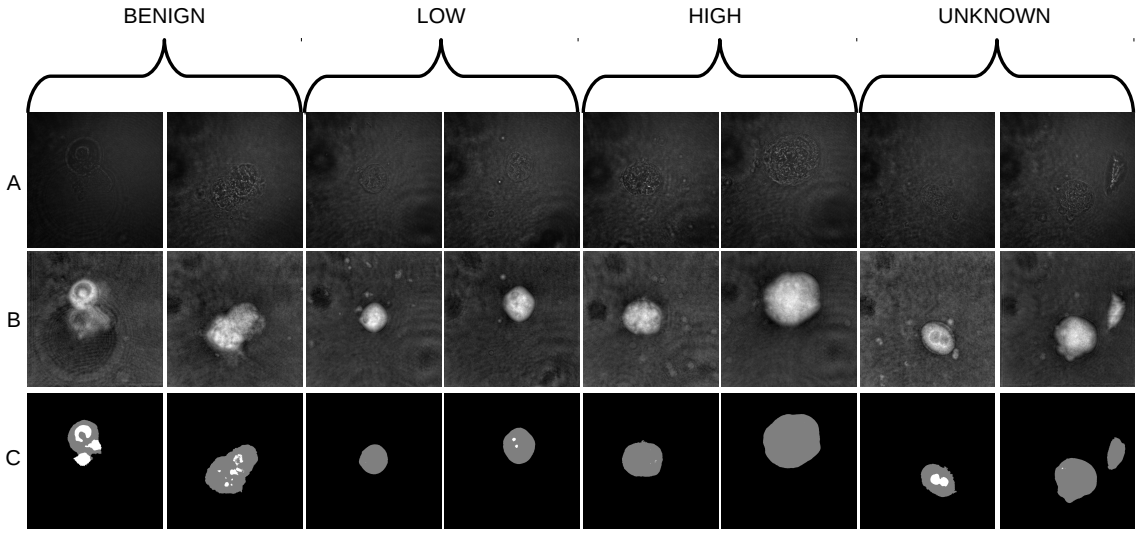

Figure 9.20 Example organoids assigned to cluster 0. Row A: hologram, row B: in-focus unwrapped phase reconstruction, row C: mask with maximum of three values for background, non-hollow 3D region, and lumen.

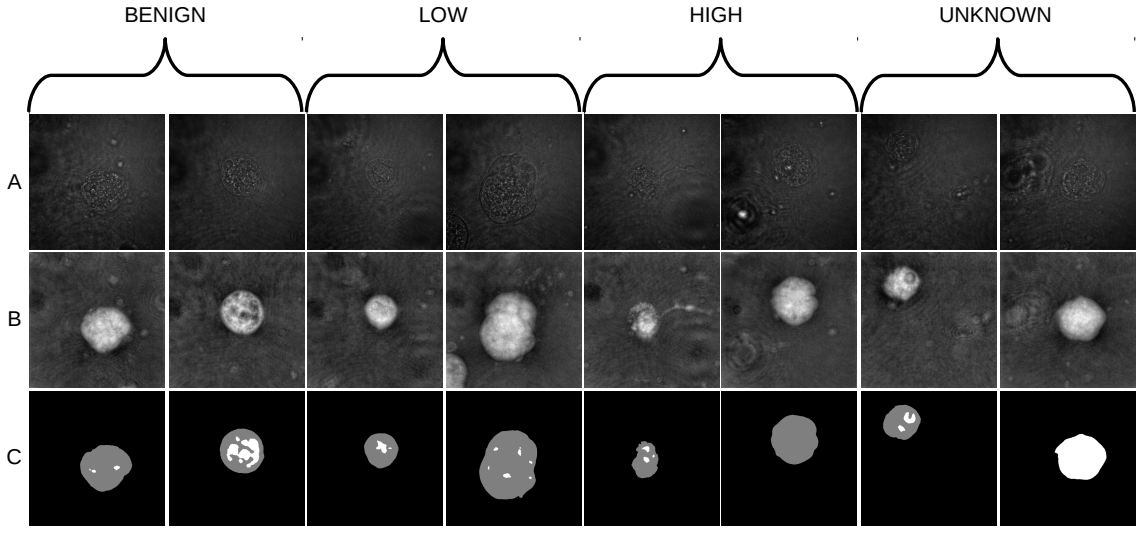

Figure 9.21 Example organoids assigned to cluster 1. Row A: hologram, row B: in-focus unwrapped phase reconstruction, row C: mask with maximum of three values for background, non-hollow 3D region, and lumen.

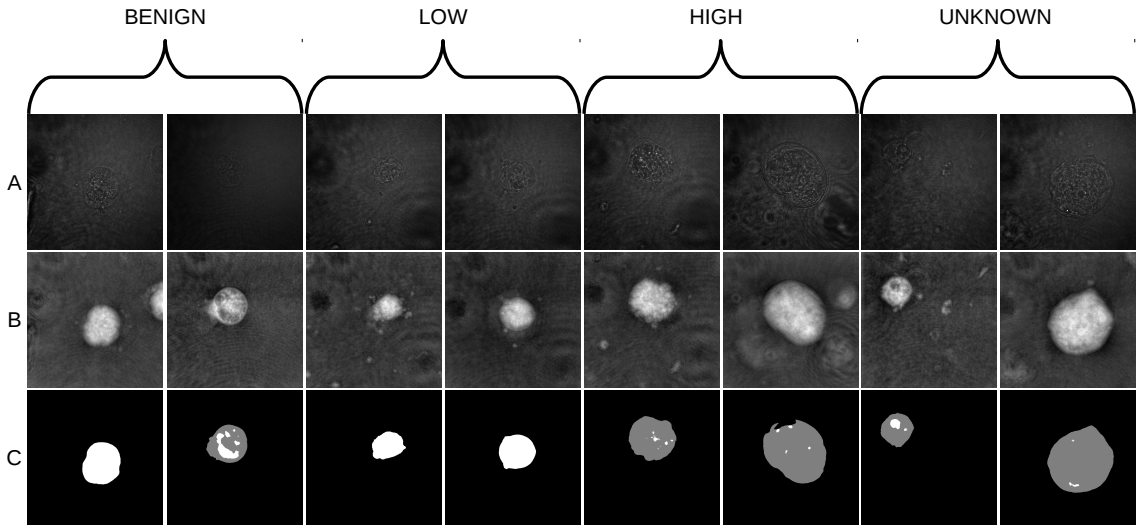

Figure 9.22 Example organoids assigned to cluster 2. Row A: hologram, row B: in-focus unwrapped phase reconstruction, row C: mask with maximum of three values for background, non-hollow 3D region, and lumen.

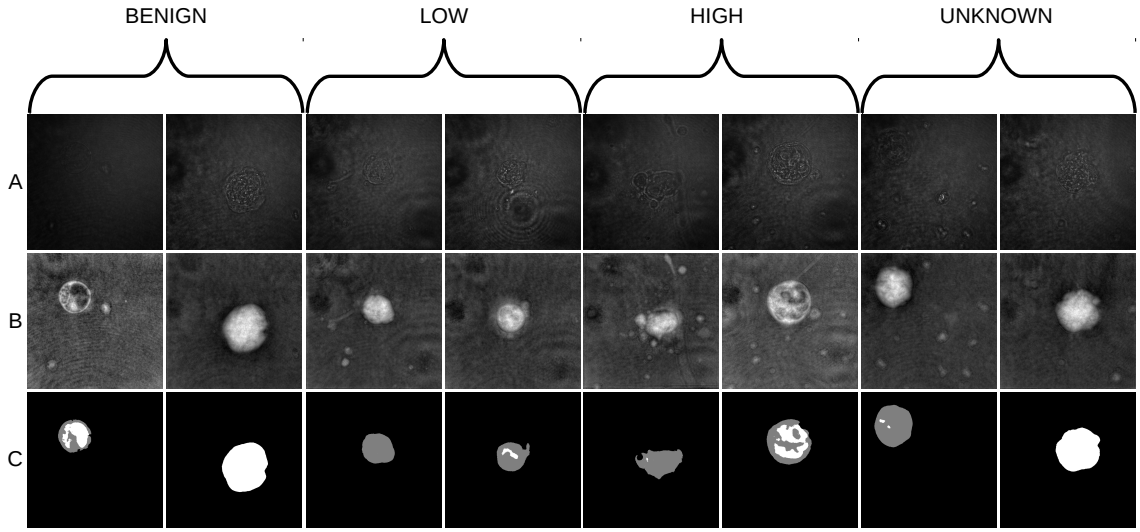

Figure 9.23 Example organoids assigned to cluster 3. Row A: hologram, row B: in-focus unwrapped phase reconstruction, row C: mask with maximum of three values for background, non-hollow 3D region, and lumen.

Table 9.9 SVM and MLP training and testing. Showing the experiment number, data used for training and testing.

Experiment	Training data	Testing data
1	Data of benign hyperplasia patients and one patient ('W5') with 'HIGH' pathological evaluation	Other patients not used in training
2	RWPE cell line (control and $KRas^{V12}$)	All the patients

9.3.3 Non-linear classification

To investigate if non-linear approaches are able to identify cancerous organoids, we trained MLP and CNN classifiers. In addition, for comparison, we trained an SVM classifier with a radial basis function kernel that performed well with the MDCK data as reported in Chapter 4. Training and testing data strategies for SVM and MLP classifiers are shown in Table 9.9.

Multiple SVM and MLP classifiers were trained and all the results reported here are averages of those runs. For both SVM experiments (Table 9.9), 1000 classifiers were trained for each. For both experiments 1 and 2, 300 MLP classifiers were trained for each. RWPE cell line was included as it is a human prostate cell line. In principal, patient derived organoids should resemble these.

The support vector machine classifiers were trained with the radial basis function kernel, the penalty parameter was 1, the kernel coefficient (γ) was 0.1, the shrinking heuristic was used, and the stopping criterion was set to 0.001.

The result of experiment 1 (Table 9.9) is shown in Fig. 9.24. The majority of the samples are in 'Benign' class as expected. However, there is no increase in 'Cancer' class from 'LOW' to 'HIGH' patients as could be expected.

The result of experiment 2 (Table 9.9) is shown in Fig. 9.25. Interestingly, based on this result, a large majority of patient samples resemble $KRas^{V12}$ organoids.

The same two experiments were repeated using MLP classifier. For experiments 1 and 2 (Table 9.9), 20 % of data was extracted for validation data. The network had a single hidden layer with 100 neurons. ReLu was used as an activation function, the regularisation parameter was 0.001, the maximum number of iterations was 10000, and the tolerance for the optimisation was 1×10^{-6}. The optimiser was limited-memory BFGS [250].

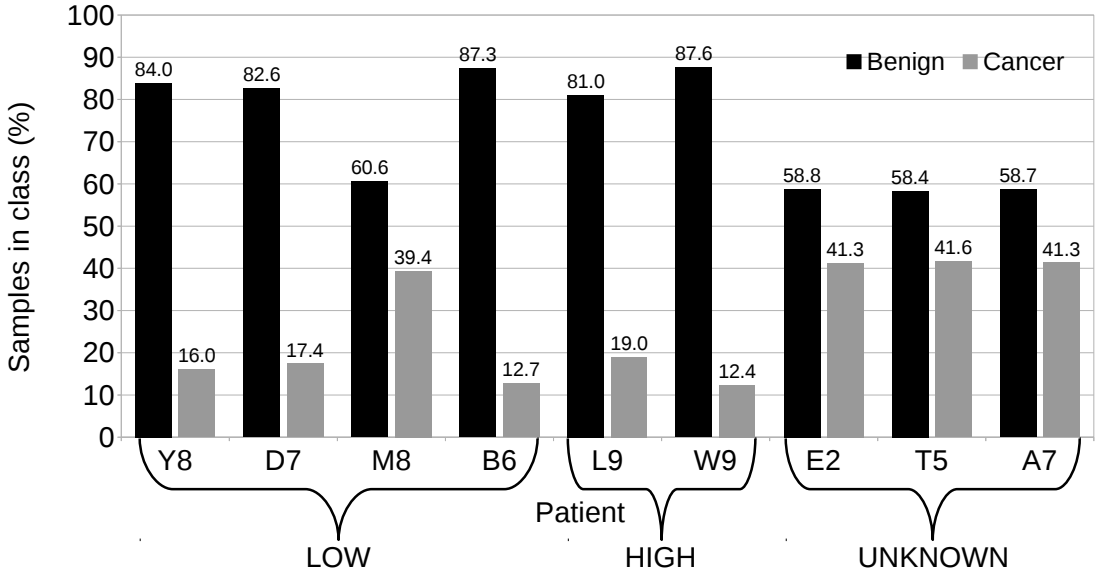

Figure 9.24 SVM result with 'BENIGN' patients and a single 'HIGH' cancer patient ('W5') in training. Predicted class (Benign or Cancer) for other patients, not used in training, was predicted and distributions are shown here.

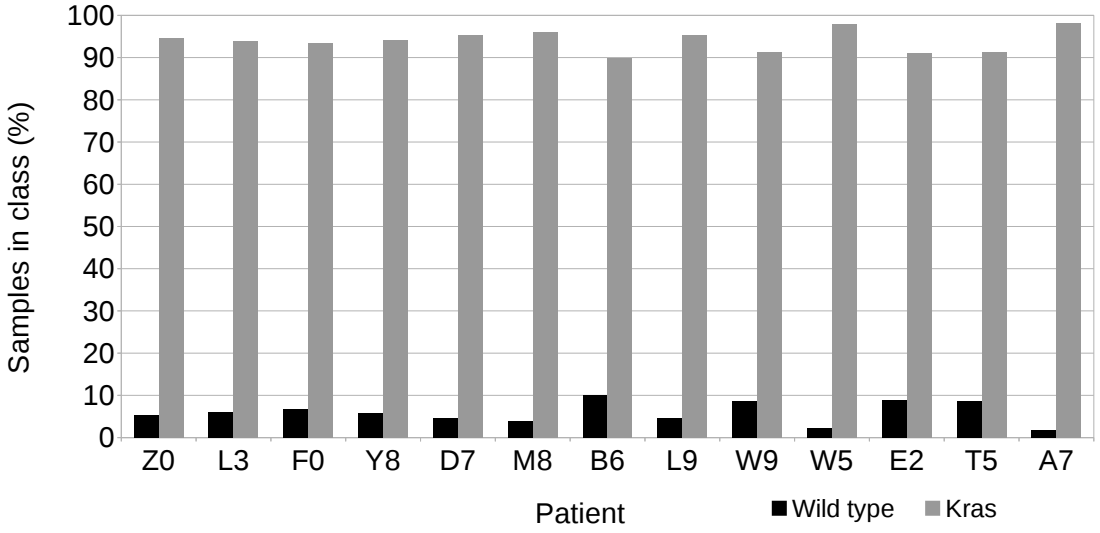

Figure 9.25 SVM result using RWPE data in training. Class (Wild type or $KRas^{V12}$) for all the patients was predicted and distributions are shown here.

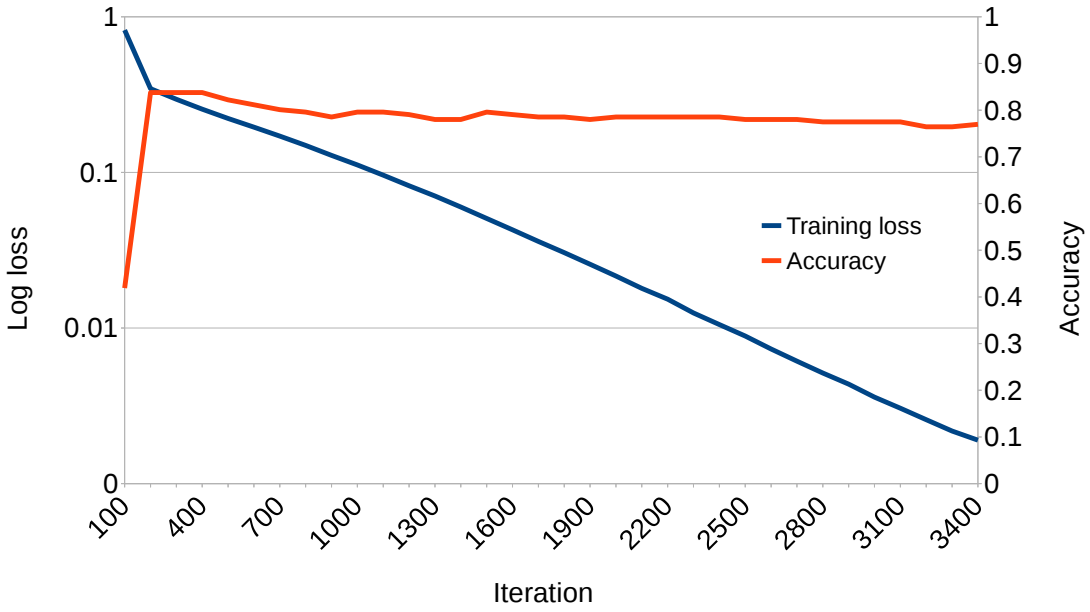

Figure 9.26 MLP loss and accuracy with 'BENIGN' patients and a single 'HIGH' patient in training.

Fig. 9.26 show training loss and validation accuracy for one of the classifiers. The training loss is reduced and converges, while accuracy remains approximately at the same level after 200 iterations. The classification of experiment 1 (Table 9.9) is shown in Fig. 9.27. There is no obvious increment in sample proportions for the Cancer class from 'LOW' to 'HIGH' patients. The split between two classes is more balanced than with the SVM classifier.

Training loss and validation accuracy for the experiment 2 (Table 9.9) of one trained classifier are shown in Fig. 9.28. After 1000 iterations the validation accuracy remains at the same level, and the training loss converges. The classification result is shown in Fig. 9.29. According this, the majority of patient samples resemble $KRas^{V12}$ samples. Three exceptions exist: 'HIGH' patient 'W5' and two 'UNKNOWN' patients 'E2' and 'A7'.

For the CNN classifier, samples from benign hyperplasia patients formed class 0 data, and samples from patient 'W5' formed class 1 data. Class 0 contained 578 samples of which 29 (5.0 %) were extracted for validation data, 20 (3.5 %) as test data, and 500 (86.5 %) samples were used for training data. Class 1 contained 304 samples that were split into training (255, 83.9 %), validation (29, 9.5 %) and test data (20, 6.6 %).

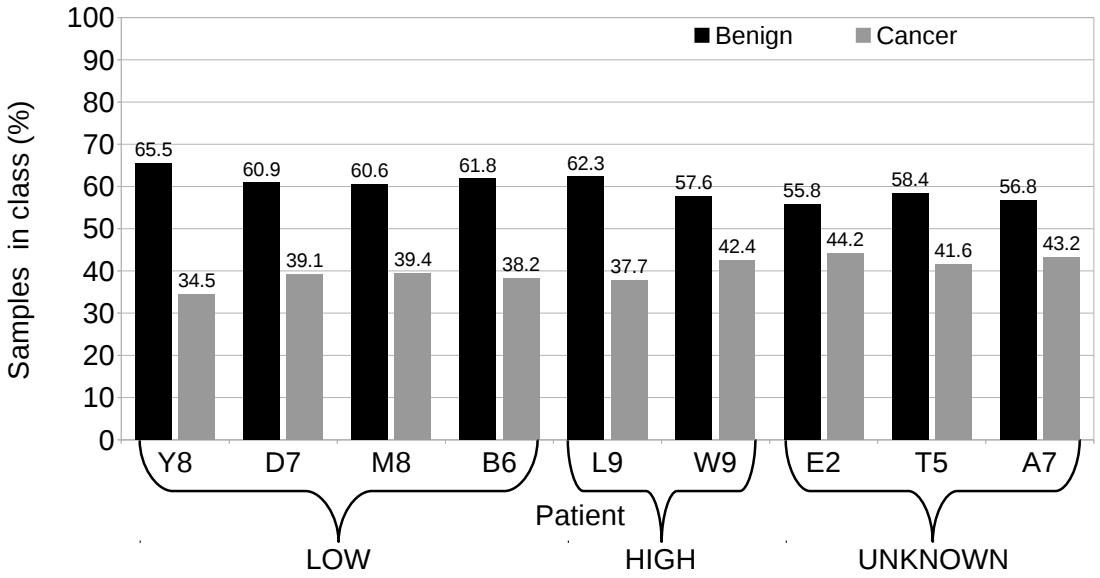

Figure 9.27 MLP result with benign patients and a single high cancer patient in training. Predicted class (Benign or Cancer) for other patients, not used in training, was predicted and the distributions are shown here.

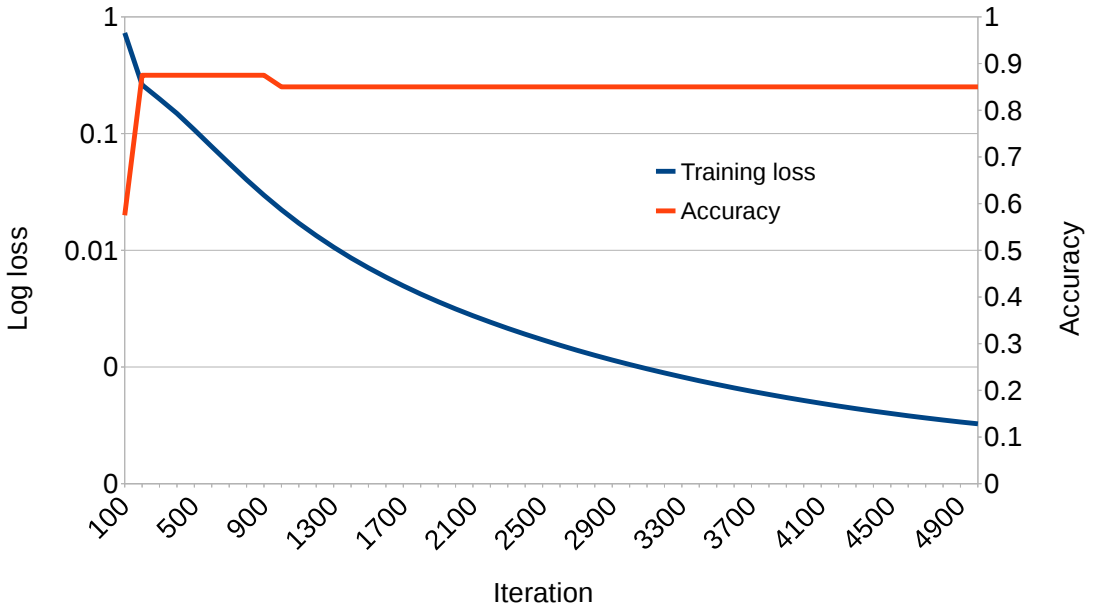

Figure 9.28 MLP loss and accuracy with RWPE training data.

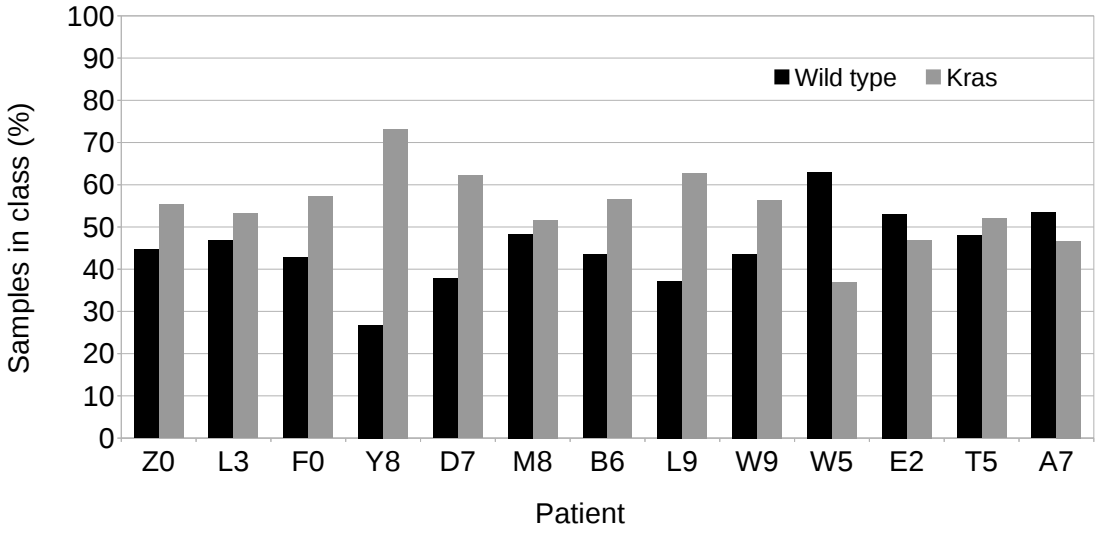

Figure 9.29 MLP result using RWPE data in training. Class (Wild type or Kras) for all the patients was predicted and distributions are shown here.

The holograms were reconstructed to the in-focus plane, phase was unwrapped, segmented and rescaled to 227×227 pixels as required by the network architecture. Each input was randomly augmented during run-time by a random combination of random rotation by 90 degrees, horizontal mirroring, and vertical mirroring. The network architecture was AlexNet (see Chapter 4 for details).

The network was trained for 100 epochs with batch size of 50 on two Nvidia GTX 1080 graphics cards using the PyTorch [248] deep learning framework. Adam was used as the optimiser [254], with a weight decay of 1×10^{-3}, and cross entropy loss was used. Due to the unbalanced datasets, weights of 1/500 (class 0) and 1/255 (class 1) for the loss function were used. An adaptive learning rate was used with an initial learning rate of 1×10^{-4} and was reduced by a factor of 10 if the validation loss failed to improve for ten consecutive epochs. Three networks were trained with the above settings.

Loss and accuracy curves are shown in Fig. 9.30. Loss is not converging or convergence is insignificant. The validation accuracy remains approximately at the same level (60 %) after 56 epochs. The average test classification accuracy of the three trained networks was 64.2 %. Other patients, not used in training, were classified next. Classification averages of the three networks are shown in Fig. 9.31. Based on these results, it is possible that the number of holograms used to train a CNN classifier was not sufficient. Due to the limited amount of training data, the CNN was not able to learn distinguishing features.

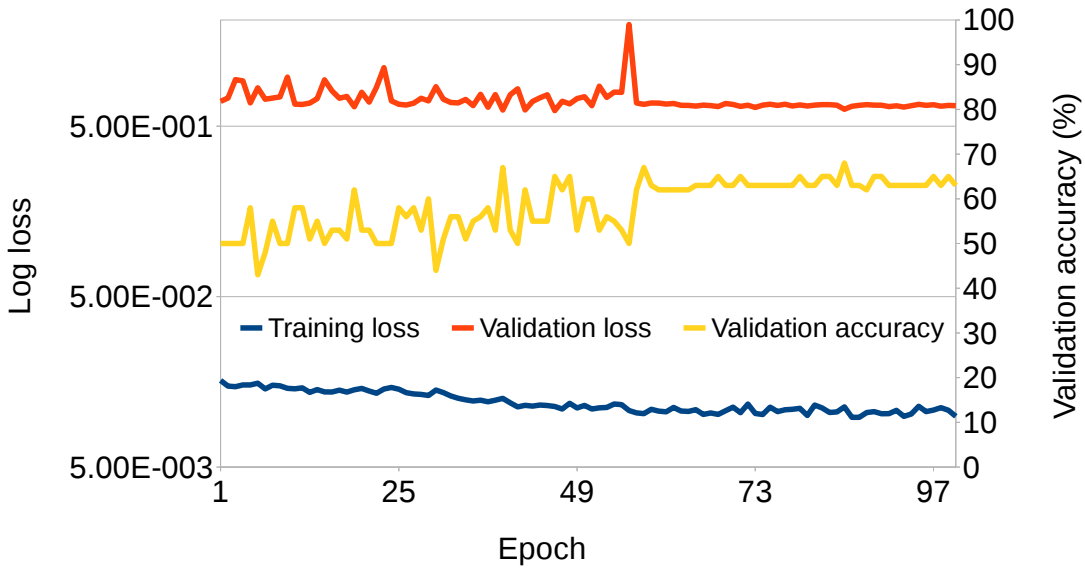

Figure 9.30 CNN loss and accuracy. Showing training and validation loss and validation accuracy.

Another explanation may be that samples in these two classes do not pose discriminating (convolutional) features and therefore the classifier fails.

9.3.4 Patient identification

The experiments in this section were conducted to study if patients can be distinguished from each other based on extracted features from the holograms only.

k-means clustering was applied to different patient combinations, where all possible two-patient-combinations were tested. The results of this is shown in Fig. 9.32. It can be observed that almost in all cases, the same cluster, 0 or 1, with both patients contains the majority of data and therefore in a majority of cases a patient can not be identified from the extracted features. However, in total 12 exceptions were found where two different clusters (0 or 1) contained more than 50 % of samples. Of these 12 cases, patients ‘L3’ (4 times), ‘Y8’ (3 times), and ‘D7’ (3 times) contribute to the 10 cases. The exceptional 12 patient combinations are marked with red squares in Fig. 9.32. The numerical values of samples in larger clusters for 12 are shown in Table 9.10.

For MLP and SVM classifiers, 20 % of the data of each patient was extracted for the testing data. The rest of the data was combined and used for training these classifiers. The

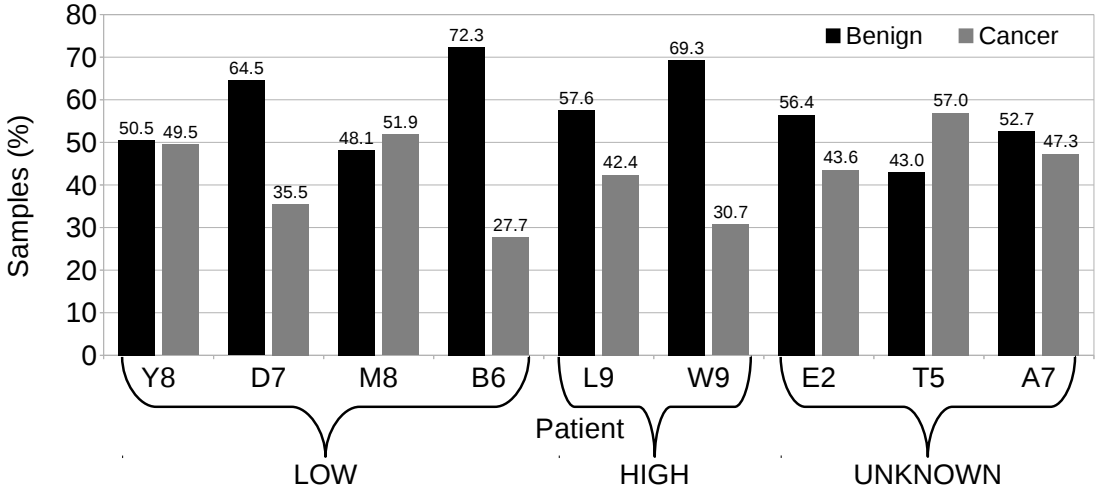

Figure 9.31 CNN classification test. Showing classification result with patient samples that were not used in training. Below the patients data is the cancer proportion evaluation of an adjacent slice given by a pathologist ('LOW', 'HIGH', 'UNKNOWN').

Table 9.10 Exceptions in k-means clustering. Showing numerical values for 12 patients where two different clusters contained more than 50 % of samples.

Patient combination	Larger cluster 1	Larger cluster 2
'Z0-B6'	51.4	73.6
'L3-Y8'	51.1	73.0
'L3-W9'	64.4	54.6
'L3-E2'	77.8	58.0
'L3-T5'	64.4	70.0
'Y8-W9'	73.9	52.3
'Y8-E2'	72.5	50.6
'Y8-A7'	95.4	57.3
'D7-M8'	74.6	56.5
'D7-E2'	61.8	51.1
'D7-A7'	83.9	62.9
'L9-A7'	86.8	51.8

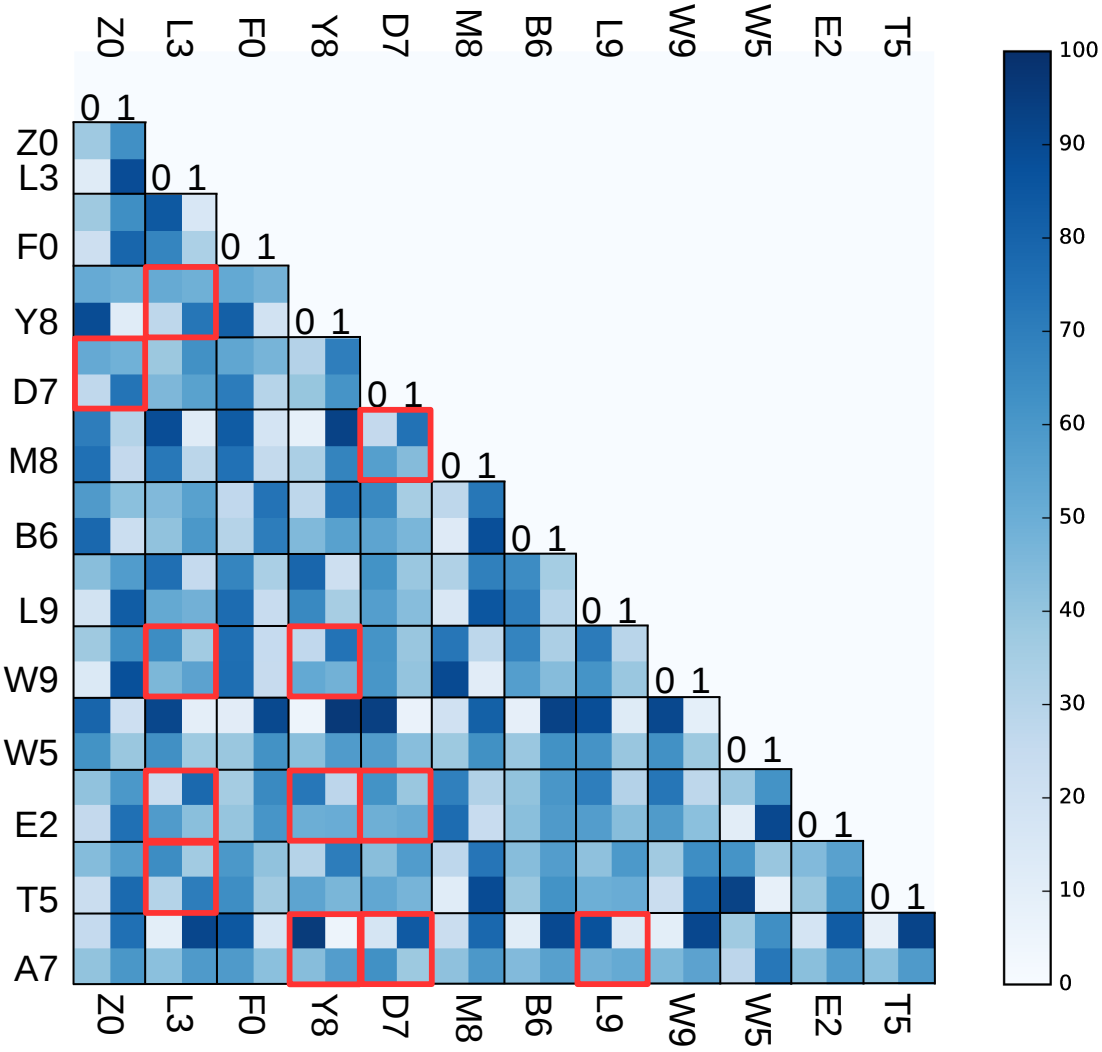

Figure 9.32 *k*-means clustering with all combinations of two patients. Each combination was clustered into two clusters in order to investigate if *k*-means clustering can be used to distinguish any two patients. Showing percentages of samples in clusters 0 and 1. Each group of 4 forms a result for one patient combination, first row above the patient id (on left) corresponds to a patient id on top. For clarity: first patient id 'Z0' is marked above patient id 'L3'; the group of 4 values on their right forms data for this patient combination 'Z0-L3'. Red squares denote patient-combinations where two different clusters contain more than 50 % of samples.

same architecture for MLP and the same settings for both classifiers as in the previous section were used.

Figure 9.33 shows the classification test result for each patient. Each row is normalised by number of samples for a particular patient and sums up to 100. For interpretation: each value on the row shows how many times a patient shown on a row label was predicted correctly (the same patient) or incorrectly (a different patient). Correct classification is shown on the diagonal. For example, patient 'Z0' was correctly classified 49 % (position $[0,0]$) of times, and as a patient 'B6' (position $[0,6]$) 10 % of the times. This figure shows that 5 out of 13 patients have highest percentage for a correct classification. Although, the predictions are distributed between multiple patients, the classification accuracy with the majority of patients is better than a random guess of $\frac{1}{13} \approx 8\%$. The weighted average classification accuracy with this classifier is 22.9 %. Figure 9.34 shows row in Fig. 9.33 as a histogram.

For Fig. 9.35, the underlining values to create Fig. 9.33 were normalised along columns instead of rows as in. For interpretation: each column sums up to 100 and each value on a column shows how many times a patient shown on a column label was predicted as the patient shown on each row label. Correct classification is shown on the diagonal. For example, patient 'Z0' was predicted correctly 28 % (position $[0,0]$) of all of the 'Z0' classification, this same patient was predicted as patient 'F0' 6 % ([position $[2,0]$]) of the time.

Given N samples from a single patient j, the probability $p(i \mid j)$ is the probability that the classifier predicts patient i conditioned on the sample being from patient j. Therefore, the probability of samples being from a patient is given by

$$p(i_1,\ldots,i_N \mid j) = \prod_{k=1}^{13} p(i_k \mid j)^{n_{i_k}} \tag{9.4}$$

where n_i is number of the N samples that are classified as the patient i. By calculating maximum a posteriori (MAP) from the classifier results we can get the best estimate for the patient as

$$j^*(i_1,\ldots,i_N) = \underset{j\in 1,\ldots,13}{\operatorname{argmax}}\, p(i_1,\ldots,i_N \mid j) \tag{9.5}$$

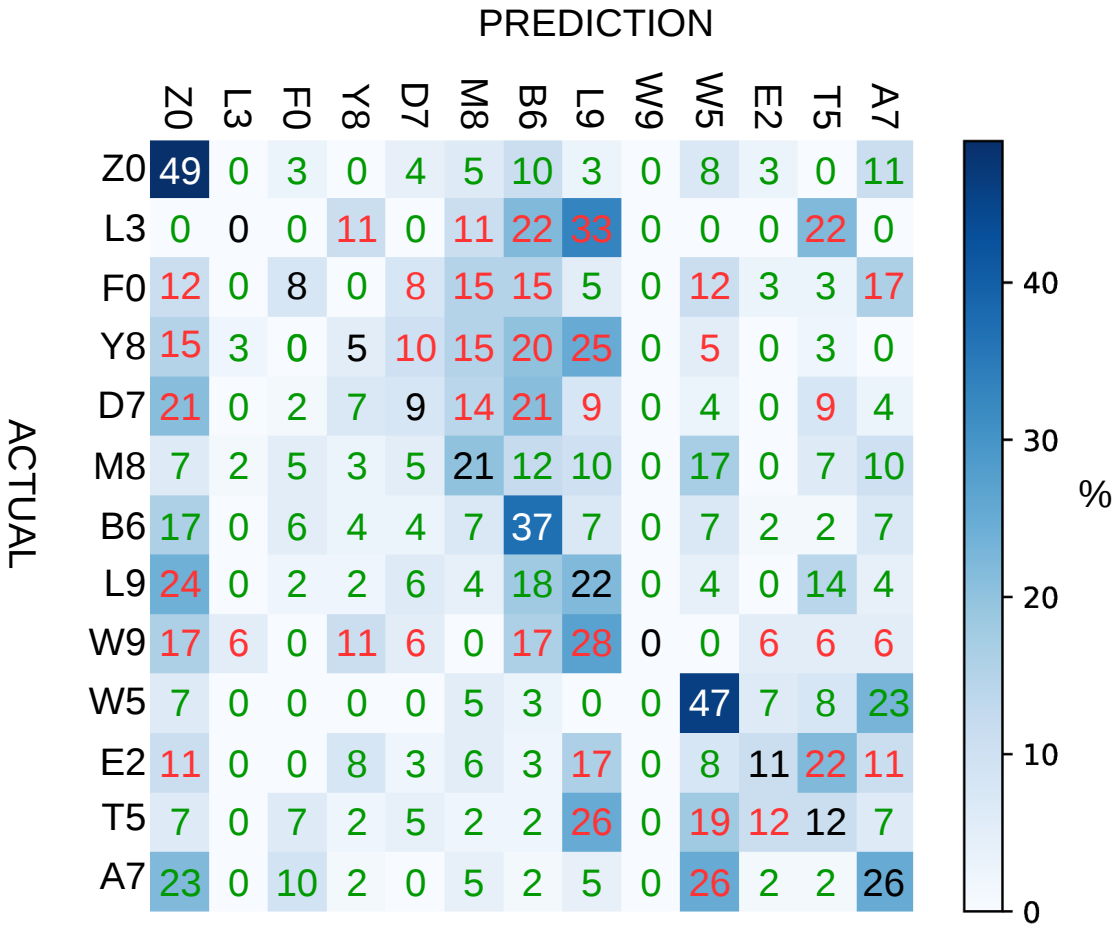

Figure 9.33 MLP result with all patient data. Diagonal numbers show classification accuracy (i.e. correct patient predicted). The red numbers appear in rows where the correct answer is not a unique mode of the distribution (e.g. patient 'L9' was correctly classified in 22 % of predictions, however this patient was misclassified as patient 'Z0' in 24 % of predictions.

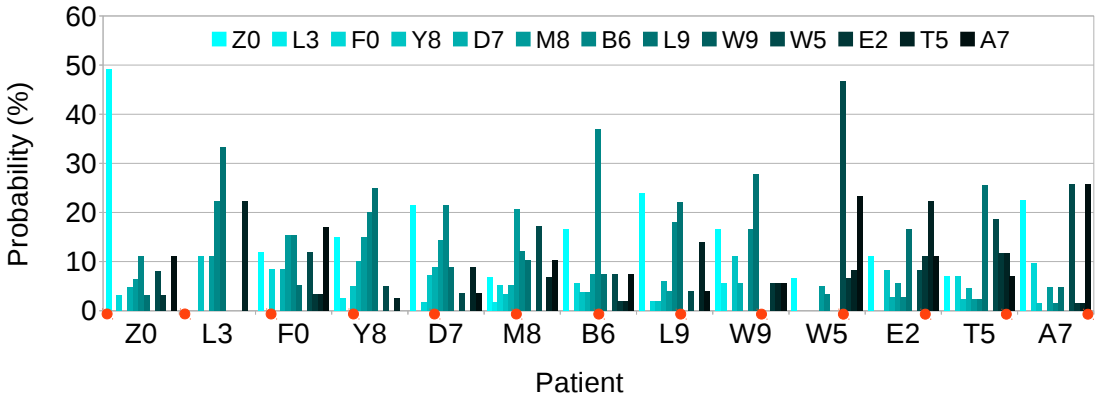

Figure 9.34 MLP probability distributions with all patients. Showing the distribution of classifications for each patient. Above each label a corresponding row from Fig. 9.33 is plotted. Red circles denote a bin for a correct patient.

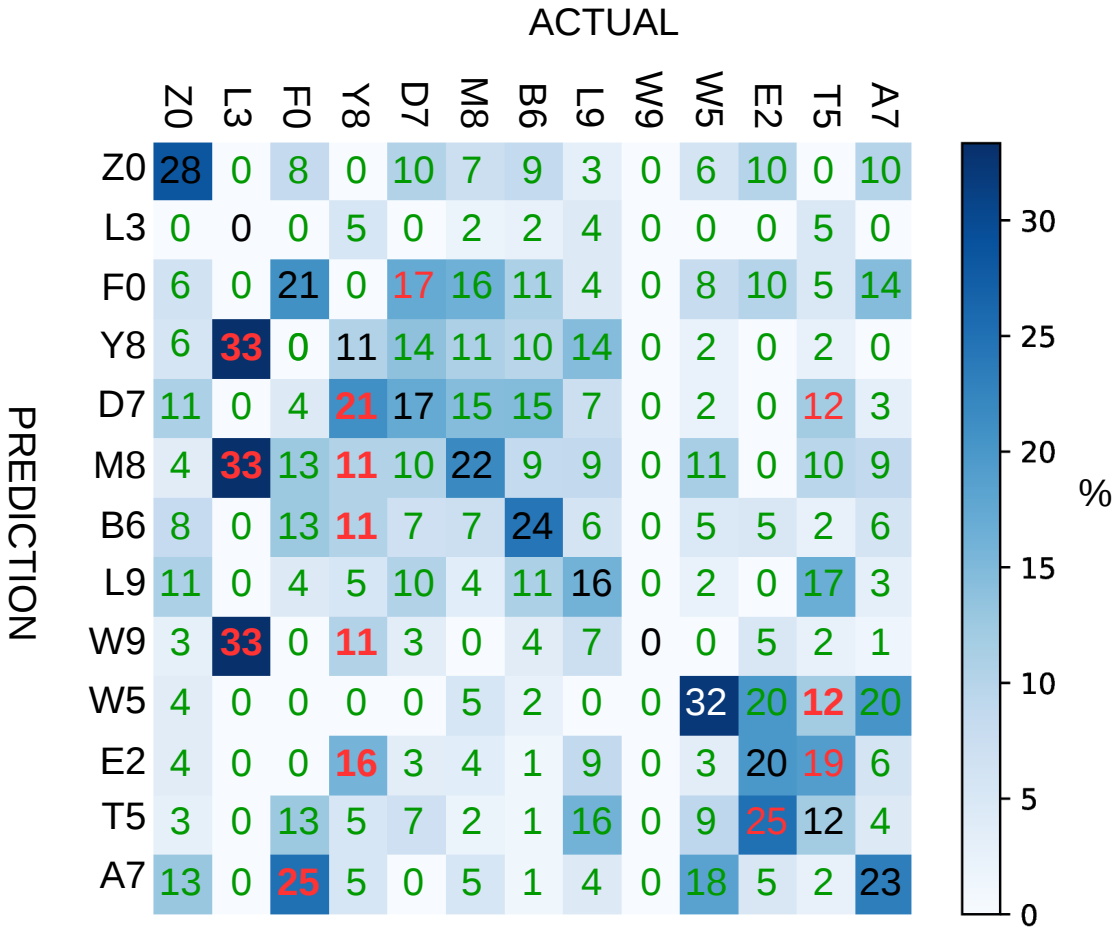

Figure 9.35 MLP result normalised along columns. See the text for more explanation.

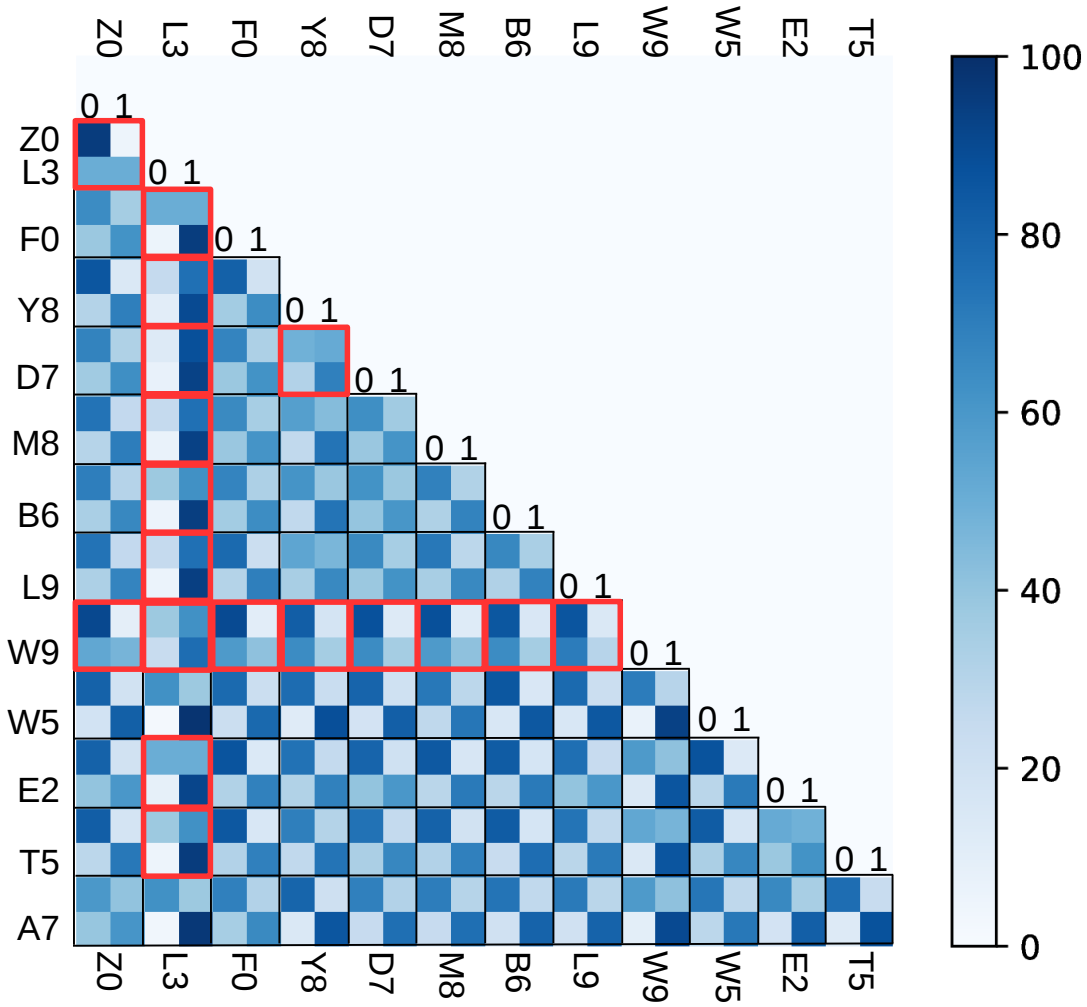

Figure 9.36 MLP classification with all combinations of two patients. Red squares denote two-patient-combinations where majority (>50 %) of both patient samples were classified to the same class.

The MLP result of training and classifying all possible two-patient-combinations is shown in Fig. 9.36. In 17 out of a total of 78 two-patient-combinations, more than 50 % of both patient samples were classified to the same class. Of these 17, two patients 'L3' and 'W9' contribute to 15 cases. The weighted classification accuracy with the MLP classifier was 74.2 %. Given any holograms of two patients, on average the correct patient can be predicted in 3 of 4 cases.

Figure 9.37 shows simulated patient identification results with different numbers of given samples. 10 to 100 (with steps of 1) MLP classification results were randomly simulated using the probability distributions (Fig. 9.34) obtained from the previously described 300 trained MLPs. In total, 10000 classification results for each patient were simulated. For each simulated result, the MAP (Eq. 9.5) was calculated using the $p(i \mid j)$

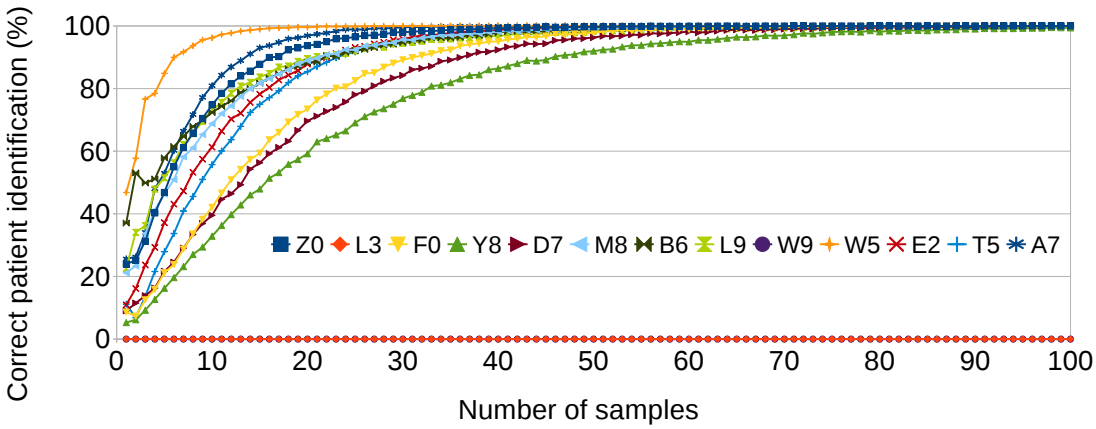

Figure 9.37 Patient prediction simulation with MLP. Showing prediction accuracy of all the patients when using different numbers of samples.

values obtained from Fig. 9.35 and the patient label was obtained. It can be observed that 11 out of 13 patients can be identified with 100 % accuracy. For these patients, the average number of samples needed to reach 90 % identification accuracy was 24 samples.

Figure 9.38 shows the classification results of using all the data with this experiment. It can be seen that the majority of patients can be identified to some extent based on the extracted data. 10 out of 13 patients have their highest percentage for the correct patient. With these patients, the test classification accuracy range is from 16 % to 50 %. The classification accuracy of each patient is better than a random guess. The weighted average classification accuracy with this classifier is 26.2 %. Figure 9.39 shows histograms of each row in Fig. 9.38.

The result of training and classifying all possible two-patient-combinations is shown in Fig. 9.41. In 5 out of the total 78 two-patient-combinations, more than 50 % of both patient samples were classified to the same class. Of these 5, patient 'L3' contributes to 3 cases. The weighted classification accuracy with the SVM classifier was 73.5 %. As with the MLP classifier, given any holograms of two patients, on average the correct patient can be predicted in 3 of 4 cases.

Figure 9.42 shows simulated patient identification results with different numbers of samples. 10 to 100 (with steps of 1) SVM classification results were randomly simulated using probability distributions (Fig. 9.39) obtained from the previously described 1000 trained SVMs. In total, 10000 classification results for each patient were simulated. For each simulated result, the MAP (Eq. 9.5) was calculated using the $p(i \mid j)$ values obtained

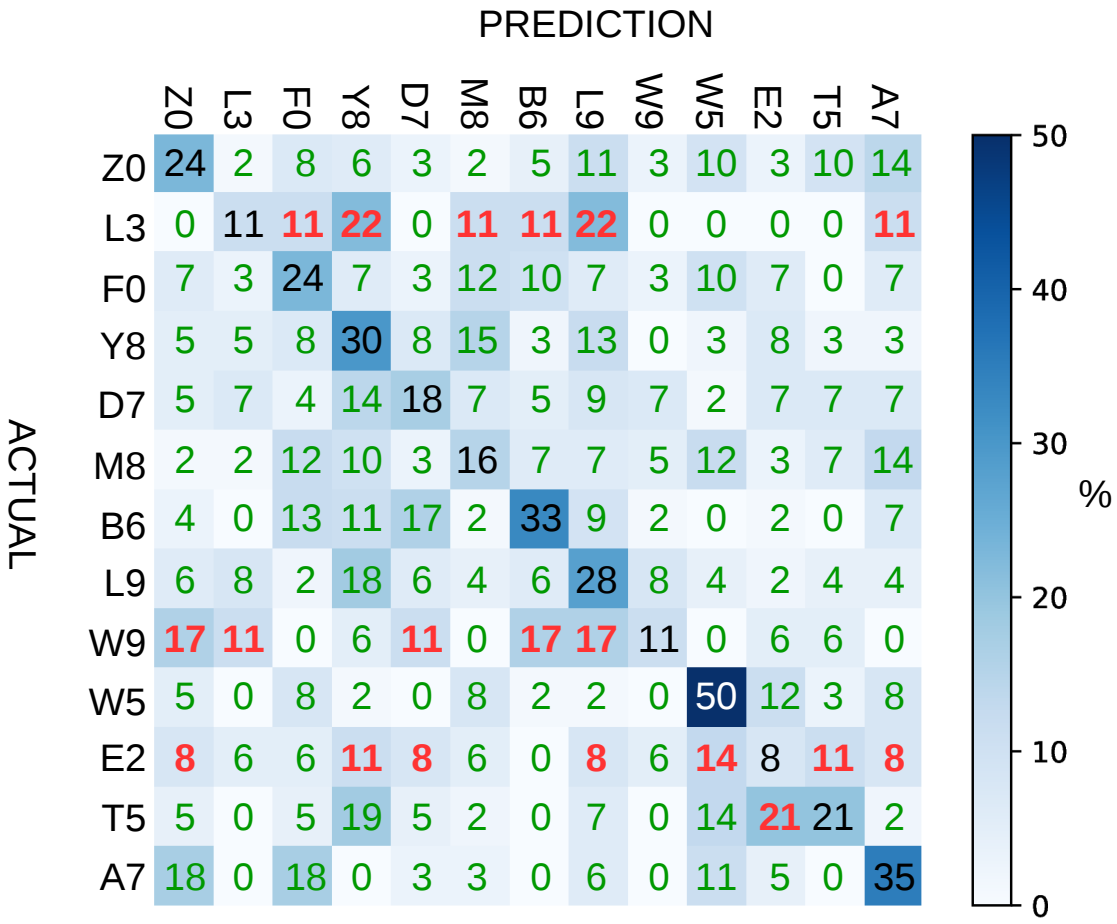

Figure 9.38 SVM result with all patient data. Black numbers (diagonal) show classification accuracy (i.e. correct patient predicted). The red numbers appear in rows where the correct answer is not a unique mode of the distribution.

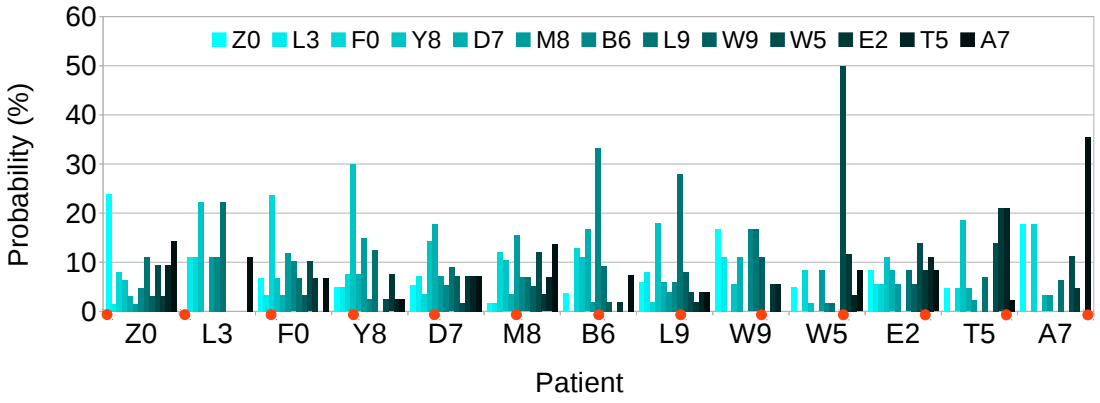

Figure 9.39 SVM probability distributions with all patients, showing the distribution of classifications for each patient. Above each label a corresponding row from Fig. 9.38 is plotted. Red circles denote a bin for a correct patient.

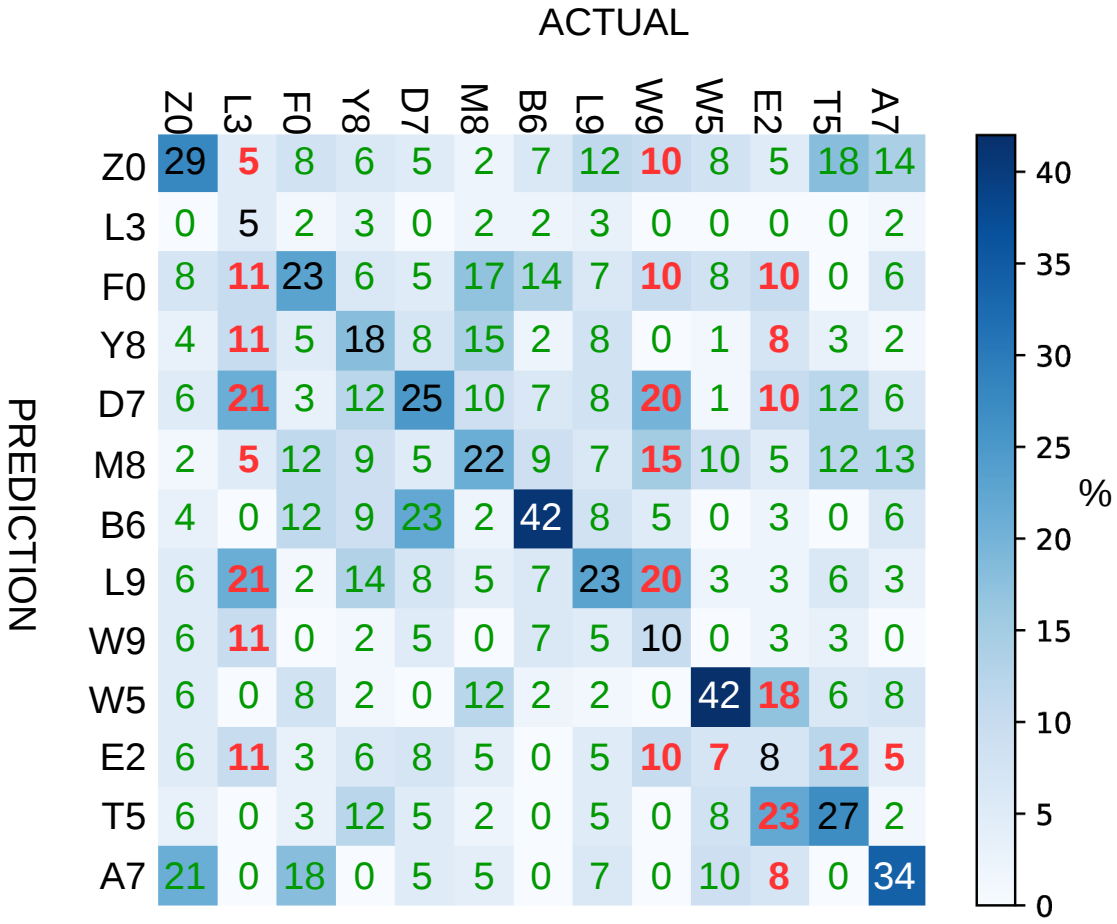

Figure 9.40 SVM result normalised along columns. See the text for more explanation.

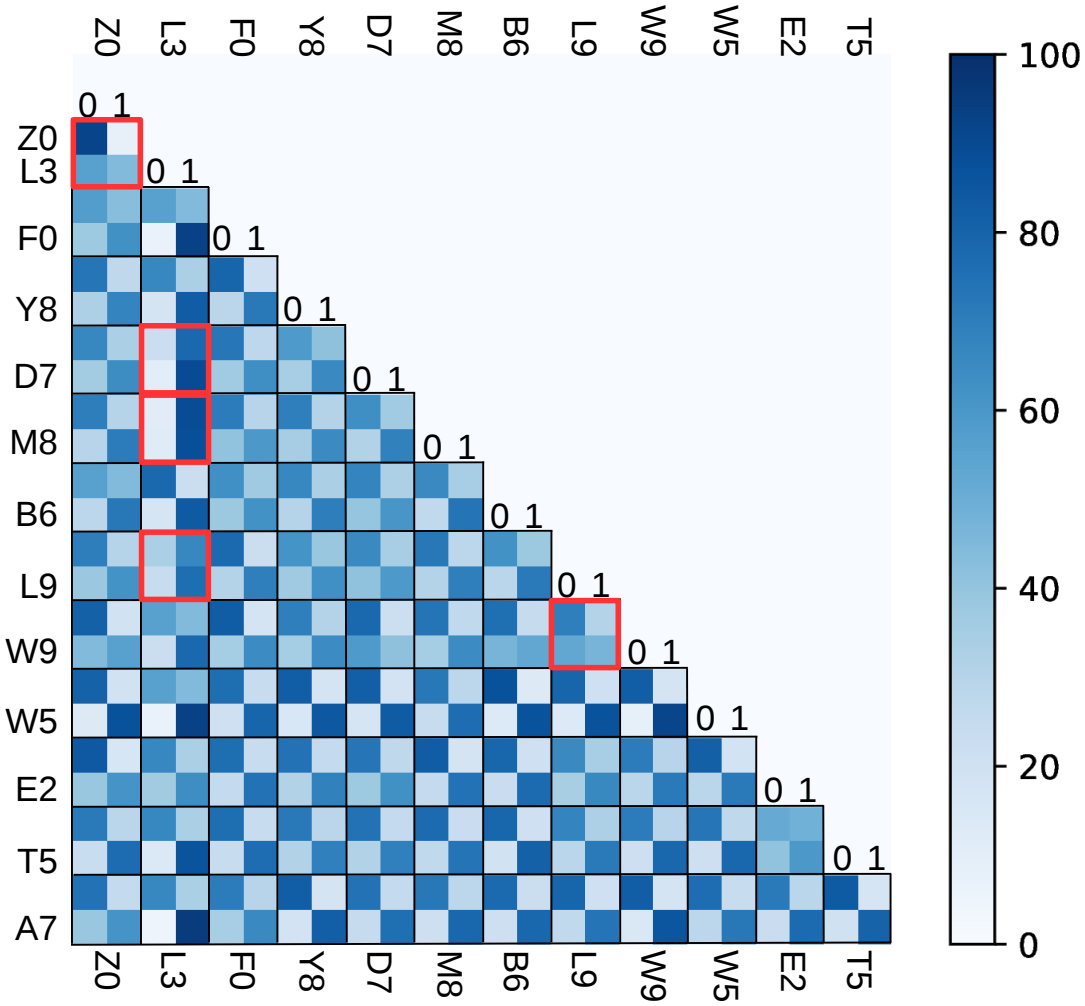

Figure 9.41 SVM classification with all combinations of two patients. Red squares denote two-patient-combinations where the majority (>50 %) of both patient samples were classified to the same class.

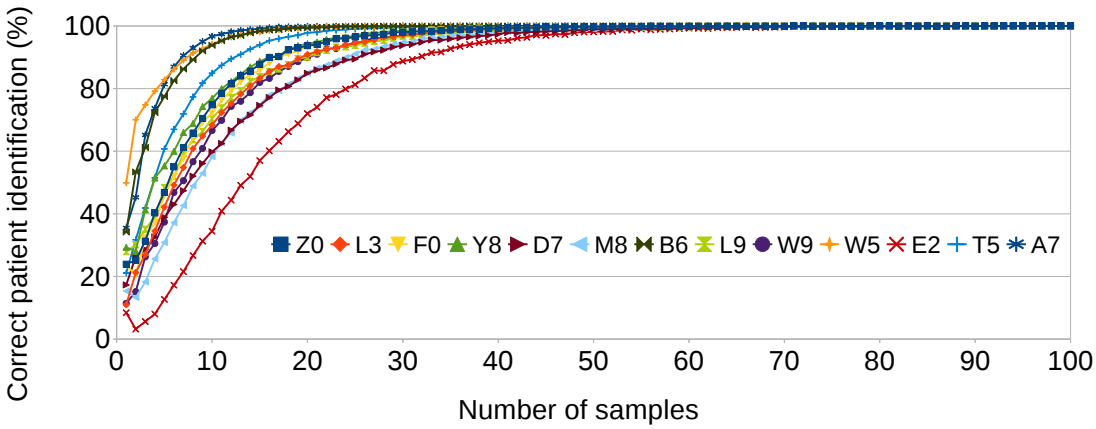

Figure 9.42 Patient prediction simulation with SVM. Showing prediction accuracy of all the patients when using different numbers of samples.

from Fig. 9.40 and the patient label was obtained. It can be observed that each of the 13 patients can be identified with 100 % accuracy. For these patients, the average number of samples needed to reach 90 % identification accuracy was 18 samples.

It should be noted that two patients 'L3' and 'W9' that received 0 % and 11 % classification accuracies with MLP and SVM, respectively, and performed poorly in patient identification, had the smallest amount of data available (in total 45 and 88 organoids, respectively).

9.4 Recommendations for future research

Working with real patient data has given us a chance to learn not only from the data but also the procedures related to obtaining the patient samples. We noticed and witnessed multiple issues or problems that should be improved or fixed in the future. For this reason we are giving the following recommendations:

1. We had a lack of control samples for our experiments. Samples should be gathered from biopsies taken for diagnosis rather than from prostatectomy patients, so that if a diagnosis of prostate cancer is not given, we can be somewhat confident that our control data comes from healthy tissue. An alternative for control data would be organ donors.

2. Enough tissue should be gathered so that cell sorting is possible and loss of the cells of interest during preparation is affordable.

3. The organoids from a patient sample should be uniformly sampled for imaging, rather than selected by a person, to avoid one source of bias in the hologram data.

4. Pathological evaluation (if possible) should be performed on the same tissue sample that is used for further analysis, rather than a neighbouring slice, to ensure correct labelling.

5. In the order of a thousand organoids for each patient should be imaged due to the large variability between organoids from a single patient.

9.5 Conclusions

This chapter contained experiments performed using time-lapse data of cell lines, and real patient data. When classifying time-lapse data with a combination of PCA and k-means clustering, there was no clear evidence that adding the rates of change of features will improve performance.

The other part of the chapter used patient data, where multiple different clustering scenarios were investigated with data extracted from holograms of patient samples. The extracted data was used in different ways; different partitions of patient data were used to investigate if the standard k-means clustering algorithm can be used to identify tumourigenic samples.

Clustering all of the patient data into two clusters solved the 'Clustering consistency problem'. It was observed that patients 'W5' and 'A7' had a majority of samples in cluster 1, while other patients had a large majority of samples in cluster 0. Unfortunately, cancer evaluation was not performed by a pathologist for patient 'A7'. Possibly this split implies differences in cancer proportions, however we can not draw too broad conclusions from these experiments. In order to solve the 'Cluster relevance problem', three 'BENIGN' patients and patient 'W5' (with a 'HIGH' cancer evaluation) were clustered into two clusters. Splits of samples between two clusters of 'BENIGN' patients differed greatly from the split of patient 'W5'. These fixed cluster centres were used to predict clusters for other patients.

Having fixed cluster centres solved the 'Cluster relevance problem'. Each cluster centre was now defined. However, this approach did not solve the 'Cluster indexing problem'. To solve this problem, PCA was used to reduce data dimensionality. It was observed that there are cluster centres close to each other for 12 out of 13 patients, and the data inside this region should have something in common. From the existing data, it could not be concluded what this mutual factor would be. This mutual factor is not directly related to a cancer evaluation given by one of the pathologists.

The 'Index mapping problem' remains untested. With the existing data and analysis, it is not possible to assign labels to different cluster indices. This will be solved when a proper control dataset is available as defined in Section 9.4.

Another set of experiments was performed by using SVM, MLP and CNN classifiers. With SVM and MLP classifiers, multiple classifiers were trained using two different datasets: a RWPE cell line and patient data. By using all 'BENIGN' patients and one 'HIGH' patient in training, and testing with the other patients, there was no linear trend with either classifier in samples classified as 'Cancer' from 'LOW' to 'HIGH' patients, as could be expected. With MLP classifiers, from the results of using extracted features of the RWPE cell line in training and patient data in testing, it can be concluded somewhat surprisingly that most of the patient samples resemble $KRas^{V12}$ samples. With SVM classifiers, the splits between the two classes are more balanced, but still most of the patient samples were classified as $KRas^{V12}$.

With the CNN classification, it was investigated if patient samples can be classified to be either cancerous or benign based on convolutional features. For this experiment, holograms of samples from 'BENIGN' hyperplasia patients formed data for class 0 and holograms of samples from a single patient with 'HIGH' pathological cancer evaluation formed data for class 1. The networks were not able to learn distinguishing features from the existing data, and from the results, we can not make too broad conclusions.

Problems with these different algorithms are partially due to the fact that the derived patient samples are highly heterogeneous, each sample consists of multiple cell types. In addition, pathological evaluation was performed on an adjacent tissue slice. It is possible that cells extracted from different parts of an organ do not display the same characteristics

and the split between benign and cancer cells is different. One major problem with the data is the lack of real control data. All of the patients had a cancer diagnosis and even though some of the organoids might be healthy, there are no correct labels associated with them. Another identified problem with the data is a possible source of bias due to the fact that all the imaged organoids were chosen by a person.

The most interesting observation from this chapter is that each patient can be uniquely identified from their organoid extracted morphological and phase features. It seems remarkable that a patient could be identified from their phenotype. It would seem to have profound implications for medical data privacy, but this is beyond the scope of the thesis.

10 Summary, discussion and future work

10.1 Summary

Cancer is a true global health challenge concerning millions of people every year. Early detection and diagnoses are two key factors for reducing mortality caused by cancer. For some cancer types there are existing screening programmes, and many cancers can be diagnosed by conducting medical tests and using medical imaging modalities. None of them utilise living cells where it would be possible to detect cancer and estimate its aggressiveness.

In this thesis, we chose DHM as the imaging modality that works at the cellular level. Imaging modality augmentation allowed high-throughput capturing of holographic data of living and moving 3D cysts. Despite the critical importance of this hardware for this thesis and work, the main contribution of this thesis was in the design, implementation, and testing of different software solutions. These solutions are summarised here in an order they would be used for a fully-functional automated high-throughput system. The first of these allows detection of the presence of an object in the field of view. This increases the total throughput of the system as only those holograms that contain an object should be processed and analysed further. This heuristic approach was parametrised using analysis of convolutional neural network classification results. The performance of the heuristic approach was at the same level as the more sophisticated and time-consuming deep learning approach.

A solution that finds the in-focus depth of an object in a hologram was introduced. Two different CNN approaches, regression and classification, were shown to work with the

holographic data. These approaches used amplitude information only. The greatest benefit of these approaches, if compared to other traditional solutions that need to reconstruct a stack of images, was the fact that the runtime is constant. It was shown that these solutions perform well, providing accurate in-focus depth predictions. In addition, an experiment was realised where the amount of needed data for a CNN was investigated. It was shown that a good performance can be obtained with a relatively small amount of data. We showed that reducing data by 60 % does not affect accuracy. This can be an important benefit for other domains as well. Having enough but not too much data can reduce learning times significantly.

A segmentation algorithm using a CNN approach was shown to work well with holographic data. It was shown that the in-focus phase reconstructions as inputs provide the most accurate segmentation results. Ground-truth masks, for training, were made semi-automatically using outputs from a specially designed heuristic segmentation approach. The heuristic approach used the phase slope information of closely packed phase wraps present at the outer layer of cysts. As the hologram is a projection along the optical axis, this layer is formed of tens of cells between the basement and apical membranes that cause phase wraps in the wrapped phase image. The CNN approach was able to segment cysts and hollow lumen with high accuracy. This allowed accurate analysis of lumen, unlike using the heuristic segmentation approach, as a single lumen may be filled with cells or other biological material at some 3D regions while other regions are hollow.

From the segmented cysts, quantitative features was extracted. It was shown that data can be extracted, that is equivalent to other 3D microscopy techniques. In addition, it was shown that quantitative phase-related features, such as density can be extracted. To this end, we developed a novel metric, lumen index, that quantifies the lumen. The quantity of a lumen is the most representative feature when biologists qualitatively analyse cysts' potentiality for tumourigenesis. This metric was formulated based on a close collaboration with biologists and the requirements for the software came from the field of biology.

Extracted features were used with different conventional non-deep classifiers to classify cysts as either healthy or oncogenic. SVMs, random forests and multilayer perceptrons represented conventional classifiers. These classifiers were trained with a different number

of extracted features that were added in the order given by the analysis of variance. We showed that by adding more, but not necessarily all, extracted features, classification accuracy could be improved. The performance of conventional classifiers with holograms that were captured at the same time as the training data was excellent. Problems arose when data was captured at a different time, under different laboratory conditions, showing that classification based on extracted features only was not reliable enough. To confront this challenge, deep learning approaches were applied. An exhaustive search of well-defined subset of shallow CNN architectures was executed and performance was compared with a well-known deeper CNN model. We trained 62 different CNN architectures that were built using simple building blocks. We were able to achieve accuracy within 10 % of AlexNet by using approximately 1 % of its learning parameters. This is a significant finding showing that extremely shallow models can provide excellent performance for some difficult two-class classification tasks. Shallow and computationally simple models are also preferred for mobile and internet-of-things devices.

An experiment measuring the importance of the non-overlapping background of cyst reconstructions was carried out. It was shown that objects need to be segmented with great care as the background can contain a "fingerprint" that a CNN can learn, and classify based on this information alone. As another notable and novel result, we showed that adding a priori hand-crafted information from holograms of cysts into a CNN improves network performance of shallow networks. With deeper models containing more learning parameters, hand-crafted features are less important.

It was shown that hardware modifications allowed time-lapse data capture of living and moving 3D cysts. It was shown that different quantitative metrics can be followed over long periods of time. A combination of PCA and *k*-means for sample discrimination was investigated. It was shown that adding change-rate of different metrics does not necessarily improve discrimination. Real patient tissue samples were imaged with the DHM. Different approaches using the *k*-means clustering algorithm were attempted. It was shown that the hologram data is extremely challenging to work with and off-the-shelf *k*-means cannot be used ‘as-is’ for any reliable discrimination tasks with the existing data. Different non-linear classifiers were trained to classify patient data. With SVM and MLP classifiers it was

shown that patients can be identified based on extracted features of organoids only. A large set of problems with the existing patient data were identified and discussed.

In summary, this thesis provides guidelines on how it is possible to implement a system that can be used for high-throughput analysis of living 3D cysts. These guidelines cover hardware, software, and biology, with an emphasis on software. The solutions invented in this thesis can be combined and used together to provide a fully automated analysis system.

10.2 Discussion

The rigorous goal set at the beginning of this project was high-throughput morphological signature establishment of tumour cells with the use of digital holographic microscopy. We wanted to reduce the software gap that prevents using DHM for automated high-throughput analysis system. These solutions, when developed, could enable wider use of DHM technology. In answer to this need, we were able to develop and implement various different algorithms that together with the mechanical and optical hardware enable an automated high-throughput system as summarised in the previous section. The solutions in this thesis can be used to identify cancer-specific features to some extent. For the prostate cancer patient samples, identification of tumourigenic samples is still an open question. This is due to the limited amount of available patient samples and heterogeneous sample feature set. In addition, as imaged organoids were selected manually by a person, it is possible that our data is biased. It is possible that the chosen organoids were chosen based on some visually appealing feature. Those organoids might be the most normal or alternatively the most abnormal looking examples; we do not know this.

When starting this project, all developed solutions were model-based. In collaboration with biologists, we attempted to build models that describe cysts. At an early stage, it was noted that the data we had was noisy, normal samples looking tumorigenic and vice versa, and software needed to be more robust for reliable results. This observation led to introducing machine learning. However, these approaches were still relying strongly on extracted features. Quantitative features were extracted and used with different conventional machine learning approaches. These approaches performed quite well with the data,

however, the performance was not as good as required. This reflection steered the research in the direction of learning-based approaches as these approaches were used actively with a great success in many scientific fields. With deep learning, we were able to achieve good and remarkable results within many different applications.

Interestingly, throughout the thesis, we were able to combine these two different approaches, model-based and learning-based. Throughout the thesis we have identified this theme as the paradigms theme. We were able to transfer knowledge between the two approaches both ways within different applications. For the field of view solution (Chapter 5), we were able to learn from a CNN which features are the most distinguishing if there is an object in the field of view. This knowledge was incorporated to the model-based solution. For the CNN segmentation (Chapter 7), the initial ground-truth masks were created using a model-based approach. This is truly a novel approach, typically these masks are made completely manually using a very laborious process. For shallow networks (Chapter 8), we were able to improve the performance of shallow CNN models by integrating a-priori knowledge (model-based) into the network. Model-based and learning-based approaches can support and improve performance each other. Each approach can bring useful advantage for another approach.

Another identified theme has been referred to as simplification throughout the thesis. This theme was present in multiple different sections throughout the thesis. The number of extracted features for conventional classifiers can be reduced while retaining good performance. For a CNN, the amount of data can be reduced significantly without decreasing classification accuracy or network performance. Shallow networks are another example of the theme in this thesis. The shallow networks results show that with a fraction of AlexNet's parameters, CNNs are able to perform well in some hard binary classification tasks. Another example of the theme was the study of the importance of non-overlapping background in classification. Through a model-based approach in the experiment, it was found out that background needs to be removed completely for reliable CNN classification.

Is there future for digital holographic microscopy? Is there future for deep learning? Is there future for deep learning within holographic microscopy?

By searching the PubMed database [300], each year approximately 70 papers in the field of digital holography are published. The exponential growth in the number of deep learning publications is remarkable. Deep learning is used in an extremely wide spectrum of applications in various scientific fields. This trend has been seen also in the field of digital holographic microscopy. Over the past few years, more and more interesting applications using deep learning have been published by multiple different research groups across the world. As for the technology itself, DHM has advantages over many other imaging techniques that make it a good candidate for high-throughput analysis. The current trend with 3D samples seems to be to use different tomographic techniques. Still, DHM without tomography is more appropriate for high-throughput analysis. DHM is a single-shot technique allowing freely moving samples in the field of view. In addition, processing a single hologram is faster than processing a 3D stack. The advantages of DHM and increased interest in the use of deep learning in the field, give confidence that DHM will be an important technology, and deep learning will be used within DHM.

10.3 Future work

The system described in this thesis can be used to monitor living cysts over long periods of time and the system is able to analyse thousands of cysts from a single petri dish. Besides morphological measurements, this allows time-lapse imaging that can be used to estimate cancer aggressiveness by monitoring the behaviour of cysts. Further analysis needs to be performed on extracted time-lapse data in order to invent suitable metrics.

An interesting aspect of future work is in the use of real patient data where cell samples derived from cancer patients are cultured and analysed. Different treatments can be applied to multiple samples derived from the same patient. This would be one step closer to personalised medicine – a trend towards which medicine, in general, is moving. Real patient samples, as shown also in this thesis, provide additional challenges. In the future, possibly some fluorescent marker can be used to mark only the cells of interest, and the same FOVs can be captured with two different imaging modalities (e.g. DHM and epifluorescence microscope). The patient tissue samples analysed in this thesis were obtained from patients going through a radical surgical operation. In the future, our

collaborators will obtain these tissue samples through biopsy. One advantage of this is that with a high probability biopsies contain also healthy tissue that can be collected, cultured, and used for machine learning. This would form a real control dataset, that we did not have. Our high-throughput analysis system is not restricted to any specific cancer type, or any specific cell line - in the future different types of cancers and cell lines should be investigated.

The current trend in deep learning applications is to train a single network to do multiple tasks. In the future, one interesting possibility would be to investigate and design a network that is able to take a real-valued hologram as an input and output multiple different metrics and/or images. One example could be a network that with a hologram input, outputs multiple images e.g. a segmentation mask and a segmented in-focus phase reconstruction.

References

[1] World Health Organization, "WHO/Europe | Cancer - Data and statistics," http://www.euro.who.int/en/health-topics/noncommunicable-diseases/cancer/data-and-statistics, accessed: 2017-03-21.

[2] J. De Bono and A. Ashworth, "Translating cancer research into targeted therapeutics," *Nature*, vol. 467, no. 7315, pp. 543–549, 2010. [Online]. Available: https://doi.org/10.1038/nature09339

[3] N. R. Council *et al.*, *Toward precision medicine: building a knowledge network for biomedical research and a new taxonomy of disease*. National Academies Press, 2011.

[4] D. Gonzalez de Castro, P. Clarke, B. Al-Lazikani, and P. Workman, "Personalized cancer medicine: molecular diagnostics, predictive biomarkers, and drug resistance," *Clinical Pharmacology & Therapeutics*, vol. 93, no. 3, pp. 252–259, 2013. [Online]. Available: https://doi.org/10.1038/clpt.2012.237

[5] F. S. Collins and H. Varmus, "A new initiative on precision medicine," *New England Journal of Medicine*, vol. 372, no. 9, pp. 793–795, 2015. [Online]. Available: https://doi.org/10.1056/NEJMp1500523

[6] P. C. Nowell, "The clonal evolution of tumor cell populations," *Science*, vol. 194, no. 4260, pp. 23–28, 1976.

[7] L. Song, C. Varma, J. Verhoeven, and H. J. Tanke, "Influence of the triplet excited state on the photobleaching kinetics of fluorescein in microscopy," *Biophysical Journal*, vol. 70, no. 6, pp. 2959–2968, 1996. [Online]. Available: https://doi.org/10.1016/S0006-3495(96)79866-1

[8] R. M. Martin, H. Leonhardt, and M. C. Cardoso, "DNA labeling in living cells," *Cytometry Part A*, vol. 67, no. 1, pp. 45–52, 2005. [Online]. Available: https://doi.org/10.1002/cyto.a.20172

[9] E. Cuche, F. Bevilacqua, and C. Depeursinge, "Digital holography for quantitative phase-contrast imaging," *Optics Letters*, vol. 24, no. 5, pp. 291–293, 1999. [Online]. Available: https://doi.org/10.1364/OL.24.000291

[10] D. Gabor, "A new microscopic principle," *Nature*, vol. 161, no. 4098, pp. 777–778, 1948. [Online]. Available: https://doi.org/10.1038/161777a0

[11] T. Mitchell, *Machine Learning*. McGraw Hill, 1997.

[12] K. Pearson, "LIII. On lines and planes of closest fit to systems of points in space," *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, vol. 2, no. 11, pp. 559–572, 1901. [Online]. Available: https://doi.org/10.1080/14786440109462720

[13] R. A. Fisher, "The statistical utilization of multiple measurements," *Annals of Eugenics*, vol. 8, no. 4, pp. 376–386, 1938. [Online]. Available: https://doi.org/10.1111/j.1469-1809.1938.tb02189.x

[14] A. L. Samuel, "Some studies in machine learning using the game of checkers," *IBM Journal of Research and Development*, vol. 3, no. 3, pp. 210–229, July 1959. [Online]. Available: https://doi.org/10.1147/rd.33.0210

[15] F. Rosenblatt, "The perceptron a perceiving and recognizing automaton," tech. rep., Technical Report 85-460-1, Cornell Aeronautical Laboratory, 1957. 2, Tech. Rep., 1957.

[16] T. Cover and P. Hart, "Nearest neighbor pattern classification," *IEEE transactions on information theory*, vol. 13, no. 1, pp. 21–27, 1967. [Online]. Available: https://doi.org/10.1109/TIT.1967.1053964

[17] D. Psaltis and N. Farhat, "Optical information processing based on an associative-memory model of neural nets with thresholding and feedback," *Optics Letters*, vol. 10, no. 2, pp. 98–100, Feb 1985. [Online]. Available: https://doi.org/10.1364/OL.10.000098

[18] Y. Frauel and B. Javidi, "Neural network for three-dimensional object recognition based on digital holography," *Optics Letters*, vol. 26, no. 19, pp. 1478–1480, 2001. [Online]. Available: https://doi.org/10.1364/OL.26.001478

[19] A. E. Shortt, T. J. Naughton, and B. Javidi, "Compression of optically encrypted digital holograms using artificial neural networks," *Journal of Display Technology*, vol. 2, no. 4, pp. 401–410, Dec 2006. [Online]. Available: https://doi.org/10.1109/JDT.2006.884693

[20] U. S. Kamilov, I. N. Papadopoulos, M. H. Shoreh, A. Goy, C. Vonesch, M. Unser, and D. Psaltis, "Learning approach to optical tomography," *Optica*, vol. 2, no. 6, pp. 517–622, 2015. [Online]. Available: https://doi.org/10.1364/OPTICA.2.000517

[21] B. Schneider, J. Dambre, and P. Bienstman, "Fast particle characterization using digital holography and neural networks," *Applied Optics*, vol. 55, no. 1, pp. 133–139, Jan 2016. [Online]. Available: https://doi.org/10.1364/AO.55.000133

[22] G. Barbastathis, A. Ozcan, and G. Situ, "On the use of deep learning for computational imaging," *Optica*, vol. 6, no. 8, pp. 921–943, Aug 2019. [Online]. Available: http://www.osapublishing.org/optica/abstract.cfm?URI=optica-6-8-921

[23] J. W. Goodman, *Introduction to Fourier Optics*. Roberts and Company Publishers, 2005.

[24] E. N. Leith and J. Upatnieks, "Wavefront reconstruction with continuous-tone objects," *Journal of Optical Society of America A*, vol. 53, no. 12, pp. 1377–1381, 1963. [Online]. Available: https://doi.org/10.1364/JOSA.53.001377

[25] G. W. Stroke and D. G. Falconer, "Attainment of high resolutions in wavefront-reconstruction imaging," *Physics Letters*, vol. 13, no. 4, pp. 306–309, 1964.

[26] B. R. Brown and A. W. Lohmann, "Complex spatial filtering with binary masks," *Applied Optics*, vol. 5, no. 6, pp. 967–969, Jun 1966. [Online]. Available: https://doi.org/10.1364/AO.5.000967

[27] J. W. Goodman and R. Lawrence, "Digital image formation from electronically detected holograms," *Applied Physics Letters*, vol. 11, no. 3, pp. 77–79, 1967. [Online]. Available: https://doi.org/10.1063/1.1755043@apl.2019.APLCLASS2019.issue-1

[28] J. W. Goodman, *Introduction to Fourier Optics, McGaw-Hill Physical and Quantum Electronics Series*. McGraw-Hill Book Co., New York, New York, 1968.

[29] L. P. Yaroslavskii and N. S. Merzliakov, "Methods of digital holography," *Moscow Izdatel Nauka*, vol. 1, 1977.

[30] U. Schnars and W. Jüptner, "Direct recording of holograms by a CCD target and numerical reconstruction," *Applied Optics*, vol. 33, no. 2, pp. 179–181, 1994. [Online]. Available: https://doi.org/10.1364/AO.33.000179

[31] I. Yamaguchi and T. Zhang, "Phase-shifting digital holography," *Optics Letters*, vol. 22, no. 16, pp. 1268–1270, Aug 1997. [Online]. Available: https://doi.org/10.1364/OL.22.001268

[32] I. Yamaguchi, T. Matsumura, and J. ichi Kato, "Phase-shifting color digital holography," *Optics Letters*, vol. 27, no. 13, pp. 1108–1110, Jul 2002. [Online]. Available: https://doi.org/10.1364/OL.27.001108

[33] O. Matoba, T. J. Naughton, Y. Frauel, N. Bertaux, and B. Javidi, "Real-time three-dimensional object reconstruction by use of a phase-encoded digital hologram," *Applied Optics*, vol. 41, no. 29, pp. 6187–6192, Oct 2002. [Online]. Available: http://ao.osa.org/abstract.cfm?URI=ao-41-29-6187

[34] L. Onural, F. Yaras, and H. Kang, "Digital holographic three-dimensional video displays," *Proceedings of the IEEE*, vol. 99, no. 4, pp. 576–589, 2011. [Online]. Available: https://doi.org/10.1109/JPROC.2010.2098430

[35] W. S. Haddad, D. Cullen, J. C. Solem, J. W. Longworth, A. McPherson, K. Boyer, and C. K. Rhodes, "Fourier-transform holographic microscope," *Applied Optics*, vol. 31, no. 24, pp. 4973–4978, Aug 1992. [Online]. Available: https://doi.org/10.1364/AO.31.004973

[36] K. Boyer, J. Solem, J. Longworth, A. Borisov, and C. Rhodes, "Biomedical three-dimensional holographic microimaging at visible, ultraviolet and X-ray wavelengths," *Nature Medicine*, vol. 2, no. 8, pp. 939–941, 1996. [Online]. Available: https://doi.org/10.1038/nm0896-939

[37] T. Zhang and I. Yamaguchi, "Three-dimensional microscopy with phase-shifting digital holography," *Optics Letters*, vol. 23, no. 15, pp. 1221–1223, Aug 1998. [Online]. Available: https://doi.org/10.1364/OL.23.001221

[38] U. Schnars, "Direct phase determination in hologram interferometry with use of digitally recorded holograms," *Journal of Optical Society of America A*, vol. 11, no. 7, pp. 2011–2015, Jul 1994. [Online]. Available: https://doi.org/10.1364/JOSAA.11.002011

[39] E. Cuche, P. Marquet, and C. Depeursinge, "Simultaneous amplitude-contrast and quantitative phase-contrast microscopy by numerical reconstruction of Fresnel off-axis holograms," *Applied Optics*, vol. 38, no. 34, pp. 6994–7001, Dec 1999. [Online]. Available: https://doi.org/10.1364/AO.38.006994

[40] H. J. Kreuzer, N. Pomerleau, K. Blagrave, and M. H. Jericho, "Digital in-line holography with numerical reconstruction," in *International Conference on Optical Metrology*. International Society for Optics and Photonics, 1999, pp. 65–74.

[41] K. M. Molony, B. M. Hennelly, D. P. Kelly, and T. J. Naughton, "Reconstruction algorithms applied to in-line Gabor digital holographic microscopy," *Optics Communications*, vol. 283, no. 6, pp. 903–909, 2010.

[42] J. P. Ryle, S. McDonnell, and J. T. Sheridan, "Lensless multispectral digital in-line holographic microscope," *Journal of Biomedical Optics*, vol. 16, no. 12, pp. 126 004–12 600 417, 2011. [Online]. Available: https://doi.org/10.1117/1.3659681

[43] F. Charrière, A. Marian, F. Montfort, J. Kuehn, T. Colomb, E. Cuche, P. Marquet, and C. Depeursinge, "Cell refractive index tomography by digital holographic microscopy," *Optics Letters*, vol. 31, no. 2, pp. 178–180, Jan 2006. [Online]. Available: https://doi.org/10.1364/OL.31.000178

[44] F. Charrière, N. Pavillon, T. Colomb, C. Depeursinge, T. J. Heger, E. A. Mitchell, P. Marquet, and B. Rappaz, "Living specimen tomography by digital holographic microscopy: morphometry of testate amoeba," *Optics Express*, vol. 14, no. 16, pp. 7005–7013, 2006. [Online]. Available: https://doi.org/10.1364/OE.14.007005

[45] W. Choi, C. Fang-Yen, K. Badizadegan, S. Oh, N. Lue, R. R. Dasari, and M. S. Feld, "Tomographic phase microscopy," *Nature Methods*, vol. 4, no. 9, p. 717, 2007. [Online]. Available: https://doi.org/10.1038/nmeth1078

[46] G. Dardikman, M. Habaza, L. Waller, and N. T. Shaked, "Video-rate processing in tomographic phase microscopy of biological cells using CUDA," *Optics Express*, vol. 24, no. 11, pp. 11 839–11 854, May 2016. [Online]. Available: https://doi.org/10.1364/OE.24.011839

[47] J. Kostencka, T. Kozacki, A. Kuś, B. Kemper, and M. Kujawińska, "Holographic tomography with scanning of illumination: space-domain reconstruction for spatially invariant accuracy," *Biomedical Optics Express*, vol. 7, no. 10, pp. 4086–4101, Oct 2016. [Online]. Available: https://doi.org/10.1364/BOE.7.004086

[48] D. Jin, R. Zhou, Z. Yaqoob, and P. T. C. So, "Tomographic phase microscopy: principles and applications in bioimaging," *Journal of Optical Society of America B*, vol. 34, no. 5, pp. B64–B77, May 2017. [Online]. Available: https://doi.org/10.1364/JOSAB.34.000B64

[49] Y. Cotte, F. Toy, P. Jourdain, N. Pavillon, D. Boss, P. Magistretti, P. Marquet, and C. Depeursinge, "Marker-free phase nanoscopy," *Nature Photonics*, vol. 7, no. 2, p. 113, 2013. [Online]. Available: https://doi.org/10.1038/nphoton.2012.329

[50] T. M. Kreis and W. P. O. Jüptner, "Suppression of the dc term in digital holography," *Optical Engineering*, vol. 36, no. 8, pp. 2357–2360, 1997. [Online]. Available: https://doi.org/10.1117/1.601426

[51] G.-L. Chen, C.-Y. Lin, M.-K. Kuo, and C.-C. Chang, "Numerical suppression of zero-order image in digital holography," *Optics Express*, vol. 15, no. 14, pp. 8851–8856, Jul 2007. [Online]. Available: https://doi.org/10.1364/OE.15.008851

[52] P. Zhang, R. Li, and J. Li, "Segmentation of holographic images using the level set method," *Optik-International Journal for Light and Electron Optics*, vol. 123, no. 2, pp. 132–136, 2012. [Online]. Available: https://doi.org/10.1016/j.ijleo.2011.02.023

[53] E. Cuche, P. Marquet, and C. Depeursinge, "Spatial filtering for zero-order and twin-image elimination in digital off-axis holography," *Applied Optics*, vol. 39, no. 23, pp. 4070–4075, 2000. [Online]. Available: https://doi.org/10.1364/AO.39.004070

[54] X. He, C. V. Nguyen, M. Pratap, Y. Zheng, Y. Wang, D. R. Nisbet, R. J. Williams, M. Rug, A. G. Maier, and W. M. Lee, "Automated Fourier space region-recognition filtering for off-axis digital holographic microscopy," *Biomedical Optics Express*, vol. 7, no. 8, pp. 3111–3123, Aug 2016. [Online]. Available: https://doi.org/10.1364/BOE.7.003111

[55] P. Ferraro, S. D. Nicola, A. Finizio, G. Coppola, S. Grilli, C. Magro, and G. Pierattini, "Compensation of the inherent wave front curvature in digital holographic coherent microscopy for quantitative phase-contrast imaging," *Applied Optics*, vol. 42, no. 11, pp. 1938–1946, Apr 2003. [Online]. Available: https://doi.org/10.1364/AO.42.001938

[56] T. Colomb, F. Montfort, J. Kühn, N. Aspert, E. Cuche, A. Marian, F. Charrière, S. Bourquin, P. Marquet, and C. Depeursinge, "Numerical parametric lens for shifting, magnification, and complete aberration compensation in digital holographic microscopy," *Journal of Optical Society of America A*, vol. 23, no. 12, pp. 3177–3190, Dec 2006. [Online]. Available: https://doi.org/10.1364/JOSAA.23.003177

[57] C. Zuo, Q. Chen, W. Qu, and A. Asundi, "Phase aberration compensation in digital holographic microscopy based on principal component analysis," *Optics Letters*, vol. 38, no. 10, pp. 1724–1726, 2013. [Online]. Available: https://doi.org/10.1364/OL.38.001724

[58] G. Goldstein and K. Creath, "Quantitative phase microscopy: automated background leveling techniques and smart temporal phase unwrapping," *Applied Optics*, vol. 54, no. 16, pp. 5175–5185, Jun 2015. [Online]. Available: https://doi.org/10.1364/AO.54.005175

[59] S. Liu, Q. Lian, Y. Qing, and Z. Xu, "Automatic phase aberration compensation for digital holographic microscopy based on phase variation minimization," *Optics Letters*, vol. 43, no. 8, pp. 1870–1873, 2018. [Online]. Available: https://doi.org/10.1364/OL.43.001870

[60] T. Nguyen, V. Bui, V. Lam, C. B. Raub, L.-C. Chang, and G. Nehmetallah, "Automatic phase aberration compensation for digital holographic microscopy based on deep learning background detection," *Optics Express*, vol. 25, no. 13, pp. 15 043–15 057, 2017. [Online]. Available: https://doi.org/10.1364/OE.25.015043

[61] D. C. Ghiglia, G. A. Mastin, and L. A. Romero, "Cellular-automata method for phase unwrapping," *Journal of Optical Society of America A*, vol. 4, no. 1, pp. 267–280, Jan 1987. [Online]. Available: https://doi.org/10.1364/JOSAA.4.000267

[62] D. C. Ghiglia and L. A. Romero, "Robust two-dimensional weighted and unweighted phase unwrapping that uses fast transforms and iterative methods," *Journal of Optical Society of America A*, vol. 11, no. 1, pp. 107–117, Jan 1994. [Online]. Available: https://doi.org/10.1364/JOSAA.11.000107

[63] D. Kerr, G. H. Kaufmann, and G. E. Galizzi, "Unwrapping of interferometric phase-fringe maps by the discrete cosine transform," *Applied Optics*, vol. 35, no. 5, pp. 810–816, Feb 1996. [Online]. Available: https://doi.org/10.1364/AO.35.000810

[64] T. J. Flynn, "Two-dimensional phase unwrapping with minimum weighted discontinuity," *Journal of Optical Society of America A*, vol. 14, no. 10, pp. 2692–2701, 1997. [Online]. Available: https://doi.org/10.1364/JOSAA.14.002692

[65] M. A. Schofield and Y. Zhu, "Fast phase unwrapping algorithm for interferometric applications," *Optics Letters*, vol. 28, no. 14, pp. 1194–1196, Jul 2003. [Online]. Available: https://doi.org/10.1364/OL.28.001194

[66] Y. Li, J. Zhu, and W. Shen, "Phase unwrapping algorithms, respectively, based on path-following and discrete cosine transform," *Optik-International Journal for Light and Electron Optics*, vol. 119, no. 11, pp. 545–547, 2008. [Online]. Available: https://doi.org/10.1016/j.ijleo.2007.03.003

[67] S. Liu, W. Xiao, F. Pan, F. Wang, and L. Cong, "Complex-amplitude-based phase unwrapping method for digital holographic microscopy," *Optics and Lasers in Engineering*, vol. 50, no. 3, pp. 322 – 327, 2012. [Online]. Available: http://doi.org/10.1016/j.optlaseng.2011.11.006

[68] B. Tayebi, M. R. Jafarfard, F. Sharif, Y. S. Song, D. Har, and D. Y. Kim, "Large step-phase measurement by a reduced-phase triple-illumination interferometer," *Optics Express*, vol. 23, no. 9, pp. 11 264–11 271, May 2015. [Online]. Available: https://doi.org/10.1364/OE.23.011264

[69] Z. Cheng, D. Liu, Y. Yang, T. Ling, X. Chen, L. Zhang, J. Bai, Y. Shen, L. Miao, and W. Huang, "Practical phase unwrapping of interferometric fringes based on unscented Kalman filter technique," *Optics Express*, vol. 23, no. 25, pp. 32 337–32 349, Dec 2015. [Online]. Available: https://doi.org/10.1364/OE.23.032337

[70] P. Girshovitz and N. T. Shaked, "Fast phase processing in off-axis holography using multiplexing with complex encoding and live-cell fluctuation map calculation in real-time," *Optics Express*, vol. 23, no. 7, pp. 8773–8787, Apr 2015. [Online]. Available: https://doi.org/10.1364/OE.23.008773

[71] D. Claus, G. Pedrini, and W. Osten, "Iterative phase retrieval based on variable wavefront curvature," *Applied Optics*, vol. 56, no. 13, pp. F134–F137, May 2017. [Online]. Available: https://doi.org/10.1364/AO.56.00F134

[72] B. Horn, *Robot vision*. MIT press, 1986.

[73] W. K. Pratt, *Image Segmentation*. John Wiley & Sons, Inc., 2002, pp. 551–587.

[74] M. DaneshPanah and B. Javidi, "Segmentation of 3D holographic images using bivariate jointly distributed region snake," *Optics Express*, vol. 14, no. 12, pp. 5143–5153, Jun 2006. [Online]. Available: https://doi.org/10.1364/OE.14.005143

[75] L. Ma, H. Wang, Y. Li, and H. Jin, "Numerical reconstruction of digital holograms for three-dimensional shape measurement," *Journal of Optics A: Pure and Applied Optics*, vol. 6, no. 4, p. 396, 2004. [Online]. Available: https://doi.org/10.1088/1464-4258/6/4/016

[76] C. P. McElhinney, J. B. McDonald, A. Castro, Y. Frauel, B. Javidi, and T. J. Naughton, "Depth-independent segmentation of macroscopic three-dimensional objects encoded in single perspectives of digital holograms," *Optics Letters*, vol. 32, no. 10, pp. 1229–1231, May 2007. [Online]. Available: https://doi.org/10.1364/OL.32.001229

[77] K. M. Molony, J. P. Ryle, S. McDonnell, J. T. Sheridan, and T. J. Naughton, "Segmentation and visualization of digital in-line holographic microscopy of three-dimensional scenes using reconstructed intensity images," in *SPIE Optical Engineering + Applications*. International Society for Optics and Photonics, 2009, pp. 74 431F–74 431F. [Online]. Available: https://doi.org/10.1117/12.826832

[78] F. Yi, I. Moon, B. Javidi, D. Boss, and P. Marquet, "Automated segmentation of multiple red blood cells with digital holographic microscopy," *Journal of Biomedical Optics*, vol. 18, no. 2, pp. 026 006–026 006, 2013. [Online]. Available: https://doi.org/10.1117/1.JBO.18.2.026006

[79] E. Malkiel, J. N. Abras, and J. Katz, "Automated scanning and measurements of particle distributions within a holographic reconstructed volume," *Measurement Science and Technology*, vol. 15, no. 4, p. 601, 2004. [Online]. Available: https://doi.org/10.1088/0957-0233/15/4/001

[80] A. E. Mallahi, C. Minetti, and F. Dubois, "Automated three-dimensional detection and classification of living organisms using digital holographic microscopy with partial spatial coherent source: application to the monitoring of drinking water resources," *Applied Optics*, vol. 52, no. 1, pp. A68–A80, Jan 2013. [Online]. Available: https://doi.org/10.1364/AO.52.000A68

[81] J. Sheng, E. Malkiel, and J. Katz, "Digital holographic microscope for measuring three-dimensional particle distributions and motions," *Applied Optics*, vol. 45, no. 16, pp. 3893–3901, Jun 2006. [Online]. Available: https://doi.org/10.1364/AO.45.003893

[82] L. Orzó, Z. Göröcs, A. Fehér, and S. Tőkés, "In-line hologram segmentation for volumetric samples," *Applied Optics*, vol. 52, no. 1, pp. A45–A55, Jan 2013. [Online]. Available: https://doi.org/10.1364/AO.52.000A45

[83] F. Dubois, C. Minetti, O. Monnom, C. Yourassowsky, J.-C. Legros, and P. Kischel, "Pattern recognition with a digital holographic microscope working in partially coherent illumination," *Applied Optics*, vol. 41, no. 20, pp. 4108–4119, Jul 2002. [Online]. Available: https://doi.org/10.1364/AO.41.004108

[84] B. Javidi, I. Moon, S. Yeom, and E. Carapezza, "Three-dimensional imaging and recognition of microorganism using single-exposure on-line (SEOL) digital holography," *Optics Express*, vol. 13, no. 12, pp. 4492–4506, Jun 2005. [Online]. Available: https://doi.org/10.1364/OPEX.13.004492

[85] B. Javidi, S. Yeom, I. Moon, and M. Daneshpanah, "Real-time automated 3D sensing, detection, and recognition of dynamic biological micro-organic events," *Optics Express*, vol. 14, no. 9, pp. 3806–3829, May 2006. [Online]. Available: https://doi.org/10.1364/OE.14.003806

[86] I. Moon, M. Daneshpanah, B. Javidi, and A. Stern, "Automated three-dimensional identification and tracking of micro/nanobiological organisms by computational holographic microscopy," *Proceedings of the IEEE*, vol. 97, no. 6, pp. 990–1010, 2009. [Online]. Available: https://doi.org/10.1109/JPROC.2009.2017563

[87] I. Moon and B. Javidi, "3-D visualization and identification of biological microorganisms using partially temporal incoherent light in-line computational holographic imaging," *IEEE transactions on medical imaging*, vol. 27, no. 12, pp. 1782–1790, 2008. [Online]. Available: https://doi.org/10.1109/TMI.2008.927339

[88] A. Anand, V. K. Chhaniwal, N. R. Patel, and B. Javidi, "Automatic identification of malaria-infected RBC with digital holographic microscopy using correlation algorithms," *IEEE Photonics Journal*, vol. 4, no. 5, pp. 1456–1464, 2012. [Online]. Available: https://doi.org/10.1109/JPHOT.2012.2210199

[89] R. Liu, A. Anand, D. K. Dey, and B. Javidi, "Entropy-based clustering of embryonic stem cells using digital holographic microscopy," *Journal of Optical Society of America A*, vol. 31, no. 4, pp. 677–684, Apr 2014. [Online]. Available: https://doi.org/10.1364/JOSAA.31.000677

[90] P. Memmolo, L. Miccio, M. Paturzo, G. Di Caprio, G. Coppola, P. A. Netti, and P. Ferraro, "Recent advances in holographic 3D particle tracking," *Advances in Optics and Photonics*, vol. 7, no. 4, pp. 713–755, 2015. [Online]. Available: https://doi.org/10.1364/AOP.7.000713

[91] M. DaneshPanah and B. Javidi, "Tracking biological microorganisms in sequence of 3D holographic microscopy images," *Optics Express*, vol. 15, no. 17, pp. 10 761–10 766, Aug 2007. [Online]. Available: https://doi.org/10.1364/OE.15.010761

[92] Y.-S. Choi and S.-J. Lee, "Three-dimensional volumetric measurement of red blood cell motion using digital holographic microscopy," *Applied Optics*, vol. 48, no. 16, pp. 2983–2990, Jun 2009. [Online]. Available: https://doi.org/10.1364/AO.48.002983

[93] P. Langehanenberg, L. Ivanova, I. Bernhardt, S. Ketelhut, A. Vollmer, D. Dirksen, G. Georgiev, G. von Bally, and B. Kemper, "Automated three-dimensional tracking of living cells by digital holographic microscopy," *Journal of Biomedical Optics*, vol. 14, no. 1, pp. 014 018–014 018, 2009. [Online]. Available: https://doi.org/10.1117/1.3080133

[94] S. J. Lee, K. W. Seo, Y. S. Choi, and M. H. Sohn, "Three-dimensional motion measurements of free-swimming microorganisms using digital holographic microscopy," *Measurement Science and Technology*, vol. 22, no. 6, p. 064004, 2011. [Online]. Available: https://doi.org/10.1088%2F0957-0233%2F22%2F6%2F064004

[95] M. H. Sohn, K. W. Seo, Y. S. Choi, S. J. Lee, Y. S. Kang, and Y. S. Kang, "Determination of the swimming trajectory and speed of chain-forming dinoflagellate Cochlodinium polykrikoides with digital holographic particle tracking velocimetry," *Marine Biology*, vol. 158, no. 3, pp. 561–570, 2011. [Online]. Available: https://doi.org/10.1007/s00227-010-1581-7

[96] L. Leal-Taixé, M. Heydt, S. Weiße, A. Rosenhahn, and B. Rosenhahn, "Classification of swimming microorganisms motion patterns in 4D digital in-line holography data," in *Joint Pattern Recognition Symposium*. Springer, 2010, pp. 283–292. [Online]. Available: https://doi.org/10.1007/978-3-642-15986-2_29

[97] J. Persson, A. Mölder, S.-G. Pettersson, and K. Alm, "Cell motility studies using digital holographic microscopy," *Microscopy: Science, Technology, Applications and Education. Microscopy Series*, vol. 4, pp. 1063–1072, 2010.

[98] X. Yu, J. Hong, C. Liu, M. Cross, D. T. Haynie, and M. K. Kim, "Four-dimensional motility tracking of biological cells by digital holographic microscopy," *Journal of Biomedical Optics*, vol. 19, no. 4, pp. 045 001–045 001, 2014. [Online]. Available: https://doi.org/10.1117/1.JBO.19.4.045001

[99] L. Miccio, P. Memmolo, F. Merola, S. Fusco, V. Embrione, A. Paciello, M. Ventre, P. Netti, and P. Ferraro, "Particle tracking by full-field complex wavefront subtraction in digital holography microscopy," *Lab on a Chip*, vol. 14, no. 6, pp. 1129–1134, 2014. [Online]. Available: https://doi.org/10.1039/C3LC51104A

[100] E. Darakis, T. Khanam, A. Rajendran, V. Kariwala, T. J. Naughton, and A. K. Asundi, "Microparticle characterization using digital holography," *Chemical Engineering Science*, vol. 65, no. 2, pp. 1037–1044, 2010. [Online]. Available: https://doi.org/10.1016/j.ces.2009.09.057

[101] B. Kemper, A. Bauwens, A. Vollmer, S. Ketelhut, P. Langehanenberg, J. Müthing, H. Karch, and G. von Bally, "Label-free quantitative cell division monitoring of endothelial cells by digital holographic microscopy," *Journal of Biomedical Optics*, vol. 15, no. 3, pp. 036 009–036 009, 2010. [Online]. Available: https://doi.org/10.1117/1.3431712

[102] P. Bon, J. Savatier, M. Merlin, B. Wattellier, and S. Monneret, "Optical detection and measurement of living cell morphometric features with single-shot quantitative phase microscopy," *Journal of Biomedical Optics*, vol. 17, no. 7, pp. 0 760 041–0 760 047, 2012. [Online]. Available: https://doi.org/10.1117/1.JBO.17.7.076004

[103] N. Pavillon, J. Kühn, C. Moratal, P. Jourdain, C. Depeursinge, P. J. Magistretti, and P. Marquet, "Early cell death detection with digital holographic microscopy," *PLoS One*, vol. 7, no. 1, p. e30912, 2012. [Online]. Available: https://doi.org/10.1371/journal.pone.0030912

[104] J. Balvan, A. Krizova, J. Gumulec, M. Raudenska, Z. Sladek, M. Sedlackova, P. Babula, M. Sztalmachova, R. Kizek, R. Chmelik, and M. Masarik, "Multimodal holographic microscopy: distinction between apoptosis and oncosis," *PLoS One*, vol. 10, no. 3, p. e0121674, 2015. [Online]. Available: https://doi.org/10.1371/journal.pone.0121674

[105] F. Yi, C.-G. Lee, and I.-K. Moon, "Statistical analysis of 3D volume of red blood cells with different shapes via digital holographic microscopy," *Journal of the Optical Society of Korea*, vol. 16, no. 2, pp. 115–120, Jun 2012.

[106] T.-W. Su, L. Xue, and A. Ozcan, "High-throughput lensfree 3D tracking of human sperms reveals rare statistics of helical trajectories," *Proceedings of the National Academy of Sciences*, vol. 109, no. 40, pp. 16 018–16 022, 2012. [Online]. Available: https://doi.org/10.1073/pnas.1212506109

[107] P. Girshovitz and N. T. Shaked, "Generalized cell morphological parameters based on interferometric phase microscopy and their application to cell life cycle characterization," *Biomedical Optics Express*, vol. 3, no. 8, pp. 1757–1773, Aug 2012. [Online]. Available: https://doi.org/10.1364/BOE.3.001757

[108] J. Kühn, E. Shaffer, J. Mena, B. Breton, J. Parent, B. Rappaz, M. Chambon, Y. Emery, P. Magistretti, C. Depeursinge, P. Marquet, and G. Turcatti, "Label-free cytotoxicity screening assay by digital holographic microscopy," *ASSAY and Drug Development Technologies*, vol. 11, no. 2, pp. 101–107, 2013. [Online]. Available: https://doi.org/10.1089/adt.2012.476

[109] Y. Wang, Y. Yang, D. Wang, L. Ouyang, Y. Zhang, J. Zhao, and X. Wang, "Morphological measurement of living cells in methanol with digital holographic microscopy," *Computational and Mathematical Methods in Medicine*, vol. 2013, 2013. [Online]. Available: http://dx.doi.org/10.1155/2013/715843

[110] T. Pitkäaho, M. Niemelä, and V. Pitkäkangas, "Partially coherent digital in-line holographic microscopy in characterization of a microscopic target," *Applied Optics*, vol. 53, no. 15, pp. 3233–3240, May 2014. [Online]. Available: https://doi.org/10.1364/AO.53.003233

[111] B. Rappaz, B. Breton, E. Shaffer, and G. Turcatti, "Digital holographic microscopy: a quantitative label-free microscopy technique for phenotypic screening," *Combinatorial Chemistry & High Throughput Screening*, vol. 17, no. 1, pp. 80–88, 2014.

[112] D. Bettenworth, P. Lenz, P. Krausewitz, M. Brückner, S. Ketelhut, D. Domagk, and B. Kemper, "Quantitative stain-free and continuous multimodal monitoring of wound healing in vitro with digital holographic microscopy," *PLoS One*, vol. 9, no. 9, p. e107317, 2014. [Online]. Available: https://doi.org/10.1371/journal.pone.0107317

[113] P. Memmolo, L. Miccio, F. Merola, O. Gennari, P. A. Netti, and P. Ferraro, "3D morphometry of red blood cells by digital holography," *Cytometry part A*, vol. 85, no. 12, pp. 1030–1036, 2014. [Online]. Available: https://doi.org/10.1002/cyto.a.22570

[114] F. Yi, I. Moon, and Y. H. Lee, "Three-dimensional counting of morphologically normal human red blood cells via digital holographic microscopy," *Journal of Biomedical Optics*, vol. 20, no. 1, pp. 016 005–016 005, 2015. [Online]. Available: https://doi.org/10.1117/1.JBO.20.1.016005

[115] K. Jaferzadeh and I. Moon, "Quantitative investigation of red blood cell three-dimensional geometric and chemical changes in the storage lesion using digital holographic microscopy," *Journal of Biomedical Optics*, vol. 20, no. 11, pp. 111 218–111 218, 2015. [Online]. Available: https://doi.org/10.1117/1.JBO.20.11.111218

[116] B. Rappaz, I. Moon, F. Yi, B. Javidi, P. Marquet, and G. Turcatti, "Automated multi-parameter measurement of cardiomyocytes dynamics with digital holographic microscopy," *Optics Express*, vol. 23, no. 10, pp. 13 333–13 347, May 2015. [Online]. Available: https://doi.org/10.1364/OE.23.013333

[117] C. M. Bishop, *Pattern Recognition and Machine Learning*. Springer, 2006.

[118] Y. LeCun, B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard, and L. D. Jackel, "Backpropagation applied to handwritten zip code recognition," *Neural Computation*, vol. 1, no. 4, pp. 541–551, 1989. [Online]. Available: https://doi.org/10.1162/neco.1989.1.4.541

[119] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015. [Online]. Available: https://doi.org:/10.1038/nature14539

[120] R. Olga, D. Jia, S. Hao, K. Jonathan, S. Sanjeev, M. Sean, H. Zhiheng, K. Andrej, K. Aditya, B. Michael, B. A. C., and L. Fei-Fei, "ImageNet large scale visual recognition challenge," *International Journal of Computer Vision*, vol. 115, no. 3, pp. 211–252, 2015. [Online]. Available: https://doi.org/10.1007/s11263-015-0816-y

[121] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Advances in Neural Information Processing Systems*, vol. 25, pp. 1097–1105, 2012.

[122] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2015, pp. 1–9. [Online]. Available: https://doi.org/10.1109/CVPR.2015.7298594

[123] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2014.

[124] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016, pp. 770–778. [Online]. Available: https://doi.org/10.1109/CVPR.2016.90

[125] S. Zagoruyko and N. Komodakis, "Wide residual networks," *arXiv preprint arXiv:1605.07146*, 2016.

[126] G. E. Hinton, N. Srivastava, A. Krizhevsky, I. Sutskever, and R. R. Salakhutdinov, "Improving neural networks by preventing co-adaptation of feature detectors," *arXiv preprint arXiv:1207.0580*, 2012.

[127] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," *arXiv preprint arXiv:1502.03167*, 2015.

[128] D. Ciresan, A. Giusti, L. M. Gambardella, and J. Schmidhuber, "Deep neural networks segment neuronal membranes in electron microscopy images," in *Advances in Neural Information Processing Systems*, 2012, pp. 2843–2851.

[129] E. Shelhamer, J. Long, and T. Darrell, "Fully convolutional networks for semantic segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 4, pp. 640–651, 2016. [Online]. Available: https://doi.org/10.1109/TPAMI.2016.2572683

[130] V. Badrinarayanan, A. Kendall, and R. Cipolla, "Segnet: A deep convolutional encoder-decoder architecture for image segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 12, pp. 2481–2495, 2017. [Online]. Available: https://doi.org/10.1109/TPAMI.2016.2644615

[131] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *International Conference on Medical Image Computing and Computer-assisted Intervention*. Springer, 2015, pp. 234–241. [Online]. Available: https://doi.org/10.1007/978-3-319-24574-4_28

[132] L.-C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A. L. Yuille, "Semantic image segmentation with deep convolutional nets and fully connected crfs," *arXiv preprint arXiv:1412.7062*, 2014.

[133] L.-C. Chen, G. Papandreou, F. Schroff, and H. Adam, "Rethinking atrous convolution for semantic image segmentation," *arXiv preprint arXiv:1706.05587*, 2017.

[134] L.-C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A. L. Yuille, "Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 40, no. 4, pp. 834–848, 2018. [Online]. Available: https://doi.org/10.1109/TPAMI.2017.2699184

[135] L.-C. Chen, Y. Zhu, G. Papandreou, F. Schroff, and H. Adam, "Encoder-decoder with atrous separable convolution for semantic image segmentation," in *European Conference on Computer Vision*. Springer, 2018, pp. 833–851. [Online]. Available: https://doi.org/10.1007/978-3-030-01234-2_49

[136] P. Sermanet, D. Eigen, X. Zhang, M. Mathieu, R. Fergus, and Y. LeCun, "Overfeat: Integrated recognition, localization and detection using convolutional networks," *arXiv preprint arXiv:1312.6229*, 2013.

[137] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2014, pp. 580–587.

[138] R. Girshick, "Fast r-cnn," in *Proceedings of the IEEE international conference on computer vision*, 2015, pp. 1440–1448.

[139] S. Ren, K. He, R. Girshick, and J. Sun, "Faster r-cnn: Towards real-time object detection with region proposal networks," in *Advances in neural information processing systems*, 2015, pp. 91–99.

[140] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 779–788.

[141] P. Y. Simard, D. Steinkraus, and J. C. Platt, "Best practices for convolutional neural networks applied to visual document analysis," in *ICDAR*, vol. 3. Citeseer, 2003, pp. 958–962.

[142] B. E. Boser, I. M. Guyon, and V. N. Vapnik, "A training algorithm for optimal margin classifiers," in *Proceedings of the fifth annual workshop on Computational learning theory*. ACM, 1992, pp. 144–152. [Online]. Available: https://doi.org/10.1145/130385.130401

[143] C. Cortes and V. Vapnik, "Support-vector networks," *Machine Learning*, vol. 20, no. 3, pp. 273–297, 1995. [Online]. Available: http://dx.doi.org/10.1007/BF00994018

[144] J.-P. Vert, K. Tsuda, and B. Schölkpf, "A primer on kernel methods," in *Kernel methods in computation biology*, B. Schölkpf, K. Tsuda, and J.-P. Vert, Eds. Cambridge, Massachusetts: The MIT Press, 2004, ch. 1, pp. 1–42.

[145] M. Pal, "Multiclass approaches for support vector machine based land cover classification," *arXiv preprint arXiv:0802.2411*, 2008.

[146] J. A. Hartigan, *Clustering algorithms*. John Wiley & Sons, Inc, 1975.

[147] L. Breiman, "Random forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001. [Online]. Available: https://doi.org/10.1023/A:1010933404324

[148] ——, "Bagging predictors," *Machine Learning*, vol. 24, no. 2, pp. 123–140, 1996. [Online]. Available: https://doi.org/10.1007/BF00058655

[149] N. Orlov, J. Johnston, T. Macura, L. Shamir, and I. Goldberg, *Computer vision for microscopy applications*. INTECH Open Access Publisher, 2007.

[150] F. Ghaznavi, A. Evans, A. Madabhushi, and M. Feldman, "Digital imaging in pathology: whole-slide imaging and beyond," *Annual Review of Pathology: Mechanisms of Disease*, vol. 8, pp. 331–359, 2013. [Online]. Available: https://doi.org/10.1146/annurev-pathol-011811-120902

[151] J. Griffin and D. Treanor, "Digital pathology in clinical use: where are we now and what is holding us back?" *Histopathology*, vol. 70, no. 1, pp. 134–145, 2017. [Online]. Available: https://doi.org/10.1111/his.12993

[152] A. H. Beck, A. R. Sangoi, S. Leung, R. J. Marinelli, T. O. Nielsen, M. J. van de Vijver, R. B. West, M. van de Rijn, and D. Koller, "Systematic analysis of breast cancer morphology uncovers stromal features associated with survival," *Science Translational Medicine*, vol. 3, no. 108, pp. 108ra113–108ra113, 2011. [Online]. Available: https://doi.org/10.1126/scitranslmed.3002564

[153] N. Linder, J. Konsti, R. Turkki, E. Rahtu, M. Lundin, S. Nordling, C. Haglund, T. Ahonen, M. Pietikäinen, and J. Lundin, "Identification of tumor epithelium and stroma in tissue microarrays using texture analysis," *Diagnostic Pathology*, vol. 7, no. 1, p. 22, 2012. [Online]. Available: https://doi.org/10.1186/1746-1596-7-22

[154] E. Cosatto, P.-F. Laquerre, C. Malon, H.-P. Graf, A. Saito, T. Kiyuna, A. Marugame, and K. Kamijo, "Automated gastric cancer diagnosis on h&e-stained sections; ltraining a classifier on a large scale with multiple instance machine learning," in *SPIE Medical Imaging*. International Society for Optics and Photonics, 2013, pp. 867 605–867 605. [Online]. Available: https://doi.org/10.1117/12.2007047

[155] A. A. Cruz-Roa, J. E. A. Ovalle, A. Madabhushi, and F. A. G. Osorio, "A deep learning architecture for image representation, visual interpretability and automated basal-cell carcinoma cancer detection," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2013, pp. 403–410. [Online]. Available: https://doi.org/10.1007/978-3-642-40763-5_50

[156] E. Rexhepaj, M. Agnarsdóttir, J. Bergman, P.-H. Edqvist, M. Bergqvist, M. Uhlén, W. M. Gallagher, E. Lundberg, and F. Ponten, "A texture based pattern recognition approach to distinguish melanoma from non-melanoma cells in histopathological tissue microarray sections," *PLoS One*, vol. 8, no. 5, p. e62070, 2013. [Online]. Available: https://doi.org/10.1371/journal.pone.0062070,

[157] H. Wang, A. Cruz-Roa, A. Basavanhally, H. Gilmore, N. Shih, M. Feldman, J. Tomaszewski, F. Gonzalez, and A. Madabhushi, "Mitosis detection in breast cancer pathology images by combining handcrafted and convolutional neural network features," *Journal of Medical Imaging*, vol. 1, no. 3, p. 034003, 2014. [Online]. Available: http://dx.doi.org/10.1117/1.JMI.1.3.034003

[158] J. S. Lewis, S. Ali, J. Luo, W. L. Thorstad, and A. Madabhushi, "A quantitative histomorphometric classifier (QuHbIC) identifies aggressive versus indolent p16-positive oropharyngeal squamous cell carcinoma," *The American Journal of Surgical Pathology*, vol. 38, no. 1, p. 128, 2014. [Online]. Available: https://doi.org/10.1097/PAS.0000000000000086

[159] J. Arevalo, A. Cruz-Roa, V. Arias, E. Romero, and F. A. González, "An unsupervised feature learning framework for basal cell carcinoma image analysis," *Artificial Intelligence in Medicine*, vol. 64, no. 2, pp. 131–145, 2015. [Online]. Available: https://doi.org/10.1016/j.artmed.2015.04.004

[160] J. Xu, X. Luo, G. Wang, H. Gilmore, and A. Madabhushi, "A deep convolutional neural network for segmenting and classifying epithelial and stromal regions in histopathological images," *Neurocomputing*, vol. 191, pp. 214–223, 2016. [Online]. Available: https://doi.org/10.1016/j.neucom.2016.01.034

[161] K.-H. Yu, C. Zhang, G. J. Berry, R. B. Altman, C. Ré, D. L. Rubin, and M. Snyder, "Predicting non-small cell lung cancer prognosis by fully automated microscopic pathology image features," *Nature Communications*, vol. 7, 2016. [Online]. Available: https://doi.org/10.1038/ncomms12474

[162] Y. Yamamoto, A. Saito, A. Tateishi, H. Shimojo, H. Kanno, S. Tsuchiya, K. ichi Ito, E. Cosatto, H. P. Graf, R. R. Moraleda, R. Eils, and N. Grabe, "Quantitative diagnosis of breast tumors by morphometric classification of microenvironmental myoepithelial cells using a machine learning approach," *Scientific Reports*, vol. 7, 2017. [Online]. Available: https://doi.org/10.1038/srep46732

[163] A. Cruz-Roa, H. Gilmore, A. Basavanhally, M. Feldman, S. Ganesan, N. N. Shih, J. Tomaszewski, F. A. González, and A. Madabhushi, "Accurate and reproducible invasive breast cancer detection in whole-slide images: A deep learning approach for quantifying tumor extent," *Scientific Reports*, vol. 7, 2017. [Online]. Available: https://doi.org/10.1038/srep46450

[164] T. G. Dietterich, R. H. Lathrop, and T. Lozano-Pérez, "Solving the multiple instance problem with axis-parallel rectangles," *Artificial Intelligence*, vol. 89, no. 1, pp. 31–71, 1997. [Online]. Available: https://doi.org/10.1016/S0004-3702(96)00034-3

[165] D. C. Cireşan, A. Giusti, L. M. Gambardella, and J. Schmidhuber, "Mitosis detection in breast cancer histology images with deep neural networks," in *International Conference on Medical Image Computing and Computer-assisted Intervention*. Springer, 2013, pp. 411–418. [Online]. Available: https://doi.org/10.1007/978-3-642-40763-5_51

[166] B. Neumann, M. Held, U. Liebel, H. Erfle, P. Rogers, R. Pepperkok, and J. Ellenberg, "High-throughput RNAi screening by time-lapse imaging of live human cells," *Nature Methods*, vol. 3, no. 5, pp. 385–390, 2006. [Online]. Available: https://doi.org/10.1038/nmeth876

[167] S. Osowski, R. Siroic, T. Markiewicz, and K. Siwek, "Application of support vector machine and genetic algorithm for improved blood cell recognition," *IEEE Transactions on Instrumentation and Measurement*, vol. 58, no. 7, pp. 2159–2168, 2009. [Online]. Available: https://doi.org/10.1109/TIM.2008.2006726

[168] J. W. Han, T. P. Breckon, D. A. Randell, and G. Landini, "The application of support vector machine classification to detect cell nuclei for automated microscopy," *Machine Vision and Applications*, vol. 23, no. 1, pp. 15–24, 2012. [Online]. Available: https://doi.org/10.1007/s00138-010-0275-y

[169] R. Stotzka, R. Männer, P. H. Bartels, and D. Thompson, "A hybrid neural and statistical classifier system for histopathologic grading of prostatic lesions," *Analytical and quantitative cytology and histology/the International Academy of Cytology [and] American Society of Cytology*, vol. 17, no. 3, pp. 204–218, 1995.

[170] M. Binder, H. Kittler, A. Seeber, A. Steiner, H. Pehamberger, and K. Wolff, "Epiluminescence microscopy-based classification of pigmented skin lesions using computerized image analysis and an artificial neural network," *Melanoma Research*, vol. 8, no. 3, pp. 261–266, 1998.

[171] M. V. Boland and R. F. Murphy, "A neural network classifier capable of recognizing the patterns of all major subcellular structures in fluorescence microscope images of HeLa cells," *Bioinformatics*, vol. 17, no. 12, pp. 1213–1223, 2001. [Online]. Available: https://doi.org/10.1093/bioinformatics/17.12.1213

[172] E. Jurrus, A. R. Paiva, S. Watanabe, J. R. Anderson, B. W. Jones, R. T. Whitaker, E. M. Jorgensen, R. E. Marc, and T. Tasdizen, "Detection of neuron membranes in electron microscopy images using a serial neural network architecture," *Medical Image Analysis*, vol. 14, no. 6, pp. 770–783, 2010.

[173] N. Sharma, A. K. Ray, S. Sharma, K. Shukla, S. Pradhan, and L. M. Aggarwal, "Segmentation and classification of medical images using texture-primitive features: Application of BAM-type artificial neural network," *Journal of Medical Physics*, vol. 33, no. 3, p. 119, 2008. [Online]. Available: https://dx.doi.org/10.4103%2F0971-6203.42763

[174] D. Cireşan, A. Giusti, L. M. Gambardella, and J. Schmidhuber, "Deep neural networks segment neuronal membranes in electron microscopy images," *Advances in Neural Information Processing Systems*, vol. 25, pp. 2843–2851, 2012.

[175] P. Prentašić, M. Heisler, Z. Mammo, S. Lee, A. Merkur, E. Navajas, M. F. Beg, M. Šarunić, and S. Lončarić, "Segmentation of the foveal microvasculature using deep learning networks," *Journal of Biomedical Optics*, vol. 21, no. 7, p. 075008, 2016. [Online]. Available: http://dx.doi.org/10.1117/1.JBO.21.7.075008

[176] H. Rezaeilouyeh, A. Mollahosseini, and M. H. Mahoor, "Microscopic medical image classification framework via deep learning and shearlet transform," *Journal of Medical Imaging*, vol. 3, no. 4, p. 044501, 2016. [Online]. Available: http://dx.doi.org/10.1117/1.JMI.3.4.044501

[177] O. Z. Kraus, J. L. Ba, and B. J. Frey, "Classifying and segmenting microscopy images with deep multiple instance learning," *Bioinformatics*, vol. 32, no. 12, pp. i52–i59, 2016. [Online]. Available: https://doi.org/10.1093/bioinformatics/btw252

[178] P. Eulenberg, N. Koehler, T. Blasi, A. Filby, A. E. Carpenter, P. Rees, F. J. Theis, and F. A. Wolf, "Deep learning for imaging flow cytometry: Cell cycle analysis of Jurkat cells," *bioRxiv*, p. 081364, 2016. [Online]. Available: https://doi.org/10.1038/ncomms10256

[179] O. Dürr and B. Sick, "Single-cell phenotype classification using deep convolutional neural networks," *Journal of Biomolecular Screening*, p. 1087057116631284, 2016. [Online]. Available: https://doi.org/10.1177%2F1087057116631284

[180] G. Gopakumar, K. H. Babu, D. Mishra, S. S. Gorthi, and G. R. K. S. Subrahmanyam, "Cytopathological image analysis using deep-learning networks in microfluidic microscopy," *Journal of the Optical Society of America A*, vol. 34, no. 1, pp. 111–121, Jan 2017. [Online]. Available: http://doi.org/10.1364/JOSAA.34.000111

[181] O. Z. Kraus, B. T. Grys, J. Ba, Y. Chong, B. J. Frey, C. Boone, and B. J. Andrews, "Automated analysis of high-content microscopy data with deep learning," *Molecular Systems Biology*, vol. 13, no. 4, p. 924, 2017. [Online]. Available: https://doi.org/10.15252/msb.20177551

[182] A. Abdolmanafi, L. Duong, N. Dahdah, and F. Cheriet, "Deep feature learning for automatic tissue classification of coronary artery using optical coherence tomography," *Biomedical Optics Express*, vol. 8, no. 2, pp. 1203–1220, 2017. [Online]. Available: https://doi.org/10.1364/BOE.8.001203

[183] S. P. K. Karri, D. Chakraborty, and J. Chatterjee, "Transfer learning based classification of optical coherence tomography images with diabetic macular edema and dry age-related macular degeneration," *Biomedical Optics Express*, vol. 8, no. 2, pp. 579–592, Feb 2017. [Online]. Available: http://doi.org/10.1364/BOE.8.000579

[184] Y. Rivenson, Z. Göröcs, H. Günaydin, Y. Zhang, H. Wang, and A. Ozcan, "Deep learning microscopy," *Optica*, vol. 4, no. 11, pp. 1437–1443, 2017. [Online]. Available: https://doi.org/10.1364/OPTICA.4.001437

[185] Y. Rivenson, H. Wang, Z. Wei, Y. Zhang, H. Gunaydin, and A. Ozcan, "Deep learning-based virtual histology staining using auto-fluorescence of label-free tissue," *arXiv preprint arXiv:1803.11293*, 2018.

[186] R. Farjam, H. Soltanian-Zadeh, R. A. Zoroofi, and K. Jafari-Khouzani, "Tree-structured grading of pathological images of prostate," in *Medical Imaging*. International Society for Optics and Photonics, 2005, pp. 840–851. [Online]. Available: https://doi.org/10.1117/12.596068

[187] L. E. Boucheron, Z. Bi, N. R. Harvey, B. Manjunath, and D. L. Rimm, "Utility of multispectral imaging for nuclear classification of routine clinical histopathology imagery," *BMC Cell Biology*, vol. 8, no. 1, p. S8, 2007. [Online]. Available: https://doi.org/10.1186/1471-2121-8-S1-S8

[188] T. R. Jones, A. E. Carpenter, M. R. Lamprecht, J. Moffat, S. J. Silver, J. K. Grenier, A. B. Castoreno, U. S. Eggert, D. E. Root, and P. Golland, "Scoring diverse cellular morphologies in image-based screens with iterative feedback and machine learning," *Proceedings of the National Academy of Sciences*, vol. 106, no. 6, pp. 1826–1831, 2009. [Online]. Available: https://doi.org/10.1073/pnas.0808843106

[189] D. H. Theriault, M. L. Walker, J. Y. Wong, and M. Betke, "Cell morphology classification and clutter mitigation in phase-contrast microscopy images using machine learning," *Machine Vision and Applications*, vol. 23, no. 4, pp. 659–673, 2012. [Online]. Available: https://doi.org/10.1007/s00138-011-0345-9

[190] H. C. Koydemir, Z. Gorocs, D. Tseng, B. Cortazar, S. Feng, R. Y. L. Chan, J. Burbano, E. McLeod, and A. Ozcan, "Rapid imaging, detection and quantification of Giardia lamblia cysts using mobile-phone based fluorescent microscopy and machine learning," *Lab on a Chip*, vol. 15, no. 5, pp. 1284–1293, 2015. [Online]. Available: https://doi.org/10.1039/C4LC01358A

[191] B. T. Grys, D. S. Lo, N. Sahin, O. Z. Kraus, Q. Morris, C. Boone, and B. J. Andrews, "Machine learning and computer vision approaches for phenotypic profiling," *Journal of Cell Biology*, pp. jcb–201 610 026, 2016. [Online]. Available: https://doi.org/10.1083/jcb.201610026

[192] T. B. Lannin, F. I. Thege, and B. J. Kirby, "Comparison and optimization of machine learning methods for automated classification of circulating tumor cells," *Cytometry Part A*, vol. 89, no. 10, pp. 922–931, 2016. [Online]. Available: https://doi.org/10.1002/cyto.a.22993

[193] J. Friedman, T. Hastie, and R. Tibshirani, "Additive logistic regression: a statistical view of boosting (with discussion and a rejoinder by the authors)," *The Annals of Statistics*, vol. 28, no. 2, pp. 337–407, 2000. [Online]. Available: https://doi.org/doi:10.1214/aos/1016218223

[194] A. Feizi, Y. Zhang, A. Greenbaum, A. Guziak, M. Luong, R. Y. L. Chan, B. Berg, H. Ozkan, W. Luo, Y. Wu, and A. Ozcan, "Rapid, portable and cost-effective yeast cell viability and concentration analysis using lensfree on-chip microscopy and machine learning," *Lab on a Chip*, vol. 16, no. 22, pp. 4350–4358, 2016. [Online]. Available: https://dx.doi.org/10.1039/C6LC00976J

[195] H. S. Park, M. T. Rinehart, K. A. Walzer, J.-T. A. Chi, and A. Was, "Automated detection of P. falciparum using machine learning algorithms with quantitative phase images of unstained cells," *PLos One*, vol. 11, no. 9, p. e0163045, 2016. [Online]. Available: https://doi.org/10.1371/journal.pone.0163045

[196] D. Roitshtain, L. Wolbromsky, E. Bal, H. Greenspan, L. L. Satterwhite, and N. T. Shaked, "Quantitative phase microscopy spatial signatures of cancer cells," *Cytometry Part A*, vol. 91, no. 5, pp. 482–493, 2017. [Online]. Available: https://doi.org/10.1002/cyto.a.23100

[197] Y. Jo, S. Park, J. Jung, J. Yoon, H. Joo, M. hyeok Kim, S.-J. Jo, M. C. Choi, S. Y. Lee, and Y. Park, "Holographic deep learning for rapid optical screening of anthrax spores," *bioRxiv*, p. 109108, 2017. [Online]. Available: https://dx.doi.org/10.1126/sciadv.1700606

[198] Y. Rivenson, Y. Zhang, H. Günaydın, D. Teng, and A. Ozcan, "Phase recovery and holographic image reconstruction using deep learning in neural networks," *Light: Science & Applications*, vol. 7, no. 2, p. 17141, 2018. [Online]. Available: https://doi.org/10.1038/lsa.2017.141

[199] Y. Wu, Y. Rivenson, Y. Zhang, Z. Wei, H. Günaydin, X. Lin, and A. Ozcan, "Extended depth-of-field in holographic imaging using deep-learning-based autofocusing and phase recovery," *Optica*, vol. 5, no. 6, pp. 704–710, 2018. [Online]. Available: https://doi.org/10.1364/OPTICA.5.000704

[200] Z. Göröcs, M. Tamamitsu, V. Bianco, P. Wolf, S. Roy, K. Shindo, K. Yanny, Y. Wu, H. C. Koydemir, Y. Rivenson *et al.*, "A deep learning-enabled portable imaging flow cytometer for cost-effective, high-throughput, and label-free analysis of natural water samples," *Light: Science & Applications*, vol. 7, no. 1, p. 66, 2018. [Online]. Available: https://doi.org/10.1038/s41377-018-0067-0

[201] Y. Wu, A. Calis, Y. Luo, C. Chen, M. Lutton, Y. Rivenson, X. Lin, H. Ceylan Koydemir, Y. Zhang, H. Wang *et al.*, "Label-free bio-aerosol sensing using mobile microscopy and deep learning," *ACS Photonics*, vol. 0, no. 0, p. null, 2018. [Online]. Available: https://doi.org/10.1021/acsphotonics.8b01109

[202] Z. Ren, Z. Xu, and E. Y. Lam, "Learning-based nonparametric autofocusing for digital holography," *Optica*, vol. 5, no. 4, pp. 337–344, 2018. [Online]. Available: https://doi.org/10.1364/OPTICA.5.000337

[203] M. D. Hannel, A. Abdulali, M. O'Brien, and D. G. Grier, "Machine-learning techniques for fast and accurate feature localization in holograms of colloidal particles," *Optics Express*, vol. 26, no. 12, pp. 15 221–15 231, 2018. [Online]. Available: https://doi.org/10.1364/OE.26.015221

[204] G. Zhang, T. Guan, Z. Shen, X. Wang, T. Hu, D. Wang, Y. He, and N. Xie, "Fast phase retrieval in off-axis digital holographic microscopy through deep learning," *Optics Express*, vol. 26, no. 15, pp. 19 388–19 405, 2018. [Online]. Available: https://doi.org/10.1364/OE.26.019388

[205] J. Lee, W. Jeong, K. Son, W. Jeon, and H. Yang, "Autofocusing using deep learning in off-axis digital holography," in *Imaging and Applied Optics 2018 (3D, AO, AIO, COSI, DH, IS, LACSEA, LS&C, MATH, pcAOP)*. Optical Society of America, 2018, p. DTh1C.4. [Online]. Available: https://doi.org/10.1364/DH.2018.DTh1C.4

[206] T. Nguyen, V. Bui, and G. Nehmetallah, "3D optical diffraction tomography using deep learning," in *Imaging and Applied Optics 2018 (3D, AO, AIO, COSI, DH, IS, LACSEA, LS&C, MATH, pcAOP)*. Optical Society of America, 2018, p. DW2F.4. [Online]. Available: https://doi.org/10.1364/DH.2018.DW2F.4

[207] Z. Luo, A. Yurt, R. Stahl, A. Lambrechts, V. Reumers, D. Braeken, and L. Lagae, "Pixel super-resolution for lens-free holographic microscopy using deep learning neural networks," *Optics Express*, vol. 27, no. 10, pp. 13 581–13 595, May 2019. [Online]. Available: https://doi.org/10.1364/OE.27.013581

[208] C. Trujillo and J. Garcia-Sucerquia, "Automatic detection and counting of phase objects in raw holograms of digital holographic microscopy via deep learning," *Optics and Lasers in Engineering*, vol. 120, pp. 13–20, 2019. [Online]. Available: https://doi.org/10.1016/j.optlaseng.2019.02.010

[209] Z. Ren, Z. Xu, and E. Y. Lam, "End-to-end deep learning framework for digital holographic reconstruction," *Advanced Photonics*, vol. 1, no. 1, p. 016004, 2019.

[210] H. Byeon, T. Go, and S. J. Lee, "Deep learning-based digital in-line holographic microscopy for high resolution with extended field of view," *Optics & Laser Technology*, vol. 113, pp. 77–86, 2019. [Online]. Available: https://doi.org/10.1016/j.optlastec.2018.12.014

[211] T. H. Nguyen, S. Sridharan, V. Macias, A. K. Balla, M. N. Do, and G. Popescu, "Prostate cancer diagnosis using quantitative phase imaging and machine learning algorithms," in *SPIE BiOS*. International Society for Optics and Photonics, 2015, pp. 933 619–933 619. [Online]. Available: https://doi.org/10.1117/12.2080321

[212] C. L. Chen, A. Mahjoubfar, L.-C. Tai, I. K. Blaby, A. Huang, K. R. Niazi, and B. Jalali, "Deep learning in label-free cell classification," *Scientific Reports*, vol. 6, p. 21471, 2016. [Online]. Available: http://dx.doi.org/10.1038/srep21471

[213] B. Guo, C. Lei, H. Kobayashi, T. Ito, Y. Yalikun, Y. Jiang, Y. Tanaka, Y. Ozeki, and K. Goda, "High-throughput, label-free, single-cell, microalgal lipid screening by machine-learning-equipped optofluidic time-stretch quantitative phase microscopy," *Cytometry Part A*, 2017. [Online]. Available: https://doi.org/10.1002/cyto.a.23084

[214] S. Li, M. Deng, J. Lee, A. Sinha, and G. Barbastathis, "Imaging through glass diffusers using densely connected convolutional networks," *Optica*, vol. 5, no. 7, pp. 803–813, Jul 2018. [Online]. Available: http://www.osapublishing.org/optica/abstract.cfm?URI=optica-5-7-803

[215] A. Sinha, J. Lee, S. Li, and G. Barbastathis, "Lensless computational imaging through deep learning," *Optica*, vol. 4, no. 9, pp. 1117–1125, Sep 2017. [Online]. Available: https://doi.org/10.1364/OPTICA.4.001117

[216] S. Li and G. Barbastathis, "Spectral pre-modulation of training examples enhances the spatial resolution of the phase extraction neural network (PhENN)," *Optics Express*, vol. 26, no. 22, pp. 29 340–29 352, Oct 2018. [Online]. Available: http://www.opticsexpress.org/abstract.cfm?URI=oe-26-22-29340

[217] S. Li, A. Sinha, J. Lee, and G. Barbastathis, "Cell imaging by phase extraction neural network (PhENN)," in *Imaging and Applied Optics 2018 (3D, AO, AIO, COSI, DH, IS, LACSEA, LS&C, MATH, pcAOP)*. Optical Society of America, 2018, p. CW2E.6. [Online]. Available: http://www.osapublishing.org/abstract.cfm?URI=COSI-2018-CW2E.6

[218] L. Zheng, K. Yu, S. Cai, Y. Wang, B. Zeng, and M. Xu, "Lung cancer diagnosis with quantitative DIC microscopy and a deep convolutional neural network," *Biomedical Optics Express*, vol. 10, no. 5, pp. 2446–2456, May 2019. [Online]. Available: https://doi.org/10.1364/BOE.10.002446

[219] Ovizio Imaging Systems NV/SA, "Machine learning explained," 2020, [Online; accessed 24-May-2020]. [Online]. Available: http://www.ovizio.com/_dbfiles/lacentrale_files/500/505/Ovizio-Infographic-machine-learning.pdf

[220] Lyncée tec SA, "Koala acquisition & analysis software," 2020, [Online; accessed 24-May-2020]. [Online]. Available: https://www.lynceetec.com/koala-acquisition-analysis/

[221] F. Montfort, Y. Emery, E. Solanas, E. Cuche, N. Aspert, P. Marquet, C. Joris, J. Kühn, and C. Depeursing, "Surface roughness parameters measurements by digital holographic microscopy (dhm)," in *Third International Symposium on Precision Mechanical Measurements*, vol. 6280. International Society for Optics and Photonics, 2006, p. 62800V.

[222] B. Janicke, A. Kårsnäs, P. Egelberg, and K. Alm, "Label-free high temporal resolution assessment of cell proliferation using digital holographic microscopy," *Cytometry Part A*, vol. 91, no. 5, pp. 460–469, 2017.

[223] 4-Deep Inc., *Octopus Software User Guide*, 2nd ed. Halifax Nova Scotia, Canada: 4-Deep Inc., 2018.

[224] ——, *Swordfish Software User Guide*, 2nd ed. Halifax Nova Scotia, Canada: 4-Deep Inc., 2018.

[225] ——, *Stingray Software User Guide*, 2nd ed. Halifax Nova Scotia, Canada: 4-Deep Inc., 2018.

[226] R. I. Freshney and M. G. Freshney, *Culture of epithelial cells*. John Wiley & Sons, 2004, vol. 10.

[227] E. Haug, O. Sand, Ø. V. Sjaastad, and K. C. Toverud, *Ihmisen fysiologia*. WSOY, 1995.

[228] A. Manninen, "Epithelial polarity–generating and integrating signals from the ECM with integrins," *Experimental Cell Research*, vol. 334, no. 2, pp. 337–349, 2015. [Online]. Available: https://doi.org/10.1016/j.yexcr.2015.01.003

[229] B. Tanos and E. Rodriguez-Boulan, "The epithelial polarity program: machineries involved and their hijacking by cancer," *Oncogene*, vol. 27, no. 55, pp. 6939–6957, 2008. [Online]. Available: https://doi.org/10.1038/onc.2008.345

[230] K. Zhang, S. Myllymäki, P. Gao, R. Devarajan, V. Kytölä, M. Nykter, G. Wei, and A. Manninen, "Oncogenic K-Ras upregulates ITGA6 expression via FOSL1 to induce anoikis resistance and synergizes with αv-class integrins to promote EMT," *Oncogene*, vol. 36, no. 41, p. 5681, 2017. [Online]. Available: https://doi.org/10.1038/onc.2017.177

[231] A. A. Samatar and P. I. Poulikakos, "Targeting RAS–ERK signalling in cancer: promises and challenges," *Nature Reviews Drug discovery*, vol. 13, no. 12, pp. 928–942, 2014. [Online]. Available: https://doi.org/10.1038/nrd4281

[232] T. Sakuma, M. A. Barry, and Y. Ikeda, "Lentiviral vectors: basic to translational," *Biochemical Journal*, vol. 443, no. 3, pp. 603–618, 2012. [Online]. Available: https://doi.org/10.1042/BJ20120146

[233] A. Y. Liu and L. D. True, "Characterization of prostate cell types by CD cell surface molecules," *The American Journal of Pathology*, vol. 160, no. 1, pp. 37–43, 2002. [Online]. Available: https://doi.org/10.1016/S0002-9440(10)64346-5

[234] W. C. R. Fund. (2018) Prostate cancer statistics. [Online]. Available: https://www.wcrf.org/dietandcancer/cancer-trends/prostate-cancer-statistics

[235] W. H. Organisation. (2018) Cancer key facts. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/cancer

[236] M. J. Barry, "Prostate-specific–antigen testing for early diagnosis of prostate cancer," *New England Journal of Medicine*, vol. 344, no. 18, pp. 1373–1377, 2001.

[237] F. C. Hamdy, J. L. Donovan, J. A. Lane, M. Mason, C. Metcalfe, P. Holding, M. Davis, T. J. Peters, E. L. Turner, R. M. Martin *et al.*, "10-year outcomes after monitoring, surgery, or radiotherapy for localized prostate cancer," *New England Journal of Medicine*, vol. 375, no. 15, pp. 1415–1424, 2016. [Online]. Available: https://doi.org/10.1056/NEJMoa1606220

[238] R. Tal, H. H. Alphs, P. Krebs, C. J. Nelson, and J. P. Mulhall, "Erectile function recovery rate after radical prostatectomy: A meta-analysis," *The Journal of Sexual Medicine*, vol. 6, no. 9, pp. 2538–2546, 2009.

[239] V. Ficarra, G. Novara, R. C. Rosen, W. Artibani, P. R. Carroll, A. Costello, M. Menon, F. Montorsi, V. R. Patel, J.-U. Stolzenburg, H. V. der Poel, T. G.Wilson, F. Zattoni, and A. Mottrie, "Systematic review and meta-analysis of studies reporting urinary continence recovery after robot-assisted radical prostatectomy," *European Urology*, vol. 62, no. 3, pp. 405–417, 2012.

[240] F. Pampaloni, E. G. Reynaud, and E. H. Stelzer, "The third dimension bridges the gap between cell culture and live tissue," *Nature Reviews Molecular Cell Biology*, vol. 8, no. 10, pp. 839–845, 2007. [Online]. Available: https://doi.org/10.1038/nrm2236

[241] T. H. Nguyen, M. E. Kandel, M. Rubessa, M. B. Wheeler, and G. Popescu, "Gradient light interference microscopy for 3D imaging of unlabeled specimens," *Nature Communications*, vol. 8, no. 1, p. 210, 2017. [Online]. Available: https://doi.org/10.1038/s41467-017-00190-7

[242] J. Huisken, J. Swoger, F. D. Bene, J. Wittbrodt, and E. H. Stelzer, "Optical sectioning deep inside live embryos by selective plane illumination microscopy," *Science*, vol. 305, no. 5686, pp. 1007–1009, 2004. [Online]. Available: https://doi.org/10.1126/science.1100035

[243] T. Deckers, T. Lambrechts, S. Viazzi, G. N. Hall, I. Papantoniou, V. Bloemen, and J.-M. Aerts, "High-throughput image-based monitoring of cell aggregation and microspheroid formation," *PLoS One*, vol. 13, no. 6, p. e0199092, 2018. [Online]. Available: https://doi.org/10.1371/journal.pone.0199092

[244] E. Marques, T. Peltola, S. Kaski, and J. Klefström, "Phenotype-driven identification of epithelial signalling clusters," *Scientific Reports*, vol. 8, no. 1, p. 4034, 2018. [Online]. Available: https://doi.org/10.1038/s41598-018-22293-x

[245] R. Barer, "Determination of dry mass, thickness, solid and water concentration in living cells," *Nature*, vol. 172, no. 4389, p. 1097, 1953. [Online]. Available: https://doi.org/10.1038/1721097a0

[246] S. Van der Walt, J. L. Schönberger, J. Nunez-Iglesias, F. Boulogne, J. D. Warner, N. Yager, E. Gouillart, and T. Yu, "scikit-image: image processing in python," *PeerJ*, vol. 2, p. e453, 2014. [Online]. Available: https://doi.org/10.7717/peerj.453

[247] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, "Scikit-learn: Machine learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.

[248] A. Paszke, S. Gross, S. Chintala, G. Chanan, E. Yang, Z. DeVito, Z. Lin, A. Desmaison, L. Antiga, and A. Lerer, "Automatic differentiation in pytorch," in *Advances in Neural Information Processing Systems*. NIPS, 2017.

[249] V. Härmä, H.-P. Schukov, A. Happonen, I. Ahonen, J. Virtanen, H. Siitari, M. Åkerfelt, J. Lötjönen, and M. Nees, "Quantification of dynamic morphological drug responses in 3D organotypic cell cultures by automated image analysis," *PLoS One*, vol. 9, no. 5, p. e96426, 2014. [Online]. Available: https://doi.org/10.1371/journal.pone.0096426

[250] R. H. Byrd, P. Lu, J. Nocedal, and C. Zhu, "A limited memory algorithm for bound constrained optimization," *SIAM Journal on Scientific Computing*, vol. 16, no. 5, pp. 1190–1208, 1995. [Online]. Available: https://doi.org/10.1137/0916069

[251] P. Memmolo, M. Iannone, M. Ventre, P. A. Netti, A. Finizio, M. Paturzo, and P. Ferraro, "On the holographic 3D tracking of in vitro cells characterized by a highly-morphological change," *Optics Express*, vol. 20, no. 27, pp. 28 485–28 493, 2012. [Online]. Available: https://doi.org/10.1364/OE.20.028485

[252] M. Lyu, C. Yuan, D. Li, and G. Situ, "Fast autofocusing in digital holography using the magnitude differential," *Applied Optics*, vol. 56, no. 13, pp. F152–F157, May 2017. [Online]. Available: https://doi.org/10.1364/AO.56.00F152

[253] K. He, X. Zhang, S. Ren, and J. Sun, "Delving deep into rectifiers: Surpassing human-level performance on imagenet classification," in *Proceedings of the IEEE International Conference on Computer Vision*, 2015, pp. 1026–1034.

[254] D. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.

[255] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, D. Batra *et al.*, "Grad-CAM: Visual explanations from deep networks via gradient-based localization," in *ICCV*, 2017, pp. 618–626.

[256] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 9, no. 1, pp. 62–66, 1979. [Online]. Available: https://doi.org/10.1109/tsmc.1979.4310076

[257] C. S. Vikram and M. L. Billet, "Far-field holography at non-image planes for size analysis of small particles," *Applied Physics B*, vol. 33, no. 3, pp. 149–153, 1984. [Online]. Available: http://dx.doi.org/10.1007/BF00688521

[258] L. Onural and M. T. Özgen, "Extraction of three-dimensional object-location information directly from in-line holograms using Wigner analysis," *Journal of Optical Society of America A*, vol. 9, no. 2, pp. 252–260, 1992. [Online]. Available: https://doi.org/10.1364/JOSAA.9.000252

[259] F. Soulez, L. Denis, C. Fournier, Éric Thiébaut, and C. Goepfert, "Inverse-problem approach for particle digital holography: accurate location based on local optimization," *Journal of Optical Society of America A*, vol. 24, no. 4, pp. 1164–1171, Apr 2007. [Online]. Available: https://doi.org/10.1364/JOSAA.24.001164

[260] S.-H. Lee, Y. Roichman, G.-R. Yi, S.-H. Kim, S.-M. Yang, A. van Blaaderen, P. van Oostrum, and D. G. Grier, "Characterizing and tracking single colloidal particles with video holographic microscopy," *Optics Express*, vol. 15, no. 26, pp. 18 275–18 282, Dec 2007. [Online]. Available: https://doi.org/10.1364/OE.15.018275

[261] F. C. Cheong, B. J. Krishnatreya, and D. G. Grier, "Strategies for three-dimensional particle tracking with holographic video microscopy," *Optics Express*, vol. 18, no. 13, pp. 13 563–13 573, Jun 2010. [Online]. Available: https://doi.org/10.1364/OE.18.013563

[262] J. Fung, K. E. Martin, R. W. Perry, D. M. Kaz, R. McGorty, and V. N. Manoharan, "Measuring translational, rotational, and vibrational dynamics in colloids with digital holographic microscopy," *Optics Express*, vol. 19, no. 9, pp. 8051–8065, Apr 2011. [Online]. Available: https://doi.org/10.1364/OE.19.008051

[263] M. Seifi, L. Denis, and C. Fournier, "Fast and accurate 3D object recognition directly from digital holograms," *Journal of Optical Society of America A*, vol. 30, no. 11, pp. 2216–2224, Nov 2013. [Online]. Available: https://doi.org/10.1364/JOSAA.30.002216

[264] A. Yevick, M. Hannel, and D. G. Grier, "Machine-learning approach to holographic particle characterization," *Optics Express*, vol. 22, no. 22, pp. 26 884–26 890, Nov 2014. [Online]. Available: https://doi.org/10.1364/OE.22.026884

[265] J. Gillespie and R. A. King, "The use of self-entropy as a focus measure in digital holography," *Pattern Recognition Letters*, vol. 9, no. 1, pp. 19–25, 1989. [Online]. Available: http://dx.doi.org/10.1016/0167-8655(89)90024-X

[266] P. Ferraro, G. Coppola, S. D. Nicola, A. Finizio, and G. Pierattini, "Digital holographic microscope with automatic focus tracking by detecting sample displacement in real time," *Optics Letters*, vol. 28, no. 14, pp. 1257–1259, Jul 2003. [Online]. Available: https://doi.org/10.1364/OL.28.001257

[267] M. Liebling and M. Unser, "Autofocus for digital Fresnel holograms by use of a Fresnelet-sparsity criterion," *Journal of Optical Society of America A*, vol. 21, no. 12, pp. 2424–2430, Dec 2004. [Online]. Available: https://doi.org/10.1364/JOSAA.21.002424

[268] F. Dubois, C. Schockaert, N. Callens, and C. Yourassowsky, "Focus plane detection criteria in digital holography microscopy by amplitude analysis," *Optics Express*, vol. 14, no. 13, pp. 5895–5908, 2006. [Online]. Available: https://doi.org/10.1364/OE.14.005895

[269] P. Langehanenberg, B. Kemper, D. Dirksen, and G. von Bally, "Autofocusing in digital holographic phase contrast microscopy on pure phase objects for live cell imaging," *Applied Optics*, vol. 47, no. 19, pp. D176–D182, Jul 2008. [Online]. Available: https://doi.org/10.1364/AO.47.00D176

[270] J. Dohet-Eraly, C. Yourassowsky, and F. Dubois, "Fast numerical autofocus of multispectral complex fields in digital holographic microscopy with a criterion based on the phase in the Fourier domain," *Optics Letters*, vol. 41, no. 17, pp. 4071–4074, Sep 2016. [Online]. Available: https://doi.org/10.1364/OL.41.004071

[271] G. He, W. Xiao, and F. Pan, "Automatic focus determination through cosine and modified cosine score in digital holography," *Optical Engineering*, vol. 56, no. 3, p. 034103, 2017. [Online]. Available: http://dx.doi.org/10.1117/1.OE.56.3.034103

[272] Z. Ren, N. Chen, and E. Y. Lam, "Automatic focusing for multisectional objects in digital holography using the structure tensor," *Optics Letters*, vol. 42, no. 9, pp. 1720–1723, May 2017. [Online]. Available: https://doi.org/10.1364/OL.42.001720

[273] Y. Jia, E. Shelhamer, J. Donahue, S. Karayev, J. Long, R. Girshick, S. Guadarrama, and T. Darrell, "Caffe: Convolutional architecture for fast feature embedding," in *Proceedings of the 22nd ACM international conference on Multimedia*. ACM, 2014, pp. 675–678. [Online]. Available: https://doi.org/10.1145/2647868.2654889

[274] S. J. Pan, Q. Yang *et al.*, "A survey on transfer learning," *IEEE Transactions on Knowledge and Data Eengineering*, vol. 22, no. 10, pp. 1345–1359, 2010. [Online]. Available: https://doi.org/10.1109/TKDE.2009.191

[275] T. Pitkäaho, A. Manninen, and T. J. Naughton, "Performance of autofocus capability of deep convolutional neural networks in digital holographic microscopy," in *Digital Holography and Three-Dimensional Imaging*. Optical Society of America, 2017, p. W2A.5. [Online]. Available: http://www.osapublishing.org/abstract.cfm?URI=DH-2017-W2A.5

[276] L. S. Davis, A. Rosenfeld, and J. S. Weszka, "Region extraction by averaging and thresholding," *IEEE Transactions on Systems, Man, and Cybernetics*, no. 3, pp. 383–388, 1975. [Online]. Available: https://doi.org/10.1109/TSMC.1975.5408419

[277] L. S. Davis, "A survey of edge detection techniques," *Computer Graphics and Image Processing*, vol. 4, no. 3, pp. 248–270, 1975. [Online]. Available: https://doi.org/10.1016/0146-664X(75)90012-X

[278] M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: Active contour models," *International Journal of Computer Vision*, vol. 1, no. 4, pp. 321–331, 1988. [Online]. Available: https://doi.org/10.1007/BF00133570

[279] S. N. Sulaiman and N. A. M. Isa, "Adaptive fuzzy-K-means clustering algorithm for image segmentation," *IEEE Transactions on Consumer Electronics*, vol. 56, no. 4, 2010. [Online]. Available: https://doi.org/10.1109/TCE.2010.5681154

[280] L. G. Ugarriza, E. Saber, S. R. Vantaram, V. Amuso, M. Shaw, and R. Bhaskar, "Automatic image segmentation by dynamic region growth and multiresolution merging," *IEEE Transactions on Image Processing*, vol. 18, no. 10, pp. 2275–2288, 2009. [Online]. Available: https://doi.org/10.1109/TIP.2009.2025555

[281] P. Campadelli, D. Medici, and R. Schettini, "Color image segmentation using hopfield networks," *Image and Vision Computing*, vol. 15, no. 3, pp. 161–166, 1997. [Online]. Available: https://doi.org/10.1016/S0262-8856(96)01121-3

[282] S. Ji and H. Park, "Image segmentation of color image based on region coherency," in *Proceedings of the 1998 International Conference on Image Processing, ICIP. Part 2 (of 3)*, vol. 1. IEEE, 1998, pp. 80–83. [Online]. Available: https://doi.org/10.1109/ICIP.1998.723425

[283] C. Huang, "Pattern image segmentation using modified Hopfield model," *Pattern Recognition Letters*, vol. 13, pp. 345–353, 1999. [Online]. Available: https://doi.org/10.1016/0167-8655(92)90032-U

[284] M. Sammouda, R. Sammouda, N. Niki, and K. Mukai, "Segmentation and analysis of liver cancer pathological color images based on artificial neural networks," in *Image Processing, 1999. ICIP 99. Proceedings. 1999 International Conference on*, vol. 3. IEEE, 1999, pp. 392–396. [Online]. Available: https://doi.org/10.1109/ICIP.1999.817142

[285] Y.-S. Lo and S.-C. Pei, "Color image segmentation using local histogram and self-organization of Kohonen feature map," in *Image Processing, 1999. ICIP 99. Proceedings. 1999 International Conference on*, vol. 3. IEEE, 1999, pp. 232–235. [Online]. Available: https://doi.org/10.1109/ICIP.1999.817107

[286] F. Yi, I. Moon, and B. Javidi, "Automated red blood cells extraction from holographic images using fully convolutional neural networks," *Biomedical Optics Express*, vol. 8, no. 10, pp. 4466–4479, 2017. [Online]. Available: https://doi.org/10.1364/BOE.8.004466

[287] P. Memmolo, G. D. Caprio, C. Distante, M. Paturzo, R. Puglisi, D. Balduzzi, A. Galli, G. Coppola, and P. Ferraro, "Identification of bovine sperm head for morphometry analysis in quantitative phase-contrast holographic microscopy," *Optics Express*, vol. 19, no. 23, pp. 23 215–23 226, Nov 2011. [Online]. Available: https://doi.org/10.1364/OE.19.023215

[288] T. G. Team. (2019) GNU image manipulation program. [Online]. Available: https://www.gimp.org/

[289] W. Jeon, W. Jeong, K. Son, and H. Yang, "Speckle noise reduction for digital holographic images using multi-scale convolutional neural networks," *Optics Letters*, vol. 43, no. 17, pp. 4240–4243, 2018. [Online]. Available: https://doi.org/10.1364/OL.43.004240

[290] L. Jin, S. Gao, Z. Li, and J. Tang, "Hand-crafted features or machine learnt features? together they improve RGB-D object recognition," in *Multimedia (ISM), 2014 IEEE International Symposium on*. IEEE, 2014, pp. 311–319. [Online]. Available: https://doi.org/10.1109/ISM.2014.56

[291] W. Li, S. Manivannan, S. Akbar, J. Zhang, E. Trucco, and S. J. McKenna, "Gland segmentation in colon histology images using hand-crafted features and convolutional neural networks," in *Biomedical Imaging (ISBI), 2016 IEEE 13th International Symposium on*. IEEE, 2016, pp. 1405–1408. [Online]. Available: https://doi.org/10.1109/ISBI.2016.7493530

[292] H. Lee, H. Hong, J. Kim, and D. C. Jung, "Deep feature classification of angiomyolipoma without visible fat and renal cell carcinoma in abdominal contrast-enhanced CT images with texture image patches and hand-crafted feature concatenation," *Medical Physics*, 2018. [Online]. Available: https://doi.org/10.1002/mp.12828

[293] D. T. Nguyen, T. D. Pham, N. R. Baek, and K. R. Park, "Combining deep and handcrafted image features for presentation attack detection in face recognition systems using visible-light camera sensors," *Sensors*, vol. 18, no. 3, p. 699, 2018. [Online]. Available: https://doi.org/10.3390/s18030699

[294] S. Hosseini, S. H. Lee, and N. I. Cho, "Feeding hand-crafted features for enhancing the performance of convolutional neural networks," *arXiv preprint arXiv:1801.07848*, 2018.

[295] S. Wu, Y.-C. Chen, X. Li, A.-C. Wu, J.-J. You, and W.-S. Zheng, "An enhanced deep feature representation for person re-identification," in *Applications of Computer Vision (WACV), 2016 IEEE Winter Conference on*. IEEE, 2016, pp. 1–8. [Online]. Available: https://doi.org/10.1109/WACV.2016.7477681

[296] K. Simonyan, A. Vedaldi, and A. Zisserman, "Deep inside convolutional networks: Visualising image classification models and saliency maps," *arXiv preprint arXiv:1312.6034*, 2013.

[297] D. Arthur and S. Vassilvitskii, "k-means++: The advantages of careful seeding," in *Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms*. Society for Industrial and Applied Mathematics, 2007, pp. 1027–1035.

[298] S. Koivusalo, "α1- and αv-integrins in prostate cancer: digital holographic microscopy-based analysis of patient-derived prostate cancer organoids," Master's thesis, University of Oulu, PL 8000, 90014 Oulun yliopisto, Finland, 11 2018.

[299] Z. Richards, T. McCray, J. Marsili, M. L. Zenner, J. T. Manlucu, J. Garcia, A. Kajdacsy-Balla, M. Murray, C. Voisine, A. B. Murphy *et al.*, "Prostate stroma increases the viability and maintains the branching phenotype of human prostate organoids," *iScience*, vol. 12, pp. 304–317, 2019. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2589004219300306

[300] N. C. for Biotechnolopgy Information, "PubMed," https://www.ncbi.nlm.nih.gov/pubmed/.

[301] Qt Group. (2019) Qt software development framework. [Online]. Available: https://www.qt.io

[302] The HDF Group. (2000-2010) Hierarchical data format version 5. [Online]. Available: http://www.hdfgroup.org/HDF5

A

Imaging experiments

Table A.1 Conducted imaging experiments. Cell line: MDCK, Madine Darby canine kidney; RWPE, human prostate epithelial; MFC10a, mammary gland breast epithelial. Type describes possible manipulation aimed at a specific gene: wild, no treatment done; KRasV12, oncogenic G12V-mutant of K-Ras proto-oncogene; KO after the type defines different gene knockouts. Amount describes how many holograms were captured during the imaging experiment.

ID	Captured on	Cell line	Type	Amount
0	10/2013	MDCK	wild	25
1	10/2013	MDCK	$KRas^{V12}$	26
2	4/2015	MDCK	wild	101
3	4/2015	MDCK	$KRas^{V12}$	101
4	5/2015	MDCK	wild	205
5	5/2015	MDCK	$KRas^{V12}$	156
6	12/2015	RWPE	wild	23
7	12/2015	RWPE	wild inhibitor	19
8	12/2015	RWPE	$KRas^{V12}$	11
9	12/2015	RWPE	$KRas^{V12}$ inhibitor	7
10	12/2015	RWPE	wild	43
11	12/2015	RWPE	$KRas^{V12}$	6
12	1/2016	RWPE	wild	15
13	1/2016	RWPE	$KRas^{V12}$	16
14	4/2016	MFC10A	wild	43
15	4/2016	MFC10A	genetically transformed	42
16	8/2016	RWPE	wild (20X)	50
17	8/2016	RWPE	$KRas^{V12}$ (20X)	63
18	8/2016	MFC10A	wild and genetically transformed (10X)	108
19	8/2016	MFC10A	wild and genetically transformed (20X)	32
20	8/2016	MFC10A	wild and genetically transformed	263
21	8/2017	MDCK	wild	77615
22	8/2017	MDCK	$KRas^{V12}$	129663
23	8/2017	MDCK	wild	147110
24	8/2017	MDCK	wild	18377
25	9/2017	RWPE	wild	150
26	9/2017	RWPE	α-1-KO	150
27	11/2017	RWPE	wild (20X)	1682
28	11/2017	RWPE	wild	1654
29	11/2017	RWPE	P10-KO (20X)	1681
30	1/2018	MDCK	$KRas^{V12}$	149256
31	1/2018	MDCK	wild	14633
32	1/2018	MDCK	aV	40032
33	1/2018	MDCK	$KRas^{V12}$	50

B

Tools created

B.1 Control, processing and analysis software "Phaser"

A software application, called Phaser, was designed and implemented during the project. This software is an interface between the microscope and the user. It can be used to design imaging experiments, process and analyse holograms, and visualise results. The graphical user interface (GUI) was implemented in Qt [301] and using the PySide bindings for Python. Below is a collection of screen-shots of different widgets that are part of the software (Figs B.1 through B.4).

B.2 HDF file processing software "HDF Viewer"

The processed hologram information was saved in Hierarchical Data Format (HDF) that was developed originally by National Center for Supercomputing Applications for storing scientific data [302]. This flexible data format allows one to save data in multiple different formats. HDF does not have the same restrictions as common image formats that require frames to have the same dimension and bit depth for each pixel. Besides this, each dataset (images in our case) can contain meta data as attributes. We developed the application "HDF Viewer" for faster processing of holographic data. By using this tool, it was possible to quickly visualise data e.g. to make sure that segmentation was done properly. Figure. B.5 shows the GUI of the software.

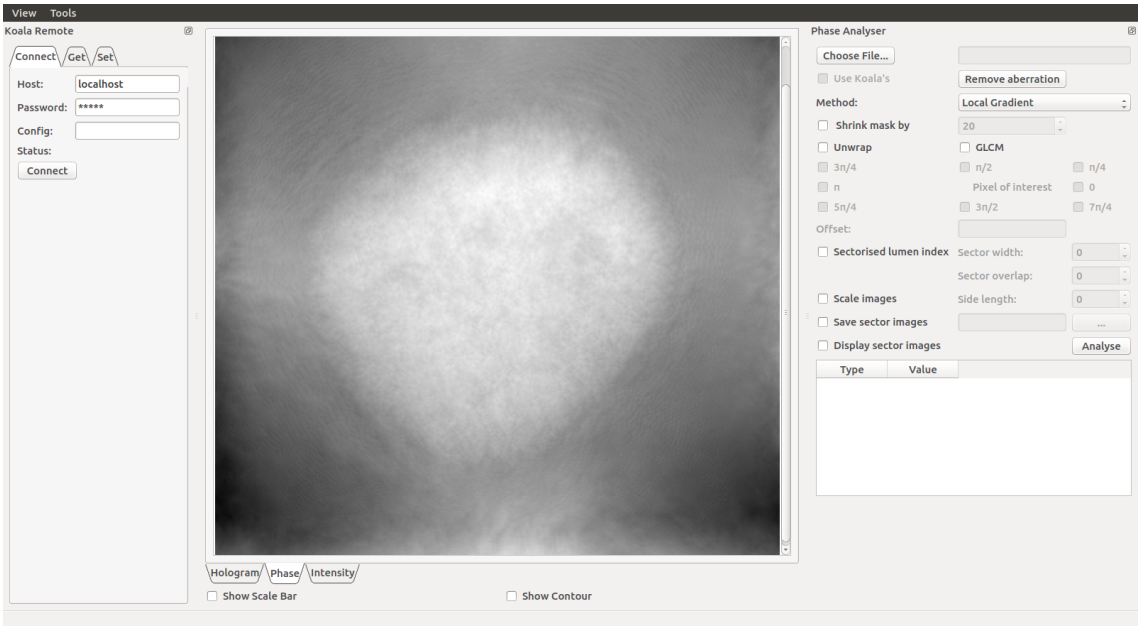

Figure B.1 Main window of the GUI. The centre portion can display holograms, reconstructions, or other visualisations. On the left-hand side is a widget (Koala Remote) that is used to connect Phaser to the software that controls the microscope. On the right-hand side is a widget (“Phase Analyser”) that is used to analyse holograms, extract information from holograms, and display extracted information and results of analysis.

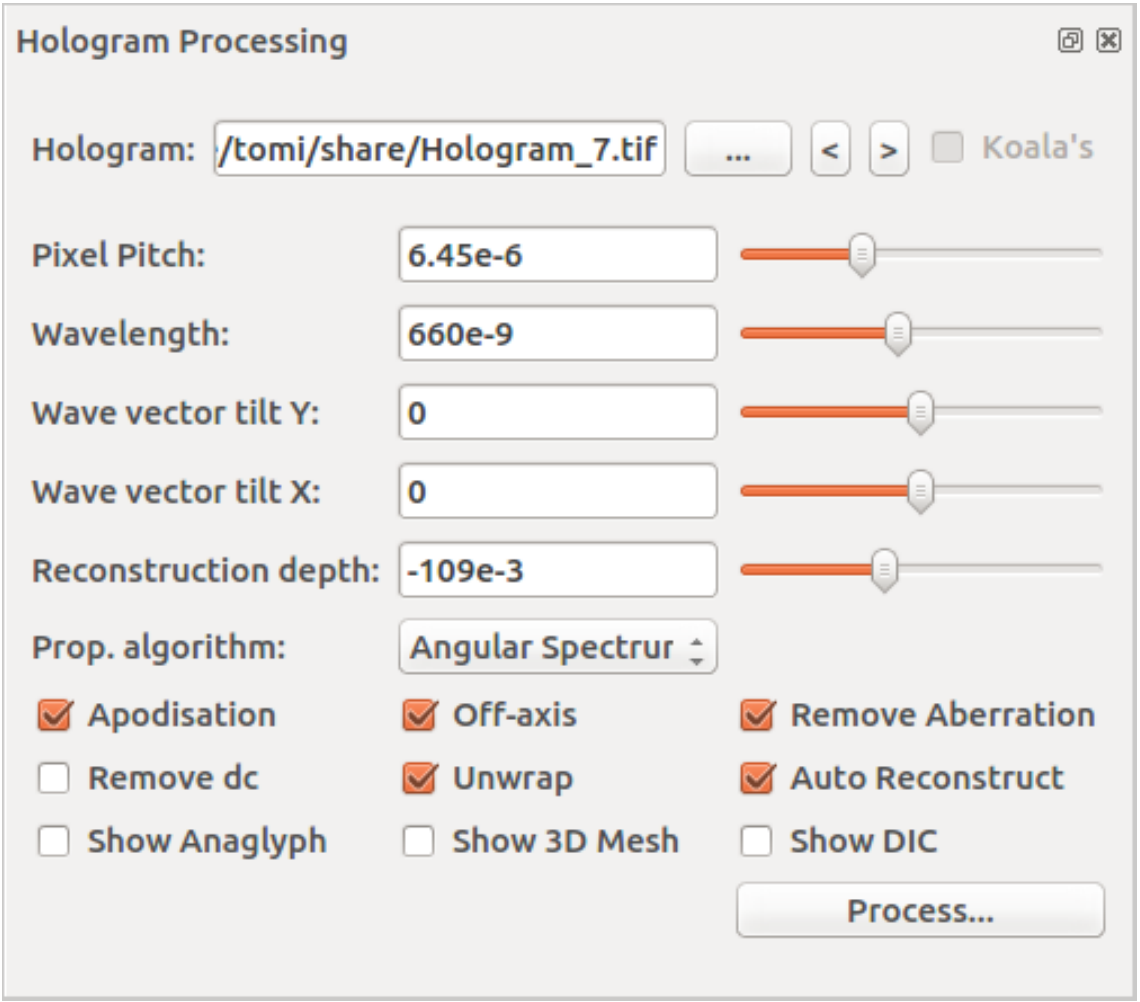

Figure B.2 Hologram processing widget. The widget that is used to process holograms: propagation, aperture apodisation, off-axis processing, aberration removal and dc removal, and a choice of visualisations.

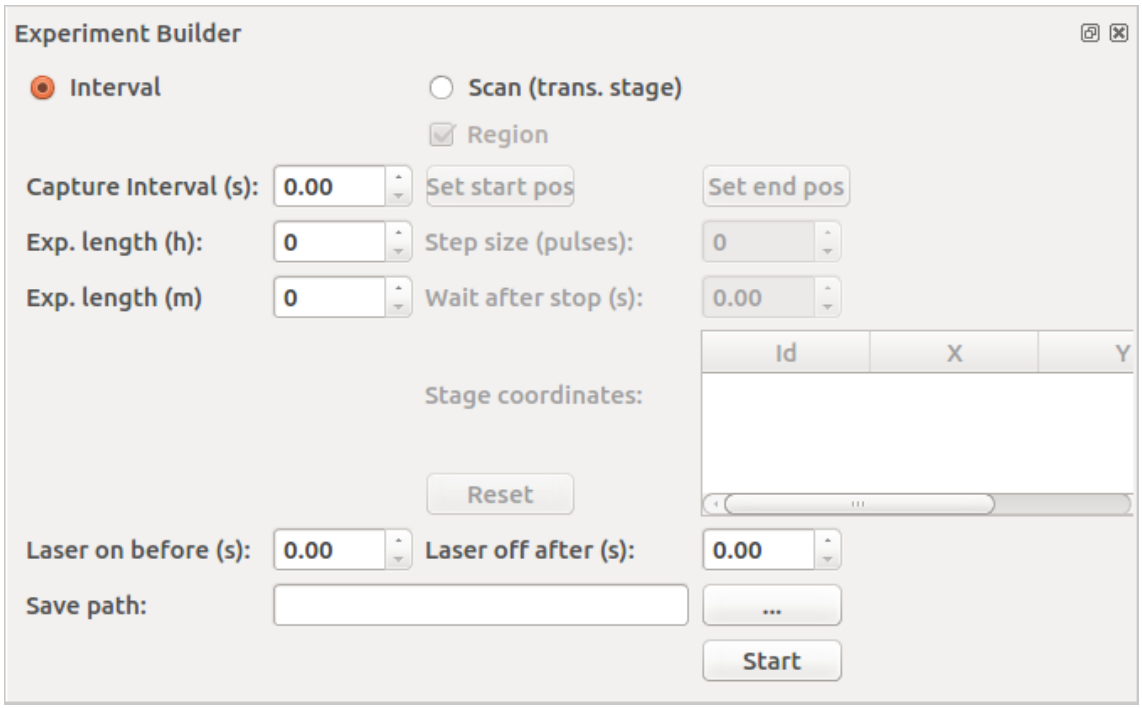

Figure B.3 Experiment builder widget. The widget that is used to program imaging experiments: point or region based scanning.

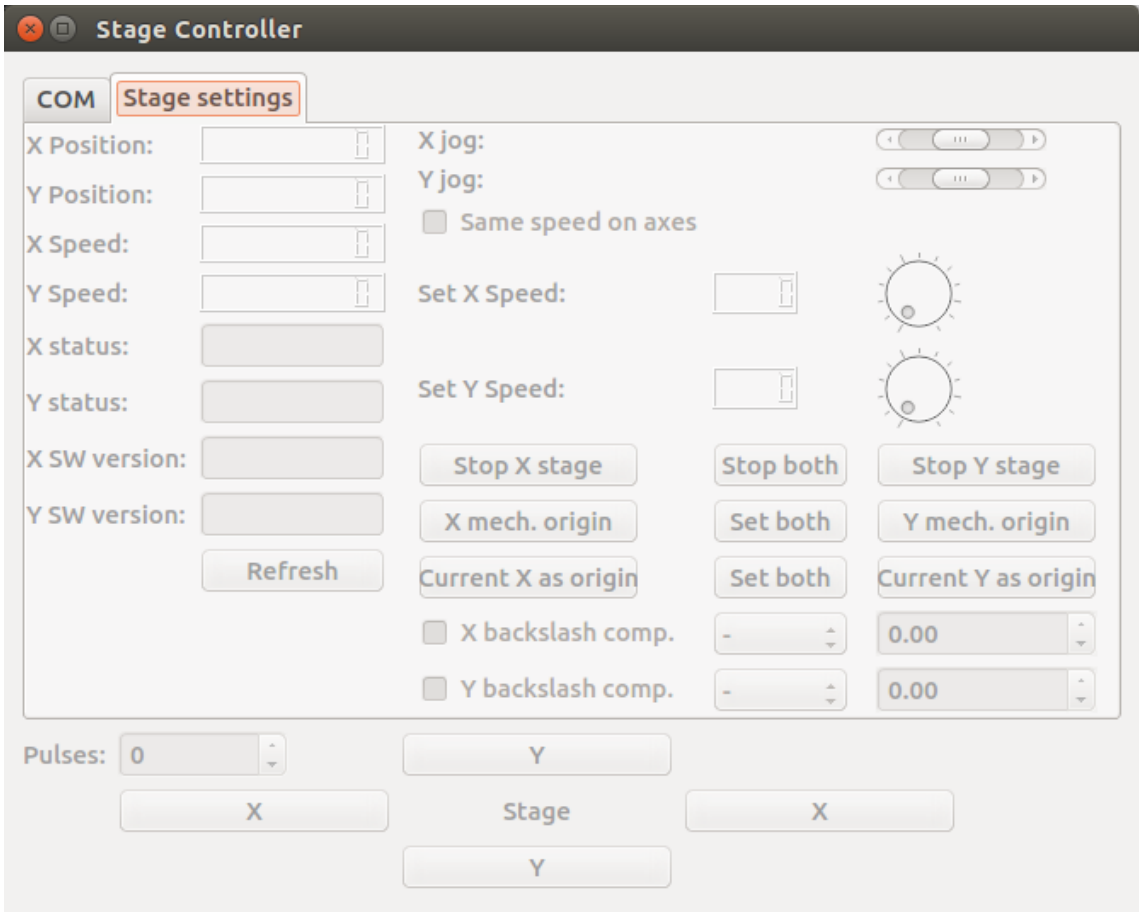

Figure B.4 Translation stage widget. The widget that is used to control the motorised translation stages.

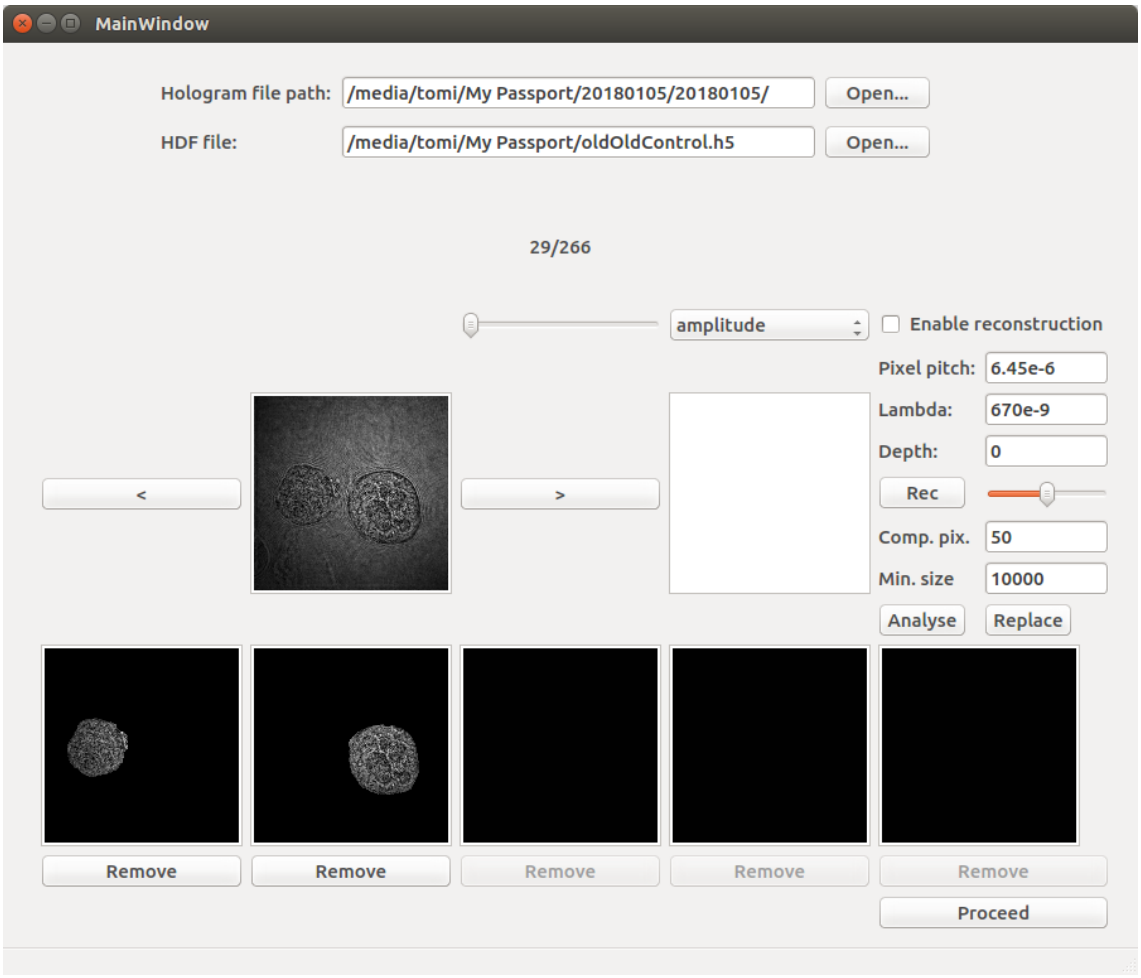

Figure B.5 HDF Viewer GUI. Showing reconstruction parameters, amplitude reconstruction and two segmented cysts.