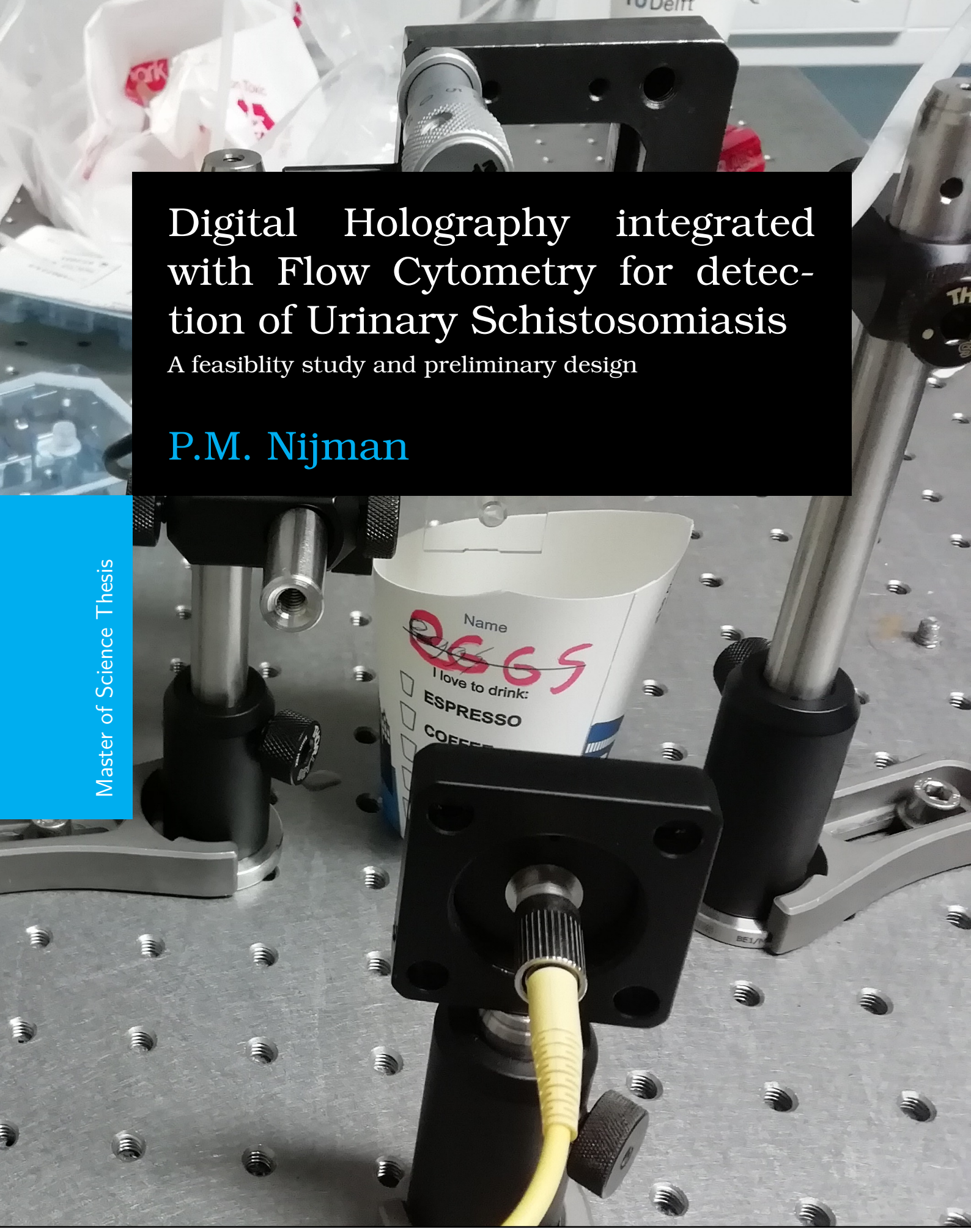


Digital Holography integrated with Flow Cytometry for detection of Urinary Schistosomiasis

A feasiblity study and preliminary design

P.M. Nijman

Master of Science Thesis

TU Delft Delft University of Technology

Delft Center for Systems and Control

Digital Holography integrated with Flow Cytometry for detection of Urinary Schistosomiasis

A feasiblity study and preliminary design

Master of Science Thesis

For the degree of Master of Science in Systems and Control at Delft University of Technology

P.M. Nijman

November 27, 2019

Faculty of Mechanical, Maritime and Materials Engineering (3mE) · Delft University of Technology

Delft University of Technology
Department of
Delft Center for Systems and Control (dcsc)

The undersigned hereby certify that they have read and recommend to the Faculty of Mechanical, Maritime and Materials Engineering (3mE) for acceptance a thesis entitled

Digital Holography integrated with Flow Cytometry for detection of Urinary Schistosomiasis

by

P.M. Nijman

in partial fulfillment of the requirements for the degree of

Master of Science Systems and Control

Dated: November 27, 2019

Supervisor(s):

Prof.dr. G.V. Vdovin

Ir. T.E. Agbana

Reader(s):

Dr.ir. J.C. Diehl

Abstract

Schistosomiasis is an intravascular infection with major public health consequences in developing countries. It is one of the major Neglected Tropical Diseases with more than 240 million people infected and 800 million people at risk in 2015, mostly in sub-Saharan Africa. It is caused by trematode parasites of the genus Schistosoma, in this report the focus was on Schistosomiasis Haematobium since it is the most prevalent form of the disease.

One of the limiting factors of the control program is the standard diagnostic procedure set by World Health Organization, which is based on counting the parasite's eggs in a person's urine. Examination by microscopy requires the use of expensive microscopes, is prone to human errors and inconsistency, is time consuming, and uses filters which are often not available. The research objective was identified from these shortcomings:

> *"Develop a low-cost, smart diagnostic method for Schistosomiasis Haematobium based on detecting eggs in urine by combining lensless imaging and flow cytometry, and developing Artificial Intelligence models for automated detection."*

In-line planar wavefront digital holography was identified as the most suitable lensless imaging method. A sample will be analyzed by the following repetitive procedure: (1)mechanically press the piston of a syringe by a small volume, (2)wait for the flow to stop, (3)record a hologram, (4)detect eggs. The implemented egg detection procedure consisted of a series of image processing algorithms: (1)apply Foreground detection, (2)localize the moving objects with a Blob detector, (3)locally reconstruct the hologram at the found locations, (4)classify the reconstruction as egg or not egg. The imaging method provided accurate reconstructions of eggs and the object detection algorithm was able to locate moving objects with sufficient accuracy and computational time. On the other hand, the lab and field tests showed that the data set of the classifier did not contain enough images to train a generalized model and that the local reconstruction and classification takes increasingly more time during analysis. As of now the method is an order of magnitude slower than an expert microscopist.

The diagnostic method is not yet able fulfill the research objective. However, there are some promising aspects such as the low-cost imaging method, fast object detection algorithm, and absence of sample preparation which makes further research worthwhile.

Table of Contents

List of Figures

List of Tables

Preface

The dissertation "Digital holography integrated with flow cytometry for detection of urinary schistosomiasis", is written to fulfill the graduation requirements of the Systems & Control Master program at the Technical University of Delft. It is a feasibility study and preliminary design of a new diagnostic method.

The research was performed under supervision of Temitope Agbana and professor doctor Gleb Vdovin, and in collaboration with the Smart Optical Diagnosis Of urinary Schistosomiasis (SODOS) group. I chose the topic in early 2017 because of its social relevance and the potential impact it could have on developing countries with limited resources. Before this I was not aware of the existence of Schistosomiasis, and even now I find it surprising how problems like these are not discussed more openly in western society.

The scope of the research was broad, covering many interesting topics such as tropical diseases, lensless imaging, and computer Vision. It suffices to say that it has been eye-opening in many ways.

I am happy to have been a part of this field of research as it has provided me ample of experience in a variety of subjects and put me in contact with some great people. The road has not been one without scientific and personal hurdles, but I would not want to have had it any other way.

Enjoy your reading!

Acknowledgements

First and foremost I would like to thank my supervisors for helping me throughout this research. The discussions with Temitope has been very helpful in the experimental development and data analysis of this research. I am also grateful to Prof. Gleb for providing me with many new insights in the course of this research. I thank you both for this opportunity and I wish you the best in the future.

I have had the pleasure of collaborating with two industrial design students. Although I did not spend a lot of time with Mirte Vendel, meetings with her were always pleasant. I spent a lot of time with Max Hoeboer in the later stages of my research, during which I got a look into to the embodiment and interaction design of a diagnostic device. Max' work ethic has been very inspiring, and I really enjoyed collaborating with him.

Ben, Rens, Chris, Roger, Folkert, Daniel, and other fellow master students: I really enjoyed the cups of coffee and beers we shared.

There were periods during my Thesis where I was subjected to some of the hardest personal problems I have faced in life. Although many people supported me in this process and deserve recognition, the people mentioned below stand out especially.

Sébastien and Victor helped me to put things into perspective and make the best out of this situation.

My brothers brought joy into my life when I was down on luck and had to move back in with my parents.

And most of all I would like to express my gratitude towards my dear parents. Without their unconditional love and support I would have never been able to finish this project. Thanks for being there when I needed you the most.

Delft, University of Technology — P.M. Nijman
November 27, 2019

Chapter 1

Introduction

Schistosomiasis Haematobium (SH) affects more than 100 million people worldwide. Poor hygiene conditions, expensive and inaccessible diagnostic procedures are the major reasons for the uncontrollable spread of this Neglected Tropical Disease (NTD). To provide accurate and effective drug prescription to infected patients, the infection load must be well defined. This thesis reports on the development of a new optical diagnostic method that will potentially provide label free quantitative diagnosis of SH. This chapter provides a quick overview of SH and also outlines the goal and objective of this report. First the biological characteristics of the disease are presented along with the current state of the art of the diagnostic methods. Next, the development of new quantitative diagnostic methods for SH is researched. The chapter concludes by presenting the research objective and by proving a brief layout of what will be discussed in each chapter of this report.

Figure 1-1: A male and female pair of Schistosoma.[1]

1-1 Schistosomiasis

Schistosomiasis, also called bilharzia, is a major intravascular infection with major public health consequences in developing countries. It is one of the major NTDs with more than 240 million people infected and 800 million people at risk in 2015. Disease burden assessment for schistosomiasis has shown that the annual number of disability adjusted life years (DALYs) list is around 70 million, with around 300.000 deaths annually in Africa.[15]

Schistosomiasis is caused by trematode parasites of the genus Schistosoma, of which 3 major species cause the disease in humans: Schistosoma mansoni, Schistosoma japonicum, and Schistosoma haematobium. People contract the disease by contact with water, hence it prevails in areas of poverty where people do not have direct access to sanitary facilities. Local inhabitants wash themselves in rivers and lakes as shown in Figure 1-2, which results in the high prevalence mentioned earlier.[2] In this report the focus is on SH since it is the most prevalent of the three.[2] Schistosoma have a complex life-cycle which is schematically depicted in Figure 1-3, a distinct part of the life cycle of Schistosoma haematobium is that the eggs are laid in the bladder and exit the body along with urine.

Figure 1-2: Kids playing and bathing in water[2]. Waters like these are often filled with cercariae which will hatch into the schistosoma blood flukes in case they manage to infect a person.

The World Health Organisation (WHO) has been putting effort into controlling Schistosomiasis with the end goal of eventually eradicating the disease. The current strategy for control aims to prevent morbidity in later life through regular treatment with praziquantel, which is currently the only recommended drug for treatment of Schistosomiasis. Praziquantel kills the parasitic worms but does not prevent reinfection, therefore continuous treatment is needed in endemic areas. Infection rates and intensity of infection have reduced since the introduction of control programs in endemic countries[3]. However, detecting low intensity infections has been difficult especially for people that have already been treated. This makes it hard to identify the prevalence of SH in endemic regions.

1-2 Diagnostic standard for SH

There are several ways to determine whether a person is infected with SH, most of which rely on the presence of some substance in urine. The current standard for in-field Point of Care (POC) diagnostics set by the WHO relies on counting eggs in a persons urine. There are diagnostic methods reliant on other factors in urine such as anti body DNA or worm DNA however these methods are not accepted by the WHO as a standard. Therefore this section will only elaborate on the standard diagnostic method set by the WHO.[2]

As shown in Section 1-1, a person infected with SH has the parasite's eggs in their urine. The eggs from Schistosoma haematobium are approximately $135[\mu m] \times 50[\mu m]$ in size and

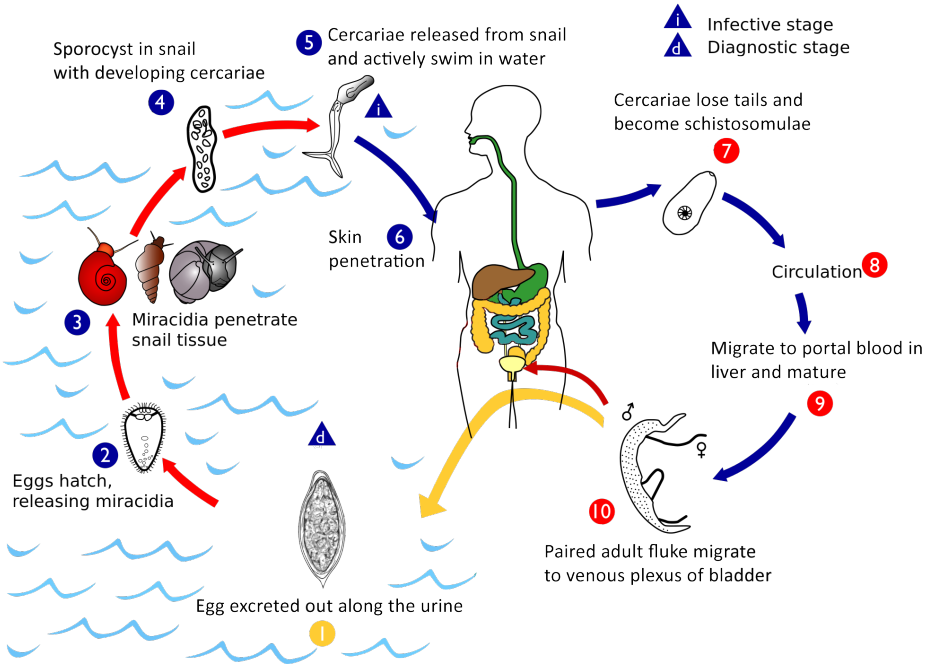


Figure 1-3: Visualisation of the lifecycle of SH[3].

are unique in shape, making them a key indicator for presence of the disease[16]. The WHO dictates that the infection intensity is quantified as the amount of eggs present in a person's urine[3]. The infection intensity is determined according the amount of eggs present in 10 mL of urine. The absence of eggs is noted as uninfected, if there are between 0 and 50 eggs it is diagnosed as a light infection, any egg count beyond 50 is diagnosed as a heavy infection. Since the amount of eggs in urine can vary throughout the day often multiple samples are examined for a more accurate diagnosis. Due to its simplicity, syringe filtration is preferred in community and school surveys as it requires less equipment than urine concentration with centrifugation[15].

A study done in Ghana by A. Koukounari et al concluded that the best diagnostic test for the detection of the prevalence of SH out of 5 tested methods was examination by microscopy. The study tested diagnosis with the following methods: Urine-antigen detection, Serology anti-IgG, Ultrasound examination, and parasitological examination through microscopy and haematuria. With microscopy they were able to reach a sensitivity of 92.5% and a specificity of 97.9% from a single urine sample passed between 10:00 and 14:00, which is the time of optimum egg passage. [17][18]

To give an impression of what the eggs look like, Figure 1-4 shows a Schistosoma haematobium egg stained with iodine. Egg counting is performed in the following manner:

1. A midday urine sample is collected.
2. 10 mL of urine is drawn from the sample.
3. A nylon, paper, or polycarbonate filter($12 - 20[\mu m]$ pore size) is used to filter the eggs from the sample [3].

4. The eggs are stained by using Lugol's iodine[2].
5. The filter is put under a microscope.
6. A trained professional counts the number of eggs.

Figure 1-4: Schistosoma hamatobium egg stained with iodine [4]

The current diagnostic standard comes with the following drawbacks which makes it unsuitable for diagnosing large populations:

- Filters are not always available in field settings and often have to be imported from Western countries.
- Expensive equipment (microscope) is necessary.
- Human expertise needed for reliable diagnoses which is not commonly available in low resource settings with high disease prevalence.
- The method is prone to Human error due to a combination of large volume of samples and fatigue, and lack of concentration.

These drawbacks emphasize the need for a low cost, smart diagnostic device for infield use.

1-3 Previous work on diagnostics for SH

In 2014, Bogoch et al. experimented with a mobile phone microscope for the diagnosis of Schistosomiasis mansoni, SH, and soil-transmitted helminths. They concluded that the mobile phone microscope had only modest sensitivity with their experimental set-up. They also mention that the development of portable diagnostic technologies that can be used at point-of-sample collection will enhance diagnostic coverage in clinical and epidemiological settings[19].

In 2015, Ephraim et al. experimented with a mobile phone-mounted Foldscope and reversed-lens CellScope combined with paper towel filtration for SH diagnosis. These experiments were performed in order to support the movements to affordable diagnostic methods in low resource settings. Locally produced paper towels were used as filters in order to save on conventional filters which are generally not available in those regions. They reported that the mobile phone-mounted Foldscope had limited sensitivity but excellent specificity for the diagnosis of SH. The reversed-lens CellScope showed similar results. Based on the report, the need for robust, simple, and inexpensive diagnostic tests in resource-constrained settings is cogent. Their tests with toilet paper filtration showed that it was not possible to detect SH by filtering the urine with single-ply paper towels. The paper is not smooth enough which causes the light to refract, making it impossible to distinguish the eggs[20].

There have been several studies on automating the process of counting the eggs in recorded images of samples. In 2017, Holmström et al researched automated detection using a deep learning-based algorithm in combination with a digital microscopy scanner. The eggs were visually identified for training of the neural network. Usage of the neural network showed a sensitivity ranging from $83.3\% - 100\%$ in the test set of manually labeled helminth eggs, however the images only contained up to 12 eggs. The study showed the imaging performance of a mobile, digital microscope was sufficient for visual detection of soil-transmitted helminths and Schistosoma haematobium. Furthermore, it also showed that deep learning-based image analysis can be used for the automated detection of SH[21]. The analyzed images were obtained from 4 $20\mu L$ samples which were examined under a microscope, so the method is not suited to replace the current diagnostic standard since the sample preparation is too complex. In 2016, Q. Li explored an automatic detection method for Schistosomiasis Japonicum miracidia, consisting of a real-time video detection algorithm. The method showed better accuracy and efficiency than human eye recognition for Schistosomiasis japonicum miracidia, but is not aplicable for diagnosing Schistosomiasis Japonicum[22]. Linder et al have investigated low cost on-chip imaging to see if using this is a feasible method in order to diagnose Schistosomiasis. This research showed that Schistosomiasis can be detected using an inexpensive imaging device such as a webcam. It also showed that computer vision algorithms can be used for detection, removing the human intervention. The images of the eggs are not up to the standard as can be seen in Figure 1-5, the absence of reconstruction limits the image quality and the sensors have to be replaced after every use because it is based on contact imaging (more on this in Section 2-1). However, it is a study which shows that imaging sensors can be used to image eggs directly without lens[5].

In 2012, Gao et al. applied an automatically microscopic scanning system in the etiological diagnosis of Schistosomiasis in order to partially automate the counting of eggs. The movement of the slide under the microscope was automated, but the egg counting was still performed manually. In the control group, professionals analyzed 20 intestinal Schistosomiasis samples by manually moving the slide. The sensitivity, specificity, and overall accuracy of the diagnosis of Schistosomiasis of the automatically microscopic scanning system was higher than that of the traditional manual microscopy detection. However, human expertise and sample preparation were still required for this method of egg counting.[23]

In short, there have been several efforts to make the quantitative diagnosis based on egg counting more affordable and/or less labor intensive. However, so far the efforts have not resulted in a method which could improve on conventional microscopic egg counting done by a trained individual to an extend to where it could be replaced.

Figure 1-5: Image of an SH egg which was put on top of the imaging chip[5].

1-4 Research objective

This section will emphasize the research gap and present the reader with the research objective of this report. First the diagnostic technology gap is elaborated on and next some methods are introduced which could be used to fill the technology gap.

1-4-1 Diagnostic technology gap

Current parasitological tests are subject to shortcomings which leave room for improvement. The need for a novel diagnostic method is clearly stated by Kosala G. A. D. Weerakoon[15]:

> *"It is a sobering thought that despite extensive efforts, as indicated earlier, the global disease burden of schistosomiasis still remains unacceptably high. This persistence of the disease despite massive and integrated control programs over the last few decades may be due in part to the lack of accurate diagnostic tools for case detection and community screening in areas where schistosomiasis is endemic.*
> *...*
> *It is imperative to develop more effective approaches for the prevention, control, and elimination of schistosomiasis. Morbidity reduction and parasite elimination are the two main pillars of current control programs. Effective diagnosis plays a key role in control strategies, with wide applications in case detection in areas with a high prevalence as well as those with a low prevalence, where the main aims are elimination of infection, evaluation of disease intensity, and assessment of therapeutic responses as well as the overall effectiveness of the interventions employed."*

Two methods that raise interest for diagnosing SH through detection of eggs are lensless imaging and flow cytometry. As mentioned in Chapter 2, lensless imaging can be a low cost

and robust alternative to lens based microscopy, while flow cytometry will not require any of the sample preparation methods mentioned in Section 1-2 since the sample can administered directly from the syringe. The combination of these two techniques can make diagnosing SH low cost and user friendly, therefore this application should be investigated. Combining these methods for quantitative diagnosis will reduce the materials needed for field testing since sample preparation materials will be reduced.

Quantitative field testing is a labor intensive procedure which requires a person's concentration throughout the day in order to accurately count eggs. There have been examples of artificial intelligence based counting of red blood cells, which hints at its application to egg counting[24][25][26]. Automating the counting procedure will give the following advantages:

- No need for technical expertise for in-field testing.
- Save on training cost and time.
- Time savings, ability to do other tasks.
- Removal of human error.
- Improved data collection.

1-4-2 Research objective

This chapter has discussed the shortcomings of diagnosing SH, what should be done to address these, and present lensless imaging and flow cytometry as a possible solution. Combining this leads to the following research objective:

> *"Develop a low-cost, smart diagnostic method for SH based on detecting eggs in urine by combining lensless imaging and flow cytometry, and developing Artificial Intelligence (AI) models for automated detection."*

This objective naturally splits up into several minor objectives since digital holography and flow cytometry have not been applied in this field before, and it concerns the development of a new method for this application. The objective can be partitioned as shown below:

- Verify whether a combination of lensless imaging and Flow cytometry can be used to image SH eggs in flowing urine samples.
- Develop a model for automated detection and estimation of infection load.
- Validate the method according to medical standards.
- Implement the developed method on a mobile device suitable for in field testing.

These steps give an idea on how the research will be performed during this research.

1-4-3 Future effects

The value of research does not depend purely on the results, but rather on the implications those results have on future research. Successfully completing the research objective will add the following knowledge to the field of research:

- Digital holography can be combined with flow cytometry for analyzing the presence of SH eggs in urine.
- Artificial intelligence can alleviate the responsibility of manually counting and quantifying the SH eggs in a given infected sample.
- Field testing for SH will be cheaper and less time intensive as the new method will be simple to use, cheap to maintain, and not need any sample pre-processing.
- The result serves as a catalyst for other research related to NTD diagnosis automation.

1-5 Report layout

How is the research objective, mentioned in the previous section, handled in this report? This section provides insight about the structure of this report.

First, Chapter 2 will provide background information on the methods used to design the new diagnostic method. Concepts such as lensless imaging, flow cytometry, and machine vision and their application to this topic will be discussed. Then Chapter 3 will present the experimental results of the validation tests of the aforementioned concepts, including a resolution test of the holographic microscope and benchmark results of the detection algorithm. Next, Chapter 4 presents a case study in which a prototype based on the detection algorithm will be used for field tests in Ivory Coast. The results of these field tests will be analyzed in this chapter. Finally the report concludes with Chapter 5 which contains a discussion on the results of the research and recommendations for its continuation.

Chapter 2

Diagnostic Method Proposal

The previous chapter presented the research objective and along with it concepts such as lensless imaging, flow cytometry, and Artificial Intelligence (AI) were introduced. This chapter presents the background information on these concepts and shows how they can be applied to fulfill the research objective. In other words: A new diagnostic method is proposed. First, Section 2-1 presents a short review of lensless imaging methods and selects the most appropriate one for the purpose of this research. Next, Section 2-2 discusses flow cytometry and how it can be applied in order to record urine flow. The egg detection algorithm is then presented in Section 2-3, followed by a discussion on how a diagnosis can be performed in Section 2-4. Finally the Chapter is concluded in Section 2-5.

Figure 2-1: Sample preparation and examination by microscope, the research objective is to eliminate this procedure[6].

2-1 Lensless Imaging

Lensless imaging is an emerging field of research for high resolution imaging in low-resource settings. This Section will briefly discuss the various lens-less imaging methods and decide which method is most applicable in this case.

2-1-1 Digital Holographic Microscopy

Digital holography is an imaging technique where the magnitude and phase information of an object wave is captured as an interference pattern. In this report only inline holography is considered since it is the least complex and does not require any lenses, a schematic of this imagin method is shown in Figure 2-2.

For an in-line holographic setup only one source of illumination is used, the portion of the illumination unperturbed by the object is considered to be the reference wave $R(x, y, z)$. A small portion of the illumination hits the transparent object, causing some of the illumination light to scatter, which constitutes the object wave $O(x, y, z)$. The measured interference between the un-scattered reference wave and the scattered object wave is called the hologram, which is shown in Equation (2-1). The interference appears as a fringe pattern perpendicular to the direction of propagation. The most important terms of this equation are the last two, which represent the dominant interference terms. The hologram is reconstructed by illuminating the recorded hologram with the reference wave.[27]

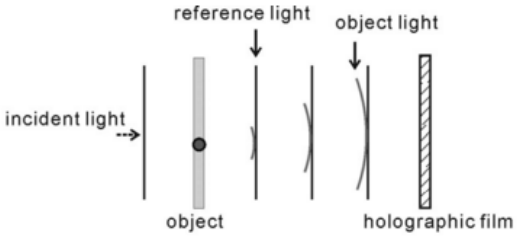


Figure 2-2: Depiction of the setup for obtaining a Gabor hologram[7]

$$|R + O|^2 = |R^2| + |O^2| + R * O + O * R \tag{2-1}$$

In digital holography the hologram is recorded on an electronic image sensor-array, while the reconstruction step is done numerically. This yields a digital image which ideally contains both phase and magnitude information of the object. Since the reference wave is necessary for reconstruction this has to be simulated properly when doing numerical reconstruction. Measuring the phase distribution of an arbitrary wave with the required precision for high resolution imaging is problematic. Therefore an illumination source with known wavefront should be used, this is usually a spherical or planar wavefront. Both these types are being used for holographic microscopy, but their merits and reconstruction methods differ. The wave propagation algorithms which are used to reconstruct digital holograms are based on the scalar diffraction theory and are the numerical solutions of either the Fresnel-Kirchhoff, or the Rayleigh-Sommerfeld diffraction integrals. The latter out of the two being considered the most accurate because it does not rely on any approximations unlike the Fresnel-Kirchhoff diffraction integral which uses the Fresnel approximation. [7]

The Rayleigh-Sommerfeld diffraction integrals results in a reconstruction algorithm which uses the angular spectrum method shown in equation (2-2). In this equation $h(x, y, z)$ represents the propagation kernel which depends on the reference wavefront, $E_1(x, y, 0)$ the scalar electric field in the hologram plane, and $E_2(x, y, z)$ the scalar electric field in the object plane.

$$E_2(x, y, z) = \mathcal{F}^{-1}\{\mathcal{F}\{h(x, y, z)\} \cdot \mathcal{F}\{E_1(x, y, 0)\}\} \tag{2-2}$$

Equation (2-3) shows the distance condition for no aliasing when using Angular Spectrum Method (ASM) while Equation (2-4) shows the distance condition for no aliasing when using Fresnel-Kirchhoff diffraction integrals[7]. In these equations z represents the distance between object and sensor, λ the illumination wavelenght, Δ_x the pixel pitch, n the image width in pixels, and n' the width of the area containing nonzero pixel values. It is evident that the ASM is better suited for short distances, as the propagation distance has to be smaller than some value.[7]

$$z \leq \frac{\sqrt{4\Delta_x^2 - \lambda^2}}{2\lambda}(n - n')\Delta_x \tag{2-3}$$

$$z \geq \frac{2n\Delta_x^2}{\lambda} \tag{2-4}$$

The fringes contain information on the transparency of the object as well as the thickness, as these properties affect the intensity of light as well as the phase. With this information it becomes possible to image transparent objects without labeling them. Over the years this has caused the rise of Digital Holographic Microscope (DHM), also called 3D microscopy or quantitative phase microscopy, which can be used to image microscopic biological specimen which are often transparent. [28] During the reconstruction of in-line holograms the object wave and the conjugate object wave represent optical fields traveling in opposite directions. As a result, upon propagation of the recorded hologram toward the object plane, one of these terms converges to an image of the object, while the other further diverges, forming a weaker defocused twin-image related artifact that is concentric with the actual image. This spatial artifact is caused by the fact that the image sensor only measures the intensity of the light, and thus, the phase of the scalar electric field at the hologram plane is lost. There are methods to remove the twin-image artifact, however; these require extensive computational time, control over the object, or a more complex setup.[29] An important drawback of the discussed holographic system is the need for transparent samples since these microscopes work in transmission mode.

There are two ways, based on the illumination wave, in which digital in-line holography can be applied in order to capture images of microscopic particles. These will be discussed in the following two paragraphs.

Digital in-line holographic microscopy with spherical wavefront illumination

In this method, a pinhole with a size smaller than $1.22 * \lambda$ is illuminated in order to create a coherent spherical wavefront. The spherical wavefront is used to illuminate the target object, where the distance between aperture and object(z_1) is smaller than the distance between object and sensor(z_2). The spherical wave in combination with the distance between object and sensor leads to a fringe magnification of $M = 1 + z_2/z_1$. Numerical reconstruction

has shown to produce images of specimen with sub micron resolution, fringe magnification overcomes the sampling limit introduced by the pixel size of the sensor.[27] The object wave has a high diffraction angle when it hits small objects/details, the spatial resolution is then limited by the sensor dimensions due to relatively large distance between the object and sensor planes. This also limits the Field of View (FOV) since the scattered light of objects located outside of the center is not captured by the sensor. Consequently the whole sensor surface needs to be coherently illuminated, requiring a large spatial coherence diameter as well as the use of a narrow band light source in order to get sufficient temporal coherence. On the other hand, an increase in spatial and temporal coherence can cause unwanted interference terms due to the reflective surfaces in the system and is a source of multiple reflection and speckle noise. LEDs can also be used for illumination, as demonstrated by Repetto [27], but their coherence length is in the micrometer range so the spectral characteristics of the used LED for an optical set-up have to be considered. One practical limitation of these systems is that pinholes of such small size tend to get blocked by dust or other particles easily, and the proper adjustment of such pinhole to a focused laser beam requires high precision and could also require relatively expensive optical and mechanical elements.[27]

The strict requirements of the illumination light combined with the limited FOV due to fringe magnification limit applicability for the purpose of this research.

Digital in-line holographic microscopy with planar wavefront illumination

In this method the illumination light consists of a planar wavefront. A key difference compared to spherical wavefront illumination is that a smaller sample to sensor distance is required. The fringe magnification of this method is approximately 1, which results in a FOV equal to the sensor area. Therefore this method is suited for applications that demand for a large FOV and high throughput. Any light that hits the sample can be considered to be a plane wave because the distance between the illumination source and the sample is much larger than the distance between sample and sensor and, therefore the digital reconstruction does not have the same sampling and aliasing related issues that spherical wavefronts have. Additionally, spatial and temporal coherence of the light source can be significantly lower due to smaller path length between sample and sensor. The main limiting factor for spatial resolution is the physical pixel size and not the width of the sensor.

Biomedical use of this platform has already been demonstrated for blood analysis, imaging antibody microarrays, semen analysis, and the detection of waterborne parasites [27]. The pixel size at the detector array now place an important role in spatial resolution, it may cause under sampling of high frequency fringes of a lense free hologram. Another limitation can be considered the need for computation in order to reconstruct images. The application of planar wavefront digital holography for imaging Schistosomiasis Haematobium (SH) eggs has already been shown by T.E. Agbana [8], of which the resulting image is shown in Figure 2-3. It has shown to provide sufficient resolution in order to visually identify the eggs, and can be applied with low cost off-the-shelf components.

2-1-2 Shadow Imaging

Shadow imaging is a method where images are obtained by recording the diffraction pattern of an incoherent or partially incoherent illuminated object. The main difference between this

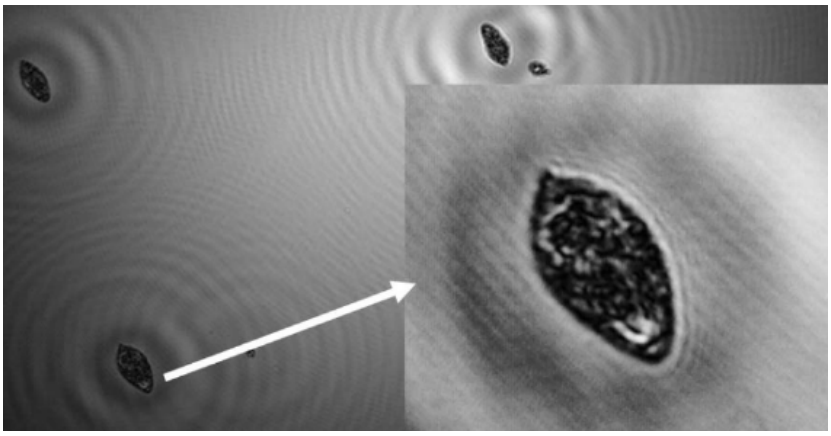


Figure 2-3: Reconstruction of Schistosoma haematobium eggs made from a holographic recording[8].

method and that of Digital Holographic Microscopy is that the diffraction between the object and detector plane is left unprocessed. The recording is not identical to the object itself since diffraction is unavoidable. However, the image does contain characteristic information about the object so it can still be of use for classification. Shadow imaging has similarities with planar wavefront holography since the FOV is equal to the active sensor area. However the non coherent and uncontrolled nature of the illumination prevent the use of wave propagation algorithms to digitally undo the effect of diffraction. The spatial resemblance of the object to its shadow strongly depends on:

- 3D nature of the object since only 2D shadow is measured without the ability to digitally refocus or reconstruct
- the sample to the detector distance
- pixel size

The main advantage of shadow imaging is its simplicity for screening a large FOV and the reduced computational power needed compared to other lens-free method. On the other hand, the microscopic spatial features cannot be distinguished from the shadow because of diffraction. To mitigate this shortcoming and contain high microscopic images using incoherent shadow imaging one has to significantly decrease the sample sensor distance in order to minimize the effect optical diffraction. The specimen could be put directly onto the sensor surface this is also called contact imaging. A contact imaging geometry can only image object directly at the surface and as such not suited for imaging objects in a 3D geometry. Additionally only a limited amount of samples can be analyzed by one sensor if they are put directly on the surface.[27]

2-1-3 Lens-free Fluorescent Imaging on a Chip

Fluorescent imaging works by exciting an object with a specific wavelength, this light is then absorbed by the fluorophores present in the material. The light emitted by the fluorophores is

captured while the excitation light is filtered out. Since no lenses are involved in this imaging modality , the fluorescent emission of cells or particles will diverge, potentially causing overlap at the sensor plane. This also limits the attainable resolution. These problems can be worked around by applying reconstruction algorithms as well as utilizing specialized optical elements such as nanostructured surfaces or tapered fiber optic face plates.[27]

2-1-4 Method selection

This section has presented several methods for lensless imaging. Shadow imaging does not provide sufficient resolution nor is it capable of imaging 3 dimensional space. Fluorescent imaging does not provide the spatial resolution for identifying the distinct features of the eggs. In-line Digital holography is a lensless imaging method capable of imaging microscopic specimen with a resolution limited by sensor area for spherical wavefront illumination and pixel size for planar wavefront illumination. Spherical wavefront holographic microscopy requires a coherent wavefront and sufficient distance between the sensor and the object. The planar wavefront digital holography can be performed with simple LED's and a very short sensor to object distance which means it requires less space. Additionally it has been shown that the eggs can be reconstructed with sufficient resolution. Lensfree on chip holography lends itself to make compact cost-effective and mechanically robust architecture which can especially useful to build microscopes for field use in low resource settings.

2-2 Flow Cytometry

This section provides the information needed for implementation of a flow cytometry system for recording a urine sample. First background information is presented on the basics of flow cytometry. Next the sample administration method and its requirements are discussed. Lastly the recording method will be considered.

2-2-1 Background information

Flow cytometry is a technique that employs an optical-electronic detection device to analyze the physical and chemical properties of microscopic particles suspended in a liquid medium. Generally, a light beam is directed through a continuous flow of suspended particles which are often marked with fluorescent substances. The light beam is scattered by the suspended particles and subsequently captured by sensors. Flow cytometry is routinely used in clinical diagnostics, biotechnology, and research[30].

The main components of a flow cytometer are a flow chamber, a light source, a detector, and a digital analogical converter which processes the relevant parameters. In the flow cytometer, the particles of interest are carried to the laser interception point in a fluid stream. Any suspended particle or cell from $0.2 - 150[\mu m]$ in size is suitable for analysis. The portion of the fluid where the sample particles are is called the sample core[9].

The sample core is usually injected into a stream of sheath fluid within the flow chamber. The flow chamber is designed to cause the sample core to be focused in the center of the fluid to ensure it is hit by the light source, this can be seen in Figure 2-4. [9]

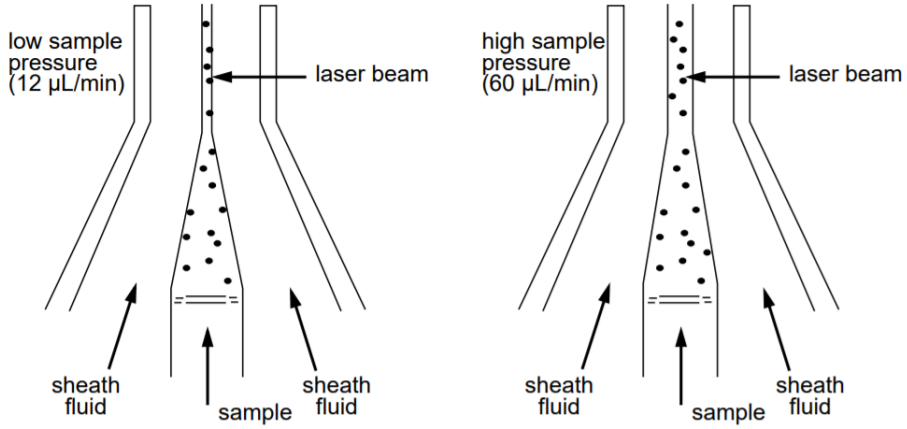


Figure 2-4: Hydrodynamic focusing of the sample core through a flow cell[9]

Imaging Flow Cytometry (IFC) combines the high-throughput, multi-parameter capabilities of conventional flow cytometry with morphological and spatial information. [31] M. Doan et al note that IFC has many potential clinical uses especially when it comes to diagnostics. Furthermore, they mention that deep learning is well suited for diagnostic tasks such as identifying colon cancer cells and disease progression. So far IFC is primarily used in research rather than clinical practice. This is caused by the variation of results, manual tuning, and interpretation. These issues might be overcome with machine learning approaches. [31] The literature hints at the validity of flow cytometry to the application of this research. However, as M. Vendel pointed out sample preservation might be preferred. Therefore combining the flow with sheath fluid is not desired. The effect of this on the flow properties of the sample core should be assessed[32]. If this turns out to be a problem, concessions should be made for example by using saline solution as sheath fluid.

2-2-2 Sample administration through the flow cell

In order to reduce human interaction the sample should be administered automatically without human contact. This ensures consistent recording for every sample. This can be achieved by using a stepper or servo motor to automatically pump urine samples through a flow cell using a standard syringe. The following points should be accounted for when implementing this system, as they are necessary to obtain an accurate hologram:

- The occurence of air bubbles should be minimized by airtight seals
- Air bubbles that have appeared should be removed from view automatically
- Control of the syringe should be steady such that the image quality is not affected by vibrations

2-2-3 How to record and flow

Next to the method of sample administration, the way to record the flow should also be considered. The recording method influences the amount of data generated and the amount

of objects that can be captured. The following two recording methods are considered.

Continuous recording

By recording a continuous video of the urine flowing through the flow channel most objects are captured in multiple frames while moving through the flow channel. This method has a better chance at capturing an egg at the right orientation, with the terminal spine being in view. Additionally the eggs can be tracked through multiple frames to ensure that no eggs are double counted. The downside of this method is that a large amount of data is generated which slows down the analysis of the video. Diagnostic time is important in a Point of Care (POC) setting since it limits the population size that can be diagnosed in a single day. Application of this method would require implementation of tracking algorithms as well as optimization of computing time by selecting the proper microcomputer. This was deemed out of the scope of this research, therefore this method will not be used.

Piece wise recording

By neglecting fluid friction, the complete volume of the syringe can be captured by moving the piston of the syringe by the amount of volume in view of the sensor and subsequently capturing an image. Doing this iteratively until the syringe is empty yields a batch of images containing the entire volume of the sample. Ideally, this means that any egg which may have been present in the sample has been captured. The advantage of this method is that it is the fastest way to capture the whole volume of the sample. However, the negligence of friction does not hold true causing the urine to flow slow at the boundaries of the flow cell and fast in the middle. Some eggs flowing near the boundaries maybe captured over multiple frames while eggs flowing in the middle might be missed completely. This can be compensated by adjusting the flow speed but one of these two problems will always exist. This is the preferred method as the implementation is the least complex whilst still showing the practicality of flow cytometry.

2-3 Egg detection

This section will elaborate on how the eggs can be found in the images in an efficient manner. First, the order of computational steps is laid out, next the method of object localization is presented and finally the classification of objects is discussed.

2-3-1 Algorithm layout

It is necessary to look into the order in which the computational steps are performed in order to design an efficient detection algorithm. There are several computational steps which need to be performed before a urine sample has been analyzed completly. These are in no particular order:

- Flow sample and record video

- Make image reconstructions across several planes over the depth of the flow channel
- Determination of object locations
- Classification of objects

The time it takes to determine a diagnosis is essential for a diagnostic device [32]. Therefore, these steps should be performed in a way in which computational time is minimized. Simply reconstructing whole images will take too long so the reconstruction and classification process will have to be performed in a different way. The eggs are expected to be sparsely represented in a sample, so unnecessary reconstructions and classifications should be avoided. Therefore, some processing needs to be done before the eggs can be counted.

For the diagnosis, the heaviest computational steps should be limited in the amount of times they are run. Reconstruction is unavoidable and should therefore be performed on a smaller area. The same goes for object location determination, which would take too long when performed on the full resolution video. By considering this and the computational analysis presented in Chapter 3, the following optimal algorithm flow is obtained:

1. Record interference pattern.
2. Keep original pattern and create a reduced resolution version.
3. Pick out the moving objects from the reduced resolution version of the pattern.
4. Locally reconstruct the original frame at the object location.
5. Classify the reconstructed object.

The reason of this order of operations has been chosen is in order to apply the heavy computational algorithms such as image reconstruction and classification on smaller images which saves time. From this it can be inferred that reconstruction and classification should be used last, and that potential objects should be selected from the fringe patterns in the raw image data and not from the reconstructions. The following subsections are dedicated to elaborating on how steps $3-5$ in the list above will be realized.

2-3-2 Object location determination

The following paragraphs discuss the methods that are applied in order to localize the sparse amount of objects present in the urine sample.

Foreground detection

The recorded images will have a view of the flow-cell and the urine that is flowing through it. The eggs will appear as fringe patterns occasionally moving over the view, which means that there is a relatively small area of interest compared to the FOV. The detection of the moving Foreground (FG) in a scene is an active field of research, as shown by T. Bouwmans [33]. By using these methods it becomes easier to find the areas of interest since moving

objects are distinguished clearly from the stationary Background (BG). There are several methods which can be used to perform this task, all with their specific advantages in certain situations. However, these will not be looked into extensively because implementing all these and finding the most optimal method is outside of the scope of this research. Instead, the methods provided by the OpenCV library will be used for the initial algorithm design[11]. The literature points at the Gaussian Mixture Model (GMM) being the most applied method for similar applications over the years [33]. This method is available for use in the OpenCV library, and as such it will be used to model the BG of the recorded frames. The other methods of the OpenCV library mainly focus on FG detection with shadows or dynamic scenes, both of which do not apply in this case. The GMM method is a pixel based BG subtraction method which involves a pixel-wise decision on whether it belongs to the BG or FG. Each pixel is modeled by a mixture of K Gaussian distributions. The probability that a certain pixel has a value of $\mathbf{x}_N$ at a time N is described by Equation (2-5)[34][35]. It is because of this pixel-wise decision that a downsized video is being used for detecting the foreground, as this can significantly reduce the computational load of this step.

$$p(\mathbf{x}_N) = \sum_{j=1}^{K} w_j \eta(\mathbf{x}_N; \theta_j) \tag{2-5}$$

Where w_k is the weight parameter of the k^{th} Gaussian component and $\eta(\mathbf{x}; \theta_j)$ the Normal distribution of the k^{th} component. For the method being implemented according to P. KaewTraKuPong [34], the K distributions are ordered based on a fitness value and the First B distributions are used as a model of the background of the scene where B is estimated according to Equation (2-6).

$$arg_b \quad min(\sum_{j=1}^{b} w_j > T) \tag{2-6}$$

Where the threshold T is the minimum fraction of the BG in the model, or the minimum prior probability that the background is in the frame. BG subtraction is then performed by marking any pixel that is more than 2.5 standard deviations away from any of the B distributions as a FG pixel. The Gaussian component that matches the test value will then be updated[34]. The result is a binary FG mask which indicates which pixels are marked as FG. The FG mask marks the fringe patterns caused by moving objects while ignoring stationary fringe patterns due to dents and objects that are stuck. It is therefore well suited for determining the exact location of eggs as they move accros the sensor area.

Blob Detection

Blob detection methods are aimed at detecting regions of images that differ in properties such as brightness or color compared to their surrounding region. The FG mask consist of a black image with white spots at the location of moving objects, therefore blob detection is suitable for the determining the locations of these areas. The three most common methods depend on the Laplacian of Gaussian, difference of Gaussians, or the Determinant of Hessian (DOH). Out of these three, the latter has been shown to have scaling properties as well as a lower computational load. The former two perform better on less distinct regions however this is not required because of the FG properties that were mentioned earlier. Therefore a DOH blob detector will be used to locate the moving objects in the images.

A DOH blob detector finds regions of interest by looking at the determinant of the local Hessian H_σ according to Equation (2-7). In this equation σ denotes the scale parameter of the image, H_σ denotes the Hessian matrix at a specific image location in level σ and ∂ denotes the second order derivative of the image along the axis indicated in subscript. A scale space is built by blurring the image repeatedly with a Gaussian function in order to achieve scale invariance. If the value computed by Equation (2-7) is above a certain positive threshold the region is labeled as a blob.

$$\sigma^4 \cdot det(H_\sigma) = \sigma^4 \cdot (\partial_{xx} \cdot \partial_{yy} - \partial_{xy}^2) \tag{2-7}$$

The original image will be reconstructed locally at the locations found by the blob detector. These reconstructions will be fed to classifier after some processing depending on which classifier is being used, this processing will be discussed next.

2-3-3 Classifier input features

The reconstructions will show an object surrounded by the semi-white background of the flow-cell similar to the reconstruction in Figure 2-3. Adding to this, not every area of the egg is of interest when classifying it. Some classifiers are able to deal with this reconstruction just fine, however; other classifiers may not and will require some preprocessing of the data to remove redundant information. For this reason the dimensionality of the data will be reduced. A useful method for this is Principal Component Analysis (PCA), which is a method to reduce the amount of features while retaining as much information as needed[36]. G. Şengül has shown that eggs can be classified by using the co-occurrence matrix as features [37]. Both of these methods shall be tested as input feature generators for the classifiers that do not perform well with the redundant information that is present in the reconstructions. The implementation of these methods will be discussed in the following paragraphs, starting with the general image alterations.

General image alterations

All the pixel values will be remapped to be in a range between $0 - 255$ regardless of which classifier is being used in order to enhance the contrast. After reconstruction the images will be of relatively high resolution in terms of classifications ($> 100 \times 100$pixels). Although this resolution results in a detailed image, it will also require a large amount of training data. For instance if the image is fed directly to a classifier and at least one training sample is required for each feature, at least 10000 training samples would be preferred (this depends on which classifier is being used). In order to account for this, the images will be reduced in size for classifiers that have the image as input.[38]

PCA features

PCA is a statistical method which uses orthogonal transformation in order to convert a set of, possibly correlated, variables into a set of linearly uncorrelated variables called principal components. This is done by determining the set of singular value - singular vector pairs from the covariance of the set. Each singular value then represents a portion of the variance of the

dataset. The datapoints can then be represented by the variance along each singular vector. At this point the amount of singular value - singular vector pairs is equal to the amount of variables in each observation. However, most of the variance of the set is often represented by a portion of the pairs. Therefore, much of the data can be retained when dropping some singular value-singular vector pairs. This is called dimensionality reduction. This procedure is performed on $n \times n$ gray-scale images as follows:

1. Flatten each image into a n^2 length array.

2. Subtract the mean of each pixel value from the corresponding pixel in each image.

3. Determine the $n^2 \times n^2$ covariance matrix of the pixel values.

4. Perform Singular Value Decomposition (SVD) on the covariance matrix (Equation (2-8)). The singular vectors in U represent the axes of maximum variance, while the singular values in Σ represent the variance of the set along that axis.

5. Select the singular value - singular vector pairs which represent about $90 - 95\%$ of the variance. The reduced set of singular vectors is called $U_{reduced}$.

6. Principal components can be determined by multiplying new flattened images with the singular vectors according to Equation (2-9). The resulting values are called the principal components and represent the image with less information needed.

$$C = U\Sigma W^T \tag{2-8}$$

$$Components = I_{flattened} \cdot U_{reduced} \tag{2-9}$$

The variance accounted by every principal component should be analysed in order to select the amount of components. This can be done by ordering the singular values contained in Σ by their magnitude. By plotting the accumulated variance, starting from the highest variance, over the amount of singular values a visualization is obtained which shows the total variance accounted for by the singular vectors when representing the original data. An example of this is shown in Figure 2-5. With this dimensionality reduction a significant portion of the variance of the data can be compressed in a few parameters. The amount of components should be selected such that $90 - 95\%$ of the variance is accounted for. This way the distinguishing features are retained within less data, resulting in less training data needed and faster classification[36]. The amount of components that are dropped from the singular value - singular vector pairs are a tuning parameter which can be changed in order to control the model accuracy and run time.

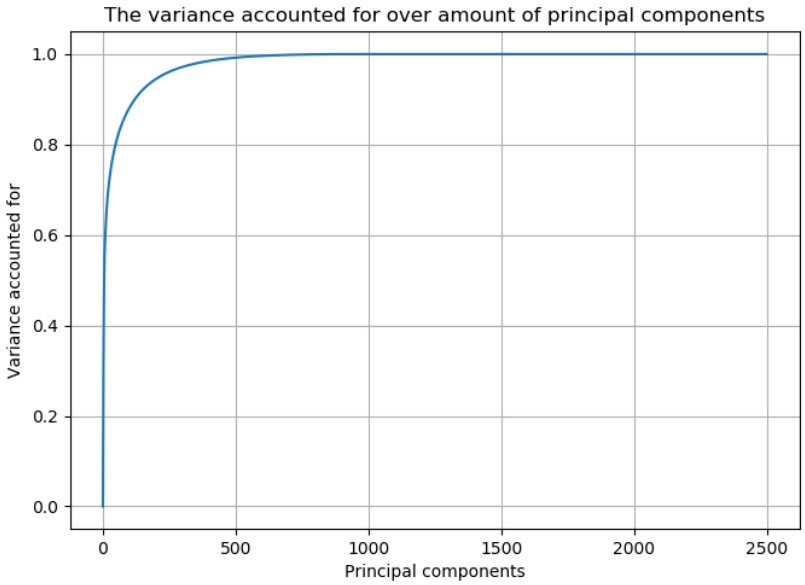


Figure 2-5: The amount of variance accounted for in the data set by the number of singular vectors, starting from the singular vectors that account for the most variance. The dataset consisted of images with a resolution of 50×50 pixels

.

Co-occurence matrix features

A co-occurrence matrix is defined as the distribution of co-occurring pixel values at a given offset within an image. By looking at the way pixels appear next to each other, a measure of texture is obtained. The resulting matrix is often sparse and can be used to generate features using other methods. However, the matrix entries themselves can also be used as features[37]. The list below shows how a co-occurrence matrix can be obtained from an image.

1. Rescale the pixel values as integers between 0 and a desired value(k).
2. Analyze the images pixels for neighboring values according to equation(2-10).
3. flatten the $k+1 \times k+1$ array into a $(k+1)^2$ array. This array can be used as features for a classifier.

$$C_{\Delta x,\Delta y}(i,j) = \sum_{x=1}^{n}\sum_{y=1}^{m} \begin{cases} 1, & \text{if } I(x,y) = i \text{ and } I(x+\Delta x, y+\Delta y) = j \\ 0, & \text{otherwise} \end{cases} \tag{2-10}$$

This method of feature generation adds another parameter, The pixel values described in step 1 in the list above, which can be tuned in order to optimize the model's accuracy and runtime.

2-3-4 Classification

There exist a variety of classifiers, each with a specific purpose[39]. The selection of a classifier for the algorithm prototype will be based on classifiers in similar research. There has been research on the classification of red blood cells, which are round in shape[40][26]. Additionally, recently there have been efforts to classify Schistosoma eggs as well as mentioned in Chapter 1[41][21]. The most commonly applied classifiers in these kinds of research are Support Vector Machine (SVM) and Convolutional Neural Network (CNN). These classifiers are widely used, and expecially CNN are flexible and can be tailored to fit almost any application regarding image classification. Lastly, a Random Forest Classifier (RFC) will be tested since it is a flexible algorithm which sees application in many different fields including image classification[42]. These classifiers are briefly discussed in the following three paragraphs. There are many more methods which could be tested but they will not be considered for this prototype[43].

SVM

SVMs are techniques for data classification first mentioned by Cortes and Vapnik in 1995[44]. The goal of an SVM is to produce a model predicting the target values of the input data. Input data of any kind is flattened into a vector which represents a point in a hyperspace, in this case the hyperspace will consist of pixel values, principal components, or the co-occurence matrix. The SVM then draws a hyperplane between the set of points labeled as egg and the set of points labeled as not egg[41][40]. An SVM can be tailored by selection of the kernel: linear, polynomial, or Radial Basis Function (RBF) as well as by tuning the hyperparameters γ and C[44]. C is the penalty term for the error during optimization and γ the kernel coefficient. The γ parameter represents the variance of the set, with low values representing a high variance within the data. When γ is very small, the model will be similar to a linear model. On the other hand, when γ is very large the model will overfit. Of all the kernels, RBF is the most used because of it's flexibility. Therefore RBF shall be used as the kernel function during this initial implementation[39]. γ And C are then the relevant tuning parameters which will affect the performance.

RFC

A RFC consist of a large number of individual decision trees that operate as an ensemble. Each tree returns a classification with a final classification being the class with the most votes. The idea behind the random Forest classifier is that a decision made based on many simple, inaccurate classifiers will be accurate. RFC make use of bagging and feature randomness when building each individual tree to try to create an uncorrelated forest of trees whose prediction by majority vote is more accurate than that of any individual tree. The main hyperparameters are the amount of trees in a forest, the number of decisions made in a tree, and the amount of features that are considered for each tree. Later on the RFC implementations will make use of the principal components and co-occurrence matrix as input features.

CNN

CNNs are a class of supervised deep learning techniques. CNNs are mainly used for image classification and are therefore suited for recognizing Schistosoma haematobium eggs as well. Adding to this: CNNs have been applied to detection of Schistosoma eggs and similar cases in the past, though these efforts have not made it past the testing stages [24][26][21]. Designing an optimal CNN for this is a complex procedure which requires expertise on the subject. For the sake of simplicity the CNN that will be implemented in this thesis will be based on other implementations that are similar in complexity. There are several options which fall outside of this scope such as MobileNets introduced by A.G Howard in 2017[45] or Faster R-CNN[46], which should be considered in a later design stage. One of the standard databases is called MNIST[47], and is frequently used to train new classifier models without spending effort on gathering data. The dataset consists of low-resolution grayscale images of handwritten digits. These digits are relatively simple in shape, much like the eggs that have to be recognized. There are examples of CNN implementations using this dataset which achieve an out of sample accuracy of more than 98%. Several of these implementations will be tested for detecting eggs in order to assess the applicability of CNN to classify reconstructed images, these will use the pixel values as inputs. [47] Three models were selected for testing, their architechtures and accuracies are shown in Table 2-1. Details of these models can be found in appendix 5-2-5.

Table 2-1: Three CNN models that were used as examples for testing the applicability of CNN for detecting eggs. All of these architectures will have their input and output layers modified to suit the purpose of classifying eggs. Details of these models can be found in appendix 5-2-5.

Model name	Architecture
Yalcin [47]	$Conv2D - MaxPooling2D - Flatten - Dense - Dropout - Dense$
Ghouzam [48]	$Conv2D - Conv2D - MaxPooling2D - Conv2D - Conv2D - MaxPooling2D - Dropout - Flatten - Dense - Dropout - Dense$
Keras [49]	$Conv2D - Conv2D - MaxPooling2D - Dropout - Flatten - Dense - Dropout - Dense$

2-4 Diagnosis

The algorithm designed so far in this chapter finds results on a microscopic level: an x amount of eggs and y amount of other objects. This section will discuss how these results should be interpreted in order to draw a diagnosis comparable to the World Health Organisation (WHO) standard. First the general procedures for evaluating new diagnostic test is laid out. Then the complications of basing a diagnosis on classifier count is presented, and a solution is proposed.

2-4-1 Diagnostic evaluation

In diagnostics, the accuracy of a newly developed method is usually compared with an independently established standard diagnosis, usually this is called the "Gold standard". As the name implies, the gold standard should provide a full certainty diagnosis, but this is

rarely actually the case. Table 2-2 is an example of a comparison between a golden standard and another method. Such a table can be used to determine statistical characteristics of a diagnostic method. The two most important metrics are a test's sensitivity and specificity, shown in Equation 2-11, which are independent of prevalence within the test population[50]. They represent the probability of a correct positive and negative diagnosis respectively(true positive and negative rates).

$$\begin{aligned} sensitivity &= \frac{a}{a+c} \\ specificity &= \frac{d}{b+d} \\ accuracy &= \frac{a+d}{n} \end{aligned} \tag{2-11}$$

Table 2-2: Table comparing the results of an evaluated method compared to the gold standard.

	Gold standard		
Evaluated method	positive	negative	total
positive	a	b	$a+b$
negative	c	d	$c+d$
total	$a+c$	$b+d$	$a+b+c+d=n$

The sensitivity and specificity of a test determines its usefulness. For diagnosis, having just a high sensitivity is alright because a positive outcome will be reliable. On the other hand, having just a high specificity makes it useful for screening because a negative outcome will likely be true and therefore rules out people who do not have the disease [50].

2-4-2 Diagnostic problem

The analysis of the sample results in an egg count and not egg count, however; the classifier is not perfect. In fact, Table 2-2 can be used for a classifier as well except now it contains the correct and incorrect classifications, this is called a confusion matrix. Depending on the classifier and the training data there is a different accuracy for positive and negative predictions. For positive predictions there is a probability that not all of those predictions are actually true, the same goes for a negative prediction. Consider the following hypothetical situation:

> The algorithm has 95% accuracy and detects 50 eggs and 950 other objects. Finding 50 eggs in a sample usually means that the patient is heavily infected[3]. However, the algorithm seems to have found plenty of other objects. No direct diagnosis can be drawn from this since the 50 eggs that were detected could be false positives, judging from the algorithm's performance.

As described above, the diagnosis drawn from the result of the algorithm is not equivalent to the WHO procedure as described in Chapter 1. High accuracy (99%+) can provide accurate counts most of the time but outliers should be considered nonetheless, such as when a lot of

blood or dust is present in the sample. By applying statistical methods an estimate can be made of the egg count with a certain Confidence Interval (CI) based on the confusion matrix of the classifier.

2-4-3 Statistical method for automated diagnosis

Consider a classifier trained with a diverse, uncorrelated sample pool with a sensitivity p and a specificity q determined from an independent data set. This classifier has a probability p of correctly detecting an egg and probability $1-q$ of mistaking a different object for an egg (type 1 error). Therefore the egg count resulting from a sample analysis is not the true egg count present in the sample. A CI of the actual amount of eggs would provide better insight in the quality of the test. The egg count can be modeled as a Poisson binomial distribution where there are n eggs that can be classified as eggs with probability p and m other objects that be classified as eggs with probability $1-q$. The variance of the outcome of this test can be calculated according to Equation (2-12).[51]

$$Var = \sigma^2 = n \cdot p(1-p) + m \cdot (1-q)q \tag{2-12}$$

Although we cannot determine the real egg count, we can use the outcome of the test in conjunction with Equation (2-12) to determine a CI of what the actual egg count may be. The CI of the egg count can be calculated according to Equation (2-13). In this equation $n_{classifier}$ represents the egg count found by the algorithm, c a multiplier based on what CI is being used (z value), and $\sigma_{classifier}$ the standard deviation based on the algorithm output (calculated using Equation (2-12)).[52]

$$CI = n_{classifier} \pm c \cdot \sigma_{classifier} \tag{2-13}$$

The magnitude of this CI provides a measure of reliability of the test with large values indicating an unreliable outcome. This is shown by the simulations shown in Table 2-3. How this should be interpreted will have to be derived from actual tests.

Table 2-3: Table showing the confidence intervals of simulated test outcomes from a classifier with 0.96 sensitivity and 0.91 specificity (values arbitrarily chosen). Lower bounds have been rounded down and upper bounds have been rounded up.

Test outcome	CI actual egg count								
Eggs /Not eggs	90%			95%			99%		
20/10	17	–	23	17	–	23	16	–	24
50/50	45	–	55	45	–	55	43	–	57
50/200	42	–	58	41	–	59	38	–	62
50/1000	34	–	66	32	–	68	26	–	74
20/1000	5	–	35	2	–	38	−4	–	44

2-5 Method overview

This chapter presented the background information needed for the implementation of a low cost, smart diagnostic device for SH. First, in-line planar wavefront digital holography was

found to be the most applicable lensless imaging method when combined with flow cytometry. Next, the initial diagnostic algorithm was discussed. The moving objects will be found by a combination of FG detection and blob detection. The objects are then locally reconstructed in order to obtain images for the classifier. These images will then be fed to the classifier, of which three types will be tested: SVM, RFC, and CNN. Finally the result of these classifiers will be interpreted in order draw a diagnosis. A schematic overview can be seen in Figure 2-6The next step is to verify the use of digital holography and flow cytometry experimentally, and to test the potential results on the implemented algorithm.

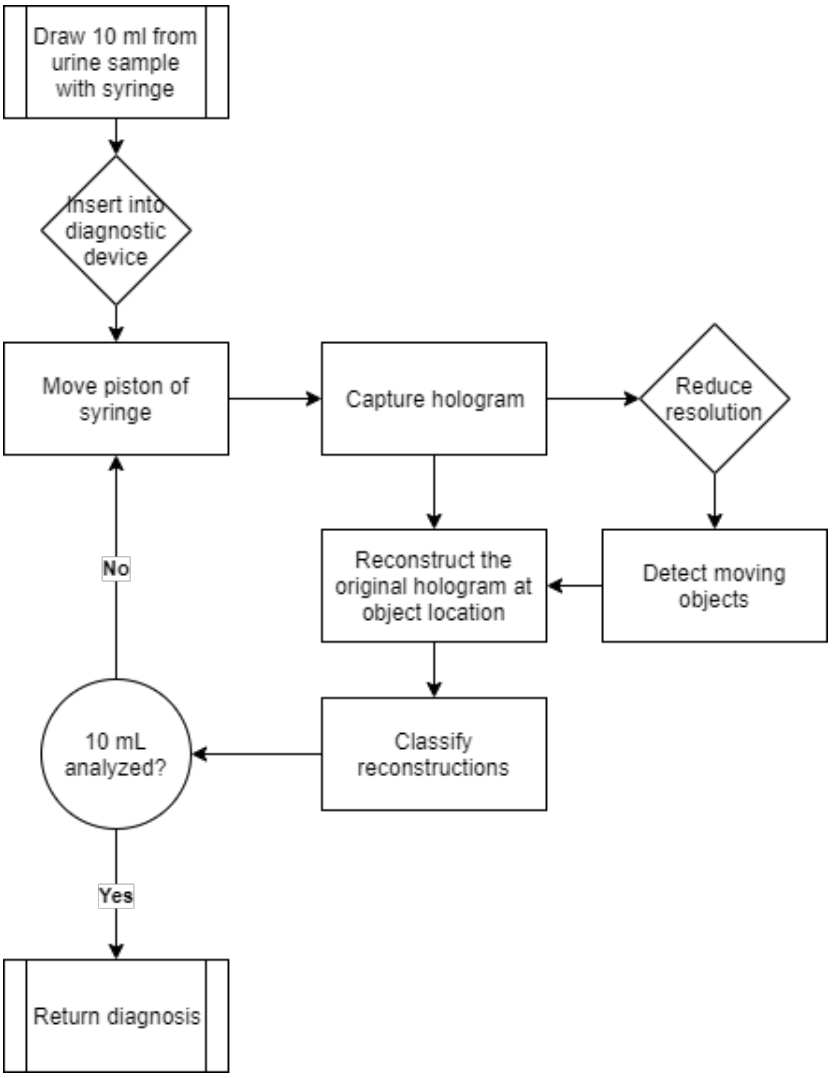


Figure 2-6: Simplified block diagram of the proposed diagnostic method.

Chapter 3

Diagnostic Method Verification

Chapter 2 presented a possible automatic diagnostic method for Schistosomiasis Haematobium (SH). In this chapter, a preliminary version of this method is implemented for verification purposes. To start off, Section 3-1 will verify the optical setup by showing that SH eggs can be imaged in flow using a DHM. Section 3-2 follows by presenting the result of the implementation of the object detection algorithm. Next, the performance of the various classifiers that were proposed is discussed in Section 3-3. The various steps of the algorithm are benchmarked in Section 3-4. Finally the observation are discussed in Section 3-5.

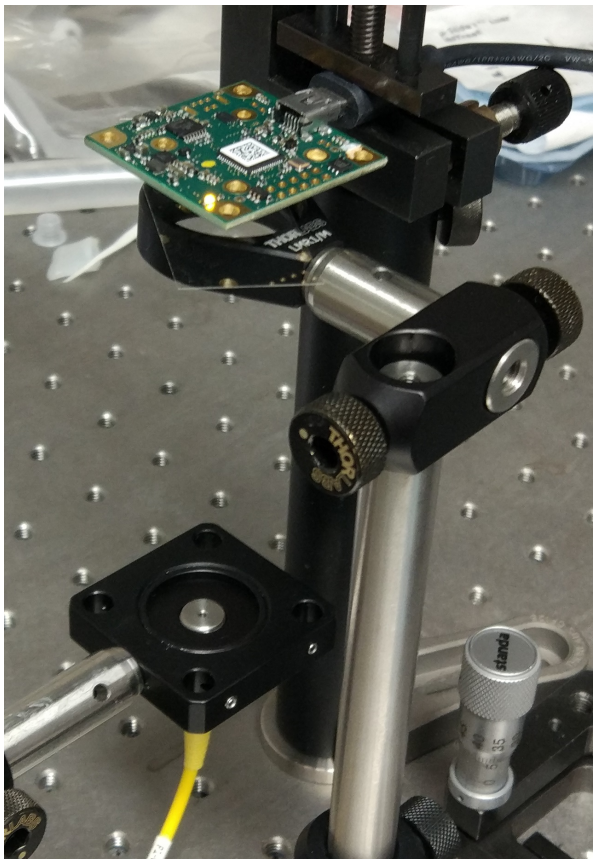


Figure 3-1: In-line Digital Holographic Microscope (DHM) setup, from top to bottom: sensor chip, object on glass slide, illumination source.

3-1 Recording method

This section presents the experimental results obtained from the setup combining a DHM and a flow cytometer, which constitutes the recording method. First the experimental setups are described in detail, then the optimal performance of the DHM is determined, and lastly the proposed recording method is verified for its applicability for imaging SH eggs.

3-1-1 Setup specifications

Two setups were built: one in order to determine the attainable resolution of the DHM, and one to verify whether SH eggs can be recorded in flow. Both setups used the same illumination source and sensor chip: A S1FC635 Fiber-Coupled Laser Source from Thorlabs[53] and a IDS UI-1492LE-M Complementary Metal-Oxide Semiconductor (CMOS)[54] respectively. The specifications of these instruments are listed in Table 3-1.

Table 3-1: The specifications of the holographic imaging setup[14].

Property	Value
Sensor resolution[pixels]	3840×2748
Sensor dimensions[mm]	6.413×4.0589
Pixel pitch Δ_p $[\mu m]$	1.67
Framerate [fps]	3.2
Minimum exposure time[ms]	0.340
Laser wavelength λ [nm]	635
Laser spectral linewidth [nm]	2.5
Maximum output power [mW]	8.0
Setpoint resolution[mW]	0.01

Resolution test setup

The resolution test setup consists of the illumination source and sensor chip mentioned before, with a stationary object placed in between. The first object to be tested is a USAF1951 resolution target which is commonly used to determine the resolution of optical instruments[10]. Afterwards a sample containing SH eggs in saline solution between two microscopic glass slides shall be used as an object to verify the resolution. This setup is shown in Figure 3-2. The minimum object - sensor distance of this setup is physically limited to $4[mm]$.

Flow imaging setup

The flow imaging setup is similar to the resolution test setup but has a flow cell at the object location. The flow cell is provided by ibidi[55] and has a channel width of $5[mm]$ and depth of $0.8[mm]$. A sample of SH eggs in saline solution shall be administered by hand using a syringe. The flow is stopped once eggs are in view of the sensor after which an image is captured. The sample has to be imaged when it is stationary, recording while the sample was still flowing resulted in distorted images.

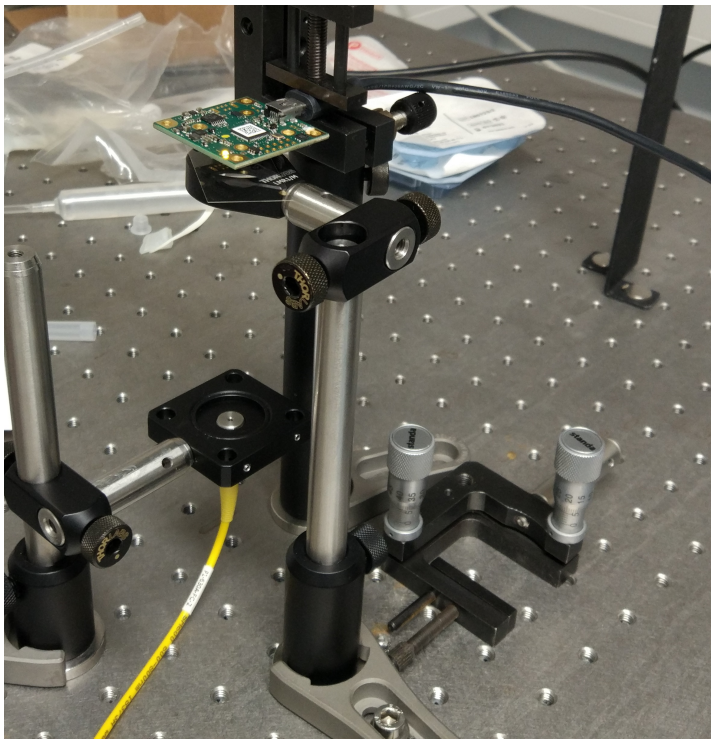
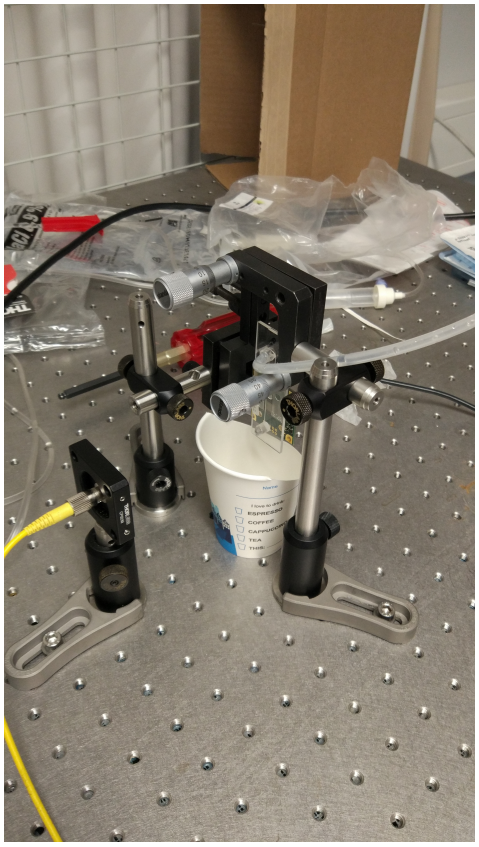


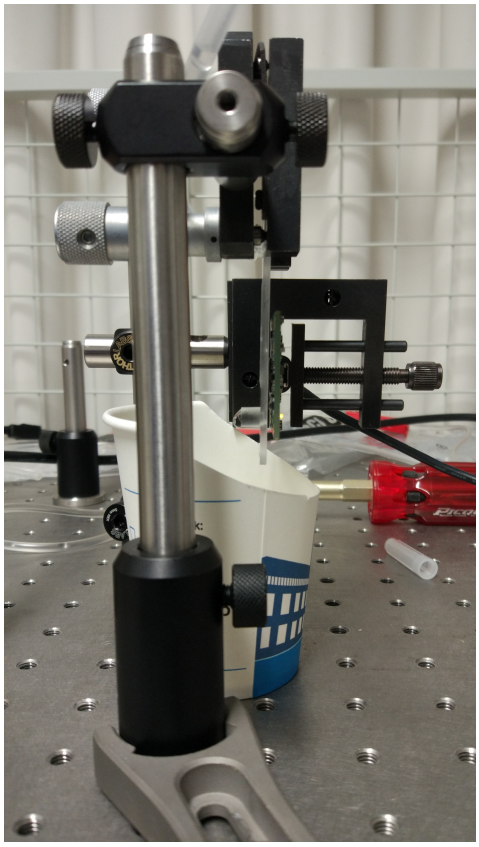
Figure 3-2: The initial configuration of the DHM setup for imaging stationary samples. In this image the object consists of SH eggs between two microscopic glass slides.

The exit of the flow cell will have a cup below in order to catch the sample for potential reuse. An airtight seal between the syringe - tube, and tube - flow cell is required to prevent air bubbles from entering the flow. Additionally, the setup is built such that the flow is descending at all times to prevent air bubbles from getting stuck in the system. The entire flow setup can be seen in Figures 3-3a and 3-3b. At first the flow shall be administered by hand, however; to avoid any inconsistencies and mistakes this process shall be automated in later stages.

In Chapter 2 it was determined that the sample will be imaged by the piece wise recording of fluid elements. Therefore it is important to know the amount of volume in view by the sensor, this is determined by the channel and sensor dimensions. The area in view of the sensor is $5.0000[mm] \times 4.0589[mm]$ and the depth is $0.8000[mm]$, which combines to a volume in view of $16.236[mm^3] = 0.016236[mL]$. The sample will be analyzed by pressing the piston of the syringe by this amount of volume and taking an image when the content of the flow cell is stationary. This is performed repeatedly until $10[mL]$ of the sample has been administered by the syringe. The syringe will be controlled by a stepper motor which ensures that every step is the same. In order to analyze the entire volume at least $\frac{10}{0.016236} = 616[-]$ images will need to be taken. This recording method shall be implemented in the prototype presented in Chapter 4.



(a) Global view of the setup.



(b) Side view of the setup.

Figure 3-3: Two views of the flow setup. The sample is administered by a syringe through a silicone tube, it is then caught in a coffee cup after flowing through the flow cell.

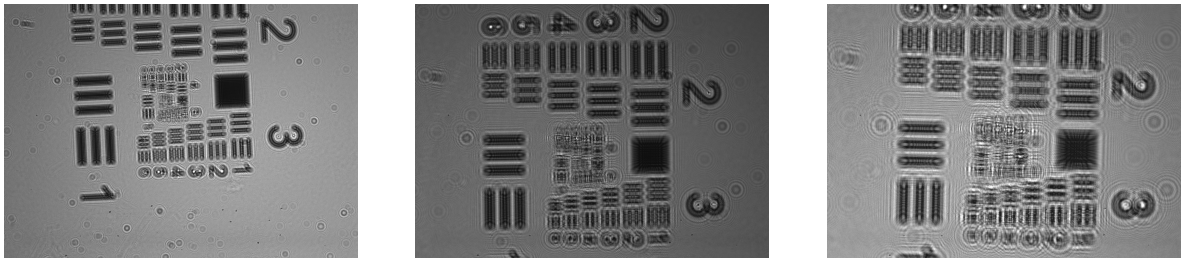
3-1-2 DHM resolution test

The next two paragraphs will verify the use of a DHM for the new diagnostic method. First the spatial resolution shall be determined, then an actual sample will be examined in order to verify the capability to reconstruct SH eggs.

USAF1951 resolution test

The USAF1951 resolution target contains a pattern which will be imaged, consisting of 9 groups each containing 6 line elements. The spatial resolution of the setup can be determined from the smallest discernible element. The target was put in front of the sensor at a distance of 4$[mm]$, 10$[mm]$, and 18$[mm]$. Multiple distances were used in order to verify whether a closer distance will result in better resolution. The recorded fringe patterns are shown in Figures 3-4a - 3-4c.

With the fringe patterns captured, the next step is to reconstruct them by using Equation (2-2). When reconstructing the images the aim is to be able to discern the smallest possible



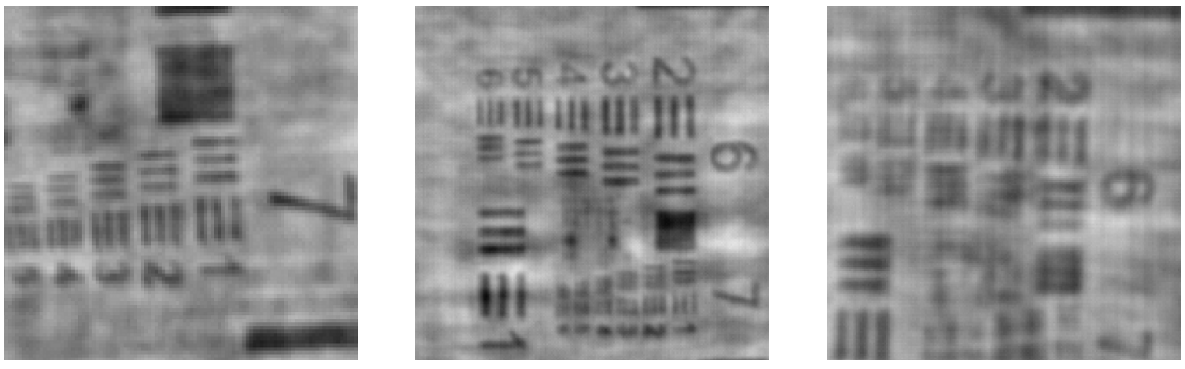
(a) $4mm$ **(b)** $10mm$ **(c)** $18mm$

Figure 3-4: Captured fringe patterns of the USAF 1951 resolution target[10]. The fringes become more apparent when recording and a larger distance.

line. For reconstruction the object - sensor distance should be known within micron precision. The optimal reconstruction distance was found according to the following iterative procedure:

1. Initiate range of reconstruction distance, e.g. $0 - 20[mm]$.
2. Reconstruct the image at several planes over the range.
3. Select the sharpest reconstruction.
4. Narrow the range of reconstruction down around the distance of the selected reconstruction.
5. return to step 2 until the differences in the reconstruction quality is negligible. Generally this occurs when the range of distances is within $20[\mu m]$, which is less than the thickness of an egg.

This procedure was performed for each of the holograms, and resulted in the reconstructions shown in Figures 3-5a - 3-5c. As can be seen, the reconstruction of the image taken at $4[mm]$



(a) $4mm$ **(b)** $10mm$ **(c)** $18mm$

Figure 3-5: Reconstructed images of the USAF 1951 resolution target[10]. The images have been zoomed in to show the smallest discernible elements. Element 5 of group 7 which can be seen in figure **(a)** has a line thickness of $2.46[\mu m]$.

distance shows the thinnest lines(group 7, element 5) which have a width of $2.46[\mu m]$[10].

As mentioned in Chapter 1 the eggs are about $135[\mu m] \times 50[\mu m]$ in size, so they will appear with sufficient detail to be recognized by eye when using the DHM. The full image of the reconstruction is shown in Figure 3-6, which is a reconstructed version of Figure 3-4a. Some fringes are still visible which is caused by twin image contamination, however; this effect does not contaminate Figures 3-5a-3-5c to the point of hindering recognition. Next the setup will be used to examine a sample containing actual SH eggs.

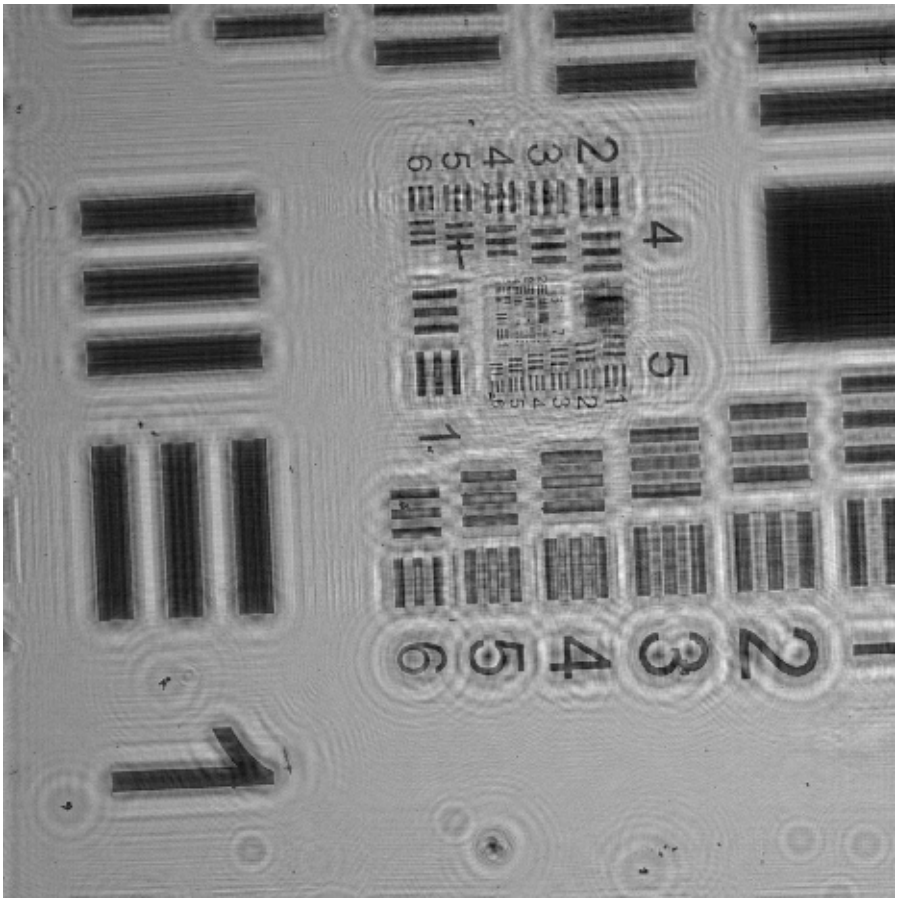


Figure 3-6: Reconstruction of the fringe pattern shown in figure 3-4a. Note that the elements of group 4 and beyond were very distorted in the original image, while the elements up to group 7 can be distinguished in the reconstruction.

SH sample examination

The previous paragraph showed that a spatial resolution of $2.46[\mu m]$ can be achieved with the set up, this will now be verified by examining SH eggs. A microscopic pipette is used to extract a small volume of saline solution containing the SH eggs. This sample was then put between two microscopic glass slides. Putting the sample in front of the sensor and illuminating it produced the image shown in Figure 3-7.

The sample is reconstructed according to the same procedure as explained earlier, resulting in the image shown in Figure 3-8. The eggs have a length of about 50 pixels in the reconstruction,

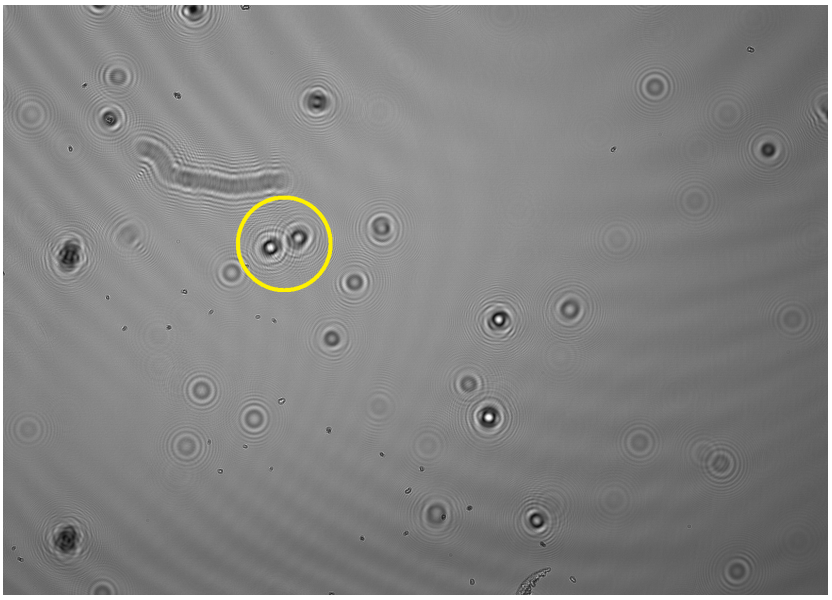


Figure 3-7: Fringe pattern of a sample containing SH eggs. The reconstruction will be shown on the location with the yellow encirclement. Fringe patterns resulting from a SH egg have a particular shape which is different from other objects observed in the sample such as dust.

supporting the obtained resolution of $2.46[\mu m]$ found earlier with help of the USAF 1951 resolution target. The reconstruction shows sufficient resolution to recognize the SH egg by eye.

3-1-3 Flow test

The time it takes to record the sample partly determines the diagnostic time. This is limited by the cross-sectional area of the flow cell, area of the sensor, and the data-rate of the camera (recording a lower quality image can result in a higher imaging rate assuming the data-rate stays constant). The flow cell and sensor cannot be changed, however the imaging rate can be changed. In order to account for this and obtain optimal image quality the following factors will be tested:

- The resolution: recording at a lower resolution makes it possible to have higher frame-rates.
- Camera - object distance: this influences spatial aliasing and resolution as shown in Section 2-2.

These factors will be analyzed by recording at two different settings: The resolution was put at 3840×2748 and 1920×1080 pixels, and camera - object distance was put at $1[mm]$ and $10[mm]$. Images were recorded under circumstances described by each of the mentioned settings, all of these are shown in Figures 3-9a - 3-9d.

The fringes are more concentrated with a low object - camera distance, which is in line with the theory presented in Section 2-1. However, the fringe frequency is not high enough for

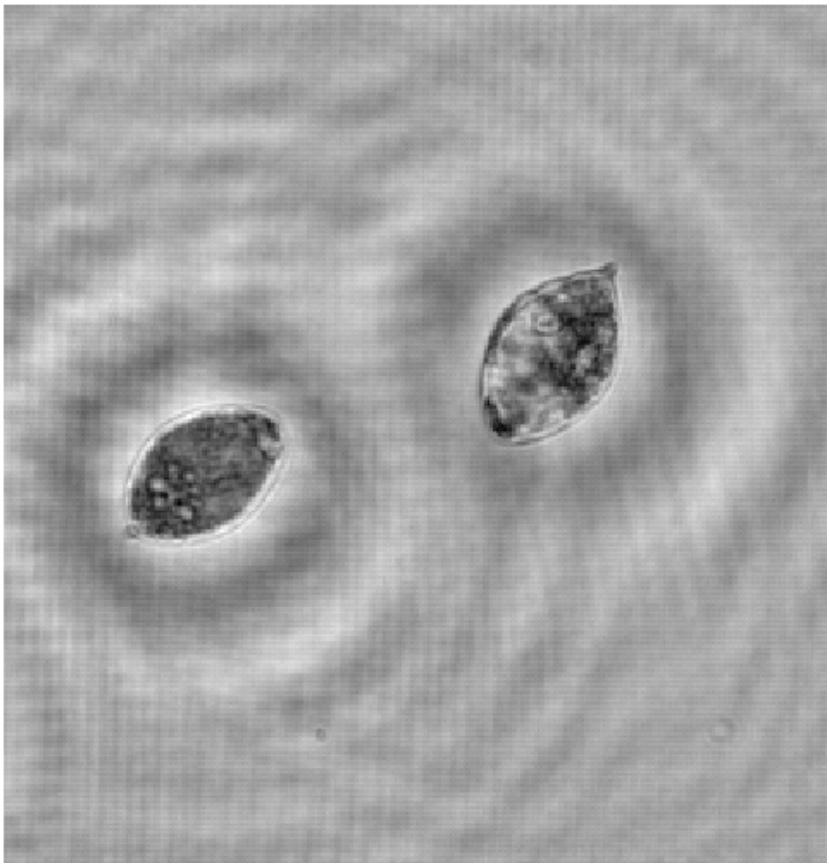


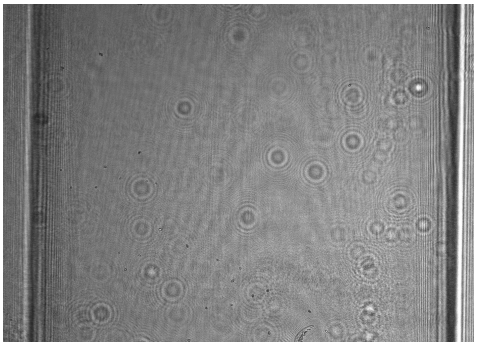
Figure 3-8: Reconstructed image containing two SH eggs caught between two microscopic glass slides. The image was reconstructed from Figure 3-7 at the yellow encircled area. The detail of the reconstruction is comparable to that of conventional microscopy (see Figure 1-4).

the occurrence of spatial aliasing. Each frame has been zoomed in on an egg location and reconstructed, the result of this is shown in Figures 3-10a - 3-10d. These images can be used for a comparative analysis in order to select the optimal recording circumstances for the reconstruction of SH eggs.

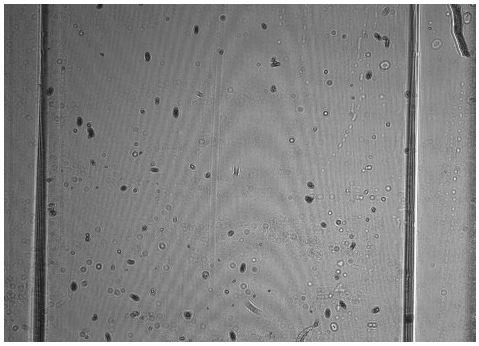
From the reconstructions in Figures 3-10a - 3-10d can be seen that Figure 3-10b contains the sharpest reconstruction of an egg. The two images taken at $1[mm]$ distance were all sharper than their counterparts taken at $10[mm]$ distance. Furthermore, recording at higher resolution shows more detail which is expected. However, recording at low resolution can result in a higher maximum flow rate so this should be kept in mind at a later design stage. For this proof of concept the images shall be recorded at the maximum resolution with minimal distance between object and sensor.

3-2 Moving object detection

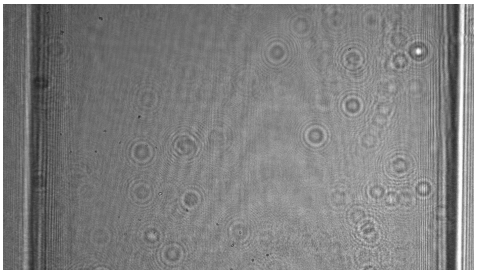
For testing this implementation, images were recorded using the automatic sample administration method mentioned in Section 3-1. The following 4 paragraphs show the implementation Foreground (FG) detection, blob detection, and local reconstruction to these images.



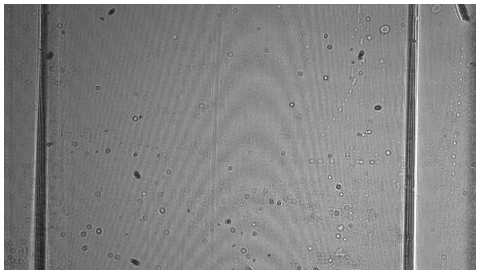
(a) High resolution, 0.08 mW, 10 mm



(b) High resolution, 0.08 mW, 1 mm



(c) Low resolution, 0.08 mW, 10 mm



(d) Low resolution, 0.08 mW, 1 mm

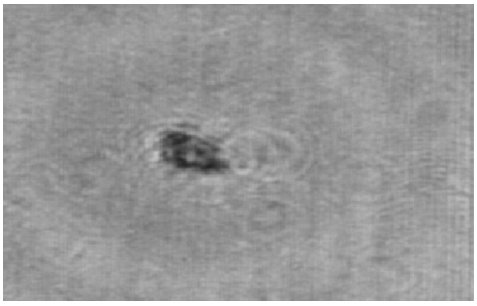
Figure 3-9: Recorded holograms of samples with different setup settings. Recording and a short distance results in distinct fringes however the spatial frequency is reduced which might cause aliasing to affect the quality of the reconstruction.

3-2-1 Resolution reduction

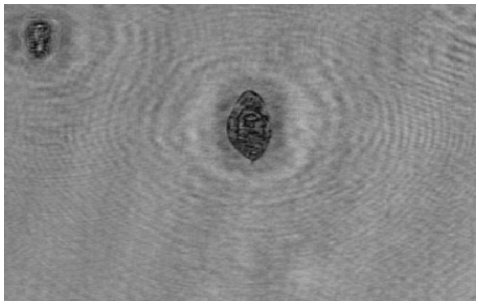
The methods used find the object locations by analyzing the pixel values of an image. Since the resolution of the sensor is $3840 \times 2748[pixels]$ these computational steps can negatively affect the time it takes to analyze a sample. Therefore the resolution shall be decreased 16-fold to $960 \times 687[pixels]$ by averaging the pixel values. The objects shall be found in these smaller images, their location is then traced back in the original image in order to perform local reconstruction.

3-2-2 FG detection

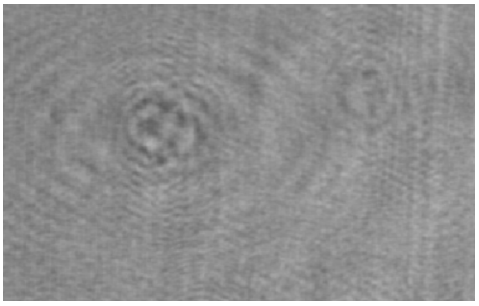
A Gaussian Mixture Model (GMM) was used to extract the moving objects from the stationary background. This resulted in a black canvas with white spots at the locations of moving fringe patterns as shown in Figure 3-11b, this is known as the FG mask. The fringes appearing on the FG mask are speckled which can be solved by dilating the image as shown in Figures 3-12a and 3-12b. Dilation enlarges the area of bright spots at the cost of amplifying the areas of the FG that were caused by noise and vibrations. However, this prevents the blob detector from detecting a fringe pattern coming from a single object as multiple objects.



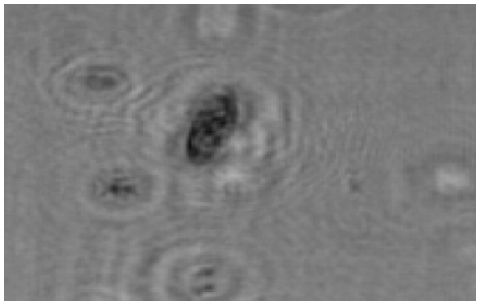
(a) High resolution, 0.08 mW, 10 mm



(b) High resolution, 0.08 mW, 1 mm



(c) Low resolution, 0.08 mW, 10 mm



(d) Low resolution, 0.08 mW, 1 mm

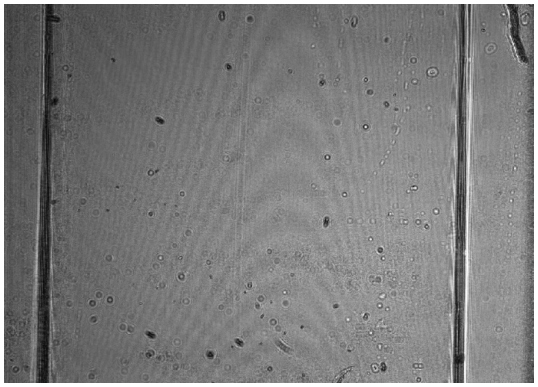
Figure 3-10: Reconstructions of the snapshots shown in Figures 3-9a - 3-9d.

3-2-3 Blob detection

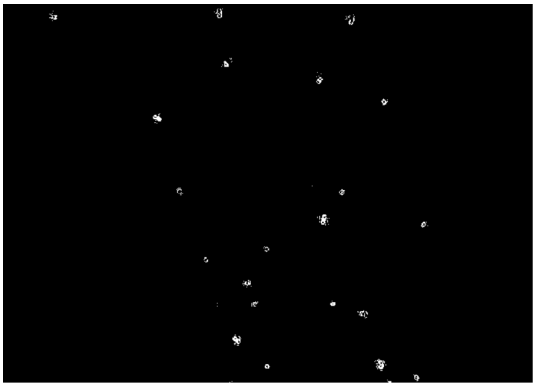
A blob detector [56] was used to automatically locate the white spots in the dilated FG mask. Figure 3-13b shows an example of this, all the white spots have been found and encircled automatically. The coordinates found by the blob detector correspond to the locations on the resized image. These are first translated back to the original image before reconstruction.

3-2-4 Local reconstruction

A clear image of the reconstruction at the object location is desired. The propagation distance is needed in order to obtain a sharp reconstruction. Since the object freely flows through the flow cell there is no exact knowledge of this distance, only that it is located along the depth of the flow cell. This is made clear in Figure 3-14, where eggs appear at different distances from the sensor. This means they have to be reconstructed at different distances for them to appear as sharp images. The least complex way of making sure a sharp image of the object exists is to reconstruct over several distances across the depth of the channel. This should be done in a way in which at least one of the reconstructions of an egg would be a sharp image. The depth of the channel that is being used is $800[\mu m]$ while the average egg width, its shortest dimension, is $50[\mu m]$. According to these values at least 17 images are needed to have part of the egg exactly on the reconstruction plane. However, the egg does not have to be reconstructed perfectly for it to be recognizable by eye or by a classifier. Hence, less



(a) Original video frame.



(b) FG mask extracted from the videoframe.

Figure 3-11: Example of the application of FG detection, resulting in a distinct FG masks[11]. This application makes it possible to detect moving fringe patterns while disregarding stationary ones. This prevents unnecessary reconstructions and classification of stationary fringe patterns.

reconstruction planes can be used to find a sharp enough image of the object. To verify this, a fringe pattern from an egg in a video-frame was reconstructed over 3,5, and 9 propagation distances (Reconstruction distance pitches of $400, 200, 100[\mu m]$). The sharpest and second sharpest reconstructions were selected and are shown in Figures 3-15a until 3-15f.

The figures show that having more than 5 reconstruction planes over the channel depth of $800[\mu m]$ does not show significant improvement. Therefore, the objects found will be reconstructed for 5 different distances (distance pitch $200[\mu m]$).

3-3 Object classification

So far the algorithm is able to locate particles in urine and reconstruct them locally. This section presents several classifier implementations which will tell whether these objects are SH eggs or not. Classifiers require training data as well as validation data, the collection of these is discussed first. This is followed by the feature generation implementations. The classifiers are then described and evaluated on their performance, concluding with a comparison.

3-3-1 Data acquisition

The initial training dataset consists of approximately 1600 manually labeled reconstructions, which are then rotated $90, 180, 270$ degrees and mirrored. This resulted in a total dataset size of approximately 12800 labeled images. This total dataset was imbalanced because there were more images labeled as eggs than not eggs, this was accounted for by discarding the surplus of images after shuffling the dataset. This resulted in a total of 10000 labeled training images, 90% of which is used to train the models and the remaining 10% is used to verify model accuracy.

The models will be tested with an independent dataset after training in order to test if the model is generalized. A dataset consisting of 1000 labeled images from a different recorded

(a) closeup of FG mask

(b) FG mask after morphological dilation.

Figure 3-12: Example of the application of morphological dilation to a FG mask. the fringes that appeared speckled are connected after dilation.

sample was created for this purpose. This set will be used to test the model accuracy after verification has been done, and will provide a more accurate representation of the model accuracy since it is completely independent of the training dataset.

3-3-2 Feature generation

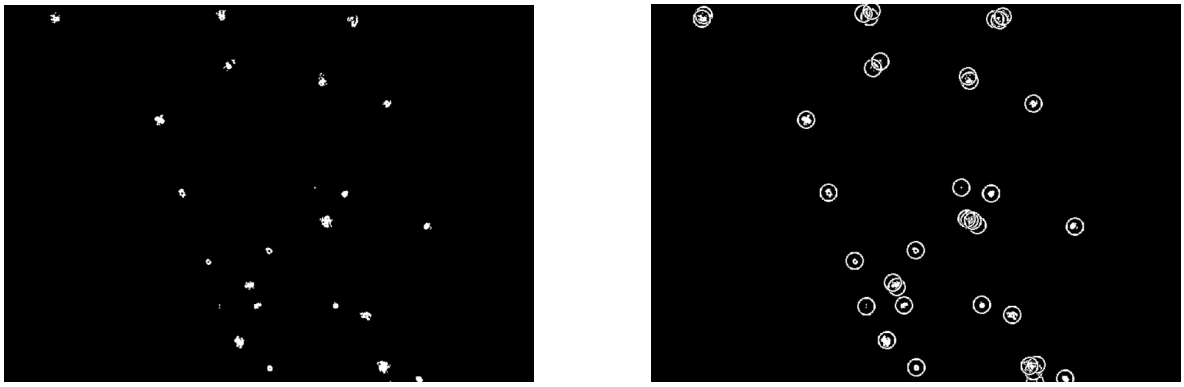
The reconstructed images will contain an object surrounded by a transparent background as shown in Figure 3-16. The Convolutional Neural Network (CNN) and Support Vector Machine (SVM) can both use images like this as input, however Random Forest Classifier (RFC) needs some input processing which was described in Section 2-3. The paragraphs below describe the parameters used during this processing, the SVM and RFC use these processed images as input.

Principal Component Analysis (PCA)

The images are resized to 50×50 pixels for the PCA. Covariance matrix matrix was determined for the complete set of images. The singular vectors in U resulting from the Singular Value Decomposition (SVD) of this covariance matrix then spanned the variance of the images. The variance accounted for by the sets of singular vectors is shown in Figure 2-5. Figure 3-17b shows that the image reconstructed from 120 principal components is similar to the original image.

Co-occurrence matrix

Similar to the principal component input features, the images were resized to 50×50 pixels. Next, the pixel values were scaled as integers between $0 - 9$. The co-occurence matrix was then calculated according to the procedure described earlier in section 2-3.



(a) FG mask after morphological dilation.

(b) Objects found by blob detector are encircled.

Figure 3-13: Example of the application of blob detection on a processed image. The objects that were found are encircled, this figure shows that not a single white spot has been overlooked.

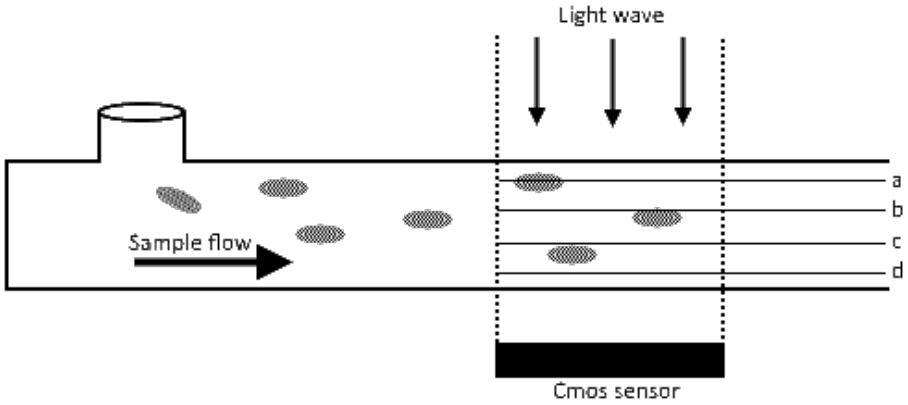


Figure 3-14: Depiction of way eggs appear at different depths in the flow channel. Note that each reconstruction plane (a, b, c, d) only overlaps with one egg each.

3-3-3 SVM

The SVM was implemented with pixel values, principal components, and the co-occurrence matrix as inputs. The classifier was optimized separately for the input modes, the model parameters can be found in Table 3-2 and the resulting performance is presented in Table 3-5.

3-3-4 RFC

The RFC was implemented with principal components and the co-occurrence matrix as inputs. The classifier was optimized for both inputs separately, the model parameters can be found in Table 3-3 and the resulting performance is presented in Table 3-5.

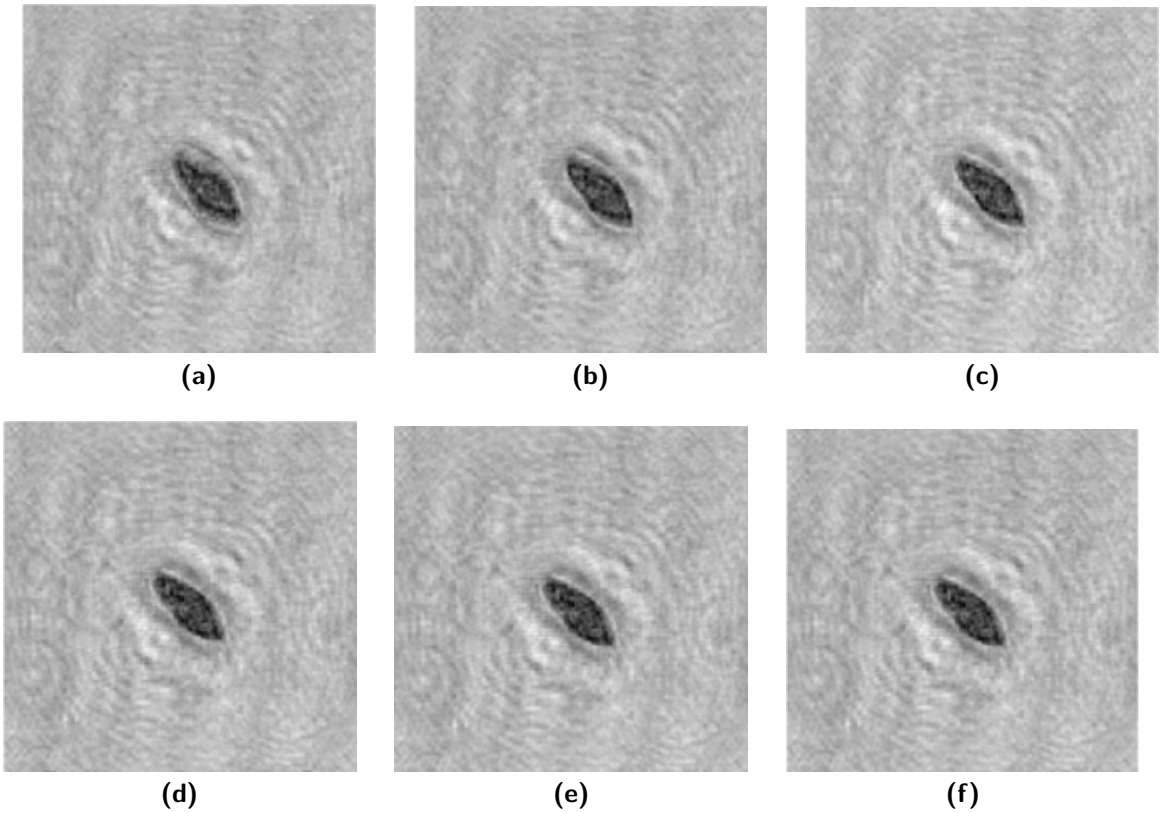


Figure 3-15: Sharpest egg reconstructions at reconstruction distance pitches $400[\mu m](a, d)$, $200[\mu m](b, e)$, and $100[\mu m](c, f)$ with d, e, and f being the sharpest. This shows that the eggs do not have to overlap directly with the reconstruction planes to result in reconstructions.

3-3-5 CNN

The accuracy of a CNN depends on various factors such as the architecture, learn rate, and the amount of epochs. For simplicity the learn rate and architecture are kept constant when evaluating the methods. The CNNs will be trained for 3 and 10 epochs and have their performance on the training and validation set evaluated. The best model will be selected for further testing. Table 3-4 shows the result of this, and presents the test and validation accuracy.

The table shows that the model based on Y. Ghouzam's model performs the best of the three tested models, reaching a validation accuracy of 96% after only 10 epochs. Another feature that stands out is that Ghouzam's model does not have a disproportionate difference between false-positives and false-negatives[48]. Ghouzam's model will be perturbed by changing layers in order to see whether the accuracy can be improved.

The layers have been perturbed by changing kernel sizes, dropout rates, pool sizes, layer densities, and dropping out complete layers. However, with no more than 6 training epochs the accuracy does not seem to increase.

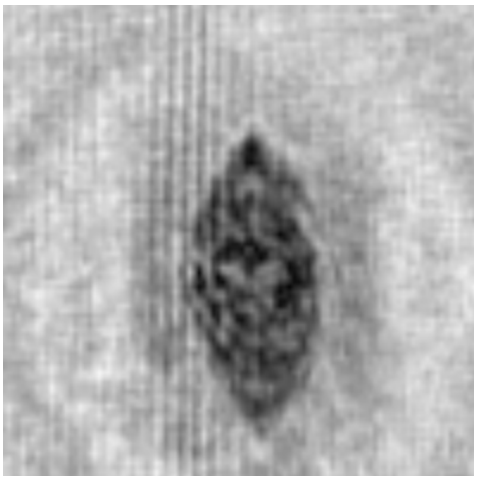


Figure 3-16: A randomly picked egg from the training dataset. This is $100 \times 100[pixels]$ and will be rotated and mirrored in order to augment the dataset. The resolution will be reduced for some classifiers in order to speed up training and require less training data.

Table 3-2: Model parameters of the various SVM with different inputs.

Model input	C	γ
Pixel values	10^3	10^{-6}
Principal components	10	10^{-6}
Co-occurrence matrix	100	10^{-4}

3-3-6 Comparison

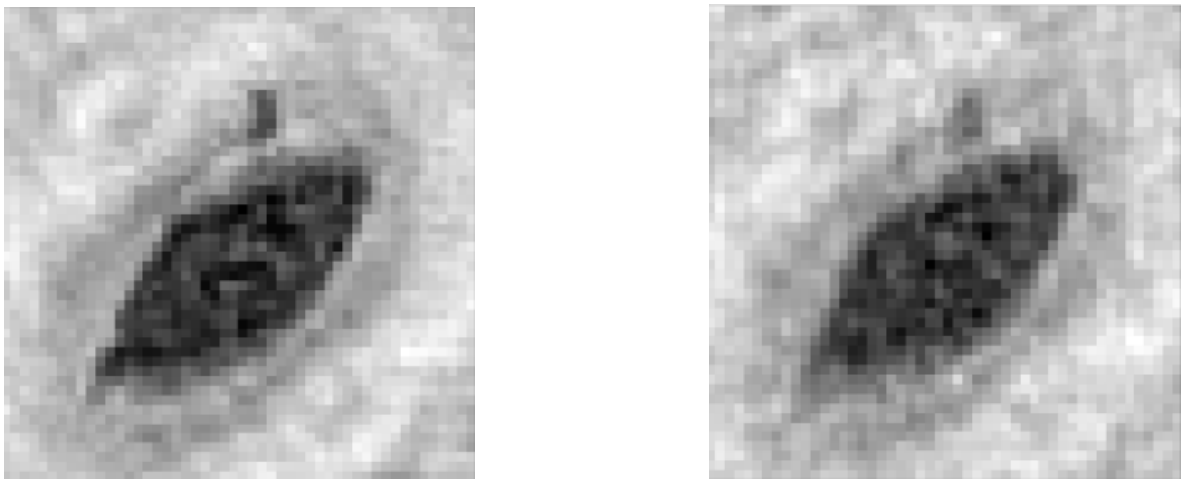
So far this section has presented a range of classifiers and feature descriptors. This subsection shows the performance of the classifiers on the dataset of SH eggs. The best models were determined by varying the hyperaparameters and model structures, with their results shown in Table 3-5.

The table shows that the RFC with the co-occurence matrix as input features has the best performance on the validation and test set. As will be shown in the next section, the RFC performs the worst when it comes to computational time and is therefore not suited for this diagnostic method.

The CNN models perform similar to the SVMs on validation accuracy but have better test accuracy, which is a more reliable indicator of a model's performance[43]. The most accurate classifier when ruling out RFC is the CNN based on Ghouzam's model[48], which is then the most suited classifier for further testing. The model shall be retrained in the next Chapter with more representative training data from the prototype.

3-4 Algorithm benchmark

Finally, the complete algorithm shall be evaluated on its computational efficiency. Table 3-6 shows the environment on which the algorithms were designed and tested.



(a) **(b)**

Figure 3-17: Showcase of the effect of reduction of dimensionality. The original image **(a)** was converted to 120 components, the image was then reconstructed from these values **(b)**.

Table 3-3: The amount of trees used for the two RFC.

Model input	Decision trees
Principal components	20
Co-occurrence matrix	10

3-4-1 Results

Figures 3-18a and 3-18b presents the computational time of reconstruction and classification for various window sizes in order to give an impression of the magnitude of computational time. Simply reconstructing whole images will take too long so the reconstruction and classification process will have to be performed in a different way. To give an example, a video for diagnosis generally contains 616 images of 3840×2748 pixels. Extrapolating from Figure 3-18a, it will take approximately 4.4 seconds to reconstruct such an image with the specified setup. At least 5 reconstructions are needed for each captured image in order to have a sharp view of every object present in the flow cell, so reconstruction will take 22.5 seconds for each frame. This means it will take $616 \times 22.5 =$ **13860 seconds**, or **3.9 hours**, just to get a clear view of one sample using a desktop computer. This is without actually finding the objects and classifying them.

Image resolution reduction

The input of the object detection algorithm is a shrank version of the recorded frame. The effectiveness of this was tested by running the algorithm with various amounts of window shrinking. The results of these tests are shown in Table 3-7 which shows its effect on computational time as well as the amount of objects found.

The table shows that there is a significant time-save when reducing the resolution, but the amount of objects found reduces as well. This reduction of found objects could either be caused

Table 3-4: CNN classifier model verification and validation accuracies for several trained models with different amounts of epochs and shuffled training dataset.

Model	Model accuracy	
	Verification	Validation
Yalcin (3 epochs)	0.9329	0.9115
Yalcin (10 epochs)	0.9454	0.9300
Ghouzam(3 epochs)	0.8465	0.8236
Ghouzam (10 epochs)	0.9416	0.9645
Keras (3 epochs)	0.8584	0.9339
Keras (10 epochs)	0.9611	0.9136

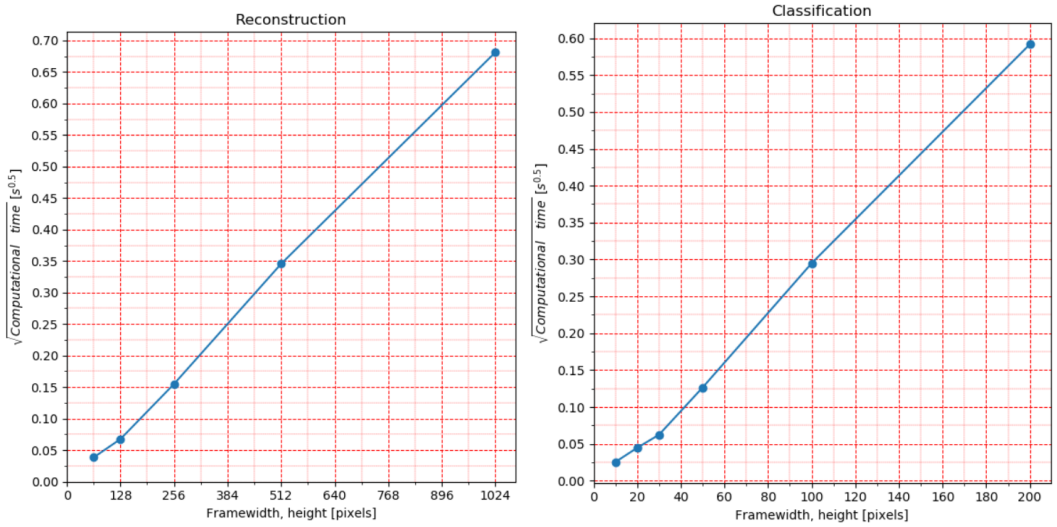
Table 3-5: Classifier performances on the prepared dataset. Each of the CNN models were trained for 10 epochs, while the other classifiers were trained with different amounts of data in order to ensure the models did not overfit.

Classifier	Model accuracy	
	Validation	Test
Yalcin CNN	0.9454	0.930
Ghouzam CNN	0.9416	0.965
Keras CNN	0.9611	0.914
SVM pixel values	0.948	0.92
SVM PCA	0.930	0.902
SVM co-occurence	0.944	0.943
RFC PCA	0.965	0.835
RFC co-occurence	0.998	0.973

by the algorithm missing detections because of the size reduction, or the algorithm finding a single object multiple times because of the fringe patterns being detected as multiple objects. The detected objects were encircled in the video to see which case is the most prominent. The video was analyzed, and the latter case seems to be happening: high resolution fringe patterns cause multiple detections of a single object. Figures 3-19a-3-19d shows an area of a snapshot in the video which is representative for explaining the variation in object counts. It can be seen that the same objects are detected, but at high resolution the objects are detected multiple times. Judging from the images, the resolution reduction of 16× has some double counts but no missed detections.

Object detection

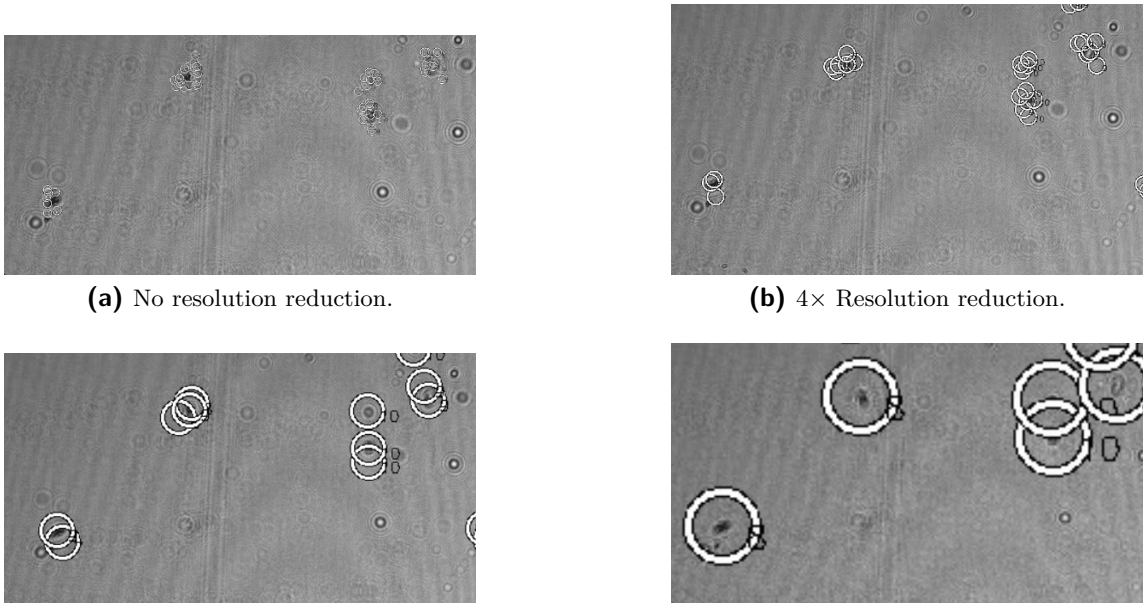
Table 3-8 shows the computational time of the algorithm designed in this report compared with a sliding window algorithm. Both perform similar tasks: they generate images which can be fed into the reconstruction algorithm and then into the classifier. The method proposed in this report takes longer to complete for each frame, but produces less frames for reconstruction and classification. The time it takes to find the objects in a video of 625 frames is about 27 seconds for the method described in this report, while it takes the sliding window only 0.5 seconds. On the other hand, the images fed into the reconstruction and classification algorithms are respectively $0 - 10000$ compared to > 300000. Figures 3-18a and 3-18b show



(a) Square root of reconstruction time over input image dimension.

(b) Square root of classification time over input image dimension.

Figure 3-18: Two plots showing the square root of the computational time of two essential operations that have to be performed when counting eggs. The square root of the computational time has been used as y-scale in order to show that the relation between computational time and amount of input pixels is approximately linear.



(a) No resolution reduction.

(b) 4× Resolution reduction.

(c) 16×Resolution reduction.

(d) 64× Resolution reduction.

Figure 3-19: Showcase of the effect of reducing the image size before determining the object locations. Having a high resolution seems to result in more double counts. The objects found in every frame can all be distinguished in the image that was reduced 64×, showing that the moving objects can be detected even after significant resolution reduction.

Table 3-6: Specs of the desktop computer and design environment used for the initial design and bench-marking of the algorithm

Component	Product	Specifications
Operating system	Windows 10 pro	
Motherboard	MSI B350 PC Mate	
CPU	AMD ryzen 3 1200	AM4, 3.1GHz
GPU	GeForce GTX 1050 Ti	4GB, GDDR5, 7.008GHz
RAM	Crucial Ballistix Sport LT	8GB DDR4
Python IDE	Pycharm Community edition	
Python version	3.7	All packages up to date as of 16-5-2019

Table 3-7: Time taken by the object detection algorithm for various amounts of window shrinking as mentioned in the previous section. There seems to be a disparity in the objects found between the various resolution reductions. Note that an actual sample will contain less eggs than the sample used while making these recordings.

size reduction	Computational time for 10 frames[ms]	Objects found
64	115.6	199
16	518.5	303
4	2239	713
1(No reduction)	7666	1988

that reconstruction and classification of a $256 \times 256[pixel]$ image takes respectively 0.0225 and 0.09 seconds. For this particular analysis it is assumed that only one reconstruction and classification is needed for every window. This results in the sliding window taking at least $300000 \times 0.60 = 180000[s]$ compared to an estimated analysis time of $500 \times 0.6 + 27 = 327[s]$ for object detection(assuming 5 reconsructions and classifications per window/object). This approximation shows that using an object detection algorithm before reconstruction and classification can cut computational time by several orders of magnitude.

Table 3-8: Computational time taken per frame by a sliding window(256×256 pixel resolution, 128 pixel stride) and by the foreground detector + blob detector. The implemented algorithm takes about 50 times longer for generating the object images than it takes to move a sliding window over each image. The flip-side is that only a fraction of the images are generated for analysis.

Operation/Algorithm	Computational time [ms]	results
Foreground detection + blob detection	43	$0 - 100$ images to analyze
Sliding window 256×256 pixels	0.86	> 500 images to analyze

Classification

Table 3-9 shows the computational time of reconstructing and classifying 100 objects as well as the computational time of a single classification. The computational time of the CNN and SVM implementation are of similar orders of magnitude, while the RFCs take significantly longer. The RFC will take more than one second if it has to classify 5 frames for every object,

which is too long. A urine sample from a person with a heavy infection load (> 50 eggs) can take more than 1 minute just to reconstruct the local images and classify them. Diagnostic time is crucial when doing mass surveys, and so the RFC, although being the most accurate, is not suitable for implementation in the actual algorithm. It is for this reason that a CNN was selected in the previous section.

Table 3-9: The time taken by reconstruction($256 \times 256[pixels]$) and classification using 50×50 pixel windows for the classifier/feature input for 100 objects, 5 reconstructions and classification per object. A total of 500 reconstructions are performed during each simulation which takes approximately $11.5[s]$.

Classifier	Total computational time [s]	Mean time per classification [ms]
CNN (Ghouzam)	30.19	37.4
SVM pixel values	30.43	37.9
SVM PCA	23.35	23.7
SVM co-occurence	25.47	27.9
RFC PCA	132.4	242
RFC co-occurence	132.8	243

3-5 Discussion

This chapter presented the implementation and initial performance evaluation of a new diagnostic method. each step has been verified to be working but the exact performance remains unknown, this will be tested in the next chapter. This section discusses the considerations made when designing the algorithm and concludes with a brief reflection on the chapter.

3-5-1 Other algorithm considerations

The algorithm presented in this chapter is a result of iterative procedure where various algorithms and combinations were tested. The end result was considered to be the most efficient in computational time while accurately finding the objects and classifying them. Below are some of the the ideas which were tested but were found to hinder the performance of the diagnostic method.

Phase reconstruction

Holographic reconstruction results in an image with complex values, the phase of this complex number represents the thickness and material of the object the light is passed through while the magnitude represents the impedance of the object. The classifier uses the reconstruction magnitude as input since this the closest resemblance to how the eggs are viewed through a microscope by the human eye. Using a phase image was considered, however this was not applied since accurate phase Retrieval requires iterative reconstruction procedures .[57]

Automatic focusing of the reconstructions

Right now the algorithm produces 5 reconstructions for every found object, all of which are then fed to the classifier. A first order difference focusing algorithm was tested to see if it was possible to automatically find a sharp reconstruction. The focusing algorithm needed at least 15 reconstructions in order to work for some objects, while it failed to find a sharp reconstruction for other objects. An image focusing algorithm requires more reconstructions and takes more than double the computational time when it is performed. Therefore the use of such algorithms them was omitted.

Continuous recording of the flow cell

The diagnostic method records the content of the flow cell step wise: the syringe is pressed and stopped, an image is then recorded. Continuous recording of the flow while taking images was attempted however this resulted in blurry images. Therefore the step wise recording method was used.

Classification of fringe patterns

The classifier can also so have the fringe pattern as input which would circumvent reconstruction. This would increase the computational speed of the algorithm, however; The World Health Organisation (WHO) requires the eggs to be identified by eye and therefore at this stage reconstruction cannot be omitted. If this were to be implemented, reconstruction should still need to be used for labeling the training data. The training data will then consist of the fringe patterns instead of reconstructions. This will not be considered as it is out of the scope of this preliminary design, but should be considered in the future.[2]

3-5-2 Reflection

The diagnostic method is able to record the sample, find particles of interest, reconstruct them locally, and classify them as egg or not egg. The algorithm was evaluated on its computational speed and not on its capability to find eggs in a sample. Although it performed the latter task adequately judging from the images. The exact amount of eggs present in the sample was unknown and could therefore not be used to test the performance. This will be done in the next chapter where the algorithm will be implemented on a prototype device. This prototype is then subject to a field test where its performance is compared to that of an expert egg counter.

Chapter 4

Prototype validation study

The previous chapter was dedicated to verifying the imaging method and the egg detection algorithm. These are to be integrated into a single device which can quantitatively diagnose urinary Schistosomiasis. The capability to perform this task will be evaluated in this chapter. First Section 4-1 will elaborate on how the algorithm is integrated with the prototype device and present results of laboratory tests. The prototype is then brought to the field for testing where its performance is compared to that of an expert microscopist, which is discussed in section 4-2. Finally the Laboratory and in-field performance of the prototype is reviewed in Section 4-3.

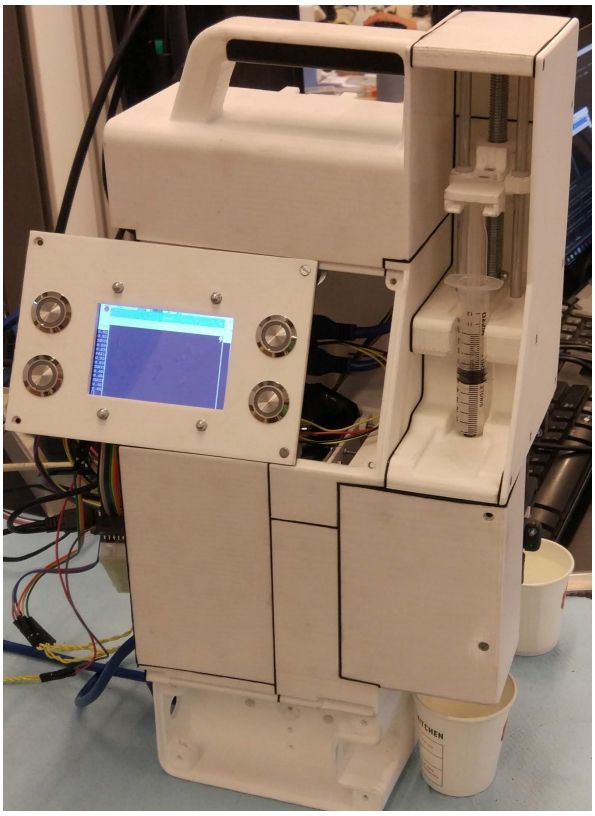


Figure 4-1: Working prototype assembly.[12]

(a) Rendering of the prototype. **(b)** The actual prototype.

Figure 4-2: Showcase of the prototype used for testing the prototype-algorithms interaction and performance[12]. Note the syringe holder, which is situated above the recording chamber.

4-1 Prototype integration

The algorithm was implemented on a prototype device which uses a Raspberry pi 3 Model B, Figures 4-2a and 4-2b show a rendering and picture of the prototype[12]. The sample is analyzed by consecutively moving the syringe with a stepper motor and recording a hologram, this method was developed with support of Section 3-2.

The holograms will be captured by the prototype but the images will not yet be analysed, images will be analyzed on the the computer setup mentioned in Chapter 3. The chamber in which the holograms are recorded is insulated from all outside light sources, the only illumination hitting the the sensor comes from the monochromatic illumination source within the chamber itself. An image captured by the prototype can be seen in Figure 4-3. Variation in environmental light did not show observable differences in the holograms. For more information on a prototype see M. Hoeboer's report.[12]

The prototype was tested with Schistosomiasis Haematobium (SH) eggs in saline solution in order to find the range of reconstruction distances. Figure 4-4b shows an egg that was reconstructed from a hologram capture with the prototype. The mean reconstruction distance was determined from reconstructing 100 eggs, and turned out to be 3.0$[mm]$.

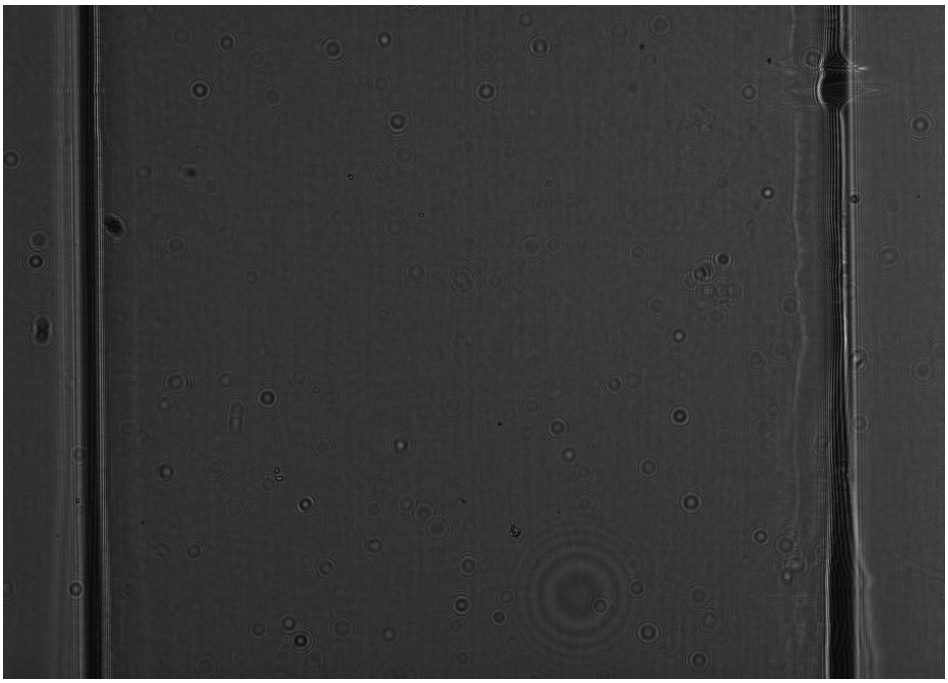
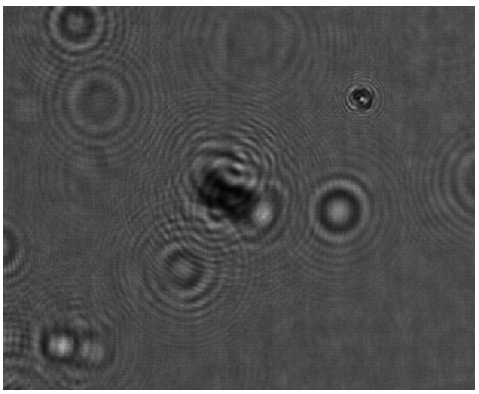
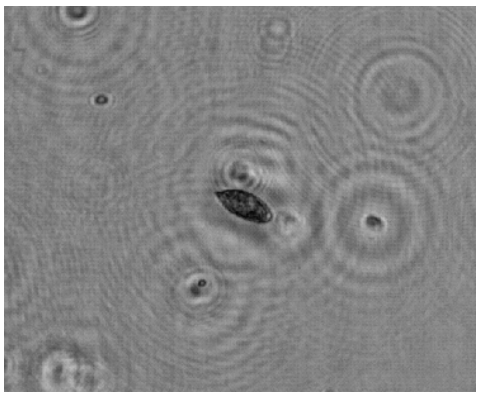


Figure 4-3: The flow cell as captured by the sensor within the imaging chamber of the prototype.



(a) Hologram of an egg captured by the prototype



(b) Reconstruction

Figure 4-4: Showcase of the reconstruction performance of the prototype. The reconstruction shows more detail than the verification tests shown in chapter 3.

4-1-1 Data acquisition

A sample containing approximately 12000 SH eggs in saline solution was used by a lab technician to create new samples with various approximate egg counts mixed with either saline solution or urine from an uninfected person. Some of these samples were used for creating a training and testing dataset for the classifier while the others were used for testing the algorithm. Table 4-1 shows the samples with their approximate amount of eggs, the fluid it was mixed with, and the purpose of the sample. Each sample measured $12[mL]$ which is equal

Table 4-1: Samples containing SH eggs. The samples will be used to train the classifier and analyse its performance in the lab.

ID	Eggs	Fluid	Purpose
1	200-300	Saline	Analysis
2	400	Saline	Analysis
3	400	Saline	Analysis
4	0	Urine	Analysis
5	200-300	Urine	Analysis
6	200-300	Urine	Analysis
7	400	Urine	Analysis
8	400	Urine	Testing
9	1000	Urine	Test data
10	1000	Urine	Training
11	2000	Urine	Training
12	2000	Urine	Training

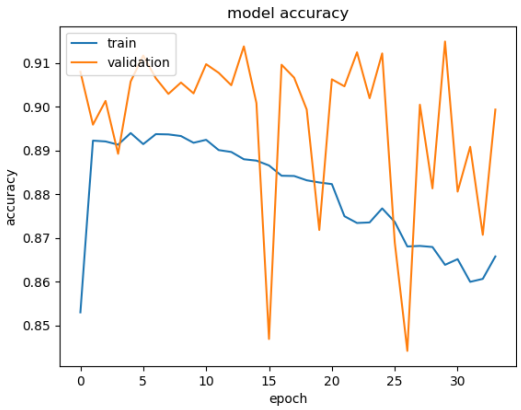
to the $10[mL]$ defined by the World Health Organisation (WHO) standard, the remaining accounts for dead volume within the device. Training and testing data was generated by using the prototype to capture holograms of the entire sample and running the algorithm. The reconstructions were examined by the author and labelled manually, instead of classifying automatically. This resulted in a training and test dataset of which the numbers are shown in Table 4-2. The data was shuffled and balanced in order to have a generalized dataset with equal amounts of eggs and not eggs.

Table 4-2: The amount of training and test data available to the Convolutional Neural Network (CNN). The training data will further be split into training and validation data.

Dataset	Eggs	Not eggs
Training	12087	12208
Testing	352	9016

4-1-2 Classifier evaluation

Ghouzam's CNN model was trained on the dataset with a training-validation split of 2:1, the training results are shown in Figures 4-5a and 4-5b. The model was then tested on the test data whose results are shown in Table 4-3. The training curves show that the maximum accuracy is reached within 15 epochs and that there are fluctuations in the validation accuracy. This is an indicator of a lack of generalization. This can be explained by the fact that there is not enough data available to represent the variation in egg transparency, size, shape, orientation, and depth. This is further proven by Table 4-3 which shows a sensitivity of 0.506, specificity of 0.986, and overall accuracy of $0,968$ has been achieved on the independent data set. The purpose of the classifier is to provide an accurate eggs count, for which it needs both a high sensitivity and specificity. A data set is needed with more representative examples of eggs in this setting in order to achieve this goal, however this data is not available.



(a) Classifier training and validation accuracy.

(b) Classifier training and validation loss.

Figure 4-5: Accuracy and loss charts during training of the CNN. The amount of epochs should be selected such that the loss is minimal and the training and validation accuracy like closely together.

Table 4-3: Confusion matrix of the classified test data, from an independent sample. The test has a sensitivity of 0.506, specificity of 0.986, and overall accuracy of $0,968$ on this data set.

	Labeled		
CNN	egg	not egg	total
egg	178	126	304
not egg	174	8890	9064
total	352	9016	9368

4-1-3 Lab test processing

The samples mentioned in Table 4-1 were analyzed with the algorithm in order to test the computational time for a complete diagnosis. The results are shown in Table 4-4, as expected the egg counts do not correspond to the estimated values. The computational times are significantly lower than the alternative methods mentioned in Section 3-4, but not low enough to compete with human experts[2]. During sample analysis the reconstruction and classification time seemed to increase significantly over the duration. For example, when analyzing sample 6(Table 4-4) the computational time for reconstructing and classifying 5 frames increased from $0.127[s], 0.167[s]$ to $0.211[s], 0.761[s]$ after analyzing 240 images. Efforts to keep these times constant were unsuccessful. This can be explained by the memory handling of Python, and can be prevented by implementing the algorithm on embedded code such as c++. However, this is out of the scope of this research. The object detector consisting of the Foreground (FG) detector and blob detector was assessed as well. The recorded hologram and the FG mask were put side by side with the found objects encircled in both images as shown in Figures 4-6a and 4-6b. The algorithm is able to find minor changes in the image due to the fringes caused by moving objects. The first 100 holograms of sample 1 mentioned in Table 4-4 were analyzed by eye for missed detections. No notable objects were missed,

Table 4-4: The results of analyzing the samples recorded in the lab. The amount of eggs found by the algorithm is lower than the approximate amount of eggs present in each sample.

ID	Eggs	Eggs found	Not eggs found	Time taken $[s]$
1	200-300	3	371	387
2	400	3	487	418
3	400	0	606	660
4	0	0	1276	1542
5	200-300	0	7255	9821
6	200-300	1	7202	10776
7	400	1	4616	5210

and the object detector was able to outperform the lab technician. The reconstructions were observed manually after they were found by the object detector. Eggs can appear clear or dark depending on their age, so this was done in order to verify that both of these are detected properly. Examples of the found eggs are shown in Figures 4-7a and 4-7b, the fringe patterns corresponding to these eggs are all found by the object detector. Therefore the discrepancy in egg counts can be explained by the shortcomings of the classifier mentioned before, and the increased flow-rate along the center of the flow-cell. The latter causes eggs to flow past the sensor without being captured. These discrepancies shall be discussed further in Section 4-3.

4-2 Field test processing

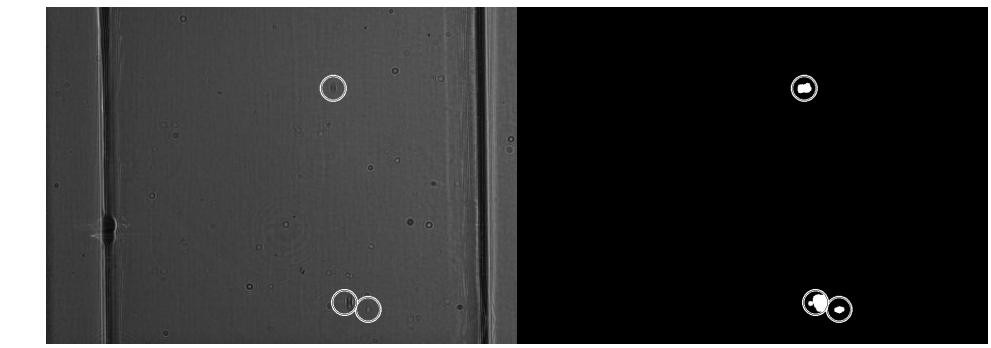
The prototype was taken to the field in Azaguié, Ivory Coast. In this village the local population was screened for SH by using the prototype. One of the goals of this field test was to see how the diagnostic algorithm would perform on urine from infected individuals in the relevant environment. First the data collection method shall be discussed, after which the most notable results are presented.

4-2-1 Data collection method

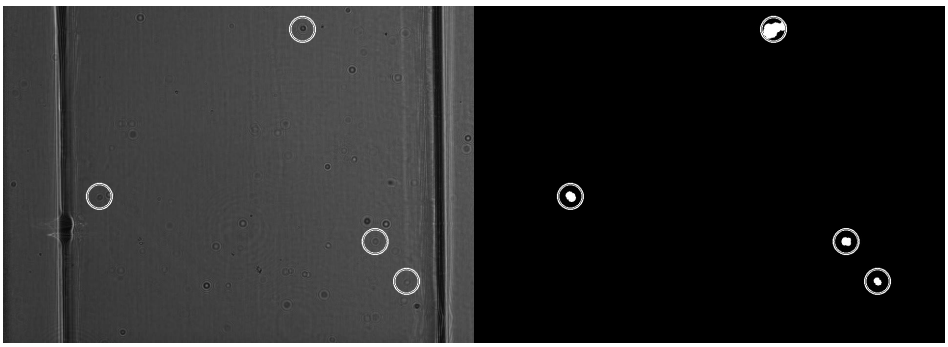
A 12$[mL]$ urine sample of each person was collected and subsequently drawn in a syringe which had to be connected to the prototype afterwards. The urine was caught in a cup after analysis, after which it was examined by a professional microscopist following the WHO standard. [12][2]A total of 95 measurements were completed by the prototype which was analyzed by the algorithm afterwards. The result of the diagnosis performed by the microscopist can be found in Appendix 5-2-5. Some of these results will be discussed in the paragraph below.

4-2-2 Results

The previous section already indicated that reconstruction and classification takes too much time, this became more apparent when analyzing an infected person's urine. The most heavily infected individual had 173 eggs in their urine. The corresponding images recorded by the prototype contained significantly more moving particles compared to the recording in the lab,



(a)



(b)

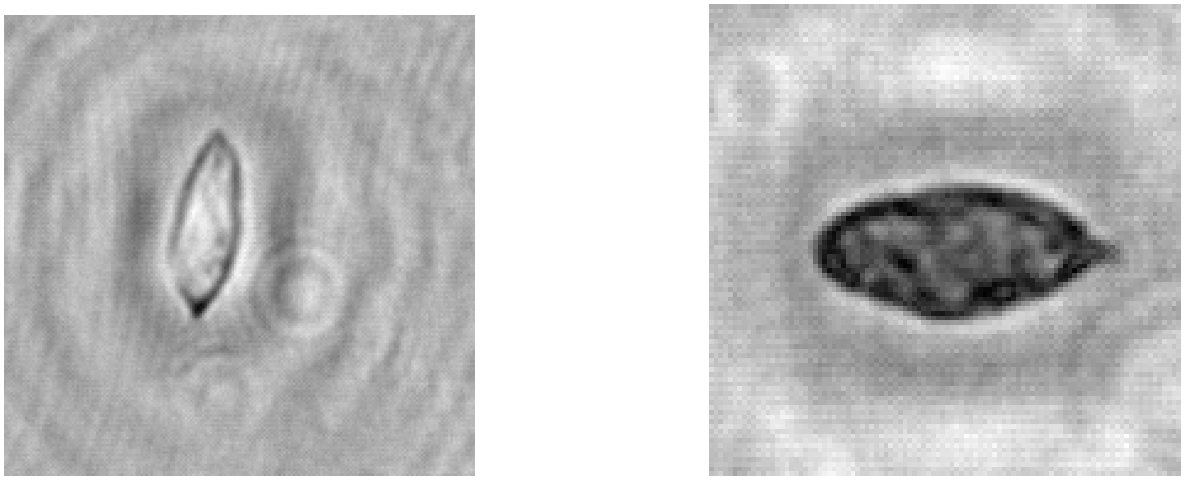
Figure 4-6: Side-by-side comparison of the recorded hologram and the detected foreground of a sample prepared in the lab. The locations of moving objects are indicated by circles.

this can be observed in Figures 4-9a and 4-9b. In order to prevent the same memory related issues as stated earlier only the first 100 recorded images were analyzed, the result of this is presented in Table 4-5. Only $\frac{1}{6}$th of the sample was analyzed which took over 50 minutes, and no eggs were found by the classifier even though the person was heavily infected according to the WHO standard. The presence of eggs in the recorded holograms was confirmed by manual inspection, however these eggs are classified as not eggs by the CNN. One of the recorded holograms in the field is shown in Figure 4-8a which happens to contain an egg, this egg was found by the object detector and is shown in Figure 4-8b however this egg was classified as not being an egg. This means that the eggs come in view of the hologram, are found by the object detector, but are misclassified resulting in an incorrect egg count.

Table 4-5: The result of analyzing the first 100 images of sample 159. This was a person who was considered to be heavily infected, yet zero eggs were found by the algorithm.

ID	Eggs	Eggs found	Not eggs found	Time taken[s]
159	173	0	4297	3232

The reconstruction and classification parts of the algorithm were turned off in order to see how the object detection part performs. The sample containing 173 eggs along with samples containing 105, 46, and 0 eggs were analyzed completely with this reduced algorithm. This resulted in an 'object count' and 'measurement time taken' as shown in Table 4-6. The



(a) **(b)**

Figure 4-7: Examples of the variation of transparency of the eggs. The transparent and not transparent eggs are detected by the algorithm.

algorithm finds more objects in the samples of infected people compared to the samples of the lab and the uninfected individual. This affects the recorded hologram and the amount of objects found as shown in Figures 4-9a and 4-9b. However, no conclusions regarding the infection with SH can be drawn from this observation since the sample size is too small. Significant time difference can be observed between analysing a sample with and without

Table 4-6: The result of analyzing 4 samples for the amount of objects found and the time taken.

ID	Eggs	Objects found	Time taken [s]
108	46	13390	186
110	0	465	182
114	105	19272	182
159	173	21641	159

object reconstruction and classification. The algorithm is able to locate all the objects in the recorded images in approximately 3 minutes, even when the sample is contaminated (See Figure 4-9a and 4-9b compared to Figure 4-6a and 4-6b).

4-3 Performance reflection

The diagnostic method was tested in the lab and in the field. Several observations were done regarding the performance, these will be discussed in this section.

4-3-1 Performance measures

The new diagnostic method which was evaluated in this chapter should also be assessed by the performance measures presented in Section 2-4. Table 4-4 show the result of analyzing

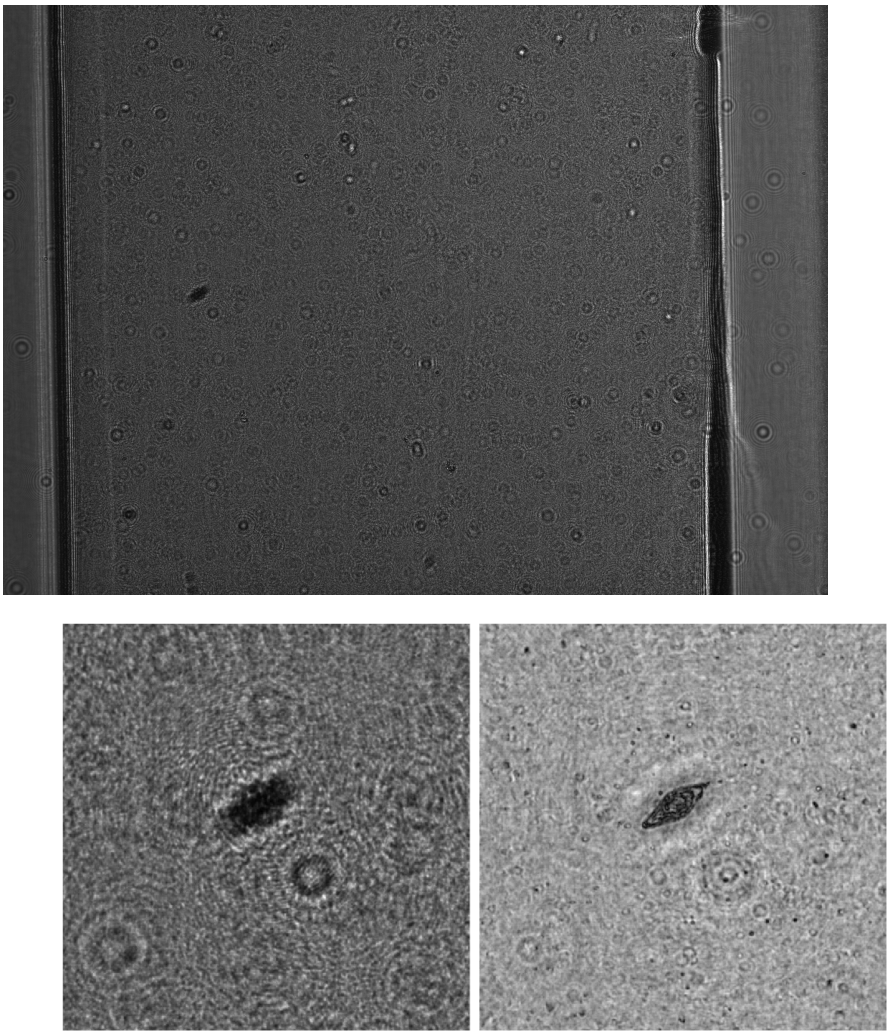
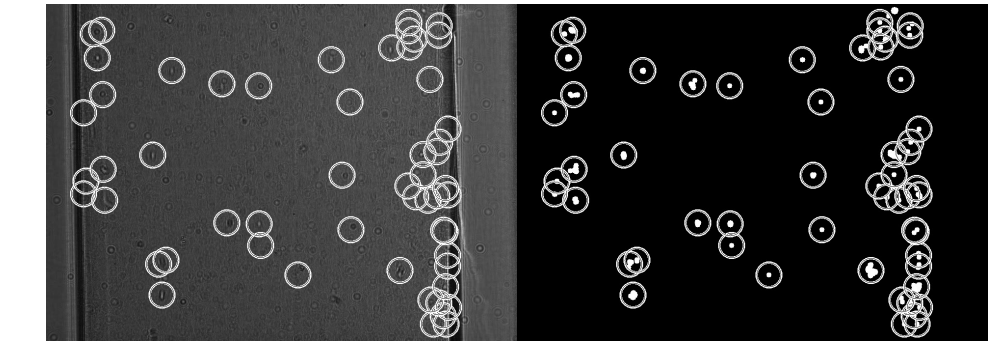


Figure 4-8: Hologram containing an egg, recorded during a field test(a), local fringe pattern and corresponding reconstruction of the egg (b).

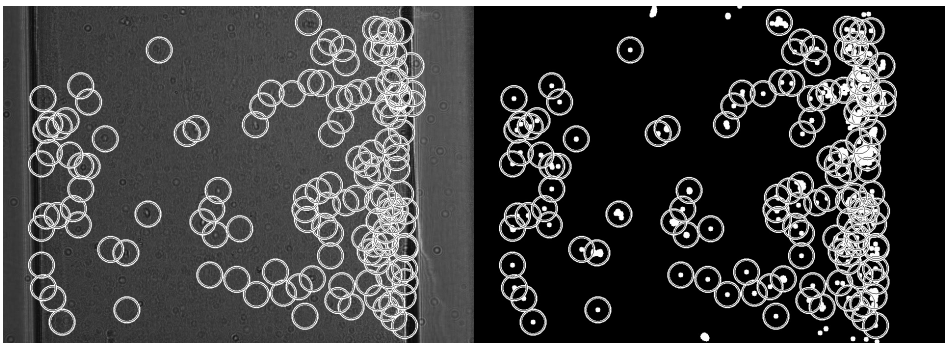
7 samples, the confusion matrix shown in Table 4-7 summarizes the outcome of the test. With these results the diagnostic method has a sensitivity of 66.7% and a specificity of 100%, however; this observation does not have value because of the following reasons:

- Low volume of tested samples ($n = 7$).
- Discrepancy between the amount of eggs found by the algorithm compared to the estimated amount present in the sample.
- Low classifier performance measures result in an unreliable egg count (SN 0.506, SP 0.986, acc $0,968$).

Therefore this new diagnostic method does not yet have sufficient performance in order to replace or supplement the current diagnostic standard.



(a)



(b)

Figure 4-9: Side-by-side comparison of the Recorded hologram and the detected foreground of a sample from an infected person. The locations of moving objects are indicated by circles.

4-3-2 Sample recording method

The combination of digital holography and flow cytometry provided a method for lensless imaging of SH eggs without sample preparation. The terminal spine which is characteristic of the eggs is visible in the reconstructions of their respective fringe patterns. A lot of data is generated for one diagnosis, this does not pose a problem with the right computational equipment and memory management.

4-3-3 Computational time

There is a memory leak which causes a buildup of memory usage, resulting in some computational steps taking longer as a diagnosis progresses. The diagnostic time ended up being significantly longer than expected. In Section 4-1 it was indicated that classifying 5 reconstructions at the start of sample analysis takes 0.167 seconds, which is 0.033 seconds per reconstructed image. If reconstruction is omitted and the fringe patterns of the found objects are classified directly, it will take up to $0.033 \cdot 21641 + 180 = 894[s]$ to analyze a sample, which is an order of magnitude less than the diagnostic time indicated in tables 4-5 and 4-4. This indicates that the object reconstruction and classification part of the algorithm is the limiting factor of diagnostic time. Additionally the analysis was run on the desktop computer, not the Raspberry Pi 3 that is on board of the prototype. The analysis would take longer than

Table 4-7: Confusion matrix of the results from the lab tests

	Truth		
New method	positive	negative	total
positive	4	0	4
negative	2	1	3
total	6	1	7

indicated in this chapter if it were run on the prototype.

4-3-4 Object detection

During the lab tests the object detection system was able to find and reconstruct every notable moving object in all of the holograms. In some of the field test samples there were orders of magnitude more objects present in the holograms compared to the lab tests, which gave the foreground detector little information to learn the difference between a background and a foreground pixel. This indicates that the the foreground detector should be tuned differently depending on the quality of the sample.

4-3-5 Missed egg detections

The fluid dynamics were neglected when determining how much volume is pressed by the syringe between each capture. The fluid experiences friction near the boundaries, it moves slower the nearer it is to these boundaries(see Figure 4-10). The fluid has a higher velocity when moving away from the boundary causing some eggs to pass by the sensor without being in view during capture of the hologram. On the other hand eggs moving near the boundary are observed over multiple different frames as they move slower. Whether this effect balances out and results in a accurate account has not been researched yet as it requires a boundary layer analysis [13].

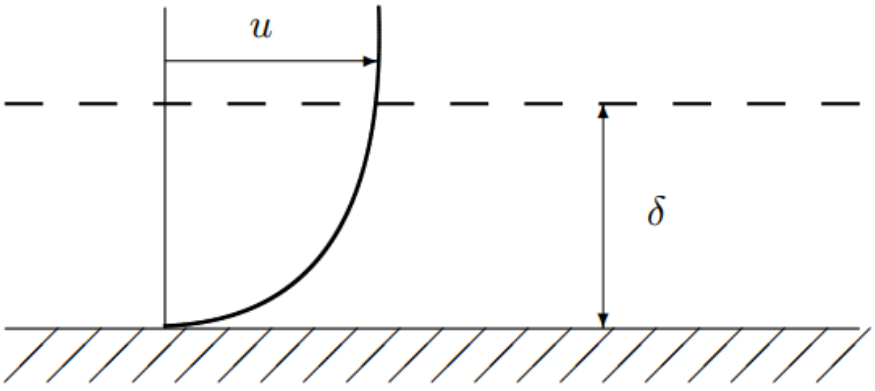


Figure 4-10: Illustration of the boundary layer caused by friction as a fluid moves over a stationary surface, u represents the local flow velocity and δ the distance from the surface.[13]

4-3-6 Non generalized classifier

The training dataset of the classifier was obtained from a few samples which contained many eggs. As such it did not contain information on how the eggs appear in actual infected people. Although many eggs were used, the sample size was not yet big enough to represent the variation in egg size, shape, orientation, and transparency. Additionally, there appear moe particles around the eggs in contaminated urine (see Figure 4-8b), these particles are not present in the training data set. For this reason many eggs that were found by the object detector were classified as not eggs during field and lab tests.

The diagnostic algorithm is not yet able to fulfill the research objective. However the shortcomings mentioned above could be circumvented by collecting a larger, generalized data set and by implementing the algorithm on embedded code with proper memory management.

Chapter 5

Conclusion

This chapter will conclude the report. The first Section will briefly summarize the topics that were discussed which is followed by a final conclusion on. Finally recommendations for future research regarding the discussed topics will be mentioned in section 5-2.

5-1 Reflection

This report presented the preliminary development of a low-cost smart diagnostic method for Schistosomiasis Haematobium (SH), focusing on improving and replacing the current World Health Organisation (WHO) standard for diagnosis. Chapter 1 looked at the current state of of diagnostics of SH. The WHO diagnostic standard was described and its shortcomings were identified. Additionally, the latest research regarding the development of smart diagnostic methods was looked into. Finally the research objective was identified:

> *"Develop a low-cost, smart diagnostic method for SH based on detecting eggs in urine by combining lensless imaging and flow cytometry, and developing Artificial Intelligence (AI) models for automated detection."*

Chapter 2 then followed by presenting the background information needed to implement this new diagnostic method. In-line planar wavefront digital holography was identified as the most suitable lensless imaging method. A sample will be analyzed by the following repetitive procedure: mechanically press the piston of a syringe by a small volume, wait for the flow to stop, record a hologram, detect eggs. The proposed egg detection procedure consisted of a series of image processing algorithms:

- Apply Foreground (FG) detection in order to label moving objects in the recorded holograms.
- Localize the objects with a Blob detector.

- Locally reconstruct the hologram at the found locations.
- Classify the reconstruction as egg or not egg.

Finally the diagnosis can be provided as a confidence interval of the account depending on the classifier performance.

These methods were verified in Chapter 3. The Digital Holographic Microscope (DHM) proved to have sufficient resolution for imaging the eggs in a flow cell. The egg detection algorithm was then implemented on a desktop computer. Finally diagnostic algorithm was bench-marked to ensure that the diagnostic time was within reasonable limits.

Chapter 4 served as the final test for the algorithm. The detection algorithm was implemented on a prototype device which automatically analyzed the urine. This device was first tested in a lab with cultivated egg samples. After this, the prototype was brought to Ivory Coast for a field test where its performance was compared to an expert microscopist.

The sample recording method has been shown to be an effective alternative to the current diagnostic standard, the DHM was able to accurately reconstruct eggs within the flow cell. An entire sample could be imaged automatically with the only sample preparation being: drawing the sample with a syringe. No filtration, staining, or human interaction was needed besides this. Additionally, The object detection architecture was able to efficiently find all the microscopic objects within the sample. Even objects much smaller than the eggs were found to be present in urine. A basic classifier implementation showed promising results with an accuracy of 0.96.

Testing the algorithm on a prototype further validated the imaging method as eggs were found and reconstructed accurately. On the other hand, the lab tests showed that the data set of the classifier did not contain enough images to train a generalized model and that the local reconstruction and classification takes increasingly more time during analysis. The field tests further validated this observation by showing that the diagnostic algorithm is an order of magnitude slower than an expert microscopist. On the other hand, the object detection algorithm was able to locate moving objects with sufficient accuracy and computational time.

The research objective turned out to be more complex than initially thought, the developed diagnostic method is not yet able fulfill this objective completely. However, the algorithm has some promising aspects such as the imaging method, object detection algorithm, and lack of sample preparation method which could make further research worthwhile.

The diagnostic method will not be able to outperform a field microscopist under the considered circumstances. However if some optimizations are made to the algorithm, the device could be used to diagnose SH in the absence of appropriately trained people.

5-2 Recommendations

The research objective was not fulfilled completely, however; it is interesting to look at how the research could be continued or which aspects could see other promising uses.

5-2-1 Imaging method

One of the limiting factors of the method is the amount of holograms that have to be captured. Using a flow cell with a larger depth allows for a higher volume in view of the sensor. Capturing images was one of the most time-consuming parts of the analysis so this can decrease the diagnostic time significantly.

The boundary layer within the flow cell should be analyzed with the goal of preventing missed detections. By finding out how the urine sample behaves as it is moving through the flow cell, an optimal flow speed and piston displacement can be determined such that all of the fluid has been in view of the sensor.

5-2-2 Object detection

The FG detector was optimized for the samples recorded in the lab. The samples recorded in the field contained more particles for which the FG detector was not prepared. This showed that the FG detector should be able to adapt to the varying amounts of contamination it can encounter in the field.

Many of the objects found by the object detector were significantly smaller than eggs. An additional selection stage could be added which looks at the local fringe frequency in order to judge whether an object is of sufficient size. Small particles cause a larger diffraction angle, so the size of objects can be estimated prior to reconstruction.

5-2-3 Classification

The diagnostic method reconstructs objects locally over multiple planes and classifies each of these reconstructions. However the classifier can be trained on the fringe patterns instead, foregoing the reconstruction step in the algorithm. This will save on time since the reconstruction step takes time and a classifier will only have to classify one image for each object instead of every reconstruction plane. The information content in the image is retained during reconstruction therefore the classifier should be able to reach similar accuracy when using the fringe pattern as input.

The classifier was trained on a dataset that did not come from infected people, thus the data is not representative for what the algorithm finds when it analyzes an infected person's urine. A diagnostic method based on a classifier should be trained on a dataset obtained from real infected people with a variety of infection intensities and gender.

The classifier only uses the magnitude of the reconstruction. One of the properties of digital holography is that both magnitude and phase information of the object can be retrieved. This could be capitalized on in the future as it provides more information to the classifier resulting in a higher potential accuracy.

5-2-4 Implementation

The reconstruction and classification of the objects took increasingly more time over the duration of a diagnosis. This could be improved on by implementing the algorithm on embedded code with proper memory management.

Another way to decrease diagnostic time is to increase the computational power which will allow for a higher frame rate. It can then become possible to capture a continuous video instead of separate images. By applying a tracking algorithm such as a Kalman filter the object can be observed across multiple frames allowing for a better view and preventing double counts and missed detections.

5-2-5 Other applications

This diagnostic method was specifically designed for the detection of SH. It is able to find microscopic particles and reconstruct them with details up to to 2.5[μm]. With this amount of detail the device could also be applied in other Fields such as the analysis of bodies of water. For example nanoplankton and microplankton have sizes of $2 - 20[\mu m]$ and $20 - 200[\mu m]$ respectively. [58]Digital holography in combination AI has proven to be a low-cost method for imaging transparent samples at various depths simultaneously. Therefore the application to other diagnostic methods in low-resource settings should be researched.

One of the problems observed during a few tests was that some urine samples have a lot of contamination. With these samples the algorithm found many objects in every frame which had to be reconstructed and classified, which was time consuming. This can be circumvented by assessing the urine sample before and only using the device for clean urine samples.

One of the the main advantages of this device was the absence of a sample preparation method. The drawback of this was that a huge amount of data has to be analyzed. However a filtered sample could also so be analyzed by using a DHM setup, as the the lensless imaging method is low cost and easy to implement. It should be noted that the filter will induce additional fringe patterns which will have to be compensated for. Digital holography could also be applied to analyze a stationary volume containing urine, which would also circumvent sample preparation.

Bibliography

[1] "Trustees of the natural history museum," 2017.

[2] W. H. Organization 2018.

[3] F. Abdel-Wahab, D. M. Amin, D. F. Arfaa, D. B. Blas, D. C. Gang, D. T. El-Khoby, D. P. Jordan, A. Oyediran, D. M. Porto, D. L. Rey, D. B. Sellin, and P. F. V. Lichtenberg, "Public health impact of schistosomiasis: disease and mortality," *WHO Technical Report Series*, 1993.

[4] C. for Disease Control and Prevention, "Laboratory identification of parasites of public health concern, accessed on 5-3-2019,"

[5] E. Linder, A. Grote, S. Varjo, N. Linder, M. Lebbad, M. Lundin, V. Diwan, J. Hannuksela, and J. Lundin, "On-chip imaging of schistosoma haematobium eggs in urine for diagnosis by computer vision," *PLOS*, 12 2013.

[6] B. S. et al., "Field evaluation of a schistosome circulating cathodic antigen rapid test kit at point-of-care for mapping of schistosomiasis endemic districts in the gambia," *PLOS ONE*, vol. 12, 8 2017.

[7] T. Poon and J. Liu, "Introduction to modern digital holography," *Cambridge*.

[8] T. Agbana, H. Gong, A. Amoah, V. Bezzubik, M. Verhaegen, and G. Vdovin, "Aliasing, coherence, and resolution in a lensless holographic microscope," *Optics Letters*, vol. 42, 6 2017.

[9] B. Dickinson and Company, "Introduction to flow cytometry: A learning guide," 12 2002.

[10] A. image Inc., "Basic resolution testing,"

[11] G. Bradski, "The OpenCV Library," *Dr. Dobb's Journal of Software Tools*, 2000.

[12] M. Hoeboer, "Smart optical diagnosis of schistosomiasis," *TU Delft*, 2019.

[13] A. Veldman, "Boundary layers in fluid dynamics," *Rijksuniversiteit Groningen*, 2012.

[14] IDS, "Ui-1492le-m datasheet,"

[15] K. Weerakoon, G. Gobert, P. Cai, and D. McManusa, "Advances in the diagnosis of human schistosomiasis," 2015.

[16] J. Jewsbury, "Variation in egg shape in schistosoma haematobium," *Department of Zoology and Applied Entomology, Imperial College, London, England*, 7 1967.

[17] A. Koukounari, J. Webster, C. Donnelly, B. Bray, J. Naples, K. Bosompem, and C. Shiff, "Sensitivities and specificities of diagnostic tests and infection prevalence of schistosoma haematobium estimated from data on adults in villages northwest of accra in ghana," *Am J Trop Med Hyg.*, 3 2009.

[18] J. Utzinger, S. L. Becker, L. van Lieshout, G. J. van Dam, , and S. Knopp, "New diagnostic tools in schistosomiasis," 2015.

[19] I. Bogoch, J. Coulibaly, J. Andrews, J. Keiser, J. Stothard, E. N'goran, and J. Utzinger, "Evaluation of portable microscopic devices for the diagnosis of schistosoma and soil-transmitted helminth infection," 4 2014.

[20] R. Ephraim, E. Duah, J. Cybulski, M. Prakash, M. D'Ambrosio, D. Fletcher, J. Keiser, J. Andrews, and I. Bogoch, "Diagnosis of scistosoma haematobium infection with a mobile phone-mounted foldscope and a reversed-lens cellscope in ghana," *Cambridge University Press*, 2015.

[21] O. Holmström, "Point-of-care mobile digital microscopy and deep learning for the detection of soil-transmitted helminths and schistosoma haematobium," *Global Health Action*, 2017.

[22] M. Zhao, N. Liu, and Q. Li, "Blurred video detecion algorithm based on support vector machine of schistosoma japonicum miracidium," 11 2016.

[23] Y. Gao, Z. Zong, P. Yang, H. Zhang, W. Y. S. Yang, D.H. Wei, W. Zhou, and Y. Wang, "Automatic microscopic scanning system in etiological diagnosis of schistosomiasis," *Chinese journal of schistosomiasis control*, 12 2012.

[24] J. Hung, A. Goodman, S. Lopes, G. Rangel, D. Ravel, F. Costa, M. Duraisingh, M. Marti, and A. Carpenter, "Applying faster r-cnn for object detection on malaria images," 4 2018.

[25] N. Mahmood and M. Mansor, "Red blood cells estimation using hough transform technique," *Signal & Image Processing: An International Journal*, vol. 3, 4 2012.

[26] G. Gopakumar, M. Swetha, G. Siva, and G. Subrahmanyam, "Convolutional neural network-based malaria diagnosis from focus stack of blood smear images acquired using custom-built slide scanner," 2017.

[27] Z. Göröcs and A. Ozcan, "On-chip biomedical imaging," *NIH-PA*, vol. 6, 2013.

[28] T. Tahara, X. Quan, R. Otani, Y. Takaki, , and O. Matoba, "Digital holography and its multidimensional imaging applications: a review," *Oxford University Press: Microscopy*, 2 2018.

[29] L. Denis, C. Fournier, T. Fournel, and C. Ducottet, "Twin-image noise reduction by phase retrieval in in-line digital holography," *Wavelets XI, SPIE's Symposium on Optical Science and Technology*, 8 2005.

[30] P. Errante, P. Ebbing, F. Rodrigues, R. Ferraz, and N. Silva, "Flow cytometry: a literature review,"

[31] M. Doan, I. Vorobjev, P. Rees, A. Filby, O. Wolkenhauer, A. Goldfeld, J. Lieberman, N. Barteneva, A. Carpenter, and H. Hennig, "Diagnostic potential of imaging flow cytometry," 2018.

[32] M. Vendel, "The ec: a device to diagnose urinary schistosomiasis in ghana," 8 2018.

[33] T. Bouwmans, "Traditional and recent approaches in background modeling for foreground detection: An overview," *Elsevier*, 4 2014.

[34] P. KaewTraKuPong and R. Bowden, "An improved adaptive background mixture model for real time tracking with shadow detection,"

[35] Z. Zivkovic, "Improved adaptive gaussian mixture model for background subtraction," *Intelligent and Autonomous Systems Group*, 2004.

[36] M. Mudrova and A. Prochazka, "Principal component analysis in image processing," *Institute of Chemical Technology, Prague*, 2006.

[37] G. Şengül, "Classification of parasite egg cells using gray level cooccurence matrix and knn," *Biomedical Research(India)*, vol. 27, 2016.

[38] S. J. Raudys and A. K. Jain, "Small sample size effects in statistical pattern recognition: recommendations for practitioners," *IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE*, vol. 13, 1991.

[39] D. Lu and Q. Weng, "A survey of image classification methods and techniques for improving classification performance," *International Journal of Remote Sensing*, 3 2007.

[40] J. Akrimi, A. Suliman, L. George, and A. Ahmad, "Classification red blood cells using support vector machine,"

[41] M. Zhao, N. Liu, and Q. Li, "Blurred video detection algorithm based on support vector machine of schistosoma japonicum miracidium," 12 2016.

[42] A. B. et al., "Image classification using random forests and ferns," 2008.

[43] S. Nath, J. Kar, G. Mishra, S. Chakraborty, and N. Dey, "A survey of image classification methods and techniques," *2014 International Conference on Cotrol, Instrumentation, Communication and Computational Technologies*, 2014.

[44] C.Cortes and C. Vapnik, "Support-vector network," *Machine Learning*, 1995.

[45] A. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreeto, and H. Adam, "Mobilenets: Efficient convolutional neural networks for mobile vision appications," 4 2017.

[46] G. R., "Fast r-cnn,"

[47] O. Yalcin, "Image classification in 10 minutes with mnist dataset." https://towardsdatascience.com/image-classification-in-10-minutes-with-mnist-dataset-54c35b77a38d 2018.

[48] Y. Ghouzam, "Introduction to cnn keras." https://www.kaggle.com/yassineghouzam/introduction-to-cnn-keras-0-997-top-6, 2017.

[49] F. Chollet, "Keras." https://keras.io/examples/mnist_cnn/, 2016.

[50] M. Campbell and J. Freeman, "Diagnostics tests: Statistical evaluation to determine te suitability for diagnosis," 12 2008.

[51] Y. Hong, "On computing the distribution function for the poisson binomial distribution," *Computational Statistics & Data Analysis*, vol. 59, 2013.

[52] U. of Oxford, "Lecture 7: Confidence interval and normal approximation,"

[53] Thorlabs, "S1fc635 fiber-coupled laser source,"

[54] IDS, "Ui-1492le cmos sensor,"

[55] Ibidi, "Ibidi: Cells in focus,"

[56] 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," 2014.

[57] H. Zhang, T. Stangner, K. Wiklund, and M. Andersson, "Object plane detection and phase retrieval from single-shot holograms using multi-wavelength in-line holography," *Appl. Opt.*, vol. 57, pp. 9855–9862, Nov 2018.

[58] V. Dyomin, I. Polovtsev, and A. Davydova, "Fast recognition of marine particles in underwater digital holography," 11 2017.

Glossary

List of Acronyms

NTD	Neglected Tropical Disease
WHO	World Health Organisation
DHM	Digital Holographic Microscope
CNN	Convolutional Neural Network
SVM	Support Vector Machine
CMOS	Complementary Metal-Oxide Semiconductor
ASM	Angular Spectrum Method
DALYs	disability adjusted life years
IFC	Imaging Flow Cytometry
GMM	Gaussian Mixture Model
BG	Background
FG	Foreground
RBF	Radial Basis Function
PCA	Principal Component Analysis
SVD	Singular Value Decomposition
RFC	Random Forest Classifier
FOV	Field of View
AI	Artificial Intelligence
DOH	Determinant of Hessian

POC Point of Care

CI Confidence Interval

SH Schistosomiasis Haematobium

Appendix A: CNN Models

Convolutional neural Network Models that were tested for the classification of Schistosoma haematobium eggs.

Yalcin

[47]

```
 model = tf.keras.models.Sequential()
model.add(tf.keras.layers.Conv2D(32, kernel_size=(3, 3),
activation='relu',
input_shape=(clfsize, clfsize, 1)))
model.add(tf.keras.layers.Conv2D(64, (3, 3), activation='relu'))
model.add(tf.keras.layers.MaxPooling2D(pool_size=(2, 2)))
model.add(tf.keras.layers.Dropout(0.25))
model.add(tf.keras.layers.Flatten())
model.add(tf.keras.layers.Dense(128, activation='relu'))
model.add(tf.keras.layers.Dropout(0.5))
model.add(tf.keras.layers.Dense(1, activation='sigmoid'))
model.compile(optimizer='adam', loss='binary_crossentropy', metrics=['accuracy'])
```

Ghouzam

[48]

```
 model = tf.keras.models.Sequential()

model.add(tf.keras.layers.Conv2D(filters=32, kernel_size=(5, 5), padding='Same',
activation='relu', input_shape=(clfsize, clfsize, 1)))
model.add(tf.keras.layers.Conv2D(filters=32, kernel_size=(5, 5), padding='Same',
activation='relu'))
```

```
model.add(tf.keras.layers.MaxPool2D(pool_size=(3, 3)))
model.add(tf.keras.layers.Dropout(0.25))

model.add(tf.keras.layers.Conv2D(filters=64, kernel_size=(3, 3), padding='Same',
activation='relu'))
model.add(tf.keras.layers.Conv2D(filters=64, kernel_size=(3, 3), padding='Same',
activation='relu'))
model.add(tf.keras.layers.MaxPool2D(pool_size=(2, 2), strides=(2, 2)))
model.add(tf.keras.layers.Dropout(0.25))

model.add(tf.keras.layers.Flatten())
model.add(tf.keras.layers.Dense(256, activation="relu"))
model.add(tf.keras.layers.Dropout(0.5))
model.add(tf.keras.layers.Dense(1, activation="sigmoid"))

optimizer = tf.keras.optimizers.RMSprop(lr=0.001, rho=0.9, epsilon=1e-08, decay=0.0)
model.compile(optimizer=optimizer, loss="binary_crossentropy", metrics=["accuracy"])
```

Keras

[49]

```
 model = tf.keras.models.Sequential()
model.add(tf.keras.layers.Conv2D(32, kernel_size=(3, 3),
activation='relu',
input_shape=(clfsize,clfsize,1)))
model.add(tf.keras.layers.Conv2D(64, (3, 3), activation='relu'))
model.add(tf.keras.layers.MaxPooling2D(pool_size=(2, 2)))
model.add(tf.keras.layers.Dropout(0.25))
model.add(tf.keras.layers.Flatten())
model.add(tf.keras.layers.Dense(128, activation='relu'))
model.add(tf.keras.layers.Dropout(0.5))
model.add(tf.keras.layers.Dense(1, activation='sigmoid'))

model.compile(loss=tf.keras.losses.binary_crossentropy,
optimizer=tf.keras.optimizers.Adadelta(),
metrics=['accuracy'])
```

Appendix B: Field test results

Urine ID	Received	Collected at	Obtained at		Duration manual	Eggs
101	Yes	23/06/2019 - 24/06/2019		###		0
102	Yes	23/06/2019 - 24/06/2019		###		2
103	Yes	23/06/2019 - 24/06/2019		###	04:53	21
104	Yes	23/06/2019 - 24/06/2019		###	04:20	1
105	Yes	23/06/2019 - 24/06/2019		###	04:20	1
106	Yes	23/06/2019 - 24/06/2019		###	05:10	3
107	Yes	23/06/2019 - 24/06/2019		###	04:32	4
108	Yes	23/06/2019 - 24/06/2019		###	02:36	46
109	Yes	23/06/2019 - 24/06/2019		###	02:30	7
110	Yes	23/06/2019 - 24/06/2019		###	01:02	0
111	Yes	23/06/2019 - 24/06/2019		###	01:29	1
112	Yes	23/06/2019 - 24/06/2019		###	01:19	1
113	Yes	23/06/2019 - 24/06/2019		###	01:41	6
114	Yes	23/06/2019 - 24/06/2019		###	03:12	105
115	Yes	23/06/2019 - 24/06/2019		###	02:11	9
116	Yes	25/06/2019 - 26/06/2019		###	02:24	0
117	Yes	23/06/2019 - 24/06/2019		###	04:17	4
118	Yes	23/06/2019 - 24/06/2019		###	03:30	0
119	Yes	23/06/2019 - 24/06/2019		###	02:41	0
120	Yes	23/06/2019 - 24/06/2019		###	02:45	0
121	Yes	24/06/2019 - 25/06/2019		###	03:57	1
122	Yes	24/06/2019 - 25/06/2019		###	03:34	1
123	Yes	24/06/2019 - 25/06/2019		###	03:11	0
124	Yes	24/06/2019 - 25/06/2019		###	03:57	2
125	Yes	24/06/2019 - 25/06/2019		###	04:09	0
126	Yes	24/06/2019 - 25/06/2019		###	03:35	0
127	Yes	24/06/2019 - 25/06/2019		###	02:47	0
128	Yes	24/06/2019 - 25/06/2019		###	04:02	0
129	Yes	24/06/2019 - 25/06/2019		###	03:18	0
130	Yes	24/06/2019 - 25/06/2019		###	03:20	0
131	Yes	24/06/2019 - 25/06/2019		###	03:22	0
132	Yes	24/06/2019 - 25/06/2019		###	02:44	0
133	Yes	24/06/2019 - 25/06/2019		###	02:45	0
134	Yes	24/06/2019 - 25/06/2019		###	03:22	0
135	Yes	24/06/2019 - 25/06/2019		###	03:20	0
136	Yes	24/06/2019 - 25/06/2019		###	02:50	0
137	Yes	24/06/2019 - 25/06/2019		###	03:50	1
138	Yes	24/06/2019 - 25/06/2019		###	02:54	0
139	Yes	24/06/2019 - 25/06/2019		###	03:17	0
140	Yes	24/06/2019 - 25/06/2019		###	02:34	0
141	Yes	25/06/2019 - 26/06/2019		###	02:27	0
142	Yes	25/06/2019 - 26/06/2019		###	02:36	0
143	Yes	25/06/2019 - 26/06/2019		###	02:30	0
144	Yes	25/06/2019 - 26/06/2019		###	03:30	4
145	Yes	25/06/2019 - 26/06/2019		###	03:04	0
146	Yes	25/06/2019 - 26/06/2019		###	03:05	0
147	Yes	25/06/2019 - 26/06/2019		###	03:10	0

148	Yes	25/06/2019 - 26/06/2019	###	03:03	0
149	Yes	25/06/2019 - 26/06/2019	###	03:17	4
150	Yes	25/06/2019 - 26/06/2019	###	02:47	0
151	Yes	25/06/2019 - 26/06/2019	###	03:32	0
152	Yes	25/06/2019 - 26/06/2019	###	02:14	0
153	Yes	25/06/2019 - 26/06/2019	###	02:25	0
154	Yes	25/06/2019 - 26/06/2019	###	02:41	0
155	Yes	25/06/2019 - 26/06/2019	###	03:11	0
156	Yes	25/06/2019 - 26/06/2019	###	02:40	0
157	Yes	25/06/2019 - 26/06/2019	###	02:15	0
158	Yes	25/06/2019 - 26/06/2019	###	02:54	0
159	Yes	25/06/2019 - 26/06/2019	###	03:28	173
160	Yes	25/06/2019 - 26/06/2019	###	02:55	1
161	Yes	25/06/2019 - 26/06/2019	###	02:19	0
162	Yes	25/06/2019 - 26/06/2019	###	02:20	0
163	Yes	25/06/2019 - 26/06/2019	###		
164	Yes	25/06/2019 - 26/06/2019	###		
165	Yes	26/06/2019 - 27/06/2019	###	04:54	3
166	Yes	26/06/2019 - 27/06/2019	###	02:08	0
167	Yes	23/06/2019 - 24/06/2019	###	02:04	0
168	Yes	26/06/2019 - 27/06/2019	###	04:59	0
169	Yes	26/06/2019 - 27/06/2019	###	04:30	0
170	Yes	26/06/2019 - 27/06/2019	###	02:10	0
171	Yes	26/06/2019 - 27/06/2019	###	03:10	0
172	Yes	26/06/2019 - 27/06/2019	###	02:08	0
173	Yes	26/06/2019 - 27/06/2019	###	02:30	0
174	Yes	26/06/2019 - 27/06/2019	###	02:10	0
175	Yes	26/06/2019 - 27/06/2019	###	02:10	0
176	Yes	26/06/2019 - 27/06/2019	###	02:10	0
177	Yes	26/06/2019 - 27/06/2019	###		
178	Yes	26/06/2019 - 27/06/2019	###		
179	Yes	26/06/2019 - 27/06/2019	###		
180	Yes	26/06/2019 - 27/06/2019	###		
181	Yes	27/06/2019 - 28/06/2019	###	02:11	0
182	Yes	27/06/2019 - 28/06/2019	###	02:12	0
183	Yes	27/06/2019 - 28/06/2019	###	02:02	0
184	Yes	27/06/2019 - 28/06/2019	###	02:23	0
185	Yes	27/06/2019 - 28/06/2019	###	02:29	0
186	Yes	27/06/2019 - 28/06/2019	###	02:08	0
187	Yes	27/06/2019 - 28/06/2019	###	02:10	0
188	Yes	27/06/2019 - 28/06/2019	###	02:12	0
189	Yes	27/06/2019 - 28/06/2019	###	02:31	1
190	Yes	27/06/2019 - 28/06/2019	###	02:20	0
191	Yes	27/06/2019 - 28/06/2019	###		
192	Yes	27/06/2019 - 28/06/2019	###	02:10	0
193					
194					
195					

196					
197					
198					
199					
200					
002	Yes	26/06/2019 - 27/06/2019	###	02:38	0
004	Yes	26/06/2019 - 27/06/2019	###	03:09	32
008	Yes	26/06/2019 - 27/06/2019	###	02:08	10
009	Yes	26/06/2019 - 27/06/2019	###		60
021	Yes	26/06/2019 - 27/06/2019	###	02:14	27
028	Yes	26/06/2019 - 27/06/2019	###	02:24	30
029	Yes	26/06/2019 - 27/06/2019	###	02:09	39
033	Yes	26/06/2019 - 27/06/2019	###	02:40	61
066	Yes	26/06/2019 - 27/06/2019	###	02:40	19
067	Yes	26/06/2019 - 27/06/2019	###	02:35	0