UNIVERSITY OF
BIRMINGHAM

Vision Guided Automation for Intra-cytoplasmic Sperm Injection

By

Ferhat Sadak

A thesis submitted to

University of Birmingham

For the degree of

DOCTOR OF PHILOSOPHY

School of Engineering

University of Birmingham

Birmingham, United Kingdom

November 2020

ABSTRACT

Biological cell injection is an effective technique in which a foreign material is directly introduced into the target cell. Intracytoplasmic Sperm Injection (ICSI) is a microinjection technique which is used for infertility treatment. In this technique, a single sperm cell is directly injected into an oocyte using micropipettes. The operations in this application are manually controlled by an embryologist and more importantly, this reduces the accuracy, repeatability, and consistency of the operation. Therefore, the full automation is a prerequisite for microinjection operations, particularly in ICSI application. This thesis focuses on enhancing the microinjection procedure by developing vision-guided processes prior to and during the operation. Initially, a vision-controlled technique was proposed to align the injection and holding pipettes in three orthogonal axes which is essential for successful microinjection. To conduct reliable injection, the vibrational displacement of the injection pipette's tip needs to be evaluated and improved before the operations continue further. A novel vision-based sensor was developed to measure the displacement changes at the tip in three orthogonal axes. By employing the developed vision sensor, the effect of injection speed on vibrational displacement creation was analysed to determine the value of various injection parameters, such as force fluctuation, and penetration force on cell damages. An ultimate automation task is required in microinjection to position the randomly located biological cell within the Petri dish to the system's field of view. The proposed technique fills a gap in the literature by proposing a real-time cell recognising and positioning system that can be employed with different types of biological cells at various maturation stages, as well as with different microscope types that are being used in microinjection applications.

ACKNOWLEDGMENT

I would like to express my sincere gratitude and appreciation to my supervisor, Dr. Mozafar Saadat, for his support and motivation during this research. Dr. Saadat has provided me the opportunity to be a part of his research team and provided me the freedom to explore new ideas and methods. Without his advice, I could not have reached my full potential and abilities for my Ph.D. study.

I would like to acknowledge and thank my colleague and a fantastic friend, Dr. Amir Mohammad Hajiyavand, for his advice on my studies which improved my research work and widen my horizon from many aspects during the journey of my Ph.D. study. I would also like to thank all my friends and colleagues in Mechanical Engineering. It has been a great pleasure to meet you all.

I am grateful to the Republic of Turkey-Ministry of National Education for their financial support during my Ph.D. study.

Considering having a big family, I would like to count each of my family members one by one in a name for all their endless supports during this long journey. Special thanks to my parents, Bekir and, İpek, my sisters, Hülya and Melek, my brothers, Ömer, and Hakan for all their great support and encouragement throughout my life.

Finally, I would like to dedicate my Ph.D. thesis to my wife, Tuğçe. I am forever indebted to her endless and unrequited love and support during this study.

Table of contents

Page

List of figures

List of tables

Abbreviation

IUI - Intrauterine Insemination

IVF - In Vitro Fertilization

ICSI - Intracytoplasmic Sperm Injection

ART - Assisted Reproductive Technology

FMAs - Focus Measurement Algorithms

YOLO - You Only Look Once

DOF - Degrees of Freedom

FOV - Field of View

HCDA - Hough Cell Detection Algorithm

CNN - Convolutional Neural Networks

CLAHE - Contrast-limited Adaptive Histogram Equalization

MFAR - Maximum Force Amplitude Ratio

IoU - Intersection over Union

AP - Average Precision

fps - frame rate

SGDM - Stochastic Gradient Descent with Momentum Optimizer

TP - True Positive

TN - True Negative

FP - False Positive

FN - False Negative

Chapter 1: Introduction

1.1 Background

Infertility is defined as the inability of conceiving for a couple trying regular unprotected sexual intercourse. According to World Health Organization (WHO) data collected through the Demographic and Health Survey (DHS), 186 million women were infertile in developing countries excluding China in 2004, which translates into one in every 4 couples [1]. There are various infertility treatments available globally, namely intrauterine insemination (IUI), in vitro fertilization (IVF), and intracytoplasmic sperm injection (ICSI).

ICSI, which is first introduced in 1992 [2], is an Assisted Reproductive Technology (ART) where a single sperm is manually selected, immobilized, and then injected into an oocyte using an injection micropipette. ICSI is mainly used where the cause of infertility is due to male factor deficiencies. The illustration of the ICSI process is shown in Figure 1-1.

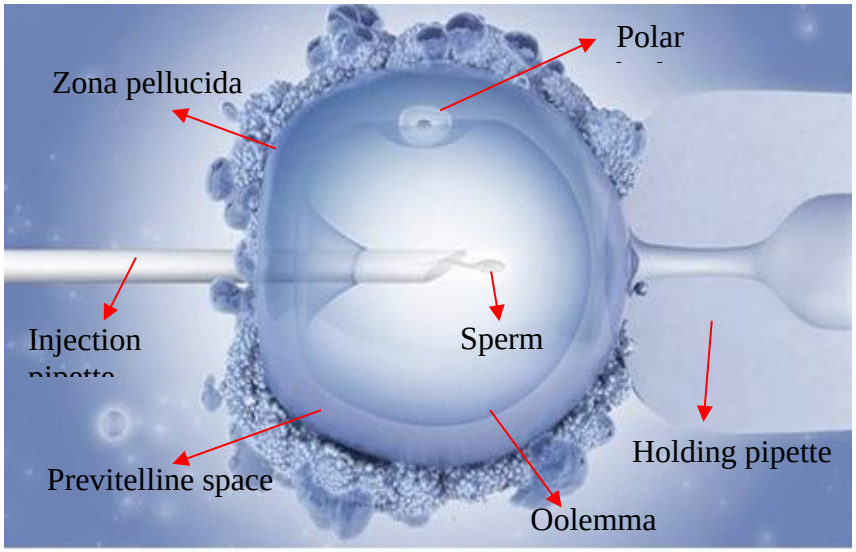


Figure 1-1 Intracytoplasmic Sperm Injection (ICSI) procedure [3]

Microinjection is a technique in which a foreign material such as proteins, DNA, sperms, etc. is injected using a fine needle. The microinjection technique is effectively used in ICSI operation to inject a sperm cell into the oocyte due to its advantages of causing minimal harm to the oocyte [4] in comparison to other techniques, such as electroporation [5] and gene guns [6]. However, a manually operated ICSI process demonstrates deficiency in the fulfillment of robust injection operation.

The current ICSI technique requires manual handling from a trained embryologist, and the results are not satisfactory [7]. Manually handled ICSI operation cause inconsistency, high-cost, low accuracy, and unreliability across the operators as it relies on the human-controlled operation. Automated microinjection systems eliminate the need for training of embryologists, reduce the risk of contamination, and increase the success of the microinjection operation. Therefore, there are various manual operation aspects involved in the microinjection process which require full automation [8].

There are clinical and non-clinical factors that affect the outcome of the fertilization rate in ICSI. The clinical factors are mainly sperm source, female patient's age, and ovarian reserve, etc. [9]. However, this study only investigates the potential deficiencies caused to the ICSI system from the engineering point of view. The core of inconsistencies in ICSI operation is due to human involvement. For this reason, automation of the microinjection procedure has become a significant focus for the researchers due to the low success rate in manual microinjection operations [10, 11]. The engineering challenges in the ICSI operation are the alignment of the injection and holding pipettes in three orthogonal axes, the vibrational displacement sensing at the tip of the injection pipette, and the automated positioning of the egg to the system's field of view. These challenges need to be conducted automatically before the microinjection procedure starts in ICSI. In this thesis, vision-based solutions are proposed

to deal with the aforementioned challenges mentioned.

Visual sensing in the microinjection procedure refers to the capability of the system to visually perceive the environment without requiring a physical interaction to extract useful information and conduct various tasks. These various vision-based tasks could be a measurement, object location extraction, object detection, and object positioning to guide the microinjection system. The data collected by the developed vision system can improve the processing speed, accuracy, and consistency of the task performed by the operator as well as reduce the cost. This is particularly a demand in biomedical applications where cameras are a significant component within the system, such as ICSI.

Practical usage of physical sensors is limited to be used towards the automation demand of ICSI application. It is hypothesised that vision-based sensing platforms are strong candidate among the sensor types to overcome automation challenges in the field of microinjection and manipulations applications, particularly for ICSI.

This thesis is aimed at developing a novel vision-based sensing system to be utilised in a fully automated microinjection system and enhance the current deficiencies in the automation of the microinjection process within ICSI to help practitioners in eliminating the variability of the injection procedure [12].

1.2 **Motivation**

The conventional method of egg injection is not efficient since it involves a human operator using a joystick for motion control and a microscope for visual feedback. To advance this system, there are several reasons to make auto-alignment of the injection and holding pipette a necessity. First of all, if the alignment was not conducted accurately by an operator, this

would cause failing in penetration of the oocyte membranes, such as zona pellucida and oolemma [13]. The consequence of inaccurate penetration would cause damage to the oocyte membrane, which directly affects the fertilization chance. The damage to the injection pipette may also require the replacement of the injection tool, adding extra cost to the system. In addition to cost, replacement of the injection tool would cause an issue in the system operation if the developed vision system relies on the length or the template of the injection pipette, such as template matching technique. This computationally expensive method relies on the template that is obtained before the injection, and if the injection pipette is replaced, the template should be captured again which is time-consuming and not efficient.

The employment of physical sensors in healthcare technologies is limited, particularly in microinjection systems such as in ICSI. This is because the physical sensors are commonly prone to be affected by the temperature, provide low sensitivity and precision, and low signal-to-noise ratio. In addition to these issues, the interface of the physical sensors is highly challenging due to the limited space and inappropriate assembly location within ICSI setups. Therefore, the usage of vision-based sensing systems provides great advantages in the field of microinjection applications as it is a non-contact measurement and non-invasive as well as easy to set up and within the system.

Additionally, if the system does not receive feedback from the sensor regarding the displacement changes at the tip of the injection pipette, the resolution of the motorized stage cannot be effectively used to control the injection pipette movement. The lack of sensing in the displacement of the injection pipette tip would also cause the failure of the position and force control strategies that are developed for microinjection operations [14]. While the effect of vibrational displacement within the operation of the system is significant, excess displacement would also cause damage to the oocyte membrane and spoil the structure of the

cell. This directly affects the success rate of the injection operation or causes failure [15].

The fully automated microinjection systems in the literature always assumed that the biological sample is already placed in the field of view of the system. Cell positioning task is commonly conducted through a joystick controlling the stage underneath of the Petri dish, or the dish is manually maneuvered by hand. This is not practical and requires full automation. This is because this manual task relies on human operator ability through a vision system, which affects the repeatability, reliability, and consistency of the operation undesirably.

Visual sensing provides useful and accurate information regarding the dynamic environment in the system. This motivates this study to recognise and evaluate the wide practical usage of vision-based sensing for automation in ICSI. This is aimed at minimizing human involvement and increasing the success of microinjection operation, mainly in ICSI.

This thesis targets to fulfill the requirement of microinjection automation using visual sensing. This includes micropipette auto-alignment and pipette's tip vibrational detection in three orthogonal axes, and finally automated positioning of the biological cell to the microscope's field of view as a part of automation for microinjection.

1.3 **Research aim, objectives, and methodology**

The aim of this thesis is the development of a novel vision-based automation for microinjection technology for biological applications, especially for ICSI.

The objectives of this thesis are as follows:

- To develop a robust technique to align the injection and holding micropipettes in three orthogonal axes. The technique should be insensitive to noise for the injection and

holding pipettes at random positions in the image plane. The developed technique aimed to be within the tolerance of 15 µm when the injection pipette is inserted into the holding pipette.

- To develop a visual sensor to measure the vibration displacement of the tip of the injection pipette in three orthogonal axes and various magnifications. The role of vibration displacement sensing in microinjection operations will be analysed. The vibration displacement sensing platform is aimed to have one-pixel resolution.
- To automatically position the biological cells to a predefined position within the microscope's field of view in real-time. The developed technique will be generalized in other microinjection setups to demonstrate its effectiveness. The centre point of the biological cell aimed to be located within the center coordinate of the field of view of the microscope with a maximum of 10% error when the full operation is conducted.

In this research, a novel vision detection software was developed for automating the microinjection operation within ICSI. Figure 1-2 illustrates the graphical overview of how the proposed methodology is achieved and conducted within this study. The full operation consists of three main sub-operations, namely, auto-alignment of the injection and holding pipette, system stability check through vibrational displacement measurement, and development of a zebrafish embryo positioning system.

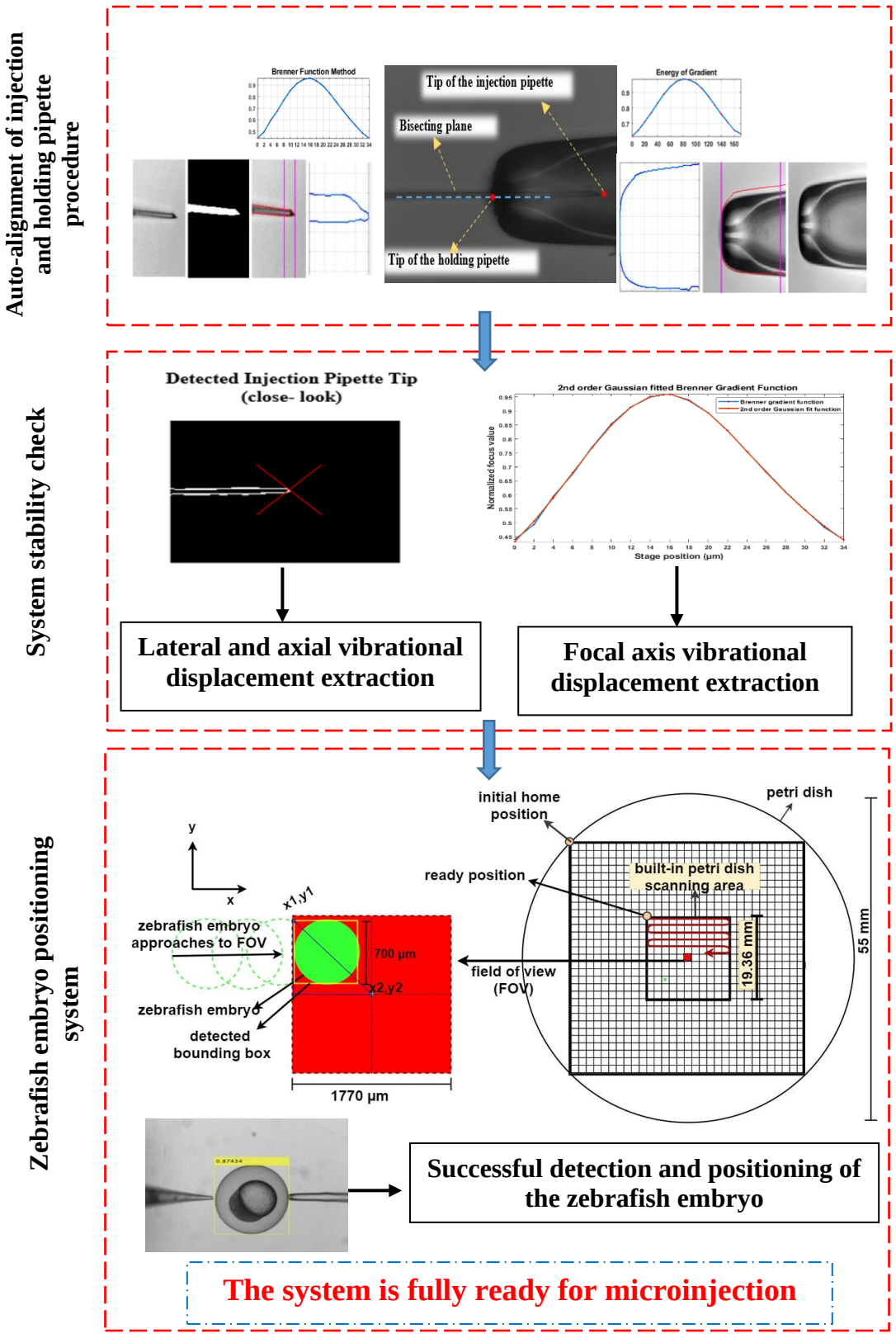


Figure 1-2 A graphical overview of the developed vision-based automated microinjection operations

MATLAB software was used throughout this research as a strong platform for the developed algorithm. Mainly Image Processing and Neural Network Toolbox are utilised to conduct the research. Conventional image processing techniques were employed for aligning the injection and holding pipettes, as well as extracting the vibrational displacement of the injection pipette tip in three orthogonal axes. Convolutional Neural Networks (CNNs) were used to automatically recognise and manipulate the zebrafish embryo to the desired position within the field of view of the microscope.

The zebrafish embryo was used as a biological model. This is because the zebrafish embryo is a suitable biological model that is commonly used in bioscience due to its similar structure to humans [16]. Besides, it is easy to retrieve which makes the zebrafish embryo to be a suitable model in experiments. The in-house captured zebrafish embryo and representation of its structure is shown in Figure 1-3.

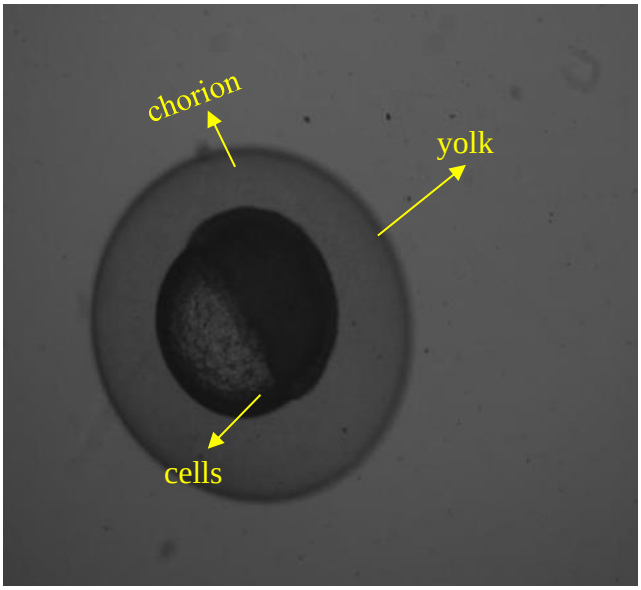


Figure 1-3 The structure of zebrafish embryo (in-house image)

1.4 **Thesis structure**

This thesis is divided into 6 chapters. The structure of the thesis is as follows:

Chapter 2 covers the literature review of the published researches. This chapter provides a comprehensive background for the research demonstrated in the following chapters.

Chapter 3 proposes a method to align the injection and holding pipettes in three orthogonal axes prior to injection operation. In this part of the study, various Focus Measurement Algorithms (FMAs) were analysed and evaluated based on the ranking methodology. Later, the evaluation results were analysed and experimental validations were presented.

Chapter 4 reports the unique development of the vision sensor to track the vibration displacement in three orthogonal axes at the tip of the injection pipette during the egg injection using a microscopic camera. The proposed vision sensor operates between magnifications of 4x to 40x. Considering the cell injection speed is a dynamical injection factor contributing to the creation of cell deformation, this chapter also covers the effect of injection speed on lateral vibration displacement occurring at the tip of the injection pipette. The experimental results link the values of egg deformation, force fluctuation, and penetration force to demonstrate the effect of vibration on reducing the biological cell damages. This chapter demonstrates how the sensing of vibrational displacement at the tip of the injection pipette is a decisive parameter and limits the cell injection speed during the operation.

Chapter 5 covers the methodology proposing the development of vision-guided automation for the recognition and positioning of the randomly placed zebrafish embryo withing the Petri dish to the field of view of the microscope. In this chapter, eight different pre-trained neural networks have been used as a backbone with you only look once (YOLOv2) method through

transfer learning. Based on the analysis and evaluation of the proposed method, the positioning of the zebrafish embryo within the field of view has been successfully illustrated. Chapter 6 concludes this thesis by providing a detailed analysis of chapters 3, 4, and 5 and correlations between them. This chapter also suggests future works towards the development of vision-based automation in the field of microinjection and micromanipulation applications.

Chapter 2: Literature Review

2.1 Introduction

The current microinjection operation as a part of the ICSI application, is conducted manually by the operator using joysticks. This operation is conducted using manual micromanipulators to control the movement of the injection and holding pipettes as well as Petri-dish, depending on the system setup. The injection procedure is a delicate task that needs to be conducted precisely by a skilful embryologist. For instance, an injection procedure must be conducted through a specific orientation of the oocyte where polar-body at the upper location of the oocyte (named as 12 o'clock) or at the lower location of the oocyte (named as 6 o'clock) to avoid damages to the polar body of the oocyte and be as far as possible from the polar body. Any damage to the polar body directly results to failure of the injection procedure since it causes deficiency and complications to the embryo. Hence, the injection operation requires visual control including accurate detection of the injection and holding pipettes and their alignment in three orthogonal axes to have full control in microinjection. Also, the vibration displacement sensing by the vision system needs to be conducted to ensure a safe injection. As a part of the full automation of the procedure, the randomly located egg within the Petri dish needs to be located to the system's field of view automatically to eliminate human involvement. This requires an accurate vision-based egg recognition and positioning system. There has been little effort in the literature in the vision-based microinjection automation operation. The following section is divided into three parts which address the relevant literature in object detection and autofocusing, vision-based displacement sensing, and biological cell detection and positioning.

2.2 Object detection and autofocusing

Intracytoplasmic sperm injection (ICSI), first introduced in 1992, is an assisted reproductive technology in which a single spermatozoon is injected into the oocyte using a glass injection micropipette [2]. There has been considerable demand to automate the ICSI processes to enhance their success and survival rate [17–21]. This is because the manual operation is conducted by a trained embryologist, relying on the visual information through the microscope. However, manual operation is time-consuming and may cause oocyte damage, with the consequence of a low success rate. To automate the oocyte injection task, the injection pipette must be aligned with the holding pipette at a desired position prior to the ICSI procedure to ensure the correct injection. This is key to accurately conducting the fully automated ICSI process and increasing system efficiency. This is because the tip of the injection and holding pipette can be easily damaged during the ICSI process if the injection pipette is not aligned correctly to the holding pipette. The micropipettes could fail to penetrate the zona pellucida or oolemma membrane of the oocyte [13]. This would cause a serious injury to the oocyte membrane or require a replacement of the micropipettes. Therefore, this makes accurate focusing a necessity.

Autofocusing can be examined using passive and active methods depending on the way of measuring the distance between the lens and the object. Active autofocusing measures the distance externally using ultrasonic or infrared waves, whilst passive autofocusing measures the sharpness of the images captured at different focal positions and finds the peak point of the obtained focus curve [22]. Active autofocusing demonstrates an effective sensing mechanism under different lighting conditions. However, such a sensing technique may struggle to focus well through glass due to high infrared or ultrasound reflectivity [23].

Passive auto-focusing relies on the image information and it is not essential to consider reflectivity since there is no external sensor included. It also decreases the cost markedly. Consequently, image-based passive auto-focusing will be discussed in this thesis since it does not require any additional device such as a range finding sensor as in active autofocusing [24].

A variety of focusing algorithms have been proposed such as Brenner gradient, Tenengrad gradient, and Energy gradient to adjust the experimental sample in an accurate position in the focal zone of the microscope [25-27]. Since the injection and holding pipette present a unimodal focus curve, focusing algorithm methods can be utilized. However, the selection of an appropriate focus algorithm for a fully automated ICSI process remains deficient in previous researches since the selection of the optimal autofocusing algorithm greatly depends on the specific application [28]. Previous literature has discussed finding the focal distance of the micro-objects as it is a challenging task in micro applications where the depth of focus is a crucial parameter and plays an essential role.

Template matching is a technique in the field of computer vision for searching a portion of a sourced image which matches a template image obtained in advance. Template matching technique has recently been proposed for an automatic oocyte injection process to find the depth of micropipettes [29]. Based on the technique, when the error minimized between the focal position and template obtained before, the focused image is obtained. The limitation of this method is that a new template is needed to feed the proposed technique if the micropipette is changed. Additionally, the matching process of the injection pipette contains all possible locations in the sourced image, which increases the computational cost of the technique. Microfingers are micro-grippers which are made from a hollow glass tube and used for gripping and handling biological cells [30, 31]. The depth of field for the tip of the

micro-fingers can be searched by using colour information along with template matching techniques as the shape of the finger is fixed during the injection task [32]. Colour information helps to find the correct direction of the focal plane. However, any method for positioning the micro-objects using template matching technique is not an effective way to track the micro-objects without a clear morphology [33]. On the other hand, any method using the template matching technique is computationally expensive [34].

The inner space of the holding pipette can also be utilized for focal plane estimation. When the inner space of the holding pipette has the highest width value, then it is claimed that the image is the focal plane [35]. Due to the large set of pre-recorded image collection before applying the image processing algorithm, utilization of the width of the holding pipette is not practical. This is due to changes in image processing algorithm parameters, and the image collection task would need to be repeated.

Contrast measurement greatly affects the accuracy of the focusing procedure but the efficiency of the FMA depends on the peak search algorithm used. A variety of search strategies have been proposed in the literature to drive the focus motor in the desired position. Global search is the most widely used method. In this method, all focal positions are scanned, and the peak point is detected. Therefore, there is no possibility of detecting a wrong focal position. Conversely, this method is not practical when a high focus range is needed for a specific application as it requires high computation time. Fibonacci and particle swarm optimization with Gaussian jump are studied in the literature in which suitability of the search algorithm depends on the shape of the focus curve [36, 37]. Since the result of the developed auto-alignment algorithm produces a unimodal curve after evaluation of the 12 different focus measure algorithms, a Fibonacci search algorithm has been found to be an appropriate search method and is implemented into the developed algorithm for auto-alignment of the

injection pipette task. This method is further improved using curve fitting to eliminate local optimum disturbance [38]. Implementation of a Fibonacci search algorithm requires a motorized stage to be driven with a high resolution for the precise alignment of the pipettes. Suitable focus algorithms and searching methods are two essential parameters for a successful auto-alignment of the injection pipette process.

2.3 Vision-based displacement sensing

Physical sensors are commonly used to measure the displacement of the objects in the micro-world in various applications such as micro-manipulation and micro-gripping [39, 40]. However, these types of sensors have limited measurement distance and are sensitive to object spot size and the object's material. Existing micromanipulators for biological egg injection are not equipped with any displacement sensors as it is not possible to interface due to the inappropriate assembly location which causes a high amount of noises interacting with the real data. Considering the limitation of the sensor accommodation in intracytoplasmic sperm injection (ICSI), interfacing three sensors for each access would not be practical. On the other hand, since the ICSI operation is conducted when the egg and sperm are in an aqueous medium, this affects the accurate measurement of the displacements for both contact and-non-contact physical sensors. Additionally, sensor interfacing greatly increases the cost of micromanipulation setup. Vision-based displacement sensing is more suitable to use in an aqueous medium [41]. The outstanding advantages of the vision-based displacement sensor are its low cost, easy to implement, high reliability and accuracy [42], and measurement of displacement in three orthogonal axes at once by a single camera.

Automated microinjection has demonstrated rapid improvement in various fields such as

ICSI which offers a precise motion for the injection pipette during egg injection [29]. Injection speed and trajectory of the injection pipette are two main injection parameters that contribute toward a successful egg injection [43]. The displacement of the injection pipette is affected significantly by cell injection speed and its trajectory [44, 45]. Besides, the magnitude of the vibration at the tip of the injection pipette increases since the injection procedure requires acceleration and deceleration of the injection pipette [46]. Produced vibration at the tip of the injection pipette is out of control and negatively affects the accuracy of the egg injection operation. The injection pipette is driven by a motorized stage during egg injection and the resolution of the motorized stage is not effectively utilized if the undesired displacement changes are not measured correctly. Regardless of the amount of vibration induced at the tip of the injection pipette in any direction during egg injection, uncontrolled vibration is an undesired dynamic factor which may affect the control strategies developed, particularly in the ICSI operation, such as position and force control [14]. On the other hand, induced undesirable vibration at the tip may cause damage to the membrane of the oocyte and lead to degeneration. If this damage cannot be healed effectively, abnormal growth occurs and this causes the failure of the operation [2, 15]. Overall, vibration is an undesired dynamic factor preventing the injection operation to be conducted in a stable environment. Therefore, minimizing the vibration at the tip of the injection pipette is essential in the ICSI operation.

Analysis of vibration in the ICSI procedure is very challenging due to the complexity of the ICSI machine and the impractical use of displacement sensors. Analytical and numerical models have been proposed in the literature to analyse the vibration at the drawn section of the injection pipette [47]. There are various sources of internal and environmental parameters which induce the vibration on the injection pipette. Hence, it is not safe to rely only on

numerical and analytical analysis as it does not cover all existing dynamic parameters which affect the displacement of the injection pipette tip. In another study, three photonic displacement sensors were implemented into the cell manipulation system to measure the vibration at the tip of the injection pipette holder for each vibrational direction as reported in [48]. However, due to the tiny diameter of the injection pipette tip, the injection pipette could not be present in their experiment; instead, it has been detected manually in 2 dimensions by using a high-speed camera.

A vision-based sensor for vibrational detection of the injection pipette tip in three dimensions requires an accurate tip recognition of the injection pipette. Image-based detection of the injection pipette tip position in the X, Y axes, and focus estimation in the Z-axis is one of the steps to automate the ICSI process [33]. As the injection pipette must be located under the microscope prior to the cell injection procedure, Wu et al. proposed a vision-aided injection pipette tip detection algorithm [49]. In their study, Canny edge detection and Hough transform were utilized for the detection of the injection pipette tip. However, the disadvantage of using the Hough transform is that it can only be applied if the object has a regular shape, such as lines or circular shapes [50, 51]. The template-matching technique has also been proposed for micro-object tip tracking purposes in micro-robotics applications [29, 52]. However, locating micro-objects using the template matching technique is not an efficient way since this method requires clear morphology [33]. Then again, using a template matching method is also computationally expensive [34]. Li et al. developed a technique to detect the micro-motion of the 3 degrees of freedom (DOF) precision positioning stage and then a nanometer-level accuracy was achieved by the proposed micro-vision imaging system to extract the in-plane displacements data [53, 54]. Zhang et al. developed a robust rotation-invariant displacement method for the micro-nano positioning system [55, 56]. Theoretically,

0.001-pixel accuracy was achieved in their study. However, none of these types of vision-based precision positioning stage micro-motion detection helps to measure the injection pipette displacements in three orthogonal axes in ICSI operation and, in particular, it does not provide information in focal axis changes at the tip of the micropipette which makes accurate vibration displacement measurement in three-dimensional space a necessity in the field of ICSI.

2.4 **Biological cell detection and positioning**

A biological microinjection procedure is required to deploy a single foreign material within biological cells using micropipettes. The positioning of the biological cells to the field of view (FOV) of the microscope in the microinjection task is conducted through a heuristic approach by an embryologist. Therefore, the localisation of the biological cell is unreliable, inconsistent, high cost, and time-consuming. For instance, automation of the Intracytoplasmic sperm injection (ICSI) is an example of biological microinjection application that has been explored by various research groups, such as auto-alignment of micropipettes [57], automated sperm immobilization [58], and insertion of sperm into the oocyte [13]. Human involvement affects the repeatability and accuracy of the task across operators. Automation of this procedure minimises the requirement of any manual handling to locate the egg to the desired position. Hence, automation is desirable to conduct reliable microinjection [8].

There is a number of researches that have been conducted towards automation of the microinjection procedure in the literature, such as positioning [59, 60] and injection of biological cells [43], three-dimensional microinjection pipette tip displacement measurement

[61], and identification of the specific structure of the biological cells [35, 50, 62]. Wang et al. proposed the first automated microinjection system in 2007 [63] in which the zebrafish embryos were placed and immobilized in 5 x 5 array-like in-house designed and held by vacuum-based embryo holding device. The device is placed on an XY controllable stage to allow precise movement of the cells into the field of view. However, randomised positioning of the zebrafish embryo has not been studied as the holder's positions had been defined to the system. In the same year, a semi-automated microinjection system was proposed in [64]. In their study, a charge-coupled device (CCD) camera was employed to acquire the video, and then the process was teleoperated by an operator through a joystick. The XY stage was incorporated in their study to move the petri dish along with wells to bring them into the FOV of the microscope during the process via joystick. However, this process relies on a human-based-operator who provides positional feedback for visual servoing and controls the XY stage by joystick. This manual task influences the repeatability of the injection task undesirably. In 2008, a vision-servo system was proposed in [34] for automated cell injection purpose, however, automated positioning of the biological cells to the field of view of microscope was not included in their study.

Fast, accurate, and reliable biological cell positioning to the FOV of the microscope requires the detection of the cell. The traditional image processing techniques have been used commonly to detect biological cells. In the first attempt, the morphological operator and Bayesian estimation technique were utilized for cell detection and tracking purposes [65]. The main drawback of this method is that it is not suitable for real-time applications as it is computationally expensive. Hough cell detection algorithm (HCDA) is also commonly used for the detection of biological cells [34]. However, it is computationally expensive and required 22 seconds to execute on a single image. This limits its usage for fully automated

microinjection systems. Additionally, the Hough transform is a suitable method that can be used as long as the detected objects present a regular shape, such as lines and circles [51]. Deep learning techniques enhance the detection procedure compared to conventional techniques as it reduces the expensive computational time as well as detecting irregular shapes which is common among the biological cells [66].

Recently, deep learning techniques are a hot topic to solve difficult automation problems, particularly in medical applications such as CT, MRI [67], blood cell count [68], identification of vesicular and electron transport proteins [69, 70], and detection of diabetic retinopathy in retinal fundus photographs [71]. Deep learning is a subset of machine learning in which inspired by human brain structure which is called artificial neural networks. These neural networks consist of many nodes that each perform a simple task in a collaborative way to each other. It simulates the behaviour of the human brain nervous system and makes a meaningful prediction. Each node has a numeric weight that is being tuned during the training process. Deep learning consists of various categories and Convolutional Neural Networks (CNN) is the widely used one in image processing applications. CNNs use a number of layers to extract the image feature information and then convert the data into a high dimensional and non-linear space. The performance of CNNs in object category classification and detection on hundreds of object categories and millions of images was successfully shown in [72]. Hence, it is remarkable to investigate the implementation of deep learning algorithms to automatically locate the zebrafish embryo in the FOV of the microscope prior to the microinjection task.

R-CNN was the first deep-learning-based model for object detection [73]. In this method, a selective search algorithm was employed [74] and the region of interest was generated. Later, CNN was used to check whether the region of interest is related to the background or the

object itself. Subsequently, Fast R-CNN [75] and Faster R-CNN [76] were proposed. Although their object detection accuracy is promising, it is still computationally expensive and not suitable for real-time applications. To increase the detection accuracy and speed of the region based detection methods, Redmon et al. proposed a CNN based object detection method, called you only look once (YOLO) [77]. In contrast to region based algorithms such as Faster R-CNN, the CNN-based YOLO algorithm looks at the image at once and predicts the bounding boxes and class probabilities in a single evaluation. Since the whole detection procedure is in a single network, it can be optimized end to end directly and detection accuracy can be improved. In 2017, Redmon et al. proposed YOLOv2, the improved version of the YOLO method [78]. In this detection model, the detection speed further improved while keeping the detection accuracy stable. Hence, the YOLOv2 framework was selected in this study due to its outstanding detection accuracy and speed.

The FOV of the microscope is limited and the maneuver of the biological cells to the FOV requires several trials by moving the petri dish while checking the cell visibility under the microscope. This time consuming and laborious work is a part of the automation procedure of microinjection systems and it requires full automation.

2.5 **Conclusion**

Microinjection is the most suitable technique to introduce a single sperm cell into the egg using a fine needle. To eliminate human involvement, numerous approaches were proposed in the literature as a part of automation in microinjection. Visual sensing is the primary sensor that is feasible to be used in ICSI and reduces the potential challenges of physical sensors. The various technical aspect of microinjection procedure within the context of automation

from a visual sensing point of view was discussed in this chapter.

Initially, injection and holding pipette detection as well as their autofocus procedure were reported. The effective approach for measuring the degree of focus for injection and holding pipette is the selection of Focus Measurement Algorithms (FMAs). This is because the optimal selection of the FMAs is the system-specific and requires detailed evaluation.

Physical sensor usage in the microinjection procedure is not feasible. The role of vibration displacement sensing in microinjection was reported. The coordinate extraction at the tip of the injection pipette requires an accurate tip position recognition. Here, different methods of injection pipette tip coordinate extraction were discussed along with advantages and disadvantages.

Currently, egg positioning to the system's field of view is conducted manually and requires maneuver of the egg with several trials. Automation of this procedure requires an accurate egg recognition and manipulation system. Different types of egg recognition approaches were discussed in detail along with their advantages and disadvantages. Additionally, the conventional image processing and deep learning techniques in recognition of objects in medical applications were reported.

Chapter 3: Three-Dimensional Auto-Alignment of the ICSI Pipette

3.1 Introduction

Intracytoplasmic sperm injection (ICSI) is an assisted reproductive technology used in infertility treatment, where a single sperm is selected and immobilised using a glass injection pipette and is inserted directly into the cytoplasm of an oocyte under the microscope. The auto-alignment of the injection pipette is a prerequisite for any proposed automated oocyte injection procedure. In this chapter, an auto-alignment procedure is proposed. This technique requires the positioning of the injection and holding pipettes in the three orthogonal axes under microscopy which is complex. The existing methods are system-specific and require appropriate algorithms. In this chapter, twelve commonly used focusing algorithms are evaluated to verify the optimal one for ICSI application. These algorithms were assessed by measuring focusing error, range, number of false maxima, and the width of the curve. The focus level for each pipette is calculated by the algorithm using focus measure functions (FMA). The Fibonacci search algorithm is employed for controlling the z-axis of the motorised stage to obtain the focal plane of the injection and holding pipette.

3.2 System Design

3.2.1 System Configuration

Figure 3-1 shows the system setup which consists of a Best-scope 2700 inverted microscope, a microscopy camera (BASLER Camera acA1300-200um - Python1300), a holding micropipette with a 120 µm outer diameter (Smiths Medical-Wallace WHP-120B-30, USA)

and a 5.5-µm spiked tip ICSI injection micropipette (Smiths Medical-Wallace WIC-55H-30, USA). Images are captured through a 40x objective lens in maximum lighting conditions. Captured images are processed by an Intel® Core™ i5-6500 CPU @3.2 GHz (4 CPUs) host computer. The 8MT167 3-DOF motorized stage from Standa is employed to control the position of the holding and injection pipettes with a resolution of less than 1 µm. The produced vibration during the alignment procedure is damped using an anti-vibrational table.

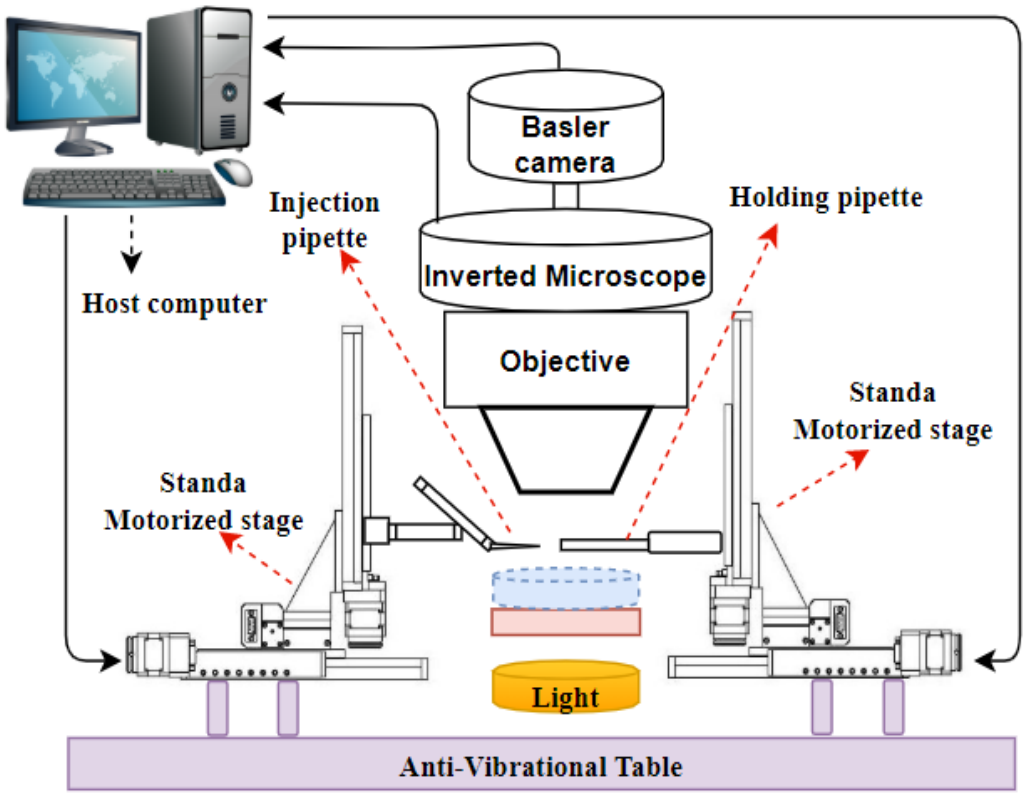


Figure 3-1 System setup for the auto-alignment procedure

3.2.2 Method Overview

During the auto-alignment operation, the pipettes are initially detected using the visual detection algorithm and then are manipulated automatically using the information received

from the focus curve and focus algorithm. A focus curve demonstrates the focusing position of the pipette based on a calculated focus value obtained from focusing algorithms. The behaviour of the focus curve plays a crucial role in the alignment of the injection and holding pipettes. Ideally, the focus curves obtained from focusing algorithms should be unimodal and sharp at the top and long-tailed. However, some focus algorithms may present various local maxima or the same focus value at different focal positions. This causes an error in the accurate measurement of the focus point. Therefore, evaluation of the focus algorithms is essential in finding a suitable algorithm to implement for the auto-alignment of the pipettes. In this study, different focusing functions were compared quantitatively, and focus curves were normalized to obtain the most accurate algorithm for each injection and holding pipette. The Fibonacci search algorithm was then implemented to drive the injection and holding pipettes in Z-axis independently to position them in the focal plane.

In the following section, 12 different focus algorithms are described briefly which are used to obtain the optimum focus curve.

3.2.3 Autofocus Functions

A variety of focus algorithms have been proposed in the literature to find the optimal algorithm for a specific application. The ideal focus curve is defined as having the sharpest image at the maximum value of the focus curve. This value decreases once the images demonstrate further distance positions from the focal plane. To obtain the initial position of the pipettes, an automated stage moved the pipettes in Z-axis until the micropipettes become partially visible in the field of view. The focalization procedure of the pipettes starts after the immediate detection of the pipette. To measure the focus values by implementing each

algorithm, the stage moved in an increment of 2 and 10 μm for the injection and holding pipette respectively, and scanned the pipettes. The focus curves are then obtained based on the algorithm results. In this study, twelve different focus algorithms were examined, evaluated, and compared for each pipette separately to find the optimum algorithm demonstrating a good distribution of the values in a curve and to conclude a single focal plane. The measured values obtained from the algorithms are correlated to the position of the stage. Consequently, the focal level was calculated based on the home position and the differences between each incremental movement of the stage which was linked to the focus curve.

This is the first report on the extensive evaluation of the Focus Measure Functions (FMA) in the field of oocyte manipulation and sperm injection. The FMA used in this study can be classified as derivative-based, statistical-based, and histogram-based algorithms.

Derivative-based Algorithms

Derivative-based algorithms assume that focused images hold more high-frequency content than unfocused images. Hence, the larger intensity changes provide sharper images. The advantage of the derivative-based focus algorithms is to provide highly focused accuracies. However, derivative based algorithms are prone to be affected by high- frequency noise.

1. *Thresholded gradient*

The sharpness of an image is proportional to the high frequency of an image and high pass filters are widely used to detect the high frequency of the portions of an image [79]. The Thresholded gradient algorithm computes the first difference of the pixel intensities and then accumulates if it reaches above the specified threshold. Therefore, this algorithm reveals a

larger gradient in the image. The focalization value is given in equation (3.1)[80].

$$F_{th\_grad} = \sum_{M}\sum_{N} |g(i,j+1) - g(i,j)| \tag{3.1}$$

$$\text{While} \quad |g(i,j+1) - g(i,j)| \geq \upsilon \tag{3.2}$$

where g (i, j) represents the gray level intensity of pixel (i, j), and υ the gradient threshold. In this algorithm, M x N image is used where M and N present the height and width of an image respectively in pixels.

2. Modified Laplacian

A modified Laplacian as a focus measure function is computed at a point (i, j) in a window around (i, j), which is greater than a threshold value [81]. Equation (3.3) demonstrates how this method uses the discrete approximation of the Laplacian to compute the second derivatives in the horizontal and vertical direction of an image.

$$F(i,j) = \sum_{x=i-N}^{i+N} \sum_{y=j-N}^{j+N} ML(x,y) \; for \; ML(x,y) \geq T_1 \tag{3.3}$$

where N represents the window size used to compute the focus measure function. N is taken normally 1 or 2 as the modified Laplacian method uses very small window size such as 3-by-3 or 5-by-5.

3. *Tenengrad gradient*

In this algorithm, convoluted images are used along with Sobel operators to compute the first difference as in equation (3.4)[26].

$$F_{tenengrad} = \sum_{height} \sum_{width} (S_x(x,y)^2 + S_y(x,y)^2) \tag{3.4}$$

where $S_x(x,y)$ and $S_y(x,y)$ are the convoluted images with Sobel operators.

4. *Brenner gradient*

The Brenner gradient algorithm calculates the square difference of each pixel between its two neighbours on the image used and then sums them together using the equation (3.5) [25].

$$F_{brenner} = \sum_{x,y} i(x+1,y) - i(x-1,y))^2 \tag{3.5}$$

with $|i(x+1), y) – i(x-1), y)| > \alpha$, where i (x, y) is the intensity at pixel (x, y), α is the threshold of the intensity difference.

5. *Energy of gradient*

Gradient energy of a single pixel demonstrates a certain difference between focus and unfocused image. The Energy of gradient focus measure function is computed as below equations (3.6) [27, 82].

$$F_{energy_{grad}} = \sum_{x}\sum_{y}(g_x^2 + g_y^2) \tag{3.6}$$

where $g_x(x,y) = g_i(x+1,y) - g_i(x,y)$

And $g_y(x,y) = g_i(x,y+1) - g_i(x,y)$

6. Spatial frequency

This is a modified version of the Energy of the gradient algorithm. In this algorithm, M x N image is used where M and N present the number of rows and columns respectively, along with gray value F (j, k) at position (j, k) [83]. The row and column frequencies are given in equation (3.7) and (3.8).

$$R_F = \sqrt{\frac{1}{MN}\sum_{J=0}^{M-1}\sum_{k=1}^{N-1}[F(j,k) - F(j,k-1)]^2} \tag{3.7}$$

And

$$C_F = \sqrt{\frac{1}{MN}\sum_{k=0}^{N-1}\sum_{J=1}^{M-1}[F(j,k) - F(j-1,k)]^2} \tag{3.8}$$

The total frequency is calculated as in equation (3.9).

$$Spatial\ frequency = \sqrt{{R_F}^2 + {C_F}^2} \tag{3.9}$$

Statistical Algorithms

Statistical-based algorithms takes the advantage of various image statistics, such as variance and correlation as the surface descriptor to measure the focus level. Generally, they are less sensitive to noise than derivative-based algorithms.

7. Gray level variance

While sharp images show high gray level variance, the images associated with blurring have low gray level variance. The simple standard definition of image variance is adopted into the equation given below. The equation (3.10) is used as a focus measure function using gray level variance [26].

$$\sigma^2 = \frac{1}{N^2} \sum_{x=1}^{N} \sum_{y=1}^{N} [I(x,y) - \mu]^2 \tag{3.10}$$

where μ is the mean of the gray level. The aim here is to maximize the σ^2 value since increases in σ^2 value will let us obtain higher gray level variance, therefore sharper image.

8. Tenengrad variance

Another focus measurement alternative to get the gradient information is to calculate the

gradient magnitude. In this direction, Tenengrad focus measure is calculated by the following equation (3.11) [84].

$$F_{tenengrad_{variance}} = \sum_{m}^{M}\sum_{n}^{N}[S(m,n) - S']^2 \quad (3.11)$$

$$and\ for\ S(m,n) > T$$

where T is the threshold, M and N are the height and width of an image, respectively. S' is the mean of magnitudes which is given by equation (3.12).

$$S' = \frac{1}{NM}\sum_{m}^{M}\sum_{n}^{N} S(m,n) \quad (3.12)$$

9. Variance of Laplacian

This method calculates the variance of the absolute values which provides a new focus measurement [84]. The equation (3.13) is demonstrated as:

$$F_{laplacian\_variance} = \sum_{m}^{M}\sum_{n}^{N}[|L(m,n)| - L']^2 \quad (3.13)$$

where L (m, n) is the convolution of the input image I (m, n) with the Laplacian operator L which can be approximated using the following mask:

$$L = \frac{1}{6}\begin{pmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{pmatrix} \tag{3.14}$$

and L' is the mean of absolute values given by

$$L' = \frac{1}{NM}\sum_{m}^{M}\sum_{n}^{N}|L(m,n)| \tag{3.15}$$

10. Vollath's correlation

This focus measure function is based on autocorrelation function and is computed as in equation (3.16) [80].

$$F_{vollath's} = (\sum_{i=1}^{M-1}\sum_{j=1}^{N} g(i,j).g(i+1,j)) - (\sum_{i=1}^{M-2}\sum_{j=1}^{N} g(i,j).g(i+2,j)) \tag{3.16}$$

The advantage of this method is that it does not depend on a threshold while some of the focus function does.

11. Image Curvature

Image curvature can be utilized as a focus measure since the curvature is higher in sharp images than blurred images [85]. In this method, the gray values are considered as a 3D surface (x, y, t(x, y). First of all, the quadratic equation $f(x,y) = kx + ly + mx^2 + ny^2$ is used to approximate the surface. After that the least square approximation technique is used

with t_0 and t_2 to calculate the coefficients (k, l, m, n).

$$t_0 = \begin{pmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{pmatrix} \quad and \quad t_2 = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{pmatrix} \tag{3.17}$$

Finally, the coefficients are combined and focus measure is obtained as in equation (3.18).

$$F_{curvature} = |k| + |l| + |m| + |n| \tag{3.18}$$

Histogram-based Algorithms

Histogram-based algorithms utilises the advantages of histograms to analyse the distribution and frequency of image intensities to obtain the in-focus image. This is an advantageous technique as the most in-focus images histogram demonstrates occurrences in many bins. However, for some instances in-focus images start with occurrences in a few bins and this should be considered to satisfy the application needs.

12. Histogram entropy

Let P (I) represents the frequency of the gray level I. The histogram entropy is defined by equation (3.19) [26].

$$E = -\sum_{I} P(I) \ln[P(I)] \qquad P(I) \neq 0 \tag{3.19}$$

'E' reaches its maximum value when all P (I) are equal, and minimum values when P (I) = 0 for all but one value of I. In this case, sharper images have smaller entropy than blurred

images. Therefore, minimizing the value of E will give the sharper image. This algorithm has also been evaluated and compared in previous researches [80, 86].

3.3 Auto-alignment of the Injection and Holding Pipette

Focalization provides key information in Z-axis for automatic levelling of the pipettes to the focus. In the following section, the focalization of the injection and holding pipettes and their alignment with each other will be given.

3.3.1 Autofocusing of the Injection and Holding Pipette

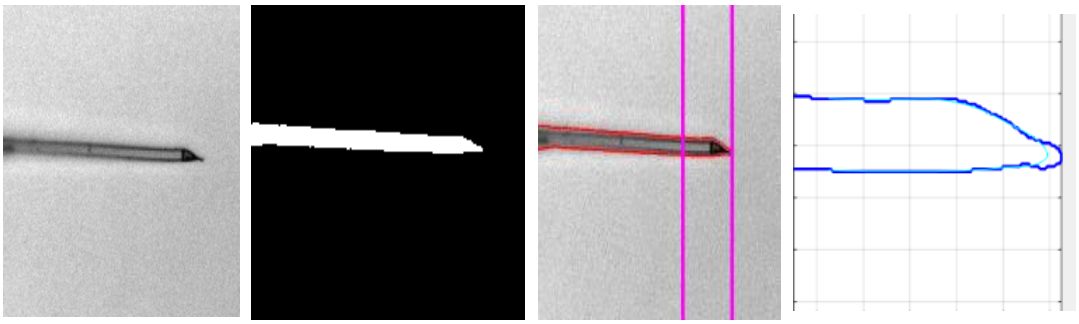
For the focalization of the injection pipette, a set of 18 images was captured for the injection pipette by scanning the movement of the pipette among Z-axis in a range between 0 to 34 micrometers using 2 micrometers increments. The initial position is obtained as the most blurred image captured below the focal plane, and the final destination position represents the most blurred image obtained above the focal plane. This means that any distance above or below this level brings the pipettes fully out of focus.

A similar image acquisition process is also conducted for the holding pipette focalization. Similarly, at the beginning of the operation, a set of 18 images was captured in a range between 0 to 170 micrometers by scanning with a 10 micrometers increment. Under the assumption of achieving 100% accurate focalization for the injection pipette, it should be possible to insert the injection pipette into the holding pipette without causing any bending when it goes through the holding pipette as long as the holding pipette is in alignment as well. The inner diameter of the holding and injection pipette is 30µ and 5.5 µ, respectively.

In this case, the tolerance of the developed auto-alignment process is 12.25 µ.

Digital images are generally exposed by noise. The level of noise is propagated during the image acquisition and transmission process. Therefore, pre-processing is often required to prevent the negative influence of the noise on developed algorithm performance [87]. A Gaussian filter as a part of pre-processing is implemented to reduce the noises on the captured image for the injection and holding pipette. Subsequently, the adaptive image threshold using first order statistics is used to obtain the binary image of the injection pipette. This process is called thresholding. Employing binary image procedure is not practical for the holding pipette since it has an inner reflection under the microscope particularly when the higher magnification is utilised. In this case, the active contour method is implemented on the holding pipette along with a defined focus window where all focus algorithms are applied [88].

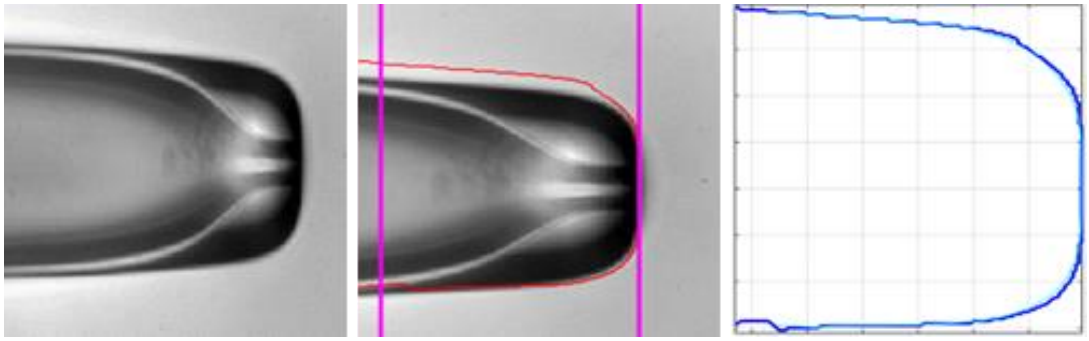
The boundaries of the injection pipette are detected by using the obtained binary image. However, the level of smoothness of the boundary demonstrates irregular data points at some fragment of the boundaries for both injection and holding pipettes. This can be seen in Figures 3-2 and 3-3. Hence, Kernel smoother, which is a statistical technique widely used in engineering, is utilized on the boundary of the injection and holding pipettes to obtain better data visualization. This is because the Kernel smoother represents the set of irregular data points on the boundary of pipettes as a smooth line. The extracted boundary is taken as a reference to plot two different lines which divide the boundary of the injection pipette at the ratio of 1:4 of the extracted boundary from the tip of the injection pipette. The mask is created within the region of interest where all focus algorithms are applied. Figure 3-2 represents the procedure of the pipette marginal detection and indicates the region of the interest extraction procedure for the injection pipette.



(a) **(b)** **(c)** **(d)**

Figure 3-2 Region of Interest extraction (a) Original injection pipette image (b) Final Binary Image (c) Image with a boundary at the ratio of 1:4 divided by lines (d) Extraction of Region of Interest

The mask is created in the focus window after implementing the active contour method on the holding pipette with a defined focused window. The region of interest for the holding pipette is extracted at the ratio of 3:4 from the tip of the holding pipette. The ratio selection to extract the region of interest for both injection and holding pipettes was selected to provide sufficient information to the FMA methods for their analysis. The implemented image processing technique is sufficient for holding pipette focalization due to its size. The obtained data variations on the recorded images using lower than 10 µm increments are not significantly different. Consequently, 10 µm is selected to be the optimum increment, and also this increment is within the tolerance. Figure 3-3 represents the region of interest extraction for analysis of the focus algorithms in-focus plane.



(a) **(b)** **(c)**

Figure 3-3 Region of Interest extraction (a) Original holding pipette image (b) Holding pipette contour detection and its division by lines at the ratio of 3:4 (c) Extraction of the region of interest

3.3.2 Curve Fitting and Fibonacci Search Algorithm

Curve fitting is the process of building a curve that has the best fit to the data points. In order to minimize the number of images captured to estimate the focal position of the injection and holding pipette, autofocusing is carried out by the mathematical function that accurately mimics the focus curve. It is essential to have sufficient data points near the focal plane with the purpose of having accurate auto-alignment of the pipettes. Global maxima and minima are defined as the largest and the smallest value of the mathematical function obtained from the curve fitting, respectively. The obtained focal position could be a global maximum or minimum depending on the FMA used. However, in this study, the focus curve is inverted to global maxima if the FMA produces a global minimum focal point. Here, the second-order Gaussian fit is implemented to the focus curves to obtain the focus function of the focus curve [81].

In order to search the focal plane of the injection and holding pipette precisely, a Fibonacci search algorithm was employed to find the global maximum point of the focus curve function.

The search interval was introduced to the Fibonacci search algorithm; it successively narrows the search interval until a satisfactory approximation for focus is achieved [36]. A Fibonacci search algorithm was an appropriate algorithm as long as the curve shows the unimodal property. Under the unimodality assumption of the focus curves for holding and injection pipettes, it has been proven that the Fibonacci search algorithm is the optimal algorithm [89].

3.3.3 Alignment of the Injection Pipette to the Holding Pipette

Before the alignment operation started, the pipettes were randomly inserted into the pipette holders. This caused no exact identification of the pipette tips initial position. Each position after initial detection had a corresponding focus value on the graph. Consequently, the proposed algorithm calculated the differences between the focal position and the current position values. These differences provide sufficient information to the developed algorithm to understand whether the tip of the pipettes is below or above the focal plane. Then the system determined the accurate corresponding positions and their differences. Once the pipettes were located to the focal plane, the coordinate of the pipette tips in X and Y-axis were calculated. The pixel coordinate of the tip of the pipettes was obtained by a set of image processing algorithms. This was achieved by obtaining the binary image of the pipettes using the adaptive image threshold and their blob detection once they were in focus. Blob detection is a computer vision method that detects the regions in the image by distinguishing them in terms of pixel mean intensity, perimeter and the area of the blob. Here the pipettes were distinguished from each other in terms of their blob pixel area. After detection of the pipettes in binary form, on the far right of the pixel of the injection pipette, and the far left pixel of the holding pipette, they were both detected and reported. Obtaining the binary form of the

holding pipette is only practical when it is in focus. Then, the pipettes are travelling only in the XY plane to grasp the oocyte and bring it to the predefined position and then conduct the injection when all are in focus.

The schematic 3D view of the injection and holding pipettes is given in Figure 3-4. As demonstrated, the levelling of the pipettes was conducted in Z-axis. However, the final grasping and injection are conducted in XY plane.

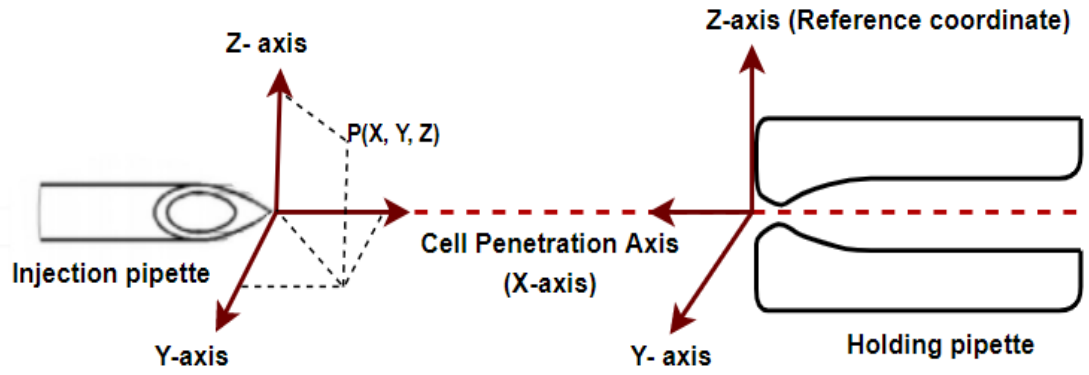


Figure 3-4 Schematic 3D view of the injection and holding pipette

3.4 **Results and Discussion**

This section presents and discusses the obtained results from auto-focusing algorithms for the auto- focalization procedure of the pipettes for the oocyte injection task. In this section, the results for the automatic alignment of both the holding and injection pipettes are presented separately. Then the operational confirmation will be experimentally demonstrated by inserting the injection pipette into the holding pipette. This fully illustrates that both pipettes are in absolute focus and within the same XY plane.

3.4.1 Focus Measurement Algorithm Results

Figure 3-5 shows the results of focus curves obtained from twelve focusing algorithms for auto-focusing of the injection pipette. In each graph, the normalized focus measure is demonstrated versus pipette position reported by the motorized stage.

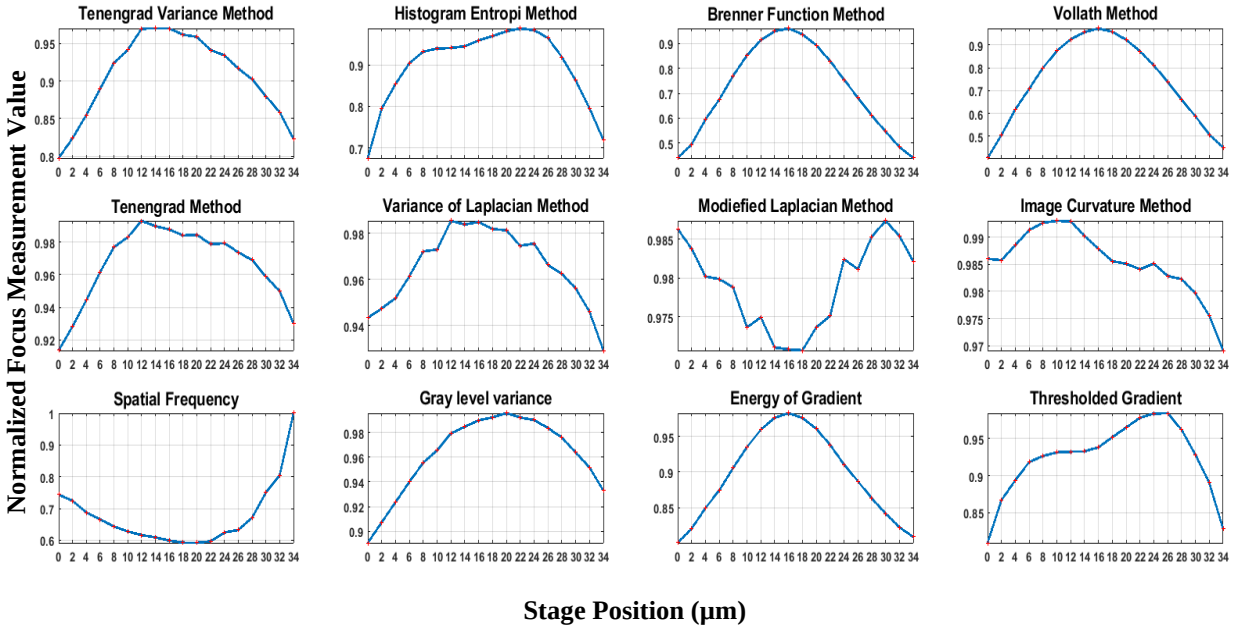


Figure 3-5 Auto-focusing of the Injection pipette based on different focus measurement algorithms

Based on the results in Figure 3-5, Energy of gradient, Vollath correlation, Gray level variance, and Brenner gradient methods produced reasonable focus curves as the rest of the methods produced false local maxima. Additionally, producing a unimodal curve is essential to enable the Fibonacci search algorithm to find the peak point of the curve accurately.

The Focalization of the holding pipette is also essential for the auto-alignment task for an oocyte injection procedure. Figure 3-6 presents the results of the focus curves obtained from focusing algorithms for the auto-focusing of the holding pipette. All focus values are normalized and plotted versus stage position.

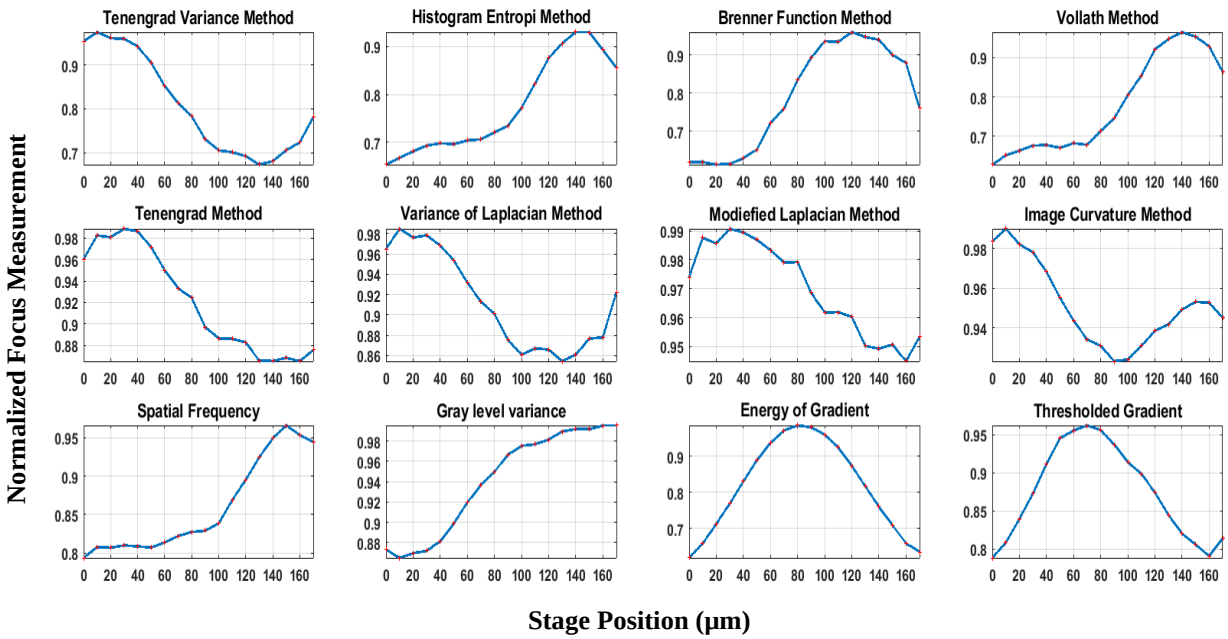


Figure 3-6 Auto-focusing of the Holding pipette based on different focus measurement algorithms

From the data in Figure 3-6, it is apparent that very few FMA were able to produce a sensible focus curve. This is due to the light reflection of the inner space of the holding pipette where there was not any such noise as in the injection pipette. Only the Energy of gradient method was able to produce a gradable focus curve.

The performance of the focus curves is evaluated based on focusing error, range, number of false maxima, and the width of the curve [90]. Error is measured from the differences between the determined best focus positions manually and the peak position of the curve obtained from the focus algorithm. A smaller error value gives a more accurate focus position of the injection and holding pipettes. Range criteria measure the distance between two adjacent local minima of the peak points in the focus curve. The larger range of focus curve will help the searching algorithm to find the global peak without being trapped between two minima points during the search. A number of false maxima describe any other peak points

of the curve apart from the global maximum. The width of the focus curve illustrates the sharpness of the peak as the sharper curve at the peak point has given a more accurate focus position to drive injection and holding pipette. In this study, 50% of the height of the curve is evaluated as the width of the curve.

Based on the results obtained from Figures 3-5 and 3-6, focus evaluation parameters are demonstrated in Table 3-1. The desired value for error, number of false maxima, and the width of the curve at %50 is 0, while the desired value for range criteria is 1. For the evaluation of the FMA, the difference between the desired criterion's value and the value of the focus curve for each criterion is taken as a coordinate. Maximum range criteria are subtracted from the range value of each FMA for calculation of the overall performance. Hence, the desired distance coordinate of the focus curve is [0 0 0 0] since 4 different criteria are used in this study. The overall score is calculated as a Euclidean distance for the focus curve obtained. If the given criteria are not applicable to the focus curve obtained, maximum focus criterion value is taken into account for Euclidean distance calculation. All values for each criterion are normalized and given within a range of between 0 and 1 so that each evaluation parameter maintains the equal weight for the overall performance comparison. The lower overall score presents the better performance of the focusing algorithm which is essential to have higher accuracy Z-axis focalization of the micropipettes.

Table 3-1 Injection and holding pipette's focus curves overall evaluations

EMA	Error		Range		Number of false maxima		The width of the curve at %50	
	Injection pipette	Holding pipette	Injection pipette	Holding pipette	Injection pipette	Holding pipette	Injection pipette	Holding pipette
Thresholded gradient	0.55	0.11	1	0.94	0.4	0.2	1	1
Energy of gradient	0	0	1	1	0	0	0.857	0
Gray-level variance	0.22	1	1	0.058	0	0.6	1	1
Modified Laplacian	0.77	0.55	0.058	0.058	0.8	1	1	1
Image Curvature	0.22	0.77	1	0.058	0.8	0.6	1	1
Tenengrad	0.22	0.22	1	0.176	0.6	0.6	0.857	1
Tenengrad variance	0.22	0.77	1	0.058	1	0.6	0.821	1
Variance of Laplacian	0.22	0.77	1	0.0176	0.8	0.8	0.785	1
Vollath's correlation	0	0.66	1	0.058	0	0.4	0.785	1
Brenner gradient	0	0.66	1	0.176	0	0.8	0.714	1
Spatial frequency	1	0.77	0.058	0.058	1	0.4	1	1
Histogram entropy	0.33	0.66	1	0.058	0.4	0.4	0.928	1

As it is pointed out in Table 3-1 for the evaluation of the injection pipette, Brenner gradient, Vollath correlation, and Energy of gradient demonstrated 100% accuracy while Gray-level variance illustrated 88% accuracy. The tolerance assumption for holding pipette alignment was conditioned with 100% accuracy of the injection pipette. Hence, 88% of accuracy has not been considered. Among the methods giving 100% accuracy, the only difference was obtained in the width of their curves at 50%. This is a significant parameter that needs to be

considered for FMA evaluations. This is because the lower width at 50% of the curve not only demonstrates how sharp the peak point of the curve is but also illustrates a higher focus measurement range. This is an essential feature to be obtained for the auto-alignment of the pipettes. A distinguishable focus measurement value between different focus levels at a small distance to each other helps the auto-alignment procedure remarkably, particularly around the focal plane. Brenner gradient provided the smallest width of the curve at 50% among all focus methods while providing 100% accuracy, full range with no false maxima as shown in Table 3-1. Hence, Brenner gradient demonstrated better performance than Vollath correlation and Energy of gradient method based on the calculation of the Euclidean method.

Figure 3-6 makes it evident that most of the FMA were exposed with a high noise for holding pipette focalization. This is due to the inner reflection of the holding pipette as stated previously. The 11 out of 12 FMA algorithms were failed with the exception of Energy of gradient method. The result of the Thresholded gradient method also produced a gradable curve. However, the focus measurement value at the last frame was always higher than the previous frame. This causes a disturbance in the Fibonacci search algorithm to find the peak point of the curve. The correct detection of the peak point of the curve of the Energy gradient method was 100%. This is in decent agreement with the tolerance of the holding pipette alignment. Overall, the Energy of gradient algorithm was found to be the most suitable for holding pipette focalizations for the purpose of the auto-alignment procedure.

The Energy of gradient and Brenner gradient methods are both derivative-based FMA. This type of algorithm tends to be affected by noise while providing better accuracy in comparison to statistical algorithms [90]. With the aid of a comprehensive evaluation of the FMA, decent accuracy for focalization of the pipettes was achieved whilst selected algorithms were not affected by any type of disturbance during their operation. The feature of the insensitivity of

the noise for the developed auto-alignment process while providing high accuracy will be well proved in the following section.

Computational time is also an essential parameter to be considered to increase the productivity of the developed algorithm. Computational time is highly dependent on the quality of the image and the number of pixels processed. For the analysis of the computational time, 8 sets of image collection for holding and injection pipette focalization were captured separately at random positions in the XYZ plane. Each set consists of 18 images for both pipettes and was run 15 times and then the averaged computation time was recorded. Recorded computation times include the entire computation to run the algorithm. Subsequently, the average computation time of 8 sets was recorded along with its standard deviation. Table 3-2 illustrates the average computational time per frame for performing the auto-alignment of the pipettes.

Table 3-2 Computational time analysis for auto-alignment of the pipettes in overall

FMA	Averaged Computation Time per Frame (s)	Standard Deviation
Brenner Gradient	1.159	0.0106
Energy of Gradient	1.717	0.0252

The alignment procedure calculation was conducted only once and the operational distance was calculated based on that particular frame which does not require additional image acquisition for any rest of the images within each trial.

3.4.2 Validation Test Results for ICSI Insertion to the Holding Pipette

In this section, the effectiveness of the developed image-based auto-alignment of the

Injection pipette is shown experimentally. It is motivated to conduct the insertion of the injection pipette within the holding pipette to illustrate the accuracy of levelling of the pipettes. This is a reasonable test to visually validate the levelling of micropipettes. This proved the operational accuracy of the algorithm in focusing the pipettes. Each pipette is randomly positioned at different levels in XYZ and programmed to be inserted. A total of 100 experiments have been conducted to prove this hypothesis. All of the insertions were conducted without any failure. The results proved the high reliability of the algorithm in different stages of pipette detection, operation calculation, and manipulation. Figures 3-7 and 3-8 demonstrate an example of the different positioning of pipettes and insertion experiments.

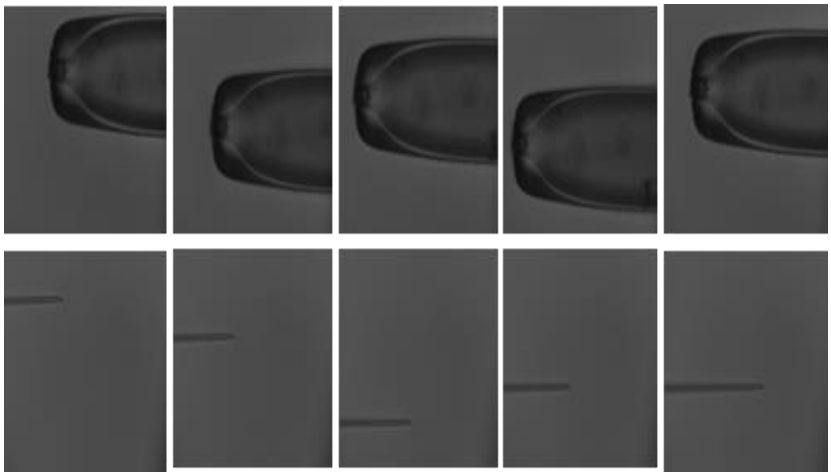


Figure 3-7 Visual demonstration of randomly placed injection and holding pipettes at various locations to check system performance in auto-alignment. Both pipettes are randomly moved in X and Y axes in the image plane, and then their focal position is randomly positioned on Z-axis within the visual field of the microscope.

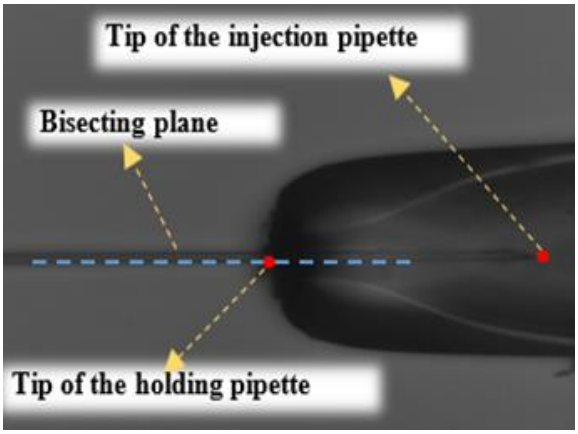


Figure 3-8 Auto-alignment of the pipettes for oocyte injection

The experimental results illustrate that the developed image-processing algorithm can perform the auto-alignment procedure successfully regardless of the pipettes' positions. The developed method can perform an auto-alignment even if the position of injection and holding pipette images are not within the recognition ranges.

In order to assess the trade-off between the sensitivity of the focalization of the pipettes to the noises and also their reliability at the same focal levels at different positions on the image plane, another set of experiments was carried out. In this experiment, the pipettes were randomly positioned at 8 different locations. The number of focus levels for each set is limited where the pipettes are in focus. For each position, 18 images were captured above and below the focal point with a 10 and 2-micrometers increment to each other for holding and injection pipettes, respectively. Figure 3-9 demonstrates the normalized focus value versus the set number.

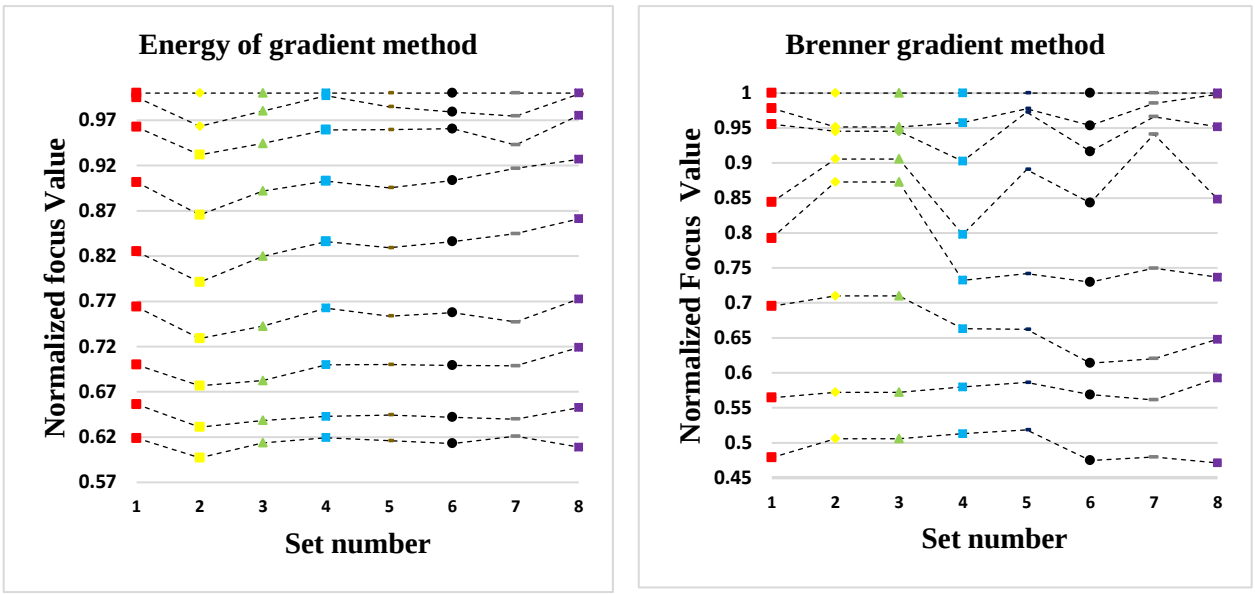


Figure 3-9 Focus value variations of the injection and holding pipette at different positions

In Figure 3-9, each colour represents one set of experiments and each set is evaluated within itself. Hence, there is no continuity between sets in terms of their data points. No intersections between the indicated dashed lines were observed for the pipettes as illustrated in Figure 3-9, which clearly demonstrates the reliability of the algorithm for an oocyte injection operation. This also demonstrates that the values are only dependent on the image set and are not a general measure for each level.

This study demonstrates the first reporting of auto-alignment of the pipettes used in ICSI with a combined computational time for pipette detection and alignment of 3s. This is in contrast to a much larger corresponding value of 14s reported in a similar study conducted for holding pipette [35], hence offering a considerable reduction in computational cost. Additionally, the results indicate 100% accuracy in both detection and alignment, enhanced reliability, and insensitivity to noise at different positions in the image plane.

3.5 Conclusion

This study offers a technique to automatically align the injection and holding pipette under microscopy imaging for ICSI operation. A ranking methodology was implemented on twelve different focus measure algorithms to determine the most appropriate focus algorithm in ICSI operation. The Fibonacci search algorithm has been implemented into the developed technique to drive the z-axis of the motorized stage to the focus position. It has been demonstrated that the Brenner gradient and Energy of gradient algorithms have superior performance for the injection and holding pipette focalization respectively. The results show that the auto-alignment tasks were achieved with high accuracy, reliability, and demonstrated insensitivity to noises at different positions in the image plane.

In the following chapter, a new vision-based sensor, operating at various magnifications, will be proposed to measure the vibration displacement of the tip of the injection pipette in three orthogonal axes.

Chapter 4: Vision-Based Sensor for Three-Dimensional Vibrational Motion Detection in Biological Cell Injection

4.1 Introduction

Intracytoplasmic sperm injection (ICSI) is an infertility treatment where a single sperm is immobilised and injected into the egg using a glass injection pipette. Minimising vibration in three orthogonal axes is essential to have precise injector motion and full control during the egg injection procedure. Vibration displacement sensing using physical sensors in ICSI operation is challenging since the sensor interfacing is not practically feasible. This study proposes a non-invasive technique to measure the three-dimensional vibrational motion of the injection pipette by a single microscope camera during egg injection. The contrast-limited adaptive histogram equalization (CLAHE) method and blob analyses technique were employed to measure the vibration displacement in axial and lateral axes, while the actual dimension of the focal axis was directly measured using the Brenner gradient algorithm as a focus measurement algorithm. The proposed algorithm operates between the magnifications range of 4× to 40× with a resolution of half a pixel. Experiments using the proposed vision-based algorithm were conducted to measure and verify the vibration displacement in axial and lateral axes at various magnifications. Additionally, the effect of injection speed on lateral vibration displacement was measured experimentally and was used to determine the values for egg deformation, force fluctuation, and penetration force. It has been demonstrated successfully that visual sensing has played a key role in identifying the limitation of the egg injection speed created by lateral vibration displacement of the injection pipette tip.

4.2 Materials and Methods

This section introduces the developed measurement technique for the three-dimensional vibration displacement of the injection pipette during egg injection. The developed measurement technique is divided into 2 main sections. The first part of the image-processing algorithm measures the tip displacement in lateral and axial directions while the other part of the algorithm calculates the displacement in the focal axis. A vibration term was used precisely to explain the mechanical oscillation of the tip of the micropipette. In the following sections, the system configuration will be explained.

4.2.1 System Configuration and Calibration

Figure 4-1 illustrates the schematic diagram of the three-dimensional egg injection setup. This setup is composed of an inverted microscope (Best-scope 2090), and a microscopy camera (BASLER Camera acA1300-200Um - Python1300), which can take up to 200 frames per second. The injection unit is firmly assembled on a 3-DOF motorized stage (Standa 8MT167) with an accuracy of 1 µm to move the injection pipette towards the desired positions using the BASLER camera as a vision sensor. The positional control of the motorized stage and image processing algorithms are hosted by a computer. Captured images are processed by an Intel® Core™ i5-6500 CPU @3.2 GHz (4 CPUs) host computer. The entire trajectory motion of the injection pipette during injection is captured and transferred to the host computer as a series of frames. Each frame is then analysed by the in-house developed image-processing algorithm in three orthogonal axes.

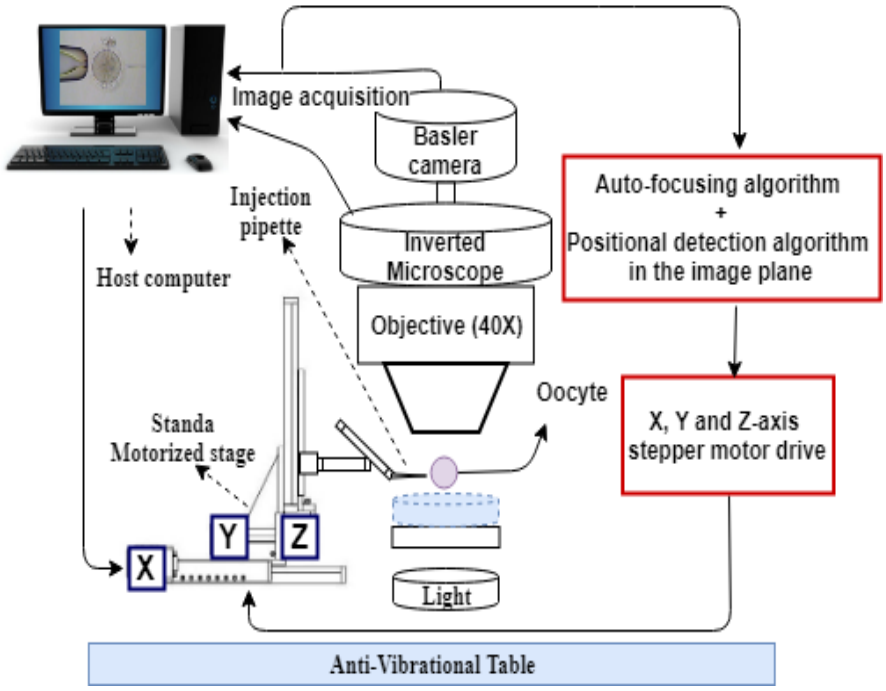


Figure 4-1 Schematic diagram of the cell-injection system setup.

The dimensional detail of the injection pipette used for the cell-penetration operation as well as its schematic 3D view is given in Figure 4-2.

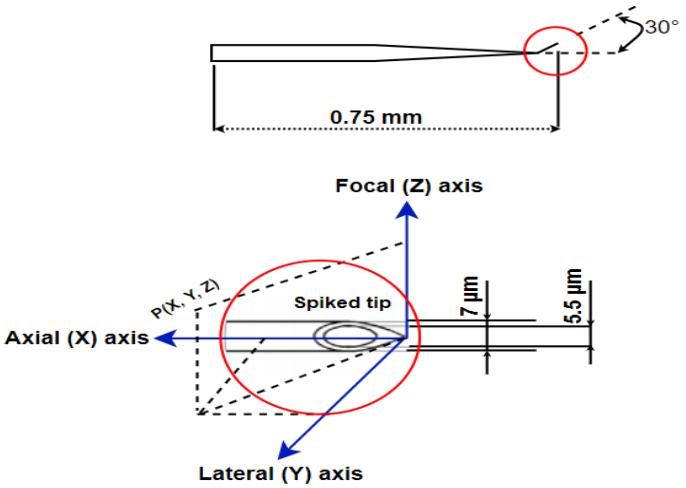


Figure 4-2. The dimensions of the injection pipette and its schematic 3D view.

The technique developed for the positional detection of the injection pipette in axial and lateral axes requires system calibration before the injection procedure starts. In this study, the outer diameter of the injection pipette is utilized as a dimensional reference to find the pixel size by using ImageJ software. The pixel size obtained will be employed into the developed Matlab algorithm as an input to report the positional changes in the X and Y axes versus time in real-world dimensions.

4.2.2 Detection of Positional Changes in Image Plane

In this section, the algorithm development procedure will be presented for extracting the 2-dimensional vibrational displacement information of the injection pipette tip. The steps involved in the development of the image-processing algorithm are pre-processing, filtering and thresholding, edge detection, and tip position measurement of the injection pipette. The details of these steps will be presented and discussed in the following part of the study.

Pre-Processing

The acquired images captured by the microscope camera require pre-processing for its further analysis. To increase the efficiency of the developed algorithm, the size of all the images acquired during egg injection was reduced to 50%. The captured image is in the eight-bit grayscale type where intensity level varies from 0 to 255. RGB, which stands for Red, Green and Blue, images are converted to grayscale within the developed algorithm. A grayscale image is then used for the positional detection of the ICSI pipette. Contrast-limited adaptive histogram equalization (CLAHE) can only be applied to the grayscale image. CLAHE is used to improve the contrast in images while avoiding amplifying the existing noise in the image.

This is a particularly significant feature to highlight the ICSI pipette for its further analysis as the developed measurement technique relies on the detection of the ICSI pipette accurately. In this study, the CLAHE method is implemented as pre-processing. In CLAHE, the input image is divided into small rectangular areas, which are called tiles, and then the local histogram for each tile is adjusted by enhancing their contrast [91]. The histogram is a graphical way of showing the frequency of occurrence at different color intensities in the image. Each bar in the histogram represents the total number of pixels at a particular intensity level changing from 0 to 255; 0 indicates the black in the histogram while 255 indicates white. In general, a balanced gray level is desirable in the histogram, which makes peaks in the middle and tapers off towards the edges. The comparison of gray-level distribution for the image of the ICSI procedure and processed image by the CLAHE method are shown in Figure 4-3.

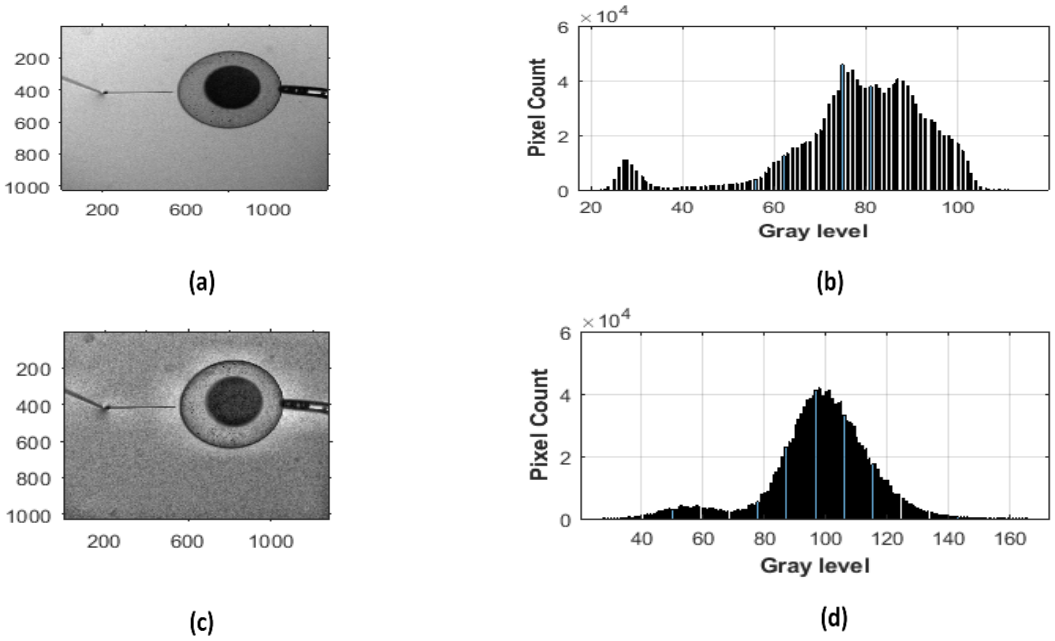


Figure 4-3 Enhancement of gray-scale intracytoplasmic sperm injection (ICSI) image by the contrast-limited adaptive histogram equalization (CLAHE) method and its histograms (a**) Original ICSI image; (**b**) histogram of the original ICSI image; (**c**) contrast-limited adaptive histogram equalization (CLAHE) image; (**d**) histogram of CLAHE image.**

The histogram of the original ICSI pipette image as demonstrated in Figure 4-3b illustrates that the image is overexposed. That means the input for image segmentation is not sufficient since the input image is not able to detect the edges in comparison to higher contrast input. On the other hand, the histogram of the image obtained by the CLAHE method illustrates a much-improved exposed image with evenly distributed gray-level bars and higher saturated tones. Hence, the CLAHE method is implemented successfully into the developed image processing algorithm as pre-processing to not lose any data on the image and lead to any failure of the injection pipette tip positional detection in axial and lateral axes.

Filtering and Thresholding

Filtering is a significant step to enhance the performance of the developed algorithm. The level of noise increases during the image acquisition and transmission process. Hence, it is essential to reduce the negative effect of the noises [87, 92]. A Gaussian filter is a linear filter that performs a weighted average of surrounding pixels chosen based on Gaussian function. Consequently, a Gaussian filter is utilized to reduce the noises on the ICSI image. The two-dimensional Gaussian filter used on ICSI image can be demonstrated in Equation (4.1).

$$G(x,y) = \frac{1}{\sqrt{2\pi\sigma}} \exp\left(-\frac{x^2+y^2}{2\sigma^2}\right) \tag{4.1}$$

where σ indicates the standard deviation of the distribution for the Gaussian filter and x and y illustrate the distance from the origin in the horizontal and vertical axes, respectively.

The grayscale image needs to be converted into the binary image as a part of the thresholding

procedure. The binary image is presented only as 0 and 1 for the background pixel and the highest intensity in the image plane. Adaptive thresholding is a common solution when the variations in illumination is a consideration. In this method, the threshold value is computed for every single pixel. This algorithm calculates the thresholding value for a small region, and then it applies different thresholding values for the other region of the image. Therefore adaptive thresholding provides better robustness changes in illumination [93]. Hence, the adaptive image threshold is employed to obtain the binary image of the ICSI process.

Binary images can be noisy, and any scattered pixels are required to be removed from the image for further processing. The Matlab *bwareafilt* function is utilized to eliminate the small pixels around the injection when holding the pipette and a zebrafish egg.

Edge Detection

The result of the binary image is then used to obtain the contour of the entire ICSI process through Canny edge detection [94]. In this method, the input image is converted into a set of curves to extract the remarkable features on the image. These remarkable features are the contour of the objects in the plane that it is aimed to detect. This detection relies on sudden intensity changes in the image. Hence, initially obtained curves data are transformed into the lines at the end of the procedure.

Measurement

Blob detection is a method to detect the regions in the image by distinguishing properties such as mean intensity, area, diameter, perimeter, centroid, etc. in comparison to surrounding regions. Each blob is labelled in order to make a measurement on the ICSI process. In our

case, in the ICSI process, the developed vision-based vibration-sensing technique detects two blobs, one is for injection pipette and one is for the combination of holding the pipette and the egg. The properties of these two blobs are extracted, which are mean intensity, area, perimeter, centroid, and diameter of each blob. Pixel intensity and the area of the blobs were sufficient to distinguish them from each other. The properties of the injection pipette, i.e. its area and pixel intensity, are defined within the developed Matlab algorithm to keep only the injection pipette on the image plane by isolating the rest of the objects which is the combination of holding pipette and the egg. These properties should be calibrated if the magnification of the microscope changes as it affects the value of the properties of the injection pipette. Afterward, the developed vision-based displacement sensor ignores any blobs which do match with the properties of the injection pipette and just keeps the injection pipette for further analysis. Subsequently, the far-right pixel of the binary image is detected and reported in the X and Y axis in pixels. The pixel size obtained from the calibration operation is employed to convert the pixel measurement into actual dimensions.

Figure 4-4 demonstrates the steps involved in the positional detection of the injection pipette tip in the image plane.

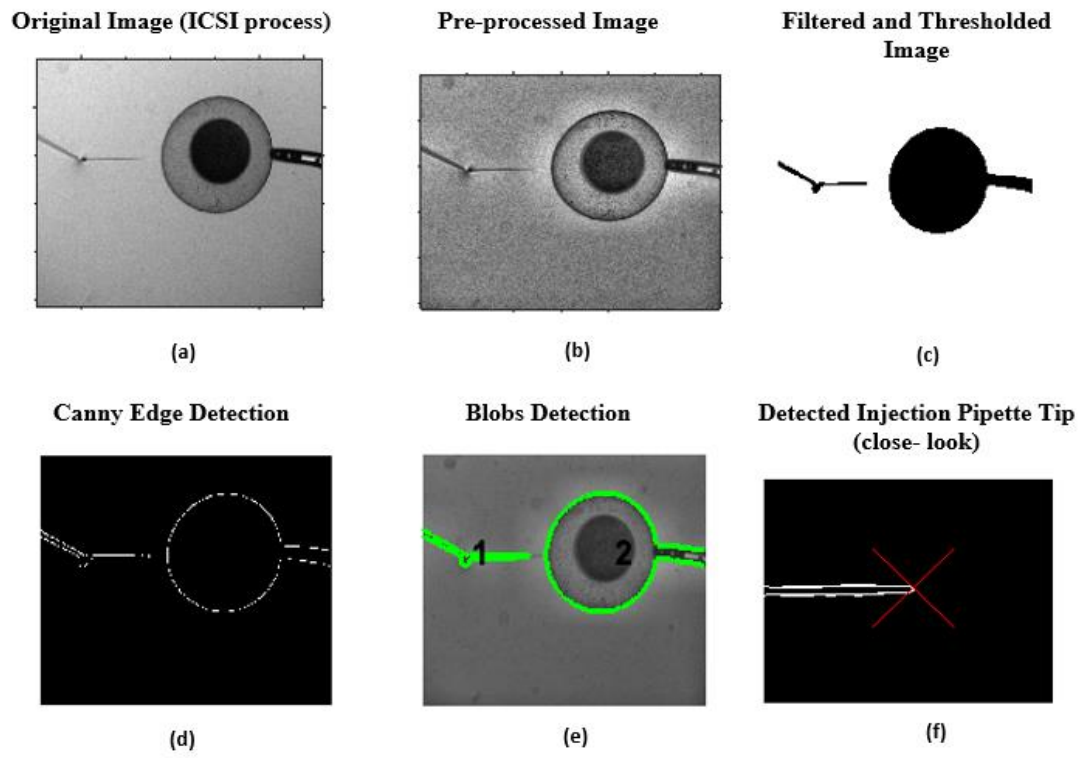


Figure 4-4 Steps in vision-based positional detection of the injection pipette: (a**) original ICSI image captured by the camera; (**b**) image after implementation of the CLAHE method; (**c**) ICSI image after application of Gaussian filter and adaptive thresholding; (**d**) extraction of the contour information of the ICSI image using canny edge detection algorithm; (**e**) recognition of the egg and injection pipette and distinguishing them based on their blobs feature; (**f**) detection of the injection pipette tip indicated by the red cross.**

4.2.3 The Detection of Focalization Position Changes

The technique for the focalization of the injection pipette tip has been conducted previously by our research group and the details can be found in [57]. In this study, the focalization procedure of the injection pipette tip is utilized to find the positional changes in Z-axis at the time of cell injection. In order to obtain the focal position of an object under a microscope, focus algorithms are widely used [26, 27]. These algorithms provide information regarding the degree of focus for the number of images taken by varying the focus lens position. The Brenner gradient algorithm was selected as the focusing algorithm after the evaluation of the

12 different focus algorithms examined for this application. In the Brenner gradient algorithm, the square difference of each pixel between its two neighbours are calculated and then they are summed using Equation (3.5)[25].

To locate the injection pipette tip to the focal plane under a microscope, the movement of the injection pipette is scanned in Z-axis in a range between 0 to 34 μm with an increment of 2 μm. Hence, a set of 18 images was attained. The initial position presents the most blurred image under the focal plane, while the final destination position demonstrates the most blurred one above the focal plane.

Since the level of noise increases during the image acquisition process, it requires pre-processing to enhance the performance of the algorithm. The noise on the images is eliminated by using a Gaussian filter as a part of pre-processing. Then, the adaptive image threshold using first-order statistics is utilized and the binary image of the injection pipette was obtained. Obtaining a binary image of the injection pipette without any scattered points is essential since boundary extraction of the injection pipette is fully dependant on the obtained binary image. However, as it is visible in the obtained binary image of the injection pipette in Figure 3-2, at some fraction of the boundary of the injection pipette, irregular data points result. In order to prevent any irregular data points, particularly at the tip of the injection pipette, a kernel smoother was employed to set the irregular data points on the boundary of the injection pipette as a smooth line.

The extracted boundary is divided by two lines at the ratio of 1:4 starting from the tip of the injection pipette. The created mask in this area is considered as a region of interest for the image-processing algorithm. Finally, the Brenner gradient algorithm is implemented in the region of interest extracted. Figure 3-2 has already illustrated the marginal detection of the injection pipette image and its region of the interest extraction procedure.

Figure 4-5 demonstrates the pattern of the normalized focus measurement values versus stage position. As it is shown, the injection pipette presents a unimodal curve that enables Brenner gradient algorithm to be utilized. Also, the maximum value of the normalized focus value corresponds to the sharpest image where the stage position is in the focal plane. Subsequently, the second-order Gaussian fit is implemented to the curves obtained from the Brenner gradient method for the injection pipette [81]. In order to search the global maximum of the curves precisely, a Fibonacci search algorithm is employed to the procedure into the defined focus range [36]. In this method, the range of search is defined for the algorithm and the focus range was narrowed successively until the focal plane is obtained. It has been proved that the Fibonacci search algorithm was found to be an optimal search algorithm under the assumption of the unimodal property of the focus curve [81]. Figure 4-5 clearly demonstrates that there is a peak where the tip of the injection pipette is entirely in focus.

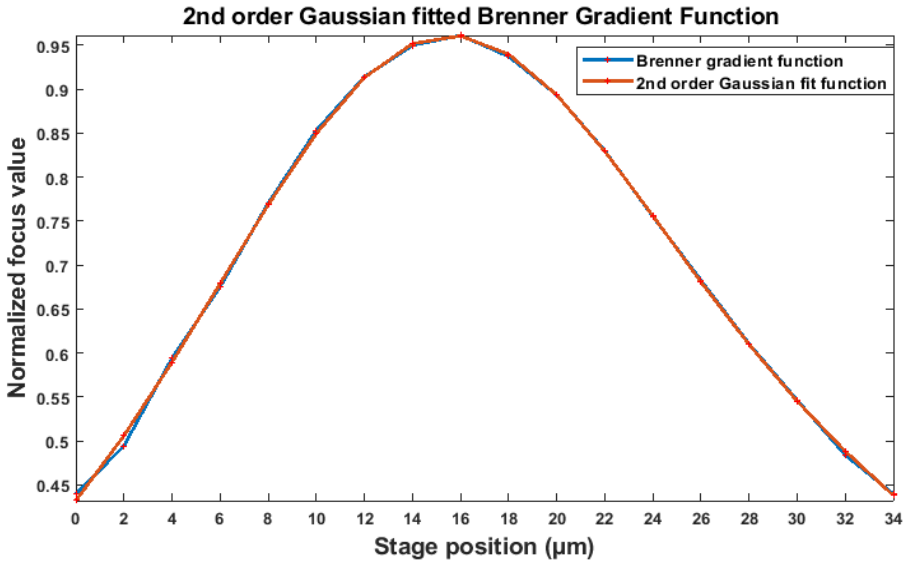


Figure 4-5 Representation of the focus curve obtained from the Brenner gradient method.

In Figure 4-5, each stage position has got a corresponding normalized focus value. Initially, a set of images saved at the time of egg injection are analysed based on its normalized focus value. Each frame is labelled with a normalized focus value obtained from the Brenner gradient algorithm. The reliability of the Brenner gradient method at the same focal level at different XY plane was previously verified for this particular application in our previous publication [57]. Since the equation of the curve was obtained after the implementation of the Gaussian fit method, each focus value will have the corresponding stage position between 0 to 34 µm. As the above and below the focal plane is identical, by being independent to the direction selection, the right-hand side of the curve was taken into account for the implementation of the normalized focus values to be implemented into the focus function. Therefore, these normalized focus values are converted into the real dimension by the algorithm developed. Finally, each position of the frame in the Z-axis is subtracted from the focal point of the focus curve. The differences are reported as displacement in Z-axis versus time. This technique is capable of measuring the vibration displacement while the injection pipette is in the field of view under the microscope. Based on the assessment of the Brenner gradient algorithm in our previous work, this algorithm is reliable and efficient to report the same pipette position in Z-axis by being independent to the X and Y axes in the field of ICSI. Figure 4-6 demonstrates the block diagram of the developed algorithms to calculate the vibration displacement of the injection pipette tip in axial, lateral, and focal axes.

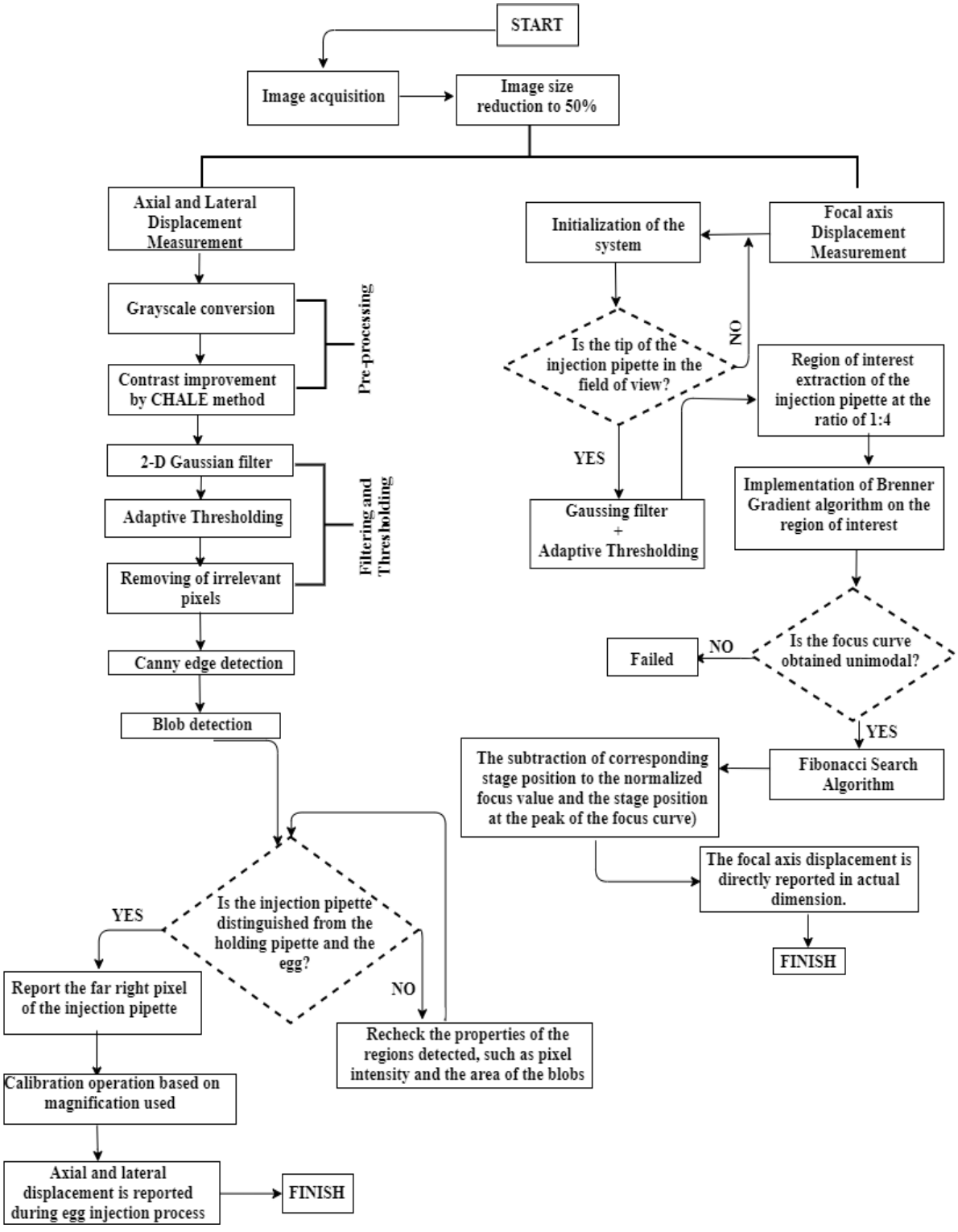


Figure 4-6 Block diagram of the vibration displacement sensing procedure by the developed vision algorithms in 3 dimensions.

4.3 Results and Discussion

In the previous sections, the method of a three-dimensional vision-based vibration displacement measurement method was demonstrated. Here, the algorithm is examined and the results are presented in the following section. For evaluating the functionality of the algorithm, the vibration was introduced randomly to the system to demonstrate the effectiveness of the algorithm developed. Subsequently, the motion of the injection pipette was analysed by the algorithm developed. The results obtained from the algorithm have then been validated by using ImageJ software. This software lets the user select the desired pixel value in the XY plane manually. Hence, the tip coordinate of the injection pipette in the XY plane was compared with the results obtained manually from ImageJ software. A total number of 100 images were captured sequentially during the egg injection. These image acquisition procedures were conducted separately at four different magnifications, as shown in Figure 4-7.

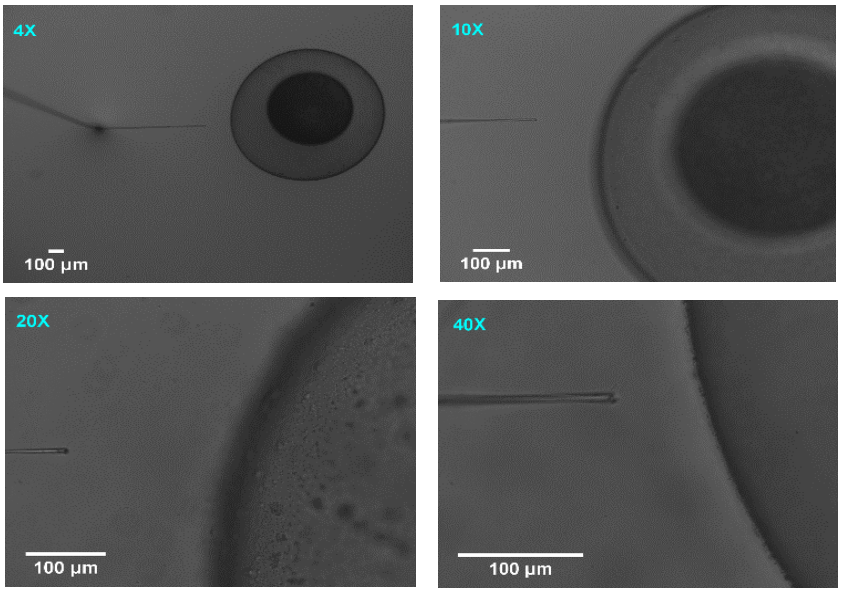


Figure 4-7 Egg-injection procedure at different magnifications.

Before the experiment was conducted, the initial position of the injection pipette tip was recorded in each magnification. This value is subtracted from each reported coordinates from the vision algorithm during egg injection which took 0.5 seconds for all levels of magnification. Then, the magnitude of vibration displacement was measured and reported in axial and lateral axes by the developed vision algorithm.

To validate the results obtained from the developed algorithm, the ImageJ software is employed. This is a manual image processing program which is developed to analyse and process the images. For this reason, the images were focused on the pixel levels, and the final pixel which demonstrated the tip of the pipette was manually selected. Figure 4-8 illustrates the original and focused zoom used in ImageJ. Figure 4-8 illustrates that different magnifications provide different focused zoom. The yellow square shows the area of interest for manual selection and the red cross on the images shows the point which was selected for the tip of the pipette. By increasing the magnification, the possibility of a manual selection of the most appropriate pixel, which represents the tip of the pipette, increases.

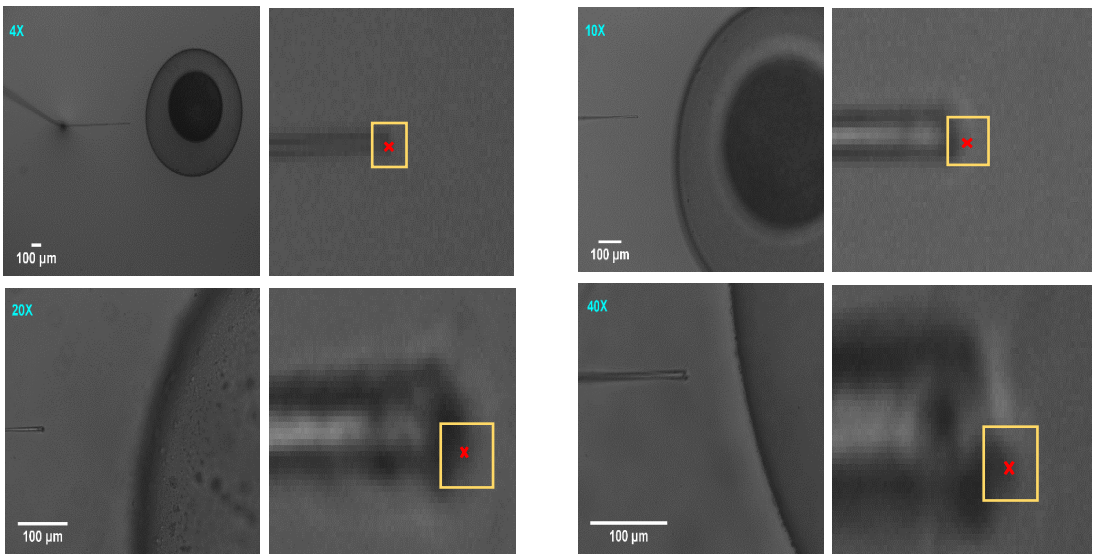


Figure 4-8 Pixel selection by ImageJ software.

This method of comparison demonstrates the accuracy of the developed algorithm compared to manual selection. It is noted that there may be some human errors contributing to these results, however, integrating other types of physical sensors would not be practically possible. On the other hand, this comparison has higher validity due to matching the results based on an image which eliminates other environmental contributing factors. Figure 4-9 illustrates the differences in results obtained manually and by the developed algorithm both laterally and axially.

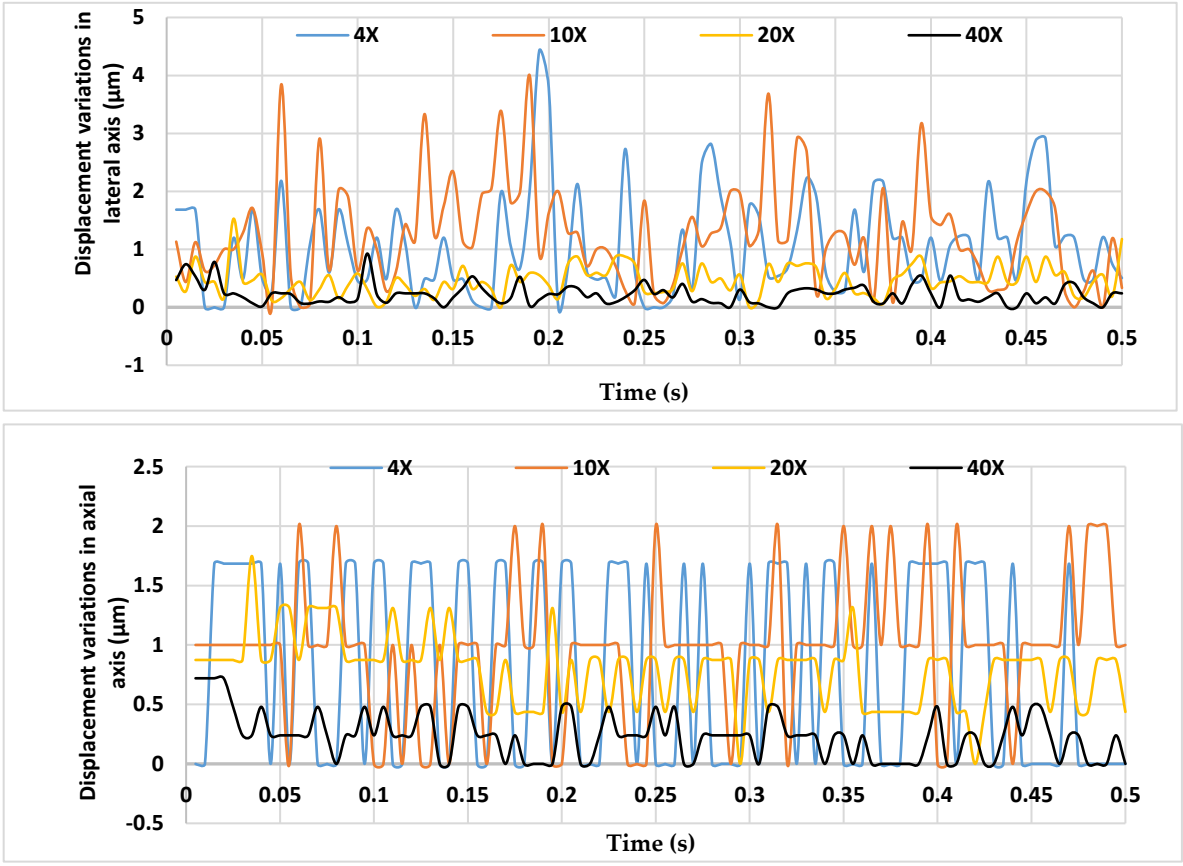


Figure 4-9 Displacement variations between developed vision sensor and Image J software in axial and lateral axes.

The detail of the average axial and lateral vibration displacement variations at each

magnification is shown in Table 4-1. The level of standard deviation is high as expected due to the vibration introduced.

Table 4-1 The average variation between Image J software and vision algorithm.

Level of Magnification	Axial Axis (μm) and Its Standard Deviation (SD)(μm)		Lateral Axis (μm) and Its Standard Deviation (SD)(μm)	
4×	0.808	0.841	1.05	0.881
10×	0.970	0.576	1.25	0.890
20×	0.790	0.302	0.46	0.264
40×	0.235	0.195	0.22	0.173

As the comparison results demonstrate in Table 4-1, the variations between the developed algorithm and the ImageJ software were both decreased when the level of magnification increased. This is due to the changes in pixel sizes. Considering the total tip size which is 7 μm, the area of interest provides only 4 allowed pixels for manual selection in 4×, however, the same area of interest provides 42 allowed pixels in 40× which hugely decreases the potential errors of manual selection. This can be confirmed in Figure 4-9 as well which shows the variations decrease by 71% and 79% from 4× to 40× for axial and lateral vibration displacement, respectively. Overall, the developed vision algorithm demonstrated consistency and its reliability with the results obtained by ImageJ software.

The displacement of the injection pipette tip in the Z-axis was obtained with the utilization of the Brenner gradient method. A number of 100 images were captured sequentially during injection at different magnifications. Later, their normalized focus value was introduced to the focus function obtained from the Brenner algorithm. The corresponding stage position for each focus value was obtained. As shown in Figure 4-5, the stage position at the peak of the curve is known. Therefore, the stage position at each frame was subtracted from the

position where the tip position of the injection pipette has a peak in the curve. The displacement variations in Z-axis versus time were plotted at different magnifications in Figure 4-10.

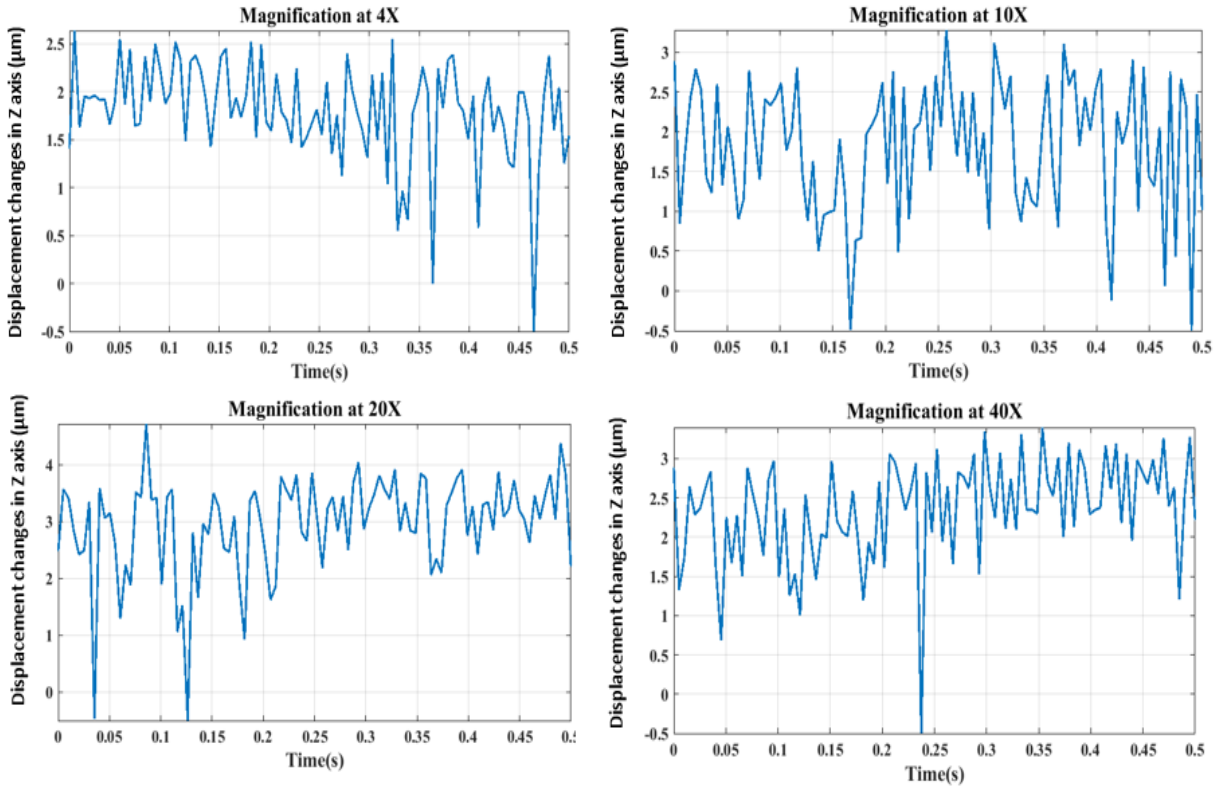


Figure 4-10 Injection tip position changes in Z-axis using the Brenner gradient method

As it is shown in Figure 4-10, the vibration-induced injection pipette position has changed in the Z-axis by its focus value at different magnifications. As previously stated, the right-hand side of the focus curve was used for the implementation of the normalized focus values which were implemented into the focus function. Hence, the resultant vibration displacement in the focal axis is only positive. The displacement of the injection pipette is directly obtained as shown in Figure 4-10 using the Brenner gradient algorithm as a focus measurement.

This study assists in the need for the vibration-sensing process accurately in three orthogonal

axes in injection procedures. Currently, the effect of vibration at the tip of the injection pipette in the ICSI procedure is not studied in the literature. This study will try to study the viability of the vibration analysis to realise the role of vibration displacement in three dimensions induced by any sort of internal and external dynamic parameters and whether that may contribute to the success rate of the ICSI operation. For this purpose, another set of experiments was conducted to demonstrate the key role of lateral vibration displacement in ICSI.

As the algorithm had been validated, it was then used to evaluate the effect of egg-injection parameters such as speed on the vibration. In the previous publication, the authors demonstrated the variations in forces and deformation creation during egg injection [45]. In this research, the potential contribution of the demonstrated factors is evaluated in vibration creation which may be a cause of egg damage during the penetration. Injection speed was selected as the main contributing factor in potential damage to the cell. So, the various speeds were considered for the vibrational measurements. Then the results obtained were correlated to the previous results. The total travel distance was assigned to be 100 μm for each injection speed. Figure 4-11 illustrates the lateral vibrational displacement results at various speeds.

As illustrated in Figure 4-11, the injector starts with some vibrations, but it moves steadily. However, at the end of the motion procedure, the lateral vibration displacements dramatically increased. This moment is considered as the injection time. This is due to the acceleration and deceleration maneuver of the injector. On the other hand, the vibration displacements were increased by increasing the speed as expected in the hypothesis.

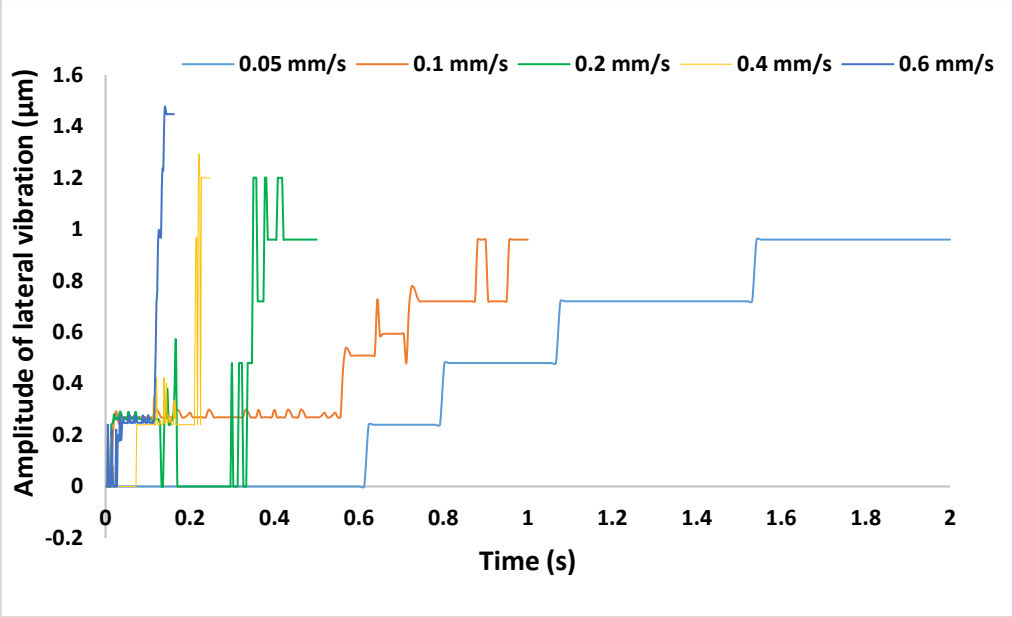
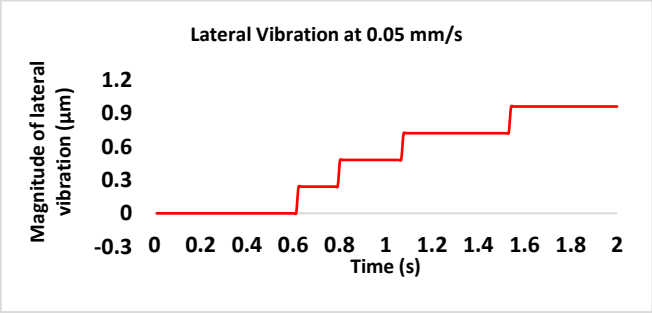
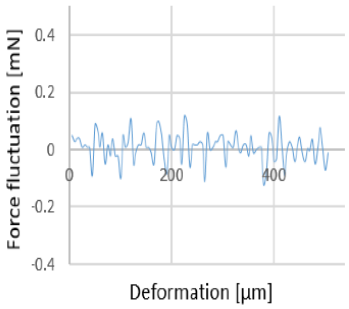
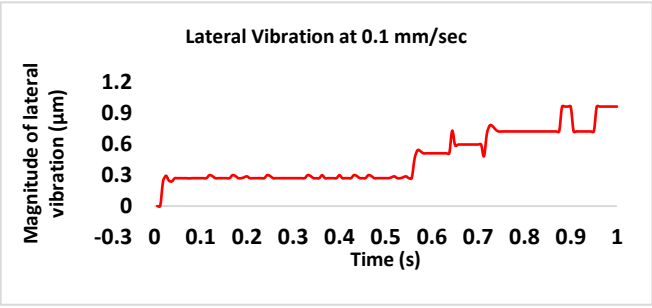


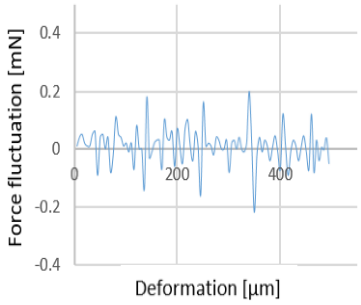
Figure 4-11 The effect of egg-injection speed on vibration creation

Figure 4-12 illustrates the direct comparison of the lateral vibration displacement, force, and deformation at each speed.









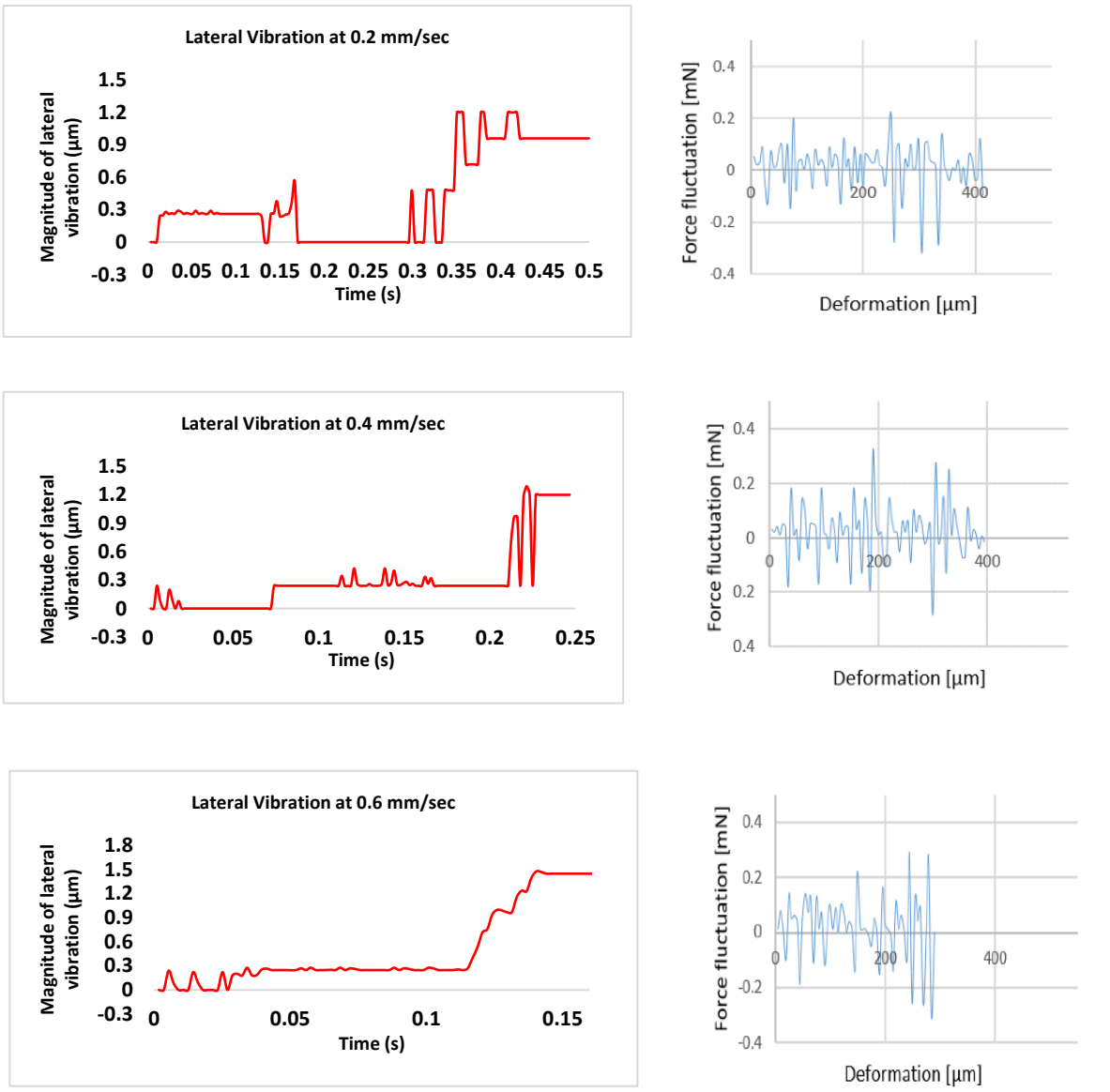


Figure 4-12 Lateral vibration displacement, force fluctuation, and egg deformation at various injection speeds.

Figure 4-13 illustrates the comparison between the force amplitude, penetration force, induced deformation, and also increases in vibration. The maximum force amplitude ratio (MFAR) is defined as the ratio of the maximum force amplitude resultant during injection to the maximum penetration force in percentage. This demonstrates the force fluctuations during the injection. 0.05 mm/s injection speed is considered as the reference to calculate the

increase in vibration displacement. As shown in Figure 4-13, the vibration displacement did not increase significantly at a lower speed, however, it increased significantly after 0.4 mm/s. The results demonstrate that force fluctuation increases while vibration displacements increases, however, it causes less deformation to the cell. The sudden force fluctuation was increased by 58% while lateral vibration displacement increased by up to 54% when the injection speed increased from 0.05 mm/s to 0.6 mm/s. The rate of increase for lateral vibration displacement and force fluctuations is consistent which demonstrates the reliability of the developed vision algorithm for vibration sensing. In the meantime, the egg deformation was reduced to approximately ~34%. Although the egg deformation was reduced, it has been ascertained that high lateral vibration displacement preventing egg injection should be conducted under stable conditions. The egg injection operation is challenging after 0.6 mm/s due to sudden increased lateral vibration displacement (up to 54%) which causes damage to the egg. Lateral vibration displacement is a decisive dynamic parameter in the egg-injection operation which limits the injection speed although increases in speed decrease the egg deformation in the general trend. This demonstrates that injection operation requires speed optimization considering the induced vibration displacement as well as egg deformation. Overall, the correlation between the lateral vibration displacement, force fluctuation, egg deformation, and injection speed has demonstrated the critical role of visual vibration sensing in the ICSI operation.

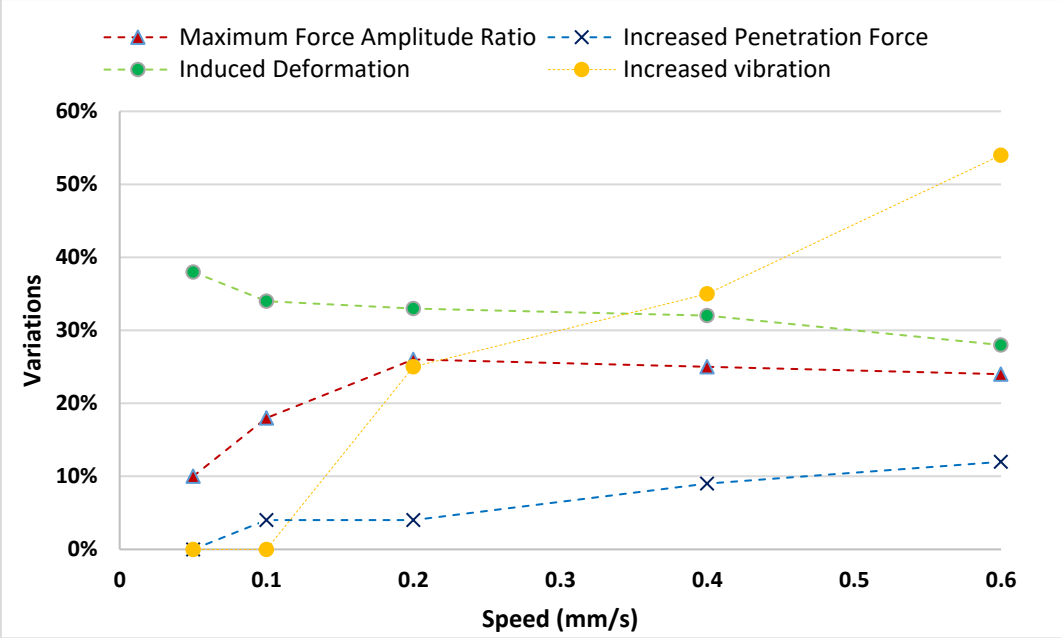


Figure 4-13 The impact of injection speed on lateral vibration displacement, penetration force, egg deformation, and force amplitude speeds.

4.4 Conclusions

In this chapter, the vision-based three-dimensional vibrational motion detection technique was proposed for the application of biological cell injection. The advantage of this technique is that it uses only one microscope camera to obtain vibrational displacement information in three orthogonal axes. The developed vibration displacement measurement algorithm demonstrated consistency with the results obtained by ImageJ software. The vibration displacement occurred in the focal axis was also measured by the Brenner gradient algorithm. To show the significance of the developed vision algorithm, the effect of egg injection speed on lateral vibration displacement creation was investigated. It has been shown that the increase in injection speed resulted in an increase in the lateral vibration displacement at the tip of the injection pipette, which is up to 54% while the egg deformation has decreased to approximately 34%. It has further been demonstrated that the feasibility of egg injection is

limited after the injection speed of 0.6 mm/s. The results show that lateral vibration displacement is a decisive parameter that needs to be measured in the egg injection task as it limits the injection speed. The proposed three-dimensional vibration sensor enhances the vibrational displacement measurement within the microinjections, specifically for ICSI.

The next chapter will discuss an automated zebrafish embryo positioning system to the predefined position within the microscope's field of view in real-time using deep learning.

Chapter 5: Real-time Deep Learning-based Image Recognition for Applications in Automated Positioning and Injection of Biological Cells

5.1 Introduction

Biological cell injection is an effective method in which a foreign material is directly introduced into a biological cell. Since human involvement reduces the success rate of the biological microinjection procedure, an extensive research effort has been made towards its automation. The accurate positioning of a randomly placed biological cell in the microscope's field of view is a prerequisite for any automated injection procedure. Vision is the primary source for visual servoing in microinjection applications. For this reason, a visual sensing system is required to recognise, calculate, and manipulate the cell to the desired position. In this study, eight different pretrained neural networks were analysed and used as a backbone for the YOLOv2 object detection method, and the optimal network was evaluated based on mean Intersection over Union (IoU) accuracy, Average Precision (AP) at different thresholds, and frame rate (fps) in our dataset. Ten different sets of experiments were conducted to examine the algorithm by verifying the zebrafish embryo gradual presence within the field of view to bring the zebrafish embryo to the predefined position. Later, the generalization of the proposed solution was verified in a different dataset from the real microinjection setup.

5.2 Materials and Methods

5.2.1 System Configuration

Figure 5-1 represents the system configuration which consists of the motorized manipulator (STANDA 8MT167) to control the position of the injection pipette. The holding pipette is attached firmly to the micropipette holder. The vision system in the setup configuration is a combination of a microscopy camera (BASLER Camera acA1300-200um - Python1300), an inverted microscope (Bestscope 2090), objective, and lighting. The camera is directly connected to the host computer and run by a MATLAB software (2019b) interface. Hence, the sequence of captured images is directly provided to neural networks as input for detection and classification purposes in real-time. The 4x objective at the maximum lighting condition was used throughout the study as this was the optimal condition to experiment with zebrafish. The motorized XY stage (SUTTER MP-78) with a controller (MP-285) was employed to hold and position the petri dish to automatically bring the zebrafish embryos into the FOV of the microscope. All process was conducted by 3.2 GHz host computer.

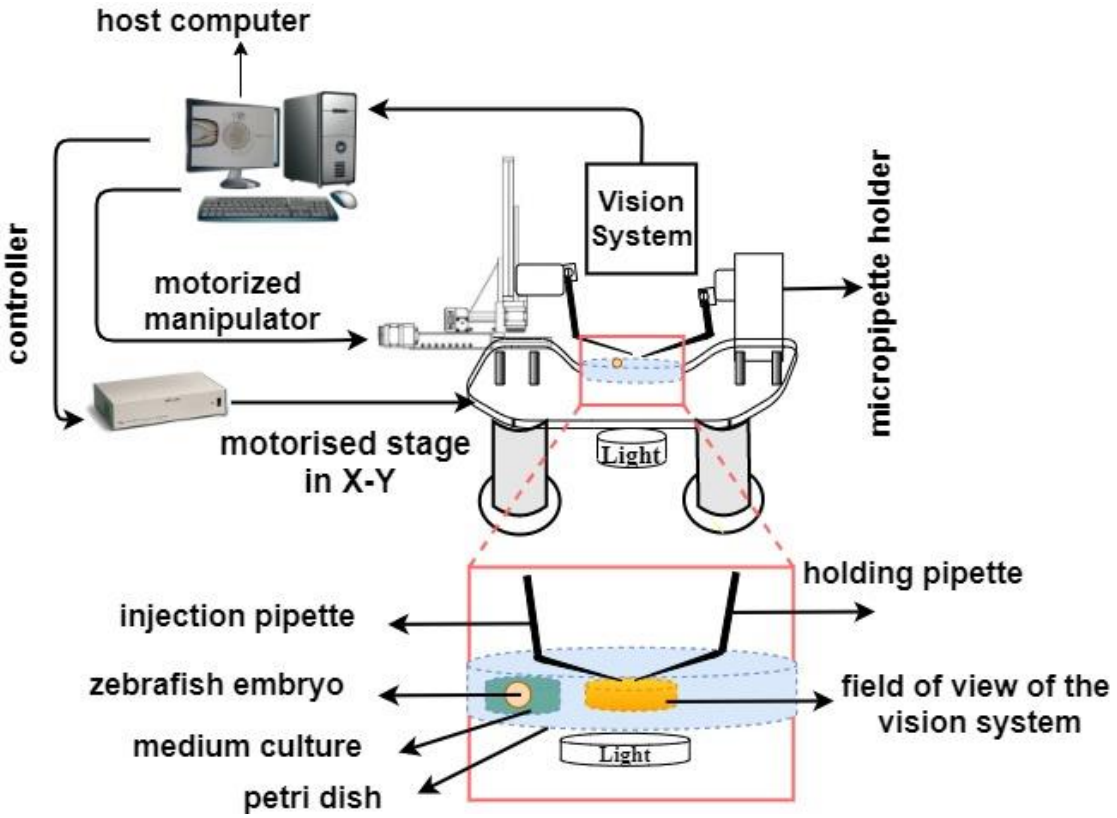
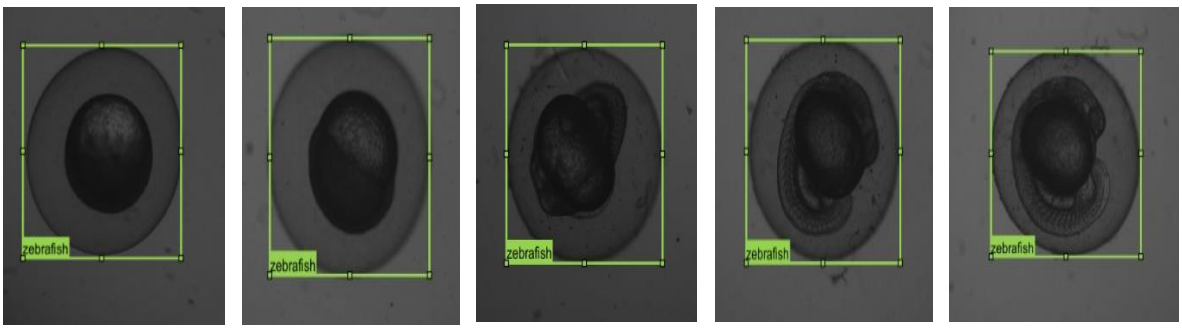


Figure 5-1 Schematic view of the system configuration

5.2.2 Data Collection, Labelling, and Augmentation Process

In this study, the zebrafish embryo was used as an experimental model. Zebrafish embryo has a similar genetic structure to humans in terms of mechanical response and is easy to retrieve to be used in research as the size of the zebrafish embryo is about 0.7 mm. The zebrafish embryos were treated with standard embryo preparation procedures [95]. The in-house created zebrafish embryo dataset includes 700 images which include 60% of training and 40% of test datasets where the number of images used for training and testing purposes is 420 and 280 respectively. This ratio is sufficiently practical to show the performance of the trained neural network as well as its generalization considering the size of our dataset.

The test set is used for the evaluation of the trained neural network model. The data was shuffled for training purposes using *randperm*, a MATLAB built-in function, which generates a random permutation of the data provided. The dataset collection was conducted during the first 24 hours after the immediate collection of fresh zebrafish. In this case, the zebrafish images were captured at different developmental stages to increase the learning capability of the neural network at any developmental stage of the zebrafish. During the data collection, the illumination intensity was also varied. The zebrafish embryo images were then labelled as it is shown in Figure 5-2 and used as an input to train the neural network models within this study. MATLAB Ground Truth Labeller application was used for labelling procedure. The sample zebrafish embryo dataset and image labelling are shown in Figure 5-2.



(a) **(b)** **(c)** **(d)** **(e)**

Figure 5-2 The sample zebrafish embryo datasets labelling and various stages of the zebrafish embryo development to enhance the datasets. Developmental stages are illustrated as follows; (a) Immediate after collection. (b) after 6 h. (c) after 12 h. (d) after 18 h. (e) after 24 h.

Data augmentation was implemented to enlarge the dataset and increase the learning capability of the neural network. This eliminates the overfitting on the dataset during the training procedure, as well as increase the performance of the neural network [96]. In our study, contrast, saturation, and the brightness of the images were altered slightly using

MATLAB built-in function *jitterColorHSV* to model various attributes of visual perception. The scale factors selected for contrast, saturation, and brightness are 0.2, 0.1, and 0.2, respectively. Additionally, Gaussian noise was introduced to simulate a realistic environment for the zebrafish embryo dataset. Finally, random reflection transformation that flips the input image with 50% probability in X and Y axes, scaling transformation that resizes the input image randomly with a scale factor range between [1, 3] in X and Y axes, randomized rotation transformation in the range of [0, 360], and translation transformation that shifts the image in X and Y-axis in the image plane with a selected distance between [-10, 10] in pixels were employed to simulate various operational conditions. Implementation of more than one factors has been examined in our developed algorithm to validate the reliability of the proposed system. Online augmentation, which is also called augmentation on the fly, approach was implemented to generate the augmented samples. In this computer storage-friendly method, the required transformations are applied within the mini-batches which then have been fed to the neural network models during the training procedure. 420 images were augmented at each epoch. As the number of 30 epochs were used during the training, 12,600 augmented sample images were generated from the training dataset and all used during the training procedure. Sample augmented images from the dataset are shown in Figure 5-3.

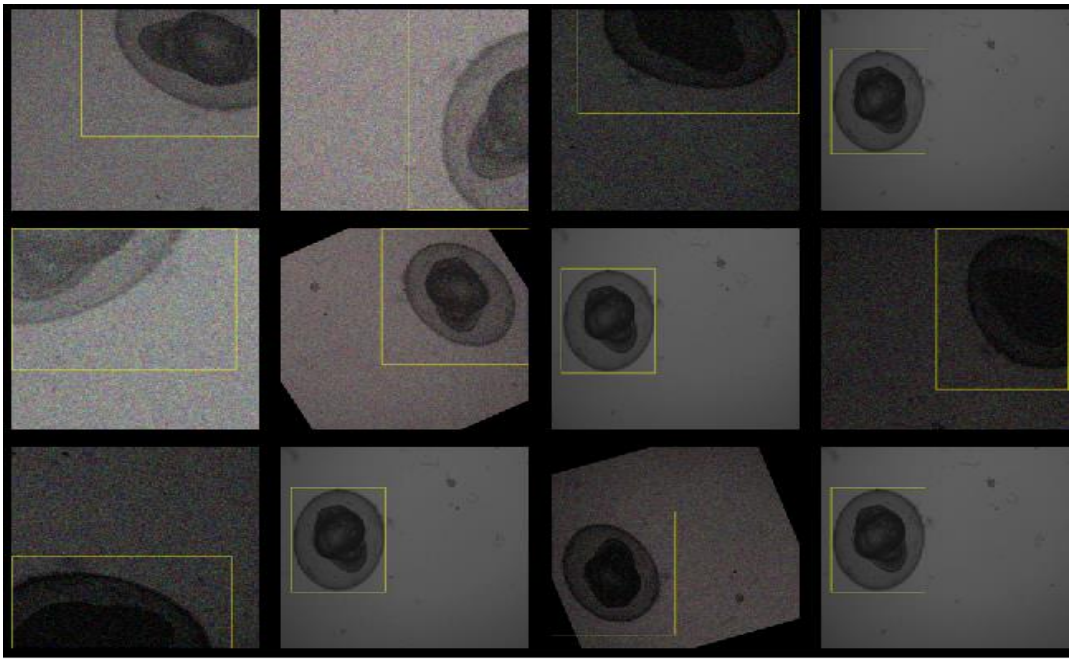


Figure 5-3 Random data augmentation of the zebrafish embryo from the dataset

5.2.3 Description of YOLOv2 Algorithm

You only look once (YOLO) is an end-to-end object detection method that was proposed by Redmon et al. for real-time processing in 2016 [77]. YOLOv2 is the improved version of the YOLO algorithm [78]. In this version, the accuracy of the YOLO algorithm was improved while speeding up the algorithm which makes it an outstanding choice in real-time applications. The significant difference in YOLOv2 is that it does not fix the anchor parameters, instead, YOLOv2 learns the anchor parameters of the ground truth bounding boxes making use of k-means clustering on the input data. During the training of the neural network, the weights are changed which changes the shape of the data that is sent to the next layer of the neural network. However, the next layer desires to see similar data pattern which makes it sensible to learn. Therefore, batch normalization is used to normalize the data before sending it to the next layer of the neural network. Hence, YOLOv2 performs batch

normalization after each layer and this highly improves the performance of the YOLOv2 algorithm. Another advantage of the YOLOv2 over the original YOLO is to accept various and higher resolution input images for the neural network. For instance, while the original YOLO accept 224 x 224 input images, YOLOv2 accepts a range of input images, such as 448 x 448, 608 x 608 etc. In YOLOv2, the input images are also randomly resized during the training so that the neural network can learn with objects in which the size is not fixed in various cases.

In YOLOv2, the input image is divided into S x S grid cells. Each of the grid cells is responsible to predict 5 bounding box coordinates, class probability, and a confidence score for each anchor. Bounding box coordinates consist of x, y, w, and h parameters. While x and y present the abscissa and ordinate of the centre pixel of the bounding box relative to the bounds of the grid cell, w and h indicate the width and height of the bounding box with respect to the whole image. Class probability decides the class of the object in the current grid cell probably belongs to. The confidence score represents the confidence that an object exists in the current grid cell. The confidence score is defined as follow;

$$Confidence = \Pr(obj)\ x\ IoU \tag{5.1}$$

Where Pr(obj) indicates the probability of the object falling into the current grid cell, Intersection over union (IoU) represents the rate of the area of overlap and area of union between the ground truth defined manually and ground truth predicted. Finally, the bounding boxes which possess less confidence score than the defined threshold are removed automatically using Non-Maximum Suppression (NMS). Simply, NMS is a post-processing algorithm that detects only once for each object by merging all bounding boxes for the same

object.

Multi-part loss function was used in YOLOv2 as it is shown in Equation 5.2. The loss function consists of three parts. These are bounding box, confidence, and class loss [78].

$$
\begin{aligned}
Loss = \ & \lambda_{coord} \sum_{i=0}^{S^2} \sum_{j=0}^{B} \mathbb{1}_{ij}^{obj} \left[(x_i - \hat{x}_i)^2 + \ (y_i - \hat{y}_i)^2 \right] \\
& + \ \lambda_{coord} \sum_{i=0}^{S^2} \sum_{j=0}^{B} \mathbb{1}_{ij}^{obj} \left[\left(\sqrt{w_i} - \sqrt{\widehat{w}_i} \right)^2 + \left(\sqrt{h_i} - \sqrt{\hat{h}_i} \right)^2 \right] \\
& + \sum_{i=0}^{S^2} \sum_{j=0}^{B} \mathbb{1}_{ij}^{obj} \left(C_i - \hat{C}_i \right)^2 + \ \lambda_{noobj} \sum_{i=0}^{S^2} \sum_{j=0}^{B} \mathbb{1}_{ij}^{noobj} \left(C_i - \hat{C}_i \right)^2 \\
& + \sum_{i=0}^{S^2} \mathbb{1}_{i}^{obj} \sum_{c \epsilon clases} (p_i(c) - \hat{p}_i(c))^2
\end{aligned}
\tag{5.2}
$$

where x_i and y_i represents the centre location of the bounding box, while w_i and h_i denote the width and height of the bounding box. C_i is the confidence score of the bounding box which shows whether there is an object. Later, $p_i(c)$ is the classification loss of the bounding box. All circumflex values, $\hat{x}_i, \hat{y}_i, \widehat{w}_i, \hat{h}_i, \hat{C}_i, \hat{p}_i(c)$ represent the predicted values of $x_i, y_i, w_i, h_i, C_i, p_i(c)$, respectively. S x S grid cell is stated as S^2 in the loss function. B represents the bounding boxes. $\mathbb{1}_{ij}^{obj}$ denotes whether there is an object, i.e. it value is 1, otherwise, it is 0. $\mathbb{1}_{ij}^{noobj}$ indicates that there is no object, i.e. its value is 1, else it is 0. $\mathbb{1}_{i}^{obj}$ represents that there is an actual object, i.e. its value is 1 if there is no actual object predicted, its value is 0. λ_{coord} is a constant which shows the weight of the coordinate loss while λ_{noobj} is the weight of the loss of bounding boxes when there is no object.

Yolov2 consists of 30 layers, of which 22 layers are convolutional layers and 5 layers are the

max-pooling layers. It is noted that each convolutional layer within the network includes batch normalization and Leaky Relu activation function, except the 30th layer. YOLOv2 also has 2 route layers to merge the layers, while the reorganization layer at the 26th layer is responsible for reshaping the output tensor so that the number of row and column of the tensor can match with the next output tensor and results in a concatenated layer with the size of 13 x 13 x 3072.

5.2.4 Transfer Learning

Transfer learning is an effective deep-learning technique that enables developers to use a pretrained classification neural networks that have already learned important features on huge datasets, such as ImageNet [72], as a starting point to learn a new task. ImageNet contains more than a million data samples with the capability of a 1000 class label. In this study, widely used convolutional neural networks were employed as a base network for the YOLOv2 to be able to detect the zebrafish embryo in the petri-dish in real-time. MATLAB Deep Network Designer app was used to visualize the pretrained neural networks. Next, the final classification layer of the pretrained neural networks was replaced with YOLOV2 detection layers to learn features specific to the zebrafish embryo dataset. The order of the detection layers of the YOLOv2 is Convolution, Batch Normalization, Relu, Batch Normalization, Relu, Convolution, YOLOv2 Transform Layer, and YOLOV2 Output Layer. This interface can be used as a guide to be used in other programming languages, such as Python and C. In the following section, the evaluated CNNs will be briefly discussed.

Resnet-50 and Resnet-18

Resnet-50 is a powerful deep neural network that demonstrated excellent performance on other recognition tasks and received first place in ImageNet detection, ImageNet localization, Coco detection, and Coco segmentation in ILSVRC and COCO 2015 competition [97]. Resnet-50 consists of 50 layers deep with an image size 224-by-224. Many researchers attempt to build deeper neural networks as the deeper neural network gives higher accuracy. However, when the network starts converging, due to the vanishing gradient effect, the accuracy degrades and saturates quickly. Resnet-50 is a residual network that skips the convolutional layers by using shortcut connections. Hence, the residual network adds the value of the previous block to the end of the block, then this never causes a small derivative that can approach zero and making the training ineffective. A small derivative is not desired. This is because the initial layers of the neural network are very significant input data, and if the weights and bias are not updated effectively at each iteration, the accuracy of the network will be low. In this way of Resnet-50, the neural network also handles the degrading accuracy problem, particularly in deeper networks. Resnet-18 is like Resnet50 with a reduced number of layers. Resnet-18 includes only 5 convolutional stages, which is a trade-off between computational cost and accuracy of the neural network.

MobileNet-v2

MobileNet-v2 consist of 53 deep layers with an image size 224-by-224 [98]. The unique side of MobileNet-v2 is that it employs depthwise separable convolution instead of standard convolution. By the cost of little accuracy, its computational cost was dropped about 87.5%. Hence its implementation as a pretrained network depends on the application requirements

in terms of a trade-off between accuracy and speed. Additionally, MobileNet-v2 also introduces linear bottlenecks between layers and shortcut connections between the bottlenecks. As it is explained in the Resnet-50 model, the shortcut connections help to give faster training and better accuracy.

Darknet-19

Darknet19 is the base feature extractor neural network that is being used for YOLOv2. Darknet-19 consists of 19 convolution layers and 5 max-pooling layers with an image size 256-by-256 [99]. Therefore, it requires less operation to process while providing high accuracy. In darknet-19, the full connection of layers removed, and batch normalization was applied to each layer. The details on YOLO v2 neural network architecture is already given in section 5.2.3.

SqueezeNet

SqueezeNet consists of 18 deep layers with an image size 227-by-227 [100]. SqueezeNet aims to create a small network that can easily fit into computer memory. Fire modules which include squeeze and expand phases are used in SqueezeNet. The squeeze phase uses a 1x1 filter size, while the expand phase utilizes 1x1 and 3x3 with a Rectified linear unit (ReLU). The size of parameters is only 0.5 MB in SqueezeNet, while AlexNet produces 240 MB. On the other hand, these two neural network models provided the same accuracy in the ImageNet image classification validation dataset.

Inception-v3 and GoogleNet

Inception-v3, which was an enhanced model of GoogleNet [101], has 48 deep layers with an image size 299-by-299 [102]. GoogleNet was the winner of the ILSVRC competition in 2014. Inception-v3 was proposed as a better version of GoogleNet in terms of computational efficiency and a reduced number of parameters. The originality of Inception v3 in comparison to GoogleNet is that it contains multiple small convolutions instead of having a larger filter size, and in the end, it concatenates them. To prevent overfitting, Inception v3 also utilizes label smoothing and batch regularization in auxiliary classifiers with RMSProp optimizers.

AlexNet

AlexNet is a simple CNN that consist of 8 deep layers with an image size 227-by-227 [103]. The simple structure of AlexNet makes it easy to train and optimize. AlexNet was proposed in the ImageNet Large Scale Visual Recognition Challenge (ILSVRC) in 2012. In AlexNet CNN architecture, the first four layers are followed by the pooling layer, while the fifth layer is followed by three fully connected layers. The usage of the ReLU non-linearity layer and the dropout regularization technique make AlexNet successful. Since the ReLU function is a half-wave rectifier as it is shown in Equation 5.3, this speeds up the training process as well as prevents overfitting. In the dropout technique, the input or hidden neurons are set to be zero stochastically to improve the co-adaptions of neurons.

$$f(x) = \max(x, 0) \tag{5.3}$$

5.2.5 Implementation of Data Training

In this study, the size of the input image was set to be 480 x 480 for all neural network

analyses. The configuration of the training hyperparameters was set to be as follows; the number of epochs was set to be 30 using stochastic gradient descent with momentum optimizer (SGDM), the initial learning rate was set as 0.001, and this learning rate was decreased 10 times after 15^{th} epoch. The momentum of 0.9 was used, while the mini-batch size was set to be 8. The training was conducted by a computer with an Intel® Core™ i7-8700 CPU @3.2 GHz (6 CPUs) and 32 GB of memory, GEFORCE RTX 2080 GPU with 12G memory. This study was entirely conducted with the same computer to make a sensible comparison. The transfer-learning-based neural network models using the YOLOv2 was written in MATLAB framework [104].

5.2.6 Performance Evaluation of Trained Models

The analysed trained neural networks were then evaluated by numerous parameters. These are mean Intersection over Union (IoU), average precision (AP), and average frame rate per second (fps). IoU is an evaluation metric that is used to measure the accuracy of the object detector based on a specific dataset. IoU is computed between the ground-truth defined manually before the data training procedure and the predicted bounding box generated by the detector. Hence, IoU is defined as the rate of the area of overlap and area of union. The mathematical formula of the IoU calculation of ground-truth bounding box, B_{gt} , versus predicted bounding box, B_p ,is given for zebrafish embryo in Figure 5-4.

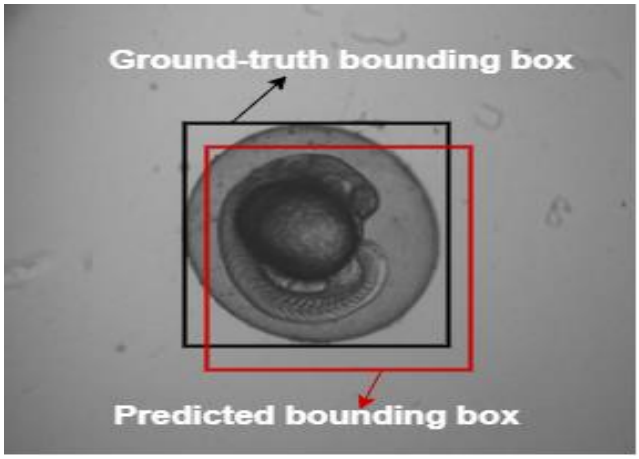


Figure 5-4 Schematic view of ground truth and predicted bounding box for IoU calculation

The calculation of IoU is therefore given in Equation 5.4:

$$IoU = \frac{B_{gt} \cap B_p}{B_{gt} \cup B_p} \tag{5.4}$$

Most of the performance evaluation parameters rely on four basic decision scores. These scores are True positive (TP), true negative (TN), false positive (FP), and false-negative (FN). These evaluation metrics are widely used in various studies in the literature [105, 106]. TP and FP depend on whether IoU is greater or less than the defined threshold. TP is defined as the number of positive samples correctly classified, while TN is the number of correctly predicted negative class. FP is described as the number of incorrectly predicted positive class, and FN is indicated as the number of negative samples incorrectly predicted.

Precision and Recall are an efficient way of measuring the quality of prediction. Their values are between 0 and 1. Precision, also called positive predictive value, is defined as the rate of TP to the number of predicted positives:

$$precision = \frac{TP}{TP + FP} \tag{5.5}$$

Recall is also known as hit rate or sensitivity. Recall is expressed as how many of the true positives were recalled:

$$recall = \frac{TP}{TP + FN} \tag{5.6}$$

Accuracy is given as the ratio of correctly predicted class to the total class:

$$accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{5.7}$$

AP can be calculated based on Recall – Precision curves. AP is the area under the Recall-Precision curves and this value is between 0 to 1. The greater the AP value provides better detection accuracy.

Finally, the average frame rate that can be processed in a second (fps) is measured to show the suitability of the neural network for positioning the zebrafish embryo to the FOV of the microscope in real-time.

5.3 **Results and Discussion**

The Loss Curves of the trained neural networks provide useful information and data which shows how the learning rate is well selected as well as whether the model overfits. Figure 5-5 represents the training loss converging curves during the learning period of eight different

neural network models. Since 420 images were used in the training dataset with a mini-batch size of 8, each epoch consists of approximately 52 iterations, and approximately 1600 iterations were conducted during the training procedure of the neural network models. The first 100 iterations out of approximately 1,600 iterations have been illustrated to present the behaviour of the neural network during the training procedure which was sufficient for its examination. The iterations beyond the 100 and to 1,600 behave similarly to what is illustrated from 10 to 100 iterations. As it is shown in Figure 5-5, the training loss starts around 20 for AlexNet, and approximately 30 for Resnet-50 and Darknet-19, while the training loss was started above 80 for the rest of the neural networks. Later, the training loss trend made a sudden decrease to lower values and then reached a stable loss until the end of the iteration. Among the all neural network models, the loss of Resnet-50 demonstrated the fastest decrease and reached to a minimum, followed by AlexNet. Although Darknet-19 was decreased to its minimum considerably fast, it has demonstrated slightly undesirable fluctuations during training. Resnet50 kept its lowest loss value and stability during training in comparison to other neural network models which demonstrates its consistency.

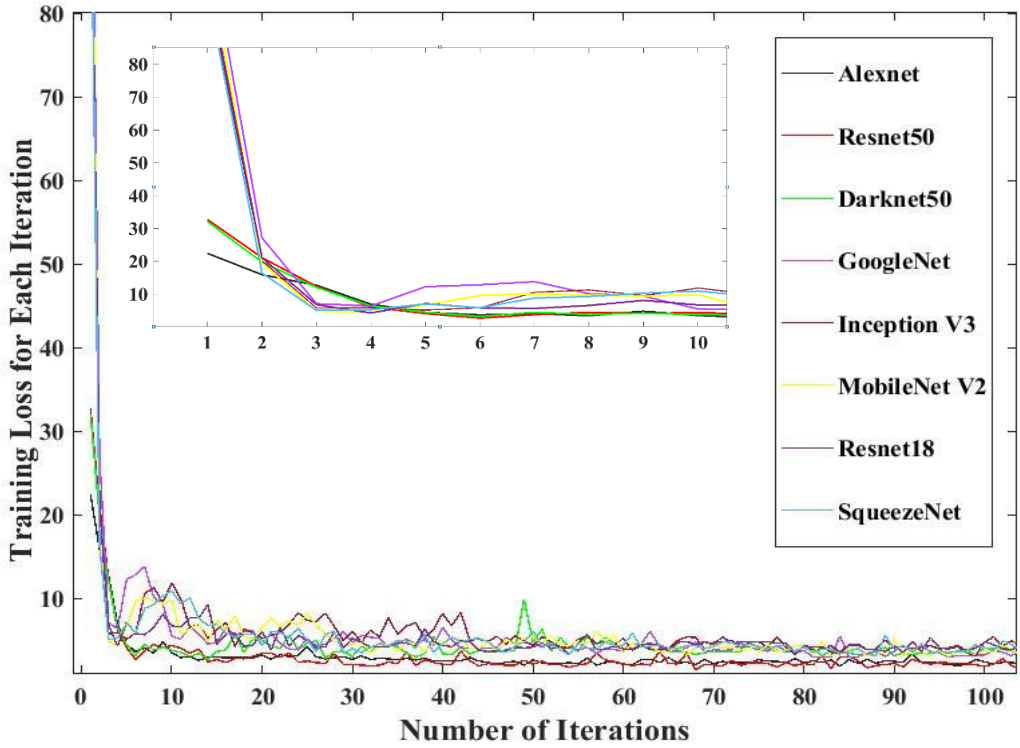


Figure 5-5 Training loss vs the number of iterations for pre-trained network models with YOLOv2 as well as close-look between the iterations 0 to 10

Average precision (AP) is a significant evaluation metric to show the performance of a neural network model. Table 1 indicates neural network models comparison using AP based on their precision-recall curves at 50%, 70%, and 85% IoU threshold, the training duration and the detection frame rate along with its standard deviation. Generally, True Positive of the detection is based on over 50% IoU threshold. However, our comparison technique was specially designed to see the performance of the neural network models, particularly at higher thresholds. This provides the essential platform to compare the models more transparently and help to distinguish between the models to obtain the most optimal one. In table 5-1, it is shown that all neural networks achieved over 95% detection accuracy at 50% IoU threshold, while the detection accuracy has demonstrated different patterns at 70% IoU threshold. The

comparison between the average precision of the ResNet-50 to the other closest network is almost the same for 50% and 70% IoU (1% for DarkNet-19 and 1.5% for ResNet-18 respectively). However, this figure increased to 53% difference (GoogleNet) for 85% IoU. Resnet-50 has demonstrated superior performance in comparison to other neural network models at 50% and 70% IoU threshold, with 100% and 99% detection accuracy respectively. The strength of the Resnet-50 had a clear effect on the detection accuracy at 85% IoU threshold. Resnet-50 is the best neural network model that is survived at 85% IoU threshold with a 66% detection accuracy as compared to models showing the closest performance, such as GoogleNet and Mobilenet-v2 with a 43% and 36% respectively. This is quite an important feature to have, as this high detection accuracy even if at higher thresholds, gives a hint about how the model will respond in localization of the bounding box.

Table 5-1 Neural network models comparison based on AP at 50%,70%, and 85% IoU threshold and average frame rate

Method	AP (%) at 50% IoU threshold	AP (%) at 70% IoU threshold	AP (%) at 85% IoU threshold	Average FPS	The standard deviation for FPS	Training duration (hours)
Resnet-50	0.9964	0.9857	0.6602	25.53	0.06	12.5
Darknet-19	0.9857	0.9399	0.1498	37.44	0.11	6.25
Resnet-18	0.9714	0.9714	0.1106	61.56	0.10	3.75
MobileNet-v2	0.9714	0.9678	0.3583	66.66	0.12	6.5
GoogleNet	0.9750	0.9640	0.4317	49.10	0.08	4.25
AlexNet	0.9780	0.8306	0.1141	115.05	0.04	1.5
Inception-v3	0.9559	0.7112	0.0447	33.96	0.10	5.25
SqueezeNet	0.9835	0.8476	0.0658	109.56	0.10	2

To meet the high-speed performance in the automated positioning of the zebrafish embryo task in real-time, average frame rate performance was also analysed for each method and presented in Table 5-1. Five sets of 20 images were recorded from the zebrafish embryo. Each set was run 10 times and the average fps score for all sets was then recorded. Subsequently, the final average fps score was calculated for 5 sets along with their standard deviation. Resnet-50 has shown the slowest frame rate performance with an average of approximately 26 fps, among all neural networks model. The highest frame rate was achieved by AlexNet with an average of nearly 115 fps; however, its detection accuracy is unsatisfactorily poor at 85% IoU threshold with approximately 11%. This detection accuracy directly affects its performance in localization of the boundary boxes as well as leading the misdetection. Detection accuracy is as significant as bounding box accuracy. This is because the technique proposed in this thesis aimed for auto-positioning of the zebrafish embryo which relies on its detection accuracy as well as its predicted bounding box accuracy. If the localization of the bounding box is poor, the embryo could only partially be detected, and the vision system would misguide the XY motorized stage. This would also lead mispositioning of the zebrafish embryo to the centre of the FOV. To avoid this, the neural network model is expected to provide high bounding box overlap accuracy while keeping the detection accuracy high. This study has conducted a comprehensive analysis of Intersection over Union (IoU) of all neural network models. The distribution of the IoU values was plotted against each image in the test set as it is presented in Figure 5-6. The number of images in each test set is 280. IoU threshold was set to be 0.5 for all neural network models and it was plotted in orange line in Figure 5-6. The mean IoU accuracy was plotted for each neural network model in red line. Figure 5-6 makes evidence that except Resnet-50 and Darknet-19 models, the rest of the models show a high amount of failed detection rates. The highest mean IoU accuracy

was found to be 86% for Resnet-50 and MobileNet-v2, while the lowest mean IoU was recorded as 56% and 57% for Inception-v3 and SqueezeNet respectively. However, MobileNet-v2 held 10 failed detection in the test, which is high as compared to the Resnet-50 model. The darknet-19 model has also demonstrated only 3 failed detection out of 280; however, it is average localization accuracy is only 76% and it is far behind in comparison to Resnet-50 model's average localization accuracy of 86%. Resnet-50 were failed only at the 190^{th} test image, just like all other neural network models failed. This single failed detection possibly related to the low quality of the image or have limited pixel data, i.e. the zebrafish embryo is partially visible in the test set. Overall, the Resnet-50 pretrained neural network model used as a backbone for YOLOv2 has demonstrated the highest detection and mean IoU accuracy among all analysed pretrained models. Although its frame rate is the slowest in comparison to other models, Resnet-50 with YOLOv2 was selected as an optimal neural network structure to conduct the fully automated zebrafish embryo positioning. Before conducting an experimental test for automatic zebrafish embryo positioning, an attempt was made towards optimization of the Resnet-50 with the YOLOv2 model to speed up the detection frame rate while keeping the detection and mean IoU accuracy maintained.

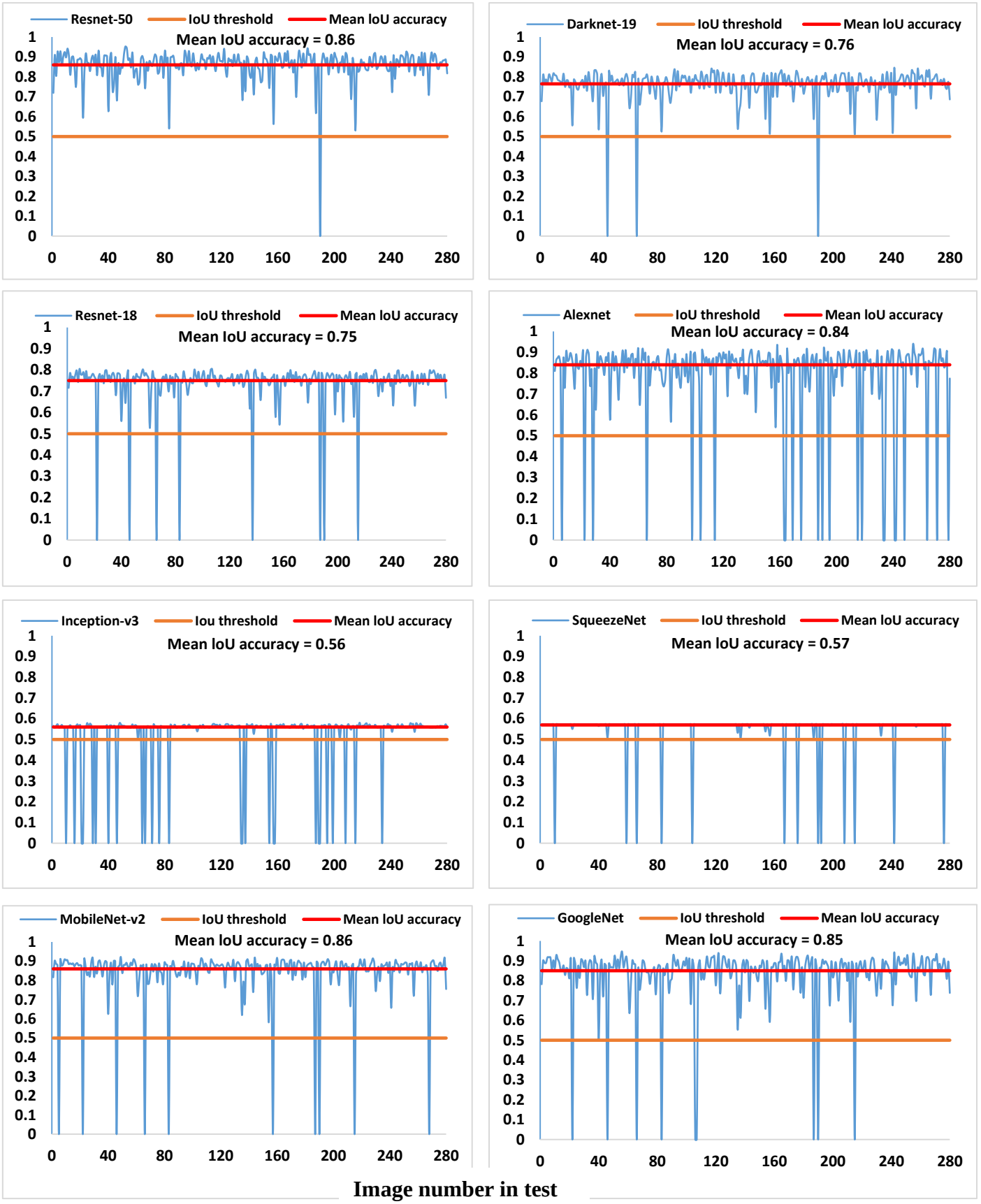


Figure 5-6 IoU analysis of neural network models in the test set

The input image size was reduced to 416 by 416 from 480 by 480 to improve the frame rate of the detector. In this case, Resnet-50 with a 416 by 416 input image provided 33 fps as compared to 26 fps previously which illustrates the increase of almost 27% in processing time. While keeping the detection accuracy the same at 50% and 70% IoU threshold, the detection accuracy also increased from 66% to 70% at 85 IoU threshold. Mean IoU accuracy was also improved from 86%±6% to 89%±6%. Reduction in resolution as a result of resizing the images does not affect the operation of the proposed technique. Overall, YoloV2 with Resnet-50 model demonstrated an outstanding detection and localisation accuracy while keeping the detection speed highly adequate for real-time application.

5.3.1 The Operational Assessment of the Automated Zebrafish Embryo Detection and Positioning

In this section, the automated zebrafish embryo positioning technique is proposed and evaluated. This section is divided into two parts. In the first part, the experimental zebrafish embryo detection procedure by YOLOv2 with Resnet-50 was investigated and verified since the positioning to the FOV relies on whether the zebrafish embryo is entirely in the FOV of the microscope. In the second part, the experimental automation procedure is given in detail. The initial experiment was conducted to verify the zebrafish embryo's full presence in the FOV. Ten different embryos were selected and moved to the FOV gradually. 60 images were captured for each embryo at an increment of approximately 13.2 µm movement. Confidence score changes were continuously recorded with respect to scanning distance as shown in Figure 5-7 along with the portions of zebrafish embryo visibility in percentage. The scanning was conducted continuously during the operation. However, the aim distance covers from

the point that the zebrafish embryo enters the FOV until it is completely visible. To validate the results, the entrance of the embryo and full visibility of it is confirmed by manual visual confirmation, and the accuracy of the measured distance was evaluated using ImageJ software. In Figure 5-7, each colour represents one set of experiments, and the IoU threshold was set to be 0.5 for each set. Once at most 30% of the zebrafish embryo enters the FOV, the developed vision system starts detecting procedure. While the scanning distance increases, the confidence score recorded by the algorithm rises as well up to the point that the zebrafish embryo illustrates full visibility. After this point, the operation was conducted using the information received from the coordinates and the size of the bounding box.

During the validation of the algorithm by using zebrafish as a biological model, the zebrafish embryo was detected at approximately 30% of its full size (~200µm). The XY stage continues to move while scanning the embryo until the algorithm reports above 89% confidence score. A further movement of 100µm is conducted to ensure the full visibility of the embryo within the FOV. This operation was conducted automatically by XY motorized stage to confirm the embryo is completely in the FOV. This further movement of the zebrafish does not affect the positioning task to the centre of FOV since the bounding box coordinate is in real-time tracking. Once the zebrafish embryo is in the FOV entirely, its confidence is score is checked for final confirmation. If the confidence score is lower than the defined threshold, the operation would automatically be brought to halt and a warning message appears for the users and mentioning “an inappropriate cell has been recognised”. Additionally, any abnormalities would affect the confidence score which can then be isolated from the healthy one when the warning message appears. However, the user can instruct the system by only confirming the procedure to conduct the rest of the operation. In this case, the CS drops to a lower value than the defined threshold.

This threshold is set to be above 80%. The lowest confidence score was observed as 89% in our experiment and it satisfactorily confirms its full presence in FOV. This threshold can be adjusted as required specifically for other types of microinjection applications. Finally, the vision system was effective to recognise zebrafish embryo full presence within the field of view of the microscope. The next task was to achieve the appropriate positioning of the zebrafish embryo to the centre of the FOV.

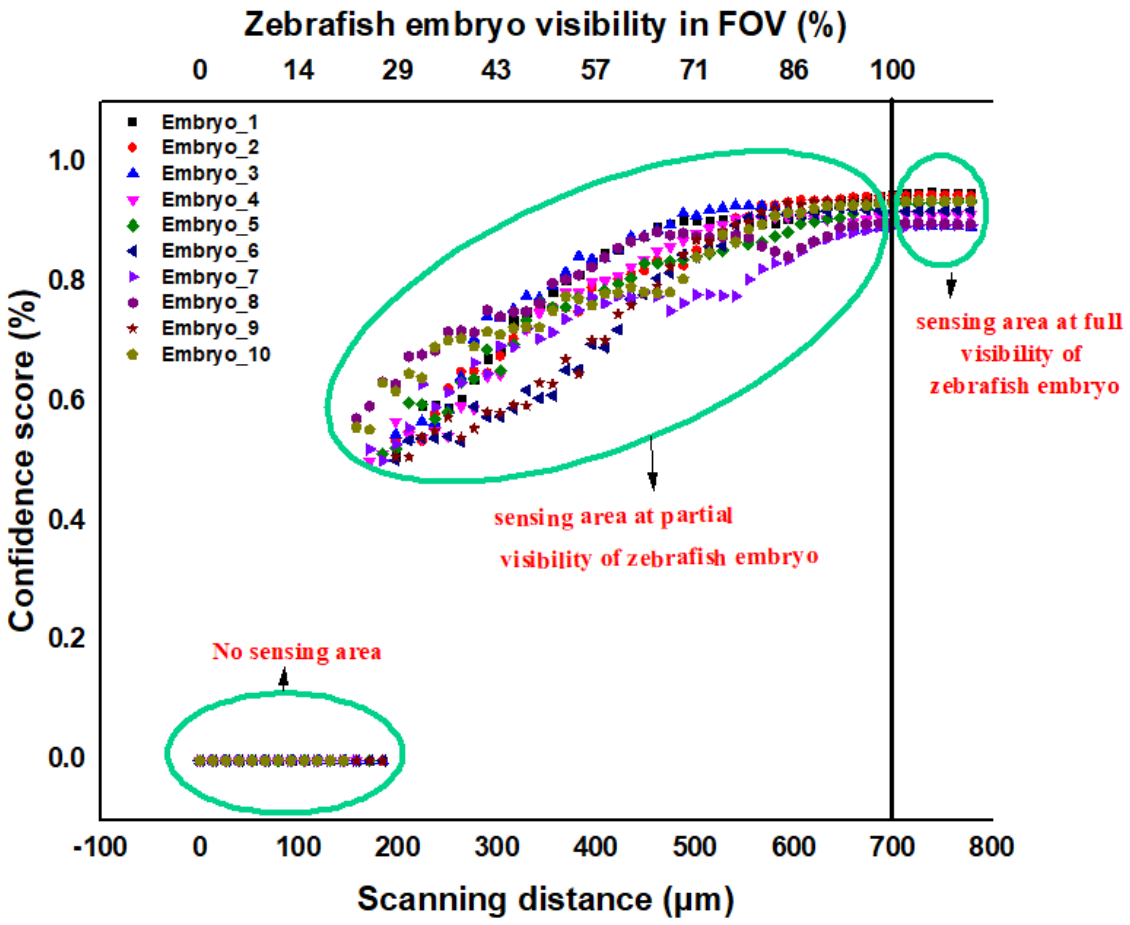


Figure 5-7 Zebrafish embryo scanning by XY motorized stage and its detection process based on the confidence score

The schematic view of the zebrafish embryo positioning is given in Figure 5-8. The built-in petri dish, just like a centre well dish, was employed to reduce the total scanning area by

87.4%. This will help to reduce the total operation time. This is because the petri dish was moved immediately from the initial home position to the ready position where the positioning procedure starts. Later, a serpentine movement was set for the XY motorized stage which starts from the ready position and moves to the right side of the scanning area. Once the first row of the scanning area is travelled, it moves down and continues scanning to the left side until the moment that the zebrafish embryo is completely visible under the microscope. As it is mentioned previously, the proposed technique can detect the zebrafish embryo while it enters gradually to the FOV of the microscope. Once it is fully in the FOV and confirmed by the YOLOv2 with the Resnet-50 algorithm, the vision algorithm commands to stop the motorized stage. Although the zebrafish embryo is in the field of view, it is expected to bring it to the aimed position which is the centre of FOV. The proposed technique benefited from the coordinate of the detected bounding box by YOLOV2 with Resnet-50. As it is shown in the description of the YOLOv2 method, each grid cell predicts bounding box coordinates. YOLOv2 with Resnet-50 model in MATLAB framework predicts the location of the object and returns as *'bboxes'* as an output argument. The output includes the coordinate and the size of the bounding box in the form of *[x, y, width, height]* where x and y indicate the coordinate of the top left corner of the bounding box. The centre position of the FOV is also known and fixed at each operation. The motorized stage moves from (x1, y1) coordinate to (x2, y2) coordinate as it is shown in Figure 5-8 as a part of the coarse tuning of zebrafish embryo positioning.

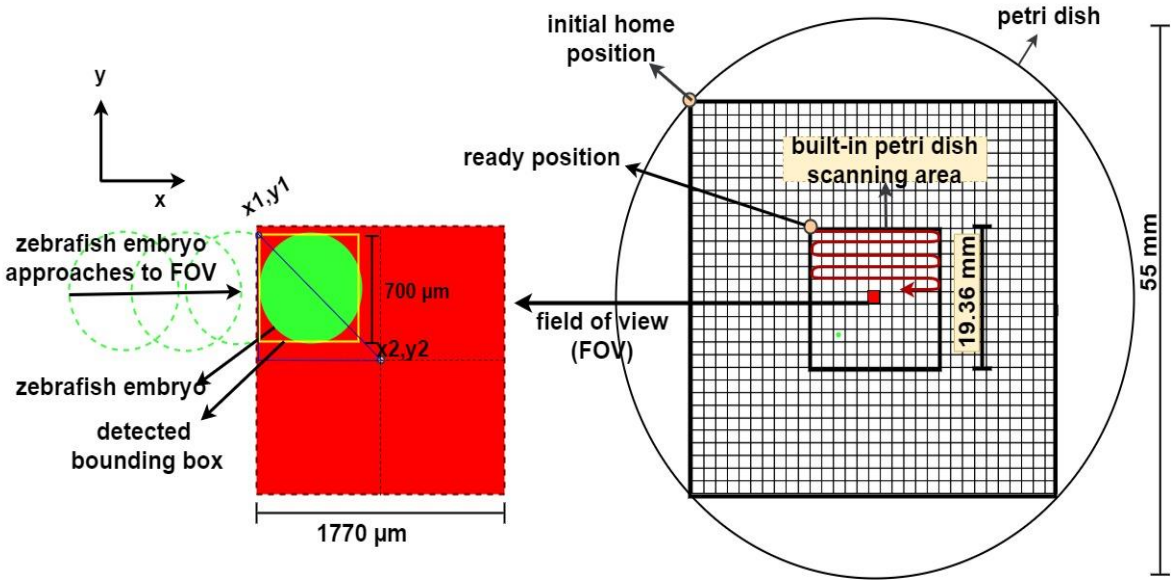


Figure 5-8 Schematic view of the positioning of the zebrafish embryo

Considering the 89% mean IoU accuracy, the width of the bounding box is almost 90% in similarity to the diameter of the zebrafish embryo. Since the embryo size can be slightly different from each other at each experimental trial, the proposed technique relied on the width and height of the detected bounding box with a total error of 10%. The fine-tuning operation has conducted the calibration by adjusting the bounding box coordinates to the centre of FOV using height and width information and hence positioning the embryo by the centre of FOV. The full operational flow diagram is presented in Figure 5-9.

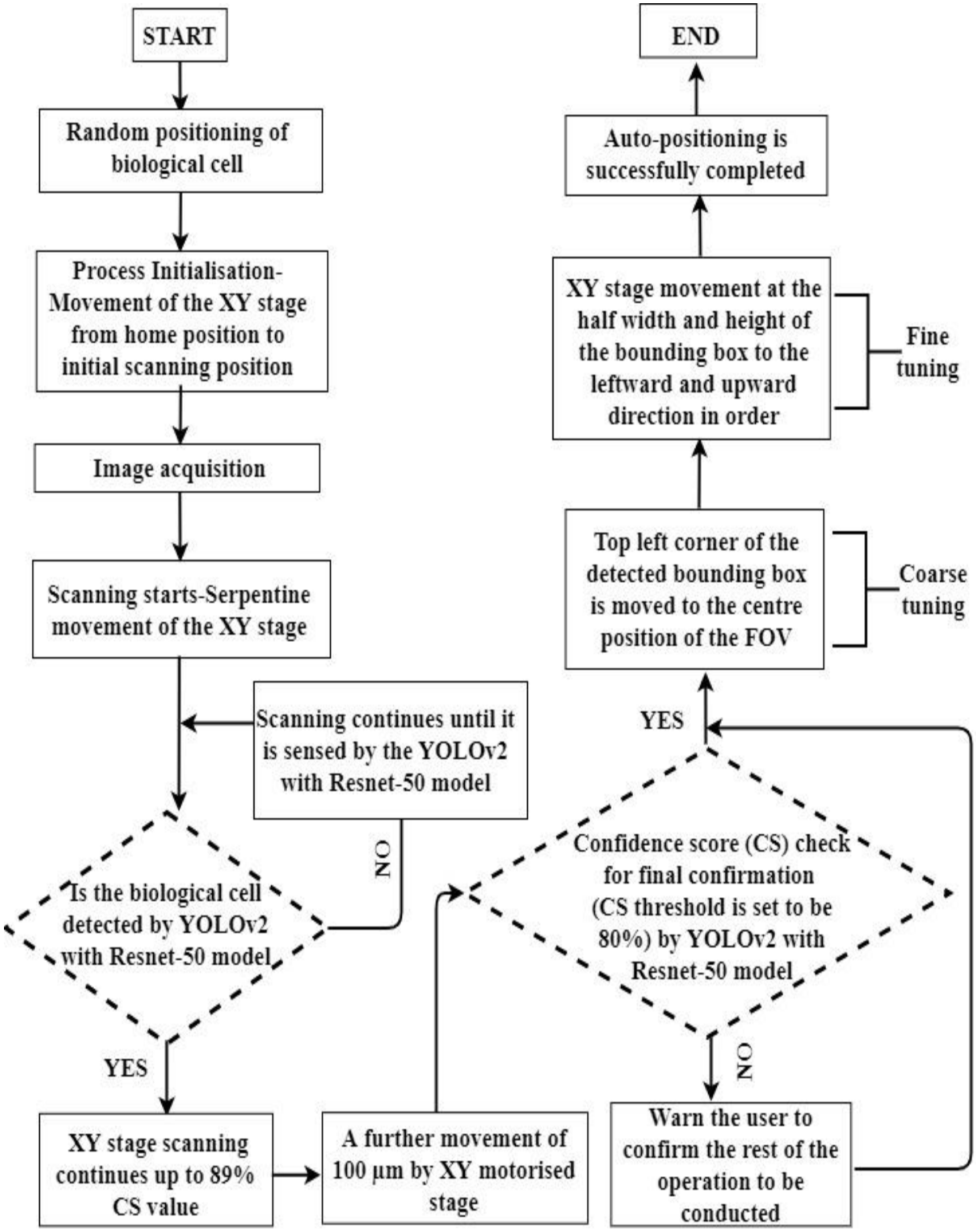


Figure 5-9 Operational flow chart of automated positioning of a biological cell

Our hypothesis was verified on a total of 100 experiments and it has shown its repeatability in automated zebrafish embryo positioning. All the zebrafish embryo positioning to the centre of FOV was conducted without any failure. Experimental results demonstrated that the technique proposed can achieve automated zebrafish embryo positioning effectively in real-time regardless of the initial position of the zebrafish embryo.

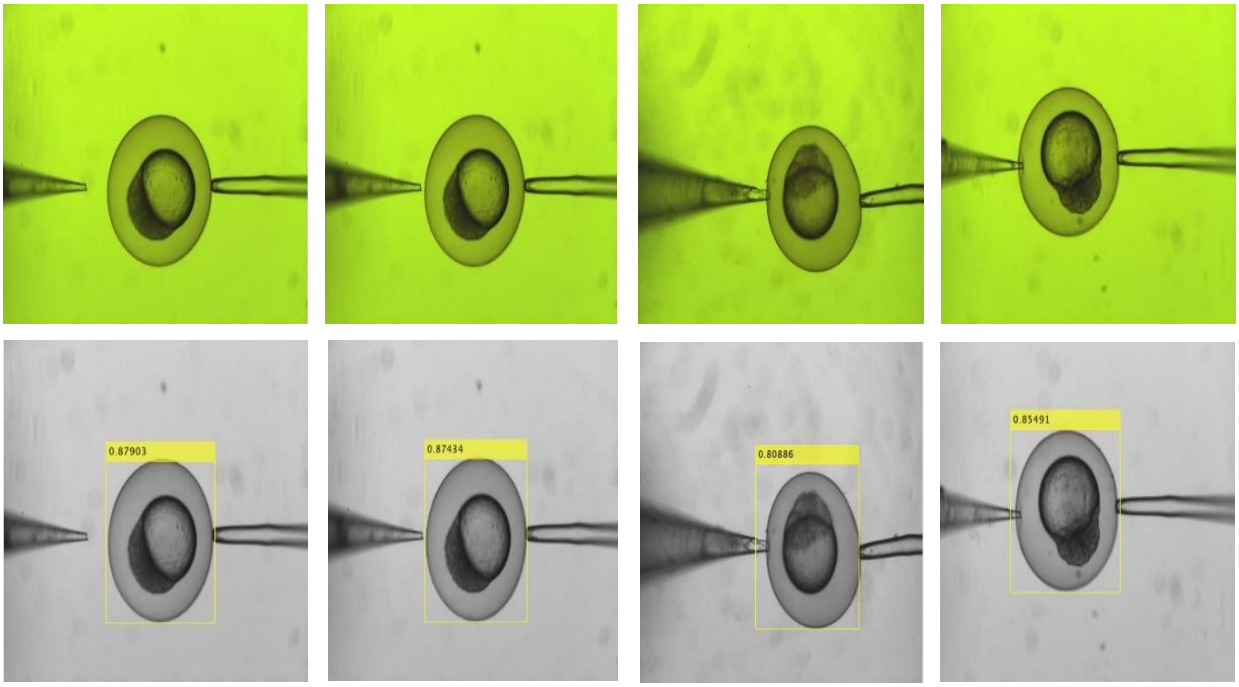
The proposed solution for this specific application is novel which hasn't been previously reported. In [107], an additional macro camera was used in conjunction with a micro camera to position the microgripper and zebrafish embryo to the field of view. This adds complexity and cost to the system. The template matching and grid detection algorithm were also utilised which is computationally expensive and requires an averagely of 30 seconds to bring the microgripper to the field of view. However, it was only suggested that the same method and algorithm could be used for zebrafish embryo positioning, though the results are not demonstrated. Due to the high computational time, it makes the real-time operation not feasible. In contrast, the proposed solution in this study was conducted in real-time using the transfer-learning powered Resnet-50 with the YOLOv2 technique using only a microscope camera. Our solution enhances the manipulating operation by adding real-time control (33 fps) and positional confirmation which increases the repeatability and accuracy in lower operational time.

5.3.2 Microinjection case study

Intracytoplasmic sperm injection (ICSI) is an assisted reproductive application where microinjection technology is effectively used. In ICSI, a holding pipette is used to hold a female egg in place by negative suction pressure, and then a single sperm cell is selected,

immobilised, and is directly introduced to the female egg using an injection micropipette. In this case study, a microinjection setup was designed similarly to the current ICSI system to evaluate the feasibility of the proposed algorithm on the zebrafish embryo recognition and its positioning to the centre of the FOV.

Despite employing different stages for the maturation of zebrafish embryo, the data set is strongly enhanced to detect the various samples with a detection accuracy of 100%. This technique can be generalized to any type of biological cell at various maturation stages in different microinjection applications. To validate this claim, additional automated positioning of the biological cells on different microinjection image sets was considered and, a dataset was used from our previous publication as shown in Figure 5-10 [45]. In this configuration, holding and injection pipettes are also involved to examine the impact of the manipulators' existence within the images. This microinjection setup configuration is different from the current setup that is used in the training section. BestScope (BUC2B-500C) camera was used with a 640*480 pixel input size at 30 fps. The dataset includes 180 images in total from the microinjection setup. Differently, the single image consists of the injection pipette on the left side, the holding pipette on the right side with a zebrafish embryo. The input image was also resized to 416 by 416 pixels and then it was converted to grayscale to match with the neural network input shape. The Yolov2 with Resnet-50 model was run on this dataset to investigate its consistency with the test set that is examined previously. The sample images as well as its detection were illustrated in Figure 5-10.



(a) **(b)** **(c)** **(d)**

Figure 5-10 Sample dataset from a microinjection experimental setup as well as its detection procedure. (a) the micropipette is away from the zebrafish embryo (b) the micropipette is almost in contact with the zebrafish embryo (c) the micropipette is in contact (d) the micropipette just starts puncturing the zebrafish embryo

Yolov2 with Resnet-50 model was demonstrated 89%±3% mean IoU accuracy with a 100% detection accuracy on the dataset as shown in Figure 5-11. Although the dataset has additional objects such as holding and injection pipette, it was sufficient to classify and detect the zebrafish embryo successfully without any failure. This result has shown consistency with our test set as the previous IoU and detection accuracy was reported 89%±6% mean IoU accuracy with 100% detection accuracy in the test set. Hence, the proposed solution is promising to be used in any type of microinjection setup as it is completely independent of the camera used in the system configuration. Additionally, the solution proposed for the automated positioning of the biological cells can also be used for the classification and detection of the cells during the microinjection process.

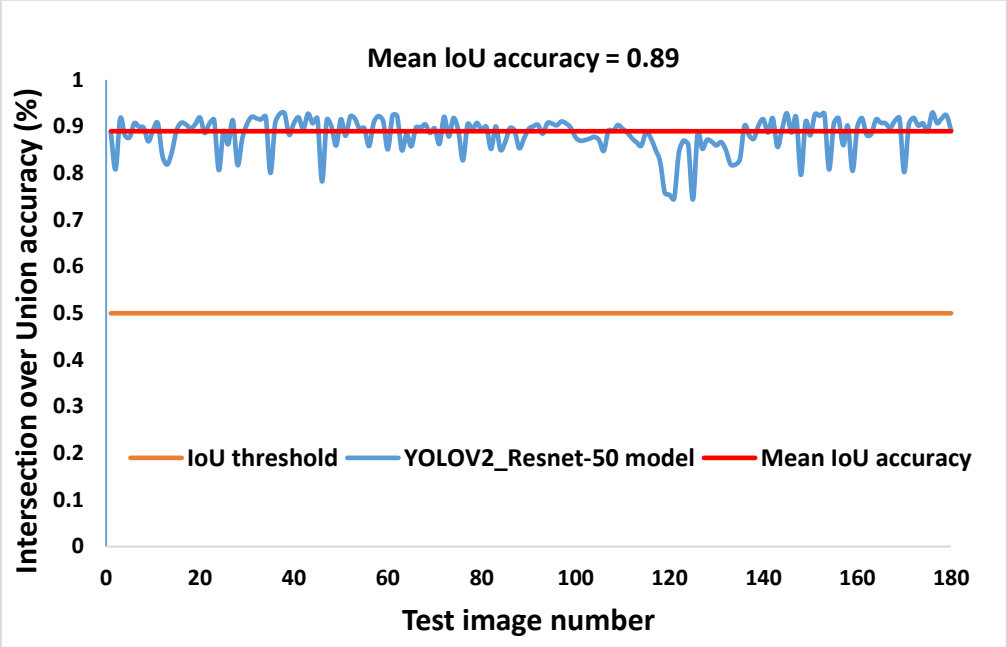


Figure 5-11 IoU analysis in the different microinjection setup dataset

5.4 **Conclusion**

This chapter proposes a technique to automatically recognise and position the randomly placed zebrafish embryo as a model of a biological cell to the centre of the FOV of the microscope. After a comprehensive evaluation of eight different pretrained neural network models as a backbone for the YoloV2 through transfer learning, YOLOv2 with Resnet-50 model demonstrated the optimal performance for the application of biological cell injection. It has demonstrated 89% mean IoU accuracy and 100% detection accuracy at an average of 33 frame rate in our test set. Zebrafish embryo detection operation was conducted on 100 sets of experiments successfully during the scanning process of the XY motorized stage. The proposed technique benefited from the coordinate and the size of the detected bounding box and locate the zebrafish embryo to the centre of the field of view in real-time. The experimental results demonstrated the proposed technique is reliable, repeatable, and

provides a 100% success rate regardless of the initial position of the zebrafish embryo at various maturation stages in the petri dish. The proposed solution was also tested and verified on different data from a real microinjection setup. Real-time detection and positioning of the biological cells can help embryologists save time, reduce the impact of subjective parameters in microinjection operations. This technique is not limited to microinjection applications and it can also be generalized to any type of biological cell at different maturation stages.

Chapter 6: Conclusion

6.1 Introduction

This thesis proposes a vision-guided methodology for the automation of the three main operations prior to the biological cell injection procedure. These operations are the alignment of the injection and holding pipette in three-dimensional, development of a vision-based sensor to measure the tip of the injection pipette displacement changes in three dimensions, and automated positioning of the zebrafish embryo to the centre of the microscope's field of view.

In the following sections, the developed methodology and main results along with the demonstration of the relation between the main chapters, the contribution of this thesis, and the future directions of this study are provided.

6.2 Summary of the thesis

Initially, the alignment of the injection and holding pipettes is conducted in three orthogonal axes in chapter 3. This is a prerequisite and the first step in microinjection operation. This provides a robust technique for auto-alignment of the injection and holding pipette prior to microinjection, specifically developed for ICSI application. The appropriate focusing algorithms have been selected to find an optimal focusing algorithm, specifically for ICSI. Twelve different focusing algorithms were evaluated based on focusing error, range, the number of false maxima, and width of the curve that each algorithm produces. Based on the evaluation of twelve focusing algorithms examined for injection and holding pipette, the

Brenner gradient has shown the highest overall performance for injection pipette focalization, while the Energy of gradient has demonstrated the highest overall performance for holding pipette focalization. It has been shown that the Fibonacci search algorithm is a suitable algorithm to drive the injection and holding pipettes to the global maximum of their focus curve. The reliability, accuracy, and insensitivity to the noise were further analysed and supported with a total of 100 experiments. The provided technique is almost 5 times faster than the available technique in previous studies.

In the developed auto-alignment technique, the initialization of the system starts before the alignment procedure to position the injection and holding pipette somewhere between the sensing ranges in the focal direction. Hence the developed technique was capable of focusing the injection and holding pipette and align them. This system does not require any additional operation, in case the injection pipette is being replaced. On the other hand, the system automatically aligns them to ensure the system is ready to start the injection. Hence, this would not damage the injection pipette during injection and conduct the operation safely.

The alignment of the injection and holding pipette procedure is not solely sufficient to start the microinjection operation. The vibrational displacement of the tip of the injection pipette needs to be checked before the operations continue further. This is because when injection starts, the acceleration and deceleration of the injection pipette magnify the vibration at the tip. This additional and undesired vibration causes uncontrollable forces to the cell membrane during the injection. This additional force may cause damage to the cell membrane or dislocate the biological cell hold by the holding pipette. It has been demonstrated that lateral vibration amplitude must be within the threshold that the system allows, which depends on the size of the biological sample used in the research.

After an accurate alignment procedure, the vibrational displacement of the injection pipette

in three orthogonal axes was measured to ensure the system is safe to continue to the injection procedure. Since the integration of the displacement sensors is not feasible in ICSI setup, initially a vision-based vibrational displacement sensor was developed for this purpose as demonstrated in chapter 4. The contribution to the development of vision-sensor to the literature is significant as it helps to clarify the role of the vibrational displacement sensing in ICSI. Therefore, after sensing the displacement of the injection pipette tip, further analysis was conducted to realize the effect of injection speed on the vibrational displacement creation. Later this information was used to determine the cell deformation, force fluctuation, and penetration force. Here, it was shown that visual sensing of the displacement sensing within ICSI operation is a crucial parameter that needs to be clarified. This is because the experimental results have demonstrated that injection speed is a contributing dynamical factor in ICSI operation due to its effect on the creation of the vibrational displacement. The increase in injection speed from 0.05 mm/s to 0.6 mm/s has led to an increase in lateral vibrational displacement by 54% while the egg deformation was decreased by 34%. Although the cell deformation was decreased, the system did not allow to conduct the operation in stable conditions. The cell injection operation was challenging after 0.6 mm/s injection speed due to high vibrations. Hence, the vibrational displacement is a decisive dynamical factor that directly affects the speed determination of the injection operation.

The final step towards the automation of the microinjection procedure is a real-time deep learning-based cell recognition and positioning framework as illustrated in chapter 5. This study provides an embryologist friendly approach for an automated egg positioning system to enhance automated microinjection operation experience. The systems proposed in the literature always assumed that the egg is already placed in the system's field of view. However, this manual operation needs to be automated as it is controlled by the operator

through a joystick.

In this study, eight different pre-trained neural networks were employed as a backbone for YOLOv2 method through transfer learning. The pre-trained models were evaluated based on mean Intersection over Union (IoU), Average Precision (AP) at different thresholds, and frame rate (fps) within the in-house produced zebrafish embryo dataset. Resnet-50 has shown superior performance with 89% mean IoU accuracy with 100% detection accuracy at an average of 33 fps. Its gradual presence was also verified with 100 different sets of experiments. The positioning of the zebrafish embryo to the defined position has performed successfully with a 100% success rate in all experimental trials. A dataset from a different microinjection setup which includes 180 images in total was also used to verify the performance of the proposed system. Yolov2 with Resnet-50 model was illustrated 89%±3% mean IoU accuracy with 100 % detection accuracy. The capability of the proposed system was sufficient to recognize the zebrafish embryo when the holding and injection pipette was included in the dataset which shows its robustness to the external disturbance in the image plane.

This technique is reliable, repeatable, and provides a 100% success rate for the randomly positioned zebrafish embryo at various maturation stages. This technique will help the embryologist to save time and eliminate subjective parameters in the injection operation. The employment of this technique will also contribute to the literature not only in microinjection applications but also it will guide the researchers to position any type of biological cells for various biomedical applications.

The entire vision-based automation towards fully automated microinjection operation was shown in Figure 1-2. Here, the proposed vision-based automation of the microinjection system is divided into three sections, namely auto-alignment of the injection and holding

pipette, system stability check through the vibrational displacement of the injection pipette tip, and zebrafish embryo positioning system. This graphical overview of the entire system was shown to demonstrate the relation between each system developed and how they complete each other to conduct a fully automated injection operation.

6.3 Contributions of this thesis

This thesis concentrated on the development of the vision system to improve automated microinjection operation and increase the success of ICSI procedure. The research contributions of this thesis are as follow:

- A novel technique was proposed to align the injection and holding pipettes in three orthogonal axes. It has been shown that Brenner gradient and Energy of gradient have superior performance to extract the depth position of the injection and holding pipettes.
- A new vision-based sensor was developed to measure the vibration displacement of the tip of the injection pipette in three orthogonal axes to ensure safe microinjection. The proposed sensor is operating at various magnifications.
- A new deep learning-based technique was built to automatically position the zebrafish embryo to the predefined position within the microscope's field of view and in real-time prior to injection operation. The developed vision-based positioning system can work with a different type of biological cells at various maturation stages using different branded microscopes.

6.4 **Future Directions**

In this thesis zebrafish embryo was used as a biological model; however, zebrafish embryo does not contain a polar body where a human oocyte does. The polar body is an intracellular structure where it is located in perivitelline space which is between the oolemma and zona pellucida membranes. The polar body consists of chromosomes where microinjection should be avoided from it as far as possible. Hence, a precise polar body detection is required procedure in the automation of microinjection systems since the biological cell needs to be manipulated based on the position of the polar body so that the system can avoid touching the polar body to prevent permanent damage to the cell. Additionally, the human oocyte is around seven times smaller than the zebrafish embryo in size. Although the developed system can be adapted to the human or other animal cells, this should be considered in the development of vision-based automated systems for ICSI.

A new vision-based sensor was developed to measure the tip of the micropipette displacement in three orthogonal axes. It is essential to minimize the measured displacement changes by adjusting the dynamical injection parameters, such as injection speed. Hence, an active vibration control scheme that dynamically reacts to vibration created at the tip of the micropipette can be isolated. This is essential to reduce the cell damages due to excessive vibration while having full control over the injection speed. This is an essential feature that an ICSI system can have to reduce the cell deformation by increasing the injection speed without introducing vibration on the cell membrane.

The full integration of the entire system is also required to provide meaningful communication between each developed vision systems. The signal delay is one of the significant issues in the integration of vision systems that need to be dealt with due to the

signal delays at each operation. Hence, the full system integration must be conducted in the future direction. In addition to this, clinical validation must be checked by the medical community to ensure the reliability of the proposed fully automated microinjection system.

References

[1] S. O. Rutstein and I. H. Shah, "Infecundity, Infertility, and Childlessness in Developing Countries. DHS Comparative Reports No.9," *ORC Macro, Meas. DHS*, 2004.

[2] G. Palermo, H. Joris, P. Devroey, and A. C. Van Steirteghem, "Pregnancies after intracytoplasmic injection of single spermatozoon into an oocyte," *Lancet*, vol. 340, no. 8810, pp. 17–18, 1992.

[3] "Fertility Treatments." [Online]. Available: https://www.newlife-ivf.co.uk/fertility-treatments. [Accessed: 04-Sep-2020].

[4] Y. Zhao, H. Sun, X. Sha, L. Gu, Z. Zhan, and W. J. Li, "A review of automated microinjection of zebrafish embryos," *Micromachines*, vol. 10, no. 1, p. 7, 2018.

[5] K. Sakaki, N. Dechev, R. D. Burke, and E. J. Park, "Development of an Autonomous Biological Cell Manipulator With Single-Cell Electroporation and Visual Servoing Capabilities," *IEEE Trans. Biomed. Eng.*, vol. 56, no. 8, pp. 2064–2074, 2009.

[6] M. T. Lin, L. Pulkkinen, J. Uitto, and K. Yoon, "The gene gun: current applications in cutaneous gene therapy," *Int. J. Dermatol.*, vol. 39, no. 3, pp. 161–170, 2000.

[7] M. Bungum, P. Humaidan, M. Spano, K. Jepson, L. Bungum, and A. Giwercman, "The predictive value of sperm chromatin structure assay (SCSA) parameters for the outcome of intrauterine insemination," *Hum. Reprod.*, vol. 19, no. 6, pp. 1401–1408, 2004.

[8] L. S. Mattos, E. Grant, R. Thresher, and K. Kluckman, "Blastocyst microinjection automation," *IEEE Trans. Inf. Technol. Biomed.*, vol. 13, no. 5, pp. 822–831, 2009.

[9] S. Friedler, A. Raziel, D. Strassburger, M. Schachter, Y. Soffer, and R. Ron-El, "Factors influencing the outcome of ICSI in patients with obstructive and non-obstructive azoospermia: A comparative study," *Hum. Reprod.*, vol. 17, no. 12, pp. 3114–3121, 2002.

[10] E. Taverna, C. Haffner, R. Pepperkok, and W. B. Huttner, "A new approach to manipulate the fate of single neural stem cells in tissue," *Nat. Neurosci.*, vol. 15, no. 2, pp. 329–337, 2012.

[11] W. Wang, Y. Sun, M. Zhang, R. Anderson, L. Langille, and W. Chan, "A system for high-speed microinjection of adherent cells," *Rev. Sci. Instrum.*, vol. 79, no. 10, p. 104302, 2008.

[12] G. Shull, C. Haffner, W. B. Huttner, S. B. Kodandaramaiah, and E. Taverna, "Robotic platform for microinjection into single cells in brain tissue," *EMBO Rep.*, vol. 20, no. 10, pp. 1–16, 2019.

[13] Z. Lu, X. Zhang, C. Leung, N. Esfandiari, and R. F. Casper, "Robotic ICSI (intracytoplasmic sperm injection)," *IEEE Trans. Biomed. Eng.*, vol. 58, no. 7, pp. 2102–2108, 2011.

[14] G. Wang and Q. Xu, "Design and development of a piezo-driven microinjection system with force feedback," *Adv. Robot.*, vol. 31, no. 23–24, pp. 1349–1359, 2017.

[15] J. M. Dumoulin *et al.*, "Embryo development and chromosomal anomalies after ICSI: effect of the injection procedure.," *Hum. Reprod.*, vol. 16, no. 2, pp. 306–312, 2001.

[16] F. R. Khan and S. S. Alhewairini, "Zebrafish (Danio rerio) as a model organism," in *In Current Trends in Cancer Management*, IntechOpen, 2018.

[17] H. B. Huang, D. Sun, S. Member, J. K. Mills, and S. H. Cheng, "Robotic Cell Injection System With Position and Force Control : Toward Automatic Batch Biomanipulation," vol. 25, no. 3, pp. 727–737, 2009.

[18] Y. H. Anis, M. R. Holl, and D. R. Meldrum, "Automated Selection and Placement of Single Cells Using Vision-Based Feedback Control," *IEEE Trans. Autom. Sci. Eng.*, vol. 7, no. 3, pp. 598–606, 2010.

[19] R. Kumar, A. Kapoor, and R. H. Taylor, "Preliminary Experiments in RobotMuman Cooperative Microinjection," in *Proceedings of the 2003 IEEE/RSJ lnll. Conference on lntelligent Robots and Systems*, 2003, no. October.

[20] Q. Zhao, M. Sun, M. Cui, J. Yu, Y. Qin, and X. Zhao, "Robotic cell rotation based on the

minimum rotation force," *IEEE Trans. Autom. Sci. Eng.*, vol. 12, no. 4, pp. 1504–1515, 2015.
[21] Z. Nan and Q. Xu, "Multiple-cell recognition and path planning for robotic microinjection system," in *Proceedings of the 36th Chinese Control Conference*, 2017, pp. 6691–6696.
[22] R. Redondo *et al.*, "Autofocus evaluation for brightfield microscopy pathology," *J. Biomed. Opt.*, vol. 17, no. 3, pp. 036008-1-036008–8, 2012.
[23] N. Kehtarnavaz and H. Oh, "Development and real-time implementation of a rule-based auto-focus algorithm," *Real time Imaging*, vol. 9, pp. 197–203, 2003.
[24] S. K. Nayar, M. Watanabe, and M. Noguchi, "Real-Time focus range sensor," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 18, no. 12, pp. 1186–1198, 1996.
[25] J. Brenner, B. Dew, J. Horton, T. King, P. Neurath, and W. Selles, "An automated microscope for cytologic research a preliminary evaluation," *J. Histochem. Cytochem.*, vol. 24, no. 1, pp. 100–111, 1976.
[26] E. Krotkov, "Focusing *," *Int. J. Comput. Vis. 1(3)*, pp. 223–237, 1988.
[27] M. Subbarao, T. Choi, and A. Nikzad, "Focusing techniques," *Opt. Eng.*, vol. 32, no. 11, pp. 2824–2837, 1993.
[28] Y. Qiu *et al.*, "NIH Public Access," *Anal Cell Pathol(Amst)*, vol. 36, no. 0, pp. 37–44, 2013.
[29] Y. Sun and B. J. Nelson, "Biological Cell Injection Using an Autonomous MicroRobotic System," *Int. J. Rob. Res.*, vol. 21, no. 10/11, pp. 861–868, 2002.
[30] T. Tanikawa, T. Arai, and N. Koyachi, "Development of Small-sized 3 DOF Finger Module in Micro Hand for Micro Manipulation," in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 1999, pp. 876–881.
[31] C. Nguyen, K. Ohara, E. Avci, T. Takubo, Y. Mae, and T. Arai, "Automated Micromanipulation for a Microhand with All-In-Focus Imaging System," in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2011, pp. 427–432.
[32] A. Suzuki, T. Arai, Y. Mae, K. Inoue, and T. Tanikawa, "Automated micro handling," in *IEEE International Symposium on Computational Intelligence in Robotics and Automation. Computational Intelligence in Robotics and Automation for the New Millennium (Cat. No. 03EX694)*., 2003, pp. 348–353.
[33] J. Liu *et al.*, "Locating end-effector tips in robotic micromanipulation," *IEEE Trans. Robot.*, vol. 30, no. 1, pp. 125–130, 2014.
[34] Y. Zhang, K. K. Tan, and S. Huang, "Vision-servo system for automated cell injection," *IEEE Trans. Ind. Electron.*, vol. 56, no. 1, pp. 231–238, 2008.
[35] Z. Wang, C. Feng, W. T. Ang, S. Yih, M. Tan, and W. T. Latt, "Autofocusing and polar body detection in automated cell manipulation," *IEEE Trans. Biomed. Eng.*, vol. 64, no. 5, pp. 1099–1105, 2017.
[36] E. P. Krotkov, *Active computer vision by cooperative focus and stereo*. New York: Springer Science & Business Media, 2012.
[37] I. M. Bahadur and J. K. Mills, "Robust autofocusing in microscopy using particle swarm optimization," in *IEEE International Conference on Mechatronics and Automation*, 2013, pp. 213–218.
[38] Y. Xiong and S. A. Shafer, "Depth from focusing and defocusing," in *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, 1993, pp. 68–73.
[39] Y. Wei and Q. Xu, "An overview of micro-force sensing techniques," *Sensors Actuators A Phys.*, vol. 234, pp. 359–374, 2015.
[40] T.-P. Dao *et al.*, "Analysis and optimization of a micro - displacement sensor for compliant microgripper," *Microsyst. Technol.*, vol. 23, no. 12, pp. 5375–5395, 2017.
[41] H. Zang, X. Zhang, B. Zhu, and S. Fatikow, "Recent advances in non-contact force sensors used for micro / nano manipulation," *Sensors Actuators A Phys.*, vol. 296, pp. 155–177, 2019.
[42] J. J. Lee and M. Shinozuka, "A vision-based system for remote sensing of bridge displacement," *Ndt E Int.*, vol. 39, no. 5, pp. 425–431, 2006.
[43] Y. Kimura and R. Yanagimachi, "Intracytoplasmic Sperm Injection in the Mouse '," *Biol.*

Reprod., vol. 52, no. 4, pp. 709–720, 1995.
[44] F. Sadak, M. Saadat, A. M. Hajiyavand, and G. Nomicos, "Vibrational Analysis during Cell Injection in ICSI Operation," in *2018 International Conference on Manipulation, Automation and Robotics at Small Scales (MARSS)*, 2018, pp. 1–6.
[45] A. M. Hajiyavand, M. Saadat, A. Abena, F. Sadak, and X. Sun, "Effect of Injection Speed on Oocyte Deformation in ICSI," *Micromachines*, vol. 10, no. 4, p. 226, 2019.
[46] E. Avci, C. Nguyen, K. Ohara, M. Kojima, Y. Mae, and T. Arai, "Towards High-Speed Automated Micromanipulation," in *2013 IEEE International Conference on Robotics and Automation*, 2013, pp. 1718–1723.
[47] M. Karzar-jeddi, N. Olgac, and T. Fan, "Dynamic response of micropipettes during piezo-assisted intracytoplasmic sperm injection," *Phys. Rev.*, vol. 84, no. 4, p. 041908, 2011.
[48] K. Ediz and N. Olgac, "Microdynamics of the piezo-driven pipettes in ICSI," *IEEE Trans. Biomed. Eng.*, vol. 51, no. 7, pp. 1262–1268, 2004.
[49] Q. Wu *et al.*, "Integration of autopatching with automated pipette and cell detection in vitro," *J. Neurophysiol.*, vol. 116, no. 4, pp. 1564–1578, 2016.
[50] M. Saadat, A. M. Hajiyavand, and A. S. Bedi, "Oocyte Positional Recognition for Automatic Manipulation in ICSI," *Micromachines*, vol. 9, no. 9, p. 429, 2018.
[51] T. Kasaya, H. T. Miyazaki, S. Saito, K. Koyano, T. Yamaura, and T. Sato, "Image-based autonomous micromanipulation system for arrangement of spheres in a scanning electron microscope," *Rev. Sci. Instrum.*, vol. 75, no. 6, pp. 2033–2042, 2004.
[52] B. Tamadazte, N. L. Piat, and S. Dembele, "Robotic Micromanipulation and Microassembly Using Monoview and Multiscale Visual Servoing," *IEEE/ASME Trans. Mechatronics*, vol. 16, no. 2, pp. 277–287, 2011.
[53] H. Li, X. Zhang, B. Zhu, Y. Lu, and H. Wu, "Micro-motion detection of the 3-DOF precision positioning stage based on iterative optimized template matching," *Appl. Opt.*, vol. 56, no. 34, pp. 9435–9443, 2017.
[54] H. Li, B. Zhu, Z. Chen, and X. Zhang, "Realtime in-plane displacements tracking of the precision positioning stage based on computer micro-vision," *Mech. Syst. Signal Process.*, vol. 124, pp. 111–123, 2019.
[55] X. Zhang, X. Zhang, H. Wu, J. Gan, and H. Li, "A high accuracy algorithm of displacement measurement for a micro-positioning stage," *Aip Adv.*, vol. 7, no. 5, p. 055301, 2017.
[56] X. Zhang, X. Zhang, H. Wu, H. Li, and J. Gan, "A robust rotation-invariance displacement measurement method for a micro- / nano-positioning system," *Meas. Sci. Technol.*, vol. 29, no. 5, p. 055402, 2018.
[57] F. Sadak, M. Saadat, and A. M. Hajiyavand, "Three Dimensional Auto-Alignment of the ICSI Pipette," *IEEE Access*, vol. 7, no. 1, pp. 99360–99370, 2019.
[58] Z. Zhang *et al.*, "Robotic immobilization of motile sperm for clinical intracytoplasmic sperm injection," *IEEE Trans. Biomed. Eng.*, vol. 66, no. 2, pp. 444–452, 2018.
[59] X. Zhang, C. Leung, Z. Lu, N. Esfandiari, R. F. Casper, and Y. Sun, "Controlled positioning of biological cells inside a micropipette," in *IEEE International Conference on Robotics and Automation*, 2012, pp. 2779–2784.
[60] K. Kim, X. Liu, Y. Zhang, and Y. Sun, "Nanonewton force-controlled manipulation of biological cells using a monolithic MEMS microgripper with two-axis force feedback," *J. Micromechanics Microengineering*, vol. 18, no. 5, p. 055013, 2008.
[61] F. Sadak, M. Saadat, and A. M. Hajiyavand, "Vision-based sensor for three-dimensional vibrational motion detection in biological cell injection," *Sensors*, vol. 19, no. 23, p. 5074, 2019.
[62] J. Liu *et al.*, "Automated robotic measurement of 3-D cell morphologies," *IEEE Robot. Autom. Lett.*, vol. 2, no. 2, pp. 499–505, 2016.
[63] W. Wang, X. Liu, D. Gelinas, B. Ciruna, and Y. Sun, "A fully automated robotic system for microinjection of zebrafish embryos," *PLoS One*, vol. 2, no. 9, 2007.

[64] L. Mattos, E. Grant, R. Thresher, and K. Kluckman, “New developments towards automated blastocyst microinjections,” in *Proceedings - IEEE International Conference on Robotics and Automation*, 2007, pp. 1924–1929.
[65] W. Z., B. G. Hu, L. C. Liang, and Q. Ji, “Cell detection and tracking for micromanipulation vision system of cell-operation robot,” in *In Smc 2000 conference proceedings. 2000 ieee international conference on systems, man and cybernetics.’cybernetics evolving to systems, humans, organizations, and their complex interactions*, 2000, pp. 1592–1597.
[66] N. O’Mahony *et al.*, “Deep Learning vs. Traditional Computer Vision,” in *Science and Information Conference*, 2019, pp. 128–144.
[67] P. F. Christ *et al.*, “Automatic Liver and Tumor Segmentation of CT and MRI Volumes using Cascaded Fully Convolutional Neural Networks,” *arXiv Prepr. arXiv1702.05970*, 2017.
[68] M. M. Alam and M. T. Islam, “Machine learning approach of automatic identification and counting of blood cells,” *Healthc. Technol. Lett.*, vol. 6, no. 4, pp. 103–108, 2019.
[69] N. Q. K. Le, E. K. Y. Yapp, N. Nagasundaram, M. C. H. Chua, and H. Y. Yeh, “Computational identification of vesicular transport proteins from sequences using deep gated recurrent units architecture,” *Comput. Struct. Biotechnol. J.*, vol. 17, pp. 1245–1254, 2019.
[70] N. Q. K. Le, E. K. Y. Yapp, and H. Y. Yeh, “ET-GRU: Using multi-layer gated recurrent units to identify electron transport proteins,” *BMC Bioinformatics*, vol. 20, no. 1, p. 377, 2019.
[71] V. Gulshan *et al.*, “Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs,” *JAMA - J. Am. Med. Assoc.*, vol. 316, no. 22, pp. 2402–2410, 2016.
[72] O. Russakovsky *et al.*, “ImageNet Large Scale Visual Recognition Challenge,” *Int. J. Comput. Vis.*, vol. 115, no. 3, pp. 211–252, 2015.
[73] R. Girshick, J. Donahue, T. Darrell, and J. Malik, “Rich feature hierarchies for accurate object detection and semantic segmentation,” in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2014, pp. 580–587.
[74] J. R. R. Uijlings, K. E. A. van de Sande, T. Gevers, and A. W. M. Smeulders, “Selective Search for Object Recognition,” *Int. J. Comput. Vis.*, vol. 104, no. 2, pp. 154–171, 2013.
[75] R. Girshick, “Fast R-CNN,” in *Proceedings of the IEEE International Conference on Computer Vision*, 2015, pp. 1440–1448.
[76] S. Ren, K. He, R. Girshick, and J. Sun, “Faster R-CNN: Towards real-time object detection with region proposal networks,” *Adv. Neural Inf. Process. Syst.*, pp. 91–99, 2015.
[77] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, “You only look once: Unified, real-time object detection,” in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2016, pp. 779–788.
[78] J. Redmon and A. Farhadi, “YOLO9000: Better, Faster, Stronger,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 7263–7271.
[79] L. C. Chiu and C. S. Fuh, “An efficient auto focus method for digital still camera based on focus value curve prediction model,” *J. Inf. Sci. Eng.*, vol. 26, no. 4, pp. 1261–1272, 2010.
[80] A. Santos, C. Ortiz De Solorzano, J. J. Vaquero, J. . M. Pena, N. Malpica, and F. Pozo, “Evaluation of autofocus functions in molecular cytogenetic analysis,” *J. Microsc.*, vol. 188, no. 3, pp. 264–272, 1997.
[81] K. Navar, “Shape from focus,” Carnegie Mellon University, 1989.
[82] W. Huang and Z. Jing, “Evaluation of focus measures in multi-focus image fusion,” *Pattern Recognit. Lett.*, vol. 28, no. 4, pp. 493–500, 2007.
[83] A. M. Eskicioglu and P. S. Fisher, “Image quality measures and their performance,” *IEEE Trans. Commun.*, vol. 43, no. 12, pp. 2959–2965, 1995.
[84] J. L. Pech-Pacheco, G. Cristobal, J. Chamorro-Martinez, and J. FernandezValdivia, “Diatom autofocusing in brightfield microscopy: a comparative study,” in *International Conference on Pattern Recognition.*, 2000, pp. 314–317.
[85] F. S. Helmli and S. Scherer, “Adaptive Shape from Focus with an Error Estimation in Light

Microscopy," in *Proceedings of the 2nd International Symposium on Image and Signal Processing and Analysis. In conjunction with 23rd International Conference on Information Technology Interfaces (IEEE Cat.*, 2001, pp. 188–193.
[86] X. . Liu, W. H. Wang, and Y. Sun, "Dynamic evaluation of autofocusing for automated microscopic analysis of blood smear and pap smear," *J. Microsc.*, vol. 227, no. 1, pp. 15–23, 2007.
[87] M. T. Yildirim, A. Basturk, and M. E. Yuksel, "Impulse noise removal from digital images by a detail-preserving filter based on type-2 fuzzy logic," *IEEE Trans. fuzzy Syst.*, vol. 16, no. 4, pp. 920–928, 2008.
[88] M. Kass, A. Witkin, and D. Terzopoulos, "Snakes : Active contour models," *Int. J. Comput. Vis.*, vol. 1, no. 4, pp. 321–331, 1988.
[89] J. Kiefer, "Sequential minimax search for a maximum," *Am. Mathmatical Soc.*, vol. 4, no. 3, pp. 502–506, 1953.
[90] Y. U. Sun, S. Duthaler, and B. J. Nelson, "Autofocusing in computer microscopy : selecting the optimal focus algorithm," *Microsc. Res. Tech.*, vol. 65, no. 3, pp. 139–149, 2004.
[91] K. Zuiderveld, "Contrast Limited Adaptive Histogram Equalization," in *Graphics Gems IV.San Diego: Academic Press Professional*, Academic Press, Inc., 1994, pp. 474–485.
[92] G. Deng and L. W. Cahill, "Adaptive Gaussian Filter For Noise Reduction and Edge Detection," in *IEEE Conference Record Nuclear Science Symposium and Medical Imaging*, 1993, pp. 1615–1619.
[93] D. Bradley and G. Roth, "Adaptive Thresholding using the Integral Image," *J. Graph. Tools*, vol. 12, no. 2, pp. 13–21, 2007.
[94] J. Canny, "A Computational Approach to Edge Detection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 8, no. 6, pp. 679–698, 1986.
[95] M. Westerfield, *The Zebrafish Book. A Guide for the Laboratory Use of Zebrafish (Danio rerio)*, 5th ed. Univ. of Oregon Press, 2007.
[96] L. Perez and J. Wang, "The effectiveness of data augmentation in image classification using deep learning," *arXiv Prepr. arXiv1712.04621*, 2017.
[97] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2016, pp. 770–778.
[98] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L. C. Chen, "MobileNetV2: Inverted residuals and linear bottlenecks," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2018, pp. 4510–4520.
[99] J. Redmon, "Darknet: Open Source Neural Networks in C." [Online]. Available: https://pjreddie.com/darknet.
[100] F. N. Iandola, S. Han, M. W. Moskewicz, K. Ashraf, W. J. Dally, and K. Keutzer, "SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and <0.5MB model size," in *arXiv preprint arXiv:1602.07360.*, 2016.
[101] C. Szegedy, S. Reed, P. Sermanet, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," in *In Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015, pp. 1–9.
[102] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the Inception Architecture for Computer Vision," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2016, pp. 2818–2826.
[103] A. Krizhevsky, I. Sutskever, and G. E.Hinton, "ImageNet classification with deep convolutional neural networks," *Adv. Neural Inf. Process. Syst.*, pp. 1097–1105, 2012.
[104] Mathworks, "Matlab and Deep Learning Toolbox Release 2019b." Natick, Massachusetts, United States.
[105] N. Q. K. Le, Q. T. Ho, and Y. Y. Ou, "Incorporating deep learning with convolutional neural networks and position specific scoring matrices for identifying electron transport proteins," *J.*

Comput. Chem., vol. 38, no. 23, pp. 2000–2006, 2017.
[106] D. T. Do and N. Q. K. Le, "Using extreme gradient boosting to identify origin of replication in Saccharomyces cerevisiae via hybrid features," *Genomics*, vol. 112, no. 3, pp. 2445–2451, 2020.
[107] L. Su *et al.*, "Macro-to-micro positioning and auto focusing for fully automated single cell microinjection," *Microsyst. Technol.*, 2020.