

DOCTORAL THESIS

Microfluidic Droplet Detection,
Classification and Quality
Assessment for Embedded
Flow Cytometry Systems

Fariha Afrin

TALLINN UNIVERSITY OF TECHNOLOGY
DOCTORAL THESIS
57/2025

Microfluidic Droplet Detection, Classification and Quality Assessment for Embedded Flow Cytometry Systems

FARIHA AFRIN



TALLINN UNIVERSITY OF TECHNOLOGY
School of Information Technologies
Thomas Johann Seebeck Department of Electronics

The dissertation was accepted for the defence of the degree of Doctor of Philosophy (Information and Communication Technology) on 17th June 2025

Supervisor: Professor Yannick Le Moullec, PhD
Thomas Johann Seebeck Department of Electronics
Tallinn University of Technology, Tallinn, Estonia

Co-supervisor: Senior Research Fellow Tamás Pardy, PhD
Thomas Johann Seebeck Department of Electronics
Tallinn University of Technology, Tallinn, Estonia

Opponents: Assistant Professor Luca Reggiani, PhD
Dipartimento di Elettronica, Informazione e Bioingegneria (DEIB)
Politecnico, Milano, Italy

Associate Professor András József Laki, PhD
Faculty of Information Technology
Pázmány Péter Catholic University, Budapest, Hungary

Defence of the thesis: 25th August 2025, Tallinn

Declaration:

Hereby I declare that this doctoral thesis, my original investigation and achievement, submitted for the doctoral degree at Tallinn University of Technology, has not been submitted for any academic degree elsewhere.

Fariha Afrin

signature

Copyright: Fariha Afrin, 2025
ISSN 2585-6898 (publication)
ISBN 978-9916-80-364-6 (publication)
ISSN 2585-6901 (PDF)
ISBN 978-9916-80-365-3 (PDF)
DOI <https://doi.org/10.23658/taltech.57/2025>
Printed by Koopia Niini & Rauam

Afrin, F. (2025). *Microfluidic Droplet Detection, Classification and Quality Assessment for Embedded Flow Cytometry Systems* [TalTech Press]. <https://doi.org/10.23658/taltech.57/2025>

TALLINNA TEHNIKAÜLIKOOL
DOKTORITÖÖ
57/2025

**Tilkade tuvastamine mikrofluidikas,
nende klassifitseerimine ja kvaliteedi
hindamine sardsetes voolutsütomeetria
süsteemides**

FARIHA AFRIN

Contents

LIST OF PUBLICATIONS	7
OTHER RELATED PUBLICATIONS	8
AUTHOR'S CONTRIBUTIONS TO THE PUBLICATIONS	9
Abbreviations.....	10
1 INTRODUCTION.....	12
1.1 Development of Imaging Flow Cytometry and Research Motivation	12
1.2 Challenges in Portable IFC Devices	15
1.3 Problem Statement, Research questions and Research Hypothesis	16
1.4 Thesis Contributions	17
1.5 Organization of the Thesis	18
2 SOME BACKGROUND ON BLUR-FREE SINGLE DROPLET IMAGE ACQUISITION, QUALITY ASSESSMENT AND DROPLET CLASSIFICATION IN IFC	19
2.1 Background on Blur-Free Single Droplet Image Acquisition.....	20
2.2 Background on Image Quality Assessment for IFC.....	20
2.2.1 Performance Metrics for IQA	22
2.3 Background on Droplet Classification	22
2.3.1 Performance Metrics for classification	23
2.4 Conclusion	23
3 BLUR-FREE SINGLE-IMAGE ACQUISITION IN IMAGING FLOW CYTOMETRY	25
3.1 Overview	25
3.2 Method	28
3.3 Selected results of the single droplet image acquisition	33
3.4 Conclusion on Blur-Free Single-Image Acquisition in Imaging Flow Cytometry	35
4 AUTOMATED DROPLET IMAGE QUALITY ASSESSMENT FOR PORTABLE MICROFLU- IDIC IMAGING FLOW CYTOMETRY	36
4.1 Overview	36
4.2 Method	38
4.2.1 Architecture for Microfluidic Droplet Image Quality Assessment	38
4.2.2 Training	38
4.2.3 Implementation	40
4.3 Selected results on automated droplet image quality assessment for portable microfluidic IFC.....	40
4.4 Conclusion on automated droplet image quality assessment for portable microfluidic IFC	42
5 CLASSIFICATION OF MICROFLUIDIC DROPLET ON RESOURCE CONSTRAINED PLATFORM	43
5.1 Overview	44
5.2 Method	45
5.2.1 Dataset Preparation	46
5.2.2 Modified Architecture.....	46

5.3	Overview of results of the classification of microfluidic droplet on resource-constrained platforms.....	47
5.4	Conclusion on the classification of microfluidic droplet on resource constrained platform.....	48
6	CONCLUSION AND FUTURE WORK	50
6.1	Summary	50
6.2	Perspectives	52
	List of Figures	53
	List of Tables	55
	References.....	56
	Acknowledgements	63
	Abstract.....	64
	Kokkuvõte	65
	Appendix 1.....	67
	Appendix 2	81
	Appendix 3	87
	Appendix 4	99
	Appendix 5	103
	Curriculum Vitae	111
	Elulookirjeldus.....	112

LIST OF PUBLICATIONS

The present Ph.D. thesis is based on the following publications that are referred to in the text by Roman numbers.

- I (ETIS 1.1) **F. Afrin**, K. Pärnamets, Y. Le Moullec, A. Udal, A. Koel, T. Pardy, T. Rang, "Embedded Blur-Free Single-Image Acquisition Pipeline for Droplet Microfluidic Imaging Flow Cytometry (IFC)," IEEE Access, vol. 12, pp. 92431-92441, 2024, DOI: <https://doi.org/10.1109/ACCESS.2024.3421637>

- II (ETIS 3.1) **F. Afrin**, Y. Le Moullec, T. Pardy, "Microfluidic Droplet Classification through Tuned Convolutional Neural Network on a Resource Constrained Platform", 2024 19th Biennial Baltic Electronics Conference (BEC2024), 2-4 October 2024, Tallinn, Estonia, DOI: <https://doi.org/10.1109/BEC61458.2024.10737958>

- III (ETIS 1.1) **F. Afrin**, Y. Le Moullec, T. Pardy, T. Rang, "Lightweight CNN-based Microfluidic Droplet Classification for Portable Imaging Flow Cytometry". Proceedings of the Estonian Academy of Science, 74, 2S, 302–311, 2025. DOI: <https://doi.org/10.3176/proc.2025.2S.05>.¹

- IV (ETIS 5.2) R. Jõemaa, **F. Afrin**, N. Gyimah, K. Ashraf, K. Pärnamets, T. Pardy, "Cogni-Flow: Integrated Modular System For Automated Droplet Microfluidic Bioanalysis", EUROSENSORS XXXVI, 01-04 September 2024, Debrecen, Hungary

- V (ETIS 3.1) **F. Afrin**, D. Ndubuisi Ezechukwu, Y. Le Moullec, T. Pardy, T. Rang, A. Koel, "CNN-Transformer Hybrid Model Towards Automated Droplet Image Quality Assessment of Portable Imaging Flow Cytometer". To appear in the proceedings of IEEE EUROCON 2025 - 21st International Conference on Smart Technologies, 4-6 June 2025, Gdynia, Poland.

¹This is an extended version of Publication II. This extended version includes additional contribution and material that increase the depth and breadth of the work, including deployment on an additional board, comparisons, and insights. See Chapter 4 for further details.

OTHER RELATED PUBLICATION

During my PhD studies, I contributed to the CogniFlow-Cyte "Cognitronic Lab-on-a-Chip System for Highly-Automated Flow Cytometry" project (Estonian Science Agency ETAg PRG620 grant). In this context, I collaborated with other project members and I contributed to the following paper (which is not part of this PhD thesis):

(ETIS 1.1) K. Pärnamets, T. Pardy, A. Koel, T. Rang, O. Scheler, Y. Le Moullec, **F. Afrin**, "Optical detection methods for high-throughput fluorescent droplet microflow cytometry", *Micromachines* 12 (3), 345, 2021

AUTHOR'S CONTRIBUTIONS TO THE PUBLICATIONS

- I In Publication I, I was the first author. I developed the lightweight imaging pipeline, including conceptualization of the hardware solution and development of a lightweight algorithm for detecting, counting, and acquiring single object images from video streams. I designed and executed experiments using videos of fast-moving droplets with high flow rates. I analyzed the data to assess counting accuracy and processing time on various platforms (desktop PC, Raspberry Pi-4, and Nvidia Jetson Nano), and contrasted the results with existing studies. I drafted the manuscript. I incorporated my supervisors' and co-authors' feedback in the revisions.
- II In Publication II, I was the first author. I developed the customized YoloV4-Tiny model for microfluidic droplet classification on a Raspberry Pi-5 single board computer (SBC). I created the custom dataset from our microfluidic setup, trained the neural network, and implemented it on the SBC. I evaluated the model's performance and analyzed the results to demonstrate the model's suitability for portable devices. I created the figures and tables, and drafted the manuscript. I incorporated my supervisors' feedback in the revisions.
- III In Publication III, I was the first author. In addition to the work presented in Publication II, I extended the state of the art overview, provided more details about the customization of the model and its deployment. Moreover, as an additional significant contribution, I trained and deployed a YoloV5s model on a MaixCam board. Furthermore, I conducted a comparative analysis against seven other ML model-platform combinations and I analyzed the results. I drafted the paper and revised it based on my supervisor's and co-authors' feedback.
- IV In Publication IV, I was the second author. As a co-author, I contributed to developing and implementing the YoloV4 and YoloV4-Tiny based algorithms for object classification. I also contributed in achieving the suitable accuracy for single-cell detection in droplets towards its integration in the existing microfluidic system. I also contributed to the preparation of the paper and poster.
- V In Publication V, I was the first author. I generated a new dataset of good and poor quality microfluidic droplet images. I compared and evaluated the selected NR-IQA approaches on this dataset. I developed the novel hybrid model combining a modified MobileNet with a transformer block, specifically for microfluidic droplet IQA. I evaluated the hybrid model's performance; I also deployed and tested it on a Raspberry PI-5 and analyzed the results. I drafted the paper and incorporated feedback from supervisors during revisions.

Abbreviations

2D	Two-Dimensional
3D	Three-Dimensional
AI	Artificial Intelligence
ANN	Artificial Neural Network
BRISQUE	Blind/Referenceless Image Spatial Quality Evaluator
CNN	Convolutional Neural Network
CL	Convolutional layer followed by batch normalization and the Leaky-ReLU activation function
CSP	Cross-stage Partial
DOS	Distribution of Opinion Score
DL	Deep Learning
DRL	Deep Reinforcement Learning
EMD	Earth Mover's Distance
FACS	Fluorescence-Activated Cell Sorting
FC	Flow Cytometry
FPS	Frame Per Second
FR-IQA	Full Reference IQA
GPU	Graphics Processing Unit
IFC	Imaging Flow Cytometry
IoU	Intersection Over Union
IQA	Image Quality Assessment
IS100	ImageStream System 100
IS-X	ImageStream Instruments (IS-X)
MAP	Mean Average Precision
MF	MicroFluidic
ML	Machine Learning
MOS	Mean Opinion Score
MSE	Mean Squared Error
NasNet	Neural architecture search Network
NIQE	Naturalness Image Quality Evaluator
NR-IQA	No-Reference IQA
PC	Personal Computer
PDMS	Polydimethylsiloxane
PIQE	Perception-bases Image Quality Evaluator
PLCC	Pearson Linear Correlation Coefficient
PoC	Point-of-Care
PSNR	Peak Signal-to-Noise Ratio
RPI-4	Raspberry Pi-4
RPI-5	Raspberry Pi-5
ReLU	Rectified Linear Unit
RQ	Research Question
RR-IQA	Reduced-Reference IQA
SBC	Single Board Computer
SRCC	Spearman Rank-Order Correlation Coefficient
SSIM	Structural Similarity Index Measure
STD	Standard Deviation
TOPS	Tera Operations Per Second

WSCNet	Weakly Supervised Cell Counting Network
Yolo	You Only Look Once
YoloV4	A version of Yolo that balances speed and accuracy, implemented with Darknet
YoloV4-Tiny	Compressed version of YoloV4 designed to train on machines that have lower computing power
YoloV5	A successor to YoloV4 with architectural improvements and optimizations for speed and accuracy, implemented with PyTorch
YoloV5-s	A small version of YoloV5

1 INTRODUCTION

1.1 Development of Imaging Flow Cytometry and Research Motivation

Flow cytometry (FC) is a powerful technology in biomedical research to characterise the complex phenotypes of the target samples, for example, micro-organisms or individual cells. Imaging flow cytometry (IFC) combines traditional flow cytometry with high throughput image acquisition, providing spatial information about cellular features. IFC allows simultaneously analyzing multiple parameters, e.g. cell morphology, protein localization, and fluorescence intensity in a single-cell. IFC captures high-resolution images of cells flowing through the system, enabling researchers and practitioners to visualize and quantify cellular characteristics that were previously difficult to assess using conventional flow cytometry alone. IFC has a wide range of applications, e.g. immunology, cancer research, and cell biology.

A short history of FC and IFC is depicted in Figure 1 and summarized in what follows.

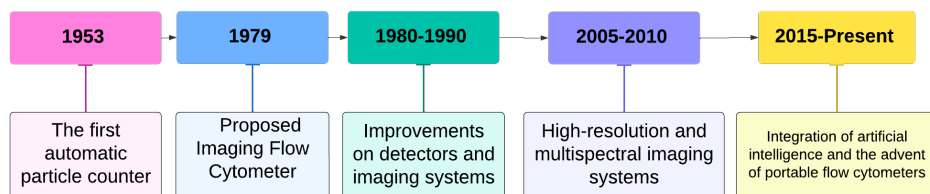


Figure 1: Evolution of Imaging Flow Cytometry, from the first automatic particle counter to the integration of AI/ML and the advent of portable flow cytometers

The groundwork of traditional flow cytometry began with the invention of the Coulter Counter by Wallace H. Coulter in 1953 [1]. The Coulter Counter was the first automated device to count and measure the size of particles suspended in a fluid, using changes in electrical impedance. In the late 1960s, Leonard Herzenberg and his colleagues introduced fluorescence-activated cell sorting (FACS) [2], another milestone of FC that allowed to label cells with fluorescent markers and sort them based on their fluorescence intensity.

IFC itself was first proposed in 1979 [3] and then first commercially developed by Amnis Corporation [4]. However, for nearly two decades, IFC faced significant challenges due to technological limitations, especially in terms of image resolution compared to fluorescence microscopes. In the 1980s and 1990s [5], efforts focused on improving detectors and imaging systems.

Then in 2005, an IFC instrument, the ImageStream System 100 (IS100) was introduced by Amnis, and subsequently the second generation of ImageStream instruments (the IS-X) was launched in 2009 [6]. The IS100 combined flow cytometry's speed and phenotyping capabilities with microscopy's detailed imagery and functional insights, enabling simultaneous quantification of multiple fluorescent emissions and scattered light. The IS-X, ten times faster than its predecessor, significantly improved both sample (image) throughput and image quality, enabling faster and more detailed cellular analysis. Its modularity made it adaptable to a wider range of budgets and research needs, and thus more competitive against traditional flow cytometry systems. These advancements made IFC a powerful tool for biomedical research, particularly in cellular analysis.

Until approximately 2010, research and studies in IFC focused on enabling the analysis of more complex biological systems with higher accuracy, and around 2015, artificial intelligence (AI) –in particular machine learning (ML)– began to play an increasingly important role in enhancing image analysis. ML-driven algorithms improved the ability to identify cellular features and classify cell types based on complex patterns.

The most recent developments in IFC focus on image analysis capabilities using ML [7] and AI [8], [9] as well as the portability [10]. Portable IFC instrumentation devices, suitable for use in resource-constrained environments, are now being developed for point-of-care (PoC) diagnostics [11] and global health applications. These devices use lightweight hardware components (such as miniaturized MF chips and detector), and use deep learning models for real-time data analysis, to bring the power of IFC outside the traditional laboratory environment. The target is to automate the classification of cellular objects and extract more information from image data than was previously possible with manual methods.

As depicted in Figure 2, a typical IFC device based on a microfluidic (MF) droplet chip has three main modules: A. a sample pumping module, B. a droplet generation module which includes an excitation and detection lighting system, and C. an evaluation module.

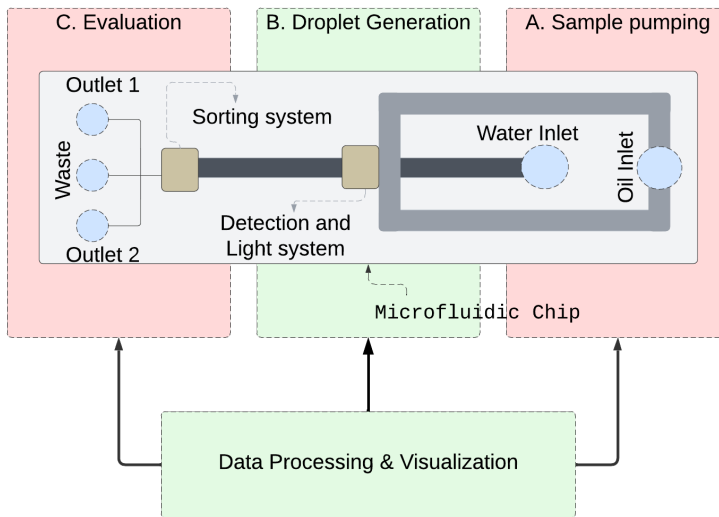


Figure 2: Simplified diagram of a MF chip based IFC device

Using MF droplets in imaging flow cytometer holds great promise as this can perform typical laboratory operations by flowing two immiscible fluids, typically oil and water, within the microscale channels of the MF chip. Unlike continuous flow systems, droplet-based systems focus on creating few hundreds to thousands of uniform droplets in a second [12]. Each droplet acts as tiny, isolated reaction chamber, which is particularly advantageous in biological and chemical studies requiring precisely controlled microenvironments. Droplet formation in microfluidics is typically achieved through specialized geometries and flow designs, such as flow-focusing, T-junction, and co-flow, that facilitates accurate control over droplet size, generation rate and composition. Over the past decade, MF droplet generation and encapsulating cells in droplets has been used for cell analysis, and medical diagnostics.

The principles of droplet generation and cell encapsulation are illustrated in Figure 3, and actual examples of droplet images captured when flowing inside a microfluidic channel (images extracted from the dataset used in later chapters of this thesis) as shown in Figure 4.

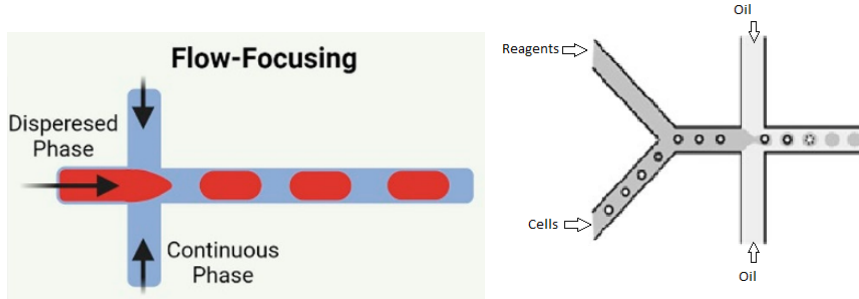


Figure 3: An example of MF droplet generation (Left) and cell encapsulation (Right). Reproduced from [13]

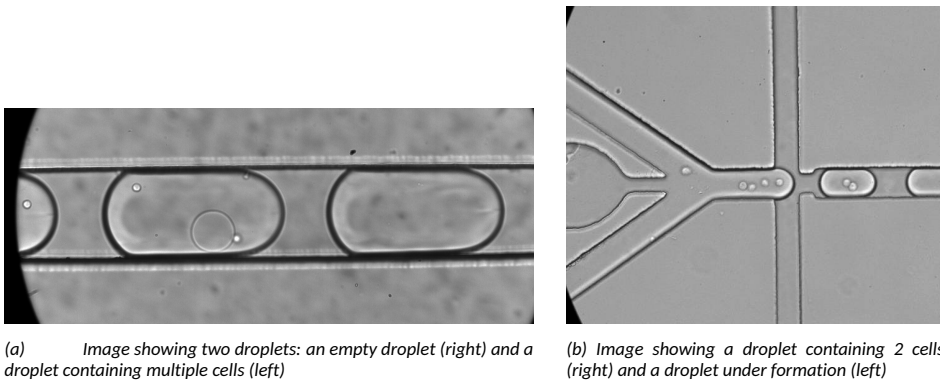


Figure 4: Images of MF droplets flowing inside a microfluidic channel (images extracted from the dataset used in later chapters of this thesis).

In recent years, microfluidic droplet data analysis has gained significant attention in fields such as biomedical research, drug discovery, and diagnostic testing. The analysis of microfluidic object in a resource constrained platform at higher throughput could lead to the development of portable imaging flow cytometry.

The research presented in this PhD thesis focuses on the generation module in IFC, specifically targeting the image quality assessment, droplet detection, tracking and classification, through image analysis techniques suitable for resource constrained platforms. This involves investigating methods to accurately identify and monitor droplet within the portable IFC.

The motivation of this research, therefore, is to design efficient microfluidic droplet image assessment and classification in resource constrained platform focusing on module droplet generation.

1.2 Challenges in Portable IFC Devices

IFC devices have gained significant attention over the past decade due to their applicability in high-throughput and high content screening [14], [15]. The research trends in this domain of object detection and classification have evolved rapidly, driven by advancements in imaging technologies and embedded systems design. Recent studies have emphasized precise detection and classification using ML, especially deep learning methodologies [16], [17]. Additionally, there is a growing focus on conducting analyses on embedded platforms, facilitating the development of portable applications [18], [19].

This PhD thesis revolves around a resource-efficient pipeline consisting of three primary elements:

1. **Blur-Free Image Acquisition:** This component is responsible for acquiring good quality, blur-free images of microfluidic droplets without necessitating additional hardware.
2. **Image Quality Assessment:** This component is responsible for evaluating droplet image quality automatically prior to classification, allowing only suitable images to be further processed.
3. **Droplet Classification:** Following image quality assessment, the pipeline aims to classify MF droplets into three distinct categories: empty droplets, droplets containing a single cell, and droplets containing multiple cells.

Implementing the above components for blur-free image acquisition, image quality assessment, and droplet classification on resource-constrained devices presents several challenges due to the high computational demands associated both with conventional signal processing and deep learning models.

A first challenge is to avoid using complicated camera synchronization for acquiring blur-free single droplet images. Conventional hardware setups and corresponding algorithms are not suitable for obtaining droplet images that are clear and suitable for subsequent classification, even within the constraints of embedded devices.

Next, achieving fast inference times using DL models while maintaining classification accuracy presents a considerable obstacle. Indeed, most existing research predominantly relies on powerful desktop personal computers (PCs) and powerful graphics processing units (GPUs), which are not suitable for portable applications.

The challenges on resource constrained platform are induced from limited computational power, memory, and power consumption, which can significantly impact the performance and accuracy of classification models. Resource-constrained platforms often lack the computational capacity to support complex ML/deep learning algorithms. This limitation compels to use simplified models for droplet classification. Additionally, the available memory on such platforms can restrict the size of the models that can be deployed, limiting the complexity and number of features that can be utilized for droplet classification. Moreover, droplet classification algorithms must be optimized to minimize energy usage, which can often conflict with the need for more sophisticated and computationally intensive techniques. Hence, if the model is not selected carefully, it may inadequately capture the dynamics of MF droplets, leading to potential classification errors and reduced accuracy.

Overall, addressing these challenges requires innovative approaches in model selection, optimization, and the efficient use of hardware resources, ensuring that droplet classification can be effectively implemented even in constrained environments.

1.3 Problem Statement, Research questions and Research Hypothesis

IFC is a very valuable tool for various biomedical and diagnostic applications. However, as discussed above, the current systems that perform droplet imaging and classification often rely on high-performance computing platforms, which are not suitable for portable, resource-constrained environments. To summarize, the key challenges include: 1) The difficulty of capturing blur-free images in a dynamic, fast-moving MF environment, 2) the high computational cost of the models, which limits their deployment on embedded systems or single-board computers, and 3) the need for a lightweight image quality assessment mechanism to filter out low-quality images, ensuring the reliability of classification models in resource-constrained environments.

The problem this thesis addresses is how to design a blur-free image acquisition, image quality assessment and a lightweight classification model, all of which can operate efficiently on resource-constrained platforms such as Raspberry Pi. The general workflow is shown in Figure 5.

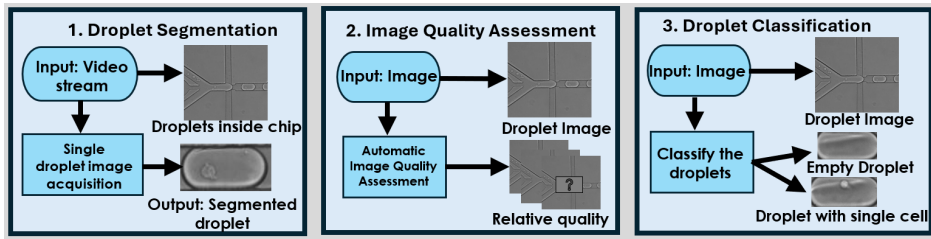


Figure 5: An overview of research task

Segmenting blur-free droplets from the video is crucial step for single droplet analysis as it reduces the possibilities of capturing duplicate image of the same droplet. The second task, i.e. automatic image quality assessment, plays an important role to filter out unnecessary data such as blurry images or low contrast droplet images. This process ensures reliable data analysis. Lastly, classification is important to the analytical process in categorizing specific cellular entities. This overall workflow is designed for portable application. The research presented in this PhD thesis builds upon the following three research questions:

- RQ1: How to design a lightweight and effective image processing pipeline for detecting, counting and acquiring blur-free segmented droplet images suitable for application-specific resource-constrained platforms?
- RQ2: Can deep neural network replace conventional signal processing for achieving effective image quality assessment on a resource-constrained platform?
- RQ3: How to design a custom droplet classifier (presence or absence of cells inside the droplet) suitable for rapid inference on a resource-constrained platform while maintaining high classification accuracy?

The central hypothesis of this research is that it is possible to develop a resource-efficient pipeline for microfluidic droplet analysis that can deliver high classification accuracy and real-time performance on resource-constrained platforms by optimizing deep learning models, reducing image acquisition blur, and implementing lightweight image quality assessment algorithms. Specifically:

- A deep learning model, such as Yolov4-tiny, can be compressed and tuned to run on low-power devices without significant loss of accuracy in droplet classification.
- Computational methods can be developed to ensure blur-free image acquisition in resource-limited environments, reducing the reliance on expensive hardware for motion stabilization.
- An image quality assessment mechanism can be implemented on a resource-constrained platform, which could later be used to filter low-quality images and improve the overall performance of the classification pipeline.

1.4 Thesis Contributions

This thesis makes several contributions to the field of resource-efficient image processing and microfluidic droplet analysis:

Contribution 1 (addressing RQ1): Contrary to most existing systems that rely on complex hardware setups, I propose an image processing pipeline that utilizes minimum hardware and a lightweight algorithm for detecting, counting, and acquiring single object images. Implementation results show that it is suitable for single object image acquisition in IFC on an embedded portable platform

Contribution 2 (addressing RQ2): Image Quality Assessment on Embedded Platforms: I propose a lightweight deep learning solution that replaces conventional signal processing for image quality assessment on embedded platforms. I designed a new lightweight CNN-Transformer hybrid model and implemented it on Raspberry Pi 5. This approach enables objective IQA within strict computational and memory constraints. The result establishes DNN as an effective and resource-efficient alternative for automated, portable IFC systems.

Contribution 3 (addressing RQ3): In order to bridge the gap between high-throughput requirements and the limitations of resource-constrained platforms, I developed a new customized YoloV4-tiny model incorporating model compression and architectural modification (e.g. reduced number of filters from convolutional layers). Results show that the proposed approach enables rapid inference while maintaining high classification accuracy for categorizing droplets containing no, one, or multiple cells.

Table 2 provides an overview of the mapping between the contributions, RQs, and papers. The papers are available in full in Appendices I-V in this thesis.

Table 2: Summary of Contributions in Relation to Research Papers

RQ	Contributions	Publication I	Publication II	Publication III	Publication IV	Publication V
RQ1 (Lightweight droplet images acquisition)	C1	✓			✓	
RQ2 (Image quality assessment)	C2					✓
RQ3 (Droplet classifier on embedded platforms)	C3		✓	✓	✓	

1.5 Organization of the Thesis

This PhD thesis is structured into five chapters;

1. **Introduction:** The introductory chapter provides an overview of the historical background of IFC, problem statement as well as research questions.
2. **Chapter 2:** begins with the background of single droplet image acquisition, image quality assessment and MF droplet classification. Then provides the overview of existing research for portable IFC. Note: the more specific state of the art work related to the individual contributions are presented in the following respective chapters.
3. **Chapter 3:** discusses the overview and challenges in blur free single droplet image acquisition.
4. **Chapter 4:** discusses the overview and challenges for automatic IQA in IFC and suitable model for portable system.
5. **Chapter 5:** discusses the overview and challenges for automatic droplet classification, showcasing proposed model for different embedded platform.
6. **Chapter 6:** provides the summary of the main contributions, presents the conclusion of the thesis as well as the future work.

The framework for the Ph.D. thesis including all the chapters (Chpt.) and corresponding publications is shown in Figure 6.

The chapters provide an overview of the work and results; the reader is encouraged to read the aforementioned papers, provided in appendices, for additional details and results.

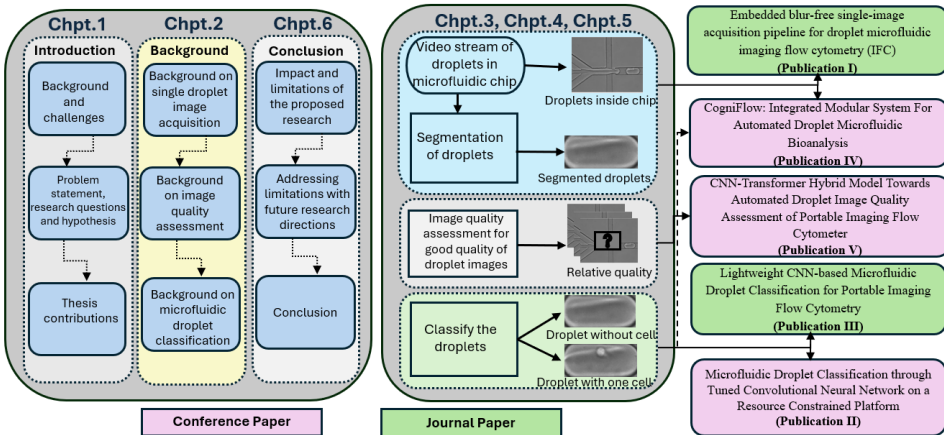


Figure 6: Framework for the Ph.D thesis.

2 SOME BACKGROUND ON BLUR-FREE SINGLE DROPLET IMAGE ACQUISITION, QUALITY ASSESSMENT AND DROPLET CLASSIFICATION IN IFC

As illustrated in Figure 7, in MF droplet IFC, acquiring blur-free single droplet image, performing image quality assessment², and performing droplet classification are important parts of the analysis. Each part has its own challenges and thus directly influence the reliability and precision of data analysis in e.g. biomedical applications. In addition, embedded platforms needed to realized point-of-care services also face challenges in each part, requiring careful selection of models or solution. This chapter provides the fundamental background necessary to understand these challenges and their implications within portable IFC research.

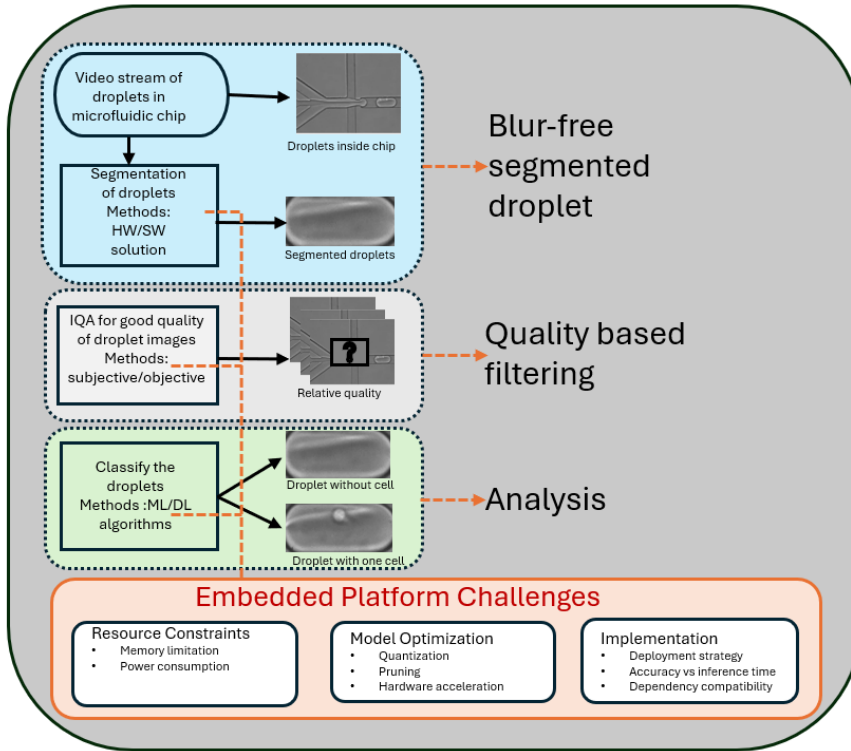


Figure 7: Illustration of MF Droplet Analysis with Embedded Platform Challenges.

Note that while the following sections offer a broad overview of these topics, the specific state-of-the-art methodologies corresponding to each key contribution of this dissertation are presented in their dedicated chapters. The objective of this chapter is to establish the essential background required to develop an understanding towards achieving the PhD thesis' objective.

²Since there is no standard for image quality in this domain, additional expert's opinion was considered as part of the PRG620 project under which this thesis was conducted

2.1 Background on Blur-Free Single Droplet Image Acquisition

Research in droplet-based cell analysis has focused on the development of multiple scientific fields, i.e. cellular immunology [20], drug screening [21], etc. As previously stated, the fluid flow variation in MF IFC leads to motion blur if not synchronized with camera properties. Traditionally high speed cameras and complex optical setups have been used to capture blur-free images from rapid flow. For example, time delay integration CCD camera has been used in high-throughput IFC for blur-free fluorescence images of flowing cells [22]. Stroboscopic illumination can effectively freeze the object motion to obtain blur-free images, for example in [23] the authors used stroboscopic illumination to capture blur-free images with a CMOS camera. In [24], a polygon mirror scanner that moves along with cells to freeze the motion of the blood cells for further classification. However, these approaches are complicated and expensive and requires special optical modules that makes it unsuitable for portable applications.

Software-based methods, including computer vision and deep-learning based single object image acquisition obtained popularity due to their adaptability and lower hardware dependency. Classical segmentation techniques, such as Otsu's algorithm has been used for urine cell segmentation in [25], but this type of algorithm struggle with complex backgrounds and non-uniform lighting conditions. More recent studies have leveraged CNNs for single object image acquisition and provides higher accuracy compared to traditional methods. In [26], the authors used U-Net architecture that takes the features from multiple layers into account and provides good localization for cell level segmentation and counting. Deep CNN has been used to segment fluorescent images of cell nuclei as well as phase images of the cytoplasm of individual bacterial and mammalian cells [27]. Despite these advances, single object image acquisition challenges persist in embedded platforms because of resource constraints. Moreover, further optimizations, such as quantization and pruning, are necessary to make deep-learning models suitable for embedded devices.

Although the scientific literature review shows that the blur-free image can be acquired by developing dedicated hardware system or ML/DL based algorithm for single droplet image acquisition, these are either not suitable for portable platform or not developed for blur-free segmented object image. That is why it is necessary to go beyond the state of the art to develop a pipeline which can segment each droplet and acquire the image without any duplication.

2.2 Background on Image Quality Assessment for IFC

IQA is a key element in image quality evaluation and different types of IQA methodologies are widely applied in image generation and processing. Acquired images from IFC is usually assessed by humans who are familiar with the system. Very limited research has been done to assess the quality of de-blurred or reconstructed images, although those are not fully dedicated for automated IFC system. With the development of computer graphics and digital photography, algorithms have been created to automate the assessment task and explored widely for generic images.

In Figure 15, the hierarchical classification of IQA is represented. The IQA methodologies are mainly classified into subjective and objective; subjective IQA is based on human evaluator assessment, while objective IQA provides quantitative assessment without needing any human evaluation.

These are further classified into three categories: (1) Full-Reference (FR), (2) Reduced-

Reference (RR), and (3) No-Reference (NR) IQA approaches.

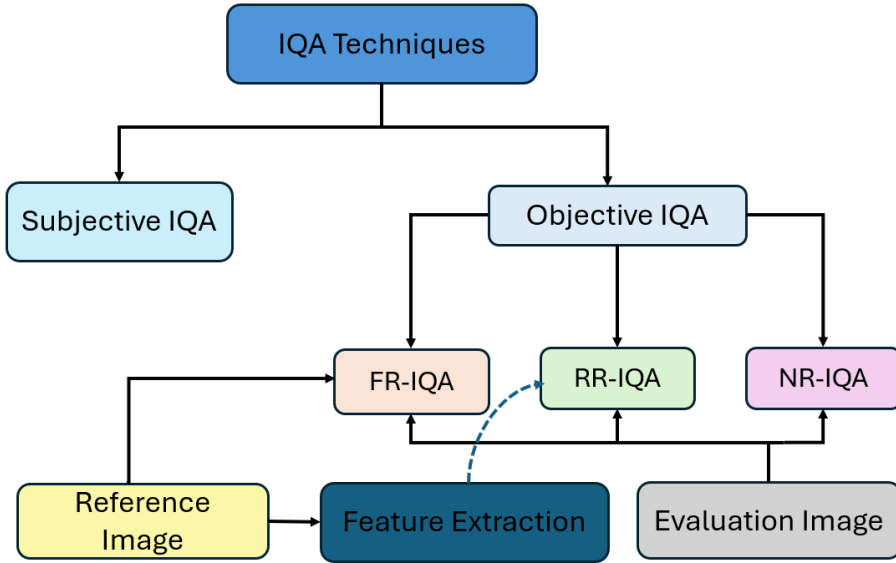


Figure 8: Hierarchical classification of IQA methodologies.

- **Full Reference IQA (FR-IQA):** This approach requires a high-quality reference image to compare with the target image. It involves computing similarity measures such as Peak Signal-to-Noise Ratio (PSNR) [28] and Structural Similarity Index (SSIM) [29]. While FR-IQA provides accurate quality assessment, its dependency on reference images makes it impractical in real-time applications like IFC.
- **Reduced-Reference IQA (RR-IQA):** RR-IQA requires only partial information or specific features from a reference image. It extracts key statistical or structural features from both the reference and evaluation images to estimate quality. This approach offers a balance between accuracy and practicality but still requires some reference data.
- **No-Reference IQA (NR-IQA):** NR-IQA assesses image quality without requiring a reference image. It analyzes inherent features such as sharpness, contrast, and noise to infer image quality. This makes NR-IQA particularly suitable for real-time applications, where reference images are unavailable.

Automated IQA is critical for ensuring consistent and accurate image-based analyses in IFC. NR-IQA methods, such as blind/referenceless image spatial quality evaluator (BRISQUE) [30], natural image quality evaluator (NIQE) [31], [32], and perception based image quality evaluator PIQE [33], have been used to evaluate image quality. However, these metrics often show limited correlation with human perceptual scores. Most of the NR-IQA model are developed based on the publicly available datasets i.e Laboratory for Image & Video Engineering (LIVE) [34], Tampere Image Database (TID2013) [35], KonIQ-10k [36]. LIVE and KonIQ-10k database contains natural distortion datasets, there are other database as well but mostly contains synthetically distorted datasets.

2.2.1 Performance Metrics for IQA

IQA is associated with performance metrics such as Pearson Linear Correlation Coefficient (PLCC), Spearman Rank-Order Correlation Coefficient (SROCC), Kendall Rank Order Correlation Coefficient (KROCC), and Root Mean Squared Error (RMSE). The mostly used performance metrics are PLCC and SROCC, as defined below:

- PLCC: evaluates how well the relationship between two variables can be described by a straight line. The coefficient ranges from -1 to +1, where:
 - +1 indicates a perfect positive linear correlation
 - 0 indicates no linear correlation
 - -1 indicates a perfect negative linear correlation

$$r = \frac{\sum (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum (X_i - \bar{X})^2} \sqrt{\sum (Y_i - \bar{Y})^2}} \quad (1)$$

PLCC is computed using (Eq. 1): where:

- X_i and Y_i are individual data points.
 - $\bar{X}$ and $\bar{Y}$ are the mean values of X and Y .
 - The numerator represents the covariance between X and Y .
 - The denominator normalizes the result using the standard deviations.
- SROCC measures the strength and direction of monotonic relationships between two variables that means it assesses how well the relationship between two datasets can be described by a monotonic function. If one variable increases, the other tends to increase (or decrease) consistently. SROCC can be expressed with (Eq. 2):

$$\rho = 1 - \frac{6 \sum d_i^2}{n(n^2 - 1)} \quad (2)$$

where:

- d_i is the difference between the rankings of the two variables for the i th observation.
- n is the number of data points.

2.3 Background on Droplet Classification

As previously stated, automated classification of microfluidic droplets is essential for various biomedical and chemical applications, including cell encapsulation and single-cell analysis. ML and DL can be used for object classification. A ML based decision maker is used for tumor cell classification in [37]. ML based classification approaches through feature extraction can achieve reasonable but are limited to handle complex characteristics of objects.

Recent advances in DL algorithms and AI has revolutionized object classification, with CNNs achieving SOTA performance. Depending on the available labeled training data, images or objects can be classified into two or more classes using the DL method. A trained ResNet18 model was implemented under the TensorRT framework on the NVIDIA GeForce GTX 1080 TI GPU to classify bead or cell images, total data processing including detection,

tracking within 3ms [38]. Human white blood cell classification and sorting was achieved using FPGA, Nvidia Quadro RTX A6000 GPU with an optimized custom CNN UNet model [39]. In [40], the authors used a faster RCNN architecture to identify and count malaria infected blood cells in bright field image. Unfortunately, the application of DL algorithms to high-throughput IFC in embedded platform faces significant challenges because of the resource constraints.

2.3.1 Performance Metrics for classification

Classification performance evaluation is important to ensure the accuracy and reliability of the classification models. The most commonly used performance metrics in classification tasks are accuracy, precision, recall, F1 score, as described below:

- **Accuracy:** The ratio of correctly classified objects to the total number of objects.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (3)$$

where:

- TP (True Positive): Correctly classified positive objects.
 - TN (True Negative): Correctly classified negative objects.
 - FP (False Positive): Incorrectly classified negative objects as positive.
 - FN (False Negative): Incorrectly classified positive objects as negative.
- **Precision:** Measures the proportion of correctly predicted positive instances among all predicted positive instances.

$$Precision = \frac{TP}{TP + FP} \quad (4)$$

- **Recall (Sensitivity):** Measures how well the model captures positive instances.

$$Recall = \frac{TP}{TP + FN} \quad (5)$$

- **F1-score:** measures the model accuracy by representing the balance between precision and recall and calculated from:

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (6)$$

2.4 Conclusion

This chapter concisely reviewed the key aspects of blur-free single droplet image acquisition, IQA, and droplet classification within the context of MF IFC. The challenges in acquiring blur-free single droplet images were discussed, highlighting the limitations of existing hardware and software-based methods. For IQA, different methodologies, including FR, RR, and NR approaches, were introduced, along with commonly used performance metrics i.e. PLCC and SROCC. The importance of automatic IQA for improving portable MF imaging applications was emphasized. Furthermore, classification techniques based on ML and DL were discussed, with a specific focus on CNN-based methods that offer better accuracy in classification. The evaluation metrics for classification models, such as

accuracy, precision, recall, and F1-score were also included to provide a comprehensive understanding about the model performance.

Despite advancements in these areas, challenges persist, particularly in the optimization of deep learning models for resource-constrained embedded platforms. To address these challenges, the rest of this thesis focuses on developing lightweight algorithms and models that can efficiently operate on portable and resource constrained devices while maintaining sufficient accuracy. By addressing these challenges, the development of portable MF droplet IFC imaging solutions can be improved, benefiting applications in biomedical research and single-droplet analysis.

3 BLUR-FREE SINGLE-IMAGE ACQUISITION IN IMAGING FLOW CYTOMETRY

This chapter presents a lightweight imaging pipeline designed to acquire blur-free single-droplet images in portable IFC. The proposed method requires minimal hardware, enabling efficient single-droplet image acquisition and droplet counting on embedded platforms. The methodology uses color-based detection and Euclidean distance tracking, which is lightweight, making it suitable solution for portable devices.

This chapter is based on the following publications:

Publications I:

Afrin, Fariha & Pärnamets, Kaiser & Le Moullec, Yannick & Udal, Andres & Koel, Ants & Pardy, Tamas & Rang, Toomas, Embedded Blur-Free Single-Image Acquisition Pipeline for Droplet Microfluidic Imaging Flow Cytometry (IFC), IEEE Access, vol. 12, pp. 92431-92441, 2024, DOI: <https://doi.org/10.1109/ACCESS.2024.3421637>

Publication IV:

Jöemaa, Rauno & Afrin, Fariha & Gyimah, Nafisat & Ashraf, Kanwal & Pärnamets, Kaiser & Pardy, Tamas, Cogni-Flow: Integrated Modular System For Automated Droplet Microfluidic Bio-analysis, EUROSENSORS XXXVI, 01-04 September 2024, Debrecen, Hungary.

3.1 Overview

As mentioned in the previous chapters, IFC droplet images are critical for e.g. drug discovery and medical diagnostics. However, high flow speeds within microfluidic channels often lead to motion blur, making it challenging to extract meaningful information from the images. This is because when the camera's exposure time is not well synchronized with the flow rate of the object, it results in either motion blur (if the object moves too fast) or noise (if the exposure time is too long) [41], [42]. Traditional solutions to this problem typically involve high-speed imaging, advanced optical components, and high-end GPU that, while effective, significantly increase both the complexity and cost of the system.

The summarized workflow and challenges in blur-free single droplet image acquisition for IFC are schematized in Figure 9.

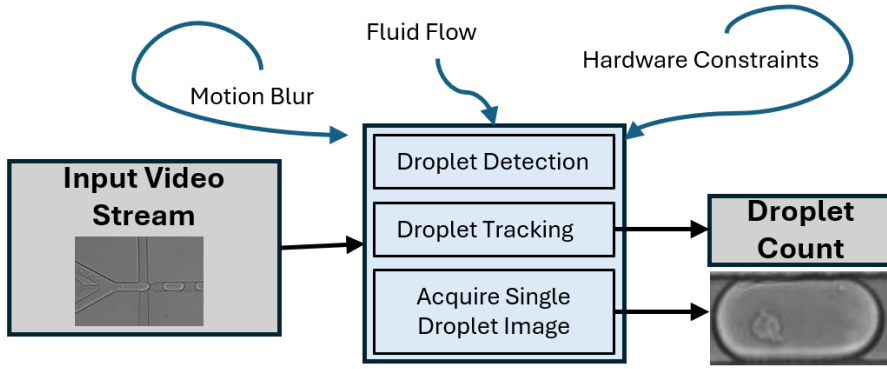


Figure 9: Summarized Workflow and Challenges in Blur-Free Single Droplet Image Acquisition for Imaging Flow Cytometry.

Controlling the fluid flow rate along with a velocity detection subsystem to synchronize the camera's exposure time with the movement of the cells is an existing solution. For ex-ample, the Commercial ImageStream platform focuses on improving the optical setup of the flow cytometry systems as well as uses a precise pump [43]. Another approach is the use of parallelized microchannels to reduce the flow speed, thus ensuring better control over object imaging in high-throughput environments [44], [45]. The optical time-stretch imaging method is another significant advancement in high-speed imaging for IFC. This technique captures high-quality single object image in a highly dynamic environment by exploiting spatial and temporal dispersion [46], [47]. Another method, virtual motion-freezing fluorescence imaging (VIFFI), has been developed to achieve high-throughput imaging by freezing particle motion through an ultra-fast shutter, allowing the capture of blur-free images even at high speeds [48], [49]. However, these solutions require additional hardware with complex optical setups, which makes them complex and costly [50],[51].

In addition to developing a dedicated hardware set-up, the integration of ML and DL techniques has significantly improved the analysis and acquisition of single-object images. These approaches are widely used to process and classify images after acquisition. For instance, image reconstruction can be applied without integrating additional hardware, yet this process first identifies the blur and then processes the data [52]. Although ML and DL have enhanced capabilities in acquisition and further analysis, their computational complexity and hardware demands remain significant challenges. Most studies focus on improving accuracy, often overlooking the critical trade-off between speed, accuracy, and hardware requirements. For example, a portable flow cytometer that uses deep learning to detect *Giardia lamblia* cysts in water samples was reported, but data processing still relies on desktop PCs, limiting the applicability in portable IFC [10].

The number of research papers directly comparable with our work are limited; the closest ones and their key features are summarized in Table 3.

Table 3: Comparison of different techniques and the proposed method. Each key element (complexity, portability, etc.) used for the comparison is scored on a scale ranging from "+" to "+++". (Table reproduced from publication 1)

Techniques	Complexity	Portability	Throughput/accuracy	Cost	Summary of the methodologies
Virtual-freezing [49]	+++	+	Throughput >10,000 cells/s; classification accuracy 95.3%	+++	Acquires blur free single cell image by using a sophisticated system that includes a flow-controlled microfluidic chip, a speed-controlled polygon scanner, and a series of precise timing control circuits.
Deep learning [52]	+	+	96.6% classification accuracy	++	Data processing runs on the computer
Machine learning [53]	+++	++	Detection rate 93.8%; classification rate 99.8%	+++	Optofluidic detection at the edge device relies in the time trace
Deep learning [10]	+	+++	Autofocus on a single object in <7 ms	++	Data processing runs on the computer
OpenCV [19]	+	+++	Classification and counting accuracy are above 95%	+	The detection and counting system operate after saving 10 videos (RPI4 platform). However, the processing time for each object detection is not reported in the paper.
Proposed method	+	+++	Counting accuracy is 100% for the tested videos	+	Acquires blur free single droplet image by using a lightweight pipeline on embedded platform

In [49] virtual-freezing technique uses a sophisticated hardware system, including a flow-controlled microfluidic chip, a speed-controlled polygon scanner, and precise timing circuits to capture blur-free single-cell images at high throughput (>10,000 cells/s) with 95.3% classification accuracy. Another work [52] applies deep learning-based classification and achieves 96.6% accuracy rate but the data processing is performed on a computer. Despite their effectiveness, they are highly complex, expensive, and lacks portability.

In [53] a machine learning model was used for edge-device classification, achieving 99.8% classification accuracy with a 93.8% detection rate. The method relies on precise time-trace data, making it highly complex but moderately portable. Another [25] deep learning-based method, focusing on autofocus capabilities (≈ 7 ms per object) rather than just classification and data processing runs on a computer.

In [10], OpenCV is used to perform classification and counting with an accuracy of

>95% on a RPi-4. While it is moderately complex, it is highly portable and cost-effective, making it a practical alternative. However, the processing time per object detection is not reported. Additionally, many studies focus on post-experimental data processing, which is often impractical in real-time applications due to the limited memory capacity of embedded systems such as single-board computers (SBCs). This makes the application of real-time, edge-based solutions challenging, particularly in high-throughput IFC, where large numbers of images must be captured and processed simultaneously.

The main goal of this part of this PhD thesis is to acquire blur-free single droplet images without any camera synchronization to handle the complexity and computational cost challenge. This method introduces a lightweight pipeline on an embedded platform to achieve blur-free single droplet imaging with 100% accuracy in tested videos. The proposed method offers high portability, low cost, and low complexity, making it a promising alternative to traditional IFC imaging solutions.

3.2 Method

The proposed pipeline was first tested with simulated two-phase flow videos, and then with real-world experimental fluorescent microscopy recordings. To provide an understanding of these testing phase, a detailed explanation is provided here. The core microfluidic chip made of a polydimethylsiloxane (PDMS), schematically represented in Figure 10(a) (Figure reproduce from publication 1). This chip's design is based on a previously validated microfluidic platform documented in reference [54]. The chip's architecture incorporates several components: water and oil inlets, a gas spring connector (labeled as "air inlet" in Figure 10(a)), a flow-focusing junction measuring 90 μm in width and 100 μm in height, and a single outlet channel. The filtration systems integrated into the oil inlets, specifically to capture and remove any contaminants potentially present in the mineral oil (Sigma Aldrich 330779), thereby preventing channel obstruction and ensuring consistent fluidic flow.

Simulated videos are generated using COMSOL Multiphysics version 5.6, implementing a sophisticated two-phase flow model previously established in [54]. The simulation parameters are carefully calibrated to maintain water and oil flow rates at precisely 10 $\mu\text{L}/\text{min}$ and 67.7 $\mu\text{L}/\text{min}$, respectively. This 6.77:1 ratio of oil to water flow rate serves a critical function: it ensures sufficient spatial separation between generated droplets. Within the COMSOL visualization framework (illustrated in Figure 12) the liquid phases are represented through a color-coded system - red and blue for the respective phases, with green indicating interfacial regions where the relative water concentration approaches 50%.

The experimental validation phase uses a significantly more complex setup. Deionized water is precisely formulated with 10 $\mu\text{g}/\text{ml}$ fluorescein isothiocyanate-dextran (FITC) to enable fluorescent visualization. This fluorescent aqueous solution and oil are introduced into the fabricated PDMS chip through their respective inlets using calibrated 3 mL syringes. Flow rate control is achieved through a high-precision syringe pump (SpinSplit Netpump, manufactured by SpinSplit LLC, Budapest), with the notable absence of the air inlet in this experimental configuration. The optical detection system, depicted in Figure 10(b), uses a 488 nm laser light source (Sharp GH04850B2G equipped with specialized focusing optics sourced from AliExpress) positioned strategically at the post-junction region where droplet formation occurs, as described in reference [31]. This positioning is critical for optimal excitation of the FITC fluorophore. The dynamic droplet flow through the PDMS microchannels is captured using a high-speed Basler Ace camera operating at an

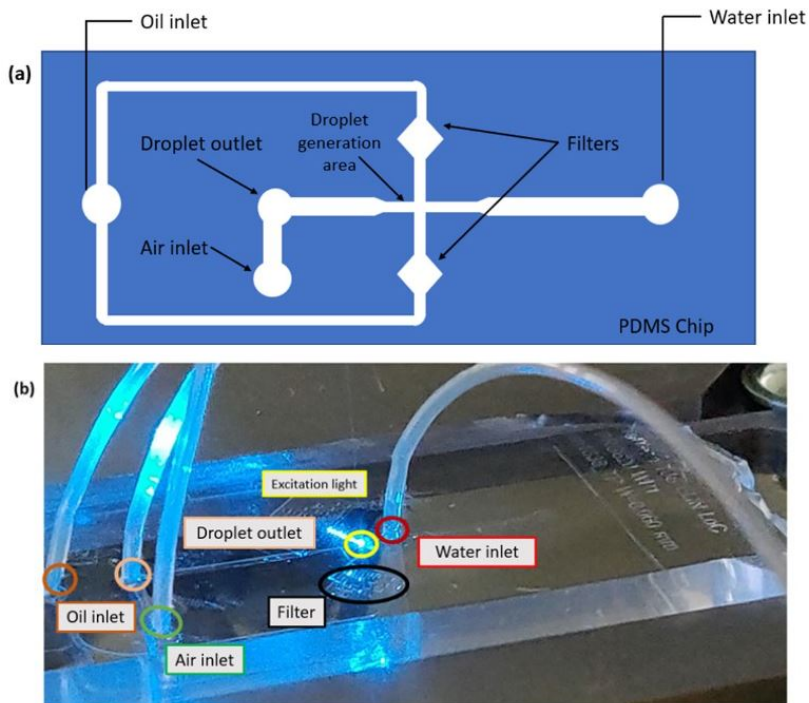


Figure 10: Droplet generation. (a) Schematic of the PDMS chip (not drawn to scale). (b) Experimental setup showing all (water, oil, air) inlets, the droplet outlet, the filter, and the excitation laser light. The power supply for driving the laser, the computer for controlling the syringe pump, and the Basler Ace camera are not visible here. Figure reproduced from Publication 1 (Figure 1).

exceptional temporal resolution of 1000 frames per second (fps), enabling detailed analysis of rapid droplet formation and movement dynamics. This dual-validation approach, combining both computational simulation and experimental verification, provides a robust framework for assessing the proposed pipeline's performance across different operational scenarios and conditions.

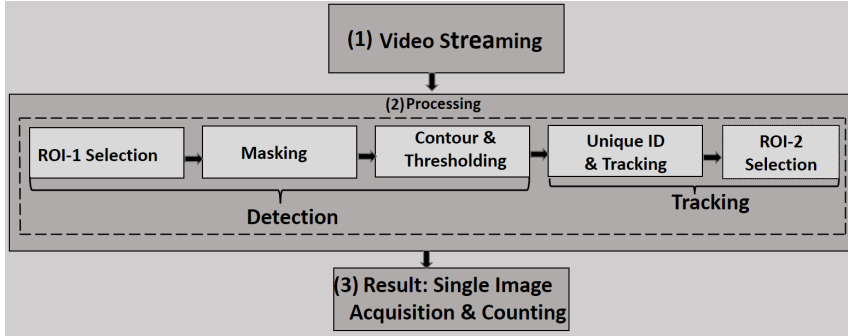


Figure 11: Block diagram of the single-image acquisition process flow, consisting of (1) video streaming, which occurs at 1000 fps; (2) processing, which involves detecting droplets based on their color and tracks; and (3) result acquisition, which involves acquiring a single image and counting the number of droplets using the tracking result without any duplication. Figure reproduced from Publication I (Figure 2).

Figure 11 presents a detailed block diagram of the single-image acquisition work flow, consisting of three main stages: (1) video streaming, (2) processing, and (3) result acquisition.

The pipeline was developed in Python 3.8 with OpenCV 4.5.5, incorporating custom color-based detection with Euclidean distance tracker rather than Background Subtraction Mixture of Gaussians-2 (BS-MOG2). This design choice was made because color-based detection demonstrated superior robustness against intensity variations in experimental environments, whereas BS-MOG2 required specific history (100) and threshold (50) values and performed poorly in non-simulated conditions. The pipeline begins with **video streaming**, droplet movement through microfluidic channels. Then the **processing** stage which forms the core of the system and is divided into two main components:

1. **Detection:** This includes:

- **ROI-1 Selection:** Defining an initial region of interest for droplet detection
- **Masking:** Implementing color-based detection using HSV color space
- **Contour Thresholding:** Applying a threshold value (100 pixels) to exclude noise and identify valid droplet contours

2. **Tracking:** This includes:

- **Unique ID & Tracking:** Assigning and maintaining unique identifiers to each droplet using a Euclidean distance tracker algorithm
- **ROI-2 Selection:** Defining a secondary region of interest for counting. More details are included later in Figure 12.

The tracker maintains positional proximity (less than 10 pixels) between frames which helps to prevent duplication. The algorithm calculates the Euclidean distance

as follows [55]:

$$D = \min \left(\sqrt{(a_j - a_i)^2 + (b_j - b_i)^2} \right) \quad (7)$$

where D represents the Euclidean distance between previous and current tracks, a denotes the x-coordinate of the track box, b denotes the y-coordinate of the track box, j represents prediction boxes in the current frame, and i represents prediction boxes in previous frames. This tracking approach is particularly effective in the initial section of the microfluidic channel, where droplets flows in a single flow pattern without occlusion.

3. **Result acquisition:** this includes capturing single images of each droplet (that can be exported in various formats (.tiff, .png, .gif, .jpeg, .bmp) for subsequent analysis) and accurately counting the total number of droplets without any duplication. The detailed process can be understood from the following pseudo code.

The memory usage, starting near 0 MIB and increasing gradually to approximately 60 MIB. This represents the loading of the video processing libraries, initialization of data structures, and establishment of the processing pipeline. The maximum memory utilization remains below 100 MIB, indicating efficient resource utilization suitable for deployment on resource-constrained platforms like SBCs. This memory profile complements the previously discussed processing pipeline, representing that the implementation is not only effective but also resource-efficient, aligning with the overall goal of developing lightweight image processing solutions for microfluidic applications.

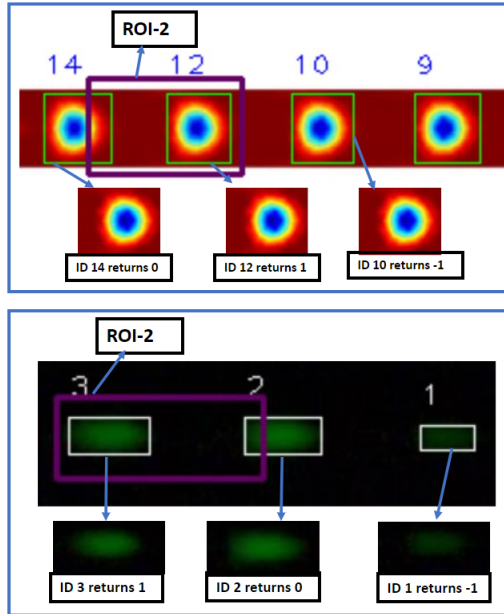


Figure 12: Three different conditions for accurate counting (Top: COMSOL simulated two-phase flow video, bottom: experimental fluorescence video): (i) if the green bounding box is outside ROI-2, the operation returns -1; (ii) if the bounding box is exactly on the boundary of ROI-2, it returns 0; and (iii) if the bounding box is inside ROI-2, it returns +1, which is the right condition to store the location. Figure reproduced from Publication 1 (Figure 5).

Algorithm 1 Pseudo Code for Proposed Pipeline. Reproduced from Publication I(Algorithm 1)

Data: Video frame

Result: Acquire single droplet image and count

```
1. tracker = EuclideanDistTracker() // Initialize Euclidean distance tracker
2. create_directory('data') // Create a directory named 'data' for saving droplet images
3. object_count = set() // Initialize droplet count set
4. Process frames in a loop
while True:
    frame = read_frame(cap) // Read a frame from the video
    Hsv_img = convert_to_hsv(frame) // Convert the frame to HSV
    mask_green = create_green_mask(Hsv_img, Low_Green, High_Green) // Create
mask using color thresholding
    contours = find_contours(mask_green) // Find contours in the mask
    detections = detect_objects(contours) // Detect droplets based on contour area
    boxes_ids = tracker.update(detections, 1) // Update droplet tracking using the
Euclidean distance tracker
5. Process each detected and tracked object
    for box_id in boxes_ids:
        x, y, w, h, obj_id = box_id
        draw_rectangle(frame, x, y, w, h) // Draw bounding box
        draw_text(frame, str(obj_id), x, y) // Text on the frame
        result = point_in_polygon(roi_polygon, (x, y)) // Check if the droplet is inside the
defined counting area
        if result:
            object_count.add(obj_id) // Update droplet_count based on the object ID
            save_droplet_image(frame, obj_id) // Save image of the detected droplet
            total_count, objects_in_second = calculate_count(object_count, fps_count) //
Calculate total droplet count and droplets per second
            display_count(frame, total_count) // Display the droplet count on the frame
6. Check for user input to exit the loop
    key = wait_for_key()
    if key == 27:
        break
7. Cleanup and release the video capture object
    release_video_capture(cap) // Release the video capture object
    destroy_all_windows() // Destroy all open windows
End of Algorithm
```

Figure 12 illustrates the implementation of region-based logic for accurate droplet counting across COMSOL-simulated video with color-coded droplets (numbered 9-14) and experimental fluorescence videos with green bounding boxes identifying droplet positions (numbered 1-3). The figure demonstrates three distinct scenarios that determine the counting operation:

1. When the droplet bounding box is positioned outside ROI-2, the operation returns -1, indicating the droplet should not be counted.
2. When the droplet bounding box is positioned completely inside of ROI-2, the operation return +1, representing the correct condition for storing the droplet location.
3. When the droplet bounding box is positioned overlapping with ROI-2, the operation returns 0, indicating the droplet should not be counted.

This conditional logic ensures that each droplet is counted exactly once as it passes through the designated region of interest, preventing duplicate counts.

3.3 Selected results of the single droplet image acquisition

As previously discussed, two algorithmic combinations were tested: (1) color-based detection with Euclidean distance tracker, and (2) BS-MOG2 with Euclidean distance tracker. These combinations were evaluated using two types of video data, simulated and experimental, on three hardware platforms: a desktop PC, a Raspberry Pi 4 (RPi4), and a NVIDIA Jetson Nano.

The analysis shows that the color-based detection with Euclidean distance tracker approach significantly outperforms the BS-MOG2 with Euclidean distance tracker approach across all platforms, offering faster processing times and demonstrating greater suitability for portable IFC applications. For simulated video data, the color-based method achieved processing times as low as 10 milliseconds on the Jetson Nano, whereas BS-MOG2 required approximately 30 milliseconds on the same platform. More detail results are shown in Publication I.

This performance analysis validates the selection of color-based detection as the preferred algorithm for resource-constrained platforms, aligning with the goal of acquiring single droplet image in portable IFC.

Table 4: Comparison of Most-closely related results vs. our work. Table reproduced from Publication I (Table 4).

References	Techniques	Platform	Processing time (ms)	Counting accuracy
[56]	Deep learning	High-end NVIDIA V-100 GPU	33	N/A
[49]	Sophisticated hardware setup	Desktop PC	N/A	N/A
[10]	Deep learning	Nvidia RTX 2080 GPU	~34	N/A
[19]	OpenCV	Raspberry Pi	N/A	95.9% to 99.7%
This work	Lightweight pipeline using color-based detection and Euclidean distance tracker	Jetson Nano	10	Counting accuracy is 100% for the tested videos

Moreover, Table 4 contrasts our proposed lightweight pipeline and other most-closely-related state-of-the-art work for droplet detection and counting based on utilized platform, processing time and accuracy.

- **Processing Time:** Our lightweight pipeline implemented on Jetson Nano achieves a processing time of 10 ms per frame, which is approximately 3.3-3.4 times faster than deep learning approaches running on high-end GPUs ([56], [10]) that require 33-34 ms per frame respectively. Copy
- **Counting Accuracy:** While many previous works ([56], [49], [10]) do not report counting accuracy, openCV-based detection and classification implementations on Raspberry Pi platforms ([19]) reported 95.9% to 97.7% accuracy. While our approach obtains counting accuracy more than 95%.
- **Platform:** Our work specifically targets resource-constrained edge computing devices (Jetson Nano), whereas most comparable studies utilize high-performance GPUs or desktop computers, making our solution more suitable for portable IFC.

This comparative analysis represents the balance between performance and resource efficiency achieved by our pipeline, particularly for deployment in portable IFC.

Additionally, the pipeline was also evaluated using an external reference brightfield microscopy video obtained with a similar chip design [57], successfully demonstrating single-droplet detection, image acquisition and counting shown in Figure 13.

This evaluation indicates the generalization and robustness of the developed pipeline across different devices and video types.

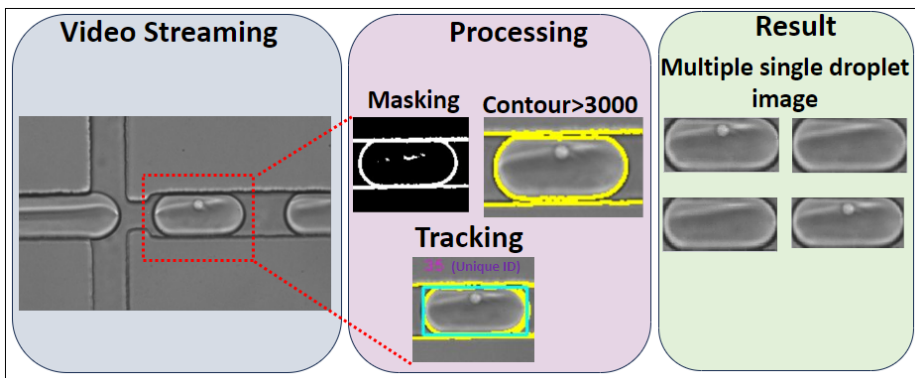


Figure 13: Single droplet image acquisition in brightfield microscopy (Left to right: video streaming, processing, and result). Masking (white boundary), contouring (yellow boundary) and tracking (green rectangle) shows only for one droplet from the video. Multiple saved droplet images are shown in result. Figure reproduced from Publication 1 (Figure 8).

3.4 Conclusion on Blur-Free Single-Image Acquisition in Imaging Flow Cytometry

Single-droplet image acquisition plays an important role in extracting meaningful information about individual micro-particles. Due to the high throughput and rapid flow rate in IFC, most research efforts have focused on developing hardware-based solutions to capture blur-free single images. In contrast, the goal of this work was achieved with lightweight pipeline which, requires minimum hardware, making it suitable for portable IFC application.

The pipeline employs a color-based detection algorithm that efficiently acquires each single-droplet images while simultaneously counting the total number of droplets from a single-line high-throughput droplet flow. The extracted images can be seamlessly integrated into other algorithms for further analysis, such as classification, or any other morphology analysis. Droplet detection presents challenges due to the variation of fluid flow in the portable platform, but the proposed pipeline effectively mitigates these issues, ensuring reliable detection, counting and single droplet image acquisition.

This fully automated system detects and counts droplets while capturing blur-free single-droplet images within 2ms. The initial phase requires a one-time manual setup for identifying the object's color range and defining the threshold for mask creation. The pipeline was validated using external reference data. Additionally, it was tested on a desktop PC and two embedded platforms, comparing the processing times of different object detection algorithms. The Jetson Nano SBC achieves a detection rate of 100 droplets per second (DPS), whereas a desktop PC processes droplets approximately five times faster (500 DPS).

4 AUTOMATED DROPLET IMAGE QUALITY ASSESSMENT FOR PORTABLE MICROFLUIDIC IMAGING FLOW CYTOMETRY

This chapter explores the suitability of IQA techniques and the challenges involved in dataset preparation within the context of a portable LoC microfluidic IFC system. The primary objective is to assess the applicability of IQA methodologies for portable IFC applications. The evaluation results presented in paper IV demonstrates the proposed model's effectiveness, exhibiting good correlation coefficients with human Mean Opinion Score (MOS) and inference capabilities on RPI-5 resource constrained platform.

This chapter is based on the following paper:

Publication V:

Afrin, Fariha, Ndubuisi Ezechukwu, Dismas, Le Moullec, Yannick, Pardy, Tamas, Rang, Toomas, "CNN-Transformer Hybrid Model Towards Automated Droplet Image Quality Assessment of Portable Imaging Flow Cytometer," Accepted to be presented at and published in the proceedings of IEEE EUROCON 2025 - 21st International Conference on Smart Technologies, 4-6 June 2025, Gdynia, Poland.

4.1 Overview

As previously discussed, IFC plays important role in biomedical imaging by integrating the high-throughput capabilities of flow cytometry with detailed morphological analysis through imaging [58]. However, ensuring good-quality image acquisition remains a critical challenge due to variations in lighting conditions and motion blur [52], [59]. Poor image quality can lead to inaccurate analysis in subsequent processing steps, potentially affecting the accuracy.

Furthermore, in IFC, image quality is typically assessed subjectively, which is inadequate for automated workflows in analysis. Implementing automated IQA is essential to ensure consistent and precise data interpretation. A reliable IQA system in portable platform directly impacts the accurate analysis of microparticles such as cells within droplets, thereby enhancing the robustness and reproducibility of IFC-based analysis where computational efficiency is a primary constraint.

Figure 14 represents a flow chart depicting an automated IQA process for IFC. The process begins with an input image, which undergoes a quality assessment. If the image meets the predefined quality criteria, it proceeds to further analysis; otherwise, it is discarded, thereby reducing unnecessary analysis and improving the reliability of IFC-based applications.

Traditional subjective assessment of image quality lacks consistency and reproducibility, motivating the need for automated IQA methods. Recent research has increasingly focused on developing model based on CNN for no-reference image quality assessment (NR-IQA) [60] due to their capability to learn complex features directly from the image without requiring a reference image, while maintaining strong correlation with human opinion scores.

In the field of natural image processing, extensive research has been conducted on NR-IQA methods, particularly for analyzing complex structures, with multiple studies comparing their performance across various datasets [61]. However, despite these advancements, no dedicated research has been undertaken specifically for IFC applications

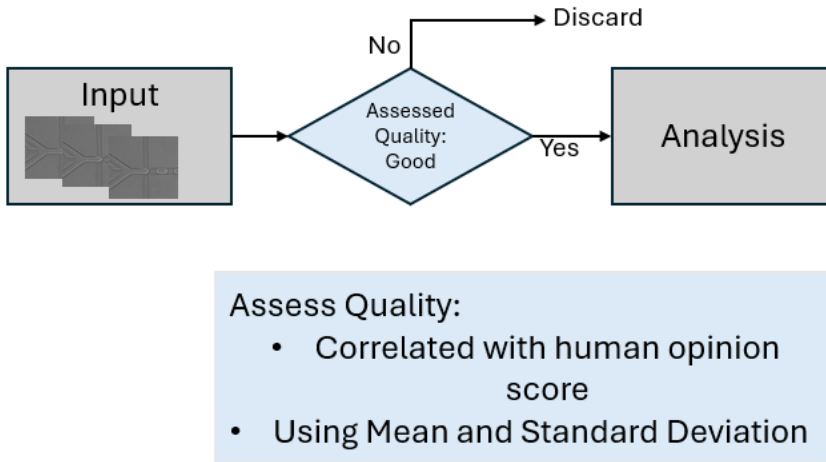


Figure 14: Automated Image Quality Assessment for Imaging Flow Cytometry.

and existing research primarily focuses on general image restoration and deblurring quality assessment [62], [52]. This leaves a significant research gap in the applicability of CNN-based IQA models for IFC, particularly their deployment on portable platforms. Addressing this gap could significantly enhance the accuracy and efficiency of IFC analysis in real-world, portable applications, reinforcing the need for further advancements in automation, computational optimization, and real-time processing in constrained environments. Thus, human-centric IQA techniques that can better align with subjective assessments in IFC are highly desirable.

Quality assessment criteria are typically based on correlation with human opinion scores and statistical measures such as mean and standard deviation. While MOS is a widely used metric for evaluating perceived image quality, it remains insufficient for capturing the full extent of subjective diversity in human assessments. To address this limitation, the Distribution of Opinion Scores (DOS) provides a more comprehensive and objective evaluation, ensuring greater consistency and reliability in subsequent image analysis.

In [63], the authors emphasize the importance of DOS acquisition, aiming to comprehensively evaluate image quality by automatically predicting DOS rather than relying solely on MOS. Building upon this approach, this chapter aims to obtain both the quality score and DOS, ensuring strong correlation with subjective assessments, thereby enhancing the robustness of automated image quality evaluation in IFC.

In [62], the authors introduced a U-Net-based CNN model for ovarian cell image restoration aimed at addressing the challenges posed by high flow speeds and reduced exposure times in IFC. The goal to enhance image quality under these conditions by compensating for the distortions caused by motion blur and limited exposure. The restored images were then assessed using the Blind/Referenceless Image Spatial Quality Evaluator (BRISQUE), the Naturalness Image Quality Evaluator (NIQE), and the Perception-based Image Quality Evaluator (PIQE). These are useful in quantifying image quality, but have shown limitations when compared to human perceptual scores.

Image deblurring is an important research in IFC [64], researchers developed a deep learning-based classifier combined with a deblurring system for red blood cell analysis [65]. The image quality was assessed with peak signal-to-noise ratio (PSNR) and struc-

tural similarity index measure (SSIM), both of which require pristine reference images to quantify blur and structural distortions. However, in real-world imaging scenarios, high-quality reference images may not always be available, limiting the applicability of these metrics.

In [52], gradient-based method was used for the quantitative assessment of deblurred images which estimates image sharpness by analyzing edge differences, making it particularly effective for fluorescence images. However, this is not suitable for low-gradient images with weak contrast and inherent blurriness, leading to inaccurate sharpness measurements. These limitations further emphasize the need for developing IQA method for IFC application, particularly in portable imaging environments.

4.2 Method

The initial selection of the models focused on architectures with low computational requirements to ensure compatibility with embedded platforms while maintaining high performance. Three lightweight CNN-based models were selected and evaluated on our dataset: MobileNet [66], EfficientNet, and NasNet [67]. Among these, MobileNet demonstrated comparatively better performance; however, more optimization was necessary to make it suitable for assessment applications. The whole workflow integrates data pre-processing, training, and deployment in embedded platform specifically optimized for portable droplet MF IFC applications.

4.2.1 Architecture for Microfluidic Droplet Image Quality Assessment

Figure 15 represents an end-to-end framework for predicting the quality of MF droplet images through a hybrid deep learning approach. The framework begins with data processing, where original droplet images undergo multiple augmentation techniques (motion blur, gaussian blur, contrast variation, rotation transformation, and cropping operations) to enhance the dataset diversity and model robustness. This enables the model to generalize effectively under various image-quality conditions encountered in real-world MF experiments.

4.2.2 Training

The produced dataset was split into training and validation dataset:

- 90% of the processed data was allocated as training set;
- 10% of the processed data was allocated as validation set.

As the total number of images in the dataset is not large, this split ensures sufficient data for model training while keeping adequate data for validation, which helps to assess generalization performance.

The architectural design (bottom section of the figure) represents neural network structure with several key components:

1. **Input Layer:** Accepts $224 \times 224 \times 3$ image tensors, which is standardized dimension.
2. **Feature Extraction Backbone:** Uses MobileNetV2 pre-trained on ImageNet with frozen initial layers.
3. **Global Average Pooling:** Reduces spatial dimensions to create a vector, making it suitable for input to the transformer block.

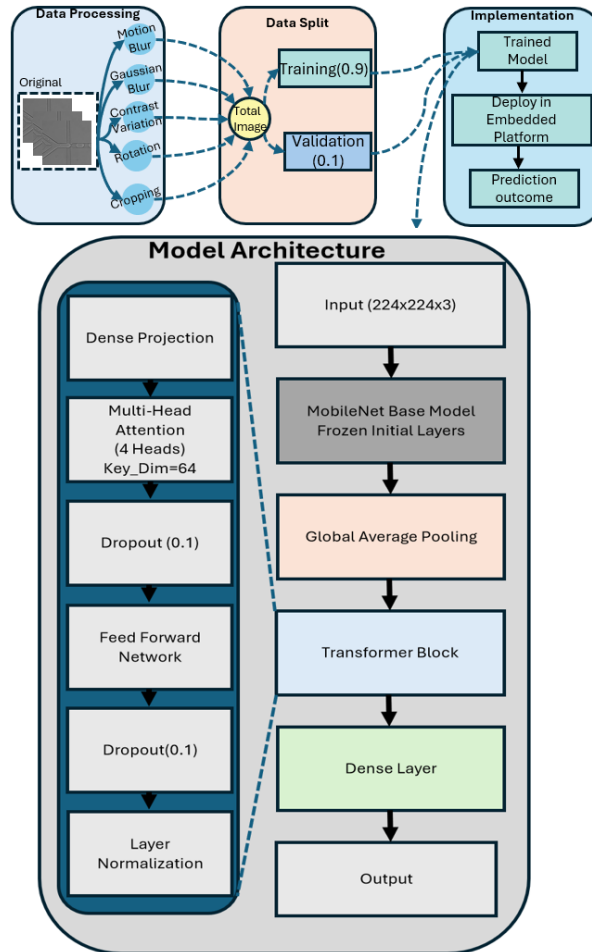


Figure 15: Proposed framework for predicting the quality of MF droplet images. Expanded version of Figure 1 in Publication V.

4. **Transformer Block:** The transformer block consists of a multi-head self-attention mechanism with four attention heads, and a feed-forward network with 512 units followed by dropout for regularization to capture long-range dependencies between image regions, which is particularly valuable for assessing overall image quality.

The expanded view of the 'Transformer Block' further details:

- Initial dense projection takes the data into a higher-dimensional space;
 - Multi-headed attention mechanism with 4 heads allows the model to focus on different parts of the input simultaneously;
 - Dropout regularization (0.1) for overfitting prevention;
 - Feed-forward network with non-linear activations;
 - Second dropout layer (0.1) for additional regularization;
 - Final layer normalization for stable training.
5. **Dense Layer:** Provides final non-linear transformations before classification.
 6. **Output Layer:** Produces the final quality prediction score.

4.2.3 Implementation

For implementation, TensorFlow along with Keras version 2.10.0 was used, running on an Intel 12th Gen Core i9-12900K processor (3.2 GHz, 16 cores). Upon completing the training process, the model was converted to TensorFlow Lite, an optimized version of TensorFlow designed for mobile, embedded, and edge devices. This conversion ensured efficient performance and seamless compatibility with the RPI-5 SBC.

The overall framework represents the process of accurate quality assessment of MF droplet images with minimal computational resources, making it suitable for integration into portable MF system.

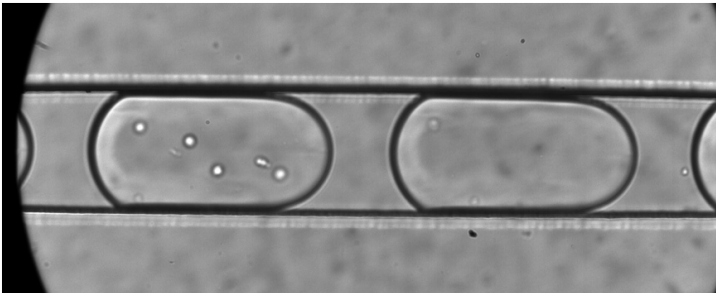
4.3 Selected results on automated droplet image quality assessment for portable microfluidic IFC

As discussed earlier, MobileNet showed comparatively better results, yet not suitable for successful IQA. The modification and fine-tuning struggled with generalization. Attempts to improve generalization through hyperparameter tuning, transfer learning, and dropout regularization did not lead to a significant improvement. To overcome this limitation, a lightweight transformer block was integrated into the modified MobileNet architecture, which effectively minimized the generalization gap.

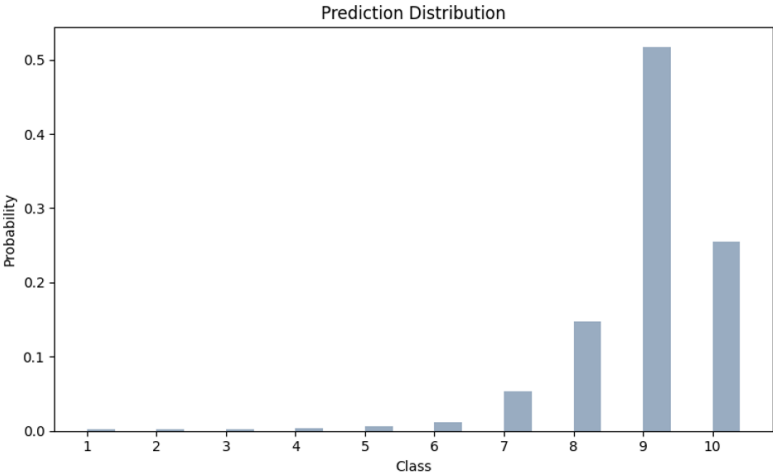
Table 5 compares the complexity (number of parameters) and performance of different IQA models using three evaluation metrics: SRCC, PLCC, and EMD. In [70], the SRCC and PLCC scores are highest: 0.88 and 0.86, respectively, but it has a very high parameter count, 152.45M, making it computationally expensive. In [71], [68] achieve high SRCC and PLCC values (0.86 and 0.84) with lower complexity compared to [70]. AlexNet [69], a well-known deep learning model, underperforms compared to other models, with a SRCC of 0.61. The proposed model has significantly lower complexity as it has 4.05M parameters, while achieving a reasonable performance (SRCC: 0.73, PLCC: 0.75). The EMD value (0.11), potentially indicating considerable performance. This table highlights the trade-off

Table 5: Comparison of Model Complexity (number of parameters) and Performance with State-of-the-Art IQA Models Based on SRCC, PLCC, and EMD. Reproduced from Publication V (Table 1).

Model/Method	Number of Parameters (in millions (M))	SRCC	PLCC	EMD
BRISQUE [30]	-	0.66	0.68	-
TTL-IQA [68]	-	0.86	0.84	-
AlexNet [69]	60	0.61	0.67	-
TReS [70]	152.45	0.88	0.86	-
HyperIQA [71]	27.37	0.86	0.84	-
Proposed	4.05	0.73	0.75	0.11



(a)



(b)

Figure 16: (a) Test image (example of a good image quality), and (b): Class prediction distribution of the image ($MOS \pm STD = 8.881 \pm 1.095$). Reproduced from Publication V (Figure 6).

between model complexity and performance, showcasing the efficiency of the proposed model implemented in MF droplet IQA compared to larger architectures.

Figure 16 consists of two subfigures illustrating the evaluation of image quality and its corresponding prediction distribution for proposed model. Figure 16(a) shows a droplet image representing a good-quality test image, as indicated through sharp boundaries, well-defined droplet borders, and minimal noise artifacts. Figure 16(b) represents the class prediction distribution for the image quality assessment model. The horizontal axis represents quality classes (ranging from 1 to 10), while the vertical axis represents the probability of the image belonging to each class. Class 9 exhibits the highest probability in distribution plot, followed by class 8 and class 10, suggesting the model is most confident in classifying the image as higher categories. $MOS \pm STD$ is 8.881 ± 1.095 , suggesting that the model rates this image as high quality.

Finally, the memory consumption distribution of different components within the proposed model is evaluated when processing an input resolution of 224×224 with a batch size of 1. Modified MobileNet is the most intensive part of the model and consumes largest portion of memory which is (91.0%) as it is responsible for feature extraction. Then GAP consumes 8.26% of memory usage. Transformer block and dense layer collectively consumes less than 1% of the 56.39 MB. The achievable inference time on RPI-5 is 110 milliseconds which corresponds 9 FPS. This result highlights the feasibility of deploying the model on resource-constrained portable systems.

4.4 Conclusion on automated droplet image quality assessment for portable microfluidic IFC

This chapter presented a hybrid IQA model for portable MF IFC. To address the challenges of IQA in embedded environments, three different lightweight CNN-based models were evaluated: MobileNet, EfficientNet, and NasNet. Because of the limited dataset, transfer learning was incorporated during model training. MobileNet performed best, yet suffered with generalization issue even after applying dropout regularization.

To mitigate these limitations, a transformer block was integrated, enabling the model to capture global features and reduce the gap between training and validation loss. Early stopping was also used to minimize computational costs. The proposed modified MobileNet-Transformer hybrid model stabilized significantly faster, converging at approximately 110 epochs. The model achieved a SRCC exceeding 70%, demonstrating strong correlation with human quality assessments. Furthermore, we deployed the model in RPI-5 and measured the inference time 110 ms, which yields 9 FPS.

Overall, the proposed hybrid IQA has the potential to significantly assess the quality of the MF droplet image. The findings in this chapter provide important insights into selecting NR-IQA techniques for portable IFC applications. This contribution could lead to advancement in automatic image assessment in portable IFC as well as could be used to get feedback for adjustment to camera parameters to ensure good quality image acquisition.

5 CLASSIFICATION OF MICROFLUIDIC DROPLET ON RESOURCE CONSTRAINED PLATFORM

This chapter discusses microfluidic IFC droplet classification through the empirical optimization of a deep learning architectures specifically customized for resource-constrained computational environments. The chapter addresses the research gap, challenges, and strategies involved in developing an efficient model capable of classifying microfluidic droplets. This work focuses on the need to balance speed and accuracy while meeting the constraints imposed by limited computational resources, making it a contribution to portable IFC applications.

The work presented in this chapter constitutes an advancement in translating traditionally resource-intensive deep learning approaches into deployable solutions for point-of-care devices for microfluidic droplet applications where computational infrastructure is severely limited. The work bridges the considerable gap between laboratory-based high-performance computing implementations and portable microfluidic systems. The empirical optimization strategies explore model compression techniques, parameter-efficient architectural modifications, and inference acceleration validated against experimental microfluidic datasets to ensure robustness. The analysis presented in this chapter provides valuable insights for researchers and practitioners working at the intersection of microfluidics, IFC, DL, and portable devices.

This chapter is based on the following publications:

Publication II:

Afrin, Fariha & Le Moullec, Yannick & Pardy, Tamas, Microfluidic Droplet Classification through Tuned Convolutional Neural Network on a Resource Constrained Platform. 19th Biennial Baltic Electronics Conference (BEC2024), Tallinn, Estonia, October 2–4 2024. DOI: 10.1109/BEC61458.2024.10737958.

Publication III:

Afrin, Fariha & Le Moullec, Yannick & Pardy, Tamas & Rang, Toomas, Lightweight CNN-based Microfluidic Droplet Classification for Portable Imaging Flow Cytometry. Accepted for publication in a future issue of the Proceedings of the Estonian Academy of Science. 2025.

Publication IV:

Jõemaa, Rauno & Afrin, Fariha & Gyimah, Nafisat & Ashraf, Kanwal & Pärnamets, Kaiser & Pardy, Tamas, Cogni-Flow: Integrated Modular System For Automated Droplet Microfluidic Bio-analysis, EUROSENSORS XXXVI, 01-04 September 2024, Debrecen, Hungary.

Note: Publication III extends Publication II with: a) 13 new references (of which 9 in the state of the art, Section 1 in Publication III); b) a more detailed description of composition, preprocessing, augmentation, and characteristics of the dataset (Sections 2.1 and 2.2 in Publication III); c) expanded explanation of the model customization and optimization, including new Figure 2 and text (Section 2.3 in Publication III), and d) additional model (YoloV5-s) and target platforms (RPI-5 and MaixCam) and a wider and deeper comparison of various ML model deployments, including physical size, power requirements, and hardware capabilities (Sections 3.4 and 3.5 in Publication III). The extended experimental results in Publication III show that YoloV5-s on MaixCam achieves significantly faster clas-

sification time due to its onboard hardware accelerator at the cost of lower accuracy due to quantization. This wider analysis highlights the importance of addressing the trade-off between computational efficiency and classification accuracy.

5.1 Overview

Accurate microfluidic droplet classification is essential research for analyzing or sorting the target particles in IFC, providing valuable insights for biomedical applications. Recent advancements in DL have significantly transformed object detection and classification across various domains, including IFC, thanks to its automatic feature extraction and better performance than traditional approaches.

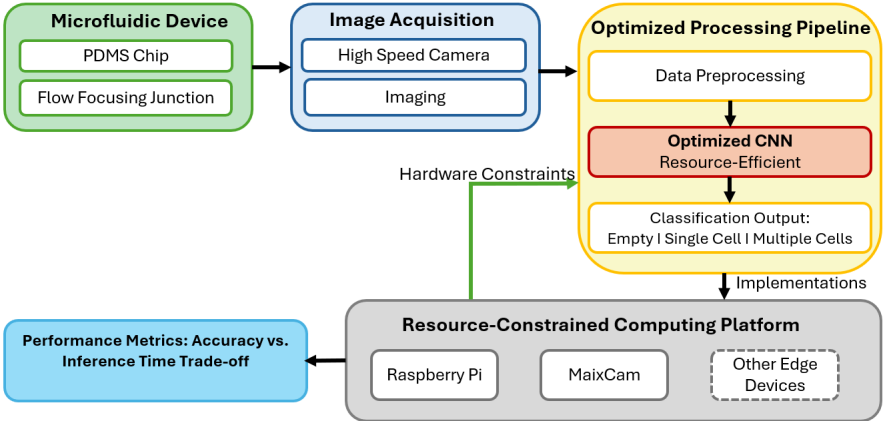


Figure 17: Visual representation of microfluidic droplet classification on resource constraint platforms with optimized CNN.

Existing methods are mostly based on DL, which includes single stage [72] and two-state detector [73], and are developed for lab-grade application using powerful GPUs. Faster and accurate droplet detection and classification implementation in resource-constrained platforms may significantly advance the field towards portability. However, the computational complexity and memory consumption poses challenges for their deployment on resource-constrained platforms. Despite the potential of CNNs to enhance droplet classification in IFC, existing methods are computationally expensive and rely on powerful GPUs, limiting their applicability in portable systems.

This work focuses on improving a single layer CNN model to classify three different types of droplets (empty droplet, droplet with single cell, droplet with multiple cells), making it suitable to meet the IFC requirement (better accuracy and smaller inference time) on a resource constrained platform. Prior research has explored CNN-based algorithms for droplet and cell classification, achieving significant accuracy improvements but often relying on desktop or GPU-based systems. For instance, the Weakly Supervised Cell Counting Network (WSCNet) achieved over 89% accuracy but required a desktop computer [74]. Similarly, other works have demonstrated high accuracy in cell classification using transfer learning [75] and optimized models on high-end GPUs [38]. However, these approaches are not suited for portable applications due to their computational requirements.

Figure 17 illustrates the complete representation from the microfluidic device through image acquisition to the optimized processing pipeline for droplet classification.

This chapter presents a customized Yolov4-tiny model, optimized for deployment on

resource-constrained devices, thereby addressing the limitations in conventional high-resource detection methodologies. The work includes three principal contributions that collectively advance the state-of-the-art in portable microfluidic analysis systems. The contributions are as follows:

- **Dataset Generation:** Creation of a microfluidic droplet dataset comprising 975 microfluidic droplet images extracted from experimental video sequences. This dataset incorporates augmentation techniques to enhance the data diversity and model robustness. The dataset generation establishes an important foundation for model optimization, providing sufficient variance while maintaining domain specificity.
- **Model Optimization:** Reduction of convolutional filters by 20% and implementation of batch processing (6 images per batch) to improve inference speed while maintaining high accuracy. This balanced reduction preserves essential feature extraction capabilities while significantly reducing the inference time.
- **Portable Deployment:** Successful deployment of the optimized YoloV4-tiny model on Raspberry Pi-4 and Raspberry Pi-5, demonstrating significant improvements in inference speed and resource efficiency. Moreover, for comparison purposes, results of a re-trained (with the same dataset) YOLOv5-s model on RPI-5 and MaixCam are presented.

5.2 Method

The workflow is composed of three steps as shown in Figure 18.

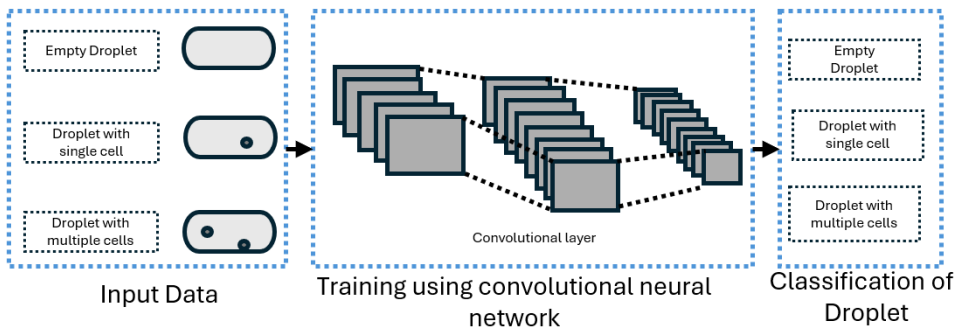


Figure 18: The proposed conceptual workflow for droplet classification.

The workflow begins with input data consisting of microscopic images of droplets, CNN architecture to classify microfluidic droplets based on their cellular content, categorizes them into three distinct classes:

- Empty droplets (containing no cells)
- Droplets with a single cell
- Droplets with multiple cells

This automated approach eliminates the need for manual inspection and classification based on their cellular content utilizing modified YOLOv4-tiny architecture optimized for resource-constrained environments while maintaining high classification accuracy.

5.2.1 Dataset Preparation

The dataset was derived from two experimental videos recorded at 15 frames per second, capturing droplet formation at a T-junction in a microfluidic chip. To enhance model robustness, data augmentation techniques were applied, including:

- Contrast variation
- Apply blur
- Rotational transformations

These augmentations expanded the initial dataset to 975 images. Annotation was performed using the "makesense.ai" platform, where bounding boxes were precisely marked around droplets.

5.2.2 Modified Architecture

Based on YOLOv4-tiny, a new lightweight object classifier with a 20% reduction in convolutional filters was developed (see Figure 19). This architectural modification resulted in a remarkable 36.16% decrease in model size while preserving detection accuracy.

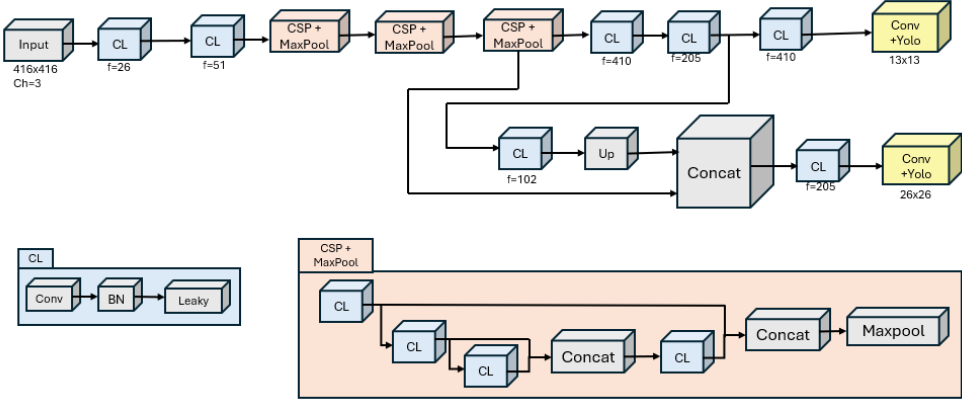


Figure 19: Proposed customized YOLOv4-tiny architecture: the backbone is made of CL blocks and CSP + MaxPool blocks, the detection head is mainly composed of CL blocks and two YOLO heads. Figure reproduced from Publication III (Figure 2).

The principal module in the architecture is the Cross-stage Partial (CSP), which consists of a Convolutional layer followed by Batch normalization and the Leaky-RELU activation function and skip connections for gradient flow optimization. Three CSP modules progressively extract hierarchical features from input images. For the detection head, YOLOv4-tiny uses two heads at dual scales:

- 13×13 feature maps targeting larger objects
- 26×26 feature maps optimized for smaller objects

A key observation in the proposed approach is that in the original YOLOv4-tiny, which employs 512 and 256 convolutional filters in various layers, not all parameters contribute significantly to prediction accuracy, allowing for empirical filter reduction without compromising performance. To reduce inference time, the number of filters in the convolutional layers was reduced by 20%.

The model was trained using the Darknet framework on Google Colaboratory Pro with a T4 GPU and training hyperparameters included a learning rate of 0.00261, decay of 0.0005, momentum of 0.9, and batch size of 64.

This empirical observation indicates filter reduction does not impact on steady convergence and sacrifice model performance for microfluidic droplet classification.

5.3 Overview of results of the classification of microfluidic droplet on resource-constrained platforms

The optimized model achieved a classification accuracy of over 98% and an inference time of 13 ms on the Raspberry Pi-5, resulting 76 FPS. The performance metrics included a precision of 0.95, recall of 1, and a mean average precision (mAP) of 99.95% at an IoU threshold of 0.5. These results indicate the reliability and effectiveness in identifying different microfluidic droplets.

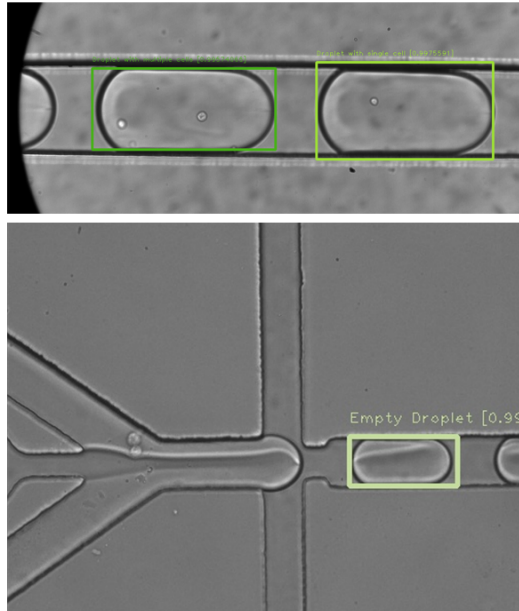


Figure 20: Visual examples of microfluidic droplet classification results after 6000 training iterations, where the prediction probability is above 99%. Figure reproduced from Publication III (Figure 4).

Figure 20 represents validation examples where the top image shows the classification of a droplet containing a single cell, highlighted with a blue bounding box, with a prediction confidence of approximately 99%. The bottom image shows the classification of an empty droplet, marked with a light green bounding box with a prediction confidence 99%. This indicates the model is capable of distinguishing different types of droplets accurately. Moreover, the resource utilization evaluation on Raspberry Pi 5 demonstrates that the modified model operates with an average CPU usage of 70% and an average memory usage of 60%. These results confirm that the model is computationally efficient and suitable for deployment on resource-constrained platforms like embedded devices. This work also discusses the YoloV4-tiny model classification performance metrics against seven other models/platforms. Regarding the first five works in Table 6, one-to-one comparisons are

not possible due to different applications and different computational power. Regarding resource-constrained platforms, to balance inference processing time and accuracy, model optimizations, hardware-aware deployment strategies, and empirical comparisons are used. In addition to the custom YoloV4-tiny deployed on RPI-5, YoloV5-s was retrained and deployed on the RPI-5 and MaixCam boards.

The MaixCam board is a low-cost, compact edge device with an embedded neural processing unit. It is based on a LicheeRV Nano module itself based on a Sophgo Technologies SG2002 system on chip (SoC). The board has a primary 700 MHz RISC-V C906 core, secondary boot-selectable 1 GHz RISC-V C906 core or 1 GHz ARM A53 core, and 256 MB on-chip RAM. The SG2002 features a TPU with performance up to 1 Tera Operations Per Second (TOPS) @INT8.

We used the same dataset described earlier; empirical experiments with hyperparameters, e.g. batch size and learning found best validation accuracy for batch size = 4 and learning rate = 0.0001, yielding validation accuracy = 0.966 at epoch #120.

Table 6 contrasts several closest-related droplet and cell classification methods across different platforms. Further details about the MaixCam and other platforms can be found in Publication III. Existing works do not target the same applications as that of this PhD work, and most existing studies implementation are on powerful desktop PCs and/or GPUs. In [76], 99.3% mAP at 0.5 and 74.3 ms processing time is obtained using YoloV5 on Intel Core i7-12650H, whereas the method proposed in this chapter yields 99.95% mAP and 13 ms processing time implementation on RPi-5. Other works include Mask-RCNN on Nvidia RTX 2080 Ti GPU that achieves 73% mAP for cell detection [77]. Also, CNN implementations on high-end GPUs (Tesla K80, K40c) achieves 95.7 and 99% mAP respectively with lower processing time, but they are developed for different, non-portable applications [78], [79].

When comparing resource-constrained deployments, the custom YoloV4-tiny on RPi5 performs well with a 13 ms inference time and 99.95% mAP@0.5. The MaixCam running YoloV5-s achieves a faster inference time of 5.34 ms but with a lower mAP of 55.09% due to INT8 TPU quantization. The YoloV5-s model on RPi5, while achieving 92.10% mAP@0.5, has a high inference time of 208.5 ms, making it less favorable. Performance scaling between MaixCam and RPi5 shows MaixCam is about 39.1 times faster but suffers a 40.2% accuracy drop. A more detailed discussion of the results can be found in Publication III.

The table shows that the modified YoloV4-tiny model successfully meets the research objectives of maintaining high classification accuracy while enabling deployment on resource-constrained devices for microfluidic droplet classification. Its deployment on RPi5 offers a good balance of accuracy and processing time for portable IFC devices, despite higher power consumption compared to MaixCam.

5.4 Conclusion on the classification of microfluidic droplet on resource constrained platform

The successful development of a compressed YoloV4-tiny model for microfluidic droplet classification represents a contribution in portable microfluidic droplet classification applications. By reducing convolutional filter counts by 20% while maintaining classification accuracy above 99%, the proposed approach effectively addresses the challenges between computational efficiency and classification performance on resource-constrained devices.

Table 6: Combinations of ML Models and Their Deployments. Table reorganized from Publication III (Table 3).

Reference	Model	Platform	Object	mAP (%)	Average inference processing time per image
[76]	YoloV5	Intel Core i7-12650H	Droplet	99.3@0.5	74.3 ms
[77]	Mask-RCNN	Nvidia RTX 2080 Ti GPU	Cells	73	Not reported in the paper
[78]	CNN	NVIDIA Tesla K80 GPU	Cells	95.7	2.2 ms
[80]	CNN	High-end PC (assumed)	Cells	93-99	Not reported in the paper
[79]	CNN	Nvidia Tesla K40c GPU	Cells	99	Not reported in the paper
This work	Customized YoloV4-tiny	RPI5	Droplet	99.95@0.5	13 ms
This work	YoloV5-s	RPI5	Droplet	92.10@0.5	208.5 ms
This work	YoloV5-s	MaixCam	Droplet	55.09@0.5	5.34 ms

The experimental results validate that the proposed approach can reliably distinguish between empty droplets and those containing single or multiple cells, achieving 99.95% mAP with a processing time of 13ms on a RPI-5 platform; this demonstrates that the proposed deep learning approach can be successfully deployed on Single Board Computers (SBCs) without sacrificing classification reliability.

The methodology developed here not only contributes to the specific field of microfluidic droplet classification but also represents the adaptability of optimized deep learning architectures to resource-limited environments across various applications in PoC devices.

6 CONCLUSION AND FUTURE WORK

6.1 Summary

The outset for this PhD was the increasing interest for the high-content screening capabilities of IFC, a high-throughput imaging technique used in immunology, cancer research, and cell biology. It was noted that recent research focuses on precise detection and classification using machine learning, particularly deep learning; at the same time, there are increasing opportunities for developing IFC portable devices for PoC applications, leading to the need for implementing IFC on embedded platforms. However, it was also noted that implementing droplet classification models in resource-constrained environments faces several key challenges:

1. Limited computational resources, including limited processing power and memory, which constrain the complexity of models that can be implemented.
2. The need for energy-efficient solutions, which often conflicts with the implementation of advanced, computationally intensive techniques.
3. Challenges in obtaining clear, blur-free images in the dynamic environment of microfluidic systems.
4. The requirement for streamlined, resource-efficient mechanisms to assess image quality or to analyze the image such as object classification.

These constraints significantly affect the performance and precision of the models when deployed on resource-constrained devices. The majority of current research targets high-computational powered hardware that is not suitable for portable devices. Consequently, achieving rapid inference while preserving accuracy presents a considerable challenge. To overcome these hurdles, it was necessary to develop new solutions in areas such as model selection, optimization techniques, and the efficient exploitation of available hardware resources.

To address these challenges, this PhD thesis has made three contributions to the field of IFC, i.e. Contribution 1) blur-free image acquisition, Contribution 2) image quality assessment, and Contribution 3) droplet classification, each corresponding to specific research question(s).

Answers to the research questions, posed in the introduction of the thesis, are answered below.

- RQ1 was "How to design a lightweight and effective image processing pipeline for detecting, counting and acquiring blur-free segmented droplet images suitable for application-specific resource-constrained platforms?"

This question was addressed through Contribution 1. The proposed design comprises key novelties: i.e. a novel lightweight algorithm for detecting, counting, and acquiring single object images from video streams and techniques to address issues of blur and duplicate droplet image acquisition. The experimental evaluation using videos of fast-moving droplets with fluid flow rates up to $67.7 \mu\text{L}/\text{min}$ demonstrates a 100% counting accuracy on tested videos, as well as processing times per droplet of 2 ms (500 DPS) on Desktop PC, 25 ms (40 DPS) on Raspberry Pi-4, and 10 ms (100 DPS) on Nvidia Jetson Nano. The Jetson Nano implementation (100 DPS) improves performance over existing similar studies and proved suitable for the target application.

Therefore, this lightweight pipeline enables efficient single object image acquisition in IFC on embedded portable platforms, addressing limitations of current complex, non-portable hardware setups.

- RQ2 was "Can deep neural network replace conventional signal processing for achieving effective image quality assessment on a resource-constrained platform?"

This question was addressed through Contribution 2. A NR-IQA approach techniques within the context of a portable lab-on-chip microfluidic droplet IFC device was proposed. The key novelty is the developed new hybrid model comprising a modified MobileNet followed by a transformer block, specifically for microfluidic droplet IQA. Competitive performance metrics are illustrated in terms of SRCC: 0.73 and PLCC: 0.75, as well as inference time on a RPI-5: 110 ms, translating to 9 FPS.

These results demonstrate that deep neural networks can effectively replace conventional signal processing for image quality assessment in resource-constrained environments. The trade-off between accuracy and computational resources is favorable, with the hybrid model achieving good correlation coefficients while maintaining a reasonable inference time on a low-power device.

- RQ3 was "How to design a custom droplet classifier (presence or absence of cells inside the droplet) suitable for rapid inference on a resource-constrained platform while maintaining high classification accuracy?"

This question was addressed through Contribution 3, with key novelties being the novel approach to classifying microfluidic droplets on resource-constrained platforms, i.e. 1) the development of a customized YoloV4-tiny model for rapid inference while maintaining high classification accuracy, 2) the creation of a custom dataset of 975 images from real-life microfluidic experimental setups, and 3) the classification of droplets into three categories (empty, single cell, multiple cells). The results demonstrate an inference time of 13 ms per classification on RPI5 with mAP@0.5 of 99.95%. Deployment of a YoloV5s model on the low-cost MaixCam board yields an inference time of 5.34 ms (faster than on RPI-4 due to onboard tensor processing unit) with mAP@0.5 of 55.09% (lower than on RPI-5 due to quantization).

These results show that careful system design can enable robust droplet classification performance even on devices with limited resources, successfully addressing the challenge of droplet classification in portable IFC.

To sum-up, this PhD thesis has made contributions to the field of IFC, particularly in the context of resource-constrained environments. The results illustrates that the contributions can address the challenges of blur-free image acquisition, image quality assessment, and droplet classification on portable platforms. The lightweight imaging pipeline (Contribution 1) provides a foundation for efficient image acquisition in portable IFC devices. The novel hybrid model for image quality assessment (Contribution 2) demonstrates the potential of deep learning techniques in resource-limited settings. Finally, the customized YoloV4-tiny model for droplet classification (Contribution 3) showcases the ability to achieve high accuracy and rapid inference on resource constrained devices.

The achieved results open the way for further research, as exemplified in the perspectives presented in the next section.

6.2 Perspectives

Several directions for future research could be pursued beyond the scope of the current work. Regarding lightweight imaging for IFC, this thesis has demonstrated competitive results on desktop and resource constrained platforms, but some applications may require higher scaling of the throughput while maintaining image quality. Addressing this challenge would require research beyond the scope of this thesis, i.e. novel methodological and algorithmic approaches. One promising option exemplified by [81], would be developing deep learning approaches that can analyze diverse datasets acquired from IFC, i.e. moving beyond simple image features to more sophisticated pattern recognition. Moreover, the development of reactive flow cytometry systems as explored by [82] could enable real-time feedback control over cell populations.

When it comes to IQA, our model established a baseline for NR-IQA with good correlation coefficients but also identified the need to develop domain-specific quality metrics. Future research would require a significant interdisciplinary research effort between computer vision specialists and biologists to investigate quality metrics specifically tailored to biological relevance. A promising direction would be the integration of profiling techniques that analyze multiple cellular parameters simultaneously, considering not just traditional image quality metrics but also cellular morphology and protein localization [83].

Finally, this thesis demonstrated real-world feasibility on resource constrained platforms, but the rapidly evolving landscape of edge AI hardware presents both new opportunities and challenges. Hardware accelerators such as the Hailo-8 HAT for RPI5 with up to 26 TOPS capability [84] or the Sony IMX500 with its stacked sensor structure for edge AI processing [85] are promising platforms for reducing processing time while potentially maintaining accuracy. Leveraging these technologies and its inherent hardware-software co-optimizations require complex design space exploration and optimization strategies that represent a distinct research project rather than a mere extension of the current work.

List of Figures

1	Evolution of Imaging Flow Cytometry, from the first automatic particle counter to the integration of AI/ML and the advent of portable flow cytometers	12
2	Simplified diagram of a MF chip based IFC device	13
3	An example of MF droplet generation (Left) and cell encapsulation (Right). Reproduced from [13]	14
4	Images of MF droplets flowing inside a microfluidic channel (images extracted from the dataset used in later chapters of this thesis).	14
5	An overview of research task	16
6	Framework for the Ph.D thesis.....	18
7	Illustration of MF Droplet Analysis with Embedded Platform Challenges. ...	19
8	Hierarchical classification of IQA methodologies.	21
9	Summarized Workflow and Challenges in Blur-Free Single Droplet Image Acquisition for Imaging Flow Cytometry.	26
10	Droplet generation. (a) Schematic of the PDMS chip (not drawn to scale). (b) Experimental setup showing all (water, oil, air) inlets, the droplet outlet, the filter, and the excitation laser light. The power supply for driving the laser, the computer for controlling the syringe pump, and the Basler Ace camera are not visible here. Figure reproduced from Publication I (Figure 1).	29
11	Block diagram of the single-image acquisition process flow, consisting of (1) video streaming, which occurs at 1000 fps; (2) processing, which involves detecting droplets based on their color and tracks; and (3) result acquisition, which involves acquiring a single image and counting the number of droplets using the tracking result without any duplication. Figure reproduced from Publication I (Figure 2).	30
12	Three different conditions for accurate counting (Top: COMSOL simulated two-phase flow video, bottom: experimental fluorescence video): (i) if the green bounding box is outside ROI-2, the operation returns -1; (ii) if the bounding box is exactly on the boundary of ROI-2, it returns 0; and (iii) if the bounding box is inside ROI-2, it returns +1, which is the right condition to store the location. Figure reproduced from Publication I (Figure 5).....	31
13	Single droplet image acquisition in brightfield microscopy (Left to right: video streaming, processing, and result). Masking (white boundary), contouring (yellow boundary) and tracking (green rectangle) shows only for one droplet from the video. Multiple saved droplet images are shown in result. Figure reproduced from Publication I (Figure 8).....	35
14	Automated Image Quality Assessment for Imaging Flow Cytometry.	37
15	Proposed framework for predicting the quality of MF droplet images. Expanded version of Figure 1 in Publication V.....	39
16	(a)Test image (example of a good image quality), and (b): Class prediction distribution of the image ($MOS \pm STD = 8.881 \pm 1.095$). Reproduced from Publication V (Figure 6).....	41
17	Visual representation of microfluidic droplet classification on resource constraint platforms with optimized CNN.	44
18	The proposed conceptual workflow for droplet classification.	45

19	Proposed customized YoloV4-tiny architecture: the backbone is made of CL blocks and CSP + MaxPool blocks, the detection head is mainly composed of CL blocks and two Yolo heads. Figure reproduced from Publication III (Figure 2).	46
20	Visual examples of microfluidic droplet classification results after 6000 training iterations, where the prediction probability is above 99%. Figure reproduced from Publication III (Figure 4).....	47

List of Tables

2 Summary of Contributions in Relation to Research Papers 17

3 Comparison of different techniques and the proposed method. Each key
element (complexity, portability, etc.) used for the comparison is scored
on a scale ranging from "+" to "+++". (Table reproduced from publication I) 27

4 Comparison of Most-closely related results vs. our work. Table repro-
duced from Publication I (Table 4). 34

5 Comparison of Model Complexity (number of parameters) and Perfor-
mance with State-of-the-Art IQA Models Based on SRCC, PLCC, and EMD.
Reproduced from Publication V (Table 1). 41

6 Combinations of ML Models and Their Deployments. Table reorganized
from Publication III (Table 3). 49

References

- [1] W. H. Coulter, "High speed automatic blood cell counter and cell size analyzer: proceedings of the national electronics conference," 1957.
- [2] L. A. Herzenberg, D. Parks, B. Sahaf, O. Perez, M. Roederer, and L. A. Herzenberg, "The history and future of the fluorescence activated cell sorter and flow cytometry: a view from stanford," *Clinical chemistry*, vol. 48, no. 10, pp. 1819–1827, 2002.
- [3] V. Kachel, G. Benker, K. Lichtnau, G. Valet, and E. Glossner, "Fast imaging in flow: a means of combining flow-cytometry and image analysis.," *Journal of Histochemistry & Cytochemistry*, vol. 27, no. 1, pp. 335–341, 1979.
- [4] T. A. More, B. Dalal, R. Devendra, P. Warang, A. Shankarkumar, and P. Kedar, "Applications of imaging flow cytometry in the diagnostic assessment of red cell membrane disorders," *Cytometry Part B: Clinical Cytometry*, vol. 98, no. 3, pp. 238–249, 2020.
- [5] Z. Darzynkiewicz, E. Bedner, X. Li, W. Gorczyca, and M. R. Melamed, "Laser-scanning cytometry: a new instrumentation with many applications," *Experimental cell research*, vol. 249, no. 1, pp. 1–12, 1999.
- [6] P. Rees, H. D. Summers, A. Filby, A. E. Carpenter, and M. Doan, "Imaging flow cytometry," *Nature Reviews Methods Primers*, vol. 2, no. 1, p. 86, 2022.
- [7] S. Ota, I. Sato, and R. Horisaki, "Implementing machine learning methods for imaging flow cytometry," *Microscopy*, vol. 69, no. 2, pp. 61–68, 2020.
- [8] P. Pozzi, A. Candeo, P. Paiè, F. Bragheri, and A. Bassi, "Artificial intelligence in imaging flow cytometry," *Frontiers in Bioinformatics*, vol. 3, p. 1229052, 2023.
- [9] R. Subramanian, R. Tang, Z. Zhang, V. Joshi, J. N. Miner, and Y.-H. Lo, "Multimodal nash prognosis using 3d imaging flow cytometry and artificial intelligence to characterize liver cells," *Scientific reports*, vol. 12, no. 1, p. 11180, 2022.
- [10] Z. Göröcs, D. Baum, F. Song, K. de Haan, H. C. Koydemir, Y. Qiu, Z. Cai, T. Skandakumar, S. Peterman, M. Tamamitsu, *et al.*, "Label-free detection of giardia lamblia cysts using a deep learning-enabled portable imaging flow cytometer," *Lab on a Chip*, vol. 20, no. 23, pp. 4404–4412, 2020.
- [11] S. Zhang, Z. Li, and Q. Wei, "Smartphone-based cytometric biosensors for point-of-care cellular diagnostics," *Nanotechnology and Precision Engineering*, vol. 3, no. 1, pp. 32–42, 2020.
- [12] S. Sohrabi, M. K. Moraveji, *et al.*, "Retracted article: Droplet microfluidics: fundamentals and its advanced applications," *RSC advances*, vol. 10, no. 46, pp. 27560–27574, 2020.
- [13] Z. Chen, S. Kheiri, E. W. Young, and E. Kumacheva, "Trends in droplet microfluidics: From droplet generation to biomedical applications," *Langmuir*, vol. 38, no. 20, pp. 6233–6248, 2022.
- [14] G. S. Elliott, "Moving pictures: imaging flow cytometry for drug development," *Combinatorial chemistry & high throughput screening*, vol. 12, no. 9, pp. 849–859, 2009.

- [15] Y. Wang, T. Yoshihara, S. King, T. Le, P. Leroy, X. Zhao, C. K. Chan, Z.-H. Yan, and S. Menon, "Automated high-throughput flow cytometry for high-content screening in antibody development," *SLAS DISCOVERY: Advancing Life Sciences R&D*, vol. 23, no. 7, pp. 656–666, 2018.
- [16] T. E. Hybel, S. H. Jensen, M. A. Rodrigues, T. E. Hybel, M. N. Pedersen, S. H. Qvick, M. H. Enemark, M. Bill, C. A. Rosenberg, and M. Ludvigsen, "Imaging flow cytometry and convolutional neural network-based classification enable discrimination of hematopoietic and leukemic stem cells in acute myeloid leukemia," *International Journal of Molecular Sciences*, vol. 25, no. 12, p. 6465, 2024.
- [17] A. Cohen, M. Dudaie, I. Barnea, F. Borrelli, J. Běhal, L. Miccio, P. Memmolo, V. Bianco, P. Ferraro, and N. T. Shaked, "Label-free imaging flow cytometry for cell classification based on multiple interferometric projections using deep learning," *Advanced Intelligent Systems*, vol. 6, no. 1, p. 2300433, 2024.
- [18] P.-C. Chen and J. Lawrensen, "Improving a smartphone based droplet flow cytometry system with micro lens arrays integrated optofluidic chip," *Sensors and Actuators A: Physical*, vol. 367, p. 115080, 2024.
- [19] R. Wang, X. Huang, X. Xu, J. Sun, S. Zheng, X. Ke, J. Yao, W. Han, M. Wei, J. Chen, *et al.*, "A standalone and portable microfluidic imaging detection system with embedded computing for point-of-care diagnostics," *IEEE Sensors Journal*, vol. 22, no. 6, pp. 6116–6123, 2022.
- [20] J.-u. Shim, R. T. Ranasinghe, C. A. Smith, S. M. Ibrahim, F. Hollfelder, W. T. Huck, D. Klenerman, and C. Abell, "Ultrarapid generation of femtoliter microfluidic droplets for single-molecule-counting immunoassays," *ACS nano*, vol. 7, no. 7, pp. 5955–5964, 2013.
- [21] Z. Su, J. He, P. Zhou, L. Huang, and J. Zhou, "A high-throughput system combining microfluidic hydrogel droplets with deep learning for screening the antisolvent-crystallization conditions of active pharmaceutical ingredients," *Lab on a Chip*, vol. 20, no. 11, pp. 1907–1916, 2020.
- [22] E. D. Diebold, B. W. Buckley, D. R. Gossett, and B. Jalali, "Digitally synthesized beat frequency multiplexing for sub-millisecond fluorescence microscopy," *Nature Photonics*, vol. 7, no. 10, pp. 806–810, 2013.
- [23] G. Holzner, B. Mateescu, D. van Leeuwen, G. Cereghetti, R. Dechant, S. Stavrakis, and A. DeMello, "High-throughput multiparametric imaging flow cytometry: toward diffraction-limited sub-cellular detection and monitoring of sub-cellular processes," *Cell Reports*, vol. 34, no. 10, 2021.
- [24] N. Nitta, T. Sugimura, A. Isozaki, H. Mikami, K. Hiraki, S. Sakuma, T. Iino, F. Arai, T. Endo, Y. Fujiwaki, *et al.*, "Intelligent image-activated cell sorting," *Cell*, vol. 175, no. 1, pp. 266–276, 2018.
- [25] S. Lv, Y. Chu, P. Zhang, S. Ma, M. Zhao, Z. Wang, Y. Gu, and X. Sun, "Improved efficiency of urine cell image segmentation using droplet microfluidics technology," *Cytometry Part A*, vol. 99, no. 7, pp. 722–731, 2021.

- [26] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *Medical image computing and computer-assisted intervention—MICCAI 2015: 18th international conference, Munich, Germany, October 5-9, 2015, proceedings, part III* 18, pp. 234–241, Springer, 2015.
- [27] D. A. Van Valen, T. Kudo, K. M. Lane, D. N. Macklin, N. T. Quach, M. M. DeFelice, I. Maayan, Y. Tanouchi, E. A. Ashley, and M. W. Covert, "Deep learning automates the quantitative analysis of individual cells in live-cell imaging experiments," *PLoS computational biology*, vol. 12, no. 11, p. e1005177, 2016.
- [28] U. Sara, M. Akter, and M. S. Uddin, "Image quality assessment through fsm, ssim, mse and psnr—a comparative study," *Journal of Computer and Communications*, vol. 7, no. 3, pp. 8–18, 2019.
- [29] L. Wang, "A survey on iqa," *arXiv preprint arXiv:2109.00347*, 2021.
- [30] A. Mittal, A. K. Moorthy, and A. C. Bovik, "No-reference image quality assessment in the spatial domain," *IEEE Transactions on image processing*, vol. 21, no. 12, pp. 4695–4708, 2012.
- [31] A. Mittal, R. Soundararajan, and A. C. Bovik, "Making a "completely blind" image quality analyzer," *IEEE Signal processing letters*, vol. 20, no. 3, pp. 209–212, 2012.
- [32] Z. Zhang, G. Dai, X. Liang, S. Yu, L. Li, and Y. Xie, "Can signal-to-noise ratio perform as a baseline indicator for medical image quality assessment," *IEEE Access*, vol. 6, pp. 11534–11543, 2018.
- [33] N. El Abbadi, E. A. Al-Zubaidi, and H. S. Razzaq, "Image quality assessment tools," *Journal of Xi'an University of Architecture and Technology*, vol. 12, no. 3, pp. 1260–1276, 2020.
- [34] H. R. Sheikh, M. F. Sabir, and A. C. Bovik, "A statistical evaluation of recent full reference image quality assessment algorithms," *IEEE Transactions on image processing*, vol. 15, no. 11, pp. 3440–3451, 2006.
- [35] N. Ponomarenko, O. Ieremeiev, V. Lukin, K. Egiazarian, L. Jin, J. Astola, B. Vozel, K. Chehdi, M. Carli, F. Battisti, *et al.*, "Color image database tid2013: Peculiarities and preliminary results," in *European workshop on visual information processing (EUVIP)*, pp. 106–111, IEEE, 2013.
- [36] H. Lin, V. Hosu, and D. Saupe, "Koniq-10k: Towards an ecologically valid and large-scale iqa database," *arXiv preprint arXiv:1803.08489*, 2018.
- [37] D. Pirone, A. Montella, D. G. Sirico, M. Mugnano, M. M. Villone, V. Bianco, L. Miccio, A. M. Porcelli, I. Kurelac, M. Capasso, *et al.*, "Label-free liquid biopsy through the identification of tumor cells by machine learning-powered tomographic phase imaging flow cytometry," *Scientific Reports*, vol. 13, no. 1, p. 6042, 2023.
- [38] K. Lee, S.-E. Kim, J. Doh, K. Kim, and W. K. Chung, "User-friendly image-activated microfluidic cell sorting technique using an optimized, fast deep learning algorithm," *Lab on a Chip*, vol. 21, no. 9, pp. 1798–1810, 2021.

- [39] R. Tang, L. Xia, B. Gutierrez, I. Gagne, A. Munoz, K. Eribez, N. Jagnandan, X. Chen, Z. Zhang, L. Waller, *et al.*, "Low-latency label-free image-activated cell sorting using fast deep learning and ai inferencing," *Biosensors and Bioelectronics*, vol. 220, p. 114865, 2023.
- [40] J. Hung and A. Carpenter, "Applying faster r-cnn for object detection on malaria images," in *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*, pp. 56–61, 2017.
- [41] E. Shechtman, Y. Caspi, and M. Irani, "Increasing space-time resolution in video," in *Computer Vision—ECCV 2002: 7th European Conference on Computer Vision Copenhagen, Denmark, May 28–31, 2002 Proceedings, Part I* 7, pp. 753–768, Springer, 2002.
- [42] D. Bradley, B. Atcheson, I. Ihrke, and W. Heidrich, "Synchronization and rolling shutter compensation for consumer video camera arrays," in *2009 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, pp. 1–8, IEEE, 2009.
- [43] E. K. Zuba-Surma, M. Kucia, A. Abdel-Latif, J. W. Lillard, and M. Z. Ratajczak, "The imagestream system: a key step to a new era in imaging.," *Folia histochemica et cytobiologica*, vol. 45, no. 4, pp. 279–290, 2007.
- [44] A. S. Rane, J. Rutkauskaitė, A. deMello, and S. Stavrakis, "High-throughput multi-parametric imaging flow cytometry," *Chem*, vol. 3, no. 4, pp. 588–602, 2017.
- [45] E. Schonbrun, S. S. Gorthi, and D. Schaak, "Microfabricated multiple field of view imaging flow cytometry," *Lab on a Chip*, vol. 12, no. 2, pp. 268–273, 2012.
- [46] A. K. Lau, A. H. Tang, J. Xu, X. Wei, K. K. Wong, and K. K. Tsia, "Optical time stretch for high-speed and high-throughput imaging—from single-cell to tissue-wide scales," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 22, no. 4, pp. 89–103, 2015.
- [47] B. Guo, C. Lei, H. Kobayashi, T. Ito, Y. Yalikun, Y. Jiang, Y. Tanaka, Y. Ozeki, and K. Goda, "High-throughput, label-free, single-cell, microalgal lipid screening by machine-learning-equipped optofluidic time-stretch quantitative phase microscopy," *Cytometry Part A*, vol. 91, no. 5, pp. 494–502, 2017.
- [48] K. Goda, A. Ayazi, D. R. Gossett, J. Sadasivam, C. K. Lonappan, E. Sollier, A. M. Fard, S. C. Hur, J. Adam, C. Murray, *et al.*, "High-throughput single-microparticle imaging flow analyzer," *Proceedings of the National Academy of Sciences*, vol. 109, no. 29, pp. 11630–11635, 2012.
- [49] H. Mikami, M. Kawaguchi, C.-J. Huang, H. Matsumura, T. Sugimura, K. Huang, C. Lei, S. Ueno, T. Miura, T. Ito, *et al.*, "Virtual-freezing fluorescence imaging flow cytometry," *Nature communications*, vol. 11, no. 1, p. 1162, 2020.
- [50] D. V. Voronin, A. A. Kozlova, R. A. Verkhovskii, A. V. Ermakov, M. A. Makarkin, O. A. Inozemtseva, and D. N. Bratashov, "Detection of rare objects by flow cytometry: imaging, cell sorting, and deep learning approaches," *International journal of molecular sciences*, vol. 21, no. 7, p. 2323, 2020.

- [51] W. E. Ortyn, B. E. Hall, T. C. George, K. Frost, D. A. Basiji, D. J. Perry, C. A. Zimmerman, D. Coder, and P. J. Morrissey, "Sensitivity measurement and compensation in spectral imaging," *Cytometry Part A: The Journal of the International Society for Analytical Cytology*, vol. 69, no. 8, pp. 852–862, 2006.
- [52] Y. Wang, Z. Huang, X. Wang, F. Yang, X. Yao, T. Pan, B. Li, and J. Chu, "Real-time fluorescence imaging flow cytometry enabled by motion deblurring and deep learning algorithms," *Lab on a Chip*, vol. 23, no. 16, pp. 3615–3627, 2023.
- [53] V. Ganjalizadeh, G. G. Meena, M. A. Stott, A. R. Hawkins, and H. Schmidt, "Machine learning at the edge for ai-enabled multiplexed pathogen detection," *Scientific Reports*, vol. 13, no. 1, p. 4744, 2023.
- [54] K. Pärnamets, A. Udal, A. Koel, T. Pardy, N. Gyimah, and T. Rang, "Compact empirical model for droplet generation in a lab-on-chip cytometry system," *IEEE Access*, vol. 10, pp. 127708–127717, 2022.
- [55] Y. Zhao, X. Zhou, X. Xu, Z. Jiang, F. Cheng, J. Tang, and Y. Shen, "A novel vehicle tracking id switches algorithm for driving recording sensors," *Sensors*, vol. 20, no. 13, p. 3638, 2020.
- [56] M. Durve, A. Tiribocchi, F. Bonaccorso, A. Montessori, M. Lauricella, M. Bogdan, J. Guzowski, and S. Succi, "Droptack—automatic droplet tracking with yolov5 and deepsort for microfluidic applications," *Physics of Fluids*, vol. 34, no. 8, 2022.
- [57] K. Fischer, A. Lulla, T. Y. So, P. Pereyra-Gerber, M. I. Raybould, T. N. Kohler, T. S. Kaminski, J. C. Yam-Puc, R. Hughes, F. Leiß-Maier, *et al.*, "Microfluidics-enabled fluorescence-activated cell sorting of single pathogen-specific antibody secreting cells for the rapid discovery of monoclonal antibodies," *bioRxiv*, pp. 2023–01, 2023.
- [58] S. Dimitriadis, L. Dova, I. Kotsianidis, E. Hatzimichael, E. Kapsali, and G. S. Markopoulos, "Imaging flow cytometry: Development, present applications, and future challenges," *Methods and Protocols*, vol. 7, no. 2, p. 28, 2024.
- [59] H. Wang, C. Jin, Y. Feng, D. Qi, Y. Sa, and X.-H. Hu, "Quantitative assessment of image motion blur in diffraction images of moving biological cells," *Optical Engineering*, vol. 55, no. 2, pp. 023103–023103, 2016.
- [60] L. Kang, P. Ye, Y. Li, and D. Doermann, "Convolutional neural networks for no-reference image quality assessment," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 1733–1740, 2014.
- [61] Y. Niu, Y. Zhong, W. Guo, Y. Shi, and P. Chen, "2d and 3d image quality assessment: A survey of metrics and challenges," *IEEE Access*, vol. 7, pp. 782–801, 2018.
- [62] Z. Wang, J. Zhou, Q. Liu, L. Wang, and X. Su, "Multiplex imaging flow cytometry with content-aware image restoration deep learning enhances cell imaging quality," in *Fifteenth International Conference on Information Optics and Photonics (CIOP 2024)*, vol. 13418, pp. 872–878, SPIE, 2024.
- [63] A. Liu, J. Wang, J. Liu, and Y. Su, "Comprehensive image quality assessment via predicting the distribution of opinion score," *Multimedia Tools and Applications*, vol. 78, pp. 24205–24222, 2019.

- [64] F. Zhang, C. Lei, C.-J. Huang, H. Kobayashi, C.-W. Sun, and K. Goda, "Intelligent image de-blurring for imaging flow cytometry," *Cytometry Part A*, vol. 95, no. 5, pp. 549–554, 2019.
- [65] A. Kurenkov, A. Kussanova, and N. S. Barteneva, "Advancing precision single-cell analysis of red blood cells through semi-supervised deep learning using database of patients with post-covid-19 syndrome," in *Imaging, manipulation, and analysis of biomolecules, cells, and tissues XXII*, vol. 12846, p. 1284602, SPIE, 2024.
- [66] R. Patel and A. Chaware, "Mobilenet architecture and its application to computer vision," IET, 2021.
- [67] S. Joshi, M. Manu, and A. Mittal, "A review of the evolution and applications of convolutional neural network (cnn)," in *2023 2nd International Conference on Edge Computing and Applications (ICECAA)*, pp. 1109–1114, IEEE, 2023.
- [68] X. Yang, F. Li, and H. Liu, "Ttl-iqa: Transitive transfer learning based no-reference image quality assessment," *IEEE Transactions on Multimedia*, vol. 23, pp. 4326–4340, 2020.
- [69] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," *Communications of the ACM*, vol. 60, no. 6, pp. 84–90, 2017.
- [70] S. A. Golestaneh, S. Dadsetan, and K. M. Kitani, "No-reference image quality assessment via transformers, relative ranking, and self-consistency," in *Proceedings of the IEEE/CVF winter conference on applications of computer vision*, pp. 1220–1230, 2022.
- [71] S. Su, Q. Yan, Y. Zhu, C. Zhang, X. Ge, J. Sun, and Y. Zhang, "Blindly assess image quality in the wild guided by a self-adaptive hyper network," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 3667–3676, 2020.
- [72] H. Wang, Y. Yu, Y. Cai, X. Chen, L. Chen, and Y. Li, "Soft-weighted-average ensemble vehicle detection method based on single-stage and two-stage deep learning models," *IEEE Transactions on Intelligent Vehicles*, vol. 6, no. 1, pp. 100–109, 2020.
- [73] D. Demetriou, P. Mavromatidis, P. M. Robert, H. Papadopoulos, M. F. Petrou, and D. Nicolaides, "Real-time construction demolition waste detection using state-of-the-art deep learning methods; single-stage vs two-stage detectors," *Waste Management*, vol. 167, pp. 194–203, 2023.
- [74] X. Zhou, Y. Mao, M. Gu, and Z. Cheng, "Wscnet: Biomedical image recognition for cell encapsulated microfluidic droplets," *Biosensors*, vol. 13, no. 8, p. 821, 2023.
- [75] A. Kensert, P. J. Harrison, and O. Spjuth, "Transfer learning with deep convolutional neural networks for classifying cellular morphological changes," *SLAS Discovery: Advancing Life Sciences R&D*, vol. 24, no. 4, pp. 466–475, 2019.
- [76] J. Xu, W. Fan, J. Madsen, G. P. Tanev, and L. Pezzarossa, "Ai-based detection of droplets and bubbles in digital microfluidic biochips," in *2023 Design, Automation & Test in Europe Conference & Exhibition (DATE)*, pp. 1–6, IEEE, 2023.

- [77] M. Ghafari, D. Mailman, P. Hatami, T. Peyton, L. Yang, W. Dang, and H. Qin, "A comparison of yolo and mask-rcnn for detecting cells from microfluidic images," in *2022 International Conference on Artificial Intelligence in Information and Communication (ICAIIIC)*, pp. 204–209, IEEE, 2022.
- [78] Y. Li, A. Mahjoubfar, C. L. Chen, K. R. Niazi, L. Pei, and B. Jalali, "Deep cytometry: deep learning with real-time inference in cell sorting and flow cytometry," *Scientific reports*, vol. 9, no. 1, p. 11088, 2019.
- [79] N. Meng, E. Y. Lam, K. K. Tsia, and H. K.-H. So, "Large-scale multi-class image-based cell classification with deep learning," *IEEE journal of biomedical and health informatics*, vol. 23, no. 5, pp. 2091–2098, 2018.
- [80] Y. Suzuki, K. Kobayashi, Y. Wakisaka, D. Deng, S. Tanaka, C.-J. Huang, C. Lei, C.-W. Sun, H. Liu, Y. Fujiwaki, et al., "Label-free chemical imaging flow cytometry by high-speed multicolor stimulated raman scattering," *Proceedings of the National Academy of Sciences*, vol. 116, no. 32, pp. 15842–15848, 2019.
- [81] P. Rees, H. Summers, A. Filby, A. Carpenter, and M. Doan, "Imaging flow cytometry: a primer," *Nature Review Methods Primers*, vol. 2, no. 86, 2022.
- [82] F. Delvigne and J. A. Martinez, "Advances in automated and reactive flow cytometry for synthetic biotechnology," *Current Opinion in Biotechnology*, vol. 83, p. 102974, 2023.
- [83] G. P. Way, H. Saimem, S. Shave, R. Kasprowicz, and N. O. Carragher, "Evolution and impact of high content imaging," *SLAS Discovery*, vol. 28, no. 7, pp. 292–305, 2023. High-Content Imaging and Informatics: 3rd Annual Joint Special Issue with the Society for Biomolecular Imaging and Informatics.
- [84] Raspberry Pi Ltd, "Raspberry pi ai hat+ product brief," October 2024. <https://datasheets.raspberrypi.com/ai-hat-plus/raspberry-pi-ai-hat-plus-product-brief.pdf>.
- [85] Raspberry Pi Ltd, "Raspberry pi ai camera product brief," September 2024. <https://datasheets.raspberrypi.com/camera/ai-camera-product-brief.pdf>.

Acknowledgements

I express my warmest gratitude to my PhD advisors, Dr. Yannick Le Moullec, Dr. Tamás Pardy, and Dr. Ants Koel for their support, mentorship, guidance and feedback throughout my PhD journey.

I also share my gratitude with Dr. Toomas Rang for inviting me to join the Lab-on-a-Chip and Microfluidics team.

I also thank Mr. Laur Lemendik and Dr. Alar Kuusik, former and current Heads of Thomas Johann Seebeck Department of Electronics at Tallinn University of Technology, for giving me the opportunity to conduct my PhD work in this department.

I am grateful to all the colleagues and other PhD students who contributed to creating a supportive and nurturing environment.

I also thank Assistant Prof. Tomasz Kaminski, University of Warsaw, Poland, for providing me with the validation data.

Last but certainly not least, I thank my family for their support and encouragements throughout my PhD journey.

I acknowledge and deeply appreciate financial support from:

- Estonian Science Agency ETAg project Cognitronic Lab-on-a-Chip System for Highly-Automated Flow Cytometry (CogniFlow-Cyte, grant PRG620)
- Estonian Education and Youth Board project Artificial Intelligence, Edge Computing and IoT Solutions in Distributed Systems (AloT*5G, grant ÕÜF11)
- Estonian IT Academy project Sustainable Artificial Internet of Things (SAIoT, grant TEM-TA138).

Abstract

Microfluidic Droplet Detection, Classification and Quality Assessment for Embedded Flow Cytometry Systems

Imaging flow cytometry (IFC) is a high-throughput technique with wide applications in immunology, cancer research, and cell biology thanks to its high-content screening capabilities. Recent research has focused on precise detection and classification using machine learning, particularly deep learning, on computationally powerful setups. In parallel, there is also an increasing interest for developing portable applications in the field, such as point of care, which has triggered an increasing demand for implementing IFC systems on embedded platforms. However, implementing droplet classification models in resource-constrained environments faces several key challenges, in particular limited computational power and memory, energy efficiency requirements, difficulties in capturing blur-free images in dynamic microfluidic environments, and the need for lightweight image quality assessment mechanisms. These limitations adversely impact model performance and accuracy, which is a significant obstacle to achieving fast inference times while maintaining acceptable accuracy levels. This PhD thesis addresses these challenges through three contributions.

First, the thesis introduces a novel, lightweight imaging pipeline for acquiring good quality, blur-free images of microfluidic droplets without additional hardware, making it suitable for portable applications. The proposed lightweight algorithm effectively detects, counts, and acquires single object images from video streams by addressing issues of motion blur and duplicate droplet image acquisition. Experimental evaluation using videos of fast-moving droplets showed 100% counting accuracy and throughput of 500 droplets per second (DPS) on Desktop PC, 40 DPS on Raspberry Pi-4, and 100 DPS on Nvidia Jetson Nano; the latter outperforms existing similar studies.

Second, the thesis proposes a new no reference image quality assessment (NR-IQA) hybrid model comprising a modified MobileNet model followed by a transformer block, specifically developed for microfluidic droplet IQA. Results show that the proposed hybrid model yields competitive performance metrics (Spearman rank-order correlation coefficient: 0.73; Pearson linear correlation coefficient: 0.75), and inference time on a Raspberry Pi-5 of 110 ms (9 frames per second).

Third, the thesis presents a new customized YoloV4-tiny model for rapid inference on resource-limited platforms. This model enables rapid inference while maintaining high classification accuracy for categorizing droplets as empty, containing a single cell, or containing multiple cells. The results show an inference time of 13 ms per classification and mean Average precision at intersection over union (mAP@0.5) of 99.95% on a Raspberry Pi-5. Additional training and deployment of a YoloV5s model showed the potential for even faster processing through specialized hardware acceleration at the cost of reduced performance (5.34 ms, mAP@0.5: 55.09% on the MaixCam board).

The results of this PhD thesis shows that through careful system design and model optimization, robust performance in microfluidic droplets based IFC can be achieved on platforms with limited computational resources, which paves the way for portable applications in the field.

Kokkuvõte

Tilkade tuvastamine mikrofluidikas, nende klassifitseerimine ja kvaliteedi hindamine sardsetes voolutsütomeetria süsteemides

Kujutise fikseerimise voolutsütomeetria (Imaging flow cytometry: IFC) on efektiivsel monitoorimisel põhinev võimekas tehnoloogia kasutamiseks immunoloogias, vähiuuringutes, ja rakubioloogias. Käesolev uuring on fokuseeritud tilkade täpselt detekteerimiseks ja klassifitseerimiseks kasutades masinõpet koos osalise süvaõppega suure võimekusega arvutuslikes lahendustes. Valdkonnas on täheldatav aina kasvav huvi mobiilsete (kaasaskantavate) realisatsioonide arendamiseks hooldussõlmede loomise näol, mis on omakorda käivitunud suurema vajaduse realiseerida IFC süsteeme sardsetele platvormidele. Siiski, realiseerides tilkade klassifitseerimise lahendused piiratud vahenditega keskkondades põrkutakse mitmetele võtmetähtsusega väljakutsetele, nagu piiratud arvutuslik võimsus ning mäluah, energiatarbe efektiivsuse nõuded, raskused saada hägustuseta kujutisi dünaamilistes mikrofluidika keskkondades, ning vajadus lihtsate kujutise kvaliteedi hindamise mehhanismide järele. Loetletud piirangud mõjutavad realisatsioonide koostamist ja täpsust, mis on oluline takistus saavutamaks kiireid liideseid vastuvõetava täpsusega. Käesolev dokortöö panustabki mainitud väljakutsetele lahenduste otsimisele läbi allpool käsitletud kolme temaatika.

Esiteks. Dokortöö tutvustab uudset ja lihtsat torujuhtmes olevat tilkade kujutist saamiseks hea kvaliteediga hägustuseta mikrofluidika tilkade kujutist kasutamata täiendavat riistvara võimaldamaks mobiilset (kaasaskantavat) rakendust. Väljapakutud algoritm detekteerib, loendab ning fikseerib üksikute tilkade kujutised terviklikust videovoost lahendades ka liikumisest tingitud hägustumise ning tilkade duplitseerimisest tekkinud olukordi. Kiiresti liikuvate tilkade videote eksperimentaalne hindamine näitas 100% hindamise täpsust 500 tilka sekundis (Droplet per second: DPS) olevat läbilaskevõimet personaalarvuti puhul, 40 DPS Raspberry Pi-4 puhul ning 100 DPS Nvidia Jetson Nano puhul; viimane ületab teadaolevaid teisi sarnaste uuringute tulemusi.

Teiseks. Dissertatsioonis tutvustatakse uut kujutise kvaliteeti mittemõjutavat kvaliteedi hindamise hübriidset rakendust (No Reference-Image Quality Assessment: NR-IQA), mis põhineb modifitseeritud MobileNet mudelil, millele järgneb spetsiaalselt mikrofluidika tilkade (IQA) tarbeks konstrueeritud muundurplokki. Saadud tulemused näitavad, et hübriidne mudel pakub võrreldavat võimekuse meetrikat (Spearman korrelatsioonikordaja: 0.73, Pearsoni lineaarne korrelatsioonikordaja: 0.75 ning liidese aeg Raspberry Pi-5: 110ms (9 kaadrit sekundis),

Kolmandaks. Dissertatsioonis pakutakse välja uus YoloV4 lihtne sobitatud mudel piiratud ressurssidega platvormidele kiirete liideste tarbeks. Mudel võimaldab kiireid liideseid säilitades klassifitseerimise täpsuse tilkade kategoriseerimises, nagu tilka ei ole (tühi), sisaldab ühte tilka, või sisaldab mitut tilka. Tulemused näitavad liidese viidet 13ms klassifikaatori kohta keskmistatud täpsusega ristumiskohal (mAP@0.5) 99.95% kasutades Raspberry Pi-5 lahendust. Koos täiendava treeninguga näitas YoloV5 mudeli rakendus potentsiaali veelgi kiiremaks protsessi toimumiseks kui kasutati spetsialiseeritud riistvaralist kiirendit vähendatud võimekuse juures (5.34 ms, mAP@0.5: 55.09%, MaixCam plaat).

Doktoridissertatsioonis saavutatud tulemused näitavad, et läbi hoolsa süsteemi disaini ja mudeli optimeerimise on võimalik saavutada IFC-l põhinev suur jõudlikus mikrofluidika tilkade määratlemisel limiteeritud arvutusliku võimekusega platvormidel, mis sillutab teed valdkondlikele kaasaskantavatele rakendustele.

Appendix 1

I

F. Afrin, K. Pärnamets, Y. Le Moullec, A. Udal, A. Koel, T. Pardy, T. Rang, "Embedded Blur-Free Single-Image Acquisition Pipeline for Droplet Microfluidic Imaging Flow Cytometry (IFC)," in *IEEE Access*, vol. 12, pp. 92431-92441, 2024, DOI: <https://doi.org/10.1109/ACCESS.2024.3421637>

Received 23 May 2024, accepted 19 June 2024, date of publication 1 July 2024, date of current version 12 July 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3421637

RESEARCH ARTICLE

Embedded Blur-Free Single-Image Acquisition Pipeline for Droplet Microfluidic Imaging Flow Cytometry (IFC)

FARIHA AFRIN¹, (Member, IEEE), KAISER PÄRNAMEETS¹, (Member, IEEE),
YANNICK LE MOULLEC¹, (Senior Member, IEEE), ANDRES UDAL², (Member, IEEE),
ANTS KOEL¹, (Member, IEEE), TAMAS PARDY^{1,3}, (Member, IEEE),
AND TOOMAS RANG^{1,3}, (Senior Member, IEEE)

¹Thomas Johann Seebeck Department of Electronics, Tallinn University of Technology, 19086 Tallinn, Estonia

²Department of Software Science, Tallinn University of Technology, 19086 Tallinn, Estonia

³Department of Chemistry and Biotechnology, Tallinn University of Technology, 19086 Tallinn, Estonia

Corresponding author: Fariha Afrin (fariha.afrin@taltech.ee)

This work was supported in part by Eesti Teadusagentuur (ETAg) under Grant PRG620, in part by Estonian Research Council project under Grant PSG897, and in part by Estonian Education and Youth Board Project Artificial Intelligent Internet of Things (AIoT)*5G under Grant ÕÜF11.

ABSTRACT Good quality of single droplet image acquisition in imaging flow cytometry (IFC) is crucial for a wide range of biological analyses. Recently, there have been significant advances in droplet microfluidic data analysis; however, acquiring blur-free single object images is still a great challenge because of the tradeoff between high flow rate and hardware setup complexity and cost. State-of-the-art hardware setups for blur-free single image acquisition are often complex, cumbersome, and not portable, limiting their suitability for point-of-care diagnostics. Moreover, motion blur and duplicate droplet image acquisition can occur with flow rate variation. To address these issues, this paper proposes a lightweight imaging pipeline for acquiring blur-free single droplet images for portable applications; this pipeline is capable of acquiring every single droplet image. While most of the existing literature focuses on complex hardware setups, utilizing high frame rate cameras that are not cost effective and complex optical solutions for droplet focusing, our pipeline utilizes minimum hardware and a lightweight algorithm for detecting, counting, and acquiring single object images from the video stream. The proposed pipeline was evaluated experimentally using videos of fast-moving droplets in which the input fluid flow rate was as high as 67.7 $\mu\text{L}/\text{min}$. The proposed pipeline achieves 100% counting accuracy on the tested videos and 2 ms, 25 ms and 10 ms processing time for each droplet on a desktop PC, single-board computer Raspberry Pi-4, and Nvidia Jetson Nano, respectively. This yields a maximum of 500, 40, and 100 blur free detected droplets per second (DPS), respectively. The Jetson Nano implementation, achieving 100 DPS with processing time of 10 ms, is faster than existing similar studies and fast enough for the target application. The results suggest that the proposed lightweight pipeline is suitable for efficient single object image acquisition in IFC on an embedded portable platform.

INDEX TERMS Droplet, image acquisition, imaging flow cytometry, microfluidic, single-board computer.

I. INTRODUCTION

Imaging Flow Cytometry (IFC) is a combination of optical image acquisition, often involving microscopy, and flow cytometry; IFC is a key enabling technology in the drug

discovery [1], [2] and medical diagnosis fields [3], [4]. IFC enables diverse morphology analyses, which are performed on a single-object image using a fully automated or partially automated system in laboratory or point of care (POC) devices [5]. Hence, a high-quality single-object image is indispensable for the accurate analysis of the specific object, as well as for machine learning/deep learning training [6], [7].

The associate editor coordinating the review of this manuscript and approving it for publication was Mahmoud Al Ahmad¹.

The analytical results are used in several biomedical applications [8], such as early disease detection and circulating tumor cell detection [9], [10].

However, the process of acquiring a blur-free single-object image is challenging in high throughput IFC, especially if the camera settings are not synchronized with the flow of the object. For example, motion blur can happen if the object flow is too fast relative to the camera exposure time, and noisy images can be captured if the camera exposure time is too long [11], [12].

Some existing studies have focused on developing a dedicated optical system for acquiring single-object images at a high throughput and speed which makes the whole system bulky and expensive [13], [14]. Other examples such as the commercial IFC system “ImageStream” [15] as well as some other IFC systems, achieve cell focusing by using a precise pump to control the fluid flow rate and a velocity detection subsystem to control the time-delay integration readout [1]. Parallelized microchannels are also used to decrease the flow speed for a specific throughput and to obtain a single bright field object image [16], [17].

Moreover, the optical time-stretch imaging method, which consists of a complex optical setup, enables single-image acquisition and exploits spatial and temporal dispersion [18], [19]. In [20], the particle flow, controlled for fixed speed and motion, is frozen by the ultrafast shutter speed of the camera to obtain blur-free images. In [21], an optomechanical virtual motion freezing fluorescence imaging (VIFFI) method was developed for high throughput (> 10,000 cells/s) imaging. However, it has a complex hardware setup that consists of an excitation beam scanner, a speed-controlled polygon scanner, and a series of timing control circuits in order to increase the exposure time of the image sensor.

The integration of these additional hardware systems for controlling the flow increases the complexity and decreases the cost efficiency. Image reconstruction can be applied without integrating additional hardware, yet this requires blur identification and data preprocessing [22].

Other studies focused on applying machine learning or deep learning algorithms to acquire and then analyse single-droplet images. A demonstration of TensorFlow feasibility in classifying red blood is presented in [23]. AI assisted pathogen detection with 99.8% classification accuracy at the edge device was proposed in [24].

It should be noted that most of the research work focuses on detection accuracy but overlooks the trade-off between speed and hardware requirements. Although a portable flow cytometer that uses deep learning to detect Giardia lamblia cysts in water samples is reported, its data processing module still ran on a desktop PC [25]. Moreover, these are post-experimental processes and are challenging to apply in a single-board computer (SBC) platform in real time applications since data acquisition and storing large number of images for further morphology analysis are limited by memory capacity.

The closest work to ours is deemed to be the portable and computer vision-based detection platform reported in [26] which can automatically classify the results from saved data without the need of professionals. A Raspberry Pi-4 SBC has been used for computing and a control module is used for controlling the data processing. However, the software code complexity, detection per second and processing time are not reported nor discussed in the paper.

To sum up, the number of research papers directly comparable with our work are limited; the closest ones and their key features are summarized in Table 1 (in addition, comparative results are provided in Table 3, see Section III).

TABLE 1. State-of-the-art for blur free image acquisition and portable IFC vs. our work. complexity, portability, and cost are rated from low/no to high/yes (I.E. + to +++).

Techniques	Complexity	Portability	Throughput/accuracy	Cost	Summary of the methodologies
Virtual-freezing [21]	+++	+	Throughput >10,000 cells/s; classification accuracy 95.3%	+++	Acquires blur free single cell image by using a sophisticated system that includes a flow-controlled microfluidic chip, a speed-controlled polygon scanner, and a series of precise timing control circuits.
Deep learning [22]	+	+	96.6% classification accuracy	++	Data processing runs on the computer
Machine learning [24]	+++	++	Detection rate 93.8%; classification rate 99.8%	+++	Optofluidic detection at the edge device relies in the time trace
Deep learning [25]	+	+++	Autofocus on a single object in ~7 ms	++	Data processing runs on the computer
OpenCV [26]	+	+++	Classification and counting accuracy are above 95%	+	The detection and counting system operate after saving 10 videos (RPi4 platform). However, the processing time for each object detection is not reported in the paper.
Proposed method	+	+++	Counting accuracy is 100% for the tested videos	+	Acquires blur free single droplet image by using a lightweight pipeline on embedded platform

To summarize, existing IFC systems are for laboratory grade applications and exhibit high complexity and high cost in order to perform droplet detection. Moreover, the captured blur-free images are transferred to desktop PCs for analysis, which is not suitable for portable devices. The development of a lightweight and reliable pipeline remains a crucial open issue for reducing the software code complexity so that it can run on an embedded platform. To bridge this gap, we built

a cost-effective pipeline which can acquire blur-free images and count each droplet automatically on a portable platform, without a laboratory environment.

A. CONTRIBUTION

The main goal of this work is to acquire blur-free single droplet images without any camera synchronization to handle the complexity and computational cost challenge [27], [28].

Existing systems rely on complex hardware setups comprising high frame rate cameras that are not cost effective and complex optical solutions for droplet focusing. In contrast, we present a pipeline for acquiring single-droplet images that utilizes minimum hardware and a lightweight algorithm for detecting and acquiring single object images from video stream. In addition to single-image acquisition, our pipeline can count the total number of generated droplets. The proposed pipeline adopts color-based detection and a Euclidian distance tracker for droplet tracking in real time.

Consequently, our pipeline can be deployed on different platforms ranging from high-performance device (desktop PC) to hardware-accelerated device (Nvidia Jetson Nano) and to resource-constrained device (RPI4) for acquiring blur free single object image and counting the total number of objects. Experiments using video of fast-moving droplets (the input fluid flow rate was as high as $67.7 \mu\text{L}/\text{min}$) show that the proposed pipeline achieves 2 ms, 25 ms and 10 ms processing time for each droplet on Desktop PC, Raspberry Pi-4 and Nvidia Jetson Nano, respectively. This yields a maximum of 500 DPS on a desktop PC and 100 DPS on a resource-constrained platform, which is fast enough to be utilized for high volume sample analysis where droplet generation rate lies to 100 to 500 Hz in IFC [29]. As counting and acquisition of single droplet image are both accomplished using lightweight detection and tracking algorithm, it is insensitive to camera synchronization. As multiple droplets can be detected simultaneously, the overall setup does not require an additional subsystem such as velocity detector to capture images of droplets under high flow rate. We also conducted experiments using external reference data (see Section III-C) to validate the proposed pipeline.

This proposed pipeline can automatically acquire blur-free single droplet image and subsequent droplet counting on resource constrained embedded platform at a rate of 100 DPS. Therefore, we believe that the pipeline presents a promising solution for future POC diagnostics, more specifically in object classification in droplets, e.g. for antimicrobial susceptibility testing, where the significantly higher frequency droplet-based assays could significantly contribute to sepsis prevention. At present, droplet-based imaging flow cytometry platforms primarily rely on high-end, high-power instrumentation (e.g. GPUs), which limits embedded or portable applications and scalability due to high associated component costs.

The rest of this paper is organized as follows: Section II presents the experimental setup including the materials used,

the proposed detection and tracking method and its complexity analysis, and the process to acquire blur-free single droplet image acquisition along with counting. Section III presents the experimental results, and their analysis and Section IV provides a comparison against the closest related works and discussion thereof. Finally, Section V summarizes the key findings and limitations of this work and outlines possible future work.

II. MATERIALS AND METHODS

A. DROPLET GENERATION VIDEOS

The proposed pipeline has been tested firstly with simulated two-phase flow videos and secondly with real-world experimental fluorescent videos. Before explaining how these simulated and real-world videos are obtained, we first briefly introduce the physical system; the microfluidic droplets are generated inside a polydimethylsiloxane (PDMS) chip (schematic shown in Figure 1(a)). The microfluidic chip design is based on the real-life microfluidic chip [30] used in previous experiments. The chip contains water and oil inlets, a gas spring connector (denoted air inlet in Figure 1(a)), a flow-focusing junction with a width of $90 \mu\text{m}$ and a height of $100 \mu\text{m}$, and an outlet. The oil inlets are equipped with filters that capture any particles that may be present in mineral oil (Sigma Aldrich 330779) in order to prevent the microchannels from becoming clogged.

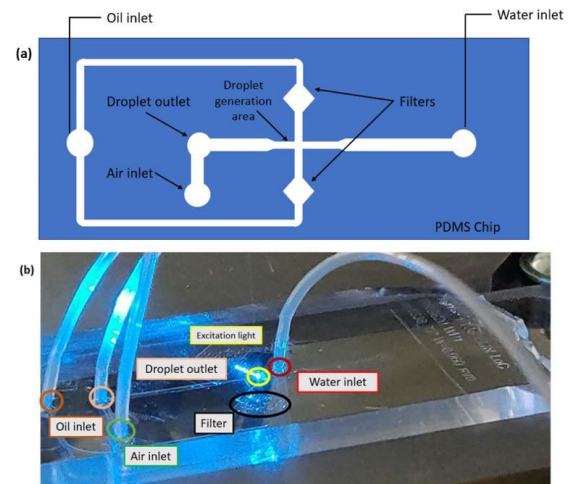


FIGURE 1. Droplet generation. (a) Schematic of the PDMS chip (not drawn to scale). (b) Experimental setup showing all (water, oil, air) inlets, the droplet outlet, the filter, and the excitation laser light. The power supply for driving the laser, the computer for controlling the syringe pump, and the Basler Ace camera are not visible here.

Firstly, the simulated videos are obtained from a two-phase flow simulation model (implemented in COMSOL Multi-physics version 5.6 in our previous work [30]), where water and oil flows are maintained at $10 \mu\text{L}/\text{min}$ and $67.7 \mu\text{L}/\text{min}$, respectively, to generate droplets. The flow rate of oil is 6.77 times greater than that of water; this helps to keep

the droplets separated (relatively large distances between droplets make it easier for the simulation to avoid “jetting” in the channel). The liquid phases in COMSOL (see Fig. 3(A)) are displayed as red and blue. The central color, green, indicates the region where the relative concentration of water is 50%.

Secondly, for the experimental video, deionized water mixed with 10 $\mu\text{g/ml}$ fluorescein isothiocyanate- dextran (FITC) and oil through corresponding inlets are used in the fabricated PDMS chip channel. The mixture and oil are injected through the chip inlet via a 3 mL syringe. The flow rates of the water and oil fluid are maintained by using a syringe pump (SpinSplit Netpump, SpinSplit LLC, Budapest) and air inlet has not been used in our experiment. As shown in Figure 1(b), a 488 nm laser light (Sharp GH04850B2G with focusing optics from AliExpress) is used as the excitation light source; it is located right after the junction where the droplets from [31].

In the experiments, the droplet flow through a PDMS chip is captured by a Basler Ace camera at 1000 frames per second (fps).

B. DROPLET DETECTION AND TRACKING

To implement the imaging pipeline, custom code utilizing color-based detection and a Euclidean distance tracker algorithm are combined. The pipeline was developed in Python 3.8 using OpenCV 4.5.5.

The acquired imaging results were exported automatically in suitable formats (e.g., .tiff, .png, .gif, .jpeg, and .bmp), and they could be used for further data analysis involving, for example, classification, or morphology analysis. The pipeline is also able to simultaneously count the total number of droplets during the time of detection.

Figure 2 shows the three main steps of the pipeline: (1) video streaming; (2) processing; and (3) result acquisition. The pipeline begins by acquiring the video input in real time and then processes it to obtain single-droplet images. This processing step can be divided into two main parts: detection and tracking. During tracking, the pipeline can also count the total number of detected droplets. Finally, the results are

saved. What follows presents the details of the detection and tracking methods implemented in our pipeline.

We conducted experiments with both the color-based detection and Background Subtraction Mixture of Gaussians-2 (BS-MOG2) methods. The color-based detection method is more efficient than BS-MOG2 algorithm because it does not fail if there is intensity variation in the experiments [32], [33]. Indeed, BS-MOG2 requires history and threshold input values, which in our case were selected to be 100 and 50, respectively, for the best optimization of the results. This history value defines how many last frames would affect the background model. BS-MOG2 creates a binary mask, where 0 represents the background and 1 represents the foreground, which works well for simulated video but fails in an experimental environment [34]. On the other hand, in our experiments, color-based detection works well for both environments. Mask was created by identifying the high and low range of the object color using HSV (Hue Saturation Value) range. To exclude the noise in both algorithms, the contours of the moving objects on the mask were passed through a threshold value (100 pixels), see Figure 5. Any objects above the threshold value were delimited by a bounding box. Once the object was detected inside the bounding box, each center position of the bounding rectangle was inserted in a single array.

Next, all the positions were fed to the tracker to update the unique ID for each droplet; a Euclidean distance tracker algorithm has been used to track each droplet. When a new droplet enters into region of interest-1 (ROI-1), the pipeline checks the central distance between consecutive droplets. An object generally follows a trajectory made up of points that are very close to each other. The closer the previous point is to the object, the greater the probability that the previous position belongs to it; in this case the tracker considers a droplet having a small distance (< 10 pixels) as the same object instead of new object. Therefore, it does not change the unique ID for this specific droplet. The algorithm calculates the Euclidean distance as follows [35]:

$$D = \min(\sqrt{(a_j - a_i)^2 + (b_j - b_i)^2}) \quad (1)$$

where D is the Euclidean distance of the previous tracks and the current tracks, a is the x- coordinate of the track box, b is the y-coordinate of the track box, j represents all the prediction boxes in current frame, and i represents all the prediction boxes in the previous frames.

The main purpose of tracking is to ensure precise droplet counting by preventing duplicate counting of droplets and to perform single-image acquisition at the same time. The tracking yields every single-droplet image without any duplication and counts the total number of droplets. The process of tracking a droplet consists of assigning it a unique ID that it retains until it leaves ROI-1. The algorithm performed well near the droplet generation area, where the single flow of droplets does not become occluded. However, after a short while, when the droplets reach the end of the channel,

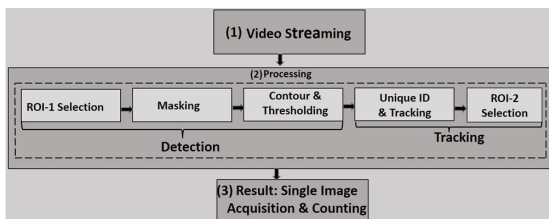


FIGURE 2. Block diagram of the single-image acquisition process flow, consisting of (1) video streaming, which occurs at 1000 fps; (2) processing, which involves detecting droplets based on their color and tracks; and (3) result acquisition, which involves acquiring a single image and counting the number of droplets using the tracking result without any duplication.

they become occluded and lose their unique IDs. To solve this issue, and to save images as well as count the number of droplets, another region, ROI-2 (shown in Figure. 3), has been selected.

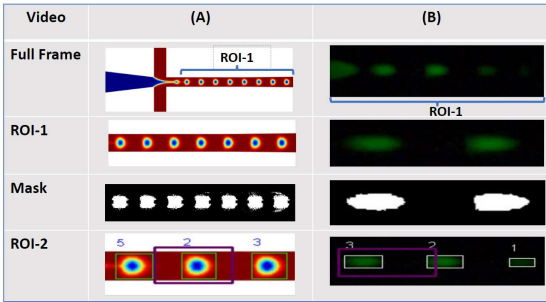


FIGURE 3. Detailed video processing diagram showing the full frame, ROI-1, mask, and ROI-2. (A) COMSOL-simulated 2-phase flow video and (B) experimental fluorescence video captured by the Basler Ace camera at 1000 fps. For both cases, full frame shows the droplet generation inside the chip channel; ROI-1 represents only the specific small portion of the channel where the droplet flow can be observed; mask shows the object in white and the background in black (the droplet is flowing from left to right); and ROI-2 represents the small initial portion of the channel from the left side (purple region) used to obtain true counting.

C. IMAGING PIPELINE'S PSEUDO CODE AND COMPLEXITY ANALYSIS

This section illustrates the imaging pipeline's pseudo code and its time and space complexity analysis.

D. TIME AND SPACE COMPLEXITY ANALYSIS

The time complexity of the pipeline depends on the dominant factors (video processing loop, masking, contour detection, contour filtering, and object tracking) contributing to the overall running time.

The primary loop iterates over each frame (denoted by n) of the video. The time complexity associated with each pixel in the frame can be approximated, with the big O notation, as $O(m \times p)$ where m and p , are the frame dimensions. The time complexity for contour filtering depends on the number of contours and expressed as $O(c \times f)$ where c is the number of contours and f represents the time complexity associated with the operation within the contouring filter. $O(t \times g)$ is the time complexity for tracking where t is the number of tracked object and g represents the time complexity associated with the operations within the object tracking loop. Hence, the overall time complexity can be expressed as $O(n \times (m \times p) + c \times f + t \times g)$.

The space complexity for each frame (input data) is constant as it is processed one at a time, and the total space complexity for handling video frames is negligible. The Euclidean distance tracker maintains a dictionary to store the center positions of tracked droplets and an integer to keep track of droplet IDs. If s is the number of tracked droplets, then the space complexity of the tracker is $O(s)$. As the counting area polygon has a constant number of

Pseudo Code for Proposed Pipeline

Data: Video frame

Result: Acquire single droplet image and count

```

1. tracker = EuclideanDistTracker()
Initialize Euclidean distance tracker
2. create_directory('data') // Create a directory named 'data'
for saving droplet images
3. object_count = set() // Initialize droplet count set
4. Process frames in a loop
while True:
    frame = read_frame(cap) // Read a frame from the video
    Hsv_img = convert_to_hsv(frame) // Convert the frame to
    HSV
    mask_green = create_green_mask(Hsv_img, Low_Green,
    High_Green) // Create mask using color thresholding
    contours = find_contours(mask_green) // Find contours in
    the mask
    detections = detect_objects(contours) // Detect droplets
    based on contour area
    boxes_ids = tracker.update(detections, 1) // Update droplet
    tracking using the Euclidean distance tracker
5. Process each detected and tracked object
for box_id in boxes_ids:
    x, y, w, h, obj_id = box_id
    draw_rectangle(frame, x, y, w, h) // Draw bounding box
    draw_text(frame, str(obj_id), x, y) // Text on the frame
    result = point_in_polygon(roi_polygon, (x, y)) // Check if the
    droplet is inside the defined counting area
    if result:
        object_count.add(obj_id) // Update droplet_count based on
        the object ID
    save_droplet_image(frame, obj_id) // Save image of the
    detected droplet
    total_count, objects_in_second = calculate_counts(object_count, fps_count) // Calculate total
    droplet count and droplets per second
    display_counts(frame, total_count) // Display the droplet
    count on the frame
7. Check for user input to exit the loop
key = wait_for_key()
if key == 27:
    break
8. Cleanup and release the video capture object
release_video_capture(cap) // Release the video capture
object
destroy_all_windows() // Destroy all open windows
End of Algorithm

```

vertices the space complexity is $O(1)$. Moreover, if the total number of unique droplets detected is denoted by q , the space complexity is $O(q)$. A dictionary is needed to store images of detected droplets, so the space complexity depends on the number of detected objects $O(q)$. Various variables used for masking, contour detections, and bounding boxes have space

complexity which is proportional to the size of the data, and these are temporary, hence do not contribute significantly to overall space complexity. Therefore, the overall space complexity is dominated by the Euclidean distance tracker i.e. $O(s)$ and the sets used for counting objects i.e. $O(q)$. Finally, the overall space complexity can be expressed as $O(s+q+I)$.

Memory usage in Figure 4 shows that it does not exhibit any significant spikes or drops throughout the overall processing. It takes approximately 95 Mebibytes (MiB) maximum.

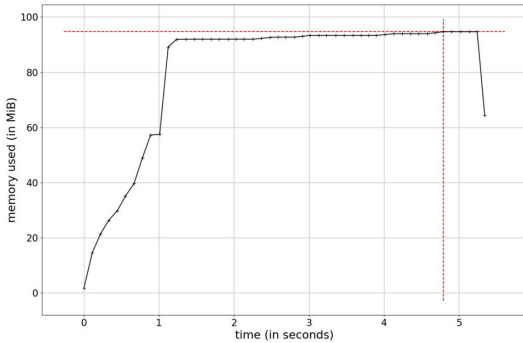


FIGURE 4. Memory consumption as a function of time. The horizontal axis shows the time of overall processing for whole video and the vertical axis shows memory consumption trend in Mebibytes.

The time complexity of the closest related paper [25] depends on the dominant factors (Background difference method, Gaussian filtering, Binarization, Opening and Expansion processing). For their case, let's assume b is the number of pixels, so time complexity would be $O(b)$ for background difference. If v is the size of gaussian kernel, time complexity would be for gaussian filtering would be $O(b \times v^2)$. The time complexity for binarization depends on the number of pixels in each image, hence $O(b)$. If h is the number of pixels in the structural element time complexity for opening and expansion process would be $2 \times O(b \times h)$. The overall time complexity would be $O(b) + O(b \times v^2) + O(b) + 2 \times O(b \times h)$ which simplifies to $O(b \times (1 + v^2 + 2 \times h))$.

The memory required for storing the difference image, filtered image, binarized image, opening operation and expansion processing is proportional to the size of each frame, resulting $O(b)$ where b is the total number of pixels in each frame. Table 2 shows the comparison of computational complexity of the proposed work and the closest related paper. The complexities are not provided in earlier works and hence are derived based on materials available in their paper.

In our proposed method, the dominant terms are n , m and p , while the second and third terms are constants and multiplied by f and g , respectively. On the other hand, the time complexity of the method in [26] involves the product of four variables which makes it more computationally complex, while its space complexity is lower.

TABLE 2. Comparison of computational complexity with closest related paper.

Reference	Time Complexity	Space Complexity
OpenCV [25]	$O(b \times (1 + v^2 + 2 \times h))$	$O(b)$
Proposed Method	$O(n \times (m \times p) + c \times f + t \times g)$	$O(s+q+I)$

E. SINGLE-DROPLET IMAGE ACQUISITION AND COUNTING

In this process, a single-droplet image is acquired by automatically cropping each detected droplet from the bounding box. The location of one corner (instead of all four corners) of the bounding box is stored in a dictionary to reduce the complexity of the computation, and the object image is cropped when the particle is exactly inside of a specific ROI-2 (shown in Figure 5).

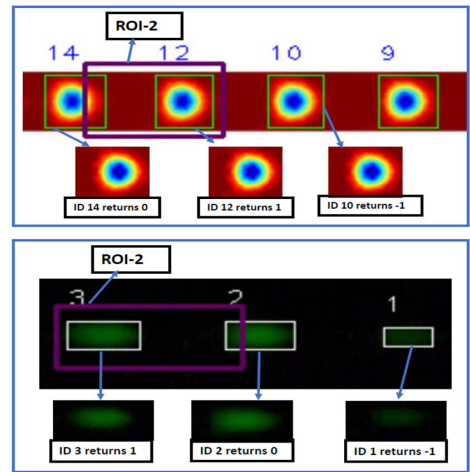


FIGURE 5. Three different conditions for accurate counting (Top: COMSOL simulated two-phase flow video, bottom: experimental fluorescence video): (i) if the green bounding box is outside ROI-2, the operation returns -1; (ii) if the bounding box is exactly on the boundary of ROI-2, it returns 0; and (iii) if the bounding box is inside ROI-2, it returns +1, which is the right condition to store the location.

The geometric operation has been conducted for three different conditions: (i) if the bounding box is outside ROI-2, the operation returns -1; (ii) if the bounding box is exactly on the boundary of ROI-2, it returns 0; and (iii) if the bounding box is inside ROI-2, it returns +1.

The location is stored only if the operation returns +1; otherwise, the algorithm does not take any action. The algorithm checks whether a box with the same unique ID has already been stored or not. If the box has already been stored, the algorithm will not store it again, thus preventing duplication. Although the main contribution of this algorithm

is to acquire a blur-free single-droplet image from a fast flow, the counting of droplets is also performed in near real time. This entire process does not require any additional hardware to acquire single-droplet images or to perform counting.

F. OVERVIEW OF THE COMPUTING PLATFORMS

We have implemented our pipeline on three different platforms, ranging from high-performance device (desktop PC) to hardware-accelerated device (Nvidia Jetson Nano) and to resource-constrained device (RPI4 Model B).

1. Desktop PC with an Intel i5-10210U processor (4 cores, 8 threads) running at 1.6 GHz and with 16 GB RAM. The OS is Windows 10.
2. Raspberry Pi 4, based on quad-core ARM Cortex-A72 processor clocked at 1.5GHz. It has 4 GB of RAM and is used in headless mode.
3. The Jetson Nano is also a compact device (NVIDIA Jetson NANO Developer Kit); it has a 128-core Maxwell GPU and a quad-core ARM A57 CPU. 1.434 GB of LPDDR4 [36].

We first tested the droplet image acquisition and droplet counting on the Windows-based PC and then on the two embedded platforms.

III. RESULTS

To validate the proposed pipeline, we performed experiments using A) 2-phase simulated video, and B) experimental fluorescent droplet generation video (see Section II-A). The flow rate is 100 $\mu\text{l}/\text{min}$, and the camera frame rate is 1000 fps. The fluid flow is of high throughput, which can easily create motion blur when images are captured using a camera. Image quality degradation does not take place if the flow rate is low, but it can occur for a high throughput. The proposed pipeline can handle moving droplet videos to obtain images of good quality.

A. DROPLET DETECTION AND TRACKING

Contour area selection from the mask resulted in successful droplet detection. White pixel values greater than 100 were considered above the threshold value for accurate droplet area detection. Values below the threshold were considered as noise and were not included in further processing.

B. SINGLE-DROPLET IMAGE ACQUISITION AND COUNTING

Every detected droplet was tracked using its unique ID and counted. The total droplet count increases whenever a new object enters ROI-2. A 0.05 s video that contains 18 droplets was used for the experiment, and in each experiment 18 images were properly recorded. These numbers were validated through manual counting by three human observers. In addition, we observed that the processing time of our pipeline in two platforms for fast-moving droplet detection, counting and single-image capture.

The main goal of our proposed pipeline is capturing a single-droplet image from the near real-time fluid flow.

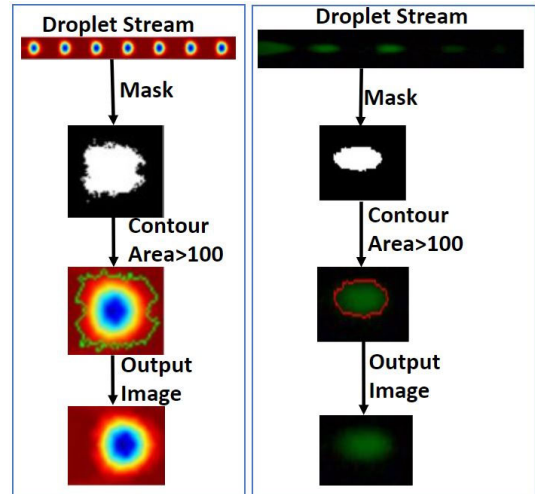


FIGURE 6. Manual thresholding if the contour area is greater than 100. Left: green contour for CFD (COMSOL) simulated video; right: red contour for experimental fluorescence video boundary. An area meeting the threshold criterion is considered to be an object, resulting in precise droplet detection.

The pipeline is suitable for capturing images of moving objects, not for capturing images of static objects. Since it can detect, count, and save each droplet successfully, it is suitable for any other similar single streaming IFC application.

The execution times for a single-droplet handling for the two different types of videos (simulated and experimental) using the two different algorithms (color-based and BS-MOG2) on the three platforms (desktop PC, RPI4, and Jetson Nano) are shown in Figure 7.

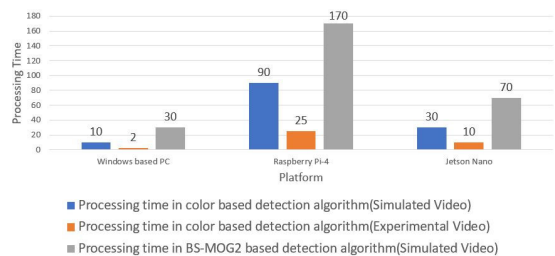


FIGURE 7. Comparison of single-frame processing time, where the horizontal axis represents the platforms, and the vertical axis represents the single-frame processing time in milliseconds(ms). The color corresponds to the algorithm and video type used. The color-based detection algorithm takes 10 ms, 90 ms, and 30 ms on the three different platforms, respectively, while the background subtraction algorithm takes 30 ms, 170 ms, and 70 ms for the simulated video. Experimental video yields the best result using color-based detection algorithm, which is 2 ms, 25 ms, and 10 ms on the desktop PC, Raspberry Pi 4, and Jetson Nano, respectively.

The color-based detection algorithm yields the minimum processing time for the experimental video on every platform; it takes 2 ms, 25 ms, and 10 ms on the desktop PC, Raspberry

Pi 4, and Jetson Nano, respectively. The same algorithm takes 3 to 5 times more time for the simulated video than it does for the experimental video (10 ms, 90 ms, and 30 ms, respectively).

Note that Figure 7 does not show results for BS-MOG2 on the experimental data; despite being adaptive and robust, the BS-MOG2 algorithm does not work on the experimental video. The reason is that BS-MOG2 is sensitive to Gaussian distribution of pixel and illumination variation. In this specific microfluidic droplet applications, the background is complex and dynamic because of the liquid flows and interactions. The single-droplet processing times with the BS-MOG2 algorithm on the simulated video are 30 ms, 170 ms, and 70 ms, for the Windows desktop PC, Raspberry Pi 4, and Jetson Nano, respectively.

Next, Table 3 shows the DPS and corresponding processing time for the two different algorithms on the two types of videos in three different platforms. As expected, due to its higher computational power, the desktop PC achieves the highest (500) DPS, with the color-based detection algorithm applied to the simulated video. However, it is also noticed that 100 DPS (10 ms to acquire one droplet image) can be obtained for the same algorithm and video with the Jetson Nano, which is considered as high throughput for the target applications.

TABLE 3. Comparison of DPS on the three different platforms.

Algori thm	Data	Platform	Detected droplets per second (DPS)	Corresponding processing time (ms)
Color based detecti on	COM SOL Simul ated Video	Windows- based desktop PC	100	10
		Raspberry Pi-4	11	90
		Jetson Nano	33	30
	Experi mental Video	Windows- based PC	500	2
		Raspberry Pi-4	40	25
		Jetson Nano	100	10
BS- MOG 2 based detecti on	COM SOL Simul ated Video	Windows- based PC	33	30
		Raspberry Pi-4	5	171
		Jetson Nano	14	70

C. TEST WITH EXTERNAL REFERENCE DATA

The pipeline was also tested with external reference brightfield microscopy video obtained with a similar chip design [37]. The microfluidic channel width was 50μm; the flow rates were controlled using neMESYS 290N syringe

pumps (Cetoni) and gas-tight syringes (Hamilton) connected to PTFE tubing, and the generation of droplet was monitored using a high-speed Mini UX-100 camera (Photron). Single object image acquisition and counting works well using our proposed pipeline on this external reference data. Figure 8 illustrates the single droplet image acquisition in brightfield microscopy.

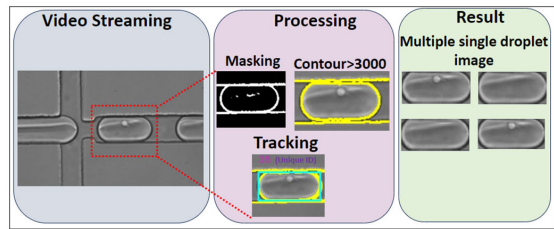


FIGURE 8. Single droplet image acquisition in brightfield microscopy (Left to right: video streaming, processing, and result). Masking (white boundary), contouring (yellow boundary) and tracking (green rectangle) shows only for one droplet from the video. Multiple saved droplet images are shown in result.

For automatic single droplet image acquisition and counting, the high and low range of the object color using HSV range needed to identify once to create mask. Then contour area selection from the mask was set with a threshold of 3000. This implies that white pixel values exceeding 3000 were taken into account the precise droplet area detection. Pixel values below this threshold were considered as noise and excluded from subsequent processing.

Each detected droplet was tracked using its unique ID and counted. The execution times for a single-droplet handling for brightfield microscopy image acquisition were not significantly different than that for the two-phase CFD simulation and fluorometry images as the overall process went through same steps.

IV. COMPARISON AND DISCUSSION

To summarize, previous similar (i.e. closest related, but not directly comparable) studies for acquiring blur free images, detecting, and counting methodologies are either complex and not developed for being deployed into portable devices ([21], [37], [24]), or meant for other types of applications ([24]), or they do not provide all performance details [25]). In contrast, our proposed solution is able to acquire blur-free images, detect and count accurately each droplet automatically, making it suitable for portable platform.

These closest related works are summarized in Table 4 and further discussed.

Blur-free droplet image acquisition, detection, and counting the total number of droplets in an embedded platform has emerged as essential research towards POC technology. As indicated previously, the closest related works shown in Table 4 are not directly comparable, but for reference their platform suitability, processing time, and accuracy metrics were looked into. Detection and counting of microfluidic

TABLE 4. Most-closely related results vs. our work.

References	Techniques	Platform	Processing time (ms)	Counting accuracy
[38]	Deep learning	High-end NVIDIA V-100 GPU	33	N/A
[21]	Sophisticated hardware setup	Desktop PC	NA	N/A
[25]	Deep learning	Nvidia RTX 2080 GPU	~34	N/A
[26]	OpenCV	Raspberry Pi	N/A	95.96 to 99.7%
This work	Lightweight pipeline using color-based detection and Euclidean distance tracker	Jetson Nano	10	Counting accuracy is 100% for the tested videos

droplets is obtained in [38] by utilizing YoloV5 which is fast enough to meet the requirement of imaging flow cytometry. However, this method has been optimized (processing time: 33 ms) for powerful GPU to meet the requirements of their specific application. A sophisticated hardware setup has been developed in [21] which can handle more than 10,000 cells/s throughput and acquire blur free images of cells (processing time not reported in their paper). However, the setup is for laboratory grade imaging flow cytometer. A deep learning based miniaturized imaging flow cytometer is developed for waterborne parasite detection which takes approximately 34 ms for autofocusing, color reconstruction and detection. Similar research was conducted targeting an embedded platform [26] by utilizing OpenCV for blood cell detection and classification but have not provided their processing time.

Contrary to existing systems that rely on additional hardware setup and complex techniques for blur-free image acquisition, object detection and counting which are often costly and not suitable for portable devices, our proposed pipeline is optimized for portability. It utilizes minimal hardware and lightweight color-based detection and Euclidean distance tracker for droplet tracking in real time algorithm for detecting, counting, and acquiring single object images from video stream. It requires approximately three times less processing time than the most closely related existing studies [25], [38], making it well suited for embedded platform. Memory complexity analysis of the proposed pipeline indicates a memory usage of 95 MiB, with most of it dedicated for storing the unique identification number for each droplet. The characteristics of the proposed work eliminate the need for complex and expensive hardware, thereby democratizing access to droplet microfluidic IFC.

V. CONCLUSION

Single-droplet image acquisition is important for extracting valuable information about a given particle. Because of the high throughput and fast flow of IFC, researchers have focused on developing hardware systems to capture blur-free single images. The pipeline proposed in this study enables near real-time image acquisition without the integration of any additional hardware components; video streaming is performed at 1000 fps in this paper.

The pipeline, which consists of a color-based detection algorithm, is capable of acquiring high quality single droplet image and counting the total number of droplets from a near real-time, single-line high throughput droplet flow. The pipeline’s output can be fed into any other algorithm for further analysis; it can be used to perform tasks such as classification, segmentation, or morphology analysis. Droplet detection is challenging when the droplets have a high velocity and occlusions happen after some period, but the proposed pipeline is reliable enough to solve these issues.

The complete process is automated, and it can detect and count droplets, as well as acquire blur free single-droplet image in near real time. Identification of low and high color ranges of the object to create mask as well as setting manual threshold is used only once in initial phase. Despite the constraint of limited data, the proposed pipeline was tested using external data for validation. The pipeline was tested on a desktop PC and two embedded platforms, and the processing times of two different object detection algorithms were compared. The achievable droplet detection per second value is 100 DPS on the Jetson Nano SBC. The processing time on desktop PC is approximately five times faster; hence, it yields 500 DPS. For future work, an extension to this work would be to adapt auto-thresholding and the deployment of the pipeline in other embedded platforms to assess and compare their performances. Moreover, real time-analysis, exploration of hardware and software optimizations for possibly improving the performance and/or efficiency of the implementation is a future research direction and extension of the proposed work.

ACKNOWLEDGMENT

The authors would like to thank to Assist. Prof. Tomasz Kaminski, University of Warsaw, Poland, for providing us with the validation data.

REFERENCES

[1] D. A. Basiji, W. E. Ortyn, L. Liang, V. Venkatachalam, and P. Morrissey, “Cellular image analysis and imaging by flow cytometry,” *Clinics Lab. Med.*, vol. 27, no. 3, pp. 653–670, Sep. 2007, doi: 10.1016/j.cll.2007.05.008.

[2] Y. Han, Y. Gu, A. C. Zhang, and Y.-H. Lo, “Review: Imaging technologies for flow cytometry,” *Lab Chip*, vol. 16, no. 24, pp. 4639–4647, 2016, doi: 10.1039/c6lc01063f.

[3] B. S. Edwards and L. A. Sklar, “Flow cytometry: Impact on early drug discovery,” *SLAS Discovery*, vol. 20, no. 6, pp. 689–707, Jul. 2015, doi: 10.1177/1087057115578273.

[4] M. Ding and B. S. Edwards, “High-throughput flow cytometry in drug discovery,” *SLAS Discovery*, vol. 23, no. 7, pp. 599–602, Aug. 2018, doi: 10.1177/2472555218778250.

- [5] T. Nguyen, V. A. Chidambara, S. Z. Andreasen, M. Golabi, V. N. Huynh, Q. T. Linh, D. D. Bang, and A. Wolff, "Point-of-care devices for pathogen detections: The three most important factors to realise towards commercialization," *TrAC Trends Anal. Chem.*, vol. 131, Oct. 2020, Art. no. 116004.
- [6] A. K. Pomerantz, F. Sari-Sarraf, K. J. Grove, L. Pedro, P. J. Rudewicz, J. W. Fathman, and T. Krucker, "Enabling drug discovery and development through single-cell imaging," *Exp. Opinion Drug Discovery*, vol. 14, no. 2, pp. 115–125, Feb. 2019, doi: [10.1080/17460441.2019.1559147](https://doi.org/10.1080/17460441.2019.1559147).
- [7] J. Xu, D. Zhou, D. Deng, J. Li, C. Chen, X. Liao, G. Chen, and P. A. Heng, "Deep learning in cell image analysis," *Intell. Comput.*, vol. 2022, pp. 1–15, Sep. 2022, doi: [10.34133/2022/9861263](https://doi.org/10.34133/2022/9861263).
- [8] M. Doan, I. Vorobjev, P. Rees, A. Filby, O. Wolkenhauer, A. E. Goldfeld, J. Lieberman, N. Barteleva, A. E. Carpenter, and H. Hennig, "Diagnostic potential of imaging flow cytometry," *Trends Biotechnol.*, vol. 36, no. 7, pp. 649–652, Jul. 2018, doi: [10.1016/j.tibtech.2017.12.008](https://doi.org/10.1016/j.tibtech.2017.12.008).
- [9] Z. Liu, W. Guo, D. Zhang, Y. Pang, J. Shi, S. Wan, K. Cheng, J. Wang, and S. Cheng, "Circulating tumor cell detection in hepatocellular carcinoma based on karyoplasmic ratios using imaging flow cytometry," *Sci. Rep.*, vol. 6, no. 1, p. 39808, Dec. 2016, doi: [10.1038/srep39808](https://doi.org/10.1038/srep39808).
- [10] A. Lopresti, F. Malergue, F. Bertucci, M. L. Liberatoscioli, S. Garnier, Q. DaCosta, P. Finetti, M. Gilabert, J. L. Raoul, D. Birnbaum, C. Acquaviva, and E. Mamessier, "Sensitive and easy screening for circulating tumor cells by flow cytometry," *JCI Insight*, vol. 4, no. 14, Jul. 2019, Art. no. e128180, doi: [10.1172/jci.insight.128180](https://doi.org/10.1172/jci.insight.128180).
- [11] E. Shechtman, Y. Caspi, and M. Irani, "Increasing space-time resolution in video," in *Computer Vision—ECCV (Lecture Notes in Computer Science)*, vol. 2350, A. Heyden, G. Sparr, M. Nielsen, and P. Johansen, Eds., Berlin, Germany: Springer, 2002, pp. 753–768, doi: [10.1007/3-540-47969-4_50](https://doi.org/10.1007/3-540-47969-4_50).
- [12] D. Bradley, B. Atcheson, I. Ihrke, and W. Heidrich, "Synchronization and rolling shutter compensation for consumer video camera arrays," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. Workshops*, Miami, FL, USA, Jun. 2009, pp. 1–8, doi: [10.1109/CVPRW.2009.5204340](https://doi.org/10.1109/CVPRW.2009.5204340).
- [13] D. V. Voronin, A. A. Kozlova, R. A. Verkhovskii, A. V. Ermakov, M. A. Makarkin, O. A. Inozemtseva, and D. N. Bratashov, "Detection of rare objects by flow cytometry: Imaging, cell sorting, and deep learning approaches," *Int. J. Mol. Sci.*, vol. 21, no. 7, p. 2323, Mar. 2020, doi: [10.3390/ijms21072323](https://doi.org/10.3390/ijms21072323).
- [14] W. E. Ortyen, B. E. Hall, T. C. George, K. Frost, D. A. Basiji, D. J. Perry, C. A. Zimmerman, D. Coder, and P. J. Morrissey, "Sensitivity measurement and compensation in spectral imaging," *Cytometry A*, vol. 69, no. 8, pp. 852–862, Aug. 2006, doi: [10.1002/cyto.a.20306](https://doi.org/10.1002/cyto.a.20306).
- [15] E. K. Zuba-Surma, M. Kucia, A. Abdel-Latif, J. W. Lillard, and M. Z. Ratajczak, "The ImageStream System: A key step to a new era in imaging," *Folia Histochemica et Cytobiologica*, vol. 45, no. 4, p. 13, 2007.
- [16] A. S. Rane, J. Rutkauskaitė, A. de Mello, and S. Stavrakis, "High-throughput multi-parametric imaging flow cytometry," *Chem. vol.* 3, no. 4, pp. 588–602, Oct. 2017, doi: [10.1016/j.chempr.2017.08.005](https://doi.org/10.1016/j.chempr.2017.08.005).
- [17] E. Schonbrun, S. S. Gorthi, and D. Schaak, "Microfabricated multiple field of view imaging flow cytometry," *Lab Chip*, vol. 12, no. 2, pp. 268–273, 2012, doi: [10.1039/c1lc20843h](https://doi.org/10.1039/c1lc20843h).
- [18] A. K. S. Lau, A. H. L. Tang, J. Xu, X. Wei, K. K. Y. Wong, and K. K. M. Tsia, "Optical time stretch for high-speed and high-throughput imaging—From single-cell to tissue-wide scales," *IEEE J. Sel. Topics Quantum Electron.*, vol. 22, no. 4, pp. 89–103, Jul. 2016, doi: [10.1109/JSTQE.2015.2512978](https://doi.org/10.1109/JSTQE.2015.2512978).
- [19] B. Guo, C. Lei, H. Kobayashi, T. Ito, Y. Yalikun, Y. Jiang, Y. Tanaka, Y. Ozeki, and K. Goda, "High-throughput, label-free, single-cell, microalgal lipid screening by machine-learning-equipped optofluidic time-stretch quantitative phase microscopy," *Cytometry A*, vol. 91, no. 5, pp. 494–502, May 2017, doi: [10.1002/cyto.a.23084](https://doi.org/10.1002/cyto.a.23084).
- [20] K. Goda, A. Ayazi, D. R. Gossett, J. Sadasivam, C. K. Lonappan, E. Sollier, A. M. Fard, S. C. Hur, J. Adam, C. Murray, C. Wang, N. Brackbill, D. Di Carlo, and B. Jalali, "High-throughput single-microparticle imaging flow analyzer," *Proc. Nat. Acad. Sci. USA*, vol. 109, no. 29, pp. 11630–11635, Jul. 2012, doi: [10.1073/pnas.1204718109](https://doi.org/10.1073/pnas.1204718109).
- [21] H. Mikami, M. Kawaguchi, C. J. Huang, H. Matsumura, T. Sugimura, K. Huang, C. Lei, S. Ueno, T. Miura, T. Ito, and K. Nagasawa, "Virtual-freezing fluorescence imaging flow cytometry," *Nature Commun.*, vol. 11, no. 1, p. 1162, Mar. 2020, doi: [10.1038/s41467-020-14929-2](https://doi.org/10.1038/s41467-020-14929-2).
- [22] Y. Wang, Z. Huang, X. Wang, F. Yang, X. Yao, T. Pan, B. Li, and J. Chu, "Real-time fluorescence imaging flow cytometry enabled by motion deblurring and deep learning algorithms," *Lab Chip*, vol. 23, no. 16, pp. 3615–3627, 2023.
- [23] R. K. R. Baskaran, A. Link, B. Porr, and T. Franke, "Classification of chemically modified red blood cells in microflow using machine learning video analysis," *Soft Matter*, vol. 20, no. 5, pp. 952–958, 2024.
- [24] V. Ganjalizadeh, G. G. Meena, M. A. Stott, A. R. Hawkins, and H. Schmidt, "Machine learning at the edge for AI-enabled multiplexed pathogen detection," *Sci. Rep.*, vol. 13, no. 1, p. 4744, Mar. 2023.
- [25] Z. Göröcs, D. Baum, F. Song, K. de Haan, H. Ceylan Koydemir, Y. Qiu, Z. Cai, T. Skandakumar, S. Peterman, M. Tamamitsu, and A. Ozcan, "Label-free detection of giardia lamblia cysts using a deep learning-enabled portable imaging flow cytometer," *Lab Chip*, vol. 20, no. 23, pp. 4404–4412, Nov. 2020.
- [26] R. Wang, X. Huang, X. Xu, J. Sun, S. Zheng, X. Ke, J. Yao, W. Han, M. Wei, J. Chen, H. Gao, J. Guo, and L. Sun, "A standalone and portable microfluidic imaging detection system with embedded computing for point-of-care diagnostics," *IEEE Sensors J.*, vol. 22, no. 6, pp. 6116–6123, Mar. 2022.
- [27] S.-Y. Chiu, C.-C. Chiu, and S. S.-D. Xu, "A background subtraction algorithm in complex environments based on category entropy analysis," *Appl. Sci.*, vol. 8, no. 6, p. 885, May 2018, doi: [10.3390/app8060885](https://doi.org/10.3390/app8060885).
- [28] N. G. El-Gamal, H. E. Moustafa, and F. E. Z. Abou-Chadi, "A new combination method for background subtraction in video sequences," in *Proc. 8th Int. Conf. Inform. Syst. (INFOS)*, May 2012, pp. 1–25.
- [29] A. Lashkaripour, C. Rodriguez, L. Ortiz, and D. Densmore, "Performance tuning of microfluidic flow-focusing droplet generators," *Lab Chip*, vol. 19, no. 6, pp. 1041–1053, 2019.
- [30] K. Parnamets, A. Udai, A. Koel, T. Pardy, N. Gyimah, and T. Rang, "Compact empirical model for droplet generation in a lab-on-chip cytometry system," *IEEE Access*, vol. 10, pp. 127708–127717, 2022, doi: [10.1109/ACCESS.2022.3226623](https://doi.org/10.1109/ACCESS.2022.3226623).
- [31] *DiY 12x40 mm 5.6 mm To18 Laser Diode Housing Case W/405 nm 450 nm 650 nm 780 nm 980 nm Collimating Lens*. Accessed: Jan. 15, 2023. [Online]. Available: http://www.aliexpress.com/item/1005003509604067.html?src=ibdm_d03p0558e02r02&sk=&aff_platform=&aff_trace_key=&af=&cv=&cn=&dp=
- [32] Z. Zivkovic, "Improved adaptive Gaussian mixture model for background subtraction," in *Proc. 17th Int. Conf. Pattern Recognit.*, Cambridge, U.K., 2004, pp. 28–31, doi: [10.1109/ICPR.2004.1333992](https://doi.org/10.1109/ICPR.2004.1333992).
- [33] I. Iszaidy, R. Ngadiran, R. B. Ahmad, N. Ramli, M. I. Jais, and V. Vijayaravesswari, "An analysis of background subtraction on embedded platform based on synthetic dataset," *J. Phys., Conf.*, vol. 1755, no. 1, Feb. 2021, Art. no. 012042, doi: [10.1088/1742-6596/1755/1/012042](https://doi.org/10.1088/1742-6596/1755/1/012042).
- [34] Y. Xu, H. Ji, and W. Zhang, "Coarse-to-fine sample-based background subtraction for moving object detection," *Optik*, vol. 207, Apr. 2020, Art. no. 164195, doi: [10.1016/j.jlco.2020.164195](https://doi.org/10.1016/j.jlco.2020.164195).
- [35] Y. Zhao, X. Zhou, X. Xu, Z. Jiang, F. Cheng, J. Tang, and Y. Shen, "A novel vehicle tracking ID switches algorithm for driving recording sensors," *Sensors*, vol. 20, no. 13, p. 3638, Jun. 2020, doi: [10.3390/s20133638](https://doi.org/10.3390/s20133638).
- [36] *NVIDIA Jetson NANO Developer Kit*.
- [37] F. Hollfelder, "Microfluidics-enabled fluorescence-activated cell sorting of single pathogen-specific antibody secreting cells for the rapid discovery of monoclonal antibodies," to be published.
- [38] M. Durve, A. Tiribocchi, F. Bonaccorso, A. Montessori, M. Lauricella, M. Bogdan, J. Guzowski, and S. Succi, "DropTrack—Automatic droplet tracking with YOLOv5 and DeepSORT for microfluidic applications," *Phys. Fluids*, vol. 34, no. 8, Aug. 2022, Art. no. 082003.



FARIHA AFRIN (Member, IEEE) received the B.Sc. degree in electrical and electronics engineering from International Islamic University Chittagong, Bangladesh, in 2013, the M.Sc. degree in electrical and electronics engineering from American International University-Bangladesh, in 2017, and the master's degree in communicative electronics from Tallinn University of Technology, Estonia, in 2020, where she is currently pursuing the Ph.D. degree with the Thomas Johann Seebeck

Department of Electronics. Her research interests include fluorescence and morphology detection in flow cytometry.



KAISER PÄRNAMETS (Member, IEEE) was born in Tallinn, Estonia, in January 1989. He received the B.Sc. and M.Sc. degrees in engineering from Tallinn University of Technology, in 2013 and 2016, respectively, where he is currently pursuing the Ph.D. degree. From 2014 to 2017, he was a Hardware Engineer at the electronics industry. Since 2018, he has been an Electronics Lecturer for bachelor's and master's students. His research interests include lab-on-a-chip and microfluidics.



YANNICK LE MOULLEC (Senior Member, IEEE) received the M.Sc. degree from Université de Rennes I, France, in 1999, and the Ph.D. and HDR (accreditation to supervise research) degrees from Université de Bretagne Sud, France, in 2003 and 2016, respectively. From 2003 to 2013, he was a Postdoctoral Researcher, an Assistant Professor, and an Associate Professor with the Department of Electronic Systems, Aalborg University, Denmark. He joined the Thomas Johann Seebeck Department of Electronics, Tallinn University of Technology, Estonia, first as a Senior Researcher, from 2013 to 2016, and a Professor, since 2017. He has supervised or co-supervised 60 M.Sc. theses and 15 Ph.D. theses. He has been involved in more than 20 projects, including five as a PI, a co-PI, or a co-main applicant. His research interests include embedded systems and wireless systems, including the IoT and the application thereof. He is a member of the IEEE Sustainable ICT Technical Community and the IEEE Circuits and Systems Society.



ANDRES UDAL (Member, IEEE) received the Ph.D. degree in electronical engineering from Tallinn University of Technology, Tallinn, Estonia, in 1999. From 1990 to 1995, he was acting as a Visiting Researcher and a TCAD Software Designer at Silvaco Int., Santa Clara, CA, USA; Darmstadt University of Technology, Germany; and Uppsala University, Sweden. He is currently a Senior Research Fellow with the Department of Software Science, Laboratory for Proactive Technologies, Tallinn University of Technology.



ANTS KOEL (Member, IEEE) was born in Tallinn, Estonia, in August 1962. He received the Engineering Diploma degree in industrial electronics from Tallinn Polytechnic Institute, Tallinn, Estonia, in 1985, and the master's and Ph.D. degrees from Tallinn University of Technology, in 1998 and 2014, respectively. His research interests include semiconductor electronics and devices, TCAD simulations, and circuit design. He was the Chairperson of the Steering Committee of the IEEE-Sponsored Baltic Electronics Conference, in 2020 and 2022, participated in ERA Chair projects COEL and SGSOLAR, led electronics in-kind contribution projects of Estonia in ESS.



TAMAS PARDY (Member, IEEE) received the M.Sc. degree in info-bionics engineering from Pázmány Péter Catholic University, Budapest, Hungary, in 2014, and the Ph.D. degree in electronics and telecommunication from Tallinn University of Technology, Tallinn, Estonia, in 2018. He is currently a Senior Researcher with Tallinn University of Technology. He has supervised one Ph.D. thesis and three M.Sc. theses and has authored or co-authored 20 scientific publications. His research interests include flow- and temperature-control of lab-on-a-chip devices.



TOOMAS RANG (Senior Member, IEEE) received the Ph.D. degree in semiconductor electronics from the Hungarian Academy of Sciences, in 1981. He is currently a Professor Emeritus with the Thomas Johann Seebeck Department of Electronics and a Senior Research Fellow with the Department of Chemistry and Biotechnologies, Tallinn University of Technology, Estonia. He has supervised 15 Ph.D. students and has fulfilled the PI position in several European and domestic research and industrial projects. His current research interest includes applied microfluidics approaches for lab-on-chip applications. He is the Initiator and today the honor Chairperson of the Baltic Electronics Conference (BEC) Series, in 1987, supported by IEEE, since 1996.

...

Appendix 2

II

F. Afrin, Y. Le Moullec, T. Pardy, "Microfluidic Droplet Classification through Tuned Convolutional Neural Network on a Resource Constrained Platform", 2024 19th Biennial Baltic Electronics Conference (BEC2024), 2-4 October 2024, Tallinn, Estonia, DOI: <https://doi.org/10.1109/BEC61458.2024.10737958>³

³Note: The authors provided the following text after publication of the paper: Acknowledgement: "The authors thank Assistant Prof. Tomasz Kaminski, University of Warsaw, Poland, for providing us the droplet image dataset for training the droplet classification model."

Microfluidic Droplet Classification through Tuned Convolutional Neural Network on a Resource Constrained Platform

Fariha Afrin

Thomas Johann Seebeck Department of
Electronics
Tallinn University of Technology
Tallinn, Estonia
fariha.afrin@taltech.ee

Yannick Le Moullec

Thomas Johann Seebeck Department of
Electronics
Tallinn University of Technology
Tallinn, Estonia
yannick.lemoullec@taltech.ee

Tamas Pardy

Thomas Johann Seebeck Department of
Electronics
Tallinn University of Technology
Tallinn, Estonia
tamas.pardy@taltech.ee

Abstract—Microfluidic droplet classification is an essential research step in imaging flow cytometry. Although deep learning algorithms have proven effective for detecting and classifying microfluidic droplets in benchtop setups, their implementation in portable devices presents a significant challenge. The computational complexity of deep learning needed to meet the speed and accuracy requirements of imaging flow cytometer often exceed the capabilities of portable resource-constrained devices, making it difficult to transition from benchtop setups to field-deployable devices. To address this challenge, we present a tuned YoloV4-tiny model mapped onto a Raspberry Pi-5 single board computer. The neural network is trained on 878 images from our own dataset consisting of 975 images from two videos recorded on our microfluidic experimental setup. The performance is measured in terms of inference time and mean average precision. Our implementation is able to classify three different types of droplets within 13 ms while maintaining an accuracy rate of more than 98% accuracy.

Keywords—Droplet, droplet classification, neural networks, resource constrained platform, single board computer

I. INTRODUCTION

Microfluidic droplet classification can bring phenomenal insight to biomedical research and related applications [1], [2]. Recent advances in deep learning have paved the way for real time automated, reliable, and accurate object classification. For example, convolutional neural networks (CNNs), a class of deep learning models, have led to advancements specifically for object detection and classification in applications ranging from biomedical imaging [3], [4] to autonomous driving [5]. These models have the potential to automate and advance the droplet classification stage in imaging flow cytometry (IFC) since they are capable of learning complex features and patterns from large datasets. However, existing methods for droplet classification are developed for benchtop (lab grade) applications which use e.g. powerful Graphics Processing Unit (GPU).

In the study [6], researchers developed a CNN-based algorithm named Weakly Supervised Cell Counting Network (WSCNet) to classify cell-encapsulated droplets. The proposed method significantly improves on traditional image classification accuracy (reaching more than 89%) and exhibits robustness under different lighting conditions; yet it relies on a desktop computer for operation. Another research effort aims to enhance the accuracy and efficiency of detecting

microfluidic droplet content in liquid biopsy workflows through CNN [7]. The automatic classification system based on CNN achieves 96% accuracy in classifying droplets.

Transfer learning with deep CNNs is used for classifying cellular morphological changes achieving high accuracy between 95% and 97% [8]. This approach reduces the need for extensive dataset collection and labeling, which is a significant bottleneck in developing deep learning models for specific microfluidic applications. An optimized image activated cell sorter based on deep learning model under the TensorRT29 framework was deployed on NVIDIA GeForce GTX 1080 TI GPU [9]. The system is specifically tailored to classify and sort polystyrene beads and cells in real time.

Although there have been significant advancements in the microfluidic object classification field, there remains a gap in the literature regarding the optimization of models for implementing them on portable, resource-constrained devices. Most of the existing research prioritize achieving high classification accuracy at the expense of high computational complexity, rendering these approaches largely unsuitable for portable systems.

Our study addresses this gap by optimizing a deep learning-based droplet classification system for portable devices; our results show the feasibility of classification model on resource-constrained devices. We demonstrate that our deep learning-based classifier can achieve high accuracy and speed, enabling analysis of microfluidic droplets containing cells. Furthermore, we tested our system using a new, unseen testing dataset, highlighting its robustness. By implementing deep learning techniques into the resource constrained platform, we aim to facilitate resource-efficient droplet classification, paving the way for advancements in portable microfluidic applications.

The main steps and contributions of this research are listed as follows.

1) We have generated a new dataset of 975 microfluidic droplets images. The dataset is created from two videos recorded on our microfluidic experimental setup.

2) We have developed a droplet classification method (building upon YoloV4-tiny model) with high accuracy and inference speed, yet lightweight for resource-constrained implementation. High accuracy is achieved via data-augmentation and high inference speed is achieved by

reducing the number of filters in the convolutional layers by -20% and via batch processing (6 images per batch).

3) We have implemented and deployed our proposed method on two resource-constrained devices, i.e. single board computers (SBC), i.e., first on Raspberry Pi-4 and second on Raspberry Pi-5 (achieving high accuracy (98%) and inference speed (13 ms)), which shows the feasibility of our proposed solution on resource-constrained devices.

The rest of this paper is structured as follows: Section II provides an overview of the materials and method, including the implementation of the model onto the resource-constrained devices. Section III presents the results, including the performance comparative analysis. Finally, Section IV concludes the paper and outlines some potential future directions for research in this area.

II. MATERIALS AND METHOD

A. Dataset

Training data acquisition is a significant bottleneck in developing microfluidic object detection and classification due to the lack of data availability. The microfluidic experimental video stream is recorded using a camera placed directly above the microfluidic chip where droplets are formed at a T-junction of two channels. Two different experimental videos with different conditions were used to extract the images (example images are shown in Fig. 1) for model training. Each image contains one or two microfluidic droplets. Observing the droplet activity, images are extracted from the videos at 15 frames per second (fps) instead of the default 1 fps and saved in the JPG format.

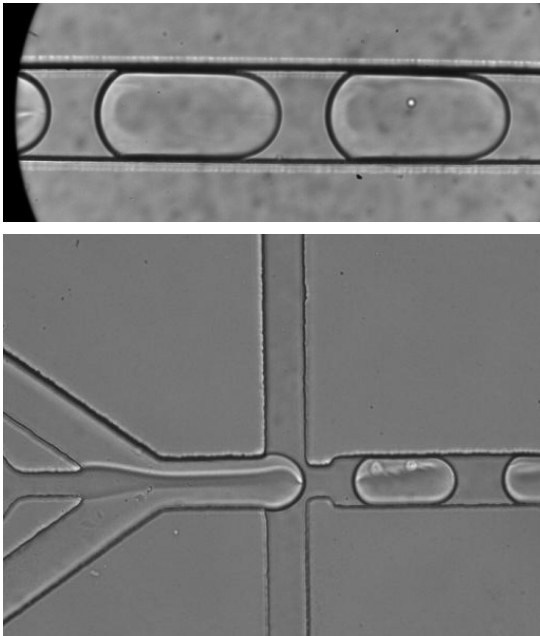


Fig. 1. Picture of droplets flowing inside the microfluidic channel. Top: an empty droplet (left) and droplet containing a single cell (right). Bottom: a droplet containing multiple cells.

CNN performance can deteriorate if the amount of dataset is not diverse enough. To achieve higher performance (accuracy and robustness), we have applied a data

augmentation technique to increase the number of different images and reduce the over-fitting problem. In the data augmentation method, the number of different images is increased by changing the contrast from 0.4 to 1.6, applying gaussian blur from 0.3 to 0.9 and from 1.1 to 2.5, as well as applying vertical and horizontal rotation. Different contrast and noise were applied to create a model robust in varying experimental conditions. This yielded a total of 975 images, from which 878 images were used for training and validation, and 97 images were used for testing.

B. Image Annotation

Image annotation consists in identifying and noting down the location of all objects of interest (in this case droplets). We used 878 images (including augmented images) as the training dataset and approximately 10% of these (87) images were used as validation dataset. To annotate all the images based on a rectangle bounding box, we used the “makesense.ai” online tool and manually labelled all the images. YoloV4-tiny bounding box annotation format is $[x,y,w,h]$, where (x,y) is the centroid, w is the width, and h is the height of the bounding box.

C. CNN and Implementation

Given the needs formulated by a bioanalytical expert, we decided to classify the images into three classes: (i) empty droplet, (ii) droplet with a single cell, and (iii) droplet with multiple cells. Only droplets which are fully visible in an image are considered for classification. We took CNN model YoloV4-tiny as a basis for our method. Then we tuned it by reducing the number of filters in the convolutional layers to improve the inference speed without significant reduction of accuracy. Compressed filters could reduce the model size, hence improving inference speed.

We trained the model using the ‘Darknet’ framework in Google Colaboratory Pro (T4 GPU was configured with Python 3.10.12, CUDA12.2, cuDNN 8.9.6, OpenCV 4.8.0). The trained model takes an image as input and estimates a confidence score for each detected droplet within the image. The performance was measured by calculating the accuracy, mean average precision and loss in validation dataset. The loss for the trained model for 6000 iterations is shown in Fig. 2; the values do not measure the model’s actual performance yet. The best performance was obtained at 6000 iterations, where the loss is 0.09.

For training, we set the final learning rate to be 0.00261, decay 0.0005, momentum 0.9, batch size 64, subdivisions of 8, max batches 6000, according to default configuration, and the filters are reduced by 20% in all convolutional layers. The tuned value of 20% was found empirically, as a trade-off between complexity and accuracy. This reduced the weight file size by -36.16% (from 22.4 MB to 14.3 MB). After training completion, the model was deployed on two SBCs, i.e. Raspberry Pi-4 and Raspberry Pi-5. Inference time measured in RPI-5 was reduced by -35% (from 20 ms to 13 ms). The results are obtained using 6 batches of images instead of a single image. We have tested the inference speed for 1 to 16 batches of images. It was observed that the processing time was improved until 6 batches of images, but not for larger batches. Then all the values are averaged, and the inference time is calculated for a single input.

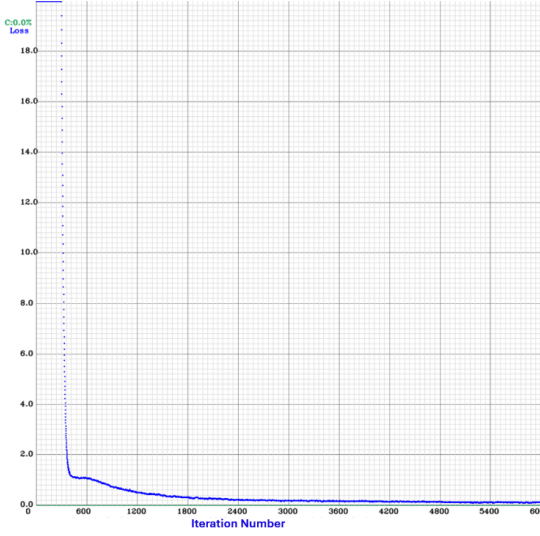


Fig. 2. YoloV4-tiny with compressed filters training results up to 6000 iterations, where the average loss reaches 0.09.

III. IMPLEMENTATION RESULTS

In this section, we present the results of our proposed deep learning-based microfluidic droplet classification system, specifically designed for portable device implementation. Our evaluation focuses on three main aspects: classification accuracy, processing speed, and resource utilization on portable, resource-constrained hardware.

A. Droplet Classification

Our study assessed droplet classification performance with 416x416 pixel input size and non-maximum-suppression rate at 0.7 test datasets. Fig. 3 shows examples of droplet detection results on original images from our test dataset.

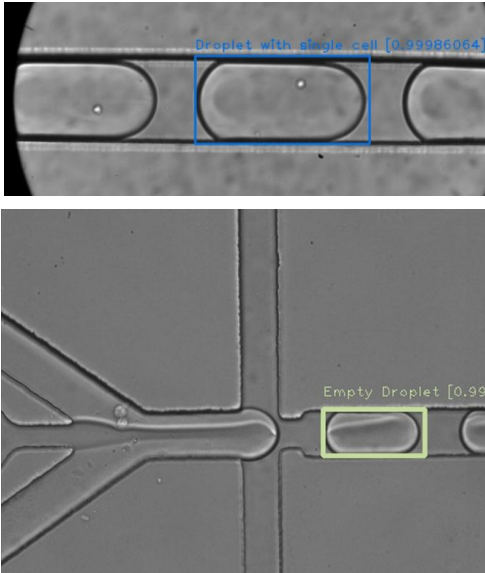


Fig. 3. Examples of microfluidic droplet classification results after 6000 training iterations, where the prediction probability is above 99%.

B. Model Performance

Performance metrics are shown in Table I and calculated using Equations 1 and 2, where TP is true positives, FP is false positives, and FN is false negatives.

$$\text{Precision} = \frac{TP}{TP+FP} \quad (1)$$

$$\text{Recall} = \frac{TP}{TP+FN} \quad (2)$$

TABLE I. CLASSIFICATION PERFORMANCE METRICS (TEST DATASET = 97 IMAGES)

Metric	Performance value
TP	112
FP	6
FN	0
Precision	0.95
Recall	1
mAP@0.50	99.95 %

The high precision (0.95) and recall (1) indicate that the model is highly reliable and effective at correctly identifying different droplet types with minimal false classification. The mAP@0.50 is the mean average precision, where AP is calculated using an intersection over union (IoU) threshold of 0.5. IoU indicates the overlap of the predicted and the ground truth bounding boxes. The high mean of average precision (mAP) at 99.95% highlights our model's robustness.

Our portable implementation was tested both on Raspberry Pi-4 and Raspberry Pi-5 SBCs (both with 8 Gb RAM and running on a 64-bit operating system). One of the key requirements for real-time applications is the ability to process images at high speed. To evaluate the processing speed, we executed tests for the model using corresponding test dataset on targeted devices. By running the trained model, we observed that the processing time is 20 times faster (13 ms) with the Raspberry Pi-5 than with the Raspberry Pi-4 (265 ms).

On Raspberry Pi-5, our system achieved a classification inference time of 13 ms, which yields 76 fps. This rate was obtained using a batch of 6 images and is sufficient for practical applications.

Resource utilization was a critical aspect of our evaluation, given the constraints of portable devices. We monitored CPU and memory usage during the classification tasks to ensure the system's efficiency. The results are presented in Table II. As can be seen in the table, the average CPU usage and memory usage were 70% and 60%, respectively. These numbers indicate that the proposed method is sufficiently lightweight and suitable for portable application as CPU and memory usage are within the operational limits.

TABLE II. RESOURCE UTILIZATION ON RASPBERRY PI-5

Resource	Average Usage
CPU Usage	70%
Memory Usage	60%

C. Comparison with similar studies

In this section, we compare our results with the closest-related existing research, as shown in Table III. Note that i) existing works do not target the same applications as our, and ii) most existing studies' implementation are on powerful desktop PCs and/or GPUs.

TABLE III. CLASSIFICATION PERFORMANCE METRICS

Reference	Method	Platform	Object	mAP (%)	Processing time
[10]	YoloV5	Intel Core i7-12650H	Droplet	99.3@0.5	74.3 ms
[11]	Mask-RCNN	Nvidia RTX 2080 Ti GPU	Cells	73	-
[12]	CNN	NVIDIA Tesla K80 GPU	Cells	95.7	2.2ms
[13]	CNN	-	Cells	93~99	-
[14]	CNN	Nvidia Tesla K40c GPU	Cells	99	-
This study	Tuned YoloV4-tiny	Raspberry Pi-5	Droplet	99.95@0.5	13 ms

Although a direct comparison is not possible, some insights can be derived from the table. In particular, it can be noted that our tuned implementation with YoloV4-tiny on Raspberry Pi-5 achieves better processing speed (13 ms) than that of [10] with YoloV5 on an Intel Core i7 (74.3 ms for single class), while mAP accuracy values are also comparable (99.95%@0.5 vs. 99.3%@0.5). High-end GPUs like the Nvidia Tesla K80 [12] and Nvidia Tesla K40c GPU in [14] can achieve faster processing speed, and higher mAP, but they are developed for different, non-portable applications. Overall, the results in Table III shows that the tuned YoloV4-tiny on a Raspberry Pi-5 has significant accuracy and reliability on a resource-constrained platform, making it suitable for field-deployable devices.

IV. CONCLUSION

In this paper, we presented microfluidic image-based droplet classification implemented on a resource constrained platform using tuned YoloV4-tiny model. Data augmentation results in high accuracy and robustness; optimizing the model by reducing the number of filters in the convolutional layers by -20% and applying batch processing (6 images per batch) yields high processing speed. 878 images from our own droplet dataset consisting of 975 images were used to train the model. The solution demonstrates the capability to classify the droplets within 13 ms with higher accuracy of more than 98% when deployed on the Raspberry Pi-5 SBC.

For the sake of completeness, we note the following:

1) As mentioned in Section III, we also tested the model on a Raspberry Pi-4 SBC. However, the inference time was much higher (265 ms) while the other performance indicators were not degraded. However, the Raspberry Pi-5 has notable higher power requirements (5V, 5A) compared to Raspberry Pi-4 (5V, 3A). If power (e.g. heat constraints) and/or energy (e.g. battery-powered) are stringent application constraints, one might trade-

off the higher speed of Raspberry Pi-5 for the lower consumption of the Raspberry Pi-4.

2) We also converted the model to TensorFlow and subsequently to TensorFlow Lite (TFLite). However, the resulting inference time for a single image was more than 1 s. This indicates that the proposed model maintains high accuracy yet the performance in terms of inference speed is limited to TFLite framework and Raspberry Pi-4.

In future work, the processing time could be further reduced by applying optimization techniques such as pruning and quantization to achieve real-time or near-real time classification speeds; the trade-off between this speed gain, power consumption and the resulting decrease in accuracy should be assessed carefully.

ACKNOWLEDGMENT

This work was supported by the Estonian Science Agency ETAg grant number PRG620. Additionally, the authors acknowledge support from the Estonian Education and Youth Board project AIoT*5G grant number ÕÜF11 and Estonian IT Academy project SAIoT Sustainable Artificial Internet of Things.

REFERENCES

- [1] T. N. D. Trinh, H. D. K. Do, N. N. Nam, T. T. Dan, K. T. L. Trinh, and N. Y. Lee, "Droplet-Based Microfluidics: Applications in Pharmaceuticals," 2023.
- [2] Z. Chen, S. Kheiri, E. W. K. Young, and E. Kumacheva, "Trends in Droplet Microfluidics: From Droplet Generation to Biomedical Applications," 2022.
- [3] Z. Liu, "A survey on applications of deep learning in microscopy image analysis," *Computers in Biology and Medicine*, 2021.
- [4] G. P. Rutkowski, I. Azizov, E. Unmann, M. Dudek, and B. A. Grimes, "Microfluidic droplet detection via region-based and single-pass convolutional neural networks with comparison to conventional image analysis methodologies," 2022.
- [5] G. Doğan, "A new CNN-based semantic object segmentation for autonomous vehicles in urban traffic scenes," *International Journal of Multimedia Information Retrieval*, 2024.
- [6] X. Zhou, Y. Mao, M. Gu, and Z. Cheng, "WSCNet: Biomedical Image Recognition for Cell Encapsulated Microfluidic Droplets "2279," 2023.
- [7] G. Soldati *et al.*, "Microfluidic droplets content classification and analysis through convolutional neural networks in a liquid biopsy workflow," 2018.
- [8] A. Kensert, P. J. Harrison, and O. Spiuth, "Transfer Learning with Deep Convolutional Neural Networks for Classifying Cellular Morphological Changes," 2019.
- [9] K. Lee, S.-E. Kim, J. Doh, K. Kim, and W. K. Chung, "User-friendly image-activated microfluidic cell sorting technique using an optimized, fast deep learning algorithm," *Lab on a Chip*, 2021.
- [10] J. Xu, W. Fan, J. Madsen, G. P. Tanev, and L. Pezzarossa, "AI-Based Detection of Droplets and Bubbles in Digital Microfluidic Biochips," 2023.
- [11] M. Ghafari *et al.*, "A Comparison of YOLO and Mask-RCNN for Detecting Cells from Microfluidic Images," 2018.
- [12] Y. Li, A. Mahjoubfar, C. L. Chen, K. R. Niazi, L. Pei, and B. Jalali, "Deep Cytometry: Deep learning with Real-time Inference in Cell Sorting and Flow Cytometry," Aug. 2019. doi: 10.1038/s41598-019-47193-6.
- [13] Y. Suzuki *et al.*, "Label-free chemical imaging flow cytometry by high-speed multicolor stimulated Raman scattering," 2019.
- [14] N. Meng, E. Y. Lam, K. K. Tsia, and H. K.-H. So, "Large-Scale Multi-Class Image-Based Cell Classification With Deep Learning," *IEEE JOURNAL OF BIOMEDICAL AND HEALTH INFORMATICS*, vol. 23, no. 5, 2019.

Appendix 3

III

F. Afrin, Y. Le Moullec, T. Pardy, T. Rang, "Lightweight CNN-based Microfluidic Droplet Classification for Portable Imaging Flow Cytometry". Proceedings of the Estonian Academy of Science, 74, 2S, 302–311, 2025. DOI: <https://doi.org/10.3176/proc.2025.2S.05>



Proceedings of the
Estonian Academy of Sciences
2025, 74, 2S, 302–311

<https://doi.org/10.3176/proc.2025.2S.05>

www.eap.ee/proceedings
Estonian Academy Publishers

EMBEDDED IMAGING FLOW CYTOMETRY

RESEARCH ARTICLE

Received 16 December 2024
Accepted 2 April 2025
Available online 9 June 2025

Keywords:

droplet, droplet classification,
neural networks, resource constrained
platform, single board computer

Corresponding author:

Fariha Afrin
fariha.afrin@taltech.ee

Citation:

Afrin, A., Le Moullec, Y., Pardy, T. and
Rang, T. 2025. Lightweight CNN-based
microfluidic droplet classification for
portable imaging flow cytometry.
*Proceedings of the Estonian Academy of
Sciences*, 74(2S), 302–311.
<https://doi.org/10.3176/proc.2025.2S.05>

© 2025 Authors. This is an open
access article distributed under the
terms and conditions of the Creative
Commons Attribution (CC BY) license
(<http://creativecommons.org/licenses/by/4.0>).

Lightweight CNN-based microfluidic droplet classification for portable imaging flow cytometry

Fariha Afrin, Yannick Le Moullec, Tamas Pardy and
Toomas Rang

Thomas Johann Seebeck Department of Electronics, Tallinn University of Technology
(TalTech), Ehitajate tee 5, 19086 Tallinn, Estonia

ABSTRACT

Classifying microfluidic droplets is an essential step in imaging flow cytometry. While deep learning algorithms can detect and classify such droplets in benchtop laboratory settings, their deployment on portable devices remains challenging because the computational requirements often exceed the capabilities of compact, resource-limited devices. This hinders the transition from stationary lab setups to field-deployable instruments. To tackle this issue, we introduce a customized YoloV4-tiny model deployed on a Raspberry Pi-5 (RPi5) single-board computer. Our neural network is trained using 878 images from a custom dataset of 975 images, derived from two videos captured with real-life microfluidic experimental setup. We evaluate performance based on inference time and mean average precision. Our system successfully classifies three distinct droplet types (no cell, one cell, multiple cells) within 13 ms, achieving a 99.95% mean average precision at an intersection over union threshold of 0.5 (mAP@0.5). We also compare the classification performance metrics of our customized YoloV4-tiny model against seven other combinations of machine learning models and platforms, including a recent low-cost, highly compact edge device with tensor processing unit capabilities, specifically, the MaixCam board with LicheeRV Nano module (SOPHGO SG2002) running a YoloV5-s model. Compared to this proposed customized YoloV4-tiny on the RPi5, the YoloV5-s on MaixCam achieves a significantly shorter classification time (5.34 ms) owing to its onboard tensor processing unit but suffers from a lower mAP@0.5 of 55.09% due to quantization. Our work shows that carefully designed systems can achieve a balance between speed and accuracy, enabling robust performance even on resource-limited devices and paving the way for microfluidic droplet classification in portable imaging flow cytometry.

Introduction

Microfluidic droplet classification offers remarkable potential for biomedical research and applications [1,2]. In parallel, recent progress in machine learning (ML) and deep learning has enabled (near) real-time, automated, reliable, and precise object classification. In particular, convolutional neural networks (CNNs), a subset of deep learning models, have led to advances in object detection and classification in various fields, including biomedical imaging [3,4]. Given their ability to extract complex features and patterns from image datasets, CNN models show promise in automating and enhancing the droplet classification phase in imaging flow cytometry (IFC).

Nevertheless, current ML-based droplet classification methods are predominantly designed for benchtop (laboratory-grade) applications, typically relying on high-performance graphics processing units (GPUs). Implementing ML-based droplet classification on embedded/edge devices remains challenging due to: (i) resource constraints (limited computational power, memory, and energy compared to laboratory-grade systems), (ii) the need to customize, adapt, and optimize large models and complex algorithms to run efficiently on resource-constrained devices, (iii) requirements for possible (near) real-time processing and speed vs accuracy trade-off on low-power devices; and (iv) the potential availability and exploitation of specialized hardware, such as low-power tensor processing units (TPUs) and neural processing units (NPU). This paper is an extended version of [5], providing additional references, more details about model customization and deployment, as well as additional deployment comparisons.

In [6], researchers developed a CNN-based algorithm called the weakly supervised cell counting network (WSCNet) to classify cell-encapsulated droplets. The proposed method significantly improved the accuracy of traditional image

classification (approx. 89%) and exhibited robustness under different lighting conditions; however, WSCNet relied on a desktop computer for operation.

Another research effort aimed to improve the accuracy and efficiency of detecting microfluidic droplet contents in liquid biopsy workflows using CNN [7]. This CNN-based automatic classification system achieved 96% precision in droplet classification but was not designed for resource-constrained platforms.

Transfer learning with deep CNNs has been used to classify cellular morphological changes, achieving high accuracy between 95% and 97% [8]. Although this approach reduces the need for extensive dataset collection and labeling, which is a significant bottleneck in the development of deep learning models for specific microfluidic applications, it is not suitable for resource-constrained platforms.

In [9], researchers developed an optimized image-activated cell sorter based on a deep learning model specifically tailored to classify and sort polystyrene beads and cells in real time. However, it was deployed using the TensorRT29 framework on an NVIDIA GeForce GTX 1080 Ti GPU.

The authors of [10] developed a ML-based computer vision solution for real-time detection of droplets and bubbles. Their approach used the YoloV5 framework with custom pre- and post-processing techniques, trained on a dataset of 5115 images. Their results demonstrated real-time speed and high accuracy in detecting and differentiating between droplets and unwanted bubbles, but the method required a relatively high-end PC.

In [11], researchers explored the application of computer vision and deep learning techniques to automate the analysis of yeast cell replicative lifespans. They compared Yolo and Mask R-CNN in terms of their efficacy in detecting and analyzing yeast cells from microfluidic images. They found that Yolo demonstrated superior sensitivity in cell detection, while Mask R-CNN provided more detailed information on cell sizes. However, their work was implemented on a desktop GPU (RTX 2080 Ti).

The authors of [12] developed a deep learning pipeline for high-throughput, label-free cell classification, using CNNs to process raw measurement signals and enabling low-latency inference suitable for real-time cell sorting applications. While their method demonstrated over 95% accuracy in label-free classification of specific types of white blood cells and epithelial cancer cells, their system was deployed on a desktop GPU (Tesla K80).

In [13], researchers proposed a label-free chemical IFC that combines pulse pair-resolved wavelength-switchable Stokes laser, multicolor stimulated Raman scattering (SRS) microscopy, and a 3D acoustic focusing microfluidic chip, supported by deep learning algorithms. They achieved a throughput of approximately 140 cells/s; however, they did not indicate the platform used (we assume that it was a high-end PC).

The research presented in [14] applied CNNs for processing large-scale datasets of label-free cell images for high-throughput cell classification. The authors compared CNN performance against k-nearest neighbors (kNN) and support

vector machine (SVM) methods. Their CNN-based approach yielded over 99% accuracy in identifying multiple cell types based on label-free bright-field images. However, they used a desktop GPU (Tesla K40c).

The work presented in [15] introduced a rapid and label-free antimicrobial susceptibility testing method for colistin, combining deep learning with droplet microfluidics. The DropDeepL AST method used a deep learning-powered approach for sensitive detection of bacterial growth in droplets, achieving 100% categorical agreement with the reference broth microdilution method for colistin susceptibility profiles. However, the paper did not specify the type of platform on which the model was trained or deployed; we assume that it was not a resource-constrained platform.

A deep learning-augmented T-junction droplet generation system was presented in [16]. The study used finite element analysis to simulate droplet production and its dynamics, followed by ML algorithms to estimate droplet characteristics based on input parameters. This approach enabled preselecting designs with comparable microfluidic configurations within the studied range. Nevertheless, the specific platform used for training and deploying the model was not disclosed; however, our assumption is that it was not a resource-constrained platform.

Finally, [17] developed a droplet-based microfluidic platform to detect peptides that are self-secreted by yeast. They used ML-based image processing techniques to analyze fluorescence emitted by single yeast cells in droplets, yielding high-throughput analysis and characterization of agonistic peptides. However, the paper did not specify the training or deployment platform; we also assume it was not a resource-constrained platform.

Despite the notable progress in microfluidic droplet object analysis, a research gap remains in the scientific literature related to model optimization for portable IFC devices. Many research efforts have focused on maximizing classification accuracy, often at the cost of increased computational demands, making them impractical for portable platforms. Our research contributes to bridging this gap by refining a deep learning-based droplet classification system for use in portable devices, demonstrating its viability even with resource limitations.

To address this issue, we have built a portable droplet classification system that leverages the YoloV4-tiny model. This model was chosen for its balance between efficiency and accuracy; YoloV4 remains well suited for embedded systems due to its compact architecture, which allows for rapid inference without compromising detection capabilities. We specifically selected YoloV4-tiny because it offers a significant reduction in computational requirements compared to its full-size counterpart, making it ideal for deployment on resource-constrained devices such as the Raspberry Pi-5 (RPi5) [18]. Furthermore, we identified the need to customize this model for our specific use case. Customization is necessary to enhance the model's performance in classifying droplets within our unique microfluidic setup. By fine-tuning the model on our custom dataset of microfluidic droplets, we ensure that it can accurately distinguish between different droplet types (no cell, one cell, multiple cells). This customization process

allows us to optimize the model's architecture and parameters, resulting in improved accuracy and faster inference times specifically for our droplet classification task. This also aligns with prior work on optimizing ML models for energy-efficient applications, where ML has been used in low-power applications [19]; in a similar vein, our approach explores the feasibility of deploying deep learning models on resource-limited embedded systems by balancing accuracy, inference speed, and computational complexity.

We evaluate performance in terms of inference time and mean average precision. On the RPi5, our system successfully classifies three distinct droplet types (no cell, one cell, multiple cells) in 13 ms, while maintaining over 98% accuracy. We compare the classification performance metrics of our customized YoloV4-tiny model against seven other models/platforms, including a recent, low-cost and highly compact edge device with TPU capabilities. The main steps and contributions of this research are as follows:

- We generate a new custom dataset of 975 microfluidic droplets images. The dataset is created from two videos recorded in microfluidic experimental setup.
- We introduce a droplet classification method (based on the YoloV4-tiny model) that offers high accuracy and inference speed, while remaining lightweight for resource-constrained implementation. High accuracy is achieved through data augmentation, and high inference speed is achieved by reducing the number of filters in the convolutional layers by 20% and applying batch processing (six images per batch).
- Our neural network is trained using 878 images extracted from the custom dataset of 975 images. We first implement and deploy our proposed approach on two resource-constrained single-board computers (SBCs): initially on a Raspberry Pi-4B (RPi4) with a BCM2711 chip and 8 GB RAM, and then on an RPi5 with a BCM2712 SoC and 8 GB RAM, achieving high accuracy (98%) and fast inference (13 ms). We then compare our results with those obtained on other platforms, including the compact and

low-cost MaixCam board [20] with an SG2002 SoC featuring a TPU¹ and 256 MB RAM.

We describe our customized model, built on deep learning principles, and how it attains both high accuracy and rapid performance, allowing for effective analysis of microfluidic droplets containing cells. Furthermore, we evaluate our system using a new and previously unexamined test dataset, highlighting its robustness. By integrating deep learning methods into a resource-limited platform, our work supports the development of resource-efficient droplet classification, thereby advancing the field of portable microfluidic technology.

The remainder of this paper is organized as follows. Section 2 describes the materials and method used in this work, particularly focusing on model implementation for resource-constrained devices. Section 3 presents the results, including a comparative performance analysis. Section 4 concludes the paper and suggests potential future research.

2. Materials and method

2.1. Dataset

The acquisition of training data is a major bottleneck in advancing microfluidic object detection and classification due to the scarcity of available datasets. In our work, a camera positioned directly above the microfluidic chip records the experimental video stream, capturing droplets formed at the T-junction of two channels. To extract images for model training, we used two distinct experimental videos [21] recorded under varying conditions (illustrative still images extracted from these videos are presented in Fig. 1). The images were retrieved by capturing video frames at a rate of 15 frames per second (fps), rather than the standard 1 fps, and were stored in JPG format. Each image contains either one or two microfluidic droplets, and each droplet may contain either no cell, a single cell, or multiple cells.

Next, to detect and classify such droplets in images, we consider an ML-based approach based on CNNs. However,

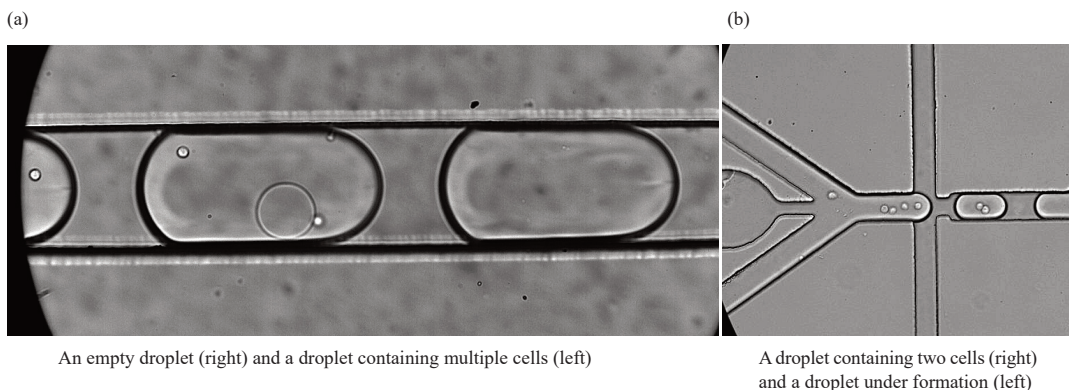


Fig. 1. Images of droplets flowing inside the microfluidic channel. Original resolutions are 1024×416 pixels for (a) and 560×512 pixels for (b); both are resized to 416×416 pixels during the training phase of the proposed customized YoloV4-tiny model.

¹ Similar to an NPU for low-precision operations (INT8) used in edge AI inference.

while CNNs are powerful tools for image classification, their performance can be negatively impacted by a lack of diversity in the training dataset. Without sufficient variety, the model may struggle to generalize well to new, unseen data. This may lead to overfitting, where the model performs well on training data but poorly on new examples. To address this issue, the two above-mentioned video recordings offer some diversity by capturing variations in illumination, contrast, motion blur, and object positioning, thereby helping to reduce overfitting.

However, to further combat the risk of overfitting, we implemented a data augmentation strategy to synthetically expand the size and diversity of the training dataset by creating modified versions of existing images.

We applied the following augmentation techniques:

- Contrast modification: the image contrast was adjusted within a range of 0.4 to 1.6, simulating variations in lighting conditions and image quality that may occur in real-world scenarios.
- Gaussian blur: two ranges of Gaussian blur were applied: 0.3 to 0.9 and 1.1 to 2.5. This technique mimics different levels of focus or image clarity, helping the model become more robust to variations in image sharpness.
- Vertical and horizontal rotations: these transformations enable the model to recognize objects regardless of their orientation within the image.

The augmentation process significantly expanded the dataset to a total of 975 images. This expanded dataset was then split into two parts, i.e., training and validation set, and testing set, as follows:

- Training and validation set: 878 images (approx. 90% of the total 975 images), with 791 images (approx. 90% of the 878 images) used for training and 87 images (approx. 10% of the 878 images) for validation². This validation set plays a vital role in assessing the model's generalization capabilities and helps prevent overfitting during the training process.
- Testing set: 97 images (approx. 10% of the total 975 images).

This split ensures a substantial amount of diverse data for training and validation while reserving a separate set for final testing to evaluate the model's performance on unseen data.

2.2. Image annotation

The process of image annotation is a crucial step in preparing datasets for object detection tasks, such as identifying droplets in IFC systems. This procedure requires recognizing and precisely marking the locations of all target objects within each image.

To conduct the annotation, we leveraged the Make Sense AI tool [22], which provides an intuitive interface for manual labeling and allows for meticulous, high-precision annotation of each image. We opted for rectangular bounding boxes as our annotation method, because they effectively capture the spatial extent of droplets within the images. The annotation process consisted of the following steps:

- Uploading images to the makesense.ai platform;
- Carefully examining each image for droplets;

- Drawing rectangular bounding boxes around each identified droplet;
- Verifying the accuracy of annotations through multiple reviews;
- Exporting the annotation data in the required format (.txt files).

The annotation format we adopted aligns with the YoloV4-tiny architecture requirements, using the [x, y, w, h] convention. In this format, (x, y) represents the center point of the bounding box, providing the focal point of the detected object; w denotes the width of the bounding box, capturing the horizontal extent of the droplet; and h stands for the height of the bounding box, representing the vertical extent of the droplet.

2.3. Proposed customized CNN model and deployment

In response to the requirements set forth by a bioanalytical specialist, we categorized the droplets into three groups: (i) empty droplets, (ii) droplets containing a single cell, and (iii) droplets containing multiple cells. Only droplets that are entirely visible in an image are factored into the classification process.

Our approach leverages the YoloV4-tiny model as a foundation. We implemented a refinement process to optimize the model for our specific use case of microfluidic droplet detection, with a focus on inference speed without significantly compromising accuracy, as outlined in what follows.

The original YoloV4-tiny architecture consists of 21 convolutional layers, organized into a series of cross-stage partial network (CSPNet) blocks. A key modification in our customization process involved a careful reduction in the number of filters in these convolutional layers, as shown in Fig. 2. Through a series of ablation studies, we empirically determined that a 20% reduction in filter count provides a suitable trade-off between model complexity and accuracy, as explained below. Each cross-stage partial (CSP) module consists of a convolutional layer followed by batch normalization and the Leaky-RELU activation function, collectively referred to as CL. The module also features skip connections, which help achieve an optimal balance between detection efficiency and accuracy. Three CSP modules in the backbone progressively extract the image features. For the detection head, two heads are used for detecting larger and smaller objects. The detection section consists of one CL block and a convolutional layer, followed by the detection layer. The feature maps are upsampled and combined with residual connections from the same feature map resolution for the detection head at the second scale. The first scale, with a feature map size of 13×13 , is used for larger objects, while the second scale, with a feature map size of 26×26 , targets smaller objects.

To ensure that the accuracy remained within acceptable bounds, we used an iterative process of filter reduction and performance evaluation. Through a series of ablation studies, we empirically tested different filter reduction levels (10%, 20%, and 30%). We used a validation set to monitor the

² A 10% validation split is commonly used for microfluidic applications with relatively small datasets, such as in [4], which used 786 images.

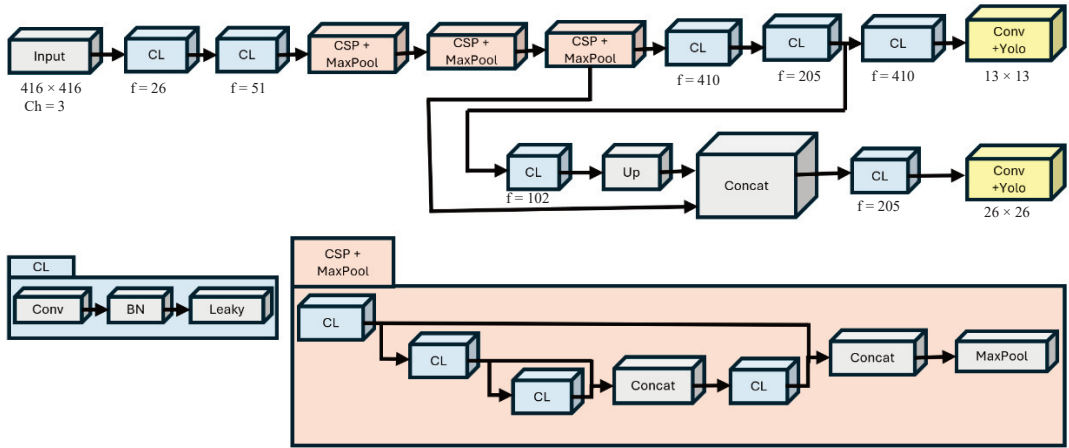


Fig. 2. Proposed customized YoloV4-tiny architecture. The network accepts a three-channel input ($ch = 3$), and its backbone comprises a series of convolutional layers (CL) and cross-stage partial (CSP) blocks combined with MaxPooling. Skip connections and upsampling (Up) are used for multi-scale feature fusion via concatenation. The number of filters is indicated by f , and the final stage includes two detection heads. Abbreviations: Conv – convolutional, BN – batch normalization.

model's performance after each round of filter reduction, ensuring that any accuracy drop remained minimal. Based on previous studies and our experimental results, a 20% filter reduction typically results in a 1–5% decrease in mean average precision (mAP), while batch processing may introduce an additional 0.5–2% loss due to batch normalization effects. In our case, empirical evaluation showed that the accuracy drop remained within this expected range, making the trade-off acceptable for our specific application. The 20% filter reduction resulted in a significant 36.16% decrease in the weight file size, from 22.4 MB to 14.3 MB.

In practice, this filter reduction strategy serves four purposes, listed below.

First, by reducing the number of filters, we significantly decrease the total number of parameters in the model. This compression result in a smaller model size, which is crucial for deployment on resource-constrained edge devices used in portable microfluidic systems.

Second, fewer filters result in fewer computations during the forward pass of the network. This directly translates to faster inference times, which are critical for real-time droplet detection and classification in flow cytometry applications.

Third, reducing the model's complexity through filter reduction can help mitigate overfitting, especially when working with limited datasets, which are common in specialized scientific applications, such as microfluidic droplet analysis.

Fourth, the reduced model size requires less memory during both training and inference, making it more suitable for deployment on devices with limited RAM.

Then, we fine-tuned the hyperparameters to optimize the performance of our customized YoloV4-tiny model. After extensive experimentation, we settled on a final learning rate of 0.00261, which was determined through a cyclical learning rate test to find an optimal balance between convergence speed and stability. This learning rate was coupled with a

weight decay rate of 0.0005 to prevent overfitting and improve generalization.

Finally, we set a momentum of 0.9 in the stochastic gradient descent optimizer, which helped accelerate convergence and mitigate oscillations during training. The size of the batch of images was set to 64 to balance memory constraints and the need for stable gradient estimates. This batch size was further divided into eight subdivisions and the training process was configured to run for a maximum of 6000 batches, following the default settings recommended for YoloV4-tiny.

For training, we used the Darknet framework [23] on Google Colaboratory Pro, with a T4 GPU configured with a software environment including Python 3.10.12, CUDA 12.2, cuDNN 8.9.6, and OpenCV 4.8.0. The trained model processes each input image to assign a confidence score to each detected droplet. We assessed performance by evaluating accuracy and mean average precision on the validation dataset. As depicted in Fig. 3, the model achieved optimal performance at 6000 iterations, reaching a training loss value of 0.09.

3. Deployment results

This section first outlines the outcomes of our proposed system for classifying microfluidic droplets, designed particularly for integration into portable devices. Our analysis emphasizes three primary aspects: classification accuracy, inference processing time, and resource consumption on a resource-constrained device. The section then presents a comparison with similar studies.

3.1. Droplet classification performance evaluation

We evaluated the performance of droplet classification using an input size of 416×416 pixels and a non-maximum-suppression threshold set to 0.7 for our test datasets. Figure 4 illustrates examples of droplet detection outcomes on images from our test dataset.

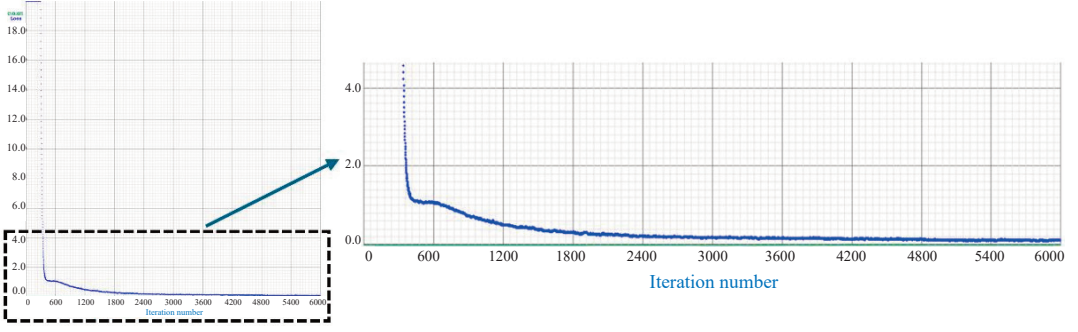


Fig. 3. Left: overview of the training loss function for YoloV4-tiny with compressed filters, trained for up to 6000 iterations, at which point the average loss reaches 0.09; right: zoomed-in view showing the convergence of the training loss function.

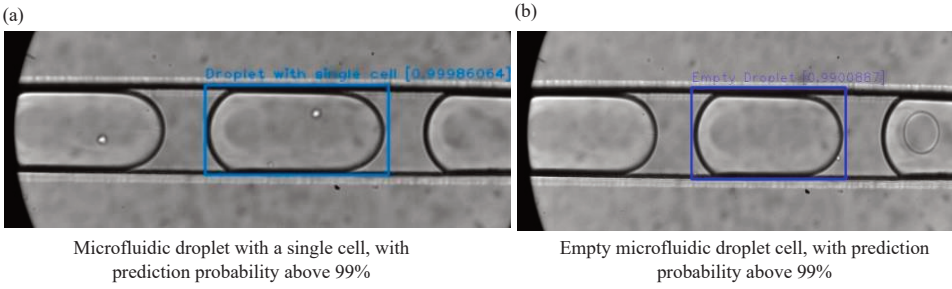


Fig. 4. Examples of microfluidic droplet classification results after 6000 training iterations, with prediction probability above 99%.

3.2. Custom model performance

Table 1 presents the performance metrics of our customized YoloV4-tiny model, which are derived using Eqs (1) and (2). In these equations, *TP* represents true positives, *FP* signifies false positives, and *FN* stands for false negatives:

$$\text{Precision} = TP / (TP + FP), \quad (1)$$

$$\text{Recall} = TP / (TP + FN). \quad (2)$$

The model demonstrates substantial reliability and effectiveness in accurately distinguishing various droplet types, as shown by its high precision of 0.95 and perfect recall of 1, resulting in minimal misclassification. The mAP at an intersection over union (IoU) threshold of 0.5, denoted as mAP@0.5, is a reflection of the accuracy, where IoU measures the over-

lap between predicted and actual bounding boxes. The robustness of the model is significant, with a notable mAP of 99.95%.

Our implementation was evaluated on both RPi4 and RPi5 SBCs, each equipped with 8 GB of RAM and running on a 64-bit operating system. High-speed image processing is crucial for real-time applications. To assess processing efficiency, we conducted tests with the model using the appropriate test dataset on the specified devices. The test results showed that the processing duration on the RPi5 was 20 times faster (13 ms) compared to RPi4 (265 ms).³ On the RPi5, the model achieved a classification inference time of 13 ms, equating to 76 fps; this speed, obtained with a six-image batch, suffices for practical applications. Given these superior results with the RPi5, the RPi4 was not considered for further analysis.

As noted above, these results were obtained using batches of six images rather than a single image. We evaluated inference speed across batch sizes ranging from 1 to 16 images and observed that processing time improved up to a batch size of six, beyond which no further improvements were noted. Subsequently, all results were averaged, and inference time was calculated for a single input image.

Beside inference processing time, resource utilization was an important part of our assessment due to the limitations of portable devices. We tracked CPU and memory consumption during classification tasks to guarantee the system's performance efficiency. As shown in Table 2, the mean CPU usage

Table 1. Classification performance metrics of the proposed customized YoloV4-tiny model (test dataset = 97 images of droplets, each containing zero, one, or multiple cells)

Metric	Performance value (%)
TP	112
FP	6
FN	0
Precision	0.95
Recall	1
mAP@0.5	99.95

³ We also transformed our customized model into TensorFlow and then TensorFlow Lite (TFLite) to try to reduce the inference processing time. However, the inference processing time for a single image exceeded 1 s on RPi4.

Table 2. Resource utilization of the proposed customized YoloV4-tiny model on RPi5

Resource	Average usage
CPU usage	70%
Memory usage	60%

and memory consumption on the RPi5 were 70% and 60%, respectively. These values suggest that the proposed approach is adequately lightweight, making it applicable for portable use, since CPU and memory consumption remain within practical limits.

3.3. Comparison with similar studies

In this section, we juxtapose the classification performance metrics of our optimized YoloV4-tiny model results against seven other models/platforms. It should be noted that direct one-to-one comparisons are not feasible because (i) previous works do not focus on the same applications as ours, and (ii) many of these studies were implemented on high-performance desktop PCs and/or GPUs. Nevertheless, such a comparison helps get a better understanding of the trade-offs between accuracy and inference processing time.

To balance inference processing time and accuracy, we leveraged model optimizations, hardware-aware deployment strategies, and empirical comparisons. Notably, we deployed our customized YoloV4-tiny on an RPi5, and also retrained YoloV5-s using our dataset for deployment on the MaixCam board. This additional training and deployment were conducted as part of this work and are briefly described below, prior to presenting the overall comparison table.

3.4. Training and deployment of the YoloV5-s model onto the MaixCam board

The MaixCam board [20] is a prime example of a recent (July 2024), low-cost (approx. 34 EUR), and highly compact (22.86×35.56 mm; see Fig. 5) edge device with an embedded neural processing unit. The MaixCam board is based on the LicheeRV Nano module [24], built around a SOPHGO SG2002 system on chip (SoC) [25]; it features a primary 700 MHz RISC-V C906 core and a secondary (boot-selectable) 1 GHz RISC-V C906 core or 1 GHz ARM A53 core, along with 256 MB of on-chip RAM. Notably, the SG2002 features a TPU capable of up to 1 tera operations per second (TOPS) @INT8, which should help reduce the inference processing time.

We used the same dataset described earlier in the paper. We converted the labeled dataset from a Darknet-compatible format (.jpg and .txt files) to a VOC-compatible format (.jpg and .xml), and uploaded the data to the MaixHub environment [26] for training the model and generating a format compatible with the SG2002 chip on the MaixHub board (.mud and .cvi files).

We also experimented with several hyperparameters, such as batch size and learning rate. The best validation accuracy results that we obtained empirically were with a batch size of four and a learning rate of 0.0001, which yielded a validation accuracy of 0.966 at epoch #120.

3.5. Contrasting combinations of ML models and their deployments

As mentioned earlier, a direct comparison of the different combinations of ML models and their deployments on various hardware targets is not feasible; however, some insights can be derived from Table 3.

Firstly, when compared to desktop deployments, it can be clearly seen that our customized YoloV4-tiny model deployed on the RPi5 delivers a smaller inference processing time of 13 ms, compared to the 74.3 ms reported in [10] for single-class detection using YoloV5 on an Intel Core i7, along with comparable mAP accuracy values ($99.95\% @ 0.5$ vs $99.3\% @ 0.5$, respectively). It is also worth noting that studies using high-end PCs or GPUs ([12–14]) achieved both high mAP and low processing times (where such metrics were reported); however, the GPUs used – such as the NVIDIA Tesla K80 in [12] and the Nvidia Tesla K40c in [14] – are dual-slot PCIe cards ($267 \text{ mm} \times 111 \text{ mm}$), with a computational power of approximately $8.74 \text{ TFLOPS} @ \text{FP32}$ / $2.91 \text{ TFLOPS} @ \text{FP64}$ and $4.29 \text{ TFLOPS} @ \text{FP32}$ / $1.43 \text{ TFLOPS} @ \text{FP64}$, respectively. These boards have power requirements of up to 375 W and 245 W, respectively. Such specifications clearly position these solutions for non-portable applications due to the physical size and power requirements of their processing units.

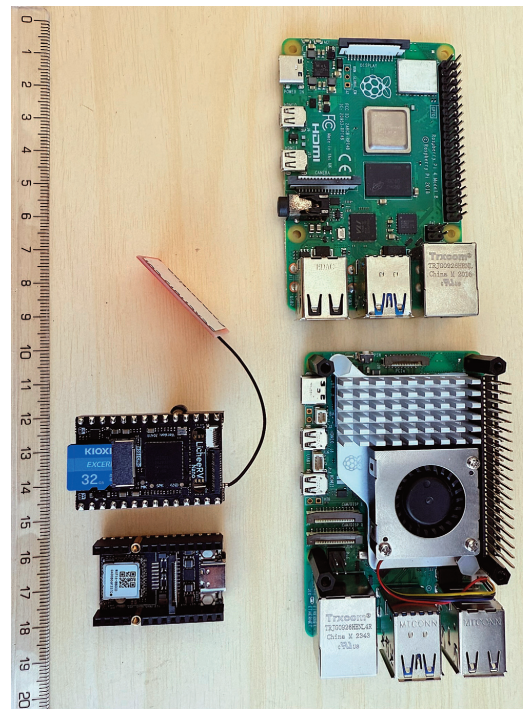


Fig. 5. Boards used in this work. Clockwise from the top: RPi4 SBC (used only up to Section 3.2 in this paper), RPi5 SBC with heatsink/fan, and MaixCam edge devices (top side with SG2002 SoC, bottom side with WiFi chip), along with a metric ruler. As can be seen, the MaixCam is a highly compact board owing to its minimal connectivity options (USB and WiFi), yet featuring a SoC with a 1 TOPS@INT8 TPU.

Table 3. Combinations of machine learning models and their deployments

Reference	Model	Platform	Object	mAP, %	Average inference processing time per image
[12]	CNN	NVIDIA Tesla K80 GPU	Cells	95.7	2.2 ms
[13]	CNN	High-end PC (assumed)	Cells	93–99	Not reported in the paper
[14]	CNN	Nvidia Tesla K40c GPU	Cells	99	Not reported in the paper
[11]	Mask-RCNN	Nvidia RTX 2080 Ti GPU	Cells	73	Not reported in the paper
[10]	YoloV5	Intel Core i7-12650H	Droplet	99.3@0.5	74.3 ms
This work	Customized YoloV4-tiny	RPi5	Droplet	99.95@0.5	13 ms
This work	YoloV5-s	RPi5	Droplet	92.10@0.5	208.5 ms
This work	YoloV5-s	MaixCam	Droplet	55.09@0.5	5.34 ms

In contrast, the RPi5 measures only $85.6 \text{ mm} \times 56.5 \text{ mm}$ and has a power requirement of up to 25 W only, providing approximately up to 12 GFLOPS@FP32/CPU, 20 GFLOPS@FP32/GPU, and 750 MFLOPS@FP64/CPU. (Note: The RPi4's Cortex-A72 CPU lacks efficient FP64 acceleration and native GPU acceleration, so Yolo is executed as FP32/CPU.) The MaixCam features even more compact specifications: $22.86 \times 35.56 \text{ mm}$ in size, up to 2.5 W power consumption, approximately 2.8–4 GFLOPS@FP32, 250 MFLOPS@FP64, and 1 TOPS@INT8/TPU (Yolo is accelerated on this TPU).

Secondly, when comparing the resource-constrained deployments, it can be seen that the MaixCam running the YoloV5-s model achieves a smaller inference processing time (5.34 ms) than the customized YoloV4-tiny on the RPi5 (13 ms); however, this comes at the cost of a much lower mAP of 55.09%. This lower score stems from the more aggressive quantization required to map the model onto the INT8 TPU of the SG2002 SoC on the MaixCam board.

We also trained and deployed the YoloV5-s model for the RPi5; while it achieved a mAP@0.5 of 92.10%, the classification time was high at 208.5 ms, indicating that this combination is not favorable.

Due to a yet unsolved issue with the toolchain, we were unable to convert our customized YoloV4-tiny model for the MaixCam. Instead, we provide estimated performance numbers, as explained below. Using the YoloV5-s results, we can derive an approximate performance scaling factor between the MaixCam and RPi5 platforms:

- Speedup factor (MaixCam vs RPi5 for YoloV5-s): $208.5 \text{ ms (RPi5)} / 5.34 \text{ ms (MaixCam)} \approx 39.1\times$. This reflects that the MaixCam's TPU accelerates inference significantly compared to the RPi5's CPU-based inference.
- Accuracy drop (YoloV5-s on MaixCam vs. RPi5): from 92.10@0.5 to 55.09@0.5 $\approx 40.2\%$. This reflects that the MaixCam's TPU has precision limitations due to lower-bit computation, i.e., only INT8 (MaixCam) instead of FP32 (RPi5).
- Since YoloV4-tiny is structurally similar to YoloV5-s, we assume the same performance ratio applies. Using the $39.1\times$ speedup factor yields $13 \text{ ms} / 39.1 \approx 0.33 \text{ ms}$. This suggests that the MaixCam TPU could theoretically process our customized YoloV4-tiny model in under 1 ms; however, real-world constraints (e.g., memory access, TPU overhead) would likely increase inference time to around 1–2 ms. On the other hand, we expect the mAP as

$99.95 \times (100 - 40.2) \approx 59.8\%$; i.e., running our customized YoloV4-tiny model on the MaixCam TPU might reduce accuracy to around 60%, which is not favorable despite the small processing time.

Besides the above accuracy and processing time performance results, it should also be noted that the RPi5 is an SBC with power requirements of up to 25 W (5 V, 5 A). This is not necessarily a major concern; for example, our portable setup [27] uses the RPi as a common platform, as this allows implementing additional functionalities on a single board. On the other hand, for applications where power and/or energy requirements are more stringent, a board such as the MaixCam – with an approximately $10\times$ lower power requirement of 2.5 W (5 V, 500 mA) – would be more suitable, if the cost of lower accuracy is acceptable.

To sum up, the above results illustrate that our customized YoloV4-tiny on the RPi5 offers a good trade-off among the different tested combinations, offering significant accuracy with competitive processing time on a resource-limited platform, which makes it suitable for portable IFC devices.

4. Conclusion

This work demonstrated the feasibility of effective droplet classification in IFC on a resource-constrained device using a customized YoloV4-tiny model. A new dataset of droplet images was created from videos recorded on our existing setup, with improved accuracy and robustness achieved through data augmentation. The model's inference processing time was reduced by cutting the convolutional layer filters by 20% and using batches of six images. Our proposed system can accurately classify droplets in 13 ms, achieving an accuracy surpassing 99% when running on an RPi5 SBC. Additional experiments with YoloV5-s on the compact MaixCam board, featuring an SG2002 SoC with a TPU, illustrated how a significantly smaller classification time (5.34 ms) must be traded off for accuracy (mAP@0.5) due to quantization.

Future work could focus on decreasing inference processing time on resource-constrained platforms by applying more sophisticated pruning and quantization techniques, aiming for real-time or near-real-time classification while preserving accuracy. Moreover, we could explore other Yolo versions; while larger models (e.g., YoloV4, YoloV5-m/l/x) might offer higher accuracy, they would increase computational load, limiting real-time feasibility. Conversely, YoloV7-

tiny and YoloV8-nano may enhance accuracy while remaining efficient for edge deployment, though their optimization for our target hardware requires further investigation.

While we experimented with three different platforms, it should be noted that the landscape of edge AI hardware is rapidly evolving. For example, the recent Hailo-8 HAT for the RPi5 promises up to 26 TOPS (INT8). These developments offer promising avenues for further decreasing processing time while balancing accuracy; however, it is also essential to consider the maturity of these new solutions. Their integration (including the necessary software tools) into specialized scientific applications, such as IFC devices, will require additional careful evaluation and testing. We will investigate these issues in our future work.

Data availability statement

All data and code supporting this work are available at <https://github.com/taltechloc/sw-open-droplet-analysis>.

Acknowledgments

This work was supported by the Estonian Science Agency ETAg project “Cognitronic Lab-on-a-Chip System for Highly-Automated Flow Cytometry” (CogniFlow-Cyte, grant No. PRG620). The authors also acknowledge support from the Estonian Education and Youth Board project “Artificial Intelligence, Edge Computing and IoT Solutions in Distributed Systems” (AIoT*5G, grant No. ÕÜF11), and the Estonian IT Academy project “Sustainable Artificial Internet of Things” (SAIoT, grant No. TEMTA138). The authors thank Assistant Prof. Tomasz Kaminski, University of Warsaw, Poland, for providing the droplet image dataset for training the droplet classification model. The publication costs of this article were partially covered by the Estonian Academy of Sciences.

References

- Trinh, T. N. D., Do, H. D. K., Nam, N. N., Dan, T. T., Trinh, K. T. L. and Lee, N. Y. Droplet-based microfluidics: applications in pharmaceuticals. *Pharmaceuticals*, 2023, **16**(7), 937. <https://doi.org/10.3390/ph16070937>
- Chen, Z., Kheiri, S., Young, E. W. K. and Kumacheva, E. Trends in droplet microfluidics: from droplet generation to biomedical applications. *Langmuir*, 2022, **38**(20), 6233–6248. <https://doi.org/10.1021/acs.langmuir.2c00491>
- Liu, Z., Jin, L., Chen, J., Fang, Q., Ablameyko, S., Yin, Z. et al. A survey on applications of deep learning in microscopy image analysis. *Comput. Biol. Med.*, 2021, **134**, 104523. <https://doi.org/10.1016/j.compbiomed.2021.104523>
- Rutkowski, G. P., Azizov, I., Unmann, E., Dudek, M. and Grimes, B. A. Microfluidic droplet detection via region-based and single-pass convolutional neural networks with comparison to conventional image analysis methodologies. *Mach. Learn. Appl.*, 2022, **7**, 100222. <https://doi.org/10.1016/j.mlwa.2021.100222>
- Afrin, F., Le Moulec, Y. and Pardy, T. Microfluidic droplet classification through tuned convolutional neural network on a resource constrained platform. In *2024 19th Biennial Baltic Electronics Conference (BEC), Tallinn, Estonia, 2–4 October 2024*. IEEE, 2024, 1–4.
- Zhou, X., Mao, Y., Gu, M. and Cheng, Z. WSCNet: biomedical image recognition for cell encapsulated microfluidic droplets. *Biosensors*, 2023, **13**(8), 821. <https://doi.org/10.3390/bios13080821>
- Soldati, G., Del Ben, F., Brisotto, G., Biscontin, E., Bulfoni, M., Piruska, A. et al. Microfluidic droplets content classification and analysis through convolutional neural networks in a liquid biopsy workflow. *Am. J. Transl. Res.*, 2018, **10**(12), 4004–4016.
- Kensert, A., Harrison, P. J. and Spjuth, O. Transfer learning with deep convolutional neural networks for classifying cellular morphological changes. *SLAS Discov.*, 2019, **24**(4), 466–475. <https://doi.org/10.1177/2472555218818756>
- Lee, K., Kim, S.-E., Doh, J., Kim, K. and Chung, W. K. User-friendly image-activated microfluidic cell sorting technique using an optimized, fast deep learning algorithm. *Lab Chip*, 2021, **21**(9), 1798–1810. <https://doi.org/10.1039/D0LC00747A>
- Xu, J., Fan, W., Madsen, J., Tanev, G. P. and Pezzarossa, L. AI-based detection of droplets and bubbles in digital microfluidic biochips. In *2023 Design, Automation & Test in Europe Conference & Exhibition (DATE), Antwerp, Belgium, 17–19 April 2023*. IEEE, 2023, 1–6. <https://doi.org/10.23919/DATE56975.2023.10136887>
- Ghafari, M., Mailman, D., Hatami, P., Peyton, T., Yang, L. and Dang, W. A comparison of YOLO and Mask-RCNN for detecting cells from microfluidic images. In *2022 International Conference on Artificial Intelligence in Information and Communication (ICAIIIC), Jeju Island, Korea, 21–24 February 2022*. IEEE, 2022, 204–209.
- Li, Y., Mahjoubfar, A., Chen, C. L., Niazi, K. R., Pei, L. and Jalali, B. Deep cytometry: deep learning with real-time inference in cell sorting and flow cytometry. *Sci. Rep.*, 2019, **9**, 11088. <https://doi.org/10.1038/s41598-019-47193-6>
- Suzuki, Y., Kobayashi, K., Wakisaka, Y. and Ozeki, Y. Label-free chemical imaging flow cytometry by high-speed multicolor stimulated Raman scattering. *Proc. Natl. Acad. Sci. U.S.A.*, 2019, **116**(32), 15842–15848. <https://doi.org/10.1073/pnas.1902322116>
- Meng, N., Lam, E. Y., Tsia, K. K. and So, H. K.-H. Large-scale multi-class image-based cell classification with deep learning. *IEEE J. Biomed. Health Inform.*, 2019, **23**(5), 2091–2098. <https://doi.org/10.1109/jbhi.2018.2878878>
- Riti, J., Sutra, G., Naas, T., Volland, H., Simon, S. and Perez-Toralla, K. Combining deep learning and droplet microfluidics for rapid and label-free antimicrobial susceptibility testing of colistin. *Biosens. Bioelectron.*, 2024, **257**, 116301. <https://doi.org/10.1016/j.bios.2024.116301>
- Ahmadvpour, A., Shojaeian, M. and Tasoglu, S. Deep learning-augmented T-junction droplet generation. *iScience*, 2024, **27**(4), 109326. <https://doi.org/10.1016/j.isci.2024.109326>
- Asama, R., Liu, C. J. S., Tominaga, M., Cheng, Y.-R., Nakamura, Y., Kondo, A. et al. Droplet-based microfluidic platform for detecting agonistic peptides that are self-secreted by yeast expressing a G-protein-coupled receptor. *Microb. Cell Factories*, 2024, **23**, 104. <http://dx.doi.org/10.1186/s12934-024-02379-0>
- Raspberry. *Raspberry Pi Hardware*. <https://www.raspberrypi.com/documentation/computers/raspberry-pi.html> (accessed 2024-12-13).
- Khan, S. Z., Le Moulec, Y. and Alam, M. M. An NB-IoT-based edge-of-things framework for energy-efficient image transfer. *Sensors*, 2021, **21**(17), 5929. <https://doi.org/10.3390/s21175929>
- Speed. *MaixCam – Fast Development for AI Vision and Audio Projects*. <https://wiki.speed.com/hardware/en/maixcam/index.html> (accessed 2024-12-13).
- De Jonghe, J., Kaminski, T. S., Morse, D. B., Tabaka, M., Ellermann, A. L., Kohler, T. N. et al. spinDrop: a droplet microfluidic platform to maximise single-cell sequencing information content. *Nat. Commun.*, 2023, **14**, 4788. <http://dx.doi.org/10.1101/2023.01.12.523500>
- Skalski, P. *makesense.ai*. <https://github.com/SkalskiP/makesense> (accessed 2024-12-14).
- Redmon, J. *Darknet Open Source Neural Network Framework*. <https://github.com/pjreddie/darknet> (accessed 2024-12-14).

24. Sipeed. *LicheeRV Nano*. https://wiki.sipeed.com/hardware/en/lichee/RV_Nano/1_intro.html (accessed 2024-12-13).
25. SOPHGO. *SG200X Hardware*. <https://github.com/sophgo/sophgo-hardware/tree/master/SG200X> (accessed 2024-12-13).
26. Sipeed. *MaixHub*. <https://maixhub.com/> (accessed 2024-12-14).
27. Jõemaa, R., Gyimah, N., Ashraf, K., Pärnamets, K., Zajt, A. and Scheler, O. CogniFlow-Drop: integrated modular system for automated generation of droplets in microfluidic applications. *IEEE Access*, 2023, **11**, 104905–104929.

Konvolutsioonilistel närvivõrkudel (CNN) põhinev mikrofluidsete tilkade klassifitseerimine portatiivsetes voolutsütomeetrites

Fariha Afrin, Yannick Le Moullec, Tamas Pardy ja Toomas Rang

Tilkade klassifitseerimine on oluline aspekt pilditöötlust sisaldavate voolutsütomeetrite arendamisel. Süvaõppe algoritmid on võimelised avastama ja klassifitseerima tilku suuremõõtmelistes laboriseadmetes, kuid sarnase tehnoloogia rakendamine portatiivsetes seadmetes kujutab endast suurt väljakutset, kuna nende arvutusvõimsus ei vasta kompaktsete seadmete arvutusvõimekusele. See on oluline takistus üleminekul statsionaarsetelt laboriseadmetelt välioludes kasutatavatele portatiivsetele lahendustele. Takistuse ületamiseks tutvustame artiklis kohandatud YoloV4-tiny mudelit, mida on rakendatud Raspberry Pi-5 (RPi5) platvormil.

Närvivõrgupõhist lahendust treeniti 878 erineva kujutise abil, mis pärinesid 975 kujutisega kohandatud andmehulgast. See andmehulk koguti meie loodud reaalse eksperimentaalse mikrofluidikaseadmega. Tulemusi hindasime interferentsiaja ja keskmise täpsuse (mAP – *mean average precision*) alusel. Loodud lahendus suutis edukalt klassifitseerida kolme selgelt eristatavat olukorda (tilk puudub, üks tilk, mitu tilka) 13 ms jooksul, saavutades keskmise täpsuse 99,95% lävendiga 0.5 (mAP@0,5). Samuti võrdlesime kohandatud YoloV4-tiny mudelit seitsme masinõppemudeli (ML) ja platvormi kombinatsiooniga, sealhulgas uusima, kompaktses ja soodsa tensoritöötlusega tippseadmega (MaixCam plaat koos LicheeRV Nano mooduli / SOPHGO SG2002-ga), mis kasutab YoloV5 algoritmi. YOLOv4-tiny RPi5 lahendust võrdlesime YOLOv5-s mudeli ja MaixCam platvormi kombinatsiooniga. Tänu lisatud tensoritöötluse algoritmile lühenes klassifitseerimise aeg 5,34 ms-ni, saavutades keskmise täpsuse 55,09% juhtudest lävendiga 0,5 (mAP@0,5). Täpsuse protsentuaalne vähenemine on tingitud kvantimisest. Uuring näitab, et süsteemi täpse disainimise abil on võimalik saavutada tasakaal täpsuse ja kiiruse vahel, võimaldades mikrofluidsete tilkade usaldusväärset klassifitseerimist ka piiratud arvutusvõimekusega portatiivsetes voolutsütomeetrites.

Appendix 4

IV

R. Jõemaa, **F. Afrin**, N. Gyimah, K. Ashraf, K. Pärnamets, T. Pardy, "Cogni-Flow: Integrated Modular System For Automated Droplet Microfluidic Bio-analysis", EUROSENSORS XXXVI, 01-04 September 2024, Debrecen, Hungary

CogniFlow: Integrated Modular System For Automated Droplet Microfluidic Bioanalysis

Rauno Jõemaa¹, Fariha Afrin¹, Nafisat Gyimah¹, Kanwal Ashraf¹, Kaiser Pärnamets¹, Lucas Giese²,
Mathieu Rocancourt², Tamás Pardy¹

¹ Tallinn University of Technology, 5 Ehitajate tee, 19086 Tallinn, Estonia,

² ENSEA, 6 Av. Du Ponceau, 95000 Cergy, France

tamas.pardy@taltech.ee

Summary: Droplet microfluidics uses discrete, chemically isolated droplets to multiplex reactions in bioanalysis. Droplets also allow process control exceeding traditional means, as well as continuous flow microfluidics. Here we present the concept of CogniFlow, an integrated, modular droplet microfluidic instrumentation platform, which automates complete droplet generation, imaging and detection workflows. The platform is demonstrated in cell encapsulation and detection. It is our hope that CogniFlow can significantly lower the entry barrier to instrumenting existing droplet bioanalytical workflows.

Keywords: droplets, microfluidics, instrumentation, control, cognitronics

Introduction

Droplets allow excellent process control by chemically isolating pico- to nanoliter range reaction volumes (like tiny test tubes), allowing selection of samples from a pool of millions of non-specific targets, which would be nearly impossible with conventional tools (e.g. well-plate stations) [1], [2]. This has opened the way to novel applications in microfluidic bioanalysis and biotechnology [3]. However, biology teams desiring to automate their droplet-based workflows today primarily load functions onto highly complex microfluidic chips, and/or use combinations of general-purpose off-the-shelf instruments, resulting in highly complex, expensive systems, which are difficult to scale and replicate, take significant time to build and are not energy-efficient [4], [5]. In this poster, we propose a technology platform that can address these challenges and reduce development time and cost associated with instrumenting droplet bioanalytical and biotechnology applications. In our previous paper we presented the droplet generation module of the platform [6]. In this poster, we present the integrated system, including droplet imaging and detection.

Novelty

CogniFlow offers the following novelties:

Transferability: modules are compact and portable, can be swapped, but the system has its standards. System can be reconfigured (e.g. different filters, lenses and light sources, camera vs. photodiode etc.), but workflows are easily replicated between labs.

Scalability: Low cost (<1k€-2k€/module, using 3D printed parts, low-cost electronics and optics) and standardization ensure easy replication of results in different labs, as well as scaling up throughput. Despite the low cost, performance of thus-far demonstrated components has been on par with comparable systems [6].

Efficiency: modules have power consumption ~10-20W, enabling battery operation. Wireless communication via ECAL [7] further reduces device footprint and increases portability. Optimizations to flow control reduce reagent waste.

Total automation: CogniFlow has 3 main modules: 1. Droplet generation/encapsulation [6], 2. Droplet imaging and 3. Embedded object classification algorithm. These cover a complete droplet-based imaging or light intensity analysis workflow.

Demonstrations in literature meeting the majority of the aforementioned criteria are rare [1], [2].

System modules and results

Droplet generation [6]: the first demonstrated module of the platform generated droplets in 50-200 µm range with 5-10% coefficient of variability at up to 1kdps (1000 droplets per second). We controlled pressure drop and droplet size in a piezoelectric pumping setup with rapid and precise controller response, but a simple and inexpensive setup. On the poster, we mention this module for the sake of completeness.

Droplet imaging: the second module of the system presented here for the first time is based on an affordable optical setup using a Basler

acA640-750uc and a cheap 20x magnification lens to achieve ~750 fps imaging rate [8]. The integrated system (Fig. 1) is built on the same framework as the first module, by swapping internal components. The footprint is 22 x 33 x 40 cm. Imaging system output with encapsulated cells is shown in Fig. 2.

Object detection (Fig. 3): a droplet segmentation and object classification algorithm based on YOLOv4 and TinyYOLOv4 is presented, which is capable of detecting single cells in droplets with 85% accuracy on an image stream of droplets in flow provided by our academic partners. The algorithm can run on an NVIDIA® Jetson Nano™ AI accelerator. The AI accelerator board fits on top of the electronics stack of the second module (Fig 1. left), and while it will double system power consumption, the total will still be well below what PD3.0+ compliant power bank can provide.

Between the two hardware and the one software module presented here, the following solutions are in common: both hardware modules rely on a Raspberry Pi 4 for control and communication. Wireless communication is implemented via ECAL, as presented in [6], however, for experiments, a wired Ethernet connection is also possible. Power electronics is the same between module 1 and 2, as is the mechanical frame, which consists of 3D printed parts and commonly used metal fasteners.



Figure 1: Integrated droplet imaging module of the CogniFlow platform. The mechanical frame and enclosure (not shown here) are shared between all platform modules, optics and electronics stages are changed according to the purpose.

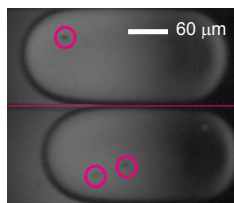


Figure 2: Droplets imaged with the droplet imaging module of CogniFlow using an affordable optical setup. Highlighted objects are algae cells encapsulated. Reagents used for droplet generation are described in [6].

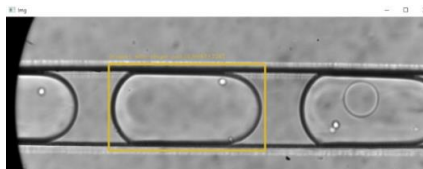


Figure 3: Object detection algorithm demonstration. Reference video of droplets with microbeads and cells encapsulated was kindly provided by University of Warsaw to test the algorithm on an independent dataset.

References

- [1] C. Huang, Y. Jiang, Y. Li, and H. Zhang, 'Droplet Detection and Sorting System in Microfluidics: A Review', *Micromachines (Basel)*, vol. 14, no. 1, p. 103, Dec. 2022, doi: 10.3390/mi14010103.
- [2] M. Gantz, S. Neun, E. J. Medcalf, L. D. van Vliet, and F. Hollfelder, 'Ultra-high-Throughput Enzyme Engineering and Discovery in In Vitro Compartments', *Chem Rev*, vol. 123, no. 9, pp. 5571–5611, 2023, doi: 10.1021/acs.chemrev.2c00910.
- [3] T. S. Kaminski, O. Scheler, and P. Garstecki, 'Droplet microfluidics for microbiology: techniques, applications and challenges', *Lab Chip*, vol. 16, no. 12, pp. 2168–2187, 2016, doi: 10.1039/C6LC00367B.
- [4] Y. Zhou, Z. Yu, M. Wu, Y. Lan, C. Jia, and J. Zhao, 'Single-cell sorting using integrated pneumatic valve droplet microfluidic chip', *Talanta*, vol. 253, p. 124044, Feb. 2023, doi: 10.1016/J.TALANTA.2022.124044.
- [5] Atrandi Biosciences, 'Styx | Atrandi Biosciences'. Accessed: Feb. 03, 2024. [Online]. Available: <https://atrandi.com/styx>
- [6] R. Jõemaa *et al.*, 'CogniFlow-Drop: Integrated Modular System for Automated Generation of Droplets in Microfluidic Applications', *IEEE Access*, vol. 11, pp. 104905–104929, 2023, doi: 10.1109/ACCESS.2023.3316726.
- [7] K. Ashraf, Y. Le Moullec, T. Pardy, and T. Rang, 'Design of Cyber Bio-analytical Physical Systems: Formal methods, architectures, and multi-system interaction strategies', *Microprocess Microsyst*, vol. 97, Mar. 2023, doi: 10.1016/J.MICPRO.2023.104780.
- [8] K. Parnamets, A. Koel, T. Pardy, and T. Rang, 'Open Source Hardware Cost-Effective Imaging Sensors for High-Throughput Droplet Microfluidic Systems', *Proceedings of the 2022 26th International Conference Electronics, ELECTRONICS 2022*, 2022, doi: 10.1109/IEEECONF55059.2022.9810383.

Appendix 5

V

F. Afrin, D. Ndubuisi Ezechukwu, Y. Le Moullec, T. Pardy, T. Rang, A. Koel, "CNN-Transformer Hybrid Model Towards Automated Droplet Image Quality Assessment of Portable Imaging Flow Cytometer". To appear in the proceedings of IEEE EUROCON 2025 - 21st International Conference on Smart Technologies, 4-6 June 2025, Gdynia, Poland.

CNN-Transformer Hybrid Model Towards Automated Droplet Image Quality Assessment of Portable Imaging Flow Cytometer

Fariha Afrin*, Dismas Ndubuisi Ezechukwu[†], Yannick Le Moullec*,
Tamas Pardy*, Toomas Rang* Ants Koel[‡]

*Thomas Johann Seebeck Department of Electronics, Tallinn University of Technology, Tallinn, Estonia
{fariha.afrin, yannick.lemoullec, tamas.pardy, toomas.rang}@taltech.ee

[†]Department of Electrical and Computer Engineering, University of Texas at Dallas, Texas, USA
dismas.ezechukwu@utdallas.edu

[‡]Estonian Military Academy, Tartu, Estonia
ants.koel@taltech.ee

Abstract—In imaging flow cytometry (IFC), image quality obtained from the camera is generally assessed subjectively, which is not suitable for automated processes. Automated image quality assessment (IQA) is essential for consistent and precise data analysis and interpretation in IFC, directly impacting the reliable classification of microparticles inside droplets. However, despite its importance, no dedicated research has been carried out to check the suitability of different techniques towards objective IQA. This work investigates the performance of no reference-IQA (NR-IQA) techniques within the context of a portable lab on a chip microfluidic droplet IFC device. While NR-IQA is suitable for scenarios lacking reference images, inferring image quality solely from images of various quality remains challenging. To resolve this, we first generate good and poor quality microfluidic droplet images and then compare the results obtained from selected NR-IQA approaches. Second, we propose a hybrid model comprising of a modified MobileNet followed by a transformer block specifically developed for microfluidic droplet IQA. Results show that this hybrid model yields competitive performance metrics, i.e. Spearman rank-order correlation coefficient (SRCC) and Pearson linear correlation coefficient (PLCC) of 0.73 and 0.75, respectively. The achievable inference time on a Raspberry PI-5 (RPI-5) is 110 ms, translating to 9 frames per second (FPS). Our work aims to provide valuable insight and guidance for researchers and practitioners in selecting appropriate techniques for ensuring reliable droplet IQA in portable IFC.

Index Terms—Convolutional Neural Network, Image Quality Assessment, Imaging Flow Cytometry, RPI-5, Transformer

I. INTRODUCTION

IFC has emerged as a powerful tool for analyzing cellular characteristics and functions at the single-cell level. IFC combines the high-throughput capabilities of traditional flow cytometry with the imaging capabilities of microscopy [1]. Numerous studies have shown that the detection and classification of droplets using machine learning (ML) or deep learning (DL) algorithms achieve high accuracies [2]. However, most existing models are trained and evaluated using only good quality images, typically assessed by human experts. In real-world IFC environments, image quality varies due to different

lighting conditions and fluid flows. This can result in issues such as blurriness and inconsistent contrast, potentially leading to false positive results [3] and reduced model reliability. Since the accuracy of IFC-based analysis and interpretation depends significantly on the quality of the acquired images [1], it is therefore essential to perform automated IQA to filter out poor-quality images at the initial stage, ensuring that only good-quality images are used in subsequent analytical steps in the imaging pipeline.

Automatic IQA can be performed using objective IQA methods, which are categorized into full-reference (FR-IQA), reduced-reference (RR-IQA), and no-reference (NR-IQA) approaches [4]. Among these, NR-IQA is particularly advantageous in practical applications, as it evaluates image quality without requiring a pristine reference image. This makes NR-IQA especially suitable for scenarios where reference images are not available, such as in real-time imaging systems and high-throughput analysis.

Recent research is increasingly focused on developing convolutional neural network (CNN)-based models for NR-IQA [4], due to their ability to learn complex features directly from image data. However, existing research primarily focuses on image restoration and measuring the quality of de-blurred images. Therefore, a significant research gap remains regarding the applicability of CNN-based models for IFC and their deployment on portable platforms. Addressing this gap can potentially increase the accuracy and efficiency of IFC analysis in real-world, portable settings, highlighting the need for further advancements in automation and portability.

For example, in [5], the authors used a U-net based network for image restoration, addressing the challenges associated with high flow speed and reduced exposure time. The quality of the restored images was assessed using blind/referenceless image spatial quality evaluator (BRISQUE), naturalness image quality evaluator (NIQE) and perception-based image quality evaluator (PIQE) metric. However, the results of these objective IQA methods are not fully correlated with human percep-

tual scores and their corresponding prediction distributions, highlighting the need for improved IQA techniques.

In [6], the researchers developed a deep learning based classifier and de-blurring system for red blood cell analysis. To evaluate the image quality, they used peak signal-to-noise ratio (PSNR) and structural index similarity (SSIM) to estimate blur and structural similarity. However, these IQA metrics require pristine reference images, which may not always be available in real-world imaging scenarios.

In [7], a tenengrad gradient function was used for the quantitative assessment of de-blurred images. This gradient-based method evaluates image sharpness by analyzing the difference in edges, making it particularly well-suited for fluorescence images, where the edge differences are strong. However, it is less suitable for low-gradient images because of insufficient contrast and inherent blurriness, which can lead to inaccurate sharpness measurements.

Recent studies indicate that hybrid architectures that combine complementary designs can boost model performance by leveraging the unique strengths of each sub-architecture. For example, in [8], a hybrid model comprising a CNN with a vision transformer was used for medical image classification, resulting in improved accuracy and generalization. Building on this concept, we propose a novel hybrid model for automated droplet IQA, specifically designed for implementation on embedded platforms. Our model predicts quality scores based on blurriness and contrast, ensuring a strong correlation with human perceptual scores. Unlike traditional approaches that focus solely on mean opinion score (MOS) and standard deviation (STD), our model also considers the distribution of the scores. By capturing the variability in IQA, our approach provides a more comprehensive and robust evaluation framework, making it well-suited for portable applications in IFC. While IQA processes can be resource-intensive and time-consuming due to manual data labeling, our selection of the sub-architectures for the hybrid model and optimization strategies can significantly reduce computational costs, human resource requirements, and processing time.

II. PROPOSED METHOD

The presentation of the proposed hybrid architecture and resulting model for droplet IQA is divided into four subsections: A) model architecture, B) loss function, C) dataset preparation, D) implementation setup, and E) evaluation metrics.

A. Model Architecture

Our model selection prioritized architectures with low computational parameters to ensure suitability for embedded platforms while maintaining strong performance. Three different CNN-based lightweight models (MobileNet, EfficientNet, and NasNet) were selected for training and evaluation on our dataset. Among them, MobileNet achieved the best performance but required further improvements and fine-tuning. The proposed quality predictor model architecture consists of two key components, i.e. i) a modified MobileNet (inspired from Neural Image Assessment (NIMA) [9]) serving as the

base feature extractor, followed by ii) a transformer block for global feature learning. The architecture is illustrated in Fig. 1, showing both the convolutional and transformer-based layers.

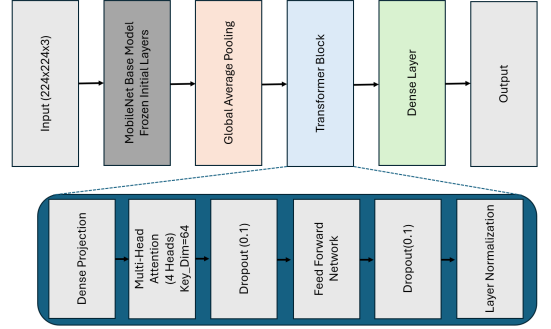


Fig. 1: Proposed architecture for predicting the quality of microfluidic droplet images.

To preserve the pre-trained features, the initial layers of the MobileNet are frozen, while the last layers were fine-tuned for improved learning. The model was initialized with ImageNet weights to accelerate convergence, particularly beneficial for limited datasets. The feature maps produced by MobileNet are then passed through a Global Average Pooling (GAP) layer to convert the output into a vector, making it suitable for input to the transformer block. The transformer block consists of a multi-head self-attention mechanism with four attention heads, and a feed-forward network with 512 units followed by dropout for regularization. To reduce the model complexity, the embedding dimension of the transformer block is set to 256. Subsequent to the transformer block, a dense layer with 128 units and Gaussian Error Linear Unit (GELU) activation further refines the features, and another dropout layer is added to prevent overfitting. The final output layer is a softmax-activated dense layer with 10 neurons, which corresponds the number of classes and produces the model's classification predictions. The model is regularized using L2 regularization and dropout to ensure generalization and avoid overfitting. This hybrid model is specifically designed for deployment on embedded, resource-constrained platforms to enable assessing the microfluidic droplet image quality on portable devices.

The proposed model predicts the distribution of ratings for a given image, and the mean is defined as μ :

$$\mu = \sum_{i=1}^N p_i \cdot i \quad (1)$$

where N is the total number of classes, p_i is the probability of the quality score falling into class i , and i is the class index.

The standard deviation of the score is computed as σ :

$$\sigma = \sqrt{\sum_{i=1}^N p_i \cdot (i - \mu)^2} \quad (2)$$

where μ is the mean score computed above, and $(i - \mu)^2$ represents the squared deviation from the mean.

B. Loss Function

We use the Earth Mover’s Distance (EMD) [10] training loss function; it measures the dissimilarity between two probability distributions by quantifying the minimum effort required to transform one distribution into the other. The EMD is defined as the root-mean-square of the differences between the cumulative distribution functions (CDFs) of the two distributions.

Given the ground truth probability mass function $\mathbf{p} = [p_1, p_2, \dots, p_N]$ and the estimated probability mass function $\hat{\mathbf{p}} = [\hat{p}_1, \hat{p}_2, \dots, \hat{p}_N]$ over the ordered classes $\{1, 2, \dots, N\}$, the EMD between $\mathbf{p}$ and $\hat{\mathbf{p}}$ is computed as:

$$EMD(\mathbf{p}, \hat{\mathbf{p}}) = \left(\frac{1}{N} \sum_{k=1}^N |\text{CDF}_{\mathbf{p}}(k) - \text{CDF}_{\hat{\mathbf{p}}}(k)|^2 \right)^{\frac{1}{2}} \quad (3)$$

where:

- $\text{CDF}_{\mathbf{p}}(k)$ and $\text{CDF}_{\hat{\mathbf{p}}}(k)$ are the cumulative distribution functions of $\mathbf{p}$ and $\hat{\mathbf{p}}$ at point k , respectively.
- The CDF at each point k is the cumulative sum of the distribution values up to and including k , i.e., $\text{CDF}_{\mathbf{p}}(k) = \sum_{i=1}^k p_i$ and similarly for $\hat{\mathbf{p}}$.
- The absolute difference between the CDFs at each point k is squared, summed over all k , and averaged over the total number of points N .
- Finally, the square root of this average is taken to obtain the EMD.

C. Dataset Preparation

It is to be noted that the existing IQA models are mostly trained and evaluated on publicly available datasets, for example TID2013 [11] or KADID-10k [12] databases; however, these differ significantly from microfluidic images. Therefore, in our work we used microfluidic droplet dataset¹ as well as approximately 1.9% data from Roboflow²; the dataset preparation involved several steps, i.e. collecting microfluidic droplet images and then categorizing them into ten distinct quality classes. The quality scale ranges from 1, representing the worst quality, to 10, representing the highest quality. A t-distributed Stochastic Neighbor Embedding (t-SNE) algorithm was used to visualize the class vote distributions for 917 images, depicted in Fig. 2. As shown in the figure, each image is associated with vote counts corresponding to the 10 different classes and each class is separable by distinct clusters with different colors and marker shapes.

Few data points are located farther from their respective clusters, which is due to the spread of vote distribution to the nearest classes.

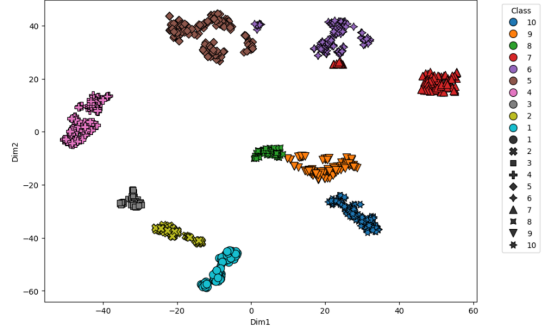


Fig. 2: t-SNE visualization of the latent representations of different classes.

To represent real-world conditions, three types of distortions were introduced: motion blur, gaussian blur, and contrast variation. Motion blur represents the blur due to droplet motion, Gaussian blur represents the uniform blur, and contrast variation emulates varying lighting conditions. Due to the limited availability of naturally distorted images, synthetic distortion techniques were applied to augment the dataset, ensuring sufficient variability for training the model.

Four independent human annotators labeled the images; they were provided with clear visual instructions representing the ten distinct categories of image quality. However, it was observed that with only four annotators, the resulting distribution of the voting lacked sufficient granularity to accurately reflect majority opinions on image quality. To achieve a robust and reliable distribution, it is recommended that the number of annotators be significantly higher than the number of classes.

Each distortion type was considered equally significant in the IQA. A maximization equation was applied to derive the final quality scores from three different distortions, ensuring a balanced evaluation across all distortion types. To expand the dataset and improve labeling efficiency, data augmentation and generation techniques were applied using the “Roboflow” online platform³. This strategy not only increased the dataset’s volume but also introduced variability, enhancing the model’s robustness. After augmentation, ambiguous and duplicate images were manually reviewed and removed, and any inaccurate labels were corrected to maintain the integrity of the dataset.

Finally, the dataset was split into training and validation sets with a 90:10 ratio, resulting in 917 training images and 101 validation images. This split ensures that the model has a sufficient amount of data for training while retaining a representative validation set to effectively assess its performance.

D. Implementation Setup

For implementation, we used TensorFlow with Keras version 2.10.0 on an Intel 12th Gen Core i9-12900K processor (3.2 GHz, 16 cores). After training, the model was converted

¹The authors thank Assistant Prof. Tomasz Kaminski, University of Warsaw, Poland, for providing us with data based on their microfluidic droplet setup [13]

²Available: <https://universe.roboflow.com/le-qgwow/droplets-avvwz>. [Accessed: Dec. 19, 2024]

³<https://roboflow.com/>

to TensorFlow Lite, i.e. a TensorFlow version optimized for mobile and embedded/edge devices to ensure compatibility and efficient performance on the RPI-5 single board computer.

E. Evaluation Metrics

For performance evaluation, we computed two widely used metrics: SRCC and PLCC. PLCC measures the linear correlation coefficient between the ground truth scores and the predicted scores, while SRCC measures the rank correlation between the ground truth scores and the predicted scores. Both SRCC and PLCC values range from 0 to 1; the closer the result is to 1, the better the prediction performance. We also computed EMD which measures the closeness of the ground truth and predicted rating distribution; the closer the value to 0, the better the performance. During training, the fine-tuned learning rate was 10^{-5} , and the batch size was set to 32.

III. RESULTS

We compare the proposed model with the modified version of MobileNet (inspired from NIMA) that has been fine-tuned specifically for the IQA application. The fine-tuning process began with the NIMA MobileNet, and we gradually adjusted the parameters to optimize its performance.

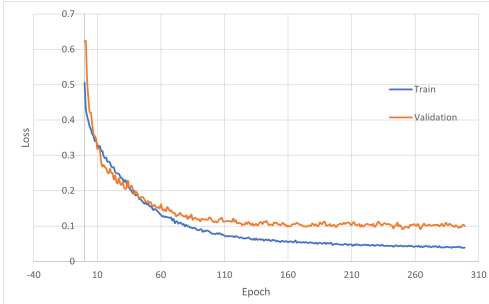


Fig. 3: Training and validation loss graph of modified MobileNet for predicting the quality of microfluidic droplet image.

As can be seen in Fig. 3, the modified MobileNet initially exhibited challenges in generalization, as indicated by the gap between the training and validation loss curves. Efforts to mitigate this issue through hyper-parameter tuning, leveraging transfer learning, as well as adding dropout regularization, did not yield a significant reduction in the generalization gap.

To address this, we hybridized the architecture by integrating a lightweight transformer block, which effectively minimized the gap. As a result, our proposed model performance improved, achieving an approximate 60% reduction (estimated at epoch 210) in the training-validation loss gap.

Indeed, Fig. 4 depicts that the validation loss follows the training loss and the loss remains almost unchanged after 110 epochs. This suggests that the training process converges around 110 epochs.

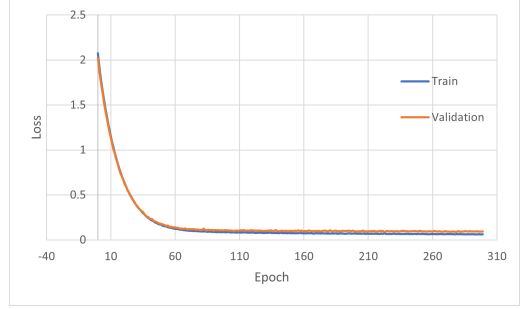


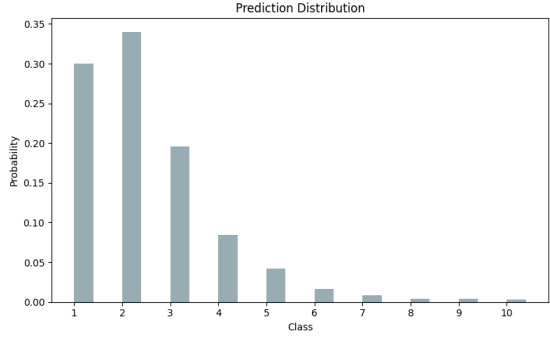
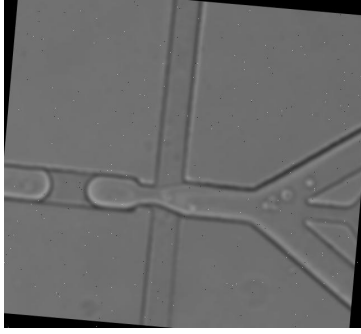
Fig. 4: Training and validation loss graph of proposed model for predicting the quality of microfluidic droplet image.

We then compare the complexity and performance of our model with several state-of-the-art works [14]–[18]. However, it is important to note that direct one-to-one comparisons are not entirely feasible because these works were trained on different types of images from publicly available datasets and are implemented on high-performance desktop PCs and/or GPUs. Nevertheless, this comparison provides valuable insights into the potential of our model in relation to existing methods. The results are presented in Table I and are discussed below.

TABLE I: Comparison of Model Complexity and Performance with State-of-the-Art IQA Models Based on SRCC, PLCC, and EMD

Model/Method	Parameters (in millions (M))	SRCC	PLCC	EMD
BRISQUE [14]	-	0.66	0.68	-
TTL-IQA [15]	-	0.86	0.84	-
AlexNet [16]	60	0.61	0.67	-
TReS [17]	152.45	0.88	0.86	-
HyperIQA [18]	27.37	0.86	0.84	-
Proposed	4.05	0.73	0.75	0.11

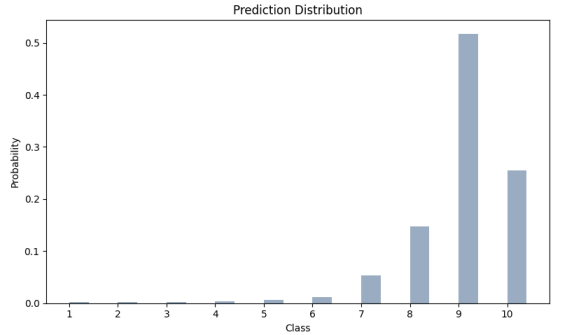
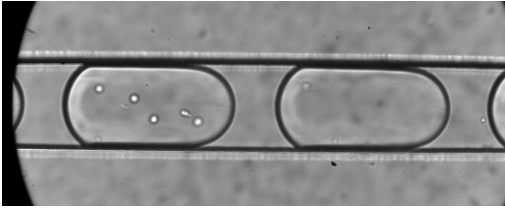
BRISQUE [14] is a NR-IQA method whose SRCC and PLCC values are 0.66 and 0.68 respectively (depicted in Table I), estimated using the KADID10K [12] database. Table I also shows that TTL-IQA [15], AlexNet [16], TReS [17] and HyperIQA [18] exhibit varying degrees of performance in TID2013 [11]; the most notable model, TReS [17], demonstrates superior performance with the highest SRCC and PLCC scores of 0.88 and 0.86, respectively, though it requires a substantial 152.45M parameters. In contrast, our proposed method, while achieving more moderate SRCC (0.73) and PLCC (0.75) scores, operates with significantly fewer parameters at just 4.05M, suggesting better computational efficiency. The other methods (TTL-IQA, AlexNet, and HyperNet) show varying degrees of performance, with SRCC values ranging from 0.61 to 0.86 and PLCC values from 0.67 to 0.84. Our proposed method is the only one to report an EMD score (0.11), potentially indicating considerable performance. Although an absolute direct comparison is not possible, the



(a)

(b)

Fig. 5: (a): Test image 1 (example of a poor image quality), and (b): Class prediction distribution of the image ($MOS \pm STD = 2.384 \pm 1.460$).



(a)

(b)

Fig. 6: (a): Test image 2 (example of a good image quality), and (b): Class prediction distribution of the image ($MOS \pm STD = 8.881 \pm 1.095$).

above suggests a trade-off between model complexity (parameters) and performance metrics, with the proposed method potentially offering a competitive result despite not achieving the highest accuracy scores.

Fig. 5 depicts the performance of the proposed model for a poor image quality (Fig. 5(a)) and corresponding prediction distribution (Fig. 5(b)) across the ten distinct classes. The horizontal axis denotes the class labels (ranging from 1 to 10), while the vertical axis indicates the corresponding probabilities assigned by the proposed model. Class 2 exhibits the highest probability, followed by class 1 and class 3, suggesting the model is most confident in classifying the image as lower categories. As expected, the decreasing probability values for the remaining classes (4 to 10) indicate lower confidence.

On the contrary, Fig. 6 depicts the performance of the proposed model for a good quality of image (Fig. 6(a)) and corresponding prediction distribution (Fig. 6(b)) across the ten distinct classes. The horizontal axis denotes the class labels (ranging from 1 to 10), while the vertical axis indicates the

corresponding probabilities assigned by the proposed model. As expected, the low probability values for the initial classes (1 to 7) indicate lower confidence. On the other hand, class 9 exhibits the highest probability, followed by class 8 and class 10, suggesting the model is most confident in classifying the image as higher categories, as expected.

Finally, we assess the memory consumption of the proposed model when deployed on the RPI-5 using batch size 1 and image resolution 224×224 , as depicted in Fig. 7. Out of the total memory usage of 56.39 MB, the modified MobileNet consumes 91% due to feature maps and the storage of activation functions across various layers. The transformer block is the second largest memory consumer, utilizing 8.26% of the 56.39 MB for global feature mapping. The remaining dense, GAP layers collectively consumes less than 1% of the 56.39 MB. The achievable inference time on the RPI-5 is 110 ms, translating to 9 frames per second (FPS). This deployment result with the RPI-5 demonstrates the suitability

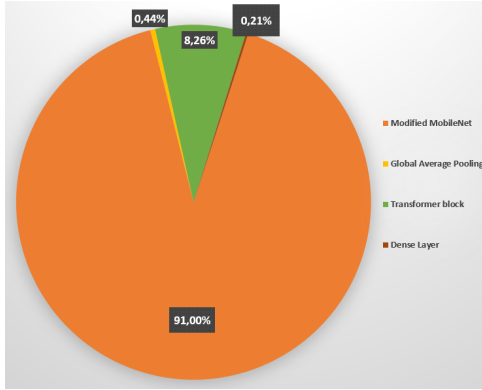


Fig. 7: Memory consumption of the proposed model for an input resolution 224×224 with a batch size 1.

of the proposed model for portable, resource constrained devices. The model could potentially be deployed on even smaller devices such as the MaixCam embedded board (1 GHz ARM A53 core and 256 MB on-chip RAM).

IV. CONCLUSION

In this work, three different CNN-based models (MobileNet, EfficientNet, NasNet) suitable for embedded platforms were tested. Transfer learning was used during the model training because of the limited dataset. Among the three different models, MobileNet demonstrated best performance. To mitigate the encountered overfitting issue, dropout regularization was applied but did not improve the results. We then hybridized the model by integrating a transformer block to enable the model to learn the global features and reduce the generalization issue. Additionally, early stopping was used to reduce the computational costs.

The version based solely on the modified MobileNet had a significant loss gap of 0.06, which indicates that the model has generalization issue. On the other hand, the proposed hybrid model (modified MobileNet and Transformer block) stabilizes much more rapidly at around 110 epochs. The spearman rank correlation coefficient is larger than 70%, indicating satisfactory performance. Inference time on RPI-5 is 110ms. These results pave the way for further research, which could involve further optimization of the model and architecture for better inference time, as well as the collection of more data to enhance the model's robustness. Furthermore, automated IQA could provide real-time feedback on image quality immediately after acquisition, enabling discarding of poor-quality images or adjustments to camera parameters to ensure high-quality of images.

ACKNOWLEDGMENT

This work was supported by the Estonian Research Council ETAg grant number PRG620. Additionally, the authors acknowledge support from the Estonian Education and Youth

Board project AIoT*5G grant number ÕÜF11. The authors thank Assistant Prof. Tomasz Kaminski, University of Warsaw, Poland, for providing us the corresponding data used to prepare the dataset to train the model for droplet image quality assessment.

REFERENCES

- [1] Lei, Cheng, et al. "High-throughput imaging flow cytometry by optofluidic time-stretch microscopy," *Nature protocols* 13.7 (2018): 1603-1631.
- [2] Srikanth, Sangam, S. Kumar Dubey, A. Javed, and S. Goel, "Droplet based microfluidics integrated with machine learning," *Sensors and Actuators A: Physical* 332 (2021): 113096.
- [3] Zhang, Fangzheng, et al. "Intelligent image de-blurring for imaging flow cytometry," *Cytometry Part A* 95.5 (2019): 549-554.
- [4] J. Kim, A. -D. Nguyen and S. Lee, "Deep CNN-Based Blind Image Quality Predictor," in *IEEE Transactions on Neural Networks and Learning Systems*, vol. 30, no. 1, pp. 11-24, Jan. 2019, doi: 10.1109/TNNLS.2018.2829819.
- [5] Wang, Zhiwen, J. Zhou, Q. Liu, L. Wang, and X. Su, "Multiplex imaging flow cytometry with content-aware image restoration deep learning enhances cell imaging quality," in *Fifteenth International Conference on Information Optics and Photonics (CIOP 2024)*, vol. 13418, pp. 872-878, SPIE, 2024.
- [6] Kurenkov, Andrey, A. Kussanova, and Natasha S. Barteneva, "Advancing precision single-cell analysis of red blood cells through semi-supervised deep learning using database of patients with post-COVID-19 syndrome," in *Imaging, manipulation, and analysis of biomolecules, cells, and tissues XXII*, vol. 12846, p. 1284602, SPIE, 2024.
- [7] Wang, Yiming, et al. "Real-time fluorescence imaging flow cytometry enabled by motion deblurring and deep learning algorithms," *Lab on a Chip* 23.16 (2023): 3615-3627.
- [8] O. Manzari, H. Ahmadi, H. Kashiani, S. Shokouhi, A. Ayatollahi, "MedViT: A robust vision transformer for generalized medical image classification," *Computers in Biology and Medicine*, vol. 157, 2023, 106791, doi: 10.1016/j.combiomed.2023.106791.
- [9] H. Talebi and P. Milanfar, "NIMA: Neural Image Assessment," in *IEEE Transactions on Image Processing*, vol. 27, no. 8, pp. 3998-4011, Aug. 2018, doi: 10.1109/TIP.2018.2831899.
- [10] L. Hou, C. Yu, and D. Samaras, "Squared earth mover's distance-based loss for training deep neural networks," *arXiv preprint arXiv:1611.05916* (2016).
- [11] N. Ponomarenko et al., "Color image database TID2013: Peculiarities and preliminary results," *European Workshop on Visual Information Processing (EUVIP)*, Paris, France, 2013, pp. 106-111.
- [12] H. Lin, V. Hosu and D. Saupe, "KADID-10k: A Large-scale Artificially Distorted IQA Database," *2019 Eleventh International Conference on Quality of Multimedia Experience (QoMEX)*, Berlin, Germany, 2019, pp. 1-3, doi: 10.1109/QoMEX.2019.8743252.
- [13] J. Jonghe, et al., "spinDrop: a Droplet Microfluidic Platform to Maximise Single-Cell Sequencing Information Content," *Nature communications*, 14(1), 4788, DOI: 10.1038/s41467-023-40322-w
- [14] A. Mittal, A. K. Moorthy and A. C. Bovik, "No-Reference Image Quality Assessment in the Spatial Domain," in *IEEE Transactions on Image Processing*, vol. 21, no. 12, pp. 4695-4708, Dec. 2012, doi: 10.1109/TIP.2012.2214050.
- [15] X. Yang, F. Li and H. Liu, "TTL-IQA: Transitive Transfer Learning Based No-Reference Image Quality Assessment," in *IEEE Transactions on Multimedia*, vol. 23, pp. 4326-4340, 2021, doi: 10.1109/TMM.2020.3040529.
- [16] A. Krizhevsky, I. Sutskever, and G. Hinton, "ImageNet classification with deep convolutional neural networks," *Commun. ACM* 60, 6 (June 2017), 84-90, doi:org/10.1145/3065386
- [17] Golestaneh, S. Alireza, S. Dadsetan, and K. Kitani, "No-reference image quality assessment via transformers, relative ranking, and self-consistency," in *Proceedings of the IEEE/CVF winter conference on applications of computer vision*, pp. 1220-1230, 2022.
- [18] S. Su, et al., "Blindly assess image quality in the wild guided by a self-adaptive hyper network," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 3667-3676, 2020.

Curriculum Vitae

1. Personal data

Name	Fariha Afrin
Date and place of birth	17 January 1991, Dhaka
Nationality	Bangladesh

2. Contact information

E-mail	fariha.afrin@taltech.ee
--------	-------------------------

3. Education

2020–Present	Tallinn University of Technology Thomas Johann Seebeck Department of Electronics, PhD
2018–2020	Tallinn University of Technology Communicative Electronics, MSc
2015–2017	American International University-Bangladesh Electrical & Electronic Engineering, MSc
2009–2013	International Islamic University Chittagong Electrical & Electronics Engineering, BSc

4. Language competence

Bengali	Native
English	Fluent
Estonian	Basic
Hindi	Basic

5. Professional employment

2020–2020	Tallinn University of Technology, Junior Researcher
2017–2018	Holoteq, Doha, Qatar, Team Lead
2014–2017	Infomedia, Dhaka, Bangladesh, Assistant Product Manager

Elulookirjeldus

1. Isikuandmed

Nimi	Fariha Afrin
Sünniaeg ja -koht	17 Jaanuar 1991, Dhaka
Kodakondsus	Bangladesh

2. Kontaktandmed

E-post	fariha.afrin@taltech.ee
--------	-------------------------

3. Hariduskäik

2020–praeguseni	Tallinna Tehnikaülikool, PhD
2018–2020	Tallinna Tehnikaülikool Kommunikatiivne elektroonika, MSc
2015–2017	Ameerika rahvusvaheline ülikool - Bangladesh Elektri- ja elektroonikatehnika, MSc
2009–2013	Chittagongi rahvusvaheline islamiülikool Elektri- ja elektroonikatehnika, BSc

4. Keelteoskus

Bengali keel	Emakeel
Inglise Keel	kõrgtase
Eesti Keel	Põhiline
Hindi keel	Põhiline

5. Teenistuskäik

2020–2020	Tallinna Tehnikaülikool, Nooremteadur
2017–2018	Holoteq, Doha, Katar, Meeskonna juht
2014–2017	Infomedia, Dhaka, Bangladesh, Tootejuhi assistent

ISSN 2585-6901 (PDF)
ISBN 978-9916-80-365-3 (PDF)