Biology and Computation for Improving Access to Health Screening

Jason Stuart Hoffman

A dissertation

submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

University of Washington

2025

Reading Committee:

Shwetak N. Patel, Chair

Chris Thachuk

Luis Ceze

Program Authorized to Offer Degree:

Computer Science and Engineering

University of Washington

Abstract

Biology and Computation for Improving Access to Health Screening

Jason Stuart Hoffman

Chair of the Supervisory Committee:

Shwetak N. Patel

Computer Science and Engineering

Many patients who need medical care are underserved by today's systems, leaving them unable to get adequate warning when health conditions require medical attention. Computing technology can aid in reducing this healthcare gap by commoditizing sensors capable of providing diagnostic and prognostic information in an affordable and usable form factor. Ubiquitous sensing technologies, such as those built into modern smartphones, can be considered more accessible or affordable because the hardware can be re-used for many purposes, effectively amortizing the cost of each function over the full utility of the device. However, the generalized purpose designed into these sensors introduces unintended noise into signals which otherwise contain useful health information. Creatively designed data gathering techniques that leverage chemical and medical understanding of the underlying biology and physiology can detect signals that can be useful for health screening. In this thesis, I demonstrate these techniques in four projects over two areas: (1) blood-oxygen saturation and (2) hemoglobin concentration for health sensing in vivo using an unmodified smartphone camera, and (3) passively gathering samples of viral particles and (4) capacitance-based detection of specific nucleic acids for environment sensing in vitro. In order for technology to deliver on the promise to aid in healthcare, a physiological understanding of the molecular components can be combined with computational techniques, such as machine learning and signal processing, to reveal the underlying signals of interest, improving access to health screening for the growing proportion of the population with smartphones.

Acknowledgments

My life's driving force has been to have a positive impact in the world by "paying forward" the privilege I have been afforded. To this day, I do not know from where exactly this force has emanated, but I must acknowledge the impact of my community, family, friends, and teachers on my dreams and aspirations. The very nature of privilege means that people and systems have conspired to give me a chance to achieve my dreams, and thus I am grateful for the space here to acknowledge their past and continued impact on my life and my appreciation for their care. I have so many people in my life to thank for me getting to this point, and I hope I can describe them all, though their impact is so woven into my daily life, that it's sometimes indistinguishable from it. I would not have been able to follow my dreams without the people around me who love me and support me, and I am appreciative of their dedication on my behalf. Since this is an academic document, I want to begin by acknowledging my teachers, coaches, mentors, and friends in the academic space, leading from my early years to today. As I look back in time, various influences weave their way into my life, including my colleagues at Microsoft, professors at USC, teachers in elementary through high school, and my friends and family along the way.

Most recently, my research advisors and professors at UW and beyond have enabled this next step that turned me into a researcher who could produce this thesis. Shwetak Patel, my main advisor and PI of the Ubicomp Lab, has created an incredibly vibrant, welcoming space in his lab on the top of the Paul G. Allen Center for Computer Science and Engineering. I had met a few lab members in 2017 and 2018 that led me to realize this would be a very special place to work and pursue further education. I'd like to give a special shout-out to Lillian de Greef and Morelle Arian, who introduced me to the Ubicomp Lab and invited me to lab lunches before I was a student at UW. This allowed me to meet mentors like Edward Wang, Alex Mariakakis, Tien-jui Lee, and CJ Park. Each of you gave of your time and energy to help me learn and grow, and I will be forever grateful. In addition, I would like to thank Gabe Cohn, Sidhant Gupta, and Desney Tan for encouraging me, as a coworker, to make the right decision (for me, in hindsight) to leave Microsoft to pursue my PhD in such a wonderful environment. I would not be here without any of you, and all of you deserve credit for creating the special culture we have in Ubicomp Lab and for ultimately granting me access to this space that has changed my life. This is not to mention the current members of the Ubicomp Lab who have continued to evolve the culture and keep producing wonderful research in a collaborative and encouraging environment. Thank you for all the good times and the collaboration opportunities.

At UW, I encountered another special group of folks in the Molecular Information Systems Lab (MISL). Beginning with Georg Seelig, moving through Luis Ceze and Jeff Nivala, and finishing with Chris Thachuk, I've been fortunate to work with nearly every advisor in the MISL lab as I attended UW CSE through a time of transitions. From lunch meetings to Zoom meetings and finally back to lunch meetings, I have learned so much from you and all our fellow students about biology and different ways of thinking about computing. Thank you also to David Ward and Gwen Roote for managing the lab space, as well as Matthew Hirano and Zoe Derauf for helping all of us by managing the lab spaces in the Nano Engineering Sciences building, where I did a lot of my wetlab work. We could not do the research we do without your efforts behind the

scenes, on top of all the research you contribute to. The fellow students and postdoctoral scholars in MISL Lab have created that special environment where we can work on cutting edge biotechnology and apply computing principles. Finally, other labs and their members, especially the ICTD Lab and the Security and Privacy Lab and more generally the various HCI labs in CSE, have welcomed me into their spaces and seminars and had a profound impact on the way I think about research, and I appreciate your openness and friendship.

All my labmates have impacted me in unique ways in both the MISL and Ubicomp Labs, but I have had the privilege to work even more closely with a few. It was a pleasure to work on significant projects and co-author with Anandghan Waghmare, Varun Viswanath, Matthew Hirano, and Daniel Gordon, without whom this thesis would not be possible. My medical collaborators, like Dr. Matthew Thompson, Dr. Margaret Rosenthal, Dr. David Suskind, Dr. Kirk Beach, and Dr. Lorena Casas-Wright, also provided incredibly important insights and perspectives, and I look forward to continuing to work with you if I am fortunate enough to do so. I'd also like to give special thanks to all the undergraduate researchers who entrusted me with their research experience, especially folks who worked on their theses or other significant works with me, including Bo Liu, Ellen Xu, Caiwei Tian, Zoe Derauf, Hannah Lee, and Derek Zhu. I look forward to seeing where your careers take you, and I hope you maintain the optimism and curiosity you demonstrated in our collaborations.

That leads me to a group who supported me in research and other steps in my career when I was an undergraduate in Biomedical Engineering at USC and a fresh young Program Manager at Microsoft. When I was just learning about the magic of research, I was lucky to encounter Heidi Gensler during her PhD work, and her continued investment in my growth over a period of years left an impact on me that I sought to pass on to my mentees when I finally started in my PhD. My professors at USC, especially my recommendation letter-writers who still remembered me five years later when I showed up at USC and told them I wanted to re-enter academia, enabled me to take this unique career path. This includes professors who taught me in classes that fascinated me, including Physics with Gene Bickers, Digital Logic with Mark Redekopp, and Biomedical Engineering Senior Projects with Jean-Michel Maarek. This also, of course, includes my mentors who supported me in research at USC and at Microsoft, including in the USC BioMEMS Lab with Ellis Meng and the Surface Hub team (including the Perceptive Pixel startup team that was integrating into Microsoft) with Jeff Han. My managers, mentors, and coworkers at Microsoft, such as Rob Huryn, Dave Kearney, Ryan Rutter, Siri Velauthipillai, and Martin Hall, harnessed my passions and gifted me with responsibilities I was lucky to encounter as a 23-year-old. I still apply many of these lessons today as a researcher with industry experience. I also want to thank all my coworkers from that time, but especially Andy Hurley, Jenny Chang, and Ali Hajy for forming closer bonds of friendship that lasted beyond employment. As I enter this next stage in the intersection between academia and industry, other managers and collaborators during graduate school come to mind, including Matt Whitehill, Michael Brandwein, and Laura Dorsey, for paving paths and providing new opportunities that, I hope, will result in impactful work that helps people long-term.

In other thanks, I want to call out to champions of educational outreach within the Allen School who, like me, believe that our duty at a public institution of higher education goes beyond the boundaries of campus and into the community that surrounds us. Anyone who joined me in the efforts to promote more outreach by PhD students in the Allen School, especially the leaders - Hannah Potter, Jialin Li, Vicente Arroyos, Jerry Cao, Fernanda Jardim, Chloe Dolese-Mandeville, Jeremy Munroe, Juliet Quebatay, and Lauren Bricker - deserve special thanks for enabling more PhD students to connect with K-12 students in local schools. In addition, I want to shout out to Vicente Arroyos and Kyle Johnson, who started AVELA as undergrads and continue to provide an admirable model of what it means to spread the benefits of higher education to a wide

range of folks in the community. I aspire for my work to have as much impact as theirs. This also leads me more broadly to think about those who collaborated with me on the Anti-Racism Proposals we wrote in 2020 that brought many of these issues to the attention of UW CSE admininistration, which eventually influenced many of the aspects of the UW DEI strategic plan. This unsung work is not often recognized, but when I think back to my first year, when I was trying to pretend that I wasn't scared about potential repercussions of speaking up, I can be thankful that I found the right allies at UW, and I hope that UW continues to be a place that values different opinions in uncertain times. On that note, I also want to acknowledge the impact of the work of all our student union organizers, as another thankless task that sets up the infrastructure for taking action that supports positive change. I've learned a lot from various people throughout my time at UW, opened up to new perspectives, and supported in ways that I can, and I've also found analogies from here to all the forms of systemic oppression folks face. My research work and personal perspective echo from that age-old struggle. In the US capitalistic system, we can't benefit from these healthcare innovations unless we have support through employers, and they are not always aligned with our interests, so it's organizations like unions that will continue providing those benefits until our system changes.

Community is very important to me, and I've been lucky to establish myself in such a vibrant one in Seattle. I was lucky to be in a somewhat unique position to transfer to graduate school within a city where I had already lived for five years. However, my community starts from my home base in Kansas City where I grew up and formed a strong sense of the fierce Midwest friendliness that I value in myself and friends. This extends from my wonderful teachers at a fantastic public institution, Blue Valley North High School, and the feeder elementary and middle schools that formed my foundation in K-12 education. Specifically, teachers like Michelle Buche (then Radio), Deb Hotujac, Teresa Hogan, and Mark Chonko were especially encouraging of my curiosity and tenacity to learn. All my classmates along the way also aided in creating a wonderful environment to thrive and grow from a curious little boy who loved reading Harry Potter and Lord of the Rings to someone who now feels comfortable seriously discussing the finer points of non-fiction books in social justice book club and academic discussions. I value the continued friendship of my USC community around the world, as well, including Jared, Max, Steve, Marisa, Schessa, Claire, Nick, Terrence, Gabby, Robin, and others. They have known me during years of enormous growth and change, and I appreciate that they still want to hang out when I pop into town. In Seattle, my community of friends has helped me grow into the person I am today, from ultimate frisbee teams, soccer teams, and regular meet-ups at Chuck's for the past ten years. Of course, we have fun and general hilarity with going to city spots and events, including camping, tournaments, festivals, pride and other events, driving along the great tree-lined highways into the mountains of the Pacific Northwest. From the ultimate frisbee team, Bullet Train, which gave me a fast, warm community those early few years in Seattle to "Social Justice Book Club" where I get to bridge the academic side of my life with my frisbee friends and stretch my brain with new perspectives, I have strengthened relationships I truly value along the way. Some special friends who have been with me virtually all the way since I moved to Seattle include Maggie, Andy, Alex, Erik, Brian, Lena, Iris, Patrick, Tamara, Aldan, Alicia, Jeremy, Michael, Tess, Joe, Leah, and many more. From strengthening our friendships (and our muscles at pod workouts and lifting sessions), as well as being a sounding board for each other through stresses and daily life, people who intersect multiple areas of my life understand me in a more in-depth way than anyone else, and you are all special and critical in pushing me to be the best version of myself I can be.

At this juncture, I would be remiss to ignore the undeniable and irreplaceable effect of closer life partners as we journey through this world. Annie and I conquered USC Biomedical Engineering together alongside our vibrant ultimate frisbee community. Joelle and I survived the COVID-19 pandemic together, completing our intimate pandemic pod of care and fun with our good friends Maria and Grace. Miranda and I learned

about life together and opened each other's minds to new ways of thinking and living while she was in Seattle. Each have had their lasting impacts on me and my perspective, and brought enormous joy and shared growth opportunities. Relationships, especially romantic ones, are dynamic, and some are meant to come and go for different times in our lives, while others are meant to endure. Thus, I am extremely grateful for my new partner, Zoe Geiger, who has become a steadying force in my life. I wouldn't have been able to finish out this PhD if not for the lucky occurrence of our encounter and budding new love. Zoe and her family, and the warmth with which they have received me, is strengthening my roots in Seattle and giving me hope for a future of adventures, surmounting challenges, and living a happy life that brings more joy into the world for everyone.

Finally, I wouldn't be anywhere without the foundation of my family. It often pains me that we're so far apart and I can't fully bridge the distance all the time, but I appreciate all the times we do get to spend together and the opportunities you have afforded me to go find a place in the upper-left United States where I get to live my fullest life and all the opportunities it affords, one of which is this thesis and my PhD work and future career. My aunts, uncles, cousins, and dearly beloved late grandparents, in Kansas City, Detroit, Colorado, New Jersey, Florida, and California, all deserve thanks for being the best family I could ask for, though a few deserve special thanks in the context of my career journey. My Aunt Deb Glueck and Uncle Joe Hoffman (and their children, my cousins Sam, Becca, and Abe), enabled my love of skiing and the outdoors, and did not succeed in convincing me not to pursue a PhD while giving me an example of what academic life could be like from within my extended family. Also, my Uncle Jeff Frankenstein has always been an inspiration to me, as despite also consistently telling me to do things differently than he did, his efforts to pique my curiosity at every turn - including showing me how to code my first calculator using Power Builder on Windows 95 - first opened my eyes to the magic of learning the language of computing. More immediately, my brother Michael is the only person in the world who uniquely knows what it's like to be in my family, and we share a bond and an understanding that I'm grateful for when we talk for hours on the phone or give each other knowing glances when a shared memory is triggered. We care for each other deeply, and we'll be in this journey of life together for the long term. My father Jim gave me his calm, analytical perspective and passion for his family, not only hereditarily, but also through the way he demonstrated these values and how he holds himself in good times and bad. He holds our family together tightly and never hesitates to express that he's proud of me, and it's with this firm base that I can move forward. My mother Michelle is an unwavering support and it's no wonder I can feel her love and care across continents. Even when she is (subtly or not) arguing against something that I know is best for me, I know it comes from a place of love and I always begrudgingly take it into account. I know it's not her ultimate preference to have me living so far away from home, but I also believe she understands that sometimes you have to loosen your grip on the things you care about in order to see them fluorish to their wildest dreams. The significant medical bent to my career stems greatly from my dad's position in electronic health records at Cerner for 25 years, and my mom's constant attention she pays to all of our medical needs, and even though they didn't explicitly tell me to pay attention to that, I believe that this subtly impacted my thinking as I made decisions in my career. I know this thesis and the next steps I take in my career will continue to make them proud.

To all my teachers,
for creating a safe space where my curiosity could only grow
and for leading me down a captivating trail I couldn't help but follow.

Contents

List of Figures

List of Tables

Chapter 1

Introduction and Motivation

1.1 Introduction: The Potential of Ubiquitous Computing

A large portion of the potential for the computing revolution to improve healthcare remains unrealized. The Electronic Health Record (EHR), remote "telehealth" interaction with doctors, and a plethora of highly accurate purpose-built diagnostics and treatment systems have made significant strides towards improving people's opportunities to access the medical information they need and receive care for their conditions. However, too many of the solutions are gate-kept behind barriers to access, such as high cost of care and low availability of appointments with the people trained to use these solutions (namely, health care practitioners). General-purpose sensor systems, leveraging concepts of ubiquitous computing, show promise in enabling anyone with a smartphone to access care more readily at their convenience with devices they already own (namely, smartphones). Unfortunately, though, many of these promises remain unfulfilled despite enormous growth in academic publications and industrial investment in the area. Great ideas turn into research projects but stagnate when met with the considerable validation cost and practical concerns associated with implementation in healthcare settings.

1.1.1 Problem Statement

Specifically, many patients who need healthcare are underserved by today's medical systems. The reasons for this include, but are not limited to, our limits of understanding medicine and biology, progress on technological development, ethical concerns, accuracy compared to the current standard of care, and practical cost considerations. My thesis illustrates my attempts to make progress in addressing the latter issues of healthcare access; namely, reducing the cost of obtaining information about the body's physiological state for diagnostic purposes. Doing so has involved leveraging current understanding of biology and ubiquitous technologies to create accurate diagnostics that near the accuracy required by medical standards and available on devices that people already own with no, or at least minimal, modifications.

1.1.2 Thesis Statement

My research projects, and the associated learnings, have led me to posit the following thesis statement:

Computational techniques, like machine learning, can be combined with physiological knowledge to

reveal the clinically-relevant signal out of the noise inherent in ubiquitous commodity sensors.

This thesis is pictorially summarized in an illustration in Figure 1.1, representing the ways we try to amplify a signal from a molecule of interest to be detectable by sensors that are already ubiquitously available in the computing world.

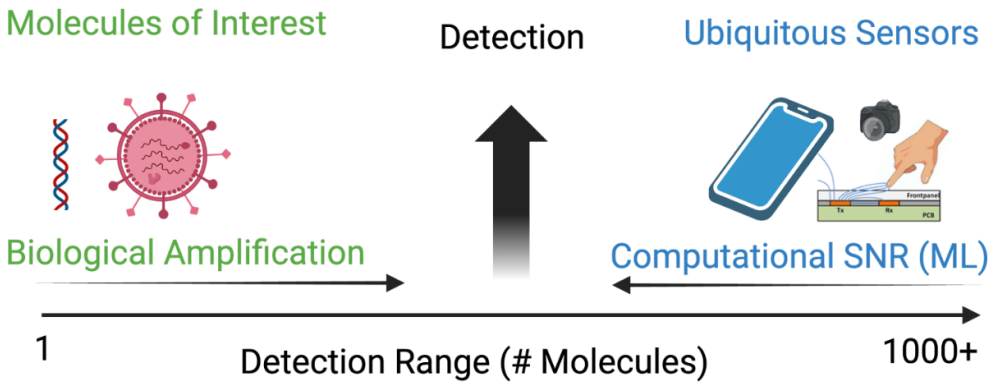


Figure 1.1: The availability of molecular detection can increase if we use biological knowledge to amplify the signal from small molecules and also computational techniques to sort through the noise inherent in commodity sensors.

1.1.3 Research Questions

In this dissertation, I explore the facets of this thesis in four projects motivated by the following research questions:

1. What are important aspects of data collection for accurate-enough results for a clinically relevant health diagnostic, given a physiological explanation for why computational methods like machine learning may work?
2. What computational techniques can we use to improve accuracy if it is difficult to gather a clinically-relevant spread of training data?
3. How can we passively leverage the built environment to reveal data for public health?
4. How can we more accessible bring clinical lab testing approaches to the field and bring the computing world closer to molecular data?

1.2 Motivation: Improving Access to Healthcare

Today, with a smartphone, one can perform a large portion of their banking, communication, transportation, socialization, shopping, and other related needs. Some of these tasks can be completed using only the phone, but most also connect with existing infrastructure over the internet or leverage systems, such as banking databases and medical systems, to accomplish things more conveniently, faster, or more accurately. Leveraging the fact that a user already owns a smartphone to do everything else, the field of ubiquitous computing demonstrates that the marginal cost of doing another thing with the same device is low. The

purported result is that doing that new thing is more affordable and thus more accessible. This opens the door to more convenient forms of healthcare, such as telehealth scenarios [Wosik et al., 2020; Mahtta et al., 2021].

On the medical front, medical professionals and systems are increasingly over-burdened, resulting in ballooning wait times, reduced quality of care, and increased costs. These costs can be seen from both an economical (high insurance or copay rates) and physical (worsened condition due to delays in treatment) perspective [Duong and Vogel, 2023; Mæstad et al., 2010]. This is especially true in non-specialty areas that serve a wide variety of patients, such as Family, Internal, or Emergency Medicine. So, what are the blockers that are preventing us from leveraging the promises of ubiquitous technologies to provide high-quality care to all? Incentive structures for how medicine is developed and delivered play a large role. This becomes apparent in the ways insurance payer systems can be misaligned from the patients they serve, the medical research developments that are funded based on their potential for future profit, and even in the incentives for medical students who are picking their specialty [Rosenthal, 2017]. Incentive structures and high-quality purpose-built medical diagnostics and treatments will continue to be researched and developed, as their applications often have clear incentive structures for benefiting patients and inventors in the end. However, the incentive structure for developing lower-cost care options is not always as clear, even though it could help many people.

Improving in all of these areas could help to ameliorate these problems, but I have chosen to pursue solutions by building technology that is enabled by computational techniques that attempts to fill the gap in providing quality care at a low price point. In order to target these computational solutions for the widest possible audience, I have focused on solutions that can be used by people with smartphones, which make up a wide and growing proportion of the population. Many projects in the Ubicomp community and beyond have used smartphones to detect various health markers [Wang et al., 2016, 2017a; Mannino et al., 2018a; Bui et al., 2020a; Ding et al., 2019]. In addition, medical care comes in many forms, including systems where people may not have access to smartphones, but these often include Community Health Workers (CHWs) in the health system. These care workers are increasingly equipped with smartphones, while they do not necessarily have access to a large array of medical equipment, especially when providing care via home visits [Yadav et al., 2021; Brown and Mickelson, 2019; Organization et al., 2020]. Therefore, a diagnostic developed to perform with good accuracy on a wide range of smartphones would increase access for the widest possible set of people.

A specific example can be illustrative, such as the case of sensing blood oxygen saturation (or SpO_2), which I will describe in further detail with my project in Chapter 3. SpO_2 is a common measure taken at most preventative care visits and can indicate underlying respiratory problems, and there are 3 distinct categories of diagnostic methods available. The most accurate is an Arterial Blood Gas (ABG) reading of the amount of oxygen in a sample of blood using an ABG machine (gray box in Figure 1.2). Another category, which is slightly less accurate but is more convenient, is the pulse oximeter, which uses light absorption to compute the proportion of hemoglobin molecules in your blood vessels carrying oxygen during the reading (yellow box in Figure 1.2). Third is the category which I have focused on with my research, which is smartphone-based SpO_2 readings (blue box in Figure 1.2). These readings may not provide as much sensitivity and specificity, but I argue that they have enough utility to be helpful in medical care in triage and screening. By utility, I mean to include the concept of sensitivity, while also factoring in additional important aspects of using the diagnostic, including availability, portability, affordability, and general ability to access the diagnostic. This chart requires a somewhat estimative form of dimensionality reduction in order to explain, but it is illustrative of the value of the smartphone-based approaches I have pursued in my research.

So, given all this, why are our general-purpose sensors in smartphones not used for more of healthcare?

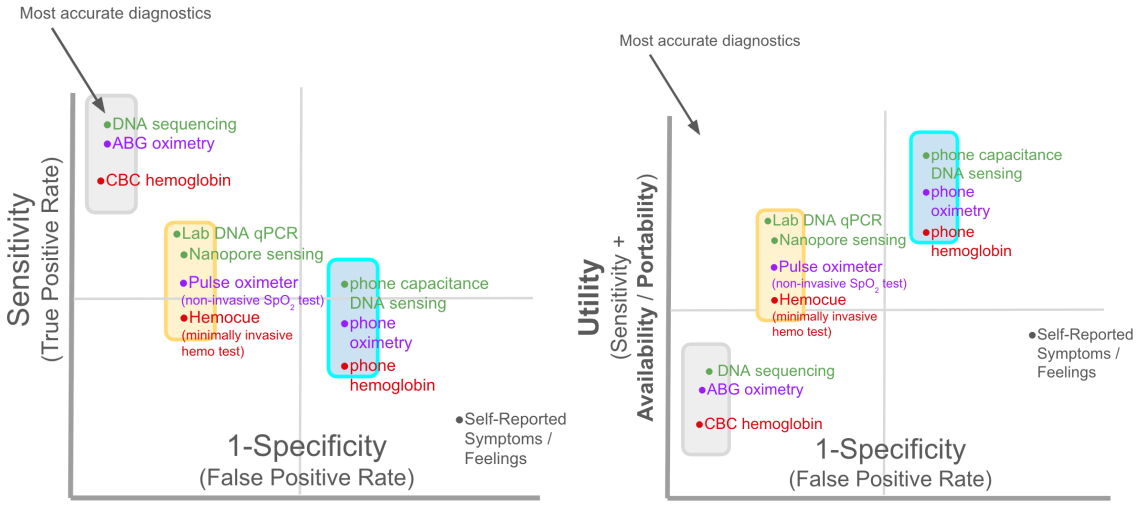


Figure 1.2: The projects in my thesis focusing on moving diagnostics up the utility scale while remaining accurate enough for use. The most accurate diagnostics appear in the top left of each graph, with the graph on the right incorporating the concept of utility to illustrate the value of smartphone-based diagnostic systems described in this thesis.

As mentioned previously, there are factors that limit our understanding of biology, thus limiting our understanding of disease itself. Much research is being done to work towards understanding disease, but I also see a gap in our ability to widely diagnose diseases in ways we already understand. I pursued contributing to answering these questions throughout my PhD. I focused on metrics often used as primary screening for infectious disease severity or those that are treatable, like hemoglobin as a metric for diet-based anemia [Hoffman et al., 2022b; Mannino et al., 2018a]. And finally, in addition to leveraging smartphones for direct diagnostic tests, I sought to bring more advanced testing, such as that performed by specific viral presence detection tests, to wider use via linking passive sensors to smartphone-based detection devices for use in the field by a variety of people [Hoffman et al., 2022a].

1.3 Methodology: Biological Knowledge Informs Computational Methods

In order to investigate these research questions and make progress on proving this thesis, I investigated two projects that demonstrated that computational techniques could be used to sort through the noise and two projects that delved deeper into how we sample and understand signals produced by biological molecules. I pushed the boundaries on both sides of my thesis illustration in Figure 1.1 in order to better understand the full picture of what we can do to leverage ubiquitous sensors for sensing health-relevant biomarkers.

Through these projects, a few principles emerged. (1) A deep physiological understanding of the underlying molecules was needed in order to design systems to detect them. Oftentimes, prior techniques were leveraged and tweaked to fit the needs of widely available sensors on "ubiquitous" devices. (2) When ubiquitous sensors were used, the general-purpose nature of the sensors caused more noise that was present in purpose-built devices, requiring the use of computational techniques to parse out the relevant signal. In a few of these projects, the main technique was machine learning, but other techniques, such as signal processing or simple thresholding, were sufficient for others. Finally, (3) when the in-built sensors in ubiquitous devices were not sufficient for accurately reading the health-relevant signal of interest, a minimal attachment

Category	Project	Chapter
Health Screening	Smartphone SpO_2 Sensing (SpO_2)	3
in vivo	Smartphone Hemoglobin Sensing (Hgb)	4
Environmental Detection	Passive Viral RNA Detection	5
in vitro	Smartphone DNA Capacitance Sensing	6

Table 1.1: Summary of Projects

was necessary to adapt the sensor to be able to gather the relevant data, which allows practical detection for a low cost. The accuracy gained compared to the cost of the attachment was considered in motivating its inclusion in the solution.

My thesis will describe the projects, the findings, and their implications for attempting to answer the core question of how to use engineering to increase the number of people who have access to healthcare. These answers solve a small subset of all the problems that prevent a wider swath of people from accessing healthcare that can support their quality of life and longevity, but some of the lessons can apply to a wider set of those solutions, as well.

1.4 Contributions: Research Projects in Computational Health

In all, my thesis contributes the following four projects as an answer to the research questions posed previously. Together, these demonstrate a few health-relevant biological analytes (oxegenated hemoglobin, total hemoglobin, and presence of nucleic acid sequences) which may be detectable using software on a modern smartphone with zero hardware modifications or attachments. In this section is a summary of those contributions to set the stage for the following chapters, and a more full-featured summary of the work can be seen in Appendix A.

Project 1. SpO2 sensing (in vivo) using unmodified smartphone via PPG analysis and ML

We showed that we could somewhat accurately predict blood-oxygen saturation (SpO_2) over a medically relevant range of healthy and ill patients. We used a deep learning CNN model and custom software app on the ISO standard clinical development test which artificially lowers the subjects' blood oxygen saturation down to 70%. We achieved an MAE=5.0%, while the requirement for FDA-approved devices based on the standard is 3.5%. For screening purposes, we also demonstrated a simultaneous sensitivity and sensitivitiy of 80% at predicting a person was desaturated [Hoffman et al., 2022b].

In my article on smartphone camera oximetry, I collaborated with my co-authors to gather PPG data for pulse oximetry using a smartphone camera on a clinically relevant range of SpO_2 levels for the first time (70%-100%). We used a technique called a varied fractional inspired oxygen (ViO_2) technique in order to gather this data safely. We then applied a deep learning model to result in an overall mean average error (MAE) of 5.0% SpO_2. This compares favorably to industry standards for commercial pulse oximeters, which require an accuracy of less than 3.5% for approved reflectance PPG devices. We also compared this result to a traditional fitted ratio-of-ratios model, and found that deep learning performed better in inferring the blood oxygen saturation for new subjects. We believe this is due to the ability of deep learning models to recognize patterns and focus on the signal despite the noise introduced by the commodity sensors. We also applied thresholding on the final result to give the user an 80% sensitivity and specificity for determining if they were below a healthy range of blood oxygen saturation, indicating its usefulness in health screening

scenarios. Finally, we contributed to the literature by sharing the dataset and starter code open source with the community, and others have begun to build upon our work.

Project 2. Hemoglobin sensing (in vivo) using unmodified smartphone on color change in fingernails

We demonstrated acccurately sensing hemoglobin using an unmodified smartphone. Based on recently published research, an ML model should be able to correlate color of the blood under the fingernail (captured using an unmodified smartphone) with the hemoglobin level [Gordon et al., 2024].

I also applied similar ideas to a project to detect hemoglobin concentration. Inspired by prior efforts to detect hemoglobin noninvasively using a smartphone [Wang et al., 2016; Mannino et al., 2018a], I and my collaborators understood that hemoglobin levels impact the red tone of the blood, and we attempted to gather data objectively via a smartphone camera. We gathered data on a large number of users, but were not able to achieve our goals of collecting data from people displaying a wide range of ground truth hemoglobin levels and skin tones. Therefore, when we applied machine learning techniques to infer blood hemoglobin levels, we were not able to achieve highly accurate results. Using some techniques for augmenting the dataset improved those results, indicating that this problem may be tractable by gathering the right data, but overally we furthered the hypothesis that the right data is needed to be gathered in order to make an accurate system.

Project 3. Passive viral fabric sensors (in vitro) of SARS-CoV-2 RNA in public buses

We showed that cheap, disposable sensors made of fabric could be used to passively collect evidence of airborne disease and then brought to the lab for detection. We combined previously used protocols to extract viral particles, kill and break them open to release the RNA, and then sense them using qPCR [Hoffman et al., 2022a].

The goal of the SARS-CoV-2 project was to develop a passive sensing system for gathering data on COVID spread using the in-built HVAC systems in public transit buses. We understood the sample of interest (SARS-CoV-2 RNA) and the gold standard method of detection, and we developed a protocol for collecting those samples and transporting them back to lab. We detected traces of SARS-CoV-2 RNA on 14% of buses during the pandemic timeframe, but the procedure was cost- and time- prohibitive to roll out on a larger scale, indicating the need for more accessible field-based methods of detection.

Project 4. Nucleic acid reaction detection (in vitro) for in-field molecular-computer interfacing using the ubiquitous capacitive touchscreen of smartphones

In the interest of developing a way for the previous viral sensing method to scale and be more easily accessed outside of the lab, we tested aspects of a phone's capacitive touchscreen sensor as a detector for reaction output. While being theoretically accessible to anyone with a smartphone, this solution would also provide a direct molecular->computer interface to allow this data to be immediately connected to all the resources in a modern computing device. This would enable automatic connection to online geographically-tagged tracking databases, other phone sensors like location, time, and user, and processing with on-device computational techniques including ML. Previous works have shown disease diagnostics using molecular computing [Lopez et al., 2018]. This work builds on prior work on DNA capacitive sensing to envision and begin to enable a future where nucleic acid diagnostics can be performed on anyone's phone [Lee et al., 2018; Stagni et al., 2006].

Contribution

Overall, with my work I hope to illuminate a path for increasing access to healthcare, attempting to begin to ameliorate the growing healthcare divide with this work and my future career.

Chapter 2

Background and Related Work

In this chapter, I lay the foundation for the field, myself, and this overall thesis. This document's structure, my personal positionality as a researcher and engineer, and my understanding of the latest research in the field impact the way I interpret my results and can provide helpful background for understanding my results and conclusions.

2.1 Background

2.1.1 Document Structure

This dissertation document contains a full set of my research and work that relates to this thesis statement, conducted during my PhD studies at the University of Washington under Dr. Shwetak Patel from 2019-2025. It does not contain an exhaustive set of all the work I have done in this area, nor all the research I conducted during this time. However, I hope it is educational, and this section will help the reader get their bearings among this much text.

The first two chapters of the dissertation, including the previous chapter and the current chapter, are the Introduction and the Related Work, which establish the overall basis, motivation, context, and background for this research work as it stands in 2025. This is an evolving field, and thus these represent point-in-time results from this period. I expect this field to continue evolving, and I hope that others expand beyond this work in the future. Following these introductory sections are in-depth chapters for individual projects that contribute to the overall thesis. Each chapter contains a background section at the beginning that connects the project findings to the overall thesis and a conclusion section at the end that summarizes the learnings and how the current project relates to the others in contributing to the thesis. These background and conclusion sections are also copied into Appendix A near the end of the thesis in order, which allows the reader who is already familiar with my published works to read through the full narrative of the thesis in order, which may help with getting a better flavor for the thesis overall. Finally, the last chapter is dedicated to conclusions and reflections that tie together the overall ideas and paint a picture of where I see this field evolving in the future.

2.1.2 Positionality Statement

My personal position as a budding engineer and researcher has influenced my work a great deal. I acknowedge that there are many ways by which to pursue improvement to healthcare, and I have chosen a technical angle for a variety of reasons, both societal and personal. I find the projects fascinating, gratifying

to complete, and I'm hopeful that they will have impact. At the same time, I have to acknowledge that my perspective is influenced by the fact that today's society compels each of us to build "employable skills" through experience, which has thus influenced me to pick a career that has potential to continue to provide economic value to society for which I can be compensated. I do not believe that all health problems can be solved via technology; in fact, many problems cannot, and most are solved by a combination of technology alongside political and human resources. My cultural background, as a white, Jewish, male-presenting descendant of European immigrants, influences my choices on a daily basis and I must remain cognizant of this bias alongside any conclusions I might draw from my research.

2.2 Related Work

Leveraging widely available sensors can enable more access to healthcare if we build on prior work and push to understand how to connect the digital world with the biological. The fields of ubiquitous computing and synthetic biology have revealed several important things that have influenced the direction of this thesis proposal. From the rise of touchscreen smartphone usage impacting solutions for moblie health sensing to the continued control we're building in using nucleic acid material for repeatable design of DNA computers, the vision of using commodity sensors in mobile phones for health sensing continues to emerge.

2.2.1 Phone-Based Health Sensing

Many fields, including ubiquitous computing, have responded to the need for more accessible diagnostics to give us many advances in phone-based health sensing. While attachments for smartphones of various sizes are possible, advancements in health sensing that leverage the sensors that are already built-in to smartphone are becoming more common, as they promise to benefit users without asking them purchase and carry around new hardware.

Many of the phone's built-in sensors have been repurposed for health sensing. The accelerometer, typically used to automatically rotate the screen, has been shown to be able to monitor things passively, from steps to sleep, though the accuracy ranges widely [Pan and Lin, 2015; Höchsmann et al., 2018; Case et al., 2015; Liu et al., 2015; Saeb et al., 2017]. The microphone, typically used for voice recording for phone calls, video recording, and voice control, has been given enhanced usage as a detector for lung function based on breathing sounds and TB detection based on cough sounds [Larson et al., 2012a; Sharma et al., 2024]. The RF antennas, typically used to send and receive data over wifi, cellular networks, and bluetooth, have been repurposed to sense activity in a room [Zakaria et al., 2023; Liu et al., 2019]. The thermistor in phones, typically used to detect whether internal components are maintaining safe temperatures, has been used to detect fevers, as well [Breda et al., 2023].

This thesis will focus on two particular sensors: the camera and the capacitive touchscreen sensor. Therefore, more background is given about using those sensors for healthcare in the following sections.

Cameras

Cameras are included in nearly all modern smartphones, as photography, videography, and video calling have become indispensable features for users, and they offer an interesting health sensing avenue due to the deep multidimensionality of the input. As such, much healthcare equipment purpose-built for gathering health-relevant data involves some element of light or wave sensing, including pulse oximeters, x-rays, computed tomography (CT) scans, and others [Welch, 2005; Kalender, 2006; Tremper, 1989a]. Not all of

these can be emulated by sensors built into a phone, but the ones that use wavelengths close to the visible spectrum could potentially do so.

In any case, cameras function to gather and detect light in different visible wavelengths of red, green, and blue (RGB), bringing that data from the environment into the phone's processor after passing through a series of lenses, sensor, Image System Processor (ISP), and internal data lines. Different external lenses or attachments can be used to modify the light's path into the RGB sensor, such as those for transferring the light to a wavelength that more closely resembles that of purpose-built devices [Bui et al., 2020a]. The camera has been used for contact sensing and remote sensing of bilirubin for infant jaundice, capillary refill time, and other health metrics [de Greef et al., 2014; Strutt et al., 2023]. Different approaches using the same sensor occasionally offer competing results, such as the case with using camera video PPG [Wang et al., 2016] or still images of fingernails [Mannino et al., 2018a] to sense hemoglobin levels. Photoplethysmography (PPG) has been of special interest, as it is used to determine heart rate, heart rate variability (HRV), respiration rate, blood-oxygen saturation, among others. In addition, in research remote photoplethysmography (rPPG) methods have shown quite good accuracy at predicting heart rate (HR) and heart rate variability (HRV), though usage is quite low [Poh et al., 2010].

Sensing fluorescense of chemical reactions using smartphones has often involved semi-complex filters or attachments for positioning the phone just right or for making sure the light is filtered adequately. [Cao et al., 2016; Hunt et al., 2021; Ning et al., 2021; Shah et al., 2018]. Colorimetry in DNA reactions is starting to become more sensitive, though is still limited to certain sequences [Dong et al., 2022; Berk et al., 2021].

The issues with the smartphone camera is that, while the overall architecture of digital image capture sensors has become fairly standard, there are still many different design decisions that manufacturers make to differentiate their product. Adjustments are made at several steps in the process, from the diodes that capture light, to the filters and lenses that are selected, to the hardware-based ISPs, to the software that processes it. There are differences in every hardware vertical that can be manufactured, but camera may be the most varied due to its obvious visual difference.

Capacitive Touchscreens

Compared to cameras, touchscreen sensors have received relatively less attention for health sensing. After a few years of competition, the mobile phone industry has converged on the capacitive touchscreen as the standard for widely sold multi-purpose devices. These sensors enable a wider array of interaction on a small surface, while a display under the touch sensor can change to provide different buttons as the use requires. Thus, capacitive touchscreen sensors have also become nearly universal in smartphones today. Other industries have increasingly included capacitive touchscreens as users have become familiar with the mode of interacting, including wearables, cars, laptops, and other various appliances.

Capacitance, in general, refers to the capacity of an object to hold an electric charge and can be sensed in a couple different ways that fall broadly into two categories: self capacitance and mutual capacitance [Grosse-Puppendahl et al., 2017]. For sensing self capacitance, a system can be designed where a single electrode simultaneously transmits (Tx) a signal, often an oscillating sine wave, and receives (Rx) the impulse of that signal [Nam et al., 2021a]. When an object enters the field that signal reaches, the signal is impacted and changed in its magnitude and phase and that change is sensed by the electrode. For sensing mutual capacitance a system is designed with at least 2 electrodes, so at least one can act as the Tx source and other electrodes can act as Rx receivers. When an object (such as your finger) enters the area of the field, multiple Rx electrodes are impacted in varying quantities based on the amount of charge that object absorbs and its location relative to the electrodes.

Capacitive touchscreen sensors on smartphones have been optimized over iterations of industrial re-

sponse to customer demand to accurately sense the location and movement of human fingers. A capacitive touchsreen leverages mutual capacitance to sense the location of a user's finger, and it is most often composed of a grid of extended electrodes, stretched like wires in a checkerboard pattern. The placement of the electrodes allows the Rx lines to sense different perturbations based on their location relative to the finger and enables localization of the input. While capacitive sensors of the type built for smartphones has not been specifically used for sensing biological molecules, it has been used for health sensing more broadly and custom capacitive sensors have been used for biology in some medically relevant scenarios [Waghmare et al., 2023b; Kimbahune et al., 2021; Horstmann et al., 2021; Honrado and Dong, 2014; Nam et al., 2021b].

2.2.2 Computational Techniques for Health Sensing

As described, there is interest in sensing health-relevant signals using commodity sensor for its potential to increase access, but unhelpful noise is often present in the raw signals gathered due the sensor design, placement, and opaque variation in manufacturing that leads to differences in signals. Various techniques have been employed to resolve the health-relevant pattern out of the noisy signals. In the camera case, theories and algorithmic principles based on the reflectance of light on surfaces have been modeled into a tunable set of transforms for videos of skin that can fairly reliably be interpreted as the heart rate [Wang et al., 2017c].

In certain cases, such as those for detecting blood-oxygen saturation using light absorption, physical properties were studied and modeled to determine how physical laws relate to physiological conditions, such as the case of the Beer-Lambert Law for light absorption of different wavelengths in different tissues [Tremper, 1989b]. In addition, parameters can be fitted through calibration processes, such as those used in pulse oximetry ratio-of-ratios method [Sinex, 1999; Welch, 2005]. In general, when there is a strong physiological reason for a signal being present in a data stream, better results ensue.

2.2.3 Machine Learning

One of the most common group of techniques used for deriving health-based signals from commodity sensors is machine learning (ML). Situations where ML has been shown to be helpful in health sensing begin with a physiological basis for the pattern that ML. The ML model or architecture is often chosen with this in mind. A simple multi-linear regression model used to correlate the color of a person's fingernails to their hemoglobin level [Mannino et al., 2018a].

Deep learning, as a branch of machine learning, has rapidly gained wide use in recent years as its efficacy has been proven for solving problems where a large amount of labeled training data is available [LeCun et al., 2015]. These models are commonly referred to as neural networks (NNs) due to the way their architecture is inspired by the action potential propagation of signals in the mammalian nervous system. Transformers are an even more recent development which involve tuning only the attention aspect of neural nodes, and which led to the architectures that support very recent Large Language Models (LLMs) [Vaswani et al., 2017]. Deep learning techniques now vary widely, but those have been used for health sensing in the case of blood-oxygen saturation and in heart rate detection from a noisy video signal [Ding et al., 2019; Liu et al., 2020].

The danger of using ML if not chosen carefully is that it can find patterns in parts of data. Explainable AI (xAI) methods can be useful in determining what are the differentiators in a data input that caused the trained model to output a result. However, without an underlying explanation for why ML may work for a given problem, the solution may find patterns in data unrelated to health of the patient, risking improper diagnosis and medical error [Erickson et al., 2017]. Machine learning (ML) is utilized as one of a number

of computational techniques to attempt to delineate a pattern in signals. When used in conjunction with commodity sensors, ML can aid in sorting through noise that is present in the readings from these sensors.

2.2.4 Molecular Reaction Output and DNA Computing

To understand areas where computational techniques can be effective in interpreting health signals, it's helpful to start from an understanding of the underlying biology. Because molecules are so small, the output of molecular reactions must be transduced to an observable signal in order for us to understand our work, diagnostics, and sensing. Building sensing tools that enable remote care involves transducing signals from molecules, which are not readily visible to the naked eye, to a sensor that can read it.

The most common output method in laboratory environments is based on fluorescence output, in which a molecule with fluorescent properties in a specific wavelength is attached to a strand in such a way that it will change in positioning relative to a co-located quencher molecule, and produce light in that wavelength [Smith et al., 1986; Chiang et al., 1996]. This requires a custom machine which both applies energy in the fluorophore's absorption spectrum and reads a narrow band of light corresponding to the fluorophore's emmission spectrum. Multiple fluorophores may be used simultaneously to achieve multiplexing, but the capacity is limited due to spectral overlap [Chen et al., 2007].

Another popular method, colorimetric molecular output, involves adapting the reaction to interact with other molecules such that the color of something changes observably to the human eye. Other projects have investigating improving the accuracy by making this somewhat subjective reading more objective and automatic by using a smartphone camera [Park et al., 2020].

Lab Detection of Molecules

One way to split the types of molecular output methods is whether the availability of those testing tools is limited to a lab space or available more widely. This is helpful, as it allows us to look at which tools are available for only scientists who have access to lab spaces vs which tools are available for a wider set of people due to their cost and availability. A large amount of development has brought many advancements to devices for lab use that allow ever-increasing view into biological phenomena. These are helpful for those with access to a lab, and training, as well as the resources to purchase.

Most sensitive tests are performed in a lab space, specially designed and outfitted with specialized equipment for a controlled environment for sample processing. A major example, which has been used for decades but gained wide popularity for its usage starting in 2020 in response to the COVID-19 pandemic, quantitative PCR (qPCR) is the standard for sensitive nucleic acid tests [Rallu et al., 2006; Gayet-Ageron et al., 2013; Malorny et al., 2003]. However, there are many other devices commonly found in a wetlab that biologists can use to determine what's happening in a chemical reaction based on fluorescence, including plate readers, gel visualizers. All have their purposes and places in scientific research settings, but qPCR has remained the gold standard for nucleic acid-based diagnostic tests, despite the high cost of readers.

For medical tests, a CBC often is used to gather a relatively large quantity of blood and test directly for components, such as hemoglobin, among others. For blood-oxygen saturation, a purpose-built device called a pulse oximeter is used, and an arterial blood gas (ABG) device can be used to directly analyze blood, but requires an invasive arterial catheter to be inserted, and thus is only used in hospital and critical care settings.

Field Detection of Molecules

Principles of these lab devices have proved applicable to out-of-lab solutions, though the tradeoffs must be adjusted for increased accesss, including accessibility or availability based on what devices and training

users can be expected to have. These principles adapters or modifications to biology

When diagnostics or molecular sensing is needed in the field, systems have been designed to be more portable, less power-hungry, and less labor-specific [Panpradist et al., 2019, 2021b; Zhang et al., 2022]. In addition, the portability of DNA sequencers based on nanopore devices is improving, but the cost per test is still high and the compute power needed to identify bases is larger than mobile devices typically support [Eisenstein, 2012]. For these reasons, pushing towards more portability can increase availability of these technologies.

While this work will not fully test the field deployability of these methods, it can demonstrate directions for more field usage by building a system that has a reasonable accuracy and strives for low-cost and ease-of-use.

DNA Computing

DNA computing, as a sub-field of molecular computing that leverages techniques of synthetic biology, refers to the use of deoxyribonucleic acid (DNA) molecules to construct the core elements required for computing [de Silva and Uchiyama, 2007]. Leveraging the Watson-Crick base-pairing of individual nucleotides, in which Guanine (G) pairs with Cytosine (C) and Adenine (A) pairs with Thymine (T), DNA strands can be designed and synthesized to bind in pre-planned ways. This effectively results in somewhat deterministic aggregate effects of the stochastic biological processes that govern biology [Zhang and Seelig, 2011; Simmel et al., 2019]. The circuit analogy stems from the way that these DNA strands can be co-located to form the same set of logic gates that are commonly constructed from transistors to form electronic computers today. Computing has also been demonstrating using other biological molecules, including proteins but this thesis focuses on those using DNA.

When data is biological in nature, such as the case in the genetic code with viral presence or drug resistance status, then it becomes efficient and theoretically sensible to perform the computation in situ. The alternatives effectively incur an extra cost to prepare and sequence the DNA, then use the computational cycles to perform the computation in silico. This has been proposed to have positive environmental and cost impacts due to its highly parallel nature and natural processes that maintain it [Ceze et al., 2019; Liu et al., 2016]. For example, one can design a disease classifier that emulates machine learning in DNA by designing a set of probes and protocols which give different outputs based on the presence of viral or bacterial genetic material [Lopez et al., 2018].

By leveraging principles of DNA computing, output mechanisms can be adaptable that is adaptable or programmable with different DNA sequences as primers. A simplified version. For more complicated circuits, such as those involved in DSD reactions, software systems have been designed that speed up the process of designing DNA circuits by simulating elements of the equilibrium structure or kinetics of reactions based on sequence and environment [Fornace et al., 2020; Zadeh et al., 2011; Zhang et al., 2018; Badelt et al., 2020; Schaeffer et al., 2015]. This may be analogous to early versions of circuit programmers, such as HDL simulators, if the molecular programming field continues to grow, though many components are still in early stages of development.

In all of these molecular reactions, the same output limitations apply as discussed earlier. If we continue to develop the output of these reactions to enable molecular output to more easily be transduced to the digital world, the analogy can continue to form the beginnings of what may be called a Molecular Programming Unit (MPU) to go along with the Graphics Processing Unit (GPU) and newer Tensor Processing Units (TPU) that are being developed today, effectively enabling more information processing to be done efficiently and locally [Owens et al., 2008; Oh and Jung, 2004; Jouppi et al., 2017].

Chapter 3

Measuring Blood Oxygen Saturation Non-Invasively Using a Smartphone Camera

3.1 Background, Summary, and Related Work

Background: Balanced Data Collection Inspired by Health Validation Studies

In an effective machine learning pipeline for biomarker sensing, the first step is gathering high quality data for training that model. In my work on sensing blood oxygen saturation using a smartphone camera, I noticed that prior studies into this had used a convolutional neural network (CNN), but did not gather a clinically relevant range of data. Because we have a known physical model (the Beer-Lambert Law) for sensing the proportion of hemoglobin molecules in blood carrying oxygen, we can understand why machine learning can be applied in this case to sort through the noise inherent in using a commodity sensor.

In my project, the commodity smartphone camera emulated the purpose-built sensor of a pulse oximeter, but the general purpose nature of the camera added some unhelpful noise into the signal. Machine learning, particularly a relatively shallow convolutional neural network, proved to be a successful computational technique to sift through that noise efficiently on a large scale over thousands of data points. In order to demonstrate accuracy of the model on a clinically relevant dataset, I, along with my collaborators at UCSD and SMU, gathered data for that model from test subjects in an induced hypoxemia study design, called a Varied Fractional Inspired Oxygen (FiO_2) study. This study was modeled on the requirements for clearing a pulse oximeter device for medical use, which allows us to understand the medical relevance of a potential diagnostic tool for which the physics and physiology is already fairly well understood, which allows us to use the computational technique of machine learning to create a potentially clinically-relevant diagnostic.

Summary

Hypoxemia, a medical condition that occurs when the blood is not carrying enough oxygen to adequately supply the tissues, is a leading indicator for dangerous complications of respiratory diseases like asthma, COPD, and COVID-19. While purpose-built pulse oximeters can provide accurate blood-oxygen saturation (SpO_2) readings that allow for diagnosis of hypoxemia, enabling this capability in unmodified smartphone cameras via a software update could give more people access to important information about their health. Towards this goal, we performed the first clinical development validation on a smartphone camera-based

SpO_2 sensing system using a varied fraction of inspired oxygen (FiO_2) protocol, creating a clinically relevant validation dataset for solely smartphone-based contact PPG methods on a wider range of SpO_2 values (70%-100%) than prior studies (85%-100%). We built a deep learning model using this data to demonstrate an overall MAE=5.00% SpO_2 while identifying positive cases of low SpO_2<90% with 81% sensitivity and 79% specificity. We also provide the data in open-source format, so that others may build on this work.

Related Work

Smartphone-based SpO_2 monitors, especially those that rely only on built-in hardware with no modifications, present an opportunity to detect and monitor respiratory conditions in contexts where pulse oximeters are less available. Smartphone-based solutions for monitoring blood oxygen saturation have been explored previously, employing various solutions used to gather and stabilize the PPG signal [Carni et al., 2016], augment the IR-filtered broad-band camera sensor [Bui et al., 2020b], and filter the resultant signal for noise or outlier correction [Ding et al., 2018]. Some solutions require extra hardware, such as a color filter or external light source [Mendelson and Ochs, 1988; Carni et al., 2016; Tayfur and Afacan, 2019; Scully et al., 2011; Bui et al., 2020b], whereas others rely only on the in-built smartphone hardware and employ software techniques to process the PPG signal [Tomlinson et al., 2018; Nemcova et al., 2020; Ding et al., 2018; Lamonaca et al., 2015; Sun et al., 2021; Kateu et al., 2022]. These prior works indicate that there is potential for smartphone-based SpO_2 monitors to fill gaps in access to care, but lack validation data on a full range of clinically relevant SpO_2 levels. Prior evaluation techniques for these smartphone-based studies have been limited to a minimum of 80% SpO_2 using techniques such as breath-holding, which is limited to short durations of data collection due to participant discomfort, limiting the clinical applicability of the findings. The US Food and Drug Administration (FDA) recommends cleared reflectance pulse oximeter devices achieve $< 3.5\%$ error across the full range of clinically relevant data of 70%-100% [Luks and Swenson, 2020; Food et al., 2013]. To our knowledge, our study is the first to evaluate unmodified smartphone-based pulse oximetry on this range of SpO_2 data using a Varied Fractional Inspired Oxygen (Varied FiO_2) study procedure, as shown in Fig. 3.1.

Blood-oxygen saturation, reported as SpO_2 percentage, is one of a number of health measures used by clinicians to assess cardiovascular function, reporting the proportion of hemoglobin in the blood currently carrying oxygen. This ratio can be directly measured from samples of arterial blood using an Arterial Blood Gas (ABG) analysis device. However, obtaining and analyzing arterial blood samples is invasive and can be technically difficult; therefore, it is limited to in-clinic hospital or outpatient lab cases. As a result, clinicians typically rely on the convenience of noninvasive measures of SpO_2 using FDA-cleared, purpose-built devices called pulse oximeters, consisting of a finger clip and readout screen (Fig. 3.1b). Pulse oximeters typically perform oxygenation measurement via transmittance photoplethysmography (PPG) sensing at the finger tip, clamping around the end of the finger and transmitting red and IR light via LEDs [Welch et al., 1990]. By measuring the resultant ratio of light reception on the other side of the finger, the devices leverage the Beer-Lambert Law to estimate the absorption properties of the blood, using calibrated curves based on empirical data to infer blood oxygen saturation [Bui et al., 2020b]. This device allows clinicians to noninvasively monitor SpO_2 for single (spot-check) or continuous measures.

While baseline SpO_2 level varies slightly (typically $96\% - 98\%$ at sea level in otherwise healthy individuals), deviations of 5% or more below these levels can be a sign of more serious cardiopulmonary disease. Respiratory illnesses, such as asthma, chronic obstructive pulmonary disease (COPD), pneumonia, and COVID-19, can cause significant decreases in SpO_2, hypoxemia (low blood oxygen), and potentially hypoxia (low tissue oxygen). Hypoxia can lead to serious complications, such as organ damage to vital organs like the brain or kidneys, and even death, if uncorrected or occurring acutely for an extended pe-

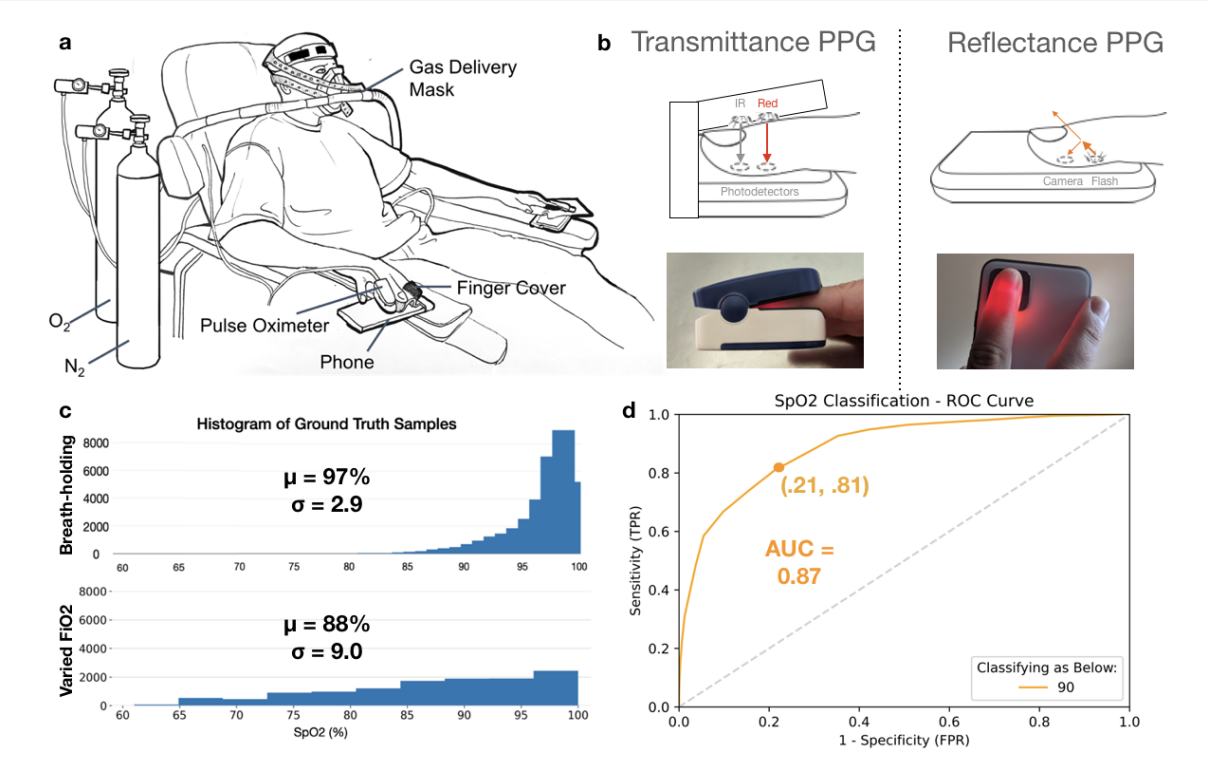


Figure 3.1: Varied FiO_2 study using an unmodified smartphone camera. **a** Illustration of the experimental setup of the varied FiO_2 experiment conducted for this study. The subject breathes a controlled mixture of oxygen and nitrogen to slowly lower the SpO_2 level over a period of 13-19 minutes **b** During the study, one finger was placed over a smartphone camera with flash on to record light response via Reflectance PPG, while a second finger was placed in the fingerclip of a tight-tolerance pulse oximeter acting as a transfer standard, which emits Red and IR light reports SpO_2 via Transmittance PPG. **c** Comparison of histograms of a breath-holding study dataset, adapted with permission from Ding et al. [Ding et al., 2018], and the histogram of the ground truth distribution from our varied FiO_2 experiment dataset reveal that more clinically relevant data spread was gathered using this protocol than prior work. **d** Classification results for the smartphone method reveal that 79% of cases of hypoxemia (defined as a low SpO_2 below 90%) were detected using this method. Illustration and images are the authors'.

riod of time [Bickler et al., 2017]. Repeated measurements of SpO_2 can be used to monitor for changes in the severity of a wide range of cardiopulmonary conditions such as asthma and COPD [Siddiqui and Morshed, 2018], and indicate potential presence of other illnesses including Idiopathic Pulmonary Fibrosis, Congestive Heart Failure, Diabetic Ketoacidosis, and pulmonary embolism [Wilson, 2012; Kline et al., 2003; Zisman et al., 2007]. Pulse oximetry also has prognostic value; for example, an SpO_2 level below 90% SpO_2 has been correlated to increased in-hospital mortality rates for COVID-19 patients [Xie et al., 2020] and levels below 95% associated with complications from community-acquired pneumonia [Moore et al., 2019] or complications in patients with diagnosed pulmonary embolism [Kline et al., 2003]. Thus, determining whether a patient's blood oxygen saturation is below a threshold would likely be valuable in an accessible early warning screening tool to indicate that further attention from a clinician is needed. Usability of smartphone screening tools has also been explored but generally shown that accuracy is poor due to the lack of clinical validation and user experience challenges [Alexander et al., 2017; Modi et al., 2021; Hudson et al., 2012].

In this study, we take a step towards SpO_2 monitoring using the unmodified camera on a smartphone. Our hypothesis was that, by training a model using data from a varied FiO_2 study, we could accurately predict SpO_2 on a wider range of clinically relevant SpO_2 levels (70%-100%) than prior smartphone-based studies. Our analysis reveals that a convolutional neural network (CNN) model evaluated on this range is able to achieve, on average, a Mean Absolute Error (MAE) of 5.00% (σ=1.90) SpO_2 in predicting a new subject's SpO_2 level, after it has been trained only on other subjects' labeled data. To assess potential hypoxemia screening capability, we show that this corresponds to an average sensitivity and specificity of 81% and 79% respectively in classifying a new subject's SpO_2 as below 90%. This work builds on a growing tradition of using ubiquitous mobile devices as decision support tools in healthcare, indicating the need for health care consultation [Kanakasabapathy et al., 2017; Laksanasopin et al., 2015]. Smartphones are widely owned because of their multi-purpose utility, and contain increasingly powerful sensors, including a camera with a LED flash [Steinhubl et al., 2015; Gambhir et al., 2018; Topol, 2019]. Researchers have used sensors in off-the-shelf smartphone devices to assess many physiological conditions, including detecting voice disorders [Mehta et al., 2012], tracking pulmonary function [Mehta et al., 2012; Larson et al., 2012b], assessing infertility [Kanakasabapathy et al., 2017], measuring hemoglobin concentration [Mannino et al., 2018b; Wang et al., 2017b], and estimating changes in blood pressure [Wang et al., 2018; Chandrasekaran et al., 2012]. Alongside these results, we share the data from the Varied FiO_2 study with the community, so others may build on this work.

3.2 Results: Hypoxemia Classification with a Smartphone

3.2.1 SpO_2 prediction performance

Our convolutional neural network (CNN) achieved an average MAE of 5.00% (σ=1.90) SpO_2 when trained and evaluated via leave-one-out cross validation (LOOCV) across the range of 70%-100% of data from the varied FiO_2 study (Fig. 3.2). An average correlation of R^2=0.61 (σ=0.15) is observed between the model predictions and readings from the ground truth reference pulse oximeter. The average root mean squared error (A_{rms}) is 5.55% (σ=1.89) across all subjects in this range, which is 2.05% higher than the ISO 80601-2-61:2017 standard of 3.5% for reflectance pulse oximeter devices to be cleared for clinical use [Food et al., 2013]. Bland-Altman analysis demonstrates the performance of the CNN relative to a tight-tolerance fingerclip pulse oximeter in LOOCV. The SpO_2 values predicted by the learned model near the Limits of Agreement (LOA) reported in previous studies of clinical and non-clinical pulse oximeters, while evaluating on a wider range of SpO_2 levels [Luks and Swenson, 2020; Kelly et al., 2001; Muñoz et al., 2008; Lipnick

Subject #	1	2	3	4	5	6	Mean ± SD
MAE	8.56	3.33	4.12	4.51	6.35	3.14	5.00 ± 1.90
A_{rms}	9.03	3.76	4.96	4.88	6.94	3.74	5.55 ± 1.89
R^2	0.52	0.73	0.65	0.41	0.50	0.87	0.61 ± 0.15
μ / LOA	-1.77 / 12.59	-0.36 / 8.22	-0.64 / 10.62	-4.57 / 9.06	2.30 / 11.81	0.75 / 5.79	-0.72 ± 2.13 / 9.68 ± 2.29

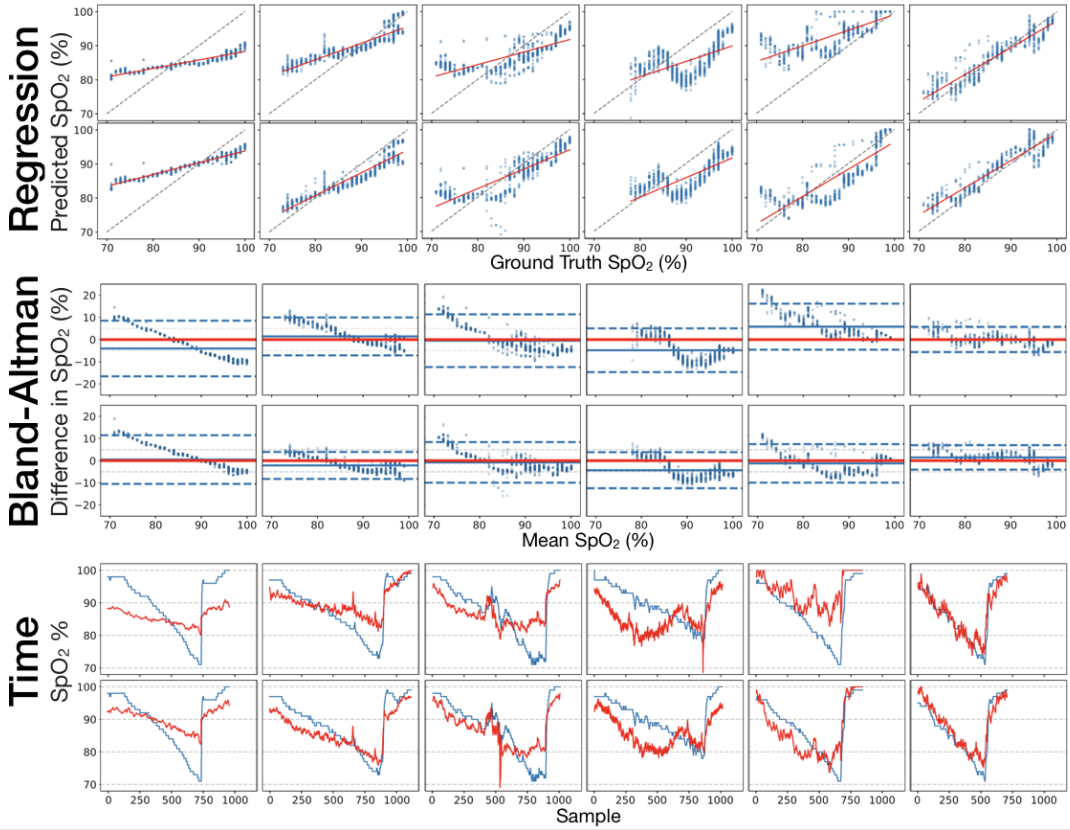
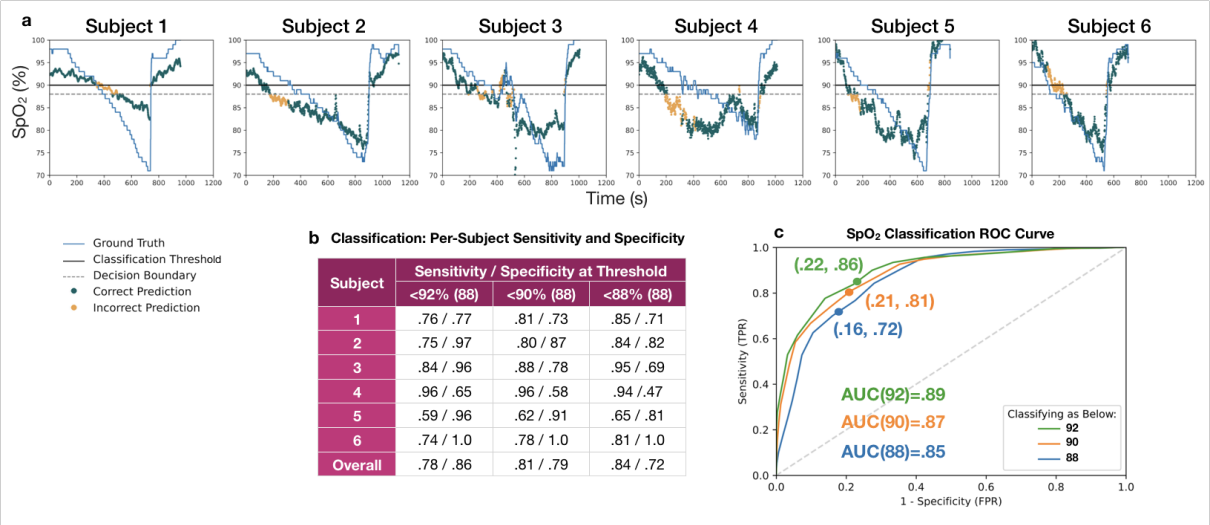


Figure 3.2: Regression results, Bland-Altman comparison, and time series data from the varied FiO_2 study. MAE averages to 5.00% (σ=1.90) over all 6 subjects in the study. R^2 correlation averages to 0.61 over the full range of data gathered. The average difference (μ) and limits of agreement (LOA) average to -0.72 and 9.68. **Table:** MAE and Bland-Altman statistics for CNN evaluation by LOOCV for each subject (n=6) in the study. **Regression:** Predictions from smartphone data plotted against associated ground truth SpO_2 data collected via standalone pulse oximeter. Left hand is on top and right hand is on bottom. **Bland-Altman:** Bland-Altman plots displaying the spread of predictions against ground truth. **Time:** Plots of direct performance analysis of regression results. Model predictions (in red) and ground truth readings (in blue) for the 6 subjects in the FiO_2 study plotted against time of study.

et al., 2016]. Considering that the ground truth measurements from pulse oximeters exhibit similar variance to these results, this indicates that the model has learned features in the PPG signal that are common across subjects and the model is not simply mean-tracking. On the other hand, for Subjects 2, 3, and 5, the negative trend in predictions and mean difference above the limits of agreement for most ground truth values in the range 70% - 80% SpO_2 reveals that the model is consistently over-predicting on SpO_2 samples below 80%. Notably, this is the first study to observe model performance below 85%, as no prior work has demonstrated that smartphone-based sensing systems may perform poorly in this range.



Subject	Sensitivity / Specificity at Threshold		
	<92% (88)	<90% (88)	<88% (88)
1	.76 / .77	.81 / .73	.85 / .71
2	.75 / .97	.80 / 87	.84 / .82
3	.84 / .96	.88 / .78	.95 / .69
4	.96 / .65	.96 / .58	.94 /.47
5	.59 / .96	.62 / .91	.65 / .81
6	.74 / 1.0	.78 / 1.0	.81 / 1.0
Overall	.78 / .86	.81 / .79	.84 / .72

Figure 3.3: Classification results for the system. a Classifications overlaid on ground truth for each subject with a 90% classification threshold and 88% decision boundary. **b** Summary statistics for classification across subjects shows that classification performed better on certain patients, and overall achieved a 81% sensitivity and 79% specificity rate at sensing whether a subject fell below a 90 % SpO_2 level **c** ROC curves for the classification of low SpO_2, produced by thresholding the regression model. Classification accuracy decreases as the classification goal is shifted lower, from 92% to 90% to 88%. The classification decision boundary was varied to produce curves for all 3 classification goals, with each point plotted as the average test classification False Positive Rate and True Positive Rate for all LOOCV combinations. The points that are labeled on each curve are those closest to (0,1) for each classification threshold. The Area Under the Curve (AUC) is .87 for the 90% threshold SpO_2 level classification.

3.2.2 Classification of hypoxemia

Rather than simply inferring an estimate for SpO_2, a smartphone-based tool could be valuable for screening for low SpO_2, indicating whether or not further medical attention is needed. To explore the potential of using an unmodified smartphone camera oximeter system as a screening tool for hypoxemia, we calculated the classification accuracy of our model in providing an indication of whether an individual has an SpO_2 level below three different thresholds: 92%, 90%, and 88%. A pulse oximetry value below 90% SpO_2 is a common threshold used to indicate the need for medical attention [Jones et al., 2003], but other thresholds could be valuable clinically. Thus, we evaluate the ability of our system to classify samples from our test set by thresholding the regression result from our model at different decision boundaries and comparing it to whether the ground truth pulse oximeter simultaneously reports less than the threshold value. We compute sensitivity (true positive rate) and specificity (true negative rate) across all combinations of LOOCV to compute an average result. This experiment simulates the scenario where a smartphone screens a subject it has never seen before, as the model was trained only on 5 other subjects from the dataset.

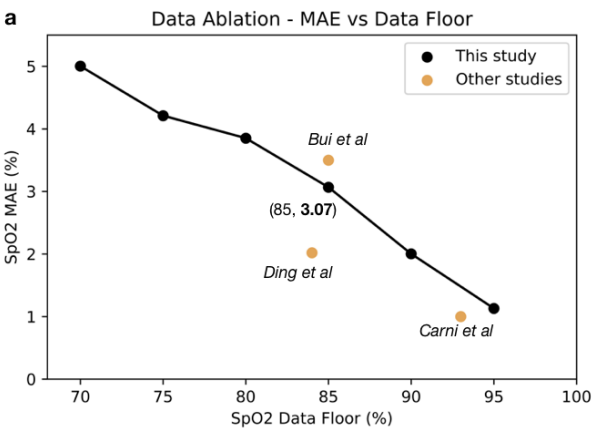
The results of this classification analysis can be seen in Fig.3.3. For classifying SpO_2<90%, on average across all 6 test subjects, our model attains a sensitivity of 81% for correctly classifying the positive samples in our dataset of suspected hypoxemia, while maintaining a specificity of 79%. For classifying a subject as below SpO_2<92%, specificity increases to 86% with a sensitivity of 78%. Not all combinations of test and train subjects displayed the same level of accuracy. In order to visualize classification accuracy across our entire dataset, we varied the classification decision boundary for three classification thresholds that may be clinically relevant, (92%, 90%, and 88%), and averaged the results across all 6 combinations of LOOCV. The results of varying the decision boundary are plotted on the ROC curve in Fig. 3.3c. For the SpO_2<90% classification threshold, the highest accuracy (defined as the closest point to (0,1) on the ROC curve) occurred when the classification decision boundary was set to 88% SpO_2. A decision boundary of 90% on the regression result for the SpO_2<90% classification task resulted in 92% sensitivity at identifying hypoxemic cases alongside 35% false positives (sensitivity of 92% and specificity of 65%).

Classification on individual subjects can be seen in Fig. 3.3a. The model achieved the best performance on Subject 4, with a sensitivity=88% and specificity=78%, reporting correctly 88% of the time when the subject had a dangerous SpO_2 level. Subject 1 displayed the lowest sensitivity to specificity tradeoff of 81% to 73%. As noted in Discussion, the subject had significantly thickened skin on their fingers. Even though the regression for this test subject produces a $MAE = 8.56\%$, the classification result indicates that the tool could still be helpful in determining whether or not the user should seek medical attention. Overall, this classification result indicates that the current model is insufficient for medical use, but further research, including collecting a larger data set from a wider array of subjects, may improve the accuracy in the future.

3.2.3 Data ablation

To understand how the accuracy of our model compares to previously published smartphone-based pulse oximetry systems, we study how excluding subsets of the dataset affects the accuracy. Due to the larger range evaluated in this study compared to prior studies, the overall MAE is not as low as prior studies. However, a data ablation study reveals that, as subsets of the data with lower associated ground truth SpO_2 readings are removed, the accuracy of our model nears that of other published work. Notably, none of these proof-of-concept works were evaluated on data where a statistically significant portion of the SpO_2 evaluation data was below 85%, whereas in our varied FiO_2 dataset, the minimum SpO_2 value included is 70% and the mean of all ground truth SpO_2 levels is 87.1% (See Fig. 3.1c).

We train and evaluate our machine learning models against a similar dataset to these proof-of-concept



b

SpO_2 Data Floor	MAE	A_{rms}	μ / LOA
70%	5.00	5.55	-0.72 / 9.68
75%	4.21	4.67	-0.65 / 8.04
80%	3.85	4.20	-0.83 / 7.15
85%	**3.07**	3.38	-0.60 / 5.38
90%	2.00	2.30	-0.51 / 3.55
95%	1.13	1.30	-0.33 / 1.85

c

Study	MAE	Range (SpO_2 %)
This study	**3.07**	85% - 100%
Ding et al	2.02	84% - 100%
Bui et al	3.50	85% - 94%
Carni et al	1.00	93% - 98%

d

Study Method	# Subjects	# Samples	Mean / Std	Range	MAE (85-100)	MAE (70-100)
Breath-holding (*Ding et al* (1))	39	~6000	97 / 2.9	82-100	2.02	N/A
Varied FiO_2 (this work)	6	~12000	88 / 9.0	61-100	3.07	5.00

Figure 3.4: Data ablation study. As shown by a data ablation analysis, our model achieves increased accuracy at smaller ranges of data, such as that in prior studies evaluating above 80% SpO_2 using a breath-holding study technique. **a** Accuracy of our model improves when ablating our data to remove data below a floor of lowest ground truth SpO_2 readings. **b** Accuracy statistics from the ablation analysis show that A_{rms} and Limits of Agreement improve with a higher data floor. **c** Mean Absolute Error (MAE) of prior works in smartphone-based SpO_2 sensing that perform on datasets with SpO_2 values in the range of 85% to 100%. When the range of the data in our work is reduced to a similar range, we achieve comparable accuracy to prior work. Note that Bui et al used attachments on the smartphone to enhance the photoplethysmographic signal for inference while Ding et al and the present work use an unmodified smartphone camera [Ding et al., 2018; Bui et al., 2020b; Carni et al., 2016]. **d** Sample statistics and MAE results for this varied FiO_2 study are compared to a recent breath-holding study using smartphone cameras and deep learning [Ding et al., 2018].

works using a data ablation technique. We first subsample our dataset so that we only include samples with ground truth SpO_2 above a floor threshold. We then retrain and evaluate our models to calculate a sub-sampled MAE. Varying across possible thresholds, we observe a negative linear correlation between the minimum SpO_2 value included and the resultant mean absolute error, as can be seen in Fig. 3.4a. That is, as we reduce the range of SpO_2 values in our training and testing dataset, our models perform more accurately. To directly compare to the performance of prior work from Ding et al. and Bui et al. (Fig. 3.4b), we set a SpO_2 threshold of 85%. While Ding et al. report a range of 73%-100%, their dataset shows that only .6% of all samples are below 85% (Fig.3.1c), so we report this as a practical floor of 85% for comparison purposes. At a floor SpO_2 value of 85%, our model performs nearly as well as prior work with a mean absolute error of 3.06%. With this analysis, we can be confident that our techniques are at least as reliable as prior works, and likely benefit from the larger range of training examples.

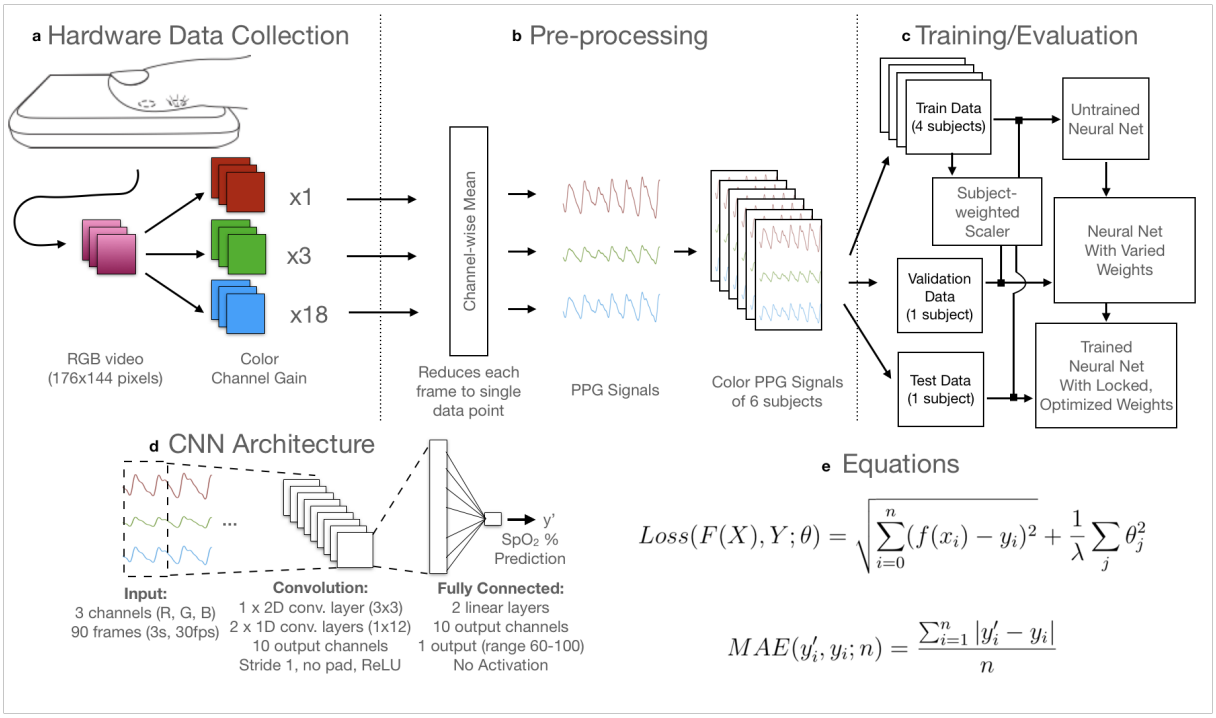


Figure 3.5: Signal extraction and deep learning pipeline a PPG signal extraction occurs after collecting video data from the smartphone camera, applying empirically determined per-channel gains to ensure that each channel is within a usable range (no clipping or saturating). Gains for the R, G, and B channels were empirically determined and held constant throughout all subjects to avoid clipping or biasing towards one channel. **b** Pre-processing of the data extracts the PPG signal for each channel by computing the average pixel value of each frame. The mean of each channel value across the entirety of each frame was used. **c** Training and evaluation was performed using Leave-One-Out Cross-Validation (LOOCV) by using 5 subjects' data as the training set, holding one of these subject's data as the validation set for optimizing the model, and then evaluating the trained model on one test subject. **d** The deep learning model is constructed of 3 convolutional layers and 2 linear layers operating on the input of 3 seconds of RGB video data (90 frames for 3s at 30fps). The output is a prediction of the current blood-oxygen saturation (SpO_2 %) of the individual, which was evaluated using Mean Absolute Error (MAE) compared to the ground truth standalone pulse oximeter reading. **e** Equations for Loss and MAE that were used in training and evaluating the model.

3.3 Discussion: Machine Learning Data Representation

The classification results from this study indicate a direction to consider for enabling more accessible screening for hypoxemia via unmodified smartphones. Considering the unique positioning of smartphones in the pockets of billions of people worldwide, it would be useful to not only reproduce the function of a pulse oximeter in software, but also to provide an initial screen for clinically significant low SpO_2 levels. The COVID-19 pandemic highlighted this need for an affordable remote oxygen desaturation detection tool that can be accurately and safely used for initial screening and monitoring, informing users whether or not they should seek expert medical attention. This potential is important to consider, as software applications are already being used in this manner even when those applications have not cleared the FDA regulatory requirements [Jordan et al., 2020; digiDoc Technologies, 2013 (accessed August 10, 2020]. Our system is the

first unmodified smartphone camera sensor to report accuracy at levels below 85% SpO_2, and it achieved relatively high sensitivity (81%) and specificity (79%) when classifying subjects with SpO_2 below 90%.

This SpO_2 prediction pipeline, including smartphone hardware, custom software application, data processing, deep learning and evaluation, is summarized in Fig. 3.5. Overall, CNN modeling worked well on this input data, learning a function that approximates the data in a non-linear fashion.

We designed our CNN architecture with three goals in mind. First, we chose the overall number of layers in our model such that there are enough affine computations that the model could learn to approximate the ratio-of-ratios model traditionally used in purpose-built pulse oximeters. Second, we included convolutional layers to provide time-invariance to remain robust to inputs that start at different phases of the heartbeat. Third, the ReLU activation function provided non-linearity, allowing the model to learn features embedded in our wide-band input data, which has more noise than the narrow-band input of purpose-built pulse oximeters traditionally used with a ratio-of-ratios model. Linear layers finally regress from these high-level features to the SpO_2 prediction.

We investigated other types of models, as well, but were not able to achieve better performance than our CNN model. While the ratio-of-ratios approach from [Nemcova et al., 2020] has some predictive power on this full range of smartphone camera data (average MAE=7.12, σ=1.64), it does not infer SpO_2 as accurately as our CNN model (Supplementary Figure 1). This is likely because our CNN model, having more parameters and derived features, handles noise in the signal better than the ratio-of-ratios model. Even so, the fact that the ratio-of-ratios model showed some correlative power (average R^2=.21, σ=0.20) is encouraging, suggesting that it and the CNN model could have modeled a similar underlying phenomenon. Future work may leverage gradient-weighted activation mappings to further investigate this relationship.

Statistically, our study does not indicate that this smartphone method of measurement and deep learning approach is ready to be used as a medical device comparable with current pulse oximeters, but further studies could be conducted to develop the method and validate for medical use. A Wilcoxon signed-rank test indicates that our observed MAE differences are large enough to reject a null hypothesis that the measures are equivalent with p=0.03, even though the sample size is small (n=6). The ISO standard 80601-2-61:2017 for safety of pulse oximeter devices indicates that at least 10 subjects with diverse skin tones should be tested and result in a root-mean-squared error (A_{rms}) below 3.5%, which indicates that more subjects should be tested before we can determine whether this method is accurate enough for clinical use [Food et al., 2013; Clinimark, 2010 (accessed August 9, 2020; Batchelder and Raley, 2007]. In addition, an Arterial Blood Gas (ABG) measurement should be used as ground truth for comparison, and a single model would need to be trained prior to testing on these 10 subjects, rather than the LOOCV procedure that was used in this study.

In addition, we investigated whether heart rate (HR) or respiration rate (RR), which are correlated with acute drops in SpO_2, were major contributing factors to model accuracy. We found that encoding the input data as 3 beats at 60bpm, effectively removing heart rate as a discernable feature from the input data, only reduced the accuracy of the model by 0.35 to an average MAE=5.35 (σ=2.20), indicating that HR was not a major contributing factor to model performance (Supplementary Figure 3). RR was not encoded in the input data, as 3 seconds is not enough time to see a single breathing cycle for subjects resting in a reclined position. Overall, this level of performance on a relatively small test subject sample (n=6 subjects with s=12108 total samples) indicates that the model accuracy could increase if more training samples were gathered from further varied FiO_2 experiments, representing a larger range of potential users of the system.

For this study, camera settings were locked during data gathering by presetting auto-balancing and manually enhancing color gain, which are unique steps in our data collection system relative to prior works in this area. Camera image capture is variably exposed based on three factors: exposure time, sensor sensitivity, and aperture. For RGB cameras used in smartphones, all three color channels typically use the same expo-

sure time and aperture settings. Even though the Bayer filter pattern of CMOS camera sensors is designed to sense twice the green light photons per area, it is sometimes not possible to measure all three channels with high dynamic range simultaneously. Both oxygenated and deoxygenated hemoglobin have a significantly higher absorption coefficient in the blue and green wavelengths than for the red wavelengths by about two orders of magnitude. Thus, it would not be possible to measure all three wavelengths simultaneously under the same exposure. If the hardware sensor's sensitivity to a particular color is too high or too low, pixel values for that color may clip by recording the minimum or maximum value of 0 or 255. Because phones use an 8-bit precision scheme for storing pixel data, the pixels will all be rounded to 0 and small changes in that color will be lost. In our application, red is the most dominant color, and prior work has shown that with the use of white balance presets for incandescent light, the tones between blue and green can be amplified [Karlen et al., 2013]. Software advancements in smartphone image processing pipelines now provide more independent control of each color channel's exposure through independent per-channel amplifier gain settings. By having control of independent amplifier gain settings, we can balance the exposure settings to amplify the blue and green channels, as shown in Fig. 3.5a.

We see particularly aberrant performance on subject 1 with MAE=8.56. We suspect this is due to exacerbated tissue noise on the subject's fingers from thickened skin, which is not represented in the rest of the training data. This subject was noted to be the only subject in the study with noticeable calluses on their fingertips, and the subject indicated this was due to sports. We investigate the data obtained from this subject more closely in Fig. 3.6b and observe that the PPG signals for subject 1 show nearly 50% dampened oscillations (AC signal component) and 50% higher average value (DC signal component) than other subjects. We hypothesize that these abnormal features are a result of the calluses. Specifically, an abnormally thick layer of tissue on the finger would absorb more light in the blue and green spectra. Because our device's sensor has fixed sensitivity, the abnormally attenuated light in the blue and green spectra results in poor measurement of the pulsatile blood and altered spread in color channel values. With a small training set of 4 subjects including no other examples of subjects with fingertip calluses, the model cannot learn to account for these tissue differences. We anticipate the model could learn to account for tissue abnormalities if trained on more subjects or if adaptive gain settings were employed to gather data that ensured a similar oscillation amplitude in the AC signal for the input data collected by the smartphone.

From this limited dataset, we are unable to make definitive conclusions regarding the effect of skin tone or sex on smartphone pulse oximetry. Our test subjects included 1 subject with a dark skin tone (subject 2 identified as African-American) and 5 subjects with a light skin tone (all other subjects identified as Caucasian), as seen in Fig. 3.6c. Our model does not appear to perform differently based on skin tone with this limited dataset, as the results for subject 2 fell in a similar range as other subjects, as seen in Fig. 3.2. However, it has been shown that standalone pulse oximeters, such as the one used as the ground truth in our dataset, can produce decreased accuracy on patients with darker skin tones [Feiner et al., 2007; Gottlieb et al., 2022]. Based on our study, we do not claim any findings around model performance based on skin tone, but that should be evaluated in future studies. Our model also does not perform differently on any subset of our 3:3 female:male sex split. Analyzing performance of our model to users of different skin tones and biological sexes is important, but will require further work to understand.

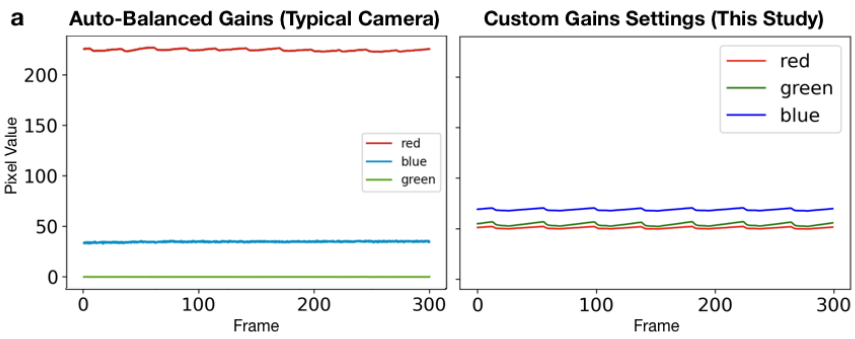
Our results, in this pilot study of 6 subjects, provide a positive indication that a smartphone could be used to assess risk of hypoxemia without the addition of extra hardware in the future. In order to validate and enable this, we would recommend gathering more data with a smartphone in varied FiO_2 studies that induce hypoxemia to increase the training data variety and the accuracy of the deep learning model. With an improved model, we could set up user studies in which the app is used in conjunction with a standalone pulse oximeter to measure the accuracy of the software-based solution in real-world scenarios. Usability of

the smartphone-based measure could be explored to further enhance the clinical applicability of the findings [Alexander et al., 2017; Hudson et al., 2012]. Further analysis could involve processing or preparing this type of dataset differently, including exploring the use of different ROIs for signal extraction [Karlen et al., 2012] and gathering the data differently to study the effect of preset camera gains. Additionally, different phone models have different camera sensor configurations and thus the cross-device compatibility of a model should be tested. We would also like to see what others in the community can do with the open-source FiO_2 data that we are providing alongside this paper. More development and testing could allow this tool to become beneficial for low-cost clinical management of individuals with chronic respiratory conditions, such as COPD, as well as acute respiratory diseases like COVID-19.

3.4 Methods: A New Dataset for Smartphone Oximetry

3.4.1 Study Design and Data Collection

Study Design 6 healthy test subjects were recruited and enrolled to participate in a varied FiO_2 study to evaluate the efficacy of using unmodified smartphone cameras in pulse oximetry. The varied FiO_2 study was performed using the varied fractional inspired oxygen protocol administered by a clinical validation laboratory, Clinimark, which is a group that performs validation services for medical devices [Clinimark, 2010 (accessed August 9, 2020]. This experiment was approved by the Institutional Review Board at Clinimark. Written informed consent for each participant was obtained prior to commencing the test procedure. Six subjects were administered controlled fractional mixtures of medical grade oxygen-nitrogen in a controlled hospital setting for 14-19 minutes. The subjects rested comfortably in a reclined position while the gas mixture was given to induce hypoxemia in a stair-stepped manner. The mixture of oxygen was started at 18-21% and was adjusted downwards in a stair-stepped manner every 1-2 minutes. The goal was to maintain the subjects' SpO_2 level on a plateau for 30 seconds, with 4 discrete levels within the following ranges: above 93 (subject's resting SpO_2 level breathing room air), 89-93, 80-88, 70-79. During this time, the subjects' fingers were instrumented with multiple transmittance pulse oximeter clips and two smartphone devices, with the smartphone device on the index finger of each hand. In the controlled hospital setting, the ambient light was kept at a constant level of a controlled fluorescent white light and the position of the smartphone did not change between subjects. The ground truth data was recorded using multiple purpose-built pulse oximeters, including a tight-tolerance finger clip pulse oximeter acting as a transfer standard, the Masimo Radical-7, which has a tested A_{rms} of 2% between 70%-100% and 3% between 60%-80% SpO_2 and was used as ground truth in this evaluation [Masimo, 2020 (accessed August 10, 2020, 2021 (accessed November 5, 2021]. Variation in ground truth between the tight-tolerance pulse oximeter chosen as a transfer standard (Masimo Radical-7) and other reference pulse oximeters (which were placed on different fingers on different hands) averaged less than 1, measured in the absolute value of the mean difference between all samples, but varied widely in Limits of Agreement, highlighting the differences between approved pulse oximeters (Supplementary Figure 5). Subject characteristics and data statistics can be seen in Fig. 3.6c. Subject observations were recorded, including the observation that one subject, Subject 5 in the analysis, had particularly callused hands. A trained administrator monitored the subjects' vitals, including SpO_2, pulse rate, $EtCO_2$, respiration rate, ECG rhythm, and FiO_2, for any abnormalities and would intervene if deemed necessary. The subjects were informed that they could stop the test at any time. The target for minimum SpO_2 level was 70%, as that is the level above which the ground truth pulse oximeter was validated as accurate, but some subjects' SpO_2 level drifted below 70% briefly during testing and data in that range was excluded from the study (Supplementary Figure 4).



c

Subject	Sample Statistics (SpO_2 %)				Test Subject Characteristics			
	Mean	Median	Range	Samples	Length (s)	Skin Tone	Sex	Age
1	87.1	91	65-100	2180	1090	Light	Male	31
2	88.7	89	73-99	2244	1122	Dark	Male	34
3	86.7	90	66-100	2132	1066	Light	Female	23
4	90.3	90	78-100	2030	1015	Light	Male	20
5	85.8	87	66-99	1854	927	Light	Female	24
6	83.9	85	61-99	1668	834	Light	Female	23
Overall	87.2	89	61-100	12108	1009	1D / 5L	3F / 3M	20-34

Figure 3.6: Analysis of collected data. Visualization of PPG data, derived from smartphone videos, reveal the effects of camera gains settings and skin tissue differences on the input signal for our deep learning model. **a** PPG signal using auto-balance from a prior study [Ding et al., 2018] vs custom empirically determined gain settings from this study. In the left image, the green channel is clipped so that the dynamic range becomes so low that the AC variation in the signal cannot be observed. In the right image, the pulsation is visible in all three channels. This shows how standard smartphone camera settings, designed for photography, can reduce the information available to smartphone-based systems for accurate SpO_2 sensing. **b** Skin tissue aberrations (such as calluses seen in Subject 1's fingers) can affect the quality of data available for SpO_2 sensing. At left, the raw data in the red, blue, and green channels for Subject 1 are dampened and the oscillating portion of the signal cannot be observed at a resolution of 300 frames. At right, the oscillations can be clearly seen for Subject 2 at the same resolution. This abnormality is likely due to Subject 1's callused tissue on the fingers. **c** Subject breakdown for the FiO_2 study and ground truth data statistics (in SpO_2 %) for each subject. The average difference between mean and median for each subject is 1.58, showing minimal skew. Skin tone was recorded based on appearance of the skin on the subjects' hand. The average length of each subject's test run is about 16 minutes.

Smartphone Device Configuration and Setup We collected camera oximetry data with a Google Nexus 6P, recording video at 30 frames per second in a custom video capture application developed in Java using Android Studio. The device was specifically configured so that camera exposure settings in the camera hardware did not change throughout the entire study. Color gains were set to 1x for the red channel, 3x for the green channel, and 18x for the blue channel. These gains were chosen empirically by manually analyzing the impact of gain value adjustments on 20 healthy individuals to find gain values that avoided data loss due to compression and obtained optimal signal quality (see Fig. 3.6a). The Android camera2 API was used to set a target framerate of 30 fps and the phones were plugged in and kept cool with ice packs so the framerate did not dip below 30 fps during the recording. During the varied FiO_2 study, because the device could overheat from recording continuous video with flash enabled for more than 1 minute, we placed clay ice packs around the device to keep its temperature down for the 14-19 minute duration of the study. The ice packs were placed strategically to avoid contact with the hand.

Data pre-processing For each hand on each subject, we recorded an ordered list of n RGB image frames, each with 176×144 pixels. To obtain a PPG signal, we computed the mean pixel value for each color channel and obtained a $3 \times n$-shaped matrix of values. Each hand of each subject is treated as a unique subject in the display of results. We divide the data into samples for each 1-second (30 frames) window, combining the 3 seconds (90 frames) of sample RGB data centered on 1 ground truth SpO_2 reading as one sample. This provides over 8000 training examples (4 subjects) to our models, with about 2000 samples (1 subject) held out for both the cross-validation and test set for each configuration of LOOCV. Samples under 70% SpO_2 are removed prior to training and validation, as the samples gathered below 70% were a result of incidental over-shooting of the intended study range of 70-100% and were not represented in every subject.

3.4.2 Machine Learning and Models

Convolutional neural network We applied a CNN machine learning model, detailed in Fig. 3.5. We designed and trained a network with three convolutional layers followed by two fully connected layers. For the first convolution, we treat the RGB channel components of our signals as a second dimension and use kernel sizes of 3×3 with no padding. We normalize and standardize both training and validation datasets based on a weighted channel-wise mean and standard deviation of the training dataset, where the weights are scaled by the length each subject's data collection. A 3 second sample (90 samples at 30 Hz) was chosen as input based on our intuition that it would provide enough input data for the model to see multiple heartbeats for inference, remaining robust to brief sources of noise from movement, while also keeping in mind usability by keeping the length of required recording to a short length. This input choice was validated via hyperparameter search, which optimized model structure, input data, and regularization parameters based on mean cross-validation set loss across all rounds of LOOCV. The model is trained using the Adam optimizer with a learning rate of 0.00001 (with a rate decay by 0.1 after 80 epochs) and an L2 regularization of strength 0.1, using LOOCV with five subjects in the training set (with one held out for cross-validation) and tested on the remaining subject after optimization. We optimize model weights on the cross-validation subject with Mean Squared Error (MSE) as our loss function and report the accuracy of the results by computing the MAE (Fig.3.5e) on the test subject. Hyperparameters were selected in a hyperparameter search using the LOOCV method, selecting those parameters based on the lowest average cross-validation MAE prior to final model training. The model is built and trained using the PyTorch library. In addition to testing a CNN model, we attempted to model the data and infer SpO_2 accurately using a few different models, including a linear regression model and a ratio-of-ratios model. We achieved the highest

accuracy when the CNN was applied, so we conducted our analysis on that model, but also report our results from the ratio-of-ratios model in Supplementary Figure 1.

Model Benchmark For comparison, we implemented and applied a ratio-of-ratios model as described in [Nemcova et al., 2020] to our data. We apply a variation of the technique in [Nemcova et al., 2020], applying their technique to each 3 second sample from our data where the PPG signal is stable. To do this, we first segmented each beat and extracted the slopes and heights of systolic peaks. Partial beats on the edge of the sample window were dropped. The average of the slopes and heights from the red and green color channels were used to calculate the SpO_2 using eq. 1 from [Nemcova et al., 2020] with the appropriate absorption coefficients. Finally, we fit a linear regression from the calculated SpO_2 values and ranges of the RGB channels to the ground truth. We analyzed the regression performance of this model, and the results can be seen in Supplementary Figure 1.

3.4.3 Statistical Analysis and Data Availability

Statistical analysis We identified and evaluated two potential usage scenarios for a software-based oximetry solution on a standalone smartphone: (1) as a replacement for traditional pulse oximeters by regressing a continuous SpO_2 value, and (2) as an at-home screening tool to inform the need for a follow-up with a physician by classifying regression results as below a particular threshold.

We explored the first scenario of pulse oximetry measurement by performing a regression analysis, comparing our smartphone measurement to a purpose-built pulse oximeter with error and Bland-Altman metrics. In our performance assessment, we evaluated models using Leave-One-Subject-Out cross validation (LOOCV). Specifically, we evaluated six validation splits, holding one subject out as the test set in each split (Fig.3.5c) and averaging the test MAE for the overall reported MAE (Fig.3.2). Signed rank test and skew were computed using the statsmodels library in Python 3.8.

For Bland-Altman analysis, we see minimal skew in our differences with skew=0.64. Therefore, we calculate Limits of Agreement assuming a normal distribution by computing 1.96 times the standard deviation of the difference between the ground truth and predictions of the validation dataset [Bland and Altman, 1999]. We visually examined the ground truth distributions of the splits to ensure there was not a heavy imbalance in the dataset. We experimented with upsampling t he minority SpO_2 range within each training batch. In practice, this increases the weight of mistakes that are made on examples in the minority SpO_2 range. By weighting the minority SpO_2 range mistakes, the model should learn features that improve performance on these examples. However, this had little effect on the performance of the model, so most of our experiments were performed without upsampling. We compared the performance of algorithms using Mean Absolute Error and reported A_{rms} and R^2.

We explored the second scenario of hypoxemia screening by performing a classification analysis, thresholding the ground truth recordings below 3 different SpO_2 levels (92%, 90%, and 88%) and comparing it to our thresholded regression result. We examined the true positive (sensitivity) and true negative (specificity) rates at different screening decision boundaries (92%, 90%, and 88%) to illustrate the potential performance of the system for use in hypoxemia screening. To interrogate the potential to adjust this decision boundary to bias towards sensitivity or specificity, we varied the decision boundary across the range of 70%-100% and plotted ROC curves for each subject using LOOCV.

3.5 Conclusion: Representative Data Collection in Machine Learning

In machine learning for health sensing applications, training and testing the resultant health tool on data that is representative of the target population in both healthy and sick states is needed to demonstrate that such a diagnostic tool would be useful in healthcare situations. My study was the first to measure accuracy of smartphone pulse oximetry on a range of ground truth readings significantly below 90%, the threshold below which many clinicians agree further medical attention and treatment is necessary. Prior studies in smartphone-based pulse oximetry used different methods to gather training data, which did not allow this analysis. Those studies revealed that machine learning may be possible, but mine was the first to test this hypothesis all the way down to below 70%.

Designing the data collection based on the medical regulatory requirements worked in this case to produce good results at ranges that were previously unexplored, paving the path for future work that may increase the accuracy of these methods to be usable in medical practice. It also paves the way for looking at other biomarkers in a similar manner, such as those described in the following chapters of my thesis. Next, we'll look at how smartphone cameras can be configured to noninvasively detect the concentration of hemoglobin in blood.

Chapter 4

Measuring Blood Hemoglobin Concentration Non-Invasively Using a Smartphone Camera

4.1 Background, Summary, and Related Work

Background: Data Augmentation for ML Training in Health

As we saw in the previous chapter, theory and practice show that machine learning models for healthcare perform best for patients when the physiological underpinning is well-understood, the sensor (general-purpose or designed) can measure the expected signal, and a fully representative spread of data is gathered to train the model. In those cases, we can be confident that the trained model fully represents the potential set of patients who could use the diagnostic tool. Most importantly, we can be confident that the diagnostic tool can perform accurately not only in cases when the patient is healthy, but also in cases when the patient is sick and needs medical attention. This is crucial for establishing an adequate sensitivity vs specificity tradeoff for all patients who may use that diagnostic (or utility vs specificity as described in Figure 1.2). However, practical considerations in study design and patient recruitment can intervene, giving cases in which the relevant range of data is difficult to obtain. For example, if the treatment is highly accessible or the lack of treatment can be very dangerous, the treatment is necessarily applied quickly, helping the patient to recover and simultaneously reducing our ability to gather training data that would be helpful in developing the new diagnostic.

This was our experience in the case of pediatric anemia. I, along with my collaborators at MyOr, were interested in trying to build a smartphone-based screening tool for the pediatric population. Treatments, such as iron supplements, can be relatively affordable and available in many cases, which means that pediatric patients can benefit greatly from low-cost anemia screening. It also means that it can be difficult to gather training data from patients exhibiting anemia, as they are often already treated, especially in the hospital setting in which we gathered the data. Even though the physiological explanation, that hemoglobin is the major component of blood that reflects red light, was fairly convincing, and prior research showed that simple machine learning techniques, like a multi-linear regression model, can give a good result, we could not gather an adequate range of data for this study. Therefore, we looked to machine learning literature and employed data augmentation techniques, particularly SMOTE, which enabled us to balance the training data for our model. In this study, we showed that more accurate hemoglobin sensing across the full range

of patients in a pediatric setting may be plausible and demonstrated that computational techniques can be applied to data from commodity sensors to reveal the signal of interest out of the noise to improve access to healthcare.

Summary

We conducted a study to determine whether data collected from a smartphone camera can be used to detect anemia in a pediatric population.

HEMO-AI (Hemoglobin Easy Measurement by Optical Artificial Intelligence), a clinical study carried out from December 2020 to February 2023, recruited patients from the Pediatric Emergency Department, Pediatric Inpatient Department and Pediatric Hematology Unit of the Haemek Medical Center, Afula, Israel. A population-based sample of 823 patients aged 6 months to 18 years who had undergone a venous blood draw for a complete blood count since being admitted to the hospital were enrolled. Patients with total leukonychia, nailbed darkening or discoloration due to medication, nail clubbing, clinically indicated jaundice, subungual hematoma, nailbed lacerations, avulsion injuries, or nail polish applied on fingernails were not eligible for study recruitment. Video and images of the patients' hand placed in a collection chamber were collected using a smartphone camera.

823 samples, 531 from a 12.2 megapixel camera and 256 from a 12.2 megapixel camera, were collected. 26 samples were excluded by the study coordinator for irregularities. 97% of fingernails and 68% of skin samples were successfully identified by a post-trained machine learning model. Separate models built to detect anemia using images taken from the Pixel 3 had an average precision of 0.64 and an average recall of 0.4, whereas models built using the Pixel 6 had an average precision of 0.8 and an average recall of 0.84. Further supplementation of training data with synthetic data boosted the precision of the latter to 0.84 and the average recall to 0.87.

This study lays the groundwork for the future evolution of non-invasive, pain-free, and accessible anemia screening tools tailored specifically for pediatric patients. It identifies important sample collection parameters and design, provides critical algorithms for the pre-processing of fingernail data, and reports an initial capability to detect anemia with 87% sensitivity and 84% specificity.

Related Work

The emergence of artificial intelligence (AI) in healthcare has begun to transform diagnostics, and as of October 2022, the FDA has approved 521 artificial intelligence and machine learning-enabled medical devices [Joshi et al., 2024]. Nearly 3% of these approved devices are hematology devices and include several devices for the analysis of blood taken for a complete blood count (CBC). To date, an entirely non-invasive AI-enabled device for blood component monitoring has yet to receive FDA approval.

Anemia is defined as either a reduced absolute number of circulating red blood cells (RBCs), a diminution in the volume of occupied RBCs, or an insufficiency of the RBCs to meet physiological oxygen-carrying capacity needs.2 It is the most widespread hematological condition globally and affected 42% of children under 5 years of age in 2016 [Chaparro and Suchdev, 2019]. Anemia is associated with negative health outcomes in the pediatric population, including delayed cognitive and behavioral development and functioning [McCann and Ames, 2007; Beard, 2003]. Clinical symptoms of anemia include fatigue, shortness of breath, palpitations, and clinical pallor [Gallagher, 2022].

Anemia is most often diagnosed through a venous blood draw in the clinical setting, relying on several blood parameters, including hemoglobin (Hb), hematocrit, RBC count, mean corpuscular volume, and others. Whereas Hb often provides the first clue as to an individual's status, the other measurements, coupled

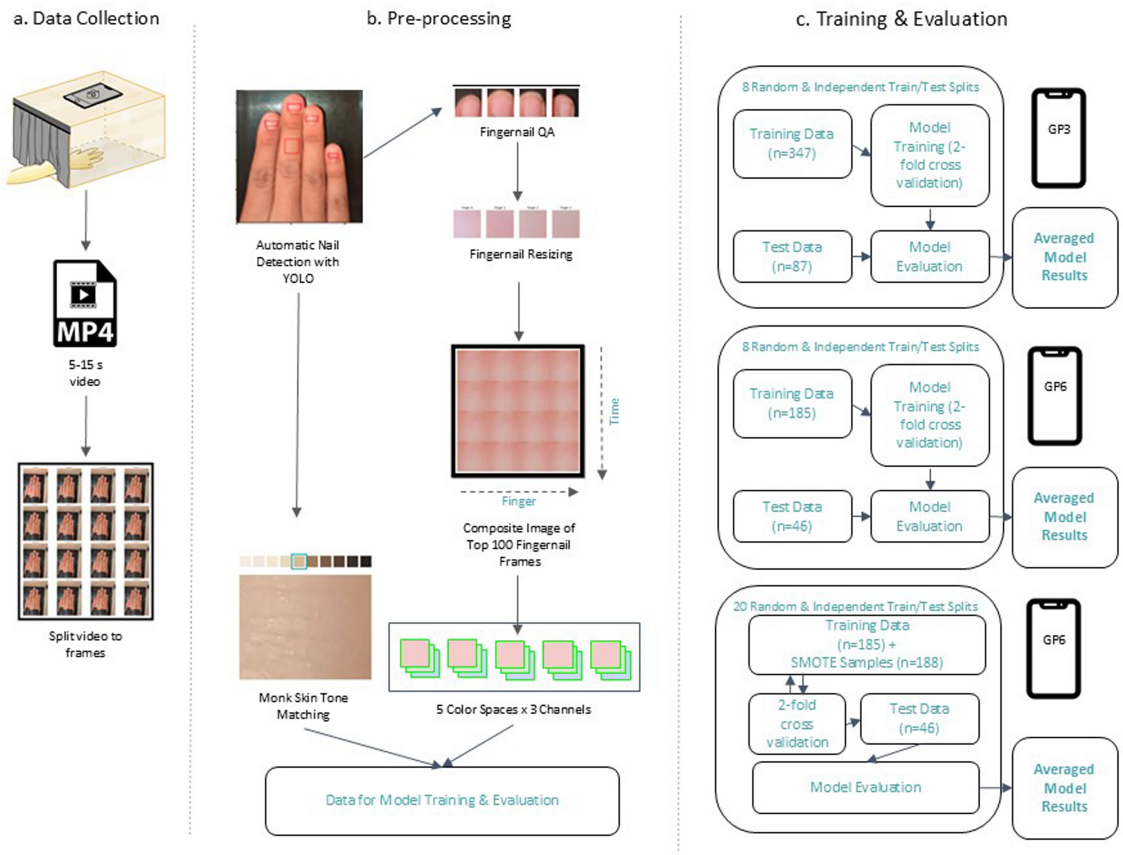


Figure 4.1: Overview of data collection and processing pipeline. (a) Signal extraction occurs by collecting 5–15 s of video on a smartphone camera from the patient's hand situated in a collection chamber. The video is then parsed to individual frames. (b) Individual frames are put through a set of pre-processing steps. The post-trained YOLO model detects a swatch of skin from the middle finger and the four fingernails from the index, middle, ring, and little fingers. The skin swatch is matched to its corresponding monk skin tone. The fingernail images undergo a quality scoring and resizing pipeline, at the end of which a composite image of the top 100 fingernail frames is generated. The gain from 15 different color channels is extracted. (c) Three distinct models were trained and evaluated. The upper panel depicts the development of a model using images collected with a GP3, the middle panel depicts the development of a model using images collected with a GP6, and the lower panel depicts the development of a model using images collected with a GP6, which was supplemented with SMOTE-generated images for the training of the model.

with a detailed clinical history, are used in diagnostic algorithms to pinpoint the cause of anemia [Tefferi, 2004; Moscheo et al., 2022]. The World Health Organization's recommendations on the use of Hb concentration for diagnosing anemia, which has been widely accepted since its publication, provide a summary table outlining blood Hb levels to diagnose anemia stratified by age, sex, and pregnancy. Importantly, it states that children from 6 months to 14 years of age are non-anemic if their blood Hg level is higher than 11–12 grams per deciliter [Organization et al., 2011].

The bulk of anemia screening is carried out in the healthcare setting through the CBC, a routine medical procedure requiring a venous blood draw and specialized machinery. At scale, the CBC is the most accurate and cost-effective anemia screening methodology. However, it necessitates an invasive and painful blood draw by a trained professional, does not provide instantaneous results, and relies on expensive machinery and trained technicians [Wemyss et al., 2023]. This makes it particularly inaccessible in underdeveloped countries. The HemoCue, introduced in 1980, attenuates, but does not alleviate, these pain points, as it requires blood from a finger prick and specialized machinery, yet can provide results in about 1 minute [Cohen and Seidl-Friedman, 1988]. A further advancement in attenuating some of the pain points associated with the CBC came with the introduction of the Masimo Radical-7 SpHb Station which allows for entirely non-invasive blood measurement, albeit at the cost of precision [Causey et al., 2011]. Importantly, the unique absorption spectra of Hb and its derivatives were characterized nearly a century ago [Horecker, 1943].

Recognizing the need for a noninvasive and accessible tool for anemia detection and taking into account the known colorimetric element of anemia (i.e., pallor at distinctive anatomical sites), a flurry of studies have explored the feasibility of using a smartphone for blood Hb measurements in the past few years [Wemyss et al., 2023]. Many of these studies have employed machine learning because it can create an objective function out of previously subjective measures in multifactorial scenarios, such as that in anemia screening where blood coloration in fingernails, noise from ambient lighting, skin tone, and age can all play a role. Each study addresses the inherent combined challenges of ambient lighting, feature selection, color channels, and model input slightly differently, and together paint a promising picture of the future of anemia detection using smartphones. However, studies suffer from small sample sizes, ranging from 29 to 344 individuals, and all but one include data solely from adults. Therefore, there remains a need to develop a smartphone-based anemia screening technology suited for the pediatric population using a robust and representative dataset with broad applicability to the greater population.

4.2 Methods: Augmenting Unbalanced Data

4.2.1 Participant recruitment and selection criteria

Patients were recruited from the Pediatric Emergency Department (PED), Pediatric Inpatient Department (PD), and Pediatric Hematology Unit (PHU) of the Haemek Medical Center, Afula, Israel from December 2020 until February 2023. Written informed consent was obtained by parents/guardians of enrolled participants according to the approved protocols and procedures received from the Helsinki Committee of the Haemek Medical Center (study name: HEMO-AI, approval number 0150-20-EMC). The study was prospectively registered on www.clinicaltrials.gov (Identifier: NCT04573244) on 15 September 2020, prior to subject recruitment. Enrollment criteria can be found in this article's Supplementary Methods Section.

4.2.2 Data Collection

Patients were instructed to place their hands in a customized chamber, described in the Supplementary Methods Section and in Supplementary Figure C.1, in a splayed position for data collection. An Android-native

application (app) was designed and developed for the purpose of the study data collection, described further in the Supplementary Methods Section. The app was loaded onto a phone (GP3) with a 12.2 megapixel rear-facing camera and on a separate phone (GP6) with a 50 megapixel rear-facing camera.

4.2.3 Data Preprocessing

Fingernails from patient images were manually tagged and used to post-train an object-detection classifier, YOLOv8 (You Only Look Once), to detect patient fingernails from a set of images [Redmon et al., 2016]. A histogram representing the color intensity was constructed for 15 color channels and was input into the machine learning model. Additionally, a 50×50 pixel area of skin from the middle finger was detected and automatically matched to its corresponding monk skin tone (MST) swatch which was input into the machine learning model along with the relative error of detected MST to account for inter-level variability [Schumann et al., 2023]. Figure C.1(a) and (b) depicts the data preprocessing workflow and further information on the data preprocessing pipeline can be found in this article's Supplementary Methods Section.

4.2.4 Anemia Classification Algorithms

XGBoost, a scalable gradient-boosted decision tree machine learning algorithm15 was trained on 80% of the clinical dataset and then tested on the remaining 20% of the dataset [Chen and Guestrin, 2016]. A twofold cross-validation was performed for each model iteration whereby the training set was split in half to optimize the model hyperparameters. Eight independent test-train splits were conducted and eight separate classification models were built to observe the reproducibility and consistency of the model when trained with different samples. The Synthetic Minority Over-sampling Technique (SMOTE) was used to generate additional artificial anemia examples for the train set [Chawla et al., 2002]. Figure C.1(c) depicts the training and evaluation pipelines. Model hyperparameters included a maximum tree depth of 3, a learning rate of 0.5, a Logloss loss function, and grid search for hyperparameter tuning. Model development was performed using Sklearn in the Python 3.8 object-oriented programming language. Performance evaluation metrics included recall (sensitivity), accuracy, precision (positive predictive value), f1 score, and area under the curve (AUC) of the receiver operating characteristic (ROC) curve.

4.3 Results: Anemia Classification with a Smartphone

4.3.1 Clinical study results

We collected 823 clinical samples from pediatric patients who visited the PED, and PHU from patients who were hospitalized in the PD at HaEmek Medical Center. About 36 samples were excluded by the study coordinator for irregularities, including fingernails that were photographed with fingernail polish. Samples were collected using a GP3 from December 2020 until June 2022 (n=531) and then using a GP6 from June 2022 until February 2023 (n=256). Figure 4.2 depicts a CONSORT diagram presenting sample collection. Most samples were collected in the PED from patients as they underwent a CBC, with an additional $\sim$10% of samples gathered in the PHU and PD from patients whose CBC had been previously recorded and were found to be anemic. Anemic patients were significantly older than non-anemic patients in the first round of sample collection ($p<0.01$, Table 4.1), whereas anemic patients were significantly younger than non-anemic patients in the second round of sample collection ($p<0.0001$, Table 4.1). No significant differences were noted for sex between anemic and non-anemic patients in either round of sample collection. Histograms depicting the distribution of both data sets are presented in Supplementary Figure 4.2.

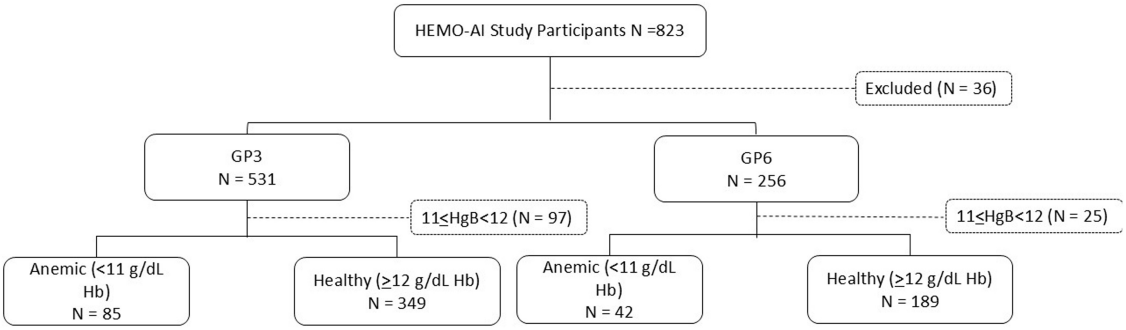


Figure 4.2: Flow diagram of patients enrolled in the Hemoglobin Easy Measurement by Optical Artificial Intelligence (HEMO-AI) study.

4.3.2 Fingernail Detection

We developed a model to autonomously distinguish fingernails and skin sections from other sections of the parsed images, serving as essential inputs for the subsequent algorithm designed to quantify blood Hb concentrations. A pre-trained machine learning YOLO model was post-trained using 329 tagged images of the patients' fingernails and representative skin section. The model's performance was then tested on 44 images. The model successfully identified 97% of fingernails and 68% of skin samples. Figure 4.3(c) and (d) depicts the confusion matrix and the associated precision-confidence curve for the fingernail detection process along with a set of representative training (Figure 4.3(a)) and representative test samples (Figure 4.3(b)). Fingernails detected from images or videos taken with the GP3 were each 20×20 pixels (400 pixels total), whereas fingernails detected from images or video taken with the GP6 were 80×80 pixels (6400 pixels total).

4.3.3 Anemia Classification Model

About 8500 fingernail images from 85 anemic patients and 34,900 fingernail images from 349 non-anemic patients were taken with a GP3 and separately, 4200 fingernail images from 42 anemic patients and 18,900 fingernail images from 189 non-anemic patients taken with a GP6 were used to train and validate eight independent iterations of the XGBoost model. Importantly, all 100 images from each individual were pooled together and presented as one composite image to either the train or test dataset once per model iteration. Samples with a blood Hb level below 11 g/dL were classified as anemic, whereas samples above 12 g/dL were classified as non-anemic. The mean accuracy for the model using images taken from the GP3 phone was 0.51 with an average precision of 0.64, average recall of 0.4 and an average ROC AUC of 0.53. The mean accuracy for the model using images taken from the GP6 phone was 0.79 with an average precision of 0.8, average recall of 0.84 and an average ROC AUC of 0.81. Supplementary Table C.1 presents the metrics of each of the eight independent test-train iterations. To train the model on a more balanced anemia/non-anemia sample set, we used the SMOTE computational technique to generate new artificial anemia samples for the training set (n=118). We trained 20 independent iterations with corresponding random and independent train-test splits of the XGBoost model. The mean accuracy for the model using images taken from the GP6 phone was 0.85 with an average precision of 0.84, average recall of 0.87, and an average ROC AUC of 0.93. Supplementary Table C.2 presents the independent metrics of each of the 20 test-train iterations and Figure 4.4 presents the average ROC curve with the corresponding mean AUC and CI. The mean sensitivity of these models is 87% and the mean specificity is 84%.

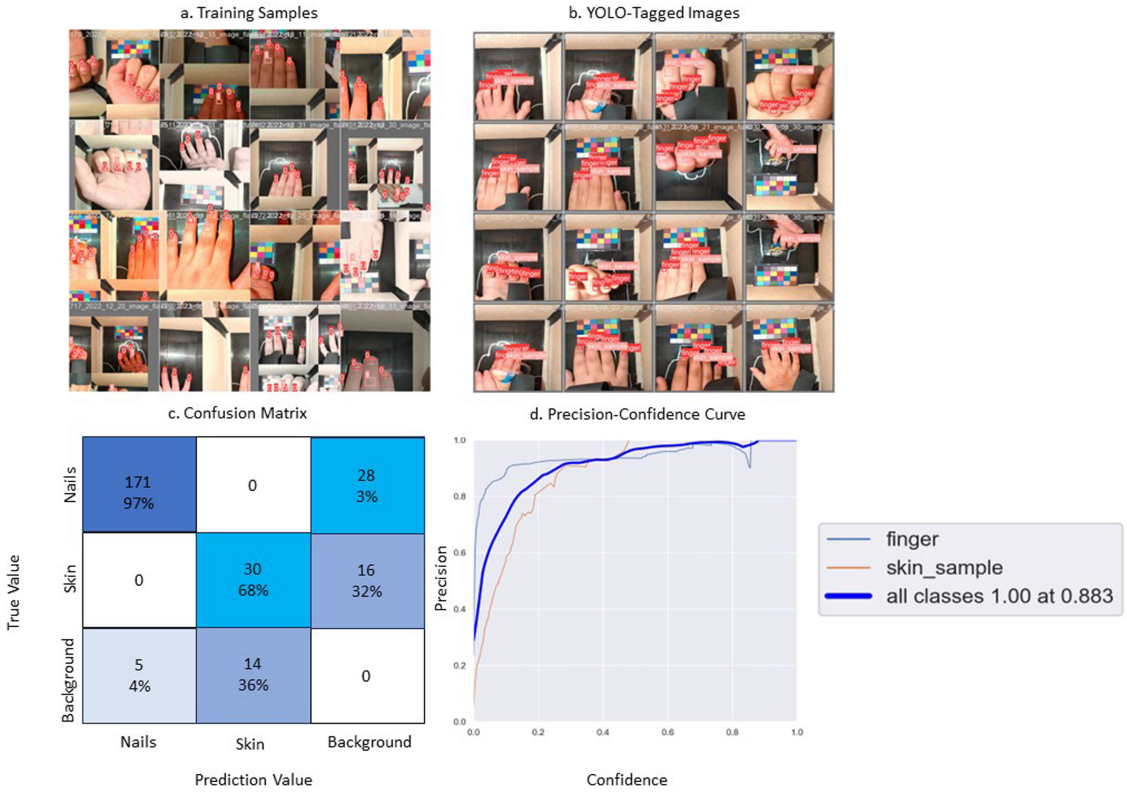


Figure 4.3: Automatic fingernail detection. (a) A set of representative training samples for the post-training of the YOLO fingernail and skin detection model. (b) A set of representative labeled samples tagged by the post-trained YOLO fingernail and skin detection model. (c) Confusion matrix for the fingernail, skin, and background detection process. (d) Precision recall curve for the model is shown. Precision represents the positive predictive value of the algorithm, whereas recall represents the sensitivity of the model.

	GP3					GP6				
	>12 g/dL Hb n=349		<11 g/dL Hb n = 82		p-value	>12 g/dL Hb n = 189		<11 g/dL Hb n = 42		p-value
Age (median [IQR])	12	[8.00, 15.00]	8	[2.25, 18.50]	0.207	12	[8.00, 14.00]	5	[2.00, 10.75]	<0.001
Sex = male (%)	213	(61.0)	49	(59.8)	0.931	109	(57.7)	20	(47.6)	0.31
WBC (median [IQR])	8.77	[6.81, 12.09]	9.54	[6.51, 14.46]	0.211	9.11	[6.78, 11.97]	10.52	[6.54, 15.83]	0.418
NEUT (median [IQR])	66.2	[54.00, 79.60]	55.6	[48.43, 69.02]	<0.001	66.4	[54.25, 81.30]	65.9	[53.23, 77.57]	0.505
LYM (median [IQR])	24.4	[12.75, 33.70]	29.9	[18.32, 39.40]	0.001	24	[11.40, 32.95]	22.9	[13.10, 31.30]	0.699
MPXI (median [IQR])	0.8	[1.80, 3.40]	3.6	[1.93, 5.82]	0.004	9.65	[6.40, 12.57]	10.15	[6.62, 12.03]	0.966
MONO (median [IQR])	4.9	[3.85, 5.90]	5.1	[4.20, 6.70]	0.039	5.3	[4.20, 6.30]	4.5	[3.60, 5.80]	0.069
EOS (median [IQR])	1.4	[0.60, 2.60]	1.5	[0.60, 2.87]	0.674	1.5	[0.50, 3.10]	0.95	[0.40, 1.58]	0.02
BASO (median [IQR])	0.4	[0.25, 0.60]	0.6	[0.40, 0.80]	<0.001	0.4	[0.30, 0.60]	0.4	[0.20, 0.60]	0.505
LUC (median [IQR])	1.8	[1.20, 2.50]	2.3	[1.67, 3.30]	<0.001	1.5	[1.00, 2.10]	1.7	[1.30, 2.70]	0.06
RBC (median [IQR])	4.86	[4.59, 5.06]	3.69	[3.27, 4.45]	<0.001	4.89	[4.65, 5.17]	4.26	[3.54, 4.71]	<0.001
HCT (median [IQR])	39.3	[37.70, 41.62]	29.4	[25.70, 32.00]	<0.001	39.7	[38.20, 42.10]	31.35	[28.08, 32.27]	<0.001
HB (median [IQR])	13.2	[12.60, 13.90]	9.7	[8.43, 10.40]	<0.001	13.1	[12.50, 13.90]	10.3	[9.00, 10.70]	<0.001
MCV (median [IQR])	82.5	[79.20, 85.40]	77.2	[70.90, 83.00]	<0.001	82.3	[79.10, 85.50]	74.9	[67.17, 76.95]	<0.001
MCH (median [IQR])	27.7	[26.60, 28.90]	25.3	[22.50, 27.30]	<0.001	27.1	[26.20, 28.40]	23.7	[21.08, 26.05]	<0.001
MCHC (median [IQR])	33.5	[32.90, 34.10]	32.4	[31.40, 33.30]	<0.001	33.1	[32.30, 33.80]	32.75	[31.42, 33.30]	0.007
RDW (median [IQR])	13.6	[13.10, 14.10]	16.6	[15.20, 18.40]	<0.001	14	[13.50, 14.60]	15.5	[15.03, 16.40]	<0.001
HDW (median [IQR])	2.54	[2.43, 2.73]	2.99	[2.74, 3.26]	<0.001	2.52	[2.40, 2.73]	2.82	[2.69, 3.35]	<0.001
MICRO (median [IQR])	1.7	[1.00, 3.80]	9.45	[4.25, 20.30]	<0.001	2.1	[1.10, 3.85]	9.95	[6.83, 27.08]	<0.001
MACRO (median [IQR])	0.1	[0.00, 0.20]	0	[0.00, 0.20]	0.017	0.1	[0.00, 0.30]	0	[0.00, 0.10]	<0.001
HYPO (median [IQR])	0.8	[0.40, 1.60]	4.8	[2.22, 11.20]	<0.001	2.3	[0.88, 5.80]	6.4	[2.35, 18.65]	<0.001
HYPER (median [IQR])	0.7	[0.45, 1.10]	0.9	[0.50, 1.55]	0.062	0.4	[0.20, 0.80]	0.65	[0.40, 1.33]	0.008
PLT (median [IQR])	256	[216.00, 306.50]	324	[242.00, 487.00]	<0.001	263	[216.00, 313.00]	282.5	[190.25, 385.00]	0.305
MPV (median [IQR])	7.5	[7.00, 8.00]	7.4	[6.75, 7.95]	0.232	7.9	[7.40, 8.47]	7.3	[6.50, 8.15]	0.004

Table 4.1: Demographic and clinical characteristics of anemic and non-anemic patients.

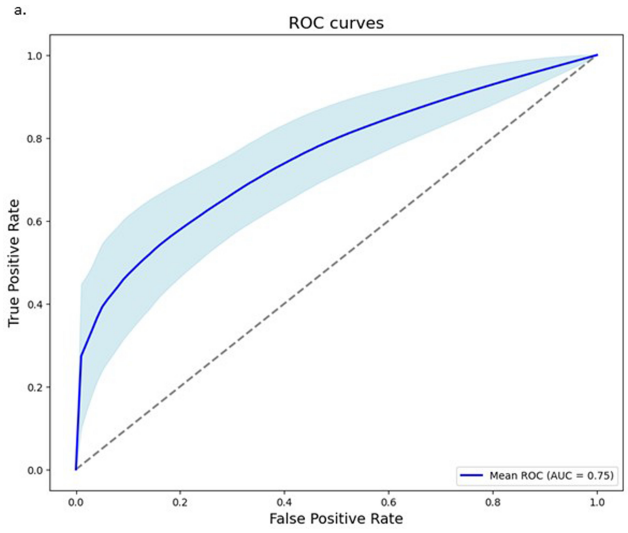


Figure 4.4: Anemia detection model performance. (a) Mean area under the receiver operating characteristic curve of 0.75 for the XGBoost model with 95% CIs representing the variation of 20 independent model runs.

4.4 Discussion: Machine Learning Techniques and Applications

The results of our study provide a framework for the continued development of a non-invasive, pain-free, and accessible anemia screening tool for the pediatric patient that does not require specialized and costly machinery or expertise. We honed machine learning models that automatically detect fingernails of all sizes while in motion as a crucial preprocessing step for widespread adoptability of any smartphone-based anemia classification algorithm. The pipeline parses video to images and extracts color channel histogram data and statistics from each finger in each frame of the video. Anemia detection algorithms were sensitive to image quality, as measured by pixel density, and performed well only when trained and tested on images taken with the GP6 (released in October 2021) and not from the GP3 (released in October 2018). The sensitivity (87%) and specificity (84%) of our anemia classification pipeline provide an attractive alternative for anemia screening in the context of routine well-baby visits, annual checkups, or in underserved populations.

Whereas our anemia screening algorithms are first developed on a large dataset of pediatric patients, other tools developed for the non-pediatric population provided important benchmarks for the development of our pipeline. The vast majority of studies undertaken to date use photographs of the eye, including the sclera, retinal fundus, or lower eyelid [Lobbes et al., 2019; Mitani et al., 2020; Dimauro et al., 2018]. Others used fingernail or fingertip images [Mannino et al., 2018a; Hasan et al., 2021]. Many also grapple with the effect of ambient light on the quality of the photograph, a feature more easily controlled when photographing the hand as opposed to the facial region. The design and employment of a specialized photo collection chamber for our study (Supplementary Figure C.1) was essential for controlling for ambient lighting conditions, thereby ensuring consistency and minimizing extraneous variables, which enhances the accuracy and repeatability of the image data obtained. The materials used to construct the specialized photo collection chamber—wood, masking tape, and black cloth—are economical and portable, making it an optimal tool for utilization in underserved populations and adaptable to a myriad of settings, thereby democratizing access to this technology.

Previous studies count anemic patients with inherited forms of anemia (i.e., sickle cell anemia or thalassemia) as the majority of anemic patients in their cohort. This restricts the model's generalizability, as the idiosyncrasies and unique characteristics of this cohort may not accurately reflect the heterogeneity and diverse manifestations of the condition in the broader population. Furthermore, this limits the potential intended purpose of the technology, as regulatory bodies will most likely restrict the intended purpose of the device to match that of the study population.

In our anemia detection pipeline, we made the deliberate choice to use videos over static images, a decision that significantly enriched our data pool for machine learning models. By parsing each video into hundreds of individual frames, we were able to extract a multitude of data points from a single sample. Crucially, we implemented a strict policy of single representation for each individual's samples: the images parsed from a particular person's video were pooled and utilized only once per k-split model, in either the training set or test set, but not both. This approach ensured that each individual's samples were exclusively represented once in either the test set or the training set, thereby preserving the integrity and variability of the datasets. This shift to video data was enabled through the post-training of the YOLO machine learning model to specifically identify fingernails and a skin swatch from the finger in each frame. This approach not only expanded our data repository but also helped to enhance the overall performance of our anemia detection models by enabling more nuanced and varied data for training. In general, drawbacks of using videos over images can include poorer image quality and noise due to compression, which were outweighed by the advantages of videos in our sample set and pipeline.

Our pipeline recognizes the need to account for skin tone when utilizing images to screen for anemia, given the known racial bias in pulse oximetry measurements [Sjoding et al., 2020]. Variations in skin pig-

mentation can significantly influence the accuracy of color-based assessments, introducing potential bias and disparity in diagnosis. Neglecting to account for this variable could inadvertently perpetuate the issues experienced with pulse oximetry, wherein darker skin pigmentation has been linked to less accurate measurements. The introduction of the MST and the matching of an individual's skin tone to a correlative MST (Figure 4.1) provides the groundwork for an integrative approach to ensure equitable and accurate anemia detection across diverse populations [Schumann et al., 2023].

Limitations of our study include the relatively small sample size of the data set used to train the algorithms and the non-normal distribution of ages and significant differences of ages between anemic and non-anemic patients. The former stems from our decision to recruit patients from the general population, mostly in the PED, and not from more anemia-prone clinical settings. This decision was made considering the notable phenotypic abnormalities of RBCs in sickle-cell anemic and thalassemic patients, which we hypothesized would affect the colorimetric properties of the blood and the notable pallor of cancer patients undergoing chemotherapy, a phenomenon not directly related to blood Hb levels. Algorithms trained with such patient samples may learn to classify individuals based on the unique physiological presentations mentioned, and not by blood Hb levels as mentioned above. To address the limitations of sample size and age differences, we trained eight separate and independent algorithms using eight random train-test splits. The results and metrics of the algorithms were consistent and indicative of a robust methodology for classifying anemia using the input data. Furthermore, we utilized the SMOTE, a computational tool often used in the field of health informatics to address the issue of imbalance in datasets by generating artificial examples from the minority class (anemic patients). This is achieved by pinpointing characteristics within the minority data points and creating new instances that, while similar, exhibit minor variations, resulting in a more evenly distributed dataset for better medical research outcomes. Importantly, SMOTE was used only on images in the training set, thereby improving the training capabilities of the model without hampering its evaluation.

Another limitation of our study stems from the deliberate omission of samples with Hgb values between 11 and 12 g/dL (GP3=97/531, GP6=25/256). While this exclusion limits the clinical application of the algorithm, it was taken in consideration of the WHO recommendations, which demarcate anemia at 11.0 g/dL, 11.5 g/dL, or 12.0 g/dL for different pediatric populations. As there were no enough data to create age- and gender-specific algorithms, the 11–12 g/dL range was excluded. With additional data collection, we could establish models divided by age and gender, and their corresponding clinical cutoffs for anemia. We could also build and train independent models for different skin tones. This would facilitate the inclusion of samples within the aforementioned Hb range, thereby enhancing the model's inclusivity and accuracy.

Our research offers promising insights that lay the groundwork for the future evolution of non-invasive, pain-free, and accessible anemia screening tools tailored specifically for pediatric patients. Our approach harnesses state-of-the-art technology to potentially enable care equity regardless of gender, skin tone, or age. Given the sensitivity of this population to pain and the profound impact anemia can have on their development, the advantages of a non-invasive classification pipeline are multifaceted. This method mitigates the discomfort and fear often associated with invasive CBCs, enhancing the patient experience, adherence to screening protocols, and, consequently, early detection rates. Our results also indicated that the performance of anemia detection algorithms was sensitive to image quality, as defined by pixel density. Therefore, a potential avenue for future research lies in optimizing these algorithms for diverse imaging capabilities of various smartphone models, as current models performed well exclusively with images captured by GP6, but not GP3. As we move forward, we aim to address the limitations of our current model, broaden its applicability, and thereby democratize access to efficient and painless anemia screening, potentially transforming the prognosis for children worldwide.

4.5 Conclusion: Augmenting Machine Learning Data to Indicate Paths Forward for Diagnostics

When real-world circumstances prevent the collection of an ideal dataset for representing the states of the predicted biomarker of interest, machine learning data augmentation techniques can provide an indicator for paths forward. In this study, I worked with a startup company to redesign the data collection smartphone application for the Google Pixel 6 Pro and advocated for gathering data across the full spread of anemic and non-anemic test subjects. However, despite our best efforts, our collaborators at the hospital were unable to collect enough samples from test subjects with anemia to balance the dataset. Data augmentation was helpful in this case, showing that reasonable synthetic data could be sprinkled into the training process to produce a more accurate model overall, illuminating the path forward towards a future where non-invasive anemia screening may be helpful in identifying candidates for affordable treatment.

For real-world health diagnostics, it's likely that a combination of good data collection, processing, and modeling techniques are helpful. Ultimately, a combination of these data gathering and data augmentation approaches add up to the most effective computational approach for building affordable health diagnostics based on sensors built into ubiquitous computing devices. Also, these approaches typically work well for sensing biomarkers which already have a well-developed physiological basis for how they interact with sensors. I demonstrated the computational side of my thesis statement depicted in Figure 1.1 in these first two projects on non-invasive, in vivo health biomarker sensing and learned lessons that I applied to my next set of two projects on in vitro environment sensing.

In cases where the current detection modes are not as well-adapted to direct smartphone sensing, an approach that begins with building understanding of the underlying biology is appropriate for beginning to develop a smartphone sensor-based diagnostic. We will explore this side of the thesis in Chapter 6, beginning with the motivations for how such a diagnostic could be useful for broadly scalable population sensing of pandemic safety in Chapter 5.

Chapter 5

Passively Monitoring for the Presence of Viral RNA in Public Transit Infrastructure

5.1 Background, Summary, and Related Work

Background: Passively Collecting Viral Samples by Leveraging Infrastructure

Applying scalable sensing techniques to other areas that impact health, such as environmental pathogen monitoring, became critical during the global COVID-19 pandemic. Data about whether a certain molecule was located in a specific location defined whether people left their houses and interacted with each other. In order to contribute to the solution for this and future pandemics, we attempted to passively monitor the spread of the virus via a creative passive sampling technique that could theoretically provide a scalable map of the spread of the virus.

I worked with King County Metro to place air filters, which we dubbed "passive fabric sensors", into the existing HVAC air filtration systems of public transit buses. Between the passive sampling and subsequent lab extraction and amplification, we demonstrated that the two steps needed for this system are possible, but I noticed that the latter was more limiting in its scalability than the former. Our method to place these sensors and then gather those samples was relatively low-touch and scalable, but following that up with standard laboratory techniques for RNA extraction and RT-qPCR was laborious, time-intensive, and required specialized equipment in a well-equipped research laboratory. The goal of sensing the virus in a moving vehicle attached to routing and timing data for the bus route has the potential to build a live map of the risk of viral spread in different areas of the city. However, the difficulty of the second step, extracting and affirmatively sensing the viral samples via RT-qPCR, reduces the scalability of the overall process, preventing the accomplishment of these broader goals without enormous investment. As we will see in Chapter 6, this experience inspired the final project in my thesis, and the combination of these two projects demonstrates the left side of my thesis in Figure 1.1. Understanding the biology to effectively amplify its signal is the second critical piece for improving access to screening for health-relevant biomarkers, and will be explored in the next two chapters.

Summary

Affordably tracking the transmission of respiratory infectious diseases in urban transport infrastructures can inform individuals about potential exposure to diseases and guide public policymakers to prepare timely responses based on geographical transmission in different areas in the city. Towards that end, we designed

and tested a method to detect SARS-CoV-2 RNA in the air filters of public buses, revealing that air filters could be used as passive fabric sensors for the detection of viral presence. We placed and retrieved filters in the existing HVAC systems of public buses to test for the presence of trapped SARS-CoV-2 RNA using phenol-chloroform extraction and RT-qPCR. SARS-CoV-2 RNA was detected in 14% (5/37) of public bus filters tested in Seattle, Washington, from August 2020 to March 2021. These results indicate that this sensing system is feasible and that, if scaled, this method could provide a unique lens into the geographically relevant transmission of SARS-CoV-2 through public transit rider vectors, pooling samples of riders over time in a passive manner without installing any additional systems on transit vehicles.

Related Work

The global pandemic of COVID-19 has exceeded 50 million reported cases in the US and 270 million confirmed cases worldwide as of December 20th, 2021 [Centers for Disease Control and Prevention, 2020; World Health Organization and others, 2020]. The virus causing COVID-19, SARS-CoV-2 , is primarily transmitted through airborne respiratory droplets via face-to-face contact [Buonanno et al., 2020; Zhang et al., 2020; Wiersinga et al., 2020] with asymptomatic or pre-symptomatic infected individuals. [Bai et al., 2020; Rothe et al., 2020; Yu et al., 2020; Hu et al., 2020b; Yu and Yang, 2020; Huff and Singh, 2020]. Therefore, disease monitoring via viral presence testing is essential for managing potential outbreaks. Current disease monitoring is focused primarily on testing individual members of the population. However, frequent widespread testing across the entire population can be cost-prohibitive in many communities, even with pooled testing [Augenblick et al., 2020]. While this resource intensive sampling strategy is useful for capturing the overall presence of a disease, alternative environmental sampling can serve as a warning sign of early-stage disease presence in a community prior to symptomatic patients testing positive [Daughton, 2020; World Health Organization, 2020].

One example of passive viral sensing is testing for SARS-CoV-2 in community wastewater plants [Wurtzer et al., 2020; Wu et al., 2020; Daughton, 2020; Peccia et al., 2020; Medema et al., 2020]. However, wastewater surveillance methods suffer from a fixed, coarse granularity since sampling happens far downstream from the individual source. Leveraging multiple environmental sampling techniques through additional infrastructural media, such as public transit, can make viral monitoring more robust. Wastewater sampling has been explored for commercial aircraft and cruise ships [Ahmed et al.], but these approaches cannot be extended to public transport vehicles without wastewater management facilities. Public transit such as buses, light rails, and trains may be valuable targets for surveillance sampling, since they are linked to the population's geospatial mobility. The United Nations estimated >50% global population lives in urban centers [Neiderud, 2015], such as Seattle, where nearly 50% of urban commuters use public transit [of Transportation, 2019].

Viral particles expelled from the respiratory system of an infected individual can circulate through the air into Heating, Ventilation, and Air Conditioning (HVAC) systems, and have been detected in air filters in hospitals treating infected individuals [Kim et al., 2016; Ong et al., 2020; Guo et al., 2020; Chirico et al., 2020; Horve et al., 2020], suggesting a similar approach for public transit. The HVAC system in King County Metro buses, which involve an air intake, MERV-rated filter, and recirculation, pull air from the front or rear passenger-seating areas on buses and are designed to be running all-day long to provide fresh air for passengers. Therefore, they represent an opportunity for passive sensing if viral presence can be detected in the system. Prior work has examined risk of transmission for passengers on buses, trains, and airplanes at local, national, and international scale [Hu et al., 2020a; Zheng et al., 2020; Zhao et al., 2020; Browne et al., 2016; Hoehl et al., 2020; Shen et al., 2020; Luo et al., 2020; Bae et al., 2020; Yang et al., 2020; Kong et al., 2020; Kasper et al.; Harvey et al.; Silcott et al., 2020]; however, prior work has

not explored community monitoring on public transit. One potential reason for this is the cost and time associated with known sampling methods with adequate Limits of Detection (LOD) to sense the low number of copies of virus expected in filters without employing active systems of collection, such as environmental swabbing or vacuum-like Personal Environmental Monitor (PEM) equipment [Moreno et al.]. Rapid and inexpensive RNA extraction methods have detected 10-20 copies/reaction, which may be above the viral copies recovered from passive HVAC systems in non-concentrated settings outside of hospitals [Panpradist et al., 2021c]. Additionally, virus particles can remain viable for 7 days on porous surfaces, like air filters, and 3 days on non-porous surfaces, like metal hand-grips [Aboubakr et al., 2020; Matson et al., 2020]. Therefore, air filters may accumulate and maintain virus over a longer time than swabbed surfaces, capturing data from more individuals with a single sample, enabling pooled testing. Notably, viral presence does not imply infectivity, as a disinfected surface may still hold a dead virus with RNA that can be detected while the virus is no longer capable of infecting someone.

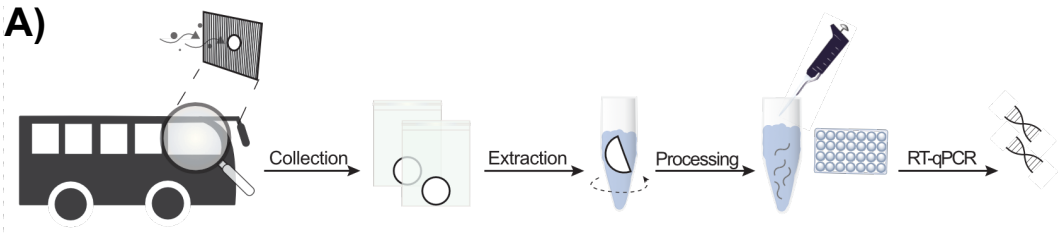
Here, we explore the feasibility of passive surveillance sampling in public buses by installing fabric sensors in vehicle air filtration systems. We demonstrate that sensitive methods of detection can be used to detect small virus copy numbers from samples collected from bus filters, using viral lysis, RNA extraction, and RNA detection via reverse transcription quantitative polymerase chain reaction (RT-qPCR) in a combination not proven in prior literature. In contrast to prior work studying building air filters [Rosario et al., 2018; Horve et al., 2020], we show that a novel combination of methods enables passive, scalable sampling while maintaining the potential for finer-grained community spread monitoring in localized areas via known bus routes. A requirement for scalability of this method is leveraging the already-operating HVAC systems in the bus for sensing while maintaining high analytical sensitivity and specificity for low concentration environments. Thus, we evaluated our in-house method in samples collected from actively circulating buses from August 2020 to March 2021 to demonstrate the detection of SARS-CoV-2 RNA in real-world environments, and we present herein an analysis of how this method may be related to citywide cases for future disease monitoring use cases.

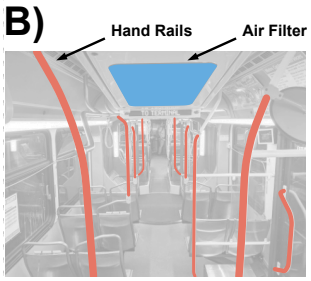
5.2 Methods: Passive Sampling and Environment Detection via In-Lab qPCR

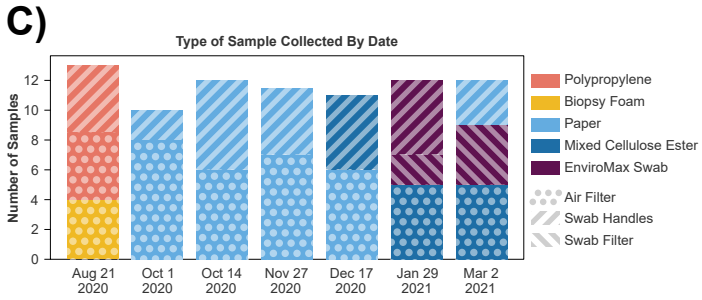
5.2.1 Sample Collection from Public Buses

Between August 2020 and March 2021, environmental samples were collected from 15 actively deployed buses in the Seattle King County Metro fleet (Figure 5.1A). Bus selection was narrowed down to the main bus depot that serviced the Downtown Seattle area, which has the highest ridership. Individual buses were selected to be sampled via a convenience sampling approach based on which buses could be made available at the depot on a regular basis between 7:00-9:00 AM for sample retrieval.

Air filters and environmental swabs were used to capture samples on buses. For air filtration testing, four different materials were tested as supplementary air filters: foam biopsy pads (22-038-221, Fisher Scientific), PolyPro fabric (25PPMB, CanvasETC), mixed cellulose ester filters (A020A025A, Thomas Scientific), and paper filters (Figure 5.1). Supplementary 5 cm^2 filters were placed in front of the existing air filter in the bus HVAC system in both the front and rear of the bus. Mixed cellulose ester filters were preferably used when they were available due to prior literature confirming their effectiveness at capturing and retaining biological material [Bartlett et al., 1997; Junter and Lebrun, 2017]. On filter collection day, four different environmental swab materials were also tested: PolyPro fabric, microporous paper separators, mixed cellulose ester filters, and EnviroMax Swabs. Swabbing was performed by running one swab across common hand-hold areas (Figure 5.1B in red) in the front and rear of the bus. When EnviroMax swabs were available, a second swab of the front face of the main bus filter (Figure 5.1B in blue) was performed.







D)

Filter Type \ Bus Location	All Air Filter	Front Air Filter	Rear Air Filter	Swab - Rear Filter	Swab - Bus Handholds	Overall (% pos, CT mean +/- std)
MCE (Mixed Cellulose Ester)	18.2% (2/11) 37.3 +/- 0.04	N/A	18.2% (2/11) 37.3 +/- 0.04	N/A	0% (0/5) N/A	12/5% (2/16) 37.3 +/- .04
Paper	7.4% (2/27) 38.1 +/- 0.1	12.5% (2/16) 38.1 +/- 0.1	0% (0/11) N/A	N/A	62.5% (10/16) 35.0 +/- 1.8	30.0% (12/43) 35.5 +/- 2.0
EnviroMax Swab	N/A	N/A	N/A	33.3% (2/6) 36.3 +/- 1.1	20.0% (1/5) 32.4 +/- N/A	27.2% (3/11) 35.0 +/- 2.4
Polypropelene cloth	25.0% (1/4) 39.7 +/- N/A	50.0% (1/2) 39.7 +/- N/A	0% (0/2) N/A	N/A	40.0% (2/5) 38.6 +/- 1.1	33.3% (3/9) 39.0 +/- 1.0
Biopsy Foam	0% (0/3) N/A	0% (0/2) N/A	0% (0/1) N/A	N/A	N/A	0% (0/3) N/A
Overall (% pos, CT mean +/- std)	11.1% (5/45) 38.1 +/- 1.0	15.0% (3/20) 38.6 +/- 0.9	8.0% (2/25) 37.3 +/- 0.04	33.3% (2/6) 36.3 +/- 1.1	41.9% (13/31) 35.4 +/- 2.2	24.4% (20/82) 36.1 +/- 2.2

Figure 5.1: Detection of the samples collected from the metro bus using our in-house extraction protocol. (**A**) Workflow for passive sensing SARS-CoV-2 RNA including sample collection, sample transfer from papers or swabs, RNA extraction, and RT-qPCR for detection. (**B**) Sampling occurred via two methods in different areas of the bus. We collected supplementary pre-filters after more than 7 days of being installed inside the HVAC systems of actively-used metro buses (blue). We also swabbed commonly-touched surfaces on the bus (red). (**C**) Sample types and collection methods used during the course of the study. (**D**) Positivity rate and average CT value breakdown by collection material and location. Sampling from both air filters as well as surfaces returned traces of SARS-CoV-2 . Swabs from bus handholds made up the majority of SARS-CoV-2 detections with 42% positivity rate (13/31), while materials placed in air filters had the lowest positivity rate at 11% (5/45).

All bagged samples were placed in a plastic secondary container, which was wiped with bleach-based disinfectant, and transported to an approved lab facility. All procedures involving the untreated filter material were performed in a BSL2-certified Class II A2 biosafety cabinet. All types of filters that were used are shown in Figure 5.1C. Due to safety-related lab space and chemical SOP limits for phenol-chloroform isoamyl extraction, a maximum of n=6 buses (2 replicates for each of the two methods - filter and swab) could be tested in a single experiment. Notably, passenger safety measures were in effect for public transit throughout the sampling period in response to the COVID-19 pandemic. This included a mask mandate, which required all passengers to wear masks while riding the bus, as well as nightly spraying and wiping of all commonly-touched surfaces, such as seats, handrails, payment readers, stop-request pull cords, and door handles, with the Virex quaternary ammonium disinfectant (04332, Diversey).

5.2.2 Detection of SARS-CoV-2 RNA in Filters

Sample extraction for testing was performed within the same day of the sample collection from metro buses. Detection of SARS-CoV-2 RNA consisted of the following steps: viral extraction and lysis, RNA isolation via phenol-chloroform isoamyl extraction, and RNA detection via RT-qPCR (Figure 5.1A).

RNA Extraction and Isolation

Filters collected from buses were cut into 2-cm^2 pieces. Two pieces, considered sample replicates, were placed into microcentrifuge tubes containing 200 lysis buffer (50mM EDTA pH 8.0, 250mM Tris-HCl pH 8.0, 50mM NaCl, 1% (w/v) SDS) [Miura et al., 2011]. The tubes were placed on a foam tube rack attached to a vortexer and agitated for 15 minutes, at high speed, at room temperature. After vortexing, 600 TRIzol was added to each tube, pipette-mixed 10 times, and then the resultant 800 solution was transferred into a new tube. The solutions were incubated at room temperature for 5 minutes to allow complete dissociation of viral particles into the upper media and inactivation of any potentially remaining active virus in the solution. The RNA was then isolated from protein and DNA following the standard TRIzol phase separation procedure [Rio et al., 2010]. Precipitation of RNA was carried out by adding 1 200-proof ethanol and 1 RNA-grade glycogen (R0551, ThermoFisher) to each tube, followed by 1-minute vortexing. Each tube was then incubated overnight at -20. Following overnight precipitation, the supernatant was discarded and residual ethanol was allowed to evaporate. The RNA pellet was washed following the standard TRIzol RNA wash procedure and subsequently re-suspended in 8 nuclease-free water.

Detection of SARS-CoV-2 RNA Using RT-qPCR

Each 8 TRIzol isolation product was assayed with TaqPath 1-step RT-qPCR (A15299, ThermoFisher Scientific) in 20 reactions. We used probes from the CDC SARS-CoV-2 qPCR probe assay targeting two regions in the N gene, designated N1 and N2 (10006713, Integrated DNA Technologies), one for each sample replicate. To avoid cross-contamination, the reactions were loaded into non-adjacent wells in a 96-well plate on ice at a separate bench from where the RNA isolation step was performed. Wells also were covered with Parafilm between loading samples. RT-qPCR was carried out on a Quantstudio 3 (ThermoFisher) using the CDC-recommended protocol [for Disease Control and Prevention, 2020].

5.2.3 Positivity Determination

Positive results were determined by amplification before a specified PCR cycle threshold (CT). Raw RT-qPCR amplification data were loaded via a custom Python script (Supplementary Information) that deter-

mined CT based on a common threshold across all samples (50000 RFU) which delineated between amplified samples and non-amplified samples in control experiments. For environmental samples from buses, a positive result for a bus was declared if one replicate from one of the sample methods from that bus had a CT <40. Positive and negative controls experiments were conducted by dripping AccuPlex enveloped RNA reference material (0505-0126, Seracare, Milford, MA) onto unused filter, and the full extraction and detection method was performed to confirm that positive results were a result of viral presence in the air filter.

5.3 Results: Lab-Based Detection of Viral Particles

5.3.1 Results on Environmental Samples from Buses

Out of 82 samples (164 total with replicates) tested, 24% (20/82, 95% CI: 16-35) of samples tested positive for SARS-CoV-2 RNA (Figure 5.1D), indicating viral presence but not necessarily infectivity. Samples collected by the swabbing method showed the highest positivity rate at 42% (13/31, 95% CI: 25-61) while filters had the lowest positivity rate at 11% (5/45, 95% CI: 4-24). 1 replicate amplified in all positive bus samples (Table D.3). This indicates that the viral particles may not be distributed evenly across the sampling material or that the sample methods may be sensing viral presence from different signal sources (i.e. riders who breathe may not touch the railing). Most positive samples had SARS-CoV-2 RNA near or below the LOD of the RT-qPCR assay (Figure D.2) and thus were confirmed correct product sizes by fragment analysis (Figure D.10).

5.3.2 Method Comparison

We compared our TRIzol-based RNA extraction method with a more commonly used column-based RNA extraction from Qiagen. After dividing five filters in half and processing in parallel, we found a bus positivity rate of 60% (3/5) and 80% (4/5) (Table D.4) in TRIzol-based and column-based methods, respectively. Interestingly, our TRIzol-based method did not yield positive results from any samples collected by EnviroMax swabs, which were positive with the column-based method. We observed black particle residues in samples using EnviroMax swabs (Figure D.11), which were filtered out by the column-based extraction. These residues ended up in the RT-qPCR reactions when EnviroMax samples were extracted by the TRIzol-based method, which may have inhibited the RT-qPCR reaction. On the other hand, the column-based method displayed 0% positivity rate on all air filter samples, while the Trizol-based method detected 40% (2/5) positivity on the same air filters. We hypothesize that the debris broken from filters could interfere with the binding of SARS-CoV-2 RNA to the silica membrane in the columns or the SARS-CoV-2 RNA might remain trapped in the columns. These filters are made from mixed cellulose ester which have different surface properties and porous structures from those of polyurethane foam structures (EnviroMax swabs) which reportedly had a high release efficiency even with minimal agitation [Panpradist et al., 2014].

All air filters were installed on buses for more than 7 days, and thus can represent pooled samples of all riders for the prior 7 days (Figure 5.2). One exception is that, for one sampling date on October 14, 2021, filters were installed and collected in one day. In one-day testing, 0 filters returned positive, indicating that one day may not be enough filter exposure time to build up a detectable viral load. However, a relatively high rate (60%) of swab samples returned positive, which may be attributable to lack of surface decontamination mid-day. Metro cleans buses nightly, and the morning sampling period for all 2-week samples occurred the morning following the decontamination, in between which no riders would have ridden the bus.

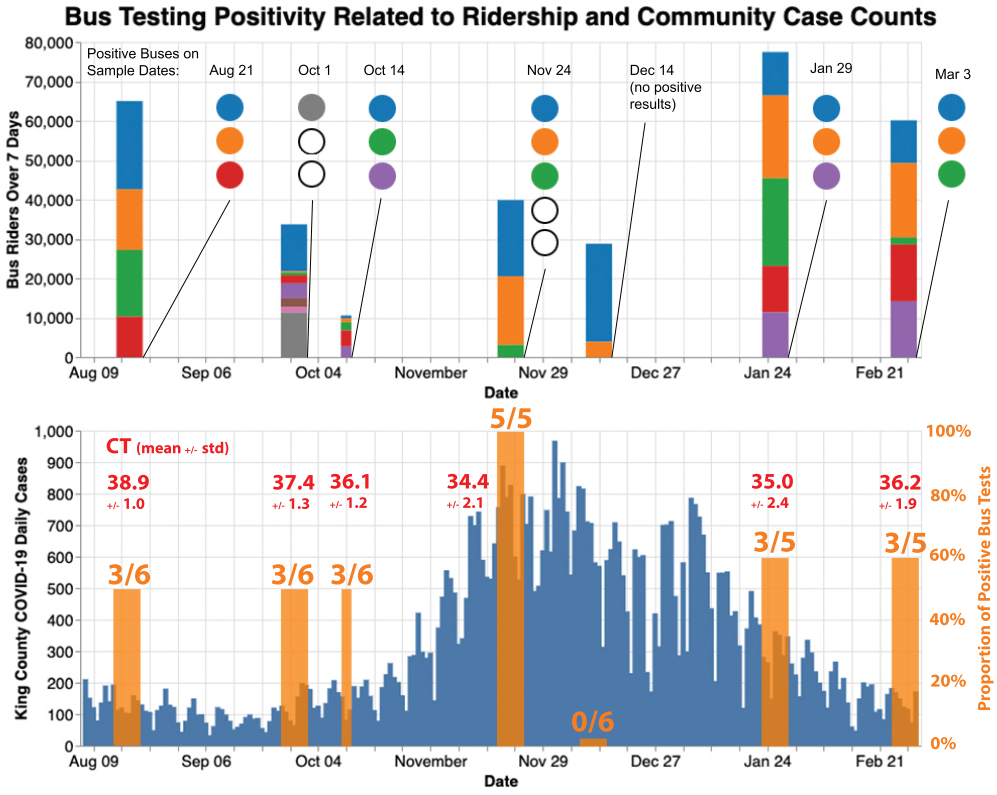


Figure 5.2: Total metro bus riders compared to positive sample results. Top chart shows total riders per bus per 7-day period of filter installation before sampling. Each color denotes a unique bus that week. The color of associated circles denotes a positive result from that bus. An empty circle denotes a positive sample for a bus with 0 riders during that week. **Bottom** chart shows new cases of SARS-CoV-2 in King County (blue) superimposed with the proportion of buses sampled that week returning positive results (orange). CT values for positive results are listed by date (red), and showed a -0.687 Pearson's correlation when compared to King County ridership D.5

5.3.3 Results Compared to Population Testing

Figure 5.2 shows the bus testing results juxtaposed with the SARS-CoV-2 case counts and bus ridership counts for the 7 sampling periods. While the sample size is too low to demonstrate positive correlation, we do see that a higher proportion of buses sampled return positive results when SARS-CoV-2 cases in King County were high, with the exception of December 14th, which showed no positive results. This was likely due to the fact that only 2/6 buses sampled had any ridership during that week, reducing exposure and decreasing likelihood of detection. October 1st and November 24th also returned positive results for buses which did not leave the station (zero riders), denoted by the empty circles in Figure 5.2. This finding indicates that some results in this study may be a signal of continued viral RNA presence after more than 7 days or infected maintenance workers who entered a bus during the sample period.

Metro's ridership data was compared to bus positivity results to determine whether ridership was correlated with the positivity rate for buses, and a small Pearson's Correlation of 0.255 was observed. King County population testing data was also compared to average CT value of positive results, and a Pearson's Correlation of -0.687 was observed (Figure D.5). The negative trend indicates that the strength of the detected signal on buses may increase as more people test positive in the city. The correlation observed may be reduced due to the small sample size and infection control measures taken by King County Metro. This included required rider masking and nightly surface disinfection (see Methods) throughout the entire sampling period, likely resulting in an increased number of false negative results due to a decreased number of viral particles escaping the mask of infected riders and reduced viral presence on hand rails. Ventilation and virus collection on filters may have been affected by driver and rider behavior, as well, as it is the driver's choice whether to activate the ventilation system (though most drivers do) and passengers may open the windows if they choose, which would modulate ventilation and likelihood of viral capture on filters. Additionally, the population of metro riders from which the testing was sampled is not fully representative of the overall population from which individual testing was performed.

5.3.4 Control Validation

In control experiments, extracting spike control from filters yielded 6/9 (66.7%) replicates and 7/9 (77.8%) replicates of spike controls directly in solution amplifying and displaying positive detection using our method, while only 1/18 (5.6%) of the replicates of negative controls returned positive with a high CT value of 37.3 (Figure D.3). Based on a PCR standard curve from the same experiment (Figure D.4), the extraction efficiency is approximately 33% (average of 134 out of 400 copies) for filter extraction controls and 115% (average of 459 out of 400 copies, which is within the error range for PCR extraction) for direct in solution extraction controls (Table D.2). This suggests that about one-third of viral material may be lost during the filter extraction step, but not much is lost during RNA extraction. This may have resulted in some false negatives in bus testing but demonstrates that positive detection results were likely a result of viral material collected from bus sampling.

5.4 Discussion: Potential for Onsite Detection

Here we show that passive infrastructure-mediated sensing of viral presence may be feasible. By leveraging a passive viral sensing method such as ours in parallel with other environmental and individual testing methods, epidemiologists could inexpensively monitor a community to identify locales of transmission and estimate case numbers within local regions. Our method may be more valuable when cases within the general community are low, leveraging the fabric sensors to passively pool respiratory droplet samples from

riders temporally over the filter installation period and spatially over bus routes. This monitoring would be enabled by sufficient resourcing to enable daily sampling and testing of a subset of bus routes to ensure adequate coverage, which was not possible with the small study team in this proof-of-concept study.

This study is limited by its sample size (n=39 total buses). Sample filters were placed and recovered manually by the research team and metro collaborators, which could be scaled by larger research teams (Table D.5). False negatives in air filters may have been caused by the limited coverage of the sample media over the air vent and air currents diverting around the sample filter media, resulting in lower positivity rate than swab samples. In addition, mask mandates were in effect for riders during the sample period, likely reducing the number of viral particles expelled into the air by breathing of infected riders landing on the filters. Finally, Metro's nightly cleaning process is designed to reduce the overall viral load in the bus, even if it did not completely remove the viral presence signal. Considering these effects, the small viral loads (Figure D.2) of some samples are unsurprising, but may fall below the typical LOD for many commercial PCR kits, including our chosen Taqpath 1-step kit [for Disease Control and Prevention, 2020]. Alternative sensitive assays that amplify multiple regions in SARS-CoV-2 may be useful to detect these samples with low concentrations [Kline et al., 2021; Food and Administration, 2021]. We also note that our method does not necessarily identify the risk to bus riders, but rather the presence of inactive SARS-CoV-2 RNA. Viral viability tests and more frequent sampling are needed to understand the risk to riders.

5.4.1 Future Work

Future research into scalable, sensitive viral detection for environmental samples would enhance this approach. Studies evaluating filter placement, size, and material, incorporating further control experiments in simulated environments, could further validate the sensitivity of the method and optimal materals for sampling [Buonanno et al., 2020; Holmgren et al., 2010]. City-wide deployments enabled by scalable detection methods, such as rapid diagnostic lateral flow detection for on-site detection enabled by miniaturized amplification devices [Panpradist et al., 2021c,a] or sequencers [Cardozo et al., 2019], could gather more data, enabling network analysis techniques to study probability of SARS-CoV-2 transmission on a neighborhood level. This method could be adapted and deployed to provide an early signal of community outbreak of SARS-CoV-2 , or other viruses transmitted by respiratory droplets, when case counts are low in the population. This could provide a relatively inexpensive early warning system and ongoing monitoring insight into the local routes of viral transmission for current and future respiratory pathogens.

5.5 Conclusion: Passive Sampling and Non-Passive Molecular Output

In its current state, qPCR is the gold standard for highly sensitive assays, such as the one that was needed for detection of SARS-CoV-2 after passive environmental sampling. However, it is not a very efficient method of detection, despite efforts to scale it up and bring it into the field. After designing and implementing a viral surveillance system leveraging public transit infrastructure to assist with early warning for spread of the COVID-19 pandemic, I identified that the main gap which prevents large-scale rollout of this method is the molecular readout throughput. Improvements in conduction of the molecular signal to the digital world would allow for scaling this method, as well as others, and answering finer-grained questions about the spread of the virus. I developed a passive sampling methodology, but the output mode of the results was far from passive.

Thus, I started developing a method that could imaginably be brought directly onsite for testing for the presence of specific nucleic acid targets. With this, we could process many more samples and begin

to answer questions about disease spread in certain geographic locations, precise timing of when the virus appeared in the filters, and more. This insight led me on a journey to discover more scalable ways to solve the detection problem, which can be broken down to a biological understanding and molecular output problem. The results of that investigation, and how they apply to my thesis, are shared in the next chapter.

Chapter 6

Detecting the Output of Nucleic Acid Reactions Using Capacitive Touchscreen Sensors on Ubiquitous Smartphones

6.1 Background, Summary and Related Work

Background: Improving Access to Viral Sensing via the Smartphone Capacitive Touchscreen Sensor

For detection of viral pathogens, the gold standard for sensitivity and quantification is nucleic-acid based amplification methods, such as RT-qPCR, but those techniques are not as affordable and available to all due to their reliance on lab-based equipment and environments. As we saw in the previous chapter, this detection method is one example of a range of techniques that enable the amplification of the signal that a virus is present. These are all based on the positive recognition of a certain sequence of nucleotides (the target or template) by a designed sequence of nucleotides (the primer), which triggers a chain reaction of replication of that sequence (catalyzed by a polymerase), and quantified by increasing accumulation of a signal from an attached fluorescent molecule. This requires a readout method, often (as in the case of RT-qPCR) necessitating specialized lab equipment and energy input in order to provide the environment to trigger this chain reaction and detect its output. While research efforts are underway to miniaturize and portabalize this equipment, alternative methods are being developed that obviate the equipment and provide acceptable sensitivity for the application [Panpradist et al., 2021b,a].

Building on my experience from previous projects in using smartphones to detect health biomarkers and detecting viruses using highly sensitive lab equipment, I reviewed the literature, the capabilities of the smartphone's various sensors, and I turned towards capacitance sensing of nucleic acid reactions for a potential future look at how we may detect the presence of viruses in the future. The capacitance sensor of most modern smartphones is used to sense and localize the touch of a user's finger, and has converged to a fairly standard design, despite some variations. In this chapter, I build on this standard design to explore the potential for sensing the minute chemical changes associated with DNA hybridization and polymerase-catalyzed extension. At this time, we are not yet at the stage where direct sensing of nucleic acid strands is clearly feasible, so I, along with my Ubicomp Lab and MISL Lab collaborators, built a low-cost tool that can be used in the lab to interrogate the capacitance characteristics of DNA solutions. We showed some potential signs that nucleic acid isothermal amplification reactions may be detectable at certain frequencies via a technique that is related to the mutual capacitance grid of electrodes in modern smartphones. The

process of doing so illustrates the full circle of development of biomarker sensing using commodity sensors, starting from understanding the biology of the underlying molecule to attempt to amplify it.

Summary

Capacitive sensing of the output of nucleic acid molecules has the potential to improve access to health diagnostics and synthetic biology research. Prior studies have shown it may be possible to sense DNA with ITO-patterned capacitive sensing devices similar to the touchscreen, but these studies require surface preparation that make it incompatible with touchscreen smartphone usage. Applying principles of ubiquitous computing to this problem, we envision developing a low-cost, easy-to-use sticker attachment that enables detection of changes in controllable DNA reactions, opening up the sensing method for health diagnostics, environmental detection, and affordable research in the future. We take steps towards this vision in building an affordable lab-based capacitance sensor that uses a Vector Network Analyzer (VNA) approach to interrogating the capacitance of a solution, and we design and test a simple polymerase extension reaction, showing that there may be potential to this approach.

Related Work

Capacitive Sensing of Biological Output Capacitance has been used as a sensor for chemical reactions for many applications in biology, though the minute and sensitive nature of the output has resulted in most projects focusing on custom capacitive sensors. They have been designed for various applications in health, including CTCs, bacterial cells, [Nakanishi et al., 2019]. So far, these have been custom capacitive sensors requiring external power for the VNA.

Recent work has pushed in the direction of using the smartphone touchscreen to sense molecules, relying on an intermediate circuit involving a Voltage Controlled Oscillator (VCO) component to generate frequency-modulated (FM) signals to convey the biological reaction output to the touchscreen [Waghmare et al., 2023b].

Capacitive Sensing of DNA Work has also gone into measuring the capacitance of DNA, building reporter systems that change capacitance in response to an input [Stagni et al., 2006].

More recently, researchers have pushed towards demonstrating capacitive touchscreen-like sensors, in which an ITO patterned grid of electrodes was functionalized with synthesized DNA probes that were designed to bind to viral DNA targets [Lee et al., 2018]. When the probe encountered its intended target, such as a sample with H1N1 virus or M1 influenza DNA, the capacitance at that point in the electrode grid increased, as measured by pulse count. Other DNA reading techniques via capacitance leveraged a field-effect detection or an oscillating wave impulse [Fritz et al., 2002; Stagni et al., 2006].

DNA Strand Displacement (DSD) was introduced in Chapter 2 as a method for building deterministic pathways for DNA to interact that acts as a circuit. Input and output can determined based on specific sequences of DNA "coded" as synthesized DNA strands. The output can be amplified via molecules, such as methyl blue, or via known molecular methods discussed in the next section [Farjami et al., 2010].

Isothermal Amplification Amplification techniques involving enzymes and temperature adjustments are often used to increase the size of a signal from nucleic acid-based reaction. A subset of these techniques can operate at a single held temperature, including recombinase polymerase amplification (RPA), Loop-Mediated Isothermal Amplification, and others [Dhama et al., 2014; Lee, 2017; Li et al., 2019; Lobato and O'Sullivan, 2018]. This has been combined with CRISPR-based targeting, as well [Broughton et al., 2020;

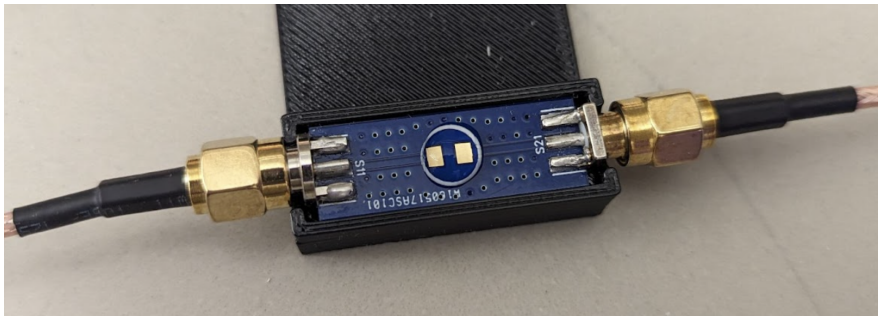


Figure 6.1: The board we designed leaves the circuit open until liquid solution (about 30 L) is placed in the circle. At that point, the signals generated by the VNA run through the solution and the response is measured on the S11 and S21 ports.

Gootenberg et al., 2017]. If the output method can detect changes in concentration, a simpler version of DSD detection or polymerase extension could be used to trigger a change in concentration, or entropy, when a target molecule is recognized by a primer [Hellyer and Nadeau, 2004; Collins et al., 1988; J. Toley et al., 2015].

Smartphones are also known generate heat, as there is a reason the thermistor and cooling system is included to prevent [Kang et al., 2019; Breda et al., 2023]. This may be a factor that can be leveraged in a future experiment to generate heat from the phone by running the processor to generate heat [Dai et al., 2018].

Vector Network Analyzer (VNA) for Biology Commonly used to calibrate antennas, a vector network analyzer (VNA) allows interrogation of charge-related characteristics of a Device Under Test (DUT). It functions by generating an oscillating signal, such as a sine wave, in a sweep of frequencies that is sent into the DUT from its input port (S11) and reading the resulting signal that transfers through the DUT from one of the output ports (S21). The other output port is actually the same as the input port (S11) and it reads the reflected signal. When the transferred and reflected signal is read at each frequency, a unique signature based on capacitance characteristics of the DUT can be inferred.

In previous works, this VNA method has been used for interaction, showing that electrodes formed into a ring device on a finger can detect unique signal changes as the hand moves, enabling hand interaction for virtual reality (VR) environments [Waghmare et al., 2023a]. This has been proposed for sensing biological phenomena [Sahishnavi et al., 2024]. VNA-on-a-chip designs have been created to make this sensing drastically more accessible, such as the case of detecting circulating tumor cells on a small chip [Nakanishi et al., 2019].

6.2 Methods: Raw Touchscreen Data for Biological Applications

6.2.1 Hardware Design: VNA DNA Board for Sensing Capacitance of Solutions

We used a LiteVNA to sense the response of aqueous solutions containing different concentrations of DNA [ZeenKo, 2022]. We designed and built a simple signal pass-through PCB based around a DNA test area, which contained two gold-plated electrodes that form an open circuit until aqueous solution is placed in the area using a pipette. As shown in Figure 6.2, the VNA frequency sweep is injected via Port 1, and the

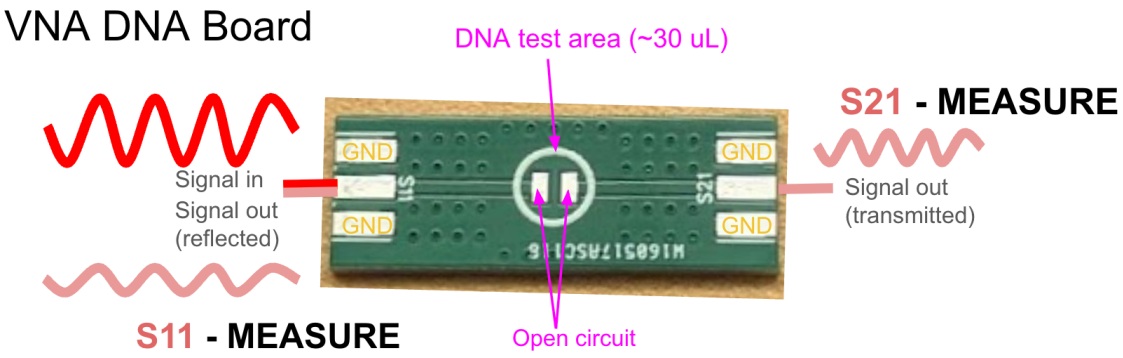


Figure 6.2: The signals measured at each frequency of the VNA freqency sweep are as follows: S21 transmitted signal (or the signal measured at Port 2 sourced from Port 1) and S11 reflected signal (or the signal measured at Port 1 sourced from Port 1)

resultant transmitted signal (S21) is measured from Port 2 and the reflected signal (S11) is measured from Port 1. The frequency sweep was set to sweep at 50 points between

6.2.2 Test Reaction: Polymerase Extension Reaction

Our reaction involved a targeted polymerase extension reaction. We designed an 18 bp primer to recognize one end of a synthetic 90bp DNA oligo as our target (sometimes referred to as template). We included equal concentration of a reaction buffer (NEBuffer r3.1, New England Biolabs) and used an isothermally active polymerase (BST 3.0, New England Biolabs) to produce an extension reaction when the target was introduced. We added dNTPs to the mix as our limiting reactant so that changes in concentration of dNTP are significant compared to the starting concentration. We validated this against negative controls in which an isotonic buffer or a buffer solution containing the wrong 90bp DNA oligo to compare behavior of the reactions. Our oligos were ordered from IDT and used PAGE Gel purification for the longer strands.

Experimental Protocol - Using the VNA Sensor

Our experimental process involves running a reaction and using a pipette to drop a 30 uL size into the DNA test area labeled in Figure 6.2. Care must be taken to have the drop fill the fiducial circle and fully cover the 2 test pads to achieve consistent measurements. We can set the measurement length to a specific number of seconds, and we chose 30 seconds to see a stable result for reactions already complete. For reactions that we expect to be actively occurring while on the VNA board, we set the length of measurement to 10 minutes, and we designed a 3D printed semi-spherical cap, which prevented evaporation of the drop while the reaction occurred.

Data Processing

Custom software in Python was developed for reading data with the LiteVNA system [ZeenKo, 2022]. We developed a codebase with multiple stages of processing. We calculated the magnitude of the signal at a given timepoint by computing the square-root of the sum of squares of the imaginary and real portion of the as shown in Equation 6.1.

$$|S| = \sqrt{I^2 + Q^2} \tag{6.1}$$

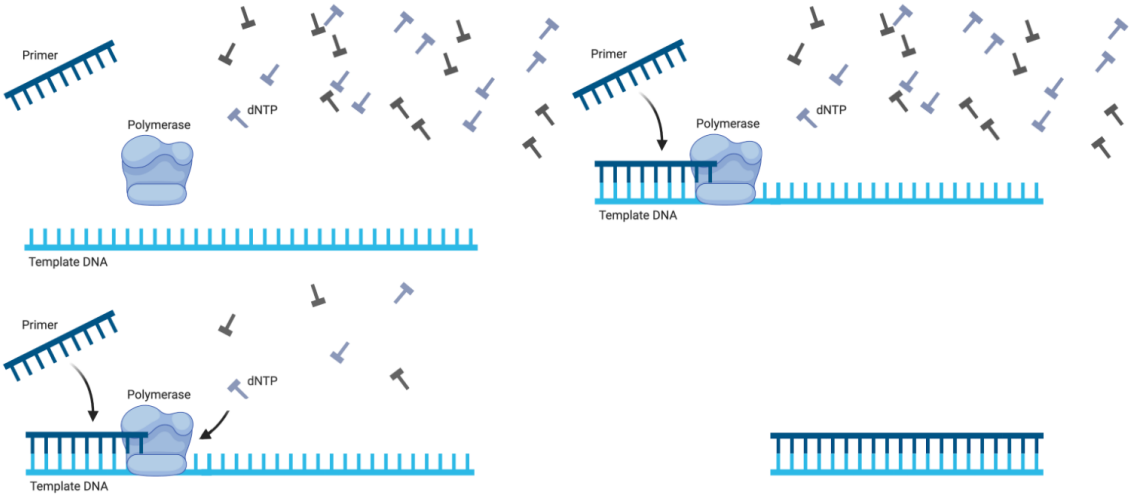


Figure 6.3: Polymerase extension reaction in 4 stages. When a matching template is introduced into a starting mix of primer, polymerase, and excess dNTPs, the primer recognizes the template (or target) and triggers polymerase extension, which consumes dNTPs to produce double-stranded DNA. Images generated using Biorender.

where I=real (in-phase) and Q=imaginary (Quadrature) and the output P is plotted in decibels (dB) by taking the logarithm of the resulting output to amplify small changes, as shown in Equation 6.2.

$$P = 20 * log_{10}(|S|) \tag{6.2}$$

where P is plotted, such as in Figure 6.8. The output of the sensor is a set of frequencies in a range of 50 bins specified prior to the experiment, and we set our measurement frequencies to the range of 50 bins between 1 kHz to 500 kHz (in steps of 9.98 kHz), because smartphone touchscreens are often in that range. We elected to analyze the results at a single frequency of 80.84 KHz because that frequency showed the most separation between different trials.

Currently, the signal is visually inspected at the desired frequency to determine whether a difference in the capacitance is detected, but future iterations would include automatic threshold algorithms to determine differences, which would allow for automatic analysis in scenarios like diagnostics.

6.2.3 Evaluation

We have designed a new board that interfaces with the a card-sized portable VNA (LiteVNA-64) and gives a space to place a liquid solution onto it, as shown in Figure 6.1. It begins with an open circuit with 2 pads, and after some liquid is deposited, that closes the circuit, and the solution effectively becomes the DUT.

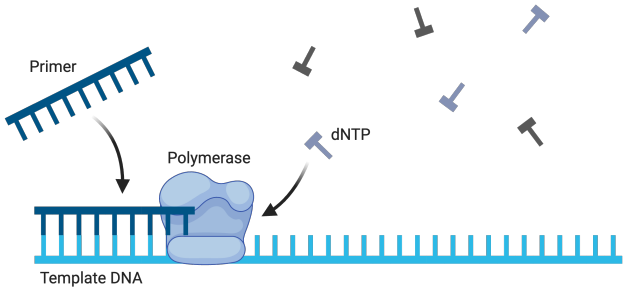


Figure 6.4: A polymerase acts to extend the complementary strand of DNA, beginning from the spot where the primer was bound and pulling in dNTPs to form the other ssDNA. This effectively changes the concentration of total molecules in solution if, and only if, the target template is located by the primer. Image generated using Biorender.

6.3 Results: Feasibility of Touchscreen-Based Capacitance for Sensing Nucleic Acids

Feasibility Results - dNTPs Concentration Differentiation

We have collected several data points doing experiments with an output to the VNA sensor that have shown that we believe this is feasible. Figure 6.5 shows that different concentrations of ssDNA are distinguishable from each other at certain frequencies. In addition, it shows that deoxynucleotide triphosphates (dNTPs) of different concentrations are distinguishable from each other, as well as distinguishable from ssDNA of the same number of base pairs.

This final fact, that dNTPs of the same number of base pairs are distinguishable from ssDNA, reveals that we may be able to design a reaction that converts dNTPs to ssDNA and the signal will probably be visible on the VNA.

Product Results: Polymerase Extension Reaction

For the first test of our system with the polymerase extension reaction, we measured the capacitance after running the polymerase extension reaction for 20 minutes at 65 ∘C. The solution was then placed on the VNA sensor and measured for 30 seconds. Results from the VNA sensor can be seen in Figure 6.6 and were verified via gel as shown in Figure 6.7.

Dynamic Results: Polymerase Extension Reaction

The second test of our system involved testing a dynamic reaction

Figure 6.8

Gel Verification

In order to compare the VNA results to the lab standard results, we ran an E-gel to verify the reaction was happening and saw expected results when the Positive Control primer was used. These gels are seen in Figure 6.7 (correlated with the VNA results in Figure 6.6) and Figure 6.9 (correlated with the VNA results in Figure 6.8).

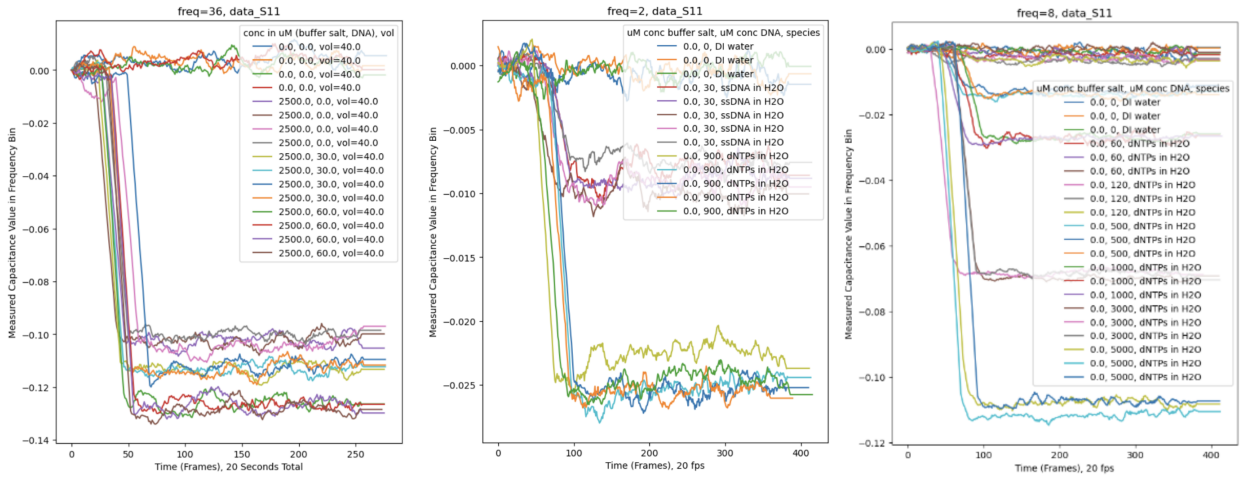


Figure 6.5: Feasibility data demonstrates that different concentrations of ssDNA and dNTPs can be delineated using our VNA setup. **Left:** 3 concentrations of 30 bp ssDNA in ionic solution are differentiable from each other and from DI water. **Center:** dNTPs of 30x the concentration of 30 bp ssDNA strands are differentiable from each other, indicating that this VNA setup works as an "entropy detector" since both solutions have the same physical molecules and they are simply in a different form. **Right:** Different concentrations of dNTPs are differentiable from DI water above a threshold of 500 μM dNTPs.

Negative Control Primer Verification

In order to check the validity of our results and analyze whether our chosen sequence is sufficient to validate this sensor, we ran a flipped experiment where we kept the same input sequence targets (PC and NC), but used the primer that was designed to recognize the NC sequence, which we'll call the "B" primer. As shown in Figure 6.11, the gel showed a strong signal for both the PC and NC template strands in this case, indicating potential problems with our sequence design of the positive control primer. At the same time, the VNA result in Figure 6.10 came back with a similar difference to the previous result with the positive control primer in Figure 6.6. This is inconclusive, and we were unable to investigate further at this time, but we discuss what this could mean for our results, and potential future work, in the Discussion section of this chapter.

Results Taring

Figure 6.12

6.4 Discussion: More Ubiquitous Biological Sensing Challenges and Possibilities

While the path to using a touchscreen capacitive sensor for DNA sensing is still not clear at the time of this work, we made a few contributions that may aid in the development of such a sensing system. We developed a VNA-based biological sample sensor, developed a protocol for its use, developed code for processing and visualizing the data output, and analyzed the output of a few experiments to demonstrate its potential use in further research. Further development and testing using this system could contribute to development of

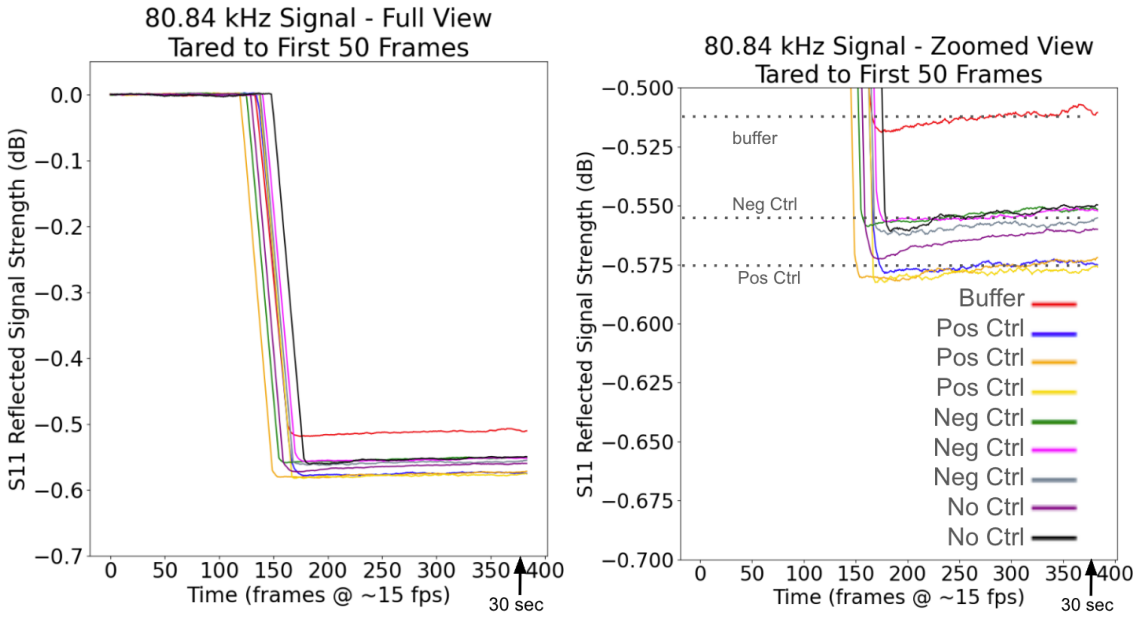


Figure 6.6: Testing the output of positive and negative control solutions on the VNA after running in an incubator for 20 minutes at 60 °C. The positive control has a slightly different reflected signal at 80.84 kHz than the

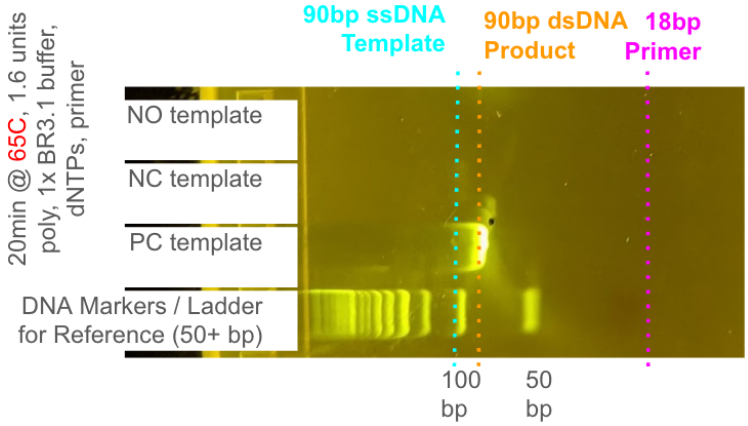


Figure 6.7: Results from the experiment shown in Figure 6.6

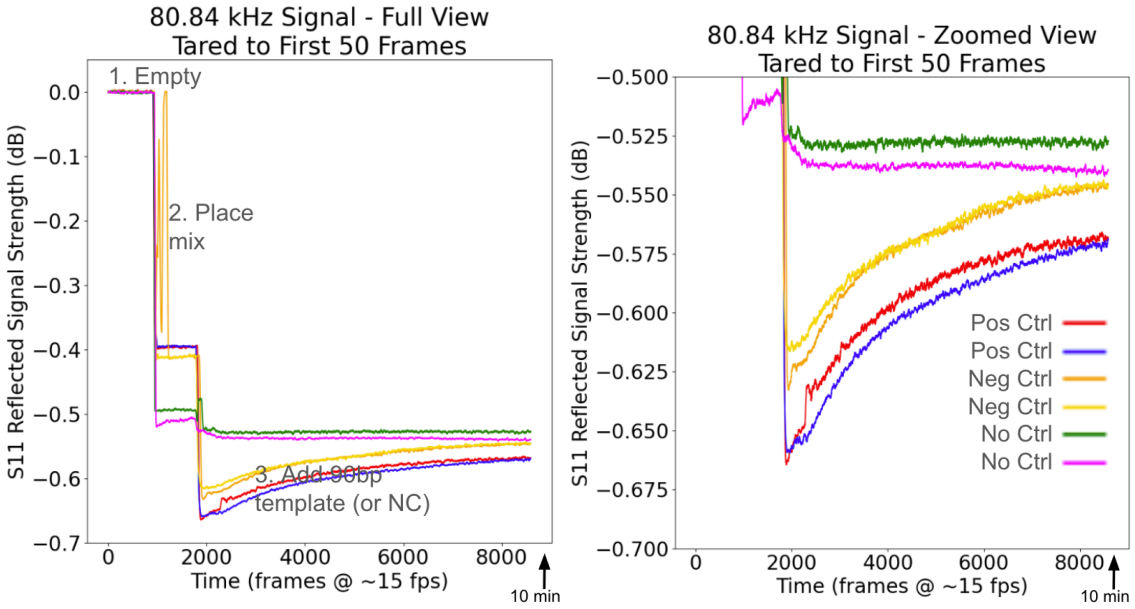


Figure 6.8: Dynamic reaction on VNA without heat over 10 minutes. The positive control shows a different signal than the negative control and no control strands.

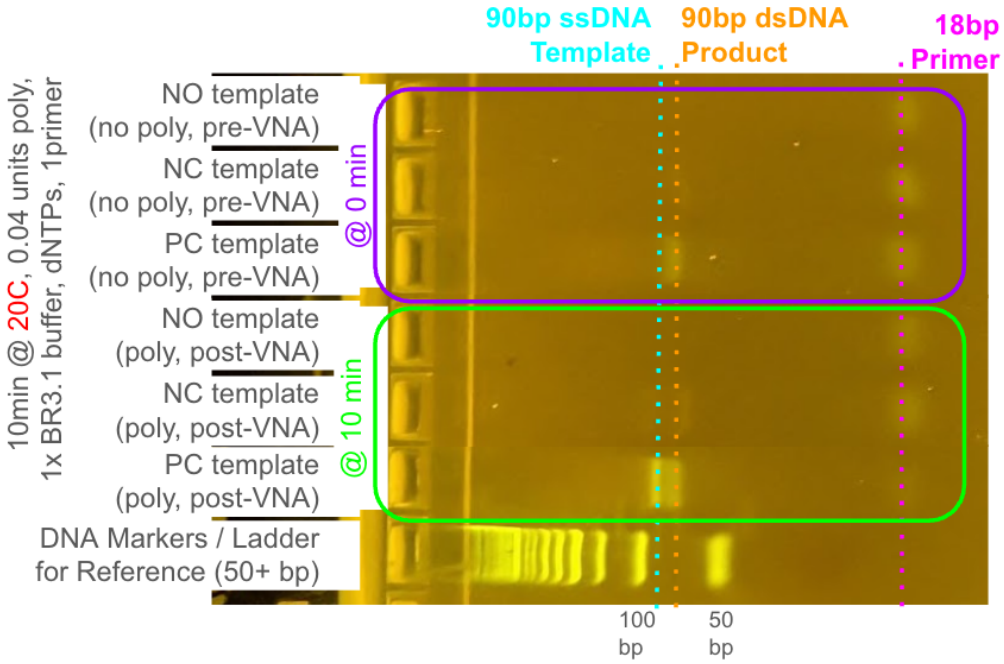


Figure 6.9: Results from the experiment shown in Figure 6.8

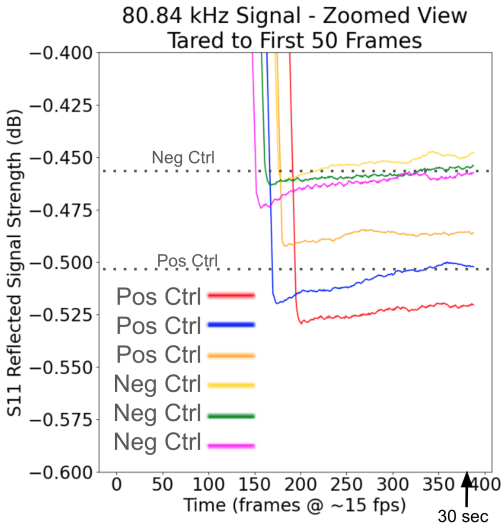


Figure 6.10: Negative Control Primer VNA Result

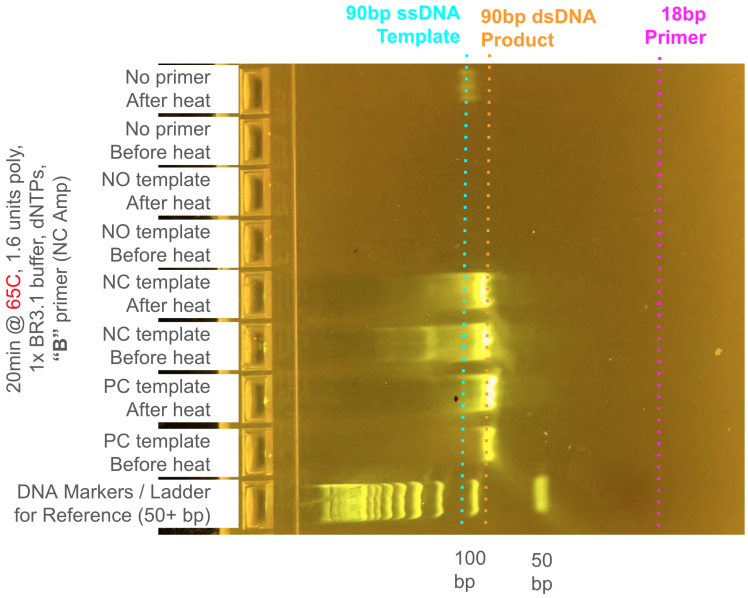


Figure 6.11: Gel results from the negative control template experiment indicate there may be some self-annealing occurring in the positive control template.

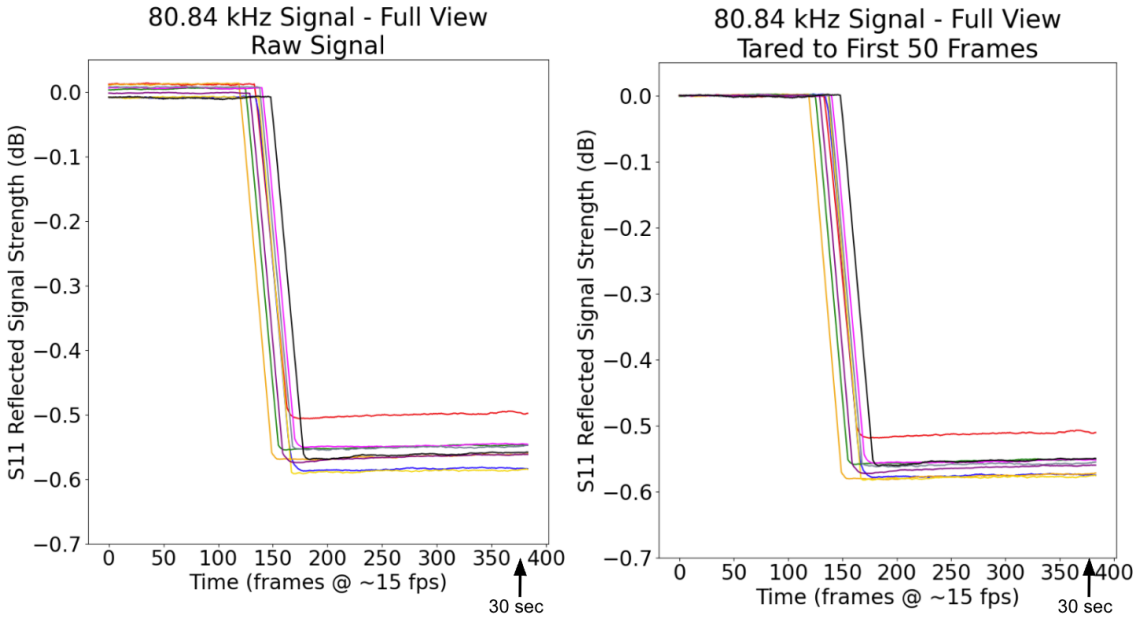


Figure 6.12: The sensor needs to be tared (or set a baseline to 0) before analyzing results. All VNA results int his chapter (including Figure 6.6 and Figure 6.8) are tared by setting the average of the initial 50 frames to 0 in the method illustrated in this figure, but other methods could be explored in future work.

usable in-field sensors leveraging the capacitive touchscreen of smartphones, but we have not proven this yet.

6.4.1 Potential Conclusions

The results identified in the present study are not clear enough to indicate affirmatively that the nucleic acid is being detected by this method.

While it appears that we may be seeing a consistent repeatable difference in signal across replicates in this study, this difference could also be explained by a number of experimental condition factors that do not relate to the desired experimental output. For example, the results observed could be due to slight differences in ionic charge of the buffer based on slight differences in dilution procedure or specified equipment error for pipettes and Nanodrop used in the experiment. Efforts were made to align these solutions as close as possible in their ionic charge for the buffers used, but the differences appear to be within the error range of pipetting and diluting techniques.

The results in the dynamic reaction case do show a difference, but the reaction was not run long enough to fully see the end of the reaction. The behavior of the solution, though it was interesting to see in Figure 6.8 such a difference in the "NO template" case compared to the "NC template" and "PC template" cases, when a similar buffer was inserted at step 3.

6.4.2 Sequence Design Limitations

Results in Figure 6.11 indicate that our original target sequences may have some self-dimerization or other issues that may render our experimental results insufficient to answer our experimental questions. Our sequence was chosen and modified from a sequence seen in literature and spot-checking in NUPACK 4.0

and Benchling showed little signifant self-dimerization [Lee et al., 2018; Fornace et al., 2020]. However, sequence issues are always possible, and further work and testing with designed sequences would enable us to verify this system more fully [Zhang et al., 2018].

6.4.3 Sensitivity Limitations

If further testing is conducted and these results hold up repeatably, then the sensitivity of the reaction should be considered for its applicability to diagnostics. At the current stage, no firm conclusions can be drawn, but it's important to note that the concentrations being used would not be relevant for diagnostics without a nucleic acid amplification step (such as PCR). The most sensitive diagnostics operate at the Limit of Detection (LOD) around 10 copies, and that correlates to a femto-molar range concentration at the volume we're testing, compared to the orders of magnitude larger range we're testing. We saw this sensitivity in our usage of RT-qPCR for viral detection on public transit infrastructure in Chapter 5, and the current state of capacitance-based diagnostics falls short of that standard [Hoffman et al., 2022a]. This detection method could be combined with amplification to enable it to be used, but this would necessitate the required lab equipment and the benefits of this approach would be lost. Therefore, more work needs to be done in order to validate the approach and assess the sensitivity more closely, but this work does not include enough data to indicate that one way or the other. Still, as indicated in the Future Work section of this chapter, there may be other molecular reporter-type systems to design (such as the addition of methyl blue) to amplify the capacitance changes associated with a change in nucleic acid conformation. We are at the early stages of the development of this type of output, so it's unclear whether the limitations of this approach outweigh the potential benefits, but more work to stabilize and characterize the system will provide more clarity.

6.4.4 DNA Sensing Mechanism

Multiple models have been proposed for sensing DNA via its effect on capacitance sensors, and we present our own model in Figure 6.13, based on past work and our preliminary results with our new VNA DNA sensor [Fritz et al., 2002; Stagni et al., 2006]. If we are sensing the change in conformation of the nucleic acid material, we believe it could be due to the change in overall entropy of the system as dNTPs are consumed by the polymerase and formed into a dsDNA product. This reduction in entropy removes molecules floating in solution and attenuates the capacity of the solution to pass a signal from one port to the other. More tests need to be done to confirm this.

6.4.5 Future Work

This work opens up many pathways for future work to contribute to our understanding of capacitance sensing of nucleic acids, working towards sensing systems that may be compatible with the capacitive touchscreens sensors on modern smartphones.

The next set of future work is experimental, with many different combinations of solutions and DNA that could enable us to characterize the ability of our VNA sensor to detect changes in nucleic acids. For example, new sets of target strands and associated primers could be designed that would enable us to see a clear extension reaction occurring in lab-standard methods, such as E-Gels, which would enable us to verify what we are detecting with the VNA. In addition, using new methods to dilute the samples. Finally, comparing with modifications to the input molecules, such as heat-killing the polymerase after mixing (but before extension) would allow us to compare two conditions with identical buffer concentrations. Different

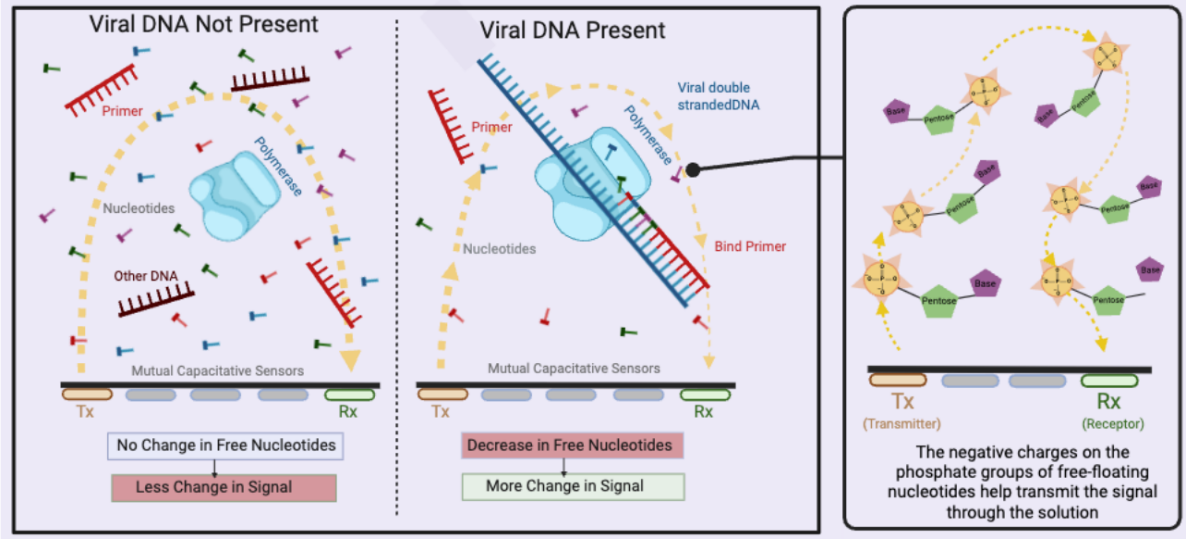


Figure 6.13: Proposed mechanism for VNA sensing of DNA presence in solution

types of reactions could be tested with our new device, including those that work with functionalization of probes directly to the sensor to enhance the signal via proximity.

Finally are modifications to sensor device itself, in either the hardware or the computational software that is used to process through the noise to reveal the signal of interest. It would be interesting to build a microfluidics-based sample holder, which could control more precisely the volume of liquid being placed on the sensor. In addition, we have begun to pursue temperature control for the VNA sensor, though it is not reliably operational at the time of this writing. On the processing side, machine learning could be a useful computational technique for further increasing our signal-to-noise ratio. Prior work used a method of functionalization the recognition or probe strands, and that could be tested if a different reaction were used and a redesign of the board were conducted to measure across the gap differently.

One set of future work is on the data analysis pipeline. There are many factors at play with our sensor, including which frequencies to analyze and even which frequency bins to collect at experiment time. Looking beyond just the magnitude of signal change into the complex and real parts of the signal separately, may reveal insights we were unable to report on with the present study. Repeat testing is critical to ensure robustness, so any experiment that works should be repeated and analyzed in an aggregate manner that illustrates the average and error range to give confidence that the sensor will work into the future. Finally, the aspect of taring (or zero-ing) the sensor is an interesting topic that deserves further study. Is the baseline for the experiment when the sensor is empty (Step 1 in 6.8) and an open circuit is being interrogated? Or is the baseline the moment after the mix is placed (Step 2 in 6.8) and the system is waiting to encounter a target? We could apply machine learning to the entire frequency sweep of the VNA, or to a smaller set of frequencies, which would allow characteristics of the signal that are not apparent to the human eye to be recognized and used to classify results. However, it's important to note that this should not be pursued until we have a better sense of what our VNA sensor is recognizing, as applying machine learning arbitrarily to medical data has a tendency to recognize coincidental markers not associated with the clinically-relevant signal.

There are many interesting directions for future work enabled by this new sensor, and I hope that others may pick up our designs and continue development to fully prove out this tool.

6.5 Conclusion: Towards More Ubiquitous Sensing of Nucleic Acid Reactions

As mentioned in Chapter 2, there are multiple options for molecular reaction output, but I chose to explore capacitive output in my research due to its association with a sensor available on all modern smartphones. While this is not yet proven in full on commodity capacitive sensors, I have demonstrated detection of small changes in solution contents on a lab-based development kit. The experiments I undertook, alongside my collaborators, to test the limits of these commodity sensors with the goal of detection of specific DNA sequences associated with particular viruses demonstrated the final piece of my thesis. We need to develop a biological understanding of the molecules of interest if we hope to sense them on ubiquitous sensors, such as the nearly globally saturated set of touchscreen smartphones that people are already using in their daily lives.

The future work items listed in this chapter also illustrate the thesis. Pursuing these items will enable us to move closer to detection from both sides of the sensing axis in Figure 1.1. We will need to understand capacitance-based nucleic acid sensing systems from the biology side, while also attempting to read the signal through the noise; noise that will likely grow when we move the sensing to the ubiquitous smartphone touchscreen. I hope that future researchers may pick up on this thread to further develop this work for more complex nucleic acid reactions, including DNA computation, and move closer to a reaction that can be sensed by the sensors already in use in modern phones to enable touchscreen interfaces.

In the next chapter, I will conclude my thesis by speculating and reflect upon the broader themes associated with ubiqitous computing and biotechnology. If we can solve problems such as the ones described in the preceding chapters, it could further democratize the methods of biotechnology to more scientists and members of the general population, accelerating the growth of biological research and increasing access to these screening tools and important health information for more people in the future.

Chapter 7

Conclusions and Reflections

7.1 Background: Lessons from Ubiquitous Computing to Biotech

In this chapter, I conclude my thesis, reflecting on learnings from my PhD research and speculating about how it may apply to future research in the health realm of the ubiquitous computing field. My four thesis projects, two in vivo sensing of blood-based health biomarkers and two in vitro nucleic acid sensing, demonstrate ways we can build solutions that enable improved access to sensing health-relevant signals using devices we already have in our pockets. The aim my work has been (and will continue to be) to empower more people to gather important information about their health so that they may act on it and improve their chances at a heatlhy life. More work will be needed in order to bring these innovations to impact healthcare in the wider world, and I find it likely that we will need to enlist the help of an increasing number and broader range of people in order to accomplish these goals. In my experience across computer science and biotechnology, I am struck by the difference in the approachability of computing technology compared to biotechnology. This leads me to envision the opportunity for bridging that gap, which will be critically important as we continue to push both of these fields forward.

7.2 Projects Takeaways for Thesis

In analyzing the four projects together, a picture of how we can pursue research that applies ubiquitous technologies to improve access to healthcare emerges. Both the Introduction in Chapter 1 and also Appendix A give good summaries of my learnings, but here they are summarized before diving into Reflections of what this means for the fields looking into the future.

The four thesis projects taught me different things about developing diagnostics based on ubiquitous commodity sensors. Depicted in Figure 7.1, the two projects on the right (SpO_2 and hemoglobin sensing) using the smartphone camera to sense biomarkers in vivo showed that computational techniques could be used to improve the signal-to-noise ratio and elicit the clinically-relevant signal of interest when the biological and physical basis for those signals is well understood. The two projects on the left (passive viral and capacitance nucleic acid sensing) showed that, in cases where the physics of how the molecule relates to the sensing mechanism is not clear yet, more work should be focused on understanding that signal, so that we may potentially amplify it to be brought into the range of the commodity sensors.

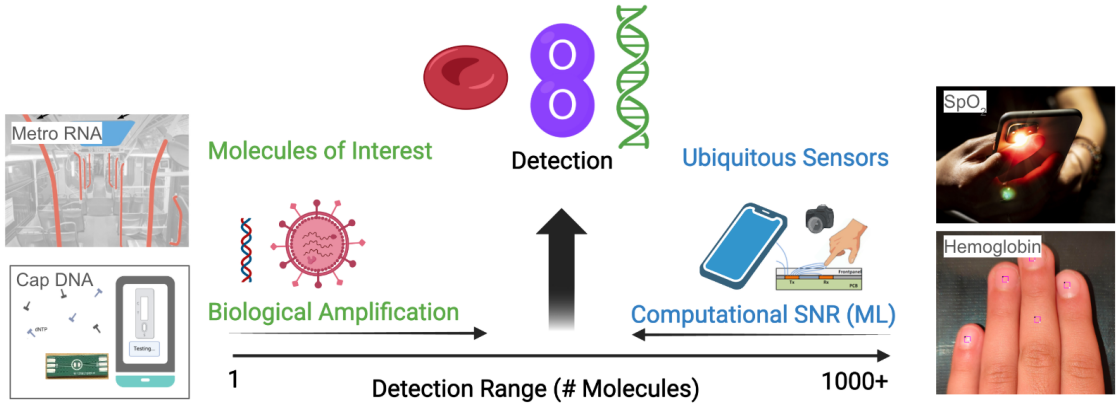


Figure 7.1: The four thesis projects taught me different things about developing diagnostics based on ubiquitous commodity sensors. The two on the right (SpO_2 and hemoglobin) showed that computational techniques, specifically machine learning, can be useful in health sensing using commodity sensors. The two on the left (passive viral and capacitance nucleic acid sensing) showed that the underlying biology needs to be understood before these computing techniques can be effectively applied.

SpO_2

When gathering data for a machine learning based approach to health screening using ubiquitous sensors, leveraging our physiological knowledge and gathering a representative dataset across the range of values that we want to predict is effective in increasing accuracy to the point where it would be medically relevant for screening.

Hemoglobin

In certain cases, we may have the physiological basis for reasoning that machine learning is possible to objectively parse through the noise associated with ubiquitous sensors, but the data that is representative of the range we want to screen for may not always be plausible to to gather. In these cases, different computational techniques, including data balancing, data augmentation, and more complex models may be used in order to increase the accuracy enough to be usable in screening. This can be helpful in indicating a direction to pursue for applications in healthcare scenarios.

Metro Detection of SARS-CoV-2

Data that can be used for environmental monitoring can be gathered by passively leveraging infrastructure present in the public transport systems in a city. However, when it comes to viral monitoring, the lab-based techniques available for sensitive detection for these passive methods are not typically scalable enough for a feasible approach to gather enough data for relatively passive monitoring.

Capacitive Sensing

Before sensing of molecules of interest can be accomplished, the physical and physiological explanation of how that molecule may affect readings on the sensor must be well understood. We can develop tools that emulate the sensor to enhance our understanding of the molecule biologically. In the case of nucleic acids,

we can use these learnings to amplify the signal from the biological level and move towards the possibility of ubiquitous, field-based detection that can be used in passive environmental monitoring and accessible medical screening.

7.3 Reflections on Ubiquitous Biotechnology

My experience pursuing interdisciplinary research at the intersection of biology and computing technology has allowed me some insights into promising directions for further research, particularly in applying ubiquitous computing to biotechnology, but also in applying the process of developing computing to the process of developing biotech.

What is Ubiquitous Computing?

The field of ubiquitous computing within computer science was envisioned by Mark Weiser in 1991 alongside his famous quote describing the destination of all technologies: "The most profound technologies are those that disappear. They weave themselves into the fabric of everyday life until they are indistinguishable from it." [Weiser, 1991]. We have witnessed this vision come to fruition in many cases, not only with ubiquitous smartphones that are now irreplaceable in daily life, but also in the vast number of sensors, displays, controllers, and chips that embed computing "smartness" and connectivity into everyday objects [Rose et al., 2015]. The network of devices and applications continues to grow, and we, ubiquitous computing researchers, continue to develop new ways to solve problems by leveraging these devices. Many of these technological methods, based mostly on the transistor-based digital computer, are now just part of our everyday life, and problems can be solved in new ways, in areas like healthcare, environmental, interaction, accessibility, communication, safety, and many more.

This computing revolution was driven in part by Moore's Law and the reduction in cost and size of computing power [Schaller, 2002]. However, it would be a mistake to look at this trend in a vacuum. Equally, if not more, important was the development of useful applications and uses for computers that drove the consumer demand to enable companies to continue pursuing the realization of Moore's Law. Users saw plentiful reasons to buy computers and weave them into their lives because it solved problems in new and convenient ways. By now, the smartphone has become a virtual necessity in modern life for its multi-purpose utility. This not only makes the case for application-driven development, but also makes the case for adding new uses for existing devices to effectively reduce the overall cost of that function. Users can effectively amortize the cost over each of the functions it serves in their life.

And as more users felt it necessary to have digital computing devices, Moore's Law and manufacturing scale brought the cost down for those devices, to the point where hobbyists and hackers could afford to own more and build with them, democratizing the development so that a plethora of ideas and groups fluorished, as summarized by Walter Isaacson in *The Innovators* [Isaacson, 2014]. This benefited other digital circuit based technologies with a feedback loop increasing their ubiquity. This can be analogized to where we can aspire to be in biotechnology, as well.

7.3.1 Molecular Computing and the MPU

At this time, however, we cannot yet say the same about the affordability and ubiquity of biotechnologies. Progress and breakthroughs have improved human and environmental health, but it is not the case that the technologies of biology are ubiquitous for a wide majority of the population. We may learn biology and organic chemistry in school, but there are not yet the at-home applications, nor the affordable tools

to enable most families to have a "personal lab" (PL) in the way that personal computers "PCs" became ubiquitous. There are some very practical reasons for this, such as the fact that safety is critical in working with chemicals of biology and chemistry. But many of those melt away, at least for specific cases like molecular programming, if we can develop accessible tools and affordable methods of reading biology, such as using the powerful sensors in our smartphones. Some examples of research in that direction exist today, and many more developments are needed to hit a critical acceleration point, like we saw with personal computing.

This hit home in the middle of my PhD work with the COVID-19 pandemic. Because of the devastation of the pandemic and the public health effort to "flatten the curve", virtually everyone in the world learned how to use a one-bit molecular computer. COVID-19 rapid tests became ubiquitous for a time when public health concerns required us to live our lives based on whether a certain molecule was in our respiratory system [Rai et al., 2021]. While we fervently hope nothing like that will ever happen again, experts warn that dangerous pathogens will likely become more common as we head into the future [Harrington et al., 2021; Cunningham and Hopkins, 2023]. In order to be ready, we need to continue to develop more accessible and convenient modes of testing for viral pathogens and other biomarkers, and this will likely accelerate the path we're already on to have biotechnology "weave itself into the fabric of everyday life until it is indistinguishable from it." COVID-19 tests became a part of everyday life due to a pressing need in 2020. What will be next?

Molecular computing has some theoretical advantages over digital computing, but the main drawback that prevents that from making a difference is the bandwidth of information transfer from the molecular realm to the digital realm. Sequencing can almost be considered an exception, but it does require time-consuming sample prep and compute-heavy processing, effectively doing all the compute "work" digitally in silico rather than offloading the molecular work to the processing unit. If the molecular world is to become accessible to a wider swatch of the population, we should seek to perform workloads that make sense "in vivo" in molecular form, and then find ways to rapidly transfer the output of that computing. As illustrated in Figure 7.2, this effectively increases the bandwidth with which we can transfer information from the molecular world to the digital.

Cross-Disciplinary Application-Driven Development

In order to drive this forward, we can take lessons for the development of the molecular computing field from the development of the digital computer. We saw in the twentieth century that a cross-disciplinary environment (like Bell Labs), driven by a strong urgency to solve an application (the telephone) can spur development of new technoloties [Gertner, 2012]. Claude Shannon developed digital logic, and the team of Shockley, Bardeen, and Brattain worked together to develop the transistor which could be combined to implement digital logic. Over time, these innovations led to further development benefiting from other teams of scientists putting together previously disparate ideas to solve real-world problems. In this context, there's a chance that COVID-19 was that initial developmental spur, with cross-disciplinary teams of scientists coming together to solve tough problems, like we tried to do with the Passive Viral RNA Detection project in Chapter 5. My experience in cross-disciplinary teams has been revelatory, and I believe that some of the best innovations come from people who think differently about a problem collaborating together, speaking enough of each other's language to be able to translate across the sociotechnical gap, and taking bold steps to try to solve problems in new ways.

In my research, I have viewed this as the process of solving problems within the context of computing becoming uqiquitous. I have focused on health, and that has led me to a combination of computer science and biotechnology. Biology is a technology (biotech) that has wound itself into daily life so much that it is

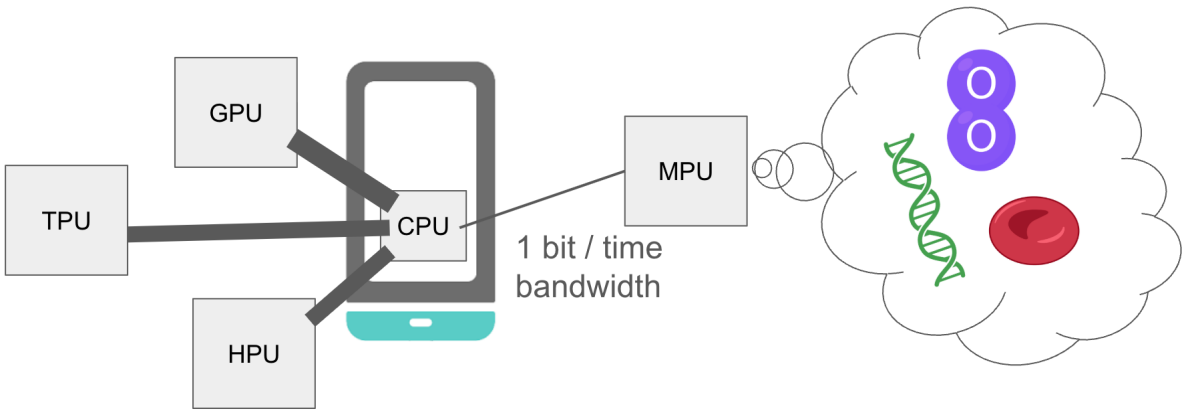


Figure 7.2: The Molecular Programming Unit (MPU) and molecular reactions in general, when applied in ubiquitous contexts such as phones, have relatively low bandwidth to communicate output to the Central Processing Unit (CPU) compared to other more well-developed processing units, like Graphics Processing Units (GPUs), Tensor Processing Units (TPUs), and Holographic Processing Units (HPUs). All development of these was driven by their applications. The bandwidth of the MPU to the CPU will likely be defined by sensors on the smartphone, which will connect with molecular data in new ways.

life. As others have noted, biology's natural information store is DNA, and the analogues between biological processes and computing have inspired a growing body of research. I believe that molecular computing has the foundation to undergo a revolution in the next few decades, and it will be due to the cross-disciplinary collaboration of multiple people and the drive of applications that solve problems.

As an important aside, I'm not ignorant of the fact that I write this at a time when scientific funding and collaboration is being severely undercut in the United States, and it's not without irony that I try to make these conjectures that seemed true only a year ago, but I hope that the political wave subsides the other way and we do see these teams work together again in four years and beyond. I have been fortunate to benefit from wonderful people across the scientific spectrum who lent me their time, energy, and enthusiastic teaching, and my projects would not have been possible without this collaborative environment.

Affordability and Access

One of the applications that is highly important to me is bringing down the cost and increasing affordable access to healthcare. As Rosenthal puts in her book *An American Sickness*, the problem in America is not only whether we can biologically accomplish something, but whether our systems of corporations, employment, and intertwined health insurance qualifications will allow us to use it [Rosenthal, 2017]. My personal belief is that healthcare is a human right, and political shifts need to continue in motion to achieve that reality. Until then, there is opportunity for us, as computing technologists, to create solutions that empower people to take control of their health. Many of my projects work towards that aim, as giving people at-home diagnostics using the devices they already own effectively provides screening low cost, if not effectively free. This is the promise of digital computing in a world where computers are woven into our lives: software can provide medical diagnostics at scale. My thesis works towards those aims, but we must acknowledge the various levels of systems at play to understand how to fully realize that promise.

More affordable access to biotech will likely spurn new developments that we can only dream of; however, I would like to speculate a bit at the end of my thesis and six years of delving into both. If you can

order DNA from a synthesizing company and read data about it on a device you already have, then you can reduce the start-up costs for getting into biotechnology. If more people are in biotechnology, we start to see the flywheel of application-driven democratization we saw with digital computing. I hope that we can avoid some of the pitfalls of the digital economy, such as the notion of *Winners Take All* seen in software "tech" companies when their products gain a small edge and they take the majority of the profits [Giridharadas, 2019]. If biotechnology truly benefits the world, then it should be designed by the majority of people and serve the majority of the people. It should not be designed by global elite with the goal of lining the pockets of CEOs who promise to pay it forward one day, as they rarely do, according to his book.

Outreach and Education

Another major category of applications that I will continue to seek with my work is educational outreach. Education is not the same for everyone, both around the world and also between different areas of the United States. This lack of parity or equity in the education system slows us all down from scientific progress. I feel privileged to have been afforded the opportunities I have to pursue my dream and try to help people with my skills. At the same time, I see potential of many students being undermined everyday in schools where funding is not appropriate or attention is not being paid to all the students. I think it's incredibly important, especially at a publicly funded institution, to enable our innovations and research to positively impact all people in the region and state. I have pursued projects that lend themselves to that well in trying to create inspiring demos and travel to schools to show it off and make it seem more accessible. In addition, I continue to choose projects that I feel can positively impact people who may not have access to healthcare, and I think this will drive both more education and more healthy lives, to fully realize the potential of as many as possible.

Relatedly, I think it's incredibly important to consider that certain health problems have been shown to affect different populations differentially based on socioeconomic and geographical factors [Farmer, 2001]. If we are privileged enough to be able to choose what projects we work on, it is important to choose to spend our efforts to create solutions that can aid those who are underserved by current medical systems. In the best case, we can empower communities to solve their problems in their own way, and I will carry this attitude into my future work as I look toward my next chapter.

7.3.2 Future Work

Many questions were answered with my thesis, but many more questions were generated and remain to be answered. This is the process of science research, and I especially look forward to others building on my work. I made code from my projects open source and attempted to describe my steps in enough detail so others can learn and continue building. I'm going to stay in the space and continue refining my skills and putting my efforts towards building up the potential of concepts of ubiquitous computing applied to biology.

I remain very curious about the pathway for ubiquitous computing innovation to move forward and impact healthcare. This can happen in many ways, whether it's a publication that changes thinking about what is possible or a deployment of a new innovation that is used by people to improve their health. I will continue to pursue these questions throughout my life, and I leave this thesis with a few suggestions for broader work to build on to see these aims achieved.

Statistical Analyses

One aspect that became clear to me as I pursued this cross-disciplinary work was that the statistical methods, or the language of communication of results, differs between fields. Some translation is possible, but we

really should seek to understand the rigorous methods of, for example, the medical field if we're trying to apply computing to medicine. Many scientists are doing so, but it became especially clear to me that machine learning is not going to have the impact in healthcare we see in its potential if we do not go further and validate our studies more fully on statistically significant sample sizes or, of course, on a clinically relevant range of training and evaluation data. As we saw in the hemoglobin sensing work in Chapter 4, data augmentation methods may be good enough to illuminate a path forward for a diagnostic, but they are not enough to replace a lack of full coverage in the data we gather.

Adjusted Specificity

One topic that seems to be indicated as a path for future research is the concept of adjusted specificity. Is there is a level of specificity of diagnostics that can provide a "good enough" balance between utility and specificity for the application? Adjusted sensitivity has been discussed in prior work in the Ubicomp health field in demonstrating that diagnostics. But the following question would be: if we maintain that level of sensitivity by trading off specificity, how much specificity can we sacrifice if the diagnostic has high utility? Health diagnostics are typically optimized for sensitivity and specificity, but, as we established in Chapter 1, utility can be important in providing more accessible tools for diagnostics, as shown in Figure 1.2. I have not fully investigated this question with this thesis work, but I believe it is a topic that deserves further study, especially in relation to resource-constrained situations or in the context of implementation science, where the cost of deploying a new medical innovation is considered in the full calculus of whether a medical intervention is an improvement over the current standard of care.

GlucoScreen

I have been afforded the opportunity to build on these concepts and this work by applying them in an entrepreneurial context. A project on which I am a co-author, but which does not feature in this thesis, is GlucoScreen, and this project has many features that I have striven to pursue. [Waghmare et al., 2023b] This involves sensing a different molecule, glucose concentration in blood, but I will apply many of these concepts that I have learned. I want to see how far I can take this project in providing the type of benefit we describe in our motivation sections of research papers, and I am very excited to see how it goes!

7.4 Conclusion: Ubiquitous Democratizatized Biotech

If we are to reap the benefits promised by ubiquitous computing technology applied to healthcare, we need to build technologies that can be used by more people. Taking lessons from the development of the computer science field, including democratization and wide use of applications of technology, seems to indicate the way forward. If we are driven by clearly valuable applications, we can accelerate development. If we can focus the aims of the research on elements that impact the widest possible set of people, we can democratize the innovations so that all can benefit. My research at the intersection of computer science and biomedical engineering can pave a path towards that future, and I hope to continue developing applications that enable more people to access these exciting innovations and improve healthcare along the way.

Bibliography

Hamada A Aboubakr, Tamer A Sharafeldin, and Sagar M Goyal. 2020. Stability of sars-cov-2 and other coronaviruses in the environment and on common touch surfaces and the influence of climatic conditions: A review. *Transboundary and emerging diseases*.

Warish Ahmed, Paul M. Bertsch, Nicola Angel, Kyle Bibby, Aaron Bivins, Leanne Dierens, Janette Edson, John Ehret, Pradip Gyawali, Kerry A. Hamilton, Ian Hosegood, Philip Hugenholtz, Guangming Jiang, Masaaki Kitajima, Homa T. Sichani, Jiahua Shi, Katja M. Shimko, Stuart L. Simpson, Wendy J. M. Smith, Erin M. Symonds, Kevin V. Thomas, Rory Verhagen, Julian Zaugg, and Jochen F. Mueller. Detection of SARS-CoV-2 RNA in commercial passenger aircraft and cruise ship wastewater: A surveillance tool for assessing the presence of COVID-19 infected travellers. 27(5).

John C Alexander, Abu Minhajuddin, and Girish P Joshi. 2017. Comparison of smartphone application-based vital sign monitors without external hardware versus those used in clinical practice: a prospective trial. *Journal of clinical monitoring and computing*, 31(4):825–831.

Ned Augenblick, Jonathan T Kolstad, Ziad Obermeyer, and Ao Wang. 2020. Group testing in a pandemic: The role of frequent testing, correlated risk, and machine learning. Technical report, National Bureau of Economic Research.

Stefan Badelt, Casey Grun, Karthik V. Sarma, Brian Wolfe, Seung Woo Shin, and Erik Winfree. 2020. A domain-level DNA strand displacement reaction enumerator allowing arbitrary non-pseudoknotted secondary structures. *Journal of The Royal Society Interface*, 17(167):20190866. Publisher: Royal Society.

Sung Hwan Bae, Heidi Shin, Ho-Young Koo, Seung Won Lee, Jee Myung Yang, and Dong Keon Yon. 2020. Early Release - Asymptomatic Transmission of SARS-CoV-2 on Evacuation Flight - Volume 26, Number 11—November 2020 - Emerging Infectious Diseases journal - CDC.

Yan Bai, Lingsheng Yao, Tao Wei, Fei Tian, Dong-Yan Jin, Lijuan Chen, and Meiyun Wang. 2020. Presumed Asymptomatic Carrier Transmission of COVID-19. *JAMA*, 323(14):1406–1407.

Marilyn S Bartlett, Sten H Vermund, Robert Jacobs, Pamela J Durant, Margaret M Shaw, James W Smith, Xing Tang, Jang-Jih Lu, Baozheng Li, Shaoling Jin, et al. 1997. Detection of pneumocystis carinii dna in air samples: likely environmental risk to susceptible persons. *Journal of clinical microbiology*, 35(10):2511–2513.

Paul B Batchelder and Dena M Raley. 2007. Maximizing the laboratory setting for testing devices and understanding statistical output in pulse oximetry. *Anesthesia & Analgesia*, 105(6):S85–S94.

John Beard. 2003. Iron deficiency alters brain development and functioning. *The Journal of nutrition*, 133(5):1468S–1472S.

Kimberly L. Berk, Steven M. Blum, Vanessa L. Funk, Yuhua Sun, In-Young Yang, Mark V. Gostomski, Pierce A. Roth, Alvin T. Liem, Peter A. Emanuel, Michael E. Hogan, Aleksandr E. Miklos, and Matthew W. Lux. 2021. Rapid visual authentication based on dna strand displacement. *ACS Applied Materials & Interfaces*, 13(16):19476–19486.

Philip E Bickler, John R Feiner, Michael S Lipnick, Paul Batchelder, David B MacLeod, and John W Severinghaus. 2017. Effects of acute, profound hypoxia on healthy humans: implications for safety of tests evaluating pulse oximetry or tissue oximetry performance. *Anesthesia & Analgesia*, 124(1):146–153.

J Martin Bland and Douglas G Altman. 1999. Measuring agreement in method comparison studies. *Statistical methods in medical research*, 8(2):135–160.

Joseph Breda, Mastafa Springston, Alex Mariakakis, and Shwetak Patel. 2023. FeverPhone: Accessible Core-Body Temperature Sensing for Fever Monitoring Using Commodity Smartphones. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 7(1):3:1–3:23.

James P. Broughton, Xianding Deng, Guixia Yu, Clare L. Fasching, Venice Servellita, Jasmeet Singh, Xin Miao, Jessica A. Streithorst, Andrea Granados, Alicia Sotomayor-Gonzalez, Kelsey Zorn, Allan Gopez, Elaine Hsu, Wei Gu, Steve Miller, Chao-Yang Pan, Hugo Guevara, Debra A. Wadford, Janice S. Chen, and Charles Y. Chiu. 2020. CRISPR–Cas12-based detection of SARS-CoV-2. *Nature Biotechnology*, 38(7):870–874. Bandiera_abtest: a Cg_type: Nature Research Journals Number: 7 Primary_atype: Research Publisher: Nature Publishing Group Subject_term: Clinical microbiology;Infectious-disease diagnostics;Infectious diseases;SARS-CoV-2;Viral infection Subject_term_id: clinical-microbiology;infectious-disease-diagnostics;infectious-diseases;sars-cov-2;viral-infection.

Suzana Brown and Alan Mickelson. 2019. Why some well-planned and community-based ictd interventions fail. *Information Technologies & International Development*, 15:13.

Annie Browne, Sacha St-Onge Ahmad, Charles R. Beck, and Jonathan S. Nguyen-Van-Tam. 2016. The roles of transportation and transportation hubs in the propagation of influenza and coronaviruses: A systematic review. *Journal of Travel Medicine*, 23(1).

Nam Bui, Anh Nguyen, Phuc Nguyen, Hoang Truong, Ashwin Ashok, Thang Dinh, Robin Deterding, and Tam Vu. 2020a. Smartphone-Based SpO $_2$ Measurement by Exploiting Wavelengths Separation and Chromophore Compensation. *ACM Transactions on Sensor Networks*, 16(1):1–30.

Nam Bui, Anh Nguyen, Phuc Nguyen, Hoang Truong, Ashwin Ashok, Thang Dinh, Robin Deterding, and Tam Vu. 2020b. Smartphone-based spo2 measurement by exploiting wavelengths separation and chromophore compensation. *ACM Transactions on Sensor Networks (TOSN)*, 16(1):1–30.

Giorgio Buonanno, Lidia Morawska, and Luca Stabile. 2020. Quantitative assessment of the risk of airborne transmission of SARS-CoV-2 infection: Prospective and retrospective applications. *medRxiv*, page 2020.06.01.20118984.

Z Cao, HY Tseng, K Salvante, P Nepomnaschy, and MA Parameswaran. 2016. High sensitivity fluorescence detection using smart phone cameras. In *2016 IEEE SENSORS*, pages 1–3. IEEE.

Nicolas Cardozo, Karen Zhang, Katie Doroschak, Aerilynn Nguyen, Zoheb Siddiqui, Karin Strauss, Luis Ceze, and Jeff Nivala. 2019. Multiplexed direct detection of barcoded protein reporters on a nanopore array. *bioRxiv*, page 837542.

Domenico Luca Carni, Domenico Grimaldi, Paolo F Sciammarella, Francesco Lamonaca, and Vitaliano Spagnuolo. 2016. Setting-up of ppg scaling factors for spo2% evaluation by smartphone. In *2016 IEEE International Symposium on Medical Measurements and Applications (MeMeA)*, pages 1–5. IEEE.

Meredith A. Case, Holland A. Burwick, Kevin G. Volpp, and Mitesh S. Patel. 2015. Accuracy of Smartphone Applications and Wearable Devices for Tracking Physical Activity Data. *JAMA*, 313(6):625–626.

Marlin Wayne Causey, Seth Miller, Andrew Foster, Alec Beekley, David Zenger, and Matthew Martin. 2011. Validation of noninvasive hemoglobin measurements using the masimo radical-7 sphb station. *The American journal of surgery*, 201(5):592–598.

Centers for Disease Control and Prevention. 2020. Cdc covid data tracker. Accessed: 2021–12-20.

Luis Ceze, Jeff Nivala, and Karin Strauss. 2019. Molecular digital data storage using DNA. *Nature Reviews Genetics*, 20(8):456–466. Publisher: Nature Publishing Group.

Vikram Chandrasekaran, Ram Dantu, Srikanth Jonnada, Shanti Thiyagaraja, and Kalyan Pathapati Subbu. 2012. Cuffless differential blood pressure estimation using smart phones. *IEEE Transactions on Biomedical Engineering*, 60(4):1080–1089.

Camila M Chaparro and Parminder S Suchdev. 2019. Anemia epidemiology, pathophysiology, and etiology in low-and middle-income countries. *Annals of the new York Academy of Sciences*, 1450(1):15–31.

Nitesh V Chawla, Kevin W Bowyer, Lawrence O Hall, and W Philip Kegelmeyer. 2002. Smote: synthetic minority over-sampling technique. *Journal of artificial intelligence research*, 16:321–357.

Tianqi Chen and Carlos Guestrin. 2016. Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, pages 785–794.

Yeechi Chen, Keiko Munechika, and David S. Ginger. 2007. Dependence of Fluorescence Intensity on the Spectral Overlap between Fluorophores and Plasmon Resonant Single Silver Nanoparticles. *Nano Letters*, 7(3):690–696. Publisher: American Chemical Society.

P. W. Chiang, W. J. Song, K. Y. Wu, J. R. Korenberg, E. J. Fogel, M. L. Van Keuren, D. Lashkari, and D. M. Kurnit. 1996. Use of a fluorescent-PCR reaction to detect genomic sequence copy number and transcriptional abundance. *Genome Research*, 6(10):1013–1026. Company: Cold Spring Harbor Laboratory Press Distributor: Cold Spring Harbor Laboratory Press Institution: Cold Spring Harbor Laboratory Press Label: Cold Spring Harbor Laboratory Press Publisher: Cold Spring Harbor Lab.

Francesco Chirico, Angelo Sacco, Nicola Luigi Bragazzi, and Nicola Magnavita. 2020. Can Air-Conditioning Systems Contribute to the Spread of SARS/MERS/COVID-19 Infection? Insights from a Rapid Review of the Literature. *International Journal of Environmental Research and Public Health*, 17(17).

Clinimark. 2010 (accessed August 9, 2020). *Pulse Oximetry Product Development Testing*. http://www.clinimark.com/TABS%20Testing%20&%20Services/Pulse%20Oximetry/Pulse%20Oximetry_Product%20Development%20Testing.html.

Alan R Cohen and Jessica Seidl-Friedman. 1988. Hemocue® system for hemoglobin measurement: evaluation in anemic and nonanemic children. *American journal of clinical pathology*, 90(3):302–305.

Mary Collins, Edward F. Fritsch, Marian S. Ellwood, Steven E. diamond, Jon I. Williams, and J. Grant Brewen. 1988. A novel diagnostic method based on DNA strand displacement. *Molecular and Cellular Probes*, 2(1):15–30.

Neil Cunningham and Susan Hopkins. 2023. Lessons identified for a future pandemic. *Journal of Antimicrobial Chemotherapy*, 78(Supplement_2):ii43–ii49.

Yuting Dai, Tao Li, Benyong Liu, Mingcong Song, and Huixiang Chen. 2018. Exploiting Dynamic Thermal Energy Harvesting for Reusing in Smartphone with Mobile Applications. In *Proceedings of the Twenty-Third International Conference on Architectural Support for Programming Languages and Operating Systems*, ASPLOS '18, pages 243–256, New York, NY, USA. Association for Computing Machinery.

Christian G. Daughton. 2020. Wastewater surveillance for population-wide Covid-19: The present and future. *Science of The Total Environment*, 736:139631.

Kuldeep Dhama, K Karthik, Sandip Chakraborty, Ruchi Tiwari, Sanjay Kapoor, Amit Kumar, and Prasad Thomas. 2014. Loop-mediated isothermal amplification of DNA (LAMP): a new diagnostic tool lights the world of diagnosis of animal and human pathogens: a review. *Pakistan journal of biological sciences*, 17(2):151–166.

Giovanni Dimauro, Danilo Caivano, and Francesco Girardi. 2018. A new method and a non-invasive device to estimate anemia based on digital images of the conjunctiva. *Ieee Access*, 6:46968–46975.

Xinyi Ding, Damoun Nassehi, and Eric C Larson. 2018. Measuring oxygen saturation with smartphone cameras using convolutional neural networks. *IEEE journal of biomedical and health informatics*, 23(6):2603–2610.

Xinyi Ding, Damoun Nassehi, and Eric C. Larson. 2019. Measuring Oxygen Saturation With Smartphone Cameras Using Convolutional Neural Networks. *IEEE Journal of Biomedical and Health Informatics*, 23(6):2603–2610. Conference Name: IEEE Journal of Biomedical and Health Informatics.

Centers for Disease Control and Prevention. 2020. Cdc 2019-novel coronavirus (2019-ncov) real-time rt-pcr diagnostic panel. `https://www.fda.gov/media/134922/download`. Accessed: 2021–04-30.

Jiantong Dong, Michael P O'Hagan, and Itamar Willner. 2022. Switchable and dynamic g-quadruplexes and their applications. *Chemical Society Reviews*.

Diana Duong and Lauren Vogel. 2023. Overworked health workers are "past the point of exhaustion". *Canadian Medical Association Journal*, 195(8):E309–E310.

Michael Eisenstein. 2012. Oxford Nanopore announcement sets sequencing sector abuzz. *Nature Biotechnology*, 30(4):295–296. Publisher: Nature Publishing Group.

Bradley J. Erickson, Panagiotis Korfiatis, Zeynettin Akkus, and Timothy L. Kline. 2017. Machine Learning for Medical Imaging. *RadioGraphics*, 37(2):505–515. Publisher: Radiological Society of North America.

Elaheh Farjami, Lilia Clima, Kurt V. Gothelf, and Elena E. Ferapontova. 2010. DNA interactions with a Methylene Blue redox indicator depend on the DNA length and are sequence specific. *Analyst*, 135(6):1443–1448. Publisher: Royal Society of Chemistry.

Paul Farmer. 2001. *Infections and inequalities: The modern plagues*. Univ of California Press.

John R Feiner, John W Severinghaus, and Philip E Bickler. 2007. Dark skin decreases the accuracy of pulse oximeters at low oxygen saturation: the effects of oximeter probe type and gender. *Anesthesia & Analgesia*, 105(6):S18–S23.

U.S. Food and Drug Administration. 2021. Emergency use authorization (eua) of the amazon multi-target sars-cov-2 real-time rt-pcr test. https://www.fda.gov/media/151456/download. Accessed: 2021–10-25.

US Food, Drug Administration, et al. 2013. Pulse oximeters—premarket notification submissions [510 (k) s]: guidance for industry and food and drug administration staff. *US Department of Health and Human Services*.

Mark E. Fornace, Nicholas J. Porubsky, and Niles A. Pierce. 2020. A Unified Dynamic Programming Framework for the Analysis of Interacting Nucleic Acid Strands: Enhanced Models, Scalability, and Speed. *ACS Synthetic Biology*, 9(10):2665–2678.

Jürgen Fritz, Emily B. Cooper, Suzanne Gaudet, Peter K. Sorger, and Scott R. Manalis. 2002. Electronic detection of DNA by its intrinsic molecular charge. *Proceedings of the National Academy of Sciences*, 99(22):14142–14146. Publisher: Proceedings of the National Academy of Sciences.

Patrick G Gallagher. 2022. Anemia in the pediatric patient. *Blood, The Journal of the American Society of Hematology*, 140(6):571–593.

Sanjiv Sam Gambhir, T Jessie Ge, Ophir Vermesh, and Ryan Spitler. 2018. Toward achieving precision health. *Science translational medicine*, 10(430).

Angèle Gayet-Ageron, Stephan Lautenschlager, Béatrice Ninet, Thomas V. Perneger, and Christophe Combescure. 2013. Sensitivity, specificity and likelihood ratios of PCR in the diagnosis of syphilis: a systematic review and meta-analysis. *Sexually Transmitted Infections*, 89(3):251–256. Publisher: BMJ Publishing Group Ltd Section: Epidemiology.

Jon Gertner. 2012. *The idea factory: Bell Labs and the great age of American innovation*. Penguin.

Anand Giridharadas. 2019. *Winners take all: The elite charade of changing the world*. Vintage.

Jonathan S. Gootenberg, Omar O. Abudayyeh, Jeong Wook Lee, Patrick Essletzbichler, Aaron J. Dy, Julia Joung, Vanessa Verdine, Nina Donghia, Nichole M. Daringer, Catherine A. Freije, Cameron Myhrvold, Roby P. Bhattacharyya, Jonathan Livny, Aviv Regev, Eugene V. Koonin, Deborah T. Hung, Pardis C. Sabeti, James J. Collins, and Feng Zhang. 2017. Nucleic acid detection with CRISPR-Cas13a/C2c2. *Science*, 356(6336):438–442. Publisher: American Association for the Advancement of Science Section: Reports.

Daniel Gordon, Jason Hoffman, Keren Gamrasni, Yotam Barlev, Alex Levine, Tamar Landau, Ronen Shpiegel, Avishai Lahad, Ariel Koren, Carina Levin, et al. 2024. Artificial intelligence-enabled non-invasive ubiquitous anemia screening: The hemo-ai pilot study on pediatric population. *Digital Health*, 10:20552076241297057.

Eric Raphael Gottlieb, Jennifer Ziegler, Katharine Morley, Barret Rush, and Leo Anthony Celi. 2022. Assessment of racial and ethnic differences in oxygen supplementation among patients in the intensive care unit. *JAMA Internal Medicine*.

Lilian de Greef, Mayank Goel, Min Joon Seo, Eric C. Larson, James W. Stout, James A. Taylor, and Shwetak N. Patel. 2014. Bilicam: using mobile phones to monitor newborn jaundice. In *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing - UbiComp '14 Adjunct*, pages 331–342, Seattle, Washington. ACM Press.

Tobias Grosse-Puppendahl, Christian Holz, Gabe Cohn, Raphael Wimmer, Oskar Bechtold, Steve Hodges, Matthew S. Reynolds, and Joshua R. Smith. 2017. Finding Common Ground: A Survey of Capacitive Sensing in Human-Computer Interaction. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, CHI '17, pages 3293–3315, New York, NY, USA. Association for Computing Machinery.

Zhen-Dong Guo, Zhong-Yi Wang, Shou-Feng Zhang, Xiao Li, Lin Li, Chao Li, Yan Cui, Rui-Bin Fu, Yun-Zhu Dong, Xiang-Yang Chi, Meng-Yao Zhang, Kun Liu, Cheng Cao, Bin Liu, Ke Zhang, Yu-Wei Gao, Bing Lu, and Wei Chen. 2020. Aerosol and Surface Distribution of Severe Acute Respiratory Syndrome Coronavirus 2 in Hospital Wards, Wuhan, China, 2020. *Emerging Infectious Diseases*, 26(7).

Walter N Harrington, Christina M Kackos, and Richard J Webby. 2021. The evolution and future of influenza pandemic preparedness. *Experimental & molecular medicine*, 53(5):737–749.

Abigail P. Harvey, Erica R. Fuhrmeister, Molly E. Cantrell, Ana K. Pitol, Jenna M. Swarthout, Julie E. Powers, Maya L. Nadimpalli, Timothy R. Julian, and Amy J. Pickering. Longitudinal Monitoring of SARS-CoV-2 RNA on High-Touch Surfaces in a Community Setting. 8(2):168–175.

Md Kamrul Hasan, Md Hasanul Aziz, Md Ishrak Islam Zarif, Mahmudul Hasan, Mma Hashem, Shion Guha, Richard R. Love, and Sheikh Ahamed. 2021. Noninvasive Hemoglobin Level Prediction in a Mobile Phone Environment: State of the Art Review and Recommendations. *JMIR mHealth and uHealth*, 9(4):e16806.

Tobin J Hellyer and James G Nadeau. 2004. Strand displacement amplification: a versatile tool for molecular diagnostics. *Expert Review of Molecular Diagnostics*, 4(2):251–261. Publisher: Taylor & Francis _eprint: https://doi.org/10.1586/14737159.4.2.251.

Sebastian Hoehl, Onur Karaca, Niko Kohmer, Sandra Westhaus, Jürgen Graf, Udo Goetsch, and Sandra Ciesek. 2020. Assessment of SARS-CoV-2 Transmission on an International Flight and Among a Tourist Group. *JAMA Network Open*, 3(8):e2018044–e2018044.

Jason S. Hoffman, Matthew Hirano, Nuttada Panpradist, Joseph Breda, Parker Ruth, Yuanyi Xu, Jonathan Lester, Bichlien H. Nguyen, Luis Ceze, and Shwetak N. Patel. 2022a. Passively sensing SARS-CoV-2 RNA in public transit buses. *Science of The Total Environment*, 821:152790.

Jason S. Hoffman, Varun K. Viswanath, Caiwei Tian, Xinyi Ding, Matthew J. Thompson, Eric C. Larson, Shwetak N. Patel, and Edward J. Wang. 2022b. Smartphone camera oximetry in an induced hypoxemia study. *npj Digital Medicine*, 5(1):1–10. Publisher: Nature Publishing Group.

Helene Holmgren, Evert Ljungström, Ann-Charlotte Almstrand, Björn Bake, and Anna-Carin Olin. 2010. Size distribution of exhaled particles in the range from 0.01 to 2.0 μm. *Journal of Aerosol Science*, 41(5):439–446.

Carlos Honrado and Tao Dong. 2014. A capacitive touch screen sensor for detection of urinary tract infections in portable biomedical devices. *Sensors*, 14(8):13851–13862.

B. L. Horecker. 1943. THE ABSORPTION SPECTRA OF HEMOGLOBIN AND ITS DERIVATIVES IN THE VISIBLE AND NEAR INFRA-RED REGIONS. *Journal of Biological Chemistry*, 148(1):173–183.

Sebastian Horstmann, Cassi J Henderson, Elizabeth AH Hall, and Ronan Daly. 2021. Capacitive touchscreen sensing-a measure of electrolyte conductivity. *Sensors and Actuators B: Chemical*, 345:130318.

Patrick F. Horve, Leslie Dietz, Mark Fretz, David A. Constant, Andrew Wilkes, John M. Townes, Robert G. Martindale, William B. Messer, and Kevin Van Den Wymelenberg. 2020. Identification of SARS-CoV-2 RNA in Healthcare Heating, Ventilation, and Air Conditioning Units. *medRxiv*, page 2020.06.26.20141085.

Maogui Hu, Hui Lin, Jinfeng Wang, Chengdong Xu, Andrew J. Tatem, Bin Meng, Xin Zhang, Yifeng Liu, Pengda Wang, Guizhen Wu, Haiyong Xie, and Shengjie Lai. 2020a. The risk of COVID-19 transmission in train passengers: An epidemiological and modelling study. *Clinical Infectious Diseases: An Official Publication of the Infectious Diseases Society of America*.

Zhiliang Hu, Ci Song, Chuanjun Xu, Guangfu Jin, Yaling Chen, Xin Xu, Hongxia Ma, Wei Chen, Yuan Lin, Yishan Zheng, Jianming Wang, Zhibin Hu, Yongxiang Yi, and Hongbing Shen. 2020b. Clinical characteristics of 24 asymptomatic infections with COVID-19 screened among close contacts in Nanjing, China. *Science China. Life Sciences*, 63(5):706–711.

Jacqueline Hudson, SM Nguku, Jules Sleiman, Walter Karlen, GA Dumont, CL Petersen, CB Warriner, and J Mark Ansermino. 2012. Usability testing of a prototype phone oximeter with healthcare providers in high-and low-medical resource environments. *Anaesthesia*, 67(9):957–967.

Hanalise V. Huff and Avantika Singh. 2020. Asymptomatic transmission during the COVID-19 pandemic and implications for public health strategies. *Clinical Infectious Diseases: An Official Publication of the Infectious Diseases Society of America*.

Brady Hunt, Samuel S Streeter, Alberto J Ruiz, M Shane Chapman, and Brian W Pogue. 2021. Ultracompact fluorescence smartphone attachment using built-in optics for protoporphyrin-ix quantification in skin. *Biomedical Optics Express*, 12(11):6995–7008.

C. Höchsmann, R. Knaier, J. Eymann, J. Hintermann, D. Infanger, and A. Schmidt-Trucksäss. 2018. Validity of activity trackers, smartphones, and phone applications to measure steps in various walking conditions. *Scandinavian Journal of Medicine & Science in Sports*, 28(7):1818–1827. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/sms.13074.

Walter Isaacson. 2014. *The innovators: How a group of hackers, geniuses, and geeks created the digital revolution*. Simon and Schuster.

Kevin Jones, Paul Cassidy, Jeremy Killen, and Helen Ellis. 2003. The feasibility and usefulness of oximetry measurements in primary care. *Primary Care Respiratory Journal*, 12(1):4–6.

Taylor B Jordan, Cody L Meyers, Walter A Schrading, and John P Donnelly. 2020. The utility of iphone oximetry apps: a comparison with standard pulse oximetry measurement in the emergency department. *The American journal of emergency medicine*, 38(5):925–928.

Geeta Joshi, Aditi Jain, Shalini Reddy Araveeti, Sabina Adhikari, Harshit Garg, and Mukund Bhandari. 2024. Fda-approved artificial intelligence and machine learning (ai/ml)-enabled medical devices: an updated landscape. *Electronics*, 13(3):498.

Norman P. Jouppi, Cliff Young, Nishant Patil, David Patterson, Gaurav Agrawal, Raminder Bajwa, Sarah Bates, Suresh Bhatia, Nan Boden, Al Borchers, Rick Boyle, Pierre-luc Cantin, Clifford Chao, Chris Clark, Jeremy Coriell, Mike Daley, Matt Dau, Jeffrey Dean, Ben Gelb, Tara Vazir Ghaemmaghami, Rajendra Gottipati, William Gulland, Robert Hagmann, C. Richard Ho, Doug Hogberg, John Hu, Robert Hundt, Dan Hurt, Julian Ibarz, Aaron Jaffey, Alek Jaworski, Alexander Kaplan, Harshit Khaitan, Daniel Killebrew, Andy Koch, Naveen Kumar, Steve Lacy, James Laudon, James Law, Diemthu Le, Chris Leary, Zhuyuan Liu, Kyle Lucke, Alan Lundin, Gordon MacKean, Adriana Maggiore, Maire Mahony, Kieran Miller, Rahul Nagarajan, Ravi Narayanaswami, Ray Ni, Kathy Nix, Thomas Norrie, Mark Omernick, Narayana Penukonda, Andy Phelps, Jonathan Ross, Matt Ross, Amir Salek, Emad Samadiani, Chris Severn, Gregory Sizikov, Matthew Snelham, Jed Souter, Dan Steinberg, Andy Swing, Mercedes Tan, Gregory Thorson, Bo Tian, Horia Toma, Erick Tuttle, Vijay Vasudevan, Richard Walter, Walter Wang, Eric Wilcox, and Doe Hyun Yoon. 2017. In-Datacenter Performance Analysis of a Tensor Processing Unit. In *Proceedings of the 44th Annual International Symposium on Computer Architecture*, ISCA '17, pages 1–12, New York, NY, USA. Association for Computing Machinery.

Guy-Alain Junter and Laurent Lebrun. 2017. Cellulose-based virus-retentive filters: a review. *Reviews in Environmental Science and Bio/Technology*, 16(3):455–489.

Bhushan J. Toley, Isabela Covelli, Yevgeniy Belousov, Sujatha Ramachandran, Enos Kline, Noah Scarr, Nic Vermeulen, Walt Mahoney, Barry R. Lutz, and Paul Yager. 2015. Isothermal strand displacement amplification (iSDA): a rapid and sensitive method of nucleic acid amplification for point-of-care diagnosis. *Analyst*, 140(22):7540–7549. Publisher: Royal Society of Chemistry.

Willi A Kalender. 2006. X-ray computed tomography. *Physics in medicine & Biology*, 51(13):R29.

Manoj Kumar Kanakasabapathy, Magesh Sadasivam, Anupriya Singh, Collin Preston, Prudhvi Thirumalaraju, Maanasa Venkataraman, Charles L Bormann, Mohamed Shehata Draz, John C Petrozza, and Hadi Shafiee. 2017. An automated smartphone-based diagnostic assay for point-of-care semen analysis. *Science translational medicine*, 9(382):eaai7863.

Soowon Kang, Hyeonwoo Choi, Sooyoung Park, Chunjong Park, Jemin Lee, Uichin Lee, and Sung-Ju Lee. 2019. Fire in Your Hands: Understanding Thermal Behavior of Smartphones. In *The 25th Annual International Conference on Mobile Computing and Networking*, MobiCom '19, pages 1–16, New York, NY, USA. Association for Computing Machinery.

Walter Karlen, J Mark Ansermino, Guy A Dumont, and Cornie Scheffer. 2013. Detection of the optimal region of interest for camera oximetry. In *2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 2263–2266. IEEE.

Walter Karlen, Joanne Lim, J Mark Ansermino, Guy Dumont, and Cornie Scheffer. 2012. Design challenges for camera oximetry on a mobile phone. In *2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 2448–2451. IEEE.

Matthew R. Kasper, Jesse R. Geibe, Christine L. Sears, Asha J. Riegodedios, Tina Luse, Annette M. Von Thun, Michael B. McGinnis, Niels Olson, Daniel Houskamp, Robert Fenequito, Timothy H. Burgess, Adam W. Armstrong, Gerald DeLong, Robert J. Hawkins, and Bruce L. Gillingham. An Outbreak of Covid-19 on an Aircraft Carrier. page NEJMoa2019375.

Firmin Kateu, Gentian Jakllari, and Emmanuel Chaput. 2022. Smartphox: Smartphone-based pulse oximetry using a meta-region of interest. In *2022 IEEE International Conference on Pervasive Computing and Communications (PerCom)*, pages 130–140. IEEE.

A-M Kelly, R McAlpine, and E Kyle. 2001. How accurate are pulse oximeters in patients with acute exacerbations of chronic obstructive airways disease? *Respiratory medicine*, 95(5):336–340.

Sung-Han Kim, So Young Chang, Minki Sung, Ji Hoon Park, Hong Bin Kim, Heeyoung Lee, Jae-Phil Choi, Won Suk Choi, and Ji-Young Min. 2016. Extensive Viable Middle East Respiratory Syndrome (MERS) Coronavirus Contamination in Air and Surrounding Environment in MERS Isolation Wards. *Clinical Infectious Diseases: An Official Publication of the Infectious Diseases Society of America*, 63(3):363–369.

Sanjay Kimbahune, Sujit Shinde, Karan Bhavsar, Avik Ghose, Sundeep Khandelwal, and Arpan Pal. 2021. Heart rate monitoring using capacitive touchscreen sensing. In *Proceedings of the Workshop on Body-Centric Computing Systems*, pages 13–17.

Enos C Kline, Nuttada Panpradist, Ian T Hull, Qin Wang, Amy K Oreskovic, Peter D Han, Lea M Starita, and Barry R Lutz. 2021. Multiplex target-redundant rt-lamp for robust detection of sars-cov-2 using fluorescent universal displacement probes. *medRxiv*.

Jeffrey A Kline, Jackeline Hernandez-Nino, Craig D Newgard, Dana N Cowles, Raymond E Jackson, and D Mark Courtney. 2003. Use of pulse oximetry to predict in-hospital complications in normotensive patients with pulmonary embolism. *The American journal of medicine*, 115(3):203–208.

Dechuan Kong, Yuanping Wang, Lu Lu, Huanyu Wu, Chuchu Ye, Abram L. Wagner, Jixing Yang, Yaxu Zheng, Xiaohuan Gong, Yiyi Zhu, Bihong Jin, Wenjia Xiao, Shenghua Mao, Chenyan Jiang, Sheng Lin, Ruobing Han, Xiao Yu, Peng Cui, Qiwen Fang, Yihan Lu, and Hao Pan. 2020. Clusters of 2019 coronavirus disease (COVID-19) cases in Chinese tour groups. *Transboundary and Emerging Diseases*, n/a(n/a).

Tassaneewan Laksanasopin, Tiffany W Guo, Samiksha Nayak, Archana A Sridhara, Shi Xie, Owolabi O Olowookere, Paolo Cadinu, Fanxing Meng, Natalie H Chee, Jiyoon Kim, et al. 2015. A smartphone dongle for diagnosis of infectious diseases at the point of care. *Science translational medicine*, 7(273):273re1–273re1.

Francesco Lamonaca, Domenico Luca Carnì, Domenico Grimaldi, Alfonso Nastro, Maria Riccio, and Vitaliano Spagnolo. 2015. Blood oxygen saturation measurement by smartphone camera. In *2015 IEEE International Symposium on Medical Measurements and Applications (MeMeA) Proceedings*, pages 359–364. IEEE.

Eric C. Larson, Mayank Goel, Gaetano Boriello, Sonya Heltshe, Margaret Rosenfeld, and Shwetak N. Patel. 2012a. SpiroSmart: using a microphone to measure lung function on a mobile phone. In *Proceedings of the 2012 ACM Conference on Ubiquitous Computing*, UbiComp '12, pages 280–289, New York, NY, USA. Association for Computing Machinery.

Eric C Larson, Mayank Goel, Gaetano Boriello, Sonya Heltshe, Margaret Rosenfeld, and Shwetak N Patel. 2012b. Spirosmart: using a microphone to measure lung function on a mobile phone. In *Proceedings of the 2012 ACM conference on ubiquitous computing*, pages 280–289.

Yann LeCun, Yoshua Bengio, and Geoffrey Hinton. 2015. Deep learning. *Nature*, 521(7553):436–444.

Joon Young Lee, Byoung Yeon Won, and Hyun Gyu Park. 2018. Label-Free Multiplex DNA Detection Utilizing Projected Capacitive Touchscreen. *Biotechnology Journal*, 13(2):1700362. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/biot.201700362.

Patricia L. M. Lee. 2017. DNA amplification in the field: move over PCR, here comes LAMP. *Molecular Ecology Resources*, 17(2):138–141. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/1755-0998.12548.

Jia Li, Joanne Macdonald, and Felix von Stetten. 2019. Review: a comprehensive summary of a decade development of the recombinase polymerase amplification. *Analyst*, 144(1):31–67. Publisher: Royal Society of Chemistry.

Michael S Lipnick, John R Feiner, Paul Au, Michael Bernstein, and Philip E Bickler. 2016. The accuracy of 6 inexpensive pulse oximeters not cleared by the food and drug administration: the possible global public health implications. *Anesthesia & Analgesia*, 123(2):338–345.

Jian Liu, Hongbo Liu, Yingying Chen, Yan Wang, and Chen Wang. 2019. Wireless sensing for human activity: A survey. *IEEE Communications Surveys & Tutorials*, 22(3):1629–1645.

Shuopeng Liu, Wenqiong Su, and Xianting Ding. 2016. A Review on Microfluidic Paper-Based Analytical Devices for Glucose Detection. *Sensors (Basel, Switzerland)*, 16(12):2086.

Wanyu Liu, Bernd Ploderer, and Thuong Hoang. 2015. In Bed with Technology: Challenges and Opportunities for Sleep Tracking. In *Proceedings of the Annual Meeting of the Australian Special Interest Group for Computer Human Interaction*, OzCHI '15, pages 142–151, New York, NY, USA. Association for Computing Machinery.

Xin Liu, Josh Fromm, Shwetak Patel, and Daniel McDuff. 2020. Multi-task temporal shift attention networks for on-device contactless vitals measurement. *Advances in Neural Information Processing Systems*, 33:19400–19411.

Ivan Magriñá Lobato and Ciara K. O'Sullivan. 2018. Recombinase polymerase amplification: Basics, applications and recent advances. *TrAC Trends in Analytical Chemistry*, 98:19–35.

Hervé Lobbes, Julien Dehos, Bruno Pereira, Guillaume Le Guenno, Laurent Sarry, and Marc Ruivard. 2019. Computed and subjective blue scleral color analysis as a diagnostic tool for iron deficiency: a pilot study. *Journal Of Clinical Medicine*, 8(11):1876.

Randolph Lopez, Ruofan Wang, and Georg Seelig. 2018. A molecular multi-gene classifier for disease diagnostics. *Nature Chemistry*, 10(7):746–754. Number: 7 Publisher: Nature Publishing Group.

Andrew M Luks and Erik R Swenson. 2020. Pulse oximetry for monitoring patients with covid-19 at home. potential pitfalls and practical guidance. *Annals of the American Thoracic Society*, 17(9):1040–1046.

Lei Luo, Dan Liu, Xinlong Liao, Xianbo Wu, Qinlong Jing, Jiazhen Zheng, Fanghua Liu, Shigui Yang, Hua Bi, Zhihao Li, Jianping Liu, Weiqi Song, Wei Zhu, Zhenghe Wang, Xiru Zhang, Qingmei Huang, Peiliang Chen, Huamin Liu, Xin Cheng, Miaochun Cai, Pei Yang, Xingfen Yang, Zhigang Han, Jinling Tang, Yu Ma, and Chen Mao. 2020. Contact Settings and Risk for Transmission in 3410 Close Contacts of Patients With COVID-19 in Guangzhou, China. *Annals of Internal Medicine*.

Dhruv Mahtta, Marilyne Daher, Michelle T. Lee, Saleem Sayani, Mehdi Shishehbor, and Salim S. Virani. 2021. Promise and Perils of Telehealth in the Current Era. *Current Cardiology Reports*, 23(9):115.

Burkhard Malorny, Panayotis T. Tassios, Peter Rådström, Nigel Cook, Martin Wagner, and Jeffrey Hoorfar. 2003. Standardization of diagnostic PCR for the detection of foodborne pathogens. *International Journal of Food Microbiology*, 83(1):39–48.

Robert G. Mannino, David R. Myers, Erika A. Tyburski, Christina Caruso, Jeanne Boudreaux, Traci Leong, G. D. Clifford, and Wilbur A. Lam. 2018a. Smartphone app for non-invasive detection of anemia using only patient-sourced photos. *Nature Communications*, 9(1):4924. Number: 1 Publisher: Nature Publishing Group.

Robert G Mannino, David R Myers, Erika A Tyburski, Christina Caruso, Jeanne Boudreaux, Traci Leong, GD Clifford, and Wilbur A Lam. 2018b. Smartphone app for non-invasive detection of anemia using only patient-sourced photos. *Nature communications*, 9(1):1–10.

Masimo. 2020 (accessed August 10, 2020). *Masimo Radical-7*. https://www.masimo.com/products/continuous/radical-7/.

Masimo. 2021 (accessed November 5, 2021). *Radical-7 Pulse CO-Oximeter Operator's Manual*. https://techdocs.masimo.com/globalassets/techdocs/pdf/lab-5475l_master.pdf.

M Jeremiah Matson, Claude Kwe Yinda, Stephanie N Seifert, Trenton Bushmaker, Robert J Fischer, Neeltje van Doremalen, James O Lloyd-Smith, and Vincent J Munster. 2020. Effect of environmental conditions on sars-cov-2 stability in human nasal mucus and sputum. *Emerging infectious diseases*, 26(9):2276.

Joyce C McCann and Bruce N Ames. 2007. An overview of evidence for a causal relation between iron deficiency during development and deficits in cognitive or behavioral function. *The American journal of clinical nutrition*, 85(4):931–945.

Gertjan Medema, Leo Heijnen, Goffe Elsinga, Ronald Italiaander, and Anke Brouwer. 2020. Presence of sars-coronavirus-2 rna in sewage and correlation with reported covid-19 prevalence in the early stage of the epidemic in the netherlands. *Environmental Science & Technology Letters*, 7(7):511–516.

Daryush D Mehta, Matias Zanartu, Shengran W Feng, Harold A Cheyne II, and Robert E Hillman. 2012. Mobile voice health monitoring using a wearable accelerometer sensor and a smartphone platform. *IEEE Transactions on Biomedical Engineering*, 59(11):3090–3096.

Yitzhak Mendelson and Burt D Ochs. 1988. Noninvasive pulse oximetry utilizing skin reflectance photoplethysmography. *IEEE Transactions on Biomedical Engineering*, 35(10):798–805.

Akinori Mitani, Abigail Huang, Subhashini Venugopalan, Greg S Corrado, Lily Peng, Dale R Webster, Naama Hammel, Yun Liu, and Avinash V Varadarajan. 2020. Detection of anaemia from retinal fundus images via deep learning. *Nature biomedical engineering*, 4(1):18–27.

Takayuki Miura, Yoshifumi Masago, Daisuke Sano, and Tatsuo Omura. 2011. Development of an Effective Method for Recovery of Viral Genomic RNA from Environmental Silty Sediments for Quantitative Molecular Detection. *Applied and Environmental Microbiology*, 77(12):3975–3981. Publisher: American Society for Microbiology Section: Methods.

Aashna M Modi, Renee D Kiourkas, Jie Li, and J Brady Scott. 2021. Reliability of smartphone pulse oximetry in subjects at risk for hypoxemia. *Respiratory Care*, 66(3):384–390.

Michael Moore, Beth Stuart, Mark Lown, Ann Van den Bruel, Sue Smith, Kyle Knox, Matthew J Thompson, and Paul Little. 2019. Predictors of adverse outcomes in uncomplicated lower respiratory tract infections. *The Annals of Family Medicine*, 17(3):231–238.

Teresa Moreno, Rosa María Pintó, Albert Bosch, Natalia Moreno, Andrés Alastuey, María Cruz Minguillón, Eduard Anfruns-Estrada, Susana Guix, Cristina Fuentes, Giorgio Buonanno, Luca Stabile, Lidia Morawska, and Xavier Querol. Tracing surface and airborne SARS-CoV-2 RNA inside public buses and subway trains. 147:106326.

Carla Moscheo, Maria Licciardello, Piera Samperi, Milena La Spina, Andrea Di Cataldo, and Giovanna Russo. 2022. New insights into iron deficiency anemia in children: a practical review. *Metabolites*, 12(4):289.

X Muñoz, F Torres, G Sampol, J Rios, S Martí, and E Escrich. 2008. Accuracy and reliability of pulse oximetry at different arterial carbon dioxide pressure levels. *European Respiratory Journal*, 32(4):1053–1059.

Ottar Mæstad, Gaute Torsvik, and Arild Aakvik. 2010. Overworked? On the relationship between workload and health worker performance. *Journal of Health Economics*, 29(5):686–698.

Taiki Nakanishi, Shunya Murakami, Atsuki Kobayashi, Md. Zahidul Islam, and Kiichi Niitsu. 2019. A 40-GHz Fully-Integrated CMOS-Based Biosensor Circuit with an On-Chip Vector Network Analyzer for Circulating Tumor Cells Analysis. In *2019 IEEE Nordic Circuits and Systems Conference (NORCAS): NORCHIP and International Symposium of System-on-Chip (SoC)*, pages 1–7.

Hyoungsik Nam, Ki-Hyuk Seol, Junhee Lee, Hyeonseong Cho, and Sang Won Jung. 2021a. Review of Capacitive Touchscreen Technologies: Overview, Research Trends, and Machine Learning Approaches. *Sensors (Basel, Switzerland)*, 21(14):4776.

Hyoungsik Nam, Ki-Hyuk Seol, Junhee Lee, Hyeonseong Cho, and Sang Won Jung. 2021b. Review of capacitive touchscreen technologies: Overview, research trends, and machine learning approaches. *Sensors*, 21(14):4776.

Carl-Johan Neiderud. 2015. How urbanization affects the epidemiology of emerging infectious diseases. *Infection Ecology & Epidemiology*, 5(1).

Andrea Nemcova, Ivana Jordanova, Martin Varecka, Radovan Smisek, Lucie Marsanova, Lukas Smital, and Martin Vitek. 2020. Monitoring of heart rate, blood oxygen saturation, and blood pressure using a smartphone. *Biomedical Signal Processing and Control*, 59:101928.

Bo Ning, Tao Yu, Shengwei Zhang, Zhen Huang, Di Tian, Zhen Lin, Alex Niu, Nadia Golden, Krystle Hensley, Breanna Threeton, et al. 2021. A smartphone-read ultrasensitive and quantitative saliva test for covid-19. *Science advances*, 7(2):eabe3703.

Kyoung-Su Oh and Keechul Jung. 2004. GPU implementation of neural networks. *Pattern Recognition*, 37(6):1311–1314.

Sean Wei Xiang Ong, Yian Kim Tan, Po Ying Chia, Tau Hong Lee, Oon Tek Ng, Michelle Su Yen Wong, and Kalisvar Marimuthu. 2020. Air, Surface Environmental, and Personal Protective Equipment Contamination by Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) From a Symptomatic Patient. *JAMA*.

World Health Organization et al. 2011. Haemoglobin concentrations for the diagnosis of anaemia and assessment of severity. vitamin and mineral nutrition information system. *(No Title)*.

World Health Organization et al. 2020. What do we know about community health workers? a systematic review of existing reviews. *What do we know about community health workers? A systematic review of existing reviews*.

John D. Owens, Mike Houston, David Luebke, Simon Green, John E. Stone, and James C. Phillips. 2008. GPU Computing. *Proceedings of the IEEE*, 96(5):879–899. Conference Name: Proceedings of the IEEE.

Meng-Shiuan Pan and Hsueh-Wei Lin. 2015. A Step Counting Algorithm for Smartphone Users: Design and Implementation. *IEEE Sensors Journal*, 15(4):2296–2305. Conference Name: IEEE Sensors Journal.

Nuttada Panpradist, Ingrid A. Beck, Justin Vrana, Nikki Higa, David McIntyre, Parker S. Ruth, Isaac So, Enos C. Kline, Ruth Kanthula, Annie Wong-On-Wing, Jonathan Lim, Daisy Ko, Ross Milne, Theresa Rossouw, Ute D. Feucht, Michael Chung, Gonzague Jourdain, Nicole Ngo-Giang-Huong, Laddawan Laomanit, Jaime Soria, James Lai, Eric D. Klavins, Lisa M. Frenkel, and Barry R. Lutz. 2019. OLA-Simple: A software-guided HIV-1 drug resistance test for low-resource laboratories. *EBioMedicine*, 50:34–44.

Nuttada Panpradist, Enos Kline, Robert G Atkinson, Michael Roller, Qin Wang, Ian T Hull, Jack H Kotnik, Amy Oreskovic, Crissa Bennett, Daniel Leon, et al. 2021a. Harmony covid-19: a ready-to-use kit, low-cost detector, and smartphone app for point-of-care sars-cov-2 rna detection. *medRxiv*.

Nuttada Panpradist, Bhushan J Toley, Xiaohong Zhang, Samantha Byrnes, Joshua R Buser, Janet A Englund, and Barry R Lutz. 2014. Swab sample transfer for point-of-care diagnostics: characterization of swab types and manual agitation methods. *PloS one*, 9(9):e105786.

Nuttada Panpradist, Qin Wang, Parker S. Ruth, Jack H. Kotnik, Amy K. Oreskovic, Abraham Miller, Samuel W. A. Stewart, Justin Vrana, Peter D. Han, Ingrid A. Beck, Lea M. Starita, Lisa M. Frenkel, and Barry R. Lutz. 2021b. Simpler and faster Covid-19 testing: Strategies to streamline SARS-CoV-2 molecular assays. *EBioMedicine*, 64:103236.

Nuttada Panpradist, Qin Wang, Parker S Ruth, Jack H Kotnik, Amy K Oreskovic, Abraham Miller, Samuel WA Stewart, Justin Vrana, Peter D Han, Ingrid A Beck, et al. 2021c. Simpler and faster covid-19 testing: Strategies to streamline sars-cov-2 molecular assays. *EBioMedicine*, 64:103236.

Chunjong Park, Alex Mariakakis, Jane Yang, Diego Lassala, Yasamba Djiguiba, Youssouf Keita, Hawa Diarra, Beatrice Wasunna, Fatou Fall, Marème Soda Gaye, Bara Ndiaye, Ari Johnson, Isaac Holeman, and Shwetak Patel. 2020. Supporting Smartphone-Based Image Capture of Rapid Diagnostic Tests in Low-Resource Settings. In *Proceedings of the 2020 International Conference on Information and Communication Technologies and Development*, ICTD2020, pages 1–11, New York, NY, USA. Association for Computing Machinery.

Jordan Peccia, Alessandro Zulli, Doug E. Brackney, Nathan D. Grubaugh, Edward H. Kaplan, Arnau Casanovas-Massana, Albert I. Ko, Amyn A. Malik, Dennis Wang, Mike Wang, Daniel M. Weinberger, and Saad B. Omer. 2020. SARS-CoV-2 RNA concentrations in primary municipal sewage sludge as a leading indicator of COVID-19 outbreak dynamics. *medRxiv*, page 2020.05.19.20105999.

Ming-Zher Poh, Daniel J McDuff, and Rosalind W Picard. 2010. Non-contact, automated cardiac pulse measurements using video imaging and blind source separation. *Optics express*, 18(10):10762–10774.

Praveen Rai, Ballamoole Krishna Kumar, Vijaya Kumar Deekshit, Indrani Karunasagar, and Iddya Karunasagar. 2021. Detection technologies and recent developments in the diagnosis of covid-19 infection. *Applied microbiology and biotechnology*, 105:441–455.

Fabien Rallu, Peter Barriga, Carole Scrivo, Valérie Martel-Laferrière, and Céline Laferrière. 2006. Sensitivities of Antigen Detection and PCR Assays Greatly Increased Compared to That of the Standard Culture Method for Screening for Group B Streptococcus Carriage in Pregnant Women. *Journal of Clinical Microbiology*, 44(3):725–728. Publisher: American Society for Microbiology.

Joseph Redmon, Santosh Divvala, Ross Girshick, and Ali Farhadi. 2016. You only look once: Unified, real-time object detection. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 779–788.

Donald C Rio, Manuel Ares, Gregory J Hannon, and Timothy W Nilsen. 2010. Purification of rna using trizol (tri reagent). *Cold Spring Harbor Protocols*, 2010(6):pdb–prot5439.

Karyna Rosario, Noah Fierer, Shelly Miller, Julia Luongo, and Mya Breitbart. 2018. Diversity of DNA and RNA Viruses in Indoor Air As Assessed via Metagenomic Sequencing. *Environmental Science & Technology*, 52(3):1014–1027. Publisher: American Chemical Society.

Karen Rose, Scott Eldridge, Lyman Chapin, et al. 2015. The internet of things: An overview. *The internet society (ISOC)*, 80(15):1–53.

Elisabeth Rosenthal. 2017. *An American Sickness: How Healthcare Became Big Business and How You Can Take It Back*. Penguin. Google-Books-ID: MAqyDAAAQBAJ.

Camilla Rothe, Mirjam Schunk, Peter Sothmann, Gisela Bretzel, Guenter Froeschl, Claudia Wallrauch, Thorbjörn Zimmer, Verena Thiel, Christian Janke, Wolfgang Guggemos, Michael Seilmaier, Christian Drosten, Patrick Vollmar, Katrin Zwirglmaier, Sabine Zange, Roman Wölfel, and Michael Hoelscher. 2020. Transmission of 2019-nCoV Infection from an Asymptomatic Contact in Germany. *The New England Journal of Medicine*, 382(10):970–971.

Sohrab Saeb, Thaddeus R. Cybulski, Stephen M. Schueller, Konrad P. Kording, and David C. Mohr. 2017. Scalable Passive Sleep Monitoring Using Mobile Phones: Opportunities and Obstacles. *Journal of Medical Internet Research*, 19(4):e6821. Company: Journal of Medical Internet Research Distributor: Journal of Medical Internet Research Institution: Journal of Medical Internet Research Label: Journal of Medical Internet Research Publisher: JMIR Publications Inc., Toronto, Canada.

Bhartipudi Sahishnavi, Samriddhi Agarwal, Shameer Basha, Naveen Dasari, Andleeb Zahra, Prabhakar Bhimalapuram, Syed Azeemuddin, and Zia Abbas. 2024. An inductor-less, cost-effective On-chip CMOS VNA for bio-molecule detection. In *2024 IEEE International Symposium on Circuits and Systems (ISCAS)*, pages 1–5. ISSN: 2158-1525.

Joseph Malcolm Schaeffer, Chris Thachuk, and Erik Winfree. 2015. Stochastic Simulation of the Kinetics of Multiple Interacting Nucleic Acid Strands. In *DNA Computing and Molecular Programming*, Lecture Notes in Computer Science, pages 194–211, Cham. Springer International Publishing.

Robert R Schaller. 2002. Moore's law: past, present and future. *IEEE spectrum*, 34(6):52–59.

Candice Schumann, Femi Olanubi, Auriel Wright, Ellis Monk, Courtney Heldreth, and Susanna Ricco. 2023. Consensus and subjectivity of skin tone annotation for ml fairness. *Advances in Neural Information Processing Systems*, 36:30319–30348.

Christopher G Scully, Jinseok Lee, Joseph Meyer, Alexander M Gorbach, Domhnull Granquist-Fraser, Yitzhak Mendelson, and Ki H Chon. 2011. Physiological parameter monitoring from optical recordings with a mobile phone. *IEEE Transactions on Biomedical Engineering*, 59(2):303–306.

Kamal G Shah, Vidhi Singh, Peter C Kauffman, Koji Abe, and Paul Yager. 2018. Mobile phone ratiometric imaging enables highly sensitive fluorescence lateral flow immunoassays without external optical filters. *Analytical chemistry*, 90(11):6967–6974.

Manuja Sharma, Videlis Nduba, Lilian N. Njagi, Wilfred Murithi, Zipporah Mwongera, Thomas R. Hawn, Shwetak N. Patel, and David J. Horne. 2024. TBscreen: A passive cough classifier for tuberculosis screening with a controlled dataset. *Science Advances*, 10(1):eadi0282. Publisher: American Association for the Advancement of Science.

Ye Shen, Changwei Li, Hongjun Dong, Zhen Wang, Leonardo Martinez, Zhou Sun, Andreas Handel, Zhiping Chen, Enfu Chen, Mark H. Ebell, Fan Wang, Bo Yi, Haibin Wang, Xiaoxiao Wang, Aihong Wang, Bingbing Chen, Yanling Qi, Lirong Liang, Yang Li, Feng Ling, Junfang Chen, and Guozhang Xu. 2020. Community Outbreak Investigation of SARS-CoV-2 Transmission Among Bus Riders in Eastern China. *JAMA Internal Medicine*.

Tasnuba Siddiqui and Bashir I Morshed. 2018. Severity classification of chronic obstructive pulmonary disease and asthma with heart rate and spo2 sensors. In *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 2929–2932. IEEE.

David Silcott, Sean Kinahan, Joshua Santarpia, Blake Silcott, Ryan Silcott, Peter Silcott, Braden Silcott, Steven Distelhorst, Vicki Herrera, Danielle Rivera, et al. 2020. Transcom/amc commercial aircraft cabin aerosol dispersion tests. Technical report, NEBRASKA UNIV AT OMAHA OMAHA.

A. Prasanna de Silva and Seiichi Uchiyama. 2007. Molecular logic and computing. *Nature Nanotechnology*, 2(7):399–410. Publisher: Nature Publishing Group.

Friedrich C. Simmel, Bernard Yurke, and Hari R. Singh. 2019. Principles and Applications of Nucleic Acid Strand Displacement Reactions. *Chemical Reviews*, 119(10):6326–6369. Publisher: American Chemical Society.

James E Sinex. 1999. Pulse oximetry: Principles and limitations. *The American Journal of Emergency Medicine*, 17(1):59–66.

Michael W Sjoding, Robert P Dickson, Theodore J Iwashyna, Steven E Gay, and Thomas S Valley. 2020. Racial bias in pulse oximetry measurement. *New England Journal of Medicine*, 383(25):2477–2478.

Lloyd M. Smith, Jane Z. Sanders, Robert J. Kaiser, Peter Hughes, Chris Dodd, Charles R. Connell, Cheryl Heiner, Stephen B. H. Kent, and Leroy E. Hood. 1986. Fluorescence detection in automated DNA sequence analysis. *Nature*, 321(6071):674–679. Publisher: Nature Publishing Group.

Claudio Stagni, Carlotta Guiducci, Luca Benini, Bruno Ricco, Sandro Carrara, Bruno Samori, Christian Paulus, Meinrad Schienle, Marcin Augustyniak, and Roland Thewes. 2006. CMOS DNA Sensor Array With Integrated A/D Conversion Based on Label-Free Capacitance Measurement. *IEEE Journal of Solid-State Circuits*, 41(12):2956–2964.

Steven R Steinhubl, Evan D Muse, and Eric J Topol. 2015. The emerging field of mobile health. *Science translational medicine*, 7(283):283rv3–283rv3.

Jonathan Strutt, Girish Narayanswamy, Chunjong Park, Devesh Sarda, Sixuan Wu, Matthew Thompson, Lauren Harvey, Rachel Hedstrom, Amy Kodet, Shwetak Patel, et al. 2023. Capapp: Smartphone-based capillary refill index assessment in healthy children. In *Frontiers in Biomedical Devices*, volume 86731, page V001T09A009. American Society of Mechanical Engineers.

Zhiyuan Sun, Qinghua He, Yuandong Li, Wendy Wang, and Ruikang K Wang. 2021. Robust non-contact peripheral oxygenation saturation measurement using smartphone-enabled imaging photoplethysmography. *Biomedical Optics Express*, 12(3):1746–1760.

İsmail Tayfur and Mustafa Ahmet Afacan. 2019. Reliability of smartphone measurements of vital parameters: A prospective study using a reference method. *The American journal of emergency medicine*, 37(8):1527–1530.

digiDoc Technologies. 2013 (accessed August 10, 2020). *Pulse Oximeter Application*. https://www.digidoctech.no/products-1.html.

Ayalew Tefferi. 2004. Practical algorithms in anemia diagnosis. In *Mayo Clinic Proceedings*, volume 79, pages 955–956. Elsevier.

Sarah Tomlinson, Sydney Behrmann, James Cranford, Marisa Louie, and Andrew Hashikawa. 2018. Accuracy of smartphone-based pulse oximetry compared with hospital-grade pulse oximetry in healthy children. *Telemedicine and e-Health*, 24(7):527–535.

Eric J Topol. 2019. A decade of digital medicine innovation. *Science translational medicine*, 11(498).

Seattle Department of Transportation. 2019. 2019 seattle center city commute mode split survey. https://www.commuteseattle.com/wp-content/uploads/2020/09/2019-Mode-Split-Final-Report.pdf. Accessed: 2021-05-05.

Kevin K Tremper. 1989a. Pulse oximetry. *Chest*, 95(4):713–715.

Kevin K. Tremper. 1989b. Pulse Oximetry. *CHEST*, 95(4):713–715. Publisher: Elsevier.

Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. 2017. Attention is All you Need. In *Advances in Neural Information Processing Systems*, volume 30. Curran Associates, Inc.

Anandghan Waghmare, Youssef Ben Taleb, Ishan Chatterjee, Arjun Narendra, and Shwetak Patel. 2023a. Z-Ring: Single-Point Bio-Impedance Sensing for Gesture, Touch, Object and User Recognition. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, pages 1–18, Hamburg Germany. ACM.

Anandghan Waghmare, Farshid Salemi Parizi, Jason Hoffman, Yuntao Wang, Matthew Thompson, and Shwetak Patel. 2023b. GlucoScreen: A Smartphone-based Readerless Glucose Test Strip for Prediabetes Screening. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 7(1):30:1–30:20.

Edward J. Wang, William Li, Junyi Zhu, Rajneil Rana, and Shwetak N. Patel. 2017a. Noninvasive hemoglobin measurement using unmodified smartphone camera and white flash. In *2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 2333–2336, Seogwipo. IEEE.

Edward J Wang, William Li, Junyi Zhu, Rajneil Rana, and Shwetak N Patel. 2017b. Noninvasive hemoglobin measurement using unmodified smartphone camera and white flash. In *2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 2333–2336. IEEE.

Edward Jay Wang, William Li, Doug Hawkins, Terry Gernsheimer, Colette Norby-Slycord, and Shwetak N. Patel. 2016. HemaApp: noninvasive blood screening of hemoglobin using smartphone cameras. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing - UbiComp '16*, pages 593–604, Heidelberg, Germany. ACM Press.

Edward Jay Wang, Junyi Zhu, Mohit Jain, Tien-Jui Lee, Elliot Saba, Lama Nachman, and Shwetak N Patel. 2018. Seismo: Blood pressure monitoring using built-in smartphone accelerometer and camera. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, pages 1–9.

Wenjin Wang, Albertus C. den Brinker, Sander Stuijk, and Gerard de Haan. 2017c. Algorithmic Principles of Remote PPG. *IEEE Transactions on Biomedical Engineering*, 64(7):1479–1491. Conference Name: IEEE Transactions on Biomedical Engineering.

Mark Weiser. 1991. The computer for the 21 st century. *Scientific american*, 265(3):94–105.

Jim Welch. 2005. Pulse Oximeters. *Biomedical Instrumentation & Technology*, 39(2):125–130. Publisher: AAMI.

JP Welch, R DeCesare, and D Hess. 1990. Pulse oximetry: instrumentation and clinical applications. *Respir Care*, 35(6):584–597.

Thomas Alan Wemyss, Miranda Nixon-Hill, Felix Outlaw, Anita Karsa, Judith Meek, Christabel Enweronu-Laryea, and Terence S Leung. 2023. Feasibility of smartphone colorimetry of the face as an anaemia screening tool for infants and young children in ghana. *Plos one*, 18(3):e0281736.

W. Joost Wiersinga, Andrew Rhodes, Allen C. Cheng, Sharon J. Peacock, and Hallie C. Prescott. 2020. Pathophysiology, Transmission, Diagnosis, and Treatment of Coronavirus Disease 2019 (COVID-19): A Review. *JAMA*.

Val Wilson. 2012. Diagnosis and treatment of diabetic ketoacidosis. *Emergency Nurse*, 20(7).

World Health Organization. 2020. Status of environmental surveillance for SARS-CoV-2 virus.

World Health Organization and others. 2020. Coronavirus disease (covid-2019) situation reports. Accessed: 2021–12-20.

Jedrek Wosik, Marat Fudim, Blake Cameron, Ziad F Gellad, Alex Cho, Donna Phinney, Simon Curtis, Matthew Roman, Eric G Poon, Jeffrey Ferranti, Jason N Katz, and James Tcheng. 2020. Telehealth transformation: COVID-19 and the rise of virtual care. *Journal of the American Medical Informatics Association*, 27(6):957–962.

Fuqing Wu, Amy Xiao, Jianbo Zhang, Xiaoqiong Gu, Wei Lin Lee, Kathryn Kauffman, William Hanage, Mariana Matus, Newsha Ghaeli, Noriko Endo, Claire Duvallet, Katya Moniz, Timothy Erickson, Peter Chai, Janelle Thompson, and Eric Alm. 2020. SARS-CoV-2 titers in wastewater are higher than expected from clinically confirmed cases. *medRxiv*, page 2020.04.05.20051540.

Sebastien Wurtzer, Vincent Marechal, Jean-Marie Mouchel, Yvon Maday, Remy Teyssou, Elise Richard, Jean Luc Almayrac, and Laurent Moulin. 2020. Evaluation of lockdown impact on SARS-CoV-2 dynamics through viral genome quantification in Paris wastewaters. *medRxiv*, page 2020.04.12.20062679.

Jiang Xie, Naima Covassin, Zhengyang Fan, Prachi Singh, Wei Gao, Guangxi Li, Tomas Kara, and Virend K Somers. 2020. Association between hypoxemia and mortality in patients with covid-19. In *Mayo Clinic Proceedings*. Elsevier.

Deepika Yadav, Prerna Malik, Kirti Dabas, and Pushpendra Singh. 2021. Illustrating the gaps and needs in the training support of community health workers in india. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, pages 1–16.

Naibin Yang, Yuefei Shen, Chunwei Shi, Ada Hoi Yan Ma, Xie Zhang, Xiaomin Jian, Liping Wang, Jiejun Shi, Chunyang Wu, Guoxiang Li, Yuan Fu, Keyin Wang, Mingqin Lu, and Guoqing Qian. 2020. In-flight transmission cluster of COVID-19: A retrospective case series. *Infectious Diseases*, 0(0):1–11.

Ping Yu, Jiang Zhu, Zhengdong Zhang, and Yingjun Han. 2020. A Familial Cluster of Infection Associated With the 2019 Novel Coronavirus Indicating Possible Person-to-Person Transmission During the Incubation Period. *The Journal of Infectious Diseases*, 221(11):1757–1761.

Xingxia Yu and Rongrong Yang. 2020. COVID-19 transmission through asymptomatic carriers is a challenge to containment. *Influenza and Other Respiratory Viruses*, 14(4):474–475.

Joseph N. Zadeh, Conrad D. Steenberg, Justin S. Bois, Brian R. Wolfe, Marshall B. Pierce, Asif R. Khan, Robert M. Dirks, and Niles A. Pierce. 2011. NUPACK: Analysis and design of nucleic acid systems. *Journal of Computational Chemistry*, 32(1):170–173.

Camellia Zakaria, Gizem Yilmaz, Priyanka Mary Mammen, Michael Chee, Prashant Shenoy, and Rajesh Balan. 2023. SleepMore: Inferring Sleep Duration at Scale via Multi-Device WiFi Sensing. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, 6(4):193:1–193:32.

ZeenKo. 2022. *LiteVNA*. `https://www.zeenko.tech/litevna`.

David Yu Zhang and Georg Seelig. 2011. Dynamic DNA nanotechnology using strand-displacement reactions. *Nature Chemistry*, 3(2):103–113.

Jinny X. Zhang, John Z. Fang, Wei Duan, Lucia R. Wu, Angela W. Zhang, Neil Dalchau, Boyan Yordanov, Rasmus Petersen, Andrew Phillips, and David Yu Zhang. 2018. Predicting DNA hybridization kinetics from sequence. *Nature Chemistry*, 10(1):91–98.

Karen Zhang, Yuan-Jyue Chen, Delaney Wilde, Kathryn Doroschak, Karin Strauss, Luis Ceze, Georg Seelig, and Jeff Nivala. 2022. A nanopore interface for higher bandwidth DNA computing. *Nature Communications*, 13(1):4904. Publisher: Nature Publishing Group.

Renyi Zhang, Yixin Li, Annie L. Zhang, Yuan Wang, and Mario J. Molina. 2020. Identifying airborne transmission as the dominant route for the spread of COVID-19. *Proceedings of the National Academy of Sciences of the United States of America*, 117(26):14857–14863.

Shi Zhao, Zian Zhuang, Jinjun Ran, Jiaer Lin, Guangpu Yang, Lin Yang, and Daihai He. 2020. The association between domestic train transportation and novel coronavirus (2019-nCoV) outbreak in China from 2019 to 2020: A data-driven correlational report. *Travel Medicine and Infectious Disease*, 33:101568.

Ruizhi Zheng, Yu Xu, Weiqing Wang, Guang Ning, and Yufang Bi. 2020. Spatial transmission of COVID-19 via public and private transportation in China. *Travel Medicine and Infectious Disease*, 34:101626.

David A Zisman, David J Ross, John A Belperio, Rajan Saggar, Joseph P Lynch III, Abbas Ardehali, and Arun S Karlamangla. 2007. Prediction of pulmonary hypertension in idiopathic pulmonary fibrosis. *Respiratory medicine*, 101(10):2153–2159.

Appendix A

Supplementary Information: Chapter Summaries

This supplementary section contains the Background and Conclusion sections from each of the project chapters (Chapters 3-6), which can allow the reader to trace the full trail of the thesis through all of my projects. If the reader is already familiar with my published work, they may read this section instead of all those chapters. They can start with Introduction and Related Work (Chapters 1 and 2), then read this appendix section (Appendix A), and finally read the thesis Conclusion (Chapter 7). This will allow the reader to obtain a flavor of my thesis findings in summary form in a continuous format, without re-reading content with which they are already familiar.

A.1 Chapter 3 Summary: Measuring Blood Oxygen Saturation Non-Invasively Using a Smartphone Camera

Background: Balanced Data Collection Inspired by Health Validation Studies

In an effective machine learning pipeline for biomarker sensing, the first step is gathering high quality data for training that model. In my work on sensing blood oxygen saturation using a smartphone camera, I noticed that prior studies into this had used a convolutional neural network (CNN), but did not gather a clinically relevant range of data. Because we have a known physical model (the Beer-Lambert Law) for sensing the proportion of hemoglobin molecules in blood carrying oxygen, we can understand why machine learning can be applied in this case to sort through the noise inherent in using a commodity sensor.

In my project, the commodity smartphone camera emulated the purpose-built sensor of a pulse oximeter, but the general purpose nature of the camera added some unhelpful noise into the signal. Machine learning, particularly a relatively shallow convolutional neural network, proved to be a successful computational technique to sift through that noise efficiently on a large scale over thousands of data points. In order to demonstrate accuracy of the model on a clinically relevant dataset, I, along with my collaborators at UCSD and SMU, gathered data for that model from test subjects in an induced hypoxemia study design, called a Varied Fractional Inspired Oxygen (FiO_2) study. This study was modeled on the requirements for clearing a pulse oximeter device for medical use, which allows us to understand the medical relevance of a potential diagnostic tool for which the physics and physiology is already fairly well understood, which allows us to use the computational technique of machine learning to create a potentially clinically-relevant diagnostic.

See Chapter 3 for the full details on this project.

A.1.1 Conclusion: Representative Data Collection in Machine Learning

In machine learning for health sensing applications, training and testing the resultant health tool on data that is representative of the target population in both healthy and sick states is needed to demonstrate that such a diagnostic tool would be useful in healthcare situations. My study was the first to measure accuracy of smartphone pulse oximetry on a range of ground truth readings significantly below 90%, the threshold below which many clinicians agree further medical attention and treatment is necessary. Prior studies in smartphone-based pulse oximetry used different methods to gather training data, which did not allow this analysis. Those studies revealed that machine learning may be possible, but mine was the first to test this hypothesis all the way down to below 70%.

Designing the data collection based on the medical regulatory requirements worked in this case to produce good results at ranges that were previously unexplored, paving the path for future work that may increase the accuracy of these methods to be usable in medical practice. It also paves the way for looking at other biomarkers in a similar manner, such as those described in the following chapters of my thesis. Next, we'll look at how smartphone cameras can be configured to noninvasively detect the concentration of hemoglobin in blood.

A.2 Chapter 4 Summary: Measuring Blood Hemoglobin Concentration Non-Invasively Using a Smartphone Camera

Background: Data Augmentation for ML Training in Health

As we saw in the previous chapter, theory and practice show that machine learning models for healthcare perform best for patients when the physiological underpinning is well-understood, the sensor (general-purpose or designed) can measure the expected signal, and a fully representative spread of data is gathered to train the model. In those cases, we can be confident that the trained model fully represents the potential set of patients who could use the diagnostic tool. Most importantly, we can be confident that the diagnostic tool can perform accurately not only in cases when the patient is healthy, but also in cases when the patient is sick and needs medical attention. This is crucial for establishing an adequate sensitivity vs specificity tradeoff for all patients who may use that diagnostic (or utility vs specificity as described in Figure 1.2). However, practical considerations in study design and patient recruitment can intervene, giving cases in which the relevant range of data is difficult to obtain. For example, if the treatment is highly accessible or the lack of treatment can be very dangerous, the treatment is necessarily applied quickly, helping the patient to recover and simultaneously reducing our ability to gather training data that would be helpful in developing the new diagnostic.

This was our experience in the case of pediatric anemia. I, along with my collaborators at MyOr, were interested in trying to build a smartphone-based screening tool for the pediatric population. Treatments, such as iron supplements, can be relatively affordable and available in many cases, which means that pediatric patients can benefit greatly from low-cost anemia screening. It also means that it can be difficult to gather training data from patients exhibiting anemia, as they are often already treated, especially in the hospital setting in which we gathered the data. Even though the physiological explanation, that hemoglobin is the major component of blood that reflects red light, was fairly convincing, and prior research showed that simple machine learning techniques, like a multi-linear regression model, can give a good result, we could not gather an adequate range of data for this study. Therefore, we looked to machine learning literature and employed data augmentation techniques, particularly SMOTE, which enabled us to balance the training data for our model. In this study, we showed that more accurate hemoglobin sensing across the full range

of patients in a pediatric setting may be plausible and demonstrated that computational techniques can be applied to data from commodity sensors to reveal the signal of interest out of the noise to improve access to healthcare.

See Chapter 4 for the full details on this project.

A.2.1 Conclusion: Augmenting Machine Learning Data to Indicate Paths Forward for Diagnostics

When real-world circumstances prevent the collection of an ideal dataset for representing the states of the predicted biomarker of interest, machine learning data augmentation techniques can provide an indicator for paths forward. In this study, I worked with a startup company to redesign the data collection smartphone application for the Google Pixel 6 Pro and advocated for gathering data across the full spread of anemic and non-anemic test subjects. However, despite our best efforts, our collaborators at the hospital were unable to collect enough samples from test subjects with anemia to balance the dataset. Data augmentation was helpful in this case, showing that reasonable synthetic data could be sprinkled into the training process to produce a more accurate model overall, illuminating the path forward towards a future where non-invasive anemia screening may be helpful in identifying candidates for affordable treatment.

For real-world health diagnostics, it's likely that a combination of good data collection, processing, and modeling techniques are helpful. Ultimately, a combination of these data gathering and data augmentation approaches add up to the most effective computational approach for building affordable health diagnostics based on sensors built into ubiquitous computing devices. Also, these approaches typically work well for sensing biomarkers which already have a well-developed physiological basis for how they interact with sensors. I demonstrated the computational side of my thesis statement depicted in Figure 1.1 in these first two projects on non-invasive, in vivo health biomarker sensing and learned lessons that I applied to my next set of two projects on in vitro environment sensing.

In cases where the current detection modes are not as well-adapted to direct smartphone sensing, an approach that begins with building understanding of the underlying biology is appropriate for beginning to develop a smartphone sensor-based diagnostic. We will explore this side of the thesis in Chapter 6, beginning with the motivations for how such a diagnostic could be useful for broadly scalable population sensing of pandemic safety in Chapter 5.

A.3 Chapter 5 Summary: Passively Monitoring for the Presence of Viral RNA in Public Transit Infrastructure

Background: Passively Collecting Viral Samples by Leveraging Infrastructure

Applying scalable sensing techniques to other areas that impact health, such as environmental pathogen monitoring, became critical during the global COVID-19 pandemic. Data about whether a certain molecule was located in a specific location defined whether people left their houses and interacted with each other. In order to contribute to the solution for this and future pandemics, we attempted to passively monitor the spread of the virus via a creative passive sampling technique that could theoretically provide a scalable map of the spread of the virus.

I worked with King County Metro to place air filters, which we dubbed "passive fabric sensors", into the existing HVAC air filtration systems of public transit buses. Between the passive sampling and subsequent lab extraction and amplification, we demonstrated that the two steps needed for this system are possible, but I noticed that the latter was more limiting in its scalability than the former. Our method to place these

sensors and then gather those samples was relatively low-touch and scalable, but following that up with standard laboratory techniques for RNA extraction and RT-qPCR was laborious, time-intensive, and required specialized equipment in a well-equipped research laboratory. The goal of sensing the virus in a moving vehicle attached to routing and timing data for the bus route has the potential to build a live map of the risk of viral spread in different areas of the city. However, the difficulty of the second step, extracting and affirmatively sensing the viral samples via RT-qPCR, reduces the scalability of the overall process, preventing the accomplishment of these broader goals without enormous investment. As we will see in Chapter 6, this experience inspired the final project in my thesis, and the combination of these two projects demonstrates the left side of my thesis in Figure 1.1. Understanding the biology to effectively amplify its signal is the second critical piece for improving access to screening for health-relevant biomarkers, and will be explored in the next two chapters.

See Chapter 5 for the full details on this project.

A.3.1 Conclusion: Passive Sampling and Non-Passive Molecular Output

In its current state, qPCR is the gold standard for highly sensitive assays, such as the one that was needed for detection of SARS-CoV-2 after passive environmental sampling. However, it is not a very efficient method of detection, despite efforts to scale it up and bring it into the field. After designing and implementing a viral surveillance system leveraging public transit infrastructure to assist with early warning for spread of the COVID-19 pandemic, I identified that the main gap which prevents large-scale rollout of this method is the molecular readout throughput. Improvements in conduction of the molecular signal to the digital world would allow for scaling this method, as well as others, and answering finer-grained questions about the spread of the virus. I developed a passive sampling methodology, but the output mode of the results was far from passive.

Thus, I started developing a method that could imaginably be brought directly onsite for testing for the presence of specific nucleic acid targets. With this, we could process many more samples and begin to answer questions about disease spread in certain geographic locations, precise timing of when the virus appeared in the filters, and more. This insight led me on a journey to discover more scalable ways to solve the detection problem, which can be broken down to a biological understanding and molecular output problem. The results of that investigation, and how they apply to my thesis, are shared in the next chapter.

A.4 Chapter 6 Summary: Detecting the Output of Nucleic Acid Reactions Using Capacitive Touchscreen Sensors On Ubiquitous Smartphones

Background: Improving Access to Viral Sensing via the Smartphone Capacitive Touchscreen Sensor

For detection of viral pathogens, the gold standard for sensitivity and quantification is nucleic-acid based amplification methods, such as RT-qPCR, but those techniques are not as affordable and available to all due to their reliance on lab-based equipment and environments. As we saw in the previous chapter, this detection method is one example of a range of techniques that enable the amplification of the signal that a virus is present. These are all based on the positive recognition of a certain sequence of nucleotides (the target or template) by a designed sequence of nucleotides (the primer), which triggers a chain reaction of replication of that sequence (catalyzed by a polymerase), and quantified by increasing accumulation of a signal from an attached fluorescent molecule. This requires a readout method, often (as in the case of RT-qPCR) necessitating specialized lab equipment and energy input in order to provide the environment to trigger this chain reaction and detect its output. While research efforts are underway to miniaturize and

portabalize this equipment, alternative methods are being developed that obviate the equipment and provide acceptable sensitivity for the application [Panpradist et al., 2021b,a].

Building on my experience from previous projects in using smartphones to detect health biomarkers and detecting viruses using highly sensitive lab equipment, I reviewed the literature, the capabilities of the smartphone's various sensors, and I turned towards capacitance sensing of nucleic acid reactions for a potential future look at how we may detect the presence of viruses in the future. The capacitance sensor of most modern smartphones is used to sense and localize the touch of a user's finger, and has converged to a fairly standard design, despite some variations. In this chapter, I build on this standard design to explore the potential for sensing the minute chemical changes associated with DNA hybridization and polymerase-catalyzed extension. At this time, we are not yet at the stage where direct sensing of nucleic acid strands is clearly feasible, so I, along with my Ubicomp Lab and MISL Lab collaborators, built a low-cost tool that can be used in the lab to interrogate the capacitance characteristics of DNA solutions. We showed some potential signs that nucleic acid isothermal amplification reactions may be detectable at certain frequencies via a technique that is related to the mutual capacitance grid of electrodes in modern smartphones. The process of doing so illustrates the full circle of development of biomarker sensing using commodity sensors, starting from understanding the biology of the underlying molecule to attempt to amplify it.

See Chapter 6 for the full details on this project.

A.4.1 Conclusion: Towards More Ubiquitous Sensing of Nucleic Acid Reactions

As mentioned in Chapter 2, there are multiple options for molecular reaction output, but I chose to explore capacitive output in my research due to its association with a sensor available on all modern smartphones. While this is not yet proven in full on commodity capacitive sensors, I have demonstrated detection of small changes in solution contents on a lab-based development kit. The experiments I undertook, alongside my collaborators, to test the limits of these commodity sensors with the goal of detection of specific DNA sequences associated with particular viruses demonstrated the final piece of my thesis. We need to develop a biological understanding of the molecules of interest if we hope to sense them on ubiquitous sensors, such as the nearly globally saturated set of touchscreen smartphones that people are already using in their daily lives.

The future work items listed in this chapter also illustrate the thesis. Pursuing these items will enable us to move closer to detection from both sides of the sensing axis in Figure 1.1. We will need to understand capacitance-based nucleic acid sensing systems from the biology side, while also attempting to read the signal through the noise; noise that will likely grow when we move the sensing to the ubiquitous smartphone touchscreen. I hope that future researchers may pick up on this thread to further develop this work for more complex nucleic acid reactions, including DNA computation, and move closer to a reaction that can be sensed by the sensors already in use in modern phones to enable touchscreen interfaces.

In the next chapter, I will conclude my thesis by speculating and reflect upon the broader themes associated with ubiqitous computing and biotechnology. If we can solve problems such as the ones described in the preceding chapters, it could further democratize the methods of biotechnology to more scientists and members of the general population, accelerating the growth of biological research and increasing access to these screening tools and important health information for more people in the future.

A.5 Chapter 7 Summary: Conclusions and Reflections

Background: Lessons from Ubiquitous Computing to Biotech

In this chapter, I conclude my thesis, reflecting on learnings from my PhD research and speculating about how it may apply to future research in the health realm of the ubiquitous computing field. My four thesis projects, two in vivo sensing of blood-based health biomarkers and two in vitro nucleic acid sensing, demonstrate ways we can build solutions that enable improved access to sensing health-relevant signals using devices we already have in our pockets. The aim my work has been (and will continue to be) to empower more people to gather important information about their health so that they may act on it and improve their chances at a heatlhy life. More work will be needed in order to bring these innovations to impact healthcare in the wider world, and I find it likely that we will need to enlist the help of an increasing number and broader range of people in order to accomplish these goals. In my experience across computer science and biotechnology, I am struck by the difference in the approachability of computing technology compared to biotechnology. This leads me to envision the opportunity for bridging that gap, which will be critically important as we continue to push both of these fields forward.

See Chapter 7 for the full details on this project.

A.5.1 Conclusion: Ubiquitous Democratizatized Biotech

If we are to reap the benefits promised by ubiquitous computing technology applied to healthcare, we need to build technologies that can be used by more people. Taking lessons from the development of the computer science field, including democratization and wide use of applications of technology, seems to indicate the way forward. If we are driven by clearly valuable applications, we can accelerate development. If we can focus the aims of the research on elements that impact the widest possible set of people, we can democratize the innovations so that all can benefit. My research at the intersection of computer science and biomedical engineering can pave a path towards that future, and I hope to continue developing applications that enable more people to access these exciting innovations and improve healthcare along the way.

Appendix B

Supplementary Information: Phone Camera Oximetry

Supplementary Materials

Smartphone Camera Oximetry in an Induced Hypoxemia Study

Authors: Jason S. Hoffman,[1*†] Varun Viswanath,[2,3†] Caiwei Tian,[1] Xinyi Ding,[4] Matthew J. Thompson,[5] Eric C. Larson,[4] Shwetak N. Patel,[1,6] Edward Wang[2,3]

Affiliations:
[1]Paul G. Allen School of Computer Science and Engineering, University of Washington
[2]Department of Electrical and Computer Engineering, University of California San Diego
[3]The Design Lab, University of California San Diego
[4]Department of Computer Science, Southern Methodist University
[5]Department of Family Medicine, University of Washington
[6]Department of Electrical and Computer Engineering, University of Washington

[*]To whom correspondence should be addressed; E-mail: jasonhof@cs.washington.edu.
[†]These authors contributed equally to this work.

B.1 Supplementary Information

B.1.1 Benchmark Ratio-of Ratios Model Results

Subject #	1	2	3	4	5	6	Mean ± SD
MAE	10.58	6.87	7.39	9.54	6.78	7.31	7.12 ± 1.64
A_{rms}	13.28	8.52	10.58	10.92	8.49	9.41	10.20 ± 1.66
R^2	0.17	0.06	0.02	0.04	0.47	0.52	0.21 ± 0.20
μ / LOA	6.38 / 22.83	-0.79 / 16.63	-5.38 / 17.85	-9.07 / 11.90	2.05 / 16.15	3.51 / 17.12	1.24 ± 5.16 / 17.08 ± 3.2

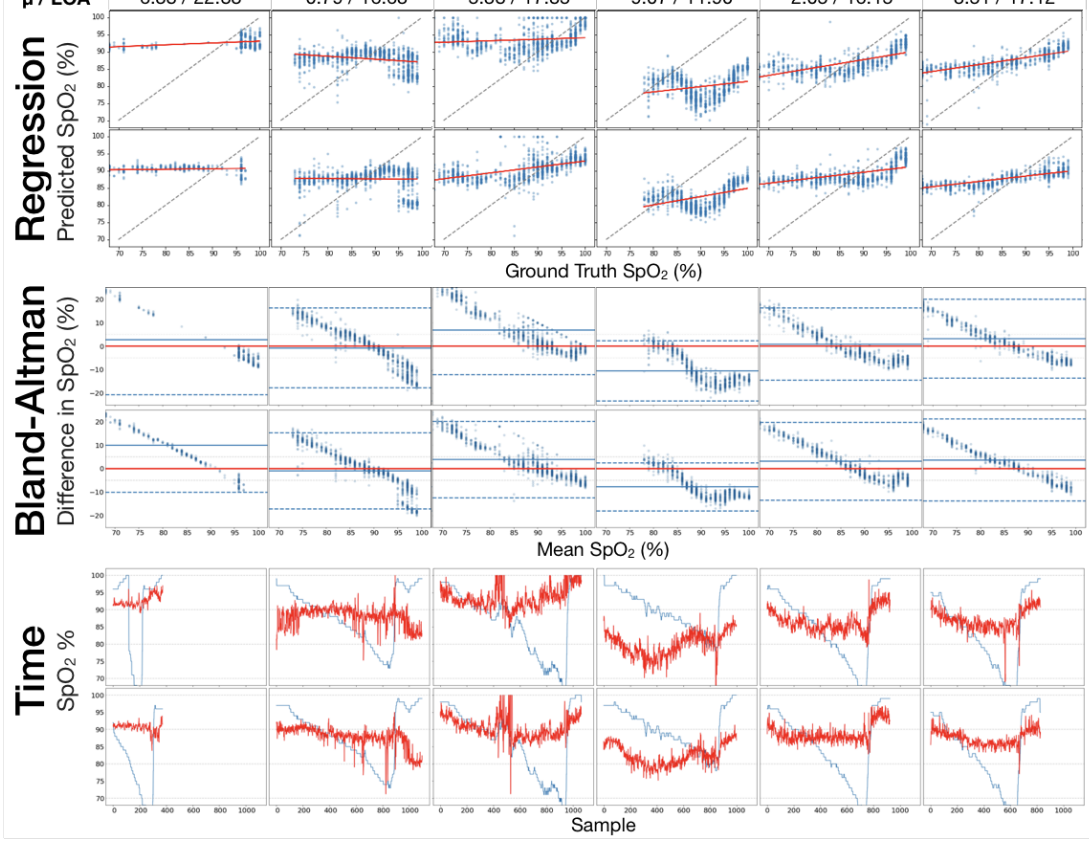


Figure B.1: Benchmark Ratio-of-Ratios Model Results. Applying the ratio-of-ratios model from [Nemcova et al., 2020] shows that this model does find a pattern in the RGB data gathered from a smartphone, but overall does not perform as well as the CNN model. The average MAE increases by 2.12 to 7.12 (σ=1.64). On certain portions of the data, such as the portion of Subject 1's data where the peaks are dampened due to calluses attenuating the signal, the slope of the PPG rise falls below 1 and the ratio-of-ratios model is undefined, and thus those samples are dropped from this analysis.

B.1.2 Including training data <70% SpO_2

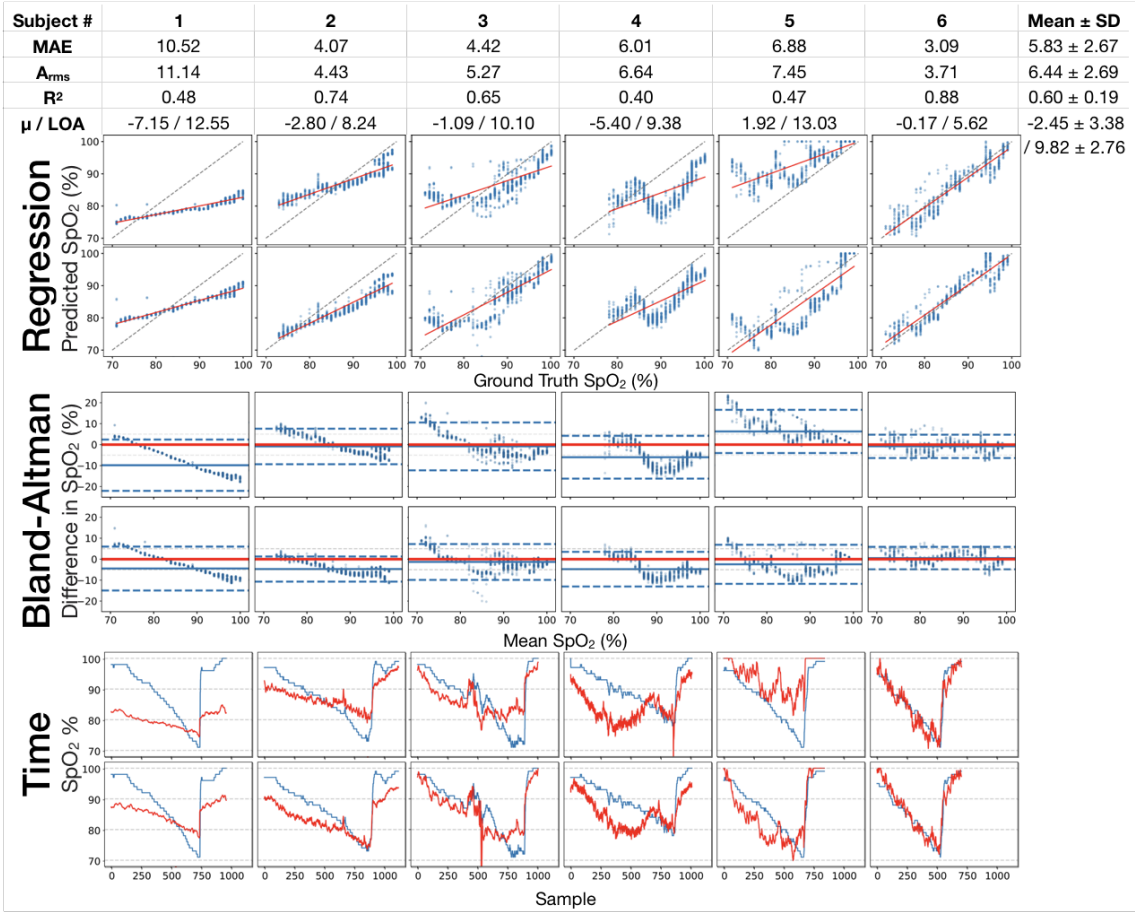


Figure B.2: Including training data <70% SpO_2. Including all data in training the model, including data below 70% SpO_2 ground truth, reduces accuracy of SpO_2 inference. This is due to the fact that the ground truth pulse oximeter used as a transfer standard for comparison in this study is not validated to be accurate below 70% SpO_2, as it is only required to be validated in the 70% to 100% SpO_2 range, as per the ISO standard 80601-2-61:2017 [Food et al., 2013]. Increased error in the training target for training examples below 70% negatively impacts training, reducing the overall accuracy of the resulting model and validity of using the model for inferring SpO_2 below 70%. Additionally, the study was designed so that the target for minimum SpO_2 level was 70%, and thus, only 2/6 subjects had a significant distribution of data below 70% to train and evaluate. Due to these factors, the analysis of data collected above 70% SpO_2 is performed throughout the paper.

B.1.3 Regression Results after normalizing heart rate

Subject #	1	2	3	4	5	6	Mean ± SD
MAE	9.14	3.65	4.02	5.47	6.47	3.33	5.35 ± 2.01
A_{rms}	9.69	4.03	4.92	6.00	7.00	3.86	5.92 ± 2.01
R^2	0.51	0.70	0.53	0.28	0.43	0.82	0.54 ± 0.18
μ / LOA	-3.62 / 12.65	-1.85 / 7.43	-0.46 / 10.24	-5.29 / 8.22	2.52 / 12.43	0.21 / 5.28	-1.41 ± 2.55 / 9.37 ± 2.67

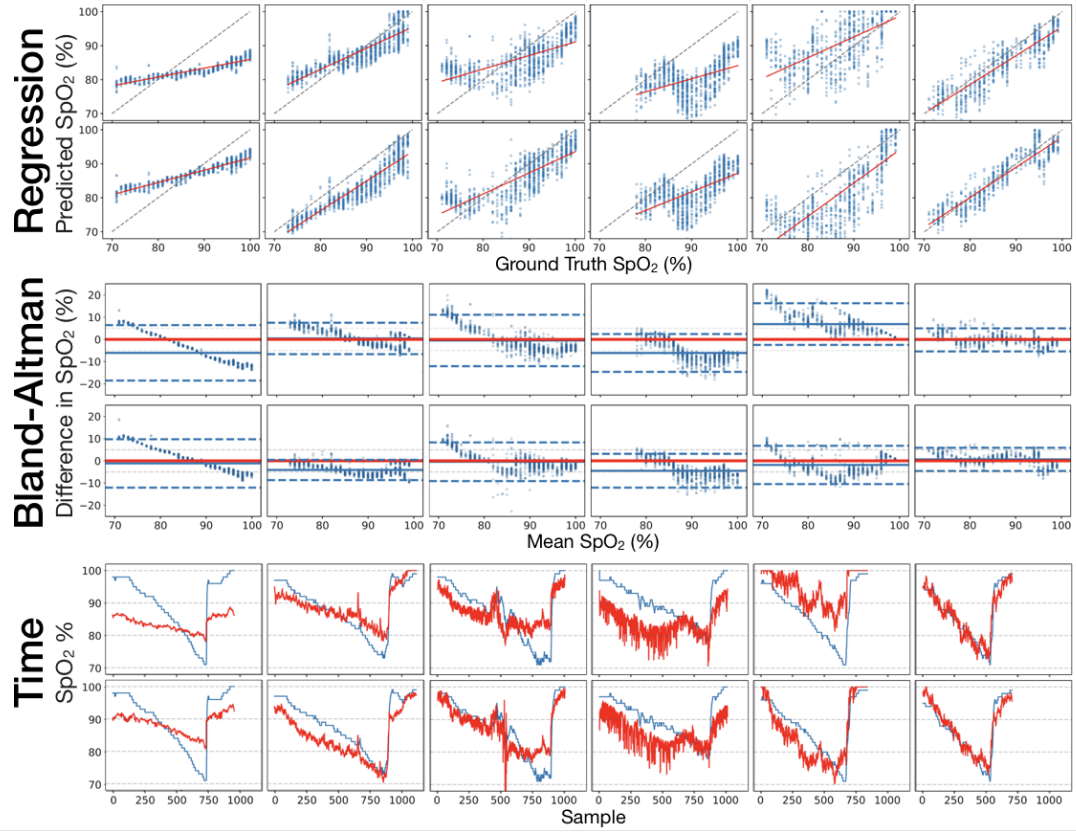


Figure B.3: Removing heart rate. After resampling the data to detect and remove the heartrate, and training the model on 3 beats of 60bpm PPG data, the MAE increases by 0.35 to 5.35 (σ=1.90) and R^2 decreases by 0.07 to 0.54, and additional spread in the predictions is observable for a couple test subjects. This indicates that heart rate may be contributing to the model predictions; however, this cannot explain the full predictive power of the model, as some of this increased error is likely contributed by the HR detection and resampling process.

B.1.4 Subject histograms

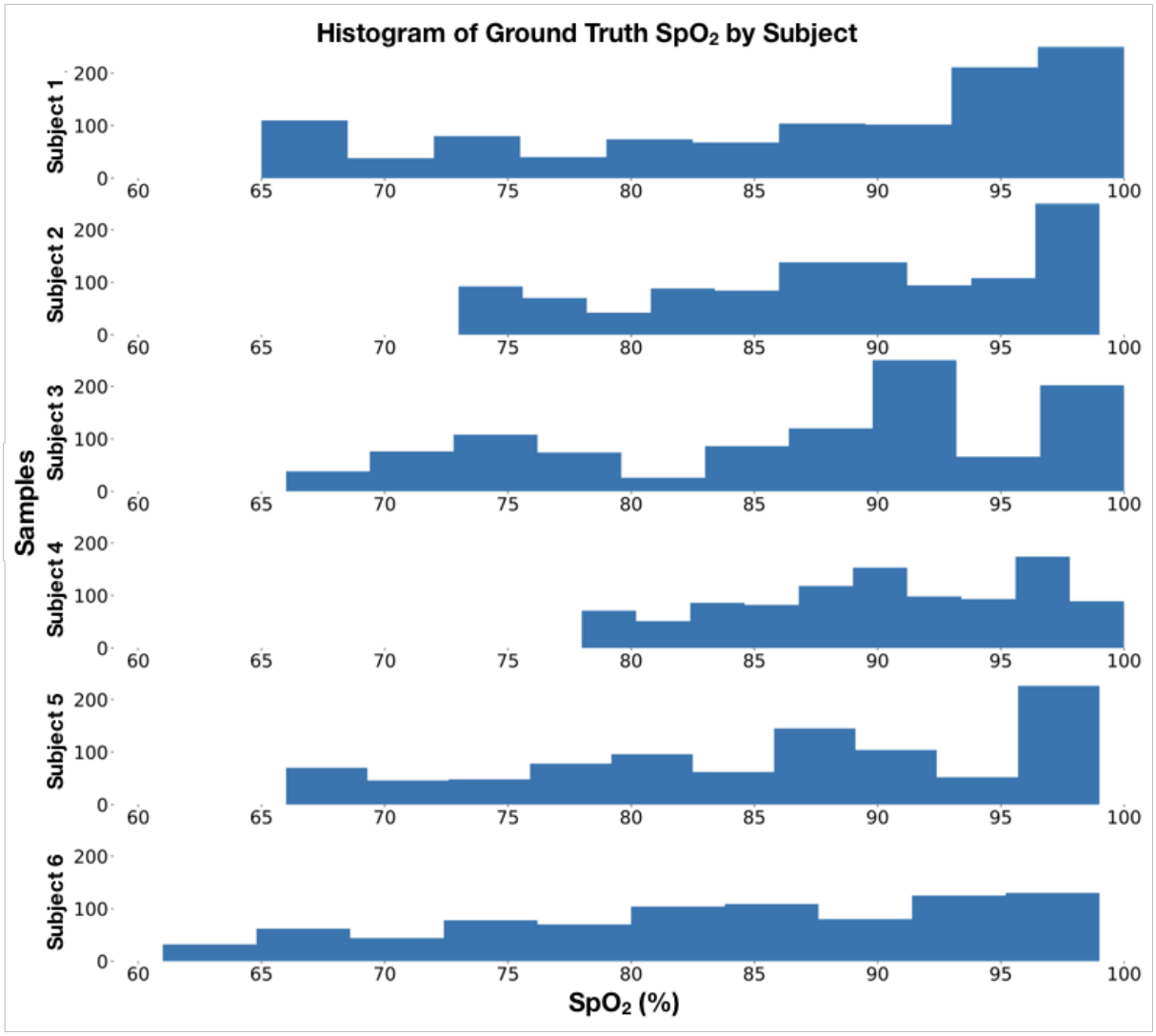


Figure B.4: Subject Histograms. Histograms of ground truth samples gathered from from all 6 test subjects in the study show that all but one subject have a significant spread of data between 70-100%, while only half of the subjects have more than a couple samples below 70%, the minimum SpO_2 level defined in the study protocol.

B.1.5 Ground truth variation

Ground Truth Pulse Oximeter Device	Difference compared to Masimo Radical 7	
	Mean Diff	Std Diff
3900P TT+	0.01	1.79
Nellcor N-600X 2	0.70	2.94
Nellcor N-600X 4	-0.38	1.59

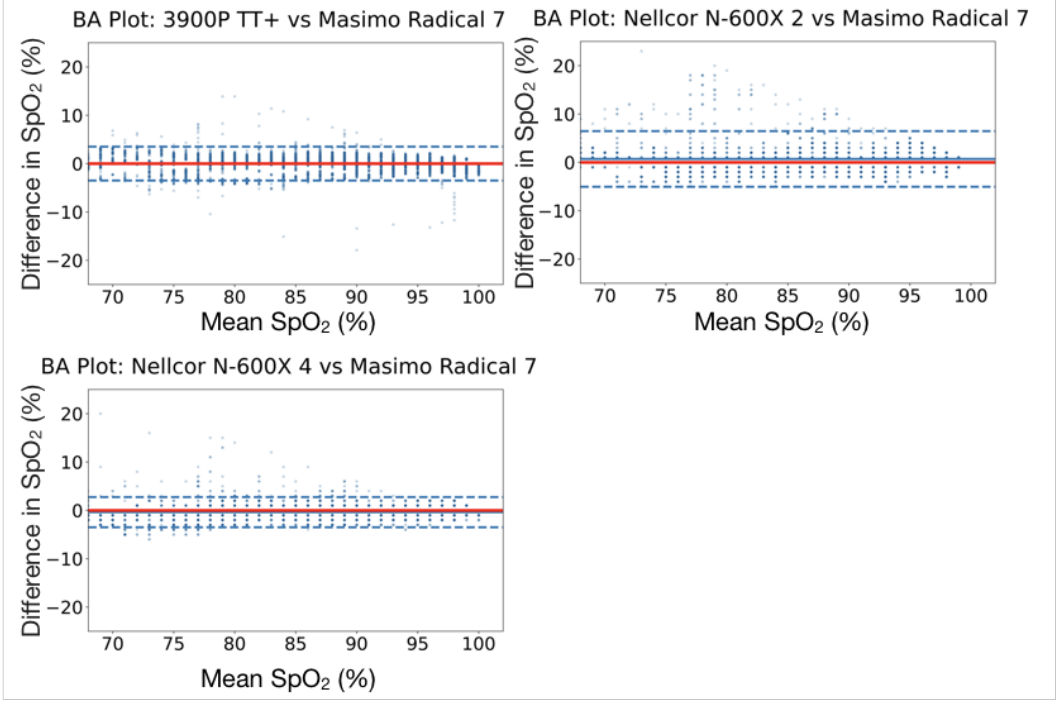


Figure B.5: Ground Truth Variation. Difference plots show that there is some variation in the other pulse oximeters that measured SpO_2 level on each subject, compared to the tight tolerance reference pulse oximeter that was used to measure ground truth in this study (Masimo Radical-7). The variation, measured by mean difference was between -0.38 and 0.7, with Standard Difference between 1.59 and 2.94. This variation may have been due to the error range of the pulse oximeters, slight differences in timing between the readings of the pulse oximeters, or differences in value of SpO_2 between the subjects' different hands or fingers.

B.1.6 Open source data table

Subject in Study	Label in Data		File names	
	Num	Label	Left Hand Video Right Hand Video	Info File Ground Truth Data File
1	1	100001	100001-1487003054311-0-1487003073393.mp4 100001-1487003016146-0-1487003016393.mp4	100001-1487003016146.info 100001.csv
2	2	100002	100002-1487006909985-0-1487006910227.mp4 100002-1487006911581-0-1487006911849.mp4	100002-1487006911581.info 100002.csv
3	3	100003	100003-1487010210734-0-1487010211003.mp4 100003-1487010212332-0-1487010212594.mp4	100003-1487010212332.info 100003.csv
4	4	100004	100004-1487016701771-0-1487016702172.mp4 100004-1487016703619-0-1487016703979.mp4	100004-1487016703619.info 100004.csv
5	5	100005	100005-1487019989533-0-1487019989854.mp4 100005-1487019992000-0-1487019992346.mp4	100005-1487019992000.info 100005.csv
6	6	100006	100006-1487023968612-0-1487023968853.mp4 100006-1487023971109-0-1487023971496.mp4	100006-1487023971109.info 100006.csv

Figure B.6: Open Source Data. Info on filenames in the open source data associated with subject numbers in this study.

Supplementary Data 1. Zip file of open source camera oximetry data

A Github repository has been set up with a Zip file of the data collected in this study, documentation, and example data-loading code:

https://github.com/ubicomplab/oximetry-phone-cam-data.

Appendix C

Supplementary Information: Phone Hemoglobin Sensing

C.1 Supplementary Methods

C.1.1 Enrollment Criteria

Patients aged 6 months to 18 years who had undergone a venous blood draw for a CBC since being admitted to the hospital no more than 24 hours prior to study enrollment were eligible for enrollment. Patients with total leukonychia, nailbed darkening or discoloration due to medication, nail clubbing, clinically indicated jaundice, subungual hematoma, nail bed lacerations, avulsion injuries on both hands or nail polish applied on fingernails were not eligible for study recruitment.

C.1.2 Data Collection

Patients were instructed to place their hands in a customized chamber in a splayed position for data collection. The customized chamber was designed to create a study setting whereby the patient's hand could be photographed and filmed with a smartphone without any ambient lighting and with the distance and angle of the camera and fingernails held at a constant. Supplementary Figure C.1 provides the chamber's schematics and depicts the chamber in the study setting. An Android-native application (app) was designed and developed for the purpose of the study data collection. The app was programmed to capture four separate photographs and two 15-second video clips of an individual's hands along with the image metadata. The following phone and camera settings were held constant: exposure, focus distance, white balance setting, gain (red, green and blue separately), and flash. Accelerometer shakiness was held steady using the sample collection chamber depicted in Supplementary Figure C.1. Finally, the app recorded patient ID, sex, age and all CBC results.

C.1.3 Data Preprocessing

Fingernails from patient images were manually tagged and used to post-train an object-detection classifier, YOLOv8 (You Only Look Once), to detect patient fingernails from a set of images13. Video clips were automatically detected and parsed by frame and the post-trained YOLO model was applied to the detected frames for automatic fingernail detection. Each fingernail image was cropped to 50x50 pixels by cropping the edges of the fingernail and using only the center of the fingernail. A quality assessment pipeline that

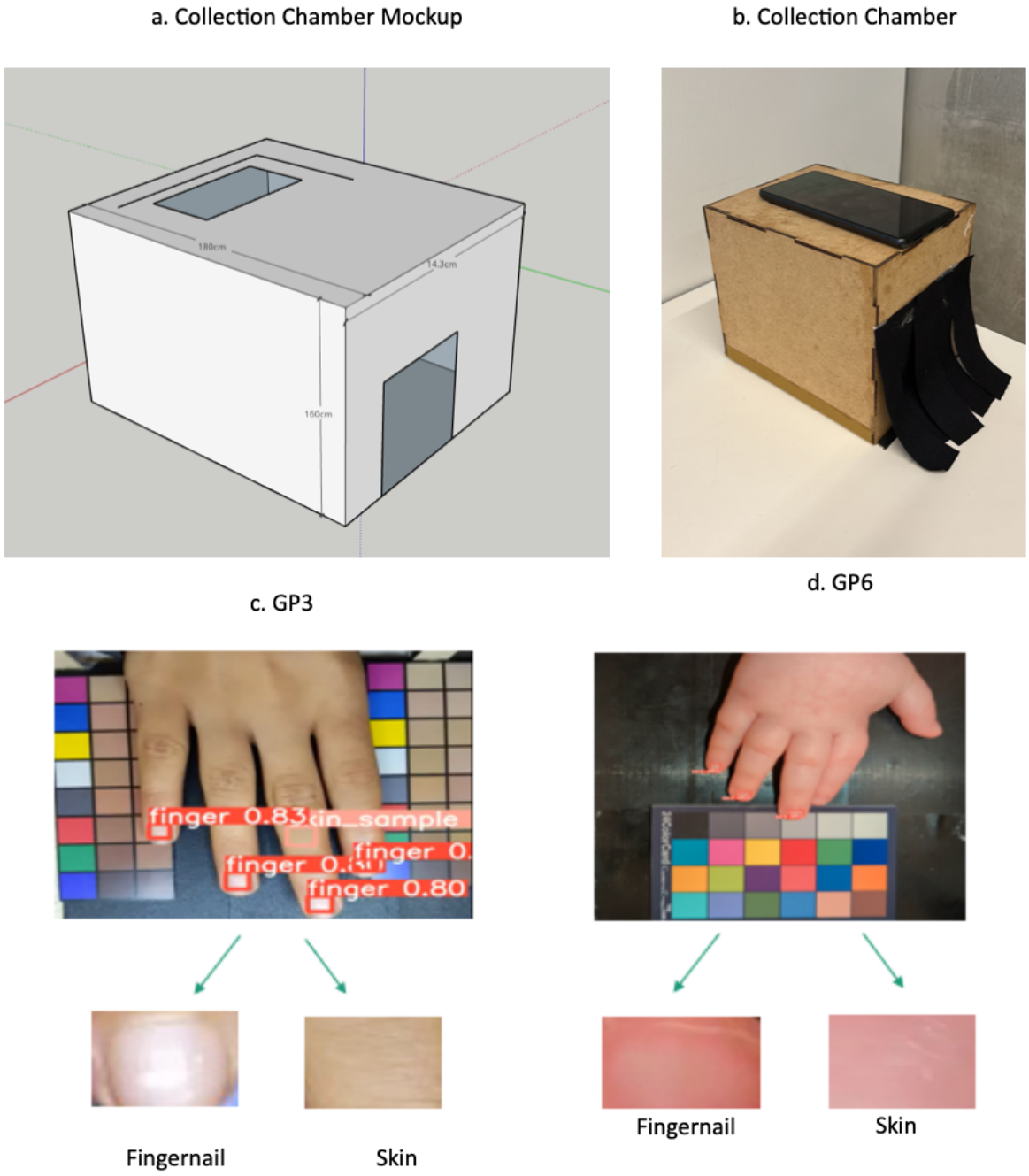


Figure C.1: a. Schematics of collection chamber with aperture for smartphone camera. **b.** Collection chamber in the study setting **c.** Hand, fingernail and skin image from GP3. Fingernails and skin detected from images or videos taken with the GP3 were each 20X20 pixels (400 pixels total), **d.** Hand, fingernail and skin image from GP6. Fingernails and skin detected from images or video taken with the GP6 were 80X80 pixels (6,400 pixels total).

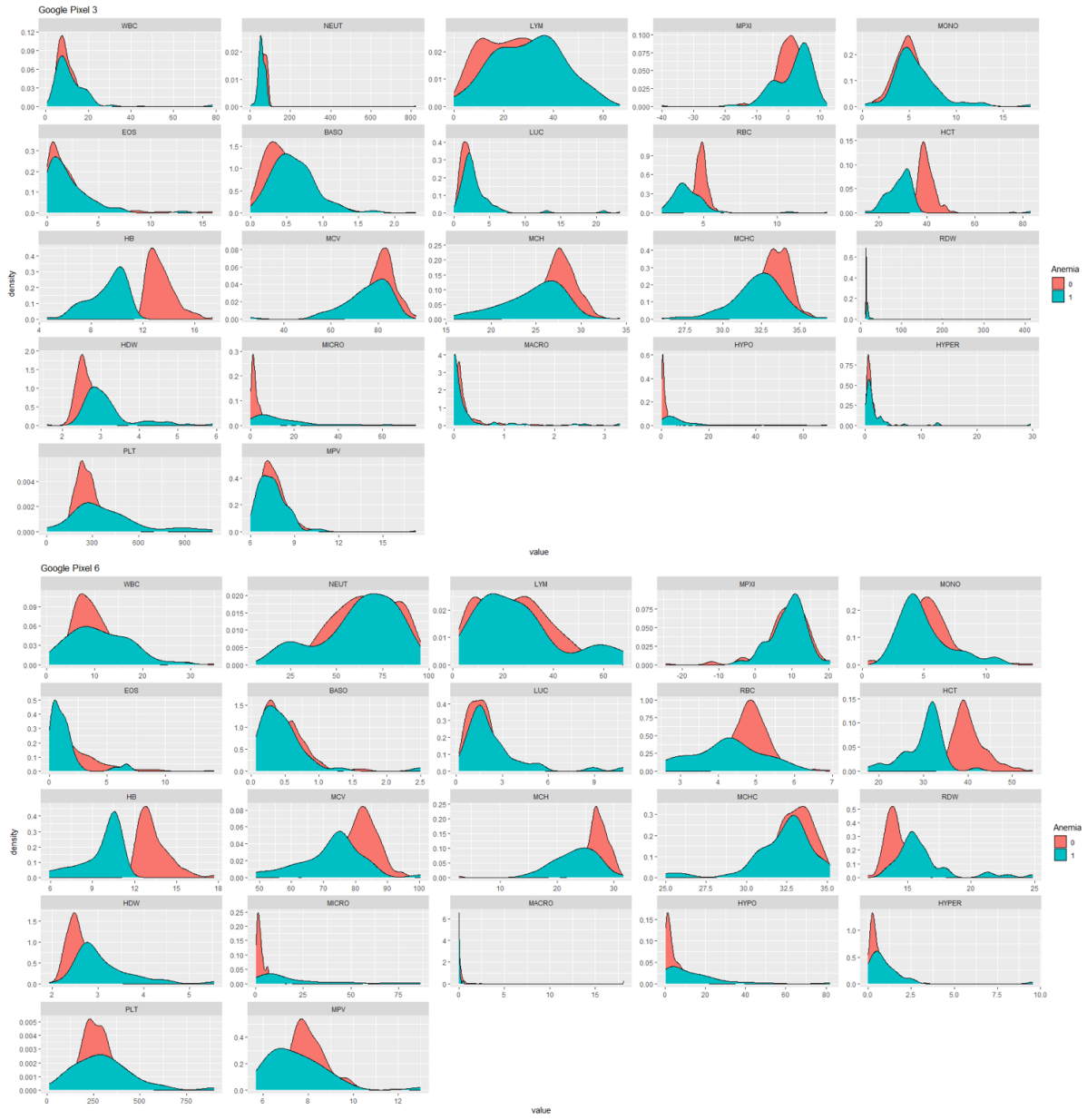


Figure C.2: Histograms depicting the distribution of blood hemoglobin levels in the cohort is presented above. Samples taken using the GP3 are depicted in the upper panel, and samples taken using the GP6 are depicted in the lower panel.

	GP3					GP6				
Model	Accuracy	Precision	Recall	f1	ROC AUC	Accuracy	Precision	Recall	f1	ROC AUC
1	0.5	0.5	0.4	0.44	0.48	0.8	0.86	0.86	0.86	0.76
2	0.6	1	0.33	0.5	0.83	0.9	1	0.86	0.92	0.86
3	0.5	0.5	0.4	0.44	0.48	0.8	0.8	0.8	0.8	0.92
4	0.4	0.5	0.17	0.25	0.38	0.7	0.6	0.75	0.67	0.75
5	0.5	0.67	0.33	0.44	0.54	0.8	0.83	0.83	0.83	0.92
6	0.5	0.43	0.75	0.55	0.33	0.7	0.71	0.83	0.77	0.58
7	0.6	1	0.43	0.6	0.71	0.9	0.86	1	0.92	0.92
8	0.5	0.5	0.4	0.44	0.48	0.7	0.75	0.6	0.67	0.8
Mean	0.51	0.64	0.4	0.46	0.53	0.79	0.8	0.82	0.81	0.81
STD	0.06	0.23	0.16	0.1	0.17	0.08	0.12	0.11	0.1	0.12

Table C.1: XGBoost Model Performance for the Classification of Anemia From Images Taken With a GP3 and GP6

averaged the blur (Laplacian variance), saturation and brightness was applied. This was calculated by dividing the Laplacian variance of each image by 100 (the default threshold), the average saturation by 0.15 and the average brightness by 0.15. A weighted between 0 and 1 was output for each. The composite quality assessment score was calculated by averaging the weighted individual scores. The 100 highest-graded images from each individual were selected and combined into a single composite image. Composite image pixels were extracted and converted into the following 5 color spaces; RGB, HLS, LAB, LUV and HSV. A histogram representing the color intensity was constructed for each of the 15 color channels and was input into the machine learning model. Additionally, a 50x50 pixel area of skin from the middle finger was detected and automatically matched to its corresponding Monk skin tone (MST) swatch which was input into the machine learning model along with the relative error of detected MST to account for inter-level variabiality14. Figure C.1a and C.1b depicts the data preprocessing workflow.

Model	Accuracy	Positive Predictive Value	f1	ROC AUC	Sensitivity	Specificity
1	0.86	0.8	0.87	0.92	0.95	0.78
2	0.86	0.71	0.83	0.94	1	0.79
3	0.84	0.94	0.83	0.9	0.74	0.95
4	0.91	0.88	0.91	0.96	0.95	0.86
5	0.8	0.83	0.77	0.87	0.71	0.87
6	0.86	0.95	0.88	0.93	0.81	0.94
7	0.84	0.87	0.85	0.93	0.83	0.85
8	0.86	0.85	0.88	0.93	0.92	0.79
9	0.95	1	0.96	0.99	0.92	1
10	0.84	0.86	0.84	0.9	0.83	0.86
11	0.86	0.95	0.88	0.96	0.81	0.94
12	0.82	0.85	0.85	0.91	0.85	0.78
13	0.84	0.74	0.83	0.99	0.94	0.77
14	0.93	0.92	0.94	0.98	0.96	0.9
15	0.8	0.75	0.8	0.92	0.86	0.74
16	0.86	0.83	0.86	0.93	0.9	0.83
17	0.84	0.85	0.87	0.93	0.88	0.78
18	0.77	0.69	0.69	0.8	0.69	0.82
19	0.82	0.78	0.82	0.95	0.86	0.78
20	0.89	0.83	0.88	0.95	0.95	0.83
Mean	0.85	0.84	0.85	0.93	0.87	0.84
STD	0.04	0.08	0.06	0.04	0.08	0.07

Table C.2: XGBoost Model Performance for the Classification of Anemia from Images Taken With GP6 Using the SMOTE Technique

Appendix D

Supplementary Information: Passively Monitoring the Presence of RNA in Public Transit Infrastructure

D.1 Supplementary Information

Passively Sensing SARS-CoV-2 RNA in Public Transit Buses

Jason S. Hoffman[a], Matthew Hirano[b], Nuttada Panpradist[c], Joseph Breda[a], Parker Ruth[a,c], Yuanyi Xu[d,e], Jonathan Lester[f], Bichlien H. Nguyen[f], Luis Ceze[a], Shwetak N. Patel[a,b]

[a]Paul G. Allen School of Computer Science and Engineering, University of Washington, Seattle, WA, USA
[b]Department of Electrical and Computer Engineering, University of Washington
[c]Department of Bioengineering, University of Washington
[d]Department of Microbiology, University of Washington
[e]Department of Chemistry, University of Washington
[f]Microsoft Research, Redmond, WA, USA

D.1.1 Testing in-house extraction protocol on contrived samples

In each RT-qPCR assay, we included a negative control of two replicates containing nuclease-free water (AAJ71786AE, Fisher Scientific) that underwent the TRIzol isolation steps in parallel to the samples. Each assay also included three positive controls containing 10, 100, or 1000 copies of purified RNA controls including the N gene regions of SARS-CoV-2 (102024, Twist Biosciences, San Francisco, CA). To test the efficiency of TRIzol isolation of RNA from viral envelopes, the TRIzol extraction and RT-qPCR method was performed on AccuPlex enveloped RNA reference material (0505-0126, Seracare, Milford, MA), and results are displayed in Figure D.1 below. In addition, a spike control experiment was performed with MCE filters to assess the extraction efficiency of the method, and results are displayed in Figure D.3.

PCR curves from controls testing

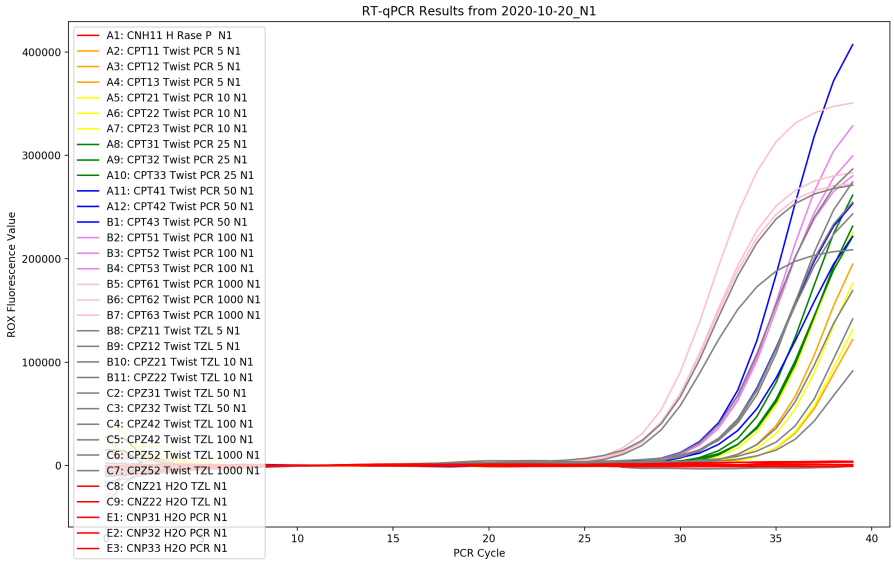


Figure D.1: Controls PCR curves for N1 gene.

Table of results from controls testing

Solution	Sol Name	Copies	CT Ave	CT 1 Res	CT 2 Res	CT 3 Res
Direct RNA PCR (Twist)	RNA PCR	5	36.10	no amp	35.4	36.79
Direct RNA PCR (Twist)	RNA PCR	10	35.72	35.8	34.68	36.68
Direct RNA PCR (Twist)	RNA PCR	25	34.37	34.47	34.08	34.55
Direct RNA PCR (Twist)	RNA PCR	50	33.07	33.17	33.77	32.27
Direct RNA PCR (Twist)	RNA PCR	100	32.46	32.36	32.49	32.52
Direct RNA PCR (Twist)	RNA PCR	1000	29.14	28.83	29.32	29.28
RNA through Trizol Extraction (Twist)	RNA TZL	5	37.28	37.28	no amp	N/A
RNA through Trizol Extraction (Twist)	RNA TZL	10	36.00	36.44	35.55	N/A
RNA through Trizol Extraction (Twist)	RNA TZL	50	33.20	33.23	33.17	N/A
RNA through Trizol Extraction (Twist)	RNA TZL	100	32.85	32.37	33.33	N/A
RNA through Trizol Extraction (Twist)	RNA TZL	1000	29.52	29.37	29.66	N/A
Spike Control Virus Trizol Extraction (AccuPlex)	Accu TZL	10	34.71	33.32	35.6	35.22
Spike Control Virus Trizol Extraction (AccuPlex)	Accu TZL	25	33.88	34.04	32.72	34.87
Spike Control Virus Trizol Extraction (AccuPlex)	Accu TZL	50	33.00	32.4	33.38	33.21
Spike Control Virus Trizol Extraction (AccuPlex)	Accu TZL	100	32.23	33.32	31.56	31.82
Spike Control Virus Trizol Extraction (AccuPlex)	Accu TZL	1000	28.62	28.14	29	28.73

Table D.1: CT results for controls. All controls above 10 copies/sample amplified in our testing. For copy numbers below the LOD of the PCR reagents (10 copies/sample), amplification still occurred intermittently.

qPCR Standard Curve

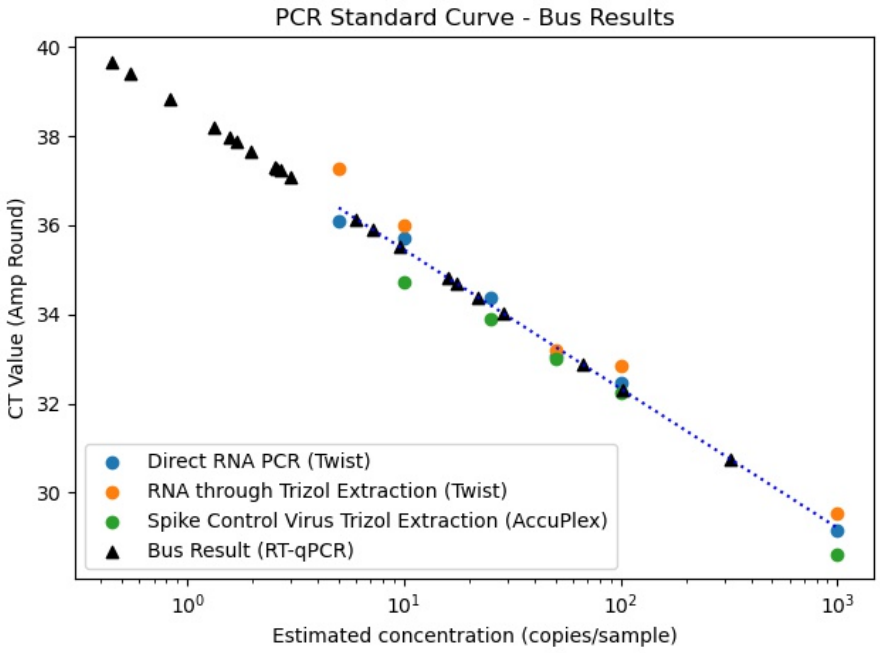


Figure D.2: Creating a standard curve from the controls samples and fitting the CT results from bus runs onto that standard curve reveals most of the bus results contained a small estimated copy number of below 10 viral cells per 8 uL of extracted sample.

D.1.2 Extraction Efficiency - Drip Control Experiment

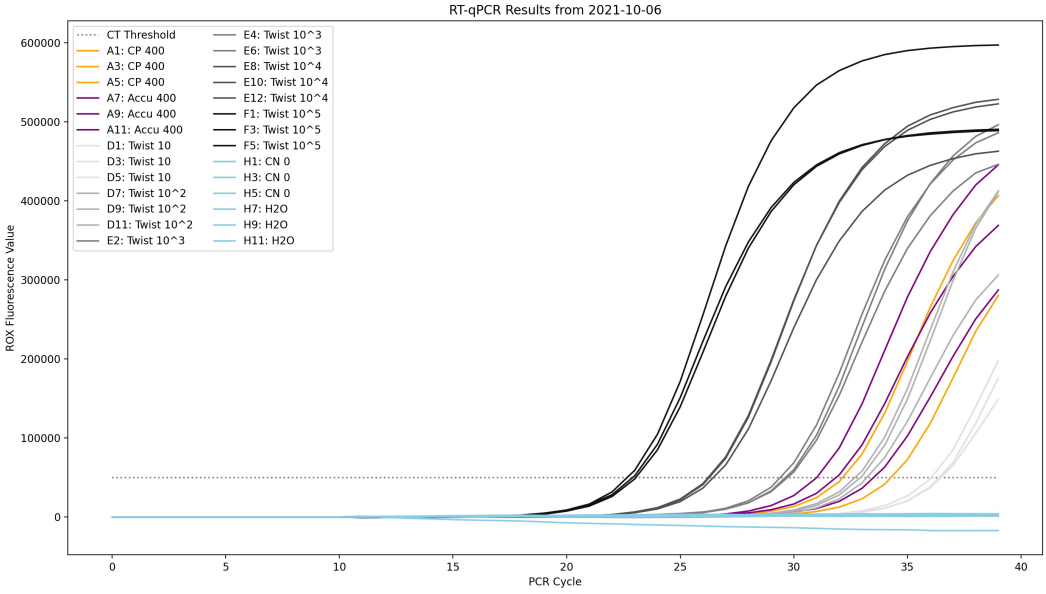


Figure D.3: Controls results from spike control experiment with filters to test extraction efficiency after dripping control onto filters.

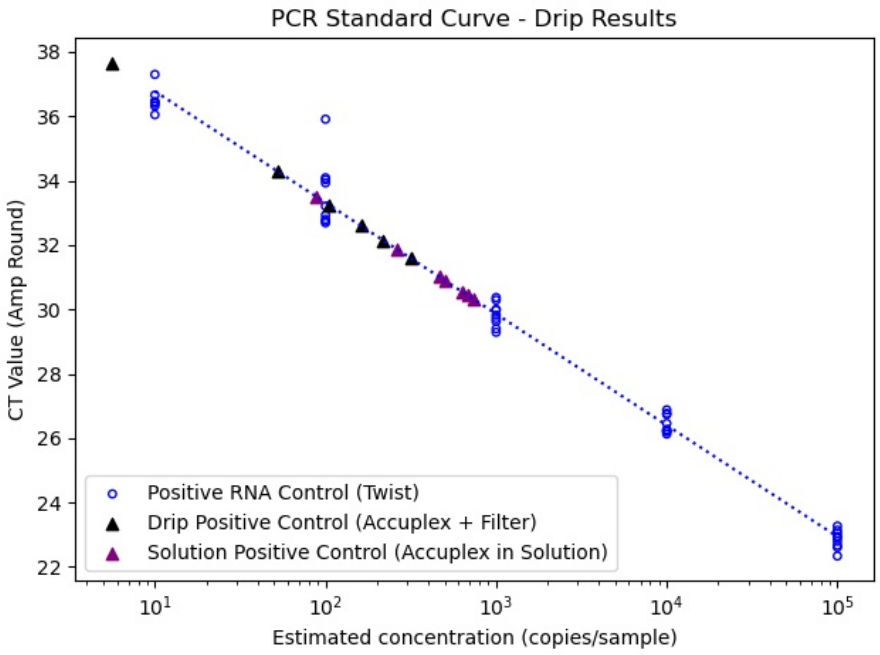


Figure D.4: Creating a standard curve from the drip controls experiment and fitting the CT results from extraction drip controls reveals that we lose about half of the copies placed from filter extraction and half again from RNA extraction.

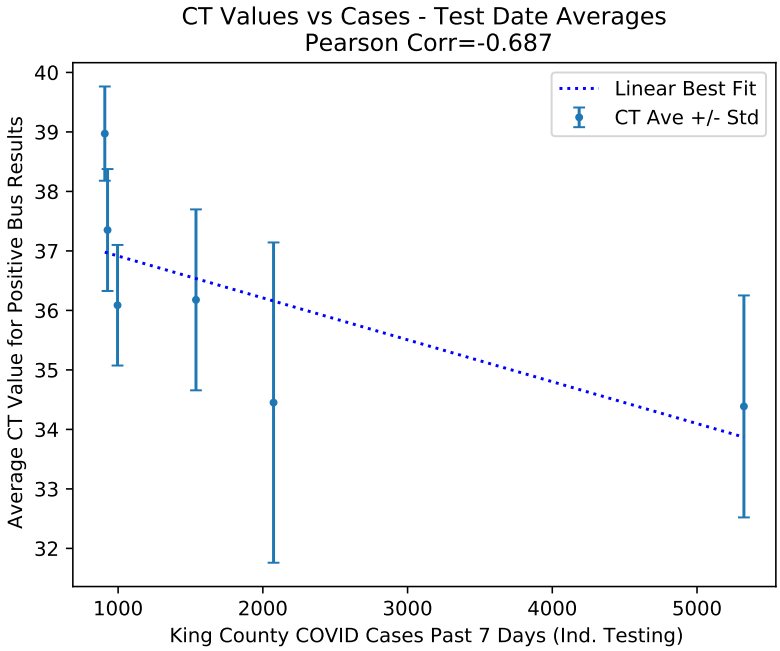


Figure D.5: CT results for positive results from buses were compared the sum of daily cases in King County for the past 7 days, and a Pearson's Correlation of -0.687 was observed.

Solution	Copies in Sol	CT Rep 1 (N1)	CT Rep 2 (N1)	CT Rep 3 (N1)	CT Average	CT St Dev	Copies Rep 1	Copies Rep 2	Copies Rep 3	Copies Extracted	Extraction Efficiency
CP Drip on Filter	400	32.15	No detect	34.27	33.21	1.50	173.42	N/A	40.32	106.87	26.72%
CP Direct in Sol	400	33.51	31.01	31.86	32.13	1.27	68.02	380.01	211.72	219.92	54.98%
Twist RNA	10	36.39	36.05	36.46	36.30	0.22					
Twist RNA	100	33.21	32.93	32.69	32.94	0.26					
Twist RNA	1000	29.4	29.63	29.72	29.58	0.17					
Twist RNA	10000	26.46	26.27	26.23	26.32	0.12					
Twist RNA	100000	22.91	22.67	23.05	22.88	0.19					
CN Extraction	0	No detect	No detect	No detect	N/A	N/A					
CN PCR	0	No detect	No detect	No detect	N/A	N/A					
Solution	**Copies in Sol**	**CT Rep 1 (N1)**	**CT Rep 2 (N1)**	**CT Rep 3 (N1)**	**CT Average**	**CT St Dev**	**Copies Rep 1**	**Copies Rep 2**	**Copies Rep 3**	**Copies Extracted**	**Extraction Efficiency**
CP Drip on Filter	400	No detect	31.58	37.66	34.62	4.30	N/A	317.94	5.55	161.75	40.44%
CP Direct in Sol	400	No detect	30.31	No detect	30.31	N/A	N/A	740.51	N/A	740.51	185.13%
Twist RNA	10	No detect	36.66	36.32	36.49	0.24					
Twist RNA	100	34.09	35.91	32.8	34.27	1.56					
Twist RNA	1000	29.82	29.29	30.29	29.80	0.50					
Twist RNA	10000	26.21	26.21	26.13	26.18	0.05					
Twist RNA	100000	22.61	22.34	22.8	22.58	0.23					
CN Extraction	0	No detect	No detect	No detect	N/A	N/A					
CN PCR	0	No detect	No detect	No detect	N/A	N/A					
Solution	**Copies in Sol**	**CT Rep 1 (N1)**	**CT Rep 2 (N1)**	**CT Rep 3 (N1)**	**CT Average**	**CT St Dev**	**Copies Rep 1**	**Copies Rep 2**	**Copies Rep 3**	**Copies Extracted**	**Extraction Efficiency**
CP Drip on Filter	400	33.24	No detect	32.6	32.92	0.45	105.29	N/A	161.24	133.27	33.32%
CP Direct in Sol	400	30.44	30.55	30.9	30.63	0.24	679.12	631.17	499.99	603.43	150.86%
Twist RNA	10	36.44	No detect	37.3	36.87	0.61					
Twist RNA	100	34.03	33.93	32.74	33.57	0.72					
Twist RNA	1000	30	29.98	30.37	30.12	0.22					
Twist RNA	10000	26.88	26.77	26.74	26.80	0.07					
Twist RNA	100000	23.14	22.95	23.27	23.12	0.16					
CN Extraction	0	No detect	37.26	No detect	N/A	N/A					
CN PCR	0	No detect	No detect	No detect	N/A	N/A					

n	Totals		
9	pos % (tot drip ctrl)	6	66.7%
9	pos % (tot sol ctrl)	7	77.8%
18	pos % (neg ctrl)	1	5.6%

Table D.2: CT results for drip control experiment. 66.7% (6/9) of positive filter extraction drip controls amplified and 5.6% (1/18) of negative controls amplified. Efficiency of extraction is calculated based on standard curve generated within the experiment of a PCR dilution ladder of RNA controls (Twist).

Legend Key

For Figures A7-B10:

Color, Well Number (96-well plate): Unique ID, Sample Description (including bus vehicle ID number and sample location or control type), Gene Target.

Examples:

- A4 = Column A, Row 4 of a 96-well PCR plate
- BFN12 = Bus sample, Front bus air filter, (N) Mixed Cellulose Ester filter material, 1st sample from that bus, 2nd replicate from that sample
- 4552 = Bus vehicle ID 4552
- CP12 = Control, Positive, 1st control sample, 2nd replicate of that control

Key:

- G = Polypropylene Fabric
- F = Biopsy Foam
- P = Paper
- N = Mixed Cellulose Ester
- E = EnviroMax Swab
- S = Swab Rails
- BS = Swab filter
- B, F, FS, BF = Air filter sample

PCR curves from bus testing

The method was also validated on environmental samples collected from buses. The procedure detected RNA from SARS-CoV-2 using both the air filter and the swab method. The results can be seen below in Figures S2-S5.

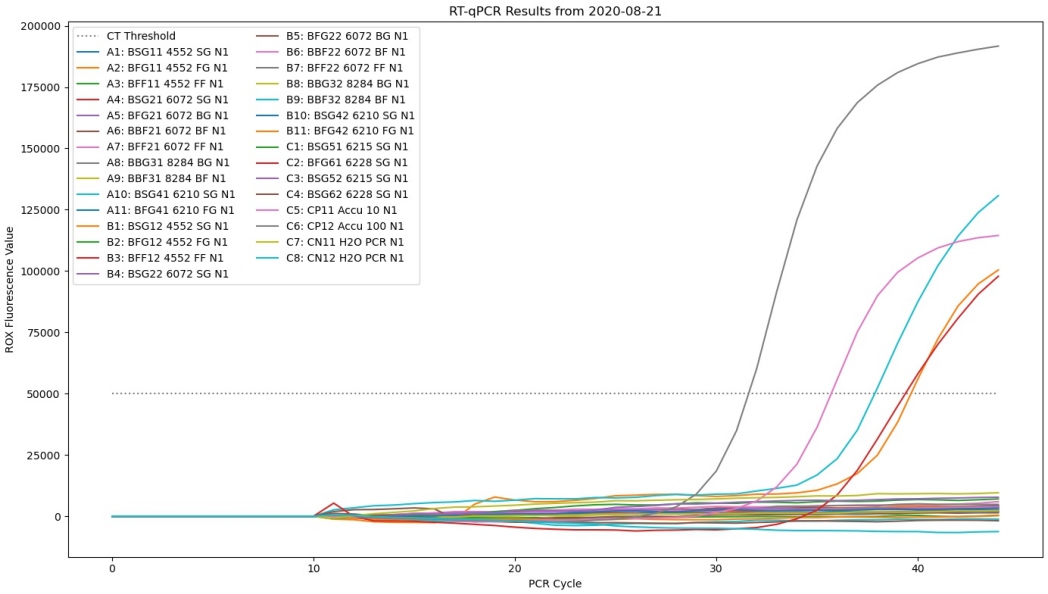


Figure D.6: Bus PCRs from first positive test result in Aug 2020.

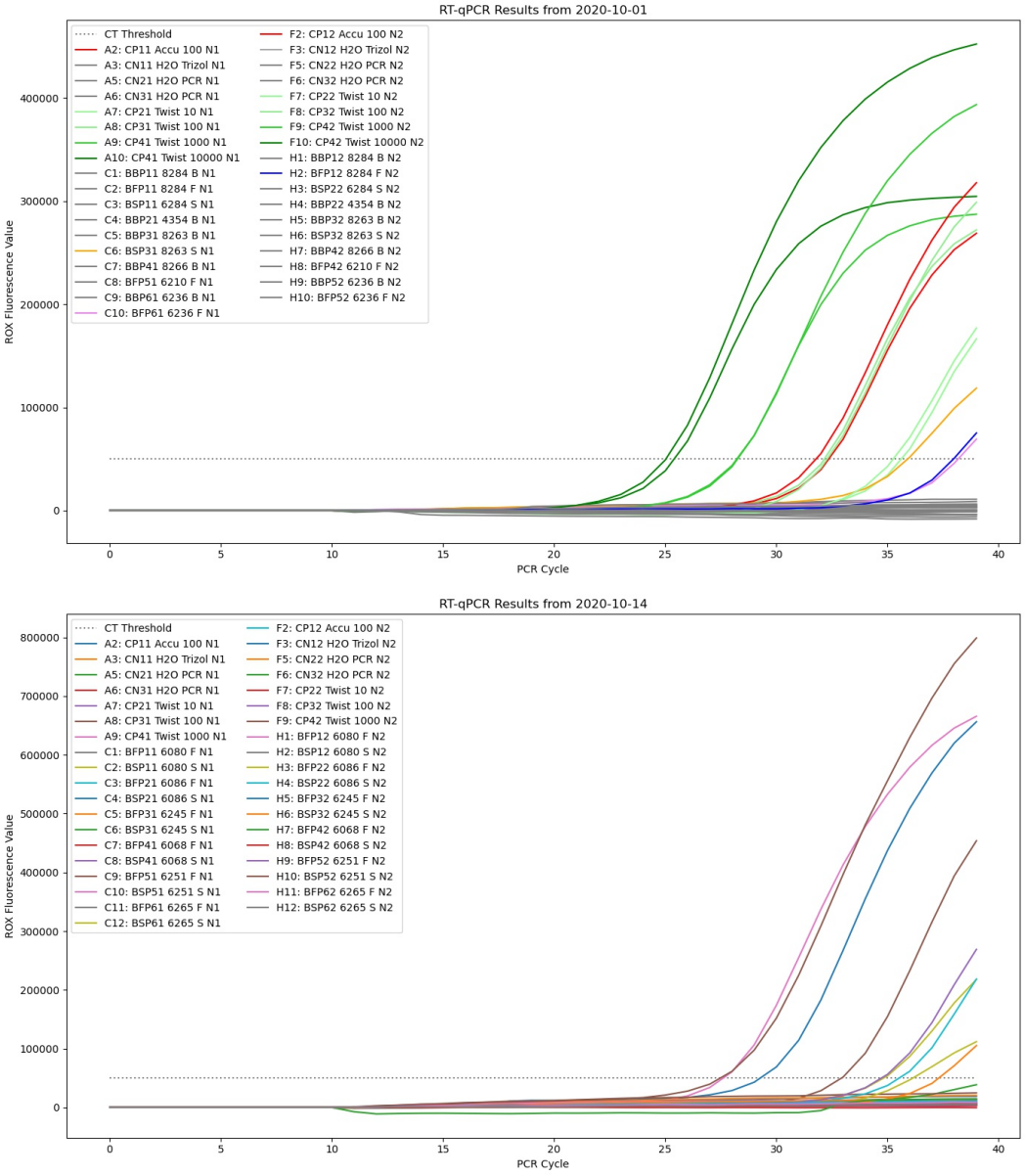


Figure D.7: Bus PCRs from collection dates Oct 2020.

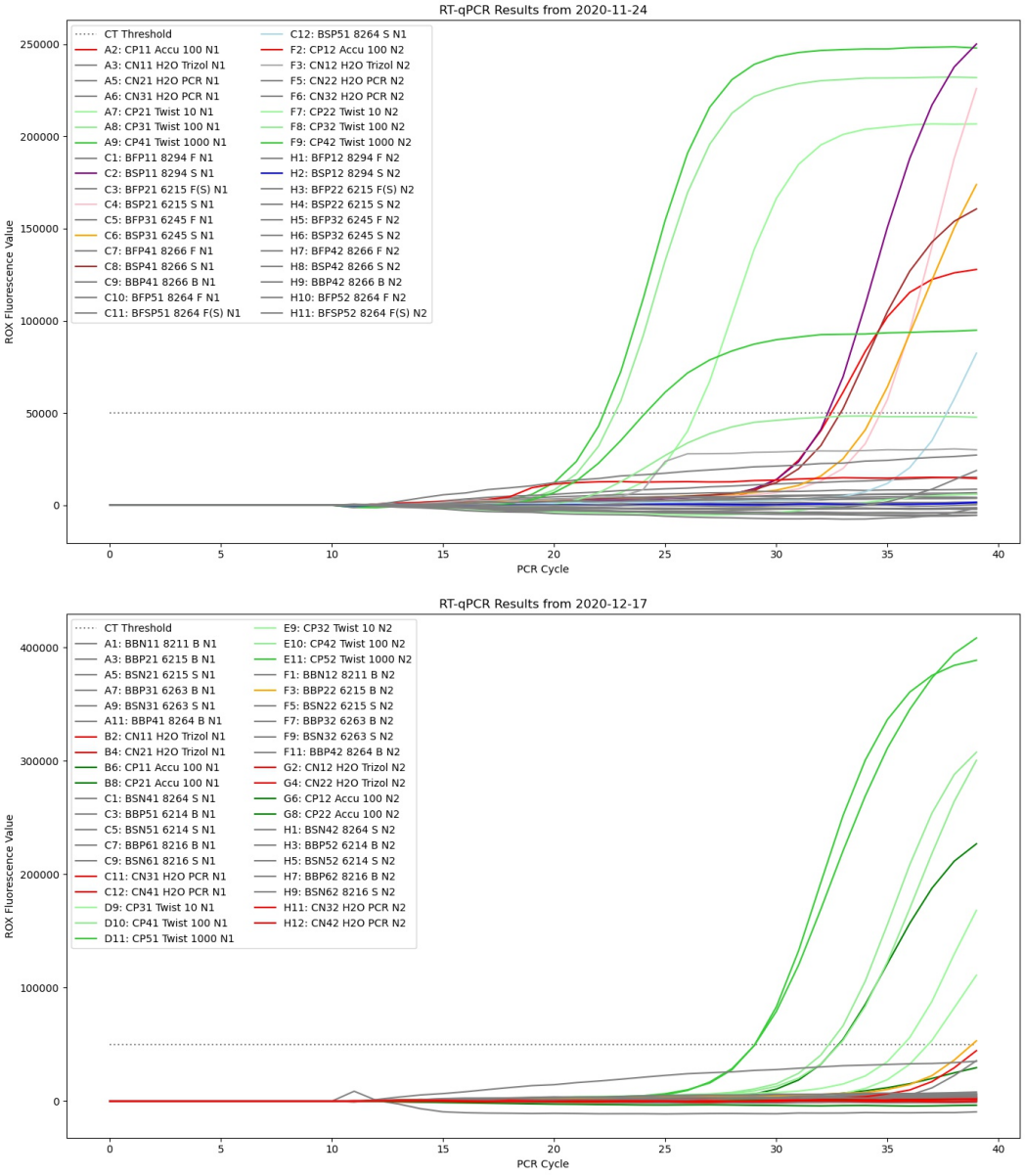


Figure D.8: Bus PCRs from collection dates Nov-Dec 2020.

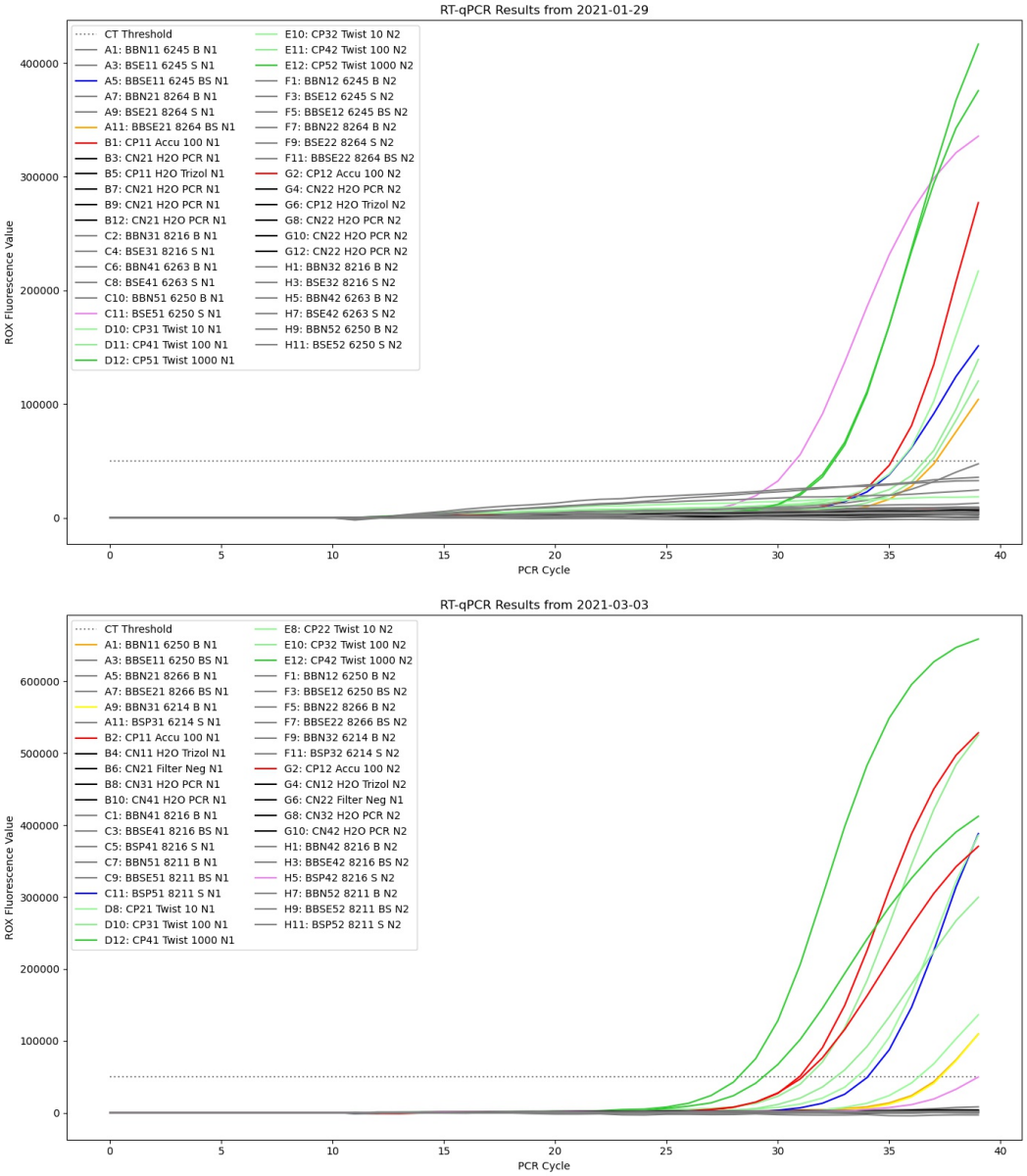


Figure D.9: Bus PCRs from collection dates in Jan-March 2021.

Table of results from bus testing

Date Collected	Bus Number	Filter Location	Filter Material	1 Result	2 Result	CT @ 50000
				N1	N1	All N1 gen
8/21/20	4552	Swab	Gram	No	No	
	4552	Front	Gram	Yes	No	39.67
	4552	Front	Foam	No	No	
	6072	Swab	Gram	No	No	
	6072	Back	Gram	No	No	
	6072	Front	Foam	No	No	
	8284	Back	Gram	No	No	
	8284	Back	Foam	No	No	
	6210	Swab	Gram	Yes	No	37.86
	6210	Front	Gram	No	No	
	6215	Swab	Gram	No	No	
	6228	Swab	Gram	No	Yes	39.39
pos % (all bus)	6	3	50.0%			
pos % (bus filters)	4	1	25.0%			
				N1	N2	N1 and N2
10/1/20	8284	Back	Paper	No	No	
	8284	Front	Paper	No	Yes	37.97
	8284	Swab	Paper	No	No	
	4354	Back	Paper	No	No	
	8263	Back	Paper	No	No	
	8263	Swab	Paper	Yes	No	35.91
	8266	Back	Paper	No	No	
	8266	Front	Paper	No	No	
	6210	Front	Paper	No	No	
	6236	Back	Paper	No	No	
	6236	Front	Paper	Yes	No	38.18
pos % (all bus)	6	3	50.0%			
pos % (bus filters)	6	2	33.3%			
				N1	N2	N1 and N2
10/14/20	6080	Front	Paper	No	No	
(1 day - AM)	6080	Swab	Paper	Yes	No	34.82
	6086	Front	Paper	No	No	
	6086	Swab	Paper	No	No	
	6245	Front	Paper	No	No	
	6245	Swab	Paper	No	Yes	
	6068	Front	Paper	No	No	
	6068	Swab	Paper	No	No	
	6251	Front	Paper	No	No	
	6251	Swab	Paper	No	No	
	6265	Front	Paper	No	No	
	6265	Swab	Paper	Yes	No	36.14
pos % (all bus)	6	3	50.0%			
pos % (bus filters)	6	0	0.0%			
				N1	N2	N1 and N2
11/24/20	8294	Front	Paper	No	No	
5 weeks	8294	Swab	Paper	Yes	No	32.31
	6215	Front (smal	Paper	No	No	
	6215	Swab	Paper	Yes	No	34.69
	6245	Front	Paper	No	No	
	6245	Swab	Paper	Yes	No	34.39
	8266	Front	Paper	No	No	
	8266	Swab	Paper	Yes	No	32.87
	8266	Back	Paper	No	No	
	8264	Front	Paper	No	No	
	8264	Front (smal	Paper	No	No	
	8264	Swab	Paper	Yes	N/A	37.66
pos % (all bus)	5	5	100.0%			
pos % (bus filters)	5	0	0.0%			

Date Collected	Bus Number	Filter Location	Filter Material	1 Result	2 Result	CT @ 50000
				N1	N2	
12/16/20	8211	Back	Nitro	No	No	
	6215	Back	Paper	No	No	
	6215	Swab	Nitro	No	No	
	6263	Back	Paper	No	No	
	6263	Swab	Nitro	No	No	
	8264	Back	Paper	No	No	
	8264	Swab	Nitro	No	No	
	6214	Back	Paper	No	No	
	6214	Swab	Nitro	No	No	
	8216	Back	Paper	No	No	
	8216	Swab	Nitro	No	No	
pos % (all bus)	6	0	0.0%			
pos % (bus filters)	6	0	0.0%			
pos % (bus- used)	2	0	0.0%			
				N1	N2	
1/29/21	6245	Back	Nitro	No	No	
	6245	Swab	EMax	No	No	
	6245	Back Swab	EMax	Yes	No	35.51
	8264	Back	Nitro	No	No	
	8264	Swab	EMax	No	No	
	8264	Back Swab	EMax	Yes	No	37.09
	8216	Back	Nitro	No	No	
	8216	Swab	EMax	No	No	
	6263	Back	Nitro	No	No	
	6263	Swab	EMax	No	No	
	6250	Back	Nitro	No	No	
	6250	Swab	EMax	Yes	No	32.42
pos % (all bus)	5	3	60.0%			
pos % (bus filters)	5	0	0.0%			
				N1	N2	
3/2/21	6250	Back	Nitro	Yes	No	37.22
	6250	Back Swab	EMax	No	No	
	8266	Back	Nitro	No	No	
	8266	Back Swab	EMax	No	No	
	6214	Back	Nitro	Yes	No	37.28
	6214	Swab	Paper	No	No	
	8216	Back	Nitro	No	No	
	8216	Back Swab	EMax	No	No	
	8216	Swab	Paper	No	No	
	8211	Back	Nitro	No	No	
	8211	Back Swab	Emax	No	No	
	8211	Swab	Paper	Yes	No	34.03
pos % (all bus)	5	3	60.0%			
pos % (bus filters)	5	2	40.0%			

Table D.3: Full set of results from buses tested at King County Metro. All buses, replicates, and samples are recorded, alongside whether or not they amplified in PCR testing and also their CT values.

D.1.3 Verification of Sample via native-PAGE

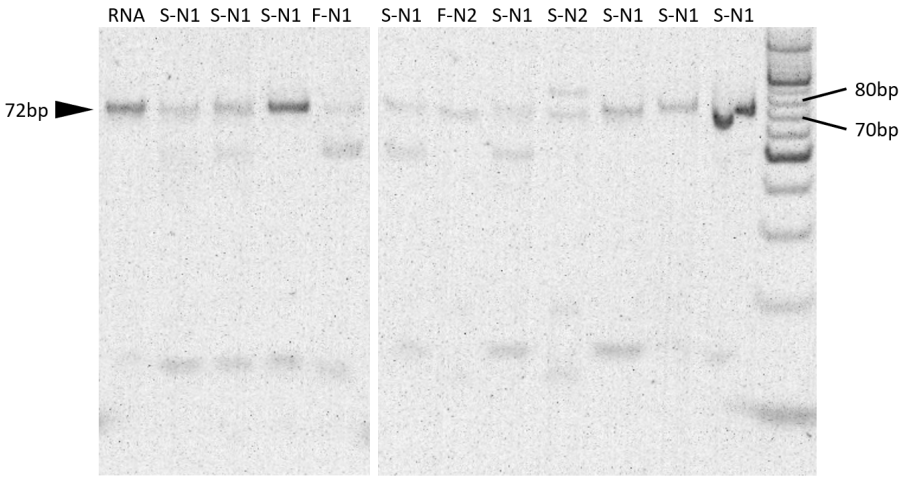


Figure D.10: Randomly selected positive PCR samples were analyzed using a 12 percent polyacrylamide gel. The arrowhead indicates the expected size of the PCR product from both N1 and N2 probes. Labels differentiate between air filter and swab samples, as well as an RNA positive control. O'RangeRuler 10bp DNA Ladder (SM1313, ThermoFisher)

D.1.4 Method comparison to Qiagen column-based RNA extraction kit

Column-based extraction kits (52906, Qiagen, Hilden, Germany) were used to extract environmental samples collected on March 2nd, 2021. Each 5 RNA product was assayed with TaqPath 1-step RT-qPCR N1 and N2 assays and compared to its counterpart replicates processed by TRIzol extraction (Figure D.4).

Date Collected	Bus ID	Filter Location	Filter Medium	Result 1	Result 2	CT @ 50000 ROX	Amplified?	N1-1	N1-2	N2-1	N2-2
3/2/2021	6250	Back	Nitro	Yes	No	37.22	No	N/A	N/A	N/A	N/A
	6250	Back Swab	EMax	No	No		Yes	36.7	36.6	38.27	35.44
	8266	Back	Nitro	No	No		No	N/A	N/A	N/A	N/A
	8266	Back Swab	EMax	No	No		Yes	N/A	37.76	N/A	39.42
	6214	Back	Nitro	Yes	No	37.28	No	N/A	N/A	N/A	N/A
	6214	Swab	Paper	No	No		No	N/A	N/A	N/A	N/A
	8216	Back	Nitro	No	No		No	N/A	N/A	N/A	N/A
	8216	Back Swab	EMax	No	No		Yes	N/A	36.96	N/A	N/A
	8216	Swab	Paper	No	No		No	N/A	N/A	N/A	N/A
	8211	Back	Nitro	No	No		No	N/A	N/A	N/A	N/A
	8211	Back Swab	Emax	No	No		Yes	N/A	N/A	N/A	38.47
	8211	Swab	Paper	Yes	No	34.03	No	N/A	N/A	N/A	N/A

Table D.4: A comparison was performed on 3/2/21 to understand how the results of our method compared to the silica-based extraction method. Based on these limited results, it appears that the sediments introduced by swab samples may interfere with precipitation and/or amplification of SARS-CoV-2 RNA using this method; whereas samples with very high CT (and thus low copy numbers) were not detected using Qiagen kits, indicating that this method may be more sensitive to low copy numbers when the sample is relatively clean, as is the case with passive air filter sampling.

Extraction Contaminants

The filter swab samples were the siltiest, returning from the bus depot with the darkest brown color. The differing ways that filter methods of TRIzol-based and column-based RNA extraction interacted withthese dirt particles likely caused difference in results for the two methods. TRIzol-based extraction resulted in positive results that column-based method missed, while the column-based method resulted in more positives in especially dirty samples.

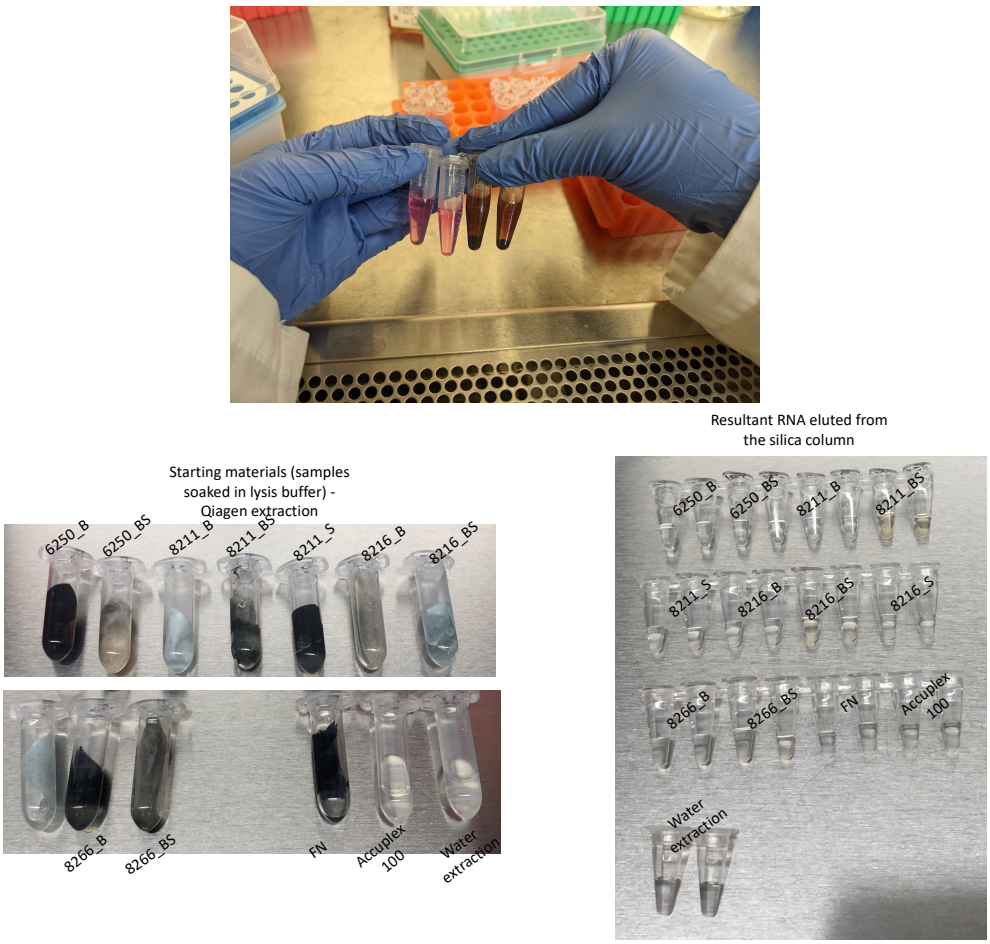


Figure D.11: Swab samples' appearance after our extraction method suggests that dirt and debris could interfere with PCR results, increasing number of false negatives for especially dirty samples. **Top** Resulting tubes after extraction by TRIzol-based extraction (our method) for 2 samples from air filters (two on left) and two samples from filter swabs (two on right). The filter swab samples were especially dirty before extraction, and some of that dirt ended up in the post-extracted material. **Bottom** Resulting tubes after extraction by column-based method (Qiagen). In column-based extraction, even the dirty bus swab samples remained clear after extraction, suggesting that they filtered different content than the TRIzol-based method.

Total Cost

	Total	Sampling Materials	Lab Materials
Total cost (for one test run):	**$341.66**	$58.21	$283.45
Cost / bus:	**$56.94**	$9.70	$47.24

Lab Materials

Category	Item	Qty / sample	Qty	Unit	Bulk qty	Price / bulk	Cost / Unit	Total cost / experiment
Consumable	Pipette tips	N/A	562	tips	960	$130	$0.14	$76.10
Consumable	Tubes	N/A	148	tubes	250	$55	$0.22	$32.56
Consumable	PCR plate	N/A	1	plate	25	$157	$6.28	$6.28
Consumable	PCR plate film	N/A	1	film	100	$43	$0.43	$0.43
Consumable	Gloves	N/A	10	gloves	1000	$268	$0.27	$2.68
Reagent	Lysis buffer	0.2	6.4	mL	500	$111.84	$0.22	$1.43
Reagent	Trizol	0.8	25.6	mL	200	$374	$1.87	$47.87
Reagent	Chloroform	0.12	3.84	mL	500	22.14	$0.04	$0.17
Reagent	Glycogen (RNA grade)	1	42	uL	200	$81.25	$0.41	$17.06
Reagent	200-proof Ethanol	2	84	mL	2000	$109	$0.05	$4.58
Reagent	Water (Mol Bio Grade)	N/A	10	mL	1000	$28.34	$0.03	$0.28
Reagent (PCR)	PCR mastermix	5	210	uL	500	$204	$0.41	$85.68
Reagent (PCR)	Primers	1	42	uL	500	$99	$0.20	$8.32
Total								$283.45

Sampling Materials

Category	Item	Qty	Unit	Bulk qty	Price for bulk	Cost / Unit	Total cost
Sample Material	Filters max	10	Filter	100	$114	$1.14	$11.40
Sample Material	Swabs	15	Swab	100	$190.72	$1.91	$28.61
Sample Material	Cleaning wipes	30	wipes	225	$11.97	$0.05	$1.60
Sample Material	Gloves	40	gloves	100	$18	$0.18	$7.20
Sample Material	Baggies	30	bags	300	$7.12	$0.02	$0.71
Sample Material	Secondary containers	1	Container	6	$52.19	$8.70	$8.70
Total							$58.21

Table D.5: Cost breakdown of lab materials. Note that these reagent and supplies costs were based on the small scaled purchased and did not include the labor cost.

D.1.5 Cost Analysis of Method

We have shown that passive filters with minimal installation requirements are capable of capturing extremely low quantities of viral particles. However, the concentration step with TRIzol-extraction makes our procedure more time-consuming and costly per-sample than other methods, such as column-purification and amplification of nucleic acids for Disease Control and Prevention [2020]. This paper details the results of 82 individual samples separately tested across 45 buses. In theory, by combining samples from the same bus in the concentration step, or by limiting the number of locations to install filters in each bus, the effort to assay multiple buses may be reduced in. Ultimately, while our method does not include cost-prohibitive or difficult-to-obtain materials, the cost of scalability must be considered alongside the increased sensitivity when comparing to alternative detection methods (Figure D.5).

D.1.6 Code Availability

Code for processing PCR results will be made public on Github at: https://github.com/jasonhof/pcr_utils.

Appendix E

Supplementary Information: Detecting Nucleic Acid Reactions Using Capacitive Touchscreen Sensors on Ubiquitous Smartphones

E.1 Challenges

Despite this feasibility data, the results of this project did not fully prove whether sensing nucleic acids with the touchscreen is possible. The technical challenge still exists, and we have identified three challenges to continue pursuing on this project.

Challenge 1

The first challenge concerns the fact that the data collected from the mutual capacitance sensors, like the smartphone touchscreen DK, is not clear yet. We are building algorithms and protocols to gather and process data differently to understand the nature of the differences between this sensor and other capacitive sensors, such as the VNA. However, there's still a chance that some aspect of the smartphone sensor (sensitivity, frequency, layout, etc) may not be amenable to sensing biological material.

Mitigation 1

If this is the case, our proposed mitigation would be to focus on developing an inexpensive USB-C attachment based on the freqencies and sensitivity revealed by the VNA. At the same time, we can characterize why the smartphone screen problem is difficult.

Challenge 2

The second risk is that the results we're seeing today with the VNA may not replicate well when other needed reactants are thrown into the mix. Our feasibility data thus far has involved creating the products of reactions and putting them on the sensor to simulate what an output would look like after we design the full system. This has shown good results, but the full system will include components, such as primers and

enzymes, that are not currently being measured, and there's a chance we may not be able to separate result from noise.

Mitigation 2

If this turns out to be the case, we can adjust the protocol to add extra experimental steps to clear the product reaction mix of those confounding materials, either by filtering or extracting the product into a new solution mix. This will complicate the reaction steps, reducing the wide availability of the results, but will still point us towards the steps needed to produce a more ubiquitous capacitance sensor for molecular reactions.

Challenge 3

The proposed targets of the proof-of-concept reaction may not entirely work as planned. This could be due to unforeseen folding of the nucleic acid targets, or safety restrictions on working with certain materials. We will follow our university's safety guidelines, and we have worked with viral control particles before, but there's still a chance that some confounding factor may pop up.

Mitigation 3

If this is the case, we can pivot to a different DNA target to show feasibility quickly, such as HIV drug resistance or synthetic target. There are many such diagnostics designed, and we can read the background of various options to choose a suitable target.

E.2 VNA Sensor Supplementary Data

Using a standard lab technique of an electrophoretic gel (QIAxcel), we verified that a reaction with our designed polymerase extension reaction is working as expected, as shown in Figure E.1.

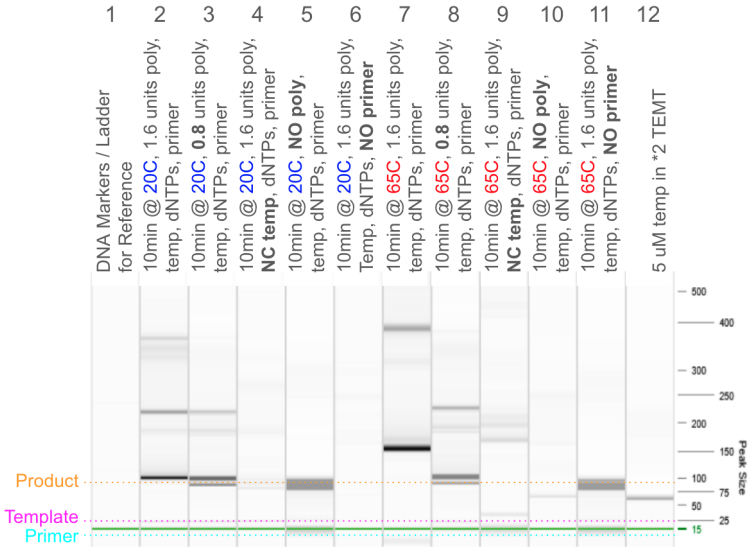


Figure E.1: An electrophoretic gel (QIAxcel) was run to verify the reaction that was expected to occur on the VNA.

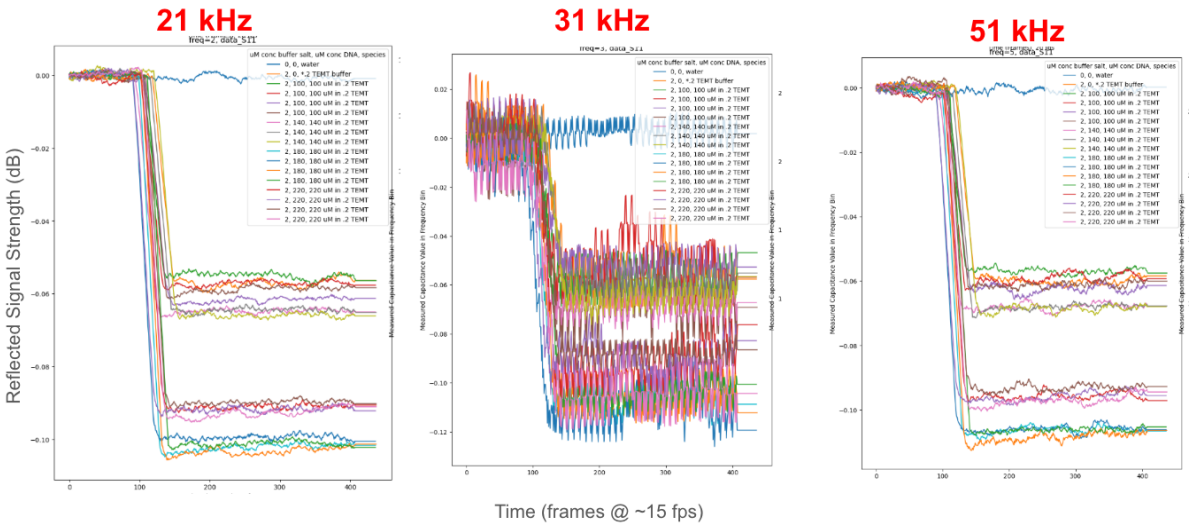


Figure E.2: The differences between solutions were more clear at certain frequencies than others, indicating that certain reactants within the solutions were more responsive to certain frequencies of input impulses.

One key aspect is the frequency of the impulse wave that is used for sensing in mutual capacitance touchscreens. DKs for mutual capacitance touchscreens, such as the Microchip unit that we have, operate in the frequency range of 20-200 kHz. Based on this range, we analyzed the response of the different contents of solutions at different frequencies of input impulse response, as shown in Figure E.2. This data would help inform the design of a capacitive touchscreen sensor to improve the signal that might be differentiable.

E.3 Results on Touchscreen

Experiments with mutual capacitance touchscreens have not been fruitful in delineating a difference between different concentrations of solution, using ssDNA or salt buffer. Thus far, we have tested on a Google Pixel 8 touchscreen and a touchscreen dev kit (DK) (Microchip). The results from the mutual capacitance dev kit are shown in Figure E.3. In the future, researchers could further tests on this to attempt to understand the limitations of the touchscreen as it relates to the VNA in terms of freqency, sensitive range, area, and processing software.

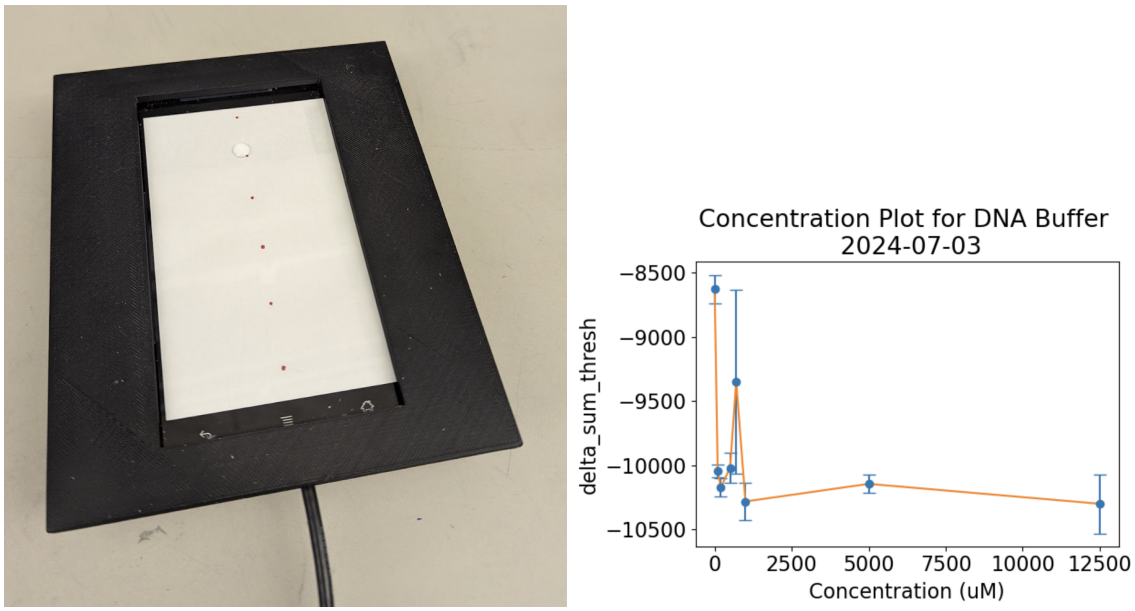


Figure E.3: Data gathered using a mutual capacitance dev kit (DK) (Microchip) shows that it may not be detecting capacitance in the same way as the VNA setup. Investigating this further through using the VNA may enable us to understand how to implement this on smartphones in the future.