

The proliferation of ubiquitous technologies like smartphones and smartwatches is increasing the number of sensors that travel with us at all times. Motion, audio, and visual sensor data can be analyzed to provide suggestions catered to individuals at their convenience. These sensors can also operate in the background and intervene when a serious situation arises. **My dissertation work focuses on the application of machine learning and computer vision on data from the smartphone's built-in sensors to create mobile apps that improve access to health screening and safety tools.** I work in the areas of ubiquitous computing and human-computer interaction, but my research is highly interdisciplinary; I collaborate with clinicians, designers, and informaticists. Because my work is so diverse, I have the ability to conduct clinical, survey, and prototype-focused user studies.

As my colleagues and I note in an upcoming article on the challenges of smartphone-based health sensing [6], uncontrolled environments and device heterogeneity add noise to sensor data, breaking assumptions that can be made in laboratory settings. Reliability is critical when sensors are being used to keep us healthy and safe at all times, so **my research approach often involves hardware and software compromises.** I work along the continuum of software-only and software-hardware solutions, making tradeoffs between scalability and reliability depending on the application. Although my dissertation work revolves around smartphones, my skills and methods generalize to any sensor package. There may be future sensor packages designed specifically for the problems I examine, but the findings from my research can be used to inform their designs.

The applications I have built are one-time tests that have their own uncertainty. Repeated measurements can increase statistical confidence and uncover trends, leading to better outcomes and possible clinical breakthroughs in the case of health screening. Nevertheless, Bayesian reasoning states that test results are not definitive, but rather act on probabilistic priors. I believe that user-specific priors can be estimated using ambient sensing and survey instruments. As such, **my long-term research vision entails combining longitudinal data collection with ambient sensing and survey instruments to produce a more accurate representation of the user's condition through Bayesian reasoning.** Below, I describe how my research enables repeated and in-situ measurements through existing technology as a first step towards my research vision.

PH.D. RESEARCH

Smartphone-based Health Sensing [2,3,7]

Currently, healthcare is most effective when providers see patients in their clinics. Ubiquitous smartphone-based health sensing can help patients be proactive in their own care, lessen clinical resource burdens, encourage adherence to treatment regimens outside of clinics, and provide relevant interventions at the most appropriate time. My broad skillset in computer vision and machine learning can apply to many health-related domains, but my work in this space focuses on symptoms that manifest in the eyes.

One attribute I have explored is the pupillary light reflex (PLR)—the manner in which a person's pupils react to a light stimulus. Emerging evidence has shown that an impaired PLR may be a useful biomarker of traumatic brain injuries (TBIs) like sports-related concussions. Professional sports teams can afford to have trained physicians on the sidelines deciding whether or not an athlete who has suffered a concussion can return to their game; however, teams with less funding must rely on a school nurse or volunteer who does not have the same expertise or tools at their disposal.

PupilScreen [3] is a smartphone app that achieves similar functionality as a clinical pupillometer—a device that uses an infrared camera to measure pupil size over time—at a fraction of the cost. PupilScreen constricts a person's pupils with the smartphone's flash and records the pupils' response through the camera. The video is processed using a fully convolutional neural network, producing per-pixel pupil segmentation for each frame. Those measurements are collated with timestamps to generate the same pupil diameter-over-time curves that a pupillometer would (Figure 1, bottom). Keeping the lighting stimulus consistent across scenarios is important for ensuring that the test is repeatable. Compromising scalability for the sake of consistency, I designed a 3D-printable box that blocks out ambient light and controls the distance between the smartphone and the person's eyes (Figure 1, top). We evaluated PupilScreen through a study on 42 individuals with a normal PLR to ensure that the algorithm could properly capture non-trivial pupil dynamics. PupilScreen was able to track pupil diameter with a median error of 0.30 mm, which is between the accuracy of a pupillometer (0.23 mm) and human observers (0.50 mm). We tested PupilScreen on six patients with TBI in a short follow-up study. When clinicians were given PupilScreen's



Figure 1. (top) The PupilScreen box controls lighting conditions for measuring the pupillary light reflex. **(bottom)** An annotated example of pupil constriction.

output alone, they were almost perfect when distinguishing responsive and unresponsive pupillary light reflexes. We are currently conducting a larger study through emergency departments in Thailand to collect more data from patients with suspected TBI.

Another symptom I have explored during my dissertation work is jaundice—yellow discoloration of the skin and eyes caused by a buildup of bilirubin in the bloodstream. Adults can become jaundiced as a result of complications of the liver or pancreas, such as alcoholism and pancreatic cancer. Bilirubin is only measured through a blood draw when a clinician suspects a relevant disease. The manifestation of jaundice can lead to a blood draw, but it is only obvious in severe states. Accessible jaundice detection can help undiagnosed individuals catch serious conditions early and enable diagnosed individuals monitor their own condition.

BiliScreen [2] analyzes the color of a person’s sclera through a smartphone photograph. The algorithm starts with sclera segmentation, which is non-trivial since jaundice changes the sclera’s color profile. BiliScreen segments the sclera using an automated variant of GrabCut [8]. The sclera’s color is transformed into a feature vector that is fed into a random forest regression model that estimates a person’s bilirubin level. Repeatable color measurement is of the utmost importance in BiliScreen, yet pictures can be taken in various lighting conditions. Here again, I developed two potential accessories that trade scalability for reliability. The first option was the same head-mounted box that was used for PupilScreen. The second option was a pair of paper glasses with colored squares around their rims that serve as known color references for calibration (Figure 2). We found that BiliScreen performed better with the box than with the glasses in a 70-person clinical trial, achieving a sensitivity of 89.7% and a specificity of 96.8%.



Figure 2. The BiliScreen glasses include squares along the rims for color calibration.

Public Safety [4,5]

I examine issues of public safety from the perspective of situational impairments [9]—contextual phenomena that impede a person’s ability to perform a task. Smartphones are often involved in situational impairments because we carry them with us at all times, but smartphones can also detect these impairments through their built-in sensors.

One situational impairment I investigated is inebriation. Portable breathalyzers are the status quo for measuring inebriation, but they are typically used after a drunk driver has been caught—rarely to prevent a drunk person from driving in the first place. Ubiquitous inebriation testing connected to a smart-locking mechanism could help prevent impaired people from getting behind the wheel. **Drunk user interfaces** [5] (DUIs) measure a combination of task-specific performance metrics and sensor-derived features to estimate a person’s blood alcohol level (BAL). For instance, performance on a typing task can be characterized by typing error rate and how off-center the user hits each intended key. Features within and across different DUIs are processed by a random forest regression to estimate BAL. It is important to note that DUIs assess side effects of inebriation rather than BAL directly; other factors like fatigue and learning can also affect a person’s performance on DUIs. Because of these confounds, we collected data from 14 participants in a 5-day longitudinal study. Participants used DUIs at different BALs, but the same time of day to control for fatigue and account for learning. We were able to estimate BAL with a mean absolute error of $0.005\% \pm 0.007\%$. As a point of comparison, the breathalyzer used as a ground truth had a self-reported tolerance of 0.005%.

Although smartphones are a great mechanism for addressing safety, they can sometimes be the cause of safety issues. For example, a pedestrian who is reading email on their smartphone must balance their focus between their smartphone and the physical space in front of them, imposing cognitive load that can lead to frustration or a pedestrian accident.

SwitchBack [4] alleviates cognitive burden by allowing the user to resume their smartphone task more efficiently. SwitchBack uses the front-facing camera to track the user’s attention and gaze. Although precise gaze-tracking is difficult in mobile contexts with small screens, larger gaze direction changes (saccades) are more robust to detect. With that in mind, we targeted a common activity where saccades are prominent: reading. As a person reads from line-to-line, their gaze occasionally jumps from right to left. SwitchBack counts how many lines the user has read, using information like typical reading speeds, rough vertical gaze position, and the text’s layout to adjust that estimate. When the user looks away and then returns their gaze back to the screen, SwitchBack highlights the last line of text the user

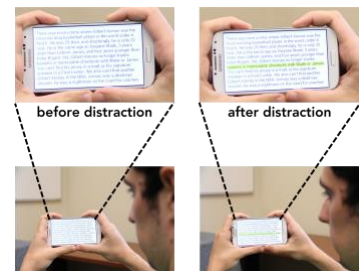


Figure 3. SwitchBack guides a person’s attention back to an article by counting the number of lines read.

read to guide their attention back to where they left off (Figure 3). To simulate a common pedestrian scenario, we conducted a 17-person study where participants were asked to read news articles and perform a secondary distraction task while walking on a treadmill. We found that SwitchBack improved average reading speed by 7.7%, demonstrating quicker transitions between contexts.

FUTURE RESEARCH PLANS

Moving from In-Clinic Testing to Self-Testing

I am part of a multi-organizational team working with the Bill and Melinda Gates Foundation on the first-ever over-the-counter rapid diagnostic test (RDT) for influenza. Our contribution to this goal is a mobile app that automatically interprets an RDT and checks for procedural adherence using computer vision. For instance, we plan on confirming whether a nasal swab is sufficiently pressed against a tube during sample transfer by applying optical flow on a video recording of the process.

The shift from in-clinic testing to self-testing for influenza will produce a structured platform for my research vision. We will be able to use ambient sensing underneath our RDT interpretation app to estimate Bayesian priors and probe other symptoms. The prior will account for the likelihood of transmission, which is a function of travel, social network, temperature, and humidity. The large network of users and GPS data will also enable us to estimate local prevalence.

Even if we carry out Bayesian reasoning on the user's behalf, it is up to them to determine how they should account for test results in their decision-making. For example, the user may ignore a positive test result because they either do not trust the test or are not confident in how they self-administered the test. I am developing a survey instrument that examines health-sensing tools through the lens of the Health Belief Model, a well-supported psychological model for health behavior change [1]. The instrument allows developers to examine how factors like sensitivity, specificity, and interface design affect users' decisions without deploying an actual app. I plan to further explore these questions further through surveys and interviews with various stakeholders.

Facilitating Health Sensing Innovation

In my experience, the development of novel health-sensing applications can be broken into three questions: (1) whether the phenomenon is relevant to a medical measurement (e.g., scleral jaundice to bilirubin level); (2) whether sensors can detect the phenomenon (e.g., smartphone camera); and (3) whether an algorithm can measure the phenomenon. I believe we can build annotation tools that clinical researchers without programming expertise can use to investigate the first two questions, turning algorithm development into teaching-by-demonstration. I plan on surfacing common features that clinical researchers want to capture. After aligning those features with the clinician's desired output, I will build a system that automatically selects a relevant algorithm and machine learning models to produce a proof-of-concept system. For instance, color selection on a heat map can lead to a custom image filter, and audio timestamps can produce training data for a recurrent neural network. I plan on engaging with clinical researchers to learn about the features they would like to capture, and I will use my vast experience in mobile health to identify the problems that are well-suited for teaching-by-demonstration.

Cross-Context Mobile Sensing

The diverse contexts in which people use smartphones add noise to sensor-derived datasets, often causing machine learning models to overfit. Discriminative adversarial networks (DANs) have the power to mitigate these confounds rather than requiring controlled environments or calibration as I have in the past. A DAN's loss function can be constructed to reward information about the label and penalize information about the data's source. I plan on exploring this approach to improve generalizability across contexts.

1. Nancy K Janz and Marshall H Becker. 1984. The Health Belief Model: A Decade Later. *Health Education Quarterly* 11, 1: 1–47. <https://doi.org/10.1177/109019818401100101>
2. **Alex Mariakakis**, Megan A Banks, Lauren Phillipi, Lei Yu, James Taylor, and Shwetak N Patel. 2017. BiliScreen: Smartphone-Based Scleral Jaundice Monitoring for Liver and Pancreatic Disorders. *Proc. IMWUT '17* 1, 2: 20:1–20:26. <https://doi.org/10.1145/3090085>
3. **Alex Mariakakis**, Jacob Baudin, Eric Whitmire, Vardhman Mehta, Megan A Banks, Anthony Law, Lynn Mcgrath, and Shwetak N Patel. 2017. PupilScreen: Using Smartphones to Assess Traumatic Brain Injury. In *Proc. IMWUT '17*, 81:1–81:27. <https://doi.org/10.1145/3131896>
4. **Alex Mariakakis**, Mayank Goel, Md Tanvir Islam Aumi, Shwetak N Patel, and Jacob O Wobbrock. 2015. SwitchBack: Using Focus and Saccade Tracking to Guide Users' Attention for Mobile Task Resumption. In *Proc. CHI '15*, 2953–2962. <https://doi.org/10.1145/2702123.2702539>
5. **Alex Mariakakis**, Sayna Parsi, Shwetak N Patel, and Jacob O Wobbrock. 2018. Drunk User Interfaces: Determining Blood Alcohol Level through Everyday Smartphone Tasks. In *Proc. CHI '18*, 1–13. <https://doi.org/10.1145/3173574.3173808>
6. **Alex Mariakakis**, Edward Wang, Mayank Goel, and Shwetak Patel. 2019. Challenges in Realizing Smartphone-based Health Sensing. *To appear in IEEE Pervasive Computing*, 1–10.
7. **Alex Mariakakis**, Edward Wang, Shwetak N Patel, and Joanne C Wen. 2016. A Smartphone-based System for Assessing Intraocular Pressure. In *Proc. EMBC '16*.
8. Carsten Rother, Vladimir Kolmogorov, and Andrew Blake. 2004. Grabcut: Interactive foreground extraction using iterated graph cuts. *ACM Transactions on Graphics (TOG)* '04 23, 3: 309–314. Retrieved May 21, 2016 from <http://dl.acm.org/citation.cfm?id=1015720>
9. A Sears, M Lin, J Jacko, and Y Xiao. 2003. When computers fade: Pervasive computing and situationally-induced impairments and disabilities. In *Proc. HCI International '03*, 1298–1302.