UNIVERSITÄT ZU LÜBECK

TECHNISCHE HOCHSCHULE LÜBECK

Master thesis

Topic: Assessing data quality and integrity of optical measurements in a portable blood analyzer using machine learning.

Summary:
Mobox is an advanced blood analyzer developed for critical care and emergency medicine, offering robust and reliable results powered by an AI-based optical sensor. It is compact and lightweight, enabling rapid point-of-care diagnostics and providing medical professionals with critical information for life-saving decisions.
The device utilizes a cost-effective and disposable microfluidic chip for blood analysis.
Blood samples are introduced into the chip, which is inserted into the device through a front-facing slot. By integrating a miniaturized spectrometer and an AI-powered optical sensor driven by convolutional neural networks (CNNs), mobox achieves portability and practicality by eliminating the need for thermal stabilization, calibration lamps, and additional optical components.
Mobox utilizes transmittance spectra of whole blood samples as inputs for its CNN based models. However, the lack of a system to validate the quality of these spectra poses a significant challenge to ensure accurate and reliable performance. As an IVD Class C device under the EU regulation on in vitro diagnostic medical devices (IVDR), mobox must comply with stringent regulatory standards to ensure both performance and safety. This classification necessitates robust measures to mitigate any potential risks, including those associated with unvalidated spectra.
The measurement of blood parameters, generally known as CO-oximetry, plays a crucial role in assessing the oxygenation and composition of blood. CO-oximetry involves the measurement of blood parameters such as total hemoglobin (tHb), oxygenated hemoglobin (O2Hb), deoxygenated hemoglobin (HHb), carboxyhemoglobin (COHb), and methemoglobin (MetHb). This technique is particularly susceptible to inaccuracies caused by air bubbles in the blood sample or on the test chip, as well as hemolyzed blood. These errors can significantly impact CO-oximeter readings, potentially leading to incorrect treatment decisions.
To address these issues, this thesis focuses on evaluating and enhancing the reliability of the data used in the models, ensuring high-quality inputs for robust and consistent performance. By identifying and discarding compromised samples, such as those affected by air bubbles or hemolysis, before they contribute to inaccurate results, this approach enhances the reliability of the device and minimizes the risk of delivering incorrect readings that could impact patient care. The practical component of this thesis involved acquiring a comprehensive dataset through experimental adjustment of blood samples and test chips.
Therefore, typical errors in real-world scenarios, such as hemolysis and air bubbles, were deliberately introduced. A newly designed flow cell enabled the controlled generation of air

bubbles, while the path length of the flow cell was evaluated to confirm the consistency of the manufacturing process.
Various hemolysis techniques were also assessed, including chemical methods using 0.01% Triton X-100 and Triton Lysis Buffer, temperature fluctuations, blood freezing, and mechanical stress by repeated centrifugation. Among these methods, freezing proved to be the most effective for inducing controlled hemolysis, delivering consistent and reproducible results.
Using a dataset of 158 blood samples with diverse hematocrit (Hct) levels, spanning from 0.1 to 0.6, and oxygenation levels, ranging between 40% and 100%, a CNN-based machine learning model was developed and evaluated by 5-fold cross-validation. The model demonstrated strong performance in distinguishing good blood samples from those affected by errors, achieving a mean precision of 0.94, recall of 0.95, an F1-score of 0.95, and an overall accuracy of 0.93. These results highlight the model's ability to reliably classify samples, minimizing false positives and false negatives.
Future efforts will focus on refining measurement methodologies, expanding the dataset, and validating the model's scalability to ensure its effectiveness in diverse clinical settings.

Author: Maria Angela Qato
Supervisor: Prof. Dr.-Ing. Stefan Müller, M. Sc. Benjamin Kern
Date of submission: 8/01/2024