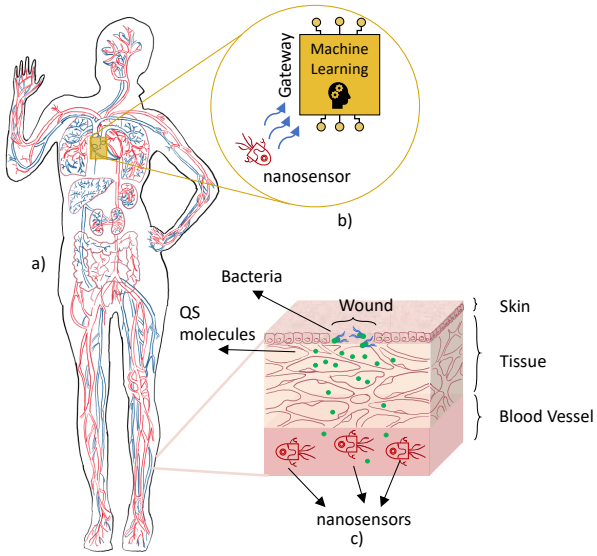
Datasheet for the BloodVoyagerS Dataset

1 Motivation

- This Dataset is created to enable researchers to explore travelling paths of nanoparticles in the human circulatory system (HCS).
- The dataset is created by the Telecommunication Network Group, Michigan State University, and Lakeside Labs research groups.
- This research was supported in part by the project MAMOKO funded by the German Federal Ministry of Education and Research (BMBF) under grant number 16KIS0917 as well as the project NaBoCom funded by the German Research Foundation (DFG) under grant number DR 639/21-1.



2 Composition

- The instances are travelling time and similarity of concentration levels of flowing nanoparticles. See an example of collected samples for the center body in Fig. 1.
- The travelling time refers to the time elapsed for a nanoparticle when travelling through these closed-loops starting and ending at the left Heart.
- Similarity of the concentration level refers to the correspondence of the concentration level, as recorded by the nanoparticle, and the expected one for the given circuit. It is computed as a cross-correlation.
- Data is split for the different 14 circuits labeled as the capillaries, e.g., "Head", "Shoulders", etc. In addition, all sequences are labeled as *delay_* when referring to time, and as *con_cross* when referring to the concentration level. For example, samples related to the travelling time through the Head are labeled as "delay_Head", while the cross concentration metric is labeled as "cross_conc_Head".
- The total of instances are the ones produced by 1000 nanobots in $300\,\mathrm{s}$. Depending on the travelled circuit it will be the total of those. For example, through the Head, there are $1\,712\,000$ samples for the travelling time and similarity of concentration level.
- This dataset provides metrics according to a person with $1.72\,\mathrm{m}$ height and $69\,\mathrm{kg}$ weight [1]. Further details related to the dimension of vessels are given in [2]

3 Collection

- Samples are collected from simulated nanoparticles according to the blood flow in the human circulatory system (HCS).
- The simulation is provided by the BloodVoyagerS (BVS) framework [1] where samples are identified according to 14 closed-loops addressing the main capillaries: Head, Shoulders, Upper Arms, Elbows,

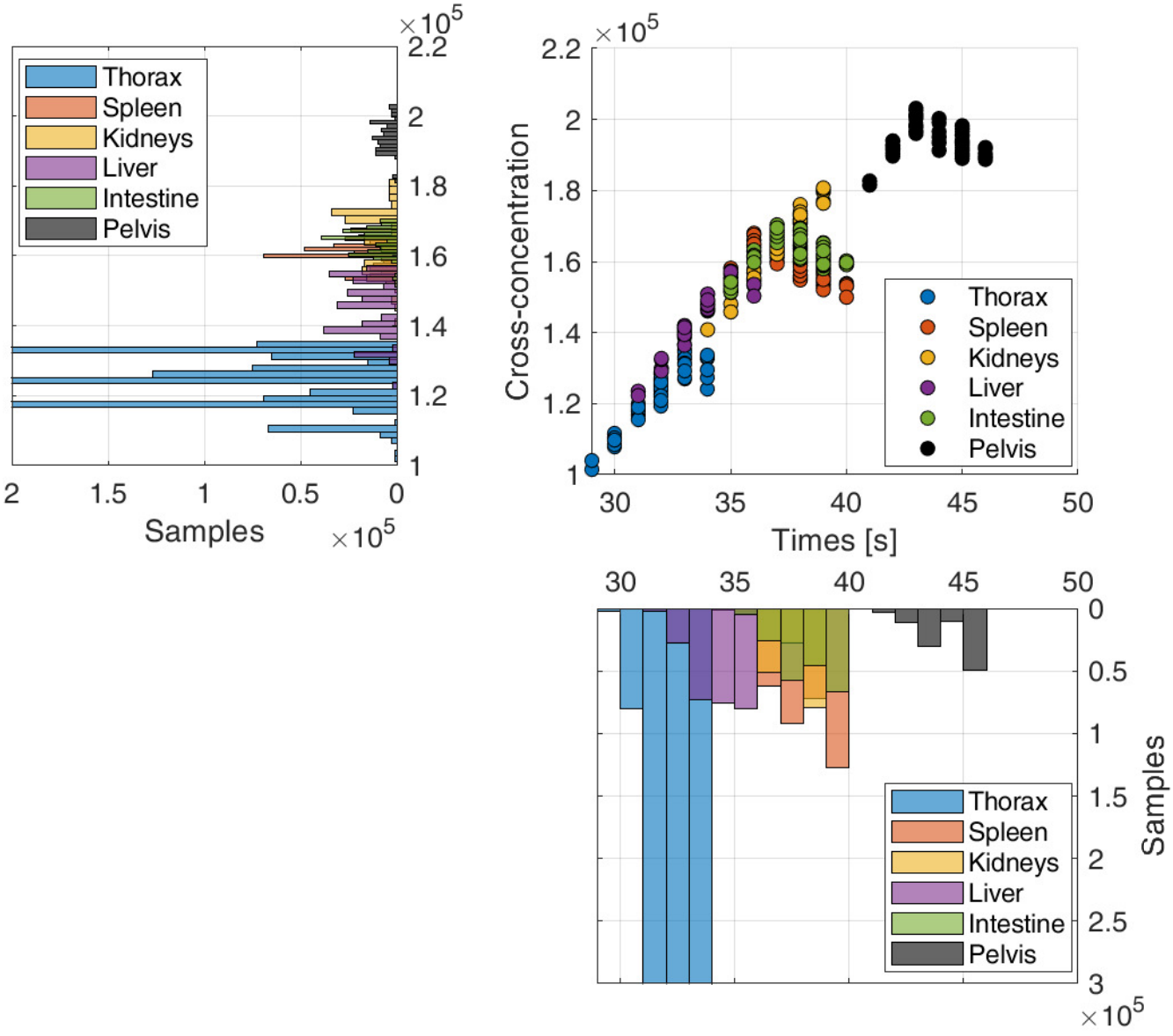


Figure 1: Representation of the collected data.

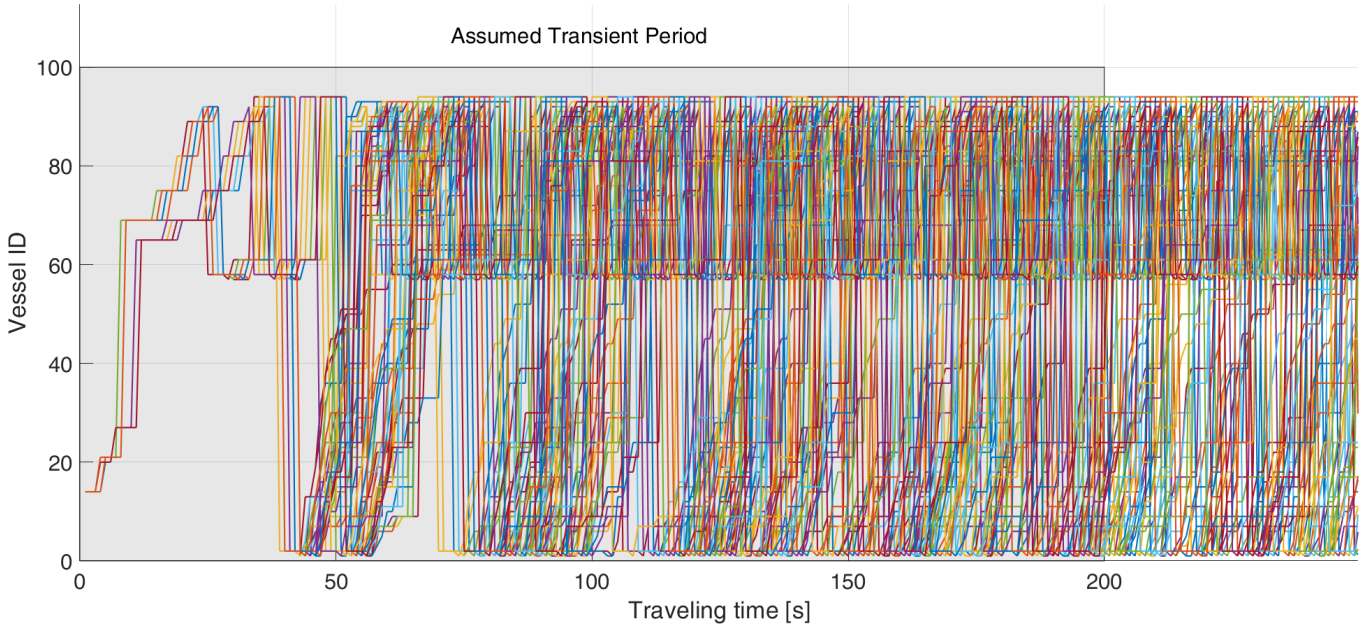


Figure 2: Traveling vessel IDs with time followed by nanosensors (100 here represented).

Hands, Thorax, Spleen, Kidneys, Liver, Intestine, Pelvis, Hips, Knees, and Feet. The flow of nanoparticles may be visualized in this Lübeck Universität Link.

- The data is inferred from the vessel ID provided by the simulator regarding each flowing nanoparticle. Where the vessel ID is detailed at this table (first column) GitHub Link.

4 Preprocessing/Cleaning/Labeling

- Data here provided is collected after removing $200\,\mathrm{s}$ transient period (cf. Fig. 2).
- During the transient period, the distribution of nanosensors is still not stable in the body, thereby introducing some bias regarding its concentration with time.
- The length of the transient period has been identified visually after plotting the concentration level per circuit and identifying the initial time interval where the nanoparticles were flowing together from the injection point.
- The removal of samples is directly controlled by the variable "t_initial" in "code_file.mlx" (Matlab code) at GitHub Link.

5 Uses

- This dataset has been used to produce the results presented in the paper [3]. In this paper, we developed a model to enhance localization and detection capabilities of abnormalities by flowing nanosensors. However, this work exploited only the time-domain data, i.e., the travelling time of nanosensors along with the HCS.
- This dataset may provide the concentration level on particular vessel segments with potential use in the study of the effectiveness of treatment and detection of abnormalities in the human body.
- This dataset is limited to the physiological parameters provided in [2], that is a person with $1.72\,\mathrm{m}$ height and $69\,\mathrm{kg}$ weight.

6 Distribution

- The dataset is publicly available at this GitHub Link.

References

[1] R. Geyer, M. Stelzner, F. Büther, and S. Ebers, "BloodVoyagerS: Simulation of the Work Environment of Medical Nanobots," in *5th ACM International Conference on Nanoscale Computing and Communication (NANOCOM 2018)*, Reykjavík, Iceland: ACM, Sep. 2018, 5:1–5:6.

[2] R. Geyer, "Konzeption und erste Umsetzung eines Körpersimulators zur realistischenModellierungder Arbeitsumgebung von medizinischen Nanobots," Master's Thesis, Universität zu Lübeck, Lübeck, Germany, Mar. 2017.

[3] J. T. Gómez, A. Kuestner, K. Pitke, J. Simonjan, B. D. Unluturk, and F. Dressler, "A Machine Learning Approach for Abnormality Detection in Blood Vessels via Mobile Nanosensors," in *Proceedings of the 19th ACM Conference on Embedded Networked Sensor Systems*, 2021, pp. 596–602.