Machine Learning-Driven Localization of Infection Sources in the Human Cardiovascular System

Saswati Pal, Jorge Torres Gómez, *Senior Member, IEEE*, Lisa Y. Debus, *Graduate Student Member, IEEE*, Regine Wendt, Florian-Lennert Lau, Cyrus Khandanpour, Malte Sieren, Stefan Fischer, and Falko Dressler, *Fellow, IEEE*

***Abstract*—In vivo localization of infection sources is essential for effective diagnosis and targeted disease treatment. In this work, we leverage machine learning models to associate the temporal dynamics of biomarkers detected at static gateway positions with different infection source locations. In particular, we introduce a simulation that models infection sources, the release of biomarkers, and their decay as they flow through the bloodstream. From this, we extract time-series biomarker data with varying decay rates to capture temporal patterns from different infection sources at specific gateway positions. We then train a stacked ensemble model using LightGBM and BernoulliNB to analyze biomarker time-series data for classification. Our results reveal that higher biomarker degradation rates significantly reduce the localization accuracy by limiting the biomarker signal detected at the gateways. A fivefold increase in decay rate lowers the mean cross-validation accuracy from $\sim$92 % to $\sim$66 %. This effect is more pronounced for infection sources located farther from the gateways, e.g., the kidneys. Due to the longer distance, more biomarkers degrade before reaching the wrist-located gateways, leading to a substantial decline in classification performance.**

***Index Terms*—IoBNT, Localization, Machine Learning, Biomarker Decay, Human Circulatory System**

I. Introduction

In recent years, the emerging idea of the Internet of Bio-Nano-Things (IoBNT) has offered a vision of enhanced healthcare, with potential advancements in early disease detection, monitoring, and treatment [1]. The concept could facilitate the monitoring of biochemical signals released from infection sources, enabling their localization. Infection sources, such as tumor cells, release specific biochemical signals called biomarkers into the bloodstream, indicating the presence of an infection [2]. These biomarkers related to cancer can be categorized as either extracellular or intracellular. Extracellular biomarkers include circulating tumor cells, circulating tumor Deoxyribonucleic Acid (DNA), protein markers, and exosomes. Intracellular biomarkers include telomerase and microRNAs. The biomarkers disperse from the infection source and flow through the circulatory system. In the context of molecular communication (MC), they act as information carriers released by biological transmitters, e.g., infection sites, and are transported through the bloodstream (channel) to passive receivers (gateways). Using the temporal profiles collected at these receivers, the identification of the biomarker source is interpreted as an MC-based localization task [3]. A representative example of technological advances for blood sample collectors is researched in [4]. Such receiver devices can provide temporal profiles of circulating biomarkers using biofunctionalized field-effect transistors. However, temporal profiles from a biomarker signal do not inherently reveal the specific organ of origin. To address this limitation, it is essential to characterize and map biomarker release patterns across different organs.

To generate organ-specific biomarker profiles, we employ the BloodVoyagerS (BVS) framework [5] that models blood flow dynamics and the mobility of nanoscale particles within the human circulatory system (HCS). BVS provides a schematic, graph-based representation of the human vascular network. It allows configurable placement of transmitters and receivers along the vasculature. Based on this framework, we configure the behavior of static infection sources and simulate the release and transport of biomarkers within the HCS. Biomarkers undergo exponential degradation in the human body [6], and we incorporate this dynamic into the simulator, following our previous work [7]. Next, we employ a data-driven machine learning (ML)-based classification approach to address the complexity of localizing infection sources in the vascular network. In such networks, traditional analytical methods are limited in their ability to reverse-map observed biomarker concentrations to their sources. By learning temporal patterns from biomarker data observed at gateways, ML enables the classification of the infection source.

Our main contributions can be summarized as follows:

- We extended our simulation tool to support infection sources, the release of biomarkers from these sources, and their mobility along the vessels,
- We implement and train our ML models to classify the location of infection sources based on temporal biomarker dynamics,
- We evaluate the effect of different biomarker decay rates on the accuracy of ML models.

Saswati Pal, Jorge Torres Gómez, Lisa Y. Debus, and Falko Dressler are with the School for Electrical Engineering and Computer Science, TU Berlin, Berlin, Germany, email: {pal, torres-gomez, debus, dressler}@ccs-labs.org.

Regine Wendt, Florian-Lennert Lau, and Stefan Fischer are with the Institute of Telematics, University of Lübeck, Germany, e-mail: {regine.wendt, f.lau, stefan.fischer}@uni-luebeck.de.

Cyrus Khandanpour and Malte Sieren are with the University Cancer Center Schleswig-Holstein, Lübeck, Germany, e-mail: {Cyrus.Khandanpour, Malte.Sieren}@uksh.de.

Reported research was supported in part by the projects NaBoCom funded by the German Research Foundation (DFG) under grant numbers DR 639/21-2 and FI 605/21-2 and IoBNT funded by the German Federal Ministry of Research, Technology and Space (BMFTR) under grant 16KIS1986K.

II. Related Work

For the localization of disease sites in the HCS, Simonjan et al. [8] proposed a distance tracking approach that employs conceptual bionanosensors equipped with an inertial measurement unit (IMU) composed of nanoscale accelerometers and gyroscopes. However, this method is limited by the data constraints of bionanosensors and the potential for inaccurate IMU readings due to the complexity of blood flow patterns. An alternative approach leverages the concept that specific body regions can be identified by their environmental properties as a fingerprint [9]. Thus, nanobots are enabled determine their position through local pattern recognition. The nanobots can ascertain their position within the human body and relay this information to external systems. However, this method relies on unique combinations of protein-coding genes for each tissue as the basis for the fingerprints. Identifying such combinations can be challenging as tissues exhibit similar gene expression patterns, potentially reducing localization accuracy.

To localize abnormalities within the HCS, our previous work proposed a framework combining unsupervised and supervised ML models [10]. We used a Markov model to evaluate the nanosensor distribution and movement in the HCS. The ML models were trained on data from BVS to improve the accuracy of abnormality localization. Furthermore, Pascual et al. [11] developed an analytical model to represent raw data for flow-guided localization, where the data is influenced by the communication and energy constraints of nanodevices. The model was validated against a simulator, showing similarity in results across different scenarios and performance metrics.

The localization of infection sources within the body, a complex physiological environment, is challenging to formulate through analytical approaches. Biomarkers may be transported through different blood vessels, each following a distinct path. As a result, predicting their exact location at any given time is difficult. Circumventing the challenges of predicting the source location, we resort to ML-based methods due to a variety of inherent advantages. In contrast to classical analytic models, trained ML modules learn organ-specific signal signatures automatically, allowing them to discriminate among different source organs; scale naturally across heterogeneous physiological pathways; and generalize far better to unseen operating conditions – e.g., new flow rates, pathway geometries, or noise levels.

III. Modeling and Implementation

We first provide details on the system model and the implementation of biomarker dynamics in BVS. We then give detailed insights into the selected ML classification frameworks.

A. System Model

In our system model, biomarkers are continuously released by a static infection source located in one of three regions: the head, thorax, or kidneys. These biomarkers travel through the HCS via blood flow to the detection points, referred to as gateways. In this work, we consider two such gateways located at the blood vessels beneath the left and right wrist. Of course, the model in general can support other configurations of infection locations or gateway positions. Gateways are modeled as transparent, static observers that passively count biomarkers flowing through designated vessel segments, specified by their vessel ID in the simulator. The biomarker counts collected at the gateways are used for subsequent analysis and source localization.

For each simulation, a single infection source is considered at a time, with biomarkers released from one location (i.e., head, thorax, or kidney) and observed at both gateways. At the end of each simulation, biomarker time-series data are recorded at each gateway, enabling analysis of their temporal distribution from a given infection source. The generated data is analyzed to identify patterns in biomarker counts, facilitating the classification and localization of the infection source within the body.

B. Implementing Biomarker Dynamics in BVS

The BVS framework [5] integrates a human body model into ns-3, modeling the movement of nanobots within a simplified cardiovascular system. The human model in BVS bases its blood vessel lengths on a $1.72\,\mathrm{m}$, $69\,\mathrm{kg}$ woman, covering 94 major blood vessels connecting organs to realistically simulate biomarker mobility within the HCS. We extended BVS to integrate infection sources, their release of biomarkers, and the decay of emitted biomarkers as they move along the HCS.

The simulation environment is implemented as a C++ extension of the ns-3 and operates in discrete time, with a fixed step size of $10\,\mathrm{ms}$ and a total simulation duration of $500\,\mathrm{s}$. Key configurable parameters include the total simulation duration; the time interval between successive simulation steps, which determines the temporal resolution; the number of infection sources; the specific vessel segment for infection sources; and the number of biomarkers released from each source.

BVS simulates the mobility of biomarkers based on laminar flow, with vessel-specific velocity profiles derived from physiological parameters. The emission of biomarkers is modeled as a continuous point source. The vessel walls are considered rigid and impermeable boundaries, restricting biomarker movement strictly to the vessel lumen, without modeling interactions such as permeability, adhesion, or penetration. In large and mid-sized blood vessels, such as arteries and veins, the flow regime is predominantly laminar due to relatively low Reynolds numbers under normal conditions [12]. This approach simplifies simulation of biomarker transport while remaining consistent with observed hemodynamics. To focus on large-scale transport and source localization, we excluded diffusion, vessel wall interactions, temperature effects, and biochemical reactions. Gateways are modeled as passive, transparent observers, counting biomarkers passing through specific vessel segments. Reception noise and signal interference were similarly omitted to isolate core signal dynamics. Table I summarizes the key parameters relevant to molecular communication.

The simulator models a single, stationary infection source per simulation run, emitting one type of biomarker from a predefined location. At the start of the simulation, the IDs of the infection source blood vessels are provided as input by the user. Each ID corresponds to a specific blood vessel associated with

TABLE I: Simulation parameters used in BVS

Parameter	Description
Blood velocity	Modeled as laminar flow; values vary by vessel type (e.g., aorta $\approx$ 20 cm/s, veins $\approx$ 2–4 cm/s, capillaries $\approx$ 0.02 cm/s).
Infection source type	Modeled as a point source, continuously releasing biomarkers into the bloodstream.
Receiver type	Gateways are modeled as transparent observers, passively counting biomarkers that pass through their location.
Channel model	BVS simulates laminar, stream-based flow within blood vessels, modeling biomarker transport as advection rather than diffusion. Biomarkers are carried passively along vessel-specific velocity profiles, consistent with convective transport observed in the HCS.
Diffusion & Temperature	Since particle transport is governed by laminar blood flow rather than Brownian motion, BVS does not account for diffusion or temperature effects.
Péclet number	The flow-dominant transport dynamics of the simulation implies a high Péclet number.

an organ or region in the human body. A corresponding method is invoked to model the release of biomarkers from the infection source within the specified blood vessel over a set duration. For the release, biomarkers are placed either individually or in groups at a fixed location within a blood vessel. Each biomarker is then randomly assigned to one of the parallel internal streams of the vessel to simulate a spatially distributed but localized release within that region. The simulation triggers an initial release of the biomarkers and schedules subsequent releases at an interval of one second throughout the simulation. This approach effectively simulates an infection source in the vessel, continuously releasing biomarkers for the specified simulation duration.

We implement the decay of biomarkers by modeling their gradual reduction over time in the simulation following an exponential decay process [7]. At each time step, the simulator computes the number of biomarkers to decay based on the current total number of circulating biomarkers. As a result, the decay scales proportionally with the biomarker count at any given time. The number of biomarkers removed per time step is given by, $N(t) = e^{-r_{\text{decay}}\Delta t}N(t-\Delta t)$, where r_{decay} is the decay rate and $N(t)$ is the current biomarker count. The simulator then randomly selects blood vessels from the global map and further selects random streams within those vessels. In the bloodstream, biomarkers are randomly removed from the simulation. The decay process is rescheduled to run every time step, allowing continuous biomarker decay over time. This approach effectively models the natural degradation dynamics in a biological system, ensuring realistic simulation behavior.

C. Classification Framework

For training and evaluation, we use biomarker data collected from multiple simulation runs. Each time series is labeled with the ID of the blood vessel corresponding to the infection source, following the mapping in [5]. The resulting dataset consists of 240 samples, each representing time-series biomarker data over $500\,\text{s}$ of simulation time. We analyze the time-series data to capture temporal dependencies arising from physiological transport. The dataset is then split into training and testing sets using an $\frac{80}{20}$ stratified split, to ensure that all infection source classes are proportionally represented in both subsets. We employ a stacking ensemble model with LightGBM [13] and Bernoulli Naive Bayes (BernoulliNB) [14] as base classifiers. LightGBM, a gradient boosting framework, handles complex feature interactions in structured data, while BernoulliNB offers a probabilistic perspective that performs well on sparse features. Together, these models are well-suited for capturing nonlinear patterns in biomarker time-series data that arise from complex vascular topology, branching dynamics, degradation, or failure of the signal to reach gateways. The LightGBM classifier is initialized with predefined parameters to balance model complexity and generalization. To further refine the model, we perform randomized hyperparameter search using 5-fold stratified cross-validation. The best parameter configuration is used to train the final classifier.

TABLE II: Summary of ML parameters

Component	Parameter	Value / Description
Base	boosting_type	dart, gbdt
Classifiers	learning_rate	0.005, 0.01
LightGBM /	n_estimators	500, 1000
BernoulliNB	max_depth	3, 5
	reg_alpha, reg_lambda	1.0, 5.0
	Hyperp. Tuning	RandomizedSearchCV (5-fold)
Meta	Model	Random Forest
Classifier	n_estimators	300
	max_depth	5
Ensemble	Method	StackingClassifier
Strategy	Cross-validation	StratifiedKFold (n_splits=5)
Data Split	Train-Test Ratio	80% train / 20% test
	Stratification	Yes (to maintain class balance)

We adopt a stacking ensemble where the outputs of LightGBM and BernoulliNB are combined using a Random Forest meta-classifier [15]. Random Forest is chosen for its robustness in capturing nonlinearities between base model outputs and effective generalization across diverse data. The ensemble model is then trained and evaluated using 5-fold stratified cross-validation to ensure consistent performance across classes. Although higher fold values were tested, no significant performance improvement was observed beyond 5 folds. This is because increasing the number of folds results in smaller validation sets, which tend to have higher variance in evaluation scores. Meanwhile, the incremental increase in training set size results in minimal bias reduction, which is disproportionate to the additional computational cost. Table II summarizes the key parameters of our ML model.

IV. Evaluation

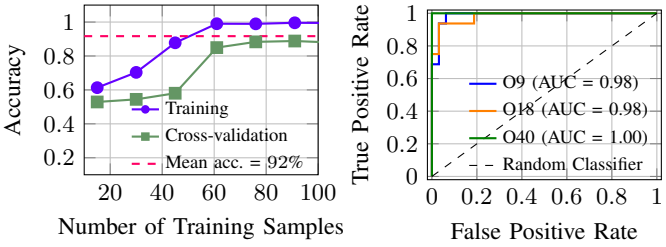
In the following, we present the results and discuss the impact of biomarker decay on the performance of ML classification models. We conducted simulations for 500 s, ensuring a continuous release of 100 biomarkers per second from the infection source and observed by the gateways. The gateways perform measurements within the vessel segments of the axillaris vein, located in the right and left arms, corresponding to vessel IDs 74 and 75, respectively. Different infection sources, head, thorax, and kidney, correspond to blood vessel segments O9, O18, and O40, respectively, in the BVS simulator. These segments are treated as distinct classes in the classification model for source identification. The biomarkers circulated within the HCS throughout the simulation and were assumed to decay exponentially [16, Eq. (60)], i.e., their degradation follows the rule $e^{-r_{\text{decay}}t}$, where r_{decay} is the decay rate as 0.007, 0.015 and $0.035\,\text{s}^{-1}$. We evaluated model performance using classification accuracy and Receiver Operating Characteristic (ROC) curves, offering insights into both overall accuracy and class-wise classification. All simulation data[1] and the Python code used for data processing[2] and model evaluation are made publicly available under the CC BY and MIT licenses. We also provide a supplementary file with biomarker distribution profiles and results for the 70/30 train/test split.

We highlight that this decay model is a heuristic approximation and not derived from a detailed physical formulation of biomarker transport and degradation. This modeling aligns with the assumptions of the BVS framework, which simulates biomarker transport under laminar flow without accounting for turbulence, diffusion, or complex boundary interactions. Accordingly, exponential decay is applied uniformly to estimate net biomarker loss, as localized effects like cardiac turbulence have minimal impact at the wrist-located gateway.

A. Results

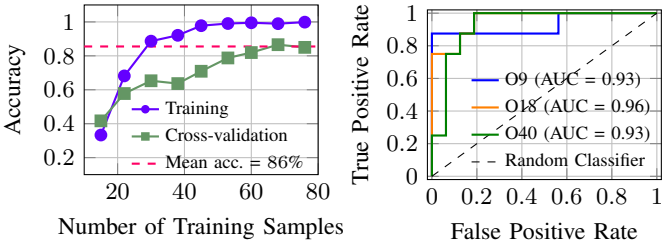
Fig. 1 illustrates the performance of the ensemble classification model under a decay rate of $r_{\text{decay}} = 0.007$. Fig. 1a presents the learning curve, showing classification accuracy on both training and validation sets as a function of the number of training samples. The model demonstrates improved generalization with increasing data, achieving a mean cross-validation accuracy of 92%, as indicated by the dashed red line. Fig. 1b shows the ROC curve for each class, O9, O18, and O40, with area under the curve (AUC) values of 0.98, 0.98, and 1.00, respectively. The results highlight the ability of the model to effectively differentiate among infection source classes based on the temporal patterns of biomarker data.

Fig. 2 depicts the performance of the ensemble classification model for a decay rate of $r_{\text{decay}} = 0.015$. Fig. 2a shows the learning curve depicting training and validation accuracy as a function of the number of training samples. As the decay rate increases, the mean cross-validation accuracy drops to 86%. However, the upward trend indicates that the performance improves as the training data increases. Fig. 2b illustrates the



(a) Accuracy. (b) ROC curve.

Fig. 1: Performance of the ensemble model for $r_{\text{decay}} = 0.007$.

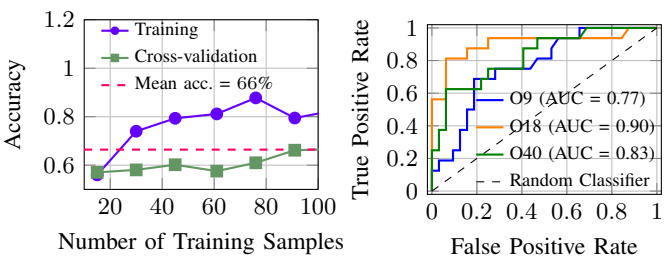


(a) Accuracy. (b) ROC curve.

Fig. 2: Performance of the ensemble model for $r_{\text{decay}} = 0.015$.

ROC curves for the three infection source classes, O9, O18, and O40, with respective AUC values of 0.93, 0.96, and 0.93. These values reflect good class separability, with a modest reduction in performance compared to the $r_{\text{decay}} = 0.007$ scenario. Overall, the results demonstrate that the ensemble model maintains robust classification performance even under increased biomarker decay conditions.

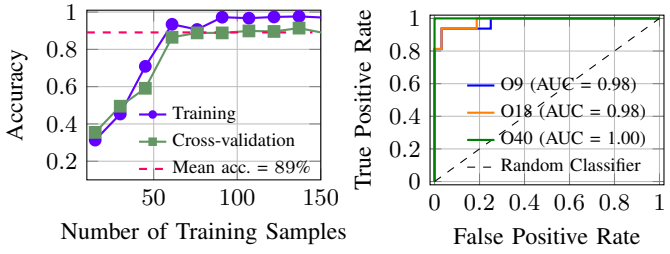
Fig. 3 illustrates the performance of the ensemble classification model under a higher biomarker decay rate of $r_{\text{decay}} = 0.035$. As the number of training samples increases, the training accuracy improves, reaching 90%, as shown in Fig. 3a. However, the validation accuracy levels off at a relatively low level, resulting in a mean cross-validation accuracy of 66%. The gap between training and validation curves indicates increased overfitting and reduced generalization when biomarkers degrade more rapidly. Fig. 3b shows the ROC curves, where O18 remains clearly distinguishable, while reduced separability for the other classes indicates increased classification uncertainty as decay increases. These results indicate that faster biomarker degradation significantly reduces accuracy, particularly for sources located at a greater distance from the gateways. This highlights the influence of decay dynamics on model



(a) Accuracy. (b) ROC curve.

Fig. 3: Performance of the ensemble model for $r_{\text{decay}} = 0.035$.

[1] https://doi.org/10.5281/zenodo.15303191
[2] https://github.com/tkn-tub/BVS-Localization



(a) Accuracy. (b) ROC curve.

Fig. 4: Performance of the ensemble model for linear decay.

performance.

We also simulated linear degradation by removing a constant 60 biomarkers per time step, resulting in a steady decrease independent of the current count. Fig. 4 evaluates the performance of the ensemble model when subjected to a linear decay scenario. Fig. 4a shows the learning curve, where both training and validation accuracies increase with the number of training samples and eventually converge. A mean cross-validation accuracy of 89% demonstrates that the model maintains effective generalization across folds, closely aligning with the 92% achieved under the exponential decay model with $r_{\text{decay}} = 0.007$. The high AUC values in Fig. 4b indicate robust classification performance under linear decay, suggesting that the consistent reduction in biomarker count over time preserves the distinct temporal pattern of data.

B. Discussion

The analysis of our results reveals that biomarker decay significantly influences the performance of ML models for infection source localization. Specifically, as shown in Figures 1a, 2a and 3a, the slower biomarker decay leads to more accurate model performance, as indicated by higher accuracy and more consistent loss curves. This is because the slower decay of the biomarkers allows the model to better capture the temporal dynamics of biomarker release, which is crucial for accurately classifying and thereby localizing the infection source. In contrast, faster decay introduces more noise and reduces the effectiveness of the model. In this scenario, the biomarker count drops too rapidly, making it difficult for the model to distinguish temporal patterns. The analysis therefore emphasizes the importance of understanding and accounting for the natural degradation of biomarkers within the HCS. It is essential to accurately model the biological processes to be able to classify and localize infection sources effectively.

V. Conclusions

In this work, we leveraged ML models to localize infection sources in the HCS by analyzing time-series biomarker data. To obtain the biomarker data, we extended the open-source simulator BVS, which models infection sources and biomarker release, decay, and mobility through the blood vessels. We generated time-series biomarker data by simulating the release of biomarkers from specific organs, i.e., the head, thorax, and kidneys. Biomarkers were observed at two locations in the simulator (left and right arm blood vessels), and each data sample was labeled according to the source organ releasing the biomarkers. We employed a multi-class classification approach based on the temporal dynamics of the biomarkers to identify the organ that was releasing the biomarkers. We evaluated the impact of biomarker decay on the performance of the ML models in the context of infection source localization. Our results highlight the importance of accurately modeling biomarker behavior to improve the effectiveness of ML-based localization methods.

References

[1] I. F. Akyildiz, M. Ghovanloo, U. Guler, T. Ozkaya-Ahmadov, A. F. Sarioglu, and B. D. Unluturk, "PANACEA: An Internet of Bio-NanoThings Application for Early Detection and Mitigation of Infectious Diseases," *IEEE Access*, vol. 8, pp. 140 512–140 523, Jan. 2020.

[2] L. Jiang, X. Lin, F. Chen, et al., "Current research status of tumor cell biomarker detection," *Microsystems & Nanoengineering*, vol. 9, no. 1, p. 123, Oct. 2023.

[3] N. Etemadi, M. Farahnak-Ghazani, H. Arjmandi, M. Mirmohseni, and M. Nasiri-Kenari, "Abnormality Detection and Localization Schemes Using Molecular Communication Systems: A Survey," *IEEE Access*, vol. 11, pp. 1761–1792, 2023.

[4] A. Abdali and M. Kuscu, "Frequency-Domain Model of Microfluidic Molecular Communication Channels With Graphene BioFET-Based Receivers," *IEEE Transactions on Communications*, vol. 72, no. 8, pp. 4564–4576, Aug. 2024.

[5] 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)*, Reykjavík, Iceland: ACM, Sep. 2018, 5:1–5:6.

[6] S. Meng, D. Tripathy, E. P. Frenkel, et al., "Circulating Tumor Cells in Patients with Breast Cancer Dormancy," *Clinical Cancer Research*, vol. 10, no. 24, pp. 8152–8162, Dec. 2004.

[7] S. Pal, J. Torres Gómez, R. Wendt, S. Fischer, and F. Dressler, "Age of Information-based Abnormality Detection with Decay in the Human Circulatory System," *IEEE Transactions on Molecular, Biological and Multi-Scale Communications*, vol. 10, no. 3, pp. 487–492, Sep. 2024.

[8] J. Simonjan, B. D. Unluturk, and I. F. Akyildiz, "In-body Bionanosensor Localization for Anomaly Detection via Inertial Positioning and THz Backscattering Communication," *IEEE Transactions on NanoBioscience*, vol. 21, no. 2, pp. 216–225, Apr. 2022.

[9] R. Wendt, F.-L. A. Lau, L. Unger, and S. Fischer, "Proteome Fingerprinting as a Localization Scheme for Nanobots," in *10th ACM International Conference on Nanoscale Computing and Communication (NANOCOM)*, Coventry, United Kingdom: ACM, Sep. 2023.

[10] J. Torres Gómez, A. Kuestner, J. Simonjan, B. D. Unluturk, and F. Dressler, "Nanosensor Location Estimation in the Human Circulatory System using Machine Learning," *IEEE Transactions on Nanotechnology*, vol. 21, pp. 663–673, Oct. 2022.

[11] G. Pascual, F. Lemic, C. Delgado, and X. Costa-Pérez, "Analytical Modelling of Raw Data for Flow-Guided In-body Nanoscale Localization," in *IEEE International Conference on Machine Learning for Communication and Networking (ICMLCN)*, Stockholm, Sweden: IEEE, May 2024, pp. 428–433.

[12] C. G. Caro, T. J. Pedley, R. C. Schroter, and W. A. Seed, *The Mechanics of the Circulation*, 2nd ed. Cambridge, UK: Cambridge University Press, 2012.

[13] D. Wang, Y. Zhang, and Y. Zhao, "LightGBM: An Effective miRNA Classification Method in Breast Cancer Patients," in *International Conference on Computational Biology and Bioinformatics (ICCBB)*, Newark, NJ: ACM, Oct. 2017, pp. 7–11.

[14] G. Singh, B. Kumar, L. Gaur, and A. Tyagi, "Comparison between Multinomial and Bernoulli Naïve Bayes for Text Classification," in *International Conference on Automation, Computational and Technology Management (ICACTM)*, London, United Kingdom: IEEE, Apr. 2019.

[15] A. Parmar, R. Katariya, and V. Patel, "A Review on Random Forest: An Ensemble Classifier," in *International Conference on Intelligent Data Communication Technologies and Internet of Things (ICICI)*, Coimbatore, India: Springer, Dec. 2018, pp. 758–763.

[16] V. Jamali, A. Ahmadzadeh, W. Wicke, A. Noel, and R. Schober, "Channel Modeling for Diffusive Molecular Communication - A Tutorial Review," *Proceedings of the IEEE*, vol. 107, no. 7, pp. 1256–1301, Jul. 2019.

Saswati Pal is a post-doctoral researcher with the Telecommunication Networks Group at the Technical University Berlin. She received her Ph.D. in 2023 from the Indian Institute of Technology Kharagpur, where she previously worked in the Department of Computer Science and Engineering as a Senior Research Fellow. She received her Master's degree in Electronics and Communication from National Institute of Technology Jalandhar, India in 2016. Her current research interests include Molecular Communication, Internet of Things, and Wireless Body Area Networks.

Jorge Torres Gómez is a Senior Researcher with the Telecommunication Networks Group at the Technical University of Berlin. He specializes in wireless and molecular communications, focusing on topics such as the Internet of Bio-Nano-Things, Age of Information, Neural Networks applied to Nanonetworks. Jorge Torres has led several projects, including a BMBF-funded initiative on IoBNT and a DFG-funded project on in-body nanocommunications. With over a decade of academic and research experience, he has contributed to areas like machine learning-assisted communication systems, energy-efficient IoT, and educational methodologies. He has served as a TPC Co-chair for the ACM NanoCom and leads educational initiatives as the Head of IEEE Germany Section's Educational Activities.

Lisa Y. Debus (Student Member, IEEE) received her B.Sc. in Computer Engineering from TU Berlin in 2022 and completed the consecutive M.Sc. in 2024. She is currently pursuing her Ph.D. degree with the Telecommunication Networks Group, Department of Telecommunication Systems, Technical University of Berlin, Berlin. Her research interests include Molecular Communication, the Internet of Bio-Nano Things, and Semantic Communication.

Regine Wendt (née Geyer) received her Bachelor's degree in Molecular Life Science in 2014 and her Master's degree in Computer Science with a focus on applications in Bioinformatics in 2017 at the University of Lübeck, Germany. In 2025, she obtained her doctoral degree at the Institute of Telematics, University of Lübeck, where she continues her research as a postdoctoral researcher. She is also a long-standing member of the Nano Group Lübeck, founded in 2016. Her research centers on the use of nanodevices in precision medicine, with a particular focus on simulating the cardiovascular system as a working environment for medical nanobots and developing strategies to optimize their functionality for diagnostic and therapeutic applications.

Florian Lau received his Bachelor's in computer science in 2013 and his Master's degree in computer science focused on Computational Complexity Theory in 2015 respectively. Since 2016 he worked as a PhD student at the University of Lübeck. He is a head of the Nano Group Lübeck, which was founded in 2016. In 2020 he received his PhD which was awarded best PhD thesis in 2020 in the area of distributed systems by KuVS. He is since then working as a postdoctoral researcher at the University of Lübeck. His work focuses on DNA-based Self-Assembly Systems and their applications to Nanorobots and Nanonetworks.

Cyrus Khandanpour has been professor of Translational Cancer Research at the University of Lübeck since July 2022. He holds his professorship at the Department of Hematology and Oncology, University Medical Center Schleswig-Holstein, Lübeck Campus, where he heads the Department of Translational Cancer Research, Multiple Myeloma, Amyloidosis, and Histiocytosis. Previously, he was a department head and senior physician at the University Hospital Münster, where he established several clinical studies. His research focuses on transcription factors in leukemia and lymphomas, as well as advancing precision oncology and molecular tumor boards.

Malte Sieren is an assistant professor and senior physician in Radiology at the University Hospital Schleswig-Holstein, Campus Lübeck, where he leads the Artificial Intelligence in Radiology group. He studied medicine at the University of Lübeck and received his doctoral degree in 2019 and his habilitation in Radiology in 2024. His research focuses on AI-assisted imaging and image-guided therapies, particularly for aortic and cardiovascular diseases. Dr. Sieren has authored more than 55 peer-reviewed publications, holds two patents, and has received several national and international awards, including the JVIR Distinguished Laboratory Investigation Award and the DRG Wachsmann Innovation Prize.

Stefan Fischer is a full professor in Computer Science at the University of Lübeck, Germany, and the director of the Institute for Telematics. He got his diploma degree in Information Systems and his doctoral degree in Computer Science from the University of Mannheim, Germany, in 1992 and 1996, respectively. After a postdoctoral year at the University of Montreal, Canada, he held positions at the International University in Germany as an assistant professor and at the Technical University of Braunschweig as an associate professor, until he joined Lübeck University in 2004. His research interest is focused on network and distributed system structures such as ad-hoc and sensor networks, Internet of Things, Smart Cities and especially molecular communications on the nano level. He has (co-)authored more than 220 scientific books and articles.

Falko Dressler is full professor and Chair for Data Communications and Networking at the School of Electrical Engineering and Computer Science, TU Berlin. He received his M.Sc. and Ph.D. degrees from the Dept. of Computer Science, University of Erlangen in 1998 and 2003, respectively. Dr. Dressler has been associate editor-in-chief for IEEE Trans. on Mobile Computing and Elsevier Computer Communications as well as an editor for journals such as IEEE/ACM Trans. on Networking, IEEE Trans. on Network Science and Engineering, Elsevier Ad Hoc Networks, and Elsevier Nano Communication Networks. He has been chairing conferences such as IEEE INFOCOM, ACM MobiSys, ACM MobiHoc, IEEE VNC, IEEE GLOBECOM. He authored the textbooks Self-Organization in Sensor and Actor Networks published by Wiley & Sons and Vehicular Networking published by Cambridge University Press. He has been an IEEE Distinguished Lecturer as well as an ACM Distinguished Speaker. Dr. Dressler is an IEEE Fellow as well as an ACM Distinguished Member. He is a member of the German National Academy of Science and Engineering (acatech). He has been serving on the IEEE COMSOC Conference Council and the ACM SIGMOBILE Executive Committee. His research objectives include adaptive wireless networking (sub-6GHz, mmWave, visible light, molecular communication) and wireless-based sensing with applications in ad hoc and sensor networks, the Internet of Things, and Cyber-Physical Systems.