XAI-Enhanced Bilateral Molecular Communication: Revealing Cancer Microenvironment Dynamics via Extracellular Tumor Vesicles

Osman Tugay Basaran, Jorge Torres Gómez, and Falko Dressler,
School of Electrical Engineering and Computer Science, TU Berlin, Germany
{basaran, torres-gomez, dressler}@ccs-labs.org

***Abstract*—Molecular communications, particularly via extracellular vesicles (EVs), play a critical role in the tumor microenvironment by influencing processes such as anti-tumor signaling and cellular responses like migration. EVs function as signaling pathways mediators, affecting the target cells and initiating feedback loops that establish a bilateral communication system. Evaluating the distance between the two cells remains challenging due to the complexity of these interactions, especially in closed-loop systems. Overcoming this challenge, we train a neural network (NN) using features extracted from raw data, including the numerical differentiation of received molecules and the amplitude and location of resulting peaks. Our major contribution is to provide a plausible explanation of the NN operation. By applying explainable artificial intelligence (XAI) frameworks, such as Shapley values and manual permutation importance, we provide deeper insights and proof of correctness for the NN-based distance estimator. While existing work often relies on local interpretable model-agnostic explanations (LIME) and individual conditional expectation (ICE) for model interpretability, these methods have limitations in capturing the complex feature interactions and nonlinearities inherent in biological systems like molecular communication. Our findings underline the potential of XAI in making complex molecular interactions more transparent, providing critical understanding of the tumor microenvironment, and assisting in the development of more targeted cancer treatments.**

***Index Terms*—Molecular communication, cell-two-cell communications, machine learning, neural networks, explainable AI**

I. Introduction

Molecular communication (MC) is an emerging interdisciplinary field focusing on the exchange of biochemical signals at the nanoscale, crucial for understanding cellular interactions, particularly in the context of precision medicine for cancer treatment [1]. In the context of cancer cells, one of the primary channels for such communication is through EVs. EVs, including exosomes and microvesicles, are lipid-bound particles released by cells to transmit molecular information such as proteins, lipids, and ribonucleic acid (RNA) between cells. This process influences critical functions in the tumor microenvironment, including cancer initiation, progression, immune response modulation, inflammation, and apoptosis.

The transmission of molecular signals via EVs is often modeled as a communication process between a transmitter (donor cell) and a receiver (target cell). In the tumor microenvironment, EVs serve as external stimuli that can activate or inhibit receptor pathways in target cells, leading to either amplification or attenuation of further signaling responses. These interactions frequently operate in a bilateral feedback loop, where the EV-mediated signal induces further EV release, reinforcing the communication process. The understanding of these dynamic interactions, particularly the frequency response of EV internalization and release, is essential for uncovering how malignant transformations and cellular responses are orchestrated in cancer.

While traditional models focus on the physical and biochemical aspects of EV-mediated communication, there is a growing need to incorporate advanced machine learning techniques to capture the intricate, non-linear relationships between cellular variables. AI, particularly NNs [5], has shown promise in modeling complex biological systems due to its ability to learn patterns and relationships from data without explicit programming [6]. NNs can model the nonlinear dependencies and interactions inherent in biological systems, making them suitable for studying MC mechanisms in cancer.

However, a significant limitation of traditional AI models is their lack of interpretability. The "black-box" nature of NNs means that while they can make accurate predictions, they do not provide insights into the underlying mechanisms or rationale behind those predictions. This lack of transparency is a critical barrier in fields like cancer research, where understanding findings' biological significance and implications is essential. Explainable Artificial Intelligence (XAI) [7–9] has emerged as a solution to this problem, aiming to make AI models more transparent and their predictions more understandable to humans. XAI techniques enable researchers to interpret and trust AI models by explaining their decisions. In the context of MC and cancer research, XAI can help uncover the mechanisms by which EVs influence cellular behavior and interactions, thereby providing valuable insights that can inform therapeutic strategies.

In this study, we introduce XAI techniques to enhance the interpretability of AI-based MC models. We focus on estimating the distance between immune cells and cancer cells based on the number of exchanged EVs. Distance knowledge conveys a significant role in later localizing and actuating over the tumor cell. Specifically, we develop a NN to model the cell-to-cell communication link between T-cells and tumor cells. Cancer cells naturally release EVs that reach immune cells, and upon detection, T-cells respond by releasing EVs back

TABLE I: Previous XAI research on molecular communication

XAI Studies	Neural Network	Year	Focus	Data type	XAI Method
Explainability of NN-based Detectors in MIMO Molecular Channels [2]	Feedforward	2023	Symbol Detection	Synthetic	LIME, ICE
Explainability of NNs for Symbol Detection in Molecular Communication Channels [3]	Feedforward	2023	Symbol Detection	Synthetic	LIME, ICE
Explainable Asymmetric Auto-Encoder for End-to-End Learning of IoBNT Communications [4]	Autoencoder	2024	Symbol Detection	Synthetic	Linear Modelling
Our Study	Feedforward	2024	Distance Estimation	Synthetic	Permutation Importance, Shapley

into the medium. This exchange creates a feedback loop, with the number of released and received EVs depending on the distance between the cells.

Building upon previous research that models this closed-loop vesicle exchange in [1], we address the challenge of intertwined dependencies in the analytical expressions governing EV exchange, which involve complex terms related to the spatial positions of cells within the extracellular matrix. These dependencies make it difficult to derive closed-form expressions for the distance between cells. To overcome this, we employ a NN that learns the relationship between the number of induced released vesicles by the T-cell and the distance to the tumor cell.

Furthermore, by applying XAI techniques to our NN model, we enhance its interpretability, enabling us to understand how different input features contribute to the output predictions. This not only increases the transparency of the model but also provides valuable insights into the biological processes underlying MC in the tumor microenvironment. Such insights are crucial for cancer research, as they can inform the development of novel therapeutic interventions that target specific aspects of EV-mediated communication. The core outcomes of our study are summarized as new contributions ("C") and new findings ("F") as follows:

C1. We present a NN model that accurately estimates the distance between immune cells and cancer cells based on vesicle exchange data. This model addresses the challenges posed by the complex dependencies in analytical expressions, offering a practical solution for studying cell-to-cell communication.

C2. We integrate XAI methods into our NN model, enhancing its interpretability. This allows researchers to understand the model's decision-making process, bridging the gap between AI predictions and biological understanding.

F1. Our NN model successfully estimates the distance between the T-cell and tumor cell within the range of $2-10\,\mu\text{m}$, achieving a relative error of $3.3\,\%$.

F2. Our XAI method indicates that the peak's time coordinate is the most important feature in the model, as permuting its values leads to a substantial increase in prediction error while shuffling the peak amplitude has a negligible effect.

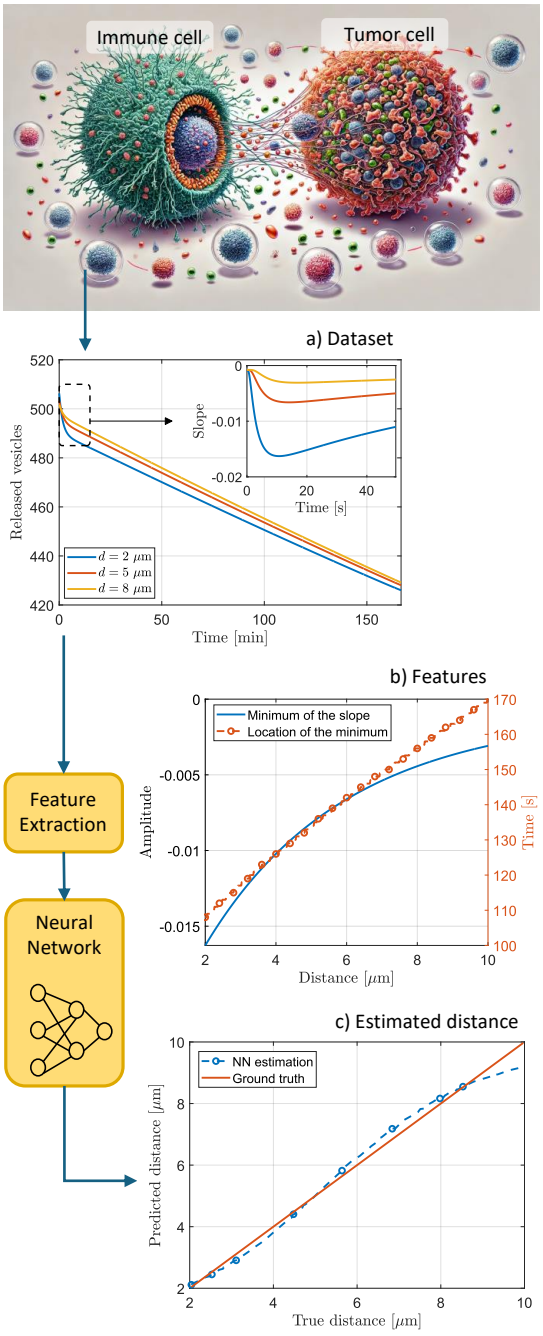

Fig. 1: Diagram for estimating the distance among cells using a feedforward NN.

II. System Model

A. Background

Very few recent studies have applied machine learning models, specifically NNs, to improve symbol detection in MC systems, addressing challenges in evaluating end-to-end channel models and reducing bit error rates [2, 3]; as can be seen in Table I. Recognizing the lack of transparency in these models, XAI methods such as local interpretable model-agnostic explanations (LIME) [10], partial dependence plots (PDP) [11], and ICE [12] plots have been employed to interpret NN-based symbol detectors in MC channels, revealing underlying mechanisms like threshold detection based on molecule counts. However, these studies focus on symbol detection, and there is a gap in applying XAI techniques to distance estimation in MC systems, particularly within the context of cancer research—a gap that our work aims to address.

B. Data Collection

The dataset to train and test the NN module is created with the code provided by the authors in [1], which models the reaction-diffusion propagation of EVs in the extracellular medium. The dataset comprises the number of released vesicles with time and distance. This number of EVs is evaluated as the superposition of the natural release mechanism (see [1, Eq. (39)]) and the induced number upon the reception of vesicles from the tumor cell. As illustrated in Figure 1 a), the number of released vesicles by the immune cell decreases linearly with time but also with distance. This behavior makes the peak concentration of EVs a feature to explore when estimating the distance between the immune and the tumor cells.[1]

C. NN Estimator

Mathematical diffusion models are used to predict how the concentration of vesicles changes as they move away from the donor cell. However, finding exact solutions to these models is often challenging because they involve complex terms representing the precise locations of both the vesicle-releasing (donor) cell and the receiving (target) cell within the extracellular matrix. The concentration of vesicles at a certain time and distance depends on several factors, including the rate at which vesicles are released, the diffusion coefficient (which indicates how quickly vesicles spread through the medium), and the specific position of the donor cell. Estimating the distance between cells by solving these models is particularly difficult due to the complexity of working backward through the equations—known as an inverse problem—and the inherently random nature of vesicle diffusion and reception processes. NNs offer a powerful alternative by learning the complex relationship between vesicle exchange patterns and intercellular distances from data.

[1]We provide open access to the dataset code and the NN estimator in https://github.com/tkn-tub/NN_molecular_communications/tree/main/Section_III_A_distance_estimator

TABLE II: NN parameters for the distance estimation

Parameter	Value
Number of Layers	1
Total of nodes	2
Activation function	Sigmoid
Loss tolerance	1×10^{-9}

We implement a simple architecture for estimating the distance with a feature extraction block and a feedforward NN; see its parameters in Table II. Following the developments in [13], the feature extraction block evaluates the slope of the number of released vesicles with time (see the subfigure in Figure 1 a)) and outputs the slope minimum (Feature 1) and its location ,i.e, time coordinate (Feature 2), see Figure 1 b). The NN, comprised of a single hidden layer and 2 nodes, is trained and tested with these two inputs only, which renders a low-complex model for the distance estimation. As depicted in Figure 1 c), this architecture accurately estimates the distance between the immune and cancer cells. This simple NN accurately estimate intercellular distances within the $2-10\,\mu\mathrm{m}$ range, achieving a relative error of $3\,\%$.

III. Proposed XAI-Enhanced Bilateral Molecular Communication

As presented in the previous section, neural can model the nonlinear dependencies without explicit analytical formulations, making them suitable for capturing the essence of bilateral MC. However, these models often function as "black boxes," providing accurate predictions without revealing the underlying decision-making processes. By incorporating XAI techniques, we can dissect the NN's internal workings to interpret how different input features influence the output predictions. With proposed XAI methods, we can quantify the contribution of each input feature x_i to the output y, measuring a relevancy score as [14]:

$$R_i = x_i \frac{\partial y}{\partial x_i} \tag{1}$$

A. Manual Permutation Importance

In the context of our NN model estimating the distance between immune cells and cancer cells based on the number of exchanged vesicles, understanding the contribution of each input feature to the model's predictions is crucial. This is where the Manual Permutation Importance method becomes invaluable. It quantifies the importance of each feature by measuring the decrease in model performance when the feature's values are randomly shuffled, thereby breaking the relationship between the feature and the true outcome. Let us consider a trained NN model $f(\mathbf{X})$ that predicts the distance between cells based on input features $\mathbf{X} = [X_1, X_2, \ldots, X_p]$, where p is the total number of features. The true target variable is y, representing the actual distance between immune and cancer cells. First, we compute the original performance of the model

on a validation dataset, using a suitable error metric such as the mean squared error (MSE):

$$\text{Perf}_{\text{baseline}} = \frac{1}{N}\sum_{j=1}^{N}\left(y_j - f(\mathbf{X}_j)\right)^2 \tag{2}$$

where N is the number of samples, and $\mathbf{X}_j$ is the feature vector for sample j. Next, for each feature X_i, we create a permuted version X_i^{perm} by randomly shuffling its values:

$$X_i^{\text{perm}} = \text{Permute}(X_i). \tag{3}$$

Then construct a new dataset $\mathbf{X}^{\text{perm}}$ by replacing X_i with X_i^{perm}:

$$\mathbf{X}^{\text{perm}} = \left(X_1, X_2, \ldots, X_i^{\text{perm}}, \ldots, X_p\right) \tag{4}$$

We evaluate the model's performance on the permuted dataset:

$$\text{Perf}_{\text{permuted}} = \frac{1}{N}\sum_{j=1}^{N}\left(y_j - f(\mathbf{X}_j^{\text{perm}})\right)^2. \tag{5}$$

The importance of feature X_i is quantified as the difference in performance before and after permutation:

$$I(X_i) = \text{Perf}_{\text{permuted}} - \text{Perf}_{\text{baseline}}. \tag{6}$$

A larger value of $I(X_i)$ indicates a greater impact on the model's performance, signifying that feature X_i is more important for accurate predictions. By applying Manual Permutation Importance, we can rank the input features based on their influence on the model's predictive accuracy. This not only enhances the interpretability of the NN but also provides valuable biological insights. For instance, identifying that certain features, such as the amplitude of vesicle concentration peaks, have a higher importance could suggest their critical role in the communication between cells. Understanding feature importance helps validate the model and ensures that it aligns with known biological phenomena. It also aids in uncovering new patterns or relationships that may warrant further investigation. By integrating this XAI method, we bridge the gap between predictive performance and model transparency, which is essential for the acceptance and application of AI models in cancer research.

B. Explainability through Shapley Values

The Shapley Value for a feature quantifies its average marginal contribution to the prediction over all possible subsets of features. Let's $f(\mathbf{x})$ be the model's prediction function, $\mathbf{x} = [x_1, x_2, \ldots, x_p]$ be the input feature vector and $N = \{1, 2, \ldots, p\}$ be the set of all feature indices. The Shapley Value ϕ_i for feature x_i is defined as [15]:

$$\phi_i = \sum_{S \subseteq N \setminus \{i\}} \frac{|S|!(p - |S| - 1)!}{p!}\left[f_{S\cup\{i\}}(\mathbf{x}_{S\cup\{i\}}) - f_S(\mathbf{x}_S)\right] \tag{7}$$

where S is any subset of the feature indices excluding i, $|S|$ is the number of elements in S, $f_S(\mathbf{x}_S)$ is the model's prediction using features in subset S (other features are set to a baseline value). The term $f_{S\cup\{i\}}(\mathbf{x}_{S\cup\{i\}}) - f_S(\mathbf{x}_S)$ represents the marginal contribution of feature x_i when added to subset S. The weights $\frac{|S|!(p-|S|-1)!}{p!}$ ensure that each subset's contribution is weighted fairly. Calculating exact Shapley Values requires evaluating all 2^{p-1} subsets, which is computationally infeasible for large p. To overcome this, approximation methods such as sampling or algorithms like SHAP (SHapley Additive exPlanations) [15] are used to efficiently estimate Shapley Values. By utilizing tools like SHAP, we can compute approximate Shapley Values for our NN model:

1) **Compute Baseline Prediction**: Establish a baseline prediction $E[f(\mathbf{x})]$ using baseline feature values.
2) **Estimate Shapley Values**: Calculate ϕ_i for each feature x_i using the model and the data.
3) **Interpret Results**: Analyze the Shapley Values to understand each feature's contribution and interactions.

The model's prediction can then be expressed as:

$$f(\mathbf{x}) = E[f(\mathbf{x})] + \sum_{i=1}^{p}\phi_i, \tag{8}$$

where $E[f(\mathbf{x})]$ is the expected prediction at baseline, and $\sum_{i=1}^{p}\phi_i$ represents the sum of the contributions from all features. Applying Shapley Values in our problem allows us to, *i) Quantify Feature Contributions:* Determine how much each feature contributes to the prediction of intercellular distance, *ii) Understand Feature Interactions:* Reveal interactions between features that may correspond to biological phenomena, *iii) Enhance Model Interpretability:* Provide transparent and fair explanations for the model's predictions.

IV. Experiment Evaluations

In this section, we present a comprehensive analysis of our model's interpretability using several explainability techniques, starting with baseline models such as LIME and ICE and then compared with our implemented Manual Permutation Importance and Shapley Value estimation methods. Starting with LIME, model works by locally approximating the black-box model with a simpler linear model. It focuses on individual predictions and highlights feature importance in those specific instances. In Figure 2; all three observations, Feature 2 (peak time-coordinate) dominates the explanation, contributing significantly to the prediction. Feature 1 (peak's height) shows almost no impact. This suggests that LIME's focus on local perturbations might mislead the interpretation, as it may amplify or underplay the importance of a feature without generalizing across the entire dataset. Vesicle dynamics involve interactions that vary greatly depending on cell proximity and environmental factors. Therefore, a local explanation, such as LIME, might miss out on the holistic importance of a feature across all instances. Since LIME simplifies the model locally, it fails to provide insights into the global feature importance that is critical for accurate biological interpretation.

ICE plots visualize how the model's prediction changes when a particular feature value is varied while other features remain constant. ICE result for Feature 1 demonstrates almost

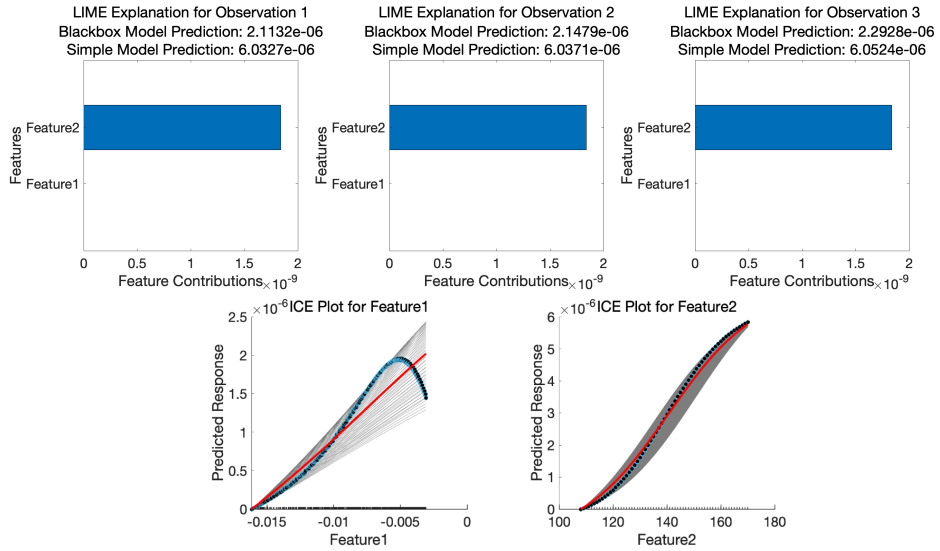

Fig. 2: Feature Interpretability Analysis: LIME-Based Explanations for Selected Observations (Top Row) and Individual Conditional Expectation (ICE) Plots Depicting Feature-1 and Feature-2 Influence on Model Predictions (Bottom Row)

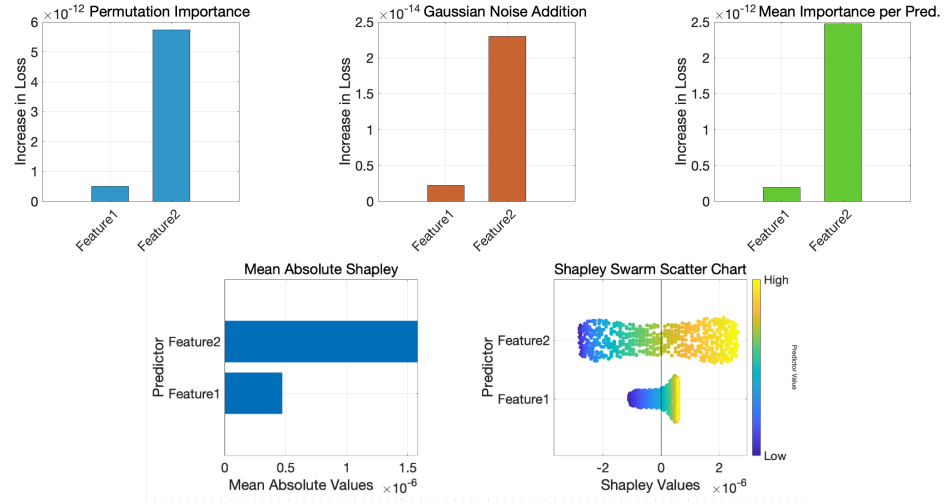

Fig. 3: Feature Importance Analysis: Manual Permutation Importance for Evaluating Loss Sensitivity (Top Row) and Shapley Value-Based Insights into Feature Contributions (Bottom Row) for our feedforward NN Model

linear changes in predicted response as Feature 1 varies. This suggests that Feature 1 contributes very little to the overall prediction, with most predictions staying close to the same response for different values of Feature 1. Feature 2's ICE plot shows a clear non-linear increase in predicted response as the feature value increases. This indicates that Feature 2 plays a significant role in the prediction, with a curved relationship typical of biological processes.

However, ICE still does not account for interactions between Feature 1 and Feature 2. The isolated analysis could miss important joint effects in your model, which are often key in MC dynamics. On the contrary, manual permutation importance highlight these interactions. As Figure 3 displays, Feature 2 shows an increase in loss of approximately 6×10^{-12}, which is significantly higher than the loss increase from Feature 1 ($< 1 \times 10^{-12}$). This indicates that Feature 2 is the most important feature in the model, as permuting its values leads to a substantial increase in prediction error, while shuffling Feature 1 has a negligible effect. Besides, Shapley values provide a numerical breakdown of how much each feature contributes to the model's output. With a mean Shapley value of 1.5×10^{-6} for Feature 2 compared to 0.4×10^{-6} for Feature 1, Shapley values provide a deeper understanding of how features contribute to predictions, far surpassing LIME's local approximations. All results were verified with a $p - value < 0.001$, indicating strong statistical confidence in

the dominance of Feature 2 in our model's decision-making process.

The various explainability methods highlight the "smart" use of the features by the NN. Coincidentally, in free diffusion, there is a direct relation between traveling time and the square of the distance as $t_{\text{peak}} \propto \frac{d^2}{D}$, which is independent of the peak's amplitude; see [16, below Eq. (2.8)]. In a way, the NN architecture filters the entry related to the peak and evaluates a direct relation between the peak time and distance.

Besides, we concluded that vesicle dynamics involve interactions that vary greatly depending on cell proximity and environmental factors. Therefore, a local explanation, such as LIME, might miss out on the holistic importance of a feature across all instances. Since LIME simplifies the model locally, it fails to provide insights into the importance of global features that are critical for accurate biological interpretation. Moreover, molecular data often has non-linear relationships, and LIME's linear approximation could misrepresent these interactions. For instance, if a non-linear interaction between two features influences vesicle exchange, LIME would not capture this dependency properly. Also in biological systems like MC, the interaction between features (e.g., environmental conditions and vesicle production rates) is often more important than individual feature effects. ICE overlooks this, providing only partial insights into the overall model behavior. These results highlight the global importance of Feature 2, which was not fully captured by LIME or ICE due to their focus on local or isolated feature analysis. Clear numerical distinction in both Permutation Importance (loss increase) and Shapley Values (contribution) provides a mathematically precise understanding of the relative importance of Feature 2 and Feature 1. LIME and ICE lack this distinction because they offer localized or single-feature insights without capturing feature interactions.

V. Conclusion and Future Research Directions

In this study, we developed a NN model to estimate the distance between immune cells and cancer cells based on the number of EVs. By integrating XAI techniques—specifically Shapley Values and Manual Permutation Importance—we enhanced the interpretability of our model, providing valuable insights into the underlying biological processes of MC within the tumor microenvironment. Our main finding is that NN model uses the peak's arrival time to estimate the distance between the immune and tumor cells, disregarding the peak value. This result is in direct relation to the expected performance in free-diffusion MC channels.

Future work could involve exploring advanced NN architectures to capture more intricate patterns in vesicle exchange data. Implementing deeper networks with additional layers or utilizing architectures such as Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) may improve the model's ability to learn spatio-temporal dependencies inherent in biological processes. However, increasing model complexity necessitates robust XAI methods to maintain interpretability. Also, focus on integrating experimental data from in vitro and in vivo studies to refine the model parameters and validate its accuracy in real-world scenarios. Additionally, applying the model to clinical datasets could help assess its utility in predicting patient-specific tumor-immune dynamics, potentially guiding personalized treatment strategies and improving clinical outcomes.

Acknowledgment

This work has been funded by the German Federal Ministry of Education and Research (BMBF, Germany) as part of the 6G Research and Innovation Cluster 6G-RIC under Grant 16KISK020K, the project IoBNT under Grant 16KIS1986K.

References

[1] M. Zoofaghari, F. Pappalardo, M. Damrath, and I. Balasingham, "Modeling Extracellular Vesicles-Mediated Interactions of Cells in the Tumor Microenvironment," *IEEE Transactions on NanoBioscience*, vol. 23, no. 1, pp. 71–80, Jan. 2024.

[2] X. Li, "Explainability of NN-based Detectors in MIMO Molecular Channels," Master's Thesis, School of Electrical Engineering and Computer Science, Berlin, Germany, Jul. 2023.

[3] J. Torres Gómez, P. Hofmann, F. H. P. Fitzek, and F. Dressler, "Explainability of Neural Networks for Symbol Detection in Molecular Communication Channels," in *7th Workshop on Molecular Communications (WMC 2023)*, Erlangen, Germany, Apr. 2023.

[4] R. Khanzadeh, S. Angerbauer, J. Torres Gómez, P. Hofmann, F. Dressler, F. H. P. Fitzek, A. Springer, and W. Haselmayr, "Explainable Asymmetric Auto-Encoder for End-to-End Learning of IoBNT Communications," in *IEEE International Conference on Machine Learning for Communication and Networking (ICMLCN 2024)*, Stockholm, Sweden: IEEE, May 2024.

[5] W. S. McCulloch and W. Pitts, "A logical calculus of the ideas immanent in nervous activity," *The Bulletin of Mathematical Biophysics*, vol. 5, no. 4, pp. 115–133, Dec. 1943.

[6] M. Bartunik, J. Kirchner, and O. Keszocze, "Artificial intelligence for molecular communication," *it - Information Technology (itIT)*, vol. 65, no. 4-5, pp. 155–163, Aug. 2023.

[7] C. Paris, J. Moore, and W. Swartout, "Explanations in knowledge systems: design for explainable expert systems," *IEEE Expert*, vol. 6, no. 3, 58â€"64, Jun. 1991.

[8] R. Confalonieri, L. Coba, B. Wagner, and T. R. Besold, "A historical perspective of explainable Artificial Intelligence," *WIREs Data Mining and Knowledge Discovery*, vol. 11, no. 1, Jan. 2021.

[9] L. Longo, M. Brcic, F. Cabitza, J. Choi, R. Confalonieri, J. D. Ser, R. Guidotti, Y. Hayashi, F. Herrera, A. Holzinger, R. Jiang, H. Khosravi, F. Lecue, G. Malgieri, A. Páez, W. Samek, J. Schneider, T. Speith, and S. Stumpf, "Explainable Artificial Intelligence (XAI) 2.0: A manifesto of open challenges and interdisciplinary research directions," *Information Fusion*, vol. 106, p. 102 301, Jun. 2024.

[10] M. T. Ribeiro, S. Singh, and C. Guestrin, ""Why Should I Trust You?": Explaining the Predictions of Any Classifier," in *22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, San Francisco, CA: ACM, Aug. 2016.

[11] B. M. Greenwell, B. C. Boehmke, and A. J. McCarthy, "A Simple and Effective Model-Based Variable Importance Measure," arXiv, cs.IT, May 2018.

[12] A. Goldstein, A. Kapelner, J. Bleich, and E. Pitkin, "Peeking Inside the Black Box: Visualizing Statistical Learning With Plots of Individual Conditional Expectation," *Journal of Computational and Graphical Statistics*, vol. 24, no. 1, pp. 44–65, Jan. 2015.

[13] F. Güleç and B. Atakan, "Distance estimation methods for a practical macroscale molecular communication system," *Elsevier Nano Communication Networks*, vol. 24, p. 100 300, May 2020.

[14] G. Montavon, A. Binder, S. Lapuschkin, W. Samek, and K.-R. Müller, *Layer-Wise Relevance Propagation: An Overview*. Springer, 2019.

[15] S. M. Lundberg and S.-I. Lee, "A Unified Approach to Interpreting Model Predictions," in *31st International Conference on Neural Information Processing Systems (NIPS 2017)*, Long Beach, CA: Curran Associates Inc., Dec. 2017, pp. 4768–4777.

[16] H. C. Berg, *Random Walks in Biology*. Princeton University Press, 1993.