

Supplementary Information

Enabling Multiple Cellular Enumeration Applications of Bioparticle Sensing Platform using Machine Learning

Authors: Muhammad Nabeel Tahir¹, Brandon K Ashley¹, Jianye Sui¹, Mehdi Javanmard¹, and Umer Hassan^{*1,2}

¹ Department of Electrical and Computer Engineering, School of Engineering, Rutgers, The State University of New Jersey, United States of America.

² Global Health Institute, Rutgers, The State University of New Jersey, New Brunswick, United States of America.

*Corresponding Author: umer.hassan@rutgers.edu (UH)

Device setup and data acquisition details:

Figure 1 shows an overview of the system employed in the study, and Section II-D reports briefly on the impedance and video data acquisition protocols, along with the equipment used. The cells alone or conjugated with metal oxide micro particles ($3\mu\text{m}$) were passed through a microfluidic channel shown in Figure S1. The channel consists of three electrodes, each of width $100\mu\text{m}$ and spaced $150\mu\text{m}$ apart. A $30\mu\text{m}$ narrow focus regions were introduced in the $100\mu\text{m}$ wide channel to concentrate the electrical field lines in a small region, with the assumption that the barcoded particles will be affected by the frequencies in the input signal. The channel height was kept constant at $22\mu\text{m}$. A high-speed camera, as discussed in Section II-D, records videos of cells passing through the microfluidic channel. The impedance and simultaneous video recording are controlled by the LabVIEW program, and impedance pulses of cells passing through the channel and their respective video frames are matched based on the timestamps. Once the timestamps of the cells in the video and impedance data are matched, the video frames are manually inspected to quantify the number of particles attached to a cell. Figure S2 (A) shows the captured cells using a high-speed camera with or without the conjugated particles. And Figure S2(B) shows the impedance pulses recorded for each cell for four frequency components.

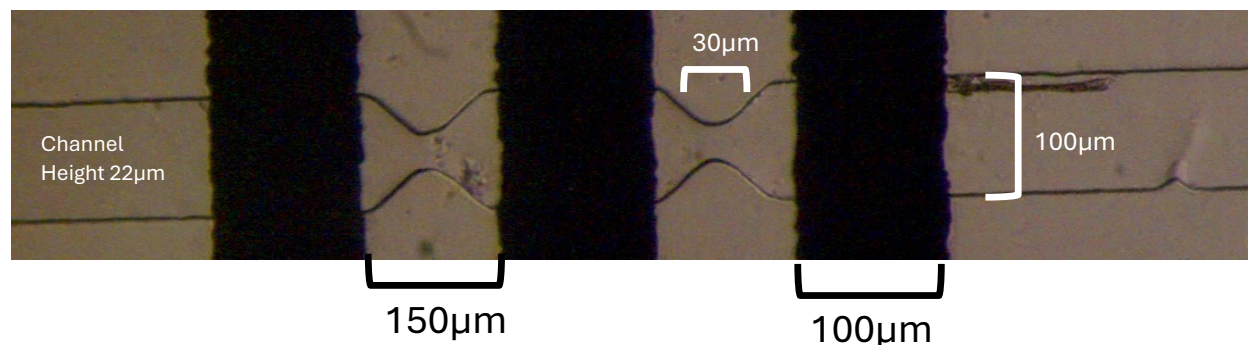


Figure S1: Microfluidic channel dimensions and details.

Table S1: Data distribution of the individual class in all the samples.

Sample	Cells Alone	10nmCD11b1	10nmCD11b2	10nmCD11b3	20nmCD66b1	20nmCD66b2	20nmCD66b3
A	352	263	332	338	290	407	321
B	92	171	123	295	159	278	432
C	97	137	131	166	101	214	278
Total	541	571	586	799	550	899	1031

ISF= Isolation Forest

LOF= Local Outlier Factor

Table S2: Application 1 performance metric analysis.

Data Type	TP	FP	FN	TN	Precision	Recall	F1	Accuracy
Sample A Original Data	140	1	11	42	99.3%	92.7%	95.9%	93.8%
Sample A Pruned with ISF	129	5	5	45	96.3%	96.3%	96.3%	94.6%
Sample A Pruned with LOF	143	1	6	34	99.3%	96.0%	97.6%	96.2%
Sample A Pruned with Mahalanobis	103	2	3	31	98.1%	97.2%	97.6%	96.4%
Sample B Original Data	90	2	3	8	97.8%	96.8%	97.3%	95.2%
Sample B Pruned with ISF	88	0	2	8	100.0%	97.8%	98.9%	98.0%
Sample B Pruned with LOF	86	0	3	9	100.0%	96.6%	98.3%	96.9%
Sample B Pruned with Mahalanobis	61	1	5	8	98.4%	92.4%	95.3%	92.0%
Sample C Original Data	65	0	3	13	100.0%	95.6%	97.7%	96.3%
Sample C Pruned with ISF	67	1	1	8	98.5%	98.5%	98.5%	97.4%
Sample C Pruned with LOF	68	0	2	7	100.0%	97.1%	98.6%	97.4%
Sample C Pruned with Mahalanobis	52	3	0	6	94.6%	100.0%	97.2%	95.1%
Combined Samples Original Data	287	7	15	67	97.6%	95.0%	96.3%	94.2%
Combined Samples Pruned with ISF	280	5	18	54	98.3%	93.9%	96.1%	93.6%
Combined Samples Pruned with LOF	266	10	20	61	96.4%	93.0%	94.7%	91.6%
Combined Samples Pruned with Mahalanobis	203	4	9	36	98.1%	95.7%	96.9%	94.8%

Table S3: Application 2 performance metric analysis.

Data Type	TP	FP	FN	TN	Precision	Recall	F1	Accuracy
Sample A Original Data	166	3	11	27	98.2%	93.8%	96.0%	93.2%
Sample A Pruned with ISF	144	6	10	36	96.0%	93.5%	94.7%	91.8%
Sample A Pruned with LOF	148	7	8	33	95.5%	94.9%	95.2%	92.3%
Sample A Pruned with Mahalanobis	105	4	4	29	96.3%	96.3%	96.3%	94.4%
Sample B Original Data	125	2	7	11	98.4%	94.7%	96.5%	93.8%
Sample B Pruned with ISF	119	6	3	10	95.2%	97.5%	96.4%	93.5%
Sample B Pruned with LOF	120	4	2	12	96.8%	98.4%	97.6%	95.7%
Sample B Pruned with Mahalanobis	91	0	1	8	100.0%	98.9%	99.5%	99.0%
Sample C Original Data	89	0	4	12	100.0%	95.7%	97.8%	96.2%
Sample C Pruned with ISF	81	0	5	13	100.0%	94.2%	97.0%	94.9%
Sample C Pruned with LOF	89	0	2	8	100.0%	97.8%	98.9%	98.0%
Sample C Pruned with Mahalanobis	64	0	4	6	100.0%	94.1%	97.0%	94.6%
Combined Samples Original Data	357	18	17	62	95.2%	95.5%	95.3%	92.3%
Combined Samples Pruned with ISF	349	14	10	59	96.1%	97.2%	96.7%	94.4%
Combined Samples Pruned with LOF	364	6	15	47	98.4%	96.0%	97.2%	95.1%
Combined Samples Pruned with Mahalanobis	247	10	5	35	96.1%	98.0%	97.1%	94.9%

Table S4: Application 3 performance metric analysis.

Data Type	N	Micro Acc.	Macro P	Macro R	Macro F1	Weighted P	Weighted R	Weighted F1
Sample A Original Data	341	75.1%	80.4%	77.4%	78.4%	75.9%	75.1%	74.8%
Sample A Pruned with ISF	329	76.3%	80.0%	77.5%	78.3%	77.0%	76.3%	76.2%
Sample A Pruned with LOF	329	76.6%	78.2%	75.4%	76.0%	77.6%	76.6%	76.2%
Sample A Pruned with Mahalanobis	238	76.9%	79.5%	76.6%	77.5%	77.3%	76.9%	76.5%
Sample B Original Data	233	74.2%	69.6%	64.5%	66.5%	73.9%	74.2%	73.9%

Sample B Pruned with ISF	222	71.2%	74.1%	73.2%	73.2%	71.5%	71.2%	70.7%
Sample B Pruned with LOF	222	71.2%	67.8%	68.1%	67.9%	71.5%	71.2%	71.3%
Sample B Pruned with Mahalanobis	155	61.9%	53.4%	65.4%	56.9%	61.8%	61.9%	61.4%
Sample C Original Data	170	87.1%	87.8%	76.9%	80.5%	87.3%	87.1%	86.5%
Sample C Pruned with ISF	161	87.6%	82.7%	88.5%	84.8%	88.4%	87.6%	87.7%
Sample C Pruned with LOF	161	88.2%	88.9%	86.1%	87.1%	88.8%	88.2%	88.1%
Sample C Pruned with Mahalanobis	119	85.7%	89.6%	76.1%	81.4%	85.6%	85.7%	85.3%
Combined Samples Original Data	748	77.4%	80.0%	75.4%	77.1%	78.4%	77.4%	77.5%
Combined Samples Pruned with ISF	710	75.9%	75.4%	78.7%	76.8%	76.0%	75.9%	75.9%
Combined Samples Pruned with LOF	710	73.9%	75.0%	77.8%	75.3%	76.6%	73.9%	74.0%
Combined Samples Pruned with Mahalanobis	488	71.7%	70.6%	74.2%	72.2%	71.8%	71.7%	71.7%

Table S5: Application 4 performance metrics analysis.

Data Type	N	Micro Acc.	Macro P	Macro R	Macro F1	Weighted P	Weighted R	Weighted F1
Sample A Original Data	194	64.9%	69.9%	65.1%	64.5%	73.0%	64.9%	65.9%
Sample A Pruned with ISF	184	77.7%	76.4%	76.6%	76.4%	78.5%	77.7%	78.0%
Sample A Pruned with LOF	184	69.6%	75.1%	68.6%	68.9%	76.1%	69.6%	69.7%
Sample A Pruned with Mahalanobis	139	73.4%	72.3%	72.0%	72.0%	73.5%	73.4%	73.3%
Sample B Original Data	103	69.9%	68.2%	68.7%	67.6%	70.9%	69.9%	69.6%
Sample B Pruned with ISF	98	75.5%	73.6%	76.1%	74.6%	75.3%	75.5%	75.3%
Sample B Pruned with LOF	98	60.2%	55.1%	51.4%	52.5%	60.2%	60.2%	59.7%
Sample B Pruned with Mahalanobis	75	82.7%	79.9%	82.8%	81.0%	83.2%	82.7%	82.7%
Sample C Original Data	81	84.0%	83.2%	80.8%	80.6%	86.5%	84.0%	84.1%
Sample C Pruned with ISF	77	85.7%	84.5%	84.7%	84.2%	86.2%	85.7%	85.6%
Sample C Pruned with LOF	77	87.0%	87.6%	88.2%	87.1%	88.8%	87.0%	87.2%

Sample C Pruned with Mahalanobis	61	70.5%	68.5%	69.9%	68.3%	73.8%	70.5%	71.3%
Combined Samples Original Data	376	60.4%	67.8%	61.7%	60.5%	67.9%	60.4%	59.5%
Combined Samples Pruned with ISF	357	79.3%	77.8%	78.6%	78.1%	79.7%	79.3%	79.4%
Combined Samples Pruned with LOF	357	75.1%	76.9%	75.3%	75.6%	76.8%	75.1%	75.4%
Combined Samples Pruned with Mahalanobis	253	77.5%	76.3%	76.8%	76.5%	77.9%	77.5%	77.6%

Table S6: Application 5 performance metrics analysis.

Data Type	N	Micro Acc.	Macro P	Macro R	Macro F1	Weighted P	Weighted R (=Acc)	Weighted F1
Sample A Original Data	207	67.1%	75.2%	67.9%	66.0%	77.0%	67.1%	66.2%
Sample A Pruned with ISF	196	81.1%	81.1%	81.1%	81.0%	81.1%	81.1%	81.0%
Sample A Pruned with LOF	196	82.7%	85.1%	82.4%	82.9%	85.0%	82.7%	83.0%
Sample A Pruned with Mahalanobis	141	78.7%	76.4%	76.8%	76.4%	79.9%	78.7%	79.1%
Sample B Original Data	145	83.4%	86.0%	79.4%	81.6%	85.6%	83.4%	83.8%
Sample B Pruned with ISF	138	77.5%	69.4%	66.5%	67.4%	77.3%	77.5%	77.0%
Sample B Pruned with LOF	138	85.5%	81.5%	78.8%	80.0%	84.9%	85.5%	85.1%
Sample B Pruned with Mahalanobis	100	72.0%	65.5%	59.5%	59.7%	72.3%	72.0%	70.6%
Sample C Original Data	105	81.9%	82.2%	74.3%	74.8%	85.2%	81.9%	81.2%
Sample C Pruned with ISF	100	86.0%	86.4%	85.7%	86.0%	86.4%	86.0%	86.1%
Sample C Pruned with LOF	99	90.9%	90.0%	93.1%	91.1%	91.7%	90.9%	91.0%

Sample C Pruned with Mahalanobis	74	89.2%	82.9%	82.2%	82.4%	89.6%	89.2%	89.2%
Combined Samples Original Data	454	71.4%	76.8%	71.9%	71.4%	77.4%	71.4%	70.8%
Combined Samples Pruned with ISF	432	85.9%	86.7%	85.3%	85.9%	86.1%	85.9%	85.9%
Combined Samples Pruned with LOF	432	82.9%	83.1%	82.0%	82.4%	83.5%	82.9%	83.0%
Combined Samples Pruned with Mahalanobis	297	81.1%	78.8%	77.9%	78.3%	80.8%	81.1%	80.9%

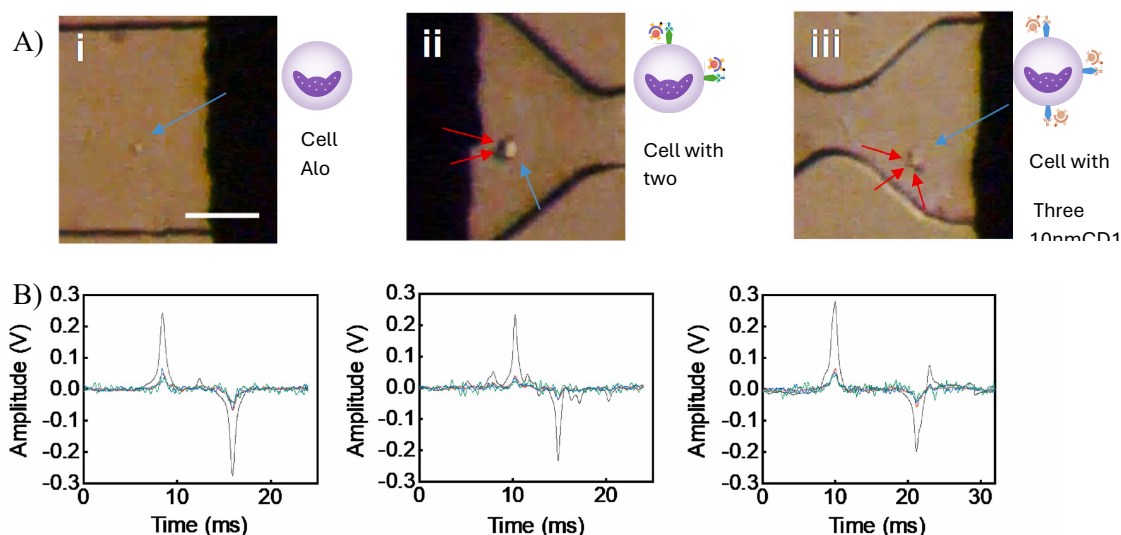


Figure S2: (A) Microscope images (scale bar = 35 μm) captured from high-speed camera video recordings for neutrophils (blue arrows) without or with conjugated MOJPs (red arrows)[1]. (B) Blood cells and (MOJP) electrical responses over four recorded voltage frequencies at 500 kHz (gray), 7.5 MHz (red), 8.3 MHz (blue), and 9 MHz (green). (B) Respective Bipolar Impedance signals [1].

References:

[1] B. K. Ashley, J. Sui, M. Javanmard, and U. Hassan, "Multi-modal sensing with integrated machine learning to differentiate specific leukocytes targeted by electrically sensitive hybrid particles," *Biosensors and Bioelectronics*, vol. 241, p.115661, 2023. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0956566323006036>