


Application of k-Nearest Neighbors Method for Drug Concentration and Cardiotoxicity Classification Using Extracellular Field Potentials and Reconstructed Action Potentials of Cardiac Cells

M. O Shpotak^f,  [0000-0002-4706-7603](https://orcid.org/0000-0002-4706-7603)

N. H. Ivanushkina^s, PhD Assoc.Prof.,  [0000-0001-8389-7906](https://orcid.org/0000-0001-8389-7906)

Department of Electronic Engineering

National technical university of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute"  [00syn5v21](https://ror.org/00syn5v21)
Kyiv, Ukraine

Abstract—Micro-electrode array (MEA) systems are important for measuring extracellular field potentials (FP) of cardiac cells, which is a crucial step in cardiotoxicity assessment. However, without modification, the MEA system is only capable of recording FPs. This limits the number of parameters for cardiotoxicity assessment only to FP parameters, while the action potential (AP) parameters remain unused. To address this issue the MEA systems are often modified to use electro-optoporation to record the local extracellular APs (LEAPs), which allows to reliably quantify the AP morphology. As an alternative to MEA modification and cell membrane stimulation the AP can be reconstructed mathematically. This study explores how using additional parameters from reconstructed action potentials (RAPs), derived from FPs, can improve the accuracy of k-NN machine learning models for drug concentration and potential cardiotoxicity classification. The k-NN classifier was trained using combinations of FP and RAP parameters. The k-NN models were evaluated using five-fold stratified cross-validation and cross-channel validation. Their performances were compared using error rate, macro precision, macro recall and macro F1 score accuracy metrics. The results indicated that incorporating RAP parameters into the feature set increased the F1 score of k-NN model for DMSO concentration classification by up to 10.78% compared to the training set with only FP features.

Keywords — extracellular field potentials; reconstructed action potentials; machine learning; k-nearest neighbours; cardiotoxicity; classification; feature selection.

I. INTRODUCTION

Cardiac electrophysiology has traditionally relied on manual interpretations of data, often leading to time-consuming and potentially subjective analyses [1]. The introduction of Lab-on-a-Chip (LOC) and multi-electrode array (MEA) technologies allowed to record detailed extracellular field potential (FP) data from multiple cardiac cells with high throughput as opposed to patch-clamp technology, which recorded data from a single cell [2]. The large amounts of high-dimensional data may require an alternative data processing approach [3].

Cardiotoxicity assessment, a cornerstone in drug safety evaluation, requires precise and efficient analysis methodologies to navigate the extensive array of

features these advanced technologies present. Machine learning (ML) models can process and interpret large datasets generated by LOC and MEA systems, encompassing a multitude of features that describe the FPs of cardiac cells [4]. The number of features can range from one to dozens, including various parameters of FPs such as amplitudes, durations and frequencies. ML models can be designed to focus either on the specific features or on the patterns of FP cycles [5], providing a comprehensive view of cardiac responses to pharmacological agents.

The integration of ML in cardiotoxicity assessment presents several advantages. Firstly, it automates the analysis process, significantly reducing the time and labor involved in manual interpretations [6]. Secondly,



these models can be improved over time with exposure to more data, enhancing their predictive accuracy.

There are multiple approaches in cardiotoxicity classification using MEA systems with different benefits and disadvantages. In some papers, researchers use manually determined parameters like FP duration (FPD), peak FP amplitude (FPA) or short-term variability (STV) to classify the cardiotoxicity risk in recorded signals. This approach requires signal processing to determine the parameters and an expert to separate them into classes, but it also allows to employ more simple and interpretable ML algorithms like k-Nearest Neighbours (k-NN), Support Vector Machines or Random Forest [7]. The downside of the parameter-based classification is that different data is sensitive to different parameters. This problem requires additional data processing step in feature selection to choose the best possible features and discard the insignificant ones.

Cardiotoxicity assessment with MEA systems can be performed using only FP recordings, but over time researchers developed techniques to acquire other types of electrophysiological data like the local extracellular action potentials (LEAP) that could be used for high throughput cardiac AP measurements [8]. Approaches like this use electro- [9] or optoporation [10] to increase the permittivity of the cell membrane. This allows the extracellular electrode to record the potential from the merged extracellular and intracellular spaces. However, these methods require modification of MEA systems with a laser for optoporation and with stimulating electrodes for electroporation. As an alternative to invasive methods the AP can also be reconstructed mathematically from the extracellular FPs recorded with unmodified MEA systems. The reconstructed APs (RAPs) not only allows to expand the number of parameters for cardiotoxicity assessment [11], but also can be used to expand the number of features for ML applications.

This study focuses on integrating RAPs and their parameters into ML models to automate drug concentration and cardiotoxicity assessment, with potential applications in LOC and MEA systems. The choice of the k-NN model in this paper comes from its ease of implementation compared to other models and its ability to adapt well to diverse data distributions without imposing strong assumptions.

II. MATERIALS AND METHODS

The dataset for this study comprised FP recordings of human-induced pluripotent stem cell-derived cardiomyocytes (hiPSC-CMs), obtained using a micro electro coaxial guide (uECG) system, a platform for monitoring cardiac electrophysiology [12]. The dataset encompassed the effects of Dimethylsulfoxide (DMSO) and Sotalol on cardiac cells. Signals for the DMSO-treated group were recorded using 2 channels; each channel recorded 7 signals: one from control group and 6 with different

concentrations of DMSO (from 0.1 to 0.6%). The signals with the concentration of 0.2% DMSO were discarded because of significant amount of artifacts, resulting in total of 6 signals per 2 channels. In contrast, the Sotalol dataset had 1 control group signal and 6 signals with different drug concentrations (1, 3, 7.5, 15, 30 and 60 nM), but each recording was captured from a single electrode.

Data preparation involved processing of FP recordings to ensure applicability of the dataset for ML training, feature extraction and class categorization. Processing steps involved identification and removal of artefacts, filtering of 50 Hz power supply noise using a complex filter based on wavelet denoising and PCA [11]. Feature extraction involved period separation and quantification of key FP parameters, such as interspike interval (ISI), FP duration (FPD) and maximum FP amplitude (FPA). Additionally, FP recordings with multiple channels (DMSO-treated group) were used to mathematically reconstruct the AP using the approach from [11] allowing to consider such parameters as RAP durations at 50, 70 and 90% of the repolarization (RAPD50, RAPD70 and RAPD90) for model training (Fig. 1).

For the cardiotoxicity assessment, both *in vitro* studies and clinical trials rely on electrophysiological parameters to provide biophysical information about cardiac cells. The parameters measured in *in vitro* experiments using patch-clamp or MEA systems are associated with corresponding parameters in the clinical Electrocardiogram (ECG):

- Action Potential Duration (APD): in cardiac cells, the duration of the AP is a time period during which the cell's AP goes through the phases of depolarization and repolarization. APD is associated with the QT interval observed in ECGs. The QT interval represents the time taken for ventricular depolarization and repolarization. Prolongation or shortening of the QT interval is a key marker for the assessment of drug-induced

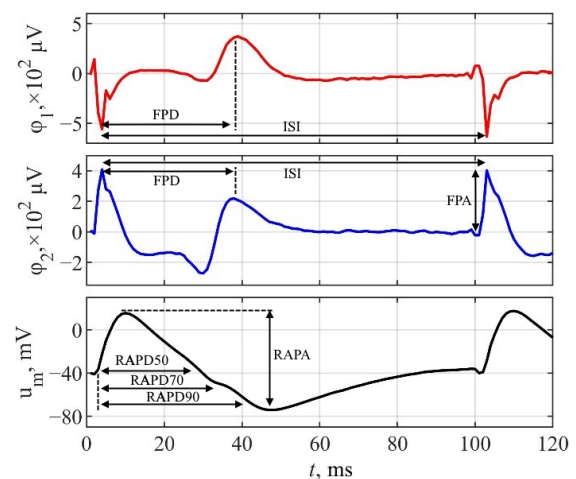


Fig. 1 Waveforms of FPs (ϕ_1 and ϕ_2), RAP (u_m) and their parameters

cardiotoxicity, indicating potential risks for arrhythmia [13].

- Field Potential Duration (FPD): in *in vitro* cardiotoxicity assays like MEA systems, FPD has high correlation with APD [14]. And just like with APD, the prolongation or shortening of FPD, is associated with prolongation or shortening of the QT interval.
- Interspike Interval (ISI): in *in vitro* experiments the ISI measures the time between consecutive cardiac depolarization events, just like the RR interval in ECGs, which denotes the time between successive heartbeats. Alterations in the RR interval (and ISI) can indicate changes in heart rate and rhythm, providing essential information on the compound's effect on cardiac pacing and rhythm stability [15].

Each parameter measured in cardiac electrophysiology studies has significance in assessing cardiac function and potential toxicity. In [16] DMSO was studied using parameters from both FP and AP, such as Resting Membrane Potential (RMP), peak AP amplitude (APA), APD and corrected FPD (FPD_c) using Fredericka's formula:

$$FPD_c = \frac{FPD}{ISI^{1/3}}.$$

The study [16] found that concentrations of DMSO greater than 1% resulted in significant changes in these electrophysiological parameters. At concentrations of 0.3% DMSO, the waveforms of both AP and FP were recorded as irregular.

In [17] the assessment of Sotalol's cardiac toxicity was focused mainly on FPD and the signals' shape. The study found that with increasing concentrations of Sotalol, there was a prolongation of the repolarization wave peak, which corresponds to the prolongation of ECG T-wave. The prolongation of the repolarization wave (or T-wave prolongation) is indicative of changes in the cardiac repolarization process. Prolonged repolarization can lead to arrhythmias and is a critical factor in evaluating drug-induced QT interval prolongation, a well-known marker of cardiotoxicity. It was also noted that the offset point of the FPD remained similar across different Sotalol concentrations. This observation suggests that while Sotalol affects the duration of repolarization, it does not significantly alter the overall cycle length of the cardiac action potential. These conclusions marked FPD as an important parameter for cardiotoxicity assessment of Sotalol.

The categorization into "high risk" and "low risk" cardiotoxicity groups for DMSO was achieved by using threshold values for FPD_c, RAP durations (RAPD) and their STVs, based on established cardiotoxicity criteria and the observed effects of the drugs at different concentrations [18]. Recordings that showed significant

deviations from the control group values, such as significant prolongations or shortening in FPD_c or RAPD (±10%), or substantial variability between subsequent values (±90%), were classified into the 'high risk' group. Conversely, recordings with parameters within the normal ranges were categorized as 'low risk'.

k-NN is a classic method in supervised ML which depends on the number of neighbouring datapoints (k), the distance measure, and the distance weighting system. For this study, the Euclidean distance was chosen as the preferred metric for the numerical experiments, based on its demonstrated high accuracy in similar applications in [19]. The training set was based on a selection of manually identified parameters from the FP and AP signals. Permutation feature importance was employed to evaluate the impact of different features on k-NN's accuracy, demonstrating the potential redundancy in overly extensive feature sets. Therefore, only the most influential RAPD parameters were selected and combined into differences RAPD₉₀₋₅₀ and RAPD₉₀₋₇₀, which represent the specific phases of the RAP repolarization.

The performance of models trained on the datasets was assessed using established metrics of accuracy, macro specificity, macro sensitivity and macro F1-score:

- Error Rate is calculated as the ratio of incorrectly predicted observations to the total instances in the validation set:

$$Error\ Rate = \frac{FP + FN}{TP + TN + FP + FN},$$

where TP – True Positives, TN – True Negatives, FP – False Positives, FN – False Negatives.

- Macro precision is calculated by averaging the precision of each class. Precision for each class is the ratio of true positives (correctly predicted positive observations) to the total predicted positives (both true positives and false positives) for that class:

$$Macro\ Precision = \frac{1}{N} \sum_{i=1}^N Precision_i,$$

where N is the number of classes and

$$Precision_i = \frac{TP_i}{TP_i + FP_i}.$$

- Macro recall is the average recall calculated separately for each class. Recall for each class is the ratio of true positives to the actual positives (true positives plus false negatives) for that class:

$$Macro\ Recall = \frac{1}{N} \sum_{i=1}^N Recall_i,$$

where $Recall_i = \frac{TP_i}{TP_i + FN_i}$.



- F1 score is a harmonic mean of precision and recall, it computes how many times a model made a correct prediction across the entire validation set. Macro F1 score averages all per-class F1 scores:

$$\text{Macro F1 Score} = \frac{1}{N} \sum_{i=1}^N F1_i$$

$$F1 \text{ Score}_i = 2 \cdot \frac{\text{Recall}_i \times \text{Precision}_i}{\text{Recall}_i + \text{Precision}_i}$$

where

The high recall values indicate the models' capability to correctly identify true instances of specific class, a critical aspect in the context of cardiotoxicity assessment where missing a true instance can have significant implications. On the other hand, the precision values reflect the models' effectiveness in accurately classifying instances into their respective classes, ensuring that the predictions are reliable and minimizing the risk of false alarms.

III. RESULTS

Fig. 2 shows categorization of DMSO into cardiotoxicity risk groups based on normalized FPD_c and STV_{FPD_c} thresholds, DMSO is widely known to be a non-toxic compound [20], but the data shows that at the highest concentration of 0.6%, when the FPD_c and STV_{FPD_c} cross the thresholds the waveform contains FP alterations, which may be a marker of cardiotoxicity. On the other hand, Sotalol is a known antiarrhythmic agent [21], it is used in treating certain types of cardiac arrhythmias, like atrial fibrillation [22] and ventricular tachycardia [23], but its capacity to prolong the QT interval at higher doses can increase the risk of a potentially life-threatening arrhythmia called Torsades de Pointes; additionally, as a beta-blocker, Sotalol can have other cardiac effects such as bradycardia. Comparison of FPD_c and STV_{FPD_c} data for Sotalol shows that both FPD_c and STV_{FPD_c} cross the cardiotoxicity risk threshold at the highest drug concentrations of 30 and 60 nM.

Reconstructing the waveform of the APs from FP recordings affected by DMSO compound using approach proposed in [11] allowed to determine parameters like $RAPD_{50}$, $RAPD_{70}$, $RAPD_{90}$ and their differences $RAPD_{90-50}$ and $RAPD_{90-70}$, which allowed to further evaluate the effect of DMSO on the electrophysiology of cardiac cells. Fig. 4 supplements Fig. 1 with information about the duration of a more specific phase of the RAP that significantly deviates from the control group.

An evaluation was conducted on the performance of k-NN models for classifying drug concentrations with various combinations of electrophysiological features (Fig. 5). Assessment metrics included Error Rate, Macro Precision, Macro Recall, and Macro F1 score. Results for DMSO and Sotalol drug concentration classification using

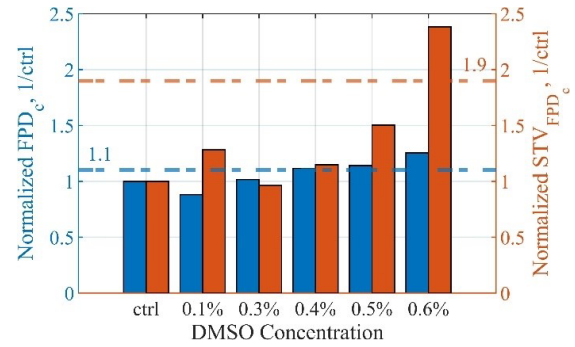


Fig. 2 Comparison of normalized FPD_c and its STV_{FPD_c} for different DMSO concentrations, where color-coded dashed lines represent high cardiotoxicity risk threshold for the respective parameters

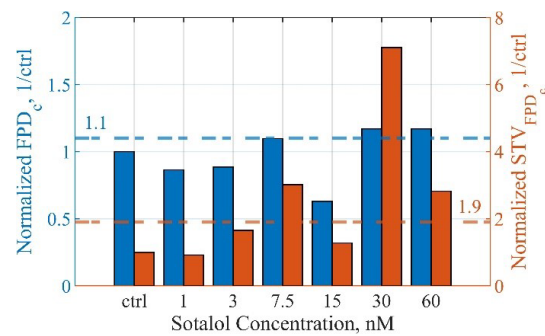


Fig. 3 Comparison of normalized FPD_c and its STV_{FPD_c} for different Sotalol concentrations, where color-coded dashed lines represent high cardiotoxicity risk threshold for the respective parameters

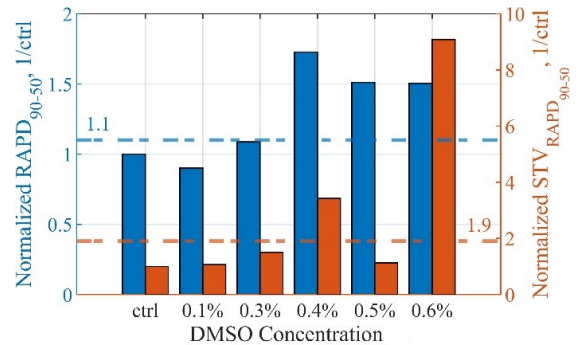


Fig. 4 Comparison of normalized $RAPD_{90-50}$ and its $STV_{RAPD_{90-50}}$ for different DMSO concentrations, where color-coded dashed lines represent high cardiotoxicity risk threshold for the respective parameters

different feature combinations are presented in Table 1 and Table 2.

For DMSO concentration classification the number of neighbors $k = 15$ and for Sotalol concentration classification $k = 8$. The k values were chosen based on the size of the dataset and accuracy evaluation.

For DMSO concentration classification the $RAPD_{90-50}$ was selected because it showed the highest accuracy among other RAPD features. The dataset was balanced to represent all classes equally. The model was validated

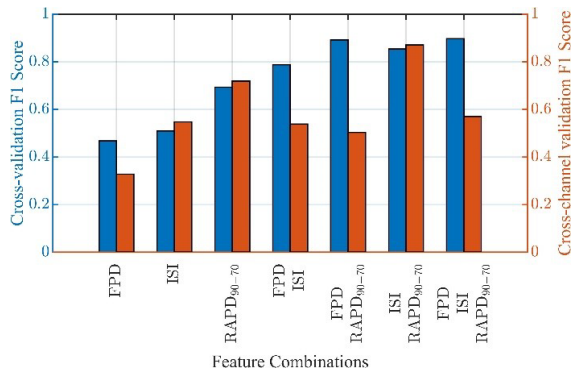


Fig. 5 Macro F1 score from 5-fold cross validation and cross-channel validation for different combinations of FP and RAP features for k-NN model for DMSO concentration classification

using 5-fold cross-validation and cross-channel validation. In cross-channel validation the model was trained on the dataset based on the first channel using 5-fold cross-validation and then each fold was validated individually using the dataset based on the second channel.

For Sotalol concentration classification no AP could be reconstructed, so only FP features were used for classification.

For cardiotoxicity risk classification the classes were formed using threshold information from Fig. 2 and Fig. 3. The highest concentration of DMSO (0.6%) was labeled as having high risk of cardiotoxicity, while the rest of the signals were labeled as having low risk of cardiotoxicity. The dataset was balanced using stratified cross-validation to represent all variations of features from signals with different DMSO concentration. The number of neighbors $k=11$ according to the dataset size and accuracy evaluation. The model was validated using stratified 5-fold cross-validation and cross-channel validation.

Table 3 shows the metrics of the k-NN model for DMSO cardiotoxicity risk classification. For this model the RAP parameter $RAPD_{90-70}$ showed the highest contribution to accuracy among the other RAP parameters.

IV. DISCUSSION

The k-NN models, with a Macro F1 score of 0.9 for drug concentration and an F1 score of 0.93 for cardiotoxicity risk, show high accuracy in their respective classifications. For drug concentration classification increasing the number of features results in increased classification accuracy. Certain features are particularly effective at distinguishing different drug concentrations. The models incorporating $RAPD_{90-50}$ show higher Macro F1 scores of 0.85, 0.88, and 0.9, compared to a score of 0.79 for the model without $RAPD_{90-50}$ (Table 1). Some other features like FPA can significantly vary in FPs recorded from different channels, which may have negative impact on the model generalizability. In 1-feature models that use FPA the k-fold cross-validation shows Macro F1 score

of 0.67, but the cross-channel validation shows the score of 0.33; cross-channel validation for models that use combinations of features and include FPA also shows a lower Macro F1 score compared to the models that exclude FPA (Fig. 5).

TABLE 1 THE 5-FOLD CROSS VALIDATION ACCURACY METRICS OF K-NN MODEL FOR DMSO CONCENTRATION CLASSIFICATION USING DIFFERENT FEATURE COMBINATIONS

Used Features	Error Rate	Macro Precision	Macro Recall	Macro F1 score
FPD	0.4657	0.4213	0.5353	0.4713
ISI	0.5176	0.5316	0.4817	0.5053
RAPD90-50	0.3265	0.7062	0.6742	0.6896
FPD, ISI	0.2029	0.7996	0.7952	0.7974
FPD, RAPD90-50	0.1176	0.8834	0.8817	0.8825
ISI, RAPD90-50	0.151	0.8595	0.8494	0.8544
FPD, ISI, RAPD90-50	0.0961	0.906	0.9045	0.9052

TABLE 2 THE 5-FOLD CROSS VALIDATION ACCURACY METRICS OF K-NN MODEL FOR SOTALOL CONCENTRATION CLASSIFICATION USING DIFFERENT FEATURE COMBINATIONS

Used Features	Error Rate	Macro Precision	Macro Recall	Macro F1 score
FPA	0.2945	0.6887	0.7084	0.6982
FPD	0.2233	0.7768	0.7496	0.7626
ISI	0.384	0.6236	0.6189	0.6212
FPA, FPD	0.0758	0.92	0.922	0.9208
FPA, ISI	0.1566	0.8626	0.8508	0.8566
FPD, ISI	0.1699	0.8343	0.8236	0.8289
FPA, FPD, ISI	0.0671	0.9475	0.9287	0.9379

TABLE 3 THE 5-FOLD CROSS VALIDATION ACCURACY METRICS OF K-NN MODEL FOR DMSO CARDIOTOXICITY RISK CLASSIFICATION USING DIFFERENT FEATURE COMBINATIONS

Used Features	Error Rate	Precision	Recall	F1 score
FPA	0.328	0.6802	0.6737	0.6769
FPDc	0.12	0.8852	0.879	0.8821
RAPD90-70	0.238	0.7774	0.7626	0.7699
FPA, FPDc	0.12	0.8842	0.8807	0.8824
FPA, RAPD90-70	0.186	0.8287	0.8139	0.8212
FPDc, RAPD90-70	0.094	0.9141	0.9095	0.9118
FPA, FPDc, RAPD90-70	0.064	0.9423	0.9356	0.9389

The differences between RAPD values, such as $RAPD_{90-70}$ underscore the specific segments of the RAP, each correlating with predominant ionic currents. For instance, the $RAPD_{90-70}$ interval is used in quantifying the modulation of I_{kr} and I_{ks} currents. If a drug does not specifically target these currents, the variation in $RAPD_{90-70}$ might be minimal. Consequently, including $RAPD_{90-70}$ as a feature in the model would have a limited impact on enhancing its accuracy.

The inclusion of the RAP feature results in a marginal improvement in the F1 score for the DMSO cardiotoxicity risk classification model, increasing from 0.88 to 0.93 (Table 3). For the DMSO concentration classification model, using the RAP feature leads to a more significant improvement in the Macro F1 score, from 0.79 to 0.9 (Table 1), compared to using only FP features. It is

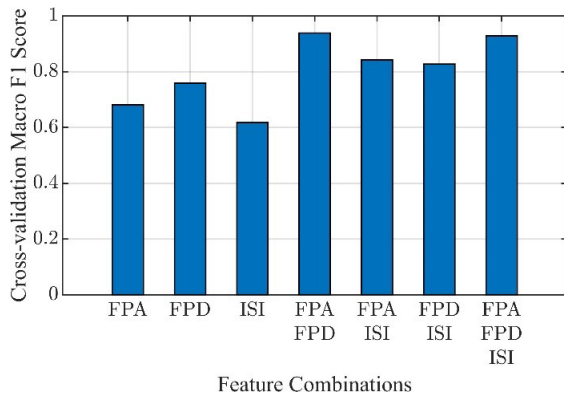


Fig. 6 Macro F1 score from 5-fold cross validation for different combinations of FP features for k-NN model for Sotalol concentration classification

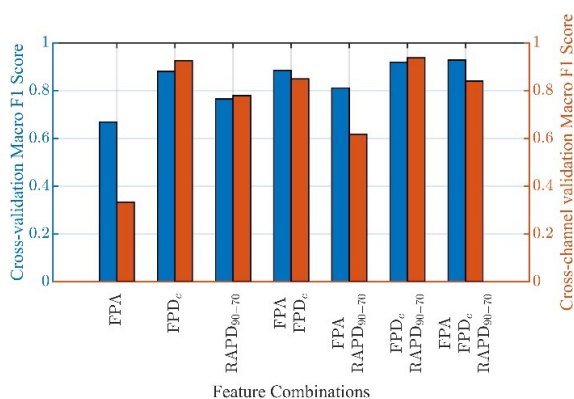


Fig. 7 Macro F1 score from 5-fold cross validation and cross-channel validation for different combinations of FP and RAP features for k-NN model for DMSO cardiotoxicity risk classification

important to note that the cross-channel validation might show a bias towards models that incorporate the RAP feature, potentially displaying enhanced accuracy. This bias could be attributed to the RAP features' integration of data from both channel 1 and 2 during the reconstruction process.

The paper's focus on particular drug concentrations (DMSO and Sotalol) and electrophysiological features may limit generalizability across different pharmacological compounds. Future research could extend these findings to a broader spectrum of drugs and explore the integration of additional FP and RAP related features that capture different aspects of drug-cell interactions, along with experimenting with other models and their hyperparameters to further refine the concentration and cardiotoxicity risk classification performance.

CONCLUSIONS

The investigation into using additional parameters from reconstructed action potentials (RAPs) derived from field potentials (FPs) for drug concentration and cardiotoxicity risk classification with k-Nearest Neighbors (k-NN) algorithm indicates enhanced model accuracy. The F1 score for the cardiotoxicity risk model increased from 0.88 to 0.93, while the Macro F1 score for the concentration model rose from 0.79 to 0.9 compared to the models that didn't include the RAP features.

The k-NN model used a combination of manually selected parameters from FP and RAP data. The initial number of neighbors k was selected based on the dataset size. Then, the models were fine-tuned by evaluating the performance for different values of k in the vicinity of the initial guess and selecting the k that showed the highest accuracy.

The accuracy of the DMSO concentration classification model was evaluated through 5-fold cross-validation and cross-channel validation. The inclusion of RAP parameters, specifically RAP duration (RAPD), in the feature set resulted in improvements of F1 score up to 10.78% compared to using only FP features. However, the influence of RAP parameters such as RAPD, and differences like RAPD₉₀₋₇₀ and RAPD₉₀₋₅₀, as features is contingent upon how the drug affects the phases they represent.

Training ML models for classification tasks, such as determining drug concentrations or assessing cardiac toxicity, necessitates a significant amount of data, including FP or AP recordings. The use of RAPs to expand the dataset features offers a viable alternative for MEA systems lacking electro- or optoporation capabilities. The proposed AP reconstruction approach enables the application of RAPs in various ML models like k-NN, which rely on fixed parameters for features. Incorporation of RAPs into ML algorithms offers an approach for assisting in the automation of cardiotoxicity assessment in MEA and Lab-on-a-Chip systems, which serves as a contribution to more efficient drug safety evaluations.

ACKNOWLEDGMENTS

The authors would like to express their gratitude to the researchers at Politecnico di Milano, especially A. Redaelli and R. Visone, for sharing the recorded FP signals from uECG device [12].

REFERENCES

- [1] K. W. Johnson et al., "Artificial Intelligence in Cardiology", *J. Amer. College Cardiol.*, vol. 71, no. 23, pp. 2668–2679, Jun. 2018. DOI: [10.1016/j.jacc.2018.03.521](https://doi.org/10.1016/j.jacc.2018.03.521)
- [2] P. P. Kanade et al., "MEA-integrated cantilever platform for comparison of real-time change in electrophysiology and contractility of cardiomyocytes to drugs", *Biosens. Bioelectron.*, p. 114675, Sep. 2022. DOI: [10.1016/j.bios.2022.114675](https://doi.org/10.1016/j.bios.2022.114675)
- [3] H. Chen, O. Engkvist, Y. Wang, M. Olivecrona, and T. Blaschke, "The rise of deep learning in drug discovery", *Drug Discov. Today*, vol. 23, no. 6, pp. 1241–1250, Jun. 2018. DOI: [10.1016/j.drudis.2018.01.039](https://doi.org/10.1016/j.drudis.2018.01.039)
- [4] J. M. Rivera-Arbeláez et al., "Automated assessment of human engineered heart tissues using deep learning and template matching for segmentation and tracking", *Bioeng. & Translational Medicine*, Apr. 2023. DOI: [10.1002/btm2.10513](https://doi.org/10.1002/btm2.10513)




- [5] W. Guo et al., "Review of machine learning and deep learning models for toxicity prediction", *Exp. Biol. Medicine*, Dec. 2023. DOI: [10.1177/15353702231209421](https://doi.org/10.1177/15353702231209421)
- [6] L. Pu, M. Naderi, T. Liu, H.-C. Wu, S. Mukhopadhyay, and M. Brylinski, "eToxPred: a machine learning-based approach to estimate the toxicity of drug candidates", *BMC Pharmacol. Toxicol.*, vol. 20, no. 1, Jan. 2019. DOI: [10.1186/s40360-018-0282-6](https://doi.org/10.1186/s40360-018-0282-6)
- [7] D. Pan, B. Li, and S. Wang, "Establishment and validation of a torsade de pointes prediction model based on human iPSC-derived cardiomyocytes", *Exp. Therapeutic Medicine*, vol. 25, no. 1, Dec. 2022. DOI: [10.3892/etm.2022.11760](https://doi.org/10.3892/etm.2022.11760)
- [8] H. B. Hayes et al., "Novel method for action potential measurements from intact cardiac monolayers with multiwell microelectrode array technology", *Scientific Rep.*, vol. 9, no. 1, Aug. 2019. DOI: [10.1038/s41598-019-48174-5](https://doi.org/10.1038/s41598-019-48174-5)
- [9] B. Duckert, M. Fauvart, P. Goos, T. Stakenborg, L. Lagae, and D. Braeken, "High-definition electroporation: Precise and efficient transfection on a microelectrode array", *J. Controlled Release*, vol. 352, pp. 61–73, Dec. 2022. DOI: [10.1016/j.jconrel.2022.10.001](https://doi.org/10.1016/j.jconrel.2022.10.001)
- [10] M. Dipalo et al., "Intracellular action potential recordings from cardiomyocytes by ultrafast pulsed laser irradiation of fuzzy graphene microelectrodes", *Sci. Advances*, vol. 7, no. 15, Apr. 2021, Art. no. eabd5175. DOI: [10.1126/sciadv.abd5175](https://doi.org/10.1126/sciadv.abd5175)
- [11] N. G. Ivanushkina, K. O. Ivanko, M. O. Shpotak, and Y. V. Prokopenko, "Reconstruction of action potentials of cardiac cells from extracellular field potentials", *Radioelectronics Commun. Syst.*, vol. 65, no. 7, pp. 354–364, Jul. 2022. DOI: [10.3103/s073527222090047](https://doi.org/10.3103/s073527222090047)
- [12] R. Visone et al., "Micro-electrode channel guide (μ ECG) technology: An online method for continuous electrical recording in a human beating heart-on-chip", *Biofabrication*, vol. 13, no. 3, p. 035026, Apr. 2021. DOI: [10.1088/1758-5090/abe4c4](https://doi.org/10.1088/1758-5090/abe4c4)
- [13] M. Malik, "Drug-induced qt/ctc interval shortening: Lessons from drug-induced qt/ctc prolongation", *Drug Saf.*, vol. 39, no. 7, pp. 647–659, Mar. 2016. DOI: [10.1007/s40264-016-0411-3](https://doi.org/10.1007/s40264-016-0411-3)
- [14] L. G. J. Tertoolen, S. R. Braam, B. J. van Meer, R. Passier, and C. L. Mummery, "Interpretation of field potentials measured on a multi electrode array in pharmacological toxicity screening on primary and human pluripotent stem cell-derived cardiomyocytes", *Biochem. Biophysical Res. Commun.*, vol. 497, no. 4, pp. 1135–1141, Mar. 2018. DOI: [10.1016/j.bbrc.2017.01.151](https://doi.org/10.1016/j.bbrc.2017.01.151)
- [15] A. A. Kondratyev, J. G. C. Ponard, A. Munteanu, S. Rohr, and J. P. Kucera, "Dynamic changes of cardiac conduction during rapid pacing", *Amer. J. Physiol.-Heart Circulatory Physiology*, vol. 292, no. 4, pp. H1796–H1811, Apr. 2007. DOI: [10.1152/ajpheart.00784.2006](https://doi.org/10.1152/ajpheart.00784.2006)
- [16] S.-W. Hyun, B.-R. Kim, S.-A. Hyun, and J.-W. Seo, "The assessment of electrophysiological activity in human-induced pluripotent stem cell-derived cardiomyocytes exposed to dimethyl sulfoxide and ethanol by manual patch clamp and multi-electrode array system", *J. Pharmacolog. Toxicolog. Methods*, vol. 87, pp. 93–98, Sep. 2017. DOI: [10.1016/j.vascn.2017.03.003](https://doi.org/10.1016/j.vascn.2017.03.003)
- [17] P. Pradhapan, J. Kuusela, J. Viik, K. Aalto-Setälä, and J. Hyttinen, "Cardiomyocyte MEA Data Analysis (CardioMDA) – A Novel Field Potential Data Analysis Software for Pluripotent Stem Cell Derived Cardiomyocytes", *PLoS ONE*, vol. 8, no. 9, Sep. 2013, Art. no. e73637. DOI: [10.1371/journal.pone.0073637](https://doi.org/10.1371/journal.pone.0073637)
- [18] T. Kaneko et al., "On-chip in vitro cell-network pre-clinical cardiac toxicity using spatiotemporal human cardiomyocyte measurement on a chip", *Scientific Rep.*, vol. 4, no. 1, Apr. 2014. DOI: [10.1038/srep04670](https://doi.org/10.1038/srep04670)
- [19] I. Saini, D. Singh, and A. Khosla, "QRS detection using K-Nearest Neighbor algorithm (K-NN) and evaluation on standard ECG databases", *J. Adv. Res.*, vol. 4, no. 4, pp. 331–344, Jul. 2013. DOI: [10.1016/j.jare.2012.05.007](https://doi.org/10.1016/j.jare.2012.05.007)
- [20] J. Galvao, B. Davis, M. Tilley, E. Normando, M. R. Duchon, and M. F. Cordeiro, "Unexpected low-dose toxicity of the universal solvent DMSO", *FASEB J.*, vol. 28, no. 3, pp. 1317–1330, Dec. 2013. DOI: [10.1096/fj.13-235440](https://doi.org/10.1096/fj.13-235440)
- [21] N. Augustin, C. Alvarez, and J. Kluger, "The Arrhythmogenicity of Sotalol and its Role in Heart Failure: A Literature Review", *J. Cardiovascular Pharmacol.*, Publish Ahead of Print, May 2023. DOI: [10.1097/FJC.0000000000001439](https://doi.org/10.1097/FJC.0000000000001439)
- [22] H. Lenhoff, H. Jarnbert-Petersson, B. Darpo, P. Tornvall, and M. Frick, "Mortality and ventricular arrhythmias in patients on d,l-sotalol for rhythm control of atrial fibrillation - A nationwide cohort study", *Heart Rhythm*, Aug. 2023. DOI: [10.1016/j.hrthm.2023.08.019](https://doi.org/10.1016/j.hrthm.2023.08.019)
- [23] J. Larson, L. Rich, A. Deshmukh, E. C. Judge, and J. J. Liang, "Pharmacologic Management for Ventricular Arrhythmias: Overview of Anti-Arrhythmic Drugs", *J. Clin. Medicine*, vol. 11, no. 11, p. 3233, Jun. 2022. DOI: [10.3390/jcm11113233](https://doi.org/10.3390/jcm11113233)

Надійшла до редакції 12 січня 2024 року

Прийнята до друку 11 березня 2024 року



Використання методу k-найближчих сусідів для класифікації концентрації препаратів та ризику кардіотоксичності з використанням потенціалів позаклітинних полів та реконструйованих потенціалів дії серцевих клітин

М. О. Шпотак^f,  [0000-0002-4706-7603](https://orcid.org/0000-0002-4706-7603)

Н. Г. Іванушкіна^s, к.т.н. доцент,  [0000-0001-8389-7906](https://orcid.org/0000-0001-8389-7906)

Кафедра електронної інженерії

Національний технічний університет України

"Київський політехнічний інститут імені Ігоря Сікорського"  [00syn5v21](https://ror.org/00syn5v21)

Київ, Україна

Анотація—Системи з мікроелектродними решітками (МЕР) важливі для вимірювання позаклітинних потенціалів поля (ПП) клітин серця, що є важливим кроком в оцінці кардіотоксичності. Однак, без модифікації БЕР система здатна реєструвати лише потенціали поля. Це обмежує кількість параметрів для оцінки кардіотоксичності лише параметрами ПП, в той час як параметри потенціалу дії (ПД) залишаються невикористаними. Для вирішення цієї проблеми БЕР системи модифікують, щоб використовувати електро- або оптопорацію для реєстрації локальних позаклітинних потенціалів дії (ЛППД), що дозволяє отримувати сигнали з достовірною морфологією ПД. З іншого боку, існує альтернатива модифікації МЕР систем, що дозволяє уникнути стимуляції клітинної мембрани — математична реконструкція ПД.

У цьому дослідженні вивчається, як використання додаткових параметрів реконструйованих потенціалів дії (РПД), отриманих з ПП, може підвищити точність таких моделей машинного навчання як k-найближчих сусідів (k-NN) для класифікації концентрацій лікарських препаратів та ризику їхньої кардіотоксичності.

Класифікатор k-NN було натреновано на комбінаціях параметрів ПП та РПД. Перевірка моделей була проведена за допомогою p-залежної перехресної валідації та міжканальної валідації. Якість k-NN моделей була оцінена за допомогою таких метрик точності як частота помилок, макро влучність, макро повнота та макро F1-міра.

Результати показали, що включення РПД параметрів до набору ознак підвищило F1-міру моделі k-NN для класифікації концентрації Dimethylsulfoxide (DMSO) до 10.78% порівняно з моделями, які були натреновані виключно на ознаках з ПП.

Ключові слова — позаклітинні потенціали поля; реконструйовані потенціали дії; машинне навчання; k-найближчих сусідів; кардіотоксичність; класифікація; вибір ознак.

