

1 PREPARED FOR SUBMISSION TO JINST
2 25TH INTERNATIONAL WORKSHOP ON RADIATION IMAGING DETECTORS
3 30. 06. - 04. 07. 2024
4 LISBON

5 Machine learning models for single-particle 6 classification with Timepix 3 detector

7 K. Sykorova,^{a,1} Z. Buk,^b M. Cepek,^c C. Granja,^a C. Oancea,^a L. Marek^a

8 ^aADVACAM s.r.o., Prague, Czech Republic

9 ^bDepartment of Theoretical Computer Science, Czech Technical University, Prague, Czech Republic

10 ^cDepartment of Applied Mathematics, Czech Technical University, Prague, Czech Republic

11 E-mail: katerina.sykorova@advacam.cz

12 ABSTRACT: Semiconductor hybrid pixel detectors with Timepix 3 chips developed by Medipix col-
13 laboration at CERN can simultaneously measure deposited energy and time of arrival of individual
14 particle hits in all 256×256 pixels with $55\mu m$ pitch size. Leveraging the single-particle detection
15 sensitivity of these chips, there is a potential to develop algorithms for classifying detected single
16 particles into distinct categories corresponding to different particle types. In this study, various
17 machine learning models are introduced, such as neural networks or gradient boosted decision
18 trees, designed to facilitate the classification of single particle detections. These models are trained
19 and tested on an extensive database of experimental data obtained from controlled radiation source
20 experiments, allowing for robust performance across various scenarios.

21 KEYWORDS: Particle tracking detectors, Particle identification methods

22 ARXIV EPRINT: [1234.56789](https://arxiv.org/abs/1234.56789)

¹Corresponding author.

23 Contents

24	1 Introduction	1
25	2 Particle detection	1
26	2.1 Timepix 3 detector	1
27	2.2 Particle clusters	2
28	3 Single particle event classification	2
29	3.1 Machine learning algorithms	3
30	3.2 Results and model comparison	4
31	4 Electron - photon classification	6
32	4.1 Unsupervised machine learning technique	6
33	4.2 Distribution interpretation and results	6
34	5 Conclusions and outlook	7

35 1 Introduction

36 Classifying single particle events into distinct classes associated with different particle types is
37 essential for tasks like improving the signal-to-noise ratio in physical analyses or assessing con-
38 tributions from various radiation sources. For instance, in dosimetry, distinct components of a
39 radiation field can produce varying biological effects or impacts on electronic systems.

40 Diverse hardware and software techniques exist for detecting and distinguishing between types
41 of particle radiation, each offering different spatial and temporal resolutions, energy ranges, and
42 versatility. This study utilizes the single-particle detection capabilities of semiconductor hybrid
43 pixel detectors equipped with Timepix 3 ASIC chips [1], which can measure different radiation
44 field components. Machine learning algorithms are then applied to classify these measured single-
45 particle events.

46 2 Particle detection

47 2.1 Timepix 3 detector

48 The Timepix 3 ASIC chips, based on CMOS technology, were developed by the Medipix collab-
49 oration at CERN [1]. Each chip is highly segmented, containing 256×256 pixels with a pitch
50 size of $55 \mu\text{m}$, resulting in a sensitive area of $14 \times 14 \text{ mm}^2$ per chip. Timepix 3 chips enable the
51 simultaneous measurement of deposited energy (measured as Time over Threshold, or ToT) and the
52 time of signal arrival (ToA) for each pixel. Data can be read using either frame readout, where the

entire pixel matrix is read at once, or data-driven readout, where data is transmitted immediately after a pixel is hit, while the rest of the chip remains sensitive.

The machine learning models for single-particle classification presented here were built for AdvaPiX/MiniPix Timepix 3 devices [2], developed and produced by Advacam s.r.o, equipped with a $500\text{ }\mu\text{m}$ thick silicon sensor and the Timepix 3 ASIC chip. The entire pixel matrix is assumed to be calibrated, with a minimum energy threshold set to 3 keV for energy measured in a single pixel, and the bias voltage fixed at 80 V .

2.2 Particle clusters

Semiconductor hybrid pixel detectors with Timepix 3 ASIC chips can detect various types of ionizing radiation, including photons, electrons (and positrons), protons, alpha particles, and other heavy ions. When a particle interacts with the sensor, it is immediately converted into an electrical signal. Typically, this interaction produces a signal spread across multiple pixels, which can be combined into a cluster representing the detected particle event. Since different types of particles interact with the sensor through different physical processes, they produce clusters with distinct features, such as shape and energy distribution within the cluster. Figure 1 shows a selection of clusters corresponding to different particle types and energies.

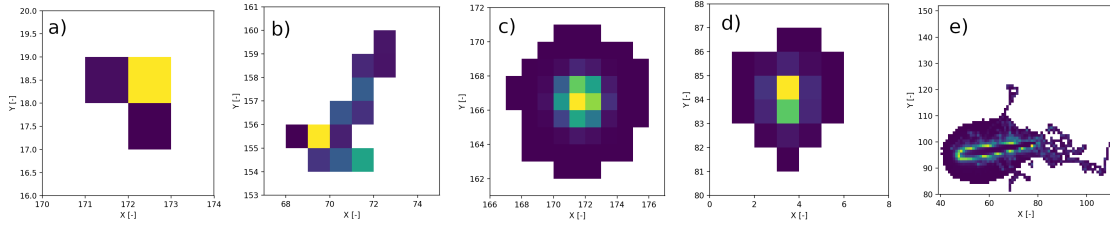


Figure 1: Clusters corresponding to different particle types: a) X-ray with energy 160 keV , b) electron with kinetic energy 21 MeV , c) alpha particle with kinetic energy 3.5 MeV , d) proton with kinetic energy 70 MeV , e) iron nuclei with kinetic energy 1 GeV per nucleon.

3 Single particle event classification

Morphological characteristics and spectral parameters of clusters form a set of cluster features which can be used to develop an algorithm that classifies detected radiation into classes associated with different particle types. In creating this algorithm, the experience gained from a previously developed classification algorithm, which categorized particles into three classes (protons, electrons & photons, and ions) as introduced in ref. [3]. Our goal was to further improve and expand upon these concepts.

Clusters corresponding to single particle events were classified into 5 distinct classes called **electrons & photons**, **alpha particles**, **heavy nuclei** (except alpha), **low energy protons** ($E \lesssim 100\text{ MeV}$) and **high energy protons** ($E \gtrsim 100\text{ MeV}$). Several machine learning methods were employed to facilitate particle event classification:

- Bidirectional Long Short-Term Memory recurrent Neural Network (BiLSTM),

- Fully connected Neural Network (FNN),
- Gradient boosted Decision Trees (CatBoost).

The models are detailed in the following section 3.1. The first method analyzes the time-ordered sequence of pixels within clusters, the latter two evaluate various cluster features, such as size, linearity, roundness, average energy deposited per pixel, etc. The feature dependent approaches thus require expert knowledge and the additional preprocessing step.

The models were trained using a database of reference experimental (calibration) data, summarized in Table 1. All datasets were manually cleaned to remove unwanted particle background. For each class, the training dataset was evenly composed of the available experimental data, and dataset weighting was applied to account for the varying sizes of data across different classes.

Table 1: Summary of experimental datasets used for training machine learning classification algorithms.

Class label	Radiation source	Approx. Source Energy [<i>MeV</i>]
electrons & photons	^{90}Sr decay	0.549, 2.3
	electron	8, 16, 21
	^{60}Co decay	0.31, 1.17, 1.33
	X-ray	0.033, 0.049, 0.083, 0.160
alpha particles	He	20, 38
	^{241}Am decay	2.2 – 5
heavy nuclei	^{52}Fe	230, 350, 1000 per nucleon
	^{28}Si	490 per nucleon
low energy protons	proton	8, 16, 26, 34, 70
high energy protons	proton	150, 225

3.1 Machine learning algorithms

BiLSTM recurrent Neural Network

A bidirectional Long Short-Term Memory (BiLSTM) network [4] is used for processing sequences by leveraging two LSTM layers: one reads the sequence from start to end, and the other from end to start, capturing both past and future dependencies. In this application, the input is a sequence representing the time-ordered pixel trajectory of a particle cluster. Each entry in the sequence consists of four parameters: pixel coordinates, signal ToA, and deposited energy. The BiLSTM processes this multidimensional sequence to classify it into one of five classes, considering the full context of the particle’s trajectory and behavior over time.

The advantage of this method is that it requires no preprocessing of the sequence, such as calculating additional cluster features or relying on expert knowledge. The BiLSTM can directly learn relevant features from the raw data, capturing both spatial and temporal patterns automatically.

Fully connected Neural Network

The fully connected neural network (FNN) [5] (feedforward network [6]) classifies particle events based on a vector of particle cluster features, which is generated through preprocessing evaluating morphological characteristics and spectral parameters of clusters. The network then processes this input through multiple dense layers to perform classification into one of the five classes.

Gradient Boosted Decision Trees CatBoost

CatBoost is a specific application of gradient boosting approach with decision trees as base predictors. The gradient boosting is a process of iteratively constructing a sequence of predictors, in our case decision trees. More details on the algorithm is described in [7]. The particle event classification by the CatBoost model is build on top of the precalculated vectors of particle cluster features.

3.2 Results and model comparison

The model's performance was assessed during training using the validation subset of data from Table 1. From this process, the best performing Gradient Boosted Decision Tree, Bi-LSTM Recurrent Neural Network, and Fully Connected Neural Network models were selected. The models were then tested on a separate subset, comprising 20% of the total experimental data, which was not used during training or validation. The confusion matrices for these evaluations are shown in Figure 2.

All models produced comparable and highly satisfactory results, with the ratio of true positives—correct predictions of the cluster class—being close to 100% across all classes. However, some minor patterns in misclassification were observed. Notably, in over 1% of cases, the Bi-LSTM and CatBoost tended to misclassify high-energy protons as alphas. Conversely, the FNN was more likely to misclassify alphas as high-energy protons and low-energy protons as electrons or photons, see Figure 2.

The models were further compared based on their inference times, which represent the average time required to make a single prediction. The results showed that the best inference time for the BiLSTM model was 0.14 ms, while the Fully Connected Neural Network achieved 0.013 ms, and CatBoost excelled with an inference time of just 0.0048 ms. This indicates that feature-dependent approaches enabled faster predictions. However, the BiLSTM model has the advantage of not requiring additional data preprocessing to extract cluster features.

Lastly, it's important to note that machine learning models trained on particle cluster feature vectors can help determine which cluster features are most significant. This, in turn, highlights the key morphological characteristics and spectral parameters that are most relevant for classifying single particle events. To achieve this, the SHAP technique was used to evaluate the contribution of individual features to the model's final output. This technique compares the model's predictions with and without each specific cluster feature [8].

Table 2 contains the three most important cluster features (those with the highest SHAP values) for both the Fully Connected Neural Network and CatBoost Classifier. These features include LET (Linear Energy Transfer of the particle in the detector), Height (the maximum energy deposited in a single pixel of the cluster), EpixStd (the standard deviation of the pixel energy distribution),

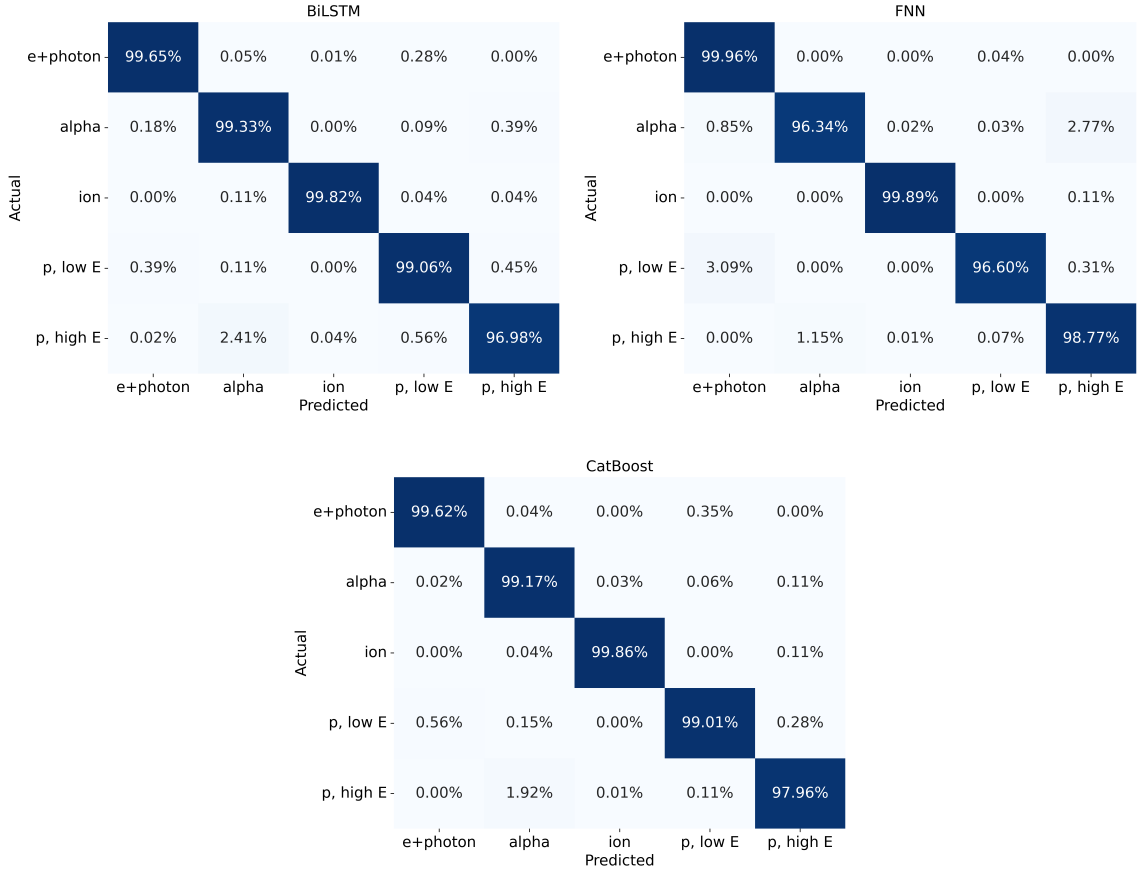


Figure 2: Confusion matrices for the top-performing Bi-LSTM Recurrent Neural Network, Fully Connected Neural Network, and Gradient Boosted Decision Tree models, evaluated on the test subset of the experimental data.

Table 2: Feature importance measured by normalized mean SHAP value over 2500 samples (500 for each class). SHAP values are normalised into range between 0 and 1 for each example and then mean value over the whole dataset is calculated. The values between Full Connected Neural Network and Catboost are not directly comparable.

(a) SHAP Values for Fully Connected Neural Network.

Feature	SHAP
LET	0.376
StdAlong	0.373
Height	0.207

(b) SHAP Values for CatBoost Classifier.

Feature	SHAP
StdAlong	0.619
EpixStd	0.588
LET	0.559

143 and StdAlong (the standard deviation of pixel positions along the cluster main axis, weighted by
 144 the deposited energy). For the last feature, each pixel center is treated as a point on a flat rigid
 145 body, weighted by the square root of its energy. The cluster axis is determined as one of the main
 146 rotational axes of this rigid body, minimizing its moment of inertia.

4 Electron - photon classification

Separating the electron and photon components of a radiation field is an intriguing challenge with potential applications, such as in electron microscopy. One of the main difficulties in distinguishing between electrons and photons arises since high-energy photons (with energies above approximately 100 keV) primarily interact with the silicon sensors through the Compton effect. This interaction produces electrons within the detector sensitive area, resulting in signals that are nearly indistinguishable from those generated by primary electrons.

4.1 Unsupervised machine learning technique

In this study, the first attempt to tackle this problem is presented. The Gaussian Mixture unsupervised machine learning technique was applied to a mixed dataset of electron and photon experimental data with varying energies, see Table 3. The Gaussian Mixture model assumed that the data consists of two Gaussian distributions, ideally associated with electrons and photons, with unknown parameters. The goal was for the algorithm to identify these two distributions (Distribution 0, Distribution 1) without any supervision. The unsupervised approach offered the machine learning model more flexibility compared to supervised algorithms, where data were labelled as “electrons” and “photons”, which struggled due to the similarity between signals from high-energy photons and electrons.

The Gaussian Mixture models were evaluated on the experimental data summarized in Table 3, and the hyperparameters were adjusted to achieve the best distinction between the electron and photon datasets.

Table 3: A summary of the electron and photon experimental data with varying energies, divided into two distributions (Distribution 0 and Distribution 1).

Radiation source	Energy [MeV]	Distribution 0	Distribution 1
^{90}Sr	0.549, 2.3 (beta)	14.0%	86.0%
e	8	0.0%	100.0%
e	16	0.0%	100.0%
e	21	0.0%	100.0%
^{60}Co	0.31 (beta); 1.17, 1.33 (gamma)	43.3%	56.7%
X-ray	0.033	99.7%	0.3%
X-ray	0.049	99.4%	0.6%
X-ray	0.083	97.3%	2.7%
X-ray	0.160	98.9%	1.1%

4.2 Distribution interpretation and results

Figure 3 presents a random sample of 500 clusters from Distribution 0 and Distribution 1, taken from experimental datasets from Table 3, along with their corresponding energy and size spectra. Distribution 0 mainly consists of small clusters, up to 4 pixels in size, with energies below approximately 200 keV. In contrast, Distribution 1 primarily includes larger, wiggly-shaped clusters with higher energies.

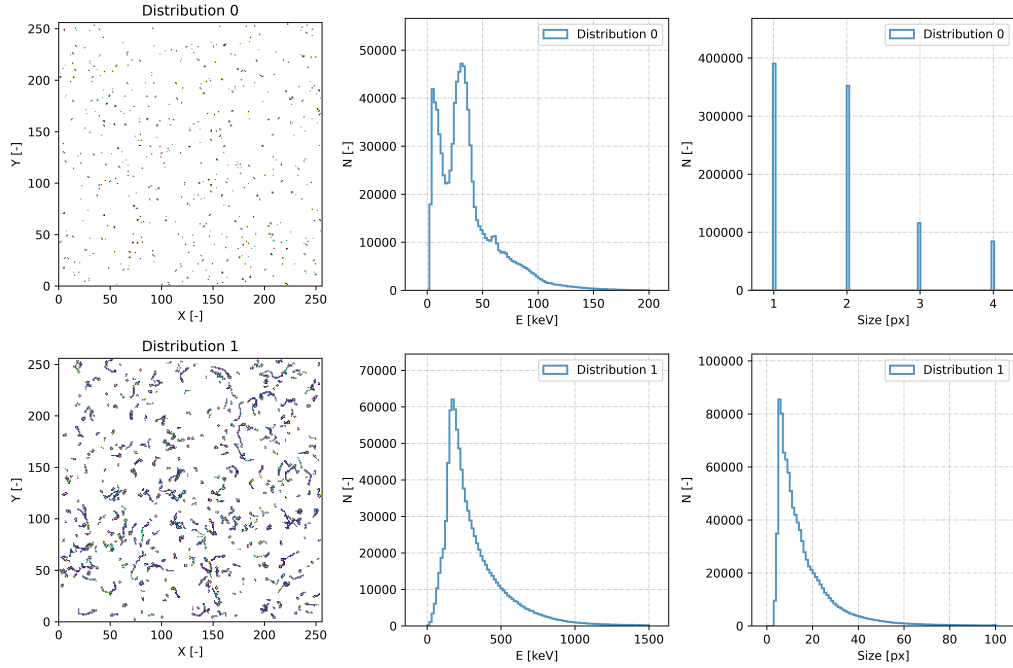


Figure 3: (Left) Random sample of 500 clusters taken from experimental datasets from Table 3, (Middle) their corresponding energy spectra and (Right) size spectra for Distribution 0 and Distribution 1.

Table 3 summarizes the performance of the most promising model on experimental datasets. All electron data with energies above approximately 5 MeV belong to Distribution 1, while nearly all X-ray photons with energies below 200 keV reside to Distribution 0. This indicates that the model effectively distinguishes between high-energy electrons ($O(10 \text{ MeV})$ energy range and above) and low-energy X-ray photons. However, the algorithm struggles in intermediate energy ranges. For example, in the case of the ^{90}Sr beta source, electrons with low deposited energy were found in Distribution 0, resulting in 14% of events being labeled incorrectly. Similarly, for the ^{90}Co source, the expected beta-to-gamma radiation ratio is 1:1 [9], but the model produced a 1:2 ratio when splitting the dataset into Distribution 0 and Distribution 1 events.

5 Conclusions and outlook

Methods for classifying single particle events were developed for semiconductor hybrid pixel detectors with the Timepix3 ASIC chip. These events were categorized into five classes: electrons & photons, alpha particles, heavy nuclei (except alpha), low-energy protons ($E \lesssim 100 \text{ MeV}$), and high-energy protons ($E \gtrsim 100 \text{ MeV}$). Three machine learning models were used: BiLSTM, FNN, and CatBoost. BiLSTM required no preprocessing but was slower, while FNN and CatBoost offered faster inference with cluster feature preprocessing. The SHAP technique identified the most important cluster features for particle classification as LET (linear energy transfer), the maximum energy deposited per pixel, the spread of energy distribution across pixels, and the weighted spread of pixel positions along the cluster's main axis. Additionally, the task of distinguishing single

particle events associated with electrons and photons was addressed by using an unsupervised machine learning technique. The model successfully differentiated between high-energy electrons and low-energy photons, but encountered difficulties in the intermediate energy ranges.

Future work will focus on refining classification methods and integrating them into TraX Engine software [10] developed by Advacam s.r.o.

Acknowledgments

The authors would like to thank EVOPTIMA s.r.o. for the provision of computational resources, which significantly contributed to the completion of this research. Work at Advacam was partially performed in frame of Contract No. 4000130480/20/NL/GLC/hh from the European Space Agency.

References

- [1] T. Poikela, *Timepix3: a 65K channel hybrid pixel readout chip with simultaneous ToA/ToT and sparse readout*, JINST **9** (2014) no.05, C05013
- [2] C. Granja et al, *MiniPIX Timepix3 — a miniaturized radiation camera with onboard data processing for online characterization of wide-intensity mixed-radiation fields*, JINST **17** (2022), C03019
- [3] M. Lukas, G. Carlos, J. Jan, I. Jan, T. Daniel, V. Marco and O. Cristina, *Data Processing Engine (DPE): data analysis tool for particle tracking and mixed radiation field characterization with pixel detectors Timepix*, JINST **19** (2024) no.04, C04026 doi:10.1088/1748-0221/19/04/C04026
- [4] A. Graves, S. Fernández, J. Schmidhuber, *Bidirectional LSTM networks for improved phoneme classification and recognition*, International conference on artificial neural networks, Springer, (2005), pp. 799–804.
- [5] S. Haykin, *Neural Networks: A Comprehensive Foundation* (2 ed.). Prentice Hall. ISBN 0-13-273350-1, (1998)
- [6] I. Goodfellow, Y. Bengio, A. Courville, *Deep Learning*. MIT Press (2016).
- [7] L. Prokhorenkova, G. Gusev, A. Vorobev, A. Dorogush, A. Gulin, *CatBoost: unbiased boosting with categorical features*, Advances in neural information processing systems, **31** (2018).
<https://proceedings.neurips.cc/paper/2018/file/14491b756b3a51daac41c24863285549-Paper.pdf>
- [8] M. Scott, Su-In. Lee, *A unified approach to interpreting model predictions.*, Advances in neural information processing systems **30** (2017): 4765-4774.
- [9] International Atomic Energy Agency (IAEA). (n.d.), *Interactive chart of nuclides*,
<https://www-nds.iaea.org/relnsd/vcharthtml/VChartHTML.html>
- [10] C. Oancea, L. Marek et al, *TraX Engine: Data Processing of Timepix Detectors Used in Space, Medical, and Imaging Applications*, proceedings of the IWORID2024 conference (under revision)