Experimental Assessment of PCA and DT Classification for Streamlined Position Reconstruction in Anger Cameras

Beatrice Pedretti, *Student Member, IEEE,* Susanna Di Giacomo, *Student Member, IEEE,* Luca Buonanno, *Member, IEEE,* Ilenia D'Adda, *Student Member, IEEE,* Carlo Alaimo, Marco Carminati, *Senior Member, IEEE* and Carlo Fiorini, *Senior Member, IEEE*

***Abstract*—In this paper we present a study of machine learning (ML) algorithms to simplify the computation of the planar scintillation coordinates in Anger Cameras for emission tomography applications. Two ML-based techniques for data inference and one technique for speed-up the training procedure are explored within the framework of a multimodal SPECT scanner. Firstly, the use of Principal Component Analysis (PCA), a dimensionality reduction algorithm, is explored to reduce the computational complexity of maximum-likelihood statistical estimation method. The analysis indicates a ∼3-fold reduction of computational complexity for typical Anger Camera architectures (with 72 channels). Secondly, the estimation of the scintillation coordinates is formulated as a classification problem, addressed by means of a Decision Tree (DT) classifier. No degradation of the achievable intrinsic spatial resolution (∼1.2 mm FWHM) of the detection module was observed when applying PCA (reducing from 72 to 25 components). The DT classifier was trained on experimental data obtained using a parallel-hole collimator: again no degradation of spatial resolution is observed and the computation cost is reduced by more than two orders of magnitude. Finally, in order to overcome the limits of a cumbersome training procedure involving the translation of the collimator, data augmentation was successfully leveraged for the generation of artificial data.**

***Index Terms*—Medical Imaging, Machine Learning, Scintillator, SiPM**

I. Introduction

MACHINE learning (ML) is increasingly recognized as a powerful tool to automatically extract meaningful information from a large amount of data in many fields, including the medical imaging domain. In nuclear medical imaging, ML is widely applied at different levels: starting from the detection module, where it is used to localize the position (planar or tridimensional) of the interaction point of a γ photon in the scintillator or improve the time-of-flight (ToF) estimation [1], up to tomographic image reconstruction [2]–[5], and artifact and attenuation correction stages.

This work places itself in the first framework, presenting the experimental characterization of the position sensitivity achievable when implementing ML algorithms in Anger Cameras at detector level. Our goal is to assess the trade-off between the deterioration of spatial resolution versus the reduction of the number of output signals and processing power when leveraging ML in monolithic crystals. Two ML techniques were considered: Principal Component Analysis (PCA), an unsupervised learning technique which aims at performing a reduction of the dimensionality of the original feature space, and an ensemble of fast Decision Tree (DT) classifiers, possibly combined with PCA and easily executed in embedded digital platforms.

The working principle of scintillation-based γ-ray detectors in PET and SPECT is the detection of high-energy γ photons as a consequence of their interaction with the scintillation crystal. The burst of secondary photons generated in the visible spectrum is then detected by an array of photodetectors coupled to the crystal. Today the technology for detector development is steadily transitioning from conventional Photomultiplier Tubes towards solid-state Silicon Photomultipliers (SiPMs) detectors [6], with evident advantages in terms of compactness and compatibility with magnetic fields, which is pivotal for multimodal imaging [7].

There are two major architectures for detectors: pixelated and monolithic ones. Pixelated architectures have been for a long time the standard for PET detectors and they are characterized by independent pixels, each one corresponding to a single scintillator crystal. The scintillators can be coupled 1:1 with the photodetectors or, more commonly, the scintillator array can be coupled to a less dense array of photodetectors. The achievable spatial resolution is therefore roughly determined by the size of the pixel. The design of high-resolution pixelated scanners (for preclinical imaging) requires a high 2D pixel density, at the expense of an increase in fabrication costs and system complexity.

An alternative solution consists in monolithic detectors, also known as Anger Cameras [8], [9], where a single monolithic scintillator crystal is coupled to an array of photodetectors. These detectors have gained more and more attention during the last years, for both SPECT and PET applications [1], [10]–[12]. Besides requiring a smaller number of electronics channels per unit area, they both facilitate the determination of the depth-of-interaction (DOI) [10], [12]–[15] and unbind the minimum spatial resolution from the pixel size. In fact, in this case the spatial resolution is related to the statistical fluctuations occurring in the physical process of light sharing among

This work did not involve human subjects or animals in its research.

B. Pedretti, S. Di Giacomo, L. Buonanno, I. D'Adda, C. Alaimo, M. Carminati and C. Fiorini are with Dipartimento di Elettronica, Informazione e Bioingegneria, Politecnico di Milano, Milano (MI), 20133, Italy and with Istituto Nazionale di Fisica Nucleare (INFN), Sez. di Milano, Milano (MI), 20133 Italy (e-mail: carlo.fiorini@polimi.it).

pixels. Consequently, the achievable resolution significantly depends on the reconstruction algorithm. The main drawback of a continuous scintillator is that it limits the maximum thickness of the crystal.

During the years, many techniques have been proposed for estimating the interaction coordinates of γ photons in monolithic detectors: the first centroid-based methods [9] have been replaced by statistical approaches such as the Maximum-Likelihood Estimation (MLE) algorithm, which was demonstrated to provide better positioning performances and reduced edge biases [16].

Different ML-based reconstruction strategies have been proposed for the estimation of the position of interaction of high-energy photons in PET monolithic detectors [1]. Artificial Neural Networks (ANNs) have been implemented for estimating both the 2D and 3D (with DOI information) interaction coordinates in monolithic detectors [17]–[21]. Decision tree-based methods like Gradient Tree Boosting have been implemented for 2D [22] and 3D reconstruction [23] in monolithic PET detectors; since they basically implement a set of binary decisions, they have been observed to provide a computationally relaxed algorithm for fast event processing, which is also sustainable by the memory allocation available in commercial FPGAs.

A common feature of state-of-the-art reconstruction techniques is that the signal information from all the photodetectors coupled to the monolithic crystal is required; this can represent an important limitation in terms of computational effort, especially since research in the field of PET scanners is moving towards the design of scanners with a large field of view (FoV) [24], up to whole body, with a consequent increase in the number of signals to be acquired and processed.

In this context, different multiplexing techniques have been proposed throughout the years, in order to reduce the number of output channels to process [25]–[29]. However, it has been observed that reducing the number of output signals from the photodetectors implies a reduction of the amount of information available to the algorithm, thus being detrimental for the spatial resolution achieved by the detection module [28].

In order to overcome this issue, recent studies have proposed the implementation of PCA-based multiplexing schemes, which offer the possibility to reduce the number of output channels, by keeping the original information content almost unchanged. From a simulation study evaluating the positioning performances of different multiplexing strategies, the PCA transformation of the original channels proved to guarantee the best results in terms of spatial resolution [29].

This work aims at exploring from an experimental point-of-view the performance of a SiPM-based Anger Camera implementing PCA feature analysis combined to MLE. Furthermore, the performance of a specifically trained ensemble of DT classifiers for position estimation is assessed. Finally, a new technique, based on Data Augmentation (DA), for the enrichment of the training dataset acquisition is also proposed, trading-off spatial resolution and complexity of the calibration procedure.

The paper is organized as follows: Sec. II introduces the clinical module used for the experimental measurements, while Sec. III and Sec. IV introduce, respectively, the Maximum-Likelihood Estimation algorithm for position sensitivity, and the implementation of PCA for feature reduction. A discussion of the results for MLE combined with PCA is presented in Sec. V. Sec.VI introduces the use of Decision Tree-based classifiers for position sensitivity in Anger Cameras, while Sec. VII, Sec. VIII, and Sec. IX describe respectively the experimental characterization of DTs, the use of DA for training dataset acquisition, and the related discussion. Conclusions are drawn in Sec. X.

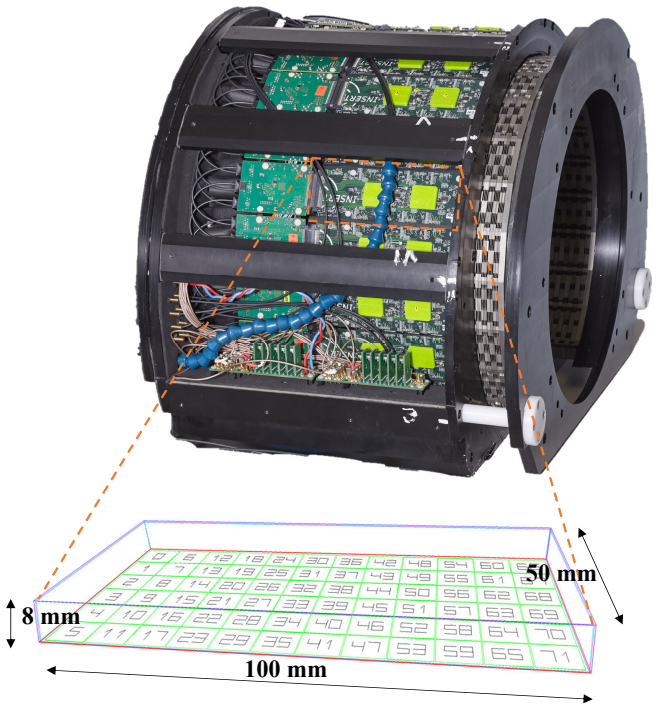

Fig. 1. Picture of the INSERT clinical instrument. This MRI-compatible insert is composed of 20 clinical modules, each equipped with 2 ASICs for signal processing and a DAQ board for digitization. Data are sent to a central unit via an optical fibers bundle. The INSERT clinical modules couple a 10 cm×5 cm×8 mm CsI(Tl) scintillation crystal to an array of 72 SiPMs macro-pixels, each one having an 8 mm×8 mm area.

II. Experimental Setup: the INSERT System

A single detection module of the INSERT clinical scanner for SPECT/MRI imaging (Fig. 1) was used in this study to test the position sensitivity achieved using machine learning algorithms [30], [31].

The goal of the INSERT project is to develop a custom stationary SPECT apparatus that can be used as an insert for commercially available MRI scanners, for simultaneous structural and functional imaging.

The spatial resolution of the detection module is important because the INSERT camera is spatially constrained inside the bore of the MRI scanner (of ∼50 cm diameter). The minification effect leads to a 10 mm spatial resolution of the image (comparable to the one of a stand-alone SPECT system), starting from an intrinsic spatial resolution of the camera as small as 1 mm.

The SPECT system is composed of 20 modules, each one coupling a 50 mm×100 mm×8 mm CsI(Tl) scintillator to an array of 4 mm×4 mm RGB-HD SiPM tiles ([32], [33]). The detectors readout scheme merges the SiPMs pixels into groups of 4, for a total of 72 macro-pixels having 8 mm side each. This represents a compromise between system complexity and spatial resolution ([34], Fig. 1).

The clinical module reaches a spatial resolution of 1.04±0.18 mm full width at half maximum (FWHM) within a useful field of view (UFOV) of 90 mm×40 mm, using MLE algorithm [35]. This result will be taken as a benchmark to compare the performance of the ML-based methods proposed in this work.

III. Maximum Likelihood Estimation Algorithm for the INSERT System

MLE is a statistical estimation method that uses labeled data to infer unknown parameters of a process, such as the γ photon interaction coordinates in the scintillator, or, equivalently, the generated number of photoelectrons from a γ photon interaction in the scintillator [29], [36]–[42].

Statistical methods are based on the prior definition of a mathematical model of the detector. Specifically, the MLE algorithm requires the knowledge of the average response of each individual photodetector as a function of the event position, the so-called Light Response Functions (LRFs).

The MLE algorithm estimates the event coordinates (x, y) by maximizing the following log-likelihood function:

$$\ln L(x, y, N_{ph}) = \sum_{j=1}^{D} [n_j \ln(LRF_j(x, y) \cdot N_{ph}) + - LRF_j(x, y) \cdot N_{ph}] \quad (1)$$

where N_{ph} is the amount of optical photons produced by the local energy deposition processes: it is assumed constant in our application of emission tomography with a single photopeak, D is the number of photodetectors, and n_j is the number of photons detected by the j-th photodetector.

For the case of the INSERT clinical module, if the logarithm is implemented as a look-up table (LUT), the likelihood computation of a single point within the grid where the LRFs are defined is given by 72 multiply-and-add operations, followed by a subtraction (for the maximum value search).

It is apparent from (1) how important is an accurate knowledge of the LRFs. The LRF computation method used for INSERT module is based on an iterative procedure, which exploits measured scintillation events produced by an uncollimated γ source that uniformly illuminates the camera [43]–[45].

The MLE algorithm implementation led to an intrinsic spatial resolution in the crystal of ∼1.2 mm FWHM considering a 90 mm×40 mm UFOV. This value is slightly worse than the one reported in [35] probably due to the aging of the module in the last four years.

Despite acknowledging its state-of-the-art performance in terms of position sensitivity, we highlight that the application of MLE algorithm to data collected from the INSERT clinical instrument requires the storage of 125×10^3 LRFs data (a square grid of LRFs is defined in the 5 cm×10 cm scintillation crystal plane, having 200 μm pitch in both directions), and $72\times125\times10^3 = 9\times10^6$ multiply-and-add operations per event (plus the maximum value search complexity).

In the INSERT module, each LRF, representing the average response of the 72 macro-pixels for a given scintillation position, is an entry of the LUT and contains 72 elements. If real-time processing is targeted, for each event all the LRFs need to be fetched from the memory at a very demanding transfer rate above 130 GB/s, for a modest count rate of 10 kcps. If the logarithm is implemented as a look-up table, this constraint is worsened of a factor 2. Some solutions have been proposed in literature to reduce the overall number of operations, such as splitting the MLE search in two rounds [46], but the number of operations per event would still be as large as 1.5×10^5.

IV. Principal Component Analysis for Dimensionality Reduction

One could wonder if, for estimating the position of γ-ray interaction in monolithic scintillators, the use of all SiPM signals is strictly necessary.

Since the signals are intrinsically subject to statistical fluctuations (due to multiple sources including: the generated optical photons collection, their scattering in the crystal, the SiPM cross-talk and dark counts, the electronics readout noise, etc. [47]), it might be beneficial to implement an appropriate transformation of the statistical feature space, in order to reduce the dimensionality of the problem, and to extract relevant features from the statistical variables distribution.

Dimensionality reduction techniques are sets of methods allowing the transformation of high-dimensional data into a meaningful representation of reduced dimensionality, expressing the data in such a way that their similarities and differences are highlighted.

With this work, we want to find a reduced set of features that allows us to infer spatial information without (or with minimum) accuracy loss.

Dimensionality reduction techniques applications are reported in [48]–[50], while a more general discussion is reported in [51], [52]. PCA performs dimensionality reduction by embedding the data into a linear subspace of lower dimensionality. PCA is a consolidated tool for the unsupervised analysis of data variance, since it allows the identification of a new, orthonormal reference system which maximizes the variance of data across a subset of dimensions (the principal components, PCs).

A. Principal Components Computation

The PC computation is based on an iterative algorithm, which starts from finding the first component as the one maximizing the explained variance in the data. At each subsequent iteration, the next PC is computed as the one which maximizes the residual variance, adding the constraint of orthogonality with respect to all the PCs previously found. The estimation of the optimal number of PCs is usually performed by looking at

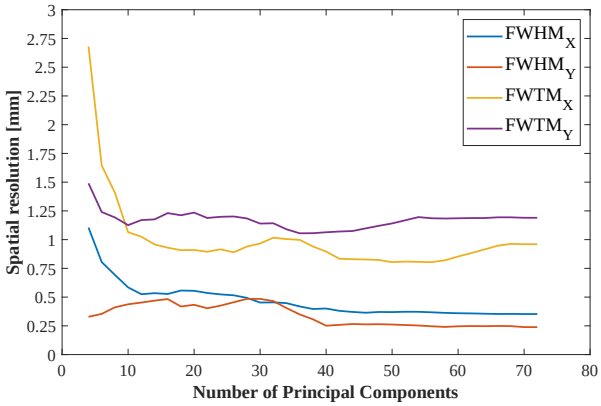

Fig. 2. MLE performance for the reconstruction of single-spot simulated events (FWHM and FWTM for both x and y coordinates), for varying numbers of principal components.

a scree plot (percentage of explained variance as a function of the number of PCs); the number of components corresponding to the "elbow" of the scree plot is usually chosen as dimension of the new feature space [53].

The position of each observation in the new coordinate system of PCs can be calculated as the linear combination of the values of the original variables. For a typical Anger Camera (with less than 100 channels) the cost of this algorithm is negligible with respect to the MLE computational cost.

In a real scenario, the computation of PCs has to be performed on a proper calibration dataset containing events interacting all over the detector surface. An experimental flood irradiation dataset was used in this work for both the LRF estimation and the computation of the d PCs as the eigenvectors of the dataset covariance matrix, where d represents the dimension of the new feature space and is a tunable parameter.

Once the PCs are computed, they can be used for the projection of collected data from SiPMs (either in real-time or in post-processing). If the detector module is affected by a gain drift during its lifespan or some pixel fails, it will be necessary to repeat the measurement.

In order to perform MLE after PCA, both the LRFs and the event to reconstruct have to be projected into the reduced feature space. While the projection of an event in the principal components space is obtained by a simple scalar product between the observation and the set of PCs, obtaining LRFs_{PCA}, namely the equivalent of the LRFs in the transformed space, required a different approach.

In order to enable the computation and the allocation of the LRFs 3D continuous curves, a proper binning of the (x, y) coordinates is performed; the length of the crystal along the x and y direction is divided respectively into 506 and 258 discrete bins, with a pixel dimension of 0.2 mm. Consequently, the LRFs estimation process leads to the definition of 72 different matrices with dimension 258×506.

In order to obtain the LRFs_{PCA}, a new dataset was created, where rows represented the 258×506 bins, while columns represented the values assumed by each LRF for a specific bin; this operation allowed the creation of a 72-dimensional dataset which could be then projected into the new principal components space. After the projection in the principal components space, the LRFs have been composed again into a d-dimensional object defined inside the same 258×506 bins matrix, which finally constitutes the LRFs_{PCA}.

Fig. 2 depicts the position sensitivity expressed as FWHM and FWTM (full width at tenth maximum) detailing the MLE algorithm performance with varying numbers of PCs considered for the computation, for the 8 mm×8 mm SiPM array of the clinical INSERT module. Training and validation datasets were acquired from single-spot simulations performed in ANTS2 environment (a simulation package developed for Anger Camera type detectors [54]), where data correspond to the number of photoelectrons generated in each SiPM pixel of the Anger Camera, event per event. The elaboration procedure that was adopted is the same as the one explained in the following subsection (Sec. IV-B) for the experimental measurements. The simulation results show that ∼25 PCs are sufficient for our target spatial resolution.

This outcome has been confirmed also by the experimental measurements: as shown in the following subsection, we found that a subset of 20-25 components is sufficient to describe the geometrical feature of light spreading of experimental data on the SiPM matrix. This reduced set of components allows a reduction of the LUT size, leading to reduced computation and bandwidth requirements for the MLE algorithm in the event reconstruction process (3×10^6 multiply-and-add operations per event if 25 PCs are used). The computation of PCs has a computational complexity corresponding to $N_{PC}\times N_{ch}$ multiply-and-add operations, where N_{PC} is the number of PCs and N_{ch} is the number of detector channels (original feature space). If N_{PC} is chosen equal to 25 and the original number of channels is 72, the number of operations to recall the PCs is in the order of 10^3. The increase in computational complexity, when PCA and MLE are cascaded, is therefore negligible.

B. PCA Experimental Characterization

In order to compare previous experimental results of the INSERT module (obtained using the SiPM signal values as statistical variables in MLE) with the ones obtained with PCA (using the extracted PCs as statistical variables for MLE), we have evaluated the spatial resolution of an INSERT clinical module for the two cases. For the examined technique, experimental results were extracted when using all available PCs or a finite subset of components.

During the preliminary calibration phase, the LRFs of the clinical module were experimentally computed by means of a flood irradiation of the scintillation crystal, and using the iterative algorithm exploited in [34], [43]. After computing the LRFs, they were projected into the principal components space and used to build the new LUT.

Fig. 3 shows the intensity of the first 6 principal components over the 72 macro-pixels of the INSERT module. It is interesting to note that the intensity of these features underlies a geometrical description of the light distribution (e.g., the first component indicates the derivative along one axis of the crystal, while the second component represents the second order momentum along the same direction).

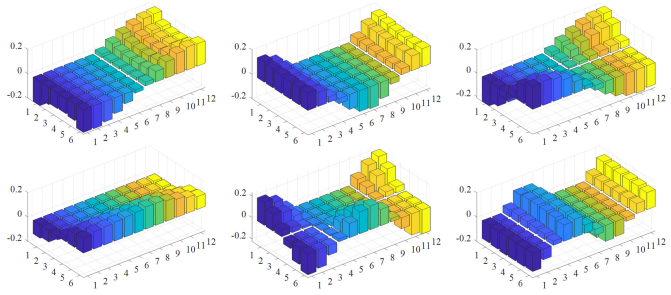

Fig. 3. Intensity distribution of the first 6 principal components over the 72 pixels of the INSERT module. It is apparent how the PCs express geometrical features of the light distribution over the detector surface (e.g., the first component indicates the derivative along one axis of the crystal, while the second component represents the second order momentum along the same direction).

To test the accuracy of the MLE algorithm with PCA, a new dataset was acquired where the gamma photons were collimated using a 2 mm-pitch parallel holes collimator. The reconstructed images are shown in Fig. 4 (only a central FOV of 60 mm×30 mm is here reported), together with the unidimensional profile of the reconstructed spots. For each spot, the full widths at half maximum of the Gaussian curves that fit the histograms of the reconstructed x and y coordinates were calculated, and the spatial resolution was quoted as the mean FWHM over the two coordinates, considering a UFOV of 90 mm×40 mm. When using 25 components, the worsening of the spatial accuracy is negligible (with an average FWHM among all points within the UFOV of 1.25 mm±0.1 mm). Instead, if only 10 components are used, the spatial resolution is degraded (1.55±0.25 mm).

V. Discussion on MLE with PCA

The implementation of PCA previous to MLE in modern Anger Cameras allows a reduction of dimensionality and, consequently, of the required computational complexity. A number of PCs equal to 25 is sufficient, for the INSERT clinical module Anger Camera, to achieve the target spatial resolution FWHM≈1.2 mm.

Since the increase in computational complexity when PCA and MLE are cascaded is negligible, by applying PCA the computational complexity of MLE algorithm only scales linearly with the dimensionality of the problem (the number of features), and with the number of LRFs. A 2.9-fold reduction of the number of operations is foreseen, if the number of components used for the estimation goes from 72 to 25. The reduction is even more significant for Anger Cameras with a larger number of channels.

It is interesting to note that this approach does not decrease the Anger Camera hardware architecture complexity. The number of channels to convert (and therefore the ASIC front-end channels, the ADC number of channels, etc.) is in fact not changed. The number of channels to convert might be reduced with an analog implementation of PCA, using for example a memristor-based cross-point array for matrix-vector-multiplication [55].

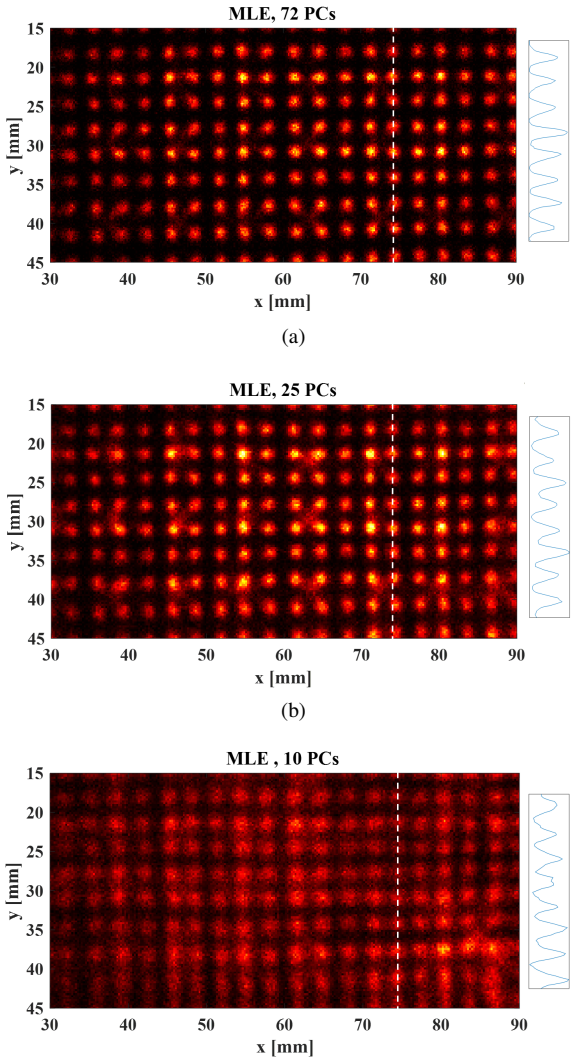

Fig. 4. Reconstructed images (only a central FOV of 60 mm×30 mm is here reported), applying MLE with PCA, of a 2 mm-pitch grid, using (a) all 72 principal components (average spots FWHM=1.22 mm±0.08 mm), (b) using 25 principal components (average spots FWHM=1.25 mm±0.1 mm), and (c) using 10 principal components (FWHM=1.55±0.25 mm). FWHM was determined as the average across the two directions, inside a UFOV of 90 mm×40 mm. The profiles crossing the centers of a line of collimator holes (dashed white line) are also reported. Smoothing (a 10-sample moving average) is applied to the profiles to clean their visualization.

VI. Decision Trees

Decision Trees can be exploited in the machine learning framework as classifiers, implementing models used in supervised learning problems to generate predictions starting from data. More specifically, DTs infer data using a sequence of conditional statements [56].

The problem of reconstructing the 2D position of interaction in a gamma camera can be addressed as a classification problem, considering the sampled values from SiPM macropixels as the feature array, and dividing the (x, y) interaction coordinates in a finite set of possible positions (classification outcomes, also referred to as classes).

Therefore, in this section therefore we present the use of DTs as an alternative to MLE as a reconstruction algorithm

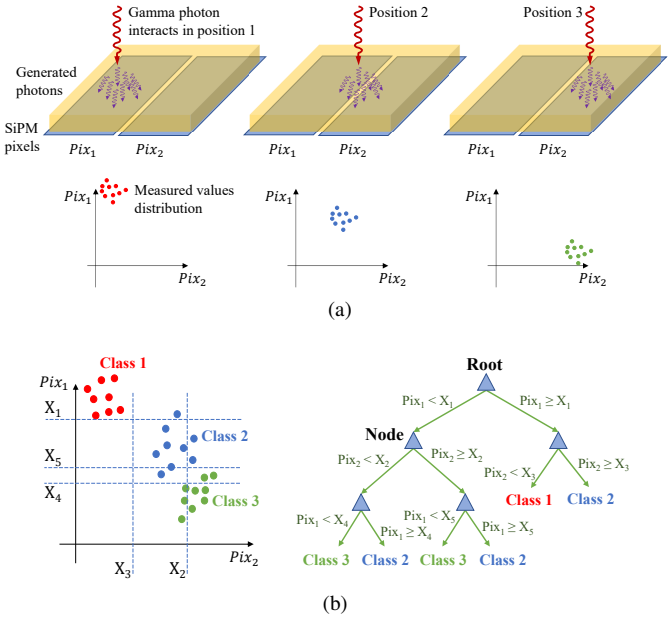

Fig. 5. (a) Scintillation events whose geometrical efficiency is higher for Pix1 (or for Pix2) will tend to have larger value of Pix1 (or Pix2) feature sampled value. (b) The DT training leads to the definition of a sequence of partitions of the training dataset in the feature space that opportunely describes the class distribution of the training dataset, allowing for classification of new data.

for the γ-ray position of interaction in an Anger Camera. The inputs of the DT are the signals produced by the front-end electronics during the scintillation event. The outputs of the classifier are the scintillation coordinates.

A. Decision Tree Training

The training of a decision tree is the process of defining its nodes (boundary conditions) and leafs (classification outcomes). In order to fulfill this task, it is essential to use a reliable training dataset, which is a dataset containing observations whose target value is known. Given a particular dataset, the decision tree structure is defined by the training algorithm.

Since the classification task aims to the identification of positions of interaction starting from data distribution in the pixel space, we need to build a dataset where we have the SiPM macro-pixels signals converted values, for a large number of scintillation events. Also, we need for each event the exact information (with an accuracy smaller of 1 mm, our target resolution) of the (x, y) position of interaction. This dataset will be used to build the DT classifier that allows for the inference of new acquired data.

A simplified version of the classification problem for the Anger Camera is depicted in Fig. 5, showing the relation between training dataset collection and DT structure. Since it is not possible to imagine how the variable space is divided for the 72-dimensional case, we introduced a problem with only 2 pixels (therefore, with a 2D feature space) and 3 classes. Since only 3 classes are used for training, the position sensitivity obtained here is limited to the separation of events in the left side from events in the middle and from events in the right side of the scintillator.

Any scintillation event whose geometrical efficiency is higher for Pixel 1 (or for Pixel 2) will tend to have larger value of *Pix1* (or *Pix2*) feature sampled value, as shown in Fig. 5a. If we use a collimated gamma photon beam to build the training dataset, we can then conveniently divide the feature space in such a way that the DT prediction of new events (under the hypothesis that the new events to infer are sufficiently similar to the training dataset in the feature space) will be correct. During the training phase, while partitioning the training dataset in the feature space, the algorithm is identifying a suitable ensemble of node conditions (Fig. 5b), which will lead to the classification once the leaf is reached.

The training procedure is discussed in the following with greater detail; anyway, we would like to stress the importance of building an accurate training dataset, without which it would impossible to achieve an accurate classification, even with modern, sophisticated training procedures.

VII. Decision Trees: Experimental Results

In order to verify the achievable accuracy when using DT classifiers, we collected a training dataset and a test dataset using the INSERT clinical module. The a-priori knowledge of the position of interaction in the scintillator was built using a parallel hole collimator, as explained in the following. The sub-mm precision shift of the parallel hole collimator for building the training dataset might be avoided leveraging Data Augmentation (DA) (Sec. VIII).

Once the training dataset was built, it was possible to train a decision tree classifier to be used for inferring test dataset positions of interaction. The quality of the reconstruction using the DT was compared to the one of the MLE algorithm evaluating the average FWHM of the reconstructed spots within a UFOV of 90 mm$\times$40 mm.

A collimated ^{57}Co dataset was acquired for 16 different position of the parallel-hole collimator (Fig. 6a). A custom setup was built where the collimator was moved on the two axes exploiting a pair of uniaxial translators, in order to fully automate the training dataset acquisition process. A large number of dataset blocks was generated (1250 spots $\times$ 16 collimator positions = 20×10^3), each one associated to a single hole position over the CsI(Tl) crystal through which γ-ray beams reach the crystal, and therefore to a unique position in the scintillator. In order to perform the DT training task, every dataset block was associated to a unique class.

The pitch between the collimator holes is 2 mm (hole diameter: 0.2 mm), while the distance between adjacent classes was set to 500 μm (Fig. 6b). Even though the use of the parallel hole collimator allows to simplify the complexity of the experimental measurement needed to retrieve the training dataset (the collimator is only moved 16 times, instead of the 20×10^3 times that would be needed using a single hole collimator), for each position of the parallel hole collimator data corresponding to multiple classes were collected. Therefore, it was necessary to separate the classes corresponding to each collimator aperture from one another.

In order to separate the classes, one preliminary position estimation was performed by means of the MLE algorithm,

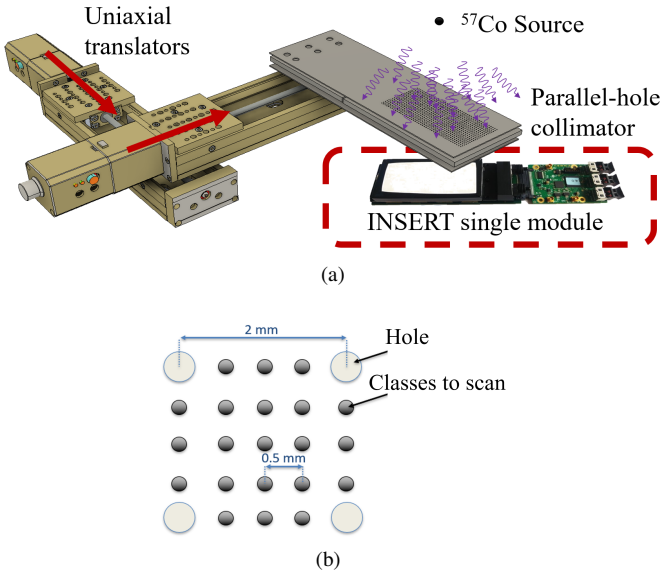

Fig. 6. (a) Experimental setup for the automated procedure of training dataset collection. Two uniaxial translators move the parallel-hole collimator. Data corresponding to individual spots are identified and labeled with their irradiation position, in order to use the label as a knowledge a priori over the training dataset. (b) Collimator holes, distributed on a 2 mm pitch grid, are translated 16 times in the training phase, on a grid having a 0.5 mm pitch.

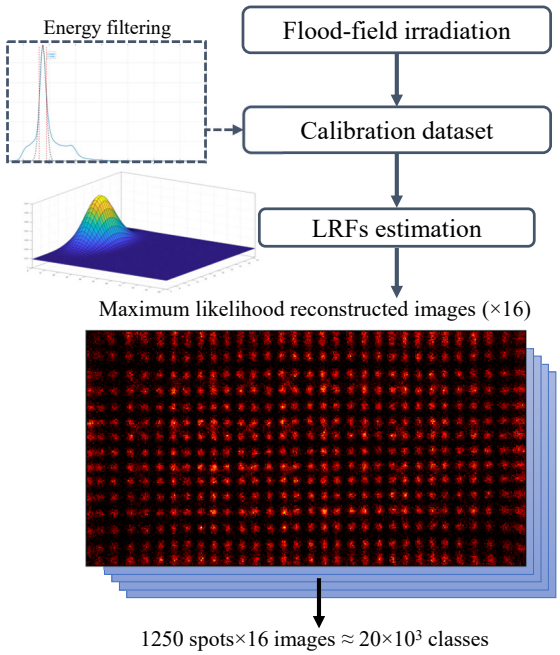

Fig. 7. The 16 dataset corresponding to a single position of the parallel-hole collimator are reconstructed using MLE algorithm. The forward model consists in light response functions generated from a simple flood field irradiation through an iterative procedure. The 16 datasets are divided in 1250 spots each, applying k-means clustering algorithm.

where the forward model consists in light response functions generated from a simple flood field irradiation through an iterative procedure (Fig. 7) as introduced in Sec. III. Once the image corresponding to a single position of the grid was acquired, spots were separated from one another using the k-means unsupervised learning clustering algorithm [57].

This preliminary procedure, albeit including the computationally intensive MLE algorithm, allows for the creation of a DT-based estimator that, after the training, offers several benefits in terms of effort sustained by the computation unit, as discussed in Sec. IX. In fact, once the training is completed, the reconstruction of the scintillation position for any new test set is performed by the DT classifier with a sequence of conditional statements.

The training dataset processing, the DT training phase and its optimization, and the test dataset image reconstruction were performed in MATLAB environment.

As anticipated, the basic idea of DT training for the reconstruction of the (x, y) scintillation coordinates of the γ photon in Anger camera consists in a discrete classification problem, where each class corresponds to a specific position in the crystal. The scintillator surface is virtually divided into C classes, subjecting the achievable spatial resolution to the inter-class distance and therefore to the number of defined classes. While the minimum achievable spatial resolution decreases when decreasing the distance between two adjacent classes, choosing a too dense pixelation on the other hand unavoidably leads to an increase in computational complexity and required memory resources.

The problem of localization of the γ-photon scintillation coordinate was therefore solved by implementing a cascade of DT classifiers. Specifically, a global DT (GDT) was implemented to first select a macro-region of the crystal where the target class was found, and then a local DT (LDT) for each crystal macro-region was used to identify the correct position of interaction of the γ-ray beam. All LDTs were trained with the same set of hyperparameters. In order to avoid cascading misclassification errors from the GDT to the LDTs, macro-regions overlap along the borders, so that if the GDT selects a wrong region for classes in the proximity of the region border, correct classification is still viable.

The estimated objective function (misclassification error) for the LDTs training hyperparameters, obtained in MATLAB environment during the training procedure, is reported in Fig. 8. While the LDT leaf size was directly optimized ($\sim$200), because a larger number of node (or splits) in the tree implies a larger memory occupation (data latency is not an issue, being proportional simply to the logarithm of the tree size), the maximum number of splits was set equal to the minimum value that allowed accuracy optimization within 1% from the absolute optimum, ($\sim 3\times10^4$). *Twoing rule* [58] was selected as splitting criterion for both GDT and LDTs.

To ease the training procedure, PCA, introduced in Sec. IV, was applied on the training dataset thus obtained. In other words, before growing the tree, the corresponding input training dataset underwent feature reduction, resulting in a new set of features to be used for training procedure. The first 25 principal components were used as training dataset, according to the elbow method. It was empirically observed by the authors that when PCA is not applied, the DT classifier for position sensitivity in the scintillation crystal is prone to overfitting, leading to the generation of trees having substantial size, therefore requiring a memory allocation incompatible

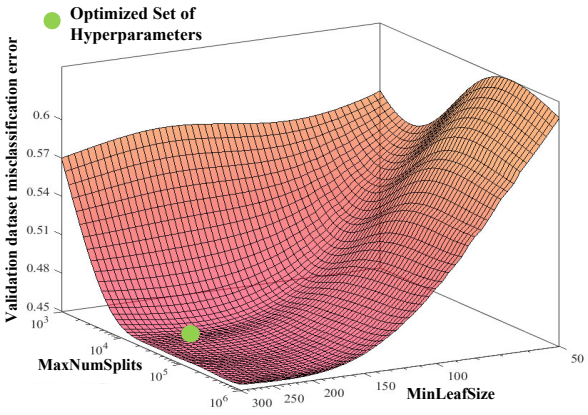

Fig. 8. Misclassification error as a function of the DTs hyperparameters; the error is minimized in the DT training procedure. The number of splits is reduced (within 1% from the absolute optimum), in order to reduce the FPGA memory occupation.

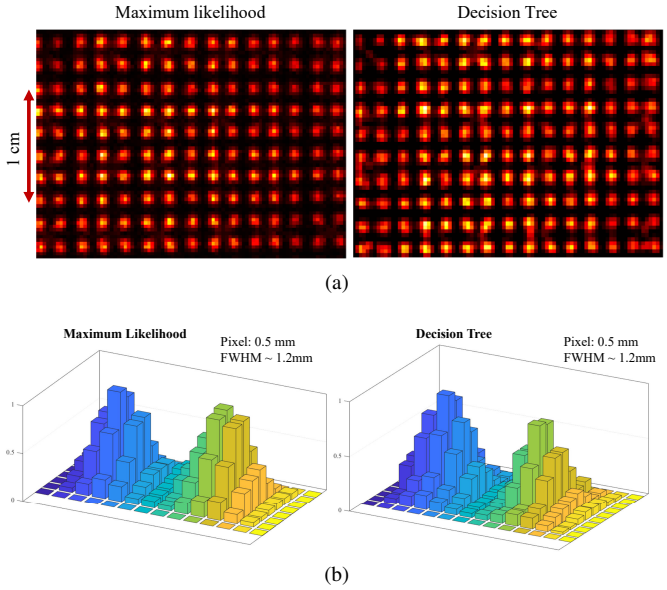

Fig. 9. (a) Reconstructed images of a flood irradiation through the collimator with 2 mm pitch between holes, using the MLE and the DT algorithm, with a 0.5 mm binning of the image. The FWHM of the reconstructed spot is similar. (b) Image showing the event distribution over two adjacent parallel holes positions.

with an FPGA-embedded solution. Fig. 9 shows the comparison between two image reconstruction of the test dataset (Fig. 9a) and of two individual spots belonging to the grid (Fig. 9b), using the MLE method and the DT classification algorithm. In order to compare the experimental results, the MLE algorithm pixelation is set equal to 500 μm, the same as the DT reconstructed image pixelation, which is currently limited by the spatial frequency of the experimentally obtained classes. The spots are very well separated in both cases and the experimentally computed mean resolution of $\sim$1.2 mm FWHM (reported in Tab. I for (x, y) directions) is comparable for both algorithms. The DT algorithm seems to introduce a further correction in the distortion effect with respect to the MLE method thanks to the training procedure: the deviation of the reconstructed spot center from the true position, computed

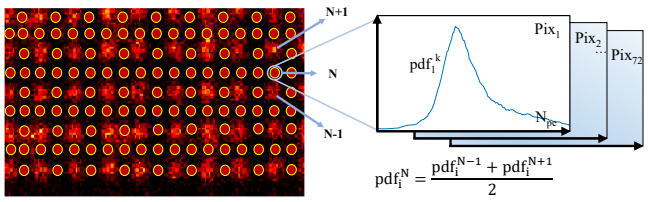

Fig. 10. Data Augmentation (DA) algorithm. When spots are well separated one from another in the grid, it can be assumed that normalized local spectra on each pixel are a good approximation of the pdfs of the value sampled on each pixel, for a given position of interaction (the spot). The pdf of the artificial point N is computed from a linear interpolation of the distributions of the two adjacent spots (N-1) and (N+1). Circled spots are the virtual ones created by DA.

for experimental data as the RMSE of the Euclidean distance between the true and the predicted (x,y) positions of the spot center, shows a $0.84\,\mathrm{mm}$ error inside a 90 mm$\times$40 mm UFOV for the MLE algorithm, and a $0.6\,\mathrm{mm}$ error for DTs.

VIII. Data Augmentation for Decision Tree Training

The quality of the training dataset and, in particular, the carefulness of its acquisition procedure, determine the capability of the system to correctly infer information from new data.

The acquisition procedure shown in Sec.VII grants a high quality of the DT training dataset: since, ideally, the γ photons interacting with the scintillators can pass the collimator only from the parallel holes, the interaction position uncertainty is given, event per event, by the 0.2 mm holes diameter. The holes position, moved by the mechanical setup, is known with negligible error. The main drawback of this procedure is that the duration: for a measurement with a 200 kBq point source, translating the collimator with a 0.5 mm step requires, as previously stated, 16 acquisitions for a total of about 4 hours. Secondly, it is necessary to operate a dedicated experimental setup, where the collimator motion is automated and controlled by software. Finally, even if the DT classification algorithm allows for a reduction of computational complexity, there is no benefit in terms of spatial resolution when using this complex training procedure: the spatial resolution obtained experimentally is approximately equal to 1.2 mm for both the ML-based and the MLE algorithm (Tab. I). It has to be highlighted that the latter only needs a quick flood measurement for LRFs computation [35], while the former requires a complete training dataset. However, albeit facing an intensive training phase with respect to the MLE algorithm, the DT inference appeared to be less burdensome, with the data latency and memory bandwidth requirements reported in Sec.IX.

To overcome the limits of the training dataset acquisition, we introduce here a custom Data Augmentation (DA) algorithm aimed at the interpolation of artificial data corresponding to intermediate classes between the parallel-hole collimator positions.

The generation of artificial data ensured a class density, and therefore an image pixelation density, higher than the one of the irradiated spots. In order to augment the dataset, the

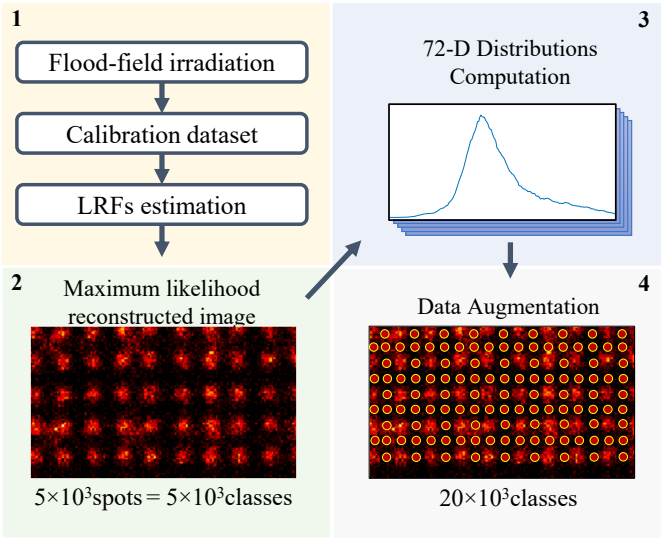

Fig. 11. Workflow for the generation of the augmented dataset. Once LRFs are extracted from a flood-field irradiation, a reconstructed single grid is used for the computation of the 72-D distributions, each ensemble of 72 pdfs being associated to a single spot. The distributions are then averaged between adjacent spots, in order to artificially generate events to be used for training of the DT. For the sake of clarity, in the sample image shown here the sampling density is augmented of a factor 4, going from 2 mm to 1 mm. This experimental procedure was applied starting from a 1 mm sampling, obtaining 0.5 mm distance between classes.

measured values distributions for each position of interaction have to be computed; to this end, we measured the local spectra associated to individual macro-pixels. For a sufficiently large number of observations, we can assume that local spectra are a good approximation of the probability density function (pdf) of the value sampled on each pixel for a given position of interaction (the spot). If we compute the local profiles (i.e. the signal histogram collected by each SiPM) for each irradiation position, we therefore obtain 72 pdfs for each point of interaction (Fig. 10).

Under the assumption of smoothness of data distributions between adjacent spots, we can infer that pdfs in intermediate positions are the linear interpolation of close-by known distributions (Fig. 10). We used the computed set of pdfs to generate the artificial part of the training dataset by exploiting an inverse transform sampling technique [59]. Let us assume that the characteristic length for variations of the data distribution (in the 72 SiPM macro-pixels space) is larger than 0.5 mm and 1 mm; the pdf for an arbitrary intermediate position can be approximated by a linear interpolation of the experimentally collected data distribution.
In other words, if the characteristic length over which the 72-D pdfs over macro-pixels change is longer than 1 mm, we can use the irradiated points as extrapolated samples of the 72-D pdfs and approximate the distribution in intermediate points using linear interpolation.
Once the data distribution is known for any arbitrary point of interaction, the corresponding data can be generated and used for training a machine learning classifier.

As a first attempt, we generated a training dataset for DT classifier starting from a grid of experimentally irradiated

TABLE I
SPATIAL RESOLUTION IN TERMS OF FWHM, 90 MM×40 MM UFOV

Reconstruction method	**FWHM$_x$** [mm]	**FWHM$_y$** [mm]
Maximum Likelihood	1.20 ± 0.17	1.24 ± 0.19
Decision Trees	1.18 ± 0.25	1.27 ± 0.26
DTs + DA (0.25 mm step)	1.08 ± 0.28	1.11 ± 0.18
DTs + DA (0.5 mm step)	1.36 ± 0.20	1.44 ± 0.22

CORRECTED SPOT POSITION

spots (similar to the one presented in Sec.VII) having a 0.5 mm pitch in order to improve spatial resolution; this was obtained with 16 acquisitions with the translating collimator. A denser spatial sampling was obtained by interpolating the 0.5 mm pitch experimental spots, leading to an augmented dataset comprising of both experimental and intermediate artificial irradiation spots having 0.25 mm pitch between one another. This augmented dataset was subsequently used to train the classifier (Fig. 11), resulting in a mean FWHM equal to 1.1 mm (x and y values are reported in Tab. I) for the augmented dataset with 0.25 mm quantization.

A second analysis was then performed with the aim of easing the training dataset acquisition procedure, rather than focusing on the improvement of spatial resolution: starting from a setup similar to the one in Section VI, the 2 mm pitch collimator was moved with 1 mm steps, covering the entire detector surface with only 4 acquisitions, instead of 16, and leveraging DA to generate spots with a 0.5 mm distance among each other; this leads us to a dataset comparable to the one obtained experimentally with the procedure presented in Sec. VII, with the difference of an acquisition time reduced by 75%. The same results could theoretically be achieved with a 1 mm pitch collimator that does not need to be translated, effectively overcoming the need for a troublesome calibration and acquisition procedure and still obtaining a dataset comparable to the one used for the training of the original decision trees, with negligible error. As before, a DT classifier capable of identifying the position of interaction with 0.5 mm pitch between classes was trained. The experimental results, however, reported a modest spatial resolution loss, with the average FWHM over both x and y directions increasing from ∼1.2 mm to ∼1.4 mm; detailed results reporting mean FWHM and uncertainty along both directions are shown in Tab. I.

IX. DISCUSSION ON DECISION TREE AND DATA AUGMENTATION

The use of ML algorithms for classification and regression is today an appealing alternative to classical statistical methods for data inference. As the quality of the outcome is largely dependent on the training phase of the algorithms, it is essential to identify robust procedures for data collection. The simplification of the training procedure is a crucial target

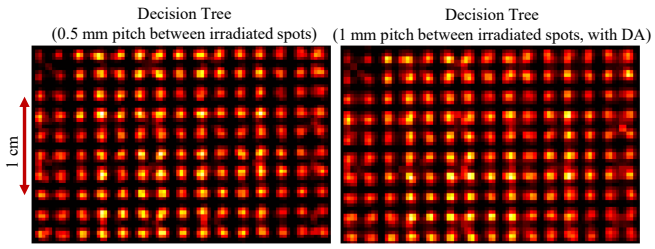

Fig. 12. Reconstructed grid with DT classifier. (left) The training dataset is generated using the moving collimator with a 0.5 mm pitch, for a total of 16 acquisitions (FWHM∼1.2 mm). (right) The training dataset is generated starting from 4 grid acquisitions with 1 mm pitch and using DA to obtain a 0.5 mm quantization between experimental and interpolated spots (FWHM∼1.4 mm).

for a widespread diffusion of ML algorithms for position reconstruction in Anger cameras.

The experimental results showed that a modest spatial resolution loss is present when implementing DT classification. However, in order to achieve comparable spatial resolution, a very dense training dataset is necessary. It can achieved by means of DA, which can be leveraged to speed-up the training procedure.

The benefits offered by DT with respect to MLE in terms of computational complexity and memory bandwidth are distinct: a cascade of only 10-15 subtractions is needed for classification with DT, allowing therefore a data latency smaller than few hundreds of ns with standard FPGAs, while less than 1 kB of data (the DT predictors) per event have to be fetched from memory. Assuming that condition values are stored as single precision floats, predictors are stored within a single byte, and that children nodes arrays are stored using 2 bytes, for a 10^5 nodes DT, about 1 MB of memory is sufficient to store a DT such as the one used here, which is compatible with low-end commercial FPGAs.

Furthermore, we experimentally demonstrated that the implementation of PCA prior to DT classification is essential to obtain the reported position sensitivity, in order to keep the number of required nodes compatible with a simple FPGA or microcontroller implementation, therefore inferiorly limiting the gain in terms of computational complexity introduced by the DT classifier. Even in this case, since the number of multiply-and-add operations needed to recall the PCs is in the order of 10^3 for the INSERT clinical module, the gain in terms of computational complexity with respect to the MLE algorithm is larger than 10^2.

In conclusion, it is interesting to note how the classification latency might benefit from an analog hardware implementation of PCA, using for example a memristor-based cross-point array for matrix-vector-multiplication [55], as suggested in Sec. V for MLE leveraging PCA algorithm.

X. Conclusions

We have reported the study of two machine learning techniques, that can be used as mathematical tools for the reduction of computational complexity, and a linear interpolation technique for the generation of a training dataset. Firstly, the PCs scoring leads to a subset of 20-25 features that can be used as input for a standard statistical reconstruction algorithm, such as MLE, while retaining the spatial resolution capability of the system. Secondly, the procedure for collection of a training dataset was introduced and the data were used to generate DT classifiers, aiming at real-time inference of the position of interaction in the scintillation crystal. The DT classifier did not show performance loss with respect to the well consolidated MLE implementation. Lastly, the generation of artificial data was performed thanks to a Data Augmentation algorithm, leveraging the inverse transform sampling technique for the generation of interpolated data following the probability density functions of the values sampled on each pixel of the SPECT module. While a modest spatial resolution loss is present for the same quantization step, the training procedure is made simpler to implement and faster. Furthermore, it can be operated simultaneously on multiple modules. This work envisions the possibility to adopt ML-based data inference algorithms and to move the computational complexity towards the edge of a γ imaging system, with potential benefits in terms of relaxation of data bandwidth, reduction of system complexity, and increase of count rate.

Acknowledgements

All authors declare that they have no known conflicts of interest in terms of competing financial interests or personal relationships that could have an influence or are relevant to the work reported in this paper.

References

[1] K. Gong, E. Berg, S. R. Cherry, and J. Qi, "Machine learning in PET: from photon detection to quantitative image reconstruction," *Proceedings of the IEEE*, vol. 108, no. 1, pp. 51–68, 2019.

[2] A. S. Lundervold and A. Lundervold, "An overview of deep learning in medical imaging focusing on MRI," *Zeitschrift für Medizinische Physik*, vol. 29, no. 2, pp. 102–127, 2019.

[3] G. Wang, J. C. Ye, K. Mueller, and J. A. Fessler, "Image reconstruction is a new frontier of machine learning," *IEEE transactions on medical imaging*, vol. 37, no. 6, pp. 1289–1296, 2018.

[4] M. N. Wernick, Y. Yang, J. G. Brankov, G. Yourganov, and S. C. Strother, "Machine learning in medical imaging," *IEEE signal processing magazine*, vol. 27, no. 4, pp. 25–38, 2010.

[5] H. Greenspan, B. Van Ginneken, and R. M. Summers, "Guest editorial deep learning in medical imaging: Overview and future promise of an exciting new technique," *IEEE transactions on medical imaging*, vol. 35, no. 5, pp. 1153–1159, 2016.

[6] F. Simon, "Silicon photomultipliers in particle and nuclear physics," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 926, pp. 85–100, 2019.

[7] M. Carminati, G. L. Montagnani, M. Occhipinti, A. Kuehne, T. Niendorf, K. Nagy, A. Nagy, M. Czeller, and C. Fiorini, "SPECT/MRI INSERT compatibility: Assessment, solutions, and design guidelines," *IEEE Transactions on Radiation and Plasma Medical Sciences*, vol. 2, no. 4, pp. 369–379, 2018.

[8] H. Anger, "Use of a gamma-ray pinhole camera for in vivo studies," *Nature*, vol. 170, no. 4318, pp. 200–201, 1952.

[9] H. O. Anger, "Sensitivity, resolution, and linearity of the scintillation camera," *IEEE Transactions on nuclear science*, vol. 13, no. 3, pp. 380–392, 1966.

[10] M. C. Maas, D. R. Schaart, D. J. J. van der Laan, P. Bruyndonckx, C. Lemaitre, F. J. Beekman, and C. W. E. van Eijk, "Monolithic scintillator PET detectors with intrinsic depth-of-interaction correction," *Physics in Medicine and Biology*, vol. 54, no. 7, pp. 1893–1908, mar 2009. [Online]. Available: https://doi.org/10.1088/0031-9155/54/7/003

[11] C. W. Lerche, J. Benlloch, F. Sanchez, N. Pavon, B. Escat, E. Gimenez, M. Fernandez, I. Torres, M. Gimenez, A. Sebastia *et al.*, "Depth of/spl gamma/-ray interaction within continuous crystals from the width of its scintillation light-distribution," *IEEE Transactions on Nuclear Science*, vol. 52, no. 3, pp. 560–572, 2005.

[12] T. Ling, T. K. Lewellen, and R. S. Miyaoka, "Depth of interaction decoding of a continuous crystal detector module," *Physics in Medicine and Biology*, vol. 52, no. 8, pp. 2213–2228, mar 2007. [Online]. Available: https://doi.org/10.1088/0031-9155/52/8/012

[13] W. C. Hunter, H. H. Barrett, and L. R. Furenlid, "Calibration method for ML estimation of 3D interaction position in a thick gamma-ray detector," *IEEE transactions on nuclear science*, vol. 56, no. 1, pp. 189–196, 2009.

[14] H. T. van Dam, S. Seifert, R. Vinke, P. Dendooven, H. Löhner, F. J. Beekman, and D. R. Schaart, "A practical method for depth of interaction determination in monolithic scintillator PET detectors," *Physics in Medicine and Biology*, vol. 56, no. 13, pp. 4135–4145, jun 2011. [Online]. Available: https://doi.org/10.1088/0031-9155/56/13/025

[15] I. D'Adda, A. J. Morahan, M. Carminati, K. Erlandsson, M. Ljungberg, B. F. Hutton, and C. Fiorini, "A Statistical DOI Estimation Algorithm for a SiPM-Based Clinical SPECT Insert," *IEEE Transactions on Radiation and Plasma Medical Sciences*, vol. 6, no. 7, pp. 771–777, 2022.

[16] L. A. Shepp and Y. Vardi, "Maximum likelihood reconstruction for emission tomography," *IEEE transactions on medical imaging*, vol. 1, no. 2, pp. 113–122, 1982.

[17] P. Bruyndonckx, C. Lemaitre, D. Van Der Laan, M. Maas, D. Schaart, W. Yonggang, Z. Li, M. Krieguer, and S. Tavernier, "Evaluation of machine learning algorithms for localization of photons in undivided scintillator blocks for PET detectors," *IEEE Transactions on Nuclear Science*, vol. 55, no. 3, pp. 918–924, 2008.

[18] P. Bruyndonckx, S. Léonard, S. Tavernier, C. Lemaître, O. Devroede, Y. Wu, and M. Krieguer, "Neural network-based position estimators for PET detectors using monolithic LSO blocks," *IEEE Transactions on Nuclear Science*, vol. 51, no. 5, pp. 2520–2525, 2004.

[19] S. Tavernier, P. Bruyndonckx, S. Leonard, and O. Devroede, "A high-resolution PET detector based on continuous scintillators," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 537, no. 1-2, pp. 321–325, 2005.

[20] P. G. De Acilu, I. Sarasola, M. Canadas, R. Cuerdo, P. R. Mendes, L. Romero, and C. Willmott, "Study and optimization of positioning algorithms for monolithic PET detectors blocks," *Journal of Instrumentation*, vol. 7, no. 06, p. C06010, 2012.

[21] Y. Wang, W. Zhu, X. Cheng, and D. Li, "3D position estimation using an artificial neural network for a continuous scintillator PET detector," *Physics in Medicine & Biology*, vol. 58, no. 5, p. 1375, 2013.

[22] F. Müller, D. Schug, P. Hallen, J. Grahe, and V. Schulz, "Gradient tree boosting-based positioning method for monolithic scintillator crystals in positron emission tomography," *IEEE Transactions on Radiation and Plasma Medical Sciences*, vol. 2, no. 5, pp. 411–421, 2018.

[23] F. Müller, D. Schug, P. Hallen, J. Grahe, and V. Schulz, "A novel DOI positioning algorithm for monolithic scintillator crystals in PET based on gradient tree boosting," *IEEE Transactions on Radiation and Plasma Medical Sciences*, vol. 3, no. 4, pp. 465–474, 2018.

[24] X. Zhang, J. Zhou, S. R. Cherry, R. D. Badawi, and J. Qi, "Quantitative image reconstruction for total-body PET imaging using the 2-meter long EXPLORER scanner," *Physics in Medicine and Biology*, vol. 62, no. 6, pp. 2465–2485, feb 2017. [Online]. Available: https://doi.org/10.1088/1361-6560/aa5e46

[25] S. Siegel, R. W. Silverman, Y. Shao, and S. R. Cherry, "Simple charge division readouts for imaging scintillator arrays using a multi-channel PMT," *IEEE Transactions on Nuclear Science*, vol. 43, no. 3, pp. 1634–1641, 1996.

[26] X. Li, C. Lockhart, T. K. Lewellen, and R. S. Miyaoka, "Study of PET detector performance with varying SiPM parameters and readout schemes," *IEEE Transactions on Nuclear Science*, vol. 58, no. 3, pp. 590–596, 2011.

[27] Y. C. Shih, F. W. Sun, L. R. MacDonald, B. P. Otis, R. S. Miyaoka, W. McDougald, and T. K. Lewellen, "An 8×8 row-column summing readout electronics for preclinical positron emission tomography scanners," in *IEEE Nuclear Science Symposium Conference Record*, 2009.

[28] R. S. Miyaoka, T. Ling, and T. K. Lewellen, "Effect of number of readout channels on the performance of a continuous miniature crystal element (cMiCE) detector," in *IEEE Nuclear Science Symposium Conference Record*, 2006.

[29] L. Pierce, W. Hunter, D. Haynor, L. MacDonald, P. Kinahan, and R. Miyaoka, "Multiplexing strategies for monolithic crystal PET detector modules," *Physics in Medicine & Biology*, vol. 59, no. 18, p. 5347, 2014.

[30] P. Bouziotis and C. Fiorini, "SPECT/MRI: dreams or reality?" *Clinical and Translational Imaging*, vol. 2, no. 6, pp. 571–573, 2014.

[31] A. D. Butt, Z. Nyitrai, P. Busca, C. Fiorini, K. Nagy, and M. Occhipinti, "Development of an MRI-compatible cooling unit for SPECT/MRI detection modules," in *2015 IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC)*. IEEE, 2015, pp. 1–3.

[32] V. Regazzoni, F. Acerbi, G. Cozzi, A. Ferri, C. Fiorini, G. Paternoster, C. Piemonte, D. Rucatti, G. Zappalà, N. Zorzi *et al.*, "Characterization of high density SiPM non-linearity and energy resolution for prompt gamma imaging applications," *Journal of Instrumentation*, vol. 12, no. 07, p. P07001, 2017.

[33] C. E. Aalseth, F. Acerbi, P. Agnes, I. Albuquerque, T. Alexander, A. Alici, A. Alton, P. Ampudia, P. Antonioli, S. Arcelli *et al.*, "Cryogenic Characterization of FBK RGB-HD SiPMs," *Journal of Instrumentation*, vol. 12, no. 09, p. P09030, 2017.

[34] M. Carminati, I. D'Adda, A. Morahan, K. Erlandsson, K. Nagy, M. Czeller, B. Tölgyesi, Z. Nyitrai, A. Savi, P. van Mullekom *et al.*, "Clinical SiPM-based MRI-compatible SPECT: Preliminary characterization," *IEEE Transactions on Radiation and Plasma Medical Sciences*, vol. 4, no. 3, pp. 371–377, 2019.

[35] M. Occhipinti, M. Carminati, P. Busca, A. Butt, G. Montagnani, P. Trigilio, C. Piemonte, A. Ferri, A. Gola, T. Bukki *et al.*, "Characterization of the detection module of the INSERT SPECT/MRI clinical system," *IEEE Transactions on Radiation and Plasma Medical Sciences*, vol. 2, no. 6, pp. 554–563, 2018.

[36] R. M. Gray and A. Macovski, "Maximum a posteriori estimation of position in scintillation cameras," *IEEE Transactions on Nuclear Science*, vol. 23, no. 1, pp. 849–852, 1976.

[37] J. Joung, R. S. Miyaoka, S. Kohlmyer, and T. K. Lewellen, "Implementation of ML based positioning algorithms for scintillation cameras," *IEEE Transactions on Nuclear Science*, vol. 47, no. 3, pp. 1104–1111, 2000.

[38] H. H. Barrett, W. C. Hunter, B. W. Miller, S. K. Moore, Y. Chen, and L. R. Furenlid, "Maximum-likelihood methods for processing signals from gamma-ray detectors," *IEEE transactions on nuclear science*, vol. 56, no. 3, pp. 725–735, 2009.

[39] L. Pierce, S. Pedemonte, D. DeWitt, L. MacDonald, W. Hunter, K. Van Leemput, and R. Miyaoka, "Characterization of highly multiplexed monolithic PET/gamma camera detector modules," *Physics in Medicine & Biology*, vol. 63, no. 7, p. 075017, 2018.

[40] S. Pedemonte, L. Pierce, and K. Van Leemput, "A machine learning method for fast and accurate characterization of depth-of-interaction gamma cameras," *Physics in Medicine & Biology*, vol. 62, no. 21, p. 8376, 2017.

[41] S. J. Wilderman, N. H. Clinthorne, J. A. Fessler, and W. L. Rogers, "List-mode maximum likelihood reconstruction of Compton scatter camera images in nuclear medicine," in *1998 IEEE Nuclear Science Symposium Conference Record. 1998 IEEE Nuclear Science Symposium and Medical Imaging Conference (Cat. No. 98CH36255)*, vol. 3. IEEE, 1998, pp. 1716–1720.

[42] M. Fontana, J.-L. Ley, D. Dauvergne, N. Freud, J. Krimmer, J. M. Létang, V. Maxim, M.-H. Richard, I. Rinaldi, and É. Testa, "Monitoring ion beam therapy with a Compton camera: simulation studies of the clinical feasibility," *IEEE Transactions on Radiation and Plasma Medical Sciences*, vol. 4, no. 2, pp. 218–232, 2019.

[43] V. Solovov, V. Belov, D. Y. Akimov, H. Araujo, E. Barnes, A. Burenkov, V. Chepel, A. Currie, L. DeViveiros, B. Edwards *et al.*, "Position reconstruction in a dual phase xenon scintillation detector," *IEEE Transactions on Nuclear Science*, vol. 59, no. 6, pp. 3286–3293, 2012.

[44] A. Morozov, V. Solovov, F. Alves, V. Domingos, R. Martins, F. Neves, and V. Chepel, "Iterative reconstruction of detector response of an Anger gamma camera," *Physics in Medicine & Biology*, vol. 60, no. 10, pp. 4169–4184, 2015.

[45] A. Morozov, F. Alves, J. Marcos, R. Martins, L. Pereira, V. Solovov, and V. Chepel, "Iterative reconstruction of SiPM light response functions in a square-shaped compact gamma camera," *Physics in Medicine and Biology*, vol. 62, no. 9, p. 3619, 2017.

[46] S. Pedemonte, A. Gola, A. Abba, and C. Fiorini, "Optimum real-time reconstruction of Gamma events for high resolution Anger camera with the use of GPGPU," in *2009 IEEE Nuclear Science Symposium Conference Record (NSS/MIC)*. IEEE, 2009, pp. 3388–3394.

[47] F. Acerbi and S. Gundacker, "Understanding and simulating sipms," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 926, pp. 16–35, 2019.

[48] M. Ringnér, "What is principal component analysis?" *Nature biotechnology*, vol. 26, no. 3, pp. 303–304, 2008.

[49] B. M. Kandel, D. J. Wang, J. C. Gee, and B. B. Avants, "Eigenanatomy: Sparse dimensionality reduction for multi-modal medical image analysis," *Methods*, vol. 73, pp. 43–53, 2015.

[50] C. Parmar, J. D. Barry, A. Hosny, J. Quackenbush, and H. J. Aerts, "Data analysis strategies in medical imaging," *Clinical cancer research*, vol. 24, no. 15, pp. 3492–3499, 2018.

[51] I. Guyon and A. Elisseeff, "An introduction to variable and feature selection," *Journal of machine learning research*, vol. 3, no. Mar, pp. 1157–1182, 2003.

[52] G. T. Reddy, M. P. K. Reddy, K. Lakshmanna, R. Kaluri, D. S. Rajput, G. Srivastava, and T. Baker, "Analysis of dimensionality reduction techniques on big data," *IEEE Access*, vol. 8, pp. 54 776–54 788, 2020.

[53] R. B. Cattell, "The scree test for the number of factors," *Multivariate Behavioral Research*, vol. 1, no. 2, pp. 245–276, 1966, pMID: 26828106. [Online]. Available: https://doi.org/10.1207/s15327906mbr0102_10

[54] A. Morozov, V. Solovov, R. Martins, F. Neves, V. Domingos, and V. Chepel, "ANTS2 package: simulation and experimental data processing for Anger camera type detectors," *Journal of Instrumentation*, vol. 11, no. 04, p. P04022, 2016.

[55] Z. Sun, G. Pedretti, E. Ambrosi, A. Bricalli, W. Wang, and D. Ielmini, "Solving matrix equations in one step with cross-point resistive arrays," *Proceedings of the National Academy of Sciences*, vol. 116, no. 10, pp. 4123–4128, 2019.

[56] L. Rokach and O. Maimon, *Decision Trees*. Boston, MA: Springer US, 2005, pp. 165–192. [Online]. Available: https://doi.org/10.1007/0-387-25465-X_9

[57] J. MacQueen *et al.*, "Some methods for classification and analysis of multivariate observations," in *Proceedings of the fifth Berkeley symposium on mathematical statistics and probability*, vol. 1, no. 14. Oakland, CA, USA, 1967, pp. 281–297.

[58] M. Kayri and İ. Kayri, "The comparison of gini and twoing algorithms in terms of predictive ability and misclassification cost in data mining: an empirical study," *International Journal of Computer Trends and Technology*, vol. 27, no. 1, pp. 21–30, 2015.

[59] X. An, A. Artemyev, V. Angelopoulos, S. Lu, P. Pritchett, and V. Decyk, "Fast inverse transform sampling of non-gaussian distribution functions in space plasmas," *Journal of Geophysical Research: Space Physics*, vol. 127, no. 5, pp. 1–10, 2022.