

A Novel DOI Positioning Algorithm for Monolithic Scintillator Crystals in PET Based on Gradient Tree Boosting

Florian Müller¹, David Schug¹, Patrick Hallen, Jan Grahe, and Volkmar Schulz

Abstract—Monolithic crystals are examined as an alternative to segmented scintillator arrays in positron emission tomography (PET). Monoliths provide good energy, timing, and spatial resolution including intrinsic depth of interaction (DOI) encoding. DOI allows reducing parallax errors (radial astigmatism) at off-center positions within a PET ring. We present a novel DOI-estimation approach based on the supervised machine learning algorithm gradient tree boosting (GTB). GTB builds predictive regression models based on sequential binary comparisons (decision trees). GTB models have been shown to be implementable in FPGA if the memory requirement fits the available resources. We propose two optimization scenarios for the best possible positioning performance: One restricting the available memory to enable a future FPGA implementation and one without any restrictions. The positioning performance of the GTB models is compared with a DOI estimation method based on a single DOI observable (SO) comparable to other methods presented in literature. For a 12 mm high monolith, we achieve an averaged spatial resolution of 2.15 mm and 2.12 mm FWHM for SO and GTB models, respectively. In contrast to SO models, GTB models show a nearly uniform positioning performance over the whole crystal depth.

Index Terms—Depth of interaction (DOI), FPGA, gradient tree boosting (GTB), machine learning, monolithic scintillator, positron emission tomography (PET).

I. INTRODUCTION

POSITRON emission tomography (PET) is a functional imaging technique with a variety of applications in both preclinical as well as clinical research and practice [1], [2]. Two 511 keV gamma particles originating from a positron-electron annihilation are registered by radiation detectors arranged in a ring geometry. State-of-the-art radiation detectors consist of scintillation crystals (e.g., BGO, LSO, and LYSO) converting the gamma particles to optical photons and photosensor arrays with multiple channels detecting the optical photons.

Manuscript received August 15, 2018; revised October 6, 2018; accepted November 20, 2018. Date of publication November 30, 2018; date of current version July 1, 2019. This work was supported by the German Research Foundation under Project SCHU 2973/2-1. (Corresponding author: Florian Müller.)

The authors are with the Department of Physics of Molecular Imaging Systems, Institute for Experimental Molecular Imaging, RWTH Aachen University, 52074 Aachen, Germany (e-mail: florian.mueller@pmi.rwth-aachen.de).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TRPMS.2018.2884320

The key challenge in PET detector instrumentation is to detect the gamma particles with high sensitivity and with good spatial, energy, and time resolution. The knowledge of depth of interaction (DOI) reduces the parallax error (radial astigmatism) at off-center positions within a PET ring [3]. DOI information is especially important for PET systems with a small ring diameter such as preclinical devices or organ-specific applications imaging the brain or female breast [4]–[6]. Recently, the advantages of DOI information have also been experimentally validated for a clinical case with a 70 cm diameter tomographic setup [7]. Furthermore, DOI is expected to also improve the performance of long-axial field-of-view PET systems (total body PET) [8].

Several detector concepts have been demonstrated in literature which can be mainly divided up into two groups: 1) pixelated and 2) monolithic detectors. Most state-of-the-art clinical PET scanners are operated with pixelated detector designs. These detectors consist of arrays of scintillator needles and are read out employing light-sharing techniques or one-to-one coupling. None of the currently used clinical detectors for whole-body PET provide DOI information [9]. Several adaptations of pixelated detectors have been presented to obtain DOI information. Among others, multiple-layer designs with small shifts between each layer [10], light-sharing techniques encoding the DOI information into the light distribution [11], combination of scintillation layers with different properties (phoswich detectors) [12], subsurface laser engraving [13], and dual-sided readout [14] were proposed. However, often these concepts come at the cost of other key performance parameters or introduce additional complexity and cost to the detector design.

As an alternative, monolithic detectors consist of scintillators without any segmentation coupled directly to a photosensor array. Monoliths are able to provide good spatial, timing, and energy resolutions as widely shown in [15]–[19]. Furthermore, monolithic detectors enable intrinsic DOI encoding. For example, the 3-D position can be estimation based on a fit of a model of the light distribution [20]. Instead of fitting to a model, maximum likelihood searches based on 3-D training data have also been successfully demonstrated [6], [18]. A different approach employed an analytic expression based on the attenuation of the scintillator and the expected light-transport [21], [22]. After fitting two free parameters, the ratio of event energy to the maximum local intensity was utilized as DOI observable. A method based on a single DOI

observable only including the photon counts—the sum of the squared pixel intensities (SPIs)—was applied to both single and dual-sided read-out [15], [23]. The expected distribution of events according to the attenuation was simulated and then matched to the distribution of the DOI observable. The main principle of this method testing several DOI observables was first published in [24]. Besides the presented methods, also neural network estimators based on training data acquired by a side irradiation of the scintillator were shown [25]. To widely translate monolithic scintillators into applications with a large number of detectors, time-efficient, and easy calibration methods are required. Furthermore, all employed algorithms such as position estimation need to be scalable for a large number of detectors.

Recently, we presented a novel planar positioning algorithm based on the supervised machine learning technique gradient tree boosting (GTB) [26]. Here, we present the adaptation and application of this algorithm to DOI calibration by a side irradiation of the scintillator. GTB builds a set of sequential binary decisions (decision trees) which are evaluated as simple comparisons with two possible outcomes. The algorithm handles different sets of input features and their combinations as well as partially missing data. Trained GTB models are shown to be implementable in FPGA if the memory requirement (MR) does not exceed the available resources [27], [28]. Besides a general description of the hyperparameter tuning of GTB models, we present two optimization scenarios to find the best possible positioning performance: one restricting the available memory to enable a future FPGA implementation and one without any restrictions. The positioning performance is compared to a DOI-approach using a single DOI observable similar to [15] as well as results presented in literature.

II. MATERIALS

As the same materials and detector were used as presented in [26], only a brief description of the single components is given in the corresponding sections. We utilized the Philips Digital Photon Counting (PDPC) technology evaluation kit with two sensor tiles of DPC 3200-22 photon counter as a coincidence setup. For calibration, a fan beam collimator providing adjustable beam widths for both coincidence and detector under study was employed. The fan beam collimator consists of a fixed bottom shielding, a tool housing up to two sources, and two adjustable top shielding units defining the beam (see Fig. 1). The whole setup was operated in a light-tight temperature chamber at approx. 5 °C sensor temperature.

A. Photodetector

We used an array, referred to as tile, of 4×4 digital silicon photomultiplier (dSiPM) DPC 3200-22 of PDPC [29], [30]. Each DPC is an independent trigger region consisting of four pixels with a pitch of 4 mm resulting in a total of 64 pixels per tile. Every DPC provides a customizable two-level trigger scheme: After the first trigger signal is generated, the second, higher threshold needs to be reached within a given

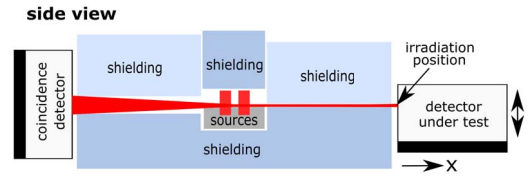


Fig. 1. Sketch of the setup containing the fan beam collimator, coincidence detector, and detector under test. For DOI calibration, a side irradiation of the detector under test is performed. The irradiation position is determined by the feed-back loop of the stepper motor.

time interval as well. For our settings, on average 2.33 and 17 photons need to be detected to reach the first and second trigger threshold. In case the second trigger condition is met, the integration phase starts and all four pixels of the DPC are read out afterward. The collected information of a DPC is referred to as a hit. The entity of all hits corresponding to one gamma particle interaction is called a cluster as described in more detail in Section III-A. For a single gamma particle interaction, not all 16 DPCs of the tile will certainly output hit data. This leads to clusters with missing hits as reported in Section IV-A, especially if the photon densities are low. Further information regarding the photosensor can be found in [31]–[34].

B. Scintillator Crystal and Wrapping

We studied a monolithic LYSO scintillator of dimensions $32 \text{ mm} \times 32 \text{ mm} \times 12 \text{ mm}$ matching the active sensor area of the tile. The monolith was wrapped in highly reflective Teflon tape (Klinger, Idstein, Germany) and coupled to the photosensor using a two-component dielectric silicon gel (Sylgard 527, Dow Corning, Midland, Michigan, USA). We chose the reflective wrapping to achieve a high light-output of the scintillator. An optical simulation studying the influence of several scintillator wrappings (e.g., black tape) on the position performance of GTB-models can be found in [35]. To register coincidences, we employed a 12 mm-high pixelated array with 1 mm pitch also utilized in [36] and [37].

C. Collimator Setup

The setup contained a fan beam collimator equipped with two ^{22}Na sources of approximately 10 MBq each and an electrically driven translation stage (LIMES 90, Owis, Staufen im Breisgau, Germany) (see Fig. 1). All measurements were conducted at a sensor temperature of approx. 5 °C achievable on system level as demonstrated in [38]. The fan beam collimator reached a coincidence rate of 199 Hz for slit widths of 5 and 0.25 mm for the coincidence detector and the detector under study, respectively. We chose a much larger slit width for the coincidence detector to avoid losing coincidence events due to geometrical effects and uncertainties of the collimator setup. The slit width of 0.25 mm for the detector under study translates to a beam width of 0.42 mm FWHM at the crystal surface determined by stepping a crystal edge through the beam [26]. The translation stage was connected to a control PC and the current position of the detector under test was

recorded. This irradiation position was employed as reference in later analysis.

III. METHODS

A. Data Acquisition and Preprocessing

The data acquisition aimed to create datasets for establishing and testing DOI positioning models based on a side irradiation of the detector. However, a side irradiation leads to an exponential distribution of the gamma particle interactions along the propagation direction according to Lambert–Beer’s law. For the employed setup (see Fig. 1), we expected an exponential distribution along the x -direction and a uniform distribution along the y -direction. To achieve reliable DOI positioning models and estimates of the positioning performance, datasets with uniformly distributed events in both x - and y -direction are strongly beneficial. Otherwise, the exponential distribution may mislead machine learning-based algorithms during training of DOI models and weights the positioning performance achieved over the whole crystal volume nonproportional. Therefore, we need a planar positioning algorithm to estimate the gamma interaction position of the events obtained with the side irradiation. Subsequently, the full calibration process included irradiations of the crystal’s top surface for planar positioning calibration and a side irradiation for DOI positioning calibration. Both x - and y -direction of the crystal were irradiated at parallel lines with a pitch of 0.75 mm for the planar positioning calibration. The side irradiation was performed with a pitch of 0.25 mm.

All conducted measurements shared the initial preprocessing based on a tool developed by Schug *et al.* [36]: the collected optical photons of one gamma particle interaction create up to 16 hits in each of the detectors. All hits associated to one gamma particle interaction were identified and merged to clusters. We applied a cluster window of 40 ns. The timestamp of the earliest hit determined the timestamp of the cluster. Afterward, coincident clusters were searched using a sliding coincidence window of 20 ns. Pixels missing in a cluster were marked with a negative value to distinguish them from zero photon counts.

Clusters with a total photon count fewer than 700 photons were rejected to remove noisy events from the data. Based on the total photon distribution, this photon cut translates to an energy cut of around 300 keV. No further quality cuts were applied to the data. Accepted clusters are called events.

In the next step, GTB-based planar positioning models were established employing the irradiations of the crystal’s top surface as described in [26]. All events acquired by the side irradiation were positioned along the planar directions. Then, a subset uniformly distributed along x - and y -direction was chosen for every irradiation position of the side irradiation as motivated above. The resulting data were split up into three data sets: 1) training data for building the DOI positioning models; 2) validation data for optimization of the hyperparameter of the DOI positioning models; and 3) test data for the final evaluation of the positioning performance. Each data set contained 10 000 events per irradiation position of the side irradiation. For the training data, both the binning of the training

data as well as the number of events per irradiation position were varied later on aiming for short calibration times. We selected a minimum number of 100 training events per irradiation position due to practical reasons: for 100 training events, the required time for moving the translation stage and operations of the control PC are already the dominating factor compared to an irradiation time of around 0.5 s.

B. Performance Parameters

All following positioning performance parameters are based on the positioning error distribution (irradiation position–estimated position) calculated for every single irradiation position. If a single value is given, the performance parameter is averaged for all irradiation positions.

- 1) *Bias Vector*: The bias vector is defined as the mean positioning error for a given position. Due to the edge effects found in monolithic crystals, the bias vector magnitude distribution is non-Gaussian. 50th and 90th percentile of the bias vector magnitude distribution (Bias_{50} and Bias_{90}) are given to account for both the central part and tails of the distribution.
- 2) *Spatial Resolution (SR)*: The SR is defined as the FWHM of the positioning error distribution. We calculate the SR in accordance to the NEMA NU 4-2008 procedure [39]. The SR is not corrected for the finite beam width of the collimator.
- 3) *Root Mean Squared Error (RMSE)*: The RMSE is the root of the mean squared positioning error.
- 4) *Mean Absolute Error (MAE)*: The MAE is the mean of the absolute positioning error.
- 5) *Percentile Distance d_x* : The percentile distance is defined as the distance enclosing the given percentile x of all events around an irradiation position. We report the 50th and 90th percentile distance.
- 6) *Score of Distance 1.5 mm*: The score of distance 1.5 mm is the fraction of events which are assigned a position within 1.5 mm around the irradiation position.

While MAE, RMSE, percentile distances, and the score value are sensitive to bias effects, SR is not prone to a global shift of the positioning error distribution.

C. Single Observable DOI Estimation

This DOI estimation method utilizes a single observable (SO) calculated based on the measured light distribution as demonstrated in [23] and [24]. In contrast to methods presented in literature, the observable must not include more than the four pixels of one DPC to avoid jitter caused by missing hits in an event. We defined a set of possible observables and examined the correlation with the DOI position. The set included: 1) the fraction of the highest photon count to the total photon count of the hottest DPC and 2) the sum of the SPIs of the hottest DPC. The crystal was divided into equally sized segments with their own calibration to account for differences in the detector response. The number of segments ranged from 1 to 32 along both planar directions.

To match the observable to the DOI position, multiple methods such as lookup tables and fits to polynomials of

higher order are possible. For this paper, we employed isotonic regression (IR). IR minimizes the mean squared error for monotone data without assuming any form of target function [40], [41]. Thus, this method is preferable if no physical model is present.

As this method is a benchmark to compare the GTB-based DOI estimation models, only results of the best performing observable and segmentation are shown. The averaged performance parameters as well as the spatial distributions of bias vector, MAE and SR are reported. Tests reducing the number of training events and irradiation positions aiming for short calibration times are out of the scope of this paper and are not presented.

D. Gradient Tree Boosting DOI Estimation

A detailed description of the GTB algorithm for planar positioning is given in [26]. Thus, only the main characteristics of the algorithm and hyperparameters used later on are described.

As part of supervised machine learning techniques, GTB utilizes training data with known irradiation positions to establish predictive regression models. The algorithm handles missing data for training and evaluation and can be used with arbitrary input features [42]. GTB builds a set of sequential binary decisions (decision trees) which are evaluated as simple comparisons with two possible outcomes [43], [44]. The ensemble is trained in an additive manner: the first decision tree is based on the given irradiation position. Every following decision tree is trained on the positioning error (irradiation position–estimated position) of the previous ensemble. We employed RMSE as training loss of the objective function.

Four hyperparameters of GTB models are of particular importance for the following optimizations.

- 1) Number of decision trees of the ensemble.
- 2) *Maximum Depth*: The maximum number of comparisons in a single decision tree.
- 3) *Learning Rate*: The learning rate multiplies the positioning error of the already established ensemble with a constant factor less or equal 1 for training of the next decision tree as described in [26]. This allows to reduce the number of decision trees while reducing the highest achievable positioning performance [45].
- 4) *Features of the Input Set*: In addition to the 64 raw photon counts, further features motivated by the physical properties of the problem can be added to improve the positioning performance. Used features and input sets are defined in the following section.

Evaluation of trained GTB models is possible in CPU-, GPU-, and FPGA-based architectures [28]. All listed architectures should allow for real-time event processing. This paper focusses on the possibility of an FPGA implementation because FPGAs are widely employed at several points in the current and future architecture developed in this group [38], [46]. FPGA-based processing combined with the data acquisition allows to significantly decrease the amount of data which needs to be sent to the connected server reducing the bandwidth requirements. Trained GTB models are implementable in FPGA if the MR does not exceed the available

resources [27], [28]. The MR of a single decision tree of a GTB model can be estimated by

$$\text{MR}(d) = (2^d - 1) \cdot 11 B + 2^d \cdot 6 B \quad (1)$$

with d the maximum depth [26].

We present a general optimization protocol examining the influence of the single hyperparameters. Furthermore, two optimization scenarios of GTB models for the best possible positioning performance are demonstrated: One restricting the available memory to enable a future FPGA implementation and one without any restrictions.

1) *General Optimization Process*: We adapted the developed optimization protocol presented in [26] to the DOI problem. First, a suitable start point was searched testing similar settings found for the planar positioning as in [26]. Aiming for short calibration times, we studied the influence of the binning of the irradiation positions as well as of the number of training events per irradiation position. The binning of the irradiation positions ranged from 0.25 to 3 mm and the number of events per irradiation position from 100 to 10 000. Then, always one of the hyperparameters introduced in the previous section was varied keeping all other parameters constant. We tested maximum depths from 4 to 12, learning rates from 0.05 to 0.7 and three different input sets defined below. The parameter ranges were chosen according to the results of the planar positioning optimization and suggestions found in [26] and [44]. In general, GTB models were trained for 1000 decision trees and tested with the validation data set.

To study the influence of the input sets, we validated three possible sets of input features.

- 1) *Raw Data*: the 64 photon counts.
- 2) *Raw Data and Calculated Features (CFs)*: CFs included the index numbers of the hottest pixel and DPC, first and second moment of the light distribution, both defined DOI observables, the total photon sum and projections of the photon counts along both planar directions.
- 3) *Raw data, CF and estimated planar interactions positions*. In contrast to input set 3), input sets 1) and 2) do not require the estimated planar interaction position for evaluating events.

Thus, GTB models trained with these input sets can be fully parallelized together with the planar positioning models which could be beneficial for processing data of a full PET system. We present results for the different input sets for all tested maximum depths ranging from 4 to 12. To represent models with low and high MRs and study the influence of the input sets during the training process, ensembles of 50 and 1000 decision trees were validated.

For all optimization steps, we chose to present the averaged MAE as validation metric to account for bias effects and allow a direct comparison with other publications [15], [23]. Additionally, an overlay and a difference plot of the spatial distribution of the MAE for the varied binning of irradiation positions are shown. The spatial distribution indicates if GTB models have working regression capabilities: for a working regression model, no bias toward the irradiation positions used for training should be observed.

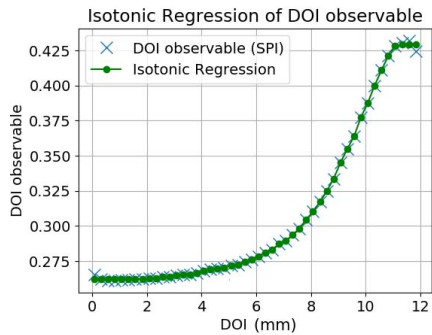


Fig. 2. Course of the DOI observable SPI and IR. The DOI observable monotonely increases with the DOI position. DOI position 0 mm represents the top surface of the crystal and 12 mm the photosensor.

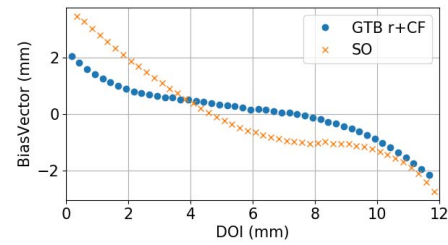
TABLE I

OVERVIEW OF AVERAGED PERFORMANCE PARAMETERS FOR THE BEST FOUND SO MODEL AND HIGH PERFORMANCE GTB MODELS FOR ALL THREE INPUT SETS [RAW DATA (R); RAW DATA AND CFS (R + CF); RAW DATA, CALCULATED FEATURES, AND ESTIMATED PLANAR INTERACTION POSITIONS (R + CF + POS)]

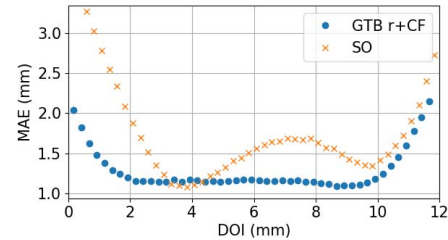
	SO		GTB	
		r	r+CF	r+CF+pos
RMSE / mm	2.22	1.86	1.81	1.81
MAE / mm	1.73	1.32	1.28	1.28
SR / mm	2.15	2.26	2.12	2.09
d_{50} / mm	1.35	0.92	0.84	0.88
d_{90} / mm	3.60	3.01	2.91	2.90
Bias ₅₀ / mm	1.06	0.54	0.55	0.55
Bias ₉₀ / mm	2.62	1.74	1.67	1.66
Score of 1.5 mm	0.54	0.70	0.71	0.72

2) *High-Performance Optimization*: No memory restrictions were applied to select the best-performing DOI positioning models. We elected to pick those models with the best averaged RMSE value for all three tested input sets as the RMSE is used as loss function during model training as well. First, the minimum averaged RMSE was searched for every GTB model trained during the general optimization process. In case the minimum RMSE was found for the maximum trained number of decision trees (1000 trees), we continued training until the respective GTB model tend to overfit. The GTB model was assumed to be overfitting, if the RMSE started to worsen. Second, the best averaged RMSE of all models was searched. The averaged performance parameters of these GTB models are presented. Furthermore, the spatial distributions of bias vector, MAE, and SR are plotted for the best-performing model with raw data and CFs as input.

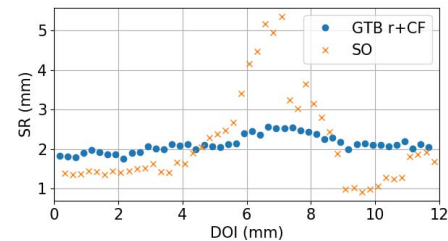
3) *Memory-Requirement Optimization*: To enable an FPGA implementation, the GTB model with the best possible positioning performance for a given memory restriction is searched. We trained the GTB models for combinations of maximum depth and learning rate while an empirically chosen convergence criterion determined the number of decision trees. No further decision trees were added if the averaged MAE did not improve more than 0.0001 mm/decision tree. For high learning rates, the chosen convergence criterion may be too conservative and takes action in the overfitting regime. Thus, the found positioning performance is compared to those



(a)



(b)



(c)

Fig. 3. Spatial distribution of three performance parameters for the best performing SO models (denoted as SO) and GTB model with raw data and CFs as input (denoted as GTB r + CF). DOI position 0 mm represents the top surface of the crystal and 12 mm the photosensor. (a) Distribution of the bias vector. (b) Distribution of the MAE. (c) Distribution of the SR.

found for the high-performance optimization. In case the found number of decision trees is larger than those of the high-performance model, the cut gets discarded and the high-performance model of same maximum depth and learning rate is selected.

IV. RESULTS

A. Data Acquisition

The number of read out DPCs per gamma interaction followed a Gaussian distribution with a mean of 9.8 DPCs and a standard deviation of 1.8 DPCs. An irradiation time per irradiation position of around 5 min was required to measure 10 000 events uniformly distributed along the planar directions.

B. Single Observable DOI Estimation

Both tested DOI observables led to similar results. In general, the sum of SPIs performed a bit better (approx. 3%). Also, the segmentation along the planar directions improved the positioning performance with an optimum at eight segments along both directions. The difference of positioning performance between the best and worst DOI model was less than 2.5%. Fig. 2 shows the course of the chosen

DOI observable. Assuming a monotone behavior of the DOI observable is justified. An averaged MAE and SR of 1.73 and 2.15 mm FWHM were achieved (see Table I). Fig. 3 displays the spatial distribution of bias vector, MAE and SR. Over the whole crystal depth, a bias vector ranging from approx. ± 3 mm close to the edges to approx. 1 mm in the central area is observed. The MAE deteriorates toward the edges and shows an additional decrease of the positioning performance between 4 and 10 mm DOI position. The SO method achieves a SR better than 2 mm FWHM for 0 to 4.5 mm and 8.9 to 12 mm DOI position. However, the positioning performance significantly worsens for the other DOI positions to a maximum of 5.4 mm FWHM.

C. Gradient Tree Boosting DOI Estimation

1) *General Optimization Process:* We found a maximum depth of 10 and a learning rate of 0.1 with raw data of 5000 training events at 1 mm binning of the calibration grid as a suitable start point. The full optimization process is displayed in Fig. 4.

Fig. 4(a) shows the averaged MAE against the number of decision trees and different pitches of the calibration grid. In general, the MAE improves with an increasing number of decision trees. Only the model based on 3 mm binning shows slight overfitting effects starting at around 400 decision trees. The positioning performance difference is less than 13% between the 0.25 mm and 3 mm calibration grid evaluated at their respective optimum. For the calibration grids of 0.25 mm and 1 mm, a deviation of the positioning performance of 6% is found. The course of the spatial distribution of the MAE is generally the same for all tested calibration grids [see Fig. 4(b)]: the MAE is nearly constant in the central region of the crystal and deteriorates toward the crystal edges. Significant higher order effects such as additional periodicity in the spatial distribution of the MAE start to get visible for the 3 mm calibration grid. These effects are not visible for the 1 mm calibration grid; the MAE worsens globally. However, the positioning performance difference close to the crystal edges is less dominant [see Fig. 4(c)]. Due to reasons of clarity not shown in Fig. 4(a)–(c), further pitches of the calibration grid beside the shown ones were tested—namely 0.5, 0.75, 1.5, and 2 mm. All obtained results are in congruence with the plotted pitches of the calibration grid and follow the behavior of the shown curves of averaged and spatial MAE.

Fig. 4(d) displays the averaged MAE against the number of decision trees for different numbers of training events per irradiation position. The positioning performance increases if more events are employed for training the GTB models. However, only small improvements can be achieved after a sufficient amount of training events is present. Increasing the number of training events per irradiation position from 2500 to 5000 or 10000 improves the positioning performance less than 3.4% or 6%, respectively.

The maximum depth strongly influences the positioning performance for ensembles with a low number of decision trees [see Fig. 4(e)]; increasing the maximum depth improves the averaged MAE. Using higher maximum depths

than 10 show no significant enhancement. The positioning performance difference between the GTB models with different maximum depths vanishes for ensembles with a high number of decision trees and converges to an optimum.

Increasing the learning rate, the optimum of the GTB models is found for ensembles with less decision trees before overfitting occurs [see Fig. 4(f)]. For example, overfitting starts after less than ten decision trees for learning rate 0.7 while no overfitting is observed up to 1000 decision trees for learning rate 0.05. At the same time, the best possible positioning performance decreases for higher learning rates.

The influence of different input sets at several maximum depths is shown for ensembles of 50 and 1000 decision trees in Fig. 4(g) and (h), respectively. For the GTB model with 50 decision trees, the input sets including the CFs and planar position improve the averaged MAE for all tested maximum depths. The difference between the MAE gets smaller increasing the maximum depth. The maximum positioning performance gain by additionally adding the estimated planar interaction positions is smaller than 0.7% for every maximum depth. The behavior regarding the benefit of adding CFs and estimated planar interaction positions is similar for GTB models with 1000 decision trees. In contrast to the GTB models with fewer decision trees, the averaged MAE slightly deteriorates for higher maximum depths as the models are already in the overfitting regime.

2) *High-Performance Optimization:* The averaged performance parameters of the best found GTB models for all three input sets are displayed in Table I. A GTB model of maximum depth 8, learning rate 0.05 and around 1550 decision trees performed best for raw data as input. The best models with inputs containing the CFs for both excluding and including the planar interaction position were found at maximum depth 6, learning rate 0.05, and around 1000 decision trees. These correspond to MRs of 6.7 and 1.1 MB according to (1), respectively. Furthermore, adding CFs improves the positioning performance by 1.6% to 8% compared to raw data as input while r_{50} and SR are affected most. No significant deviation of the positioning performance is induced by adding the planar interaction position.

Fig. 3 shows the spatial distribution for bias vector, MAE, and SR for raw data and CFs as input set. The bias vector points toward the crystal center close to both edges. In the region ranging from 2 to 10 mm, a bias vector magnitude below 1 mm is observed. The MAE deteriorates toward the crystal edges and is nearly constant at around 1.2 mm for DOI positions between 2 and 10 mm. The SR is nearly constant over the whole crystal and shows a maximum drop of approx. 24.7% close to 6 to 8 mm.

3) *Memory-Requirement Optimization:* Fig. 5 shows the averaged MAE against the MR of GTB models. The number of decision trees was determined by the convergence criterion for all three input sets. The lines represent models with the maximum depth denoted next to them while every marker indicates a tested learning rate. The plot allows choosing the best performing GTB models for given restrictions on the memory. High learning rates generally led to models with low MRs and

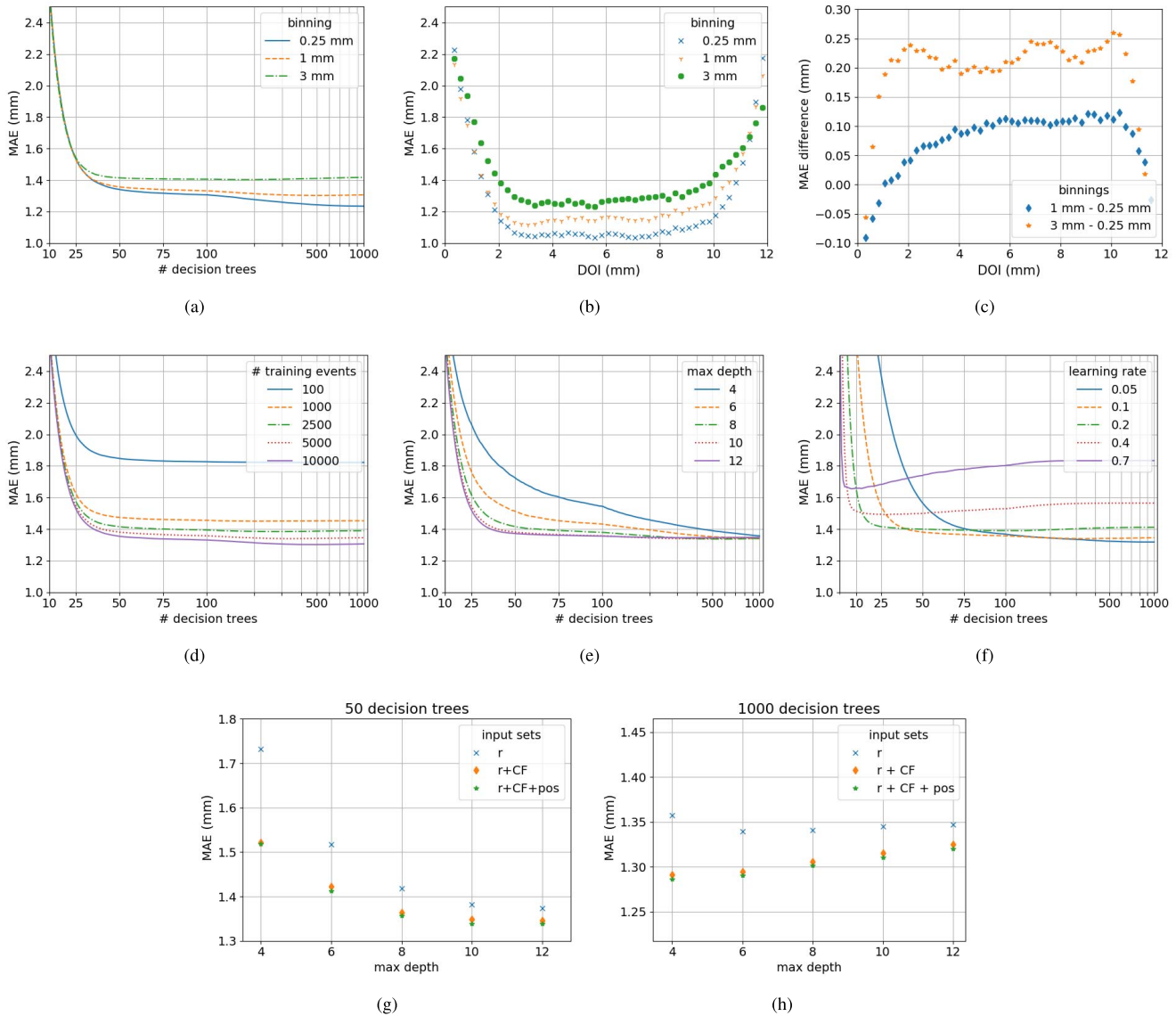


Fig. 4. General optimization process for GTB models. We employed a maximum depth of 10, learning rate 0.1, raw data of 5000 training events per irradiation position as input set and pitch of 1 mm of the calibration grid unless stated otherwise. The MAE is calculated employing the validation data set with 10 000 test events and 0.25 mm pitch of the irradiation grid. For the figures which display the number of decision trees, the abscissa is linearly scaled up to 100 decision trees and logarithmically afterwards. In case a spatial distribution is shown, DOI position 0 mm represents the top surface of the crystal and 12 mm the photosensor. Plots regarding the binning of the calibration grid show only selected pitches of the calibration grid due to reasons of clarity. (a) Averaged MAE against the number of decision trees and binning of the calibration grid. (b) Spatial distribution of the MAE for different pitches of the calibration grid. (c) Difference of the spatial distribution of the MAE for the tested calibration grids. (d) Averaged MAE against number of decision trees and number of training events per irradiation position. (e) Averaged MAE against number of decision trees and maximum depth. (f) Averaged MAE against number of decision trees and learning rate. (g) and (h) Averaged MAE against maximum depth and different input sets [raw data (r); raw data and CFs (r + CF); raw data, calculated features, and estimated planar interaction positions (r + CF + pos)]. Ensembles of 50 decision trees and 1000 decision trees are displayed, respectively.

vice versa. Adding CFs to the input is beneficial for the positioning performance for models of both low and high MR. The positioning performance improves from 0.1% to 0.9% comparing the corresponding GTB models with and without estimated planar interaction position of same maximum depth and learning rate.

V. DISCUSSION

The photon densities in the employed monolithic crystal are too low to generate a significant amount of events with all 16

hits present. This emphasizes the need for calibration and positioning algorithms able to deal with missing hit information. Although the presented results are obtained with the DPC photosensor, both DOI estimation methods can be applied to all kinds of photosensors.

The SO model can be successfully employed as a DOI estimator. In contrast to lookup table-approaches as employed in [15], [23], and [24], the IR model allows a continuous DOI estimation. In general, no huge influence on the positioning performance was observed for varying the segmentation of the scintillator because the main differences of the photon

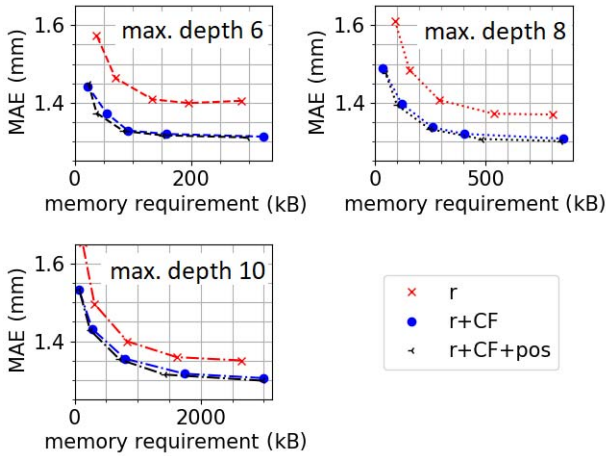


Fig. 5. Memory-requirement optimization: MAE against MR for all three input sets [raw data (r); raw data and CFs (r + CF); raw data, calculated features, and estimated planar interaction positions (r + CF + pos)]. For reasons of clarity, only selected maximum depths are shown. Adding the CF to the input of the GTB models improves the positioning performance and allows GTB models with lower MRs.

detection efficiency between single DPCs are canceled out by the normalization of the DOI observable. Here, SO models with a segmentation matching the pixel pitch performed best. For more segments, the number of training events per segment decreases below 150 per irradiation position which increases the uncertainty of the IR. Thus, a finer segmentation may lead to an improved performance in case more training events are available. The positioning performance deteriorates for all performance parameters in the central DOI region (see Fig. 3). We assume that the variation of the light distribution is lower in this area reducing the specificity of the used DOI observable. A similar course of the positioning performance employing the same DOI observable is reported for a 22 mm thick monolithic scintillator [15]. The deterioration in the central DOI region is significantly reduced if a dual-sided read-out is utilized [23]. In contrast to this paper, the calculation of the DOI observable in [15] is based on fully read out events. The authors presented an adaptation and performance of their method for up to four missing DPCs. Thus, the observed deterioration could be intensified as the DOI observable needs to be limited to one DPC to handle events with missing hits. Van Dam *et al.* [24] studied a monolithic scintillator of identical height as used in this paper testing multiple DOI observables. For all DOI observables, a SR ranging from 1 to 5 mm FWHM is reported while a SR better than 2 mm FWHM is only shown within 2 mm distance to the photosensor. As far as the SR alone allows a comparison, the presented SO approach performs better than the presented estimation by Van Dam *et al.* and will be used as one benchmark for the GTB models.

GTB models were successfully adapted to the DOI estimation problem. The general optimization process shows the influence of the hyperparameters and represents a protocol applicable to further scintillator geometries as well. GTB models show well-working regression capabilities allowing to significantly reduce the number of irradiation positions required for training [see Fig. 4(a)–(c)]. Thus, we were able

to reduce the binning of the calibration grid to 1 mm corresponding to 12 irradiation positions for a full calibration. Furthermore, GTB creates DOI positioning models with a comparably small amount of training data in the order of a few thousand events per irradiation position [see Fig. 4(d)]. Based on these results, we chose 5000 events per irradiation position for all trained models including the high-performance optimization. The selected calibration grid and number of training events per irradiation position lead to a calibration time of less than 30 min for the side irradiation a single detector block. This calibration time seems practical for calibrating a large number of detectors as well. If calibration time is a critical issue, the GTB models allow a further reduction of irradiation positions and training events without compromising much on the positioning performance.

After a sufficient maximum depth is reached, a further increase of the maximum depth did not yield to a significantly better positioning performance [see Fig. 4(e)]. In case a single decision tree corrected the given input during the training process with a smaller depth than the allowed maximum depth, all further added nodes are based on statistical fluctuations found in the training data. As these additional nodes do not contribute to the general predictivity of the GTB model, the positioning performance is not improved.

Maximum depth and learning rate are the most important hyperparameters to tune both the positioning performance and MR of the GTB models. As the MR is $\mathcal{O}(2^d)$ with d the maximum depth, this is the most important hyperparameter for an FPGA implementation. High learning rates allow a significant reduction of the number of decision trees at the cost of a decreased positioning performance [see Fig. 4(f)].

Adding CFs to the input features improves the positioning performance for GTB models with both small and high number of decision trees [see Fig. 4(g) and (h)]. The physically motivated features such as the defined DOI observable are easily interpretable and allow faster learning. Also, geometrical information as encoded in the projections has a high information content which is not directly accessible by the raw data. This also leads to less complex GTB models with smaller MR as shown for the high-performance optimization. Due to differences in the light response of the detector, the GTB models need to distinguish between different planar interaction positions. Thus, it is of interest if the GTB models benefit from the planar interaction position concerning a better positioning performance or reduction of required memory. As a consequence, adding the planar interaction position would require a sequential positioning process in a system architecture. As shown in Fig. 4(g) and (h), only a small positioning performance difference of less than 0.7% can be observed when adding the planar interaction position to the calculated features. The same observation holds true for the high-performance and memory-requirement optimization (see Table I and Fig. 5). The estimated planar interaction position does not provide a significant higher information content than the already added CFs such as center of gravity.

Both presented optimization scenarios focus on different use-cases. In case the CFs are added to the input set, the

high-performance optimization leads to GTB models of around 1 MB MR which are easy to handle for modern computers. However, the MR is the limiting factor for currently available FPGA in case no external memory is added. The memory-requirement optimization allows selecting the best-performing model for given resources. For example, the MR can be reduced down to 91 kB with a reduction of the MAE of less than 3% compared to the high-performance model requiring 1 MB. This MR can be easily handled by currently available FPGA families.

Comparing the SO models and high-performance GTB models, the GTB models have a better positioning performance of up to 50% for the averaged performance parameters (see Table I). Particularly, the bias vector and bias-sensitive performance parameters are improved; the averaged SR shows no significant difference. However, the differences get clearly visible comparing the spatial distribution of the performance parameters (see Fig. 3): In contrast to the SO model, the GTB models provide a nearly uniform positioning performance over the whole crystal depth. For example, the SR ranges between 1.9 and 2.5 mm FWHM for GTB models and between 1.0 and 5.3 mm for SO models. The bias-sensitive performance parameters such as MAE [see Fig. 3(b)] stays constant over the whole crystal except close to the edges for GTB models. As a main difference between SO and GTB, the GTB algorithm has access and utilizes all available information of the raw data and calculated features. To our best knowledge, the GTB models achieve the best reported positioning performance for DOI estimation with respect to averaged performance parameters and uniformity for the tested scintillator geometry. For example, Li *et al.* [20] achieve an averaged SR of 2.6 mm FWHM (corrected for finite beam width) using a parametric fitting model for a 10 mm thick scintillator. As already discussed, Van Dam *et al.* [24] presented a SR ranging from 1 mm to 5 mm not providing an uniform positioning performance. Comparable to our results, Wang *et al.* [25] demonstrated an averaged SR of around 2 mm employing a neural network. The reported SR deteriorated after applying a correction for bias effects. In contrast to the GTB method, the monolithic crystal was segmented into 35 cuboids with their own neural network DOI estimation models. Thus, the planar interaction position needs to be estimated before the DOI position.

VI. CONCLUSION

We presented two DOI positioning methods based on a side irradiation conducted with a fan beam collimator. Both SO and GTB models handle missing hit information. Due to the regression capabilities of the GTB models, the required time for the side irradiation can be reduced to less than 30 min without compromising too much on the positioning performance. The GTB model is able to employ the raw photon counts as input set. Adding physically motivated features to the raw data improves the positioning performance and allows models with smaller MRs. For DOI estimation, the GTB models do not require information of the planar interaction position. Thus, a full parallelization of both planar and DOI positioning is

feasible on system level. The developed memory optimization process allows training GTB models suitable for future FPGA implementation of the algorithm. In contrast to the SO models and other methods presented in literature, GTB models provide a nearly uniform positioning performance over the whole crystal depth. We achieved an averaged MAE of 1.28 mm and SR of 2.12 mm FWHM for the 12 mm-high crystal, respectively. Future research will translate and evaluate the GTB algorithm to monolithic scintillators of higher depth. Furthermore, alternatives replacing the side irradiation by means of optical simulation to create the training data are under investigation.

ACKNOWLEDGMENT

The authors would like to thank the open-source software packages NumPy [47], Matplotlib [48], and ROOT [49] heavily used in the analysis of the presented data.

REFERENCES

- [1] M. E. Phelps, *PET: Molecular Imaging and Its Biological Applications*. New York, NY, USA: Springer, 2004.
- [2] R. Myers, "The biological application of small animal PET imaging," *Nucl. Med. Biol.*, vol. 28, no. 5, pp. 585–593, Jul. 2001, doi: [10.1016/S0969-8051\(01\)00213-X](https://doi.org/10.1016/S0969-8051(01)00213-X).
- [3] G. B. Saha, *Basics of PET Imaging Basics of PET Imaging*. Cham, Switzerland: Springer Int., 2016, doi: [10.1007/978-3-319-16423-6](https://doi.org/10.1007/978-3-319-16423-6).
- [4] A. J. Gonzalez *et al.*, "A PET design based on SiPM and monolithic LYSO crystals: Performance evaluation," *IEEE Trans. Nucl. Sci.*, vol. 63, no. 5, pp. 2471–2477, Oct. 2016, doi: [10.1109/TNS.2016.2522179](https://doi.org/10.1109/TNS.2016.2522179).
- [5] M. S. Lee, K. Y. Kim, G. B. Ko, and J. S. Lee, "Prototype pre-clinical PET scanner with depth-of-interaction measurements using single-layer crystal array and single-ended readout," *Phys. Med. Biol.*, vol. 62, no. 10, pp. 3983–3996, May 2017, doi: [10.1088/1361-6560/aa64c7](https://doi.org/10.1088/1361-6560/aa64c7).
- [6] S. Krishnamoorthy *et al.*, "Performance evaluation of the MOLECUBES β -CUBE—A high spatial resolution and high sensitivity small animal PET scanner utilizing monolithic LYSO scintillation detectors," *Phys. Med. Biol.*, vol. 63, no. 15, Jul. 2018, Art. no. 155013, doi: [10.1088/1361-6560/aacec3](https://doi.org/10.1088/1361-6560/aacec3).
- [7] G. Borghi, V. Tabacchini, R. Bakker, and D. R. Schaart, "Sub-3 millimeter, near-200 picosecond TOF/DOI-PET imaging with monolithic scintillator detectors in a 70 cm diameter tomographic setup," *Phys. Med. Biol.*, vol. 63, no. 15, pp. 1–14, Jul. 2018, doi: [10.1088/1361-6560/aad2a6](https://doi.org/10.1088/1361-6560/aad2a6).
- [8] X. Zhang, R. D. Badawi, S. R. Cherry, and J. Qi, "Theoretical study of the benefit of long axial field-of-view PET on region of interest quantification," *Phys. Med. Biol.*, vol. 63, no. 13, 2018, Art. no. 135010, doi: [10.1088/1361-6560/aac815](https://doi.org/10.1088/1361-6560/aac815).
- [9] S. Vandenberghe, E. Mikhaylova, E. D'Hoe, P. Mollet, and J. S. Karp, "Recent developments in time-of-flight PET," *EJNMMI Phys.*, vol. 3, no. 1, p. 3, Dec. 2016, doi: [10.1186/s40658-016-0138-3](https://doi.org/10.1186/s40658-016-0138-3).
- [10] M. Ito *et al.*, "A four-layer DOI detector with a relative offset for use in an animal PET system," *IEEE Trans. Nucl. Sci.*, vol. 57, no. 3, pp. 976–981, Jun. 2010, doi: [10.1109/TNS.2010.2044892](https://doi.org/10.1109/TNS.2010.2044892).
- [11] M. S. Lee and J. S. Lee, "Depth-of-interaction measurement in a single-layer crystal array with a single-ended readout using digital silicon photomultiplier," *Phys. Med. Biol.*, vol. 60, no. 16, pp. 6495–6514, Aug. 2015, doi: [10.1088/0031-9155/60/16/6495](https://doi.org/10.1088/0031-9155/60/16/6495).
- [12] J. Seidel, J. J. Vaquero, S. Siegel, W. R. Gandler, and M. V. Green, "Depth identification accuracy of a three layer phoswich PET detector module," *IEEE Trans. Nucl. Sci.*, vol. 46, no. 3, pp. 485–490, Jun. 1999, doi: [10.1109/23.775567](https://doi.org/10.1109/23.775567).
- [13] H. Uchida *et al.*, "A novel single-ended readout depth-of-interaction PET detector fabricated using sub-surface laser engraving," *Phys. Med. Biol.*, vol. 61, no. 18, pp. 6635–6650, Sep. 2016, doi: [10.1088/0031-9155/61/18/6635](https://doi.org/10.1088/0031-9155/61/18/6635).
- [14] H. G. Kang *et al.*, "A dual-ended readout detector using a meantime method for SiPM TOF-DOI PET," *IEEE Trans. Nucl. Sci.*, vol. 62, no. 5, pp. 1935–1943, Oct. 2015, doi: [10.1109/TNS.2015.2449891](https://doi.org/10.1109/TNS.2015.2449891).

- [15] G. Borghi, V. Tabacchini, and D. R. Schaart, "Towards monolithic scintillator based TOF-PET systems: Practical methods for detector calibration and operation," *Phys. Med. Biol.*, vol. 61, no. 13, pp. 4904–4928, Jul. 2016, doi: [10.1088/0031-9155/61/13/4904](https://doi.org/10.1088/0031-9155/61/13/4904).
- [16] H. T. Van Dam, G. Borghi, S. Seifert, and D. R. Schaart, "Sub-200 ps CRT in monolithic scintillator PET detectors using digital SiPM arrays and maximum likelihood interaction time estimation," *Phys. Med. Biol.*, vol. 58, no. 10, pp. 3243–3257, 2013, doi: [10.1088/0031-9155/58/10/3243](https://doi.org/10.1088/0031-9155/58/10/3243).
- [17] P. Bruyndonckx *et al.*, "Evaluation of machine learning algorithms for localization of photons in undivided scintillator blocks for PET detectors," *IEEE Trans. Nucl. Sci.*, vol. 55, no. 3, pp. 918–924, Jun. 2008, doi: [10.1109/TNS.2008.922811](https://doi.org/10.1109/TNS.2008.922811).
- [18] R. Marcinkowski, P. Mollet, R. Van Hoken, and S. Vandenberghe, "Sub-millimetre DOI detector based on monolithic LYSO and digital SiPM for a dedicated small-animal PET system," *Phys. Med. Biol.*, vol. 61, no. 5, pp. 2196–2212, 2016, doi: [10.1088/0031-9155/61/5/2196](https://doi.org/10.1088/0031-9155/61/5/2196).
- [19] A. J. Gonzalez *et al.*, "The MINDView brain PET detector, feasibility study based on SiPM arrays," *Nucl. Instrum. Methods Phys. Res. A Accelerators Spectr. Detectors Assoc. Equip.*, vol. 818, pp. 82–90, May 2016, doi: [10.1016/j.nima.2016.02.046](https://doi.org/10.1016/j.nima.2016.02.046).
- [20] X. Li, C. Lockhart, T. K. Lewellen, and R. S. Miyaoka, "A high resolution, monolithic crystal, PET/MRI detector with DOI positioning capability," in *Proc. Conf. Annu. IEEE Eng. Med. Biol. Soc.*, vol. 2008, pp. 2287–2290, 2008, doi: [10.1109/IEMBS.2008.4649654](https://doi.org/10.1109/IEMBS.2008.4649654).
- [21] A. González-Montoro *et al.*, "Detector block performance based on a monolithic LYSO crystal using a novel signal multiplexing method," *Nucl. Instrum. Methods Phys. Res. A Accelerators Spectr. Detectors Assoc. Equip.*, to be published, doi: [10.1016/j.nima.2017.10.098](https://doi.org/10.1016/j.nima.2017.10.098).
- [22] A. González-Montoro *et al.*, "Performance study of a large monolithic LYSO PET detector with accurate photon DOI using retroreflector layers," *IEEE Trans. Radiat. Plasma Med. Sci.*, vol. 1, no. 3, pp. 229–237, May 2017, doi: [10.1109/TRPMS.2017.2692819](https://doi.org/10.1109/TRPMS.2017.2692819).
- [23] G. Borghi, B. J. Peet, V. Tabacchini, and D. R. Schaart, "A 32 mm x 32 mm x 22 mm monolithic LYSO:Ce detector with dual-sided digital photon counter readout for ultrahigh-performance TOF-PET and TOF-PET/MRI," *Phys. Med. Biol.*, vol. 61, no. 13, pp. 4929–4949, Jun. 2016, doi: [10.1088/0031-9155/61/13/4929](https://doi.org/10.1088/0031-9155/61/13/4929).
- [24] H. T. Van Dam *et al.*, "A practical method for depth of interaction determination in monolithic scintillator PET detectors," *Phys. Med. Biol.*, vol. 56, no. 13, pp. 4135–4145, 2011, doi: [10.1088/0031-9155/56/13/025](https://doi.org/10.1088/0031-9155/56/13/025).
- [25] Y. Wang, W. Zhu, X. Cheng, and D. Li, "3D position estimation using an artificial neural network for a continuous scintillator PET detector," *Phys. Med. Biol.*, vol. 58, no. 5, pp. 1375–1390, Mar. 2013, doi: [10.1088/0031-9155/58/5/1375](https://doi.org/10.1088/0031-9155/58/5/1375).
- [26] 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 Trans. Radiat. Plasma Med. Sci.*, vol. 2, no. 5, pp. 411–421, Sep. 2018, doi: [10.1109/TRPMS.2018.2837738](https://doi.org/10.1109/TRPMS.2018.2837738).
- [27] R. Kułaga and M. Gorg, "FPGA implementation of decision trees and tree ensembles for character recognition in Vivado HLS," *Image Process. Commun.*, vol. 19, nos. 2–3, pp. 71–82, 2015, doi: [10.1515/ipc-2015-0012](https://doi.org/10.1515/ipc-2015-0012).
- [28] B. Van Essen, C. Macaraeg, M. Gokhale, and R. Prenger, "Accelerating a random forest classifier: Multi-core, GP-GPU, or FPGA?" in *Proc. IEEE 20th Int. Symp. Field Program. Custom Comput. Mach.*, Apr. 2012, pp. 232–239, doi: [10.1109/FCCM.2012.47](https://doi.org/10.1109/FCCM.2012.47).
- [29] C. Degenhardt *et al.*, "The digital silicon photomultiplier—A novel sensor for the detection of scintillation light," in *Proc. IEEE Nucl. Sci. Symp. Conf. Rec.*, 2009, pp. 2383–2386, doi: [10.1109/NSSMIC.2009.5402190](https://doi.org/10.1109/NSSMIC.2009.5402190).
- [30] T. Frach *et al.*, "The digital silicon photomultiplier—Principle of operation and intrinsic detector performance," in *Proc. IEEE Nucl. Sci. Symp. Conf. Rec.*, 2009, pp. 1959–1965, doi: [10.1109/NSSMIC.2009.5402143](https://doi.org/10.1109/NSSMIC.2009.5402143).
- [31] D. Schug *et al.*, "First evaluations of the neighbor logic of the digital SiPM tile," in *Proc. IEEE Nucl. Sci. Symp. Med. Imag. Conf. Rec.*, Oct. 2012, pp. 2817–2819, doi: [10.1109/NSSMIC.2012.6551642](https://doi.org/10.1109/NSSMIC.2012.6551642).
- [32] V. Tabacchini, V. Westerwoudt, G. Borghi, S. Seifert, and D. R. Schaart, "Probabilities of triggering and validation in a digital silicon photomultiplier," *J. Instrum.*, vol. 9, no. 6, Jun. 2014, Art. no. P06016, doi: [10.1088/1748-0221/9/06/P06016](https://doi.org/10.1088/1748-0221/9/06/P06016).
- [33] R. Marcinkowski, S. Espana, R. Van Hoken, and S. Vandenberghe, "Effects of dark counts on digital silicon photomultipliers performance," in *Proc. IEEE Nucl. Sci. Symp. Med. Imag. Conf. (NSSMIC)*, Oct. 2013, pp. 1–6, doi: [10.1109/NSSMIC.2013.6829323](https://doi.org/10.1109/NSSMIC.2013.6829323).
- [34] D. R. Schaart, E. Charbon, T. Frach, and V. Schulz, "Advances in digital SiPMs and their application in biomedical imaging," *Nucl. Instrum. Methods Phys. Res. A Accelerators Spectrometers Detectors Assoc. Equip.*, vol. 809, pp. 31–52, Feb. 2016, doi: [10.1016/j.nima.2015.10.078](https://doi.org/10.1016/j.nima.2015.10.078).
- [35] J. Grahe, F. Müller, D. Schug, P. Hallen, and V. Schulz, "Optical Monte-Carlo simulation to evaluate monolithic PET detector concepts," in *Proc. IEEE Nucl. Sci. Symp. Med. Imag. Conf.*, Atlanta, GA, USA, 2017, pp. 1–6, doi: [10.1109/NSSMIC.2017.8532632](https://doi.org/10.1109/NSSMIC.2017.8532632).
- [36] D. Schug *et al.*, "Data processing for a high resolution preclinical PET detector based on Philips DPC digital SiPMs," *IEEE Trans. Nucl. Sci.*, vol. 62, no. 3, pp. 669–678, Jun. 2015, doi: [10.1109/TNS.2015.2420578](https://doi.org/10.1109/TNS.2015.2420578).
- [37] N. Gross-Weege, D. Schug, P. Hallen, and V. Schulz, "Maximum likelihood positioning algorithm for high-resolution PET scanners," *Med. Phys.*, vol. 43, no. 6, pp. 3049–3061, 2016, doi: [10.1118/1.4950719](https://doi.org/10.1118/1.4950719).
- [38] B. Weissler *et al.*, "A digital preclinical PET/MRI insert and initial results," *IEEE Trans. Med. Imag.*, vol. 34, no. 11, pp. 2258–2270, Nov. 2015, doi: [10.1109/TMI.2015.2427993](https://doi.org/10.1109/TMI.2015.2427993).
- [39] *Performance Measurements of Small Animal Positron Emission Tomographs*, document SP NU 4-2008, NEMA, Nairobi, Kenya, 2008.
- [40] S. M. Kakade, V. Kanade, O. Shamir, and A. Kalai, "Efficient learning of generalized linear and single index models with isotonic regression," in *Advances in Neural Information Processing Systems 24*, J. Shawe-Taylor, R. S. Zemel, P. L. Bartlett, F. Pereira, and K. Q. Weinberger, Eds., Curran Assoc., 2011, pp. 927–935. [Online]. Available: <http://papers.nips.cc/paper/4429-efficient-learning-of-generalized-linear-and-single-index-models-with-isotonic-regression.pdf>
- [41] F. Pedregosa *et al.*, "Scikit-learn: Machine learning in Python," *J. Mach. Learn. Res.*, vol. 12, pp. 2825–2830, Feb. 2011.
- [42] T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," in *Proc. 22nd ACM SIGKDD Int. Conf. Knowl. Disc. Data Min.*, 2016, pp. 785–794, doi: [10.1145/2939672.2939785](https://doi.org/10.1145/2939672.2939785).
- [43] S. B. Kotsiantis, "Decision trees: A recent overview," *Artif. Intell. Rev.*, vol. 39, no. 4, pp. 261–283, 2013, doi: [10.1007/s10462-011-9272-4](https://doi.org/10.1007/s10462-011-9272-4).
- [44] A. Natekin, A. Knoll, M.-O. Gewaltig, and O. Michel, "Gradient boosting machines, a tutorial," *Front. Neurobot.*, vol. 7, p. 21, Dec. 2013, doi: [10.3389/fnbot.2013.00021](https://doi.org/10.3389/fnbot.2013.00021).
- [45] J. H. Friedman, "Greedy function approximation: A gradient boosting machine," *Ann. Stat.*, vol. 29, no. 5, pp. 1189–1232, 2001.
- [46] P. Gebhardt *et al.*, "FPGA-based singles and coincidences processing pipeline for integrated digital PET/MR detectors," in *Proc. IEEE Nucl. Sci. Symp. Med. Imag. Conf. (NSSMIC)*, 2012, pp. 2479–2482, doi: [10.1109/NSSMIC.2012.6551565](https://doi.org/10.1109/NSSMIC.2012.6551565).
- [47] S. van der Walt, S. C. Colbert, and G. Varoquaux, "The NumPy array: A structure for efficient numerical computation," *Comput. Sci. Eng.*, vol. 13, no. 2, pp. 22–30, Mar./Apr. 2011, doi: [10.1109/MCSE.2011.37](https://doi.org/10.1109/MCSE.2011.37).
- [48] J. D. Hunter, "Matplotlib: A 2D graphics environment," *Comput. Sci. Eng.*, vol. 9, no. 3, pp. 90–95, May/Jun. 2007, doi: [10.1109/MCSE.2007.55](https://doi.org/10.1109/MCSE.2007.55).
- [49] R. Brun and F. Rademakers, "ROOT—An object oriented data analysis framework," *Nucl. Instrum. Methods Phys. Res. A Accelerators Spectrometers Detectors Assoc. Equip.*, vol. 389, nos. 1–2, pp. 81–86, Apr. 1997, doi: [10.1016/S0168-9002\(97\)00048-X](https://doi.org/10.1016/S0168-9002(97)00048-X).