Physics in Medicine & Biology



**PAPER** • **OPEN ACCESS**

# Sub-millimeter precise photon interaction position determination in large monolithic scintillators via convolutional neural network algorithms

To cite this article: M Kawula *et al* 2021 *Phys. Med. Biol.* **66** 135017

IOP Publishing Phys. Med. Biol. 66 (2021) 135017 https://doi.org/10.1088/1361-6560/ac06e2

Physics in Medicine & Biology

IPEM Institute of Physics and Engineering in Medicine

OPEN ACCESS

RECEIVED
21 February 2021

REVISED
11 May 2021

ACCEPTED FOR PUBLICATION
1 June 2021

PUBLISHED
2 July 2021



PAPER

# Sub-millimeter precise photon interaction position determination in large monolithic scintillators via convolutional neural network algorithms

**M Kawula[1], T M Binder[1,2], S Liprandi[1], R Viegas[1,3], K Parodi[1] and P G Thirolf[1,*]**

[1] Department of Medical Physics, Ludwig-Maximilians-Universität München, Garching b. München, Germany
[2] KETEK GmbH, Munich, Germany
[3] University of Coimbra, Portugal
[*] Author to whom any correspondence should be addressed.

**E-mail:** maria.kawula@physik.uni-muenchen.de **and** Peter.Thirolf@physik.uni-muenchen.de

**Keywords:** beam range monitoring, Compton camera, hadron therapy, monolithic scintillator, neural networks, radiation detection, spatial resolution

## Abstract

In this work, we present the development and application of a convolutional neural network (CNN)-based algorithm to precisely determine the interaction position of $\gamma$-quanta in large monolithic scintillators. Those are used as an absorber component of a Compton camera (CC) system under development for ion beam range verification via prompt-gamma imaging. We examined two scintillation crystals: $LaBr_3$:Ce and $CeBr_3$. Each crystal had dimensions of 50.8 mm $\times$ 50.8 mm $\times$ 30 mm and was coupled to a 64-fold segmented multi-anode photomultiplier tube (PMT) with an 8 $\times$ 8 pixel arrangement. We determined the spatial resolution for three photon energies of 662, 1.17 and 1.33 MeV obtained from 2D detector scans with tightly collimated $^{137}$Cs and $^{60}$Co photon sources. With the new algorithm we achieved a spatial resolution for the CeBr3 crystal below 1.11(8) mm and below 0.98(7) mm for the LaBr3:Ce detector for all investigated energies between 662 keV and 1.33 MeV. We thereby improved the performance by more than a factor of 2.5 compared to the previously used categorical average pattern algorithm, which is a variation of the well-established k-nearest neighbor algorithm. The trained CNN has a low memory footprint and enables the reconstruction of up to $10^4$ events per second with only one GPU. Those improvements are crucial on the way to future clinical *in vivo* applicability of the CC for ion beam range verification.

## 1. Introduction

A therapeutic hadron beam used in particle therapy stops inside the patient's body, rendering the information on the dose deposition not as easily accessible as it is the case with photon irradiation. Among the numerous range verification techniques presently under study, there are those that rely on the detection of an ionoacoustic signal originating from thermal expansion following localized heating in the dose deposition of a pulsed ion beam (Kellnberger *et al* 2016), secondary charged particles (in case of carbon ion beams) (Gwosch *et al* 2013) or delayed annihilation $\gamma$-rays such as in positron emission tomography (PET) (Zhu and El Fakhri 2013). Finally, there are methods based on the detection of prompt $\gamma$-rays such as multi-slit camera (Smeets *et al* 2016), prompt-gamma spectroscopy (Verburg and Seco 2014), prompt-gamma timing (Golnik *et al* 2014), $\gamma$-PET, also called triple-$\gamma$ PET or whole gamma imaging (Lang *et al* 2014, Manzano *et al* 2015, Yoshida *et al* 2020), or Compton camera (CC) based imaging systems (Krimmer *et al* 2015, Polf *et al* 2015, Llosá *et al* 2016, Aldawood *et al* 2017). In this paper the focus is set on the latter approach.

The basic components of a CC system are two detectors: a scatterer and an absorber. Depending on the energy, the initial $\gamma$-ray undergoes Compton scattering in the first detector and should then be fully absorbed in the second one. The goal of a single-event reconstruction is to determine the so-called Compton cone, which is a

collection of all possible locations from which the detected $\gamma$-quantum could have been emitted. By sequentially detecting multiple $\gamma$-rays originating from the same point, it is possible to get the source location at the intersection of the reconstructed Compton cones.

The use of monolithic scintillators as an absorber component is an alternative to pixelated detectors, as being well-established in PET scanners, providing better energy resolution, higher sensitivity as well as cheaper and easier manufacturing. All these advantages, however, come at the cost of a more complex derivation of spatial information that has to be carried out with the help of designated algorithms.

Several algorithms have been proposed so far, notably the k-nearest neighbors (kNN) and categorical averaged pattern (CAP) (van Dam *et al* 2011) algorithms. The main principle of both are lookup tables, here called reference libraries, containing the position dependent detector response, i.e. the scintillation light amplitude distribution across the (segmented) photosensor array. During the reconstruction process, the light amplitude distribution of an unknown event is compared with all reference library entries. Then, according to specific rules, the photon interaction position is determined by the maximum of a 2D histogram, which contains a selectable number of k most similar events. Although both algorithms achieve satisfactory spatial resolution, these methods are not applicable in real-time practice due to their long computational times.

In this paper we propose an algorithm for determining the position of $\gamma$-ray interactions in a monolithic scintillation crystal, which is based on supervised machine learning involving convolutional network networks (CNN). We show the development steps starting from the network architecture optimization, through the determination of the optimal database size and suggest an effective training schedule. Our algorithm is applicable to all types of detectors where a photon interaction position can be associated with a characteristic light amplitude distribution. The new method allows improving the positioning accuracy by a factor of about 2.5 compared to the CAP algorithm, has a low memory footprint and opens up prospects for real-time application.

## 2. Materials

The CC prototype under development in our group for ion beam range verification via prompt-gamma imaging is a two-stage detector system. The scatterer consists of a stack of six double-sided silicon strip detectors, which provide direct information on the deposited energy as well as interaction positions of both, the incident $\gamma$-ray and the scattered electron. The second component, which is the focus of this work, is a single thick monolithic scintillation crystal read out by a multi-anode PMT that absorbs the scattered photon. A characteristic light distribution pattern registered by the PMT correlates with the photon interaction position inside the crystal.

Two monolithic scintillation crystals with dimensions of 50.8 mm $\times$ 50.8 mm $\times$ 30 mm were examined as potential candidates for the absorber component: $LaBr_3$:Ce and $CeBr_3$. The crystals were wrapped in a reflective material which, despite a more inhomogeneous spatial response of the detector compared to an absorbing coating, has a significantly better energy and time resolution and thus a better over all performance (Aldawood *et al* 2015). Both crystals show hygroscopic properties and were therefore protected from humidity within an aluminum housing. Both scintillation materials exhibit short decay times of 16–19 ns, high light yield in the range of 6.3–6.8 $\times$ $10^4$ ph $MeV^{-1}$ and an excellent relative energy resolution of 3.5%–4.1% at 662 keV. While $LaBr_3$:Ce carries an intrinsic radioactivity of about 2 Bq $cm^{-3}$, $CeBr_3$, being free of this disadvantage, shows a higher signal-to noise ratio (Saint-Gobain 2016, Scionix 2018).

We used a 64-fold segmented multi-anode PMT with an 8 $\times$ 8 pixel arrangement from Hamamatsu (models H8500C and H12700 Hamamatsu 2015), which has been shown to be superior when compared to a higher segmented 256-channel PMT (Viegas 2018). Each pixel has a size of 5.8 mm $\times$ 5.8 mm, the outer dimension of the PMT array is 49 mm $\times$ 49 mm, therefore slightly smaller than the size of the scintillation crystals. The photosensor and absorber were coupled with an optical grease whose refractive index is similar to that of the absorber material and the photomultiplier entrance window made of borosilicate glass. The signals of the 64 PMT channels are sent to adapter boards via four 16-pin coaxial ribbon cables and transferred further into four Constant Fraction Discriminator modules via LEMO cables. The MCFDs amplify the energy signals and create amplitude-independent logical signals that act as individual gates for the charge digitizer modules. Both types of signals are then fed into charge-to-digital converter modules that integrate the incoming signals over the duration of the individual gates, providing energy information of each channel. In addition to the 64 signals, a signal from the PMT sum-dynode is fed to a separate MCFD module, whose common logic output is sent further to a TRIVA 5 VME trigger unit (Hoffman 2009) as trigger for the data acquisition system, as well as to a Quad Coincidence Logic Unit, that splits this master gate into identical copies which are sent to each MQDC module, opening an acquisition time window to ensure synchronized data acquisition across all modules. The acquired data is sent to the control PC via a VME frontend CPU (RIO-3 operated under the real-time operational system LynxOS). A block of the data acquisition system is presented in figure 1.

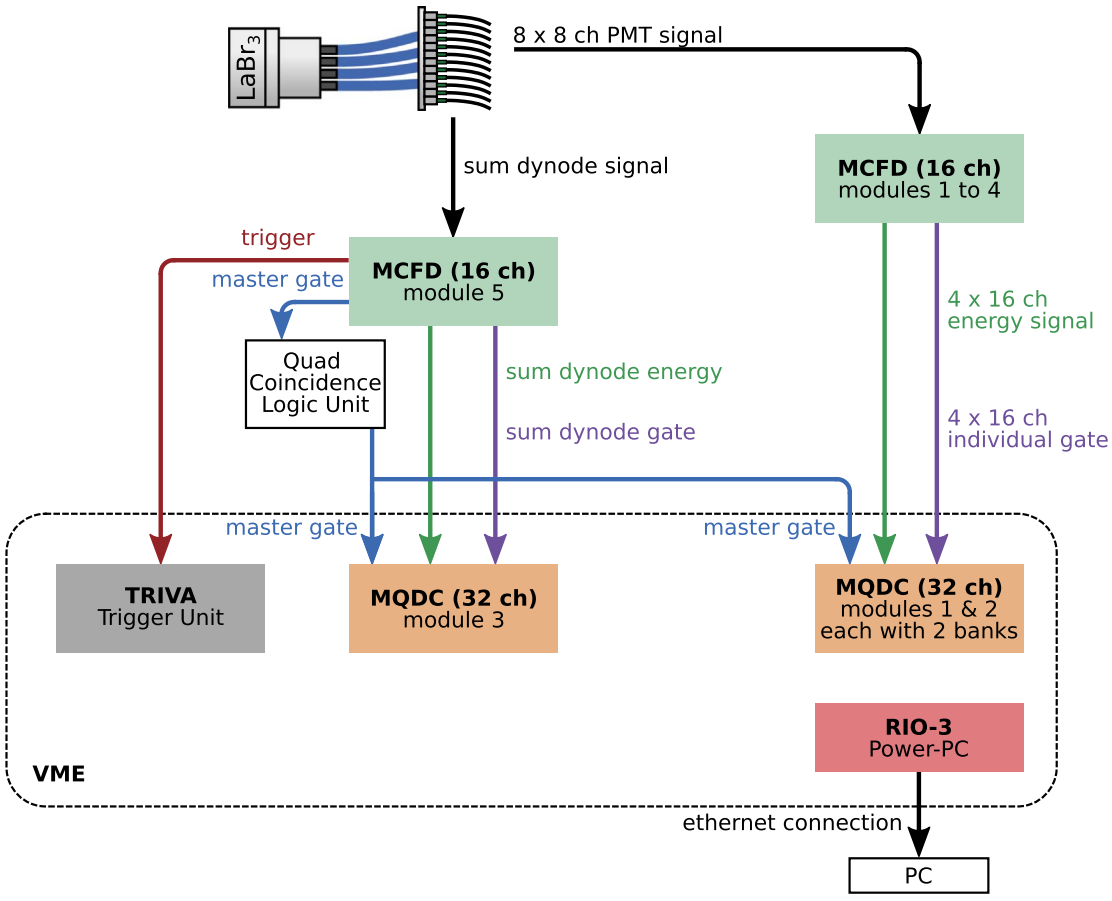


**Figure 1.** Block diagram of the absorber detector signal read out and data acquisition electronics.

## 3. Methods

### 3.1. Data acquisition

The training of a CNN requires an extensive collection of labeled data (here referred to as events), which are in our case 2D light amplitude distributions registered by a PMT and associated to their respective irradiation position coordinates. We adapted existing reference libraries, which served previously as lookup tables for position reconstruction using the kNN and CAP algorithms. The data acquisition of the light amplitude distributions was performed by irradiating the crystal's front surface with tightly collimated $\gamma$-ray sources: $^{137}$Cs emitting 662 keV $\gamma$-rays and $^{60}$Co emitting two coincident photons of energies of 1.17 and 1.33 MeV. Our collimator was a 10 cm long DENSIMENT rod with a central 1 mm diameter channel, arranged pointing perpendicular to the absorber surface. The crystal was irradiated on a 2D grid of 102 $\times$ 102 reference positions with a 0.5 mm step size (via a scanning system with remotely controlled and motorized $(x, y)$ translation stages). For the $LaBr_3$:Ce crystal 800 photopeak events originating from the $^{137}$Cs irradiation and 600 events from $^{60}$Co irradiation (separately for both cobalt energies) were collected at each position. The generation of the aforementioned $^{137}$Cs library (source activity 72 MBq) of 600 $\times$ 102 $\times$ 102 entries took about 7 d of continuous measurement time. The activity of the $^{60}$Co source was significantly lower (27 MBq), thus the acquisition of 600 $\times$ 102 $\times$ 102 photopeak events required about 18 d uninterrupted data acquisition. In order to separate scattered events from photopeak events, an energy gate was set to select the photopeak energy. Due to the absence of internal radioactivity for the $CeBr_3$ and therefore the higher signal-to-noise ratio, we were able to reduce (compared to the $LaBr_3$:Ce crystal irradiation) the number of reference library entries to 600 and 480 for the irradiation with $^{137}$Cs and $^{60}$Co sources, respectively. A sketch of the experimental setup is presented in figure 2(a).

### 3.2. Data preprocessing

Each event entry belonging to a reference library consists of a two-dimensional, 8 $\times$ 8 feature vector, representing the light amplitude distribution in each channel of the PMT, and a target vector indicating the $(x, y)$ coordinates of the irradiation position. The target vectors were constructed using the one-hot, also known as

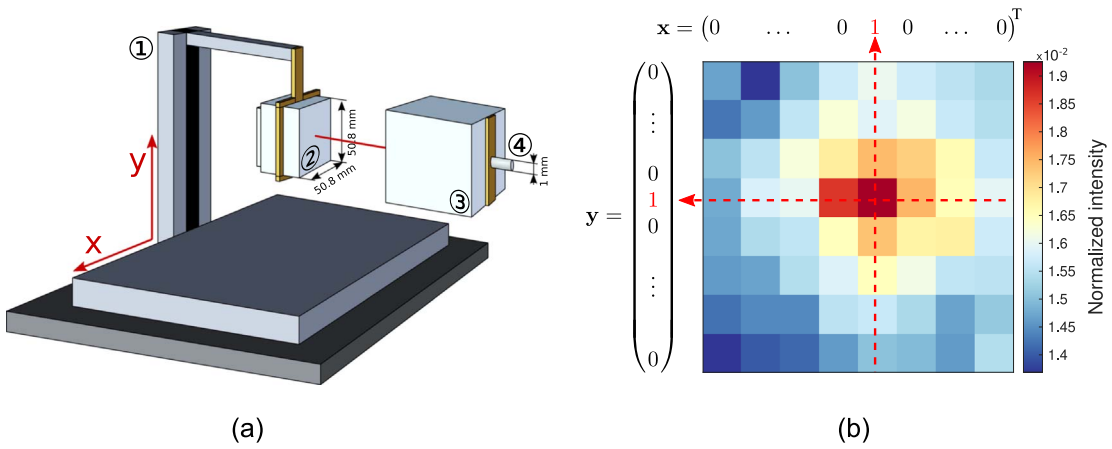


**Figure 2.** (a) Sketch of the setup for the 2D detector scan using collimated $\gamma$-ray sources. The scintillation crystal ② is attached to a motorized translation stage ① while the collimator ④ together with its surrounding lead shielding block ③ is kept stationary. (b) Schematic light intensity distribution map with indicated interaction point (intersection of the red dashed arrows) and corresponding one-hot-encoded target vectors.

dummy, encoding scheme (Alkharusi 2012). All input vectors were normalized such that the sum over all PMT channels is equal to one. An exemplary light amplitude distribution, averaged over 400 events registered at the same irradiation position, together with the corresponding target vectors is presented in figure 2(b).

Approximately 20%–30% of the reference events were set aside to serve later as a test set being a proxy for the new unlabeled data. The remaining 70%–80% were used as training and validation data to optimize the model parameters.

### 3.3. Workflow

We categorized the task of finding the interaction position of a $\gamma$-ray, based on the PMT response, as an image, i.e. pattern, recognition problem. CNN building blocks were used to design a new reconstruction algorithm, while the large datasets of labeled data enabled us to carry out supervised learning.

CNN algorithms are known for their large number of uncorrelated hyperparameters that must be set prior to the training. An attempt of optimizing them all simultaneously would be equivalent to finding the absolute minimum of a multidimensional function, which due to computational constraints is not possible to be carried out in practice. For that reason, the process of searching their optimum combination was divided into several stages with the following priorities: (1) model specific hyperparameters (number of convolutional layers and kernels), (2) database size (3) optimizer hyperparameters (learning rate and number of epochs for training).

### 3.4. Architecture design

The first stage of the hyperparameter optimization, performed to determine the number of convolutional layers and $3 \times 3$ kernels building the so-called convolutional block, was carried out through the random search method, meaning that several highly dissimilar network architectures were investigated to determine a smaller sub-space for the second stage of optimization, the grid search. After every convolution rectified linear unit (ReLU) activation (Gonzalez and Woods 2018) and batch normalization (Ioffe and Szegedy 2015) were applied. To prevent the image from shrinking, each layer was zero-padded before performing convolution operation. To generate predictions, separately for the $x$ and $y$ coordinates, the network was split into two output blocks, each made up of two fully connected layers with softmax (Gonzalez and Woods 2018) as the final activation function. The output had the form of two 102-dimensional probability vectors, originating from the 102 reference library irradiation positions along each axis. The coordinate with the maximum probability value was considered as the reconstructed interaction position.

### 3.5. Training

We carried out the training using a categorical cross-entropy loss function, Adam optimizer and introduced a data split into mini-batches of size 400. Even though the Adam optimizer provides a separate learning rate ($lr$) for each network parameter, which is adapted as learning unfolds, we found it beneficial to divide the training into several phases, gradually decreasing the initial learning rate for the optimizer as follows: $lr = \{10^{-5}, 5 \times 10^{-6}, 10^{-6}\}$ for approximately 100, 50 and 50 epochs, respectively.

We prepared separate networks, for each energy of the incident $\gamma$-ray and each crystal. However, to exploit common features of the light distributions, data from all libraries were included at the beginning of the training.

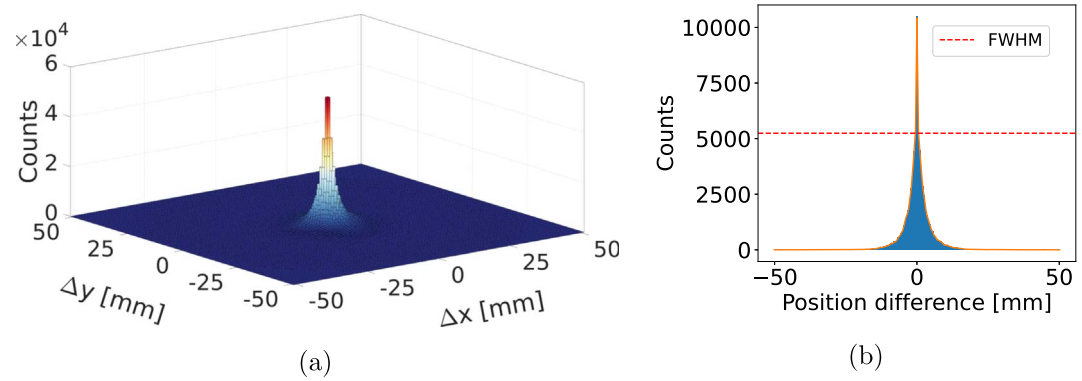


**Figure 3.** (a) Two-dimensional error histogram. (b) An exemplary projection of the error histogram onto the *x* axis. The orange curve presents the linear interpolation between adjacent bins. The red horizontal line indicate the FWHM level.

The final parameter tuning was carried out with the events of the specific energy and for the dedicated crystal only.

### 3.6. Database optimization

In theory, the neural network performance should improve with increasing database size, but eventually a saturation point is reached and the model does no longer benefit from new training examples. Since reference library acquisition is a lengthy process (7 and 18 d to collect $^{137}$Cs and $^{60}$Co data, respectively), it is advantageous to determine the minimum database size required to achieve the desired model accuracy.

Therefore, the training was carried out for 13 different numbers of photopeak events per irradiation position:

$$n_{\mathrm{epp|train}} = \{10,\ 50,\ 100,\ 150,\ 200,\ 250,\ 300,\ 350,\ 400,\ 450,\ 500,\ 550,\ 600\},$$

while additional 50 events per position were used as validation set. Model weights and biases were initialized with the identical set of random numbers.

Six different values of the number of photopeak events per irradiation position $n_{\mathrm{epp|test}}$ were examined, in order to determine the minimum number of test events that produce a smooth histogram, suitable for full width at half determination:

$$n_{\mathrm{epp|test}} = \{1,\ 5,\ 10,\ 15,\ 20,\ 25\}.$$

### 3.7. Determination of the algorithm spatial resolution

Interaction positions of events from the test set were predicted by the neural network and compared to the ground-truth. The differences between these two were used to fill a 2D error histogram creating a narrow peak similar in shape to a 2D Gaussian or Lorentzian. The spatial resolution of the algorithm was calculated as the average FWHM of the histogram projections along both axes. An exemplary error histogram and its projection in *x* direction are depicted in figure 3.

In order to quantify the statistical uncertainty of the spatial resolution determination, the test dataset was divided into five subgroups and for each of them a separate error histogram with its FWHM was calculated. From the resulting five data points an average FWHM and its standard deviation were determined. The systematic uncertainty has many sources. Firstly, both hygroscopic absorber crystals are encapsulated in an aluminum housing and the exact position of the crystal edges has to be determined indirectly by the so-called 'edge scan', i.e. 1D scans in *x* and *y* directions, respectively, followed by sigmoid fits to the resulting intensity profiles, with a precision of about $\pm$0.3 mm (Liprandi 2018). Secondly, the beam is not infinitesimally narrow, but collimated by a 10cm long tube with a 1 mm diameter opening. Assuming for simplicity, that all photon interactions take place at the crystal front surface, which is 0.2–0.3 mm away from the rear end of the collimator, we have experimentally determined the beam spot size to have approximately 1.2 mm diameter (Binder 2017) with a Gaussian intensity profile.

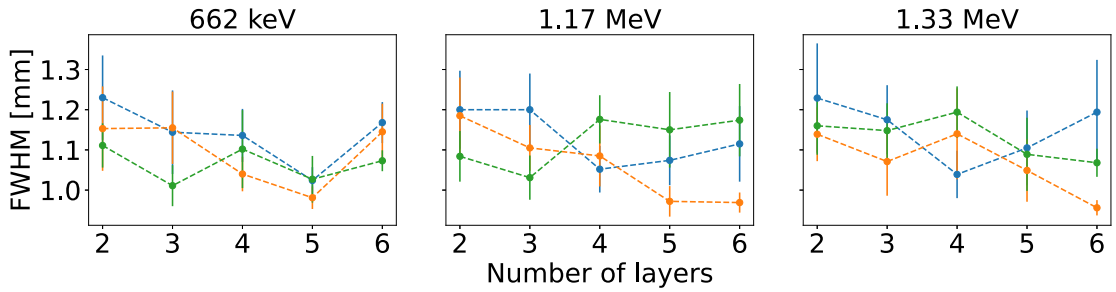


**Figure 4.** Performance of the CNN algorithm for three different energies of incident $\gamma$-rays, as a function of the number of layers and kernels (two-blue, three-orange, four-green) for the $LaBr_3$:Ce crystal. The spatial resolution was calculated as FWHM of the error histogram.

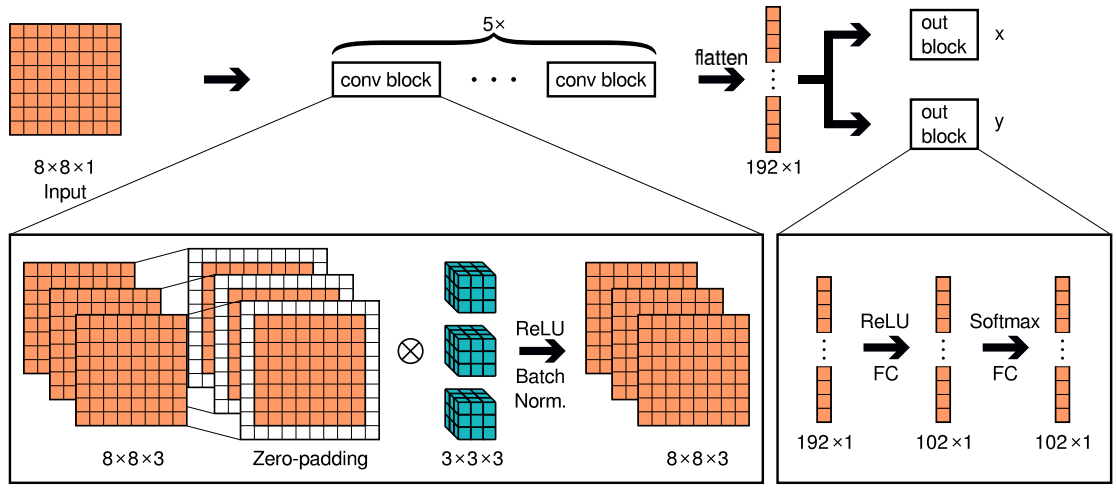


**Figure 5.** Sketch of the best performing CNN architecture (for more details see text).

## 4. Results

### 4.1. Architecture design

Using a random search method we determined the upper limit of the number of convolutional layers and $3 \times 3$ kernels of $n_{conv} = 6$ and $n_{kern} = 4$, respectively. More complex architectures did not lead to further improvements. The second stage of optimization was performed employing the grid search technique for all combinations of the following parameter values: $n_{conv} = \{2, 3, 4, 5, 6\}$ and $n_{kern} = \{2, 3, 4\}$. Figure 4 present the corresponding outcomes of the optimization referring to the $LaBr_3$:Ce crystal and the three investigated photon energies, illustrated by showing the resulting spatial resolution as expressed by the FWHM value of the error histogram as introduced in section 3.7.

There is no clear criterion for an unambiguous selection of the best set of parameters. Nonetheless, the architecture comprising five convolutional layers of three kernels with ReLU activation and batch normalization almost always provided the best results, was easy to train, and therefore was selected to serve as the final model. Figure 5 shows the sketch of the CNN architecture.

The same training was carried out for the $CeBr_3$ scintillator. A summary of the algorithm performance, expressed by the achieved spatial resolution via the FWHM of the error histogram, using the best final CNN architecture for both crystals as a function of the initial $\gamma$-ray energies is presented in figure 6.

### 4.2. Crystal performance

Figure 7 presents the spatial resolution achieved with the CNN algorithm as a function of the distance from the crystal central axis (0, 0) for three geometrical regions: central ($|x|, |y| \leqslant 8.5$ mm), intermediate (8.5 mm $< |x|, |y| \leqslant 17$ mm) and outer (17 mm $< |x|, |y| \leqslant 25.5$ mm).

It can be observed that the spatial resolution improves towards the crystal's central axis. This can be attributed to the less frequent scattering of scintillation light if compared to the crystal edges, distorting the light distribution on the PMT grid and enhancing absorption. The behavior of both $^{60}$Co curves is very similar,

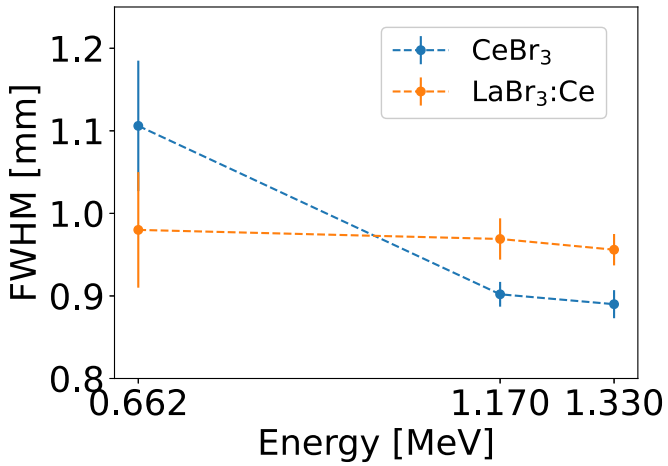


**Figure 6.** Spatial resolution achieved with the CNN algorithm (five convolutional layers, three kernels) as a function of the initial $\gamma$-ray energy, for the two scintillator crystals $LaBr_3$:Ce and $CeBr_3$.

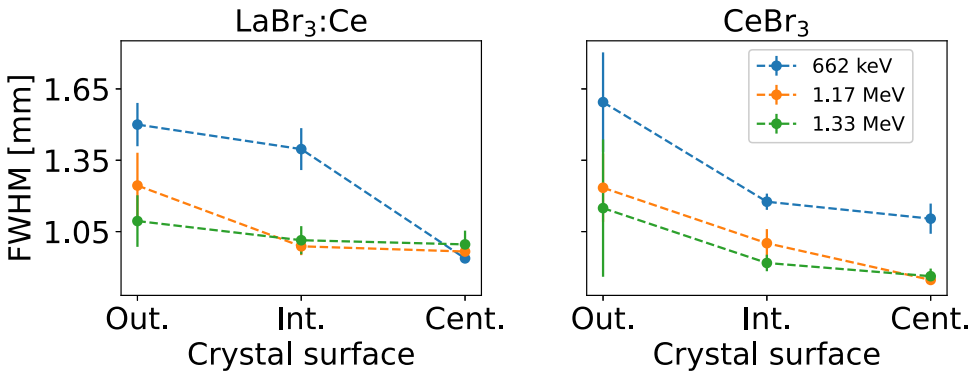


**Figure 7.** Spatial resolution achieved with the CNN algorithm for the $LaBr_3$:Ce and $CeBr_3$ monolithic scintillators as a function of the distance of the irradiation position from the crystal's central axis: outer, intermediate and central. The results for three investigated energies of the incident $\gamma$-rays are presented.

because of their small energy difference of only 0.16 MeV. The network performance for $^{137}$Cs is inferior to the higher energies of $^{60}$Co due to the lower photon energies. In fact, less energetic $\gamma$-rays tend to interact further away from the PMT's entrance window, increasing the likelihood of scattering and attenuation of the scintillation light on its path to the photosensor.

### 4.3. Database optimization

We determined the optimum number of events per irradiation position required for an efficient training to be $n_{\mathrm{epp|train}} = 400$ (350 for the training set and 50 for validation). Using smaller databases resulted in overfitting or led to suboptimal performance. Above that point the spatial resolution (FWHM) did not improve further.

For the optimization of the test set, error histograms corresponding to $n_{\mathrm{epp|test}} \leqslant 15$ suffered from too low statistics, so the use of a linear interpolation to determine the FWHM of these peaks could not be justified. It is necessary to use a validation dataset with $n_{\mathrm{epp|test}} \geqslant 20$. To get the uncertainty of the spatial resolution determination we used five sets with $n_{\mathrm{epp|test}} = 20$ events.

Table 1 summarizes the comparison between the CAP and CNN algorithms for the $LaBr_3$:Ce scintillation crystal concerning the computational time, memory footprint and the best spatial resolution achieved with both techniques.

Apart from the superior positioning accuracy, the CNN algorithm is up to $10^4$ times faster and has a much lower memory footprint compared to the CAP algorithm. This can be understood by analyzing the working principles of the CAP algorithm, which assumes comparisons of the unknown event with all entries from the extensive reference library being time and memory consuming. In contrary, the CNN requires a single forward pass through a fairly simple network.

**Table 1.** Quantitative comparison between the CAP and CNN algorithms for the $LaBr_3$:Ce scintillation crystal. The training time, reconstruction speed per one GPU (Quadro P1000), memory requirements and the best spatial resolution achieved with both techniques are given.

| Algorithm | Training time (h) | Reconstruction speed (events $s^{-1}$) | Memory footprint (GB) | Spatial resolution for $LaBr_3$:Ce at 1.33 MeV (mm) |
|---|---|---|---|---|
| CAP | — | 1 | 10 | 2.9(1) |
| CNN | 10 | $10^4$ | <1 | 0.96(2) |

## 5. Discussion

The determination of the photon interaction position inside a monolithic scintillation crystal has been performed so far using algorithms like kNN or its optimized version CAP. However, those suffered from long computational times and high memory requirements. This motivated us to develop an alternative reconstruction procedure that works faster and does not sacrifice the algorithm positioning performance. Even though the method presented here was developed for two large 50.8 mm × 50.8 mm × 30 mm monolithic $LaBr_3$:Ce and $CeBr_3$ scintillation crystals, coupled to 64-fold segmented photosensors with an 8 × 8 pixel arrangement, it can be applied to all kinds of detectors where a photon interaction position is associated with a characteristic light amplitude distribution.

The newly designed CNN consists of five convolutional layers, each comprising three 3 × 3 kernels with ReLU activation, batch normalization and zero padding applied. To generate predictions, separately for the $x$ and $y$ coordinates, the network was split into two output blocks, each made up of two fully connected layers with softmax as the final activation function. The coordinates with the highest probability are considered as the reconstructed position. Using more complicated architectures increased the training complexity and did not lead to an improvement in the performance.

Employing this method, we achieved an accuracy of the position determination reaching an optimum value of 0.9(±0.2)mm for $CeBr_3$ at $^{60}$Co photon energies of 1.17 and 1.33 MeV and under the same conditions for $LaBr_3$:Ce a resolution of 0.96 (±0.02)mm. This accuracy is superior by more than a factor of 2.5 compared to the best value of 2.9(0.1) mm achieved by the CAP algorithm for the $LaBr_3$:Ce scintillator at 1.3 MeV (Liprandi 2018). The lack of internal radioactivity in $CeBr_3$ explains its better performance if compared to $LaBr_3$:Ce. We observed a trend of slightly better performance for increasing photon energies which can be explained, firstly, by the larger penetration power at higher $\gamma$-ray energies, causing more energetic photons to interact closer to the PMT entrance window and secondly, by more secondary scintillation photons generated by higher energetic primary photons.

To the best of our knowledge, the CNN algorithm as presented here achieves the best reported spatial resolution for 2D position determination inside large monolithic scintillators.

On the computational side the new method is compatible with both CPUs and GPUs and allows to determine the interaction position of about $10^4$ events per second using a single GPU node (Nvidia Quadro P1000), which, when considering multi-mode instrumentation, will be sufficient for a potential real-time application in prompt-gamma imaging. This is a significant improvement over the CAP algorithm, that required a computer cluster and was able to reconstruct only about one event per second per single CPU node.

Exploiting large datasets of labeled data, that were originally acquired to serve as lookup tables for the kNN and CAP algorithms, we carried out supervised training. At each of 102 × 102 irradiation positions 460–800 photopeak events were collected and each of them was assigned two labels, corresponding to the $x$ and $y$ coordinates. Considering this task as a regression problem, in contrast to the classification paradigm, required a long training and did not lead to satisfying results.

An extensive database is crucial to provide a representative training set to train models that generalize well and are resistant to overfitting. Yet, the data acquisition is a time consuming process (takes up to a couple of weeks) and if possible should be kept at minimum. For this reason we set a lower threshold of 400 and 100 (5 × 20) events per irradiation position for training and testing, respectively. This facilitates a fast training without overfitting and smooth error histograms necessary for FWHM determination.

The accuracy of the network is best in the central area of the crystal and deteriorates slightly towards the outside regions. This reflects the larger probability of scattering and attenuation of the scintillation light at the scintillator edges compared to the more central interactions.

The error histograms derived from the comparison of calculated and ground-truth interaction positions exhibit a non-Gaussian shape (see figure 3) and thus their FWHM is not linearly related to the standard deviation $\sigma$ (as in the case of normal distribution where $\mathrm{FWHM} = 2\sqrt{2\ln 2}\,\sigma$). However, the FWHM of the error histogram still serves as a valid measure of the network performance, enabling a quantitative comparison

between different models. The performance of the previously utilized CAP and kNN algorithms was determined in the same way (also based on non-Gaussian peaks with the corresponding error histograms). Therefore, it is possible to perform a direct comparison between the previous, well-established methods and the new CNN algorithm.

We designed an optimum training schedule. Despite using the Adam optimizer, we suggest a gradual lowering of the initial learning rate value to enable faster convergence. In the beginning, while the parameters are still far from the optimum values, a larger learning rate value enables reaching bigger changes, while with a lower learning rate at the end of the training more precise adjustments are possible. We have noticed a benefit from adapting transferred learning, and instead of random parameter initialization we reused partially adjusted weights and biases. We adapted models trained on data originating from an irradiation of a second crystal of the same dimensions or including different $\gamma$-ray energies. The benefit lies in the reduction of training time as well as in a final model more robust and less prone to noisy data.

In our present study the CNN was trained with perpendicularly impinging gamma-ray beamlets. It will be left to further computational and experimental studies to test our assumption if the incidence angle of Compton-scattered photons does not influence the performance of the presented CNN algorithm.

The present reconstruction method can be implemented directly to determine the interaction position inside the here described absorber component of the CC for gamma quanta energies between 662 keV and 1.33 MeV. The interaction position reconstructed here is used later to calculate the Compton cones and their intersection in order to determine the origin of the primary gamma quanta. To make the algorithm applicable also for higher or lower energies, an additional training with events of corresponding energies should be performed.

## 6. Conclusions

We have developed a new CNN algorithm that provides an efficient way to determine precisely photon interaction positions in large monolithic scintillators, while preserving all benefits of the monolithic crystal (good energy resolution and timing). The algorithm outperforms previously used and well-established methods in terms of spatial resolution, memory requirements and speed. Those three factors are of particular importance for a future clinical application of the CC for ion beam range verification via prompt-gamma imaging.

To improve further the accuracy of the interaction position determination, which is currently limited by the $\gamma$-source diameter given by the collimator diameter of 1 mm, our future work will focus on performing a new reference scan using a collimator with an opening of 0.6 mm. Additionally, we will concentrate on deriving the depth of interaction information by incorporating into the training data a new reference library consisting of events originating from an irradiation of the crystal's side surfaces, thus extending the preset 2D methodology into three-dimensions. Further work should focus also on scaling the presented method to higher gamma energies preferably up to the multi-MeV range which is relevant for prompt-gamma Compton imaging in particle therapy ion beam range verification.

## Acknowledgments

We wish to thank Dr Ben Hoyle for his contribution to this project by introducing us to the topic of convolutional neural networks as well as Martin Rädler for help in creating figures for this manuscript. This work was supported by the DFG Cluster of Excellence MAP (Munich-Centre forAdvanced Photonics).

## ORCID iDs

M Kawula https://orcid.org/0000-0002-5039-8278
K Parodi https://orcid.org/0000-0001-7779-6690

## References

Aldawood S *et al* 2015 Comparative characterization study of a $LaBr_3$ (Ce) scintillation crystal in two surface wrapping scenarios: absorptive and reflective *Front. Oncol.* **5**
Aldawood S *et al* 2017 Development of a Compton camera for prompt-gamma medical imaging *Radiat. Phys. Chem.* **140** 190–7
Alkharusi H 2012 Categorical variables in regression analysis: a comparison of dummy and effect coding *Int. J. Educ.* **4** 202–10
Binder T M 2017 Evaluation of new components for the absorber detector of the Garching Compton camera prototype *Master's Thesis* Ludwig-Maximilians-Universität München
Golnik C *et al* 2014 Range assessment in particle therapy based on prompt $\gamma$-ray timing measurements *Phys. Med. Biol.* **59** 5399–422
Gonzalez R and Woods R 2018 *Digital Image Processing* 3rd (University of Tennessee: Pearson)

Gwosch K et al 2013 Non-invasive monitoring of therapeutic carbon ion beams in a homogeneous phantom by tracking of secondary ions *Phys. Med. Biol.* **58** 3755–73
Photonics, Hamamatsu 2015 Flat panel type multianode PMT assembly *H8500 series/ H10966 series*
Hoffman J 2009 TRIVA, VME trigger module https://gsi.de/fileadmin/EE/Module/TRIVA/triva5_5.pdf
Ioffe S and Szegedy C 2015 Batch normalization: accelerating deep network training by reducing internal covariate shift arXiv:1502.03167
Kellnberger S et al 2016 Ionoacoustic tomography of the proton Bragg peak in combination with ultrasound and optoacoustic imaging *Sci. Rep.* **6** 29305
Krimmer J et al 2015 Development of a Compton camera for medical applications based on silicon strip and scintillation detectors *Nucl. Instrum. Methods Phys. Res.* A **787** 98–101
Lang C et al 2014 Sub-millimeter nuclear medical imaging with high sensitivity in positron emission tomography using $\beta+\gamma$ coincidences *J. Instrum.* **9** P01008
Liprandi S 2018 Development and performance evaluation of detectors in a Compton camera arrangement for ion beam range monitoring in particle therapy *PhD Thesis* Ludwig-Maximilians-Universität München
Llosá G et al 2016 First images of a three-layer Compton telescope prototype for treatment monitoring in hadron therapy *Front. Oncol.* **6** 14
Manzano L G et al 2015 Xemis: a liquid xenon detector for medical imaging *Nucl. Instrum. Methods Phys. Res.* A **787** 89–93
Polf J C et al 2015 Imaging of prompt gamma rays emitted during delivery of clinical proton beams with a Compton camera: feasibility studies for range verification *Phys. Med. Biol.* **60** 7085–99
Saint-Gobain 2016 Lanthanum bromide and enhanced lanthanum bromide (https://www.crystals.saint-gobain.com/sites/imdf.crystals.com/files/documents/lanthanum-material-data-sheet.pdf)
Scionix 2018 High resolution low background $CeBr_3$ scintillators https://scionix.nl/wp-content/uploads/2017/07/CeBr3-scintillation-detectors.pdf
Smeets J et al 2016 Experimental comparison of knife-edge and multi-parallel slit collimators for prompt gamma imaging of proton pencil beams *Front. Oncol.* **6** 156
van Dam H T et al 2011 Improved nearest neighbor methods for gamma photon interaction position determination in monolithic scintillator pet detectors *IEEE Transactions on Nuclear Science* **58** 2139–47
Verburg J M and Seco J 2014 Proton range verification through prompt gamma-ray spectroscopy *Phys. Med. Biol.* **59** 7089–106
Viegas R 2018 Optimization study of the performance of a $LaBr_3$ monolithic scintillator in a Compton camera system *Master's Thesis* University of Coimbra
Yoshida E et al 2020 Whole gamma imaging: a new concept of PET combined with Compton imaging *Phys. Med. Biol.* **65** 125013
Zhu X and El Fakhri G 2013 Proton therapy verification with PET imaging *Theranostics* **3** 731–40