Concept Development of an On-Chip PET System

C. Clement¹, G. Birindelli¹, M. Pizzichemi²,³, F. Pagano²,³, M. Kruithof-de Julio⁴, A. Rominger¹, S. Ziegler⁵, E. Auffray² and K. Shi¹

Abstract—Organs-on-Chips (OOCs), microdevices mimicking in vivo organs, find growing applications in disease modeling and drug discovery. With the increasing number of uses comes a strong demand for imaging capabilities of OOCs. Positron Emission Tomography (PET) would be ideal for OOC imaging, however, current PET systems have insufficient spatial resolution for this task. In this work, we propose the concept of an On-Chip PET system capable of imaging OOCs. Our system consists of four detectors arranged around the OOC device. Each detector is made of two monolithic Lutetium–yttrium oxyorthosilicate (LYSO) crystals and covered with Silicon photomultipliers (SiPMs) on multiple surfaces. We use a Convolutional Neural Network (CNN) trained with data from a Monte Carlo Simulation (MCS) to predict the first gamma-ray interaction position inside the detector from the light patterns that are recorded by the SiPMs on the detector’s surfaces. With the Line of Responses (LORs) created by the predicted interaction positions, we reconstruct with Simultaneous Algebraic Reconstruction Technique (SART). The CNN achieves a mean average prediction error of 0.78 mm in the best configuration. We use the trained network to reconstruct an image of a grid of 21 point sources spread across the field-of-view and obtain a mean spatial resolution of 0.53 mm. We demonstrate that it is possible to achieve a spatial resolution of almost 0.5 mm in a PET system made of multiple monolithic LYSO crystals by directly predicting the scintillation position from light patterns created with SiPMs. We observe that CNNs from the ResNet family perform better than those from the EfficientNet family and that certain surfaces encode significantly more information for the scintillation-point prediction than others.

I. BACKGROUND

OOCs are microdevices that mimic in vivo organs. They contain 3D tissue cultures connected by microfluidic channels [1]. OOCs have sparked the interest of researchers in the past decade, especially for the drug discovery and development process [2].

With the growing numbers of uses for OOCs comes an increasing demand for novel measurement capabilities. Monitoring metabolism and other physiologic processes in OOCs is critical to refine the technology to closely resemble in vivo physiology and promote its application in new biological models. PET would be the ideal candidate for OOC imaging due to its ability to retrieve in vivo information about metabolism and molecular pathways [3]. However, current imaging devices for measuring PET tracer uptake are inadequate for the task of OOC imaging due to their limited spatial resolution [4].

In recent years, there has been a trend in pre-clinical PET research towards using monolithic instead of pixelated crystals as detectors to increase the spatial resolution. The resolution in monolithic crystals is not inherently limited by the pixel size but can be improved with more advanced readout schemes and data processing methods. The key to increasing the resolution is to predict the first gamma-ray interaction position in the detector as precisely as possible [5].

The work by Wang et al. [6] introduced a monolithic PET detector system that can estimate gamma-ray interaction positions with Neural Networks (NNs). The input data were created with a simplified readout scheme with signals from a Photomultiplier Tube (PMT) on one side of the crystal. The system achieves spatial and Depth of Interaction (DOI) resolutions of 2.0 mm.

Sanaat and Zaidi [7] presented another approach to estimate the DOI in a monolithic crystal using a NN. They trained a multilayer perceptron that outputs the 3D gamma-ray interaction position with data from a MCS and reached a spatial resolution of 1.54 mm in the x-y plane.

Jaliparthi et al. [8] developed AnnPET, a monolithic annular PET system consisting of a single annulus shaped LYSO crystal with SiPM arrays attached to its outer surfaces. They employed a ten-layer CNN to estimate the gamma-ray interaction position and reached single dimension Mean Absolute Error (MAE) values between 0.42 mm and 0.54 mm. When using the trained network for reconstruction, they achieved Full Width at Half Maximum (FWHM) values between 0.71 and 0.80 mm.

In this work, we propose an On-Chip PET system to make functional imaging of OOCs possible. The novelties presented in this work are twofold. First, we design a scanner made up of four detectors that consist of two glued-together monolithic crystals each. Second, we train a CNN directly with the light pattern images that emerge on the surfaces of the detectors to predict the first scintillation positions inside the detectors. We optimize the design of the system with a MCS to create datasets of light pattern images emerging on the surfaces of the detectors through scintillation. With these datasets, we train and evaluate CNNs that predict the first interaction positions of the gamma rays inside the detector.
II. METHODS

A. Monte Carlo Simulation

We model the interaction of the proposed system with a back-to-back gamma source in a MCS built with the Geant4 Application for Emission Tomography (GATE) tool [9]. GATE enables the creation of MCSs in the field of nuclear medicine through a macro language that controls the experimental settings such as the geometry, physics processes, surface treatments, and source setup.

1) Geometry: We use the most generic system in GATE, the scanner to set up the geometry. The scanner is defined as a box-shaped volume and placed in the world volume. The purpose of the scanner volume is to encapsulate our proposed PET system. The box-shaped detector volume is placed inside the scanner volume and repeated four times with a ring repeater around the z-axis. Inside the detector volume, the box-shaped crystal volume is placed and repeated two times with a linear repeater. Table I contains the lengths and materials of each volume. All dimensions of the detector are chosen such that arrays of commercially available SiPMs fit on the surfaces.

| Name    | Type | Lengths [mm] | Material |
|---------|------|--------------|----------|
| world   | box  | 126.3        | Vacuum   |
| scanner | box  | 114.8        | Vacuum   |
| detector| box  | 52.2, 13.1, 104.4 | Epoxy   |
| crystal | box  | 52.0, 13.0, 52.0 | LYSO    |

2) Physics & Cuts: As a physics list, the Electromagnetics (EM) constructor with option four is chosen, which uses the most accurate standard and low-energy models available in Geant4. We set a cut of 0.1 mm for gammas, electrons, and positrons in the scanner volume.

3) Surfaces: We use Geant4’s unified model to define the surfaces in the simulation. The surfaces between the detector and crystal volumes are dielectric-dielectric ones with a ground finish and a sigmaalpha value of 0.01 corresponding to a typical polished crystal. Their specular lobe constant is set to 1.0.

4) Source: We add a source that emits 511 keV back-to-back gamma particles with an activity of 1,000 Bq to the simulation setup. To create the training dataset, the position of the source is sampled from a box volume with dimensions of 10.0 mm x 26.0 mm x 76.0 mm resembling the size of an OOC device. To evaluate the spatial resolution of the system, another dataset is created where 21 point sources are arranged in a 7 x 3 grid with a distance of 10 mm in between each source.

Figure 1 depicts the simulation setup viewed from the front. The volume from which the source position is sampled to create the training dataset is shown in green.

B. Dataset Creation

We post-process the hits output files from the GATE simulation runs to create the datasets used to train the CNNs. For each primary event, in which two back-to-back gamma-rays are emitted, zero to two samples of the training dataset are created. The amount of created samples depends on the number of detectors in which the gamma-rays interact. If there is a scintillation event in at least one detector, the position of the first interaction of the gamma ray as well as the corresponding light patterns that emerge on the surfaces of the detectors are saved. To create the light patterns, we simulate the SiPMs that are attached to the surfaces of the detector, which come from Hamamatsu’s S14161 series with a sensitive-area size of 3 mm [10].

One example sample consisting of five light pattern images with the scintillation point is depicted in Figure 2.

Fig. 1. Simulation setup viewed from the front. The yellow volumes represent the crystals, the blue ones the epoxy layer around the crystal, and the green one the volume from which the source position is sampled to create the training dataset.

Fig. 2. One sample consisting of five light pattern images. The light patterns are recorded with the SiPMs on all surfaces of the detector except the front one. The light patterns are padded such that they all have the same square shape. The green dot is the scintillation position of the gamma-ray inside the crystal projected onto each surface.
C. Scintillation Position Prediction

With the created dataset described in the previous section, we train a CNN that predicts the gamma-ray interaction position inside the crystal. The input to the network are the stacked light patterns recorded with SiPMs on the detectors’ surfaces. With the light pattern images as input, the network should predict the gamma-ray interaction positions.

In this work, we investigate the influences of the network architecture and the number of surfaces covered with SiPMs on the scintillation-position prediction performance. For the network architecture, we selected variants from the ResNet [11] and EfficientNet [12] families with different depths, as these architectures have shown great performances on image data. The parameters given in Table II are used for every experiment run.

TABLE II
PARAMETERS USED FOR EVERY EXPERIMENT RUN.

| Parameter          | Value                     |
|--------------------|---------------------------|
| Dataset Splits     | 800k train, 200k val, 10k test |
| Input Tensor Size  | Cx32x32                   |
| Loss Function      | Mean Absolute Error (MAE) |
| Optimizer          | Adam [13]                 |
| Learning Rate      | 3e-4                      |
| Batch Size         | 256                       |

D. Reconstruction

We evaluate the reconstruction performance of our proposed system with a grid of point sources, as described in section II-A.4. We perform the following steps for SART:

1) Load all predicted pairs of scintillation positions.
2) Compute the distance and angle of the LOR to the origin for each LOR that is defined by the pair of predicted scintillation positions.
3) Create the sinogram from the LORs by computing the 2D histogram of the distances and angles with a bin size of 400 in both dimensions.
4) Generate the corresponding reconstructed image by running five iterations of SART [14] [15].

III. RESULTS & DISCUSSION

A. Scintillation Position Prediction

1) Network Architecture: With the first set of CNN training runs, we determined the best network architecture among variants from the ResNet and EfficientNet families with different depths. Table III shows the prediction performances of the different network architectures on the validation dataset. ResNet50 and ResNet18 are the architecture that perform best with a MAE value 0.80 mm closely followed by ResNet101 and ResNet152. The three architectures from the EfficientNet family achieve results between 0.85 and 1.00 mm. We observe that the ResNet architectures is better suited for predicting the first scintillation position. A depth of 50 seems to be the perfect trade-off between number of network parameters and overfitting as shaller and deeper networks perform slightly worse.

2) Combination of Light Patterns as Input: With the second set of CNN training runs, we determined the optimal combination of input light patterns. All possible combinations of the five surfaces, in total 31, are evaluated. The training parameters are those from Table II. Table IV contains the prediction results. The best ten combinations achieve MAE values between 0.79 mm and 0.80 mm. The worst performances are achieved with the top and bottom surfaces with a MAE of around 3.9 mm. These results indicate that the light patterns from certain surfaces, especially the back one, encode significantly more information about the scintillation position than others. This behaviour can be expected as the back surface is the largest of all surfaces and leads to a practical implication for the design of the prototype: we do not need to cover all five surfaces with SiPMs to achieve a good scintillation-position prediction performance.

TABLE III
SCINTILLATION-POSITION PREDICTION RESULTS OF DIFFERENT NETWORK ARCHITECTURES ON THE VALIDATION DATASET.

| Network Architecture | Mean Absolute Error (MAE) [mm] |
|----------------------|--------------------------------|
| ResNet50             | 0.80                           |
| ResNet18             | 0.80                           |
| ResNet101            | 0.82                           |
| ResNet152            | 0.82                           |
| EfficientNet-B0      | 0.85                           |
| EfficientNet-B4      | 0.91                           |
| EfficientNet-B7      | 1.00                           |

TABLE IV
SCINTILLATION-POSITION PREDICTION RESULTS OF DIFFERENT SURFACES AS INPUT. THE BEST PERFORMING COMBINATION FOR EVERY NUMBER OF LIGHT PATTERNS IS SHOWN IN BOLD.

| #  | Surface Count | Surfaces     | Mean Absolute Error (MAE) |
|----|---------------|--------------|---------------------------|
| 1  | 3             | ba-bo-le     | 0.79                      |
| 2  | 3             | ba-le-ri     | 0.79                      |
| 3  | 4             | ba-le-ri-to  | 0.80                      |
| 4  | 3             | ba-bo-to     | 0.80                      |
| 5  | 3             | ba-le-to     | 0.80                      |
| 6  | 5             | ba-bo-le-ri-to | 0.80                   |
| 7  | 3             | ba-bo-ri     | 0.80                      |
| 8  | 1             | ba           | 0.80                      |
| 9  | 2             | ba-bo        | 0.80                      |
| 10 | 3             | ba-ri-bo     | 0.80                      |

3) Best Result on Test Dataset: We trained a network with the best performing design choices from the two experiment runs - the ResNet50 architecture and the back-bottom-left surface combination - with the training dataset consisting of one million samples. This network achieves a MAE of 0.78 mm on the test dataset and is used for the reconstruction.

B. Reconstruction

The SART steps from section II-D are performed to create the sinogram and reconstructed image shown in Figure 3.
The mean FWHM value taken over the 21 point sources is 0.53 mm with a standard deviation of 0.14 mm. The FWHM values are computed by averaging the peak half widths of vertical and horizontal profiles drawn through the sinogram. The profiles are Gaussian filtered with a sigma of 0.5. These results show a good reconstruction performance over the entire OOC device.

IV. CONCLUSION

In this work, we designed a system consisting of four detectors where each is made up of two monolithic LYSO crystals with SiPMs attached to multiple surfaces. We generated training, testing, and reconstruction datasets with a MCS of the system and observed that the ResNet50 architecture achieved the best scintillation-position prediction results with a MAE value of 0.78 mm on the test dataset. With the trained network, we reconstructed a grid of point sources using SART and reached a mean FWHM value of 0.53 mm.

The results from the scintillation-position prediction and reconstruction demonstrate the capability of our system to achieve a resolution of almost 0.5 mm for a large field-of-view using out-of-the-box reconstruction methods. As next steps, we will shift our focus from interaction-position prediction to reconstruction to achieve an even better resolution. We will develop a list mode-based reconstruction method incorporating the geometrical priors that our system and OOC devices are constrained by.

Some limitations of this work are that we have simulated the detector in Vacuum and without any plastic or other attenuating material around the source. Furthermore, we did not take into account an energy window and dead time of the electronics.

Our results showed practical implications that play a crucial role in the next project steps, where we are going to build a prototype of the proposed system. We observed that the back surface encodes significantly more information about the scintillation position compared to other surfaces. Therefore, we concluded that not all five surfaces of each detector need to be covered with SiPMs. This reduces the number of channels that need to be read out individually.

REFERENCES

[1] L. A. Low, C. Mummery, B. R. Berridge, C. P. Austin, and D. A. Tagle, “Organs-on-chips: Into the next decade,” Nature Reviews Drug Discovery, vol. 20, pp. 345–361, May 2021.
[2] E. W. Esch, A. Bahinski, and D. Huh, “Organs-on-chips at the frontiers of drug discovery,” Nature Reviews. Drug Discovery, vol. 14, pp. 248–260, Apr. 2015.
[3] S. S. Gambhir, “Molecular imaging of cancer with positron emission tomography,” Nature Reviews. Cancer, vol. 2, pp. 683–693, Sept. 2002.
[4] G. Pratx, K. Chen, C. Sun, L. Martin, C. M. Carpenter, P. D. Olcott, and L. Xing, “Radioluminescence Microscopy: Measuring the Heterogeneous Uptake of Radiotracers in Single Living Cells,” PLOS ONE, vol. 7, p. e46285, Oct. 2012.
[5] T. Jones and D. Townsend, “History and future technical innovation in positron emission tomography,” Journal of Medical Imaging (Bellingham, Wash.), vol. 4, p. 011013, Jan. 2017.
[6] Y. Wang, W. Zhu, X. Cheng, and D. Li, “3D position estimation using an artificial neural network for a continuous scintillator PET detector,” Physics in Medicine and Biology, vol. 58, pp. 1375–1390, Feb. 2013.
[7] A. Sanaat and H. Zaidi, “Accurate estimation of depth of interaction in PET on monolithic crystal coupled to SiPMs using a deep neural network and Monte Carlo simulations,” in 2019 IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC), pp. 1–3, Oct. 2019.
[8] G. Jaliparthi, P. F. Martone, A. V. Stolin, and R. R. Raylman, “Deep residual-convolutional neural networks for event positioning in a monolithic annular PET scanner,” Physics in Medicine & Biology, vol. 66, p. 145008, July 2021.
[9] S. e. a. Jan, “GATE: A simulation toolkit for PET and SPECT,” Physics in Medicine and Biology, vol. 49, pp. 4543–4561, Oct. 2004.
[10] H. Photonics, “S14160/S14161 series,” June 2020.
[11] K. He, X. Zhang, S. Ren, and J. Sun, “Deep Residual Learning for Image Recognition,” arXiv:1512.03385 [cs], Dec. 2015.
[12] M. Tan and Q. V. Le, “EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks,” arXiv:1905.11946 [cs, stat], Sept. 2020.
[13] D. P. Kingma and J. Ba, “Adam: A Method for Stochastic Optimization,” arXiv:1412.6980 [cs], Jan. 2017.
[14] A. H. Andersen and A. C. Kak, “Simultaneous algebraic reconstruction technique (SART): A superior implementation of the art algorithm,” Ultrasonic Imaging, vol. 6, pp. 81–94, Jan. 1984.
[15] S. van der Walt, J. L. Schönberger, J. Nunez-Iglesias, F. Boulogne, J. D. Warner, N. Yager, E. Gouillart, T. Yu, and t. scikit-image contributors, “Scikit-image: Image processing in Python,” PeerJ, vol. 2, p. e453, June 2014.