Hadron therapy range verification via Machine-Learning aided prompt-gamma imaging

J. Balibrea-Correa, J. Lerendegui-Marco, V. Babiano-Suarez, C. Domingo-Pardo, I. Ladarescu, C. Guerrero, T. Rodríguez-González, M. C. Jiménez-Ramos

Abstract—The aim of this work is to demonstrate the capability of the i-TED Compton imager combined with ML-techniques for enhanced performance range verification in prompt gamma-ray monitoring. Prompt Gamma monitoring constitutes a promising technique for range verification in hadron therapy treatments. Hadron Therapy with protons introduces advantages with respect to the conventional radiotherapy because of the maximization of the energy deposition (dose) at the Bragg peak. i-TED is an advanced array of Compton cameras originally designed for neutron-capture nuclear experiments. However, due to its large detection efficiency, fast response, high time resolution, compactness, low sensitivity to neutron-induced backgrounds and image resolution, i-TED shows also an excellent performance for medical purposes such as PG monitoring. Furthermore, aiming at improved quality Compton images in the high-energy gamma-ray range characteristic of Hadron Therapy, a novel Machine Learning methodology has been developed and applied for identification of full-energy events. To that purpose, a detailed GEANT4 Monte Carlo study simulating a clinical irradiation has been performed. Finally, we present first results of the novel Machine Learning methodology applied to an experiment performed at the cyclotron of the CNA facility, Spain, where two i-TED modules were simultaneously operated as in-beam PET and Compton imagers.

I. INTRODUCTION

Hadron therapy, and in particular proton therapy, has become a popular treatment for certain type of cancers [1]. This is reflected on more than 220000 patients treated worldwide by 2019 and the increasing number of hadron therapy centers, with a total of 111 centers in operation, 39 under construction and 29 in planning stage in 2020 [2], [3].

This work was supported by the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovative programme (ERC Consolidator Grant Project HYMNS, grant agreement 681740). The authors acknowledge partial support from the Spanish MICINN grants and FPA2017-83946-C2-1-P and PID2019-104714GB-C21.

J. Balibrea-Correa IFIC (CSIC-University of Valencia) Paterna 46980 Spain (e-mail: javier.balibrea@ific.uv.es).
J. Lerendegui-Marco IFIC (CSIC-University of Valencia) Paterna 46980 Spain (e-mail: jroge.leredegui@ific.uv.es).
V. Babiano-Suarez, IFIC (CSIC-University of Valencia) Paterna 46980 Spain (e-mail: babiano@ific.uv.es).
I. Ladarescu, IFIC (CSIC-University of Valencia) Paterna 46980 Spain (e-mail: ladarescu@ific.uv.es).
C. Guerrero, (University of Seville) Seville, 41004 Spain (e-mail: cguerreror4@us.es).
T. Rodríguez-González (University of Seville) Seville, 41004 Spain (e-mail: mrodriguezg@us.es).
M.C Jiménez-Ramos (Centro Nacional de Aceleradores), Seville, 41004 Spain (e-mail: mcjyr@us.es).

Proton therapy presents advantages with respect to conventional radiotherapy because of the maximization of the dose in the vicinity of the Bragg peak in opposite to conventional radiotherapy. Therefore, an improved targeting of the tumor region can be attained, thereby minimizing the dose received by the neighboring tissues and reducing the long-term secondary effects on the patients [4]. Nowadays there are still some limitations on the technique associated to uncertainties on anatomical changes, patient setup errors and particle stopping power unreliability in different materials. Such limitations lead to an increase of the safety margins and thus to a significant reduction of the potential benefits of proton therapy [5].

In this context, several experimental proton range verification methodologies are under development, mainly based on the Prompt Gamma (PG) rays produced in nuclear reactions during the proton irradiations and monitoring $\beta^+$ decay of byproduct unstable nuclei short time after the irradiation [6]. The latter is usually made by using a commercial PET scan, which can be performed only a minutes after the proton irradiation. Because of the time gap between the proton irradiation and the PET scan, the biological wash-out constrains the precision of such monitoring technique. On the other hand, PG monitoring offers the advantage of the high spatial correlation with the dose distribution [7]. However, there are still important challenges for implementing PG monitoring in clinical conditions, such as the limited detection efficiency [8], [9], large counting rates registered by the detection systems [10] and small signal compared to the background present in the irradiation room [11], [12], [13].

In this contribution we discuss the applicability of the i-TED Compton imager [14], [15] in order to overcome some of the limitations and challenges discussed before. Specifically, i-TED consists of an array of four two-plane Compton cameras, which cover a large solid angle and offer an intrinsically small neutron sensitivity. Each of the four i-TED modules is based on five largest-size commercially available monolithic LaCl₃ crystals of 50x50 mm², thus featuring a sensitive volume of 1150 cm³ with 3D position sensitive capabilities for the full i-TED array. For further details the reader is referred to [14], [15], [16], [17], [18]. This Compton imager was originally developed for neutron-capture time-of-flight nuclear experiments [19], where the radiative capture channel of interest is very weak compared to the $\gamma$-ray neutron-induced backgrounds. Other requirements for neutron capture cross-section experiments comprise an excellent time-resolution and large dynamic range to cope with $\gamma$-ray energies from a few
hundreds of keV up to 7-8 MeV. These experimental challenges are comparable to those encountered in HT treatments and proton range-monitoring, where the PG γ-ray transitions are mainly 2.3 MeV (14N), 4.4 MeV (12C), 5.25 MeV (16O) and 6.1 MeV (16O) [20]. Also, the large neutron induced backgrounds are a common aspect for both types of application [21].

One of the critical aspects for an accurate Compton image and ion-range verification using a two plane Compton Camera as i-TED is the correct identification of those events where the incoming γ-ray has been completely absorbed, hereafter referred to as full-energy events. This critical point is easily explained from the Compton angle (θc) formula for two-planes Compton cameras described by Eq. 1

$$\cos \theta_c = 1 - \frac{m_e^2 c^2 E_s}{(E_s + E_a)E_a}$$ (1)

where $E_s$ and $E_a$ are the deposited energy in the scatter and absorber planes, respectively, and $m_e^2 c^2$ is the rest mass of the electron. Eq. 1 is correct only for full-energy events. Any event detected in time-coincidence between both detection planes that does not hold $E_s = E_a + E_a$ will lead to an incorrect $\theta_c$ and consequently to a deterioration of the final image and of the proton-range assessment. A hardware solution to this problem is the use of three or more detection layers [13], [22], [23], however at the cost of detection efficiency, which is one of the present requirements and limitations [8].

In this work we have explored the use of a Machine Learning (ML) algorithm, specifically a deep neural network classifier, aiming at suppressing non full-energy events for the Compton imaging reconstruction. In the last part of this work we apply this ML classifier to a real experiment performed at CNA (Seville) to validate the results obtained by means of Monte Carlo (MC) simulations.

II. MATERIAL AND METHODS

In order to perform the training and evaluation of the ML classifier, dedicated intensive MC simulations based on GEANT4 10.6 toolkit [24] were performed using the QGSP_INCLXX_HP physics list [25]. A figure of the MC setup is sketched in Fig. 1. The setup consists of four i-TED modules placed at 50 mm from the frontal face of a 10 × 10 × 20 cm³ cubic phantom. For each event registered in the sensitive detectors, the position and deposited energy was recorded for a posterior reconstruction in the same way as the experimental data is treated. During the MC reconstruction, experimental effects such as resolutions in deposited energy and γ-ray interaction position were included according to the values obtained from a laboratory characterization [18], [26].

For the ML classifier, dedicated MC simulations for the training stage were computed with γ-rays energies from 0.2 MeV up to 7 MeV. The γ-rays were emitted randomly within the phantom, thus replicating the field of view of all i-TED modules during a real application. In this specific case, air was the material used for the phantom. The ML discrimination used is a deep neural network implemented in TENSORFLOW [27]. The deep neural network classifier consists of an stack of four layers with 256 units each and RELU as activation function. The output layer is made of a single unit with SIGMOID as activation function. The model was compiled using BINARY CROSS-ENTROPY as loss function, usually applied to this type of classification problem [28].

As an input for the ML classifier, we used the deposited energy and position reconstruction in each detection plane of the independent i-TED modules ($E_s, x_s, y_s, z_s, E_a, x_a, y_a, z_a$) and the Compton probability given by the Klein-Nishima Compton differential cross-section [29], assuming energy $E_s = E_a + E_a$, i.e., the registered event is a full-energy event. This last feature was introduced aiming at a further improvement of the ML classifier accuracy. Based on the best performance in terms of accuracy, the ML classifier was trained in several energy ranges, covering from 0.2 to 7 MeV in intervals of 0.5 MeV, leading to 14 independent deposited energy ranges within the range of interest. Validation MC simulations were performed using PMMA and water materials as a replacement for the aforementioned air phantom. In this case, a 180 MeV proton beam with gaussian shape ($\sigma=3$mm) impinging on the center of the phantom was used, as indicated in Fig. 1. The goal of these MC simulations is to investigate the potential of the ML classifier and its impact in the quality of the reconstructed Compton images.

Fig. 2 shows a 2D representation from the MC calculation of the γ-rays emitted from the phantom as a function of depth in water. The end-point (~10 cm) corresponds to the Bragg peak position. A similar result is obtained for PMMA material. One can clearly appreciate some of the γ-ray lines best suited for proton range verification: 2.3 MeV (14N), 4.4 MeV (12C), 5.25 MeV (16O) and 5.25 MeV (18O)}.
MeV ($^{15}$O) and 6.1 MeV ($^{16}$O) [20].

For the Compton imaging reconstruction we have used an analytical inversion of the Compton imaging formula based on an infinite spherical harmonics series developed by Tomotani and Hisarawa in 2002 [30]. The approximate solution given a position in the image space, $\mathbf{s}$, is described by:

$$f(\mathbf{s}) \approx \int_{\cos \theta_{\text{min}}}^{\cos \theta_{\text{max}}} d\cos \theta_c \int_S d\mathbf{t} k^{-1}(\mathbf{t}; p; \cos \theta_c) g(\mathbf{t}; \cos \theta_c),$$

(2)

where $\mathbf{t}$ is the unit vector into the projection space, $\theta_{\text{min}}$ and $\theta_{\text{max}}$ are the minimum and maximum Compton scattering angles that the Compton camera can measure, $g(\mathbf{t}; \cos \theta_c)$ is the projection data in the image space and $k^{-1}(\mathbf{t}; p; \cos \theta_c)$ is the inversion kernel described as $k^{-1}(\mathbf{t}; p; \cos \theta_c) = \sum_{n=-N}^{N} \frac{2n+1}{4 \pi H_n} P_n(\cos \theta_c) P_n(\mathbf{s} \cdot \mathbf{t})$, with $H_n$ given by

$$H_n = \int_{\cos \theta_{\text{min}}}^{\cos \theta_{\text{max}}} \sigma(\cos \theta_c) P_n^2(\cos \theta_c) d\cos \theta_c.$$  

(3)

In the formula, $N_{\text{max}}$ is the maximum number of terms involved in the series, $P_n$ is the Legendre polynomial of order $n$ and $\sigma(\cos \theta_c)$ is the Klein-Nishima Compton differential cross-section [29]. The drawback of this algorithm is the large computational cost required to reconstruct a Compton image. Because of this, $H_n$ was pre-computed for a wide range of $\gamma$-ray energies and Compton scattering angles corresponding to the minimum and maximum values that can be registered by the detection system experimentally. Thus, $H_n$ values were collected in a table format for its posterior use, saving a crucial computational time. Additionally, the algorithm was implemented for GPUs using the CUDA toolkit [31]. With the latter the calculation was speed-up by a factor 121 with respect to the single-threaded version of the same code.

III. PRELIMINARY RESULTS FROM MONTE CARLO SIMULATIONS

The ML classifier evaluation was performed using the dedicated validation MC configuration explained in section II.

Fig. 2. 2D histogram of $\gamma$-ray energy emitted from the phantom and depth in water from the validation MC simulations.

Fig. 3. Peak to No peak gain factor ratio as a function of the add-back deposited energy. The black and red lines represent, respectively, the results including only $\gamma$-ray events and $\gamma$-ray and neutron events.

Fig. 4. Compton images 1D projections along the proton beam axis combining the four main transitions in PG. See text for details.

Fig. 3 shows the Peak- to No-peak gain factor ratio as a function of the add-back energy, $E_s + E_a$, spanning from 0.2 up to 7 MeV. This magnitude is defined as the ratio between the number of full energy events identified correctly by the ML classifier and the total full energy events present in the same add-back energy period. The results including only $\gamma$-rays are shown with a black line, whereas the contribution from both $\gamma$-rays and neutron interactions are displayed by the red line. These results demonstrate the improvement attained when the ML classifier is applied, both with and without the neutron contributions. The performance varies as a function of the add-back energy, $E_s + E_a$, with a maximum gain-factor value of 1.9 at $\sim$1.8 MeV, which decreases down to a gain-factor of 1.35 at 7 MeV. This results suggest that a two-plane Compton camera, in combination with specific ML algorithms, can indeed provide a competitive solution for ion-range verification, without compromising detection efficiency for a true real-time performance.

The attainable improvement is even more apparent when this methodology is applied to Compton imaging reconstruction.
Fig. 5. Picture of the experimental setup. The different parts such as detectors, samples holder and others are labelled for a comprehensive understanding of the experimental arrangement.

Fig. 4 displays the Compton images projected along the proton beam axis. The three solid curves correspond to the images using all registered events (red), only full-energy $\gamma$-ray events (purple), and after the ML classification has been applied (green). The true proton-depth profile is displayed by the dashed black line. The distal position can be fully recovered after the ML classifier has been applied to the data, thereby delivering results which are remarkably close to the ideal case of only full-energy events.

IV. PRELIMINARY RESULTS FROM EXPERIMENTAL DATA

A first i-TED in-beam test experiment was carried out at the Cyclotron facility of the Centro Nacional de Aceleradores CNA in Seville [32]. The main aim of this beam-time was the study of $\beta^+$ emitters production cross sections [33], [34]. However, it turned out to be also a convenient experiment to simultaneously check the suitability of the i-TED detection device for PG monitoring and in-beam PET imaging. The CNA Cyclotron delivers a 18 MeV proton beam to the dedicated radiobiological experimental line with currents up to $\sim 80\ \mu A$ [35], [36]. Fig. 5 shows a part of the experimental setup, which was placed in a dedicated bunker. The i-TED detection system consisted of two Compton modules (i-TED.A and i-TED.B), placed front to front, and covering a $100 \times 100\ \text{mm}^2$ PET Field of View (FoV). A sample holder was placed between both i-TED modules fully aligned with entrance of the proton beam. Five thin samples or targets $(41.2 \times 41.2 \times 0.8\ \text{mm}^3)$ were situated in the slots of the sample holder on a regular pitch of 1.6 cm. At the end of the holder a thick graphite layer (2 mm thick) was placed in order to fully stop the beam and record the delivered proton current. A remotely movable PLA matrix, acting as $\beta^+$ converter, completed the experimental assembly.

The experiment was carried out in two separated stages defined as beam-on and beam-off phases. During the first stage, the $\beta^+$ converter is removed from the proton beam path, thus irradiating with protons all the thin layers under study. During the beam-off phase the PLA converter was inserted into the samples holder, thereby filling up the gap between thin layers and reducing the mean free path of the $e^+$ particles from the nuclear decays. Several configurations using different materials were investigated, however, more details will be reported in a dedicated article, which includes also a more detailed analysis and results [37]. For the sake of space limitation, here we report only on the application of the ML classifier (Sec. II) to one of the measured configurations, where PMMA was used for the targets in the sample holder. Regarding Compton imaging, owing also to extension limitation, only the results for i-TED.B will be shown although similar results were obtained for the i-TED.A module.

In summary, the goals of the ML classifier for this application were intended to:

- Study the capability of the ML classifier for rejecting non full-energy events with real data and experimental setup.
- Experimentally validate the results obtained in Sec. III by means of MC simulations.
- Identify weak aspects and elucidate further developments and improvements.

Fig. 6 displays the add-back deposited-energy spectra registered with i-TED.B during the configuration studied in this work. The red line shows the result for the unfiltered data, i.e. before the ML classifier is applied, while the blue line displays the spectrum obtained after the classification is performed.

The unfiltered spectrum shows a clear $\gamma$-ray transition at 4.4 MeV corresponding to the thick graphite layer, where the $^{12}\text{C}(p,p')^{12}\text{C}$ reactions dominate. The first and second escape peaks from the 4.4 MeV $\gamma$-rays can be also appreciated in the same spectrum. On top of the Compton continuum a further $\gamma$-ray transition at 720 keV can be observed, which corresponds to the $^{10}\text{C}$ decay. Additionally, 511 keV $\gamma$-rays are clearly
identified from the production of unstable $\beta^+$ nuclei in the different materials along the proton irradiations.

After the ML-based classification is applied the first and second escape peaks are strongly suppressed, as it is demonstrated by the blue-spectrum in Fig. 6. Furthermore, the large Compton continuum from the same 4.4 MeV $\gamma$-ray line is strongly reduced, as it can be appreciated in the region from 1.5 to 3.8 MeV. The peaks at 720 keV and 511 keV become also very prominent in the ML-filtered (blue) spectrum.

In summary, the large fraction of events that would distort the reconstructed image have been significantly suppressed. As a drawback, we have identified discontinuities in the shape of the filtered spectra corresponding with some discontinuities in the performance of the classifiers as a function of the add-back deposited energy. The latter aspect will be improved for the final analysis of the data.

The 2D-Compton images reconstructed from the unfiltered and filtered data of i-TED.B are displayed in the top and bottom panels of Fig. 7, respectively. The images were reconstructed using the methodology described in Sec. II and normalized to the maximum for comparison purposes. The beam direction is from right to left in the image x-axis, being the position of the thick graphite layer in the left (negative) side of the 2D histogram. The position of the thick graphite layer is clearly observed in the latter 2D images owing to the very intense 4.4 MeV $\gamma$-ray yield. Remarkably, the peak of the image from unfiltered events is broader and slightly shifted towards the negative part of the image compared to the image reconstructed from ML-filtered events. A further difference can be appreciated in the central part of the image: While there is not clear structure in the unfiltered image, the ML-treated image shows an structure compatible with the radiation emission from the other thin layers present during the irradiation. A clear visualization of the latter seems to be hindered by the much larger emission yield from the 2 mm thick graphite layer, when compared to that of the 0.8 mm thick PMMA layers.

These differences become more apparent when the x-axis projection is made, as displayed in Fig. 8. The projections from unfiltered and filtered data are displayed by the red and blue lines, respectively. The shift, broadening and central part structure are compatible with the conclusions drawn from the MC calculations dicussed in Sec.III.

V. CONCLUSIONS AND OUTLOOK

Enhancing the accuracy of ion-range monitoring may help to improve hadron therapy cancer treatments by reducing the safety margins, enlarging the potential benefits and extending its applicability to other types of diseases.

Proton range verification can be made via PG monitoring, i.e. detecting and imaging several $\gamma$-ray prompt transitions emitted along the nuclear reactions that take place, particularly at the Bragg peak. For a two plane Compton Camera, as the i-TED imager presented here, one of the critical aspects is the identification of full energy events, defined as the $\gamma$-ray events depositing all their energy in both detection planes.

In order to identify these events a Machine Learning algorithm, and specifically a deep neural network classifier was implemented in TENSORFLOW. It consisted in a stack of 4 layers with 256 units each using RELU as activation function.
function. The output of the classifier is given by one unit using SIGMOID as activation function. The model was compiled using BINARY CROSS-ENTROPY. The training of the model was made based on GEANT4 Monte Carlo simulations using an air $10 \times 10 \times 20 \text{ cm}^3$ phantom that covered the detectors field of view. In order to obtain the best performance in terms of accuracy, the training of the classifier was made in 0.5 MeV energy intervals, from 0.2 up to 7 MeV, thereby covering all the important $\gamma$-ray lines in PG. For validating the trained neural network additional GEANT4 MC simulations were performed replacing the air material by PMMA and by water.

The gain factor displayed in Fig. 3 after applying the ML classifier spans from a maximum value of 1.9 at $1.8 \text{ MeV}$, which decreases down to a gain-factor of 1.35 at 7 MeV. These results suggest that a two-plane Compton camera in combination with specific ML algorithms can provide a good performance both in terms of image reconstruction capability and efficiency.

Empowered by the good performance of the classifier we have applied it to a pilot experiment performed at CNA, where two i-TED imagers were utilized. Preliminary results shown in Fig. 6 demonstrate a strong suppression of the first and second escape peaks of the 4.4 MeV $\gamma$-ray line, and of the Compton continuum as well. Additionally, full-energy events of the 511 keV and 720 keV $\gamma$-ray lines corresponding to the decay of $\beta^+$ particles and excited state of $^{13}$C are significantly improved. At this moment, still some discontinuities are observed because of the training methodology used. This aspect will be improved in the final analysis of the data.

The results obtained for the reconstructed Compton images of the experimental data (Fig. 7 and Fig. 8) show the improvement obtained after applying the Machine Learning classifier.

Finally, additional experiments using 4 i-TED modules have been performed recently at the Heidelberg Ion Beam Therapy Center (HIT) where this methodology will be applied and further developed.

ACKNOWLEDGMENT

We thank our colleagues from the Electronics Unit and the Mechanics Unit at IFIC for their professional work and support.

REFERENCES


