



HAL
open science

Extended sources reconstructions by means of coded mask aperture systems and deep learning algorithm

G. Daniel, O. Limousin

► **To cite this version:**

G. Daniel, O. Limousin. Extended sources reconstructions by means of coded mask aperture systems and deep learning algorithm. Nuclear Inst. and Methods in Physics Research, A, 2021, 1012, 10.1016/j.nima.2021.165600 . insu-03745696

HAL Id: insu-03745696

<https://insu.hal.science/insu-03745696>

Submitted on 2 Aug 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

1 Extended sources reconstructions by means of coded 2 mask aperture systems and deep learning algorithm

3 G. Daniel¹, O. Limousin²

4 ¹ Université Paris-Saclay, CEA, Service de Thermo-hydraulique et de Mécanique des Fluides, 91191, Gif-sur-Yvette, France

5 ² IRFU, CEA, Université Paris-Saclay, Gif-sur-Yvette, France and Université Paris Diderot, AIM, Sorbonne Paris Cité, CEA,
6 CNRS, Gif-sur-Yvette, France

7

8 *Abstract*— Diagnostics and monitoring of radiological scenes are critical to the field of
9 nuclear safety and here, the localization of radioactive hotspots is mandatory and remains a
10 critical challenge. In order to perform gamma-ray imaging, one main method relies on
11 indirect imaging by means of coded mask aperture associated with a position sensitive
12 gamma-ray detector and a dedicated deconvolution algorithm. However, the deconvolution
13 problem is non-injective, which implies limitations of the reconstruction performance,
14 especially for spatially extended radioactive sources with respect to the angular resolution.
15 In this paper, we present and evaluate a new method based on a deep learning algorithm
16 with a convolutional neural network to overcome this limitation, in comparison with a
17 classical iterative algorithm. Our deep learning algorithm is trained on simulated data of
18 extended sources that may imply an intrinsic regularization of the neural network. We test it

19 on real data acquired with a gamma camera system based on Caliste, a CdTe detector for
20 high-energy photons.

21 *Index Terms*— CdTe, coded mask aperture, gamma imaging, deep learning, machine
22 learning, convolutional neural networks.

23 1 Introduction

24 Nuclear instrumentation is vital to nuclear safety and security, in order to provide accurate
25 information on unknown radiological environments, such as post-accident situations or in
26 the context of decommissioning and dismantling. In particular after the Fukushima accident
27 in 2011, a new generation of gamma cameras was developed, whose objective is to provide
28 a gamma-ray image of radiological scenes and to localize radioactive hotspots.

29 Due to the physical properties of gamma-ray photons, direct imaging using lenses or mirrors
30 cannot be used in compact systems such as gamma cameras. Moreover, direct imaging leads
31 to narrow fields of view. Consequently, indirect imaging is performed and two main
32 methods exist. The first one relies on the Compton scattering of high-energy photons [1],
33 [2], [3], [4], which is interesting for its extremely large field of view from 2π to 4π
34 steradians and can typically be used for photons with energies above 250 keV. The second
35 method is based on coded mask aperture imaging [5], [6], [7], which consists of setting a
36 coded mask aperture in front of a position sensitive gamma-ray detector and reconstructing
37 the position of the radioactive sources by means of the deconvolution of the projected image
38 onto the detector, called the “shadowgram”. This method can be used for photons with
39 sufficiently low energy to be absorbed by the mask, which depends on the mask thickness.

40 This method provides a typical angular resolution from 2° to 5° in existing systems. The
41 angular resolution is in fact limited only by the geometry of the system, i.e. the relative size
42 of elementary pixels and mask patterns and the distance between the mask and the detector.
43 Different mask patterns have been designed in the literature, such as URA (Uniformly
44 Redundant Array) and MURA (Modified Uniformly Redundant Array) [8] or random masks
45 [9]. URA and MURA masks are advantageous for background noise subtraction, while they
46 are likely to reconstruct false sources, also called ghosts. On the other hand, random masks
47 reconstruct fewer ghost artefacts but are more sensitive to the background noise.

48 However, coded mask aperture imaging suffers from limitations when the angular size of
49 the sources is large compared to the angular size of the elements of the mask projected on
50 the detector. Such sources are named extended sources in the following. This limitation is
51 the consequence of the deconvolution problem, which is a non-injective inverse problem.
52 The classical algorithms, such as MLEM (Maximum Likelihood Expectation Maximization)
53 [10], cannot overcome this issue and new methods must be explored.

54 Moreover, the coded mask aperture method cannot be used when the energy of the emitted
55 photons is too high to interact by photoelectric effect in the mask. This can be partially
56 solved by increasing the thickness of the mask, but a thick mask suffers from a strong
57 vignetting effect. Eventually, sources out of the mask field of view contribute to the
58 background noise and affect the reconstruction quality, thus the sensitivity. To solve the
59 latter problem, shielding can be added to the camera, but this can make it heavy and non-
60 portable.

61

62 In recent years, deep learning algorithms based on neural networks have been used in many
63 fields such as image or speech recognition and have outperformed classical methods. These
64 algorithms have been applied in gamma-ray imaging with coded mask aperture systems
65 [11], [12] on simulated data of gamma-ray point sources with a CZT detector leading to
66 promising results. The major interest of algorithms based on neural networks lies on the one
67 hand, in their speed of execution for real-time applications once they are properly trained,
68 and on the other hand, in their potential to overcome the problem of extended sources
69 reconstruction.

70 In this paper, we demonstrate the application of a Convolutional Neural Network (CNN)
71 dedicated to the image reconstruction of extended radioactive gamma-ray sources. We train
72 our CNN and compare its performance with MLEM, a reference algorithm for the
73 deconvolution problem of coded mask aperture imaging. We evaluate the results using
74 simulated data and we also perform tests on real data acquired with a gamma camera
75 prototype equipped with a coded mask and Caliste-HD, a fine pitch CdTe imaging
76 spectrometer [13]. We conclude in favour of this algorithm, which outperforms MLEM in
77 terms of reconstruction performances and computation time.

78

79 The paper is structured as follows:

80

81 Section 2 focuses on the detection system properties. Section 3 presents the mathematical
82 formulation of the problem and the deconvolution limitations with MLEM. We report the
83 architecture of our CNN and our training method in section 4. Finally, Section 5 presents
84 the evaluation of the performances of our algorithm and shows test results on real data.

85 2 A Gamma Camera prototype based on Caliste-HD associated with a Coded Mask 86 Aperture

87 2.1 *Caliste-HD: A CdTe-based detector for high-energy photons*

88 Caliste-HD [13] is a miniature pixelated detector devoted to high-energy photon imaging
89 spectroscopy. The sensor consists of a 1 mm thick Schottky CdTe crystal. The pixelated
90 electrode pattern is made of 16×16 pixels with $625 \mu\text{m}$ pitch ($525 \mu\text{m}$ pixel and $100 \mu\text{m}$
91 gap). The total detection area is 1 cm^2 . Each pixel is a tiny independent spectrometer. The
92 pixel pattern is used to record the shadowgram of our coded aperture. The pixel array is
93 associated with a stack of ultra-low noise readout electronics, the IDeF-X HD ASICs [14].
94 The whole detector is assembled into a 3D module using 3D PLUS technology [13], [15].

95
96 Thanks to its unique design, Caliste-HD reaches a spectrometric resolution of 670 eV
97 FWHM (Full Width at Half Maximum) at 60 keV and 4.7 keV FWHM at 662 keV . Its
98 dynamic range extends from 2 keV to 1 MeV . Energy resolution is important in coded mask
99 aperture imaging when spectral separation of different isotopes is required in the
100 reconstruction.

102 The characteristics of Caliste-HD are summarized in Table 1

103 TABLE I

104 CHARACTERISTICS OF CALISTE-HD

PARAMETER	CALISTE-HD
Number of pixels	256 (16×16)
Pixel Pitch	625 μm
Crystal thickness	1 mm
Dimension without CdTe	10×10×16.5 mm ³
Power Consumption	200 mW
Energy Range	2 - 1000 keV
FWHM at 60 keV	670 eV
FWHM at 662 keV	4.7 keV

105

106 The optimal spectral and noise performance of the detector is obtained at a temperature of -
107 10°C, which is reached by using an embedded compact cooling system with Peltier
108 modules.

109 This miniature detector and compact cooling system are the basic components of a portable
110 gamma camera prototype with a weight of 1 kg, presented in

111 Fig. 1. This gamma camera is dedicated to nuclear safety applications such as radiological
112 environment analysis and is planned to be developed in an industrial version by 3D PLUS
113 company.

114 *2.2 The Coded Mask Aperture geometry*

115 Our gamma camera is equipped with a coded mask aperture in order to perform radioactive
116 source localization. The geometry of the mask is given in Fig. 2. Opaque elements are made
117 of Tantalum while transparent elements are simply holes. A point source placed at infinity
118 casts a shadow of the mask pattern on the detector plane, as the photons are stopped by
119 opaque elements while they move directly to the plane through transparent elements.

120 The high atomic number ($Z = 73$) and density (16.4) of Tantalum are of interest in order to
121 absorb high-energy photons. The thickness of the mask is adaptive from 0.3 mm to 3 mm,
122 by stacking up to ten 0.3-mm thick unit layers. This flexibility is an advantage to manage
123 the vignetting effect. In fact, thick masks are efficient at high energy but the effective size
124 of the transparent elements becomes smaller and smaller as the sources move away from the
125 optical axis. In the following and for the sake of simplicity, the total thickness of our mask

126 is 0.9 mm, i.e. three layers. In this configuration, the total probability of interaction in the
127 opaque elements of the mask at the 60 keV line of ^{241}Am is 99%.

128 Our mask is an Optimized Random Aperture mask composed of 46×46 elements, with a
129 unit size of 1×1 mm². Still, the randomness is constrained by mechanical considerations,
130 such as the fact that an opaque element cannot be surrounded only by transparent elements.
131 Thus, the mask is self-supporting. Its open fraction is 40%. This is a miniaturised version of
132 the mask initially developed for the purpose of the ECLAIRs spaceborne instrument
133 devoted to Gamma-Ray Burst detection [9].

134 The mask is set at 5.8 cm from the detector. From geometrical considerations, the Fully
135 Coded Field of View is consequently 34.5° and the total Field of View is up to 51.5° . In this
136 work, we use the total Field of View. From simple geometrical considerations of the
137 projections of the mask, according to the mask element unit size and the detector pixel
138 pitch, for a source at infinity from the gamma camera, the reconstructed image, called the
139 “sky”, is discretised with 75×75 positions, with a resolution of 0.7° .

140 3 Formulation of the reconstruction problem, Classical algorithm and Limitations

141 *3.1 Reconstruction problem*

142 In order to pose the problem mathematically, we consider that we register a hit map Y ,
143 where Y_i represents the number of photons detected in each detector pixel i . Y is a vector of
144 size 256 (16×16). We want to reconstruct the position on the sky, from which the recorded

145 photon was emitted. For this purpose, we designate the sky intensity map by X , where X_j is
146 the number of photons emitted by position j . X is a vector of size 5625 (75×75).

147 Finally, we introduce M , the mask matrix, where M_{ij} contains the expected projection of the
148 mask on the detector pixel i for a source at position j . From a probabilistic point of view, Y
149 follows a Poisson probability law as described in Equation (1).

$$Y \sim \text{Poisson}(M \cdot X + \varepsilon) \quad (1)$$

150 where ε is a perturbation term, which takes into account the background counts, false events
151 due to electronic noise and a potential dead pixel map. Note that a first difficulty arises from
152 the fact that the components of the term ε depend on measurement conditions that are
153 generally unknown.

154 On the other hand, the probabilistic aspect of the problem creates inherent noise in the
155 reconstruction due to the Poissonian counting fluctuations, especially for low photon count
156 statistics.

157 Eventually, the reconstruction problem is not injective. Indeed, the dimension of the
158 reconstructed sky X is larger than the dimension of the measurement Y . This leads to the
159 limits of the reconstruction of extended sources that we will show in paragraph 3.3.

160 3.2 *Maximum Likelihood Expectation Maximisation (MLEM)*

161 A classical way to solve this ill-posed inverse problem in a probabilistic paradigm is to use
162 an algorithm that maximizes the probability $p(Y|X)$ of observing the data Y given the

163 reconstructed image X , also known as the likelihood. It is equivalent to solving the problem
164 given in Equation (2).

$$\hat{X} = \operatorname{argmax}_X(p(Y|X)) \quad (2)$$

165 MLEM is a method initially developed in the framework of tomographic reconstruction [10]
166 and provides an iterative algorithm when the data follows a Poisson law, in order to find the
167 optimal value \hat{X} . For each iteration n , Equation (3) gives the update formula.

$$\hat{X}_j^{n+1} = \hat{X}_j^n \sum_i \frac{p(i|j)Y_i}{\sum_{j'} p(i|j')\hat{X}_{j'}^n} \quad (3)$$

168 $p(i|j)$ indicates the probability that a detected photon emitted at position j will be detected
169 in the detector's pixel i . Actually, $p(i|j)$ is given by the mask matrix element M_{ij} .

170 For initialisation, \hat{X}^0 can be chosen as a uniform strictly positive image. The algorithm
171 ensures the conservation of the number of photons N_{phot} given in Equation (4), for each
172 step $n > 0$.

$$\sum_j \hat{X}_j^n = \sum_i Y_i = N_{\text{phot}} \quad (4)$$

173 This algorithm has a low sensitivity to the unknown perturbations ε in Equation (1), such as
174 dead pixels or background uniformity. However, there is no clear stopping criterion to
175 determine when the algorithm reaches the convergence, which can lead to the overfitting of
176 data.

177 Finally, we will show in the paragraph 3.3 that MLEM algorithm does not overcome the
178 limitations identified for the reconstruction of extended sources.

179 *3.3 Limitations for extended sources*

180 The deconvolution problem (1) is non-injective, which means that one observation Y can be
181 associated to several sky maps X . This phenomenon is illustrated in Appendix 7.

182 Fig. 3 shows a simulation of an extended square source with 10^8 detected photons without
183 background in the simulation, and the corresponding image reconstruction with MLEM
184 algorithm. The simulation is a direct application of the mask projection, combined with a
185 random draw of the number of photons detected according to a Poisson distribution, in each
186 pixel. In this simulation, the photons are fully absorbed by the opaque elements of the mask.
187 Although this is an ideal case, the MLEM algorithm is not able to locate or reconstruct the
188 simulated source.

189 Since MLEM is based only on the likelihood of the data, it does not embed any
190 regularization property. A regularization term properly designed could favour solutions with
191 compact sources. Regularization can be done by adopting a Bayesian paradigm through a
192 MAPEM (Maximum A Posterior Expectation Maximization) algorithm [16], which consists
193 of maximizing the a posteriori probability $p(X|Y)$, proportional to $p(X).p(Y|X)$. The prior
194 $p(X)$ is then seen as the regularisation term. On the one hand, the difficulty arises from the
195 mathematical formulation of this prior in order to favour compact sources. On the other
196 hand, the resolution then requires the use of another algorithm because the algorithm given

197 in Equation (3) cannot be applied. To the best of our knowledge, there is no study in the
198 literature, which introduces such a method to reconstruct extended sources with coded mask
199 aperture imaging.

200 In Section 4, we propose a method using Convolutional Neural Networks to address this
201 problem.

202 4 Use of a Convolutional Neural Network for image reconstruction

203 4.1 *Architecture of the CNN*

204 The performance of deep learning algorithms in the domain of image processing, and
205 recently for coded mask [11] [12] aperture imaging of gamma-ray point sources, show that
206 the study of these methods is relevant to the problem of extended sources.

207 We propose a Convolutional Network Architecture in order to test the reconstruction of
208 extended sources by means of deep learning algorithm. The architecture is presented in Fig.
209 4.

210 The rationale for the use of convolution layers is given by the fact that the input image, the
211 hit map, is expected to be locally correlated through the projection of the mask patterns. The
212 number of layers has been tuned in order to ensure the learning of the algorithm. Our first
213 tests on shallower networks show that numerous convolution layers and fully-connected
214 layers are needed, otherwise the learning is blocked and does not converge. We do not apply
215 pooling operations since the input size is small.

216 We use dropout as a regularisation technique. We leave the dropout active during testing to

217 approximate a Bayesian neural network by means of Monte-Carlo (MC) dropout method
218 [17]. This method has been developed to provide an uncertainty estimate on the result, but
219 our preliminary tests have shown that it gives promising results on our problem. At test
220 time, we run the neural network on the same example several times, typically 100 times,
221 and we determine the median of the results, which we consider to be the final answer of the
222 network.

223 We point out that many relevant architectures can be tested alternatively. Performance of
224 uncertainty quantification by MC Dropout can be evaluated, but these studies are out the
225 scope of this paper.

226 *4.2 Online training with synthetic data*

227 Training data are required to use machine learning algorithms. We chose to create simulated
228 data since simulations allows us to control exactly the shape, the position and the intensity
229 of the sources.

230 To create extended sources with different shapes, we use the *random_shapes* function of the
231 *scikit-image* [18] Python package. This allows us to create random elementary shapes:
232 circles, rectangles and triangles. We also apply random rotations and resizing's of the
233 shapes for each training example. The intensity of the emission is considered to be uniform
234 over the shape. This shape X is the output that the neural network must predict. The
235 simulated image is normalized in norm 1: its elements sum up to 1.

236 Given the shapes, we compute the corresponding projections Y_p on the detector. For each

237 example in the training set, we draw the number of simulated photons N_{phot} between 10 and
238 10^7 photons, logarithmically following equation (5) in order not to penalize low counting
239 statistics.

$$N_{\text{phot}} = 10^{7x+1}, x \sim \text{Unif}([0,1]) \quad (5)$$

240 We then draw a hit map with N_{phot} photons corresponding to the computed projection Y_p .
241 We add a uniform noise on the detector with an intensity between 0 and the intensity of the
242 simulated source. We also randomly disable up to two pixels of the detector to anticipate a
243 partially malfunctioning pixel array and take real experimental conditions into account in
244 the training. The resulting hit map Y is normalized by its maximum value and corresponds
245 to the input of the neural network. The normalization must be applied for each example
246 presented to the network.

247 One training example is illustrated in

248 Fig. 5, which shows the simulated shapes and the corresponding simulated hit map on the
249 detector.

250 The simulation is fast since it takes about 8 ms to create one example, regardless of the
251 number of simulated photons, which is an advantage over real measurements.

252 A more realistic Monte-Carlo simulation, such as a Geant4 [19] code, can be considered to
253 take into account the energy of the photons and the detailed interactions in the surrounding
254 environment. However, the computation time will be much longer and will depend on the
255 number of simulated photons. In Section 5, we show that such a detailed simulation is not

256 necessary at low energy, when opaque elements of the mask are fully efficient in stopping
257 the photons. At higher energy, Compton scattering becomes the dominant interaction
258 process and a complete Monte-Carlo simulation is necessary, which is beyond the scope of
259 this paper.

260 As the number of possible configurations is infinite (shapes, positions, intensities of
261 sources, number of photons, random noise...), we decided not to create a fixed pre-
262 generated database with a limited number of examples, but we rather use “online” learning,
263 meaning that we simulate examples at the learning phase. This has the advantage of
264 showing a large number of different examples to the network and avoids memory storage
265 issues.

266 During the learning phase, at each iteration, we generate 1000 examples, divided into 900
267 examples for the training set and 100 examples for the validation set, to check the training.
268 Then we compute a learning step of one epoch out of 900 training examples. We repeat this
269 process for 200 000 iterations, so that the network learns with approximately 200 million
270 examples.

271 The loss function used for training is the binary cross-entropy function and the optimizer is
272 the ADAM algorithm [20] with a learning rate of 10^{-4} and moments $\beta_1 = 0.9$ and $\beta_2 =$
273 0.999 . The training is made on NVIDIA GeForce RTX 2080Ti GPU. We use *Keras* [21]
274 Library for implementation in Python. In this configuration, our learning phase lasts one
275 week.

276 5 Results

277 5.1 Performance evaluation on simulated data

278 5.1.1 Performance assessment in an ideal case

279 Before confronting our CNN network with real data, we quantitatively evaluate its
280 performance on the simulated database and compared results with the MLEM approach to
281 solve the coded mask aperture image reconstruction problem. We evaluate the results
282 relatively to the size of the extended sources. For that purpose, we built a test set with
283 square shapes of different sizes ranging from 0.7° (i.e. 1 pixel on the sky) up to 20.6° (30
284 pixels on the sky), with 1000 examples for each size. Note that the largest source nearly
285 corresponds to 40% of the total field of view. In this section, the simulated hit map is
286 calculated in ideal conditions: no background, no disabled pixels, large statistics with 10
287 million photons simulated. Although these conditions are not realistic, applying the
288 deconvolution with both CNN and MLEM reconstruction techniques allows us to easily
289 intercompare both method results. For illustration, one example is shown in Fig. 6, which
290 represents a simulated square shaped source, the corresponding simulated hit map and the
291 reconstructed images obtained with our CNN and MLEM.

292 To quantify the performance, we use the standard metric Intersection over Union (IoU),
293 which is given by Equation (6).

$$\text{IoU} = \frac{\text{Area}(\text{True} \cap \text{Recons})}{\text{Area}(\text{True} \cup \text{Recons})} \quad (6)$$

294 Area($\text{True} \cap \text{Recons}$) is the area covered by both the reconstructed positions and the
295 expected simulated positions. Area($\text{True} \cup \text{Recons}$) is the area covered by at least the
296 reconstructed positions or the expected simulated positions. A perfect reconstruction would
297 give an IoU of 1 while a value of 0 would indicate a total absence of intersection between
298 the reconstructed image and the real source position.

299 The reconstructed positions are selected for each algorithm as the positions with an intensity
300 above a threshold. If \hat{X} is the reconstructed image, we choose this threshold to be 20% of
301 $\max(\hat{X})$. This threshold is set empirically and can be optimized. The choice of this
302 threshold does not favour one method over the other.

303 Fig. 7 shows the results for the CNN and the MLEM algorithm.

304 When the size of a source corresponds to one (point sources) or two pixels on the sky, the
305 image reconstruction is perfect for both algorithms.

306 MLEM performance drops drastically as the sources get more extended. IoU is found to be
307 lower than 0.1 for 5° extended sources, while the performance of the CNN decreases much
308 more slowly and IoU remains higher than 0.8.

309 After IoU for MLEM reaches a minimum, it begins to increase slowly with the source size
310 when it exceeds 8° of angular size. This effect is expected and due to the fact that the larger
311 the source is, the more likely we are to randomly reconstruct positions within the expected
312 source region. The red curve in Fig. 7 corresponds to the fraction of the field of view
313 covered by the simulated and known area of the source. We notice that the IoU of MLEM

314 follows the trend of this quadratic curve, which means that a large part of the IoU is
315 attributed to randomness.

316 This quantification study shows that the CNN is able to reconstruct a large part of the
317 sources at the right position in the sky, even if they are extended in angular area up to a fifth
318 of the Field of View, while MLEM is not able to perform such a reconstruction when the
319 size of the source exceeds 3° .

320 We point out that the CNN does not contain an explicit regularization term that would
321 favour the reconstruction of compact sources. However, the CNN does have regularization
322 properties due to the training database we designed, which contains only examples of
323 compact extended sources. This is a significant advantage of the CNN over the MLEM
324 algorithm. The latter would require the introduction of an explicit prior in order to transform
325 it into a MAPEM algorithm and then, the corresponding maximization algorithm must be
326 created.

327 This study is carried out considering an ideal case, with only one of the available shapes
328 included in the training set of our neural network. In paragraph 5.1.2 and 5.2, we
329 demonstrate that our CNN is successfully applied to other shapes, even if they are not
330 included in the training simulations.

331 *5.1.2 Image reconstruction in the case of a randomly generated extended source*

332 In this section, we illustrate the performance of the algorithms in non-ideal cases. Fig. 8
333 shows the reconstruction for a shape that is not included in the training simulation set. Its

334 intensity is not uniform, which is also not learned in the training simulations. With 50 000
335 photons and by adding noise of the same intensity as the source, the extended source is
336 successfully located by the CNN. The localization is correct and the shape is roughly
337 restored, while the MLEM algorithm fails to provide any information.

338 *5.2 Reconstruction test on real data acquired with a detector*

339 We test our algorithms on real data acquired with our gamma camera. Our laboratory is not
340 equipped with extended radioactive sources, but only with point sources. In order to mimic
341 an extended source, we place three point-sources on a rotating blade that rotates slowly and
342 thus mimic a disc shape after some acquisition time.

343 The sources are ^{241}Am with activities of 312, 393 and 440 kBq. The sources are placed so as
344 to mimic a disk shaped source of 6.5° in diameter. The source is nearly in the center of the
345 field of view, intentionally slightly off-axis. We set our rotating system 1 m away from the
346 detector. One can consider the source to be placed at an infinite distance for our imaging
347 system according to our geometry. The dose rate at detector level is thus 27 nSv/h, the same
348 order of magnitude as a typical level of natural background radiation.

349 For illustration purpose, we superimpose the gamma image obtained by our algorithms with
350 an optical image caught from an optical camera associated to our gamma camera. A
351 calibration process has been previously computed to realize this superposition.

352 Fig. 9 shows the result after a 2 hours and 35 minutes long acquisition. Gamma ray events
353 are selected in the energy range from 15 to 62 keV, which corresponds to the dynamic of

354 interest for ^{241}Am . Our CNN is able to locate the source and restore its shape and size as a
355 disk shape of 5° in diameter while MLEM is unable to give exploitable information. The
356 IoU is found to be 0.72 and 0.15 for CNN and MLEM respectively.

357 We make another acquisition by moving the sources away from the centre of the field of
358 view. In addition, we occult part of the source by placing a 5 cm thick triangular lead brick
359 in front of the sources, so that the shape is no longer a disk but a circular sector. Photons
360 with an energy lower than 60 keV are entirely absorbed by the lead. The acquisition time is
361 1 h 30 and the energy selection is again between 15 keV and 62 keV.

362 Fig. 10 shows the result of the reconstruction with both algorithms. Our CNN locates the
363 source position and its shape accurately. MLEM still fails to provide any exploitable
364 information. The IoU is found to be 0.76 and 0.15 for CNN and MLEM respectively.

365 These tests were performed with ^{241}Am sources so that the photons interacting with opaque
366 elements of the mask are totally absorbed. In future studies, other radioelements, such as
367 ^{137}Cs are considered to test and evaluate the reconstruction performance. To do this, the
368 simulation process may need to be refined using a more complex simulation code, such as
369 Geant4, as we describe in section 4.2.

370 6 Conclusions

371 The localization of extended radioactive sources with coded mask aperture imaging is a
372 challenge for algorithmic processing since the deconvolution problem is not injective. Even
373 using ideal data, with very high counting statistics and without any noise or perturbation,

374 the classical algorithm MLEM is unable to solve this problem when the source is large
375 compared to the imaging system angular resolution. A regularization technique, such as in
376 MAP-EM algorithm, is thus necessary.

377 We demonstrate that convolutional neural networks have the capacity to process this type of
378 data and to reconstruct the position of the sources and their shape. Our CNN has been
379 successfully tested on ideal simulated data as well as on noisy simulated data and,
380 importantly on real data. We can interpret their ability to reconstruct the sources through a
381 regularization process, which is intrinsic to training data since our CNN has been trained to
382 reconstruct only compact extended sources.

383 Further studies are needed to evaluate the ability of deep learning algorithms to handle
384 multiple point sources or multiple extended sources. In addition, the findings of this study
385 apply to low-energy gamma radioactivity below 100 keV. For higher energy gamma-ray
386 emission such as photons emitted by ^{137}Cs or ^{60}Co , the study and tests will be performed
387 again, including full treatment of scattering both within the detector and in the inactive
388 materials surrounding it. Non-uniform background noise, due to radiation sources outside
389 the field of view, should also be studied. It is also relevant to work on uncertainty
390 assessment and dose rate quantification with these machine learning techniques.

391 7 Appendix

392 The objective deconvolution problem given by equation (1) is to reconstruct a sky map X of
393 size 75×75 from a detector hit map Y of size 16×16 . Because of the dimensions of X

394 and Y , the problem is non-injective: several sky maps X can lead to the same observation Y .

395 This property implies limitations especially in the case of extended sources reconstruction.

396 Fig. 11 illustrates this phenomenon:

- 397 • In Fig. 11a, we simulate an extended source X with a 10° square shape and we
398 calculate the corresponding projection Y by the mask on the detector in Fig. 11b.
399 Here, we intentionally consider an ideal case as we do not introduce any
400 perturbation ε and we calculate the shadowgram without taking into account
401 Poisson's law.
- 402 • In Fig. 11c, we solve the least square problem given in Equation (7) in order to find
403 the optimal positions \hat{X} of the sources, with a positivity constraint on the result.

$$\hat{X} = \operatorname{argmin}_{X \geq 0} (\|Y - M \cdot X\|_2^2) \quad (7)$$

404 where $\|\cdot\|_2$ is the Euclidean norm.

405 The resulting reconstruction solution in Fig. 11c is clearly different from the expected
406 simulated extended source in Fig. 11a. However, when reprojecting this result through the
407 mask onto the detector, one finds the image in Fig. 11d, which is identical to the expected
408 projection, the relative difference being as low as 10^{-15} . Even in an ideal case, the expected
409 source cannot be found.

410 8 Acknowledgement

411 The authors wish to thank 3D PLUS and CEA for supporting this work within the
412 framework of the joint laboratory ALB3DO (Advanced Lab for 3D Detection Device
413 Developments).

414

415 9 References

416

- [1] S. Takeda, Y. Ichinohe, K. Hagino, H. Odaka, T. Yuasa, S. Ishikawa, T. Fukuyama, S. Saito, T. Sato, G. Sato, S. Watanabe, M. Kokubun, T. Takahashi, M. Yamaguchi, H. Tajima, T. Tanaka, K. Nakazawa, Y. Fukazawa and T. Nakano, Applications and Imaging Techniques of a Si/CdTe Compton Gamma-Ray Camera, Physics Procedia, Volume 37, Pages 859-866, 2012.
- [2] C. G. Wahl, W. R. Kaye, W. Wang, F. Zhang, J. Jaworski, A. King, Y. A. Boucher and Z. He, The Polaris-H imaging spectrometer, NIM-A, Volume 784, Pages 377-381, 2015.
- [3] L. Sinclair, P. Saull, D. Hanna, H. Seywerd, A. MacLeod and P. Boyle, Silicon Photomultiplier-Based Compton Telescope for Safety and Security (SCoTSS), IEEE Transactions on Nuclear Science, Volume: 61, Issue: 5, 2014, pp. 2745 - 2752.
- [4] P. Saull, L. Sinclair, H. Seywerd, D. Hanna, P. Boyle and A. MacLeod, First demonstration of a Compton gamma imager based on silicon photomultipliers, Nuclear Instruments and Methods in Physics Research Section A, Volume 679, Pages 89-96, 2012.
- [5] G. Montemont, P. Bohuslav, J. Dubosq, B. Feret, O. Monnet, O. Oehling, L. Skala, S. Stanchina, L. Verger and G. Werthmann, NuVISION: a Portable Multimode Gamma Camera based on HiSPECT Imaging Module, IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC), 2017.
- [6] Mirion Technologies, iPIX - Ultra Portable Gamma-Ray Imaging System (datasheet), 2019.
- [7] G. Skinner, Imaging with coded-aperture masks, NIM, Volume 221, Issue 1, Pages 33-40, 1984.

- [8] S. Gottesman and E. Fenimore, New family of binary arrays for coded aperture imaging, *Applied Optics*, Vol. 28, Issue 20, pp. 4344-4352, 1989.
- [9] S. Schanne, B. Cordier, J.-L. Atteia, O. Godet, C. Lachaud and K. Mercier, The ECLAIRs GRB-trigger telescope on-board the future mission SVOM, *Proceedings of science*, 2014.
- [10] L. Shepp and Y. Vardi, Maximum Likelihood Reconstruction for Emission Tomography, *IEEE Transactions on Medical Imaging*, Vol. 1, Issue 2, 1982.
- [11] R. Zhang, P. Gong, X. Tang, P. Wang, C. Zhou, X. Zhu, L. Gao, D. Liang and Z. Wang, Reconstruction method for gamma-ray coded-aperture imaging based on convolutional neural network, *Nuclear Inst. and Methods in Physics Research, A*, Volume 934, p. 41-51., 2019.
- [12] R. Zhang, X. Tang, P. Gong, P. Wang, C. Zhou, X. Zhu, D. Liang and Z. Wang, Low-noise reconstruction method for coded-aperture gamma camera based on multi-layer perceptron, *Nuclear Engineering and Technology*, 2020.
- [13] A. Meuris, O. Limousin, O. Gevin, F. Lugiez, I. Le Mer, F. Pinsard, M. Donati, C. Blondel, A. Michalowska, E. Delagnes, M. Vassal and F. Soufflet, "Caliste HD: a New Fine Pitch Cd(Zn)Te Imaging Spectrometer from 2 keV up to 1 MeV", *IEEE Nuclear Science Symposium Conference Record*, 2011.
- [14] A. Michalowska, O. Gevin, O. Lemaire, F. Lugiez, P. Baron, H. Grabas, F. Pinsard, O. Limousin and E. Delagnes, IDeF-X HD : A low power multi-gain CMOS ASIC for the readout of Cd(Zn)Te detectors, *IEEE Nuclear Science Symposium and Medical Imaging Conference*, NSS/MIC p. 1556-1559 5874037, 2010.
- [15] 3D PLUS, "3D PLUS," [Online]. Available: <https://www.3d-plus.com/>.
- [16] E. Levitan and G. Herman, A Maximum a Posteriori Probability Expectation Maximization Algorithm for Image Reconstruction in Emission Tomography, *IEEE Transactions on Medical Imaging*, Vol. 6, Issue 3, 1987.
- [17] Y. Gal and Z. Ghahramani, Dropout as a Bayesian Approximation: Representing Model Uncertainty in Deep Learning, *Proceedings of the 33 rd International Conference on Machine Learning*, New York, NY, USA, JMLR: W&CP volume 48, 2016.
- [18] S. Van der Walt, J. Schönberger, J. Nunez-Iglesias, F. Boulogne, J. Warner, N. Yager, E. Gouillart, T. Yu and the scikit-image contributors, scikit-image: Image processing in Python, *PeerJ* 2:e453, 2014.
- [19] S. Agostinelli, J. Allison, K. Amako, J. Apostolakis, H. Araújo, P. Arce and e. al., Geant4 - a simulation toolkit, *Nuclear Instruments and Methods in Physics Research A*, Volume 506, Issue 3, Pages 250-303, 2003.
- [20] D. Kingma and J. Ba, Adam: A Method for Stochastic Optimization, *Conference paper at the 3rd International Conference for Learning Representations*, San Diego, 2015.
- [21] F. Chollet, "Keras," Github repository, 2015. [Online]. Available: <https://github.com/fchollet/keras>.

417

418

419

420 10 List of figures

421

422

423

424

425

426

427

428

429

430

431

432

Electronics and
firmware

Camera head
containing the
Caliste detector



Optical stereoscopic camera

Coded mask aperture

433

Fig. 1

434

435

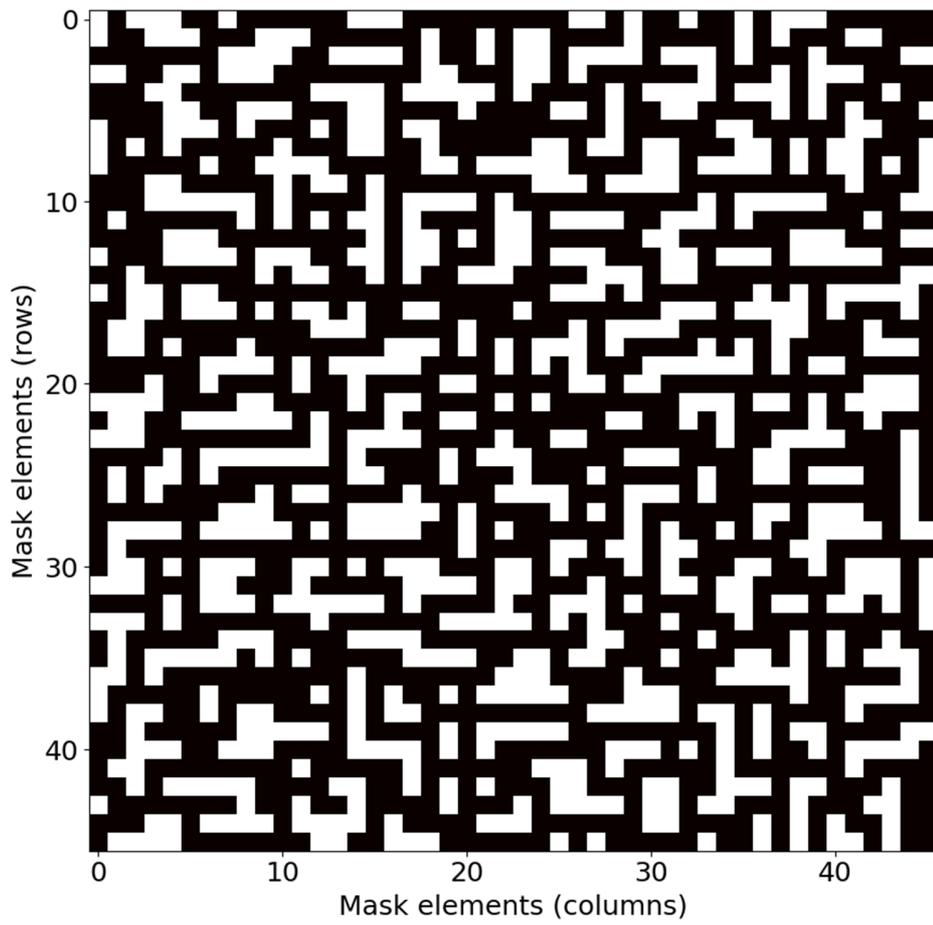


Fig. 2

436

437

438

439

440

441

442

443

444

445

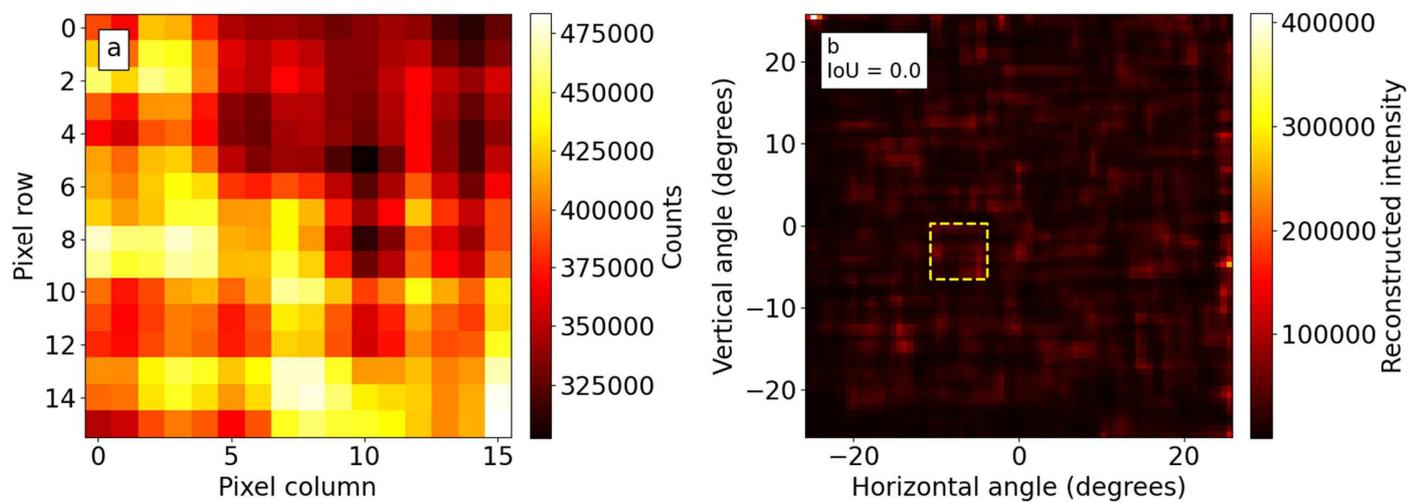
446

447

448

449

450



451

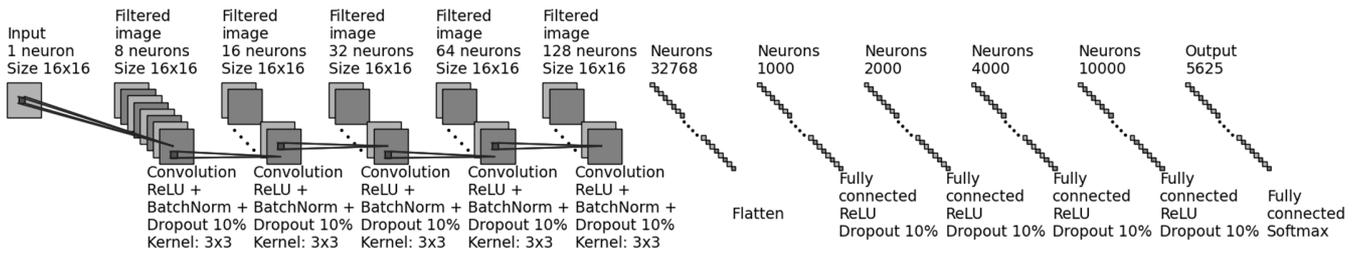
452

453

454

455

Fig. 3

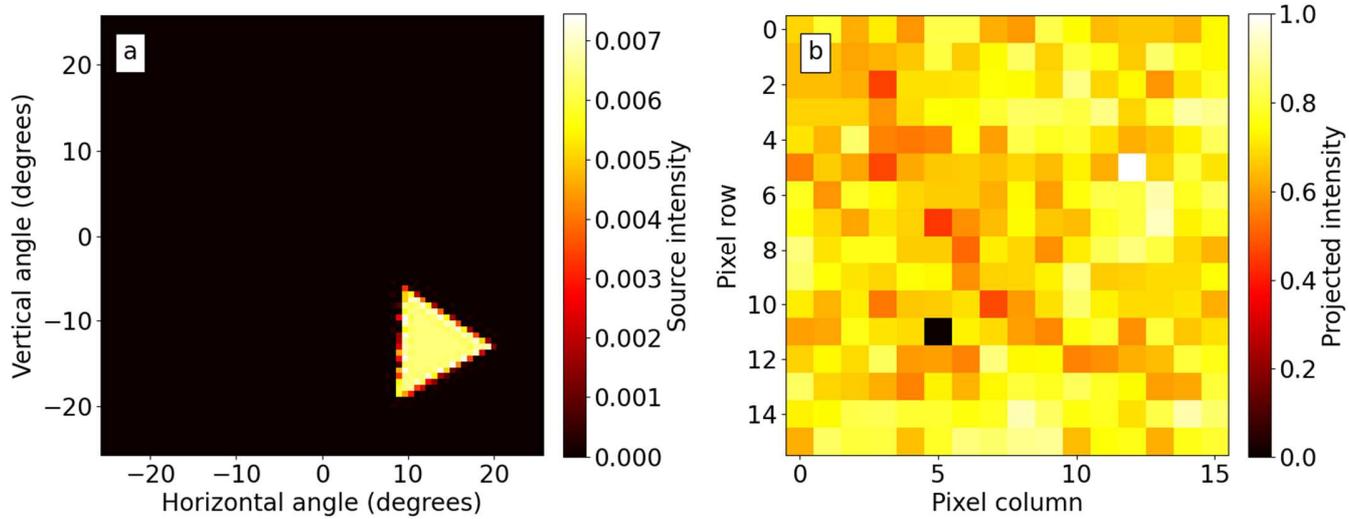


456

457

Fig. 4

458



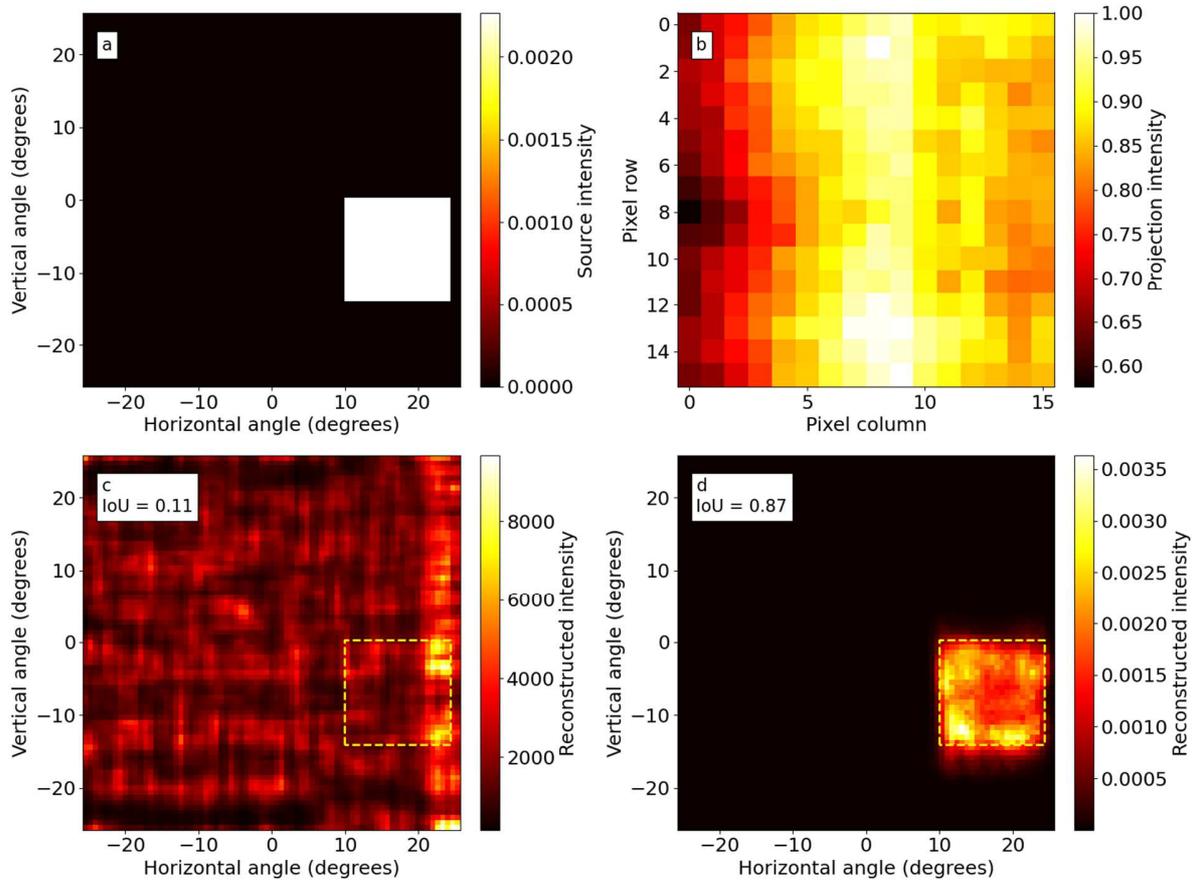
459

460

461

Fig. 5

462



463

464

Fig. 6

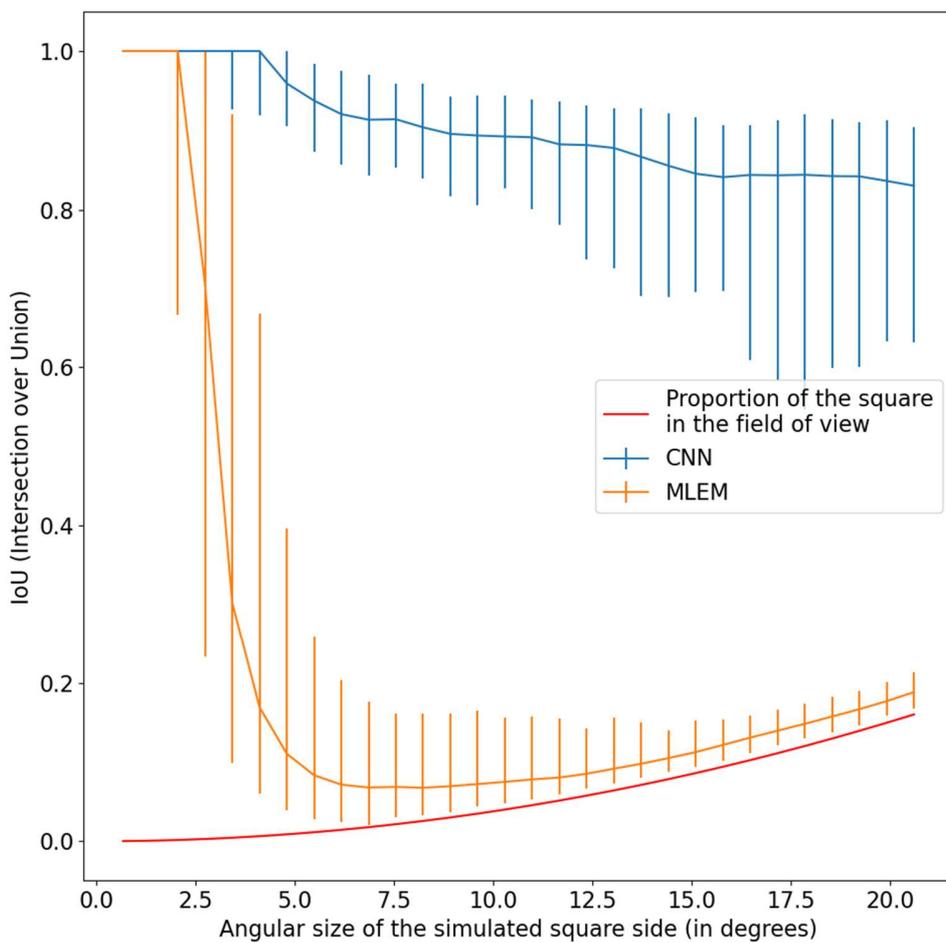


Fig. 7

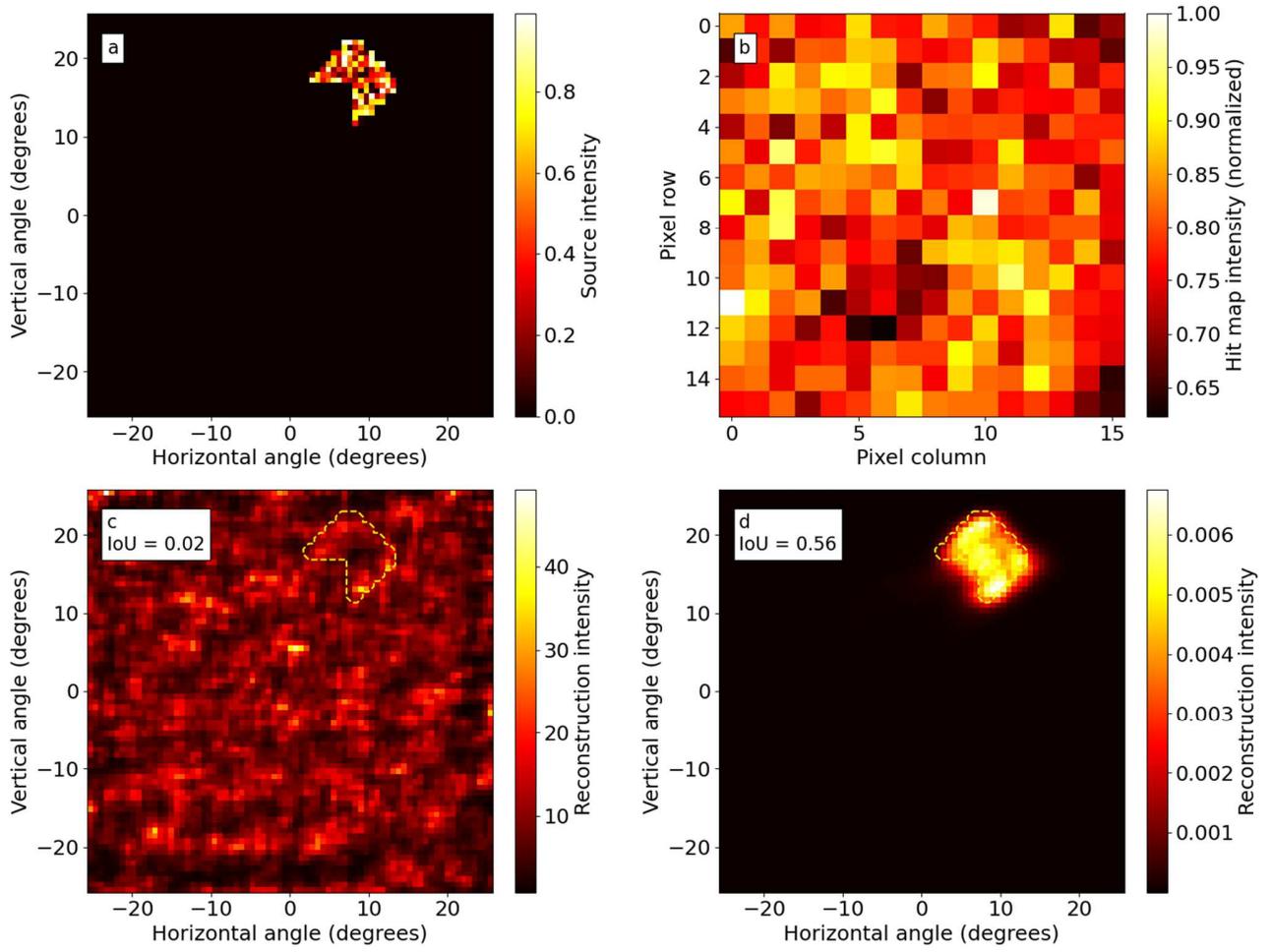
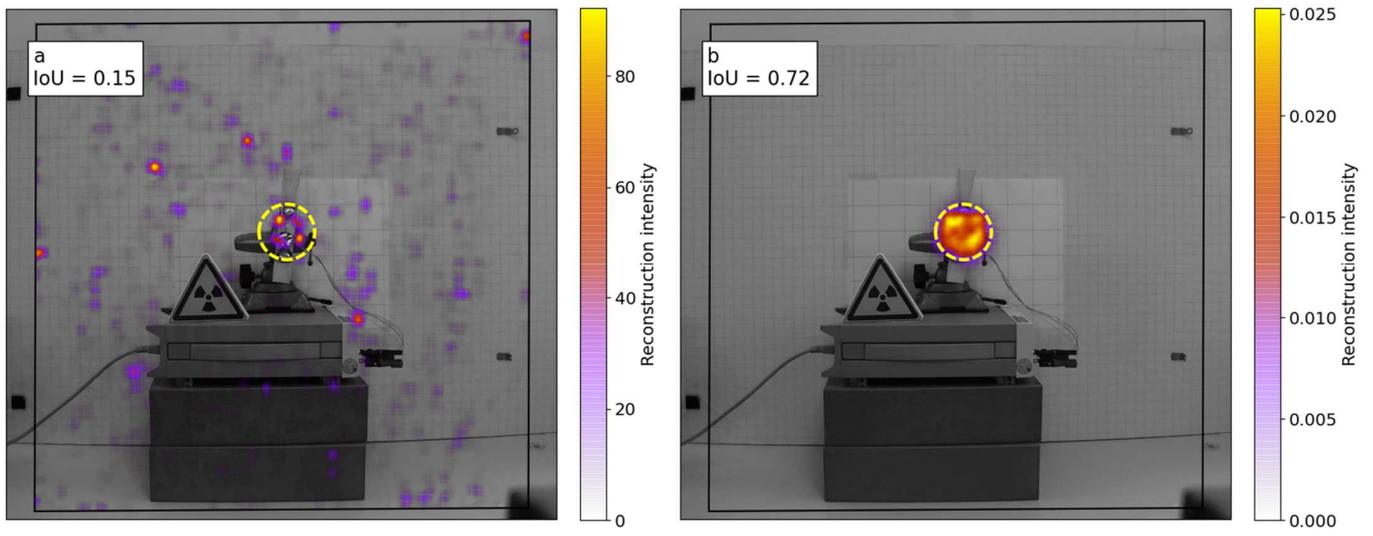


Fig. 8

471



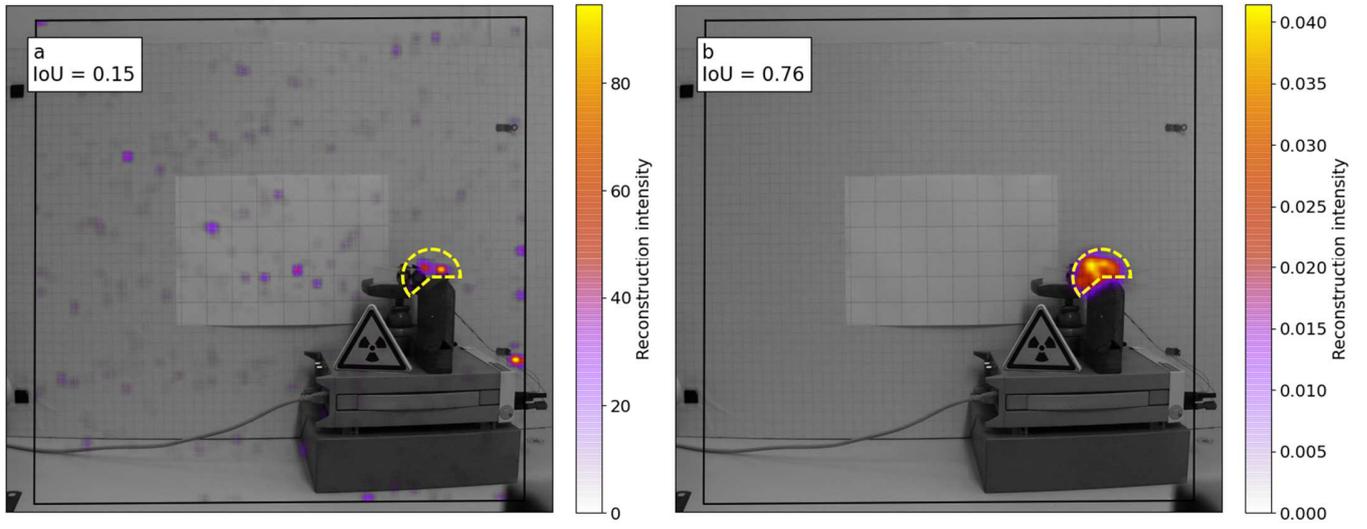
472

473

Fig. 9

474

475



476

477

Fig. 10

478

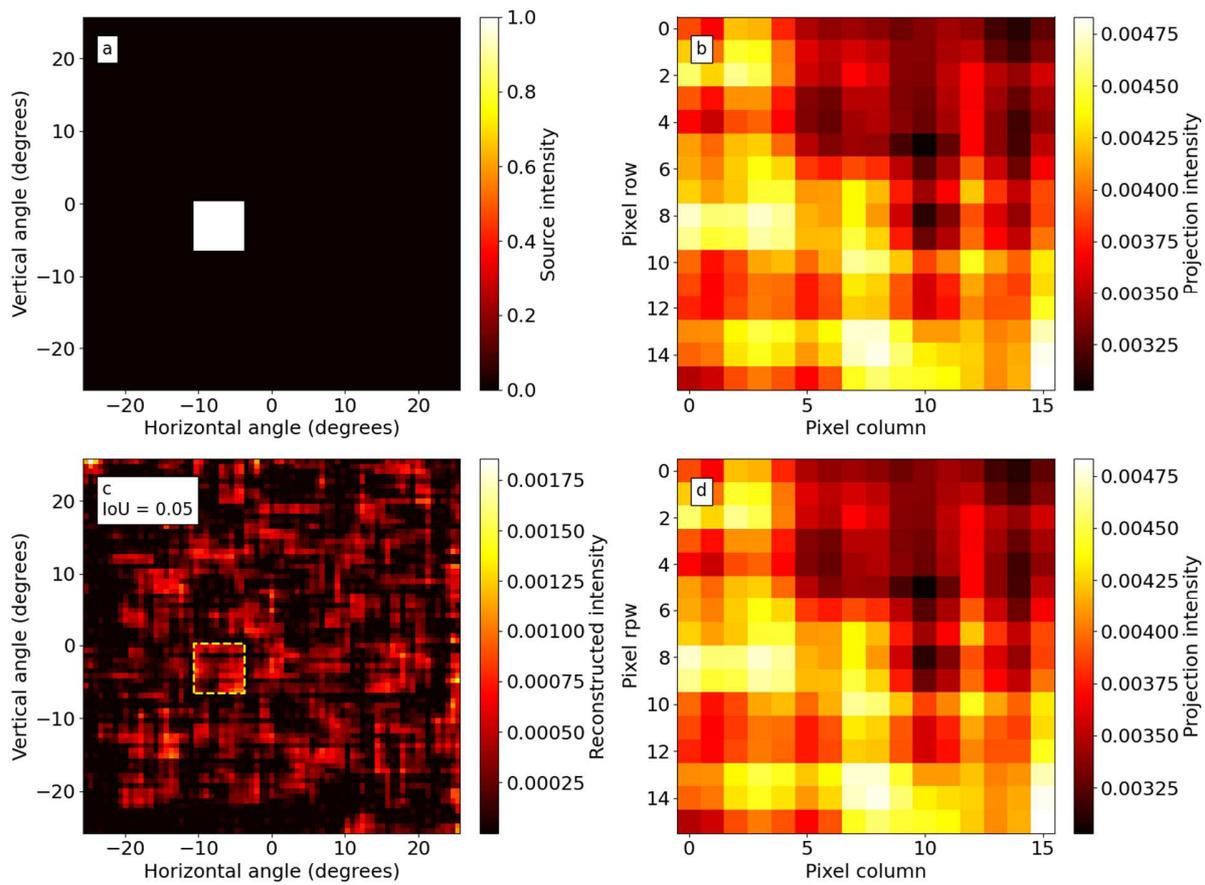


Fig. 11

479

480

481

482 11 List of figures captions

483

484 Fig. 1: Gamma camera prototype equipped with a coded mask aperture and an optical
485 stereoscopic camera

486 Fig. 2: Random coded mask aperture. The black elements are the opaque elements, the
487 white elements are the empty elements or holes.

488 Fig. 3: a. Projection of the source with 10^8 photons. The photons interacting with the
489 opaque elements of the mask are fully absorbed. b. Reconstruction with MLEM algorithm,
490 the yellow dashed square indicates the simulated source location.

491 Fig. 4: CNN architecture

492 Fig. 5: a. Simulated source with normalization by the sum – Output of the training example
493 b. Simulated hit map with normalization by the maximum – Input of the training example.
494 10 000 simulated photons, one disabled pixel.

495 Fig. 6: a. Simulated source square, 10 million photons, no noise. b. Normalized hit map. c.
496 Reconstruction MLEM. d. Reconstruction CNN. The yellow dashed squares indicate the
497 source position.

498 Fig. 7: Performances of CNN and MLEM according to the IoU metric relatively to the size
499 of the simulated square.

500 Fig. 8: a. Simulated source, random shape, 50 000 photons with noise. b. Normalized hit
501 map. c. Reconstruction MLEM. d. Reconstruction CNN. The yellow dashed lines indicate
502 the source position.

503 Fig. 9: Real test acquisition with ^{241}Am sources, 2 h 35 mn acquisition duration.

504 a. MLEM reconstruction. b. CNN reconstruction. The black square is the coded mask field
505 of view. The yellow dashed circles are the expected shape of the disk source.

506 Fig. 10: Real test acquisition with ^{241}Am sources, 1 h 30 acquisition time.

507 a. MLEM reconstruction. b. CNN reconstruction. The black square is the coded mask field
508 of view. The yellow dashed lines are the expected shape of the circular section source.

509 Fig. 11: a. Simulated squared shape extended source, 10° by side. b. Computed shadowgram
510 on the detector plane. c. Reconstruction by the least square method, the yellow dashed
511 square indicates the expected position of the simulated source (a). d. Reprojection of the
512 reconstruction solution (c) on the detector.