

## Article

# Optical to Planar X-ray Mouse Image Mapping in Preclinical Nuclear Medicine Using Conditional Adversarial Networks

Eleftherios Fysikopoulos <sup>1,2,\*</sup>, Maritina Rouchota <sup>1,2</sup>, Vasilis Eleftheriadis <sup>2</sup>, Christina-Anna Gatsiou <sup>2</sup>, Irinaios Pilatis <sup>2</sup>, Sophia Sarpaki <sup>2</sup>, George Loudos <sup>2</sup>, Spiros Kostopoulos <sup>1</sup> and Dimitrios Glotsos <sup>1</sup>

<sup>1</sup> Biomedical Engineering Department, University of West Attica, 12210 Athens, Greece; mrouchota@bioemtech.com (M.R.); skostopoulos@uniwa.gr (S.K.); dimglo@uniwa.gr (D.G.)

<sup>2</sup> BIOEMTECH, Lefkippos Attica Technology Park, N.C.S.R. Democritos, 15343 Athens, Greece; vasilis.eleftheriadis@bioemtech.com (V.E.); cgatsiou@bioemtech.com (C.-A.G.); ipilatis@bioemtech.com (I.P.); ssarpaki@bioemtech.com (S.S.); george@bioemtech.com (G.L.)

\* Correspondence: lefteris.fysikopoulos@gmail.com or lfysikop@bioemtech.com

**Abstract:** In the current work, a pix2pix conditional generative adversarial network has been evaluated as a potential solution for generating adequately accurate synthesized morphological X-ray images by translating standard photographic images of mice. Such an approach will benefit 2D functional molecular imaging techniques, such as planar radioisotope and/or fluorescence/bioluminescence imaging, by providing high-resolution information for anatomical mapping, but not for diagnosis, using conventional photographic sensors. Planar functional imaging offers an efficient alternative to biodistribution ex vivo studies and/or 3D high-end molecular imaging systems since it can be effectively used to track new tracers and study the accumulation from zero point in time post-injection. The superimposition of functional information with an artificially produced X-ray image may enhance overall image information in such systems without added complexity and cost. The network has been trained in 700 input (photography)/ground truth (X-ray) paired mouse images and evaluated using a test dataset composed of 80 photographic images and 80 ground truth X-ray images. Performance metrics such as peak signal-to-noise ratio (PSNR), structural similarity index measure (SSIM) and Fréchet inception distance (FID) were used to quantitatively evaluate the proposed approach in the acquired dataset.

**Keywords:** molecular preclinical imaging; image-to-image translation; PET; SPECT; deep learning; pix2pix; cGAN; X-ray



**Citation:** Fysikopoulos, E.; Rouchota, M.; Eleftheriadis, V.; Gatsiou, C.-A.; Pilatis, I.; Sarpaki, S.; Loudos, G.; Kostopoulos, S.; Glotsos, D. Optical to Planar X-ray Mouse Image Mapping in Preclinical Nuclear Medicine Using Conditional Adversarial Networks. *J. Imaging* **2021**, *7*, 262. <https://doi.org/10.3390/jimaging7120262>

Academic Editors: George Kastis, Anastasios Gaitanis and Nicholas E. Protonotarios

Received: 25 October 2021

Accepted: 30 November 2021

Published: 3 December 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Drug discovery aims in identifying novel candidate treatments against diseases. It is a lengthy (it takes about 15 years) and costly (about 1–1.5 billion euros) process involving a laborious series of different stages, including preclinical studies [1,2].

A crucial part in the preclinical evaluation of drug discovery is performed on the basis of biodistribution ex vivo studies that aim in monitoring the interaction of radioisotope- or fluorescence-labeled candidate drugs on animal subjects invasively. However, this invasive method, although a gold standard, raises economical and ethical issues due to the large number of required animals [3]. Non-invasive methods, such as molecular imaging, have been proposed, aiming in improving animal handling and catalyzing the necessary time required for biodistribution ex vivo studies. Such methods include imaging using Positron Emission Tomography (PET), Single Photon Emission Tomography (SPECT), Computed Tomography (CT) and Magnetic Resonance Imaging (MRI), similar to those used in clinical practice. It is well proven that small animal imaging speeds up the whole process, increases accuracy and decreases cost [4–6].

Planar molecular imaging is another alternative since it can effectively be used to track a new tracer and study the accumulation from zero point in time post-injection. Two-

dimensional fluorescence/bioluminescence optical imaging is a well-established method for this purpose [7,8], offering high sensitivity, simplicity and decreased cost while having low depth penetration, limited clinical translation and being nonquantitative [1,2]. On the other hand, both SPECT and PET radioisotopes planar imaging provide quantitative information of high spatial resolution while also retaining simplicity, low cost and high throughput properties at the expense of access to radioactivity [9,10]. In order to provide simultaneous anatomic information, both techniques have been used, along with X-ray imaging, in commercial systems (In-Vivo MS FX PRO, Bruker) and prototypes [11], although increasing the cost and complexity, including radiation protection.

The aim of this work is to explore a potential solution for generating adequately accurate synthetic morphological X-ray images by translating standard photographic images of mice using deep learning. Such an approach would benefit planar molecular imaging studies since the acquired photographic images have poor resolution in terms of anatomical information. By superimposing adequately high-resolution synthesized anatomical images on real functional mouse images, it would be possible to generate a preliminary estimation of the drug distribution within the body of the animal using conventional photographic sensors, thus reducing cost and complexity by eliminating the need for X-ray imaging of the animal.

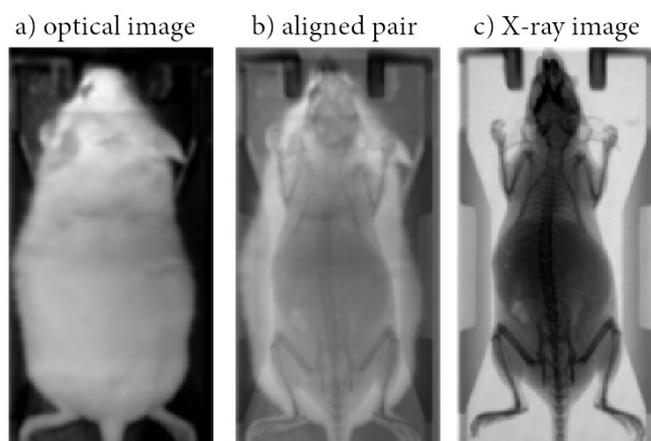
For this purpose, we trained and tested a well-known conditional generative adversarial network for image-to-image translation (pix2pix cGAN) [12,13] in a dataset of 780 paired photographic/X-ray mice images. Conditional generative adversarial networks (cGANs) have been used previously for image synthesis in molecular imaging applications [14–17]. The preliminary results of the current study have been presented by our group elsewhere [18]. Compared to our previous study, we have properly balanced, extended and finalized the dataset in order to optimize the prediction in different potential inputs. Moreover, the achieved performance of the pix2pix network using different loss functions in the current dataset is presented. Finally, two commercial planar nuclear medicine preclinical scanners with a pre-installed photographic sensor have been used for further evaluation by superimposing the acquired 2D nuclear images with the corresponding trained network outputs in real-time during in vivo experiments.

## 2. Methodology

A pix2pix cGAN was trained to map conventional photographic mouse images to X-ray scans [12]. Pix2pix is a common framework based on conditional Generative Adversarial Networks (cGANs), which predict pixels from pixels in any dataset, in which the aligned image pairs vary in the visual representation, but the renderings for, e.g., edges, stay the same (Figure 1). The above concept fits well to the presented dataset, as the field of view is clearly defined and the dimensions and weight of the mice usually involved in preclinical studies are similar. The methodology of the current work consists of 2 stages: (a) Data collection and preprocessing; (b) Modeling and performance evaluation.

### 2.1. Data Collection and Preprocessing

A dataset of 780 input/ground truth images has been acquired in order to train and test the pix2pix network. A photographic mouse image and the corresponding X-ray scan of the same anesthetized animal are referred to as input/ground truth images, respectively. The input images have been acquired using two commercial planar scanners for SPECT and PET radioisotope imaging, which contain a standard photographic sensor that provides an optical image of the animal as an anatomical mapping solution (eye's series, BIOEMTECH, Greece) (Figure 1a). The X-ray images have been acquired with an X-ray tube (Source-Ray Inc, US) and a CMOS detector (C10900D, Hamamatsu, Japan), both mounted on a prototype PET/SPECT/X-ray system (Figure 1c) [11]. Both systems are adequate for mice imaging providing a useful field of view (FOV) of 50 mm × 100 mm.



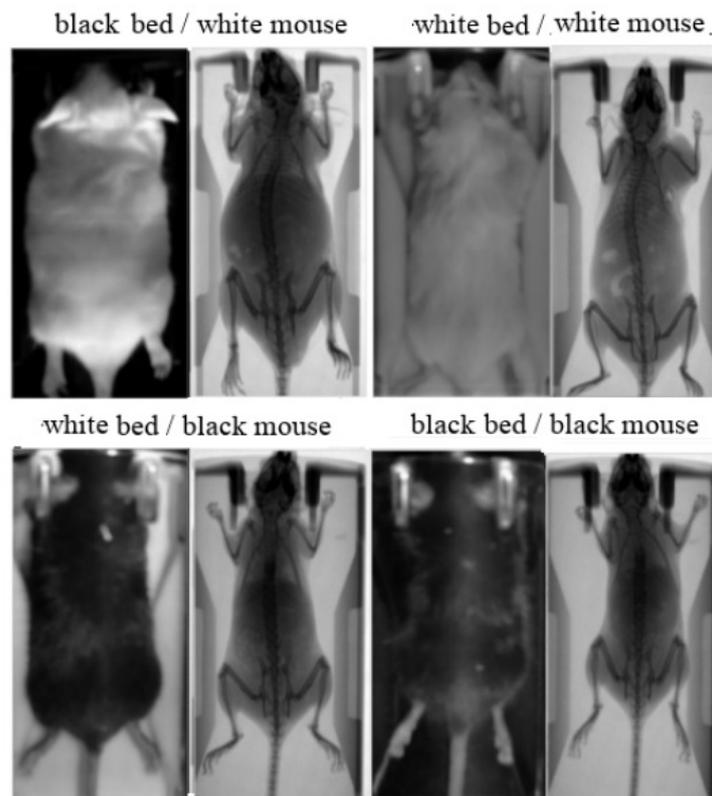
**Figure 1.** Indicative optical image acquired using a conventional photographic sensor located in BIOEMTECH eye’s series radioisotope screening tools (a); Indicative X-ray image acquired in a prototype PET/SPECT X-ray system (c) used as ground truth; Aligned pair (b).

We acquired 5 input/ground truth images of each animal in different poses upon the hosting bed. Given that small mice used in preclinical studies have similar dimensions and weight ( $\sim 20\text{--}40$  g), we assume that the aforementioned methodology does not affect the training procedure and minimizes the number of laboratory animals used in the study. In all cases, animals were anesthetized with isoflurane and kept warmed during the scans. All animal procedures were approved by the General Directorate of Veterinary Services (Attica Prefecture, Athens, Greece) and the Bioethical Committee of the Institution (Permit number: EL 25 BIOexp 04).

The study involved 78 white and 78 black Swiss albino mice, leading to a total number of 780 input/ground truth images. The mice were classified into two groups: (i) a group of 70 white and 70 black mice in order to collect the 700 paired images for training; (ii) a group of 8 white and 8 black mice in order to collect the 80 paired images for test and validation. Except mouse color, the method was evaluated against different animal hosting beds (plastic bed with white color; plastic bed with black color), which leads to different backgrounds in the input image. Four indicative input/ground truth pairs used for training and/or testing are illustrated in Figure 2. The paired images have been properly aligned prior to the training procedure, having  $512 \times 1024$  pixels resolution, corresponding to the  $50\text{ mm} \times 100\text{ mm}$  field of view (Figure 1b). The detailed characteristics of the dataset are summarized in Table 1. The ratio between the test and train images was kept higher than 10% in each occasion, and the individual mice of the training set are separate from the individual mice of the test set.

**Table 1.** Train and test dataset detailed characteristics.

Mouse Color	Bed Color	Train	Test	Test/Train (%)
White	Black	260	30	11.5
White	White	90	10	11.1
Black	White	260	30	11.5
Black	Black	90	10	11.1
<b>Total</b>		<b>700</b>	<b>80</b>	



**Figure 2.** Aligned image pairs used for training. Photographic input image (**left**); Corresponding X-ray scan used as ground truth (**right**).

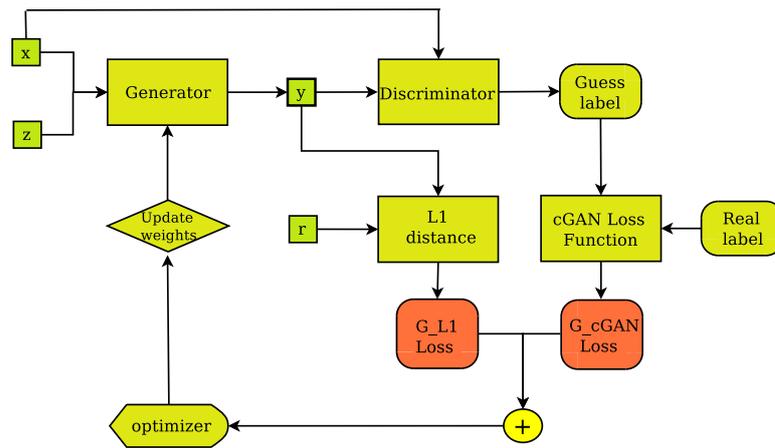
## 2.2. Modeling and Performance Evaluation

Pix2pix is a cGAN (conditional Generative Adversarial Network) deep learning network that can be used for image-to-image translation. A cGAN is designed to create synthetic images  $y$  from a set of known images  $x$  and a randomly created noise vector  $z$  as follows:

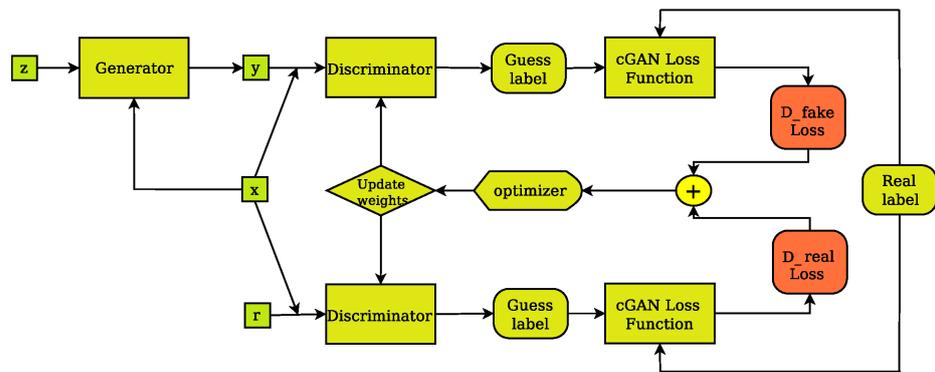
$$G : [x, z] \rightarrow y \quad (1)$$

It consists of two main sub-networks, the generator ( $G$ ) and the discriminator ( $D$ ). The generator (Figure 3) is designed to create real-like, fake images, whereas the discriminator (Figure 4) is designed to classify images as real or fake. During training, the generator tries to improve its performance by updating its parameters based on input from the discriminator. This enables the generator to create better, more realistic fake images. The discriminator, on the other hand, improves its performance as a stand-alone network by learning to separate fake images from their real counterparts. The above ideas can be expressed mathematically in terms of a loss function as follows:

$$LOSS_{cGAN}(G, D) = E_{x,y}[\log D(x, y)] + E_{x,z}[\log(1 - D(x, G(x, z)))] \quad (2)$$



**Figure 3.** The pix2pix Generator’s training layout. The generator creates output image (y) from input image (x) and random noise vector (z) and improves its performance by receiving feedback from the discriminator, as well as regarding the degree of fakeness of the synthetic image (y) compared to the ground truth (r).



**Figure 4.** The cGAN discriminator’s training layout. The discriminator compares the input(x)/ground truth(r) pair of images and the input(x)/output(y) pair of images and outputs its guess about how realistic they look. The weights vector of the discriminator is then updated based on the classification error of the input/output pair (D fake Loss) and the input/target pair (D real Loss).

The discriminator and the generator are adversaries in the sense that the generator attempts to minimize the Loss function, i.e., to fool the discriminator by producing real-like synthetic images, whereas the discriminator attempts to maximize the Loss function, i.e., expose imposter images. Training is complete when the generator is able to create fake, synthesized images, which are very difficult to be identified as fake by the discriminator. The pix2pix cGAN is further evolution of the above idea. Pix2pix is an implementation of the cGAN where the generation of an image is conditional on a specific target image instead of a set of target images. This way, a pix2pix generator (Figure 3) is not only trained to fool the discriminator but also to synthesize images that are as close as possible to the ground truth pair of each input image. This can be expressed mathematically by incorporating the idea of L1 distance between the paired generated image of the known input and the specific paired target image into Equation (2), as follows:

$$LOSS_{pix2pix} = LOSS_{cGAN}(G, D) + \lambda \times L_{L1}(G) \tag{3}$$

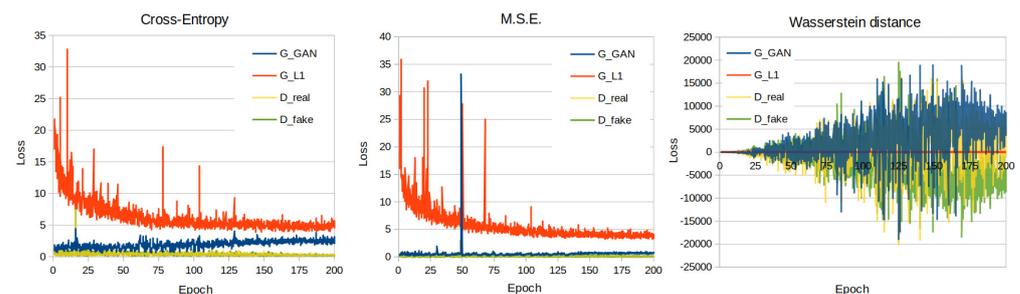
where  $\lambda$  is a regularization parameter.

In this study, we have utilized the PyTorch implementation [13] of the pix2pix algorithm as presented in [12]. In that implementation, the discriminator is designed as a PatchGAN, whereas the generator is designed as a U-Net. During the training process of all the models presented in this work, we used the default parameter and hyperpa-

parameter values of the PyTorch pix2pix implementation [13]. A list of values of important adjustable parameters for the training of the pix2pix models is illustrated in Table 2. In the present dataset, the cross entropy, which was used as the default cGAN loss function in the pix2pix implementation of Isola et al. [12], was evaluated against the mean squared error (MSE) (squared L2 norm) loss used in LSGANs [19], the Wasserstein distance loss and wGANs [20]. The loss curves of the presented models' training process are illustrated on Figure 5.

**Table 2.** Values of important adjustable training parameters and hyperparameters.

Parameter	Value
Learning rate	0.0002
Beta 1 parameter for the optimizer (adam)	0.5
Beta 2 parameter for the optimizer (adam)	0.999
Maximum epochs	200
Lambda ( $\lambda$ ) weight for the L1 loss	100
Generator layers	8
Discriminator layers	3
Load size	512
Mini batch size	1



**Figure 5.** Training loss curves of the cross entropy, M.S.E. and Wasserstein distance loss function models.

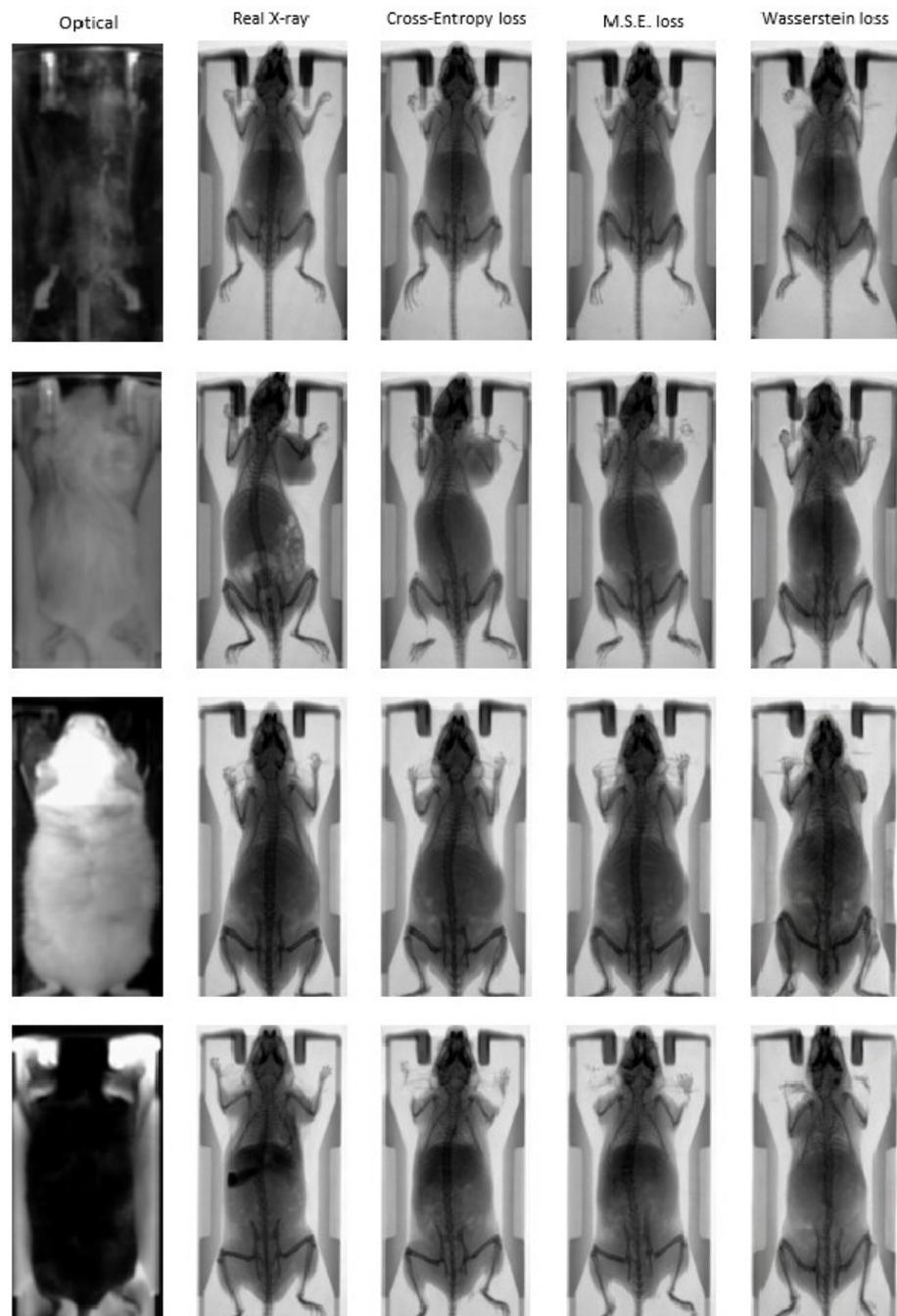
The performance evaluation of the pix2pix models in the present dataset has been performed using three image quality metrics: (a) peak signal-to-noise ratio (PSNR), (b) structural similarity index measure (SSIM) and (c) Fréchet inception distance (FID). PSNR is an expression for the ratio between the maximum possible signal power value and the power of distorting noise that affects the quality of its representation [21]; SSIM compares local patterns of pixel intensities that have been normalized for luminance and contrast [22]; FID is a metric focused on assessing the quality of images created by generative models and is considered one of the most widely accepted image quality metrics for image generation [23,24]. FID compares the distribution of the generated images with the distribution of the set of the target images and captures the similarity between the two sets. The aforementioned metrics have been used previously to assess the quality of cGAN-generated images quantitatively in comparison to other metrics, such as mean absolute error (MAE) and mean square error (MSE), which, in some cases, are not appropriate for evaluating the results of the cGAN approach [25–28].

### 3. Results

#### 3.1. Quantitative Evaluation

Figure 6 illustrates indicative optical to X-ray translations in the test dataset for the four distinct combinations of mouse and bed color that the pix2pix network has been trained on. The input optical image and the ground truth are presented alongside synthetic X-ray images generated by the pix2pix network that was trained using three different loss functions: (a) Cross Entropy; (b) MSE; (c) Wasserstein distance. The results demonstrate the

ability of the trained networks to efficiently map the photographic mouse image to an X-ray scan in terms of animal mapping, but not for diagnosis, as the information of the actual structure of mouse cannot be generated. However, the renderings and the morphological characteristics are reproduced in an adequate manner compared to the ground truth X-ray images, demonstrating the feasibility of the proposed approach for animal mapping.



**Figure 6.** Indicative “fake” X-ray images from the pix2pix trained network using different loss functions: Cross Entropy (3rd column); MSE (4th column); Wasserstein distance (5th column). The input photographic images and the corresponding ground truth images are presented in the first two columns.

Quantitative evaluation of network outputs compared to ground truth images has been performed using three performance metrics: (a) PSNR; (b) SSIM; (c) FID in the test dataset. According to the calculated values presented in Table 3, the pix2pix network trained using Cross Entropy as the cGAN loss function is able to generate X-ray images that are more accurate translations of the photographic image to the real X-ray scan compared to MSE and/or Wasserstein distance loss functions.

**Table 3.** Metrics of the different cGAN loss functions tested.

cGAN Loss Function	PSNR ↑	SSIM ↑	FID ↓
Cross entropy	21.923	0.771	85.428
MSE	21.954	0.770	90.824
Wasserstein distance	17.952	0.682	162.015

Table 4 presents the calculated metrics of the pix2pix Cross Entropy model on the different combinations of mouse and bed color. According to the FID values, the photographic images that have different colors of bed than the colors of the mouse are translated into images that resemble the ground truth better. This can be explained by (a) the greater percentage of such cases within the dataset, meaning the network is better trained on such cases, and (b) the greater contrast of the mouse against the background, which makes it easier for the network to distinguish the different subjects.

**Table 4.** Metrics of the pix2pix Cross Entropy model on the different combinations of mouse and bed color.

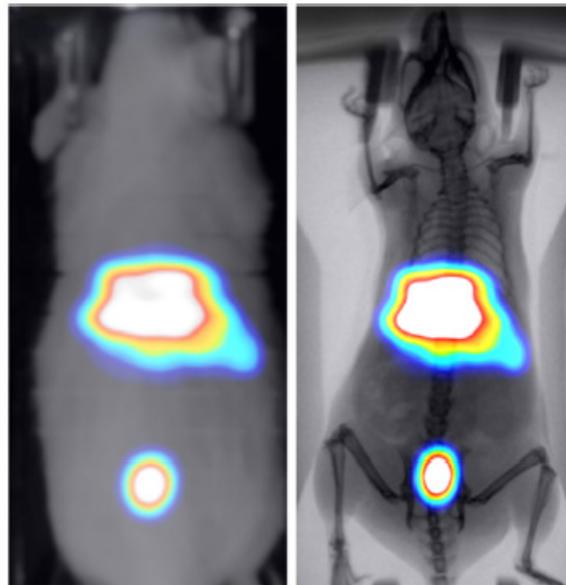
Mouse Color	Bed Color	PSNR ↑	SSIM ↑	FID ↓
black	white	22.808	0.791	112.948
black	black	22.894	0.794	151.006
white	black	21.196	0.750	109.116
white	white	20.270	0.743	163.056

### 3.2. Animal Mapping during In Vivo Molecular Imaging Experiments

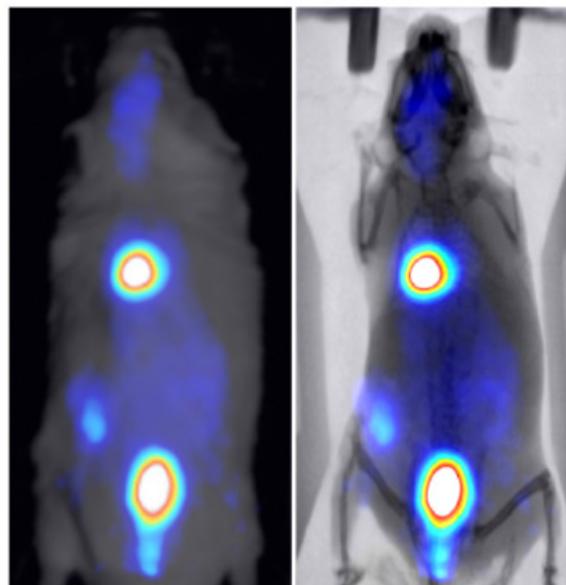
The proposed trained network was used for anatomical mouse mapping in two, proof of concept, nuclear molecular imaging experiments. Two planar preclinical scanners for SPECT and PET isotopes imaging have been used (eye's series, BIOEMTECH, Greece). The scanners come with a pre-installed photographic sensor, which is used to superimpose standard photography of the mouse with the acquired nuclear image for anatomical mapping purposes. The trained with the Cross Entropy loss function pix2pix network was used to generate an X-ray scan from the acquired photographic image in real-time during in vivo experiments. The output was then superimposed in real-time with the corresponding functional (PET or SPECT) image to provide high-resolution information for anatomical mapping.

Two healthy mice were administered through tail vein injection with 286 uCi/50 uL of technetium-99m ( $^{99m}\text{Tc}$ ) and 30 uCi/100 uL of 2-deoxy-2- $^{18}\text{F}$ fluoro-D-glucose ( $^{18}\text{F}$  FDG) to study the kinetics of these widely used tracers in BIOEMTECH's  $\gamma$ -eye scintigraphic system and  $\beta$ -eye planar coincidence imaging system, respectively. Figure 7 shows the  $^{99m}\text{Tc}$  nuclear image fused with the optical one provided in the  $\gamma$ -eye system and with the predicted X-ray image produced from the pix2pix network. The nuclear image shows the clear targeting of the tracer and the biodistribution in the liver, spleen and bladder, as the main organs of accumulation. Figure 8 shows the  $^{18}\text{F}$ FDG nuclear image fused with the optical one provided in the  $\beta$ -eye system and with the predicted X-ray image produced from the pix2pix network. The nuclear image shows the accumulation of the compound in the brain, heart, liver, intestines and bladder, as expected. In both cases, the produced

X-ray superimposed with the corresponding nuclear image provides an anatomical map of the small animal that enhances the overall image information.



**Figure 7.**  $^{99m}\text{Tc}$ -MDP-labelled nuclear image of a healthy mouse fused with the optical image provided in the  $\gamma$ -eye scintigraphic system (left) and the X-ray produced from the pix2pix trained network (right).



**Figure 8.**  $^{18}\text{F}$ -FDG nuclear image of a healthy mouse fused with the optical image provided in the  $\beta$ -eye planar coincidence imaging system (left) and the X-ray produced from the pix2pix trained network (right).

#### 4. Discussion and Conclusions

Image-to-image translation techniques have been used in medical imaging for several tasks, including segmentation, denoising, super-resolution, modality conversion and reconstruction [27]. In the current work, we present an off-the-shelf approach for generating adequately accurate synthetic morphological X-ray images by translating standard photographic mice images. Artificially produced X-ray mouse images can be superimposed with functional radioisotope or fluorescence/bioluminescence 2D images to enhance overall

anatomical information. Although not suitable for diagnosis, such information can be used for animal mapping in preclinical planar molecular imaging systems without adding the costs and complexity of radiation sources, radiation protection and animal irradiation during repeated scans.

A well-known cGAN network (pix2pix) that learns a mapping from the input image to the output image has been trained in a 700 input (mouse photography)/ground truth (mouse X-ray scan) dataset and evaluated in a 80 input/ground truth dataset. The pix2pix network has been evaluated against three different loss functions: (a) Cross Entropy, (b) MSE and (c) Wasserstein distance, with the first two achieving similar performance superior to the Wasserstein distance. However, the results in all cases show that the network predicts an X-ray image with sufficient accuracy and that the calculated metrics are comparable with those presented in other studies that have evaluated pix2pix networks. In [28], four approaches including the pix2pix network, which was used as a baseline, were evaluated on a well-known dataset primarily presented in [29]. The SSIM and PSNR values of the pix2pix model on that dataset were calculated equal to 0.286 and 12.868, respectively. In [30], a pix2pix network was used as a baseline and was evaluated on datasets presented in [31,32]. The SSIM and FID values of the pix2pix model on the [31] dataset were calculated equal to 0.770 and 66.752, respectively, while FID values on three datasets presented in [32] were 96.31 (handbag dataset), 197.492 (shoes dataset) and 190.161 (clothes dataset). In [33], four pix2pix variations were evaluated on two datasets presented in the same study and the original pix2pix network was used as a baseline. The FID scores of pix2pix on MR-to-CT scan and CT-to-MR scan translation tasks were 122.6 and 90.8, respectively. Although different problems are presented in the aforementioned studies, achieved metrics (Table 3) are indicative of the success of the approach in our dataset.

Two proof-of-concept nuclear molecular imaging *in vivo* experiments have been conducted in order to demonstrate the efficacy of the method in terms of enhancing anatomical information in functional imaging preclinical studies. For that purpose, two commercial planar radioisotope imaging systems with a pre-installed conventional photographic sensor have been used (eye's series, BIOEMTECH, Greece). The produced X-ray superimposed with the corresponding nuclear image provides an accurate morphological map of the small animal that is used to better identify the organs that the studied compound has accumulated.

Future work is oriented in expanding our dataset with data augmentation techniques in order to improve the generalization of our method and hyperparameter tuning in order to further tailor the behavior of the pix2pix network to the final dataset. Taking into account that the primary information comes from photographic sensors and given that small mice used in preclinical studies have similar dimensions and weight (~20–40 g), geometric and/or photometric transformations may expand the dataset without losing the generalization of the approach, thus providing better performance metrics.

**Author Contributions:** Conceptualization, E.F. and G.L.; methodology, E.F., V.E., S.K. and D.G.; software, V.E., C.-A.G.; validation, V.E., E.F. and D.G.; formal analysis, E.F.; investigation, E.F. and V.E., D.G.; resources, E.F., M.R. and G.L.; data curation, I.P., V.E., M.R. and S.S.; writing—original draft preparation, E.F. and D.G.; writing—review and editing, E.F., G.L., D.G., S.K. and V.E.; visualization, E.F. and V.E.; supervision, E.F.; project administration, D.G.; funding acquisition, E.F., D.G. and G.L. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research is co-financed by Greece and the European Union (European Social Fund-ESF) through the Operational Program “Human Resources Development, Education and Lifelong Learning 2014–2020” in the context of the project “Deep learning algorithms for molecular imaging applications” (MIS-5050329).

**Institutional Review Board Statement:** The study was conducted according to the guidelines of the Declaration of Helsinki and approved by General Directorate of Veterinary Services (Athens, Attica Prefecture, Greece) and the Bioethical Committee of the Institution (Permit number: EL25 BIOexp 04, date of approval: 15 January 2020).

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Not applicable.

**Acknowledgments:** We thank Giorgos Toliadis, assistant professor at the Czech Technical University in Prague, for helpful discussions.

**Conflicts of Interest:** The authors declare the following financial interests/personal relationships that may be considered as potential competing interests with the work reported in this paper: co-authors Vasilis Eleftheriadis, Irinaios Pilatis, Christina-Anna Gatsiou and Sophia Sarpaki are currently coworkers of BIOEMTECH. Co-authors Eleftherios Fysikopoulos, Maritina Rouchota and George Loudos are general partners of BIOEMTECH.

## References

1. Willmann, J.; van Bruggen, N.; Dinkelborg, L.; Gambhir, S. Molecular imaging in drug development. *Nat. Rev. Drug Discov.* **2008**, *7*, 591–607. [[CrossRef](#)] [[PubMed](#)]
2. Pysz, M.A.; Gambhir, S.S.; Willmann, J.K. Molecular imaging: Current status and emerging strategies. *Clin. Radiol.* **2010**, *65*, 500–516. [[CrossRef](#)] [[PubMed](#)]
3. Kagadis, G.; Ford, N.; Karbanatidis, D.; Loudos, G. *Handbook of Small Animal Imaging*; CRC Press: Boca Raton, FL, USA, 2016.
4. Paul, S.M.; Mytelka, D.S.; Dunwiddie, C.T.; Persinger, C.C.; Munos, B.H.; Lindborg, S.R. How to improve RD productivity: The pharmaceutical industry's grand challenge. *Nat. Rev. Drug Discov.* **2010**, *9*, 203–214. [[CrossRef](#)] [[PubMed](#)]
5. US Food and Drug Administration. *US Department of Health and Human Services; Estimating the Maximum Safe Starting Dose in Initial Clinical Trials for Therapeutics in Adult Healthy Volunteers; Guidance for Industry*: Rockville, MD, USA, 2015.
6. Mannheim, J.G.; Kara, F.; Doorduyn, J.; Fuchs, K.; Reischl, G.; Liang, S.; Verhoye, M.; Gremse, F.; Mezzanotte, L.; Huisman, M.C. Standardization of Small Animal Imaging-Current Status and Future Prospects. *Mol. Imaging Biol.* **2018**, *20*, 716–731. [[CrossRef](#)] [[PubMed](#)]
7. Stuker, F.; Ripoll, J.; Rudin, M. Fluorescence Molecular Tomography: Principles and Potential for Pharmaceutical Research. *Pharmaceutics* **2011**, *3*, 229–274. [[CrossRef](#)] [[PubMed](#)]
8. Debie, P.; Lafont, C.; Defrise, M.; Hansen, I.; van Willigen, D.M.; van Leeuwen, F.; Gijssbers, R.; D'Huyvetter, M.; Devoogdt, N.; Lahoutte, T.; et al. Size and affinity kinetics of nanobodies influence targeting and penetration of solid tumours. *J. Control Release* **2020**, *317*, 34–42. [[CrossRef](#)] [[PubMed](#)]
9. Georgiou, M.; Fysikopoulos, E.; Mikropoulos, K.; Fragogeorgi, E.; Loudos, G. Characterization of g-eye: A low cost benchtop mouse sized gamma camera for dynamic and static imaging studies. *Mol. Imaging Biol.* **2017**, *19*, 398. [[CrossRef](#)] [[PubMed](#)]
10. Zhang, H.; Bao, Q.; Vu, N.; Silverman, R.; Taschereau, R.; Berry-Pusey, B.; Douraghy, A.; Rannou, F.; Stout, D.; Chatziioannou, A. Performance evaluation of PETbox: A low cost bench top preclinical PET scanner. *Mol. Imaging Biol.* **2011**, *13*, 949–961. [[CrossRef](#)] [[PubMed](#)]
11. Rouchota, M.; Georgiou, M.; Fysikopoulos, E.; Fragogeorgi, E.; Mikropoulos, K.; Papadimitroulas, P.; Kagadis, G.; Loudos, G. A prototype PET/SPET/X-rays scanner dedicated for whole body small animal studies. *Hell. J. Nucl. Med.* **2017**, *20*, 146–153. [[PubMed](#)]
12. Isola, P.; Zhu, J.-Y.; Zhou, T.; Efros, A. Image-to-image translation with conditional adversarial networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Proceedings, Honolulu, HI, USA, 21–26 July 2017.
13. Paszke, A.; Gross, S.; Massa, F.; Lerer, A.; Bradbury, J.; Chanan, G.; Killeen, T.; Lin, Z.; Gimelshein, N.; Antiga, L.; et al. PyTorch: An Imperative Style, High-Performance Deep Learning Library. In Proceedings of the Advances in Neural Information Processing Systems, Vancouver, BC, Canada, 8–14 December 2019; Volume 32.
14. Gong, K.; Yang, J.; Larson, P.E.Z.; Behr, S.C.; Hope, T.A.; Seo, Y.; Li, Q. MR-based Attenuation Correction for Brain PET Using 3D Cycle-Consistent Adversarial Network. *IEEE Trans. Radiat. Plasma Med. Sci.* **2021**, *5*, 185–192. [[CrossRef](#)] [[PubMed](#)]
15. Amyar, A.; Ruan, S.; Vera, P.; Decazes, P.; Modzelewski, R. RADIOGAN: Deep Convolutional Conditional Generative Adversarial Network to Generate PET Images. In Proceedings of the 7th International Conference on Bioinformatics Research and Applications (ICBRA), Berlin, Germany, 13–15 September 2020.
16. Denck, J.; Guehring, J.; Maier, A.; Rothgang, E. Enhanced Magnetic Resonance Image Synthesis with Contrast-Aware Generative Adversarial Networks. *J. Imaging* **2021**, *7*, 133. [[CrossRef](#)] [[PubMed](#)]
17. Ouyang, J.; Chen, K.T.; Gong, E.; Pauly, J.; Zaharchuk, G. Ultra-low-dose PET reconstruction using generative adversarial network with feature matching and task-specific perceptual loss. *Med. Phys.* **2019**, *46*, 3555–3564. [[CrossRef](#)]
18. Fysikopoulos, E.; Rouchota, M.; Eleftheriadis, V.; Gatsiou, C.-A.; Pilatis, I.; Sarpaki, S.; Loudos, G.; Kostopoulos, S.; Glotsos, D. Photograph to X-ray Image Translation for Anatomical Mouse Mapping in Preclinical Nuclear Molecular Imaging. In Proceedings of the 2021 International Conference on Medical Imaging and Computer-Aided Diagnosis (MICAD 2021), Birmingham, UK, 25–26 March 2021.
19. Mao, X.; Li, Q.; Xie, H.; Lau, R.Y.K.; Wang, Z.; Smolley, S.P. Least Squares Generative Adversarial Networks. In Proceedings of the 2017 IEEE International Conference on Computer Vision (ICCV), Venice, Italy, 22–29 October 2017.
20. Arjovsky, M.; Chintala, S.; Bottou, L. Wasserstein GAN. *arXiv* **2017**, arXiv:1701.07875.

21. Sara, U.; Akter, M.; Uddin, M. Image Quality Assessment through FSIM, SSIM, MSE and PSNR—A Comparative Study. *J. Comput. Commun.* **2019**, *7*, 8–18. [[CrossRef](#)]
22. Wang, Z.; Bovik, A.C.; Sheikh, H.R.; Simoncelli, E.P. Image quality assessment: From error visibility to structural similarity. *IEEE Trans. Image Process.* **2004**, *13*, 600–612. [[CrossRef](#)] [[PubMed](#)]
23. Heusel, M.; Ramsauer, H.; Unterthiner, T.; Nessler, B.; Hochreiter, S. GANs Trained by a Two Time-Scale Update Rule Converge to a Local Nash Equilibrium. In Proceedings of the Advances in Neural Information Processing Systems, Long Beach, CA, USA, 4–9 December 2017.
24. Benny, Y.; Galanti, T.; Benaim, S.; Wolf, L. Evaluation Metrics for Conditional Image Generation. *Int. J. Comput. Vis.* **2021**, *129*, 1712–1731. [[CrossRef](#)]
25. Salimans, T.; Goodfellow, I.; Zaremba, W.; Cheung, V.; Radford, A.; Chen, X. Improved techniques for training gans. In Proceedings of the Advances in Neural Information Processing Systems, Barcelona, Spain, 5–10 December 2016.
26. Kupyn, O.; Budzan, V.; Mykhailych, M.; Mishkin, D.; Matas, J.; DeblurGAN, J. Blind motion deblurring using conditional adversarial networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–22 June 2018.
27. Kaji, S.; Kida, S. Overview of image-to-image translation by use of deep neural networks: Denoising, super-resolution, modality conversion, and reconstruction in medical imaging. *Radiol. Phys. Technol.* **2019**, *12*, 235–248. [[CrossRef](#)] [[PubMed](#)]
28. Yoo, J.; Eom, H.; Choi, Y. Image-to-image translation using a crossdomain auto-encoder and decoder. *Appl. Sci.* **2019**, *9*, 4780. [[CrossRef](#)]
29. Cordts, M.; Omran, M.; Ramos, S.; Rehfeld, T.; Enzweiler, M.; Benenson, R.; Franke, U.; Roth, S.; Schiele, B. The cityscapes dataset for semantic urban scene understanding. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 26–30 June 2016.
30. Albahar, B.; Huang, J. Guided Image-to-Image Translation With Bi-Directional Feature Transformation. In Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), Seoul, Korea, 27–28 October 2019.
31. Liu, Z.; Luo, P.; Qiu, S.; Wang, X.; Tang, X. DeepFashion: Powering Robust Clothes Recognition and Retrieval with Rich Annotations. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 26–30 June 2016.
32. Xian, W.; Sangkloy, P.; Agrawal, V.; Raj, A.; Lu, J.; Fang, C.; Yu, F.; Hays, J. TextureGAN: Controlling Deep Image Synthesis with Texture Patches. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Salt Lake City, UT, USA, 18–22 June 2018.
33. Paavilainen, P.; Akram, S.U.; Kannala, J. Bridging the Gap Between Paired and Unpaired Medical Image Translation. In Proceedings of the MICCAI Workshop on Deep Generative Models, Strasbourg, France, 1 October 2021.