

Comprehensive Review of AI Integration in Nuclear Medicine – Current Techniques, Limitations, and Future Innovations

Andrea García-Tejedor Bilbao-Goyoaga
University Medical Center (UMC) Utrecht
Utrecht University (UU)
Utrecht, Netherlands
a.garcia-tejedor@students.uu.nl

Abstract—The integration of Artificial Intelligence (AI) methods in Positron Emission Tomography (PET) and Single-Photon Emission Computed Tomography (SPECT) presents unique challenges and opportunities in healthcare. Despite the slower adoption of AI in medical fields compared to other domains, recent advancements have showcased its potential to revolutionize diagnostic precision and therapeutic innovation. This literature review explores the integration of AI in nuclear imaging, focusing on the applications in photon detection, image reconstruction, and post-processing, as well as in further image analysis where segmentation and radiomics play an important role. Specific examples making use of different Machine Learning, and Deep Learning techniques such as Convolutional Neural Networks (CNNs) and Generative Adversarial Networks (GANs) have been reviewed, demonstrating their ability to either outperform the conventional methods in extracting information from images or to automatize those that are tedious and time-consuming for clinicians. Despite the great results obtained in research, many limitations keep these methods still a step behind in their commercialization. This review aims to provide insights into the current AI applications in nuclear imaging that address challenges such as data complexity, standardization, and lack of explainability, along with the expectations of future directions for research and clinical implementation.

Index Terms—Medical Imaging, Nuclear Medicine, PET, SPECT, AI, Radiomics, Explainable AI

I INTRODUCTION

Background

From the first Computer-Aided Diagnosis systems to the introduction of Convolutional Neural Networks (CNNs) that allowed for the extraction of features from images, the application of Artificial Intelligence (AI) methods in the medical field has been notably slower than into the rest of engineering domains [1][2]. This disparity can be attributed to healthcare’s unique challenges: the complexity of medical data, diverse imaging modalities, varied protocols, the need for large datasets, and concerns about AI precision have slowed progress throughout history. In a field where people’s health is at risk, building trust in algorithms is crucial. The need for a separate analysis of AI methods for nuclear imaging from other imaging techniques comes from

the unique characteristics of Positron Emission Tomography (PET) and Single-Photon Emission Computed Tomography (SPECT), which involve the use of radioactive tracers to visualize and analyze physiological processes at the molecular level rather than anatomical structures alone, requiring specific considerations to develop effective and accurate AI methods: the analyses on these images need to take into account quantitative data related to physiological processes, such as metabolic activity or blood flow. Additionally, images generated by these techniques usually show higher levels of noise and variability in the data compared to some other imaging techniques. Image quality correlates with the injected radioisotope dose and acquisition time [3], so AI methods need to account for these characteristics and enhance image quality to minimize radiation dose while maintaining diagnostic accuracy. On the other hand, nuclear imaging is frequently used in combination with other imaging modalities like CT or MRI to provide complementary information, and some studies involve dynamic imaging, capturing the temporal changes in tracer distribution over time. The developed AI methods might have to integrate multi-modal information or dynamic datasets, which adds dimension to the analysis.

Nowadays, thanks to the new technological advancements and in-depth research performed in the field, AI has been demonstrated to be a powerful and trustworthy tool in the field of health especially in medical imaging, introducing us into a new era of diagnostic precision and therapeutic innovation [4][5][6][7]. The application of AI methods in nuclear imaging techniques, such as PET and SPECT, is promising for revolutionizing healthcare by providing deeper insights into physiological processes at the molecular level. In the study of the integration of AI in nuclear medicine, it is crucial to address the challenges, understanding that overcoming data complexity and giving impetus to the validation and explainability of such techniques is essential [8] [9]. More detailed information on the theory behind the algorithms mentioned in this paper can be found in the Appendix section.

Objectives

Addressing these challenges, this literature review explores the integration of AI methods in nuclear medicine. The objective is to analyze of AI applications in the full pipeline of nuclear imaging, reviewing techniques in use today, their successes and limitations, as well as those in development that hold potential for implementation in research and clinical settings.

II AI APPLICATIONS IN NUCLEAR MEDICINE

Image generation

PET and SPECT acquisitions consist of transverse images containing the distribution of a radiotracer within the human body, and the devices' computer reconstructs the transverse images using filtered backprojection method or iterative reconstruction methods, that are more accurate but computationally less efficient [10]. In PET, the correction for nonuniform attenuation can be implemented on the projection data before the reconstruction process, in contrast to SPECT where such correction is intertwined with the reconstruction process and makes it more complex [11]. The spatial and temporal resolution of PET images are limited by the amount of photon counts, scanner design, and underlying physics. Although advances in statistical image reconstruction methods have mitigated some limitations, there is still a need to enhance resolution and reduce noise, and even when image quality is acceptable, the concern for lowering the radiation dose demands shorter acquisition times and reduced doses [12]. Current AI methods on nuclear image generation can be summarized in three main areas:

1 Detection systems

In PET and SPECT image formation, most detectors are based on scintillating crystals, which absorb 511-keV photons, producing a short burst of light collected by fast photodetectors. The electronics of the detection systems digitize and decode photodetector signals, measuring parameters such as total charge, arrival time, and pulse shape [12]. However, scintillation photon production is negatively affected by several effects due to limited spatial resolution; photon scatter, photon attenuation, and partial volume artifacts [13] [14]. For these detectors, machine learning algorithms have not been deeply studied for positioning estimation since traditional linear estimation methods have straightforward relations with lookup tables that give accurate results. However, machine learning can still play a crucial role in enhancing positioning accuracy, particularly in the challenging identification of intercrystal scatter events that impact both spatial and energy resolution. Methods employing Support Vector Machines (SVM) have shown promise in distinguishing intercrystal scatter, offering simplicity and avoiding time-consuming training steps [15]. Nevertheless, simulated data is used, so there are still challenges in obtaining experimental training data labels. Additionally, a neural network approach was demonstrated by

Michaud et al. [16] for the LabPET scanner, using Monte Carlo simulated data to characterize intercrystal coincidences, resulting in a notable increase in sensitivity. These emerging machine-learning approaches demonstrate their potential to improve the performance of pixelated detectors and address challenges associated with intercrystal scatter.

Monolithic detectors are another type of detector used in nuclear imaging, which is composed of an array of photodetector elements coupled to one crystal face. They measure the spatial distribution of scintillation light to estimate the position of interaction, giving advantages like higher sensitivity, intrinsic depth of interaction measurement, and improved spatial resolution [17]. However, challenges include a nonlinear relationship between position-of-interaction and light distribution, especially at crystal edges, making position estimation difficult, particularly for thick monolithic crystals. Additionally, these detectors require complex and time-consuming calibration procedures, making the scalation to a complete PET system challenging [12]. The aim of machine learning methods in this step is to better localize the position-of-interaction of the source photons from the information collected in the detectors in comparison to the conventional methods. For monolithic detectors, multilayer Artificial Neural Networks (ANNs) (usually based on Levenberg–Marquardt backpropagation) are used to estimate the 2-D or 3-D position of interaction. The training is done by using labeled data acquired through pencil-beam irradiation at various angles, where the input consists of the charge collected by each photodetector and the output is a regression or a classification array representing the distribution of positions [18]. These approaches have outperformed conventional methods such as center-of-gravity or fitting methods, reducing positioning bias at detector edges [19]. Fig. 1 illustrates a deep-learning learning-based event-positioning scheme in monolithic detectors. An alternative approach for positioning has also been described by Muller et al. [21] [22] with an algorithm based on gradient tree boosting algorithms (GTB) that can handle different sets of input features, their combinations, and partially missing data. GTB builds predictive regression models based on sequential binary comparisons (decision trees), making them computationally faster. Since GTB models are implementable in FPGA if the memory requirements allow, two positioning optimization algorithms were proposed: one with memory restriction (for future FPGA implementations) and one without. Peng et al. [23] developed a quasi-monolithic detector using thin scintillator slabs stacked and read on their sides with SiPMs. CNNs were used to accurately determine the position of interaction and detect any scattering between layers by mapping the charge collected by the SiPM to a 2D map in each layer, obtaining an average spatial resolution of 0.40 mm and increased sensitivity without compromising other key performance features. In addition to ANNs, lazy-learning Machine Learning Algorithms are used for this task, such as k-nearest neighbors (kNNs) [24]. These techniques, rather than training a model, use the information from a dataset

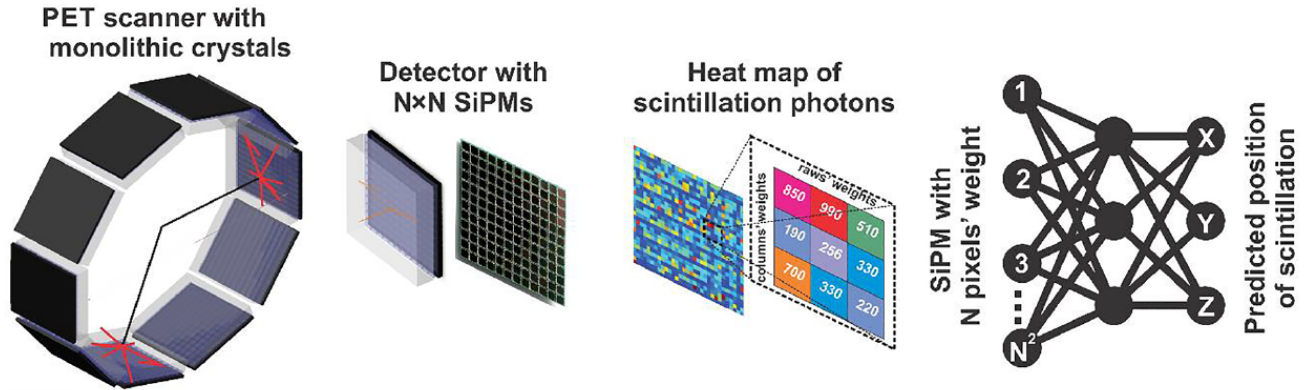


Fig. 1: Deep learning-based event positioning in monolithic detectors [20].

containing the position and outputs from the detector and determine the position of interaction by comparing the test event to the known events. However, this method requires large storage and is computationally challenging.

Besides spatial resolution, good timing resolution in the detection process is essential for the reconstructed image quality and for the accurate rejection of coincidence events in PET and SPECT imaging. Various physical processes contribute to timing uncertainty, including scintillation light generation, variable light propagation influenced by detector design, timing jitter during electrical conversion, and random noise sources [12]. Traditional linear methods using predefined thresholds for estimating time-of-interaction reduce the information contained in the photodetector signals into a single linear estimator, leading to inaccuracies in timing discrimination. Deep learning algorithms offer the potential to exploit the full information in the time-varying photodetector signals. In particular, CNNs can learn complex patterns and have translational invariance. A study by Berg et al. [25] uses CNNs to estimate Time Of Flight (TOF) directly from the pair of digitized detector waveforms for a coincident event in PET. By using ground-truth-labeled data obtained from experimental setups using photomultiplier tube-based scintillation detectors, the CNN-based approach improves timing resolution by 20% compared to traditional methods like leading edge discrimination and constant fraction discrimination. The study demonstrates that CNN depth significantly impacts timing resolution, while other network parameters have minor effects.

2 Image reconstruction

For many years, the golden standard technique in SPECT and PET image reconstruction has been iterative reconstruction [11]. In this process, extensive pre-processing and data correction are required before image reconstruction. Attenuation correction is a crucial step for accurate quantitative image reconstruction and is typically done using attenuation maps derived from coregistered CT data in SPECT or PET/CT systems. The case of SPECT or PET/MRI systems is more challenging since the signal and the attenuation coefficients are

not directly related, so methods based on atlases or segmentation are employed (e.g. see Fig. 2). However, these approaches require pairs of MR and transmission maps, which are not always available, and might suffer from quantification errors [26] [27]. Additionally, the generated images suffer from high noise levels, which affect image quality, lesion detectability, and quantitative accuracy. With the advancements in algorithms and computer power, many new machine-learning and AI-based techniques have been tested and implemented to solve the complex inverse problem of image reconstruction addressing these challenges and giving focus to improvement of image quality [14].

To save up on radiation dose in the often acquired simultaneous CT, many studies use Deep Learning methods to estimate attenuation maps from MRI scans instead. Most of these methods are based on creating pseudo-CT scans from the MRI data by training a CNN using pairs of MR images as input and CT images as labels [29]. Generally, the networks are based on U-Net architecture [30], and recently, a study by Kaviani et. al. [31] has integrated a residual U-Net-transformer regularizer into the unrolled maximum a posteriori expectation maximization (MAPEM) algorithm for PET image reconstruction, obtaining less noisy images and with better edge preservation compared to other methods, demonstrating potential for clinical use. Additionally, several studies have reported improved accuracy in PET and SPECT quantification by estimating attenuation maps by creating pseudo-CTs through augmented generative adversarial networks (GANs) using MRI images as input [32] [33] (for more information about GANs, see Appendix). Cycle-GANs, a type of unsupervised GANs, have also been used to avoid the need for matching pairs in the training process [33]. For now, GAN-based approaches to generate attenuation maps have only been applied to brain and pelvic regions, since registration errors are relatively high on most body parts and still more research needs to be done to improve metrics [28].

Some approaches aim to perform image reconstruction with various corrections using PET data only, eliminating the need for simultaneous CT or MRI acquisition. For example, by

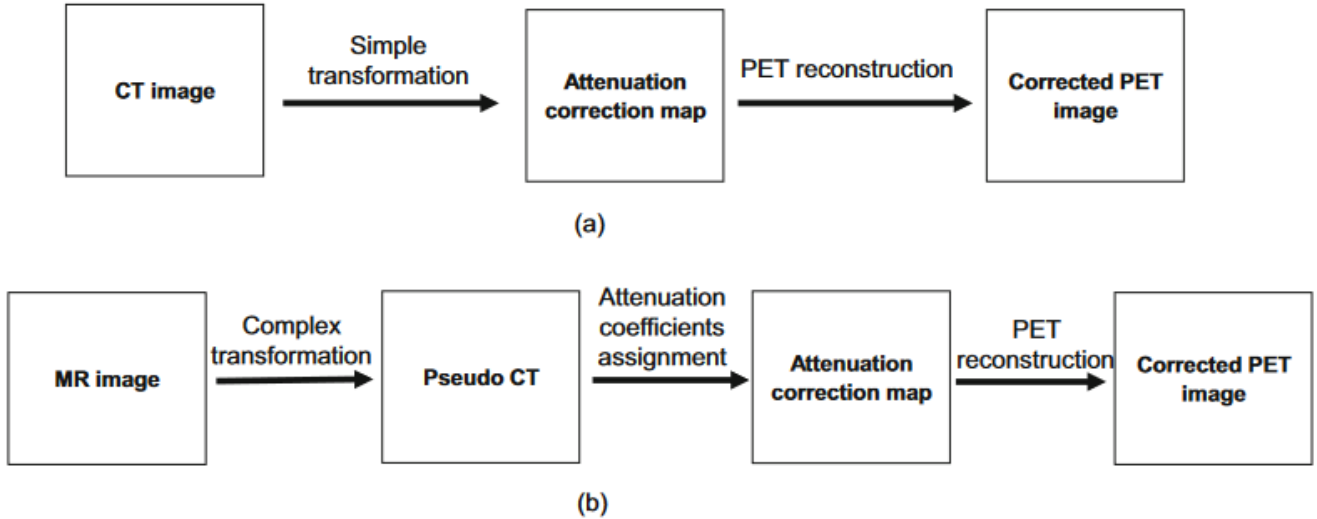


Fig. 2: The overview process of deriving attenuation correction maps from a) CT images, b) MR images [28]

training a CNN to directly map reconstructed images without attenuation correction to images with attenuation correction [34]. Yang et. al. [35] propose a CNN that performs image reconstruction while correcting for attenuation and scatter at the same time, without the need for a further post-processing denoising step. However, the results show many outliers due to the fact that PET data can only determine attenuation factors up to a constant, making it insufficient to determine the emission activity and attenuation map needed for scatter estimation. Furthermore, DeepPET [36], is a deep encoder-decoder network that can perform PET image reconstruction from sinogram data without the need of noise models and correction maps. This proposal increases the reconstruction speed and image quality over the traditional methods and has been approved by the FDA. Figure 3 shows a diagram of the architecture used in DeepPET. Recently, DPIR-Net [37] was proposed based on DeepPET and Wasserstein GAN (WGAN) applying a discriminator to the DeepPET architecture, which solved the problem of excessive smoothness and loss of details present in traditional methods. Algorithms designed for SPECT image reconstruction are significantly fewer than PET ones, although the methods utilized can be very similar [20]. Analogously to DeepPET, SPECTnet was proposed by Shao et. al. [38] to reconstruct SPECT images from the obtained projection data. The architecture of SPECTnet is composed of a neural network that compresses the data into a vector space, and a decoder that is taken from an autoencoder previously trained to recover the input image at its output. These straightforward methods have many limitations in terms of explainability, which highlights the importance of using domain-specific networks. Optimizing each step of the process at a time instead of performing the full pipeline at once, gives more control and traceability when trying to fix errors or understand the reasons why the model gives a certain output.

3 Image post-processing

Many studies keep exploring methods for post-processing alone independent from the reconstruction step. Post-processing techniques aim to enhance image quality and reduce noise, especially valuable in low-dose and fast-acquisition images (See Fig. 4). Noise can be decreased by increasing the injected radiotracer dose, as more photons decay and reach out to the detectors, but has, however, a negative impact on patients' health and comfort [20]. Traditional methods use image processing and filtering, statistical modeling, and MRI-guided noise removal among other techniques to minimize noise, but they experience issues such as loss of spatial resolution and excessive smoothing.

The first study that showed significant dose reduction in PET imaging using CNNs was conducted by Xu et al. [39]. In their research, a 2.5D U-Net with supervised residual learning was employed to transform low-dose 18F-FDG PET brain scans into full-dose equivalents, achieving a 200 dose reduction factor and demonstrating superior performance in various metrics such as signal-to-noise-ratio (SNR) compared to previous methods. Subsequent AI methods in post-processing mostly have aimed to train CNNs to directly map reconstructed noisy images, to denoised high-quality images. In the study conducted by Shiri et. al. low-dose SPECT images with reduced acquisition time per projection and the number of angular projections were mapped to full-dose ones using ResNet residual neural networks [40], demonstrating that the deep neural network recovers image quality and reduces bias in quantification metrics. Lu et. al [41] demonstrated an increase in quantitative accuracy of small lung nodules and improvement in visual image quality by using a denoising CNN based on U-Net for postprocessing that mapped 10% low-dose to standard dose deconstructed PET images. In this study, the effect of different network architectures, image dimensions, labels, and inputs concerning both noise reduction

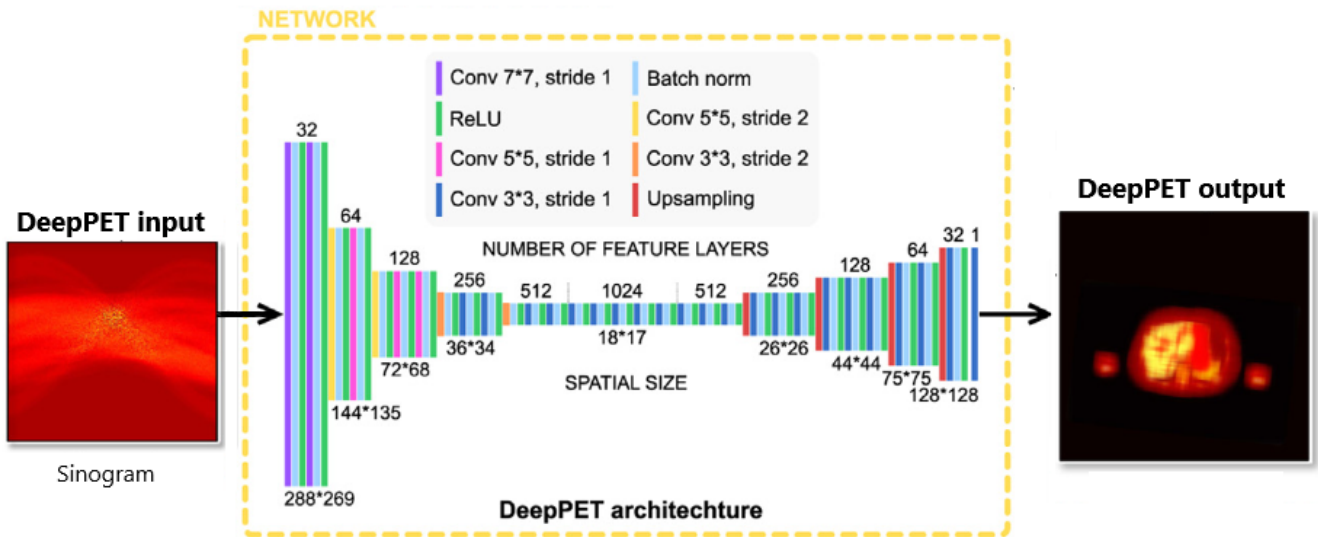


Fig. 3: Schematic architecture of DeepPET for direct PET image reconstruction using sinogram data as input and outputting PET images [36].

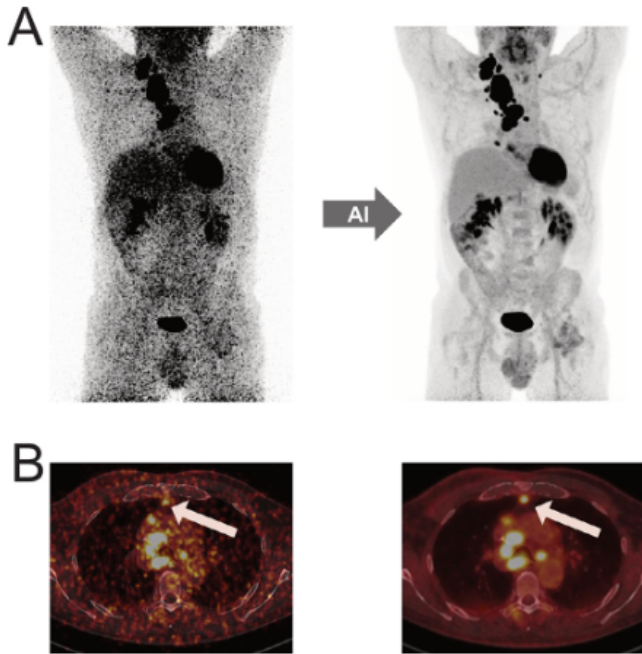


Fig. 4: PET image denoising using AI. A - Ultra-short emission time FDG-PET Maximum Intensity Projection (left) together with the standard acquisition protocol (50 times longer acquisition) (right). B - Slice-wise visualization showing an improvement in lesion detectability (arrow) [26]

performance and quantitative accuracy were also studied. The mapping has also been studied in the projection space instead of in the image space such as sinogram-based super-resolution [42]. Sanaat et al. [43] compared mappings from low-dose to full-dose using a 3D U-Net in both image space and projection space, revealing superior results in sinogram space with significantly higher PSNR and lower SUV bias.

Influenced by the limited availability of low and high quality training pairs, some studies have also included unsupervised learning and transfer-learning techniques for image denoising [44]. The deep image prior (DIP) method [45] consists of forcing the recovered image to be synthesized from a given deep architecture, and was employed as a post-processing technique for noisy whole-body PET images by Cui et al. [46] showing improved contrast-to-noise ratio compared to other approaches.

In the cases where PET/CT or PET/MR scanners are available, anatomical priors can be employed as additional input channels for PET image denoising, as well as previous PET scans of the same patient, reviewed in [47]. Chen et al. [48] employed a residual U-Net with a 2D input that incorporated multi-contrast T1, T2, FLAIR, and a 100-fold low-dose PET and outputted to full-dose images. The diagnostic accuracy for amyloid using the predicted full-dose images exhibited was comparable to the accuracy using real full-dose acquisitions. U-Net modifications have also been explored, integrating feature extractors for low-dose PET and T1-weighted MR inputs before concatenating the input data. Transfer-learning was also studied for amyloid image quality enhancement from different scanning hardware and protocols [49]. In this study, the network inputs are multi-contrast MR (T1-, T2-, and T2 FLAIR-weighted) and the ultra-low-dose PET images from one study, and the neural network parameters are transferred to another study rather than the data. Comparisons with random initialization and trained with the full dataset show better generalizability and image quality when applying transfer-learning.

While these studies demonstrate proof-of-principles, large-scale clinical evaluations are necessary before widespread clinical application. The combination of AI-based reconstruc-

tion, denoising, and advancements in PET technology, are key for the development of ultra-low-dose PET imaging with improved signal-to-noise ratios and image quality. However, further research and validation of real-world data are essential before these methods can be widely adopted in clinical practice. In addition, learning methods that can effectively generalize for different imaging protocols and clinical scenarios will be crucial for the widespread clinical application of deep learning in PET imaging.

Image segmentation

Segmentation is used in medical imaging to assign labels to pixels (in 2D) or voxels (in 3D) to separate the different regions, organs, or tissues. This step is essential for the quantification of certain tissues, for feature extraction, and for the better visualization of relevant areas. However, this is especially challenging in nuclear imaging given the poor spatial resolution and high statistical noise of PET and SPECT. Currently, manual segmentation performed by clinicians is the most common practice because of its accuracy, but it is, however, extremely time-consuming, labor-intensive, and prone to intra- and inter-observer variability, especially for 3D imaging [20]. Low-level segmentation methods have also been developed and commercialized, based on image operations (E.g thresholding and morphological operations), deformable models, and dynamic computing, but their performance is limited and user interaction is often needed as initial input or to correct the result [50]. These methods also do not consider the spatial correlations of voxels, making them sensitive to image noise, uptake inhomogeneity, and partial volume effect, and are also sensitive to the selected threshold values, leading to large differences of volume [51]. In the first MICCAI PET segmentation challenge [52], DL-based algorithms outperformed every other technique in the task of image segmentation on a dataset of 176 PET images from simulation, phantom, and clinical studies. This paved the way for other ML- and DL-based methods for single and multi-modal PET and SPECT imaging explored by recent studies mainly in the fields of oncology, cardiology, and neurology [53].

1 Oncology

In the field of oncology, the precise delineation of tumor tissue is essential for effectively treating cancer, and for avoiding irradiating organs at risk, which directly affects the survival of patients. In addition to radiation therapy, accurate tumor segmentation is important for further quantitative analysis and radiomics. PET imaging with FluoroDeoxyGlucose (FDG-PET), often integrated with CT, is a very powerful imaging tool for cancer detection, which is promising for its improved clinical diagnosis, assessment of prognosis, treatment planning, and for and monitoring progress after treatment [54]. Lung cancer is the most studied application of this technique. A study by Ikushima et. al. [55] presents a framework for ML-based delineation of gross tumor volumes (GTVs) of lung cancer patients, utilizing an optimum contour selection

(OCS) method. The framework involves training a machine-learning classifier with image features surrounding GTV contours identified by radiation oncologists. An SVM is employed to extract initial GTV regions, with subsequent refinement using the OCS method to select the best contour. The efficacy of the framework was evaluated in 14 lung cancer cases, showing promising results with an average Dice similarity Coefficient (DSC) of 0.777, demonstrating a great potential to aid oncologists in delineating GTV regions accurately. Deep learning, fully convolutional networks (DFCN) were studied by Zhong et.al [56] for the simultaneous tumor cosegmentation on dual-modality non-small cell lung cancer (NSCLC) in PET and CT images. This method utilizes two coupled 3D U-Nets with an encoder-decoder architecture to leverage both PET and CT information. The results outperformed methods based on PET or CT alone in different evaluation metrics, demonstrating the promise of cosegmentation for enhancing tumor delineation and integrating with multi-modality imaging tools for future clinical trials. The use of FDG PET/CT imaging in the evaluation of head and neck cancer has significantly risen in recent decades [57], together with the development of deep CNN methods to contour the GTV [58]. A study by Guo et. al [59] explores the use of a Dense-Net architecture to perform GTV delineation of PET/CT images of head and neck cancer patients undergoing radiation therapy. The evaluation on a dataset of 250 HNC patients using manually delineated GTV contours as ground truth, demonstrated that the multi-modality network outperforms both single-modality approaches and the compared 3D U-Net framework, achieving higher Dice coefficients and reducing prediction variability. Fig. 5 shows an example of the result. Recently, the second edition of the HEAd and neCK TumOR (HECKTOR) challenge [60] was held to obtain the best performance on the automatic segmentation of the GTV of head and neck oropharyngeal primary tumors in FDG-PET/CT images. Most of the approaches to the challenge were Deep Learning methods based on U-Net architectures. The winner's work used a well-tuned patch-based 3D self-configuring U-Net called nn-UNet [61] with standard pre-processing and training, a dynamically adjusted learning rate and they used the Squeeze and Excitation (SE) normalization [62], which was key for the algorithm robustness.

For general oncological image segmentation, a recent study by Andrade et. al showed a very complete analysis of multi-modal medical Transformers [63]. The researchers conducted controlled studies comparing various models, including hybrid CNN-ViT encoder-based models, pure Transformer encoder networks, and CNN-based models. The results show that a hybrid model is the most effective, outperforming other models in most ranking strategies, but the performance varies depending on the task and dataset used for evaluation. Although hybrid approaches incorporating both CNN and Transformer components generally outperform pure Transformer-based models, it needs to be taken into account that spatial downsampling influences Transformers' ability to capture long-range dependencies. The study recommends using the nn-UNet pipeline as a starting point for evaluating

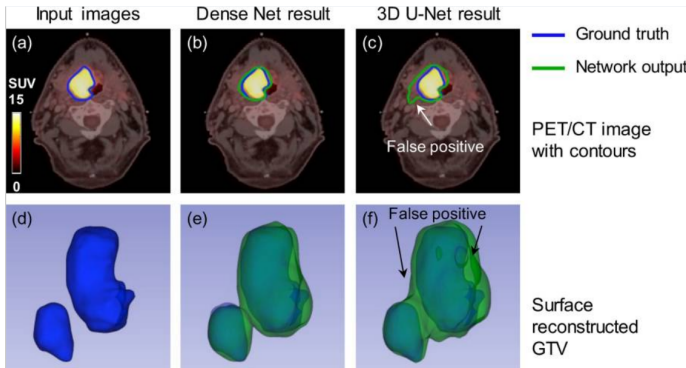


Fig. 5: Comparison of the segmentation results from the (b) multi-modality Dense-Net and (c) 3D U-Net. (a) Input image; (d, e, f) corresponding 3D visualizations of (a, b, c), respectively [59].

image segmentation architectures since it is a well-established and robust pipeline.

2 Cardiology

Segmentation steps are also crucial in the field of heart imaging. For example, for the assessment of left ventricular (LV) function, myocardial perfusion SPECT (MPS) has been one of the most important imaging modalities, and the performance of the assessment is affected by the accuracy of the segmentation. CNNs based on V-Net architecture have been used to automatically segment LV myocardium by delineating its endocardial and epicardial surface, showing great results in various metrics using manual delineations as ground truth, which demonstrated the feasibility of using learning methods in accurately quantifying LV myocardium volume change over the cardiac cycle [64]. The use of U-Net architecture was also proposed by Zhang et. al. [65] for the same task, to assess LV function in patients with coronary artery disease. The correlation of the predicted LV volume and ground truth was good with minimal mean relative error, and the correlation with the commercial software is fair, but it can only be considered a reference, and not a golden standard since other post-processing techniques are used in the commercial software, which leads to discrepancies. A similar approach with 3D U-Net architecture has also been studied by Piri et. al. [66] to segment the heart from ^{18}F -sodiumfluoride (NaF)-PET/ CT scans, with the final objective of assessing the severity of atherosclerosis affecting the coronary arteries and the heart, which gave a negligible difference in the estimations with manual segmentations, and in a much shorter time.

3 Neurology

Segmentation plays a critical role in various aspects of neuroimaging analysis, providing valuable insights into brain structure, function, and pathology. Several studies have highlighted its benefits. For instance, the effectiveness of segmenting gray matter in ictal-interictal perfusion SPECT and interictal ^{18}F -FDG-PET scans for patients with medically refractory

epilepsy has been demonstrated. This segmentation approach enhances the detection of small cortical areas exhibiting ^{18}F -FDG-PET hypometabolism, while also eliminating hyperperfused seizure propagation pathways within white matter during ictal perfusion SPECT imaging [67]. Masking gray matter can reduce the influence of partial volume effect and retention in white matter and cerebrospinal fluid [68]. In some other applications, the segmentations are performed in MRI or CT images that are registered into PET and SPECT. In the case of dementia [53], it has been proven that errors in MRI-based segmentations can lead to errors in PET quantitation. In most studies, the segmentations are performed either manually or with techniques like gradient-based segmentation, region growing (such as fuzzy-c-means), statistical algorithms, machine learning, and texture-based segmentations, but deep-learning based methods have also been recently proposed to overcome their limitations, but its use is very scarce yet. GANs have been proposed by Oh et. al [69] to segment white matter in ^{18}F -FDG PET/CT images, given its ability to measure changes in glucose metabolism in the brain that allows the diagnosis of neurodegenerative diseases at early stages: changes in white matter volume have been linked to aging, psychosis, multiple sclerosis, and Alzheimer's disease, with white matter hyperintensities associated with increased risk of vascular dementia and cognitive decline. In this study, the images were first preprocessed to obtain only brain region and coregistered to MRI images because of their anatomical accuracy, and then a generator was trained to produce a segmentation map from the ^{18}F -FDG-PET/CT image that resembled the real segmentation map, while the discriminator was trained to differentiate between the generated segmentation map and the real one. Results show excellent performance in generating segmentation maps compared to other commonly used deep-learning methods, but clinical implications should be further studied. Another study by the same authors, proposes W_{hyper} -GAN [70], a similar method for White Matter Hyperintensity segmentation. A feasibility study by Pasini et. al. [71] uses deep learning models, specifically 3D U-Net and V-Net networks, to automatically segment diagnostic regions associated with Alzheimer's disease in ^{18}F -FDG PET scans. The dataset includes volumes from controls, Alzheimer's disease, and mild cognitive impairment patients, with ground truth manually generated by expert users corresponding to the delineation of six regions including temporal lobes, parietal lobes, and frontal lobes. Results based on the evaluation metrics (DCS, overlapping area coefficient (AOC), and extra area coefficient (EAC)) indicate a significant improvement, with 3D U-Net achieving the best performance, even when hypometabolic regions were present.

Radiomics

Radiomics, first introduced in 2012 [72], is a rapidly growing area of quantitative image analysis, that is defined as the high-throughput extraction of large amounts of image features from radiographic images, to associate them to clinical or biological endpoints [73]. These features can serve as

non-invasive biomarkers for tumor characterization, prognostic classification, and prediction of treatment response, advancing precision medicine. In PET and SPECT imaging, biological processes are quantitatively expressed by the spatial distribution of radiotracer uptake, so radiomics can be used in these modalities to quantify such uptake. The field of radiomics can currently be divided into two areas: traditionally handcrafted radiomics and deep-learning radiomics. The traditional approach pipeline consists of image acquisition and reconstruction, possible image postprocessing, volume of interest (VOI) segmentation, and predefined feature extraction. The extracted features quantitatively describe shape, intensity, and texture. Geometric or morphological (shape) features outline lesion characteristics such as size, intensity features quantify tracer uptake such as maximum, peak, or mean SUV and texture features capture spatial heterogeneity in voxel values. These features are analyzed in statistical analysis or machine learning models that are trained for patient classification. In the handcrafted radiomics pipeline example shown in Fig. 6 A), a VOI is manually or semi-automatically defined, then, features are extracted and selected and are later introduced in a statistical or machine learning model. On the other hand, the deep learning radiomics pipeline 6 B) does not require VOI delineation since the architecture automatically processes the images in their raw form extracting complex features to perform feature selection and classification.

pathological state with features from CT, PET, fused PET/CT, and combined CT and PET features, demonstrating the accuracy of PET/CT radiomics for the preoperative prediction of clinical and pathological stages for patients with esophageal cancer. Radiomics was studied by Oikonomou. et. al. [76] for the prediction of response to stereotactic radiotherapy for lung cancer. In this study, principal component analysis was performed on homogeneity, size, maximum intensity, mean and median gray level, standard deviation, entropy, kurtosis, skewness, morphology, and asymmetry among others for CT and PET-derived features to determine which ones accounted for most of the variability to predict overall survival.

A hybrid model including many-objective radiomics (MaO-radiomics) and a 3D CNN was proposed by Chen et al. [77] for the differentiation of malignant from suspicious non-malignant lymph nodes in patients with head and neck cancer. In this approach, the outputs of both models were fused through an evidential reasoning approach, obtaining higher accuracy than with radiomics alone. The handcrafted selected set of features is finite and might not entirely represent the optimal quantification approach for the desired task, but deep learning algorithms are based on the raw data, where the extracted features by the hidden layers are more abstract and unconstrained, giving a more generalizable model, however, larger datasets are required. CNNs eliminate of the time-consuming lesion identification and annotation step but do not necessarily outperform handcrafted radiomics in combination with classification machine learning. However, they are generally preferred methods by researchers and radiologists because it is more user-friendly, required less data handling, and suffer less selection bias, as demonstrated in the study by Wang et al. [78] for the classification of mediastinal lymph nodes of NSCLC. In the specific field of nuclear imaging, DL applications on radiomics are infrequent, but there are some examples of studies comparing CNNs with handcrafted radiomic analyses. In a study performed by Ypsilantis et al. [79] a CNN trained directly from PET scans was used to predict response to neoadjuvant chemotherapy in oesophageal cancer, and it was compared to an approach based on statistical classifiers using over 100 quantitative imaging descriptors, including texture features and SUVs. Results showed that for this application CNNs outperformed the handcrafted radiomic approach.

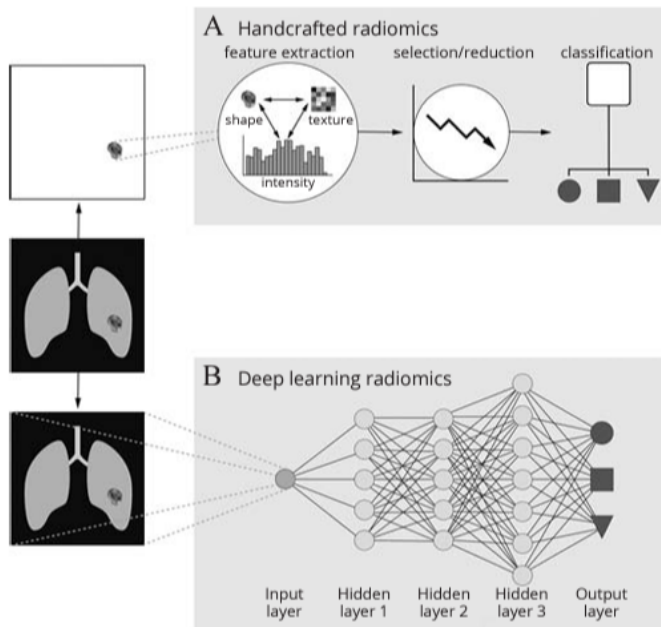


Fig. 6: A) Handcrafted and B) deep learning radiomic pipeline [74].

In the field of oncology, radiomic shape, intensity, and texture features can be computed to quantitatively analyze tracer uptake, as well as the stratification of histology, tumor grades or stages, and clinical outcome [74]. A recent feasibility study by Lei et. al. [75] has used logistic regression to classify the T stage, lymph node metastasis, and

While most applications in the literature mentioned in this paper focus on oncological developments, radiomics has also been studied in other medical fields. Non-oncology applications of handcrafted radiomics can mostly be found in SPECT and some applications include the prediction of coronary artery calcification in [99mTc]-sestamibi SPECT myocardial perfusion scans [80] and outcome prediction in Parkinson's disease by evaluating the extraction of radiomic features from longitudinal dopamine transporter SPECT [81].

III WHAT CAN WE EXPECT IN CLINIC?

In recent years, Artificial Intelligence methods have been exponentially studied in the field of nuclear imaging, making significant strides in modern healthcare. Some of these methods aim to improve and increase the accuracy of the outcomes from the current used methods, reducing human intervention if possible, while others enable procedures that were previously impossible. As AI-based solutions continue to evolve, their translation into clinical practice is evaluated in parallel. The implementation and commercialization of the explored techniques are more likely to happen for those that replace the previously existing methods since they would not completely change the current frameworks: less standardization, protocol redefinition, and staff retraining are required.

The integration of machine learning algorithms, particularly CNNs, in PET detectors, is promising for improving the timing and spatial resolution, essential for image quality in nuclear imaging. The advancement of these technologies is possible thanks to the development of fast and affordable waveform digitizers, which are expected to be incorporated into PET scanners' system electronics, particularly in front-end electronics for better position and timing estimation. Studies envision the implementation of machine learning algorithms that can simultaneously estimate position, energy, and time-of-interaction from photodetector waveforms, potentially giving a better performance, due to the physical dependencies of these parameters. Furthermore, there is a proposal to adapt CNNs to include waveforms from multiple photodetectors, for the estimation of not only timing but also position-of-interaction within the detectors. CNNs and GANs are the biggest technological trends in the context of image reconstruction and post-processing, aiming for "ultra low-dose" PET imaging, by allowing high quality image generations from minimal tracer activity and possibly eliminating the need for simultaneous CT scans, which increases the dose. For these algorithms to be implemented in clinical practice, their accuracy needs to be proven in a clinical context by performing large-scale studies. Apart from increasing the number of scanners, a reduction in the image acquisition time could be key for increasing the number of examinations per scanner and thus the availability of data for large scale studies.

Despite their potential to increase accuracy and save clinicians from performing arduous tasks, deep learning approaches are used for very specific applications in research, however, they are not that common in commercialized products because of the lack of regulations for which these models can be trustfully used. Guidelines have been recently proposed by the European Association of Nuclear Medicine (EANM) [82] for image acquisition, reconstruction to generate more homogenic quantification of tracer uptake

and also for radiomic analysis so that obtained results are more reliable, repeatable, and reproducible. Another of the big limitations is the difficulty of standardizing the data storage, which affects also the integration of information coming from different modalities and institutes and makes it very hard to homogenize the processing pipelines. Data should be collected according to the FAIR Guiding Principles, created for scientific data management, based on Findable, Accessible, Interoperable, and Reusable data. Although models can become more generalizable by adding more data into the training stage, this comes together with a loss of interpretability, which is an obstacle, especially in the health domain. Clinicians are sometimes unwilling to implement certain radiomic algorithms whose capabilities they do not fully understand since the vast amount of parameters in a deep learning model are not directly interpretable. Therefore, there is a need to provide the end-users with an explanation for a decision-making algorithm. Explainable AI (XAI) techniques have been lately explored for this purpose. In the context of nuclear imaging, one of the most useful applications would be the generation of heat maps that highlight the parts of the image that are the most relevant for making a certain decision. For example, for the prediction of survival of cancer patients; if an XAI model is capable of indicating the voxels that affect prognosis the most, that would give researchers a better understanding of the disease and could help build more targeted therapies with better survival expectations. Explainable AI methods are still far from being widely used in clinical applications, but they hold promise for solving the problem of lack of interpretability and trust in deep models. Finally, education is key for the transmission of knowledge to next-generation scientists and clinicians to make become familiar and grow-up with AI and therefore enhance further its potential adoption and development.

In summary, there are still many challenges in the implementation of AI-based solutions in nuclear imaging, including also ethical and legal aspects. The techniques that hold promise for being shortly present in clinics are the ones whose frameworks are already established, as well as the ones that clinicians are willing to adopt. If data is collected and managed correctly with the proposed guidelines for standardization, and the algorithms become more robust and transparent for end-users, the implementation of AI algorithms in the everyday nuclear imaging pipelines will happen sooner.

IV CONCLUSION

The integration of Artificial Intelligence methods into the field of nuclear medicine presents significant opportunities for revolutionizing healthcare by providing deeper insights into physiological processes at the molecular level. Despite historical challenges, including the complexity of medical data, diverse imaging modalities, and concerns about AI precision and trust, recent technological advancements and research have demonstrated the potential of AI as

a powerful and trustworthy tool in this field. Recent advancements in AI methods, particularly in PET detectors and image reconstruction, aim to improve image quality and reduce radiation exposure. CNNs and Generative Adversarial Networks (GANs) show promise in achieving "ultra low-dose" PET imaging, potentially eliminating the need for simultaneous CT scans. However, their clinical implementation requires rigorous validation through large-scale studies. For image analysis, CNNs have mostly been reviewed as an approach to perform image segmentation in different health fields, to reduce variability in the results and save clinicians from performing arduous and time-consuming tasks. Radiomics, a rapidly growing field of quantitative image analysis, offers non-invasive biomarkers for tumor characterization, prognostic classification, and treatment response prediction. Traditional handcrafted radiomics and deep learning radiomics present distinct approaches, each with its advantages and challenges. While traditional methods offer interpretability, deep learning algorithms leverage raw data for more abstract features, potentially leading to more generalizable models. However, larger datasets and efforts in data standardization are required to maximize their potential in clinical settings.

Challenges in implementing AI-based solutions in nuclear imaging include regulatory concerns, data standardization, interoperability, and the interpretability of deep learning models. Explainable AI techniques hold promise in addressing the lack of interpretability and trust in deep models, potentially facilitating their adoption in clinical practice. Despite these challenges, the translation of AI algorithms into everyday nuclear imaging pipelines is expected to occur as frameworks become established, and algorithms become more robust, transparent, and aligned with clinicians' needs. The adoption of the proposed guidelines for data standardization and the ongoing efforts in research, together with the development of educational programs for scientists and clinicians, are crucial for accelerating this transition and exploiting the full potential of AI in nuclear imaging.

Bibliography

- [1] S Kevin Zhou et al. "A review of deep learning in medical imaging: Imaging traits, technology trends, case studies with progress highlights, and future promises". In: *Proceedings of the IEEE* 109.5 (2021), pp. 820–838.
- [2] Vivek Kaul, Sarah Enslin, and Seth A Gross. "History of artificial intelligence in medicine". In: *Gastrointestinal endoscopy* 92.4 (2020), pp. 807–812.
- [3] Takaaki Yoshimura et al. "Medical Radiation Exposure Reduction in PET via Super-Resolution Deep Learning Model". In: *Diagnostics* 12.4 (2022).
- [4] Reabal Najjar. "Redefining radiology: a review of artificial intelligence integration in medical imaging". In: *Diagnostics* 13.17 (2023), p. 2760.
- [5] Ryuji Hamamoto et al. "Application of artificial intelligence technology in oncology: Towards the establishment of precision medicine". In: *Cancers* 12.12 (2020), p. 3532.
- [6] Anil K Philip et al. "Artificial Intelligence and Precision Medicine: A New Frontier for the Treatment of Brain Tumors". In: *Life* 13.1 (2022), p. 24.
- [7] Daniel S Kermany et al. "Identifying medical diagnoses and treatable diseases by image-based deep learning". In: *cell* 172.5 (2018), pp. 1122–1131.
- [8] Ahmad Waleed Salehi et al. "A Study of CNN and Transfer Learning in Medical Imaging: Advantages, Challenges, Future Scope". In: *Sustainability* 15.7 (2023), p. 5930.
- [9] Tyler J Bradshaw et al. "Nuclear medicine and artificial intelligence: best practices for algorithm development". In: *Journal of Nuclear Medicine* 63.4 (2022), pp. 500–510.
- [10] H Malcolm Hudson and Richard S Larkin. "Accelerated image reconstruction using ordered subsets of projection data". In: *IEEE transactions on medical imaging* 13.4 (1994), pp. 601–609.
- [11] Jerrold T Bushberg and John M Boone. *The essential physics of medical imaging*. Lippincott Williams & Wilkins, 2011.
- [12] Kuang Gong et al. "Machine learning in PET: from photon detection to quantitative image reconstruction". In: *Proceedings of the IEEE* 108.1 (2019), pp. 51–68.
- [13] Georges N El Fakhri et al. "Respective roles of scatter, attenuation, depth-dependent collimator response and finite spatial resolution in cardiac single-photon emission tomography quantitation: a Monte Carlo study". In: *European journal of nuclear medicine* 26 (1999), pp. 437–446.
- [14] Philipp Ritt. "Recent developments in SPECT/CT". In: *Seminars in Nuclear Medicine*. Vol. 52. 3. Elsevier, 2022, pp. 276–285.
- [15] Eiji Yoshida et al. "Inter-crystal scatter identification for a depth-sensitive detector using support vector machine for small animal positron emission tomography". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 571.1-2 (2007), pp. 243–246.
- [16] Jean-Baptiste Michaud et al. "Sensitivity increase through a neural network method for LOR recovery of ICS triple coincidences in high-resolution pixelated-detectors PET scanners". In: *IEEE Transactions on Nuclear Science* 62.1 (2014), pp. 82–94.
- [17] Dennis R Schaart et al. "A novel, SiPM-array-based, monolithic scintillator detector for PET". In: *Physics in Medicine & Biology* 54.11 (2009), p. 3501.
- [18] Peter Bruyndonckx et al. "Comparison of nonlinear position estimators for continuous scintillator detectors in PET". In: *2006 IEEE Nuclear Science Symposium Conference Record*. Vol. 4. IEEE, 2006, pp. 2518–2522.

- [19] Peter Bruyndonckx et al. “Evaluation of machine learning algorithms for localization of photons in undivided scintillator blocks for PET detectors”. In: *IEEE Transactions on Nuclear Science* 55.3 (2008), pp. 918–924.
- [20] Hossein Arabi et al. “The promise of artificial intelligence and deep learning in PET and SPECT imaging”. In: *Physica Medica* 83 (2021), pp. 122–137.
- [21] Florian Müller et al. “Gradient tree boosting-based positioning method for monolithic scintillator crystals in positron emission tomography”. In: *IEEE Transactions on Radiation and Plasma Medical Sciences* 2.5 (2018), pp. 411–421.
- [22] Florian Müller et al. “A novel DOI positioning algorithm for monolithic scintillator crystals in PET based on gradient tree boosting”. In: *IEEE Transactions on Radiation and Plasma Medical Sciences* 3.4 (2018), pp. 465–474.
- [23] Peng Peng et al. “Compton PET: a simulation study for a PET module with novel geometry and machine learning for position decoding”. In: *Biomedical physics & engineering express* 5.1 (2018), p. 015018.
- [24] Radosław Marcinkowski et al. “Sub-millimetre DOI detector based on monolithic LYSO and digital SiPM for a dedicated small-animal PET system”. In: *Physics in Medicine & Biology* 61.5 (2016), p. 2196.
- [25] Eric Berg and Simon R Cherry. “Using convolutional neural networks to estimate time-of-flight from PET detector waveforms”. In: *Physics in Medicine & Biology* 63.2 (2018), 02LT01.
- [26] Robert Seifert et al. “Artificial intelligence and machine learning in nuclear medicine: future perspectives”. In: *Seminars in nuclear medicine*. Vol. 51. 2. Elsevier. 2021, pp. 170–177.
- [27] David Izquierdo-Garcia et al. “An SPM8-based approach for attenuation correction combining segmentation and nonrigid template formation: application to simultaneous PET/MR brain imaging”. In: *Journal of Nuclear Medicine* 55.11 (2014), pp. 1825–1830.
- [28] Imene Mecheter et al. “MR image-based attenuation correction of brain PET imaging: review of literature on machine learning approaches for segmentation”. In: *Journal of Digital Imaging* 33 (2020), pp. 1224–1241.
- [29] Fang Liu et al. “Deep learning MR imaging-based attenuation correction for PET/MR imaging”. In: *Radiology* 286.2 (2018), pp. 676–684.
- [30] Xiao Han. “MR-based synthetic CT generation using a deep convolutional neural network method”. In: *Medical physics* 44.4 (2017), pp. 1408–1419.
- [31] Sanaz Kaviani et al. “Image reconstruction using UNET-transformer network for fast and low-dose PET scans”. In: *Computerized Medical Imaging and Graphics* 110 (2023), p. 102315.
- [32] Hossein Arabi et al. “Novel adversarial semantic structure deep learning for MRI-guided attenuation correction in brain PET/MRI”. In: *European journal of nuclear medicine and molecular imaging* 46 (2019), pp. 2746–2759.
- [33] Jelmer M Wolterink et al. “Deep MR to CT synthesis using unpaired data”. In: *Simulation and Synthesis in Medical Imaging: Second International Workshop, SASHIMI 2017, Held in Conjunction with MICCAI 2017, Québec City, QC, Canada, September 10, 2017, Proceedings 2*. Springer. 2017, pp. 14–23.
- [34] Helena Van Hemmen et al. *A deep learning-based approach for direct whole-body PET attenuation correction*. 2019.
- [35] Jaewon Yang et al. “Joint correction of attenuation and scatter in image space using deep convolutional neural networks for dedicated brain 18F-FDG PET”. In: *Physics in medicine & biology* 64.7 (2019), p. 075019.
- [36] Ida Häggström et al. “DeepPET: A deep encoder-decoder network for directly solving the PET image reconstruction inverse problem”. In: *Medical image analysis* 54 (2019), pp. 253–262.
- [37] Zhanli Hu et al. “DPIR-Net: Direct PET image reconstruction based on the Wasserstein generative adversarial network”. In: *IEEE Transactions on Radiation and Plasma Medical Sciences* 5.1 (2020), pp. 35–43.
- [38] Wenyi Shao, Steven P Rowe, and Yong Du. “SPECT-net: a deep learning neural network for SPECT image reconstruction”. In: *Annals of Translational Medicine* 9.9 (2021).
- [39] Junshen Xu et al. “200x low-dose PET reconstruction using deep learning”. In: *arXiv preprint arXiv:1712.04119* (2017).
- [40] Isaac Shiri et al. “Standard SPECT myocardial perfusion estimation from half-time acquisitions using deep convolutional residual neural networks”. In: *Journal of Nuclear Cardiology* 28.6 (2021), pp. 2761–2779.
- [41] Wenzhuo Lu et al. “An investigation of quantitative accuracy for deep learning based denoising in oncological PET”. In: *Physics in Medicine & Biology* 64.16 (2019), p. 165019.
- [42] Kye Young Jeong et al. “Sinogram-based super-resolution in PET”. In: *Physics in Medicine & Biology* 56.15 (2011), p. 4881.
- [43] Amirhossein Sanaat et al. “Projection space implementation of deep learning-guided low-dose brain PET imaging improves performance over implementation in image space”. In: *Journal of Nuclear Medicine* 61.9 (2020), pp. 1388–1396.
- [44] Jianan Cui et al. “PET image denoising using unsupervised deep learning”. In: *European journal of nuclear medicine and molecular imaging* 46 (2019), pp. 2780–2789.
- [45] Victor Lempitsky, Andrea Vedaldi, and Dmitry Ulyanov. “Deep image prior”. In: *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. IEEE. 2018, pp. 9446–9454.
- [46] Jianan Cui et al. “Populational and individual information based PET image denoising using conditional

- unsupervised learning”. In: *Physics in Medicine & Biology* 66.15 (2021), p. 155001.
- [47] Cameron Dennis Pain, Gary F Egan, and Zhaolin Chen. “Deep learning-based image reconstruction and post-processing methods in positron emission tomography for low-dose imaging and resolution enhancement”. In: *European Journal of Nuclear Medicine and Molecular Imaging* 49.9 (2022), pp. 3098–3118.
- [48] Kevin T Chen et al. “Ultra-low-dose 18F-florbetaben amyloid PET imaging using deep learning with multi-contrast MRI inputs”. In: *Radiology* 290.3 (2019), pp. 649–656.
- [49] Kevin T Chen et al. “Generalization of deep learning models for ultra-low-count amyloid PET/MRI using transfer learning”. In: *European journal of nuclear medicine and molecular imaging* 47 (2020), pp. 2998–3007.
- [50] Paul Suetens et al. “Image segmentation: methods and applications in diagnostic radiology and nuclear medicine”. In: *European journal of radiology* 17.1 (1993), pp. 14–21.
- [51] Paul Blanc-Durand et al. “Automatic lesion detection and segmentation of 18F-FET PET in gliomas: a full 3D U-Net convolutional neural network study”. In: *PLoS One* 13.4 (2018), e0195798.
- [52] Mathieu Hatt et al. “The first MICCAI challenge on PET tumor segmentation”. In: *Medical image analysis* 44 (2018), pp. 177–195.
- [53] Abir Baazaoui et al. “A survey of PET image segmentation: applications in oncology, cardiology and neurology”. In: *Current Medical Imaging* 12.1 (2016), pp. 13–27.
- [54] Isaac Shiri et al. “Decentralized distributed multi-institutional pet image segmentation using a federated deep learning framework”. In: *Clinical Nuclear Medicine* 47.7 (2022), pp. 606–617.
- [55] Koujiro Ikushima et al. “Computer-assisted framework for machine-learning-based delineation of GTV regions on datasets of planning CT and PET/CT images”. In: *Journal of radiation research* 58.1 (2017), pp. 123–134.
- [56] Zisha Zhong et al. “Simultaneous cosegmentation of tumors in PET-CT images using deep fully convolutional networks”. In: *Medical physics* 46.2 (2019), pp. 619–633.
- [57] Mohammad S Sadaghiani, Steven P Rowe, and Sara Sheikhabaehi. “Applications of artificial intelligence in oncologic 18F-FDG PET/CT imaging: a systematic review”. In: *Annals of Translational Medicine* 9.9 (2021).
- [58] Bin Huang et al. “Fully automated delineation of gross tumor volume for head and neck cancer on PET-CT using deep learning: a dual-center study”. In: *Contrast media & molecular imaging* 2018 (2018).
- [59] Zhe Guo et al. “Gross tumor volume segmentation for head and neck cancer radiotherapy using deep dense multi-modality network”. In: *Physics in Medicine & Biology* 64.20 (2019), p. 205015.
- [60] Vincent Andrearczyk et al. “Automatic head and neck tumor segmentation and outcome prediction relying on FDG-PET/CT images: findings from the second edition of the HECKTOR challenge”. In: *Medical Image Analysis* 90 (2023), p. 102972.
- [61] Fabian Isensee et al. “nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation”. In: *Nature methods* 18.2 (2021), pp. 203–211.
- [62] Andrei Iantsen, Dimitris Visvikis, and Mathieu Hatt. “Squeeze-and-excitation normalization for automated delineation of head and neck primary tumors in combined PET and CT images”. In: *Head and Neck Tumor Segmentation: First Challenge, HECKTOR 2020, Held in Conjunction with MICCAI 2020, Lima, Peru, October 4, 2020, Proceedings 1*. Springer. 2021, pp. 37–43.
- [63] Gustavo Andrade-Miranda et al. “Multi-modal medical Transformers: A meta-analysis for medical image segmentation in oncology”. In: *Computerized Medical Imaging and Graphics* 110 (2023), p. 102308.
- [64] Tonghe Wang et al. “A learning-based automatic segmentation and quantification method on left ventricle in gated myocardial perfusion SPECT imaging: A feasibility study”. In: *Journal of Nuclear Cardiology* 27 (2020), pp. 976–987.
- [65] Yangmei Zhang et al. “An automatic segmentation method with self-attention mechanism on left ventricle in gated PET/CT myocardial perfusion imaging”. In: *Computer Methods and Programs in Biomedicine* 229 (2023), p. 107267.
- [66] Reza Piri et al. ““Global” cardiac atherosclerotic burden assessed by artificial intelligence-based versus manual segmentation in 18F-sodium fluoride PET/CT scans: Head-to-head comparison”. In: *Journal of Nuclear Cardiology* 29.5 (2022), pp. 2531–2539.
- [67] Kathryn C Elkins et al. “Utility of gray-matter segmentation of ictal-Interictal perfusion SPECT and interictal 18F-FDG-PET in medically refractory epilepsy”. In: *Epilepsy research* 130 (2017), pp. 93–100.
- [68] Teng Zhang et al. “Spatial normalization and quantification approaches of PET imaging for neurological disorders”. In: *European Journal of Nuclear Medicine and Molecular Imaging* 49.11 (2022), pp. 3809–3829.
- [69] Kyeong Taek Oh et al. “Semantic segmentation of white matter in FDG-PET using generative adversarial network”. In: *Journal of digital imaging* 33 (2020), pp. 816–825.
- [70] Kyeong Taek Oh et al. “Segmentation of white matter hyperintensities on 18 F-FDG PET/CT images with a generative adversarial network”. In: *European Journal of Nuclear Medicine and Molecular Imaging* 48 (2021), pp. 3422–3431.
- [71] Elena Pasini et al. “Convolution Neural Networks for the Automatic Segmentation of 18F-FDG PET Brain as an Aid to Alzheimer’s Disease Diagnosis”. In: *Electronics* 11.14 (2022), p. 2260.

- [72] Philippe Lambin et al. “Radiomics: extracting more information from medical images using advanced feature analysis”. In: *European journal of cancer* 48.4 (2012), pp. 441–446.
- [73] Michele Avanzo et al. “Machine and deep learning methods for radiomics”. In: *Medical physics* 47.5 (2020), e185–e202.
- [74] Wyanne A Noortman et al. “Nuclear medicine radiomics in precision medicine: Why we can’t do without artificial intelligence”. In: *QJ Nucl. Med. Mol. Imaging* 64 (2020), pp. 278–290.
- [75] Xiyao Lei et al. “Preoperative prediction of clinical and pathological stages for patients with esophageal cancer using PET/CT radiomics”. In: *Insights into Imaging* 14.1 (2023), p. 174.
- [76] Anastasia Oikonomou et al. “Radiomics analysis at PET/CT contributes to prognosis of recurrence and survival in lung cancer treated with stereotactic body radiotherapy”. In: *Scientific reports* 8.1 (2018), p. 4003.
- [77] Liyuan Chen et al. “Combining many-objective radiomics and 3D convolutional neural network through evidential reasoning to predict lymph node metastasis in head and neck cancer”. In: *Physics in Medicine & Biology* 64.7 (2019), p. 075011.
- [78] Hongkai Wang et al. “Comparison of machine learning methods for classifying mediastinal lymph node metastasis of non-small cell lung cancer from 18 F-FDG PET/CT images”. In: *EJNMMI research* 7 (2017), pp. 1–11.
- [79] Xenia Fave et al. “Delta-radiomics features for the prediction of patient outcomes in non-small cell lung cancer”. In: *Scientific reports* 7.1 (2017), p. 588.
- [80] Saeed Ashrafinia et al. “Radiomics analysis of clinical myocardial perfusion stress SPECT images to identify coronary artery calcification”. In: *medRxiv* (2021), pp. 2021–01.
- [81] Arman Rahmim et al. “Improved prediction of outcome in Parkinson’s disease using radiomics analysis of longitudinal DAT SPECT images”. In: *NeuroImage: Clinical* 16 (2017), pp. 539–544.
- [82] Roberto C Delgado Bolton et al. *The QuADRANT study: current status and recommendations for improving uptake and implementation of clinical audit of medical radiological procedures in Europe—the nuclear medicine perspective*. 2023.
- [83] Antonia Creswell et al. “Generative adversarial networks: An overview”. In: *IEEE signal processing magazine* 35.1 (2018), pp. 53–65.
- [84] Li Tao et al. “Pseudo CT image synthesis and bone segmentation from MR images using adversarial networks with residual blocks for MR-based attenuation correction of brain PET data”. In: *IEEE Transactions on Radiation and Plasma Medical Sciences* 5.2 (2020), pp. 193–201.

Appendix

Theoretical concepts of some Artificial Intelligence methods

Artificial Intelligence covers all machine performances that are usually done by humans, and the ones that require automated learning are classified as Machine Learning. Figure A.1 shows the domains Artificial Intelligence.

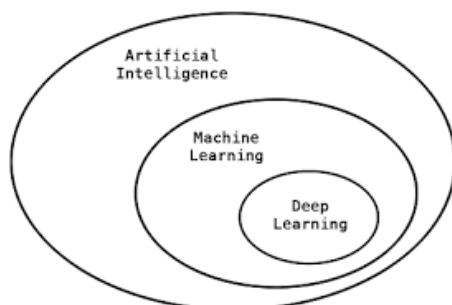


Fig. A.1: Domains in Artificial Intelligence. Deep Learning is a subset of Machine Learning that is a subset of Artificial Intelligence.

Unlike classical programming, which uses data and several ordered rules to get answers, **Machine Learning** makes use of data and answers to generate rules and then uses those rules to generate new answers. This can be very useful in problems where the specifications for generating an output are not very clear. The machine is trained rather than explicitly programmed. **Deep Learning** is the subset of methods in Machine Learning that are composed of neural networks with a large number of layers and parameters.

Figure A.2 shows a diagram of a perceptron which is the fundamental computational block for neural networks, and is based on the comparison of a biological neuron with an artificial one. Given n inputs x_1, \dots, x_n , the nucleus of the perceptron performs a weighted sum along with a bias term, which shifts the decision boundary away from the origin and does not depend on any input value.

$$z(\mathbf{x}, \mathbf{w}, b) = b + \mathbf{w} \cdot \mathbf{x} \quad (1)$$

Where $\mathbf{w} \cdot \mathbf{x}$ represents the dot product $\sum_{i=1}^n x_i w_i$. Additionally, activation functions are applied in each artificial neuron to generate non-linearities between the layers.

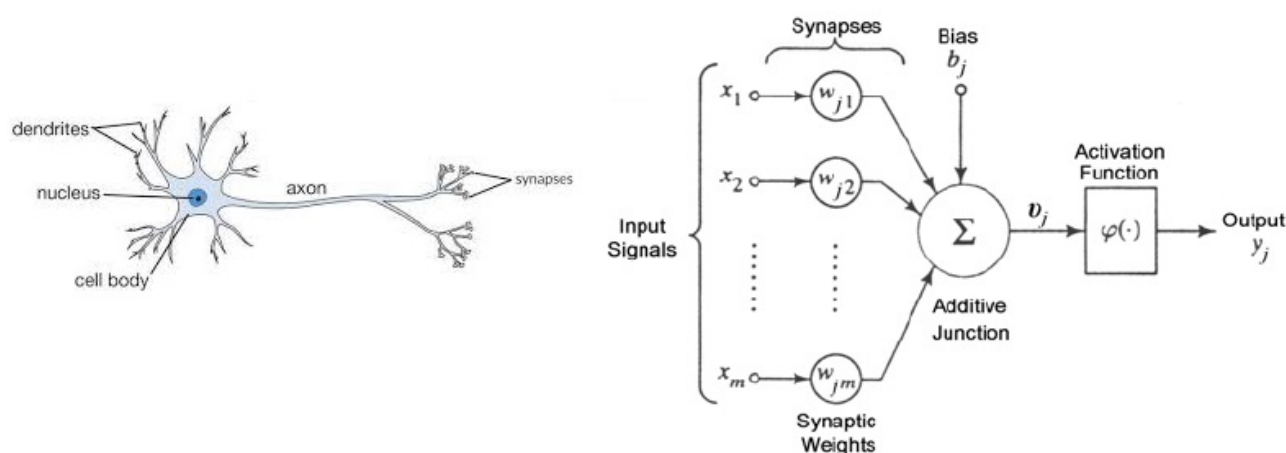


Fig. A.2: Simplified diagram of a biological neuron (left): the information enters through the dendrites, is processed in the nucleus and outputted through the axon. An artificial neuron (right). In this analogy, the biological units are replaced with mathematical ones

However, a single perceptron itself is not capable of finding complex patterns. Fortunately, we can expand this idea to create a multilayer perceptron model, commonly known as a basic **Artificial Neural Networks (ANNs)**. For this, several vertical

layers of these neurons are considered, each one taking an input (x) and using their outputs (y) as the input of the next layer of perceptrons, as indicated in Fig A.3:

$$x^l = y^{l-1} \tag{2}$$

When every neuron in a layer is connected to every neuron in the next layer, it is known as a fully connected layer. When the ANNs are composed of several hidden layers, they are considered **Deep Neural Networks**.

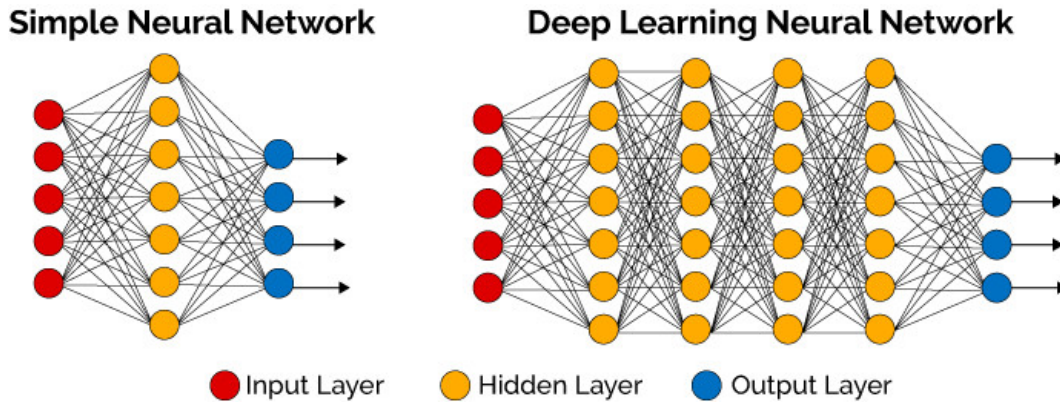


Fig. A.3: Simple Neural Network (left) and Deep Neural Network (right). It is precisely the number of hidden layers what provides the "deep" sense in Deep Learning.

Networks receive a set of input data along with the corresponding correct outputs (Ground-Truth). The algorithm learns by comparing its actual output with the correct output to find errors, and modifies the model accordingly by adjusting the weights and biases values in the network. The function that evaluates the error of the score is called *loss function*. The process of learning is precisely the process of finding the right weights and biases that minimize the defined loss function, and Figure A.4 shows a diagram of such process.

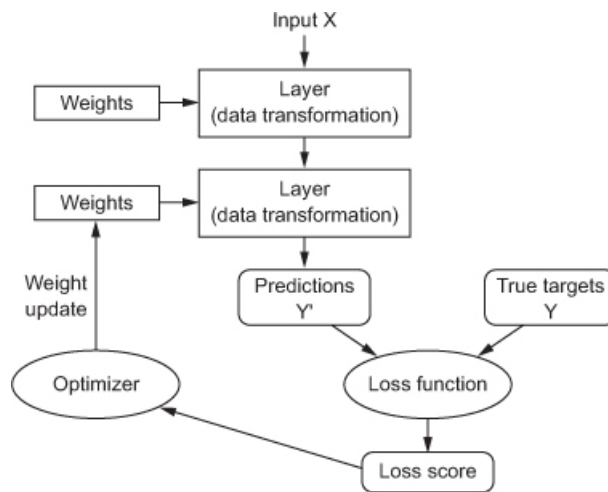


Fig. A.4: Diagram of the learning process of a Neural Network

Deep Learning based Neural Networks have been mentioned in this literature review for their wide use in medical imaging regarding classification or regression tasks. Using fully connected neural networks when working with image data (especially medical image data that is usually 3D) is unpractical due to the huge amount of weights it would have to handle. However, **Convolutional Neural Networks (CNNs)** are specifically designed to work on images, and their fundamental advantage is that they learn local patterns in their input feature space while densely connected layers learn global patterns. These networks constitute a blank model of millions of parameters that can be trained to recognize complex patterns and extract incredibly abstract features from the input to perform image classification, regression or segmentation tasks. CNNs use convolutional layers constituted by *kernels* and works by sliding in steps of the *stride* size the kernels and performing the *convolution operation*:

$$y_j = g(b + K_j \otimes x) \tag{3}$$

⊗ denoting the convolution operation. U-Nets, as shown in Figure A.5 are type of CNNs that are constituted con convolution and deconvolution layers and are used for generating outputs that are also images such as the example of DeepPET mentioned in 3.

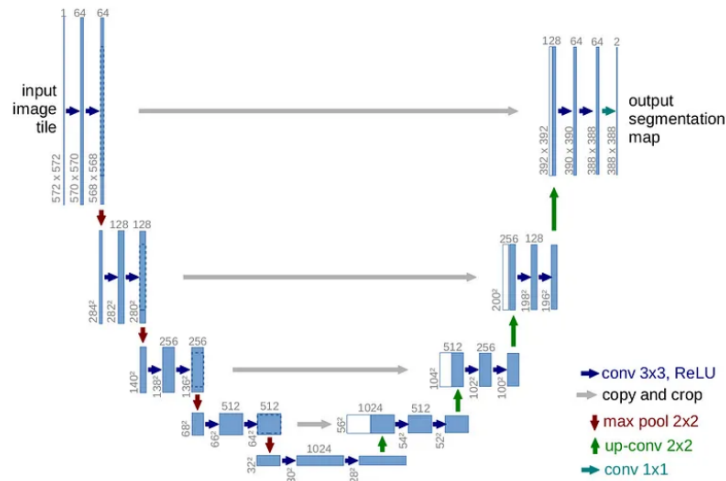


Fig. A.5: U-Net architecture

Generative Adversarial Networks (GANs)

GANs are a type of artificial intelligence model that consists of two neural networks, a generator and a discriminator, trained simultaneously through adversarial training using pairs of images. They have mostly been mentioned in this paper for the generation of pseudo-CTs using MRI as input in the process of image reconstruction, where the CT information is needed to create attenuation maps. The MR images serve as input to the generator, and the corresponding real CT images act as the ground truth during training. The discriminator learns to differentiate between real CT images and synthetic ones generated by the generator, and the generator learn to create pseudo-CT images from MRI [83].

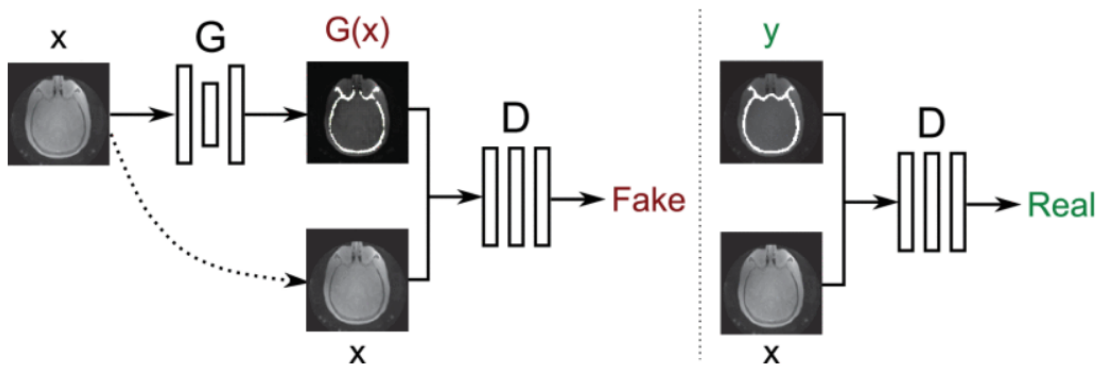


Fig. A.6: Illustration for a conditional GAN (cGAN). “G” represents the generator, “D” represents the discriminator, “x” represents the input, “y” represents the label, and “G(x)” represents the output from the generator [84].