



# Applications of machine learning and deep learning in SPECT and PET imaging: General overview, challenges and future prospects

Carmen Jimenez-Mesa<sup>a</sup>, Juan E. Arco<sup>a,b</sup>, Francisco Jesus Martinez-Murcia<sup>a</sup>, John Suckling<sup>c</sup>,  
Javier Ramirez<sup>a</sup>, Juan Manuel Gorriz<sup>a,c,\*</sup>

<sup>a</sup> Department of Signal Theory, Networking and Communications, University of Granada, 18010, Spain

<sup>b</sup> Department of Communications Engineering, University of Malaga, 29010, Spain

<sup>c</sup> Department of Psychiatry, University of Cambridge, Cambridge CB21TN, UK

## ARTICLE INFO

### Keywords:

Deep Learning  
Diagnosis  
Machine Learning  
PET  
Sample size  
SPECT

## ABSTRACT

The integration of positron emission tomography (PET) and single-photon emission computed tomography (SPECT) imaging techniques with machine learning (ML) algorithms, including deep learning (DL) models, is a promising approach. This integration enhances the precision and efficiency of current diagnostic and treatment strategies while offering invaluable insights into disease mechanisms. In this comprehensive review, we delve into the transformative impact of ML and DL in this domain. Firstly, a brief analysis is provided of how these algorithms have evolved and which are the most widely applied in this domain. Their different potential applications in nuclear imaging are then discussed, such as optimization of image acquisition or reconstruction, biomarkers identification, multimodal fusion and the development of diagnostic, prognostic, and disease progression evaluation systems. This is because they are able to analyse complex patterns and relationships within imaging data, as well as extracting quantitative and objective measures. Furthermore, we discuss the challenges in implementation, such as data standardization and limited sample sizes, and explore the clinical opportunities and future horizons, including data augmentation and explainable AI. Together, these factors are propelling the continuous advancement of more robust, transparent, and reliable systems.

## 1. Introduction

In recent years, the integration of machine learning (ML), a branch of Artificial Intelligence (AI) that includes deep learning (DL), with medical imaging techniques has emerged as a promising approach to enhance the accuracy and effectiveness of diagnostic and prognostic strategies [1, 2]. Among these imaging modalities, Single-Photon Emission Computed Tomography (SPECT) and Positron Emission Tomography (PET) play a crucial role in providing valuable insights into various diseases by capturing functional and molecular information [3,4].

The integration of ML and DL algorithms with SPECT and PET imaging offers several advantages [5–7]. Firstly, these algorithms can aid in automating the analysis of large volumes of imaging data, allowing for more efficient and objective interpretation [8,9]. Secondly, ML and DL algorithms can learn complex patterns and relationships within the imaging data, enabling the identification of disease-specific biomarkers and the development of predictive models for early detection and prognosis [10–12]. Furthermore, these algorithms hold the promise of

enhancing image reconstruction, reducing noise, and mitigating artifacts, thereby elevating both the overall image quality and its interpretability [13–15]. Specifically, the incorporation of ML algorithms into SPECT/PET imaging systems offers compelling opportunities. A major advantage is the potential to increase diagnostic accuracy, allowing for more efficient and accurate patient assessments while minimizing the required doses of radiopharmaceuticals (radiopharmaceutical optimization). Additionally, ML models can advance personalized medicine by leveraging a multitude of patient-specific biomarkers, enabling customized treatment strategies and predictive treatment outcome assessments, ultimately enhancing the overall quality of life for patients.

When these algorithms are applied in the medical field, they are usually referred to as Computer-Aided Diagnosis (CAD) systems. CAD systems have been developed to assist clinicians by providing computerized tools with pattern recognition and prediction capabilities [16–19]. It should be noted that CAD systems are not exclusively confined to diagnosis; they are also employed at various stages of the

\* Corresponding author at: Department of Signal Theory, Networking and Communications, University of Granada, 18010, Spain.

E-mail address: [gorriz@ugr.es](mailto:gorriz@ugr.es) (J.M. Gorriz).

<https://doi.org/10.1016/j.phrs.2023.106984>

Received 24 July 2023; Received in revised form 4 October 2023; Accepted 4 November 2023

Available online 7 November 2023

1043-6618/© 2023 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC license (<http://creativecommons.org/licenses/by-nc/4.0/>).

medical imaging process, which may encompass preprocessing, reconstruction, or image segmentation, among others. ML techniques applied in these systems range from linear and low-dimensional schemes to more advanced deep neural network (NN) architectures that facilitate feature selection, extraction, and classification [20,21]. These techniques enhance accuracy and recognition rates when analyzing complex patterns in high-dimensional contexts, even in the presence of subgroups within different conditions [22–24]. DL, particularly NN models, has gained popularity in the medical field, but their increased complexity does not always translate into noticeable improvements in performance. Instead, it often leads to higher computational costs and reduced model interpretability. This poses challenges, including clinicians' reluctance to adopt methods lacking physical interpretation, diminishing their capacity for understanding, and concerns regarding the opacity and learning capacity of the algorithms [25,26]. To establish CAD systems as a standard for clinical support, it is crucial to address these issues along with other challenges that exist, such as the limited availability of labeled and annotated datasets, standardization of imaging protocols, and the need for validation and regulatory approval.

To address these challenges, continuous improvement of existing algorithms is imperative. Efforts are being made to enhance the interpretability of algorithmic decision-making systems. Additionally, techniques like transfer learning and data augmentation are being applied to address the limitation of available samples, optimizing model performance. Moreover, the exploration of multimodal data fusion is gaining traction, harnessing the power of multiple imaging techniques to augment the information available for each patient. The ultimate goal is to develop systems capable of providing real-time decision support, enabling on-the-spot clinical decision-making.

This work aims to provide a comprehensive overview of the current state-of-the-art in the application of ML and DL techniques in SPECT and PET imaging. We will explore the challenges, opportunities, and future prospects of this field. Moreover, we will discuss the use of these algorithms to improve the diagnostic accuracy, predictive modeling, and personalized treatment planning in the context of SPECT and PET imaging. The insights gained from this review can guide future research and development efforts in utilizing machine learning algorithms to further enhance the diagnostic capabilities and clinical utility of SPECT and PET imaging technologies.

## 2. Machine learning in nuclear imaging

The evolution of ML and DL techniques has been remarkable over the past few decades, leading to significant advancements [9,27,28]. In the realm of ML, there has been a shift from more traditional approaches, such as linear methods and decision trees, to more sophisticated and flexible techniques, such as nonlinear ML algorithms and ensembles methodologies [29,30]. One of the most widely used algorithms is support vector machines (SVM) and its kernel diversity [31]. SVMs are well-suited for medical imaging tasks as they can effectively handle nonlinear relationships and accommodate large feature spaces, exhibiting promising results in terms of accuracy and robustness [10,32,33].

Ensemble techniques, such as bagging [34], boosting [35] and random forests [36] show improved capabilities in handling complex datasets and extracting relevant features [12,37].

Despite the good performance of these algorithms, the emergence of DL marked a major milestone in the field of ML. In particular, one of the algorithms that have revolutionised this field the most are convolutional neural networks (CNNs) [38]. These networks stood out for their architecture: convolutional layers that detect local patterns in images using filters, and pooling layers that reduce dimensionality, significantly reducing the number of parameters compared to traditional NNs. An example of CNN is illustrated in Fig. 1, where convolutional layers are applied to extract relevant features and dense layers to perform specific tasks, such as classification or regression. Due to the application of filters, CNNs intelligently adapt to various features within an image. For example, specific filters may emphasize regions with high metabolic activity, enabling the network to concentrate on understanding the distribution of this activity. Additionally, it is common to generate features that detect edges, allowing the network to focus on recognizing the image's shape and contours. Therefore, these deep NNs have demonstrated exceptional abilities to learn high-level patterns and representations from data, leading to improvements in accuracy and generalization in image processing tasks [39–41].

From this architecture, more complex architectures have emerged. For example, siamese networks are a NN architecture with two identical branches processing two different input objects [42]. Their main function is to compare the similarity or difference between these two elements. These networks can therefore play a valuable role in the processing and analysis of molecular images by facilitating the comparison, alignment and improvement of the quantitative accuracy of these images [43–45].

A widely used CNN architecture for medical image segmentation tasks is the U-Net [46,47]. The U-Net architecture consists of a U-shaped design, with a downsampling path and an upsampling path connected by skip connections. The skip connections between the downsampling and upsampling paths allow the combination of low-level and high-level information during the reconstruction phase. This connection helps preserve fine spatial details and provides global context for accurate segmentation. U-Net is commonly used for semantic segmentation of medical images, where the goal is to identify and delineate specific structures or ROIs in the image, such as organs, tumors, and lesions [48, 49]. In particular, this architecture is also useful for employing low-dose PET imaging, thus reducing the radiation to which the patient is subjected [50], or directly to synthesize PET scans from Magnetic Resonance Imaging (MRI) scans [51].

Another widely used architecture is autoencoder (AE) [52]. Its basic structure consists of a NN trained to reconstruct its own input. It comprises two stages. The first stage, called the encoder, takes the original input and transforms it into a lower-dimensional latent representation. The other stage is called the decoder, which takes the latent representation generated by the encoder and reconstructs it in a way that closely resembles the original input. For the design of these models, convolutional layers are often applied to extract the reduced-dimension features

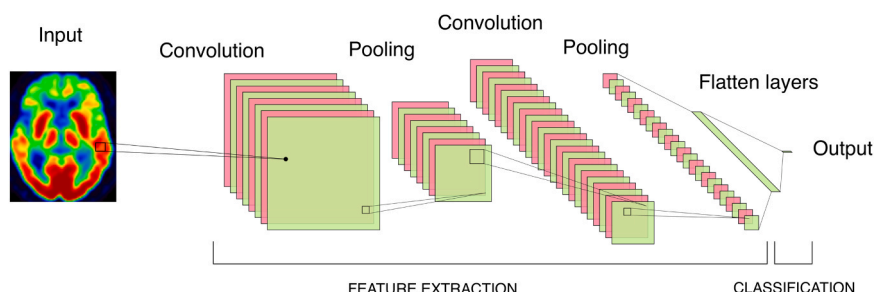


Fig. 1. Overview of a convolutional neural network.

(latent space). Therefore, it is a very useful architecture for dimensionality reduction (reducing the computational cost of the CAD system process) [53], noise removal in images (and hence improving their quality) [54], segmentation of regions of interest (ROIs) [55] or even data generation using Variational Autoencoders (VAEs) [56]. This modality introduces a stochastic component that makes it easier to generate new data in the latent space.

An architecture that has a similar goal to VAEs is the generative adversarial network (GAN) [57]. GANs have gained attention for their ability to generate realistic data samples by training a generator network to produce data that is indistinguishable from real samples, while a discriminator network aims to differentiate between real and generated data. Therefore, GANs have emerged as a valuable tool in neuroimaging, enabling tasks such as image synthesis or correction, and data augmentation [58–60]. Although GANs require large and diverse datasets and careful optimization, they hold great promise for advancing neuroimaging research and personalized medicine.

Another perspective is to study the existing relationships between data instead of directly extracting features from them. In this case, a newly designed DL model to employ is a graph convolutional network (GCN) [61]. GCN was designed to work with structured data in the form of graphs, where nodes and the relationships between them are fundamental. It uses graph convolution operations to propagate information between nodes and each node updates its representation based on the representations of its neighboring nodes in the graph. Its application in medical imaging is growing steadily to study brain connectivity and establish brain disorders prediction. So far, its primary application has been in structural and functional MRI [62–64].

As can be observed, the development of DL models is ongoing in the state of the art. An emerging trend in DL models is the utilization of attention-based models [65], which is gaining increasing popularity. These models selectively process the part of the data which contains the most relevant information in order to make predictions about the output. Among the various DL techniques in this domain, one family that particularly stands out is transformers [66]. Transformers have the capability to process sequential data in a highly efficient manner. To do this, they are composed of stackable encoders and decoders and employ an attention mechanism that allows them to consider all parts of a sequence simultaneously, facilitating the capture of long-range dependencies. These properties led to transformers revolutionising natural language processing (NLP) with models such as BERT (Bidirectional Encoder Representations from Transformers) [67] and GPT (Generative Pre-trained Transformer) [68]. As transformers demonstrated their

effectiveness in NLP, they were adapted to computer vision applications [69]. The Vision Transformer (ViT) architecture [70] enabled image processing by dividing images into patches and treating them as sequences of tokens. This allow to capture complex relationships and dependencies in, for example, medical imaging, resulting in advancements in image segmentation, fusion, classification, synthesis or object detection [71–73]. Although still in early stages, the integration of transformers in neuroimaging shows promising potential for advancing brain research and improving diagnostic capabilities.

All these algorithms are implemented in CAD systems by following the fundamental steps shown in Fig. 2. The first step in the process involves preprocessing the brain images (or different features) to ensure standardized and optimal system performance. Subsequently, the system is trained to accomplish specific tasks, such as categorizing conditions, identifying ROIs, segmenting images, or generating predictions. In this stage, it is considered supervised learning if the data is labeled, or unsupervised learning if there are no labels. In the former case, the model's goal is to learn a function that maps input data to existing labels (e.g., classification), while in the latter, the objective is to group the data into categories or clusters based on similarities (e.g., segmentation). Finally, the performance of the system is evaluated in a final step to assess its performance in terms of accuracy, sensitivity or specificity, among others.

### 3. Applications in SPECT and PET imaging

As previously mentioned, ML (and DL) has found various applications in the field of SPECT and PET imaging, offering significant advancements in the analysis and utilization of these imaging techniques for diagnostic and clinical purposes. This section will discuss the different aspects in which ML has potential applicability in SPECT and PET, which are summarized in Fig. 3, and the advantages that these algorithms offer.

#### 3.1. Optimization of image acquisition and preprocessing

ML has contributed to the improvement of image acquisition, artifact correction, attenuation/scatter correction, and reconstruction in PET and SPECT, leading to higher-quality images and more accurate interpretation [74]. Additionally, several segmentation methods have also been proposed. Table 1 summarizes some of the most recent works in these areas.

An elementary step in obtaining high-quality scans is their

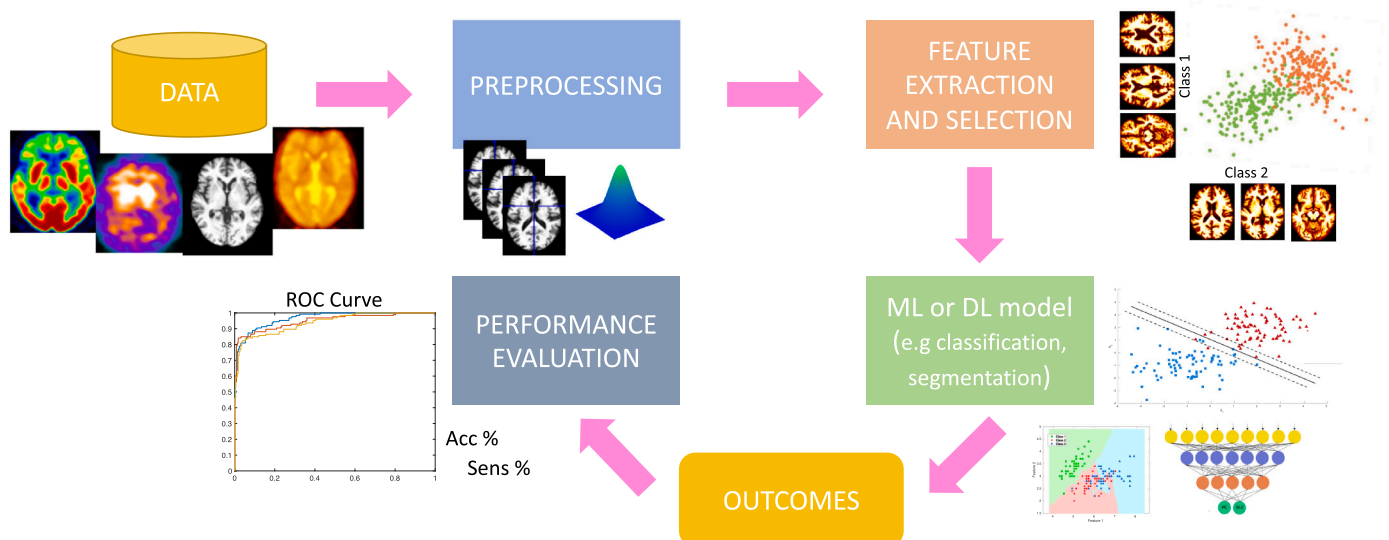


Fig. 2. Typical flowchart of the stages included in a CAD system.

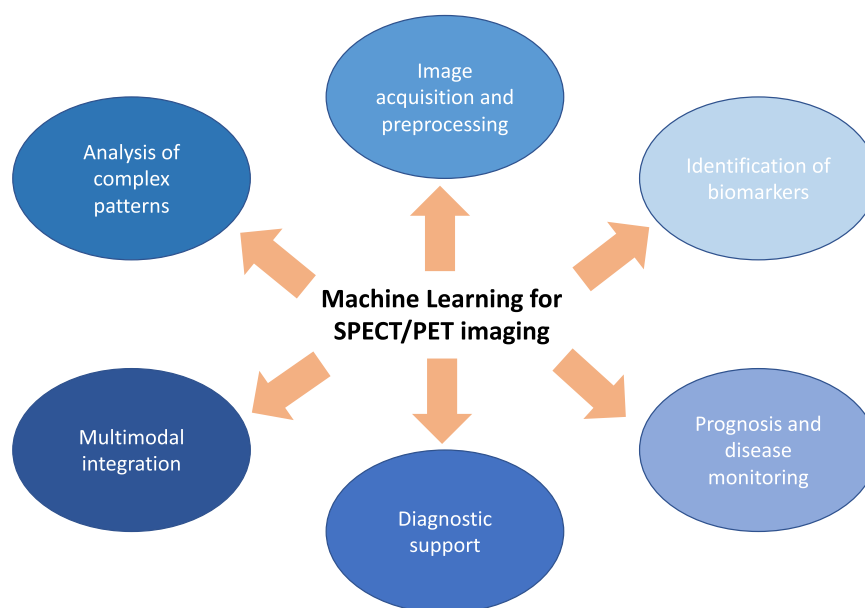


Fig. 3. Key applications of machine learning in PET/SPECT imaging.

Table 1

Summary of recent published works on the application of ML/DL in nuclear image acquisition and preprocessing.

Work	Year	Modality	Radiotracer	Region	Approach	Model
<b>Image acquisition</b>						
[75]	2018	PET			Prediction interaction position	Gradient Tree Boosting
[76]	2018	PET			Prediction interaction position	CNN
[77]	2020	PET			Prediction interaction position	MLP
[78]	2021	PET			Prediction interaction position	Residual-CNN
[79]	2023	SPECT	<sup>99m</sup> Tc-tetrofosmin	Chest	Projections generation	DuDoSS
<b>Image denoising and reconstruction</b>						
[80]	2020	PET			Sinogram data denoising	CNN
[81]	2019	PET	<sup>68</sup> Ga-PRGD2, <sup>18</sup> F-FDG	Whole-body	Image denoising	U-net
[72]	2021	SPECT	<sup>123</sup> I-FP-CIT	Brain	Image generation	ViT
[59]	2018	PET	<sup>18</sup> F-FDG	Brain	Low-dose imaging generation	GAN
[82]	2019	PET	<sup>18</sup> F-FDG	Chest	Denoising in low-dose imaging	U-net
[13]	2022	PET	<sup>11</sup> C-PiB, <sup>18</sup> F-FE-PE2I	Brain	Denoising in low-dose imaging	CNN
[83]	2023	SPECT MPI	<sup>99m</sup> Tc-sestamibi	Chest	Denoising in low-dose imaging	GAN
[84]	2020	PET	<sup>18</sup> F-FDG	Chest	Denoising in low-dose imaging	GAN
[56]	2021	PET	<sup>18</sup> F-DOPA	Brain	Fast dynamic imaging generation	VAE-GAN
[85]	2022	SPECT MPI	<sup>99m</sup> Tc-sestamibi	Chest	Predict standard-dose from low-dose images	GAN
<b>Attenuation/scatter or artifact correction</b>						
[60]	2020	PET	<sup>18</sup> F-FDG	whole-body	Attenuation correction	GAN
[86]	2019	PET	<sup>18</sup> F-FDG	Brain	Attenuation correction	AE
[87]	2019	PET	<sup>18</sup> F-FDG	Brain	ASC	CNN
[88]	2023	SPECT	<sup>99m</sup> Tc-3PRGD2	Chest	Attenuation correction	U-net
[89]	2021	PET	<sup>82</sup> Rb	Chest	Motion correction	LSTM
[90]	2020	PET	<sup>18</sup> F-FDG	Chest	Motion correction	Transformer
<b>Image segmentation</b>						
[48]	2018	PET	<sup>18</sup> F-FET	Brain	Glioma's lesion detection and segmentation	U-net
[91]	2019	PET	<sup>18</sup> F-FDG	Whole-body	Cervical tumor segmentation	CNN
[49]	2022	PET/CT	<sup>18</sup> F-FDG	Chest	Lung lesion segmentation	U-net
[92]	2020	SPECT	<sup>99m</sup> Tc	Whole-body	Segmentation of thorax bone	U-net/CNN

MPI: myocardial perfusion imaging, LSTM: long short-term memory, DuDoSS: deep-learning-based dual-domain sinogram synthesis is a model based on residual dense networks.

acquisition process. This step is dependent on the scanner's instrumentation and their associated parameters (e.g., timing and energy resolution or localization of the interaction position) which affect factors such as sensitivity, spatial resolution, and signal-to-noise ratio of the scanner. For example, Sanaat and Zaidi [77] proposed the use of a positioning algorithm based on a multilayer perceptron (MLP) to predict the localization of the interaction position inside a monolithic crystal in order to enhance the spatial resolution.

Traditional reconstruction methods may be limited in terms of image quality, noise reduction, and quantitative accuracy. ML algorithms,

particularly those based on DL, have been recently proposed to cope with this task [6]. Both CNNs [80,81] and GANs [84,85] have been employed to reduce noise and enhance the quality of PET and SPECT images. These models learn patterns of noise and are capable of generating cleaner and sharper images, enabling more precise interpretation and improved anomaly detection. For example, Lu et al. [82] analyze the possibility of reducing radiation dose in PET imaging while maintaining image quality and accuracy. It was found that 3D U-Net effectively reduced image noise and controlled bias, even for small lung nodules, using low count down-sampled data to generate standard dose PET



images.

Moreover, ML algorithms can be used to correct attenuation artifacts in PET and SPECT images [93,94]. These models can learn the relationship between the original images and attenuation maps, applying corrections to obtain more accurate and quantitatively precise images. For example, in [87] it is proposed joint attenuation and scatter correction (ASC) in image space for non-corrected PET without the need for conventional attenuation map generation and time-consuming scatter correction. Motion artifacts and other artifacts can also be corrected in PET and SPECT through ML algorithms, improving image quality and reducing the impact of involuntary movements during image acquisition [95].

Automated segmentation of anatomical structures in PET and SPECT images is another crucial step in quantitative and qualitative analysis. ML algorithms have been utilized to segment specific ROIs in the images, such as tumors or specific brain areas [96], enabling more precise analysis and robust feature extraction. One algorithm that has gained importance in this field is the U-Net. Based on this algorithm, in [48], a CAD system for lesion detection and segmentation of gliomas in Amino-acids PET is proposed. The algorithm successfully detected all lesions in the dataset. Specifically, the voxel-level segmentation achieved a sensitivity of 88% and a specificity of 99 %.

### 3.2. Multimodal integration

ML techniques have facilitated and can take advantage of the benefits of the integration of multimodal data in PET and SPECT, such as combining these imaging modalities with MRI [97,98] or computerized tomography (CT) [49,99]. This integration provides a more comprehensive and enhanced view of the conditions under analysis and overcoming limitations of individual modalities. Moreover, it harnesses both functional information (e.g., SPECT/PET) and anatomical information (e.g., MRI/CT), yielding several benefits, including improved image quality (higher resolution and precision), enhanced diagnostic capabilities due to a wealth of patient information, early detection of abnormalities, streamlined treatment planning, and the ability to conduct more comprehensive and in-depth clinical research on various diseases and medical conditions. Nonetheless, the integration of multiple medical imaging modalities carries certain disadvantages. These drawbacks include technical complexity and the cost associated with acquiring and maintaining specialized equipment. Moreover, the image fusion process can be time-consuming, which is not ideal in emergency situations. Some patients may experience discomfort due to longer scan times, and in hybrid systems like PET/CT, there is a concern about increased radiation exposure. Additionally, managing large datasets generated by integration may require substantial resources. Therefore, the decision to use multimodal imaging should be based on the specific clinical or research objectives and the available resources. Some recent works related to the fusion of the most common image modalities are reported in Table 2.

Various approaches are employed to integrate these modalities, which can be categorized into four groups [115]. Multimodal Fusion involves visually combining multiple modalities to generate a single composite image that presents information from both. These images can be acquired using different instruments or the same one, and coregistration is necessary to align them accurately. The Multi-Focus Fusion technique is utilized for merging images captured with varying focal lengths. In this approach, regions of interest (ROIs) are designated in anatomical images (such as MRI), and these ROIs are applied to PET images to quantify metabolic activity in specific regions. Multi-Temporal Fusion pertains to the fusion of images of the same modality captured at different moments, facilitating the detection of changes in the subject over time. Lastly, Multi-View Fusion involves the simultaneous acquisition of images of the same modality under differing conditions, providing a broader spectrum of information. For more in-depth insights into these techniques, please read [115].

**Table 2**

Summary of recent published works related to multimodal fusion.

Work	Year	Radiotracer	Region	Goal	Model
<b>PET/CT</b>					
[100]	2021	$^{18}\text{F}$ -FDG	Whole-body	Full-dose synthesis from low-dose imaging	GAN
[101]	2023	$^{18}\text{F}$ -FDG	Upper-body	Prognosis	Ensemble
[102]	2021	$^{18}\text{F}$ -FDG	Upper body	Diagnostic classification	SVM / LDA
[103]	2022	$^{18}\text{F}$ -FDG	Whole-body	ROIs identification	U-net / CNN
[104]	2021	$^{18}\text{F}$ -FDG	Upper body	Prognosis	Residual CNN
<b>PET/MRI</b>					
[45]	2023	$^{18}\text{F}$ -FDG	Brain	ROIs identification	Siamese NN
[105]	2023	$^{18}\text{F}$ -FDG	Whole-body	ROIs identification	Random forest
[106]	2021	$^{18}\text{F}$ -FDG	Brain	Diagnostic classification	CNN
[107]	2023	$^{18}\text{F}$ -FDG	Brain	Diagnostic classification	SVM / RF
<b>SPECT/CT</b>					
[108]	2023	$^{99\text{m}}\text{Tc}$ -Sestamibi	Renal scintigraphy	Diagnostic classification	XGBoost
[109]	2021	$^{99\text{m}}\text{Tc}$ -GSA	Whole-body	Prognosis	SVM / RF
[110]	2023	MPI (REFINE [111])	Chest	Prognosis	XGBoost
[112]	2022	$^{68}\text{Ga}$ -PSMA-HEBD-CC	Whole-body	Prediction of dosimetry	RF regression/MLP
<b>SPECT/MRI</b>					
[54]	2021	$^{123}\text{I}$ -FP-CIT	Brain	Image reconstruction	AE
[98]	2023	$^{99\text{m}}\text{Tc}$ -Sestamibi	Brain	Data fusion modeling	CNN
[113]	2022	$^{123}\text{I}$ -FP-CIT	Brain	Disease progression analysis	PCA / K-Means
[114]	2021	$^{123}\text{I}$ -FP-CIT	Brain	Diagnostic classification	Ensemble

RF: random forest, LDA: linear discriminant analysis

An example multimodal data fusion in neuroimaging is illustrated in Fig. 4. Both SPECT and MRI scans are preprocessed (coregistered and overlaid) to obtain the multimodal fusion scan. In this field, the fusion offers significant advantages, such as correction of distortions, improved temporal and spatial resolution, and the integration of structural and functional information. Moreover, there has been an impact of multimodal data fusion on the transition of neuroimaging diagnosis from qualitative analysis to quantitative evaluation, addressing photon attenuation, scatter, and partial volume effects in PET and SPECT quantification. For example, the attenuation effect present in PET/MRI systems may be a challenge. A DL model proposed to mitigate this is presented by I. Shiri et al. in [86], where a convolutional AE is applied to directly correct the attenuation effect in PET images. Regarding the partial volume effect (PVE), Quarantelli et al. [116] developed an Integrated software called PVELab to conduct PVE correction for brain PET/SPECT scans and their coregistered segmented MRI.

Several studies have shown that this fusion of imaging modalities results in increased diagnostic accuracy. For example, Polikar et al. [117] conducted an study related to Alzheimer's Disease (AD) where a binary classification (AD vs. HC) was analyzed by means of MRI, PET and electroencephalogram (EEG) data fusion. In this case, an accuracy of 85.55% is obtained applying an ensemble of DL classifiers (multilayer perceptrons, MLPs) based on sum majority voting. Other multimodal data fusion example is the work developed by Song et al. [106], where the gray matter tissue area of MRI scans and FDG-PET images are fused

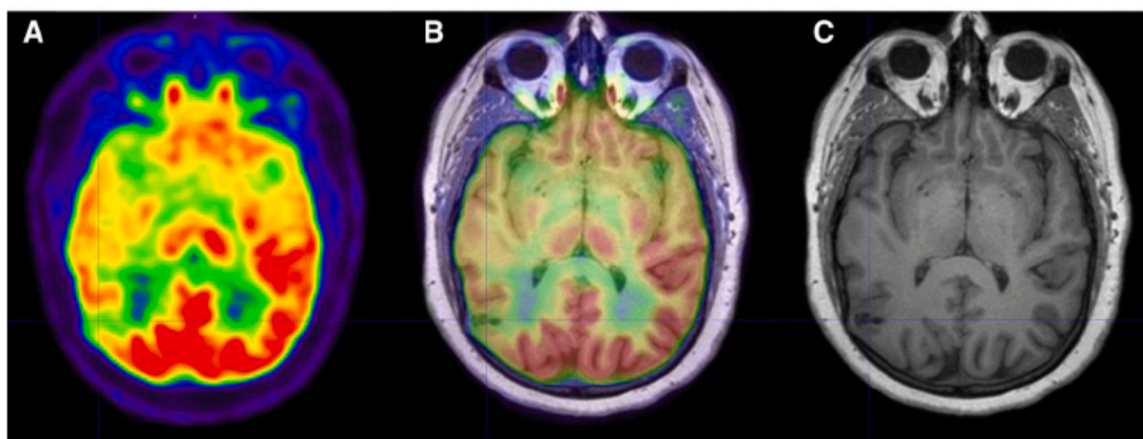


Fig. 4. Multimodal fusion scan (B) from SPECT (A) and MRI (C) scans of the same individual.

by registration and mask coding, denominating the method as “GM-PET”. Then, a 3D CNN and a 3D Multi-Scale CNN was applied in binary classification and multi-classification tasks to evaluate the effectiveness of the method, achieving accuracies up to 94.11% in the AD-vs-HC task or 71.52% in the AD-vs-MCI-vs-HC task. Arco et al. [45] proposed the use of a siamese neural network to fuse information from MRI and PET. This enabled to quantify similarities between both modalities, detecting the role of different brain regions in the development of AD according to structural and functional changes.

As can be inferred, multimodal integration is an area of potential future growth which may uncover valuable insights and improving diagnostic accuracy. In fact, most recent studies are tending to apply this approach, for example by merging PET and CT data [49,99] or PET/-SPECT with MRI [45,98].

### 3.3. Identification of biomarkers and analysis of complex patterns

The search for new biomarkers to model neurodegenerative diseases

is a challenging task that has implications for both the development of more effective diagnostic support systems and the quantitative measurement of neuronal loss and other physiological changes [118]. Previous studies addressing the characterization of diseases like Parkinson’s Disease (PD) have primarily focused on genetic and biochemical markers [119–121]. However, ML techniques have emerged as powerful tools in identifying specific biomarkers that are less invasive and do not require specialized laboratory units [122]. This advancement provides valuable insights into the underlying mechanisms of diseases and contributes to personalized treatment approaches. In the context of SPECT and PET imaging, when examining the radiological characteristics of the striatum region in healthy control subjects (HC) and PD patients, it becomes apparent that the shape of the region with higher uptake of the radioligand  $^{123}\text{I}$ -Ioflupane itself can potentially serve as an early disease progression marker, even before clear cognitive impairment becomes evident [119,123]. Following this line, Castillo-Barnes et al. [124] proposed to extract morphological features from isosurfaces of the scans at various intensity levels. Then, these features were classified by means

**Table 3**  
Summary of recent published works related to clinical Applications of nuclear imaging.

Work	Year	Modality	Radiotracer	Region	Approach	Model
<b>Biomarkers identification/Pattern analysis</b>						
[124]	2020	PET	$^{123}\text{I}$ -FP-CIT	Brain	Features from isosurfaces	SVM
[125]	2017	PET	$^{18}\text{F}$ -FET	Brain	Textural features	Clustering
[126]	2019	PET/CT	$^{18}\text{F}$ -FDG	Chest	Features from polar maps	CNN
[122]	2023	SPECT MPI	$^{99\text{m}}\text{Tc}$ -MIBI	Chest	Features from polar maps	AE
[127]	2023	SPECT/MRI	$^{123}\text{I}$ -FP-CIT	Brain	$^{123}\text{I}$ -FP-CIT uptake prediction on MRI	CNN
[128]	2021	PET	$^{11}\text{C}$ -MET	Brain	Low-dimensional feature extraction	PCA
[129]	2023	SPECT MPI	Tl-201	Chest	Low-dimensional feature extraction	AE
[53]	2021	PET	$^{18}\text{F}$ -FDG	Brain	Feature ranking	AE
[130]	2022	SPECT/MRI	$^{123}\text{I}$ -FP-CIT	Brain	ROIs significance	SPM/SAM
[131]	2020	PET	$^{18}\text{F}$ -FDG	Brain	ROIs significance	SSP/SPM
<b>Diagnostic support systems</b>						
[12]	2018	SPECT	$^{123}\text{I}$ -FP-CIT	Brain	Diagnostic classification	SVM
[132]	2019	PET	$^{18}\text{F}$ -FDG	Brain	Diagnostic classification	InceptionV3
[133]	2022	PET	$^{18}\text{F}$ -FDG	Brain	Diagnostic classification	CNN
[134]	2022	SPECT	$^{123}\text{I}$ -FP-CIT	Brain	Diagnostic classification	CNN
[135]	2022	SPECT	$^{123}\text{I}$ -FP-CIT	Brain	Diagnostic classification	Clustering
[136]	2019	PET/CT	$^{18}\text{F}$ -FDG	whole-body	Classification of uptake patterns for cancer	CNN
[108]	2023	SPECT/CT	$^{99\text{m}}\text{Tc}$ -MIBI	Chest	Tumor characterization and diagnostic	XGBoost
[137]	2023	PET/CT	$^{18}\text{F}$ -FDG	Upper body	Cancer detection	Random forest
[138]	2022	PET/CT	$^{18}\text{F}$ -FDG	Chest	Nodules classification	XGB/MLP
<b>Disease prognosis and monitoring</b>						
[99]	2023	PET/CT	$^{18}\text{F}$ -FDG	Upper body	Predictor of cancer prognosis	Random Survival Forest
[139]	2020	PET	$^{18}\text{F}$ -FDG	Whole-body	Predictor of cancer prognosis	Random Survival Forest
[140]	2023	PET/CT	$^{18}\text{F}$ -FDG	whole-body	Predictor of cancer prognosis	Logistic regression
[141]	2020	PET/CT	$^{18}\text{F}$ -FDG	Upper body	Prognostic value of radiomics	Random Forest
[142]	2022	PET/CT	$^{18}\text{F}$ -FDG	whole-body	Prediction of tumor progression	Naive Bayes
[143]	2022	SPECT	$^{123}\text{I}$ -FP-CIT	Brain	Disease progression modeling	SVM

MPI: myocardial perfusion imaging, SSP: statistical surface projections

of SVM and Naïve Bayesian algorithms. The results indicated the reliability of these biomarkers, particularly those related to sphericity, center of mass, number of vertices, 2D-projected perimeter, and 2D-projected eccentricity. Other works are summarized in Table 3.

In line with biomarker identification, these algorithms also allow for the analysis of complex patterns and extraction of relevant features in PET and SPECT images, revealing hidden relationships and enhancing the understanding of neurological disorders. Feature extraction is a commonly applied approach in CAD systems as a preliminary step to classification. This extraction is performed through dimensionality reduction, which transforms a high-dimensional dataset into a lower-dimensional representation. In more traditional ML systems, this is achieved by applying techniques such as Principal Component Analysis (PCA) or Partial Least Squares (PLS) [128,144,145]. DL-based systems, such as AE, are also applied for this approach [53,129]. Furthermore, it is important to note that DL models that employ convolutional layers extract features within these layers, which are subsequently used in the algorithm's decision-making process [126,134].

Another technique used for analyzing neuroimaging patterns is the generation of brain maps. Statistical brain mapping is a straightforward method to detect ROIs when comparing two groups, such as HC and PD subjects. It helps identify relevant patterns between these groups, making it particularly useful for exploratory analyses. The standard approach for statistical inference maps is the widely-used Statistical Parametric Mapping (SPM) software [146], which employs classical statistics (e.g., general linear model) to generate maps. However, this approach relies on various assumptions and has limitations [147,148]. As a result, researchers have been exploring alternative approaches. One proposed method is data-driven supervised learning techniques, such as those utilized in ML approaches [149]. Although ML techniques were not originally developed for brain mapping hypothesis testing [150–153], they can provide confidence intervals in classifying image patterns and produce maps of statistical significance [154]. To address this, a non-parametric approach called Statistical Agnostic Mapping (SAM) was proposed by Gorriz et al. [155]. SAM employs concentration inequalities and provides confidence intervals by assessing the upper

bounds of the actual error in binary classification problems and by using simple significance tests of a population proportion within it. It was found that this novel framework based on Statistical Learning Theory (SLT) provides similar activation maps than the ones obtained by the voxel-wise SPM but in a region-wise level. A comparison of both methods is shown in Fig. 5 for PD vs HC contrast from PET images.

The study conducted by Jimenez-Mesa et al. is an example of the applicability of these statistical maps [130]. Both, SPECT images and MRI scans are compared to assess the usefulness of these imaging modalities for Parkinson's disease. Despite the existing state-of-the-art proposal to apply MRI for PD, this study concluded that MRI is not a very reliable source of information for the diagnosis of PD, regardless of the classification results reported in previous studies [97,156].

### 3.4. Diagnostic support

Applying ML algorithms in CAD systems offers a significant advantage in enhancing diagnostic accuracy. ML algorithms excel at identifying disease-specific features and constructing predictive models, resulting in earlier and more precise detection of various conditions, including neurological disorders and cardiac diseases, among others [7, 157,158]. Ultimately, these advancements significantly benefit patient care and contribute to improved treatment outcomes. Some recent examples of these algorithms used to support diagnosis are shown in Table 3.

Studies based on the more standard ML techniques usually apply ensemble methodologies, i.e. multiple models are combined to improve the overall prediction or classification performance. It also tends to happen that not only features extracted from the image are used, but also other clinical or demographic data. An example of both situations is the work of Castillo et al. [159]. An ensemble methodology, based on SVM, was employed in this study. In addition to SPECT data, cerebrospinal fluid (CSF), RNA, and serum tests were used as features, achieving an accuracy of up to 96% when comparing PD patients and HC.

On the other hand, in DL-based studies, algorithms usually extract

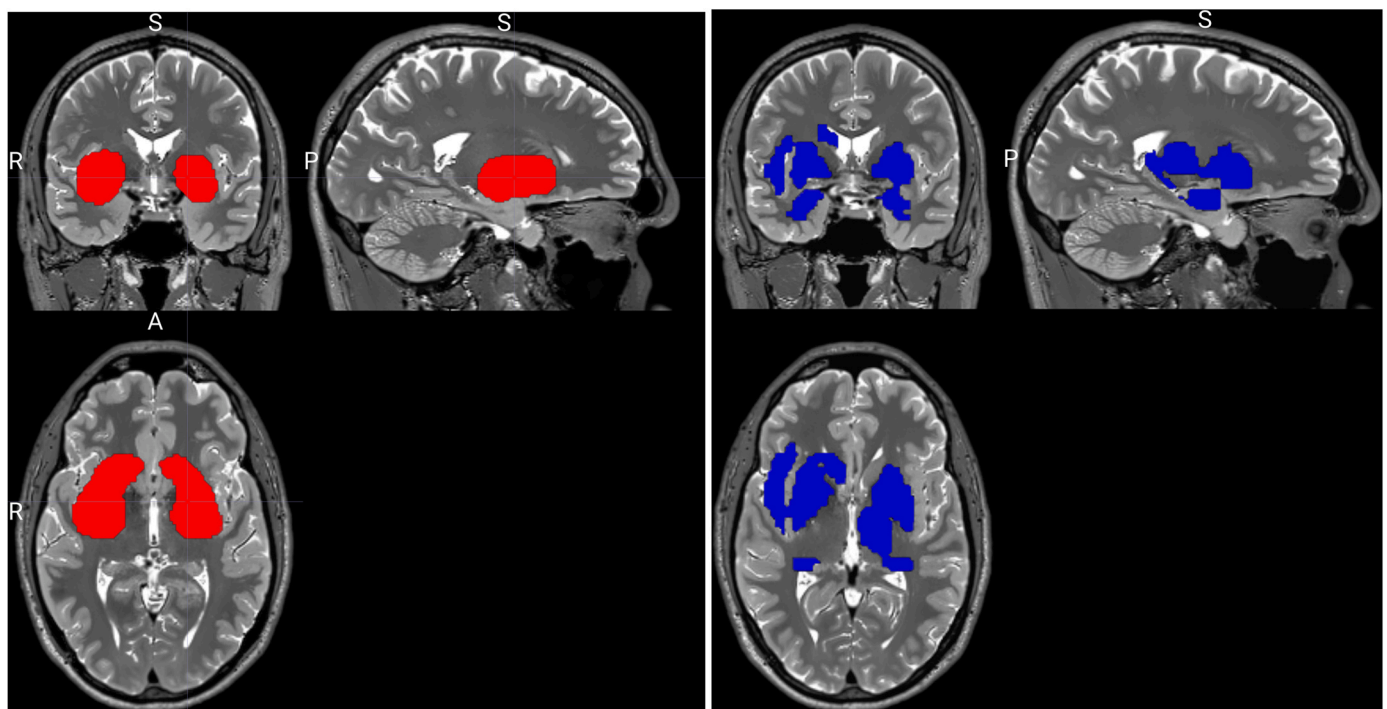


Fig. 5. Performance comparison between SPM (red, left) and SAM (blue, right) when contrasting PD vs. HC samples of a SPECT dataset. While the former uses a voxel-wise approach, the latter uses an atlas-based region-wise approach.

the relevant information directly from the image. For example, in the study conducted by Ding et al. [132], the InceptionV3 architecture [159] was applied to predict early diagnosis of AD using  $^{18}\text{F}$ -FDG PET imaging. The model achieved impressive results, with an area under the receiver operating characteristic (ROC) curve of 0.98. Furthermore, the algorithm achieved a specificity of 82% and a sensitivity of 100%, indicating its ability to accurately identify individuals with AD. Similarly, Etmiani et al. [133] utilized a 3D CNN model for predicting the diagnosis of dementia with Lewy bodies, AD, and mild cognitive impairment (MCI) using  $^{18}\text{F}$ -FDG PET imaging. This was a multiclass classification problem, and the model achieved excellent performance. The area under the ROC curve was reported as 96.2% for dementia with Lewy bodies, 96.4% for AD, 71.4% for mild cognitive impairment due to AD (MCI-AD), and 94.7% for HC.

### 3.5. Prognosis and disease monitoring

ML algorithms have proven effective not only in establishing diagnoses but also in predicting the prognosis of diseases and tracking disease progression over time. These capabilities are invaluable for clinical decision-making, as clinicians can predict the likelihood of disease progression, estimate the rate of decline in motor and cognitive functions, and identify patients who may benefit from specific treatment strategies or interventions. Some works focused on this purpose are shown in Table 3.

In order to assess the prognosis of a given condition, quantitative features are usually analyzed and their value is assessed by considering different factors (clinical assessments, genetic markers, and demographic information, etc.). For example, Zhong et al. [99] proposed several radiomics features from  $^{18}\text{F}$ -FDG PET/CT scans and clinically derived features for successfully predicting prognosis of multiple myeloma patients. For the same purpose, Morvan et al. [139] applied a ML model based on Random Survival Forest (RFS) and variable importance (VIMP) using radiomics, conventional (e.g., standardized uptake value or SUV) and clinical biomarkers, reducing the error of the predicted progression.

Furthermore, ML algorithms can track disease progression longitudinally by analyzing serial SPECT or PET scans taken at different time points. By comparing imaging features and biomarkers over time, these algorithms can provide valuable insights into the evolution of the

disease, detect subtle changes in neurobiology, and assess the effectiveness of therapeutic interventions. For this, several neuroimaging initiatives are collecting longitudinal data from patients, such as the Parkinson's Progression Markers Initiative (PPMI) (www.ppmi-info.org), which primary objective is to identify new biomarkers that can measure the progression of PD and develop new treatments; the Alzheimer's Disease Neuroimaging Initiative (ADNI) [160] and the Dominantly Inherited Alzheimer Network (DIAN) [161], both related to the study of AD for early diagnosis and better understanding of the disease. These databases are continuously updated, with the inclusion of new participants in the studies and the addition of more recent samples from existing participants, which makes them widely used repositories [24, 143, 162, 163].

## 4. Challenges, opportunities and future prospects

As we have seen, ML holds great promise for advancing SPECT and PET imaging, but it also presents certain challenges. Understanding these challenges is crucial for harnessing the full potential of ML in this domain. The opportunities are many, and the future is promising. The main issues addressed in this section are illustrated in Fig. 6.

### 4.1. Challenges

Among the problems faced by ML in PET/SPECT imaging domains, four could be highlighted: data standardization, limited data availability, evaluation and interpretability. These challenges need to be addressed to increase confidence in CAD systems as diagnostic and prognostic support methods.

#### 4.1.1. Data standardization

Data standardization is related to the acquisition protocols of PET/SPECT scans. Variations in imaging protocols, data acquisition, and preprocessing techniques pose significant challenges in the development and deployment of ML models across different clinical sites. These differences can lead to inconsistencies in the data, affecting the performance and generalizability of ML algorithms. For instance, variations in image resolution, signal intensity, or image reconstruction methods can introduce biases and confounding factors that impact the accuracy of the models. To address these challenges, standardization efforts are crucial.

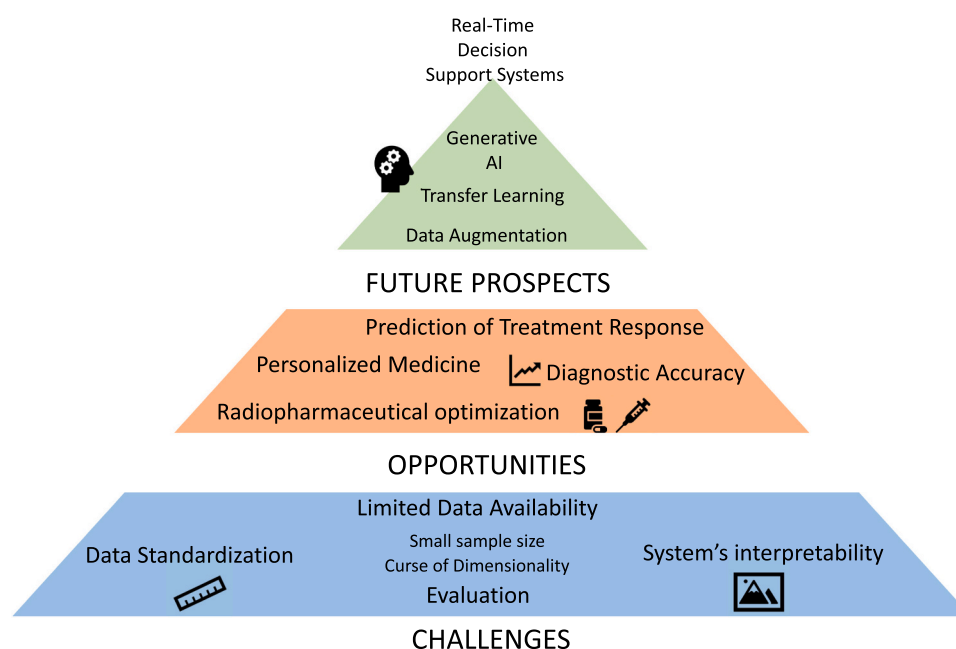


Fig. 6. Challenges, opportunities and future prospects about the application of ML in PET/SPECT imaging.



Standardization involves the establishment of uniform guidelines, protocols, and procedures that ensure data consistency and compatibility across different imaging centers. This includes standardizing imaging acquisition parameters, data preprocessing steps, and quality control measures. Fortunately, as we have already discussed, there are more and more techniques that try to deal with image acquisition and processing protocols [80,81,84,85,95].

Standardization efforts also extend to data sharing and collaboration among research institutions. By creating centralized databases and repositories that follow standardized data formats and annotations, researchers can pool larger and more diverse datasets. This data sharing not only promotes collaboration but also improves the generalizability and reliability of ML models. Moreover, standardization may also increase the use of transfer learning techniques (which will be discussed in Section 4.3), as ML models can be trained more effectively on one site and effectively transferred to other sites with minimal performance degradation. Therefore, standardization efforts play a crucial role in overcoming the heterogeneity of data across different clinical sites in neuroimaging and other medical domains. By ensuring consistency and compatibility, standardization paves the way for more robust and reliable ML models that can be deployed across diverse healthcare settings.

#### 4.1.2. Limited sample sizes

Closely related to this challenge is the limited data availability in the radiology field (and in the field of neuroimaging itself [11,164]). This issue is related to the differences between samples from different clinical centers and the difficulty of obtaining clinical data, which leads to a problem known as the “Small Sample Size Problem” [165,166]. Neuroimaging studies often rely on data from a relatively small number of subjects, typically ranging from fewer than a hundred to rarely exceeding a thousand. Such limited sample sizes introduce statistical uncertainty, leading to imprecise measurements, variable effect sizes, and an increased risk of false positives and false negatives in disease detection [167,168]. Therefore, these statistical limitations collectively weaken the statistical power of the studies. Although larger sample sizes would be ideal, practical implementation presents challenges, including difficulties in participant recruitment and conducting multi-center studies with diverse clinical conditions and acquisition protocols [169]. Initiatives like the Human Connectome Project (HCP) [170] and UK Biobank [171] aim to collect data from thousands of participants, offering potential solutions. However, for highly complex DL models such as transformers, the current number of samples in these repositories would still be insufficient. Another approach gaining popularity is the use of data augmentation techniques [172–174], which will be discussed in section 4.3. For example, Lu et al. [82] employed synthetic data for the initial training of the U-net model. Subsequently, real samples are utilized to fine-tune the pre-trained model.

In relation to this problem, the literature also refers to the “Curse of Dimensionality” [175]. It arises from the imbalance between the number of samples (sample size) and the number of features used in computational systems. Neuroimaging data, such as brain imaging scans, encompass a vast number of voxels, often reaching millions, with additional features derived from anatomical and physiological measurements [176,177]. To mitigate the “Curse of Dimensionality”, feature selection and feature extraction techniques have been widely employed in traditional ML-based CAD systems [22,178–180]. However, such approaches are less meaningful in DL, as DL architectures inherently perform feature extraction and subsequent classification [181,182]. Precisely, the capacity of NN architectures to generalise is currently under discussion [183,184], as their complexity increases, they may even learn from corrupted data (permutation of data labels) [185]. Thanapol et al. [186] applied data augmentation during the training of a CNN to test its effect on the overfitting and generalization ability of the model. In this context, frameworks to assess the reliability of classification performance in neuroimaging is a current issue [187,188]. C. Jimenez-Mesa et al. [162] proposed a non-parametric framework to

assess such statistical significance of the ML models. In this study, not only is the generalizability of the models assessed, but there is also a suggestion that the validation method employed in neuroimaging may not be well suited to the small sample size, as discussed by Isaksson et al. [189]. Therefore, an alternative method based on resubstitution with upper-bound correction is proposed.

This issue is also encountered in voxel-based morphometry studies, where classical techniques tend to produce numerous false positives or rely heavily on sample size [147]. Thus, recent ML-based methods opt for mapping based on ROIs and dimensionality reduction techniques [155,190].

#### 4.1.3. Evaluation

A limited sample size is one of the major challenges when designing CAD systems. Intimately tied to this challenge is the method employed to assess the implemented algorithms. The effectiveness of our approach significantly hinges on how we partition the data, whether through training and test sets, cross-validation, substitution, or other means. To illustrate, when dealing with a initially small sample set and employing k-Fold cross-validation, we encounter substantially reduced subgroups characterized by considerable result variability within each [168]. Therefore, the sample size, the generalization ability, and the performance evaluation approach of the implemented ML techniques are closely related to the reliability of the CAD system.

While robust systems are typically obtained with large datasets and low generalization error, concerns arise with models based on DL that may still exhibit low generalization error even with corrupted data or manipulated labels [185,191]. Detecting and measuring these issues becomes essential, leading to the need for methods to assess the generalization capacity of NNs [183,184,186,192]. The reliability of classification results in neuroimaging has already been studied, with a focus on evaluating if the estimated accuracy surpasses chance performance [187,188]. Permutation methods are commonly employed for this purpose [193]. Additionally, the literature proposes statistical power and type I error analysis methods [194,195], including comparisons between cross-validation (CV) and resubstitution methods [162]. CV may not be effective for small sample sizes, leading to controversial predictive power of the classifier [168,189,196]. Hence, the application of a resubstitution approach with an upper bound correction [196,197] provides a worst-case accuracy estimate comparable to that of the CV test subset. The advantage lies in utilizing the complete dataset without dividing it into folds, thus enhancing the reliability of accuracy estimation.

Another significant challenge emerges when DL models, initially trained on one dataset, demonstrate less-than-optimal performance when applied to a different dataset. This phenomenon, known as dataset shift or domain adaptation, underscores the complexity of assessing a DL model’s robustness and generalization capabilities [198]. The inherent variability between datasets, stemming from differences in acquisition protocols, demographics, and other factors, makes it difficult to establish a one-size-fits-all evaluation framework. Consequently, researchers are faced with the task of developing evaluation strategies that not only account for the model’s performance within its training dataset but also its adaptability to novel data sources [199–201]. This aspect of model evaluation remains an ongoing challenge in the field of deep learning for medical applications.

#### 4.1.4. Interpretability

In addition to the crucial aspect of reliability, another significant challenge in the application of ML models, especially complex ones like DL, is the issue of interpretability. ML algorithms, particularly deep NNs, are known for their black-box nature, where they make predictions without providing explicit explanations for their decisions. This lack of interpretability poses challenges in understanding the underlying reasoning and factors driving the model’s predictions. The interpretability of ML models is of paramount importance in various domains,

including neuroimaging and healthcare. Clinicians and medical professionals need to have confidence in the model's predictions and understand the rationale behind them before they can fully trust and adopt these technologies for clinical decision-making. Additionally, regulatory authorities often require transparent and interpretable models to ensure patient safety and ethical considerations.

Addressing the interpretability challenge is an active area of research in ML and DL that has a lot of future research ahead of it. Various techniques, such as model visualization [202], feature importance analysis [203], and attention mechanisms [204], aim to shed light on the internal workings of complex models and provide explanations for their predictions. These techniques could be encompassed in the concept of eXplainable Artificial Intelligence (XAI) [205]. XAI techniques aim to enhance the transparency and understandability of complex models, enabling researchers and clinicians to gain insights into the underlying mechanisms of disorders and treatment responses [206,207]. These techniques allow for the identification of key features, e.g. regions in brain images, that contribute most to the model's predictions. This helps in understanding the relationships between specific ROIs and diseases, as well as the impact of certain features on diagnostic or prognostic decisions. XAI methods can generate visual explanations, such as saliency maps [204,208] or Grad-CAMs [202], highlighting the regions that are most relevant for the model's decision-making process. Therefore, these techniques, which make ML models more transparent and interpretable, enhance the trustworthiness and adoption of ML models in clinical applications, since clinicians are more likely to embrace and integrate ML-based technologies into their practice if they can comprehend how the model arrived at a specific diagnosis or treatment recommendation [206]. For example, M. Nazari et al. [209] applied an explainable method to clinically analyzed the CNN-based classification of DAT-SPECT in patients with clinically uncertain parkinsonian syndromes. Moreover, these techniques are not only useful for classification assessment or pattern detection in imaging, but can even be applied to cognitive testing [210], which could lead to a reduction in the number of tests to be undertaken by the patient.

#### 4.2. Opportunities

Throughout this work, we have already seen that the advantages offered by the application of ML and DL models in radiology are many and intriguing. In this section, the most relevant clinical opportunities are highlighted.

An evident trend in the literature is the consistent improvement in diagnostic accuracy due to ML algorithms [132,157,211], with reported accuracies exceeding 95% in studies related to conditions like PD and AD [12,133]. Consequently, ML algorithms hold substantial promise for elevating the precision of SPECT and PET imaging by uncovering subtle patterns and features that often challenge human interpretation, ultimately leading to more precise and dependable diagnoses. Furthermore, strategies to overcome the challenge of small sample sizes, as discussed in Section 4.1, encompass a range of approaches such as dimensionality reduction techniques [22,178] and refinements in validation methodologies [162,196], all aimed at bolstering model generalization.

Another important breakthrough is emerging in personalized medicine. ML models play a pivotal role in this paradigm, which represents a cutting-edge approach that holds great promise for revolutionizing healthcare, by harnessing the power of diverse patient data to develop personalized treatment plans that cater to individual needs and characteristics. In this sense, the identification of patient-specific biomarkers through ML models is of great interest [120,122]. These models have the ability to efficiently process and analyze vast amounts of data from different sources, including imaging scans, genetic profiles, and clinical records. Therefore, by integrating such a diversity of data, ML models can uncover intricate patterns and correlations that might be otherwise overlooked, leading to a comprehensive understanding of a patient's unique health profile. These biomarkers can be indicative of disease

progression, treatment response, or potential adverse effects as reported in the works discussed in Section 3.3.

Additionally, it has been discussed that ML models can predict treatment response, allowing physicians to make informed decisions about the most suitable therapeutic approaches for each patient [99, 139]. These predictions are based on comprehensive analyses of past patient data and clinical outcomes, enabling healthcare providers to identify the most effective treatments for specific disease subtypes or patient population.

Finally, another opportunity to highlight is the radiopharmaceutical optimization. Radiopharmaceutical optimization through ML techniques represents a significant advancement in the field of SPECT and PET imaging, offering the potential to enhance the overall effectiveness and safety of these diagnostic procedures, e.g. changing the dose planning or administration schedules. Radiopharmaceuticals are essential components of nuclear medicine imaging, and their proper utilization is crucial for accurate disease diagnosis and patient management. Moreover, it has been seen in Section 3.1 that ML techniques can help address acquisition and reconstruction challenges, such as motion artifacts [95], attenuation artifacts [87] or denoising [82,85], which could allow a reduction in the dose of radiopharmaceutical required due to better image quality.

#### 4.3. Future prospects

In this section, the emerging trends related to the application of ML and DL algorithms in SPECT/PET modalities are further discussed. Specifically, the domains that we consider to have the potential for the greatest growth and relevance for clinical research are transfer learning and data augmentation, generative AI and real-time decision support systems.

Transfer learning and data augmentation are already mentioned research areas that will be of great relevance in addressing the reliability of CAD systems for SPECT and PET imaging. By leveraging pre-trained models on large datasets from related domains, ML algorithms can be enhanced even with limited SPECT and PET data, leading to improved diagnostic accuracy and robustness. Moreover, data augmentation methods offer a valuable solution to address the challenge of limited data. By artificially increasing the size and diversity of the dataset, data augmentation enables more comprehensive ML training.

The use of synthetic samples generated through data augmentation can indeed enhance the robustness of ML models by introducing additional variability and diversity into the training data. However, it is crucial to acknowledge that such artificially generated variability may not always accurately represent the true clinical population. As a result, the reliability and clinical relevance of the ML model's predictions might be compromised. The presence of synthetic patterns in the augmented dataset may lead the ML model to prioritize these patterns during the learning process, potentially overshadowing or missing important patterns present in real samples. Consequently, the model's performance on real data may be hindered, limiting its clinical utility within the context of a CAD system. Further progress is therefore expected in this field to improve the reliability of the systems that use data augmentation.

Algorithms within Generative AI, like GANs and transformers, are likely to be the ones to continue the progress in the creation of synthetic images that complement data obtained from patients. For example, they can already generate standard scans from low-dose scanners [85,100] or provide information similar to what would be acquired through multi-modal integration [60,81]. This reduces both the time patients are exposed to imaging and the overall cost [72]. Transformers, specifically, are expected to have a significant role, although addressing the issue of limited data, possibly through techniques like domain adaptation, is necessary. In summary, we anticipate increasingly comprehensive, transparent, and interpretable CAD systems that integrate information from various sources.

These advances would not be possible without the improvement of

hardware and computing technology, which have played a crucial role in the development and adoption of ML and DL techniques [212,213]. The exponential growth in computational power has enabled efficient training and execution of more complex and deeper models. The use of graphics processing units (GPUs) and distributed systems has further accelerated the performance of ML algorithms, allowing for massive parallel processing. These evolution of both ML and DL algorithms and the hardware and software related to their implementation will make it possible to apply CAD systems that allow real-time decision support to be established in situ. This would enable clinicians to assist in rapid diagnosis and treatment planning, improving efficiency.

## 5. Conclusions

In conclusion, this work has highlighted the significant impact of ML and DL techniques in the field of SPECT and PET imaging. The integration of ML and DL algorithms has led to improvements in diagnostic accuracy by enabling earlier and more precise detection of diseases. These algorithms have demonstrated the ability to learn complex patterns and extract relevant features from SPECT and PET images, leading to enhanced understanding of the condition under analysis (e.g. neurological disorders) and the identification of condition-specific biomarkers.

The use of ML and DL in SPECT and PET imaging has also facilitated the prediction of disease prognosis and tracking of disease progression over time. This information can be invaluable for clinicians in making informed decisions and providing personalized patient care. Furthermore, ML and DL algorithms have contributed to the optimization of image acquisition, reconstruction, and denoising processes, leading to higher-quality images and more accurate interpretations. This has improved the overall reliability and clinical utility of SPECT and PET imaging.

The integration of multimodal data, such as combining PET/SPECT images with MRI, has been made possible through ML and DL techniques, providing a more comprehensive and enhanced understanding of diseases.

Despite the remarkable advancements, several challenges need to be addressed. These include the limited availability of labeled and annotated datasets, the need for interpretability and explainability of DL models, standardization of imaging protocols, and the generalization of models with limited sample sizes.

Looking ahead, there are several promising avenues for future research. Further development of transfer learning techniques, XAI approaches, generative AI and multimodal integration can enhance the accuracy and clinical applicability of ML and DL in SPECT and PET imaging. Additionally, continued efforts in data sharing, standardization, and collaboration between researchers and clinicians will be crucial for advancing the field.

Overall, the applications of ML and DL in SPECT and PET imaging have demonstrated their potential to revolutionize the field, improving diagnostic accuracy, personalized medicine, and patient outcomes. Continued research and technological advancements in this domain hold great promise for advancing the capabilities of SPECT and PET imaging and transforming clinical practice in the future.

## CRedit authorship contribution statement

Carmen Jimenez-Mesa: Conceptualization, Investigation, Writing-Original draft preparation, Writing - Review & Editing. Juan E. Arco: Writing - Review & Editing. Francisco Jesus Martinez-Murcia: Writing - Review & Editing. John Suckling: Supervision, Writing - Review & Editing. Javier Ramirez: Supervision, Writing - Review & Editing. Juan Manuel Gorriz: Conceptualization, Supervision, Writing - Review & Editing.

## Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Acknowledgments

This publication is funded by the grant RYC2021-030875-I funded by MCIN/AEI/10.13039/501100011033 and by the European Union NextGenerationEU/PRTR. The research was also supported by grants PID2022-137629OA-I00 and PID2022-137451OB-I00, funded by the MCIN/AEI/10.13039/501100011033 and by FSE+. Work by C. Jimenez-Mesa and J.E. Arco was funded by Spanish Ministry of Universities under FPU 18/04902 and Margarita Salas respectively.

## References

- [1] S. O'Sullivan, F. Jeanquartier, C. Jean-Quartier, A. Holzinger, D. Shiebler, P. Moon, C. Angione, Developments in AI and machine learning for neuroimaging. Artificial Intelligence and Machine Learning for Digital Pathology, Springer International Publishing, 2020, pp. 307–320.
- [2] K. Suzuki, L. Zhou, Q. Wang, Machine learning in medical imaging, *Pattern Recognit.* 63 (2017) 465–467.
- [3] T. Bateman, Advantages and disadvantages of PET and SPECT in a busy clinical practice, *J. Nucl. Cardiol.* 19 (2012) 3–11.
- [4] S.C. Vaz, F. Oliveira, K. Herrmann, P. Veit-Haibach, Nuclear medicine and molecular imaging advances in the 21st century, *Br. J. Radiol.* 93 (1110) (2020), 20200095.
- [5] E.H. Weissler, T. Naumann, T. Andersson, R. Ranganath, O. Elemento, Y. Luo, D. F. Freitag, J. Benoit, M.C. Hughes, F. Khan, et al., The role of machine learning in clinical research: transforming the future of evidence generation, *Trials* 22 (1) (2021) 1–15.
- [6] H. Arabi, A. AkhavanAllaf, A. Sanaat, I. Shiri, H. Zaidi, The promise of artificial intelligence and deep learning in pet and spect imaging, *Phys. Med.* 83 (2021) 122–137.
- [7] B. Palumbo, et al., Artificial intelligence techniques support nuclear medicine modalities to improve the diagnosis of parkinson's disease and parkinsonian syndromes, *Clin. Transl. Imaging* 9 (2020) 19–35.
- [8] M.P. McBee, O.A. Awan, A.T. Colucci, C.W. Ghobadi, N. Kadom, A.P. Kansagra, S. Tridandapani, W.F. Auffermann, Deep learning in radiology, *Acad. Radiol.* 25 (11) (2018) 1472–1480.
- [9] J. Waring, C. Lindvall, R. Umeton, Automated machine learning: Review of the state-of-the-art and opportunities for healthcare, *Artif. Intell. Med.* 104 (2020), 101822.
- [10] A. Rojas, J. Górriz, J. Ramírez, I. Illán, F. Martínez-Murcia, A. Ortiz, M.G. Rfo, M. Moreno-Caballero, Application of empirical mode decomposition (emd) on datscan spect images to explore parkinson disease, *Expert Syst. Appl.* 40 (7) (2013) 2756–2766.
- [11] H.G. Schnack, R.S. Kahn, Detecting neuroimaging biomarkers for psychiatric disorders: sample size matters, *Front. Psychiatry* 7 (2016), 50.
- [12] D. Castillo-Barnes, J. Ramírez, F. Segovia, F.J. Martínez-Murcia, D. Salas-Gonzalez, J.M. Górriz, Robust ensemble classification methodology for i123-ioflupane spect images and multiple heterogeneous biomarkers in the diagnosis of parkinson's disease, *Front. Neuroinformatics* 12 (2018), 53.
- [13] R.S. Daveau, I. Law, O.M. Henriksen, S.G. Hasselbalch, U.B. Andersen, L. Anderberg, L. Højgaard, F.L. Andersen, C.N. Ladefoged, Deep learning based low-activity pet reconstruction of [11c] pib and [18f] fe-pe2i in neurodegenerative disorders, *Neuroimage* 259 (2022), 119412.
- [14] T.-A. Song, F. Yang, J. Dutta, Noise2void: unsupervised denoising of pet images, *Phys. Med. Biol.* 66 (21) (2021), 214002.
- [15] A. Singh, S. Sengupta, V. Lakshminarayanan, Explainable deep learning models in medical image analysis, *J. Imaging* 6 (6) (2020), 52.
- [16] D.B. Dwyer, P. Falkai, N. Koutsouleris, Machine learning approaches for clinical psychology and psychiatry, *Annu. Rev. Clin. Psychol.* 14 (2018) 91–118.
- [17] M. Alber, A. BuganzaTepole, W.R. Cannon, S. De, S. Dura-Bernal, K. Garikipati, G. Karniadakis, W.W. Lytton, P. Perdikaris, L. Petzold, et al., Integrating machine learning and multiscale modeling—perspectives, challenges, and opportunities in the biological, biomedical, and behavioral sciences, *NPJ Digit. Med.* 2 (1) (2019), 115.
- [18] T. Carlson, E. Goddard, D.M. Kaplan, C. Klein, J.B. Ritchie, Ghosts in machine learning for cognitive neuroscience: Moving from data to theory, *NeuroImage* 180 (2018) 88–100.
- [19] A.N. Nielsen, D.M. Barch, S.E. Petersen, B.L. Schlaggar, D.J. Greene, Machine learning with neuroimaging: evaluating its applications in psychiatry, *Biol. Psychiatry.: Cogn. Neurosci. Neuroimaging* 5 (8) (2020) 791–798.
- [20] C. Davatzikos, Machine learning in neuroimaging: progress and challenges, *NeuroImage* 197 (2019) 652–656.
- [21] Y. LeCun, Y. Bengio, G. Hinton, Deep learning, *Nature* 521 (2015) 436–444.
- [22] L. Khedher, J. Ramírez, J.M. Górriz, A. Brahim, F. Segovia, A. Disease Neuroimaging Initiative, Early diagnosis of Alzheimer's disease based on partial



- least squares, principal component analysis and support vector machine using segmented MRI images, *Neurocomputing* 151 (2015) 139–150.
- [23] S. Vieira, W.H.L. Pinaya, A. Mechelli, Using deep learning to investigate the neuroimaging correlates of psychiatric and neurological disorders: Methods and applications, *Neurosci. Biobehav. Rev.* 74 (2017) 58–75.
  - [24] D. Castillo-Barnes, L. Su, J. Ramirez, D. Salas-Gonzalez, F.J. Martinez-Murcia, I. A. Illan, F. Segovia, A. Ortiz, C. Cruchaga, M.R. Farlow, C. Xiong, N.R. Graff-Radford, P.R. Schofield, C.L. Masters, S. Salloway, M. Jucker, H. Mori, J. Levin, J. M. Gorris, D.I.A.N. (DIAN), Autosomal dominantly inherited alzheimer disease: analysis of genetic subgroups by machine learning, *Inf. Fusion* 58 (2020) 153–167.
  - [25] H. Choi, Deep learning in nuclear medicine and molecular imaging: current perspectives and future directions, *Nucl. Med. Mol. Imaging* 52 (2017) 109–118.
  - [26] A.M. Chekroud, N. Koutsouleris, The perilous path from publication to practice, *Mol. Psychiatry* 23 (2017) 24–25.
  - [27] K. Sharifani, M. Amini, Machine learning and deep learning: a review of methods and applications, *World Inf. Technol. Eng. J.* 10 (07) (2023) 3897–3904.
  - [28] J. Górriz, I. Álvarez-Illán, A. Álvarez-Marquina, J. Arco, M. Atzmueller, F. Ballarín, E. Barakova, G. Bologna, P. Bonomini, G. Castellanos-Dominguez, et al., Computational approaches to explainable artificial intelligence: advances in theory, applications and trends, *Inf. Fusion* 100 (2023), 101945.
  - [29] C.M. Bishop, N.M. Nasrabadi, *Pattern Recognition and Machine Learning*, 4, Springer, 2006.
  - [30] T.G. Dietterich, Ensemble methods in machine learning. Multiple Classifier Systems, Lecture Notes in Computer Science, Springer Berlin Heidelberg, 2000, pp. 1–15.
  - [31] B. Schölkopf, A.J. Smola, F. Bach, et al., Learning with kernels: support vector machines, regularization, optimization, and beyond, MIT Press, 2002.
  - [32] G. Fung, J. Stoeckel, SVM feature selection for classification of SPECT images of Alzheimer's disease using spatial information, *Knowl. Inf. Syst.* 11 (2006) 243–258.
  - [33] X. Gao, C. Chu, Y. Li, P. Lu, W. Wang, W. Liu, L. Yu, The method and efficacy of support vector machine classifiers based on texture features and multi-resolution histogram from 18f-fdg pet-ct images for the evaluation of mediastinal lymph nodes in patients with lung cancer, *Eur. J. Radiol.* 84 (2) (2015) 312–317.
  - [34] L. Breiman, Bagging predictors, *Mach. Learn.* 24 (1996) 123–140.
  - [35] Y. Freund, R.E. Schapire, Experiments with a new boosting algorithm. IJML'96: Proceedings of the Thirteenth International Conference on International Conference on Machine Learning, 1996.
  - [36] L. Breiman, Random forests, *Mach. Learn.* 45 (2001) 5–32.
  - [37] Z. Hu, Y. Wang, X. Zhang, M. Zhang, Y. Yang, X. Liu, H. Zheng, D. Liang, Super-resolution of pet image based on dictionary learning and random forests, *Nucl. Instrum. Methods Phys. Res. Sect. A: Accel., Spectrometers Detect. Assoc. Equip.* 927 (2019) 320–329.
  - [38] S. Mohapatra, T. Swarnkar, J. Das, Deep convolutional neural network in medical image processing. Handbook of Deep Learning in Biomedical Engineering, Elsevier, 2021, pp. 25–60.
  - [39] J. Gu, Z. Wang, J. Kuen, L. Ma, A. Shahroudy, B. Shuai, T. Liu, X. Wang, G. Wang, J. Cai, T. Chen, Recent advances in convolutional neural networks, *Pattern Recognit.* 77 (2018) 354–377.
  - [40] F.J. Martinez-Murcia, J.M. Górriz, J. Ramirez, A. Ortiz, Convolutional neural networks for neuroimaging in Parkinson's disease: Is preprocessing needed? *Int. J. Neural Syst.* 28 (2018), 1850035.
  - [41] A. Krizhevsky, I. Sutskever, G.E. Hinton, ImageNet classification with deep convolutional neural networks, *Commun. ACM* 60 (2017) 84–90.
  - [42] G. Koch, R. Zemel, R. Salakhutdinov, et al., Siamese neural networks for one-shot image recognition, vol. 2. IJML Deep Learning Workshop, Lille, 2015.
  - [43] A.S. Yousif, Z. Omar, U.U. Sheikh, An improved approach for medical image fusion using sparse representation and siamese convolutional neural network, *Biomed. Signal Process. Control* 72 (2022), 103357.
  - [44] J.E. Arco, A. Ortiz, D. Castillo-Barnes, J.M. Górriz, J. Ramirez, Quantifying inter-hemispheric differences in parkinson's disease using siamese networks. Artificial Intelligence in Neuroscience: Affective Analysis and Health Applications: 9th International Work-Conference on the Interplay Between Natural and Artificial Computation, IWINAC 2022, Puerto de la Cruz, Tenerife, Spain, May 31–June 3, 2022, Proceedings, Part I, Springer, 2022, pp. 156–165.
  - [45] J.E. Arco, A. Ortiz, N.J. Gallego-Molina, J.M. Górriz, J. Ramirez, Enhancing multimodal patterns in neuroimaging by siamese neural networks with self-attention mechanism, *Int. J. Neural Syst.* 33 (04) (2023), 2350019.
  - [46] O. Ronneberger, P. Fischer, T. Brox, U-net: Convolutional networks for biomedical image segmentation. Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5–9, 2015, Proceedings, Part III 18, Springer, 2015, pp. 234–241.
  - [47] G. Du, X. Cao, J. Liang, X. Chen, Y. Zhan, Medical image segmentation based on u-net: a review, in: *Journal of Imaging Science & Technology*, 64, 2020.
  - [48] P. Blanc-Durand, A. Van Der Gucht, N. Schaefer, E. Itti, J.O. Prior, Automatic lesion detection and segmentation of 18f-fet pet in gliomas: a full 3d u-net convolutional neural network study, *PLoS One* 13 (4) (2018), e0195798.
  - [49] N.E. Protonotarios, I. Katsamenis, S. Sykiotis, N. Dikaio, G.A. Kastis, S. N. Chatziioannou, M. Metaxas, N. Doulamis, A. Doulamis, A few-shot u-net deep learning model for lung cancer lesion segmentation via pet/ct imaging, *Biomed. Phys. Eng. Express* 8 (2) (2022), 025019.
  - [50] L. An, P. Zhang, E. Adeli, Y. Wang, G. Ma, F. Shi, D.S. Lalush, W. Lin, D. Shen, Multi-level canonical correlation analysis for standard-dose pet image estimation, *IEEE Trans. Image Process.* 25 (7) (2016) 3303–3315.
  - [51] S. Hu, J. Yuan, S. Wang, Cross-modality synthesis from mri to pet using adversarial u-net with different normalization. 2019 International Conference on Medical Imaging Physics and Engineering (ICMIPPE), IEEE, 2019, pp. 1–5.
  - [52] D. Charte, F. Charte, S. García, M.J. delJesus, F. Herrera, A practical tutorial on autoencoders for nonlinear feature fusion: taxonomy, models, software and guidelines, *Inf. Fusion* 44 (2018) 78–96.
  - [53] P.M. Tuan, T.-L. Phan, M. Adel, E. Guedj, N.L. Trung, Autoencoder-based feature ranking for alzheimer disease classification using pet image, *Mach. Learn. Appl.* 6 (2021), 100184.
  - [54] W. Shao, S.P. Rowe, Y. Du, Spectnet: a deep learning neural network for spect image reconstruction, *Ann. Transl. Med.* 9 (9) (2021).
  - [55] X. Wu, L. Bi, M. Fulham, J. Kim, Unsupervised positron emission tomography tumor segmentation via gan based adversarial auto-encoder. 2020 16th International Conference on Control, Automation, Robotics and Vision (ICARCV), IEEE, 2020, pp. 448–453.
  - [56] A. Sanaat, E. Mirsadeghi, B. Razeghi, N. Ginovart, H. Zaidi, Fast dynamic brain pet imaging using stochastic variational prediction for recurrent frame generation, *Med. Phys.* 48 (9) (2021) 5059–5071.
  - [57] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, Y. Bengio, Generative adversarial nets, *Adv. Neural Inf. Process. Syst.* 27 (2014).
  - [58] R. Wang, V. Bashyam, Z. Yang, F. Yu, V. Tassopoulou, S.S. Chintapalli, I. Skampardon, L.P. Sreepada, D. Sahoo, K. Nikita, et al., Applications of generative adversarial networks in neuroimaging and clinical neuroscience, *Neuroimage* (2023), 119898.
  - [59] Y. Wang, B. Yu, L. Wang, C. Zu, D.S. Lalush, W. Lin, X. Wu, J. Zhou, D. Shen, L. Zhou, 3D Conditional generative adversarial networks for high-quality pet image estimation at low dose, *Neuroimage* 174 (2018) 550–562.
  - [60] X. Dong, Y. Lei, T. Wang, K. Higgins, T. Liu, W.J. Curran, H. Mao, J.A. Nye, X. Yang, Deep learning-based attenuation correction in the absence of structural information for whole-body positron emission tomography imaging, *Phys. Med. Biol.* 65 (5) (2020), 055011.
  - [61] J. Zhou, G. Cui, S. Hu, Z. Zhang, C. Yang, Z. Liu, L. Wang, C. Li, M. Sun, Graph neural networks: a review of methods and applications, *AI Open* 1 (2020) 57–81.
  - [62] S. Parisot, S.I. Ktena, E. Ferrante, M. Lee, R. Guerrero, B. Glocker, D. Rueckert, Disease prediction using graph convolutional networks: application to autism spectrum disorder and Alzheimer's disease, *Med. Image Anal.* 48 (2018) 117–130.
  - [63] H. Jiang, P. Cao, M. Xu, J. Yang, O. Zaiane, Hi-gcn: a hierarchical graph convolution network for graph embedding learning of brain network and brain disorders prediction, *Comput. Biol. Med.* 127 (2020), 104096.
  - [64] K. Qin, D. Lei, W.H. Pinaya, N. Pan, W. Li, Z. Zhu, J.A. Sweeney, A. Mechelli, Q. Gong, Using graph convolutional network to characterize individuals with major depressive disorder across multiple imaging sites, *EBioMedicine* 78 (2022).
  - [65] M.-H. Guo, T.-X. Xu, J.-J. Liu, Z.-N. Liu, P.-T. Jiang, T.-J. Mu, S.-H. Zhang, R. R. Martin, M.-M. Cheng, S.-M. Hu, Attention mechanisms in computer vision: a survey, *Comput. Vis. Media* 8 (3) (2022) 331–368.
  - [66] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A.N. Gomez, Ł. Kaiser, I. Polosukhin, Attention is all you need, *Adv. Neural Inf. Process. Syst.* 30 (2017).
  - [67] J. Devlin, M.-W. Chang, K. Lee, K. Toutanova, Bert: Pre-training of deep bidirectional transformers for language understanding, arXiv preprint arXiv: 1810.04805, 2018.
  - [68] A. Radford, K. Narasimhan, T. Salimans, I. Sutskever, et al., Improving language understanding by generative pre-training, 2018.
  - [69] S. Khan, M. Naseer, M. Hayat, S.W. Zamir, F.S. Khan, M. Shah, Transformers in vision: a survey, *ACM Comput. Surv. (CSUR)* 54 (10s) (2022) 1–41.
  - [70] A. Dosovitskiy, L. Beyer, A. Kolesnikov, D. Weissenborn, X. Zhai, T. Unterthiner, M. Dehghani, M. Minderer, G. Heigold, S. Gelly, et al., An image is worth 16×16 words: transformers for image recognition at scale, arXiv preprint arXiv: 2010.11929, 2020.
  - [71] F. Shamshad, S. Khan, S.W. Zamir, M.H. Khan, M. Hayat, F.S. Khan, H. Fu, Transformers in medical imaging: a survey, *Med. Image Anal.* (2023), 102802.
  - [72] S. Watanabe, T. Ueno, Y. Kimura, M. Mishina, N. Sugimoto, Generative image transformer (git): unsupervised continuous image generative and transformable model for [123 i] fp-cit spect images, *Ann. Nucl. Med.* 35 (2021) 1203–1213.
  - [73] W. Tang, F. He, Y. Liu, Y. Duan, Matr: Multimodal medical image fusion via multiscale adaptive transformer, *IEEE Trans. Image Process.* 31 (2022) 5134–5149.
  - [74] K. Gong, E. Berg, S.R. Cherry, J. Qi, Machine learning in pet: from photon detection to quantitative image reconstruction, *Proc. IEEE* 108 (1) (2019) 51–68.
  - [75] F. Müller, D. Schug, P. Hallen, J. Grahe, V. Schulz, A novel doli positioning algorithm for monolithic scintillator crystals in pet based on gradient tree boosting, *IEEE Trans. Radiat. Plasma Med. Sci.* 3 (4) (2018) 465–474.
  - [76] P. Peng, M.S. Judenhofer, A.Q. Jones, S.R. Cherry, Compton pet: a simulation study for a pet module with novel geometry and machine learning for position decoding, *Biomed. Phys. Eng. Express* 5 (1) (2018), 015018.
  - [77] A. Sanaat, H. Zaidi, Depth of interaction estimation in a preclinical pet scanner equipped with monolithic crystals coupled to sips using a deep neural network, *Appl. Sci.* 10 (14) (2020), 4753.
  - [78] G. Jaliparthi, P.F. Martone, A.V. Stolin, R.R. Raylman, Deep residual-convolutional neural networks for event positioning in a monolithic annular pet scanner, *Phys. Med. Biol.* 66 (14) (2021), 145008.
  - [79] X. Chen, B. Zhou, H. Xie, T. Miao, H. Liu, W. Holler, M. Lin, E.J. Miller, R. E. Carson, A.J. Sinusas, et al., Dudos: deep-learning-based dual-domain sinogram synthesis from sparsely sampled projections of cardiac spect, *Med. Phys.* 50 (1) (2023) 89–103.



- [80] S. Lu, J. Tan, Y. Gao, Y. Shi, Z. Liang, Prior knowledge driven machine learning approach for pet sinogram data denoising, vol. 11312. Medical Imaging 2020, SPIE, 2020, pp. 1064–1068. vol. 11312.
- [81] J. Cui, K. Gong, N. Guo, C. Wu, X. Meng, K. Kim, K. Zheng, Z. Wu, L. Fu, B. Xu, et al., Pet image denoising using unsupervised deep learning, Eur. J. Nucl. Med. Mol. Imaging 46 (2019) 2780–2789.
- [82] W. Lu, J.A. Onofrey, Y. Lu, L. Shi, T. Ma, Y. Liu, C. Liu, An investigation of quantitative accuracy for deep learning based denoising in oncological pet, Phys. Med. Biol. 64 (16) (2019), 165019.
- [83] J. Sun, H. Jiang, Y. Du, C.-Y. Li, T.-H. Wu, Y.-H. Liu, B.-H. Yang, G.S. Mok, Deep learning-based denoising in projection-domain and reconstruction-domain for low-dose myocardial perfusion spect, J. Nucl. Cardiol. 30 (3) (2023) 970–985.
- [84] L. Zhou, J.D. Schaefferkoetter, I.W. Tham, G. Huang, J. Yan, Supervised learning with cyclegan for low-dose fdg pet image denoising, Med. Image Anal. 65 (2020), 101770.
- [85] N. AghakhanOlia, A. Kamali-Asl, S. HaririTabrizi, P. Geramifard, P. Sheikhzadeh, S. Farzanefar, H. Arabi, H. Zaidi, Deep learning-based denoising of low-dose spect myocardial perfusion images: quantitative assessment and clinical performance, Eur. J. Nucl. Med. Mol. Imaging (2022) 1–15.
- [86] I. Shiri, P. Ghafarian, P. Geramifard, K.H.-Y. Leung, M. Ghelichoghli, M. Oveisi, A. Rahmim, M.R. Ay, Direct attenuation correction of brain pet images using only emission data via a deep convolutional encoder-decoder (deep-dac), Eur. Radiol. 29 (2019) 6867–6879.
- [87] J. Yang, D. Park, G.T. Gullberg, Y. Seo, Joint correction of attenuation and scatter in image space using deep convolutional neural networks for dedicated brain 18f-fdg pet, Phys. Med. Biol. 64 (7) (2019), 075019.
- [88] H. Xing, T. Wang, X. Jin, J. Tian, J. Ba, H. Jing, F. Li, Direct attenuation correction for 99mtc-3prgd2 chest spect lung cancer images using deep learning, Front. Oncol. 13 (2023), 1165664.
- [89] L. Shi, Y. Lu, N. Dvornik, C.A. Weyman, E.J. Miller, A.J. Sinusas, C. Liu, Automatic inter-frame patient motion correction for dynamic cardiac pet using deep learning, IEEE Trans. Med. Imaging 40 (12) (2021) 3293–3304.
- [90] T. Li, M. Zhang, W. Qi, E. Asma, J. Qi, Motion correction of respiratory-gated pet images using deep learning based image registration framework, Phys. Med. Biol. 65 (15) (2020), 155003.
- [91] L. Chen, C. Shen, Z. Zhou, G. Maquilan, K. Albuquerque, M.R. Folkert, J. Wang, Automatic pet cervical tumor segmentation by combining deep learning and anatomic prior, Phys. Med. Biol. 64 (8) (2019), 085019.
- [92] Q. Lin, M. Luo, R. Gao, T. Li, Z. Man, Y. Cao, H. Wang, Deep learning based automatic segmentation of metastasis hotspots in thorax bone spect images, PLoS One 15 (12) (2020), e0243253.
- [93] T. Wang, Y. Lei, Y. Fu, W.J. Curran, T. Liu, J.A. Nye, X. Yang, Machine learning in quantitative pet: a review of attenuation correction and low-count image reconstruction methods, Phys. Med. 76 (2020) 294–306.
- [94] J.S. Lee, A review of deep-learning-based approaches for attenuation correction in positron emission tomography, IEEE Trans. Radiat. Plasma Med. Sci. 5 (2) (2020) 160–184.
- [95] C. Catana, Motion correction options in pet/mri, in: Seminars in Nuclear Medicine, 45, Elsevier, 2015, pp. 212–223.
- [96] J.E. Iglesias, M.R. Sabuncu, Multi-atlas segmentation of biomedical images: a survey, Med. Image Anal. 24 (1) (2015) 205–219.
- [97] R. Martins, et al., Automatic classification of idiopathic parkinson's disease and atypical parkinsonian syndromes combining [11c]raclopride PET uptake and MRI grey matter morphometry, J. Neural Eng. 18 (2021), 046037.
- [98] C. Panigrahy, A. Seal, C. Gonzalo-Martín, P. Pathak, A.S. Jalal, Parameter adaptive uni-linking pulse coupled neural network based mri–pet/spect image fusion, Biomedical, Signal Process. Control 83 (2023), 104659.
- [99] H. Zhong, D. Huang, J. Wu, X. Chen, Y. Chen, C. Huang, 18f-fdg pet/ct based radiomics features improve prediction of prognosis: multiple machine learning algorithms and multimodality applications for multiple myeloma, BMC Med. Imaging 23 (1) (2023) 87.
- [100] A. Sanaat, I. Shiri, H. Arabi, I. Mainta, R. Nkoulou, H. Zaidi, Deep learning-assisted ultra-fast/low-dose whole-body pet/ct imaging, Eur. J. Nucl. Med. Mol. Imaging 48 (2021) 2405–2415.
- [101] Y. Onozato, T. Iwata, Y. Uematsu, D. Shimizu, T. Yamamoto, Y. Matsui, K. Ogawa, J. Kuyama, Y. Sakairi, E. Kawakami, et al., Predicting pathological highly invasive lung cancer from preoperative [18f] fdg pet/ct with multiple machine learning models, Eur. J. Nucl. Med. Mol. Imaging 50 (3) (2023) 715–726.
- [102] Y. Han, Y. Ma, Z. Wu, F. Zhang, D. Zheng, X. Liu, L. Tao, Z. Liang, Z. Yang, X. Li, et al., Histologic subtype classification of non-small cell lung cancer using pet/ct images, Eur. J. Nucl. Med. Mol. Imaging 48 (2021) 350–360.
- [103] D. Wallis, M. Soussan, M. Lacroix, P. Akl, C. Duboucher, I. Buvat, An [18f] fdg-pet/ct deep learning method for fully automated detection of pathological mediastinal lymph nodes in lung cancer patients, Eur. J. Nucl. Med. Mol. Imaging (2022) 1–8.
- [104] W. Mu, L. Jiang, Y. Shi, I. Tunali, J.E. Gray, E. Katsoulakis, J. Tian, R.J. Gillies, M. B. Schabath, Non-invasive measurement of pd-1 status and prediction of immunotherapy response using deep learning of pet/ct images, J. Immunother. Cancer 9 (6) (2021).
- [105] P. Katiyar, J. Schwenck, L. Frauenfeld, M.R. Divine, V. Agrawal, U. Kohlhofer, S. Gattidis, R. Kontermann, A. Königsrainer, L. Quintanilla-Martínez, et al., Quantification of intratumoural heterogeneity in mice and patients via machine-learning models trained on pet–mri data, Nat. Biomed. Eng. (2023) 1–14.
- [106] J. Song, J. Zheng, P. Li, X. Lu, G. Zhu, P. Shen, An effective multimodal image fusion method using mri and pet for alzheimer's disease diagnosis, Front. Digit. Health 3 (2021), 637386.
- [107] A. Shukla, R. Tiwari, S. Tiwari, Alzheimer's disease detection from fused pet and mri modalities using an ensemble classifier, Mach. Learn. Knowl. Extr. 5 (2) (2023) 512–538.
- [108] M.E. Klontzas, E. Koltsakis, G. Kalarakis, K. Trpkov, T. Papatthomas, A. H. Karantanas, A. Tzortzakakis, Machine learning integrating 99mtc sestamibi spect/ct and radiomics data achieves optimal characterization of renal oncogenic tumors, Cancers 15 (14) (2023) 3553.
- [109] M. Nakajo, M. Jinguji, A. Tani, D. Hirahara, H. Nagano, K. Takumi, T. Yoshiura, Application of a machine learning approach to characterization of liver function using 99m tc-gsa spect/ct, Abdom. Radiol. 46 (2021) 3184–3192.
- [110] A. Feher, K. Pieszk, R. Miller, M. Lemley, A. Shanbhag, C. Huang, L. Miras, Y.-H. Liu, A.J. Sinusas, E.J. Miller, et al., Integration of coronary artery calcium scoring from ct attenuation scans by machine learning improves prediction of adverse cardiovascular events in patients undergoing spect/ct myocardial perfusion imaging, J. Nucl. Cardiol. 30 (2) (2023) 590–603.
- [111] P.J. Slomka, J. Betancur, J.X. Liang, Y. Otaki, L.-H. Hu, T. Sharir, S. Dorbala, M. DiCarli, M.B. Fish, T.D. Ruddy, et al., Rationale and design of the re gistry of f ast myocardial perfusion i maging with ne xt generation spect (refine spect), J. Nucl. Cardiol. 27 (2020) 1010–1021.
- [112] S. Xue, A. Gafita, C. Dong, Y. Zhao, G. Tetteh, B.H. Menze, S. Ziegler, W. Weber, A. Afshar-Oromieh, A. Rominger, et al., Application of machine learning to pretherapeutically estimate dosimetry in men with advanced prostate cancer treated with 177lu-psma i& t therapy, Eur. J. Nucl. Med. Mol. Imaging 49 (12) (2022) 4064–4072.
- [113] M.R. Salmanpour, M. Shamsaei, G. Hajianfar, H. Soltanian-Zadeh, A. Rahmim, Longitudinal clustering analysis and prediction of parkinson's disease progression using radiomics and hybrid machine learning, Quant. Imaging Med. Surg. 12 (2) (2022), 906.
- [114] M.R. Salmanpour, M. Shamsaei, A. Saberi, G. Hajianfar, H. Soltanian-Zadeh, A. Rahmim, Robust identification of parkinson's disease subtypes using radiomics and hybrid machine learning, Comput. Biol. Med. 129 (2021), 104142.
- [115] Y.-D. Zhang, Z. Dong, S.-H. Wang, X. Yu, X. Yao, Q. Zhou, H. Hu, M. Li, C. Jiménez-Mesa, J. Ramirez, et al., Advances in multimodal data fusion in neuroimaging: overview, challenges, and novel orientation, Inf. Fusion 64 (2020) 149–187.
- [116] M. Quarantelli, K. Berkouk, A. Prinster, B. Landeau, C. Svarer, L. Balkay, B. Alfano, A. Brunetti, J.-C. Baron, M. Salvatore, Integrated software for the analysis of brain pet/spect studies with partial-volume-effect correction, J. Nucl. Med. 45 (2) (2004) 192–201.
- [117] R. Polikar, C. Tilley, B. Hillis, C.M. Clark, Multimodal eeg, mri and pet data fusion for alzheimer's disease diagnosis. 2010 Annual International Conference of the IEEE Engineering in Medicine and Biology, IEEE, 2010, pp. 6058–6061.
- [118] A. Vitale, R. Villa, L. Ugga, V. Romeo, A. Stanzione, R. Cuocolo, et al., Artificial intelligence applied to neuroimaging data in parkinsonian syndromes: actuality and expectations, Math. Biosci. Eng. 18 (2021) 1753–1773.
- [119] J.-H. Kang, B. Mollenhauer, C.S. Coffey, J.B. Toledo, D. Weintraub, D.R. Galasko, D.J. Irwin, V. Van Deerlin, A.S. Chen-Plotkin, C. Caspell-Garcia, et al., Csf biomarkers associated with disease heterogeneity in early parkinson's disease: the Parkinson's progression markers initiative study, Acta Neuropathol. 131 (2016) 935–949.
- [120] K. Gwinn, K.K. David, C. Swanson-Fischer, R. Albin, C.S. Hillaire-Clarke, B.-A. Sieber, C. Lungu, F.D. Bowman, R.N. Alcalay, D. Babcock, et al., Parkinson's disease biomarkers: perspective from the ninds parkinson's disease biomarkers program, Biomark. Med. 11 (6) (2017) 451–473.
- [121] S. Saiki, T. Hatano, M. Fujimaki, K.-I. Ishikawa, A. Mori, Y. Oji, A. Okuzumi, T. Fukuhara, T. Koinuma, Y. Imamichi, et al., Decreased long-chain acylcarnitines from insufficient  $\beta$ -oxidation as potential early diagnostic markers for Parkinson's disease, Sci. Rep. 7 (1) (2017), 7328.
- [122] Z. He, X. Zhang, C. Zhao, X. Ling, S. Malhotra, Z. Qian, Y. Wang, X. Hou, J. Zou, W. Zhou, A method using deep learning to discover new predictors from left-ventricular mechanical dyssynchrony for crt response, J. Nucl. Cardiol. 30 (1) (2023) 201–213.
- [123] A. Garg, S. Appel-Cresswell, K. Popuri, M.J. McKeown, M.F. Beg, Morphological alterations in the caudate, putamen, pallidum, and thalamus in parkinson's disease, Front. Neurosci. 9 (2015), 101.
- [124] D. Castillo-Barnes, F.J. Martinez-Murcia, A. Ortiz, D. Salas-Gonzalez, J. Ramírez, J.M. Górriz, Morphological characterization of functional brain imaging by isosurface analysis in Parkinson's disease, Int. J. Neural Syst. 30 (09) (2020), 2050044.
- [125] S. Kebir, Z. Khurshid, F.C. Gaertner, M. Essler, E. Hattingen, R. Fimmers, B. Scheffler, U. Herrlinger, R.A. Bundschuh, M. Glas, Unsupervised consensus cluster analysis of [18f]-fluoroethyl-l-tyrosine positron emission tomography identified textural features for the diagnosis of pseudoprogression in high-grade glioma, Oncotarget 8 (5) (2017), 8294.
- [126] R. Togo, K. Hirata, O. Manabe, H. Ohira, I. Tsujino, K. Magota, T. Ogawa, M. Haseyama, T. Shiga, Cardiac sarcoidosis classification with deep convolutional neural network-based features using polar maps, Comput. Biol. Med. 104 (2019) 81–86.
- [127] Y.J. Bae, B.S. Choi, J.-M. Kim, W.A. Ai, I. Yun, Y.S. Song, Y. Nam, S.J. Cho, J. H. Kim, Deep learning regressor model based on nigrosome mri in parkinson syndrome effectively predicts striatal dopamine transporter-spect uptake, Neuroradiology (2023) 1–9.
- [128] C. Martens, O. Debeir, C. Decaestecker, T. Metens, L. Lebrun, G. Leurquin-Sterk, N. Trotta, S. Goldman, G. Van Simaey, Voxelwise principal component analysis of dynamic [s-methyl-11c] methionine pet data in glioma patients, Cancers 13 (10) (2021), 2342.

- [129] A. Higaki, N. Kawaguchi, T. Kurokawa, H. Okabe, T. Kazatani, S. Kido, T. Aono, K. Matsuda, Y. Tanaka, S. Hosokawa, et al., Content-based image retrieval for the diagnosis of myocardial perfusion imaging using a deep convolutional autoencoder, *J. Nucl. Cardiol.* 30 (2) (2023) 540–549.
- [130] C. Jimenez-Mesa, D. Castillo-Barnes, J.E. Arco, F. Segovia, J. Ramirez, J. M. Górriz, Analyzing statistical inference maps using MRI images for Parkinson's disease. *Artificial Intelligence in Neuroscience: Affective Analysis and Health Applications*, Springer International Publishing, 2022, pp. 166–175.
- [131] D. Moreno-Ajona, E. Prieto, F. Grisanti, I. Esparragosa, L. SánchezOrdúz, J. GállegoPérez-Laraya, J. Arbizu, M. Riverol, 18f-fdg-pet imaging patterns in autoimmune encephalitis: impact of image analysis on the results, *Diagnostics* 10 (6) (2020), 356.
- [132] Y. Ding, J.H. Sohn, M.G. Kawczynski, H. Trivedi, R. Harnish, N.W. Jenkins, D. Lituev, T.P. Copeland, M.S. Aboian, C. Mari Aparici, et al., A deep learning model to predict a diagnosis of alzheimer disease by using 18f-fdg pet of the brain, *Radiology* 290 (2) (2019) 456–464.
- [133] K. Etminani, A. Soliman, A. Davidsson, J.R. Chang, B. Martínez-Sanchis, S. Byttner, V. Camacho, M. Bauckneht, R. Stegeran, M. Ressler, et al., A 3d deep learning model to predict the diagnosis of dementia with lewy bodies, Alzheimer's disease, and mild cognitive impairment using brain 18f-fdg pet, *Eur. J. Nucl. Med. Mol. Imaging* (2022) 1–22.
- [134] J. Hathaliya, R. Parekh, N. Patel, R. Gupta, S. Tanwar, F. Alqahtani, M. Elghatwary, O. Ivanov, M.S. Raboaca, B.-C. Neagu, Convolutional neural network-based parkinson disease classification using spect imaging data, *Mathematics* 10 (15) (2022), 2566.
- [135] H. Khachnaoui, N. Khelifa, R. Mabrouk, Machine learning for early Parkinson's disease identification within swedd group using clinical and datscan spect imaging features, *J. Imaging* 8 (4) (2022) 97.
- [136] L. Sibille, R. Seifert, N. Avramovic, T. Vehren, B. Spottiswoode, S. Zuehlsdorff, M. Schäfers, 18f-fdg pet/ct uptake classification in lymphoma and lung cancer by using deep convolutional neural networks, *Radiology* 294 (2) (2020) 445–452.
- [137] A. Girard, L. Dercle, H. Vila-Reyes, L.H. Schwartz, A. Girma, M. Bertaux, C. Radulescu, T. Lebre, O. Delcroix, M. Rouanne, A machine-learning-based combination of criteria to detect bladder cancer lymph node metastasis on [18f] fdg pet/ct: a pathology-controlled study, *Eur. Radiol.* 33 (4) (2023) 2821–2829.
- [138] Y.S. Salihoğlu, R.U. Erdemir, B.A. Püren, S. Özdemir, Ç. Uyulan, T.T. Ergüzel, H. O. Tekin, Diagnostic performance of machine learning models based on 18f-fdg pet/ct radiomic features in the classification of solitary pulmonary nodules, *Mol. Imaging Radionucl. Ther.* 31 (2) (2022), 82.
- [139] L. Morvan, T. Carlier, B. Jamet, C. Bailly, C. Bodet-Milin, P. Moreau, F. Kraeber-Bodéré, D. Mateus, Leveraging rsf and pet images for prognosis of multiple myeloma at diagnosis, *Int. J. Comput. Assist. Radiol. Surg.* 15 (2020) 129–139.
- [140] M. Nakajo, K. Kawaji, H. Nagano, M. Jinguji, A. Mukai, H. Kawabata, A. Tani, D. Hirahara, M. Yamashita, T. Yoshiura, The usefulness of machine learning-based evaluation of clinical and pretreatment [18f]-fdg-pet/ct radiomic features for predicting prognosis in hypopharyngeal cancer, *Mol. Imaging Biol.* 25 (2) (2023) 303–313.
- [141] Y. Toyama, M. Hotta, F. Motoi, K. Takanami, R. Minamimoto, K. Takase, Prognostic value of fdg-pet radiomics with machine learning in pancreatic cancer, *Sci. Rep.* 10 (1) (2020), 17024.
- [142] M. Nakajo, M. Jinguji, A. Tani, E. Yano, C.K. Hoo, D. Hirahara, S. Togami, H. Kobayashi, T. Yoshiura, Machine learning based evaluation of clinical and pretreatment 18 f-fdg-pet/ct radiomic features to predict prognosis of cervical cancer patients, *Abdom. Radiol.* (2022) 1–10.
- [143] J.A. Simón-Rodríguez, F.J. Martínez-Murcia, J. Ramírez, D. Castillo-Barnes, J. M. Górriz, Modelling the progression of the symptoms of parkinsons disease using a nonlinear decomposition of 123i fp-cit spect images. *International Work-Conference on the Interplay Between Natural and Artificial Computation*, Springer, 2022, pp. 104–113.
- [144] B. Borroni, D. Anchisi, B. Paghera, B. Vicini, N. Kerrouche, V. Garibotto, A. Terzi, L. Vignolo, M. DiLuca, R. Giubbini, et al., Combined 99mtc-ecd spect and neuropsychological studies in mci for the assessment of conversion to ad, *Neurobiol. Aging* 27 (1) (2006) 24–31.
- [145] F. Segovia, J.M. Górriz, J. Ramírez, I. Álvarez, J.M. Jiménez-Hoyuela, S.J. Ortega, Improved Parkinsonism diagnosis using a partial least squares based approach, *Med. Phys.* 39 (2012) 4395–4403.
- [146] W.D. Penny, K.J. Friston, J.T. Ashburner, S.J. Kiebel, T.E. Nichols, *Statistical Parametric Mapping: the Analysis of Functional Brain Images*, Elsevier, 2011.
- [147] A. Eklund, T.E. Nichols, and H. Knutsson, Cluster failure: why fmri inferences for spatial extent have inflated false-positive rates, *Proceedings of the National Academy of Sciences*, 113, 7900–7905, 2016.
- [148] J.D. Rosenblatt, M. Vink, Y. Benjamini, Revisiting multi-subject random effects in fmri: advocating prevalence estimation, *NeuroImage* 84 (2014) 113–121.
- [149] I. Kim, A. Ramdas, A. Singh, L. Wasserman, Classification accuracy as a proxy for two-sample testing, *Ann. Stat.* 49 (1) (2021) 411–434.
- [150] K. Friston, Ten ironic rules for non-statistical reviewers, *Neuroimage* 61 (4) (2012) 1300–1310.
- [151] P.T. Reiss, Cross-validation and hypothesis testing in neuroimaging: an irenic comment on the exchange between friston and lindquist et al, *NeuroImage* 116 (2015) 248–254.
- [152] M.A. Lindquist, B. Caffo, C. Crainiceanu, Ironing out the statistical wrinkles in “ten ironic rules, *NeuroImage* 81 (2013) 499–502.
- [153] K. Friston, Sample size and the fallacies of classical inference, *NeuroImage* 81 (2013) 503–504.
- [154] J. Gorriz, et al., A connection between pattern classification by, machine learning and statistical inference with the general linear model, *IEEE J. Biomed. Health Inform.* (2021) 1.
- [155] J. Gorriz, C. Jimenez-Mesa, R. Romero-Garcia, F. Segovia, J. Ramirez, D. Castillo-Barnes, F. Martinez-Murcia, A. Ortiz, D. Salas-Gonzalez, I. Illan, C. Puntonet, D. Lopez-Garcia, M. Gomez-Rio, J. Suckling, Statistical agnostic mapping: a framework in neuroimaging based on concentration inequalities, *Inf. Fusion* 66 (2021) 198–212.
- [156] O. Cigdem, I. Beheshti, H. Demirel, Effects of different covariates and contrasts on classification of parkinson's disease using structural MRI, *Comput. Biol. Med.* 99 (2018) 173–181.
- [157] H. Liu, J. Wu, E.J. Miller, C. Liu, Y. Liu, Y.-H. Liu, Diagnostic accuracy of stress-only myocardial perfusion spect improved by deep learning, *Eur. J. Nucl. Med. Mol. Imaging* 48 (2021) 2793–2800.
- [158] J. Zhang, Mining imaging and clinical data with machine learning approaches for the diagnosis and early detection of parkinson's disease, *NPJ Parkinson's Dis.* 8 (1) (2022), 13.
- [159] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, Going deeper with convolutions, In: *Proceedings of the IEEE conference on computer vision and pattern recognition*, 1–9, 2015.
- [160] L. Friedman, G.H. Glover, Report on a multicenter fmri quality assurance protocol, *J. Magn. Reson. Imaging*.: Off. J. Int. Soc. Magn. Reson. Med. 23 (6) (2006) 827–839.
- [161] J.C. Morris, P.S. Aisen, R.J. Bateman, T.L. Benzinger, N.J. Cairns, A.M. Fagan, B. Ghetti, A.M. Goate, D.M. Holtzman, W.E. Klunk, E. McDade, D.S. Marcus, R. N. Martins, C.L. Masters, R. Mayeux, A. Oliver, K. Quaid, J.M. Ringman, M. N. Rossor, S. Salloway, P.R. Schofield, N.J. Selsor, R.A. Sperling, M.W. Weiner, C. Xiong, K.L. Moulder, V.D. Buckles, Developing an international network for Alzheimer's research: the dominantly inherited alzheimer network, *Clin. Investig.* 2 (2012) 975–984.
- [162] C. Jimenez-Mesa, J. Ramirez, J. Suckling, J. Vöglein, J. Levin, J.M. Gorriz, A non-parametric statistical inference framework for deep learning in current neuroimaging, *Inf. Fusion* 91 (2023) 598–611.
- [163] A.M. Fagan, C. Xiong, M.S. Jasielec, R.J. Bateman, A.M. Goate, T.L.S. Benzinger, B. Ghetti, R.N. Martins, C.L. Masters, R. Mayeux, J.M. Ringman, M.N. Rossor, S. Salloway, P.R. Schofield, R.A. Sperling, D. Marcus, N.J. Cairns, V.D. Buckles, J. H. Ladenson, J.C. Morris, D.M. Holtzman, Longitudinal change in CSF biomarkers in autosomal-dominant Alzheimer's disease, *Sci. Transl. Med.* 6 (2014) 226ra30.
- [164] D. Szucs, J.P. Ioannidis, Sample size evolution in neuroimaging research: an evaluation of highly-cited studies (1990–2012) and of latest practices (2017–2018) in high-impact journals, *NeuroImage* 221 (2020), 117164.
- [165] K.S. Button, J.P.A. Ioannidis, C. Mokrysz, B.A. Nosek, J. Flint, E.S.J. Robinson, M. R. Munafò, Power failure: why small sample size undermines the reliability of neuroscience, *Nat. Rev. Neurosci.* 14 (2013) 365–376.
- [166] J.E. Desmond, G.H. Glover, Estimating sample size in functional mri (fmri) neuroimaging studies: statistical power analyses, *J. Neurosci. Methods* 118 (2) (2002) 115–128.
- [167] C.S. Carter, T.A. Lesh, D.M. Barch, Thresholds, power, and sample sizes in clinical neuroimaging, *Biol. Psychiatry: Cogn. Neurosci. Neuroimaging* 1 (2) (2016) 99–100.
- [168] G. Varoquaux, Cross-validation failure: Small sample sizes lead to large error bars, *Neuroimage* 180 (2018) 68–77.
- [169] P.A. Bandettini, *fMRI, ch. Twenty-six Controversies and Challenges*, MIT Press, 2020, pp. 163–214.
- [170] D.C.V. Essen, S.M. Smith, D.M. Barch, T.E. Behrens, E. Yacoub, K. Ugurbil, The WU-minn human connectome project: an overview, *NeuroImage* 80 (2013) 62–79.
- [171] K.L. Miller, F. Alfaro-Almagro, N.K. Bangerter, D.L. Thomas, E. Yacoub, J. Xu, A. J. Bartsch, S. Jbabdi, S.N. Sotiropoulos, J.L.R. Andersson, L. Griffanti, G. Douaud, T.W. Ockell, P. Weale, I. Dragonu, S. Garratt, S. Hudson, R. Collins, M. Jenkinson, P.M. Matthews, S.M. Smith, Multimodal population brain imaging in the UK biobank prospective epidemiological study, *Nat. Neurosci.* 19 (2016) 1523–1536.
- [172] J. Nalepa, M. Marcinkiewicz, M. Kawulok, Data augmentation for brain-tumor segmentation: a review, *Front. Comput. Neurosci.* 13 (2019).
- [173] Q. Lin, C. Cao, T. Li, Z. Man, Y. Cao, H. Wang, dsPIC: a deep spect image classification network for automated multi-disease, multi-lesion diagnosis, *BMC Med. Imaging* 21 (1) (2021) 1–16.
- [174] H. Zhang, Z. Huang, and Z. Lv, Medical image synthetic data augmentation using gan, In: *Proceedings of the 4th International Conference on Computer Science and Application Engineering*, 1–6, 2020.
- [175] J.M. Gorriz, J. Ramirez, J. Suckling, I.A. Illan, A. Ortiz, F.J. Martinez-Murcia, F. Segovia, D. Salas-Gonzalez, S. Wang, Case-based statistical learning: a non-parametric implementation with a conditional-error rate SVM, *IEEE Access* 5 (2017) 11468–11478.
- [176] *Handbook of Statistics*, ch. in: R.R. Krishnaiah, L.N. Kanal (Eds.), Dimensionality and Sample Size Considerations in Pattern Recognition Practice, North-Holland, 1982, pp. 825–855.
- [177] K.L. Moulder, B.J. Snider, S.L. Mills, V.D. Buckles, A.M. Santacruz, R.J. Bateman, J.C. Morris, Dominantly inherited alzheimer network: facilitating research and clinical trials, *Alzheimer's Res. Ther.* 5 (2013) 1–7.
- [178] C. Jimenez-Mesa, I.A. Illan, A. Martin-Martin, D. Castillo-Barnes, F.J. Martinez-Murcia, J. Ramirez, J.M. Gorriz, Optimized one vs one approach in multiclass classification for early Alzheimer's disease and mild cognitive impairment diagnosis, *IEEE Access* 8 (2020) 96981–96993.
- [179] S. Rathore, M. Habes, M.A. Ifikhar, A. Shacklett, C. Davatzikos, A review on neuroimaging-based classification studies and associated feature extraction

- methods for Alzheimer's disease and its prodromal stages, *NeuroImage* 155 (2017) 530–548.
- [180] X. Hao, Y. Bao, Y. Guo, M. Yu, D. Zhang, S.L. Risacher, A.J. Saykin, X. Yao, L. Shen, Multi-modal neuroimaging feature selection with consistent metric constraint for diagnosis of Alzheimer's disease, *Med. Image Anal.* 60 (2020), 101625.
- [181] F.J. Martinez-Murcia, A. Ortiz, J.-M. Gorriz, J. Ramirez, D. Castillo-Barnes, Studying the manifold structure of Alzheimer's disease: a deep learning approach using convolutional autoencoders, *IEEE J. Biomed. Health Inform.* 24 (2020) 17–26.
- [182] M. Khodatars, A. Shoeibi, D. Sadeghi, N. Ghaasemi, M. Jafari, P. Moridian, A. Khadem, R. Alizadehsani, A. Zare, Y. Kong, A. Khosravi, S. Nahavandi, S. Hussain, U.R. Acharya, M. Berk, Deep learning for neuroimaging-based diagnosis and rehabilitation of autism spectrum disorder: a review, *Comput. Biol. Med.* 139 (2021), 104949.
- [183] C. Zhang, S. Bengio, M. Hardt, B. Recht, O. Vinyals, Understanding deep learning (still) requires rethinking generalization, *Commun. ACM* 64 (3) (2021) 107–115.
- [184] Z. Tu, F. He, and D. Tao, Understanding generalization in recurrent neural networks, In: *International Conference on Learning Representations*, 2019.
- [185] M. Ojala, G.C. Garriga, Permutation tests for studying classifier performance. 2009 Ninth IEEE International Conference on Data Mining, IEEE, 2009.
- [186] P. Thanapol, K. Lavangnananda, P. Bouvry, F. Pinel, F. Leprevost, Reducing overfitting and improving generalization in training convolutional neural network (CNN) under limited sample sizes in image recognition, In: *2020 - 5th International Conference on Information Technology (InCIT)*, IEEE, 2020.
- [187] J.D. Rosenblatt, Y. Benjamini, R. Gilron, R. Mukamel, and J.J. Goeman, Better-than-chance classification for signal detection, *Biostatistics*, 2019.
- [188] F. Pereira, T. Mitchell, M. Botvinick, Machine learning classifiers and fMRI: a tutorial overview, *NeuroImage* 45 (2009) S199–S209.
- [189] A. Isaksson, M. Wallman, H. Göransson, M. Gustafsson, Cross-validation and bootstrapping are unreliable in small sample classification, *Pattern Recognit. Lett.* 29 (2008) 1960–1965.
- [190] H. Zhu, T. Li, B. Zhao, Statistical learning methods for neuroimaging data analysis with applications, *Annu. Rev. Biomed. Data Sci.* 6 (2023) 73–104.
- [191] K. Kawaguchi, L.P. Kaelbling, Y. Bengio, Generalization in deep learning, *arXiv preprint arXiv:1710.05468*, 2017.
- [192] B. Neyshabur, S. Bhojanapalli, D. McAllester, N. Srebro, Exploring generalization in deep learning, *Adv. Neural Inf. Process. Syst.* 30 (2017) 5947–5956.
- [193] P. Golland, F. Liang, S. Mukherjee, D. Panchenko, Permutation tests for classification. *Learning Theory*, Springer Berlin Heidelberg, 2005, pp. 501–515.
- [194] A. Gretton, K.M. Borgwardt, M.J. Rasch, B. Schölkopf, A. Smola, A kernel two-sample test, *J. Mach. Learn. Res.* 13 (1) (2012) 723–773.
- [195] E. Olivetti, S. Greiner, P. Avesani, Induction in neuroscience with classification: issues and solutions. *Lecture Notes in Computer Science*, Springer Berlin Heidelberg, 2012, pp. 42–50.
- [196] J.M. Gorriz, J. Ramirez, J. Suckling, On the computation of distribution-free performance bounds: application to small sample sizes in neuroimaging, *Pattern Recognit.* 93 (2019) 1–13.
- [197] V.N. Vapnik, *Statistical Learning Theory*, John Wiley and Sons, Inc, New York, 1998.
- [198] A. Farahani, S. Voghoei, K. Rasheed, and H.R. Arabnia, A brief review of domain adaptation, *Advances in data science and information engineering: proceedings from ICDATA 2020 and IKE 2020*, 877–894, 2021.
- [199] J. Zhang, Z. Ding, W. Li, and P. Ogunbona, Importance weighted adversarial nets for partial domain adaptation, In: *Proceedings of the IEEE conference on computer vision and pattern recognition*, 8156–8164, 2018.
- [200] X. Li, Y. Gu, N. Dvornek, L.H. Staib, P. Ventola, J.S. Duncan, Multi-site fmri analysis using privacy-preserving federated learning and domain adaptation: abide results, *Med. Image Anal.* 65 (2020), 101765.
- [201] M. Yousefnezhad, D. Zhang, A.J. Greenshaw, R. Greiner, Editorial: multi-site neuroimage analysis: domain adaptation and batch effects, *Front. Neuroinformatics* 16 (2022).
- [202] R.R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, D. Batra, Grad-CAM: visual explanations from deep networks via gradient-based localization. 2017 IEEE International Conference on Computer Vision (ICCV), IEEE, 2017.
- [203] M.T. Ribeiro, S. Singh, and C. Guestrin, “Why should I trust you?”: Explaining the predictions of any classifier, In: *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, San Francisco, CA, USA, August 13–17, 2016, 1135–1144, 2016.
- [204] K. Simonyan, A. Vedaldi, and A. Zisserman, Deep inside convolutional networks: visualising image classification models and saliency maps, *arXiv preprint arXiv:1312.6034*, 2013.
- [205] D. Gunning, M. Stefik, J. Choi, T. Miller, S. Stumpf, and G.-Z. Yang, XAI—explainable artificial intelligence, *Science Robotics*, 4, 2019.
- [206] E. Tjoa, C. Guan, A survey on explainable artificial intelligence (xai): toward medical xai, *IEEE Trans. Neural Netw. Learn. Syst.* 32 (11) (2020) 4793–4813.
- [207] B.H. Van der Velden, H.J. Kuijff, K.G. Gilhuijs, M.A. Viergever, Explainable artificial intelligence (xai) in deep learning-based medical image analysis, *Med. Image Anal.* (2022), 102470.
- [208] J. Adebayo, J. Gilmer, M. Muelly, I. Goodfellow, M. Hardt, B. Kim, Sanity checks for saliency maps, *Adv. Neural Inf. Process. Syst.* 31 (2018).
- [209] M. Nazari, A. Kluge, I. Apostolova, S. Klutmann, S. Kimiaei, M. Schroeder, R. Buchert, Explainable ai to improve acceptance of convolutional neural networks for automatic classification of dopamine transporter spect in the diagnosis of clinically uncertain parkinsonian syndromes, *Eur. J. Nucl. Med. Mol. Imaging* (2022) 1–11.
- [210] C. Jiménez-Mesa, J.E. Arco, M. Valenti-Soler, B. Frades-Payo, M.A. Zea-Sevilla, A. Ortiz, M. Ávila-Villanueva, D. Castillo-Barnes, J. Ramírez, T. delSer-Quijano, C. Carnero-Pardo, and J.M. Gorriz, Using explainable artificial intelligence in the clock drawing test to reveal the cognitive impairment pattern, *Int. J. Neural Syst.*, 2023.
- [211] B. Palumbo, M.L. Fravolini, T. Buresta, F. Pompili, N. Forini, P. Nigro, P. Calabresi, N. Tambasco, Diagnostic accuracy of parkinson disease by support vector machine (svm) analysis of 123i-fp-cit brain spect data: implications of putaminal findings and age, *Medicine* 93 (27) (2014).
- [212] B. Huang, Y. Huan, L.D. Xu, L. Zheng, Z. Zou, Automated trading systems statistical and machine learning methods and hardware implementation: a survey, *Enterp. Inf. Syst.* 13 (1) (2019) 132–144.
- [213] K. Berggren, Q. Xia, K.K. Likharev, D.B. Strukov, H. Jiang, T. Mikolajick, D. Querlioz, M. Salinga, J.R. Erickson, S. Pi, et al., Roadmap on emerging hardware and technology for machine learning, *Nanotechnology* 32 (1) (2020), 012002.