Exploring the Usage of U-Net-based Deep Learning Networks for Reduction of Brain PET Scan Time in Alzheimer's Disease Patients

by

Louis Lee

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Applied Science
in
Electrical and Computer Engineering

Waterloo, Ontario, Canada, 2022

Author's Declaration

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

In medical imaging, positron emission tomography (PET) is an imaging technique that uses radiotracers to tag and investigate biological processes. A high-quality PET scan requires a high dosage of such tracers and/or a long scan time in a PET machine, both of which can be sources of discomfort for the patient. In this work, a potential solution based on deep learning is explored, such that PET scans obtained with shorter scan times can be denoised to minimize image quality loss in brain PET scans of Alzheimer's disease patients. Using the open ADNI database, 215 brain PET studies of Alzheimer's disease patients using 18F-Florbetapir radiotracer were obtained. Each study contains four sequences of 5-minute scans, and the average of these scans is taken to be the true noiseless image. 203 studies were used to train a U-Net based neural network, using a single 5-minute scan as input and the full 20-minute scan as the ground truth. The U-Net neural network is 18 convolutional layers deep, separated into an encoder and a decoder, where each 2-layer pair in the encoder is concatenated to its corresponding parallel pair in the decoder. The first convolutional layers in the encoder have 64 filters, with the number of filters doubling at each encoding depth up to 1024. The decoder halves the number of filters at each convolutional layer pair, and a final convolution layer collapses the number of filters down to 1, generating the output image. A pipeline was developed to obtain the quantitative metrics for the network's performance by generating the neural network outputs from single 5-minute scans of the validation studies and obtaining the standardized value uptake ratio values across 56 regions of interest. Combined with a qualitative analysis by a single nuclear medicine physician, the outputs from the neural network with a single 5-minute scan input are comparable both qualitatively and quantitatively to the full 20-minute scans.

Acknowledgements

Data collection and sharing for this project was funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant U01 AG024904) and DOD ADNI (Department of Defense award number W81XWH-12-2-0012). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: AbbVie, Alzheimer's Association; Alzheimer's Drug Discovery Foundation; Araclon Biotech; BioClinica, Inc.; Biogen; Bristol-Myers Squibb Company; CereSpir, Inc.; Cogstate; Eisai Inc.; Elan Pharmaceuticals, Inc.; Eli Lilly and Company; EuroImmun; F. Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc.; Fujirebio; GE Healthcare; IXICO Ltd.; Janssen Alzheimer Immunotherapy Research & Development, LLC.; Johnson & Johnson Pharmaceutical Research & Development LLC.; Lumosity; Lundbeck; Merck & Co., Inc.; Meso Scale Diagnostics, LLC.; NeuroRx Research; Neurotrack Technologies; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; Servier; Takeda Pharmaceutical Company; and Transition Therapeutics. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health (www.fnih.org). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Therapeutic Research Institute at the University of Southern California. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California.

Dedication

This is dedicated to the endless cups of coffee I have consumed during my time here.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

Alzheimer's disease is a neurodegenerative pathology associated with memory and motor loss, affecting millions of senior adults[1]. While its treatment is still under heavy research, one of the biological hallmarks of its early onset is the presence of beta amyloid plaques and misfolded tau protein tangles in the brain [2]. The detection of these compounds is done by injecting the patient with a radiotracer that binds selectively to these compounds, tagged with a radioactive isotope that decays and emits radiation that can be detected using technologies such as single-photon emission computed tomography (SPECT) or positron-emission tomography (PET). However, the atomic nature of this scanning technique means even with the most sensitive detectors, the resulting image from the detectors can often be noisy. Reducing this noise puts the imaging technicians in a dilemma between increasing the radiotracer dosage (which increases the patient's radiation exposure level) or increasing the scan time (which can be uncomfortable for the patient). The former also presents an ethical problem of introducing internal radiation to the patient solely for tomographic purposes.

This thesis presents work done on a possible solution to denoise PET scans using deep learning techniques. It begins by laying the necessary groundwork, including the physical

foundations of radioactive decay, how these decay events are captured by a detector, and how this knowledge can be used to build a selectively targeting medical imaging machine. It also provides a brief background on machine learning and deep learning, and its applications in a tomographic environment.

This thesis then covers the details of the work done, including data acquisition, preprocessing, and construction of a deep learning neural network. It explores different techniques and networks that have been attempted, and their shortcomings. The results then follow, showing both a quantitative and a qualitative assessment of the performance of the trained neural network.

Finally, the thesis ends with a dedicated discussion section, where the author reflects on the work and comments on the decisions made throughout the work. A further commentary on the marriage of tomography and deep learning, as well as an outlook on where the work done can be improved, closes out the thesis.

Chapter 2

Background

This work is heavily based on a widely used medical imaging technique known as positron emission tomography (PET). In PET, selectively binding biological agents are tagged with an isotope that emits radiation, and the technique relies on the detection and localization of these radiation events in volumetric space. PET offers the imaging technicians the freedom to choose the binding agent to selectively target and image desired tissues and biological processes, at the expense of requiring usage of potentially harmful radioactive processes internal to the patient's body. A full appreciation of the challenges presented by the tomography, the origins of such challenges, and the potential solution offered in this thesis necessitates an understanding of the physical processes that govern this technique. To that effect, this chapter of the thesis includes an overview of radioactive decay, its use in the creation of radioactive binding agents (named radiotracers), and the mechanism through which a PET scanner can reconstruct a 3-dimensional model from the detection of these radiation events.

In this chapter, we discuss the background information of the work, such as radioactive decay, Alzheimer's disease, and the methodology underlying PET.

2.1 Underlying Physical Processes

Naturally occurring isotopes, and especially synthetically created ones, are often unstable and will spontaneously decay into more stable forms. The process in which a radioactive decay results in a new isotope is generally classified into two groups[1] —

- **α-decay**, in which two protons and two neutrons are released; and
- **β-decay**, in which a proton converts into a neutron (or vice versa) and emits an electron (positron) in the process to conserve the total charge.

In the case where a proton is changed into a neutron, an electron is emitted to account for the loss of charge. This process is further classified as a β^--decay, and likewise, the conversion of a neutron into a proton to emit a positron is a β^+-decay. Notably, a positron is an antiparticle and an antimatter analogue to the electron. When the two come in contact, they undergo an annihilation process, releasing two photons:

$$\mathrm{e}^- + \mathrm{e}^+ \longrightarrow \gamma + \gamma \tag{2.1}$$

During the electron-positron annihilation, the mass of both particles are each converted to 511keV of energy, following Einstein's famous equation $E = mc^2$. The emitted photons have no net linear momentum, so they travel in opposite directions (Figure 2.1). This has important ramifications in imaging applications, as the detection of these two photons allows the annihilation process to be localized in space. This is especially invaluable in an imaging situation where a chaotic number of photons are being released in various directions; if we can isolate two photons along a line, then we can identify where they originated.

[1] Other forms of decay such as γ-decay exist, but these do not alter the mass number of the isotope, or are daughter processes to the two forms of decay listed.

The decay events are discrete and are characterized by half-lives $t_{1/2}$, the time it takes a sample to decay to half its amount. For applications in imaging, one would need to consider the isotope of choice to ensure that $t_{1/2}$ is not too short that the radioisotope decays to undetectable amounts before the imaging is complete, nor too long that the decay events are infrequent. Some common examples of a radioisotope used in medical imaging are ^{11}C, ^{13}N, and ^{18}F, with $t_{1/2}$ of 20.34, 9.97, and 109.74 minutes respectively. The relatively short half-lives of these isotopes in an imaging environment means that they are often synthesized directly within the lab and immediately applied [3, 4], and also has the benefit that it keeps the radiation exposure time short for the patient.

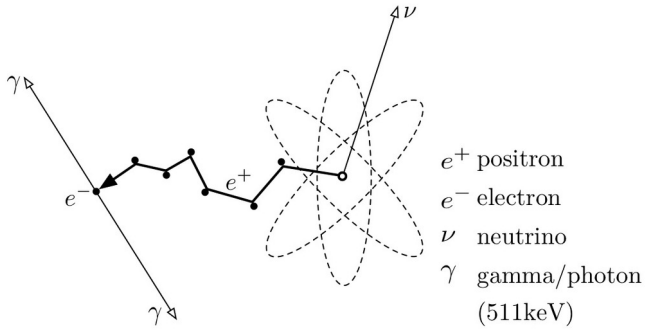


Figure 2.1: **A diagram illustrating the process of electron-positron annihilation.** A positron is emitted from the radioisotope and collides with a free electron, emitting two photons that travel in opposite directions. Adapted from [5].

2.1.1 Radiotracers and Application in Medical Imaging

If observing two photons with opposite momenta allows localization of the annihilation event, then a potential imaging application is to isolate the radioisotopes (and thus events) to a desired region. Radioisotopes provide us with the means of detecting events, but how do we selectively and spatially localize them in an imaging environment?

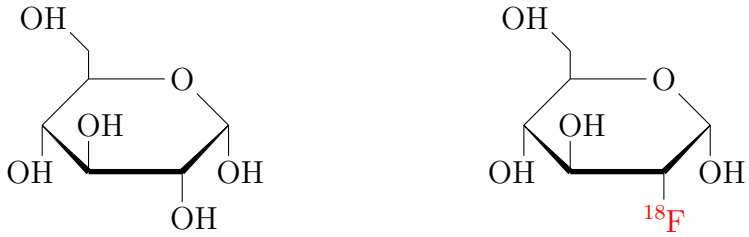


Figure 2.2: **The Haworth projections of (left) glucose and (right) ^{18}F-fluorodeoxyglucose (FDG).** ^{18}F-FDG is a glucose analog, where one of the hydroxyl groups has been swapped to a radioactive ^{18}F isotope (red). Biological processes with glucose uptake will take up FDG as well, and the photons emitted from the ^{18}F decay can be detected following the uptake. This is especially important in imaging tumours, which take up an irregularly high amount of glucose.

In the context of medical imaging, this is achieved through usage of organic ligands and compounds that are metabolically taken up or bound to biological processes. For example, a glucose analog such as fluorodeoxyglucose would selectively bind to tumours with higher glucose uptake (Figure 2.2). By taking such a compound and replacing a group with a radioactive isotope that undergoes β^+-decay, the emitted photons from the isotope's radioactive decay can be regionally identified. These compounds are referred to as radioactive tracers, or radiotracers.

Radiotracers have a profound impact in the field of medical imaging due to their selective binding. Perhaps the most pivotal of its roles is in nuclear imaging, where following the binding of the ligands, the photons emitted by the isotope can be used to localize the binding sites. This allows the detection of the presence and the extent of the biological tissues of interest, and as such radiotracers have seen applications in oncology such as detection of breast cancer [6], brain tumours [7, 8], and leukemia [9]. Outside nuclear imaging, radiotracers have also piqued interest in theranostics, where the usage of controlled radiation from radiotracers is investigated to combine diagnosis and therapy [10, 11].

2.2 Positron Emission Tomography

A radiotracer combines a binding agent with a signal generator. The only remaining task is to capture and detect these signals, and localize them in volumetric space. One attempt at this form of radioactive tomography is the single-photon emission computed tomography (SPECT) that detected the emitted gamma rays from radioisotopes directly. These detectors use a camera that captures the emitted gamma rays at different angles, and uses the data to reconstruct a 3D model. SPECT offers a low-cost solution to tomography, at the cost of poor resolution of the scan [12].

Alternatively, a technician could make use of the fact that a positron-electron annihilation event yields two photons with opposite momenta. A requirement for such a radiotracer is that the radioisotope undergoes a β^+ decay such that it emits a positron, which contacts a free electron to annihilate. Since organisms in nature are not typically made of antimatter, free electrons are abundant in any imaging environment. Chemically, this means that the mass number of the isotope must be less than that on the periodic table. This way, it converts a proton to a neutron, changing the species of the isotope while leaving the mass number the same, and to conserve charge, a positron is emitted. The application where these *positron-emitting* radiotracers are used to obtain a medical image is called positron emission tomography (PET). As mentioned in Section 2.1, the most common positron-emitting isotopes used in PET are ^{11}C, ^{13}N, and ^{18}F. Some common radiotracers used in PET as well as medical imaging in general are listed in Table 2.1.

Unlike SPECT, which uses a mobile camera, a PET detector surrounds the patient with scintillation crystals that catch an incoming 511keV photon from a β^+-decay and breaks it into thousands of photons around 1eV, putting it within visible light range (Figure 2.3). The advantage PET has over SPECT is that with each event, two photons are emitted simultaneously, and thus two crystals will be capturing photons coincidentally. These coincidence events can be used to isolate where the photons originated (Figure 2.4). With

Isotope	Half-life	Tracer	Application
^{15}O	122s	$H_2^{15}O$	Cerebral blood flow
		$^{15}O_2$	Oxygen metabolism
^{18}F	110m	^{18}F-2-deoxyglucose (^{18}FDG)	Glucose metabolism
		^{18}F-6-fluorodopa (^{18}F-dopa)	Dopamine storage
^{11}C	20.5m	^{11}C-SCH23390	Dopamine D_1 receptors
		^{11}C-Raclopride	Dopamine D_2 receptors
		^{11}C-Flumazenil	Central benzodiazepine binding
		^{11}C-Diprenorphine	Opioid binding
		^{11}C-Methionine	Cellular amino acid uptake
		^{11}C(R)-PK11195	Activated microglia
^{123}I	13.2h	Ioflupane	Diagnosis of Parkinson's Disease
		$Na^{123}I$	Diagnosis of hyperthyroidism

Table 2.1: **A list of common radiotracers in medical imaging and the metabolisms they target.** Each radiotracer is tagged with a positron-emitting isotope, indicated on the right column. Adapted from [13].

enough coincidence events, the data obtained can be used to reconstruct a 3D image with greater fidelity than SPECT.

Figure 2.3: **A typical PET scanner facility.** Many modern PET scanners combine PET and computed tomography (CT), which helps image the anatomical structures, to form PET/CT machines capable of identifying and locating abnormalities in bodily functions. Public domain image adapted from [14].

2.2.1 Imperfections and Data Correction

While PET offers a relatively less invasive tomographic technique, it presents us with physical and ethical problems.

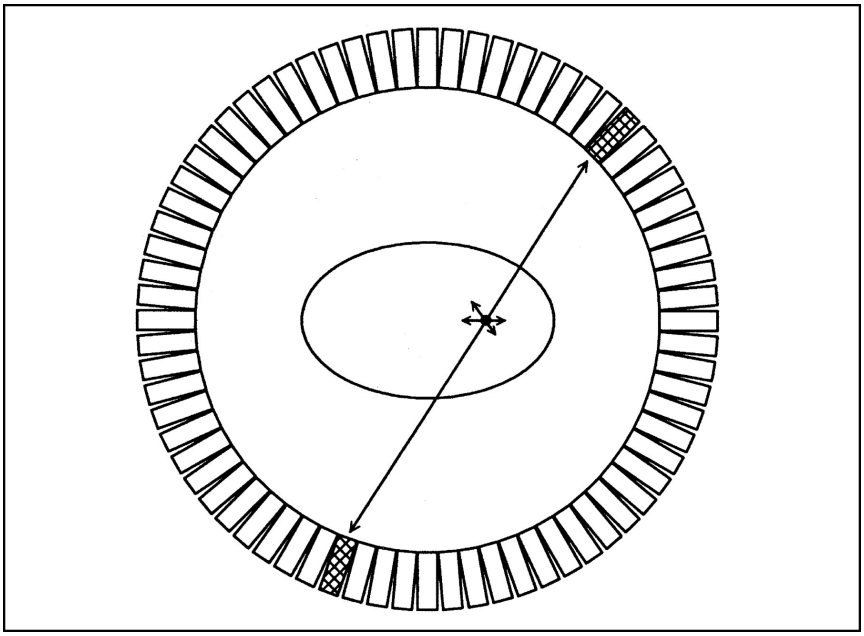


Figure 2.4: **A cross-section of a PET scanner.** To catch the photon emission, many detectors surround the patient. When an annihilation event occurs, the emitted photons travel along a line of response (LOR) until it is detected by two detectors simultaneously. This coincidence event can be used to localize the origin of the photons. Adapted from [15].

The line of response detection of a coincidence event as illustrated in Figure 2.4 is purely ideal (called a *true* coincidence), and realistically, other factors and types of detected signals degrade the overall quality of the PET image (Figure 2.5).

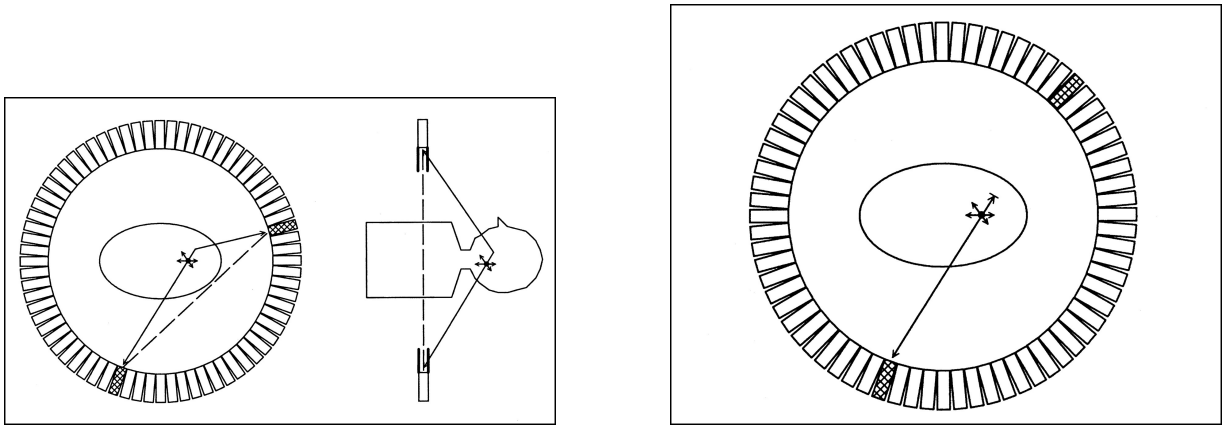
American physicist Arthur H. Compton showed in 1923 that a momentum-carrying photon striking a free electron (or one that is loosely bound to an atom) results in an inelastic collision that changes the kinetic energy and wavelength of the photon [16], a phenomenon now known as Compton scattering. This effect is inevitable in PET scanners, where the emitted photons scatter against the free electrons inside the patient, altering its trajectory before it reaches the detector. The apparent line of response seen from the

scattering is erroneous, and can even appear as if the signal came from outside the body (Figure 2.5a). If the PET scanner counts this coincidence event as usual, then this effect contributes to the overall noise and quality of the PET image. The effects of scattering in PET can be partially remedied through algorithmic approaches [17].

If a photon undergoes numerous scatter interactions (such as if it passes through denser tissues and materials) then it may be *attenuated* and never reach the detector (Figure 2.5b). By nature of PET, the two emitted photons must be unattenuated for their coincidence event to be successfully counted. However, the two photons in opposite directions may go through different densities of medium and interaction within the patient. As with scattering, this results in a noisy image and incorrect coincidence event counts. While the noise is difficult to remedy, the attenuation effects can be corrected by doing a "blank" scan of the patient body and determining the correction factor [15].

An ethical issue that arises with PET is that it relies on radioactivity internal to the body, and at high doses, the ionizing radiation can lead to potential health problems [18]. On the other hand, the emitted radiation is very small, and even the most sensitive machines will yield noisy PET scans if the radiotracer dosage is inadequately low. A typical PET scan procedure attempts to mitigate this issue by retaining a low dosage while using a long scan time. The body is scanned for multiple passes in succession, and the average of the resulting set of images is taken to minimize the environmental noise and artifacts (Figure 2.6). While this procedure results in a higher quality image, the increased scan time in a claustrophobic scanning bed may lead to patient discomfort, as well as introduce motion artifacts that may result in inaccurate registration [19]. Thus, there is merit in exploring methods to convert a noisy single-pass PET scan and generate a clean PET image with reduction in noise.

Earlier attempts at denoising PET scans were algorithmic in nature, involving iterative reconstruction [20, 21] and highly constrained backprojection (HYPR) [22, 23]. Algorithmic



(a) (b)

Figure 2.5: **Erroneous signals potentially detected by a PET scanner. (a)** The emitted photons may scatter before reaching the external ring of detectors, resulting in an incorrect line of response. **(b)** The energy of the photon is attenuated before it reaches a detector, resulting in only one photon from the annihilation event being detected. Adapted from [15].

approaches are ambitious, but in image processing they often require complex modelling of the nature of the noise, leading to computationally intensive algorithms. Furthermore, a digital representation of an image is prone to transformative variations such as rotation and stretching, which algorithms may struggle to capture without the aid of a normalization procedure such as registration.

2.3 Machine Learning

Machine learning (ML) is a term that encompasses a large branch of modern computer science and engineering, with perhaps a loose definition as a consequence. In his original 1959 paper [24], Arthur L. Samuel writes that a "learned" machine must be:

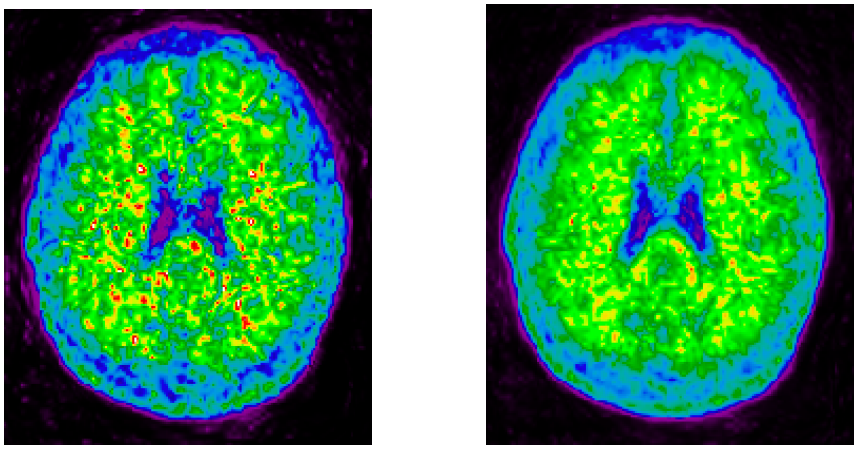


Figure 2.6: **Comparison between (left) a single-pass and (right) a multiple-pass PET scan.** A single pass in a PET machine results in a noisy scan with artifacts. Combining and averaging multiples of these scans results in a reduction in the artifacts. However, this comes at the expense of increasing the PET scan time.

> programmed so that it will learn to play a better game of checkers than can be played by the person who wrote the program.

Machine learning is a subset of artificial intelligence, that makes use of a large database to train a computer for future data. With each datum, the computer adjusts itself to slowly approach the contextually optimal solution[2], which parallels how humans would learn in a natural setting. While ML is still ultimately an algorithm, it differs from the "traditional" algorithms in that with each iteration, the internal hidden representation of the problem is changing slightly. As a result, it is less sensitive to variations in the database than the algorithmic counterpart, and is capable of adapting to new data points. The advent of this technique, paired with the modern era phenomenon of ever-increasing amount of available information, has paved the way for a new potential approach to denoising images.

[2]At least, optimal for the training database given.

2.3.1 Deep Learning

Under the umbrella of machine learning, deep learning is one of many approaches that use large databases to train a machine [25] (Figure 2.7). In deep learning, multiple layers are chained together to form a network of linear weights. The term "deep" refers to the multitude of these layers, which progressively transform the data into a more complex feature space. In the context of image processing, the input image is simply represented as a raw tensor of pixel values. As this tensor flows through the network, however, its representation becomes more abstract, and the deeper layers may be able to identify more abstract features such as edges and segmentation of regions of interest. Thus, deep learning sees many applications in image processing, such as handwritten digit recognition [26], traffic and street recognition for autonomous vehicles [27], and medical image segmentation [28, 29].

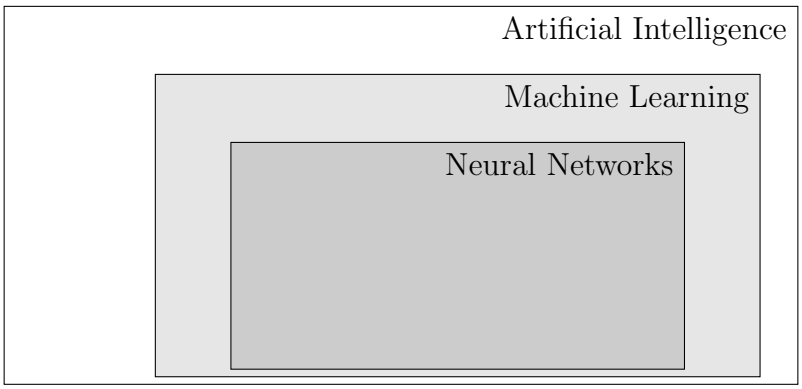


Figure 2.7: **An illustration of the enveloping nature of artificial intelligence and deep learning ecosystem.** Machine learning is a branch of artificial intelligence, and one particular machine learning approach is deep learning, which uses layers of nodes that combine together to form a neural network.

The inputs and outputs can be modelled and represented as an n-dimensional tensor of real values, and their relations can be characterized in the form of linear combinations.

In deep learning, these relations are represented as nodes and chained together to form a neural network; the purpose of the nodes is to store the linear weights in transforming from one tensor to the next, and it is these weights that are optimized during the training process (Figure 2.8). For imaging applications, the input is often represented as a tensor of raw pixel values with $n = 3$, where the third dimension is for the coloured channels. The size of this third dimension can be 1 for grayscale (and effectively making the input tensor 2-dimensional) or 3 for coloured images, where each pixel can be broken down into its red, green, and blue components. The shape of the output tensor is more sensitive to the context of the problem; image classification tasks such as handwritten digit recognition may only require a single valued output that represents the classification, whereas image segmentation and image reconstruction tasks need the output to be a different representation of the input image, so in such cases the output tensor may be the same shape as the input tensor.

A denoising task is, in essence, an image reconstruction task. By using the full, averaged PET scans as the ground truth, one could train a neural network to take a single scan and reconstruct a denoised approximation to the ground truth. This could allow reduction in the scan time by a factor equal to the number of passes (in our case, 4) without significantly compromising the quality of the scans. Previous works have involved using generative adversarial networks (GAN) to implement a generator and a discriminator network that compete against each other to train a network capable of generating denoised images [30, 31]. In this work, we use a simpler model involving only a convolutional neural network, leading to lower algorithmic cost.

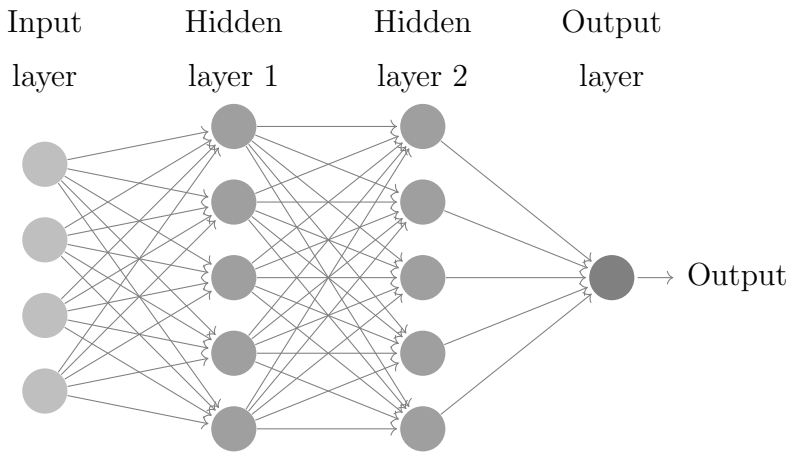


Figure 2.8: **An abstract representation of a neural network.** Each circle represents a node which holds linear weights. At each layer, the node combines the information from the previous layer, combines them into a new tensor using the internal weightings, then passes them forward to the next layer. The actual weightings are hidden from the end user. In this particular network, the nodes are connected densely; each node is connected to every node in the next layer. However, this is not always necessary, and the connections between the layers is one of many freedoms granted to the neural network designer.

2.4 Alzheimer's Disease

Alzheimer's disease (AD) is a neurodegenerative pathology that is characterized by memory loss, and in severe cases, impaired motor functions. It is the most common cause of dementia, affecting more than 5 million Americans [32], and is the sixth leading cause of death in America, fifth among Americans aged 65 or over [33]. Biologically, the brains of AD patients show an abnormally large amount of beta-amyloid (Aβ) plaque buildup, as well as tangles of misfolded tau proteins. While it is unknown if these structures play an active role in AD, identification of these structures is required for an AD diagnosis today [34].

Both the Aβ plaques and the tau protein tangles can be detected through the usage of radiotracers. Aβ plaque buildup occurs outside and in-between neurons, whereas tau protein tangles occur within the neurons. Thus, imaging the tau proteins still presents a challenge [35] but may present superior results to amyloid imaging once those challenges are overcome [36]. Many organic ligands can bind to Aβ, such as florbetaben, flutemetamol, and florbetapir, all of which can be tagged with the radioisotope ^{18}F [37]. These tracers can be used in a PET scanner to bind to the plaque buildup, allowing an early detection of AD.

In this chapter, we covered the physical foundations of a PET scanner, the tomographic mechanisms behind a radioactive ligand, and the shortcomings of such an imaging technique. In the following chapter, we will use AD as a specific application to describe a deep learning approach to potentially reduce scan time in brain PET imaging with minimal loss to the image quality.

Chapter 3

Materials and Methods

This chapter details the data acquisition and preprocessing, as well as the neural network design.

3.1 Data Acquisition

Data used in the preparation of this thesis were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). The ADNI was launched in 2003 as a public-private partnership, led by Principal Investigator Michael W. Weiner, MD. The primary goal of ADNI has been to test whether serial magnetic resonance imaging (MRI), positron emission tomography (PET), other biological markers, and clinical and neuropsychological assessment can be combined to measure the progression of mild cognitive impairment (MCI) and early Alzheimer's disease (AD).

From the ADNI database, 215 ^{18}F-Florbetapir brain PET studies grouped under the AD cohort were retrieved, with 12 of those studies reserved for validation. The average age among the studies is 75 years, distributed among 127 males and 88 females. (Table 3.1).

Statistic	Value
Average age	75
Males	127
Females	88

Table 3.1: **Statistics for the study subjects.** All studies were obtained from ADNI.

Each study includes anonymized metadata regarding the acquisition, as well as the three-dimensional magnetization-prepared rapid acquisition with gradient echo (MPRAGE). The MPRAGE contains four consecutive 5-minute PET scans of the patient, storing the raw detected coincidence events on a per-voxel basis, allowing a model reconstruction of the brain. For the studies, each PET MPRAGE data is accompanied by a brain MRI scan obtained from the same patient.

3.1.1 Preprocessing

The database is stored as a DICOM file format, a medical imaging communication standard that sees widespread use [38]. The DICOM files embed the raw spatial coincidence count data, as well as numerous data fields, such as patient ID and information regarding the PET machine used to obtain the data. However, many of these fields are optional and often unnecessary, and the format has drawn criticism for the lack of consistency in the labeling of these fields [39]. Thus, as part of the preprocessing, the network will ignore all of these data fields, except the fields that directly contribute to the spatial reconstruction of the brain PET scan, such as slice thickness, pixel spacing, and the number of rows and columns (Figure 3.1).

For each study, the raw voxel data is extracted from the MPRAGE data, saving them as individual axial image slices. For consistency between the PET scans, all slices are

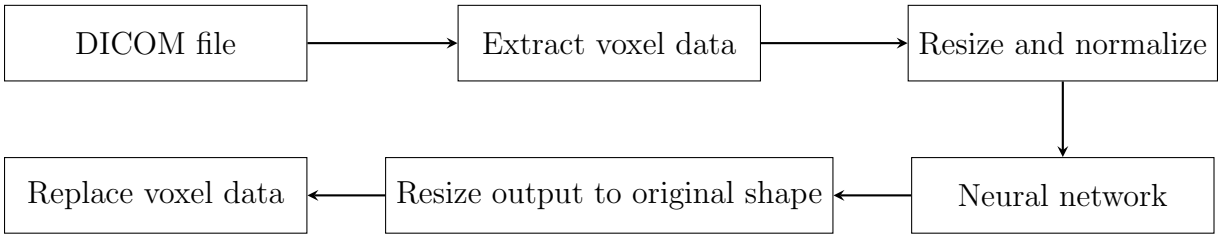


Figure 3.1: **A flowchart outlining the preprocessing and generation of new DICOM files from the original.** The voxel data is extracted from the DICOM files, resized and normalized, and becomes the input to the neural network. The network reconstructs a denoised $256 \times 256 \times 1$ tensor, which is re-shaped to the original DICOM file's dimensions, and the pixel data is replaced, leaving all of the embedded metadata unaltered.

resized to a 256×256 array. The voxel data contains the counts of coincidence events (Section 2.2) which can be in the orders of magnitude too large for optimal neural network training. Thus, each slice is normalized to values between 0 and 1. The normalization factor is determined by the largest voxel value found within the slice. Typically, DICOM pixel values are stored as 16-bit signed integers, and the maximal count per spatial position is 32767. However, through analysis of all of the ADNI data, some DICOM files were found with voxel values exceeding this; for such studies, the voxel values were taken to be 16-bit unsigned integers instead, and used the normalization factor 65535.

The preprocessing procedure ends with the resizing and normalization, to minimize mutation of the raw data. The neural network should ideally train on the most fundamental expression of the PET scans as possible, rather than a reinterpreted representation of it. Since the MRI data is not used during the neural network training and validation, no preprocessing is required for them. After preprocessing, each study yields $4n$ images ready for training and validation, where n is the number of axial slices stored in that PET scan. These slices are augmented through reflections in the x- and y-axes, as well as a half-rotation.

3.2 U-Net

The neural network architecture is designed as a U-Net [40] composed of several convolutional layers (Figure 3.2). During the encoding phase, the network begins with the $256 \times 256 \times 1$ input tensor, an axial slice from a single 5-minute scan, which passes through a pair of 2D convolutional (conv) layers with 64 filters and 3x3 kernel size. Each conv layer is followed by a ReLU activation layer, and after the conv layer couplet, the tensor passes through a batch normalization layer and a 2×2 max pooling layer. The encoding phase is described by 4 such pairs of conv layers, each one decreasing the dimensions of the tensor in the x- and y- directions and doubling the number of filters. At the apex of the U-Net, the tensor has dimensions $16 \times 16 \times 1024$. The decoding phase reverts this tensor back to the original dimensions of the image by concatenating each decoding conv layer with the corresponding parallel encoding conv layer. Unlike the encoding phase, the decoding conv layers double the dimensions of the image while halving the number of filters at each couplet, until the tensor has the dimensions $256 \times 256 \times 64$. Then, a final conv layer with 1 filter yields an output image with identical dimensions to the input, $256 \times 256 \times 1$. The network uses the Adam optimizer for its low computational cost and high efficiency with a large dataset [41]. The hyper-parameters for the optimizer follow the suggested values from [41]. However, raw voxel data in a DICOM format PET scan can yield a very large number of coincidence events, especially in high-dose and high-scan time images, resulting in voxels with very small values following normalization. To account for this, the learning rate α was reduced to slow the adjustment rate of the learned parameters. The hyper-parameters for the optimizer are shown in Table 3.2.

With the input of an axial slice from a single 5-minute scan, the network output is compared with the axial slice of the full 20-minute scan at the same location, and the structural similarity index measure (SSIM) is taken as the error function for the network due to its prominence for distorted imagery [42]. The network had a total number of

Parameter	Value
α	1×10^{-5}
β_1	0.9
β_2	0.999
ϵ	1×10^{-7}

Table 3.2: **Hyper-parameters chosen for the Adam optimizer in the neural network.** With the exception of the learning rate α, all hyper-parameters follow the recommend values from [41].

31,042,369 trainable parameters and trained on Google Colab for 300 epochs with a batch size of 128.

An earlier iteration of this work attempted to use this U-Net as a generator network for a generative adversarial network. However, the high computational overhead and poor quality denoising led us to pursue a simpler model. Approaching simpler solutions has many benefits especially in a medical imaging context; these will be further discussed in Chapter 5.

3.3 Standardized Uptake Value Ratio

Training a neural network, and some may argue computing in general, is a very objective process, where algorithms, equations, and formulae govern the impartial evaluation of results. Thus, metrics such as SSIM and mean squared error see heavy use in neural network error calculation and performance validation; a training neural network has a framework with which to compare its generated output with some ground truth. Yet, these metrics have little meaning for most everyday applications, and in some contexts, they may not even correctly indicate the better solution. Consider the medical imaging

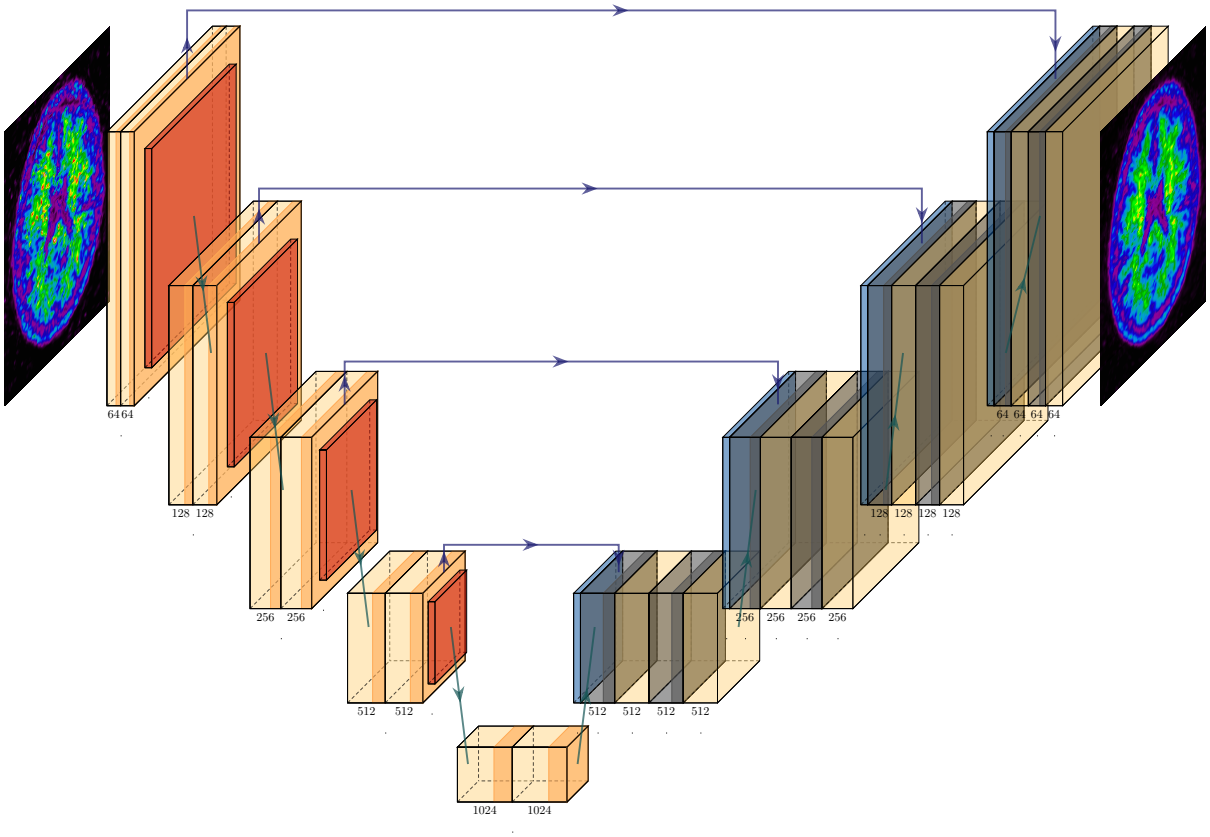


Figure 3.2: **Schematic of U-Net structure.** The encoding phase comprises the left half of the structure, and the decoding phase follows on the right. The network deconstructs the initial image into numerous low-resolution features, then attempts to reconstruct a noiseless version of the input.

context in particular, and what we desire from a denoising neural network. If a generated image (that more closely resembles the ground truth than the raw image according to such metrics) has imaging artifacts or corruptions that affect technicians and physicians to make decisions they would otherwise not, can we truly call it a "better" image?

The quest to find an ideal approach to data analysis in medical imaging is an ongoing struggle [43]. Qualitative metrics such as visual analysis by experts is more reflective of the quality of the images in practice, but is prone to bias and inconsistency among experts. On the other hand, relying solely on pixel-based numerical data for evaluation is myopic and may fail to capture the holistic nature of imaging — that it is a *visual* representation of data. In this work, we use one approach that aims to bridge this gap, by using standardized uptake value ratios (SUVR) [44], which use template masks to calculate the normalized tracer uptake value per pixel in regions of interest (ROI). The SUVR is of particular interest in nuclear medicine because, with the right tracer and ROI, the ratio provides a good indication of presence and stage of pathologies. Furthermore, SUVR metrics see common use in ^{18}F-based PET imaging for Alzheimer's disease [45, 46].

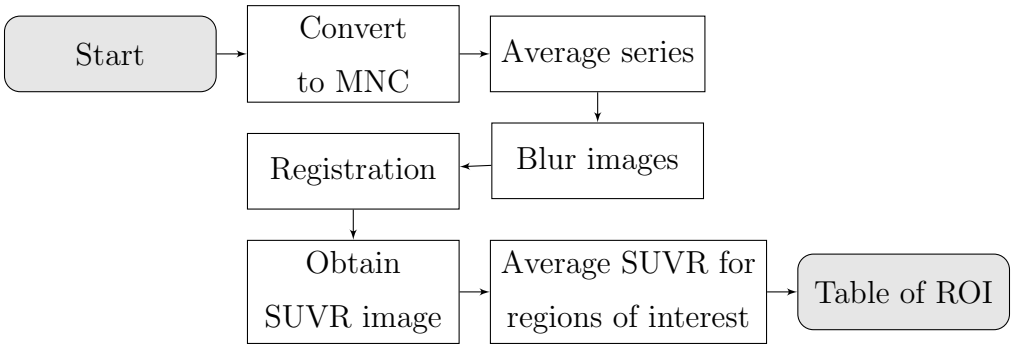


Figure 3.3: **A diagram outlining the pipeline operation.** Each raw PET scan will be parsed on a per-slice basis, passing through this pipeline to calculate the standardized uptake value ratio of that slice for each of the 56 regions of interest.

For obtaining the SUVR metrics, masks corresponding to 56 ROIs were selected. To

impose a normalization standard to fit the masks, a pipeline was developed using the open-source Medical Imaging NetCDF (MINC) tool kit developed at the McConnell Brain Imaging Centre [43]. The purpose of the pipeline (outlined in Figure 3.3) is to register or align the PET scans spatially, apply the masks, calculate the uptake values, and normalize their values relative to all pixels in the image. The pipeline was implemented in bash script for automation; the implementation details follow.

3.3.1 SUVR Pipeline

The ADNI database allows users to download data in DICOM or in Neuroimaging Informatics Technology Initiative (NIfTI) format. The downloaded datum combines all 4 scans into numerous image scans, so the individual scans in a sequence must be separated. The file names contain the sequence number of each image, so the preliminary steps for the pipeline is regex matching to separate these scans. Before we can use the MINC toolkit, the files must be in MNC format. The toolkit provides a `dcm2mnc` command for this purpose. For each PET scan, the `dcm2mnc` command is run on each of the 4 runs, resulting in 4 MNC files that each represent a single run.

The next step is to average the 4 MNC files together for final normalization of the uptake values. This is done through the `mincaverage` command in the toolkit. Typically, the images are blurred at this step to remove minor noise and aid in registration. However, blurring the averaged images led to inaccurate registrations for the majority of the studies, presumably due to ADNI database having already applied a blur to the uploaded images. Thus, the blurring step is skipped in the pipeline.

The images can then properly undergo the registration process, which takes the input image and calculates the mathematical transformations required to match a given template. The transformation from the raw PET scan to the corresponding MRI scan is first calculated, then the linear and nonlinear transformations from the MRI scan to the tem-

plate is obtained. These three transformations are concatenated together, and the PET scan is resampled with these transformations applied. The resulting images are stored as a new MNC file.

From the resulting MNC files, the pipeline can calculate the SUVR of the 56 regions of interest. Using pre-existing masks, the base SUV is calculated for (in our case) AV45, and using `mincmath` tool, the transformed MNC file is normalized to this SUV value to get a SUVR image. Finally, the masks for the ROIs are applied to the SUVR images to obtain a SUVR value for each region, and compiled into a comma-separated value file (Table 4.1).

While the pipeline is straightforward and easy to implement, it is computationally intensive. In particular, the calculation of the linear and non-linear transform matrices bottlenecks the runtime of the script. Furthermore, the variance in each PET scan and attempting to automate their transformation to the same spatial orientation means errors are inevitable. Earlier iterations of the pipeline had terribly warped images resulting from the pipeline (Figure 3.4), and even the later iterations had images that were only partially registered.

The following chapter summarizes the results as obtained from the final version of the neural network, including both quantitative metrics and visual representations. The chapter afterwards discusses the results, using them to reflect on the role of ML in imaging applications and the problems introduced through its usage.

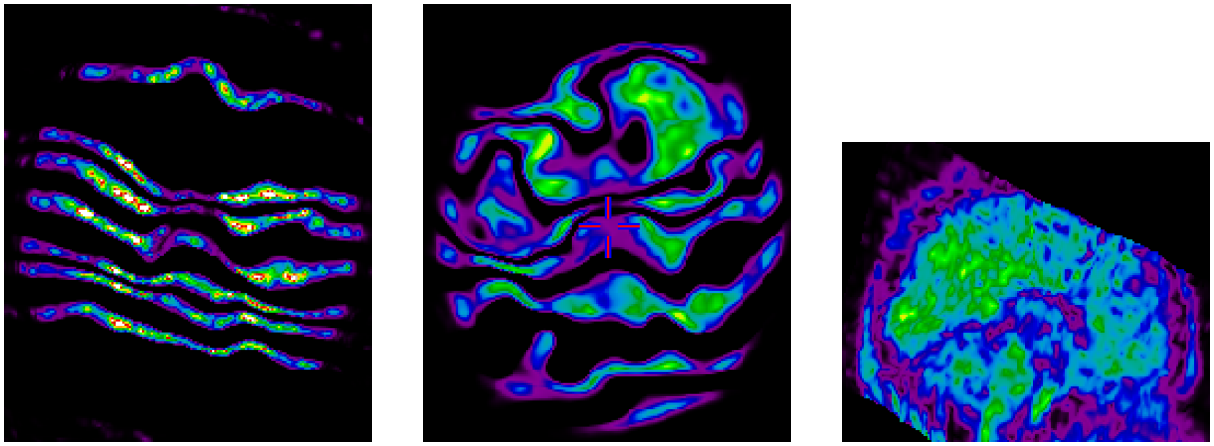


Figure 3.4: **Erroneously registered PET scans from the pipeline.** The automated nature of the pipeline inevitably leads to some PET scans to be not registered properly. Typically, these images must be corrected manually by an expert. **(Left, middle)** The pipeline has many parameters to tweak, such as blurring radius. Earlier iterations of the pipeline led to wildly off-target registrations such as these. **(Right)** Some PET scans result in an incomplete registration, where some slices are "cut off" in the final registration. The result is seen here in the sagittal reconstruction of the axial slices post-registration.

Chapter 4

Results

For quantitative evaluation of the neural network performance, the standardized uptake value ratio (SUVR) normalized to cerebellar grey matter was obtained in 56 regions of interest for each of the 16 validation studies. For each validation study, one of the 5-minute scans was extracted, and a neural network output of this scan was obtained by running each slice through the trained neural network kernel and coalescing the resulting output slices. The full 20-minute scan and the neural network generated output are all run through a pipeline that registers the PET scans to the MRI scans obtained from ADNI, resamples with the registration, and obtains the SUVR from each of the 56 masks that each represent a region of interest (ROI). We also obtained a visual qualitative assessment from a nuclear medicine physician.

SUVR values from 56 ROIs were obtained for the full 20-minute scan as well as from the U-Net-generated image from a single 5-minute scan, and the absolute differences in SUVRs were under 0.1 in most regions. Table 4.1 shows the SUVR values obtained from one of the validation studies. Qualitatively, Figure 4.1 shows a comparison of three sample studies obtained from the validation set, with the full results from the 16 validation targets shown in Figure 4.2. Overall quality appears to be promising when compared to that of

the 20-minute scan. However, dynamic range may be higher and some artifacts may have been emphasized. The next Chapter provides a discussion on these results.

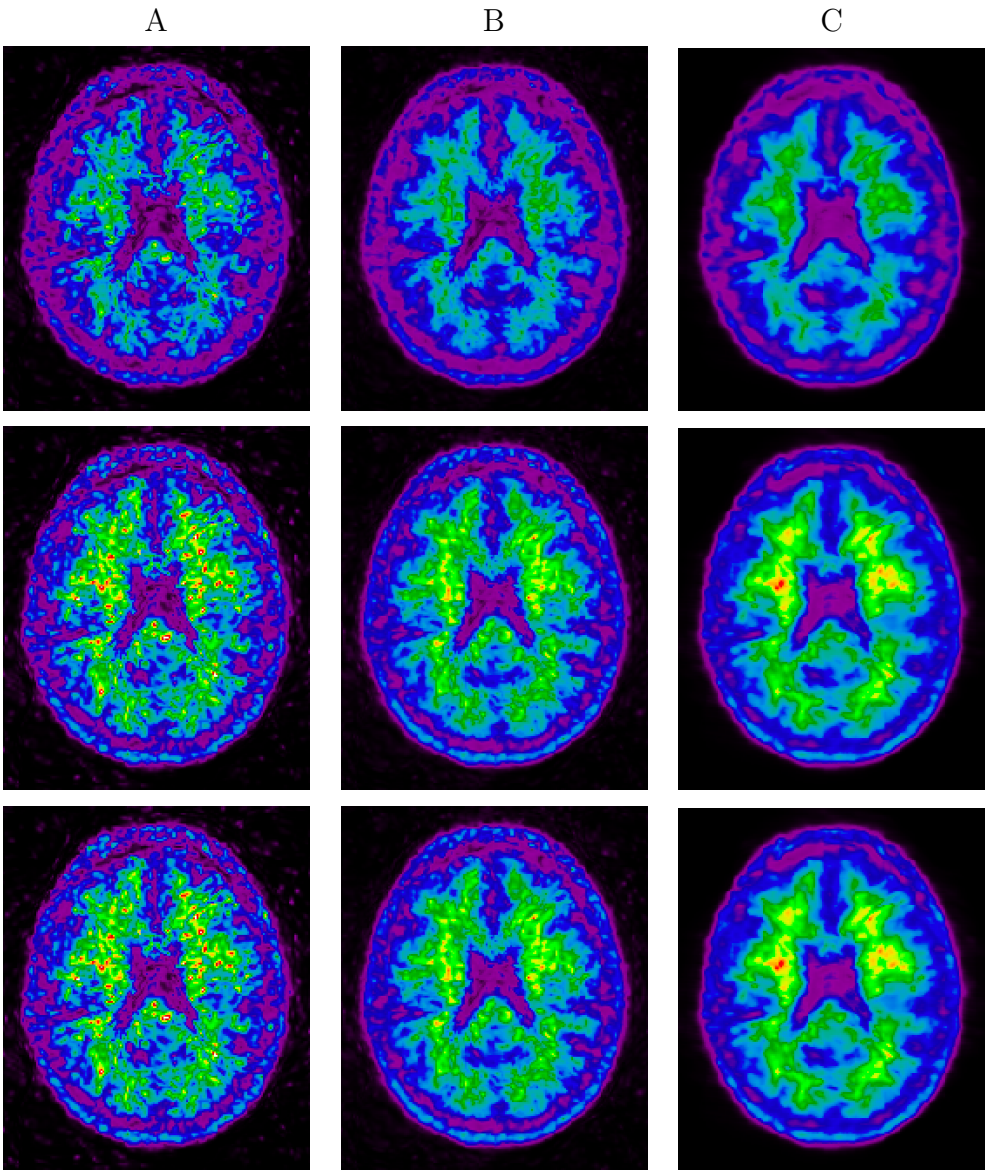
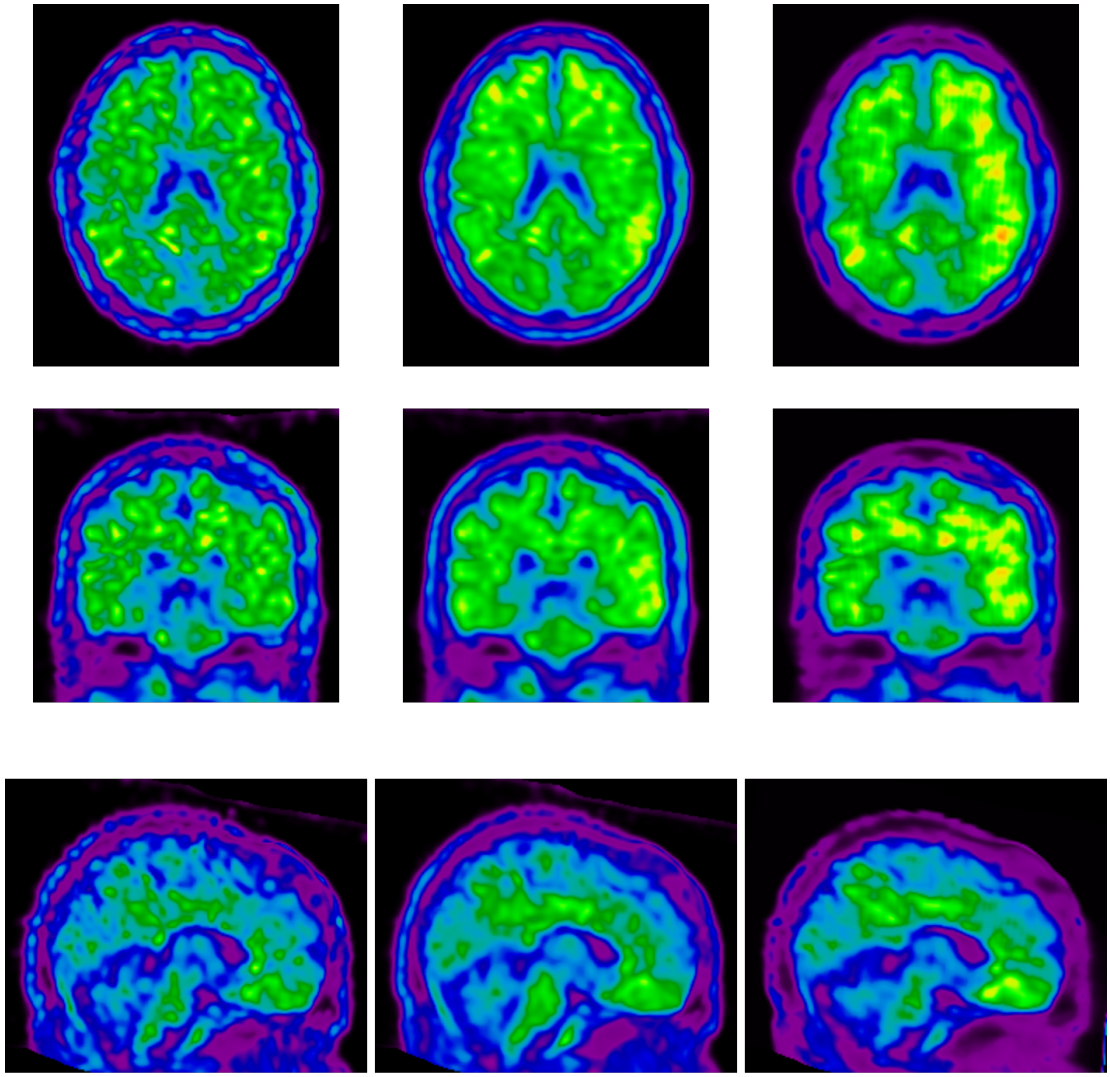
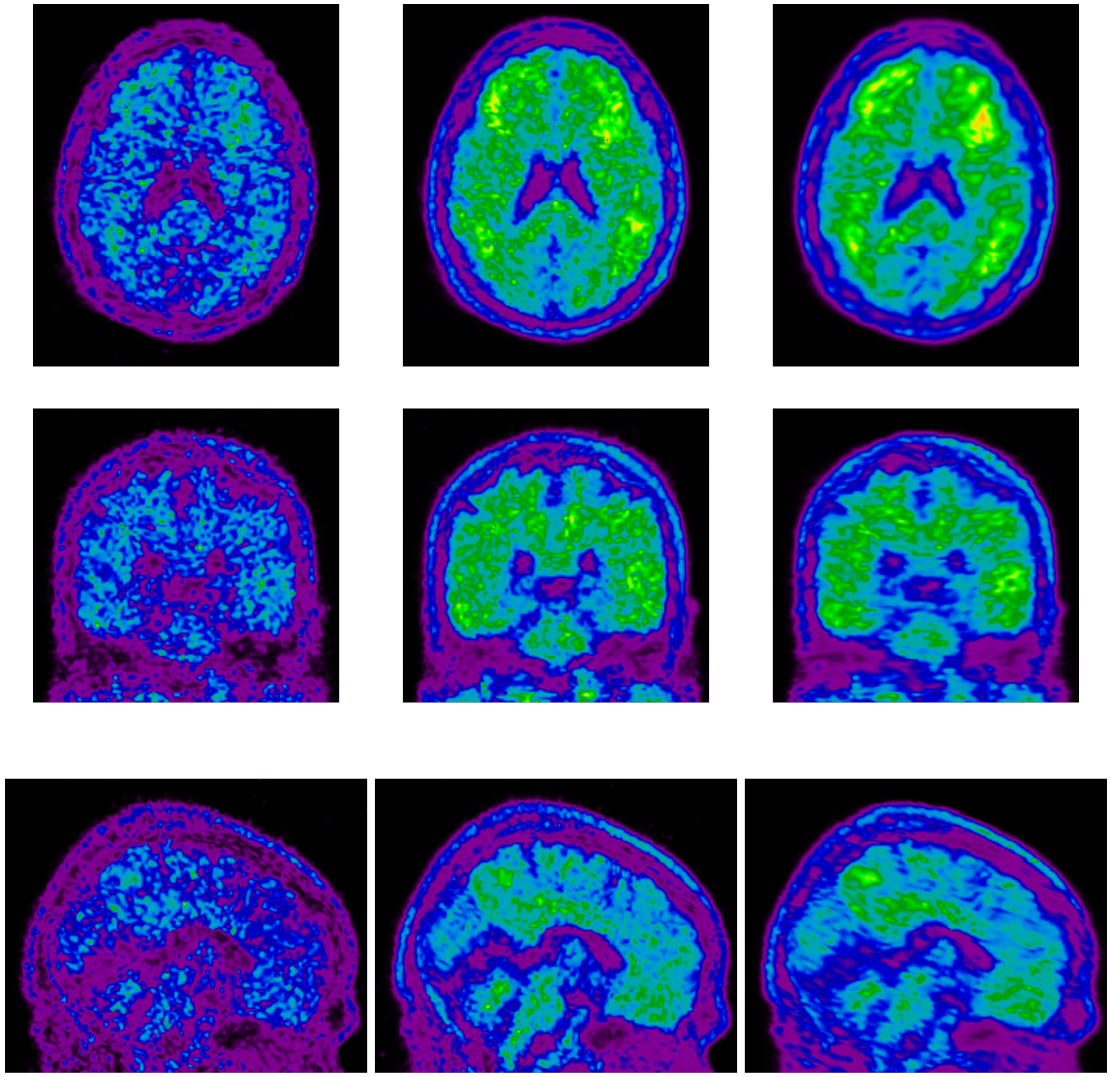


Figure 4.1: **Comparison of 3 different validation subjects and the PET scan slices.** Each row represents a different validation subject, with **(A)** a single scan, **(B)** full scan, and **(C)** a single scan run through the neural network being compared. All slices were obtained at the same spatial location.



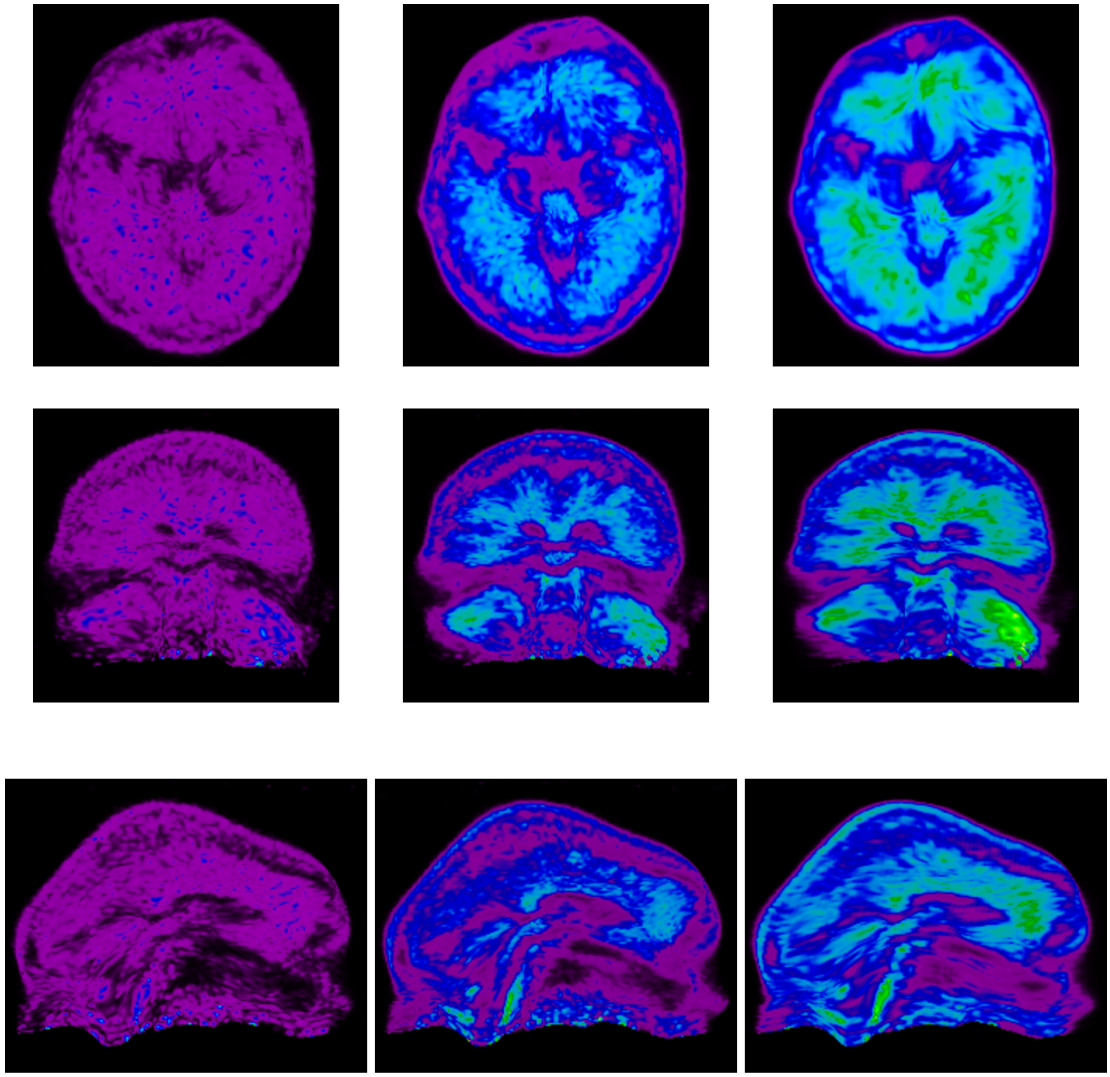
(a)

Figure 4.2: **The 16 validation studies used to visually inspect the performance of the U-Net.** Each figure shows the brain PET scan as obtained from (left column) a single-pass 5-minute PET scan, (middle column) a full 20-minute scan, and (right column) the neural network generated output from using the single-pass 5-minute scan. The resulting figures are also separated into (top) axial, (middle) coronal, and (bottom) sagittal slices.



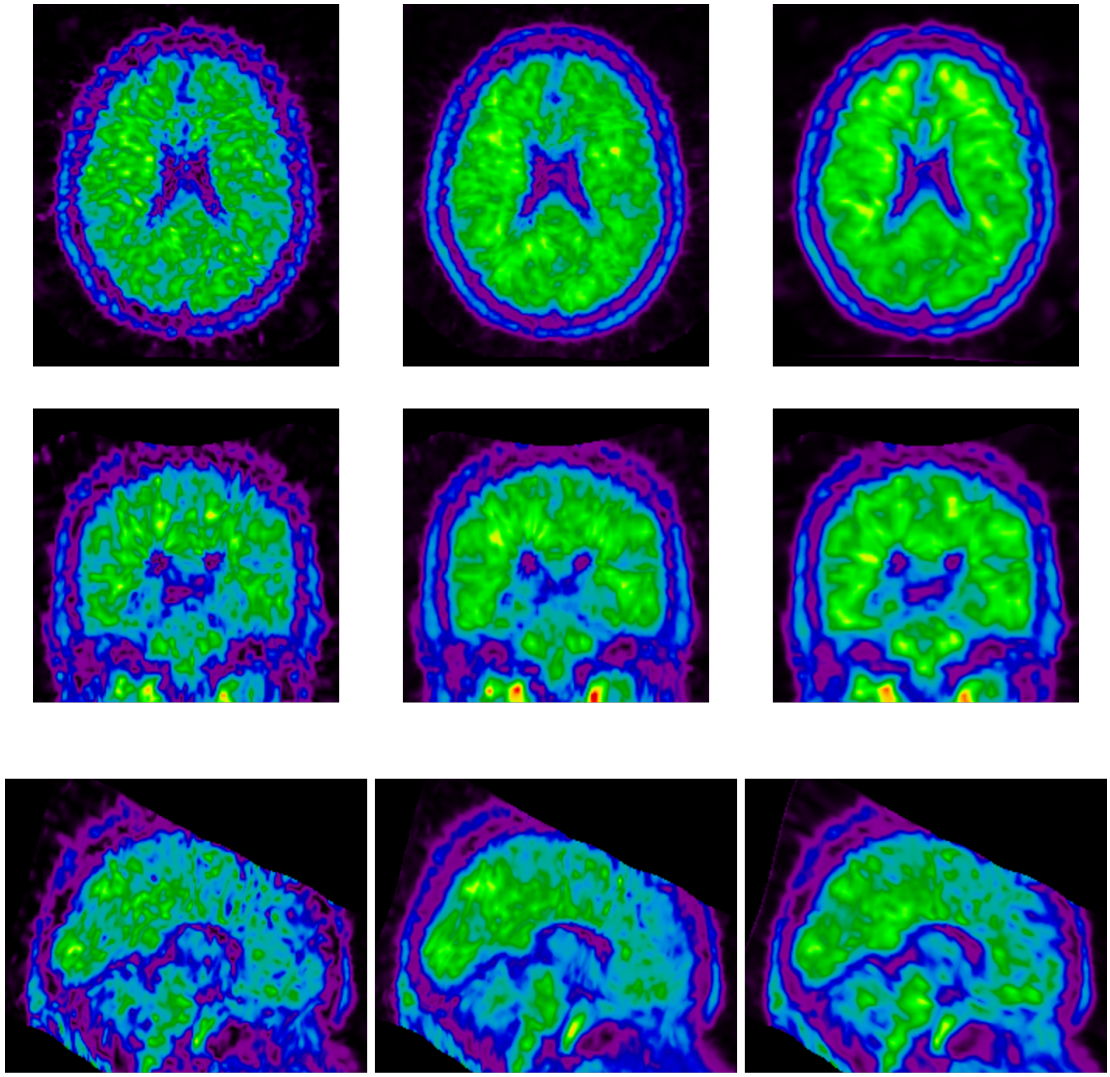
(b)

Figure 4.2: **(continued) The 16 validation studies used to visually inspect the performance of the U-Net.** Each figure shows the brain PET scan as obtained from (left column) a single-pass 5-minute PET scan, (middle column) a full 20-minute scan, and (right column) the neural network generated output from using the single-pass 5-minute scan. The resulting figures are also separated into (top) axial, (middle) coronal, and (bottom) sagittal slices.



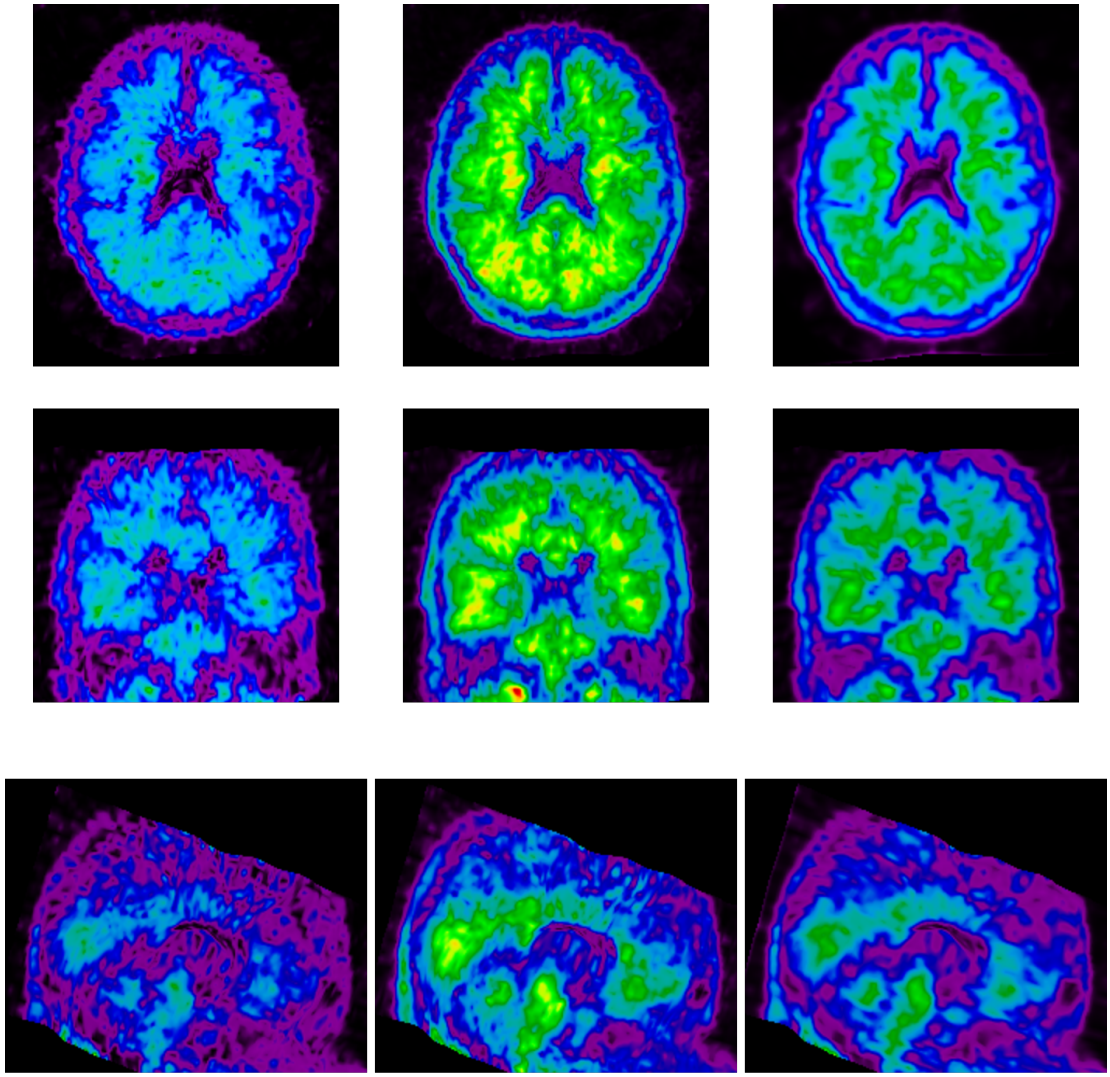
(c)

Figure 4.2: **(continued) The 16 validation studies used to visually inspect the performance of the U-Net.** Each figure shows the brain PET scan as obtained from (left column) a single-pass 5-minute PET scan, (middle column) a full 20-minute scan, and (right column) the neural network generated output from using the single-pass 5-minute scan. The resulting figures are also separated into (top) axial, (middle) coronal, and (bottom) sagittal slices.



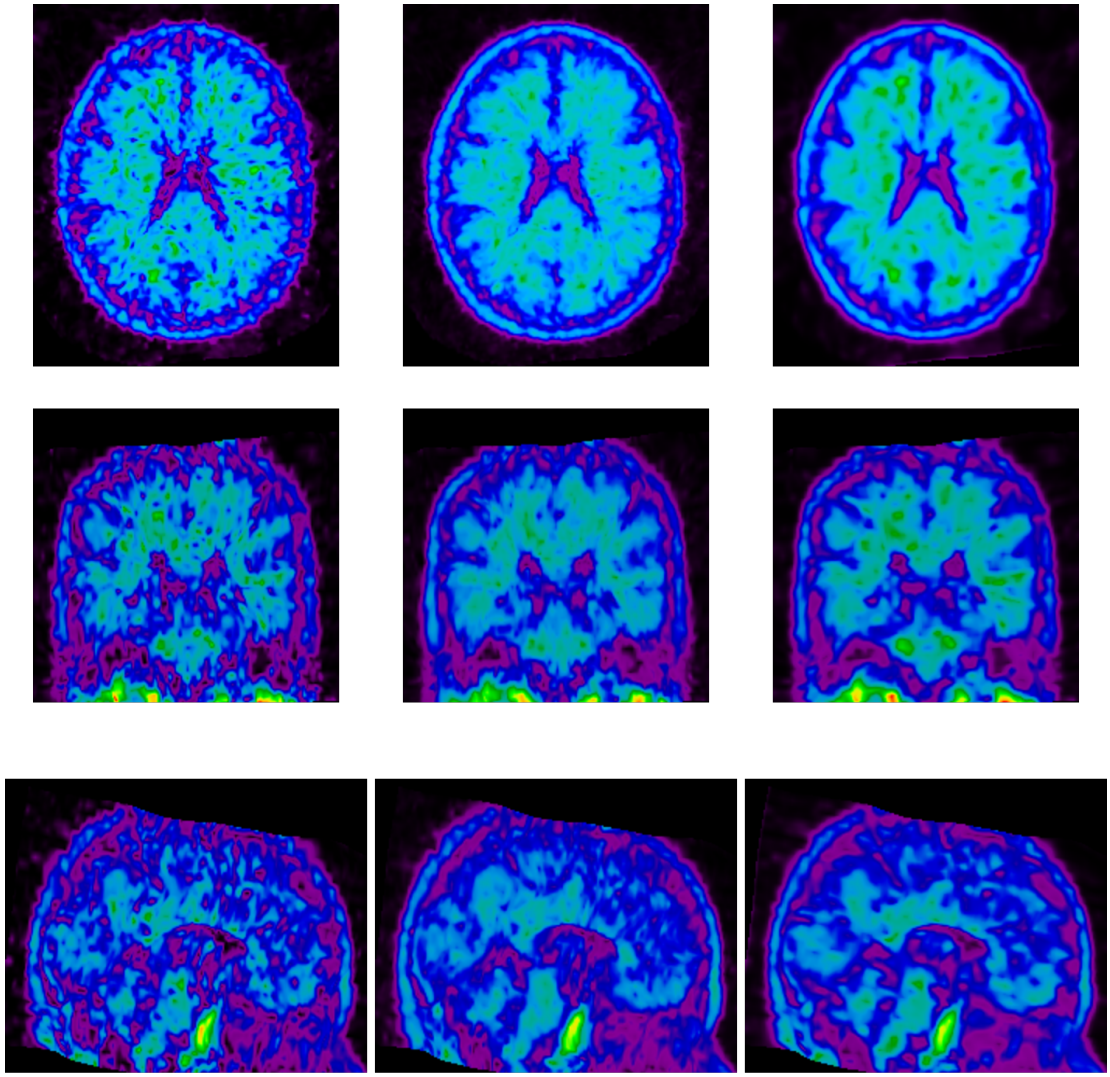
(d)

Figure 4.2: **(continued) The 16 validation studies used to visually inspect the performance of the U-Net.** Each figure shows the brain PET scan as obtained from (left column) a single-pass 5-minute PET scan, (middle column) a full 20-minute scan, and (right column) the neural network generated output from using the single-pass 5-minute scan. The resulting figures are also separated into (top) axial, (middle) coronal, and (bottom) sagittal slices.



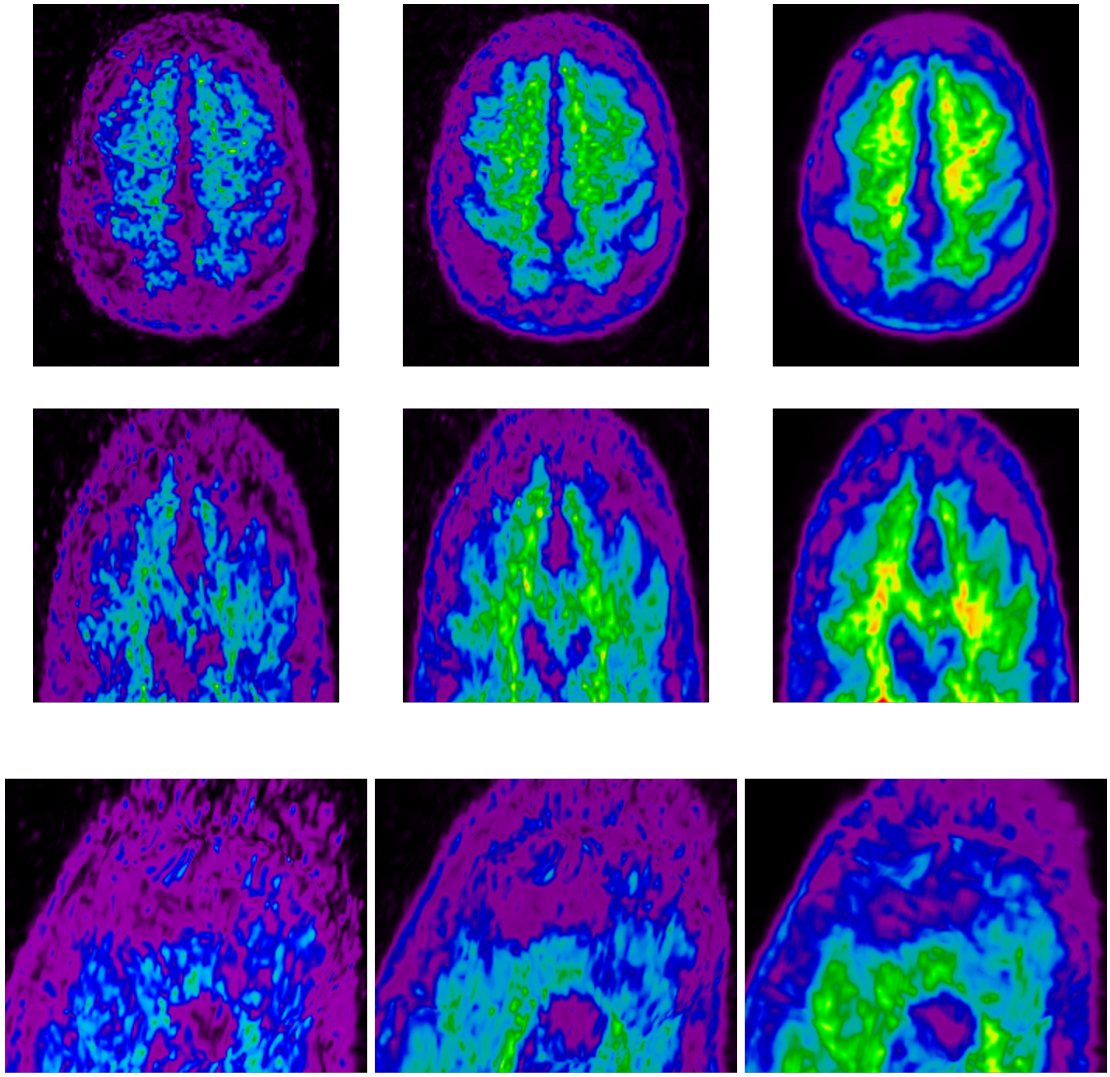
(e)

Figure 4.2: **(continued) The 16 validation studies used to visually inspect the performance of the U-Net.** Each figure shows the brain PET scan as obtained from (left column) a single-pass 5-minute PET scan, (middle column) a full 20-minute scan, and (right column) the neural network generated output from using the single-pass 5-minute scan. The resulting figures are also separated into (top) axial, (middle) coronal, and (bottom) sagittal slices.



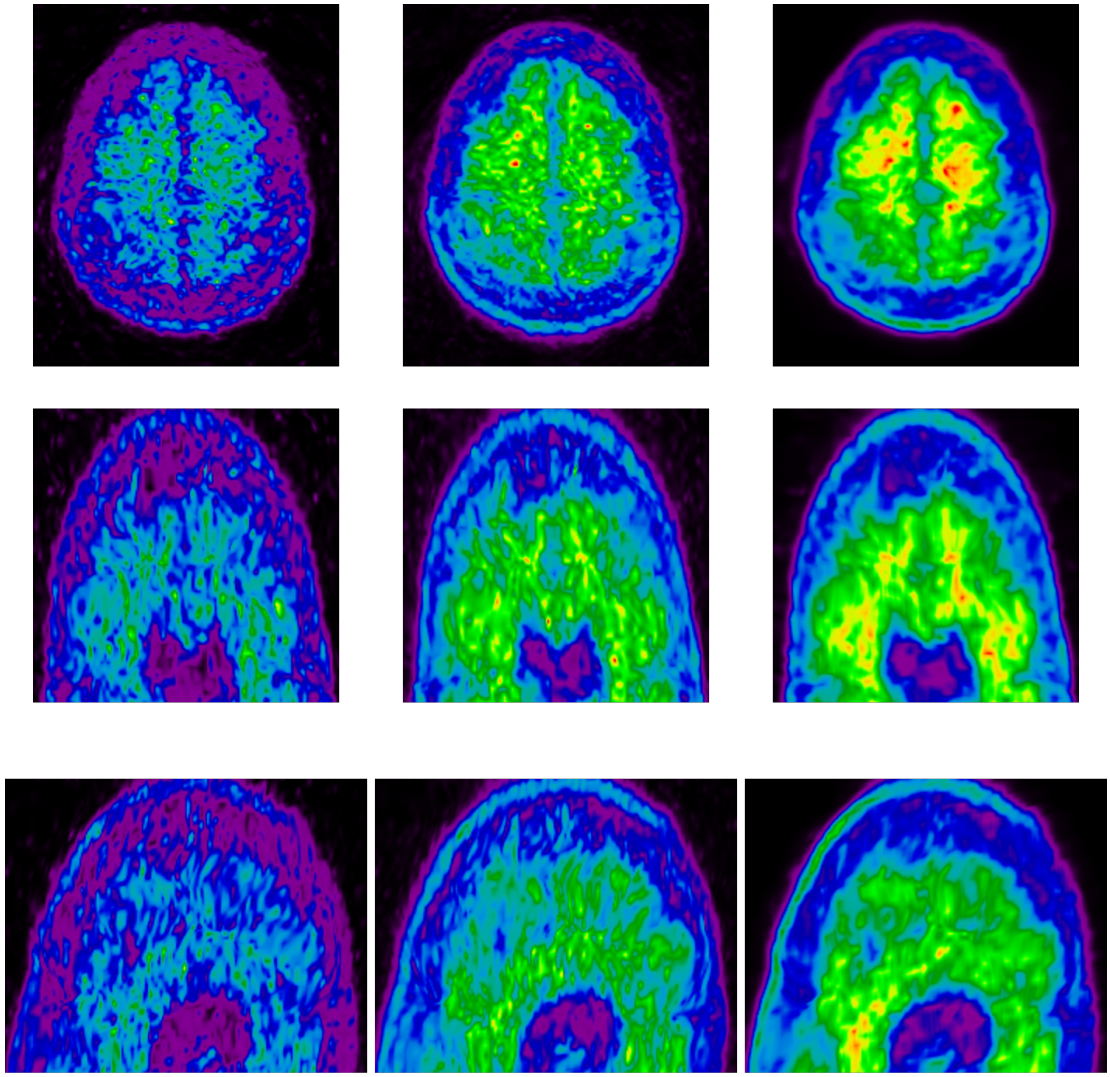
(f)

Figure 4.2: **(continued) The 16 validation studies used to visually inspect the performance of the U-Net.** Each figure shows the brain PET scan as obtained from (left column) a single-pass 5-minute PET scan, (middle column) a full 20-minute scan, and (right column) the neural network generated output from using the single-pass 5-minute scan. The resulting figures are also separated into (top) axial, (middle) coronal, and (bottom) sagittal slices.



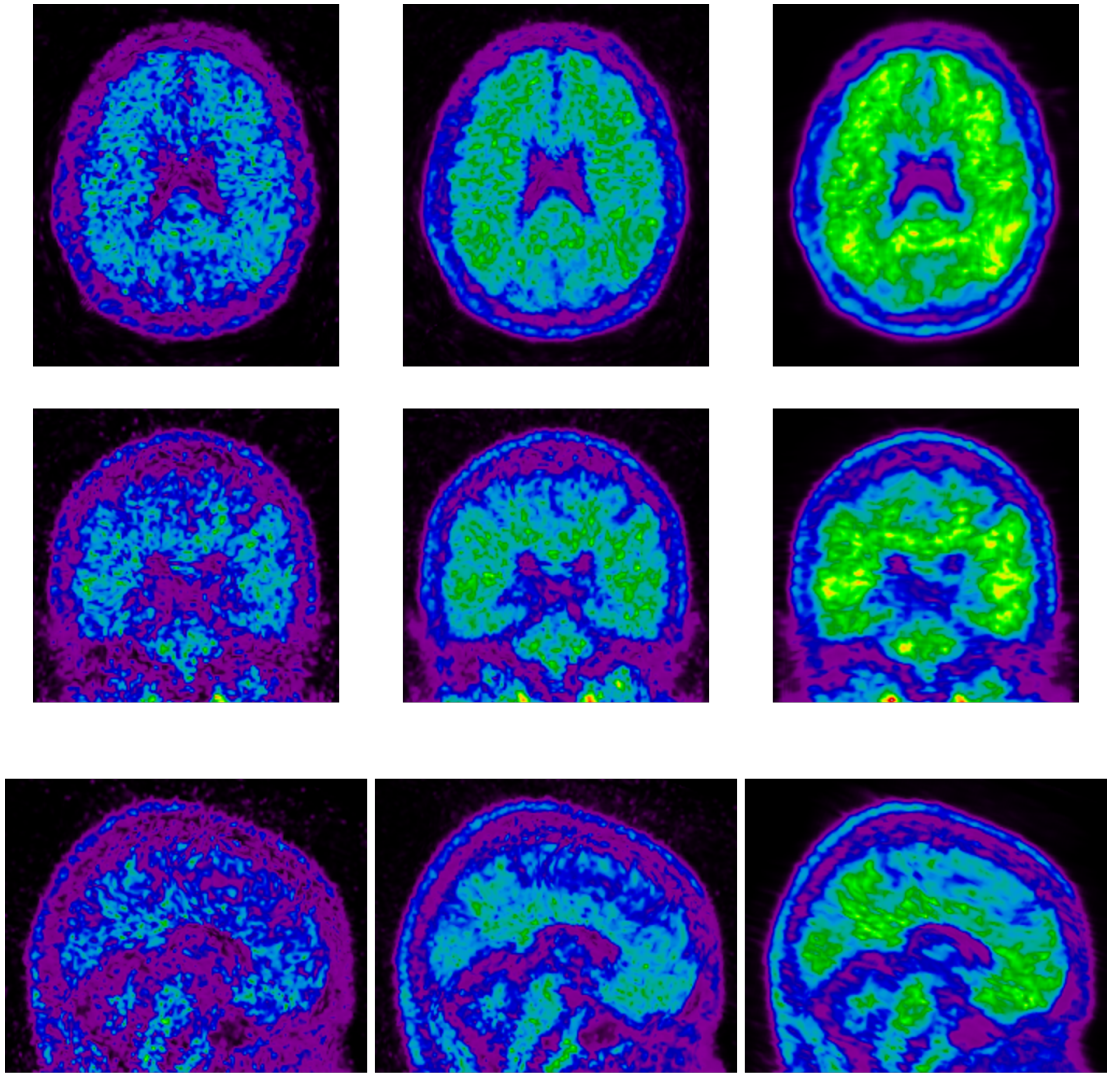
(g)

Figure 4.2: **(continued) The 16 validation studies used to visually inspect the performance of the U-Net.** Each figure shows the brain PET scan as obtained from (left column) a single-pass 5-minute PET scan, (middle column) a full 20-minute scan, and (right column) the neural network generated output from using the single-pass 5-minute scan. The resulting figures are also separated into (top) axial, (middle) coronal, and (bottom) sagittal slices.



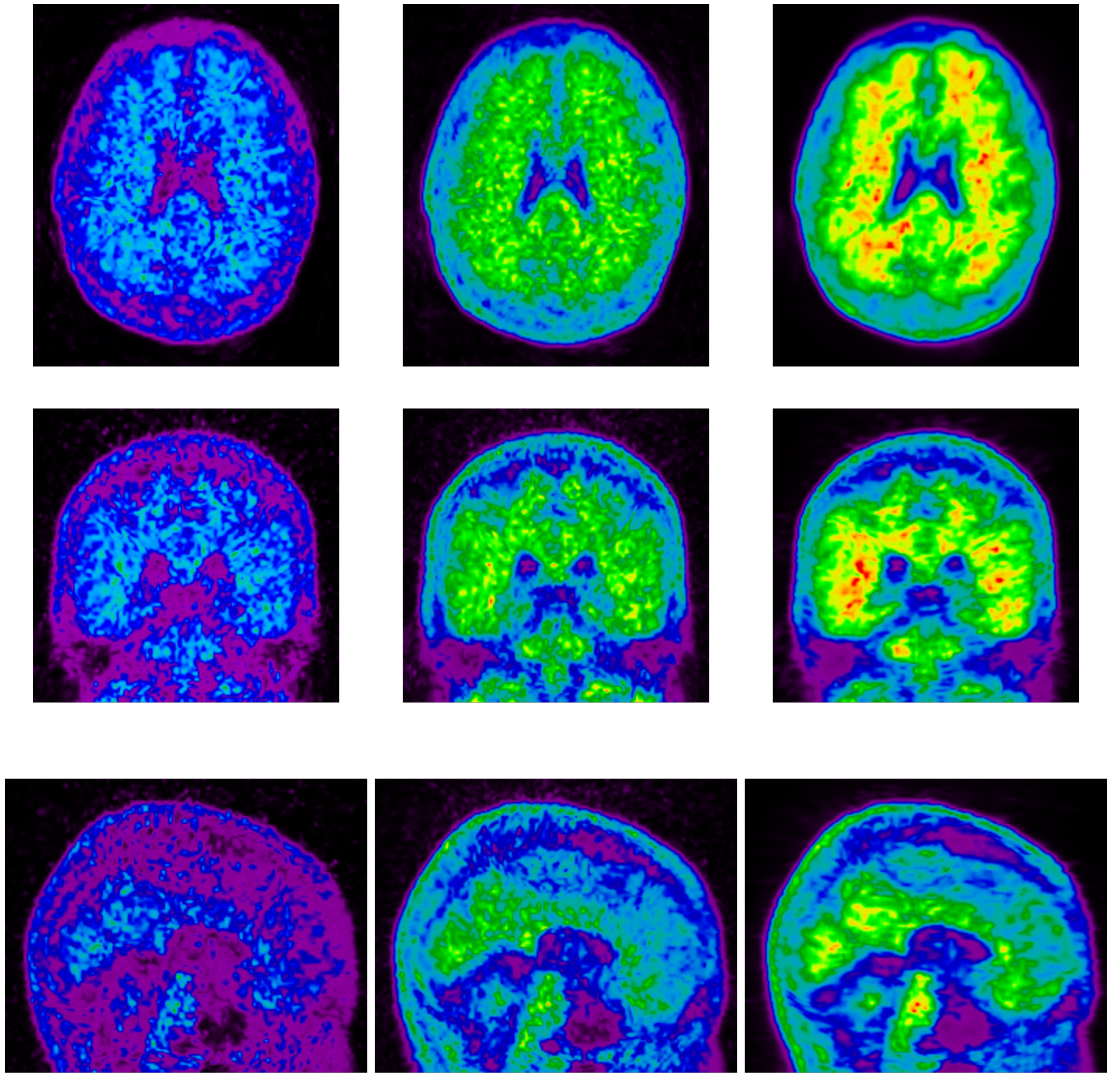
(h)

Figure 4.2: **(continued) The 16 validation studies used to visually inspect the performance of the U-Net.** Each figure shows the brain PET scan as obtained from (left column) a single-pass 5-minute PET scan, (middle column) a full 20-minute scan, and (right column) the neural network generated output from using the single-pass 5-minute scan. The resulting figures are also separated into (top) axial, (middle) coronal, and (bottom) sagittal slices.



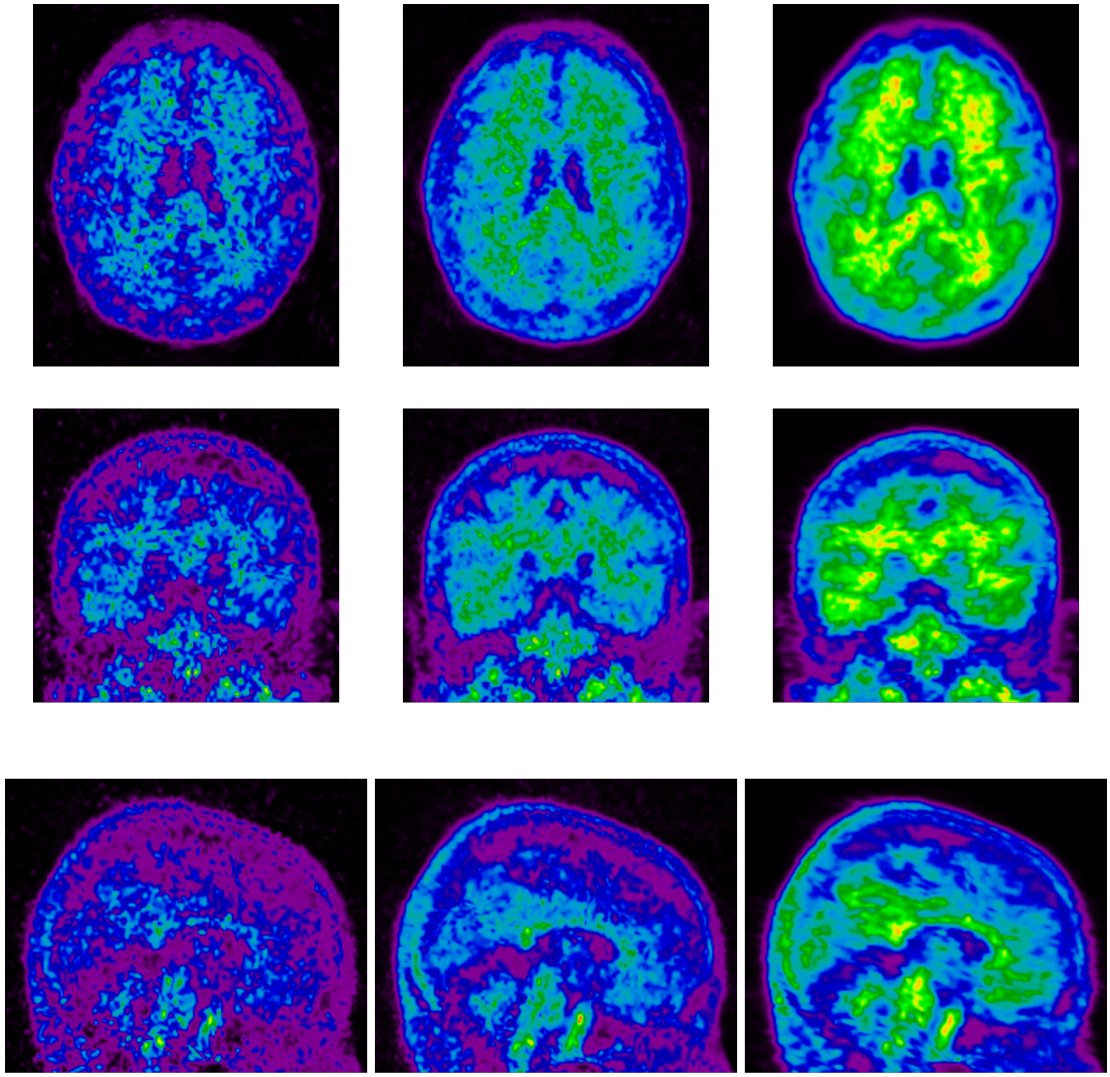
(i)

Figure 4.2: **(continued) The 16 validation studies used to visually inspect the performance of the U-Net.** Each figure shows the brain PET scan as obtained from (left column) a single-pass 5-minute PET scan, (middle column) a full 20-minute scan, and (right column) the neural network generated output from using the single-pass 5-minute scan. The resulting figures are also separated into (top) axial, (middle) coronal, and (bottom) sagittal slices.



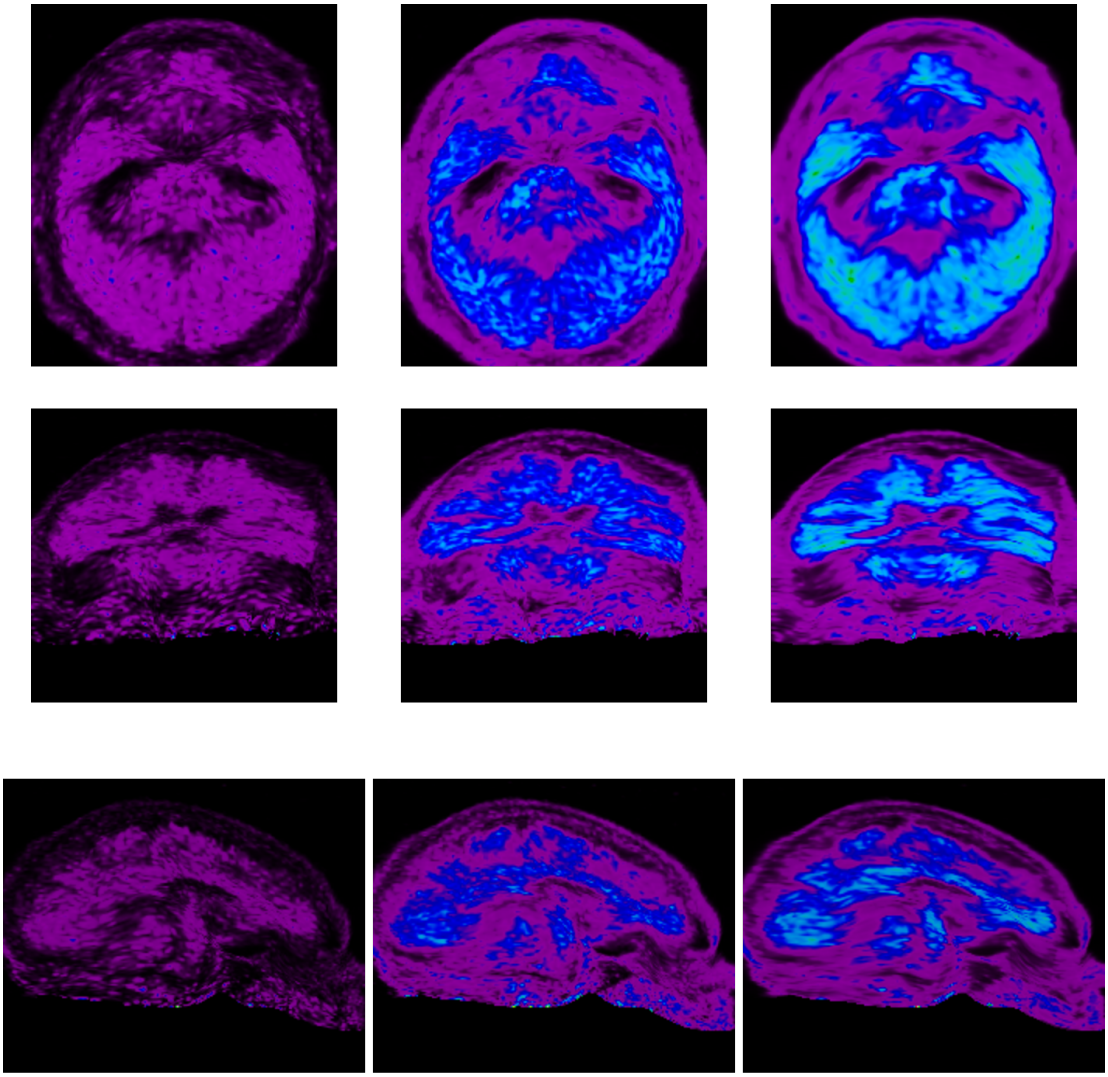
(j)

Figure 4.2: **(continued) The 16 validation studies used to visually inspect the performance of the U-Net.** Each figure shows the brain PET scan as obtained from (left column) a single-pass 5-minute PET scan, (middle column) a full 20-minute scan, and (right column) the neural network generated output from using the single-pass 5-minute scan. The resulting figures are also separated into (top) axial, (middle) coronal, and (bottom) sagittal slices.



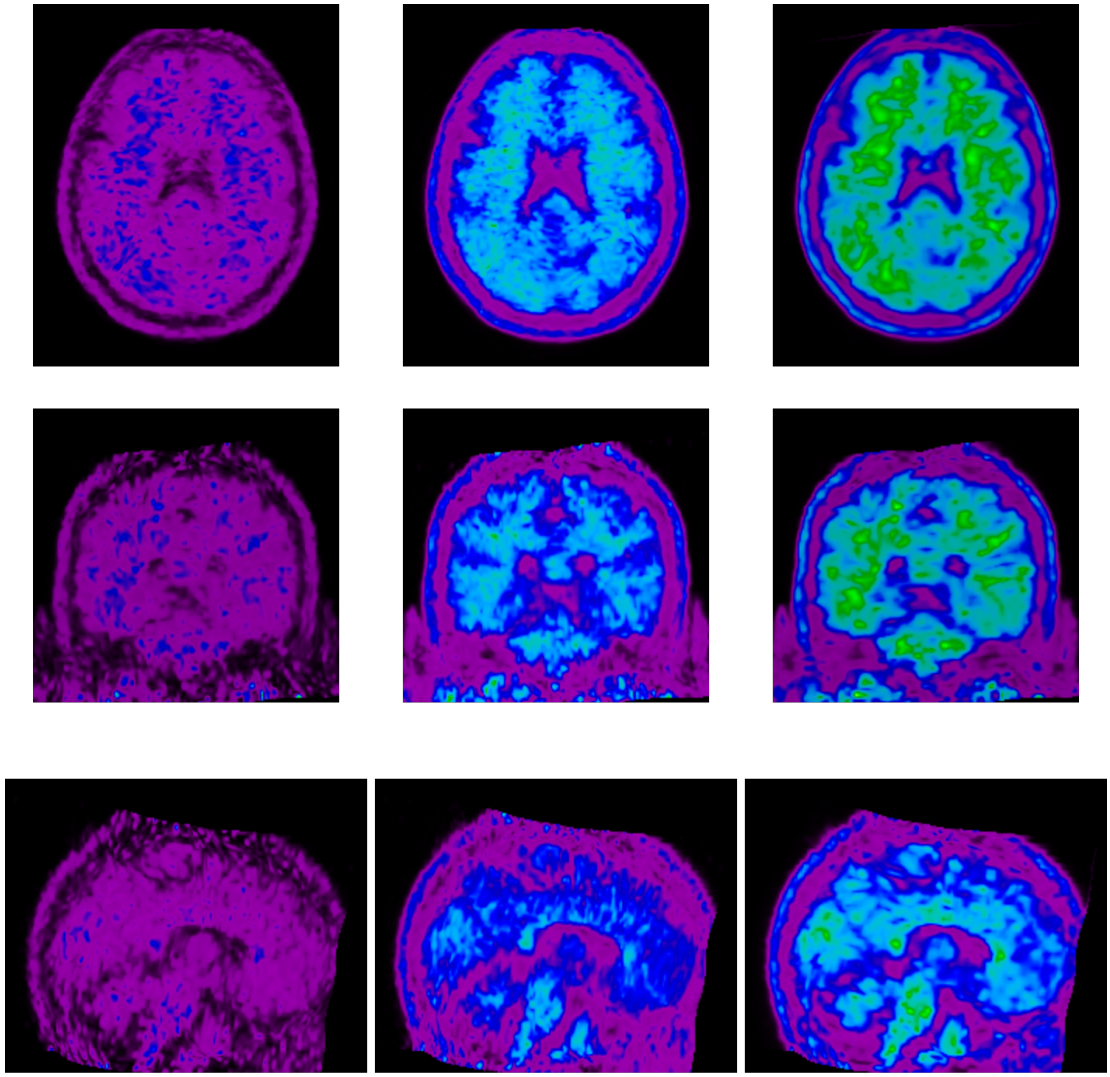
(k)

Figure 4.2: **(continued) The 16 validation studies used to visually inspect the performance of the U-Net.** Each figure shows the brain PET scan as obtained from (left column) a single-pass 5-minute PET scan, (middle column) a full 20-minute scan, and (right column) the neural network generated output from using the single-pass 5-minute scan. The resulting figures are also separated into (top) axial, (middle) coronal, and (bottom) sagittal slices.



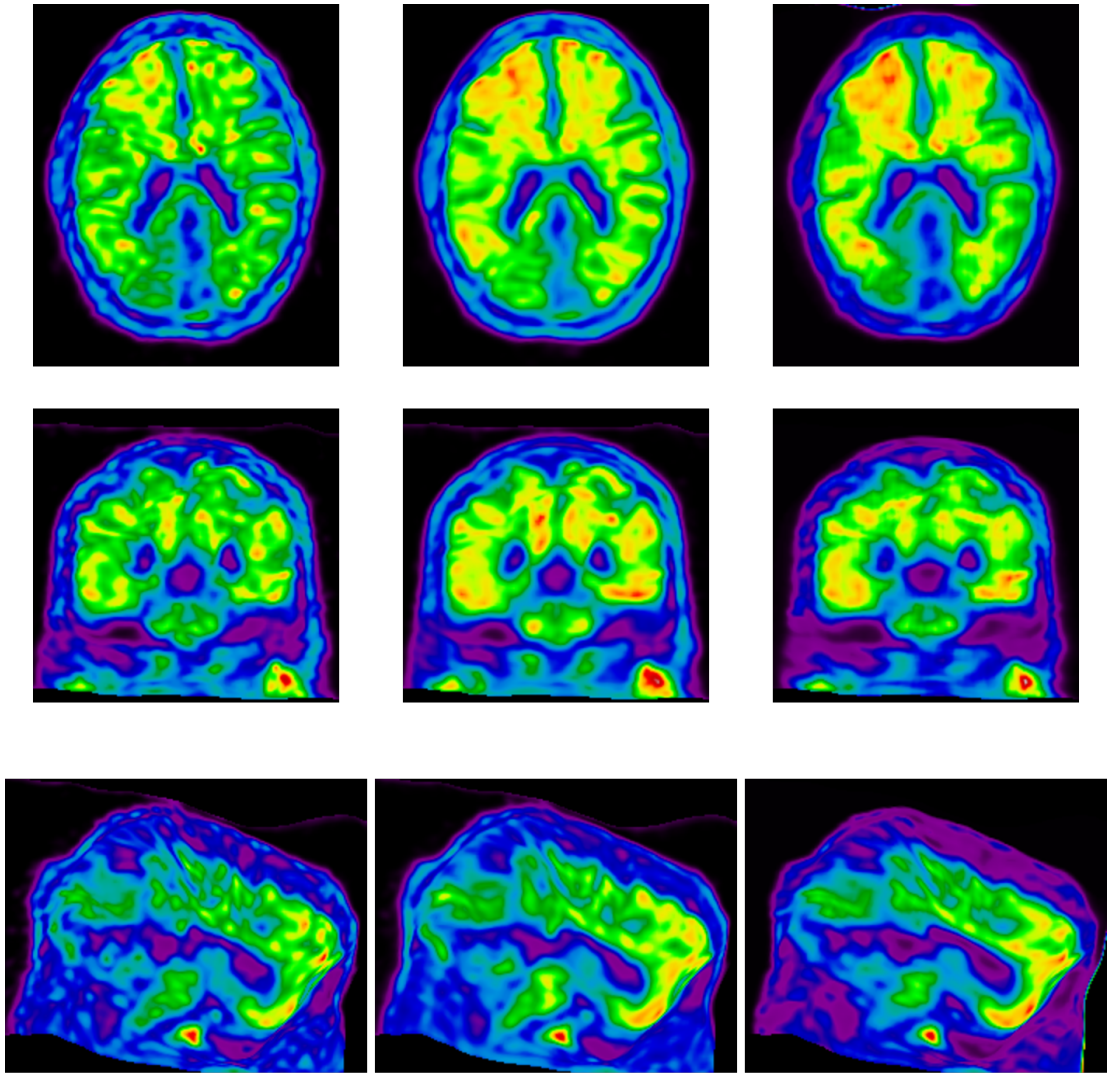
(l)

Figure 4.2: **(continued) The 16 validation studies used to visually inspect the performance of the U-Net.** Each figure shows the brain PET scan as obtained from (left column) a single-pass 5-minute PET scan, (middle column) a full 20-minute scan, and (right column) the neural network generated output from using the single-pass 5-minute scan. The resulting figures are also separated into (top) axial, (middle) coronal, and (bottom) sagittal slices.



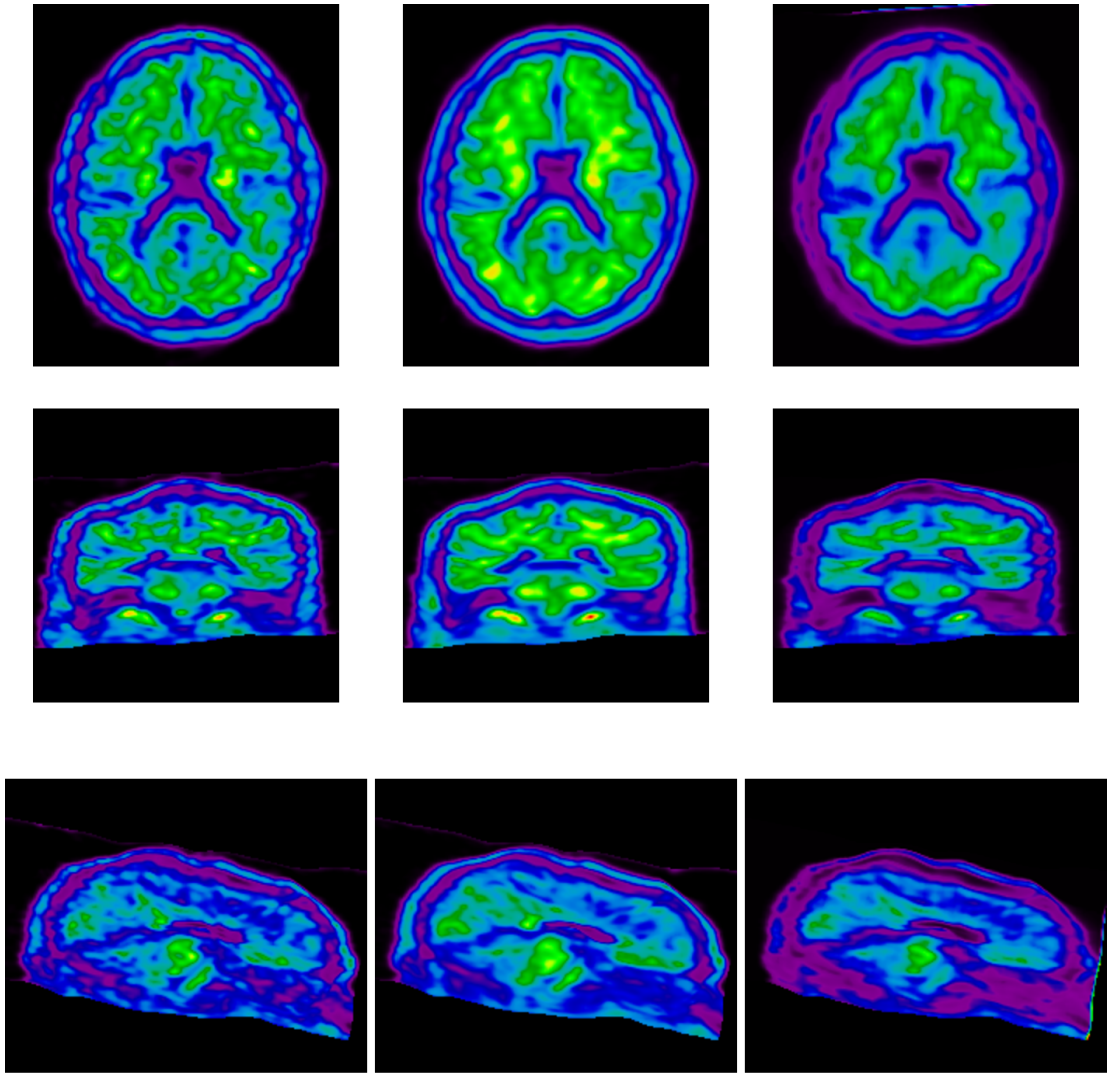
(m)

Figure 4.2: **(continued) The 16 validation studies used to visually inspect the performance of the U-Net.** Each figure shows the brain PET scan as obtained from (left column) a single-pass 5-minute PET scan, (middle column) a full 20-minute scan, and (right column) the neural network generated output from using the single-pass 5-minute scan. The resulting figures are also separated into (top) axial, (middle) coronal, and (bottom) sagittal slices.



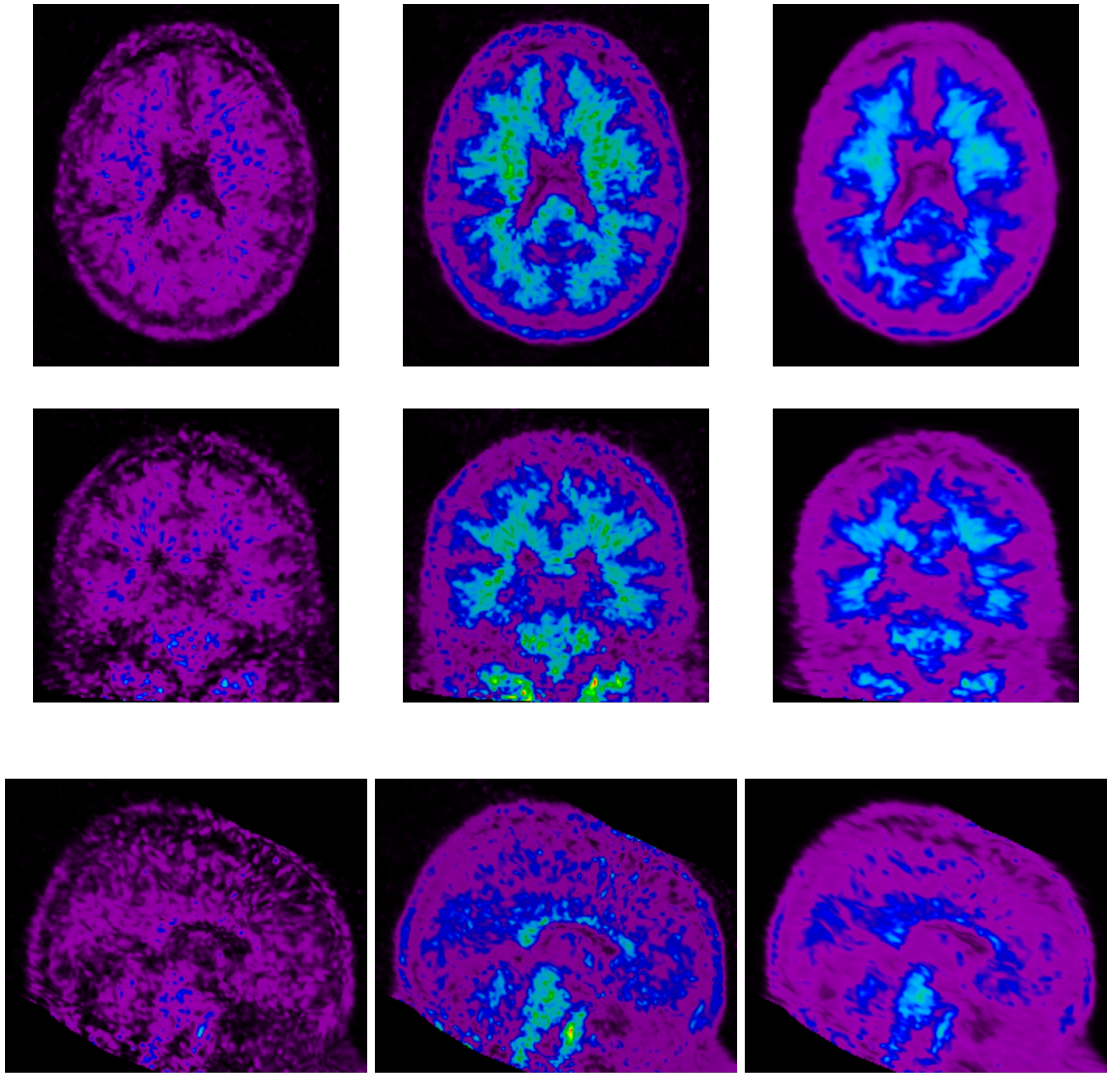
(n)

Figure 4.2: **(continued) The 16 validation studies used to visually inspect the performance of the U-Net.** Each figure shows the brain PET scan as obtained from (left column) a single-pass 5-minute PET scan, (middle column) a full 20-minute scan, and (right column) the neural network generated output from using the single-pass 5-minute scan. The resulting figures are also separated into (top) axial, (middle) coronal, and (bottom) sagittal slices.



(o)

Figure 4.2: **(continued) The 16 validation studies used to visually inspect the performance of the U-Net.** Each figure shows the brain PET scan as obtained from (left column) a single-pass 5-minute PET scan, (middle column) a full 20-minute scan, and (right column) the neural network generated output from using the single-pass 5-minute scan. The resulting figures are also separated into (top) axial, (middle) coronal, and (bottom) sagittal slices.



(p)

Figure 4.2: **(continued) The 16 validation studies used to visually inspect the performance of the U-Net.** Each figure shows the brain PET scan as obtained from (left column) a single-pass 5-minute PET scan, (middle column) a full 20-minute scan, and (right column) the neural network generated output from using the single-pass 5-minute scan. The resulting figures are also separated into (top) axial, (middle) coronal, and (bottom) sagittal slices.

	Left		Right	
	20-min.	**5-min.**	**20-min.**	**5-min.**
Angular Gyrus	1.28	1.20	1.27	1.27
Caudate	0.28	0.30	0.23	0.15
Anterior Cingulate Gyrus	0.94	0.92	0.93	0.92
Cuneus	1.02	0.98	1.28	1.29
Fusiform Gyrus	1.19	1.14	1.04	0.93
Gyrus Rectus	1.10	1.21	1.12	1.20
Hippocampus	0.91	1.00	0.95	0.96
Inferior Frontal Gyrus	1.01	1.01	0.90	0.99
Inferior Occipital Gyrus	1.45	1.12	1.28	0.98
Inferior Temporal Gyrus	1.33	1.30	1.15	1.18
Insular Cortex	0.98	0.94	0.98	1.05
Lateral Orbitofrontal Gyrus	0.86	0.81	0.78	0.89
Lingual Gyrus	1.23	1.06	1.17	0.95
Middle Frontal Gyrus	1.02	1.05	0.92	1.06

Table 4.1: **Example comparison of SUVRs for the 56 regions of interest in one patient, for full 20-minute scan and neural network output from a single 5-minute scan.** Qualitatively, the neural network output is comparable to that of the full scan in all ROIs.

	Left		Right	
	20-min.	**5-min.**	**20-min.**	**5-min.**
Middle Occipital Gyrus	1.34	1.14	1.38	1.28
Middle Orbitofrontal Gyrus	1.03	1.08	1.03	1.18
Middle Temporal Gyrus	1.39	1.32	1.24	1.36
Parahippocampal Gyrus	0.68	0.70	0.68	0.68
Postcentral Gyrus	1.07	1.10	0.85	0.96
Precentral Gyrus	1.14	1.23	0.83	0.98
Precuneus	1.06	0.97	0.86	0.86
Putamen	1.41	1.41	1.29	1.48
Superior Frontal Gyrus	0.92	1.00	0.87	0.87
Superior Occipital Gyrus	1.18	1.08	1.12	1.03
Superior Parietal Gyrus	1.11	1.10	0.89	0.94
Superior Temporal Gyrus	1.12	1.26	0.91	1.02
Supramarginal Gyrus	1.11	1.14	0.93	1.03
Posterior Cingulate	1.20	1.13	1.24	1.20

Table 4.1: **(continued) Example comparison of SUVRs for the 56 regions of interest in one patient, for full 20-minute scan and neural network output from a single 5-minute scan.** Qualitatively, the neural network output is comparable to that of the full scan in all ROIs.

Chapter 5

Discussion

This chapter details comments regarding the methodology and results of the presented work, considerations for potential future extensions to the work, as well as exploring and inviting discussion of the role of machine learning in nuclear medicine in general.

5.1 Evaluation of Results

Table 4.1 lists all the SUVR values from one validation study. These values represent the uptake value of florbetapir in each region relation to, in this case, cerebellar gray matter. Of special interest to nuclear medicine physicians are the cortical ROIs such as precuneus, anterior cingulate, and posterior cingulate [47]. The first consideration is then to verify that the SUVR values in the cortical regions are similar in value enough that clinical decisions are unaffected. The proposed threshold SUVR for classification of β-amyloid positivity using florbetapir is 1.10 [48]. In all of the aforementioned regions, the SUVR values from the neural network generated image does not deviate far enough away from the full 20-minute scan that it affects the positivity classification.

This SUVR threshold also helps indicate the similarity between the SUVR values for the full 20-minute scan and the neural network output from a single 5-minute scan. The nature of SUVR is that, in our case, the uptake values are normalized to that of cerebellar gray matter. Thus, the differences between the SUVR values as seen in the table for each ROI is negligible, provided they convey the same classification of β-amyloid positivity or negativity. In Table 4.1, the largest differences are observed in the inferior occipital gyrus and the left middle occipital gyrus. In all other regions, the SUVR difference is small and is not significant enough to change the classification. The large differences may arise from a poor registration through the pipeline, and not necessarily due to shortcomings of the neural network.

5.2 Neural Network Design

The U-Net structure was chosen for its specialized usage in image reconstruction tasks [49]. The input and output tensors of a U-Net have identical dimensions; and the structure appends the parameters learned in the encoding phase to the decoding phase to aid in reconstruction. At each layer of the encoding phase, the convolutional layers reduce the dimensions of the original image, while increasing the number of convolution filters. This is a direct result of the convolution procedure, where a small 2-dimensional window is passed over the input image and the overlapping values between the window and the image are linearly combined. The purpose of these filters, then, is to capture features within the image that would be computationally difficult to identify algorithmically; in the context of brain PET scans, these features may range from identifying the lobes to a full-on brain segmentation. The depth of the layers determines the complexity of the filters, at the expense of computational overhead. Notably, the features are learned in the internal layers of the network in the form of parameters that represent the weights of the linear combinations, and typically remain invisible. These features can then be appended

to the convolutional layers in the decoding phase, assisting in reconstruction of the original image.

In the U-Net, the dimensions are halved in both axes while the number of filters doubles. This inherently leads to a loss of information with each layer as the image progresses through the network. In the context of image reconstruction, this information loss ensures that the network is truly rebuilding the image from the learned features, rather than simply memorizing the relation between the input and output images. Because no data is effectively lost, each intermediate image representation becomes another expression of the original image, and it is difficult to say that the network trains to identify features within the image. For a U-Net, appending the parameters learned during the encoding phase to the decoding phase achieves little, as the problem decays from a reconstruction task to a re-representation one.

Compared to a GAN, a standalone U-Net has less computational cost due to not requiring a discriminator network to distinguish the generated images from the true images. Given the comparative qualitative and quantitative results, the reduction in computation cost allows for clinicians to develop and train the network locally, without having to rely on external services. This has implications in security and privacy; the usage of PET scans is inevitable in deep learning, and mass data transfers introduce a security vulnerability issue [50]. Even if the metadata for the PET images are anonymized, the nature of imaging allows a potentially personally identifiable data that cannot be obscured. A simple model could forgo any form of data transfer and omit this problem entirely. Furthermore, a simpler model has an additional advantage in that they are more robust and are less prone to mistakes due to inconsistencies in the images. Capturing finer features with a more complex model requires a larger training database, otherwise they may become sensitive to features not present during training.

5.3 Shortcomings of ML

Since information is destroyed and generated in a ML application, it can introduce anomalous artifacts to the outputs. This is especially important in a field such as tomography, where the coincidence counts solely determine the presence or absence of biological tissues and functions of interest.

5.3.1 Ethical Issues in Medical Imaging

The usage of artificial intelligence (and by extension, machine learning) in radiology has raised ethical concerns regarding the acquisition and usage of potentially confidential data [51, 52]. The nature of machine learning demands a very large dataset of pre-existing images, but behind each image is an individual with rights to privacy and consent to usage of their information. The storage, processing, and transfer of these images in pursuit of a machine learning application introduces issues that must be considered and addressed.

The widely used medical imaging format DICOM stores many identifying metadata header fields that must ideally be properly processed to remove or obscure any personally identifiable information [53]. This is further complicated by tomography machines adding proprietary metadata that may not conform to DICOM standards, making them difficult to locate and anonymize properly, leading to ill-processed images that can be maliciously used to personally identify the patient [54]. The approach to properly de-identifying radiology files is a complex topic in and of itself, necessitating discussion of matters such as how different associations would define personally identifiable data and which information can be deemed "required" and which can be obscured [55]. Following DICOM recommendations, there has been advances in development of anonymization tools, such as the Teaching File Transfer Tool (TFTT) and American College of Radiology's TRIAD software. Even if all such identifying information were to be removed from the DICOM metadata, the inherent

nature of imaging makes complete de-identification near impossible. Since a tomographic scan counts localized uptake in volumetric space and reconstructs a 3D image, it is possible to reconstruct the facial features with this information (Figure 5.1). To address this, algorithms such as skull stripping have been developed, to isolate the brain tissue from the relatively less important features such as skin, muscle, and eyes [56].

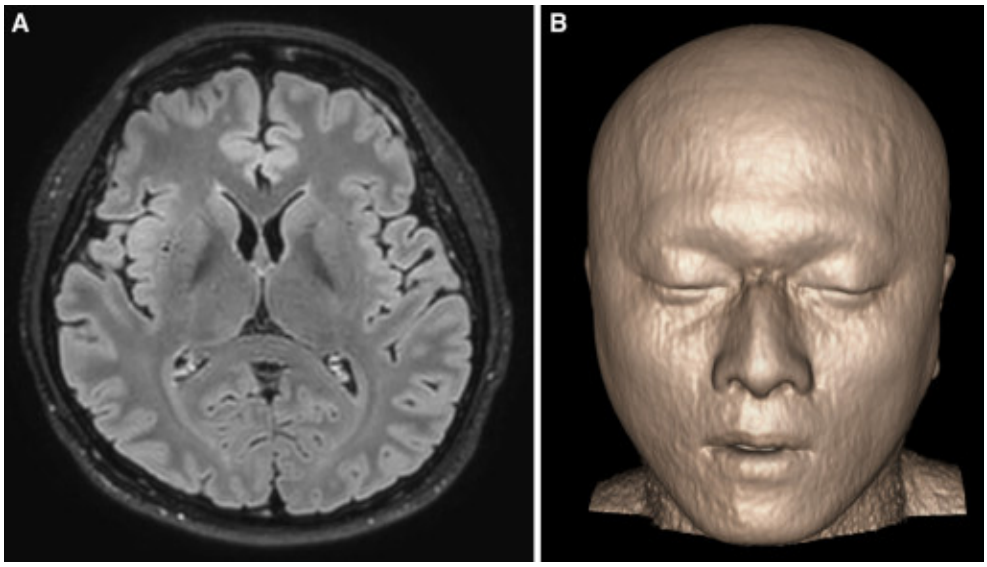


Figure 5.1: **The reconstruction of personally identifiable facial features from a tomographic scan.** (Left) An axial slice of a MRI scan. (Right) A 3-dimensional reconstruction based on the MRI slices. Adapted from [51].

5.4 Future Considerations

Relying solely on SUVR values for a quantitative evaluation is insufficient; obtaining such values relies on an automated registration pipeline, leading to inconsistencies between studies and even among regions of interest within the same study. An interesting extension to the evaluation would be, for each region of interest, to plot a graph of the SUVR value

for the full 20-minute scan against that of the neural network output and look for linear trends. This, however, would require a much larger dataset than the one used in this work. Furthermore, quantitative metrics such as SUVR can only tell part of the story. Quantifying the performance of the neural network is helpful as an objective evaluation, but ultimately an application of this type will be used in a clinical environment. The differences in metrics such as SUVR and mean squared error may not necessarily be as meaningful to nuclear medicine physicians as engineers. Thus, for a thorough analysis, a potential extension to this work is to incorporate clinical evaluations, consulting nuclear radiologists for their opinions on whether the neural network can generate clinically adequate images without affecting the decisions they would make. These would provide the insight and domain knowledge required for a full evaluation of the neural network.

Another consideration is the efficacy of the scan time reduction. In aiming for a 4-fold reduction in scan time, the input images may lose too much coincidence count information, leading to an image that is too noisy for a reconstruction task. There is merit in exploring various scan times (e.g. combining two 5-minute scans to simulate a 10-minute scan) and comparing neural networks trained at different scan times. Even a 2-fold reduction from a full 20-minute scan down to 10 minutes is a significant enough reduction in a clinical environment to warrant investigation [57].

References

[1] 2021 alzheimer's disease facts and figures. *Alzheimer's & Dementia*, 17(3):327–406, March 2021.

[2] Martina Zvěřová. Clinical aspects of alzheimer's disease. *Clinical Biochemistry*, 72:3–6, October 2019.

[3] Maurizio Conti and Lars Eriksson. Physics of pure and non-pure positron emitters for PET: a review and a discussion. *EJNMMI Physics*, 3(1):8, December 2016.

[4] K W D Ledingham, P McKenna, T McCanny, S Shimizu, J M Yang, L Robson, J Zweit, J M Gillies, J Bailey, G N Chimon, R J Clarke, D Neely, P A Norreys, J L Collier, R P Singhal, M S Wei, S P D Mangles, P Nilson, K Krushelnick, and M Zepf. High power laser production of short-lived isotopes for positron emission tomography. *Journal of Physics D: Applied Physics*, 37(16):2341–2345, August 2004.

[5] Henry Knipe and Mohamed Elbanan. Electron-positron annihilation. In *Radiopaedia.org*. Radiopaedia.org, January 2012.

[6] Amelie Heesch, Jochen Maurer, Elmar Stickeler, Mohsen Beheshti, Felix M. Mottaghy, and Agnieszka Morgenroth. Development of radiotracers for breast cancer—the tumor microenvironment as an emerging target. *Cells*, 9(10):2334, October 2020.

[7] N. A. Gomzina, D. A. Vasil'ev, and R. N. Krasikova. Use of 2-[18f]fluoroethyl bromide in synthesis of o-(2′-[18f]fluoroethyl)-l-tyrosine, a radiotracer for PET diagnostics of brain tumors. *Radiochemistry*, 49(3):299–304, June 2007.

[8] Paolo Palmisciano, Gina Watanabe, Andie Conching, Christian Ogasawara, Gianluca Ferini, Othman Bin-Alamer, Ali S. Haider, Maria Gabriella Sabini, Giacomo Cuttone, Sebastiano Cosentino, Massimo Ippolito, and Giuseppe E. Umana. The role of [68ga]ga-DOTA-SSTR PET radiotracers in brain tumors: A systematic review of the literature and ongoing clinical trials. *Cancers*, 14(12):2925, June 2022.

[9] Zixuan Zhao, Yanwen Hu, Jihui Li, Yeye Zhou, Bin Zhang, and Shengming Deng. Applications of PET in diagnosis and prognosis of leukemia. *Technology in Cancer Research & Treatment*, 19:153303382095699, January 2020.

[10] Heying Duan, Andrei Iagaru, and Carina Mari Aparici. Radiotheranostics - precision medicine in nuclear medicine and molecular imaging. *Nanotheranostics*, 6(1):103–117, 2022.

[11] Rudolf A. Werner, James T. Thackeray, Martin G. Pomper, Frank M. Bengel, Michael A. Gorin, Thorsten Derlin, and Steven P. Rowe. Recent updates on molecular imaging reporting and data systems (MI-RADS) for theranostic radiotracers—navigating pitfalls of SSTR- and PSMA-targeted PET/CT. *Journal of Clinical Medicine*, 8(7):1060, July 2019.

[12] Wilson Tang and Ayush Goel. SPECT vs PET, September 2014.

[13] Yen Tai and Paola Piccini. Applications of positron emission tomography (pet) in neurology. *Journal of neurology, neurosurgery, and psychiatry*, 75:669–76, 06 2004.

[14] Wikimedia Commons. 16-slice pet/ct, 2009.

[15] Timothy G. Turkington. Introduction to pet instrumentation. *Journal of Nuclear Medicine Technology*, 29(1):4–11, 2001.

[16] Arthur H. Compton. A quantum theory of the scattering of x-rays by light elements. *Phys. Rev.*, 21:483–502, May 1923.

[17] M. Bergström, L. Eriksson, C. Bohm, G. Blomqvist, and J. Litton. Correction for scattered radiation in a ring detector positron camera by integral transformation of the projections. *Journal of Computer Assisted Tomography*, 7(1):42–50, February 1983.

[18] R. A. J Nievelstein, H. M. E Quarles van Ufford, T. C Kwee, M. B Bierings, I Ludwig, F. J. A Beek, J. M. H de Klerk, W. P. Th. M Mali, P. W de Bruin, and J Geleijns. Radiation exposure and mortality risk from CT and PET imaging of patients with malignant lymphoma. *European radiology*, 22(9):1946–1954, 2012.

[19] Hyung Ju Lee, Hye Joo Son, Mijin Yun, Jung Won Moon, Yoo Na Kim, Ji Young Woo, and Suk Hyun Lee. Prone position [18F]FDG PET/CT to reduce respiratory motion artefacts in the evaluation of lung nodules. *European radiology*, 31(7):4606–4614, 2021.

[20] Benjamin P Fahimian, Yunzhe Zhao, Zhifeng Huang, Russell Fung, Yu Mao, Chun Zhu, Maryam Khatonabadi, John J DeMarco, Stanley J Osher, Michael F McNitt-Gray, and Jianwei Miao. Radiation dose reduction in medical x-ray CT via Fourier-based iterative reconstruction. *Medical physics (Lancaster)*, 40(3):031914–n/a, 2013.

[21] Kurt Schultz, Elizabeth George, Katherine M Mullen, Michael L Steigner, Dimitrios Mitsouras, Ericka M Bueno, Bohdan Pomahac, Frank J Rybicki, and Kanako K Kumamaru. Reduced Radiation Exposure for Face Transplant Surgical Planning Computed Tomography Angiography. *PloS one*, 8(4):e63079–e63079, 2013.

[22] Ju-Chieh (Kevin) Cheng, Julian Matthews, Vesna Sossi, Jose Anton-Rodriguez, André Salomon, and Ronald Boellaard. Incorporating HYPR de-noising within iterative PET reconstruction (HYPR-OSEM). *Physics in medicine & biology*, 62(16):6666–6687, 2017.

[23] Bradley T Christian, Nicholas T Vandehey, John M Floberg, and Charles A Mistretta. Dynamic PET denoising with HYPR processing. *Journal of Nuclear Medicine*, 51(7):1147–1154, 2010.

[24] A. L. Samuel. Some Studies in Machine Learning Using the Game of Checkers. *IBM Journal of Research and Development*, 3(3):210–229, July 1959.

[25] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton. Deep learning. *Nature*, 521(7553):436–444, May 2015.

[26] Li Deng. The MNIST database of handwritten digit images for machine learning research [best of the web]. *IEEE Signal Processing Magazine*, 29(6):141–142, November 2012.

[27] Shaun Fernandes, Dhruv Duseja, and Raja Muthalagu. Application of image processing techniques for autonomous cars. *Proceedings of Engineering and Technology Innovation*, December 2020.

[28] Isabel Vidaurre-Gallart, Isabel Fernaud-Espinosa, Nicusor Cosmin-Toader, Lidia Talavera-Martínez, Miguel Martin-Abadal, Ruth Benavides-Piccione, Yolanda Gonzalez-Cid, Luis Pastor, Javier DeFelipe, and Marcos García-Lorenzo. A deep learning-based workflow for dendritic spine segmentation. *Frontiers in Neuroanatomy*, 16, March 2022.

[29] Seth Winfree. User-accessible machine learning approaches for cell segmentation and analysis in tissue. *Frontiers in Physiology*, 13, March 2022.

[30] Kui Zhao, Long Zhou, Size Gao, Xiaozhuang Wang, Yaofa Wang, Xin Zhao, Huatao Wang, Kanfeng Liu, Yunqi Zhu, and Hongwei Ye. Study of low-dose PET image recovery using supervised learning with CycleGAN. *PloS one*, 15(9):e0238455–e0238455, 2020.

[31] Jelmer M Wolterink, Tim Leiner, Max A Viergever, and Ivana Isgum. Generative Adversarial Networks for Noise Reduction in Low-Dose CT. *IEEE transactions on medical imaging*, 36(12):2536–2545, 2017.

[32] Kevin A. Matthews, Wei Xu, Anne H. Gaglioti, James B. Holt, Janet B. Croft, Dominic Mack, and Lisa C. McGuire. Racial and ethnic estimates of alzheimer's disease and related dementias in the united states (2015–2060) in adults aged $\geq$ 65 years. *Alzheimer's & Dementia*, 15(1):17–24, 2019.

[33] Alzheimer's Association. 2019 alzheimer's disease facts and figures. *Alzheimer's & Dementia*, 15(3):321–387, 2019.

[34] Michael A. DeTure and Dennis W. Dickson. The neuropathological diagnosis of alzheimer's disease. *Molecular Neurodegeneration*, 14(1), August 2019.

[35] Antoine Leuzy, Konstantinos Chiotis, Laetitia Lemoine, Per-Göran Gillberg, Ove Almkvist, Elena Rodriguez-Vieitez, and Agneta Nordberg. Tau PET imaging in neurodegenerative tauopathies—still a challenge. *Molecular Psychiatry*, 24(8):1112–1134, August 2019.

[36] Renaud La Joie, Adrienne V. Visani, Suzanne L. Baker, Jesse A. Brown, Viktoriya Bourakova, Jungho Cha, Kiran Chaudhary, Lauren Edwards, Leonardo Iaccarino, Mustafa Janabi, Orit H. Lesman-Segev, Zachary A. Miller, David C. Perry, James P. O'Neil, Julie Pham, Julio C. Rojas, Howard J. Rosen, William W. Seeley, Richard M.

Tsai, Bruce L. Miller, William J. Jagust, and Gil D. Rabinovici. Prospective longitudinal atrophy in Alzheimer's disease correlates with the intensity and topography of baseline tau-PET. *Science Translational Medicine*, 12(524):eaau5732, January 2020.

[37] Silvia Morbelli and Matteo Bauckneht. Amyloid PET imaging: Standardization and integration with other alzheimer's disease biomarkers. In *Biomarkers for Alzheimer's Disease Drug Development*, pages 203–212. Springer New York, 2018.

[38] Peter Mildenberger, Marco Eichelberg, and Eric Martin. Introduction to the DICOM standard. *European Radiology*, 12(4):920–927, September 2001.

[39] Mark O. Gueld, Michael Kohnen, Daniel Keysers, Henning Schubert, Berthold B. Wein, Joerg Bredno, and Thomas M. Lehmann. Quality of DICOM header information for image categorization. pages 280–287, San Diego, CA, May 2002.

[40] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-Net: Convolutional Networks for Biomedical Image Segmentation. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, volume 9351 of *Lecture Notes in Computer Science*, pages 234–241. Springer International Publishing, Cham, 2015.

[41] Diederik P Kingma and Jimmy Ba. Adam: A Method for Stochastic Optimization. 2014.

[42] Zhou Wang, A.C Bovik, H.R Sheikh, and E.P Simoncelli. Image quality assessment: from error visibility to structural similarity. *IEEE transactions on image processing*, 13(4):600–612, 2004.

[43] Robert D. Vincent, Peter Neelin, Najmeh Khalili-Mahani, Andrew L. Janke, Vladimir S. Fonov, Steven M. Robbins, Leila Baghdadi, Jason Lerch, John G. Sled, Reza Adalat, David MacDonald, Alex P. Zijdenbos, D. Louis Collins, and Alan C.

Evans. MINC 2.0: A flexible format for multi-modal images. *Frontiers in Neuroinformatics*, 10, August 2016.

[44] G. Lucignani, G. Paganelli, and E. Bombardieri. The use of standardized uptake values for assessing FDG uptake with PET in oncology: a clinical perspective. *Nuclear Medicine Communications*, 25(7):651–656, July 2004.

[45] Yanxiao Li, Yee Ling Ng, Manish D. Paranjpe, Qi Ge, Fengyun Gu, Panlong Li, Shaozhen Yan, Jie Lu, Xiuying Wang, Yun Zhou, and for the Alzheimer's Disease Neuroimaging Initiative. Tracer-specific reference tissues selection improves detection of 18f-fdg, 18f-florbetapir, and 18f-flortaucipir pet suvr changes in alzheimer's disease. *Human Brain Mapping*, 43(7):2121–2133, 2022.

[46] John Seibyl, Gary Wisniewski, Denise Ferraiolo, Kenneth Marek, and George Zubal. P3–114: Quantitative evaluation of anthropomorphic brain phantoms for standardization of amyloid PET SUVrs in alzheimer's multicenter therapeutic trials. *Alzheimer's & Dementia*, 9(4S_Part_15), July 2013.

[47] V. Camus, P. Payoux, L. Barré, B. Desgranges, T. Voisin, C. Tauber, R. La Joie, M. Tafani, C. Hommet, G. Chételat, K. Mondon, V. de La Sayette, J. P. Cottier, E. Beaufils, M. J. Ribeiro, V. Gissot, E. Vierron, J. Vercouillie, B. Vellas, F. Eustache, and D. Guilloteau. Using PET with 18f-AV-45 (florbetapir) to quantify brain amyloid load in a clinical environment. *European Journal of Nuclear Medicine and Molecular Imaging*, 39(4):621–631, January 2012.

[48] Abhinay D. Joshi, Michael J. Pontecorvo, Chrisopher M. Clark, Alan P. Carpenter, Danna L. Jennings, Carl H. Sadowsky, Lee P. Adler, Karel D. Kovnat, John P. Seibyl, Anupa Arora, Krishnendu Saha, Jason D. Burns, Mark J. Lowrey, Mark A. Mintun, and Daniel M. Skovronsky and. Performance characteristics of amyloid PET with

florbetapir f 18 in patients with alzheimer's disease and cognitively normal subjects. *Journal of Nuclear Medicine*, 53(3):378–384, February 2012.

[49] Evan Shelhamer, Jonathan Long, and Trevor Darrell. Fully Convolutional Networks for Semantic Segmentation. *IEEE transactions on pattern analysis and machine intelligence*, 39(4):640–651, 2017.

[50] T.Y Alvin Liu and Neil M Bressler. Controversies in artificial intelligence. *Current opinion in ophthalmology*, 31(5):324–328, 2020.

[51] and Jacob L. Jaremko, Marleine Azar, Rebecca Bromwich, Andrea Lum, Li Hsia Alicia Cheong, Martin Gibert, François Laviolette, Bruce Gray, Caroline Reinhold, Mark Cicero, Jaron Chong, James Shaw, Frank J. Rybicki, Casey Hurrell, Emil Lee, and An Tang. Canadian association of radiologists white paper on ethical and legal issues related to artificial intelligence in radiology. *Canadian Association of Radiologists Journal*, 70(2):107–118, May 2019.

[52] Marc Kohli and Raym Geis. Ethics, artificial intelligence, and radiology. *Journal of the American College of Radiology*, 15(9):1317–1319, September 2018.

[53] Rita Noumeir, Alain Lemay, and Jean-Marc Lina. Pseudonymization of radiology data for research purposes. *Journal of Digital Imaging*, 20(3):284–295, December 2006.

[54] Stephen M. Moore, David R. Maffitt, Kirk E. Smith, Justin S. Kirby, Kenneth W. Clark, John B. Freymann, Bruce A. Vendt, Lawrence R. Tarbox, and Fred W. Prior. De-identification of medical images with retention of scientific research value. *RadioGraphics*, 35(3):727–735, May 2015.

[55] Burbridge Brent. Dicom image anonymization and transfer to create a diagnostic radiology teaching file. *International Journal of Radiology and Imaging Techniques*, 6(2), September 2020.

[56] P. Kalavathi and V. B. Surya Prasath. Methods on skull stripping of MRI head scan images—a review. *Journal of Digital Imaging*, 29(3):365–379, December 2015.

[57] Zhao Peng, Ming Ni, Hongming Shan, Yu Lu, Yongzhe Li, Yifan Zhang, Xi Pei, Zhi Chen, Qiang Xie, Shicun Wang, and X. George Xu. Feasibility evaluation of PET scan-time reduction for diagnosing amyloid-β levels in alzheimer's disease patients using a deep-learning-based denoising algorithm. *Computers in Biology and Medicine*, 138:104919, November 2021.

[58] Charles E. Kahn, John A. Carrino, Michael J. Flynn, Donald J. Peck, and Steven C. Horii. DICOM and Radiology: Past, Present, and Future. *Journal of the American College of Radiology*, 4(9):652–657, 2007.

[59] Charles E. Metz, Robert F. Wagner, Kunio Doi, David G. Brown, Robert M. Nishikawa, and Kyle J. Myers. Toward consensus on quantitative assessment of medical imaging systems. *Medical Physics*, 22(7):1057–1061, July 1995.

[60] W. Nicholson Price and I. Glenn Cohen. Privacy in the age of medical big data. *Nature Medicine*, 25(1):37–43, January 2019.

Glossary

coincidence event An event detected by a PET scanner where two photons are detected simultaneously along a line of response 7

DICOM Digital Imaging and Communications in Medicine; a medical imaging file format standard 19

isotope An atom of an element with a different number of neutrons (and thus mass number) than the standard in the periodic table. 4

ML Machine learning 12

PET Positron emission tomography 3, 7

radiotracer An organic compound tagged with a radioactive isotope, which undergoes radioactive decay to emit energy that can be detected 3, 6