MOTION CORRECTION STRATEGIES IN PET/MRI SCANNERS AND DEVELOPMENT OF A DEEP LEARNING BASED COMPUTER AIDED DETECTION SYSTEM

A THESIS SUBMITTED TO THE GRADUATE SCHOOL OF APPLIED SCIENCES OF NEAR EAST UNIVERSITY

By
ALİ IŞIN

In Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy in Biomedical Engineering

NICOSIA, 2018

Ali IŞIN: MOTION CORRECTION STRATEGIES IN PET/MRI SCANNERS AND DEVELOPMENT OF A DEEP LEARNING BASED COMPUTER AIDED DETECTION SYSTEM

Approval of Director of Graduate School of Applied Sciences

Prof. Dr. Nadire CAVUŞ

We certify this thesis is satisfactory for the award of the degree of Doctor of Philosophy in Biomedical Engineering

Examining Committee in Charge:

Prof.Dr. Bülent Bilgehan	Committee Chairman, Department of Electrical-Electronic Engineering, NEU
Assist.Prof.Dr. Dilber U. Özşahin	Supervisor, Department of Biomedical Engineering, NEU
Assoc.Prof.Dr. Kamil Dimililer	Department of Automotive Engineering, NEU
Prof.Dr. Ayşe Günay Kibarer	Department of Chemistry, Hacettepe University
Prof.Dr. Barlas Naim Aytaçoğlu	Faculty of Medicine, Girne American University

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Last name:

Signature:

Date:

To my son, Kemal Işın…

ACKNOWLEDGEMENTS

First and foremost I would like to thank my supervisor Asst. Prof. Dr. Dilber Uzun Özşahin who has shown plenty of encouragement, patience, and support as she guided me through this thesis process. I am also thankful for the contributions and comments to the teaching staff of the Department of Biomedical Engineering and also to my Family.

ABSTRACT

Diagnosis of the lung cancer, the most fatal cancer type, involves screening the patients initially by Computed Tomography (CT) for the presence of lung lesions, which can be malign or benign. However diagnosing malignancy from just CT images is not an easy task. In this regard functional imaging provided by the Positron Emission Tomography (PET) is an invaluable solution which enables non-invasive lung cancer diagnosis.

Researchers frequently develop and test proposed improvements for the PET using robust simulation environments like the GATE. Since PET scanner requires several minutes to complete the scan of a patient, natural respiratory motion of the patient is unavoidable during the lung cancer imaging. This adversely affects the overall image quality, thus motivating researchers to establish motion correction techniques for increasing the quality of the images. As the first aim of this thesis, several different motion correction techniques (based on image reconstruction) are developed and tested using a simulated torso phantom (with lung lesions) in GATE simulation environment. Obtained results clearly demonstrate the quality improvements that the correction of the respiratory motion related artifacts provide.

Additionally, radiologists need to go over large numbers of image slices manually in order to detect and diagnose lung lesions. This process is very time consuming and its performance is very dependent on the performing radiologist. Thus assisting the radiologists by developing an automated computer aided detection (CAD) system is an interesting research goal. In this regard, as the second goal of this thesis a pre-trained AlexNet (deep learning) framework is transferred to develop and implement a robust CAD system for the classification of lung images depending on whether they bear a lesion or not. High performances of 98.72% sensitivity, 98.35% specificity and 98.48% accuracy are reported as a result.

Keywords: Lung Cancer; PET; Respiratory Motion Correction; CAD; Deep Learning

ÖZET

En ölümcül kanser tipi olan akciğer kanserinin teşhisi, ilk olarak bilgisayarlı tomografi (BT) ile malign veya benign olabilecek akciğer lezyonlarının varlığını taramayı içermektedir. Bununla birlikte, sadece BT görüntülerinden maligniteyi teşhis etmek kolay bir iş değildir. Bu bağlamda, Pozitron Emisyon Tomografisi (PET) tarafından sağlanan fonksiyonel görüntüleme, invaziv olmayan akciğer kanseri teşhisini mümkün kılan çok değerli bir çözümdür.

Araştırmacılar, GATE gibi güçlü simülasyon ortamlarını kullanarak PET için önerilen iyileştirmeleri sıklıkla geliştirip test etmektedirler. PET tarayıcısının bir hastanın taramasını tamamlaması için birkaç dakika gerektiğinden, akciğer kanseri görüntülemesi sırasında hastanın doğal solunum hareketi kaçınılmazdır. Bu durum, genel görüntü kalitesini olumsuz etkileyerek, görüntü kalitesini iyileştirmek için araştırmacıları hareket düzeltme yöntemleri geliştirmeye motive etmektedir. Bu tezin ilk amacı olarak, GATE simülasyon ortamında simüle edilmiş bir gövde fantomu (akciğer lezyonları eklenerek) kullanılarak çeşitli farklı hareket düzeltme teknikleri (görüntü rekonstrüksiyonu tabanlı) geliştirilmiş ve test edilmiştir. Elde edilen sonuçlar, solunum hareketlerine bağlı artifaktların düzeltilmesinin sağladığı kalite iyileştirmelerini açıkça göstermektedir.

Ayrıca, radyologlar akciğer lezyonlarını saptamak ve teşhis etmek için çok sayıda görüntü dilimini elden taramaları gerekmektedir. Bu süreç çok zaman alıcı olup performansı gerçekleştiren radyoloğa bağlıdır. Böylece otomatik bir bilgisayar destekli algılama (CAD) sistemi geliştirerek radyologlara yardımcı olmak ilginç bir araştırma hedefidir. Bu bağlamda, bu tezin ikinci amacı olarak, akciğer görüntülerinin bir lezyon barındırıp barındırmadığı yönünde sınıflandırma yapmak üzere bir CAD sistemi geliştirmek ve uygulamak amacı ile önceden eğitilmiş bir AlexNet (derin öğrenme) çerçevesi mevcut işe aktarılmıştır. Sonuç olarak, % 98,72 duyarlılık, % 98.35 özgüllük ve % 98,48 hassasiyetle yüksek performanslar rapor edilmektedir.

Anahtar Kelimeler: Akciğer Kanseri; PET; Solunum Hareket Düzeltme; CAD; Derin Öğrenme

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

3DRT: Three Dimensional Conformal Radiotherapy

ACF: Attenuation Correction Factor

ANN: Artificial Neural Networks

AP: Anteroposterior

BGO: Bismuth Germanate

CAD: Computer Aided Detection

CC: Connected Components

CNN: Convolutional Neural Networks

CRF: Conditional Random Fields

CT: Computed Tomography

DOI: Depth of Interaction

FBP: Filtered Back Projection

FDG: Fluorodeoxyglucose

FWHM: Full Width Half Maximum

GPU: Graphical Processing Unit

GSO: Gadolinium Oxyorthosilicate

HU: Hounsfield Units

IGRT: Image Guided Radiation Therapy

IMRT: Intensity Modulated Radiation Therapy

kNN: k-Nearest Neighbor Classifier

LINAC: Linear Accelerator

LOR: Line of Response

LSO: Cerium-doped Lutetium Oxyorthosilicate

MLEM: Maximum Likelihood Expectation Maximization

MRI: Magnetic Resonance Imaging

NaI(Tl): Thallium-doped Sodium Iodide

NCAT: Non-uniform Cardiac and Torso Phantom

NECR: Noise-equivalent Count Rate

NMR: Nuclear Magnetic Resonance

NSCLC: Non-Small Cell Lung Cancer

OSEM: Ordered Subsets Expectation Maximization

PA: Posteroanterior

PCA: Principal Component Analysis

PD: Proton Density

PET: Positron Emission Tomography

PET/CT: Hybrid Positron Emission Tomography-Computed Tomography

PET/MRI: Hybrid Positron Emission Tomography-Magnetic Resonance Imaging

PLD: Public Lung Database

PVE: Partial Volume Effect

RF: Radio Frequency

RFs: Random Forests

SCLC: Small Cell Lung Cancer

SIFT: Scale Invariant Feature Transform

SiPM: Silicon Photomultiplier

SNR: Signal to Noise Ratio

SOM: Self Organizing Maps

SPECT: Single Photon Emission Tomography

SUV: Standardized Uptake Value

SVM: Support Vector Machine

T: Tesla

TE: Echo Time

TOF: Time of Flight

TR: Repetition Time

WHO: World Health Organization

XCAT: Extended Cardiac and Torso Phantom

CHAPTER 1
INTRODUCTION

Positron Emission Tomography (PET) is an established imaging technique in medicine for obtaining functional images of the human body. Due to its property to obtain functional images of the metabolic activity, PET is widely used in oncological diagnostic imaging for the detection of various cancerous tissues including detection and monitoring of tumors located in the torso region. Lung tumors are the most fatal tumor types that can be widely encountered in this region. Apart from oncological imaging, neurological imaging to diagnose disorders, cardiologic imaging to diagnose diseases and its integration as an imaging tool in radiation therapy planning can be considered as the other main uses of the PET device (Chen, 2013; Lin and Alavi, 2009; Ford et al., 2009).

PET imaging concept is based on injecting radioactive tracers to the patient. These radio-tracers release positrons which annihilate with the electrons of the tissues to generate two back to back gamma photons which can be detected by the specialized detectors of the PET device to generate diagnostic images using advanced image reconstruction algorithms. Since radiolabeled tracer molecules are coupled with molecules, like sugars, that can easily accumulate in metabolically active tissues, the detection of those released photons is used to estimate the metabolic activity of the tissues, providing a competent imaging tool for functional body imaging (Nehmeh et al., 2002).

Most commonly used radiotracer in oncological PET imaging is ^{18}F-fluorodeoxyglucose (^{18}F-FDG). ^{18}F-FDG is transported into the cells of the patient by the glucose metabolism thus enabling it to be used in imaging of the glucose metabolic activity of the patient in PET imaging. Since cells in cancer tissues divide and grow far more rapidly than normal body cells, they require higher glucose metabolic activity to provide the necessary energy in the process. Thus PET images using ^{18}F-FDG radiotracers provide to be a powerful tool in imaging and locating cancer cells. Apart from ^{18}F-FDG, some other used PET radiotracers and their corresponding involved biological processes can be seen in Table 1.1.

Table 1.1: PET radiotracers used in medicine and their corresponding involved biological processes

Name of the tracer	Involved biological process
^{18}F-FDG	Glucose metabolism
^{18}FMISO	Hypoxia
^{11}C-methionine	Cellular amino acid uptake
$H_2{}^{15}O$	Blood flow
^{18}F-dopa	Dopamine storage

Due to its design and concept, the powerful functional imaging capability of the PET device comes along with its poor spatial resolution and structural imaging capability with respect to more conventional medical imaging techniques such as Computed Tomography (CT) and Magnetic Resonance Imaging (MRI) devices. While these conventional imaging modalities lack PET`s functional metabolic imaging properties, they can provide detailed high resolution images of the anatomical structures within the patient's body. Using these highly detailed structural images to precisely locate the anatomical positions of the high metabolic activities, such as tumors, was the main driving force behind the recent development of the hybrid functional-structural imaging modalities.

PET/CT and PET/MRI (as seen in Figure 1.1) are two of these hybrid devices that are used in tumor imaging. PET/CT uses the high quality, high resolution structural hard tissue (like bones and ligaments) imaging capabilities of CT device to precisely localize high metabolic activities. In contrast, PET/MRI uses the high quality, high resolution structural soft tissue (like muscles, brain tissue, lung tissue etc.) imaging capabilities of the MRI device. Due these powerful hybrid imaging capabilities PET/CT and PET/MRI devices currently can be considered as the state of the art for detection, localization and diagnosis of the cancerous tissues in oncological imaging. Instead of using PET, CT or MRI alone, using these hybrid devices for cancer diagnosis provides far more better diagnostic accuracy (Antoch et al., 2003). Main drawbacks of these devices are their very high cost and very rare availability when compared to other medical imaging modalities. Rarity of these devices (only 1 PET/CT device was available in North Cyprus as the time of writing,

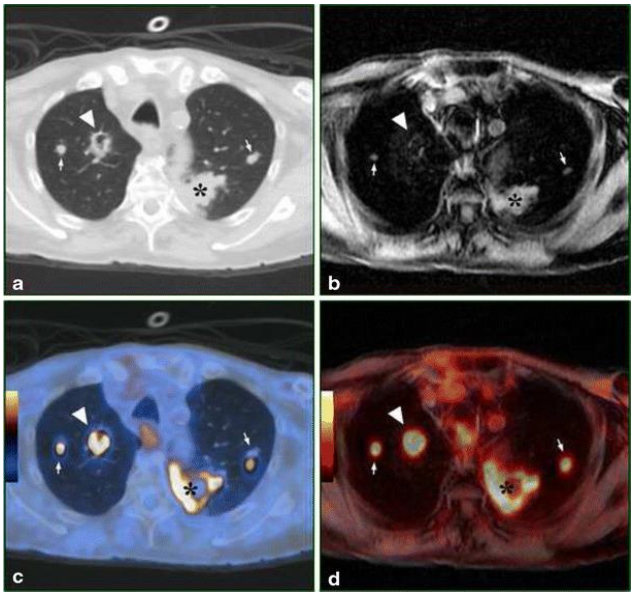


Figure 1.1: a) CT scan, b) MRI scan of lung tumors marked by arrows. Bottom row shows the corresponding c) PET/CT and d) PET/MRI scans (Appenzeller et al, 2013)

while there were none PET/MRI devices) creates the need for the development and use of powerful simulators, for the scientists to work on and develop new methods to improve the imaging modality.

1.1 Lung Cancer and PET

Lung cancer can be considered as the deadliest cancer type in the world. World Health Organization (WHO) reported 1.690.000 deaths from lung cancer in year 2015. This number is believed to increase to around 2.280.000 deaths by 2030 (http://www.who.int/healthinfo/global_burden_disease/projections2002/en/ Retrieved 10 February, 2017). In USA smoking tobacco products are the main reason behind lung cancer with 90 percent of whole cases (Alberg et al., 2007). Exposure to polluted air and genetics can be considered as other factors. Early diagnostics play a very important role in the survival rates of the lung cancer patients, so precise detection of the lung tumors by using state of the art PET/CT or PET/MRI devices that implement the best quality, best resolution imaging is very crucial.

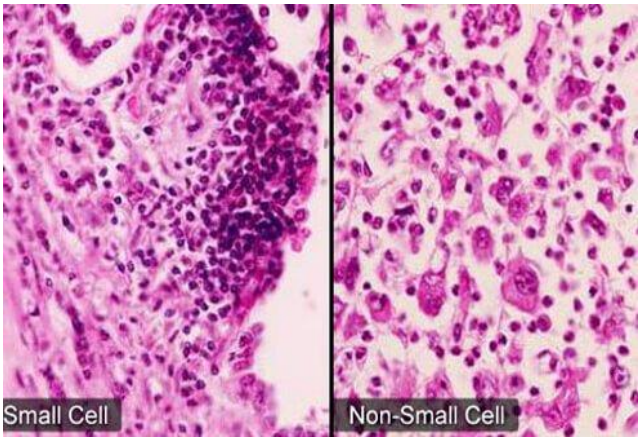


Figure 1.2: Microscopic images of SCLC (left) and NSCLC (right) (https://www.onhealth.com/content/1/lung_cancer Retrieved 7 February, 2018)

Small cell lung cancer (SCLC) and non-small cell lung cancer (NSCLC) can be considered as the two main types of lung cancer (see Figure 1.2). NSCLC is the most common type but it is less aggressive, spreading to other tissues and organs far more slowly than the less common SCLC. In SCLC, cancer cells are very small when observed under microscope and they are derived from epithelial cells to become solid tumors. On the other hand in NSCLC, the cells are bigger when compared and they are also derived from epithelial cells to become solid tumors. NSCLC can be further divided into squamous cell carcinoma, adenocarcinoma and large cell carcinoma subtypes. Most common indications of lung cancer can be considered as chest pain, loss of weight, coughing up blood and chronic coughing. Although cancer cells from other organs can also spread to the lungs as metastases, generally they are not classified as lung cancer. Some common lung metastases are breast cancer, prostate cancer, bladder cancer and colon cancer.

Apart from the type, the stages, which is basically the scale that shows how much the cancer has spread in the body, of the lung cancer is also important. SCLC can be distinguished into two main stages. In its limited stage, cancer cells are limited to the one side of lungs or lymph nodes near the lungs. Where, in its extensive stage cancer cells are spread to both lungs, to lymph nodes on the other side and even to other parts of the body. Different from SCLC, NSCLC can be distinguished into six stages. In its “occult stage”,

Table 1.2: Survival rates of the patients with NSCLC with different stages of the disease. Table adapted from (Mountain, 1997).

Stage of the disease	Survival rate in percentage after 5 years of treatment
I	%61
IIA	%34
IIB	%24
IIIA	%13
IIIB	%5
IV	%1

tumor location cannot be identified and the cancers cells can only be detected in a sputum cytology exam. In “stage 0”, cancer cells can be detected in the top layers of air passages. In “stage I” a non-spread small tumor can be detected. “Stage II” follows the stage I, where the tumor size increases and it spread to lymph nodes near the lungs. In “stage III” cancer has spread to the same side of the chest where it has started or to the opposite side of the chest or above the collar bone. Finally, in its most advanced “stage IV”, lung cancer has spread to the both sides of the lung, it can be detected in the fluid surrounding the lungs or it can be discovered in the fluid surrounding the heart. Stages II and III are also divided into two subclasses. Table 1.2 shows how the survival rate decreases dramatically in NSCLC patients when the cancer stage increases. This clearly indicates the importance of the early diagnosis of the lung cancer where PET imaging plays a crucial role.

As for the treatment options; for early stages surgery is usually the first option, sometimes followed by chemotherapy and radiotherapy. For advanced stages chemotherapy is the main option sometimes followed by radiotherapy and may be surgery. Combination of all these methods can be used together throughout the treatment. Patient’s condition and doctors directions play an important role in the determination of the treatment options.

In hospital environment, when scanning for lung cancer, the first step for the patient is to undertake conventional x-ray radiography or a more advanced CT scan. As a result of these scans, abnormalities in the lungs, i.e. lesions can be detected. If the detected lesion is

smaller than 3cm it is generally identified as a nodule. If it is larger, it is generally identified as a mass. Detection of a lesion is not always an indication of a cancer. Other diseases like tuberculosis, inflammation of the lungs and pneumonia can also lead to the formation of lesions in the lungs. In this regard proper diagnosis of the malignancy is of great importance. Even though some morphological properties of the lesions observed under conventional radiograph or CT scan can be useful to assess the malignancy of the detected lesions, usually these properties alone are not sufficient enough to diagnose the malignancy accurately (Erasmus et al., 2000). Invasive methods like biopsies or thoracoscopic surgeries can be undertaken to carry out the diagnosis more accurately (Rohren et al., 2004). Disadvantages of such procedures are that invasive procedures carry high risks and complications. So, because of these disadvantages, if these invasive methods are skipped by the doctor, standard clinical routine involves follow-up conventional radiographs or CT scans over a course of 3 to 6 months to carry out radiological assessment to diagnose malignancy of the lesion through its growth.

Since malignant lesions have an increased glucose metabolism when compared to benign lesions, ^{18}F-FDG PET scans have the ability to detect and diagnose malignant lesions early on without the risks and potential complications of the invasive methods. Because of this increased glucose metabolism, lung cancer cells accumulate ^{18}F-FDG, annihilating more electrons in the cancerous tissue, which are in turn detected by the PET detectors and thus creating bright areas on the PET images (see Figure 1.3). When coupled with the precise structural imaging provided by CT or MRI, PET scans provide clinicians with an effective non-invasive method for early detection, localization and evaluation of lung lesions (Beyer et al., 2000). This effective early diagnosis allows the patient to start treatments earlier, providing more treatment options and increasing survival rate dramatically.

Standardized uptake value (SUV) is the standard measurement for the PET images and it is calculated as follows:

$$SUV = \frac{radiotracer\ concetration}{injected\ dose/body\ weight} \tag{1.1}$$

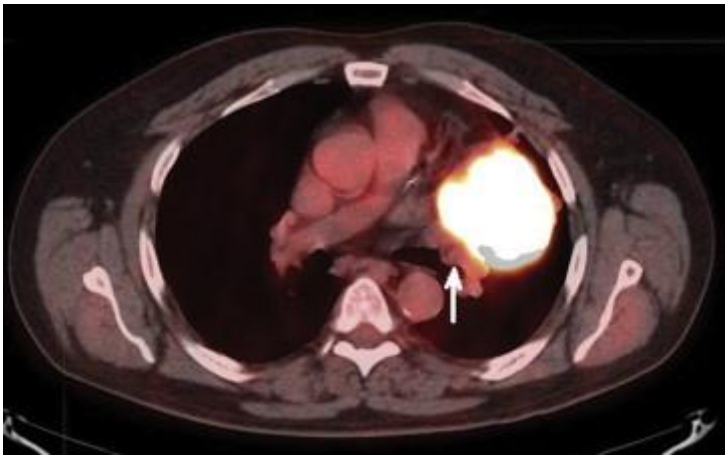
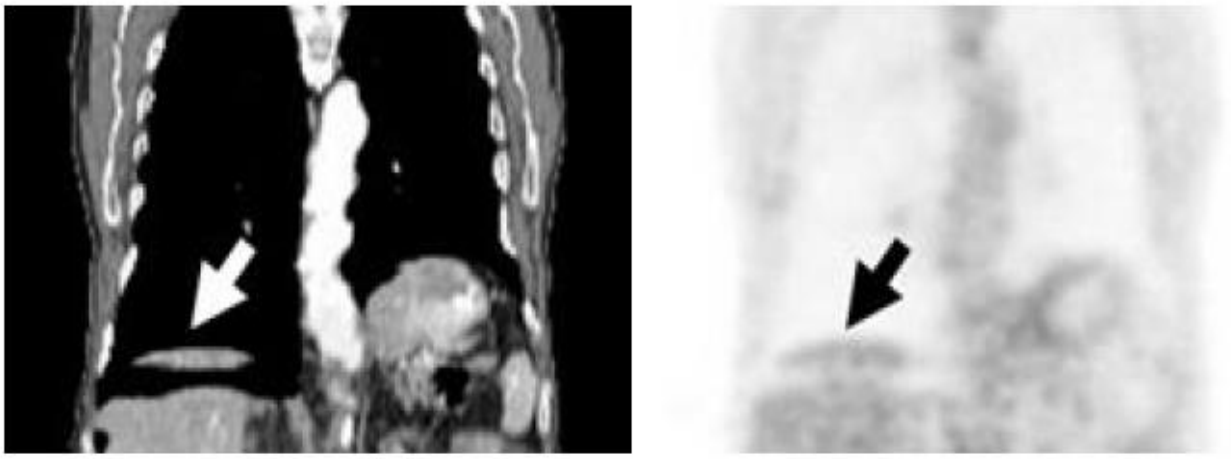


Figure 1.3: Lung cancer appears as a bright area on the ^{18}F-FDG PET scan, indicating a high metabolic area (Mahowald et al., 2015)

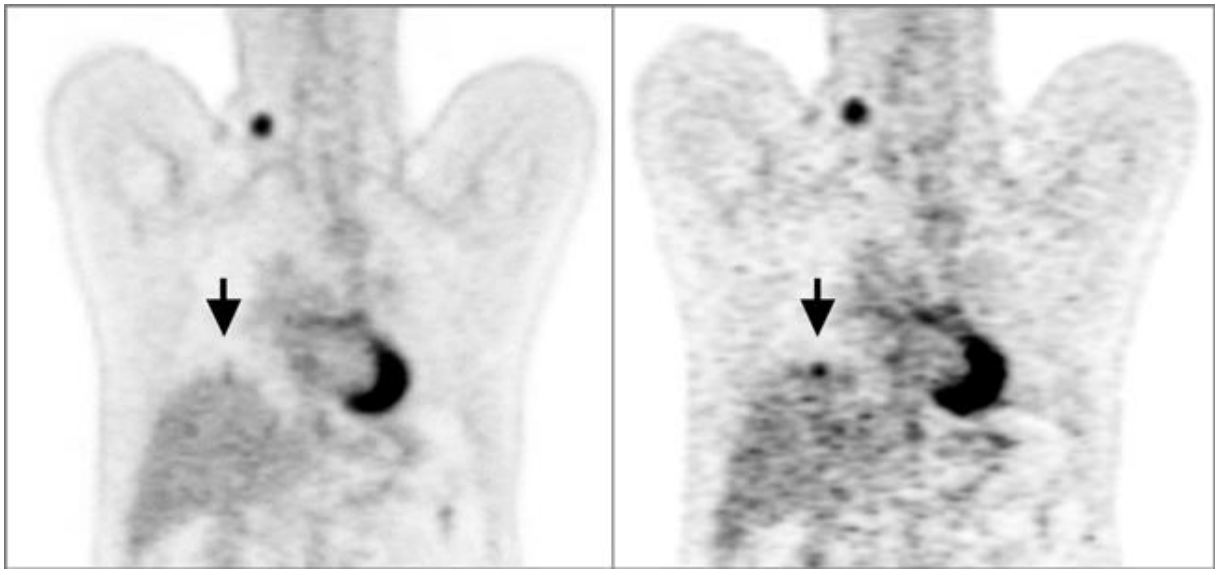
To distinguish lung cancer in PET images a SUV level of 2.5 is used. If SUV value is above 2.5 a lung lesions can be considered as malignant, i.e. cancerous. If it is below 2.5 it can be considered as benign. Although this is true, sometimes partial volume effects (PVE) in the image can cause miss diagnoses specifically for small lesions. In those cases, a small lung lesion with a SUV value below 2.5 can also be a malignant lesion. So, effective use of this measurement method along with the correct image reconstruction method is necessary for proper diagnosis.

Because of the small size of lung tumors and the low spatial resolution nature of the PET images there is always a risk of faulty diagnosis. Thus, improving PET image quality is one of the main research areas in the field. One of the main reasons of poor PET image quality in pulmonary PET imaging is the patient's natural motion, such as breathing and heart beating, during the scan. During a whole body PET scan, which last about 15 to 30 minutes, it is not possible to prevent involuntary and sometimes voluntary motion of the patient. While voluntary motions can be classified as the slight movements of the body, limbs or the head, mostly the patient carries in order to relieve pain or pressure during the scan, involuntary motions are the motions that the patient cannot control directly, such as periodic movement of the organs during natural cycles like breathing and heart beating. Because of these motions, the position of the organs can change by several centimeters



a) CT b) PET

Figure 1.4: Artifacts due to the movement of the liver during respiration (Townsend, 2008)



a) Standard PET b) Respiratory gated PET

Figure 1.5: a) PET scan with the almost non-visible blurred lesion (arrow) due to respiratory motion; b) Respiratory gated PET scan to eliminate respiratory motion blur effect. Arrow points the clearly detectable small tumor (http://depts.washington.edu/imreslab/currentResearch.html Retrieved 7 February, 2018)

during the acquisition (see Figure 1.4). Variation in the positions of the organs during the scan leads to some spread of the radioactive activity in the field of view of the device over an area proportional to the magnitude of the motion, worsening the reconstructed PET image quality and causing motion blur in the images (see Figure 1.5.a). In the special case of PET imaging for lung cancer diagnosis, eliminating the respiratory motion during breathing is the biggest challenge. As a result of the respiratory motion, several adverse effects are imposed on the lung PET images, thus creating an adverse impact on consequent clinical diagnosis. Blurring of the images, degradation of the spatial resolution and problems in attenuation correction of PET can be considered as the main effects. Due to these, reduction in lesion intensity and overestimation of lesion size can be observed (Devaraj, Cook and Hansell, 2007; Erdi et al., 2004)

Therefore, important efforts were taken by the researchers to solve the natural motion problem during lung PET acquisitions to improve the overall quality of the cancer diagnosis (see Figure 1.5.b). When carrying out research for eliminating natural motion based artifacts, one important point to consider is to reduce the exposure of patients to radioactivity which can result from the repeated intake of radioactive tracers during the repeated scans undertaken that are necessary for the development and testing of new methods. Use of powerful PET scan simulators to develop and test motion compensation techniques is an effective approach in this regard.

1.2 Respiratory Motion Correction

During the natural inspiration and expiration, physical volume of the thoracic space changes. This volume change causes the positions of most organs found in the torso region, such as lungs, liver, spleen, pancreas, kidneys, prostate and even the heart to move and shift their locations along with possible pathologies inside those organs. Generally respiration is periodic but in inspiration and expiration phases, the affected organs does not follow a similar path.

When a PET acquisition is performed, these natural respiratory motions of the organs degrade the image quality and have an adverse effect on acquired data quantification. This is an important situation which can lead to faulty clinical diagnoses because of the

introduced blurring effect and the reduction of the reconstructed images` contrast levels along with the measurement errors in radioactivity concentration (Yu et al., 2016).

Various techniques can be applied for correcting respiratory motion artifacts (Rahmin et al., 2007). Of these techniques, an accepted approach is gating. In gating technique, an external device is coupled with the PET for the registration of the respiratory motion phases in the imaged organ; then the signal obtained from this device is used divide the PET emission data into partitions that are synchronized with the various parts of the respiratory cycle (McClelland et al., 2013). Even though gating is an accepted approach it usually generates images with low signal-to-noise ratio (SNR) (Li et al., 2006) and separate devices like spirometers, chest belts with pressure sensors and optical systems for tracking body position are usually necessary for recording motion of the patient during respiration. Instead, other techniques that rely on data driven methods can also be used for providing motion characterization. After models regarding the respiratory motion are obtained, image-based registration (Fulton et al., 2002) or motion-compensated image reconstruction (Lamare et al., 2007) is used for correcting respiratory motion.

Common motion characterization techniques currently used in the field can be grouped as PET-derived techniques, MRI-based techniques and joint-prediction techniques (Catana, 2015). These techniques involve obtaining motion fields which give information about the locations of the organs throughout the different respiratory phases.

PET-derived techniques involve deriving the motion fields directly from the PET images. In these methods optical flow algorithms are implemented in which the respiratory cycle is separated into several phases and the transformations between the corresponding PET images at each cycle phase are predicted. Intensities of the image pixels after the motion are related to the optical flows (velocities) in each direction along with the variations in pixel intensities in the corresponding directions for carrying out the prediction (Dawood et al., 2008). Motion fields can also be obtained using B-spline deformable registration instead of using the optical flow technique (Bai and Brady, 2009).

In the recent years, a hybrid configuration of the PET device, the PET/MRI become more available for clinical use and researches developed various MRI-based motion characterization techniques along with it (Dutta et al., 2014). These techniques range from

simple methods like generating respiratory motion models from repeated 2D MR image acquisitions over several respiratory cycles (Wurslin et al., 2013) or using 3D self-gated radial MRI gradient echo sequences (Grimm et al., 2013) to more complex methods like tagged MRI, phase contrast MRI and pulse field gradient methods (Ozturk et al., 2003). Additionally, in MRI-based methods the disadvantage of the PET/CT in terms of the imposed extra radiation dose by its CT component is eliminated with the ionizing-radiation free imaging capabilities of the MRI.

Also the information gathered from both PET and MRI can be combined for implementing the motion characterization. While PET presents the respiratory surrogate signal for the motion model by applying principal component analysis (PCA), 2D multi-slice MRI presents the imaging input of the model (Manber et al., 2016).

After motion characterization is completed, one of the pre-reconstruction, reconstruction or post-reconstruction techniques can be implemented for carrying out the respiratory motion correction of the PET image data. Pre-reconstruction methods involve compensating the respiratory motion before reconstructing the images from the raw PET data. For example, each detected event in a pair of detectors is reassigned to another pair of detectors based on the derived motion fields (Livieratos et al., 2005). Differently, in reconstruction based motion compensation techniques the obtained motion model is incorporated in to the reconstruction algorithm to modify the PET system matrix directly. Examples exist for common reconstruction algorithms; list-mode maximum likelihood expectation maximization (MLEM) algorithm (Guerin et al., 2011) or list-mode ordered-subsets expectation maximization (OSEM) algorithm (Chun et al., 2012). In these techniques, a motion-warping operator is interpolated from the motion fields and used for modifying the original system matrix. Since all the detected events by the PET are counted in these methods, they produce images with improved image quality when compared to the conventional gating techniques. Thus they can be considered as an ideal approach. Moreover, simultaneous image reconstruction and motion characterization (which reduce the motion blur and increases the SNR) can be implemented for further improved image quality (Blume et al., 2010). Post-reconstruction techniques are another alternative for the respiratory motion compensation. Motion fields obtained either from PET or MRI can be

used to co-register the already reconstructed images at various phases of the respiratory cycle to a common reference frame (Wurslin et al., 2013; Lamare et al., 2014).

1.3 Computer Aided Tumor Detection

In addition to the aforementioned important roles of PET imaging in lung cancer diagnosis, one other main application area of PET imaging is in planning for lung cancer radiation therapy (Feng et al., 2009). During radiation therapy to treat cancerous lung tumors, implementation of advanced techniques like, image guided radiation therapy (IGRT), 3D conformal radiotherapy (3DRT), intensity-modulated radiation therapy (IMRT) and computer assisted 3D planning is of great importance. Implementation of such techniques guides radiation therapy devices, such as medical linear accelerators (LINACS), to focus destructive high radiation dose only to the unwanted tumor tissue while keeping the damage to healthy tissues at minimal levels (MacManus et al, 2009). In order to achieve this with high precision, very accurate detection and segmentation of the lung tumor tissue from the surrounding tissues is very important. With accurate segmentation, radiation therapy device can focus high radiation dose to the segmented tumor region, destroying cancer cells without causing unwanted harm to the surrounding cells of the healthy tissues.

Before the widespread of PET devices in hospitals, structural imaging devices, like CT and MRI, were mainly used to detect and segment the anatomical perimeters of the lung tumors. With PET becoming more common, it is combined with structural methods to provide additional functional information to improve the segmentation of the lung tumors from surrounding tissues. Functional information provided by the PET allows the segmentation of the functional perimeters of the tumor, providing accurate information about the active parts of the tumors. With this addition, clinicians can segment tumors, evaluate treatment responses and predict survivals with far more precision (Erdi et al., 2002).

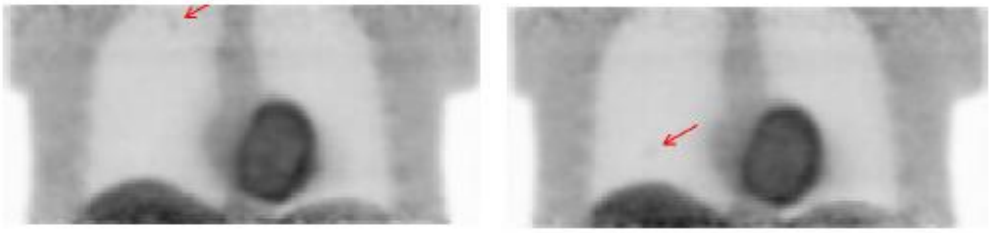
In common clinical routine, detection and segmentation of lung tumors is carried out by clinicians, mainly radiologists, manually or semi-automatically. First, doctor goes over series of PET/CT or PET/MRI images of the patient slice by slice to accurately detect the presence and the location of the lung lesions. Then proper diagnosis is carried out based on

the prior knowledge of the doctor and sometimes the additional information provided by other pathological examinations. After lung tumors are detected and diagnosed, clinician either manually delineates the tumor boundaries, i.e. manually draws the tumor regions carefully, or guides the semi-automatic systems of the imaging or therapy modalities to obtain the segmented tumor. Obtained results are manually annotated and presented. This procedure is far more same in most of the other medical pathological image segmentation tasks, like brain tumor segmentation and so on.

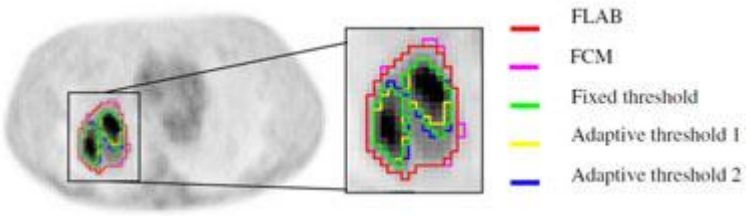
Clinician's manual involvement in these procedures means they are time consuming, subjective i.e. radiologist dependent, and highly dependent on prior knowledge, expertise and manual-visual capabilities of the clinician performing the procedure. In this regard, results of such manual segmentations are subject to errors and large intra and inter rater variability. Because of these concerns, development of robust computer aided automatic tumor detection and segmentation methods i.e. computer aided detection (CAD) systems, to provide efficient and objective detection and segmentation results, became a very interesting and popular research area in all of medical imaging fields in the recent years (Işın et al., 2016).

In lung cancer detection and segmentation, manual segmentation can be far more challenging for the clinician. Small lung lesions can cover areas as small as 2 or 3 voxel diameter on the images. Also their contrast levels can be very insignificant when compared to the contrast levels of the surrounding tissues (see Figure 1.6). These make visual detection and manual delineation by the clinician a very tough process, resulting in missed tumors during detection. Therefore, automatic detection provides invaluable assistance to the clinicians for accurately reading and analyzing oncological images, thus ensuring excellence in diagnosis and treatment of lung cancer.

In fully automatic tumor detection and segmentation techniques, no user interaction is required. Mainly, almost all well-known image processing techniques along with machine learning and artificial intelligence methods can be implemented to carry out the automatic detection and segmentations. Even in some methods prior knowledge is combined to solve the problem. Automatic detection and segmentation methods can be mainly classified as discriminative or generative methods (Işın et al., 2016).



a) Tumor on top of the right lung b) Tumor at the bottom of the right lung



c) An example automatic tumor segmentation

Figure 1.6: a) & b) PET images of an 8mm lung lesion at different locations of the lung. Arrows show lesion locations. Due to small size and insignificant contrast these lesions can be easily missed by the radiologist during manual/visual inspection. c) An example CAD system, representing segmentation performance of different techniques (http://medicalphysicsweb.org/cws/article/research/50874 Retrieved 7 February, 2018)

Discriminative methods require ground truth data to learn the relationship between input images containing the tumors with the ground truth to carry out decisions. Generally these methods involve extracting features from the images using different image processing techniques.

Deciding which features to use is of great importance in these techniques. In most cases final decision is made by using supervised machine learning techniques which, in order to perform well, require large image datasets with accurate ground truth data. In contrast, generative methods require prior knowledge, such as location and spatial extent of healthy tissues to generate probabilistic models, which carry out the final segmentation. Prior obtained maps of healthy tissues are implemented to segment the unknown tumors areas.

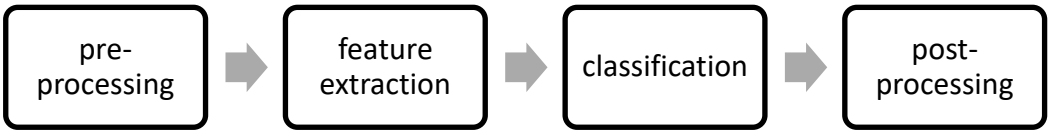


Figure 1.7: Example of a common processing pipeline for tumor detection and segmentation

Developing suitable probabilistic models by using prior knowledge is however a complicated task.

In application, almost all discriminative methods employ a similar process, named as a processing pipeline (see Figure 1.7). Typical processing pipeline for tumor detection and segmentation starts with pre-processing followed by feature extraction, classification and post-processing procedures.

In the first step of pre-processing, filtering operation to remove possible noises from the images and operations like intensity bias corrections can be carried out. In feature extraction step, most of the well-known and common image processing techniques can be implemented to extract different features to define the differences in target tumor tissues and healthy normal tissues. Many different features including, asymmetry-related features, contrast levels, intensity gradients, size information, first order statistical features, raw intensities, local image textures and edge based features can be extracted from the images for both healthy and tumor tissues, which is used to make a classification in the next step. In classification step, different types of classifiers like, artificial neural networks (ANN), k-nearest neighbor classifiers (kNN), self-organizing maps (SOM), support vector machines (SVM) and random forests (RFs) are implemented to make the decision of assigning an image pixel either to healthy tissue class or to tumor tissue class. Some applications require the results of the previous steps to be refined to increase overall detection and

segmentation performance. Techniques like conditional random fields (CRF) and connected components (CC) are used to carry out this post-processing (Işın et al., 2016).

Despite the aforementioned CAD methods for tumor detection and segmentation were very common and successful in previous years; an emerging technique of deep learning began to replace them in the recent years. As the main deep learning method, convolutional neural networks (CNN) obtained state-of-the-art performances in many of the well-known object recognition challenges (Krizhevsky et al., 2012). These marginal performances allowed deep learning methods to become highly recognized also in the field of medical image processing. In previous research, by obtaining record performances, application of the deep learning methods to the most complex medical tumor detection and segmentation tasks proved to be very effective (Işın et al., 2016). The main advantage of CNNs is that, due to their very deep, i.e. many in number, computational layers, they learn highly representative complex features directly from the input images given to them. Oppositely, in traditional automatic classification applications, features representing the differences in tissue classes need to be extracted by hand using the aforementioned image processing techniques. Extracting highly representative features from the input images to be used for the classifier has the most powerful effect on the performance of computerized tumor detection and segmentation applications. However, handcrafting these features requires high skill and knowledge. It is also very time-consuming, involving most of the work and generally selected features are not robust with respect to the variations in the image data. Since CNNs automatically learn these complex representative features, the burden of feature handcrafting is eliminated and the performance of the classification is greatly enhanced. As a result of this, instead of trying to develop better image processing techniques for better feature extraction, current research on developing CNN based techniques for tumor detection and segmentation greatly focuses on designing new and better network architectures. Figure 1.8 illustrates an example deep learning architecture for tumor detection.

Despite its clear improvements over traditional methods, implementing deep learning techniques also have some hassles. Training a deep convolutional neural network requires very large annotated training image dataset for improving performance by increasing the number of convolutional layers. Also, increasing the network depth increases the

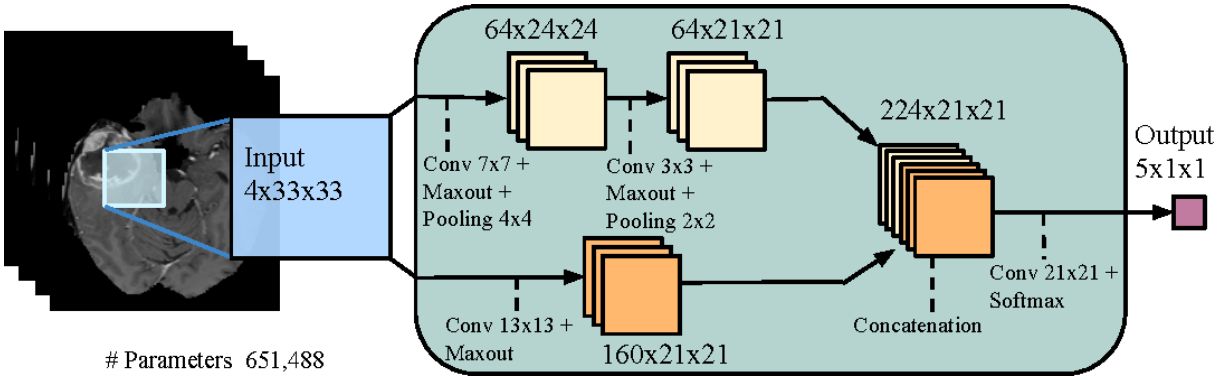


Figure 1.8: Example deep learning architecture for brain tumor detection (Havaei et al., 2017)

computational cost due to an increase in the complex operations carried out in the deep convolutional layers. In this regard, designing and training deep convolutional neural networks require powerful GPU powered computers, and even in some cases GPU powered super computers. Scarcity of PET, PET/CT and PET/MRI modalities makes the availability of such large annotated image databases even a more difficult task.

An adaptation of deep learning, namely transfer learning (and its variation transfer learning with fine tuning) (Tajbakhsh et al., 2016) presents an effective solution for this problem. In conditions where limited training image data, not enough machine learning expertise and limited computational resources are available, researchers can use transfer learning as an efficient deep learning application. Basically, transfer learning means that, a pre-trained deep learning system is imported to be used as an efficient feature extractor for the desired application in question (Işın and Ozdalili, 2017). To be more explanative, a deep learning framework, such as a convolutional neural network, that is previously trained on a large annotated general image dataset (does not need to be medical) where it has obtained high performance can be imported for a medical imaging application like lung cancer classification. This imported pre-trained CNN can be used as an automatic feature extractor for extracting highly representative features from the PET lung cancer images. Automatically extracted features are then delivered as an input into a more conventional, computationally more cost effective and easier to implement classifier for carrying out the final classification between normal healthy and cancerous tissues. One step further, one or more convolutional layers of this pre-trained network can be trained again, i.e. fine-tuned,

with the PET lung cancer image data in question. This method is then called transfer learning with fine tuning.

Implementation of such transfer learning technique to detect and segment lung cancers in PET images is feasible for developing a robust and efficient lung cancer CAD system to provide precise and objective segmentations for radiation therapy planning and to provide automatic diagnosis assistance to the clinicians. With this achievement, tough process of visual detection and manual delineation of the lung tumors by the clinicians, which can result in missed tumors, can be assisted by the automatic detection system which in turn provides invaluable assistance to the clinicians for accurately reading and analyzing oncological images, thus ensuring excellence in diagnosis and treatment of lung cancer.

Most of the previous research for developing such CAD systems using PET images for lung tumor detection relies on private PET data acquired directly from patients in clinics and the ground truths are prepared manually by expert radiologists in that institutions (Wang et al., 2017; Hanzouli-Ben et al., 2017; Kopriva et al., 2017). Unfortunately, at the time of application there were no publicly available PET lung cancer image databases with extensive ground truth present that could be used in the proposed transferred deep learning based CAD system. Due to there is only one PET/CT device available in North Cyprus and getting the acquired images evaluated by expert radiologist would be costly and would require so much time, the option of creating our own database was not available to us in this study. Using simulation data is another option but that would be not representative enough for the observable variations in real clinical cases and also would require expert annotations and delineations. In this regard, to develop a CAD system for detecting lung cancer, we decided to use lung CT databases already available with expert annotations. However, since CT lacks the functional metabolic activity information, which can be provided by the PET, detecting and deciding about the pathology of lung lesions with automated systems by only using lung CT images is a very difficult task and most of the time not possible (Baker et al., 2017). Although, morphological information of the lesions detected in CT images can give the radiologists hints about the pathology, further pathological analysis, as mentioned previously, and/or motorization of the development of the lesion with follow-up scans over a long period of time is required to achieve accurate diagnosis. So instead, detecting lung lesions/nodules without diving into pathology is a

more feasible task. At a later stage, with the availability of public PET databases or with the development of our private database, functional information of the PET can be incorporated to the system to give additional diagnosis assistance on that detected lung lesions.

1.4 Contributions

In this thesis the following contributions are made:

- Simulation of a PET/MRI device is performed in GATE environment using a computerized lung lesion induced XCAT phantom.
- Several motion correction methods are developed, incorporated into OSEM reconstruction algorithm and tested in the GATE simulation environment for the compensation of respiratory motion artifacts in PET images for lung cancer imaging.
- AlexNet deep learning framework is transferred into the medical imaging task of lung lesion detection.
- Transferred AlexNet is used as an automatic hierarchical feature extractor for extracting features from lung CT images.
- Two other non-deep learning based feature extraction methods are developed for comparing transferred deep learning method.
- Developed feature extraction methods are used in the development of a high performing CAD system for the detection of lung lesions from lung CT images.

CHAPTER 2
THEORETICAL BACKGROUND

This chapter presents the detailed technical background on the main imaging modalities used for lung cancer imaging. Technical details of deep learning methods for CAD system development are also introduced.

Imaging modalities used in medicine can be classified as the devices that provide structural information or functional information or hybrid devices which provide both types of information. Devices that give images of anatomical structures inside the body are: x-ray radiography, CT and MRI. PET provides functional information about the metabolic activities inside the patient's body where hybrid devices like PET/CT and PET MRI combines both structural and functional imaging.

2.1 Conventional X-Ray Radiography

Conventional X-ray radiography is based on the principle that, when they delivered X-rays can penetrate through the human body. While they penetrate through the human body they lose some of their energy, i.e. attenuate, due to interactions with the tissue material. Attenuation properties of the different types of tissues are different from each other. This is mainly due to their density properties. Hard tissues like bones attenuate the most, while air and fluids inside the body attenuate the least. This allows the intensity differences of the x-rays after penetration and attenuation through the body to be mapped on plain fluorescent films or on digital detectors creating the x-ray image (Bettinardi et al., 2002). Since less attenuation means the passed through x-ray has more energy, air and fluids appear darker on the x-ray image. High energy x-ray "burns" the film, turning it to darker tone. In contrast, bones and other calcified structures allow less x-ray energy to pass through them, in turn appearing white and well defined on the film. Low energy x-ray cannot "burn" the film leaving it white. Soft tissues, which have medium attenuation properties, appear grey on the x-ray image. Fat tissue is an exception, which appears little darker. See Figure 2.1 for an example X-ray image.

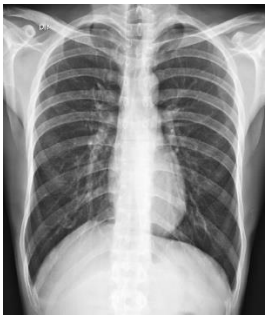


Figure 2.1: An example chest x-ray image (https://radiopaedia.org/cases/normal-chest-x-ray Retrieved 10 February, 2018)

Due to these properties, x-ray radiography is mainly used for imaging bones for the purpose of detecting fractures. Abdominal scans and scanning chest for lung cancer and pneumonia detection are among other common uses.

Generally, x-ray radiographs are taken along a defined projection angle which maps the x-ray attenuation through the body. Three projection angles of posteroanterior (PA), anteroposterior (AP) and lateral scans are commonly used for chest scans. Conventionally, x-ray images are recorded using fluorescent films. These films require processing by liquids in film baths before image can be formed. However, in modern radiography, digital electronic detectors are used to detect the incoming x-rays and turn them into electrical signals, where output signal level of each detector element is directly proportional to the x-ray energy incoming to that particular detector element. Digital images can easily be processed, stored and viewed by computers without the need for external processing.

Conventional x-ray radiography of the chest is a low cost, easy to perform and very quick imaging technique. Typical scans rarely lasts longer than 10 minutes and does not require extensive preparations of the patient. Because of these advantages it is generally preferred as an initial examination for the medical diagnostic procedures. However due to low soft tissue contrast and not including three dimensional information (image is presented as one slice of information on two dimensional plane from the chosen projection angle), proper evaluation is difficult and further scans using more advanced imaging modalities are usually required for precise diagnostics (Armstrong et al., 2010). It should also be noted that, x-rays used in radiography imaging are classified as ionizing radiation. So frequent or

lengthy exposures to x-rays are dangerous for the health and can increase the risk of cancer for the undertaking patients. Especially pregnant women are not recommended to undertake x-ray scans of any form.

Conventional x-ray radiography is performed using an x-ray device. While there are many different adaptations, like fluoroscopy devices, dental x-rays, mammography devices, angiography devices and so on, main components and working principles are far more similar.

The x-ray device is made up from four main components; the x-ray tube, the high voltage generator and film or flat panel detector. All of these components play important roles during the generation of x-rays delivered to the patient and during the formation of the final x-ray radiograph.

2.1.1 X-ray tube

In order for the x-ray device to generate medical images, an x-ray source with the following properties is required; it should produce necessary x-rays in short exposure time, it should allow user to vary the x-ray energy, it should produce x-rays in a reproducible way and it should be safe and cost effective. Despite there are other practical x-ray sources like radioactive isotopes, nuclear reactions such as fission and fusion and particle accelerators, only x-ray tubes (which are special purpose particle accelerators) meet all the aforementioned requirements.

In medical x-ray tubes there are two main parts. A cathode (negatively charged), which houses the filament that produces the free electrons, and an anode (positively charged) target where the free electrons are accelerated towards to generate x-rays.

During operation, cathode filament that is made from tungsten material is heated with an electric current, called the filament current, which causes electrons to be emitted from the filaments surface. Amount of electrons emitted is directly related to the amount of filament current applied. When very high positive voltage is applied to the anode with respect to the cathode (called the tube voltage), free electrons accumulated around the filament surface

are accelerated towards the anode target, producing a current inside the tube which is called the tube current.

Anode part of the device is a metal target electrode for the accelerated electrons to hit and it is kept at a positive voltage relative to the cathode. Accelerated electrons of the tube current hit the anode material, depositing most of their energy as heat. Small fraction of the remaining energy is emitted as x-rays which is directed and focused towards the patient.

The relationship between the tube and the filament current is directly dependent upon the tube voltage. The user can adjust the tube voltage and the filament current to generate desired x-ray energy levels for the desired medical imaging application.

2.1.2 High voltage generator

The main function of the high voltage generator is to generate and deliver current at a high voltage to the x-ray tube for the generation of the tube voltage. Due to electrical power available in hospital can be around 480 Volts at maximum, which way lower than up to 150.000 Volts required by the x-ray tube to accelerate electrons, high voltage generator is used to step up low input voltage into the required high voltage by using transformers.

Another property of the high voltage generator is that it converts alternating current produced by the transformers into direct current. The reason for this conversion is that x-ray tube operates with a direct current. If an alternating current is applied to the tube, back propagation of the electrons could occur during the part of the alternating current cycle when the cathode is positive and anode is negative. If anode is very hot at that stage, electrons can be released from the anode surface and accelerated towards the filament, which can destroy the filament causing the x-ray tube to malfunction. High voltage generator uses rectifier circuits, made up from diodes, to convert alternating current into direct current,

2.1.3 Film or flat panel detector

This is the part where the attenuated x-rays coming through the patient's body is detected and converted into image. In conventional x-ray devices films, that are similar to a

photographic film, are used to form the image. These films require development/processing with chemical fluids in order for the image to be produced. While development takes time and requires chemical fluids, the film itself is inexpensive when compared to the flat panel detectors. But it should be noted that at least one or more films are used per patient.

In modern devices, flat panel detectors are used in digital radiography to detect the incoming x-rays. In indirect flat panel detectors, x-rays react with the scintillator crystals (caesium iodide or gadolinium oxysulfide) of the detector creating visible light which in turn detected by the semiconductors, i.e. amorphous silicon photodiodes, and converted into electrical signals. There are also direct conversion detectors where x-ray energy is directly converted into electrical signal without the need of x-ray to light conversion. Level of the generated electrical signal from each detector element is directly related to the energy of the incoming x-ray to that element enabling the generation of an image by the computer electronically. Use of the flat panel detectors are more sensitive than film and enable fast imaging, plus the image data can be viewed, stored and processed quite easily. In addition it requires lower x-ray dose than the film to produce a similar quality image. However, flat panel detectors are very expensive and can be easily damaged if dropped by the user, rendering the detector and the x-ray device unusable.

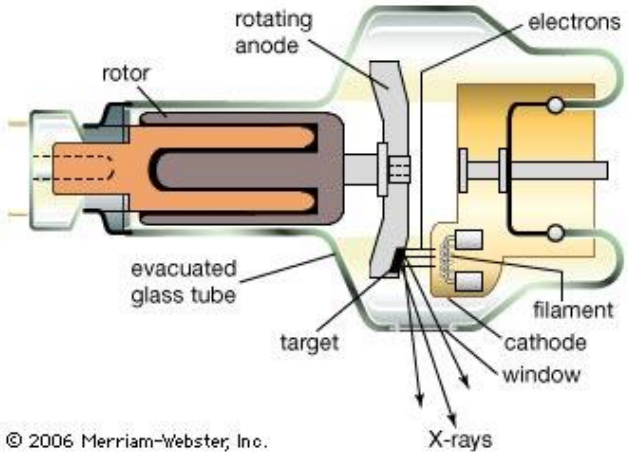


Figure 2.2: X-ray tube (http://www.wikiradiography.net/page/Physics+of+the+X-Ray+Tube Retrieved 18 March, 2018)

2.2 Computed Tomography (CT)

In conventional x-ray radiography, patient`s three dimensional anatomy is reduced into a two dimensional projection image. Intensity of a pixel on a radiograph represents the attenuation properties of the tissues and other structures within the patient along the projection line, i.e. a line between the focal spot of the tube and the point on the flat panel detector or film corresponding to that pixel point. Because of this, all anatomical information that lies parallel to the x-ray beam cannot be represented on the image. In clinical practice, to overcome this disadvantage, two images with perpendicular projection angles can be taken. For example, in chest scans a lateral projection image of the patient can be taken to provide depth information to a standard PA image. This provides better special localization for the objects that can be identified in both projections. However, for the diagnosis of complex medical pathology this technique is not sufficient.

Basic principle behind tomography is that, image of an unknown object can be obtained by taking infinite projections through that object. To provide more location information, instead of taking two projections, several projection images (to be specific, 360) can be acquired with 1-degree angular intervals around the patient's chest. With this technique, it could be possible to obtain a similar data to a chest CT scan. Although it is possible in theory, that data would present anatomical information in a way that it would be impossible for a clinician to interpret. However, if all that data is transferred into a powerful computer, the computer can reformat the data to reconstruct a chest CT examination.

Similar to x-ray radiography, CT imaging also uses x-rays as a source. Technical principles of the x-ray tube, high voltage generator and detector are similar with some modifications and improvements in designs. Difference is in that CT provides 3D image of the x-ray attenuation properties of the patient's body when compared to the 2D image of conventional radiography. In basic principle, a similar x-ray tube to radiograph emits the x-rays which are detected by the series of detectors positioned at the opposite side of the patient. Then the tube and the detectors rotate around the patient in a synchronized manner, as shown in Figure 2.3, so that attenuation information at different projections is acquired.

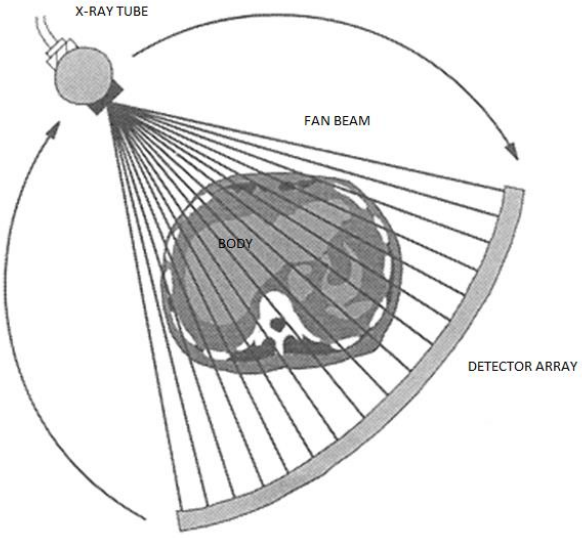
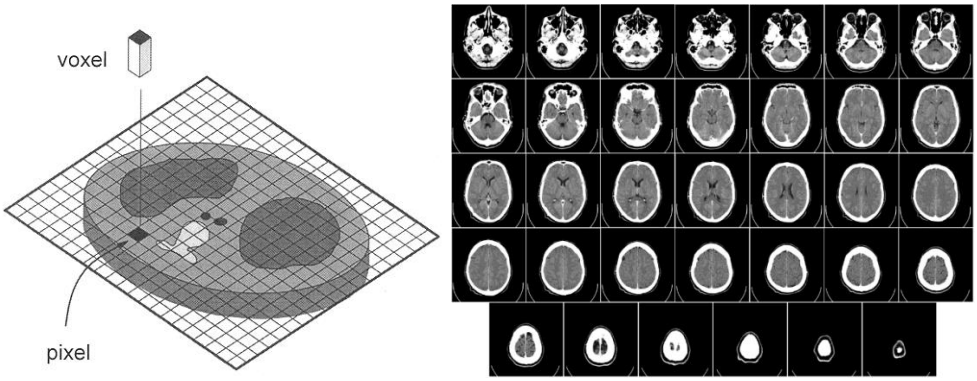


Figure 2.3: Representation of the operation of CT scanner. X-ray tube and detector array rotates around the patient in synchronized manner (Beutel et al., 2000)

All these acquired data is then processed by powerful computers and CT images are reconstructed by using reconstruction algorithms. A CT image is a picture of a very thin slice (0.5 to 10mm) through patient`s anatomy. Every two dimensional CT image slice provides information about a very thin three dimensional section of the patient. In that regard, a 2D pixel on a CT image corresponds to a 3D voxel within the patient's anatomy. Intensity of each pixel provides information about the average x-ray attenuation properties of the tissue/tissues in the corresponding voxel (see Figure 2.4).

In earlier CT designs, acquisition is carried out by using rotate-step-rotate principle. X-ray tube and the detectors rotate around the patient to acquire information related to a single CT slice and then information is translated to the computer. The patient table then moves by a step and the tube rotates again to scan the next slice. This procedure continues until the desired field is scanned completely. However in modern helical/spiral CT scanners, tube rotation and the table movement are simultaneous. With this type of movement, tube and detectors follow a helical path around the patient, as shown in Figure 2.5. Helical CT scanners cover greater volume than the earlier designs for the same acquisition time. Since it is no longer required to translate the patient table movement, total acquisition time



a) b)

Figure 2.4: a) Relationship of a CT image pixel and the corresponding voxel in the patient`s anatomy. b) Full series of CT brain scan slices (Beutel et al., 2000)

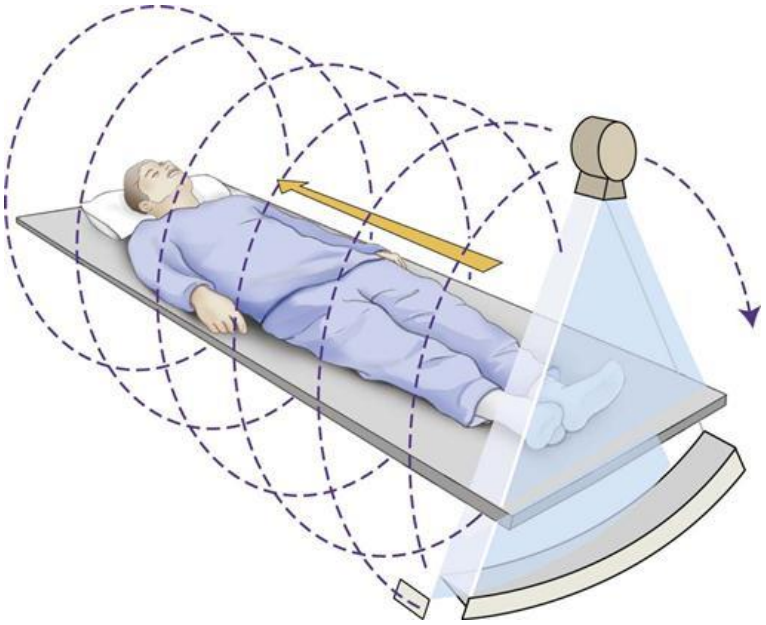


Figure 2.5: Helical path followed by the x-ray tube and the detector array through patient table movement in helical/spiral CT scanners (https://pocketdentistry.com/14-other-imaging-modalities/ Retrieved 7 February, 2018)

required to image the patient is reduced to a great extent along with the total radiation dose that the patient is exposed during the acquisition. One important parameter to consider in helical scanners is the detector pitch, which is basically the relationship between the table movement speed and the CT gantry rotation. Additionally, instead of using one detector array, multiple detector arrays can be incorporated into helical CT scanners to obtain multiple slices of images in single tube rotation (see Figure 2.6). These multi-slice/ multi-detector helical CT scanners use more than one, closely spaced detector arrays. With each tube rotation and non-helical movement, each of the detector arrays can acquire a separate image slice. When helical acquisition is added, multi-slice CT devices can achieve increased table speeds, increased pitch and increased coverage for a given period of time. This allows the tube output to be used more efficiently, scan time to be reduced to a great extent and longitudinal spatial resolution to be improved (Beutel et al., 2000).

CT scanners are frequently used in clinical environment for lung cancer diagnosis and follow ups. When any suspicious lesion or other condition is detected through x-ray radiography, CT scans are performed to provide more precise detection and diagnosis due to the previously mentioned limitations of the x-ray radiography (Alberts., 2007). CT scans have better tissue contrast than conventional x-ray radiography and provides three dimensional imaging thus more information for the clinical diagnosis. However, since repeated projections are required over a longer period, CT scanners deliver more radiation dose to the patient than the x-ray radiography. Where chest radiography has a radiation dose around 0.1 mSv, a chest CT scan radiation dose can be as high as 7 mSv (Townsend, 2008). When it is compared to the annual worldwide average dose as a result of natural background radiation, which is around 2.4 mSv, performing CT scans frequently for potential cancer candidates can have debatable benefits over potential harms due to excess radiation. To overcome this, imaging modalities like MRI, which does not impose any radiation dose to the patient, can be preferred for providing anatomical imaging for cancer diagnosis over x-ray based methods. Additionally, for imaging of the more complicated cancer types (in terms of morphology and imaging), like brain tumors, even though CT can also provide initial imaging, soft tissue contrast of the CT is not as good as the MRI`s and it fails to provide detailed information about the extent and the sub regions of the complex brain tumors.

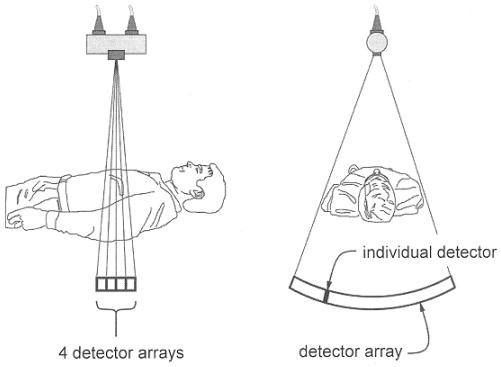


Figure 2.6: Representation of an early Multi-Slice/Multi-detector CT scanner (Beutel et al., 2000)

In the following two subsections, more technical details regarding the tomographic acquisition and reconstruction processes will be introduced briefly to provide better understanding of the concepts before moving to the introduction of the other imaging modalities used in cancer diagnosis.

2.2.1 Tomographic acquisition

In tomographic image acquisition, a single measurement of x-ray attenuation transmission made by a single element of the CT detector is referred as a ray. When several rays are transmitted through the patient with the same tube angle, this is referred as a projection or view. Parallel beam and fan beam geometries are the two main projection geometries that are commonly used in CT scanners. In parallel beam geometry all delivered rays in a projection are parallel to each other and in fan beam geometry they diverge from the tube and resembles to a fan shape. Most modern CT devices use fan beam geometry for image acquisition and reconstruction.

CT scanners take multiple transmission recordings through the patient at different projection angles. Thus a single CT image slice can involve total of 800,000 transmission records, with 800 rays delivered at each 1000 distinct projection angles. Depending on the device configurations and number of detector arrays these number can increase significantly. After the imaging of the slice is completed, patient table is moved in the "z-axis" of the device and the process is repeated for the next slice.

2.2.2 Tomographic reconstruction

With each ray measurement, CT detector measures an x-ray intensity value, I_t, which corresponds to the attenuated x-ray beam intensity through the patient. Along with this, un-attenuated intensity, I_0, is also recorded by a reference detector. Relationship of these parameters is as following;

$$\ln(I_0/I_t) = \mu t \tag{2.1}$$

where, t is the thickness of the patient and μ is the average linear attenuation coefficient along the ray. It is important to note that while I_t and I_0 depends on the CT device, μt is directly related to the anatomical properties of the patient along the measured ray path. This pre-processing calculation allows CT image to not depend heavily on device parameters, but to depend mainly on the anatomical properties. This gives the CT scanners their high clinical utility (Beutel et al., 2000).

Further, attenuation coefficient is converted into Hounsfield units (HU) and used in the representation of each pixel value on the image. HU conversion assumes the attenuation value of water as 0, attenuation value of air as -1000 and attenuation value of bone as 1000. This linear transformation is expressed as;

$$HU = \frac{\mu - \mu_{\text{water}}}{\mu_{\text{water}} - \mu_{\text{air}}} \, x \, 1000 \tag{2.2}$$

After the processing of the raw data, one of the numerous reconstruction algorithms is used to reconstruct the image. Most common one is filtered back-projection (FBP). This technique forms the CT image by reversing the image acquisition steps. The attenuation coefficient μ of each ray is projected back along the same path onto the image matrix. When data from the whole rays are back-projected, high attenuation areas reinforce each other, as the low attenuation areas do, forming up the final CT image (see Figure 2.7).

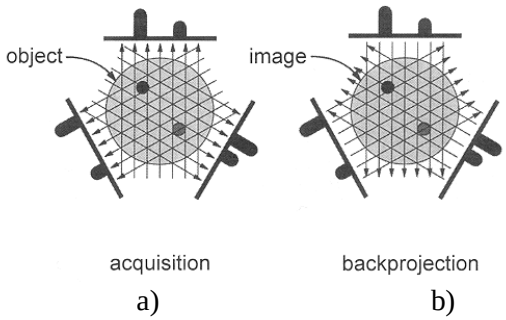


Figure 2.7: a) Acquisition phase. b) Reconstruction phase (Beutel et al., 2000)

2.3 Magnetic Resonance Imaging (MRI)

While x-ray radiography and CT depends on the use of x-ray attenuation properties of the tissues to acquire the images of the internal anatomic structures of the patients, magnetic resonance imaging depends on the atomic scale magnetism and radio frequency (RF) resonance properties of the tissues to form the images. Since, MRI uses more sensitive information of the tissue nuclei properties, rather than the attenuation properties, to form the tissue contrast in images, MRI images have far more better tissue contrast, as seen in Figure 2.8, with comparison to CT and x-ray radiography. Also since it does not rely on ionizing radiation it can be considered safer to use for imaging cancer patients. Additionally, images can be obtained in any anatomic plane without the need to move the patient.

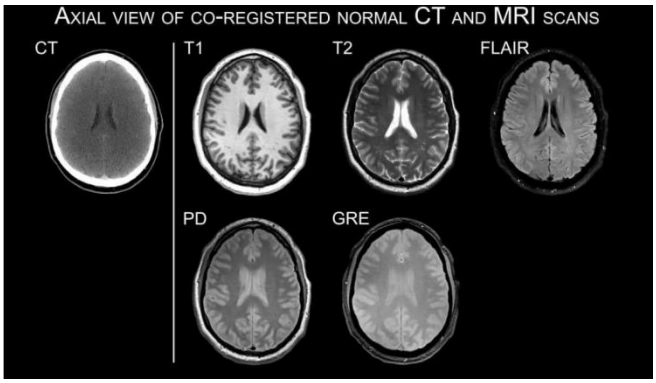


Figure 2.8: Brain CT image (left) and several MRI images for comparison (right) (https.//healthcareplex.com/mri-vs-ct-scan/ Retrieved 10 February, 2018)

Main working principle of the MRI device is based on nuclear magnetic resonance (NMR). It involves using very strong electromagnetic fields to align hydrogen nuclei in the tissues and analyzing their magnetic spin properties by delivering RF signals to excite them (Hashemi et al., 2004). Main components of the MRI scanner are; main magnet coil that is used for generating the powerful large magnetic field (from 1.5 Tesla (T) to 3T in clinical use) required for aligning the nuclei; gradient magnet coils that are required to obtain images from different anatomic planes (i.e. sagittal, coronal or transverse); RF coils to deliver RF signals required to excite hydrogen nuclei in the tissues and receive the echoes back as a result of the resonance; and a powerful computer system to reconstruct, process, store and present the final image.

2.3.1 Magnetic characteristics of the nuclei involved in MRI

The nucleus of the elements shows magnetic properties that are determined by the spin and charge distributions inherent to the proton and neutron. Positively charged protons generate a magnetic dipole when they spin. On the other hand neutrons generate a same strength magnetic field with the opposite direction. Magnetic field characteristics of the nucleus are defined by the magnetic moment. Positively charged hydrogen ion, simply referred as proton, is mainly used as a target element in MRI. It is the most common element inside the human body, due to water, and it has the largest magnetic moment. When the proton spins, which is simply referred as "spin", it acts like a tiny magnet with north and south poles (Beutel et al., 2000).

When the protons are kept under the influence of the strong external magnetic field generated by the scanner's powerful magnet, denoted as B_0, spins are either aligned with (parallel) the scanners magnetic field at a low-energy level or against (anti-parallel) the scanners magnetic field at a marginally higher energy level as it can be seen in Figure 2.9. However, at this state spins cannot be polarized statically and they wobble around the axis of the magnetic field. This is referred as precession and the frequency of this precession is given by the Larmor equation (Hashemi et al., 2004):

$$\omega_0 = \gamma B_0 \tag{2.3}$$

where ω_0 is the Larmor frequency (angular frequency of the precession) and γ is the gyromagnetic ratio. Every different atom has different gyromagnetic ratio and the gyromagnetic ratio of hydrogen atom is used in MRI imaging.

2.3.2 MRI image production

In MRI, to produce images a RF signal, generated by the RF coils of the scanner, with a frequency same as the Larmor frequency is focused on the targeted body part of the patient. Since the RF signal has the same frequency with the precession frequency of the hydrogen atom, it excites the hydrogen atoms in the targeted body area and causes them to resonate. This resonance effect causes spinning protons to gain energy, shifting low energy state protons in to the higher energy anti-parallel state. Additionally, precession of each individual proton twists by a certain angle with the implementation of the RF signal, bringing the collective precession of the protons in phase with each other. When the MRI system cuts off the RF delivery, energized protons undergo longitudinal and transverse relaxation returning back to the original state of low energy and out of phase precession. This loss of energy is released back as RF signal during the relaxation period, which in turn detected by the RF coils of the system and used to reconstruct the image (Beutel et al., 2000).

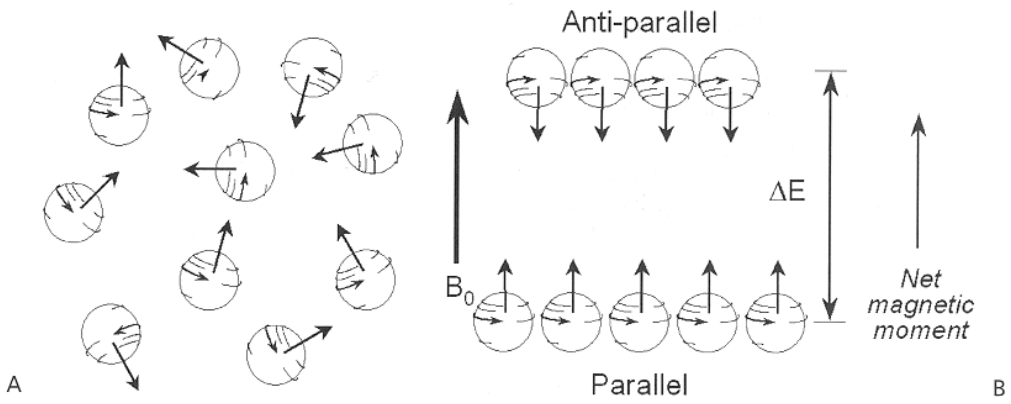


Figure 2.9: a) Protons spin naturally in the tissue. b) When the tissue is under the influence of the powerful magnetic field, spins align themselves as parallel or anti-parallel to the main magnetic field (Beutel et al., 2000).

Gradient coils of the scanner are used at this stage to encode spatial information into the RF signals, which enables 3D image reconstruction.

The final appearance of the MRI image is adjusted by controlling the timing of the delivered RF pulses to the patient, referred as the repetition time (TR) and timing of the echo signals released back from the body, referred as the echo time (TE). Two different relaxation times, T1 and T2, are measured and used along with the proton density of the corresponding tissue to decide about the final intensity of a voxel on the MRI image. T1 is the longitudinal relaxation time which is governed by the characteristics of the tissue regarding spin interaction with the lattice (molecular arrangement and structure of the tissue). It measures the time it takes the proton spins to lose externally RF induced energy and return back to the original direction of precession (which is the direction of the main magnetic field of the scanner). On the other hand, T2 is the transverse relaxation time which is governed by the characteristics of the tissue regarding spin-spin interactions which cause loss of phase alignment due to magnetic properties of the tissue. It measures the time it takes for the protons to lose their phase integrity, gained after the applied RF energy, and return back to their original out of phase precession. Both T1 and T2 reflect natural tissue characteristics (differ greatly from one tissue type to another) and are fixed for a specific tissue under a given magnetic field strength enabling the MRI scanner to distinguish easily between different types of tissues, granting excellent soft tissue contrast to the scanner (Beutel et al., 2000). Along with this, spatial resolution of the MRI is also superior when compared to x-ray based modalities.

Final MRI image comes in the form of different modalities (in other words weighted image types). "Weighting" of the image means that the contrast of the final image is heavily affected by either one of the, T1, T2 or proton density (PD) measurements. These T1-weighted, T2-weighted and PD weighted images can be obtained by the scanner by implementing different RF pulse sequences. One of the frequently used sequences is the spin-echo sequence, where proton magnetic field vectors are shifted by 90 to 180 degrees. Signal intensity of a spin-echo sequence is approximately calculated as follows:

$$S = K.[H].\left(1 - e^{-TR/T1}\right).e^{-TE/T2} \qquad (2.4)$$

where [H] is the PD and K is a factor for scaling the calculation. It can be seen from the calculation that T1 effects on the image are related to TR and T2 effects are related to TE, where [H] effects are present all the time. Accordingly, when TR and TE are kept short, image becomes T1-weighted. In contrast, when TR and TE are kept long, image becomes T2 weighted. On the other hand, when TR is long and TE is short, image becomes PD weighted image (see Figure 2.10). This multi-modality acquisition capability of the MRI allows pathological structures or other tissues that cannot appear clearly on one of the modalities to be seen clearly on other modalities, by only performing one patient scan, equipping MRI scanner with great clinical utility (see Figure 2.11). As an example, tumors tend to appear as bright signal on T2 images where on T1 images they appear darker (Armstrong et al., 2010).

Excellent soft tissue contrast of the MRI enables high clinical utility in pathological diagnosis. Diagnosis of cancer, multiple sclerosis and hematoma are among the main applications. In cancer diagnosis, MRI is used as the primary imaging modality for brain tumor imaging. Although it is also used for lung cancer imaging, it is not used as common as CT (Armstrong et al., 2010). However when combined with PET it can provide excellent soft tissue localization for the high metabolic activity areas. Additionally, contrast agents, like

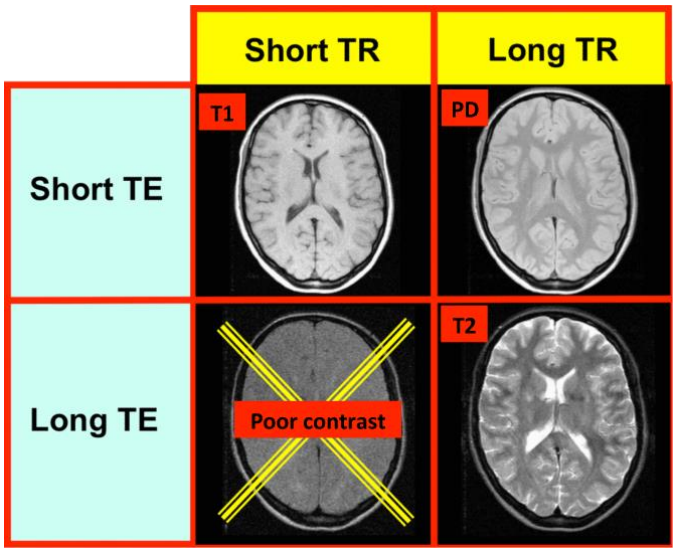
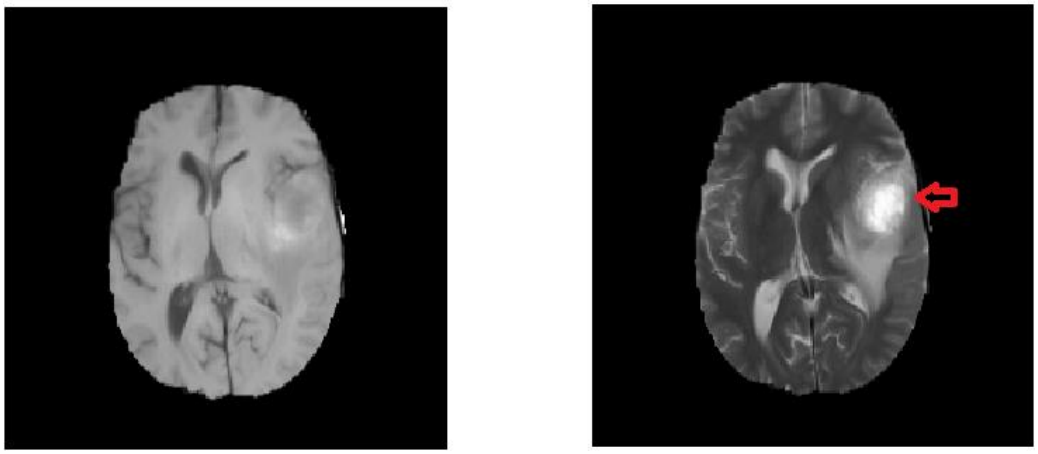


Figure 2.10: Shows the relations of TR and TE with different weighted MRI images (Beutel et al., 2000).



a) b)

Figure 2.11: a) T1 weighted brain image. b) T2 weighted brain image. Notice that the brain tumor (marked by arrow) and its surrounding edema tissue can be seen in more detail in T2 image, whereas T1 image does not provide clear details of the tumor (Işın et al., 2016).

gadolinium, that accumulate in metabolically active areas of tumors can be used to increase the MRI signal intensity received from those regions, enabling better imaging and thus enhanced diagnostic capabilities. For example, gadolinium enhanced T1-weighted imaging is used heavily for imaging brain tumors (Işın et al., 2016). Gadolinium contrast agent is accumulated in the active regions of the brain tumor, producing strong relaxation that appears as a bright signal on the T1-weighted images. This provides a higher level of diagnostic capability, providing information about not only the tumor but also its sub-compartments. If the contrast agent is not used, it is very hard for the radiologist to distinguish active and non-active sub-regions of the tumor core by just looking at plain non-contrast MRI images.

2.4 Positron Emission Tomography (PET)

In its basic definition, in PET imaging a biological radiotracer, which is a specific kind of radionuclide attached to a chemical compound, is delivered into the patient's body, where it emits a positron to annihilate with an electron in the tissue generating two back to back gamma rays that are detected by the PET scanner's detectors and processed into an image.

This process is illustrated in Figure 2.12. A specific radiotracer is selected for the desired application, so that it accumulates in the regions that the clinicians are interested in imaging. When the positron-electron annihilation process occurs at the desired anatomic region, the detected photon distribution by the detectors is reconstructed into the final image which corresponds to the radiotracer distribution inside the patient with the accumulated regions represented by high intensities (Saha, 2015).

There are many different radionuclides that can be used to emit positrons. Some of these common radionuclides are given in Table 2.1 with their half-life and energy properties. ^{18}F is the most common one and in clinical PET imaging it is generally combined with glucose based chemical compound, as introduced in Chapter 1, to form fluorodeoxyglucose or simply ^{18}F-FDG. Clinical application of the FDG PET imaging provides clinicians with a functional image that reflects the distribution of the glucose metabolism inside the targeted body part. Since cancer cells have higher metabolic activity and tend to absorb more glucose in that process than healthy cells, cancer cell regions generate higher intensity

Table 2.1: Some of the common radionuclides that can be used as positron emitters. Adapted from (Saha, 2015).

Radionuclide	Half-life	Energy of the Emitted Positron (MeV)
^{18}F	110 min	0.64
^{13}N	10 min	1.20
^{11}C	20.4 min	0.97
^{15}O	2 min	1.74
^{82}Rb	75 sec	3.55

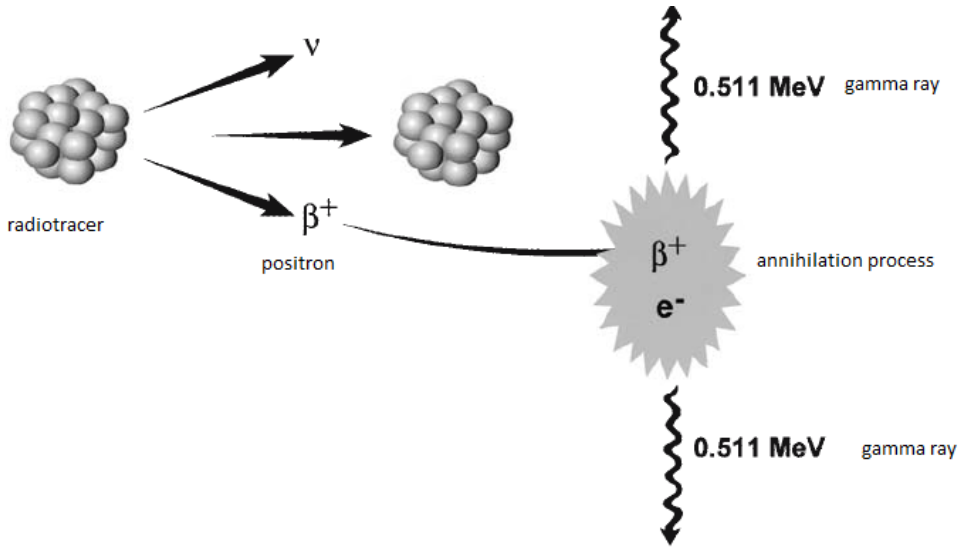


Figure 2.12: Illustration of the positron-electron annihilation process. Radiotracer emits a positron which annihilates with a nearby electron producing two back to back (180^0 opposite direction) gamma photons (Saha, 2015).

signals on the FDG-PET image providing very important information to the clinicians for cancer diagnosis.

The two back to back 511 KeV gamma photons that are generated after the annihilation process travels through the patient body and hit to a corresponding pair of detectors where recorded as a coincidence pair. The line of response (LOR) is recorded as a line between the pair of detectors where the coincidence pair is detected and the position of the occurred annihilation lies on this LOR. Total number of emitted positrons along the LOR is measured by counting the number of coincidence instances detected by the detector pair (see Figure 2.13). In the PET device, a detector ring is formed by connecting the individual detectors with each other and multiple detector rings are combined to form the whole scanner detector. The whole ring shaped PET scanner detector is used to measure all the incoming coincidence events from LORs of different angles over time (allowing dynamic imaging), then coincidence data is grouped into parallel projections for tomographic reconstruction before recording all LORs as sinograms or as list mode data. In the first

method, the number of coincidence events recorded by each detector pair are counted and presented as a histogram. In list mode method, each recorded coincidence event is accompanied by additional LOR and occurrence time data. After all these steps, one of the FBP, OSEM or MLEM reconstruction algorithms is used to form the final image of the metabolic distribution of the radiotracer.

When detecting incoming coincidence events, detectors of the PET scanner search for simultaneous (within 5 to 10 ns) gamma ray absorptions at the detectors. This time window is called the coincidence timing window. Four different types of coincidence events can be recorded (Bailey et al., 2005) as illustrated in Figure 2.14. First one is the true coincidence. In this type of event, both of the two gamma photons, generated by the annihilation, reach opposite detectors of the detector ring. Important point here is that there is no serious interaction between the photons and the surrounding tissue atoms and the detection is made within the coincidence window. Second type of event is called a scattered coincidence. This happens when one or both of the gamma photons from a single annihilation interacts with the surrounding tissue atoms and scatter, resulting in energy loss and direction change of the photon. Detecting such an event causes the LOR to shift from the actual position of the annihilation, resulting in decreased contrast and inaccurate localization in the final PET image. Third type is the random coincidence, which occurs when two different positron-electron annihilations happen almost simultaneously. If two of these photons from different annihilations are recorded within the coincidence timing window and the other two cannot reach the detectors, this event is counted as a valid event but it presents information which becomes spatially unrelated to the tracer distribution. There are also multiple coincidence events, as the final type, where similar to random events, in this type three events from the two different annihilations are detected within the coincidence timing window. In contrast to random events, these can be distinguished easily and discarded. Some times during a single event, one of the photons can be absorbed in the patient tissues or can escape from the detector. In these cases these events are considered as single lost events and discarded similarly. As a result, total recorded projection signal includes the true signal from the true events plus the noise signals from the scattered and random events. In order to obtain a clear signal, noise signals need to be estimated and removed.

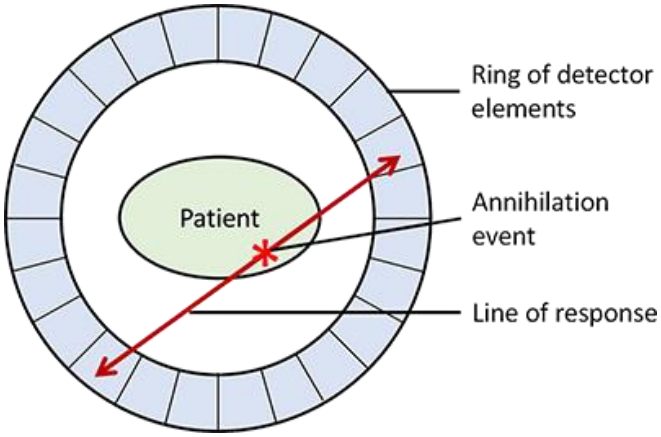


Figure 2.13: Two back to back gamma photons are detected by a pair of detectors of the scanner on a LOR as a coincidence event (https://www.radiologycafe.com/radiology-trainees/frcr-physics-notes/pet-imaging Retrieved 22 February, 2018)

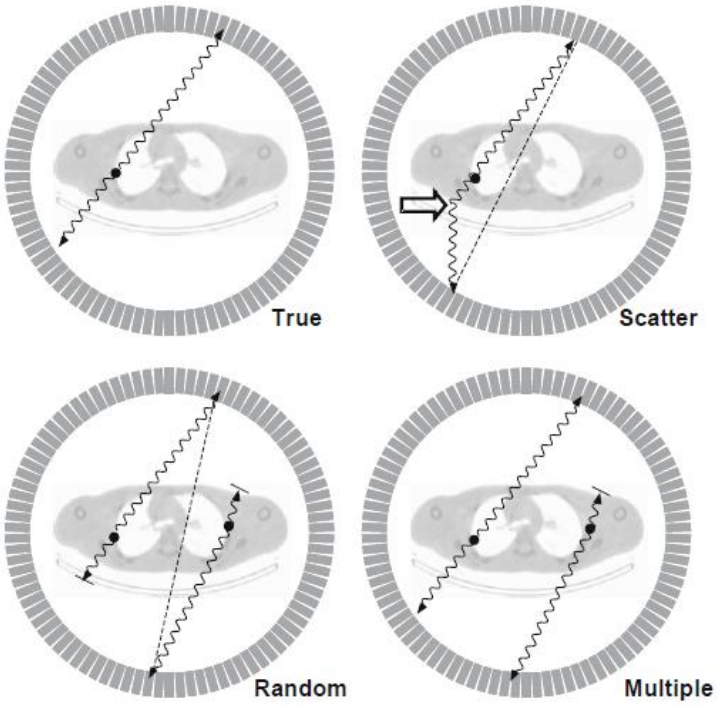


Figure 2.14: Four different types of coincidence events. The black circle is the place of annihilation. Dotted lines indicate the false assigned LORs in the case of scattered and random events (Bailey et al., 2005)

While random event estimation is usually accurate and efficient, scatter event estimation can have significant errors (Bailey et al., 2005).

Every PET scanner can acquire images in two different dimensions. In 2D acquisition mode, each individual detector ring is separated from each other by using collimation techniques. In this setup, gamma photon pairs that are traveling with large angles towards the opposite detectors that belong to different detector rings are blocked by lead or tungsten collimator leafs. Even though this technique can block some scattered and random coincidences it also blocks some true coincidences reducing the overall image quality. In contrast, in 3D acquisition mode there is no such collimation, increasing the number of detected true coincidence events, increasing the overall image quality. Else, clinician can also opt for a similar image quality like 2D acquisition, but with almost half the scan time (Strobel et al., 2007).

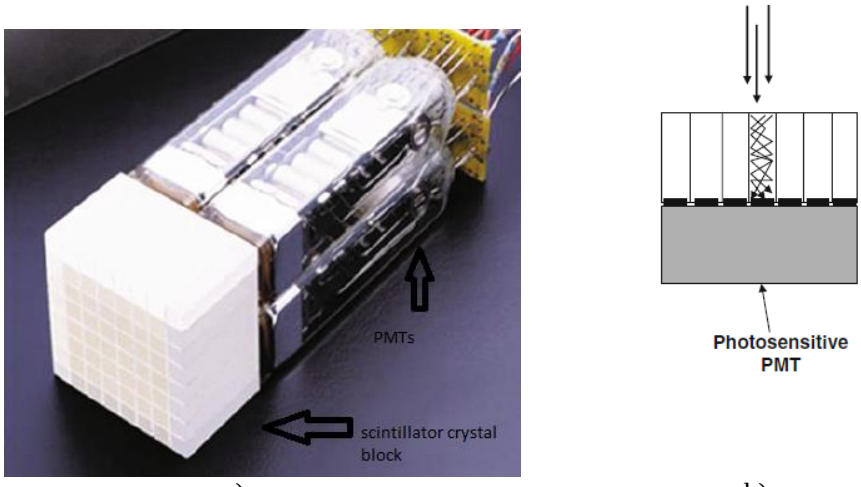
Apart from 2D/3D acquisition options, PET scanners can also obtain both static and dynamic images, even though for the imaging of lung cancers static acquisition is the preferred one. In static acquisition, a single image frame is taken over a defined time period after the distribution of the radiotracer concentration becomes far more static (approximately 20-40 mins after the injection). On the other hand, in dynamic imaging, series of image frames are acquired beginning just after the delivery of the radiotracer to the patient. Dynamic images normally tend to have inferior image quality when compared to static images but dynamic imaging can provide radioactivity distribution information over time which has uses in other clinical applications like studies involving neuropsychiatry (Gee, 2003).

It should also be taken into consideration that the gamma rays generated after the positron-electron annihilation can be also harmful for the patient if exposed over increased durations because of the ionizing properties of the gamma rays. However, diagnostic advantages of the PET device are greater so clinicians neglect the harms of a proper scan (one whole body PET scan can produce almost similar dose to a chest CT scan) for the positives of the accurate cancer diagnosis.

2.4.1 PET detector configurations, designs and materials

Earlier developed PET scanners incorporated simple gantry designs which used only two opposing detectors to detect the incoming gamma photons. Although, some later designs used partial ring designs with more detectors and even geometrically polygonal shaped designs, currently full-ring PET scanners are the standard for clinical applications. Since, PET devices do not require the imaging source to be generated by the device itself (in CT the x-ray tube and in MRI RF coils and magnets are necessary), main technology and cost of the equipment comes with the detector materials, technology and design. Due the amount of detector elements needed for producing full-ring scanners, cost of the full-ring devices are the highest. Implementation of more detector elements in full-ring system design comes with superior spatial resolution as a result of high event detection ability and coincidence count rate effectiveness when compared to other designs (Khalil, 2010).

Clinical PET scanners use photomultiplier tubes (PMT) coupled with scintillation crystals as individual detector elements (see Figure 2.15). There are also state-of-the-art semiconductor detector designs that use silicon photomultipliers (SiPM); although they allow realization of high gain with low voltage and fast response detectors with compact design, use of SiPMs instead of PMTs in producing full-ring scanners comes with optimization, signal amplification and digitalization problems and increases the cost of already very expensive PET scanner operation to a level that becomes not feasible for many institutions. Each detector block of the clinical scanners, implements a number of PMTs at the low layer to read the information generated by the previous layer made up from an array of scintillation crystals. Incoming gamma photon to the detector passes through the scintillation crystal where it is converted into light, a process called scintillation. There is optical isolation between each scintillation crystal by applying reflective material between each array element, which prevents the passage of the generated light signals from one crystal element to another, dramatically increasing detection performance of the following PMT layer. The generated light signal is then detected by the photocathode of the PMT and converted into photo-electrons which further multiplied by the dynodes of the PMT until they reach the anode part and recorded as an electrical signal.



a) b)

Figure 2.15: a) PET detector block. A block of scintillator crystal array with discrete elements is followed by four PMTs. b) illustration of the scintillation process in discrete element design (Khalil, 2010).

Crystals with high mass density, which enhances crystal stopping power for better radiation detection, are preferred as scintillator material. Additionally, light output and speed properties of the crystal are also important. Increased light output of the crystal means that the noise generated in the scintillation process is decreased and a fast crystal allows for a faster radiation detection enabling the system to use shorter coincidence timing period. Although earlier designs adopted thallium-doped sodium iodide (NaI(Tl)) crystals for scintillator layers, bismuth germanate (BGO), gadolinium oxyorthosilicate (GSO) and cerium-doped lutetium oxyorthosilicate (LSO) crystals provide better stopping power due to higher mass densities and they (especially LSO) became the preferred choice for the recent clinical PET scanners (Khalil, 2010).

However use of crystal scintillator elements in detector blocks also comes along with some intrinsic limitations. Large parallax error is one of the main limitations. Due to the natural characteristics and the minimally achievable thickness of each individual scintillation crystal, the exact entry point of the incoming photon to the crystal element cannot be distinguished over the single crystal element volume and thus depth of interaction (DOI) information becomes insufficient. Detector assumes that the photon is entering from the midpoint of the individual crystal element surface and assigns the LOR from that point to

the midpoint of the corresponding opposite detector that received the other photon of the back to back pair. This results in a clear difference between the assigned and actual LORs, creating the parallax error and causing the detector to miscalculate the exact position of the annihilation. High parallax error causes the PET scanner to have low cross-plane or horizontal resolution. Low signal purity due to high noise from scattered events and limitations in sensitivity of the scanner due to the thickness of each crystal element and the gaps formed between the detector blocks are the other main limitations. Designing organ specific PET devices or using pixelated semiconductor detectors can be a possible solution for these limitations due to the use of scintillator crystals.

2.4.2 PET scanner performance characteristics

Overall performance of a PET scanner can be measured by analyzing several characteristic parameters. Improvement of these parameters can have dramatic effects on the PET scanner performance and on the overall quality of the final diagnostic image.

System Sensitivity

Sensitivity of the PET system is a very important parameter that directly affects the overall image quality and noise percentage (signal-to-noise ratio) of a PET scan with a specific radioactivity distribution over a defined acquisition time. A scanner with high sensitivity can collect more information (event data) in a shorter time. In other words, have better detection efficiency. High sensitivity means, better signal-to-noise ratio, better counting data and better spatial resolution thus improved overall image quality (Khalil, 2010). Geometric and intrinsic properties of the scanner have direct influence on the scanner`s sensitivity. Geometric design of the detector is very important in maximizing the sensitivity of the scanner. Since sensitivity is directly related to the amount of photons detected by the detectors in a given time period, the less photons escape from the detector elements the higher becomes the sensitivity. To assure this, geometric design of the scanner should be implemented in such a way that all detector elements are tightly packed together, so that the angular coverage of the detector surface area becomes large enough and fewer photons can escape from the detector ring without getting absorbed in a detector element. This geometric sensitivity characteristic of the commercial PET scanners is

quantified by the ring-packing fraction. It is simply the ratio between the detector rings true detection area and the total circumferential detector ring area. Two different approaches can be implemented to increase the geometric sensitivity. First approach involves the implementation of narrower detector ring diameter and the second approach involves increasing the axial field coverage of the scanner. When the diameter of the detector ring is decreased, solid angle of the detectors increase, allowing the detectors to detect incoming gamma photons more efficiently. However, this implementation also introduces parallax DOI errors reducing the systems spatial resolution. Alternatively, axial field coverage of the scanner can be increased to increase the sensitivity by adding extra detector ring layers in axial direction. Even though this increases volume sensitivity by a significant margin (since more photons will be absorbed in the axial direction), adding extra detector material to the scanner increases the cost of the system dramatically (Khalil, 2010).

Different from the geometric sensitivity properties, intrinsic sensitivity of the scanner is defined by the type, composition and thickness of the scintillation crystal material used in individual detector elements. A scintillation crystal with high stopping power can stop most of the incoming gamma photons, thus providing efficient scintillation, which in turn increases detection efficiency. Therefore, as mentioned in previous section, density and effective atomic number of the preferred crystal material plays an important role in intrinsic sensitivity. Additionally, increasing the thickness of the scintillation crystal again improves the intrinsic sensitivity, however using thick crystals in detector elements also increases the parallax DOI errors. Apart from intrinsic and geometric factors, energy and time window properties also affect the overall scanner sensitivity (Khalil, 2010).

Additionally, choice of 2D or 3D acquisition modes for imaging also have an impact on system sensitivity, as mentioned in Section 2.4. In 3D acquisition since there is no collimation between detector planes, there is no limitation for the incoming photons to reach the detector plane on their direction. This enables almost five times increase in sensitivity with regards to 2D acquisition. 3D imaging also enables rapid scanning decreasing the potential motion artifacts. Tradeoff here is the increased introduction of signal noise from random events, scatter events, single events coming from the outside of the covered field and high count rates. Because of this, efficient scatter and random

correction techniques along with detectors that can provide high count rate performance should be implemented to enable efficient 3D acquisition with high sensitivity (Khalil, 2010).

Noise-equivalent count rate (NECR)

Another performance characteristic of a PET scanner is its count rate response. To obtain better quality PET images increasing the counting rates of the scanner is preferable. Although, increasing injected radioactivity dose and increasing patient scanning times can increase counting rates in turn, increasing the former also increases the exposed radiation dose of the patient which is not desired regarding safety and protection measures and increasing the latter also increases the motion artifacts and decreases the patient comfort. In addition, issues regarding to the dramatic increase in the random and scatter event rates would also arise, countering the gained image quality advantages through increasing count rates. In this regard, noise-equivalent count rate (NECR), that also takes all the above mentioned factors into consideration, is used instead to measure the count rate performance of the PET scanners (Khalil, 2010). As a result, the noise-equivalent count value that provides the highest true event counts and lowest undesired events, like random and scatter events, is considered as the final performance measure.

Coincidence timing window

Different from other nuclear imaging devices, where collimators are used to decide whether to accept an incoming photon or not, in PET imaging an electronically controlled timing window is used to decide whether an event is true and accepted or undesired and rejected. By implementing a narrow timing window, undesired contribution of random events is diminished, increasing the NECR thus the performance of the scanner. Use of fast scintillator crystal materials enabled the implementation of detectors with short coincidence timing windows, even leading to the development of advanced time-of flight (TOF) PET scanners (Khalil, 2010). TOF scanners use the information regarding the time difference measurement of the arrival of two back to back gamma photons to corresponding detector elements to determine the exact location of positron-electron annihilation. Using a scanner with very short coincidence timing window is key for precise localization using TOF.

Spatial resolution

Spatial resolution is one of the most important performance characteristics of the PET scanners that have great effect on the overall image quality; hence improving spatial resolution is of utmost importance. Having high spatial resolution provides better functional imaging of small lesions with high metabolic activity, more accurate quantitative measurements and better overall diagnostic capabilities. Detector properties, photon acollinerarity and positron range are among the main factors affecting spatial resolution in PET imaging (Khalil, 2010).

Generally, spatial resolution performance of the PET device is determined by the full width half maximum (FWHM) of the point spread function. FWHM is defined by:

$$FWHM \approx \sqrt{(d/2)^2 + b^2 + (0.0022D) + r^2} \qquad (2.5)$$

where *d* represents the width of the detector, *b* represents the secondary parameters that contribute to the loss in spatial resolution due to either photon detection process or block detector effect. Acollinearity is defined by *(0.0022D)* where *D* is the diameter of the scanner detector and finally parameter r^2 corresponds to blurring effects due to positron range.

Detector size or width *d,* is a key parameter that affects the spatial resolution. FWHM of an annihilation that is located at the mid distance between the two corresponding detectors detecting the incoming photons is equal to *d/2*. When annihilation source moves towards either one of the detectors, spatial resolution decreases. Implementing small-width scintillator crystal arrays in the detectors of the PET scanner is the key to improve spatial resolution. However it is not easy to produce small sized crystal arrays and small sized crystals can also limit the amount of light generated by the crystals for the PMTs to detect. Additionally, the production cost can increase dramatically. In this regard using scintillator crystals with improved light output and state of the art crystal cutting techniques is

necessary to manufacture small sized crystals which enable small sized detector production thus improving scanner spatial resolution. Commercial PET/CT scanners used in hospitals generally implement detector crystal sizes of 4 to 6 mm. These sizes can decrease more in dedicated and small animal scanners.

Another important factor that affects spatial resolution is the positron range. This is an unpreventable physical occurrence and directly depends on the properties of the positron source used in the radiotracer. This effect causes blurring on the reconstructed PET images. Most common clinical positron source, ^{18}F, luckily has very limited positron range effect on overall spatial resolution of the reconstructed PET images. Using higher resolution detectors by implementing small sized crystals is one measure. Taking this effect into consideration during reconstruction algorithms and even using the high magnetic field of the MRI scanner in PET/MRI scanners are among the other measures that can be considered to reduce the effect of positron range on spatial resolution (Khalil, 2010).

As mentioned earlier, acollinearity of the scanner is determined by the detector ring diameter, and its effects increase as the diameter of the scanner increases. Using narrow detector diameter can improve scanner resolution in terms of the adverse effects of acollinearity. Additionally, reducing ring diameter will also increase system sensitivity; however DOI errors will also be introduced as mentioned before. Similar to positron range effect, acollinearity effect can be taken into consideration during iterative reconstruction algorithms and its effects can be reduced by applying suitable corrections.

Parallax error, resulting from DOI errors especially in scanners with thicker crystals, also negatively affects the spatial resolution of the PET scanner by introducing blurring effects on the images. Different DOI and parallax error correction methods can be implemented to improve the spatial resolution. For example, double or more layers of scintillator crystals of different materials which coupled with two photo detectors can be used to decode depth information of the incoming photons more efficiently. This implementation reduces the effects of parallax error increasing the spatial resolution. Correcting parallax error effectively with these methods can also enable using thicker scintillator crystals which in turn improves sensitivity of the scanner.

2.4.3 Data corrections in PET

For the PET scanner images to represent functional information regarding the metabolic activities of the patient`s tissues, the coincidence data that the scanner collects contain several quantitative and qualitative properties. Quantitative properties of the collected data present numerical values that clinicians use to reach objective straightforward results regarding the data. On the other hand qualitative data that the scanner collects is represented by the radioactivity distribution images which the clinician needs to interpret for accurate diagnosis. Scanner calibrations and several data correction techniques should be applied to this collected data to ensure that the final image produced from the data represents the true radioactivity distribution inside the patient`s tissues. Normalization, dead time correction, attenuation correction, scatter and random correction are the main techniques that are used in PET data corrections.

Normalization

Unfortunately, in PET scanners sensitivities of each individual detector elements are not homogeneous. This is partly because it is not possible to assemble a scanner with detector elements that have exact same solid angles and detector pair distances. Also scintillation crystals used in every single detector element cannot have the exact same efficiency due to manufacturing reasons. Additionally electronic drifts in the PMT circuitry also add up to this non-homogeneity. This non-homogeneous sensitivity profile of the scanner results in having variations in coincidence event detection sensitivities for different LORs. If these variations are not corrected effectively artifacts, poor uniformity and increased noise is observed in the images. In this regard, in order to efficiently measure LORs with minimum geometric and electronic adverse effects, a normalization procedure must be performed to deal with the non-homogeneous detector sensitivity.

One early method for normalization requires collecting many count data for each LOR to produce a statistically accurate normalization correction factor for each detector pair with respect to the averaged acquired count data across all LORs. Since it is required to have extended acquisition times to ensure statistical accuracy of the normalization factor and biased results may be observed if the source has no uniform activity distribution, this method is not an optimal normalization solution.

Alternatively component-based normalization technique is implemented in modern clinical PET scanners. This technique accounts for both system geometry and efficiency of each individual detector pair to deal with normalization. Normalization correction factors are handled by dividing them into detector efficiency and spatial resolution components (Khalil, 2010). When compared to the previous method, this technique requires fewer counts to be acquired hence reducing the total acquisition time and can be used in normalizing 3D acquisitions. Other factors including intrinsic crystal efficiency, detector geometric profile, block detector interference, time alignment factor and count-rate-dependent block profile are added to this technique with modifications over time to improve the overall normalization performance (Khalil, 2010). Generally, with this technique normalization factors for a specific clinical PET scanner are determined in factory stage and remain constant during clinical use.

Dead time correction

Dead time is the time period in which a detector is dealing with one event where it cannot handle any more successive events reaching to the detector during that period. At high activity concentrations, probability of the source emitting simultaneous or consecutive (very close in time) photons is very high. Because of the dead time of the detector, signals from these simultaneous or consecutive photons pileup so that the output signal of the detector represents sum of all these signals rather than individual events. As a result, events with greater amplitudes than the upper energy threshold of the system or signals whose amplitudes are in the energy window can be detected. While the first type of events can be rejected by the system, second type is generally accepted but with false positron and energy determination (Khalil, 2010). Therefore at high activity levels, count rate performance of the scanner is downgraded by count losses and signal pileups. If dead time correction is not implemented, spatial resolution, signal-to-noise ratio and quantitative accuracy of the system decreases because of these effects. In modern clinical PET scanners fast scintillators coupled with front-end electronics with very fast signal processing and transferring capabilities are used in order to process larger amounts of data in shorter times to improve the digital time resolution of the system thus reducing the effects induced by the dead time of the detector. Additionally, software based methods can also be

implemented by mathematically modeling the count rate response of the PET scanner and using that model to correct count losses observed on the actual measured count rates.

Attenuation correction

When photons generated by the positron-electron annihilation travel inside the patient, they interact with the patient`s tissues along the path losing their energy and even losing the photon itself. This effect is called attenuation. Because photons that are generated inside the patient`s body need to travel through more tissue material until they reach the detectors than the photons generated outside, the radioactivity distribution inside the patient is underestimated if the attenuation effect is not corrected. Attenuation along a LOR can be calculated by the probability of a photon pair along a LOR to travel through the patient and reach both corresponding detectors, and it is given by;

$$P_l = e^{-\int_{l1+l2} \mu(x)dx}$$

$$= e^{-\int_l \mu(x)dx}, \qquad (2.6)$$

where *l1* and *l2* are the paths of the two photons, *l* is the LOR and $\mu(x)$ is the attenuation coefficient at *x*. This attenuation calculation is used as the foundation of the attenuation correction methods, and the important point of this calculation is that the attenuation is not dependent on the location along the LOR of the annihilation. The two main attenuation correction methods are the measured and the calculated attenuation correction methods.

In measured attenuation correction method, an attenuation map is generated by directly measuring attenuations of a source placed outside the patient and taking two scans first without the patient (blank scan) and second with the patient (transmission scan). As mentioned earlier since the attenuation is not dependent on the location along the LOR, placing the source outside or inside the patient does not affect the attenuation. After the scans are performed, by taking the ratio of the count rate measurements of the scan with the patient to that of without the patient, attenuation for each LOR is determined.

Attenuation correction factors (ACF) are then calculated as the conjugates of the attenuation. In order for this method to be statistically efficient many counts should be collected for each LOR and performing two separate scans requires extra time.

In calculated attenuation correction method, attenuation map that is used to make corrections on the PET data is calculated using the Equation 2.7 by assuming that the attenuation coefficients along with the shape and the structure of the patient are known. The shape and the structure of the patient is obtained either by segmenting emission/transmission scan or by using the high resolution structural image of the CT scan if hybrid PET/CT scanner is being used. Statistical noise is not a problem in this method, but efficiency of the segmentation for shape and structure determination and the attenuation coefficient assumption is very important for obtaining high performance. Since, very high spatial resolution of anatomical CT images provide excellent patient shape and structure determination, and since CT images can also be used for accurate attenuation coefficient assumption (attenuation coefficients obtained on CT, which are at a different energy level, can be transposed for PET energy level), ACFs can be calculated with high accuracy making the CT-based calculated attenuation correction the accepted method for attenuation correction in PET imaging.

Scatter correction

As mentioned earlier, scattered coincidences are among the undesired effects that corrupt the total count rate of the PET scanner. Although, scattered events can be distinguished from true events by their energy levels, due to PET detectors having limited energy resolution and the fact that some true events release only a portion of their energy in detectors, leveling them with scattered events in terms of energy, this is not an easy task for the PET scanners. In this regard, several methods that are based on different techniques are developed to correct for the scattered events.

One of these methods is dual-energy window method. In this method two different energy level windows are used to distinguish between photons with peak energy and scattered events. In both of the windows, the detected events are considered to contain both true and scattered events but it is assumed that the window with the lower energy band contains mainly scattered events. Difference between these two windows are obtained later and

scaled using phantom data to approximate the distribution of scattered events in the total patient data obtained. After the scatter event distribution is obtained it is subtracted from the peak photon window data to finalize the correction (Khalil, 2010). A slightly modified version of the dual-energy window method uses two energy windows (one standard level and one high energy level) that overlap with each other instead using two distinct ones.

In another type of method, it is considered that the distribution of the scattered events differs gradually across the field of view of the scanner. In this analytic method, a Gaussian fitting of the scattered event counts outside the patient is carried out to approximate the scattered event distribution, providing fast and smooth correction method.

Using a Monte Carlo simulation to simulate the scattered event distribution is another common approach (Barret et al., 2005). In this approach, image is initially reconstructed and the attenuation map is determined so that the scattered event distribution can be simulated based on the reconstruction. Monte Carlo simulation is an efficient method since it accounts for the radioactivity distribution, the whole course of photons from their emission to the scattering interaction with the detector or their escape from the gantry, other physical interactions and detector characteristics along with the attenuation properties. Thus Monte Carlo simulation method is considered as a very accurate method, and standard for evaluation scatter correction techniques and a standard scattered event correction method for commercial clinical PET scanners (Khalil, 2010).

Random correction

Similar to scattered events, random coincidence events also adds up as a noise to the true event signal recorded by the detectors. Thus methods are also developed to deal with and correct for the random events.

In one approach, a coincidence window with a delayed timing is implemented to measure random events directly. When this delayed window is used, coincidence events detected by a detector pair only contains random events. The probability distribution of these random events when a delayed window is used is same with the probability distribution when a usual coincidence window is used. Thus by taking the difference of the whole PET data

and the measured random event data the random correction is carried out. This is the main method used in most of the available clinical PET scanners.

In an alternative method, singles count rate for each detector is measured and the random event count rate of a detector pair is calculated using the formulae (Dale et al., 2005);

$$N_r = 2\tau N_1 N_2 \tag{2.7}$$

where, N_r is the random event count rate, 2_T is the coincidence time window, N_1 and N_2 are the singles count rates for the two corresponding detectors. This method provides statistically efficient determination of random event count rate since the singles count rates are much higher than the random count rates, but coincidence timing window for each detector pair should also be known exactly for accurate calculation.

2.4.4 PET image reconstruction

In PET image reconstruction, analytical or iterative methods are implemented to generate an image of the radioactivity distribution in the patient's body from the measured coincidence events by the scanner.

Analytical reconstruction

Analytical reconstruction methods are based on the principles of computed tomography. Two main approaches of backprojection and filtered backprojection can be applied for analytical reconstruction.

In PET imaging, a projection $p(s, \phi)$ is obtained by applying line integration of the radioactivity distribution along all parallel LORs (where s represents a specific LOR) at a specific projection angle ϕ. All the projections obtained from different angles are then organized into a sinogram. Each single projection obtained by the scanner fills a single row of a specific ϕ in the sinogram. Additionally a single point in the two-dimensional patient space $f(x, y)$ is represented by a sinusoid in the sinogram space (see Figure 2.16). Such transformation of the $f(x, y)$ into $p(s, \phi)$ at different angles in two dimensions is called

"*Radon transform"* and the "*inverse Radon transform*" can be calculated by applying backprojection techniques for the reconstruction of the image *f(x, y)* from *p(s, ϕ)*(Kak and Slaney, 1988). 3D image of *f(x, y, z)* can be further obtained by repeating 2D image acquisition for multiple axial slices *(z)*, then obtaining sinograms from *f(x, y)* for each value of *z*, later reconstructing those sinograms into 2D image slices and stacking them on the axial dimension to form the 3D image. However this method of forming 3D images is carried out by using 2D acquisition. In "fully" 3D acquisition, instead of stacking 2D images only from "direct" planes to form the 3D image as in 2D acquisition, additional information regarding the line-integrals on "oblique" planes are also acquired and directly used in image reconstruction (see Figure 2.17) (Dale et al., 2005).

In simple bakprojection reconstruction, a value of *p(s, ϕ)* is projected back into an image matrix along its corresponding LOR (see Figure 2.18). Since exact location of the value

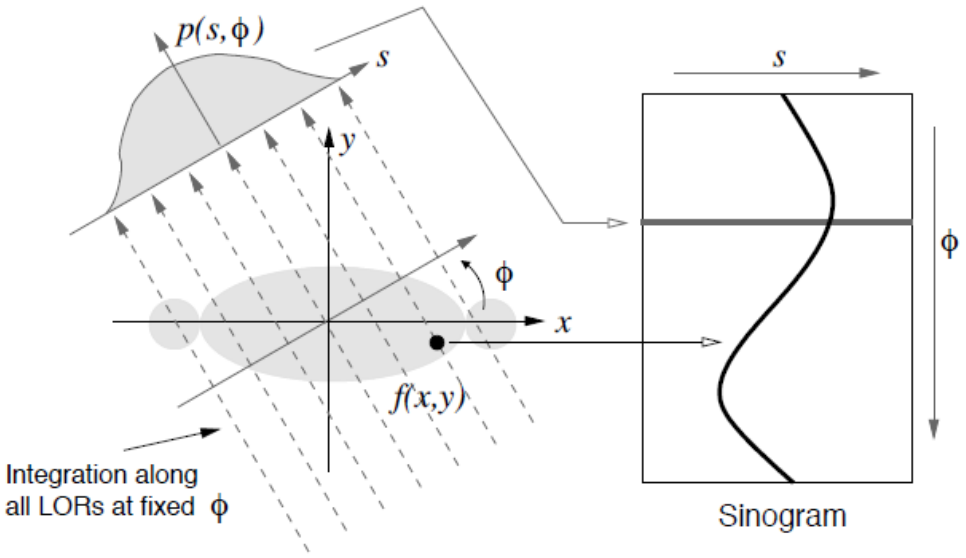


Figure 2.16: Formation of a sinogram in PET imaging (Dale et al., 2005)

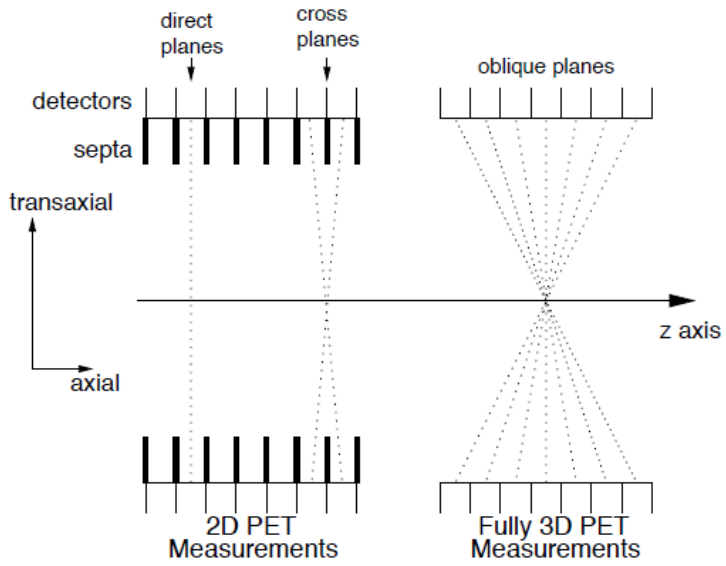


Figure 2.17: Difference between 2D and 3D PET acquisition for image reconstruction (Dale et al., 2005)

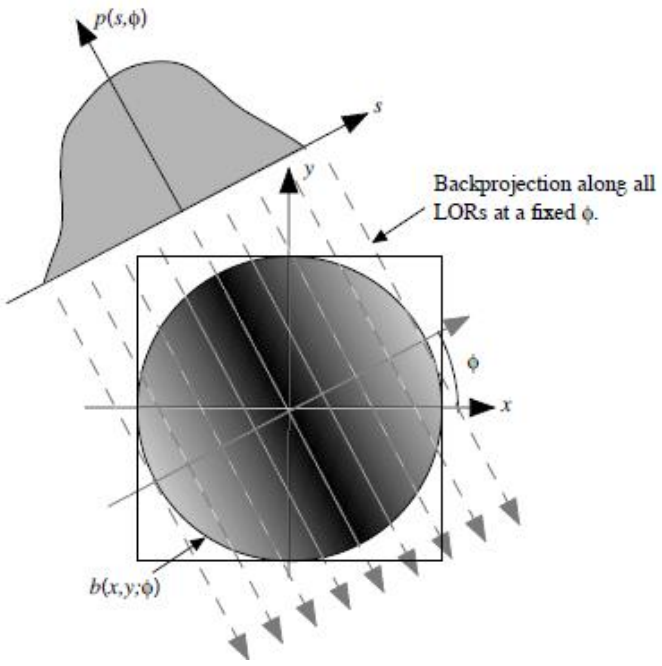


Figure 2.18: Backprojection of all the values of $p(s, \phi)$ into image matrix for a specific angle ϕ, forms the backprojection $b(x, y; \phi)$. If the procedure is repeated for all ϕ a complete 2D image slice is reconstructed (Dale et al., 2005)

along the LOR is lost during projection step, a constant value is projected back into all cells of the image matrix along that LOR. However due to this, when projections from all angles are backprojected an oversampling problem arises which causes blurs in the final reconstructed image slice (Khalil, 2010). To overcome this oversampling problem, a filtered back projection (FBP) method can be applied instead of using plain backprojection. In FBP, one dimensional Fourier transform of each obtained projections at every different projection angles are computed. Then a ramp filter is used to filter each Fourier transformed projection for eliminating the oversampling problem. As the final step, inverse Fourier transform of those filtered projections are computed before implementing the traditional backprojection to form the final image.

Reconstruction using analytical methods provide very rapid image formation and the oversampling problem intrinsic to backprojection can be overcome using FBP. However, these methods does not take into consideration that a statistical variance, which causes apparent noise in the images, is present in the data collection procedure due to the fact that each detectors measured event counts are not exactly equal to the integral. Although frequency based filtering is used to eliminate the resulting noise, a degradation in the spatial resolution of the image also results due to the filtering.

Iterative reconstruction

Second type of the image reconstruction methods are iterative methods. When compared to analytical methods, in iterative methods statistical variance and the noise structure of the measurements including scattered and random coincidence events are taken into consideration to a greater extend. Due to this nature, iterative methods can be considered as an improvement to the analytical reconstruction techniques. However increased computational cost due to the nature of iterative process and complexity of not having a direct analytical solution can be considered as the main disadvantages. In its simple explanation, iterative methods try to improve the estimate of an unknown image by repeating and updating the estimation until a satisfactory solution, i.e. most coherent with the measured PET data, is reached (Leahy and Qi, 2000). Generally, implementation of iterative methods models the photon detection process more efficiently and results in more accurate image reconstruction with improved image quality than analytical methods.

A common iterative reconstruction method used for PET image reconstruction is the Maximum Likelihood Expectation Maximization (MLEM) algorithm, where the reconstruction process is modeled as an optimization problem and solved using an expectation-maximization framework by maximizing a likelihood function (Dempster et al., 1977). In PET image reconstruction MLEM algorithm tries to solve the following iterative equation for the best image *f*;

$$f_j^{(n+1)} = \frac{f_j^{(n)}}{\sum_{i'} H_{i'j}} \sum_i H_{ij} \frac{p_i}{\sum_k H_{ik} f_k^{(n)}} \tag{2.8}$$

where $f_j^{(n+1)}$ is the next estimate of voxel *j*, $f_j^{(n)}$ is the current estimate of the voxel *j*, *H* is the system model/matrix (which relates the image to the data) and p_i is the measured intensity in the projection *i*. The whole iterative process works like follows; initial image estimation $f^{(0)}$, usually an image with a constant value in all voxels, is made as the first step. In next step, this image is forward projected to all projections ($\sum_k H_{ik} f_k^{(n)}$). Then these projected projections are compared with the actual measured projections ($\frac{p_i}{\sum_k H_{ik} f_k^{(n)}}$) to form a correction factor for every projection. Later these correction factors are backprojected into all voxels of the system matrix ($\sum_i H_{ij} \frac{p_i}{\sum_k H_{ik} f_k^{(n)}}$) to obtain an image correction factor. Finally, this image correction factor is multiplied with the initial image estimate and divided by a weighting factor based on the system matrix ($\frac{f_j^{(n)}}{\sum_{i'} H_{i'j}} \sum_i H_{ij} \frac{p_i}{\sum_k H_{ik} f_k^{(n)}}$). This final step updates the initial estimated image (Equation 2.8) and this new updated image estimate is reintroduced into the algorithm as the next image estimation and the previous steps are repeated. Algorithm continues to iterate until the difference between the two consecutive image estimations ($f^{(n+1)}$ and $f^{(n)}$) falls below a predetermined level, i.e. the estimated image converges to the maximum likelihood solution. Due to the nature of this process, every iteration requires one forward projection and one back projection. Additionally, convergence is very slow; depending on the image it can take up to fifty iterations for the MLEM algorithm to converge to an adequate solution. Together these factors increase the overall reconstruction time dramatically.

Ordered Subsets Expectation Maximization (OSEM) algorithm can be implemented in order to accelerate the convergence of the traditional MLEM, reducing the overall reconstruction time (Hudson and Larkin, 1994). In OSEM, the traditional MLEM formulation is somewhat altered as following;

$$f_j^{(n+1)} = \frac{f_j^{(n)}}{\sum_{i' \in S_b} H_{i'j}} \sum_{i \in S_b} H_{ij} \frac{p_i}{\sum_k H_{ik} f_k^{(n)}} \tag{2.9}$$

Here the whole projection data is divided into *B* subsets and the backprojection steps apply the summation over the projections belonging only to the subset S_b. In MLEM, summation is carried out over all projections and the image is updated only once per iteration. In contrast, OSEM updates the image estimation during each sub-iteration resulting in a total of *B* image updates per iteration. If *B* is set to 1, then OSEM will be same as the MLEM. Total number of projections processed per iteration determines the overall computational cost per iteration of an estimation maximization algorithm. Since this is same when applying both algorithms, OSEM and MLEM require similar computation time per iteration. On the other hand, while MLEM estimation converges once per iteration, an image estimation made by OSEM converges *B* times (with each one of them being similar to one MLEM iteration convergence) per iteration, accelerating the whole convergence process (it takes less iterations for the OSEM to converge efficiently), thus overall image reconstruction. Number of subsets, *B*, determines the overall acceleration factor of the convergence.

Additionally both MLEM and OSEM algorithms tend to produce a reconstruction with high variance, which raises the need for regularization. Especially using OSEM with increased number of subsets, thus obtaining higher convergence speed, increases this image variance (Lalush and Tsui, 2000). To solve this problem, a moderate subset number can be chosen and a Gaussian post-filtering can be applied to the OSEM reconstruction to rapidly obtain PET images with good quality.

2.5 Hybrid Systems: PET/CT and PET/MRI

As mentioned in previous sections, PET imaging provides clinicians with a functional image that represents the radiotracer distribution (which in turn represents the metabolism for the radiotracer in question) inside the patient`s body. However, for the clinicians to precisely locate the anatomical position of the high metabolic activities for better diagnosis, additional anatomical information is necessary.

One method to add anatomical information to PET imaging is by implementing software based fusion of PET images with anatomical images obtained using an another modality. In this type of implementation, PET images of the patient are registered using computer algorithms with the anatomical images obtained for the same patient with another modality at different times and at different conditions (different places etc.) (Camara et al., 2007). Since both acquisitions are carried out under different circumstances (patient movement and positioning is generally different between the two acquisitions) obtaining high registration accuracy is challenging.

In current clinical routine implementation of a PET as a single modality device is rare. Integration of PET device with a CT scanner (in the form of PET/CT) is the current standard for cancer diagnosis. However in the recent years, there is also an increase in the number of integrated PET/MRI devices used in clinical environment. These hybrid devices integrate two different modalities, one with functional and other with anatomical imaging capabilities, within a single imaging system in which both types of images are acquired successively (and in some cases simultaneously) and co-registered (hardware based image fusion) with higher accuracy when compared to the software based image fusion (Camara et al., 2007). Since patient undergoes the whole image acquisition in one combined scanner, final diagnosis is also achieved far more rapidly, saving time and increasing patient comfort.

2.5.1 PET/CT

PET/CT is a hybrid imaging modality where a PET device is in cascade with a multi-slice CT device inside the same gantry (Figure 2.19). With this integration, patient goes through first a spiral CT scan then patient table moves further inside the gantry bringing the desired

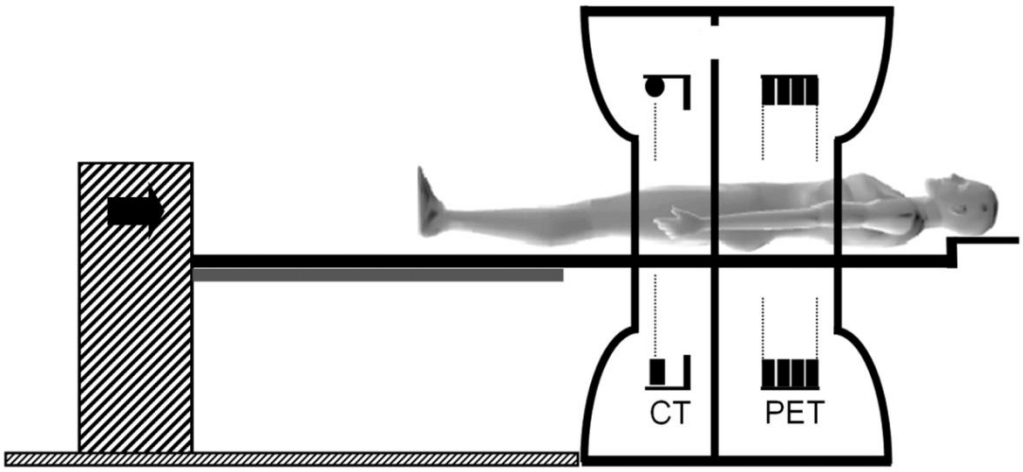


Figure 2.19: Diagram illustrating the PET/CT device. A multi-slice CT scanner is in cascade with a PET scanner in single, combined gantry setup (Beutel et al., 2000)

imaging region of the patient in line with the PET scanner detectors which carries out PET acquisition successively without the need for repositioning the patient. Although the field of views of the CT and PET are separated because of the cascade design (both CT and PET gantries cannot be overlapped), this can be solved by taking the patient table movement into consideration, which in turn allows accurate co-registration of the images without the need of a software based image fusion.

This hybrid design provides detailed anatomical information along with the functional image that allows clinicians to carry out accurate cancer diagnosis and localization. Additionally, as mentioned previously, attenuation correction is crucial for obtaining better quality PET images. In conventional PET scanners very long transmission scan is required to provide sufficient attenuation correction. However in PET/CT scanners, CT component eliminates the need for the transmission scan by providing a good quality attenuation map that is used for the attenuation correction thus reducing the overall scan time considerably.

Previous research proved that the use of hybrid PET/CT device performs better when compared to using PET or CT alone in cancer diagnosis including lung cancers (Antoch et al., 2003). One of the main challenges here is the PET image degradation caused by the

respiratory motion. In addition to blurring PET images, respiratory motion also causes problems in the co-registration process between PET and CT images and results in attenuation correction errors thus reducing overall diagnostic accuracy. In this regard, respiratory motion correction to improve the image quality is crucial for PET imaging.

2.5.2 PET/MRI

In this hybrid device, PET scanner is combined with an MRI scanner as a single complete setup. When compared to CT, MRI provides far better soft-tissue contrast that creates an advantage in the anatomical localization of tumors. Especially when imaging brain for cancer diagnosis and imaging late lung cancer patients for extra-thoracic metastases, PET/MRI provides superior performance. Another important advantage is that it eliminates the additional radiation dose delivered to the patient by the x-rays of the CT component. Combination of PET and MRI scanners is achieved in two different designs. First type of design integrates two scanners by closely fitting the PET scanner into the MRI device. This allows simultaneous acquisition of both types of images that are intrinsically matched spatially and temporally. In contrast, second type of design implements the combination by putting the two scanners in cascade but separated by a few meters distance to reduce the interference of the two scanners. Although implementation of this type of combination is technically easier, it does not provide the temporal matching of PET and MRI images as the other type of design does.

Despite, the clear advantages of PET/MRI in diagnostic capabilities, there are several technical challenges that increase the difficulty of the implementation and manufacturing, thus increasing the cost of the commercial scanners. Some commercial PET/MRI scanners can cost up to 3 times more than PET/CT scanners. One challenge is that in integrated PET/MRI design, strong magnetic field generated by the MRI scanner interferes with PET component by distorting the PMTs in the PET detectors. One measure to overcome this challenge is using SiPMs instead of PMTs in the PET detector design. SiPMs are not sensitive to MRI magnetic field but they are more expensive to implement. In cascade design this effect is avoided by the few meter distance between the two scanners and all PMTs are further individually magnetic shielded. On the other hand, PET scanner also should not cause any interference with the MRI scanner. RF coils, which are the main

signal receiving components of the MRI scanner, are manufactured differently than conventional MRI device RF coils, to avoid the interference with the electronics of the PET scanner. One another very important challenge is the attenuation correction. MRI image intensities depend on proton density and relaxation time properties of the tissues. Since there is no direct relationship between these properties and photon attenuation properties, deriving an attenuation map from the MRI scan is a difficult process. Computer algorithms that implement pattern recognition or image segmentation techniques on MRI images can be used in this regard.

2.6 Deep Learning for CAD Systems

In medicine, computer aided detection (CAD) systems are mainly developed to assist clinicians, by processing the final images obtained by the different modality scanners explained in the previous sections, in the diagnosis process by providing automatic detection of conditions that have high probability of being pathological. After these candidate conditions are detected automatically, clinicians can mainly focus their attention on those candidates, instead of the whole data, accelerating the diagnosis process, saving time and money for the medical institutions. At their current level, CAD systems should be considered as automatic detection systems that provide only assistance to the current clinical routine. Since accurate diagnosis of pathology, especially cancer, involves many different actors (radiology, pathology, nuclear medicine, surgery and others depending on the type of the disease) with high expertise, fully automating the whole diagnosis process by eliminating the need for expert clinicians is not possible at this stage.

Although there are also CAD systems developed for processing biomedical signals, main focus is on developing systems for processing biomedical images. In this regard, most of the popular image processing and traditional machine-learning techniques are applied on different medical image modality images for the purpose of automatic detection. However, as explained in Chapter 1, with the recent popularity of deep learning systems in all fields of computerized image recognition, research in medical image recognition also adopted these techniques as the current state-of-the-art. Implementation of these techniques for the cancer detection task, as medical CAD systems, obtained very high performances, which

proves to be promising for the complete integration of such systems into daily clinical routine (Işın et al., 2016).

2.6.1 Machine learning and traditional methods

Machine learning is a collection of computerized methods that are used to make predictions or determinations about a data on hand. In machine learning, computerized algorithms are used to decompose data, learn information from it and use these learnt information to make the required decisions about the data. There is no need to code software routines manually that relies on pre-determined set of instructions to carry out the decision task. Instead computers (or machines) are "trained", with special algorithms that have the ability to "learn" how to perform the task in hand, using large amounts of data. In the end "they" decide on the outcomes of the task depending on their learnt experience from the previously presented data. This learning process of machines is implemented by using mathematical techniques. When an input is given to them, machine learning algorithms try to predict the correct output mathematically. A simple illustration of a machine learning system can be seen in Figure 2.20.

There are two main types of learning process for the machine learning algorithms, namely supervised or unsupervised learning. During the supervised learning process, large amounts of input data with known outputs (ground truth/known classes/labels) are first given to the algorithm. Using these known outputs for the given input data, the mathematical model is then modified and iterated until the difference between the algorithm's predicted outputs and the actual outputs fall below a pre-determined error level. In contrast, in unsupervised learning methods, machine learning algorithm learns the relationships between the input data by itself and predicts the output class without the need for a ground truth data or any other human intervention. Table 2.2 shows some details about the differences between these two main learning algorithms.

One machine learning method that is frequently used for medical imaging is Artificial Neural Networks (ANNs). ANNs are developed by mimicking the neuron connections and organization inside the biological brain. These interconnected structure of neurons exchange data between each other by their connections and each connection have a

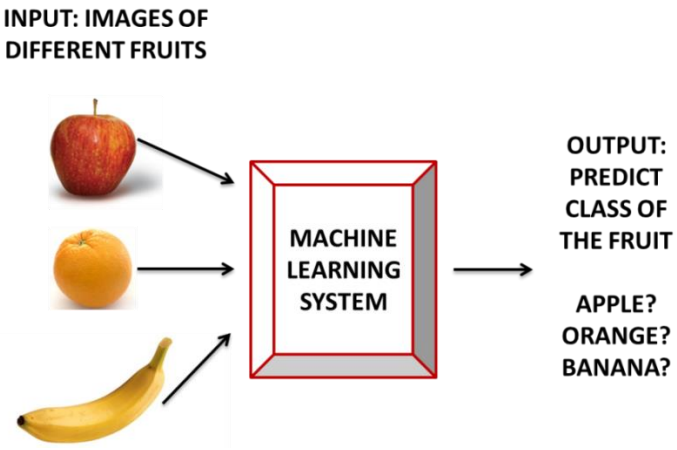


Figure 2.20: Illustration of a simple machine learning task. Machine learns, by training on large amounts of fruit image data, how to predict the class of a fruit when an input fruit image is given to it.

Table 2.2: Differences between supervised and unsupervised learning methods

Supervised Learning	**Unsupervised Learning**
Known number of output classes	Unknown number of classes
Based on a training set	No prior knowledge
Used to classify future observations	Used to understand data
Main methods: Artificial Neural Networks, Conditional Random Fields, Decision Tree Learning, Random Forests, Support Vector Machines...	Main methods: k-Means Clustering, Expectation-Maximization Algorithm, Hebbian Neural Networks, Auto Encoders…

numeric weight value which is modified while the network gains experience through iterative training. In this way ANNs adapt to the given inputs and learn the relationships between these inputs and the outputs. Since the knowledge about the input data is obtained

by the network through a learning process and connection weights are used for storing this knowledge, ANNs heavily resembles the biological brain. Generally interneuron connection weights are adjusted by comparing the predicted output and the desired output until the network predictions matches the desired target for the given inputs. In application, large amounts of such input data with known ground truths are required for effectively training ANNs by supervised learning.

Conventional ANNs normally contain one input (neuron) layer, one hidden layer (some may have more than one) and one output layer, with varying number of neurons at each layer (depending on the application) where each neuron acts like a single computational element (See Figure 2.21). All the connection weights between the neurons of these layers are updated iteratively during training process by using mathematical techniques such as the common back-propagation algorithm. Neurons on the input layer receives values directly from the data generally in the form of feature values extracted from the data using image processing techniques (data is assumed to contain medical images).

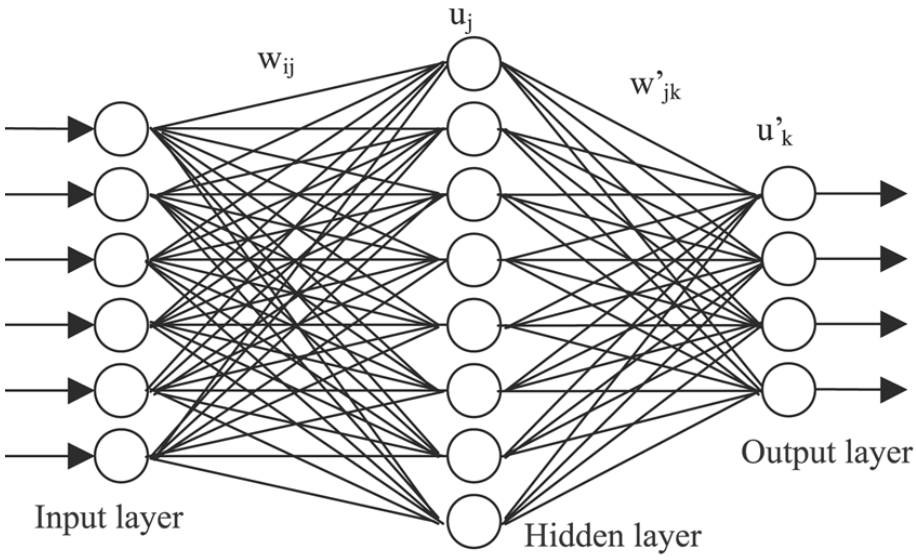


Figure 2.21: Illustration of a simple ANN structure. Connection weights between the input and hidden layer are given by W_{ij}. Where i indicates the neuron number in input layer and j indicates the neuron number in hidden layer. Similarly W'_{jk} represents the connection weights between the output layer and the hidden layer.

Input layer neurons are passive and they directly transfer input values into the network without any processing. Generally, the number of input neurons used in the input layer is equal to the number of different feature values extracted from the data. Later, every neuron in the successive hidden layer receives the output of the each input neuron. However output of the each input layer neuron is first multiplied with the connection weight between the receiving neuron of the next layer and itself. Neuron then sums all these incoming weighted input values from the neurons of the previous layer and then calculates a transfer function to produce its own output. There are several different transfer functions that can be used. One common one is the sigmoid function. Sigmoid function takes the summation of the incoming weighted inputs and transforms it into a value between zero and one. Similarly, every neuron on the output layer then receives the outputs of the each neuron of the previous hidden layer which are again multiplied with the respective connection weights, applies the summation and calculates the transfer function to generate its output. Number of the neurons in the hidden layer is generally chosen by experimentation and the number of the neurons in the output layer is generally chosen according to the number of the different classes that the ANN tries to assign the input data. Output values (estimations) received from the entire output layer neurons are then compared with the desired (actual) outputs, which are presented to the system by the user (in supervised learning), to calculate error values between them which are used in the update process of the weights for the next iteration of the learning process. Learning process continues to iterate and to update the connection weights until a desired error margin is reached. At that stage, network is assumed to have learnt the relationship between the input data and the desired outputs by storing the related knowledge in its connection weights. After the training is complete, an input data with an unknown output can be presented to the network for the network to carry out the classification using its previously learnt knowledge regarding the task in hand.

Since the main aim of the CAD part of this thesis does not involve developing a new machine learning method or new ANN architecture, but an application of a pre-trained machine learning system to the medical imaging task of lung lesion detection, further detailed explanations regarding the mathematics of the algorithms and further details of the training procedures will not be explained here. Reader may apply to other sources, if

necessary, for more detailed information about traditional machine learning methods, ANNs and training procedures.

2.6.2 Deep learning

Deep learning methods are basically machine learning techniques which have deeper architectures, meaning that they have many more layers (see Figure 2.22). Designing deeper architectures means that the required calculations during weight updating process will also be increased. Since many iterations are required until a satisfactory learning is reached, overall computational cost of the deep learning methods heavily increases. Due to this, only with the recent advancements in computational technologies, especially in GPU processing, designing and training deep learning methods became possible.

In traditional machine learning techniques, raw input images are first need to be processed with image processing techniques to extract meaningful features that have high intra-class similarity and low inter-class similarity so that the system can easily learn the differences between the output classes to effectively distinguish between them and obtain a

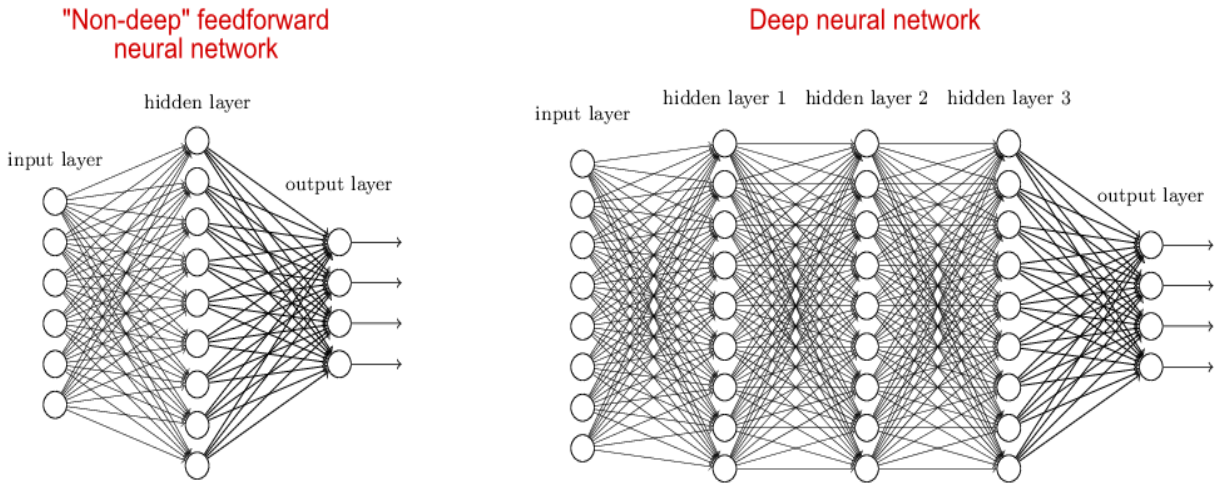


Figure 2.22: Diagram showing the difference between “non-deep” and “deep” neural network architectures. Deep learning architectures have many more hidden layers when compared to non-deep architectures (http://neuralnetworksanddeeplearning.com/chap5.html Retrieved 10 March, 2018)

satisfactory classification. This process of obtaining such features is referred as feature extraction. Performance of medical pathology detection systems that are based on traditional machine learning techniques heavily depends on the quality of these extracted features. Due to this, most of the research conducted on earlier years of CAD systems involved developing and implementing suitable image processing techniques that could extract the most efficient features representing the pathology to be detected. Many different features including, texture information, intensity information, different statistical metrics, shape and size information were extracted and tested using many different image processing methods. However extracting highly representative features from input data, especially from complex medical image data, requires very high expertise and involves most of the work regarding the development of the whole detection/classification system.

Deep learning methods eliminate the need for handcrafting features from the input data. Due to their deep architecture, deep learning methods have the ability to learn representative features directly from the data itself. This is called trainable hierarchical feature extraction (LeCun et al., 2015). Each layer of the deep network acts like a trainable feature transformation in which it transforms its input representation into a higher level representation. First layers of the network learn to extract low level features like intensity and edge information which are more common features and are shared among classes. When moved deeper into the network, higher level features are extracted from the images like object parts (combination of edges) and object models which are more specific to each class. In this regard, deep learning systems are capable of learning hierarchy of features from the input data with increased level of complexity with an increase in their depth. This presents a great advantage when compared to "shallow" methods where there is not enough number of computational layers for the method to learn to extract features efficiently without the need of any external intervention.

Even though eliminating the need for the image processing based feature extraction relatively simplifies design process of deep learning based CAD systems and increases the performance greatly (Işın et al., 2016), designing an effective deep learning architecture (choosing the most efficient method and training algorithm, choosing the most suitable number of layers for the task in hand) is the main challenge. In this regard, current deep learning research heavily focuses on developing improved architectures to improve the

detection/classification performances. There are three main types of deep learning architectures. First type is the feed-forward methods. These methods are namely, Multilayer Deep Neural Networks and Convolutional Neural Networks (CNNs). CNNs can be considered as the current state-of-the-art for medical imaging applications. They have obtained very high performances at very important and complicated medical imaging tasks in the recent years (Işın et al., 2016). CNNs mainly consist of convolutional layers. In convolutional layers, convolutional filters are used instead of traditional neurons. These filters carry out convolution operation with the input. Due to the nature of convolution operation, using convolutional filters are intrinsically efficient in image applications. Number of convolutional filters used in each convolutional layer and their dimensions are dependent on the designed architecture for a given application. Each single value of the convolutional filter matrix acts like a weight and updated independently through the learning process. Due to the complexity of convolutional filter operations when compared to the neuron operations in the traditional ANNs and since many numbers of filters in each layer and many numbers of overall convolutional layers can be implemented in deep architectures, overall complexity of the calculations throughout the training process increases heavily. This increases the computational cost and training time excessively. In some applications training times can even take several days to complete even on very powerful GPU powered computer systems. Second type of the deep learning architectures is feed-back systems. Stacked Sparse Coding systems and De-Convolutional Nets are main methods of this type. Last type includes the bi-directional systems. Deep Boltzmann Machines and Stacked Auto Encoders are the main examples.

Again since, in this thesis a deep learning system is not developed and trained from the scratch, further details about the mathematics and learning methods regarding the CNNs and other deep learning architectures will not be explained here. Again reader can apply to other sources for more detailed information regarding those topics.

2.6.3 Transfer deep learning and AlexNet

Deep learning, especially CNNs, gained increased popularity over the recent years. Its applications spread rapidly to the various fields of science. However it is not possible for every researcher in those fields to have access to powerful GPU based computers to train a

deep learning algorithm from scratch and to have the necessary machine learning expertise to design custom made efficient deep learning architectures for implementing in their research area. Additionally, deep learning systems require excessive amounts of training data (usually accompanied with ground truth data) to be trained efficiently. Obtaining such a data set may not also be possible for every situation. As introduced in Chapter 1, transfer learning (and its variation transfer learning with fine tuning) presents an effective way of deep learning application in those types of conditions. In transfer learning, a pre-trained deep learning system that is usually trained on a very intensive dataset is imported and used as a deep learning based hierarchical feature extractor for the desired application in question. Deep learning systems that are trained on the ImageNet dataset (Deng et al., 2009) are generally used for transfer learning applications. ImageNet contains a very large scale hand-annotated image database (over 14 million images belonging to over 20 thousand categories) and the deep learning systems trained on it (for the ImageNet challenge) try to correctly classify those images into one of the corresponding thousand different classes. Due to this nature of ImageNet dataset, deep learning systems trained on it can be versatile enough to be transferred and used as an efficient feature extractor for various other tasks.

AlexNet (Krizhevsky et al., 2012) is one of the CNN based deep learning systems that was trained on ImageNet dataset and obtained very high performance on the classification challenge. AlexNet was trained on a total of 1.2 million high resolution RGB images of the ImageNet dataset and learnt efficiently to classify them into a thousand different classes. It has a very complex architecture with 60 million parameters, 630 million connections and over 650.000 neurons. As it can be seen in Figure 2.23, AlexNet contains total of eight main layers. Convolutional layers make up the first five layers which contain the convolutional filters. Last three layers are conventional fully-connected layers which contain traditional neurons. Additionally first, second and fifth convolutional layers are followed by max-pooling layers. Max-pooling layers are filter layers that basically reduce the dimensions (downsample) of the previous layer's output by only taking into consideration the maximum value of the neighboring neuron group falling inside the max-pooling filter (see Figure 2.24). At the end of the network, output of the last fully

connected layer is connected to a softmax layer which generates a final distribution over the desired thousand class labels.

AlexNet receives inputs as RGB images. Acceptable input image size of the network is 227x227x3. In its first layer there are 96 convolutional filters each with a size of 11x11x3. This first layer filters the input image with 4 pixels stride (defines how the filter moves on the input space during convolution operation). All later convolution layers of the network implement 1 pixel strides. After the output of the first convolutional layer is max-pooled with a 3x3 max-pooling filter with a stride of 2 pixels, it is then fed into the second convolutional layer which contains 256 filters each with a size of 5x5x48. Again output of the second convolutional layer is similarly max-pooled and then fed into the third convolutional layer which contains 384 convolutional filters with a size of 3x3x256. Fourth convolutional layer is directly connected to the third and the fifth is directly connected to the fourth without any max-pooling. Fourth layer contains 384 filters with size of 3x3x192 and the final convolutional layer contains 256 filters with size of 3x3x192. Output of the fifth convolutional layer is also first max-pooled (with similar properties as before) and then fed into the following fully connected layers.

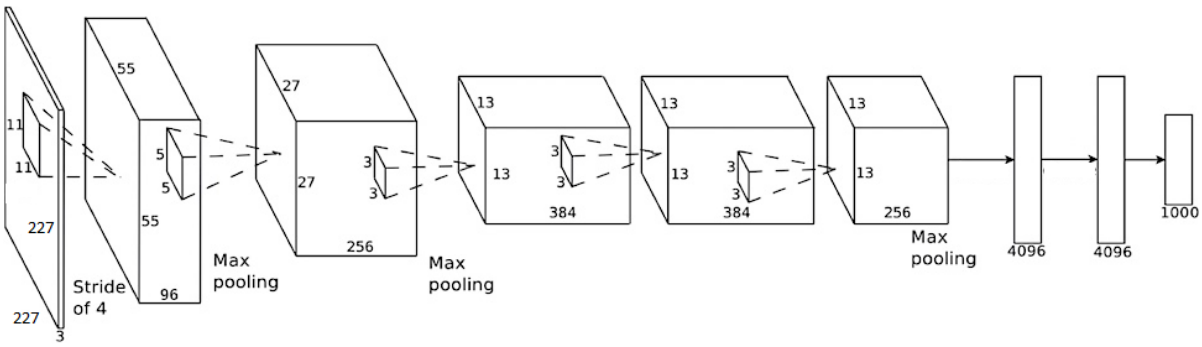


Figure 2.23: AlexNet CNN architecture. 227x227x3 sized image is fed into network, which is then processed by 5 convolutional layers and 3 fully connected layers. Max-pooling layers are not shown on the figure. See text for more detailed explanation. Image is adapted from (Krizhevsky et al., 2012).

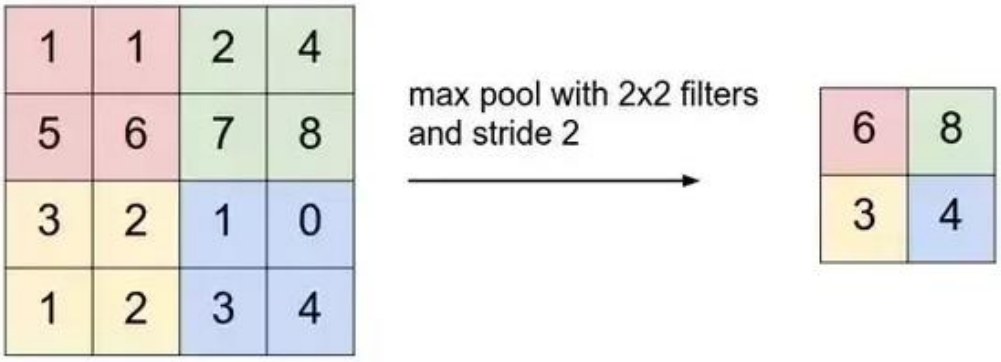


Figure 2.24: Simple illustration of a max-pooling operation. Max-pooling filter of size 2x2 takes the maximum value of the neighboring neurons of the previous layer that falls into the filter space (region with red). Then filter moves with the stride number (2 in the example) to its next location (region with green) and repeats the pooling. This continues for the whole output space of the previous layer. In the end of max pooling layer, the output of the previous layer is downsampled into new dimensions.

There are three fully connected layers, with first and second having 4096 neurons each and the last fully connected layer having 1000 neurons. Output of the last fully connected layer is then fed into a softmax layer for carrying out the final classification into 1000 different classes. Output size of a convolutional layer can be calculated by the formula;

$$output\ size = (I_S - F_S + 2P)/S\ + 1 \qquad (2.10)$$

where I_s is the input size (assuming height and width is equal), F_s is the convolutional filter size (assuming height and width is equal), S is the stride and P is the padding. From this equation; for the first convolutional layer of the AlexNet, since input size is 227, filter size is 11, stride is 4 and the padding is 0 the size of the one dimension of the output of the first convolutional layer becomes 55. Since both dimensions of the input and the filters are equal and since first layer has 96 convolutional filters, output volume will also have a depth of 96 and the overall size of the output volume of the convolutional layer becomes 55x55x96. This corresponds to a total of 290,400 neurons. Each of these neurons is connected to 11x11x3=363 weights and 1 bias value. So, in the first convolutional layer

there are 290,400x364=105,705,600 connection parameters. However, 55x55 neurons of each depth slice of the convolutional layer output volume shares the same weight set (same filter). Therefore, in the first convolutional layer there are 96 different sets of weights. Since each weight set (filter) contains 363 weights, this makes up a total of 363x96=34,848 unique weights in the first layer. If bias values are also added, there will be 34,944 unique parameters (34,848 weights plus 96 bias values). All of these weights are updated throughout the training process to learn representative features regarding the input images. Using the same logic, neuron numbers in every convolutional layer can be calculated. In this regard, AlexNet contains 290,400 neurons in its first convolutional layer, 186,624 neurons in the second, 64,896 neurons in the third, again 64,896 neurons in the fourth and 43,264 neurons in the fifth convolutional layer. Final three fully connected layers also contain 4096, 4096 and 1000 neurons respectively. This makes up over 650,000 neurons, indicating the overall complexity of such deep convolutional architecture.

Training of such deep architecture (with millions of parameters) from the scratch without overfitting is still problematic even with a very large data base like ImageNet. Therefore several considerations were taken to reduce overfitting during training of the AlexNet (Krizhevsky et al., 2012). Considering the non-availability of a similarly very large dataset of medical images for different pathologies, designing this complex deep CNN for medical CADs and training it from scratch is not possible for every researcher/institute. Additionally, training of the AlexNet on ImageNet dataset took almost six days even with parallel training on two dedicated powerful GPUs. However due to this depth and complexity of the AlexNet and since it is trained on a very large and general dataset of ImageNet, it can efficiently learn hierarchical feature maps that can be easily transferred to other applications to extract deep features, eliminating the burden of handcrafting feature sets using image processing methods and also enabling a deep learning implementation at the desired field of science without facing the aforementioned difficulties.

Trained 96 filters (feature maps) of the first convolutional layer of the AlexNet can be seen on Figure 2.25. As it can be seen from the figure, filters on the first layers of a deep CNN learns to represent simple edge features. Filters on deeper layers learn to represent more complex features, like combination of edges or object parts. These already trained filters can be transferred for another application (like medical images for our case) and they can

scan the new images for those leant features, automatically extracting features as an output (Yosinski et al., 2014). For example, if only the first convolutional layer of the AlexNet is transferred and images of a new application (again with an image size of 227x227x3) is fed into it, transferred feature maps would scan the new input image and would generate an output of 290,400 features. This way, 154,587 values of the input image is transformed into 290,400 feature values. However, transferring only the first layer of a deep network does not provide efficient feature extraction since the features learnt at the first layers of a deep learning network are simple features like edges which are not sufficient enough to provide high inter-class variability and low intra-class variability for efficient differentiation of the to be classified images. In order to provide efficient automatic feature extraction, a "deep" transfer should be carried out by transferring many layers of the pre-trained deep learning network. When moved deeper into the network, extracted features increase in complexity. In this hierarchical feature extraction, each layer of the network transforms the features outputted by the previous layer into more complex features. In this regard, if the first seven layers of the AlexNet are transferred to a new application and the output of the seventh layer is taken as extracted features, these features would be robust enough to efficiently represent the images of the new application for accurate classification.

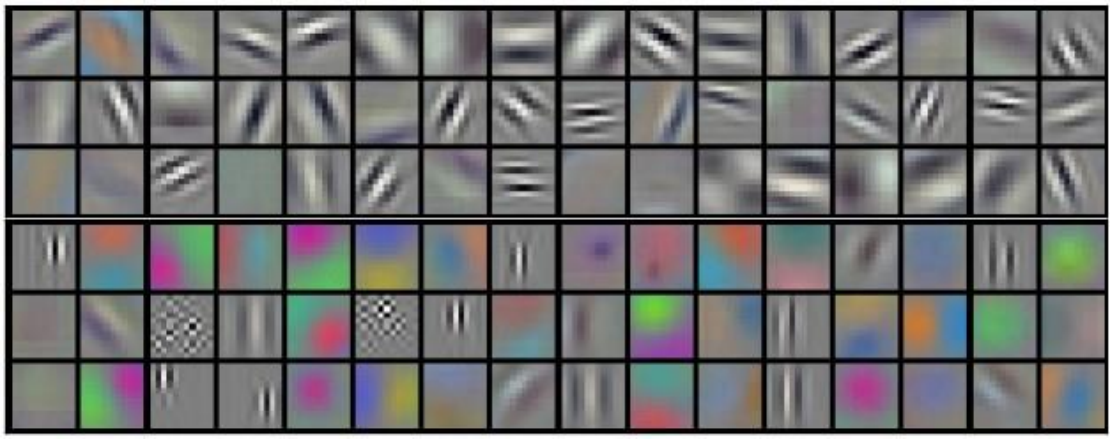


Figure 2.25: All 96 convolutional filters (with size 11x11x3) of the AlexNet's first convolutional layer after the whole training process is completed. These feature maps have learned to represent edge features with different locations and orientations. Image is taken from (Krizhevsky et al., 2012).

Up to its seventh layer, AlexNet hierarchically transforms the 154,587 values of the input image into 4096 robust feature values. Handcrafting that amount of robust features using image processing techniques requires high expertise, is very difficult and time consuming if not impossible. These automatically extracted features by the AlexNet can be then fed into a simple classifier (a SVM or a non-deep neural network) to carry out the classification of the input image into one of the desired classes of the new task. Furthermore, some of the transferred layers of the network can also be fine-tuned i.e. trained again with the new data to tailor fit the new task. How many layers to transfer from the existing pre-trained deep learning framework, whether to fine-tune or not and which layers to fine-tune are all depend on experimentation on the new task. Transfer deep learning method has proven its effectiveness previously in various fields of research and also in some other medical image and signal processing tasks (Işın and Ozdalili, 2017). Even in some medical imaging tasks, it proved to be more effective than designing a custom made deep learning network and fully training that network from the scratch (Tajbakhsh, 2016). In this thesis, due to all these advantages, transfer deep learning method is preferred and implemented to solve the medical lung lesion detection task problem.

CHAPTER 3
METHODS FOR PET RESPIRATORY MOTION CORRECTION

In this chapter, methods, which are implemented in this thesis, for the correction of motion artifacts resulting from the natural respiration of the patient during PET scan are presented in detail. Methods explained here are published in my work (Işın et al., 2017).

3.1 GATE Simulation

As mentioned several times in the previous chapters, PET devices are still rare in clinical environment and obtaining PET data regarding the FDG scans of the real patients is very expensive (due to the operating costs of the clinic regarding the device and FDG costs) and requires substantial time. Due to this, developing and evaluating new respiratory motion correction methods (or other methods regarding PET acquisition and image formation procedures) using clinical PET scanners on real patients is not an achievable task unfortunately for every researcher and institution. To overcome this limitation and to open the way for active research, powerful computerized simulators that can simulate realistic PET data can be used instead. Using these simulators, new methods can be designed and repeated experiments can be carried out for developing and evaluating effective respiratory correction methods.

For physical processes where there are significant uncertainties, Monte Carlo methods can be used for effective simulation. For PET imaging research (where there are significant uncertainties in image acquisition and other processes), Monte Carlo simulations are an established way for reproducing realistic PET data according to the chosen properties of the scanner, the patient (or patient model), and the conditions like attenuation, scatter events and random events. GEANT4 (Agostinelli et al., 2003) is a powerful Monte Carlo simulator that is mainly used for simulating the behavior of particles while traveling through any matter. It has many uses ranging from high energy physics to space physics and particle accelerators. Accordingly, GATE (Jan et al., 2004) is a specialized GEANT4 platform, which is specifically developed to create dedicated numerical simulations for medical imaging and radiotherapy solutions. Simulations regarding PET, CT and single

photon emission tomography (SPECT) modalities along with optical imaging modalities and radiotherapy simulations can be carried out effectively with little C++ programming knowledge using the flexible and evaluative GATE software. In current research, its use become almost indispensable for developing new medical imaging modality designs, more effective image acquisition methods and new image reconstruction and artifact correction techniques, especially for the medical imaging devices based on emission tomography.

For the special case of PET imaging, GATE simulation can be effectively used to generate realistic projection data and to reproduce the complete procedures of image acquisition and image formation. However, before the simulation can reproduce image acquisition and later the image formation procedures, it first requires an object (a patient or patient model) to simulate the PET image acquisition procedure on. Although researchers can use anatomic and physiological acquisitions (like CT images) obtained directly from real patients as patient models, computerized human phantoms were also developed for modeling the real patient anatomy and physiology to be specifically used in simulation studies. Using computerized human phantoms provides some advantages. Firstly, in computerized phantoms anatomical properties and the physiological functions are exactly known, which can be used as a ground truth to evaluate and improve the developed methods. Additionally, computerized phantoms can be repeatedly used for modeling different anatomies and different medical conditions providing a large and varied patient data to carry out the research on. However, when using real patients, obtainable knowledge regarding the anatomical properties and the physiological functions are heavily dependent on the used imaging device and the acquisition conditions. Also, it is not easy and practical to find and perform imaging on large numbers of patients possessing varied anatomical and medical conditions. Plus, repeated testing of different combinations of different parameters on real patients also raises ethical concerns.

In this thesis, 4D extended cardiac-torso (XCAT) phantom (Segars et al., 2010) is used as the computerized patient model to model the anatomy of the patient during respiration. XCAT is a recent generation phantom that extends the previously developed (previous generation) NCAT phantom (Segars et al., 1999). NCAT phantom is a four dimensional phantom, which includes three dimensional volumes to model realistic and detailed whole-body human anatomy plus fourth time dimension for again realistically modeling

physiological processes like cardiac and respiratory motion. As being the new generation model, the XCAT phantom improves upon the NCAT phantom by including very detailed whole-body anatomy, which contains thousands of detailed anatomical structures when compared to the limited anatomical structures (only hundred and only in torso region) of the NCAT, along with improved cardiac and respiratory motion models (see Figure 3.1 and 3.2). Due to this property it is very suitable for carrying out experiments regarding respiratory motion correction methods. In human anatomy, diaphragm is the primary muscle that regulates the regular breathing. During the inspiration phase of the respiration diaphragm muscle contracts pushing the organs located in the abdominal region downwards and outwards while at the same time rib cage moves upwards and outwards. Together, these movements during inspiration phase cause an increase in the thorax region volume inflating both of the lungs. In contrast, during the expiration phase of the respiration these movements are reversed by the relaxation of the diaphragm muscle, decreasing the thorax volume, which in turns deflates the lungs. Additionally heart, as a whole, also moves up and down due to the movement of the diaphragm throughout the respiratory cycle. During the development of the XCAT phantom, several 4D respiratory gated multi-slice CT acquisitions were performed on real subjects to simulate the respiratory model of the XCAT phantom which accurately models the movements of the aforementioned organs during the real human respiration. Similarly, several 4D cardiac-gated multi-slice CT acquisitions were also performed to simulate the cardiac model. Like the real human anatomy, in XCAT phantom cardiac and respiratory motions also interact with each other.

The XCAT program (https://olv.duke.edu/xcat Retrieved 7 February, 2016), accepts a configuration file as an input. With this configuration file many parameters regarding the desired anatomy and physiology can be set for the generation of the computerized phantom with the desired anatomical and physiological properties. Those parameters include; body size, organ size, organ labels, number of respiratory phases and the magnitude and the period of the respiratory motion.

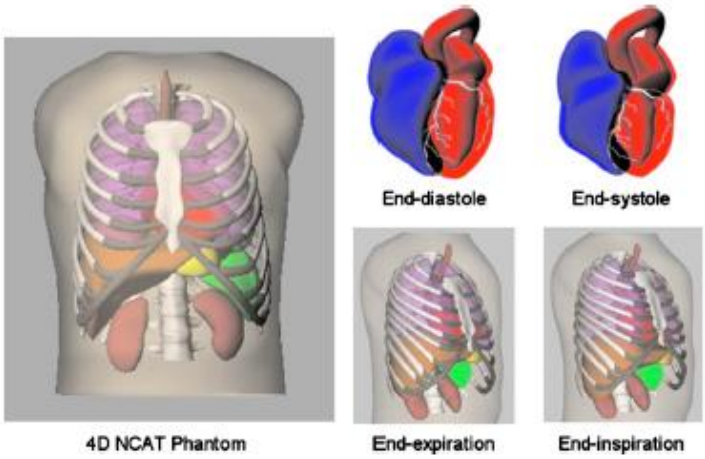


Figure 3.1: NCAT phantom (left) with cardiac (top right) and respiratory (bottom right) motion modes (Segars et al., 2010)

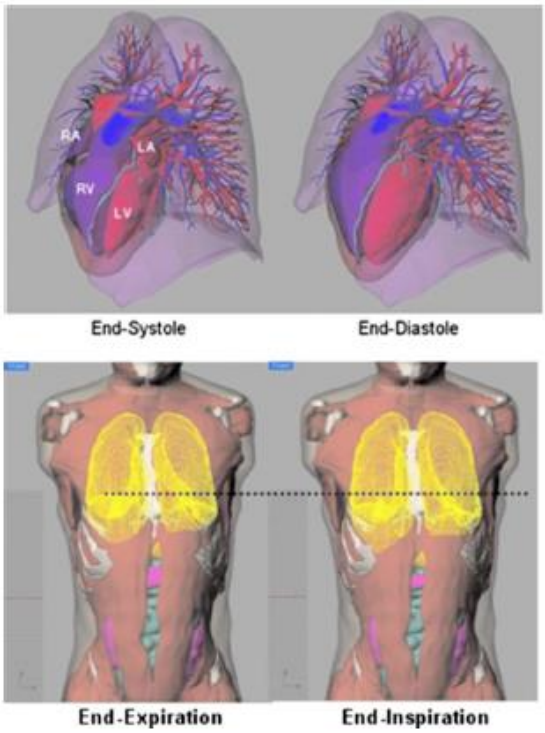


Figure 3.2: Figure shows the enhanced motion models and the whole-body anatomy of the new generation XCAT phantom (Segars et al., 2010)

Apart from these, technical parameters like; image resolution and slice thickness can also be configured. Additionally, user can place a lesion inside the lung or other tissues to observe the effects of motion on it. When this input configuration file is given, XCAT program generates a series of index (emission) and attenuation maps across a respiratory cycle (see Figure 3.3). Index map is a 3D volume, represented as image slices, of the desired region (torso in our case), where each voxel represents the label or the intensity of an organ. In other words, it gives slices of images of the torso which show the effects induced on each area and material. It is used later by the simulator to generate the activity effects. On the other hand, attenuation map is a 3D volume where each voxel represents how the corresponding anatomical volume (of that voxel) attenuates the emitted gamma photons, so the attenuation coefficients for gamma photons. It is similar to a CT image, but the attenuation map generated here is at gamma photon energy level (511 keV for XCAT attenuation map and 40-140 keV for CT image). Attenuation map is used by the simulator during the last image reconstruction phase.

After the XCAT program generates the phantom as explained above; image acquisition is simulated by the GATE simulator, by using the simulated PET data of the GATE final images for each respiratory phase are reconstructed using the OSEM reconstruction algorithm and Demons algorithm is used for performing the registration on the reconstructed PET images and for extracting the respiratory motion fields. Details of these steps are all explained in the next sections.

3.2 Motion Compensation

The pathology that is focused in this thesis is lung cancer. For the diagnosis of the lung cancer, PET scan is performed over the thorax region of the patient which is heavily suspect to respiratory motion. Designed GATE simulation is carried out in six respiratory phases and therefore a potent respiratory motion correction method should be implemented to overcome the blurring effects induced on the simulated PET images by the respiratory motion. Only by accurately estimating and correcting the motion effects, artifact free PET images can be reconstructed that allows accurate localization of the lung lesions which is of great importance for accurate clinical diagnosis of lung cancer. The main idea behind

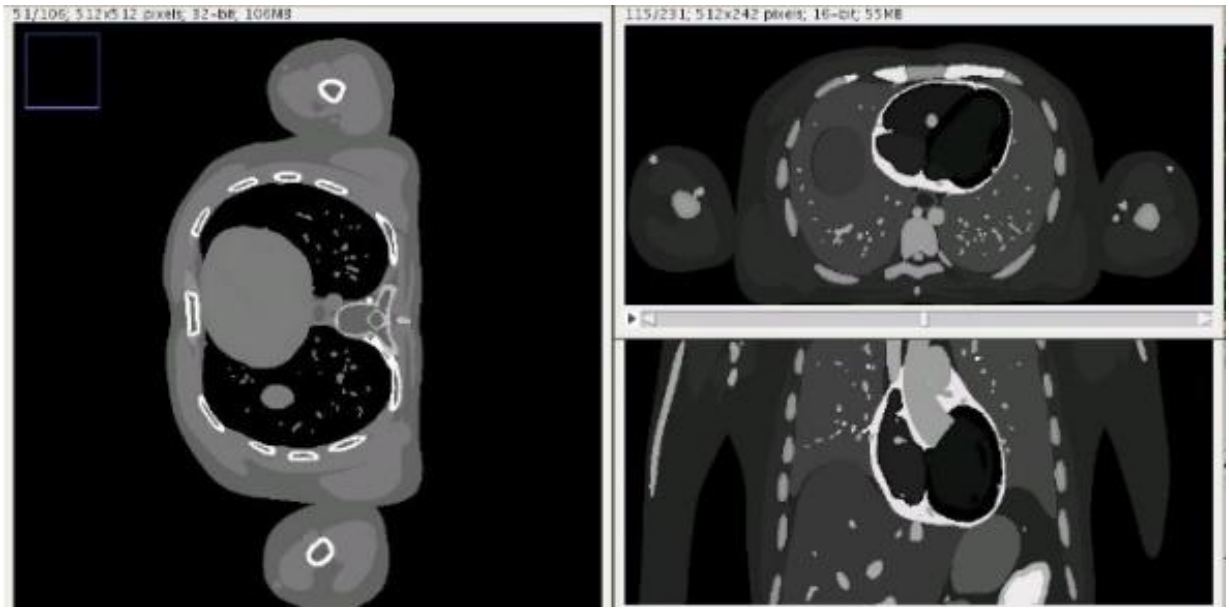


Figure 3.3: A slice of the attenuation map from transversal view (left). Transverse (top right) and coronal (bottom right) slices from the index map

the motion compensation method is to associate the intensity of the same pixel in two consecutive frames, which represent consecutive respiratory phases, and formulate a displacement vector using the difference between them. Using this method, every pixel on the PET image is assigned a 3D movement vector and motion fields are generated using all these vectors. Motion fields are then used to estimate the true intensity value of each pixel on the PET image to reduce the respiratory motion induced errors on the image.

In this thesis, Demons (Pennec et al., 1999) registration algorithm is implemented for predicting the motion fields. Demons algorithm studies the difference in intensity between the corresponding pixels in consecutive PET image frames (in time dimension). In other words, it correlates a static image/volume with a moving image/volume. To eliminate the motion effects by carrying out a motion corrected image reconstruction, three different techniques are implemented and tested to obtain the motion fields before the reconstruction. In the first technique, true motion fields that directly inform the real movement of the simulated XCAT phantom is used. This is given by the XCAT program. In the second technique, PET modality derived motions fields are used, which are obtained

from the PET images that are generated by non-attenuation corrected PET image reconstruction. And in the final technique, MRI modality derived motion fields are used. MRI motion fields are obtained from the MRI transformed version of the phantom by using MRilab program (Liu et al., 2017). After the motion fields are derived using the proposed techniques, they are integrated into the ordered subset expectation maximization based (OSEM) reconstruction algorithm (Hudson and Larkin, 1994) for the reconstruction of respiratory motion corrected PET images.

3.3 Simulation Setup

As mentioned in Section 3.1, first the patient model, i.e. the phantom is need to be generated in order for the GATE simulation to perform on. XCAT program is used to create the phantom of the thorax region (as a set of anatomical slices) with physiological respiratory movement parameters like respiratory cycle length and ribs extension length set as default. In this thesis cardiac motion is ignored to simplify the process. Six respiratory phases are used in the phantom generation. If cardiac motion phases were also included there would be total of 36 phases (6 respiratory x 6 cardiac phases) thus frames to process on. Slice resolution is set as 512x512 with a 0.1030 cm voxel size and 0.0687 thickness. These parameters are important for defining the final image quality and overall simulation speed. With these parameters XCAT generates the phantom with a total of 231 transversal image slices ranging from diaphragm to the top of the lungs covering the whole thorax region. As a result, XCAT software generates six respiratory phases both for index and attenuation maps with the defined parameters. Later these maps are post-processed. Index map is transformed into 16 bits from 32 bits and cropped for faster simulation. On the other hand, attenuation map is flipped with regards to its x-y axis to match the GATE orientation.

Total of six lung lesion are inserted to the generated phantom; Three 4 mm lesions and three 8 mm lesions in right and left lung on the same coronal slice. All the processing steps are repeated for each lesion and their index maps are combined as another copy of the index map. Also, the same process is repeated for the attenuation maps. Figure 3.4 shows sample slices of the index map for the final XCAT phantom with lung lesions inserted.

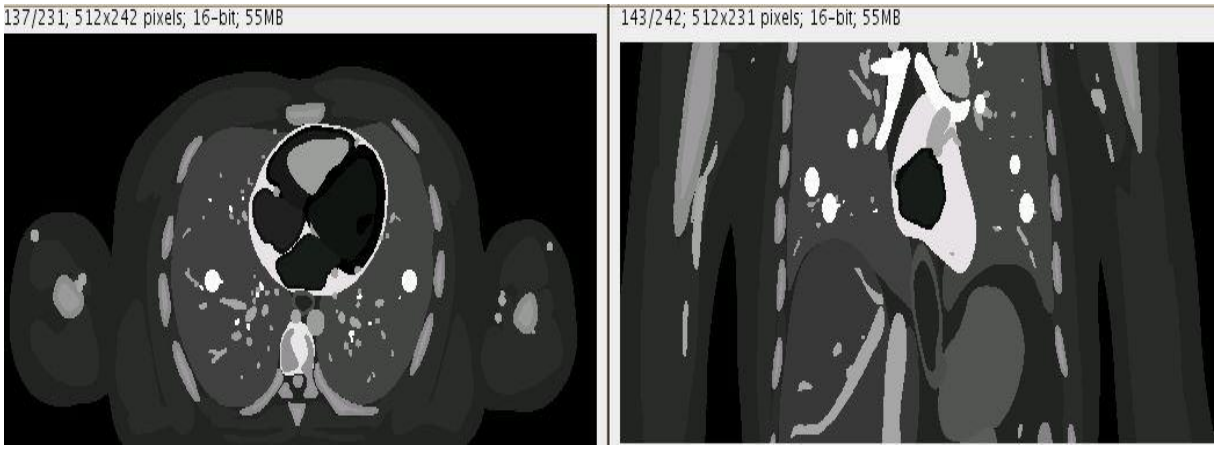


Figure 3.4: Final XCAT index map with the inserted lesions (white spheres)

After these all aforementioned steps are performed, XCAT phantoms are ready for the GATE simulation. Two separate instances of the simulation are performed, one with the six lung lesions (lesion phantom) and one without any lesion (background phantom).

The main inputs of the GATE simulator, the attenuation file and the activity file (the index map output of the XCAT is used here) are set and the simulation is carried out. As mentioned previously index maps for the lung lesions are combined and saved as another copy of the index map, so activity files of the background phantom and the lesion phantom differs in terms of contrast but their attenuation files are same. In this regard, final contrast of the images depends on lesion activity (5 Bq) and surrounding background tissue activities (1 Bq). Additionally, other technical parameters, like the time sampling, experimentation length and gamma photon energy (511 keV) are all set as the source file of the simulator. During the implementation, geometry of the PET scanner is set as the default setting; which is a Siemens mMR scanner. Also the GATE material file is set as default.

User can determine the length of the simulation and the acquisition. Since, recurrent simulation failure can happen which causes hours of accumulated simulation data to be lost, 100 seconds long simulation is divided into hundred 1 second long simulations. Therefore, each simulation has a different random seed. Outputs of these simulations are 100 list mode files, which are then combined together. Whole simulation procedure is

repeated; once for the background phantom, where there is 0 Bq activity for the lesions and common activity values for the thoracic tissues and organs; and once for the lesion phantom where lesions have 5 Bq activity. It should be noted that, it is of utmost importance to have the same number of list mode files as the output of both simulation instances for not affecting the set contrast level.

Finally, two combined list mode files are converted into two sinograms, one as the result of the lesion phantom simulation and one as the result of the background phantom simulation. Results of the both instances of the GATE simulation can be seen as sinograms in Figure 3.5. These sinograms are then joined together, and an OSEM reconstruction algorithm is implemented using the joined sinogram and the initial attenuation map to generate the final PET image.

To overcome the noise induced during the reconstruction, a median filter is used to filter every slice. After the noise removal, all of the previous simulation steps are repeated for each of the six separate respiratory phases. Then, the next step can be proceeded, which is the final reconstruction with or without the respiratory motion correction.

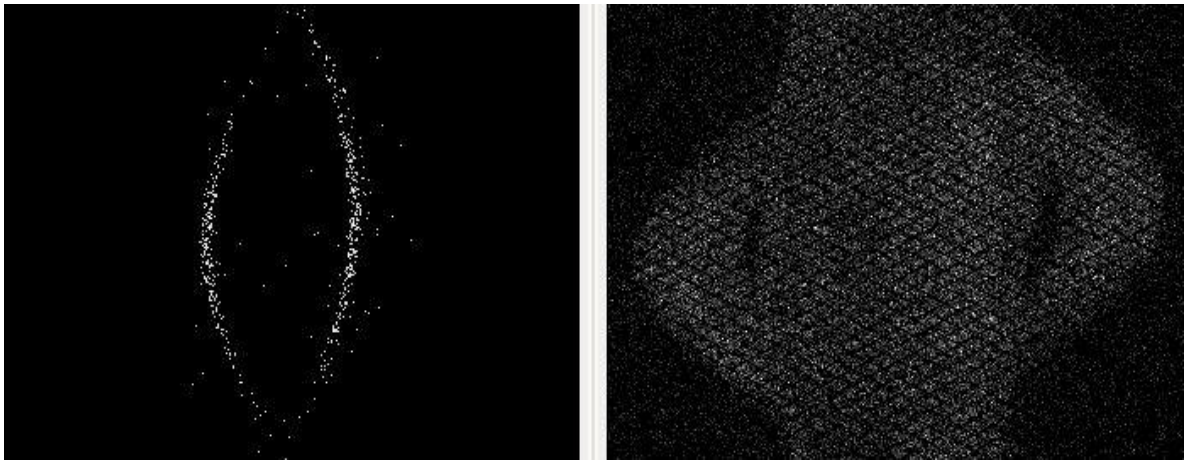


Figure 3.5: Sinogram obtained after simulating the lesion phantom (left). Sinogram obtained after simulating the background phantom (right)

3.4 Image Reconstruction

First the reference image, which will be used for comparing other methods, is reconstructed. Reconstruction of the reference frame is carried over a single respiratory phase, the first one, over a long duration to obtain a noise free image. Therefore, acquisition time in the GATE source file is set as 5000 seconds, thus stimulating a perfect motion corrected image. It should be noted that, such an acquisition time for a real patient imaging is very long. Thus implementing such a motion correction method is not possible in reality.

After the reference image is reconstructed, an un-gated image is reconstructed as the second reconstruction. In this reconstruction, sinograms of the all six respiratory phases are added and reconstructed together. Since, respiratory motion is expected over six phases, this reconstruction results in a blurred image, where the lung lesions are spread not representing the actual case. The final image obtained after this reconstruction is simply the sum of all respiratory gates while not taking the breathing phase into consideration. So, it is not a motion corrected image. The summation of all gates is done at the sinogram level.

For the third reconstruction, a reconstruction with true motion fields is implemented. True motion fields regarding the phantom are extracted using the XCAT software. To do so, mode 4 of the parameter file should be selected. Implementing this method provides the information regarding the movement of each voxel throughout the six respiratory phases. Output of this reconstruction method is expected to provide best image quality since the motion correction methods uses the actual respiratory motion information directly obtained from the phantom. In contrast, a PET derived motion fields reconstruction method is also implemented for comparison. This time, Demons registration of the reconstructed PET images in each respiratory phase is used to extract the motion fields. Additionally no attenuation correction is performed in this method.

As the final reconstruction, MRI derived motion fields reconstruction is implemented for respiratory motion correction. Mrilab program is used for applying a spin echo sequence on the XCAT phantom. Thus MRI images are obtained for the phantom. This time Demons algorithm is implemented on those MRI images of the phantom for the extraction of the

motion fields. These extracted motion fields are then incorporated to the reconstruction algorithm for reconstructing the final motion compensated image. Complete workflow of the proposed system is shown in Figure 3.6.

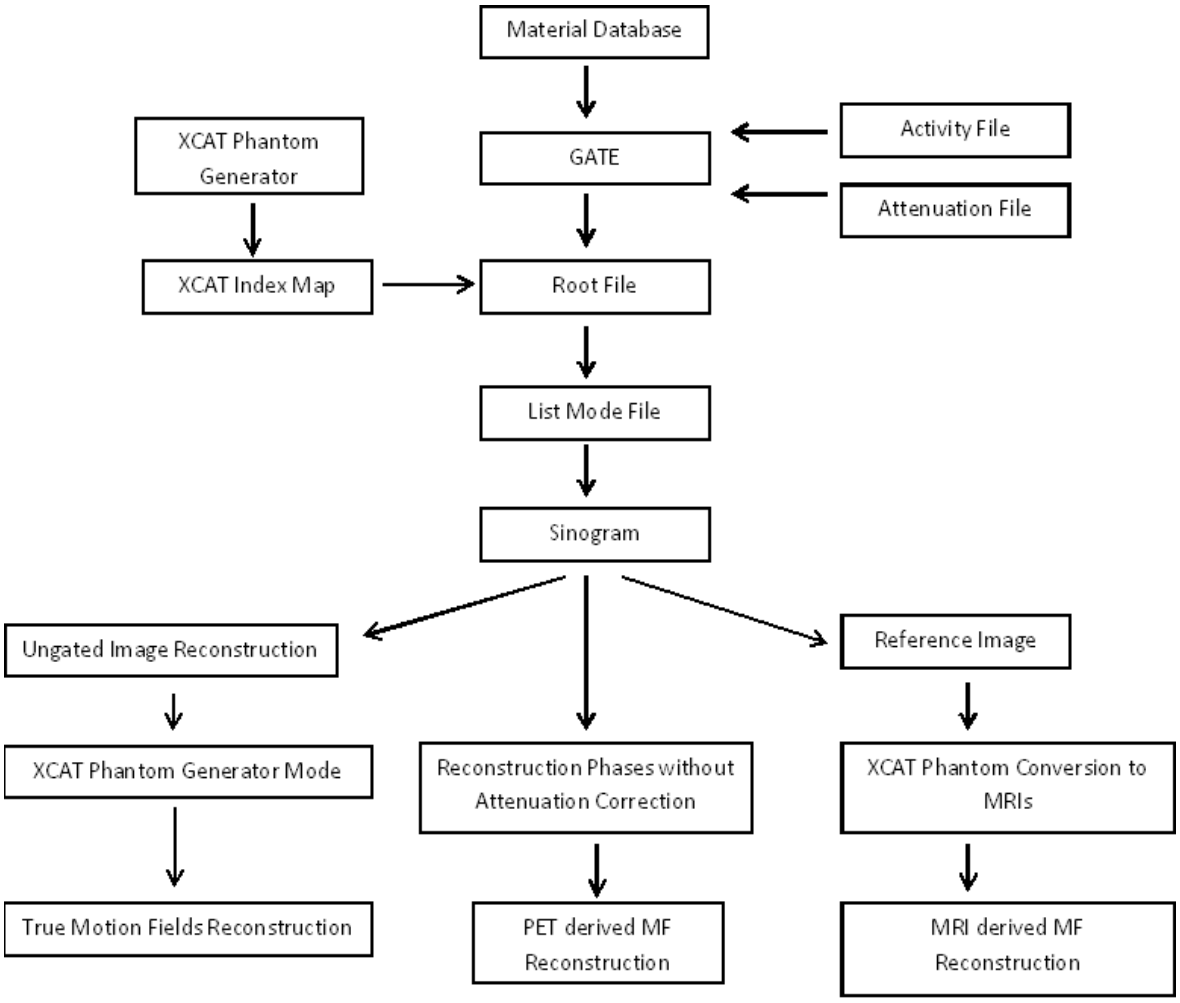


Figure 3.6: Complete workflow of the proposed system to obtain three different respiratory motion correction methods

CHAPTER 4
METHODS FOR LUNG LESION CAD SYSTEM

In this chapter, methods, which are implemented in this thesis, for the development of a deep learning based computer aided detection (CAD) system for automatically detecting lung lesions are presented in detail.

The proposed system can be considered in four main parts as represented in Figure 4.1. First, a medical lung image dataset, in which some image slices contain lung lesions, is pulled from a public database. Then, that dataset is adjusted for the next step, which is feature extraction using transferred deep learning. After the features representing the lung image slices are extracted automatically, those features are used as an input into a conventional artificial neural network (ANN) based classifier for classifying the image slices into one of the output classes; not bearing or bearing lung lesion.

This proposed method is a slight modification of my previous published work (Işın and Ozdalili, 2017), which was a successful implementation of a transfer learning method for another medical task.

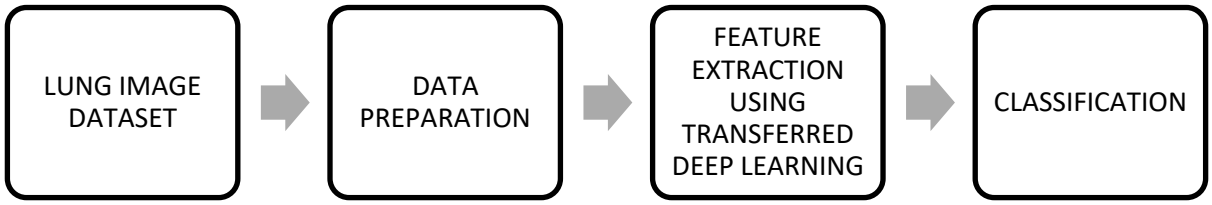


Figure 4.1: Block diagram of the proposed lung lesion detection system

4.1 Lung Lesion Image Dataset

In this thesis, due to the several reasons that are explained in Section 1.4 and Section 2.6, a CT based lung lesion image dataset is used for the development of the proposed deep learning based CAD system instead of a PET based image dataset. The Public Lung Database (PLD) (Reeves et al., 2009), which is developed for the specific purpose of aiding CAD research, is selected for the proposed system.

PLD dataset contains total of distinct 93 cases, where in each case there are varied number of CT image slices, ranging from 39 to 275 slices. Figure 4.2 shows some example CT image slices from the dataset with some containing and some not containing lung lesions. The dataset is also divided into several subsets, where each subset represents cases for different lesion types and conditions. For this thesis subsets of namely “Single Small Lesion”, “Single Large Lesion”, “Multiple Lesions” and “Sequential Scans” are selected. In Single Small Lesion subset there are total of 16 cases where CT images are acquired from patients bearing single small lung lesion volumes. The number of acquisition slices for each case is varied due to different procedures and scanners used during acquisitions of each different case. In this subset there are total of 6,433 CT lung image slices, in which 5,668 of them does not contain any lesion and 765 of them contain lesions. The second selected subset contains total of 12 cases acquired from patients bearing large lung lesion volumes. Again slice numbers vary among cases. In total, there are 2,911 slices in this subset in which 351 of them contains lung lesions. The third selected subset, “Multiple Lesions” subset contains 2 cases in its current state, but the first case contains two distinct acquisitions obtained with one month interval and the second case contains three distinct acquisitions obtained with two months intervals. In the cases of this subset there are multiple different (in terms of type and location) lesion volume observations for each case and also metastatic observations which provide a challenge for the automatic detection/classifier. In this subset there are total of 999 lung image slices, in which 220 of them bearing lung lesions. The last selected subset “Sequential Scans” contains 24 cases in which every case contains varied number of sequential scans of the same patient for the purpose of assessing therapy response. Since therapy response assessment in not an aim for the proposed CAD system, this subset is treated as a subset which contains cases with

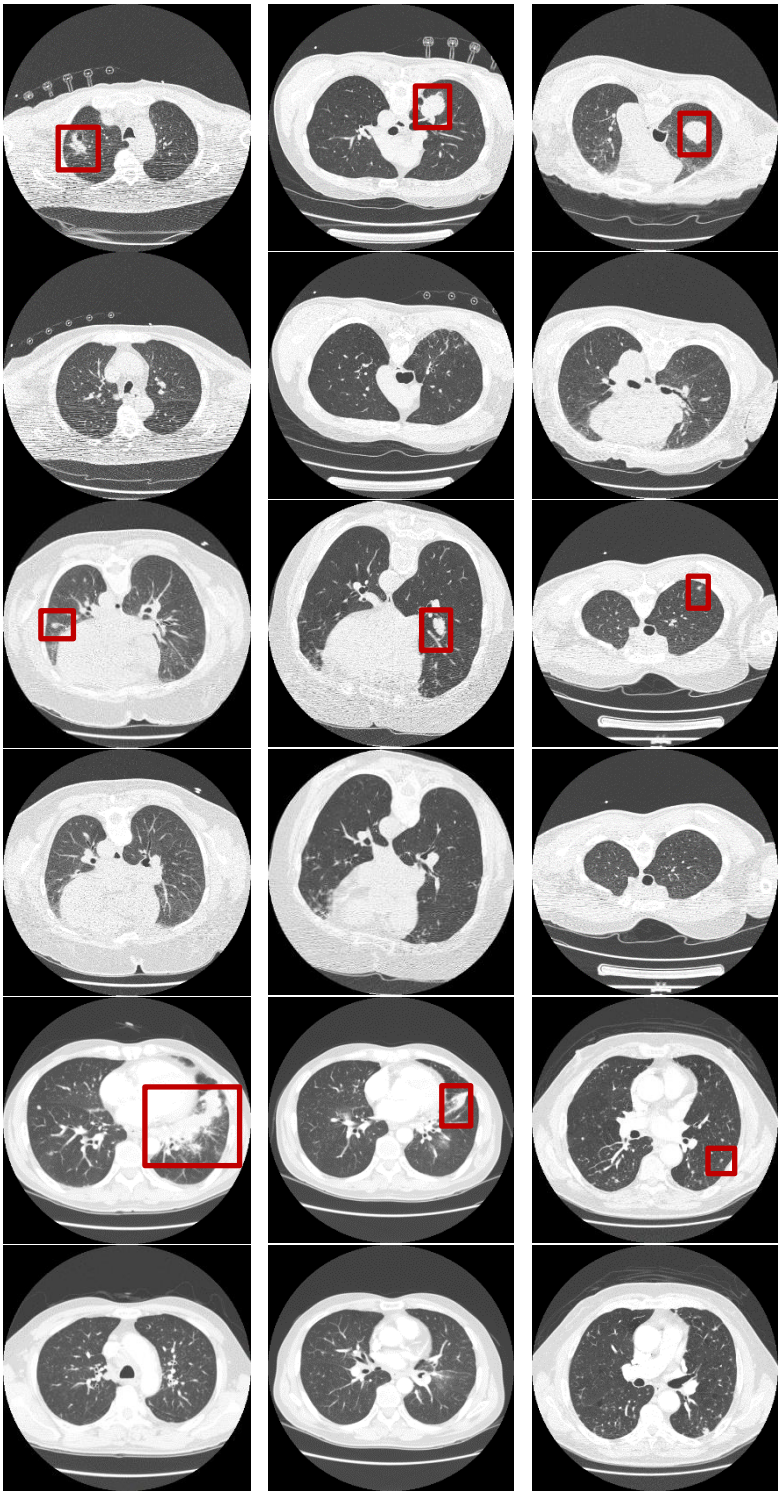


Figure 4.2: Example CT slices from PLD dataset. Images with lung lesions are marked at lesion location. It is obvious from the images that the detection of lung lesions manually from CT images requires high anatomical knowledge and expertise.

varied types of lung lesions (single large, single small, metastatic and multiple) and only one acquisition per a case is selected. Again number of image slices for each acquisition is varied among cases. In this subset total number of selected image slices is 3,532 in which 620 of them bear lung lesions and 2912 of them do not.

So the initial selected lung CT image data for the proposed CAD system development contains total of 13,875 lung image slices in which, 1,956 of them bear lung lesions and 11,919 of them do not bear any lung lesion. However, several initial slices along with several final slices of each case acquisition are image slices that usually image the non-lung tissue above and below the lungs, thus do not contain any lung lesions. 1,183 such image slices are identified in the selected data and discarded. So as a result, the final CT lung lesion image data contains total of 12,692 image slices, where 1,956 of them bear lung lesions and 10.736 of them do not bear any lung lesions. Both lesion bearing and non-bearing image sets are shuffled randomly in themselves and each one is then distributed into two subsets of training and testing datasets with 6 to 4 ratio.

As a result, 1,174 lung lesion bearing image slices are combined with 6,442 non-bearing slices to form the final training dataset of total 7,616 lung CT image slices. On the other hand, 782 lung lesion bearing and 4,294 non-bearing image slices are combined to form the final testing dataset that contains total of 5,076 lung CT image slices. Later, both training and testing datasets are randomly shuffled in themselves. In the end of this dataset formation step, obtained training and testing datasets are ready for the next step which is feature extraction using transferred deep learning.

4.2 Feature Extraction Using Transfer Deep Learning

After the lung CT images from the PLD dataset are selected, they need to be first prepared to be in line with the AlexNet which is the chosen pre-trained CNN based deep learning framework for carrying out the automatic feature extraction for the proposed CAD system.

As previously explained in every detail in Section 2.6.3, AlexNet was trained on 1.2 million high resolution RGB images of the generic ImageNet dataset and classifies those images into 1000 different image classes with a state-of-the-art performance. AlexNet`s CNN architecture contains five convolutional layers and three fully-connected layers. As

stated previously, since those layers of the AlexNet are trained on a very large annotated dataset, connection weights of those layers can be generic enough to be easily transferred and implemented even for a complex medical image classification task. Accordingly, given the lung CT images from the PLD image dataset as inputs, outputs of the deeper layers of the pre-trained AlexNet can be extracted as very representative features that provide valuable information regarding the presence of a lung lesion. Otherwise, without an automatic deep feature extractor like AlexNet, extracting similarly representative features from the complicated lung CT images using traditional feature extraction/image processing methods would be very difficult, adversely affecting the overall performance of the CAD system for the reasons that are again explained in every detail in Sections 1.4 and 2.6.

Images from the PLD image dataset are in DICOM format. First, those image slices are converted into 512x512 resolution .png format using a DICOM convertor. After the conversion, all images are adjusted into 227x227 resolution as accepted by the AlexNet. Another issue is that, AlexNet was designed for processing on RGB images. In this regard first convolutional layer of the network computes on a three dimensional input. To overcome this issue, our 227x227 sized lung CT images are reproduced three times for each image slice to imitate an RGB image. After the lung CT images of the PLD database are converted into 227x227x3 format, they are fed as inputs into the transferred AlexNet. As mentioned again in earlier related sections, deeper layers of a deep CNN architecture provide important information which can be used to classify its inputs. However, last fully connected layer of the AlexNet and the later softmax and output layers are all configured for classifying 1000 classes of the ImageNet. Although these final layers can be fine-tuned for the new lung lesion classification task, here it is preferred to extract the outputs of the previous layers and use them as learnt features to be fed into another simple classifier, which will carry out the final classification. In this regard, features extracted from the 6^{th} layer of the AlexNet, which is the first fully connected layer (will be referred as FC6) and the 7^{th} layer which is the second fully connected layer (will be referred as FC7) are used and compared in this thesis. As a result, since both FC6 and FC7 layers contain 4096 neurons each, AlexNet extracts 4096 features from the every input lung CT image slice. Obtained 4096 features representing each slice of image are then converted into feature vectors with 4096 elements ($F_v = f_1, f_2, f_3, \ldots . f_{4096}$) for all image slices. These vectors are

combined for all input image slices eventually forming a 4096 dimensional combined feature vector, i.e. since 4096 feature values are extracted from each input image slice, each input slice (or observation in general) is represented by 4096 variables/dimensions.

In this regard, for training, feature matrix size becomes 7616x4096 (7616 images/observations with 4096 dimensions each) and for testing it becomes 5076x4096. These combined feature vectors are then used as an input in the classification step which will be explained in the next section. Figure 4.3 illustrates the transfer deep learning based automatic hierarchical feature extraction procedure carried out on the Lung CT images by the AlexNet.

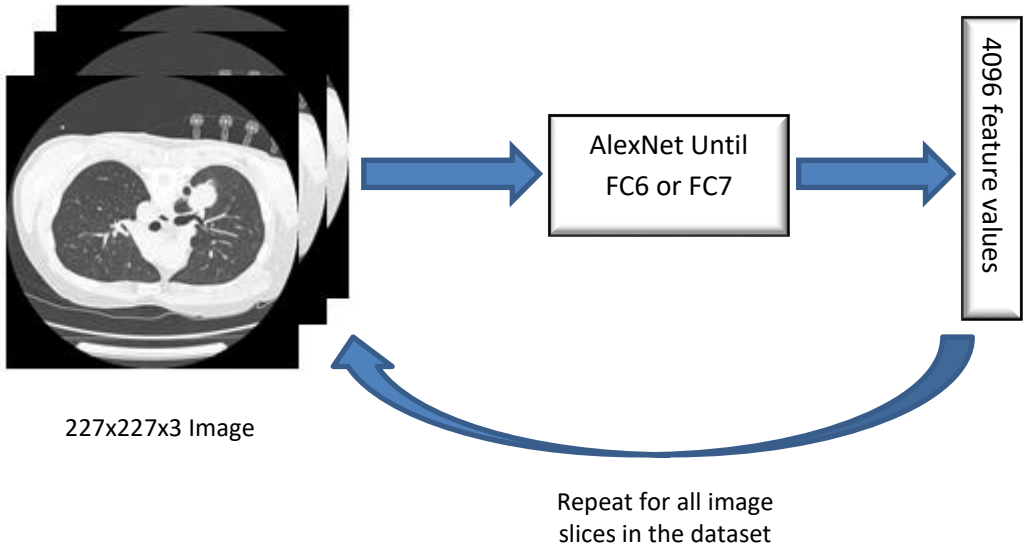


Figure 4.3: Transfer deep learning feature extraction procedure.

Additionally, to provide evaluation for the FC6 and FC7 AlexNet deep learning feature extraction, scale invariant feature transformation (SIFT) method (Lowe, 2004) is also applied to the image dataset for extracting features. SIFT extracts 128-dimensional feature vectors for each input image. Principal component analysis (PCA), which is explained in more detail in the next section, is initially performed on these feature vectors to reduce the dimensions into 64. Later, Fisher encoding (Sanchez et al., 2013) is performed on the 64-dimensional feature vectors with 32 Gaussian mixture models to generate 4096

dimensional Fisher vectors (no. of feature dimensions x no. of Gaussians x 2), as the final feature vectors which will provide an alternative to the AlexNet based automatic feature extraction. SIFT followed by fisher encoding based feature extraction method will be referred as F_SIFT from now on.

Further, manual feature extraction is also carried out from the images using traditional image processing techniques and also used as a comparison in the evaluation of the AlexNet feature extraction. As mentioned several times before in this thesis, deciding on which features to craft manually from images is a very complex and difficult task and contributes heavily to the overall classification performance of the CAD system. In this thesis, the idea that it is not easily achievable to handcraft features that represent images as efficiently as features learnt automatically by deep feature learning, especially for a complex medical imaging task, is supported. Thus, not much effort is spent for deciding on which types of features to select and hand-craft for lung lesion detection task and one of the methods already applied previously (and easily implementable) for a medical task in literature is chosen randomly and repeated (Aggarwal and Agrawal, 2012). In this regard, features regarding first and second order statistics are extracted from every lung CT image slice. There are four main extracted first order statistical features, which are mean, variance, skewness and kurtosis. Mean is the mean pixel intensity value in the image. Variance is the deviation measure of the intensities of pixels in the image from the mean pixel intensity. If the first-order histogram of the image is calculated, skewness corresponds to the asymmetry of the histogram around mean value. On the other hand, sharpness of the histogram gives the kurtosis. For the second order statistics features, contrast, angular second moment, correlation, entropy and homogeneity values are calculated from each image. While first order statistics provides features regarding the intensity (of grey level) distribution, second order statistics provide locational features. Contrast gives information regarding the local variations in gray level. Correlation gives the relationship between two pixels in different directions. While contrast takes high values for high contrast images, homogeneity is a feature that gives high values for low contrast images. Entropy is a feature that represents the smoothness of the image. For smoother images entropy is lower. On the other hand, angular second moment is another measure for local smoothness where higher the smoothness higher will be the angular second moment

measure. Second order statistics depend on distance and orientation, so mean and range of the every second order statistics feature calculated from all four directions are taken as the final feature values. Thus, there are 4 first order statistics feature values and 10 second order statistics feature values. Additionally to capture local context of the image images are rescaled to 220x220 size by ignoring some initial and late pixel rows and columns of the 227x227 images. Then 121 20x20 local image patches are extracted from each image and above explained hand-crafted feature extraction method is repeated also for every local patch. In the end 1,694 feature values extracted from the local patches are added to the 14 features extracted previously from the global image making up a total of 1,708 hand-crafted features for each lung CT image slice which will be used in the classification step to decide whether a slice contains a lung lesion or not. Features extracted using this image processing based hand-crafted feature extraction method will be referred as F_HND from now on.

4.3 Lung Lesion Detection (Classification)

Next step is the detection of the lung lesions from the lung CT images. Detection is carried out by classifying each lung CT image slice into one of the two output classes; not bearing lung lesion or bearing lung lesion. To carry out such classification, conventional and simple artificial neural networks with an architecture containing one input layer, one hidden layer and one output layer are designed and tested. All these ANNs are trained with scaled conjugate gradient back-propagation algorithm.

As explained in the previous section, transferred AlexNet method extracts 4096 features from each input image slice. In this regard, the dimension of the feature vector is 4096. This number is clearly too much for the simple architectures of our conventional ANNs which increases computational load heavily. Also it is known that ANN based classifiers usually over-fit the training data when the input dimensions increase. To solve this problem principal component analysis (PCA) is applied to select and use only the significant features as an input to the classifiers which in turn reduces the dimension of the feature vector, thus reducing computational cost and risk of over-fitting.

PCA basically computes the principal components in the given data where each principal component is a linear transformation of the original variables. When these principal components are projected back on to the original data (our combined feature vector), original data is transformed in such a way that variance of the each observation`s (image slice in our case) transformed values in the new first dimension is the highest. Variance falls when it is moved towards the last dimension and it is lowest for the last dimension. So, arranged in decreasing order of variance. If the variances of the values in each dimension of this transformed data after the PCA are added together it gives the overall variance of the original data. In other words, PCA transforms and arranges 4096 dimensional feature vector in a way that, the feature values extracted from all input image slices (which are now transformed with principal components) regarding the first feature dimension have the highest variance/distinction so have the highest representative power (for each different image slice). When we go towards the 4096^{th} feature dimension of the new transformed data, feature values extracted from all the images regarding this feature dimension have the lowest variance meaning they are more similar or shared among all images giving them far less representative power. So, to reduce the dimensions of the original 4096 dimensional feature vector, less representative dimensions of the newly transformed and arranged data (by the PCA) can be discarded. Generally, when the added variances of the certain number of initial dimensions passes a determined percentage of the overall variance of the original data (it is assumed that the features in that chosen initial dimensions are representative enough for the whole data), all the remaining dimensions are decided as not necessary and discarded. Following this, by investigating the PCA output, it is decided to use only the first 200 dimensions of the PCA transformed feature vector for the classification step. As a result feature vector dimension is reduced dramatically from 4096 to 200. Thus, PCA transformed feature matrix for training becomes 7616x200 sized and feature matrix for testing becomes 5076x200 sized. Apart from AlexNet FC6 and FC7 methods, similar PCA based dimension reduction procedure is also carried out for the F_SHIFT and F_HND methods reducing them also into same dimensionality.

After the dimension reduction is completed, a separate ANN (four in total) is designed and trained for every different feature extraction method to be tested. First two networks are for evaluation purposes of our proposed AlexNet based methods. First network, N_SIFT, takes

the feature vectors extracted by the F_SIFT method from all the training input images. Similarly second network, N_HND, takes the feature vectors extracted by the F_HND method. On the other hand, third and fourth networks, N_FC6 and N_FC7, take their input feature vectors from the proposed automatic deep learning hierarchical feature extraction methods FC6 and FC7 respectively. Since output dimensions of all these feature extraction methods are reduced to 200 dimensions, number of input neurons in the input layer of all these networks is 200. Several different hidden layer neurons are experimented and best performing number are chosen for each network and fixed for final tests. Output neuron number is two for all networks. Output vectors for training of each network are same and prepared according to the input image class distribution provided by the ground truth information of the PLD dataset. Testing output vectors are also prepared similarly to compare the results of the testing and evaluate the detection performance of the classifier. Table 4.1 shows the output classes and their corresponding output vectors.

During training default settings of the Matlab neural network and pattern recognition toolbox is used. Feature vectors extracted from each feature extraction method is fed into its corresponding network along with the corresponding output vector and trained with back propagation algorithm until a desired error level or maximum epoch number is reached or training is terminated by manual observation. Also, 10-fold cross validation is performed during the training process and the results of all folds are averaged to form the final estimation. For the neurons sigmoid function is used so their outputs are always between values 0 and 1. An output neuron is considered to fired (output 1) when its output is greater than 0.60. Else it is considered to output 0. If both output neurons fire at the same time for an input image that image is considered as lung lesion bearing. If both output neurons do not fire it is considered as healthy or not bearing a lung lesion. In such cases where only one neuron fires it is considered as a un-classification.

Table 4.1: Output vector formation for each designed network. Output vectors, with different sizes, are formed for training and testing phases.

Lung Lesion Presence	**Corresponding Output Vector**
Non-bearing (n)	[0 0]
Lesion bearing (l)	[1 1]

Training performances are recorded after the trainings of all networks are completed. After the training all of the networks are tested with the testing images and their corresponding output vectors. Again testing performance of each network is recorded for the final evaluation.

All the design, coding, training and testing procedures along with the whole procedures regarding the proposed CAD system and its comparative methods are implemented using the Matlab program, with its built in image and signal processing, AlexNet, deep learning and neural network toolboxes. Used hardware involves Intel i5-6500 3.2 GHz CPU, NVidia GeForce GTX 1060 GPU and 16 GBs of RAM which runs on Window 10 operating system.

CHAPTER 5
RESULTS

In this chapter experimentation results obtained for the methods applied in previous Chapter 3 for PET respiratory motion correction and Chapter 4 for deep learning based CAD development for the detection of lung lesions are presented in detail.

5.1 Results for Respiratory Motion Correction

After the methods explained in the Chapter 3 are implemented, obtained final reconstructed PET images are presented as results. Those PET images obtained with different motion compensation techniques are compared visually for their accuracy of lung lesion presentation.

In the obtained reference image, it can be clearly observed that the introduced lesions in the lungs are reconstructed accurately. For the reference image, there is no need to apply any noise removal method since the noise here is expected to be minimal. However, for the un-gated reconstruction, since no motion compensation method is implemented here, lung lesions are blurry along with other surrounding tissues and are not localized accurately on the PET images. Figure 5.1 presents the control (reference) image and the un-gated images. Additionally, on some images lesions can be located inaccurately on surrounding organs like the diaphragm. Also, due to blurring effect, lung lesions that are not defined in the experimentation protocol can also be observed guiding the radiologist into miss-diagnosis. Figure 5.2 shows example images of such cases. When, the Figure 5.2 is investigated, on the left top transverse slice of the torso, the blurring effect due to respiratory motion causes an artifact which looks like a secondary lung lesion in the left lung, which is not true with regards to the set phantom parameters (refer to Figure 3.4). On the right bottom coronal slice of the torso it can be observed that some lesions are reconstructed over unrelated organs, which is again not true in terms of the set phantom parameters.

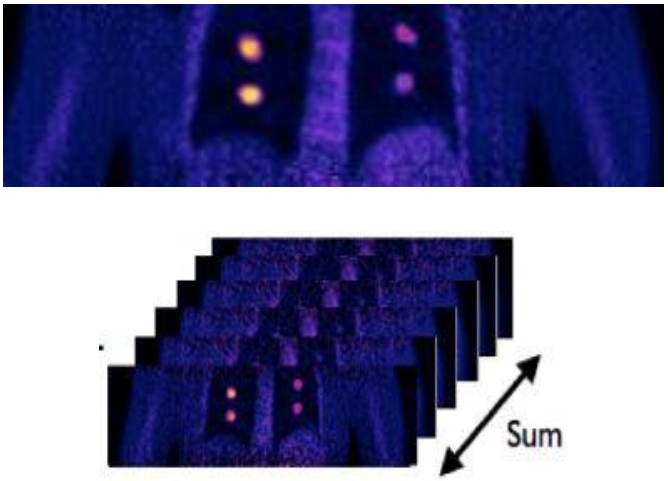


Figure 5.1: Reference image (top) and the un-gated images (bottom)

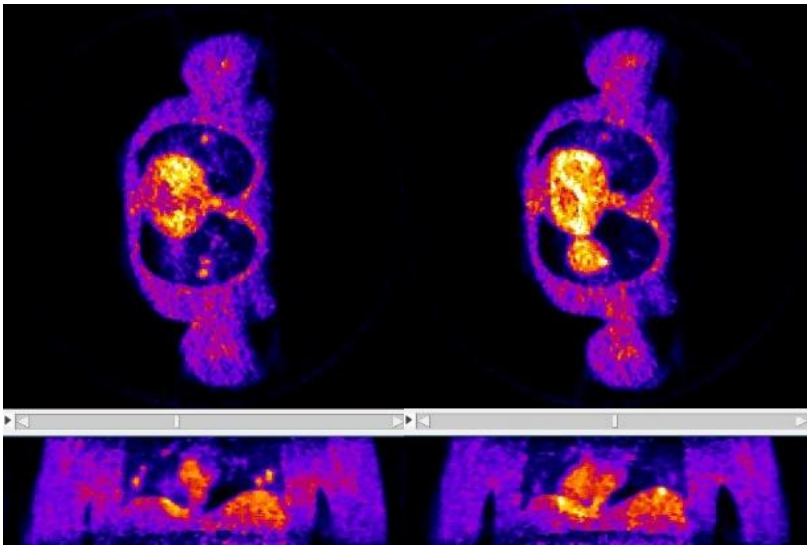


Figure 5.2: Sample PET images obtained after the un-gated reconstruction.

When the images obtained using the true motion fields are investigated, it can be seen that this method corrected the problems observed on the previous un-gated reconstruction method. Even though results obtained with this method are on par with the control image, still there are observable degradations in terms of image contrast and even in spatial resolution. It can be clearly seen from Figure 5.3 that lesions can be correctly spotted but

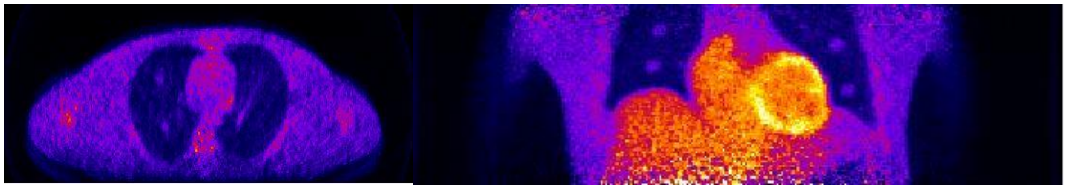


Figure 5.3: Sample results for true motion fields respiratory motion correction method

loss in contrast and resolution can cause confusion and miss-diagnosis.

Sample results obtained by using the other respiratory motion correction method, PET derived motion fields reconstruction method, are shown in Figure 5.4. If these results are compared to the results obtained with the true motion fields reconstruction method and to the reference image, it can be clearly observed that even though motion correction is applied there are still disturbing blurring effects present in the final reconstructed images. Although, this method still provided clear improvements regarding the overall PET image quality and lung lesion localization accuracy when compared to the un-gated image reconstruction.

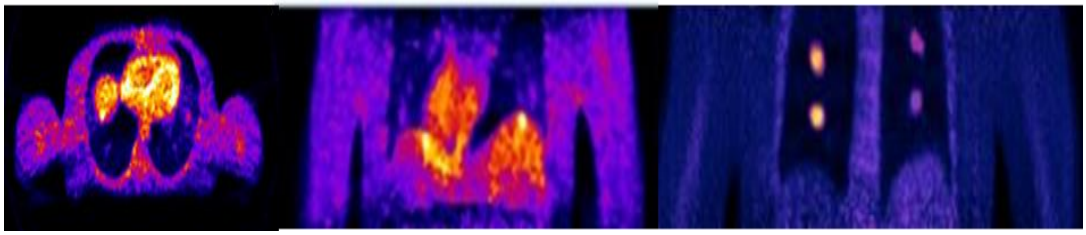


Figure 5.4: Sample images obtained after the reconstruction method with motion fields information derived from PET images

As the last respiratory motion correction technique, a reconstruction based technique that implements motion fields information obtained from MRI images is applied and the obtained results are presented in Figure 5.5. From these results it can be clearly concluded that MRI derived motion fields reconstruction method performed way better in correcting respiratory motion than the PET derived method. When the Figure 5.5 is investigated, lung lesions can be easily detected with high accuracy

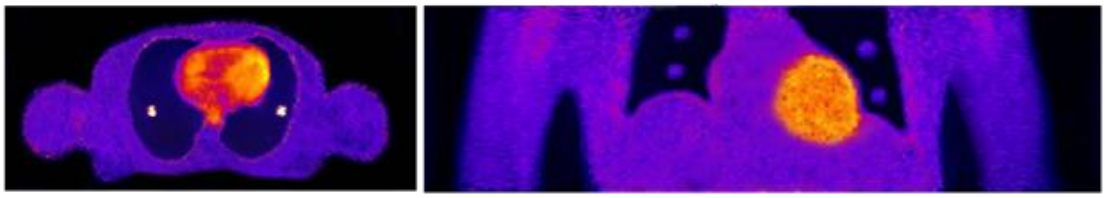


Figure 5.5: Sample results obtained after MRI derived motion fields reconstruction

without observing any noticeable blurring effect. Additionally, in terms of lesion localization, MRI derived motion fields reconstruction performed on par with the true motion fields reconstruction method. Moreover, it provides significantly better spatial and contrast resolution, thus providing improved structural and locational information which leads to better lung cancer diagnosis results.

5.2 Results for Lung Lesion CAD System

Initially, the lung CT image data from the PLD dataset is obtained, prepared and later fed into four different feature extraction methods. Dimensions of the features extracted from different methods are reduced to prevent over fitting by applying PCA. These features are then fed to their corresponding designed ANNs and trained to carry out classification of lung CT images into one of the possible classes; bearing or non-bearing a lesion.

After the training of the networks, they are tested using the testing dataset and the proposed deep learning based AlexNet feature extraction method (with its two variations) is evaluated and compared against the other two chosen test methods. Performances of the methods are evaluated using three main metrics namely; Sensitivity (Sn), Accuracy (Ac) and Specificity (Sp). They are defined by the following formulae;

$$Sn = \frac{T_p}{T_p + F_n + UN_p} \tag{5.1}$$

$$Sp = \frac{T_n}{T_n + F_p + UN_n} \tag{5.2}$$

$$Ac = \frac{T_p + T_n}{T_p + T_n + F_p + F_n + UN} \tag{5.3}$$

T_p gives the obtained true positive count, thus the correctly classified lung lesion bearing slices, T_n gives the obtained true negative count, and thus the number of non-bearing slices that are correctly classified by the method, F_p gives the obtained false positive count, thus non-bearing slices that are classified incorrectly as bearing slices. F_n is the obtained false negative count, thus lung lesion bearing slices that are classified incorrectly as non-bearing and finally *UN* is the total number of un-classifications, thus the slices that the classifier cannot assign a class label (where UN_p and UN_n are the un-classified lesion bearing and non-bearing image slices respectively). Basically, sensitivity provides a measure about the network's performance regarding the detection of the lung lesion bearing slices. On the other hand, specificity gives a measure about the performance of the network regarding the detection of the lung images that do not contain lung lesions. Accuracy, as the last measure provides information about the network's performance regarding the detection of the both cases. Together, all these three measures give an indication regarding the overall performance of the designed lung lesion CAD system. Experiments using each different method are repeated several times. Best results obtained during the performed experiments regarding each different tested method are reported. Table 5.1 shows the classification counts that each method obtained after testing with the 5076 lung CT image slices of the testing dataset. Further, Table 5.2 shows the obtained sensitivity, specificity and accuracy measures for each method, which are calculated using the count values in Table 5.1. Also Figure 5.6 illustrates those obtained results.

Table 5.1: Comparison of the obtained T_p, T_n, F_p, F_n and *UN* counts for each method in their best performing runs.

Network	**T_p**	**F_n**	**T_n**	**F_p**	**UN**		**Total Test Images**
					UN_p	**UN_n**	
N_HND	650	130	4002	285	2	7	5076
N_SIFT	695	85	4015	278	2	1	5076
N_FC6	765	12	4115	113	5	0	5076
N_FC7	772	5	4223	71	5	4	5076

Table 5.2: Comparison of the results obtained after testing the four different methods for lung lesion detection (results are recorded for the best runs). Values are given as percentages.

Network	Sn	Sp	Ac
N_HND	83.12	93.19	91.65
N_SIFT	88.87	93.50	92.79
N_FC6	97.83	95.83	96.14
N_FC7	**98.72**	**98.35**	**98.48**

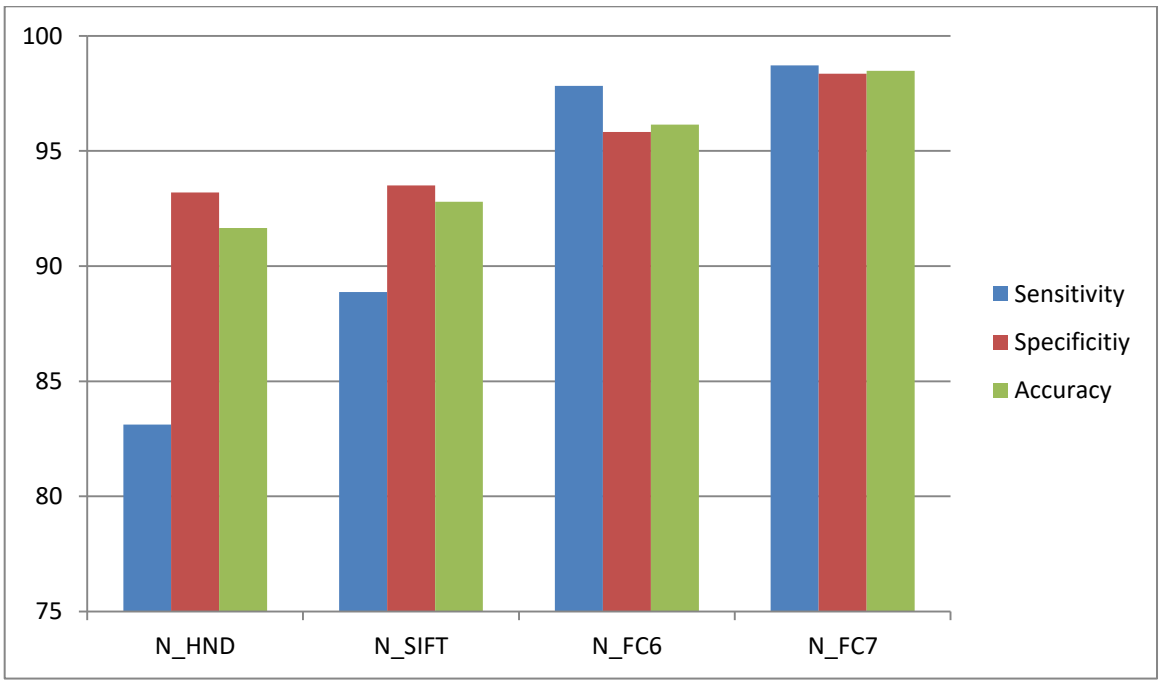


Figure 5.6: Graphical illustration of the obtained results.

CHAPTER 6
DISCUSSIONS

6.1 Discussions for Respiratory Motion Correction

In PET imaging, correction of the respiratory motion, along with other motions, is a very important factor which directly affects the quality of the obtained diagnostic PET images. Especially during the imaging of the torso region for cancer diagnosis, respiratory movement due to the natural respiration of the patient causes organs in the torso region to move, which in turn causes blurring in the functional images of the PET. Thus eliminating these motion artifacts is a necessity for the PET modality to obtain high quality images with high diagnostic capability. In this thesis as demonstrated in the results section, several methods regarding the motion correction have been developed in GATE simulation environment and tested effectively. Obtained results indicated an advantageous performance of the MRI derived motion fields method implemented at the image reconstruction phase for the respiratory motion correction. In the PET images reconstructed by this technique lung lesions can be detected easily without any noticeable blurring effect. On the other hand, other methods developed for comparison came with several errors as explained in the respective result section.

One of the main obstacles here is that, respiratory patterns of the imaged patient are not always regular. Our methods are developed using the XCAT phantom which considers respiratory patterns as regular. So simulating such cases and developing motion correction methods for them will be a future work. Additionally, muscular motions of the patient along with the motions of the beating heart also generate motion artifacts. Again in this work motion artifacts due to muscular and heart movements are not considered. Modifying the developed methods to consider these motion artifacts will also be a future aim.

Developing and testing reconstruction based respiratory motion correction techniques in a simulation environment was the main aim of the respiratory motion correction part of the

thesis. Although MRI motion fields based method obtained satisfying methods, developing improved motion field extraction algorithms along with the development of improved reconstruction methods in the future will improve the overall image quality of the PET devices providing an invaluable tool for the doctors to non-invasively diagnose lung cancers and cancer in general. Maybe deep learning methods will also have a future in PET motion fields extraction and image reconstruction methods.

6.2 Discussions for Lung Lesion CAD System

When the obtained results are investigated; it can be easily observed that the methods based on transferred AlexNet deep learning feature extraction outperformed the other methods with a very large margin, especially in terms of sensitivity. Since our task is classifying lung lesions from the CT images and since lesions have the possibility of being malignant, correctly classifying lung lesion bearing images are far more important than correctly classifying non-bearing images. Miss-classifying a cancerous lesion in a clinical routine can result in fatal consequences. On the other hand, if a non-bearing image is classified as bearing a lung lesion, this is a more tolerable situation and can be corrected by the radiologists while they re-check the detections.

Among the deep-learning based methods, N_FC7 which takes the automatic image features extracted by the 7^{th} fully connected layer of the AlexNet performed better than the N_FC6. This proves the idea that the deeper layers of a deep learning framework trained on a very large and generic dataset (like the CNN based AlexNet) can be robust enough to be easily transferred and implemented even for a very complex medical imaging task, like lung lesion detection.

Least performing method is the N_HND, which is based on hand-crafted features. As mentioned earlier, deciding what types of features to extract from the images and efficiently extracting those selected features is a complicated task and affects the overall performance of the CAD system. As expected, selected and hand-crafted features for N_HND method failed to represent lung CT images of our task effectively resulting in a non-acceptable detection performance. Especially obtained 83.12% sensitivity measure with 130 miss-classified lung lesion images (as non-bearing) is unacceptable even for

experimental purposes yet alone for clinical environment. On the other hand, N_SIFT method that is based on SIFT features, while performed better than the N_HND method still lagged way behind the deep learning based methods. As the best performing method, N_FC7 miss-classified total of 5 lung lesion bearing images out of total 782 lung lesion bearing test images. Further it un-classified 5 more. Un-classifications are more tolerable since they are not presented to the radiologist as non-bearing results. So radiologist can re-check all un-classified images (which are very few in number) and give their own decisions. Additionally, un-classifications are based on our chosen network design and these can be corrected with little modifications on the design. In the end, N_FC7 outperformed all other methods in all categories. With its very high performances of 98.72% sensitivity, 98.35% specificity and 98.48% accuracy, it proved to be an efficient method for developing a CAD system for assisting the radiologist in lung lesion detection tasks. However it should be noted that, even 5 miss-classifications can be considered relatively un-acceptable in clinical environment. Since this is not an ordinary image detection task, but a medical one regarding cancer, even if the developed methods miss-classify one single lung lesion bearing image it can cause fatal consequences for the patient. In this regard there are still room for improvements for this and whole medical CAD methods in general and at this stage they should only be considered as assistance systems for the radiologists to save their time and lower the operating costs instead of replacing the radiologist altogether.

CHAPTER 7
CONCLUSIONS

Currently, cancer is one of the most critical issues that wait a solution from the world of healthcare. Among all cancer types, lung cancer is the one that is responsible for the most deaths. In common clinical routine, CT imaging is the first step for the diagnosis of the lung lesions. Lung lesions can be malign or benign, where malign lesions develop into lung cancer. Deciding about the malignancy of the screened lung lesions is not an easy task using solely CT images. Generally patient needs to undergo several follow up CT scans to monitor the growth of the lesion in order to diagnose malignancy. Furthermore, in most cases invasive procedures like needle biopsy is also necessary to carry out pathological examination for achieving accurate diagnostics. All these procedures are not convenient for the undergoing patients. However, a more recent and advanced imaging modality, PET and its hybrid configurations of PET/CT and PET/MRI provide an invaluable solution in this regard. PET scanner can obtain metabolic functional images of the body which can enable non-invasive lung cancer diagnosis.

Although, PET imaging introduced many improvements into cancer imaging, still some areas of the modality require several improvements. PET imaging is also a very expensive modality in terms of the scanner cost and operating cost. Every patient scan requires expensive radiotracers to be injected to the patient. Additionally, these radiotracers are radioactive and extensive use of them causes harm to the patient. In this regard, it is not easy to develop and test methods on real patients to improve the PET modality. Due to this, many research efforts were carried out using powerful simulation environments like the GATE simulator. One of the areas regarding the PET scanners that need to be improved is the methods for compensating respiratory motion artifacts. PET image acquisition takes several minutes to complete. Therefore, unintentional and intentional motion of the patient during the PET scan of the torso for lung cancer imaging causes artifacts in the reconstructed image which leads to blurring of the PET image degrading the overall image quality. Therefore development of efficient motion correction methods for overcoming the respiratory motion artifacts thus improving the overall image quality is an interesting

research goal. In this regard, the first aim of this thesis is to develop and test different image reconstruction based motion correction techniques in a GATE simulation environment for the efficient correction of respiratory motion artifacts which occurs during the scanning of the torso for lung cancer imaging. To achieve this goal, first XCAT based computerized phantom of the torso with manually inserted lung lesions is developed to carry out the simulation on. After the phantom is generated, activity and attenuation files of the phantom are used in the GATE simulator to simulate sinograms regarding the activity distribution of the radiotracer inside the generated phantom torso. These sinograms are later used to reconstruct the final PET images by applying OSEM reconstruction algorithm with or without the developed motion correction methods. Results of these reconstructions are presented and compared visually. Obtained results demonstrate that, for obtaining a PET image with increased quality, there is a clear need for respiratory motion correction methods in PET scanners. Developed method that incorporates MRI derived respiratory motion fields information into the iterative reconstruction algorithm produced the best results with the highest image quality, which enhanced the quality of the final PET images with respect to the compared PET images without any applied respiratory motion correction technique.

Apart from the imaging modality, there are also other problems regarding the diagnosis of lung cancers. In current clinical routine, radiologists need to go over large numbers of lung image slices manually in order to detect and diagnose lung lesions. This process is very time consuming and its performance is very dependent on the performing radiologist. Even the same radiologist can obtain different diagnosis results when the diagnosis for the same case is repeated at different times. Additionally, variability of the diagnostic results between two different radiologists can be also very high. Thus assisting the radiologists in this complicated task of lung cancer detection by developing an automated computer aided detection (CAD) system is an interesting research goal. In this regard, as the second goal of this thesis a robust transfer deep learning (AlexNet) based CAD system is developed and implemented for the automatic detection of the lung lesions. However due to several reasons explained throughout the thesis developing such a system that operates on PET images is difficult at this stage. In this regard, a CAD system that operates on CT images is developed and tested. Unfortunately due to the inefficiency of the CT modality in

providing adequate information about the malignancy of the lung lesions, the proposed system tries to detect lung lesions in the CT images instead of detecting lung tumors, i.e. lung cancer.

To achieve this goal, first lung CT images (including lung lesions) are downloaded from the public PLD database. After the downloaded data is prepared for our task, proposed transfer deep learning based method (as two sub-methods) along with two non-deep learning based approaches is designed to extract representative features from the lung CT slices. After the feature extraction step is completed, high dimensions of the resulting feature vectors are reduced by applying the PCA method. As the final step, conventional ANN based classifiers are designed for each corresponding feature extraction method. These ANN classifiers are trained using the training dataset obtained from the PLD database and tested on the testing dataset again obtained from the same database. Obtained results are recorded and evaluated with respect to the metrics of accuracy, sensitivity and specificity. The proposed AlexNet based automatic feature extraction method obtained final performances of 98.72% sensitivity, 98.35% specificity and 98.48% accuracy. However other non-deep learning based methods which are developed and tested to evaluate the proposed method failed to achieve clinically acceptable results. Thus, it can be easily said that, an efficient CAD system which is based on an easily implementable state-of-the-art transfer deep learning automatic feature extraction method is successfully developed for the complex task of detection of lung lesions from CT images, which can provide invaluable help to the radiologists saving time and cost for the medical institutions.

CHAPTER 8
FUTURE WORK

The following issues can be considered as a future work regarding this thesis;

- Developing an MRI based motion correction method for PET imaging that includes correction strategies for muscular and cardiac motion of the patients along with the respiratory motion correction which is achieved in this thesis.
- Developing a method or in-house phantom to model irregular respiratory motion patterns of the real patients.
- Testing the proposed motion correction methods in real clinical environment.
- Implementing and testing the performance of other classifiers other than ANNs for the proposed CAD system
- Developing a CAD system which incorporates lung PET images instead of lung CT images.
- Depending on the previous point, developing a further diagnostic assistance system which can assist doctors regarding the malignancy of the detected lesions by the current proposed method.
- Depending on the previous two points, developing a system that can automatically segment and extract accurate tumor volumes from the PET images, which can assist doctors in clinical follow up and radiotherapy planning of the cancer patients.
- Developing an in-house deep learning framework from scratch based on a CNN or other methods.
- Depending on the previous point, comparing the performance of the in-house deep learning framework with the performances of the transfer deep learning method on this and other medical tasks.
- Applying current method onto other medical imaging tasks, like Parkinson's disease detection, Alzheimer's disease detection, brain tumor segmentation and so on.

REFERENCES

Abdulla S. (2017). PET Imaging. Retrieved February 22, 2018 from https://www.radiologycafe.com/radiology-trainees/frcr-physics-notes/pet-imaging

Aggarwal, N., & Agrawal, R. K. (2012). First and second order statistics features for classification of magnetic resonance brain images. *Journal of Signal and Information Processing*, *3*(02), 146.

Agostinelli, S., Allison, J., Amako, K. A., Apostolakis, J., Araujo, H., Arce, P., ... & Behner, F. (2003). GEANT4—a simulation toolkit. *Nuclear instruments and methods in physics research section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, *506*(3), 250-303.

Alberg, A. J., Ford, J. G., & Samet, J. M. (2007). Epidemiology of lung cancer: ACCP evidence-based clinical practice guidelines. *Chest*, *132*(3), 29S-55S.

Alberts, W. M. (2007). Diagnosis and management of lung cancer executive summary: ACCP evidence-based clinical practice guidelines. *Chest*, *132*(3), 1S-19S.

Antoch, G., Stattaus, J., Nemat, A. T., Marnitz, S., Beyer, T., Kuehl, H., ... & Freudenberg, L. S. (2003). Non–small cell lung cancer: dual-modality PET/CT in preoperative staging. *Radiology*, *229*(2), 526-533.

Appenzeller, P., Mader, C., Huellner, M. W., Schmidt, D., Schmid, D., Boss, A., ... & Veit-Haibach, P. (2013). PET/CT versus body coil PET/MRI: how low can you go?. *Insights into imaging*, *4*(4), 481-490.

Armstrong, P., Wastie, M., & Rockall, A. G. (2010). *Diagnostic imaging*. John Wiley & Sons.

Bai, W., & Brady, M. (2009). Regularized B-spline deformable registration for respiratory motion correction in PET images. *Physics in Medicine & Biology*, *54*(9), 2719.

Bailey, D. L., Townsend, D. W., Valk, P. E., & Maisey, M. N. (2005). *Positron emission tomography*. London: Springer.

Baker, D., Kilpatrick, J., & Chaudhry, A. (2017). Predicting Lung Cancer Incidence from CT Imagery. *Project Report, Stanford Uni*.

Barret, O., Carpenter, T. A., Clark, J. C., Ansorge, R. E., & Fryer, T. D. (2005). Monte Carlo simulation and scatter correction of the GE advance PET scanner with SimSET and Geant4. *Physics in Medicine & Biology*, *50*(20), 4823.

Bettinardi, V., Pagani, E., Gilardi, M., Alenius, S., Thielemans, K., Teras, M., & Fazio, F. (2002). Implementation and evaluation of a 3D one-step late reconstruction algorithm for 3D positron emission tomography brain studies using median root prior. *European journal of nuclear medicine and molecular imaging*, *29*(1), 7-18.

Beutel, J., Kundel, H. L., & Van Metter, R. L. (2000). *Handbook of medical imaging: Physics and psychophysics*(Vol. 1). Spie Press.

Beyer, T., Townsend, D. W., Brun, T., & Kinahan, P. E. (2000). A combined PET/CT scanner for clinical oncology. *The Journal of nuclear medicine*, *41*(8), 1369.

Blume, M., Martinez-Moller, A., Keil, A., Navab, N., & Rafecas, M. (2010). Joint reconstruction of image and motion in gated positron emission tomography. *IEEE transactions on medical imaging*, *29*(11), 1892-1906.

Camara, O., Delso, G., Colliot, O., Moreno-Ingelmo, A., & Bloch, I. (2007). Explicit incorporation of prior anatomical information into a nonrigid registration of thoracic and abdominal CT and 18-FDG whole-body emission PET images. *IEEE transactions on medical imaging*, *26*(2), 164-178.

Catana, C. (2015, May). Motion correction options in PET/MRI. In *Seminars in nuclear medicine* (Vol. 45, No. 3, pp. 212-223). Elsevier.

Chen, W. (2007). Clinical applications of PET in brain tumors. *Journal of nuclear medicine*, *48*(9), 1468-1481.

Chun, S. Y., Reese, T. G., Ouyang, J., Guerin, B., Catana, C., Zhu, X., ... & El Fakhri, G. (2012). MRI-based nonrigid motion correction in simultaneous PET/MRI. *Journal of Nuclear Medicine*, *53*(8), 1284-1291.

Dawood, M., Buther, F., Jiang, X., & Schafers, K. P. (2008). Respiratory motion correction in 3-D PET data with advanced optical flow algorithms. *Ieee transactions on medical imaging*, *27*(8), 1164-1175.

Dempster, A. P., Laird, N. M., & Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the royal statistical society. Series B (methodological)*, 1-38.

Deng, J., Dong, W., Socher, R., Li, L. J., Li, K., & Fei-Fei, L. (2009, June). Imagenet: A large-scale hierarchical image database. In *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on* (pp. 248-255). IEEE.

Devaraj, A., Cook, G. J. R., & Hansell, D. M. (2007). PET/CT in non-small cell lung cancer staging—promises and problems. *Clinical radiology*, *62*(2), 97-108.

Dineley J. (2012). Automated segmentation cuts plan errors. retrieved February 7, 2018 from http://medicalphysicsweb.org/cws/article/research/50874

Dutta, J., El Fakhri, G., Shao, X., Lorsakul, A., Guo, N., & Li, Q. (2014). Feasibility of respiratory motion compensated reconstruction using 4D PET-derived deformation fields. *Journal of Nuclear Medicine*, *55*(supplement 1), 2106-2106.

Enabling Ultra-Low Dose PET/CT Imaging. Retrieved February 7, 2018 from http://depts.washington.edu/imreslab/currentResearch.html

Erasmus, J. J., Connolly, J. E., McAdams, H. P., & Roggli, V. L. (2000). Solitary pulmonary nodules: Part I. Morphologic evaluation for differentiation of benign and malignant lesions. *Radiographics*, *20*(1), 43-58.

Erdi, Y. E., Nehmeh, S. A., Pan, T., Pevsner, A., Rosenzweig, K. E., Mageras, G., ... & Vernon, P. (2004). The CT motion quantitation of lung lesions and its impact on PET-measured SUVs. *Journal of Nuclear Medicine*, *45*(8), 1287-1292.

Erdi, Y. E., Rosenzweig, K., Erdi, A. K., Macapinlac, H. A., Hu, Y. C., Braban, L. E., ... & Yorke, E. D. (2002). Radiotherapy treatment planning for patients with non-small

cell lung cancer using positron emission tomography (PET). *Radiotherapy and Oncology*, *62*(1), 51-60.

Feng, M., Kong, F. M., Gross, M., Fernando, S., Hayman, J. A., & Ten Haken, R. K. (2009). Using fluorodeoxyglucose positron emission tomography to assess tumor volume during radiotherapy for non–small-cell lung cancer and its potential impact on adaptive dose escalation and normal tissue sparing. *International Journal of Radiation Oncology• Biology• Physics*, *73*(4), 1228-1234.

Ford, E. C., Herman, J., Yorke, E., & Wahl, R. L. (2009). 18F-FDG PET/CT for image-guided and intensity-modulated radiotherapy. *Journal of Nuclear Medicine*, *50*(10), 1655-1665.

Fulton, R. R., Meikle, S. R., Eberl, S., Pfeiffer, J., & Constable, C. J. (2002). Correction for head movements in positron emission tomography using an optical motion-tracking system. *IEEE Transactions on Nuclear Science*, *49*(1), 116-123.

Gaillard F. (2018). Retrieved February 10, 2018 from https://radiopaedia.org/cases/normal-chest-x-ray

Gee, A. D. (2003). Neuropharmacology and drug development: imaging in clinical neuroscience. *British medical bulletin*, *65*(1), 169-177.

Grimm, R., Fürst, S., Dregely, I., Forman, C., Hutter, J. M., Ziegler, S. I., ... & Block, T. (2013, September). Self-gated radial MRI for respiratory motion compensation on hybrid PET/MR systems. In *International Conference on Medical Image Computing and Computer-Assisted Intervention* (pp. 17-24). Springer, Berlin, Heidelberg.

Guerin, B., Cho, S., Chun, S. Y., Zhu, X., Alpert, N. M., El Fakhri, G., ... & Catana, C. (2011). Nonrigid PET motion compensation in the lower abdomen using simultaneous tagged-MRI and PET imaging. *Medical physics*, *38*(6Part1), 3025-3038.

Hanzouli-Ben Salah, H., Lapuyade-Lahorgue, J., Bert, J., Benoit, D., Lambin, P., Baardwijk, A., ... & Hatt, M. (2017). A framework based on hidden Markov trees for multimodal PET/CT image co-segmentation. *Medical physics*, *44*(11), 5835-5848.

Hashemi, R. H., Bradley Jr, W. G., & Lisanti, C. J. (2004). *MRI The Basics. Philadelphia, PA: Lippincott Williams & Wilkins*.

Havaei, M., Davy, A., Warde-Farley, D., Biard, A., Courville, A., Bengio, Y., ... & Larochelle, H. (2017). Brain tumor segmentation with deep neural networks. *Medical image analysis*, *35*, 18-31.

Hudson, H. M., & Larkin, R. S. (1994). Accelerated image reconstruction using ordered subsets of projection data. *IEEE transactions on medical imaging*, *13*(4), 601-609.

Işın, A., & Ozdalili, S. (2017). Cardiac arrhythmia detection using deep learning. *Procedia Computer Science*, *120*, 268-275.

Işın, A., Direkoğlu, C., & Şah, M. (2016). Review of MRI-based brain tumor image segmentation using deep learning methods. *Procedia Computer Science*, *102*, 317-324.

Işın, A., Ozsahin, D. U., Dutta, J., Haddani, S., & El-Fakhri, G. (2017). Monte Carlo simulation of PET/MR scanner and assessment of motion correction strategies. *Journal of Instrumentation*, *12*(03), C03089.

Jan, S., Santin, G., Strul, D., Staelens, S., Assie, K., Autret, D., ... & Brasse, D. (2004). GATE: a simulation toolkit for PET and SPECT. *Physics in Medicine & Biology*, *49*(19), 4543.

Kak A. C., & Slaney, M. (1988). *Principles of computerized tomographic imaging* (pp. 203-274). IEEE press.

Khalil, M. (Ed.). (2010). *Basic sciences of nuclear medicine*. Springer Science & Business Media.

Kopriva, I., Ju, W., Zhang, B., Shi, F., Xiang, D., Yu, K., ... & Chen, X. (2017). Single-Channel Sparse Non-Negative Blind Source Separation Method for Automatic 3-D Delineation of Lung Tumor in PET Images. *IEEE journal of biomedical and health informatics*, *21*(6), 1656-1666.

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems* (pp. 1097-1105).

Lalush, D. S., & Tsui, B. M. (2000). Performance of ordered-subset reconstruction algorithms under conditions of extreme attenuation and truncation in myocardial SPECT. *The Journal of Nuclear Medicine*, *41*(4), 737.

Lamare, F., Carbayo, M. L., Cresson, T., Kontaxakis, G., Santos, A., Le Rest, C. C., ... & Visvikis, D. (2007). List-mode-based reconstruction for respiratory motion correction in PET using non-rigid body transformations. *Physics in Medicine & Biology*, *52*(17), 5187.

Lamare, F., Le Maitre, A., Dawood, M., Schäfers, K. P., Fernandez, P., Rimoldi, O. E., & Visvikis, D. (2014). Evaluation of respiratory and cardiac motion correction schemes in dual gated PET/CT cardiac imaging. *Medical physics*, *41*(7).

Leahy, R. M., & Qi, J. (2000). Statistical approaches in quantitative positron emission tomography. *Statistics and Computing*, *10*(2), 147-165.

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *nature*, *521*(7553), 436.

Li, T., Thorndyke, B., Schreibmann, E., Yang, Y., & Xing, L. (2006). Model-based image reconstruction for four-dimensional PET. *Medical physics*, *33*(5), 1288-1298.

Lin, E. C., & Alavi, A. (2009). PET and PET/CT: a clinical guide.

Liu, F., Velikina, J. V., Block, W. F., Kijowski, R., & Samsonov, A. A. (2017). Fast realistic MRI simulations based on generalized multi-pool exchange tissue model. *IEEE transactions on medical imaging*, *36*(2), 527-537.

Livieratos, L., Stegger, L., Bloomfield, P. M., Schafers, K., Bailey, D. L., & Camici, P. G. (2005). Rigid-body transformation of list-mode projection data for respiratory motion correction in cardiac PET. *Physics in Medicine & Biology*, *50*(14), 3313.

Lowe, D. G. (2004). Distinctive image features from scale-invariant keypoints. *International journal of computer vision*, *60*(2), 91-110.

Lung Cancer: The Leading Cause of Cancer Death. Retrieved February 7, 2018 from https://www.onhealth.com/content/1/lung_cancer

MacManus, M., Nestle, U., Rosenzweig, K. E., Carrio, I., Messa, C., Belohlavek, O., ... & Watanabe, N. (2009). Use of PET and PET/CT for radiation therapy planning: IAEA expert report 2006–2007. *Radiotherapy and oncology*, *91*(1), 85-94.

Mahowald, M. K., Aswad, B. I., Okereke, I. C., & Ng, T. (2015). Long-term survival after pneumonectomy for primary pulmonary malignant melanoma. *The Annals of thoracic surgery*, *99*(4), 1428-1430.

Manber, R., Thielemans, K., Hutton, B. F., Wan, S., McClelland, J., Barnes, A., ... & Atkinson, D. (2016). Joint PET-MR respiratory motion models for clinical PET motion correction. *Physics in Medicine & Biology*, *61*(17), 6515.

McClelland, J. R., Hawkes, D. J., Schaeffter, T., & King, A. P. (2013). Respiratory motion models: A review. *Medical image analysis*, *17*(1), 19-42.

Mountain, C. F. (1997). Revisions in the international system for staging lung cancer. *Chest*, *111*(6), 1710-1717.

MRI vs. CT Scan Which One Is Better for Brain Imaging. Retrieved February 10, 2018 from https.//healthcareplex.com/mri-vs-ct-scan/

Nehmeh, S. A., Erdi, Y. E., Ling, C. C., Rosenzweig, K. E., Squire, O. D., Braban, L. E., ... & Humm, J. L. (2002). Effect of respiratory gating on reducing lung motion artifacts in PET imaging of lung cancer. *Medical physics*, *29*(3), 366-371.

Nielsen M. (2017). Why are deep neural networks hard to train?. Retrieved March 10, 2018 from http://neuralnetworksanddeeplearning.com/chap5.html

Other Imaging Modalities. Retrieved February 7, 2018 from https://pocketdentistry.com/14-other-imaging-modalities/

Ozturk, C., Derbyshire, J. A., & McVeigh, E. R. M. (2003). Estimating motion from MRI data. *Proceedings of the IEEE*, *91*(10), 1627-1648.

Pennec, X., Cachier, P., & Ayache, N. (1999, September). Understanding the "demon's algorithm": 3D non-rigid registration by gradient descent. In *International Conference on Medical Image Computing and Computer-Assisted Intervention* (pp. 597-605). Springer, Berlin, Heidelberg.

Physics of the X-ray Tube. Retrieved March 18, 2018 from http://www.wikiradiography.net/page/Physics+of+the+X-Ray+Tube

Projections of mortality and burden of disease, 2002-2030. Retrieved February 10, 2017 from http://www.who.int/healthinfo/global_burden_disease/projections2002/en/

Rahmim, A., Rousset, O., & Zaidi, H. (2007). Strategies for motion tracking and correction in PET. *PET clinics*, *2*(2), 251-266.

Reeves, A. P., Biancardi, A. M., Yankelevitz, D., Fotin, S., Keller, B. M., Jirapatnakul, A., & Lee, J. (2009, September). A public image database to support research in computer aided diagnosis. In *Engineering in Medicine and Biology Society, 2009. EMBC 2009. Annual International Conference of the IEEE* (pp. 3715-3718). IEEE.

Rohren, E. M., Turkington, T. G., & Coleman, R. E. (2004). Clinical applications of PET in oncology. *Radiology*, *231*(2), 305-332.

Saha, G. B. (2015). *Basics of PET imaging: physics, chemistry, and regulations*. Springer.

Sánchez, J., Perronnin, F., Mensink, T., & Verbeek, J. (2013). Image classification with the fisher vector: Theory and practice. *International journal of computer vision*, *105*(3), 222-245.

Segars, W. P., Lalush, D. S., & Tsui, B. M. (1999). A realistic spline-based dynamic heart phantom. *IEEE Transactions on Nuclear Science*, *46*(3), 503-506.

Segars, W. P., Sturgeon, G., Mendonca, S., Grimes, J., & Tsui, B. M. (2010). 4D XCAT phantom for multimodality imaging research. *Medical physics*, *37*(9), 4902-4915.

Strobel, K., Rüdy, M., Treyer, V., Veit-Haibach, P., Burger, C., & Hany, T. F. (2007). Objective and subjective comparison of standard 2-D and fully 3-D reconstructed data on a PET/CT system. *Nuclear medicine communications*, *28*(7), 555-559.

Tajbakhsh, N., Shin, J. Y., Gurudu, S. R., Hurst, R. T., Kendall, C. B., Gotway, M. B., & Liang, J. (2016). Convolutional neural networks for medical image analysis: Full training or fine tuning?. *IEEE transactions on medical imaging*, *35*(5), 1299-1312.

Townsend, D. W. (2008). Dual-modality imaging: combining anatomy and function. *Journal of Nuclear Medicine*, *49*(6), 938-955.

Townsend, D. W. (2008). Dual-modality imaging: combining anatomy and function. *Journal of Nuclear Medicine*, *49*(6), 938-955.

Wang, X. Y., Zhao, Y. F., Liu, Y., Yang, Y. K., Zhu, Z., & Wu, N. (2017). Comparison of different automated lesion delineation methods for metabolic tumor volume of 18F-FDG PET/CT in patients with stage I lung adenocarcinoma. *Medicine*, *96*(51), e9365.

Würslin, C., Schmidt, H., Martirosian, P., Brendle, C., Boss, A., Schwenzer, N. F., & Stegger, L. (2013). Respiratory motion correction in oncologic PET using T1-weighted MR imaging on a simultaneous whole-body PET/MR system. *Journal of Nuclear Medicine*, *54*(3), 464-471.

XCAT. Retrieved February 7, 2016 from https://olv.duke.edu/xcat

Yosinski, J., Clune, J., Bengio, Y., & Lipson, H. (2014). How transferable are features in deep neural networks?. In *Advances in neural information processing systems* (pp. 3320-3328).

Yu, Y., Chan, C., Ma, T., Liu, Y., Gallezot, J. D., Naganawa, M., ... & Liu, C. (2016). Event-by-event continuous respiratory motion correction for dynamic PET imaging. *Journal of Nuclear Medicine*, *57*(7), 1084-1090.

APPENDIX

CURRICULUM VITAE

ALİ IŞIN

Title: Ph.D. Cand.
Date of Birth: 20/11/1985
Email: ali.isin@neu.edu.tr, aliisin@hotmail.com
Web: http://staff.neu.edu.tr/ali.isin/
Mobile: +90 533 832 54 28

EDUCATION

Degree	Field	Institution	Date
Ph.D.	Biomedical Engineering	Near East University/ Nicosia, TRNC	2014-Present
M.Sc.	Biomedical Engineering	Near East University/ Nicosia, TRNC	2013
B.Eng.	Biomedical Engineering	Başkent University/ Ankara, Turkey	2006

RESEARCH INTERESTS

- Biomedical Imaging, Biomedical Image Processing, Biomedical Signal Processing, Machine Learning, Deep Learning.

ACADEMIC EXPERIENCE

1. **Part Time Lecturer**,
   Biomedical Engineering Department, Near East University,
   2011 – present.

2. **Supervisor,**
   11 Undergraduate Graduation Projects (Biomedical Engineering)

COURSES TAUGHT

- **Undergraduate: (2011 – Present)**
  - Biomedical Signal Processing

-Biomedical Sensors
-Biomedical Instrumentation I & II
-Biomedical Imaging
-Bioinformatics
-Introduction to Biomedical Engineering
-Medical Device Technology
-Medical Devices Maintenance Techniques

PUBLICATIONS

- **Journals:**

1. ***Işın A., Ozdalili S., "***Cardiac arrhythmia detection using deep learning", Procedia Computer Science 120, 268-275, Elsevier, 2017.

2. ***Nurcin F. V., Imanov E., Işın A., Ozsahin D. U.,*** "Lie detection on pupil size by back propagation neural network", Procedia Computer Science 120, 417-421, Elsevier, 2017.

3. ***Işın A., Ozsahin D. U., Dutta j., Haddani S., El-Fakhri G.,*** "Monte Carlo simulation of PET/MR scanner and assessment of motion correction strategies", Journal of Instrumentation 12, C03089, IOP Publishing, 2017.

4. ***Işın A., Direkoğlu C., Şah M.,*** "Review of MRI-based brain tumor image segmentation using deep learning methods", Procedia Computer Science 102, 317-324, Elsevier, 2016.

- **International Symposiums and Conferences:**

5. ***Işın A., Ozdalili S., "***Cardiac arrhythmia detection using deep learning", 9th International Conference on Theory and Application of Soft Computing, Computing with Words and Perception, ICSCCW 2017, 24-25 August 2017, Budapest, Hungary.

6. ***Nurcin F. V., Imanov E., Işın A., Ozsahin D. U.,*** "Lie detection on pupil size by back propagation neural network", 9th International Conference on Theory and Application of Soft Computing, Computing with Words and Perception, ICSCCW 2017, 24-25 August 2017, Budapest, Hungary.

7. ***Işın A., Direkoğlu C., Şah M.,*** "Review of MRI-based brain tumor image segmentation using deep learning methods", 12th International Conference on Application of Fuzzy Systems and Soft Computing, ICAFS 2016, 29- 30 August 2016, Vienna, Austria.

8. ***Işın A.,*** "Intelligent Recognition and Classification of Three Cardiac Conditions Using ECG Signals", IBMEC'15, International Biomedical Engineering Congress 2015, 12-14 March 2015, Nicosia, TRNC.

- **Published Proceeding Papers :**

9. ***Işın A.,*** " Intelligent Recognition and Classification of Three Cardiac Conditions Using ECG Signals", IBMEC 2015, BIS01, Page 48, 2015

10. ***Nurçin F. V., Işın A., Imanov E.,*** " Lie Detection on Pupil Size" IBMEC 2015, P04, Page 68, 2015

- **Books and Book Chapters:**

11. **Book:** "Using Neural Networks for the Recognition of Cardiac ECG Signals: Neural networks and ECG Recognition", Ali IŞIN, LAP LAMBERT Academic Publishing, May 6, 2013; **ISBN-10:** 3659391646, **ISBN-13:** 978-3659391644

ACADEMIC PROJECTS

1. "Motion correction strategies in PET/MRI scanners and development of a deep learning based CAD system", Nicosia, 2016-present

2. "Segmentation and semantic annotation of brain tumors from mr-images using Deep Learning", Nicosia, 2016- present

3. "Intelligent Recognition and Classification of Three Cardiac Conditions Using ECG Signals", Nicosia, TRNC, 2013

4. "Assistant Robot Development for Disabled Patients using Pioneer II Robot", University of Essex, Colchester, UK, 2007

5. "Brain Computer Interface Controlled Electric Wheel Chair", University of Essex, Colchester, UK, 2007

6. "Design of an ECG Acquisition and Processing Software in Matlab", Başkent University, Ankara, Turkey, 2006

PROFESSIONAL EXPERIENCE

1. **Senior Biomedical Engineer and Medical Devices Procurement & Tendering Expert**, Medical Devices Department, Ministry of Health, Nicosia, TRNC, 2011– Present

2. **Biomedical Engineer,** Electronics Department, Nicosia State Hospital, Nicosia, TRNC, 2009-2011

PROFESSIONAL PROJECTS

- **Directed Professional Projects (High Budget Projects):**

1. "Digitalization of Radiography Devices in Northern Cyprus State Hospitals (Project Cost: 1.000.000,00 USD "

Project Engineer, Nicosia, TRNC, 2017.

2. "Dr. Burhan Nalbantoğlu State Hospital Radiation Oncology Center Project (including installation of a Varian Linear Accelerator System) (Project Cost: 9.000.000,00 USD)", Project Engineer (for medical devices), Nicosia, TRNC, 2012 – 2016.

3. "Kyrenia State Hospital Hemodialysis Center Project (Project Cost: 350.000,00 USD)" Project Engineer, Kyrenia, TRNC, 2014

4. "Procurement and Installation of MRI devices to Nicosia and Famagusta State Hospitals (Project Cost: 2.000.000,00 USD)", Project Engineer, Nicosia/Famagusta, TRNC, 2013

5. "Procurement and Installation of Computed Tomography Devices to Nicosia and Famagusta State Hospitals (Project Cost: 1.200.000,00 USD)", Project Engineer, Nicosia/Famagusta, TRNC, 2012

6. "Procurement and Installation of a Digital Angiography System to Nicosia State Hospital (Project Cost: 800.000,00 USD)", Project Engineer, Nicosia TRNC, 2011

- **Directed Low Budget Projects:**

7. More than 50 low budget projects.

SCIENTIFIC / PROFESSIONAL SOCIETY and OTHER COMMITTEE MEMBERSHIPS

1. Member, International Association of Engineers "IAENG", Member Number: 179661, 11/2016- present

2. Representative of the Ministry of Health (responsible for Medical Devices Market), Market Inspection Committee under "Product Safety Law", 2014 – present.

3. Representative of the Ministry of Health (responsible from Medical Devices Working Group), "Free Movement of Goods in Northern Cyprus" mission of the European Commission, 2013 – present.

4. Representative of the Ministry of Health (responsible from Medical Devices Working Group), Official Committee responsible for the drafting of the legislation on "Product Safety", 2013 – present.

5. Organizing Committee Member, "IBMEC'15 International Biomedical Engineering Congress 2015", 12-14 March 2015, Near East University, Nicosia, TRNC,.

AWARDS

- "Young Researcher Award",
  Near East University`s "Science Awards in 2017", Nicosia, 2018

- "Graduated as First in Department",
  Biomedical Engineering Department, Faculty of Engineering, Başkent University, Ankara, Turkey, 2006

RECEIVED PROFESSIONAL EDUCATIONS

1. Technical, Expert and User educations, on many various medical devices ranging from Linac, MRI, CT, USG, X-Ray Devices to Surgical Instruments, Diagnostic Tools and Endoscopic-Ultrasonography,
   received directly from corresponding manufacturers/distributors between 2011-2017

2. Biomedical Calibration Education,
   received from Turkish Standards Institute, 2014

3. Series of educations and conferences on "Market Surveillance",
   received directly from European Union and European Commission between 2014-2018

4. Series of educations and conferences on "European Union's Public Tendering Norms",
   received directly from European Union and European Commission between 2013-2014

SKILLS AND ACTIVITIES

- Proficient Programming Languages: Matlab, C/C++

- Other Computer Languages: XML, Rdf, Rdfs, OWL, SPARQL, Python, Tensor Flow

- Proficient Written and Spoken Languages: Turkish, English.

- Hobbies: Football, Computer Games, Spearfishing, Sci-fi.