Introduction

The following literature review provides insight into the existing works involving the medical imaging and machine learning, dose-related noise in CT and how these have been combined. It begins by providing context in lung functional analysis with both deterministic and data-driven methods. The description, simulation, and removal of noise in X-Ray and Computed Tomography are then discussed. Gaps in this literature are then stated, and the relevant end-to-end architectures are reviewed.

Automatic Lung Functional Analysis

Automatic Lung Functional Analysis is a well-established area of study. With the recent explosion of applied machine learning research, the area has seen significant progress. [3] is an excellent example of this. It analyses a few hundred papers that used machine learning to diagnose COVID-19 in 2021. The paper [3] focusses on investigating the clinical relevance of the reviewed research papers, finding that none of them were clinically applicable for various reasons. It also provides common pitfalls and recommendations going forward specifically in the realm of automatic lung functional analysis. This is particularly relevant, in terms of how the model designed in this thesis will be developed and assessed. The number of papers reviewed in [3] demonstrates how current the study area is, and thus how important it is to critically investigate the metrics and methods used.

Generally, the approaches to this type of problem can be categorized into 1) deterministic and 2) data-driven methods. The term “deterministic” is used to refer to direct, computational methods such as those in classical computer vision, where a given input image will always provide the same output, while “data-driven” refers to models involving machine-learning-type models built from training data.

Deterministic Methods

Earlier models, such as those described in [4] and [5] provide examples of deterministic methods for lung segmentation and analysis. While [5] is more focused on the segmentation step, [4] demonstrates the effectiveness on the end goal – the task-based accuracy for the classification of lung nodules. Without any machine learning techniques, and a very limited dataset (only 38 scans), the method was quite successful, failing on only 4.9% of the test set. [5] describes the development of an automatic method for the segmentation of 3-Dimensional CT reconstructions. It introduces optimal (dynamic) thresholding and used dynamic programming to identify anterior and posterior junction lines. The paper provides a useful in-depth statistical analysis of the results that will be very useful when informing the assessment of the results of this thesis. Both [4] and [5] were relatively successful, without the use of a data-driven approach. Deterministic methods can thus contribute significantly to the end-to-end system developed here.

Data-driven Models

Modern papers are increasingly focused on the development of data-driven, machine learning models. [2] compares the use of pre-trained deep-learning image models with training from scratch for several medical imaging modalities and datasets. The results showed that the pre-trained models usually outperformed those built from scratch, particularly when the training dataset decreased in size. Deep and shallow fine-tuning was also compared, and overall, the paper shows just how useful pre-trained models can be when combined and trained with task-specific data. [6] is an example of a 3D network architecture used for lung nodule detection, similar to the popular U-net architecture [9]. [6] also discusses and compares other popular machine learning architectures and tools, including the pre-trained models mentioned in [2] used for lung functional analysis. The 3D type of architecture in [6], being popular and current, will be relevantly applied in this study. Furthermore, [6] provides a valuable, robust statistical argument to base our assessment from.

Dose and Image Quality

The dose vs image quality tradeoff is another significant aspect of study in Computed Tomography and Medical Imaging research. As described in [7] and [8], higher doses are linked to higher image quality, but also more risk to the patient, while low doses are better for the patient, but result in much more noise. The ALARA principle (As Low As Reasonably Achievable) describes the goal of gaining as much information as possible with as low radiation dose as possible. [7] details the sources and effects of image quality issues in low dose Computed Tomography as well as the effect of CT exposure on humans. Understanding the physical aspects of this relationship is necessary for the accurate simulation of low dose imaging in this thesis. [8] provides an alternative perspective, providing a more practical description of how CT scanner dose settings (included in the CT DICOM file headers) affect patients medically and the effective dose patients experience as a result. The nuances of CTDI, Effective Dose and Radiation Energy transfer is described, clarifying the concept of “low dose”. Furthermore, [8] details several metrics that are used to describe CT image quality, including Contrast-to-Noise Ratio (CNR) which is not mentioned in [7]. The description and effects of this tradeoff justifies the need to improve the modelling of low dose CT images.

**Dose-related Noise and Denoising**

Introduction

There are several ways to categorize the variety of denoising methods that have been researched for CT imaging. In this review, these have been broadly divided into 1) Projection-domain techniques, 2) Traditional Post-processing techniques and 3) Deep Learning (DL) techniques. Since many publications have investigated denoising methods, three reviews [10, 11, 12] were used to frame the discussion of denoising, with more detail being provided on several important denoising algorithms. [10] focusses on medical image noise and denoising in general, describing several modalities, including computed tomography. [11] provides a more in-depth survey of traditional CT image denoising techniques. [12] is also focused on CT image denoising, however, it primarily discusses Deep Learning approaches and relevant issues. All three reviews at least briefly mention the traditional denoising methods, and metrics used for assessment of these algorithms.

[11] broadly describes four types of noise present in CT scans. These include random noise, statistical noise (also known as quantum noise), electronic noise and roundoff errors. As described in [10], quantum noise results from statistical fluctuations of the X-ray quanta. [11] explains that if fewer quanta are detected in a measurement, the accuracy of the measurement diminishes leading to “quantum noise” in the resulting CT Image. To reduce the effects of this, one can increase the X-ray dose, thereby increasing the number of detected X-ray quanta, Of course, this also introduces increased radiation and risk to the patient. The alternative is to apply a variety of denoising techniques to the resulting noisy images. This review will focus on the effects of statistical (quantum) noise.

(figure visually showing effects of Low-Dose Noise)

The noise distribution present in CT images varies depending on several factors including the type of machine and the reconstruction algorithm used [11] and can be accurately modelled as Poisson or Additive Gaussian white noise (in the case of multi-detector CT scanners). In general, it can be modelled using a Gaussian distribution [10,11,12], thanks to an application of the Central Limit Theorem [14].

[10] provides describes four major requirements of a good denoising algorithm to be the preservation of edges, the maintaining of structural similarity, low complexity (computationally efficient) and not relying too heavily on prior databases. Over time, and for different applications, denoising algorithms achieve these to different degrees. The following section describes some of the most popular and prominent methods.

Projection-domain techniques

As described in [10, 11], CT images are taken by rotating X-ray emitter and detector pairs around a subject. The resulting measurements produce a sinogram consisting of stacked histograms of detected X-ray values for each angle around a patient. The physical process is similar to a radon transform, and through a process called back-projection, the spatial domain image is retrieved with the inverse radon transform. Projection-domain techniques are those that are applied in the Sinogram domain to suppress noise. Without the use of any processing before back-projection is applied, image is often blurry. Filtered Back-projection [15] was the first solution to this, which involved taking the two-dimensional discrete fourier transform of the sinogram and applying a ramp high-pass filter to it to remove the blur. From there, Iterative Reconstruction (IR) methods were developed [16], including Compressed Sensing [10]. IR essentially extends FBP by applying it as an iterative algorithm. It produces better quality images, particularly in cases with low-dose and/or sparse data. Understanding the algorithms in this step are useful in understanding where the output images come from and how noise propagates through several steps before being produced in the CT image.

(image of FBP process)

Traditional post-processing techniques

Traditional post-processing techniques include a variety of deterministic algorithms applied to the CT image in the spatial domain (once the inverse radon transform has been applied to the sinogram). Traditional smoothing filters smoothing filters typically update a pixel’s intensity value to be a weighted average of just the surrounding pixel intensities. Non-local Means (NLM) [17] however, computes every pixel as a weighted average of the pixels in the entire image. Each weight is calculated based on the statistical similarity and L2 distance between the given pixel’s surrounding “neighborhood” window and the other pixel’s surrounding “neighborhood” window. The result is efficient and effective denoising while preserving important clinical information [11]. NLM has the drawbacks of discarding small details and requiring a high operation time. To deal with this, the Total Variation (TV) method, originally described in [18] was developed. It works by minimizing the “Total Variation” function of the image, smoothing out noise, but maintaining edges by using a regularization term. With time, the method was further developed and improved upon [11]. Principles of NLM have been incorporated into TV in the probabilistic NLTV (PNLTV) method, which result in keeping the fine details unchanged, but keeping the denoised images sharp and smooth.

The wavelet transform [19] is a useful tool in CT image denoising. The wavelet-based denoising method essentially works by first choosing an appropriate wavelet basis and applying the wavelet transform on the image. The noise variance is then estimated, and “thresholding” is applied. Finally, the inverse wavelet transform is applied to this to obtain the noise-suppressed image [11]. Thresholding refers to the operation of updating wavelet coefficients where their values are below a certain threshold. Several papers have improved upon the method [11]. Advances include dynamic adjustments and estimations to thresholding and algorithms that define how the given thresholds apply.

(image example of wavelet transform)

According to [11], Block-matching and 3D filtering (BM3D) is the “current state of the art algorithm for denoising images corrupted by Additive White Gaussian noise (AWGN)” which is particularly relevant as most CT image noise is modelled by AWGN [10]. BM3D [12, 20] involves two main stages: 1) basic estimation (hard thresholding) and 2) refined estimation (wiener filtering). Stage 1 produces a basic estimate of the denoised image and Stage 2 uses this estimate to inform its result. In both stages, the image is divided into overlapping patches, which are then stacked together into 3D arrays. This is followed by a filtering step. For this step, Stage 1 involves thresholding in the wavelet or DCT domain and Stage 2 applies Wiener filtering using the basic estimation and the original image. Finally, the filtered patches are transformed back to the spatial domain, with overlapping patches averaged to produce the final image.

(diagram of BM3D)

Deep Learning Techniques

With the increased popularity of Deep Learning (DL) applications in recent years, several DL-based denoising strategies have been developed. An overview of these is provided in [12]. Convolutional Neural Networks (CNNs) are the most popular architecture in DL-based denoising. Making use of both convolutional layers and fully connected layers, they can extract and refine the main anatomical features of CT images and remove noise. CNNs apply non-linear transformations discerning noise from actual patters and build adaptive filters that can separate the critical features from the noise [12]. They work particularly well on high-dimensional data and can produce resulting images of high resolution. [11] and [12] both highlight the popular RED-CNN (Residual Encoder Decoder CNN) [21] and U-NET [9] architectures. These are set up with encoder-decoder structures using skip-connections. There are two main disadvantages to CNNs. Firstly, they require a significant amount of labeled training data, which is difficult to come by due to limited public medical datasets and the challenges with generating well-modelled paired noisy and noise-free images for comparison. Additionally, they are considered “black box” algorithms and can be difficult to interpret. Despite these, CNNs remain as the centre of the most effective and popular DL-based image denoising methods based on past results [12].

(images of RED-CNN & U-Net)

As described in [12], Generative Adversarial Networks (GANs) consist of a generator and discriminator network. The generator creates new data, and the discriminator discerns whether this data is real or fake. The networks are trained at the same time - the generator trying to minimize the loss function the discriminator is trying to maximize. GANs produce high resolution and realistic output images including finer details, however, require substantial input data, are computationally intensive and can generate clinically inaccurate images, which is problematic in medical applications. CycleGAN [22] is a popular architecture that learns the mapping between two domains without using paired data (in this case the mapping between denoised and noisy inputs), while WGAN [23] uses the Wasserstein distance to compare the distributions of the real and generated images to stabilize the training process. Besides GANs, several other DL-based denoising methods have been investigated, but with less popularity [12]. These include Transformer-based networks, which have shown some promise for their ability to capture spatial dependencies but require significant volumes of data and have some overall performance issues. Variational Auto-encoders (VAEs) which learn to produce a compressed representation of the data using an encoder-decoder structure, and then denoise input images [24]. ResNets, using residual connections allow training to utilize much deeper networks as they solve the “vanishing gradient problem”.

(image of GAN structure)

Metrics

The assessment of denoising algorithms, and which metrics to use is also an important consideration in CT image denoising. Several papers on CT image denoising provide some discussion of this. In many studies, results are both analyzed visually and using standard Image Quality (IQ metrics). The goal of image denoising is generally to optimize the image for clinical relevance. Often assessed by visual inspection, this is measured by qualitatively comparing the following criteria [11]:

1. artifact visibility
2. edge preservation
3. visibility of low-contrast objects
4. texture preservation

To quantitatively assess this in a standardized way, various IQ metrics have been established and used to describe the images. The reviews [10, 11, 12] provide a list of these metrics. Each has advantages and disadvantages and represents clinical relevance in differently. Table 1 provides a summary of each of these, with the most popular being SSIM (Structural Similarity Index), PSNR (Peak Signal to Noise Ratio) and MSE (Mean Squared Error) [12]. SSIM has shown to be particularly useful, being highly correlated with radiologist evaluations of diagnostic quality and low-contrast detectability, and moderately in terms of texture [26]. Many of these metrics involve the comparison between a ground-truth “high-dose” image and a noisy version of the same image, which usually requires the simulation of low-dose noise on the original higher-dose image. However, most real Low-dose CT (LDCT) images are not provided with a complimentary high-dose scan. In such cases, different metrics are used. One of the most common of which is the Entropy Difference (ED) [11]. ED measures the difference of Shannon Entropy [27] between the noisy and resultant denoised images. In Table 1, applications to Paired Samples (with ground-truth denoised images) are labelled PS and Unpaired Samples (without ground-truth denoised images) are labelled US. Various papers [10, 12, 25], affirm that standard metrics do not fully represent the suppression of noise based on human perception and the relevance of the results to clinical applications, which indicates a need to investigate standard and new ways of assessing denoising algorithms.

|  |  |  |
| --- | --- | --- |
| Metric | Description | Relevant Problem |
| SSIM | Structural Similarity Index [11] | PS |
| FSIM | Feature Similarity Index [10] | PS |
| PSNR | Peak Signal to Noise Ratio [11] | PS |
| SNR | Signal to Noise Ratio [12] | US, PS |
| MSE | Mean Squared Error [10] | PS |
| NPS | Noise Power Spectrum [11] | PS |
| IQR | Inter-Quartile Range [12] | PS |
| CCC | Concordance Correlation Coefficient [12] | PS |
| DI | Dunn’s Index [12] | US |
| IQI | Image Quality Index [11] | US |
| ED | Entropy Difference [10] | US |
| DIV | Difference in Variance [11] | US |
| GMSD | Gradient Magnitude Similarity Deviation [11] | US |
| SI | Sharpness Index [10] | US |
| EPI | Edge Preservation Index [10] | US |

Table 1: Summary of Different Metrics. Each metric may be relevant to Paired Samples (PS) or Unpaired Samples (US) in terms of their Relevant Problem

- Datasets

- Independence Test

- list available options: CTIA, Mayo etc.

- Simulation of Noise

- XCIST

- briefly describe Other papers

[end of Locate section)

Combining Task and Denoising (Focus of the Approach)

Go through several papers that are similar & list the libraries, methods and metrics that will be relevant and used in this thesis.

* Problems of Traditional Metrics & Task-based approaches
* Example Standard Framework: Denoise, then classify.
* Autoencoder model/s
* Noise-to-Noise model
* Combined Task-based solution examples

**References:**

1. (-) [2019, IEEE ICECCT] A Comparative Study of Lung Cancer Detection using Machine Learning Algorithms https://ieeexplore.ieee.org/abstract/document/8869001

2. (+) [2016, IEEE Medical Imaging] Convolutional Neural Networks for Medical Image Analysis: Full Training or Fine Tuning? (https://ieeexplore.ieee.org/document/7426826)

3. (\*) [2021, Nature] Common pitfalls and recommendations for using machine learning to detect and prognosticate for COVID-19 using chest radiographs and CT scans (https://www.nature.com/articles/s42256-021-00307-0)

4. (.) [2003, Academic Radiology] Automated lung segmentation for thoracic CT: Impact on computer-aided diagnosis (https://www.sciencedirect.com/science/article/pii/S1076633204003745)

5. (.) [2001, IEEE Transactions on Medical Imaging] Automatic lung segmentation for accurate quantitation of volumetric X-ray CT images (https://doi.org/10.1109/42.929615)

6. (+) [2024, Medical Physics] A multiscale 3D network for lung nodule detection using flexible nodule modeling

7. (\*) [2007, Journal of Nuclear Medicine Technology] Principles of CT: Radiation Dose and Image Quality (https://tech.snmjournals.org/content/35/4/213.full)

8. (+) [2002, Ped Radiol] Dose and image quality in CT (https://link.springer.com/article/10.1007/s00247-002-0796-2)

9. (+) [2021, IEEE Access] U-Net and Its Variants for Medical Image Segmentation: A Review of Theory and Applications (<https://doi.org/10.1109/ACCESS.2021.3086020>)

10. (\*) [2020, Biomed. Sig. Proc. & Control] A review on medical image denoising algorithms https://www.sciencedirect.com/science/article/pii/S1746809420301920

11. (\*) [2018, Biomed. Sig. Proc. & Control] A review on CT image noise and its denoising (https://doi.org/10.1016/j.bspc.2018.01.010)

12. (\*) [2023, Medical Physics] CT image denoising methods for image quality improvement and radiation dose reduction https://aapm.onlinelibrary.wiley.com/doi/pdf/10.1002/acm2.14270

13. (+) [2020, CVPR] Wavelet Integrated CNNs for Noise-Robust Image Classification (<https://openaccess.thecvf.com/content_CVPR_2020/papers/Li_Wavelet_Integrated_CNNs_for_Noise-Robust_Image_Classification_CVPR_2020_paper.pdf>)

14. (.) D.-H. Trinh, T.-T. Nguyen, N. Linh-Trung, An effective example-based denoising method for CT images using Markov random field, in: Proc. IEEE Int. Conf. Advanced Technologies for Communications (ATC 2014), IEEE, Hanoi, 2014, pp. 355–359.

15. L. A. Shepp and B. F. Logan, "The Fourier reconstruction of a head section," in IEEE Transactions on Nuclear Science, vol. 21, no. 3, pp. 21-43, June 1974, doi: 10.1109/TNS.1974.6499235. keywords: {Interpolation;Search methods;Fourier transforms;Bandwidth;Oscillators;Approximation algorithms;Spatial resolution},

16. Mohammadinejad P, Mileto A, Yu L, et al. CT noise-reduction methods for lower-dose scanning: strengths and weaknesses of iterative reconstruction algorithms and new techniques. RadioGraphics. 2021;41(5):1493-1508.

17. Antoni Buades, Bartomeu Coll, Jean-Michel Morel. A review of image denoising algorithms, with a new one. Multiscale Modeling and Simulation: A SIAM Interdisciplinary Journal, 2005, 4 (2), pp.490-530. ff10.1137/040616024ff. ffhal-00271141f

18. L. I. Rudin, S. Osher, and E. Fatemi, Nonlinear total variation based noise removal algorithms, Physica D, 60 (1992), pp. 259–268.

19. Mallat, S.G. A Theory for Multiresolution Signal Decomposition: The Wavelet Representation (1989) IEEE Transactions on Pattern Analysis and Machine Intelligence, 11 (7), pp. 674-693

20. 15. Zhao T, Hoffman J, Mcnitt-Gray M, Ruan D. Ultra-low-dose CT image denoising using modified BM3D scheme tailored to data statistics. Med Phys. 2019;46(1):190-198.

21. H. Chen, Y. Zhang, M.K. Kalra, F. Lin, P. Liao, J. Zhou, G. Wang, Low-Dose CT with a Residual Encoder–Decoder Convolutional Neural Network (RED-CNN), 2017 arXiv preprint arXiv:1702.00288.

(more to follow)