



Degree Project in Scientific Computing

Second cycle, 30 credits

Chest X-ray Transmission Map Reconstruction

Constrained Optimization to Invert a Family of Image Processing
Algorithms

ANDRES ALAM SANCHEZ TORRES

Chest X-ray Transmission Map Reconstruction

Constrained Optimization to Invert a Family of Image Processing Algorithms

ANDRES ALAM SANCHEZ TORRES

Date: May 30, 2025

Supervisor: Ozan Öktem

Examiner: Ozan Öktem

School of Engineering Sciences

Swedish title: Detta är den svenska översättningen av titeln

Swedish subtitle: Detta är den svenska översättningen av undertiteln

Abstract

An abstract is (typically) about 250 and 350 words (1/2 A4-page) with the following components:

- What is the topic area? (optional) Introduces the subject area for the project.
- Short problem statement
- Why was this problem worth a Bachelor's/Master's thesis project? (*i.e.*, why is the problem both significant and of a suitable degree of difficulty for a Bachelor's/Master's thesis project? Why has no one else solved it yet?)
- How did you solve the problem? What was your method/insight?
- Results/Conclusions/Consequences/Impact: What are your key results/conclusions? What will others do based on your results? What can be done now that you have finished - that could not be done before your thesis project was completed?

Choice of typeface with `\textit`, `\textbf`, and `\texttt`: x , **x**, and `x`.

Text superscripts and subscripts with `\textsubscript` and `\textsuperscript`: A_x and A^x .

Some symbols that you might find useful are available, such as: `\textregistered`, `\texttrademark`, and `\textcopyright`. For example, the copyright symbol: `\textcopyright Maguire 2022` results in ©Maguire 2022. Additionally, here are some examples of text superscripts (which can be combined with some symbols): `99mTc`, `A*`, `A\textregistered`, and `A\texttrademark` resulting in ^{99m}Tc, A*, A[®], and A[™]. Two examples of subscripts are: `H\textsubscript{2}O` and `CO\textsubscript{2}` which produce H₂O and CO₂.

You can use simple environments with `begin` and `end`: `itemize` and `enumerate` and within these use instances of `\item`.

The following commands can be used: `\eg`, `\Eg`, `\ie`, `\Ie`, `\etc`, and `\etal`: *e.g.*, *E.g.*, *i.e.*, *I.e.*, *etc.*, and *et al.*

The following commands for numbering with lowercase Roman numerals: `\first`, `\Second`, `\third`, `\fourth`, `\fifth`, `\sixth`, `\seventh`, and `\eighth`: (i), (ii), (iii), (iv), (v), (vi), (vii), and (viii). Note that the second case is set with a capital 'S' to avoid conflicts with the use of second of as a unit in the `siunitx` package.

Equations using $\backslash(xxxx \backslash)$ or $\backslash[xxxx \backslash]$ can be used in the abstract. For example: $(C_5O_2H_8)_n$ or

$$\int_a^b x^2 dx$$

Note that you **cannot** use an equation between dollar signs.

Even LaTeX comments can be handled by using a backslash to quote the percent symbol, for example: `% comment`. Note that one can include percentages, such as: 51% or 51 %.

Keywords

Nonlinear Optimization, Medical Imaging, Digital Image Processing

Choosing good keywords can help others to locate your paper, thesis, dissertation, ...and related work.

Choose the most specific keyword from those used in your domain, see for example: the ACM Computing Classification System (<https://www.acm.org/publications/computing-classification-system/how-to-use>), the IEEE Taxonomy (<https://www.ieee.org/publications/services/the-saurus-thank-you.html>), PhySH (Physics Subject Headings) (<https://physh.aps.org/>), ...or keyword selection tools such as the National Library of Medicine's Medical Subject Headings (MeSH) (<https://www.nlm.nih.gov/mesh/authors.html>) or Google's Keyword Tool (<https://keywordtool.io/>)

Formatting the keywords:

- The first letter of a keyword should be set with a capital letter and proper names should be capitalized as usual.
- Spell out acronyms and abbreviations.
- Avoid "stop words" - as they generally carry little or no information.
- List your keywords separated by commas (",").

Since you should have both English and Swedish keywords - you might think of ordering the keywords in corresponding order (*i.e.*, so that the n^{th} word in each list correspond) - this makes it easier to mechanically find matching keywords.

Sammanfattning

Enter your Swedish abstract or summary here!

Nyckelord

Canvas Lärplattform, Dockerbehållare, Prestandajustering

Nyckelord som beskriver innehållet i uppsatsen eller rapporten

Acknowledgments

I would like to thank xxxx for having yyyy. Or in the case of two authors:
We would like to thank xxxx for having yyyy.

Stockholm, May 2025
Andres Alam Sanchez Torres

Contents

1	Introduction	1
1.1	Background	1
1.2	Problem Statement	2
1.2.1	Original Problem Definition	2
1.2.2	Scientific and Engineering Challenges	3
1.3	Purpose	4
1.4	Goals	4
1.5	Thesis Structure	4
2	X-ray Transmission	5
2.1	X-ray intensity and Beer's law	5
2.2	X-ray Image acquisition	6
2.3	Application to Human Chest Imaging	7
2.4	Digital Image Processing in Radiography	8
2.4.1	Dynamic range reduction	8
2.4.2	Spatial Frequency Processing	9
2.4.3	Transmission Maps	11
3	Problem Formulation	15
3.1	Mathematical Model	15
3.1.1	Segmentation	17
3.1.1.1	Mask groups	18
3.2	Processing Operator Model	18
3.3	Optimization Approach	20
3.3.0.1	Data Fidelity Term	21
3.3.1	Regularization	21
3.3.2	Solution	22
3.4	Evaluation Methodology	23

4	Implementation	25
4.1	Data Collection and Preprocessing	26
5	Results and Analysis	28
6	Conclusions and Future work	29
6.1	Conclusions	29
6.2	Limitations	29
6.3	Future work	29
6.3.1	Immediate Extensions	29
6.3.2	Long-term Research Directions	30
6.3.3	Practical Deployment	30
6.4	Reflections	30
	References	33
A	Supporting Material	36

List of Figures

1.1	Examples of different capture conditions across datasets, taken from [2, Figure 6].	2
1.2	X-ray image acquisition pipeline. All projection radiography capture the same x-ray energies and form a latent image, which is distorted by the systems' methods to capture x-rays and subsequent processing steps. Taken from [3, p. 13]	3
2.1	Body tissue attenuation coefficients in Hounsfield units. Adapted from [7, p. 54]	8
2.2	Gradational curves applied to normalize signal over different x-ray doses.[4, Figure 3]	9
2.3	Unsharp masking algorithm [4, Figure 5]	10
2.4	Sample pair of images from chest phantom	11
2.5	Histograms of chest phantom transmission maps	12
2.6	Histograms of chest phantom transmission maps grouped by tissue groups	13
2.7	Histograms of processed transmission maps	14
3.1	Segmentation masks	18
4.1	Data preprocessing pipeline	27
A.1	Windowing function with different parameter settings.	37

List of Tables

2.1	Anatomical structures grouped by absorption coefficients . . .	13
4.1	Optimization hyperparameters	26

Listings

If you have listings in your thesis. If not, then remove this preface page.

List of acronyms and abbreviations

This document is incomplete. The external file associated with the glossary ‘acronym’ (which should be called `main.acr`) hasn’t been created.

Check the contents of the file `main.acn`. If it’s empty, that means you haven’t indexed any of your entries in this glossary (using commands like `\gls` or `\glsadd`) so this list can’t be generated. If the file isn’t empty, the document build process hasn’t been completed.

Try one of the following:

- Add `automake` to your package option list when you load `glossaries-extra.sty`. For example:

```
\usepackage[automake]{glossaries-extra}
```

- Run the external (Lua) application:

```
makeglossaries-lite.lua "main"
```

- Run the external (Perl) application:

```
makeglossaries "main"
```

Then rerun \LaTeX on this document.

This message will be removed once the problem has been fixed.

Chapter 1

Introduction

1.1 Background

Digital radiography systems and their technological advancements for better X-ray image acquisition play a vital role in enabling accurate disease diagnosis and progression monitoring. Naturally, a demand for automated image interpretation and diagnosis has emerged, followed by a need of large and diverse X-ray image datasets. Collecting such data is however a challenge given the complexity of the acquisition, annotation process, and the tendency of these datasets to have an imbalance when underrepresenting certain conditions, such as normal cases, while simultaneously overrepresenting pneumonia [1].

Another source of bias is the different capture conditions across (or even within) datasets. These can be found in images that go through different enhancing methods, device quality, and patient positioning. Some of these differences are depicted in **Figure 1.1**. And even though these differences can be mitigated by mixing datasets, this method can also lead to models discriminating on the source dataset [2]. Most importantly, any conditions that are not represented in the data can lead to existing models performing poorly.

A potential solution to these dataset biases lies in the fundamental physics of X-ray imaging. In projection radiography, all imaging systems—whether using film, computed radiography phosphor plates, or flat panel detectors—measure the spatial distribution of X-ray radiation incident on the detector after passing through the patient [3]. Different detector technologies employ various methods to convert the incoming x-rays into a visual representation, potentially losing information in the process. Furthermore,

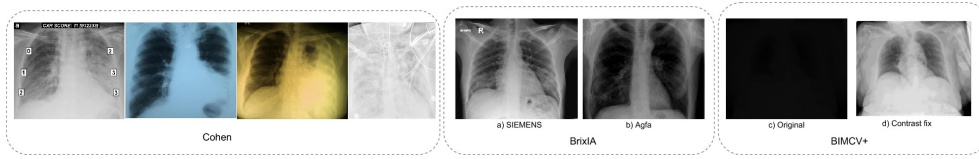


Figure 1.1: Examples of different capture conditions across datasets, taken from [2, Figure 6].

digital detectors perform image processing algorithms to increase the image quality for diagnosis, but also contribute to the differences in image appearance across different detectors.

This observation suggests a promising approach: by recovering these fundamental transmission patterns from processed clinical images, existing annotated datasets could be translated into detector-agnostic representations. This could enable Machine Learning and Deep Learning models to generalize across different imaging systems without requiring new data for each specific detector. This thesis explores the feasibility and effectiveness of transmission map recovery as a preprocessing step for improving the generalization of chest X-ray analysis models.

1.2 Problem Statement

1.2.1 Original Problem Definition

All projection radiography technologies, and their subsequent processing pipelines operate on the same *transmission maps*, which represent the X-ray radiation that passed through the patient's body. For digital radiography systems, if no processing were applied and the images captured were directly transformed into gray levels, they would appear as extremely dark due to the lack of contrast [4, p. 148]. Thus, clinical grade equipment manufacturers (e.g., Siemens, Philips, GE) apply proprietary digital image processing algorithms that include non-linear operations—likely including a combination of denoising, frequency filtering, and dynamic range adjustments. These operations form a "black-box" that varies across manufacturers, and even on an image-by-image basis.

This describes the primary goal of this project: to develop a framework that inverts the image processing transformations from high-quality, chest x-ray images. This will recover the latent images described in Figure 1.2 that will enable the extension of exiting labeled datasets via the siulation of non-

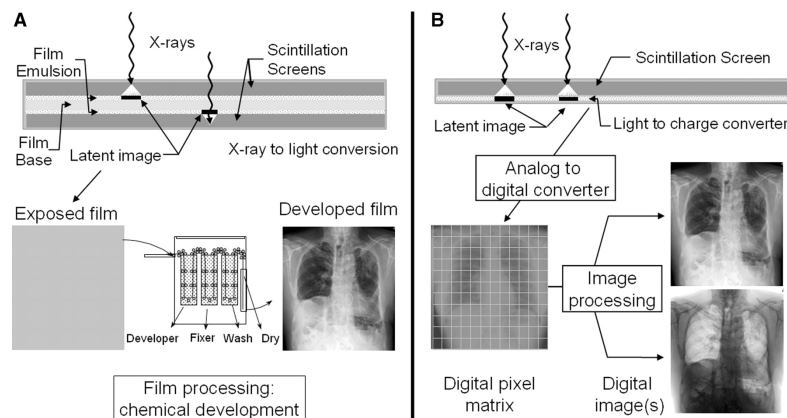


Figure 1.2: X-ray image acquisition pipeline. All projection radiography capture the same x-ray energies and form a latent image, which is distorted by the systems' methods to capture x-rays and subsequent processing steps. Taken from [3, p. 13]

represented radiography systems.

1.2.2 Scientific and Engineering Challenges

The problem to solve lies in the category of inverse problems, and has several critical challenges:

Ill-posedness A single processed image can map back to multiple possible "realistic" transmission maps. The same holds on the other way around, where a single transmission map can map back to multiple processed images. This makes the inverse problem fundamentally ill-posed.

Diversity of processing approaches Each manufacturer implements proprietary image processing pipelines that differ in their specific algorithms and parameters.

Limited ground truth We lack labeled data showing the original transmission maps corresponding to processed images.

Scaling requirements We need to process many images to build a representative dataset.

1.3 Purpose

The primary use case that motivates this project is the development of alternative X-ray detector technologies. Large chest X-ray datasets like CheXpert [5] and ChestX-ray8 [6] collect images from large and modern hospitals with extensive archives of X-ray imaging studies. While these can drive X-ray models to be used in most clinical environments, they may not be effective on resource-limited settings. These include portable, or cost-effective alternatives.

Building large datasets for these subset of technologies is unlikely, given this has been possible due to modern archiving systems (PACS) keeping long histories of X-ray studies [6, p. 3462]. Being able to translate existing datasets into representations of any other imaging system will bring existing and coming advancements in X-ray related automation to all scenarios.

1.4 Goals

Concretely, the expected outcomes of this project are the following:

1. Analyze limited X-ray transmission map data to identify and characterize the essential features that define radiologically realistic representations.
2. Characterize the diversity of image processing pipelines applied to these maps in commercial systems.
3. Develop and evaluate optimization algorithms that can recover plausible transmission maps from processed images.

1.5 Thesis Structure

The remainder of this thesis is organized as follows: Chapter 2 presents background information on X-ray transmission and digital image processing that are relevant to the problem and motivate the proposed method. Chapter 3 presents the mathematical description of the problem and the methods used to solve it. Chapter 4 includes implementation details. Chapter 5 presents experimental results, and Chapter 6, implications and directions for future work.

Chapter 2

X-ray Transmission

The lack of ground truth and prior knowledge for our problem incentivizes the analysis of the underlying physics of X-ray images, as well as further research on the types of operations that are done in the black boxes we aim to invert. This chapter presents the basic principles of X-ray transmission, the process of X-ray image acquisition, and the main goals and common algorithms that are used to process radiographic images. The knowledge presented in this section motivates the design choices made to develop the proposed method.

2.1 X-ray intensity and Beer's law

X-rays are thought of as a flux of very high-energy, electromagnetic radiation. The x-ray beam is described by a vector valued function $I(x)$. The direction of I at x is the direction of the flux at x and its magnitude,

$$I(x) = \|I(x)\| \quad (2.1)$$

is the intensity of the beam. If dS is an infinitesimal surface element at x of area $|dS|$, placed at right angles to $I(x)$, then the energy-per-unit-time passing through dS is [7, p. 56].

$$I(x)|dS|. \quad (2.2)$$

When x-rays encounter any form of matter, they are partly transmitted and partly absorbed. The fractional decrease in the intensity I of an x-ray beam as it passes through any homogeneous substance is proportional to the distance traversed x and the material encountered[8, p. 11]. The intensity I of the x-ray beam satisfies Beer's law:

$$\frac{dI}{ds} = -\mu(x)I, \quad (2.3)$$

where s is the arc-length along the straight-line trajectory of the x-ray beam. Each material encountered has a characteristic *linear attenuation coefficient* μ for x-rays of a given energy, and is dependent on the substance composition, density, and the wavelength of the x-rays [7, p. 57].

For a homogeneous medium of thickness x with linear attenuation coefficient μ , if radiation of intensity I_{in} is incident upon the medium, the transmitted intensity I_{out} is given by:

$$I_{\text{out}} = I_{\text{in}} \cdot e^{-\mu(x)} \quad (2.4)$$

2.2 X-ray Image acquisition

In the context of X-ray imaging, we measure intensities at the detector plane. Let

- $I_0(x, y)$ denote the intensity measured at detector position (x, y) in the absence of any object (the reference or flat-field measurement), and
- $I(x, y)$ denote the intensity measured at detector position (x, y) with the object present.

Then, for radiation traversing a heterogeneous medium, such as body tissues, along a ray path L from source to detector position (x, y) , the relation of the incident and transmitted intensity is given by:

$$I(x, y) = I_0(x, y) \cdot \exp \left(- \int_L \mu(s) ds \right) \quad (2.5)$$

where $\mu(s)$ represents the spatially varying attenuation coefficient along the ray path [7, p. 57].

In a real measurement, the x-ray source is turned on for a known period of time. The total energy I incident on the object along a given line l is known. The total energy, I_0 , emerging from the object along l is then measured by an x-ray detector. Integrating Beer's law we obtain [7, p. 60]

$$-\log \frac{I(x, y)}{I_0(x, y)} = \int_l \mu(s) ds, \quad (2.6)$$

where $\frac{I(x,y)}{I_0(x,y)}$ is the *transmission map*, which describes the fraction of the incident radiation that transmits through the patient at each position (x, y) , and applying the negative logarithm captures the energy absorbed by the object.

For a medium composed of n distinct homogeneous regions with attenuation coefficients $\{\mu_i\}_{i=1}^n$ and thicknesses $\{d_i\}_{i=1}^n$ along a given ray path, a simplified model can be expressed as

$$I(x, y) = I_0(x, y) \cdot \exp \left(- \sum_{i=1}^n \mu_i d_i \right). \quad (2.7)$$

Note that ideal x-ray image acquisition and the described model assume a point x-ray source, a straight line trajectory from the source through the object, and complete detection of the x-ray beam that strikes the detector. These are not realistic assumptions, however, clinical detectors apply methods to counteract nonideal conditions [3, p. 9].

2.3 Application to Human Chest Imaging

In the context of chest radiography, the human thorax can be modeled as a composition of distinct, but known tissue types, and therefore, different anatomical structures have different attenuation coefficients. Bone has a much higher attenuation coefficient than soft tissue, and different soft tissues have slightly different coefficients. More precisely, [Figure 2.1](#) shows the attenuation coefficients of common body tissue types. The table is presented in a dimensionless quantity called a Hounsfield unit, which is a measure relative to the attenuation coefficient of water, defined as [7, p. 54]

$$H_{\text{tissue}} = 1000 \cdot \frac{\mu_{\text{tissue}} - \mu_{\text{water}}}{\mu_{\text{water}}} \quad (2.8)$$

Tissue Type	Attenuation Coefficient
water	0
air	-1000
bone	1086
blood	53
fat	-61
breast tissue	9
muscle	41
soft tissue	51

Figure 2.1: Body tissue attenuation coefficients in Hounsfield units. Adapted from [7, p. 54]

2.4 Digital Image Processing in Radiography

CR systems can capture the wide attenuation differences between lungs and mediastinum due to their wide dynamic range and linear response to the incident radiation. However, it leads to a lack of contrast on the direct conversion to gray values. At the same time, image sharpness may not be as good as in screen-film due to the pixel size constraint [4, p. 148]. DR systems, despite having greater spatial resolution, they have a non-linear (S-shaped) response, leading to under- or overexposed images [9, p. 551]. Regardless of the imaging modality, digital image processing is not only used to take full advantage of the positive characteristics of the radiography systems, but also to amend the intrinsic limitations of each detector.

To overcome this, different algorithms are applied to the captured x-ray image, with the goals of:

- displaying the full range of attenuation differences in the chest,
- optimizing spatial resolution of digital chest radiographs,
- enhancing structural contrast in the lungs and mediastinum, and
- suppressing image noise [4, p. 149].

2.4.1 Dynamic range reduction

Particularly for CR systems, the dynamic range is large, only a small portion of it contains diagnostically relevant information. The image histogram can

be 'stretched' via a γ (γ) that exclude any values outside the relevant range. As depicted in [Figure 2.2](#), this improves the overall image contrast, and normalizes for different doses and patient body types [[4](#), p. 151].

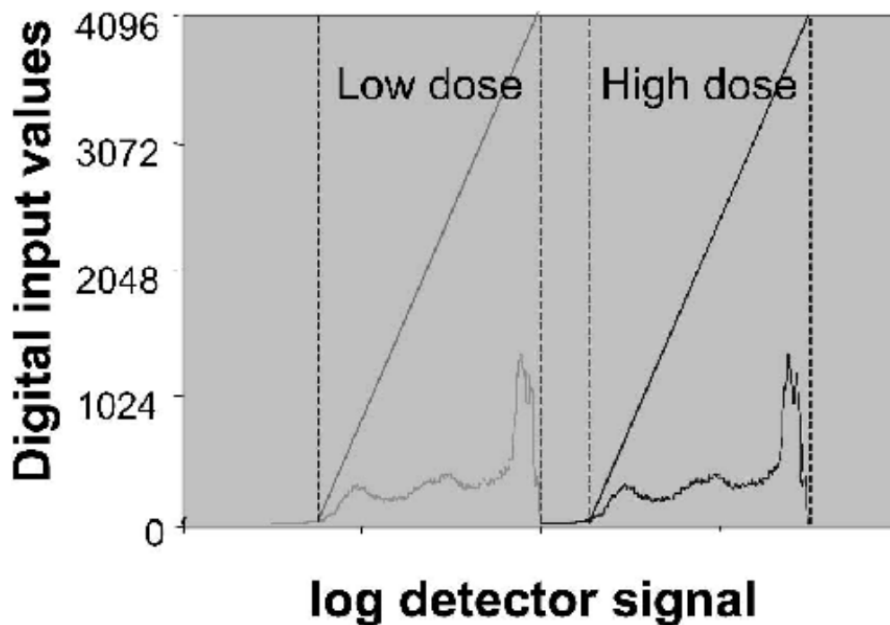


Figure 2.2: Gradational curves applied to normalize signal over different x-ray doses.[[4](#), Figure 3]

2.4.2 Spatial Frequency Processing

As a mean to improve structural contrast and achieve edge enhancement, spatial frequency algorithms are applied. Unlike dynamic range operations, these types of processings can vary drastically across manufacturers, and are driven by multiple parameters that can also change on a per-image basis. Nevertheless, despite the algorithms changing by vendor (Agfa [[Mortsel, Belgium](#)] uses MUSICA; Fuji [[Tokyo, Japan](#)] uses Gradation; and Kodak uses Tonescaping), they all share the same effect and purpose [[10](#), p. 119].

An example of a spatial frequency filter is Unsharp masking, which is a simple method for edge enhancement and sharpening of images, and consists of three core steps:

Step 1 The image is blurred using low-pass filtering, e.g. via a Gaussian filter. The wider the kernel size, the more blurred the low-pass image is, and structures lower than the kernel size are almost completely suppressed.

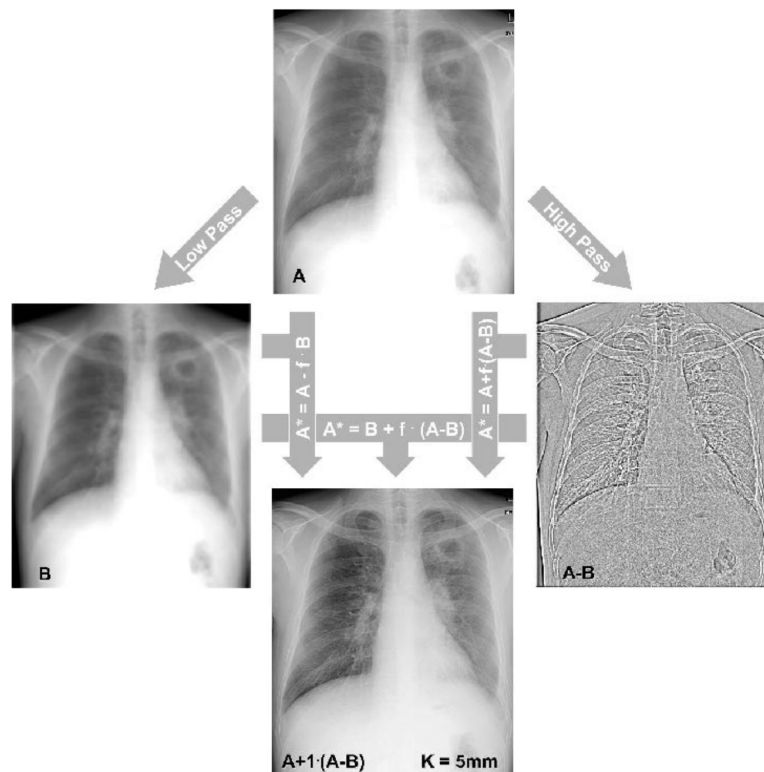


Figure 2.3: Unsharp masking algorithm [4, Figure 5]

Step 2 Subtracting the low-pass image from the original yields the high-pass information and contains those details that are suppressed on the low-pass image. If the kernel size is not too small, most sharp contours on the original image will be retained in the high-pass image.

Step 3 For the enhanced image, various weighted combinations of the original, low-pass or high-pass images can be created. A common approach is to add the weighted high-pass image to the original image, leading to an amplification of the detail information contained in it [4, p. 152-153],.

The effects of unsharp masking depend critically on the size of the filter kernel used to create the blurred image. Kernels of 20-30mm typically give the best results for general chest imaging [4, p. 153], as they enhance structures of diagnostic interest while avoiding suppression of diagnostic information.

More complex algorithms such as Multiscale processing can apply the effect of unsharp masking on multiple scales by splitting an image into many

frequency bands, enabling multiple size-specific enhancements [4, p. 156].

2.4.3 Transmission Maps

Despite the processes behind capturing X-ray images are well studied, to our knowledge, there are currently no open datasets of X-ray transmission maps. We had access to, precisely, two images taken from the same chest phantom, provided by the Hard X-ray lab at KTH, one paired with a processed image produced by a clinical-grade detector. [Figure 2.4](#) shows the obtained sample pair of transmission map and processed image.

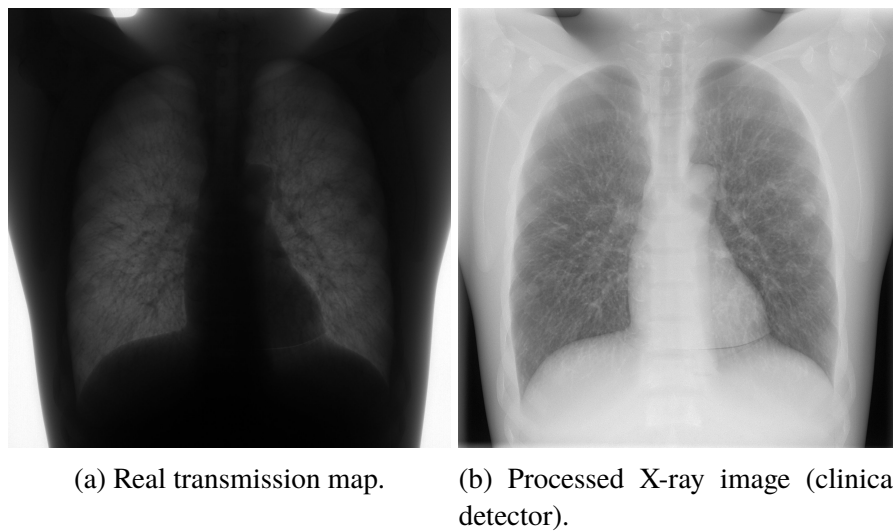


Figure 2.4: Sample pair of images from chest phantom

For this thesis, a method is proposed based on the known physics of x-ray imaging, and insights that can be taken from the analysis of the sample pair shown in [Figure 2.4](#).

Collimated area range Only a fraction of the dynamic range of a transmission map contains diagnostic information. [Figure 2.5](#) shows the histograms of the transmission map samples. It can be observed that the captured transmission values are concentrated in the range below 0.4, with the outliers corresponding to non-diagnostic relevant regions.

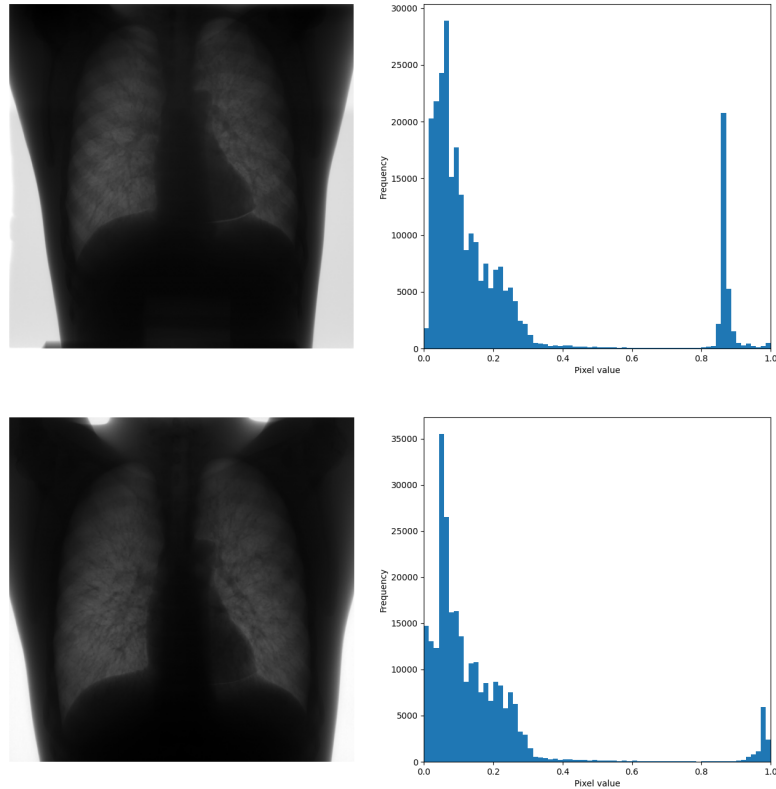


Figure 2.5: Histograms of chest phantom transmission maps

Body tissue groups by absorption coefficients [section 2.3](#) presented the absorption coefficients for the different body tissues present in the chest. The coefficients for the different soft body tissues vary slightly, by about 2% of the dynamic range of an X-ray measurement [[7](#), p. 54]. On the other hand lung (air filled) and bone regions will have a significant difference in energy absorption. This can lead to the definition of three groups of body regions that will lead to similar transmission values, described in [Table 2.1](#). It can be noted that these anatomical structures do not span the entirety of the chest. These are limited to structures we can identify via an existing segmentation model, as defined in [subsection 3.1.1](#).

Group name	Grouped labels
Bone	Left Clavicle Right Clavicle Left Scapula Right Scapula Spine
Lung	Left Lung Right Lung Left Hilus Pulmonis Right Hilus Pulmonis
Soft tissue	Heart Aorta Mediastinum Facies Diaphragmatica Weasand

Table 2.1: Anatomical structures grouped by absorption coefficients

Histogram cluster ranges Figure 2.6 shows the histograms by different tissues (see 3.1.1). Each of the histograms, despite overlapping, all of them have defined boundaries.

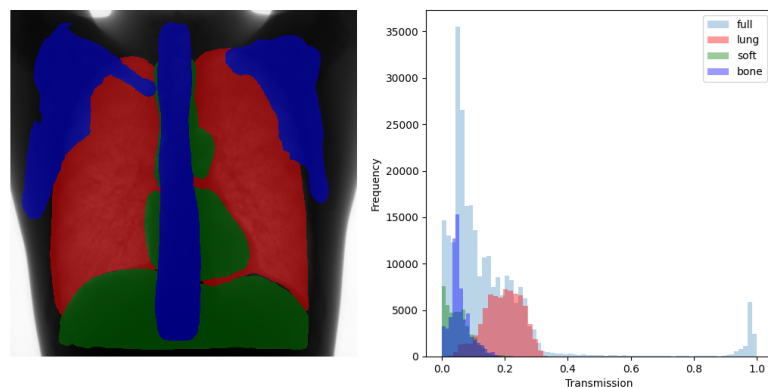


Figure 2.6: Histograms of chest phantom transmission maps grouped by tissue groups

Non-flattened histogram Figure 2.7 shows the histograms of a real pair of transmission map and processed image. Processed images tend to have a flattened histogram as consequence of contrast enhancement operations.

In contrast, the transmission map histogram has higher density in the values closer to zero.

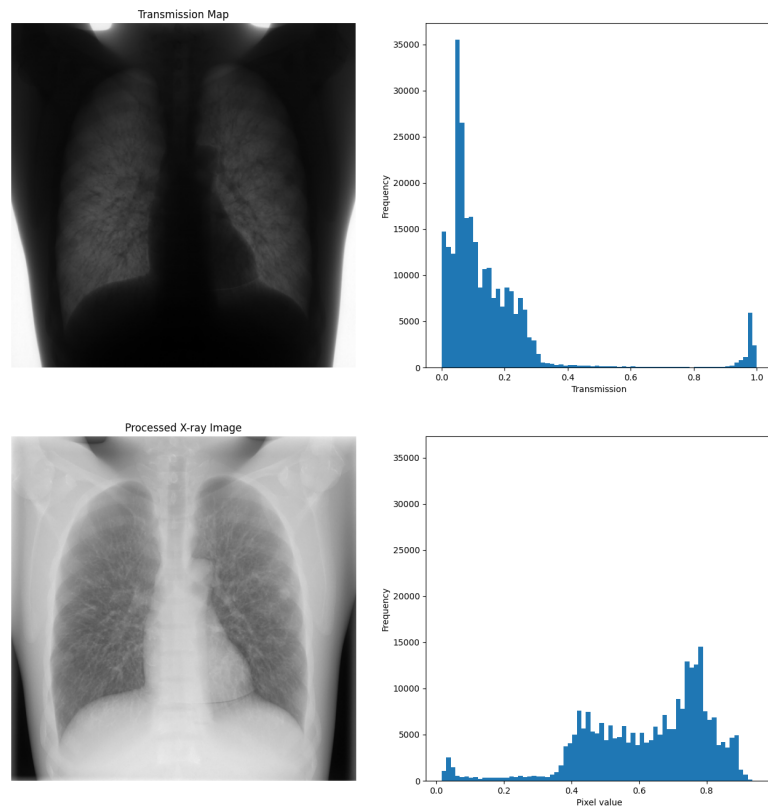


Figure 2.7: Histograms of processed transmission maps

Chapter 3

Problem Formulation

The design of the presented method is motivated by the following findings coming from prior knowledge on X-ray images and the ?? algorithms used in the field:

- Despite differences in processing algorithms across vendors, all seek the same perceptual effects [10, p. 119], and iterations of such algorithms avoid changes in the appearance that radiologists expect from X-ray images [11, p. 57].
- If we recover an appropriate image processing model that can produce images perceptually similar to the target X-ray images, we may use it as a basis to reconstruct 'realistic' transmission maps from the processed images.

Moreover, there is little to none prior information about transmission maps. Thus, if it's not possible to design a model constrained on data, it might be possible to constrain it on the transformations that map the latent images to the observed X-ray datasets.

3.1 Mathematical Model

For a dataset of observed, grayscale processed X-ray images $Y = \{y_i\}_{i=1}^N$, we aim to recover a corresponding set of transmission maps $X = \{x_i\}_{i=1}^N$, with $x_i, y_i \in [0, 1]^{r,c}$ that satisfy

$$F(x_i, \theta) = y_i. \quad (3.1)$$

Here, $F : [0, 1]^{r,c} \times \mathbb{R}^d \rightarrow [0, 1]^{r,c}$ denotes a known image processing algorithm governed by parameters $\theta \in \mathbb{R}^d$.

This can be formulated as an optimization problem: for each image y_i , we aim to find the transmission map x_i and parameters θ_i that minimize a loss function $\ell(x_i, \theta_i)$:

$$\min_{x_i, \theta_i \in \Omega_\theta} \underbrace{\mathcal{D}(y_i, F(x_i, \theta_i)) + \lambda_{\text{smooth}} R_{\text{smooth}}(x_i) + \lambda_{\text{anatomy}} R_{\text{anatomy}}(x_i)}_{=: \ell(\cdot)}. \quad (3.2)$$

\mathcal{D} is a data fidelity function measuring the discrepancy between the processed versions of our estimated transmission maps and the observed processed images, becoming zero only when Equation 3.1 is satisfied. Nevertheless, ensuring accuracy with the processed images is not a sufficient condition to solve the problem. Consider the case of a set of parameters θ_i that produce no effect on the image, for instance, applying a linear ?? that spans the original image's dynamic range, or applying unsharp masking with a factor of zero. Besides aiming to inject prior information through the operations in F , additional constraints can be imposed to ensure that the reconstructed transmission maps are physically plausible.

One expectation of the images x_i is for them to not have discontinuities or noise. An approach to promote that is to introduce a smoothness regularization term R_{smooth} , weighted by λ_{smooth} . In subsection 2.4.3 I also discussed broad constraints the images should meet, regarding the range of the transmission values, particularly in regions where the traversing tissues have close absorption coefficients. This anatomical value ranges can't be strictly enforced due to the diversity of body compositions, as well as the potential of it creating block artifacts. Therefore, these are promoted with the regularization R_{anatomy} , with corresponding weight λ_{anatomy} .

Our transformation F can be further constrained by restricting the parameters θ_i to take values that produce effects that are assumed to happen in real settings. In that case, the optimization is considered constrained, and has an admissible solution only when θ_i is in a defined domain Ω_θ . As F is chosen based on existing ?? algorithms, we can therefore constraint the parameters θ_i on expected value ranges and values recommended in the literature, and Ω_θ can be expressed as multiple inequality constraints:

$$\theta_i \in \Omega_\theta = \{\theta_i \mid c_j(\theta_i) \geq 0, j \in \mathcal{I}\}. \quad (3.3)$$

The presented model suggests that each image y_i will have a corresponding

transmission map x_i and parameters θ_i . However, the following special cases of the problem can also be considered:

Common operator parameters If we assume that there exists a single set of parameters θ (yet unknown) for all processed images, i.e. $F(x_i, \theta) = y_i$ for all $i = \{1, \dots, N\}$. Then the problem can be modeled as

$$\min_{x_i, \theta_0 \in \Omega_\theta} \mathcal{D}(y_i, F(x_i, \theta_0)) + \sum_{j=1}^K \lambda_k R_j(x_i) \quad (3.4)$$

Known parameters If we assume there is a known set of parameters θ_0 for the operator F that can model all the observed processed images, then we reduce the problem to

$$\min_{x_i} \mathcal{D}(y_i, F(x_i, \theta_0)) + \sum_{j=1}^K \lambda_k R_j(x_i). \quad (3.5)$$

The method employed in this project involved an iterative process of defining a data fidelity function and regularization functions that will recover transmission maps considered realistic (according to a predefined set of characteristics). In the following sections, we will discuss the choices made for F , \mathcal{D} , regularizations, and their effects on the reconstruction of transmission maps.

3.1.1 Segmentation

One of the core validation measures for the recovered transmission maps is the expected transmitted x-rays on different body tissues. As presented in Table 2.1, three groups of tissues can be defined that will have similar transmission values, and differ significantly across them. This will drive the choice of a regularization term based on the expected transmission values relative to the location of a pixel.

To identify these groups, the ChestXDet [12] model was used to obtain segmentation labels. Concretely, the segmentation model identifies 13 labels, which are grouped in 3 major categories. These groups and their corresponding labels are listed in Table 2.1.

3.1.1.1 Mask groups

To compute the segmentation masks, the ChestXdet model returns confidence values on a 0 to 1 range. To get binary masks, we include in our model a threshold parameter. The threshold is used to create binary masks where the values are above it, then the segmentation targets are joined according to the groups described in [Table 2.1](#). The join operation is performed by taking the logical OR of the masks.

Since the mask groups may contain overlapping regions, a difference is applied to obtain exclusive masks. This is done in an ordered manner, starting from the groups with higher absorption (bone) up to lower absorption values (lung). This ensures that each pixel is assigned to the mask that produces the higher attenuation. The complete merging operation is described in [algorithm 1](#). [Figure 3.1](#) shows a sample set of the processed segmentation masks over a CheXpert image using a threshold of 0.5.

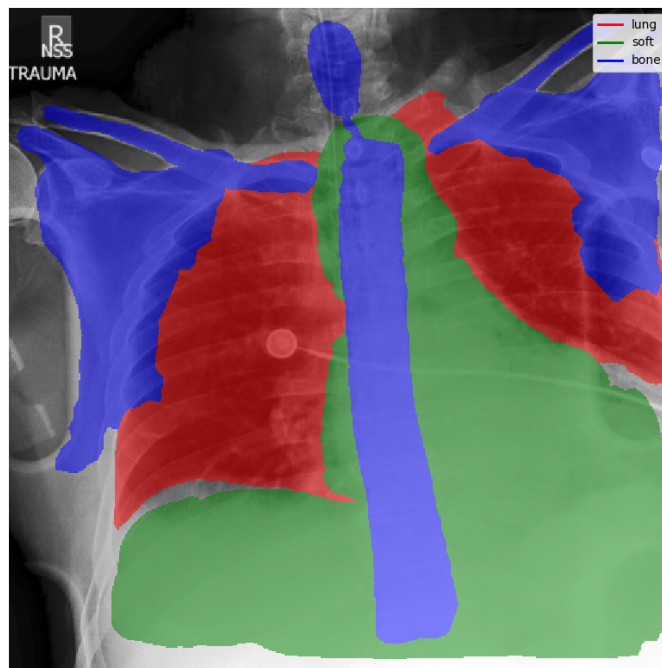


Figure 3.1: Segmentation masks

3.2 Processing Operator Model

Throughout the entire model experimentation, a fixed parameterized processing operator is used to represent a digital radiography processing pipeline.

The goal of building such function is so it can produce X-ray images with an appearance consistent to the existing image datasets. If the function and its parameters allow it, then, there is the expectation that inverting this would produce transmission maps

The choices of the steps for the operator are driven by the goals of image processing described in [section 2.4](#). These guide the design of our operator F as a composition of several transformations

$$F = F_n \circ F_{n-1} \circ \dots \circ F_1 \quad (3.6)$$

that include:

Negative logarithm $F_1(x) = -\log(x + \epsilon)$, where ϵ is a small constant to avoid numerical instability on values close to zero.

Transmission maps represent the ratio of X-ray intensities I_1 and I_0 ,

$$\frac{I_1}{I_0} = e^{-\mu x}. \quad (3.7)$$

However, processed images operate on the densities term μx , explaining the color-negative relationship between transmission maps and diagnostic images.

Windowing As a way of implementing gradational adjustment, a window function is implemented that creates an S-shape lookup table to achieve signal normalization:

$$F_2(x) = \frac{1}{1 + e^{-\gamma \frac{x-c}{w}}}. \quad (3.8)$$

c denotes the center of the sigmoid function, w is a width parameter, and γ is a steepness parameter. The effect of these parameters and how these translate into a LUT is shown in [Figure A.1](#). By definition, c and w are restricted to the domain $(0, 1)$, thus introduce box constraints on θ . For *gamma*.

Unsharp masking

$$F_3(x) = \frac{x - \alpha \cdot G_\sigma * x}{1.0 - \alpha} \quad (3.9)$$

where G_σ is a Gaussian kernel with standard deviation σ , and α is the enhance factor. The chosen formulation bounds α to $(0, 1)$. However,

since it is known that images go through some edge enhancement, the lower bound can be increased to a non-zero value to assume some edge enhancement is being applied in the processed image. Regarding the frequency band represented by σ , the implementation considers a fixed value, or a range that will capture higher-frequencies, following the recommendation of targeting a 20-30mm kernel size [4, p. 153].

Range normalization

$$F_4(x) = \frac{x - \min(x)}{\max(x) - \min(x)}, \quad (3.10)$$

as a mean to normalize values to the $[0, 1]$ range, since operations such as windowing and unsharp masking can lead to values outside this range.

Clipping

$$F_5(x) = \min(\max(x, 0), 1) \quad (3.11)$$

The parameter vector θ includes all parameters of these transformations, such as $\{c, w, \alpha, \sigma\}$.

3.3 Optimization Approach

As depicted in Equation 3.2, the transmission map recovery involves solving a constrained non-linear optimization problem. The approach taken considers Gradient Descent methods and, for the case of the θ inequality constraints, projections onto the feasible set. However, the ill-posedness nature of our problem prevents the choice of an ideal optimization method (i.e. step size selection), and hyperparameters analytically. Thus, experiments are conducted to find an appropriate loss function and a set of hyperparameters that consistently converges to 'valid' solutions.

For a recovered transmission map x_i , the concept of a 'valid' solution is defined by meeting the following criteria:

1. $F(x_i, \theta)$ is a close approximation to the observed image y_i .
2. x_i does not contain artifacts or noise.
3. x_i has suppressed high-frequency components, in contrast to y_i .
4. The transmission values of x_i at the three anatomical groups described in Table 2.1 are within their expected range.

3.3.0.1 Data Fidelity Term

The data fidelity term is a measure of how close the reconstructed image $F(x_i, \theta)$ is to the original image y_i . The function to consider is \mathcal{D} (??), defined as

$$\mathcal{D}(x, y) = \frac{1}{2} \|x - y\|_2^2. \quad (3.12)$$

MSE has been a popular choice of cost function in image restoration problems, due to its ability to favor a high \mathcal{D} (??), widely used metric for measuring image restoration quality [13, p. 1], [14, p. 191].

However, this metric has been shown to not correlate well with human's perception of image quality, and can yield the same quantity over distortions on different scales of the image [15, Figure 2]. Despite this caveat, it is still the choice to represent data fidelity due to its simplicity and computational efficiency. However, other metrics are used to evaluate and compare the performance of the model, as having accuracy in different frequency bands of our image is critical to claim the model yields a valid solution.

3.3.1 Regularization

The selection of regularization terms is guided by the soft constraints expected from the reconstructed transmission maps. These have the goals of reducing noise and artifacts from the images, as well as promoting the compliance with the expected transmission ranges at each of the absorption-similar regions.

Smoothing regularization Two alternatives of smoothing regularization are considered:

\mathcal{D} (??), defined as

$$TV(x) = \sum_{i,j} \left(|x_{i+1,j} - x_{i,j}| + \sum_{i,j} |x_{i,j+1} - x_{i,j}| \right) \quad (3.13)$$

prompiecewise smoothness in x_i , thus discouraging discontinuities in the image. Despite the fact that TV regularization can lead to over-smoothed images, it is still explored due to its low computational cost.

Segmentation-based regularization We incorporate physics-based constraints derived from anatomical segmentation to ensure realistic transmission values for each tissue type. For each of the anatomical groups discussed

in [subsection 3.1.1](#), a value range is extracted from a real transmission map, and is used as ideal bounds for recovered transmission. The formulation of this regularization consists of a square penalty for values outside their defined bounds (if any). Let $s_i^t(x, y)$ denote the binary mask for x_i of the anatomical group t , and v_{min}^t, v_{max}^t be the expected transmission bounds for said group. Then,

$$p_t(x_i, s_i^t) = \sum_{j,k} s_i^t(j, k) \cdot [\max(0, x_i(j, k) - v_{max}^t)^2 + \max(0, v_{min}^t - x_i(j, k))^2] \quad (3.14)$$

is the penalty on x_i that is zero when all the pixels within the mask s_i^t are within their respective bounds and increases quadratically as the values deviate from them. This penalty is always introduced in the loss function for all explorations.

3.3.2 Solution

For finding the solution of the minimization problem, we recurr to iterative descent methods. These refer to iterative algorithms that compute a sequence of parameters $x^{(k)}, \theta^{(k)}$ that will reduce the loss function at each iteration, i.e.

$$\ell(x^{(k)}, \theta^{(k)}) < \ell(x^{(k-1)}, \theta^{(k-1)}), \quad \text{for all } k > 0. \quad (3.15)$$

The general descent method can be described as follows:

1. $k := 0$
2. Choose initial paramaters $x^{(0)}, \theta^{(0)}$.
3. If $k > 0$ and $\ell(x^{(k)}, \theta^{(k)}) - \ell(x^{(k-1)}, \theta^{(k-1)}) < \epsilon$, stop.
4. If $k > \text{max_iterations}$, stop.
5. Compute descent directions $d_x^{(k)}, d_\theta^{(k)}$
6. Compute step sizes $\sigma_x^{(k)}, \sigma_\theta^{(k)}$
7. $x^{(k+1)} = x^{(k)} + \sigma_x^{(k)} d_x^{(k)}$
8. $\theta^{(k+1)} = \text{proj}_{\Omega_\theta}(\theta^{(k)} + \sigma_\theta^{(k)} d_\theta^{(k)})$
9. Go to 3.

The choice of descent direction is typically the negative gradient, since it points in the direction of the steepest decrease. However, the choice of a step sizes is non trivial, and we experiment with multiple known gradient descent optimizers to compute updates.

Regarding the constraints on the θ parameters, we recurr to a projected gradient descent method, where each update of θ_k is first projected into the admissible domain via

$$\text{proj}_{\Omega_\theta}(\theta_0) = \arg \min_{\theta \in \Omega_\theta} \|\theta - \theta_0\|_2. \quad (3.16)$$

Since all the specified constraints are box constraints, the minimization is satisfied when each parameter is clipped to its respective bounds.

The concrete implementation of such descent methods, as well as the choice of parameters present in the loss function and processing operator are ruled by hyperparameters fixed for the entire iterative process.

3.4 Evaluation Methodology

Due to the lack of ground truth, we rely on qualitative assessments and visual evaluations based on our validity criteria. That translates into finding a model that produces solutions that: 1) reproduce estimated processed images close to the original images, on multiple scales, and 2) satisfy the set physical constraints.

It was mentioned previously that MSE does not necessarily favor solutions that are perceptually close to the original images. There are multiple alternative loss functions designed specifically for Image Quality Assessment metrics, an alternative is a mixed loss metric \mathcal{L}^{Mix} proposed by [13] that combines ?? (??) with the ℓ_1 loss (Mean Absolute Error). ?? is an Image Quality Assessment metric that evaluates images accounting for the fact that the ?? (??) is sensitive to changes in local structure. ?? was later proposed to extend ?? to account for the fact that the scale at which local structure should be analyzed depends on the distance to the observer. This extension weighs ?? computed at different scales. Nonetheless, its demonstrated good performance in image reconstruction problems, outperforming other loss functions by correlating to other image quality indices, comes at the cost of a higher computational cost, which can scale well for Deep Learning settings, but for the current approach, it incentivates the alternative choice of ??. The L^{Mix} loss is still used as an evaluation metric to discern the best choice of hyperparameters.

For the large image datasets, we compute image quality metrics such as SSIM and PSNR between reconstructed processed images and originals, and measure an error on the transmission values that are outside the restricted physical bounds. For the limited cases where ground truth is available (e.g., from phantoms), we also use the processing model to generate alternative processed images, as well as the original paired processed images. That allows us to compare the performance of our model on ground truth.

Chapter 4

Implementation

The implementation of the proposed model consisted on building a processing operator that can produce images with a similar visual appearance to images found in datasets such as CheXpert and ChestX-ray8. As for the optimization task, *JAX* [16] and *Optax* [17] are used to implement the algorithm described in [section 3.3](#). *JAX* is a Python array computation library that provides automatic differentiation, handling the computation of gradients for the chosen loss function, and consequently, the processing operator. *Optax* is an gradient processing and optimization library that provides multiple gradient descent optimizers, ruled by additional hyperparameters, such as a learning rate.

Since the approach taken for this project does not involve learning an inverse model, but rather solve an optimization on a per-image basis, we aim to find a model and corresponding hyperparameters that can be reused across different sets of images and guarantee convergence into valid solutions. Besides optimization-related hyperparameters such as learning rates, we also consider different choices that can produce valid solutions but won't match qualitatively the expectations for transmission maps. One such choice is a regularization term (besides its corresponding weight) that promotes smoothness on the transmission map, which if not chosen properly, can lead to artifacts or over-smoothing. Overall, the hyperparameters and configurations to find are listed in [Table 4.1](#).

Hyperparameter	Description
Optimizer	Choice of optimizer, e.g. Adam, SGD, RMSProp.
Optimizer hyperparameters	Optimizer-specific, such as learning rate, momentum, weight decay, etc.
Segmentation threshold	Threshold to delimit binary segmentation masks, set to 0.5 for all results.
Frequency bands	Usharp mask is an operation that amplifies specific frequency bands, which are determined by the kernel size for the blurring operation. The processing model assumes enhancement on at least two frequency bands, low and high. The consequence of this is the exclusion of this parameter from the optimization problem, but it is rather fixed for all images.
Regularization weights	Weights for the smoothing and segmentation regularization.
Smoothing regularization	The choice among Tikhonov, Total Variation, and no regularization, to penalize discontinuities in the transmission maps.
Max steps	Maximum number of iterations for the optimization.

Table 4.1: Optimization hyperparameters

For finding hyperparameters, a Bayesian search was done, optimizing for the perceptual metric MS-SSIM. An optimal choice of parameters can later be found by picking the set that also satisfies the segmentation constraints. To implement this search, we used Weights & Biases [18] sweep functionality.

4.1 Data Collection and Preprocessing

Throughout the experimentation, besides the single transmission map-processed image pair, images were sampled from cheXpert dataset [5] and NIH Chest X-ray dataset [6]. Recall the motivation to use these datasets is to get value from the diagnostic labels they provide, which can then be tied to the simulated transmission maps. However, these labels are not used for the model.

Although we present an optimization problem that operates per-image, it is expected that multiple images will be processed in parallel to speed up the optimization process. Thus, a fixed image size is set for the implementation. Images from the datasets also do not have a standard size or aspect ratio, so all images are preprocessed by cropping by the center to keep a square

aspect ratio and resized to 512 pixels width, as well as rescaling the grayscale values to the range $[0, 1]$. The choice of the size 512 is set to match the input size of the selected segmentation model implemented in torchxrayvision [19], however, an arbitrary size can be chosen by upscaling the masks. The full data preprocessing steps are shown in Figure 4.1

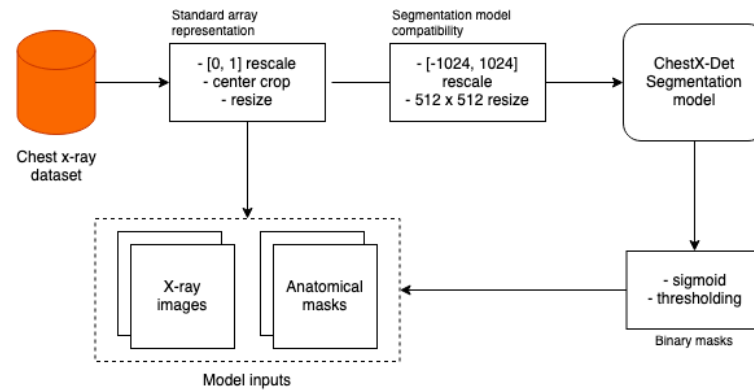


Figure 4.1: Data preprocessing pipeline

Chapter 5

Results and Analysis

Chapter 6

Conclusions and Future work

6.1 Conclusions

6.2 Limitations

Several limitations constrain the current approach:

- **Computational Scalability:** Processing large datasets remains computationally intensive, requiring days of GPU time for comprehensive datasets.
- **Hyperparameter Sensitivity:** Optimal parameters vary across image characteristics, necessitating extensive search or adaptive strategies.
- **Ground Truth Scarcity:** Limited availability of paired transmission map and processed image data restricts validation options.
- **Model Complexity:** The simplified forward model may not capture all nuances of commercial processing pipelines.

6.3 Future work

6.3.1 Immediate Extensions

1. **Adaptive Hyperparameter Selection:** Develop methods to automatically select hyperparameters based on image characteristics, potentially using meta-learning approaches.

2. **Efficient Optimization Strategies:** Explore second-order optimization methods or learned optimizers to reduce the number of iterations required.
3. **Multi-Resolution Processing:** Implement coarse-to-fine optimization strategies to improve both speed and convergence.

6.3.2 Long-term Research Directions

1. **Learned Inverse Models:** Train neural networks to directly predict transmission maps, using our optimization results as training data.
2. **Manufacturer-Specific Models:** Develop specialized forward models for different equipment manufacturers based on their processing characteristics.
3. **Uncertainty Quantification:** Incorporate Bayesian approaches to quantify uncertainty in recovered transmission maps.
4. **Real-time Processing:** Develop approximation methods enabling real-time transmission map recovery for clinical applications.

6.3.3 Practical Deployment

For practical deployment in support of low-cost X-ray system development:

- Create a cloud-based processing pipeline for batch transmission map recovery
- Develop quality metrics to automatically identify successfully recovered maps
- Build a curated dataset of transmission maps with associated metadata
- Establish validation protocols with clinical partners

6.4 Reflections

This work contributes to global health equity by enabling the development of affordable diagnostic imaging systems. The ability to generate training data for alternative X-ray detectors could significantly reduce the cost barrier for medical imaging in resource-limited settings.

The project also highlights the importance of physics-informed machine learning approaches in medical imaging, demonstrating how domain knowledge can guide optimization in under-constrained problems.

While significant computational challenges remain, the framework established here provides a foundation for future work in radiographic image analysis and affordable medical technology development.

One of the most important results is the reduction in the amount of energy required to process each packet while at the same time reducing the time required to process each packet.

The thesis contributes to the **United Nations (UN) Sustainable Development Goals (SDGs)** numbers 1 and 9 by xxxx.

In the references, let Zotero or other tool fill this in for you. I suggest an extended version of the IEEE style, to include URLs, DOIs, ISBNs, etc., to make it easier for your reader to find them. This will make life easier for your opponents and examiner.

IEEE Editorial Style Manual: https://www.ieee.org/content/dam/ieee-org/ieee/web/org/conferences/style_references_manual.pdf

Låt Zotero eller annat verktyg fylla i det här för dig. Jag föreslår en utökad version av IEEE stil - att inkludera webbadresser, DOI, ISBN osv. - för att göra det lättare för läsaren att hitta dem. Detta kommer att göra livet lättare för dina opponenter och examinator.

References

- [1] M. A. Ansari, S. Ahmad, and A. Mewada, “Mitigating risk in medical AI: Balancing X-ray datasets for reliable detection,” *Life Cycle Reliability and Safety Engineering*, May 2025. doi: 10.1007/s41872-025-00316-7 [Page 1.]
- [2] D. Arias-Garzón, R. Tabares-Soto, J. Bernal-Salcedo, and G. A. Ruz, “Biases associated with database structure for COVID-19 detection in X-ray images,” *Scientific Reports*, vol. 13, no. 1, p. 3477, Mar. 2023. doi: 10.1038/s41598-023-30174-1 [Pages ix, 1, and 2.]
- [3] J. A. Seibert and J. M. Boone, “X-ray imaging physics for nuclear medicine technologists. part 2: X-ray interactions and image formation,” *Journal of Nuclear Medicine Technology*, vol. 33, no. 1, pp. 3–18, 2005. [Online]. Available: <https://tech.snmjournals.org/content/33/1/3> [Pages ix, 1, 3, and 7.]
- [4] M. Prokop, U. Neitzel, and C. Schaefer-Prokop, “Principles of image processing in digital chest radiography,” *Journal of Thoracic Imaging*, vol. 18, no. 3, p. 148–164, Jul 2003. doi: 10.1097/00005382-200307000-00004 [Pages ix, 2, 8, 9, 10, 11, and 20.]
- [5] P. Chambon, J.-B. Delbrouck, T. Sounack, S.-C. Huang, Z. Chen, M. Varma, S. Q. Truong, C. T. Chuong, and C. P. Langlotz, “Chexpert plus: Augmenting a large chest x-ray dataset with text radiology reports, patient demographics and additional image formats,” *arXiv preprint arXiv:2405.19538*, 2024. [Pages 4 and 26.]
- [6] X. Wang, Y. Peng, L. Lu, Z. Lu, M. Bagheri, and R. M. Summers, “ChestX-Ray8: Hospital-Scale Chest X-Ray Database and Benchmarks on Weakly-Supervised Classification and Localization of Common Thorax Diseases,” in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Honolulu, HI: IEEE, Jul. 2017. doi:

- 10.1109/CVPR.2017.369. ISBN 978-1-5386-0457-1 pp. 3462–3471. [Pages 4 and 26.]
- [7] C. L. Epstein, *Introduction to the mathematics of medical imaging*, 2nd ed. Philadelphia, PA: Society for Industrial and Applied Mathematics, 2008. ISBN 9780898716429 [Pages ix, 5, 6, 7, 8, and 12.]
- [8] B. D. Cullity and S. R. Stock, *Elements of X-ray Diffraction*, 3rd ed. Harlow: Pearson Education, 2014. ISBN 978-1-292-04054-7 [Page 5.]
- [9] P. Vuylsteke and E. P. Schoeters, “Multiscale image contrast amplification (MUSICA),” M. H. Loew, Ed. doi: 10.1117/12.175090 pp. 551–560. [Online]. Available: <http://proceedings.spiedigitallibrary.org/proceeding.aspx?articleid=955751> [Page 8.]
- [10] C. E. Carter and B. L. Vealé, *Digital Radiography and PACS*, 1st ed. Mosby/Elsevier. ISBN 978-0-323-07221-2 [Pages 9 and 15.]
- [11] M. Stahl, T. Aach, and S. Dippel, “Digital radiography enhancement by nonlinear multiscale processing,” *Medical Physics*, vol. 27, no. 1, pp. 56–65, 2000. [Page 15.]
- [12] J. Lian, J. Liu, S. Zhang, K. Gao, X. Liu, D. Zhang, and Y. Yu, “A Structure-Aware Relation Network for Thoracic Diseases Detection and Segmentation,” *IEEE Transactions on Medical Imaging*, 2021. doi: 10.48550/arxiv.2104.10326. [Online]. Available: <https://arxiv.org/abs/2104.10326> [Page 17.]
- [13] H. Zhao, O. Gallo, I. Frosio, and J. Kautz, “Loss Functions for Image Restoration With Neural Networks,” vol. 3, no. 1, pp. 47–57. doi: 10.1109/TCI.2016.2644865. [Online]. Available: <http://ieeexplore.ieee.org/document/7797130/> [Pages 21 and 23.]
- [14] D. Fleet, T. Pajdla, B. Schiele, and T. Tuytelaars, Eds., *Computer Vision – ECCV 2014: 13th European Conference, Zurich, Switzerland, September 6-12, 2014, Proceedings, Part IV*, ser. Lecture Notes in Computer Science. Springer International Publishing, vol. 8692. ISBN 978-3-319-10592-5 978-3-319-10593-2. [Online]. Available: <http://link.springer.com/10.1007/978-3-319-10593-2> [Page 21.]
- [15] Z. Wang, E. Simoncelli, and A. Bovik, “Multiscale structural similarity for image quality assessment,” in *The Thrity-Seventh*

- Asilomar Conference on Signals, Systems & Computers, 2003.* IEEE. doi: 10.1109/ACSSC.2003.1292216. ISBN 978-0-7803-8104-9 pp. 1398–1402. [Online]. Available: <http://ieeexplore.ieee.org/document/1292216/> [Page 21.]
- [16] J. Bradbury, R. Frostig, P. Hawkins, M. J. Johnson, C. Leary, D. Maclaurin, G. Necula, A. Paszke, J. VanderPlas, S. Wanderman-Milne, and Q. Zhang, “JAX: composable transformations of Python+NumPy programs,” 2018. [Online]. Available: <http://github.com/jax-ml/jax> [Page 25.]
- [17] DeepMind, I. Babuschkin, K. Baumli, A. Bell, S. Bhupatiraju, J. Bruce, P. Buchlovsky, D. Budden, T. Cai, A. Clark, I. Danihelka, A. Dedieu, C. Fantacci, J. Godwin, C. Jones, R. Hemsley, T. Hennigan, M. Hessel, S. Hou, S. Kapturowski, T. Keck, I. Kemaev, M. King, M. Kunesch, L. Martens, H. Merzic, V. Mikulik, T. Norman, G. Papamakarios, J. Quan, R. Ring, F. Ruiz, A. Sanchez, L. Sartran, R. Schneider, E. Sezener, S. Spencer, S. Srinivasan, M. Stanojević, W. Stokowiec, L. Wang, G. Zhou, and F. Viola, “The DeepMind JAX Ecosystem,” 2020. [Online]. Available: <http://github.com/google-deepmind> [Page 25.]
- [18] L. Biewald, “Experiment tracking with weights and biases,” 2020, software available from wandb.com. [Online]. Available: <https://www.wandb.com/> [Page 26.]
- [19] J. P. Cohen, J. D. Viviano, P. Bertin, P. Morrison, P. Torabian, M. Guarrera, M. P. Lungren, A. Chaudhari, R. Brooks, M. Hashir, and H. Bertrand, “TorchXRayVision: A library of chest X-ray datasets and models,” in *Medical Imaging with Deep Learning*, 2022. [Online]. Available: <https://github.com/mlmed/torchxrayvision> [Page 27.]

Appendix A

Supporting Material

Algorithm 1 Algorithms for creating exclusive segmentation masks

Input: *pred* - Batch of prediction masks with shape (batch, labels, height, width)

Input: *threshold* - Optional threshold value

Output: *ordered_groups* - List of mask group identifiers

Output: *exclusive_masks* - Array of exclusive masks with shape (batch, reduced_labels, height, width)

```

1 Function BatchGetExclusiveMasks (pred, threshold) :
2   complete_masks  $\leftarrow$  []
3   foreach group  $\in$  MASK_GROUPS do
4     group_mask  $\leftarrow$  GetGroupMask(pred, group, threshold)
5     Concatenate on labels axis group_mask to complete_masks
6   end
7   exclusive_masks  $\leftarrow$  Copy(complete_masks)
8   groups_count  $\leftarrow$  len(MASK_GROUPS)
9   /* Masks are subtracted in specific order, denser groups last */
10  for i  $\leftarrow$  0 to groups_count - 1 do
11    for j  $\leftarrow$  0 to groups_count - 1 do
12      if j < i then
13        /* Subtract already exclusive masks */
14        exclusive_masks[:, i]  $\leftarrow$  SubtractMask(
15          exclusive_masks[:, i], exclusive_masks[:, j])
16      end
17      if j > i then
18        /* Subtract entire masks */
19        exclusive_masks[:, i]  $\leftarrow$  SubtractMask(
20          exclusive_masks[:, i], complete_masks[:, j])
21      else
22      end
23    end
24  end
25  return exclusive_masks

20 Function GetGroupMask (pred, group, threshold) :
21   /* Select only the labels in the group */
22   group_masks  $\leftarrow$  FilterPredMasks(pred, group)
23   normalized_masks  $\leftarrow$  Sigmoid(group_masks)
24   /* Indicator function: 1 where condition is true, 0 otherwise */
25   thresholded_masks  $\leftarrow$   $\mathbb{1}\{\textit{normalized\_masks} > \textit{threshold}\}$ 
26   joined_masks  $\leftarrow$  sum(thresholded_masks, axis = 1)
27   return clip(thresholded_masks, 0.0, 1.0)

```

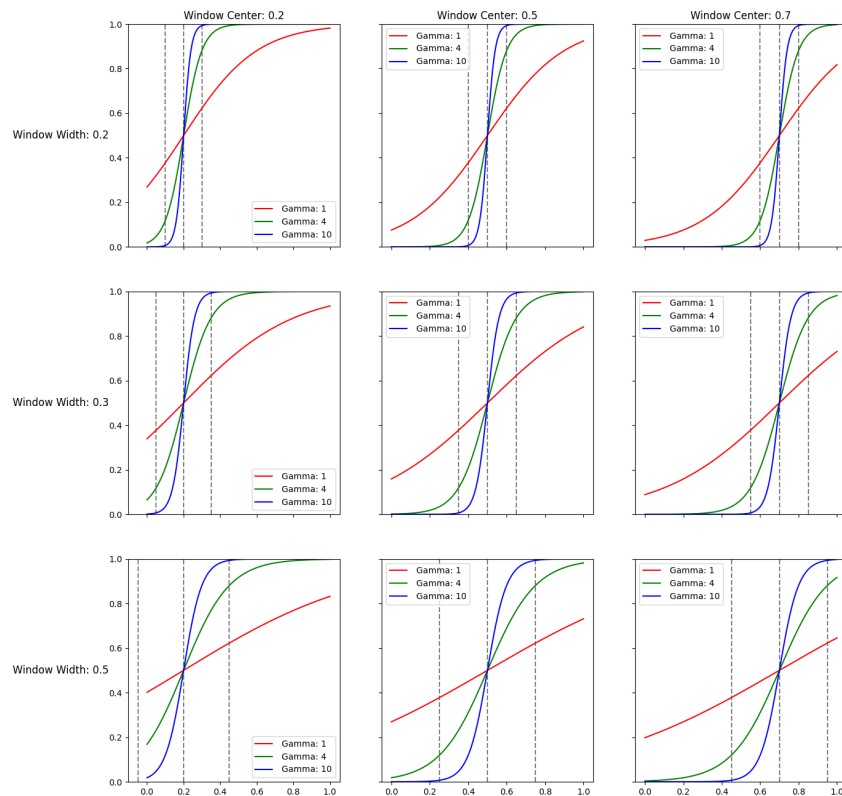


Figure A.1: Windowing function with different parameter settings.

€€€€ For DIVA €€€€

```
{
  "Author1": { "Last name": "Sanchez Torres",
    "First name": "Andres Alam",
    "Local User Id": "u100001",
    "E-mail": "aast2@kth.se",
    "organisation": { "L1": "School of Electrical Engineering and Computer Science",
    }
  },
  "Cycle": "2",
  "Course code": "SF259X",
  "Credits": "30.0",
  "Degree1": { "Educational program": ""
    , "programcode": "COSSE"
    , "Degree": "Degree of Master of Science in Engineering"
    , "subjectArea": "Scientific Computing"
  },
  "Title": {
    "Main title": "Chest X-ray Transmission Map Reconstruction",
    "Subtitle": "Constrained Optimization to Invert a Family of Image Processing Algorithms",
    "Language": "eng" },
    "Alternative title": {
      "Main title": "Detta är den svenska översättningen av titeln",
      "Subtitle": "Detta är den svenska översättningen av undertiteln",
      "Language": "swe"
    },
    "Supervisor1": { "Last name": "Öktem",
      "First name": "Ozan",
      "Local User Id": "u100003",
      "E-mail": "ozan@kth.se",
      "organisation": { "L1": "School of Engineering Sciences",
        "L2": "Mathematics" }
    },
    "Examiner1": { "Last name": "Öktem",
      "First name": "Ozan",
      "Local User Id": "u1d13i2c",
      "E-mail": "ozan@kth.se",
      "organisation": { "L1": "School of Engineering Sciences",
        "L2": "Mathematics" }
    },
    "National Subject Categories": "10201, 10206",
    "Other information": { "Year": "2025", "Number of pages": "1,36" },
    "Copyrightleft": "copyright",
    "Series": { "Title of series": "TRITA – EECS-EX", "No. in series": "2024:0000" },
    "Opponents": { "Name": "A. B. Normal & A. X. E. Normalè",
    "Presentation": { "Date": "2022-03-15 13:00"
      , "Language": "eng"
      , "Room": "via Zoom https://kth-se.zoom.us/j/ddddddd"
      , "Address": "Isafjordsgatan 22 (Kistagången 16)"
      , "City": "Stockholm"
    },
    "Number of lang instances": "2",
    "Abstract[eng ]": €€€€
  }
}
```

An abstract is (typically) about 250 and 350 words (1/2 A4-page) with the following components:

- What is the topic area? (optional) Introduces the subject area for the project.
- Short problem statement
- Why was this problem worth a Bachelor's/Master's thesis project? (*i.e.*, why is the problem both significant and of a suitable degree of difficulty for a Bachelor's/Master's thesis project? Why has no one else solved it yet?)
- How did you solve the problem? What was your method/insight?
- Results/Conclusions/Consequences/Impact: What are your key results/ conclusions? What will others do based on your results? What can be done now that you have finished - that could not be done before your thesis project was completed?

```
€€€€,
"Keywords[eng ]": €€€€,
Nonlinear Optimization, Medical Imaging, Digital Image Processing €€€€,
"Abstract[swe ]": €€€€
```

Enter your Swedish abstract or summary here!

```
€€€€,
"Keywords[swe ]": €€€€
```

Canvas Lärplattform, Dockerbehållare, Prestandajustering €€€€,

}



acronyms.tex

```
%%% Local Variables:
%%% mode: latex
%%% TeX-master: t
%%% End:
% The following command is used with glossaries-extra
\setabbreviationstyle{acronym}{long-short}
% The form of the entries in this file is \newacronym{label}{acronym}{phrase}
%                                     or \newacronym[options]{label}{acronym}{phrase}
% see "User Manual for glossaries.sty" for the details about the options, one example is shown below
% note the specification of the long form plural in the line below
\newacronym[longplural={Debugging Information Entities}]{DIE}{DIE}{Debugging Information Entity}
%
% The following example also uses options
\newacronym[shortplural={OSes}, firstplural={operating systems (OSes)}]{OS}{OS}{operating system}

% note the use of a non-breaking dash in long text for the following acronym
\newacronym{IQL}{IQL}{Independent Q28091Learning}

\newacronym{KTH}{KTH}{KTH Royal Institute of Technology}

\newacronym{LAN}{LAN}{Local Area Network}
\newacronym{VM}{VM}{virtual machine}
% note the use of a non-breaking dash in the following acronym
\newacronym{WiFi}{Wi28091Fi}{Wireless Fidelity}

\newacronym{WLAN}{WLAN}{Wireless Local Area Network}
\newacronym{UN}{UN}{United Nations}
\newacronym{SDG}{SDG}{Sustainable Development Goal}
```