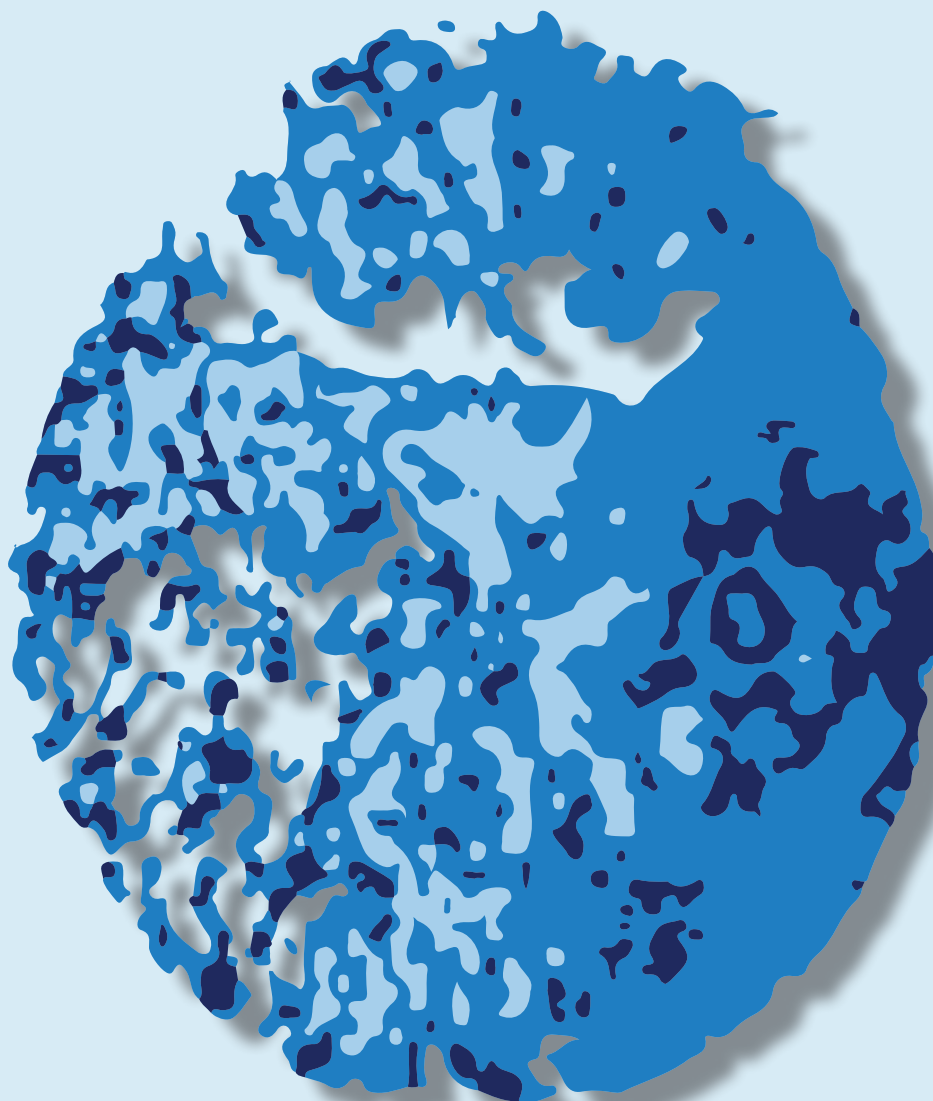


Low dose CT perfusion:

filtering the way to the clinic



Sjoerd A. M. Tunissen

**RADBOD
UNIVERSITY
PRESS**

Radboud
Dissertation
Series

Low dose CT perfusion

filtering the way to the clinic

Sjoerd A.M. Tunissen

Low dose CT perfusion: filtering the way to the clinic

Sjoerd Tunissen

Radboud Dissertation Series

ISSN: 2950-2772 (Online); 2950-2780 (Print)

Published by RADBOUD UNIVERSITY PRESS

Postbus 9100, 6500 HA Nijmegen, The Netherlands

www.radbouduniversitypress.nl

Design: Proefschrift AIO | Annelies Lips

Cover: Proefschrift AIO | Guntra Laivacuma

Printing: DPN Rikken/Pumbo

ISBN: 9789465150406

DOI: 10.54195/9789465150406

Free download at: <https://doi.org/10.54195/9789465150406>

© 2025 Sjoerd Tunissen

**RADBOUD
UNIVERSITY
PRESS**

This is an Open Access book published under the terms of Creative Commons Attribution-Noncommercial-NoDerivatives International license (CC BY-NC-ND 4.0). This license allows reusers to copy and distribute the material in any medium or format in unadapted form only, for noncommercial purposes only, and only so long as attribution is given to the creator, see <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

Low dose CT perfusion

filtering the way to the clinic

Proefschrift ter verkrijging van de graad van doctor
aan de Radboud Universiteit Nijmegen
op gezag van de rector magnificus prof. dr. J.M. Sanders,
volgens besluit van het college voor promoties
in het openbaar te verdedigen op

woensdag 21 mei 2025

om 10.30 uur precies

door

Sjoerd Antonius Marianne Tunissen

geboren op 11 november 1995

te Weert

Promotoren:

Prof. dr. I. Sechopoulos

Prof. dr. W.M. Prokop

Copromotor:

Dr. ir. E.J. Smit

Manuscriptcommissie:

Prof. dr. ir. N.J.J. Verdonschot

Prof. dr. M. Kachelrieß (Deutsches Krebsforschungszentrum, Duitsland)

Dr. ir. E.P.A. Vonken (Universitair Medisch Centrum Utrecht)

Low dose CT perfusion

filtering the way to the clinic

Dissertation to obtain the degree of doctor

from Radboud University Nijmegen

on the authority of the Rector Magnificus prof. dr. J.M. Sanders,

according to the decision of the Doctorate Board

to be defended in public on

Wednesday, May 21, 2025

at 10:30 am

by

Sjoerd Antonius Marianne Tunissen

born on November 11, 1995

in Weert

PhD supervisors:

Prof. dr. I. Sechopoulos

Prof. dr. W.M. Prokop

PhD co-supervisor:

Dr. ir. E.J. Smit

Manuscript Committee:

Prof. dr. ir. N.J.J. Verdonschot

Prof. dr. M. Kachelrieß (German Cancer Research Center, Germany)

Dr. ir. E.P.A. Vonken (University Medical Center Utrecht)

Table of Contents

General Introduction		9
Chapter 2	A CT simulation tool	25
Chapter 3	Image domain lower dose simulations	61
Chapter 4	Characterization of the 4D Similarity Filter	93
Chapter 5	GPU-based implementation of the 4D Similarity Filter	125
Chapter 6	4D Similarity Filter for ischemic stroke	137
General Discussion		205
Bibliography		215
Summaries	Engels	226
	Nederlands	227
Appendix	List of Publications	232
	Portfolio	233
	Research Data Management	234
	Curriculum Vitae	235
	Dankwoord	236



General Introduction

Computed Tomography

Computed tomography (CT) is an imaging technique that allows for the visualization of structures inside the human body. This non-invasive technique employs x-ray photons, a form of high-energy light, strong enough to pass, partly, through a human body.

During a CT scan, x rays are emitted from a source towards the specific area of the body being imaged. As they travel, most of the x rays interact with tissues in the body. The number of x rays that will interact is dependent on the energy of the x ray and the linear attenuation of the tissue the x ray is traveling through. The linear attenuation of a tissue is dependent on its density and elemental composition.

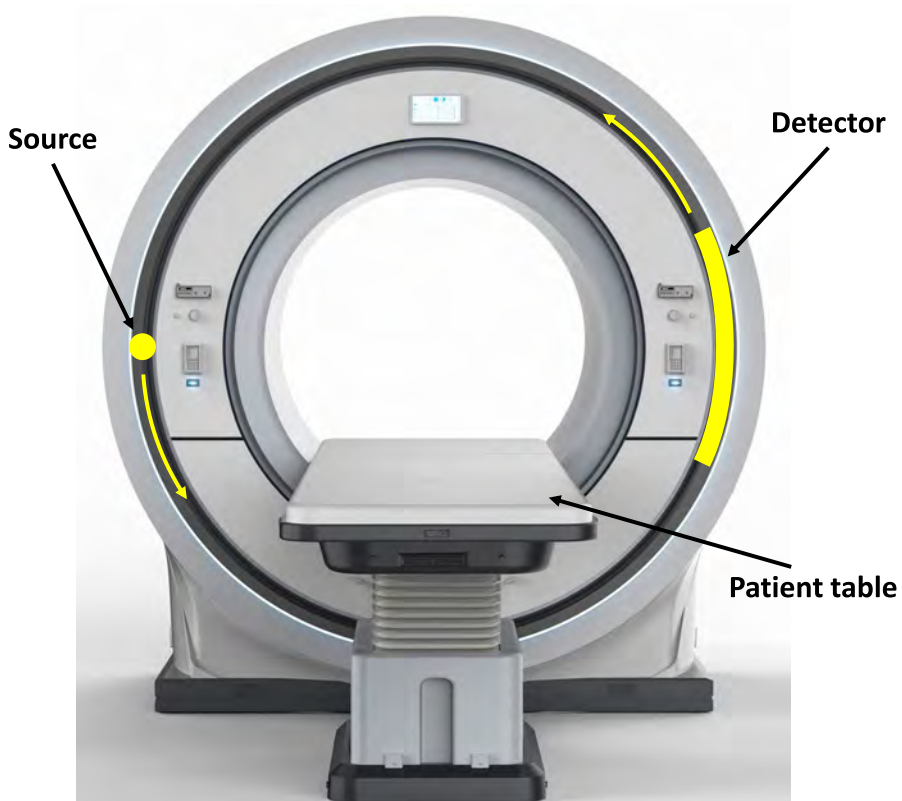


Figure 1: Illustration of a CT scanner, which shows the position and direction of movement of each component.

On the opposite side of the body, the x rays that are not stopped by the body will be detected and measured by a detector, creating an x-ray image or x-ray radiograph. In CT, this process is repeated at various angles while the source and detector rotate around the patient. Finally, a method called reconstruction combines these individual images into a cross-sectional image, creating a CT scan, providing valuable insights for diagnosis and treatment. The values in a CT image are based on the attenuation coefficient of the materials present in the voxel, and are called CT numbers, which are expressed in Hounsfield units (HU). Figure 1 provides a schematic overview of a CT acquisition and the CT systems most important components.

Currently, CT imaging is one of the most common imaging modalities in hospitals. Its prevalence continues to grow every year [1–7]. This widespread adoption can be ascribed to a number of reasons; its versatility in imaging a wide range of diseases and anatomical structures, its rapid image acquisition, and its high spatial resolution [8]. In the following paragraphs, we will delve deeper into each of them, explaining how they have helped CT move to the forefront of medical imaging.

The ability of CT to provide cross-sectional images of the body's interior, while being non-invasive, has made it a fundamental imaging tool for a wide variety of applications. It can be used for diagnosis, disease staging, and treatment response monitoring. Its utility extends even beyond these applications, since CT can even be used for image-guided biopsy or image-guided treatment.

Over the last few decades, the hardware components of CT have advanced drastically, resulting in faster scan times, higher spatial resolution, lower noise, and lower dose. Two pivotal advancements were the introduction of helical CT [9], which enables continuous scanning of the patient while the patient table is moved through the scanner, and multidetector CT [10], which uses a detector with multiple rows in the axial direction of the patient to capture a large volume of data in a single rotation. These two innovations drastically reduced scan time and motion artifacts due to the patient moving during acquisition of the CT scan. Nowadays, multidetector CT scanners can cover a volume of up to 16 cm in a single rotation [11], which takes less than one second, making CT an extremely fast-imaging technique.

An also important development over the past decades is the reduction in detector pixel size [8], which has led to high spatial resolution in CT. The in-plane resolution, or so-called slice thickness, of a CT image can nowadays be as small as 0.2 mm [12], allowing the radiologist to detect very small structures.

The short scan time and relatively little dependency on operator skills of modern CT scanners have made CT a consistent imaging modality. The scan time of CT is now so short that the negative influence of patient motion has decreased substantially, leading to more accurate images. Although the patient must be positioned properly and the protocol selection remains essential, CT imaging is less reliant on operator expertise compared to other modalities, like ultrasound. This reliability is one of the essential factors for the widespread adoption of CT as a diagnostic imaging tool.

Despite its benefits, CT imaging does have some limitations. The primary disadvantage of CT, or any x-ray-based imaging modality, is its use of ionizing radiation, which may be harmful to the patient, since it could potentially lead to cancer. This is a larger concern when scanning young patients [13,14]. Additionally, contrast in CT images is primarily based on tissue density and elemental composition differences, therefore, in principle, providing only anatomical and morphological information of the body. Thus, lacking the ability to provide functional information, such as, for instance, metabolism or blood flow. This also makes differentiation between organs challenging as they often have similar densities. Furthermore, the information on a CT image is static, since the scan is basically a snapshot in time, and therefore will not provide any information about dynamic processes in the body.

To enhance the diagnostic value of CT images, an iodinated contrast agent can be injected into the patient's bloodstream. This technique is called contrast-enhanced CT or CECT and will be explained in the next section.

Contrast-enhanced CT

Contrast-enhanced CT is a technique that uses the high attenuation coefficient of iodine in CT contrast agents that are injected into the bloodstream to demonstrate differences in perfusion between structures and within organs. Normally, the iodinated contrast is injected intravenously into the bloodstream where it will mix with the blood and flow via the heart to the arteries, organs, and then via the veins back to the heart. The timing of the CECT scan thus determines which structures and tissues are enhanced, see Figure 2. Depending on the medical question for the scan, images will be taken at a different time points after contrast injection.

When timed correctly, it is possible to depict a range of anatomies and pathologies. For example, aneurysms can be made visible [15], and the visibility of lesions can also

be improved. This ability to not only differentiate tissues based on density, but also on blood flow-related information significantly expands the diagnostic capability of CT.

CECT still only provides one static image or a snapshot in time. However, as mentioned earlier, the iodine flows with the blood throughout the body. By sampling repeatedly and rapidly, the blood flow itself could be estimated as well. This technique is called CT perfusion imaging.

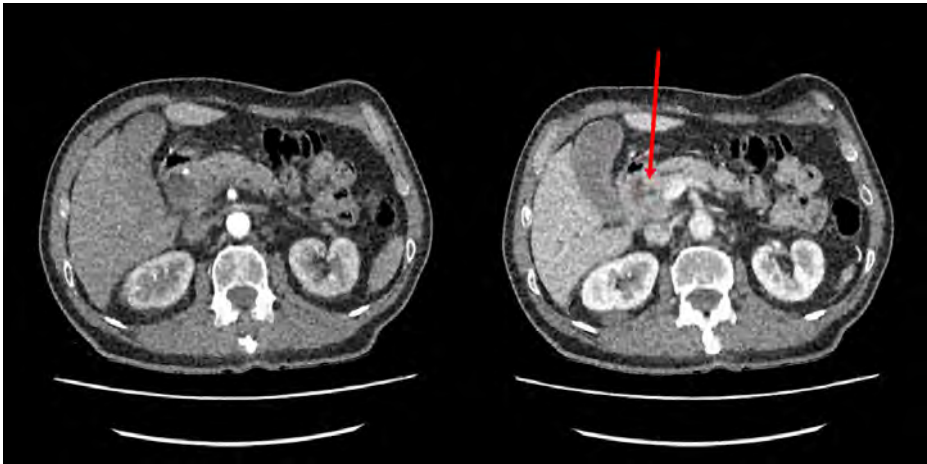


Figure 2: (left) An early arterial phase abdomen CT scan, showing the enhanced arteries, and (right) a later, venous phase abdomen CT scan, of the same patient, showing a dilated common bile duct and pancreatic duct resulting from a subtle hypodense tumor in the pancreatic head.

CT perfusion imaging

CT perfusion imaging is a dynamic imaging technique that provides information about blood flow through tissues. By acquiring multiple CT scans in a short period of time while iodinated contrast flows through the body, this technique can track the changes in tissue CT number or attenuation over time. The temporal information obtained in this way can be represented in graphs of the CT numbers over time, or so-called time attenuation curves (TACs), see Figure 3. These TACs can be processed to gain information about various parameters characterizing the perfusion of a tissue, such as the blood volume and blood flow. This functional information gives us insight into the physiological processes in the body.

To extract blood flow and blood volume estimates from the acquired CT image series the tissue residue function (RF) can be determined. The tissue RF tells us how the tissue is responding to the blood it receives.

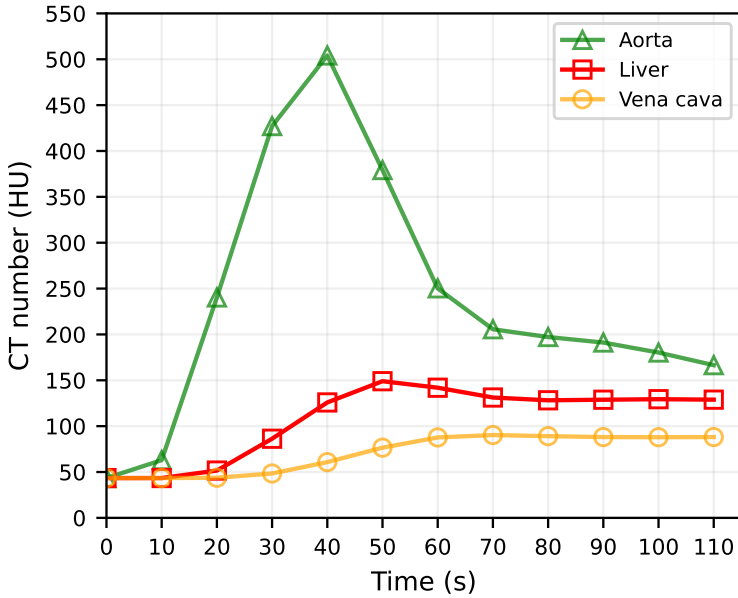


Figure 3: Example of time attenuation curves of the aorta, liver, and vena cava. The markers indicated the scan times.

The tissue RF cannot be estimated directly from the observed TAC of a tissue, since this TAC is a combination of the tissue RF and the TAC of the feeding artery, called the arterial input function (AIF). The AIF depends, among other factors, on the injection protocol and patient anatomy and physiology, whereas the tissue RF is only influenced by tissue properties. To solve this problem, it is assumed that the observed TAC of the tissue is the result of the convolution of the AIF with the tissue RF. So, by an analytical approach called deconvolution the tissue RF can be obtained. To perform the deconvolution, a mathematical technique called singular value decomposition is commonly used [16]. An alternative method to determine the tissue RF is by using a Bayesian estimation approach in combination with a vascular model. The vascular model explicitly models the capillary flow and thus the microvasculature flow dynamics. The Bayesian estimation is then used to determine the probability distribution for all model parameters [17].

With either approach, information about the tissue and vessel perfusion is determined from the estimated tissue RF and summarized in so-called perfusion maps. Several perfusion maps are generated from this analysis, including maps displaying, for example, blood volume, blood flow or mean transit time. These maps allow for easy or more intuitive visualization of functional information, see Figure 4.

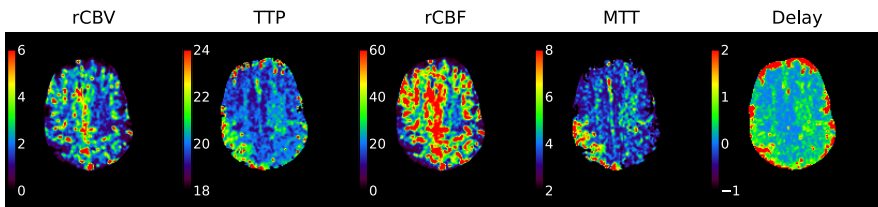


Figure 4: Five cerebral perfusion maps, all depicting different perfusion properties of the brain.

The perfusion characteristics of healthy tissue are different from those of abnormal tissue and thus making it possible to detect and characterize potential abnormalities.

One application of CT perfusion is brain imaging. When a person suffers from an ischemic stroke, the blood supply to a part of the brain is blocked. This results in a region where the blood supply is reduced in comparison to the rest of the brain. This affected region consists of a part in which the brain tissue is irreversibly damaged, i.e., the infarct core, and a part in which the tissue is salvageable and can recover after normalization of the blood supply, i.e., the penumbra.

Studies like the ones performed by Jovin et al. [18] and Albers et al. [19] have shown the potential of functional information when it comes to treating stroke. In these studies, treatment of stroke patients with symptom onset more than six hours ago was based on a mismatch between infarct core and stroke symptoms or infarct core and affected tissue. A mismatch suggests the presence of salvageable tissue, i.e., penumbra, in the brain. If a mismatch was present, patients got treated with thrombectomy even though onset of symptoms was more than six hours ago. These results showed an improved outcome, confirming the added benefit of functional imaging.

The work of Kim et al. [20] provided a clear overview of the benefits of CT perfusion in liver imaging. Currently, different molecular-targeted agents for the treatment of cancer are being developed. One type of drugs is antiangiogenic, designed to hamper the formation of new blood vessels. Since blood vessels are necessary for a tumor to provide itself with oxygen and nutrients, antiangiogenic drugs negatively impact tumor growth. However, the current clinical standard for assessing treatment response in oncology is mainly based on tumor size [21,22]. Antiangiogenic drugs target the formation of new blood vessels, an effect that does not necessarily impact tumor size directly [23]. Therefore, the addition of functional imaging, in the form of CT perfusion, to provide information on the vascularity of

the tumor could potentially help to find early imaging biomarkers for treatment response [24]. Antiangiogenic drugs are usually relatively expensive and can cause serious side effects, so early assessment of treatment response is important. In this way, the potential side effects and costs can be minimized for non-responders, which is beneficial for patients and the healthcare system.

While CT perfusion imaging holds the potential to be a powerful tool for this clinical application, it has two challenges. The first one is the assumption that the same volume of the patient is imaged in all CT scans acquired during the protocol. However, patient movement is inevitable in many cases and can cause inconsistencies in the data. To minimize the effect of movement between scans, registration algorithms have been developed to align and correct for these movements [25–27].

The second challenge, and the one this thesis focusses on, is the total radiation dose required for a CT perfusion protocol, because of the high number of CT scans involved. This relatively high radiation dose raises concerns about long-term risks and thus limits the clinical impact of CT perfusion, especially in case of patients with a good long-term prognosis. This challenge will be discussed in more detail in the next section.

Image quality vs Radiation dose

In CT imaging, image quality and radiation dose are inherently intertwined. This is due to the physics of x-ray imaging: a higher number of x-ray photons leads to lower image noise, i.e., a higher image quality. However, at the same time, a higher number of photons results in a higher radiation dose to the patient, which may be harmful. Because of this relationship, the 'As Low As Reasonably Achievable' or in short, ALARA principle is normally followed [28]. This principle says that all x-ray-based imaging should be performed such that the radiation dose administered to the patient is the minimum necessary to achieve the image quality needed for an accurate diagnosis.

CT perfusion is an imaging technique in which radiation dose plays a big role, since successive scanning leads to a total radiation dose defined by the dose per scan and the number of scans. Thus, there are two ways to keep the dose low: reducing the total number of scans acquired in a protocol and decreasing the radiation dose per individual scan.

Several efforts have been made to obtain relevant functional information while keeping the total number of scans low. One of the most notable involves performing triple-phase imaging, which are protocols performing only three CT scans over time. This technique can be used to generate simple perfusion maps of the liver, such as arterial enhancement fraction, which has been shown to closely correlate with the hepatic perfusion index obtained with a full CT perfusion protocol in an animal study [29]. The work of Kim et al. [30] and Mahnken et al. [31] showed the potential benefit of triple-phase imaging in combination with functional maps for detection and recurrence prediction of liver tumors. However, this technique falls short in providing detailed perfusion information. As explained in the previous section, the tissue RF is determined by deconvolution of the AIF, or Bayesian estimation, with the time-concentration curve of the tissue of interest. To perform this deconvolution and obtain the functional information, high temporal sampling is needed, which is not the case in triple-phase imaging.

The second way to keep the radiation dose in CT perfusion low is to decrease the dose per individual scan. This can be achieved through advancements in both hardware and software. Software-based developments hold significant but underutilized potential for reducing radiation dose while preserving the diagnostic quality of the images.

A current focus area of software development to reduce radiation dose is the development of new reconstruction methods. Traditionally, filtered back projection (FBP) was the standard reconstruction method. In the case of FBP, the spatial resolution and the noise are completely independent. However, newer reconstruction methods, such as iterative reconstruction methods and lately deep learning reconstruction, have now become the clinical standard. These methods are especially designed to reduce image noise by sacrificing some spatial resolution. With these new techniques, the independence between spatial resolution and radiation dose does not hold anymore.

There are many different types of iterative reconstructions, the most common one being iterative reconstructions with total variation regularization. Another subset of iterative reconstruction methods is called model-based iterative reconstructions (MBIR) [32]. MBIR methods use prior knowledge of the imaging process, by incorporating models that, for example, contain the noise statistics, detector characteristics, or even information on the expected anatomy. All these models provide additional information to the reconstruction algorithm, so it can result in improved image quality.

Deep learning (DL) has also been introduced to the field of reconstruction methods, offering even more opportunities to enhance image quality and thus reduce radiation dose. One way of implementing DL is in an iterative scheme, called primal-dual reconstruction [33]. In this scheme, DL blocks are used to smoothen the data without losing detail by processing it in both the projection and reconstruction domain. Another way of incorporating DL in reconstruction is in the case of sparse angular coverage. Sparse angular coverage means that the angular step between two projections is large, which can be used to reduce radiation dose. In this case, DL can be used to interpolate the projection data and thus artificially increase the angular sampling. Such a network can be trained on paired data of densely- and sparsely-sampled angular projections to train the DL network to interpolate the missing data [34].

A significant advantage of DL reconstruction over MBIR is the computational time. The computation-heavy nature of MBIR means that it is usually associated with increased computational time. However, DL reconstruction, once trained, offers fast and high-quality image reconstruction. This is a crucial advantage of DL, since in a clinical setting it can improve workflow and reduce total time for diagnosis.

In addition to reconstruction methods, there are also post-processing methods developed to enhance the image quality of low dose CT scans. DL is also a topic of interest in this field, with a plethora of network designs aimed at denoising CT data [35,36]. These networks are typically trained on paired low- and high-dose CT data, allowing the network to learn how to reduce noise while limiting the loss of spatial resolution, and thus of small details.

Alongside DL, classical image processing techniques, like bilateral filters, anisotropic diffusion filters, k-means clustering, or the Canny filter continue to play a vital role in enhancing image quality. These techniques remain particularly relevant in the case of CT perfusion, due to this modality commonly comprising of 4-dimensional data, which therefore involves a large amount of data. DL networks are not yet capable of processing this amount of data at once, due to computational and memory constraints. One could work around this constraint by, for instance, processing each CT scan separately, but then the temporal information cannot be used to aid the denoising.

Examples of classical image processing methods for denoising CT perfusion data are the work of Mendrik et al. [37]. In this work, a modified version of a bilateral filter was developed, called TIPS, to reduce image noise. The TIPS filter performs

weighted averaging of the TACs in a spatial neighborhood based on similarity and spatial distance between the TACs. Pisana et al. [38] proposed an alternative to this method in which a guiding image, spatial distance, and k-means clustering based on TACs is used to perform weighted averaging of TACs.

A common challenge for all these efforts, which some have tried to minimize, is the loss of spatial resolution. Denoising almost always comes with loss of resolution since high frequency content is sacrificed to suppress noise. This thesis focuses on addressing this issue, aiming to denoise low dose CT perfusion images with no or minimal resolution loss. The primary method evaluated and improved for this purpose is the so-called 4D similarity filter (4DSF), which will be introduced in the next section.

4D Similarity Filter

The 4DSF is a novel noise reduction method developed specially to reduce noise in image time series, such as 4-dimensional CT data, i.e., CT perfusion. Unlike conventional image filters, the 4DSF avoids the use of a spatial neighborhood, thus preserving spatial resolution.

The novelty of the 4DSF is twofold. First each CT image of the temporal sequence is filtered separately, this principle will be explained in the following paragraphs.

Let's assume we are given a 4D CT acquisition with tissues that have different TACs, so not all TACs in the set have the same underlying noiseless ground truth, but all TACs are corrupted by noise. When denoising one of the TACs, this TAC in question will be compared to all other TACs in the set and will be averaged with the N most similar TACs, because these N TACs have the highest chance of having the same underlying ground truth, and thus to be of the same tissue.

However, if the set of TACs is very large, and the number of timepoints in the TACs is not infinitely large, it is highly likely that when looking for the N most similar TACs, the algorithm will not only find curves with the same underlying ground truth but also a similar noise realization. This will cause the noise in the N most similar TACs to be correlated, which will diminish the impact of the noise reduction and potentially introduce a bias. To overcome this problem, the TACs can be filtered separately for each timepoint, so that search and averaging are decoupled.

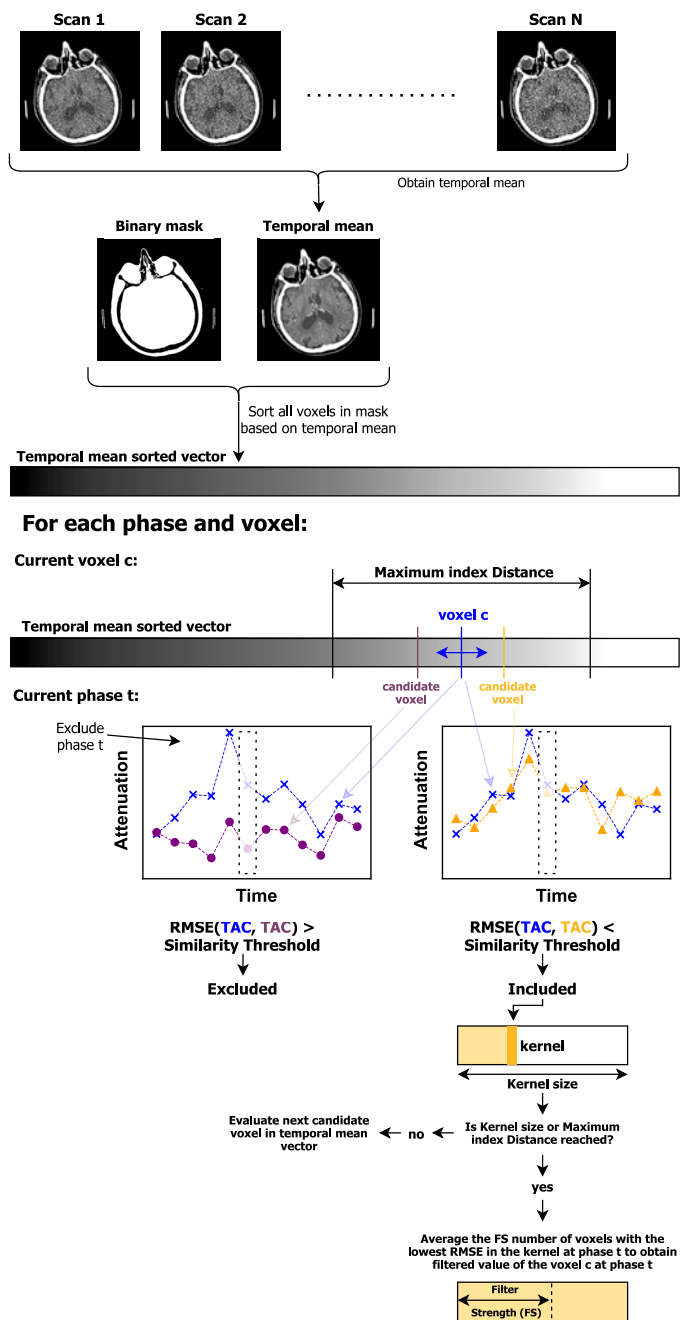


Figure 5: Schematic representation of the 4DSF, describing the steps of the filter. The binary mask is optional, but helps to reduce the number of TACs to be processed. Adapted from S. A. M. Tunissen et al., "Performance evaluation of a 4D similarity filter for dynamic CT angiography imaging of the liver", Med. Phys., doi.org/10.1002/mp.17394

To avoid this correlation, the 4DSF works as follows: when the first timepoint, i.e., the first attenuation value of the TAC, is filtered, the N most similar TACs will be determined by checking only the most similar TACs based on all other timepoints except for the first timepoint. By excluding the timepoint that is filtered from the search of similar TACs, the N TACs cannot have similar noise at the excluded timepoint. The last step in this process is to average the values of the most similar TACs only for the timepoint being filtered, so the first timepoint, hence avoiding bias. This search and averaging process is then repeated for all other timepoints.

The second novelty of the 4DSF is based on the fact that in a standard CT image the number of voxels, and thus TACs, is typically around tens of millions. In this case, computational power and time do not allow us to compare all TACs in the given set to each other. As described previously, other filtering techniques limit themselves to a spatial neighborhood, resulting in spatial resolution loss. Since 4DSF aims to reduce noise without or with minimal resolution loss, the use of a spatial neighborhood is avoided. Instead, all TACs are sorted based on their average attenuation value, i.e., temporal mean. A neighborhood based on this sorted vector, around the TAC being filtered, is then used to limit the number of calculations. A schematic representation of the 4DSF is presented in Figure 5.

Quantitative evaluation of low dose imaging

Thorough evaluation of a newly developed post-processing or reconstruction method is of utmost importance before its integration into clinical practice. For assessment of quantitative accuracy, the ground truth of the condition being studied needs to be known. Of course, patient scans do not have ground truth available, since noiseless versions of these scans are not available. Also, clinical scans need to have sufficient diagnostic value, so their image quality is normally relatively high. A common approach is to take these high-quality scans and modify them to degrade their quality, so a high and low image quality pair is created.

The most suitable methods in CT imaging to do so are the ones that operate in the projection domain, i.e., the radiographs from which the CT image is reconstructed. Notable examples of these methods are the work of Yu et al. [39] and Žabić et al. [40], who presented methods that model the noise in the projection, so that realistic noise can be added. However, since these methods work in the projection domain, they assume access to the projection domain data and the ability to reconstruct the modified data.

An alternative approach to create high- and low-quality image pairs involves algorithms that operate directly in the image domain, i.e., on the CT image itself. One example is the work of Divil et al. [41], in which an analytical method is proposed to generate noise to degrade the image quality. However, this method is only valid for analytical, and thus linear, reconstruction methods, such as FBP. With the introduction of more sophisticated reconstruction methods such as MBIR and DL reconstructions, which are normally not analytical, the application of this method becomes limited.

The limiting factor for all these methods is that only the noise level of existing data can be modified. If one wants to test a new scanning protocol with, for instance, a different tube voltage or number of scans, this is not possible with these methods.

An alternative in these cases would be to scan phantoms. The advantage of phantoms is that they can be scanned at very high dose levels without radiation dose problems, enabling the acquisition of near-perfect, "ground truth" images. For the same reason, phantoms can be scanned at different dose levels and scanner settings, providing a wide variety of test data. The main disadvantage of phantom data is their limited anatomical and functional realism.

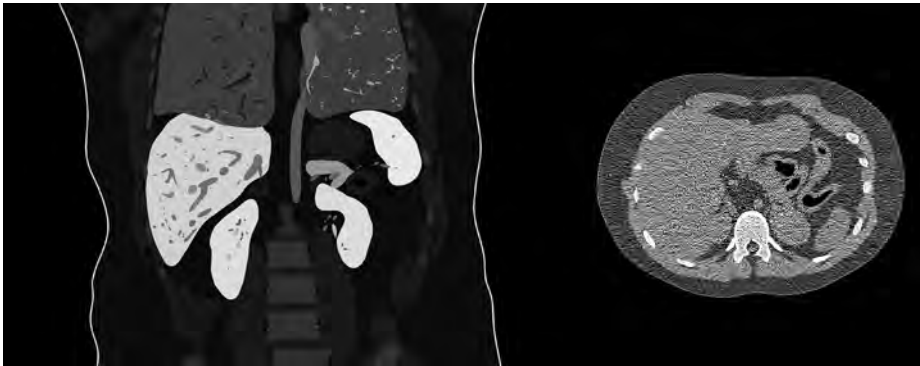


Figure 6: Example of a coronal plane view of an XCAT phantom (left), the intensities only indicate different tissues, not any physical property. A simulated CT image in the axial plane of the XCAT phantom (right).

To partly overcome this disadvantage efforts have been made in the field of digital phantoms, which are models of human bodies. One example is the work of Segars et al. [42], in which a set of digital human-based phantoms is developed. These phantoms are anatomically realistic and can, in combination with a CT simulator, provide CT images for testing purposes, see Figure 6. Since the exact shape and

composition of the phantoms is known, the ground truth underlying the CT images will be known. The mentioned CT simulator is a tool that can mimic the processes taking place in a real CT scanner, so realistic CT images of digital phantoms can be obtained [43,44]. The combination of digital phantoms and a CT simulator forms a powerful platform for testing new algorithms, without radiation exposure for patients. During this thesis such a CT simulator was developed and validated.

Aim and outline of thesis

The aim of this thesis was to characterize, validate, and, if possible, improve the 4DSF for its application in CT perfusion imaging. Our research efforts are divided into two distinct parts. The first part focuses on the development of CT image simulators that can generate realistic test data. The second part focuses on the detailed characterization of the 4DSF, exploring its behavior and limitations, ultimately leading to the development and evaluation of strategies for improving its performance.

In the first part, two tools for CT simulations are developed. Chapter 2 presents a scanner-specific CT simulator capable of generating synthetic CT scans from digital phantoms. On top of this, we validated simplifications of the simulation to reduce computation time. The second simulation tool is presented in Chapter 3 and focuses on the simulation of low dose CT images from clinical dose patient CT images when knowledge and access on the CT system and reconstruction are unavailable. The method circumvents the use of projection domain data, which is often hard to obtain, by working directly in image domain. Therefore, the method will also work for non-analytical methods.

The second part of this thesis is focused on the analysis and optimization of the 4DSF. In Chapter 4 we present a digital phantom study in which the 4DSF is characterized, and its use in combination with the TIPS filter is investigated. Finally, a suggestion on a possible implementation for 4D liver CT is given, focusing on lesion detection. Chapter 5 introduces a modified version of the 4DSF for GPU implementation. Since GPUs are much faster than CPUs, the use of the former is intriguing to accelerate the 4DSF. Chapter 6 presents another modified version of the 4DSF tailored for stroke imaging. By incorporating the unique perfusion characteristics of ischemic brain tissue, which receives blood later than healthy brain tissue, this modified version enhances the quantitative accuracy of the functional maps. The modified version is validated using a phantom study and patient cases to demonstrate its potential.



Chapter 2

A CT simulation tool

Original title: Development, validation, and simplification of a scanner-specific CT simulator

Sjoerd A. M. Tunissen, Luuk J. Oostveen, Nikita Moriakov, Jonas Teuwen,
Koen Michielsen, Ewoud J. Smit, and Ioannis Sechopoulos

Published in: Medical Physics

ABSTRACT

Background

Simulated Computed Tomography (CT) images allow for knowledge of the underlying ground truth and for easy variation of imaging conditions, making them ideal for testing and optimization of new applications or algorithms. However, simulating all processes that affect CT images can result in simulations that are demanding in terms of processing time and computer memory. Therefore, it is of interest to determine how much the simulation can be simplified while still achieving realistic results.

Purpose

To develop a scanner-specific CT simulation using physics-based simulations for the position-dependent effects and shift-invariant image corruption methods for the detector effects. And to investigate the impact on image realism of introducing simplifications in the simulation process that lead to faster and less memory-demanding simulations.

Methods

To make the simulator realistic and scanner-specific, the spatial resolution and noise characteristics, and the exposure-to-detector output relationship of a clinical CT system were determined. The simulator includes a finite focal spot size, raytracing of the digital phantom, gantry rotation during projection acquisition, and finite detector element size. Previously published spectral models were used to model the spectrum for the given tube voltage. The integrated energy at each element of the detector was calculated using the Beer-Lambert Law. The resulting angular projections were subsequently corrupted by the detector Modulation Transfer Function (MTF), and by addition of noise according to the Noise Power Spectrum (NPS) and signal mean-variance relationship, which were measured for different scanner settings. The simulated sinograms were reconstructed on the clinical CT system and compared to real CT images in terms of CT numbers, noise magnitude using the standard deviation, noise frequency content using the NPS, and spatial resolution using the MTF throughout the field of view. The CT numbers were validated using a multi-energy CT phantom, the noise magnitude and frequency were validated with a water phantom, and the spatial resolution was validated with a tungsten wire. These metrics were compared at multiple scanner settings, and locations in the field of view. Once validated, the simulation was simplified by reducing the level of subsampling of the focal spot area, rotation and of detector pixel size, and the changes in MTFs were analyzed.

Results

The average relative errors for spatial resolution within and across image slices, noise magnitude, and noise frequency content within and across slices were 3.4%, 3.3%, 4.9%, 3.9%, and 6.2%, respectively. The average absolute difference in CT numbers was 10.2 HU and the maximum was 22.5 HU. The simulation simplification showed that all subsampling can be avoided, except for angular, while the error in frequency at 10% MTF would be maximum 16.3%.

Conclusion

The simulation of a scanner-specific CT allows for the generation of realistic CT images by combining physics-based simulations for the position-dependent effects and image-corruption methods for the shift-invariant ones. Together with the available ground truth of the digital phantom, it results in a useful tool to perform quantitative analysis of reconstruction or post-processing algorithms. Some simulation simplifications allow for reduced time and computer power requirements with minimal loss of realism.

Introduction

Currently, Computed Tomography (CT) is the workhorse imaging modality in most radiology departments [3,45–47]. CT is used for screening, diagnosis, and interventional procedures, such as CT-guided biopsies or ablations [48–52]. Therefore, research on reconstruction and post-processing algorithms to increase image quality in CT, without increasing patient dose, is a growing field of interest. Examples of these efforts include developments in deep learning reconstruction of low dose CT [53], in denoising of low dose CT using Convolution Neural Networks [35,36], and in CT denoising using statistical methods [54,55].

New reconstruction or post processing algorithms are typically developed and validated using physical phantoms, which are limited in their capability to represent real human anatomy. This limits the usefulness of these phantoms when developing new algorithms. Patient images can also be used during development and validation, but aside from the ethical issues if new research-specific acquisitions are needed, patient images do not have a quantitative ground truth available, making it hard to quantify the performance of the developed algorithms [56].

Therefore, it would be beneficial to have the possibility to test these algorithms using virtual clinical trials [57]. In these, computer simulated images are generated from digital models of humans, such as the XCAT phantom [42]. These phantoms have ground truth available and include considerably realistic anatomy, making them ideal for quantitative evaluation of clinically-relevant conditions. In addition, an infinite number of different realizations of the phantoms can be generated and infinite combinations of imaging conditions/parameters can be evaluated, thus the amount of data that can be used for a virtual clinical trial is only limited by computation time and memory.

However, simulating a fully detailed CT image that incorporates all acquisition process characteristics accurately is time and memory consuming. Therefore, for specific tasks it could be beneficial to evaluate the impact on realism of different simplifications that lead to substantially shorter computation time.

Therefore, we aim to develop a simulation of a scanner-specific CT capable of generating CT images with realistic appearance considering the spatial resolution and noise characteristics of a clinical CT system [43,44]. With this scanner-specific CT simulation, a broad range of scanner parameters can be simulated, such as different tube currents, tube voltages, exposure times, bowtie filters, and focal spot sizes. In addition, once a full simulator is developed and validated, we aim to evaluate the

impact on the realism of the resulting images when different simplifications (e.g. less subsampling) of the simulator are introduced, with the aim to reduce the time and computer power necessary to simulate an image.

Methods

The scanner to be simulated in this work is a 320-row CT system (Aquilion ONE PRISM Edition, Canon Medical Systems, Otawara, Japan) installed at the Dept. of Medical Imaging of Radboudumc, Nijmegen, The Netherlands. The scanner-specific CT simulation consists of a pipeline (Figure 1) to generate CT images that are realistic and scanner-specific, based on physics-based simulations for the position-dependent effects and image corruption methods proposed by Saunders et al. for shift-invariant corruption of ideal images [58]. In this way the need for proprietary system-specific information from the vendor is minimized. However, for many of these steps, specific system characteristics need to be known, which, for this work, were obtained via measurements. The details of all steps and these measurements will be discussed in the next sections. A general simulation pipeline, with an overview of required information for each step, can be found in Appendix A (see online supplemental material).

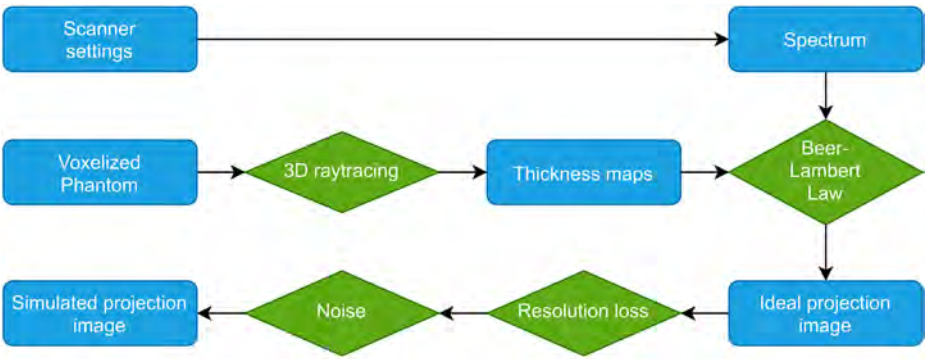


Figure 1: Scanner-specific CT simulation pipeline.

3D raytracing

To perform the 3D raytracing, the specific geometry and dimensions of the clinical CT system, including the detector pixel size, detector distance, focal spot size, focal spot angle, and focal spot distance, were used. However, these are vendor-specific and confidential, so they are not reported here. For the CT system being simulated,

the detector consists of 896 detector channels and 320 rows. The detector is curved such that all pixels in a row have the same distance to the source. The input to the 3D raytracing is a voxelized phantom, representing the object that will be imaged, with the voxel values indicating an index linking it to the material it contains. The 3D raytracing is performed for every material present in the phantom, resulting in a separate thickness map (T) for each material. The raytracing algorithm is a GPU-based pixel-driven raytracing based on the work of Moriakov et al. [59] and Syben et al. [60].

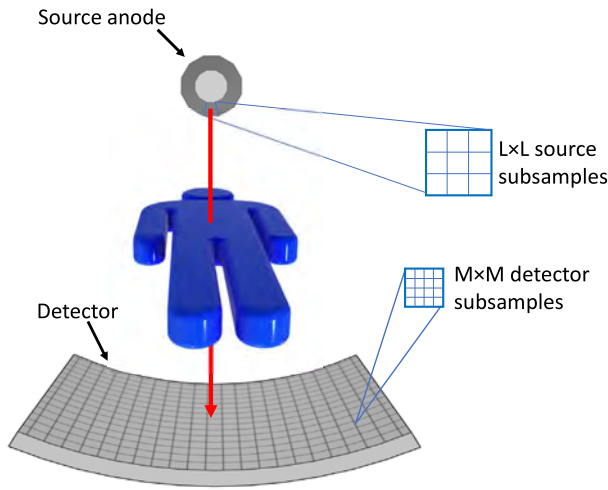


Figure 2: Schematic drawing of the focal spot and detector elements subsampling.

To account for the finite size of the focal spot, the raytracing is not performed from a single point on the focal spot, but from $L \times L$ subsamples of the focal spot, ordered in a square grid, because the shape of the focal spot is approximately square. To minimize the discretization effect on the detector, the detector elements are subsampled by $M \times M$ accordingly and are up-sampled later in the simulation process [58]. To incorporate the effect of the finite exposure time during the angular motion of the CT gantry that causes spatial resolution loss, the angular projections are also subsampled by a factor K . The focal spot and detector subsampling is depicted in Figure 2 and the angular subsampling is depicted in Figure 3.

This results in thickness maps containing the intersection length for each material from each focal spot subsample $f_{a,b}$ to the center of each subsampled detector element (x, y) at each subsampled angular projection θ_k .

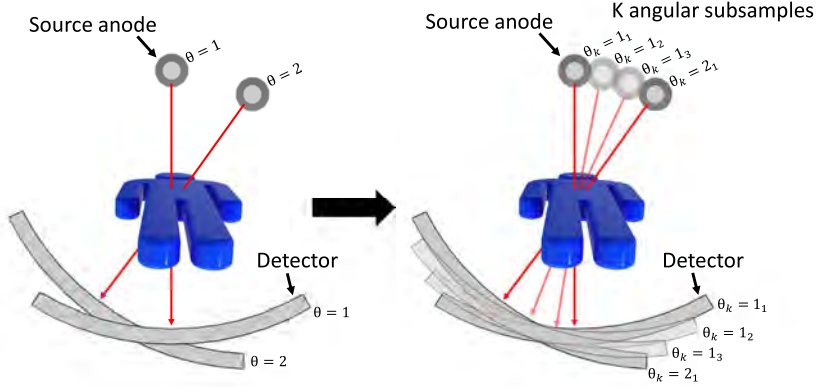


Figure 3: Schematic drawing of the angular subsampling.

To include the effect of the bowtie filters in the simulation, the shape of the bowtie filters was determined using the method described by McKenney et al. [61]. For this, the air kerma was measured with a dosimeter (10X6-0.6CT, Radcal, Monrovia, California, USA) at the center of rotation (CoR) of the CT gantry and outwards in the lateral direction from the central ray, in steps of 5 mm out to 160 mm from the CoR, while the tube remained static. The resulting air kerma measurements were used to estimate the equivalent thickness of the bowtie filter.

Primary projection images

To calculate the incident primary photon energy that is absorbed by the detector, first the x-ray spectrum (N) leaving the source must be determined. The x-ray spectrum is modeled by measuring the incident air kerma as close as possible to the detector (8 cm away) under four different attenuation conditions: no added attenuating materials, 6 mm aluminum, and 1 mm and 2 mm copper. These air kerma measurements are used to fit the spectrum model [62]. Measurements of the same attenuations are used to obtain the conversion from absorbed primary photon energy to Digital Units (DU). This conversion was determined for each bowtie filter separately.

As depicted in Figure 1, the x-ray spectrum is used together with the thickness map of each material to determine the primary photon energy absorbed by the detector, i.e., the simulated sinogram $I(f_{a,b}, x_i, y_j, \theta_k)$ according to the Beer-Lambert law [63]:

$$I(f_{a,b}, x_i, y_j, \theta_k) = \sum_e e * N_e * QE_e * \exp\left(-\sum_m \mu_{m,e} T_{m,f_{a,b}, x_i, y_j, \theta_k}\right) \quad (1)$$

where N_e is the number of photons with energy e emitted from the source, QE_e is the quantum efficiency of the detector for each energy e , $\mu_{m,e}$ is the attenuation coefficient of material m at energy e (determined using the xraydb package in python based on the work of Elam et al. [64], and the work of Boone et al. [65]), and $T_{m,f_{a,b},x_i,y_j,\theta_k}$ is the thickness map of each material m at each subsampled detector element (x_i, y_j) for each focal spot subsample $f_{a,b}$ and at each subsampled projection angle θ_k .

Spatial resolution loss

The spatial resolution characteristics of the detector are incorporated by applying the Modulation Transfer Function (MTF) of the detector to the simulated sinogram resulting from Equation 1. This is done by multiplication, in frequency domain, of the MTF with the 2D Fast Fourier Transform (FFT) of the simulated sinogram (at each subsampled projection angle) and then taking the inverse FFT. Please note that the MTF needs to be divided by the sinc function of the final detector spacing, since the sampling of the detector causes the MTF to be multiplied by the sinc function.

The detector MTF was measured in the detector row direction, since this direction does not suffer from resolution loss due to rotation, and was used in all directions, assuming it is rotationally invariant. The slanted edge method [66] was used with a tungsten edge (TX5, IBA Dosimetry, Schwarzenbruck, Germany). The tungsten edge was placed as close as possible to the detector (8 cm away) to minimize the focal spot size effect. A Lorentzian based fit [67] is used to fit the MTF. The measured edge and fitted MTF are shown in Appendix B. The fit could potentially result in values close to the zero frequency to be larger than one. These are forced to one when applying the MTF.

The other causes of spatial resolution loss, namely, the focal spot size effect and the blur caused by exposure time per angular projection, are already included in the image, as described above, by raytracing the focal spot and angular projections, including subsampling. To maximize the realism of the image simulation, these effects must be included in the raytracing step, since they are position-dependent in the field of view, and therefore this information cannot be added to the sinogram directly.

After the incorporation of the detector MTF, the simulated sinogram is binned to its real dimension, using Equation 2:

$$I(x, y, \theta) = \frac{1}{M^2 * K * L^2} \sum_{i=1}^M \sum_{j=1}^M \sum_{k=1}^K \sum_{a=1}^L \sum_{b=1}^L I(f_{a,b}, x_i, y_j, \theta_k) \quad (2)$$

where $I(x,y,\theta)$ is the sinogram after all subsamples are averaged at detector pixel x, y and angular projection θ , M is the number of detector pixel subsamples in each direction, K is the number of angular subsamples, L is the number of focal spot subsamples in each direction, $I(f_{a,b}, x_f, y_f, \theta_k)$ is the sinogram with all subsamples, $f_{a,b}$ is the focal spot subsample, x_f, y_f is the detector pixel subsample, and θ_k is the angular projection subsample.

Noise addition

To add the correct noise to the sinogram two characteristics of the noise need to be known, the mean-variance relationship of the noise signal and the noise power spectrum (NPS). Both are determined using the same scans of two water phantoms of 240 mm and 320 mm in diameter, representing the attenuation of brain and abdomen, respectively.

The mean and variance of the signal were determined in a 20×30-pixel region of interest (ROI) at approximately the center of each sinogram projection, and their averages over all projections were used as the final mean and variance. To obtain the NPSs, first a correction for image lines due to detector tiling was performed by averaging all projections and subtracting the result from each individual projection. Second, the 2D FFT of a 64×64-pixel ROI in the center of the sinogram was calculated for each projection and the square of the absolute value of these FFTs was taken. The results were averaged for all projections, resulting in 2D NPSs. Despite the anisotropic pixel size there was no significant difference between the NPS in the horizontal and vertical directions, so they were radially averaged to obtain a 1D NPS. The mean-variance relationship is dependent on the tube voltage, and bowtie filter. The shape of the NPS is dependent on the tube current, tube voltage, exposure time, and bowtie filter. Hence, both were measured at nine different tube current levels between 10 mA and 400 mA, four different tube voltage levels, 80 kV, 100 kV, 120 kV, and 135 kV, two different exposure times 0.275 s and 0.5 s, and for two different bowtie filters.

The mean-variance relationship (MV) is defined as a linear function with a positive offset (Equation 3). This offset is the electronic noise.

$$MV = a * m + b_{\text{electronic noise}} \quad (3)$$

where a is the slope of MV , m is the mean in a 3×3 pixel region (Saunders et al. [58]), and $b_{\text{electronic noise}}$ is the offset due to the electronic noise.

To get the desired noise, white noise is generated with similar spatial dimensions as the sinogram projections, as described by Saunders et al. [58]. The resulting noise is multiplied in frequency domain with the square root of the NPS (Equation 4) and is scaled by the MV in spatial domain to obtain the desired noise, which is then added to the sinogram $I(x, y, \theta)$ (Equation 5).

$$N(u, v) = \sqrt{NPS} * \mathcal{F}\{n(\mu = 0, \sigma = 1)\} \quad (4)$$

$$I_{noise}(x, y, \theta) = I(x, y, \theta) + \sqrt{MV(I(x, y, \theta))} * \left(\frac{1}{\sigma(\mathcal{F}^{-1}\{N(u, v)\})} * (\mathcal{F}^{-1}\{N(u, v)\} - \mu(\mathcal{F}^{-1}\{N(u, v)\})) \right) \quad (5)$$

where $\mathcal{F}$ is the FFT operator, n is a realization of white Gaussian noise with mean μ and standard deviation σ , $\mathcal{F}^{-1}$ is the inverse FFT, $N(u, v)$ is the colored noise image with the correct NPS, $\sigma(\mathcal{F}^{-1}\{N(u, v)\})$ is the standard deviation and $\mu(\mathcal{F}^{-1}\{N(u, v)\})$ the mean of the colored noise in image domain after inverse FFT, MV is the mean-variance relationship, I is the primary projection image after the MTF is applied and binned and I_{noise} is the projection image after the noise is added to it.

Hounsfield Unit Calibration

As is standard in CT imaging, a linear calibration was obtained to apply to all reconstructed images to correct the resulting CT numbers for different materials and densities. The applied linear correction was determined by digitally simulating a cylindrical water phantom with 5 different inserts: Teflon, Delrin, acrylic, polypropylene, and air (quality control phantom provided by Canon Medical Systems) and fitting the CT numbers of the simulation to the theoretical CT numbers. The diameter of this phantom is 190 mm, and the inserts have a diameter of 20 mm. A linear correction was obtained from the mean HU of the simulated phantom inserts and water background and their corresponding theoretical values, using Equation 6.

$$\min_{a \in \mathbb{R}} (HU_{theoretical} - a * (HU_{simulation} + b_{water}))^2 \quad (6)$$

The value of b_{water} (offset of water) was fit such that the simulated water value corresponds to the theoretical one, i.e., equal to zero. Afterwards a (slope) was fit such that the $HU_{simulation}$ after correction, had the smallest possible error against the corresponding $HU_{theoretical}$. This was done separately for each tube voltage level available in the system, and the corresponding calibration was then applied to all subsequent simulated images.

Validation of simulation

To assess the accuracy of the simulator, multiple validations were performed to validate the CT numbers of different materials, the spatial resolution, and the noise characteristics of the simulated images against images acquired with the clinical CT system. All validations were performed after reconstruction of the sinogram projections on the clinical CT system using the clinically available filtered back projection (FBP), which is based on the Feldkamp Davis Kress (FDK) algorithm [68]. For the CT number and noise characteristics validation, the number of angular subsamples K was set to 2, the number of focal spot subsamples L was set to 1, and the number of detector subsamples M was set to 2. For the resolution loss validation, the number of angular subsamples K was set to 3, the number of focal spot subsamples L was set to 3, and the number of detector subsamples M was set to 4. These subsampling factors were obtained experimentally, the details can be found in Appendix C.

The CT numbers were validated using a physical oval phantom (with 40 cm and 30 cm radii for the horizontal and vertical directions, respectively) with 15 cylindrical inserts, each of different material and of diameter 28.5 mm [69] (MECT phantom, Sun Nuclear, Middleton, WI, US). The exact dimensions and material composition of the MECT phantom were known, so we could not only image but also simulate the phantom and its image acquisition, with a tube current of 400 mA and three different tube voltage levels (100 kV, 120 kV, and 135 kV). The simulated voxel size of the phantom was $3.3 \text{ mm} \times 0.25 \text{ mm} \times 0.25 \text{ mm}$. The voxel size in the longitudinal direction was substantially larger since the phantom is constant in this direction. The measured and simulated sinograms were both reconstructed on the clinical CT system using FBP and a Field of View (FOV) of $320 \text{ mm} \times 320 \text{ mm}$ and 160 mm in the longitudinal direction. The reconstructed volume consisted of 320 slices of 512×512 pixels. The Hounsfield Units (HU) within these inserts and in the water-equivalent background was measured by averaging a squared 10×10 pixel ROI across 80 slices.

The resolution loss of the simulator was validated by imaging a $50 \text{ }\mu\text{m}$ diameter tungsten wire [70,71], both digitally and physically. This tungsten wire creates a Dirac delta function or unit impulse [72], and the point-spread function (PSF) is obtained by taking the Radon transform [73] of this signal in one direction. The MTF is then determined by calculating the FFT of the PSF. The spatial resolution was validated at 7 cm, 14 cm, and 21 cm from the isocenter, for both the digitally simulated wire and the real physical measured wire to verify the validity of the simulation of the shift-variant rotational blur and focal spot size effects. Each simulated and measured wire was reconstructed with a small FOV of $19.5 \text{ mm} \times 19.5 \text{ mm}$

of 512×512 pixels, so the PSF had enough samples. The resolution loss was checked for both focal spot sizes present in the clinical system, which will be referred to as large and small focal spots from here on, and for both the radial and tangential direction for all positions. The simulated voxel size of the phantom was $0.1432 \text{ mm} \times 0.005 \text{ mm} \times 0.005 \text{ mm}$. The voxel size in the longitudinal direction was substantially larger and set to this exact value because with a shift of one pixel per longitudinal (the direction with pixel size 0.1432 mm) step this results in the simulated wire being placed at an angle of 3 degrees. Please note that the simulations were noiseless, since noise does not influence the resolution loss.

The resolution loss in longitudinal direction (across slices) has also been validated by imaging this $50 \text{ }\mu\text{m}$ diameter of the tungsten wire both digitally and physically. The wire was placed such that the angle with the slices was 8 degrees. The slice sensitivity profile (SSP) was determined in the same way as the MTF. The simulated voxel size of the phantom was $0.005 \text{ mm} \times 0.035 \text{ mm} \times 0.005 \text{ mm}$. The voxel size in the lateral direction was substantially larger and set to this exact value because with a shift of one pixel per lateral (the direction with pixel size 0.035 mm) step this results in the simulated wire being placed at an angle of 8 degrees. Please note that the simulations were noiseless, since noise does not influence the resolution loss.

The $50 \text{ }\mu\text{m}$ diameter of the tungsten wire is relatively small compared to the detector pixel size, even when subsampled. To overcome this problem the detector subsampling M was set to 24, just for the raytracing. After the raytracing, the detector was rebinned to its original subsampling of $M=4$.

For validating the noise magnitude and frequency content, a water phantom with a radius of 320 mm was again both digitally simulated and physically measured, and the results were compared. The noise magnitude and frequency content were validated at two different tube current levels (140 mA and 400 mA), and three different tube voltage levels (100 kV , 120 kV , and 135 kV). The simulated voxel size of the phantom was $1.0 \text{ mm} \times 0.25 \text{ mm} \times 0.25 \text{ mm}$. The voxel size in the longitudinal direction was substantially larger since the phantom is constant in this direction. A volume of interest (VOI) of $64 \times 64 \times 64$ voxels was placed in the center of the water phantom images. The standard deviation of this VOI was used to validate the magnitude of the noise. To validate the noise frequency content, a 100 mm FOV was reconstructed in the center and at the periphery, approximately 120 mm from the center, of the water phantom. This smaller FOV was reconstructed, to have a smaller pixel size, making it possible to validate higher frequencies. The 2D NPS and 2D unstructured NPS of both these FOVs, were calculated in 256×256 -pixel ROIs

from across 80 slices, by determining the square of the 2D Fourier transform. In the case of the unstructured NPS the average of the 80 slices was subtracted before calculating the Fourier transform. Both the 2D NPS and 2D unstructured NPS were normalized to have an area of one, obtaining the normalized NPS (nNPS). Both the 2D nNPS and 2D unstructured nNPS were calculated to show that the simulation does not introduce any structured noise. A comparison of the nNPS at these two positions was performed to validate the changes in the noise characteristics throughout the imaging field. These 2D nNPSs at the center were also radially averaged, and again normalized to have an area of one, to obtain a 1D nNPS. To validate the frequency content across slices, the 1D nNPS was calculated across 280 slices for all pixels in a 128×128 ROI at the center, and the results were averaged.

Scatter was not included in our simulator, since the system performs scatter correction during the reconstruction process, and therefore, the benefit of adding simulated scatter would be minimal. To validate the performance of the scatter correction, the 320 mm water phantom (also used for the validation of the noise) was imaged with the standard volume scan collimation of 160 mm (equal to all measurements in this work) and with a 20 mm collimation, which is assumed to have a negligible amount of scatter. The line profile of the reconstructed water phantom images was compared for both collimations. Line profiles were obtained from these images by averaging 60 individual line profiles across 38 slices both horizontal and vertical directions for both water phantom scans and the corresponding simulation.

Simulation simplifications

The three steps of the simulator incorporating subsampling, namely, the number of angular projection subsamples K , the number of detector subsamples $M \times M$ and the number of focal spot subsamples $L \times L$, were simplified to reduce the time and computer power necessary. The angular subsamples K were set to 1, 2, and 3. The detector subsamples $M \times M$ were set to 1×1 , 2×2 , 3×3 , and 4×4 . The number of focal spot subsamples $L \times L$ was set to 1×1 , 2×2 , and 3×3 . Please note that while one of these three was reduced the other two were kept at their original value. Previous CT simulators [43,74,75] also used or optimized their subsampling, however with this analysis the impact of each individual simplification is shown.

To validate the impact of these simplifications on the realism of the simulation, the MTFs of the images resulting from the digitally simulated simplified sinograms were determined and compared to the MTFs of the physically measured sinograms. In both cases these MTFs were again determined from a tungsten wire at 7 cm, 14 cm, and 21 cm from the isocenter.

In addition, the possibility of compensating for simplifying the focal spot as being a point source by using the system MTF, i.e., the MTF measured with the edge located at the CoR, instead of the detector MTF, was also tested. These MTFs are shown in Appendix B. Finally, the possibility of simplifying the incorporation of the rotational blurring, due to the angular motion of the source and detector, was also investigated by averaging each angular projection with the subsequent one, instead of performing the angular projection subsampling.

Results

In Figure 4, one of the measurements of the MECT phantom used for the CT number validation is shown. Table 1 shows the measured and simulated CT numbers for the different materials in the MECT phantom, with the numbers corresponding to the regions in Figure 4. Table 1 also shows the maximum (bold and underlined), mean absolute, and mean error of the CT numbers in HU. As can be seen from the values in Table 1, the simulations result in a small negative bias in the CT numbers for all tube voltage levels.

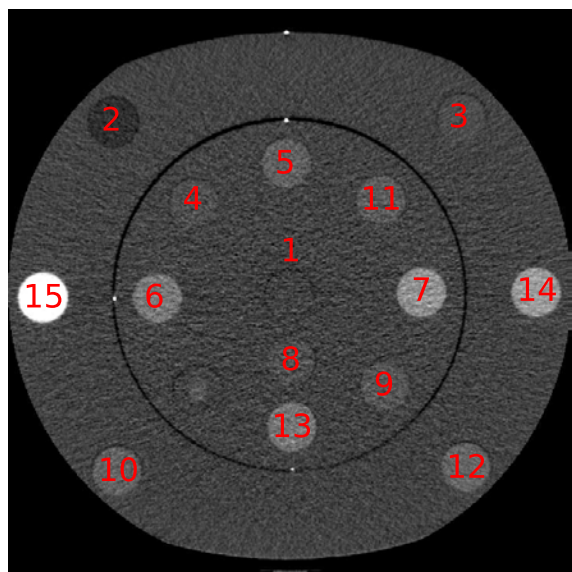


Figure 4: MECT phantom, measured with 135 kV, used for CT number validation, with a window level (WL) of 300 HU and a window width (WW) of 1000 HU.

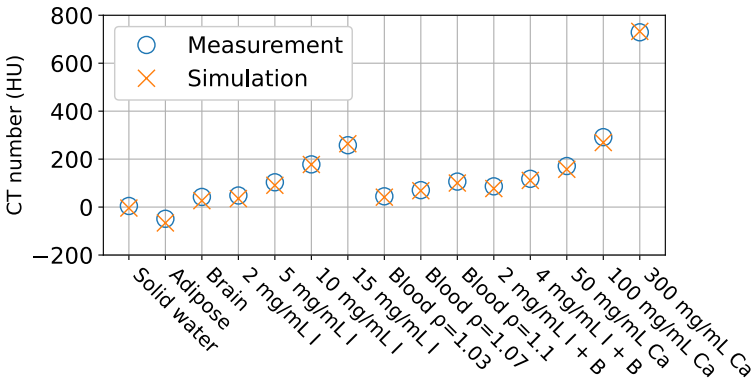


Figure 5: CT numbers of the real and simulated MECT phantom images at 135kV. Note that the B stands for blood $\rho=1.03\text{ g/cm}^3$.

Table 1: CT numbers corresponding to the regions indicated in the MECT phantom of Figure 4. The material with the maximum error in CT number is bold and underlined for each tube voltage. Note that the B in material 11 and 12 stands for blood with $\rho=1.03\text{ g/cm}^3$. The error is defined as simulation – measurement.

Material	100 kV		120 kV		135 kV	
	Measurement [HU]	Simulation [HU]	Measurement [HU]	Simulation [HU]	Measurement [HU]	Simulation [HU]
(1) Solid water	8.1	-0.6	3.9	-5.3	4.6	-4.0
(2) Adipose	-61.4	-74.7	-52.5	-69.7	-48.8	-66.2
(3) Brain	43.4	28.0	42.7	26.6	42.3	26.7
(4) 2 mg/mL I	63.6	54.1	51.6	38.6	48.3	35.4
(5) 5 mg/mL I	143.9	132.3	114.6	103.7	103.4	90.8
(6) 10 mg/mL I	267.4	263.5	206.2	203.8	178.0	177.5
(7) 15 mg/mL I	384.1	385.4	296.5	303.3	258.0	264.1
(8) Blood $\rho=1.03$	54.8	41.8	45.5	38.8	44.8	41.3
(9) Blood $\rho=1.07$	78.1	65.6	70.9	63.7	70.6	68.2
(10) Blood $\rho=1.1$	106.0	97.4	104.7	97.3	106.3	100.3
(11) 2 mg/mL I + B	107.6	95.6	92.4	82.1	86.5	77.3
(12) 4 mg/mL I + B	149.6	140.2	127.0	119.6	117.8	111.2
(13) 50 mg/mL Ca	189.8	169.9	176.3	162.4	171.2	157.7
(14) 100 mg/mL Ca	<u>335.0</u>	<u>313.7</u>	<u>302.4</u>	<u>281.7</u>	<u>291.6</u>	<u>269.1</u>
(15) 300 mg/mL Ca	882.9	886.3	776.3	779.6	728.9	732.8
Mean Absolute Error [HU]	10.9		10.2		9.4	
Mean Error [HU]	-10.3		-8.8		-8.1	

Figure 5 shows the measured and simulated CT numbers of the 135 kV case. The remaining errors in the CT numbers do not seem to have a correlation, indicating the simulation does not introduce any non-linear offset to the CT numbers. Note that, as explained before, the linear calibration is determined using a different phantom and is applied to all subsequent images.

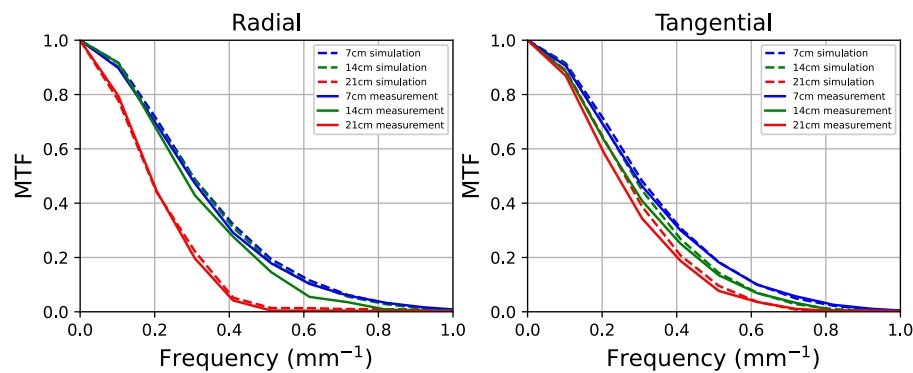


Figure 6: Modulation transfer function of measured and simulated wires in radial (left) and tangential (right) direction for the large focal spot.

The MTFs in the radial and tangential directions can be seen in Figure 6. It can be observed that the latter starts dropping when moving out of the CoR in both the measured and simulated cases. This is due to the rotation of the system introducing more blur further away from the CoR. The frequencies at 10% MTF and their relative error for the various FoV positions, directions, and focal spot sizes are listed in Table 2. The maximum error is 11.1% and the mean absolute error is 3.4%, showing that the spatial resolution characteristics in the simulated CT images are close to those of the clinical system.

The SSP of both the large and small focal spot can be seen in Figure 7. It can be observed that the resolution loss is higher for the large focal spot, as expected. Table 3 shows the frequencies at 10% SSP for measurement and simulation and their relative error for both focal spots. The mean absolute error is 3.3%, showing that the spatial resolution across slices in the simulated CT images is close to those of the clinical system.

Table 2: Frequency at 10% MTF of the measured and simulated MTF, and the relative error of these frequencies in the different directions, positions, and focal spot sizes evaluated. Relative error = (simulation – measurement)/measurement * 100%.

Focal spot size	Direction	Position [cm]	Measured frequency [mm ⁻¹]	Simulated frequency [mm ⁻¹]	Relative error [%]
Large	Tangential	7	0.62	0.62	0.3
		14	0.57	0.57	1.2
		21	0.49	0.51	3.6
	Radial	7	0.62	0.65	3.8
		14	0.56	0.63	11.1
		21	0.37	0.38	2.6
Small	Tangential	7	0.72	0.70	-3.4
		14	0.64	0.63	-1.5
		21	0.55	0.55	0.8
	Radial	7	0.77	0.74	-3.3
		14	0.70	0.73	3.6
		21	0.41	0.39	-5.7

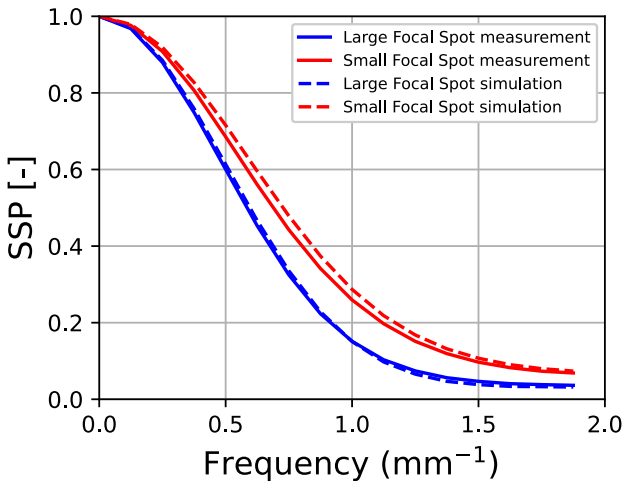


Figure 7: Slice sensitivity profile of measured and simulated wires for both focal spots present in the system.

Table 3: Frequency at 10% SSP of the measured and simulated SSP, and the relative error of these frequencies for both focal spot sizes evaluated. Relative error = (simulation – measurement)/measurement * 100%.

Focal spot size	Measured frequency [mm ⁻¹]	Simulated frequency [mm ⁻¹]	Relative error [%]
Large	1.14	1.12	-1.7
Small	1.48	1.55	4.8

Figure 8 shows images of the water phantom used to validate the noise characteristics of the simulated images. The red square indicates the ROI used for the validation of the noise magnitude. Table 4 shows the results of the noise magnitude in terms of standard deviation. The maximum and mean absolute errors were 8.5% and 4.9%.

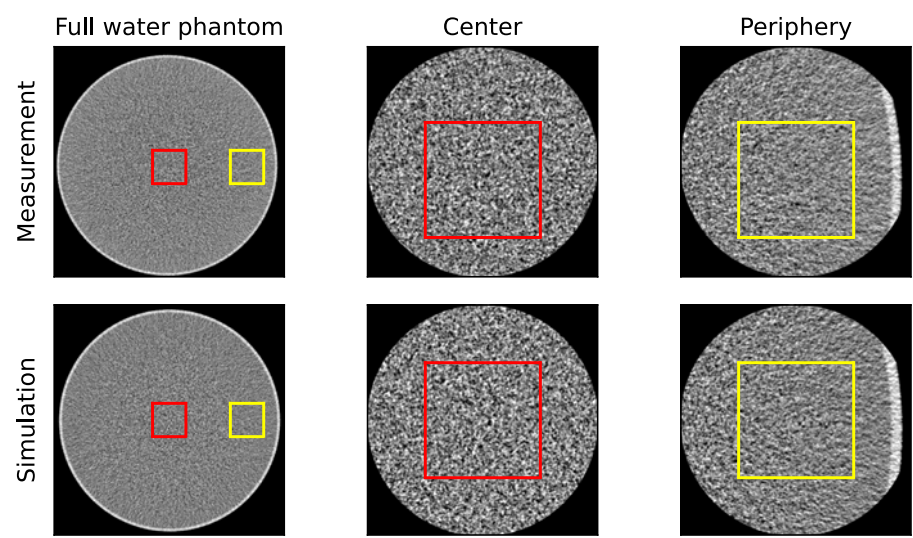


Figure 8: Water phantom used for nNPS validation (140 mA, 135 kV) with a WL of 0 HU and a WW of 400 HU. The squares indicate the ROIs used to determine the nNPSs.

The ROIs used to validate the frequency content within a slice are also indicated in Figure 8, by the red and yellow squares. Please note that the ROIs of the measurements have a slight offset in vertical direction, this offset compensates for misalignment between the measured and simulated water phantom, to ensure that the same location of the phantom was analyzed. The 2D nNPS and 2D unstructured nNPS of the noise at the center (red square in Figure 8) within a slice of the measured and simulated phantom are shown in Figure 9 and Figure 11,

respectively. It can be seen that the 2D nNPS and 2D unstructured nNPS are isotropic in the center for both measurement and simulation. The difference images only show a small overestimation at lower frequencies (white dominant ring) and underestimation at slightly higher frequencies (black dominant ring). The 2D nNPS and 2D unstructured nNPS of the noise at the periphery (yellow square in Figure 8) of the measured and simulated phantom are shown in Figure 10 and Figure 12, respectively. These results show that the nNPS is anisotropic at the periphery for both measurement and simulation, and that the degree of anisotropy in the simulated image is similar to that in the real one. The difference images only show a small overestimation in vertical direction (white dominant regions above and below the center) and underestimation in horizontal direction (black dominant regions left and right from the center). The only difference between the 2D normalized nNPSs and 2D unstructured normalized nNPSs is a small low frequency peak in the horizontal direction of the 2D nNPS of the periphery, which is not present in the 2D unstructured nNPS of the periphery. This minor peak is introduced by a small cupping artifact in the periphery of the measurement, see Figure 15.

Table 4: Noise magnitude of measured and simulated water phantom of Figure 8. Relative error = (simulation – measurement)/measurement * 100%.

Tube voltage [kV]	Tube current [mA]	σ measurement [HU]	σ simulation [HU]	Relative error [%]
100	140	128.2	140.2	8.5
	400	66.8	68.8	2.9
120	140	83.8	89.0	5.8
	400	47.0	48.1	2.2
135	140	68.4	72.8	6.1
	400	39.0	40.4	3.6

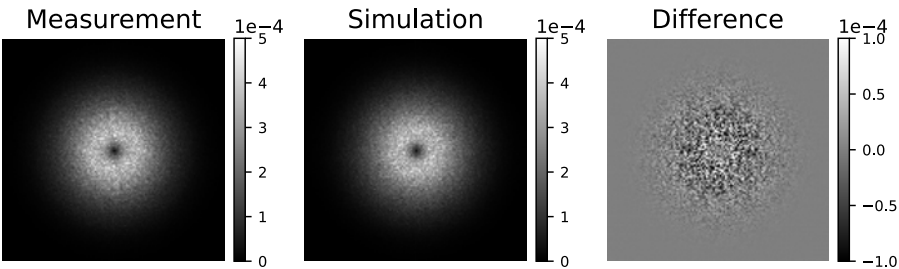


Figure 9: 2D nNPS in the center region of measured (left) and simulated (middle) water phantom, and the difference between both nNPSs (right). Difference = simulation – measurement.

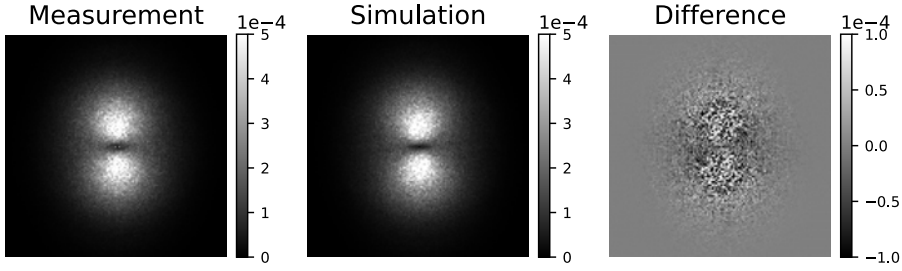


Figure 10: 2D nNPS in the periphery region of measured (left) and simulated (middle) water phantom, and the difference between both nNPSs (right). Difference = simulation – measurement.

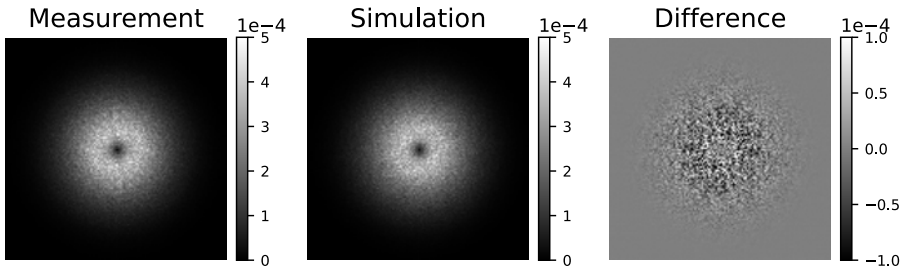


Figure 11: 2D unstructured nNPS in the center region of measured (left) and simulated (middle) water phantom, and the difference between both nNPSs (right). Difference = simulation – measurement.

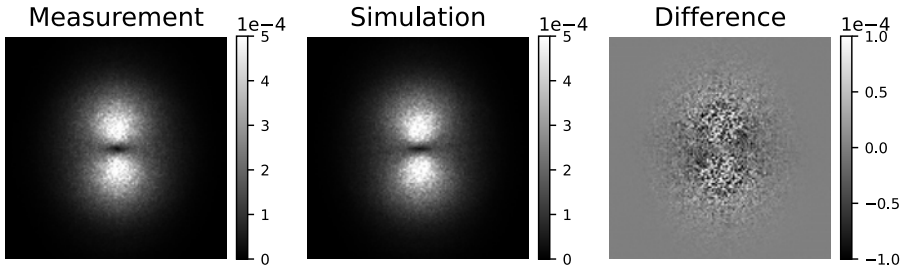


Figure 12: 2D unstructured nNPS in the periphery region of measured (left) and simulated (middle) water phantom, and the difference between both nNPSs (right). Difference = simulation – measurement.

Figure 14 shows the radially averaged nNPS and nNPS across slices in the center of both the measurement and simulation for the 135 kV and 140 mA case. The mean absolute errors of radially averaged nNPSs are summarized in Table 5, which shows that the maximum mean absolute error is 8.4% and the average mean absolute error is 3.9%. The mean absolute errors of the nNPSs in the slice direction are summarized in Table 5, and is on average 6.2% and maximum 8.8%. The plots of the other radially averaged nNPSs and nNPSs across slices are depicted in Appendix D.

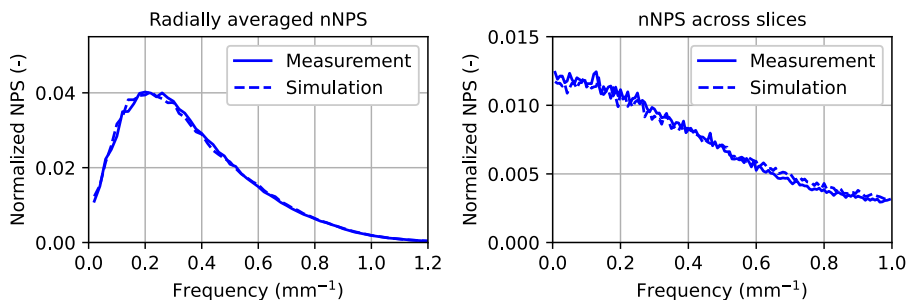


Figure 13: Radially averaged center nNPS (left) and nNPS across slices (right), for 135 kV and 140 mA.

Table 5: Difference in measured and simulated nNPS with a slice and across slices

Tube voltage	100 kV		120 kV		135 kV	
Tube current	140 mA	400 mA	140 mA	400 mA	140 mA	400 mA
% Mean absolute difference within slice	2.2	2.2	2.4	4.9	3.1	8.4
% Mean absolute difference across slices	5.6	5.7	5.2	7.3	4.8	8.8

Figure 14 indicates the region from where the line profiles of the water phantom for the scatter correction validation are obtained, with the resulting average line profiles shown in Fig 15. The line profiles show the effect of the scatter correction, resulting in a good match in HU values at the center of the phantom, but with a larger remaining error at the sides. As a result, it can be seen that the line profile of our simulation is also approximately flat, as expected, and has a maximum difference of 15 HU at the edge of the water phantom.

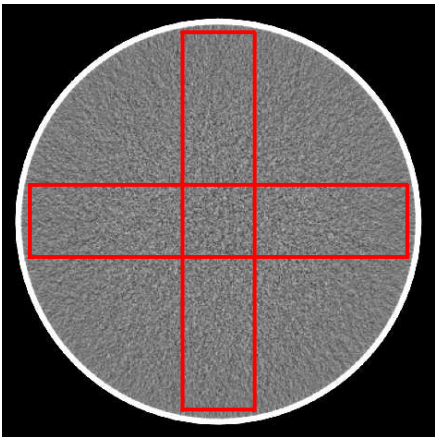


Figure 14: Averaged slices of simulated water phantom with indicated ROIs used for obtaining the line profiles.

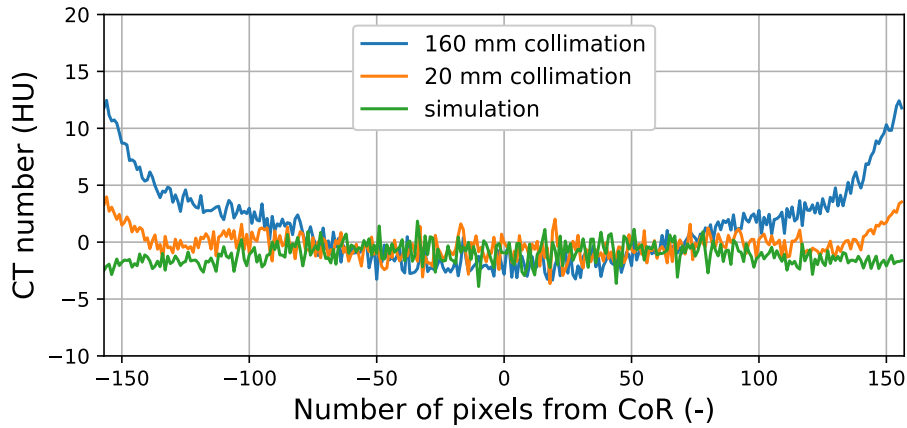


Figure 15: Line profiles of the measurement with 160 mm collimation, the measurement with 20 mm collimation, and the simulation.

Table 6: Frequency at 10% MTF for all simulation simplifications and absolute relative error of this frequency compared to the frequency of the full simulation at 10% MTF, for tangential direction.

Focal spot size	Subsampling simplification	Subsampling factor	Absolute relative error at 10% MTF			Time / memory reduction factor
			7 cm	14 cm	21 cm	
Large	focal spot	1	20.9	18.0	12.1	9
		2	7.7	8.5	4.1	2.25
	system MTF	1	12.5	11.9	7.6	9
	angular	1	2.5	6.0	16.5	3
		2	0.4	0.7	1.8	1.5
	projection averaging	1	15.2	14.5	15.3	3
	detector	1	0.4	0.0	0.2	16
		2	0.1	0.0	0.0	4
		3	0.1	0.1	0.1	1.78
	Small	focal spot	1	6.9	6.6	3.8
2			3.0	3.4	2.2	2.25
system MTF		1	0.6	1.1	0.4	9
angular		1	2.5	10.3	20.1	3
		2	0.4	1.5	2.3	1.5
projection averaging		1	16.9	12.6	9.7	3
detector		1	0.2	0.0	0.4	16
		2	0.0	0.0	0.1	4
		3	0.0	0.1	0.1	1.78

Table 6 and Table 7 show the absolute relative errors between the simplified and the fully subsampled simulation of the frequency at 10% MTF for the simulated tungsten wire at 7 cm, 14 cm, and 21 cm from the CoR. The absolute relative errors for all simplifications are shown for the tangential (Table 6) and radial (Table 7) direction. In the last column of both Table 6 and Table 7, the time and memory reduction factor are listed (the reduction factor is the same for time and memory, because the number of calculations that need to be performed scales linearly with the amount of memory in our application).

Table 7: Frequency at 10% MTF for all simulation simplifications and absolute relative error of this frequency compared to the frequency of the full simulation at 10% MTF, for radial direction.

Focal spot size	Subsampling simplification	Subsampling factor	Absolute relative error at 10% MTF			Time / memory reduction factor
			7 cm	14 cm	21 cm	
Large	focal spot	1	21.1	25.2	2.8	9
		2	11.2	10.4	1.7	2.25
	system MTF	1	12.7	16.3	3.4	9
	angular	1	0.1	0.1	0.1	3
		2	0.1	0.0	0.0	1.5
	projection averaging	1	0.6	1.0	0.4	3
	detector	1	-0.6	2.1	0.1	16
		2	0.2	0.3	0.1	4
		3	0.3	0.1	0.1	1.78
Small	focal spot	1	5.5	7.7	0.8	9
		2	3.7	3.8	0.5	2.25
	system MTF	1	1.8	0.0	1.3	9
	angular	1	0.0	0.1	0.0	3
		2	0.1	0.0	0.0	1.5
	projection averaging	1	1.1	1.3	0.1	3
	detector	1	1.2	1.3	0.2	16
		2	0.1	0.2	0.0	4
		3	0.0	0.1	0.0	1.78

The differences in the resulting MTFs are, for most cases, subtle, except for the case of not subsampling the focal spot or angular projections. The detector subsampling seems to have very little effect in both directions.

Table 6 and Table 7 also show the absolute relative errors in the frequency at 10% MTF between using the system MTF (as described in Section Simulation simplifications) and the fully subsampled focal spot simulation for the simulated tungsten wire at 7 cm, 14 cm, and 21 cm from the CoR. The results clearly show that using the system MTF improves the results when assuming the focal spot is a point source, since it reduces the error compared to the full simulation by approximately 30-40%. The results also show that averaging each angular projection with the subsequent angular projection gives worse results than when simplifying to only using a single angular projection. Therefore, to have an error of 17% or less in the frequency at 10% MTF compared to the full simulation, all subsamples can be reduced to 1, except for the angular projections, which should still be 2. In this case the maximum error is 16.3% and the time and memory consumption could be reduced by a factor of 216.

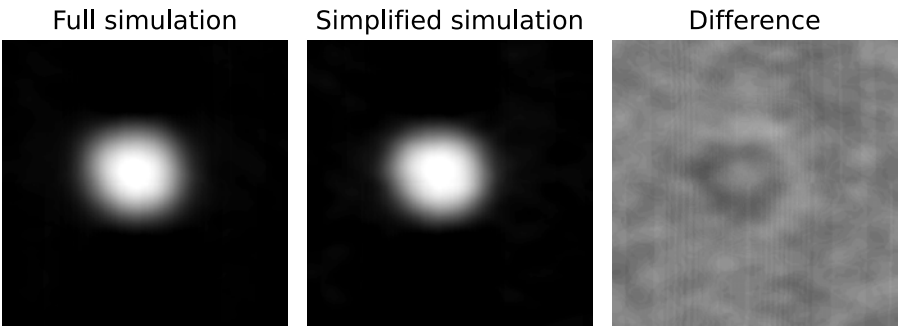


Figure 16: Noiseless simulations of a lesion at 14 cm from the CoR. (Left) Full simulation (WW: 100, WL: 85). (Middle) Simplified simulation, 1 source sample using the MTF measured in the CoR, 2 angular subsamples and 1 detector sample, time and memory consumption potentially reduced by a factor 216 (WW: 100, WL: 85). (Right) Difference between the two simulations (WW: 40, WL: 0).

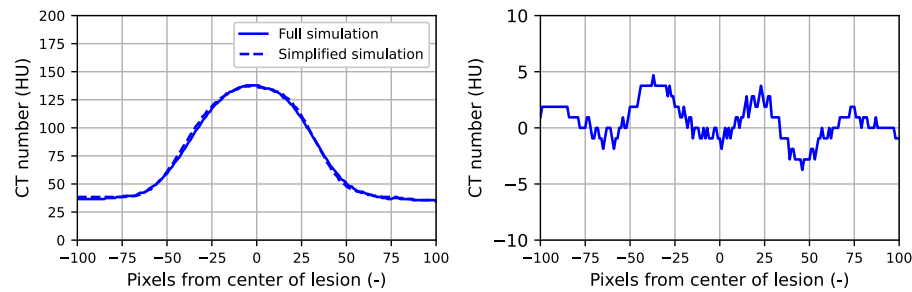


Figure 17: (Left) Line profile of the full and simplified simulation of the lesion in Figure 16. (Right) Difference between simplified simulation and full simulation (simplified simulation – full simulation).

The influence of these simplifications (resulting in 16.3% error) can be observed in Figure 16. Here a small lesion is imaged with full subsampling and simplified subsampling (1 source sample using the system MTF, 2 angular subsamples, and 1 detector sample) at 14 cm from the CoR, to show its influence. The difference image shows a minor ring but no other structural differences, indicating minimal difference between the two simulations after reconstruction. The same can be observed from the line profiles of these two reconstructed simulations plotted in Figure 17. However, for each application, the desired accuracy could differ, so the user should decide what is an acceptable error margin for their application.

Computation time

All simulations were performed on a Linux system with 128 GB RAM, AMD Ryzen Threadripper 1950X 16-core CPU, and a 48 GB Nvidia RX A600 GPU. The ray tracing and sinogram calculations were performed on the GPU. All image corruptions were performed on the CPU (due to the sinogram size). Generating all projections of 896×320 pixels for the MECT phantom, with a voxel array size of 50×1600×1600 voxels of the same voxel size as those used for validation and consisting of 15 different materials (used for CT number validation), took ~40 hours on this workstation. Please note that the time reduction factor reported with the simplification results is theoretical as some minor operations are independent of sinogram size and the calculations of an entire set of projections are done in batches due to memory constraints. As an example, by reducing the sub-sampling of the focal spot to 1, angular projections to 2, and detector pixels to 1, the simulation time for this same simulation results in ~35 minutes.

Discussion

In this work a scanner-specific CT simulation was developed and validated, combining physics-based simulations for the position-dependent effects and the shift-invariant image corruption methods described by Saunders et al. [58] for the detector effects. Therefore, minimizing the system information needed from the vendor and making it possible to perform scanner-specific CT simulations with only system-specific geometry information. In addition, the impact of simplifying the simulation process, both in terms of resulting realism and computer power requirements, was evaluated. To include all effects of the real CT system resulting in spatial resolution loss, the shift-variant impact of the finite focal spot size and of gantry rotation are modelled. The simulator is shown to generate images that match the characteristics of the real images to within an average of 3.4% and 3.3%

in terms of spatial resolution within and across image slices, respectively. In terms of noise characteristics, these matched in terms of noise magnitude (standard deviation) and noise frequency content (nNPS) within and across slices to within 4.9%, 3.9%, and 6.2%, respectively. The effect of simulation simplification was assessed, and the results showed that for general applications most simplifications, except for the angular simplification, do not cause a major decrease in realism of the simulated image (maximum error in frequency at 10% MTF of 16.3%). Therefore, the time and computer power necessary could be reduced for many applications in which this level of realism is sufficient. This would aid studies that aim to use large virtual clinical trials, since it will become feasible to generate very extensive datasets within a reasonable time frame. Also, studies about processing or reconstruction algorithms could benefit as it becomes more feasible to cover large multi-dimensional parameter spaces for (first stage) testing, after which one could choose for more realistic simulations for refinement of solutions, if needed. However, the purpose of the study must be considered to make an informed decision on the level of realism that is desired.

During the development some assumptions had to be made. One of them being the assumption of stationary behavior of the detector MTF and NPS across the entire detector. The NPSs were only measured at the detector center, since the curvature of the detector is such that the normal direction of each element is pointing to the source. Only in the direction of the rows, where there is no curvature, the normal is not pointing directly to the source, however the maximum angle is 10° and therefore assumed to have a negligible effect. Also, the MTF was only measured in the detector row direction (since in the other direction the MTF is affected by the gantry rotation), then assuming rotational symmetry between all directions. In addition, the resolution characteristics are dominated by the focal spot size and finite exposure time during the angular motion of the CT gantry, justifying the assumption of stationary and symmetric MTF across the detector. Also, the residual scatter after correction was neglected. As shown, this simplification results in a CT number difference of approximately 15 HU at the edge of a 320 mm water phantom with the widest x-ray beam collimations. The QE of the detector is based on the theoretical energy absorption of the nominal detector active layer thickness. However, even with these assumptions, the validation results point to the appropriateness of the simulations in terms of spatial resolution, noise magnitude and frequency content, and, especially, their shift variance compared to a real clinical CT system.

The developed scanner-specific CT simulation may function as a tool to facilitate virtual clinical trials to test new reconstruction algorithms and post-processing algorithms. The simulator could additionally be used to also test new components like bowtie filters or flat filters, different detectors, or a different focal spot size, as well as new acquisition protocols.

Conclusion

A scanner-specific CT simulation was developed, implemented, and validated. The validation of the performed simulations showed that it can generate images comparable to those obtained using a real clinical CT system. An analysis on simulation simplification also showed that for general applications, time and computer power can be spared without substantial loss of realism. The simulator can generate realistic scanner-specific CT images, which will aid the development of new reconstruction and post-processing algorithms by opening the possibility for virtual clinical trials.

Supplementary Material

Appendix A – General overview pipeline

The general simulation pipeline in Figure A1 gives an overview of all information or measurements required per step, so the same method can be applied for simulating other CT systems.

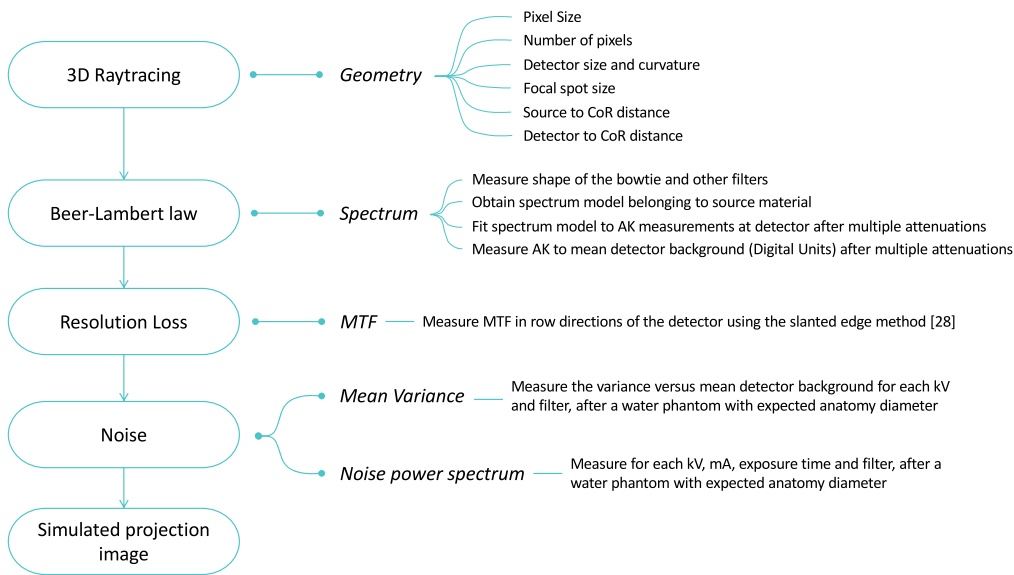


Figure A1: General pipeline presenting all required information or measurements per simulation step.

Appendix B - Modulation Transfer Functions

A tungsten edge (TX5, IBA Dosimetry, Schwarzenbruck, Germany) was used in combination with the slanted edge method [66] to measure the edge spread function in the row direction of the detector. Since this direction does not experience resolution loss due to rotation and obtaining linear domain data with the gantry remaining static was not possible. To determine the detector MTF, the tungsten edge was placed as close as possible to the detector entrance surface (8 cm away), to minimize the focal spot size effect.

The function used to fit the MTF is based on the Lorentzian fit of Siewerdsen et al. [67], defined as follows:

$$MTF(f) = \frac{1}{1 + a * f + b * f^2 + c * f^3 + d * f^4 + e * f^5} \quad (B.1)$$

where f is the spatial frequency and a, b, c, d and e are the optimizable parameters of the function.

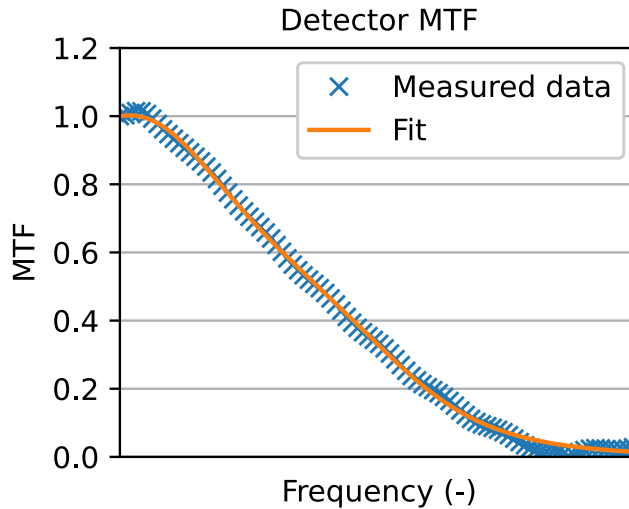


Figure B1: Measured data and fit of the detector MTF.

The data points and fitted MTF are shown in Figure B1. The R^2 of the fit is listed in Table B1. Please note that the values of the x-axis are not displayed, and the fitted parameters are not reported, since these are confidential.

The same measurements have been performed at the center of rotation (CoR) to determine the system MTF for both focal spots present in the system. The data points and fitted MTF curves are shown in Figure B2. The resulting R^2 of the fits are listed in Table B1 for both focal spot sizes. Please note that the values of the x-axes are not displayed, and the fitted parameters are not reported, since these are confidential.

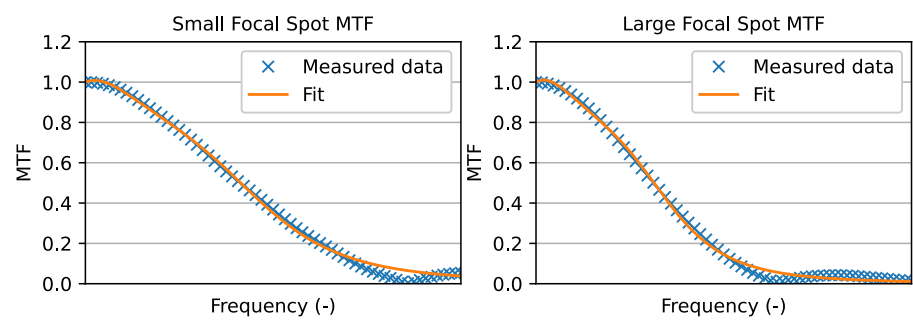


Figure B2: Measured data and fit of the system MTF of both focal spots present in the system.

Table B1: R^2 values for fits.

	R^2
Detector MTF	0.998
System MTF (small focal spot)	0.996
System MTF (large focal spot)	0.998

Appendix C – Subsampling Experiments

The number of subsamples required for the focal spot ($L \times L$), detector elements ($M \times M$), and angular projections (K) were determined experimentally. Please note that the values of the x-axes are not displayed, since these are confidential.

To determine the number of focal spot subsamples, a tungsten edge (TX5, IBA Dosimetry, Schwarzenbruck, Germany) was placed at the Center of Rotation (CoR) and imaged using the large focal spot of the system. This tungsten edge was also simulated at the CoR. The MTF in the sinogram was determined for both directions and for both measurement and simulation. Experimental validation showed that the optimal number of focal spot subsamples per direction is $L=3$, see Figure C1 and Figure C2.

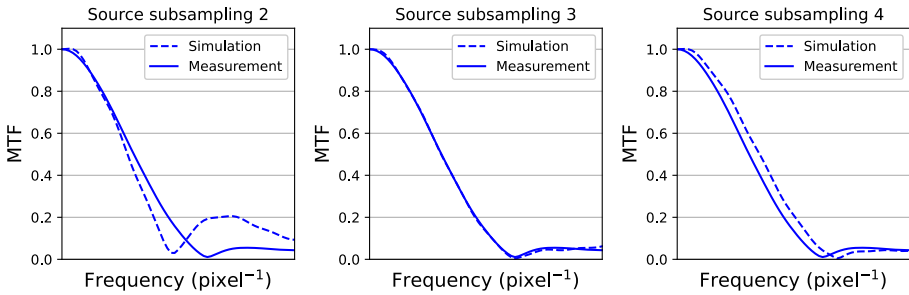


Figure C1: MTF in the row direction of the detector with, from left to right 2, 3, and 4 subsamples in each direction of the focal spot.

The number of subsamples for the detector elements was determined by placing the tungsten edge as close as possible to the detector entrance (8 cm away), to minimize the focal spot effect, and imaging it with the small focal spot present in the system. The MTF in the sinogram was determined only for the direction across detector rows since the other direction has the angular rotation effect. Experiments showed that the optimal number of detector subsamples per direction is $M=4$, see Figure C3, since lower subsamples result in the minimum of the MTF being at a too low frequency.

To determine the number of angular projections subsamples the tungsten edge was again placed as close as possible to the detector entrance (8 cm away) and imaged with the small focal spot. The MTF in the sinogram was determined only for the direction across detector columns, as this is the only direction in which the resolution loss is influenced by the angular rotation. Experimental validation showed

that the optimal number of angular projection subsamples is $K=3$, see Figure C4, since higher subsampling does not show significant improvement.

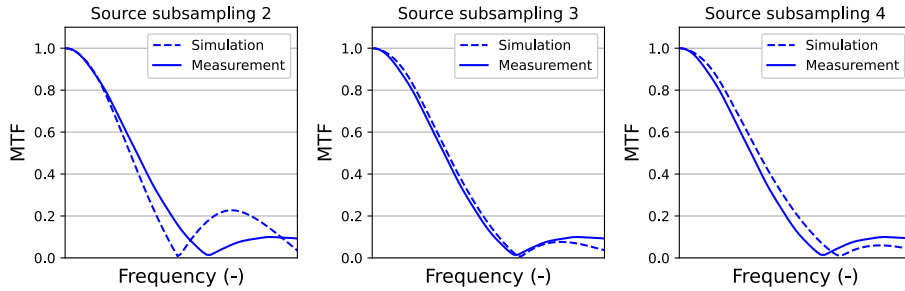


Figure C2: MTF in the column direction of the detector with, from left to right 2, 3, and 4 subsamples in each direction of the focal spot.

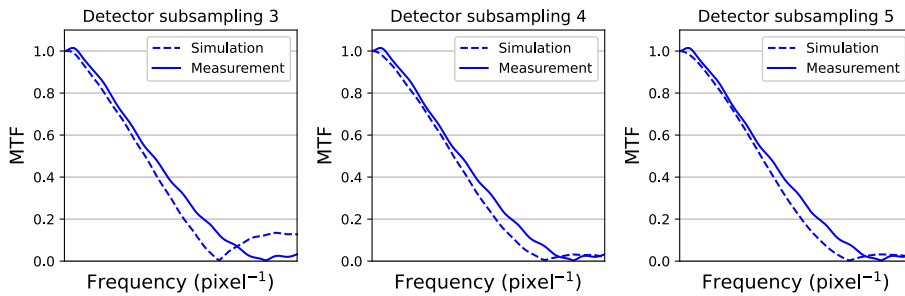


Figure C3: MTF in the row direction of the detector with, from left to right 3, 4, and 5 subsamples in each direction of the detector pixels.

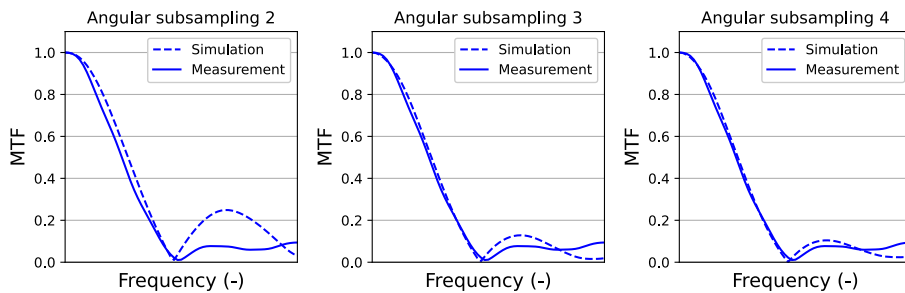


Figure C4: MTF in the column direction of the detector with, from left to right 2, 3, and 4 subsamples for each angular projection.

Appendix D – Noise Power Spectra

The radially averaged nNPS and nNPS across slices in the center of both the measured and simulated water phantom of Figure 8 of the manuscript, at tube voltage levels of 100 kV, 120 kV, and 135 kV and tube current levels of 140 mA and 400 mA are shown in Figure D1-D5. Please note that the 135 kV and 140 mA case is not shown, since it is already shown in Figure 13 of the manuscript.

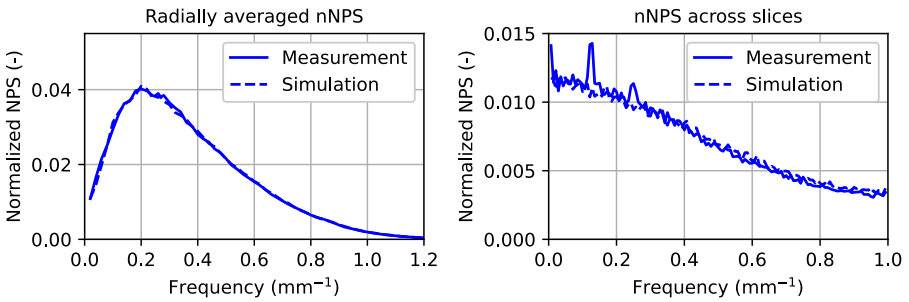


Figure D1: Radially averaged center nNPS (left) and nNPS across slices (right), for 100 kV and 140 mA.

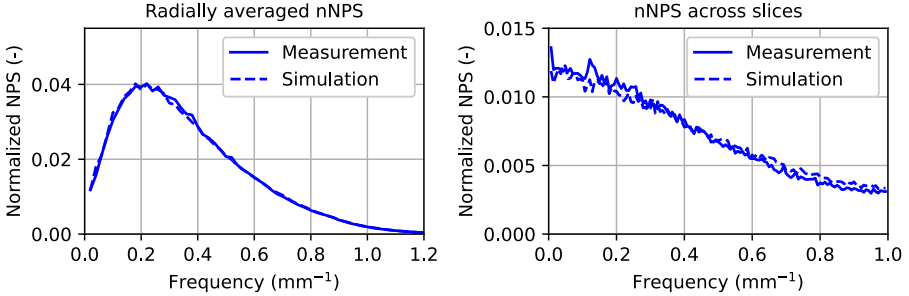


Figure D2: Radially averaged center nNPS (left) and nNPS across slices (right), for 100 kV and 400 mA.

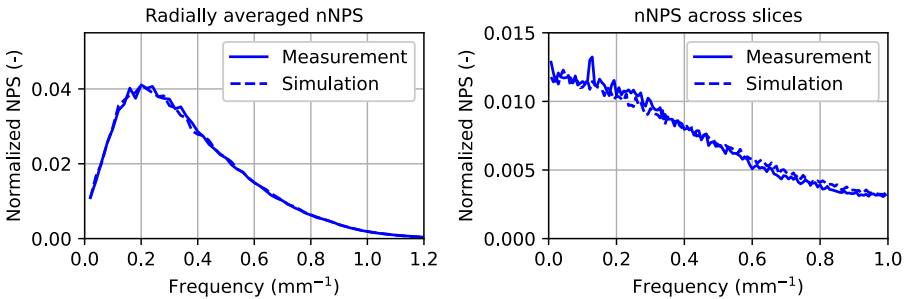


Figure D3: Radially averaged center nNPS (left) and nNPS across slices (right), for 120 kV and 140 mA.

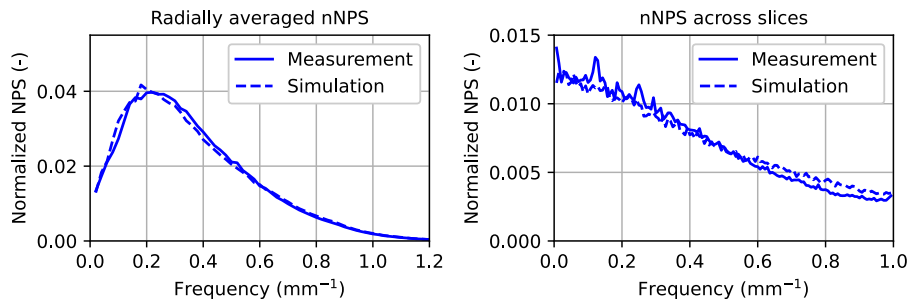


Figure D4: Radially averaged center nNPS (left) and nNPS across slices (right), for 120 kV and 400 mA.

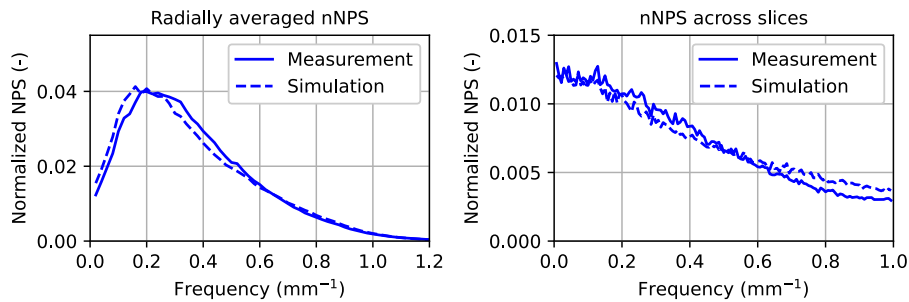


Figure D5: Radially averaged center nNPS (left) and nNPS across slices (right), for 135 kV and 400 mA.



Chapter 3

Image domain lower dose simulations

Original title: Deep learning-based low-dose CT simulator for non-linear reconstruction methods

Sjoerd A. M. Tunissen, Nikita Moriakov, Mikhail Mikerov, Ewoud J. Smit, Ioannis Sechopoulos, and Jonas Teuwen

Published in: Medical Physics

ABSTRACT

Background

Computer algorithms that simulate lower-dose Computed Tomography (CT) images from clinical-dose images are widely available. However, most operate in the projection domain and assume access to the reconstruction method. Access to commercial reconstruction methods may often not be available in medical research, making image-domain noise simulation methods useful. However, the introduction of non-linear reconstruction methods, such as iterative and deep learning-based reconstruction, makes noise insertion in the image domain intractable, as it is not possible to determine the noise textures analytically.

Purpose

To develop a deep learning-based image-domain method to generate low-dose CT images from clinical-dose CT images for non-linear reconstruction methods.

Methods

We propose a fully image domain-based method, utilizing a series of three convolutional neural networks (CNNs), which, respectively, denoise clinical-dose CT images, predict the standard deviation map of the low-dose image, and generate the noise power spectra of local patches throughout the low-dose image. All three models have U-net-based architectures and are partly or fully 3-dimensional. As a use case for this study and with no loss of generality, we use paired low-dose and clinical-dose brain CT scans. A dataset of 326 paired scans was retrospectively obtained. All images were acquired with a wide-area detector clinical system and reconstructed using its standard clinical iterative algorithm. Each pair was registered using rigid registration to correct for motion between acquisitions. The data was randomly partitioned into training (251 samples), validation (25 samples), and test (50 samples) sets. The performance of each of these three CNNs was validated separately. For the denoising CNN, the local standard deviation decrease and bias were determined. For the standard deviation map CNN, the real and estimated standard deviations were compared locally. Finally, for the noise power spectra CNN, the noise power spectra of the synthetic and real low-dose noise were compared inside and outside the skull. Two proof-of-concept denoising studies were performed to determine if the performance of a CNN- or a gradient-based denoising filter on the synthetic low-dose data vs. real data differed.

Results

The denoising network had a median decrease in noise in the cerebrospinal fluid by a factor of 1.71 and introduced a median bias of +0.7 HU. The network for standard deviation map estimation had a median error of +0.1 HU. The noise power spectrum estimation network was able to capture the anisotropic and shift-variant nature of the noise structure by showing good agreement between the synthetic and real low-dose noise and their corresponding power spectra. The two proof of concept denoising studies showed only minimal difference in standard deviation improvement ratio between the synthetic and real low-dose CT images with the median difference between the two being 0.0 and +0.05 for the CNN- and gradient-based filter, respectively.

Conclusion

The proposed method demonstrated good performance in generating synthetic low-dose brain CT scans without access to the projection data or to the reconstruction method. This method can generate multiple low-dose image realizations from one clinical-dose image, so it is useful for validation, optimization, and repeatability studies of image-processing algorithms.

Introduction

In accordance with the “as low as reasonably achievable” (ALARA) principle [76,77], low-dose CT (LDCT) scans should be performed whenever possible. In these scans, the signal-to-noise ratio (SNR) is relatively low, making their interpretation challenging. To overcome this problem the development of denoising methods to increase the SNR and improve image interpretability has been a growing research topic. Efforts in different directions have been made to denoise low dose images. Most of these efforts focus on deep learning-based denoising methods, utilizing neural networks to reduce image noise with as little structure loss as possible [35,36,78–81]. However, there are also methods that use statistical methods to reduce noise in image data [38,55].

In many cases, studies on denoising methods require pairs of low-dose CT (LDCT) and matching clinical-dose CT (CDCT) scans or other types of ground truth. This data can be used to, e.g., train deep learning-based methods or test their robustness in low dose situations but is also helpful for quantitative validation of the performance of the denoising techniques. However, paired sets of LDCT and CDCT scans, or LDCT in general, are not widely available since they are not commonly acquired during clinical work. Also, the validation of new algorithms for abnormality detection in LDCT requires LDCT data. These LDCT scans could be obtained using phantoms, but phantoms have limited anatomic variability and realism, limiting the generalizability of studies using them. Therefore, it is beneficial to be able to generate LDCT images from existing CDCT patient images since this would enlarge the available LDCT data, since CDCT patient images are normally available in larger quantities compared to LDCT patient images.

Several low-dose simulators have been developed that generate LDCT images from CDCT ones, either operating in the projection domain or the image domain. Several of the former use physics-based noise models in unprocessed projection domain [39,40,82,83]. However, in most research settings there is often no access to the unprocessed projection data, since vendors apply specific processing to their projection data, whose details are not disseminated. One could overcome this problem by forward projecting the available images to recreate the unprocessed projection data. However, this would require information about the system geometry and some of the system components, such as the (bowtie-)filter thickness. This information is, again, often not publicly available and, when possible, very challenging and time-consuming to obtain by experimental measurements. Even with access to this information, knowledge about the specific processing of

the vendor is still needed to be able to reconstruct the resulting projection data. These practical reasons limit the applicability of such projection domain methods. Therefore, in these situations, it is desirable to simulate LDCT images directly in the image domain.

Examples of image-domain methods are the noise generators of Niu et al. [84], Britten et al. [85], and Divel et al. [41]. The work of Niu et al. uses deep learning to generate LDCT scans; however, it still requires access to the reconstruction method if one wants to generate multiple LDCT realization from one CDCT. Since their method needs a new CDCT noise realization as input for the LDCT generation, which could be obtained via sinogram simulations and reconstruction. Britten et al. generate LDCT images by adding white noise that is convolved with a kernel determined using LDCT scans. The method does not require access to the reconstruction method at the time of utilization. However, the convolution kernel is shift-invariant, so the noise's position and image-content dependency are not captured. The work of Divel et al. determines the local variance and frequency content of the LDCT noise analytically. However, this method can only be applied with linear reconstruction methods, such as Filtered Back Projection (FBP), if the filter is known. In routine clinical practice, FBP has been primarily replaced by more advanced techniques such as regularized iterative- or deep learning-based reconstruction methods [86,87]. This makes analytical determination of the local variance and frequency content of the noise impossible, therefore these reconstruction methods are called non-linear in practice. In either case, even for truly linear reconstruction methods, analytical determination of these noise properties is impossible without in-depth knowledge on the reconstruction algorithm, such as the filter used and details on other processing steps.

Therefore, in this work, we present a fully image domain-based LDCT image simulator, for situations in which there is no access to projection-domain data, no knowledge about system geometry and components, and no in-depth knowledge on the reconstruction method. Therefore, our method should be seen as an alternative in this limited, but in most research settings realistic, situation and not as an improvement to projection-domain methods. The simulator aims to generate noise with shift- and content-variant magnitude and frequency, as would result from linear and non-linear reconstruction methods, without having access to information about the method itself. Our use case for this study is paired clinical-dose and low-dose brain CT scans, since images of this anatomy acquired with the same protocol were the only ones available to us in sufficient numbers.

Materials & Methods

The proposed method consists of three convolutional neural networks (CNNs) to predict the correct characteristics of the desired LDCT images. The first CNN denoises the CDCT image such that an almost noiseless image is obtained. The second network estimates the standard deviation (σ) for each voxel, i.e., voxelwise, resulting in a σ map of the LDCT noise (σ_{LD} map). The third and final CNN estimates the shift- and content-variant noise power spectra (NPS) of the LDCT noise in local patches. For linear reconstruction methods, such as FBP, using neural networks to predict the σ map and the local NPS is not needed since this can be achieved analytically [41].

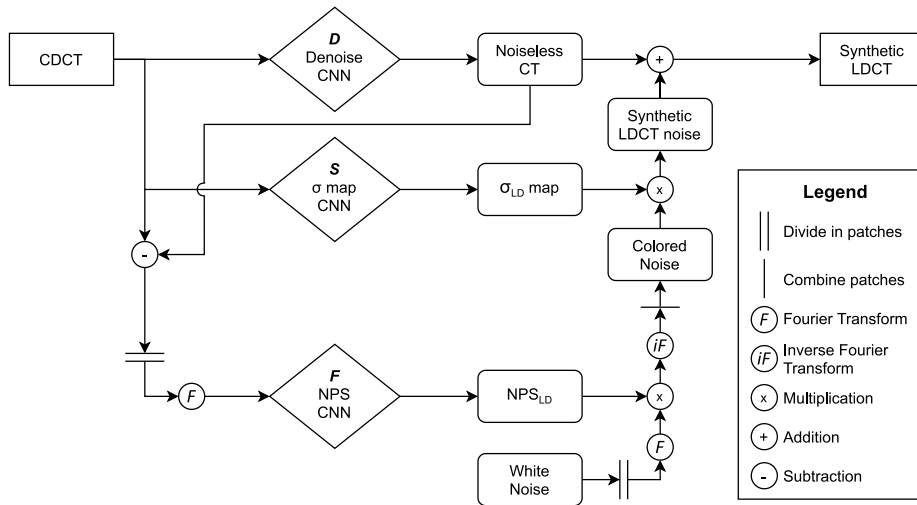


Figure 1: General overview of the proposed method showing the three CNNs and the use of the pipeline for generating synthetic LDCT images.

To create the synthetic LDCT image, a random white noise image is transformed to a corresponding LDCT noise image using the estimated local low dose NPSs and a σ_{LD} map. This ensures the creation of the correct noise correlation based on content and location in the field of view (FOV), and the correct noise magnitude based on content. A synthetic LDCT image is then obtained by adding this LDCT noise image to the noiseless image generated by the first network. Multiple realizations of the LDCT image can be generated by using different random white noise images. A schematic overview of the proposed method is shown in Figure 1.

Datasets

In this work, we use CDCT and LDCT brain images of the same patient. Ethical approval for the retrospective use of this data was obtained from the local ethical

review committee, with patient informed consent waived due to the use of only fully anonymized and already-existing clinical data. The images were obtained from the first two scans of a brain dynamic CT perfusion protocol routinely used in patients with suspected stroke. In this perfusion protocol, a clinical-dose acquisition is performed, followed by a series of low-dose scans after the intravenous injection of contrast agent. This clinical-dose scan is clinically used for offset correction of the perfusion curves. The clinical-dose and first low-dose scan are assumed to have been acquired before arrival of the contrast agent, although this assumption is sometimes violated in our data. All available patient data that followed this protocol and was acquired between 01/01/2019 and 03/30/2020 was obtained retrospectively from the electronic patient archive. Only the data with similar acquisition details were used to ensure consistency in the noise textures. The acquisition details of the scans used are listed in Table 1. All scans were acquired using the Aquilion ONE VISION Edition CT system (Canon Medical Systems, Otawara, Japan) installed at Radboud University Medical Center, Nijmegen, the Netherlands. Each LDCT and CDCT image consists of 320 slices (512×512 pixels) covering 220 mm × 220 mm with a thickness of 0.5 mm reconstructed using an iterative algorithm (AIDR 3D Iterative Reconstruction, Canon Medical Systems) [32].

The image pairs were registered by means of rigid registration using the Elastix toolbox [25]. The registration used was multi-resolution registration (4 resolutions), using Euler transform initialized in the center of gravity. The optimizer used was standard gradient descent and the metric was advanced mean squares. The used parameter file can be found in the online supplemental material. Only the head portion of the images was registered since the head rest is included in the FOV, but the head could have moved relative to it. Therefore, a mask was created for all images by assuming all CT numbers larger than -200 HU were part of the head. This assumption leads in some cases to the false positive inclusion of a small group of voxels that are part of the head rest, but the registration was not affected in these cases. In total, 326 pairs of CDCT and LDCT brain scans were available. From this dataset, 251 pairs were used for training, 25 pairs for validation, and 50 for testing the CNNs, with the partitioning done randomly at the patient level.

Table 1: Acquisition protocol used for the clinical-dose and low-dose brain CT scans used in this study.

Scan	Tube voltage [kV]	Tube current [mA]	Exposure time [s]	Reconstruction method (kernel)	Filter type
CDCT	80	400	0.5	AIDR 3D STD (FC43)	SMALL
LDCT		200			

Denoising network

We assumed that the noise contribution in the CDCT image is additive, i.e., a CDCT image can be decomposed by:

$$CDCT = GT + n_{CD} \quad (1)$$

where GT is the ground truth noiseless attenuation image and n_{CD} is the image noise introduced by the noise in the projection domain and the underlying reconstruction algorithm. Therefore, subtracting the noise from the CDCT would give a noiseless image. A similar assumption for the LDCT image yields the decomposition:

$$LDCT = GT + n_{LD} \quad (2)$$

where n_{LD} here is the image noise in the low-dose measurements.

The denoising network D is based on the method proposed by Lehtinen et al. [88]. They present a method to denoise images without the need for a noiseless GT during training. Based on their work, we can assume that the CDCT and LDCT images have the same underlying GT image, and that their noise realizations are independent, thus making the recovery of the GT possible by minimizing the mean squared error or the mean absolute error between the CDCT and LDCT.

Due to memory constraints, it is difficult to denoise an entire volume at once. Therefore, we simplify the problem and denoise the volume slice-by-slice. However, the noise correlations also span across the axial dimension and last for, in most cases, 7 or fewer slices in both directions. To provide the network with the required information from all three spatial directions, a slab of 15 axial CDCT slices (denoted as Y), with the middle one being the slice of interest, is provided to the network. If the slice of interest is in the first or last 7 slices of the volume, empty slices are added to the slab to ensure that there are 15 slices in a slab. The target is the corresponding LDCT slice. An additional 32 voxels are zero-padded to each side in the axial plane to prevent edge artifacts in the output of the network. So, for denoising the entire CDCT volume image, the network has to be ran 320 times (number of slices), since the denoising is done per individual slice. Figure 2 shows the architecture of the used partly 3D U-Net, the network architecture is based on the work of Adler et al. [36]. The main differences with the architecture used by Adler et al., are the partly-3D nature of our network, the use of instance normalization instead of batch normalization, the use of bilinear up sampling instead of pixel shuffle, and a shorter bottleneck. These differences were introduced since they are more suitable for our purpose. Please note, all convolutions have a stride of 1 and were so-called 'same' convolutions with zero padding, and all pooling layers have a stride of 2.

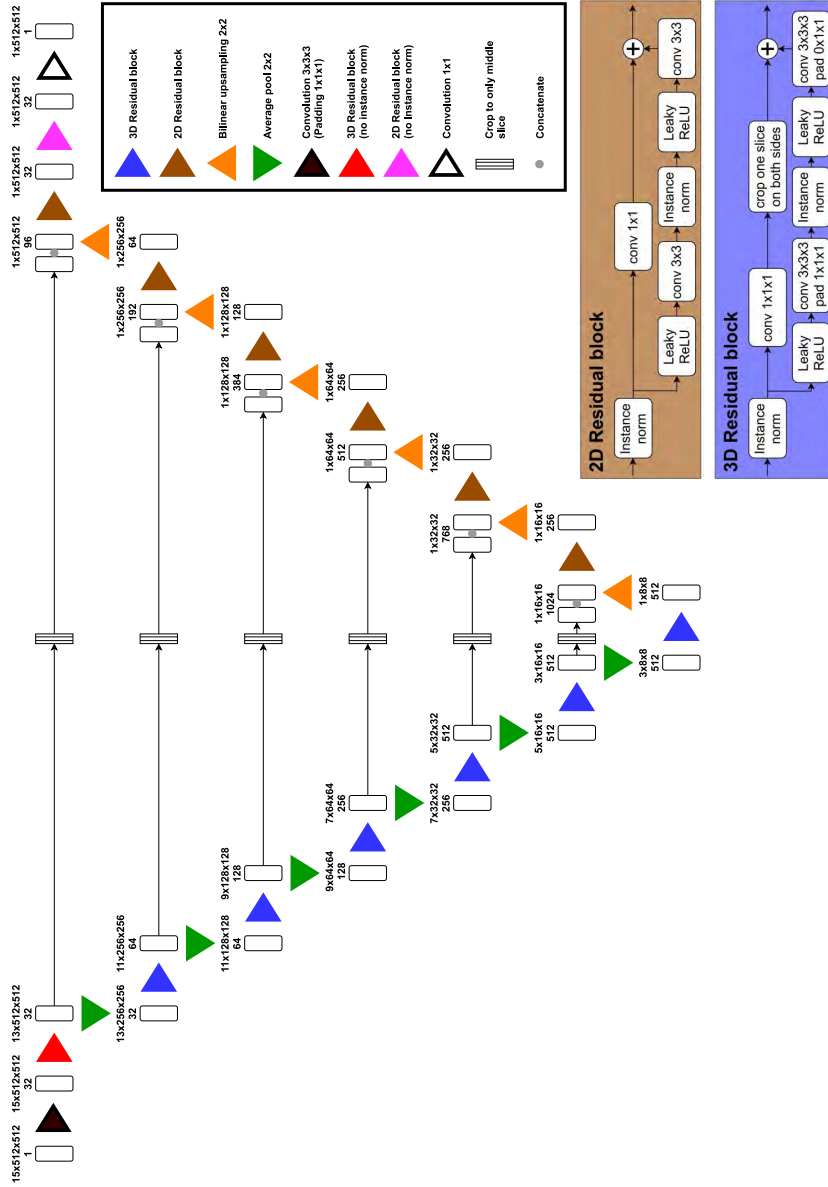


Figure 2: Schematic illustration of the CNN for denoising and for σ_o map estimation. The numbers indicate the image size and the number of feature channels. As shown, the input is a 15 slice-thick slab while the output is a single slice.

The CT numbers of the input data, in Hounsfield Units (HU), were scaled as follows:

$$v = \frac{HU + 1200}{2400} \quad (3)$$

based on the dynamic range present in the images. A mask of the reconstructed FOV is applied to Y before passing it through the network because the reconstructed FOV is circular instead of rectangular. Before calculation of the loss, the reconstruction mask is applied again, and the outer 32 voxels introduced for the zero padding are removed from the network output. The loss function was defined to be an L1-loss:

$$\mathcal{L}_{L1} = \sum |X - D_{\theta}(Y)| \quad (4)$$

where θ are the trainable parameters of the denoising network D , and X is the LDCT slice. The L1-loss was chosen because this loss is less susceptible to large errors in small areas than the L2-loss. The network was trained using the Adam optimizer [89] with a batch size of 4 due to memory constraints. The optimizer's β_1 and β_2 parameters were set to 0.5 and 0.9, respectively, and the weight decay parameter was set to 0. The learning rate α was initially $2 \cdot 10^{-4}$ and was decreased every 2 epochs with a factor $\gamma=0.9$. No warm up was used. The network was trained for 200 epochs since convergence was reached within these 200 epochs. For each sample in each epoch, a single randomly chosen slab of 15 slices per sample was used for training.

Standard deviation map network

The GT obtained for each sample, with network D , provides access to n_{CD} and during training also to n_{LD} , as given by Equations 1 and 2. The ground truth or measured σ_{LD} map, for each n_{LD} noise sample is obtained by calculating the standard deviation for each individual voxel using a $5 \times 5 \times 5$ voxel moving window. All voxels with a value higher than a threshold of 5σ are set to zero since these values are more likely to originate from registration mismatches than real noise. This 5σ threshold is determined in a $10 \times 50 \times 50$ voxel volume of interest (VOI) in the center of n_{LD} .

Due to memory constraints, it is difficult to estimate the σ_{LD} map for the entire LDCT image volume at once. Therefore, we simplify the problem and estimate the σ_{LD} map slice-by-slice. To estimate the σ_{LD} map, the same U-Net architecture (Figure 2), denoted as S , and input CDCT 15-slice slabs Y from Section 2.2, are used. The slab Y is scaled using Equation 3, before passing it through S . During training the measured σ_{LD} map is scaled using Equation 5:

$$\sigma_{LD} \text{ map}_{scaled} = \frac{\sigma_{LD} \text{ map}}{40} \quad (5)$$

where dividing by 40 scales the highest values to approximately 1 and the values inside the skull to approximately 0.4. Again 32 voxels are zero-padded, and the circular reconstruction masking was applied to the input slabs as that described in Section 2.2. This time, the loss function was a L2-loss:

$$\mathcal{L}_{L2} = \sum \left(\sigma_{LD} \text{ map}_{scaled} - S_{\psi}(Y) \right)^2 \quad (6)$$

where ψ are the trainable parameters of the network S . The training of the network was performed with the same parameters as those described in Section 2.2, except that the learning rate α was initially $1 \cdot 10^{-4}$, and the network was trained for 100 epochs.

NPS network

Although the noise structure, i.e., the NPS, differs between n_{CD} and n_{LD} , one could assume that the transformation from one to the other is constant since the noise level in projection domain should be similar for all LDCT images and for all CDCT images. Hence, it should be possible to learn this transformation. However, the noise correlation is shift and signal variant, i.e., the correlation between the noise in a specific voxel and in its surrounding voxels is dependent on where in the FOV the voxels are located and what the tissue content is in those voxels. For example, the noise close to the isocenter of the system is mostly isotropic, while it displays a strong directionality closer to the edges. Also, the noise correlation is broader, i.e., spans a larger distance, in homogeneous areas while it is reduced in areas that include sharp tissue transitions. Therefore, it is not possible to use a single NPS to characterize the entire image. Hence, the transformation from NPS_{CD} to NPS_{LD} is learned and applied patch-based, which also alleviates memory concerns. The size of these patches was set to $64 \times 64 \times 64$ voxels, from which the local NPS was obtained by calculating the squared magnitude of the discrete 3D Fourier transform. The 3D Fourier transform is used since the noise is not only correlated within each 2D slice, but also across slices.

A 3D U-Net based CNN was trained to learn the transformation from the NPS_{CD} patches to the NPS_{LD} patches. The architecture of this network, F , is depicted in Figure 3. All convolutions had a stride of 1 and were so called ‘same’ convolutions with reflection padding, and all pooling layers had a stride of 2.

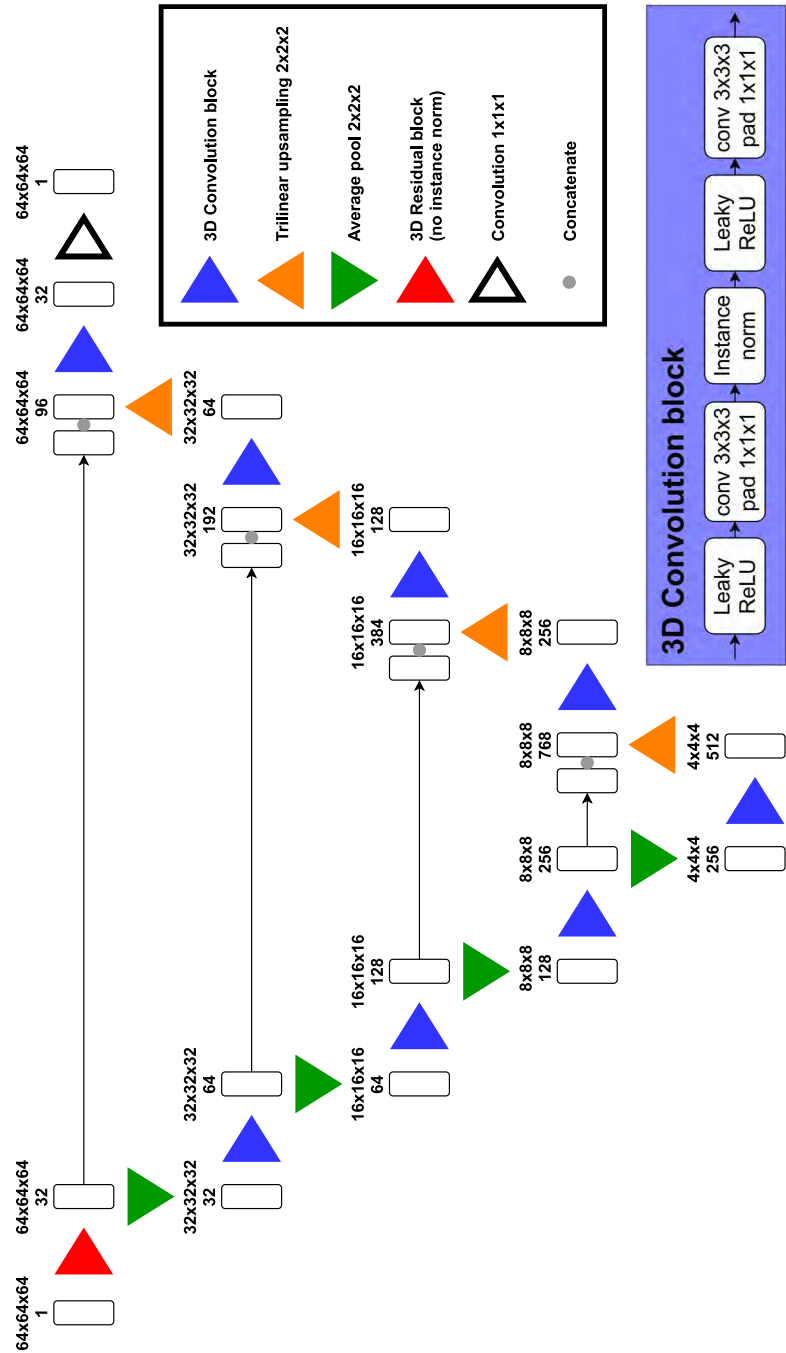


Figure 3: Schematic illustration of the CNN for NPS estimation. The numbers indicate the image size and the number of feature channels. As shown the input and output are both a $64 \times 64 \times 64$ volume.

Before predicting the NPS, a set of three masks was applied to the patch in the image domain to remove any residual registration errors. The first mask is for the circular FOV, the second mask excludes the skull, and the last mask excludes the head rest. The FOV mask was binary eroded with a $9 \times 9 \times 9$ voxel VOI since the registration sometimes moved the FOV causing rings just inside the FOV. The skull mask was obtained by setting the threshold to 100 HU in the denoised CDCT image i.e., the GT. The head rest mask was obtained by assuming that all values between -900 HU and -200 HU in the denoised image correspond to the head rest. The skull mask and head rest mask are both eroded with a $5 \times 5 \times 5$ voxel VOI to ensure the skull and head rest are fully removed. The final combined mask is obtained by subtracting the skull and head rest mask from the FOV mask, a sample from the test set can be seen in Figure 4, together with the n_{cd} before and after applying the mask.

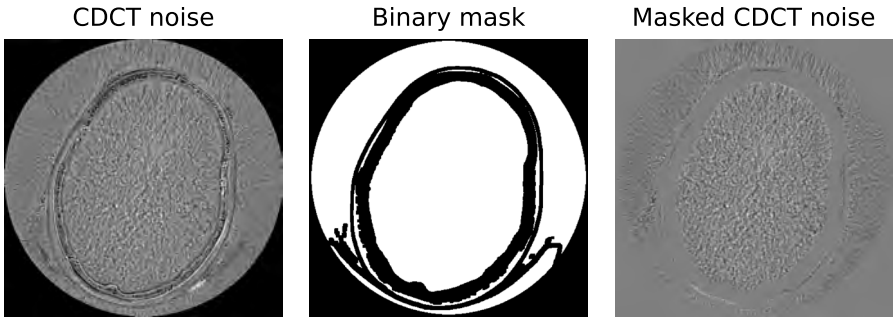


Figure 4: Test set sample CDCT image (left), corresponding binary mask for removing residual error of registration (middle), and test set sample CDCT image after applying binary mask (right).

The values of the NPS determined after applying these masks were scaled before passing them through the network as follows:

$$NPS_{scaled} = \frac{NPS}{\overline{NPS} \cdot 100} \quad (7)$$

where $\overline{NPS}$ is the average value of the NPS. Division by the average value is used to scale the NPSs to the same magnitude, which should aid the network in learning the mapping from CDCT noise NPS to LDCT noise NPS. An additional 8 voxels were zero-padded to each side of NPS_{cd} to prevent edge artifacts in the output of the network. These 8 voxels were removed from the network output. During training, more weight was given to the high frequencies, since these values are smaller and will thus contribute less to the overall loss. Equation 8 describes the weighting function:

$$W(u, v, w) = 1 + ((2 \cdot u)^2 + (2 \cdot v)^2 + (2 \cdot w)^2) \quad (8)$$

where u , v , and w are the relative frequencies in all three spatial directions, ranging from -1 to 1. The loss function was L2-loss, defined by:

$$\mathcal{L}_{L2} = \sum (W \cdot NPS_{scaled_{LD}} - W \cdot F_{\phi}(NPS_{scaled_{CD}}))^2 \quad (9)$$

where ϕ are the trainable parameters of the network F.

Again, the network was trained using the same parameters as the other two networks, but with a learning rate α initially set to $3 \cdot 10^{-5}$, which was decreased every 3 epochs with a factor $\gamma=0.33$. The network was trained for 100 epochs, every epoch one patch per sample was used for training. This patch was in a randomly chosen location, for every epoch, the only constraint is that at least one voxel of the patch has to be inside the combined mask, see Figure 4.

Low-dose CT Image Simulation

As mentioned above, the proposed method to simulate synthetic LDCT images consists of three CNN. The first CNN denoises CDCT per slice, to obtain a fully denoised CDCT volume, GT according to Equation 1. With this estimated GT and CDCT it is possible to obtain n_{CD} (Equation 1). Using the three CNNs described above, a synthetic LDCT is obtained by a fully automated pipeline that does the following:

1. A random white noise volume, the same size as CDCT, with zero mean and unit variance is generated.
2. Network F is used to estimate the NPS_{LD} per patch, with a patch overlap on each side. The patch sizes were chosen to be $64 \times 64 \times 64$ voxels, and the patch overlap was set to 25%.
3. Due to the Hermitian property of the Fourier transform, NPS values at corresponding positive and negative frequencies must be equal, which is not necessarily the case for the output of network F. Hence, the values at corresponding positive and negative frequencies are averaged to have a symmetrical NPS.
4. The estimated NPS are applied to the white noise image patch-by-patch by multiplication in the frequency domain.
5. The overlapping voxels are averaged using a trapezoidal weighting function (ω), in which the weights always add up to unity, similar to the work of Divel et al. [41]:

$$\omega(n) = \begin{cases} \frac{n}{M+1}, & 1 \leq n \leq M \\ \frac{-n+P+1}{M+1}, & P-M+1 \leq n \leq P \\ 1, & otherwise \end{cases} \quad (10)$$

where n is the current voxel, M is the number of voxels that overlap, and P is the patch size. Note that the patches at the border need to have a larger overlap to fit the FOV size of $320 \times 512 \times 512$.

6. The colored noise patches are scaled to each have zero mean and unit variance.
7. The estimated σ_{LD} map is used to scale the noise in each location to the correct magnitude, by means of multiplication.
8. To obtain the synthetic LDCT image the generated n_{LD} is added to the GT.

Validation

All three simulation steps were validated separately. The denoising network, D_θ , was validated by determining the ratio of the standard deviations of the network output GT to that of CDCT image, denoted $\sigma_{improve'}$, and by validating if the removed noise has zero mean i.e., if a bias was not introduced. To calculate $\sigma_{improve'}$, the standard deviations of the CDCT image and the GT were evaluated in the cerebrospinal fluid (CSF) of the 80 central slices, since this was considered to be the most homogeneous area in brain CT images. The CSF was identified by thresholding the denoised image between -10 HU and 40 HU, followed by binary eroding with a $9 \times 9 \times 9$ voxel VOI, and is shown in Figure 5. The bias of the noise was determined for the 80 central slices in an 80×80 VOI in the approximate center of the brain, avoiding the nasal cavity and eye sockets.

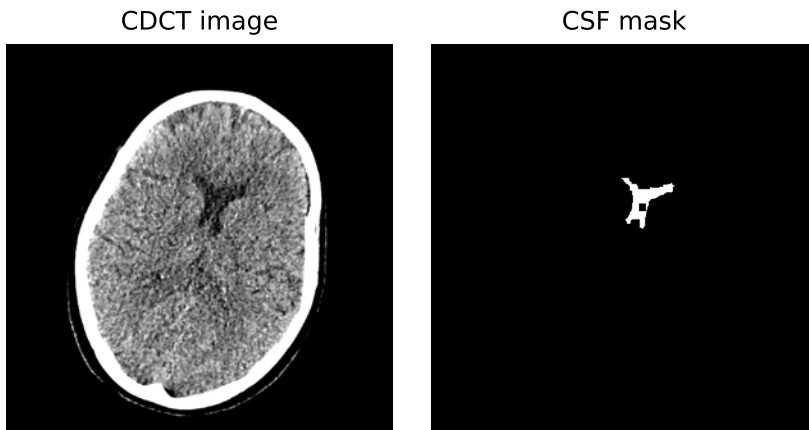


Figure 5: CDCT image with the dark area in the center of the brain being CSF, WL/WW: 50/100 (left) and corresponding CSF mask after erosion (right).

To assess the performance of the standard deviation network, S_ψ , the ROIs used for verifying the bias of the denoising network above were used here. For the measured σ_{LD} map and estimated σ_{LD} map, the mean value inside the VOI was calculated and compared to that of the test set.

To validate if the correct noise structure is present in all three spatial dimensions, different normalized NPSs (nNPSs) from within a $64 \times 64 \times 80$ voxel VOIs were calculated. Within this VOI, the average of the 2D nNPS within each of the 80 slices of 100 different noise realizations were used to validate the axial noise structure. For the noise structure in the axial direction (across slices), the 1D nNPS was determined along the 80 voxels for all voxels in one realization of the VOI.

Since the aim of the proposed method is not to outperform other methods, but simply to present an alternative to methods that need access to the projection data and/or the reconstruction algorithm, there is no comparison made to these methods. Instead, the quality of the synthetic LDCT volumes was tested by performing two proof-of-concept denoising studies. In the first study, a CNN with identical architecture and training details as the denoising network, D , was trained to denoise synthetic LDCT images. The targets for the network were the corresponding CDCT images. The training, validation, and test sets consisted of the synthetic LDCT images obtained during inference of the proposed method. The sets were partitioned in the same way as for the training of the three networks of the proposed method, i.e., 251 images for training, 25 for validation, and 50 for testing. The images were scaled using Equation 3. The network was trained with a batch size of 4 using the Adam optimizer, and L1-loss. The optimizer's β_1 and β_2 parameters were set to 0.5 and 0.9, respectively, and the weight decay parameter was set to 0. The learning rate α was initially $2 \cdot 10^{-4}$ and was decreased by a factor $\gamma=0.9$ every 2 epochs. No warm up was used. In total the network was trained for 200 epochs. For each sample in each epoch, one randomly chosen slab of 15 slices per sample was used for training. The performance of the CNN was tested on two test sets, the test set with synthetic LDCT images (50 samples) and the original test set with real counterparts (also 50 samples) of these LDCT images. For both test sets the σ is compared inside the CSF of the 80 central slices before and after denoising just like for network D . The CSF is identified by thresholding the denoised image between -10 HU and 40 HU, followed by binary eroding with a $9 \times 9 \times 9$ voxel VOI, an example of a CSF ROI is depicted in Figure 5. In this way the $\sigma_{improve}$ ratio can be obtained, and it could be verified if the CNN show similar behavior for the real vs. the synthetic LDCT images.

In the second proof-of-concept study, an Anisotropic Diffusion Filter (ADF) [90] was used for denoising the same two test sets, one with real and one with synthetic LDCT, again, each consisting of 50 images. The ADF uses the following equations:

$$\frac{\partial LDCT}{\partial t} = \nabla c \cdot \nabla LDCT + c(x, y, z, t) \Delta LDCT \quad (11)$$

$$c = \frac{1}{1 + \left(\frac{\|\nabla LDCT\|}{K} \right)^2} \quad (12)$$

Where K and ∂t are parameters of the filter, that are set to 30 and 0.2, respectively. ∇ is the gradient operator, Δ is the Laplacian operator, and c is the diffusion coefficient. Five iterations of the ADF were used to denoise the real and synthetic LDCT images. The $\sigma_{improve}$ ratio, was again compared, using the same CSF ROI as for the first proof-of-concept study. In this way, it could be verified if the results of the gradient-based filter were similar for the real and synthetic LDCT images. With these two denoising studies we aim to show that our proposed method for generating synthetic LDCT images could be used to estimate performance on real LDCT images.

Results

Denoising of CDCT

An example of a CDCT from the test set and its denoised version are shown in Figure 6. In Table 2 we report the median and interquartile ranges (IQR) of the standard deviation ratio, $\sigma_{improve}$, in the CSF, and the bias introduced by denoising. The results show that little bias is introduced, since the mean of the removed noise is close to zero. For some extreme cases a more significant bias is introduced (negative extreme -4.1, positive extreme +8.1). Visual inspection shows that this happened for cases where abnormalities were present in the images, as can be seen in Figure 7, the left image shows calcifications, and the right image shows a hemorrhage, both indicated by red arrows. The larger bias could be explained by the fact that the network is mainly trained on data without abnormalities.

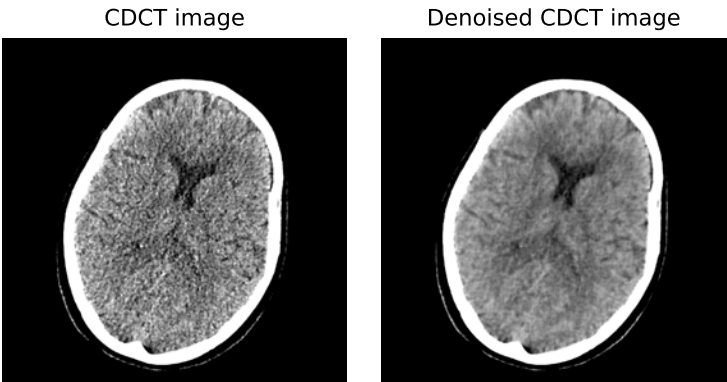


Figure 6: Single slice from CDCT volume (left), with the corresponding denoised slice of CDCT volume (right), WL/WW: 50/100.

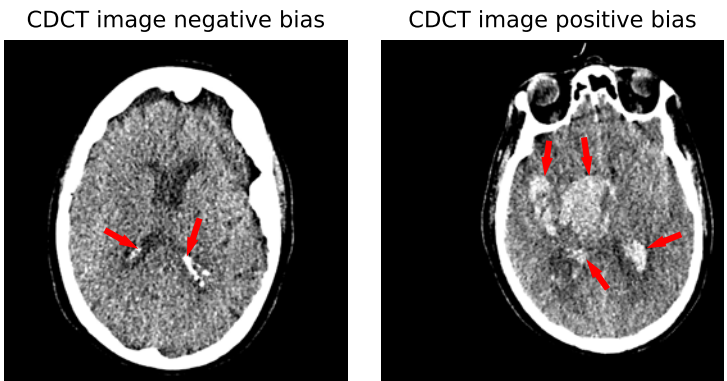


Figure 7: CDCT images with calcifications, red arrows, which resulted in the largest negative bias (left) and with a hemorrhage, red arrows, which resulted in the largest positive bias (right), WL/WW: 50/100.

Table 2: The results of denoising the CDCT images. The median and IQR of the $\sigma_{improve}$ and the introduced bias for the data of the test set. A boxplot of the results can be found in the online supplemental material.

	Mean	Median	IQR
$\sigma_{improve}$ (-)	1.78	1.71	1.61 – 1.95
Bias (HU)	+0.7	+0.7	-0.6 – +1.5

Standard deviation map

Figure 8 shows an example of a measured and estimated σ_{LD} map from the test set. Note that the measured σ_{LD} map is not as homogeneous as the latter since the former is a noisy estimate of the true underlying σ at each voxel. Further increase of the window size used to determine the measured σ would lead to introduction of errors in the σ_{LD} map at the edge of image features. The estimated σ_{LD} map is, as expected, homogeneous within the skull. A vertical line profile of the measured and estimated σ_{LD} maps shown in Figure 8 is plotted in Figure 9. The median, IQR, and difference of the averaged value in the VOI inside the skull of all measured and estimated σ_{LD} maps in the test set are reported in Table 3. The results show that network S can accurately estimate the σ of the noise in the LDCT image.

Table 3: The median σ and IQR of the σ for both the real and generated LDCT samples of the test set. Please note, the error is defined as estimated - measured. A boxplot of the results can be found in the online supplemental material.

	Mean	Median	IQR
Measurement σ (HU)	16.3	16.3	15.3 – 16.6
Estimate σ (HU)	16.2	16.1	15.6 – 16.7
Difference σ (HU)	0.0	+0.1	-0.2 – +0.3

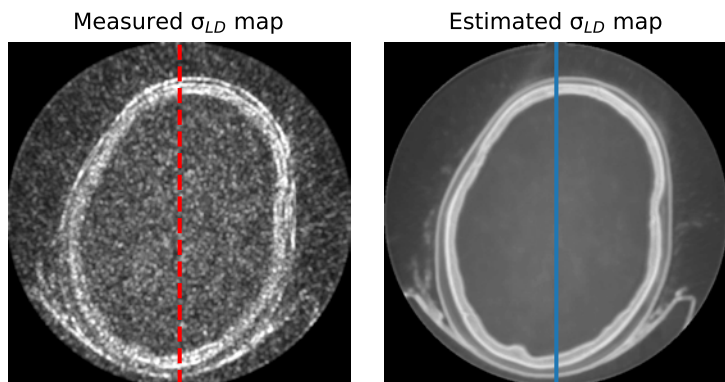


Figure 8: Measured (left) and estimated (right) σ_{LD} map, WL/WW: 20/40.

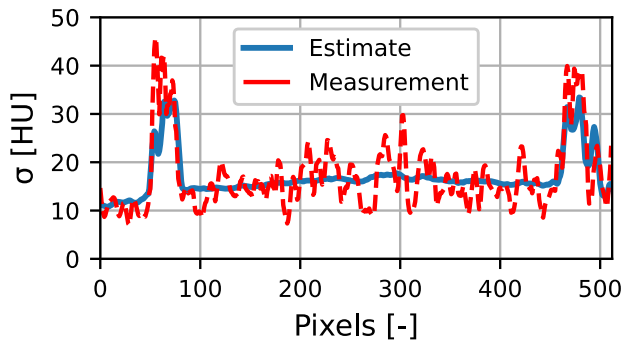
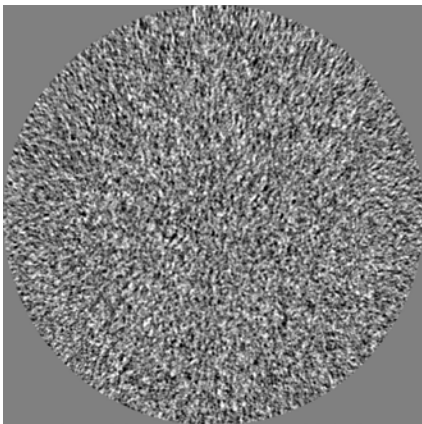


Figure 9: Line profile of measured and estimated σ_{LD} map, obtained in vertical profiles shown in Figure 8.

Noise Power Spectrum

Figure 10 shows an example of a synthetic LDCT noise image from the test set before and after scaling by the σ_{LD} map. Figure 11 shows real and synthetic LDCT noise, including zoomed-in areas of the inside of the skull and the periphery. The results show that the NPS network F is able to generate noise with visually correct location-dependent noise although the noise in the synthetic LDCT image has slightly higher frequency, since the noise looks visually a bit grainier. In addition, the real LDCT noise shows some subtle streaks in the periphery, possibly caused by beam-hardening, which are not present in the synthetic LDCT noise.

Unscaled synthetic LDCT noise



Scaled synthetic LDCT noise

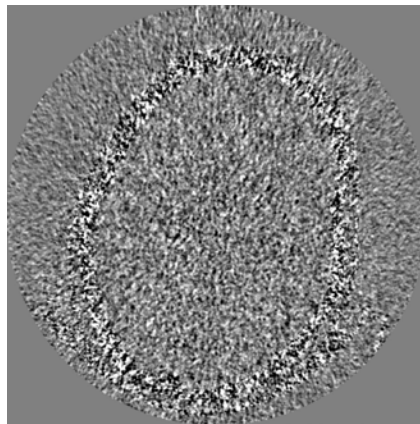


Figure 10: Unscaled synthetic LDCT noise (left) and scaled synthetic LDCT noise (right).

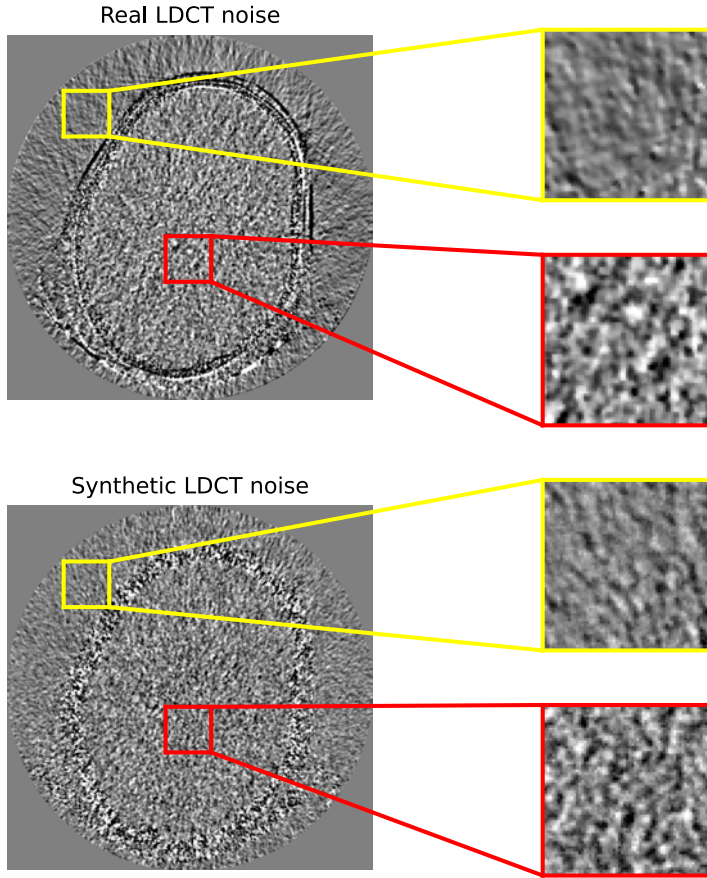


Figure 11: Real and synthetic LDCT image (left), with zoomed-in parts showing the noise texture inside and outside the skull for both real and synthetic LDCT noise (right), WL/WW: 0/80.

The 2D nNPS of these ROIs are shown in Figure 12. Please note that the nNPS of the synthetic noise is an average of 100 realizations, and therefore appears smoother than the real noise nNPS, which is derived from only 1 realization. This smoothness allows for a clearer comparison of the result in the profiles in Figure 13. The resulting 2D nNPS clearly show that our model generates different noise patterns at different locations, including different levels of anisotropy similar to the real LDCT noise. Figure 13 shows the line and circle profiles across the nNPS shown in Figure 12, the line profiles indicate that the network F slightly underestimates the noise power at low frequencies and a slightly overestimates it at high frequencies. Please note, that the y-axis of the line profile is plotted in log scale. The circle profiles suggest that the network F can accurately estimate the degree of anisotropy, since the profiles show the same trend along the full rotation.

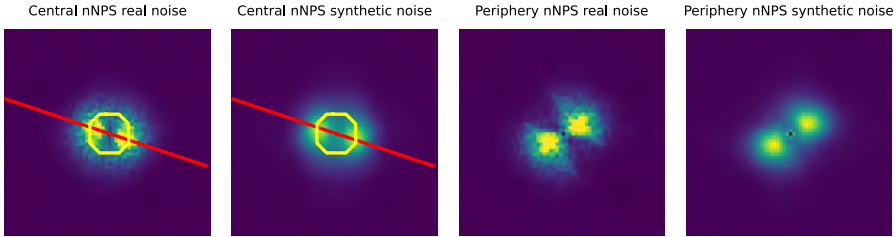


Figure 12: nNPS of real (left) and 100 synthetic (middle left) LDCT noise realizations inside the skull (red ROI) and of corresponding ROIs outside the skull (yellow ROI, middle right and right) of image in Figure 11, WL/WW: 0.002/0.004. The red line and yellow circle mark the profiles shown in Figure 13.

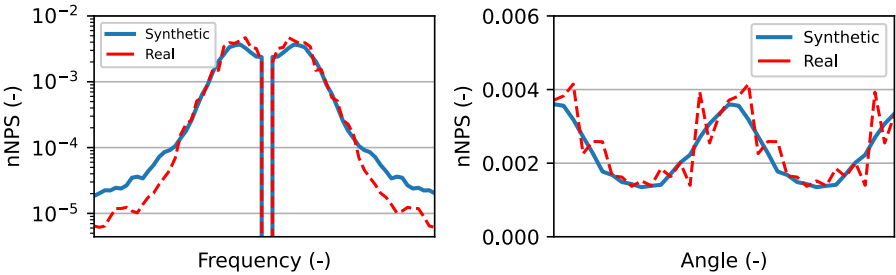


Figure 13: Line profile of red line crossing nNPS (left) and yellow circle profile around nNPS (right) depicted in Figure 12.

The axial nNPS of the red ROI of Figure 11 is shown in Figure 14. The same trend is observed as for the other two directions – slight underestimation at low frequencies and a slight overestimation at high frequencies.

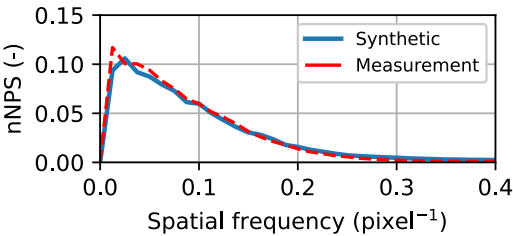


Figure 14: The synthetic and measured nNPS across slices (axial direction) of a LDCT volume image.

Figure 15 depicts three examples of a real and synthetic LDCT image pair showing good similarity. The main visual difference between them is the less clearly defined structures, e.g., the CSF indicated by red arrows, due to the denoising operation in our proposed method.

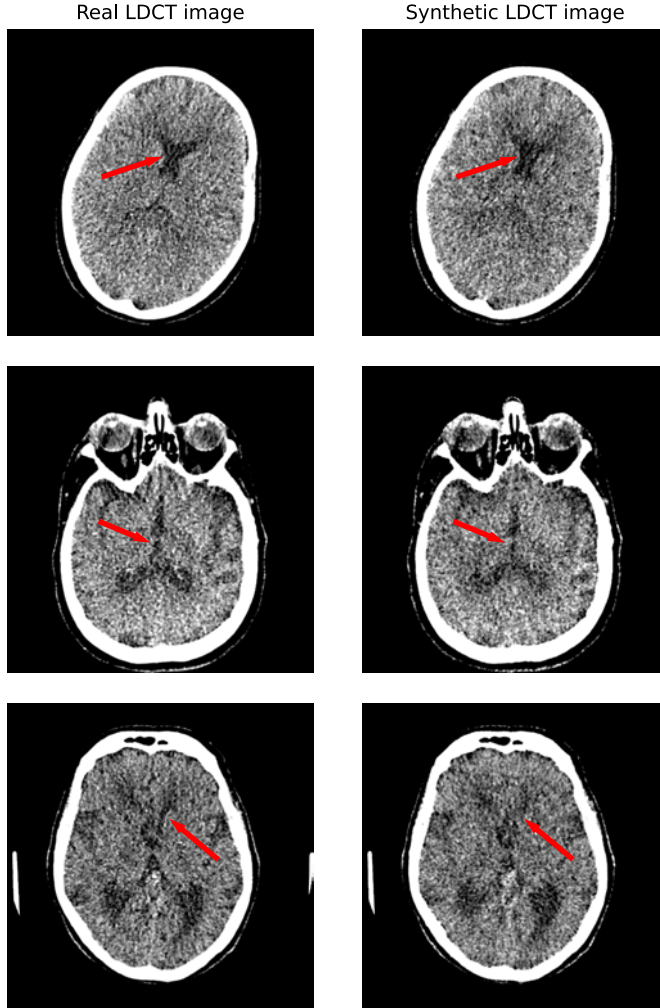


Figure 15: Single slices from three example cases of the test set, each showing the real LDCT image (left) and corresponding synthetic LDCT image (right), WL/WW: 50/100. The overall magnitude and texture of the noise can be seen to be similar between the real and synthetic images. However, as can be seen by the red arrows, the borders between the CSF and the brain matter tend to be less clearly depicted in the synthetic LDCT image.

Proof-of-concept denoising study

The results of the CNN- and ADF-based denoising studies are shown in Table 4. Both studies yield slightly better results on the synthetic test set than on the real one. The CNN denoising resulted in slightly more comparable values for $\sigma_{improve}$ than the ADF. For both studies the real test sets results in a wider IQR. However, in general the results of the synthetic and real test set are very similar.

Table 4: The median and IQR of $\sigma_{improve}$ using the denoising CNN and ADF, and the error for both the real and synthetic test set. Please note, the difference is defined as synthetic – real. A boxplot of the results can be found in the online supplemental material.

	CNN denoising			ADF denoising		
	mean	median	IQR	mean	median	IQR
$\sigma_{improve}$ real (-)	1.69	1.71	1.60 – 1.79	1.85	1.89	1.77 – 1.94
$\sigma_{improve}$ synthetic (-)	1.71	1.71	1.64 – 1.77	1.90	1.92	1.85 – 1.97
Difference (-)	+0.01	0.00	-0.02 – +0.04	+0.05	+0.05	+0.01 – +0.10

Computation time

The synthetic LDCT images were generated using a Linux system with 128 GB RAM, Intel(R) Xeon(R) Silver 4210R CPU @ 2.40GHz, and a 48 GB Nvidia RTX A600 GPU. To obtain a good computation time estimate 50 synthetic LDCT samples were generated. The generation of the full denoised CDCT volume and σ_{LD} map takes approximately 42 seconds and 44 seconds on average, respectively. The generation of a synthetic LDCT noise image takes approximately 22 seconds. Note that this also includes the time to obtain the estimated NPS_{LD} patches; thus, generating multiple synthetic LDCT noise realizations for one CDCT case is much faster. Generating 50 noise realizations, of one original CDCT case, takes a total of approximately 200 seconds, so 4 seconds on average.

Discussion

In this work, we developed and validated a method to create an arbitrary number of LDCT realizations from a given CDCT image. This method makes low dose CT data more readily available, especially in a research setting without access to every detail of the workings and algorithms of clinical CT systems. The resulting synthetic low dose CT images can be used for (pre-)training deep learning networks, testing robustness of deep learning algorithms on low dose protocols, validating statistical denoising methods, perform repeatability studies, because multiple noise realization can be obtained for one low dose CT image, and other purposes.

In contrast to earlier work, our method works completely in the image domain, does not require access to the projection data, knowledge about the reconstruction algorithm, nor information about system geometry and components. Our method works for images obtained with a non-linear reconstruction algorithm, with no given loss of generality to allow it to work on other non-linear or linear reconstruction algorithms. However, its performance with other reconstruction algorithms was not tested, given that we do not have access to paired datasets using other algorithms.

The limited information needed and the fact that it works fully in the image domain make the method easier and faster to use than projection domain methods.

We have developed and evaluated our method using LDCT and CDCT brain images. The first, denoising, network had a median decrease in σ in the cerebrospinal fluid of a factor 1.71 and introduced a median bias of +0.7 HU, indicating good noise reduction since the denoised image looks visually noiseless compared to the LDCT image and with little bias. The factor of 1.71 might appear relatively low, but it should be noted that the CDCT being denoised already has a low standard deviation, since it is a reconstruction obtained using a hybrid iterative algorithm. Also, by denoising the image further, anatomically-relevant details might be lost. The second network, for σ_{LD} map estimation, had a median error of +0.1 HU showing the ability of the network to match the σ of the real LDCT noise well. When it comes to the generation of realistic noise structure, the method only shows slight underestimation at low frequencies and a slight overestimation at high frequencies for the nNPS, and it shows excellent agreement with the real LDCT noise in degree of anisotropy. The synthetic LDCT images showed a good similarity with the real low dose data, although some edges are less sharp because of the denoising performed in our method. The denoising network D introduces a bias for cases that contain abnormalities. Therefore, for application of this method to iodine cases, the networks should be trained on a dataset containing such type of cases.

To show that the differences in noise and resolution loss are not relevant to the intended application of the simulated LDCT images, which is, for instance, to validate denoising methods, two proof-of-concept denoising studies were performed. With these two proof-of-concept denoising studies, we showed that the result of applying a filter to real data is very similar to applying it to synthetic data. Therefore, the synthesizing method appears to meet its goal of generating lower-dose images that can be used for development or testing of image processing algorithms. Also, the IQR of the $\sigma_{improve}$ in these denoising studies was larger for the real LDCT images. The resulting narrower $\sigma_{improve}$ distribution when filtering synthesized data could indicate that the variability in image noise, i.e., the structure and σ , is larger in the real noise than in the simulated noise images. This could be explained by the network trying to find an average noise structure and σ to mimic.

One of the primary assumptions in our work is that the CDCT and LDCT images represent the same underlying object. This does not completely hold as the registration of both scans is not ideal. Moreover, the denoising CNN used for obtaining the underlying ground truth or object, introduces resolution loss, as seen

in the final synthetic image (Figure 15). Therefore, small anatomical structures might be lost, and edges are smoothed. However, for the purpose of predicting post-processing performance (mainly denoising), this may be of minor importance. In case one desires to use the proposed method for a different purpose, one should be aware of these non-idealities and assess if these are crucial for the intended purpose.

A limitation of this work is that the method is currently only working for the simulation from a single clinical-dose level to a single low-dose level leading to the network's need to be retrained for different dose levels. However, if data of multiple lower dose levels is available, then an extra input parameter to the networks to indicate which lower dose level is desired could be added. In addition, this study involved the use of a single reconstruction algorithm, while it can be expected that the transformation from clinical-dose NPS to low-dose NPS might be different for other reconstruction methods. Hence, the networks should be retrained if one intends to simulate images obtained with another reconstruction algorithm. However, none of the LDCT simulation steps involve imposition of parameters specific to the reconstruction algorithm used in this study, and therefore it is expected that the methods described generalize well to other reconstruction algorithms.

Our dataset consists of only brain data. In case the method is applied to other anatomies the models should be retrained. Also, in case of, for instance, abdomen scans the variability in patient size will be larger. In addition, such body scans should probably be acquired using some form of automatic exposure control, so the noise content of images across body sizes is more consistent. However, given the current state of clinical practice, this would be expected for any contemporary dataset.

Conclusion

A low-dose CT simulation pipeline was developed and validated for a non-linear reconstruction algorithm. The simulation is fully image-domain based and uses three CNNs. The results showed good agreement in simulating both noise magnitude (σ) and noise structure. The proposed method enables the possibility to generate large datasets of LDCT images for testing post-processing algorithms, such as the ADF or other spatial filters, or for deep-learning tasks, such as abnormality detection in LDCT images. The possibility to generate multiple LDCT images from one CDCT image also enables the opportunity to perform repeatability studies. Since no reconstruction algorithm-specific assumption was made, it is expected that after re-training the method can also be used for other non-linear, as well as linear, reconstruction algorithms.

Supplementary Material

Appendix A – Registration parameters

The elastix parameters used for registration can be found below.

```
(UseDirectionCosines "true")
// ***** ImageTypes
(FixedImageDimension 3)
(MovingImageDimension 3)
(ErodeFixedMask "false")

// ***** Components
(Registration "MultiResolutionRegistration")
(Metric "AdvancedMeanSquares")
(Transform "EulerTransform")
(HowToCombineTransforms "Compose")
(Optimizer "StandardGradientDescent")

// ***** Image Sampler
(ImageSampler "RandomSparseMask" "RandomSparseMask" "RandomSparseMask")
(Interpolator "BSplineInterpolator")
(ResampleInterpolator "FinalBSplineInterpolator")
(Resampler "DefaultResampler")
(FixedImagePyramid "FixedRecursiveImagePyramid")
(MovingImagePyramid "MovingRecursiveImagePyramid")

// ***** Internal Image Type
(FixedInternalImagePixelType "float")
(MovingInternalImagePixelType "float")

// ***** Initialization
(AutomaticTransformInitialization "true")
(AutomaticTransformInitializationMethod "CenterOfGravity")
(AutomaticScalesEstimation "true")

// ***** Multi resolution
//The number of resolutions, the downsampling schedules of the pyramids are
//determined automatically. It is also possible to give the schedules manually
```

```

(NumberOfResolutions 4)
(ImagePyramidSchedule 16 16 16 8 8 8 2 2 2 1 1 1)

// ***** Optimization
//Maximum number of iterations in each resolution level:
(MaximumNumberOfIterations 500)
//Order of B-Spline interpolation used in each resolution level:
(BSplineInterpolationOrder 3)
//Number of grey level bins in each resolution level for mutual information:
(NumberOfHistogramBins 32)
//Number of spatial samples used to compute the mutual information in each
resolution level:
(NumberOfSpatialSamples 2048)
(NewSamplesEveryIteration "true")
(CheckNumberOfSamples "true")
(UseNormalization "false")
(MaximumNumberOfSamplingAttempts 0)

(NumberOfSamplesForSelfHessian 100000)
(SelfHessianSmoothingSigma 1.0)
(SelfHessianNoiseRange 1.0)

//a_k = a/(A+k+1)^alpha
(SP_a 0.0001)
(SP_alpha 0.602)
(SP_A 50.0)

// ***** Miscellaneous
(WriteTransformParametersEachIteration "false")
(WriteTransformParametersEachResolution "false")
(ShowExactMetricValue "false")

// ***** Result
//Order of B-Spline interpolation used for applying the final deformation:
(FinalBSplineInterpolationOrder 3)
(DefaultPixelValue -1060)
(WriteResultImage "true")
(ResultImageFormat "mhd")
(CompressResultImage "false")

```

Appendix B – Boxplots of Results

This appendix contains the boxplots of the results shown in Tables 2, 3, and 4.

Figure B1 shows an example boxplot indicating the meaning of each line. The end of the whiskers is the maximum and minimum or the last value falling within the 75^{th} percentile + $1.5 \cdot IQR$ upper bound or 25^{th} percentile - $1.5 \cdot IQR$ lower bound, where IQR stands for interquartile range.

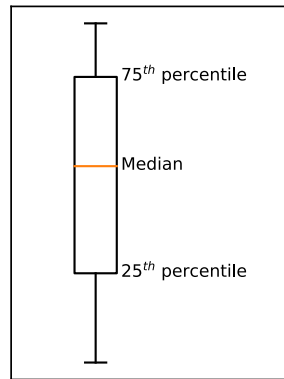


Figure B1: Example boxplot indicating the 25^{th} percentile, the median, and the 75^{th} percentile.

Figure B2 shows the boxplots of standard deviation improvement and bias introduced by the denoising network, D , on the test set (Table 2 of manuscript).

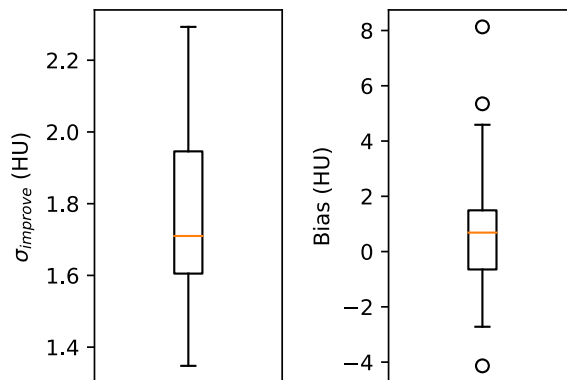


Figure B2: Boxplots of the standard deviation improvement (left) and the bias (right) of the test set results from denoising network D .

Figure B3 shows the boxplots of the measured and estimated, by network S, voxelwise standard deviation of the test set (Table 3 of manuscript).

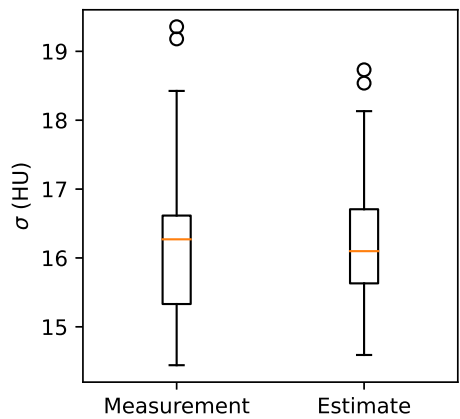


Figure B3: Boxplots of the measured (left) and estimated by network S (right) standard deviation of the test set.

Figure B4 shows the boxplots of the standard deviation improvement for the proof-of-concept denoising studies using ADF filter and CNN (Table 4 of manuscript). Indicating good agreement between the real and synthetic data for both ADF and CNN-based denoising.

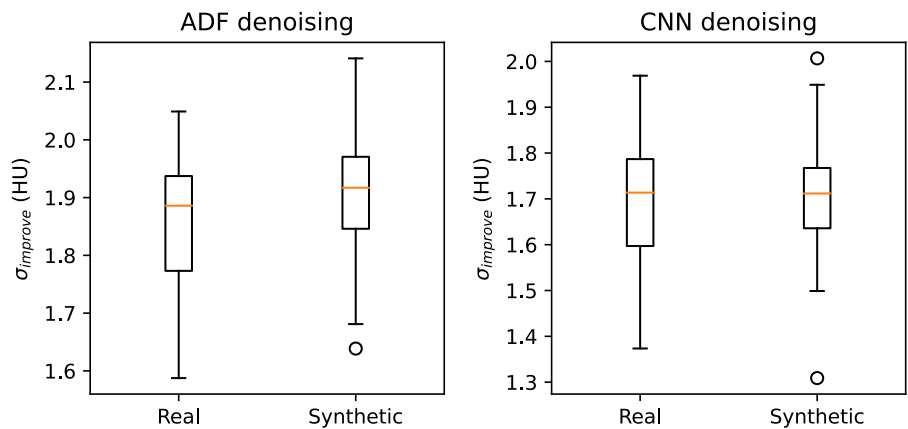


Figure B4: Boxplots of real and synthetic standard deviation improvement for ADF denoising (left), and CNN denoising (right).



Chapter 4

Characterization of the 4D Similarity Filter

Original title: Performance evaluation of a 4D similarity filter for dynamic CT angiography imaging of the liver

Sjoerd A. M. Tunissen, Ewoud J. Smit, Mikhail Mikerov, Mathias Prokop,
and Ioannis Sechopoulos

Published in: Medical Physics

ABSTRACT

Background

Dynamic CT angiography of the abdomen provides perfusion information and characteristics of the tissues present in the abdomen. This information could potentially help characterize liver metastases. However, radiation dose has to be relatively low for the patient, causing the images to have very high noise content. Denoising methods are needed to increase image quality.

Purpose

To investigate the performance, limitations, and behavior of a new 4D filtering method, called the 4D Similarity Filter (4DSF), to reduce image noise in temporal CT data.

Methods

The 4DSF averages voxels with similar time attenuation curves (TACs). Each phase is filtered individually using the information of all phases except for the one being filtered. This approach minimizes the bias towards the noise initially present in this phase. Since the 4DSF does not base similarity on spatial proximity, loss of spatial resolution is avoided.

The 4DSF was evaluated on a 12-phase liver dynamic CT angiography acquisition of 52 digital anthropomorphic phantoms, each containing one hypervascular 1cm-lesion with a small necrotic core. The metrics used for evaluation were noise reduction, lesion contrast-to-noise ratio (CNR), CT number accuracy using peak-time and peak attenuation of the TACs, and resolution loss. The results were compared to those obtained by the time-intensity profile similarity (TIPS) filter, which uses the whole TAC for determining similarity, and the combination 4DSF followed by TIPS filter (4DSF + TIPS).

Results

The 4DSF alone resulted in a median noise reduction by a factor of 6.8, which is lower than that obtained by the TIPS filter at 8.1, and 4DSF + TIPS at 12.2. The 4DSF increased the median CNR from 0.44 to 1.85, which is less than the TIPS filter at 2.59 and 4DSF + TIPS at 3.12. However, the peak attenuation accuracy in the TACs was superior for the 4DSF, with a median attenuation decrease of -34 HU compared to -88 HU and -50 HU for the hepatic artery when using the TIPS filter and 4DSF + TIPS, respectively. The median peak-time accuracy was inferior for the 4DSF filter and 4DSF + TIPS, with a time shift of -1 phases for the portal vein TAC compared to no

shift in time when using the TIPS. The analysis of the full-width-at-half-maximum (FWHM) of a small artery, showed significantly less spatial resolution loss for the 4DSF at 3.2 pixels, compared to the TIPS filter at 4.3 pixels, and 3.4 pixels for the 4DSF + TIPS.

Conclusion

The 4DSF can reduce noise with almost no resolution loss, making the filter very suitable for denoising 4D CT data for detection tasks, even in very low dose, i.e., very high noise level, situations. In combination with the TIPS filter the noise reduction can be increased even further.

Introduction

Dynamic CT angiography, also known as dCTA or 4D-CTA, aims to image the flow of injected contrast agent through the body by acquiring multiple CT images over time. From these dynamic CTA images, perfusion parameters of organs and lesions can be calculated, which can be displayed as images and are referred to as CT perfusion (CTP). Due to the repeated image acquisitions, the radiation dose used per single image acquisition must be kept low to limit the total dose of the entire protocol. However, this low radiation dose causes the individual images of a dCTA sequence to have a very high noise level and consequently very low signal-to-noise ratios (SNR). As a result, subjective radiologist interpretation and quantitative analysis of the dynamic information in these scans is challenging, therefore limiting the clinical impact of these sequences. However, this type of scans is used to detect endoleaks after endovascular aneurysm repair [91,92].

If the SNR could be increased, the resulting image sequence could provide extensive information about vascular flow patterns and organ and lesion perfusion [20,93]. In addition, an increase in SNR allows for the acquisition of a higher number of images in the protocol at the same total dose, thus achieving better temporal sampling of the perfusion dynamics. This improved information, either from higher SNR per frame or a higher temporal sampling, could potentially result in better treatment decisions and, therefore, increase the chances of a positive treatment outcome.

There have been attempts to enhance the image quality of single low-dose CT scans and of low-dose dCTA sequences. Most of these rely on advanced reconstruction algorithms involving iterative [32,86] or deep learning-based approaches [33,59,94]. However, these methods process each image separately, meaning they do not use the additional information present in the temporal dimension. Other noise-reduction methods consider the temporal dimension, like the partial temporal nonlocal means (PATEN) method [95], or the time-intensity profile similarity (TIPS) filter [37]. However, both the TIPS and PATEN filter use a spatial neighborhood making it prone to resolution loss when the noise level is high. Finally, the k-means clustering guided bilateral filter (KMGB) method [38], does suffer less from substantial spatial resolution loss. However, at very high noise levels, the difference between time attenuation curves (TAC) of similar tissues increases as the TACs are increasingly dominated by noise rather than signal. This makes k-means clustering hard and even the temporal mean will be noisy. Also, choosing the number of clusters is not a trivial task.

Recently, Smit and Prokop proposed a 4D similarity filter (4DSF) that intends to reduce the image noise in dynamic images with little to no spatial resolution loss and that still works at very high noise levels [55]. The 4DSF uses only the information in the temporal domain, ignoring any spatial information, and hence it is expected to maintain the spatial resolution of the image. In addition, by filtering each phase in the TAC separately, it is designed to avoid finding TACs with similar noise present in each phase of the TAC. The promise of the 4DSF filter has led to it being introduced for clinical use in dCTA protocols acquired with certain commercial CT systems, such as for brain perfusion and cardiac perfusion imaging applications [96–100].

However, to date, the performance of the 4DSF in CT has not been extensively evaluated. Therefore, the aim of this study is to quantitatively characterize and evaluate the performance of the 4DSF in dCTA imaging. For this work, imaging of the liver is used as a case study since multiphase imaging of the liver has been shown to result in good lesion detectability [101]. However, CT perfusion and dCTA of the liver could potentially improve liver lesion detectability further [101,102], improve lesion characterization [93,103–106], and help monitor treatment response [106,107]. The introduction of the 4DSF for liver dCTA could push these benefits into the clinical realm by achieving these clinical applications while lowering the radiation dose needed to achieve the required minimum image quality. This evaluation is performed using realistic computer simulated dCTA images of digital anthropomorphic phantoms. This provides the possibility to both evaluate multiple imaging conditions and to have the ground truth available to quantitatively characterize the filtered results.

4D Similarity Filter Description

The 4DSF is a statistical method that reduces noise in temporal CT data, such as these dCTA image sequences. The method is based on averaging the values of voxels with similar TACs, as determined by the root-mean-square error (RMSE) between the TACs. Algorithm 1 shows the pseudocode of the filter implementation.

Algorithm 1: 4D Similarity Filter

Require: Initial temporal CT image data after applying a mask I^* , with N number of voxels within the mask and T number of phases

Require: Initial temporal CT image data I

Require: 4DSF filtered temporal CT image data within the mask I^{4DSF}

Require: Number of voxels used to set the new values of the voxel being filtered, called Filter Strength (FS)

Loop over all voxels to get the current voxel c , and loop over all phases to get current phase t

```

1:   for  $c \in [1, N]$  do
2:     for  $t \in [1, T]$  do
3:       # Loop over all voxels to get the candidate voxel  $m$ 
4:       for  $m \in [1, N]$  do
5:         # Loop over all phases except for the current phase  $t$  to calculate RMSE of each
6:         # candidate voxel  $m$  with the current voxel  $c$ 
7:          $e \leftarrow 0$ 
8:         for  $p \in [1, T], p \neq t$  do
9:            $e \leftarrow e + (I_{n,p}^* - I_{m,p}^*)^2$ 
10:        end for
11:        $error_m \leftarrow \sqrt{e / (T - 1)}$ 
12:     end for
13:     # Sort the indices of all candidates sorted based on the RMSE
14:     get indices  $\{ind_1, \dots, ind_N\} | \{error_{ind_1} \leq error_{ind_2} \leq \dots \leq error_{ind_N}\}$ 
15:     # Average the values at phase  $t$  of the  $FS$  number of voxels with the lowest RMSE
16:      $I_{c,t}^{4DSF} \leftarrow mean(\{I_{ind_1,t}, \dots, I_{ind_{FS},t}\})$ 
17:   end for
18: end for
19: return  $I^{4DSF}$ 

```

There are two aspects that make the 4DSF unique. First, to find the voxels with similar TAC, the 4DSF searches the entire image for voxels similar to the one being filtered rather than restricting its search to a certain number of spatially neighboring voxels. This is because when filtering small lesions or arteries, there will only be few voxels in the neighborhood with similar TACs but there might be other lesions and arteries elsewhere in the image with similar TACs. Second, the 4DSF decouples the candidate voxel identification from the signal averaging, since, especially in high noise level situations, a single measurement is not a good representation of the underlying ground truth. To achieve this decoupling, each phase of the TAC

will be filtered separately, so the phase being filtered can be excluded from the RMSE calculation. This is because the 4DSF compares a very large number, typically thousands to tens of thousands, of voxels from the entire image, which by the very nature of the image, usually include a high level of noise. Of all these voxels, only the Filter Strength (FS) number of candidate voxels that have the most similar TACs, based on the RMSE, will be averaged. However, this FS number of voxels will not only have the same underlying ground truth TAC, but also, since the noise is so dominant, they will have a similar realization of the noise. As a result, the average value of this distribution of voxels will tend to be biased towards the noise value present in the TAC of the voxel at hand, as opposed to it being an unbiased estimate of the ground truth. By filtering each phase of the TAC separately and excluding the phase being filtered when calculating the RMSE, the TACs of the selected voxels will not necessarily have similar noise values at this excluded time point, but rather the values will be randomly distributed around the ground truth value. Thus, their average will be an unbiased estimate of the ground truth. It is important to note that this principle assumes that the noise across phases is uncorrelated, i.e., that the 4D image contains temporally uncorrelated noise. If this were not the case, the phases before and after the excluded phase would still hold information about the noise in the excluded phase. This would mean that the FS number of voxels selected would still have noise similar to that of the excluded phase.

In its theoretical form and its simplest implementation, the 4DSF has only one input parameter, the FS, which defines the number of voxels that should be averaged together to set the new value of the voxel and phase being filtered. However, evaluating all voxels to identify the ones similar to the voxel currently being filtered is computationally very demanding, therefore the real-world implementation of the 4DSF uses additional parameters and an additional input solely for reducing computation time. The additional input is a binary mask, indicating which voxels should be filtered. This excludes processing of, for example, the air and the bones. Figure 1 shows a graphical overview of this implementation, with an overview of additional filter parameters.

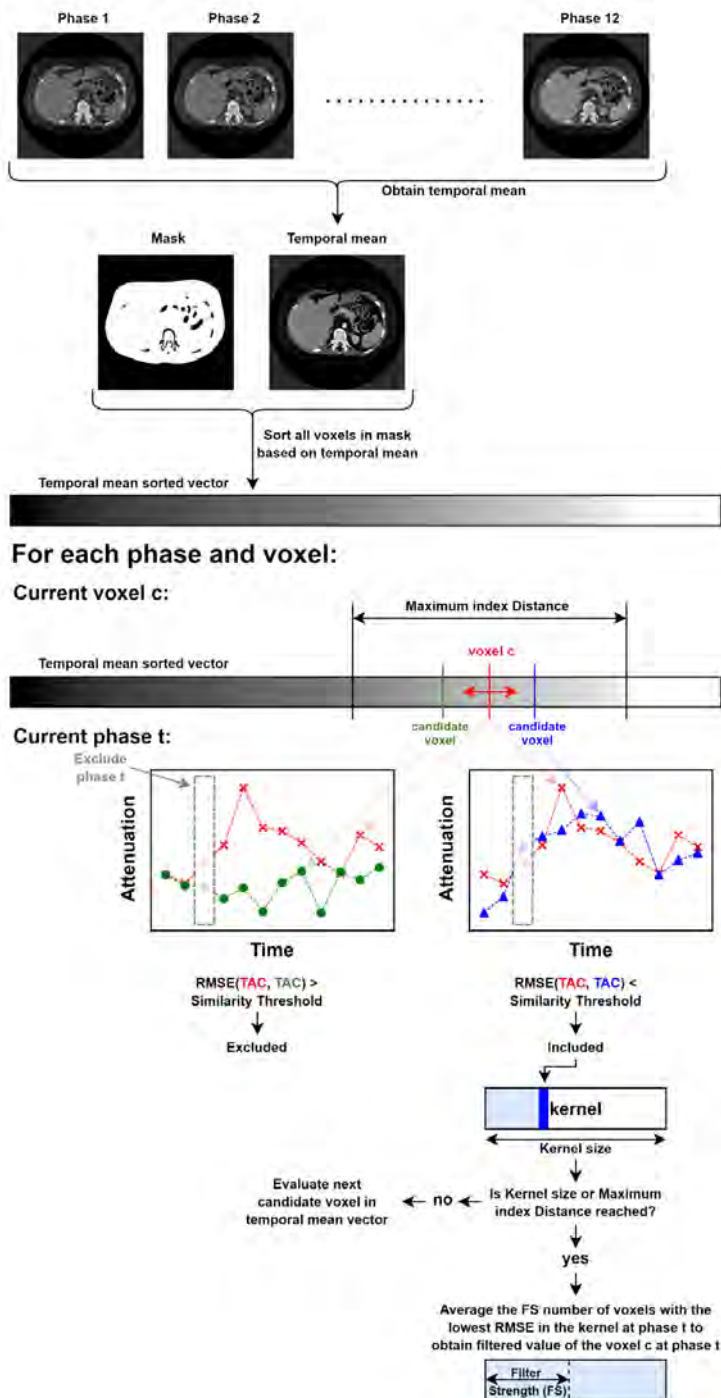


Figure 1: Graphical overview of the 4DSF, including the modifications to make it computationally feasible.

This implementation first sorts all voxels within the mask according to their temporal mean. For each voxel in the sorted vector, the principle described in Algorithm 1 is applied. However, when searching for similar voxels, only the voxels closest to the current one, in the temporal-mean sense, are considered, since voxel pairs with increasing difference in temporal mean values will be increasingly less likely to have a small RMSE between them. For computational feasibility, a maximum number of voxels is analyzed, set defined by one of the additional filter parameters, called the Maximum index Distance (MD). All voxels within the MD are checked for similarity, starting with the closest one, by determining if the RMSE between the corresponding TACs is lower than a set threshold, called the Similarity Threshold (ST). If a certain number of voxels, called Kernel Size (KS), fulfills the ST criterion before all voxels within the MD are checked, the remaining voxels are skipped. Finally, the FS number of voxels with the lowest RMSE are used to average and determine the new value of the voxel at the phase being filtered.

In case of very high noise levels in the initial images of the dCTA sequence, a pre-filter, e.g., an average filter, is used to reduce the noise in the initial images before utilizing the 4DSF. This is to reduce the possibility of including incorrect voxels (i.e., voxels of other tissues than the one being filtered) when averaging the FS best voxels. However, the pre-filtered images are only used to determine the FS best voxels, not as part of the processing of the final filtered image. Once the FS best voxels are determined, the actual averaging of the voxels in the FS set is performed with the voxel values from the initial images.

Methods

To extensively evaluate the 4DSF, anthropomorphic digital phantoms of different sizes were used to simulate hepatic dCTA image sequences, which were then processed by the 4DSF with different parameter settings. The images filtered by the 4DSF were evaluated for both image quality and quantitative accuracy to characterize the performance of the 4DSF, including the computational cost, under varying conditions and with different input parameters.

Digital Phantoms

A population of 52 4D anthropomorphic digital XCAT phantoms (Duke University, Durham, NC, USA) [42] was used for this work. See Table 1 for the details of the phantoms. The abdominal section of the phantoms was voxelized with an isotropic voxel size of 0.25 mm, with no motion included. One sphere-like liver lesion with a

10 mm diameter was generated and inserted in each of the XCAT phantoms using the work of Sauer et al. [108]. The diameter was chosen to be relatively small to test the limits of the 4DSF. The parameter settings for the lesion generation are listed in the online supplements Appendix A. The lesion represents a liver metastasis, so it is set to consist of two regions: a hyper-vascular outer layer and a necrotic inner core. The inner core is simulated via binary eroding of the initial lesion, with a 3×3×3 kernel, until the eroded part is 1/6 or less of the initial lesion volume. The lesion was added to each of the 52 phantoms, by placing it randomly within the liver.

Table 1: The details of the XCAT phantoms used in this study. For each parameter the median, minimum, and maximum are given, except for gender.

Gender (Female/Male)	Weight (kg)	Height (cm)	BMI (kg/m ²)
22/30	78.7 (52.0 – 120.0)	173 (153 - 190)	27.0 (18.2 - 38.8)

The simulated iodine concentrations over time in each tissue were determined for each XCAT phantom using the work of Sahbaee et al. [109,110], which models all organs and vessels as different compartments to determine the iodine concentration over time by volume, including the corresponding in- and outflow. The protocol simulated for this work was 100 ml of 300 mg I/ml iodinated contrast at a constant injection rate of 3.33 mL/s. For the liver lesion, which is not included in Sahbaee et al.'s work, the iodine concentration of the outer layer was set to 50% of that of the hepatic artery with a delay of 5 seconds, since, in general, liver metastases are arterially perfused [111–113]. No contrast perfusion was added to the inner core since it represents a central necrosis.

Image Simulation

A dynamic image sequence, consisting of 12 phases at 10 second intervals (i.e., total scan duration of 110 s) covering the time from pre-contrast until the delayed phase, was simulated for each of the 52 phantoms using a previously-developed and -validated CT simulator [114]. Volume scans were simulated as acquired with a tube voltage of 120 kV. The tube current-time product was set such that its total over all phases was, on average, the same as that for the clinical 4-phase protocol used at our institution, and scaled to patient size, based on water equivalent diameter [115]. On average, this resulted in a tube current-time product of 31.2 mAs and a CTDI_{vol} of 1.6 mGy per phase. No angular tube current modulation was simulated, since in this study we are not investigating, comparing, or optimizing the dose used during image acquisition.

Before reconstruction, beam-hardening correction [116] was applied to the projections. To reconstruct these images, the FDK algorithm [68] was used, with a field of view of 400 mm, a slice thickness of 0.5 mm, an image size of 512 by 512 pixels, and 320 slices.

Evaluation Conditions

For the initial evaluation of the performance of the 4DSF, all 52 phantoms were filtered, using the implementation in Figure 1, using the settings listed in Table 2. The FS was set to 100, to have a potential noise reduction factor of 10, and 100 voxels have a volume of approximately 30 mm³, which is similar to the volume of a solid sphere of 4 mm in diameter. The KS and MD were set to 30,000 and 300,000, respectively, to have enough voxels to choose from. Theoretically, the difference between two TACs with the same underlying ground truth should be $\sqrt{2} \cdot \sigma_{\text{avg}}$. Note that by setting an ST of 1,000, this parameter is in essence unused for this initial analysis.

Table 2: 4DSF parameter values for validation.

Parameter	Value
Filter Strength, FS (voxels)	100
Similarity Threshold, ST (HU)	1,000
Kernel Size, KS (voxels)	30,000
Maximum index Distance, MD (voxels)	300,000

The mask was generated by using an average filter with a 3×3×3 kernel on the first, pre-contrast, image, after which thresholding was applied to set all voxels with values between -300 HU and 300 HU to unity and all others to zero. The pre-filter used for each of the 12 images of the dCTA sequence was also an average filter with a 3×3×3 kernel.

The results were also compared to those obtained when applying the TIPS filter to the same 52 dCTA sequences. For a description of the TIPS filter and its equations, see online supplements Appendix B. This filter was chosen for comparison because it is most similar to the 4DSF, and also uses the TACs to find similar voxels. However, the TIPS filter limits the search for similar TACs to the spatial neighborhood of the voxel being filtered, while the 4DSF does this based on the temporal mean.

Since the TIPS filter works differently compared to the 4DSF, the two can also be combined. In this analysis the combination of applying the TIPS filter after 4DSF was investigated. This order seems most logical since the TIPS searches in a spatial

neighborhood, i.e., spatial gradients, which should be easier after initial noise reduction with 4DSF. The TIPS filter parameters used can be found in Appendix B in the online supplement, while the 4DSF parameters used are listed in Table 2. The TIPS filter parameters are different when the TIPS is used in combination with the 4DSF compared to when the TIPS is used on its own. Most importantly the kernel size of the TIPS filter is smaller when it is used in combination with the 4DSF, so that less smoothing is performed.

The influence of different 4DSF filter parameter settings, image acquisition settings, and lesion characteristics were also analyzed (Table 3). This was not done for the entire set of XCAT phantoms, but for the XCAT phantom of average size, in terms of water equivalent diameter (301 mm). From here on this XCAT phantom will be denoted as average XCAT. All these results will be presented in the online supplements Appendix C, except for the influence of the image acquisition settings.

Table 3: The different parameter settings, image acquisition settings, and lesion characteristics that were tested using the average XCAT.

Parameter	Values
Filter Strength, FS (voxels)	50, 100, 200, 500, 1000
Kernel Size, KS (voxels)	100, 300, 1000, 3000, 10000, 30000
Similarity Threshold, ST (HU)	25, 50, 100, 250, 500, 1000
Exposure scaling (-)	0.5, 0.75, 1.0, 1.25, 1.5
Lesion delay w.r.t. hepatic artery (s)	2, 5, 7, 10, 15
Lesion attenuation w.r.t. hepatic artery (-)	0.25, 0.375, 0.5, 0.625, 0.75
Lesion diameter (mm)	5, 10, 15, 20

The average XCAT was also used to test the 4DSF without a pre-filter and with a median pre-filter with a 3×3×3 kernel in addition to with the default average pre-filter with a 3×3×3 kernel. These results can be found in the online supplements Appendix C as well.

Analysis Metrics

The performance of the 4DSF was evaluated by determining the resulting noise reduction, lesion CNR, impact on CT number accuracy in time and attenuation, and spatial resolution loss. The noise reduction is defined as the ratio between the standard deviations (σ) in the liver parenchyma before and after use of the filter:

$$\text{Noise Reduction} = \frac{1}{T} \sum_{t=1}^T \frac{\sigma(ROI_{liver,init}(t))}{\sigma(ROI_{liver,4DSF}(t))} \quad (1)$$

where T is the total number of phases in the dCTA sequence t , is the phase number, $ROI_{liver,init}(t)$ is the region of interest (consisting of all liver parenchyma voxels) of the initial image at phase t , $ROI_{liver,4DSF}(t)$ is the region of interest after 4DSF of the t image, and $\sigma(^*)$ is the standard deviation of the region. Please note that all these tissue ROIs are binary eroded, with a $3 \times 3 \times 3$ kernel, to ensure that the influence of tissue transitions at the borders is minimized.

For each patient the contrast-to-noise ratio (CNR) of the lesion was determined before and after the application of the 4DSF. The CNR is described in Equation 2:

$$CNR = \frac{\mu(ROI_{lesion}) - \mu(ROI_{liver})}{\sqrt{\frac{1}{2}(\sigma(ROI_{lesion})^2 + \sigma(ROI_{liver})^2)}} \quad (2)$$

where ROI_{lesion} is again only the enhancing outer layer of the lesion. The CNR at all phases is determined and the maximum one is used as the CNR of the case.

The CT number accuracy was evaluated to investigate if the 4DSF introduces any bias to the attenuation or timing of the TAC by comparing the height and temporal location of the TAC peak before and after filtering. This evaluation was performed separately for the liver parenchyma, lesion outer layer, portal vein, and hepatic artery. The bias in peak attenuation was determined as:

$$bias_{att} = \max_{t \in 0, \dots, T} (\mu(ROI_{x,4DSF}(t))) - \max_{t \in 0, \dots, T} (\mu(ROI_{x,init}(t))) \quad (3)$$

where $ROI_{x,4DSF}(t)$ is the region of interest of tissue x at phase t after the use of the 4DSF, $ROI_{x,init}(t)$ is the region of interest of tissue x at phase t before the use of the 4DSF and $\mu(^*)$ is the average attenuation of the region.

The bias in timing of the TAC peak was determined as:

$$bias_{time} = \operatorname{argmax}_{t \in 1, \dots, T} (\mu(ROI_{x,4DSF}(t))) - \operatorname{argmax}_{t \in 1, \dots, T} (\mu(ROI_{x,init}(t))) \quad (4)$$

The spatial resolution loss was evaluated by comparing the full-width-at-half-maximum (FWHM) of a small hepatic artery in the average XCAT phantom. The FWHM was obtained using vertical and horizontal line profiles through the maximum of the small hepatic artery, for 10 consecutive slices. These line profiles were averaged, and a Gaussian fit was applied. The FWHM was determined using this Gaussian fit. Also, line profiles of the lesion in the average phantom before and after filtering were taken for visual comparison.

Finally, the noise texture before and after filtering were compared by means of the normalized noise power spectrum (nNPS). The 2D NPS was determined in a 55×55 voxel ROI for 40 consecutive slices and averaged, followed by rotational averaging, and normalized to unit area under the curve.

Results

Figure 2 shows one slice, and a zoomed in part, of the ground truth (no noise), initial, the TIPS filtered, and the 4DSF filtered image at phase six out of the twelve. Figure 3 shows the zoomed in subtraction images of the initial minus the 4DSF and TIPS filtered images. The results show that 4DSF visually reduces most of the noise and introduces little to no blurring, even though the initial image has high noise content. The TIPS filter also reduces most of the noise but introduces visually more blurring. The 4DSF + TIPS seems to be the cleanest image from visual assessment. The results also show that none of the filters could restore all the very small peripheral arteries in the liver.

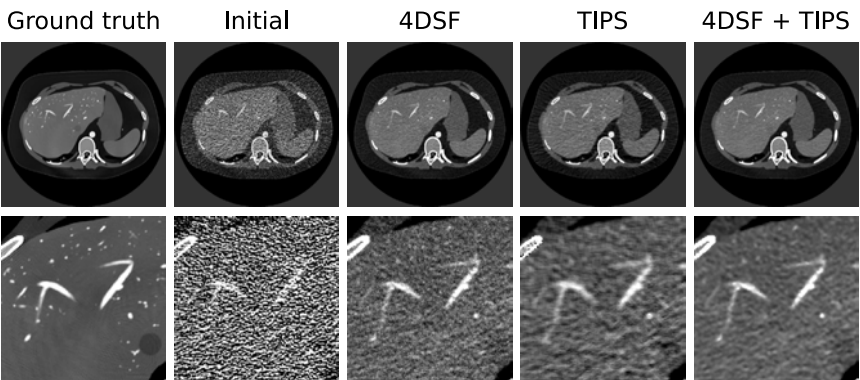


Figure 2: Ground truth, initial, and filtered results. For the 4DSF, the parameter values listed in Table 2 were used. Top row WL/WW: 150/500, bottom row WL/WW: 150/300.

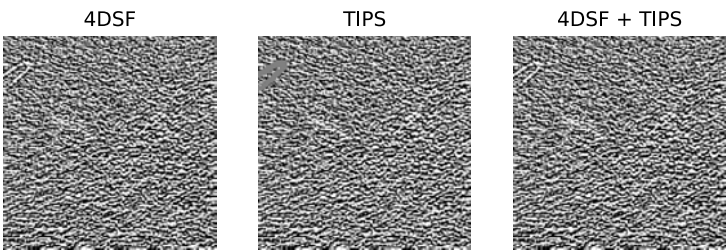


Figure 3: Subtraction images of initial image minus filtered image of Figure 2. WL/WW: 0/500.

Noise reduction & Lesion CNR

Table 4 shows the median and interquartile range (IQR) over all 52 XCAT phantoms of the standard deviation of all liver voxels before (initial) and after filtering, and of the resulting noise reduction, as defined in Equation 1. The TIPS filter, 4DSF, and 4DSF + TIPS all show excellent noise reduction capabilities, with the TIPS filter outperforming the 4DSF. In Table 5, the lesion CNR before and after filtering is shown, the TIPS filter also outperforms the 4DSF for this metric. For both metrics 4DSF + TIPS gives the best results.

Table 4: The median and IQR of the σ of the liver parenchyma determined for all 52 XCAT phantoms in the dataset.

	Median	IQR
$\sigma_{initial}$ (HU)	196.6	184.4 – 213.6
σ_{TIPS} (HU)	24.8	22.6 – 26.8
σ_{4DSF} (HU)	29.4	28.0 – 31.1
$\sigma_{4DSF+TIPS}$ (HU)	16.2	15.3 – 17.5
Noise Reduction TIPS (-)	8.1	7.8 – 8.3
Noise Reduction 4DSF (-)	6.8	6.5 – 7.0
Noise Reduction 4DSF+TIPS (-)	12.2	11.4 – 13.2

Table 5: The median and IQR of the lesion CNR determined for all 52 XCAT phantoms in the dataset.

	Median	IQR
$CNR_{initial}$ (-)	0.44	0.40 – 0.50
CNR_{TIPS} (-)	2.59	2.25 – 3.00
CNR_{4DSF} (-)	1.85	1.54 – 2.07
$CNR_{4DSF+TIPS}$ (-)	3.12	2.46 – 3.56

Time Attenuation Curve Accuracy

The CT number accuracy of the dCTA sequence after TIPS, 4DSF, and 4DSF + TIPS is shown in the histograms for the 52 XCAT phantoms in Figures 4 and 5, and summarized in Table 6. The figures show the bias in time (Figure 4) and attenuation (Figure 5). The time bias histogram of the dCTA sequences filtered by 4DSF, and by 4DSF + TIPS, indicates that in a few cases the lesion tends to have a positive time bias (delay) and in most cases the portal vein tends to have a negative time bias. In other words, the 4DSF shrinks the temporal separation between the different enhancement peaks. As can be seen in Figure 4 (left), the TIPS filter introduces very little time bias. The attenuation bias histograms indicate that for the tissues

that have a high peak, like arteries and lesions, the TIPS tends to introduce a negative attenuation bias. As opposed to the time bias, the 4DSF filter introduces a considerably smaller attenuation bias compared to the TIPS for the small and highly perfusing tissues such as the lesion and hepatic artery. The 4DSF + TIPS shows a slightly higher bias compared to only the 4DSF for the small and high peak tissues such as lesion and hepatic artery. The TACs of the average XCAT before and after the use of 4DSF are plotted in Figure 6 to show the loss in attenuation for the artery and lesion, and the bias in peak time, especially for the portal vein.

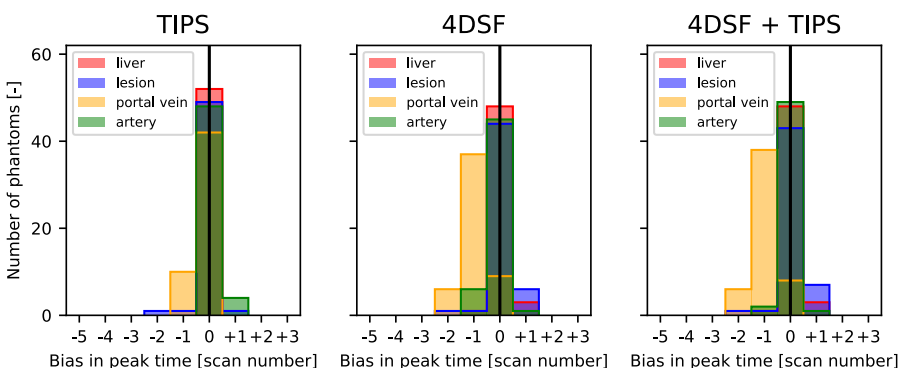


Figure 4: Histograms of CT number accuracy of all XCAT phantoms in terms of bias introduced in the TAC peak time for the dCTA sequence processed by TIPS filter (left), processed by 4DSF (middle), and processed by 4DSF + TIPS (right). Negative bias in time means that the peak is shifted forward in time after filtering.

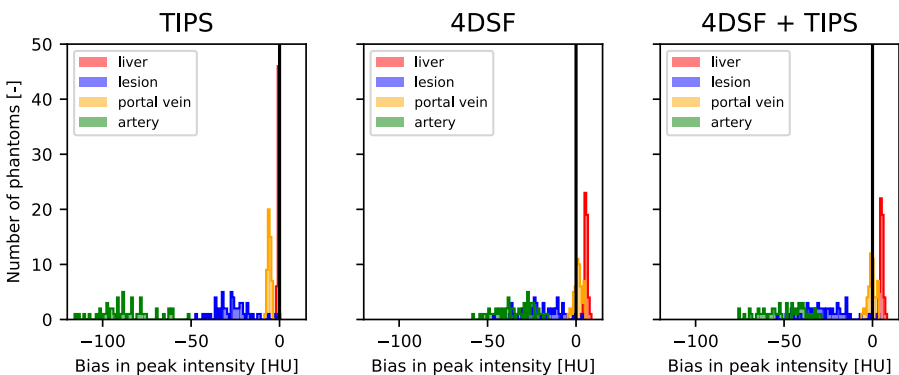


Figure 5: Histograms of CT number accuracy of all XCAT phantoms in terms of bias introduced in the peak attenuation of the TAC for the dCTA sequence processed by TIPS filter (left), processed by 4DSF (middle), and processed by 4DSF + TIPS (right). Negative bias in attenuation means that the peak is decreased after filtering and positive attenuation bias means an increase in the height of the peak after filtering.

Table 6: The median and IQR of the time and attenuation bias determined for all 52 XCAT phantoms in the dataset for the results after both the TIPS filter and 4DSF.

		Liver	Artery	Portal Vein	Lesion
Time bias (phase)	TIPS	0 (0 – 0)	0 (0 – 0)	0 (0 – 0)	0 (0 – 0)
	4DSF	0 (0 – 0)	0 (0 – 0)	-1 (-1 – -1)	0 (0 – 0)
	4DSF+TIPS	0 (0 – 0)	0 (0 – 0)	-1 (-1 – -1)	0 (0 – 0)
Attenuation bias (HU)	TIPS	-1 (-1 – -1)	-88 (-95 – -78)	-6 (-6 – -5)	-26 (-32 – -21)
	4DSF	+6 (+5 – +6)	-34 (-43 – -26)	1 (0 – 3)	-22 (-28 – -12)
	4DSF+TIPS	+5 (+5 – +6)	-50 (-61 – -41)	0 (-2 – 1)	-28 (-34 – -19)

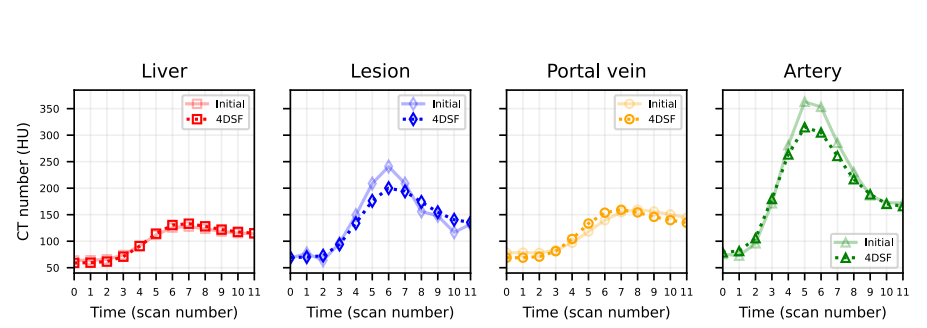


Figure 6: TACs of the average XCAT from the liver, artery, portal vein, and lesion. The transparent plot is the initial TAC and the dotted plot is the TAC after 4DSF.

Spatial Resolution

Figure 7 shows a partial CT scan of the average XCAT phantom, the yellow circle indicates the small hepatic artery used to determine the FWHM. Visually it is already clear that the resolution is different for all three filter techniques. The FWHM values for the ground truth and the three different filters are listed in Table 7. The values show almost no increase in FWHM for the techniques including 4DSF, whereas the TIPS filter did increase the FWHM by more than 35%, indicating more resolution loss.

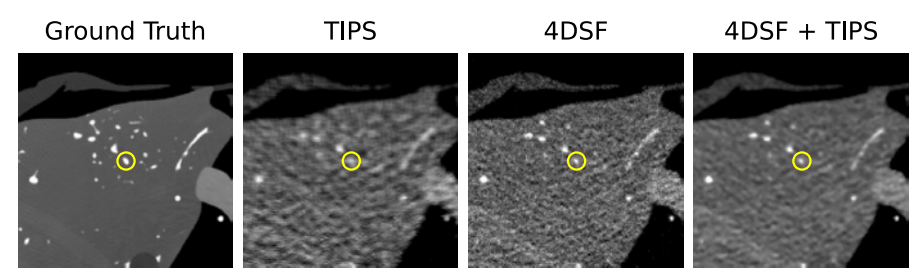


Figure 7: Zoomed in part of the liver and some of its small hepatic arteries of the sixth phase. The yellow ROI indicates of which hepatic artery the FWHM is determined. WL/WW: 150/300.

Table 7: Full Width Half Maximum values of three different filter techniques and the ground truth value for comparison.

Filter	FWHM (pixels)
Ground truth	3.1
TIPS filter	4.3
4DSF	3.2
4DSF + TIPS filter	3.4

Figure 8 shows the peak arterial phase, which is the seventh out of the twelve phases of the dCTA sequence initial, TIPS, 4DSF, and 4DSF + TIPS filtered image of the average XCAT phantom. The images show that the 4DSF-filtered image appears much sharper and has little to no blurring across edges compared to the others. The bottom row of Figure 8 shows zoomed-in portions of the top row images, highlighting the lesion. Line profiles of the voxel values between the red dots of Figure 8 are plotted in Figure 9, please note that the line profile is three pixels thick, i.e., it is an average of three line profiles next to each other, to suppress the noise for better visualization. These show that the 4DSF introduces less resolution loss compared to that resulting from applying the TIPS and TIPS post-filter, since it is the one of the three filter techniques that comes close to restoring the necrotic inner layer of the lesion. In Figure 8 it can also be seen that the noise textures after TIPS filter and TIPS post-filtering is different from the one after 4DSF only, with the noise in the latter containing more higher frequencies. The rotationally-averaged normalized noise power spectra (nNPS), obtained within the yellow ROI of Figure 8, shown in Figure 10, confirm this.

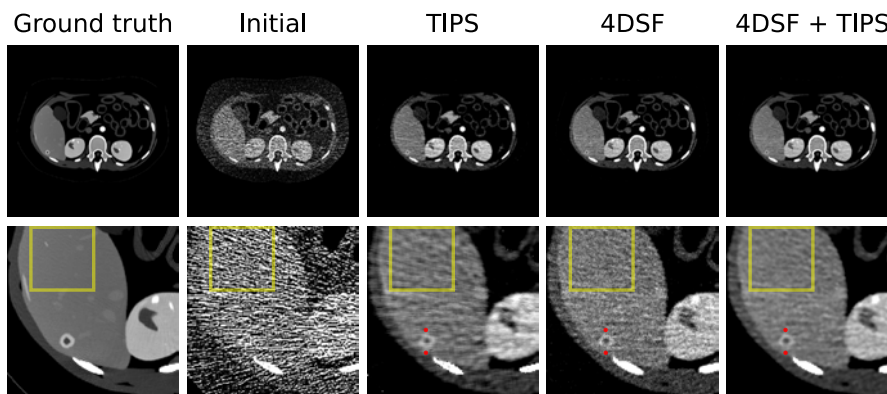


Figure 8: Results of the different filters on the average XCAT, with a lesion in the posterior of the liver. All filters result in different degrees of blurring and different noise textures, with the TIPS filter seeming to introduce the most blurring and the 4DSF the least. The red dots indicate the start and end point of the line profiles of Figure 9. The yellow ROIs indicate where the rotationally-averaged nNPS of Figure 10 were computed. WL/WW: 150/300.

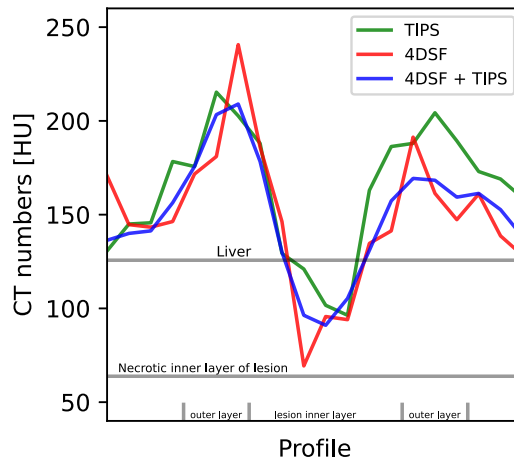


Figure 9: Line profiles of the lesion highlighted in Figure 8, showing that the 4DSF is the only filter that fully restores the voxel values of the necrotic inner layer of the lesion because it does not suffer from major resolution loss. The horizontal gray lines indicate the CT number of the liver and of the necrotic inner layer of the lesion. The gray dashes on the x-axis indicate the extent of the enhancing outer layer of the lesion and of the necrotic inner layer.

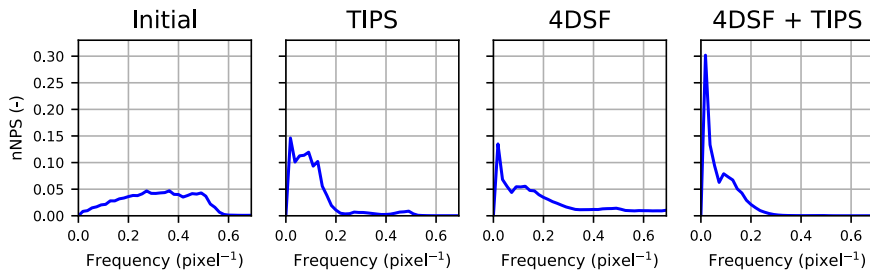


Figure 10: The nNPS (normalized to unit area under the curve) of the noise obtained from the different images within the yellow ROI of Figure 8.

Influence of image acquisition settings

Figure 11 shows the influence of different tube current levels, relative to the original tube current, called *exposure scaling* in Figure 11. The initial image noise has a minor influence on noise reduction and CNR. With decreasing initial image noise, e.g., due to increasing exposure, the resulting noise reduction decreases slightly and therefore the improvement ratio in lesion CNR is minor. Slightly more time bias seems to be introduced with increasing initial image noise. The attenuation bias shows a clear increasing trend for the arteries and a slightly increasing trend for the lesions in case of increased initial image noise. As expected, increased initial

image noise makes restoring the image more difficult, which is manifested by the introduction of larger quantitative inaccuracy in the TACs. Figure 12 shows the liver and lesion TACs for three different exposure levels, showing the decreasing difference between the lesion and liver TACs with lower exposure levels, i.e., with increasing initial noise.

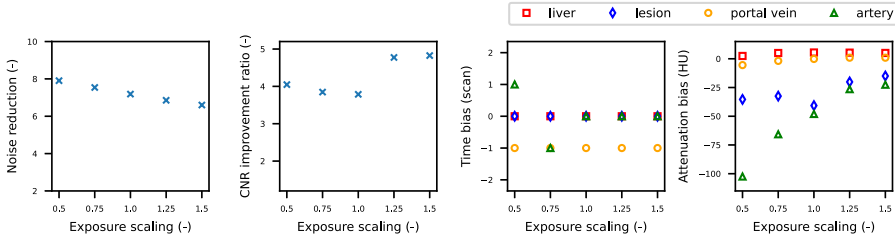


Figure 11: Influence of the initial image noise, in this case set by varying the exposure, on the four image quality metrics after filtering.

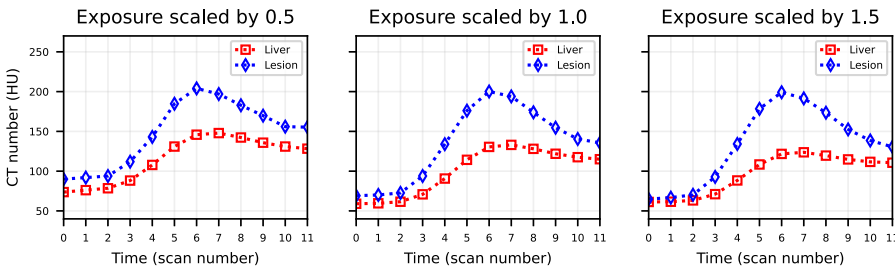


Figure 12: TACs at three different exposure levels, 0.5x the original exposure level(left), the original exposure level (middle), and 1.5x the original exposure level (right). It shows that the resulting difference between lesion and liver after filtering is decreased in images acquired with lower exposure level.

Computation time

The 4DSF was implemented in C++ and runs on the CPU only. For this study it was run on a Linux system with 128 GB RAM, Intel(R) Xeon(R) Silver 4210R CPU @ 2.40GHz. The process was multi-threaded over 32 CPU threads each on a separate core. The computation time for the standard parameter settings as listed in Table 2 was 7103 seconds. The computation time was measured for different KS and ST values and the results can be found in the online supplements Appendix C. This analysis showed that the computation time increases approximately linearly with increasing KS. So, if the KS would be set to 1,000 instead of 30,000, the computation time would be 255 seconds.

Discussion

This work presents an extensive analysis of the 4DSF on dCTA sequences of digital anthropomorphic phantoms. The performance of this filter was compared to that of the TIPS filter, the influence of different filter parameters, scanner settings, and lesion characteristics on the quality of the output image have been analyzed. By using digital anthropomorphic phantoms, it was possible to quantify the performance of the filter in terms of accuracy and introduction of bias, since the underlying truth is known.

It can be concluded that the 4DSF reduces noise and improves lesion CNR greatly, but the TIPS filter is superior for these metrics. However, the 4DSF does this with minimal spatial resolution loss and results in noise that is of higher frequency in nature. These two results are due to the filter not using information based on the spatial neighborhood of the voxel being processed. This is advantageous because if the residual noise is higher frequency in nature, this results in better visibility of low contrast abnormalities or changes [117,118]. Also, the height of the TAC peak of small highly-perfused tissues, such as hepatic artery and the lesion, were restored better by the 4DSF compared to the TIPS filter. The combination of 4DSF + TIPS filter showed the best results for noise reduction and lesion CNR. In addition, the spatial resolution loss for 4DSF + TIPS is minor compared to that resulting from the 4DSF. Finally, the height of the TAC peaks for 4DSF + TIPS were not restored as well as when applying the 4DSF alone, but better compared to TIPS filter.

The investigation on the impact of the parameters of the 4DSF showed that the FS can be increased to reduce the noise and increase the lesion CNR with only minor loss in accuracy of the TAC peak of the highly-perfused tissues, i.e., the arteries and lesions. The KS can be reduced to 10 times the FS before it affects the accuracy of the TAC peak, in case of an FS of 100. This means that for the conditions investigated in terms of image size ($512 \times 512 \times 320$ of which approximately 35% was inside the binary mask) and number of phases (12), the filter's running time is approximately 255 seconds on our system when $FS=100$.

The exposure scaling showed that the TAC peak of the small highly-perfused tissues suffered the most from decreasing exposure. Also, an increase in exposure by a factor of 1.25 seemed to have a very positive influence on the accuracy of the TAC peak. This result suggests that, for instance, reducing the number of phases acquired to nine- or ten, so that the exposure can be increased by a factor of 1.25,

resulting in the same total dose, might be a good option for increased quantitative accuracy. However, more research is needed to confirm this.

The analysis of the impact of lesion characteristics showed no clear trend in the metrics for different lesion delays. The lesion attenuation had a positive influence on the CNR of the lesion and a slightly negative influence on the attenuation bias. Finally, as expected, the larger lesions suffer less from a decrease in the height of the TAC and their CNR is easier to restore.

The results also showed that using a pre-filter, be it average or median, gives much better results for all metrics compared to no pre-filtering. The average pre-filter showed superiority compared to the median pre-filter for CNR improvement and attenuation bias with slightly lower noise reduction. Therefore, the average pre-filter is the preferred choice, based on this study.

Several limitations in this study should be kept in mind. The analysis was performed solely on digital anthropomorphic phantoms, which only represent human bodies and the iodine perfusion behavior to a certain extent. Also, the analysis was performed assuming no patient motion or perfect registration. The FDK reconstruction algorithm was used, resulting in relatively high high-frequency noise, when utilizing the filter in situations where a more sophisticated reconstruction method is used, the results might differ. We still used FDK instead of a more sophisticated reconstruction method, such as model-based reconstruction or deep learning reconstruction, because these were not available to us. Since the 4DSF is expected to work best for reconstruction methods with low spatial correlation and FDK reconstruction with a sharp kernel probably comes closest to this assumption from all commonly-available reconstructions, the presented analysis gives good insight into the capabilities of the 4DSF. In addition, the pre-filters used here do introduce spatially-correlated noise, indicating that the 4DSF works also in cases where there is increased spatial correlation of the noise, at least to a certain degree.

Because of these limitations, the performance of the 4DSF found here should be considered a ceiling, rather than what can be expected unless a highly accurate registration algorithm is applied. However, the results are still a good indication of the strong and weak points of the filter and the influence of certain variables on the resulting output.

Overall, the filter is able to achieve substantial noise reduction and CNR increase, but at the cost of reduced TAC accuracy. Therefore, the filter seems to be ideal for abnormality detection, especially in combination with the TIPS filter, i.e., 4DSF + TIPS, after applying an averaging pre-filter. However, one should be careful when performing quantitative evaluation of the filtered dCTA sequences, especially when looking at TACs of small objects, e.g., via flow or volume maps.

Conclusion

An extensive performance analysis of the 4DSF was performed on simulated 12-phase dCTA abdomen acquisition simulations. With the results of the analysis, the influence of filter parameters, scanner settings, and lesion types was determined. The strong points, being the noise reduction capability and increasing in lesion CNR, and weak points, being reducing the accuracy of the TAC peak of the highly perfusion tissues, of the filter are identified. Future use cases and protocols utilizing the 4DSF can be chosen or adjusted based on these insights, potentially leading to better or proper use of the 4DSF.

Supplementary Material

Appendix A – Lesion generation parameters

For lesion generation, a computer program provided by the CVIT was used [108]. Below the input parameters for the lesion generation computer program, as used in this work, are given.

% Base options

```
base.imgRes = 0.25;      %voxel size (mm)
base.seed = 1;          %random number generator seed
base.complexity = 1.0;   %complexity scaling, 1=normal
```

% Mass options

```
mass.lMax = 4;          %maximum spherical harmonic order [int]
mass.alpha = 5.0;       %mean mass radius (mm)
mass.meanSigma2 = 0.2;  %mean mass surface irregularity variance (mm^2)
mass.stdSigma2 = 0.04;  %mass surface irregularity standard deviation (mm^2)
mass.powerLaw = 4.0;    %covariance power law index
mass.meanLF = 500.0;    %mean number of low freq. modifications
mass.stdLF = 50.0;      %std. deviation of number of low freq. modifications
mass.meanShape = 0.8;   %mean LF shape distribution 0=spike, 1=bump
mass.stdShape = 0.1;    %std. deviation of LF shape 0=spike, 1=bump
mass.meanLFRad = 0.1;   %mean relative LF radius
mass.stdLFRad = 0.05;   %std. deviation of LF radius
mass.meanLFLen = 0.1;   %mean relative LF length
mass.stdLFLen = 0.01;   %std. deviation of LF length
mass.meanFuzzAlpha = 0.0; %mean fuzzy alpha
mass.stdFuzzAlpha = 0.0; %std. deviation of fuzzy alpha
```

% Spiculation options

```
spicule.meanInitial = 1000; %mean number of initial segments (set to zero)
spicule.stdInitial = 60; %std. deviation number of initial segments
spicule.meanNeigh = 8.98; %mean max number of neighbor segments
spicule.stdNeigh = 1.89; %std. deviation max number of neighbor segments
spicule.meanInitRad = 0.024; %mean initial relative radius
spicule.stdInitRad = 0.0053; %std. deviation initial relative radius
spicule.meanRadDec = 0.93; %mean radius decrease
spicule.stdRadDec = 0.05; %std. deviation radius decrease
spicule.meanInitLen = 0.025; %mean initial relative length
```

```

spicule.stdInitLen = 0.0025; %std. deviation initial relative length
spicule.meanLenDec = 0.95; %mean length decrease
spicule.stdLenDec = 0.05; %std. deviation length decrease
spicule.meanContProb = 0.717; %mean continue prob.
spicule.stdContProb = 0.057; %std. deviation continue prob.
%%% Assuming equi-probably symmetric and asymmetric branching
spicule.meanSymBifProb = 0.142; %mean symmetric bifurcation prob.
spicule.stdSymBifProb = 0.028; %std. deviation symm bif. prob.
spicule.meanAsymBifProb = 0.142; %mean asymmetric bifurcation prob.
spicule.stdAsymBifProb = 0.028; %std. deviation asymm bif. prob.
spicule.meanBranchAng = 6.55; %mean asymmetric bifurcation prob.
spicule.stdBranchAng = 0.62; %std. deviation asymm bif. prob.

save('tmp.mat','base','mass','spicule','-v7.3')

```

Appendix B – Tips filter

The TIPS filter is described below [37]:

$$c(\xi, a) = \exp\left(-\frac{1}{2}\left(\frac{d(\xi, a)}{\sigma_d}\right)^2\right) \quad (\text{B.1})$$

where c is a gaussian closeness function, d is the Euclidean distance in pixels, a is the voxel being filtered, ξ is a neighboring voxel, and σ_d is the parameter deciding which distance is considered close. The TIPS function p is defined as:

$$p(\xi, a) = \exp\left(-\frac{1}{2}\left(\frac{v(\xi, a)}{\sigma_v}\right)^2\right) \quad (\text{B.2})$$

$$v(\xi, a) = \frac{1}{T} \sum_{t=0}^{T-1} \left(I(\xi(x, y, z, t)) - I(a(x, y, z, t)) \right)^2 \quad (\text{B.3})$$

where v is the sum of squared differences (SSD) between the TAC, T is the number of phases, and σ_v is the parameter deciding if the SSD between the TAC is low enough.

$$I^{TIPS}(a(x, y, z, t)) = \frac{1}{n(a)} \sum_{i=-m}^m \sum_{j=-n}^n \sum_{k=-o}^o I(\xi(x+i, y+j, z+k, t)) c(\xi, a) p(\xi, a) \quad (\text{B.4})$$

$$n(a) = \sum_{i=-m}^m \sum_{j=-n}^n \sum_{k=-o}^o c(\xi, a) p(\xi, a) \quad (\text{B.5})$$

where, m , n , and o , are length of the kernel in each direction, thus half the kernel size, and I^{TIPS} is the TIPS filtered dCTA sequence. For this comparison, the parameters of the TIPS filter were set to kernel size = $5 \times 5 \times 5$, $\sigma_d = 3$ voxels, and $\sigma_v = (2 \cdot \sigma_{liver})^2$, where σ_{liver} is the standard deviation in the liver, based on the suggested parameter range and optimized via trial-and-error for liver dCTA [37]. For the TIPS post-filtering the parameters were set as followed kernel size = $3 \times 3 \times 3$, $\sigma_d = 1.5$ voxels, and $\sigma_v = (2 \cdot \sigma_{liver})^2$.

Appendix C – Additional Results

In this appendix additional results of the 4D similarity filter (4DSF) will be presented. Such as the influence of 4DSF parameters, lesion characteristics, pre- and post-filtering. Additionally, the results of a computational time analysis are given.

Influence of 4DSF Parameter

The influence of the different parameter settings for FS, KS, and TH are plotted in Figure C1. For this analysis, the improvement ratio in CNR compared to the initial dCTA sequence is given.

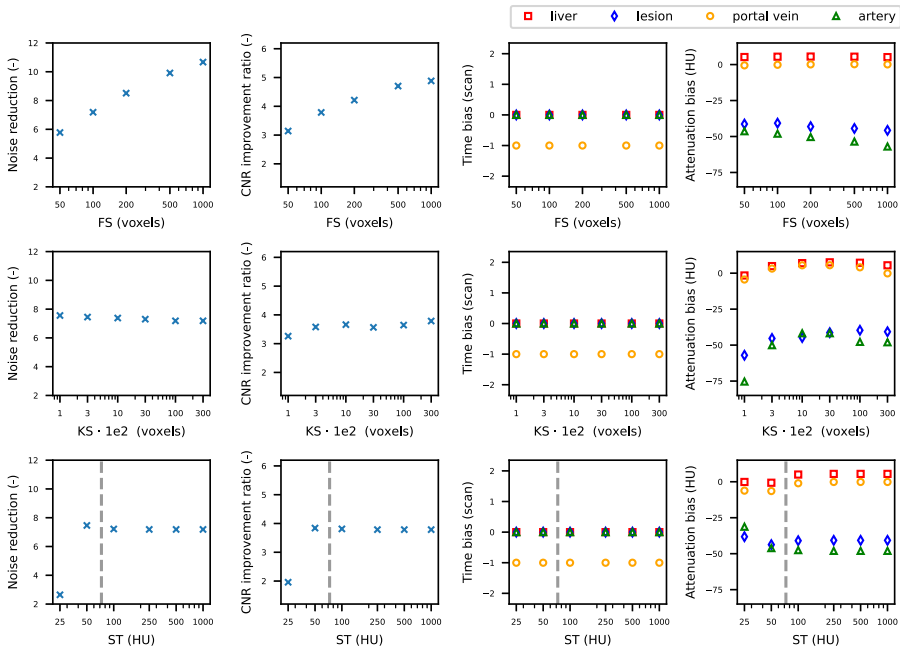


Figure C1: Influence of the 4DSF parameter settings (Filter Strength (FS), Kernel Size (KS), and Similarity Threshold (ST)) on the four image quality metrics after filtering. The vertical dashed line in the third row indicates the theoretical ST threshold.

The results show, as expected, an increase in noise reduction with increased FS, since this is the number of voxels averaged. Since the noise is reduced the lesion CNR also increases, however the magnitude of the attenuation bias in tissues with high TACs, i.e., arteries and lesions, worsen with increased FS. The KS parameter has limited to no influence on the performance metrics, as long as it is at least 10 times larger than FS, otherwise the attenuation biases in the lesions and arteries increase. Similarly, the ST parameter seems to have limited to no influence on the

performance metrics, as long as it is larger than the theoretical threshold $\sqrt{2} \cdot \sigma_{avg}$ (as explained in the manuscript), with σ_{avg} being the standard deviation of the liver after pre-filtering. The average XCAT has a σ_{avg} of 52 HU, so the theoretical ST is 73 HU, which is indicated by the dashed line in the third row of Figure C1. Below this threshold, the filter has a high chance of disregarding voxels of similar tissue types, leading to reduced filtering and thus reduced noise reduction and CNR improvement. However, reduced filtering also leads to reduced attenuation bias. In practice the number of voxels per tissue is so high that the performance only starts dropping when ST is much lower than the theoretical threshold.

Influence of lesion characteristics

Figure C2 shows the influence of different lesion characteristics on the 4DSF results. The delay in lesion enhancement compared to the liver artery has limited to no influence on the noise reduction and CNR improvement. The lesion attenuation, i.e., the iodine concentration, as expected, has a positive influence on the lesion CNR improvement, but has no influence on the magnitude of the noise reduction. The attenuation bias is slightly larger with higher lesion attenuation. The lesion size has no influence on noise reduction, as expected, and only the smallest lesion shows a decrease in CNR improvement. The attenuation bias is clearly reduced with increasing lesion size, which is logical since there will be more lesion voxels to average, so less chance of mixing with different tissues.

Influence of pre- and post-filtering

Figure C3 shows the influence of different pre-filters. The results clearly show that using no pre-filter leads to considerably worse results. Also, using an averaging pre-filter results in a lower attenuation bias, a higher lesion CNR improvement, and slightly lower noise reduction compared to using a median pre-filter. Therefore, the average pre-filter seems to be the preferred pre-filter in this situation.

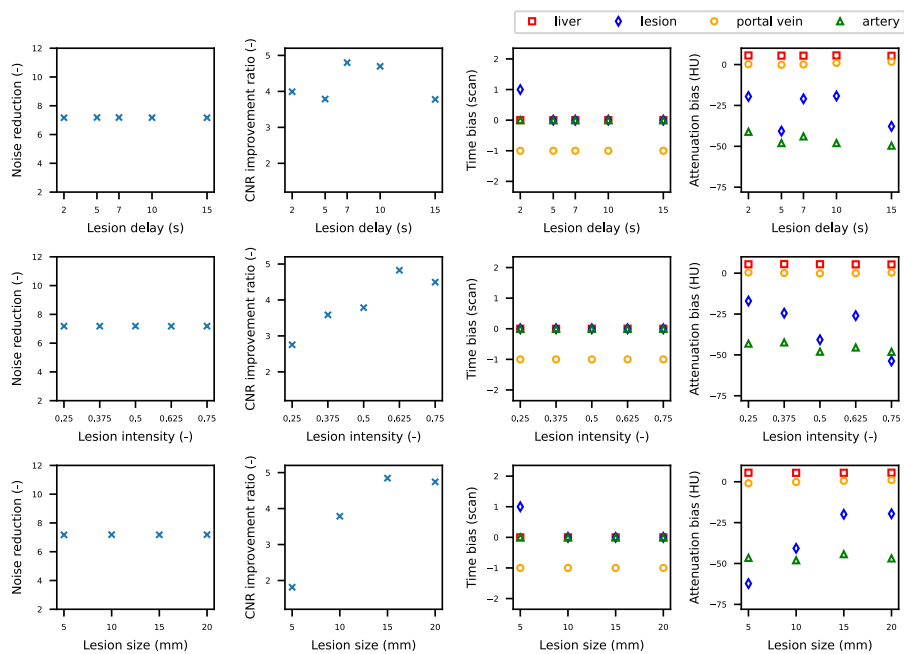


Figure C2: Influence of the lesion characteristics on the four image quality metrics after filtering.

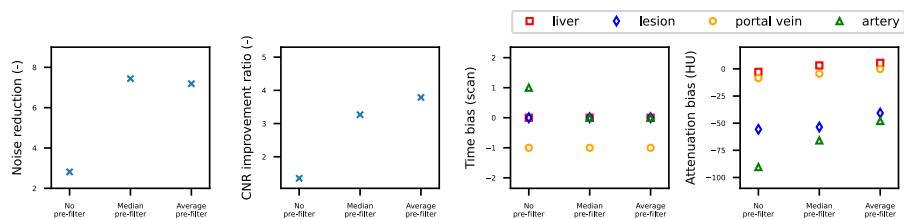


Figure C3: Plots showing all four metrics for no pre-filter, a median pre-filter, and an average pre-filter.

Computation time

The computation time was measured for different values of KS and ST and is listed in Table C1.

Table C1: Computation times of the 4DSF for different ST and KS values. The top row represents the use of the filter with the standard parameter settings as listed in Table 2, of the manuscript. The FS was 100 for all runs.

ST	KS	Time (s)	Fraction time relative to standard settings
1,000	30,000	7103	1.0
25	30,000	25069	3.53
50	30,000	14743	2.08
100	30,000	7302	1.03
250	30,000	6965	0.98
500	30,000	7143	1.01
1,000	100	38	0.005
1,000	300	104	0.015
1,000	1,000	255	0.036
1,000	3,000	649	0.091
1,000	10,000	2015	0.284

It can be seen that the time increases approximately linearly with increasing KS. Also, the value of ST does not influence the computation time as long as its value is above $\sqrt{2} \cdot \sigma_{avg}$. The TIPS filter processing time was not measured because we do not have a speed-optimized implementation of it, and therefore the comparison would not be fair.



Chapter 5

GPU-based implementation of the 4D Similarity Filter

Original title: Adaptation of a 4D noise filter for implementation on GPU

Mikhail Mikerov, Sjoerd A. M. Tunissen, Koen Michielsen, Ewoud J. Smit,
and Ioannis Sechopoulos

Submitted to: 8th International Conference on Image Formation in X-Ray Computed Tomography

ABSTRACT

Functional 4D CT imaging including contrast injection can be beneficial when evaluating pathology. However, acquisition of long image sequences implies the need to use a low dose per phase to limit total radiation dose. This calls for development of advanced processing methods that can handle the resulting high image noise efficiently. Noise filters specific for dynamic imaging can use temporal information to achieve higher noise reduction but have a high computational cost. In this work, we accelerated an existing clinical 4D noise filter by making changes to its algorithm that enabled it to run approximately 13 times faster on GPU compared to a highly optimized CPU implementation without any obvious reduction in its effectiveness.

Introduction

Despite significant progress in CT hardware and software that improved image quality and clinical decision making, standard CT provides only anatomical information of the patient condition. In many clinical situations it would be beneficial to also obtain functional information by studying the enhancement with iodinated contrast agents, especially when evaluating pathology for treatment planning or response.

Usually three- or four-phase functional imaging protocols are implemented in the clinic for patients with stroke, cardiac ischemia, or liver lesions. They can still be performed at relatively low doses but have the risk of mistimed acquisitions, resulting in missing important aspects of contrast dynamics. More densely sampled sequences would eliminate the risk of missing important points on perfusion curves.

Functional imaging with long sequences consisting of 10 to 30 phases depending on the protocol comes at the cost of an increased dose. To reduce this, each phase in the sequence needs to be acquired at considerably lower dose than a standard anatomical CT image. Noise reduction filters can then be applied as a postprocessing step to enable the extraction of useful information from otherwise too-noisy images.

Traditional denoising filters can be used in each individual phase, however they often use information from a small spatial neighborhood, leading to loss of resolution. To solve this, information from the other images in the sequence can also be used to decrease noise in the current phase. Therefore, noise filters specific for dynamic imaging that leverage the temporal information such as the 4D similarity filter have been proposed [55].

In this work, we aim to improve runtime of the 4D similarity filter by introducing a small modification to the algorithm that enables its implementation on GPU. We performed the first comparison with the exact CPU version and did not notice a performance drop.

Filter Implementation

4D Similarity Filter

Prokop and Smit introduced a 4D similarity filter to reduce noise in low dose 4D imaging [55]. The idea of this filter is to find, for each voxel v in the image at a given timepoint t , a set of voxels $\{V_c\}$ that have similar perfusion curves and use

those with the least differences to calculate an average voxel value that can then be used to update the image.

The search for similar voxels is performed by comparing the perfusion curves of the voxel of interest v (the one being updated) with a large pool of candidate voxels v_c in V_c from the entire image according to how close the temporal mean of each voxel v_c is to that of v . If the perfusion curve of a candidate voxel v_c is similar according to some threshold TH , it is saved in an array A . When the array is full, it is sorted based on the difference of perfusion curves and only the best FS candidate voxels are used to calculate the average value for this voxel. This filter can be implemented both on CPU and GPU, but the latter has not been achieved before. There exist multi-threaded CPU realizations of this filter, which are already used in clinical practice [96,97,100].

One of the drawbacks of the 4D similarity filter is long computational times since every voxel in the image needs to be processed independently. In a typical $320 \times 512 \times 512$ volume, with just 1/3 of all voxels being in the mask, the total number of voxels is around 27 million. On a 32-core CPU, each core would need to process more than 870 thousand voxels. The independence of each voxel invites use of the GPU for this task, since it has many more cores, which can process even more voxels in parallel. However, we encountered major problems when implementing the 4D similarity filter on the GPU.

Difficulties of Implementation on GPU

In general, low computational times on the GPU are achieved when it is possible to execute one thread per output element that can run independently from other threads and consists mainly of arithmetic operations with very limited memory accesses. This is due to the fact that arithmetic operations need up to 20 times fewer cycles to complete than operations on device (global) memory. If a high number of memory accesses are required to compute a single output element, several techniques can be employed. First, one wants to ensure that memory reads are coalesced, i.e., during the calculation of neighboring output elements, memory reads from the threads in the same block should be performed in accordance to memory layout. Second, if multiple threads in the same block can reuse some of the information, shared memory can be used to specify a user-defined and -maintained cache. Third, if memory reads are random but there is a high chance of having close spatial proximity, textures can be utilized due to their large cache, which is checked first before reading the data from the (slow) device memory.

Unfortunately, none of these optimization techniques are applicable in our problem since each candidate voxel can be located in a completely different part of the image, therefore making coalesced memory reads impossible. Moreover, the probability of hitting the same element is very low, which removes the advantages of using the fast texture cache. Shared memory cannot be used either since it is not known at execution time whether a candidate voxel in one of the threads will be reused in another thread within the same block and, if so, in which one specifically. Thus, it is not possible to save this value in a structured cache for future use.

Another major limiting factor is the need to sort the selected similar voxels in array A according to the difference in perfusion curves. This is difficult due to two reasons. To begin with, keeping track of sometimes thousands of data pairs (difference and voxel value) requires sufficient memory that quickly exceeds the amount of fast on-chip shared memory. Moreover, since shared memory needs to be shared between all threads in a thread block, it greatly reduces the number of voxels that can be processed in parallel. Thus, there is no other choice left other than accept writes into slow device memory. The fastest sorting algorithms require on average $\mathcal{O}(n \log n)$ comparisons, which can become a large number for typical settings of the filter. From the computational point of view this method becomes inefficient on the GPU, since memory reads and writes are done in device memory and if not all threads in a thread block (*warp*) have finished their sorting, the core cannot be released to process another block of threads.

A very straightforward approach to reduce the computational burden due to sorting at the end is to sort each element at insertion by using, e.g., min-heap structure of size FS . With this approach, one does still need to keep the differences and values of FS elements. However, FS is almost always smaller than the size of A and it can be possible to store all FS elements in fast memory. Theoretically, on average the insertion takes again $\mathcal{O}(n \log n)$ comparisons, but now the memory throughput is very high since we are working with fast shared memory only. In our case, this number will be even lower since the min-heap structure is sorted before each insertion. Nevertheless, the insertion must be performed as many times as the number of elements we wanted to have in our original array A . Furthermore, element insertion into a heap has multiple if-statements, which increases the likelihood of branching and therefore they could lead to the cores being idle more often.

Our Solution

Based on these algorithmic bottlenecks hindering an efficient implementation on GPU, we propose a simple change to the algorithm that enables us to work around these issues and attain fast computational times on the GPU at the cost of a small reduction in the numerical accuracy of the solution.

Figure 1 schematically represents our way of storing the same number of elements as in the original A with limited space requirements and a fixed number of operations ($\mathcal{O}(1)$) on insertion and reading in the worst-case scenario.

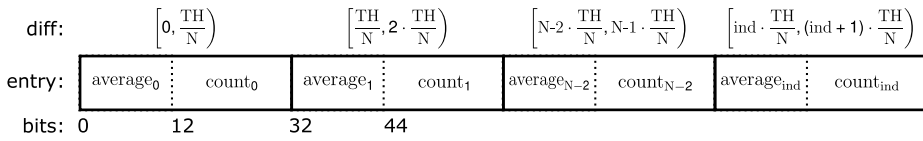


Figure 1: sorted array structure is used to store the average value of voxels within a specific range of differences, along with the corresponding voxel count required for this average calculation. The insertion and extraction operations have a worst-case performance of $\mathcal{O}(1)$.

As in the original version, a candidate voxel v_c is considered for future processing only if the difference between its and voxel v 's perfusion curves is below a given threshold TH . By carefully selecting TH , it is possible to save all good candidate values in a sorted manner with limited memory resources. Since the end-result of each kernel execution is an average of FS best candidate voxels v_c , it makes sense to store the average and corresponding count utilizing as little memory space as possible. Assuming that a typical CT image is in the range from -1024 to 3071 HU, storing each value using only positive numbers requires just 12 bits. In case of 32-bit integers it leaves us with 20 bits that can be used to store metadata corresponding to this number. For us, the only metadata we have for the average value is the number of voxels that were used to calculate it. Now we only need to store it in a sorted manner in a small data structure. We propose to create an array H of size N where each n -th bin stores the average value of the candidate voxels v_c with the differences between $n \cdot TH/N$ and $(n+1) \cdot TH/N$. For example, if $N=100$, $TH=2000$, the first bin is responsible for storing the average value of candidate voxels with the difference between 0 and 19. The size of the array N is a trade-off between the intended bin width for a certain TH and the number of threads that can run in parallel. For example, for the Nvidia A6000 GPU with 99 KB of shared memory, one can theoretically have 193 parallel threads with $N=128$ or just 96 threads with $N=256$ in each block.

Limitations

The changes introduced to the algorithm have a negative effect on the numerical accuracy of the 4D similarity filter, which is important to understand and to explore possible solutions.

To start, some accuracy is lost when calculating the average due to rounding, since only integer values can be stored. Theoretically, one could use fixed point number arithmetic to mitigate this problem, if one can accept that fewer bits will be used to represent the integer part of the number or the maximum count will be smaller. Following that, the width of the bin could be too broad so that almost all of FS best voxels can be found in one of the first bins. If the number of candidate voxels in one of the bins near the start of the array H greatly exceeds FS at some locations in the image, the noise reduction will not be homogeneous across the whole image. When it is not possible to increase the number of bins anymore due to memory constraints, one could move the array H into device memory, at the cost of increasing memory latency. Alternative solution could be to use the array H with non-uniform bin width. This way, the first bins can be made smaller to enable more granular summation of FS.

Materials & Methods

We compared the performance of our and commercial CPU implementations of the 4D similarity filter on simulated CT images of an anthropomorphic XCAT perfusion phantom. The acquisitions were simulated in low-dose mode ($\text{CTDI}_{\text{vol}} = 1.3 \text{ mGy}$ per phase) using a previously developed and validated CT simulator [114]. The images were reconstructed using an own implementation of the standard FDK algorithm. To avoid negative values and allow the custom array structure we applied the following linear scaling to arbitrary units: $AU = (HU + 3000) / 6$. In this preliminary work, we performed very simple analysis that included quantification of noise reduction in terms of signal-to-noise ratio (SNR) in a homogeneous region of the liver and estimation of the bias.

The GPU implementation has fewer and slightly different parameters than the CPU version. Even though we adjusted the parameters according to our understanding of the filter in order to achieve comparable performance, we note that both filters are not expected to lead to identical results. The size of array H was 128 bins to have a multiple of 32 to avoid bank conflicts when writing and reading from H . Mean squared error (MSE) TH was 12800, leading to each bin being 100 MSE units wide. FS was set to 100. For each voxel in the mask, 40,000 candidate voxels were checked.

Results

Figure 2 shows a slice of the original reconstructed phantom and two slices filtered using the CPU and GPU versions of the algorithm. Both results look very similar in terms of representing anatomical structures. To confirm this a subtraction image of the CPU and GPU output is shown in Figure 3.

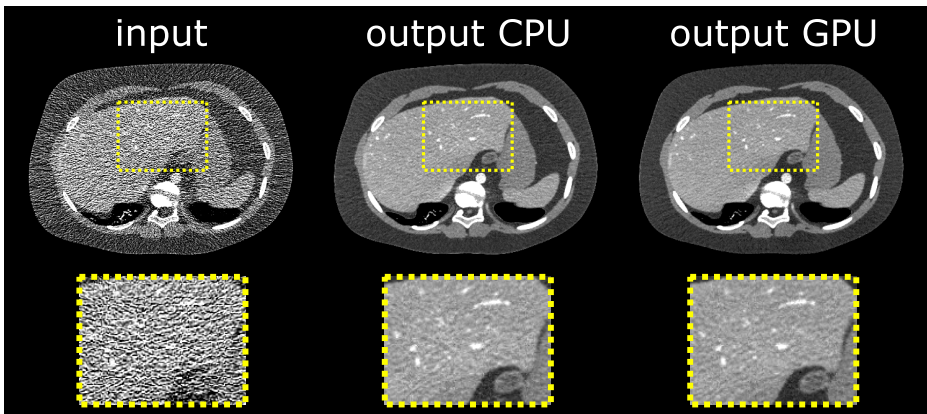


Figure 2: Comparison of exact CPU and approximate GPU 4D similarity filters. WW/WL: 200/50 HU.

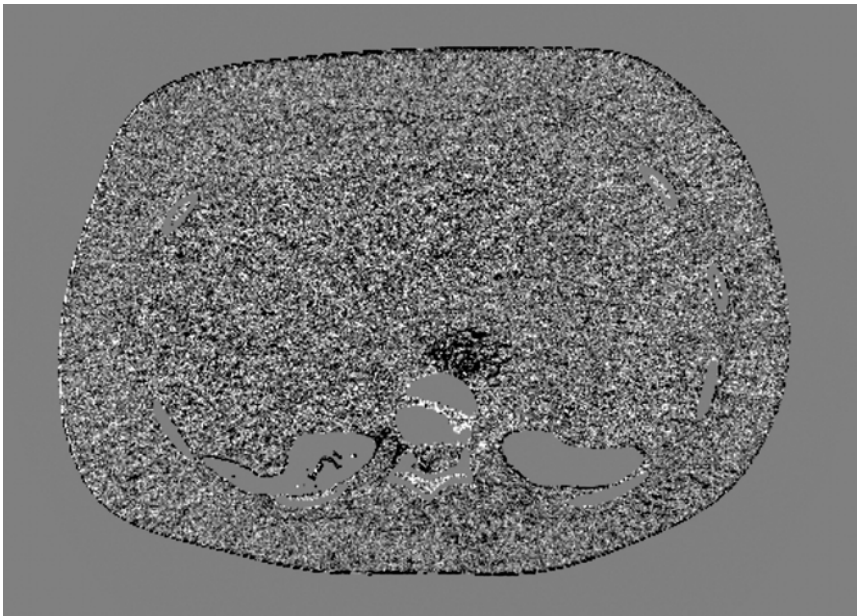


Figure 3: Subtraction image of gpu minus cpu. WL/WW: 0/100.

We calculated the SNR in the entire liver as we had perfect segmentation from the XCAT phantom. The GPU version has achieved slightly higher noise reduction than the CPU version, (SNR increased by a factor of 7.3 and 7.7, respectively). Since the noise is expected to decrease mainly as a function of the number of averaged voxels FS provided that they are sampled from the same distribution, and unlike in the CPU version, FS is not fixed, we calculated the actually used FS for each voxel in the output. Figure 4 shows the actually used FS to calculate the average for each voxel in the image. In the whole liver, we found FS to be 126 ± 19 which is higher than the intended FS used in the CPU version.

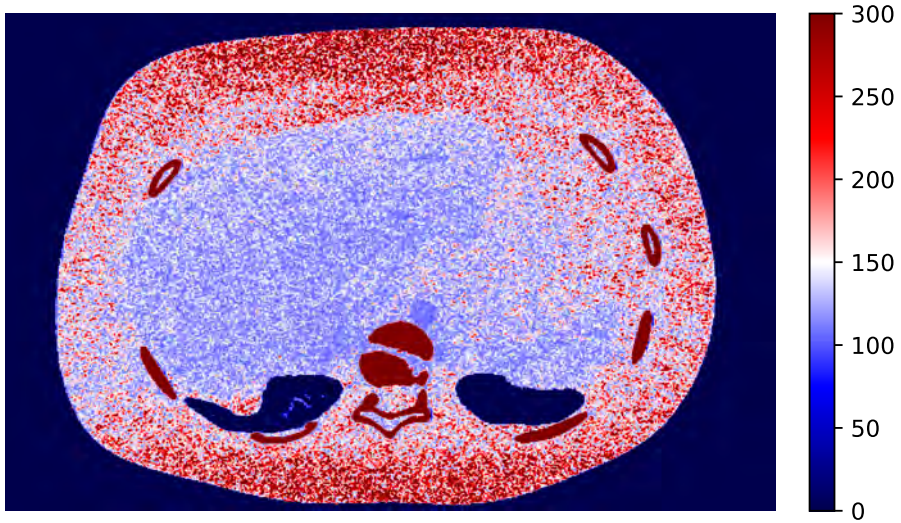


Figure 4: Actually, used FS for each voxel measured as a number of voxels used to calculate the average. The values are clipped at 300.

No visible bias is observed when comparing images filtered on CPU and GPU. To verify this we constructed the Bland-Altman plot shown in Figure 5. There is a small negative bias of -6 HU. There appears to be no clear dependency of the difference between two versions of the filter and voxel grayscale value. Moreover, we observed that both the CPU and GPU versions introduce either positive or negative bias depending on the location in the liver or other organs. There is no clear dependency on anatomy, functional or image properties.

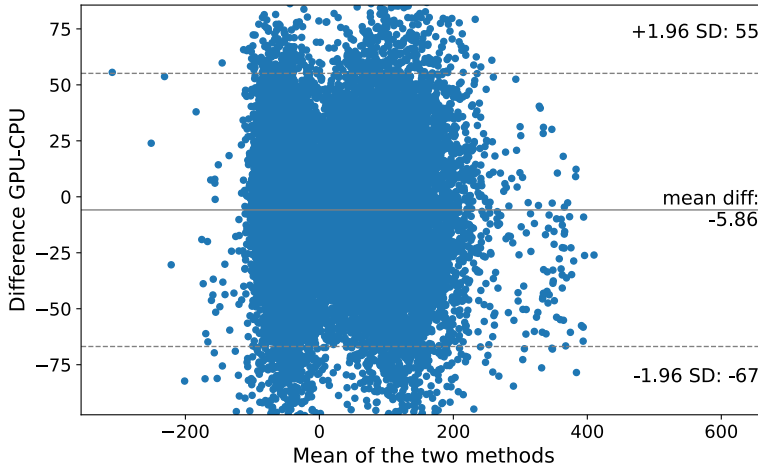


Figure 5: Bland-Altman plot comparing GPU and CPU versions of the filter. The output of the GPU version is on average lower.

The GPU result can be obtained within 44 seconds on an Nvidia A6000 GPU, which is on average 13 times faster than the multi-threaded commercial CPU implementation that is executed on 32 cores, based on the mean of three runs. The Nvidia profiler has also shown that the L1 and L2 caches are not fully utilized. In order to improve reading from device memory, we implemented memory prefetching, eight candidate voxels are loaded from the device memory into registers and are then processed. This modification improved the utilization of the caches. Better memory utilization increased compute throughput by 385% and enabled, thus, the execution of the filter within 27 seconds, which is roughly 23 times faster than the CPU version.

Discussion

We modified the original algorithm of the 4D similarity filter in order to avoid GPU-specific computational constraints. With this we were able to run it on GPU several times faster than with highly-optimized multi-threaded code on the CPU. The main change was replacing the sorting operation by using an ordered hash-like structure that can reside in fast but limited on-chip shared memory and thereby enabling $O(1)$ writing and reading. However, we needed to sacrifice accuracy to use this data structure. We performed first tests with both versions of the similarity filter to study the effect of this accuracy loss.

Overall, both versions produce very similar outputs with comparable set of parameters. The main difference is a higher level of denoising than anticipated, which is explained by the fact that FS is not a fixed parameter anymore. Therefore, it leads to different level of noise reduction in various parts of the image, as demonstrated. To counteract this effect, the bin width in the hash H could be chosen smaller, therefore enabling more granular division of candidate voxels and thus increasing the chance of using the sought number of voxels FS for averaging.

An important advantage of our implementation is significantly reduced complexity of the code. This simplification will enable us to add modifications to the algorithm and explore additional features such as looking for neighbors in a local region or using different metrics to select the voxels. Furthermore, our implementation enables the filtering of single phase in the sequence. An interesting scenario to investigate would be the application of the filter to reduce the noise in every second or third image in the sequence for calculation of approximate perfusion maps in clinical situations in which the time of diagnosis is the main concern, while more images are being processed in the background to update the perfusion maps with more accurate data.

In this work we performed an initial evaluation of how the filter is affected by the algorithm changes in a digital phantom. At the moment it is not clear whether the introduction of slightly non-uniform noise reduction throughout the image and different anatomies we found will have negative effects on the performance of radiologists or further processing of the images, e.g., to calculate perfusion maps. Based on these initial results we plan to continue with further analysis of the performance by, e.g., quantifying bias of the perfusion curves in different anatomies. Such an analysis combined with hyperparameter tuning must be done in order to ensure correct and trustworthy performance.

Conclusion and Outlook

We implemented a GPU version of the 4D similarity filter that is on average approximately 13 times faster than a highly optimized commercial multi-threaded CPU realization by introducing a change into the original algorithm to make it more GPU-friendly at the cost of lowering accuracy. Our modification to the algorithm removed a significant bottleneck for GPU implementation. Our further non algorithm-specific modification enabled further speedup by factor of 1.7 compared to the already modified algorithm.



Chapter 6

4D Similarity Filter for ischemic stroke

Original title: 4D Similarity Filter for Cerebral CT perfusion in Acute Stroke

Sjoerd A. M. Tunissen, Ewoud J. Smit, Mikhail Mikerov, and Ioannis Sechopoulos

Accepted in: Radiology

ABSTRACT

Background

Brain CT perfusion is used for diagnosing ischemic stroke and treatment planning.

Purpose

Optimizing the 4D Similarity Filter (4DSF), a noise reducing filter, for CT perfusion in patients with acute stroke.

Methods

To reduce noise in dynamic perfusion images, the 4DSF averages voxels with the most similar time-intensity curves. The commercial implementation searches for similar voxels based on temporal mean. The proposed stroke-specific 4DSF (s4DSF) searches for voxels with similar peak times, thus measuring changes in blood supply that are characteristic of acute stroke. For evaluation, nine stroke scenarios were simulated using a digital phantom. For each scenario, a 30-phase CT perfusion protocol was simulated. Perfusion analysis was performed using (i) Bayesian estimation on 4DSF-filtered (clinical standard) and s4DSF-filtered CT scans, and (ii) SVD on original (clinical standard) and s4DSF-filtered CT scans to obtain five different perfusion maps. For all maps, the contrast and root-mean-square error of the stroke and contralateral regions were compared with and without s4DSF. Improvements in contrast correctness were evaluated with a binomial test. To show clinical potential, CT perfusion scans of twelve patients admitted with suspected stroke were retrospectively retrieved. In these cases, s4DSF performance was visually assessed and compared to the clinical Bayesian and SVD methods.

Results

Compared to the clinical version, the s4DSF improved contrast correctness in 64% of the Bayesian estimation (29/45) and 73% of the SVD maps (33/45) ($P=.036$ and $P=.001$, respectively). Moreover, the root-mean-square error was lower for the s4DSF-filtered maps in 70% (63/90) of Bayesian- and 99% (89/90) of SVD-derived perfusion regions. Analysis of patient data yielded similar results, with s4DSF showing more visible and clearly delineated stroke regions.

Conclusion

The s4DSF shows better performance than current clinical filters in phantom simulations and patient cases. Therefore, application of the s4DSF could improve diagnostic accuracy of cerebral perfusion maps.

This study was supported by Canon Medical Systems Corporation, by providing technical support in the use of the Vitrea workstation so that the perfusion analysis could be performed on the images filtered by us offline, and by co-funding project FILTER, funded by Health Holland (project LSHM19020). The authors, none of whom are employees or consultants of Canon, had control of the data and the information submitted for publication at all times.

Introduction

CT perfusion (CTP) is used in acute stroke to diagnose ischemia, make treatment decisions on reperfusion therapy, and predict outcome [119–121]. Multiple prospective trials, such as the DAWN and DEFUSE 3 [122,123], demonstrated the ability of CTP to guide treatment decisions by identifying patients that are suitable for intra-arterial thrombectomy even after 6 hours since symptom onset.

Most commonly, methods for estimating perfusion maps from CTP images are based on deconvolution using singular value decomposition (SVD) [16]. However, these SVD-based methods often fail for images with low signal-to-noise ratios (SNRs) [124]. Since cerebral CTP protocols consist of multiple CT scans over time, the dose per scan must be kept relatively low, resulting in a low SNR and making it challenging to obtain accurate perfusion maps. Different variants of SVD-based methods have been proposed to minimize this problem [125], but they do not yield sufficient accuracy in low-SNR situations. Another method based on Bayesian estimation, which is more robust to low SNRs but still shows worse performance in these scenarios, has been proposed [17].

Therefore, it is important to increase the SNR of CTP scans as a preprocessing step before perfusion analysis to maximize the accuracy of the resulting perfusion maps [32,37,38,126]. One such noise reducing filter, the 4D similarity filter (4DSF) [55], has been combined with a Bayesian estimation method and made available on a clinical perfusion analysis solution (Brain Perfusion 4D, Vitrea, Vital Images, Minnetonka, MN, USA); this solution has already been shown to improve perfusion analysis [100].

In this work, an alternative implementation of the 4DSF for diagnosing ischemic stroke that aims to accurately depict differences between healthy and ischemic brain tissue is proposed and tested.

Methods

The newly proposed stroke-specific implementation of the 4DSF is referred to as the s4DSF. The s4DSF was used in combination with two perfusion methods available on the Vitrea workstation (version 7.11.0), the Bayesian estimation method and an SVD method (called SVD+). To examine whether the proposed s4DSF outperformed the current standards used clinically, it was compared with

the 4DSF-Bayesian estimation and standalone SVD methods. When using s4DSF in combination with the Bayesian estimation method, the clinically used Bayesian estimation implementation was used with the original 4DSF turned off.

Phantom data

To quantify the performance of the s4DSF and the impact of the changes proposed, a digital anthropomorphic brain phantom and perfusion model developed by Divel et al. [127] were used. In this brain phantom, the vessels, gray matter, and white matter are divided into multiple smaller regions such that local perfusion differences, including strokes, can be simulated. A previously developed and validated CT simulator [114] was used to generate CT images from this phantom. This simulator mimics the physical process of image acquisition and therefore results in realistic CT images that mimic real noise and spatial resolution limitations. More details on the phantom and images can be found in Appendix A.

A total of nine different stroke scenarios were simulated with three different peak delays of 1.5, 3.0, and 5.0 seconds and three different peak intensities of 0.2, 0.4, and 0.6 relative to the time attenuation curves (TACs) of the contralateral, healthy white matter. A 30-phase CT protocol was simulated for each scenario. All strokes were approximately 11.5 mL in size and located in the left hemisphere; none of them had an infarct, i.e., the brain tissue itself was not simulated as hypodense. Figure 1 shows a simulated CT image with and without noise.

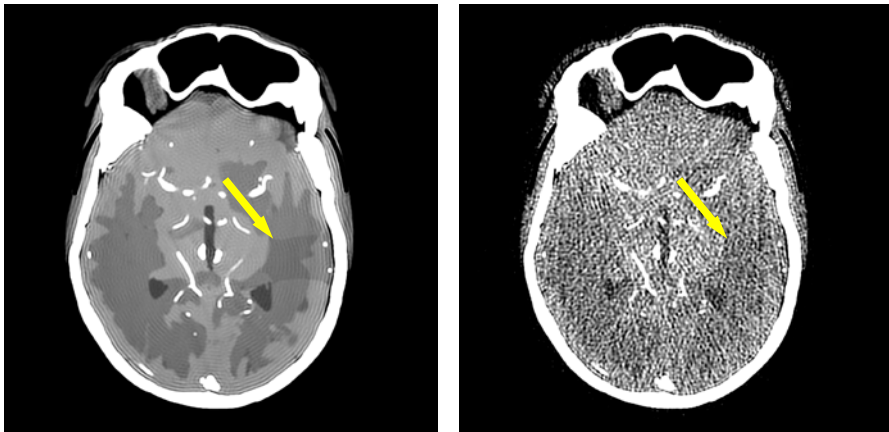


Figure 1: Reconstructions of the simulated acquisition of the phantom at the twelfth timepoint. (left) Reconstruction of the phantom without noise, i.e., the reference. (right) Reconstruction with a noise level comparable to a clinical scan in our institution. The ischemia is visible as a hypodense region in the contrast-enhanced image (indicated by the yellow arrow), since it receives less blood. Please note that the ischemic region is hardly visible on the noisy image on the right. WW/WL: 200/50.

Patient data

As an exploratory study on clinical feasibility, the CT perfusion data of twelve patients were retrospectively obtained from the imaging archive of a tertiary academic medical center. These data were obtained under an approved protocol for the use of existing clinical data obtained for clinical purposes after anonymization, with the requirement for informed consent waived (protocol numbers: CMO 2016-3045, Projects 20031, v5 and v6). The study cohort comprised of twelve patients of two groups. The first group included two consecutive patients admitted for acute stroke and subsequently underwent intra-arterial thrombectomy between January 1, 2024, and May 1, 2024, and the second group comprised of ten consecutive patients who were presented with suspected stroke and treated between August 1, 2024, and August 18, 2024. This non-continuous patient selection was due to the peer-review process. An initial evaluation of two patients was followed by the addition of ten patients at the reviewer's request. Because raw scanner data, required for this study, was not routinely archived at our institution, patients treated immediately following the initial two could not be included, necessitating the selection of a subsequent cohort of ten patients. There were no exclusion criteria.

Filter

The 4DSF searches for voxels with similar TACs and averages these voxels together, improving the SNR in the image. However, because computation time limits the number of voxels that can be assessed for similarity, the voxels are sorted on the basis of a metric; in the current version, this metric is the temporal mean and the voxels closest after sorting based on this metric are assessed. Importantly, since the tissue densities and enhancements are very similar and the image noise is relatively high, the difference between healthy and ischemic tissue in the brain is relatively small on the temporal mean images. Therefore, the temporal mean has limited discriminative ability for separating ischemic tissue from healthy tissue. In the s4DSF, the voxels are sorted on the basis of peak time. Figure 2 shows an example of the temporal mean-based values and the peak time-based values in which the peak time-based metric can clearly discriminate the stroke voxels, whereas little distinction is possible when the temporal mean-based values are used. Additionally, in the s4DSF method, only voxels that are physically close to the voxels being processed are considered. The technical details about the 4DSF are described in Appendix B.

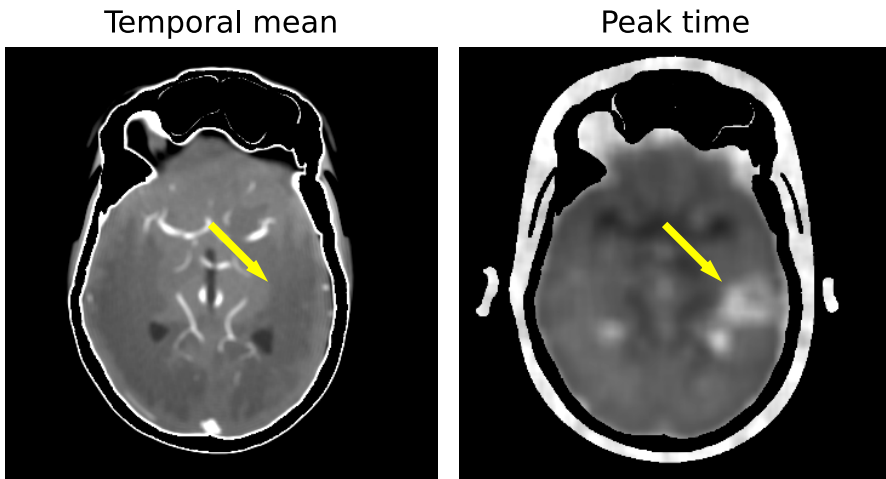


Figure 2: Images of the metrics used to sort the voxels and, consequently, determine candidates to obtain the average. (left) The temporal mean-based metric. (right) The peak time-based metric. The yellow arrow indicates the stroke area. Owing to its greater contrast between the ischemic and healthy tissues, the peak time-based metric better discriminates the stroke voxels.

Perfusion Analysis

Perfusion maps from the simulated and real patient CTP images were generated via the Vitrea 'Brain Perfusion 4D' tool. Specifically, relative cerebral blood volume (rCBV), relative cerebral blood flow (rCBF), mean transit time (MTT), time to peak (TTP), and Delay with 320 slices, each 0.5 mm thick with 512x512 pixels covering a field of view of 220 mm, were generated.

Maps using both the Bayesian estimation and SVD method were obtained. For each method, the same analysis was performed. The perfusion maps of the clinical implementation and proposed implementation using s4DSF were compared to the reference maps, i.e., the noiseless phantom data. The clinical implementation of Bayesian estimation uses the original 4DSF already in its preprocessing; therefore, when combining s4DSF with Bayesian estimation, the 4DSF step was turned off. The clinical implementation of the SVD method does not use the 4DSF during preprocessing.

The comparison was performed in two regions of interest (ROIs), defined by the known stroke (ipsilateral) and healthy (contralateral) areas of the phantom brain.

For each ROI, the average value was calculated. The contrast (difference in average value) between these two ROIs, i.e., the healthy and ischemic ROIs, was

subsequently determined. This contrast indicates the visibility of the stroke in each map. The contrast of the reference maps, i.e., the noiseless phantom data, served as the reference standard, and the contrast of the clinical and proposed methods were compared to the reference standard contrast to determine which was more accurate; this binary measure is referred to as contrast correctness.

Next, the root-mean-square error (RMSE) between the reference standard and clinical implementation was determined, as was that between the reference standard and proposed implementation. The RMSE is a combination of the systematic error and the noise level in the perfusion map and is therefore a quantitative measure of accuracy.

All these measurements were performed for all nine different stroke scenarios. Three scenarios were selected to represent mild (delay of 1.5 s and peak of 0.6), moderate (delay of 3 s and peak of 0.4) and severe perfusion deficits (delay of 5 s and peak of 0.2) and are reported in the results. The other scenarios are presented in Appendix C.

Consistent with the analysis of the simulated data, perfusion maps were obtained for the patient cases using both the Bayesian estimation and SVD methods. The clinical and proposed implementations were both used so that the effect of s4DSF on the perfusion maps could be analyzed. Since no reference standard was available for the patient data, quantitative analysis was not possible and the results were analyzed only visually.

Statistical Analysis for Phantom data

Before the contrast was calculated, differences in the values within the healthy and ischemic ROIs of the same case were evaluated via a Mann-Whitney U test, since the distributions within the ROIs were not expected to be normal and were independent. The values of the same ROIs processed with the clinical and proposed versions were compared using a Wilcoxon signed-rank test, since the distributions within the ROIs are not expected to be normal and are paired. The threshold for significance was chosen as $P < .01$ for both tests.

Finally, a binomial test was performed based on the number of perfusion maps in which the proposed method had a contrast closer to the reference standard than did the clinical method to determine whether one method outperformed the other. The threshold for significant difference was chosen as $P < .05$ for both tests. All statistical tests were performed using python package SciPy (version 1.7.3).

Results

Bayesian estimation Analysis

Figure 3 shows the perfusion maps of the 3 second delay and 0.4 relative peak intensity case for the Bayesian estimation method. The ischemic region is more visible on the TTP, MTT, and delay maps for the proposed method than those obtained with the clinical method. However, the maps of the proposed method seem slightly more blurred. The perfusion maps of all other stroke scenarios can be found in Appendix C.

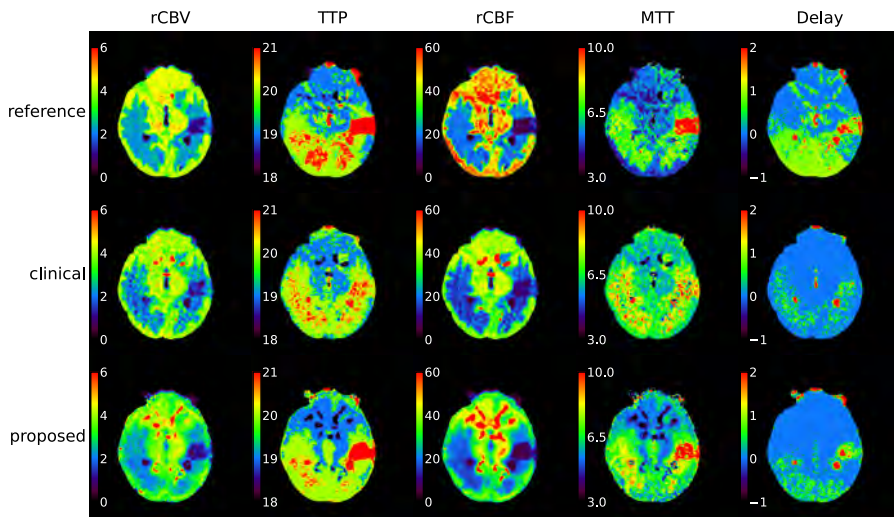


Figure 3: Perfusion maps of the Bayesian estimation method for the phantom with left-sided ischemia, with a 3 second peak delay and a peak height of 0.4 relative to healthy tissue. The top row shows the reference perfusion maps, the middle row shows the perfusion maps resulting from the clinical implementation, and the bottom row shows the perfusion maps of the proposed implementation.

In Table 1, the average values of the ischemic and healthy ROIs and their contrasts are presented for three different stroke scenarios. The results show that the contrast correctness of the proposed method was superior to that of the clinical method for most of the TTP, rCBF, and MTT maps. As expected, the proposed method was least superior in the stroke scenario with the smallest delay at a peak time of 1.5 s for Bayesian estimation. The results of the other six scenarios, presented in Tables S1 and S2 of Appendix C in the online supplementary materials, were similar. Overall, i.e., for all nine scenarios, with five perfusion maps each, 29 out of the 45 maps showed improved contrast correctness when the proposed method was used and the remaining 16 maps showed worse contrast correctness. The p value resulting from the binomial test was .036. The values of the ischemic and healthy ROIs were

significantly different for all maps according to the Mann-Whitney U test. The values of the clinical and proposed methods were significantly different for all maps except one, as indicated in Table 1, according to the Wilcoxon signed-rank test.

Table 2 presents the RMSE of both the clinical and proposed methods relative to the reference standard for each stroke scenario. The TTP, MTT, and Delay maps generated with the proposed method combined with Bayesian estimation consistently resulted in the lowest RMSE. For the rCBV and rCBF, the proposed method seemed to have slightly worse performance in terms of RMSE. The results of the other six scenarios were similar (Tables S3 and S4 of Appendix C). Overall, for all nine scenarios, with five perfusion maps each and two ROIs per map, the RMSE was reduced in 63 out of 90 analyzed ROIs for the Bayesian estimation method.

Table 1: Mean values of the reference, clinical, and proposed method-based maps for the ischemic and healthy ROIs and their contrast for the Bayesian estimation method. The contrast values closest to those of the reference standard have a gray background.

	<i>Mild stroke scenario</i> <i>Peak delay = 1.5</i> <i>Peak attenuation = 0.6</i>				<i>Moderate stroke scenario</i> <i>Peak delay = 3.0</i> <i>Peak attenuation = 0.4</i>				<i>Severe stroke scenario</i> <i>Peak delay = 5.0</i> <i>Peak attenuation = 0.2</i>			
	ischemic	healthy	contrast		ischemic	healthy	contrast		ischemic	healthy	contrast	
	Reference	1.79	3.03	1.25	1.50	3.08	1.58	0.89	0.89	3.03	2.14	
rCBV (mL/100g)	Clinical	1.93	2.85	0.92	1.58	2.84	1.27	1.03	1.03	2.43	1.40	
	Proposed	2.29	3.20	0.90	1.51	2.77	1.26	1.11	1.11	3.04	1.93	
	Reference	21.00	19.49	-1.51	22.14	19.50	-2.64	23.34	23.34	19.49	-3.85	
TTP (s)	Clinical	20.12	19.53	-0.59	20.57	19.71	-0.86	21.17	21.17	19.54	-1.63	
	Proposed	20.35	19.76	-0.59	21.46	19.73	-1.73	22.36	22.36	19.80	-2.57	
	Reference	14.20	29.58	15.38	9.88	29.70	19.82	5.97	5.97	29.34	23.37	
rCBF (mL/100g/min)	Clinical	14.26	23.20	8.94	11.47	24.74	13.27	7.35	7.35	22.02	14.67	
	Proposed	19.42	30.57	11.16	10.45	24.75	14.31	8.07	8.07	29.84	21.77	
	Reference	8.01	6.54	-1.47	10.85	6.57	-4.28	11.93	11.93	6.56	-5.37	
MTT (s)	Clinical	8.15	7.58	-0.56	8.06	7.10	-0.96	8.15	8.15	6.75	-1.41	
	Proposed	7.20	6.33	-0.87	9.21	6.81	-2.40	9.11	9.11	6.15	-2.96	
	Reference	0.84	-0.12	-0.96	1.05	-0.17	-1.22	1.85	1.85	-0.15	-2.00	
Delay (s)	Clinical	0.10	-0.67	-0.77	0.47	-0.35	-0.81	0.93	0.93	-1.14	-2.07	
	Proposed	0.17	-0.34	-0.51	0.72	-0.31	-1.04	1.47	1.47	-0.33	-1.80	

‡ Maps between which there is no significant difference

Table 2: RMSE of the clinical and proposed methods compared with the reference standard when the Bayesian estimation method is used. A lower RMSE represents better agreement with the reference maps and is highlighted with a gray background.

		<i>Mild stroke scenario</i> <i>Peak delay = 1.5</i> <i>Peak attenuation = 0.6</i>		<i>Moderate stroke scenario</i> <i>Peak delay = 3.0</i> <i>Peak attenuation = 0.4</i>		<i>Severe stroke scenario</i> <i>Peak delay = 5.0</i> <i>Peak attenuation = 0.2</i>	
		ischemic	healthy	ischemic	healthy	ischemic	healthy
rCBV (mL/100g)	Clinical	0.43	0.53	0.49	0.45	0.48	0.83
	Proposed	0.75	0.69	0.47	0.79	0.60	0.69
TTP (s)	Clinical	0.96	1.69	1.78	0.99	2.66	2.69
	Proposed	0.84	0.62	1.23	0.60	1.98	0.60
rCBF (mL/100g/min)	Clinical	3.12	8.83	10.48	7.09	10.59	10.27
	Proposed	7.29	9.27	10.07	10.72	11.08	8.89
MTT (s)	Clinical	1.71	1.57	3.72	1.08	5.22	1.23
	Proposed	1.69	0.90	2.90	0.91	4.58	0.93
Delay (s)	Clinical	0.85	4.37	1.08	2.39	2.42	6.92
	Proposed	0.80	0.87	0.95	0.83	1.64	0.73

SVD Analysis

Figure 4 shows the SVD perfusion maps for the 3 second delay and 0.4 relative peak intensity scenario. Compared with those of the clinical method, the perfusion maps of the proposed method have little noise content, considerably improving the visibility of the stroke. However, the maps of the proposed method are less well delineated. The perfusion maps of all other stroke scenarios can be found in Appendix C in the online supplementary materials.

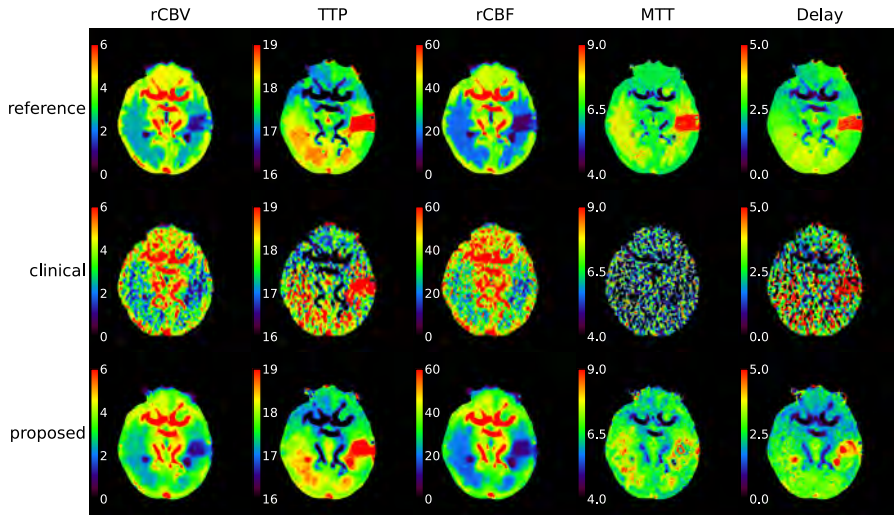


Figure 4: Perfusion maps of the SVD method for the phantom with left-sided ischemia, with a 3 second peak delay and a peak height of 0.4 relative to healthy tissue. The top row shows the reference standard perfusion maps, the middle row shows the perfusion maps obtained with the clinical implementation, and the bottom row shows the perfusion maps obtained with the proposed implementation. SVD – singular value decomposition.

Table 3: Mean values of the reference, clinical, and proposed method-based maps for the ischemic and healthy ROIs and their contrast for the SVD method. The contrast values closest to those of the reference standard have a gray background.

	Mild stroke scenario						Moderate stroke scenario						Severe stroke scenario					
	Peak delay = 1.5			Peak delay = 3.0			Peak delay = 5.0			Peak delay = 7.0			Peak delay = 9.0			Peak delay = 11.0		
	ischemic	healthy	contrast	ischemic	healthy	contrast	ischemic	healthy	contrast	ischemic	healthy	contrast	ischemic	healthy	contrast	ischemic	healthy	contrast
rCBV (mL/100g)	Reference	2.00	3.19	1.19	1.68	1.52	1.08	1.30	1.30	1.30	1.30	1.30	1.30	1.30	1.30	1.30	1.30	1.30
	Clinical	2.23	3.23	1.00	1.89	1.49	1.30	1.30	1.30	1.30	1.30	1.30	1.30	1.30	1.30	1.30	1.30	1.30
	Proposed	2.36	3.24	0.88	1.76	1.37	1.19	1.19	1.19	1.19	1.19	1.19	1.19	1.19	1.19	1.19	1.19	1.19
TTP (s)	Reference	18.86	17.68	-1.18	19.84	-2.16	20.54	17.68	-2.16	20.54	17.68	-2.16	20.54	17.68	-2.16	20.54	17.68	-2.16
	Clinical	18.91	17.54	-1.36	19.91	-2.26	21.05	17.65	-2.26	21.05	17.39	-3.66	21.05	17.39	-3.66	21.05	17.39	-3.66
	Proposed	18.44	17.64	-0.80	19.13	-1.46	20.02	17.67	-1.46	20.02	17.70	-2.31	20.02	17.70	-2.31	20.02	17.70	-2.31
rCBF (mL/100g/min)	Reference	16.45	27.86	11.41	11.46	16.17	7.63	27.62	16.17	7.63	27.70	20.07	16.17	27.70	20.07	16.17	27.70	20.07
	Clinical	23.37	30.57	7.20	24.73	12.17	16.46	36.90	12.17	16.46	28.93	12.46	16.46	28.93	12.46	16.46	28.93	12.46
	Proposed	19.16	27.50	8.33	14.72	12.27	9.04	27.00	12.27	9.04	26.07	17.03	12.27	26.07	17.03	12.27	26.07	17.03
MTT (s)	Reference	7.52	6.98	-0.54	9.22	-2.17	9.34	7.05	-2.17	9.34	7.01	-2.33	9.34	7.01	-2.33	9.34	7.01	-2.33
	Clinical	5.93	6.42	0.49	4.78	0.84	4.93	5.62	0.84	4.93	6.43	1.51	4.93	6.43	1.51	4.93	6.43	1.51
	Proposed	7.46	7.08	-0.38	7.33	-0.34	8.27	6.99	-0.34	8.27	7.16	-1.11	8.27	7.16	-1.11	8.27	7.16	-1.11
Delay (s)	Reference	3.70	2.46	-1.25	4.17	-1.80	4.93	2.37	-1.80	4.93	2.46	-2.47	4.93	2.46	-2.47	4.93	2.46	-2.47
	Clinical	1.78	1.10	-0.68	3.67	-2.70	3.94	0.97	-2.70	3.94	0.62	-3.32	3.94	0.62	-3.32	3.94	0.62	-3.32
	Proposed	2.98	2.14	-0.84	3.94	-1.80	4.13	2.13	-1.80	4.13	1.96	-2.17	4.13	1.96	-2.17	4.13	1.96	-2.17

‡ Maps between which there is no significant difference

In Table 3, the average values of the ischemic and healthy ROIs and their contrasts are presented for the three different stroke scenarios. The results show that the contrast correctness of the proposed method was superior to that of the clinical method for the Delay, rCBF, and MTT maps. The results of the other six scenarios were similar (Tables S5 and S6 of Appendix C). Overall, i.e., for all nine scenarios, with five perfusion maps each, the contrast correctness was improved in 33 out of the 45 maps and worse in 12 out of 45 maps when the proposed method was used. The p value resulting from the binomial test was .001. The values of the ischemic and healthy ROIs were significantly different for all maps according to the Mann-Whitney U test. The values of the clinical and proposed methods were significantly different for all maps except three according to the Wilcoxon signed-rank test (Tables 1 and S6).

Table 4 presents the RMSE between the reference method and both the clinical and proposed methods for each stroke scenario. The proposed method combined with SVD consistently resulted in the lowest RMSE. The results of the other six scenarios were similar (Tables S7 and S8 of Appendix C). Overall, for all nine scenarios, with five perfusion maps each and two ROIs per map, the RMSE was reduced in 89 out of 90 analyzed ROIs for the SVD method.

Table 4: RMSE of the clinical and proposed methods compared with the reference standard when the SVD method is used. A lower RMSE represents better agreement with the reference maps and is highlighted with a gray background.

		Mild stroke scenario Peak delay = 1.5 Peak attenuation = 0.6		Moderate stroke scenario Peak delay = 3.0 Peak attenuation = 0.4		Severe stroke scenario Peak delay = 5.0 Peak attenuation = 0.2	
		ischemic	healthy	ischemic	healthy	ischemic	healthy
rCBV (mL/100g)	Clinical	0.82	0.93	0.86	0.88	0.80	1.14
	Proposed	0.56	0.55	0.41	0.58	0.46	0.54
TTP (s)	Clinical	1.28	1.73	2.23	1.15	4.83	2.55
	Proposed	0.68	0.39	1.34	0.37	1.84	0.34
rCBF (mL/100g/min)	Clinical	12.09	9.48	18.1	15.02	13.69	9.86
	Proposed	4.83	6.12	5.24	5.63	4.77	6.06
MTT (s)	Clinical	2.78	1.90	5.09	2.23	5.42	2.11
	Proposed	1.03	0.74	2.58	0.67	2.77	0.72
Delay (s)	Clinical	8.45	6.87	8.12	5.44	12.47	8.86
	Proposed	1.14	1.00	1.75	0.99	5.05	1.02

Patient data

Bayesian estimation analysis

Figure 5 shows the perfusion maps of two patient cases with the Vitrea Bayesian estimation method. For both patients, the stroke was more visible and clearly delineated with the proposed method. For Patient 1, the time-based maps (i.e., TTP, MTT, and Delay) show a much better delineation of the ischemic area. For Patient 2, these maps show a deficit in the images from the proposed method, which is hard to see in the clinical images, especially on the Delay map. Patient 1 was woman age 81 and Patient 2 was a man age 73. The results of the other ten patient cases are presented in Appendix D. Out of the ten patients nine were men and one was a woman and their age ranged from 59 and 80.

SVD analysis

Figure 6 shows the perfusion maps of the two patient cases when the SVD method was used. Compared with those of the clinical method, the perfusion maps of the proposed method had little noise. For all the maps of Patient 1, with the exception of the TTP, the stroke area is much better delineated by the proposed method. For Patient 2, the stroke is barely visible on the presented slice of the rCBV, rCBF, and MTT maps of the clinical implementation; however, on the maps of the proposed implementation, the stroke is clearly visible. The results of the remaining 10 patient cases are presented in Appendix D.

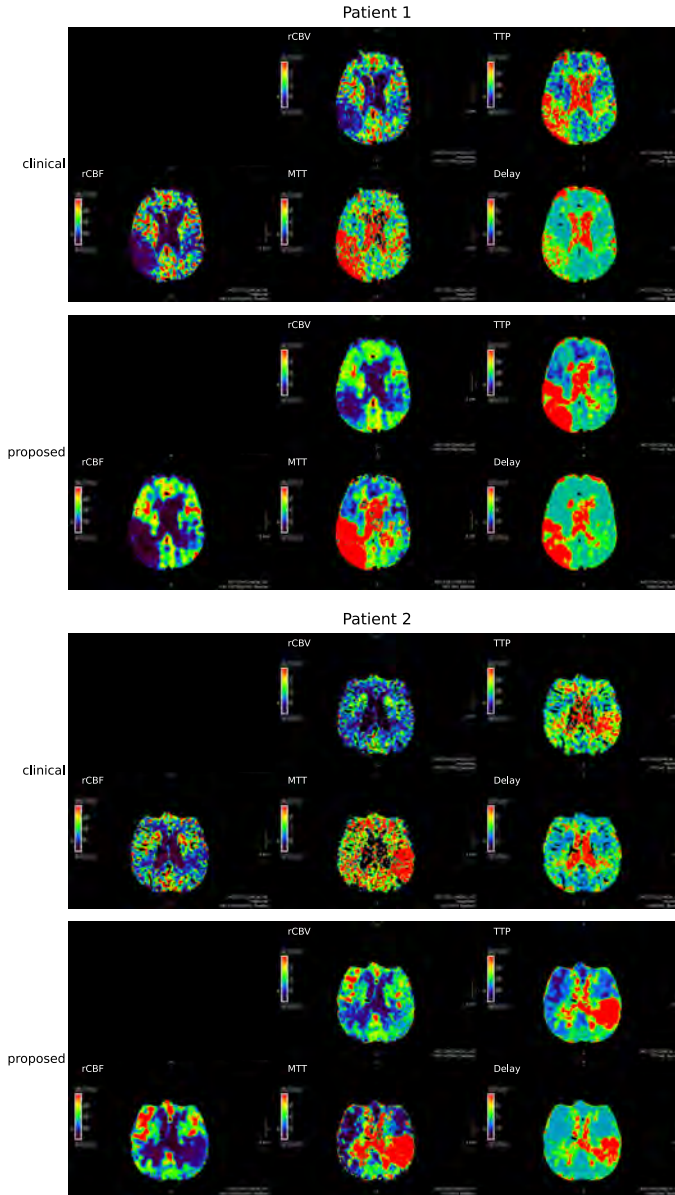


Figure 5: The perfusion maps of the patient data when using the Bayesian estimation method. The top part titled patient 1 shows the perfusion maps of an 81-year-old female with right sided middle cerebral artery occlusion presented as a wake-up stroke (NIHSS 13). The bottom part titled patient 2 shows perfusion maps of a 73-year-old male with a left sided middle cerebral artery occlusion scanned 45 minutes after onset of symptoms (NIHSS 4). For both patients the stroke is more visible and better delineated in case of the proposed method. For patient 1, the time-based maps (i.e., TTP, MTT, and Delay) show a much better delineation of the ischemic area. For patient 2 these maps show a deficit in the images from the proposed method, which is hard to see in the clinical images, especially on the Delay map.

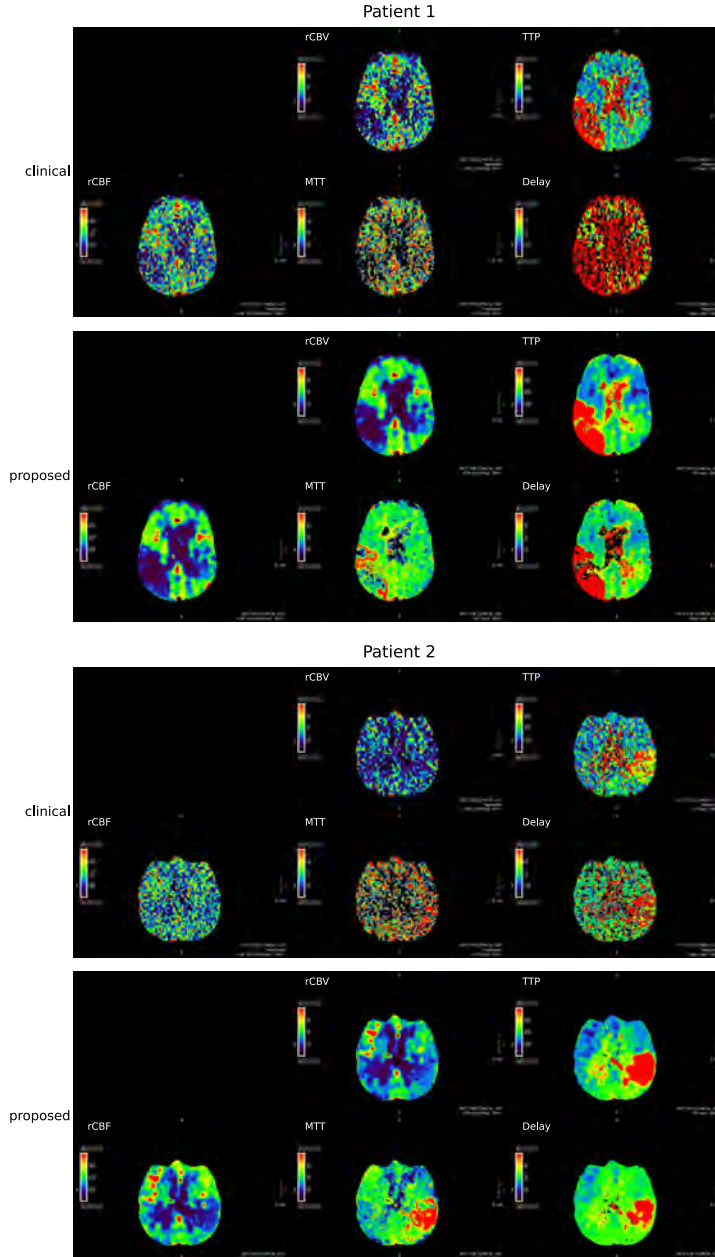


Figure 6: Perfusion maps of the patient data when the SVD method is used. The top part labeled Patient 1 shows the perfusion maps of an 81-year-old female with right-sided middle cerebral artery occlusion who presented with a wake-up stroke (NIHSS 13). The bottom part labeled Patient 2 shows perfusion maps of a 73-year-old male with left-sided middle cerebral artery occlusion scanned 45 minutes after the onset of symptoms (NIHSS 4). For both patients, the perfusion maps of the proposed method have less noise than those of the clinical method.

Discussion

Considering the characteristics of the perfusion deficits associated with stroke, this study proposed an alternative implementation of a clinical 4-dimensional noise filter, the 4DSF, for diagnosing ischemic stroke in CTP. While the clinical version of the 4DSF searches for similar voxels on the basis of tissue density and average enhancement, our proposed method searches for voxels on the basis of iodinated contrast timing, since it provides more discriminative power for the perfusion deficits associated with stroke.

As expected for the Bayesian estimation method, the proposed approach worked best for perfusion maps that relate to timing information, i.e., the TTP and MTT maps. The results of the phantom-based analysis for the Bayesian estimation method showed that the proposed s4DSF provides perfusion maps that are superior to the clinical version.

We found that both the contrast correctness of the perfusion deficit, which is important for visual assessment, and the accuracy of the perfusion values, which is important for, for example, automatic infarct core and penumbra quantification, improved. The contrast correctness improved in 64% (29/45) of the perfusion maps with $P=.036$. The rCBV maps, however, were slightly worse with the proposed implementation in terms of RMSE, whereas the contrast correctness of the perfusion deficit was similar. This has potential implications for automatic infarct core detection, but visual assessment of the rCBV maps should have results similar to those of the current clinical implementation. Owing to the central volume principle, which states that the rCBF is equal to the rCBV divided by the MTT, these moderate results for the rCBV maps also resulted in moderate results for the CBF maps. However, for the moderate and severe stroke scenarios, involving greater delays, the results of the proposed method were superior to those of the clinical method. For the mild-stroke scenario, with a shorter delay, the results were more varied.

The results of the SVD method showed that for almost all stroke scenarios, the proposed s4DSF increased the contrast correctness of the rCBF, MTT, and delay maps. The contrast correctness of the rCBV and TTP maps was more varied, with some increases and some decreases for the proposed method. The contrast correctness of the perfusion maps, when using s4DSF, were superior in 73% (33/45) with $P=.001$. Almost all these results were statistically significant; however, their clinical significance remains to be determined. The accuracy of the perfusion maps increased as the noise decreased drastically, leading to more visible and clearly delineated stroke regions. This result was expected since SVD does not

use advanced denoising as a preprocessing step. These results show that adding a noise reduction step improves the performance of SVD and that s4DSF-SVD can depict the contrast between the ischemic and contralateral sides more accurately than standalone SVD can.

The fact that the proposed s4DSF improved the results when combined with both the Bayesian estimation and SVD methods, which have very different methodologies, indicates the potential value of the s4DSF independent of perfusion estimation methods or specific pre- or postprocessing algorithms used.

The patient data could not be analyzed quantitatively since there was no reference standard available for these data. Nevertheless, the TTP, delay, and MTT maps show a greater distinction between healthy and ischemic regions and better delineations when the proposed method is used. The patient cases presented in Appendix D show that, overall, the proposed method enhances the visibility of a stroke. However, in one case, an asymmetry was observed in the perfusion maps generated with the proposed method, which could result in a false positive stroke detection, as it did not correspond with clinical symptoms. We believe that this is due to the algorithm being very sensitive to differences in the time of enhancement. Therefore, the s4DSF and color scaling for this specific algorithm must be further optimized, especially since studies have shown that color scaling is very important for image interpretation [128,129]. The fact that the patient data could only be analyzed visually is a limitation, although the observed results for the patient data agrees with those obtained with the phantom data and could be an indication that the newly proposed s4DSF is superior, as shown quantitatively with the phantom data. The actual impact of the method on diagnostic performance remains to be seen, which will require a multireader, multicase observer study for performance.

This work has some limitations. First, phantom data are only a simulation of real stroke data. Simplifications made in generating these data include simulating a 'homogeneous' stroke, whereas real strokes often consist of a core and a penumbra with different perfusion characteristics. Nonetheless, these simulations did allow for quantitative analysis. Second, the binomial test used assumes independent samples; however, the 45 samples, i.e., perfusion maps, used are from nine different stroke scenarios each resulting in five different perfusion maps. Therefore, the five maps from the same stroke scenario are neither fully independent nor fully dependent. Third, the inevitable lack of a reference standard for the patient data makes quantitative analysis of these results challenging and limited analysis to visual assessment only. However, the fact that the patient data yielded results

similar to those of the phantom study illustrated the potential of the s4DSF. Fourth, the reconstruction method used was conventional filtered back projection; thus, the noise content in the image could be lower if a more sophisticated reconstruction method had been used. While this may affect the image quality of the perfusion maps, it is not likely to affect the comparison between the clinical and proposed implementation of the 4DSF.

In summary, this work proposes a new implementation of the 4DSF optimized for stroke imaging. This results in more accurate perfusion maps in most cases. Before clinical implementation, a diagnostic performance study should be performed to confirm these early findings.

Supplementary Material

Appendix A – Data

Phantom details

In this work the digital anthropomorphic brain phantom and perfusion model developed by Divil et al. [127] was used. In this brain phantom, the vessels, gray matter, and white matter have been divided into multiple smaller regions such that local perfusion differences can be simulated, see Figure A1.

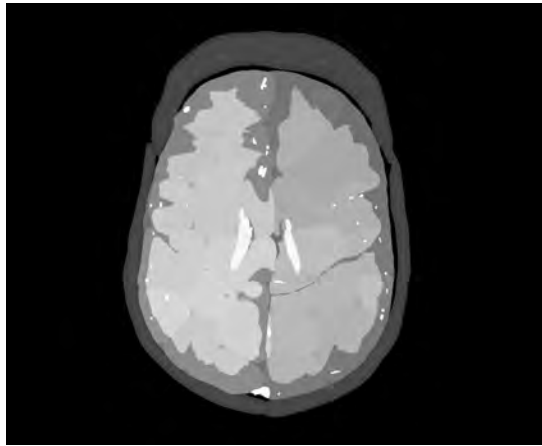


Figure A1: Brain phantom developed by Divil et al., the gray and white matter is subdivided into smaller parts, so the perfusion in each part can be different. The intensities only indicate different tissues with different perfusion characteristics, so have no physical meaning.

The amount of iodinated contrast, the concentration, and the fraction of the iodinated contrast flowing to the brain are all model inputs. It can also be indicated which artery is occluded and what the tissue parameters are for healthy and ischemic gray and white matter. These tissue parameters determine the blood flow, volume, and meant transit time of the tissue.

All gray and white matter sections that are supplied by the occluded artery will have the ischemic tissue parameters and thus will have a lower blood flow, blood volume, and mean transit time. In a real brain, the brain regions distal to the occlusion receive blood via collaterals, however in this model the occlusion is simulated as a partial occlusion.

The total iodine concentration over time in each part of the phantom can be determined using backward flow dynamics, since the blood volume, flow, and

mean transit time of each brain section and total contrast input is known. For the mathematical details, we refer the reader to the work of Divel et al. [127]. Figure A2 shows the iodine concentration over time in the ischemic white matter and in a healthy white matter brain section, for a stroke with a peak delay of 5 seconds and a peak attenuation of a factor 0.4 compared to healthy tissue.

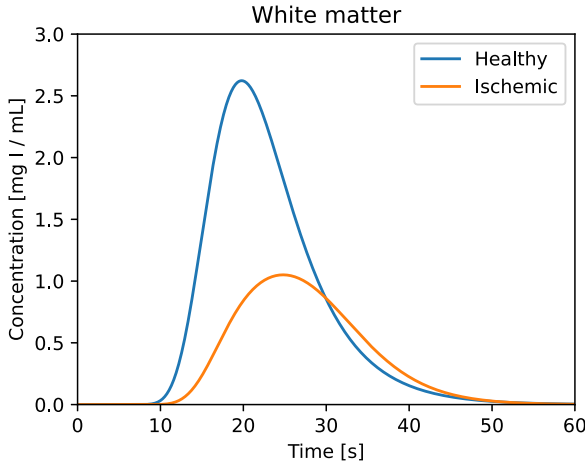


Figure A2: Iodine concentration over time for a healthy and an ischemic section of white matter in the brain phantom.

Phantom data

A previously developed CT simulation tool was used to generate a CT perfusion protocol [114]. The simulated CTP protocol consisted of 30 volume scans over 60 seconds with 2 second intervals. Each scan was acquired with a tube voltage of 80 kV and a tube current-time product of 100 mAs per scan, resulting in a $CTDI_{vol}$ of 3.3 mGy per scan. The reconstruction method used was filtered backprojection [68,116], with a field of view of 220 mm and 320 slices with a thickness of 0.5 mm. The simulations also provided access to data without noise, i.e., reference data.

Patient data

The CT perfusion protocol consisted of 19 scans, the timesteps between the first and second scan was approximately 5 seconds, followed by 12 timesteps of 2.1 seconds and lastly 5 timesteps of 5 seconds again. The tube voltage was 80 kV and the exposure of the scans was varied, the first scan was performed at 200 mAs, followed by 13 scans at 100 mAs, and lastly 5 scans at 75 mAs. The total $CTDI_{vol}$ of the protocol was 67.4 mGy. All CT scans in both protocols were reconstructed with the clinically-available filtered backprojection.

Appendix B - 4D Similarity Filter (4DSF)

The 4DSF filters each voxel in the image separately and searches for the best candidates to average this voxel with, so that the noise in the image is reduced. The 4DSF has been previously described in detail [55,130], so we only summarize it here, highlighting the overall process but focusing on the changes made for the cerebral CT perfusion application presented here.

Each timepoint on the time attenuation curve (TAC) of each voxel is analyzed separately. When searching for TACs similar to those of a specific voxel, the values for the specific timepoint being filtered is left out from the comparison. In this way, the similarity between the TAC and the identified candidate curves is not influenced by the noise present in this specific timepoint.

Because of limited time and computational resources, it is not possible to compare each voxel at each timepoint to all other voxels in the CT perfusion sequence. Thus, only a subset of voxels is compared. In the original implementation, this subset is based on the temporal mean of the TAC. For this, the entire TAC of each voxel is averaged over time and the resulting average values for all voxels are sorted in increasing order. Then, when processing each voxel its TAC is compared only to those voxels that have similar temporal mean.

However, in cerebral CT perfusion the tissue density and enhancement difference between healthy and ischemic gray or white matter is relatively small compared to the noise content of the CT scans. Therefore, the temporal mean is not the optimal metric to discriminate between ischemic and healthy tissue. In the proposed implementation of the 4DSF, the stroke 4DSF (s4DSF), the voxels are not sorted based on their temporal mean but based on their peak time.

With this modification, the algorithm of the s4DSF is as follows (Figure B1):

1. Initial noise reduction of image within each timepoint: An averaging filter with a 7x7x7 kernel is applied to each individual CT image in the perfusion sequence.
2. Determination of peak enhancement time for each voxel: The time values of the three highest-enhancement points on the TAC of each voxel are determined and averaged together, resulting in an estimate of the peak enhancement time for each voxel.
3. Creation of peak-time image: These resulting peak enhancement time values for all voxels are represented as one 3D image, denoted the peak-time image.

4. Noise reduction of peak-time image: The peak-time image is downsampled by a factor of four in each direction, filtered with an averaging filter with a $5 \times 5 \times 5$ kernel, and then upsampled back by a factor of four in each direction using trilinear upscaling.
5. Image mask to exclude air and bone: A mask is generated from the first, pre-contrast, scan by first reducing its noise (averaging filter, $3 \times 3 \times 3$ kernel) and thresholding all voxels between -300 HU and 300 HU to unity and all others to zero.
6. Masking of peak-time image: The image mask is applied to the peak-time image, resulting in an image with only soft tissue (brain, skin, etc.) remaining in the image, denoted masked-peak-time image.
7. Voxel sorting based on peak time: All voxels of the masked-peak-time image are sorted in increasing order of peak enhancement time. Therefore, this 4D image is converted into a long 1D vector of TACs, denoted the peak-time-sorted vector.

Each timepoint of the TAC of each voxel is filtered separately from here on. The voxel and timepoint being filtered are denoted the current voxel and current timepoint, respectively:

8. Identification of candidate voxels with similar peak time: The subset of 2000 voxels closest to the current voxel in the peak-time-sorted vector is selected. These identified voxels are denoted the candidate voxels.
9. Comparison of TACs between current and candidate voxels: The mean-square-error (MSE) between the TACs of each candidate voxel and the current voxel are computed. The current timepoint is excluded from this MSE calculation, so the noise present in the current timepoint does not influence the MSE.
10. Identification of similar voxels: If the MSE is below a certain threshold and the current voxel and the candidate voxels are spatially close enough to each other, then the candidate voxel is deemed a similar voxel. This is a voxel that has been identified as having a similar enough underlying perfusion process that it can be used to filter the current voxel at the current timepoint. This similar voxel is placed in a vector, denoted the MSE vector, which stores all the identified similar voxels, sorted by increasing MSE compared to the current voxel.
11. Averaging of 100 most-similar voxel values: Once all candidate voxels are checked and the similar ones are identified, the values of the current timepoint of the up to 100 (set by the filter strength) similar voxel values with the lowest MSE (i.e. the first up to 100 voxels in the MSE vector) are averaged.

If the MSE vector contains fewer than 100 voxels, then all voxels are used. The values used to average are the values of the original image, i.e., without the initial noise reduction applied in Step 1, above.

12. Setting of filtered value: The value of the current timepoint of the current voxel is set to this average value.
13. Steps 8 to 12 are repeated for every timepoint of every voxel.

Note that only Steps 2 to 4 and the spatial constraint of Step 10 are different from the standard implementation of the 4DSF.

The exact implementation of the s4DSF used in this work is a GPU-based implementation of which the details are described in the work of Mikerov et al [131]. The number of candidates checked for each voxel is 2000, the filter strength is 100, the mean squared error threshold is set to 900, and the hash length is 128. The number of candidates checked for each voxel is set relatively low for computational time purposes, since the clinical application is stroke imaging. A maximum radius of 50 voxels is used for the distance constraint. This distance constraint reduces computation time, which is very valuable for stroke imaging. It is important to note that the mean squared error between the TACs is determined after average filtering each CT image in the perfusion sequence with a 7x7x7 kernel. However, the values obtained at the timepoint being processed, to determine the running average, are of the original image.

Processing Time

The proposed s4DSF needs approximately 290 seconds, on a GPU (NVIDIA RTX A6000), to process the full 30 scan protocol. This time is specific to our protocol, and it scales approximately linearly with the number of scans in the protocol.

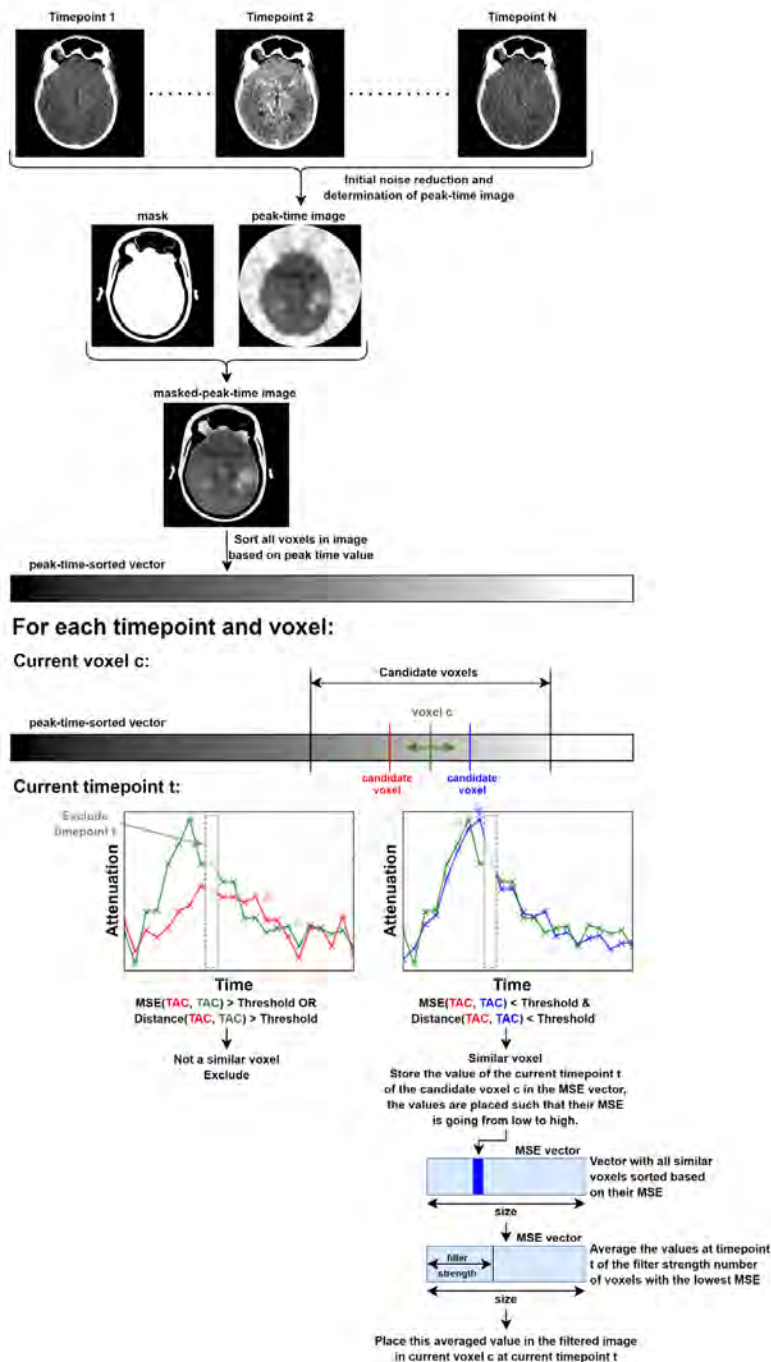


Figure B1: Schematic overview of 4DSF of ischemic stroke. Adapted from S. A. M. Tunissen et al., "Performance evaluation of a 4D similarity filter for dynamic CT angiography imaging of the liver", Med. Phys., doi.org/10.1002/mp.17394

Appendix C – Additional Results

In the main text the perfusion maps of one of the nine stroke scenarios are shown and the results of the analysis are shown for three of the nine stroke scenarios. The perfusion maps of the eight remaining scenarios and the analysis results of the six remaining scenarios are presented in this appendix, for both the Bayesian estimation and Singular Value Decomposition (SVD) method. The results are shown for the situation with (proposed) and without (clinical) use of the proposed s4DSF, and the reference situation (noiseless phantom data).

The eight scenarios of which the perfusion maps are presented are the following: a peak delay of 1.5 seconds with a peak attenuation of 0.2, 0.4, and 0.6 relative to the healthy perfusion curve, a peak delay of 3.0 seconds with a peak attenuation of 0.2, and 0.6, and a peak delay of 5.0 seconds with a peak attenuation 0.2, 0.4, and 0.6.

The six scenarios of which the analysis results are presented are the following: a peak delay of 1.5 seconds with a peak attenuation of 0.2 and 0.6 relative to the healthy perfusion curve, a peak delay of 3.0 seconds with a peak attenuation of 0.2 and 0.6, and a peak delay of 5.0 seconds with a peak attenuation 0.4 and 0.6.

Bayesian estimation

In this appendix the perfusion maps and analysis results for the Bayesian estimation method are listed. The perfusion maps of the eight scenarios are presented in Figures C1-C8.

Tables C1 and C2 present the results of the analysis of the six remaining stroke scenarios. The values of the ischemic and healthy Region of Interests (ROIs) are significantly different for all maps in Tables C1 and C2 according to the Mann-Whitney U test. The values of the clinical and proposed methods are significantly different for all maps in Tables C1 and C2 according to the Wilcoxon signed-rank test. In Tables C3 and C4 the RMSE between reference and the clinical or proposed are given.

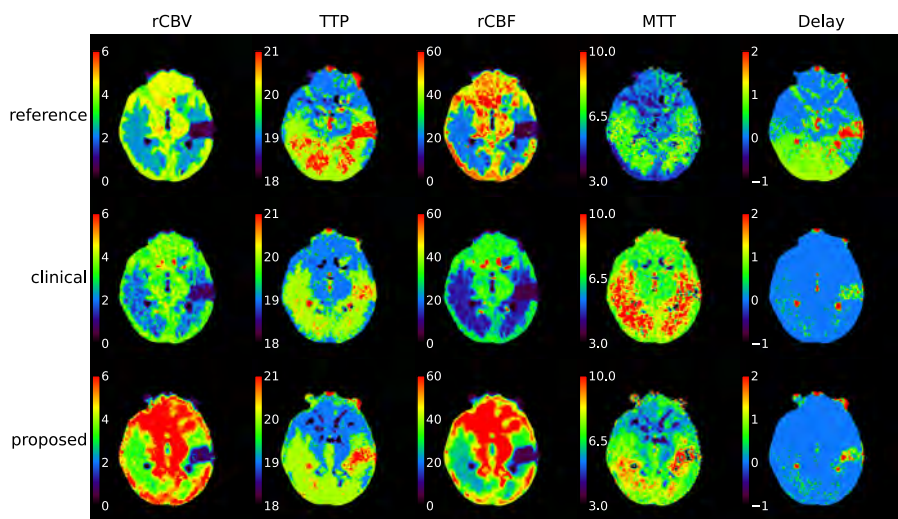


Figure C1: The perfusion maps of the Bayesian estimation method of stroke scenario with a peak delay of 1.5 seconds and a peak attenuation of 0.2.

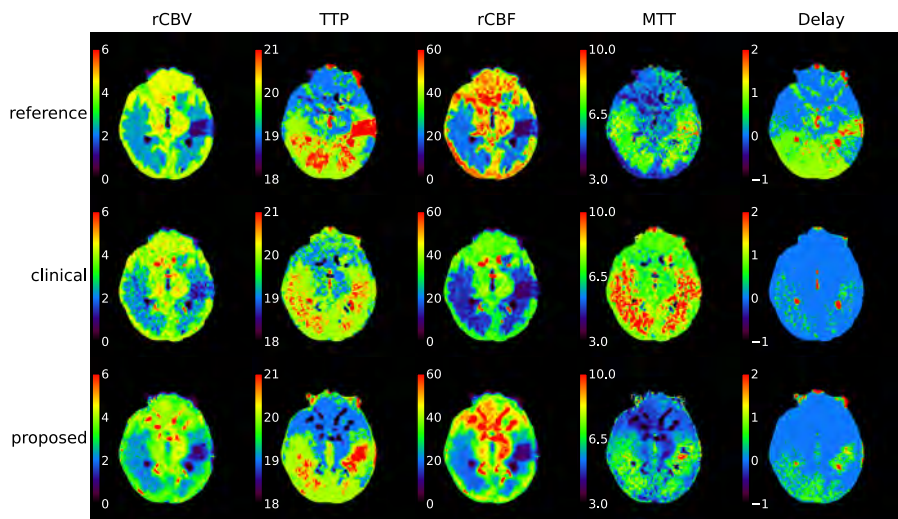


Figure C2: The perfusion maps of the Bayesian estimation method of stroke scenario with a peak delay of 1.5 seconds and a peak attenuation of 0.4.

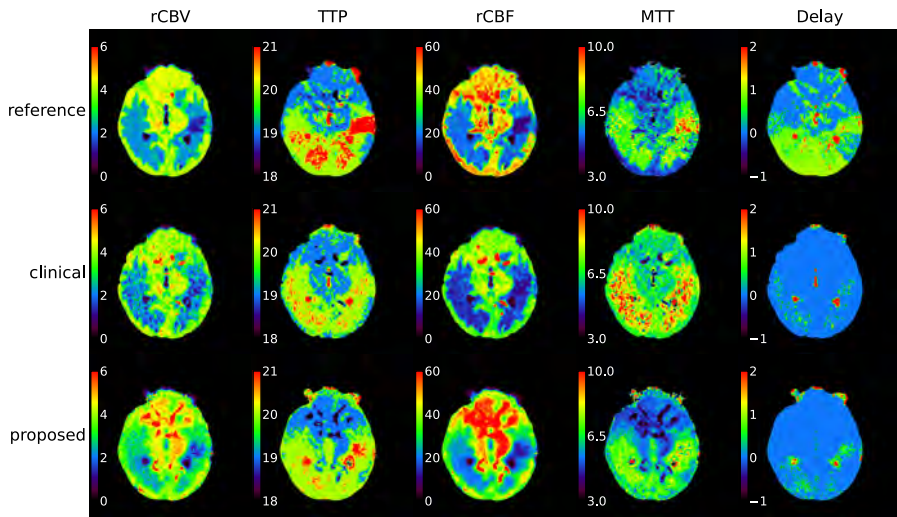


Figure C3: The perfusion maps of the Bayesian estimation method of stroke scenario with a peak delay of 1.5 seconds and a peak attenuation of 0.6.

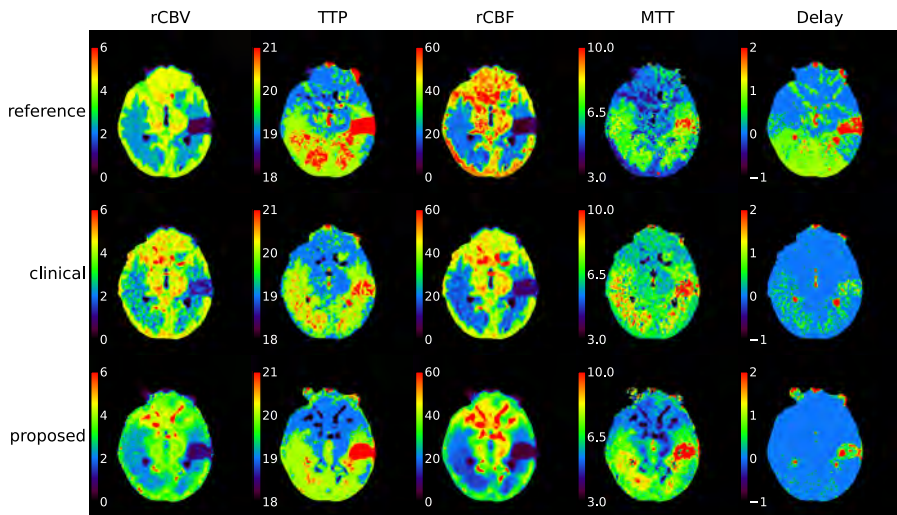


Figure C4: The perfusion maps of the Bayesian estimation method of stroke scenario with a peak delay of 3 seconds and a peak attenuation of 0.2.

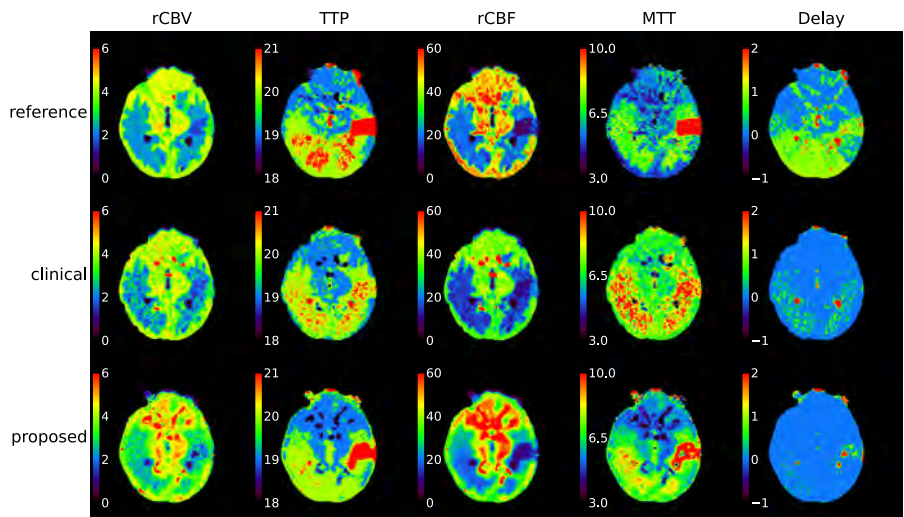


Figure C5: The perfusion maps of the Bayesian estimation method of stroke scenario with a peak delay of 3 seconds and a peak attenuation of 0.6.

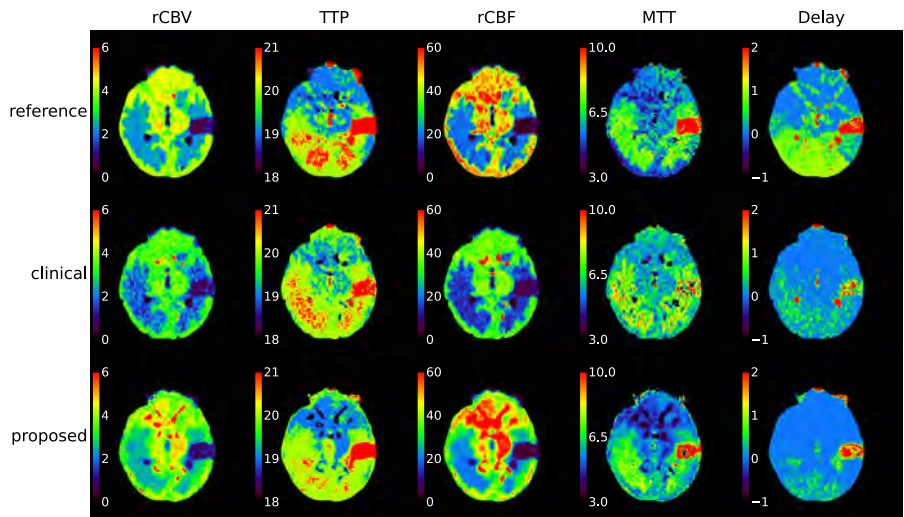


Figure C6: The perfusion maps of the Bayesian estimation method of stroke scenario with a peak delay of 5 seconds and a peak attenuation of 0.2.

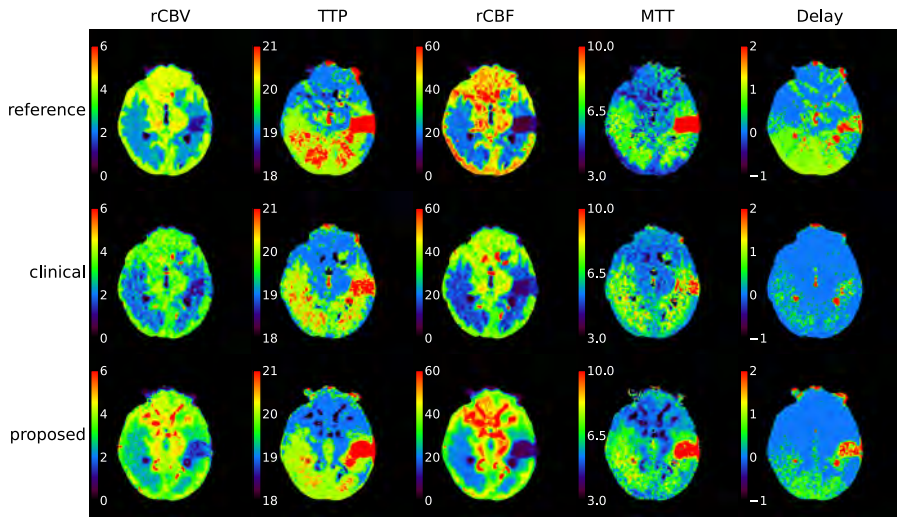


Figure C7: The perfusion maps of the Bayesian estimation method of stroke scenario with a peak delay of 5 seconds and a peak attenuation of 0.4.

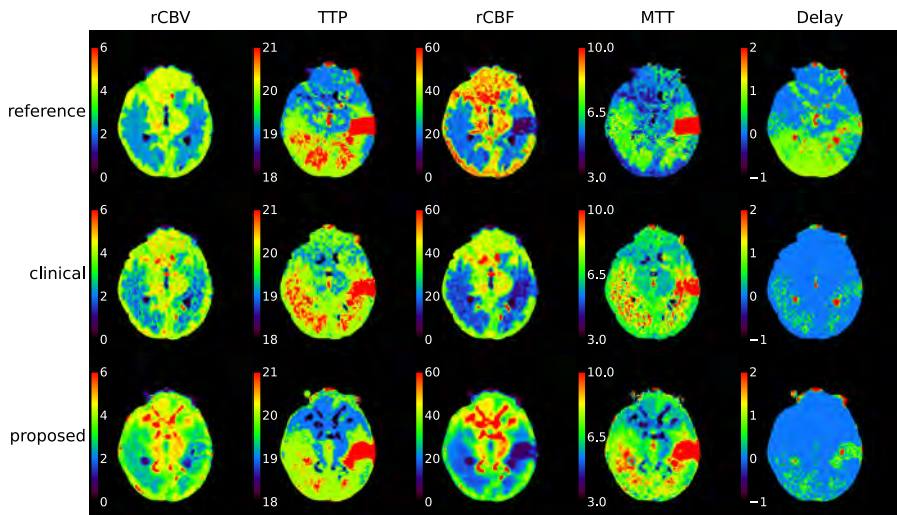


Figure C8: The perfusion maps of the Bayesian estimation method of stroke scenario with a peak delay of 5 seconds and a peak attenuation of 0.6.

Table C1: Mean values of reference, clinical, and proposed method for the ischemic and healthy ROI, and the contrast between them, when using the Bayesian estimation method. The contrast values closest to the reference have a gray background.

		Peak delay = 1.5				Peak delay = 1.5				Peak delay = 3.0			
		Peak attenuation = 0.2				Peak attenuation = 0.4				Peak attenuation = 0.2			
	Reference	ischemic	healthy	contrast	ischemic	healthy	contrast	ischemic	healthy	contrast	ischemic	healthy	contrast
rCBV (mL/100g)		0.71	3.08	2.37	1.27	3.08	1.81	0.80	3.08	2.28			
	Clinical	0.89	2.56	1.67	1.53	2.94	1.40	1.29	3.08	1.80			
	Proposed	1.64	4.17	2.53	1.57	2.83	1.26	1.20	2.74	1.55			
TTP (s)	Reference	20.90	19.50	-1.40	20.93	19.50	-1.43	22.01	19.50	-2.51			
	Clinical	20.41	19.25	-1.16	20.30	19.72	-0.58	20.77	19.36	-1.41			
	Proposed	20.48	19.67	-0.81	20.53	19.71	-0.83	21.30	19.63	-1.67			
rCBF (mL/100g/min)	Reference	7.96	29.78	21.82	11.89	29.83	17.94	6.85	29.66	22.81			
	Clinical	6.29	19.14	12.84	10.29	22.29	11.99	9.08	27.31	18.23			
	Proposed	12.42	35.70	23.28	14.23	28.05	13.82	8.02	25.10	17.07			
MTT (s)	Reference	6.38	6.55	0.18	7.34	6.55	-0.79	8.68	6.57	-2.10			
	Clinical	7.84	8.16	0.32	8.71	8.09	-0.62	8.26	6.88	-1.38			
	Proposed	8.19	7.04	-1.15	6.72	6.08	-0.64	10.13	6.59	-3.54			
Delay (s)	Reference	1.22	-0.17	-1.39	0.95	-0.17	-1.12	1.56	-0.18	-1.74			
	Clinical	0.76	-1.10	-1.86	0.12	-0.60	-0.72	0.75	-0.95	-1.70			
	Proposed	0.47	-0.34	-0.80	0.42	-0.32	-0.74	0.68	-0.35	-1.03			

Table C2: Mean values of reference, clinical, and proposed method for the ischemic and healthy ROI, and the contrast between them, when using the Bayesian estimation method. The contrast values closest to the reference have a gray background.

	Peak delay = 3.0						Peak delay = 5.0					
	Peak attenuation = 0.6			Peak attenuation = 0.4			Peak attenuation = 0.6			Peak attenuation = 0.6		
	ischemic	healthy	contrast	ischemic	healthy	contrast	ischemic	healthy	contrast	ischemic	healthy	contrast
rCBV (mL/100g)	Reference	2.14	3.03	0.89	1.68	3.03	1.35	2.49	3.03	3.03	0.54	
	Clinical	2.16	2.83	0.66	1.51	2.52	1.01	2.44	3.05		0.61	
	Proposed	2.39	3.24	0.86	1.79	3.13	1.34	2.66	3.30		0.65	
TTP (s)	Reference	22.31	19.49	-2.82	23.77	19.50	-4.28	23.94	19.49	19.49	-4.45	
	Clinical	20.41	19.46	-0.95	21.00	19.22	-1.78	21.24	19.84	19.84	-1.40	
	Proposed	21.13	19.62	-1.51	22.53	19.62	-2.91	22.38	19.78	19.78	-2.60	
rCBF (mL/100g/min)	Reference	11.31	29.67	18.36	7.29	29.26	21.98	9.97	29.47	29.47	19.50	
	Clinical	14.05	22.63	8.58	10.74	23.93	13.19	14.03	25.58	25.58	11.55	
	Proposed	14.90	29.66	14.76	10.58	28.86	18.28	11.75	26.80	26.80	15.05	
MTT (s)	Reference	12.17	6.53	-5.63	15.60	6.58	-9.03	16.61	6.54	6.54	-10.07	
	Clinical	9.28	7.62	-1.66	8.29	6.39	-1.90	10.38	7.33	7.33	-3.06	
	Proposed	9.91	6.61	-3.30	10.91	6.53	-4.38	14.45	7.42	7.42	-7.03	
Delay (s)	Reference	0.83	-0.12	-0.95	1.06	-0.15	-1.21	0.74	-0.14	-0.14	-0.88	
	Clinical	0.20	-0.81	-1.01	0.51	-1.31	-1.83	0.26	-0.61	-0.61	-0.87	
	Proposed	0.28	-0.34	-0.62	1.22	-0.28	-1.51	0.52	-0.33	-0.33	-0.84	

Table C3: RMSE of clinical and proposed with the reference when using the Bayesian estimation method. The lowest RMSE has a gray background.

		Peak delay = 1.5		Peak delay = 1.5		Peak delay = 3.0	
		Peak attenuation = 0.2		Peak attenuation = 0.4		Peak attenuation = 0.2	
		ischemic	healthy	ischemic	healthy	ischemic	healthy
rCBV (mL/100g)	Clinical	0.50	0.79	0.55	0.50	0.72	0.63
	Proposed	1.25	1.32	0.58	0.69	0.64	0.77
TTP (s)	Clinical	1.08	2.44	1.31	1.60	1.67	2.32
	Proposed	1.10	0.54	0.79	0.59	1.41	0.54
rCBF (mL/100g/min)	Clinical	10.76	13.06	10.42	9.62	10.73	7.40
	Proposed	12.64	10.59	10.37	8.79	10.26	10.30
MTT (s)	Clinical	3.26	2.11	2.44	1.95	2.69	1.18
	Proposed	2.89	1.01	1.65	0.93	3.35	0.95
Delay (s)	Clinical	1.20	6.37	2.67	4.01	1.51	6.01
	Proposed	1.25	0.81	0.89	0.91	1.50	0.71

Table C4: RMSE of clinical and proposed with the reference when using the Bayesian estimation method. The lowest RMSE has a gray background.

		Peak delay = 3.0		Peak delay = 5.0		Peak delay = 5.0	
		ischemic	healthy	ischemic	healthy	ischemic	healthy
rCBV (mL/100g)	Clinical	0.44	0.59	0.46	0.82	0.45	0.49
	Proposed	0.69	0.71	0.51	0.64	0.70	0.72
TTP (s)	Clinical	2.01	2.06	2.96	2.81	2.80	1.70
	Proposed	1.41	0.55	1.86	0.52	1.89	0.58
rCBF (mL/100g/min)	Clinical	3.97	9.63	4.57	8.98	5.74	6.91
	Proposed	5.51	8.75	5.68	8.30	5.43	9.14
MTT (s)	Clinical	3.76	1.67	8.04	1.23	6.88	1.34
	Proposed	3.47	0.94	5.81	0.91	3.95	1.33
Delay (s)	Clinical	0.81	5.34	1.10	7.32	0.80	4.18
	Proposed	0.84	0.87	1.07	0.75	0.73	0.68

SVD

In this appendix the perfusion maps and analysis results for the SVD method are listed. The perfusion maps of the eight scenarios are presented in Figures C9-C16.

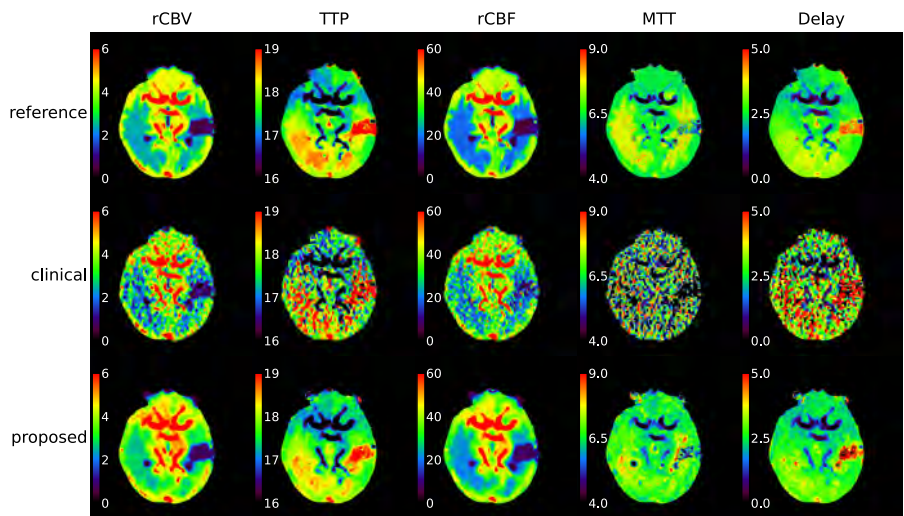


Figure C9: The perfusion maps of the SVD method of stroke scenario with a peak delay of 1.5 seconds and a peak attenuation of 0.2.

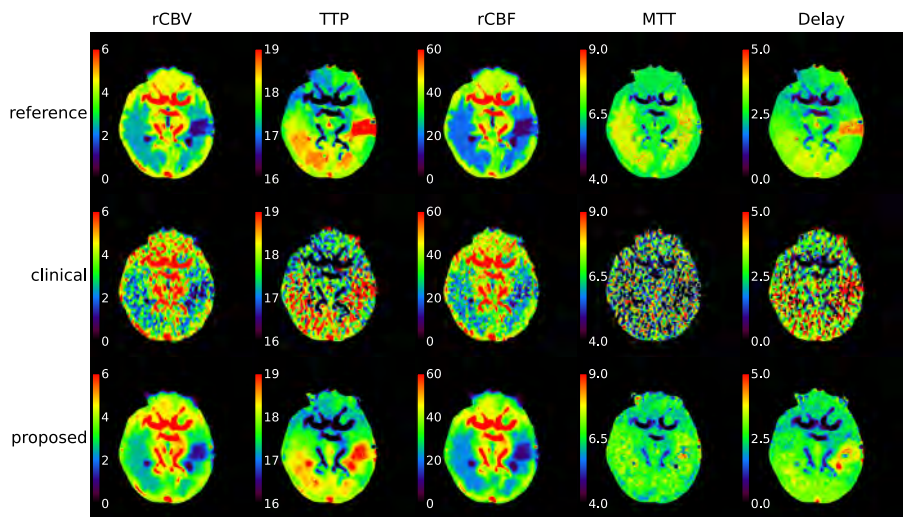


Figure C10: The perfusion maps of the SVD method of stroke scenario with a peak delay of 1.5 seconds and a peak attenuation of 0.4.

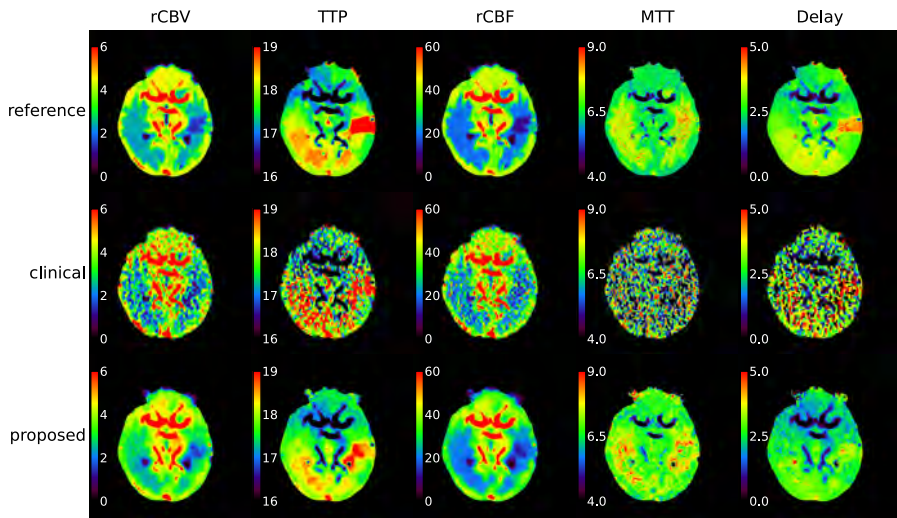


Figure C11: The perfusion maps of the SVD method of stroke scenario with a peak delay of 1.5 seconds and a peak attenuation of 0.6.

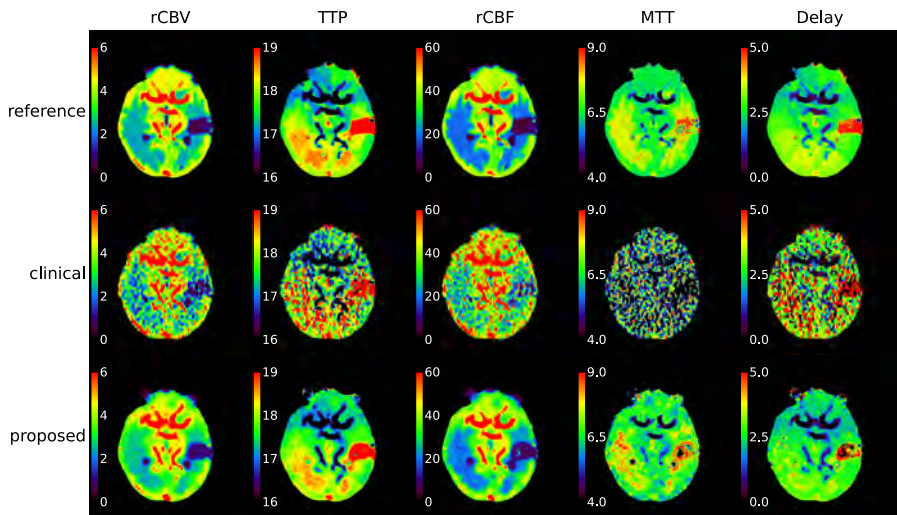


Figure C12: The perfusion maps of the SVD method of stroke scenario with a peak delay of 3 seconds and a peak attenuation of 0.2.

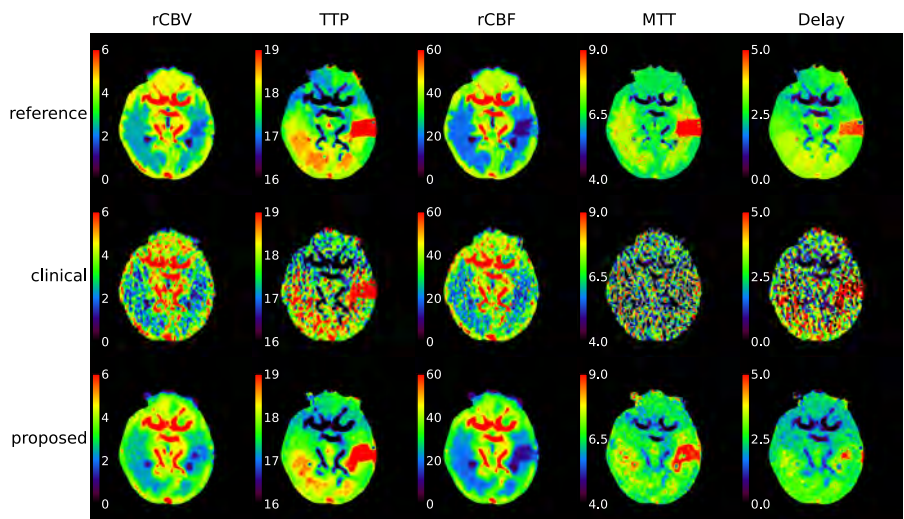


Figure C13: The perfusion maps of the SVD method of stroke scenario with a peak delay of 3 seconds and a peak attenuation of 0.6.

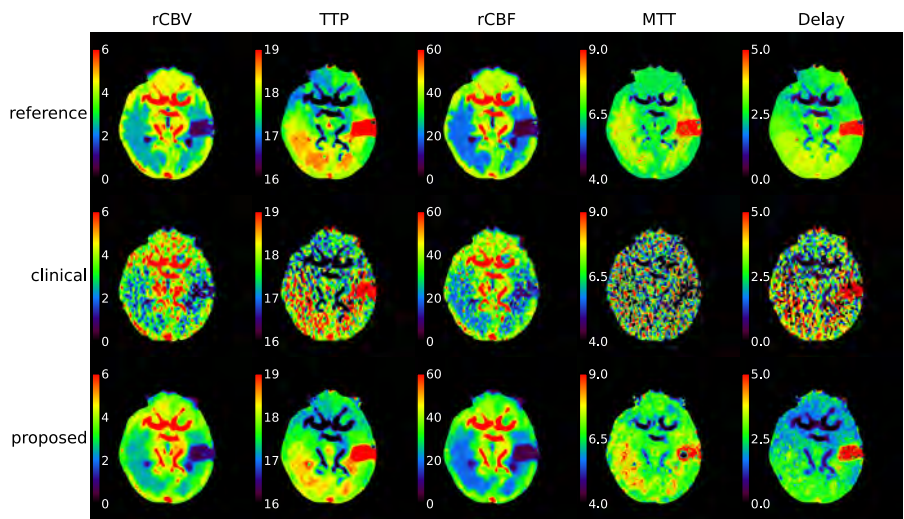


Figure C14: The perfusion maps of the SVD method of stroke scenario with a peak delay of 5 seconds and a peak attenuation of 0.2.

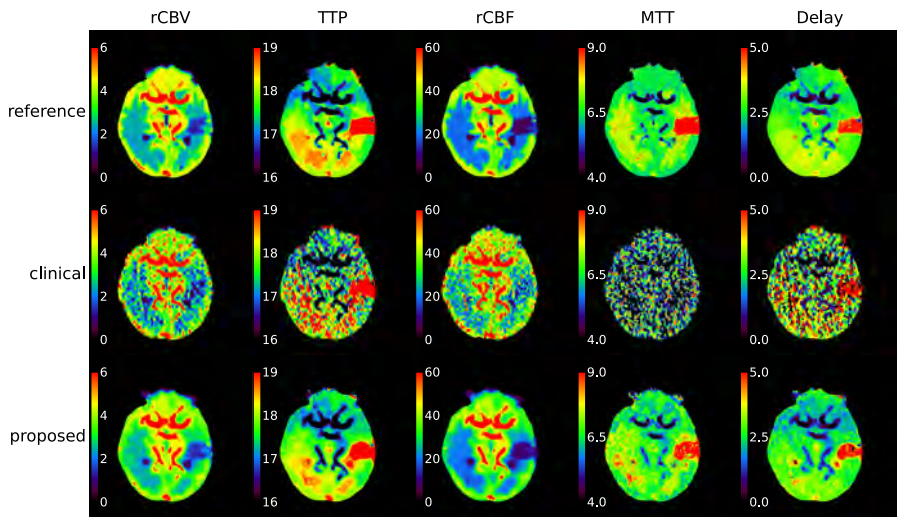


Figure C15: The perfusion maps of the SVD method of stroke scenario with a peak delay of 5 seconds and a peak attenuation of 0.4.

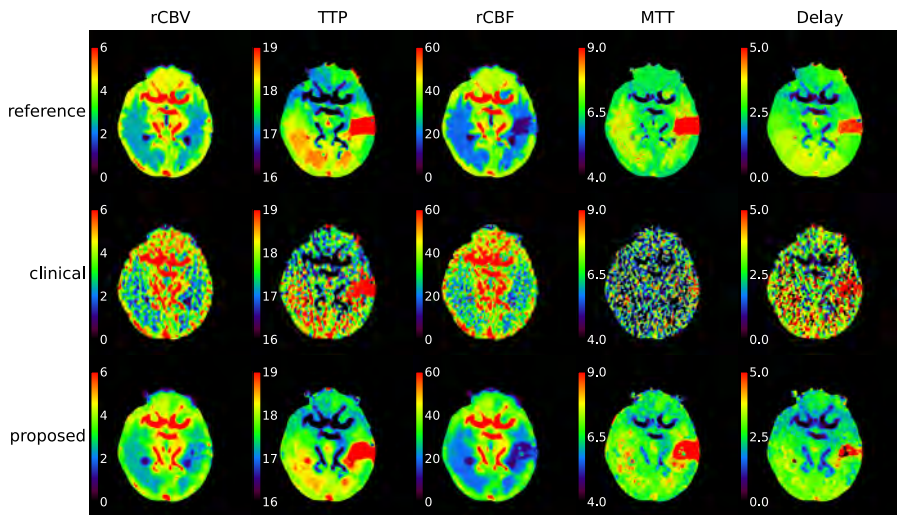


Figure C16: The perfusion maps of the SVD method of stroke scenario with a peak delay of 5 seconds and a peak attenuation of 0.6.

Tables C5 and C6 present the results of the analysis of the six remaining stroke scenarios. The values of the ischemic and healthy ROIs are significantly different for all maps in Tables C5 and C6 according to the Mann-Whitney U test. The values of the clinical and proposed methods are significantly different for all maps in Tables C5 and C6 except two, as indicated in Table C6, according to the Wilcoxon signed-rank test. In Tables C7 and C8 the RMSE between reference and the clinical or proposed are given.

Table C5: Mean values of reference, clinical, and proposed method for the ischemic and healthy ROI, and the contrast between them, when using the SVD method. The contrast values closest to the reference have a gray background.

	Peak delay = 1.5				Peak delay = 1.5				Peak delay = 3.0			
	Peak attenuation = 0.2				Peak attenuation = 0.4				Peak attenuation = 0.2			
	ischemic	healthy	contrast		ischemic	healthy	contrast		ischemic	healthy	contrast	
rCBV (mL/100g)	Reference	0.90	3.20	2.30	1.45	3.20	1.75	0.99	3.20	3.21	2.21	
	Clinical	1.10	2.92	1.82	1.70	3.37	1.67	1.29	3.21		1.92	
	Proposed	1.34	3.29	1.95	1.88	3.29	1.41	1.42	3.17		1.75	
TTP (s)	Reference	18.46	17.68	-0.78	18.71	17.68	-1.03	19.41	17.68	17.68	-1.73	
	Clinical	18.94	17.36	-1.58	18.80	17.58	-1.23	20.24	17.48		-2.76	
	Proposed	18.45	17.68	-0.77	18.46	17.70	-0.76	19.40	17.68		-1.72	
rCBF (mL/100g/min)	Reference	8.71	27.65	18.94	12.54	27.65	15.11	8.10	27.66	27.66	19.56	
	Clinical	17.55	29.06	11.51	21.22	32.94	11.72	22.27	34.26		11.99	
	Proposed	12.23	29.27	17.03	16.78	28.72	11.94	11.55	27.05		15.50	
MTT (s)	Reference	6.55	7.04	0.50	7.21	7.05	-0.16	7.88	7.04	7.04	-0.84	
	Clinical	3.95	6.08	2.13	4.92	6.20	1.28	3.68	5.74		2.05	
	Proposed	6.60	6.73	0.13	6.81	6.86	0.05	7.86	7.11		-0.76	
Delay (s)	Reference	3.60	2.38	-1.23	3.64	2.38	-1.27	4.18	2.38	2.38	-1.80	
	Clinical	1.82	1.12	-0.70	2.41	1.72	-0.69	2.25	1.70		-0.54	
	Proposed	3.25	2.28	-0.96	3.43	2.22	-1.20	1.39	2.33		0.94	

Table C6: Mean values of reference, clinical, and proposed method for the ischemic and healthy ROI, and the contrast between them, when using the SVD method. The contrast values closest to the reference have a gray background.

		Peak delay = 3.0				Peak delay = 5.0				Peak delay = 5.0			
		Peak attenuation = 0.6				Peak attenuation = 0.4				Peak attenuation = 0.6			
		ischemic	healthy	contrast		ischemic	healthy	contrast		ischemic	healthy	contrast	
rCBV (mL/100g)	Reference	2.35	3.19	0.84		1.89	3.19	1.29		2.70	3.19	0.49	
	Clinical	2.50	3.13	0.62		1.89	3.09	1.20		2.83	3.51	0.68	
	Proposed	2.32	3.09	0.78		1.89	3.16	1.27		2.58	3.15	0.58	
TTP (s)	Reference	20.11	17.68	-2.43		21.16	17.68	-3.48		21.50	17.68	-3.82	
	Clinical	20.38	17.52	-2.86		21.44	17.39	-4.05		21.68	17.58	-4.10	
	Proposed	19.47	17.64	-1.82		20.53	17.70	-2.83		20.96	17.67	-3.29	
rCBF (mL/100g/min)	Reference	14.78	27.74	12.95		10.70	27.68	16.99		13.74	27.68	13.94	
	Clinical	23.43	30.50	7.07		23.58	32.44	8.86		25.89	35.77	9.88	
	Proposed	15.99	27.49	11.50		12.45	26.36	13.91		14.72	26.83	12.11	
MTT (s)	Reference	9.84	7.00	-2.85		11.18	7.01	-4.17		12.18	7.02	-5.17	
	Clinical	6.64	6.18	-0.46		5.13	5.73	0.61		6.84	5.99	-0.84	
	Proposed	9.01	6.79	-2.22		9.48	7.15	-2.32		10.94	7.01	-3.93	
Delay (s)	Reference	4.07	2.46	-1.61		4.56	2.46	-2.10		4.42	2.46	-1.96	
	Clinical	3.25	1.32	-1.94		3.98	0.59	-3.39		5.69	1.71	-3.98	
	Proposed	3.13	2.09	-1.05		4.32	2.29	-2.03		3.42	2.35	-1.07	

‡ Maps between which there is no significant difference

Table C7: RMSE of clinical and proposed with the reference when using the SVD method. The lowest RMSE has a gray background.

		Peak delay = 1.5		Peak delay = 1.5		Peak delay = 3.0	
		Peak attenuation = 0.2		Peak attenuation = 0.4		Peak attenuation = 0.2	
		ischemic	healthy	ischemic	healthy	ischemic	healthy
rCBV (mL/100g)	Clinical	0.67	1.01	0.86	1.00	0.79	0.92
	Proposed	0.67	0.49	0.61	0.50	0.64	0.52
TTP (s)	Clinical	4.89	2.36	2.18	1.61	5.11	2.27
	Proposed	0.95	0.36	0.74	0.40	1.32	0.34
rCBF (mL/100g/min)	Clinical	14.58	9.96	13.32	10.67	19.88	13.16
	Proposed	5.71	5.62	6.10	5.47	6.09	5.44
MTT (s)	Clinical	3.57	2.23	3.27	2.03	4.90	2.21
	Proposed	1.64	0.73	1.32	0.63	2.25	0.79
Delay (s)	Clinical	11.30	8.27	9.09	6.33	12.19	8.50
	Proposed	4.00	0.90	1.15	1.02	7.54	0.91

Table C8: RMSE of clinical and proposed with the reference when using the SVD method. The lowest RMSE has a gray background.

		Peak delay = 3.0		Peak delay = 5.0		Peak delay = 5.0	
		Peak attenuation = 0.6		Peak attenuation = 0.4		Peak attenuation = 0.6	
		ischemic	healthy	ischemic	healthy	ischemic	healthy
rCBV (mL/100g)	Clinical	0.82	0.98	0.72	1.06	0.85	1.01
	Proposed	0.53	0.53	0.45	0.48	0.58	0.48
TTP (s)	Clinical	1.79	2.06	2.55	2.70	2.18	1.66
	Proposed	1.10	0.39	1.70	0.35	1.56	0.34
rCBF (mL/100g/min)	Clinical	12.82	9.30	17.78	12.78	15.52	13.16
	Proposed	3.80	5.19	3.97	5.69	4.20	5.63
MTT (s)	Clinical	4.16	2.04	6.70	2.27	6.10	2.10
	Proposed	1.74	0.67	2.67	0.81	2.31	0.78
Delay (s)	Clinical	8.49	7.05	9.67	9.12	10.16	6.73
	Proposed	1.73	1.02	4.16	0.88	4.50	0.80

Appendix D – Additional Patient Cases

In this appendix ten additional consecutive cases are presented of patients that presented at our hospital with suspected stroke. The results of the clinically and proposed Bayesian estimation and Singular Value Decomposition (SVD) methods are presented for each of these cases. The perfusion maps are depicted as generated by the clinical workstation (Vitrea) including its automatic window/leveling and alignment. Therefore, the perfusion maps of the clinical and proposed methods do not have the same window/leveling. All cases were interpreted visually by a radiologist experienced in stroke imaging (10 years of experience reading stroke scans).

Five patient cases, depicted in Figures D1, D2, D5, D6, D13, D14, D17, D18, D19, and D20, show perfusion maps on which no stroke is visible. The perfusion maps of the proposed method are much cleaner, especially for the SVD method.

Three patient cases, depicted in Figures D3, D4, D7, D8, D9, and D10, show perfusion maps on which a stroke is clearly visible. The stroke is better delineated with the proposed method. Figures D8 and D10 show two cases in which the stroke is only visible on one of the five perfusion maps for the clinical method and visible on all five perfusion maps for the proposed method.

Two patient cases, depicted in Figures D11, D12, D15, and D16, show slight asymmetries in some perfusion maps. In the patient case presented in Figures D11 and D12, this asymmetry is only visible in the TTP and MTT map (Figure D11) or TTP and Delay map (Figure D12), making it not clinically relevant. In the patient case presented in Figures D15 and D16 this asymmetry is present in five and four maps, respectively, however the clinical symptoms did not match with the region of asymmetry. Thus, this asymmetry may lead to a false positive stroke detection on the CT perfusion scan.

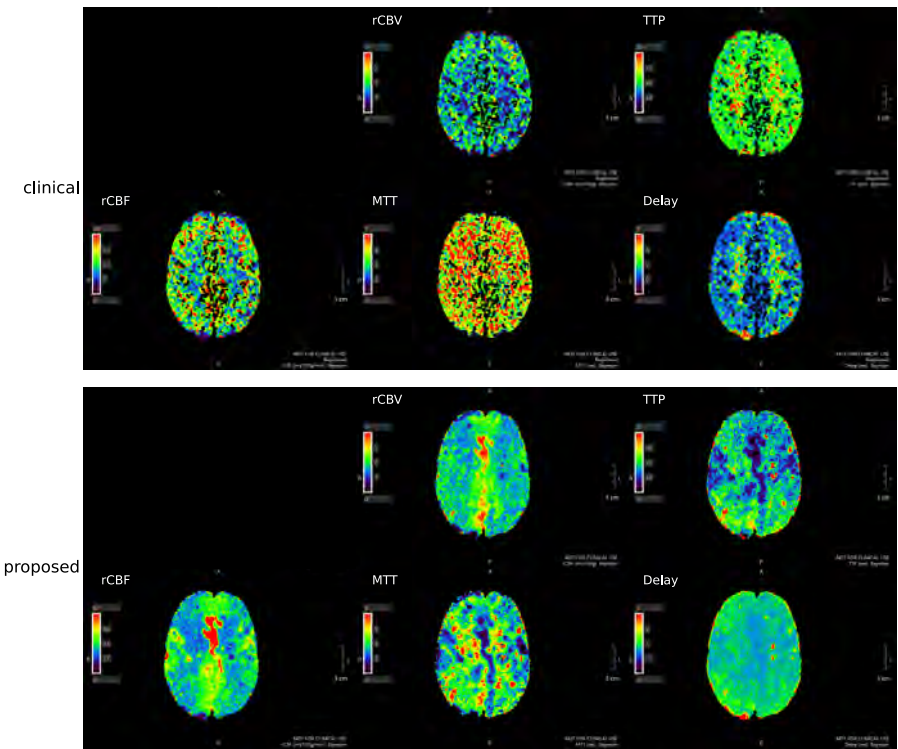


Figure D1: The perfusion maps obtained using Bayesian estimation method. The proposed method results in a much cleaner image. However, there is no stroke visible for this patient.

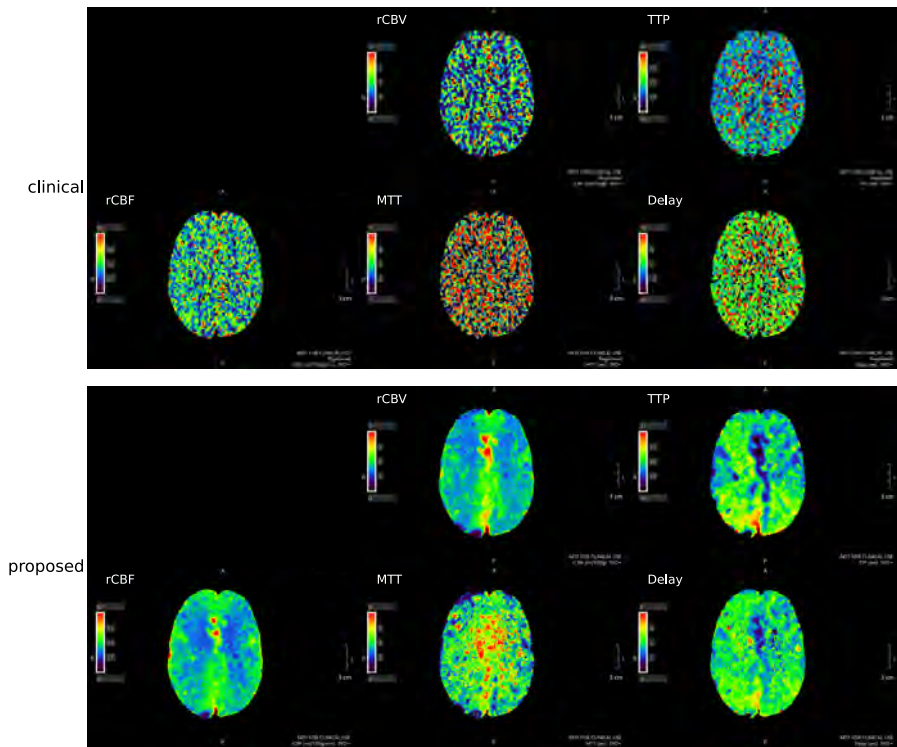


Figure D2: The perfusion maps obtained using the SVD method. The proposed method results in a much cleaner image. However, there is no stroke visible for this patient.

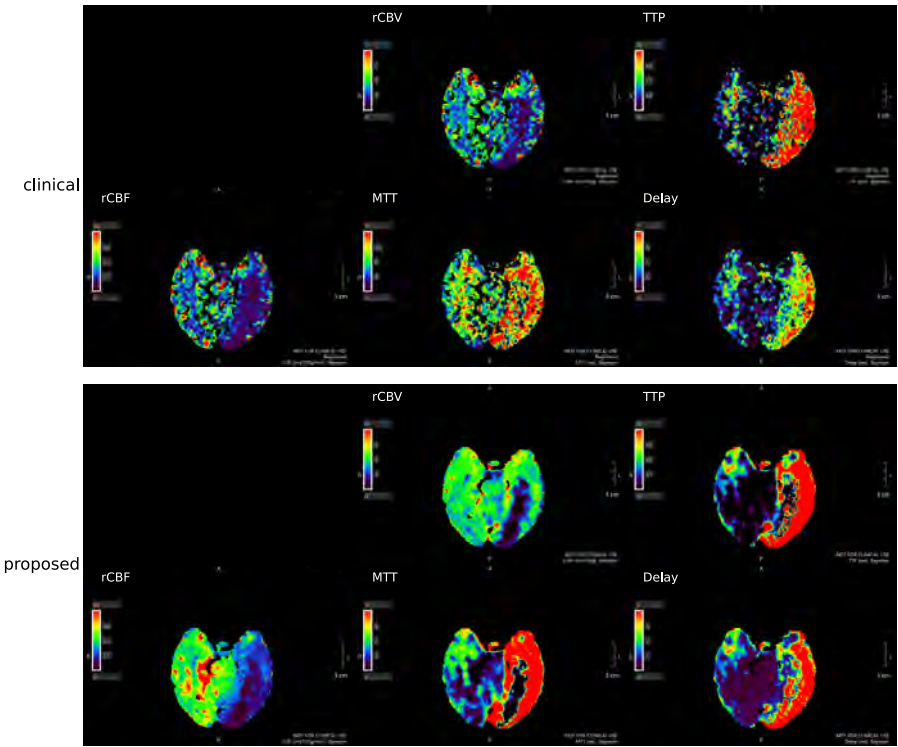


Figure D3: The perfusion maps obtained using Bayesian estimation method. The proposed method delineates the stroke better.

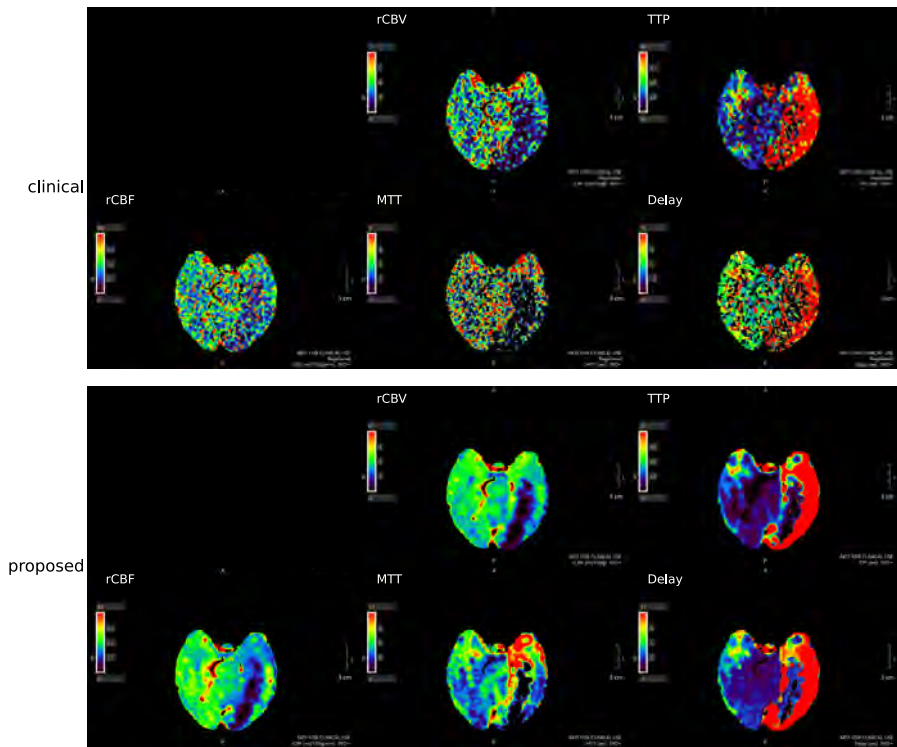


Figure D4: The perfusion maps obtained using the SVD method. The proposed method delineates the stroke better and has much lower noise content.

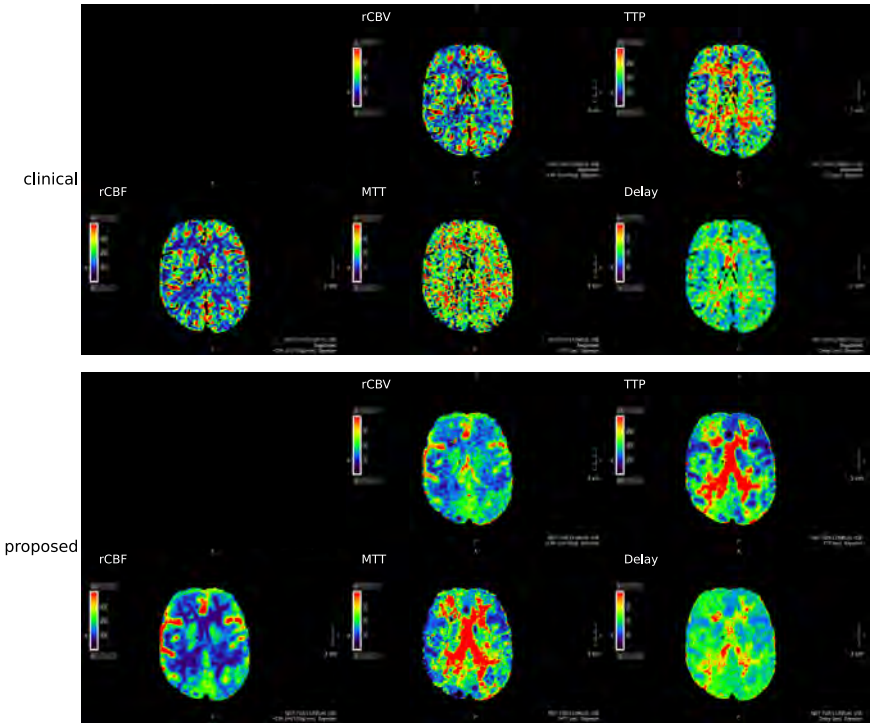


Figure D5: The perfusion maps obtained using Bayesian estimation method. The proposed method results in a much cleaner image. However, there is no stroke visible for this patient.

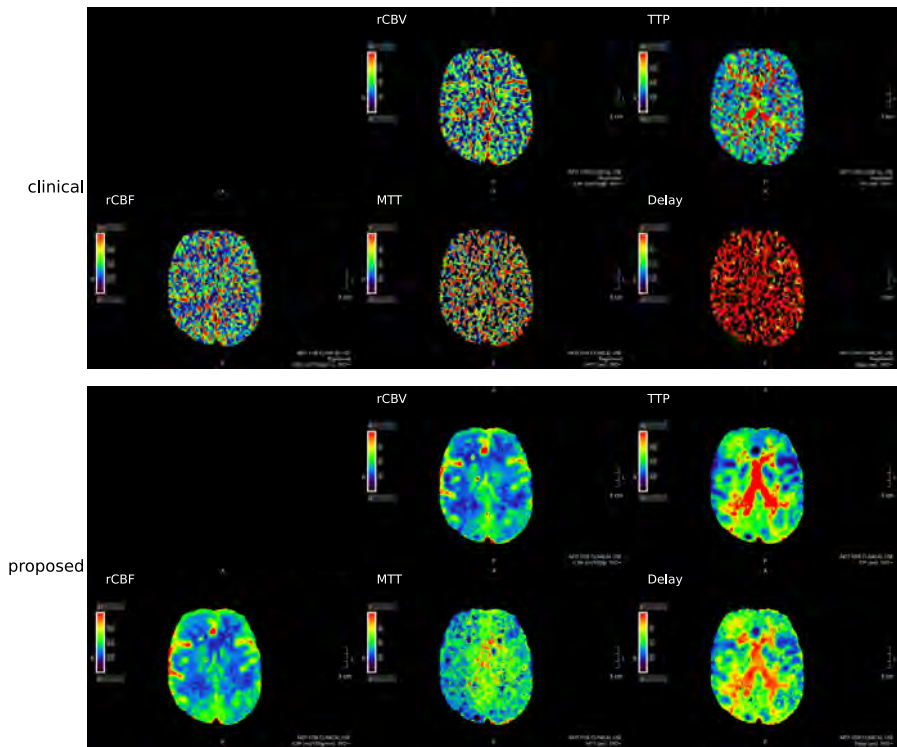


Figure D6: The perfusion maps obtained using the SVD method. The proposed method results in a much cleaner image. However, there is no stroke visible for this patient.

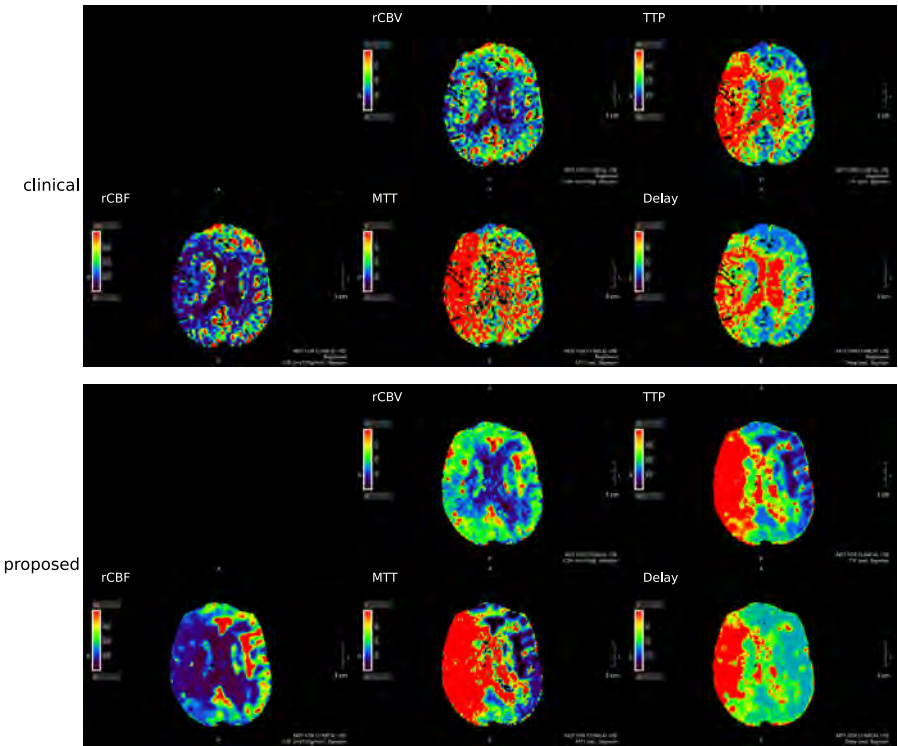


Figure D7: The perfusion maps obtained using Bayesian estimation method. The proposed method delineates the stroke better. Please note, that for this case the alignment tool did not align the clinical and proposed method in the same way.

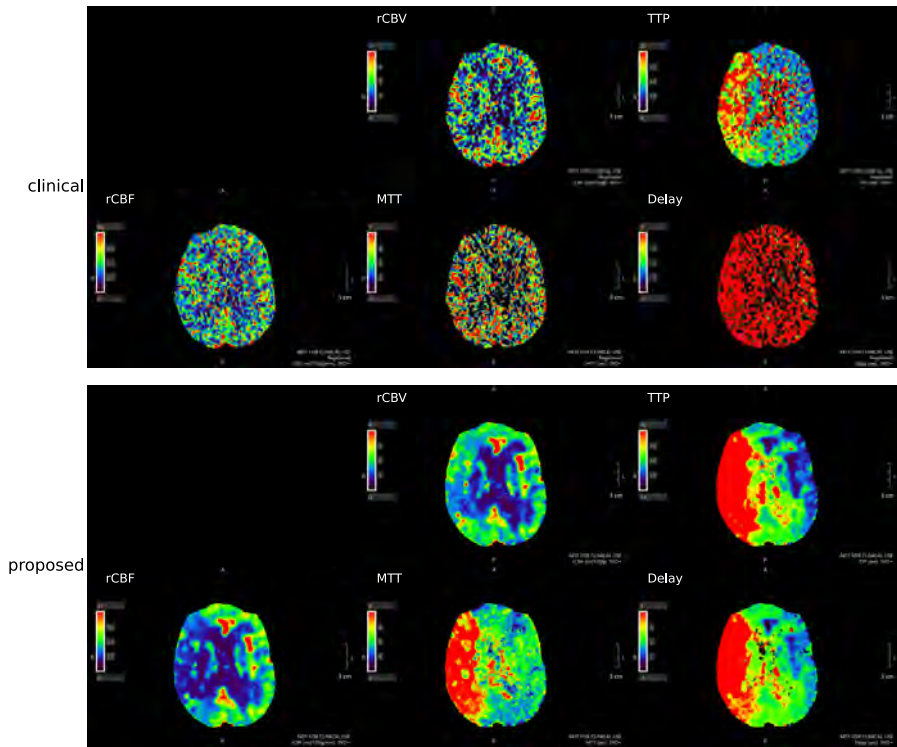


Figure D8: The perfusion maps obtained using the SVD method. The proposed method delineates the stroke better and the stroke is clearly visible in all five perfusion maps instead of only one. Please note, that for this case the alignment tool did not align the clinical and proposed method in the same way.

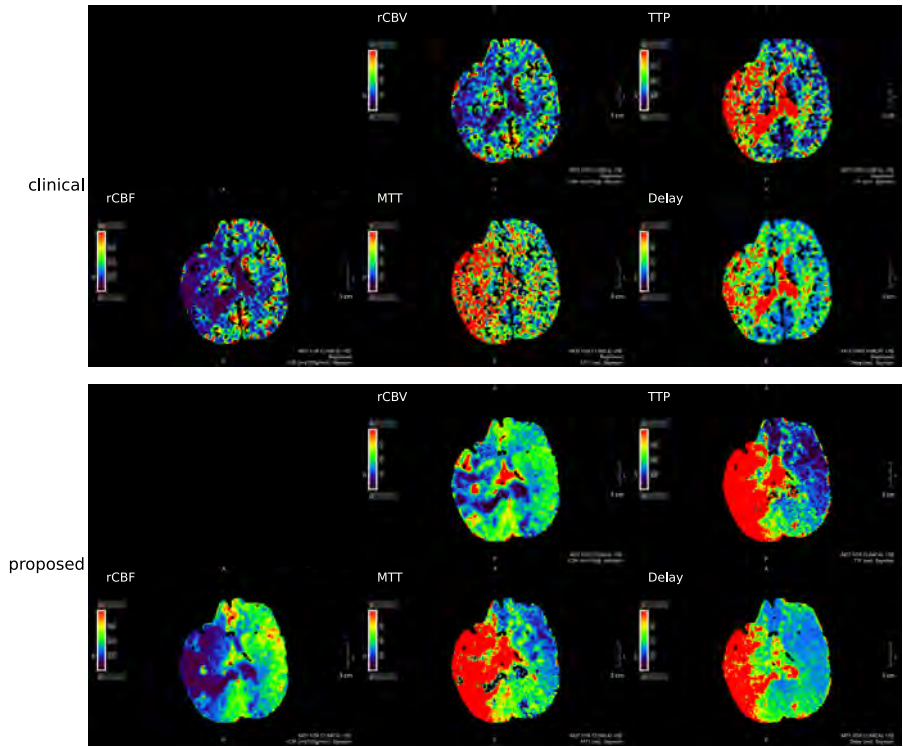


Figure D9: The perfusion maps obtained using Bayesian estimation method. The proposed method delineates the stroke better.

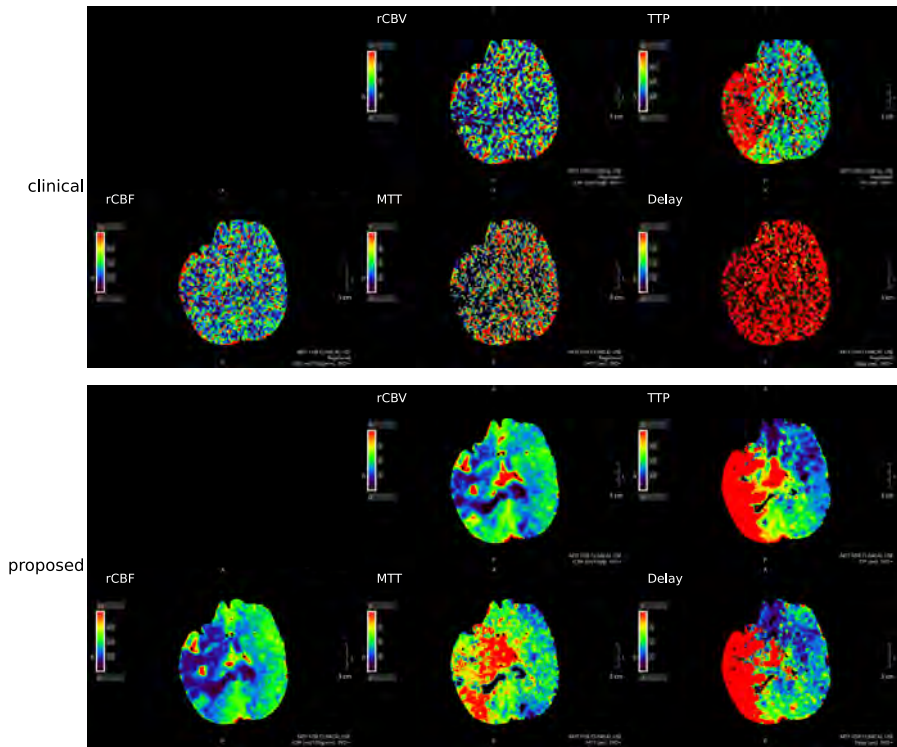


Figure D10: The perfusion maps obtained using the SVD method. The proposed method delineates the stroke better and the stroke is clearly visible in all five perfusion maps instead of only one.

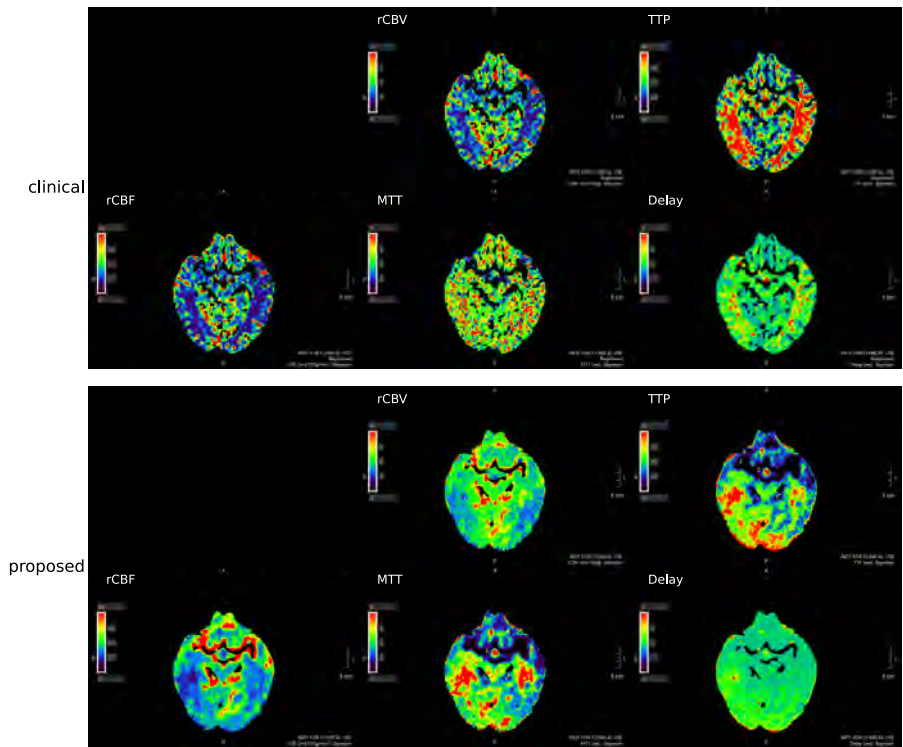


Figure D11: The perfusion maps obtained using Bayesian estimation method. The proposed method results in a cleaner image. The TTP and MTT map of the proposed method show some asymmetry, however the other maps do not show this asymmetry making it not clinically relevant.

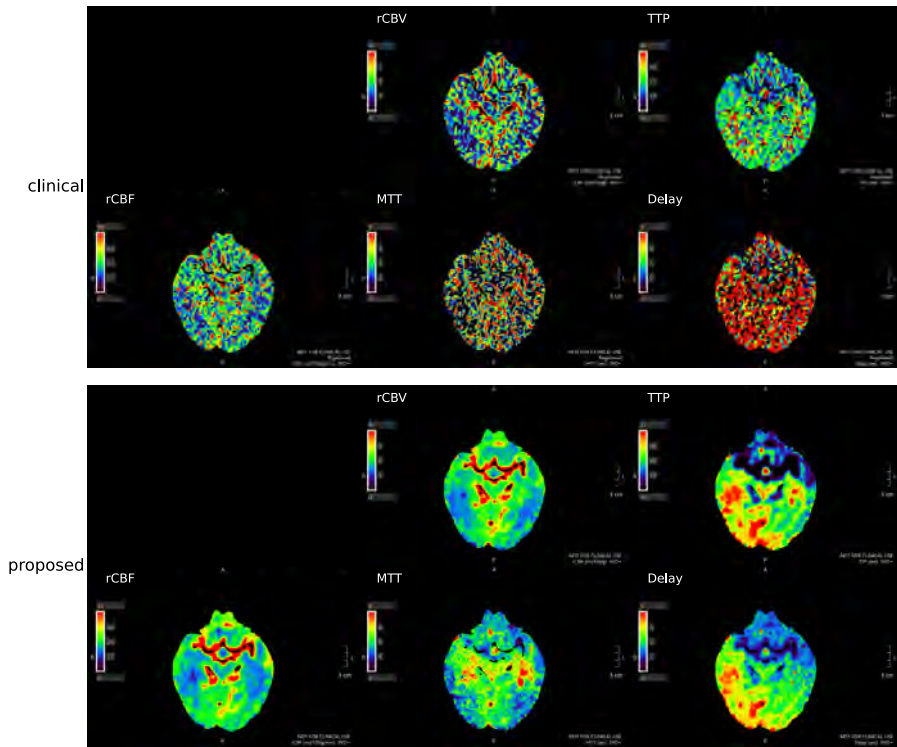


Figure D12: The perfusion maps obtained using the SVD method. The proposed method results in a cleaner image. The TTP and Delay map of the proposed method show some asymmetry, however the other maps do not show this asymmetry making it not clinically relevant.

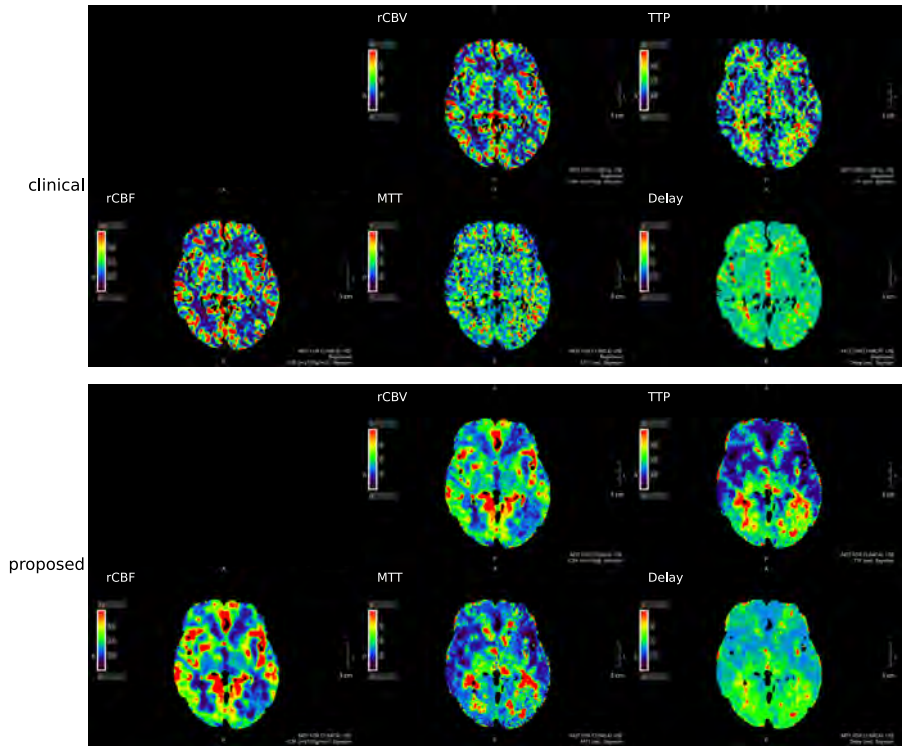


Figure D13: The perfusion maps obtained using Bayesian estimation method. The proposed method results in a cleaner image. However, there is no stroke visible for this patient.

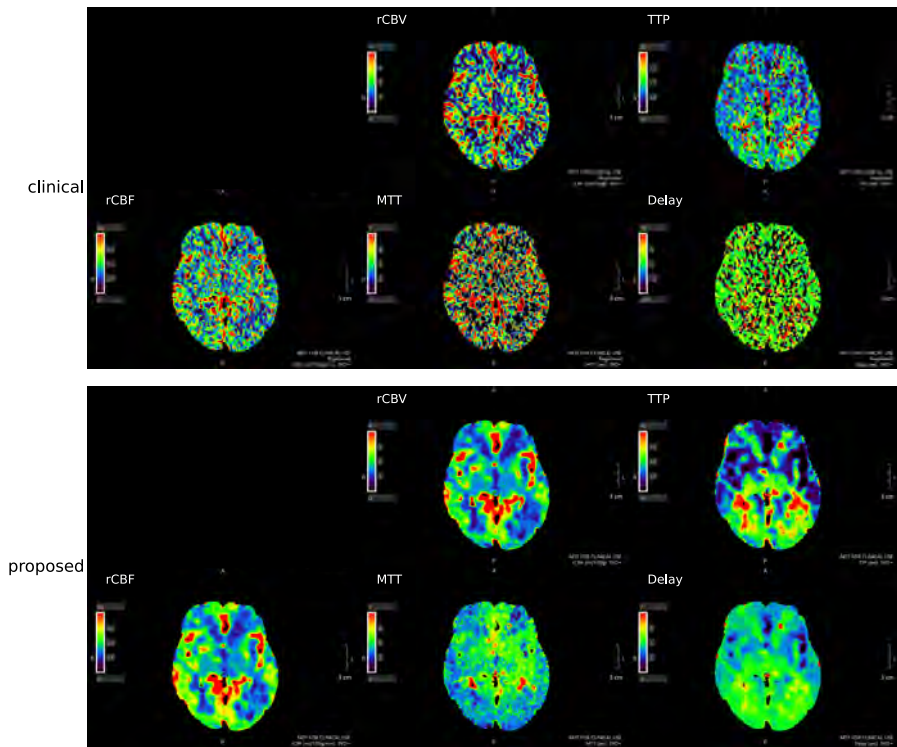


Figure D14: The perfusion maps obtained using the SVD method. The proposed method results in a much cleaner image. However, there is no stroke visible for this patient.

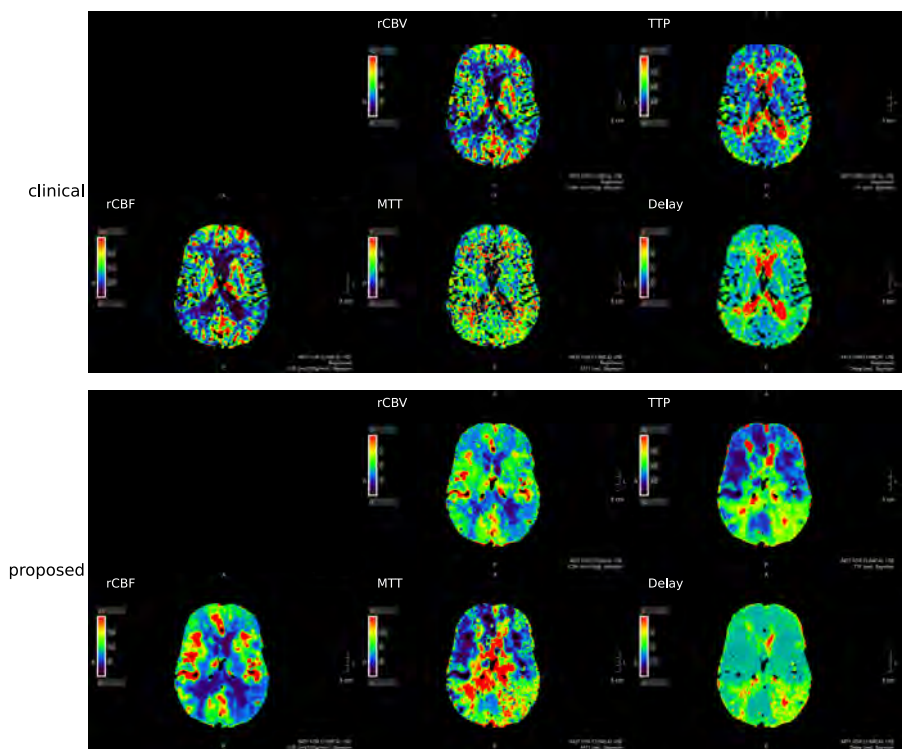


Figure D15: The perfusion maps obtained using Bayesian estimation method. The proposed method results in a cleaner image. However, this low noise content also reveals a region on the left posterior side of the brain that might be interpreted as a stroke. However, this stroke region does not correlate with the symptoms of the patient.

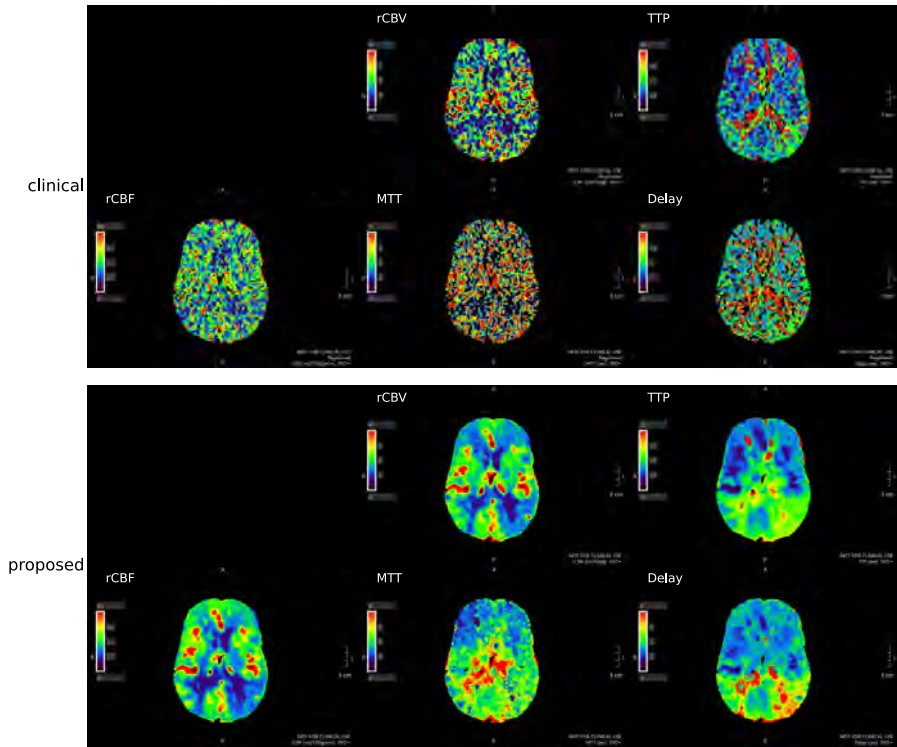


Figure D16: The perfusion maps obtained using the SVD method. The proposed method results in a cleaner image. However, this low noise content also reveals a region on the left posterior side of the brain that might be interpreted as a stroke. However, this stroke region does not correlate with the symptoms of the patient.

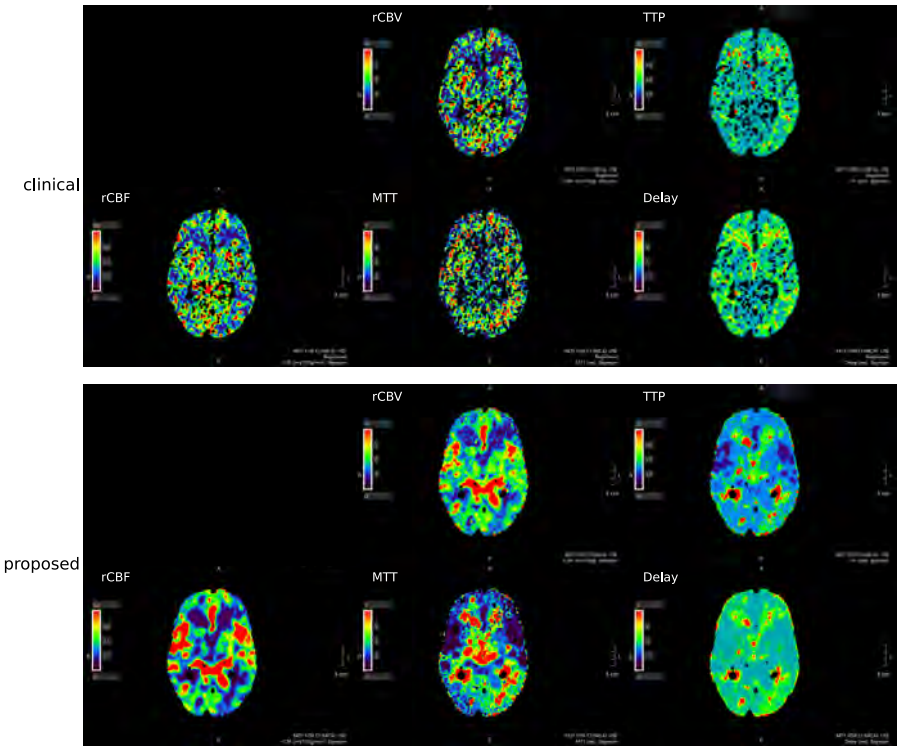


Figure D17: The perfusion maps obtained using Bayesian estimation method. The proposed method results in a cleaner image. However, there is no stroke visible for this patient.

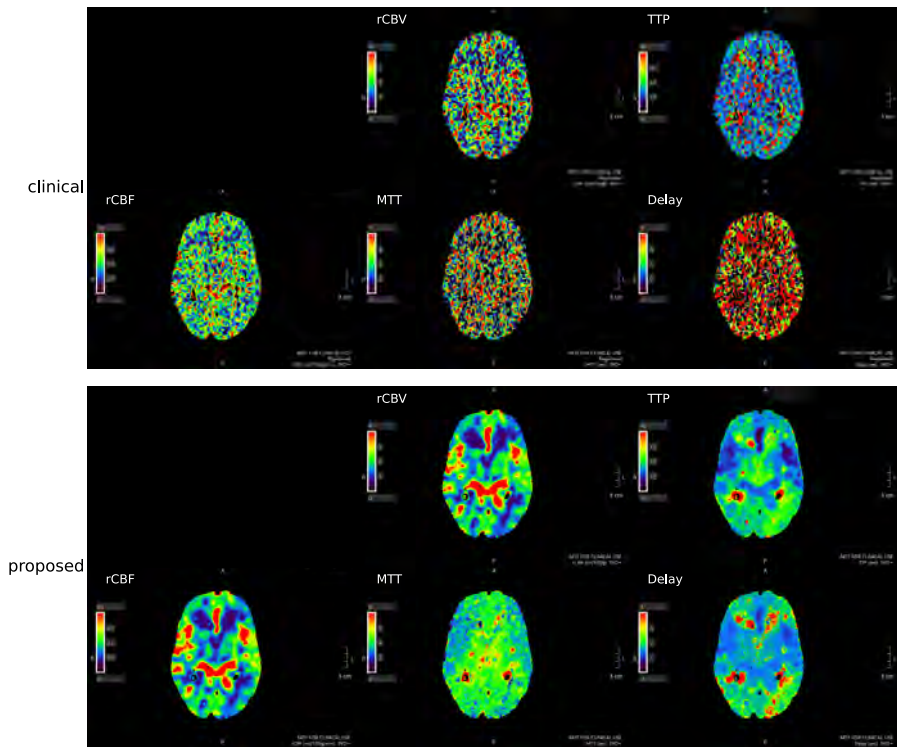


Figure D18: The perfusion maps obtained using the SVD method. The proposed method results in a cleaner image. However, there is no stroke visible for this patient.

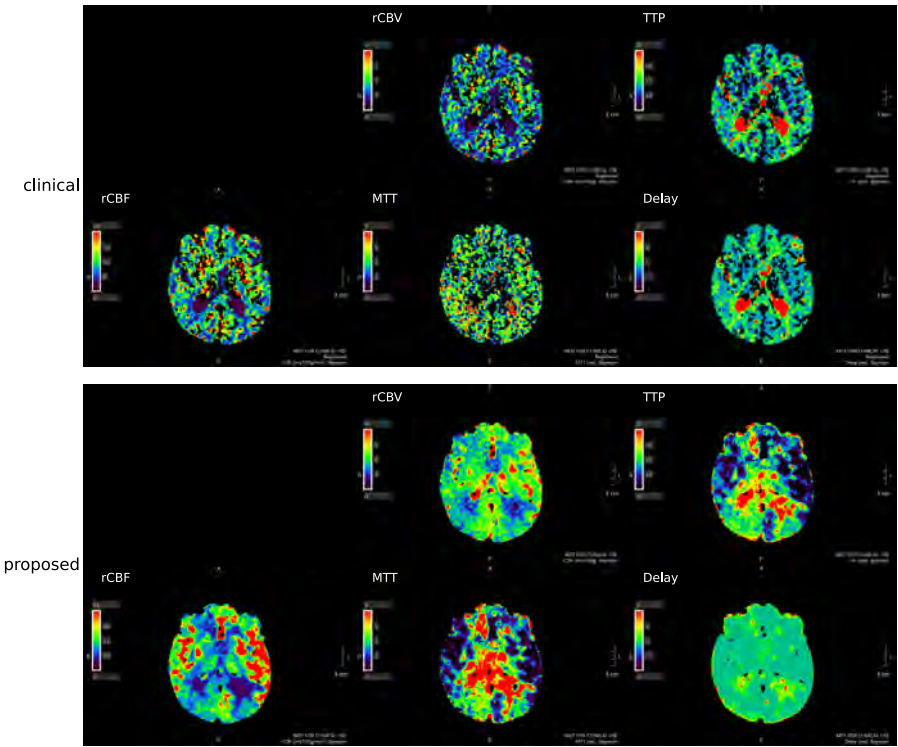


Figure D19: The perfusion maps obtained using Bayesian estimation method. The proposed method results in a cleaner image. However, there is no stroke visible for this patient.

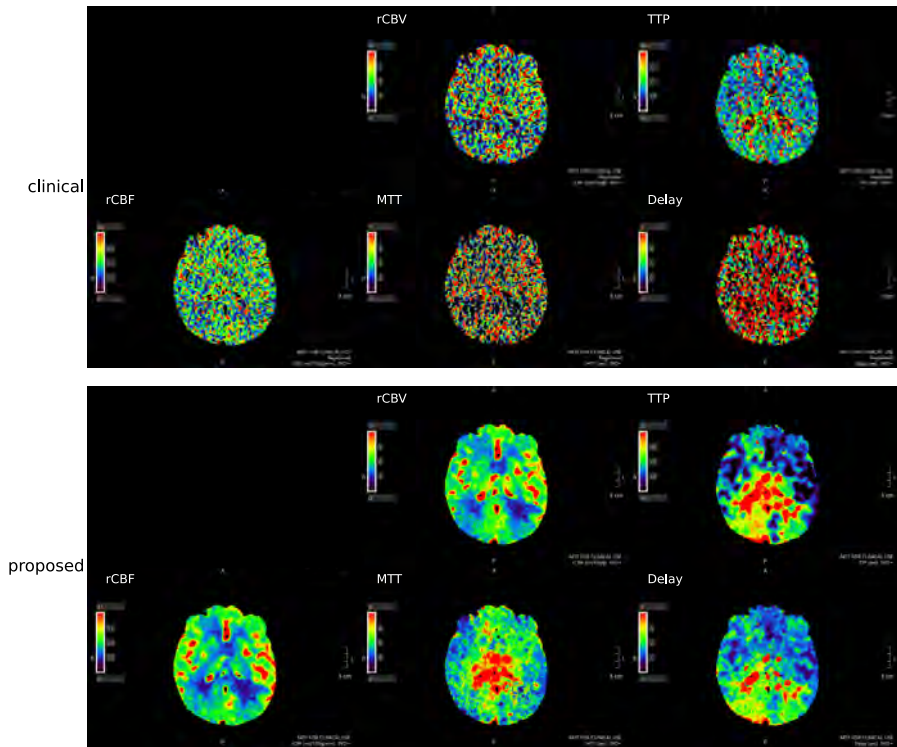


Figure D20: The perfusion maps obtained using the SVD method. The proposed method results in a cleaner image. However, there is no stroke visible for this patient.



General Discussion

In this thesis we explored the potential of the 4D Similarity Filter (4DSF) to reduce noise in CT perfusion (CTP) imaging. The first part of this thesis focused on developing simulation tools to make quantitative analysis of the 4DSF, or any other image processing method, easier. In this chapter we will first discuss the impact of the developed simulation tools, followed by a discussion of future directions for 4DSF research and implementation.

Simulation methods

In **Chapter 2**, we presented a system-specific CT simulator capable of generating realistic CT scans of digital phantoms. This simulator offers a lot of freedom when used for validation of, for instance, processing algorithms, since it allows for simulation of various acquisition settings, such as exposure or tube voltage levels. Therefore, their influence on, e.g., 4DSF performance, can be analyzed extensively without the need to image patients. In the same way, the performance and different parameter settings of the 4DSF, or other processing algorithms, can be tested. Besides avoiding the use of radiation, this type of simulation-based analysis is both quick and cost-effective, requiring only computer power and processing time. Another crucial benefit of using this type of simulations is the availability of a ground truth and thus the possibility to quantitatively assess the accuracy of a processing algorithm. This could lead to a better understanding and potentially improved version of the algorithm before entering into clinical trials.

This research field, of so-called virtual imaging trials (VIT), aims to eventually substitute for most clinical imaging trials. A complete VIT is comprised of three key components: scanner models, patient models, and observer or radiologist models. However, we believe it is still far from achieving the goal of substituting clinical trials with VITs, since the realism of each of these three components is not yet optimal or comes at a great computational cost. From here on we will only focus on the patient models and scanner models, since these are the ones utilized in this thesis.

Regarding patient models, significant improvements are still needed. For example, most digital phantoms lack intra-organ variability, meaning that each organ is modeled as consisting of homogeneous tissue, simplifying the heterogeneous nature of real organs. When studying CT perfusion, it is, of course, important to focus on the simulation of the blood flow. However, most existing phantoms lack extensive artery trees and do not model microvasculature or complex flow dynamics but rather assume uniform flow. The same is true for simulating pathology, where

the accurate modelling of diseases is still far from optimal. For example, the work of Divel et al. [127], which simulates stroke, does not incorporate collateral flow that can be present in stroke patients, thus over-simplifying reality.

In terms of scanner models, while we believe most phenomena occurring in the process of creating projections can be modelled accurately, incorporating all of them would result in excessive computational demands. This points to the need for a tradeoff between realism and computation time and power. We also showed this in **Chapter 2**, where less subsampling of, e.g., the projection angles, saves computation time, but also reduces the accuracy of the rotational blurring.

For certain applications, a high level of realism might not be essential, allowing for computational time saving. However, the level of realism required depends on the method being tested and the goal of the study being performed. Researchers must judge which factors are important for their study, since achieving complete realism for all elements of the VIT seems computationally unfeasible at the moment.

To summarize, while the current state of VITs allows for elaborate testing and optimization of image formation, processing, and analysis algorithms, they are not ready yet to replace clinical trials entirely. However, VITs provide a unique opportunity to perform quantitative analysis on the accuracy of new methods via the available ground truth. This is normally not possible in medical imaging and can, in our opinion, play a key role in the development and refinement of processing algorithms. Additionally, a VIT gives a lot of freedom while testing, since scanner settings and patient characteristics can be finely controlled. As a result, clinical test periods can be shortened, improved algorithms can be designed quicker, and clinical results can be strengthened by quantitative analysis on accuracy. **Chapter 6** is a good example of an improved algorithm design, resulting from the availability of a ground truth.

In some cases, VITs might not be sufficient or efficient in predicting performance. For instance, when training or validating deep learning (DL) tools to detect or classify lesions, the realism of the lesions needs to be as close as possible to patient cases for the DL network to learn the morphology and heterogeneous nature of real lesions and thus generalize well. For these purposes, real patient data is much better suited, since modelling all details and the heterogeneous nature of the disease is difficult.

Therefore, in **Chapter 3** we presented a method to generate low-dose CT scans from existing clinical dose CT scans. This method focuses specifically on the

situation when one does not have access to the projection data, system geometry and components, and reconstruction methods.

Methods with a similar purpose already exist, with most of them being based on projection domain data. However, in a clinical or research setting there is commonly a lack of access to this data and of details on the format of such data. In addition, vendors are usually hesitant to share such information or provide such direct access. Obtaining proprietary details about scanners and processing techniques without support of the vendor can be very time-consuming or downright impossible. The proposed method could therefore be very useful for the generation of low-dose CT scans for testing purposes. To work around restrictions imposed by vendors, the proposed method operates fully in the image domain and is non-analytical, contrary to other image-domain methods, but rather uses a signal processing-based approach. This makes the method not only suitable for filtered back projection (FBP)-based reconstruction, but also for more advanced reconstruction algorithms like DL and model-based reconstructions. The method can basically transform one type of noise to another type of noise using the noise power spectrum and voxel-wise standard deviation. This makes the proposed method ideal for testing purposes, since it is very easy to use.

If one aims for optimal realism of the simulated noise, projection domain methods are preferred, since they model the actual process and are thus expected to be more accurate. However, as explained, these methods can be very time-intensive and, in some situations, it is not even possible to use them due to lack of knowledge of the details of the imaging system. This emphasizes the usefulness of the method proposed in **Chapter 3**, which allows for the generation of low-dose testing data with real patient anatomy without access to scanner details.

Hence, we have developed two methods to generate test data for validating algorithms. As discussed, if the realism of a specific part of the anatomy, normal or diseased, is important, methods corrupting existing patient data, like the one proposed in **Chapter 3**, are more suitable over VITs. As discussed earlier, when performing classification tasks, the realism of the specific disease is of utmost importance. However, in our opinion, the disease models for VITs are not yet sophisticated enough for such classification methods to generalize well to patient data. Nevertheless, we believe these models are sufficiently accurate to test algorithms designed to improve the visibility of a disease by specific processing of the image. Therefore, either a VIT or a noise-corrupting method, like the one of **Chapter 3**, can be used in such cases. The choice between these two methods

could then be based on other factors, such as the availability of patient data, or the type of analysis to be performed on the processed data. For example, if quantitative accuracy, instead of observer performance, needs to be determined, then only the VIT approach is appropriate because it has access to the ground truth for every case.

Ideally, the method proposed in **Chapter 3** or any other image domain method would be unnecessary if computational power were unlimited and system details and access to projection data were available to everybody. However, since this is not the case, workarounds, like the one proposed in **Chapter 3**, have to be developed, which are less optimal.

In general, we think the field would benefit from a stronger vendor-researcher relationship. Especially when it comes to more technical research, since for these projects scanner and software details are often needed. During this PhD, multiple projects would not have been possible without this close vendor collaboration. However, we also noticed that there is a research interest and a company interest, which need to be balanced.

4-Dimensional Similarity Filter

The remainder of this thesis focused on the characterization, validation, and optimization of the 4DSF. At the end, we will give our view on how the 4DSF can best be implemented.

Chapter 5 introduced a slightly modified version of the 4DSF, which can be executed on a GPU. This substantially decreased computation time of the algorithm, allowing for faster or better filter results. Moreover, the GPU implementation is easier to alter in case one wants to implement a different sorting method than temporal mean or if one only wants to filter one timepoint on the time attenuation curve (TAC) instead of the entire TAC. This adaptability of the algorithm is extremely useful in a research setting. This version of the algorithm is used in **Chapter 6**.

In **Chapter 4** we presented an elaborate characterization and validation study of the 4DSF. This chapter uses the well-known XCAT phantoms in combination with the CT simulator of **Chapter 2** to simulate CT perfusion sequences of the liver. In this way, a quantitative analysis of the accuracy resulting from the 4DSF was possible. In this study the influence of different filter parameters, acquisition parameters, and lesion characteristics on image quality was analyzed.

This analysis also provided insight into the strengths and limitations of the 4DSF. The major strength was clearly the filter's ability to reduce image noise significantly, with limited resolution loss, where other noise-reducing filters showed substantial resolution loss. As demonstrated in **Chapter 4**, this limited resolution loss also leaves room for further noise reduction using classic spatial based noise reduction algorithms post 4DSF. This double filtering results in very clean, i.e., denoised, images, showing potential for detection of relatively small lesions. This double filtering could be particularly interesting in the context of low-dose CT scans for lesion characterization. This type of scans is typically used to assess morphology and to judge which part of the lesion enhances.

Although we have not performed a study on this yet, it would be interesting to see how this double filtering would perform on, for example, a four-phase liver protocol for lesion detection. Since the results of **Chapter 4** show potential possibilities to decrease dose, while keeping image quality high for detection tasks. However, needs to be confirmed in a VIT or phantom study first.

The major limitation of the 4DSF is the accuracy of the TAC after application of the 4DSF, especially for small highly-perfusing structures. The analysis showed us that the peak of these TACs is not only lowered, which is expected, but also shifted in time. So, not only the amplitude of the TAC is influenced, but also the shape. Noise reduction comes at the cost of quantitative accuracy, which is an important aspect to understand when using this filter in clinical practice.

This indicates that the optimal parameter setting is not only dependent on image content, such as initial noise level, lesion characteristics, but also on the diagnostic task at hand. Detection tasks do not per se require quantitative accuracy, while quantitative accuracy may become more important when perfusion analysis is used for lesion characterization. However, the importance of quantitative accuracy also depends on the difference in the perfusion characteristics between lesion types. This is, together with the understanding on how the filter output is influenced by the parameter settings, the most important lesson learned from the study of **Chapter 4**.

The aim of **Chapter 6** was to optimize the use of the 4DSF to accurately detect and assess acute strokes. The original aim was to use the knowledge obtained in **Chapter 4** to optimize the parameter settings of the 4DSF for stroke detection. To do this in a quantitative manner, a brain phantom was used in combination with the CT simulator developed in **Chapter 2**. While the insights of **Chapter 4** helped us to understand and improve the results of the 4DSF to some degree compared to

the clinically used version, the 4DSF still underperformed compared to what was expected and resulted in hampered stroke visibility for some cases. This shows, first of all, that the knowledge obtained in **Chapter 4** on liver CT perfusion protocols do not translate directly to stroke imaging, and secondly that for some clinical applications the original 4DSF implementation is not so suitable to reduce noise.

Chapter 4 already helped us understand how various lesion characteristics influence the image quality after 4DSF. However, in **Chapter 6** we deal with a completely different disease and a completely different anatomy. We underestimated how a different disease and anatomy influences the performance of the 4DSF. In **Chapter 4** the task is to differentiate between liver parenchyma and the vessels and tumors within it, which stand out from the liver parenchyma. However, in **Chapter 6** the task is to differentiate between healthy and ischemic brain tissue. The healthy brain tissue receives relatively little blood compared to vessels; thus, the attenuation change in the tissues is limited. During a stroke the affected brain tissue receives even less blood, however as the attenuation change is already limited relative to the noise this decrease is hardly noticeable.

As explained in the introduction, ideally each TAC is compared to all TACs in the dataset. However, this is computationally not feasible. Therefore, each TAC is only compared to a subset of TACs, which have the most similar temporal mean. This temporal mean is based on attenuation differences, which are very minimal in stroke imaging. So, the candidate TACs in the subset will be suboptimal and will result in suboptimal filtering.

To overcome this problem, we suggested an alternative version of the 4DSF basing the subset for searching similar TACs not on temporal mean, but rather on peak time, since tissue suffering stroke does not only receive less blood, but also receives it later in time. In this way the ischemic tissue could be differentiated better from the healthy tissue.

Chapter 6 taught us that optimizing the 4DSF entails more than just optimizing a set of parameters. It requires understanding of underlying disease physiology, which then dictates what the optimal modifications to the algorithm or its parameters are. Therefore, it is rather hard to give one method or parameter set that will always work.

The original goal of this thesis was to develop a DL-based method to find the optimal parameter set for each CT perfusion dataset. A DL network would be trained to predict the image quality of a CT perfusion dataset after filtering. The input of this DL network being the filter parameters and data of the CT perfusion protocol and acquisition settings, for example number of scans, time between scans, noise level, number of

voxels, reason for CT scan, etc. The optimal parameter set could then be obtained via gradient information of the DL network, thus automating the optimal parameter selection. However, as this thesis shows, image quality is influenced by many different factors, including purpose of scan, body part imaged, patient anatomy, disease type or size, among others. This suggests that CT perfusion protocol, image characteristics, and acquisition settings will probably not be enough to predict image quality. **Chapter 6** is the perfect example, since it shows that a change in the anatomy and/or disease characteristics causes the 4DSF to be suboptimal and the algorithm itself, not only its input parameters, had to be updated, compared to **Chapter 4**. This indicates that the image quality prediction does not generalize well across different anatomies and diseases, making the development of one DL network to predict optimal parameter settings nearly impossible. Therefore, we have chosen not to pursue this initial aim.

Additionally, the data for training this DL network should have been generated using the CT simulator of **Chapter 2** in combination with phantoms and disease models. However, as discussed before, disease models can be complex and accurate models do not yet exist for all diseases. This poses a challenge in obtaining a large and diverse dataset with consistent quality across various diseases. Consequently, the performance of the DL network and thus the 4DSF might be inconsistent or suboptimal when applied to patient data.

Based on the findings of this thesis, we would suggest that a systematic approach for optimization be taken when the 4DSF is introduced for a new CT perfusion protocol. We recommend to first analyze the anatomy and disease physiology to determine what differentiates the healthy tissue from the disease. This understanding should then be used to alter the 4DSF parameters and/or algorithm to optimize the performance for the disease and anatomy at hand. While making these choices, secondary factors should be considered as well, such as computational time, because perfusion results may sometimes be needed quickly, as is the case in stroke imaging.

Throughout the research performed for this thesis we have also come to understand some characteristics of the 4DSF that are not described in any of the chapters of this thesis. First, the 4DSF assumes no noise correlation between scans, i.e., no correlation between timepoints on the TAC. If this assumption is violated, the principle of preventing bias by leaving the timepoint being filtered out of the search for similar voxels becomes invalid. In such a case, we recommend to determine how many timepoints the correlation stretches over and leave this number of timepoints, instead of just one, out of the search for similar voxels. Second, since the similarity

filter's strength is to reduce image noise with minimal resolution loss, it is not advised to reconstruct images to be filtered with strong model-based or DL reconstruction, which sacrifices spatial resolution to decrease noise. This is because the resolution lost in the original image reconstruction cannot be retrieved by the 4DSF.

If enough computational power is available, it would be advisable to investigate the reconstruction of one image using FBP with a sharp kernel and one image using a strong model-based or DL reconstruction. The latter images can then be used to determine the voxels that should be averaged by the 4DSF, just like the average filter is now used in **Chapter 4**, while the FBP reconstructed images are used to retrieve the select voxel values and average them. This might lead to improved results.

For the same reason, the 4DSF achieves sub-par performance when applied to time-series MRI data, such as diffusion-weighted imaging. This is due to the inherently low noise, low spatial resolution and small matrix sizes of perfusion MRI. If one wants to enhance the 4DSF's efficacy on time-series MRI data, we advise to first explore the development of MRI sequences that yield higher resolution, which would likely lead to less correlated noise and a higher noise content. This could potentially improve the performance of the 4DSF by providing a more suitable input for its noise reduction capabilities.

To summarize, the research presented in this thesis introduced two tools to perform quantitative evaluation of processing algorithms. One using digital phantoms and the other using patient data. The latter one is the most novel, since it enables researchers with limited knowledge about their system to simulate lower dose CT data. However, the core contribution of this thesis is the investigation into the characteristics and performance of the 4DSF. The key points of this part are the following: first, the limited resolution loss of the 4DSF leaves room for further noise reduction using classic spatial-based noise reduction algorithms post-4DSF application. Second, when utilizing the 4DSF, one should be aware of the potential introduction of a modest bias to the TAC, which might influence the diagnosis, especially when performing quantitative analysis. Lastly, optimizing the parameters of the 4DSF is not a straightforward task, since disease characteristics and anatomy heavily influence the results. Therefore, one should focus on understanding the disease physiology and thus the difference between healthy and diseased tissue to optimize the algorithm by parameter selection or altering the algorithm itself.



Bibliography

1. Brenner DJ, Hall EJ, Phil D. Computed Tomography-An Increasing Source of Radiation Exposure. *N Engl J Med*. 2007;357(22):2277-2284. doi:10.1056/NEJMra072149
2. Lee TY, Chhem RK. Impact of new technologies on dose reduction in CT. *Eur J Radiol*. 2010;76(1):28-35. doi:10.1016/j.ejrad.2010.06.036
3. Smith-Bindman R, Kwan ML, Marlow EC, et al. Trends in Use of Medical Imaging in US Health Care Systems and in Ontario, Canada, 2000-2016. *JAMA*. 2019;322(9):843-856. doi:10.1001/JAMA.2019.11456
4. Bosch de Basea M, Salotti JA, Pearce MS, et al. Trends and patterns in the use of computed tomography in children and young adults in Catalonia — results from the EPI-CT study. *Pediatr Radiol*. 2016;46(1):119-129. doi:10.1007/s00247-015-3434-5
5. Westmark S, Hesselund T, Hoffmann A, et al. Increasing use of computed tomography scans in the North Denmark Region raises patient safety concern. *Eur J Radiol*. 2023;166:110997. doi:10.1016/j.ejrad.2023.110997
6. Broder J, Warshauer DM. Increasing utilization of computed tomography in the adult emergency department, 2000-2005. *Emerg Radiol*. 2006;13(1):25-30. doi:10.1007/s10140-006-0493-9
7. Dovalés ACM, Da Rosa LAR, Kesminiene A, Pearce MS, Veiga LHS. Patterns and trends of computed tomography usage in outpatients of the Brazilian public healthcare system, 2001-2011. *Journal of Radiological Protection*. 2016;36(3):547-560. doi:10.1088/0952-4746/36/3/547
8. Pelc NJ. Recent and future directions in CT imaging. *Ann Biomed Eng*. 2014;42(2):260-268. doi:10.1007/s10439-014-0974-z
9. Heiken JP, Brink JA, Vannier MW. Spiral (Helical) CT. *Radiology*. 1993;189(3):647-656. doi:10.1148/radiology.189.3.8234684
10. Kalra MK, Maher MM, Toth TL, D'Souza R, Saini S. Multidetector computed tomography technology: Current status and emerging developments. *J Comput Assist Tomogr*. 2004;28 Suppl 1:2-6. doi:10.1097/01.rct.0000120857.80935.bd
11. Siebert E, Böhner G, Dewey M, et al. 320-Slice CT neuroimaging: Initial clinical experience and image quality evaluation. *Br J Radiol*. 2009;82(979):561-570. doi:10.1259/bjr/27721218
12. Ferda J, Vendiš T, Flohr T, et al. Computed tomography with a full FOV photon-counting detector in a clinical setting, the first experience. *Eur J Radiol*. 2021;137:109614. doi:10.1016/j.ejrad.2021.109614
13. Sodickson A, Baeyens PF, Andriole KP, et al. Recurrent CT, Cumulative Radiation Exposure, and Associated Radiation-induced Cancer Risks from CT of Adults. *Radiology*. 2009;251(1):175-184. doi:10.1148/radiol.2511081296
14. Hall EJ. Lessons we have learned from our children: Cancer risk from diagnostic radiology. *Pediatr Radiol*. 2002;32(10):700-706. doi:10.1007/s00247-002-0774-8
15. Howard BM, Hu R, Barrow JW, Barrow DL. Comprehensive review of imaging of intracranial aneurysms and angiographically negative subarachnoid hemorrhage. *Neurosurg Focus*. 2019;47(6):1-13. doi:10.3171/2019.9.FOCUS19653
16. Østergaard L, Weisskoff RM, Chesler DA, Gyldensted G, Rosen BR. High resolution measurement of cerebral blood flow using intravascular tracer bolus passages. Part I: Mathematical approach and statistical analysis. *Magn Reson Med*. 1996;36(5):715-725. doi:10.1002/mrm.1910360510
17. Mouridsen K, Friston K, Hjort N, Gyldensted L, Østergaard L, Kiebel S. Bayesian estimation of cerebral perfusion using a physiological model of microvasculature. *Neuroimage*. 2006;33(2):570-579. doi:10.1016/j.neuroimage.2006.06.015

18. Jovin TG, Saver JL, Ribo M, et al. Diffusion-weighted imaging or computerized tomography perfusion assessment with clinical mismatch in the triage of wake up and late presenting strokes undergoing neurointervention with Trevo (DAWN) trial methods. *International Journal of Stroke*. 2017;12(6):641-652. doi:10.1177/1747493017710341
19. Albers GW, Lansberg MG, Kemp S, et al. A multicenter randomized controlled trial of endovascular therapy following imaging evaluation for ischemic stroke (DEFUSE 3). *International Journal of Stroke*. 2017;12(8):896-905. doi:10.1177/1747493017701147
20. Kim SH, Kamaya A, Willmann JK. CT perfusion of the liver: Principles and applications in oncology. *Radiology*. 2014;272(2):322-344. doi:10.1148/radiol.14130091
21. Therasse P, Arbuck SG, Eisenhauer EA, et al. New guidelines to evaluate the response to treatment in solid tumors. *J Natl Cancer Inst*. 2000;92(3):205-216. doi:10.1093/jnci/92.3.205
22. Eisenhauer EA, Therasse P, Bogaerts J, et al. New response evaluation criteria in solid tumours: Revised RECIST guideline (version 1.1). *Eur J Cancer*. 2009;45(2):228-247. doi:10.1016/j.ejca.2008.10.026
23. Shinagare AB, Jagannathan JP, Krajewski KM, Ramaiya NH. Liver metastases in the era of molecular targeted therapy: New faces of treatment response. *American Journal of Roentgenology*. 2013;201(1):15-28. doi:10.2214/AJR.12.9498
24. Marcus CD, Ladam-Marcus V, Cucu C, Bouché O, Lucas L, Hoeffel C. Imaging techniques to evaluate the response to treatment in oncology: Current standards and perspectives. *Crit Rev Oncol Hematol*. 2009;72(3):217-238. doi:10.1016/j.critrevonc.2008.07.012
25. Klein S, Staring M, Murphy K, Viergever MA, Pluim JPW. Elastix: A toolbox for intensity-based medical image registration. *IEEE Trans Med Imaging*. 2010;29(1):196-205. doi:10.1109/TMI.2009.2035616
26. Wyawahare M V, Patil PM, Abhyankar HK. Image Registration Techniques : An overview. *International Journal of Signal Processing, Image Processing and Pattern Recognition*. 2009;2(3):11-28.
27. Fu Y, Lei Y, Wang T, Curran WJ, Liu T, Yang X. Deep learning in medical image registration: A review. *Phys Med Biol*. 2020;65(20):0-27. doi:10.1088/1361-6560/ab843e
28. Newman B, Callahan MJ. ALARA (as low as reasonably achievable) CT 2011--executive summary. *Pediatr Radiol*. 2011;41 Suppl 2:453-455. doi:10.1007/s00247-011-2154-8
29. Kim KW, Lee JM, Kim JH, et al. CT color mapping of the arterial enhancement fraction of VX2 carcinoma implanted in rabbit liver: Comparison with perfusion CT. *American Journal of Roentgenology*. 2011;196(1):102-108. doi:10.2214/AJR.09.3971
30. Kim KW, Lee JM, Klotz E, et al. Quantitative CT Color Mapping of the Arterial Enhancement Fraction of the Liver to Detect Hepatocellular Carcinoma. *Radiology*. 2009;250(2):425-434. doi:10.1148/radiol.2501072196
31. Mahnken AH, Klotz E, Schreiber S, et al. Volumetric arterial enhancement fraction predicts tumor recurrence after hepatic radiofrequency ablation of liver metastases: Initial results. *American Journal of Roentgenology*. 2011;196(5):573-579. doi:10.2214/AJR.10.4410
32. Toshiba MS. AIDR 3D Iterative Reconstruction: Integrated, Automated and Adaptive Dose Reduction. *White Paper*. Published online 2012:1-10.
33. Adler J, Öktem O. Learned Primal-Dual Reconstruction. *IEEE Trans Med Imaging*. 2018;37(6):1322-1332. doi:10.1109/TMI.2018.2799231
34. Lee H, Lee J, Kim H, Cho B, Cho S. Deep-neural-network-based sinogram synthesis for sparse-view CT image reconstruction. *IEEE Trans Radiat Plasma Med Sci*. 2019;3(2):109-119. doi:10.1109/TRPMS.2018.2867611

35. Kang E, Chang W, Yoo J, Ye JC. Deep Convolutional Framelet Denosing for Low-Dose CT via Wavelet Residual Network. *IEEE Trans Med Imaging*. 2018;37(6):1358-1369. doi:10.1109/TMI.2018.2823756
36. Adler J, Öktem O. Deep Bayesian Inversion. *ArXiv*. Published online 2018. doi:10.48550/arXiv.1811.05910
37. Mendrik AM, Vonken EJ, Van Ginneken B, et al. TIPS bilateral noise reduction in 4D CT perfusion scans produces high-quality cerebral blood flow maps. *Phys Med Biol*. 2011;56(13):3857-3872. doi:10.1088/0031-9155/56/13/008
38. Pisana F, Henzler T, Schönberg S, Klotz E, Schmidt B, Kachelrieß M. Noise reduction and functional maps image quality improvement in dynamic CT perfusion using a new k-means clustering guided bilateral filter (KMGB). *Med Phys*. 2017;44(7):3464-3482. doi:10.1002/MP.12297
39. Yu L, Shiung M, Jondal D, McCollough CH. Development and validation of a practical lower-dose-simulation tool for optimizing computed tomography scan protocols. *J Comput Assist Tomogr*. 2012;36(4):477-487. doi:10.1097/RCT.0b013e318258e891
40. Žabic S, Wang Q, Morton T, Brown KM. A low dose simulation tool for CT systems with energy integrating detectors. *Med Phys*. 2013;40(3):031102/1-14. doi:10.1118/1.4789628
41. Divel SE, Pelc NJ. Accurate Image Domain Noise Insertion in CT Images. *IEEE Trans Med Imaging*. 2020;39(6):1906-1916. doi:10.1109/TMI.2019.2961837
42. Segars WP, Bond J, Frush J, et al. Population of anatomically variable 4D XCAT adult phantoms for imaging research and optimization. *Med Phys*. 2013;40(4):1-11. doi:10.1118/1.4794178
43. Abadi E, Harrawood B, Sharma S, Kapadia A, Segars WP, Samei E. DukeSim: A realistic, rapid, and scanner-specific simulation framework in computed tomography. *IEEE Trans Med Imaging*. 2019;38(6):1457-1465. doi:10.1109/TMI.2018.2886530
44. Caballo M, Michielsen K, Fedon C, Sechopoulos I. Towards 4D dedicated breast CT perfusion imaging of cancer: Development and validation of computer simulated images. *Phys Med Biol*. 2019;64(24). doi:10.1088/1361-6560/ab55ac
45. Hess EP, Haas LR, Shah ND, Stroebe RJ, Denham CR, Swensen SJ. Trends in Computed Tomography Utilization Rates. *J Patient Saf*. 2014;10(1):52-58. doi:10.2307/26633038
46. Broder J, Fordham LA, Warshauer DM. Increasing utilization of computed tomography in the pediatric emergency department, 2000-2006. *Emerg Radiol*. 2007;14(4):227-232. doi:10.1007/S10140-007-0618-9
47. Kocher KE, Meurer WJ, Fazel R, Scott PA, Krumholz HM, Nallamothu BK. National Trends in Use of Computed Tomography in the Emergency Department. *Ann Emerg Med*. 2011;58(5):452-462.e3. doi:10.1016/J.ANNEMERGEMED.2011.05.020
48. Mettler J, Wiest PW, Locken JA, Kelsey CA. CT scanning: Patterns of use and dose. *Journal of Radiological Protection*. 2000;20(4):353-359. doi:10.1088/0952-4746/20/4/301
49. Bach PB, Mirkin JN, Oliver TK, et al. Benefits and harms of CT screening for lung cancer: A systematic review. *JAMA - Journal of the American Medical Association*. 2012;307(22):2418-2429. doi:10.1001/jama.2012.5521
50. Koning HJ de, Aalst CM van der, Jong PA de, et al. Reduced Lung-Cancer Mortality with Volume CT Screening in a Randomized Trial. *N Engl J Med*. 2020;382(6):503-513. doi:10.1056/NEJMoa1911793
51. VanSonnenberg E, Casola G, Ho M, et al. Difficult thoracic lesions: CT-guided biopsy experience in 150 cases. *Radiology*. 1988;167(2):457-461. doi:10.1148/radiology.167.2.3357956
52. Erickson SJ, Hogan QH. CT-guided injection of the stellate ganglion: Description of technique and efficacy of sympathetic blockade. *Radiology*. 1993;188(3):707-709. doi:10.1148/radiology.188.3.8351337

53. Kang E, Min J, Ye JC. A deep convolutional neural network using directional wavelets for low-dose X-ray CT reconstruction. *Med Phys*. 2017;44(10):e360-e375. doi:10.1002/MP.12344
54. Bhadauria HS, Dewal ML. Medical image denoising using adaptive fusion of curvelet transform and total variation. *Computers and Electrical Engineering*. 2013;39(5):1451-1460. doi:10.1016/J.COMPELECENG.2012.04.003
55. Smit EJ, Prokop WM. NOISE REDUCTION IN IMAGE DATA. Published online 2017.
56. Anam C, Sutanto H, Adi K, et al. Development of a computational phantom for validation of automated noise measurement in CT images. *Biomed Phys Eng Express*. 2020;6(6). doi:10.1088/2057-1976/abb2f8
57. Abadi E, Segars WP, Tsui BMW, et al. Virtual clinical trials in medical imaging: a review. *Journal of Medical Imaging*. 2020;7(4). doi:10.1117/1.jmi.7.4.042805
58. Saunders RS, Samei E. A method for modifying the image quality parameters of digital radiographic images. *Med Phys*. 2003;30(11):3006-3017. doi:10.1118/1.1621870
59. Moriakov N, Sonke JJ, Teuwen J. LIRE: Learned Invertible Reconstruction for Cone Beam CT. *ArXiv*. Published online 2022. doi:10.48550/arXiv.2205.07358
60. Syben C, Michen M, Stimpel B, Seitz S, Ploner S, Maier AK. Technical Note: PYRO-NN: Python reconstruction operators in neural networks. *Med Phys*. 2019;46(11):5110-5115. doi:10.1002/mp.13753
61. McKenney SE, Nosratieh A, Gelskey D, et al. Experimental validation of a method characterizing bow tie filters in CT scanners using a real-time dose probe. *Med Phys*. 2011;38(3):1406-1415. doi:10.1118/1.3551990
62. Hernandez AM, Boone JM. Tungsten anode spectral model using interpolating cubic splines: Unfiltered x-ray spectra from 20 kV to 640 kV. *Med Phys*. 2014;41(4). doi:10.1118/1.4866216
63. Swinehart DF. The Beer-Lambert law. *J Chem Educ*. 1962;39(7):333-335. doi:10.1021/ed039p333
64. Elam WT, Ravel BD, Sieber JR. A new atomic database for X-ray spectroscopic calculations. *Radiation Physics and Chemistry*. 2002;63(2):121-128. doi:10.1016/S0969-806X(01)00227-4
65. Boone JM, Chavez AE. Comparison of x-ray cross sections for diagnostic and therapeutic medical physics. *Med Phys*. 1996;23(12):1997-2005. doi:10.1118/1.597899
66. Fujita H, Tsai DY, Itoh T, et al. A Simple Method for Determining the Modulation Transfer Function in Digital Radiography. *IEEE Trans Med Imaging*. 1992;11(1):34-39. doi:10.1109/42.126908
67. Siewerdsen JH, Antonuk LE, El-Mohri Y, Yorkston J, Huang W, Cunningham IA. Signal, noise power spectrum, and detective quantum efficiency of indirect-detection flat-panel imagers for diagnostic radiology. *Med Phys*. 1998;25(5):614-628. doi:10.1118/1.598243
68. Feldkamp LA, Davis LC, Kress JW. Practical cone-beam algorithm. *J Opt Soc Am*. 1984;1(6):612-619.
69. Sun Nuclear. Multi-Energy CT Phantom. Accessed January 6, 2022. https://www.sunnuclear.com/uploads/documents/datasheets/Diagnostic/MECT_Phantom_102121.pdf
70. Kayugawa A, Ohkubo M, Wada S. Accurate determination of ct point-spread-function with high precision. *J Appl Clin Med Phys*. 2013;14(4):216-226. doi:10.1120/jacmp.v14i4.3905
71. Kwan ALC, Boone JM, Yang K, Huang SY. Evaluation of the spatial resolution characteristics of a cone-beam breast CT scanner. *Med Phys*. 2007;34(1):275-281. doi:10.1118/1.2400830
72. Balakrishnan V. All about the dirac delta function(?). *Resonance*. 2003;8(8):48-58. doi:10.1007/bf02866759
73. Beylkin G. Discrete Radon Transform. *IEEE Trans Acoust*. 1987;35(2):162-172. doi:10.1109/TASSP.1987.1165108

74. De Man B, Basu S, Chandra N, et al. CatSim: a new computer assisted tomography simulation environment. In: *Medical Imaging 2007: Physics of Medical Imaging*. Vol 6510.; 2007:65102G. doi:10.1117/12.710713
75. Wu M, Fitzgerald P, Zhang J, et al. XCIST - An open access x-ray/CT simulation toolkit. *Phys Med Biol*. 2022;67(19). doi:10.1088/1361-6560/ac9174
76. Slovis TL. The ALARA concept in pediatric CT: Myth or reality? *Radiology*. 2002;223(1):5-6. doi:10.1148/radiol.2231012100
77. Uffmann M, Schaefer-Prokop C. Digital radiography: The balance between image quality and required radiation dose. *Eur J Radiol*. 2009;72(2):202-208. doi:10.1016/J.EJRAD.2009.05.060
78. Heinrich MP, Stille M, Buzug TM. Residual U-Net convolutional neural network architecture for low-dose CT denoising. *Current Directions in Biomedical Engineering*. 2018;4(1):297-300. doi:10.1515/cdbme-2018-0072
79. Chen H, Zhang Y, Kalra MK, et al. Low-Dose CT with a residual encoder-decoder convolutional neural network. *IEEE Trans Med Imaging*. 2017;36(12):2524-2535. doi:10.1109/TMI.2017.2715284
80. Yang Q, Yan P, Zhang Y, et al. Low-Dose CT Image Denoising Using a Generative Adversarial Network With Wasserstein Distance and Perceptual Loss. *IEEE Trans Med Imaging*. 2018;37(6):1348-1357. doi:10.1109/TMI.2018.2827462
81. Jin KH, McCann MT, Froustey E, Unser M. Deep Convolutional Neural Network for Inverse Problems in Imaging. *IEEE Transactions on Image Processing*. 2017;26(9):4509-4522. doi:10.1109/TIP.2017.2713099
82. Naziroglu RE, van Ravesteijn VF, van Vliet LJ, Streekstra GJ, Vos FM. Simulation of scanner- and patient-specific low-dose CT imaging from existing CT images. *Physica Medica*. 2017;36:12-23. doi:10.1016/j.ejmp.2017.02.009
83. Zeng D, Huang J, Bian Z, et al. A simple low-dose X-Ray CT simulation from high-dose scan. *IEEE Trans Nucl Sci*. 2015;62(5):2226-2233. doi:10.1109/TNS.2015.2467219
84. Niu C, Wang G, Yan P, et al. Noise Entangled GAN For Low-Dose CT Simulation. *ArXiv*. Published online 2021. <http://arxiv.org/abs/2102.09615>
85. Britten AJ, Crotty M, Kiremidjian H, Grundy A, Adam EJ. The addition of computer simulated noise to investigate radiation dose and image quality in images with spatial correlation of statistical noise: An example application to X-ray CT of the brain. *British Journal of Radiology*. 2004;77(916):323-328. doi:10.1259/bjr/78576048
86. Deák Z, Grimm JM, Treitl M, et al. Filtered back projection, adaptive statistical iterative reconstruction, and a model-based iterative reconstruction in abdominal CT: An experimental clinical study. *Radiology*. 2013;266(1):197-206. doi:10.1148/radiol.12112707
87. Singh R, Digumarthy SR, Muse V V., et al. Image Quality and Lesion Detection on Deep Learning Reconstruction and Iterative Reconstruction of Submillisievert Chest and Abdominal CT. *American Journal of Roentgenology*. 2020;214(3):566-573. doi:10.2214/AJR.19.21809
88. Lehtinen J, Munkberg J, Hasselgren J, et al. Noise2Noise: Learning image restoration without clean data. In: *35th International Conference on Machine Learning, ICML 2018*. Vol 7.; 2018:4620-4631.
89. Kingma DP, Ba JL. Adam: A method for stochastic optimization. *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings*. Published online 2015:1-15.
90. Contributors W. Anisotropic diffusion. Accessed October 16, 2023. https://en.wikipedia.org/wiki/Anisotropic_diffusion
91. Koike Y, Ishida K, Hase S, et al. Dynamic volumetric CT angiography for the detection and classification of endoleaks: Application of cine imaging using a 320-row CT scanner with 16-cm detectors. *Journal of Vascular and Interventional Radiology*. 2014;25(8):1172-1180.e1. doi:10.1016/j.jvir.2014.03.019

92. Bent CL, Jaskolka JD, Lindsay TF, Tan K. The use of dynamic volumetric CT angiography (DV-CTA) for the characterization of endoleaks following fenestrated endovascular aortic aneurysm repair (f-EVAR). *J Vasc Surg*. 2010;51(1):203-206. doi:10.1016/j.jvs.2009.07.101
93. García-Figueiras R, Goh VJ, Padhani AR, et al. CT perfusion in oncologic imaging: A useful tool? *American Journal of Roentgenology*. 2013;200(1):8-19. doi:10.2214/AJR.11.8476
94. Boedeker K. AiCE Deep Learning Reconstruction: Bringing the power of Ultra-High Resolution CT to routine imaging. [https://uk.medical.canon/wp-content/uploads/sites/8/2019/11/ White-paper-Kirsten-Boedeker.pdf](https://uk.medical.canon/wp-content/uploads/sites/8/2019/11/White-paper-Kirsten-Boedeker.pdf)
95. Li Z, Yu L, Leng S, et al. A robust noise reduction technique for time resolved CT. *Med Phys*. 2016;43(1):347-359. doi:10.1118/1.4938576
96. Kouchi T, Tanabe Y, Smit EJ, et al. Clinical application of four-dimensional noise reduction filtering with a similarity algorithm in dynamic myocardial computed tomography perfusion imaging. *International Journal of Cardiovascular Imaging*. 2020;36(9):1781-1789. doi:10.1007/s10554-020-01878-6
97. Sliwicka O, Swiderska-Chadaj Z, Snoeren M, et al. Multireader image quality evaluation of dynamic myocardial computed tomography perfusion imaging with a novel four-dimensional noise reduction filter. *Acta radiol*. 2023;64(3):999-1006. doi:10.1177/02841851221108804
98. Tsuneta S, Oyama-Manabe N, Kameda H, et al. Improvement of image quality on low-dose dynamic myocardial perfusion computed tomography with a novel 4-dimensional similarity filter. *Medicine (United States)*. 2020;99(26):E20804. doi:10.1097/MD.00000000000020804
99. Rava RA, Snyder K V., Mokin M, et al. Assessment of a Bayesian vitrea CT perfusion analysis to predict final infarct and penumbra volumes in patients with acute ischemic stroke: A comparison with RAPID. *American Journal of Neuroradiology*. 2020;41(2):206-212. doi:10.3174/ajnr.a6395
100. Murayama K, Smit EJ, Prokop M, et al. A Bayesian estimation method for cerebral blood flow measurement by area-detector CT perfusion imaging. *Neuroradiology*. 2023;65(1):65-75. doi:10.1007/s00234-022-03013-9
101. Lincke T, Zech CJ. Liver metastases: Detection and staging. *Eur J Radiol*. 2017;97:76-82. doi:10.1016/j.ejrad.2017.10.016
102. Oliver H, Jones C, Baron L, Federle P. Hypervascular Liver Metastases: Do Unenhanced and Hepatic Arterial Phase CT Images Affect Tumor Detection? *Radiology*. 1997;205(3):709-715. doi:10.1148/radiology.205.3.9393525
103. O'Connor JPB, Tofts PS, Miles KA, Parkes LM, Thompson G, Jackson A. Dynamic contrast-enhanced imaging techniques: CT and MRI. *British Journal of Radiology*. 2011;84(SPEC. ISSUE 2). doi:10.1259/bjr/55166688
104. Guyennon A. Perfusion characterization of liver metastases from endocrine tumors: Computed tomography perfusion. *World J Radiol*. 2010;2(11):449. doi:10.4329/wjr.v2.i11.449
105. Pandharipande P V., Krinsky GA, Rusinek H, Lee VS. Perfusion imaging of the liver: Current challenges and future goals. *Radiology*. 2005;234(3):661-673. doi:10.1148/radiol.2343031362
106. Oğul H, Kantarci M, Genç B, et al. Perfusion CT imaging of the liver: Review of clinical applications. *Diagnostic and Interventional Radiology*. 2014;20(5):379-389. doi:10.5152/dir.2014.13396
107. Kambadakone AR, Sahani DV. Body Perfusion CT: Technique, Clinical Applications, and Advances. *Radiol Clin North Am*. 2009;47(1):161-178. doi:10.1016/j.rcl.2008.11.003
108. Sauer TJ, Bejan A, Segars P, Samei E. Development and CT image-domain validation of a computational lung lesion model for use in virtual imaging trials. *Med Phys*. 2023;50(7):4366-4378. doi:10.1002/mp.16222

109. Sahbaee P, Segars WP, Marin D, Nelson RC, Samei E. The Effect of Contrast Material on Radiation Dose at CT: Part I. Incorporation of Contrast Material Dynamics in Anthropomorphic Phantoms. *Radiology*. 2017;283(3):739. doi:10.1148/RADIOL.2016152851
110. Sahbaee P, Abadi E, Segars WP, Marin D, Nelson RC, Samei E. The effect of contrast Material on radiation Dose at CT: Part II. A Systematic Evaluation across 58 Patient Models. *Radiology*. 2017;283(3):749-757. doi:10.1148/RADIOL.2017152852
111. Archer SG, Gray BN. Vascularization of small liver metastases. *British Journal of Surgery*. 1989;76(6):545-548. doi:10.1002/bjs.1800760607
112. Breedis C, Young G. The blood supply of neoplasms in the liver. *Am J Pathol*. 1954;30(5):969-985.
113. Taylor I, Bennett R, Sherriff S. The blood supply of colorectal liver metastases. *British Journal of Cancer* 1978 38:6. 1978;38(6):749-756. doi:10.1038/bjc.1978.283
114. Tunissen SAM, Oostveen LJ, Moriakov N, et al. Development, validation, and simplification of a scanner-specific CT simulator. *Med Phys*. 2023;51(3):1-15. doi:10.1002/mp.16679
115. Jadick G, Abadi E, Harrawood B, Sharma S, Segars WP, Samei E. A scanner-specific framework for simulating CT images with tube current modulation. *Phys Med Biol*. 2021;66(18). doi:10.1088/1361-6560/ac2269
116. Herman GT. Correction for beam hardening in computed tomography. *Phys Med Biol*. 1979;24(1):81-106. doi:10.1088/0031-9155/24/1/008
117. McCollough CH, Yu L, Kofler JM, et al. Degradation of CT low-contrast spatial resolution due to the use of iterative reconstruction and reduced dose levels. *Radiology*. 2015;276(2):499-506. doi:10.1148/radiol.15142047
118. Racine D, Brat HG, Dufour B, et al. Image texture, low contrast liver lesion detectability and impact on dose: Deep learning algorithm compared to partial model-based iterative reconstruction. *Eur J Radiol*. 2021;141:109808. doi:10.1016/j.ejrad.2021.109808
119. Wintermark M. Brain perfusion-CT in acute stroke patients. *European Radiology, Supplement*. 2005;15(4):28-31. doi:10.1007/s10406-005-0112-y
120. Mayer TE, Hamann GF, Baranczyk J, et al. Dynamic CT perfusion imaging of acute stroke. *American Journal of Neuroradiology*. 2000;21(8):1441-1449.
121. Demeestere J, Scheldeman L, Cornelissen SA, et al. Alberta stroke Program early CT score versus computed tomographic perfusion to predict functional outcome after successful reperfusion in acute ischemic stroke. *Stroke*. 2018;49(10):2361-2367. doi:10.1161/STROKEAHA.118.021961
122. Jovin TG, Saver JL, Ribo M, et al. Diffusion-weighted imaging or computerized tomography perfusion assessment with clinical mismatch in the triage of wake up and late presenting strokes undergoing neurointervention with Trevo (DAWN) trial methods. *International Journal of Stroke*. 2017;12(6):641-652. doi:10.1177/1747493017710341
123. Albers GW, Lansberg MG, Kemp S, et al. A multicenter randomized controlled trial of endovascular therapy following imaging evaluation for ischemic stroke (DEFUSE 3). *International Journal of Stroke*. 2017;12(8):896-905. doi:10.1177/1747493017701147
124. Fieselmann A, Kowarschik M, Ganguly A, Hornegger J, Fahrig R. Deconvolution-based CT and MR brain perfusion measurement: Theoretical model revisited and practical implementation details. *Int J Biomed Imaging*. 2011;2011. doi:10.1155/2011/467563
125. Wu O, Østergaard L, Weisskoff RM, Benner T, Rosen BR, Sorensen AG. Tracer arrival timing-insensitive technique for estimating flow in MR perfusion-weighted imaging using singular value decomposition with a block-circulant deconvolution matrix. *Magn Reson Med*. 2003;50(1):164-174. doi:10.1002/mrm.10522

126. Kim I, Kang H, Yoon HJ, Chung BM, Shin NY. Deep learning–based image reconstruction for brain CT: improved image quality compared with adaptive statistical iterative reconstruction-Veo (ASIR-V). *Neuroradiology*. 2021;63(6):905-912. doi:10.1007/s00234-020-02574-x
127. Divel SE, Christensen S, Segars WP, Lansberg MG, Pelc NJ. A dynamic simulation framework for CT perfusion in stroke assessment built from first principles. *Med Phys*. 2021;48(7):3500-3510. doi:10.1002/mp.14887
128. Qutbi M. Quantitative Performance Evaluation of Commonly Used Colormaps for Image Display in Myocardial Perfusion Imaging: Analysis based on Perceptual Metrics. *Mol Imaging Radionucl Ther*. 2024;33(2):94-105. doi:10.4274/mirt.galenos.2024.34711
129. Silva S, Sousa Santos B, Madeira J. Using color in visualization: A survey. *Computers and Graphics (Pergamon)*. 2011;35(2):320-333. doi:10.1016/j.cag.2010.11.015
130. Tunissen SAM, Smit EJ, Mikerov M, Prokop M, Sechopoulos I. Performance evaluation of a 4D similarity filter for dynamic CT angiography imaging of the liver. *Med Phys*. 2024;51(12):8814-8827. doi:10.1002/mp.17394
131. Mikerov M, Tunissen SAM, Michielsen K, Smit EJ, Sechopoulos I. Adaptation of a 4D noise filter for implementation on GPU. In: *8th International Conference on Image Formation in X-Ray Computed Tomography*; 2024. www.ct-meeting.org/data/ProceedingsCTMeeting2024.pdf



Summaries

Summary – English

CT perfusion is a promising imaging technique, that can provide information about blood flow in the body by obtaining a series of standard CT images. Information about the flow, i.e., perfusion, has shown to be able to improve detection of lesions, staging of diseases, and prediction of treatment response. Making CT perfusion a potentially powerful tool in the clinic. The disadvantage of CT perfusion is however, that the technique uses multiple CT scans in a short time period, and thus comes with a relatively high radiation dose for the patient. The clinical implementation of CT perfusion is limited, because of this high radiation dose. To overcome the high radiation dose, one could obtain all scans in a CT perfusion protocol at a low dose. This will however lead to high noise content and thus low image quality. This thesis tries to overcome the limitation of high noise content in low dose CT perfusion scans. We investigated and developed algorithms to reduce noise content without sacrificing details. Our focus was mainly on a novel 4-dimensional filtering technique, called 4-dimensional Similarity Filter (4DSF). The 4DSF was characterized, validated, and optimized in this thesis. To do so, we developed simulation tools to enable quantitative validation, since clinical CT perfusion data is limited.

Chapter 2 and **3** both present a method to simulate low dose CT scans. **Chapter 2** simulated the principles of the CT scanner to obtain CT images from digital phantoms. **Chapter 3** uses deep learning to transform a normal dose CT image into a low dose CT image. The simulated CT images can be used to validate processing algorithms like the 4DSF. These tools make CT data more accessible, since clinical data is not always available or easy to obtain. The biggest upside of simulated data is the availability of the ground truth, i.e., the real voxel values, which allows for quantitative analysis. **Chapter 4** presents an elaborate analysis and validation of the 4DSF for liver CT perfusion. The results of this chapter showed the potential of the 4DSF, especially combined with another filter, to greatly reduce noise. In **Chapter 5** a GPU version of the 4DSF is presented. This version reduced computation time and allowed for easier modifications to the algorithm. Lastly in **Chapter 6**, a modified version of the 4DSF is presented for ischemic stroke. This version improved the visibility of the stroke in perfusion maps, potentially leading to more accurate diagnoses.

To summarize, this thesis provides insight and understanding of the 4DSF its behavior, strengths, and weaknesses. Next to this, it gives a general approach to optimize the filters use when new CT perfusion protocols are introduced. I hope we have brought CT perfusion one step closer to clinical use with these contributions.

Samenvatting – Nederlands

CT perfusie is een veelbelovende beeldvormingstechniek, die informatie over de doorbloeding van verschillende weefsels in het lichaam kan geven door het maken van een reeks CT beelden. Informatie over de doorbloeding, of in andere woorden, perfusie, is gerelateerd aan verbeterde leasie herkenning, stadiëring van ziektes en voorspelling van de behandelrespons. Hierdoor kan CT perfusie een zeer belangrijk hulpmiddel zijn in de kliniek. Het nadeel van CT perfusie is echter dat de techniek meerdere CT scans maakt in een korte periode en dus voor relatief veel straling zorgt. In de kliniek wordt CT perfusie tot op heden weinig gebruikt, omdat de straling relatief hoog is. Om deze hoge straling tegen te gaan kan men de scans in CT perfusie protocollen verkrijgen met een lage dosis. Dit leidt echter tot veel ruis en dus een lagere beeld kwaliteit. Dit proefschrift probeert de limitaties van deze hoge ruis in de CT perfusie beelden tegen te gaan. We hebben algoritmes onderzocht en ontwikkeld om ruis te onderdrukken zonder details te verliezen. Onze focus lag voornamelijk op de nieuwe 4-dimensionale filter techniek, genaamd 4-dimensionale Similarity Filter (4DSF). De 4DSF is gekarakteriseerd, gevalideerd en geoptimaliseerd in dit proefschrift. Om dit mogelijk te maken hebben we onder andere simulatie instrumenten ontwikkeld, omdat klinische data vrij beperkt beschikbaar is en het de mogelijkheid biedt tot kwantitatieve analyse.

Hoofdstuk 2 en **3** stellen beide een methode voor om lage dosis CT beelden te simuleren. **Hoofdstuk 2** simuleert de principes van de CT scanner om CT beelden te genereren van digitale fantomen. **Hoofdstuk 3** gebruikt kunstmatige intelligentie om normale dosis CT beelden om te vormen tot lage dosis CT beelden. De gesimuleerde CT beelden kunnen gebruikt worden om algoritmes zoals de 4DSF te valideren. Deze instrumenten maken het verkrijgen van CT beelden toegankelijker, omdat klinische beelden niet altijd makkelijk te verkrijgen of beschikbaar zijn. Het grootste voordeel van gesimuleerde data is de toegang tot de grondwaarheid, oftewel, de echte voxel waarden, dit maakt het mogelijk om kwantitatieve analyses te doen. **Hoofdstuk 4** presenteert een uitgebreide analyse en validatie van de 4DSF voor lever CT perfusie. De resultaten van dit hoofdstuk laten de potentie van de 4DSF zien, in het bijzonder als de 4DSF worden gecombineerd met een andere filter om zo de ruis verder te onderdrukken. In **Hoofdstuk 5** wordt een GPU versie van de 4DSF voorgesteld. Deze versie dringt de rekentijd van het algoritme terug en maakt het makkelijker het algoritme aan te passen. Als laatste wordt in **Hoofdstuk 6** een aangepaste versie van het 4DSF voorgesteld voor herseninfarcten. Deze versie verbetert de zichtbaarheid van de beroerte in de perfusie mappen, wat mogelijk tot betere diagnostisering leidt.

Dit proefschrift biedt inzicht in de 4DSF en helpt ons het gedrag, de sterke en de zwakke punten van het filter te begrijpen. Bovendien wordt een algemene manier voorgesteld voor het optimaliseren van het filter wanneer men het wil gebruiken voor een nieuw CT perfusie protocol. Door deze bijdrages hoop ik dat we CT perfusie een stap dichterbij klinisch gebruik hebben gebracht.



Appendix

List of Publications

Papers in international Journals

- **Sjoerd A. M. Tunissen**, Luuk J. Oostveen, Nikita Moriakov, Jonas Teuwen, Koen Michielsen, Ewoud J. Smit, and Ioannis Sechopoulos. Development, validation, and simplification of a scanner-specific CT simulator. *Medical Physics*, 2024, 51 (3), 2081-2095. doi.org/10.1002/mp.16679
- **Sjoerd A. M. Tunissen**, Nikita Moriakov, Mikhail Mikerov, Ewoud J. Smit, Ioannis Sechopoulos, and Jonas Teuwen. Deep learning-based low-dose CT simulator for non-linear reconstruction methods. *Medical Physics*, 2024, 51 (9), 6046-6060. doi.org/10.1002/mp.17232
- **Sjoerd A. M. Tunissen**, Ewoud J. Smit, Mikhail Mikerov, Mathias Prokop, and Ioannis Sechopoulos. Performance evaluation of a 4D similarity filter for dynamic CT angiography imaging of the liver. *Medical Physics*, 2024, 51 (12), 8814-8827. doi.org/10.1002/mp.17394
- Luuk J Oostveen, **Sjoerd A. M. Tunissen**, and Ioannis Sechopoulos. Comparing organ and effective dose of various CT localizer acquisition strategies: a Monte Carlo study. *Medical Physics*, 2025, 52 (1), 576-584. doi.org/10.1002/mp.17447
- **Sjoerd A. M. Tunissen**, Ewoud J. Smit, Mikhail Mikerov, and Ioannis Sechopoulos. 4D Similarity Filter for Cerebral CT perfusion in Acute Stroke. *Radiology*, accepted.

Conference proceedings

- **Sjoerd A. M. Tunissen**, Luuk J. Oostveen, Nikita Moriakov, Jonas Teuwen, Koen Michielsen, Ewoud J. Smit, and Ioannis Sechopoulos. Development and validation of a clinical CT system simulator. In: *Proc. SPIE 12031, Medical Imaging 2022: Physics of Medical Imaging*, 120312G (4 April 2022). doi.org/10.1117/12.2611323
- **Sjoerd A. M. Tunissen**, Andrea Motta, Franziska Mauter, Eloy García, Oliver Diaz, John M. Boone, Ioannis Sechopoulos, and Marco Caballo. End-to-end mammographic breast density quantification with deep learning: preliminary study on simulated mammograms. In: *Proc. SPIE 12465, Medical Imaging 2023: Computer-Aided Diagnosis*, 124650K (7 April 2023). doi.org/10.1117/12.2654131
- Mikhail Mikerov, **Sjoerd A. M. Tunissen**, Koen Michielsen, Ewoud J. Smit, and Ioannis Sechopoulos. Adaptation of a 4D noise filter for implementation on GPU. *8th International Conference on Image Formation in X-Ray Computed Tomography*, in press.

Portfolio

Department: Medical Imaging
PhD period: 17/02/2020 – 31/08/2024
PhD Supervisor: prof. dr. Ioannis Sechopoulos, prof. dr. Mathias Prokop
PhD Co-supervisor: dr. ir. Ewoud J. Smit

Training activities	Hours
Courses	
• Radboudumc – Introduction Day	6
• Machine learning via Coursera from Stanford University	60
• Deep Learning Specialization via Coursera	80
• RIHS – Introduction Course for PhD candidates	15
• RU – Project management for PhD candidates	56
• RU – Scientific Writing for PhD candidates	84
• Radboudumc – Scientific Integrity	20
• RU – The Art of Presenting Science	36
Seminars	
• AXTI research meeting (2020 – 2024) *	160
• BIG-AXTI research meeting (2020 – 2024) *	40
• Medical Imaging Department research meeting (2020 – 2024) *	80
Conferences	
• 2021 AAPM Annual Meeting & Exhibition (2021) ‡	30
• SPIE Medical Imaging (2022) ‡	30
• Cancer Research Retreat (2022) ‡	12
• RIHS PhD retreat (2022) *	10
• SPIE Medical Imaging (2023) *	32
• European Congress of Radiology (2023) ‡ ‡	36
• European Congress of Radiology (2024) *	36
Teaching activities	Hours
Lecturing	
• CT lecture, Visualizing health and disease MIN02 (2020 – 2023)	24
• CT lecture, Biomedical imaging: Seeing is understanding (2022 – 2024)	10
Supervision	
• Master student supervision (Andrea Motta, Feb-Aug 2022)	78
Total	935

* Oral presentation

‡ Poster presentation

Research Data Management

In this thesis different types of data were used consisting of phantom images, simulated data, and retrospective patient images and respective imaging information.

All data has been stored on AXTI group servers belonging to the Department of Medical Imaging. All data archives, measured, raw, processed, and simulated images are stored and accessible by the associated staff members of the AXTI group. To ensure the interpretability and reusability of the data, documentation has been added to the data.

The retrospective patient images and respective imaging information used in Chapter 3 and 6 were obtained following the CMO2016-3045 (Umbrella protocol), project 20031. This is a protocol approved by the medical ethical committee, the Netherlands. In both cases the data was retrieved by using the RADNG anonymization server, to ensure no personal information that could identify individuals was included. This anonymization was only on metadata level for internal use; however, the data itself could still be identifiable via face recognition. The data will be archived for a period of 5 years before destruction, according to the CMO2016-3045 (Umbrella protocol).

The retrospective patient data is not shared publicly as it is not allowed according to the protocol we used to retrieve it. We are also not allowed to publicly share the simulated data, because of vendor agreements. We have added a metadata repository to the Radboud Data Repository about the tools we used to simulate our data, which has the following DOI: <https://doi.org/10.34973/4zgg-5j18>

Curriculum Vitae

Sjoerd Tunissen was born in Weert, The Netherlands, on November 11, 1995. He grew up in Roggel, a village in Limburg. After finishing high school in 2014, at Sint Ursula Horn, he started studying Mechanical Engineering at Eindhoven University of Technology. In 2017 he obtained his bachelor's degree and in 2020 his master's degree, both in Mechanical Engineering. His master thesis focused on heartbeat and breathing motion compensation for a robotic in-eye surgical system using optical coherence tomography (OCT) distance measurements. This project was performed in cooperation with the start-up company Preceyes B.V., located at the University campus at the time.

In February 2020, he joined the Advanced X-ray Tomographic Imaging (AXTI) group as a PhD candidate within the Department of Medical Imaging, in Radboud University Medical Center, Nijmegen. During his PhD research he focused on improving image quality of low dose Computed Tomography perfusion data. Mainly doing so by exploring the possibilities of a novel new 4-dimensional filtering method, called the 4D Similarity Filter. During these projects, he collaborated with Canon Medical Systems, supervised a master student, gave multiple lectures, and presented his work at various conferences. In October 2024, he started a new job as an Application Engineer at the Customer Support department of ASML.

Dankwoord

De afgelopen vijf jaar waren een avontuur waarin ik veel geleerd, gelachen en soms ook een beetje gehuild heb. In dit hoofdstuk wil ik graag iedereen bedanken die mij direct of indirect geholpen heeft.

Laat ik beginnen bij mijn supervisie team. Ioannis, I would like to thank you in the first place for hiring me. I can still remember the moment you called to tell me that you would like to hire me. Thank you for giving me the opportunity to learn about medical imaging and the scientific world. During the past five years I have learned so much thanks to you. The most important lesson was of course how to give a presentation, which is something that will definitely stay with me. Next to this, I really appreciate that you were always available and made time for me when I really needed it. I always felt I could count on you when it mattered, thank you for that. Ewoud, dankjewel dat ik mocht werken aan jouw 'uitvinding'. Ik heb altijd erg bewonderd hoeveel inzicht jij hebt op het technische en medische vlak. Het was altijd erg prettig om samen te werken met jou. Bedankt dat je altijd tijd voor me maakte en even checkte hoe het ging, ook op de momenten dat het niet over de inhoud ging. Jonas, jouw inhoudelijke kennis over Deep Learning was bewonderenswaardig, bedankt dat ik hiervan heb mogen leren. Mathias, the number of ideas or suggestions you could provide me in a very short time and with very little context was always something that amazed me. Thank you for the opportunity to work with you.

To our collaborators Canon Medical Systems. I am grateful for all the input you provided me with during our meetings, this was always very insightful. And thank you, mostly to Ikeda-san, for helping me with all the things I needed to know about the Canon machines/software during my research. To our other collaborators, the Center for Virtual Imaging Trials at Duke University. I would like to especially thank Ehsan and Paul for having me and helping me with all my questions about the digital phantoms. On top of that, I would like to thank Majid, Saman, Han and Mojtaba for all the lunches and Ehsan for inviting me to your house.

To my AXTI colleagues, the ones that saw my journey from close by. Let me start with the old members that were already 'senior' when I started, Marco, Joana, and Wendelien. Marco you really helped me a lot, especially in the beginning and gave me great advice. Joana thank you for always being you. It was great to always have you there (30 minutes late) when we were going to do something no matter what, a

drink or play board games. I enjoy still being in contact with you. Wendelien, thank you for always being so early, so I was not the only one in the office until 9:30.

To the new ones Gustavo, Raneim, Hanne, Maranda, and Martina. Gustavo thank you for the great story on how your bike got stolen, I will never forget this. Raneim, one of the few ones to be earlier than me in the office. It was great to get to know you and I really enjoyed all the questions you had for me. Hanne, bedankt voor een geweldig verhaal over jouw supermarkt tripje op zondagavond. Ook vond ik het leuk om met jou les te geven. Maranda, dankjewel voor al het lachen op kantoor en op ECR en voor alle koppen thee die we 's ochtends hebben gedronken bij de koffie automaat als de rest er nog niet was. Ook was het samen wielrennen en spinnen altijd gezellig. Martina, thank you for staying in the office with me when I needed help, for all the laughs, and the Italian saying about the wolfs. Thank you, for the television night at ECR with 'Wer ist es?'. To the ones abroad, Noelia and Franziska. Noelia, thank you for joining us at SPIE and visiting us in Nijmegen, it was always nice to see you. Franziska, thank you for the visits and nice talks in German, so I could practice my German a bit.

To the ones that were there for the entire journey. Olga, thank you for inviting me to all the places you have lived in Nijmegen. It was very nice to do all the Dutch language practicing with you. Sarah, dankjewel voor al je goede technische input en je hulp bij mijn appartementen jacht. Jessie, dankje dat je altijd op kantoor was en het lab altijd vulde met je goede lach en Brabantse gezelligheid. Het was fijn om nog iemand anders op kantoor te hebben met een zachte G. Luuk, degene aan wie ik altijd al mijn vragen kon stellen en wie mij altijd hielp met mijn metingen. Ik ga jouw flauwe humor niet vergeten. Mikhail, mijn vriend, bedankt dat je altijd eerlijk en recht voor zijn raap bent, voor alle keren dat je mij geholpen hebt met mijn onderzoek, en voor alle keren dat je naar mijn problemen hebt geluisterd en mij gerust hebt gesteld. Aan onze spinning en wielrenritten houd ik mooie herinneringen over. Iedere week weer het heuveltje bij Tivoli op en een snellere tijd proberen te halen of bijna overgeven na het sprinten in de spinning les. Dit zijn de dingen die ik mis aan Nijmegen; dit moeten we snel weer eensdoen in het mooie Limburg. Juan, mijn vriend en 4Daagse buddy, thanks for all the times you listened to my struggles about work or others things in life and the advice you gave me, this really helped me in hard times. I really enjoyed all the practice walks and talks we had and I am proud to say that we did 4Daagse together. And of course, it was always nice how you surprised all of us with the random Dutch words you knew. Marta, minha Amiga, tu és perfeita. Thank you for all the good times we had together, all the coffees we drank, and for organizing all the get-togethers. Our trip

together through California where we sat on the couch of Friends was the highlight of course. I am very happy I got to know you that well and thank you for helping me get through the hard times. And by the way, I still think you talk too much.

To the post-docs; Koen bedankt voor alle hulp bij de dingen die met programmeren of medische fysica te maken hadden. Daan, dankjewel voor je altijd eerlijke antwoorden en voor het hooghouden van het Nederlandse stereotype van 'direct' zijn. Je lach en energie om samen dingen te ondernemen buiten werk hebben het kantoor zeker gezelliger gemaakt. Maar zeker ook bedankt voor alle keren dat je me vertelde dat het goed zou komen. Leo, the runner and physicist. Thanks for always being open and joining for doing something fun. And thank you for reminding me that I am only an engineer and no physicist. I will never forget your competitiveness and thresh talking side, especially during the AXTI weekend. Uiteraard had ik het niet zonder Marjolein kunnen doen, dankjewel voor het altijd regelen van alles zodat wij ons op ons onderzoek konden concentreren. Lian, bedankt voor het beantwoorden van al mijn vragen over het verzamelen van data. And of course, my student Andrea, thank you for the opportunity to supervise you. It was a nice learning experience for me and most importantly great fun.

To my new colleagues, Gonçalo, Ruud, Jordi, and the others of the NTP team and Patterning group, thank you for making me feel welcome at my new job. It has honestly been a very warm and nice start.

Aan de mensen uit Limburg. Pieter, dankje voor alle wielrenritten op zondag in de corona tijd, dit moeten we snel weer oppakken. Paco, dankje voor het altijd kunnen chillen op zaterdagavond in corona tijd, de potjes schaken, toepen en andere spellen. Louis, dankje dat we altijd bij jou kunnen zitten en voor alle kampvuren. Lars, dankje voor je altijd droge humor en dat wanneer we samen zijn het altijd ouderwets gezellig. Sander, dankje voor je nuchterheid, het altijd in zijn voor wat te doen of drinken en je altijd sterke verhalen, ik ben blij dat we al zo lang vrienden zijn. Jeroen, mijn broeder van een andere moeder. Dankje dat je er altijd bent en zult zijn. Van kleins af aan zijn we al bij elkaar en het is fijn dat we nu nog steeds alles kunnen delen. Ik heb erg veel respect voor hoe jij in het leven staat, hier kan ik nog veel van leren.

Aan mijn familie. Pap, mam dankje voor alles wat jullie mij geleerd en meegegeven hebben. Pap, dankje dat je er altijd voor me bent als er iets moet gebeuren. Ik hoef maar iets te zeggen wat we misschien willen doen en jij staat al klaar met alles wat nodig is. De manier waarop jij klaar staat voor de mensen om je heen als die je iets

vragen is iets waar ik tegen op kijk. Mam, dankje dat je er altijd voor me bent als er iets is en ik altijd al mijn praat bij je kwijt kan. Ik heb enorm veel respect voor hoe jij altijd alles regelt en op de rit houdt en zorgt dat alles tip top in orde is voor pap, mij en met name Janna. Janna, dankje dat je altijd zo blij bent als ik thuis kom, door jou kan ik mijn eigen zorgen vaak vergeten. Ik ben trots op het feit dat jij mijn zusje bent en ben blij dat ik nu weer dichterbij je ben. Ook de rest van mijn familie Marian, Cor, Max, Fleur, Frieda, Joep, Jos, Doriënne en Pieter, dankje dat jullie er altijd voor me zijn en voor me klaar staan als er iets is.

Sophie, degene die mijn avontuur van dichtbij heeft meegemaakt en mij al meer dan elf jaar verblijdt. Ik weet niet hoe ik jou moet bedanken, zonder jou het ik dit avontuur misschien wel niet afgemaakt. Dankje dat je er altijd voor me was op de moeilijke momenten en je altijd precies wist wat je moest doen of zeggen om mij gerust te stellen, maar ook bedankt dat je er was op de mooie momenten. Van de afgelopen elf jaar met jou had ik geen moment willen missen. Het is fijn om te zien dat jij nu ook je plekje hebt gevonden en aan het laatste stukje van je lange weg naar dokter worden bent begonnen. Hopelijk, gaan we nu veel mooie herinneringen tegemoet in ons nieuwe huis.

