TECHNISCHE UNIVERSITÄT MÜNCHEN

Fakultät für Medizin

Institut für diagnostische und interventionelle Radiologie

Diagnostic evaluation strategies for novel CT developments

Felix Klaus Kopp

Vollständiger Abdruck der von der Fakultät für Medizin der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften (Dr. rer. nat.)

genehmigten Dissertation.

Vorsitzender:		Prof. Dr. Claus Zimmer
Prüfer der Dissertation:	1.	Priv.-Doz. Dr. Peter Noël
	2.	Prof. Dr. Franz Pfeiffer

Die Dissertation wurde am 18.07.2019 bei der Technischen Universität München eingereicht und durch die Fakultät für Medizin am 03.12.2019 angenommen.

Acknowledgments

During this doctoral project, I received an abundance of support. First of all, I want to thank Peter Noël for his great supervision. He introduced me to the interesting field of medical imaging and gave me the opportunity to participate in the doctoral program. Thank you for your excellent guidance. It was a great joy working with you for all these years. Furthermore, I would like to thank Franz Pfeiffer and Ernst Rummeny for the great and continuous support.

I want to thank my mentor Daniela Pfeiffer. She gave me great assistance in understanding the medical background and was a big help in developing the study designs. I want to thank my colleague Kai Mei for many fruitful discussions, the help with the reconstruction framework and his helpful advice during my doctoral project. A big thank you goes to Manuel Schultheiß for the great help with the installation of the machine-learning environment and the discussions concerning convolutional neural networks. I want to thank Julia Dangelmaier, Andreas Sauter and Thomas Baum for their great cooperation in so many projects. I would like to thank the physicists of the E17 chair for the great support for any technical question, my colleagues and former colleagues David Jany, Felix Gebser and Kai Scherer for the pleasant atmosphere in the office, Veronica Bodek for the great support in language editing and any kind of organizational task, and Bernhard Renger for the great help maintaining our computing servers.

Many parts of this doctoral project are based on work that was done in the context of the IMEDO project funded by the BMBF under cooperation with Philips Healthcare. I would like to thank Rolf Bippus and Heiner Daerr for their great collaboration.

Special thanks go to my parents and family who were always supportive and encouraging.

Finally, I want to express my deepest gratitude to my wife and my son. Nicole and Johannes, thank you for bringing so much joy into my life!

Vielen Dank!

Abstract

Advancements of existing imaging methods and the development of new imaging technologies require strategies for their clinical evaluation. Besides the technical requirements that a certain spatial resolution is provided or a certain limit of image noise is guaranteed, especially the diagnostic performance of radiologists is of interest. This diagnostic image quality or diagnostic value is usually investigated for specific clinical tasks. The aim of this cumulative thesis is to investigate and implement methods for the assessment of diagnostic image quality and the clinical evaluation of novel imaging methods and devices.

In the first journal publication, model observers–a strategy to perform virtual reader studies–were investigated. Neural networks and convolutional neural networks (CNNs) were used to implement novel model observers and compared to a more traditional model observer, the channelized Hotelling observer (CHO). Traditional model observers, like the CHO, are limited to easy signal detection tasks due to their mathematical definition. The goal of this work was the development of a model observer that can be applied to more realistic clinical tasks where the signal size is varying. Best performance was achieved with a CNN-based model observer that showed strong correlation with a human observer's performance in a lesion detection task with varying signal sizes (Pearson's product-moment correlation coefficient r=0.918).

A custom-made lung phantom was designed for the clinical evaluation of a spectral photon-counting CT (SPCCT) prototype. The evaluation was published in the second journal publication. A digital anthropomorphic lung phantom was created by segmentation of a healthy human lung from a clinical CT scan. The digital lung phantom was augmented with spheres and spheres with spikes to simulate benign and malignant lung nodules, respectively. A physical phantom was created by 3D-printing of the digital phantom. In the journal paper, the SPCCT prototype was evaluated regarding nodule quantification and in terms of high resolution capabilities. Compared to conventional CT in high-resolution mode (HR-CT), SPCCT showed superior results in nodule quantification (mean Dice similarity coefficient: 0.90 (SPCCT); 0.85 (HR-CT)) and volume estimation (root mean square error: 21.3 mm^3 (SPCCT); 26.4 mm^3 (HR-CT)).

In the third journal publication, simulated sparse-sampling computed tomography (SpSCT) was evaluated for examinations with suspected pulmonary embolism (PE). Four blinded radiologists participated in a reader study with 20 CT scans of patients with suspected PE. The evaluation aims to validate sparse sampling as dose reduction technology by simulating a fast pulsing x-ray tube. Technical principles of sparse sampling in the frame of CT are presented in this dissertation. Sensitivity, specificity, accuracy and area under the receiver operating characteristic curve were superior for SpSCT compared to conventional CT at investigated reduced dose levels. With SpSCT, a dose reduction by 87.5 % of the initial clinical dose was achievable for the investigated cases with high subjective image quality and high diagnostic performance.

This cumulative dissertation gives an introduction to clinical evaluation of the diagnostic value of CT imaging modalities and describes materials and methods that were essential for the realization of the above-mentioned journal publications. The contribution to the CT research community is discussed in the context of related work.

Zusammenfassung

Um Weiterentwicklungen bestehender Bildgebungsverfahren sowie Neuentwicklungen im Bereich der medizinischen Bildgebung zu evaluieren werden Strategien benötigt mit denen der klinische Nutzen bestimmt werden kann. Neben der Einhaltung von technischen Spezifikationen, wie beispielsweise eine bestimmte Bildauflösung oder einer oberen Grenze für Bildrauschen, ist insbesondere die Fähigkeit eines Radiologen eine richtige Diagnose auf den Bildern zu stellen von großem Interesse. Diese diagnostische Bildqualität oder diagnostischer Wert eines Bildes wird üblicherweise für bestimmte klinische Fragestellungen ermittelt. Ziel dieser publikationsbasierten Dissertation ist die Erforschung und Implementierung von Methoden zur Bestimmung der diagnostischen Bildqualität und die klinische Evaluierung neuartiger Bildgebungsverfahren.

In einem ersten Zeitschriftenartikel wird eine Strategie zur Durchführung virtueller radiologischer Studien, sogenannte Model Observer erforscht. Neuartige Model Observer wurden mit neuronalen Netzen und sogenannten "Convolutional Neural Networks" (CNN) implementiert und mit einem etablierten Model Observer, dem sogenannten "channelized Hotelling observer" (CHO) verglichen. Aufgrund ihrer mathematischen Definition sind etablierte Model Observer, wie der CHO, auf einfache Fragestellungen der Signalerkennung beschränkt. Ziel dieser Publikation war die Entwicklung eines neuartigen Model Observers, der auf realistischere Fragestellungen mit variierender Signalgröße angewendet werden kann. Die besten Ergebnisse wurden mit einem Model Observer erzielt, der auf einem CNN basiert. Die diagnostische Leistung korrelierte stark mit der eines menschlichen Beobachters in einer Fragestellung zur Erkennung von simulierten Leberläsionen verschiedener Größen (Pearson-Korrelation r=0.918).

Der klinische Nutzen eines spektralen photonenzählenden Computertomographen (SPCCT) wurde mit einem eigens angefertigten Lungenphantom evaluiert und in einem zweiten Zeitschriftenartikel veröffentlicht. Die Lunge eines gesunden Menschen wurde aus einer klinischen CT-Aufnahme segmentiert um ein digitales anthropomorphes Phantom eines Lungenflügels zu erstellen. Das digitale Phantom wurde mit Kugeln und Kugeln mit Spitzen erweitert um gutartige und bösartige Lungenknoten zu simulieren. Das digitale Phantom wurde mit einem 3D-Druckverfahren gedruckt. In dem Zeitschriftenartikel wurde der SPCCT-Prototyp hinsichtlich der Knotenbewertung und seiner Fähigkeiten in der hochauflösenden Bildgebung evaluiert. Im Vergleich zu einem konventionellen CT in hochauflösendem Betriebsmodus (HR-CT) erzielte der SPCCT-Prototyp bessere Ergebnisse bezüglich Knotenbewertung (mittlerer Dice similarity coefficient: 0.90 (SPCCT); 0.85 (HR-CT)) und Volumenbestimmung der Knoten (root mean square error: 21.3 mm^3 (SPCCT); 26.4 mm^3 (HR-CT)).

In einem dritten Zeitschriftentartikel wurde simulierte spärlich abgetastete Computertomographie (SpSCT) für die Erkennung von Lungenembolien bei reduzierter Strahlungsdosis untersucht. Vier Radiologen nahmen unter Verblindung des klinischen Befundes an einer Studie mit 20 CT-Aufnahmen von Patienten mit Verdacht auf Lungenembolie teil. Die Evaluierung zielt darauf, SpSCT als Technologie zur Dosisreduktion durch eine schnellschaltende Röntgenröhre zu validieren. Technische Merkmale von SpSCT werden in dieser Dissertation dargestellt. Mit SpSCT wurden für alle Dosisstufen bessere Werte für Sensitivität, Spezifität, Korrektklassifizierungsrate und Fläche unter der ROC Kurve gegenüber konventionellem CT erzielt. Für die untersuchten Fälle wäre eine Dosisreduktion um 87.5 % der ursprünglichen klinischen Dosis umsetzbar, bei hoher subjektiver Bildqualität und hoher diagnostischer Leistungsfähigkeit.

Diese publikationsbasierte Dissertation gibt eine Einführung in den Bereich der klinischen Evaluierung des diagnostischen Nutzens von Bildgebungsmodalitäten und beschreibt die Materialien und Methoden, die für die Umsetzung der eingebundenen Publikationen notwendig waren. Weiter wird der Beitrag zum CT-Forschungsgebiet im Kontext themenbezogener Arbeiten diskutiert.

List of Included Journal Publications

The present dissertation is based on the following three journal publications:

JP-I **F. K. Kopp**, M. Catalano, D. Pfeiffer, A. A. Fingerle, E. J. Rummeny, and P. B. Noël, *“CNN as model observer in a liver lesion detection task for x-ray computed tomography: A phantom study”*, Med Phys, vol. 45, no. 10, pp. 4439–4447, Oct. 2018. DOI: 10.1002/mp.13151.

JP-II **F. K. Kopp**, H. Daerr, S. Si-Mohamed, A. P. Sauter, S. Ehn, A. A. Fingerle, B. Brendel, F. Pfeiffer, E. Roessl, E. J. Rummeny, D. Pfeiffer, R. Proska, P. Douek, and P. B. Noël, *“Evaluation of a preclinical photon-counting CT prototype for pulmonary imaging”*, Sci Rep, vol. 8, no. 1, pp. 17386, Nov. 2018. DOI: 10.1038/s41598-018-35888-1.

JP-III A. P. Sauter$^+$, **F. K. Kopp**$^+$, R. Bippus, J. Dangelmaier, D. Deniffel, A. A. Fingerle, F. Meurer, D. Pfeiffer, R. Proska, E. J. Rummeny, and P. B. Noël, *“Sparse sampling computed tomography (SpSCT) for detection of pulmonary embolism: a feasibility study”*, Eur Radiol, First Online 09 May 2019. DOI: 10.1007/s00330-019-06217-5.

$^+$These authors contributed equally to this work

The above three journal publications are referred to as JP-I, JP-II and JP-III, respectively. Please refer to Chapter 5 for publication abstracts.

List of Related Publications

The author contributed also to the following subject-related journal publications (ordered by year of appearance):

J1 D. Muenzel, D. Bar-Ness, E. Roessl, I. Blevis, M. Bartels, A. A. Fingerle, S. Ruschke, P. Coulon, H. Daerr, **F. K. Kopp**, B. Brendel, A. Thran, M. Rokni, J. Herzen, L. Boussel, F. Pfeiffer, R. Proksa, E. J. Rummeny, P. Douek, and P. B. Noël, *"Spectral Photon-counting CT: Initial Experience with Dual–Contrast Agent K-Edge Colonography,"* Radiology, vol. 283, no. 3, pp. 723–728, Jun. 2016.

J2 M. R. K. Mookiah, T. Baum, K. Mei, **F. K. Kopp**, G. Kaissis, P. Foehr, P. B. Noël, J. S. Kirschke, and K. Subburaj, *"Effect of radiation dose reduction on texture measures of trabecular bone microstructure: an in vitro study,"* J Bone Miner Metab, vol. 36, no. 3, pp. 323-335, Apr. 2017.

J3 K. Mei, **F. K. Kopp**, R. Bippus, T. Köhler, B. J. Schwaiger, A. S. Gersing, A. Fehringer, A. Sauter, D. Münzel, F. Pfeiffer, E. J. Rummeny, J. S. Kirschke, P. B. Noël, and T. Baum, *"Is multidetector CT-based bone mineral density and quantitative bone microstructure assessment at the spine still feasible using ultra-low tube current and sparse sampling?,"* Eur Radiol, vol. 27, no. 12, pp. 5261–5271, Jun. 2017.

J4 K. Mei, B. J. Schwaiger, **F. K. Kopp**, S. Ehn, A. S. Gersing, J. S. Kirschke, D. Muenzel, A. A. Fingerle, E. J. Rummeny, F. Pfeiffer, T. Baum, and P. B. Noël, *"Bone mineral density measurements in vertebral specimens and phantoms using dual-layer spectral computed tomography,"* Sci Rep, vol. 7, no. 1, p. 17519, Dec. 2017.

J5 D. Muenzel, H. Daerr, R. Proksa, A. A. Fingerle, **F. K. Kopp**, P. Douek, J. Herzen, F. Pfeiffer, E. J. Rummeny, and P. B. Noël, *"Simultaneous dual-contrast multi-phase liver imaging using spectral photon-counting computed tomography: a proof-of-concept study,"* Eur Radiol Exp, vol. 1, no. 1, p. 25, Dec. 2017.

J6 S. Ehn, T. Sellerer, D. Muenzel, A. A. Fingerle, **F. K. Kopp**, M. Duda, K. Mei, B. Renger, J. Herzen, J. Dangelmaier, B. J. Schwaiger, A. Sauter, I. Riederer, M. Renz, R. Braren, E. J. Rummeny, F. Pfeiffer, and P. B. Noël, *"Assessment of quantification accuracy and image quality of a full-body dual-layer spectral CT system,"* J Appl Clin Med Phys, vol. 19, no. 1, pp. 204–217, Dec. 2017.

J7 M. R. K. Mookiah, A. Rohrmeier, M. Dieckmeyer, K. Mei, **F. K. Kopp**, P. B. Noel, J. S. Kirschke, T. Baum, and K. Subburaj, *"Feasibility of opportunistic osteoporosis screening in routine contrast-enhanced multi detector computed tomography (MDCT) using texture analysis,"* Osteoporos Int, vol. 29, no. 4, pp. 825–835, Jan. 2018.

J8 A. P. Sauter, **F. K. Kopp**, D. Münzel, J. Dangelmaier, M. Renz, B. Renger, R. Braren, A. A. Fingerle, E. J. Rummeny, and P. B. Noël, *"Accuracy of iodine quantification in dual-layer* *spectral CT: Influence of iterative reconstruction, patient habitus and tube parameters,"* Eur J Radiol, vol. 102, pp. 83–88, Mar. 2018.

J9 B. J. Schwaiger, A. S. Gersing, J. Hammel, K. Mei, **F. K. Kopp**, J. S. Kirschke, E. J. Rummeny, K. Wörtler, T. Baum, and P. B. Noël, *"Three-material decomposition with dual-layer spectral CT compared to MRI for the detection of bone marrow edema in patients with acute vertebral fractures,"* Skeletal Radiol, vol. 47, no. 11, pp. 1533–1540, May 2018.

J10 M. R. Mookiah, K. Subburaj, K. Mei, **F. K. Kopp**, J. Kaesmacher, P. Jungmann, P. Foehr, P. B. Noël, J. Kirschke, and T. Baum, *"Multidetector Computed Tomography Imaging: Effect of Sparse Sampling and Iterative Reconstruction on Trabecular Bone Microstructure,"* J Comput Assist Tomogr, vol. 42, no. 3, pp. 441–447, Jun. 2018.

J11 J. Dangelmaier, B. J. Schwaiger, A. S. Gersing, **F. K. Kopp**, A. Sauter, M. Renz, I. Riederer, R. Braren, D. Pfeiffer, A. Fingerle, E. J. Rummeny, and P. B. Noël, *"Dual layer computed tomography: Reduction of metal artefacts from posterior spinal fusion using virtual monoenergetic imaging,"* Eur J Radiol, vol. 105, pp. 195–203, Jun. 2018.

J12 A. S. Gersing, D. Pfeiffer, **F. K. Kopp**, B. J. Schwaiger, C. Knebel, B. Haller, P. B. Noël, M. Settles, E. J. Rummeny, and K. Woertler, *"Evaluation of MR-derived CT-like images and simulated radiographs compared to conventional radiography in patients with benign and malignant bone tumors,"* Eur Radiol, vol. 29, no. 1, pp. 13–21, Jun. 2018.

J13 A. S. Gersing, P. M. Jungmann, B. J. Schwaiger, J. Zarnowski, **F K Kopp**, S. Landwehr, M. Sauerschnig, G. B. Joseph, A. B. Imhoff, E. J. Rummeny, J. S. Kirschke, and T. Baum, *"Longitudinal changes in subchondral bone structure as assessed with MRI are associated with functional outcome after high tibial osteotomy,"* J ISAKOS, vol. 3, no. 4, pp. 205–212, Jun. 2018.

J14 J. Dangelmaier, D. Bar-Ness, H. Daerr, D. Muenzel, S. Si-Mohamed, S. Ehn, A. A. Fingerle, M. A. Kimm, **F. K. Kopp**, L. Boussel, E. Roessl, F. Pfeiffer, E. J. Rummeny, R. Proksa, P. Douek, and P. B. Noël, *"Experimental feasibility of spectral photon-counting computed tomography with two contrast agents for the detection of endoleaks following endovascular aortic repair,"* Eur Radiol, vol. 28, no. 8, pp. 3318-3325, Aug. 2018.

J15 N. Sollmann, K. Mei, B.J. Schwaige, rA.S. Gersing, **F.K. Kopp**, R. Bippus, C. Maegerlein, C. Zimmer, E.J. Rummeny, J.S. Kirschke, P.B. Noël, and T. Baum, *"Effects of virtual tube current reduction and sparse sampling on MDCT-based femoral BMD measurements,"* Osteoporos Int, vol. 29, no. 12, pp. 2685–2692, Aug. 2018.

J16 K. Mei, S. Ehn, M. Oechsner, **F. K. Kopp**, D. Pfeiffer, A. A. Fingerle, F. Pfeiffer, S. E. Combs, J. J. Wilkens, E. J. Rummeny, and P. B. Noël, *"Dual-layer spectral computed tomography: measuring relative electron density,"* Eur Radiol Exp, Aug 2018.

J17 D. Anitha, K. Mei, M. Dieckmeyer, **F. K. Kopp**, N. Sollmann, C. Zimmer, J. S. Kirschke, P. B. Noel, T. Baum, and K. Subburaj, *"MDCT-based Finite Element Analysis of Vertebral Fracture Risk: What Dose is Needed?,"* Clin Neuroradiol, Aug. 2018.

J18 K. Willer , A. A. Fingerle , L. B. Gromann, F. De Marco, J. Herzen, K. Achterhold, B. Gleich, D. Muenzel, K. Scherer, M. Renz, B. Renger, **F. K. Kopp**, F. Kriner, F. Fischer, C. Braun, S. Auweter, K. Hellbach, M. F. Reiser, T. Schroeter, J. Mohr, A. Yaroshenko, H. Maack, T.

Pralow, H. van der Heijden, R. Proksa, T. Koehler, N. Wieberneit, K. Rindt, E. J. Rummeny, F. Pfeiffer , and P. B. Noël, *"X-ray dark-field imaging of the human lung-A feasibility study on a deceased body,"* PLoS One, vol. 13, no. 9, Sep. 2018.

J19 D. Anitha, K. Subburaj, **F. K. Kopp**, K. Mei, P. Foehr, R. Burgkart, N. Sollmann, C. Maegerlein, J. Kirschke, Peter Noel, and Thomas Baum, *"Effect of Statistically Iterative Image Reconstruction on Vertebral Bone Strength Prediction Using Bone Mineral Density and Finite Element Modeling: A Preliminary Study,"* J Comput Assist Tomogr, Jan. 2019.

J20 A. Laugerette, B. J. Schwaiger, K. Brown, L. C. Frerking, **F. K. Kopp**, K. Mei, T. Sellerer, J. Kirschke, T. Baum, A. S. Gersing, D. Pfeiffer, A. A. Fingerle, E. J. Rummeny, R. Proksa, P. B. Noël, and Franz Pfeiffer, *"DXA-equivalent quantification of bone mineral density using dual-layer spectral CT scout scans,"* Eur Radiol, Feb. 2019.

J21 N. Sollmann, K. Mei, D. M. Hedderich, C. Maegerlein, **F. K. Kopp**, M. T. Löffler, C. Zimmer, E. J. Rummeny, J. S. Kirschke, T. Baum, and P. B. Noël, *"Multi-detector CT imaging: impact of virtual tube current reduction and sparse sampling on detection of vertebral fractures,"* Skeletal Eur Radiol, pp. 1–11, Mar 2019.

J22 A. P. Sauter, J. Hammel, S. Ehn, K. Achterhold, **F. K. Kopp**, M. A. Kimm, K. Mei, A. Laugerette, F. Pfeiffer, E. J. Rummeny, D. Pfeiffer, and P. B. Noël, *"Perfusion-ventilation CT via three-material differentiation in dual-layer CT: a feasibility study,"* Sci Rep, vol. 9, no. 1, p. 5837, Apr. 2019.

J23 N. Sollmann, K. Mei, I. Riederer, S. Schön, **F. K. Kopp**, M. T. Löffler, D. M. Hedderich, C. Maegerlein, B. Friedrich, K. Kreiser, T. Boeckh-Behrens, J. S. Kirschke, E. J. Rummeny, C. Zimmer, P. B. Noël, and T. Baum, *"Tube Current Reduction in CT Angiography: How Low Can We Go in Imaging of Patients With Suspected Acute Stroke?,"* Am J Roentgenol, Apr. 2019.

The author contributed also to the following conference abstracts (ordered by year of appearance):

C1 K. Mei, **F. K. Kopp**, A. Fehringer, F. Pfeiffer, E. J. Rummeny, J. S. Kirschke, P. B. Noël, and T. Baum, Effects of sparse sampling in combination with iterative reconstruction on quantitative bone microstructure assessment, Proc. SPIE 10132, Medical Imaging 2017: Physics of Medical Imaging, 1013244 (9 March 2017), Orlando, FL, USA, 2017 (poster presentation)

C2 **F. K. Kopp**, M. Catalano, D. Pfeiffer, E. J. Rummeny, and P. B. Noël, Evaluation of a machine learning based model observer for x-ray CT, Proc. SPIE 10577, Medical Imaging 2018: Image Perception, Observer Performance, and Technology Assessment, 105770S (7 March 2018), Houston, TX, USA, 2018 (oral presentation)

C3 **F. K. Kopp**, R. Bippus, A. P. Sauter, D. Muenzel, F. Bergner, K. Mei, J. Dangelmaier, B. J. Schwaiger, M. Catalano, A. A. Fingerle, E. J. Rummeny, and P. B. Noël, Diagnostic value of sparse sampling computed tomography for radiation dose reduction: initial results, Proc. SPIE 10573, Medical Imaging 2018: Physics of Medical Imaging, 1057340 (9 March 2018), Houston, TX, USA, 2018 (poster presentation)

C4 K. Mei, B. J. Schwaiger, **F. K. Kopp**, S. Ehn, A. S. Gersing, J. S. Kirschke, D. Munzel, A. A. Fingerle, E. J. Rummeny, F. Pfeiffer, T. Baum, and P. B. Noël, Calcium decomposition and phantomless bone mineral density measurements using dual-layer-based spectral computed tomography, Proc. SPIE 10573, Medical Imaging 2018: Physics of Medical Imaging, 105731H (9 March 2018), Houston, TX, USA, 2018 (oral presentation)

C5 K. Mei, **F. K. Kopp**, J. Hammel, B. J. Schwaiger, A. S. Gersing, T. Baum, E. J. Rummeny, and P. B. Noël, Advancements in computed tomography for musculoskeletal imaging, The Fifth International Conference on Image Formation in X-Ray Computed Tomography, pp. 310-313, Salt Lake City, UT, USA, 2018 (poster presentation)

C6 **F. K. Kopp**, R.-D. Bippus, K. Mei, A. Sauter, B. J. Schwaiger, A. S. Gersing, J. Dangelmaier, E. J. Rummeny, and P. B. Noel, Endoleak detection with ultra-low radiation exposure CT based on sparse sampling and reduced tube current, European Congress of Radiology (ECR), Vienna, Austria, March 3rd, 2018 (oral presentation)

C7 **F. K. Kopp**, M. Catalano, D. Pfeiffer, A. A. Fingerle, E. J. Rummeny, and P. B. Noël, Task-based image quality assessment of X-ray CT using convolutional neural networks, RSNA 104th Scientific Assembly and Annual Meeting, Chicago, IL, USA, November 11th, 2018 (oral presentation)

C8 **F. K. Kopp**, K. Mei, E. J. Rummeny, and P. B. Noël, Sparse-sampling computed tomography for pulmonary imaging, Proc. SPIE 10948, Medical Imaging 2019: Physics of Medical Imaging, 109482X (1 March 2019), San Diego, CA, USA, 2019 (poster presentation)

C9 A. Sauter, **F. K. Kopp**, J. Dangelmaier, D. Deniffel, F. Meurer, R. Bippus, E. J. Rummeny, and P. B. Noel, Sparse-sampling computed tomography (SpSCT) for the detection of pulmonary embolism, European Congress of Radiology (ECR), Vienna, Austria, February 27th, 2019 (oral presentation)

Contents

Acknowledgments I

Abstract III

Zusammenfassung V

List of Included Journal Publications VII

List of Related Publications IX

1 Introduction 3
1.1 Clinical Motivation . . . 3
1.2 Thesis Purpose . . . 4
1.3 Thesis Structure . . . 5

2 Technical Principles 7
2.1 Computed Tomography . . . 7
2.1.1 X-ray Generation . . . 7
2.1.2 Data Acquisition . . . 8
2.1.3 Radiation Dose . . . 10
2.2 Sparse-Sampling CT . . . 12
2.2.1 Background . . . 12
2.2.2 Principle . . . 13
2.2.3 Technical Implementation . . . 14
2.2.4 Possible Further Applications . . . 14
2.3 Photon-Counting CT . . . 16
2.3.1 Principle . . . 16
2.3.2 Advancements . . . 17
2.3.3 Challenges . . . 18
2.4 CT Reconstruction . . . 19
2.4.1 Analytical Reconstruction . . . 20
2.4.2 Iterative Reconstruction . . . 21

3 Strategies for System Evaluation 25
3.1 Reader Studies . . . 25
3.1.1 Binary Performance Classifiers . . . 25
3.1.2 Receiver Operating Characteristic . . . 26
3.2 Model Observers . . . 29
3.2.1 Channelized Hotelling Observer . . . 29
3.2.2 Machine Learning Model Observer . . . 30
3.3 Custom Phantoms for Specific Applications . . . 33
3.3.1 Design of a Custom Lung Phantom . . . 34

3.3.2 Evaluation . . . 35

4 Compliance with Ethical Standards 39

5 Comprising Journal Publications 41
5.1 Journal Publication I . . . 41
5.1.1 Abstract . . . 41
5.1.2 Author contributions . . . 42
5.2 Journal Publication II . . . 43
5.2.1 Abstract . . . 43
5.2.2 Author contributions . . . 43
5.3 Journal Publication III . . . 44
5.3.1 Abstract . . . 44
5.3.2 Author contributions . . . 44

6 Discussion 47
6.1 JP-I . . . 47
6.2 JP-II . . . 48
6.3 JP-III . . . 49
6.4 Outlook and Conclusion . . . 50

List of Abbreviations 53

List of Figures 55

List of Tables 57

Bibliography 59

1 Introduction

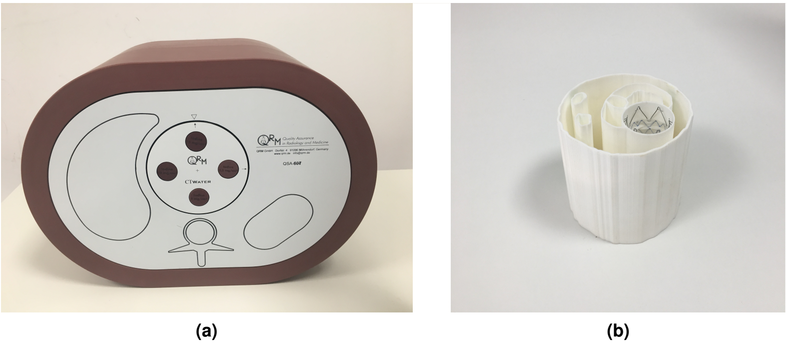
1.1 Clinical Motivation

Research and development are valuable divisions in companies producing medical imaging devices. Improvements of existing modalities as well as the development of new and innovative imaging methods guarantee the success of many companies. However, benefits and clinical usage of new developments have to be evaluated before they become available in the clinical practice. Important aspects are especially the diagnostic value and the clinical applicability. In x-ray computed tomography (CT), one additional issue is the applied radiation dose. Ionizing radiation is known for its harmful characteristics at high doses [1–3]. Therefore, image quality and diagnostic value of new developments need to be evaluated with respect to the radiation dose.

For the clinical evaluation of imaging devices, in a first step, physical and technical properties need to be met [4–6]. For a CT system, beyond others, there are usually requirements like short scan times and a sufficient detector coverage in the table direction (z-direction) for cardiac imaging or high spatial resolution for pulmonary imaging. Those physical requirements to a CT system guarantee a high quality of the acquisition data. However, image quality highly depends on the applied radiation dose and the reconstruction chain [7]. In general, a reduction of the radiation dose leads to an increased amount of noise in the acquired measurement data. With the use of conventional analytical image reconstruction, the low measurement quality is mostly linearly translated into the image space, allowing a prediction of the image quality before any reconstruction. With the introduction of iterative reconstruction, image noise and artifacts are reduced by methods that introduce non-linearity to the reconstruction chain, e.g. edge-preserving regularization in statistical iterative reconstruction. As a result, objective metrics for image quality measurements, such as the modulation transfer function (MTF) for the evaluation of the spatial resolution, may highly depend on the acquired object and specific image properties (e.g. contrast or location in the image) [8, 9].

The assessment of diagnostic image quality, or the diagnostic value of an image, is usually based on reader studies where radiologists perform a diagnostic task [10]. The diagnostic performance of the radiologists for a specific imaging modality or image quality gives the diagnostic value for the specific task. There is a high interest in metrics that allow an automated assessment of the diagnostic image quality because reader studies are very time-consuming and costly. One research area that focuses on automated task-based image quality assessment is the field of model observers, where computer models are implemented to mimic radiologists' performances [11, 12].

Another strategy to evaluate new developments for their clinical value is the investigation of phantom studies. High-precision phantoms are traditionally used to evaluate the physical and technical properties [6]. For task-based evaluations of the diagnostic value, anthropomorphic phantoms can be used (Fig. 1.1) [13–16]. Furthermore, there are commercial phantoms available that are equivalent to the human body in terms of shape and x-ray attenuation. For more specific tasks, the production of custom phantoms has the advantage that they can be designed to meet the individual needs. The broad availability of 3D printing systems has simplified the production of custom phantoms and, hence, allows the investigation of specific diagnostic tasks in phantom studies without exposing patients to ionizing radiation.



(a) **(b)**

Figure 1.1: Anthropomorphic phantoms.

(a) Commercial anthropomorphic abdomen phantom with a cylindrical insert and rods of different iodine concentrations, e.g. for the evaluation of iodine quantification with a dual-energy CT (DECT) [13, 14]. A small extension ring is added to the phantom to simulate obesity. (b) 3D-printed phantom of an aortic aneurysm simulating the aorta after an endovascular aortic repair (EVAR), e.g. for the simulation of an endoleak [15].

1.2 Thesis Purpose

The overall purpose of the present dissertation is the implementation of strategies for the clinical evaluation of new reconstruction algorithms and innovative CT systems like spectral photon-counting CT (SPCCT) and sparse-sampling CT (SpSCT) and their application.

In the first part of this doctoral project, automated evaluation of the diagnostic value is investigated. State-of-the-art model observers like the channelized Hotelling observer (CHO) use mathematical models to calculate decision variables for detection tasks and are known to fail with variable signal sizes [17]. However, in the clinical routine signals like liver lesions vary in size, shape and contrast. Machine learning (ML), on the other hand, showed great potential in image classification tasks [18–20]. Especially, convolutional neural networks were trained to perform challenging image classifications such as handwritten digit recognition with performances similar to humans [21]. Therefore, the purpose of the first part of this dissertation is the development of automated image quality assessment that is robust to signal size uncertainty using ML and compare the performance with a traditional CHO.

The second part of this doctoral project is focused on the clinical evaluation of a preclinical SPCCT prototype. In contrast to commercial clinical CT systems, where energy-integrating detectors (EID) integrate over the incoming photon flux and determine the attenuation, photon-counting detectors (PCD) measure single photons and the energy of each photon is determined. Due to its technical advancements, SPCCT offers capabilities to utilize spectral information for material decomposition and provides the possibility of advanced high-resolution imaging [22, 23]. While spectral imaging is investigated in the co-authored peer-reviewed journal publications J1, J5 and J14, the purpose of the clinical evaluation in this work is the investigation of high-resolution capabilities of the preclinical SPCCT prototype.

In a third part of this doctoral project, capabilities of SpSCT are evaluated for its clinical application. Sparse sampling is inspired by the theory of compressive sampling, or compressed sensing (CS), where the signal acquisition rate is reduced below twice the sampling frequency (Nyquist rate), asserting that signals or images can still be fully recovered [24]. For a CT system, that would mean that the radiation dose of a CT scan could be substantially reduced. However, hardware modifications are necessary to allow a blocking of the x-ray beam or a fast pulsing of the x-ray tube [25]. SpSCT has been evaluated for different clinical applications in several publications and conference abstracts that were co-authored by this author (J3, J10, J21, J15, C1, C3, C6, C8, C9). The purpose of the third part of the doctoral project is the clinical evaluation of sparse sampling for the detection of pulmonary embolism by simulating sparse sampling at different dose levels on clinical CT scans.

1.3 Thesis Structure

The technical and methodological background of the three embedded journal publications of this cumulative dissertation are presented in Chapters 2 and 3. Short summaries of the included publications JP-I, JP-II and JP-III are given in Chapter 5, and a discussion of the presented work in the context of current related work is provided in Chapter 6.

2 Technical Principles

This chapter gives a basic overview of the technical and physical background of the methodology used in this thesis.

2.1 Computed Tomography

In x-ray computed tomography (CT) a source, the so-called x-ray tube, and a detector are fixed on a rotating gantry (Fig. 2.1). X-ray photons with energies of a specific spectrum are emitted by the tube and the attenuation is measured by the detector. The acquired data is used to reconstruct volumetric images of the scanned object. In this section, the basic concept of CT is described. Interested readers find a detailed description of CT in J. Hsieh's book *Computed Tomography: Principles, Design, Artifacts, and Recent Advances [5].*

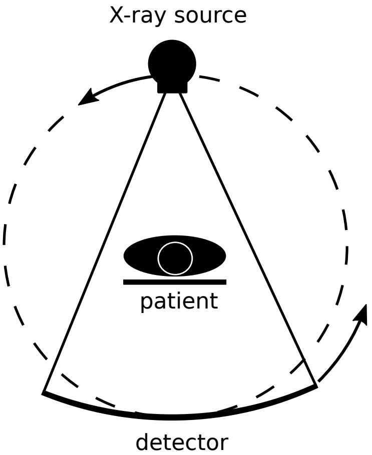


Figure 2.1: Schematic representation of a CT system.

An x-ray source and a detector rotate around an object .

2.1.1 X-ray Generation

A typical x-ray generator in a CT system consists of a cathode and a rotating anode that are installed in a vacuum-sealed atmosphere. As a result of electronic heating of the cathode, electrons are emitted and accelerated towards the rotating anode by application of a high voltage between the cathode and the anode. In clinical CT, voltages range between 80 and 140 kVp. At the interaction between the electron beam and the anode, a part of the electrons' kinetic energy is converted into x-rays [26].

With about 60 % of the electron energy, the major part is converted into heat. Therefore, a rotating anode is needed to distribute the heat on the anode plate and to guarantee a stable x-ray generation. About 39 % of the electron energy is backscattered - that corresponds to about 50 % of the electrons. With about 1 % of the electron energy, only a small part is converted to bremsstrahlung. The generated x-rays have a continuous spectrum with an upper energy limit defined by the operated voltage between cathode and anode (80 keV at 80 kV). A minor part of the electron energy is converted into characteristic x-rays. In CT imaging, the most common target material of the anode is tungsten. When the x-ray tube is operated with tungsten, about 0.1 % of the energy is converted to characteristic x-rays that correspond to tungsten. The anode emits x-rays nearly isotropically (indicated in Fig. 2.2), hence, only a small fraction of the generated x-rays leave the tube through an x-ray window and is used for imaging (about 1 %) [27]. A schematic representation of an x-ray tube is visualized in Figure 2.2.

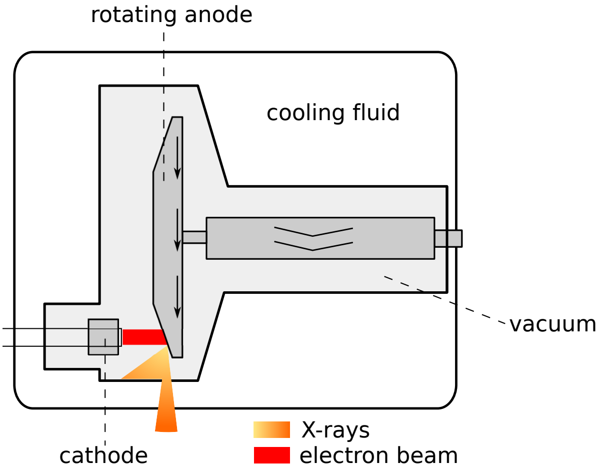


Figure 2.2: Schematic representation of an x-ray tube.

A cathode emits an electrons beam towards a rotating anode in a vacuum-sealed atmosphere. At the interaction between electrons and the cathode, parts of the electron's kinetic energy is converted into x-rays. Parts of the x-rays exit the tube through an x-ray window and are used for imaging. Graphic adapted from [26].

2.1.2 Data Acquisition

In the whole imaging chain, there are numerous technical steps necessary to guarantee a high quality of the measurements and high dose efficiency. To mention a few, the x-ray spectrum needs to be shaped by applying material filters, the x-ray beam needs to be shaped by a collimator and an anti-scatter grid is necessary to minimize the detection of scattered photons at the detector. As the detailed composition and operation of a CT system is beyond the scope of this work, interested readers are referred to literature. A detailed explanation of diagnostic x-ray sources is given in R. Behling's book *Modern diagnostic x-ray sources: technology, manufacturing, reliability* [27]. Basic concepts concerning the data acquisition are explained in this section.

Interaction with matter

Emitted x-ray photons travel through the CT aperture and are collected by the detector. On their way to the detector, photons interact with matter that is on their path. In conventional medical CT, the attenuation of the energy is measured by an energy-integrating detector (EID). EIDs and the detection process will be explained in the next section. In this section, the basic principles about the attenuation of the x-ray intensity are briefly outlined.

The energy of x-ray photons in medical CT ranges from about 20 to 140 keV. In this range, the fundamental physical mechanisms attenuating the photons are the photoelectric effect, the Compton effect and coherent scattering [5]. As a result of these interactions some photons are either scattered or absorbed. The attenuation of a monochromatic x-ray beam through a homogeneous object can be expressed by the Beer-Lambert law:

$$I = I_0 e^{-\mu L}, \tag{2.1}$$

where I is the attenuated photon intensity, I_0 is the incoming photon intensity, μ is the linear attenuation coefficient and L is the thickness of the material. The linear attenuation coefficient is the combination of the different interactions:

$$\mu = \tau + \sigma + \sigma_r, \tag{2.2}$$

where τ, σ and σ_r are the attenuation coefficients of the photoelectric effect, Compton effect and coherent scattering, respectively [5]. Higher μ values indicate a stronger attenuation.

For a non-uniform object where the x-rays pass several materials, modifications of Equation (2.1) are necessary:

$$I = I_0 e^{-\mu_1 \Delta x} e^{-\mu_2 \Delta x} ... e^{-\mu_n \Delta x} = I_0 e^{-\sum_{n=1}^{N} \mu_n \Delta x}, \tag{2.3}$$

where μ_n are the linear attenuation coefficients for several materials of thickness Δx. By division of both sides of Equation (2.3) by I_0 and by taking the negative logarithm, we obtain:

$$-\,ln\left(\frac{I}{I_0}\right) = \sum_{n=1}^{N} \mu_n \Delta x. \tag{2.4}$$

For small Δx approaching 0, the sum in Equation (2.4) becomes the integration over the length of the object:

$$p = -ln\left(\frac{I}{I_0}\right) = \int_L \mu(x) \delta x. \tag{2.5}$$

The measured projection p is the line integral of all attenuation coefficients along the path of an x-ray [5]. In CT, such line integrals are saved for each tube position and for each detector element resulting in a so-called sinogram.

Detection of X-rays

In conventional medical CT, almost all systems use solid-state scintillator detectors. X-ray photons are converted into light photons and measured by a photodiode over a defined integration time. This kind of detector is called energy-integrating detector (EID) and they are the state-of-the-art in medical CT, today. Scintillator materials are typically cesium iodide (CsI), bismuth germanate (BGO), cadmium tungstate ($CdWO_4$) and gadolinium oxysulphide (Gd_2O_2S) [28]. Figure 2.3 illustrates a schematic drawing of an EID. Each detector element consists of scintillator material surrounded by reflecting material on top of a photon detector (photodiode). Arriving x-ray radiation is converted to light inside the scintillator and absorbed by the photodiode attached to the crystal. With the absorption of the photon, an electric charge is generated that is proportional to the x-ray intensity. The electric charge is integrated over the defined integration time [28]. The electronic at the detector also generates a low electric field. Due to the concept of EIDs, where the whole energy is simply integrated, so-called electronic detector noise is present in the measurement data [5].

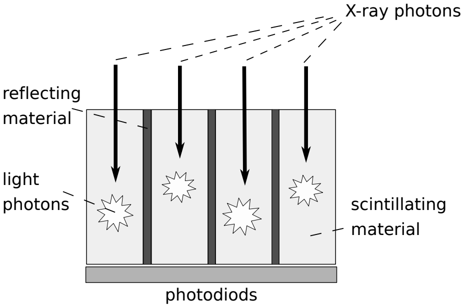


Figure 2.3: Schematic representation of an energy-integrating detector. Figure adapted from [5].

2.1.3 Radiation Dose

The radiation dose is primarily given by the operated tube voltage, the tube current and the radiation time. The more energy is absorbed by the patient, the more harmful the radiation is. Therefore, the energy of the x-ray photons, the number of x-ray photons and the duration of the radiation are the main parameters controlling the x-ray dose [29].

Tube voltage

The operated *tube voltage* defines the highest energy of the produced x-ray photons and is given as peak kilovoltage (kVp). It defines the x-ray spectrum that is used for imaging. Low-energy photons (<20 keV) are more attenuated during imaging and have a higher probability not to reach the detector. Thus, they mainly contribute to the dose applied to the patient but not to the imaging process. Therefore, the x-ray spectrum is filtered before the radiation is used for imaging (Figure 2.4). Typical filter materials are aluminum or copper [28]. On the other hand, high-energy x-ray photons have the potential to transfer more energy to the patient. Therefore, especially for small patients and pediatric CT, tube voltages of 80 or 100 kVp are used. For big patients or strong attenuating materials like bone, higher tube voltages like 120 or even 140 kVp are used.

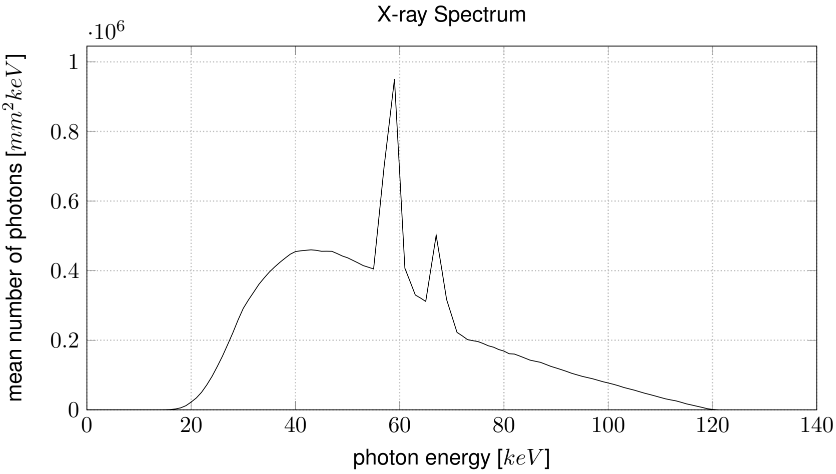


Figure 2.4: X-ray spectrum for 120 kVp.

X-ray spectrum for a tube with a tungsten anode operated with 120 kVp and with a 2.5 mm thick aluminum filter. X-ray photons with energies below about 20 keV are removed from the spectrum. The spectrum was generated with an online tool [30] that is based on References [31] and [32].

High-energy x-ray photons are less attenuated and produce a better statistic at the detector even for big patients. This results in less noise in the measurements [29]. In conventional medical CT, the tube voltage is held constant for the duration of an examination.

Tube current-time product

The *tube current* is given in milliamperes (mA) and defines the number of produced x-ray photons. In general, the tube current is proportional to the radiation dose. Keeping every other scan parameter unchanged, doubling the current e.g. from 100 mA to 200 mA doubles the x-ray output.

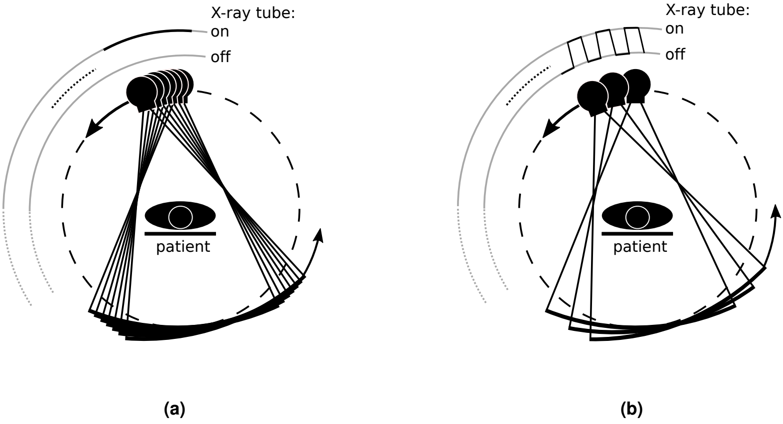
The *tube current-time product* is given in milliampere-seconds (mAs) and is the product of the operated tube current with the radiation time per rotation. This parameter is often called x-ray exposure and will be used in this work as a reference for dose. In conventional medical CT, the radiation time is equal to the rotation time because the x-ray tube is always on during a CT acquisition. For constant scan parameters, the tube current-time product is proportional to the dose. In general, short scan times are desired because motion artifacts caused by movement of the patient or organs (e.g. caused by breathing or the heart beat) are decreased. Moreover, shorter rotation times and therefore shorter radiation times allow a higher tube current at the same tube current-time product. This has the advantage that a higher x-ray photon flux results in better statistics on the detector and reduced noise in the measurements. Therefore, also the noise in reconstructed images is reduced. Typical rotation times of the gantry for one rotation are between 0.3 to 2 seconds.

The attenuation profile along a patient changes with the anatomical regions. For example, x-ray radiation along the shoulder is more attenuated than radiation passing though the chest. Therefore, *tube current modulation* (TCM) is applied in medical CT. In contrast to the tube voltage which is fixed for the duration of an acquisition, the tube current is automatically modulated during the acquisition to guarantee a nearly constant photon flux at the detector. Therefore, in regions with higher attenuation (e.g. shoulder), higher tube currents are used than in regions with lower attenuation (e.g. chest). TCM offers up to 40 % dose reduction per examination [29].

2.2 Sparse-Sampling CT

2.2.1 Background

In conventional CT, the x-ray tube is always on during a CT acquisition and the detector measures the x-ray radiation intensity. As a consequence, the radiation time for one rotation of the x-ray tube is equal to the rotation time. Low-dose CT acquisitions are performed by lowering the tube current, emitting less x-ray photons. As a result, the photon statistics at the detector is degraded and noise in the measurements as well as image noise increases. In addition, the contribution of electronic detector noise to the measurement increases. At ultra low tube currents, image quality is not diagnostic and a boundary is reached [33]. Blocking the x-ray beam at specific positions for short times would reduce the radiation time during the examination and, hence, allows higher tube currents at the same or even reduced tube current-time product. This acquisition scheme is referred to as sparse-sampling CT (SpSCT).



(a) **(b)**

Figure 2.5: Conventional CT vs. SpSCT.

(a) In conventional CT, the X-ray tube is always on during a CT examination; (b) In SpSCT, the x-ray tube can be switched off at any position resulting in a pulsed x-ray flux.

2.2.2 Principle

The idea of sparse sampling comes from the signal processing domain. The well-known concept of compressive sampling, or compressed sensing (CS), uses sparsely sampled data acquisitions to fully recover the signal or image. The initial idea is to randomly omit sampling points. The sampling rate can be reduced below twice the sampling frequency (Nyquist rate) while the signal or image is still correctly recovered [24]. For CT, periodic uniformly-spaced sampling schemes were investigated that would require a fast pulsing x-ray tube [34–36]. Figure 2.5 illustrates the schematic operating mode of SpSCT compared to conventional CT. In contrast to conventional CT where the x-ray tube is always on, in SpSCT the tube periodically changes its state between on and off. In this work, we refer to two-times sparse sampling (2-times SpS) if the tube is turned on only at every second angular tube position and four-times sparse sampling (4-times SpS) if the tube is active at every fourth tube position. 2-times SpS results in a reduction of the radiation time by 50 %, 4-times SpS reduces the radiation time by 75 %. This would reduce the radiation dose to 50 % and 25 % of a conventional CT scan, respectively, if other scan parameters remain unchanged (Table 2.1). On the other hand, SpSCT would allow a higher tube current at the same radiation dose as conventional CT (Table 2.2). This results in a better statistics at the detector and reduces noise.

sampling scheme	rotation time	radiation time	tube current	tube current-time product (dose)
conventional CT	0.5 s	0.5 s	200 mA	100 mAs
2-times SpS	0.5 s	0.25 s	200 mA	50 mAs
4-times SpS	0.5 s	0.125 s	200 mA	25 mAs

Table 2.1: Dose reduction with SpSCT.

sampling scheme	rotation time	radiation time	tube current	tube current-time product (dose)
conventional CT	0.5 s	0.5 s	200 mA	100 mAs
2-times SpS	0.5 s	0.25 s	400 mA	100 mAs
4-times SpS	0.5 s	0.125 s	800 mA	100 mAs

Table 2.2: Higher tube current with SpSCT.

However, the application of sparse sampling might introduce undersampling artifacts like aliasing or streakings and in ultra-low-dose CT there will also be an increase in noise [34]. To compensate these negative effects, advanced iterative reconstruction algorithms can be used. Statistical iterative reconstruction (SIR) is known to reduce noise and artifacts by a precise modeling of the system physics and the measurement statistics [34, 37–39]. These algorithms are computationally expensive and require long runtimes. Therefore, their availability in the clinical routine is limited, but might be extended in future, when hardware becomes more powerful.

2.2.3 Technical Implementation

In this work, sparse sampling is simulated under the assumption of a fast pulsing x-ray tube. One strategy for the realization of sparse sampling is the modification of the x-ray tube by insertion of a grid electrode between the cathode and the anode (Figure 2.6). The electron beam is blocked when a negative potential of about -12 kV is applied to the electrode. Switching the electrode off or to a slightly positive potential allows the electron beam to reach the anode. With this technique, a pulsing x-ray tube can be realized. Today, this kind of pulsing x-ray tubes are limited to interventional imaging systems and would need modifications (faster pulse speed, higher tube current) for the implementation in medical CT systems [25, 27, 40].

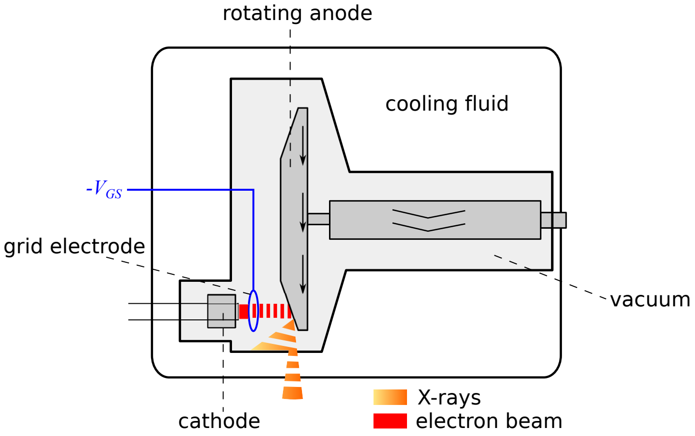


Figure 2.6: Schematic representation of a pulsing x-ray tube.

A grid electrode is added between the cathode and the rotating anode. When a negative potential V_{GS} *is applied to the electrode, the electron beam is blocked and no x-ray radiation is produced. Pulsing the electrode between 0 and about -12 kV is easier to implement than pulsing the x-ray tube itself.*

2.2.4 Possible Further Applications

Apart from a reduced radiation dose at the same operated tube current or higher possible tube currents at the same radiation dose, SpSCT would allow further improvements of the CT workflow.

Scan Planning

Today, scan planning in clinical routine is based on two-dimensional (2D) overview scans, also known as scout scans. Depending on the intended procedure, one or two 2D scans are performed. For these scans the x-ray tube is fixed at one position, while the patient table moves through, resulting in a radiograph of the patient (Figure 2.7). In general, an antero-posterior (A-P, 12 o'clock position of the tube) scan is sufficient for scan planning, however, for some procedures like imaging of the spine an additional lateral (3 o'clock position of the tube) scout scan is necessary to define the exact location of the spine.

Sparse sampling would enable three-dimensional (3D) overview scans by reducing the projection views to about 6.25 % of a conventional CT scan with about 3 mAs tube current-time product. That is in the range of 2D scout scans with about 4.5 mAs [35, 41]. 3D scout scans would allow more precise selections of the scan regions to scan, e.g. if a specific organ like the liver needs to be investigated. Safety margins at the beginning and end of the scan regions become obsolete resulting in a further reduced radiation dose. Moreover, 3D scout scans could allow automatic organ segmentation in the volume, and suggesting scan regions for a specific procedure reducing the source of error by incorrect human interaction [35, 42].

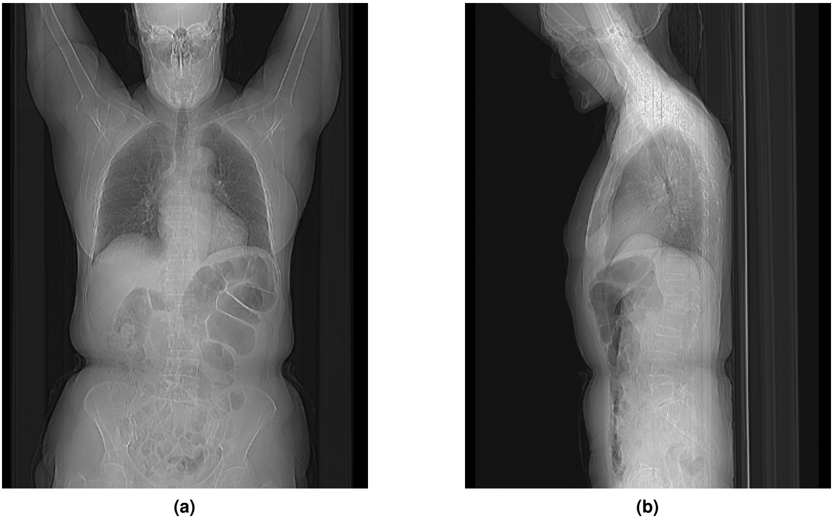

(a) **(b)**

Figure 2.7: Overview scans.

Antero-posterior (A-P) and lateral overview scans for scan planning of the spine.
(a) A-P overview scan; (b) lateral overview scan.

Advanced Dose Modulation

Tube current modulation (TCM) guarantees a better dose efficiency, because only as little dose is used as needed for a specific body region. In addition, the photon flux at the detector is held constant resulting in better image quality. In angular modulation, the tube current is modulated online during the CT scan. According to the measured attenuation at one position, the current at the position plus 180° is adjusted. This way, different tube currents can be used for the antero-posterior (A-P, 6 and 12 o'clock position) and lateral position (3 and 9 o'clock position). At the shoulder, for example, there is a higher attenuation for the lateral position than the A-P position, hence, higher tube currents are desirable [43].

More advanced z-axis modulation (along the patient) computes the modulation profile based on a 2D overview scan as shown in Figure 2.7a. Water-equivalent thickness in elliptical shape is determined for every z-position from the attenuation in the overview scan and needed tube current is calculated [43]. However, there are two challenges that could be resolved by sparse sampling. Firstly, the elliptical water-equivalent thickness is only a rough estimate of the attenuation of a patient. With sparse sampling and 3D scout scans (as described under *Scan Planning*), the exact attenuation profile and thereby an ideal modulation profile could be determined. Secondly, the realization of an ideal modulation profile requires the x-ray tube to control the electron flux very fast. Ideally, changes from low tube current to high tube current would be realizable within a few projection angles.

Conventional x-ray tubes that are used for medical CT allow a tube current slew rate between 1 and 2 % $per\ ms$. Assuming a fast rotation time of $330\ ms$ and the scan region at the shoulder, a high change in current would be desirable within a quarter rotation. For a current of 100 mA at the 12 o'clock position of the tube, within a quarter rotation ($330/4\ ms$ = $82.5\ ms$) to the 3 o'clock position only a continuous change of 165 mA would be realizable. With sparse sampling, more precise modulation could be achieved by controlling the pulse width (the duration of one pulse). For a conventional CT system with a rotation time of $330\ ms$ and 2400 projection images in one rotation, the integration time for one projection image is $138\ \mu s$. With an x-ray tube that is capable of $3\ \mu s$ pulse width, as demonstrated with a prototype by Wiedmann et al. [25], tube current-time products between about 2 % and 100 % could be achieved for each projection compared to conventional CT [44].

2.3 Photon-Counting CT

Conventional medical CT is a qualitative imaging modality where the attenuation of the x-ray radiation results in gray-scale images. With EIDs the energy of single photons is lost and therefore spectral, material specific information cannot be derived. Moreover, because EIDs simply integrate over the energy, electronic detector noise contributes to the signal and lower energy photons contribute less to the signal than higher energy photons resulting in so-called beam hardening artifacts [45]. However, with photon-counting CT, single photons are counted on the detector combined with their energy, eliminating electronic detector noise and allowing a correct weighting of low-energy photons. In addition, quantitative imaging becomes available. In this work, the basics of photon-counting CT are explained. Willemink et al. give a good overview of the current state concerning photon-counting CT for medical imaging in Ref. [22].

2.3.1 Principle

Photon-counting CT utilizes a detector which directly converts impinging x-ray photons to electric charge and captures the energy. Figure 2.8 illustrates a schematic representation of a photon-counting detector (PCD). The detector has a high bias voltage around $300\ V$ applied [46] and is made of semiconductor material, such as Si, CdTe or CdZnTe. Each x-ray photon generates an electron-hole pair at the interaction with the semiconductor and the energy is resolved. Hence, single photons can be counted by measuring the retrieved energy pulse [5]. Compared to EIDs (Figure 2.3), in PCDs the reflective coating material of each detector element is not necessary, allowing smaller detector pixel sizes. The photons are grouped with respect to their energy into so-called energy bins. The PCD used in the prototype evaluated in JP-II uses a total of five thresholds to group into the bins [47].

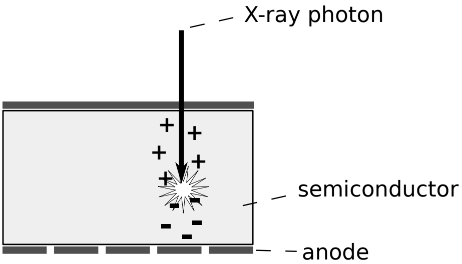


Figure 2.8: Schematic representation of a photon-counting detector.

An x-ray photon is directly converted into electric charge on the semiconductor crystal and measured by the detector. Figure adapted from [5].

2.3.2 Advancements

Photon-counting detector technology has the potential to bring several advancements to CT imaging listed below [46].

Noise and Dose Reduction

The capability to detect single x-ray photons eliminates electronic detector noise from the measurements. Especially low dose scans, e.g. pediatric CT, and scans with high attenuation as with obese patients benefit from reduced noise and a possibly reduced radiation dose. Moreover, correct weighting of low energy photons that have a higher contribution to low image contrast results in a substantial elevation of the contrast-to-noise ratio that could result in 99 % CNR improvement or 75 % dose reduction [46, 48].

Improved Spatial Resolution

Photon-counting CT is known for its capability of high resolution imaging. Figure 2.9 shows images of a spatial resolution phantom from a medical CT system (Philips IQon, , Philips Healthcare, Cleveland, OH, USA) with EIDs compared to a photon-counting CT prototype (Philips Healthcare, Haifa, Israel). There are bore holes of different sizes and spacing in the phantom representing 1.25, 1.66, 2.50, 3.33, 4.16, 5.00, 5.55, 6.25, 7.14, 8.33, 10.0 and 12.5 $line\ pair/cm$ spatial resolution. With photon-counting, the 12.5 $line\ pair/cm$ pattern can still be resolved while with the conventional CT, the pattern is washed out. In general, there are two major properties limiting the spatial resolution of EIDs. Firstly, reflecting material is coated around each detector element to guarantee that the converted light is not detected by adjacent detector pixels. To keep a certain ratio between the detector pixel size and the coating material, detector pixels should not get too small. Secondly, EIDs work better with a higher number of photons reaching each detector pixels because the contribution of electronic detector noise gets reduced and better statistics can be achieved [5, 28].

PCDs, on the other hand, require small detector pixel sizes to avoid concurrent events of arriving photons (so-called pile-up, Section 2.3.3) because arriving photons are better distributed to the detector pixels. In addition, the semiconductor layer of PCDs is not structured, meaning that there is no coating material around the detector pixels (see Figure 2.8). Each electrode functions as detector pixel with high bias voltages avoiding diffusion of the charge and ensuring a straight path to the electrodes. Thus, higher spatial resolution is achievable [45, 46].

The spatial resolution capabilities of photon-counting CT, especially for pulmonary imaging, are investigated in JP-II.

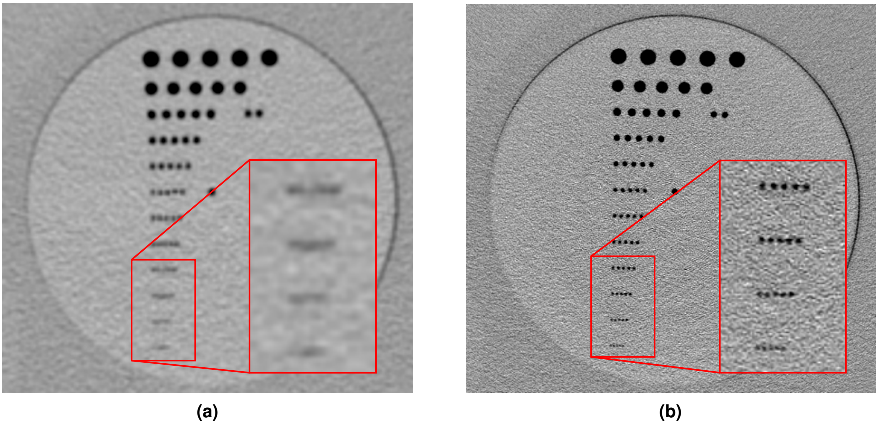

(a) (b)

Figure 2.9: CT images of a spatial resolution phantom.

Data were acquired with a medical CT system (a) and a photon-counting CT prototype (b). Both CT systems were operated with 330 mA *and 1* s *rotation time. Images were reconstructed with filtered back projection (FBP). The bottom row of the pattern represents 12.5* $line\ pair/cm$. *Level: -70 HU; window: 950 HU.*

Spectral Information

PCDs enable quantitative imaging by resolving the energy of arriving x-ray photons. Because the attenuation is material-specific and energy-dependent, material quantification becomes available (also known as material decomposition) [45]. The energy dependence of the linear attenuation coefficient of most clinically relevant materials can be described as a linear combination of the photo-electric effect and the Compton cross-sections. When materials like gadolinium (Gd) might be used as contrast agent, the so-called K-edge discontinuity also contributes to the attenuation. With photon counting, these effects can be resolved and materials can be quantified [49, 50].

Spectral capabilities of photon-counting CT were investigated in co-authored publications J1, J5 and J14 and are not directly in the scope of this cumulative thesis. Interested readers are referred to References [49] and [50] for further reading.

2.3.3 Challenges

Photon-counting CT comes with several challenges. Major challenges are addressed in the following lines [45, 46].

- *Pile-up* is the effect when two or more photons arrive within one counting rate of the detector. In this case, only one event is registered and the energies of the arriving photons are accumulated. As a result, one count with a very high energy is detected.

- *Charge sharing* is the event, when an x-ray photon hits the semiconductor layer between two detector electrodes (Figure 2.10 a). The electric charge diffuses and each of the two adjacent detector electrodes register one count with a fraction of the impinging photon.

- *Secondary photons* can be a result of the interaction between the incident photon and the semiconductor layer (Figure 2.10 b). At the interaction, characteristic fluorescent X-ray photons can be emitted that are detected by an adjacent electrode. Another effect is Compton scatter. When the photon is scattered, it changes its direction and leaves part of its energy at the incident position. The photon may then be absorbed by the same detector pixel, an adjacent detector pixel or even leave the detector.

To overcome the pile-up effect, small detector pixels with high counting rates are beneficial. Charge sharing, on the other hand, would benefit from larger detector pixel sizes. The photon-counting prototype evaluated in JP-II is equipped with detector pixels of size 500 μm x 500 μm and is capable of count-rates exceeding 15 *Mcps/pixel* [47].

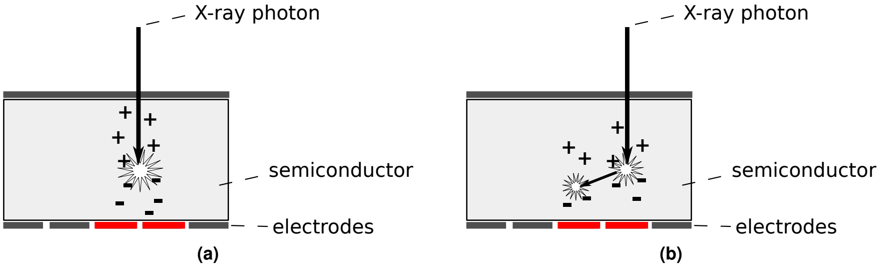


Figure 2.10: Charge sharing and secondary photons.

(a) Charge sharing: the energy of an incident x-ray photon is diffused to several detector electrodes. (b) Secondary photon: an incident x-ray photon produces a characteristic x-ray photon that is detected by an adjacent detector electrode. Red detector electrodes indicate that it counts a photon.

2.4 CT Reconstruction

CT reconstruction describes the process when CT images are generated from measured projection data. In general, reconstruction algorithms can be grouped into analytical and iterative reconstruction. In commercial CT systems, hybrid reconstruction, that is a combination of iterative and analytical reconstruction, was introduced by vendors as a third group [51]. In medical CT, these hybrid algorithms are usually referred to as iterative reconstruction. The following sections give an overview of the three groups of reconstruction algorithms.

2.4.1 Analytical Reconstruction

The most intuitive approach is simply back projecting the acquired projection measurements from each angle to the image. All pixels on the path of one ray are increased by the partial value of the measurement (Figure 2.11). After repetition for all projection angles, the object is visualized by the image. However, the reconstructed image is a blurred version of the original object (Figure 2.12).

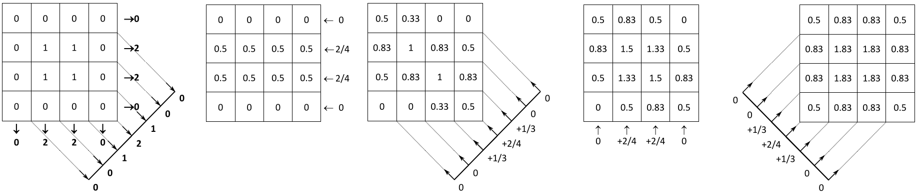
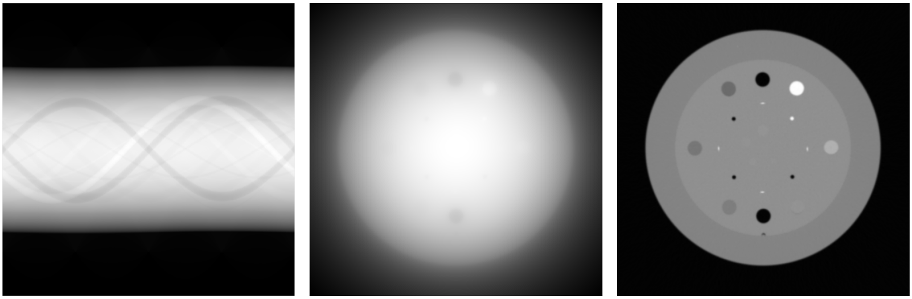

0 0 0 0 →0
0 1 1 0 →2
0 1 1 0 →2
0 0 0 0 →0
0 2 2 0
0 0 1 2 1 0 0
0 0 0 0 ← 0
0.5 0.5 0.5 0.5 ←2/4
0.5 0.5 0.5 0.5 ←2/4
0 0 0 0 ← 0
0.5 0.33 0 0
0.83 1 0.83 0.5
0.5 0.83 1 0.83
0 0 0.33 0.5
0 0 +1/3 +2/4 +1/3 0 0
0.5 0.83 0.5 0
0.83 1.5 1.33 0.5
0.5 1.33 1.5 0.83
0 0.5 0.83 0.5
0 +2/4 +2/4 0
0.5 0.83 0.83 0.5
0.83 1.83 1.83 0.83
0.83 1.83 1.83 0.83
0.5 0.83 0.83 0.5
0 0 +1/3 +2/4 +1/3 0 0

Figure 2.11: Schematic representation of a back-projection process.

The left image illustrates the true object and its forward projection for three different angles. The second, third, fourth and fifth images represent the back projection for four different projection angles.



(a) (b) (c)

Figure 2.12: Images of the sinogram (left), back projection (center) and FBP (right) of module CTP404 of a Catphan phantom.

Over the last decades, filtered back projection (FBP)–an analytical reconstruction algorithm–was used in commercial medical CT systems. FBP is a simple method that is computationally fast and, therefore, it is well suited for the clinical routine. In FBP, high-pass filters are applied to the measurements before back projection to reduce low frequencies that cause blurring in the image [5]. In general, FBP can be expressed mathematically by:

$$f(x,y) = \int_0^{\pi} d\Theta \int_{-\infty}^{+\infty} P(\omega,\Theta)|\omega| e^{j2\pi\omega(x\cos\Theta + y\sin\Theta)} d\omega, \tag{2.6}$$

where $f(x,y)$ is the reconstructed image; $P(\omega,\Theta)$ is the Fourier transform of the projection at angle θ; $|\omega|$ is the applied filter. The inner integral is the inverse Fourier transform of $P(\omega,\Theta)|\omega|$

that results in a filtered projection in the spatial domain [5].

In commercial CT systems, a variety of different filters are implemented to control the image quality of FBP. Sharp filters, like the well-known Ram-Lak filter, only filter low frequencies and let high frequencies pass producing images with sharp edges but with the drawback of high image noise. Hence, sharp filters require more dose for sufficient image quality. However, for some clinical indications, like imaging of the bone or lungs, sharp edges are essential. On the other hand, smooth filters also reduce high frequency signals to a desired degree resulting in smoother images with less image noise. These filters are preferred in low-dose CT and in examinations with focus on areas with low contrast, like liver lesion detection [26].

FBP is a reconstruction algorithm whose popularity is mainly based on the fast computation time. The requirement of low noise in the measurements leads to a relative high radiation dose–with lower radiation dose, more photons are attenuated and not reaching the detector. With increasing concerns in the society over the harmful effects of ionizing radiation to the patient, significant radiation dose reduction was requested.

2.4.2 Iterative Reconstruction

In the CT research community, substantial effort has been spent in the development of iterative reconstruction (IR) methods for many years [52–57]. There were methods that iteratively filtered the measurement, that iteratively filtered the image in image space and fully iterative methods that compared intermediate results with the acquired measurements. These methods allowed further dose reduction than FBP. However, their clinical application in commercial CT systems was not feasible for a long time due to long computation times. With the hardware advancements in computational power and the availability of high parallelizable programmable graphics processing units (GPU) and their application to IR, computation times were substantially reduced [58, 59] leading to the first introduction of IR to a commercial CT system in 2009 [60]. While the first commercial IR algorithm performed iterative de-noising of the image after a single backward projection (IRIS, iterative reconstruction in image space, Siemens Healthineers), more advanced reconstruction algorithms were introduced in the following years [51].

Hybrid Iterative Reconstruction

More advanced reconstruction algorithms were introduced that perform iterative image processing in the sinogram-space and in the image-domain. Iterative filtering of the sinogram results in reduced artifacts, then a single back projection is performed, followed by iterative filtering of the image to reduce noise (Figure 2.13) [51]. This group will be referred to as hybrid IR.

Hybrid IR already allowed a substantial reduction of the radiation dose in routine CT examinations [61–63]. iDose^4 (Philips Healthcare, Best, the Netherlands), a hybrid IR algorithm was used for evaluation of a novel model observer in JP-I. While first commercial IR algorithms only performed adaptive linear filtering in image-space and therefore, reduced the spatial resolution and are less powerful in artifact reduction, iDose^4 has the potential to substantially improve image quality by its hybrid approach. Spatial resolution uptake of 68 % at the same dose or 80 % dose reduction at same image quality were reported compared to FBP [64]. In general, the algorithm identifies projections in the measurement that have very low photon statistics and therefore, are very noisy. The projections are iteratively corrected using a model with the true photon statistics preserving edges (spatial resolution) and reducing noise as well as artifacts without significantly corrupting the intensity values. After back projection, noise is iteratively removed from the image volume. In a first step, the noise distribution in the image volume is estimated. The estimate is used throughout the iterative process to preserve the image noise power-spectrum, that guarantees a similar look and feel as FBP images, and to preserve true structures. Structural models are then fitted to the data to reduce noise without removing the true structure [64].

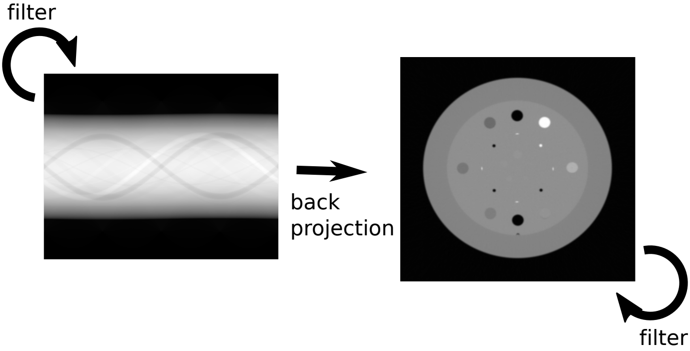


Figure 2.13: Hybrid iterative reconstruction (IR).

In hybrid IR the sinogram is filtered iteratively to reduce artifacts before back projection. After back projection the image is iteratively filtered to reduce image noise [51]. Figure adapted from [51].

Statistical Iterative Reconstruction

In model-based statistical iterative reconstruction (SIR) the reconstruction process is handled as an optimization process where the image is iteratively compared to the measurements (Figure 2.14). The idea is to minimize a cost function $\Psi(u)$ of the form:

$$\Psi(u) = -L(Au|y) + \beta R(u), \tag{2.7}$$

where L is the likelihood that the image u is a valid representation of the measurement y. A is the forward projection operation (or system matrix) and R is the regularization term. β is a parameter that controls the strength of the regularization [65, 66]. Regularization reduces noise in the reconstructed image by penalizing roughness. The regularization term can be expressed by:

$$R(u) = \sum_{j} \sum_{k \in N_j} w_k \psi(u_j - u_k), \tag{2.8}$$

where N_j is the set of neighboring voxels of voxel j, w_k is a weight according to the degree of neighboring and ψ is the potential function. Although a simple quadratic potential function will result in reduced image noise, its application also degrades edges because an edge is interpreted as roughness between a voxel and its neighbors. Therefore, edge preserving potential functions often have a threshold value δ that defines the roughness (the maximum difference to the neighboring voxels) that will be penalized, such as the Huber function [67]:

$$\Psi(t) = \begin{cases} t^2/2, & |t| \leq \delta \\ \delta|t| - \delta^2/2, & |t| > \delta. \end{cases} \tag{2.9}$$

A common algorithm to solve the optimization problem is the ordered-subset version of the separable paraboloidal surrogates algorithm published by Fessler and Erdogan, where the optimization problem is iteratively solved until convergence [56, 68].

In commercial CT systems, fully iterative reconstruction algorithms are inspired by SIR but substantially reduce the number of iterations by an optimized initialization. Iterative model reconstruction (IMR, Philips Healthcare, Best, the Netherlands), as an example, uses a Huber penalty on the projection and image data to generate noise-reduced images as initialization claiming to reduce the number of iterations by an approximate of 5 to 10 SPS iterations [65, 69].

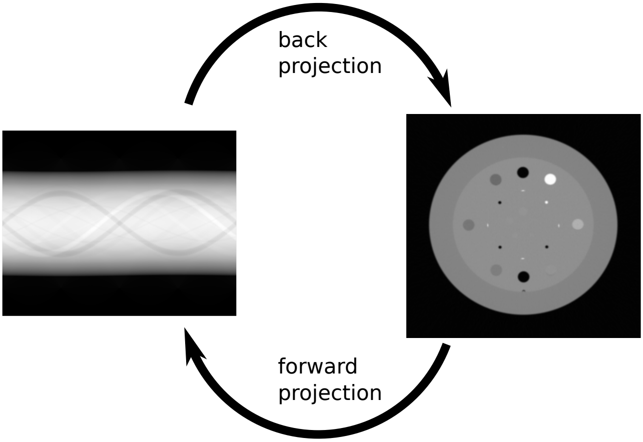


Figure 2.14: Statistical iterative reconstruction (SIR).

In SIR there is an initial guess of the image that is forward projected into sinogram space and compared with the measurement and the error sinogram is back projected to update the image. A regularization term is applied in the update process to reduce image noise [51]. Figure adapted from [51].

Non-linearity

FBP is mostly a linear reconstruction process where local resolution and noise measurements are useful tools for image quality assessment because they give a good indication about the detectability of image properties like lesions or fractures. With hybrid IR and SIR there are several steps introducing non-linearity to the reconstruction process, among others, image processing steps, non-quadratic regularization and non-negative constraints [8, 70]. As a result, the relevance of simple noise measurements and spatial resolution for the diagnostic image quality is only limited, because image quality becomes contrast dependent [9, 71] and noise is often used as a tuning parameter of IR [72]. Therefore, the CT research community spends considerable efforts on the investigation of methods for the assessment of diagnostic image quality. In this work, several approaches for the evaluation of diagnostic image quality are presented including the application of machine learning, investigation with a custom-made phantom and a reader study with human observers.

3 Strategies for System Evaluation

Different approaches for the clinical evaluation and assessment of the diagnostic value in the context of CT imaging are described in the following sections.

3.1 Reader Studies

Reader studies, also known as observer studies, are used to investigate the performance of humans for a given task. They are mostly used to compare new developments and enhancements with state-of-the-art technologies. Generally, the outcome gives an estimation about the expected performance of a given technology in clinical routine. Common imaging tasks are [73]:

- *Detection* tasks require the reader to find abnormalities in the image, but avoiding false positive findings.
- *Classification* requires the reader to name an abnormality.
- In *staging* tasks, the extend of the abnormality is determined.
- *Comparison* tasks require the reader to determine changes in structures or abnormalities.
- In *estimation* tasks, the reader is asked to determine the size, shape, intensity, etc.

A reader study can also consist of a combination of different imaging tasks. The reader study published in JP-III consists of a detection task and a classification task. Firstly, the readers are required to determine the presence of a pulmonary embolism (PE), secondly, the PE (if present) has to be classified.

To enable a finer differentiation between different image qualities, the readers are required to give a confidence rating for their decision. The binary decision (abnormality present/absent) becomes a rating on a predefined scale. More traditional figures of merit like *sensitivity*, *specificity* and *accuracy* require a binary decision for their computation. Therefore, a threshold has to be defined translating the confidence rating into a binary decision. In a receiver operating characteristic (ROC) analysis, the confidence ratings are directly used to determine the ROC curve and the area under the ROC curve (AUC) [73]. The different metrics will be explained below. In addition to the before mentioned performance metrics, subjective image quality can be assessed by readers evaluating artifacts and the image impression [63, 74].

3.1.1 Binary Performance Classifiers

The most frequently used performance metrics are based on the binary decision abnormality present/absent. If a reader rates that an abnormality is present, the rating is counted as positive– else the rating is negative. Ratings of humans are compared to the existing truth for each case. Correct positive ratings are counted as true positives (TP), incorrect positive ratings are counted as false positives (FP). True negatives (TN) and false negatives (FN) are determined analogously.

These terms are used to compute the following figures of merit that are well described in Ref. [75].

Sensitivity, also known as recall, gives the true positive rate of the study. It is defined as the fraction of TP ratings over all positive cases (TP+FN) in the study:

$$sensitivity = \frac{TP}{TP + FN}. \tag{3.1}$$

When the sensitivity is very high, almost all abnormalities are rated positive and even some negative cases may be rated as positive. Only in very rare cases abnormalities would be rated as negative. As a result, negative ratings almost definitely rule out abnormalities at high sensitivity [75].

Specificity is given by the true negative rate. Analogously to the sensitivity, it is calculated by:

$$specificity = \frac{TN}{TN + FP}. \tag{3.2}$$

When the specificity is very high, almost all control cases (cases without abnormality) are rated negative and even some cases with abnormality. However, only in very rare cases false positives occur. As a result, positive ratings almost definitely rule in the presence of abnormalities at high specificity [75].

Accuracy gives the ratio of correctly rated cases over all cases:

$$accuracy = \frac{TP + TN}{P + N}. \tag{3.3}$$

It gives the fraction of correctly classified cases in the study and is therefore an interesting metric of the overall performance.

Positive predictive value (PPV), or precision, is defined as the ratio of TP ratings over all positive ratings:

$$PPV = \frac{TP}{TP + FP}. \tag{3.4}$$

Hence, PPV gives the probability that a positive rating is truly positive.

Analogous to PPV, the *negative predictive value* (NPV) is defined as the ratio of TN ratings over all negative ratings:

$$NPV = \frac{TN}{TN + FN}. \tag{3.5}$$

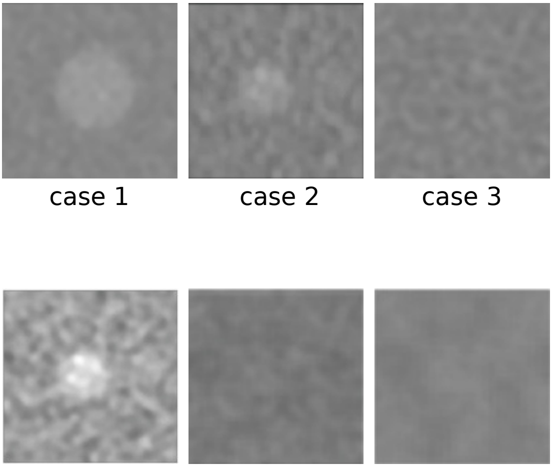
Hence, NPV gives the probability that a negative rating is truly negative [75].

3.1.2 Receiver Operating Characteristic

ROC analyses are used to quantify the performance of humans for a clinical task. It gives a measure of how good a reader can distinguish between positive and negative cases. Therefore, observers are required to give a confidence rating for a given task. Readers make their binary decision if an abnormality is present (has to be treated) with the attempt to maximize the benefits for the patient while costs and risks are minimized. Different readers may estimate the benefit-risk trade-off differently. With the ROC analysis, not the binary decision is of interest but if the reader can discriminate between an abnormality and normal cases [73]. The following section explains the generation of an ROC analysis with help of an example.

Example

We have a lesion detection task with the images in Figure 3.1. The observer is asked to read the images clinically and give a confidence rating between 1 and 100 for the presence of a lesion, where 1 is *"negative"* (definitely no lesion) and 100 is *"positive"* (lesion definitely present).



case 1 case 2 case 3

case 4 case 5 case 6

Figure 3.1: Image patches for exemplary ROC analysis.

Image patches simulate images with lesion (cases 1, 2, 4) and without lesion (cases 3, 5, 6). Simulated lesions are similar to liver lesion in a contrast enhanced CT examination of the liver.

Lets assume, the observer gives the rating shown in Table 3.1. For positive cases (with lesions) the reader gives ratings of 100, 90 and 50; for negative cases (without lesion) the reader gives ratings of 60, 30 and 10 (Table 3.2).

	case 1	case 2	case 3	case 4	case 5	case 6
rating	90	50	10	100	30	60
truth	positive	positive	negative	positive	negative	negative

Table 3.1: Exemplary rating in an ROC analysis.

score positive	score negative
50	10
90	30
100	60

Table 3.2: Ratings divided into positive and negative cases.

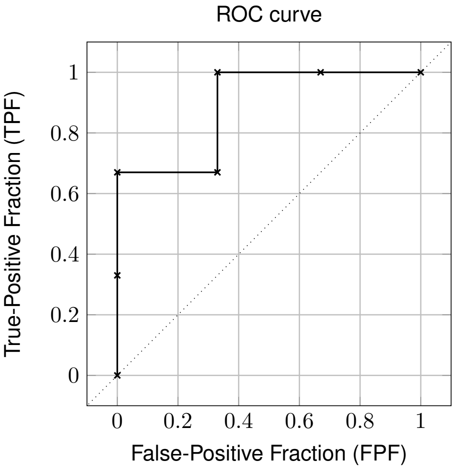


Figure 3.2: ROC curve.

ROC curve for the given example. Each data point is a (TPF, FPF)-pair for a different threshold (Table 3.3). The dotted diagonal line is the chance line.

The ROC curve is a plot of the true-positive fraction (TPF) against the false-positive fraction (FPF) corresponding to each possible threshold [73]. In our example, we get the following possible thresholds: 0, 11, 31, 51, 61, 91, 101. A threshold value of 11 means that all scores greater than 11 are considered positive (case 1, case 2, case4, case 5, case 6) and values below 11 are considered negative (case 3). Thus, we have three TP ratings (case 1, case 2, case 4), two FP ratings (case 5, case 6) and one TN rating (case 3), resulting in *TPF = 3/3* and *FPF = 2/3* for threshold 11. This procedure is repeated for each possible threshold (Table 3.3). For the generation of the ROC curve, these pairs (TPF,FPF) are plotted (Figure 3.2).

threshold	TPF	FPF
0	3/3 = 1.0	3/3 = 1.0
11	3/3 = 1.0	2/3 = 0.67
31	3/3 = 1.0	1/3 = 0.33
51	2/3 = 0.67	1/3 = 0.33
61	2/3 = 0.67	0/3 = 0
91	1/3 = 0.33	0/3 = 0
101	0/3 = 0	0/3 = 0

Table 3.3: True-positive fraction (TPF) and false-positive fraction (FPF) for each possible threshold.

The plot in Figure 3.2 is an empirical ROC curve without assumption of any distribution. There are approaches where statistical distributions are assumed that result in smother ROC curves [73]. However, the empirical ROC curve is very popular due to its easy computation. The figure of merit of an ROC analysis is the area under the ROC curve (AUC). An AUC of one indicates perfect discrimination between positive and negative cases, while an AUC of 0.5 (ROC curve would lie on the dotted chance line in Figure 3.2) suggests random rating of the cases equal to guessing. In the example provided above, the AUC is 0.89 indicating a good discrimination between positive and negative cases.

3.2 Model Observers

The planning and operation of reader studies are very time consuming and expensive. Several experts are needed to perform observer tasks on a big amount of images to create meaningful results. Therefore, high effort is spent on the investigation of quantitative image quality metrics to substitute reader studies. Anthropomorphic model observers are methods to assess task-based image quality that can be tuned to perform similar to human observers [11, 12]. In JP-I, novel model observers are implemented with machine-learning approaches and compared to a more traditional model observer, the channelized Hotelling observer (CHO). This section gives an introduction of the CHO and machine-learning model observers based on neural networks and convolutional neural networks (CNN).

3.2.1 Channelized Hotelling Observer

Model observers require training data to build an observer template. Common model observer tasks are the detection of signals in a signal known exactly/background known exactly (SKE/BKE) paradigm, meaning that both, images of the background without signal and images of the signal on top of the background are exactly known and available for the generation of the model observer.

A very popular model observer is the CHO, where images are filtered with frequency channels before computation of a decision variable. The filtering process simulates the activity of individual neurons in the visual cortex [11]. For a given image, the CHO returns a decision variable that a signal is present. The decision variable λ can be computed by

$$\lambda = \sum_{m=1}^{M} \omega_m g_m, \tag{3.6}$$

where ω is the observer template, g is the filtered image and M is the number of filter channels. In this work, Gabor channels are used because they showed good results in related work with a similar signal detection tasks [76, 77]. The implementation of the CHO is based on the description in References [76, 78]: the observer template ω is built upon the intraclass channel scatter matrix of the filtered training images when the signal is present and absent. The intraclass channel scatter matrix is the average of the channel output covariance matrix when the signal is present and absent.

Model observers showed superior results than humans in signal detection tasks. Therefore, random noise models were introduced and added to the model observers. These noise models are commonly called *internal noise*. In Ref. [79], Brankov gives an overview of different internal noise strategies. In this work, a proportional internal noise model was implemented. The decision variable λ_i with internal noise can be expressed by

$$\lambda_i = \sum_{m=1}^{M} \omega_m (g_m + \epsilon_m), \tag{3.7}$$

with

$$\epsilon_m \sim N(0, \alpha \cdot \delta_m). \tag{3.8}$$

ϵ_m is a noise vector with normal distribution of zero mean and δ_m standard deviation. α is tuning parameter to control the strength. The standard deviation reflects the roughness in signal absent images and can be derived from

$$\delta_m^2 = var\{(g_b)_m\}, \tag{3.9}$$

where var is the variance and $(g_b)_m$ are the respective channel outputs for signal-absent images.

For the evaluation of the model observer output, it is applied to a large set of images and the AUC is computed from the obtained decision variables. Therefore, model observers are metrics for task based diagnostic image quality assessment.

3.2.2 Machine Learning Model Observer

The mathematical principles of traditional model observers don't allow a high degree of signal or background uncertainty. For more realistic tasks with signals of varying sizes or shapes and lumpy backgrounds, the performance is degraded. Machine-learning algorithms, on the other hand, showed great results in challenging image classification tasks with performances similar to humans [21, 80]. Therefore, new approaches based on machine learning were investigated in literature [81–88]. In this work, neural network and CNN-based model observers were investigated. The principles of these machine learning-based model observers are described in this section and published in JP-I.

Neural Network Model Observer

The neural network model observer (NN-MO) is illustrated in Figure 3.3 and was implemented using the TensorFlow framework (TensorFlow version 1.2.1, Google Inc., Mountain View, USA). The neural-network architecture implemented in this study consists of an input layer and an output layer with a softmax function. The input image is reshaped to a 1 x M vector, with M being the total number of pixels. Each pixel x_m is an input to each neuron net_k, $k = 0, ..., K - 1$, with K being the number of neurons. In this setting, the number of neurons is equal to the confidence rating of a human observer in a reader study. In JP-I, the human observer gave a confidence rating between one and nine resulting in nine neurons in the NN-MO. After the neurons perform a linear regression, results of all neurons are normalized by a softmax function returning the likelihood that the image is of class k–equal to confidence k that the signal is present. The softmax function is given by

$$P(y|x) = \frac{e^{n_y}}{\sum_{k=1}^{K} e^{n_k}}, \tag{3.10}$$

with n_y and n_k being the outputs of neurons net_y and net_k, respectively. Weights w and biases b are optimized during training using the Adam optimizer with backpropagation [89].

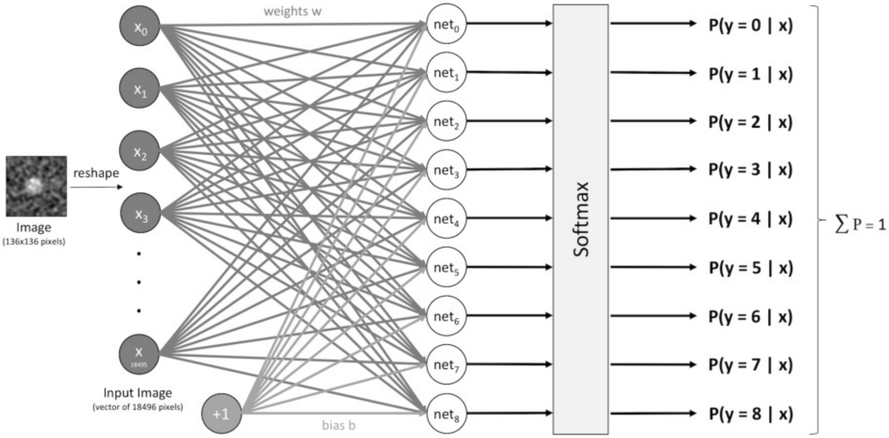

weights w
Image
(136x136 pixels)
reshape
x_0
x_1
x_2
x_3
x_{18495}
Input Image
(vector of 18496 pixels)
+1
bias b
net_0
net_1
net_2
net_3
net_4
net_5
net_6
net_7
net_8
Softmax
P(y = 0 | x)
P(y = 1 | x)
P(y = 2 | x)
P(y = 3 | x)
P(y = 4 | x)
P(y = 5 | x)
P(y = 6 | x)
P(y = 7 | x)
P(y = 8 | x)
$\sum P = 1$

Figure 3.3: Neural network-based model observer (NN-MO).

Figure is published in JP-I . Permission to copy was obtained from the publishing journal.

Convolutional Neural Network Model Observer

The architecture of the CNN model observer (CNN-MO) is based on LeNet-5 which was initially proposed for handwritten digit recognition [21]. As illustrated in Figure 3.4, the CNN-MO consists of two convolutional layers, each followed by a max-pooling layer, and two fully connected layers followed by an output layer and a softmax regression. Analogous to the NN-MO, the result is the probability P for the confidence rating of a human observer.

The first convolutional layer consists of six filters (dim: 5 x 5) that are applied to the input image of size 136 x 136. Each filter creates a feature map of the input image resulting in a volumetric output of dimensions 136 x 136 x 6. Max pooling with size 2 x 2 and a stride of 2 reduces the dimension of the feature maps to 68 x 68 x 6. Max pooling can be understood as a compression step: on each feature map, a window of size 2 x 2 slides over the map with a step size (stride) of two and the maximum value from each window position is preserved (see Figure 3.5). The second convolutional layer consists of 16 filters (dim: 5 x 5 x 6) resulting in an output of 68 x 68 x 16, succeeded by 2 x 2 max pooling. Resulting feature maps (dim: 34 x 34 x 16) are reshaped to a 1D-vector and fed into the first fully connected layer with 120 neurons. The second fully connected layer consists of 84 neurons followed by an output layer with 9 neurons and a softmax regression. Adam optimizer with backpropagation was used for training.

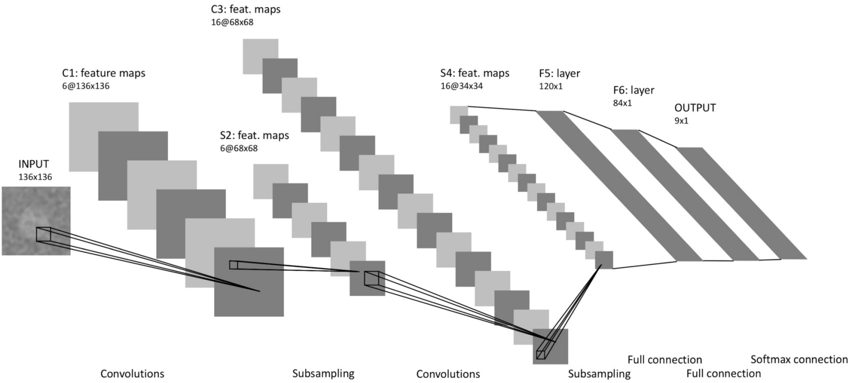
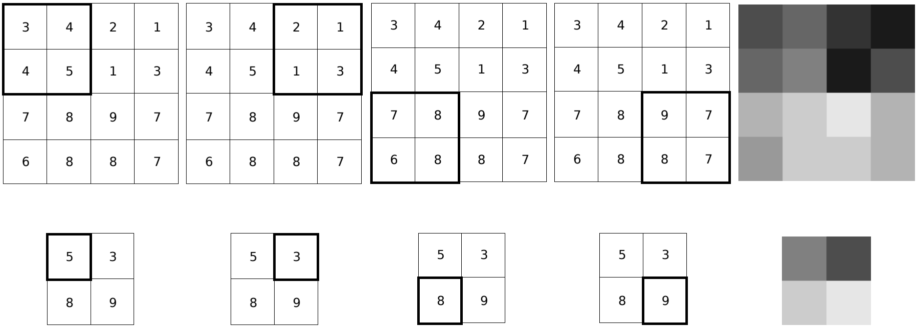

C3: feat. maps
16@68x68
C1: feature maps
6@136x136
S4: feat. maps
16@34x34
F5: layer
120x1
F6: layer
84x1
OUTPUT
9x1
S2: feat. maps
6@68x68
INPUT
136x136
Full connection
Softmax connection
Convolutions
Subsampling
Convolutions
Subsampling
Full connection

Figure 3.4: Convolutional neural network-based model observer (CNN-MO).

The architecture of the proposed CNN consists of an input layer, several hidden layers and an output layer. Filter numbers and resolution of convolutional layers are indicated on top of the respective layers. For the fully connected layers, the number of neurons is denoted. Figure is published in JP-I and adapted from [21]. Permission to copy was obtained from the publishing journal.



3	4	2	1
4	5	1	3
7	8	9	7
6	8	8	7

5	3
8	9

Figure 3.5: Max pooling of size 2 x 2 and stride 2.

A window of size 2 x 2 slides over the map with a step-size (stride) of two. The maximum value from each window position is preserved. The upper row shows the feature map before max pooling with the bold frame expressing the window position of each step. The lower row shows the feature map after max pooling with the bold frame marking the value corresponding to the window in the upper row. The images on the right show a gray scale representation of the feature map before and after max pooling (level: 5, window: 10). It can be seen that the lower image is a compressed version of the upper image.

In JP-I, labeled image data, as shown in Figure 3.1, with the confidence ratings of a radiologist as labels were used for training. The dataset was split into three sets: 80 % training, 5 %

validation and 15 % testing. The validation set is used to avoid overfitting: when the validation accuracy starts to converge or drop, the training needs to be stopped. In JP-I, the validation accuracy of the CNN-MO converged at about 70 % with a training accuracy of about 90 % at this point. Hence, training was stopped when the validation accuracy reached 70 % and the training accuracy was greater than 90 %. Analogous to the CHO, with machine learning-based model observers diagnostic image quality is evaluated by their application to a large set of images and the computation of the AUC from the predicted confidence ratings.

3.3 Custom Phantoms for Specific Applications

Another strategy to evaluate the diagnostic value of new imaging techniques is to use anthropomorphic phantoms. Due to the ionizing radiation in x-ray CT, new technologies like photon-counting CT need extensive testing before first experiments with humans are approved. For this purpose, phantom studies are very useful. With the broad availability of 3D-printers today, the creation of custom phantoms for specific applications is simplified, leading to increasing numbers of studies with custom-made phantoms [15, 90–92]. In JP-II, a custom-made lung phantom was used to evaluate a preclinical SPCCT system. The lung phantom is illustrated in Figure 3.6 and its design and evaluation is described in the following sections.

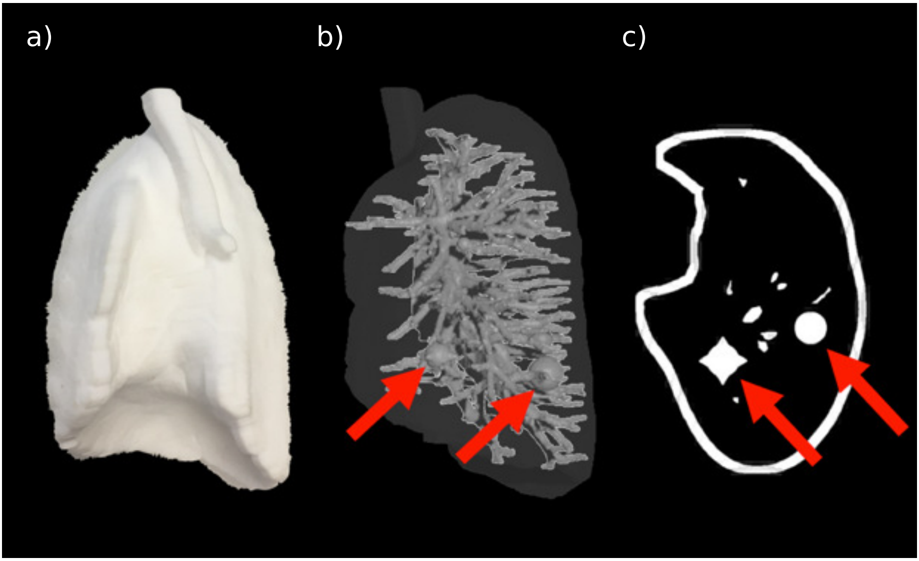


Figure 3.6: Anthropomorphic lung phantom.

a) 3D-printed lung phantom; b) 3D-rendering of the inner structure of the lung phantom; c) slice of the binary mask used for 3D printing. Red arrows in mark two inserted nodules–a sphere and a sphere with spikes representing benign and malignant nodules, respectively. Figure is published in C8. Permission to copy was obtained from the publishing journal.

3.3.1 Design of a Custom Lung Phantom

A board-certified radiologist selected a high-resolution CT acquisition of a healthy patient for the design of the lung phantom. From the scan, the right lung was manually segmented using a commercial DICOM viewer (OsiriX v.6.0 64-bit, Pixmeo SARL, Bernex, Switzerland). Further processing steps were performed using MATLAB (version R2017b, MathWorks, Massachusetts, USA). Inner structure was separated from the background by thresholding. Small parts that are not connected to the structure were removed by a region-growing algorithm that started in the boundary of the lung. Lung nodules were simulated and inserted into the lung segmentation. The experienced board-certified radiologist specified the nodule location to guarantee a realistic setup. Figure 3.7 shows a CT image of the lungs, the segmented right lung and the segmentation including simulated nodules. The simulated nodules were mathematically defined as spheres and spheres with spikes similar to the FDA lung-phantom inserts [93], representing benign and malignant nodules, respectively. Figure 3.8 illustrates the mathematical definition of the nodules. Afterwards, the digital lung phantom was loaded into an open source image processing software (ImageJ v. 2.0.0-rc-49/1.51d) and the *ImageJ 3D Viewer* plugin was used to create a surface rendering of the phantom. The surface rendering was exported as STL-file, that is required by most 3D-printing systems. A 3D-printing system with additive manufacturing technique of selective laser sintering based on polyamide was used to build the phantom. Printed vessels and surrounding walls were required to match the attenuation of human vessels. As a result of the 3D-printing technique, the lung phantom is filled with powdered polyamide. Therefore, the background in the lung has elevated intensity values compared to air. In the high-resolution scan of the lungs, intensity values of -875 HU were measured for background in the lung; a CT scan of the phantom showed about -580 HU.

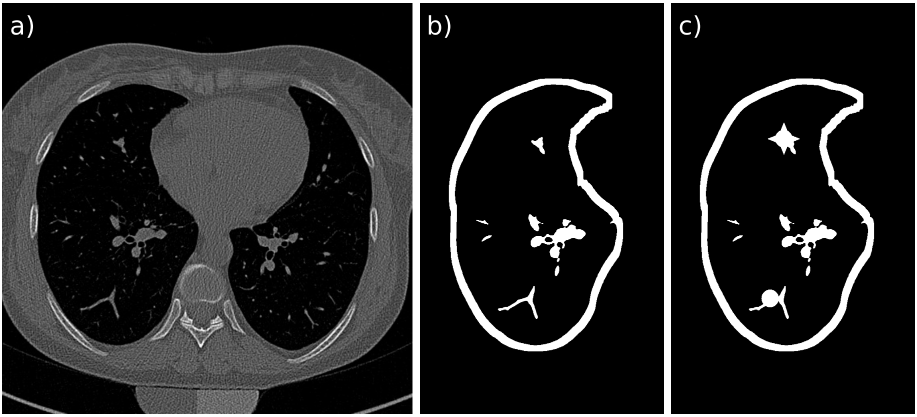


Figure 3.7: CT of the lungs and segmentation of the right lung.

a) High-resolution CT of a human thorax (level: 400 HU, window: 2400 HU); b) segmentation of the right lung after removal of small structures; c) segmentation with inserted lung nodules.

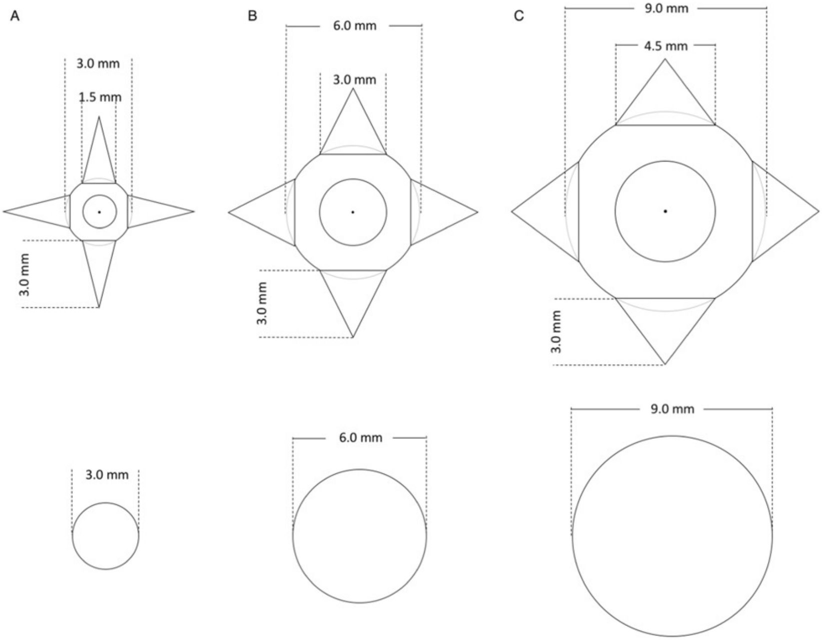


Figure 3.8: Description of the lung nodules.

Upper row illustrates spheres with spikes simulating malignant nodules; lower row illustrates spheres simulating benign nodules. Figure is published in JP-II.

3.3.2 Evaluation

Pulmonary structure inside the lung, position and shape of nodules as well as the nodule sizes are known. The availability of a digital model of the lung phantom that represents the ground truth enables visual (subjective) and quantitative evaluation. Figure 3.9 shows a comparison of the lung phantom scanned with different modalities. The green box shows the 3 mm sphere with spikes (Figure 3.8). Visually, images from SPCCT (G, H) are closer in appearance to the digital model (A, B) with sharp edges and clear boundaries. Therefore, the malignant nodule could be identified while in standard CT (C, D) and high-resolution CT (E, F) the spikes and boundaries are fuzzy. However, the visual impression of the images is a subjective evaluation that might differ from observer to observer. Therefore, shape and volume of the lung nodules are evaluated quantitatively. The methodology of the evaluation can be found in JP-II and is described below.

Nodule Segmentation

For the evaluation of nodule volume and shape, the nodule needs to be segmented. One operator selects manually the center of mass of the nodule and a circular region of interest (ROI) is drawn around the center. The diameter of the ROI is *d+1.5 mm* for spherical nodules and *d+6.5 mm* for spheres with spikes, where *d* is the size of the nodule. Based on the selected ROI, a volume of interest (VOI) is selected around the center of mass. Within the VOI, lung tissue is separated from

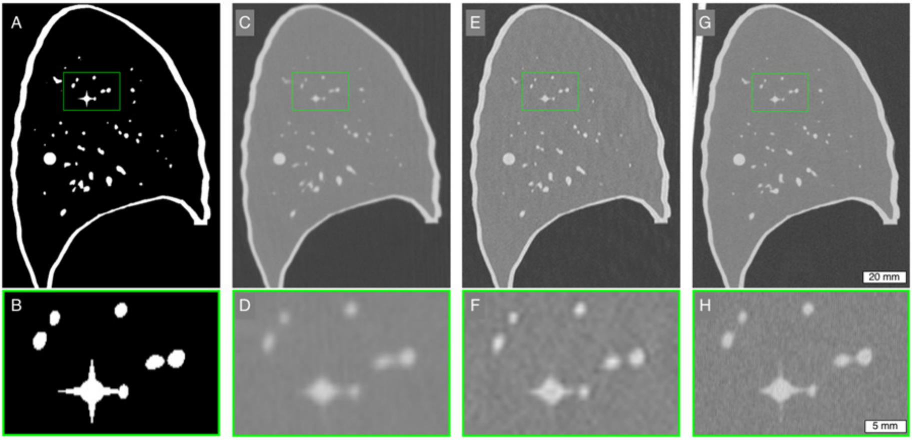


Figure 3.9: Visual comparison of lung images from different modalities.

A, B) Reference image (digital lung phantom); C, D) standard CT acquisition; E, F) CT acquisition with high-resolution protocol; G, H) SPCCT acquisition. Figure is published in JP-II.

background using the *k-means* clustering algorithm provided by MATLAB [94]. Parts that are not connected to the estimated center of mass are deleted from the segmentation by applying a 3D region-growing algorithm. In Figure 3.10, the selected ROI and the corresponding segmentation are exemplarily shown for the 6mm sphere with spikes. Due to the realistic placement of the nodules into the pulmonary structure, vessels connected to the center of mass are included within the VOI. To minimize the effect of manually selecting the center of mass in JP-II, results were reported as average values from three repeated segmentations.

Nodule Volume

The determination of the nodule size is an important task in pulmonary imaging. Size and growth of nodules are factors that are included in the classification. Nodule volume was determined based on the segmentation of each nodule by

$$V = N \cdot s, \tag{3.11}$$

where N is the number of voxels in the segmentation and s is the size of one voxel. Volumes were determined for each nodule in the digital lung phantom as reference. Determined volumes for each modality were compared to the reference with linear regression analysis and a Bland-Altman plot.

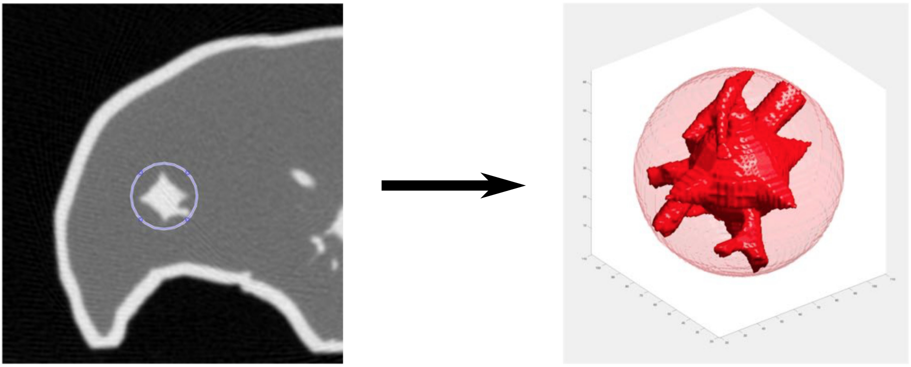


Figure 3.10: Nodule segmentation for evaluation.

The left image shows the center slice of the 6 mm sphere with spikes. A circular region of interest (ROI) with diameter 12.5 mm is manually drawn around the center of the nodule. A spherical volume of interest (VOI) based on the circular ROI is selected for segmentation. In the selected VOI, tissue is separated from background using a k-means segmentation. Parts that are inside the VOI but not connected to the center of the nodule are removed from the segmentation. The right image shows the final segmentation of the nodule (opaque red) and the surrounding VOI (light transparent red), that is used for evaluation.

Nodule Shape

Nodule shape is a crucial property in the assessment of lung nodules because it allows the determination of malignancy. In general, spherical nodules are classified as benign and nodules with proliferation, simulated by spikes, indicate malignancy. In JP-II, nodule shapes are evaluated by computing the Dice similarity coefficient (DSC) between nodule segmentations from each modality and the reference.

For a valid computation of the similarity between two segmentations, a precise registration is essential. Therefore, segmentations were resized to have an isotropic voxel size (0.14 x 0.14 x 0.14 mm^3) and rotation angles were measured with respect to the reference. Segmentations were rotated by the measured angles around the x-, y-, and z-axis. 2D cross-correlation was computed between center slices of the rotated segmentation and its reference. Precise alignment was achieved by rotating the segmentation to the maximum value of the 2D cross-correlation. The DSC was computed by

$$dsc(A, B_m) = 2 \cdot \frac{|A \cap B_m|}{|A| + |B_m|}, \tag{3.12}$$

with A being the reference, B_m is the segmentation for modality m, $\cap$ denotes the intersection of two sets and $| \cdot |$ is the cardinal of a set. The result is a value between 0 and 1, with 1 indicating perfect similarity.

4 Compliance with Ethical Standards

All investigations performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. Written informed consent was waived by the institutional review board (Ethikkommision der Medizinischen Fakultät, Technical University of Munich, Germany) as all patients were included retrospectively.

The animal experiment involving the scan of a New Zealand white rabbit with a SPCCT prototype in Lyon, France, was approved by the French Department of Education and Research. All experiments were performed in accordance with relevant guidelines and regulations.

5 Comprising Journal Publications

5.1 Journal Publication I: *CNN as model observer in a liver lesion detection task for x-ray computed tomography: A phantom study*

The publication entitled *CNN as model observer in a liver lesion detection task for x-ray computed tomography: A phantom study* was published in Medical Physics (ISSN: 2473-4209). The manuscript was authored by Felix K. Kopp, Marco Catalano, Daniela Pfeiffer, Alexander A. Fingerle, Ernst J. Rummeny, and Peter B. Noël.

5.1.1 Abstract

Purpose

The purpose of this study was the evaluation of anthropomorphic model observers trained with neural networks for the prediction of a human observer's performance.

Methods

To simulate liver lesions, a phantom with contrast targets (acrylic spheres, varying diameters, +30 HU) was repeatedly scanned on a computed tomography scanner. Image data labeled with confidence ratings assessed in a reader study for a detection task of liver lesions were used to build several anthropomorphic model observers. Models were trained with images reconstructed with iterative reconstruction and evaluated with images reconstructed with filtered backprojection. A neural network, based on softmax regression (SR-MO), and convolutional neural networks (CNN-MO) were used to predict the performance of a human observer and compared to a channelized Hotelling observer [with Gabor channels and internal channel noise (CHOi)]. Model observers were evaluated by a receiver operating characteristic curve analysis and compared to the results in the reader study. Two strategies were used to train the SR-MO and CNN-MO: A) building a separate model for each lesion size; B) building one model that was applied to lesions of all sizes.

Results

All tested model observers and the human observer were highly correlated at each lesion size and dose level. With strategy A, Pearson's product-moment correlation coefficients r were 0.926 (95% confidence interval (CI): 0.679–0.985) for SR-MO and 0.979 (95% CI: 0.902–0.996) for CNN-MO. With strategy B, r was 0.860 (95% CI: 0.454–0.970) for SR-MO and 0.918 (95% CI: 0.651–0.983) for CNN-MO. For CHOi, r was 0.945 (95% CI: 0.755–0.989). With strategy A, mean absolute percentage differences (MAPD) between the model observers and the human observer were 3.7% for SR-MO and 1.2% for CNN-MO. With strategy B, MAPD were 3.7% for SR-MO and 3.0% for CNN-MO. For the CHOi the MAPD was 2.2%.

Conclusion

Convolutional neural network model observers can accurately predict the performance of a human observer for all lesion sizes and dose levels in the evaluated signal detection task.

5.1.2 Author contributions

The first author performed the experiments, implemented the channelized Hotelling observer, neural network and convolutional neural network, and programmed the reader study assessment tool. With the help and consultation from the coauthors, the first author designed the experiment, analyzed and interpreted the data, and wrote the paper.

5.2 Journal Publication II: *Evaluation of a preclinical photon-counting CT prototype for pulmonary imaging*

The publication entitled *Evaluation of a preclinical photon-counting CT prototype for pulmonary imaging* was published in Scientific Reports (ISSN: 2045-2322). The manuscript was authored by Felix K. Kopp, Heiner Daerr, Salim Si-Mohamed, Andreas P. Sauter, Sebastian Ehn, Alexander A. Fingerle, Bernhard Brendel, Franz Pfeiffer, Ewald Roessl, Ernst J. Rummeny, Daniela Pfeiffer, Roland Proska, Philippe Douek, and Peter B. Noël.

5.2.1 Abstract

The purpose of this study was to investigate a preclinical spectral photon-counting CT (SPCCT) prototype compared to conventional CT for pulmonary imaging.

A custom-made lung phantom, including nodules of different sizes and shapes, was scanned with a preclinical SPCCT and a conventional CT in standard and high-resolution (HR-CT) mode. Volume estimation was evaluated by linear regression. Shape similarity was evaluated with the Dice similarity coefficient. Spatial resolution was investigated via MTF for each imaging system. *In-vivo* rabbit lung images from the SPCCT system were subjectively reviewed.

Evaluating the volume estimation, linear regression showed best results for the SPCCT compared to CT and HR-CT with a root mean squared error of 21.3 mm^3, 28.5 mm^3 and 26.4 mm^3 for SPCCT, CT and HR-CT, respectively. The Dice similarity coefficient was superior for SPCCT throughout nodule shapes and all nodule sizes (mean, SPCCT: 0.90; CT: 0.85; HR-CT: 0.85). 10% MTF improved from 10.1 LP/cm for HR-CT to 21.7 LP/cm for SPCCT. Visual investigation of small pulmonary structures was superior for SPCCT in the animal study.

In conclusion, the SPCCT prototype has the potential to improve the assessment of lung structures due to higher resolution compared to conventional CT.

5.2.2 Author contributions

The first author designed and fabricated the lung phantom, acquired conventional CT data, implemented the analysis software for nodule quantification, and performed the statistical analysis. With the help and consultation from the coauthors, the first author designed the experiment and wrote the manuscript.

5.3 Journal Publication III: *Sparse sampling computed tomography (SpSCT) for detection of pulmonary embolism: a feasibility study*

The publication entitled *Sparse sampling computed tomography (SpSCT) for detection of pulmonary embolism: a feasibility study* was submitted to European Radiology (ISSN: 1432-1084) and published online May 9th 2019. The manuscript was authored by Andreas P. Sauter, Felix K. Kopp, Rolf Bippus, Julia Dangelmaier, Dominik Deniffel, Alexander A. Fingerle, Felix Meurer, Daniela Pfeiffer, Roland Proska, Ernst J. Rummeny, and Peter B. Noël. Andreas P. Sauter and Felix K. Kopp equally contributed with co-first authorship.

5.3.1 Abstract

Objectives

Evaluation of sparse sampling computed tomography (SpSCT) regarding subjective and objective image criteria for the detection of pulmonary embolism (PE) at different simulated dose levels.

Methods

Computed tomography pulmonary angiography (CTPA) scans of 20 clinical patients were used to obtain simulated low-dose scans with 100%–50%–25%–12.5%–6.3%–3.1% of the clinical dose, resulting in a total of six dose levels (DL). From these full sampling (FS) data, every second (2-SpSCT) or fourth (4-SpSCT) projection was used to obtain simulated sparse sampling scans. Each image set was evaluated by four blinded radiologists regarding subjective image criteria (artifacts, image quality) and diagnostic performance (confidence, sensitivity, specificity, accuracy, and area under the curve). Additionally, the contrast-to-noise ratio (CNR) was evaluated for objective image quality.

Results

Sensitivity was 100% with 2-SpSCT and 4-SpSCT at the 25% DL and the 12.5% DL for all localizations of PE (one subgroup 98.5%). With FS, the sensitivity decreased to 90% at the 12.5% DL. 2-SpSCT and 4-SpSCT showed higher values for sensitivity, specificity, accuracy, and the area under the curve at all DL compared with FS. Subjective image quality was significantly higher for 4-SpSCT compared with FS at each dose level ($p < 0.01$, paired t-test). Only with 4-SpSCT, all examinations were rated as showing diagnostic image quality at the 12.5% DL.

Conclusion

Via SpSCT, a dose reduction down to a 12.5% dose level (corresponding to a mean effective dose of 0.38 mSv in the current study) for CTPA is possible while maintaining high image quality and full diagnostic confidence.

5.3.2 Author contributions

FKK coordinated the study, processed the reconstructed images and created the reader study assessment sheets. The co-first authors designed the experiment and analyzed the data. With the help and consultation from the coauthors, the first authors interpreted the data and wrote the manuscript.

6 Discussion

This cumulative dissertation is based on three publications providing clinical evaluation for different novel imaging procedures in the field of x-ray CT. Following sections discuss the included journal publications and give a conclusion of this work.

6.1 JP-I

CNN as model observer in a liver lesion detection task for x-ray computed tomography: A phantom study evaluates machine-learning approaches for the assessment of task-based diagnostic image quality. The idea is to implement model observers based on neural networks and CNNs that can predict the performance of a human observer in a liver lesion detection task. Apart from using machine learning as model observers, a novelty is the investigation of the performance under varying lesion sizes. Traditional model observers like the CHO will typically fail with signal size uncertainty, due to their mathematical definition. However, lesion sizes vary in the clinical routine with single observers (radiologists) assessing lesions of different sizes. Therefore, new strategies are needed for the adaption of model observers to more realistic tasks.

Brankov et al. introduced machine learning-based model observers for single photon emission computed tomography (SPECT) [81, 82, 84]. They showed that nonlinear model observers based on support vector machines (SVM) [81, 82] and a neural network [84] perform superior to CHO with the advantage of directly learning from the human's decisions. While traditional model observers are built to detect signals and their performance is decreased by adding noise to the decision–as adaption to human inefficiencies in challenging detection tasks–machine learning-based model observers can be directly trained from the human's performance. In a next step, Massanes et al. evaluated a CNN-based model observer imaging achieving "excellent agreement" with human observers. They also investigated the effect of reducing the number of training images, concluding that CNNs need a larger amount of images for excellent results [88]. However, they only investigated a single signal in different noise realizations for SPECT. In the present thesis, different signal sizes were evaluated making the detection task more realistic. In addition, the evaluation for CT images verifies the application of machine learning-based model observers for x-ray CT.

Gong et al. proposed a deep learning-based model observer (DL-MO) for CT and introduced anatomical patient background to the lesions as a follow-up to JP-I. They used the so-called transfer learning strategy, where an existing CNN that was trained on a huge amount of image data is fine tuned to a new task. The authors trained an individual DL-MO for each lesion size, dose level and reconstruction algorithm that were used in the study. DL-MOs were trained to detect lesions and internal noise was added to reduce the performance [95, 96]. The reported Pearson correlation coefficient r between DL-MOs and human observers was 0.986 (95 % confidence interval (CI): 0.950 - 0.996). That is in the same range as for strategy A in JP-I, where a separate model observer was trained for each signal size (r=0.979; 95 % CI: 0.902 - 0.996). However, Gong et al. didn't take advantage of the possibility to directly train the model observer on the data of human observers. Moreover, they built a DL-MO for each lesion and image quality. In JP-I, strategy B is used to train machine learning-based model observers that are capable to work with different lesion sizes simulating more realistic tasks.

There were limitations in JP-I. The human observer's performance is based on the results of a single radiologist in a reader study. Different readers may have different performances on the same detection task. Therefore, results obtained in this study cannot be generalized. However, aim of the study was the proof-of-concept if a machine learning-based model observer could predict the performance of a human observer. In a scenario where the performance of several radiologists need to be predicted, also several model observers could be trained–one for each human observer–as proposed by Marin et al. [87]. Another limitation is that neural networks and CNNs are known to perform better with more training data available. In realistic settings, the number of images is limited due to data privacy protection leading to a decreased number of available training images for real scenarios. There are strategies to overcome the requirement of large datasets, like transfer learning [97] and data augmentation [98, 99]. The application of such strategies could be part of future work in this research area.

In conclusion, the results of JP-I indicate that neural networks and CNNs might be suitable for human observer's performance prediction and the evaluation of diagnostic image quality. However, limitations show that further investigation is advisable.

6.2 JP-II

Evaluation of a preclinical photon-counting CT prototype for pulmonary imaging investigates the clinical applicability of SPCCT for imaging of the lung using a custom-made lung phantom. Results were compared to conventional CT in standard and high-resolution (HR-CT) mode. Higher spatial resolution could lead to an improved classification of small nodules [100] and allow to assess wall thickness of large and small airways, enabling an early detection of the destruction of alveolar architecture which is a direct effect of chronic obstructive pulmonary disease (COPD) [101, 102]. Therefore, high-resolution CT is an important topic in pulmonary imaging.

Pourmorteza et al. evaluated the high-resolution capabilities of a different SPCCT system [23]. The prototype is based on a SOMATOM Flash scanner (Siemens Healthcare, Forchheim, Germany) with a binned detector pixel size of 450 x 450 μm^2 in high-resolution mode. They reported improved image resolution compared to standard-resolution photon-counting CT. The SPCCT prototype evaluated in JP-II is based on a Brilliance iCT (Philips Healthcare, Haifa, Israel) with a detector pixel size of 500 x 500 μm^2. Improved image resolution was measured and enhanced nodule quantification was observed compared to conventional CT and HR-CT.

Zhou et al. evaluated the Siemens prototype for the quantification of lung nodules. They reported a root-mean-square error (RMSE) for nodule volume quantification between 21.6 - 28.3 mm^3 [103] and between 21.3 - 22.5 mm^3 [16]. In JP-II, the RMSE for nodule volume quantification was in the same range (21.3 - 28.5 mm^3). In addition, JP-II provides a comparison to a state-of-the-art CT system and subjectively evaluates images of an animal study.

Yanagawa et al. evaluated an ultra-high-resolution CT (U-HRCT) for lung imaging and compared it to a conventional CT [104]. The U-HRCT is equipped with a detector of similar resolution as the presented SPCCT prototype. They reported significantly improved visualization for most investigations with significantly higher overall image quality compared to the conventional CT. However, a fair comparison between U-HRCT and SPCCT would require quantitative evaluation of both systems with the same methodology and should be investigated in future work.

One limitation of the study in JP-II is the use of FBP for image reconstruction. Advanced iterative reconstruction algorithms might have improved the quantification results in the study for both, conventional CT [63] and SPCCT [105]. However, iterative reconstruction introduces non-linearity leading to limited value of the results [9]. Therefore, FBP was used to guarantee a fair comparison. Another limitation is the uncertainty in the 3D-printing process. Small errors might be introduced by interpolation when the digital model is 3D rendered and finally 3D printed. These inaccuracies might contribute to the reported RMSE between 21.3 - 28.5 mm^3 (SPCCT: 21.3 mm^3; HR-CT: 26.4 mm^3; CT: 28.5 mm^3). However, Zhou et al. reported similar RMSEs for photon-counting CT.

In conclusion, different studies indicate a great potential of high-resolution photon-counting CT with systems from different vendors. However, despite the fact that photon-counting CT systems from different vendors are based on similar concepts aiming the same goal, specific system parameters differ (like detector pixel size) and therefore, require separate evaluation of different systems. JP-II showed that the investigated SPCCT prototype allows precise lung nodule quantification and improved investigation of small pulmonary structures.

6.3 JP-III

Sparse sampling computed tomography (SpSCT) for detection of pulmonary embolism: a feasibility study evaluates sparse sampling as novel CT acquisition scheme for radiation dose reduction in examinations with suspected PE. Results on simulated data indicate that a dose reduction to 12.5 % of the currently used clinical dose would still produce diagnostic image quality for examinations with suspected PE. However, results are based on simulated data because there is no medical CT with sparse-sampling capabilities available today.

Wiedmann et al. give a summary of hardware modifications that could allow sparse-sampling CT [25]. They presented results of a multisource inverse-geometry CT prototype that is capable to pulse the x-ray flux with a pulse width on the order of 3 μs. This would allow to switch an x-ray tube on and off at arbitrary positions during one gantry rotation. However, the technique is still under investigation [106, 107] and not available in commercial medical CT systems. Chen and Muckley et al. presented first results of a sparse-sampling CT prototype with integration of a multi-slit collimator between the x-ray source and the patient [108, 109]. The x-ray flux is blocked row-wise to obtain a sparse data acquisition with reduced dose [110]. Today, presented work only proves the feasibility of such techniques and shows that more evaluation for the implementation in commercial medical CT systems is required.

Abbas et al. investigated several sampling schemes on micro-CT data [111]. Sparsely sampled data is simulated by row- or column-wise blocking of the x-ray beam for every tube position, periodically sampling one projection after every few tube positions and periodically sampling a group of projections after every few tube positions. For reconstruction, they used an optimized SIR algorithm [112, 113]. In their study, periodically sampling one projection after every few tube positions performs best reaching highest similarity to the fully sampled reference images. In JP-III, we used periodically sampling of one projection for the simulation of sparse sampling, resulting in good image quality supporting the findings by Abbas et al.

One limitation of this study is that it is based on simulations. Low tube current and sparse sampling was simulated based on real patient scans from the clinical routine. However, evaluation of reduced tube current would require additional x-ray radiation exposure to patients. Therefore, a vendor specific low-dose simulation tool was used to obtain virtually reduced tube currents [114]. The simulation tool was evaluated by Muenzel et al. [115], verifying the validity of simulated data. Sparse sampling, on the other hand, can only be simulated because there is no medical CT system available that is capable of performing sparse sampling. Another limitation is the low number of patients in the reader study. 20 patients with suspected PE were included in the study. Due to the wide range of parameters that were evaluated (6, 5 and 4 dose levels for full sampling, 2-times SpS and 4-times SpS, respectively) already a huge set of images were generated. With that amount of patients, the goal was to show the feasibility of sparse sampling for the investigation of cases with suspected PE. For generalization of the results, follow-up studies should narrow the parameters (less dose levels) and include a larger patient cohort.

In conclusion, JP-III evaluated sparse sampling for examinations with suspected PE in a reader study. The results suggest that a dose reduction of 87.5 % could be achieved without loss of subjective image quality and diagnostic performance.

6.4 Outlook and Conclusion

In this cumulative dissertation, different approaches for clinical evaluation of novel methodologies in x-ray CT were presented. While the implementation of machine learning-based model observers aims to provide automatic evaluation of diagnostic image quality, and the use of custom-made phantoms requires at least some human interaction, subjective evaluation of the diagnostic performance in a reader study relies solely on the human performance. Presented strategies for system evaluation were applied to different novel imaging techniques with the aim to estimate their clinical value.

The CNN-based model observer showed promising results in predicting the performance of a radiologist in a simulated lesion detection task with varying signal sizes. Future work should focus on their application to clinical images with anatomical background and real lesions. Beyond model observers, diagnostic image quality assessment could be trained by the subjective impression of radiologists. Lee et al. used transfer learning of a deep CNN with subjective image quality rating on lung images [116]. Whole images are rated for their visual appearance offering the advantage that it could be applied to any image without knowledge of present disease (as in model observers). However, even if radiologists approve the visual appearance of an image, it is not guaranteed that they have a good diagnostic performance on them. Thus, further research is needed for automated assessment of diagnostic image quality.

Clinical evaluation of the SPCCT prototype indicates great performance for pulmonary imaging. The investigated prototype has a limited field of view that cannot cover the whole human body. However, other prototypes exist that are capable of scanning whole human bodies. Moreover, the rapid evolution of photon-counting CT indicates a foreseeable clinical introduction within the near future [22].

With the evaluation of sparse sampling for examinations with suspected PE, the dose reduction possibility of SpSCT was presented. Although great potential of sparse sampling for low-dose CT was demonstrated, its integration into commercial medical CT systems still requires extensive research. The technical feasibility was already shown for multi-slit collimator and fast-pulsing x-ray tubes, but the performance of state-of-the-art detectors when the continuous flux is partly blocked or pulsed might be problematic. Current EIDs expect continuously impinging x-ray photons, hence, a change of this behavior needs cautious investigation. Therefore, high effort and extensive research is still required until the introduction of SpSCT to commercial CT systems.

Summing up, presented approaches and evaluations are a valuable contribution to the CT research community. While JP-I already builds a basis for ongoing research [95], JP-II and JP-III contribute to an extensive evaluation of new CT technologies that might promote their clinical introduction in the future.

List of Abbreviations

2-times SpS	two-times sparse sampling
4-times SpS	four-times sparse sampling
A-P	antero-posterior
AUC	area under the ROC curve
CHO	channelized Hotelling observer
CNN	convolutional neural network
CNN-MO	convolutional neural network-based model observer
CS	compressed sensing
CT	computed tomography
DECT	dual-energy computed tomography
DL-MO	deep learning-based model observer
DSC	Dice similarity coefficient
EID	energy-integrating detector
EVAR	endovascular aortic repair
FN	false negative
FP	false positive
HR-CT	high-resolution computed tomography
IR	iterative reconstruction
keV	kiloelectronvolt
kVp	peak kilovoltage
mA	miliampere
mAs	miliampere-second
ML	machine learning
MTF	modulation transfer function
NN-MO	neural network-based model observer

NPV	negative predictive value
PCD	photon-counting detector
PE	pulmonary embolism
PPV	positive predictive value
RMSE	root-mean-square error
ROC	receiver operating characteristic
ROI	region of interest
SIR	statistical iterative reconstruction
SKE/BKE	signal known exactly/background known exactly
SPCCT	single photon-counting computed tomography
SPECT	single photon emission computed tomography
SpSCT	sparse-sampling computed tomography
SVM	support vector machine
TCM	tube current modulation
TN	true negative
TP	true positive
U-HRCT	ultra-high-resolution computed tomography
VOI	volume of interest

List of Figures

1.1 Anthropomorphic phantoms. 4

2.1 Schematic representation of a CT system. 7
2.2 Schematic representation of an x-ray tube. 8
2.3 Schematic representation of an energy-integrating detector. Figure adapted from [5]. 10
2.4 X-ray spectrum for 120 kVp. 11
2.5 Conventional CT vs. SpSCT. 12
2.6 Schematic representation of a pulsing x-ray tube. 14
2.7 Overview scans. 15
2.8 Schematic representation of a photon-counting detector. 17
2.9 CT images of a spatial resolution phantom. 18
2.10 Charge sharing and secondary photons. 19
2.11 Schematic representation of a back-projection process. 20
2.12 Images of the sinogram (left), back projection (center) and FBP (right) of module CTP404 of a Catphan phantom. 20
2.13 Hybrid iterative reconstruction (IR). 22
2.14 Statistical iterative reconstruction (SIR). 23

3.1 Image patches for exemplary ROC analysis. 27
3.2 ROC curve. 28
3.3 Neural network-based model observer (NN-MO). 31
3.4 Convolutional neural network-based model observer (CNN-MO). 32
3.5 Max pooling of size 2 x 2 and stride 2. 32
3.6 Anthropomorphic lung phantom. 33
3.7 CT of the lungs and segmentation of the right lung. 34
3.8 Description of the lung nodules. 35
3.9 Visual comparison of lung images from different modalities. 36
3.10 Nodule segmentation for evaluation. 37

List of Tables

2.1 Dose reduction with SpSCT. 13
2.2 Higher tube current with SpSCT. 13

3.1 Exemplary rating in an ROC analysis. 27
3.2 Ratings divided into positive and negative cases. 27
3.3 True-positive fraction (TPF) and false-positive fraction (FPF) for each possible threshold. 28

Bibliography

[1] A. Sodickson, P. F. Baeyens, K. P. Andriole, L. M. Prevedello, R. D. Nawfel, R. Hanson, and R. Khorasani, “Recurrent ct, cumulative radiation exposure, and associated radiation-induced cancer risks from ct of adults,” *Radiology*, vol. 251, no. 1, pp. 175–184, 2009.

[2] D. J. Brenner, C. D. Elliston, E. J. Hall, and W. E. Berdon, “Estimated risks of radiation-induced fatal cancer from pediatric ct,” *American journal of roentgenology*, vol. 176, no. 2, pp. 289–296, 2001.

[3] D. J. Brenner and E. J. Hall, “Computed tomography—an increasing source of radiation exposure,” *New England Journal of Medicine*, vol. 357, no. 22, pp. 2277–2284, 2007.

[4] A. Ziegler, T. Köhler, and R. Proksa, “Noise and resolution in images reconstructed with fbp and osc algorithms for ct,” *Medical physics*, vol. 34, no. 2, pp. 585–598, 2007.

[5] J. Hsieh, *Computed Tomography: Principles, Design, Artifacts, and Recent Advances*. SPIE Press monograph, SPIE Press, 2003.

[6] G. Pahn, S. Skornitzke, H.-P. Schlemmer, H.-U. Kauczor, and W. Stiller, “Toward standardized quantitative image quality (iq) assessment in computed tomography (ct): A comprehensive framework for automated and comparative iq analysis based on icru report 87,” *Physica Medica*, vol. 32, no. 1, pp. 104–115, 2016.

[7] “Report 87,” *Journal of the International Commission on Radiation Units and Measurements*, vol. 12, pp. NP–NP, 04 2012.

[8] J. A. Fessler, “Assessment of image quality for the new ct: Statistical reconstruction methods,” *AAPM Interactive Session Imaging Scientific*, 2012.

[9] J. H. Pachon, G. Yadava, D. Pal, and J. Hsieh, “Image quality evaluation of iterative ct reconstruction algorithms: a perspective from spatial domain noise texture measures,” in *Medical Imaging 2012: Physics of Medical Imaging*, vol. 8313, p. 83132K, International Society for Optics and Photonics, 2012.

[10] H. H. Barrett, K. J. Myers, C. Hoeschen, M. A. Kupinski, and M. P. Little, “Task-based measures of image quality and their relation to radiation dose and patient risk,” *Physics in Medicine & Biology*, vol. 60, no. 2, p. R1, 2015.

[11] H. H. Barrett, J. Yao, J. P. Rolland, and K. J. Myers, “Model observers for assessment of image quality,” *Proceedings of the National Academy of Sciences*, vol. 90, no. 21, pp. 9758–9765, 1993.

[12] A. E. Burgess, B. Worthington, C. E. Metz, C. J. Taylor, C. R. Hill, D. C. Barber, D. E. Kuhl, D. G. Brown, K. J. Myers, M. A. Smith, P. Sharp, P. Wells, R. Brooks, and R. F. Wagner, “Report 54,” *Journal of the International Commission on Radiation Units and Measurements*, vol. os28, pp. NP–NP, 04 2016.

[13] A. P. Sauter, F. K. Kopp, D. Münzel, J. Dangelmaier, M. Renz, B. Renger, R. Braren, A. A. Fingerle, E. J. Rummeny, and P. B. Noël, “Accuracy of iodine quantification in dual-layer spectral ct: influence of iterative reconstruction, patient habitus and tube parameters,” *European journal of radiology*, vol. 102, pp. 83–88, 2018.

[14] S. Ehn, T. Sellerer, D. Muenzel, A. A. Fingerle, F. Kopp, M. Duda, K. Mei, B. Renger, J. Herzen, J. Dangelmaier, *et al.*, “Assessment of quantification accuracy and image quality of a full-body dual-layer spectral ct system,” *Journal of applied clinical medical physics*, vol. 19, no. 1, pp. 204–217, 2018.

[15] J. Dangelmaier, D. Bar-Ness, H. Daerr, D. Muenzel, S. Si-Mohamed, S. Ehn, A. A. Fingerle, M. A. Kimm, F. K. Kopp, L. Boussel, *et al.*, “Experimental feasibility of spectral photon-counting computed tomography with two contrast agents for the detection of endoleaks following endovascular aortic repair,” *European radiology*, pp. 1–8, 2018.

[16] W. Zhou, J. Montoya, R. Gutjahr, A. Ferrero, A. Halaweish, S. Kappler, C. H. McCollough, and S. Leng, “Lung nodule volume quantification and shape differentiation with an ultra-high resolution technique on a photon-counting detector computed tomography system,” *Journal of Medical Imaging*, vol. 4, no. 4, p. 043502, 2017.

[17] H. H. Barrett and K. J. Myers, *Foundations of image science*. John Wiley & Sons, 2013.

[18] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “Imagenet classification with deep convolutional neural networks,” in *Advances in neural information processing systems*, pp. 1097–1105, 2012.

[19] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, “Going deeper with convolutions,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 1–9, 2015.

[20] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 770–778, 2016.

[21] Y. LeCun, L. Bottou, Y. Bengio, P. Haffner, *et al.*, “Gradient-based learning applied to document recognition,” *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998.

[22] M. J. Willemink, M. Persson, A. Pourmorteza, N. J. Pelc, and D. Fleischmann, “Photon-counting ct: technical principles and clinical prospects,” *Radiology*, vol. 289, no. 2, pp. 293–312, 2018.

[23] A. Pourmorteza, R. Symons, A. Henning, S. Ulzheimer, and D. A. Bluemke, “Dose efficiency of quarter-millimeter photon-counting computed tomography: first-in-human results,” *Investigative radiology*, vol. 53, no. 6, pp. 365–372, 2018.

[24] E. J. Candès and M. B. Wakin, “An introduction to compressive sampling [a sensing/sampling paradigm that goes against the common knowledge in data acquisition],” *IEEE signal processing magazine*, vol. 25, no. 2, pp. 21–30, 2008.

[25] U. Wiedmann, V. B. Neculaes, D. Harrison, E. Asma, P. E. Kinahan, and B. De Man, “X-ray pulsing methods for reduced-dose computed tomography in pet/ct attenuation correction,” in *Medical Imaging 2014: Physics of Medical Imaging*, vol. 9033, p. 90332Z, International Society for Optics and Photonics, 2014.

[26] H. Alkadhi, S. Leschka, P. Stolzmann, and H. Scheffel, *Wie funktioniert CT?*, vol. 133. Springer, 2011.

[27] R. Behling, *Modern diagnostic x-ray sources: technology, manufacturing, reliability*. CRC Press, 2015.

[28] T. M. Buzug, “Computed tomography,” in *Springer Handbook of Medical Technology*, pp. 311–342, Springer, 2011.

[29] W. W. Mayo-Smith, A. K. Hara, M. Mahesh, D. V. Sahani, and W. Pavlicek, “How i do it: managing radiation dose in ct,” *Radiology*, vol. 273, no. 3, pp. 657–672, 2014.

[30] Siemens Healthcare GmbH, *Simulation of X-ray Spectra: Online tool for the simulation of X-ray Spectra*, 2019 (accessed April 24, 2019). `https://www.oem-xray-components.siemens.com/x-ray-spectra-simulation`.

[31] T. Fewell, R. Shuping, and K. Healy, “Handbook of computed tomography x-ray spectra 81-8162,” *HHS Publication (FDA)(US Government Printing Office), Washington, DC*, 1981.

[32] J. M. Boone and J. A. Seibert, “An accurate method for computer-generating tungsten anode x-ray spectra from 30 to 140 kv,” *Medical physics*, vol. 24, no. 11, pp. 1661–1670, 1997.

[33] F. K. Kopp, R. Bippus, A. P. Sauter, D. Muenzel, F. Bergner, K. Mei, J. Dangelmaier, B. J. Schwaiger, M. Catalano, A. A. Fingerle, *et al.*, “Diagnostic value of sparse sampling computed tomography for radiation dose reduction: initial results,” in *Medical Imaging 2018: Physics of Medical Imaging*, vol. 10573, p. 1057340, International Society for Optics and Photonics, 2018.

[34] X. Rui, L. Cheng, Y. Long, L. Fu, A. M. Alessio, E. Asma, P. E. Kinahan, and B. De Man, “Ultra-low dose ct attenuation correction for pet/ct: analysis of sparse view data acquisition and reconstruction algorithms,” *Physics in Medicine & Biology*, vol. 60, no. 19, p. 7437, 2015.

[35] Z. Yin, Y. Yao, A. Montillo, M. Wu, P. M. Edic, M. Kalra, and B. De Man, “Acquisition, preprocessing, and reconstruction of ultralow dose volumetric ct scout for organ-based ct scan planning,” *Medical physics*, vol. 42, no. 5, pp. 2730–2739, 2015.

[36] K. Mei, F. K. Kopp, R. Bippus, T. Köhler, B. J. Schwaiger, A. S. Gersing, A. Fehringer, A. Sauter, D. Münzel, F. Pfeiffer, *et al.*, “Is multidetector ct-based bone mineral density and quantitative bone microstructure assessment at the spine still feasible using ultra-low tube current and sparse sampling?,” *European radiology*, vol. 27, no. 12, pp. 5261–5271, 2017.

[37] J.-B. Thibault, K. D. Sauer, C. A. Bouman, and J. Hsieh, “A three-dimensional statistical approach to improved image quality for multislice helical ct,” *Medical physics*, vol. 34, no. 11, pp. 4526–4544, 2007.

[38] B. De Man and J. Fessler, “Statistical iterative reconstruction for x-ray computed tomography,” *Biomedical Mathematics: Promising Directions in Imaging, Therapy Planning and Inverse Problems*, pp. 113–40, 2010.

[39] J. S. Jorgensen, E. Y. Sidky, and X. Pan, “Quantifying admissible undersampling for sparsity-exploiting iterative image reconstruction in x-ray ct,” *IEEE Transactions on Medical Imaging*, vol. 32, no. 2, pp. 460–473, 2013.

[40] Z. F. Lu, E. L. Nickoloff, C. B. Ruzal-Shapiro, J. C. So, and A. K. Dutta, “New automated fluoroscopic systems for pediatric applications,” *Journal of applied clinical medical physics*, vol. 6, no. 4, pp. 88–105, 2005.

[41] Z. Yin, Y. Yao, A. Montillo, P. M. Edic, and B. De Man, “Feasibility study on ultra-low dose 3d scout of organ based ct scan planning,” in *The third international conference on image formation in X-ray computed tomography*, CT Meeting, 2014.

[42] T. Brosch and A. Saalbach, “Foveal fully convolutional nets for multi-organ segmentation,” in *Medical Imaging 2018: Image Processing*, vol. 10574, p. 105740U, International Society for Optics and Photonics, 2018.

[43] M. K. Kalra, M. M. Maher, T. L. Toth, B. Schmidt, B. L. Westerman, H. T. Morgan, and S. Saini, “Techniques and applications of automatic tube current modulation for ct,” *Radiology*, vol. 233, no. 3, pp. 649–657, 2004.

[44] F. K. Kopp, K. Mei, E. J. Rummeny, and P. B. Noël, “Sparse-sampling computed tomography for pulmonary imaging,” in *Medical Imaging 2019: Physics of Medical Imaging*, vol. 10948, p. 109482X, International Society for Optics and Photonics, 2019.

[45] K. Taguchi and J. S. Iwanczyk, “Vision 20/20: Single photon counting x-ray detectors in medical imaging,” *Medical physics*, vol. 40, no. 10, 2013.

[46] Marc Kachelrieß, *Neue Detektortechnologien: Photon Counting*, Deutscher Röntgenkongress 2017 (accessed May 07, 2019). `https://www.dkfz.de/en/roentgenbildgebung/ct/ct_conference_contributions/PhotonCountingCT.pdf?m=1515503812`.

[47] R. Steadman, C. Herrmann, and A. Livne, “Chromaix2: A large area, high count-rate energy-resolving photon counting asic for a spectral ct prototype,” *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 862, pp. 18–24, 2017.

[48] J. Tanguay, H. K. Kim, and I. A. Cunningham, “The role of x-ray swank factor in energy-resolving photon-counting imaging,” *Medical physics*, vol. 37, no. 12, pp. 6205–6211, 2010.

[49] E. Roessl and R. Proksa, “K-edge imaging in x-ray computed tomography using multi-bin photon counting detectors,” *Physics in Medicine & Biology*, vol. 52, no. 15, p. 4679, 2007.

[50] J. Schlomka, E. Roessl, R. Dorscheid, S. Dill, G. Martens, T. Istel, C. Bäumer, C. Herrmann, R. Steadman, G. Zeitler, *et al.*, “Experimental feasibility of multi-energy photon-counting k-edge imaging in pre-clinical computed tomography,” *Physics in Medicine & Biology*, vol. 53, no. 15, p. 4031, 2008.

[51] M. J. Willemink and P. B. Noël, “The evolution of image reconstruction for ct—from filtered back projection to artificial intelligence,” *European radiology*, vol. 29, no. 5, pp. 2185–2195, 2019.

[52] R. Gordon, R. Bender, and G. T. Herman, “Algebraic reconstruction techniques (art) for three-dimensional electron microscopy and x-ray photography,” *Journal of theoretical Biology*, vol. 29, no. 3, pp. 471–481, 1970.

[53] P. J. La Riviere, J. Bian, and P. A. Vargas, “Penalized-likelihood sinogram restoration for computed tomography,” *IEEE transactions on medical imaging*, vol. 25, no. 8, pp. 1022–1036, 2006.

[54] M. Kachelriess, O. Watzke, and W. A. Kalender, “Generalized multi-dimensional adaptive filtering for conventional and spiral single-slice, multi-slice, and cone-beam ct,” *Medical physics*, vol. 28, no. 4, pp. 475–490, 2001.

[55] A. H. Andersen and A. C. Kak, “Simultaneous algebraic reconstruction technique (sart): a superior implementation of the art algorithm,” *Ultrasonic imaging*, vol. 6, no. 1, pp. 81–94, 1984.

[56] H. Erdogan and J. A. Fessler, “Ordered subsets algorithms for transmission tomography,” *Physics in Medicine & Biology*, vol. 44, no. 11, p. 2835, 1999.

[57] K. Lange, R. Carson, *et al.*, “Em reconstruction algorithms for emission and transmission tomography,” *J. Comput. Assist. Tomogr*, vol. 8, no. 2, pp. 306–316, 1984.

[58] F. Xu and K. Mueller, “Real-time 3d computed tomographic reconstruction using commodity graphics hardware,” *Physics in Medicine & Biology*, vol. 52, no. 12, p. 3405, 2007.

[59] P. B. Noël, A. M. Walczak, J. Xu, J. J. Corso, K. R. Hoffmann, and S. Schafer, “Gpu-based cone beam computed tomography,” *Computer methods and programs in biomedicine*, vol. 98, no. 3, pp. 271–277, 2010.

[60] M. J. Willemink, P. A. de Jong, T. Leiner, L. M. de Heer, R. A. Nievelstein, R. P. Budde, and A. M. Schilham, “Iterative reconstruction techniques for computed tomography part 1: technical principles,” *European radiology*, vol. 23, no. 6, pp. 1623–1631, 2013.

[61] P. B. Noël, A. A. Fingerle, B. Renger, D. Münzel, E. J. Rummeny, and M. Dobritz, “Initial performance characterization of a clinical noise–suppressing reconstruction algorithm for mdct,” *American Journal of Roentgenology*, vol. 197, no. 6, pp. 1404–1409, 2011.

[62] P. B. Noël, B. Renger, M. Fiebich, D. Münzel, A. A. Fingerle, E. J. Rummeny, and M. Dobritz, “Does iterative reconstruction lower ct radiation dose: evaluation of 15,000 examinations,” *PloS one*, vol. 8, no. 11, p. e81141, 2013.

[63] A. Sauter, T. Koehler, A. A. Fingerle, B. Brendel, V. Richter, M. Rasper, E. J. Rummeny, P. B. Noël, and D. Münzel, “Ultra low dose ct pulmonary angiography with iterative reconstruction,” *PLoS One*, vol. 11, no. 9, p. e0162716, 2016.

[64] Philips Healthcare, *iDose4 iterative reconstruction technique - Breakthrough in image quality and dose reduction with the 4th generation of reconstruction*, Technical Whitepaper 2011 (accessed May 21, 2019). `http://incenter.medical.philips.com/doclib/enc/fetch/2000/4504/577242/577249/586938/587315/iDose4_-_Whitepaper_-_Technical_-Low_Res.pdf%3fnodeid%3d8432599%26vernum%3d-2`.

[65] K. M. Brown, S. Zabic, and T. Koehler, “Acceleration of ml iterative algorithms for ct by the use of fast start images,” in *Medical Imaging 2012: Physics of Medical Imaging*, vol. 8313, p. 831339, International Society for Optics and Photonics, 2012.

[66] P. B. Noël, T. Köhler, A. A. Fingerle, K. M. Brown, S. Zabic, D. Münzel, B. Haller, T. H. Baum, M. Henninger, R. Meier, *et al.*, “Evaluation of an iterative model–based reconstruction algorithm for low-tube-voltage (80 kvp) computed tomography angiography,” *Journal of Medical Imaging*, vol. 1, no. 3, p. 033501, 2014.

[67] J. A. Fessler, “Grouped coordinate descent algorithms for robust edge-preserving image restoration,” in *Image reconstruction and restoration II*, vol. 3170, pp. 184–195, International Society for Optics and Photonics, 1997.

[68] J. Fessler, M. Sonka, and J. M. Fitzpatrick, “Handbook of medical imaging,” *Medical Image Processing and Analysis*, vol. 2, pp. 1–70, 2000.

[69] D. Mehta, R. Thompson, T. Morton, A. Dhanantwari, and E. Shefer, “Iterative model reconstruction: simultaneously lowered computed tomography radiation dose and improved image quality,” *Med Phys Int J*, vol. 2, no. 1, pp. 147–55, 2013.

[70] G. J. Gang, K. Cheng, X. Guo, and J. W. Stayman, “Generalized prediction framework for reconstructed image properties using neural networks,” in *Medical Imaging 2019: Physics of Medical Imaging*, vol. 10948, p. 109480L, International Society for Optics and Photonics, 2019.

[71] J. D. Evans, D. G. Politte, B. R. Whiting, J. A. O’Sullivan, and J. F. Williamson, “Noise-resolution tradeoffs in x-ray ct imaging: A comparison of penalized alternating minimization and filtered backprojection algorithms,” *Medical physics*, vol. 38, no. 3, pp. 1444–1458, 2011.

[72] F. Bergner, B. Brendel, P. B. Noel, M. Dobritz, and T. Koehler, “Robust automated regularization factor selection for statistical reconstructions,” in *2nd International Conference on Image Formation in X-Ray Computed Tomography, Salt Lake City, USA, 24-27 June 2012*, 2012.

[73] “Report 79,” *Journal of the International Commission on Radiation Units and Measurements*, vol. 8, pp. 11–17, 04 2008.

[74] R. Symons, A. Pourmorteza, V. Sandfort, M. A. Ahlman, T. Cropper, M. Mallek, S. Kappler, S. Ulzheimer, M. Mahesh, E. C. Jones, *et al.*, “Feasibility of dose-reduced chest ct with photon-counting detectors: initial results in humans,” *Radiology*, vol. 285, no. 3, pp. 980–989, 2017.

[75] P. Ranganathan and R. Aggarwal, “Common pitfalls in statistical analysis: Understanding the properties of diagnostic tests–part 1,” *Perspectives in clinical research*, vol. 9, no. 1, p. 40, 2018.

[76] L. Yu, S. Leng, L. Chen, J. M. Kofler, R. E. Carter, and C. H. McCollough, “Prediction of human observer performance in a 2-alternative forced choice low-contrast detection task using channelized hotelling observer: Impact of radiation dose and reconstruction algorithms,” *Medical physics*, vol. 40, no. 4, p. 041908, 2013.

[77] A. Wunderlich and F. Noo, “Image covariance and lesion detectability in direct fan-beam x-ray computed tomography,” *Physics in Medicine & Biology*, vol. 53, no. 10, p. 2471, 2008.

[78] F. K. Kopp, M. Catalano, D. Pfeiffer, A. A. Fingerle, E. J. Rummeny, and P. B. Noël, “Cnn as model observer in a liver lesion detection task for x-ray computed tomography: A phantom study,” *Medical physics*, vol. 45, no. 10, pp. 4439–4447, 2018.

[79] J. G. Brankov, “Evaluation of the channelized hotelling observer with an internal-noise model in a train-test paradigm for cardiac spect defect detection,” *Physics in Medicine & Biology*, vol. 58, no. 20, p. 7159, 2013.

[80] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, “Rethinking the inception architecture for computer vision,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 2818–2826, 2016.

[81] J. G. Brankov, Y. Yang, L. Wei, I. El Naqa, and M. N. Wernick, “Learning a channelized observer for image quality assessment,” *IEEE Transactions on Medical Imaging*, vol. 28, no. 7, pp. 991–999, 2009.

[82] J. G. Brankov, L. Wei, Y. Yang, and M. N. Wernick, “Generalization evaluation of numerical observers for image quality assessment,” in *2006 IEEE Nuclear Science Symposium Conference Record*, vol. 3, pp. 1696–1698, IEEE, 2006.

[83] T. Marin, M. M. Kalayeh, P. H. Pretorius, M. N. Wernick, Y. Yang, and J. G. Brankov, “Numerical observer for cardiac motion assessment using machine learning,” in *Medical Imaging 2011: Image Perception, Observer Performance, and Technology Assessment*, vol. 7966, p. 79660G, International Society for Optics and Photonics, 2011.

[84] J. G. Brankov, I. El-Naqa, Y. Yang, and M. N. Wernick, “Learning a nonlinear channelized observer for image quality assessment,” in *2003 IEEE Nuclear Science Symposium. Conference Record (IEEE Cat. No. 03CH37515)*, vol. 4, pp. 2526–2529, IEEE, 2003.

[85] I. Lorente and J. G. Brankov, “Active learning for image quality assessment by model observer,” in *2014 IEEE 11th International Symposium on Biomedical Imaging (ISBI)*, pp. 1352–1355, IEEE, 2014.

[86] M. M. Kalayeh, T. Marin, and J. G. Brankov, “Generalization evaluation of machine learning numerical observers for image quality assessment,” *IEEE transactions on nuclear science*, vol. 60, no. 3, pp. 1609–1618, 2013.

[87] T. Marin, M. M. Kalayeh, F. M. Parages, and J. G. Brankov, “Numerical surrogates for human observers in myocardial motion evaluation from spect images,” *IEEE transactions on medical imaging*, vol. 33, no. 1, pp. 38–47, 2013.

[88] F. Massanes and J. G. Brankov, “Evaluation of cnn as anthropomorphic model observer,” in *Medical Imaging 2017: Image Perception, Observer Performance, and Technology Assessment*, vol. 10136, p. 101360Q, International Society for Optics and Photonics, 2017.

[89] D. P. Kingma and J. Ba, “Adam: A method for stochastic optimization,” *arXiv preprint arXiv:1412.6980*, 2014.

[90] D. Muenzel, D. Bar-Ness, E. Roessl, I. Blevis, M. Bartels, A. A. Fingerle, S. Ruschke, P. Coulon, H. Daerr, F. K. Kopp, *et al.*, “Spectral photon-counting ct: Initial experience with dual–contrast agent k-edge colonography,” *Radiology*, vol. 283, no. 3, pp. 723–728, 2016.

[91] I. Hernandez-Giron, J. M. den Harder, G. J. Streekstra, J. Geleijns, and W. J. Veldkamp, “Development of a 3d printed anthropomorphic lung phantom for image quality assessment in ct,” *Physica Medica*, vol. 57, pp. 47–57, 2019.

[92] A. J. Cloonan, D. Shahmirzadi, R. X. Li, B. J. Doyle, E. E. Konofagou, and T. M. McGloughlin, “3d-printed tissue-mimicking phantoms for medical imaging and computational validation applications,” *3D printing and additive manufacturing*, vol. 1, no. 1, pp. 14–23, 2014.

[93] The Cancer Imaging Archive (TCIA), *Phantom FDA*, 2019 (accessed June 11, 2019). `https://wiki.cancerimagingarchive.net/display/Public/Phantom+FDA`.

[94] The MathWorks, Inc, *kmeans - k-means clustering*, 2019 (accessed June 12, 2019). `https://www.mathworks.com/help/stats/kmeans.html`.

[95] H. Gong, L. Yu, S. Leng, S. K. Dilger, L. Ren, W. Zhou, J. G. Fletcher, and C. H. McCollough, “A deep learning and partial least square regression based model observer for a low-contrast lesion detection task in ct,” *Medical physics,* 2019.

[96] H. Gong, A. Walther, Q. Hu, C. W. Koo, E. A. Takahashi, D. L. Levin, T. F. Johnson, M. J. Hora, S. Leng, J. G. Fletcher, *et al.*, “Correlation between a deep-learning-based model observer and human observer for a realistic lung nodule localization task in chest ct,” in *Medical Imaging 2019: Image Perception, Observer Performance, and Technology Assessment*, vol. 10952, p. 109520K, International Society for Optics and Photonics, 2019.

[97] J. Donahue, Y. Jia, O. Vinyals, J. Hoffman, N. Zhang, E. Tzeng, and T. Darrell, “Decaf: A deep convolutional activation feature for generic visual recognition,” in *International conference on machine learning*, pp. 647–655, 2014.

[98] A. Dosovitskiy, J. T. Springenberg, M. Riedmiller, and T. Brox, “Discriminative unsupervised feature learning with convolutional neural networks,” in *Advances in neural information processing systems*, pp. 766–774, 2014.

[99] O. Ronneberger, P. Fischer, and T. Brox, “U-net: Convolutional networks for biomedical image segmentation,” in *International Conference on Medical image computing and computer-assisted intervention*, pp. 234–241, Springer, 2015.

[100] Y. Liu, H. Wang, Q. Li, M. J. McGettigan, Y. Balagurunathan, A. L. Garcia, Z. J. Thompson, J. J. Heine, Z. Ye, R. J. Gillies, *et al.*, “Radiologic features of small pulmonary nodules and lung cancer risk in the national lung screening trial: a nested case-control study,” *Radiology*, vol. 286, no. 1, pp. 298–306, 2017.

[101] D. A. Lynch, J. H. Austin, J. C. Hogg, P. A. Grenier, H.-U. Kauczor, A. A. Bankier, R. G. Barr, T. V. Colby, J. R. Galvin, P. A. Gevenois, *et al.*, “Ct-definable subtypes of chronic obstructive pulmonary disease: a statement of the fleischner society,” *Radiology*, vol. 277, no. 1, pp. 192–205, 2015.

[102] S. Milne and G. G. King, “Advanced imaging in copd: insights into pulmonary pathophysiology,” *Journal of thoracic disease*, vol. 6, no. 11, p. 1570, 2014.

[103] W. Zhou, J. Montoya, R. Gutjahr, A. Ferrero, A. Halaweish, S. Kappler, C. McCollough, and S. Leng, “Lung nodule volume quantification and shape differentiation with an ultra-high resolution technique on a photon counting detector ct system,” in *Proceedings of SPIE–the International Society for Optical Engineering*, vol. 10132, NIH Public Access, 2017.

[104] M. Yanagawa, A. Hata, O. Honda, N. Kikuchi, T. Miyata, A. Uranishi, S. Tsukagoshi, and N. Tomiyama, “Subjective and objective comparisons of image quality between ultra-high-resolution ct and conventional area detector ct in phantoms and cadaveric human lungs,” *European radiology*, vol. 28, no. 12, pp. 5060–5068, 2018.

[105] K. Mechlem, S. Ehn, T. Sellerer, E. Braig, D. Münzel, F. Pfeiffer, and P. B. Noël, “Joint statistical iterative material image reconstruction for spectral computed tomography using a semi-empirical forward model,” *IEEE transactions on medical imaging*, vol. 37, no. 1, pp. 68–80, 2017.

[106] B. De Man, J. Uribe, J. Baek, D. Harrison, Z. Yin, R. Longtin, J. Roy, B. Waters, C. Wilson, J. Short, *et al.*, “Multisource inverse-geometry ct. part i. system concept and development,” *Medical physics*, vol. 43, no. 8Part1, pp. 4607–4616, 2016.

[107] J. Baek, B. De Man, J. Uribe, R. Longtin, D. Harrison, J. Reynolds, B. Neculaes, K. Frutschy, L. Inzinna, A. Caiafa, *et al.*, “A multi-source inverse-geometry ct system: initial results with an 8 spot x-ray source array,” *Physics in Medicine & Biology*, vol. 59, no. 5, p. 1189, 2014.

[108] B. Chen, M. J. Muckley, A. Sodickson, T. O’Donnell, M. Berner, T. Allmendinger, K. Stierstorfer, T. Flohr, B. Schmidt, D. Sodickson, and R. Otazo, “First multislit collimator prototype for sparsect: Design, manufacturing and initial validation,” in *The fifth international conference on image formation in X-ray computed tomography*, CT Meeting, 2018.

[109] M. J. Muckley, B. Chen, T. O’Donnell, M. Berner, T. Allmendinger, K. Stierstorfer, T. Flohr, B. Schmidt, A. Sodickson, D. Sodickson, and R. Otazo, “Reconstruction of reduced-dose sparsect data acquired with an interrupted-beam prototype on a clinical scanner,” in *The fifth international conference on image formation in X-ray computed tomography*, CT Meeting, 2018.

[110] T. Koesters, F. Knoll, A. Sodickson, D. K. Sodickson, and R. Otazo, “Sparsect: interrupted-beam acquisition and sparse reconstruction for radiation dose reduction,” in *Medical Imaging 2017: Physics of Medical Imaging*, vol. 10132, p. 101320Q, International Society for Optics and Photonics, 2017.

[111] S. Abbas, T. Lee, S. Shin, R. Lee, and S. Cho, “Effects of sparse sampling schemes on image quality in low-dose ct,” *Medical physics*, vol. 40, no. 11, 2013.

[112] E. Y. Sidky, C.-M. Kao, and X. Pan, “Accurate image reconstruction from few-views and limited-angle data in divergent-beam ct,” *Journal of X-ray Science and Technology*, vol. 14, no. 2, pp. 119–139, 2006.

[113] E. Y. Sidky and X. Pan, “Image reconstruction in circular cone-beam computed tomography by constrained, total-variation minimization,” *Physics in Medicine & Biology*, vol. 53, no. 17, p. 4777, 2008.

[114] S. Žabić, Q. Wang, T. Morton, and K. M. Brown, “A low dose simulation tool for ct systems with energy integrating detectors,” *Medical physics*, vol. 40, no. 3, 2013.

[115] D. Muenzel, T. Koehler, K. Brown, S. Žabić, A. A. Fingerle, S. Waldt, E. Bendik, T. Zahel, A. Schneider, M. Dobritz, *et al.*, “Validation of a low dose simulation technique for computed tomography images,” *PloS one*, vol. 9, no. 9, p. e107843, 2014.

[116] J. H. Lee, B. R. Grant, J. H. Chung, I. Reiser, and M. Giger, “Assessment of diagnostic image quality of computed tomography (ct) images of the lung using deep learning,” in *Medical Imaging 2018: Physics of Medical Imaging*, vol. 10573, p. 105731M, International Society for Optics and Photonics, 2018.