Universiteit Leiden
The Netherlands

Spectral imaging and tomographic reconstruction methods for industrial applications

Zeegers, M.T.

Bibliography

[1] W. van Aarle, W. J. Palenstijn, J. Cant, E. Janssens, F. Bleichrodt, A. Dabravolski, J. De Beenhouwer, K. J. Batenburg, and J. Sijbers. "Fast and flexible X-ray tomography using the ASTRA toolbox". *Optics Express* 24.22 (2016), pp. 25129–25147 (cit. on pp. 47, 57, 63, 77, 93, 100, 107, 111, 131, 146).

[2] W. van Aarle, W. J. Palenstijn, J. De Beenhouwer, T. Altantzis, S. Bals, K. J. Batenburg, and J. Sijbers. "The ASTRA Toolbox: A platform for advanced algorithm development in electron tomography". *Ultramicroscopy* 157 (2015), pp. 35–47 (cit. on pp. 47, 57, 63, 77, 93, 107, 111, 131, 146).

[3] W. van Aarle, K. J. Batenburg, and J. Sijbers. "Automatic parameter estimation for the discrete algebraic reconstruction technique (DART)". *IEEE Transactions on Image Processing* 21.11 (2012), pp. 4608–4621 (cit. on p. 96).

[4] M. Abadi, P. Barham, J. Chen, Z. Chen, A. Davis, J. Dean, M. Devin, S. Ghemawat, G. Irving, M. Isard, et al. "TensorFlow: A system for large-scale machine learning". In: *12th USENIX Symposium on Operating Systems Design and Implementation (OSDI 16)*. (Savannah, GA, USA). 2016, pp. 265–283 (cit. on p. 35).

[5] J. F. P. J. Abascal, N. Ducros, V. Pronina, S. Rit, P. Rodesch, T. Broussaud, S. Bussod, P. C. Douek, A. Hauptmann, S. Arridge, et al. "Material decomposition in spectral CT using deep learning: A Sim2Real transfer approach". *IEEE Access* 9 (2021), pp. 25632–25647 (cit. on p. 119).

[6] J. F. P. J. Abascal, N. Ducros, and F. Peyrin. "Nonlinear material decomposition using a regularized iterative scheme based on the Bregman distance". *Inverse Problems* 34.12 (2018), p. 124003 (cit. on pp. 27, 154).

[7] A. Adler, M. Elad, and M. Zibulevsky. "Compressed learning: A deep neural network approach". *arXiv preprint arXiv:1610.09615* (2016) (cit. on p. 69).

[8] S. Akcay and T. Breckon. "Towards automatic threat detection: A survey of advances of deep learning within X-ray security imaging". *Pattern Recognition* 122 (2022), p. 108245 (cit. on pp. 35, 41, 152).

[9] S. Akcay, M. E. Kundegorski, C. G. Willcocks, and T. P. Breckon. "Using deep convolutional neural network architectures for object classification and detection within X-ray baggage security imagery". *IEEE Transactions on Information Forensics and Security* 13.9 (2018), pp. 2203–2215 (cit. on p. 68).

[10] C. Albon. *Machine learning with Python cookbook: Practical solutions from preprocessing to deep learning*. Sevastopol, CA, USA: O'Reilly Media, Inc., 2018 (cit. on p. 68).

[11] H. Alkadhi, A. Euler, D. Maintz, and D. Sahani. "Spectral imaging: Dual-energy, multi-energy and photon-counting CT" (2022) (cit. on p. 24).

[12] M. Z. Alom, T. M. Taha, C. Yakopcic, S. Westberg, P. Sidike, M. S. Nasrin, M. Hasan, B. C. Van Essen, A. A. Awwal, and V. K. Asari. "A state-of-the-art survey on deep learning theory and architectures". *Electronics* 8.3 (2019), p. 292 (cit. on p. 33).

[13] A. Alqahtani, X. Xie, and M. W. Jones. "Literature review of deep network compression". *Informatics* 8.4 (2021), p. 77 (cit. on p. 35).

[14] R. E. Alvarez and A. Macovski. "Energy-selective reconstructions in X-ray computerised tomography". *Physics in Medicine & Biology* 21.5 (1976), p. 733 (cit. on p. 26).

[15] A. H. Andersen and A. C. Kak. "Simultaneous algebraic reconstruction technique (SART): A superior implementation of the ART algorithm". *Ultrasonic imaging* 6.1 (1984), pp. 81–94 (cit. on p. 17).

[16] V. Andriiashen, R. van Liere, T. van Leeuwen, and K. J. Batenburg. "Unsupervised foreign object detection based on dual-energy absorptiometry in the food industry". *Journal of Imaging* 7.7 (2021), p. 10 (cit. on pp. 22, 40).

[17] V. Andriiashen, R. van Liere, T. van Leeuwen, and K. J. Batenburg. "CT-based data generation for foreign object detection on a single X-ray projection". *Scientific Reports* 13.1 (2023), p. 1881 (cit. on p. 153).

[18] F. van Assche, S. Vanheule, L. van Hoorebeke, and M. N. Boone. "The spectral X-ray imaging data acquisition (SpeXIDAQ) framework". *Sensors* 21.2 (2021), p. 563 (cit. on p. 21).

[19] ASTM. "G173–03: Standard tables for reference solar spectral irradiances: Direct normal and hemispherical on 37° tilted surface". *ASTM International, West Conshohocken, PA* (2003) (cit. on p. 80).

[20] H. Attouch, J. Bolte, P. Redont, and A. Soubeyran. "Proximal alternating minimization and projection methods for nonconvex problems: An approach based on the Kurdyka-Lojasiewicz inequality". *Mathematics of Operations Research* 35.2 (2010), pp. 438–457 (cit. on p. 125).

[21] N. Audebert, B. Le Saux, and S. Lefèvre. "Deep learning for classification of hyperspectral data: A comparative review". *IEEE Geoscience and Remote Sensing Magazine* 7.2 (2019), pp. 159–173 (cit. on p. 68).

[22] V. Badrinarayanan, A. Kendall, and R. Cipolla. "Segnet: A deep convolutional encoder-decoder architecture for image segmentation". *IEEE Transactions on Pattern Analysis and Machine Intelligence* 39.12 (2017), pp. 2481–2495 (cit. on p. 43).

[23] E. J. Bakker, P. B. W. Schwering, and S. P. van den Broek. "From hyperspectral imaging to dedicated sensors". In: *Targets and Backgrounds VI: Characterization, Visualization, and the Detection Process*. (Orlando, FL, USA). Vol. 4029. International Society for Optics and Photonics. 2000, pp. 312–324 (cit. on p. 66).

[24] R. Ballabriga, J. Alozy, F. Bandi, M. Campbell, N. Egidos, J. M. Fernandez-Tenllado, E. H. M. Heijne, I. Kremastiotis, X. Llopart, B. J. Madsen, et al. "Photon counting detectors for X-ray imaging with emphasis on CT". *IEEE Transactions on Radiation and Plasma Medical Sciences* 5.4 (2020) (cit. on p. 114).

[25] R. Ballabriga, J. Alozy, M. Campbell, E. Frojdh, E. Heijne, T. Koenig, X. Llopart, J. Marchal, D. Pennicard, T. Poikela, L. Tlustos, P. Valerio, W. Wong, and M. Zuber. "Review of hybrid pixel detector readout ASICs for spectroscopic X-ray imaging". *Journal of Instrumentation* 11.1 (2016), P01007 (cit. on p. 114).

[26] R. Ballabriga, M. Campbell, E. Heijne, X. Llopart, L. Tlustos, and W. Wong. "Medipix3: A 64 k pixel detector readout chip working in single photon counting mode with improved spectrometric performance". *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 633 (2011), S15–S18 (cit. on p. 21).

[27] R. F. Barber, E. Y. Sidky, T. G. Schmidt, and X. Pan. "An algorithm for constrained one-step inversion of spectral CT data". *Physics in Medicine & Biology* 61.10 (2016), p. 3784 (cit. on pp. 26, 28, 119, 132).

[28] K. J. Batenburg, W. Fortes, L. Hajdu, and R. Tijdeman. "Bounds on the quality of reconstructed images in binary tomography". *Discrete Applied Mathematics* 161.15 (2013), pp. 2236–2251 (cit. on p. 102).

[29] K. J. Batenburg, W. J. Palenstijn, P. Balázs, and J. Sijbers. "Dynamic angle selection in binary tomography". *Computer Vision and Image Understanding* 117.4 (2013), pp. 306–318 (cit. on p. 18).

[30] K. J. Batenburg and J. Sijbers. "DART: A practical reconstruction algorithm for discrete tomography". *IEEE Transactions on Image Processing* 20.9 (2011), pp. 2542–2553 (cit. on pp. 18, 61, 96, 100).

[31] K. J. Batenburg and J. Sijbers. "DART: A fast heuristic algebraic reconstruction algorithm for discrete tomography". In: *IEEE International Conference on Image Processing, 2007, ICIP 2007*. (San Antonio, TX, USA). Vol. 4. IEEE. 2007, pp. IV–133 (cit. on p. 96).

[32] D. F. Bauer, C. Ulrich, T. Russ, A.-K. Golla, L. R. Schad, and F. G. Zöllner. "End-to-end deep learning CT image reconstruction for metal artifact reduction". *Applied Sciences* 12.1 (2021), p. 404 (cit. on p. 35).

[33] I. Bayram and M. E. Kamasak. "A directional total variation". In: *2012 Proceedings of the 20th European Signal Processing Conference (EUSIPCO)*. (Bucharest, Romania). IEEE. 2012, pp. 265–269 (cit. on p. 28).

[34] A. Beck. *First-order methods in optimization*. Philadelphia, PA, USA: SIAM, 2017 (cit. on p. 128).

[35] M. Beister, D. Kolditz, and W. A. Kalender. "Iterative reconstruction methods in X-ray CT". *Physica Medica* 28.2 (2012), pp. 94–108 (cit. on p. 14).

[36] R. Bellazzini, A. Brez, G. Spandre, M. Minuti, M. Pinchera, P. Delogu, P. L. De Ruvo, and A. Vincenzi. "PIXIE III: A very large area photon-counting CMOS pixel ASIC for sharp X-ray spectral imaging". *Journal of Instrumentation* 10.01 (2015), p. C01032 (cit. on p. 21).

[37] E. van den Berg and M. P. Friedlander. "Spot - A linear-operator toolbox" (2014) (cit. on p. 146).

[38] J. Berner, P. Grohs, G. Kutyniok, and P. Petersen. "The modern mathematics of deep learning". *arXiv preprint arXiv:2105.04026* (2021) (cit. on pp. 29, 30, 34, 35).

[39] M. W. Berry, M. Browne, A. N. L., V. P. Pauca, and R. J. Plemmons. "Algorithms and applications for approximate nonnegative matrix factorization". *Computational Statistics & Data Analysis* 52.1 (2007), pp. 155–173 (cit. on p. 132).

[40] R. Bhayana, A. Parakh, and A. Kambadakone. "Material decomposition with dual-and multi-energy computed tomography". *MRS Communications* 10.4 (2020), pp. 558–565 (cit. on pp. 24, 114).

[41] F. Bleichrodt, T. van Leeuwen, W. J. Palenstijn, W. van Aarle, J. Sijbers, and K. J. Batenburg. "Easy implementation of advanced tomography algorithms using the ASTRA toolbox with Spot operators". *Numerical Algorithms* 71.3 (2016), pp. 673–697 (cit. on p. 131).

[42] F. Bleichrodt, F. Tabak, and K. J. Batenburg. "SDART: An algorithm for discrete tomography from noisy projections". *Computer Vision and Image Understanding* 129 (2014), pp. 63–74 (cit. on p. 96).

[43] J. Bolte, S. Sabach, and M. Teboulle. "Proximal alternating linearized minimization for nonconvex and nonsmooth problems". *Mathematical Programming* 146.1 (2014), pp. 459–494 (cit. on p. 125).

[44] F. G. Bossema, M. Domínguez-Delmás, W. J. Palenstijn, A. Kostenko, J. Dorscheid, S. B. Coban, E. Hermens, and K. J. Batenburg. "A novel method for dendrochronology of large historical wooden objects using line trajectory X-ray tomography". *Scientific Reports* 11.1 (2021), pp. 1–12 (cit. on p. 13).

[45] S. P. Boyd and L. Vandenberghe. *Convex optimization*. Cambridge, UK: Cambridge University Press, 2004 (cit. on pp. 191, 192).

[46] S. Boyd, N. Parikh, and E. Chu. *Distributed optimization and statistical learning via the alternating direction method of multipliers*. Hanover, MA, USA: Now Publishers Inc., 2011 (cit. on p. 195).

[47] R. Bujila, A. Omar, and G. Poludniowski. "A validation of SpekPy: A software toolkit for modelling X-ray tube spectra". *Physica Medica* 75 (2020), pp. 44–54 (cit. on p. 130).

[48] L. Bungert, D. A. Coomes, M. J. Ehrhardt, J. Rasch, R. Reisenhofer, and C.-B. Schönlieb. "Blind image fusion for hyperspectral imaging with the directional total variation". *Inverse Problems* 34.4 (2018), p. 044003 (cit. on p. 155).

[49] M. Busi, C. Kehl, J. R. Frisvad, and U. L. Olsen. "Metal artifact reduction in spectral X-ray CT using spectral deep learning". *Journal of Imaging* 8.3 (2022), p. 77 (cit. on p. 35).

[50] J.-W. Buurlage. "Real-time tomographic reconstruction". PhD thesis. Leiden University, 2020 (cit. on p. 18).

[51] T. M. Buzug. *Computed tomography: From photon statistics to modern cone-beam CT*. Berlin/Heidelberg, Germany: (1st ed.). Springer, 2008. Chap. 8, pp. 311–342 (cit. on pp. 7, 24, 61, 96).

[52] C. Cai, T. Rodet, S. Legoupil, and A. Mohammad-Djafari. "A full-spectral Bayesian reconstruction approach based on the material decomposition model applied in dual-energy computed tomography". *Medical Physics* 40.11 (2013), p. 111916 (cit. on p. 132).

[53] E. Çallı, E. Sogancioglu, B. van Ginneken, K. G. van Leeuwen, and K. Murphy. "Deep learning for chest X-ray analysis: A survey". *Medical Image Analysis* 72 (2021), p. 102125 (cit. on p. 35).

[54] J. Cammin, J. S. Iwanczyk, and K. Taguchi. "Spectral/Photon-counting computed tomography". In: *Emerging Imaging Technologies in Medicine*. Boca Raton, FL, USA: CRC Press, 2012, pp. 42–61 (cit. on p. 23).

[55] S. Carmignato, W. Dewulf, and R. Leach. *Industrial X-ray computed tomography*. Cham, Switzerland: Springer, 2018 (cit. on pp. 5, 7, 13, 18, 24).

[56] C.-I. Chang. *Hyperspectral data processing: Algorithm design and analysis*. Hoboken, NJ, USA: John Wiley & Sons, 2013 (cit. on p. 66).

[57] G. Chartrand, P. M. Cheng, E. Vorontsov, M. Drozdzal, S. Turcotte, C. J. Pal, S. Kadoury, and A. Tang. "Deep learning: A primer for radiologists". *Radiographics* 37.7 (2017), pp. 2113–2131 (cit. on p. 40).

[58] L. Chen, Z. Wei, and Y. Xu. "A lightweight spectral-spatial feature extraction and fusion network for hyperspectral image classification". *Remote Sensing* 12.9 (2020), p. 1395 (cit. on p. 69).

[59] Y. Chen, H. Jiang, C. Li, X. Jia, and P. Ghamisi. "Deep feature extraction and classification of hyperspectral images based on convolutional neural networks". *IEEE Transactions on Geoscience and Remote Sensing* 54.10 (2016), pp. 6232–6251 (cit. on p. 68).

[60] D. P. Clark and C. T. Badea. "Spectral diffusion: An algorithm for robust material decomposition of spectral CT data". *Physics in Medicine & Biology* 59.21 (2014), p. 6445 (cit. on pp. 23, 27, 118).

[61] D. P. Clark and C. T. Badea. "Hybrid spectral CT reconstruction". *PloS One* 12.7 (2017), e0180324 (cit. on p. 27).

[62] S. B. Coban, F. Lucka, W. J. Palenstijn, D. van Loo, and K. J. Batenburg. "Explorative imaging and its implementation at the FleX-ray laboratory". *Journal of Imaging* 6.4 (2020), p. 18 (cit. on pp. 45, 47).

[63] W. Cong, D. Harrison, Y. Xi, and G. Wang. "Projection decomposition for dual-energy computed tomography". *arXiv preprint arXiv:1805.05312* (2018) (cit. on p. 26).

[64] E. Cueva, A. Meaney, S. Siltanen, and M. J. Ehrhardt. "Synergistic multi-spectral CT reconstruction with directional total variation". *Philosophical Transactions of the Royal Society A* 379.2204 (2021), p. 20200198 (cit. on pp. 26–28).

[65] A. Dabravolski, K. J. Batenburg, and J. Sijbers. "A multiresolution approach to discrete tomography using DART". *PloS One* 9.9 (2014), e106090 (cit. on p. 96).

[66] Q. Dai, J.-H. Cheng, D.-W. Sun, and X.-A. Zeng. "Advances in feature selection methods for hyperspectral image processing in food industry applications: A review". *Critical Reviews in Food Science and Nutrition* 55.10 (2015), pp. 1368–1382 (cit. on p. 68).

[67] M. Danielsson, M. Persson, and M. Sjölin. "Photon-counting X-ray detectors for CT". *Physics in Medicine & Biology* 66.3 (2021), 03TR01 (cit. on p. 23).

[68] G. Demaurex and L. Sallé. "Detection of physical hazards". In: *Food Safety Management*. Elsevier, 2014, pp. 511–533 (cit. on pp. 3, 4, 152).

[69] A. M. Deshpande, A. A. Minai, and M. Kumar. "One-shot recognition of manufacturing defects in steel surfaces". *Procedia Manufacturing* 48 (2020), pp. 1064–1071 (cit. on p. 40).

[70] M. Diwakar and M. Kumar. "A review on CT image noise and its denoising". *Biomedical Signal Processing and Control* 42 (2018), pp. 73–88 (cit. on pp. 45, 61).

[71] N. Djurabekova, A. Goldberg, A. Hauptmann, D. Hawkes, G. Long, F. Lucka, and M. Betcke. "Application of Proximal Alternating Linearized Minimization (PALM) and inertial PALM to dynamic 3D CT". In: *15th International Meeting on Fully Three-Dimensional Image Reconstruction in Radiology and Nuclear Medicine*. (Philadelphia, PA, USA). Ed. by S. Matej and S. D. Metzler. Vol. 11072. International Society for Optics and Photonics. SPIE, 2019, pp. 30–34 (cit. on p. 61).

[72] W. Du, H. Shen, J. Fu, G. Zhang, and Q. He. "Approaches for improvement of the X-ray image defect detection of automobile casting aluminum parts based on deep learning". *NDT & E International* 107 (2019), p. 102144 (cit. on p. 35).

[73] Y. Dua, V. Kumar, and R. S. Singh. "Comprehensive review of hyperspectral image compression algorithms". *Optical Engineering* 59.9 (2020), p. 090902 (cit. on p. 35).

[74] J. Duchi, E. Hazan, and Y. Singer. "Adaptive subgradient methods for online learning and stochastic optimization." *Journal of Machine Learning Research* 12.7 (2011) (cit. on p. 33).

[75] N. Ducros, J. F. P. J. Abascal, B. Sixou, S. Rit, and F. Peyrin. "Regularization of nonlinear decomposition of spectral X-ray projection images". *Medical Physics* 44.9 (2017), e174–e187 (cit. on p. 119).

[76] M. C. Edwards, M. F. Stringer, et al. "Observations on patterns in foreign material investigations". *Food Control* 18.7 (2007), pp. 773–782 (cit. on p. 3).

[77] C. K. Egan, S. D. M. Jacques, T. Connolley, M. D. Wilson, M. C. Veale, P. Seller, and R. J. Cernik. "Dark-field hyperspectral X-ray imaging". *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* 470.2165 (2014), p. 20130629 (cit. on p. 118).

[78] C. K. Egan, S. D. M. Jacques, M. D. Wilson, M. C. Veale, P. Seller, A. M. Beale, R. A. D. Pattrick, P. J. Withers, and R. J. Cernik. "3D chemical imaging in the laboratory by hyperspectral X-ray computed tomography". *Scientific Reports* 5.1 (2015), pp. 1–9 (cit. on pp. 21, 114).

[79] L. Eger, S. Do, P. Ishwar, W. C. Karl, and H. Pien. "A learning-based approach to explosives detection using multi-energy X-ray computed tomography". In: *2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. (Prague, Czech Republic). IEEE. 2011, pp. 2004–2007 (cit. on p. 66).

[80] H. Einarsdóttir, M. J. Emerson, L. H. Clemmensen, K. Scherer, K. Willer, M. Bech, R. Larsen, B. K. Ersbøll, and F. Pfeiffer. "Novelty detection of foreign objects in food using multi-modal X-ray imaging". *Food Control* 67 (2016), pp. 39–47 (cit. on pp. 40, 155).

[81] H. Einarsdóttir, B. Gudmundsson, and V. Ómarsson. "Automation in the fish industry". *Animal Frontiers* 12.2 (2022), pp. 32–39 (cit. on p. 35).

[82] G. Einarsson, J. N. Jensen, R. R. Paulsen, H. Einarsdottir, B. K. Ersbøll, A. B. Dahl, and L. B. Christensen. "Foreign object detection in multispectral X-ray images of food items using sparse discriminant analysis". In: *Scandinavian Conference on Image Analysis*. (Tromsø, Norway). Ed. by P. Sharma and F. M. Bianchi. Springer, 2017, pp. 350–361 (cit. on p. 61).

[83] G. Elmasry, M. Kamruzzaman, D.-W. Sun, and P. Allen. "Principles and applications of hyperspectral imaging in quality evaluation of agro-food products: A review". *Critical Reviews in Food Science and Nutrition* 52.11 (2012), pp. 999–1023 (cit. on pp. 66, 67).

[84] C. L. Epstein. *Introduction to the mathematics of medical imaging*. Philadelphia, PA, USA: SIAM, 2007 (cit. on p. 14).

[85] A. Falkovskaya and A. Gowen. "Literature review: Spectral imaging applied to poultry products". *Poultry Science* 99.7 (2020), pp. 3709–3722 (cit. on p. 4).

[86] M. Fauvel, J. Chanussot, and J. A. Benediktsson. "Kernel principal component analysis for the classification of hyperspectral remote sensing data over urban areas". *EURASIP Journal on Advances in Signal Processing* 783194.1 (2009), pp. 1–14 (cit. on p. 68).

[87] L. A. Feldkamp, L. C. Davis, and J. W. Kress. "Practical cone-beam algorithm". *Journal of the Optical Society of America A* 1.6 (1984), pp. 612–619 (cit. on pp. 15, 61).

[88] F. Feng, S. Wang, C. Wang, and J. Zhang. "Learning deep hierarchical spatial-spectral features for hyperspectral image classification based on residual 3D-2D CNN". *Sensors* 19.23 (2019), p. 5276 (cit. on p. 69).

[89] Y.-Z. Feng and D.-W. Sun. "Application of hyperspectral imaging in food safety inspection and control: A review". *Critical Reviews in Food Science and Nutrition* 52.11 (2012), pp. 1039–1058 (cit. on pp. 3, 4).

[90] M. K. Ferguson, A. Ronay, Y. T. T. Lee, and K. H. Law. "Detection and segmentation of manufacturing defects with convolutional neural networks and transfer learning". *Smart and Sustainable Manufacturing Systems* 2.1 (2018) (cit. on p. 35).

[91] A. Förster, S. Brandstetter, and C. Schulze-Briese. "Transforming X-ray detection with hybrid photon counting detectors". *Philosophical Transactions of the Royal Society A* 377.2147 (2019), p. 20180241 (cit. on p. 21).

[92] K. Fotiadou, G. Tsagkatakis, and P. Tsakalides. "Deep convolutional neural networks for the classification of snapshot mosaic hyperspectral imagery". *Electronic Imaging* 2017.17 (2017), pp. 185–190 (cit. on p. 66).

[93] J. Frank. *Electron tomography*. New York, NY, USA: Springer, 1992 (cit. on p. 96).

[94] E. Fredenberg. "Spectral and dual-energy X-ray imaging for medical applications". *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 878 (2018), pp. 74–87 (cit. on pp. 19, 21).

[95] N. R. Fredette, A. Kavuri, and M. Das. "Multi-step material decomposition for spectral computed tomography". *Physics in Medicine & Biology* 64.14 (2019), p. 145001 (cit. on pp. 118, 119).

[96] C. Fröjdh, B. Norlin, and E. Fröjdh. "Spectral X-ray imaging with single photon processing detectors". *Journal of Instrumentation* 8.02 (2013), p. C02010 (cit. on p. 22).

[97] F. Fusseis, X. Xiao, C. Schrank, and F. De Carlo. "A brief guide to synchrotron radiation-based microtomography in (structural) geology and rock mechanics". *Journal of Structural Geology* 65 (2014), pp. 1–16 (cit. on p. 13).

[98] A. Garcia-Garcia, S. Orts-Escolano, S. Oprea, V. Villena-Martinez, and J. Garcia-Rodriguez. "A review on deep learning techniques applied to semantic segmentation". *arXiv preprint arXiv:1704.06857* (2017) (cit. on pp. 40, 41).

[99] R. Garnett. "A comprehensive review of dual-energy and multi-spectral computed tomography". *Clinical Imaging* 67 (2020), pp. 160–169 (cit. on pp. 21, 23–26, 28).

[100] *GitHub - SimonRit/OneStepSpectralCT: Matlab code for spectral CT one-step inversion. Implementation of five different methods*. `https://github.com/SimonRit/OneStepSpectralCT`. Accessed on 19 November 2021 (cit. on pp. 130, 131).

[101] L. Gjesteby, B. De Man, Y. Jin, H. Paganetti, J. Verburg, D. Giantsoudi, and G. Wang. "Metal artifact reduction in CT: Where are we after four decades?" *IEEE Access* 4 (2016), pp. 5826–5849 (cit. on p. 59).

[102] X. Glorot, A. Bordes, and Y. Bengio. "Deep sparse rectifier neural networks". In: *Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics*. (Fort Lauderdale, FL, USA). JMLR Workshop and Conference Proceedings. 2011, pp. 315–323 (cit. on p. 31).

[103] J. Gorski, F. Pfeuffer, and K. Klamroth. "Biconvex sets and optimization with biconvex functions: A survey and extensions". *Mathematical Methods of Operations Research* 66.3 (2007), pp. 373–407 (cit. on pp. 122, 193).

[104] M. Graña, M. A. Veganzons, and B. Ayerdi. *Hyperspectral remote sensing scenes*. `http://www.ehu.eus/ccwintco/index.php/Hyperspectral_Remote_Sensing_Scenes`. Accessed on 15 September 2020 (cit. on p. 77).

[105] M. Grandini, E. Bagli, and G. Visani. "Metrics for multi-class classification: An overview". *arXiv preprint arXiv:2008.05756* (2020) (cit. on p. 50).

[106] M. Graves, A. Smith, and B. Batchelor. "Approaches to foreign body detection in foods". *Trends in Food Science & Technology* 9.1 (1998), pp. 21–27 (cit. on p. 4).

[107] M. K. Griffin and H. K. Burke. "Compensation of hyperspectral data for atmospheric effects". *Lincoln Laboratory Journal* 14.1 (2003), pp. 29–54 (cit. on p. 81).

[108] Y. Guo, Y. Liu, T. Georgiou, and M. S. Lew. "A review of semantic segmentation using deep neural networks". *International Journal of Multimedia Information Retrieval* 7.2 (2018), pp. 87–93 (cit. on p. 40).

[109] R. Habel, M. Kudenov, and M. Wimmer. "Practical spectral photography". In: *Computer Graphics Forum*. Vol. 31. 2pt2. Wiley Online Library. 2012, pp. 449–458 (cit. on p. 155).

[110] M. Habermann, V. Frémont, and E. H. Shiguemori. "Feature selection for hyperspectral images using single-layer neural networks". In: *8th International Conference of Pattern Recognition Systems (ICPRS 2017)*. (Madrid, Spain). 2017, pp. 1–6 (cit. on p. 66).

[111] R. P. Haff and N. Toyofuku. "X-ray detection of defects and contaminants in the food industry". *Sensing and Instrumentation for Food Quality and Safety* 2.4 (2008), pp. 262–273 (cit. on p. 40).

[112] P. C. Hansen, J. S. Jørgensen, and W. R. B. Lionheart. *Computed tomography: Algorithms, insight, and just enough theory*. Philadelphia, PA, USA: (1st ed.). SIAM, 2021 (cit. on pp. 14, 61).

[113] P. C. Hansen and M. Saxild-Hansen. "AIR tools - A MATLAB package of algebraic iterative reconstruction methods". *Journal of Computational and Applied Mathematics* 236.8 (2012), pp. 2167–2178 (cit. on p. 131).

[114] A. Hauptmann, O. Öktem, and C. Schönlieb. "Image reconstruction in dynamic inverse problems with temporal models". *arXiv preprint arXiv:2007.10238* (2020) (cit. on p. 61).

[115] K. He, G. Gkioxari, P. Dollár, and R. Girshick. "Mask R-CNN". In: *Proceedings of the IEEE International Conference on Computer Vision*. (Venice, Italy). IEEE, 2017, pp. 2961–2969 (cit. on p. 43).

[116] Y. He, Q. Xiao, X. Bai, L. Zhou, F. Liu, and C. Zhang. "Recent progress of nondestructive techniques for fruits damage inspection: A review". *Critical Reviews in Food Science and Nutrition* (2021), pp. 1–19 (cit. on p. 40).

[117] Y. He, X. Bai, Q. Xiao, F. Liu, L. Zhou, and C. Zhang. "Detection of adulteration in food based on nondestructive analysis techniques: A review". *Critical Reviews in Food Science and Nutrition* 61.14 (2021), pp. 2351–2371 (cit. on p. 4).

[118] B. J. Heismann, B. T. Schmidt, and T. Flohr. *Spectral computed tomography*. Bellingham, WA, USA: SPIE Press, 2012 (cit. on pp. 21, 26, 119).

[119] A. A. Hendriksen, D. M. Pelt, and K. J. Batenburg. "Noise2inverse: Self-supervised deep convolutional denoising for tomography". *IEEE Transactions on Computational Imaging* 6 (2020), pp. 1320–1335 (cit. on pp. 45, 152).

[120] A. A. Hendriksen. "Deep learning for tomographic reconstruction with limited data". PhD thesis. Leiden University, 2022 (cit. on p. 35).

[121] G. T. Herman and A. Kuba. *Discrete tomography: Foundations, algorithms, and applications*. New York, NY, USA: (1st ed.). Springer, 1999 (cit. on pp. 18, 61).

[122] G. T. Herman and A. Kuba. *Advances in discrete tomography and its applications*. New York, NY, USA: Springer Science & Business Media, 2007 (cit. on p. 18).

[123] M. R. Hestenes and E. Stiefel. "Methods of conjugate gradients for solving linear systems". *Journal of Research of the National Bureau of Standards* 49 (1952), pp. 409–436 (cit. on pp. 61, 132).

[124] T. Higaki, Y. Nakamura, F. Tatsugami, T. Nakaura, and K. Awai. "Improvement of image quality at CT and MRI using deep learning". *Japanese Journal of Radiology* 37.1 (2019), pp. 73–80 (cit. on p. 35).

[125] T. Hohweiller, N. Ducros, F. Peyrin, and B. Sixou. "Spectral CT material decomposition in the presence of poisson noise: A Kullback–Leibler approach". *IRBM* 38.4 (2017), pp. 214–218 (cit. on pp. 118, 119).

[126] K. M. Holt. "Total nuclear variation and jacobian extensions of total variation for vector fields". *IEEE Transactions on Image Processing* 23.9 (2014), pp. 3975–3989 (cit. on p. 28).

[127] J. Hsieh. *Computed tomography: Principles, design, artifacts, and recent advances*. Bellingham, WA, USA, 2009 (cit. on p. 96).

[128] W. Hu, Y. Huang, L. Wei, F. Zhang, and H. Li. "Deep convolutional neural networks for hyperspectral image classification". *Journal of Sensors* 2015 (2015), p. 258619 (cit. on p. 66).

[129] H. Huang, M. Chen, and Y. Duan. "Dimensionality reduction of hyperspectral image using spatial-spectral regularized sparse hypergraph embedding". *Remote Sensing* 11.9 (2019), p. 1039 (cit. on p. 67).

[130] J. H. Hubbell and S. M. Seltzer. *Tables of X-ray mass attenuation coefficients and mass energy-absorption coefficients 1 KeV to 20 MeV for elements Z=1 to 92 and 48 additional substances of dosimetric interest*. Tech. rep. National Institute of Standards and Technology-PL, Gaithersburg, MD, USA. Ionizing Radiation Div., 1995 (cit. on pp. 57, 78, 123, 130, 184, 197).

[131] R. Huber, G. Haberfehlner, M. Holler, G. Kothleitner, and K. Bredies. "Total generalized variation regularization for multi-modal electron tomography". *Nanoscale* 11.12 (2019), pp. 5617–5632 (cit. on p. 28).

[132] M. Imani and H. Ghassemian. "An overview on spectral and spatial information fusion for hyperspectral image classification: Current trends and challenges". *Information Fusion* 59 (2020), pp. 59–83 (cit. on p. 68).

[133] N. Jaccard, T. W. Rogers, E. J. Morton, and L. D. Griffin. "Detection of concealed cars in complex cargo X-ray imagery using deep learning". *Journal of X-ray Science and Technology* 25.3 (2017), pp. 323–339 (cit. on p. 68).

[134] S. D. M. Jacques, C. K. Egan, M. D. Wilson, M. C. Veale, P. Seller, and R. J. Cernik. "A laboratory system for element specific hyperspectral X-ray imaging". *Analyst* 138.3 (2013), pp. 755–759 (cit. on p. 114).

[135] S. Jadon. "A survey of loss functions for semantic segmentation". *arXiv preprint arXiv:2006.14822* (2020) (cit. on pp. 48, 77).

[136] F. Jolivet, J. Lesaint, C. Fournier, M. Garcin, and A. Brambilla. "An efficient one-step method for spectral CT based on an approximate linear model". *IEEE Transactions on Radiation and Plasma Medical Sciences* 5.4 (2020) (cit. on p. 119).

[137] A. A. Joy, M. A. M. Hasan, and M. A. Hossain. "A comparison of supervised and unsupervised dimension reduction methods for hyperspectral image classification". In: *2019 International Conference on Electrical, Computer and Communication Engineering (ECCE)*. (Cox's Bazar, Bangladesh). IEEE. 2019, pp. 1–6 (cit. on pp. 67, 68).

[138] A. Kadu and M. T. Zeegers. *mzeegers/ADJUST: ADJUST*. Zenodo, `https://doi.org/10.5281/zenodo.7821066`. Version v1.0.2. 2023 (cit. on p. 146).

[139] J. Kaipio and E. Somersalo. *Statistical and computational inverse problems*. Vol. 160. New York, NY, USA: Springer-Verlag, 2004 (cit. on p. 129).

[140] A. C. Kak and M. Slaney. *Principles of computerized tomographic imaging*. Philadelphia, PA, USA: SIAM, 2001 (cit. on pp. 47, 98, 101, 141).

[141] K. V. Kale, M. M. Solankar, D. B. Nalawade, R. K. Dhumal, and H. R. Gite. "A research review on hyperspectral data processing and analysis algorithms". *Proceedings of the National Academy of Sciences, India Section A: Physical Sciences* 87.4 (2017), pp. 541–555 (cit. on p. 67).

[142] M. Karcaaltincaba and A. Aktas. "Dual-energy CT revisited with multidetector CT: Review of principles and clinical applications". *Diagnostic and Interventional Radiology* 17.3 (2017) (cit. on p. 114).

[143] D. Kazantsev, J. S. Jørgensen, M. S. Andersen, W. R. B. Lionheart, P. D. Lee, and P. J. Withers. "Joint image reconstruction method with correlative multi-channel prior for X-ray spectral computed tomography". *Inverse Problems* 34.6 (2018), p. 064001 (cit. on pp. 26, 28, 61, 110, 119).

[144] D. Kern and A. Mastmeyer. "3D bounding box detection in volumetric medical image data: A systematic literature review". In: *2021 IEEE 8th International Conference on Industrial Engineering and Applications (ICIEA)*. (Chengdu, China). IEEE, 2021, pp. 509–516 (cit. on pp. 45, 61).

[145] K. Kim, H. Kim, J. Chun, M. Kang, M. Hong, and B. Min. "Real-time anomaly detection in packaged food X-ray images using supervised learning" (2021) (cit. on p. 153).

[146] L. Kingma D. P.; Ba. "ADAM: A method for stochastic optimization". In: *Proceedings of the International Conference on Learning Representations*. (San Diego, CA, USA). Ed. by Y. Bengio and Y. LeCun. 2015 (cit. on pp. 33, 49, 61, 75).

[147] F. Knoll, M. Holler, T. Koesters, R. Otazo, K. Bredies, and D. K. Sodickson. "Joint MR-PET reconstruction using a multi-channel image regularizer". *IEEE Transactions on Medical Imaging* 36.1 (2016), pp. 1–16 (cit. on p. 28).

[148] R. F. Kokaly, R. N. Clark, G. A. Swayze, K. E. Livo, T. M. Hoefen, N. C. Pearson, R. A. Wise, W. M. Benzel, H. A. Lowers, R. L. Driscoll, and A. J. Klein. "USGS spectral library version 7 data: US geological survey data release". *United States Geological Survey (USGS): Reston, VA, USA* (2017) (cit. on p. 80).

[149] R. F. Kokaly, R. N. Clark, G. A. Swayze, K. E. Livo, T. M. Hoefen, N. C. Pearson, R. A. Wise, W. M. Benzel, H. A. Lowers, R. L. Driscoll, and A. J. Klein. *USGS spectral library version 7*. Tech. rep. US Geological Survey, 2017 (cit. on p. 80).

[150] O. Kouropteva, O. Okun, and M. Pietikäinen. "Selection of the optimal parameter value for the locally linear embedding algorithm". *FSKD* 2 (2002), pp. 359–363 (cit. on p. 68).

[151] M. V. Koval'chuk, E. B. Yatsishina, A. E. Blagov, E. Y. Tereshchenko, P. A. Prosekov, and Y. A. Dyakova. "X-ray and synchrotron methods in studies of cultural heritage sites". *Crystallography Reports* 61.5 (2016), pp. 703–717 (cit. on p. 13).

[152] B. Kumar, O. Dikshit, A. Gupta, and M. K. Singh. "Feature extraction for hyperspectral image classification: A review". *International Journal of Remote Sensing* 41.16 (2020), pp. 6248–6287 (cit. on pp. 67, 68).

[153] J. Kwon, J. Lee, and W. Kim. "Real-time detection of foreign objects using X-ray imaging for dry food manufacturing line". In: *2008 IEEE International Symposium on Consumer Electronics*. (Vilamoura, Portugal). IEEE, 2008, pp. 1–4 (cit. on p. 40).

[154] M. J. Lagerwerf, D. M. Pelt, W. J. Palenstijn, and K. J. Batenburg. "A computationally efficient reconstruction algorithm for circular cone-beam computed tomography using shallow neural networks". *Journal of Imaging* 6.12 (2020), p. 135 (cit. on p. 49).

[155] J. H. Lee, Y. J. Kim, and K. G. Kim. "Bone age estimation using deep learning and hand X-ray images". *Biomedical Engineering Letters* 10.3 (2020), pp. 323–331 (cit. on p. 35).

[156] T. van Leeuwen and A. Y. Aravkin. "Variable Projection for NonSmooth Problems". *SIAM Journal on Scientific Computing* 43.5 (2021), S249–S268 (cit. on p. 125).

[157] M. M. Lell and M. Kachelrieß. "Recent and upcoming technological developments in computed tomography: High speed, low dose, deep learning, multienergy". *Investigative Radiology* 55.1 (2020), pp. 8–19 (cit. on p. 35).

[158] L. Lenchik, L. Heacock, A. A. Weaver, R. D. Boutin, T. S. Cook, J. Itri, C. G. Filippi, R. P. Gullapalli, J. Lee, and e. a. Zagurovskaya M. "Automated segmentation of tissues using CT and MRI: A systematic review". *Academic Radiology* 26.12 (2019), pp. 1695–1706 (cit. on p. 44).

[159] S. Leng, M. Bruesewitz, S. Tao, K. Rajendran, A. F. Halaweish, N. G. Campeau, J. G. Fletcher, and C. H. McCollough. "Photon-counting detector CT: System design and clinical applications of an emerging technology". *Radiographics* 39.3 (2019), pp. 729–743 (cit. on pp. 23, 28).

[160] J. Leuschner, M. Schmidt, P. S. Ganguly, V. Andriiashen, S. B. Coban, A. Denker, D. Bauer, A. Hadjifaradji, K. J. Batenburg, P. Maass, and M. van Eijnatten. "Quantitative comparison of deep learning-based image reconstruction methods for low-dose and sparse-angle CT applications". *Journal of Imaging* 7.3 (2021), p. 44 (cit. on p. 35).

[161] F. Li, M. K. Ng, R. Plemmons, S. Prasad, and Q. Zhang. "Hyperspectral image segmentation, deblurring, and spectral analysis for material identification". In: *Visual Information Processing XIX*. (Orlando, FL, USA). Vol. 7701. International Society for Optics and Photonics. 2010, p. 770103 (cit. on p. 66).

[162] S. Li, H. Luo, M. Hu, M. Zhang, J. Feng, Y. Liu, Q. Dong, and B. Liu. "Optical non-destructive techniques for small berry fruits: A review". *Artificial Intelligence in Agriculture* 2 (2019), pp. 85–98 (cit. on p. 40).

[163] S. Li, W. Song, L. Fang, Y. Chen, P. Ghamisi, and J. A. Benediktsson. "Deep learning for hyperspectral image classification: An overview". *IEEE Transactions on Geoscience and Remote Sensing* 57.9 (2019), pp. 6690–6709 (cit. on p. 68).

[164] T. Li, J. Zhang, and Y. Zhang. "Classification of hyperspectral image based on deep belief networks". In: *2014 IEEE International Conference on Image Processing (ICIP)*. (Paris, France). IEEE. 2014, pp. 5132–5136 (cit. on p. 67).

[165] W. Li, F. Feng, H. Li, and Q. Du. "Discriminant analysis-based dimension reduction for hyperspectral image classification: A survey of the most recent advances and an experimental comparison of different techniques". *IEEE Geoscience and Remote Sensing Magazine* 6.1 (2018), pp. 15–34 (cit. on p. 67).

[166] Z. Li, S. Ravishankar, Y. Long, and J. A. Fessler. "DECT-MULTRA: Dual-energy CT image decomposition with learned mixed material models and efficient clustering". *IEEE Transactions on Medical Imaging* 39.4 (2019), pp. 1223–1234 (cit. on p. 119).

[167] G. Lin, A. Milan, C. Shen, and I. Reid. "Refinenet: Multi-path refinement networks for high-resolution semantic segmentation". In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. (Honolulu, HI, USA). IEEE, 2017, pp. 1925–1934 (cit. on p. 43).

[168] B. Liu, X. Yu, P. Zhang, A. Yu, Q. Fu, and X. Wei. "Supervised deep feature extraction for hyperspectral image classification". *IEEE Transactions on Geoscience and Remote Sensing* 56.4 (2017), pp. 1909–1921 (cit. on p. 68).

[169] B. Liu, Y. Li, G. Li, and A. Liu. "A spectral feature based convolutional neural network for classification of sea surface oil spill". *ISPRS International Journal of Geo-Information* 8.4 (2019), p. 160 (cit. on p. 69).

[170] J. Liu and H. Gao. "Material reconstruction for spectral computed tomography with detector response function". *Inverse Problems* 32.11 (2016), p. 114001 (cit. on p. 154).

[171] J. Long, E. Shelhamer, and T. Darrell. "Fully convolutional networks for semantic segmentation". In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. (Boston, MA, USA). IEEE, 2015, pp. 3431–3440 (cit. on p. 43).

[172] Y. Long and J. A. Fessler. "Multi-material decomposition using statistical image reconstruction for spectral CT". *IEEE Transactions on Medical Imaging* 33.8 (2014), pp. 1614–1626 (cit. on pp. 26, 132).

[173] T. van de Looverbosch, E. Raeymaekers, P. Verboven, J. Sijbers, and B. Nicolaï. "Nondestructive internal disorder detection of Conference pears by semantic segmentation of X-ray CT scans using deep learning". *Expert Systems with Applications* 176 (2021), p. 114925 (cit. on p. 41).

[174] T. van de Looverbosch, M. H. R. Bhuiyan, P. Verboven, M. Dierick, D. van Loo, J. De Beenbouwer, J. Sijbers, and B. Nicolaï. "Nondestructive internal quality inspection of pear fruit by X-ray CT using machine learning". *Food Control* 113 (2020), p. 107170 (cit. on p. 13).

[175] G. Lu and B. Fei. "Medical hyperspectral imaging: A review". *Journal of Biomedical Optics* 19.1 (2014), p. 010901 (cit. on p. 66).

[176] T. Lukić and P. Balázs. "Limited-view binary tomography reconstruction assisted by shape centroid". *The Visual Computer* 38.2 (2022), pp. 695–705 (cit. on p. 18).

[177] S. Ma, X. Zhang, C. Jia, Z. Zhao, S. Wang, and S. Wang. "Image and video compression with neural networks: A review". *IEEE Transactions on Circuits and Systems for Video Technology* 30.6 (2019), pp. 1683–1698 (cit. on p. 35).

[178] A. L. Maas, A. Y. Hannun, and A. Y. Ng. "Rectifier nonlinearities improve neural network acoustic models". In: *Proceedings of the ICML Workshop on Deep Learning for Audio, Speech and Language Processing*. (Atlanta, GA, USA). Vol. 30. 1. 2013, p. 3 (cit. on pp. 31, 73).

[179] C. Maaß, M. Baer, and M. Kachelrieß. "Image-based dual energy CT using optimized precorrection functions: A practical new approach of material decomposition in image domain". *Medical Physics* 36.8 (2009), pp. 3818–3829 (cit. on p. 26).

[180] L. van der Maaten, E. Postma, and J. van den Herik. "Dimensionality reduction: A comparative review". *Journal of Machine Learning Research* 10.66-71 (2009), p. 13 (cit. on p. 68).

[181] F. J. Maestre-Deusto, G. Scavello, J. Pizarro, and P. L. Galindo. "ADART: An adaptive algebraic reconstruction algorithm for discrete tomography". *IEEE Transactions on Image Processing* 20.8 (2011), pp. 2146–2152 (cit. on p. 96).

[182] A. Maier, H. G. Hofmann, M. Berger, P. Fischer, C. Schwemmer, H. Wu, K. Müller, J. Hornegger, J. Choi, C. Riess, A. Keil, and R. Fahrig. "CONRAD - A software framework for cone-beam imaging in radiology". *Medical Physics* 40.11 (2013), p. 111914 (cit. on p. 129).

[183] A. Maier, C. Syben, T. Lasser, and C. Riess. "A gentle introduction to deep learning in medical image processing". *Zeitschrift für Medizinische Physik* 29.2 (2019), pp. 86–101 (cit. on p. 35).

[184] K. Makantasis, K. Karantzalos, A. Doulamis, and N. Doulamis. "Deep supervised learning for hyperspectral data classification through convolutional neural networks". In: *2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*. (Milan, Italy). IEEE. 2015, pp. 4959–4962 (cit. on pp. 66, 68).

[185] H. E. Martz, C. M. Logan, D. J. Schneberk, and P. J. Shull. *X-ray imaging: Fundamentals, industrial techniques and applications*. Boca Raton, FL, USA: CRC Press, 2016 (cit. on p. 35).

[186] S. K. Mathanker, P. R. Weckler, and T. J. Bowser. "X-ray applications in food and agriculture: A review". *Transactions of the ASABE* 56.3 (2013), pp. 1227–1239 (cit. on pp. 18, 40).

[187] J. Matula, V. Polakova, J. Salplachta, M. Tesarova, T. Zikmund, M. Kaucka, I. Adameyko, and J. Kaiser. "Resolving complex cartilage structures in developmental biology via deep learning-based automatic segmentation of X-ray computed microtomography images". *Scientific Reports* 12.1 (2022), p. 8728 (cit. on p. 35).

[188] M. T. McCann, K. H. Jin, and M. Unser. "Convolutional neural networks for inverse problems in imaging: A review". *IEEE Signal Processing Magazine* 34.6 (2017), pp. 85–95 (cit. on p. 32).

[189] 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* 37.1 (2017), pp. 68–80 (cit. on pp. 119, 132).

[190] K. Mechlem, T. Sellerer, S. Ehn, D. Münzel, E. Braig, J. Herzen, P. Noël, and F. Pfeiffer. "Spectral angiography material decomposition using an empirical forward model and a dictionary-based regularization". *IEEE Transactions on Medical Imaging* 37.10 (2018), pp. 2298–2309 (cit. on pp. 118, 119).

[191] D. Mery, I. Lillo, H. Loebel, V. Riffo, A. Soto, A. Cipriano, and J. M. Aguilera. "Automated fish bone detection using X-ray imaging". *Journal of Food Engineering* 105.3 (2011), pp. 485–492 (cit. on pp. 11, 40).

[192] D. Mery. "Computer vision for X-ray testing". *Switzerland: Springer International Publishing* 10 (2015), pp. 978–3 (cit. on pp. 6–8, 22, 35, 153).

[193] D. Mery, D. Saavedra, and M. Prasad. "X-ray baggage inspection with computer vision: A survey". *IEEE Access* 8 (2020), pp. 145620–145633 (cit. on p. 13).

[194] H. S. El-Mesery, H. Mao, and A. E.-F. Abomohra. "Applications of non-destructive technologies for agricultural and food products quality inspection". *Sensors* 19.4 (2019), p. 846 (cit. on p. 4).

[195] P. Meyer, V. Noblet, C. Mazzara, and A. Lallement. "Survey on deep learning for radiotherapy". *Computers in Biology and Medicine* 98 (2018), pp. 126–146 (cit. on p. 35).

[196] P. A. Midgley and M. Weyland. "3D electron microscopy in the physical sciences: The development of Z-contrast and EFTEM tomography". *Ultramicroscopy* 96.3-4 (2003), pp. 413–431 (cit. on p. 96).

[197] J. Minnema, M. van Eijnatten, H. Der Sarkissian, S. Doyle, J. Koivisto, J. Wolff, T. Forouzanfar, F. Lucka, and K. J. Batenburg. "Efficient high cone-angle artifact reduction in circular cone-beam CT using deep learning with geometry-aware dimension reduction". *Physics in Medicine & Biology* 66.13 (2021), p. 135015 (cit. on p. 35).

[198] D. Mishra, S. K. Singh, and R. K. Singh. "Deep architectures for image compression: A critical review". *Signal Processing* 191 (2022), p. 108346 (cit. on p. 35).

[199] S. Si-Mohamed, D. Bar-Ness, M. Sigovan, D. P. Cormode, P. Coulon, E. Coche, A. Vlassenbroek, G. Normand, L. Boussel, and P. Douek. "Review of an initial experience with an experimental spectral photon-counting computed tomography system". *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 873 (2017), pp. 27–35 (cit. on pp. 40, 61, 114, 118).

[200] M. T. Mohd Khairi, S. Ibrahim, M. A. Md Yunus, and M. Faramarzi. "Noninvasive techniques for detection of foreign bodies in food: A review". *Journal of Food Process Engineering* 41.6 (2018), e12808 (cit. on pp. 4, 6, 40).

[201] M. P. Morigi, F. Casali, M. Bettuzzi, R. Brancaccio, and V. d'Errico. "Application of X-ray computed tomography to cultural heritage diagnostics". *Applied Physics A* 100.3 (2010), pp. 653–661 (cit. on p. 13).

[202] C. Mory, B. Sixou, S. Si-Mohamed, L. Boussel, and S. Rit. "Comparison of five one-step reconstruction algorithms for spectral CT". *Physics in Medicine & Biology* 63.23 (2018), p. 235001 (cit. on pp. 119, 129, 131).

[203] G. Motta, F. Rizzo, and J. A. Storer. *Hyperspectral data compression*. Berlin, Germany: Springer Science & Business Media, 2006 (cit. on p. 66).

[204] A. Mouton and T. P. Breckon. "A review of automated image understanding within 3D baggage computed tomography security screening". *Journal of X-ray Science and Technology* 23.5 (2015), pp. 531–555 (cit. on p. 13).

[205] K. P. Murphy. *Probabilistic machine learning: Advanced topics*. Cambridge, MA, USA: MIT Press, 2022 (cit. on pp. 29, 33).

[206] J. Nalepa, M. Myller, and M. Kawulok. "Validating hyperspectral image segmentation". *IEEE Geoscience and Remote Sensing Letters* 16.8 (2019), pp. 1264–1268 (cit. on pp. 66, 77).

[207] K. Narsaiah, A. K. Biswas, and P. K. Mandal. "Nondestructive methods for carcass and meat quality evaluation". In: *Meat Quality Analysis*. Ed. by A. K. Biswas and P. K. Mandal. Academic Press, 2020, pp. 37–49 (cit. on p. 40).

[208] R. A. Nasirudin, R. Tachibana, J. J. Näppi, K. Mei, F. K. Kopp, E. J. Rummeny, H. Yoshida, and P. B. Noël. "A comparison of material decomposition techniques for dual-energy CT colonography". In: *Medical Imaging 2015: Physics of Medical Imaging*. Vol. 9412. International Society for Optics and Photonics. 2015, 94124F (cit. on p. 26).

[209] F. Natterer. *The mathematics of computerized tomography*. Philadelphia, PA, USA: SIAM, 2001 (cit. on p. 14).

[210] B. M. Nicolaï, T. Defraeye, B. De Ketelaere, E. Herremans, M. L. A. T. M. Hertog, W. Saeys, A. Torricelli, T. Vandendriessche, and P. Verboven. "Nondestructive measurement of fruit and vegetable quality". *Annual Review of Food Science and Technology* 5 (2014), pp. 285–312 (cit. on p. 40).

[211] V. V. Nikitin, M. Carlsson, F. Andersson, and R. Mokso. "Four-dimensional tomographic reconstruction by time domain decomposition". *IEEE Transactions on Computational Imaging* 5.3 (2019), pp. 409–419 (cit. on p. 61).

[212] H. Noh, S. Hong, and B. Han. "Learning deconvolution network for semantic segmentation". In: *Proceedings of the IEEE International Conference on Computer Vision*. (Santiago, Chile). IEEE, 2015, pp. 1520–1528 (cit. on p. 43).

[213] I. Ordavo, S. Ihle, V. Arkadiev, O. Scharf, H. Soltau, A. Bjeoumikhov, S. Bjeoumikhova, G. Buzanich, R. Gubzhokov, A. Günther, et al. "A new pnCCD-based color X-ray camera for fast spatial and energy-resolved measurements". *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 654.1 (2011), pp. 250–257 (cit. on pp. 21, 114).

[214] N. Otsu. "A threshold selection method from gray-level histograms". *IEEE Transactions on Systems, Man, and Cybernetics* 9.1 (1979), pp. 62–66 (cit. on pp. 48, 179).

[215] W. J. Palenstijn, K. J. Batenburg, and J. Sijbers. "The ASTRA tomography toolbox". In: *13th International Conference on Computational and Mathematical Methods in Science and Engineering, CMMSE*. (Almeria, Spain). Vol. 2013. 2013, pp. 1139–1145 (cit. on p. 100).

[216] F. Palsson, J. R. Sveinsson, and M. O. Ulfarsson. "Multispectral and hyperspectral image fusion using a 3-D-convolutional neural network". *IEEE Geoscience and Remote Sensing Letters* 14.5 (2017), pp. 639–643 (cit. on p. 69).

[217] H. Pan, C. Zhou, Q. Zhu, and D. Zheng. "A fast registration from 3D CT images to 2D X-ray images". In: *2018 IEEE 3rd International Conference on Big Data Analysis (ICBDA)*. (Shanghai, China). IEEE. IEEE, 2018, pp. 351–355 (cit. on p. 41).

[218] M. E. Paoletti, J. M. Haut, J. Plaza, and A. Plaza. "A new deep convolutional neural network for fast hyperspectral image classification". *ISPRS Journal of Photogrammetry and Remote Sensing* 145 (2018), pp. 120–147 (cit. on pp. 68, 69).

[219] M. E. Paoletti, J. M. Haut, J. Plaza, and A. Plaza. "Deep learning classifiers for hyperspectral imaging: A review". *ISPRS Journal of Photogrammetry and Remote Sensing* 158 (2019), pp. 279–317 (cit. on p. 68).

[220] N. Parikh and S. Boyd. "Proximal algorithms". *Foundations and Trends in Optimization* 1.3 (2014), pp. 127–239 (cit. on p. 125).

[221] M. S. Passmore, R. Bates, K. Mathieson, V. O'Shea, M. Rahman, P. Seller, and K. M. Smith. "Characterisation of a single photon counting pixel detector". *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 466.1 (2001), pp. 202–208 (cit. on p. 120).

[222] A. Paszke, S. Gross, S. Chintala, G. Chanan, E. Yang, Z. DeVito, Z. Lin, A. Desmaison, L. Antiga, and A. Lerer. "Automatic differentiation in PyTorch". In: *Proceedings of the NIPS Autodiff Workshop*. (Long Beach, CA, USA). 2017 (cit. on pp. 35, 48, 63, 77, 93).

[223] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, and e. a. A. Luca. "PyTorch: An imperative style, high-performance deep learning library". In: *Advances in Neural Information Processing Systems*. (Red Hook, NY, USA). Ed. by H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché-Buc, E. Fox, and R. Garnett. 2019, pp. 8026–8037 (cit. on pp. 35, 48, 63, 77, 93).

[224] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. "Scikit-learn: Machine learning in Python". *Journal of Machine Learning Research* 12 (2011), pp. 2825–2830 (cit. on p. 82).

[225] D. M. Pelt. *GitHub - dmpelt/msdnet: Python implementation of the Mixed-Scale Dense Convolutional Neural Network*. `https://github.com/dmpelt/msdnet`. Accessed on 24 November 2020. 2019 (cit. on pp. 49, 75).

[226] D. M. Pelt, K. J. Batenburg, and J. A. Sethian. "Improving tomographic reconstruction from limited data using mixed-scale dense convolutional neural networks". *Journal of Imaging* 4.11 (2018), p. 128 (cit. on pp. 35, 49).

[227] D. M. Pelt and J. A. Sethian. "A mixed-scale dense convolutional neural network for image analysis". *Proceedings of the National Academy of Sciences* 115.2 (2018), pp. 254–259 (cit. on pp. 48, 49, 63, 67, 70, 71, 74, 75, 93).

[228] H. Petersson, D. Gustafsson, and D. Bergstrom. "Hyperspectral image analysis using deep learning - A review". In: *2016 Sixth International Conference on Image Processing Theory, Tools and Applications (IPTA)*. (Oulu, Finland). IEEE. 2016, pp. 1–6 (cit. on p. 68).

[229] G. Pilikos, L. Horchens, K. J. Batenburg, T. van Leeuwen, and F. Lucka. "Deep data compression for approximate ultrasonic image formation". In: *2020 IEEE International Ultrasonics Symposium (IUS)*. IEEE. 2020, pp. 1–4 (cit. on p. 35).

[230] G. Poludniowski, A. Omar, R. Bujila, and P. Andreo. "SpekPy v2.0 - A software toolkit for modelling X-ray tube spectra". *Medical Physics* 48.7 (2021) (cit. on p. 130).

[231] T. Qiao, J. Ren, Z. Wang, J. Zabalza, M. Sun, H. Zhao, S. Li, J. A. Benediktsson, Q. Dai, and S. Marshall. "Effective denoising and classification of hyperspectral images using curvelet transform and singular spectrum analysis". *IEEE Transactions on Geoscience and Remote Sensing* 55.1 (2016), pp. 119–133 (cit. on p. 67).

[232] J. Radon. "1.1 über die bestimmung von funktionen durch ihre integralwerte längs gewisser mannigfaltigkeiten". *Classic Papers in Modern Diagnostic Radiology* 5 (2005), p. 21 (cit. on p. 14).

[233] N. Rassouli, M. Etesami, A. Dhanantwari, and P. Rajiah. "Detector-based spectral CT with a novel dual-layer technology: Principles and applications". *Insights into Imaging* 8.6 (2017), pp. 589–598 (cit. on p. 24).

[234] B. Rasti, D. Hong, R. Hang, P. Ghamisi, X. Kang, J. Chanussot, and J. A. Benediktsson. "Feature extraction for hyperspectral imagery: The evolution from shallow to deep". *arXiv preprint arXiv:2003.02822* (2020) (cit. on p. 67).

[235] V. Rebuffel and J. Dinten. "Dual-energy X-ray imaging: Benefits and limits". *Insight - Non-Destructive Testing and Condition Monitoring,* 49.10 (2007), pp. 589–594 (cit. on p. 61).

[236] R. Redus, A. Huber, J. Pantazis, T. Pantazis, and D. Sperry. "Design and performance of the X-123 compact X-ray and Gamma-ray spectroscopy system". In: *2006 IEEE Nuclear Science Symposium Conference Record.* (Piscataway Township, NJ, USA). Vol. 6. IEEE. 2006, pp. 3794–3797 (cit. on p. 114).

[237] E. G. Rens, M. T. Zeegers, I. Rabbers, A. Szabó, and R. M. H. Merks. "Autocrine inhibition of cell motility can drive epithelial branching morphogenesis in the absence of growth". *Philosophical Transactions of the Royal Society B* 375.1807 (2020), p. 20190386 (cit. on p. 209).

[238] D. S. Rigie and P. J. La Rivière. "Joint reconstruction of multi-channel, spectral CT data via constrained total nuclear variation minimization". *Physics in Medicine & Biology* 60.5 (2015), p. 1741 (cit. on pp. 28, 61).

[239] D. Rong, L. Xie, and Y. Ying. "Computer vision detection of foreign objects in walnuts using deep learning". *Computers and Electronics in Agriculture* 162 (2019), pp. 1001–1010 (cit. on p. 40).

[240] 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.* (Munich, Germany). Ed. by N. Navab, J. Hornegger, W. M. Wells, and A. F. Frangi. Springer, 2015, pp. 234–241 (cit. on pp. 43, 48, 67, 71, 74).

[241] L. I. Rudin, S. Osher, and E. Fatemi. "Nonlinear total variation based noise removal algorithms". *Physica D: Nonlinear Phenomena* 60.1-4 (1992), pp. 259–268 (cit. on p. 133).

[242] P. Russo. *Handbook of X-ray imaging: Physics and technology.* Boca Raton, FL, USA: (1st ed.). CRC Press, 2017 (cit. on pp. 7, 43).

[243] J. Salazar-Vazquez and A. Mendez-Vazquez. "A plug-and-play hyperspectral imaging sensor using low-cost equipment". *HardwareX* 7 (2020), e00087 (cit. on p. 155).

[244] M. Salehjahromi, Y. Zhang, and H. Yu. "Comparison study of regularizations in spectral computed tomography reconstruction". *Sensing and Imaging* 19.1 (2018), pp. 1–23 (cit. on pp. 27, 119).

[245] M. Al-Sarayreh, M. M. Reis, W. Q. Yan, and R. Klette. "A sequential CNN approach for foreign object detection in hyperspectral images". In: *International Conference on Computer Analysis of Images and Patterns.* (Salerno, Italy). Ed. by M. Vento and G. Percannella. Springer, 2019, pp. 271–283 (cit. on pp. 40, 41).

[246] A. Sawatzky, Q. Xu, C. O. Schirra, and M. A. Anastasio. "Proximal ADMM for multi-channel image reconstruction in spectral X-ray CT". *IEEE Transactions on Medical Imaging* 33.8 (2014), pp. 1657–1668 (cit. on pp. 61, 119).

[247] O. Scharf, S. Ihle, I. Ordavo, V. Arkadiev, A. Bjeoumikhov, S. Bjeoumikhova, G. Buzanich, R. Gubzhokov, A. Gunther, R. Hartmann, et al. "Compact pnCCD-based X-ray camera with high spatial and energy resolution: A color X-ray camera". *Analytical Chemistry* 83.7 (2011), pp. 2532–2538 (cit. on p. 114).

[248] C. O. Schirra, E. Roessl, T. Koehler, B. Brendel, A. Thran, D. Pan, M. A. Anastasio, and R. Proksa. "Statistical reconstruction of material decomposed data in spectral CT". *IEEE Transactions on Medical Imaging* 32.7 (2013), pp. 1249–1257 (cit. on p. 118).

[249] C. O. Schirra, B. Brendel, M. A. Anastasio, and E. Roessl. "Spectral CT: A technology primer for contrast agent development". *Contrast Media & Molecular Imaging* 9.1 (2014), pp. 62–70 (cit. on pp. 19, 21, 28).

[250] M. Schmidt, E. Berg, M. Friedlander, and K. Murphy. "Optimizing costly functions with simple constraints: A limited-memory projected quasi-newton algorithm". In: *Artificial Intelligence and Statistics.* (Clearwater Beach, FL, USA). PMLR. 2009, pp. 456–463 (cit. on pp. 132, 146).

[251] T. G. Schmidt, B. A. Sammut, R. F. Barber, X. Pan, and E. Y. Sidky. "Addressing CT metal artifacts using photon-counting detectors and one-step spectral CT image reconstruction". *Medical Physics* (2022) (cit. on p. 27).

[252] R. Schoonhoven, A. A. Hendriksen, D. M. Pelt, and K. J. Batenburg. "LEAN: Graph-based pruning for convolutional neural networks by extracting longest chains". *arXiv preprint arXiv:2011.06923* (2020) (cit. on p. 154).

[253] D. E. Schut, K. J. Batenburg, R. van Liere, and T. van Leeuwen. "TOP-CT: Trajectory with Overlapping Projections X-ray Computed Tomography". *IEEE Transactions on Computational Imaging* 8 (2022), pp. 598–608 (cit. on p. 154).

[254] O. Semerci, N. Hao, M. E. Kilmer, and E. L. Miller. "Tensor-based formulation and nuclear norm regularization for multienergy computed tomography". *IEEE Transactions on Image Processing* 23.4 (2014), pp. 1678–1693 (cit. on p. 61).

[255] D. Sero, I. Garachon, E. Hermens, R. van Liere, and K. J. Batenburg. "The study of three-dimensional fingerprint recognition in cultural heritage: Trends and challenges". *Journal on Computing and Cultural Heritage* 14.4 (2021), pp. 1–20 (cit. on p. 13).

[256] M. Sezgin and B. Sankur. "Survey over image thresholding techniques and quantitative performance evaluation". *Journal of Electronic imaging* 13.1 (2004), pp. 146–165 (cit. on p. 40).

[257] V. Sharma, A. Diba, T. Tuytelaars, and L. van Gool. "Hyperspectral CNN for image classification & band selection, with application to face recognition". *Technical Report KUL/ESAT/PSI/1604, KU Leuven, ESAT, Leuven, Belgium* (2016) (cit. on p. 66).

[258] *Siemens Healthineers, Simulation of X-ray Spectra: Online tool for the simulation of X-ray Spectra (2018)*. `https://www.oem-xray-components.siemens.com/X-ray-spectra-simulation`. Accessed on 24 February 2020 (cit. on p. 185).

[259] A. Signoroni, M. Savardi, A. Baronio, and S. Benini. "Deep learning meets hyperspectral image analysis: A multidisciplinary review". *Journal of Imaging* 5.5 (2019), p. 52 (cit. on pp. 66, 69, 77).

[260] G. Silva, L. Oliveira, and M. Pithon. "Automatic segmenting teeth in X-ray images: Trends, a novel data set, benchmarking and future perspectives". *Expert Systems with Applications* 107 (2018), pp. 15–31 (cit. on pp. 40, 41).

[261] J. da Silva, F. Grönberg, B. Cederström, M. Persson, M. Sjölin, Z. Alagic, R. Bujila, and M. Danielsson. "Resolution characterization of a silicon-based, photon-counting computed tomography prototype capable of patient scanning". *Journal of Medical Imaging* 6.4 (2019), p. 043502 (cit. on p. 28).

[262] J. Sittner, J. R. A. Godinho, A. D. Renno, V. Cnudde, M. Boone, T. De Schryver, D. van Loo, M. Merkulova, A. Roine, and J. Liipo. "Data for: Spectral X-ray computed micro tomography: 3-dimensional chemical imaging [Data set]" (2022) (cit. on pp. 141, 147).

[263] J. Sittner, J. R. A. Godinho, A. D. Renno, V. Cnudde, M. Boone, T. De Schryver, D. van Loo, M. Merkulova, A. Roine, and J. Liipo. "Spectral X-ray computed micro tomography: 3-dimensional chemical imaging". *X-ray Spectrometry* 50.2 (2021), pp. 92–105 (cit. on pp. 21, 141).

[264] W. Skrzynski. "X-ray detectors in medical imaging". In: *Advanced X-ray Detector Technologies*. Springer, 2022, pp. 135–149 (cit. on pp. 22, 23).

[265] A. van der Sluis and H. A. van der Vorst. "SIRT-and CG-type methods for the iterative solution of sparse linear least-squares problems". *Linear Algebra and its Applications* 130 (1990), pp. 257–303 (cit. on p. 47).

[266] A. So and S. Nicolaou. "Spectral computed tomography: Fundamental principles and recent developments". *Korean Journal of Radiology* 22.1 (2021), p. 86 (cit. on p. 26).

[267] J. Spiegelberg and J. Rusz. "Can we use PCA to detect small signals in noisy data?" *Ultramicroscopy* 172 (2017), pp. 40–46 (cit. on p. 66).

[268] *Standard Solar Spectra | PVEducation (2019)*. `https://www.pveducation.org/pvcdrom/appendices/standard-solar-spectra`. Accessed on 8 July 2020 (cit. on p. 80).

[269] E. Ström, M. Persson, A. Eguizabal, and O. Öktem. "Photon-counting CT reconstruction with a learned forward operator". *IEEE Transactions on Computational Imaging* 8 (2022), pp. 536–550 (cit. on p. 35).

[270] C. H. Sudre, W. Li, T. Vercauteren, S. Ourselin, and M. J. Cardoso. "Generalised dice overlap as a deep learning loss function for highly unbalanced segmentations". In: *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*. Ed. by M. J. Cardoso, T. Arbel, G. Carneiro, T. Syeda-Mahmood, J. M. R. S. Tavares, M. Moradi, A. Bradley, H. Greenspan, J. P. Papa, A. Madabhushi, J. C. Nascimento, J. S. Cardoso, V. Belagiannis, and Z. Lu. Springer, 2017, pp. 240–248 (cit. on pp. 48, 77).

[271] C. Sun, A. Shrivastava, S. Singh, and A. Gupta. "Revisiting unreasonable effectiveness of data in deep learning era". In: *Proceedings of the IEEE International Conference on Computer Vision*. (Venice, Italy). 2017, pp. 843–852 (cit. on p. 35).

[272] D. Sun. *Hyperspectral imaging for food quality analysis and control*. San Diego, California, USA: Elsevier, 2010 (cit. on p. 66).

[273] W. Sun and Q. Du. "Hyperspectral band selection: A review". *IEEE Geoscience and Remote Sensing Magazine* 7.2 (2019), pp. 118–139 (cit. on pp. 66, 67).

[274] K. Taguchi, I. Blevis, and K. Iniewski. *Spectral, photon counting computed tomography: Technology and applications*. Boca Raton, FL, USA: CRC Press, 2020 (cit. on pp. 22, 23, 25, 27, 28, 40, 61, 114).

[275] S. Tairi, S. Anthoine, C. Morel, and Y. Boursier. "Simultaneous reconstruction and separation in a spectral CT framework". In: *Nuclear Science Symposium, Medical Imaging Conference and Room-Temperature Semiconductor Detector Workshop (NSS/MIC/RTSD), 2016*. (Strasbourg, France). IEEE. 2016, pp. 1–4 (cit. on p. 110).

[276] N. Tajbakhsh, L. Jeyaseelan, Q. Li, J. N. Chiang, Z. Wu, and X. Ding. "Embracing imperfect datasets: A review of deep learning solutions for medical image segmentation". *Medical Image Analysis* (2020), p. 101693 (cit. on p. 41).

[277] L. Tanzi, E. Vezzetti, R. Moreno, and S. Moos. "X-ray bone fracture classification using deep learning: A baseline for designing a reliable approach". *Applied Sciences* 10.4 (2020), p. 1507 (cit. on p. 35).

[278] K. Thilagavathi and A. Vasuki. "Dimension reduction methods for hyperspectral image: A survey". *International Journal of Engineering and Advanced Technology* 8 (Dec. 2018), pp. 160–167 (cit. on p. 68).

[279] R. Vaddi and M. Prabukumar. "Comparative study of feature extraction techniques for hyperspectral remote sensing image classification: A survey". In: *2017 International Conference on Intelligent Computing and Control Systems (ICICCS)*. (Madurai, India). IEEE. 2017, pp. 543–548 (cit. on p. 67).

[280] D. Valsesia and E. Magli. "High-throughput onboard hyperspectral image compression with ground-based CNN reconstruction". *IEEE Transactions on Geoscience and Remote Sensing* 57.12 (2019), pp. 9544–9553 (cit. on pp. 68, 72).

[281] J. A. T. Vasquez, R. Scapaticci, G. Turvani, M. Ricci, L. Farina, A. Litman, M. R. Casu, L. Crocco, and F. Vipiana. "Noninvasive inline food inspection via microwave imaging technology: An application example in the food industry". *IEEE Antennas and Propagation Magazine* 62.5 (2020), pp. 18–32 (cit. on p. 4).

[282] M. Veale, P. Seller, M. Wilson, and E. Liotti. "HEXITEC: A high-energy X-ray spectroscopic imaging detector for synchrotron applications". *Synchrotron Radiation News* 31.6 (2018), pp. 28–32 (cit. on pp. 21, 114).

[283] G. Vilches-Freixas, V. T. Taasti, L. P. Muren, J. B. B. Petersen, J. M. Létang, D. C. Hansen, and S. Rit. "Comparison of projection-and image-based methods for proton stopping power estimation using dual energy CT". *Physics and Imaging in Radiation Oncology* 3 (2017), pp. 28–36 (cit. on p. 25).

[284] A. S. Wang and N. J. Pelc. "Spectral photon counting CT: Imaging algorithms and performance assessment". *IEEE Transactions on Radiation and Plasma Medical Sciences* 5.4 (2020), pp. 453–464 (cit. on p. 28).

[285] G. Wang, J. C. Ye, and B. De Man. "Deep learning for tomographic image reconstruction". *Nature Machine Intelligence* 2.12 (2020), pp. 737–748 (cit. on p. 35).

[286] M. Wang. *Industrial tomography: Systems and applications*. Sawston, Cambridge, UK: Elsevier, 2015 (cit. on pp. 14, 15).

[287] X. Wang, Y. Bouzembrak, A. O. Lansink, and H. van der Fels-Klerx. "Application of machine learning to the monitoring and prediction of food safety: A review". *Comprehensive Reviews in Food Science and Food Safety* 21.1 (2022), pp. 416–434 (cit. on p. 3).

[288] R. Warr, E. Ametova, R. J. Cernik, G. Fardell, S. Handschuh, J. S. Jørgensen, E. Papoutsellis, E. Pasca, and P. J. Withers. "Enhanced hyperspectral tomography for bioimaging by spatiospectral reconstruction". *Scientific Reports* 11.1 (2021), pp. 1–13 (cit. on p. 28).

[289] T. Weidinger, T. M. Buzug, T. Flohr, S. Kappler, and K. Stierstorfer. "Polychromatic iterative statistical material image reconstruction for photon-counting computed tomography". *International Journal of Biomedical Imaging* 2016 (2016) (cit. on p. 132).

[290] R. E. Wendell and A. P. Hurter Jr. "Minimization of a non-separable objective function subject to disjoint constraints". *Operations Research* 24.4 (1976), pp. 643–657 (cit. on pp. 122, 193).

[291] *Wikipedia Commons, EM Spectrum Properties*. `https://upload.wikimedia.org/wikipedia/commons/c/cf/EM_Spectrum_Properties_edit.svg`. Accessed on 19 December 2020 (cit. on p. 5).

[292] M. J. Willemink and P. B. Noël. "The evolution of image reconstruction for CT - From filtered back projection to artificial intelligence". *European Radiology* 29.5 (2019), pp. 2185–2195 (cit. on pp. 18, 27, 28, 119).

[293] M. J. Willemink, M. Persson, A. Pourmorteza, N. J. Pelc, and D. Fleischmann. "Photon-counting CT: Technical principles and clinical prospects". *Radiology* 289.2 (2018), pp. 293–312 (cit. on p. 22).

[294] K. H. Wilm. "Foreign object detection: Integration in food production". *Food Safety Magazine* 18 (2012), pp. 14–17 (cit. on p. 40).

[295] M. D. Wilson, L. Dummott, D. D. Duarte, F. H. Green, S. Pani, A. Schneider, J. W. Scuffham, P. Seller, and M. C. Veale. "A 10 cm× 10 cm CdTe spectroscopic imaging detector based on the HEXITEC ASIC". *Journal of Instrumentation* 10.10 (2015), P10011 (cit. on p. 96).

[296] P. J. Withers, C. Bouman, S. Carmignato, V. Cnudde, D. Grimaldi, C. K. Hagen, E. Maire, M. Manley, A. Du Plessis, and S. R. Stock. "X-ray computed tomography". *Nature Reviews Methods Primers* 1.1 (2021), pp. 1–21 (cit. on pp. 28, 33).

[297] H. Wu, Q. Liu, and X. Liu. "A review on deep learning approaches to image classification and object segmentation". *Computers, Materials & Continua* 60.2 (2019), pp. 575–597 (cit. on p. 40).

[298] W. Wu, P. Chen, V. V. Vardhanabhuti, W. Wu, and H. Yu. "Improved material decomposition with a two-step regularization for spectral CT". *IEEE Access* 7 (2019), pp. 158770–158781 (cit. on pp. 118, 119).

[299] W. Wu, P. Chen, S. Wang, V. Vardhanabhuti, F. Liu, and H. Yu. "Image-domain material decomposition for spectral CT using a generalized dictionary learning". *IEEE Transactions on Radiation and Plasma Medical Sciences* 5.4 (2020) (cit. on pp. 118, 119).

[300] T. Würfl, F. C. Ghesu, V. Christlein, and A. Maier. "Deep learning computed tomography". In: *International Conference on Medical Image Computing and Computer-assisted Intervention*. (Athens, Greece). Springer. 2016, pp. 432–440 (cit. on p. 35).

[301] *X-ray Mass Attenuation Coefficients, NIST*. `https://www.nist.gov/pml/x-ray-mass-attenuation-coefficients`. Accessed on 24 February 2020 (cit. on pp. 123, 130, 184, 197).

[302] W. Xia, W. Wu, S. Niu, F. Liu, J. Zhou, H. Yu, G. Wang, and Y. Zhang. "Spectral CT reconstruction - ASSIST: Aided by Self-Similarity in Image-Spectral Tensors". *IEEE Transactions on Computational Imaging* 5.3 (2019), pp. 420–436 (cit. on p. 28).

[303] Z. Xiong, D. Sun, H. Pu, W. Gao, and Q. Dai. "Applications of emerging imaging techniques for meat quality and safety detection and evaluation: A review". *Critical Reviews in Food Science and Nutrition* 57.4 (2017), pp. 755–768 (cit. on p. 40).

[304] X. Yang, V. De Andrade, W. Scullin, E. L. Dyer, N. Kasthuri, F. De Carlo, and D. Gürsoy. "Low-dose X-ray tomography through a deep convolutional neural network". *Scientific Reports* 8.1 (2018), pp. 1–13 (cit. on p. 35).

[305] M. Yaqoob, S. Sharma, and P. Aggarwal. "Imaging techniques in agro-industry and their applications, a review". *Journal of Food Measurement and Characterization* 15.3 (2021), pp. 2329–2343 (cit. on pp. 3, 4).

[306] S. Yu, S. Jia, and C. Xu. "Convolutional neural networks for hyperspectral image classification". *Neurocomputing* 219 (2017), pp. 88–98 (cit. on p. 77).

[307] M. T. Zeegers. *A collection of 131 CT datasets of pieces of modeling clay containing stones*. Zenodo, `https://doi.org/10.5281/zenodo.5866228`. 2022 (cit. on p. 63).

[308] M. T. Zeegers. *A collection of X-ray projections of 131 pieces of modeling clay containing stones for machine learning-driven object detection*. Zenodo, `https://doi.org/10.5281/zenodo.5681008`. 2022 (cit. on p. 63).

[309] M. T. Zeegers. *mzeegers/DeepFODDataGenerator: DeepFODDataGenerator*. Zenodo, `https://doi.org/10.5281/zenodo.7825045`. Version v1.0.1. 2023 (cit. on p. 63).

[310] M. T. Zeegers. *mzeegers/DRCNN: Data Reduction CNN*. Zenodo, `https://doi.org/10.5281/zenodo.7824969`. Version v1.0.1. 2023 (cit. on p. 93).

[311] M. T. Zeegers. *mzeegers/MC-DART: MC-DART*. Zenodo, `https://doi.org/10.5281/zenodo.7824878`. Version v1.0.1. 2023 (cit. on p. 111).

[312] M. T. Zeegers, F. Lucka, and K. J. Batenburg. "A Multi-Channel DART algorithm". In: *International Workshop on Combinatorial Image Analysis*. (Porto, Portugal). Ed. by R. P. Barneva, V. E. Brimkov, and J. M. R. S. Tavares. Springer, 2018, pp. 164–178 (cit. on pp. 61, 95, 119, 209).

[313] M. T. Zeegers, D. M. Pelt, T. van Leeuwen, R. van Liere, and K. J. Batenburg. "Task-driven learned hyperspectral data reduction using end-to-end supervised deep learning". *Journal of Imaging* 6.12 (2020), p. 132 (cit. on pp. 65, 209).

[314] M. T. Zeegers, T. van Leeuwen, D. M. Pelt, S. B. Coban, R. van Liere, and K. J. Batenburg. "A tomographic workflow to enable deep learning for X-ray based foreign object detection". *Expert Systems with Applications* 206 (2022), p. 117768 (cit. on pp. 39, 209).

[315] M. T. Zeegers, A. Kadu, T. van Leeuwen, and K. J. Batenburg. "ADJUST: A Dictionary-based Joint reconstruction and Unmixing method for Spectral Tomography". *Inverse Problems* 38.12 (2022), p. 125002 (cit. on pp. 113, 209).

[316] T. Zhang, S. Zhao, X. Ma, A. P. Cuadros, Q. Zhao, and G. R. Arce. "Nonlinear reconstruction of coded spectral X-ray CT based on material decomposition". *Optics Express* 29.13 (2021), pp. 19319–19339 (cit. on p. 119).

[317] X. Zhang, Y. Zheng, W. Liu, and Z. Wang. "A hyperspectral image classification algorithm based on atrous convolution". *EURASIP Journal on Wireless Communications and Networking* 2019.1 (2019), pp. 1–12 (cit. on p. 72).

[318] Y. Zhang, X. Mou, G. Wang, and H. Yu. "Tensor-based dictionary learning for spectral CT reconstruction". *IEEE Transactions on Medical Imaging* 36.1 (2016), pp. 142–154 (cit. on pp. 118, 119).

[319] Y. Zhang and H. Yu. "Convolutional neural network based metal artifact reduction in X-ray computed tomography". *IEEE Transactions on Medical Imaging* 37.6 (2018), pp. 1370–1381 (cit. on p. 35).

[320] Y. Zhang, Y. Xi, Q. Yang, W. Cong, J. Zhou, and G. Wang. "Spectral CT reconstruction with image sparsity and spectral mean". *IEEE Transactions on Computational Imaging* 2.4 (2016), pp. 510–523 (cit. on pp. 27, 28).

[321] H. Zhao, J. Shi, X. Qi, X. Wang, and J. Jia. "Pyramid scene parsing network". In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. (Honolulu, HI, USA). IEEE, 2017, pp. 2881–2890 (cit. on p. 43).

[322] Z. Zhao, P. Zheng, S. Xu, and X. Wu. "Object detection with deep learning: A review". *IEEE Transactions on Neural Networks and Learning Systems* 30.11 (2019), pp. 3212–3232 (cit. on p. 40).

[323] J. Zhong, F. Zhang, Z. Lu, Y. Liu, and X. Wang. "High-speed display-delayed planar X-ray inspection system for the fast detection of small fishbones". *Journal of Food Process Engineering* 42.3 (2019), e13010 (cit. on p. 40).

[324] S. Zhong. "Progress in terahertz nondestructive testing: A review". *Frontiers of Mechanical Engineering* 14.3 (2019), pp. 273–281 (cit. on p. 4).

[325] Z. Zhong, W. J. Palenstijn, J. Adler, and K. J. Batenburg. "EDS tomographic reconstruction regularized by total nuclear variation joined with HAADF-STEM tomography". *Ultramicroscopy* 191 (2018), pp. 34–43 (cit. on p. 28).

[326] T. Zhou, S. Ruan, and S. Canu. "A review: Deep learning for medical image segmentation using multi-modality fusion". *Array* 3 (2019), p. 100004 (cit. on p. 35).

[327] L. Zhu, P. Spachos, E. Pensini, and K. N. Plataniotis. "Deep learning and machine vision for food processing: A survey". *Current Research in Food Science* 4 (2021), pp. 233–249 (cit. on pp. 3, 31, 33, 35, 40).

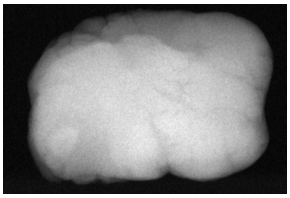
[328] X. Zhuge, W. J. Palenstijn, and K. J. Batenburg. "TVR-DART: A more robust algorithm for discrete tomography from limited projection data with automated gray value estimation". *IEEE Transactions on Image Processing* 25.1 (2016), pp. 455–468 (cit. on pp. 96, 154).

[329] A. Ziabari, S. Venkatakrishnan, M. Kirka, P. Brackman, R. Dehoff, P. Bingham, and V. Paquit. "Beam hardening artifact reduction in X-ray CT reconstruction of 3D printed metal parts leveraging deep learning and CAD models". In: *ASME International Mechanical Engineering Congress and Exposition*. Vol. 84492. American Society of Mechanical Engineers. 2020, V02BT02A043 (cit. on p. 35).

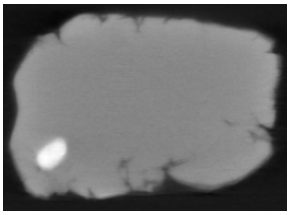
A

Appendices to Chapter 2

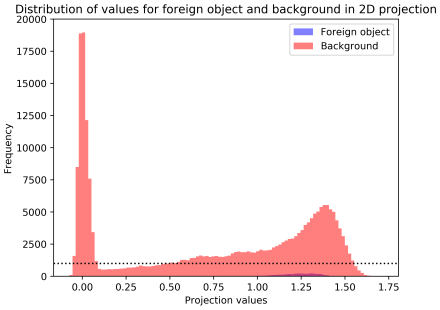
A.1 Intensity value histograms



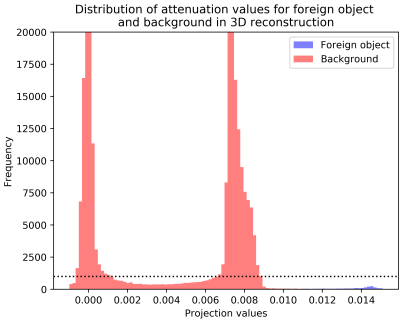
(**a**) 2D radiograph with foreign object on bottom left



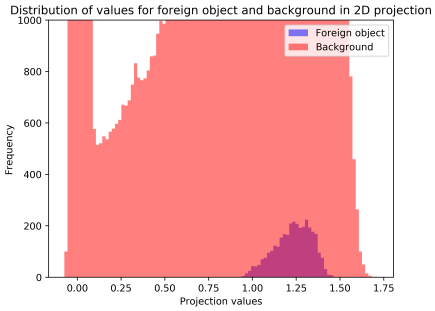
(**b**) Slice of the reconstructed 3D volume with foreign object on bottom left



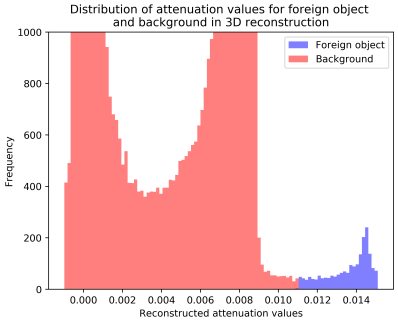
(**c**) Intensity value distribution for the 2D radiograph



(**d**) Attenuation value distribution for the slice of the reconstructed 3D object



(**e**) Intensity value distribution for the 2D radiograph (zoomed)



(**f**) Attenuation value distribution for the slice of the reconstructed 3D object (zoomed)

Figure A1: Radiograph of an object containing a foreign object (**a**) and a slice of the corresponding 3D reconstruction showing its attenuation values (**d**), indicating the difference in contrast. Additionally, histograms of intensity value distribution of the radiograph (**b**-**d**) and the attenuation value distribution of the slice of the reconstructed 3D object (**e**-**f**). In both cases, the histograms of the voxels or pixels of the foreign object are plotted separately from the other voxels or pixels. In the 3D volume, the foreign object is much easier to distinguish based on intensity values.

We compare the intensity distributions for radiographs and for a CT reconstruction of an object in Figure A1, which shows a number of statistics about the pixel and voxel intensities for object 3 (Fig. 2.6). For both approaches, the intensity value distributions are plotted and separated into values of pixel or voxels that have been marked as foreign object by the thresholding method. The 3D case has a clear separation between foreign object and the base object based on attenuation, such that a simple global threshold based on Otsu's method [214] is sufficient to segment the foreign object. On the other hand, in the 2D radiograph case, the intensity values corresponding to the foreign object locations are similar to values of the base object.

A.2 Reconstruction and ground truth similarities

In Section 2.4.8, it is verified that the direct use of generated 3D volumes results in similar ground truth projections compared to the use of the workflow, by indicating that the average Jaccard index between the ground truth pairs is 0.961. In Table A.1, the results are given in greater detail by splitting the results up for nonidentical and combined projections. In addition, we also give the MSE. We also present the similarity results for the segmentations from which the projections are generated. Lastly, results are given for the FDK and SIRT reconstruction algorithms, the latter with 200 iterations. The results indicate that by using these reconstruction algorithms the similarities between the projection pairs increase.

		Identical (%)	**Jaccard remaining**	**Jaccard overall**	**MSE remaining**	**MSE overall**
Segmentation	FDK	15	0.625	0.681	$1.34 \cdot 10^{-4}$	$1.14 \cdot 10^{-4}$
	SIRT, 100 it.	1	0.656	0.659	$1.18 \cdot 10^{-4}$	$1.17 \cdot 10^{-4}$
	SIRT, 200 it.	10	0.643	0.678	$1.28 \cdot 10^{-4}$	$1.15 \cdot 10^{-4}$
Projection	FDK	47.05	0.981	0.990	$1.57 \cdot 10^{-4}$	$0.83 \cdot 10^{-4}$
	SIRT, 100 it.	8.99	0.957	0.961	$2.97 \cdot 10^{-4}$	$3.26 \cdot 10^{-4}$
	SIRT, 200 it.	39.40	0.980	0.988	$1.59 \cdot 10^{-4}$	$0.94 \cdot 10^{-4}$

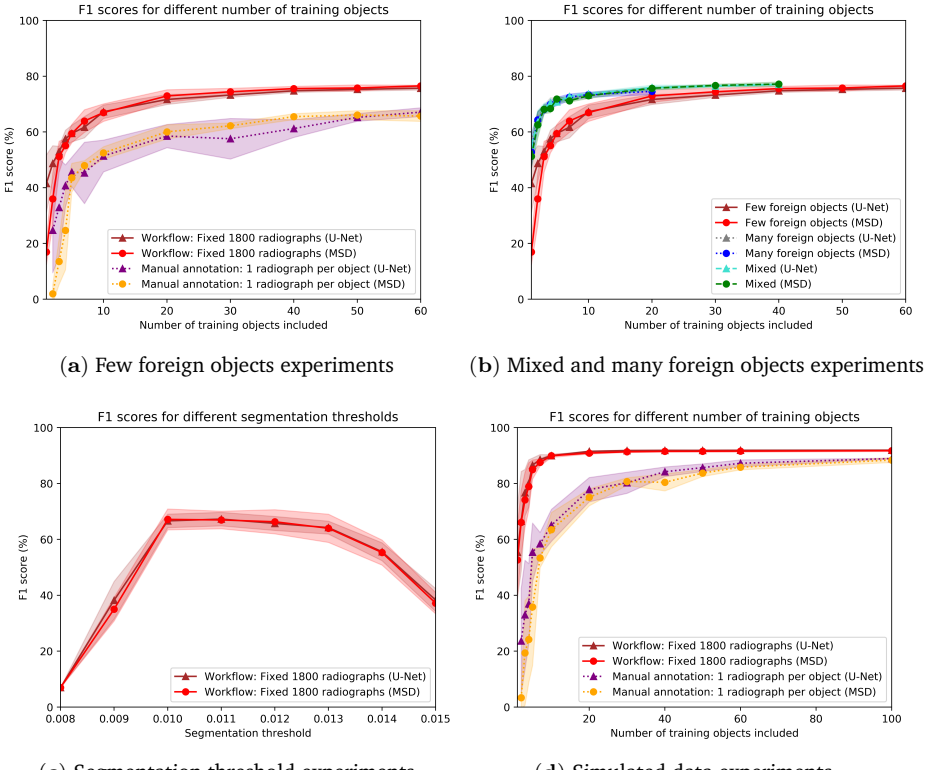
Table A.1: Similarity between ground truth volumes and the corresponding segmented volumes reconstructed from their own projections, as well as similarity between subsequent virtual projections of these volumes. The reconstruction are made over 1800 equidistant angles, and the results are averaged over these angles and 100 training objects. We measure the number of volumes that are identical to their ground truth, and the Jaccard index and the Mean Square Error (MSE) of both all examples and the nonidentical examples only.

Therefore, by adding in an even better reconstruction algorithm in terms of these similarities, an even more accurate training training set can be generated which can subsequently yield more accurate detection results than presented in Chapter 2.

A.3 Additional quality measure

In this Appendix we show the F1 scores for all experiments in Chapter 2. The F1 score is given by

$$\left(\frac{2\mathrm{TP_{FO}}}{2\mathrm{TP_{FO}} + \mathrm{FN_{BG}} + \mathrm{FN_{FO}}}\right). \tag{A.1}$$



(**a**) Few foreign objects experiments

(**b**) Mixed and many foreign objects experiments

(**c**) Segmentation threshold experiments

(**d**) Simulated data experiments

Figure A2: F1 scores for the various experiments in Chapter 2: (**a**) the standard experiment with laboratory data with few foreign objects, (**b**) the experiment with many and mixed amounts of foreign objects, (**c**) the experiment with threshold variation and (**d**) the simulated experiment. The results are shown for trained U-Net and MSD networks. The results are averaged over 5 trained networks, with a different training object order for each run. The shaded regions indicate the standard deviations.

The results for this quality metric are given in Figure A2. The graphs for all experiments are consistent with the graphs for the quality measures in Section 2.4.

B

Appendices to Chapter 3

B.1 Standard data reduction methods

In this section we briefly summarize the most common data reduction methods, used for comparison in this work: PCA (unsupervised), NMF (unsupervised) and LDA (supervised). Let $\boldsymbol{X} \in \mathbb{R}^{N_b \times (N_{\text{train}} \cdot m \cdot n)}$ be a matrix representation of $\{x_i\}_{i=1}^{N_{\text{train}}}$, where the rows represent the spectral features and the columns are the data points.

B.1.1 Principal Component Analysis

Let $\widetilde{\boldsymbol{X}}$ be the centered version of data matrix $\boldsymbol{X}$ where the means of all features are shifted to zero. Principal Component Analysis (PCA) is an unsupervised method that attempts to reduce the data $\widetilde{\boldsymbol{X}}$ to $\overline{\boldsymbol{X}} \in \mathbb{R}^{N_r \times (N_{\text{train}} \cdot m \cdot n)}$, with $N_r < N_b$ the number of components, by finding an orthogonal vector $\boldsymbol{w}$ with $||\boldsymbol{w}|| = 1$ such that the projected data $\widetilde{\boldsymbol{X}}\boldsymbol{w}$ has the highest variance. The maximization $\text{Var}(\widetilde{\boldsymbol{X}}\boldsymbol{w}) = \boldsymbol{w}^T \boldsymbol{C} \boldsymbol{w}$ yields the largest eigenvalues of the covariance matrix $\boldsymbol{C} = \text{Cov}(\widetilde{\boldsymbol{X}})$. Therefore the data matrix $\widetilde{\boldsymbol{X}}$ is multiplied by the matrix $\boldsymbol{W}$, containing the N_r largest eigenvalues of $\boldsymbol{C}$, to give $\overline{\boldsymbol{X}} = \widetilde{\boldsymbol{X}}\boldsymbol{W}$. Denote the final transformation of PCA derived from data $\widetilde{\boldsymbol{X}}$ to N_r components by $T_{\boldsymbol{X}}^{\text{PCA}_{N_r}}$. If PCA is chosen to reduce the data to N_r number of bins, then the optimization problem (3.2) becomes

$$\min_{F} \sum_{i=0}^{N^{\text{train}}} L(F(T_{\boldsymbol{X}}^{\text{PCA}_{N_r}}(x_i^{\text{train}})), y_i^{\text{train}})$$

B.1.2 Non-Negative Matrix Factorization

Let $\boldsymbol{X}^* = \boldsymbol{X} - \min(\boldsymbol{X})$ be the nonnegative matrix version of $\boldsymbol{X}$. In Non-Negative Matrix Factorization (NMF) an attempt is made to factorize the non-negative data matrix $\boldsymbol{X}^*$ into two matrices $\boldsymbol{W} \in \mathbb{R}^{N_b \times N_r}$ and $\boldsymbol{H} \in \mathbb{R}^{N_r \times (N_{\text{train}} \cdot m \cdot n)}$ in an unsupervised manner such that $\boldsymbol{X}^* = \boldsymbol{W}\boldsymbol{H}$. The matrix $\boldsymbol{H}$ will then contain the data points compressed to N_r bins, while $\boldsymbol{W}$ describes the transformation of this matrix to recover the original data matrix $\boldsymbol{X}^*$. Since the problem is not solvable in general, the matrices $\boldsymbol{W}$ and $\boldsymbol{H}$ are often approximated numerically by solving the minimization problem:

$$\min_{\boldsymbol{W},\boldsymbol{H}} ||\boldsymbol{X}^* - \boldsymbol{W}\boldsymbol{H}||_N^2$$

where $||\cdot||_N$ is usually the Frobenius norm. Denote the transformation of NMF derived from data $\boldsymbol{X}$ to N_r components by $T_{\boldsymbol{X}}^{\text{NMF}_{N_r}}$. Similar to PCA, if NMF is chosen to reduce the data to N_r number of bins, the optimization problem (3.2) becomes

$$\min_F \sum_{i=0}^{N^{\text{train}}} L(F(T_{\boldsymbol{X}}^{\text{NMF}_{N_r}}(x_i^{\text{train}})), y_i^{\text{train}})$$

B.1.3 Linear Discriminant Analysis

Let $\boldsymbol{Y} \in \mathcal{C}^{N_{\text{train}} \cdot m \cdot n}$ be the vector representation of $\{y_i\}_{i=1}^{N_{\text{train}}}$. Linear Discriminant Analysis (LDA) seeks to find a transformation W of the data such that ratio of the between-class scatter matrix $S_B(\boldsymbol{X}, \boldsymbol{Y})$ and within-class scatter matrix $S_w(\boldsymbol{X}, \boldsymbol{Y})$ is minimized:

$$\min_{\boldsymbol{W}} \frac{|\boldsymbol{W}^T S_B(\boldsymbol{X}, \boldsymbol{Y}) \boldsymbol{W}|}{|\boldsymbol{W}^T S_w(\boldsymbol{X}, \boldsymbol{Y}) \boldsymbol{W}|}$$

Intuitively, the data are projected on a lower-dimensional space that maximally separates the means of the projected class data points, while minimizing the variances within each class. Similar to PCA, this leads to an eigenvalue problem. Note that since the rank of between-class scatter matrix is at most $C - 1$, where $C = |\mathcal{C}|$ is the number of different classes in the target data $\boldsymbol{Y}$, the rank of $\boldsymbol{W}$ is at most $C - 1$ as well. This means that LDA can reduce the data to at most $N_r = |\mathcal{C}| - 1$ bins. Denote the transformation of NMF derived from data $\boldsymbol{X}$ to N_r components by $T_{\boldsymbol{X},\boldsymbol{Y}}^{\text{LDA}_{N_r}}$. Similar to the previous methods, if LDA is chosen to reduce the data to $N_r < C$ number of bins, then the optimization problem (3.2) becomes

$$\min_F \sum_{i=0}^{N^{\text{train}}} L(F(T_{\boldsymbol{X},\boldsymbol{Y}}^{\text{LDA}_{N_r}}(x_i^{\text{train}})), y_i^{\text{train}})$$

B.2 X-ray projection data computation

In this appendix, we provide further details on the computation of the simulated X-ray projections. The dataset consists of 100 2D images of size 512×512 with $N_b = 300$ spectral bins. These are simulated X-ray projections of 3D volumes of $1024 \times 1024 \times 1024$ voxels containing 120 cylinders with randomized lengths (uniformly distributed between 0.143 and 1.43 cm), thicknesses (uniformly distributed between 0.044 and 0.11 cm), angles and positions. For a schematic overview of the simulated X-ray setup, we refer to Figure 3.6. A virtual source and a virtual detector of size 1536×1536 are placed in front and behind the object respectively, and we use the ASTRA toolbox to compute the projections from this geometric setup. After this, we downscale the projections to 512×512 for computational efficiency, effectively rescaling the volume size as well. The detector pixel size is chosen to

be $s_{\text{pixel}} = 0.11$ mm, making the detector about 5.6 cm, while the voxel size is chosen to be $s_{\text{voxel}} = 0.11$ mm, making the object size about 3.75 cm. A cone beam geometry is used, where the source is placed 44 cm in front of the the center of the object, while the detector is placed 11 cm behind it. We use the National Institute for Standards and Technology (NIST) [130, 301] attenuation spectra for each associated material to compute for each ray an approximation of the number of photons in energy bin $I(E_i)$ hitting the detector. The computed quantity for each bin i with energy window E_i and $1 \leq i \leq N_b$ is given by the following:

$$I(E_i) = \int_{E_i^{\min}}^{E_i^{\max}} I_0(E) e^{-\int_\ell \mu(x,E)dx} dE \tag{B.1}$$

Here, $E_i^{\min}$ and $E_i^{\max}$ signify the energy range in bin i, $I_0(E)$ photon influx at energy E, ℓ is the ray trajectory and $\mu(x, E)$ is the attenuation at position x at energy E. This is approximated by inserting the assumption that $\mu(x, E)$ can be written as a linear combination of individual material attenuations:

$$\begin{aligned}\mu(x,E) &= \sum_{m\in\mathcal{M}} \mu_m(E)\ \alpha_m(x) \\ &= \sum_{m\in\mathcal{M}} (0.01\overline{\mu}_m(E) + 0.99\mu_{\text{polyethylene}}(E))\,\alpha_m(x)\end{aligned}$$

where $\overline{\mu}_m(E)$ is the attenuation coefficient of material $m \in \mathcal{M}$, with $\mathcal{M}$ being the set of involved materials, and $\alpha_m(x)$ the fraction of material m at position x. Inserting this into (B.1) gives:

$$\begin{aligned}I(E_i) &= \int_{E_i^{\min}}^{E_i^{\max}} I_0(E) e^{-\int_\ell \sum_{m\in\mathcal{M}} \mu_m(E)\ \alpha_m(x)dx} dE \\ &= \int_{E_i^{\min}}^{E_i^{\max}} I_0(E) e^{-\sum_{m\in\mathcal{M}} \mu_m(E) \int_\ell \alpha_m(x)dx} dE\end{aligned}$$

The integral is numerically approximated using the midpoint rule and equally sized integration bins, which gives the following:

$$I(E_i) \approx \sum_{j=1}^{N} I_0(\widetilde{E}_j) e^{-\sum_{m\in\mathcal{M}} \mu_m(\widetilde{E}_j) \int_\ell \alpha_m(x)dx} (E_{i_{\max}} - E_{i_{\min}})$$

where N is the number of integration bins, and $\widetilde{E}_j = E_{i_{\min}} + \frac{2(j-1)+1}{2n}(E_{i_{\min}} + E_{i_{\max}})$ the average energy in the j-th integration bin. The number of integration bins is set to $N = 30$ for this computation. The integral $\int_\ell \alpha_m(x)dx$ is computed using ASTRA.

The photon influx $I_0(E)$ is a product of the source spectrum $\overline{I}_0(E)$ at energy E, the exposure time t and the detector pixel size s_{pixel}:

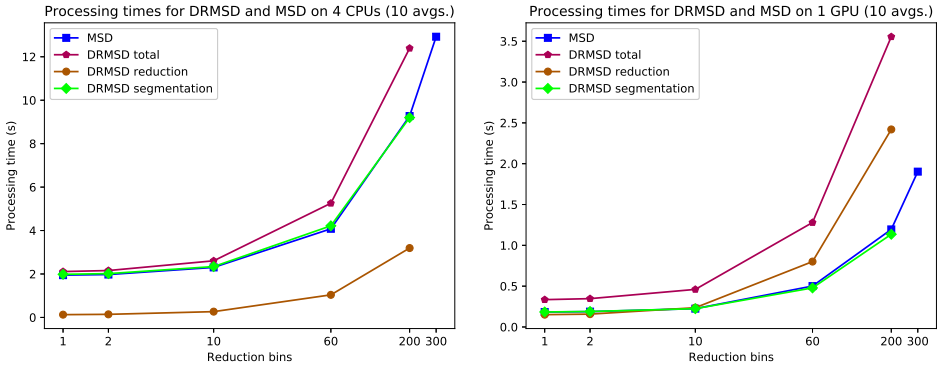
$$I_0(E) = t\overline{I}_0(E)s_{\text{pixel}}^2$$

The exposure time is chosen to be $t = 0.5$ s, and the source spectrum $\overline{I}_0$ is simulated as a radiology source spectrum for a tungsten source without filter at 70 kV, taken from Siemens Healhtineers [258]. The energy range used for this dataset is from $E_{1\min} = 14$ kV to $E_{N_{b_{\max}}} = 69$ kV, and the source spectrum including this range is given in Figure 3.8b. The final projection images in bin i are computed by dividing $I(E_i)$ by the flatfield image $I_{\text{flat}}(E_i)$ containing reference photon counts without objects

$$\frac{I(E_i)}{I_{\text{flat}}(E_i)} = \frac{I(E_i)}{\sum_{j=1}^{N} I_0(\widetilde{E}_j)(E_{i_{\max}} - E_{i_{\min}})}$$

B.3 Time comparison

In this appendix we show the measured processing time for different training setups with MSD on the generated X-ray dataset. Along with the GPU times we also include CPU times, where training is carried out on one Xeon CPU core. The processing times of the trained networks are given in Figure A1. The times for DRMSD are broken down into the data reduction part and the segmentation part. Of course, the times on the CPU cores are higher than those on the GPU core. In both cases the processing time of DRMSD reducing to 1 bin is about 7 to 8 times faster than that of MSD without any data reduction. On the CPUs the DRMSD processing time is comparable to that of MSD, with the difference increasing as the number of bins N_r increases. Note that the number of connections in both networks increase linearly with the number of reduction images N_r. When reducing



(**a**) Processing times on 4 CPUs

(**b**) Processing times on 1 GPU

Figure A1: Execution time to apply a forward pass in the trained networks on both CPUs (a) and GPUs (b). The MSD times (in blue) are added as a reference, where each data point indicates the number of input channels. These values are equivalent to the processing time on data reduced by standard methods as PCA, NMF and LDA.

the data to up to 60 bins on the GPU, the DRMSD network is less than 2 times as slow as the network on PCA, NMF and LDA reduced data. For the segmentation part the speedup of the GPU versus CPU is 87%, whereas the speedup for the data reduction part is 24%. Therefore, on the CPU the differences are smaller, but for both CPU and GPU the additional data reduction processing time is acceptable. All in all, the data show that the DRMSD can offer a processing speedup compared to MSD when accomplishing hyperspectral imaging tasks, and this conclusion could hold for some other CNN architectures as well.

B.4 Robustness

Since all experiments in this work are not averaged over multiple runs due time computation time restrictions, we assess in this appendix the stability and robustness of a number of selected experiments. Included in this selection are the experiments where we witnessed the largest variation in the test results. We compute the average, standard deviation, minimum, maximum and median values of the average class accuracy over 8 different runs. The outcomes are given in Table B.1. For

Dataset	Data type	Red. type	Red. chan.	Avg.	Std.	Min.	Max.	Median
X-ray	Noisy + Many materials	DRMSD	2	99.30	0.0883	99.13	99.42	99.31
X-ray	Noisy + Many materials	DRUNet	2	98.87	0.2937	98.36	99.29	98.97
X-ray	Noisy + Many materials	LDA + MSD	2	94.09	0.7639	92.76	95.48	93.91
X-ray	Noisy + Many materials	LDA + U-Net	2	87.24	0.6757	86.10	88.21	87.36
Remote sensing	Noisy + Overlapping	DRMSD	1	95.85	0.3642	95.28	96.34	95.86
Remote sensing	Noisy + Overlapping	DRUNet	1	94.46	0.7250	93.27	95.58	94.65
Remote sensing	Noisy + Overlapping	LDA + MSD	1	59.34	1.2328	56.27	60.60	59.78
Remote sensing	Noisy + Overlapping	LDA + U-Net	1	61.11	0.7674	60.03	62.48	61.05

Table B.1: Average, standard deviation, minimum, maximum and median of the average class accuracy for various network setups, computed over 8 runs.

each experiment, the standard deviation is at most 1.24, and for DRCNN methods this is 0.73. The difference between the minimum and maximum values is at most 3.51, and for DRMSD methods this is 1.38. From these results, we conclude that all the methods presented here are sufficiently stable, and these stability properties may be expected from the other experiments in this research as well.

C
Appendices to Chapter 5

C.1 Proof of Theorem 1

Since the convex set $\mathcal{C} = \left\{\boldsymbol{X} \in \mathbb{R}^{M\times D} \,|\, \boldsymbol{X} \geq 0,\, \boldsymbol{X}\mathbf{1} \leq \mathbf{1}, \boldsymbol{X}^T\mathbf{1} \leq \mathbf{1}\right\}$ is composed of convex set $\mathcal{C}_1 = \left\{\boldsymbol{X} \in \mathbb{R}^{M\times D} \,|\, \boldsymbol{X} \geq 0,\, \boldsymbol{X}\mathbf{1} \leq \mathbf{1}\right\}$ and convex set $\mathcal{C}_2 = \left\{\boldsymbol{X} \in \mathbb{R}^{M\times D} \,|\, \boldsymbol{X} \geq 0,\, \boldsymbol{X}^T\mathbf{1} \leq \mathbf{1}\right\}$, the indicator function $\delta_{\mathcal{C}}$ (with $\delta_{\mathcal{C}}(\boldsymbol{X}) = 0$ when $\boldsymbol{X} \in \mathcal{C}$ and $\delta_{\mathcal{C}}(\boldsymbol{X}) = \infty$ otherwise) can be expressed as

$$\delta_{\mathcal{C}}(\boldsymbol{X}) = \delta_{\mathcal{C}_1}(\boldsymbol{X}) + \delta_{\mathcal{C}_2}(\boldsymbol{X}).$$

Hence, the projection onto set $\mathcal{C}$ amounts to solving the following minimization problem

$$\mathbf{proj}_{\mathcal{C}}(\boldsymbol{Z}) = \arg\min_{\boldsymbol{X}} \left\{ \frac{1}{2}\|\boldsymbol{X} - \boldsymbol{Z}\|_F^2 + \delta_{\mathcal{C}_1}(\boldsymbol{X}) + \delta_{\mathcal{C}_2}(\boldsymbol{X}) \right\}.$$

Since the cost function is the composition of two indicator functions, we can redefine a minimization problem by introducing a new slack variable $\boldsymbol{Y}$:

$$\underset{\boldsymbol{X},\boldsymbol{Y}}{\text{minimize}} \left\{ \frac{1}{2}\|\boldsymbol{X} - \boldsymbol{Z}\|_F^2 + \delta_{\mathcal{C}_1}(\boldsymbol{X}) + \delta_{\mathcal{C}_2}(\boldsymbol{Y}) + \frac{1}{2}\|\boldsymbol{X} - \boldsymbol{Y}\|_F^2 \right\},$$

where we have penalized the slack variable $\boldsymbol{Y}$ to stay close to the original variable $\boldsymbol{X}$ using quadratic term. The optimal point of this minimization problem must satisfy the following fixed point equation:

$$\begin{aligned} \boldsymbol{X} - \boldsymbol{Z} + \partial\delta_{\mathcal{C}_1}(\boldsymbol{X}) + \boldsymbol{X} - \boldsymbol{Y} &\in \mathbf{0}, \\ \partial\delta_{\mathcal{C}_2}(\boldsymbol{Y}) + \boldsymbol{Y} - \boldsymbol{X} &\in \mathbf{0}, \end{aligned}$$

where ∂f denotes the sub-gradient of the function f. Hence, the fixed point iteration scheme to find the optimal point leads to

$$\begin{aligned} \left(\boldsymbol{I} + (1/2)\partial\delta_{\mathcal{C}_1}\right)\boldsymbol{X}_{t+1} &= \frac{\boldsymbol{Z} + \boldsymbol{Y}_t}{2}, \\ \left(\boldsymbol{I} + \partial\delta_{\mathcal{C}_2}\right)\boldsymbol{Y}_{t+1} &= \boldsymbol{X}_{t+1}, \end{aligned}$$

for $t = 1, \ldots, T$ with setting $\boldsymbol{Y}_0$ to an arbitrary vector. Since the operation $(\boldsymbol{I} + \alpha\partial\delta_{\mathcal{C}})^{-1}$ with $\alpha > 0$ is equivalent to the definition of proximal operator, we can compactly rewrite the iteration scheme as

$$\begin{aligned} \boldsymbol{Y}_{t+1} &= \left(\boldsymbol{I} + \partial\delta_{\mathcal{C}_2}\right)^{-1}\left(\left(\boldsymbol{I} + (1/2)\partial\delta_{\mathcal{C}_1}\right)^{-1}\left(\frac{\boldsymbol{Z} + \boldsymbol{Y}_t}{2}\right)\right), \\ &= \mathbf{proj}_{\mathcal{C}_2}\left(\mathbf{proj}_{\mathcal{C}_1}\left(\frac{\boldsymbol{Z} + \boldsymbol{Y}_t}{2}\right)\right). \end{aligned}$$

□

C.2 Bi-convexity of ADJUST and partial optimality

In this section, we show that the optimization problem (5.7) is bi-convex. We start with the definitions related to bi-convexity.

Definition 1 (Bi-convex set)**.** *A set $\mathcal{B} \subset \mathcal{X} \times \mathcal{Y}$ is bi-convex on $\mathcal{X} \times \mathcal{Y}$ if $\mathcal{B}_x = \{y \in \mathcal{Y} : (x, y) \in \mathcal{B}\}$ is convex for every $x \in \mathcal{X}$ and $\mathcal{B}_y = \{x \in \mathcal{X} : (x, y) \in \mathcal{B}\}$ is convex for every $y \in \mathcal{Y}$.*

Definition 2 (Bi-convex function)**.** *A function $\mathcal{F} : \mathcal{B} \to \mathbb{R}$ on a bi-convex set $\mathcal{B} \subseteq \mathcal{X} \times \mathcal{Y}$ is bi-convex if and only if for every fixed y, the function $\mathcal{F}(x, \cdot) : \mathcal{B}_x \to \mathbb{R}$ is convex on $\mathcal{B}_x$, and for every fixed x, the function $\mathcal{F}(\cdot, y) : \mathcal{B}_y \to \mathbb{R}$ is convex on $\mathcal{B}_y$.*

Definition 3 (Bi-convex optimization problem)**.** *A minimization problem of the form*

$$\underset{x,y}{\text{minimize}} \quad \mathcal{F}(x, y) \quad \text{subject to} \quad x, y \in \mathcal{B}$$

is bi-convex if the set $\mathcal{B}$ is bi-convex on $\mathcal{X} \times \mathcal{Y}$ and the objective function $\mathcal{F}$ is bi-convex on $\mathcal{B}$.

Therefore, to show bi-convexity of problem (5.7), we need to show that the constraint set $\mathcal{C}_A \times \mathcal{C}_R$ is bi-convex on $\mathbb{R}^{N\times M} \times \mathbb{R}^{M\times D}$, and the function $\mathcal{J} : \mathbb{R}^{N\times M} \times \mathbb{R}^{M\times D} \to \mathbb{R}$ is a bi-convex function.

Lemma 1. *The set $\mathcal{B} \triangleq \mathcal{C}_A \times \mathcal{C}_R$ is bi-convex on $\mathbb{R}^{N\times M} \times \mathbb{R}^{M\times D}$.*

Proof. Since the set $\mathcal{B}$ is partitioned into two independent sets $\mathcal{C}_A$ and $\mathcal{C}_R$, we only need to show that these sets are convex. The set

$$\mathcal{C}_A = \left\{ \boldsymbol{X} \in \mathbb{R}^{N\times M} \,|\, x_{ij} \geq 0, \sum_{j=1}^{M} x_{ij} \leq 1 \right\}$$

is a convex set on $\mathbb{R}^{N\times M}$ since it is an intersection of the non-negative orthant ($x_{ij} \geq 0$) with N number of hyperplanes ($\sum_{j=1}^{M} x_{ij} \leq 1$) (see 2.2.4 of [45]). Similarly, the set

$$\mathcal{C}_R = \left\{ \boldsymbol{X} \in \mathbb{R}^{M\times D} \,|\, x_{ij} \geq 0, \sum_{j=1}^{D} x_{ij} = 1, \sum_{i=1}^{M} x_{ij} \leq 1 \right\},$$

is a convex set on $\mathbb{R}^{M\times D}$ because it is an intersection of non-negative orthant ($x_{ij} \geq 0$) with M number of hyperplanes ($\sum_{j=1}^{D} x_{ij} = 1$) and D number of halfspaces ($\sum_{i=1}^{M} x_{ij} \leq 1$). Hence, from definition 1, the set $\mathcal{B} = \mathcal{C}_A \times \mathcal{C}_R$ is a bi-convex set on $\mathbb{R}^{N\times M} \times \mathbb{R}^{M\times D}$. □

Lemma 2. *The function $\mathcal{J}(\boldsymbol{A}, \boldsymbol{R}) = \frac{1}{2}\|\boldsymbol{Y} - \boldsymbol{WART}\|_F^2$ is bi-convex.*

Proof. First, we rewrite the function in the form

$$\begin{aligned}\mathcal{J}(\boldsymbol{A},\boldsymbol{R}) &= \tfrac{1}{2}\|\boldsymbol{Y}-\boldsymbol{WART}\|_F^2,\\ &= \tfrac{1}{2}\operatorname{Tr}\big((\boldsymbol{Y}-\boldsymbol{WART})(\boldsymbol{Y}-\boldsymbol{WART})^T\big) \qquad \|\boldsymbol{X}\|_F^2 = \operatorname{Tr}\big(\boldsymbol{XX}^T\big),\\ &= \tfrac{1}{2}\underbrace{\operatorname{Tr}\left(\boldsymbol{T}^T\boldsymbol{R}^T\boldsymbol{A}^T\boldsymbol{W}^T\boldsymbol{WART}\right)}_{\mathcal{P}(\boldsymbol{A},\boldsymbol{R})} - \underbrace{\operatorname{Tr}\left(\boldsymbol{Y}^T\boldsymbol{WART}\right)}_{\mathcal{Q}(\boldsymbol{A},\boldsymbol{R})} + \tfrac{1}{2}\|\boldsymbol{Y}\|_F^2.\end{aligned}$$

Hence, to show that $\mathcal{J}(\boldsymbol{A},\boldsymbol{R})$ is bi-convex, we need to show that $\mathcal{P}(\boldsymbol{A},\boldsymbol{R})$ and $\mathcal{Q}(\boldsymbol{A},\boldsymbol{R})$ are bi-convex.

We first show the bi-convexity of $\mathcal{Q}(\boldsymbol{A},\boldsymbol{R})$. To do so, fix $\overline{\boldsymbol{A}} \in \mathcal{C}_A$. Now, let $\boldsymbol{R}_1, \boldsymbol{R}_2 \in \mathcal{C}_R$ and $\lambda \in (0,1)$. Then we have

$$\begin{aligned}\lambda\mathcal{Q}(\overline{\boldsymbol{A}},\boldsymbol{R}_1) + (1-\lambda)\mathcal{Q}(\overline{\boldsymbol{A}},\boldsymbol{R}_2) &= \lambda\operatorname{Tr}\big(\boldsymbol{Y}^T\boldsymbol{W}\overline{\boldsymbol{A}}\boldsymbol{R}_1\boldsymbol{T}\big) + (1-\lambda)\operatorname{Tr}\big(\boldsymbol{Y}^T\boldsymbol{W}\overline{\boldsymbol{A}}\boldsymbol{R}_2\boldsymbol{T}\big)\\ &= \operatorname{Tr}\big(\lambda\boldsymbol{Y}^T\boldsymbol{W}\overline{\boldsymbol{A}}\boldsymbol{R}_1\boldsymbol{T}\big) + \operatorname{Tr}\big((1-\lambda)\boldsymbol{Y}^T\boldsymbol{W}\overline{\boldsymbol{A}}\boldsymbol{R}_2\boldsymbol{T}\big)\\ &= \operatorname{Tr}\big(\lambda\boldsymbol{Y}^T\boldsymbol{W}\overline{\boldsymbol{A}}\boldsymbol{R}_1\boldsymbol{T} + (1-\lambda)\boldsymbol{Y}^T\boldsymbol{W}\overline{\boldsymbol{A}}\boldsymbol{R}_2\boldsymbol{T}\big)\\ &= \operatorname{Tr}\big(\boldsymbol{Y}^T\boldsymbol{W}\overline{\boldsymbol{A}}(\lambda\boldsymbol{R}_1 + (1-\lambda)\boldsymbol{R}_2)\boldsymbol{T}\big)\\ &= \mathcal{Q}(\overline{\boldsymbol{A}}, \lambda\boldsymbol{R}_1 + (1-\lambda)\boldsymbol{R}_2)\end{aligned}$$

Hence, $\mathcal{Q}(\overline{\boldsymbol{A}},\boldsymbol{R})$ is a convex function over $\mathbb{R}^{M\times D}$ for every $\boldsymbol{A} \in \mathcal{C}_A$. Similarly, fixing $\overline{\boldsymbol{R}} \in \mathcal{C}_R$ and using an analogous deduction as above shows that

$$\lambda\mathcal{Q}(\boldsymbol{A}_1,\overline{\boldsymbol{R}}) + (1-\lambda)\mathcal{Q}(\boldsymbol{A}_2,\overline{\boldsymbol{R}}) = \mathcal{Q}(\lambda\boldsymbol{A}_1 + (1-\lambda)\boldsymbol{A}_2, \overline{\boldsymbol{R}})$$

for every $\boldsymbol{A}_1, \boldsymbol{A}_2 \in \mathcal{C}_A$ and $\lambda \in (0,1)$. Hence, $\mathcal{Q}(\boldsymbol{A},\overline{\boldsymbol{R}})$ is a convex function over $\mathbb{R}^{N\times M}$ for every $\boldsymbol{R} \in \mathcal{C}_R$. This shows that $\mathcal{Q}(\boldsymbol{A},\boldsymbol{R})$ is bi-convex.

Next, we show the bi-convexity of $\mathcal{P}(\boldsymbol{A},\boldsymbol{R})$. Thus, fix $\overline{\boldsymbol{A}} \in \mathcal{C}_A$. Now, to show that $\mathcal{P}(\overline{\boldsymbol{A}},\boldsymbol{R})$ is convex, we use the first-order condition (see 3.1.4 of [45]). Let $\boldsymbol{Q} = \overline{\boldsymbol{A}}^T\boldsymbol{W}^T\boldsymbol{W}\overline{\boldsymbol{A}}$ and $\boldsymbol{P} = \boldsymbol{TT}^T$. The first-order condition states that $\forall\, \boldsymbol{R}_1, \boldsymbol{R}_2 \in \mathbb{R}^{M\times D}$, we need

$$\begin{aligned}\mathcal{P}(\overline{\boldsymbol{A}},\boldsymbol{R}_2) &\geq \mathcal{P}(\overline{\boldsymbol{A}},\boldsymbol{R}_1) + \operatorname{Tr}\left((\boldsymbol{R}_2-\boldsymbol{R}_1)^T\nabla_{\boldsymbol{R}_1}\mathcal{P}(\overline{\boldsymbol{A}},\boldsymbol{R}_1)\right)\\ \operatorname{Tr}\big(\boldsymbol{R}_2^T\boldsymbol{QR}_2\boldsymbol{P}\big) &\geq \operatorname{Tr}\big(\boldsymbol{R}_1^T\boldsymbol{QR}_1\boldsymbol{P}\big) + 2\operatorname{Tr}\left((\boldsymbol{R}_2-\boldsymbol{R}_1)^T\overline{\boldsymbol{A}}^T\boldsymbol{W}^T\boldsymbol{W}\overline{\boldsymbol{A}}\boldsymbol{R}_1\boldsymbol{TT}^T\right)\\ \operatorname{Tr}\big(\boldsymbol{R}_2^T\boldsymbol{QR}_2\boldsymbol{P}\big) &\geq \operatorname{Tr}\big(\boldsymbol{R}_1^T\boldsymbol{QR}_1\boldsymbol{P}\big) + 2\operatorname{Tr}\left((\boldsymbol{R}_2-\boldsymbol{R}_1)^T\boldsymbol{QR}_1\boldsymbol{P}\right)\end{aligned}$$

To arrive at this condition, let us consider

$$\begin{aligned}&\operatorname{Tr}\left((\boldsymbol{R}_1-\boldsymbol{R}_2)^T\boldsymbol{Q}(\boldsymbol{R}_1-\boldsymbol{R}_2)\boldsymbol{P}\right)\\ &= \operatorname{Tr}\big(\boldsymbol{R}_1^T\boldsymbol{QR}_1\boldsymbol{P}\big) + \operatorname{Tr}\big(\boldsymbol{R}_2^T\boldsymbol{QR}_2\boldsymbol{P}\big) - \operatorname{Tr}\big(\boldsymbol{R}_1^T\boldsymbol{QR}_2\boldsymbol{P}\big) - \operatorname{Tr}\big(\boldsymbol{R}_2^T\boldsymbol{QR}_1\boldsymbol{P}\big)\\ &= \operatorname{Tr}\big(\boldsymbol{R}_1^T\boldsymbol{QR}_1\boldsymbol{P}\big) + \operatorname{Tr}\big(\boldsymbol{R}_2^T\boldsymbol{QR}_2\boldsymbol{P}\big) - \operatorname{Tr}\big(\boldsymbol{R}_1^T\boldsymbol{QR}_2\boldsymbol{P}\big) - \operatorname{Tr}\big(\boldsymbol{R}_2^T\boldsymbol{Q}^T\boldsymbol{R}_1\boldsymbol{P}^T\big)\\ &= \operatorname{Tr}\big(\boldsymbol{R}_1^T\boldsymbol{QR}_1\boldsymbol{P}\big) + \operatorname{Tr}\big(\boldsymbol{R}_2^T\boldsymbol{QR}_2\boldsymbol{P}\big) - 2\operatorname{Tr}\big(\boldsymbol{R}_1^T\boldsymbol{QR}_2\boldsymbol{P}\big).\end{aligned}$$

Since $\boldsymbol{Q}$ and $\boldsymbol{P}$ are positive semi-definite matrices, we have

$$\mathrm{Tr}\left((\boldsymbol{R}_1 - \boldsymbol{R}_2)^T \boldsymbol{Q}(\boldsymbol{R}_1 - \boldsymbol{R}_2)\boldsymbol{P}\right) \geq 0.$$

Hence, we obtain the following relation

$$\mathrm{Tr}\big(\boldsymbol{R}_1^T \boldsymbol{Q}\boldsymbol{R}_1\boldsymbol{P}\big) + \mathrm{Tr}\big(\boldsymbol{R}_2^T \boldsymbol{Q}\boldsymbol{R}_2\boldsymbol{P}\big) \geq 2\,\mathrm{Tr}\big(\boldsymbol{R}_1^T \boldsymbol{Q}\boldsymbol{R}_2\boldsymbol{P}\big),$$

which proves the first-order condition. Similarly, we can show that $\mathcal{P}(\boldsymbol{A}, \overline{\boldsymbol{R}})$ is a convex function over $\mathbb{R}^{N\times M}$ for fixed $\overline{\boldsymbol{R}} \in \mathcal{C}_R$. Hence, $\mathcal{P}(\boldsymbol{A}, \boldsymbol{R})$ is a bi-convex function.

Since $\mathcal{P}(\boldsymbol{A}, \boldsymbol{R})$ and $\mathcal{Q}(\boldsymbol{A}, \boldsymbol{R})$ are bi-convex functions, their linear combination is also a bi-convex function [103]. Hence, we prove that $\mathcal{J}(\boldsymbol{A}, \boldsymbol{R})$ is bi-convex. □

Corollary 1. *The optimization problem (5.7) is bi-convex.*

Proof. Since the cost function $\mathcal{J}(\boldsymbol{A}, \boldsymbol{R}) = \frac{1}{2}\|\boldsymbol{Y} - \boldsymbol{W}\boldsymbol{A}\boldsymbol{R}\boldsymbol{T}\|_F^2$ is bi-convex (Lemma 2) and $\mathcal{C}_A \times \mathcal{C}_R$ is a bi-convex set (Lemma 1), the optimization problem

$$\text{minimize} \quad \mathcal{J}(\boldsymbol{A}, \boldsymbol{R}) \quad \text{subject to} \quad \boldsymbol{A} \in \mathcal{C}_A,\ \boldsymbol{R} \in \mathcal{C}_R$$

is bi-convex (from Definition 3). □

Bi-convex optimization problems may have a large number of local minima as they are global optimization problems in general [103]. Since we are interested in finding a stationary point of (5.7), we define the notion of partial optimality.

Definition 4 (Partial optimality)**.** *Let $\mathcal{F} : \mathcal{X} \times \mathcal{Y} \mapsto \mathbb{R}$ be a given function and let $(x^\star, y^\star) \in \mathcal{X} \times \mathcal{Y}$. Then, $(x^\star, y^\star)$ is called a partial optimum of $\mathcal{F}$ on $\mathcal{X} \times \mathcal{Y}$, if*

$$\mathcal{F}(x^\star, y^\star) \leq \mathcal{F}(x, y^\star) \quad \forall\, x \in \mathcal{X} \quad \textit{and} \quad \mathcal{F}(x^\star, y^\star) \leq \mathcal{F}(x^\star, y) \quad \forall\, y \in \mathcal{Y}.$$

It is easy to show that a partial optimum $z^\star = (x^\star, y^\star)$ is also a stationary point of $\mathcal{F}$ in $\mathcal{X} \times \mathcal{Y}$ if $\mathcal{F}$ is differentiable at $z^\star$. Also, the converse is true [103]. Finally, the following theorem (adapted from [290]) connects the local optimality (*i.e.* stationary points) to the partial optimality:

Theorem 2. *Let $(\boldsymbol{A}^\star, \boldsymbol{R}^\star) \in \mathcal{C}_A \times \mathcal{C}_R$ be a partial optimum of $\mathcal{J}(\boldsymbol{A}, \boldsymbol{R}) = \frac{1}{2}\|\boldsymbol{Y} - \boldsymbol{W}\boldsymbol{A}\boldsymbol{R}\boldsymbol{T}\|_F^2$. Furthermore, let $\mathcal{U}(\boldsymbol{R}^\star)$ denote the set of all optimal solutions to (5.7) with $\boldsymbol{R} = \boldsymbol{R}^\star$ and let $\mathcal{V}(\boldsymbol{A}^\star)$ be the set of optimal solutions to (5.7) with $\boldsymbol{A} = \boldsymbol{A}^\star$. If $(\boldsymbol{A}^\star, \boldsymbol{R}^\star)$ is a local optimal solution to (5.7), then it necessarily holds that*

$$\mathcal{J}(\boldsymbol{A}^\star, \boldsymbol{R}^\star) \leq \mathcal{J}(\boldsymbol{A}, \boldsymbol{R}) \quad \forall\, \boldsymbol{A} \in \mathcal{U}(\boldsymbol{R}^\star),\ \boldsymbol{R} \in \mathcal{V}(\boldsymbol{A}^\star).$$

This theorem implies that the natural solution of any alternating minimization algorithm will lead to a partial optimal solution. The proof of the theorem can be found in [290].

C.3 Derivation of AAPM

First, we rephrase the original ADJUST problem in the following form

$$\begin{aligned} \underset{\boldsymbol{A},\boldsymbol{R}}{\text{minimize}} \quad & \mathcal{J}(\boldsymbol{A},\boldsymbol{R}) + \delta_{\mathcal{C}_A}(\boldsymbol{A}) + \delta_{\mathcal{C}_R}(\boldsymbol{R}), \\ \text{subject to} \quad & \boldsymbol{WART} = \boldsymbol{Y}, \end{aligned}$$

where $\delta_{\mathcal{C}}$ is an extended value function for the constraint set $\mathcal{C}$ that is 0 when constraint is satisfied and ∞ otherwise. Here, we have introduced the constraints on the misfit between simulated and true measurements in the linear form. The Lagrangian for this optimization problem reads

$$\begin{aligned} \mathcal{L}(\boldsymbol{A},\boldsymbol{R},\boldsymbol{U}) &= \mathcal{J}(\boldsymbol{A},\boldsymbol{R}) + \delta_{\mathcal{C}_A}(\boldsymbol{A}) + \delta_{\mathcal{C}_R}(\boldsymbol{R}) + \langle \boldsymbol{U}, \boldsymbol{Y} - \boldsymbol{WART} \rangle \\ &= \underbrace{\mathcal{J}(\boldsymbol{A},\boldsymbol{R}) + \langle \boldsymbol{U}, \boldsymbol{Y} - \boldsymbol{WART} \rangle}_{\triangleq \widetilde{\mathcal{J}}(\boldsymbol{A},\boldsymbol{R},\boldsymbol{U})} + \delta_{\mathcal{C}_A}(\boldsymbol{A}) + \delta_{\mathcal{C}_R}(\boldsymbol{R}) \end{aligned} \tag{C.1}$$

where $\boldsymbol{U} \in \mathbb{R}^{J \times C}$ is a Lagrange multiplier for constraint $\boldsymbol{WART} = \boldsymbol{Y}$. The Lagrange multiplier $\boldsymbol{U}$ can also be thought of as a running-sum-of-error as it captures the misfit between the true measurements and simulated measurements. The goal is to find a saddle point of this Lagrangian, since the saddle point will give the optimal solution to (5.7). The saddle point of the Lagrangian is given by

$$(\boldsymbol{A}^\star, \boldsymbol{R}^\star, \boldsymbol{U}^\star) = \arg\max_{\boldsymbol{U}} \arg\min_{\boldsymbol{A},\boldsymbol{R}} \mathcal{L}(\boldsymbol{A},\boldsymbol{R},\boldsymbol{U}).$$

It is important to note that the Lagrangian is non-differentiable due to the presence of $\delta_{\mathcal{C}_A}$ and $\delta_{\mathcal{C}_R}$. Since the min-max problem cannot be solved using a simple gradient-based iterative scheme due to non-differentiability of the Lagrangian, we need to make use of proximal alternating iterative algorithm. To derive such scheme, we approximate the Lagrangian (C.1) near point $(\boldsymbol{A}_k, \boldsymbol{R}_k, \boldsymbol{U}_k)$ using the Taylor series for the differentiable function $\widetilde{\mathcal{J}}(\boldsymbol{A},\boldsymbol{R},\boldsymbol{U})$. This approximation reads

$$\begin{aligned} \mathcal{L}(\boldsymbol{A},\boldsymbol{R},\boldsymbol{U}) &\approx \widetilde{\mathcal{L}}(\boldsymbol{A},\boldsymbol{R},\boldsymbol{U}|\boldsymbol{A}_k,\boldsymbol{R}_k,\boldsymbol{U}_k) \\ &= \widetilde{\mathcal{J}}(\boldsymbol{A}_k,\boldsymbol{R}_k,\boldsymbol{U}_k) + \\ &\quad \langle \nabla_{\boldsymbol{R}} \widetilde{\mathcal{J}}(\boldsymbol{A}_k,\boldsymbol{R}_k,\boldsymbol{U}_k), \boldsymbol{R} - \boldsymbol{R}_k \rangle + 1/(2\alpha) \|\boldsymbol{R} - \boldsymbol{R}_k\|_F^2 + \\ &\quad \langle \nabla_{\boldsymbol{A}} \widetilde{\mathcal{J}}(\boldsymbol{A}_k,\boldsymbol{R}_k,\boldsymbol{U}_k), \boldsymbol{A} - \boldsymbol{A}_k \rangle + 1/(2\beta) \|\boldsymbol{A} - \boldsymbol{A}_k\|_F^2 + \\ &\quad \delta_{\mathcal{C}_A}(\boldsymbol{A}) + \delta_{\mathcal{C}_R}(\boldsymbol{R}), \end{aligned} \tag{C.2}$$

where α and β are the Lipschitz constants of the partial gradients of $\widetilde{\mathcal{J}}(\boldsymbol{A},\boldsymbol{R},\boldsymbol{U})$ with respect to $\boldsymbol{A}$ and $\boldsymbol{R}$ respectively. This approximation leads to the following alternating scheme where we minimize with respect to the primal variables $\boldsymbol{A}$ and

$\boldsymbol{R}$, and maximize with respect to the dual variable $\boldsymbol{U}$:

$$\begin{aligned}\boldsymbol{R}_{k+1} &= \arg\min_{\boldsymbol{R}} \widetilde{\mathcal{L}}(\boldsymbol{A}, \boldsymbol{R}, \boldsymbol{U} | \boldsymbol{A}_k, \boldsymbol{R}_k, \boldsymbol{U}_k)\\ \boldsymbol{A}_{k+1} &= \arg\min_{\boldsymbol{A}} \widetilde{\mathcal{L}}(\boldsymbol{A}, \boldsymbol{R}, \boldsymbol{U} | \boldsymbol{A}_k, \boldsymbol{R}_{k+1}, \boldsymbol{U}_k)\\ \boldsymbol{U}_{k+1} &= \boldsymbol{U}_k + \rho\left(\boldsymbol{W}\boldsymbol{A}_{k+1}\boldsymbol{R}_{k+1}\boldsymbol{T} - \boldsymbol{Y}\right)\end{aligned}$$

with $k = 0, \ldots, K$, and $\rho > 0$ is the acceleration parameter. This alternating scheme requires initial values of $\boldsymbol{R}$ and $\boldsymbol{A}$, while the initial value of $\boldsymbol{U}$ can be set to $\boldsymbol{0}$. We update the dual variable $\boldsymbol{U}$ using the linearized ascent, a standard technique used by many alternating methods, *e.g.*, alternating direction method of multipliers [46]. Since the approximate Lagrangian (C.2) is composed of quadratic term and non-smooth terms for $\boldsymbol{A}$ and $\boldsymbol{R}$, we can express the iterates using proximal operations. To derive $\boldsymbol{R}$, we use the identity $\|\boldsymbol{X}+\boldsymbol{Y}\|_F^2 = \|\boldsymbol{X}\|_F^2 + \|\boldsymbol{Y}\|_F^2 + 2\langle \boldsymbol{X}, \boldsymbol{Y}\rangle$, or equivalently, $\langle \boldsymbol{X}, \boldsymbol{Y}\rangle + \frac{1}{2}\|\boldsymbol{Y}\|_F^2 = \frac{1}{2}\|\boldsymbol{X}+\boldsymbol{Y}\|_F^2 - \frac{1}{2}\|\boldsymbol{X}\|_F^2$. The derivation is now as follows:

$$\begin{aligned}\boldsymbol{R}_{k+1} &= \arg\min_{\boldsymbol{R}} \widetilde{\mathcal{L}}(\boldsymbol{A}, \boldsymbol{R}, \boldsymbol{U} | \boldsymbol{A}_k, \boldsymbol{R}_k, \boldsymbol{U}_k),\\ &= \arg\min_{\boldsymbol{R}} \left\{ \langle \nabla_{\boldsymbol{R}} \widetilde{\mathcal{J}}(\boldsymbol{A}_k, \boldsymbol{R}_k, \boldsymbol{U}_k), \boldsymbol{R} - \boldsymbol{R}_k\rangle + \tfrac{1}{2\alpha}\|\boldsymbol{R} - \boldsymbol{R}_k\|_F^2 + \delta_{\mathcal{C}_R}(\boldsymbol{R}) \right\},\\ &= \arg\min_{\boldsymbol{R}} \left\{ \tfrac{1}{\alpha}\langle \alpha\nabla_{\boldsymbol{R}} \widetilde{\mathcal{J}}(\boldsymbol{A}_k, \boldsymbol{R}_k, \boldsymbol{U}_k), \boldsymbol{R} - \boldsymbol{R}_k\rangle + \tfrac{1}{2\alpha}\|\boldsymbol{R} - \boldsymbol{R}_k\|_F^2 + \delta_{\mathcal{C}_R}(\boldsymbol{R}) \right\},\\ &= \arg\min_{\boldsymbol{R}} \left\{ \tfrac{1}{\alpha} \underbrace{\left(\langle \alpha\nabla_{\boldsymbol{R}} \widetilde{\mathcal{J}}(\boldsymbol{A}_k, \boldsymbol{R}_k, \boldsymbol{U}_k), \boldsymbol{R} - \boldsymbol{R}_k\rangle + \tfrac{1}{2}\|\boldsymbol{R} - \boldsymbol{R}_k\|_F^2 \right)}_{\text{applying the identity with } \boldsymbol{X} \triangleq \alpha\nabla_{\boldsymbol{R}} \widetilde{\mathcal{J}}(\boldsymbol{A}_k, \boldsymbol{R}_k, \boldsymbol{U}_k),\, \boldsymbol{Y} \triangleq \boldsymbol{R} - \boldsymbol{R}_k} + \delta_{\mathcal{C}_R}(\boldsymbol{R}) \right\},\\ &= \arg\min_{\boldsymbol{R}} \left\{ \tfrac{1}{2\alpha}\|\alpha\nabla_{\boldsymbol{R}} \widetilde{\mathcal{J}}(\boldsymbol{A}_k, \boldsymbol{R}_k, \boldsymbol{U}_k) + \boldsymbol{R} - \boldsymbol{R}_k\|_F^2 \right.\\ &\qquad \left. \underbrace{-\tfrac{1}{2}\|\alpha\nabla_{\boldsymbol{R}} \widetilde{\mathcal{J}}(\boldsymbol{A}_k, \boldsymbol{R}_k, \boldsymbol{U}_k)\|_F^2}_{\text{independent of } \boldsymbol{R}} + \delta_{\mathcal{C}_R}(\boldsymbol{R}) \right\},\\ &= \arg\min_{\boldsymbol{R}} \left\{ \tfrac{1}{2\alpha}\|\boldsymbol{R} - \boldsymbol{R}_k + \alpha\nabla_{\boldsymbol{R}} \widetilde{\mathcal{J}}(\boldsymbol{A}_k, \boldsymbol{R}_k, \boldsymbol{U}_k)\|_F^2 + \delta_{\mathcal{C}_R}(\boldsymbol{R}) \right\},\\ &= \mathbf{prox}_{\delta_{\mathcal{C}_R}}\left(\boldsymbol{R}_k - \alpha\nabla_{\boldsymbol{R}} \widetilde{\mathcal{J}}(\boldsymbol{A}_k, \boldsymbol{R}_k, \boldsymbol{U}_k) \right), \end{aligned} \tag{C.3}$$

where the proximal for a function $f : \mathbb{R}^n \mapsto \mathbb{R}$ reads

$$\mathbf{prox}_{\gamma f}(\boldsymbol{z}) = \arg\min_{\boldsymbol{x} \in \mathbb{R}^n} \left\{ \frac{1}{2\gamma}\|\boldsymbol{x} - \boldsymbol{z}\|_2^2 + f(\boldsymbol{x}) \right\}$$

with $\gamma > 0$. The proximal operator allows us to work with non-differentiable functions. Moreover, proximal operators for many functions have explicit expressions,

making it a very computationally-friendly tool. The proximal operator for $\delta_{\mathcal{C}}$ with $\mathcal{C} \subset \mathbb{R}^n$ takes the following form:

$$\mathbf{prox}_{\delta_{\mathcal{C}}}(\boldsymbol{z}) = \underset{\boldsymbol{x}\in\mathbb{R}^n}{\arg\min} \left\{ \frac{1}{2}\|\boldsymbol{x}-\boldsymbol{z}\|_2^2 + \delta_{\mathcal{C}}(\boldsymbol{x}) \right\}$$

Indeed, the proximal operator of a $\delta_{\mathcal{C}}$ is just an orthogonal projection of a vector onto the set $\mathcal{C}$. If the set $\mathcal{C}$ is convex, the proximal point is unique. Similar to (C.3), we can explicitly write down the update of $\boldsymbol{A}$ in terms of the proximal operator.

C.4 Gradient computations

Here we show how the gradients are computed at the final comments in Section 5.5.1. We only show the derivation of $\nabla_{\boldsymbol{A}}\widetilde{\mathcal{J}}(\boldsymbol{A},\boldsymbol{R},\boldsymbol{U})$ since the derivation of $\nabla_{\boldsymbol{R}}\widetilde{\mathcal{J}}(\boldsymbol{A},\boldsymbol{R},\boldsymbol{U})$ is very similar.

$$\begin{aligned}
&\nabla_{\boldsymbol{A}}\widetilde{\mathcal{J}}(\boldsymbol{A},\boldsymbol{R},\boldsymbol{U}) \\
&\quad= \nabla_{\boldsymbol{A}}\left(\tfrac{1}{2}\|\boldsymbol{Y}-\boldsymbol{WART}\|_F^2 + \langle \boldsymbol{U}, \boldsymbol{Y}-\boldsymbol{WART}\rangle\right) \\
&\quad= \nabla_{\boldsymbol{A}}\left(\tfrac{1}{2}\|\boldsymbol{Y}-\boldsymbol{WART}\|_F^2\right) + \nabla_{\boldsymbol{A}}\langle \boldsymbol{U}, \boldsymbol{Y}-\boldsymbol{WART}\rangle \\
&\quad\overset{*}{=} \tfrac{1}{2}\nabla_{\boldsymbol{A}}\left(\|\boldsymbol{Y}\|_F^2 + \|\boldsymbol{WART}\|_F^2 - 2\operatorname{Tr}\left(\boldsymbol{Y}^T\boldsymbol{WART}\right)\right) \\
&\qquad+ \nabla_{\boldsymbol{A}}\operatorname{Tr}\left(\boldsymbol{U}^T(\boldsymbol{Y}-\boldsymbol{WART})\right) && \langle \boldsymbol{X},\boldsymbol{Y}\rangle = \operatorname{Tr}\left(\boldsymbol{A}^T\boldsymbol{B}\right) \\
&\quad= \tfrac{1}{2}\nabla_{\boldsymbol{A}}\left(\|\boldsymbol{WART}\|_F^2\right) - \nabla_{\boldsymbol{A}}\left(\operatorname{Tr}\left(\boldsymbol{Y}^T\boldsymbol{WART}\right)\right) \\
&\qquad+ \nabla_{\boldsymbol{A}}\operatorname{Tr}\left(\boldsymbol{U}^T\boldsymbol{Y}\right) - \nabla_{\boldsymbol{A}}\operatorname{Tr}\left(\boldsymbol{U}^T\boldsymbol{WART}\right) \\
&\quad= \tfrac{1}{2}\nabla_{\boldsymbol{A}}\left(\operatorname{Tr}\left(\boldsymbol{T}^T\boldsymbol{R}^T\boldsymbol{A}^T\boldsymbol{W}^T\boldsymbol{WART}\right)\right) - \boldsymbol{W}^T\boldsymbol{YT}^T\boldsymbol{R}^T \\
&\qquad- (\boldsymbol{UTW})^T(\boldsymbol{RT})^T && \frac{\partial}{\partial \boldsymbol{X}}\operatorname{Tr}\left(\boldsymbol{AXB}\right) = \boldsymbol{A}^T\boldsymbol{B}^T \\
&\quad= \boldsymbol{W}^T\left(\boldsymbol{WART}\right)\boldsymbol{T}^T\boldsymbol{R}^T - \boldsymbol{W}^T\left(\boldsymbol{Y}\right)\boldsymbol{T}^T\boldsymbol{R}^T \\
&\qquad- \boldsymbol{W}^T\boldsymbol{UT}^T\boldsymbol{R} \\
&\quad= \boldsymbol{W}^T\left(\boldsymbol{WART}-\boldsymbol{Y}\right)\boldsymbol{T}^T\boldsymbol{R}^T - \boldsymbol{W}^T\boldsymbol{UT}^T\boldsymbol{R}.
\end{aligned}$$

In the third step (*), we use the following identity:

$$\begin{aligned}
\|\boldsymbol{X}-\boldsymbol{Y}\|_F^2 &= \operatorname{Tr}\left((\boldsymbol{X}-\boldsymbol{Y})^T(\boldsymbol{X}-\boldsymbol{Y})\right) \\
&= \operatorname{Tr}\left((\boldsymbol{X}^T-\boldsymbol{Y}^T)(\boldsymbol{X}-\boldsymbol{Y})\right) \\
&= \operatorname{Tr}\left(\boldsymbol{X}^T\boldsymbol{X} - \boldsymbol{Y}^T\boldsymbol{X} - \boldsymbol{X}^T\boldsymbol{Y} + \boldsymbol{Y}^T\boldsymbol{Y}\right) \\
&= \operatorname{Tr}\left(\boldsymbol{X}^T\boldsymbol{X}\right) - \operatorname{Tr}\left(\boldsymbol{Y}^T\boldsymbol{X}\right) - \operatorname{Tr}\left(\boldsymbol{X}^T\boldsymbol{Y}\right) \\
&\quad+ \operatorname{Tr}\left(\boldsymbol{Y}^T\boldsymbol{Y}\right) \\
&= \|\boldsymbol{X}\|_F^2 + \|\boldsymbol{Y}\|_F^2 - 2\operatorname{Tr}\left(\boldsymbol{Y}^T\boldsymbol{X}\right) && \left(\boldsymbol{Y}^T\boldsymbol{X}\right)^T = \boldsymbol{X}^T\boldsymbol{Y}
\end{aligned}$$

C.5 Dictionary matrix

In this section, we list the 42 materials that are used in the dictionary matrix $\boldsymbol{T}$ for the Disks and Shepp-Logan phantoms. The spectra are retrieved from the National Institute for Standards and Technology (NIST) [130, 301].

Mat. no.	Material name	At. no.
23	Vanadium	23
24	Chromium	24
25	Manganese	25
26	Iron	26
27	Cobalt	27
28	Nickel	28
29	Copper	29
30	Zinc	30
31	Gallium	31
32	Germanium	32
33	Arsenic	33
34	Selenium	34
35	Bromine	35
36	Krypton	36
37	Rubidium	37
38	Strontium	38
39	Yttrium	39
40	Zirconium	40
41	Niobium	41
42	Molybdenum	42
43	Technetium	43

Mat. no.	Material name	At. no.
44	Ruthenium	44
45	Rhodium	45
46	Palladium	46
47	Silver	47
48	Cadmium	48
49	Indium	49
50	Tin	50
51	Antimony	51
52	Tellurium	52
53	Iodine	53
54	Xenon	54
55	Cesium	55
56	Barium	56
57	Lanthanum	57
58	Cerium	58
59	Praseodymium	59
60	Neodymium	60
61	Promethium	61
62	Samarium	62
63	Terbium	63
64	Gadolinium	64

We plot the attenuation spectra for all dictionary elements for each bin within the selected range in Figure A1. Additionally, Figure A2 shows the spectra for a few selected materials. All of these materials have a K-edge in the considered spectral range.

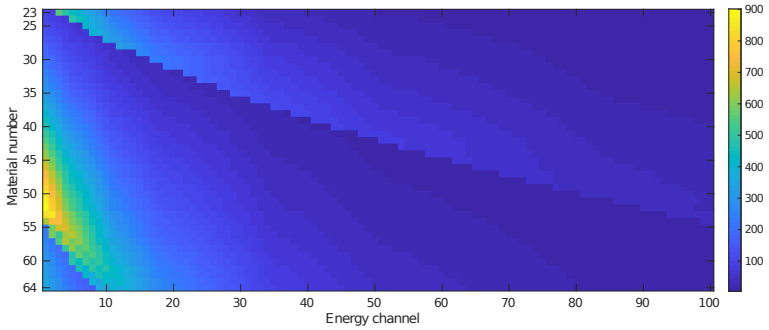


Figure A1: Dictionary matrix $\boldsymbol{T}$: Attenuation values over 100 spectral channels for 42 materials, with energies ranging from 20 keV to 119 keV.

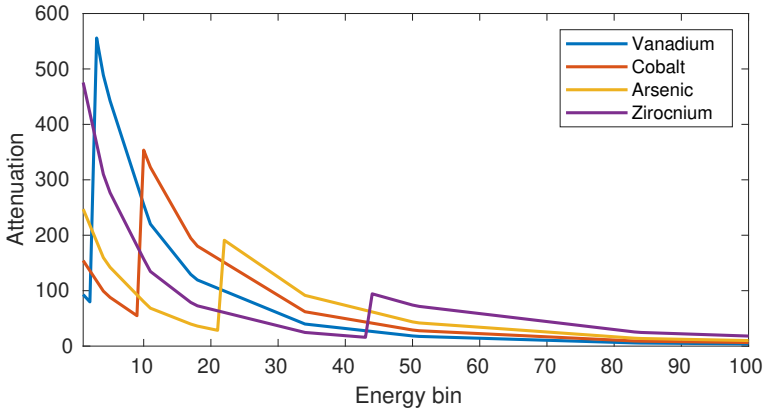


Figure A2: Attenuation values over 100 spectral channels for four materials from the dictionary matrix $\boldsymbol{T}$, with energies ranging from 5 keV to 35 keV.

C.6 Performance measures

To assess the quality of the reconstructions that ADJUST (and the comparison methods) generates, we compare the reconstructions with the ground truth. Since for the UR, RU, cJoint and ADJUST methods the best matching reconstruction of a certain channel in the ground truth may be located in a different channel in the material map matrix, a matching that minimizes the total error over the channels needs to be carried out. Let $\boldsymbol{A}^{\mathrm{GT}} \in \mathbb{R}^{N\times M}$ be the matrix containing the ground truth material maps and $\boldsymbol{A}^{\mathrm{rec}} \in \mathbb{R}^{N\times M}$ be the reconstructed material map. We compute a matrix $\boldsymbol{A}^{\mathrm{error}}$ containing the mutual errors between channels of $\boldsymbol{A}^{\mathrm{GT}}$ and $\boldsymbol{A}^{\mathrm{rec}}$, defined by

$$A_{ij}^{\mathrm{error}} = \left\| (\boldsymbol{A}_{ki}^{\mathrm{rec}})_{i\leq k\leq N} - (\boldsymbol{A}_{kj}^{\mathrm{GT}})_{1\leq k\leq N} \right\|_2$$

Given this error matrix, we use an iterative greedy approach to match the channels of the $\boldsymbol{A}^{\mathrm{GT}}$ and $\boldsymbol{A}^{\mathrm{rec}}$ matrices based on their mutual channel errors. We repeatedly compute the minimum of the error matrix and remove the possibility to match the corresponding channels. To do so, let $\mathcal{M}_0^{\mathrm{GT}} = \mathcal{M}$, $\mathcal{M}_0^{\mathrm{rec}} = \mathcal{M}$ and $\mathcal{M}_0^{\mathrm{match}} = \emptyset$. In each iteration $1 \leq l \leq M$, we compute

$$(i_l, j_l) = \underset{\substack{i \in \mathcal{M}_l^{\mathrm{rec}} \\ j \in \mathcal{M}_l^{\mathrm{GT}}}}{\arg\min} A_{ij}$$

and define $\mathcal{M}_{l+1}^{\mathrm{rec}} = \mathcal{M}_l^{\mathrm{rec}} \backslash \{i_l\}$, $\mathcal{M}_{l+1}^{\mathrm{GT}} = \mathcal{M}_l^{\mathrm{GT}} \backslash \{j_l\}$ and $\mathcal{M}_{l+1}^{\mathrm{match}} = \mathcal{M}_l^{\mathrm{match}} \cup \{(i_l, j_l\}$. Given the final channel-matching represented by $\mathcal{M}_M^{\mathrm{match}}$, we compute the following three error metrics for each $(i, j) \in \mathcal{M}_M^{\mathrm{match}}$:

- *Mean square error* (MSE) for each matched material pair:

$$\mathrm{MSE}(i,j) = \left\| (\boldsymbol{A}_{ki}^{\mathrm{rec}})_{i \leq k \leq N} - (\boldsymbol{A}_{kj}^{\mathrm{GT}})_{1 \leq k \leq N} \right\|_2^2$$

- *Peak signal-to-noise ratio* (PSNR) for each matched material pair:

$$\mathrm{PSNR}(i,j) = 10 \log_{10} \left(\left(\max_k (\boldsymbol{A}_{kj}^{\mathrm{GT}})_{1 \leq k \leq N} \right)^2 / \left\| (\boldsymbol{A}_{ki}^{\mathrm{rec}})_{i \leq k \leq N} - (\boldsymbol{A}_{kj}^{\mathrm{GT}})_{1 \leq k \leq N} \right\|_2^2 \right)$$

- *Structural similarity index* (SSIM) for each matched material pair:

$$\mathrm{SSIM}(i,j) = \left((2\mu_i \mu_j + C_1)(2\sigma_{ij} + C_2) \,/ \left(\mu_i^2 + \mu_j^2 + C_1)(\sigma_i^2 + \sigma_j^2 + C_2 \right) \right)$$

  with μ_i, μ_j and σ_i, σ_j being the means and the standard deviations of the matrices $(\boldsymbol{A}_{ki}^{\mathrm{rec}})_{i \leq k \leq N}$ and $(\boldsymbol{A}_{kj}^{\mathrm{GT}})_{1 \leq k \leq N}$ respectively, with σ_{ij} being the cross-correlation between these two matrices, and with $C_1 = (0.01L)^2$, $C_2 = (0.03L)^2$ and $L = 1$.

The averages of the MSE, PSNR and SSIM over all materials are then given by:

$$\begin{aligned}
\mathrm{MSE}_{\mathrm{avg}} &= \sum_{(i,j) \in \mathcal{M}_M^{\mathrm{match}}} \mathrm{MSE}(i,j)/M, \\
\mathrm{PSNR}_{\mathrm{avg}} &= \sum_{(i,j) \in \mathcal{M}_M^{\mathrm{match}}} \mathrm{PSNR}(i,j)/M, \\
\mathrm{SSIM}_{\mathrm{avg}} &= \sum_{(i,j) \in \mathcal{M}_M^{\mathrm{match}}} \mathrm{SSIM}(i,j)/M.
\end{aligned}$$

C.7 Numerical studies: Comparison of methods

As stated in the main chapter, we have compared ADJUST with RU, UR, and cJoint on three numerical phantoms, mainly the Shepp-Logan phantom, the Disks phantom, and the Thorax phantom. Figure A3 shows the reconstruction results (*i.e.* reconstructed spatial maps and the spectra of materials) of these methods on Disks phantom. Moreover, we also plot the performance measures of these methods per material in Figure A4.

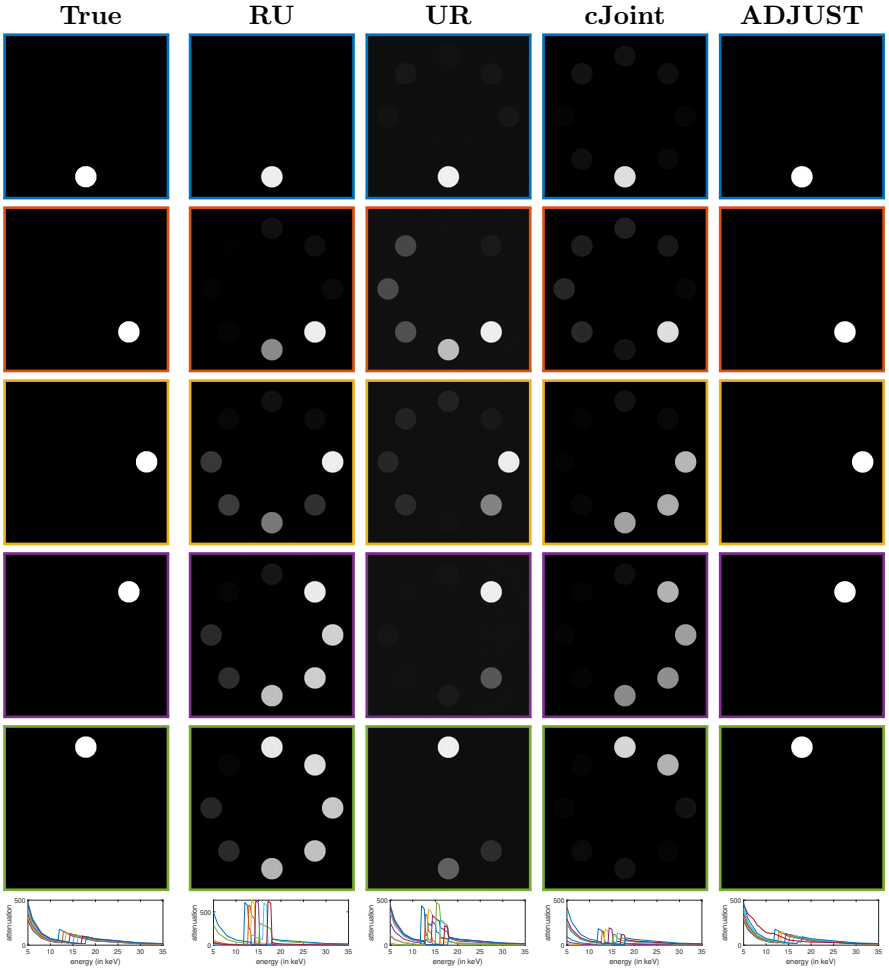


Figure A3: Visual comparison of ADJUST with RU, UR, and cJoint method on the Disks phantom. We only show the reconstructions of all disks here for the comparison. Moreover, we match the colors of the bounding box for material maps with the (recovered) spectral signatures of the materials (shown in the bottom row).

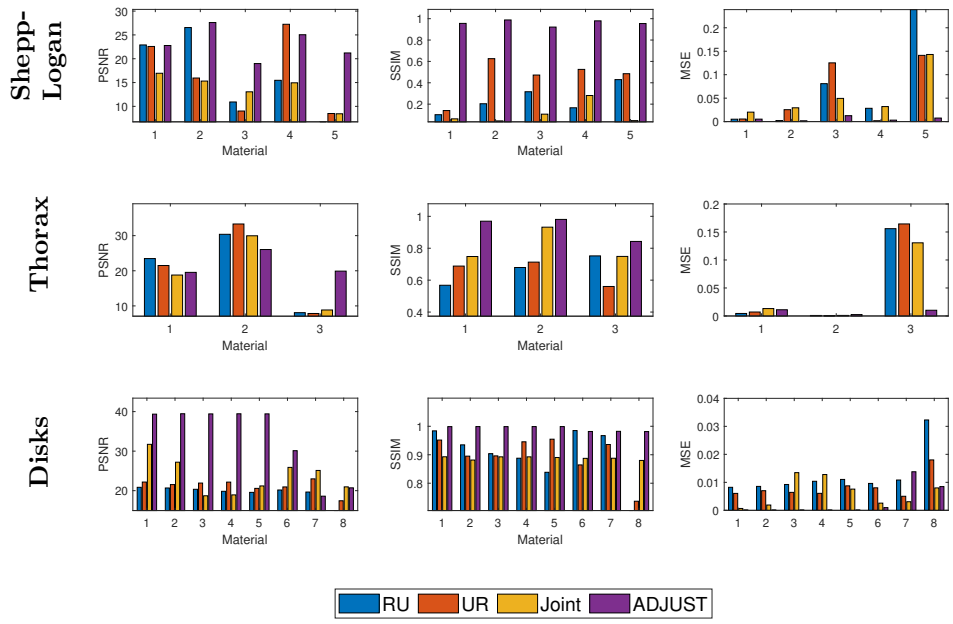


Figure A4: Performance plots showing PSNR (left column), SSIM (middle column), MSE (right column) of the reconstructed materials against the ground truth for various numerical algorithms on the phantoms.

C.8 Numerical studies: Limited measurement patterns

We consider three different types of limited measurement patterns: (*i*) *Sparse-angle*: tomographic projections from 10 equidistant angles in the range of 0 to π for 100 spectral channels, (*ii*) *Limited-view*: 60 equidistant projection angles in the limited range of $[0, 2\pi/3]$ for 100 spectral channels, (*iii*) *Sparse channels*: 60 equidistant angles between 0 and π, but with only 30 spectral channels. We test ADJUST on the two numerical spectral phantoms, *i.e.* the Shepp-Logan phantom and the Disks phantom. Figures A5 and A6 demonstrate the reconstructions of ADJUST for all three limited measurement patterns on these two phantoms.

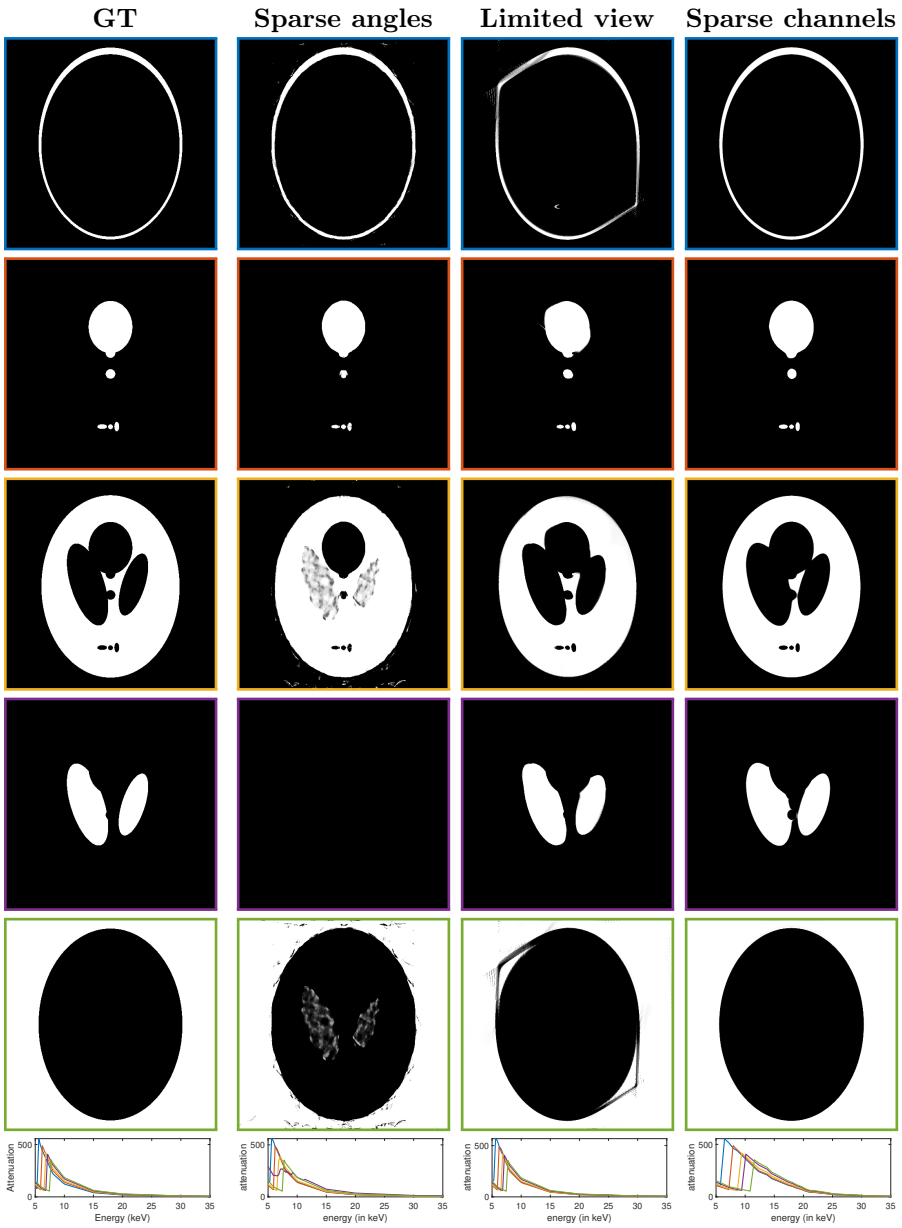


Figure A5: Results of ADJUST with sparse-angle data, limited view data and sparse spectral channels on Shepp-Logan phantom. The colors of the bounding box of material maps are matched with the spectral signatures of the materials (shown in the bottom row).

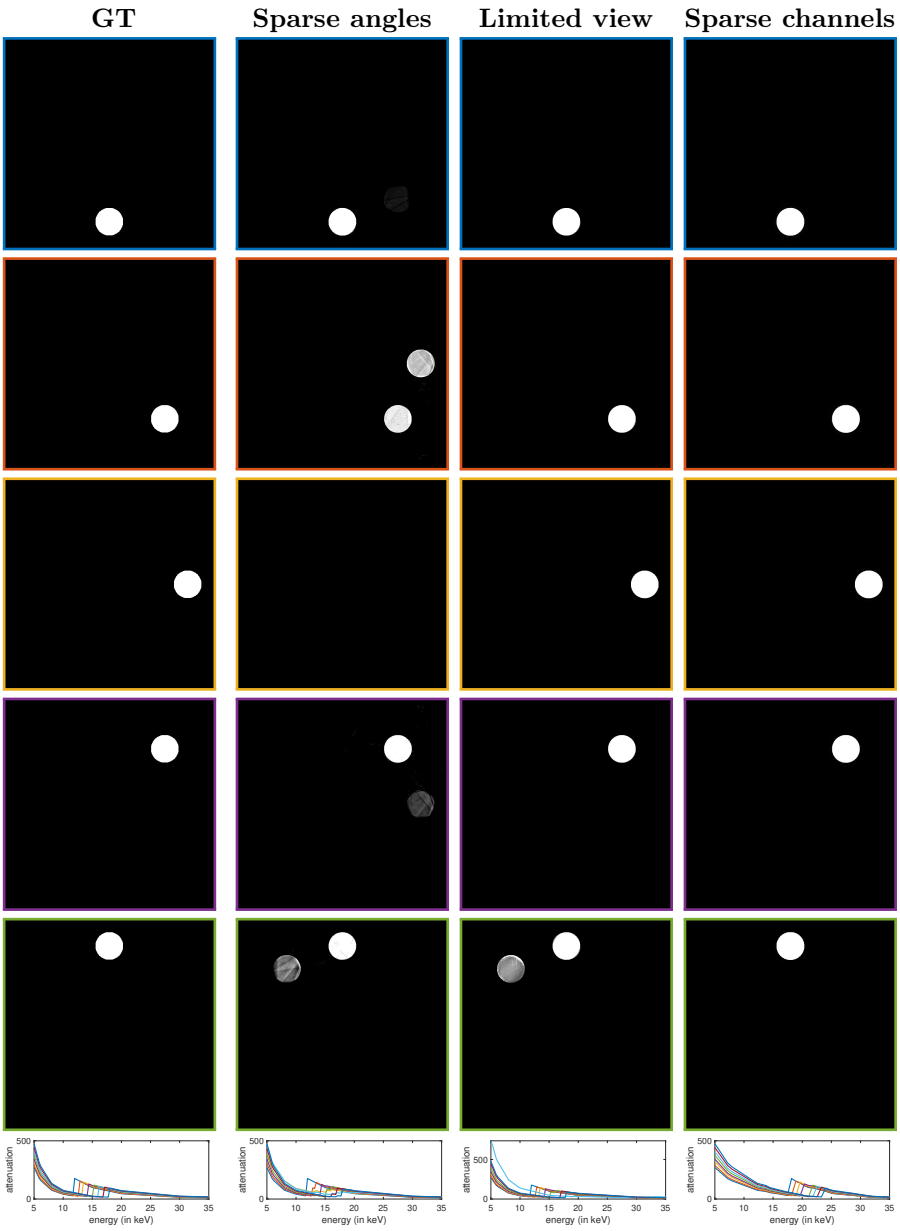


Figure A6: Results of ADJUST with sparse-angle data, limited view data and sparse spectral channels on Disks phantom. We only show material maps of first five materials. The colors of the bounding box of material maps are matched with the spectral signatures of the materials (shown in the bottom row).

C.9 Numerical studies: Mixed material phantom

We consider the Mixed Disks phantom, which consists of solid disks in an inner circle and mixed disks on an outer circle. All material mixtures are present on the outer circle. With $M = 5$ disks on the inner circle, this amounts to 10 mixed disks on the outer circle. The materials are the same as the first 5 selected materials in the Disks phantom. The ADJUST method with 2000 iterations is compared with RU, UR, and cJoint. The results of this experiment are shown in Figure A7, with the results for each material on a separate row.

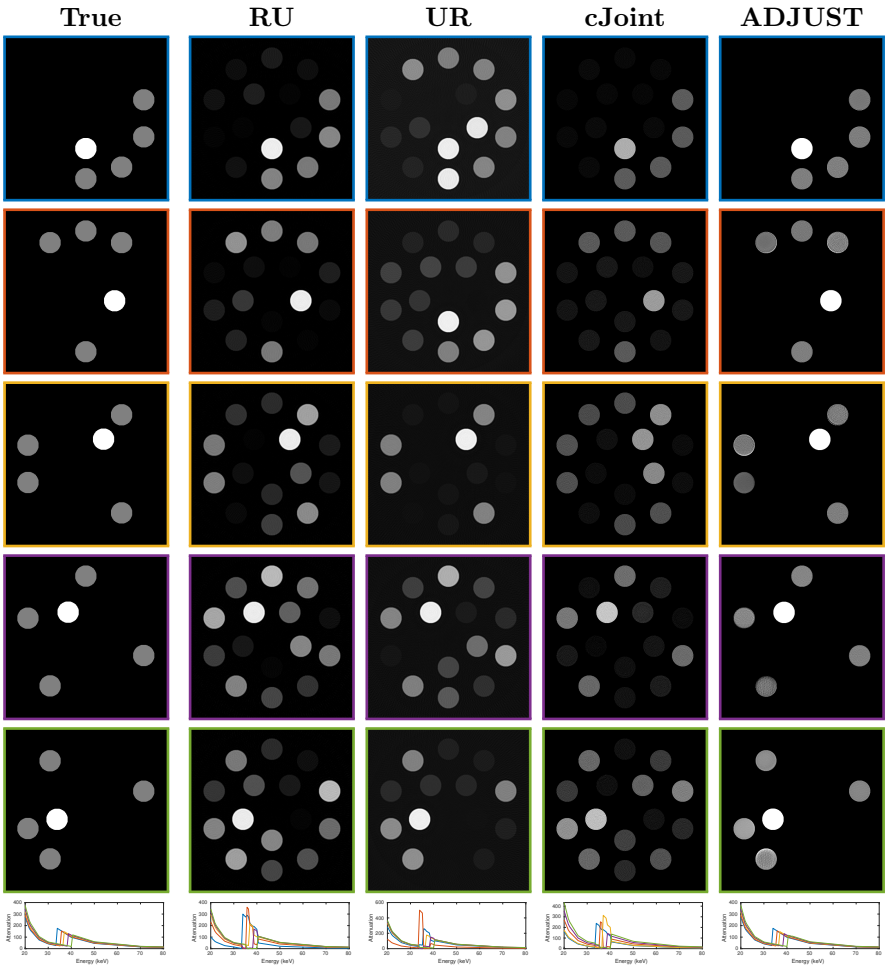


Figure A7: Comparison of various methods for spectral CT for a mixed-material Disks phantom. The materials contained in this phantom are arsenic (top row), selenium, bromine, krypton and rubidium (second-to-last row).

We see that the RU, UR, and cJoint methods are not capable of fully separating the mixtures and retrieving the disks on the inner circle. On the other hand, ADJUST nearly perfectly reconstructs the disks on the inner circle and the mixture disks on the outer circle.

C.10 Numerical studies: 3D phantom

We also apply the ADJUST algorithm to the 3D Shepp-Logan phantom to show the ability to reconstruct a 3D phantom. This 3D phantom is four times as large as the 2D Shepp-Logan phantom. The phantom is discretized on a grid of $128 \times 128 \times 128$ voxels. The phantom is shown in Figure A8. We consider 60 equidistant projection angles in the range of $[0, \pi]$ with a parallel-beam acquisition geometry. We show the visual results of the 3D material decomposition in Figure A9. The average MSE is 0.0029, the average PSNR is 26.67 and the average SSIM is 0.9763, indicating that the 3D reconstructions are almost accurate.

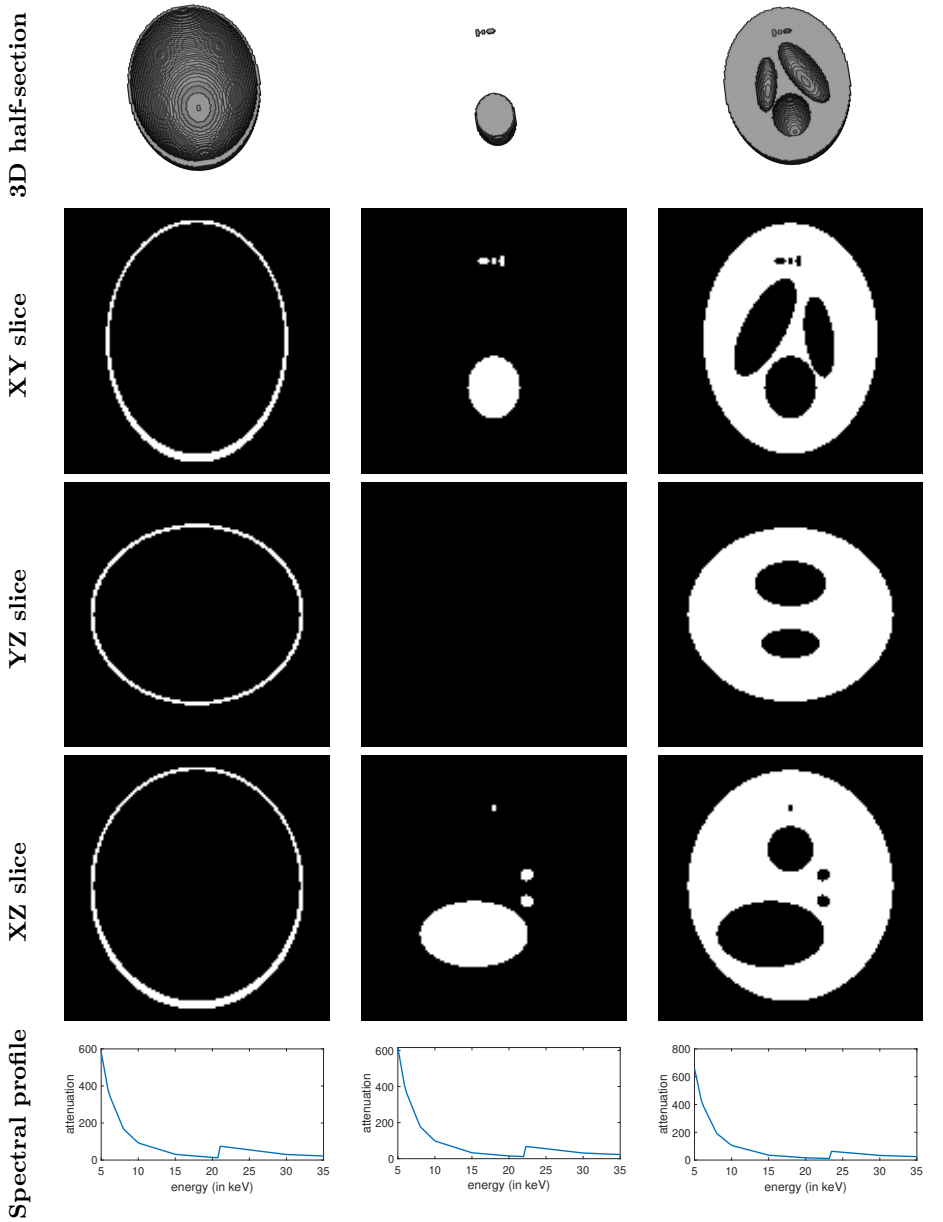


Figure A8: The true material compositions of the 3D Shepp-Logan phantom. The materials contained in this phantom are vanadium (left column), chromium (middle column), and manganese (right column).

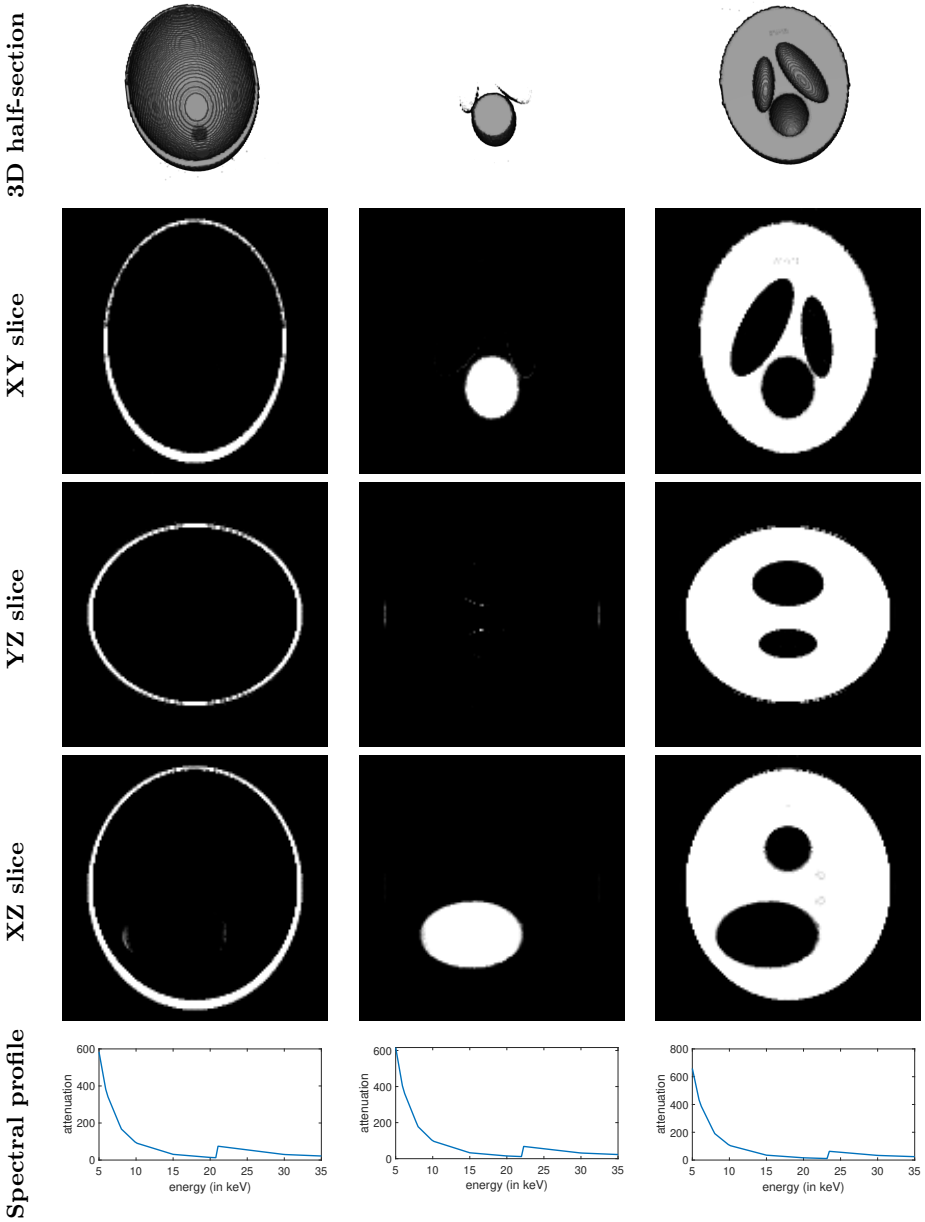


Figure A9: The reconstructed material composition of the 3D Shepp-Logan phantom from ADJUST algorithm. The materials contained in this phantom are vanadium (left column), chromium (middle column), and manganese (right column).

List of publications

Publications that are part of this dissertation:

1. M. T. Zeegers, F. Lucka, and K. J. Batenburg. "A Multi-Channel DART algorithm". In: *International Workshop on Combinatorial Image Analysis*. (Porto, Portugal). Ed. by R. P. Barneva, V. E. Brimkov, and J. M. R. S. Tavares. Springer, 2018, pp. 164–178.

2. M. T. Zeegers, D. M. Pelt, T. van Leeuwen, R. van Liere, and K. J. Batenburg. "Task-driven learned hyperspectral data reduction using end-to-end supervised deep learning". *Journal of Imaging* 6.12 (2020), p. 132.

3. M. T. Zeegers, T. van Leeuwen, D. M. Pelt, S. B. Coban, R. van Liere, and K. J. Batenburg. "A tomographic workflow to enable deep learning for X-ray based foreign object detection". *Expert Systems with Applications* 206 (2022), p. 117768.

4. M. T. Zeegers, A. Kadu, T. van Leeuwen, and K. J. Batenburg. "ADJUST: A Dictionary-based Joint reconstruction and Unmixing method for Spectral Tomography". *Inverse Problems* 38.12 (2022), p. 125002.

Publications that are not part of this dissertation:

1. E. G. Rens, M. T. Zeegers, I. Rabbers, A. Szabó, and R. M. H. Merks. "Autocrine inhibition of cell motility can drive epithelial branching morphogenesis in the absence of growth". *Philosophical Transactions of the Royal Society B* 375.1807 (2020), p. 20190386.