

生体の背景知識を活用する 機械学習と診断治療支援

中尾 恵†

一部の観測情報に基づく人体や生体臓器の状態理解は、医療従事者が日常的に直面する課題であると同時に医用工学における本質的な研究対象である。本稿では、X線画像や内視鏡カメラ画像に対して観測領域を超えた推論を実現する機械学習の枠組み、外科手術や放射線治療への利用を想定した生体医学知識を活用する機械学習の事例を紹介する。

キーワード：位置合わせ、画像生成、超観測、診断治療支援

1. まえがき

「心眼」は物事の本質を見抜く心の働き、心の目。目や耳などの感覚器で知覚できない情報を経験と想像力で推論し、見えない物の具体的な形質や挙動を把握することとされる¹⁾。一見すると非科学的にも聞こえるこの単語に強く印象を持ったのは大学院生の頃で、診断や治療に携わる臨床医学の専門家からだけでなく、生体医工学や医用画像解析を専門とする工学・情報学系の研究者からも何度か耳にした経験からである。

コンピュータビジョン分野における機械学習モデルの進展に伴い、観測された対象を認識、識別、理解する能力は2010年以降に大きく飛躍した。一方、多くは画像内に推論に必要な情報が含まれることを前提として予測に有効な特徴を抽出するものであり、観測データに含まれていない背景知識が必要となる問題では十分な予測性能を発揮できないことが多い。観測データに含まれない高次の背景知識をいかに機械学習に取り入れるか、自然言語処理分野における大規模言語モデルの性能が注目される中、「心眼」の達成は未だ多くの探索余地がある領域であるように思える。

本稿では、生体の背景知識を活用する機械学習と診断治療支援の事例として、観測領域を超えた医用画像の生成や位置合わせの研究をわれわれの取り組みを交えて紹介する。機械学習による画像生成は、欠損を有する低品質画像の改善、超解像、モダリティ変換などの幅広い用途が探索されている^{2) 3) 4)}。また、治療時に得られる低次元の画像に対して、背景知識となる情報の位置合わせを達成することによって、直接的には観測されない臓器内部の腫瘍のより正確な局在を得ることができる。本稿では、画像誘導型の

治療や手術支援を目的に、単一視点の2次元画像に対してより高次の生体情報の位置合わせを達成する最近の研究例について紹介する。

2. 機械学習に基づく医用画像の生成

医用イメージングにおける従来の画像再構成法は、信号処理と3次元幾何、計測値の統計的性質を背景として、モダリティに合わせてアルゴリズム開発⁵⁾がなされてきた。近年では、計測値からの特徴抽出や画像再構成においても機械学習が広く活用される。画像生成モデルの臨床応用は、大きく二つのカテゴリーに分類される。一つは、同一のモダリティ内における画像生成や変換であり、超解像やアーチファクト低減など入力画像の品質向上が目的とされる⁶⁾。一方は、異なるモダリティ間の画像生成や変換であり、MRI (Magnetic Resonance Imaging) 画像とCT (Computed Tomography) 画像間の相互変換、機能画像であるPET (Positron Emission Tomography) 画像からCT画像などのドメイン変換問題を対象とする⁷⁾。本稿では、画像生成モデルの臨床応用の代表例として、前者のカテゴリーに属するアーチファクト低減、欠損画素復元モデルに焦点を当てる。

CT画像における金属アーチファクトは、診断、術前計画、放射線治療などの医用画像を用いた臨床プロトコルに深刻な影響を及ぼす可能性がある。歯科用充填物や固定具など体内留置金属は、X線の透過を妨げる。金属は画像再構成時のCT値の正確な計算を阻害し、その結果としてCT画像に暗帯やストリークアーチファクトを引き起こす⁸⁾。このような画像を補正するためには、アーチファクトの除去と同時に、欠損した領域が本来有するはずのCT値と解剖学的特徴を復元する必要がある。

医用画像における欠損画素の復元問題に対して敵対的生成ネットワーク (Generative Adversarial Nets: GANs) や拡散モデルの応用が試みられている。アーチファクト低減モデルの学習に教師あり学習を用いる場合、アーチファク

† 京都大学 大学院医学研究科 人間健康科学系専攻

"Machine Learning Techniques using Background Biomedical Knowledge for Diagnosis and Therapeutic Support" by Megumi Nakao (Human Health Sciences, Graduate School of Medicine, Kyoto University, Kyoto)

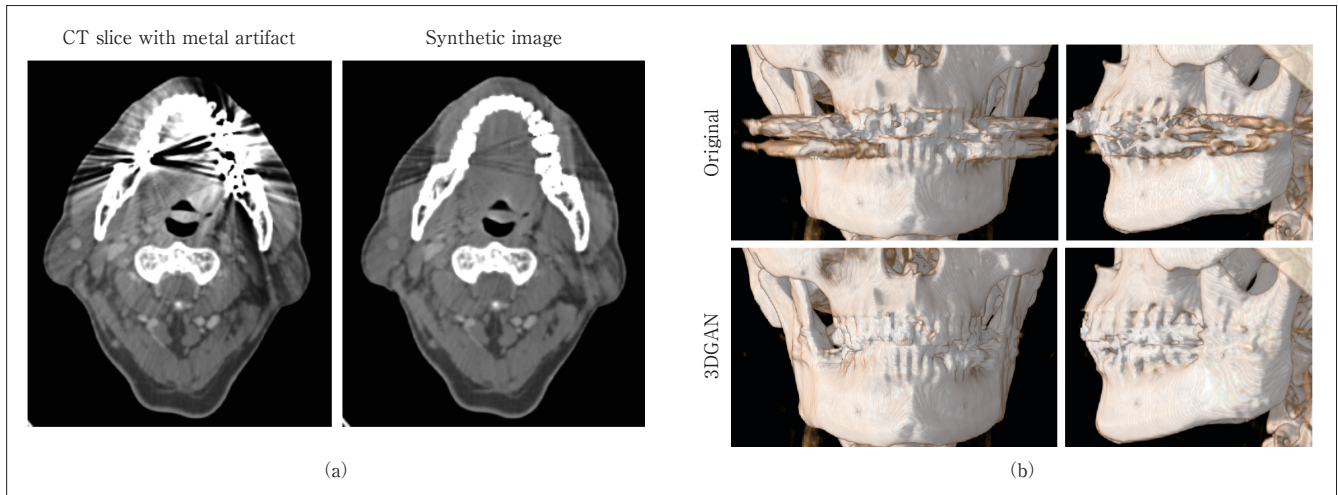
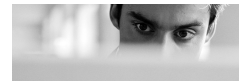


図1 敵対的学習によるCT画像内の欠損画素復元
(a) 元画像と生成画像, (b) 変換前後のボリューム可視化結果

トを有する画像に対応するアーチファクトを有しない教師画像が必要となる。しかし、臨床において一度設置した歯科金属を除去する機会はほとんどないため、アーチファクトの有無のみが異なる同一患者の実画像の集積は困難である。数値計算によって典型的な金属アーチファクトをシミュレートしたCT画像を教師画像とした教師あり学習⁹⁾が報告されているが、学習が事前に用意した金属の3次元形状と配置に依存し、アーチファクト合成画像と実画像に含まれるアーチファクトに差異が生じる点に課題が残る。実画像に生じる金属アーチファクトを直接的に学習するために、学習時に1対1のペア画像を必要としない教師なし学習の枠組みの応用も試みられた。LiaoらはCycleGANを拡張したアーチファクト低減向けモデル¹⁰⁾を提案し、アーチファクト合成画像を利用して教師あり学習と教師なし学習による低減性能を比較した。

深層学習に基づく画像変換ではアーチファクトの影響を受けた画素のみを修正し、他の正常な領域における患者本来の解剖学的構造や画像特徴は厳密に維持される必要がある。著者らは実画像に見られる複数の歯科金属から生成されるCTアーチファクトのスパース性に着目した正則化付きの3次元敵対的生成ネットワーク (3DGAN)¹¹⁾を構築した。合計361件の実患者の3次元CT画像を、歯科用金属の有無に基づいて分類し、正則化付き敵対的学習を行った。

図1(a)は、敵対的学習を用いた頭頸部CTスライス画像の画像変換の結果を示している。左は金属アーチファクトを含む元画像であり、歯科金属周辺に低輝度や高輝度の画素がさまざまな方向に発生し、画素が欠損している。右は3DGANによって得られた画像生成結果である。若干の淡いアーチファクトが残存しているものの、良好にアーチファクトが低減されつつ、本来の生体組織が示すべきCT値が復元されている。3DGANによる変換結果と20年以上の経験を持つ歯科技工士による手動補正結果との比較実験を実施し

た¹¹⁾。手動補正では、歯の表面に不規則なノイズが残存したが、提案モデルによってアーチファクト除去が可能となり、歯列の滑らかな3次元構造を復元できていた(図2)。数十枚からなる画像であるため、手動補正では52分を要したのに対し、3DGANの処理時間は13秒であった。3名の口腔外科専門医による主観評価の結果、生成画像の画質は臨床的に許容される水準に達していることが確認されている。

Cone-beam CT (CBCT)も臨床において画像品質の向上が望まれるモダリティの一つである。CBCTは小型でモビリティ性を有するイメージング装置であり、放射線治療や外科手術など治療時における迅速な計測を可能にする¹²⁾。しかし、撮像時におけるスキャン角度の制限とコーン形状のX線照射のために、CT装置よりも画質は低下し、撮像範囲も限定的である。CBCT画像に生じるアーチファクトも多様な特性を示し、画像空間全体に広範囲にわたって発生する。

筆者らは敵対的学習をCBCT画像に適用し、CT画像相当の画質改善を目指した研究¹³⁾を行った。より近年では、CBCT画像をガイドとした条件付き拡散モデルによる画質向上に取り組んでいる。同一の生体からCBCT画像に空間的に一致したCT画像を得ることは難しいため、CBCT画像の撮像原理を踏まえてアーチファクトをシミュレートした擬似CBCT画像によって学習用データベースを構築するアプローチ¹⁴⁾を採用している。図2(a)はCBCT画像をガイドとし、ノイズからターゲットとするCTスライスが生成されるプロセスを示している。図2(b)は2例のCBCTスライス画像から得られた生成画像であり、広範囲の不正確な画素が修正され、コントラストが改善されていることが確認される。

3. 2D/3D位置合わせによる重畳画像の生成

本章では、機械学習を用いた2D/3D位置合わせによる拡

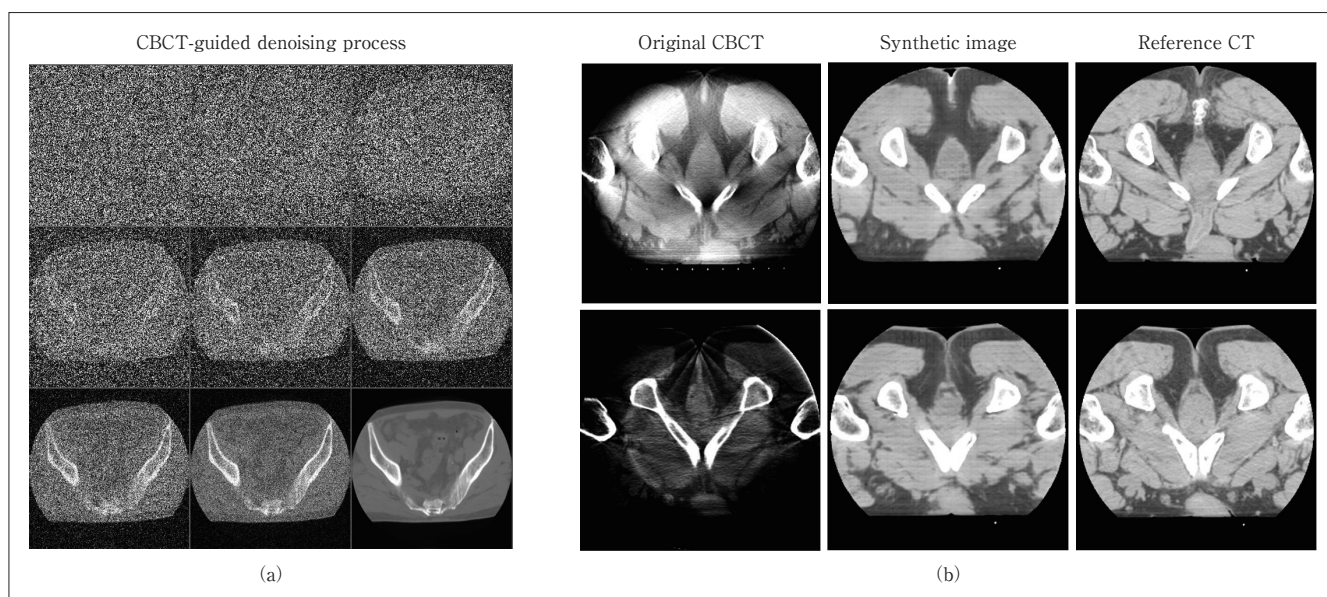


図2 拡散モデルによるCone-beam CT画像の改善, (a) ノイズからの画像生成プロセス, (b) 生成画像と参照画像

張画像の生成と、その画像誘導型の治療や外科手術支援等の臨床応用について紹介する。臨床現場においては、医療従事者と患者双方の撮像負担の軽減、治療時の医療行為を妨げない迅速なイメージング、低次元の計測データからの生体情報の可視化が期待される。

近年では、直交2枚の投影画像や単一視点のX線画像など、極めて低次元の画像群からボリュームデータや3次元形状を再構成する試み^{15) 16)}が行われている。例えば、X2CT-GAN¹⁵⁾は直交2枚の投影画像から深層学習を用いて3次元画像を生成する枠組みであり、臨床応用の可能性が検証されてきた。しかし、これらモデルでは、復元可能な3次元情報が低解像度の特徴に限られており、高解像度の画像特徴を再構成することは依然として課題である。他にも治療前の3D-CT画像を活用し、患者固有データごとに学習を行うことで、単一投影画像のみを用いた呼吸変位による画像欠損の高精度な復元を達成する研究¹⁷⁾が報告されている。

一般に、単一の投影画像から臓器の3次元形状を再構成するには、2次元の画像空間内の点からより高次元の空間の点への対応を得る必要があり、事前知識がない場合には明らかに不良設定問題である。この問題を良設定の問題として定式化するために、治療計画用に取得された3D-CTや3D-MRI画像に基づく患者固有の臓器形状を利用することが考えられる。このアプローチは2D/3D位置合わせと呼ばれ、治療中に取得された2次元投影画像から臓器の位置や変形を予測し、空間的に位置合わせてより高次元の生体情報の重畳を可能にする。

2D/3D位置合わせは医用画像解析分野において活発に研究されてきた研究分野の一つである^{18) 19)}。特に、剛体を対象とした2D/3D画像位置合わせは、X線画像における撮像対象の位置や向きを決定するパラメータの最適化問題とし

て広く研究されてきた²⁰⁾。剛体位置合わせはアフィン変換行列の最適化問題として定式化されるのに対して、可変形位置合わせは変位ベクトル場 (Displacement Vector Fields: DVF) の最適化問題として扱われる。特に変形臓器の2D/3D変形位置合わせでは、2次元画像と3次元ボリューム間の局所的な点对応を求めることが課題となる。剛体位置合わせとは異なり、可変形位置合わせではサンプリングポイントの数に比例する大規模なパラメータ最適化が必要となる。また、ボリュームデータは高密度かつ大規模であるため、解剖学的ラベルから形状モデルを構築し、それらの位置や変形を2次元画像上に重ね合わせる手法が有効と考えられる。特に、メッシュを用いた形状表現は、少数の変数で弾性特性や統計的な形状変化を表現できるため^{21) 22)}、リアルタイム性が求められるシステムにも適用しやすい。

近年、著者らは単一視点から得られる2次元投影像に対する可変形メッシュ位置合わせの枠組みとして、Image-to-Graph Convolutional Neural Network (IGCN)²³⁾を構築した。本モデルは、2次元投影像の画像特徴に基づいて平均形状モデルからの変形を学習し、臓器形状メッシュの可変形位置合わせを達成することによって重畳画像を生成する。さらに、IGCNの特徴抽出部分を改良し、画像間変換を利用した3次元変位場生成を可能にした。改良版のIGCN+は、U-netベースの変位場生成モデルとグラフ畳み込みネットワークを統合した枠組み²⁴⁾としている。図3はIGCN+を用いて単一の擬似X線画像 (Digitally Reconstructed Radiograph: DRR) (左) に対して、予測された3次元メッシュを位置合わせした重畳画像 (右) を示している。投影像では胃、十二指腸、膵臓の輪郭は視認できないものの、周辺の臓器の画像特徴との関係性を学習することによって、それぞれの臓器の位置や形状を予測できることが確認された。

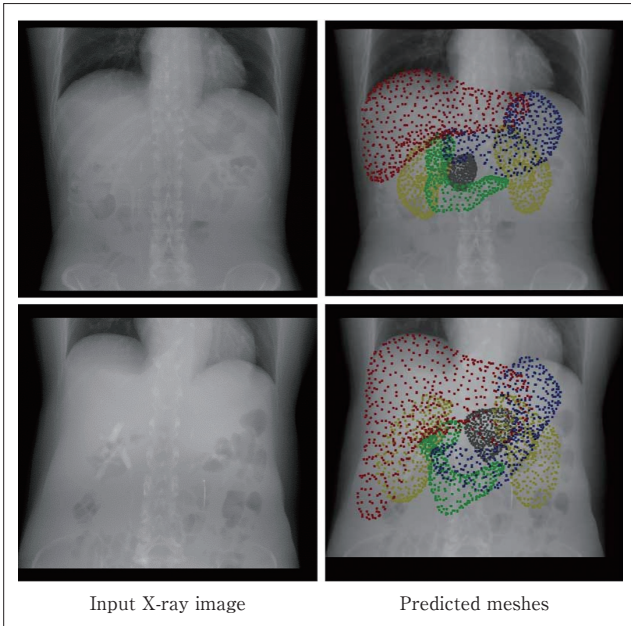
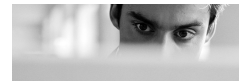


図3 擬似X線画像に対する腹部5臓器(肝臓, 胃, 十二指腸, 腎臓, 膵がん)の可変形メッシュ位置合わせ

深層学習を用いた2D/3D画像位置合わせは、画像誘導型の放射線治療や外科手術における術中ガイドへの利用が期待されている。例えば、術前に取得した患者固有の3次元CT画像から臓器モデルを生成し、それを手術中の内視鏡カメラ画像に重ね合わせることによって、拡張内視鏡画像を作成することができる²⁵⁾。図4は腹腔鏡画像や胸腔鏡画像に対する臓器形状の重畳結果を示している。臨床医学を対象とした機械学習モデルの研究においては、学習に利用可能なデータが十分に得られない点が課題である。われわれは計算機上で表現あるいはシミュレートした人工的なデータを用いて、位置合わせパラメータの学習を行うオフライン学習の枠組み²⁷⁾を探索している。具体的には、3D-CT画像から作成された患者個人の臓器メッシュを用い、

実際の手術映像からも共通に得られる臓器の形状や外観の画像特徴を表現したセマンティックラベルをさまざまに生成して、姿勢推定等の学習に用いる。

図4(上段)では26例、図4(下段)では30例の小規模な3D-CTデータベースからオフライン学習の考え方に基づいて約1万枚のセマンティックラベルをレンダリングによって生成し、位置合わせパラメータの予測タスクを学習することによって予測精度の達成を試みている。データの集積に加えて、定量評価が困難であることなど多くのハードルがあるが、臨床医学の専門家との継続的な共同研究の推進によって、臨床的に許容される水準のモデルが着実に得られつつある。

4. むすび

本稿では、背景知識を活用する機械学習の診断治療支援として、画像生成と形状位置合わせ技術に関する研究動向と事例を紹介した。臨床医学において、機械学習を活用した画像生成は追加計測による患者への追加的な負担や臨床努力を必要とせず、低次元の画像から高次元の解剖学的情報を生成できるという特長を持つ。画像誘導手術や放射線治療における新たな診断・治療応用の可能性を広げることが期待される。

(2025年2月9日受付)

〔文 献〕

- 1) Wikipedia, <https://ja.wikipedia.org/wiki/心臓>
- 2) X. Yi, E. Walia, P. Babyn: "Generative adversarial network in medical imaging: A review", *Medical Image analysis*, **58**, 101552 (2019)
- 3) T. Wang, Y. Lei, Y. Fu, J.F. Wynne, W.G. Curran, T. Liu and X. Yang: "A review on medical imaging synthesis using deep learning and its clinical applications", *J Appl Clin Med Phys*, **22**, pp.11-36 (2021)
- 4) M. Nakao: "Medical image synthesis and statistical reconstruction methods", *Advanced Medical Engineering*, **12**, pp.21-27 (2023)
- 5) J.W. Stayman, Y. Otake, J.L. Prince, A.J. Khanna and J.H. Siewerdsen: "Model-based tomographic reconstruction of objects containing known components", *IEEE Trans. Med. Imaging*, **31**, 10, pp.1837-48 (2012)

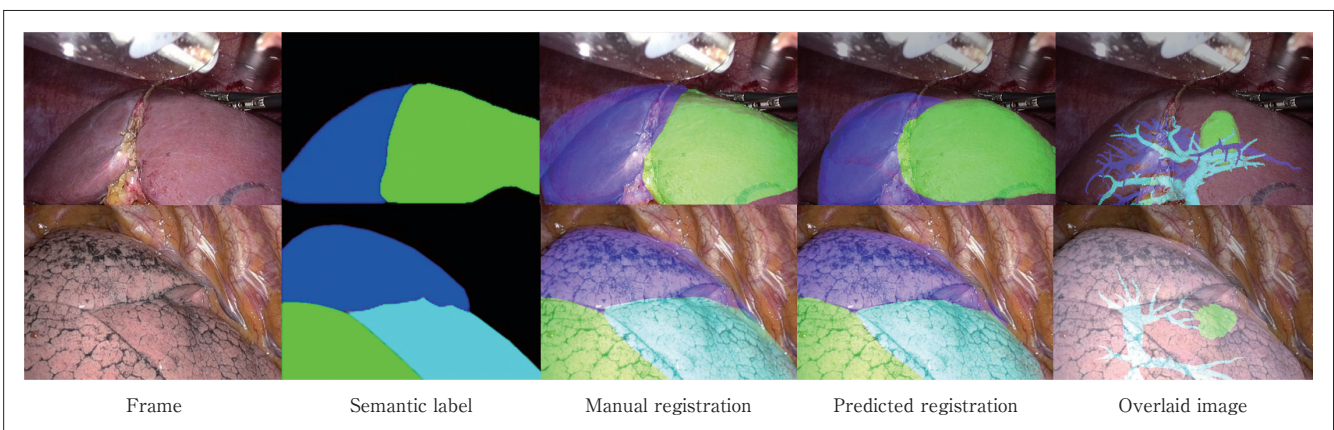


図4 オフライン学習モデルによる臓器メッシュの2D/3D位置合わせ

(上段) 腹腔鏡画像に対する肝臓形状の重畳, (下段) 胸腔鏡画像に対する虚脱肺形状の重畳. 左から手術映像の1フレーム, U2-net²⁶⁾による解剖学的ラベルの予測結果, 手動による位置合わせ, 予測された位置合わせ, 重畳画像. 臓器内部の脈管構造(青, シアン)と腫瘍(緑)が可視化されている。



- 6) C. You, G. Li, Y. Zhang, X. Zhang, H. Shan, M. Li et al: "CT Super-Resolution GAN Constrained by the Identical", Residual and Cycle Learning Ensemble (GAN-CIRCLE), IEEE Trans. Med. Imaging, **39**, 1, pp.188-203 (2020)
- 7) H. Emami, M. Dong, S.P. Nejad-Davarani and C.K. Glide-Hurst: "Generating synthetic CTs from magnetic resonance images using generative adversarial networks", Med. Phys., **45**, 8, pp.3627-3636 (2018)
- 8) L. Gjestebj, B.D. Man, Y. Jin, H. Paganetti, J. Verburg, D. Giantsoudi et al: "Metal artifact reduction in CT: Where are we after four decades?", IEEE Access, **4**, pp.5826-5849 (2016)
- 9) Y. Zhang and H. Yu: "Convolutional neural network based metal artifact reduction in X-ray computed tomography", IEEE Trans. Med. Imaging, **37**, 6, pp.1370-1381 (2018)
- 10) H. Liao, W.A. Lin, J. Yuan, S.K. Zhou and J. Luo: "Artifact disentanglement network for unsupervised metal artifact reduction", Med. Image Compt. And Comput. Assist. Interv. (MICCAI), pp.203-211 (2019)
- 11) M. Nakao, K. Imanishi, N. Ueda, Y. Imai, T. Kirita, T. Matsuda: "Regularized three-dimensional generative adversarial nets for unsupervised metal artifact reduction in head and neck CT images", IEEE Access, **8**, pp.109453-65 (2020)
- 12) M. Posiewnik, T. Piotrowski: "A review of cone-beam CT applications for adaptive radiotherapy of prostate cancer", Physica Medica, **59**, pp.13-21 (2019)
- 13) T. Hase, M. Nakao, K. Imanishi, M. Nakamura, T. Matsuda: "Improvement of image quality of cone-beam CT images by three-dimensional generative adversarial network", 43rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), pp.2843-2846 (2021)
- 14) 村橋成紀, 中村光宏, 中尾恵: "条件付き拡散モデルを用いたCone-beam CT画像の画質改善の試み", 信学技報 (MI), **123**, 411, pp.109-112 (2024)
- 15) X. Ying, H. Guo, K. Ma, J. Wu, Z. Weng and Y. Zheng: "X2CT-GAN: Reconstructing CT from Biplanar X-rays with generative adversarial networks", IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp.10611-10620 (2019)
- 16) Y. Kasten, D. Doktofsky and I. Kovler: "End-To-End convolutional neural network for 3D reconstruction of knee bones from Bi-planar X-Ray Images", Machine Learning for Medical Image Reconstruction, pp.123-133 (2020)
- 17) L. Shen, W. Zhao and L. Xing: "Patient-specific reconstruction of volumetric computed tomography images from a single projection view via deep learning", Nature Biomedical Engineering, **3**, 11, pp.880-888 (2019)
- 18) P. Markelj, D. Tomaževič, B. Likar and F. Pernuš: "A review of 3D/2D registration methods for image-guided interventions", Medical Image Analysis, **16**, 3, pp.642-661 (2012)
- 19) U. Mitrovic, Ž. Špiclin, B. Likar and F. Pernuš: "3D-2D registration of cerebral angiograms: A method and evaluation on clinical images", IEEE Trans. Med. Imaging, **32**, 8, pp.1550-63 (2013)
- 20) C.J. F. Reyneke, M. Lüthi, V. Burdin, T. Douglas, T. Vetter and T. Mutsavangwa: "Review of 2-D/3-D reconstruction using statistical shape and intensity models and X-Ray image synthesis: toward a unified framework", IEEE Reviews in Biomedical Engineering, Vol.12, pp.269-286 (2019)
- 21) M. Nakao, M. Nakamura, T. Mizowaki and T. Matsuda: "Statistical deformation reconstruction using multi-organ shape features for pancreatic cancer localization", Medical Image Analysis, Vol.67, p.101829 (2021)
- 22) J. Tokuno, T.F. Chen-Yoshikawa, M. Nakao, M. Iwakura, T. Motoki, T. Matsuda, H. Date: "Creation of a video library for education and virtual simulation of anatomical lung resection", Interactive Cardiovascular and Thoracic Surgery, **34**, 5, pp.808-813 (2022)
- 23) M. Nakao, M. Nakamura, T. Matsuda: "Image-to-graph convolutional network for deformable shape reconstruction from a single projection image", International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI), pp.259-268 (2021)
- 24) M. Nakao, M. Nakamura, T. Matsuda: "Image-to-graph convolutional network for 2D/3D deformable model registration of low-contrast organs", IEEE Trans. on Medical Imaging, **41**, 12, pp.3747-3761 (2022)
- 25) T. Oya, Y. Kadomatsu, T.F. Chen-Yoshikawa, M. Nakao: "2D/3D deformable registration for endoscopic camera images using self-supervised offline learning of intraoperative pneumothorax deformation", Computerized Medical Imaging and Graphics, 102418 (2024)
- 26) X. Qin, Z. Zhang, C. Huang, M. Dehghan, O.R. Zaiane, M. Jagersand: "U2-Net: Going deeper with nested U-Structure for salient object detection", Pattern Recognition, Vol.106, 107404 (2020)
- 27) 小林真実, 木戸口勇気, 小木曾聡, 西野裕人, 中尾恵: "オフライン学習を用いた腹腔鏡画像に対する形状位置合わせ", 信学技報 (MI), **124**, 184, pp.12-15 (2024)



中尾 恵 2003年、京都大学大学院情報学研究科博士課程修了。同大学院医学研究科助手(特任)、奈良先端科学技術大学院大学情報科学研究科助教、ハーバード大学医学部在外研究員、京都大学大学院情報学研究科准教授を経て、2022年より、京都大学大学院医学研究科人間健康科学系専攻教授。知能医工学に関する研究に従事。博士(情報学)。