

U-Net 特徴量による特発性肺線維症患者の予後予測
Use of generative adversarial network for
survival prediction in patients with interstitial lung diseases

植村 知規^{†,‡}
Tomoki Uemura

渡利 千夏[‡]
Chinatsu Watari

Näppi Janne[‡] 弘中 亨[†]
Janne Näppi Toru Hironaka

神谷 亨[†]
Tohru Kamiya

吉田 広行[‡]
Hiroyuki Yoshida

1. はじめに

突発性肺線維症(IPF: Idiopathic Pulmonary Fibrosis)は進行性線維化を伴う間質性肺疾患の主なもののひとつであり、その発症原因は不明である。米国では10万人あたり58.7人程度が発症すると推定されている[1]。経年的に呼吸機能の低下、自覚症状の悪化を来す予後不良の疾患であり、抗線維化薬の投与を行わなかった場合、予後の中央値は診断後2-5年とされている[2], [3]。しかし、個々の症例の臨床経過は多様であるため、正確な予後予測は困難である。

IPFの予後予測には性別(gender)、年齢(age)、肺機能(physiology)から算出されるGAP-Index[4], [5]が広く用いられている。しかし、IPFに向けた画像を用いた予後予測手法は未だ確立されていない。

U-Net[6]は入力画像のセマンティックセグメンテーションに有効なマルチスケールな特徴量を学習により獲得可能である。U-Netにより、パターン認識やセグメンテーション技術は大幅に向上している[7]。IPF患者の予後予測に関し、GAP-indexの一部をCT画像から得られるスコアに置き換えたCT-GAPモデルは、GAPモデルと同程度の精度であることが示されており[8]、CT画像から病状を学習し評価可能なU-NetはIPF患者の予後予測に応用可能であると推測される。U-Netから得られる特徴量は、人間により認識不可能であった新たな画像特徴を識別し、IPFの予後予測指標として利用可能であると期待される。

本研究では、U-Netから得られる画像特徴量を利用した新たな予後予測モデルを提案し、GAP-indexおよびComposite Physiologic Index (CPI)[9]との予測性能の比較を行った。

2. 関連研究

Leyら[4]は、性別(gender)、年齢(age)、努力性肺活量(FVC)および肺拡散能(D_{Lco})の2つの肺機能(physiology)のスコアリングによって算出されるGAP-indexを提案し、IPF患者の死亡率推定を行った。結果として、C-indexにおいて69.7%を達成し、GAP-indexがIPF患者の死亡率推定に有効なスコアであることを示した。また、Leeら[5]は、大規模な実験により、GAP-indexのスコア毎の生存率について検証を行った。結果として、GAP-indexがIPF患者の生存率に対して有意な予測能力を持つことを示した。Wellsら[9]は、肺機能検査によって得られる、肺拡散能(D_{Lco})、努

力性肺活量(FVC)、1秒量(FEV_1)によって算出されるCPIを提案し、CPIがIPF患者の予後予測に対する有効な指標となることを示した。これらのモデルは、患者の人口統計や肺機能検査によって算出されるスコアを基にしており、画像情報は考慮されていない。

Leyら[8]は、GAP-indexにおける肺拡散能をCT画像の読影によって得られるCT線維化スコアに置き換えたCT-GAPモデルを提案し、GAPモデルとの比較を行った。結果として、CT-GAPモデルとGAPモデルのC-indexはそれぞれ、70.3%、70.7%であり、同等の性能を持つことを示した。従って、CT画像によって得られるスコアはIPF患者の予後予測に有効である可能性があることが示された。一方で、CT-GAPモデルとGAPモデルとの間に有意差は存在せず、また、C-indexも約70%と低いことから、依然として画像解析に基づく予後予測モデルとして十分な性能は達成されていない。

3. 手法

本研究では、U-Netから得られる画像特徴量(U-Radiomics)を利用した、CT画像解析に基づく新たな予後予測モデルを提案する。予後予測モデル全体の概念図を図1に示す。提案モデルは、学習済みのU-Netに対しCT画像を入力し、U-Netにおけるエンコード部の最下層の出力を画像特徴量として得る。さらに、得られた特徴量を基にElastic-net Coxモデル[10]による生存解析を行う。以下にその詳細を述べる。

3.1 U-Net

U-Netは入力画像のセマンティックセグメンテーションを行う目的で提案された深層学習モデルである。本研究においてU-Netは、胸部CT画像を入力とし、肺野内外、正常組織および、8種類の病変領域をラベルとするセグメンテーション画像の生成を行うよう学習される。

リウマチ関連間質性肺疾患(RA-ILD: Rheumatoid Arthritis-associated Interstitial Lung Diseases)患者の胸部CT画像における病変パターンはIPFと類似することから、RA-ILD患者の胸部CT画像を用い、U-Netの学習を行った。U-Netの構造は、図1中の左部に示され、8層の畳み込み、逆畳み込み層によって構成される。学習には我々の研究機関に登録されたRA-ILD 72症例の胸部CT画像を利用した。

3.2 U-Radiomics

U-Netはエンコード部とデコード部の2つのネットワーク構造を持ち、入力画像から効率的にセグメンテーション

[†]九州工業大学 Kyushu Institute of Technology

[‡]Massachusetts General Hospital and Harvard Medical School

表 1: 各モデルにおける C-index [95% 信頼区間(CI)]

| Clinical/Radiomic Features | C-index [95% CI] |
|-----------------------------------|--------------------|
| U-radiomics | 86.0% [78.1, 93.2] |
| U-radiomics + CPI + GAP | 86.0% [78.0, 93.2] |
| Gender, age, and physiology (GAP) | 65.5% [56.1, 75.2] |
| Composite physiologic index (CPI) | 64.6% [54.6, 74.8] |

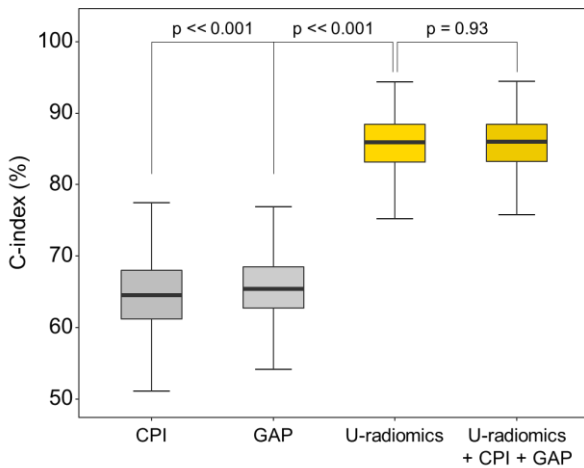


図 2: 各モデルにおける C-index のボックスプロット

また、バイアスによる影響を考慮するため、ブートストラップ法による 500 回の反復を行い、それぞれのモデルにおける C-index を求めた。

さらに、U-Radiomics と GAP, CPI モデルに関し Kaplan-Meier 生存曲線を描画し、それぞれの指標による、死亡確率に関する高リスク群と低リスク群の 2 群の分離度について、log-rank テストによる有意差検定を行った。この時、有意水準は $p = 0.05$ とした。

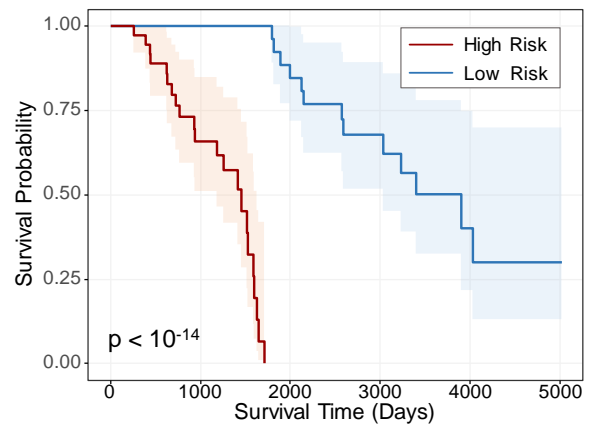
5. 実験結果

5.1 予測性能

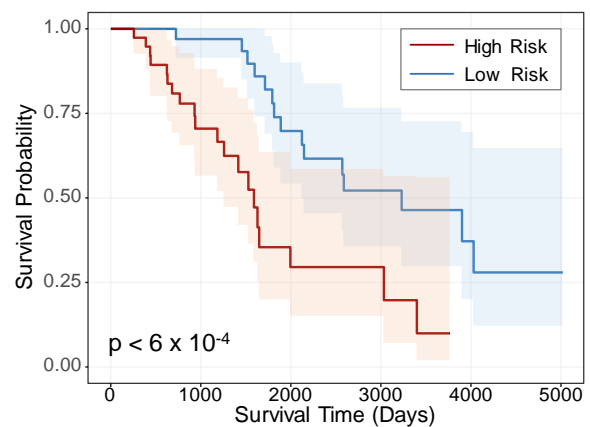
各モデルにおける C-index を表 1 および、図 2 に示す。C-index はそれぞれ、U-Radiomics: 86.0% [Confidence Interval (CI): 78.1, 93.2], GAP: 65.5% [CI: 56.1, 75.2], CPI: 64.6% [CI: 54.6, 74.8] となった。U-Radiomics と GAP, CPI の C-index の差は統計的に有意 ($p << 0.001$) であった。U-Radiomics に GAP, CPI を加えたモデルの C-index は、U-Radiomics + GAP + CPI: 86.0% [CI: 78.0, 93.2] であり、U-Radiomics との間有意差は存在しなかった ($p = 0.93$)。

5.2 Kaplan-Meier 生存曲線

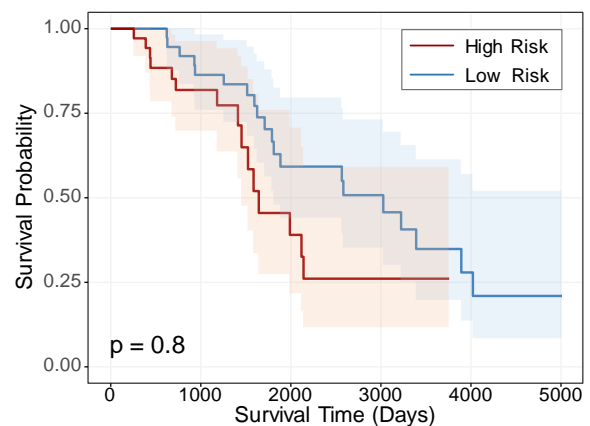
各モデルの Kaplan-Meier 生存曲線を図 3 に示す。高リスク群と低リスク群の曲線の差に関して、それぞれのモデルにおける p 値は、U-Radiomics: $p = 1.0 \times 10^{-9}$, GAP: $p = 6.0 \times 10^{-4}$, CPI: $p = 8.0 \times 10^{-1}$ であり、U-Radiomics および、GAP モデルにおいて、高リスク群と低リスク群の曲線間に有意差が存在することが示された。



(a) U-Radiomics



(b) GAP



(c) CPI

図 3: (a) U-Radiomics, (b) GAP, (c) CPI における Kaplan-Meier 生存曲線

6. 考察

U-Radiomics は 2 つの臨床バイオマーカー(GAP, CPI)によるモデルと比べ、C-index において高い性能を示し、その差は統計的に有意であった。また、生存曲線において、U-Radiomics は他のモデルと比べ、高リスク群と低リスク群をより顕著に分離可能であった。さらに、U-Radiomics

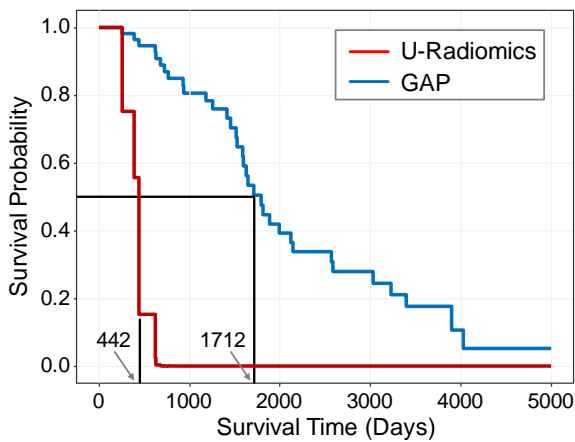


図 4: IPF 患者(死亡例:437 日)に対する U-Radiomics (赤線)、GAP (青線)による生存曲線の推定

に GAP, CPI を加えたモデル(U-Radiomics + GAP + CPI)において C-index の改善は見られないことから, U-Radiomics は GAP, CPI と比べ, IPF に対するより包括的な評価が可能であると考えられる。

診断 437 日後にイベントが起こった IPF 患者に対し, U-Radiomics, GAP により推定された生存曲線を図 4 に示す。図 4 において, 生存確率 50% の時点を予測されたイベント発生日であるとした場合, U-Radiomics は 437 日, GAP は 1712 日にイベントが発生すると予測され, U-Radiomics は GAP による予測と比べ, より正確な実イベント時間(437 日)の推定を行っている。

以上より, U-Radiomics は IPF の予後予測に対し, GAP および CPI よりも優れたバイオマーカーに成り得ると考えられる。

7. 結論

本研究では, U-Net に基づく新たな画像特徴量である U-Radiomics を用いた, CT 画像解析に基づく予後予測モデルを提案した。また, IPF 患者の胸部 CT 画像を用いた実験により, 提案モデルの性能評価を行った。U-Radiomics に基づく予後予測モデルは, 臨床バイオマーカー(GAP, CPI)と比較し, 高い性能を示した。従って, 提案モデルである U-Radiomics は, IPF 症例の予後予測に有効なバイオマーカーに成り得ることが示された。

参考文献

- [1] D. B. Esposito *et al.*, “Idiopathic Pulmonary Fibrosis in United States Automated Claims. Incidence, Prevalence, and Algorithm Validation,” *Am. J. Respir. Crit. Care Med.*, vol. 192, no. 10, pp. 1200–1207, 2015, doi: 10.1164/rccm.201504-0818OC.
- [2] G. Raghu *et al.*, “An Official ATS/ERS/JRS/ALAT Statement: Idiopathic pulmonary fibrosis: Evidence-based guidelines for diagnosis and management,” *Am. J. Respir. Crit. Care Med.*, vol. 183, no. 6, pp. 788–824, 2011, doi: 10.1164/rccm.2009-040GL.
- [3] A. Guenther *et al.*, “The European IPF registry (eurIPFreg): Baseline characteristics and survival of patients with idiopathic pulmonary fibrosis,” *Respir. Res.*, vol. 19, no. 1,

- p. 141, 2018, doi: 10.1186/s12931-018-0845-5.
- [4] B. Ley *et al.*, “A multidimensional index and staging system for idiopathic pulmonary fibrosis,” *Ann. Intern. Med.*, vol. 156, no. 10, pp. 684–691, May 2012, doi: 10.7326/0003-4819-156-10-201205150-00004.
- [5] S. H. Lee *et al.*, “Predicting survival of patients with idiopathic pulmonary fibrosis using GAP score: a nationwide cohort study,” *Respir. Res.*, vol. 17, no. 1, p. 131, Oct. 2016, doi: 10.1186/s12931-016-0454-0 [doi].
- [6] O. Ronneberger, P. Fischer, and T. Brox, “U-Net: Convolutional Networks for Biomedical Image Segmentation,” pp. 1–8, 2015, doi: 10.1007/978-3-319-24574-4_28.
- [7] M. H. Hesamian, W. Jia, X. He, and P. Kennedy, “Deep Learning Techniques for Medical Image Segmentation: Achievements and Challenges,” *J. Digit. Imaging*, vol. 32, no. 4, pp. 582–596, 2019, doi: 10.1007/s10278-019-00227-x.
- [8] B. Ley *et al.*, “Idiopathic pulmonary fibrosis: CT and risk of death,” *Radiology*, vol. 273, no. 2, pp. 570–579, Nov. 2014, doi: 10.1148/radiol.14130216.
- [9] A. U. Wells *et al.*, “Idiopathic pulmonary fibrosis: A composite physiologic index derived from disease extent observed by computed tomography,” *Am. J. Respir. Crit. Care Med.*, vol. 167, no. 7, pp. 962–969, 2003, doi: 10.1164/rccm.2111053.
- [10] N. Simon, J. Friedman, T. Hastie, and R. Tibshirani, “Regularization paths for Cox’s proportional hazards model via coordinate descent,” *J. Stat. Softw.*, vol. 39, no. 5, pp. 1–13, 2011, doi: 10.18637/jss.v039.i05.
- [11] D. R. Cox, “Regression Models and Life-Tables,” *J. R. Stat. Soc.*, vol. 34, no. 2, pp. 187–220, 1972.
- [12] R. Tibshirani, “Regression Shrinkage and Selection via the Lasso,” *J. R. Stat. Soc. Ser. B*, vol. 58, no. 1, pp. 267–288, Jun. 1996.
- [13] A. E. Hoerl and R. W. Kennard, “Ridge Regression: Biased Estimation for Nonorthogonal Problems,” *Technometrics*, vol. 12, no. 1, pp. 55–67, 1970, doi: 10.1080/00401706.1970.10488634.