

修士論文

教師なし病理画像特徴量抽出による 骨軟部腫瘍患者生存期間推定

令和 5 年度修了

三重大学大学院 工学研究科 情報工学専攻
ヒューマンコンピュータインタラクション研究室

野中 泰秀

はじめに

軟部組織腫瘍は全身の軟部組織（臓器、骨、皮膚を除いた組織で筋肉、結合組織、脂肪、血管、リンパ管、関節、神経など）に発生する腫瘍で、ほとんどが良性である。ヒトにおいて、悪性軟部腫瘍は、すべてのがんの中で約1%程度であり、主ながんと比べると、非常に稀ながんであるとされている。初期は痛みを伴わないしこりが軟部組織に発生し、徐々に大きくなっていくにつれて、神経が圧迫され痛みが生じる。

診断とその治療には、X線検査、形態学的特徴、分子所見、個々の患者の特徴を考慮する必要があるが、とりわけ病巣部細胞を観察する病理組織診断は、組織学的な悪性度の評価や、術後治療計画の決定に使用されるため、不可欠であるとされている [1]。しかし、専門とする病理医や医療機関が少ないため、診断に時間がかかってしまうことが現状である。

医用画像診断技術の進歩により、スライドガラス上の染色標本全体を高精細なデジタル画像（Whole Slide Image : WSI）に変換する技術が利用可能となった。WSI とテレコミュニケーションシステムを駆使して、全国の医療機関と共有することで、遠隔診断を行っている活用事例も存在している。悪性軟部腫瘍において、病理画像は術後治療計画を立てる際に用いられ、患者に負担のかかる化学療法を行うべきかどうかの判断に使用される。しかし、この判断は容易ではなく、病理画像に基づいた客観的な診断システムが求められている。

本研究は、病理画像から正確な生存期間を推定することで、医師が合理的な治療計画を立てる定量的な指標を得ることを目的とする。この指標によって、病理医が行う治療方法の決定にかかる時間の軽減を実現することで、医師の負担軽減につながり、患者の病状に合わせた治療の実現が見込まれる。

近年、深層学習は目覚ましい発展を遂げており、画像認識の分野に関してもその限りではない。WSI と深層学習を用いて領域抽出を行っている研究やがんの種類を自動検出する研究は多くなされているが、生存期間のような連続値に関連する研究はまだ十分になされていない。生存期間予測の研究として、WSI を用いて Graph Convolutional Network (GCN) により学習した生存リスク特徴を生存期間予測に適用することで、予測精度を

向上させたものがある [2]. 結果として, bladder urothelial carcinoma (BLCA) のデータセットに対して, Mean Absolute Error (MAE) は 123.2 日, Concordance-index (C-index) は 0.849 となり, 最も良い他のモデルに比べて MAE は約 30 日減少し, C-index は 0.3 ポイント増加した. しかし, 問題点として BLCA のような大規模なデータセットを用いていることがあげられる. 骨軟部腫瘍は希少がんのため, より少ないデータセットで高い精度を出すことが求められる.

本研究では, 教師なし学習による画像特徴量抽出と, クラス分類によって生存期間予測を行う. データセット数が少ないため, 正解ラベル分布に過度に近づきすぎることを防ぐために, 教師なし学習を事前学習として用いる. また, 生存期間のような連続値を扱う場合には, しばしば回帰問題として学習される. しかし, 分類問題に対して, 多くのデータセットが必要であるという問題点がある. そのため, 生存期間を C 個のクラスに分割し, それぞれの出力をそのクラスの所属度とすることで, クラス分類問題として学習する. 結果として, 非生存患者の生存期間の誤差は最小値で 7.69 か月だった. 生存期間が 1 年以内の予測に関しては, 予測値の最頻値が分割した正解クラスの範囲に含まれていた. しかし, 1 年以上の予測に関しては, 最頻値が正解クラスの範囲を大きく外していることや, 予測値の分布が大きく広がっていることが見られた. 今後の課題として, データ不足を解消するために, データ拡張, 他の教師なし学習の手法を検討することや, 非生存患者の生存期間だけでなく, 生存患者の観察期間を考慮して学習する手法の検討があげられる. また, 年齢, 性別, 病歴などの臨床情報と病理画像の両方を使用した, 予測精度の改善について検討することが考えられる.

目次

はじめに	i
第 1 章 緒言	1
1.1 研究背景	1
1.2 関連研究	2
1.3 研究目的	2
第 2 章 準備	4
2.1 データセット	4
2.2 ResNet18	5
2.3 Auto Encoder	6
2.4 SimCLR	7
2.5 Transformer	8
2.6 ラベル分布学習	9
2.7 Mean-Variance Loss	9
第 3 章 提案手法	11
3.1 パッチ抽出	12
3.2 事前学習	14
3.3 クラス分類モデル	17
第 4 章 実験	19
4.1 データセット	19
4.2 実験条件	20
4.3 実験結果	23
4.4 考察	27
4.5 追加実験	28

第 5 章	結言	32
5.1	まとめ	32
5.2	今後の課題	32
付録 A	付録	33
A.1	プログラム	33
A.2	プログラムの詳細	33
謝辞		34

第 1 章

緒言

1.1 研究背景

軟部組織腫瘍は全身の軟部組織（臓器、骨、皮膚を除いた組織で筋肉、結合組織、脂肪、血管、リンパ管、関節、神経など）に発生する腫瘍で、ほとんどが良性である。ヒトにおいて、悪性軟部腫瘍は、すべてのがんの中で約 1 % 程度であり、主ながんと比べると、非常に稀ながんであるとされている。悪性軟部腫瘍の発症原因については、世界中で研究されているが、よくわかっていないことも多く、一部の悪性軟部腫瘍に関しては遺伝性が認められている。症状として、初期は痛みを伴わないしこりが軟部組織に発生し、徐々に大きくなっていくにつれて、神経が圧迫され痛みが生じる。腫瘍の大きさは、良悪性を判定する上で、一つの指標となりうる。腫瘍が 5cm よりも大きい場合、組織の深部に発生して痛みや炎症を伴う場合、癒着している場合は悪性軟部腫瘍の可能性が高いと考えられる。これに対して腫瘍が 5cm 以下の場合には良性であることが多い。ただし、大きさだけで良悪性を判断することは難しいため、早期発見、早期治療が期待されている。診断とその治療には、X 線検査、形態学的特徴、分子所見、個々の患者の特徴を考慮する必要があるが、とりわけ病巣部細胞を観察する病理組織診断は、組織学的な悪性度の評価や、術後治療計画の決定に使用されるため、不可欠であるとされている [1]。しかし、専門とする病理医や医療機関が少ないため、診断に時間がかかってしまうことが現状である。悪性軟部腫瘍の治療は、主に外科手術で行われ、再発や転移を防ぐために正常な部分を含めた広範切除術が必要になる。腫瘍のできた部分によっては、正常な神経や筋肉等と腫瘍を切除しなければならず、正常な組織が失われることや場合によっては、運動機能が低下してしまうことがある。

医用画像診断技術の進歩により、スライドガラス上の染色標本全体を高精細なデジタル画像（Whole Slide Image : WSI）に変換する技術が利用可能となった。従来は顕微鏡を使用して、スライドの小部分にフォーカスし、経験豊富な病理医が標本を検討、評価してい

た．この操作は他の部分に対しても行わなければならない場合が多く，負担がかかっていた．WSI を使用することにより，病理医が光学顕微鏡を使って手作業で組織標本から得たのと同程度の情報を，多数のスライドに対して自動的に取り込むことができるため，臨床，研究，教育において多くの即時的な応用が可能な，非常に強力なツールとなると考えられている [3]．また，WSI とテレコミュニケーションシステムを駆使して，全国の医療機関と共有することで，遠隔診断を行っている活用事例も存在している．悪性軟部腫瘍において，病理画像は術後治療計画を立てる際に用いられ，患者に負担のかかる化学療法を行うべきかどうかの判断に使用される．しかし，この判断は容易ではなく，病理画像に基づいた客観的な診断システムが求められている．

1.2 関連研究

近年，深層学習は目覚ましい発展を遂げており，画像認識の分野に関してもその限りではない．WSI から悪性度の高い部分のような診断価値の高い領域を抽出する研究や，WSI から切り出したパッチ画像を用いて深層学習を行うことで，そのパッチ画像の悪性確率を求めて，WSI 中の良性と悪性部分の領域分けを行っている研究がある．また，センチネルリンパ節の WSI に対して深層学習を用いることで，転移性乳がんを自動検出するもの [4] や，The Cancer Genome Atlas から得られた WSI に対して深層畳み込みニューラルネットワーク (inception v3) を学習させ，肺がんの種類を自動的に分類するものがある [5]．

このように WSI と深層学習を用いて領域抽出を行っている研究やがんの種類を自動検出する研究は多くなされているが，生存期間のような連続値に関連する研究はまだ十分にされていない．生存期間予測の研究として，WSI を用いて Graph Convolutional Network (GCN) により学習した生存リスク特徴を生存期間予測に適用することで，予測精度を向上させたものがある [2]．結果として，bladder urothelial carcinoma (BLCA) データセットに対して，Mean Absolute Error (MAE) は 123.2 日，Concordance-index (C-index) は 0.849 となり，最も良い他のモデルに比べて MAE は約 30 日減少し，C-index は 0.3 ポイント増加した．しかし，問題点として BLCA のような大規模なデータセットを用いていることがあげられる．骨軟部腫瘍は希少がんのため，より少ないデータセットで高い精度を出すことが求められる．

1.3 研究目的

本研究では，病理画像から正確な生存期間を推定することで，医師が合理的な治療計画を立てる定量的な指標を得ることを目的とする．この指標によって，病理医が行う治療法の決定にかかる時間の軽減を実現することで，医師の負担軽減につながり，患者の病状

に合わせた治療の実現が見込まれる．本研究では，少ないデータセットに対応するために，画像のみで事前学習を行い，クラス分類で生存期間を推定する．Auto Encoder (AE), Contrastive Learning (CL), Transformer Encoder を用いた 3 つの実験と，比較のために ResNet18 を用いた実験の計 4 つを行う．

第 2 章

準備

2.1 データセット

本研究では，三重大学医学部附属病院で取得された軟部腫瘍患者 44 名の病理画像を用いる．病理画像には，患者の生存・非生存と生存期間がラベル付けされている．なお，上記画像と予後データの使用については，三重大学医学部附属病院倫理委員会の承認を得ている．図 4.9 に，本研究で使用する軟部腫瘍の H&E 染色された軟部腫瘍病理画像の例を示す．同図中の 2 つの枠は病理医が手作業で設定したものであり，外側の枠は腫瘍領域全体，内側の枠は腫瘍領域の中でも悪性度の高い領域を示している．

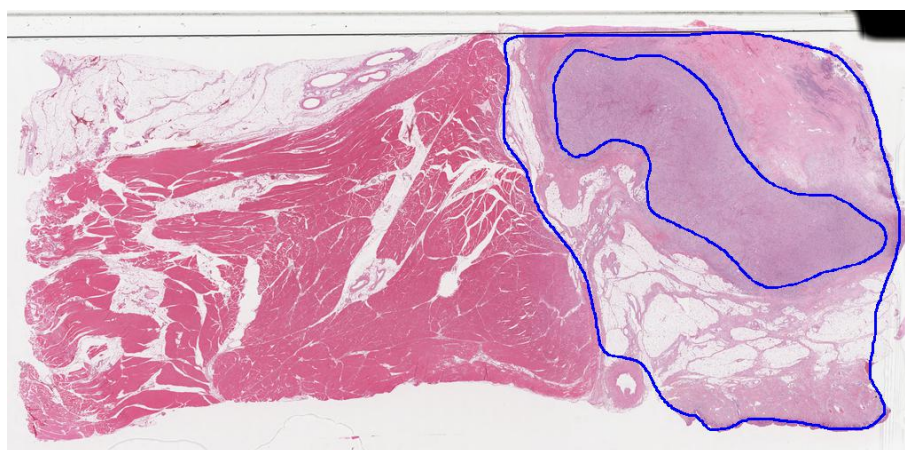


図 2.1: 病理画像のアノテーションの例

2.2 ResNet18

ResNet は畳み込みニューラルネットワークの一種である。層数が多いため詳細な特徴まで学習することができ、画像認識に強いモデルとされている。層数が多いと勾配を逆伝播させる際に最初の層まで伝わらない勾配消失の問題が起こるが、Shortcut Connection を追加することで解決している。このことにより、伝播してきた誤差を減衰することなくそのまま伝えることができる。本研究で使用する、ResNet18 [6] のアーキテクチャを図 2.2 に示す。

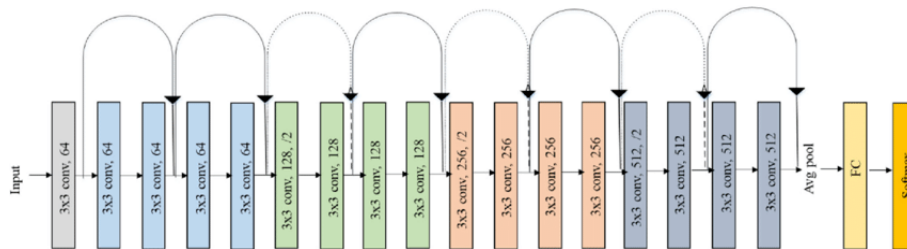


図 2.2: ResNet18

2.3 Auto Encoder

オートエンコーダは、ニューラルネットワークの仕組みの1つであり、本研究では、入力と同じ画像を復元できるように特徴抽出するモデルを使用する。任意の損失関数を用いて学習させることができるため、過学習を抑制し、より低次元の有効な特徴量を得やすいモデルを作成することが可能となる。図 2.3 に文献 [7] から引用したオートエンコーダを示す。図のエンコーダ部分では、わざと次元削減を行うことでデータを圧縮している。その際に、重要度に合わせてデータに重みづけをし、数値が低いものは除外される。学習していく過程で重みを更新していき、重要度の高い特徴のみが残るため、少ない特徴でも出力画像が復元できるようになる仕組みである。

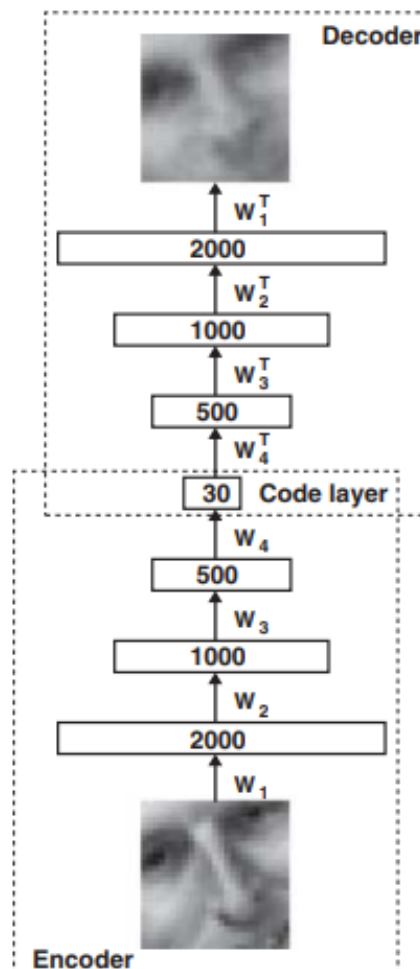


図 2.3: Auto Encoder

2.4 SimCLR

SimCLR [8] は対照学習のためのシンプルなフレームワークである。対照学習とは、正解ラベルを使用しない自己教師あり学習の1つであり、画像同士を比較して似た特徴を持つ画像は近くに、異なる特徴を持つ画像は遠くになるように学習する手法である。この手法は、コンピュータビジョンや自然言語処理（NLP）領域での活躍が期待されている。図 2.4 にアーキテクチャを示す。

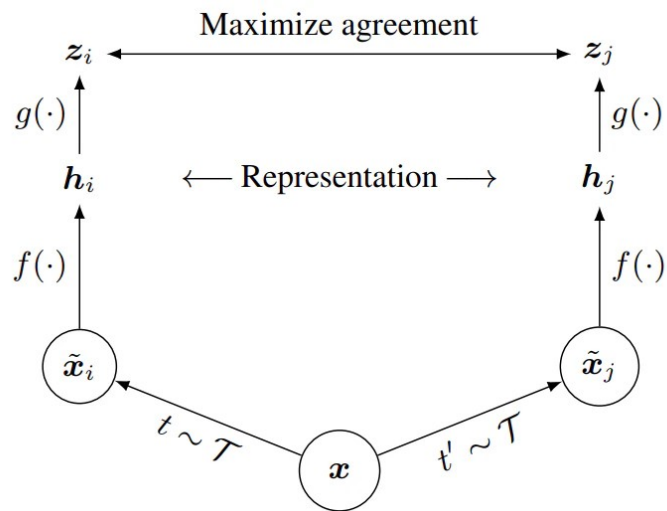


図 2.4: SimCLR

2.5 Transformer

Transformer [9] は、自然言語処理の分野でブレイクスルーを起こしたモデルである。Convolutional Neural Network (CNN) や Recurrent Neural Network (RNN) といった再帰や畳み込みの構造を用いず、Attention のみを使用している。Attention は、自然言語処理の分野で例えると、文中のある単語の意味を理解する際に、文中のどの単語に注目すればよいかを表すスコアである。近年、Transformer は自然言語処理の分野だけではなく、画像処理の分野でも使用されており、その汎用性が注目されている。図 2.5 にアーキテクチャを示す。

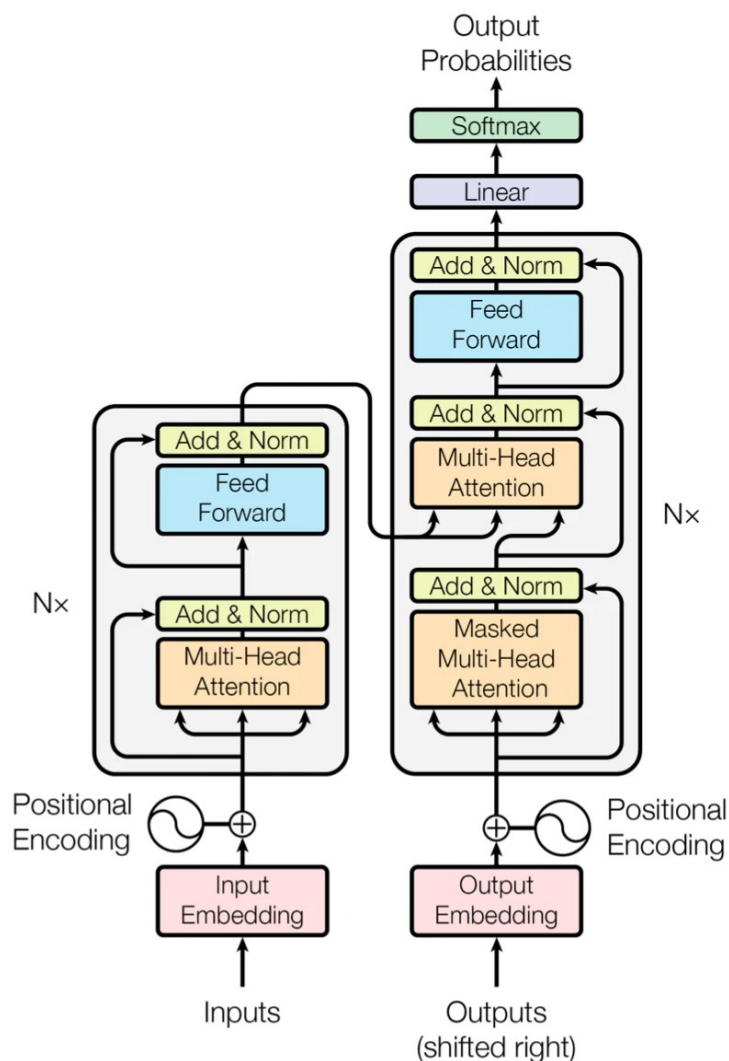


図 2.5: Transformer

2.6 ラベル分布学習

ラベル分布学習 [10] はクラス間の順序関係を考慮して学習する際に用いられる。通常、クラス分類問題には、Cross-Entropy 関数が用いられるが、正解クラスの予測確率のみを考慮するため、順序関係を持つクラス分類問題には適していない。そのため、本研究では One-hot なハードラベルから正解クラスを中心に広がりを持ったソフトラベルに変更して学習を行う。正解クラス c_t の i 番目の要素を式 2.1 に示す。

$$y_i|t = \frac{e^{-\phi(r_t, r_i)}}{\sum_{k=1}^K e^{-\phi(r_t, r_k)}} \quad (i = 1, \dots, K) \quad (2.1)$$

ここでクラスの集合を $C = c_1, \dots, c_N$ としてクラス c_i に対応する値を r_i , $K = |C|$ とする。今回のソフトラベルでは $r_i = i$, $\phi(r_i, r_j) = (r_i - r_j)^2$ としている。

2.7 Mean-Variance Loss

損失関数には、Mean-Variance Loss [11] を用いる。この関数は Mean Loss (L_m), Variance Loss (L_v), Softmax Loss (L_s) の 3 つの部分で足し合わせることで作られている。以下にその式を式 2.2 から式 2.7 に示す。

$$L = L_s + \lambda_1 L_v + \lambda_2 L_m \quad (2.2)$$

$$L_s = P_i \log \frac{P_i}{Q_i} \quad (2.3)$$

ここで、 N はバッチサイズ、 P_i は正解の分布、 Q_i は予測された分布を示す。また、 L_s は使用する正解ラベルによって変化し、ハードラベルであれば Cross Entropy Loss、ソフトラベルであれば Kullback-Leibler Divergence Loss を示す。

$$L_m = \frac{1}{2N} \sum_{i=1}^N (m_i - y_i)^2 \quad (2.4)$$

L_m はバッチごとの平均値を示す。

$$L_v = \frac{1}{N} \sum_{i=1}^N v_i \quad (2.5)$$

L_v はバッチごとの分散値を示す。

$$m_i = \sum_{j=1}^K j \cdot p_{ij} \quad (2.6)$$

$$v_i = \sum_{j=1}^K p_{ij} (j - m_i)^2 \quad (2.7)$$

m_i は i 番目のクラスの平均値, v_i は i 番目のクラスの分散値, $j \in \{1, 2, \dots, K\}$ はクラスラベル, p_{ij} はそのクラスの予測確率を示す.

第 3 章

提案手法

生存期間のような連続値を扱う場合，一般的に回帰問題として学習される．しかし，クラス分類問題に対して，データセットが多く必要であるという問題点がある．軟部腫瘍患者の数は限られているおり，データセットに限りがあるため，本研究では，クラス分類問題として生存期間を C 個のクラスに分割して，それぞれの出力をクラスの所属度として学習する．提案手法の大まかな流れを図 3.1 に示す．本手法では，前処理としてパッチ抽出を行い，事前学習した Auto Encoder (AE)，Contrastive Learning (CL)，Transformer Encoder を用いた 3 つの実験と，比較のために ResNet18 を用いた実験の計 4 つを行う．

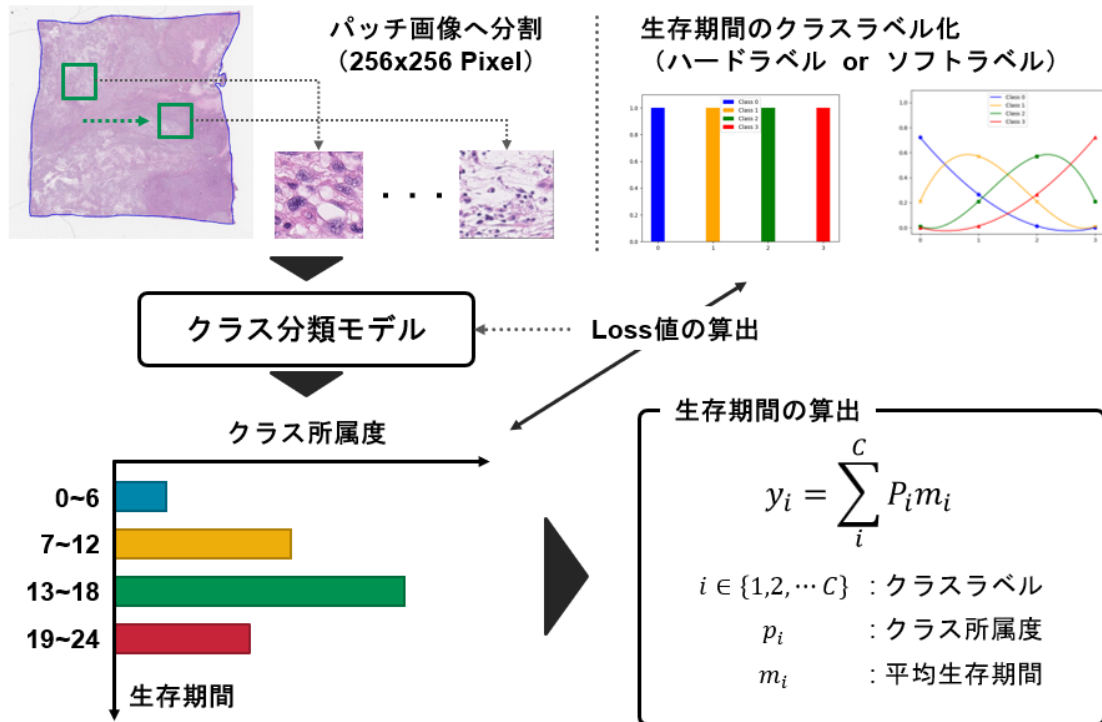


図 3.1: 提案手法の流れ

3.1 パッチ抽出

病理画像は非常に大きいため、CNN (Convolutional Neaural Network) の入力に直接使用することができない。そのため、小領域のパッチ画像に分割して用いる。パッチ抽出はスライドウィンドウ方式を用いて、サイズは 256×256 、ストライドは 256×256 で抽出する。各パッチ画像に生存・非生存と生存期間を付与して使用する。図 3.2 に切り出したパッチ画像を示す。また、表 3.1 と表 3.2 に生存、非生存の被験者ごとの、生存期間もしくは観察期間とアノテーション領域ごとのパッチ画像の枚数を示す。

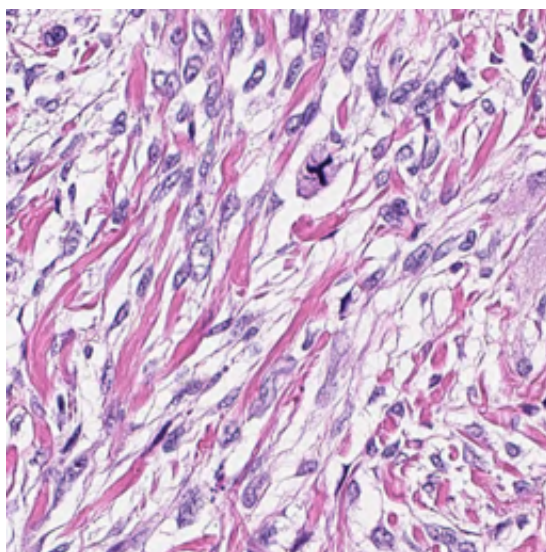


図 3.2: パッチ画像

表 3.1: 生存患者の生存期間と領域ごとの枚数

被験者 No.	観察期間	悪性度の高い領域	その他の腫瘍領域
61-5	0.983606557	48,910	113,677
50-23	2.262295082	31,569	24,407
52-5	9.672131148	23,446	19,038
42-5	13.70491803	5,997	33,826
58-7	16.55737705	48,965	19,920
47-7	20.26229508	5,018	22,655
62-2	30.49180328	60,284	39,195
37-2	30.81967213	10,643	34,802
30-4	34.2295082	1,986	18,043
116-2	36.19672131	4,883	38,789
54-6	39.14754098	7,835	14,808
46-2	54.6557377	13,483	21,315
43-11	55.21311475	15,602	25,476
115-12	58.8852459	21,023	22,594
41-12	59.3442623	1,483	22,813
45-5	59.47540984	3,794	34,748
44-8	60.16393443	16,634	17,162
39-4	61.7704918	13,914	5,297
36-3	62.2295082	23,326	1,829
118-4	63.14754098	13,082	1,845
31-2	66.2295082	10,729	23,933

表 3.2: 非生存患者の生存期間と領域ごとの枚数

被験者 No.	生存期間	悪性度の高い領域	その他の腫瘍領域
59-19	1.901639344	5,107	104,506
2-4	3.967213115	2,355	20,976
51-1	5.180327869	10,875	31,309
112-1-B	8.06557377	8,366	1,493
34-11	9.37704918	2,746	46,099
1-2	10.03278689	13,927	8,614
120-7	11.18032787	17,003	13,603
14-6	11.67213115	14,087	18,644
11-1	16.55737705	6,406	6,256
21-6	18.16393443	5,634	14,743
57-10	20.68852459	25,032	155,646
13-1	25.37704918	8,072	21,280
121-3	25.96721311	13,947	18,012
110-3	29.57377049	17,880	23,432
95-8	33.83606557	1,707	27,143
113-21	34.09836066	1,894	21,266
122-9	35.27868852	5,448	17,500
38-10	35.73770492	13,139	8,103
60-8	36.6557377	3,240	26,183
91-17	37.57377049	14,977	14,279
32-9	37.60655738	4,796	25,726
55-12	38.78688525	6,341	19,314
53-7	43.57377049	1,263	11,064

3.2 事前学習

本研究では、データセットの数が少なく、過学習を起こしやすいという問題点がある。そこで、事前学習を行うことによってこれを防ぐこととした。また、正解ラベルの分布に過度に近づきすぎることを防ぐため、学習時の学習パラメータの数を減らして画像のみを用いた教師なし学習を行う。以下で事前学習を行ったモデルの説明を行う。

3.2.1 Auto Encoder

1つ目は, Auto Encoder (AE) である. エンコーダ部分には, 未学習の ResNet18 の畳み込み層を用いて, デコーダ部分には ConvTranspose2d 層, ReLU 層, BatchNorm2d 層の3層を7セット結合したものを使用する. $3 \times 256 \times 256$ の RGB 画像を入力として用いて, $512 \times 1 \times 1$ の特徴量に圧縮して学習する. このモデルを使用する理由として, 画像の視覚特性を学習できるため, 生存期間予測に有益な特徴量を得られるのではないかと考えたからである. 本研究で使用する Auto Encoder モデルを図 3.3 に示す.

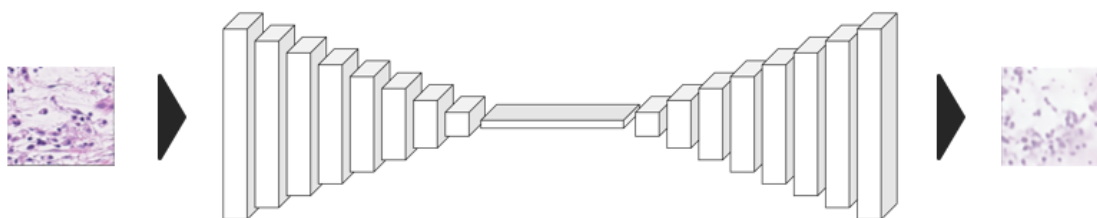


図 3.3: Auto Encoder

3.2.2 Contrastive Learning

2つ目は, Contrastive Learning (CL) を行うモデルである. これは, 未学習の ResNet18 の全結合層 (Fully-Connected ; FC) 層を, 出力が 128 次元の FC 層に変更して用いる. 入力は, $3 \times 256 \times 256$ の RGB 画像と, それにランダムな画像変換 (ガウスぼかし, クロッピング, 色変換等) を適用した画像の 2 枚を用いる. これらを正のペアとし, 正のペア以外を負のペアとする. 正のペアから得られた特徴量は \cos 類似度という指標を用いて, 最大となるように, 負のペアであれば類似度が最小となるように学習する. 以下の式 3.1 に SimCLR [8] で定義されている正のペアに対する損失関数を示す. また, 図 3.4 に本研究で使用する Contrastive Learning のモデルを示す.

$$l_{i,j} = -\log \frac{\exp(\text{sim}(z_i, z_j)/\tau)}{\sum_{k=1}^{2N} 1[k \neq i] \exp(\text{sim}(z_i, z_k)/\tau)} \quad (3.1)$$

ここで, $1[k \neq i] \in \{0, 1\}$ は $k \neq i$ のときに 1 となる定義式であり, τ は温度パラメータ, z_i は i 番目のデータのネットワークにより変換された特徴量である. $\text{sim}(z_i, z_j)$ は L_2 ノルムで正規化された z_i と z_j との内積 (コサイン類似度) を表す. データ数 N は変換フィルターにより 2 倍の $2N$ に増幅されている.

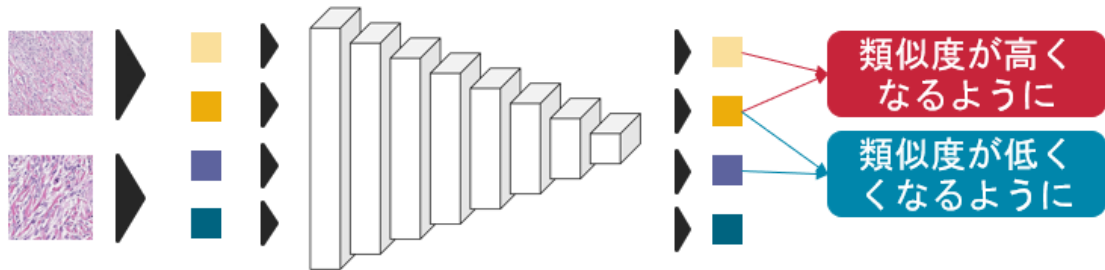


図 3.4: Contrastive Learning

3.3 クラス分類モデル

本研究では，クラス分類問題として生存期間を C 個のクラスに分割して，それぞれの出力をクラスの所属度として学習する．

まず，3.2 節で説明した事前学習済みモデルを用いてクラス分類モデルを構築する．先ほど挙げた 2 つのモデル（AE，CL）のほかに，比較対象として ImageNet と呼ばれる 100 万枚の自然画像で学習済みの ResNet18 を使用する．ResNet18 の学習では，ImageNet の画像サイズに合わせるために， $3 \times 224 \times 224$ の画像にリサイズして，平均と分散は，それぞれ ImageNet で使用されているものに合わせた．

その後，AE のデコーダ部分，CL の FC 層の部分を FC 層，Dropout 層，BatchNorm1d 層，ReLU 層の 4 層を 1 セットとして 3 セットつなげたものに置き換えて使用する．図 3.5 に本研究で使用するクラス分類モデルを示す．

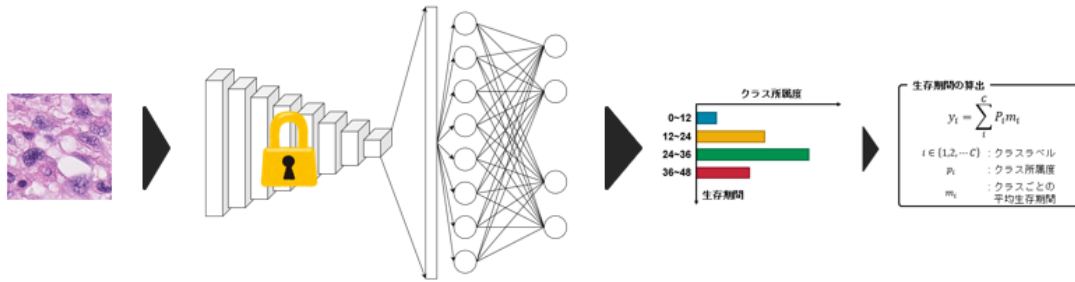


図 3.5: クラス分類モデル

また、4つ目のモデルとして、事前学習済みのモデルから得られた特徴量を Transformer Encoder に用いてクラス分類モデルを作成する。入力としてパッチ画像をまとめてスタックとし、バッチサイズ×画像のスタック数×特徴量の3次元データを用いる。このモデルを使用する目的として、病理画像全体から悪性度の高い部分とそうでない部分の移り変わりを学習することにある。悪性度の高低を学習するために、AE から得られた特徴量に対してクラスタリングを行い、位置情報として特徴量に付与する。その後、Transformer Encoder の後ろに先ほどの FC 層、Dropout 層、BatchNorm1d 層、ReLU 層の3セットを追加してクラス分類を行う。図 3.6 に本研究で使用する Transformer Encoder を用いたクラス分類モデルについて示す。

それぞれのモデルを学習した後、その出力を Softmax 関数を用いて確率に変換し、式 3.2 により生存期間を算出する。

$$\sum_i^C p_i m_i \quad (3.2)$$

ここで、 $i \in 1, 2, \dots, C$ はクラスラベルを示す。 p_i は i 番目のクラスの予測確率であり、 m_i は i 番目のクラスに含まれる学習データの平均生存期間である。

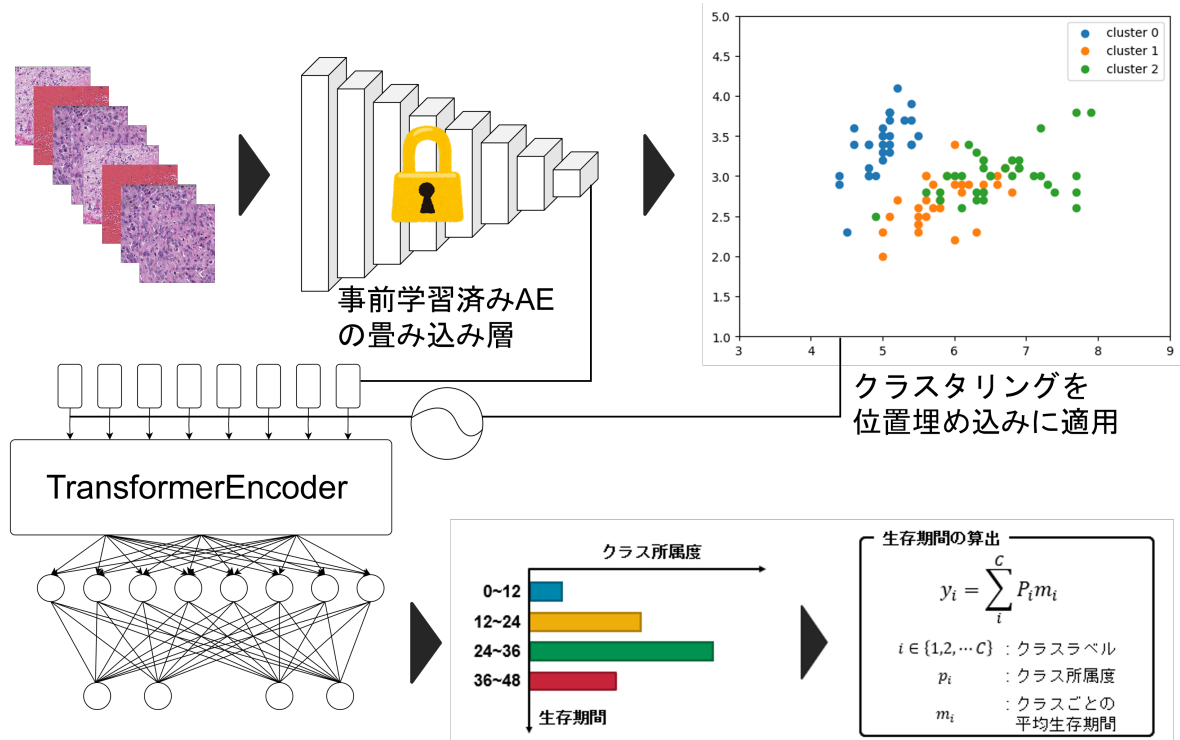


図 3.6: Transformer Encoder を用いたクラス分類モデル

第 4 章

実験

4.1 データセット

表 4.1 に患者数，病理画像枚数，分割した小領域画像の枚数をそれぞれ示す．小領域画像の枚数はラベル付けされた生存，非生存に分け，また悪性度の高い領域とそれ以外の腫瘍領域に分けて示す．

表 4.1: データセット

	生存	非生存
患者数	21	23
病理画像枚数	21	23
小領域画像枚数（高悪性度の領域）	382,606	138,748
小領域画像枚数（その他の腫瘍領域）	556,172	568,120

4.2 実験条件

AE と CL は，生存，非生存の両方を用いて，悪性度の高い領域とその他の腫瘍領域の小領域画像を使用する．その画像枚数を表 4.2 に示す．学習には，RAdam optimizer [12] を使用し，学習率は 0.001 として，バッチサイズは 32 とした．Transformer Encoder は，エンコーダの特徴量を 128，ヘッド数を 8，layer 数を 6 とした．また，クラスタリングは K-Means 法を用いて行い，クラスタ数は 4 とした．バッチサイズは 8，スタック数は 8 で学習を行う．その後の生存期間予測では，0 から 48 か月の区間を 12 か月幅で区切り，4 クラス分類を行った．学習には，RAdam optimizer を使用し，学習率は 0.0001 とした．またバッチサイズは 32 とし，非生存の小領域画像の内，悪性度の高い領域のみを使用して，各クラスの領域画像枚数が同程度になるように調整をした．AE と CL，また生存期間予測に使用した画像枚数を表 4.3 に示す．また，実験に使用する．ハードラベルとそれを式 2.1 によって変換したソフトラベルを図 4.1，図 4.2 に示す．

表 4.2: パラメータ

	バッチサイズ	最適化関数	損失関数	学習率
AE	32	RAdam	MSE	0.001
CL	32	RAdam	Contrastive Loss	0.001
Transformer Encoder	8	RAdam	Mean-Variance Loss	0.001
クラス分類モデル	32	RAdam	Mean-Variance Loss	0.0001

表 4.3: パッチ画像使用枚数

使用する病理画像の領域	AE & CL		クラス分類モデル	
	学習パッチ画像	検証パッチ画像	学習パッチ画像	検証パッチ画像
腫瘍領域全体	1,143,050	497,503	0	0
高悪性度の領域	0	0	118,708	65,494

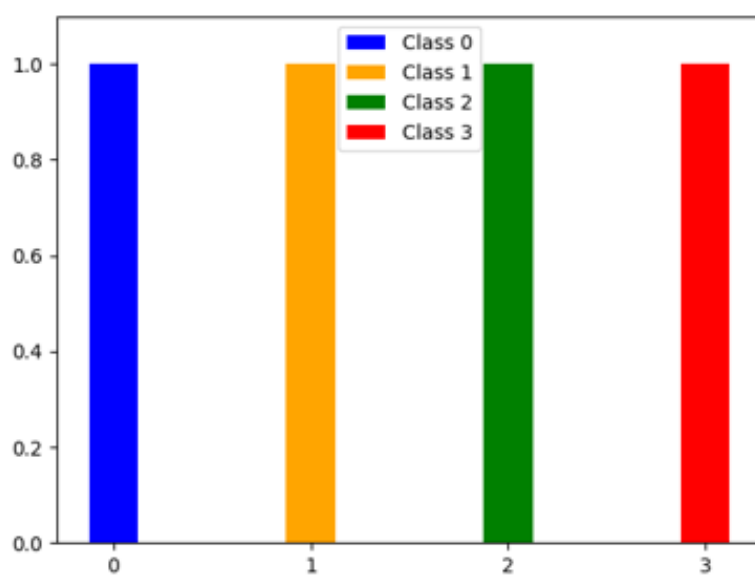


図 4.1: ハードラベル



図 4.2: ソフトラベル

4.2.1 評価方法

生存期間の評価には Mean Absolute Error (MAE) と Concordance-index (C-index) を用いる. MAE は予測値と正解値の誤差を計算するもので, 単位が変わらないため直感的に数値を理解しやすいという利点があるため使用した. また, C-index はミニバッチ内でラベル付けされた生存期間どうしを比較した際に, 予測された生存期間どうしの順序関係がどの程度一致しているかを示すものである. 以下式 4.1 と式 4.2 に MAE と C-index の式を示す.

$$MAE = \frac{1}{N} \sum_{i=1}^N |\hat{y}_i - y_i| \quad (4.1)$$

ここで, N はデータ数, y_i は真値の生存期間, \hat{y}_i は予測された生存期間を表す.

$$C_{index} = \frac{1}{K} \sum_{i=1}^P \delta_i \sum_{j: \tau_i < \tau_j} \mathbb{I}_{\hat{\tau}_i < \hat{\tau}_j} \quad (4.2)$$

ここで, K は比較可能なペアの数, P はデータ数, δ はイベントが起きたかどうか, τ は真値の生存期間, $\hat{\tau}$ は予測された生存期間を表す. また, $\mathbb{I}_{\hat{\tau}_i < \hat{\tau}_j}$ は真なら 1, 偽なら 0 を取る.

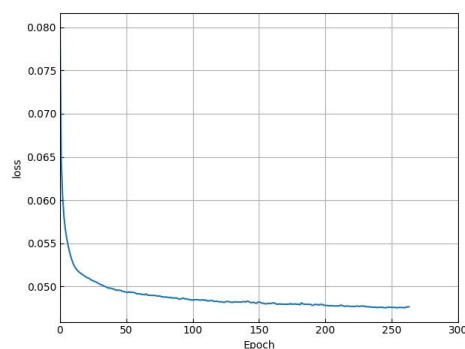
4.3 実験結果

4.3.1 事前学習

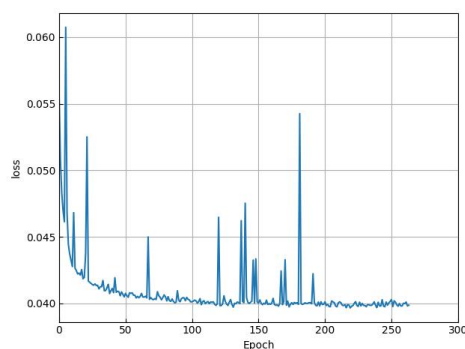
表 4.4 に AE と CL の事前学習の実験結果を示す．また，AE の事前学習の損失関数のグラフを図 4.3，CL の事前学習の損失関数のグラフを図 4.4 に示す．この結果から，AE と CL は検証データの損失値が最小となった Epoch 数 219 と Epoch 数 55 をそれぞれのクラス分類モデルで使用する事とする．

表 4.4: 事前学習（AE と CL）の実験結果

model	Loss	Train_min	Valid_min
Auto Encoder	MSE	0.047	0.039 (Epoch 219)
Contrastive Learning	ContrastiveLoss	3.05	3.42 (Epoch 55)

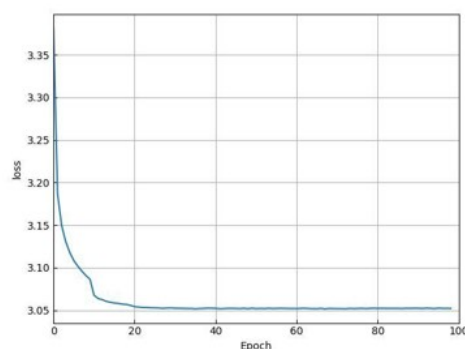


(a) 学習データの損失値 (AE)

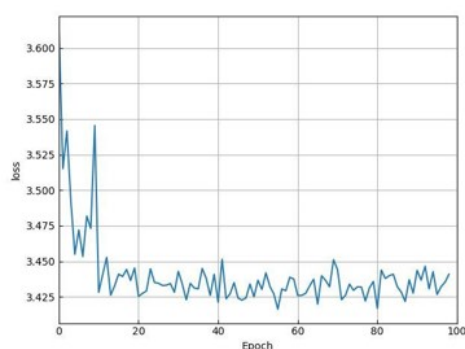


(b) 検証データの損失値 (AE)

図 4.3: AE の損失値のグラフ



(a) 学習データの損失値 (CL)



(b) 検証データの損失値 (CL)

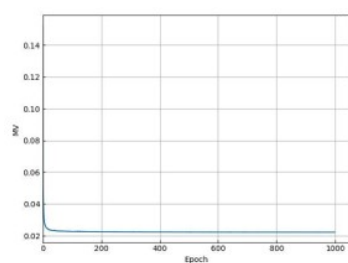
図 4.4: CL の損失値のグラフ

4.3.2 クラス分類モデル

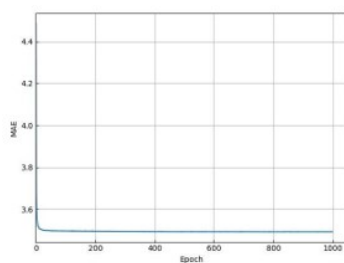
表 4.5 に学習済みの ResNet18, AE, CL の 3 つのモデルでハードラベル (Hard), ソフトラベル (Soft) をそれぞれ用いた時の MAE と C-index の値を示す. 最小の MAE は CL で学習した畳み込み層を用いてソフトラベルで学習したもので, 7.69 か月の誤差となり, 最大の C-index の値は学習済みの ResNet18 を用いてハードラベルで学習した場合で, 0.817 となった. また, CL で学習した畳み込み層を用いてソフトラベルで学習した場合の学習データと検証データの損失関数, MAE, C-index の推移を図 4.5 と図 4.6 に示す.

表 4.5: 3 つのモデルのラベルごとの比較

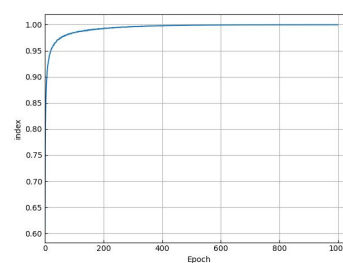
	Label	MAE (ヶ月)	C-index (Max:1)
ResNet18	Hard	8.26	0.817
	Soft	7.87	0.814
AutoEncoder	Hard	10.1	0.643
	Soft	9.98	0.658
ContrastiveLearning	Hard	8.78	0.783
	Soft	7.69	0.807
TransformerEncoder	Hard	11.1	0.502
	Soft	11.1	0.502



(a) Mean-Valiance Loss

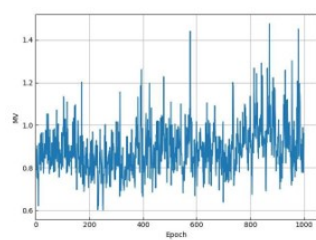


(b) Mean Absolute Error

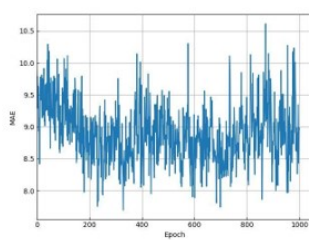


(c) C-index

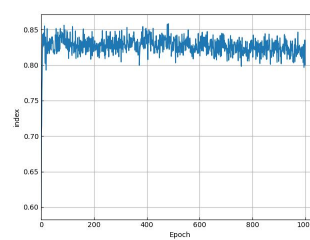
図 4.5: 学習データ



(a) Mean-Valiance Loss



(b) Mean Absolute Error



(c) C-index

図 4.6: 検証データ

4.4 考察

図 4.7 に MAE が最小値を取ったモデルの正解クラスごとに出力された予測値とその頻度のグラフを示す。クラス 0 やクラス 1（赤い分布や青い分布）の生存期間が 1 年以内のパッチ画像については、最頻値がその正解クラスの範囲に含まれており、誤差は少なくなっていると考えられる。しかし、生存期間が 1 年以上のパッチ画像については、最頻値を範囲から大きく外れていることや分布が広がりすぎていることが読み取れる。これは、病理画像中の悪性度の高い部分のみを生存期間予測に使用しているが、生存期間の長い病理画像では、その中でも悪性度のそこまで高くない部分が存在しているため、予測を大きく外すパッチ画像が出てきているのではないかと考えられる。この課題を解決するために、病理画像全体から悪性度の高低を考慮して学習することや、病理画像に対して腫瘍の大きさを考慮して学習する必要があるのではないかと考えられる。

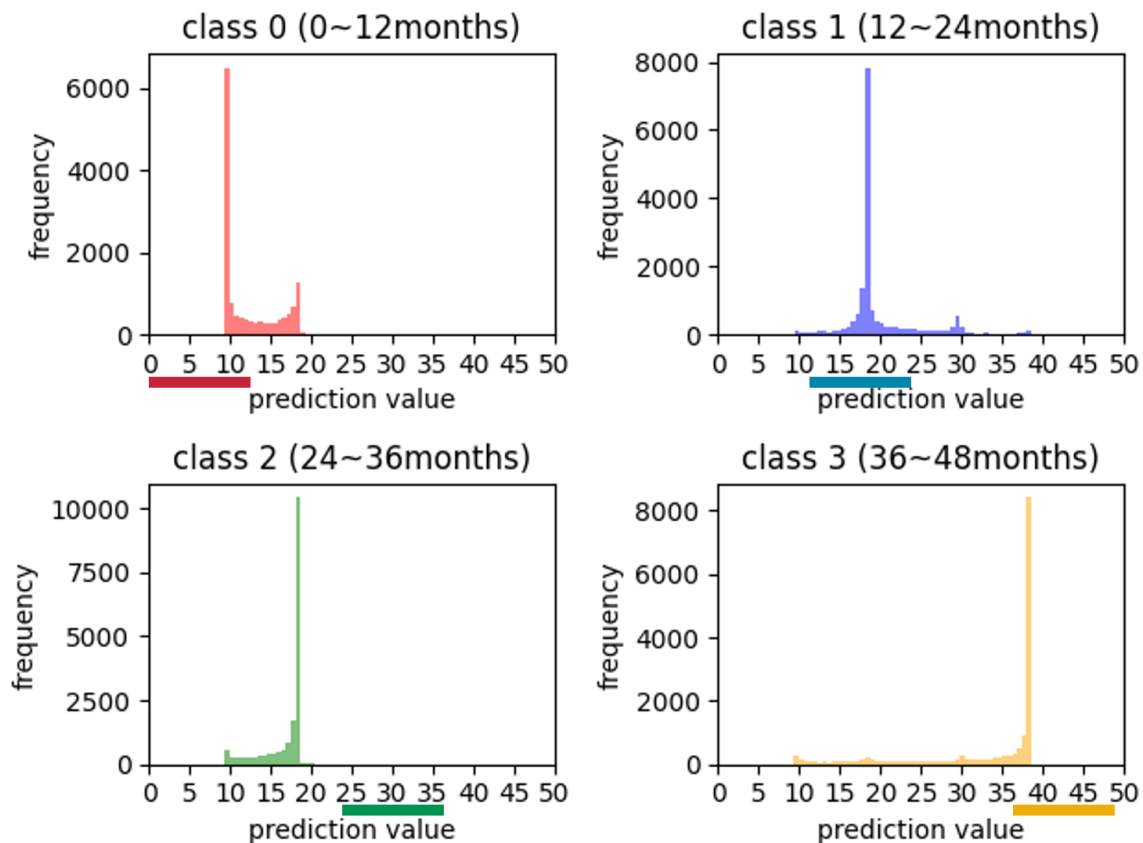


図 4.7: クラスごとの予測値が現れる頻度

4.5 追加実験

上記の実験では、クラスを正確に予測できても大きく誤差が出てしまう場合がある。例えば、クラス 0 (0 ヶ月から 12 ヶ月) が正解の場合、式 3.2 から推定誤差が 0 となる予測値は約 6 ヶ月であるが、正解の生存期間が 12 ヶ月の場合、推定誤差が 6 ヶ月となってしまう。

そこで、正解クラスをより細かく分割する。具体的には 0 ヶ月から 12 ヶ月としていたクラス 0 を 0 ヶ月から 6 ヶ月と、6 ヶ月から 12 ヶ月の 2 つのラベルに分けて計 8 つのラベルを用いた実験を行う。実験で用いるラベルの例を以下の図 4.8 に示す。

また、同様の理由で生存期間を範囲で区切るのではなく、その値から逆算してラベルを作成する実験を行う。このラベルはその値が含まれる範囲の平均生存期間とその右隣の範囲の平均生存期間からの距離の逆比を用いて作成した。具体例として正解の生存期間が 9 ヶ月の時、クラス 0 (0 ヶ月から 12 ヶ月) の範囲に含まれ、その値が含まれる範囲 (0 ヶ月から 12 ヶ月) の平均生存期間とその右隣の範囲 (クラス 1: 12 ヶ月から 24 ヶ月) の平均生存期間からの距離を計算すると、それぞれ $9 - 6 = 3$ と $18 - 9 = 9$ となる。この時、逆比を取ることによって、3 対 1 とし、クラス 0 の値を $3/4 = 0.75$ 、クラス 1 の値を $1/4 = 0.25$ としてラベルを作成する。正解の生存期間が 9 ヶ月の時のラベルを以下の図 4.9 に示す。実験結果を表 4.6 と表 4.7 に示す。

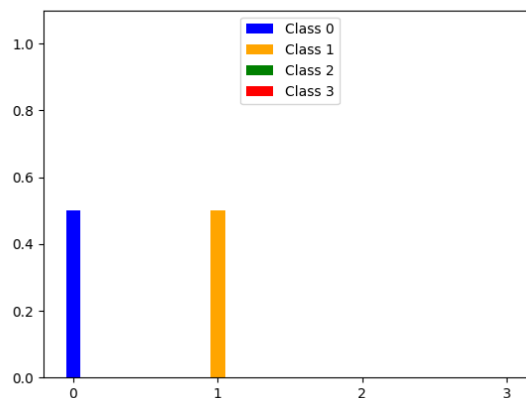


図 4.8: 正解の生存期間が 6 ヶ月から 12 ヶ月の範囲で使用するラベル

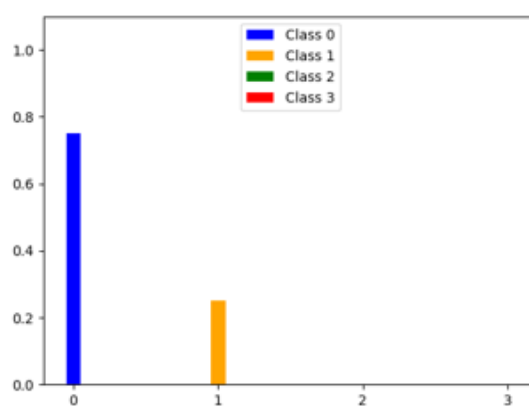


図 4.9: 正解の生存期間が 9 ヶ月の時に使用するラベル

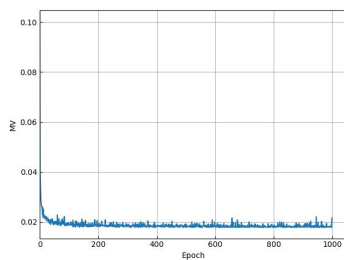
表 4.6: 範囲を狭めた時のラベルを使用した結果

	MAE (ヶ月)	C-index (Max:1)
ResNet18	5.80	0.912
Auto Encoder	6.90	0.824
Contrastive Learning	5.41	0.894
Transformer Encoder	11.1	0.573

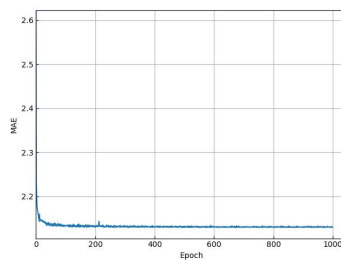
表 4.7: 値から逆算して作成したラベルを使用した結果

	MAE (ヶ月)	C-index (Max:1)
ResNet18	6.50	0.893
Auto Encoder	7.44	0.840
Contrastive Learning	6.02	9.09
Transformer Encoder	11.3	0.499

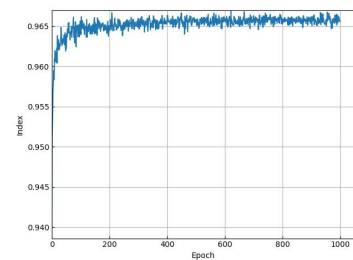
最も良い MAE の値は CL の範囲を狭めたラベルを用いた時で、最小値で 5.41 ヶ月となった。表 4.5 と比較すると Transformer Encoder を用いたモデル以外は MAE, C-index ともに良くなっており、真値と予測値との誤差の値は小さく、真値の大小関係を考慮して予測値の大小関係を予測できているといえる。CL で学習した畳み込み層を用いて範囲を狭めたラベルで学習した場合の学習データの損失関数, MAE, C-index の推移を図 4.10 と図 4.11 に示す。また図 4.12 に MAE が最小値を取ったモデルの正解クラスごとに出力された予測値とその頻度のグラフを示す。図 4.7 と比較すると、クラス 2 (緑の分布) で生存期間の短い部分によっていた分布が正解範囲に含まれるようになっており、改善していることがわかる。しかし、クラス 1 (青の分布) に関しては、正解範囲の中心からずれた位置に最頻値があり、こちらは悪化していることがわかるが大きく外していないことから結果が CL のソフトラベルを用いる場合に比べてよくなったのではないかと考えられる。今後は、生存期間の範囲をより狭めたラベルを使用していくことや値を中心として正規分布に従うように広がりを持つソフトラベルを用いる方法を試していくことがあげられる。



(a) Mean-Valiance Loss

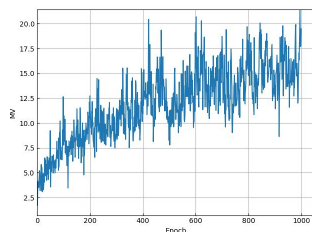


(b) Mean Absolute Error

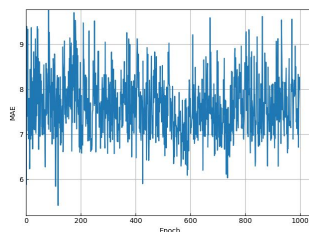


(c) C-index

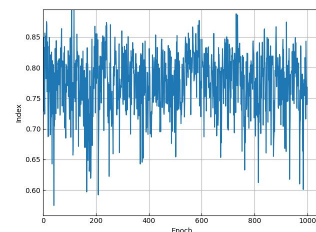
図 4.10: 学習データ



(a) Mean-Valiance Loss



(b) Mean Absolute Error



(c) C-index

図 4.11: 検証データ

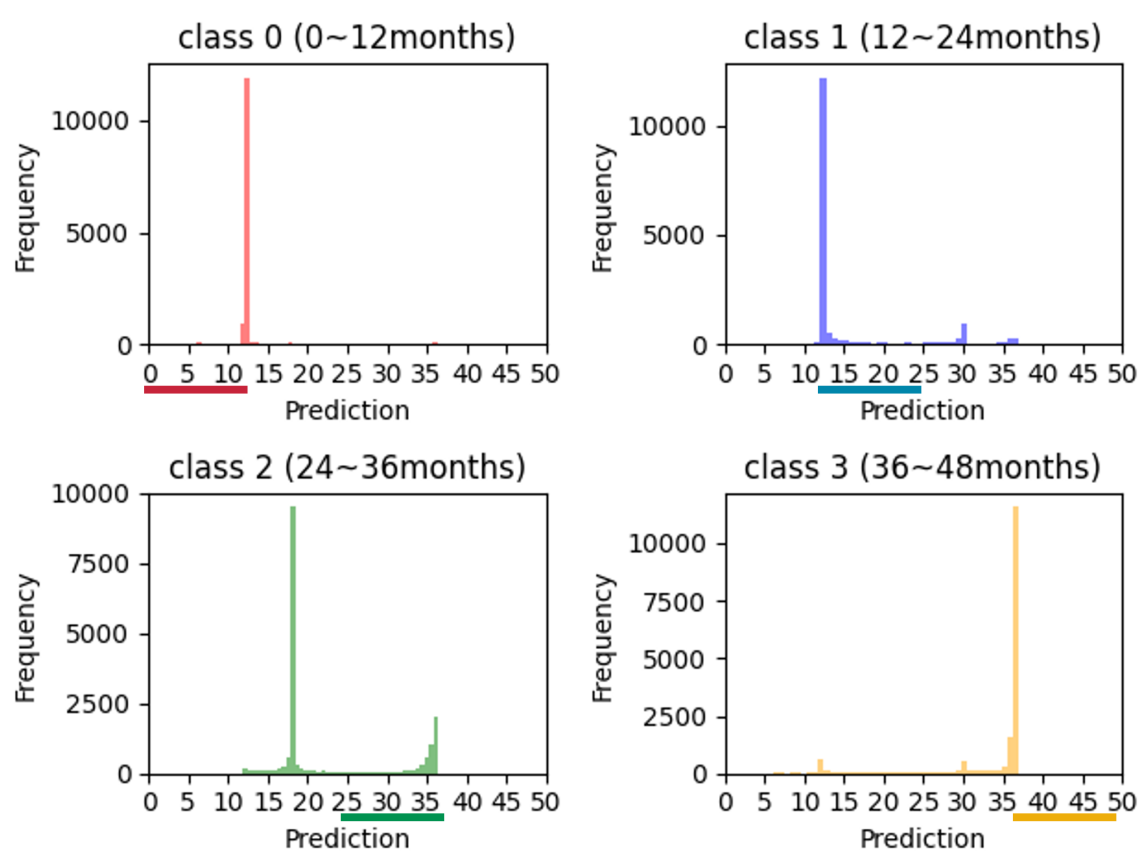


図 4.12: クラスごとの予測値が現れる頻度

第 5 章

結言

5.1 まとめ

本研究では、悪性軟部腫瘍患者の生存期間予測を様々なモデルを使用して行った。少ないデータで学習するために、事前学習とクラス分類を用いた。結果として、最も良いものは、Contrastive Learning で事前学習を行ったモデルで、非生存患者の生存期間の誤差は、最小値で 7.69 か月だった。生存期間が 1 年以内の予測に関しては、予測値の最頻値が分割した正解クラスの範囲に含まれていた。しかし、1 年以上の予測に関しては、最頻値が正解クラスの範囲を大きく外していることや、予測値の分布が大きく広がっていることが見られた。

5.2 今後の課題

今後の課題として、データ不足を解消するために、データ拡張、他の教師なし学習の手法を検討することや、非生存患者の生存期間だけでなく、生存患者の観察期間を考慮して学習する手法や、年齢、性別、病歴などの臨床情報と病理画像の両方を使用した、予測精度の改善について検討することが考えられる。

付録 A

付録

A.1 プログラム

本研究に関するプログラムはすべて以下のディレクトリ

- /net/nfs2/export/home/nonaka/root/workspace/mie-pathology

本研究に関するデータセットはすべて以下のディレクトリ

- /net/nfs2/export/dataset/mie-ortho/pathology

に存在している.

A.2 プログラムの詳細

- /net/nfs2/export//home/nonaka/root/workspace/mie-pathology/survival_time

以下に存在する survival_time.py を実行する. 詳しい実行方法は README を参照してください.

謝辞

若林哲史教授，盛田健人助教，白井伸宙助教には，的確なご指摘，助言，論文執筆をご教授いただきました．吉永みゆき事務員には日々の研究活動においてお世話になりました．心から感謝申し上げます．また，日頃からアドバイスをしていただいた研究室の先輩方，日々の活力をいただいた同期の皆様に感謝いたします．そして最後に，私の大学生活を支えてくれた家族に今一度感謝の意を示し，本論文の結びといたします．

参考文献

- [1] Ting Liu, Krishna K Narra, Phillip D Stephenson, Allie H Grossmann, Erinn Downs-Kelly, and Daniel Albertson. Pathology of soft-tissue tumors. *Sarcoma Oncology: A Multidisciplinary Approach*, p. 15, 2019.
- [2] Zhixin Xu, Seohoon Lim, Hong-Kyu Shin, Kwang-Hyun Uhm, Yucheng Lu, Seung-Won Jung, and Sung-Jea Ko. Risk-aware survival time prediction from whole slide pathological images. *Scientific Reports*, Vol. 12, No. 1, p. 21948, 2022.
- [3] Neeta Kumar, Ruchika Gupta, and Sanjay Gupta. Whole slide imaging (wsi) in pathology: current perspectives and future directions. *Journal of digital imaging*, Vol. 33, No. 4, pp. 1034–1040, 2020.
- [4] Dayong Wang, Aditya Khosla, Rishab Gargeya, Humayun Irshad, and Andrew H. Beck. Deep learning for identifying metastatic breast cancer, 2016.
- [5] Nicolas Coudray, Paolo Santiago Ocampo, Theodore Sakellaropoulos, Navneet Narula, Matija Snuderl, David Fenyő, Andre L Moreira, Narges Razavian, and Aristotelis Tsirigos. Classification and mutation prediction from non-small cell lung cancer histopathology images using deep learning. *Nature medicine*, Vol. 24, No. 10, pp. 1559–1567, 2018.
- [6] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 770–778, 2016.
- [7] Geoffrey E Hinton and Ruslan R Salakhutdinov. Reducing the dimensionality of data with neural networks. *science*, Vol. 313, No. 5786, pp. 504–507, 2006.
- [8] Ting Chen, Simon Kornblith, Mohammad Norouzi, and Geoffrey Hinton. A simple framework for contrastive learning of visual representations. In *International conference on machine learning*, pp. 1597–1607. PMLR, 2020.
- [9] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. Attention is all you need. *Advances in*

- neural information processing systems*, Vol. 30, , 2017.
- [10] Raul Diaz and Amit Marathe. Soft labels for ordinal regression. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 4738–4747, 2019.
- [11] Hongyu Pan, Hu Han, Shiguang Shan, and Xilin Chen. Mean-variance loss for deep age estimation from a face. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 5285–5294, 2018.
- [12] Liyuan Liu, Haoming Jiang, Pengcheng He, Weizhu Chen, Xiaodong Liu, Jianfeng Gao, and Jiawei Han. On the variance of the adaptive learning rate and beyond. *arXiv preprint arXiv:1908.03265*, 2019.