

2019年度 修士論文

論文題目

2種類のCNNモデルを用いた細胞核の抽出と分類

(Cell Nucleus Extraction and Classification
Using Two CNN Models)

指導教員

舟橋 健司 准教授

名古屋工業大学大学院 工学研究科 情報工学専攻

2018年度入学 30414114番

満安 佑亮

目次

第 1 章	はじめに	1
第 2 章	背景知識	4
2.1	細胞核	4
2.2	癌細胞の特徴	4
2.3	メラノーマ	6
2.4	ヘマトキシリン・エオジン染色	6
2.5	Convolutional Neural Network	7
2.5.1	畳み込み層	7
2.5.2	プーリング層	10
2.5.3	全結合層	11
2.6	U-Net	12
2.7	VGG16	13
2.8	Fine-tuning	14
2.9	Data Augmentation	14
2.10	交差検証	15
第 3 章	細胞核の抽出および分類における提案手法	16
3.1	細胞核の抽出における提案手法	16
3.1.1	モデルの構造	17
3.1.2	データセットの作成	18
3.1.3	抽出方法	19
3.1.4	後処理	20
3.2	細胞核の分類における提案手法	20
3.2.1	モデルの構造	20
3.2.2	データセットの作成	21
3.2.3	細胞核分類の処理手順	22
第 4 章	細胞核の抽出および分類実験	23
4.1	細胞核の抽出実験	23
4.1.1	モデルの評価実験	23
4.1.2	病理画像を用いた抽出実験	24
4.2	細胞核の分類実験	28
4.2.1	モデルの評価実験	29
4.2.2	病理画像を用いた分類実験	31

第 5 章 おわりに	38
謝辞	39
参考文献	40
発表論文リスト	42

第1章 はじめに

近年、医療の分野において、新しい臨床検査法や画像装置の開発、診断精度の向上によって様々な病気に対する新しい治療法が開発されるようになってきている。こうした医療の進歩には目覚ましいものがあるが、依然として癌の脅威は去っていない。世界では2番目に多い死亡原因となっており、2018年には約1800万人が癌を発症し約960万人が癌で亡くなっている。中でも日本は癌が最も多い死亡原因であり一生のうちに2人に1人は何らかの癌になると言われている。また、癌による死亡率は人口の高齢化の影響を除いた場合では減少しているが、人口の高齢化を考慮すると年々増加し続けている。そのため、癌の早期発見は今後の日本の国民の寿命の延長には欠かせないものである。一般的に癌は加齢に伴い遺伝子の異常が蓄積していくことが原因で発症するが、細胞や組織を直接診察する細胞診・組織診は癌の早期発見することにおいてとても重要な役割を担っており、正確な病理学的診断は癌の治療計画における正しい選択のために必要不可欠であると言える [1]。これらの診断の大部分は依然として病理専門医の経験と技術に依存している。また、病理専門医一人一人の力量や診断基準が異なるため、悪性度の包括的な定量化が困難であり客観性に乏しいなどの課題がある。そのため、病院間で異なる診断がなされてしまうといった患者にとって不利益な結果につながる問題も起こってしまう。さらに、増え続ける癌患者数に対して日本では診断する病理専門医の数が2019年8月7日現在で2539人 [2] しかおらず、日本の人口およそ51000人あたりに1人と極端に不足している。これにより、病理専門医1人当たりの負担が増え、適切な診断が行えず診断のばらつきや誤診の増加が懸念される問題が起きてしまう。

そこで本研究では、癌の客観的指標の作成及び病理専門医の負担軽減のための癌診断システムの構築を目的とする。この支援システムを構築した文献として文献 [3] が挙げられる。文献 [3] では高次局所自己相関 (Higher-order Local AutoCorrelation

: HLAC) 特徴を用いてシステムを構築しているが、あくまでも病理画像に画像処理を加えて誤検出を減らすためのシステムであるため診断するのは病理専門医である。したがって、病理専門医の負担は依然として改善されていない。

また、病理診断においては様々な化合物で染色された細胞核が病理診断のクライテリアとして用いられることがあり、それを定量化するためには染色された細胞核を抽出する必要がある。近年では機械学習の発展により機械学習を用いた細胞核の抽出手法である文献 [4][5] が提案されている。文献 [4] では、Genetic Programming (GP) と Simulated Annealing Programming (SAP) を組み合わせた GP-SAP を提案し、病理部位抽出のための画像処理フィルタの自動構築を行った。しかし、文献 [4] における抽出対象が大きい領域である胃底腺としているため、小さな細胞核を抽出対象にできるかどうかは不明である、かつフィルタ作成に膨大な時間がかかってしまうため実用性に乏しいといった問題がある。文献 [5] では、病理画像より細胞核を一つ一つ抽出するため特徴量を学習する Support Vector Machine (SVM) を用いて、病理画像より画素ごとの SVM スコアを示したグレースケール画像を作成し細胞核抽出を行った。しかし、文献 [5] で用いている病理画像はヘマトキシリンのみの染色であるため細胞診で一般的に使われているヘマトキシリン・エオジン (Hematoxylin-Eosin : HE) 染色された病理画像を実験対象としたときに高精度な結果になるかどうか不明である。さらに、文献 [4][5] とともに抽出の手法であるため抽出した結果で悪性度診断を行う過程まで至っていない。そこで、文献 [5] の手法を踏まえて問題を解決した手法として文献 [6] がある。文献 [6] では、AlexNet [7] と SVM を組み合わせて細胞核抽出用モデルを構築し、そのモデルよりスコア画像を作成し、元の画像とスコア画像を組み合わせることで HE 染色された病理画像に対して細胞核抽出を行った。さらに、文献 [6] の抽出結果を使った細胞核分類手法として文献 [8] がある。文献 [8] では、文献 [6] と同じように AlexNet と SVM を組み合わせて細胞核分類用のモデルを構築し、文献 [6] の結果より細胞核の分類を行った。しかし、文献 [6] は細胞核を抽出するためにスコア画像を作成するが、その画像はあくまで細胞核が存在する確率が高い領域を示しているため誤検出や抽出漏れが多くなってしまふ。また、スコア画像の作成に時間がかかるためすぐに診断が行うことができない。また、文献 [8] は良性、悪性、転移の3種類を判定しており、分類精度は全体では8割を超えているが

それぞれ個々の分類率を見ると悪性に関しては約6割の分類しかできておらず信憑性に欠けてしまう。また、癌細胞の指標を「核異形度」のみで判断しているため癌細胞を正常細胞として分類してしまっている可能性もある。

以上の課題をふまえ本稿では、文献 [6][8] と同様に細胞核に着目し、Deep Learning のアルゴリズムの一種である Convolutional Neural Network(CNN) モデルを2種類用いて抽出および分類を行う。抽出にはセグメンテーション用のモデルである U-Net[9]、分類には CNN モデルである VGG16[10] に Fine-tuning を行ったモデルを使用する手法を提案し、従来手法よりも精度の向上および高速化を図る。実験では、実際の HE 染色された組織切片サンプルから細胞核を抽出および分類を行い癌診断における本手法の有効性を評価する。

以下本稿では、2章では本研究で用いられている基礎となる背景知識について述べる。3章では細胞核の抽出および分類における提案手法について述べ、4章で提案手法を用いた細胞核の抽出および分類実験の概要と考察を述べる。5章では結論と今後の課題について述べる。

第2章 背景知識

本章では，細胞，細胞核の染色，機械学習の詳細について記述する．

2.1 細胞核

細胞核とは真核生物の細胞を構成する細胞小器官の1つであり，細胞の遺伝情報の複製と保存を行い，ほとんどすべての細胞に存在している．核内では複数の長い直鎖状のDNA分子が様々な種類のタンパク質と複合体を形成することによって染色体を形成している．これらの染色体の内部の遺伝子が核ゲノムを構成しており，細胞の機能を促進するよう構造化されている．染色体DNAは主にメッセンジャーRNAに転写されるが，この遺伝子発現の調節により細胞の活動は制御されている．この調節が狂い，細胞が異常に繁殖する状態が癌化と考えられている．また，核を作り上げている主要な構造は核膜と核マトリックスである．核膜は核全体を包む2層の脂質二重膜で，その内容物を細胞質から分離している．核マトリックスは核内部のネットワーク構造で，細胞を支える細胞骨格のように核構造の機械的支持を行っている．

2.2 癌細胞の特徴

癌化した細胞核の判断指標として以下の3つがある．

核異形度

細胞の形が正常な細胞とどのくらい異なっているかを示す度合いである．正常な細胞では細胞分裂する際に2つ細胞核に均等に分配されるため核が大きくなることはない．しかし，癌細胞の核は正常な細胞よりも大きく形がゆがむ特徴がある．

核小体

核小体は真核生物の細胞核の中に存在する、分子密度の高い領域で、rRNAの転写やリボソームの構築が行われる場所のことであり、直径1~3 μm 程である。正常の細胞には1つしか存在しないが、癌細胞の核には核小体が複数個存在している。また、核小体が明瞭であるといった特徴もある。核小体の例を図2.1に示す。

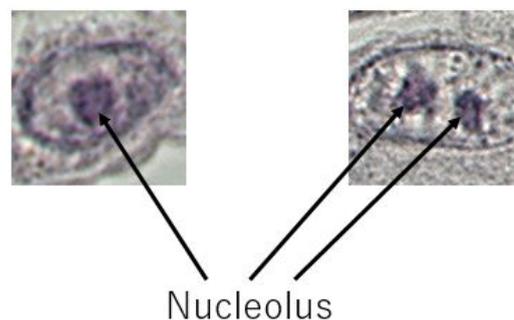


図 2.1: 核小体

N/C比

N/C比とは核と細胞質に対する面積比ないし容積比である。この比は一般的に細胞の分化度が低ければ大きく、細胞の分化度が高いものでは小さいとされている。腫瘍細胞、特に癌では核の容積が増大するのでN/C比は大きな値となる。N/C比の概要を図2.2に示す。

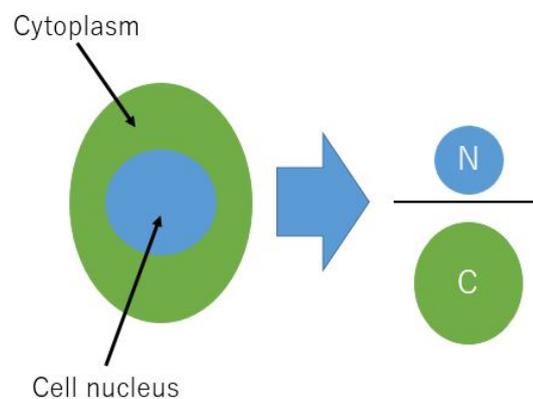


図 2.2: N/C比

2.3 メラノーマ

メラノーマ (悪性黒色腫) は皮膚がんの一種で、皮膚の色と関係するメラニンをつくるメラノサイト (色素細胞) や、ほくろの細胞 (母斑細胞) が癌化してできる癌である。発生部位としては最も多いのは足の裏だが、胴体や顔、爪など様々な部位に発生することもある。メラノーマが発生する原因については、まだよくわかっていないが、白色人種に多く発症することから、紫外線が関与していると考えられている。また、足の裏や爪などいつも刺激を受けている場所にできやすいことから、外からの物理的な刺激も関係していると考えられている。メラノーマは表在拡大型黒色腫、悪性黒子型黒色腫、末端黒子型黒色腫、結節型黒色腫の4つのグループに分けることができる。

2.4 ヘマトキシリン・エオジン染色

光学顕微鏡を用いて病理組織学的診断を行うには、元来無色の細胞あるいは組織に色彩を施し染色する必要がある。ヘマトキシリン・エオジン (HE) 染色とは、細胞および組織構造の光顕レベルでの全体像の把握を目的とする染色で、病理組織標本の最も基本的かつ重要な染色法である。ヘマトキシリンには数種類あるが基本的には細胞核を濃青紫色～藍色に、軟骨基質、粘液の一部、石灰化巣、微生物の一部、好塩基質などを青紫色～淡青色に染まる。また、エオジンで細胞質、各種線維、赤血球、好酸性物質、顆粒などを淡赤色～濃赤色に染まる。HE染色された細胞画像の一例を図2.3に示す。

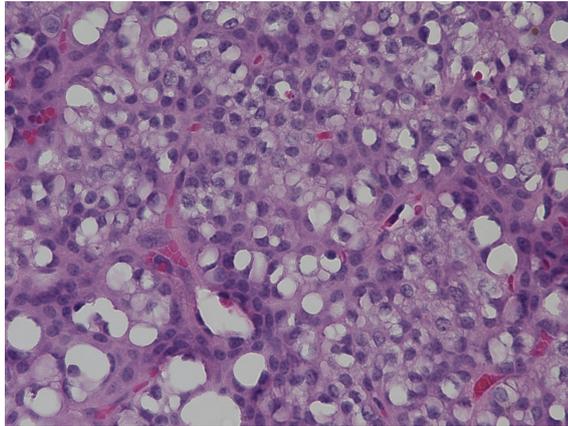


図 2.3: HE 染色された病理画像

2.5 Convolutional Neural Network

Convolutional Neural Network(CNN) は画像認識や音声認識など様々なところで用いられており, Deep Learning の登場以降, AlexNet[7], GoogLeNet[11], ResNet[12] などの高精度な手法はすべて CNN ベースのモデルとなっている. CNN の構成を図 2.4 に示す.

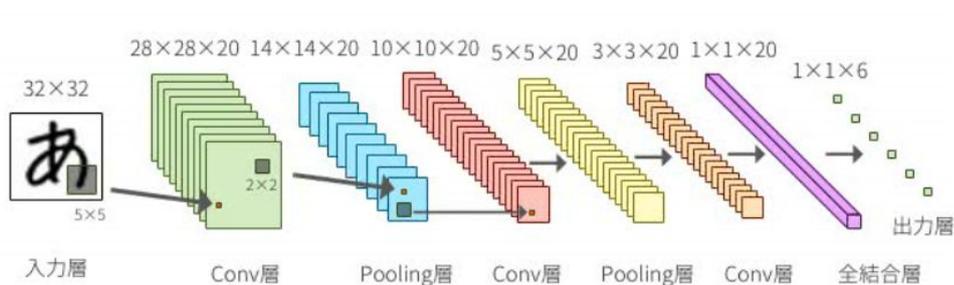


図 2.4: CNN の基本構造 (文献 [13] から引用)

CNN はニューラルネットワークのようにレイヤを組み合わせることで構成することができる. CNN の場合, 畳み込み層 (Convolution layer), プーリング層 (Pooling layer), 全結合層 (Fully connected layer : FC) で構成されている.

2.5.1 畳み込み層

畳み込みの仕組みを図 2.5 に示す.

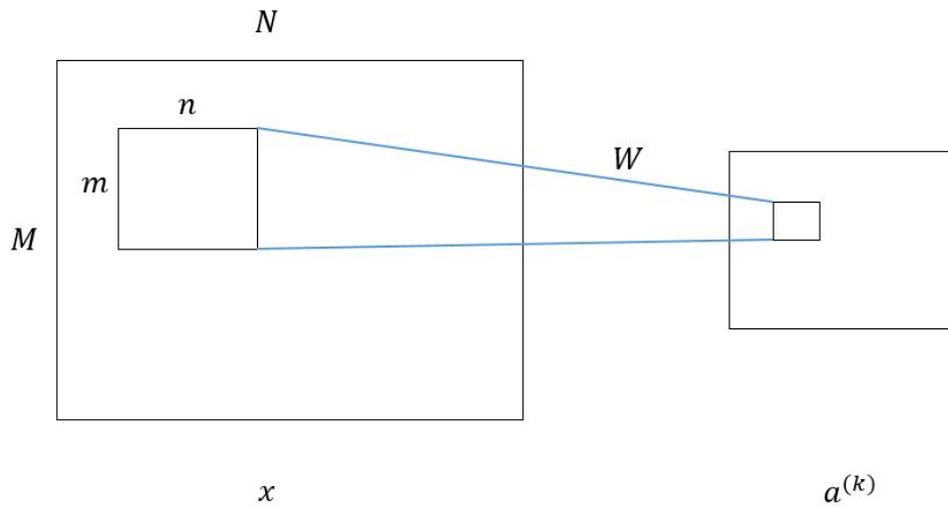


図 2.5: 畳み込みの仕組み

CNN を構成する層の一つである畳み込み層はグレースケール画像である場合、Feed-Forward は式 (2.1) で表すことができる。

$$a_{ij}^{(k)} = \sum_{s=0}^{m-1} \sum_{t=0}^{n-1} w_{st}^{(k)} x_{(i+s)(j+t)} + b^{(k)} \quad (2.1)$$

ここで、 m 、 n はカーネル (畳み込みフィルタ) のサイズ、 k はカーネルのインデックス番号、 x は入力画像データ、 $a^{(k)}$ は畳み込み層を通した後の 2 次元データ、 $w^{(k)}$ はカーネル、 $b^{(k)}$ はバイアスを表している。

次に、Backward (Backpropagation) について考える。学習すべきモデルのパラメータは $w^{(k)}$ および $b^{(k)}$ であるため、誤差関数を E で表すとすると、それぞれの勾配は式 (2.2)(2.3) のようになる。 M 、 N は入力画像のサイズである。

$$\begin{aligned} \frac{\partial E}{\partial w_{st}^{(k)}} &= \sum_{i=0}^{M-m} \sum_{j=0}^{N-n} \frac{\partial E}{\partial a_{ij}^{(k)}} \frac{\partial a_{ij}^{(k)}}{\partial w_{st}^{(k)}} \\ &= \sum_{i=0}^{M-m} \sum_{j=0}^{N-n} \frac{\partial E}{\partial a_{ij}^{(k)}} x_{(i+s)(j+t)} \end{aligned} \quad (2.2)$$

$$\begin{aligned}
\frac{\partial E}{\partial b^{(k)}} &= \sum_{i=0}^{M-m} \sum_{j=0}^{N-n} \frac{\partial E}{\partial a_{ij}^{(k)}} \frac{\partial a_{ij}^{(k)}}{\partial b^{(k)}} \\
&= \sum_{i=0}^{M-m} \sum_{j=0}^{N-n} \frac{\partial E}{\partial a_{ij}^{(k)}}
\end{aligned} \tag{2.3}$$

ここで、Backpropagation の誤差は式 (2.4) で表される。

$$\delta_{ij}^{(k)} := \frac{\partial E}{\partial a_{ij}^{(k)}} \tag{2.4}$$

式 (2.4) は前の層から逆伝播してきているのでモデルパラメータを更新することができる。畳み込み層ではこれらの処理によって更新されたパラメータによって、入力画像に対してフィルタリングが行われ目的のタスクに適した特徴が自動的に抽出される。

活性化関数

CNNでは、シグモイド関数や tanh 関数を用いるよりも ReLU 関数 (Rectified Linear Unit function)[14] を用いることが多い。Feed-Forward では ReLU 関数は式 (2.5) で表される。

$$a_{ij} = \text{ReLU}(x_{ij}) = \max(0, x_{ij}) \tag{2.5}$$

また、Backpropagation では式 (2.6) のように表される。

$$\frac{\partial E}{\partial x_{ij}} = \begin{cases} \frac{\partial E}{\partial a_{ij}} & \text{if } a_{ij} \geq 0 \\ 0 & \text{otherwise} \end{cases} \tag{2.6}$$

活性化関数として ReLU 関数がよく用いられる理由として勾配消失問題の回避が可能、学習収束が速いという点が挙げられる。畳み込みを行った後の出力に対して

活性化関数を適応したときの出力が最終的な畳み込み層の出力となり、次の層の入力となる。ReLU関数のグラフを図2.6に示す。

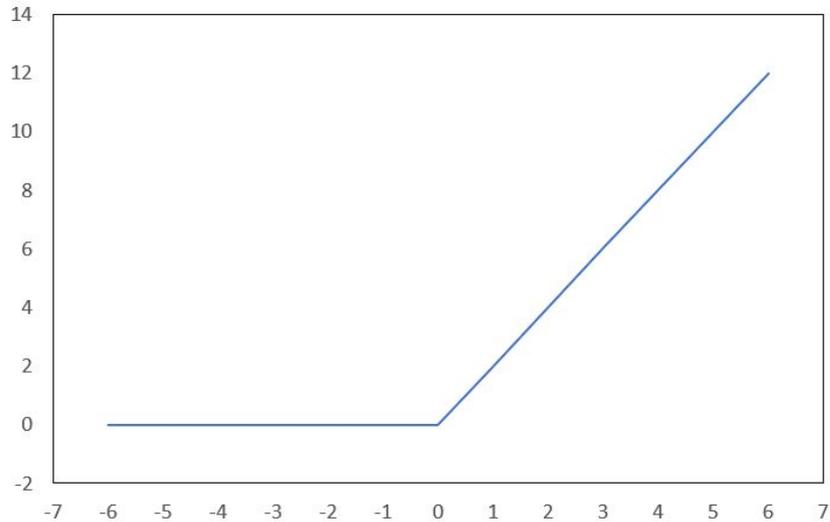


図 2.6: ReLU 関数のグラフ

2.5.2 プーリング層

プーリングには対象領域内から最大値を取る Max pooling と対象領域内から平均値を取る Average pooling がある。画像認識の分野においては主に Max pooling が使われているため本稿ではプーリング層は Max pooling とする。プーリング層では学習するパラメータを持っていないため、Feed-Forward, Backpropagation とともに各処理を行い、前後の層と繋げるのみである。Feed-Forward についての式を式 (2.7), Backpropagation についての式を式 (2.8) に示す。 l はフィルタサイズである。

$$a_{ij} = \max(x_{(li+s)(lj+t)}) \quad \text{where } s \in [0, l], t \in [0, l] \quad (2.7)$$

$$\frac{\partial E}{\partial x_{(li+s)(lj+t)}} = \begin{cases} \frac{\partial E}{\partial a_{ij}} & \text{if } a_{ij} = x_{(li+s)(lj+t)} \\ 0 & \text{otherwise} \end{cases} \quad (2.8)$$

プーリング層では入力データをより扱いやすい形に変形させるため、情報を圧縮し、Down sampling している。この処理によって抽出された特徴の微小な位置変化に対

して頑健となり，過学習を抑制し，計算コストを抑えることができる．Max pooling の例を図 2.7 に示す．

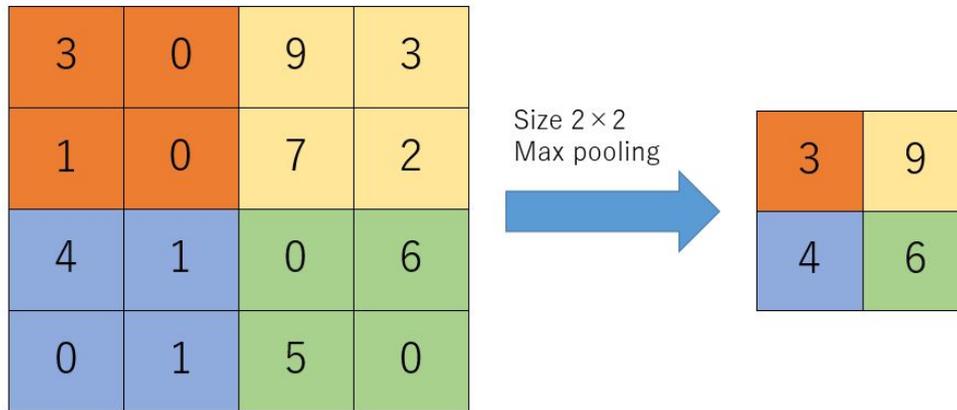


図 2.7: Max pooling の例

図 2.6 の例では 2×2 のサイズで画素値を見て各領域から最大値である 3, 9, 4, 6 の値を取り出して Down sampling を行っている．

2.5.3 全結合層

全結合層は，畳み込み層とプーリング層から取り出された特徴部分の画像データを一つのノードに結合し，活性化関数によって変換された特徴ベクトルを出力する．また、ノードの数を増やすと特徴空間の分割数が比例して増加し，各領域を特徴づける特徴変数が増加する．全結合層の構成を式 (2.9)，図 2.8 に示す．ここで， X は入力， W は各ノードとの結合重み， b はバイアス， i は画素数， j はノード数を表す．

$$h_j = f(W_x^T + b_j) \quad (2.9)$$

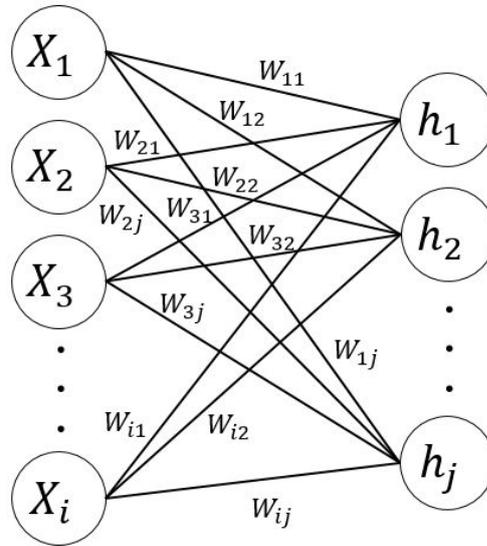


図 2.8: 全結合層の構成

2.6 U-Net

U-Net[9] は画像のセグメンテーションに特化したモデルである。特に医療画像に強みを持っており、ISBI2015における「Grand Challenge for Computer-Automated Detection of Caries in Bitewing Radiography」と「Cell Tracking Challenge」で優勝している。U-Net の構造を図 2.9 に示す。

図 2.9 の左側の下向きパスは、畳み込み層とプーリング層により、深い層ほど特徴が局所的で位置情報が曖昧になり、浅い層ほど特徴は全体的で位置情報は正確になる。一方、図 2.9 の右側の上向きパスは畳み込み層とアップサンプリングにより、特徴を保持したまま画像を大きく復元することが可能である。また、左側で使用した画像情報を右側にコピーし結合させることで、より精度の高いピクセル単位でのセグメンテーションが可能になる。

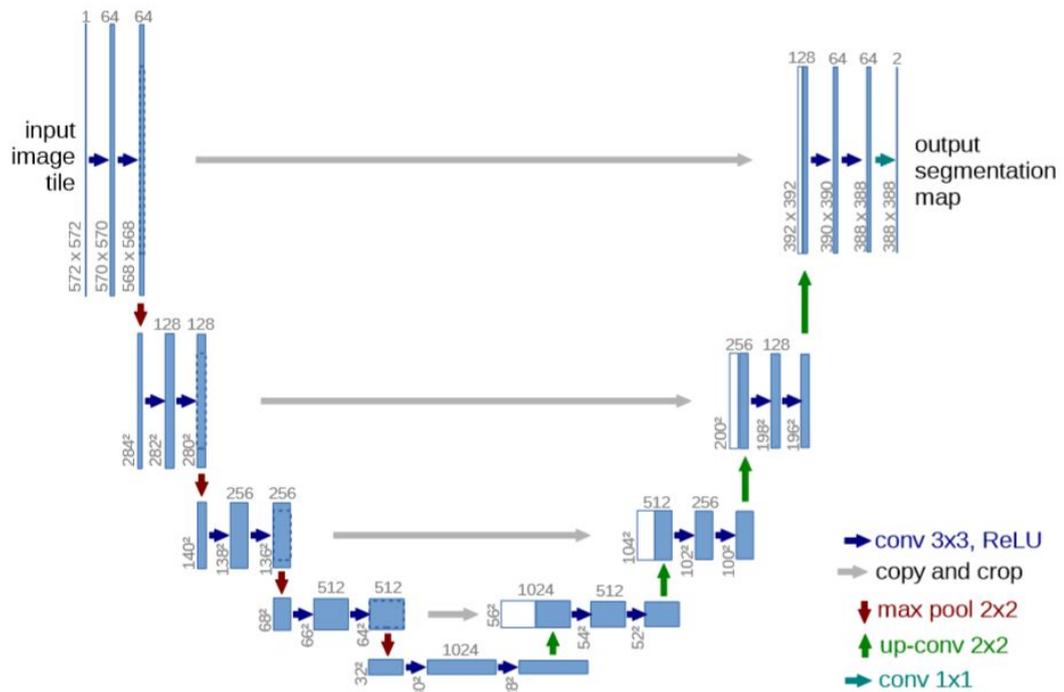


図 2.9: U-Net の構造 (文献 [9] から引用)

2.7 VGG16

VGG16[10] は、ILSVRC2014 で準優勝したモデルである。畳み込み層とプーリング層から構成されているシンプルな CNN モデルであるが畳み込み層と全結合層の重みのある層を 16 重ねている。また、畳み込み層のフィルタサイズが全て 3×3 になっており、それをプーリング層でサイズを半分にするというのを繰り返し行う構造が特徴である。この特徴により、大きいフィルタで画像を一気に畳み込むよりも小さいフィルタを何個も畳み込むことで特徴量をより良く抽出できる。VGG16 の構造を図 2.10 に示す。

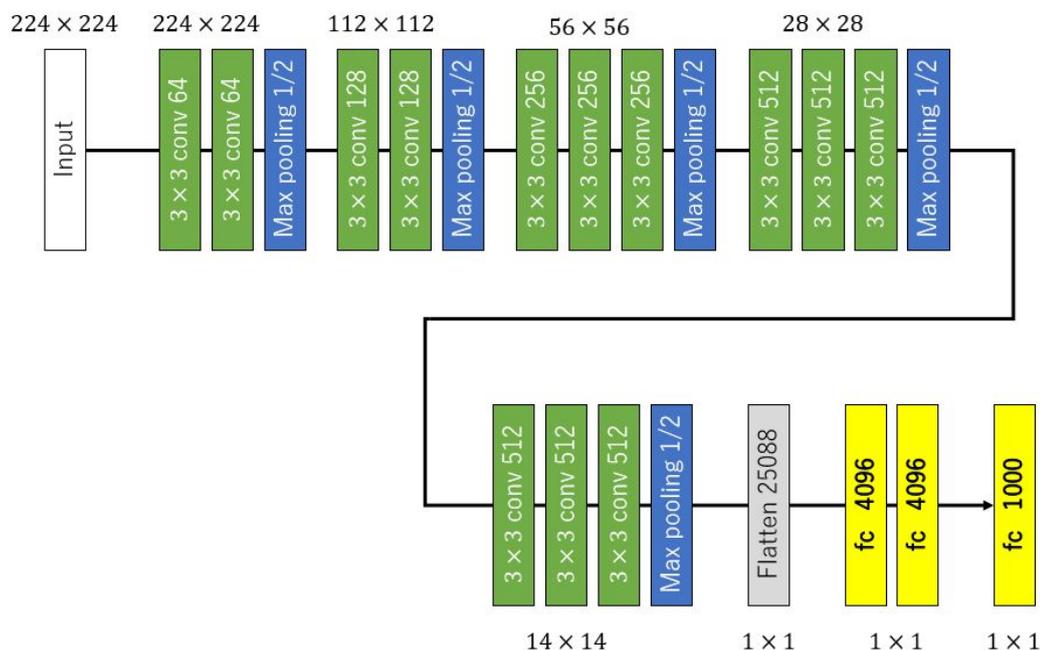


図 2.10: VGG16 の構造

2.8 Fine-tuning

規模の小さいデータセットを用いたディープラーニングに関して効果的な学習方法としてよく使用される方法として、学習済みネットワークを使用する Fine-tuning がある。学習済みネットワークとは、1000 種類のクラスで 120 万個のラベル付き画像がある ImageNet のような大規模なデータセットで学習した重みを保存したネットワークである。Fine-tuning はベースとなるネットワークの特徴を得る層を別の問題に適用できるように新しく作り直して学習させる方法である。一般的には、全結合層を新たに学習する場合と特徴を得る層まで学習する場合がある。本研究では、後者の特徴を得る層まで新しく学習する方法を用いている。

2.9 Data Augmentation

Deep Learning は、学習時に最適化するパラメータ数が多いため、数千枚、数万枚といった学習データが必要と言われている。しかし、データの取得が限られる医療画像など、十分な量の学習データを用意できないことが多々ある。Data Augmentation

とは、学習データの画像に対して平行移動、拡大縮小、回転、ノイズの付与などの処理を加えることで、データ数を人為的に増強する手法である。

2.10 交差検証

学習モデルの汎用性を示すために、学習データとテストデータを用意する必要がある。しかし、データ数が少ないと、テストデータの選び方によって、精度に大きな誤差が発生してしまう可能性がある。そのような時に交差検証が使われる。交差検証の手順を以下に示す。また、交差検証の概略図を図 2.11 に示す。

- i データセットを K 分割し K 個のブロックを作成する
- ii K 個のブロックのうち 1 個のブロックをテスト用のデータセットとし、残りのブロックを学習用のデータセットとする
- iii 学習用のデータセットで学習したモデルをテスト用のデータセットで評価する
- iv ii. から iii. を K 回繰り返し評価値の平均を出す

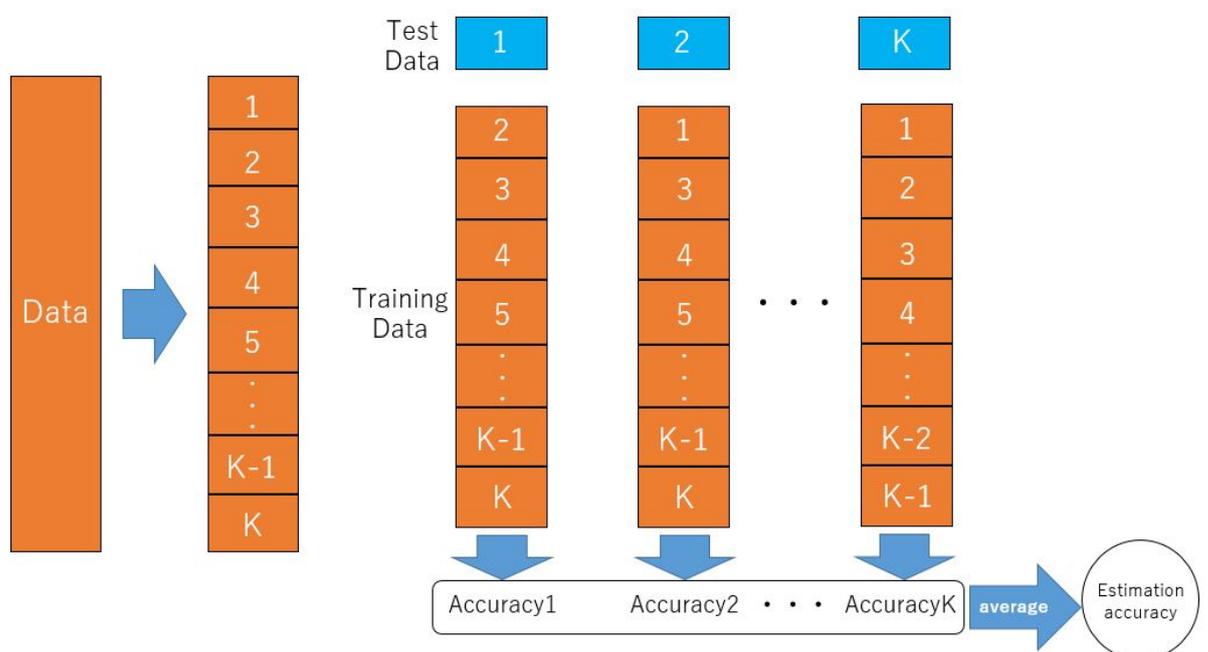


図 2.11: 交差検証の概略図

第3章 細胞核の抽出および分類における提案手法

本章では細胞核の抽出および分類の提案手法について記述する。提案手法では、医療画像のセグメンテーション用のモデルである U-Net[9] を用いて細胞核の抽出を行い細胞核抽出画像の作成を行う。さらに、その画像から CNN モデルである VGG16[10] に Fine-tuning を行い細胞核の分類を行う。以下に提案手法の流れを示す。また、概要図を図 3.1 に示す。

Step1 細胞画像およびマスク画像を用いて細胞核抽出用のモデルを学習する

Step2 細胞核抽出用のモデルを用いて細胞核抽出画像を作成する

Step3 正常細胞および癌細胞を用いて細胞核分類用のモデルを学習する

Step4 Step2 で作成した細胞核抽出画像より細胞核分類用のモデルを用いて細胞核の分類を行う

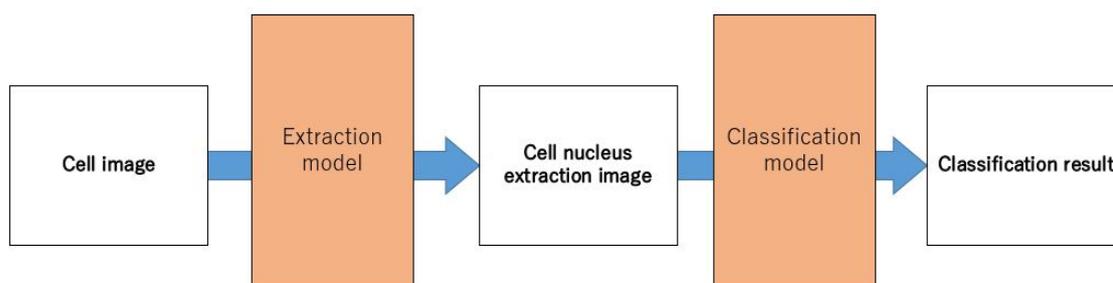


図 3.1: 提案手法の概要図

3.1 細胞核の抽出における提案手法

本節では、細胞核の抽出における提案手法について記述する。

3.1.1 モデルの構造

本手法の細胞核抽出には、医療画像のセグメンテーション用のモデルである U-Net を用いた。本手法で用いる U-Net の構造を図 3.2 に示す。

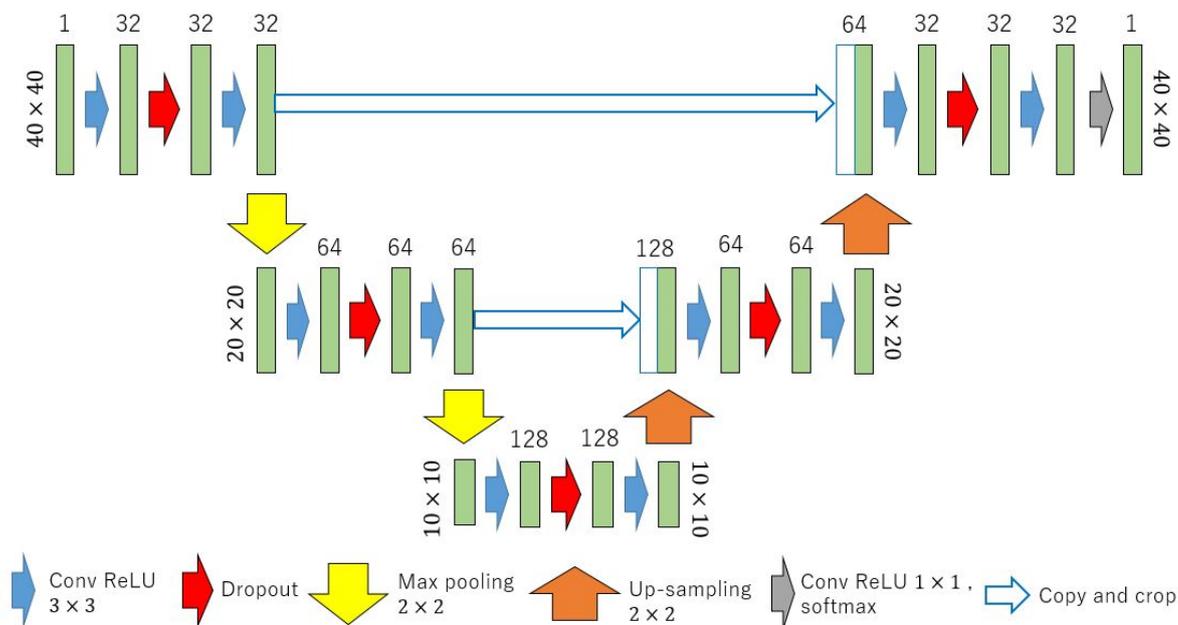


図 3.2: U-Net の構造

文献 [9] の U-Net の構造は入力サイズが 572×572 で出力サイズが 388×388 となっており、入力サイズよりも出力サイズが小さくなっている。しかし、本手法では 40×40 のパッチという小さい画像を作成し、学習用データセットとしている。そのため、U-Net における入力サイズも入力画像サイズに合わせ 40×40 としている。さらに、細胞核の分類において出力された画像を使用する関係上出力サイズも入力サイズと同様に 40×40 としている。このように入力画像と出力画像が等しいモデルのことを「Encoder-Decoder モデル」という。U-Net では左半分の畳み込みの部分が Encoder にあたり、右半分の Up-sampling の部分が Decoder にあたる。Up-sampling は、畳み込みこまれた画像を畳み込む前の状態に戻す処理のことである。

出力層直前以外の畳み込み層において出力のサイズを落とさないようにパディングを行う。また、畳み込み層では過学習を防ぎ汎化性能を高めるためデータの大きさに応じて 0 に近づけて滑らかなモデルにする L2 正則化 [15] を行う。さらに、本手

法の U-Net 全体で 50% の確率で Dropout[16] を行う層を追加している。また、本手法で使用する U-Net は学習用データセットとして扱うパッチ画像のサイズが 40×40 と小さいため文献 [9] の構造よりも層が少なくなっている。

3.1.2 データセットの作成

本手法では細胞核抽出における学習用データセットとしてパッチ画像を作成する。HE 染色されたヒトのメラノーマ病理組織画像 60 枚、遺伝子改変マウスに発症したメラノーマ組織画像 20 枚をグレースケール変換，ヒストグラム平坦化，ガンマ補正を行ったそれぞれの画像よりパッチを作成し，学習用データセットとする。また，同じ細胞画像 80 枚のマスク画像からもパッチを作成する。パッチの枚数は 190,000 枚，サイズは 40×40 である。図 3.3 に細胞画像から作成したパッチの例，図 3.4 にマスク画像から作成したパッチの例を示す。

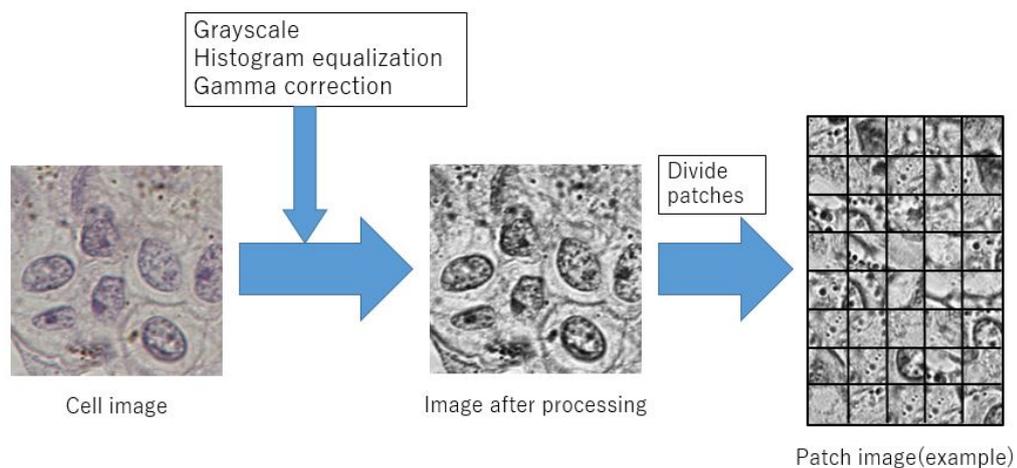


図 3.3: パッチ画像の作成 (実画像)

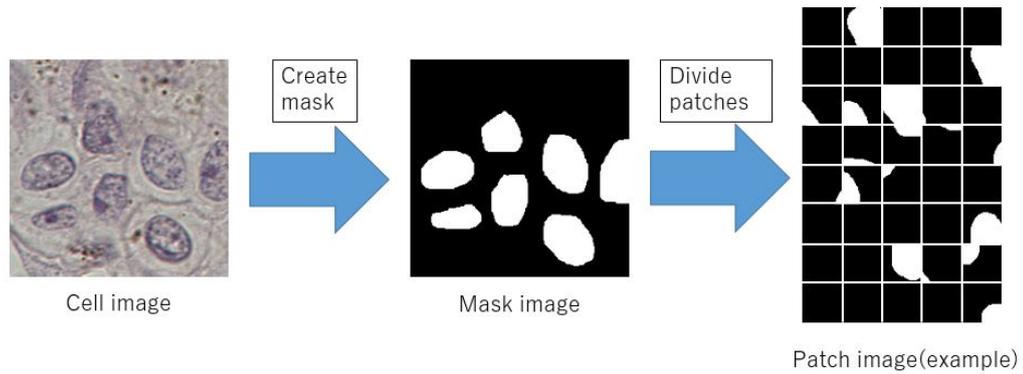


図 3.4: パッチ画像の作成 (マスク)

3.1.3 抽出方法

本手法では画素ごとではなく検出ウィンドウを設けてそれを細胞画像に対してラスタスキャンを行い，検出ウィンドウから取得した画像を入力画像とする．検出ウィンドウの大きさは学習用データセットであるパッチ画像に合わせて $40 \times 40\text{px}$ ，抽出間隔を 5px とした．モデルに通した入力画像は細胞核が抽出された状態で出力される．それをラスタスキャンが全ての画素に対して終わるまで繰り返し，全ての結果を統合し細胞画像の抽出結果画像を作成する．本手法の抽出方法を図 3.5 に示す．

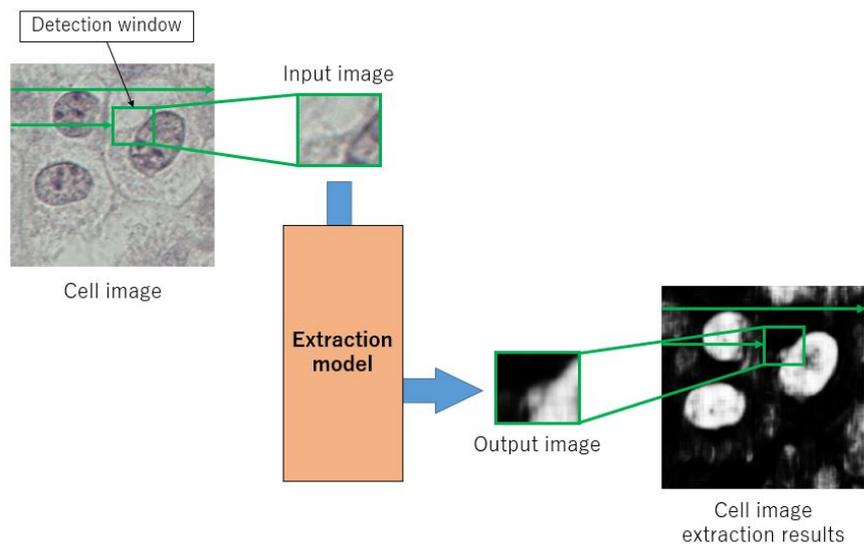


図 3.5: 抽出方法

3.1.4 後処理

作成した細胞核抽出画像にはノイズが含まれている場合がある。そのため、細胞核抽出画像には膨張処理や収縮処理を同じ回数行うモルフォロジー処理でノイズの低減を行う。また、ラベリング処理を行い抽出細胞のなかで細胞核面積が500未満のものは細胞核分類を用いるには情報量が少ないと判断しノイズとして削除した。

3.2 細胞核の分類における提案手法

本節では、抽出された細胞核の分類における提案手法について記述する。

3.2.1 モデルの構造

本手法の細胞核分類には、CNNモデルであるVGG16を用いる。しかし、VGG16は大規模データセットであるImageNetの画像分類用に作成されているため、そのままでは使用できない。そこで、学習済みネットワークを利用するFine-tuningという手法を用いてVGG16を細胞核分類用に拡張する。構築したモデルを図3.6に示す。

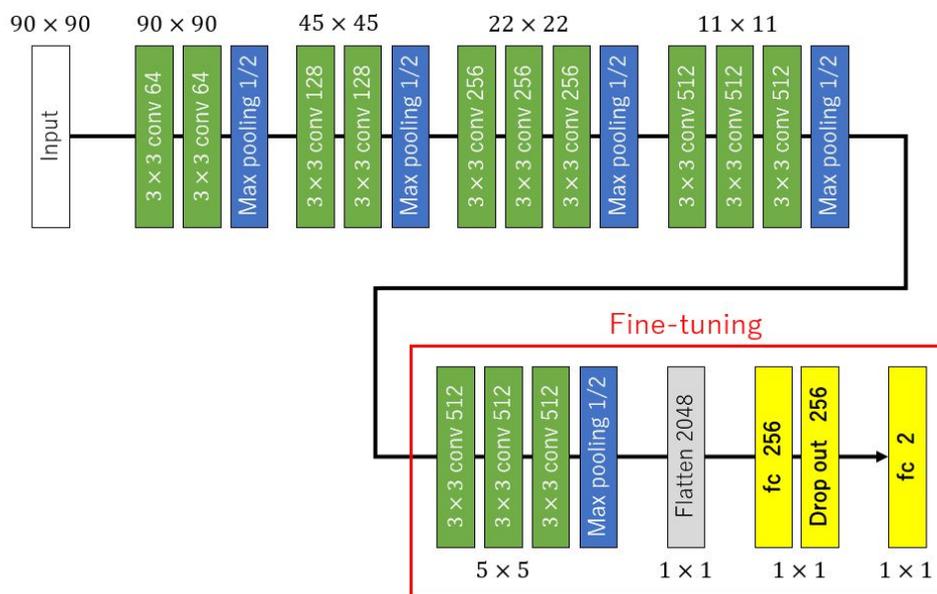


図 3.6: 細胞核分類用に構築したモデル

CNN では浅い層でエッジやブロブなど汎用的な特徴が抽出されているのに対し、深い層ほど学習データに特化した特徴が抽出される傾向がある。そこで、構築したモデルでは1~4番目のブロックをそのまま固定し ImageNet で学習された重みを使用する。また、5番目のブロックと全結合層を再学習し細胞核分類用に再調整する。

本手法で構築したモデルの全ての conv 層には、過学習を防ぎ汎化性能を高めるため特定のデータの重みを0にする事で不要なデータを削除する L1 正則化 [17]、データの大きさに応じて0に近づけて滑らかなモデルにする L2 正則化を行う。

3.2.2 データセットの作成

癌化した細胞核の判断指標として「核異形度」「核小体」「N/C比」がある。本研究で扱う細胞画像の中には重複している細胞や細胞同士の距離が近いものあることから「N/C比」を癌細胞の判断指標として扱うのは困難であると言える。そこで、目視で確認できる「核異形度」「核小体」を癌細胞の判断指標として分類用のデータセットを構築する。

肺がん、大腸がん、乳がんなど癌には様々な種類があり、特徴や発生の原因はそれぞれ異なる。そこで、本研究では分類対象を皮膚がんの一種であるメラノーマとした。メラノーマの病理画像から核異形度が比較的大きい細胞核と核小体が明瞭または複数個存在する細胞核を癌細胞、黒子などの良性腫瘍の画像から核異形度が比較的小さい細胞核と核小体が不明瞭または一個存在する細胞核を正常細胞とラベル付けを行い、分類用データセットを構築した。図 3.7 に作成したデータセットの一部を示す。なお、本研究では正常細胞とラベル付けされた細胞核画像 77 枚、癌細胞とラベル付けされた細胞核画像 98 枚をデータセットとして扱った。

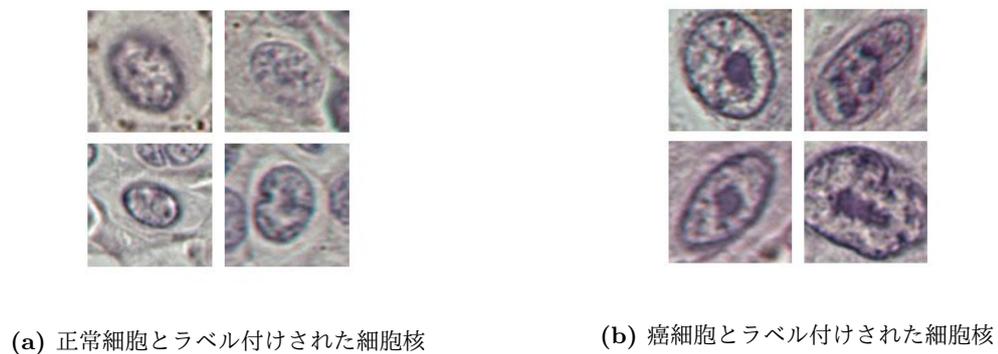


図 3.7: 分類用データセットの一例

3.2.3 細胞核分類の処理手順

細胞核の分類における処理手順を以下に示す。また、分類手法の概要を図 3.8 に示す。

Step1 抽出した細胞核にラベリング処理を行う

Step2 細胞核領域を矩形化する

Step3 矩形化された細胞核画像を入力画像として学習させる

Step4 出力された結果より正常細胞か癌細胞に分類する

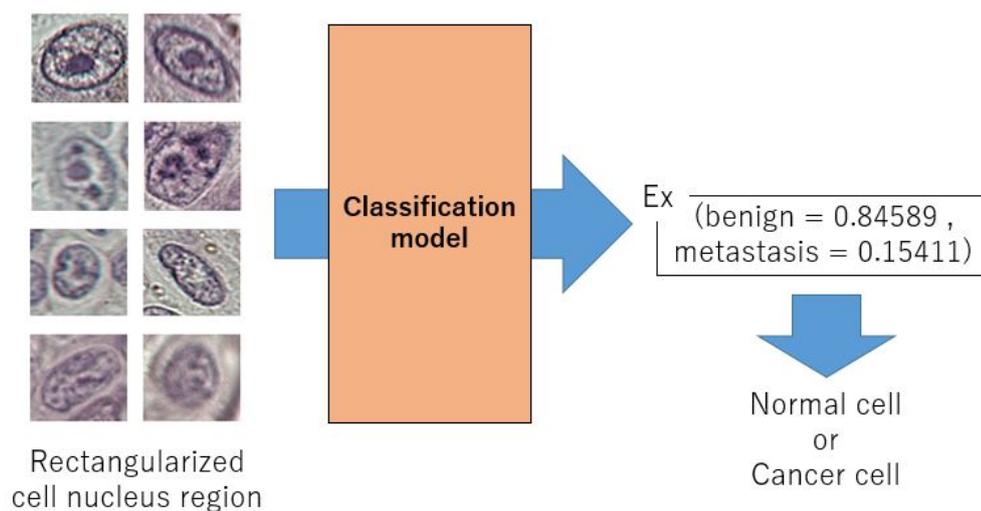


図 3.8: 分類手法の概要

第4章 細胞核の抽出および分類実験

本章では提案手法の有効性を検証するために細胞核の抽出および分類実験を行う。

4.1 細胞核の抽出実験

本節では提案手法を用いた細胞核の抽出実験を行い、提案手法の有効性を示す。

4.1.1 モデルの評価実験

抽出用モデルの汎用性を確認するため評価実験を行った。評価は10分割交差検証で抽出用モデル評価した。データセットはヒトのメラノーマ病理組織画像60枚、遺伝子改変マウスに発症したメラノーマ組織画像20枚の計80枚使用する。80枚のうち9割を学習、1割をテストとし、学習用データはパッチ画像を190,000枚作成する。比較対象として文献[6]を用いた。評価方法として、細胞核と抽出されたもののうち実際に細胞核である割合を示す Precision, 抽出された細胞核のうち正しく細胞核として抽出された割合を示す Recall, Precision と Recall の調和平均を示す F-measure をそれぞれ式(4.1)(4.2)(4.3)より求めた。

$$Precision = \frac{TP}{TP + FP} \quad (4.1)$$

$$Recall = \frac{TP}{TP + FN} \quad (4.2)$$

$$F - measure = \frac{2 \cdot Precision \cdot Recall}{Precision + Recall} \quad (4.3)$$

ここで、ここで、TP は細胞核を細胞核として抽出した場合に加算され、FP は非細胞核を細胞核として抽出し場合に加算される。FN は細胞核を非細胞核として抽

出した場合に加算される。提案手法の10分割交差検証の評価結果を表4.1、文献[6]と表4.1の精度評価をまとめたものを表4.2に示す。

表 4.1: 抽出結果 (提案手法)[%]

K	Precision	Recall	F-measure
1	76.7	80.2	78.3
2	76.2	81.5	78.7
3	50.5	67.5	57.0
4	35.3	93.8	51.1
5	57.8	77.7	65.4
6	55.0	80.9	65.2
7	85.2	72.0	77.8
8	86.5	77.9	81.8
9	76.2	80.4	78.1
10	89.1	68.6	77.2
Average	69.9	78.1	71.1

表 4.2: 精度評価のまとめ [%]

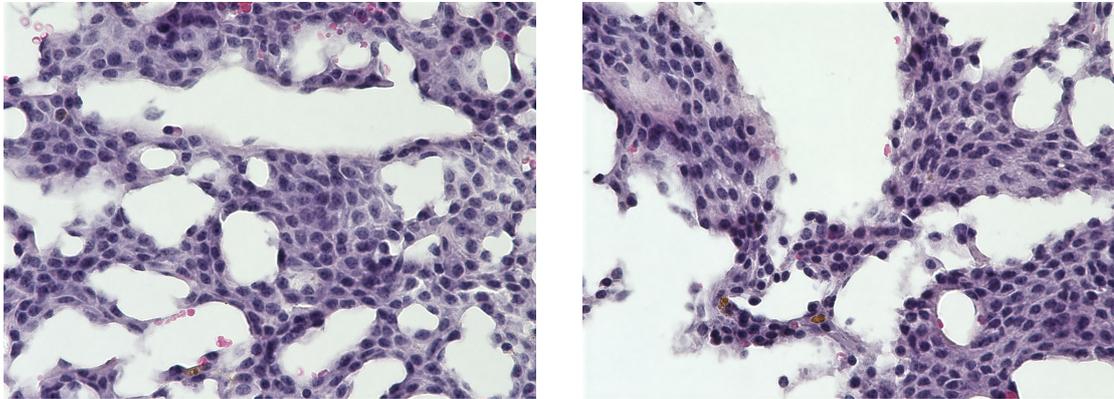
	Precision	Recall	F-measure
文献 [6]	75.7	40.1	50.1
提案手法	69.9	78.1	71.1

表 4.2 より、Recall と F-measure の評価値が提案手法の方が高いことがわかる。Precision は、文献 [6] の方が高い結果となったが細胞核抽出において抽出漏れが少ないことが重要視されることが考えられるため Recall の評価値が高い提案手法のモデルの汎用性が確認できる。

4.1.2 病理画像を用いた抽出実験

細胞核抽出における提案手法の有効性を検証するため、実際の HE 染色された細胞画像に対して、細胞核が抽出できるかどうか実験を行った。実験に用いる画像を

図4.1に示す。なお、この実験で使用する画像は、中部大学実験動物教育センターが採取したマウスの腎臓の細胞を使用し、画像のサイズは1920×1440である。



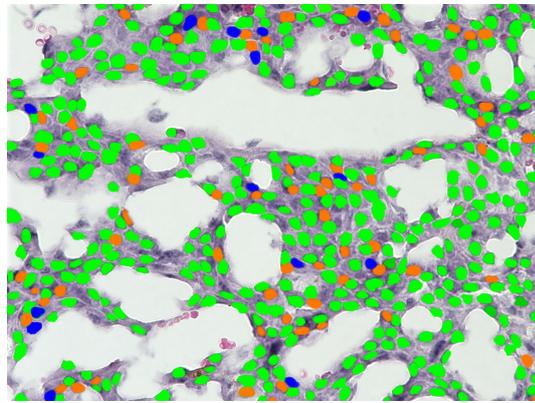
(a) 細胞画像1

(b) 細胞画像2

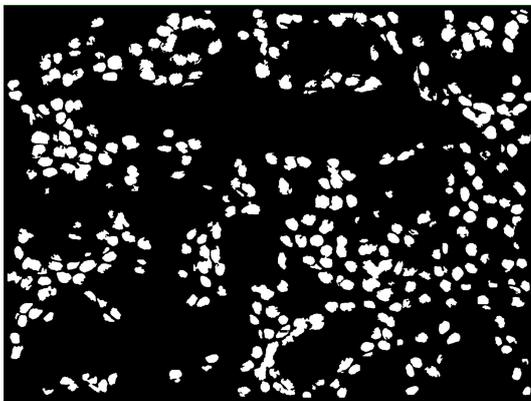
図4.1: 実験で使用する細胞画像

実験結果

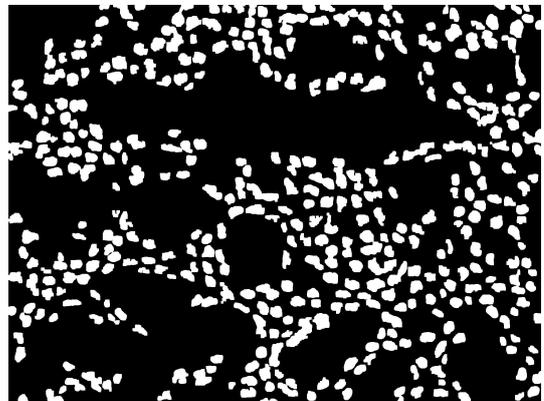
正解画像、文献[6]、提案手法によって作成された細胞核抽出画像を図4.2, 4.4に示す。また、抽出結果を切り取った一部を図4.3, 4.5に示す。なお、正解画像は目視で抽出を行ったものであり、図4.2a, 4.4aの緑色の画素は細胞核を示し、他色の画素は重複した細胞を示している。



(a) 正解画像

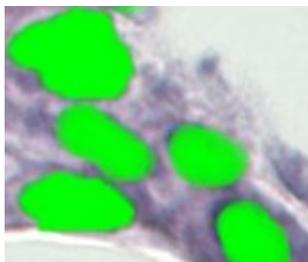


(b) 文献 [6] の抽出結果



(c) 提案手法の抽出結果

図 4.2: 細胞画像 1 の結果



(a) 正解画像 (一部)



(b) 文献 [6] の抽出結果 (一部)



(c) 提案手法の抽出結果 (一部)

図 4.3: 細胞画像 1 の結果 (一部)

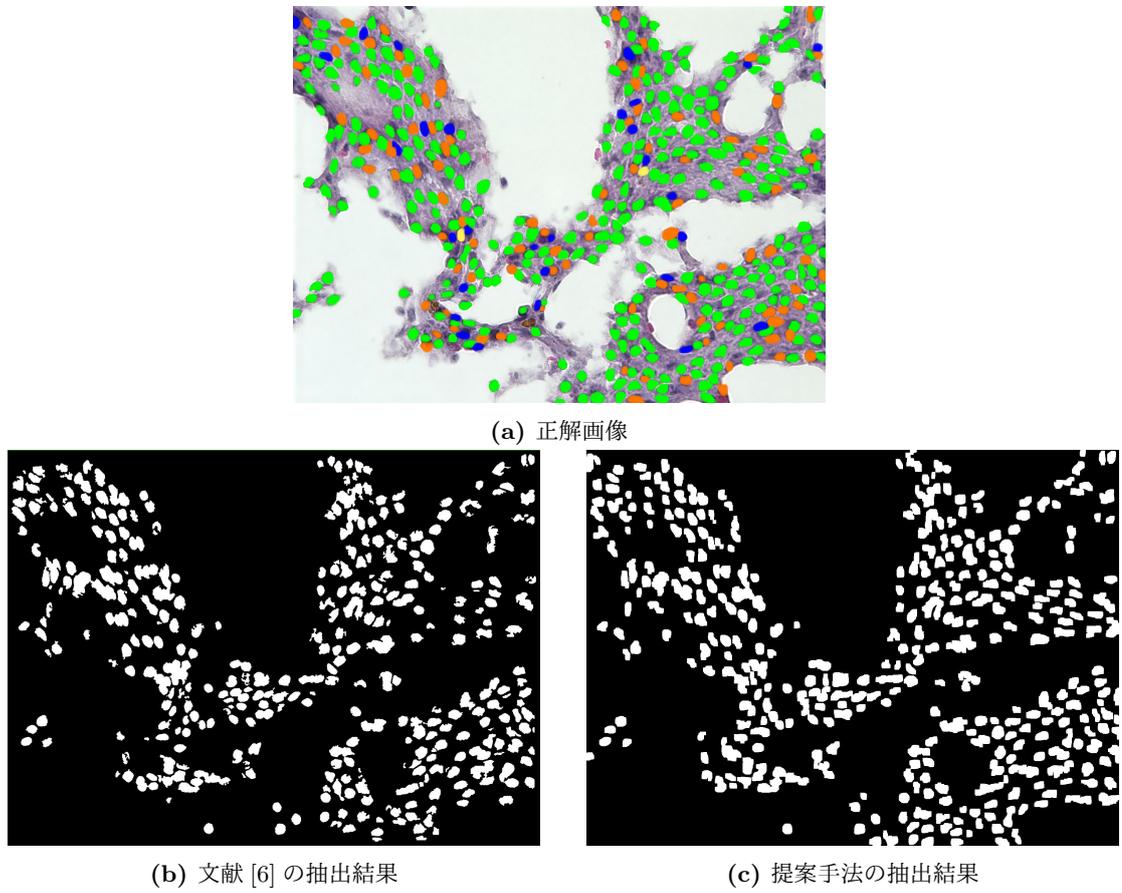


図 4.4: 細胞画像 2 の結果

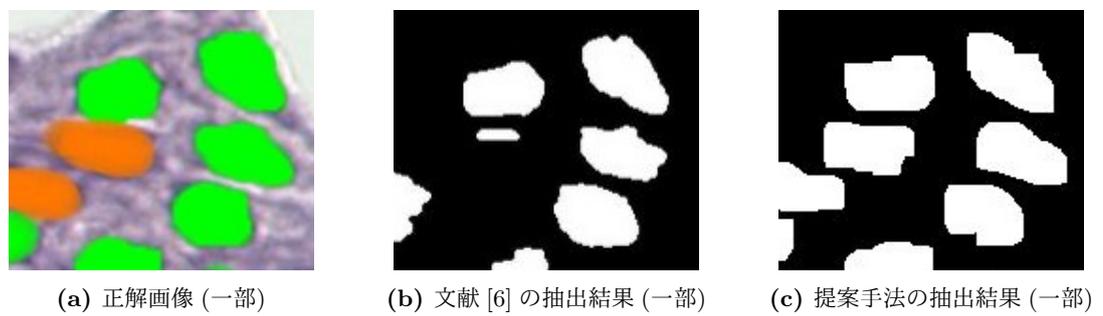


図 4.5: 細胞画像 2 の結果 (一部)

精度評価

提案手法の有効性を確認するため精度評価を行った。抽出結果の評価方法は4.1.1節と同じ方法で精度評価を行った。細胞画像それぞれに対する提案手法と文献[6]における精度評価を表4.3, 4.4に示す。

表 4.3: 精度評価 1[%]

細胞画像 1	Precision	Recall	F-measure
文献 [6]	83.5	63.0	71.8
提案手法	75.2	84.5	79.6

表 4.4: 精度評価 2

細胞画像 2	Precision	Recall	F-measure
文献 [6]	77.8	75.4	76.6
提案手法	71.0	88.7	78.9

表 4.3, 4.4 より Precision の精度が落ちてしまったが, Recall, F-measure は提案手法の方が精度の良い結果になった。特に提案手法の Recall の値に関しては80%以上と高い値であることが確認でき, 抽出漏れが少ないことを表している。細胞核の抽出において, 過検出よりも抽出漏れを少なくすることが重要であると考えられる。また, 細胞核の分類をすることを考えると, 細胞の見逃しが原因で患者に不利益な状況につながってしまうことも考えられるため, 抽出漏れが少ない本手法の有効性が確認できる。

4.2 細胞核の分類実験

本節では前節の結果より提案手法を用いた細胞核の分類実験を行い, 提案手法の有効性を示す。

4.2.1 モデルの評価実験

分類用モデルの汎用性を確認するため評価実験を行った。評価は10分割交差検証で分類用モデルの評価を行う。データセットは正常細胞とラベル付けされた細胞核画像77枚，癌細胞とラベル付けされた細胞核画像98枚の計175枚使用する。また，学習する前にトレーニングデータは回転，左右上下反転，鮮鋭化を加える Data Augmentation を行いトレーニングデータを8倍に増強する。比較対象として文献 [8] を用いた。10分割交差検証のK回目のトレーニング用画像枚数，テスト用画像枚数をそれぞれ表4.5，4.6に示す。

表 4.5: 10 分割交差検証のトレーニング用画像の枚数

K	1	2	3	4	5	6	7	8	9	10
正常細胞	69	69	69	69	69	69	70	70	70	70
癌細胞	88	88	88	88	88	88	88	88	89	89
合計	157	157	157	157	157	157	158	158	159	159

表 4.6: 10 分割交差検証のテスト用画像の枚数

K	1	2	3	4	5	6	7	8	9	10
正常細胞	8	8	8	8	8	8	7	7	7	7
癌細胞	10	10	10	10	10	10	10	10	9	9
合計	18	18	18	18	18	18	17	17	16	16

精度評価としてデータを正しく分類できた割合を示す Accuracy を求めた。ここで，正常細胞を正常細胞と分類したものを TP，正常細胞を癌細胞と分類したものを FN，癌細胞を正常細胞と分類したものを FP，癌細胞を癌細胞と分類したものを TN とする。Accuracy の式を式 (4.4) に示す。

$$Accuracy = \frac{TP + TN}{TP + FN + FP + TN} \quad (4.4)$$

提案手法における10分割交差検証の結果を表4.7に示す。また，文献 [8] と表4.7の Accuracy をまとめたものを表4.8に示す。

表 4.7: 分類結果 (提案手法)

K	TP	FP	FN	TN	Accuracy[%]
1	5	2	3	8	72.2
2	8	2	0	8	88.9
3	7	2	1	8	83.3
4	8	0	0	10	100
5	8	1	0	9	94.4
6	7	1	1	9	88.9
7	8	1	0	9	94.4
8	7	2	0	8	88.2
9	6	1	1	8	87.5
10	7	3	0	6	81.3
	Average				87.9

表 4.8: Accuracy のまとめ

	Accuracy[%]
文献 [8]	83.6
提案手法	87.9

表4.8より提案手法の方が Accuracy が高いことがわかる。したがって、提案手法におけるモデルの汎用性が確認できた。

4.2.2 病理画像を用いた分類実験

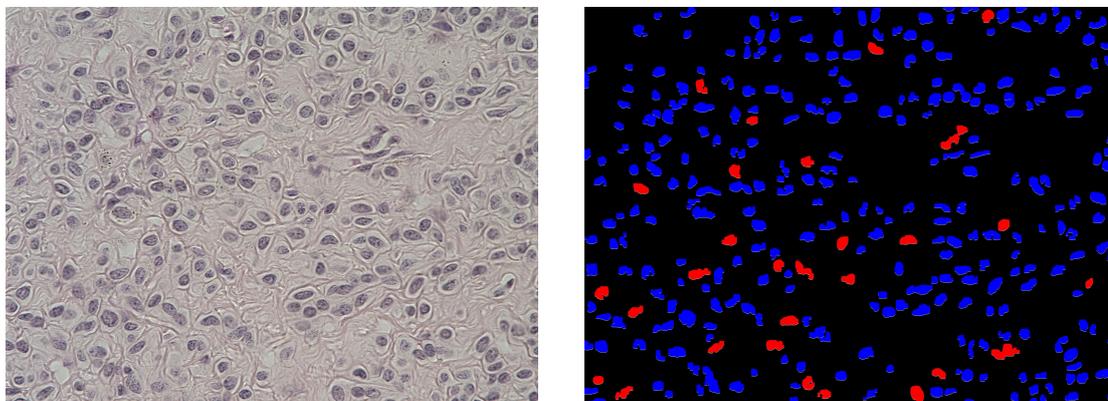
本実験には、患者から採取した皮膚組織を HE 染色した病理画像に対して、病理専門医がメラノーマかどうかを診断したものを使用する。メラノーマの判定には「良性」「悪性」「転移」の3種類ある。「良性」は良性腫瘍と診断されたもの、「悪性」は悪性腫瘍と診断されたもの、「転移」はメラノーマが他の組織に転移したもので、悪性と診断されている。また、使用する HE 染色された細胞画像は光学顕微鏡によって高倍率で撮影されたものを使用し、サイズは 1920×1440 である。

実際の細胞診において、細胞核単位で診断したときに悪性の細胞核が一定上存在したら悪性であるというような明確な判断基準は確立されていない。しかし、本実験で分類対象となってるメラノーマは特性上細胞画像内の数で判断してもよいとされている。そこで、本手法の分類基準として抽出された細胞核全体をみて良性と判断された細胞核が多ければ正常、そうでなければ癌とする。実験では、「良性」とラベル付けされている画像に対して正常と分類、「悪性」「転移」とラベル付けされている画像に対して癌と分類された場合に正しく分類されているとする。

本実験で使用する病理画像には細胞核が数千個ほど点在しているので細胞核を一つ一つ分類していくと時間がかかってしまう。そこで、本実験では病理画像と病理画像を16等分した切片画像の2種類に対して分類を行うことで、病理画像の一部を切り取った場合でも正しく分類できるかの検証を行った。

切片画像に対する分類実験

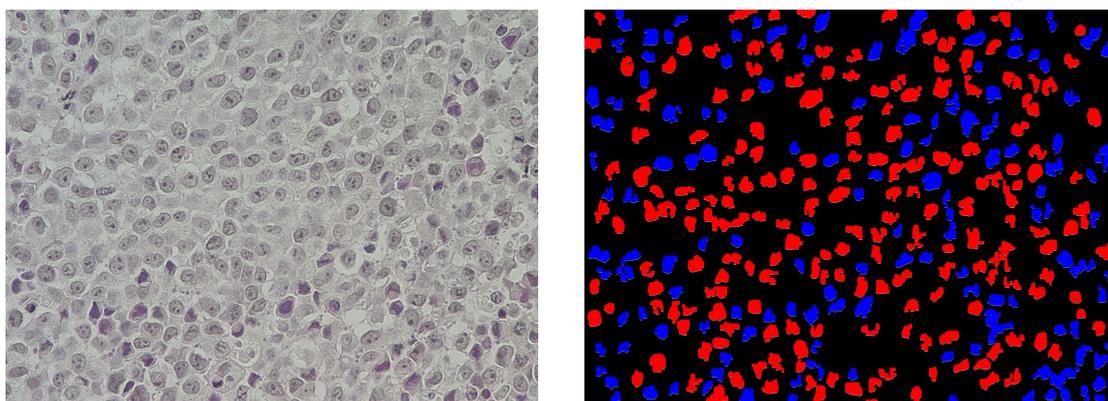
良性とラベル付けされた画像10枚、悪性とラベル付けされた画像20枚、転移とラベル付けされた画像10枚の計40枚の切片画像に対して分類実験を行った。実験に用いたそれぞれの細胞画像と分類結果の一部を図4.6, 4.7, 4.8に示す。また、分類結果の画像内における青い領域は正常と分類された細胞核、赤い領域は癌と分類された細胞核を表している。



(a) 細胞画像

(b) 分類結果

図 4.6: 良性とラベル付けされた切片画像例



(a) 細胞画像

(b) 分類結果

図 4.7: 悪性とラベル付けされた切片画像例

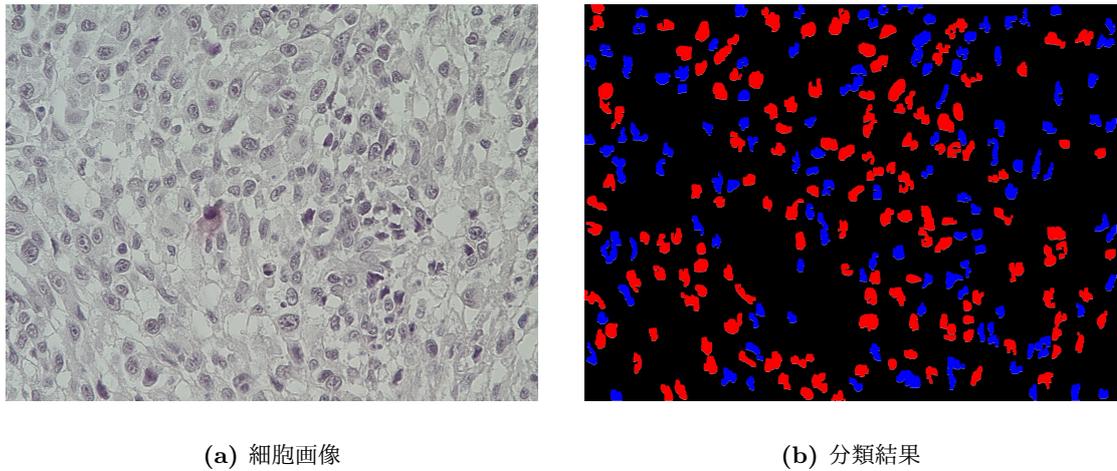


図 4.8: 転移とラベル付けされた切片画像例

各ラベルの画像に対して，正常と分類した細胞核数，悪性と分類した細胞核数，ラベル毎の正答率，全体の正答率を表 4.9 に示す．また，比較対象として文献 [8] の結果を表 4.10 に示す．

表 4.9: 切片画像に対する分類結果 (提案手法)

ラベル	正常分類数	悪性分類数	正答率	全体の正答率
良性	10	0	1.0	0.88
悪性	5	15	0.75	
転移	0	10	1.0	

表 4.10: 切片画像に対する分類結果 (文献 [8])

ラベル	正常分類数	悪性分類数	正答率	全体の正答率
良性	9	1	0.9	0.78
悪性	8	12	0.6	
転移	0	10	1.0	

表 4.9 を見ると，分類全体の正答率が 8 割を超えている．また，表 4.10 の従来手法の精度よりも全てのラベルにおいて高いことが確認できる．特に，良性と転移とラ

ベル付けされている画像は正確に分類できている。さらに、悪性とラベル付けされている画像は従来手法よりも誤分類している画像が少ないことがわかる。以上のことから従来手法よりも提案手法の有効性が確認できる。しかし、以前として悪性とラベル付けされている画像に関しては正常と分類してしまっているものもある、原因としては、悪性は正常細胞と癌細胞が入り混じっているものであるため癌細胞を正常細胞として誤分類してしまっていることが考えられる。

病理画像全体に対する分類実験

良性とラベル付けされた画像2枚、悪性とラベル付けされた画像2枚、転移とラベル付けされた画像2枚の計6枚の病理画像全体に対して分類実験を行った。実験に用いたそれぞれの病理画像の一部を図4.9, 4.10, 4.11に示す。実験の評価では、正常と分類された細胞核数、癌と分類された細胞核数のみ評価を行い、正常細胞と分類された細胞核が多い場合は正常、癌細胞と分類された細胞核が多い場合は癌とする。実験の結果を表4.11に示す。また、比較対象である文献[8]の結果を表4.12に示す。

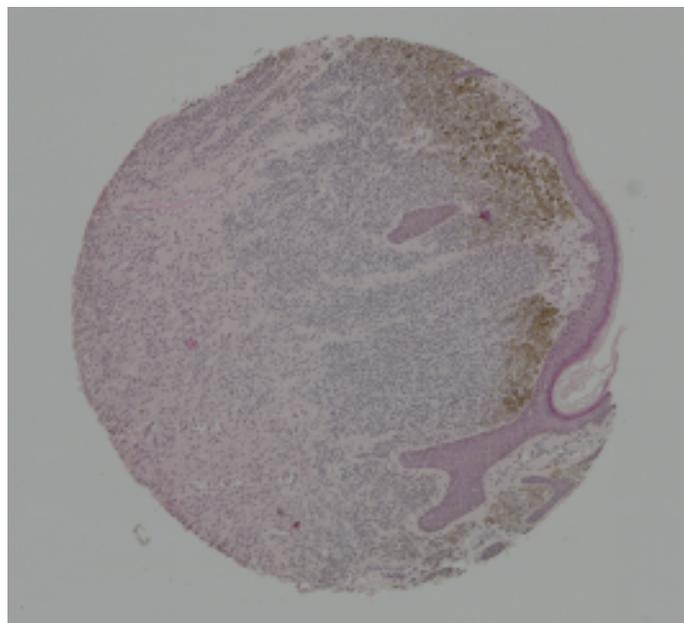


図 4.9: 良性腫瘍の病理画像全体

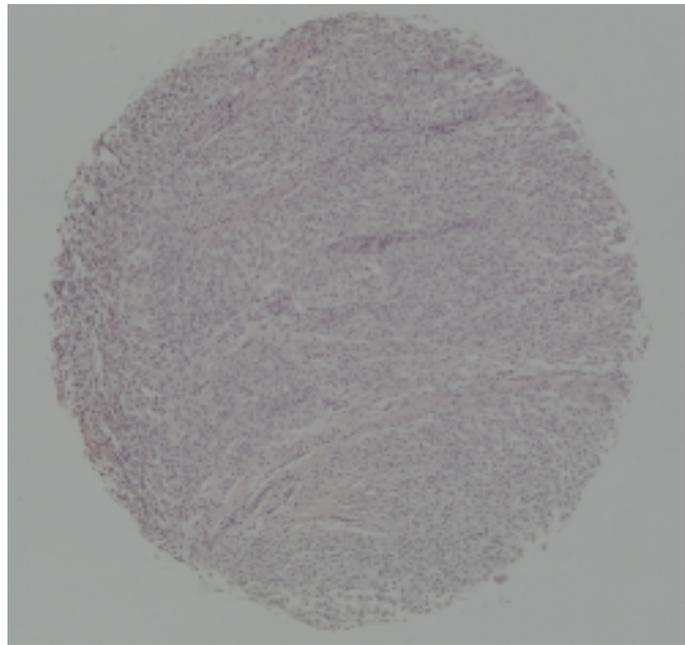


図 4.10: 悪性腫瘍の病理画像全体

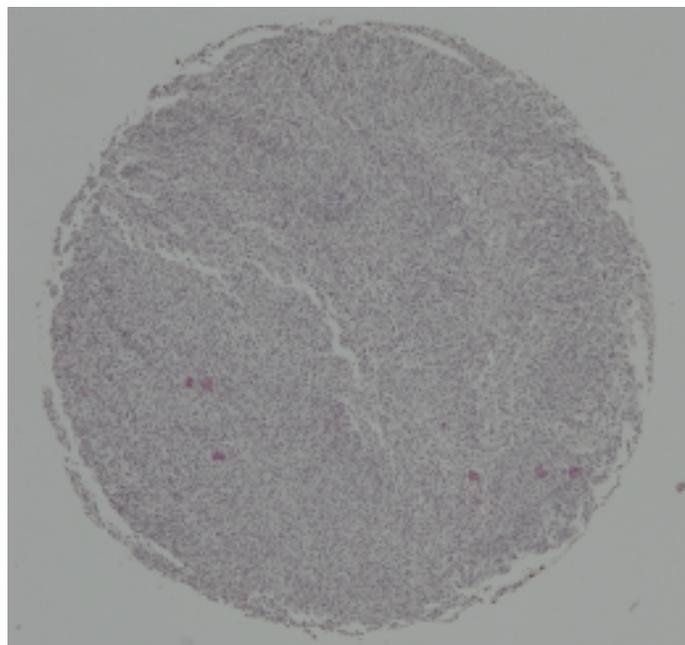


図 4.11: 転移腫瘍の病理画像全体

表 4.11: 病理画像に対する分類結果

ラベル	番号	正常細胞核数	癌細胞核数	分類結果
良性	1	1807	1038	正常
	2	2264	1723	正常
悪性	1	1414	1422	癌
	2	1886	1979	癌
転移	1	1785	2475	癌
	2	1588	1761	癌

表 4.12: 病理画像に対する分類結果 (文献 [8])

ラベル	番号	正常細胞核数	悪性細胞核数	分類結果
良性	1	1363	998	正常
	2	1557	742	正常
悪性	1	1021	934	正常
	2	754	933	癌
転移	1	412	1901	癌
	2	383	1943	癌

表 4.11 を見ると、病理画像においても切片画像の実験と同様に良性と転移は分類結果が明確になっていることが確認できる。悪性に関しては切片画像での実験では誤分類してしまっているものがあつたが病理画像全体で見るとしっかりと分類できている。また、表 4.12 の従来手法の正常と誤分類している病理画像でも提案手法では癌と分類できていることがわかる。以上より、病理画像全体においても提案手法の有効性が確認できる。

本研究ではメラノーマを分類対象としており、メラノーマの特性上病理画像内の正常細胞と癌細胞の数で癌であるか否かを判断している。しかし、癌には様々な種類があるため、その判断基準が使えるとは限らない。そこで、今後は癌のステー

ジ(進行度)を考慮した分類を加えることで、より分かりやすい分類ができると考える。癌にはステージというものがあり、ステージ0～ステージIVの5段階に分けられている。ステージごとの特徴を学習させ、細胞核が抽出された画像を学習したモデルに入れることでその画像のステージがわかり、より詳しく分類され様々な癌へ適用できると考えられる。

第5章 おわりに

本論文では、2種類のCNNモデルを用いた細胞核の抽出と分類を行う手法を提案した。

抽出においては、小さいパッチ画像を学習用データセットとするU-Netを用いて細胞核抽出画像の作成を行った。実験結果から誤抽出および従来手法よりも抽出漏れが少ない結果となり、本手法の有効性が確認できた。

分類においては、CNNモデルのVGG16に学習済みネットワークを使用し新しいモデルを作成する手法であるFine-tuningを行い、細胞核分類用にモデルを作成した。皮膚がんの一種であるメラノーマに対して細胞核抽出画像を作成し、抽出された細胞より正常細胞か癌細胞の分類を行った。実験結果から従来手法よりも正答率が向上している結果となり、本手法の有効性が確認できた。

今後の課題として細胞核抽出の分野では、重複している細胞を分離する手法の提案、抽出漏れや誤抽出をなくすためにデータセットの見直し、更なる精度向上などが挙げられる。細胞核分類の分野では、癌のステージを考慮した分類手法の提案、メラノーマ以外の癌に対しての分類手法の提案などが挙げられる。

謝辞

本研究を行うにあたり，日頃から多大なご尽力を頂き，ご指導を賜った名古屋工業大学准教授 舟橋健司先生，中部大学教授 岩堀祐之先生に感謝いたします。

また，本研究に対して御討論，御協力頂きました中部大学教授 岩本隆司先生，喬善楼先生，並びに本研究に使用した細胞画像を提供して頂きました旭川医科大学准教授 上田潤先生に深く感謝いたします。

最後に，本研究を行うにあたり多大な御指導，御協力を頂きました舟橋研究室の皆様，並びに岩堀研究室の皆様に感謝申し上げます。

参考文献

- [1] Nakhleh. R, Coffin. C, Cooper. K, “Recommendations for quality assurance and improvement in surgical and autopsy pathology” *Hum Pathol*, 37, pp. 985-988, 2006.
- [2] 日本病理医学会 “病理専門医一覧”
- [3] 栗原司, 野里博和, 坂無英徳, 高橋栄一, 寺井謙介, 徳山宣, 蛭田啓之, 古谷立美, “高次局所自己相関特徴による病理画像からの異常検出手法”, 研究報告バイオ情報学 (BIO) 2010-BIO-23(32), pp.1-6, 2010.
- [4] 藤田宗佑, 廣安知之, 渡辺章人, 三木光範, 小掠真貴, 福本学, “病理画像を用いた画像処理フィルタ構築における GP と SAP の比較”, 研究報告数理モデル化と問題解決 (MPS) 2009-MPS-75(12), pp. 1-6, 2009.
- [5] 三村 勇介, 尾崎 雄一, 一谷 修司, 平澤 宏祐, “医用画像における細胞認識技術”, KONICA MINOLTA TECHNOLOGY REPORT. Vol. 13, 2016.
- [6] Yuya Tsukada, Yuji Iwahori, Kenji Funahashi, Mami Jose, Jun Ueda, Takashi Iwamoto, “Extraction of Cell Nuclei using CNN Features”, *Procedia Computer Science*, Elsevier, Vol.112, pp.1633-1640, 2017.
- [7] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “Imagenet classification with deep convolutional neural networks”, in *Advances in neural information processing systems*, pp. 1097-1105, 2012.
- [8] Yuji Iwahori, Yuya Tsukada, Takashi Iwamoto, Kenji Funahashi, Jun Ueda, M. K. Bhuyan, “Classification of Cell Nuclei using CNN Features”, *International Conference on Intelligence Science*, Springer, pp.195-208, 2019.

- [9] 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. Springer, pp. 234-241, 2015.
- [10] K. Simonyan, A. Zisserman, “Very deep convolutional networks for large-scale image recognition”, arXiv preprint arXiv:1409.1556, 2014.
- [11] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, “Going deeper with convolutions”, in Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 1-9, 2015.
- [12] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition”, in Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 770-778, 2016.
- [13] DeepAge,
https://deepage.net/deep_learning/2016/11/07/convolutional_neural_network.html
- [14] X. Glorot, A. Bordes, and Y. Bengio, “Deep sparse rectifier neural networks”, in Proceedings of the fourteenth international conference on artificial intelligence and statistics, pp. 315-323, 2011.
- [15] Arthur E. Hoerl and Robert W. Kennard, “Ridge Regression: Biased Estimation for Nonorthogonal Problems”, American Society for Quality, Technometrics, Vol. 12, No. 1, pp. 55-67, 1970.
- [16] Nitish Srivastava, Geoffrey Hinton, Alex Krizhevsky, Ilya Sutskever, Ruslan Salakhutdinov. “Dropout: A Simple Way to Prevent Neural Networks from Overfitting.”, The Journal of Machine Learning Research, Volume 15, Issue 1, pp.1929-1958, 2014.
- [17] R. Tibshirani, “Regression shrinkage and selection via the lasso”, Journal of the Royal Statistical Society, Series B, Methodological, 58(1), pp.267-288, 1996.

発表論文リスト

1. 満安佑亮, 舟橋健司, 岩堀祐之, 小笠原尚高, 春日井邦夫, “内視鏡画像におけるアイコナル方程式に基づくポリープの形状復元”, A1-3, 医用画像, 令和元年度電気・電子・情報関係学会東海支部連合大会, 2019.
2. 満安佑亮, 岩堀祐之, 舟橋健司, 岩本隆司, 上田潤, “Fine Tuning を用いた細胞核の分類”, P237, 知的画像処理, 第 17 回情報学ワークショップ (WiNF 2019), 2019.