

平成 31 年度 修士論文

消化管腫瘍における 3 次元病理画像の
深層学習による診断支援

東京大学大学院工学系研究科
電気系工学専攻

修士課程 2 年
学籍番号 37-176484

松崎 博貴

2019 年 1 月 31 日

概要

近年になって医療データが蓄積されるシステムができてきている。医療画像をコンピュータ上に取り込むようになってから、自動で解析するシステムを構築してきた。今までは、専門医が診断を行う場合は、判断が主観的であること、ミスをする可能性があること、専門化同士でも意見が異なること、実際の臨床現場で医師はたくさんの画像を処理しなくてはいけないので、1枚にかけられる時間が限られていることから、機械による診断支援システムが必要とされている。既存の機械によるルールベースによるシステムでは解析ルール人にバイアスがかかっているということや、画像認識の著しい精度向上があるディープラーニングを利用した方法によって解析する方法に注目されている。またコンピュータの計算性能と、ビッグデータである医療画像を処理する能力が向上した背景と重なって、現在は機械学習による医療画像解析の研究が盛り上がりを見せている。

目次

第 1 章	序論	1
1.1	研究背景	1
1.2	本研究の目的	2
第 2 章	深層学習による病理画像の診断支援	5
2.1	ニューラルネットワーク	5
2.1.1	多層パーセプトロン	5
2.1.2	畳み込みニューラルネットワーク	6
2.1.3	再帰的ニューラルネットワーク	6
2.2	画像認識におけるディープラーニング	7
2.3	深層学習による 3 次元画像解析	9
2.3.1	3DCNN と Stacked Convolution	9
2.3.2	LSTM と 2DCNN の組み合わせ	9
2.4	教師なし学習	10
2.4.1	Autoencoder	10
2.4.2	Variational Autoencoder	10
2.4.3	敵対的生成ネットワーク	10
2.5	半教師あり学習	11
第 3 章	自分の研究本体を述べるところ	13
第 4 章	結論	15
4.1	少量データに対する解析手法	15
4.2	3 次元画像	15
4.3	半教師あり学習	15
4.4	今後の展望	16
付録		17
A	すごい長い証明	17
B	すごいたくさんのフィットの図	17
謝辞		19

図目次

1.1	transparent specimen	1
1.2	microscope	2
1.3	whole image analysis	2
2.1	7
2.2	8
2.3	9
2.4	10
2.5	11
2.6	12

表目次

第 1 章

序論

1.1 研究背景

胃がんは、日本でもっとも罹患数が多いがんであり，年間 13 万人が罹患し 5 万人が亡くなっている．内視鏡の生検は食道，大腸，小腸，胃などを鉗子で 2 mm 角程度の立方体で抜き出し，染色したものを 2, 3 断面にカットしてから病理診断医が顕微鏡で観察，診断している．断面のみの観察では内部に腫瘍がある場合に見落としてしまうリスクがあるため，本研究では組織透明化技術を用いて検体を丸ごと観察することを試みた．

病理医が全てを診断するには負担が大きいこと，病変の見落としリスクがあること，経験を詰んだ医師でも判断の難しい病変を見つけることが求められている．

内視鏡で見つけた病変が，腫瘍か非腫瘍かを判別したり，その悪性度を判定したりすることは生検を観察する必要がある．内視鏡検査の専門医のマンパワーが不足している．現在の日本の病理の専門医は 2483 名^[2]で，その人数に対して標本 500 万件を年間で処理している現状である．

また癌検診では，担当医が判定して専門医がダブルチェックをしているが，AI がスクリーニングを行って専門医の確認が必要な症例を絞り込む．専門医の判定は主観的であることや，専門医ごとにも判断基準がことなることが課題になる．

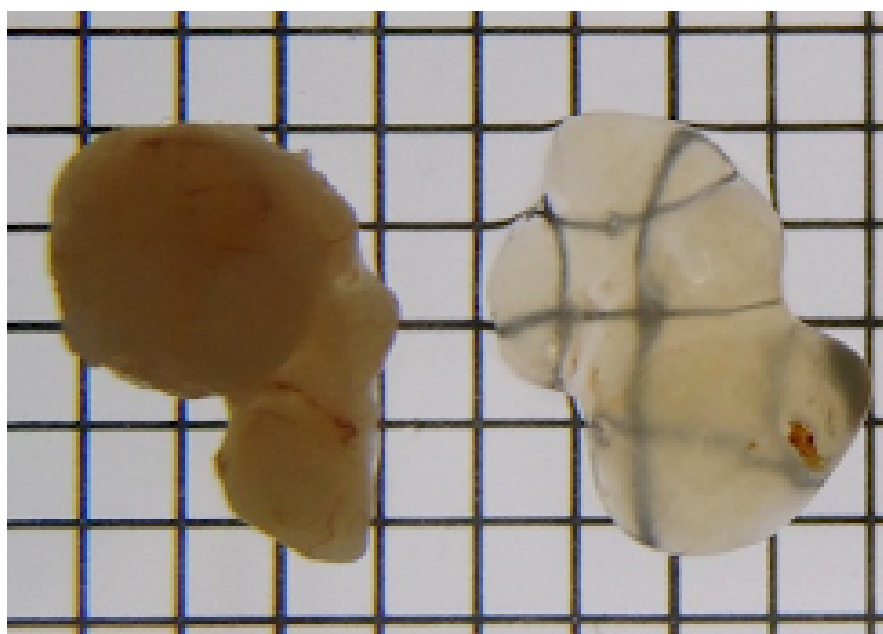


Fig. 1.1: transparent specimen

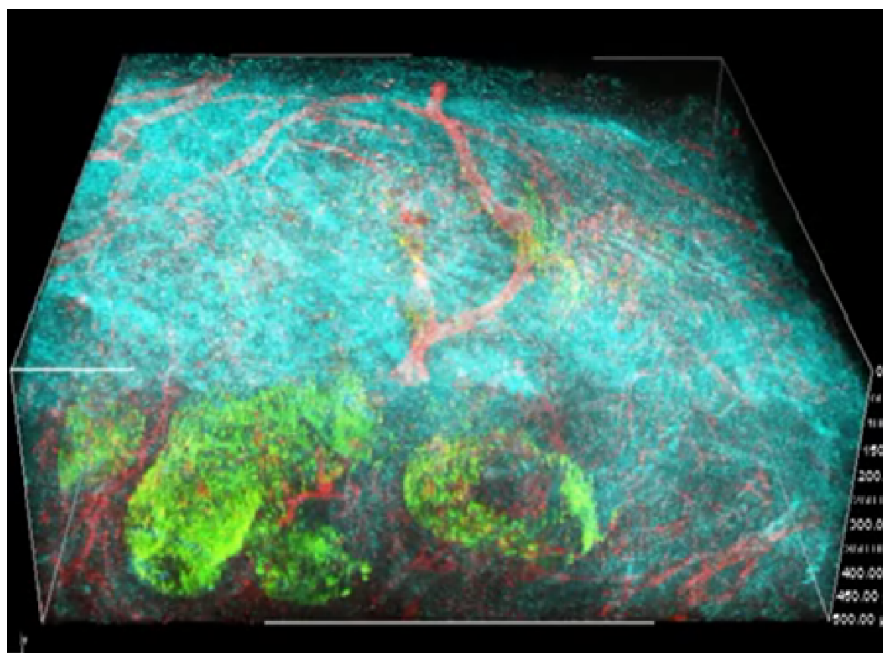


Fig. 1.2: microscope

1.2 本研究の目的

本研究の目的は内視鏡生検を透明にして深層学習を利用して癌を見落とさない診断方法を開発することである。組織透明化技術 LUCID を用いて検体を丸ごと透明化し、レーザー顕微鏡で観察することで、顕微鏡の解像度で検体の内部まで 3 次元情報として解析することができる。

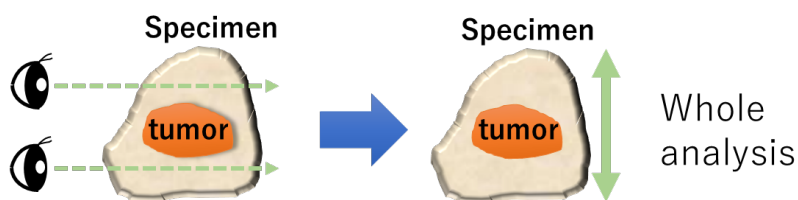


Fig. 1.3: whole image analysis

透明化されたサンプルをレーザー顕微鏡で撮影すると、従来のカットする手法に換算して 1000 カットに相当するため、得られる顕微鏡撮影像が従来の 2,3 カットに対して数 100 倍となるので、専門医が診断をするには負担が大きくなってしまう。したがって、人工知能 (Artificial Intelligence: AI) が 3 次元画像を解析して病変を検知し、専門医に提示することで病変の見落としリスクがゼロになる診断支援システムを開発することが本研究の目的である。

AI を用いた内視鏡の診断支援は生検ではなくて、内視鏡カメラの画像解析を行う事例がある。国立がん研究センターと NEC が共同で開発したシステムがある。しかし、診断を確定するには生検が必須であり、生検の病理診断にマンパワーが必要になっている。生検の AI による診断アルゴリズム開発の先行研究としては、胃がん、肺がんや乳がん、骨髄などの生検をディープラーニングで境界認識するものがあり、90% 程度の正解率を出して、専門医と同程度の精度がすでに達成されているものもあるがデータセットとして提供されているのは単純の形状のものが多い。本研究では生検の 3 次元画像を解析することで 2 次元画像では判断の難しいものを 3 次元特有の情報を用いる

ことで判定精度を上げ、機械学習を行うためには教師データを用意する必要があるが、医療データはアノテーションされていないことが多い。3次元画像を解析する先行研究としては、CTやMRI画像が上げられる。CTやMRIの場合は画像の分解能が顕微鏡像よりも低く、深さ方向にも30程度である。本研究は透明化処理した生検を蛍光顕微鏡で撮影するため分解能が高く、深さ方向にも400 700枚が撮影することができる。このためディープラーニングで解析するときに計算のメモリに乘らないという問題が生じるため解析手法を工夫して計算量がなるべく減るようにしなければいけない。今回は時系列処理で使われる手法と画像処理の手法を組み合わせることで、大きな画像サイズでも解析することができるようにした。

深層学習で解析を行う場合はデータを大量に準備する必要がある。今回のように、新しい撮影手法であったり、希少な病気であったりすると医療データを数多く集められないことがある。さらに画像データだけでなく、その画像に教師ラベルを貼る(アノテーション)必要がある。画像と、その教師ラベルのセットを数万枚と集めることができたら高い精度が出るという報告が多くあるが、少ない教師データで識別精度を上げることは深層学習において困難とされている。教師データの作成が困難であるだけでなく、少量データである場合は、その教師データが医師によってばらつきがある場合は、教師データを作成した医師の判断が大きく反映されてしまい、必ずしも真であると断言することができなくなってしまう。これらの課題があるため医療画像における深層学習を用いた診断システムは、大量にデータを用意するまでの時間や人的コストを大きく払うことになっているのが現状である。

本研究でこの課題に取り組んだ。教師ラベルを必要とせず、画像の構造的な特徴のパターンを学ぶ、教師なし学習の手法と、これまでの教師ラベルを使った学習とを組み合わせ、半(弱)教師あり学習を行った。構造的なパターンは3次元画像であることを活用し、教師なし学習については、病理画像が正常から異常になる過程には連続的な性質があることを活用して正常と異常の分布が作れるようにした。

最後にこの深層学習による腫瘍の検出結果を医師が診断する際のサポートになるように可視化する。深層学習は、よく判断の理由が分からないためブラックボックスと呼ばれ、信頼性が求められる場面では利用に対して不安視されることがある。そのため腫瘍と正常の診断の判断の理由が可視化して、どの部分に注目しているかを医師に見せる。医師の負担を減らすためのスクリーニングとして利用するために、明らかに正常のものと、腫瘍かもしれない領域とを区別して、腫瘍かもしれない部分を医師に提示する。3次元画像であるから医師が診断するには、全ての断面を確認することは現実的に不可能である。腫瘍を見落とさないように、深層学習によって、腫瘍がある確率の高い断面を提示して診断支援を行う。

本研究の目的は内視鏡生検を透明にして深層学習を利用して癌を見落とさない診断方法を開発することである。組織透明化技術 LUCID を用いて検体を丸ごと透明化し、レーザー顕微鏡で観察することで、顕微鏡の解像度で検体の内部まで3次元情報として解析することができる。

透明化されたサンプルをレーザー顕微鏡で撮影すると、従来のカットする手法に換算して1000カットに相当するため、得られる顕微鏡撮影像が従来の2,3カットに対して数百倍となるので、専門医が診断をするには負担が大きくなってしまう。したがって、AI(人工知能)が3次元画像を解析して病変を検知し専門医に提示することで病変の見落としリスク(偽陰性)がゼロになる診断支援システムを開発することが本研究の目的である。

深層学習で解析を行う場合はデータを大量に準備する必要がある。今回のように、新しい撮影手法であったり、希少な病気であったりすると医療データを数多く集められないことがある。さらに画像データだけでなく、その画像に教師ラベルを貼る(アノテーション)必要がある。画像と、その教師ラベルのセットを数万枚と集めることができたら高い精度が出るという報告が多くあるが、少ない教師データで識別精度を上げることは深層学習において困難とされている。教師データの作成が困難であるだけでなく、少量データである場合は、その教師データが医師によってばらつきがある場合は、教師データを作成した医師の判断が大きく反映されてしまい、必ずしも真であると断言することができなくなってしまう。これらの課題があるため医療画像における深層学習を用いた診断システムは、大量にデータを用意するまでの時間や人的コストを大きく払うことになっているのが現状である。

本研究でこの課題に取り組んだ。教師ラベルを必要とせず、画像の構造的な特徴のパターンを学ぶ、教師なし学

習の手法と，これまでの教師ラベルを使った学習とを組み合わせ，半（弱）教師あり学習を行った．構造的なパターンは 3 次元画像であることを活用し，教師なし学習については，病理画像が正常から異常になる過程には連続的な性質があることを活用して正常と異常の分布が作れるようにした．

第 2 章

深層学習による病理画像の診断支援

病理画像をデジタルで保存することが始まったのは数十年前になる。これによって遠隔地でも診断することができるようになったり、情報を共有することができるようになり、複数の医師で診断しミスを防止するセカンド・オピニオンが容易になった。計算機科学の分野の側面ではデータを収集することができるようになり、研究が盛んに行われることになった。その後は、様々な病理データでより改良されたアルゴリズムの提案が行われている。

細胞組織の形態を観察するための病理染色ではヘマトキシリン・エオジン染色 (HE 染色) が一般的に用いられる。細胞核を青紫色に染色し、細胞質をピンク色に染色する。正常から異常に変化していくと、細胞核が過度に増殖したり、細胞質の形が崩れたりすることで、その特徴を機械学習によって精度よく検出するための研究が行われている。

これまでは、核の形やテクスチャーからパターンマッチングなどの画像処理によって腫瘍を検出する研究されてきたが、近年になって画像処理に大きなブレイクスルーが起きたことをきっかけに、新しい手法で解析するようになってきた。そのブレイクスルーがディープラーニングである。

2.1 ニューラルネットワーク

人間の脳にはニューロンと呼ばれる神経細胞が 1000 億個以上あり、それぞれが複数のニューロンと電気信号で情報を伝達している。また脳には電気信号を受け渡すシナプスという場所がある。ニューロンとシナプスで行われる演算を模倣したアルゴリズムを作ることができれば人間のような思考や認識をコンピュータを使って再現できると考えた。そのアルゴリズムがニューラルネットワークである。

2.1.1 多層パーセプトロン

ニューラルネットワークは入力層、出力層、隠れ層から構成され、層と層の間にはニューロン同士のつながりの強さを示す重みがある。非線形問題を扱うために 1986 年 Rumelhart によって考案されたのが、パーセプトロンを複数つなぎ合わせ入力と出力以外に隠れた層を持つ多層パーセプトロン (Multi-layer perceptron: MLP) である。

ニューラルネットワークで多層パーセプトロンの層を全結合 (fully connected: FC) 層とも呼ぶ。Figure 3.1 における丸や矢印はそれぞれノード (またはニューロン) と重み (または結合) と呼び、ともに数値である。例えば画像を分類しようと思えば、各ピクセルの画素数を各ノードに入力する (28×28pix のグレースケール画像であれば 784 個のノードが必要)。データが入力層に入ってくると、その値に重みをかけ、活性化関数と呼ばれる関数を通し結果を出力する。これを繰り返し出力層に書き出す。各層の重みの値によって出力結果は異なってくる。出力層のノードは区別したいクラス数分用意し、各ノードの出力値が各クラスに属している確率を表す。学習については誤差逆伝播法を利用する。ニューラルネットワークの出力値と正解データとの比較をした時に、どれだけ正解から離れているかを評価する損失関数 (Loss) を使って、損失関数が小さくなるようにノード間の重みを勾配降下法によって更新

する。

2.1.2 畳み込みニューラルネットワーク

畳み込みニューラルネットワーク (Convolutional Neural Network: CNN) は、1998 年には既に LeNet と呼ばれるネットワークで実装されていた [3]。従来の画像認識では、画像から特徴を抽出し、それをニューラルネットワークにかけるいわゆる特徴量設計が必要で、ここをいかにうまく設計するかがポイントであったが、CNN は特徴量設計から識別までを end-to-end で行うことができることが最大の強みである。人間が物体を認識することをコンピュータにも計算させるには、画像の特徴的な部分を切り分けて数値化させる必要がある。例えば、カラー画像の場合、RGB の 3 色 (3 チャンネル) を組み合わせた画像で認識をしている。このようなフィルターの畳み込み計算を行うと、フィルターごとに異なった画像の特徴を抽出して数値化する。これが畳み込み (convolution) である。その後、画像のサイズを小さくしてコンピュータが計算コストを減らし、微小な変化に対してロバストになる仕組みしてプーリングという方法を用いる。学習には MLP と同様に誤差逆伝播法を用いる。

2.1.3 再帰的ニューラルネットワーク

再帰的ニューラルネットワーク (Recurrent Neural Network: RNN) は時系列解析や自然言語処理に利用されるニューラルネットワークである。内部にループを持つことで過去の情報を保持しておくことができる。時系列の入力 $x = (x_1, \dots, x_T)$ があつた時に、出力 $y = (y_1, \dots, y_T)$ と隠れ層のベクトル $h = (h_1, \dots, h_T)$ をそれぞれ以下の式で計算する。

$$h_t = H(W_{ih}x_t + W_{hh}h_{t-1} + b_h) \quad (2.1)$$

$$y_t = W_{ho}h_t + b_o \quad (2.2)$$

W は重みの行列であり (W_{ih} は入力と隠れ層間の重み行列), b はバイアス項である。そして H が活性化関数である。

RNN は過去の情報をどこまでさかのぼって関連性を見つけるかを判断することができないため、時系列データが長くなるほど、その長期の依存性を学習するには人が慎重にパラメータを設計する必要があるなど、学習が難しくなるという問題があつた。この問題を解決するために Long-short term memory (LSTM) が提案された。LSTM も RNN の一種ではあるため、繰り返し構造を持ち、3 つのゲートを持つ層からなっている。

$$f_t = \sigma(W_{xf}x_t + W_{hf}h_{t-1} + W_{cf}c_{t-1} + b_f) \quad (2.3)$$

$$i_t = \sigma(W_{xi}x_t + W_{hi}h_{t-1} + W_{ci}c_{t-1} + b_i) \quad (2.4)$$

$$c_t = f_t c_{t-1} + i_t \tanh(W_{xc}x_t + W_{hc}h_{t-1} + b_c) \quad (2.5)$$

$$o_t = \sigma(W_{xo}x_t + W_{ho}h_{t-1} + W_{co}c_t + b_o) \quad (2.6)$$

$$h_t = o_t \tanh(c_t) \quad (2.7)$$

ここで σ はシグモイド関数であり、次の式で定義される。

$$\sigma(x) = \frac{1}{1 + e^{-x}} \quad (2.8)$$

f_t で示される層は忘却ゲート層と呼ばれ、過去の情報で捨てるべき情報を判断する。これはシグモイド層によって行われ、0 と 1 の間の値を出力し、0 は完全に忘れる。1 は完全に維持するという意味である。 i_t や c_t で示される層は、入力ゲート層と呼ばれ、 i_t で新たに入力された情報から、どの情報を更新するかを判断し c_t で古い情報を

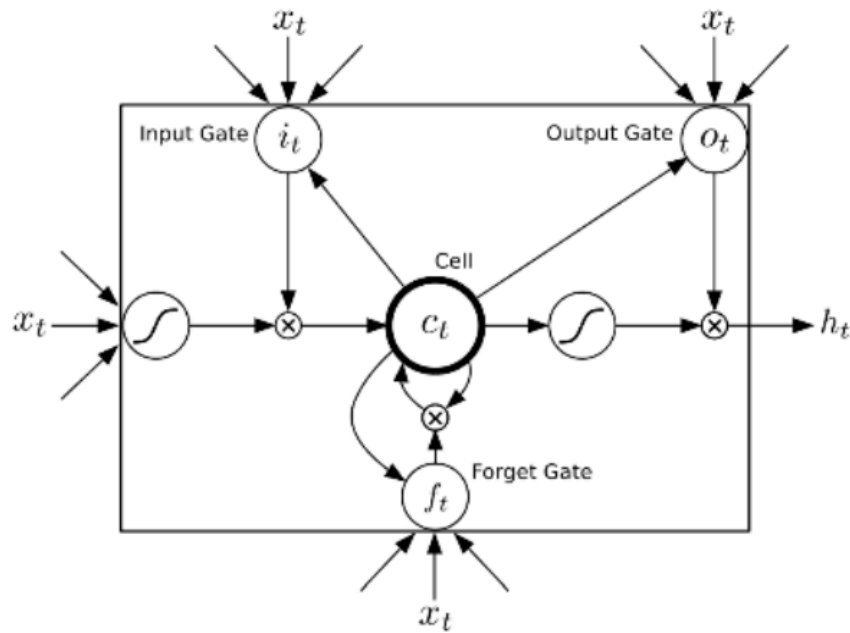


Fig. 2.1

落とし新しい情報を加え、値を更新する。最後が o_t や h_t で示される層で、出力ゲート層と呼ばれる。まず何を出すべきかを o_t のゲートで判断して、 c_t に \tanh を適用して掛けることで出力が計算される。

2.2 画像認識におけるディープラーニング

Deep Learning とは Deep Neural Network(DNN) を指すことが多い。この"Deep"とは、ニューラルネットワークの層が深いことに由来している。Figure 3.3 に画像認識タスクの精度の近年の推移を示す。これは ImageNet Large Scale Visual Recognition Challenge (ILSVRC) と呼ばれる世界的な画像認識のコンペティションである (2010 年から始まった)。カテゴリ数は 1000 クラスで、画像枚数は 120 万枚の訓練データと 15 万枚のテストデータが用意されている。2011 年と 2012 年は約 10% もの大差で AlexNet[4] が優勝している。これがディープラーニングの始まりである。AlexNet は 5 つの畳み込み層と 3 つの全結合層を持っている。2014 年には VGGNet[5] や GoogLeNet[6] が 9 割の精度を超えた。VGGNet は AlexNet(8 層) よりさらに深い構造 (19 層) であり、GoogLeNet は 22 層もある。そして 2015 年には ResNet[7] が人間の精度をも超える認識精度を達成した。ResNet は GoogLeNet よりもさらに深く 152 層もある。一般に、層を深くすることは簡単ではなく、勾配消失や過学習といった問題が起こる。CNN を複数回かけて検出を行う場合、CNN の浅い側では空間分解能はあるが抽象的な情報が少ない。深い側では意味論的な情報は取得できる (ポーズ、変形など) が空間分解能が小さいため幾何学的な情報が失われる。アーキテクチャの進化の方向は①層が深くなること ② FC 層の使用を避けること又は Inception モジュールの使用によってパラメータ数を削減すること ③ ResNet などのショートカット接続によって学習効率をあげること事前学習・転移学習を行うことでモデルの精度を上げる取り組みがある。当然ながらこの場合事前学習のデータセットと適用データとの間に類似性があると良い。

画像処理におけるディープラーニングでは大きく 3 つのタスクがあり、それぞれ、クラス分類、物体検出、セグメンテーションである。以下に詳細を述べる。

クラス分類

カテゴリ分けをすること

物体検出

物体検出とは **Bounding Box** で物体の位置とその物体の種類を特定する方法である。歴史的には幾何的情報、手動特徴量、そしてそのカスケードを利用していた。その後、HOG や SIFT など局所特徴量を抽出する方法を設計するようになったが、これは深い専門知識を必要とした。また広い範囲でオブジェクトを正確に検出する方法は、メモリ容量と処理時間に課題がある。現在は **Deep Neural Network** になりデータのみから抽象的な特徴量を複数得ることができる。一般物体認識の場合は数千のカテゴリを学習して **Top Error Rate** が 2% 以下と人間よりも認識精度が高いが、物体検出においては、今はカテゴリがせいぜい数百程度くらいまででも認識精度が人間よりも低くなってしまう。また物体検出は精度を上げるために処理に時間がかかるアルゴリズムであることが多いため、リアルタイムに物体検出を行う時は、速度と精度でトレードオフが生じてしまう。

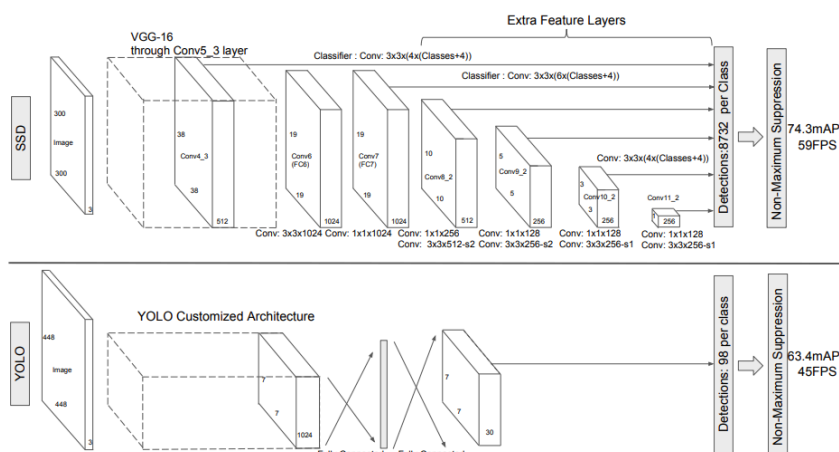


Fig. 2.2

セグメンテーション

セマンティックセグメンテーションとは、画像を画素レベルで認識することである。画像内の各画素をオブジェクトクラスに割り当てる手法である。セマンティックセグメンテーションの手法についてディープラーニング以前では、Texton Forests や、Random Forests に基づいた分類を行っていたが、物体検出と同様に CNN が登場してからは、高精度なセグメンテーションが実現するようになった。CNN を使ったセグメンテーションの手法で一般的に使われるようになったものが **Unet** である (Figure 2.3)。この Unet は文字通り U の形をしたネットワークであることが特徴で、2つのアーキテクチャーからできている。一つ目がエンコーダーのアーキテクチャーで CNN とプーリングで特徴を抽出しながら次元を削減していき、2つ目のデコーダーのアーキテクチャーで画像をセグメンテーションの結果になるように復元する。ここで問題になることが、プーリングをすることで位置情報を消してしまっているため、この位置情報を利用して画像を復元するためには、エンコーダーとデコーダーで画像サイズが同じところ同士をショートカットで接続することが Unet 構造の優れている点である。

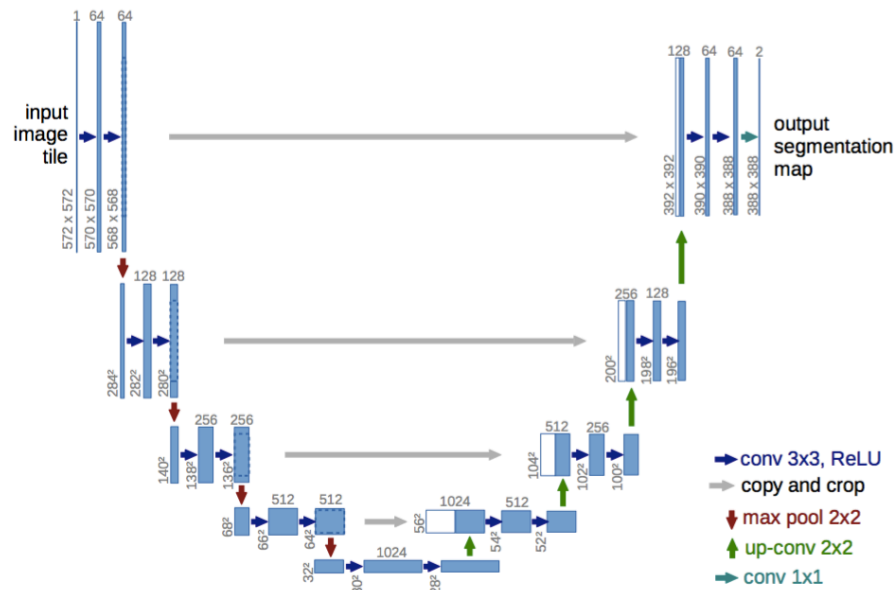


Fig. 2.3

2.3 深層学習による 3 次元画像解析

ディープラーニングを医療画像に応用するコンペティションが世界で行われているが、その半数が 3D 医療画像の解析になっているほど需要が高まっている。その理由は、現在解決しなくてはならない課題があるからである。まずは 2 次元画像と違って、処理すべきデータが大きいということである。そのため学習するパラメータをなるべく少なくする工夫がされている。また 3 次元画像には、動画または、ボリューム画像があるが、2 次元画像とその深さ方向（動画であれば、時間方向）には異方性があることから、機械学習の方法に工夫が必要になる。今まで考案されている手法として、2DCNN を拡張した 3DCNN、また CNN と時系列解析でよく用いられる LSTM を組み合わせた手法と、LSTM 内部に CNN を組み込んだ手法、それらをすべて組み合わせた手法が考案されている。LSTM の研究も盛んに行われているため、その改良モデルが数多く存在する。特に、LSTM の学習効率を上げた GRU(Gated Linear Unit) や、順方向だけでなく逆方向の時系列も計算に入れる BiLSTM が時系列解析の精度向上になっている報告がある。

2.3.1 3DCNN と Stacked Convolution

2 次元画像が深さ方向に連続している 3 次元画像の特徴を抽出するために、2 次元の CNN を拡張して、3 次元のカーネルを使って畳み込みを行う、3DCNN を利用した手法がある。

2.3.2 LSTM と 2DCNN の組み合わせ

時系列解析に使われる LSTM を用いて 3 次元の画像を解析することができる。これはよく動画の解析で行われることがある。つまりフレームごとの画像の特徴を 2 次元の CNN で計算してから時系列情報を LSTM で解析することで、画像の時系列解析を行うことができる。これを 3 次元の医療画像で CT や MRI で適用する研究も行われている。3DCNN のデメリットであったパラメータの増大を 2DCNN と LSTM の組み合わせで解決することが

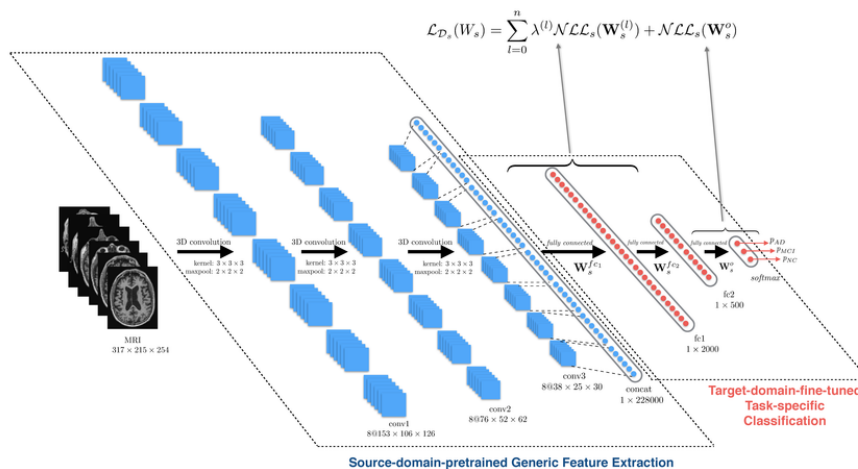


Fig. 2.4

できる。

2.4 教師なし学習

機械学習の手法には、上記で説明したように、ラベルの貼られているデータセットを用いて学習することを教師あり学習と呼び、その反対で、データセットはあっても、そのデータセットの特性を示したラベルが与えられていない場合のデータセットを用いて学習することを教師なし学習と呼ぶ。

2.4.1 Autoencoder

教師なし学習で画像の特徴を抽出する方法としてオートエンコーダがある。画像の場合におけるオートエンコーダの手法とは、ある画像から情報を圧縮する「エンコーダ」と言われる部分と、その圧縮した情報から画像を復元する「デコーダ」の二つからなる。入力とデコーダから復元された画像が同じ画像になるようにニューラルネットワークで学習させる。この学習の結果、潜在変数は似てる画像どうして近い値になるように変化し、この分布を見れば画像の分類を教師ラベルがなくても、学習を行うことができる。

2.4.2 Variational Autoencoder

本研究では、このオートエンコーダの派生である。Variational Autoencoder(VAE)を利用した。これはオートエンコーダの「エンコーダ」と「デコーダ」は同じネットワーク構造であるが、データセットの潜在変数の分布が、正規分布になるような制約を加えて学習を行う手法である。こうすることで Autoencoder の潜在変数では分布の距離に意味がないが、VAE では正規分布に埋め込まれるため、画像の類似度を分布が表現することができるところが特徴である。

2.4.3 敵対的生成ネットワーク

敵対的生成ネットワーク (Generative Adversarial Network: GAN) は 2014 年に提案された手法である。概念図のように GAN では Generator と Discriminator の 2 つのネットワークがある。Generator は訓練データと同じような画像を生成するネットワークで Discriminator は、入力されたデータが訓練データから来たものか Generator

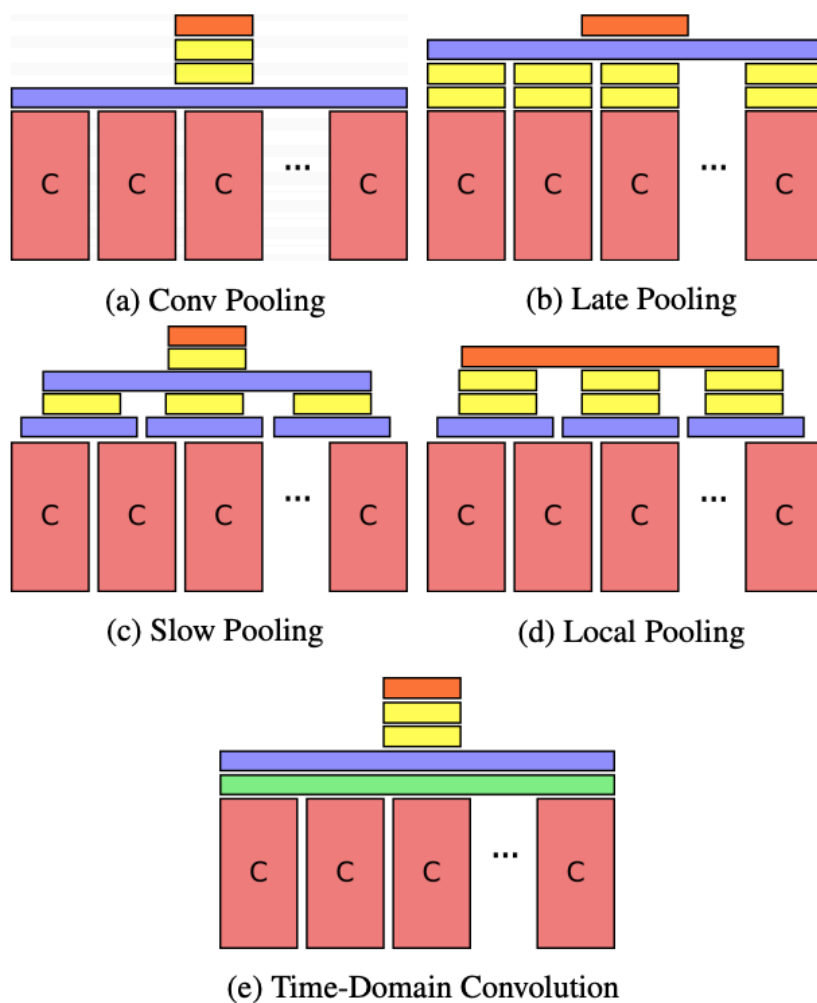


Fig. 2.5

で生成されたものかを識別するように学習する。VAE よりも GAN の方が細部まで鮮明に画像を生成することができる。しかし GAN は計算時間がかかるという問題や、Discriminator か Generator のどちらかが強くなってしまうなど、学習が安定しない問題があるため、これについての多くの研究報告がされている。

2.5 半教師あり学習

弱教師あり学習とも呼ばれる。これは、教師あり学習と教師なし学習を組み合わせる学習方法である。こうすることでデータに教師ラベルをつけているものが少数であっても、データの特徴を学習しながら少量のラベルで識別境界を決めることができる。GAN や VAE の考え方を発展させてネットワークを構築することが考えられる。

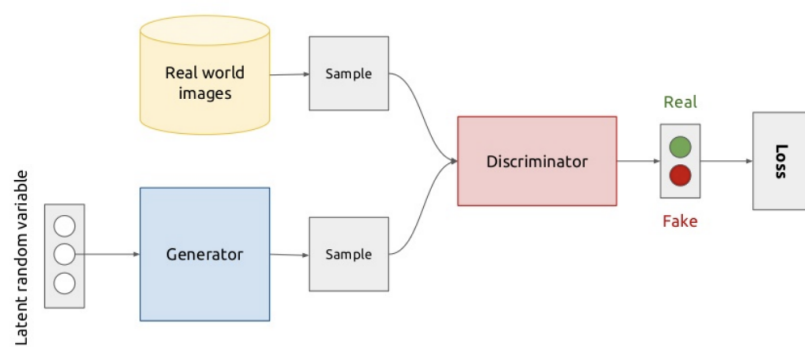


Fig. 2.6

第 3 章

自分の研究本体を述べるところ

ここは自分のやった研究を述べる章です。実際の中身に合わせて章を複数立てにする場合もあると思います。「議論」の章を別に分ける場合は、この章では得られた結果までを記述し、その結果に対する議論は「議論」の章に回すのが良いでしょう。この章は必ずしも 1 つの章のみである必要はありません。研究内容に応じて、複数の章に分割することも一般的に行われます。

修士論文で大切なことは、第 1 章や第 2 章で述べた伏線（研究の目的と動機）を回収するべく、きちんと研究内容を順序立てて書き、また自分の貢献を明確にすることです。論文全体で論理展開がきちんとしていれば良いので、必ずしも実際に行った実験などの時系列でこの章を書き進める必要はありません。また修士論文としての完成度が大切ですので、修士論文のテーマに直接関係のない自分のやったことを無理に混ぜる必要もありません。

第 4 章

結論

4.1 少量データに対する解析手法

2次元画像を入力にして、比較したのは、データ拡張、事前学習、擬似 HE 染色の 3 つの処理の有効性についてと、2次元の画像の特徴抽出に利用されるモデルの選択である。検体のサンプル数が少ない状況でかつ、サンプル間のデータのばらつきが大きい場合は、普通の学習方法では、過学習してしまい未知のサンプルの予測精度は大きく減少してしまう。それを防ぐための手法として、データ拡張によって、データの数を実感的に増やすことで学習データのバリエーションを増やすことや擬似 HE 染色による前処理によって、もとの顕微鏡撮影像よりも情報が落ちてしまうためデータがサンプル間のばらつきが少なくなるというメリットがある。さらには、世の中の一般的な手法は HE 染色なので、オープンソースされている画像が多く存在するため事前学習をすることができることで、精度を高めることができた。

4.2 3次元画像

3次元画像として MRI や CT の画像の解析例は多くあり、3DCNN を利用している場合がほとんどである。それは、MRI や CT の解像度が低いからであり、今回は深さ方向の解像度が 3 マイクロ程度であるため深さ方向に解析するとなると学習のパラメータが大きくなりすぎて学習が進まないという問題があった。これを解決するために動画解析で用いているような時系列解析を 3次元画像にも適用する。これで 2次元の畳み込みに加えて、時系列方向の変化を捉えることができるようになった。

4.3 半教師あり学習

教師ラベルが少ない場合は、教師データに対して過学習することが報告されている。これは学習したサンプル以外のデータを解析すると全く精度が出ないということになり、新しい検体に対して診断システムを利用できなくなってしまう。しかし、腫瘍と正常の形をよく見ると、腫瘍は正常だったものが乱れたり崩れていくことが分かる。正常に対して腫瘍は、異常検知をするようなことであるから、教師ラベルを貼らなくても、画像の特徴を学習することが可能であるということである。全くの教師ラベルなしでも特徴は捉えられるけれど、少量の教師データを入力することで、どれが正常で、どれが腫瘍かの認識をインストラクションするので、認識精度を向上させることができた。

4.4 今後の展望

最近の動画解析の研究で ConvLSTM という手法が出てきている。これは、LSTM の内部に CNN を入れていて、3DCNN がパラメータが大きくて学習が安定しないという問題と LSTM だけでは、空間的な相互作用が捉えきれないという問題を解決するという。これで本検体を解析することで、さらに 3 次元構造の特徴を良く捉えることができるか検討する。

GAN は、生成モデルとしては、成功しているが特徴抽出としては VAE の方が良いと言われているため VAE を利用したが、VAE の学習ではサンプルのデータの分布が正規分布に従うという過程と、Auto Encoder のように元画像に再生成できるかどうか誤差関数として導入されている。これに GAN の仕組みを組み合わせ、VAE で生成される画像がなるべく現実に近い画像となるように GAN を用いる VAE-GAN という手法で特徴抽出の精度を向上させることが期待される。

半教師あり学習は、教師ラベルがない画像と、教師ラベルのある画像を同時に使って学習することができる。教師ラベルがあれば一番精度がでるというわけではなく、人間でも判断が迷うようなものや、医師ごとに見解の異なるラベルなどが混在する場合がある。さらに、医師によってもベテランと研修医では精度に差が生じてしまう。これを解決するために、半教師あり学習のラベルの信頼度を入力に追加することができれば、信頼度によって、教師なし学習と教師あり学習の寄与率を変化させることができるはずである。

付録

「付録」(appendix)は、論文の本文に載せるには情報として邪魔もしくは必須ではないものの、読者にとって有益となるような情報を載せます。付録を必要としない論文もちろん存在しますので、そこは著者の判断です。

例えば、たくさんの観測データを様々なモデルでフィットした場合、フィット結果の絵がたくさん出てくるはずです。そのような図は本文中に大量に出されても大切な情報を見失ってしまいますので、大部分は付録に載せることが推奨されます。他には、何かしらの長い式変形や証明を載せる必要がある場合、付録に移動する場合があります。

A すごい長い証明

式 (A.1) のように、式番号がアルファベットとアラビア数字の組み合わせになるように、`LATEX` ソース中で設定してありますので、中身を眺めてみてください。

$$1 + 1 = 2 \tag{A.1}$$

B すごいたくさんのフィットの図

謝辞

「謝辞」(acknowledgments)は、修士論文を作成する上であなたを支えてくれた人への感謝を書く場所です。誰かへの感謝の気持ちを公開の場所で文書にするというのは気恥ずかしいものですし、修士論文以外ではそんなことをした経験がないかもしれませんが、投稿論文では一般的に行われます*¹。

多くの修士論文では指導教員(国立大学の法人化後は、指導教「官」とは言いません)、実験協力してくれた共同研究者、間接的に助言などをくれた研究室の他の教員・先輩・同輩・後輩、研究室の秘書さんなどに謝意を示すことが多いようです。もし奨学金をどこからか受給していたら、奨学金の出所に対しても謝辞を書いても良いでしょう。また家族・恋人・友人に対する感謝も見られますが、恋人の名前は将来隠したくなる場合もあるので注意しましょう。

感謝の気持ちを書く場所ですので、その相手に失礼のないようにしましょう。氏名の漢字の間違いや、職階の間違いが頻繁に見られます。特に助教を助教授と書き間違えたり、准教授を助教授としたりという間違いが目立ちます。

*¹ ただし、投稿論文では家族や友人への感謝はあまり書かず、研究費を出した機関や研究の協力をしてくれた研究者などを書くことが多いです。