

深層学習による材料分野の画像解析の発展

上 杉 徳 照*

1. はじめに

材料工学およびその関連分野では、微細組織の解析、製品の外観検査、破面の解析など、幅広い領域で画像解析技術が活用されてきた。材料分野における画像解析技術は、多くの場合、専門家による目視検査と手動分析に依存してきた。人による画像解析は強力かつ柔軟であり、広く受け入れられている。しかし、この方法は時間がかかり、解析者に過度な負担を強いる。さらに、個人の主観や疲労が影響するため、結果の一貫性と再現性に課題がある。加えて、解析者の属人的な暗黙知の伝承は難しい。

このような背景から、コンピュータを用いた画像解析が試みられてきた。しかし、従来のコンピュータを用いた画像解析技術には一定の限界があった。これらの問題を解決する方法として、最近では深層学習技術、特に畳み込みニューラルネットワーク (Convolutional Neural Network: CNN) が注目されている。CNNにより、材料分野における画像解析がより高精度かつ効率的に行われるようになりつつある。

本稿では、材料分野における画像解析の新しい技術として深層学習の応用に焦点を当てる。深層学習およびCNNの基本原則について解説し、材料分野における深層学習の応用と現状の問題点を概観する。また、具体的な成果事例を紹介し、この技術の実用化に向けた課題と展望について議論する。

2. 画像解析

画像解析とは、一般的にカメラで取得した画像をコンピュータで分析する技術のことで、画像データから対象に関する情報を抽出し、何らかの記述を得るプロセスを指す。例えば、明るさやコントラストの調整も画像解析だが、一様な値を設定して画像全体に処理を適用すると、部分的に明るすぎたり暗すぎたりする箇所が生じることがある。この問題に対

処するために、例えば適応的ヒストグラム平坦化 (Contrast Limited Adaptive Histogram Equalization: CLAHE)⁽¹⁾ のアルゴリズムが使用される。

適応的ヒストグラム平坦化は画像解析のアルゴリズムの一例であり、画像解析を行う際には、様々なアルゴリズムを組み合わせ、計測 (回帰)、認識 (分類)、検出、セマンティックセグメンテーション (領域分類) などのタスクを実行する。このプロセスは、まず入力画像の取り込みから始まり、前処理を施した後に、特徴量の抽出のための変換処理が行われる。その後、抽出した情報を基に、計測や分類などのタスクを行い、出力結果を得る。特徴抽出の方法には、深層学習を用いる方法と、深層学習以前の方法がある。

画像計測は、画像中の対象から数値データを算出することで、例えば外観検査での傷の数や大きさの数値化が該当する。画像認識は、画像に写っている対象が何であるかを特定することで、画像中の撮影対象をカテゴリーに分けることである。検出は、何が写っているかだけでなく、その位置も特定するものである。セマンティックセグメンテーションは、画像全体や特定部分の検出に留まらず、ピクセルレベルで各画素にラベル付けするタスクである。

画像解析の代表的なアルゴリズムを紹介しておく。画像の統計量に関するアルゴリズムには前述の適応的ヒストグラム平坦化が含まれるが、色空間に関するアルゴリズムには、モノクロ画像を適当なカラー画像に変換する疑似カラーが含まれる。フィルタリングには、空間フィルタリングと周波数フィルタリングがあり、空間フィルタを用いることで、平滑化 (ぼかし)、先鋭化、エッジ抽出などの処理が行われる。平滑化を用いれば、画像内に含まれるノイズが除去でき、エッジ抽出では有用な特徴を抽出することができる。空間フィルタリングの非線形フィルタには、有用なエッジをぼかさずに、不要なノイズを除去することができるバイラテラルフィルタ⁽²⁾ が用いられることが多い。幾何学的変換には拡大縮小、回転、鏡映などが含まれる。二値化には、一定の閾値で二値

* 大阪公立大学大学院情報学研究科学際情報学専攻; 准教授 (〒599-8531 堺市中区学園町 1-1)
Advances in Image Analysis for Materials Using Deep Learning; Tokuteru Uesugi (Department of Interdisciplinary Informatics, Graduate School of Informatics, Osaka Metropolitan University, Sakai)
Keywords: *image analysis, deep learning, machine learning, convolutional neural network, surface inspection, fractography*
2024年2月13日受理 [doi:10.2320/materia.63.371]

化する単純二値化を使うこともあるが、大津の二値化⁽³⁾や適応的二値化⁽⁴⁾が用いられる。また、二値化後の画像には膨張・収縮のようなモフォロジー処理を行うことが多い。パターン認識では、特徴抽出後の距離計算や教師あり学習の機械学習が用いられる。画像解析では深層学習以前から機械学習アルゴリズムが広く用いられていた。これらのアルゴリズムはライブラリに実装されており、すべてを最初から実装する必要は少ない。画像解析の標準的ライブラリはオープンソースのOpenCV⁽⁵⁾であり、C/C++やPythonなどで提供されている。

深層学習以前の方法では、画像から特徴量を抽出する変換処理が属人的な技術を必要とし難しかった。膨大なデータから色、輪郭、形状などの特徴を抽出する必要があり、課題に応じた特徴の選定と組み立てには多くの試行錯誤が伴っていた。さらに、複雑な形状やテクスチャの抽出は困難であり、人間のように曖昧な特徴を捉えることが難しいため、適用可能な課題が限られていた。

深層学習が導入される以前から、画像解析は材料分野にも適用されていた。特に、鋼板や鋼管などの表面性状が良好な材料においては、自動検査に画像解析技術が利用されてきた。欠陥判定の精度を高めるためには、ルールの構築が重要であり、特徴抽出後にサポートベクターマシンなどの機械学習が利用されることもあった⁽⁶⁾。

一方、鋳造品の場合、鋼板や鋼管と比較して表面が平滑ではなく、自由曲面を有している。これに加え、鋳造欠陥の種類は多岐にわたり、画像解析を用いた自動検査は困難であるが⁽⁷⁾、人間には鋳造欠陥は判別可能であるため、鋳造品の表面検査は検査員の目視に依存してきた。このような人による画像解析は非常に優れているものの、表面品質の要件を言語化して伝えることは難しい。例えば、日本産業規格では、欠陥の種類別に程度を1級から5級までの5段階で示した標準写真を準備し、肉眼での標準写真との比較による等級分類が定められている⁽⁸⁾。しかし、少数の標準写真からの比較と解釈は検査員の主観に委ねられるため、検査員ごとに等級評価が異なるという問題が生じる。

破面解析の分野でも、人による画像解析が重要な役割を果たしている。破面解析は破面に残された痕跡を調査し、破壊機構や起点を明らかにする手法であり、解析者は破断面の模様を解釈して破面を分類する。この分析には多くの経験が必要で、その経験は言語化できない暗黙知となっており、伝承は容易ではない。なお、破面解析における画像解析の適用も試みられており、画像解析での特徴抽出後にニューラルネットワークなどを用いる機械学習の利用の報告がある⁽⁹⁾。

金属組織の画像解析で一例として挙げられるのが結晶粒径の測定である。従来から、研磨しエッチング処理を施した試験片の顕微鏡画像から手動で切断法(切片法)を用いて結晶粒径が算出されてきた。画像解析では、画像を二値化して結晶粒界と交差する走査線の数をカウントする自動化手法が提案されている⁽¹⁰⁾。この方法では大津の二値化を使用し、二値化画像にモフォロジー処理を施すが、二値化は金属組織の形

状やテクスチャなどの特徴を活用しておらず、二値化された領域が結晶粒界に対応するのか、第二相粒子のようなノイズに対応するのかの分類には限界があった。

3. 深層学習の基本原則

深層学習はニューラルネットワークの層を何層にも深くした学習手法である。ニューラルネットワークは脳の仕組みを模した機械学習手法の一つであり、機械学習の主要な目標は、データから学習して正確な予測を行うモデルを構築することである。重要な点は、モデルが既知のデータの入出力関係を単に再現するだけでなく、未知の入力に対しても正確な出力を提供することである。この未知のデータに対応する能力を汎化と呼び、未知のデータに対する誤差を汎化誤差という。真の汎化誤差を直接知ることはできないため、通常は期待値として扱われる。従って、機械学習の目標は期待値としての汎化誤差の最小化と言える。機械学習についての詳細は本稿では立ち入らないが、材料分野に向けた解説記事も多数存在する⁽¹¹⁾⁽¹²⁾。

ニューラルネットワークは、微分可能な線形変換と非線形変換を組み合わせて作られた計算グラフである。計算グラフは、ノードとエッジによって構成され、ノードでは活性化関数を用いて非線形変換が行われ、エッジでは線形変換が実施される。ノードの集合を層と呼び、これらの層を複数重ねて多層化する。深層学習を含むニューラルネットワークは、全体として微分可能な連続関数であり、誤差逆伝搬法を用いて学習が行われる。学習されるパラメータとは別に、人間が事前に設定する必要のあるパラメータをハイパーパラメータと呼ぶ。ニューラルネットワークのアーキテクチャ(構造)は自由に設計できるため、調整するハイパーパラメータには層の数などが含まれる。ニューラルネットワークの基本は層間のノードが互いに全て結合した全結合層だが、画像処理では特に畳み込み層が一般的であり、畳み込み層を含む深層学習がCNNと呼ばれる。

深層学習以前の特徴抽出アルゴリズムでは人手による試行錯誤が必要だったが、深層学習の畳み込み層では画像の学習に適した特徴を学習し、全結合層で最終的な識別を行う。例えば、図1に示されるVGG16⁽¹³⁾のアーキテクチャでは、13層の畳み込み層と3層の全結合層で構成されている。入力画像に対して畳み込み層の出力結果を特徴マップと呼び、この特徴マップは次の畳み込み層の入力となる。途中でプーリング層を挟み、最終的に特徴マップは全結合層に入力される。

畳み込み層では、画像または特徴マップの入力データが2次元配置されたピクセルの集合として扱われ、隣接するピクセル間の関係を保ちながら線形変換が行われる。例えば、図2では入力データを構成する8×8の2次元ピクセル群から、フィルタと同じ大きさの3×3の小窓を切り出し、フィルタとの内積を計算して一つの値を抽出する。この小窓を画像上で1ピクセルずつスライドさせることにより、6×6の特徴マップが得られる。この操作を畳み込みと呼ぶが、実際



図1 VGG16⁽¹³⁾のアーキテクチャ。(オンラインカラー)

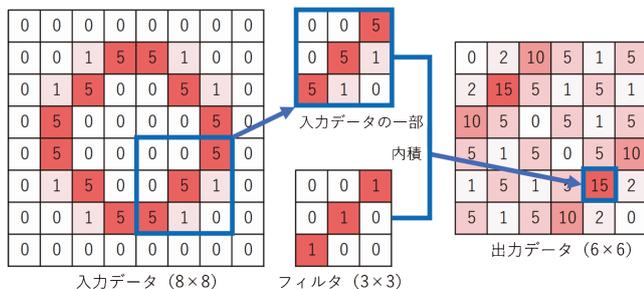


図2 畳み込みの例。(オンラインカラー)

には、画像または特徴マップが深さ(チャンネル)の次元を持つ3次元データであるため、より複雑な操作が必要となる。図2の例では、フィルタが右上から左下への線状を検出するものであり、出力された特徴マップではこの特徴を持つ箇所が強く反応している。畳み込み層で学習されるのはフィルタのパラメータであり、誤差逆伝搬法によってこれらのパラメータが学習される。

深層学習が広く注目されるきっかけとなったのは、2012年のILSVRC(ImageNet Large Scale Visual Recognition Challenge)で優勝した深層学習モデルのAlexNet⁽¹⁴⁾である。ImageNet⁽¹⁵⁾は1400万枚以上のラベル付き画像データベース

表1 代表的なCNNモデル(文献⁽¹⁷⁾より改変引用)。

| モデル | 正解率 | パラメータ数 | 層の深さ |
|-------------------|-------|--------|------|
| VGG16 | 71.3% | 1.38億 | 16 |
| VGG19 | 71.3% | 1.44億 | 19 |
| ResNet50V2 | 76.0% | 0.26億 | 103 |
| ResNet101V2 | 77.2% | 0.45億 | 205 |
| ResNet152V2 | 78.0% | 0.60億 | 307 |
| InceptionV3 | 77.9% | 0.24億 | 189 |
| InceptionResNetV2 | 80.3% | 0.56億 | 449 |
| MobileNetV2 | 71.3% | 0.04億 | 105 |
| Xception | 79.0% | 0.23億 | 81 |
| DenseNet121 | 75.0% | 0.08億 | 242 |
| DenseNet169 | 76.2% | 0.14億 | 338 |
| DenseNet201 | 77.3% | 0.20億 | 402 |

であり、ILSVRCはImageNetを用いて分類精度を競うコンペティションである。AlexNetは現在では特に深層化されたモデルではないが、このコンペティションでは特徴抽出とサポートベクターマシンを組み合わせた従来の画像解析手法の精度を大きく上回った。

AlexNetの登場以降、深層学習モデルの層の深さは飛躍的に増加した。例えば、2014年に発表されたVGG19⁽¹³⁾では19層、2017年にはDenseNet201⁽¹⁶⁾で402層の深層化が行われた。表1では、代表的なCNNモデルのImageNetにおける正解率、パラメータ数、層の深さが示されている⁽¹⁷⁾。層の深さが増加するとパラメータ数はむしろ減少することがある。例えば、VGG19での1億4千万個からDenseNet201では2千万個とパラメータは減少している。CNNを含む深層学習の表現力は、層の深さに対して指数関数的に向上するが、層の幅に対しては多項式的に増加するため、ImageNetの分類精度では、パラメータ数よりも深層化を優先したモデルの方が高い正解率を示す傾向にあるためである。ただし、深層化には誤差逆伝搬法の学習が進まないという問題があり、様々な技術的な工夫が施されている。VGG16⁽¹³⁾やVGG19⁽¹³⁾は16層または19層の比較的浅いモデルだが、ResNet50V2⁽¹⁸⁾、ResNet101V2⁽¹⁸⁾、ResNet152V2⁽¹⁸⁾は残差接続を用いて深い畳み込み層での学習を可能にした。InceptionV3⁽¹⁹⁾はInceptionモジュールごとにネットワークを分岐させてサイズの異なる畳み込みを行い、InceptionResNetV2⁽²⁰⁾ではInceptionモジュールに残差接続を用いて深層化している。MobileNetV2⁽²¹⁾はモバイル端末での使用に適したパラメータ数の削減を行っている。Xception⁽²²⁾はInceptionの改良版で、精度向上とパラメータ数削減を実現している。DenseNet121⁽¹⁶⁾、DenseNet169⁽¹⁶⁾、DenseNet201⁽¹⁶⁾はResNetの改良版で、ショートカット接続を導入することで深層化を可能にしている。

ただし、深層化されたモデルが必ずしも高い汎化性能を持つわけではない。これは機械学習におけるノーフリーランチ定理⁽²³⁾として知られている原則である。ノーフリーランチ定理によれば、どの手法も他の手法に対して常に優れている

わけではないため、課題やデータに応じて適切なアルゴリズムを試行錯誤して選択する必要がある。

表1に示されたモデルは全て画像分類のモデルだが、出力層と損失関数をタスクに合わせて設計することで、分類以外の回帰、検出、セマンティックセグメンテーションも実現可能である。分類ではクラス確率を出力するが、回帰では出力層を実数値に変更し、検出タスクでは入力画像を均等に区切ったグリッドごとにクラス確率と検出領域を出力するよう設計する。セマンティックセグメンテーションでは、ピクセルごとにクラス確率を出力する出力層が必要となる。畳み込み層は共通のモジュールとして複数のタスクに利用できるため、ImageNetの分類タスクで学習済みの表1のモデルの畳み込み層部分を取り出し、他のタスクに適用することもできる。これを転移学習と言い、データ量が少ないタスクを行う時には有効な技術である。

セマンティックセグメンテーションでは、ピクセルごとにクラス確率を出力するために、全結合層を廃止し、畳み込み層のみを用いる。代表的なモデルであるU-Net⁽²⁴⁾は、エンコーダとデコーダを繋げたUの形をしている。エンコーダは画像の特徴を抽出する役割を持ち、画像分類モデルの畳み込み層をバックボーンとしてほぼそのまま活用する。U-Netの発展形として、U-Net++⁽²⁵⁾やU-Net3+⁽²⁶⁾があり、これらのモデルではエンコーダとデコーダの繋がりがより密接になっている。さらにR2U-Net⁽²⁷⁾では、各階層に再帰構造が導入されている。

深層学習分野で広く使用されるライブラリにはTensorflow/Keras⁽²⁸⁾とPyTorch⁽²⁹⁾がある。どちらもPython言語で提供されるオープンソースのライブラリである。これらのライブラリには表1に示された学習済みモデルの大部分が利用可能である。また、GUI(Graphical User Interface)を備えた深層学習ソフトウェアも開発されているが、あらかじめ用意されたアルゴリズムのみを使用するため、精度向上に重要な前処理、転移学習、データ拡張などの拡張性に欠ける。

深層学習による画像解析での最新の進展は、CNNでなくVision Transformer(ViT)⁽³⁰⁾である。ViTは、自然言語処理で使用されているTransformerモデル⁽³¹⁾を、画像認識に応用したものである。ただし、データセットの規模が小さいと、TransformerはCNNに比べて汎化性能が低下する傾向がある。また、CNNは画像の局所的な特徴、特に物体のテクスチャを捉えていることが明らかになっているが、Transformerは画像の大域的な特徴、具体的には物体の形状を捉える⁽³²⁾。材料分野における画像解析では、データセットが小さく、微細組織や表面の複雑なテクスチャが重要な役割を果たすので、TransformerよりCNNが有利である。

4. 深層学習の応用と課題

材料分野での深層学習への適用は着実に進んでいる。鋼板の検査では、最近ではCNNを用いることで欠陥の特徴抽出を自動化し、精度を向上させることが報告されている⁽³³⁾。

表面性状が複雑で自由曲面を有する鋳造品の鋳肌検査にもCNNの有効性が示されている⁽³⁴⁾。さらに、鋳造品では良品と不良品の区別が曖昧な鋳肌検査においてもCNNの有効性が報告されている⁽³⁵⁾。

破面解析においても、最近ではCNNを用いることで破断面のより精度の高い分類が可能になっている⁽³⁶⁾。また、破面解析ではCNNを用いた破面領域のセマンティックセグメンテーションも報告されている⁽³⁷⁾。セマンティックセグメンテーションでの深層学習の利用は特に金属組織の画像解析で進んでいる。鉄鋼材料の様々な組織がセマンティックセグメンテーションで領域分割することが可能なことが報告されている⁽³⁸⁾、結晶粒界についても金属組織画像のセグメンテーションから切断法を用いて結晶粒径を自動で評価するシステムが報告されている⁽³⁹⁾。

材料分野での深層学習の適用が進展する中、深層学習に関する基本的な理解において、作り手側と受け手側の間にしばしば齟齬が見られる。特に、深層学習がどのような基準で解析結果を出すのかが理解しづらいというブラックボックス問題において顕著である。作り手側にとって、深層学習のアーキテクチャとパラメータは明確であり、これらは可視化可能であるため、ブラックボックスとは言えない。しかし、VGG19のように1億4千万個ものパラメータがある場合、これらを個別に確認することは人間にとって不可能である。一方で、深層学習を使用する者にとっては、深層学習はブラックボックスとなり、深層学習によって獲得された機能を人間が理解可能な形で説明を求めるのは自然なことである。

ブラックボックス問題は、人間にも適用可能な表現方法の問題である。例えば、画像が犬か猫かを答える問題では、多くの人は答えられるが、犬と猫を知らない人にそれを説明するのは非常に難しい。犬と猫の違いは存在するが、単純な特徴を特定するのは困難である。同様に、人間による画像解析でも、鋳肌検査のように、判断基準が曖昧で伝達が困難な場合には、標準写真との比較により等級分類を行って、深層学習の作り手側も、このような齟齬を埋める努力を進めており、説明可能な人工知能技術(Explainable AI: XAI)が期待されている。しかし、この技術は局所的な解釈しか提供せず、深層学習の内部を人間が納得する形で説明するのは依然として難しいという現実がある。

人工知能が人間の認識精度を超えたという説も、作り手と受け手の間で齟齬を生んでいる。この論点で引き合いに出されるのが、2015年のILSVRCで優勝したResNetが平均的な人間の正解率を超えたという事例である⁽⁴⁰⁾。しかし、この事例には齟齬の元となる要素がある。深層学習は大量の教師データを必要とし、その教師データは常に正しいことを前提としている。そのため、専門家によって作成された教師データは正解率100%とされており、深層学習が教師役の人間を超えることは不可能である。ただし、平均的な人間は専門家と同じくResNetに及ばない。ImageNetには、多くの珍しい犬の品種が含まれており、平均的な人間が高い正解率を得ることは難しい。この場合、人間の正解率の差は、これま

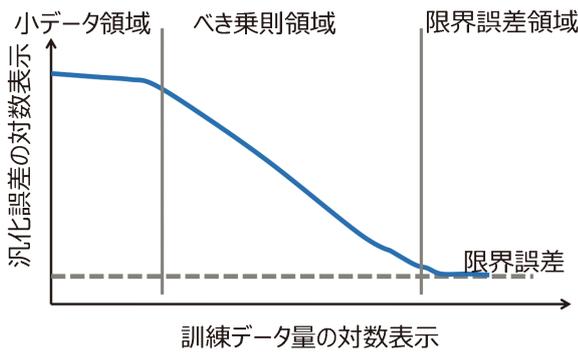


図3 データ量と汎化誤差の経験的な関係(文献⁽⁴¹⁾より改変引用)。(オンラインカラー)

で犬をどれだけ観察してきたかという学習量によって左右される。

深層学習の正解率も、人間と同じく学習量によって大きく変わる。図3で示されるデータ量と汎化誤差の関係によれば、データ量が増加すると、画像解析や自然言語処理などのタスクに関係なく、べき乗則と呼ばれる領域が観察される⁽⁴¹⁾。データ量と汎化誤差のべき指数は、タスクの難しさの指標として使用できる。したがって、ある程度のデータ量を用意し、べき指数を算出した後、目標とする汎化性能に必要なデータ量を見積もることが可能である。べき乗則の領域を超えると、モデルが改善できない限界誤差が存在することが予測される。この限界誤差の原因は、教師データでの誤ったラベル付けであり、これによって不可避な誤差が発生する。実際に、ImageNetのテストデータを調査した研究では、画像の5.83%にラベルの間違いが報告されている⁽⁴²⁾。

まとめると、人間の認識精度を超える深層学習モデルの構築は可能であるが、これは学習に使用したデータの分布範囲内に限られる。そのため、適切な正解データを十分に確保することが必須である。そして、目標の汎化性能に必要なデータ量はべき乗則から経験的に算出することができる。しかし、材料分野での実用的な画像解析において、十分なデータが利用可能な状況は稀である。限られたデータ量で高い汎化性能を達成するためには、多様な機械学習技術の活用が求められる。これらの機械学習技術には、前処理、データ拡張、転移学習が含まれる。

画像解析において、特徴量の抽出を効率化する前処理の役割は深層学習の導入前後で変わらず重要である。主な画像データの前処理手法には、画像リサイズ、グレースケール変換、ぼかし、ヒストグラム平坦化などが含まれる。画像リサイズでは画像サイズを一定の解像度に揃え、機械学習モデルが一定の入力サイズを扱えるようにする。グレースケール変換は、カラー画像をグレースケールに変換し、色情報を省略して輝度情報に注目し、計算量を削減する。逆に、グレースケールを疑似カラーに変換することで、色情報を追加することも効果的である。材料分野で扱われる画像、特に顕微鏡画像などでは、色情報が実質的に不要な場合が多く、この場合

では、グレースケール変換と疑似カラー変換が有効に機能する⁽³⁵⁾。ぼかし処理は、画像の細かなテクスチャやノイズを平滑化し、ヒストグラム平坦化は画像のコントラストを改善し、いずれも特徴量抽出を容易にすることがある。これらの前処理は、具体的なタスクやデータセットに関する知識を用いて、試行錯誤を重ねながら適用される。

データ拡張は、既存データセットに多様な操作を施し新しい訓練データを生み出す手法であり、これにより、画像データに変更を加えるため、一種の前処理とも考えられる。具体的なデータ拡張手法には、回転、平行移動、拡大・縮小、反転、明るさや色調の調整などがあり、これらはタスクやデータセットの性質に応じて選択される。例えば、対象の位置や向きが判断基準に影響しない場合では、画像の回転や平行移動を適用することが適切である。

転移学習は、異なるタスクで獲得された知識を新しいタスクへ適用し、効率的な学習を目指す手法である。転移学習では、元となるドメインから得たデータの事例、特徴、パラメータを新たな目標ドメインへ適用させる。特に画像解析分野では、大規模データセットで訓練されたモデルを基にして、新しいデータセットへのファインチューニングを行うパラメータ転移が一般的である。パラメータ転移では、事前に学習済みの畳み込み層のパラメータをそのまま転移学習として使用するか、もしくは一部の層を目標ドメインのデータに合わせてファインチューニングする。具体的には、ImageNetで事前に訓練された表1で示したようなモデルを利用し、目標ドメインのタスクに適応させることが行われる。

紹介した前処理、データ拡張、転移学習のファインチューニングは、深層学習における画像解析の一般的なアプローチであるが、これらの手法だけでは十分な汎化性能を確保できない場合もある。次節からは、著者が関わった事例を紹介する⁽³⁵⁾⁽³⁶⁾。紹介事例では、これらのアプローチを採用しているが、タスクやデータセットの特性に応じて、新たなアプローチの開発もしている。

5. 回帰と不良品分類を両立する鋳肌検査

ダクタイル鋳鉄管は水道やガス管など広範囲に使用され、鋳型遠心鋳造法で製造されている。この方法は鋳造欠陥を顕著に減少させるが、ピンホールは依然として重要な課題である。鋳肌検査において、どこまでをピンホールとするかの境界が明確でないため、検査員間で品質基準を一貫して伝達する方法が存在しない。その解決策として、日本産業規格では鋳造品の欠陥等級を五段階で示す標準写真を用いた比較法が採用されているが、この方法は定量性に欠け、検査員の主観に大きく依存するため、標準性が不足している。CNNの導入は、これらの問題を解決することが期待される。

最近、CNNは鋳肌検査でも有効性を示しているが、これまでの研究ではブローホールや割れなど明確に識別可能な欠陥に焦点を当てており、ピンホールのように不明瞭な欠陥の識別には対応していない。そこで、この研究では、ダクタイル

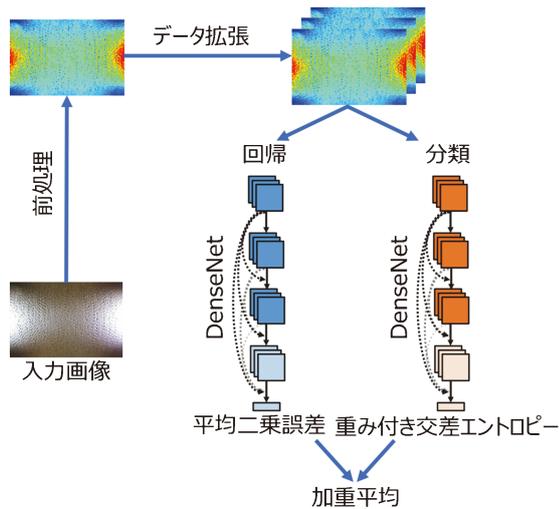


図4 提案するアンサンブル学習の概略(文献⁽³⁵⁾より改変引用)。(オンラインカラー)

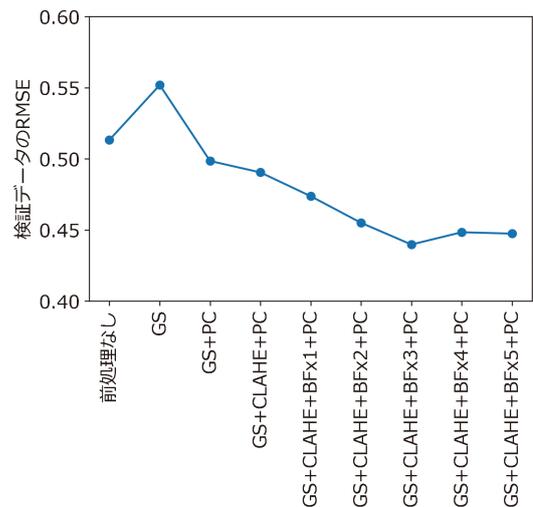


図5 前処理の検証結果(GS: グレースケール, CLAHE: 適応的ヒストグラム平坦化, BF: エッジ保存平滑化, PC: 疑似カラー, 文献⁽³⁵⁾より改変引用)。(オンラインカラー)

ル鉄管の錆検査を自動化するためのCNNベースの手法を開発し、連続値での等級判定を可能にすることで、より高度な品質管理を実現することに貢献する。回帰と不良品分類を組み合わせた新しい錆検査法を提案するだけでなく、前処理とデータ拡張において錆特有の知識を活用し、錆検査における定量性、標準性の問題を解決した⁽³⁵⁾。

提案手法では、錆等級の回帰と不良品の分類を同時に行うため、回帰と分類のアンサンブル学習法を採用した。図4に提案するアンサンブル学習の概略を示す⁽³⁵⁾。初めに錆画像を前処理して意味ある特徴を抽出し、その後、データ拡張を施して画像から汎化性能を高める。これらの画像は、CNNを用いた回帰と分類の弱学習器に入力される。ここで、ImageNetで事前に学習されたDenseNetなどを転移学習またはファインチューニングによって使用する。回帰学習器には平均二乗誤差が、分類学習器には不良品の識別を強化する目的で重み付き交差エントロピーが使用される。最後に、回帰と分類の結果を加重平均して、最終的なアンサンブル学習の出力を得る。この方法により、不良品の見逃しのない確実な分類と回帰を同時に達成することが目指される。

実製造されたダクタイル鉄管の錆画像をデータセットとして使用した。画像は一定条件下で撮影された3,481枚であり、同一の検査員によって5段階の等級に分類された。ピンホールの存在は等級によって異なり、等級1には存在せず、等級2はわずか、等級3は中程度、等級4は不良品ではないが多く、等級5は多すぎて不良品とされた。データセットは訓練、検証、テストデータに分割され、検証データでの精度評価を基に調整を行い、精度評価はテストデータを用いて行われた。クラス不均衡を解消するため、訓練データにはオーバーサンプリングが施された。精度評価では、回帰には二乗平均平方根誤差(Root Mean Squared Error: RMSE)、分類には不良品の再現率が用いられ、回帰モデルでの調整後のCNNを分類モデルにも適用した。

前処理では、画像解像度の調整が重要であった。元の画像サイズは 2736×1824 であり、複数の解像度候補から検証データを用いて最適な解像度 798×532 を選定した。リサイズ後にグレースケール変換、適応的ヒストグラム平坦化、エッジ保存平滑化、疑似カラー化を前処理として検証した結果を図5に示す⁽³⁵⁾。グレースケール後のRMSEの悪化に対し、疑似カラー化で改善された。適応的ヒストグラム平坦化とエッジ保存平滑化は錆特有の縞模様のノイズ除去に有効で、RMSE向上に貢献した。最終的に、グレースケール、適応的ヒストグラム平坦化、エッジ保存平滑化3回、疑似カラーの組み合わせが最も効果的であることが確認され、この前処理が選択された。

データ拡張の効果は訓練データに対して移動、縮小・拡大、回転、せん断、反転の操作を行い、検証データの精度評価を用いて効果を検証した。図6に検証結果を示すが、最もRMSEが優れていたデータ拡張の条件は、水平および垂直方向の反転に加えて、水平・垂直方向に0%~2%の範囲でランダムに移動、-2%~+2%の範囲でランダムに縮小・拡大、-2°~+2°の範囲でランダムに回転、0°~2°の範囲でランダムにせん断であり、この条件を最適なデータ拡張条件として採用した⁽³⁵⁾。図6には最適条件から移動、縮小・拡大、回転、せん断をそれぞれ、独立に変化させた時のRMSEも示されており、最適なデータ拡張条件を外れると、反転のみを用いた場合のRMSEまで悪化することがあった。データ拡張なし、反転のみ、最適なデータ拡張条件でのRMSEを比較すると、データ拡張なしに比べて反転のみでもRMSEは大きく向上していたが、反転以外のデータ拡張は視野の変更を伴うため、反転以外のデータ拡張による効果は限定的であった。

前処理とデータ拡張後、図7に示すCNNの比較では、

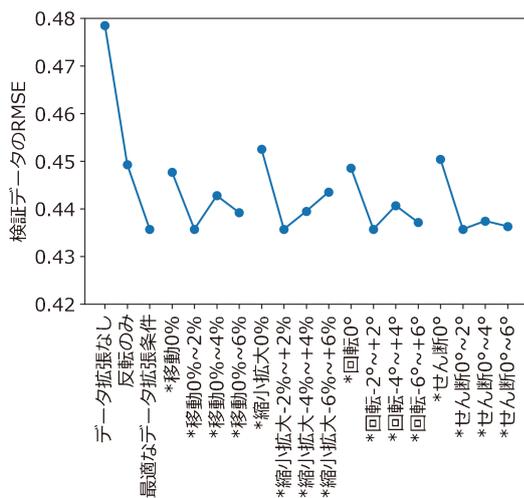


図6 データ拡張の検証結果(文献⁽³⁵⁾より改変引用)．アスタリスク(*)は最適なデータ拡張条件から部分的に変更された条件を示す．(オンラインカラー)

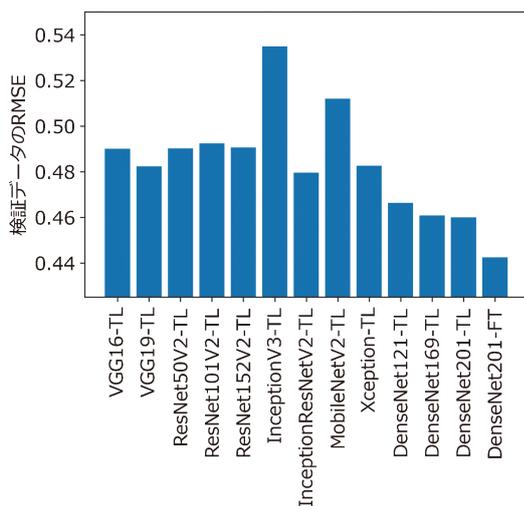


図7 CNNの検証結果(TL: 転移学習, FT: ファインチューニング, 文献⁽³⁵⁾より改変引用)．(オンラインカラー)

ImageNetで事前学習済みのDenseNet201が転移学習で最も優れたRMSEを示した⁽³⁵⁾．深層学習の効果は層の深さに指数的に依存し、DenseNet201の深い構造が良好な結果をもたらした．一部パラメータの再学習を含むファインチューニングは、DenseNet201においてRMSEがさらに改善した．そのため、分類とのアンサンブル学習ではDenseNet201のファインチューニングを使用した．

回帰モデルの損失関数を変更して分類モデルの学習を行い、回帰と分類のアンサンブル学習を実施した．不良品クラスの重みと加重平均の重みは、検証データに基づく再現率から決定された．不良品クラスの重みを上げることで、良品が不良品と分類されやすくなり、再現率は100%に達したが、正解率は低下した．アンサンブル学習では、検証データで再現率100%を達成するために、加重平均の重みを設定した．

テストデータでの推論結果は検証データと同等の汎化性能を示し、不良品分類の再現率は100%を維持した．この結果より、連続値での等級判定と不良品の見逃しのない分類を両立する鑄肌検査法の開発が達成された．今後の課題は、開発した深層学習モデルを鑄物工場に実装する自動システムとして取り入れることである．

6. 材料の種類を考慮した破面分類

金属製品の破損原因を明らかにするためには、破断面の破面解析が不可欠である．破面解析は、製品の品質評価およびその維持・向上を目的とした重要なプロセスである．このプロセスでは、解析者が破断面の模様を精査し、破壊機構に基づいて分類を行う．この作業は、専門家による観察と経験に依存した知的判断を必要とする．解析者の経験は、容易に言語化できない暗黙知として蓄積されるため、その知識の伝承は困難である．さらに、熟練した解析者が徐々に退職していく現状がある．これらの背景から、破面解析の初学者を支援し、熟練解析者の経験を引き継ぐシステムの開発が求められる．

既にCNNを用いた破面分類は報告されているが、これまでの研究では合金鋼とステンレス鋼の違いなど、材料の種類の違いを考慮せずに、全ての材料を同一の破断面クラスで分類していた．熟練した解析者は、材料の違いを暗黙知として考慮しており、この研究では、材料の種類を考慮した転移学習を用いて、CNNによる破面分類の精度を向上させることを目指した．そのため、異なる材料からの学習結果を効果的に利用する新たな転移学習手法として、ステップワイズ元クラス選択法(Stepwise Selection of Source Classes: SSSC)を提案した⁽³⁶⁾．

既存の破断面分類手法では、材料の種類を区別せずに、元ドメインのデータのみを用いた学習が行われている．ディンプル、ストライエーション、ぜい性疲労破面、へき開破面、擬へき開破面、粒界破面、ファンシェーブトパターン⁽³⁶⁾の7クラス分類において、元ドメインのみを用いた分類モデルは図8(a)に示され、材料の種類ごとに区別して学習を行った場合は、目標ドメインのみを用いた学習となり、図8(b)に示される⁽³⁶⁾．

元ドメインでの分類、目標ドメインでの分類、元クラス選択による分類の概念図を図9に示している⁽³⁶⁾．目標ドメインの分類では、事例が存在しないクラスがあるため、識別境界が不足している．しかし、元クラス選択による分類では、目標ドメインから有用な元クラスを選択することで、適切な識別境界を設定できる．クラスごとに転移させる事例の選択は、目標クラスと元クラスの組み合わせ最適化問題として扱うことができるので、提案手法であるSSSCでは、元ドメインと目標ドメインのデータを組み合わせた事例を用いて複数のモデルを学習し、目標ドメインのデータでの評価を行い、ステップワイズ法を用いて、最適な元ドメインと目標ドメインの組み合わせを探索する．

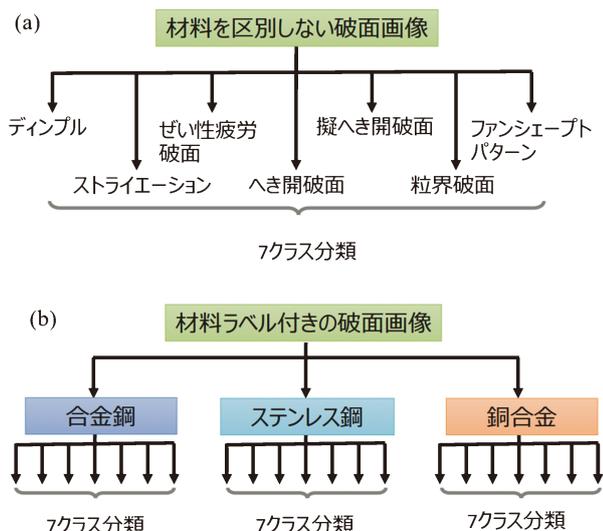


図8 (a) 元ドメインでの学習と (b) 目標ドメインでの学習 (文献⁽³⁶⁾より改変引用)。(オンラインカラー)

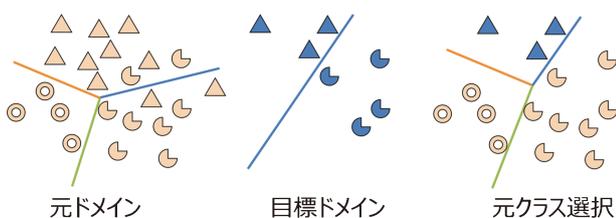


図9 元ドメインでの分類, 目標ドメインでの分類, 元クラス選択による分類(文献⁽³⁶⁾より改変引用). 図は特徴空間における識別境界と事例を表しており, 識別境界はクラス間を分離する境界であり, 事例のマーカー形状はクラスの違いを示し, 元ドメインの事例は赤色で, 目標ドメインの事例は青色で表されている。(オンラインカラー)

データセットには, 破面解析において経験豊かな専門解析者によってラベル付けされた走査型電子顕微鏡による破断面画像1033枚が含まれている. このデータセットはドメイン別, つまり材料別に分けられ, 合金鋼からは670枚, ステンレス鋼からは230枚, 銅合金からは133枚の画像が含まれている. クラス別, つまり破面形態の違いに基づく分類では, ディンプル形態の画像が231枚, ストライエーションが212枚, ぜい性疲労破面が164枚, へき開破面が127枚, 擬へき開破面が105枚, 粒界破面が161枚, ファンシエプトパターンが33枚含まれていた.

使用したデータ量は十分ではなく, また不均衡なデータであったため, オーバーサンプリングによって少数クラスの画像を複製し, データ数を増やしてクラスの不均衡を解消した. 検証データとテストデータの事例数は, それぞれのクラスについて5に設定した. 前処理としては, 画像のリサイズを行った. 図10には検証データの正解率と画像サイズ(解像度)の関係を示し, 画像サイズは360×270が最適であることが確認された⁽³⁶⁾. 早期打ち切り, ドロップアウト, バッ

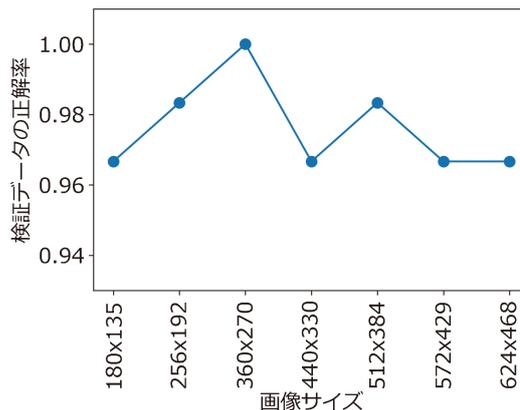


図10 正解率と画像サイズの関係(文献⁽³⁶⁾より改変引用)。(オンラインカラー)

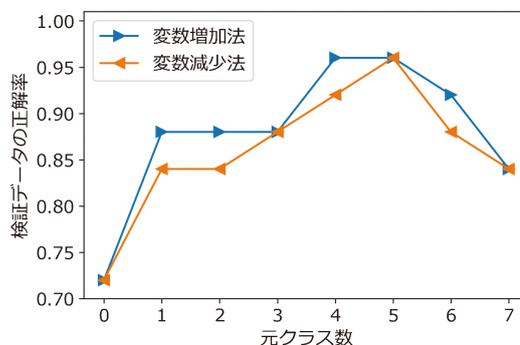


図11 ステンレス鋼での正解率と元クラス数の関係(文献⁽³⁶⁾より改変引用)。(オンラインカラー)

チ正則化といった正則化手法を用い, 活性化関数の種類やノード数などのハイパーパラメータのチューニングについては検証データを使用して行った. データ拡張は訓練データに対して施し, 画像にランダムに移動, 拡大・縮小, 回転, せん断, 反転の操作を加えた. CNNのアーキテクチャ選定では, VGG16, VGG19, ResNet50V2, ResNet101V2, ResNet152V2, InceptionV3, InceptionResNetV2, MobileNetV2, Xception, DenseNet121, DenseNet169, DenseNet201の中から, 検証データでの結果を基にVGG16を用いたファインチューニングを採用した.

前処理を施したデータとチューニングされたVGG16モデルを用いて, ステップワイズ元クラス選択法による元クラス(元ドメインのクラス)の選択を実施した. 図11には, ステンレス鋼における正解率と元クラス数の関係を, 変数増加法と変数減少法で示している⁽³⁶⁾. 元クラス数が0の場合と7の場合は, それぞれ目標ドメインのみでの学習と元ドメインのみでの学習を意味する. 元クラス数が4または5の場合に最高の正解率が観測された. 元クラスと目標クラスの選択には, 両者のデータ数の比率が影響していることが示されている. 目標クラスの訓練データが十分に多い場合は目標クラスが選択される傾向にあるが, 目標クラスの訓練データが少ない場合は元クラスが選択される傾向にあった.

表2 テストデータの正解率(文献⁽³⁶⁾より改変引用).

| ドメイン | 目標ドメインでの学習 | 元ドメインでの学習 | 提案手法SSSC |
|--------|------------|-----------|----------|
| 合金鋼 | 93.3% | 93.3% | 100.0% |
| ステンレス鋼 | 72.0% | 84.0% | 96.0% |
| 銅合金 | 79.2% | 100.0% | 100.0% |
| 平均 | 81.5% | 92.4% | 98.7% |

表2は、目標ドメインでの学習、元ドメインでの学習、そしてステップワイズ元クラス選択法を使用した場合のテストデータの正解率を比較している⁽³⁶⁾。元ドメインでの学習は材料の違いを考慮しない従来の手法で、合金鋼、ステンレス鋼、銅合金における平均正解率が92.4%である。目標ドメインのみでの学習では、材料の種類を明確に区分して直接学習を施した結果、平均正解率が81.5%に下降し、元ドメインのみでの学習結果を下回っている。しかし、ステップワイズ元クラス選択法を適用した場合の平均正解率は98.7%に達し、従来手法に比べて6.3%の改善が見られた。これにより、材料の種類を考慮する転移学習手法であるステップワイズ元クラス選択法が、破面分類の精度を著しく向上させることが示された。この研究で開発されたモデルは、すでに解析者の訓練ツールとして試験的に使用されおり、今後はこのモデルを公開し、広く利用できるようにすることを進めている。

7. おわりに

画像解析は、画像から有益な情報を得るための手段の一つに過ぎず、その結果は視覚化されて誰もが直感的に理解しやすい。深層学習を用いた画像解析技術の進歩により、GUIを搭載した深層学習ソフトウェアの開発も進み、誰もが深層学習を利用できる時代が来ることが期待される。しかし、技術の利用者と開発者の間には深層学習導入以前と同じ隔たりが存在する。特に、深層学習技術の普及において大量の画像データの準備が大きな課題であり、材料分野でのドメイン知識を活かして、限られたデータで優れたモデルを構築することが重要である。本稿では、データが少ない状況下での成果事例を示し、技術の実用化に向けた課題と展望を紹介した。材料分野における深層学習の利用は自動化への顕著な貢献が期待されるが、適切な技術選択とデータの重要性への理解が必須である。深層学習技術が、材料分野における解析手段の新たな柱組みとして確立されることを期待しつつ、本稿がその一助となることを願っている。

本稿で紹介した研究の一部は、文部科学省データ創出・活用型マテリアル研究開発プロジェクト事業JPMXP1122684766の助成を受けたものである。研究の遂行にあたり、(株)栗本鐵工所の小川耕平氏、堤親平氏、澤田健二氏、中本光二氏、(地独)大阪産業技術研究所の濱田真行氏、喜多俊輔氏、平田智丈氏から学習データの提供を受けた。鋳

肌検査のモデル構築は大阪公立大学教授の中島智晴氏との共同研究である。ここに謝意を表す。

文 献

- (1) S. M. Pizer, E. P. Amburn, J. D. Austin, R. Cromartie, A. Geselowitz, T. Greer, B. ter Haar Romeny, J. B. Zimmerman and K. Zuiderveld: *Comput. Vis., Graphics Image Process.*, **39** (1987), 355–368.
- (2) C. Tomasi and R. Manduchi: *IEEE Int. Conf. Computer Vision*, (1998), 839–846.
- (3) N. Otsu: *IEEE Trans. Sys. Man. Cyber.*, **9**(1979), 62–66.
- (4) M. Sezgin and B. Sankur: *J. Electron. Imaging*, **13**(2004), 146–168.
- (5) I. Culjak, D. Abram, T. Pribanic, H. Dzapo and M. Cifrek: *35th Int. Convent. Informat. Commun. Technol., Electron. Microelectron.*, (2012), 1725–1730.
- (6) 大重貴彦：ふえらむ, **23**(2018), 164–168.
- (7) 染次孝博, 吉村剛治, 秋田伸幸：精密工学会誌, **63**(1997), 1412–1416.
- (8) The Japanese Industrial Standards, “Visual examination and classification of surface quality for steel castings”, JIS G 0588 (1995).
- (9) M. X. Bastidas-Rodriguez, F. A. Prieto-Ortiz and E. Espejo: *Eng. Fail. Anal.*, **59**(2016), 237–252.
- (10) H. Peregrina-Barreto, I. R. Terol-Villalobos, J. J. Rangel-Magdaleno, A. M. Herrera-Navarro, L. A. Morales-Hernández and F. Manríquez-Guerrero: *Measurement*, **46**(2013), 249–258.
- (11) 上杉徳照：軽金属, **73**(2023), 104–111.
- (12) 上杉徳照：山陽特殊製鋼技報, **30**(2023), 3–20.
- (13) K. Simonyan and A. Zisserman: *The 3rd Int. Conf. Learn. Represent.*, (2015).
- (14) A. Krizhevsky, I. Sutskever and G. E. Hinton: *Adv. Neural Inform. Process. Sys.*, **2**(2012), 1097–1105.
- (15) O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, A. C. Berg and L. Fei-Fei: *Int. J. Comput. Vis.*, **115**(2015), 211–252.
- (16) G. Huang, Z. Liu, L. Van Der Maaten and K. Q. Weinberger: *30th IEEE Conf. Comput. Vis. Pattern Recognition*, (2017), 2261–2269.
- (17) S. Bianco, R. Cadène, L. Celona and P. Napolitano: *IEEE Access*, **6**(2018), 64270–64277.
- (18) K. He, X. Zhang, S. Ren and J. Sun: *Lect. Notes Comput. Sci.*, **9908**(2016), 630–645.
- (19) C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens and Z. Wojna: *IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognition*, (2016), 2818–2826.
- (20) C. Szegedy, S. Ioffe, V. Vanhoucke and A. A. Alemi: *31st AAAI Conf. Artif. Intellig.*, (2017), 4278–4284.
- (21) M. Sandler, A. Howard, M. Zhu, A. Zhmoginov and L. C. Chen: *IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognition*, (2018), 4510–4520.
- (22) F. Chollet: *30th IEEE Conf. Comput. Vis. Pattern Recognition*, (2017), 1800–1807.
- (23) D. H. Wolpert: *Neural Comp.*, **8**(1996), 1341–1390.
- (24) O. Ronneberger, P. Fischer and T. Brox: *Lect. Notes Comput. Sci.*, **9351**(2015), 234–241.
- (25) Z. Zhou, M. M. Rahman Siddiquee, N. Tajbakhsh and J. Liang: *Lect. Notes Comput. Sci.*, **11045**(2018), 3–11.
- (26) H. Huang, L. Lin, R. Tong, H. Hu, Q. Zhang, Y. Iwamoto, X. Han, Y. W. Chen and J. Wu: *IEEE Int. Conf. Acoustics, Speech Sig. Process.*, (2020), 1055–1059.
- (27) M. Z. Alom, C. Yakopcic, M. Hasan, T. M. Taha and V. K. Asari: *J. Med. Imaging*, **6**(2019), 014006.
- (28) M. Abadi *et al.*: *12th USENIX Symposium on Operating Systems Design and Implementation*, (2016), 265–283.

