

材料破断面の数値解析手法の開発

山際 謙太*1, 本田 尚*1, 佐々木 哲也*1

機械構造物の破壊に起因した事故の調査において、破断した面(破断面)の調査は原因推定に役立つ。一方で、破断面の解析には経験などが必要であるが、近年は熟練した解析者が減少しており、技術力の低下および知識伝承が問題視されている。こうした背景の中、本研究では1) 金属破断面データベースの構築と、2) ディープラーニングの活用による破断面の自動分類という2本の内容で研究を実施した。金属破断面データベースでは、破断面の写真などを収集し、観察者がコメントなどを書き込むことで、破断面の取扱及び注目すべき様相など結果を導くために必要な知識の保存及び伝承を行う環境を構築した。ディープラーニングによる破断面の自動分類では、多層型ニューラルネットワークを活用し、破断面から破壊機構を推定するための数値解析手法を開発した。こうした知識伝承を行う環境及び数値解析手法の開発を行うことで、熟練者の減少などの問題を解決していく基盤的環境の構築を実施した。

キーワード: 破断面解析, 破壊力学, 画像処理, ディープラーニング, データベース

1. はじめに

機械構造物の破損を原因とした労働災害が発生した際には、事故調査の一環として破断面の解析が行われる。この破断面の解析を“フラクトグラフィ”と呼ぶ。破断面には破壊した時の痕跡が残っていることから、フラクトグラフィは事故の原因究明に欠かざる手法である。

フラクトグラフィにより明らかにしておくことは、破壊機構、破壊の起点および最終破断部である。破壊機構を推定していくためには、走査型電子顕微鏡(SEM)などを用いてストライエーション(縞状模様)など破壊機構に対応した特徴的な模様を探していく必要がある。しかし、例えば疲労試験機を使って作成した破断面には幅の整ったストライエーションが観察されるが、実機に生じた破断面の場合は、錆または圧縮荷重の作用などによる破断面の潰れなどがあるため、破壊機構の推定が困難である場合もある。こうした模様の解釈は観察者の経験に依存してくる側面があり、結果的に客観性を欠いた結論を導くこととなる。もしくは観察者の主観により誤った結論が導かれた場合、その結果から導かれる安全対策も誤ったものとなってしまう。また、熟練した解析者は年々リタイアしており、熟練解析者が持つ知見および解析技術の伝承も問題となっている。このような状況において、解析技術を伝承する、そして解析の初心者にも数値解析などを援用してサポートする仕組みが求められている。

こうした問題点を解決するため、本研究は以下の2方針に基づいて実施された。

I. 金属破断面データベースの構築

II. ディープラーニングを活用した破断面の自動分類

Iは、破断面の解析したデータ集の構築である。実験および実機の破断面のデータを収集し、これらをインターネットで閲覧できる仕組みを構築した。この仕組みは、例えば解析初心者が実験により得られる破断面の種類を学習するなど活用することができる。また、熟練解析者が持つ知見などを保存することができ、解析技術の伝承に貢献することができる。

IIは、破断面を数値解析により自動的に破壊機構ごとに分類する手法の確立である。特に近年はディープラーニングと呼ばれる、例えば多層型ニューラルネットワークを活用した画像認識の分野が盛んに研究されている。この技術を破断面の分野に応用することで破断面のSEM画像から破壊機構を推定する手法について研究を行った。

2. 金属破断面データベースの構築

1) 金属破断面データベース構築の目的

フラクトグラフィにより破壊機構を推定するためには、

- ① 観察主体による定性的な手法
- ② 模様の幅、形状および高さなど数値解析する定量的な手法

の2種類の手法がある。金属破断面データベース(以下、データベース)は、特に①に関連した内容である。

観察主体による破壊機構推定というのは、これまでは過去に観察した破断面の例と比較することが行われてきた。例えばある実機の破壊機構が不明な破断面に対して破壊機構を推定するためには、破壊試験を行い予め破壊機構が同定されている破断面と比較して、似た模様であればその破壊機構が実機の破壊機構と同じであるという方式である。

しかし、実機の破断面と試験機により作成した破断面の違いであるが、実機の破断面の場合は錆、つぶれなど

*1 労働安全衛生総合研究所 機械システム安全研究グループ

連絡先: 〒204-0024 東京都清瀬市梅園 1-4-6

独立行政法人労働者健康安全機構労働安全衛生総合研究所

山際謙太

E-mail: yamagiwa@s.jniosh.johas.go.jp

により破断面の特徴を捉えることが難しいことがある。図 1 はその例である。上側の実機の破断面はストライエーションと呼ばれる縞状の様子が潰れて明瞭さを欠いているが、下側の試験機で作成した破断面はストライエーションが鮮明である。こうした場合に、解析の初心者が実機の破断面を観察してストライエーションであるということを判断することは困難である。

これまではこうした場合であっても、破断面の画像集 1), 2), 3), 4) を調べることでより考察することが一般的であった。しかし、掲載されている材料および破壊機構にも限りがあり、必ずしも自らが探している破断面と同じものが見つかるとは限らない。そのため、特に解析初心者では判断に不安が残ることも多かった。

こうした状況に対して、破断面の電子ファイルによるデータベースを作成することで、書籍の画像集とは異なり容量に基本的には制限の無い画像集が作成できると考え、構築を行った。解析熟練者がデータベースを構築していく時に、観察した結果およびコメントを残していくことで、熟練者の知見を蓄積していくことができる。また、解析初心者はこれをキーワードなどで検索することで、自分が調べたい破断面を容易に閲覧することができる。

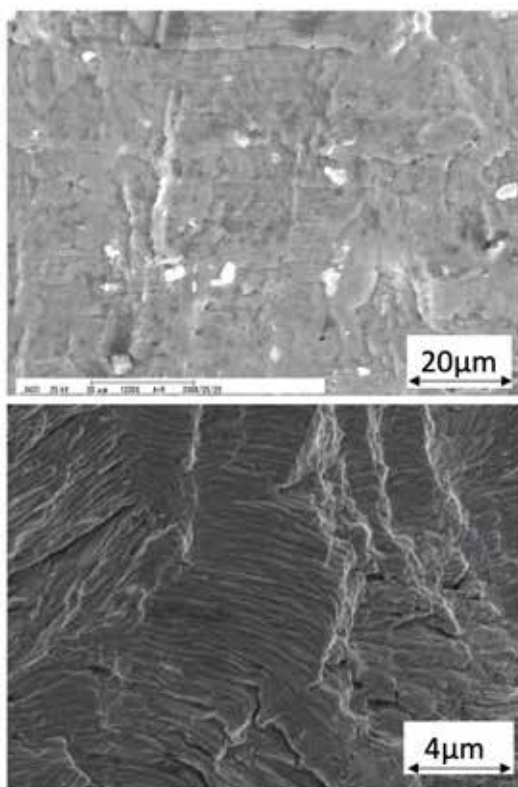


図 1 実機の破断面(上)と試験機で作成した破断面(下)の違い

なお、本データベースは(一社)日本材料学会の疲労部門委員会とフラクトグラフィ部門委員会の中でフラクトグラフィデータベース小委員会を作成し、その中でデータベースの仕様および破壊試験の条件などを審議しながら構築した。

2) 金属破断面データベースの構築

(1) 破壊試験の実施

データベースに掲載する破断面については、表 1 に示す破壊試験を実施した。主に産業機械で一般的に使用されている材料を中心に選定した。熱処理の種類については物質・材料研究機構(NIMS)が公開している金属材料技術研究所疲れデータシートの中の「機械構造用炭素鋼 S45C(0.45C)の疲れ特性データシート」に記載されている熱処理条件を参考にした。

(2) データベースの項目

データベースに収録する項目を

表 2 に示す。タイトルから破面生成までの概要については、破断面の属性に関連するものである。特に破断面生成の概要については、実験の条件または事故の状況などについて記載をする。

観察者のコメントは観察者の知見を収集したデータベースとして、観察初心者の理解を助けることができる。

破断面の外観写真については、破断面全体をカメラで撮影した画像、または細いものに関しては SEM において低倍率で撮影したものを主に掲載している。フラクトグラフィにおいて、低倍率で観察(マクロ観察)を行い、次に高倍率で観察(ミクロ観察)を行うということがルーチンワーク 2) として行われており、基本的には両方の写真を見比べながら原因を推定していくことから外観写真を掲載している。

次に破断面の SEM 写真を掲載する。SEM 写真は観察倍率と画像の解像度を掲載している。観察倍率は特に重要な要素で、この情報が無いと模様のスケールが不明になってしまう。

最後に関連項目として、実験のデータまたはそれ以外の破断面に付随した資料を掲載している。

これらの項目をまとめたものをデータシートとしてまとめた。

(3) データベースの閲覧

データベースはウェブブラウザを介して閲覧することができる。図 2 はデータシートのリストである。項目ごとにソートが可能としてある。次にリストをクリックするとデータシートを閲覧することができる。データシートの例を図 3 に示す。破断面の外観は例えばシャルピー衝撃試験の破断面は試験温度順に並べて掲載すると脆性破面率が減少していくことが容易に理解できる。これにデータシート下部には実験データ(図 3)を付属させてあるため、実験の破断面については全てでは無いがデータとセットで破断面の様子を理解することができる。

2020 年 3 月時点で 44 シートを作成して掲載を行なっている。以下の URL にアクセスすることで閲覧できる。表 1 の実験については全て終了しているが、データの掲載自体は前述のフラクトグラフィ部門委員会の委員などにより今後も継続して実施していく予定である。

フラクトグラフィデータベース

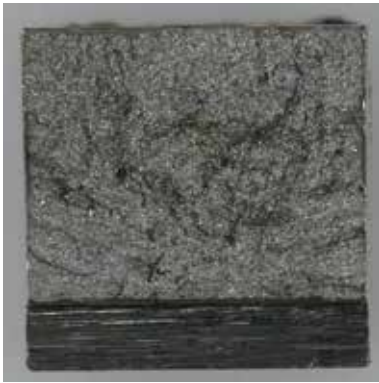
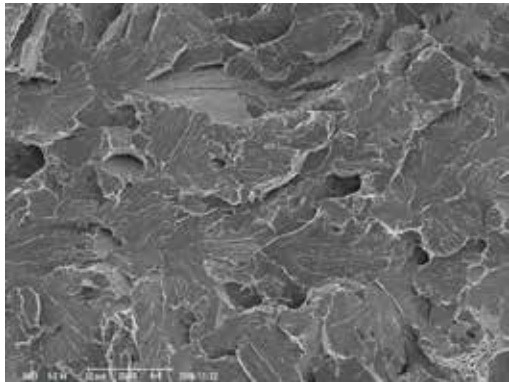
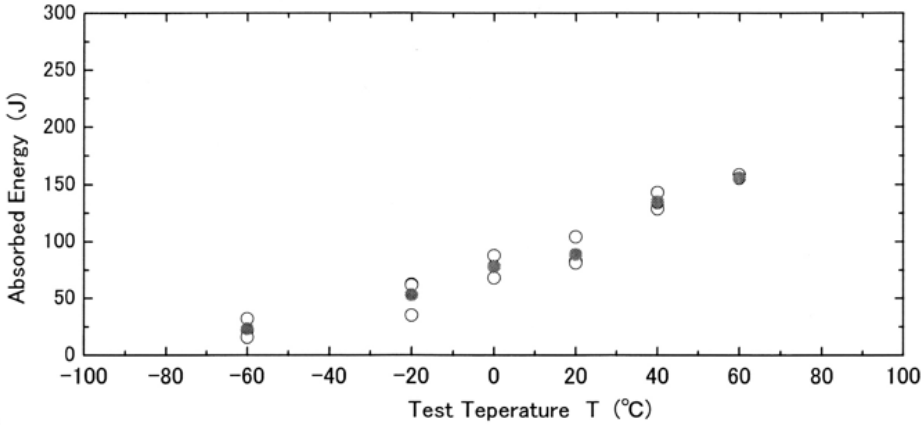
管理 (公社)日本材料学会フラクトグラフィ部門委員会

<http://www.fractography-database.org/>

表 1 破壊試験を行なった材料と実験内容

材 料		実 験			実機	
		引張 試験	シャルピー 衝撃試験	疲労試験		
炭素鋼	S25C	丸棒	済	亀裂進展(1CT)		
	S45C	熱処理無し	丸棒	済	亀裂進展(1CT) 板(片振り・両振り)	
		熱処理 1 1: 焼きならし 845°C/30分 空冷 2: 焼き入れ 845°C/30分 水冷 3: 焼き戻し 550°C/60分 水冷	丸棒	済	亀裂進展(1CT)	
		熱処理 2 1: 焼きならし 845°C/30分 空冷 2: 焼き入れ 845°C/30分 水冷 3: 焼き戻し 600°C/60分 水冷	丸棒	済	亀裂進展(1CT)	
		熱処理 3 1: 焼きならし 845°C/30分 空冷 2: 焼き入れ 845°C/30分 水冷 3: 焼き戻し 650°C/60分 水冷	丸棒	済	亀裂進展(1CT)	
		熱処理 4 1: 焼きならし 845°C/30分 空冷	丸棒	済	亀裂進展(1CT)	
	SM490A	丸棒	済	亀裂進展(1CT)		
	SCM435	丸棒				
	SCM440		済			
	STS410			丸棒		
ステンレス鋼	SUS316			亀裂進展(1CT) 丸棒		
鋳鉄	FC300	丸棒	済			
	FCD450			亀裂進展(1CT)		
	FCD400	丸棒	済	亀裂進展(1CT)		
鋳鋼	不明				有	
アルミ合金	A5083	丸棒		亀裂進展(1CT)		
アルミ鋳物	不明				有	
ニッケル基	インコネル(718)			亀裂進展(0.5CT)		
樹脂	PVC	板				

表 2 破断面データベースに掲載した項目と記載例

項目	記載例						
タイトル	シャルピー衝撃試験破面(SCM440)						
言語	日本語/英語						
作成者	山際謙太(JNIOOSH)						
材料	SCM440						
実機/実験	実験						
破壊機構	脆性破壊/延性破壊						
環境	-60~60℃の大気中						
破面生成の概要	シャルピー衝撃試験による破断面						
観察者のコメント	シャルピー衝撃試験による破断面である。材料はSCM440である。外観写真は左から試験温度が-60℃, 0℃, 20℃, 40℃, 60℃である。高温に従うにつれ、光沢のある脆性部が減り、両端のシェアリップや延性部が増えていくのがわかる。破面写真は左から脆性部, 延性部, 脆性延性境界部の破面写真である。脆性部(下側)にはリバーパターンが観察され、延性部(上側)にはディンプルが観察される。						
外観写真							
破断面 SEM 写真	 <table border="1" data-bbox="970 1249 1241 1368"> <tr> <td colspan="2">脆性破面(擬劈開)</td> </tr> <tr> <td>倍率</td> <td>2000</td> </tr> <tr> <td>画素数</td> <td>[1200, 900]</td> </tr> </table>	脆性破面(擬劈開)		倍率	2000	画素数	[1200, 900]
脆性破面(擬劈開)							
倍率	2000						
画素数	[1200, 900]						
関連資料							

Fractography Database::List

ID#	タイトル	材料	破壊機構	実験/実験#	Last Update
1	S45Cボルトの疲労破面 (1)	S45C	疲労破壊	実験	2019-08-01 23:53:04
2	STS410鋼高サイクル疲労破面	STS410	疲労破壊/延性破壊	実験	2018-08-03 21:14:18
3	オーステナイト系ステンレス鋼SUS316高サイクル疲労破面	SUS316	疲労破壊/延性破壊	実験	2018-08-05 10:07:10
4	シャルピー衝撃試験破面(S45C)	S45C	脆性破壊/延性破壊	実験	2019-08-01 23:53:08
5	オーステナイト系ステンレス鋼の高サイクル疲労破面	SUS316	高サイクル疲労破壊	実験	2019-08-01 23:53:14
6	S45C引張試験破面	S45C	延性破壊	実験	2019-08-02 21:51:18
7	SM490A引張試験破面	SM490A	延性破壊	実験	2019-08-01 23:53:23
8	A5083引張試験破面	A5083 アルミニウム合金	延性破壊	実験	2019-08-01 23:53:28
9	アルミニウム合金(A5083)の疲労破面(R=0.2)	アルミニウム合金(A5083)	疲労破壊	実験	2019-08-01 23:53:32
10	アルミ鋳物の過荷重破面	アルミニウム鋳物	延性破壊	実験	2019-08-01 23:53:41

図 2 データシートのリスト表示

概要

ID: 10 | 00110 | Last update: 2019-08-01 23:53:41

タイトル: シャルピー衝撃試験破面(S45C)

材料: S45C | 破壊機構: 延性破壊 | 実験/実験#: 実験 | Last Update: 2019-08-01 23:53:41

破断面の概要
 シャルピー衝撃試験による破断面
 破断面のコメント
 シャルピー衝撃試験による破断面である。材料はS45Cである。外観写真は右から試験速度が10°C、20°C、30°C、40°Cである。温度に依り、異なる延性破面が現れ、高剛性のアルミ製の破断面が見えていくのがわかる。破断写真は右から延性破、延性破、脆性破(延性破)の破断面である。破断面(下側)にはノッチが観察され、破断面(上側)にはアルミ製の破断面が見られる。

外観写真

<p>試験温度: 10°C</p>	<p>試験温度: 20°C</p>	<p>試験温度: 30°C</p>	<p>試験温度: 40°C</p>
<p>試験温度: 50°C</p>			

破断面 SEM 写真

<p>試験温度: 10°C</p>	<p>試験温度: 20°C</p>	<p>試験温度: 30°C</p>	
-------------------	-------------------	-------------------	--

関連資料

<table border="1"> <thead> <tr> <th>ID#</th> <th>タイトル</th> <th>材料</th> <th>破壊機構</th> <th>実験/実験#</th> <th>Last Update</th> </tr> </thead> <tbody> <tr><td>1</td><td>S45Cボルトの疲労破面 (1)</td><td>S45C</td><td>疲労破壊</td><td>実験</td><td>2019-08-01 23:53:04</td></tr> <tr><td>2</td><td>STS410鋼高サイクル疲労破面</td><td>STS410</td><td>疲労破壊/延性破壊</td><td>実験</td><td>2018-08-03 21:14:18</td></tr> <tr><td>3</td><td>オーステナイト系ステンレス鋼SUS316高サイクル疲労破面</td><td>SUS316</td><td>疲労破壊/延性破壊</td><td>実験</td><td>2018-08-05 10:07:10</td></tr> <tr><td>4</td><td>シャルピー衝撃試験破面(S45C)</td><td>S45C</td><td>脆性破壊/延性破壊</td><td>実験</td><td>2019-08-01 23:53:08</td></tr> <tr><td>5</td><td>オーステナイト系ステンレス鋼の高サイクル疲労破面</td><td>SUS316</td><td>高サイクル疲労破壊</td><td>実験</td><td>2019-08-01 23:53:14</td></tr> <tr><td>6</td><td>S45C引張試験破面</td><td>S45C</td><td>延性破壊</td><td>実験</td><td>2019-08-02 21:51:18</td></tr> <tr><td>7</td><td>SM490A引張試験破面</td><td>SM490A</td><td>延性破壊</td><td>実験</td><td>2019-08-01 23:53:23</td></tr> <tr><td>8</td><td>A5083引張試験破面</td><td>A5083 アルミニウム合金</td><td>延性破壊</td><td>実験</td><td>2019-08-01 23:53:28</td></tr> <tr><td>9</td><td>アルミニウム合金(A5083)の疲労破面(R=0.2)</td><td>アルミニウム合金(A5083)</td><td>疲労破壊</td><td>実験</td><td>2019-08-01 23:53:32</td></tr> <tr><td>10</td><td>アルミ鋳物の過荷重破面</td><td>アルミニウム鋳物</td><td>延性破壊</td><td>実験</td><td>2019-08-01 23:53:41</td></tr> </tbody> </table>	ID#	タイトル	材料	破壊機構	実験/実験#	Last Update	1	S45Cボルトの疲労破面 (1)	S45C	疲労破壊	実験	2019-08-01 23:53:04	2	STS410鋼高サイクル疲労破面	STS410	疲労破壊/延性破壊	実験	2018-08-03 21:14:18	3	オーステナイト系ステンレス鋼SUS316高サイクル疲労破面	SUS316	疲労破壊/延性破壊	実験	2018-08-05 10:07:10	4	シャルピー衝撃試験破面(S45C)	S45C	脆性破壊/延性破壊	実験	2019-08-01 23:53:08	5	オーステナイト系ステンレス鋼の高サイクル疲労破面	SUS316	高サイクル疲労破壊	実験	2019-08-01 23:53:14	6	S45C引張試験破面	S45C	延性破壊	実験	2019-08-02 21:51:18	7	SM490A引張試験破面	SM490A	延性破壊	実験	2019-08-01 23:53:23	8	A5083引張試験破面	A5083 アルミニウム合金	延性破壊	実験	2019-08-01 23:53:28	9	アルミニウム合金(A5083)の疲労破面(R=0.2)	アルミニウム合金(A5083)	疲労破壊	実験	2019-08-01 23:53:32	10	アルミ鋳物の過荷重破面	アルミニウム鋳物	延性破壊	実験	2019-08-01 23:53:41	<p>温度とシャルピーエネルギーの関係</p>	<p>温度と破断靱性の関係</p>	<p>温度と破断靱性の関係</p>
ID#	タイトル	材料	破壊機構	実験/実験#	Last Update																																																																
1	S45Cボルトの疲労破面 (1)	S45C	疲労破壊	実験	2019-08-01 23:53:04																																																																
2	STS410鋼高サイクル疲労破面	STS410	疲労破壊/延性破壊	実験	2018-08-03 21:14:18																																																																
3	オーステナイト系ステンレス鋼SUS316高サイクル疲労破面	SUS316	疲労破壊/延性破壊	実験	2018-08-05 10:07:10																																																																
4	シャルピー衝撃試験破面(S45C)	S45C	脆性破壊/延性破壊	実験	2019-08-01 23:53:08																																																																
5	オーステナイト系ステンレス鋼の高サイクル疲労破面	SUS316	高サイクル疲労破壊	実験	2019-08-01 23:53:14																																																																
6	S45C引張試験破面	S45C	延性破壊	実験	2019-08-02 21:51:18																																																																
7	SM490A引張試験破面	SM490A	延性破壊	実験	2019-08-01 23:53:23																																																																
8	A5083引張試験破面	A5083 アルミニウム合金	延性破壊	実験	2019-08-01 23:53:28																																																																
9	アルミニウム合金(A5083)の疲労破面(R=0.2)	アルミニウム合金(A5083)	疲労破壊	実験	2019-08-01 23:53:32																																																																
10	アルミ鋳物の過荷重破面	アルミニウム鋳物	延性破壊	実験	2019-08-01 23:53:41																																																																

図 3 データシートの表示 (概要と外観写真)

3. ディープラーニングを活用した破断面の自動分類

1) ディープラーニングを導入する目的

フラクトグラフィにより破断面から推定する項目で重要なのは、破壊機構、破壊の起点および最終破断部である。

破壊機構を推定していくためには、SEM などを用いてストライエーションなど破壊機構に対応した特徴的な模様を探していく必要がある。しかし、前述の図 1 のように例えば疲労試験機を使って作成した破断面には引張荷重のみ作用させることができるため幅の整ったストライエーションが観察されるが、実機に生じた破断面の場合は、錆または圧縮荷重の作用などにより対向する破断面同士が接触することで破断面の潰れなどが生じてしまう。これにより、模様が消失してしまい破壊機構の推定が困難である場合もある。また、こうした模様の解釈は観察者の経験に依存しており、客観性に欠く結果となってしまう。一方で、熟練した解析者は年々リタイヤしており、解析技術の伝承も問題となっている。このような状況において、解析の初心者をサポートする仕組みは重要である。

一方で、自動車の自動運転などではカメラの画像から自動車または歩行者などを抽出する画像処理または画像認識の技術が開発されている。こうした画像認識には機械学習が用いられており、その中でも近年はディープラーニングが積極的に用いられている。

ディープラーニングでは多層型ニューラルネットワークに膨大な量のデータを学習させ、推論用モデルを作成する⁵⁾。かつてはコンピュータの計算能力の関係から多層にすることが困難であったが、近年は容易になってきたこともあり、多層型ニューラルネットワークは様々なところで活用されてきている。その一つが画像認識の技術である。ディープラーニングを用いた場合、例えば人が書いた数字の画像を認識し、コンピュータに数値として認識させることが可能⁶⁾である。また、街の風景を撮影した写真から、自動車、歩行者または自転車などを認識し、その領域を示す⁷⁾ことも可能となっている。

これまでフラクトグラフィの分野で行われてきた機械学習では、エキスパートシステムを活用した破断面の分類などがある⁸⁾。機械学習とディープラーニングの違いの一つに、特徴量の決め方がある。これまでの機械学習は技術者が画像を解析し、画像分類に適した特徴量を考えて分類用モデルなどを構築していた。例えばディンプル¹⁾であれば、円形状の模様を捉える分類用モデルを構築していた。しかし、現実の破断面は綺麗な円形ではなく、伸長ディンプルもあり、場合によってはゴミもあり、さらには傾斜した破断面であれば画像の部位によってフォーカスが異なることもあることから、綺麗な円形となることはほとんど無い。こうした場合に、従来型の機

械学習では、例えば円形で無く楕円形のモデルを作るなど、それぞれの状況に応じた処理を考える必要があることから、画像の違いにロバストな分類モデルを構築することが困難であった。

一方でディープラーニングは、学習過程で特徴量が決められていく。そのためにはビックデータと呼ばれる膨大なデータと GPU など積んだ高速な計算機などの計算資源が必要であるが、技術者側で特徴量を決定する必要がなくロバストな分類モデルを構築することが可能である。

こうしたディープラーニングの特性を破断面の解析に応用することで、解析者の経験に依存せず破断面の性状分類が可能となると考えた。もしくは、最終的な判断は観察者が行うものの、その観察者をサポートする仕組みとしても有効であると考えられる。材料強度に関連した研究とディープラーニングの活用という点では、足立ら⁹⁾は、金属のマイクロ組織にディープラーニングを活用し、自動分類に対しての成果を挙げている。

本報では予め破断面模様が同定されている SEM 画像を多層型ニューラルネットワークに学習させた。そして、推定される破壊機構と予め同定されていた破壊機構を比較し、正答率などを検討した。これらについて報告する。

2) 破断面自動分類用多層型ニューラルネットワーク

(1) 畳み込みニューラルネットワークの活用

画像分類を行うための多層型ニューラルネットワークで最も用いられているのが畳み込みニューラルネットワーク(Convolution Neural Network: CNN)⁵⁾である。例えば手書きの 0 から 9 の数字の画像を、CNN に入力して、コンピュータ上の数値として 0 から 9 に分類する CNN などが有名である^{5),6)}。

破断面も画像であり、破壊機構に対応した模様が観察されることから CNN を活用して破断面の模様の分類を試みた。

CNN は以下の 3 種類の層を組み合わせることで構成されている。

- ① 畳み込み層
- ② プーリング層
- ③ 全結合層

これらの層の個々の役割については多くの書籍⁵⁾に記載されているので省略するが、これらを多層に組み合わせることで破断面画像の特徴を CNN が認識し、最終的には式(1)に示す Softmax 関数により各ラベルの確率が求められる。ここではラベルがディンプル、擬劈開などの破断面の模様に該当する。

$$y_i = \frac{e^{x_i}}{\sum_{j=1}^{N_{label}} e^{x_j}} \quad (1)$$

ここで N_{label} は分類に用いたラベル数, i はラベル用の添字, x_i は画像から CNN により求めた最終層の出力である。

この確率が最大値となるラベルを解析した破断面画像が該当する破断面のラベルとしている。例えばある画像の破断面模様を推定するとして, 対象となる破断面の模様がディンプル, 擬劈開そしてストライエーションの 3 分類($N_{\text{label}}=3$)であるとする。そして, CNN の計算途中の結果として, ディンプルである確率が 0.1, 擬劈開である確率が 0.3, ストライエーションである確率が 0.6 と算出されるとする。その場合, 確率が最大であるストライエーションがその画像の破断面模様として推定される。

機械学習には教師付き学習と教師無し学習と 2 種類に大別されるが, ここでは教師付き学習を選択した。予め破断面の画像と既に推定された破断面模様 (ラベル) のセットを用意する。そして, この画像群をトレーニング (training) 用と検証(validation) 用に分割する。

次に 1 回目の学習を実施する。最初にトレーニング用画像を CNN に入力して分類した結果を得る。これをラベル付けと呼ぶ。全てのトレーニング用画像に対してラベル付けを行うことをエポック(epoch)と呼ぶ。次に, 検証用の画像をネットワークに読み込ませ, ラベル付けを行う。トレーニング用, 検証用それぞれにラベル付けの結果と, 正解のラベルを比較し, 正答率を計算する。そして, トレーニング用の画像に対して不正解だったものが正解になるように誤差逆伝播法⁵⁾を用いて畳み込み層のフィルタの値を決定する。これを用いて再び次のエポックを実施する。これを繰り返して行くことで, フィルタの数字が変化し, 正答率(accuracy)が増加する。そして, 指定されたエポック数を学習した後にネットワークの構造と各層のフィルタの値をまとめた推論用モデルファイルが作成される。

CNN の性能は, 様々な評価指標があるが, ここでは正答率と誤差(loss)によって評価した。また, それぞれをトレーニング用データと検証用データにおいて求める。すなわち 4 パラメータを考慮しながら CNN の性能を評価していく。

正答率は, 正答した画像数(N_{correct})を全画像数(N_{total})で除した値(式(2))を用いた。

$$Accuracy = \frac{N_{\text{correct}}}{N_{\text{total}}} \quad (2)$$

誤差は, 交差エントロピー(式(3))を用いた。

$$loss = - \sum_{n=1}^{N_{\text{total}}} \sum_{k=1}^{N_{\text{total}}} d_{nk} \log(y_k) \quad (3)$$

ここで, y_k は式(1)の Softmax 関数の出力ベクトル, d_{nk} は正解のラベルを表すベクトルである。例えば前例の $N_{\text{label}}=3$ の事例で 10 番目の画像の正解がディンプル(最初のラベル)であれば, $d_{10}=(1,0,0)$ となる。

計算結果によっては, 正答率が高いが誤差が大きいと言うケースもある。こうした場合は, 例えば前述の $N=3$ の例においては, それぞれの模様の確率がディンプルは 0.3, 擬劈開は 0.3, ストライエーションは 0.4 といった結果の場合である。式(1)の Softmax 関数を適用した後の結果では, ストライエーションが正解であるため, 正答率は良いが, 誤差が大きいことになる。正答率が高く, かつ誤差が小さい CNN を開発することが重要であり, そのためには CNN の構造を検討する, もしくは後述するハイパーパラメータの検討などが行われる。

ディープラーニングの計算を行うためには, 専用のプログラムライブラリ(Pytorch, Caffe, Tensorflow, Keras など)が存在しているが, ここでは Google が提供している Tensorflow¹⁰⁾と Keras¹¹⁾を使用した。言語は Python 3.6 を使用した。また, ライブラリには典型的な CNN が予め用意されており, 例えば畳み込み層を数層組み合わせ合わせたシンプルな CNN から, VGG16, GoogLeNet, ResNet などの画像分類コンテスト(ILSVRC)で上位に入賞するような CNN まで準備されている。一般的には独自に CNN を開発するよりは, こうした実績のある CNN を活用する方が開発速度も速くなり簡便である。今回は GoogLeNet の中で Inception V3¹²⁾と呼ばれる CNN を活用した。ただし, 後述する過学習を抑制する観点から, 畳み込み層に L2 正則化⁵⁾を追加して CNN を構成した。

(2)解析に使用した破断面画像

破断面の画像は以下の破断面を観察した。例を図 4 に示す。図 4 はそれぞれ以下の破断面の画像である。

- ① 延性破壊させた丸棒の引張試験(ディンプル)
- ② 脆性破壊させたシャルピー衝撃試験後の破断面 (擬劈開¹⁾)
- ③ 疲労亀裂進展後の CT 試験片(ストライエーション)

ただし, 疲労に関しては, 低倍率から高倍率まで倍率を遷移させると, 低倍率ではストライエーションが観察されないが, 高倍率では観察されると様相が大きく変化する。このことから, 観察倍率が 250 倍以下の画像については「低倍率の疲労」, 250 倍以上は「中倍率の疲労」, ストライエーションが画像全体に観察されるものは「ストライエーション」とラベルをつけて分類することとした。またこれらに加えて, 粒界破壊の破断面を観察し, 合計 6 ラベルの分類とした。

さらに同一の SEM で撮影された画像のみの解析であると, その SEM のコントラストなどの特徴が解析に影響することを避けるため, Google の画像検索なども活用して, ディンプルなどのキーワードを用いて検索した結果から得られる画像も一部画像群に入れた。

ディープラーニングを行うための画像の枚数は基本的には多い方が良いとされているが, 一方で破断面の画像

¹ 金属に延性がない状態で過大荷重が作用した時に破断面に観察される多角形状の模様

の枚数を大量に用意することはスキャン時間などを考慮すると困難である。このような場合、データオーグメンテーション(水増し)と呼ばれる作業を行い、画像の枚数を増やすことがある。具体的には、画像の切り出し、回転、拡大縮小などを行い、同一の画像から複数の画像を作成することである。

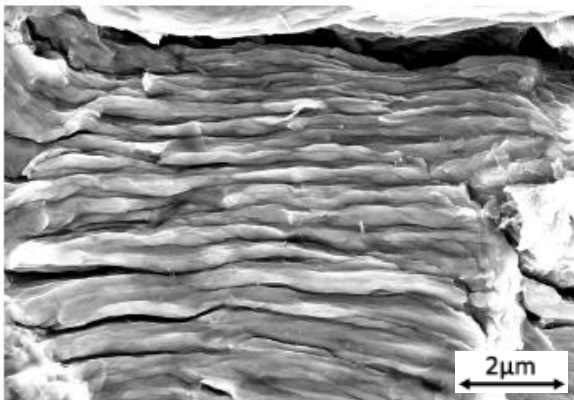
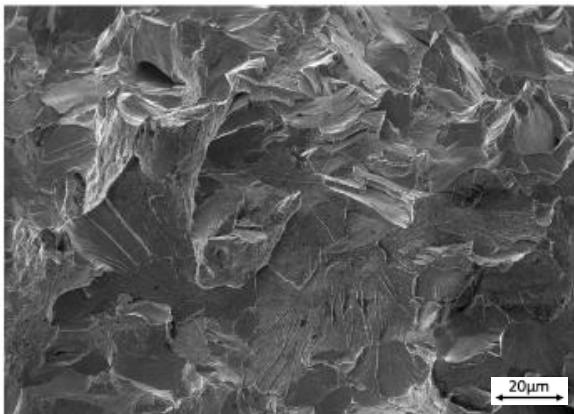
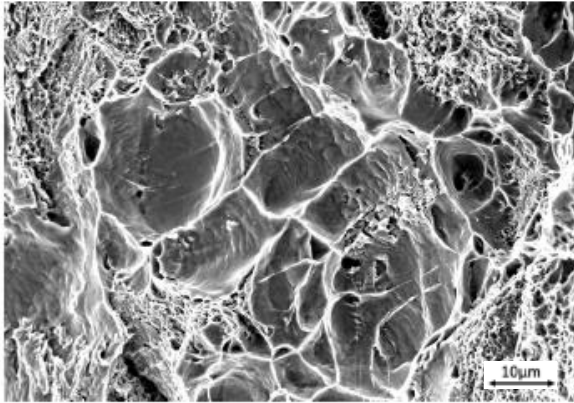


図 4 解析に使用した画像の例(上からディンプル、ファセット¹、ストライエーション)

本報で行った計算では、具体的には 2400×1800 程度の画素数を持った SEM 画像から 598×598 画素で切り出し、その画像を 90, 180 及び 270 度回転させた。さらに、その 4 種類の画像を上下に反転させた。左右の反転は同じ

画像ができるため行わない。したがって、1 枚の切り出した画像からは 8 枚の画像を作成している。これを元の画像からランダムに 64 枚程度抜き出しているため、例えば 1 枚の SEM 画像から 512 枚の画像を作成していることになる。もちろんこれは元の画像の解像度にも依存している。

解析に使用した元の画像ファイルとオーグメンテーション後の画像枚数を表 3 に示す。検証用の画像がトレーニング用画像の中に入ると、検証の意味が無くなる。したがって、オーグメンテーション前の画像の中で 8 割をトレーニング用とし、残りの 2 割を検証用と分割した上でオーグメンテーションを行なった。

表 3 ラベルと画像枚数

ラベル名	初期枚数	オーグメンテーション後の枚数(万枚)
ディンプル	416	12.2
擬劈開	400	12.5
疲労(低)	358	10.6
疲労(中)	638	21.3
ストライエーション	701	22.6
粒界割れ	240	6.1

(3)CNN に必要なハイパーパラメータの決定

計算に使用したコンピュータは HPC システムズ製である。CPU は Intel Xeon (E5-2640, clock 2.4GHz, 10 core)で 20 CPU、メモリーは 512G byte、GPU は NVIDIA 製の TITAN V を 10 枚搭載している。

CNN を実行するためには、技術者側で設定しなければならないパラメータ(ハイパーパラメータ)がある。例えば初期学習率(lr)、学習ごとに kernel は変化するが、その変化のさせ方を決定するオプティマイザ(optimizer)、およびオプティマイザ内部で必要とする係数、1 回の学習で学習させる画像枚数であるバッチサイズ(batch size)、その他にも過学習を抑制するための dropout 率、kernel 正則化率など、多くのパラメータを設定する必要がある。また、ハイパーパラメータの初期値は技術者の経験に依存することが多い。しかしながら、パラメータの数も多いことから、1 回の計算時間も数時間から長いもので数日とかかることから、最適化なパラメータセットを見つけることが非常に困難である。

例えばパラメータが 2 種類(P1, P2)として、正答率が最大となるための P1, P2 を求めることとする。しかし、正答率と P1, P2 の関数は未知(ブラックボックス関数)であることから、正答率が最大値となると証明できる P1, P2 を求める事は困難である。従って、パラメータは探索により推定するより他が無い。

¹ 金属結晶における劈開面で破壊しているものの周辺部に塑性変形があるような破断面

図 5 はパラメータの探索方法の概念図である。P1, P2 のパラメータを最適化すると仮定して、評価する値を円の面積で示している。CNN の場合、この値は正答率もしくは誤差を用いることが多い。特に検証用データの誤差を用いることが、汎化性能¹のチェックとなるためよく用いられる。こうした問題に対して、図 5 の左

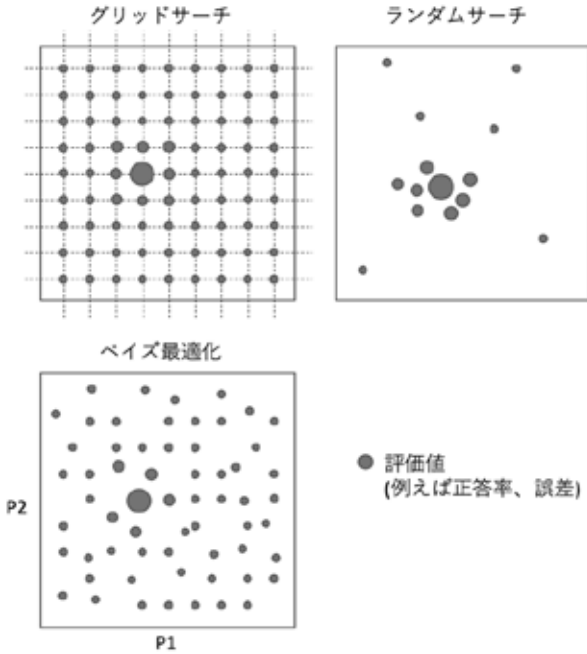


図 5 パラメータ探索の概念図

上に示すようなグリッドサーチ、または図 5 の右上に示すランダムサーチを行ってきた。

グリッドサーチは、パラメータの空間を一定の間隔で区切り、全てのパラメータの組み合わせを計算して、その中で評価する値が最大となるパラメータセットを P1, P2 の最適値とする手法である。ランダムサーチは、パラメータの空間をランダムに探索し、その中で得られた評価値の最大値となるパラメータセットを最適値とする。

そして、その中で最も性能が良いパラメータを活用していたが、もちろんこれらの手法で得られたパラメータセットが最も性能が良いという保証は無い。局所解を求めている可能性もある。

そこで、近年はベイズ最適化を活用したパラメータ最適化¹³⁾(図 5 左下)が主流となっている。

ベイズ最適化は最初の数点はランダムに決定して計算を行う。その結果とベイズの定理から確率的に評価値が最大、もしくは最小となると推定されるパラメータセットの近くを重点的に探索する。また探索した結果を元に次の探索するパラメータセットを決定していくことから、効率的なパラメータセットの検索が可能となる。

ここではベイズ最適化の手法に Optuna¹⁴⁾と呼ばれるライブラリを活用した。

3)解析結果

(1) 破断面の画像の学習によるモデルファイルの作成

図 6 に計算結果の例を示す。横軸はエポック、縦軸は上段が正答率、下段が誤差である。

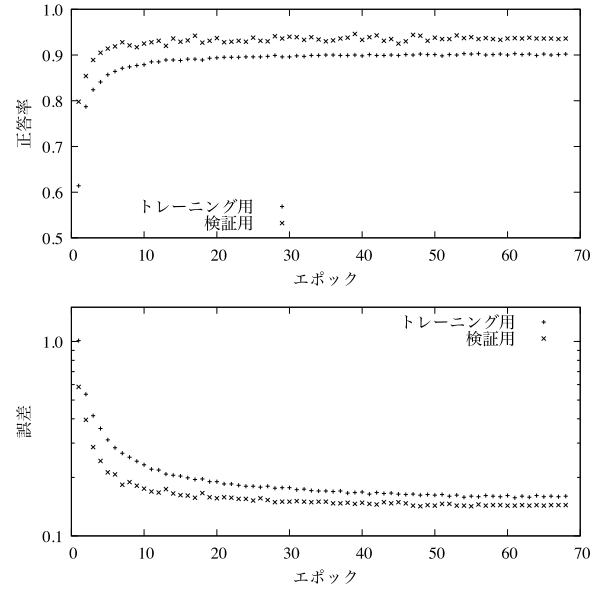


図 6 エポックと正答率および誤差の関係

約 30epochs で、トレーニング用のデータについては 90%の正答率となった。最初が 60%程度の正答率なので、学習が進展するにつれ、正答率が向上していることがわかる。一方で、テスト用のデータについては、92%程度の正答率となっている。なお、計算時間は GPU を 2 枚使用して 1epoch が約 1 時間なので、70~80 時間の計算時間である。

表 4 Optuna によるハイパーパラメータ探索により得られた値の例

パラメータ名	値
学習率	2.68×10^{-6}
Dropout 率	0.59
正則化率	9.54×10^{-8}
Optimizer	SGD
Momentum (SGD 用)	0.464
Decay (SGD 用)	1.67×10^{-6}
Nesterov (SGD 用)	無し

一方で Optuna によるハイパーパラメータの最適化にはこの計算を繰り返し行う必要がある。使用している計算機は GPU が 10 枚使用できることから、ハイパーパラメータのセットを 5 セット並列に実行した。さらにこの 5 パターンが終わった結果を考慮した上で次のハイパーパラメータのセットが決定され、次の 5 セットが計算される。この繰り返しを 6 回実施し、合計 30 セットのハイ

¹ 学習に使用していない画像などの入力に対して出せる正答率などの性能

パラメータの計算を行った。よって計算時間はおよそ 20 日程度となる。最適化の評価値には、検証用データの誤差を活用し、かつ、誤差が最小になるようにパラメータの探索を行った。表 4 にその例を示す。

ただし、これらの値は学習用データが変わればまた変化するため、これらの値で計算を行えば必ず良い結果が得られるわけではない。

CNN の計算で注意しなければいけないことに過学習がある。それはトレーニング用のデータに CNN の kernel の値が特化しすぎてしまい、検証用のデータについては正答率が上がらないといった状況が生じる。したがって、過学習が生じている場合は汎化性能が低いモデルファイルができていてと考えられ、学習したことの無い画像に対しての解析については、正答率が低くなる。

今回の計算ではトレーニング、検証用のどちらの正答率も同じように上昇し、誤差についても同じように減少していったことから、過学習になっているとは言えない。ただし、過学習であるか否かの定量的な指標は無い。

(2) 推論ソフトウェアの作成

ディープラーニングの計算を行い分類用のモデルファイルを作成した次は、これを活用してトレーニング用または検証用のデータに無い画像を実際に分類するための推論ソフトウェアを作成する必要がある。

例としてウェブブラウザを介して画像を推論用プログラムに入力し、結果を表示するプログラムの例を図 7 に示す。プログラムは Python のライブラリである Flask¹⁵⁾などを活用して構築した。

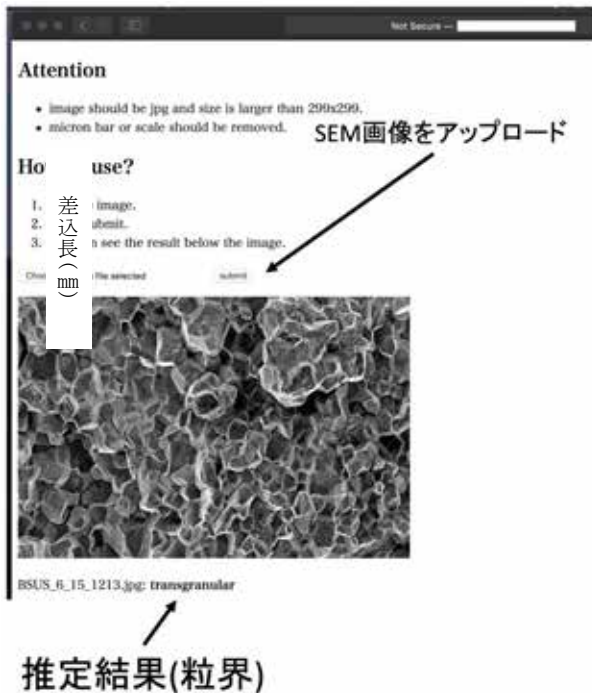


図 7 破壊機構推論プログラムの例

ここでは、画像内からランダムに 10 領域を選択し、それぞれの領域において破断面の様態を推定する。その

結果、最も数の多いラベルをその破断面の様態として決定した。上記の例では目視による判断では粒界破壊の破断面であるが、推定結果も粒界となっており、両者が一致している。

(3) 結果の信頼性に対する議論

近年は学習用データに対する信頼性と、推論結果に対する説明責任も求められるようになってきた。

例えば本報で使用した画像データと分類結果については、事前に筆者が分類していることと、Google などの画像検索の結果を入れていることなどから、その分類結果の信頼性について客観性が高いとは言えない。また、それを学習した結果であるモデルファイルの性能は、筆者らの分類能力に影響を受けていると言える。

一方で、こうした技術者の分類能力を極力排除した学習データを作成しようと考えたとき、破壊試験により既知の破壊機構による破壊を行い、その破断面のみを観察し、学習用データとする方法が考えられる。こうした方法により、破壊試験と破断面画像のセットに技術者の主観を入れない、言わば客観性の高いデータを作成することは可能である。しかし、こうした言わば教科書のような破断面に対して、実機の破断面は潰れ、汚れまたは錆などが多い。従って、実験のみで作成した学習用データと、実機の破断面とは少なからず乖離がある。そのため、こうした破壊試験のみのデータで学習することは、学習用データとして客観性は高いが、十分な汎化性能があるモデルファイルが構築できる可能性については議論として残されていると考えられる。

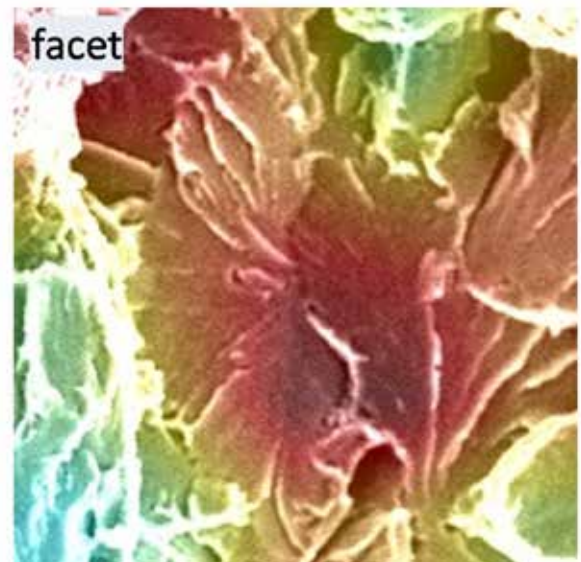


図 8 破断面の推論結果に対するヒートマップ表示

また、推論結果に対しての説明責任については、図 8 に示したヒートマップなどが提案されている¹⁶⁾。これは推論する画像の中で判断に至った領域を赤く示している。ファセットは劈開面と、その中に広がるリバーパターンが特徴であるが、図 8 では、その部分が赤く強調されており、人間による判断と近い推論がなされてい

ると言える。こうした画像を提示することで、推論に至った根拠を示すことができる。

また、これ以外では **Inference function**¹⁷⁾ と呼ばれる手法もあり、学習用画像の中から最も推論用画像に近いと判断される画像を提示する手法である。こうした様々な手法が現在も推論結果の説明のために提案されているため、こうした手法を取り込みながら破断面の推論結果を説明していくことになる。

4. まとめ

破断面の分野においては、熟練観察者の現象に伴う技術伝承が問題となっているが、本研究では次の2点を実施することにより、この問題の解決を試みた。

- ① 金属破断面データベースの構築
- ② ディープラーニングを活用した破断面の自動分類

データベースについては、熟練観察者の知見などが保存でき、それらを他の観察初心者などが閲覧することにより知識の伝承が可能となると言える。

ディープラーニングの活用については、これまでに行われてきたエキスパートシステムなどの自動分類とは全く異なる手法により自動分類を実現した。

破断面データベースは破断面解析の定性的な分野において、ディープラーニングは定量解析の分野において、それぞれこれまでに無い環境を提供することを実現したという点で高度化を実現したと言える。

謝 辞

本研究の遂行にあたっては、(公社)日本材料学会疲労部門委員会およびフラクトグラフィ部門委員会の委員の方々、横浜国立大学研究コンソーシアム「フラクトグラフィとディープラーニングの融合研究コンソーシアム(FraD)」にご参加いただいている会員の方々、東京電機大学辻研究室の学生諸君には多大なるご協力をいただきました。ここに付記し、謝辞といたします。

参 考 文 献

- 1) 小寺沢 良一(1985), 金属破断面写真集, テクノアイ.
- 2) 日本材料学会フラクトグラフィ部門委員会編(2000), フラクトグラフィ, (公社) 日本材料学会.
- 3) 西田新一(1998), フラクトグラフィと破面解析写真集, 総合技術センター.
- 4) 塩谷 義, 松尾 陽太郎, 服部 敏雄, 川田 宏之(2010), 最新 フラクトグラフィ 各種材料の破面解析とその事例, 株式会社テクノシステム.
- 5) 岡谷 貴之(2015), 深層学習 (機械学習プロフェッショナルシリーズ), 講談社.

- 6) Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner(1998), Gradient-based learning applied to document recognition, Proceedings of IEEE, Vol. 86, No. 11, pp. 2278-2324, 1998.
- 7) Kaiming He(2017), Georgia Gkioxari, Piotr Dollár, Ross Girshick, "Mask R-CNN," arXiv, <https://arxiv.org/abs/1703.06870>. (2020年7月1日閲覧)
- 8) 箕島 弘二, 山崎 伯公, 駒井 謙治郎(1994), 環境助長割れフラクトグラフィ診断エキスパートシステムの開発, 日本機械学会論文集 A 編, Vol. 60, No. 579, pp. 2518-2524, 1994.
- 9) 足立 吉隆, 田口茂樹, 弘川 奨悟(2016), ディープラーニングによる組織識別率の検証, 鉄と鋼, Vol. 102, No. 12, pp. 722-729, 2016.
- 10) Tensorflow, Google, <https://www.tensorflow.org/>. (2020年7月1日閲覧)
- 11) KERAS, <https://keras.io/>. (2020年7月1日閲覧)
- 12) Christian Szegedy, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott Reed, Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke, Andrew Rabinovich(2014), Going Deeper with Convolutions, arXiv, <https://arxiv.org/abs/1409.4842>, 2014. (2020年7月1日閲覧)
- 13) C. M. Bishop(2006), Pattern Recognition and Machine Learning, Springer, 2006.
- 14) Optuna, Preferred Networks, Inc., <https://optuna.org/>. (2020年7月1日閲覧)
- 15) Flask, <https://frask.palletsprojects.com/>. (2020年7月1日閲覧)
- 16) Ramprasaath R. Selvaraju, Michael Cogswell, Abhishek Das, Ramakrishna Vedantam, Devi Parikh, Dhruv Batra (2017), Grad-CAM: Visual Explanations from Deep Networks via Gradient-based Localization, arXiv, <https://arxiv.org/abs/1610.02391>, 2017. (2020年7月1日閲覧)
- 17) Pang Wei Koh and Percy Liang(2017), Understanding Black-box Predictions via Influence Functions, arXiv, <https://arxiv.org/abs/1703.04730>, 2017. (2020年7月1日閲覧)