

深層学習技術を利用した  
インダストリアル・エンジニアリングの効率化に関する研究

2021年12月

長崎大学大学院工学研究科

村井 浩一

# 目次

---

第1章 序論 .....	1
1.1 背景 .....	1
1.2 本研究の目的と対象範囲 .....	3
1.2.1 目的 .....	3
1.2.2 範囲 .....	3
1.3 本研究の内容 .....	4
1.3.1 作業分析結果と深層学習を利用した作業手順のリアルタイム監視 .....	4
1.3.2 回帰と分類用の深層学習を用いた製造現場での作業動作特徴抽出 .....	4
1.4 本論文の構成 .....	5
参考文献 .....	6
第2章 インダストリアル・エンジニアリングにおける課題 .....	8
2.1 インダストリアル・エンジニアリングの定義 .....	8
2.2 インダストリアル・エンジニアリングの発展過程 .....	8
2.3 インダストリアル・エンジニアリングにおける分析的アプローチ .....	10
2.3.1 時間研究 .....	11
2.3.2 動作研究 .....	13
2.4 インダストリアル・エンジニアリングにおける課題 .....	15
2.5 深層学習を用いたインダストリアル・エンジニアリングにおける課題解決 .....	17
参考文献 .....	18
第3章 関連研究および基礎理論 .....	22
3.1 製造現場でのデジタル技術を利用した改善活動に関する関連研究 .....	22
3.1.1 作業ミスの検出 .....	22
3.1.2 動作分析・時間分析 .....	26
3.2 深層学習に関する基礎理論 .....	30
3.2.1 畳み込みニューラルネット .....	30
3.2.2 ネットワークの学習 .....	33
参考文献 .....	38
第4章 作業分析結果と深層学習を利用した作業手順のリアルタイム監視 .....	40
4.1 はじめに .....	40

4.2	リアルタイム作業手順監視システムの特徴	40
4.2.1	要求条件	40
4.2.2	作業手順監視方式	41
4.2.3	ユースケース	42
4.2.4	市販の作業分析ソフトでの分析結果の利用	43
4.2.5	リアルタイム性の高いCNNモデルの採用	43
4.3	リアルタイム作業手順監視システムの構成	45
4.3.1	ハードウェア構成およびシステム構成	45
4.3.2	アプリケーション	45
4.4	評価実験	48
4.4.1	実験環境	48
4.4.2	実験対象	50
4.4.3	実験準備	50
4.4.4	実験内容	53
4.4.5	実験結果	55
4.5	考察	58
4.6	まとめと今後の課題	59
	参考文献	61

第5章	回帰と分類用の深層学習を用いた製造現場での作業動作特徴抽出	62
5.1	はじめに	62
5.2	作業動作特徴抽出システムの要件と特徴	62
5.2.1	要求条件	62
5.2.2	ユースケース	63
5.2.3	2種類のCNNモデルの採用	63
5.2.4	本システムの利用条件	64
5.3	作業動作特徴抽出システムの構成	65
5.3.1	ハードウェア構成およびシステム構成	65
5.3.2	身体や工具位置の予測処理	66
5.3.3	作業動作の分類処理	70
5.3.4	作業動作特徴の分析	72
5.4	ボルト締め作業での評価実験	72
5.4.1	実験対象	72
5.4.2	実験準備	73

5.4.3	実験内容	75
5.4.4	実験結果	76
5.5	実用性に関する評価実験	81
5.5.1	実験対象	81
5.5.2	実験準備	82
5.5.3	実験内容	83
5.5.4	実験結果	83
5.6	考察	84
5.7	まとめと今後の課題	85
	参考文献	87
第6章	まとめ	89
	参考文献	91

## 第1章 序論

### 1.1 背景

日本国内の製造業の状況について、短期的に見ると新型コロナウイルス感染症（COVID-19）の世界的な感染拡大の影響を受け、売上高および営業利益は減少傾向にある。2021年版ものづくり基盤技術の振興施策（ものづくり白書）によると、8割を超える企業がCOVID-19の感染拡大にともない業績が悪化している状況であるとされている[1]。また、東京商工リサーチが行ったCOVID-19の感染拡大による企業活動への影響に関するアンケート調査においても、2021年3月時点で一部の企業（8.2%）は既に影響が収束したと回答しているものの、大半である72.7%（大企業80.4%、中小企業71.3%）の企業が、その影響が継続していると回答している[2]。このように、COVID-19の感染拡大は、社会経済や社会生活の様々な部分で甚大な影響を及ぼしている。具体的には、需要面に関する影響や、生産・物流・配送などの供給面に関する影響をはじめ、緊急事態宣言などにもなう行動制限やリモートワークのような新しい働き方など、コロナ禍での新たな生活様式における消費者の行動変化があげられる。

COVID-19の日本国内での感染拡大は2020年の春先から始まっており、現時点においても完全な収束時期は不透明であるが、ここで一旦、COVID-19による影響を受ける前の時点の製造業の状況を把握するために、2019年度前半までの状況に目を移す。景気観や売上経常利益との相関性が高いとされている日本銀行全国企業短期経済観測調査（日銀短観）の業況判断DI推移[3]（2019年3月時点）によると、製造業全般については2008年のリーマンショック前後を底に業況は回復傾向にあり、2013年半ば以降はプラス圏内を推移している状況が続いていた。一方、国内の製造業の就業者の推移を見ると、内閣府の2019年度国民経済計算年次推計[4]に示されたとおり、1994年の1,403万人から2019年は1,076万人と四半世紀の間で23%減少しており、全産業に占める製造業就業者数の割合も1994年の20.9%から2017年は15.6%まで減少している。特に中小企業においては人材不足が顕著であり、中小企業白書2019年版における産業別従業員数過不足DI推移[5]に示されたとおり、リーマンショックの影響により過剰とされていた人材が、それ以降徐々にマイナス側（人材不足側）に拡大していることが分かる。製造業については建設業、サービス業に次いで人手不足が進んでおり、2013年第3四半期にマイナス1.8と人材不足に転じて以降、マイナス（人材不足）幅は拡大を続け、2018年第1四半期にはマイナス23.1となった。このように、2019年度前半までは、2008年のリーマンショックが与えた経済に対する悪影響が収まり、各企業ともに企業収益に改善傾向が見られる一方で、少子高齢化により生産年齢人口が減少し、人材不足という課題の深刻さが増していた時期であった。

続いて、COVID-19の流行が始まったとされる2019年度後半から2020年度前半の国内製造業の状況について見てみる。2020年3月時点における日銀短観の業況判断DI推移[6]によると、2019年第1四半期までプラス圏内を推移してきた業況が、第2四半期以降マイナス側へ落ち込む状況となった。この要因は、2018年秋以降の米中貿易摩擦による中国経済の減速や、海外経済の不透明感に対する懸念、自

然災害などの影響があげられ、その結果として、2019年は製造業の営業利益は2013年以降最低となった[7]。さらに、2020年第1四半期については、日本国内においてもCOVID-19の影響が表面化しはじめた時期でもあった。

最後に、COVID-19の影響下にある2020年度後半から2021年度に焦点をあてる。2021年3月時点の日銀短観の業況判断DI推移のグラフを見ると、COVID-19の感染拡大により11年ぶりの低水準となった2020年第2四半期以降は全般的にマイナス幅が縮小しており、大企業製造業においては2021年度にはプラス側に転じているが、その水準はまだ低い状況である[8][9]。また、中小製造業についても回復の兆候は見えるが、依然としてプラス側には至っていない状況である。実際に多くの企業がCOVID-19の影響により業績が悪化しており、このことは、2021年5月に公開された労働政策研究・研修機構（The Japan Institute for Labour Policy and Training : JILPT）による調査結果に示されたとおり、従業員規模に関わらず8割以上の企業が「悪化した」又は、「やや悪化した」と答えていることから分かる[10]。また、中小企業の課題であった人材不足について、2021年版中小企業白書内の産業別の従業員数過不足DI推移[11]によると、人材不足で推移していた状況のなかで、COVID-19の感染拡大時期には一転してプラス側（人材が過剰）に転じており、感染症拡大が雇用にも大きな影響を及ぼしていることが確認できる。しかし、2020年第4四半期以降は再びマイナス側（人材不足）に推移し、人手不足の状況が進んでいることが分かる。東京商工リサーチのアンケート調査結果[12]においても、COVID-19収束後の懸念として最も多かったのが、「経済活性化にともなう人手不足（47.5%）」とされており、人材不足に関する課題については、今後さらに増えることが予想される。

このように、日本国内の製造業の状況について、短期的に見ると新型コロナウイルス感染症の影響による業績低迷があげられるが、中期的に見ると人材不足が重要な課題となっている。さらに、現在ではCOVID-19の感染拡大をはじめ、気候変動や自然災害など、予測不能な環境の変化に対するリスクが高まっている状況である。この急激な環境変化に対応していくためにはデジタル技術が有効なツールとされ、COVID-19の流行がデジタル化に対する重要性を再認識させる1つの契機であるとも言われている[13][14]。実際に、野村総合研究所が中小企業に対して行ったCOVID-19の流行前後におけるデジタル化に対する優先度の変化に関する調査によると、優先度が「高い」または「やや高い」と答えた企業が、流行前の45.6%から流行後は61.6%と16%アップしていることが分かる[15]。既に多くの企業においてデジタル技術の活用が進められており、2021年版ものづくり白書では、JILPTによるデジタル技術を活用している企業における、活用理由とその狙いに関する2021年5月時点のアンケート結果[16]が掲載されている。このアンケート結果が示すとおり、デジタル技術は在庫管理の効率化や開発・製造などのリードタイムの削減に活用されている一方、作業負担の軽減や作業効率の改善、ベテラン技術の見える化、データ化による技能伝承の円滑化といった、人材に関わる課題の解決や人材育成にも活用されていることが分かる。つまり、製造業のデジタル化とは、ロボットの導入による自動化やITシステムの導入による業務効率化だけではなく、デジタル技術を用いた労働者の負荷軽減やムダな作業の排除などの改善業務、さらには、人材の育成や技能の伝承などにも役立てられていること

が理解できる。

## 1.2 本研究の目的と対象範囲

### 1.2.1 目的

先に述べたとおり，日本国内の製造業においてはCOVID-19による経営への悪影響を受けたものの，人材不足，人材育成については現時点でも深刻な課題とされたままであり，このことはCOVID-19収束後の大きな懸念事項とされている。ものづくりに携わる人材が不足している状況で経済成長を支えていくためには，技能伝承の視点に加え，現場レベルでの労働生産性の向上の視点，経営レベルでの労働生産性向上の視点が重要とされ[17]，製造業ではそのための手段の1つとして，インダストリアル・エンジニアリング（IE）[18]による改善活動が進められている。

そこで本研究では，製造現場での労働生産性向上を目指して，IEに対してデジタル技術の1つである深層学習の技術を用いることにより，製造現場での改善活動や人材育成を効率的に進めることができる汎用的なシステムについて開発し，その有効性を明らかにすることを目的とする。人材の育成とデジタル技術活用との関係については，2021年版ものづくり白書に掲載されたJILPTのアンケート調査結果[19]に示されたとおり，「人材の育成・能力開発の取組についてうまくいっている」と回答した企業が，デジタル技術を活用している企業では51.9%，デジタル技術を活用していない企業では41.3%と10%ほどの開きがある。つまり，デジタル技術の活用は，人材育成および能力開発に有効な手段であることが認められている[20]。しかし，中小企業を中心とした製造業においては，デジタル技術を活用できるIT技術者の育成や確保は容易ではないという課題もあげられている[21]。この問題を解決するために本研究では，デジタル技術の1つである深層学習の技術を利用した。深層学習は，従来人間が自然に行うタスクをコンピュータに学習させ，実行することができる技術であり，その学習に使用する特徴量についても，学習時にデータから自動的に抽出される仕組みとなっているため，高度なIT技術や統計に関する専門知識を用いた特徴量設計などを必要としない。これにより，IEやITに関する専門知識を持つ人材が不足している状況においても，深層学習技術がIE活動におけるタスクの一部を担当するという形で，効率的に改善活動や人材育成を進めることが可能になるものと考えられる。

### 1.2.2 範囲

IEについてはその概念の見方に二つの主流がある。1つは生産工学を主体とするもので「伝統的IE」と呼ばれる分析的なアプローチであり，他の1つは経営科学を背景とする「近代的IE」である[22]。また，適用範囲としてとらえた場合，作業やものづくり現場での生産性向上を対象としたものは「狭義のIE」と呼ばれ，生産管理や品質管理などを含む生産システム全体を対象としたものは「広義のIE」と呼ばれる[23]。いずれの手法も時代とともに発展しているが，本研究では製造現場における人材不足が重要な課題であると考え，作業者の労働生産性向上に重点を置いた分析的アプローチである「伝統的IE」且つ，「狭義のIE」を対象とする。この「伝統的IE」且つ「狭義のIE」の目的は生産性の向

上、さらにコストを改善し利潤を増大して経営効率を高めることにあり、この手法については、20世紀初頭に考案されたTaylor (F. W. Taylor, 1856~1915) による「時間研究」と、Gilbreth (F. B. Gilbreth, 1868~1924) による「動作研究」が発展、体系化されたものである[24]。本研究では伝統的且つ狭義のIEの定義のなかでも基本とされる、Taylorの「時間研究」とGilbrethの「動作研究」を対象に、その効率化を目指す。

### 1.3 本研究の内容

筆者はこれまで、IE用の作業分析ソフトの導入を通して、いくつかの企業に携わってきた。作業分析ソフトは、製造現場で撮影された動画を利用して効率的に作業の分析を行うことができ、IEにとって有効なツールである一方、製造現場での作業者の日常的な作業の状況について常に把握することは難しい。そこで本研究では、深層学習の技術を用いることにより、製造現場での作業者の作業状況や作業時の動作を把握するための手法について、汎用的な仕組みとして実現することを目指した。

#### 1.3.1 作業分析結果と深層学習を利用した作業手順のリアルタイム監視

製造現場では作業の標準化を定めることや、人による作業や設備の稼働に含まれる無効な要素を発見しその排除や改善を行うことを主な目的として、IE手法の1つである「時間研究」が利用されている。「時間研究」では製造現場をストップ・ウォッチで計測する方法や、撮影した動画を利用して分析する方法が用いられているが、これらの分析の結果から定められた標準作業に対して、製造現場の作業者が適切に作業を行っていることを確認することは容易ではない。作業時間や作業手順のミスを実タイムに検知するシステムがいくつか提案されているが、これらの多くは対象となる作業に対して専用の設計となっており、また、検知した作業ミスに対して、正しい手順での作業のやり直しまで指示するような汎用的なシステムは報告されていない。

そこで本研究では、作業分析結果と深層学習を利用して作業手順を実タイムに監視し、誤りとして検出された場合には正しい手順をティーチングすることができる、汎用的な手法について研究を行った。本研究では、作業分析結果である標準作業と標準作業時間および、分析時の作業動画から学習された深層学習モデルを用いることにより監視基準に客観性を持たせ、作業手順の誤りや作業時間の乖離が検知された際には、分析時の模範となった正しい作業手順で作業員へティーチングを行うことを可能とした。さらに、深層学習用の学習データについては、作業分析結果と作業時の動画から自動的に生成することにより、深層学習を利用する際の課題とされる学習データの準備にかかる労力と時間を大幅に削減することを可能とした。本手法について組立作業の模擬環境で評価実験を行い、有効性について明らかにした。

#### 1.3.2 回帰と分類用の深層学習を用いた製造現場での作業動作特徴抽出



製造現場の作業は多岐にわたり、作業時の動作に対する技能や作業効率を分析し、評価するための視点についても、それぞれの作業により異なる。作業者の動作を把握して分析する手法としては、IEにおける「動作研究」が用いられてきたが、この分析手法の多くは人手に頼っている。作業者の動作を自動的に把握するための手法としていくつかの研究が進められているが、従来の多くの方法については、各製造現場での作業や動作に合わせた専用のシステムが必要となり汎用化が難しい。

そこで本研究では、深層学習技術を用いることにより、カメラで撮影した動画から作業時の特徴となる動作を抽出することができる汎用的な手法について研究を行った。本手法では、回帰用のCNNによる物体の位置の予測と、分類用のCNNによる動作の分類により特徴をあらわす作業動作を抽出する。本手法について一般的な作業の1つである、スパナを利用したボルト締め作業を対象として実験を行い、有効性について明らかにした。また、溶接作業での実験を通して、本手法の汎用性および検討課題について明らかにした。

#### 1.4 本論文の構成

本論文の構成について、1章では、研究の背景である日本国内の製造業の状況および研究の目的について述べ、2章では、IEに関する手法および、IE活動における課題、本研究における深層学習を利用した課題解決へ向けたアプローチ方法について述べる。3章で関連研究および、本研究において利用した深層学習の理論と技術について述べたうえで、4章では、深層学習技術を利用した作業手順のリアルタイム監視手法に関する研究内容について述べる。5章では同じく深層学習技術を利用した作業動作の特徴抽出手法に関する研究内容について述べ、6章で結論として本論文のまとめを行う（図 1-1）。



図 1-1 本論文の構成

## 参考文献

- [1] 経済産業省, 厚生労働省, 文部科学省: 令和2年度 ものづくり基盤技術の振興施策, p.122, 入手先 <<https://www.meti.go.jp/press/2021/05/20210528002/20210528002-3.pdf>> (参照2021-06-20).
- [2] 東京商工リサーチ: 特別企画 第14回「新型コロナウイルスに関するアンケート」調査, p.1, 入手先 <[https://img03.en25.com/Web/TSR/%7B39ca8863-bf33-44ce-8ec6-1c4817983b53%7D\\_20210318\\_TSRsurvey\\_CoronaVirus.pdf](https://img03.en25.com/Web/TSR/%7B39ca8863-bf33-44ce-8ec6-1c4817983b53%7D_20210318_TSRsurvey_CoronaVirus.pdf)> (参照2021-06-20).
- [3] 日本銀行: 業況判断の推移, 短観(概要) - 2019年3月-, p.8, 入手先 <<https://www.boj.or.jp/statistics/tk/gaiyo/2016/tka1903.pdf>> (参照2021-07-20).
- [4] 内閣府: 3. 経済活動別の就業者数・雇用者数, 労働時間数(1) 就業者数, 2019年度(令和元年度) 国民経済計算年次推計(2015年基準・2008SNA), 入手先 <[https://www.esri.cao.go.jp/jp/sna/data/data\\_list/kakuhou/files/2019/2019\\_kaku\\_top.html](https://www.esri.cao.go.jp/jp/sna/data/data_list/kakuhou/files/2019/2019_kaku_top.html)> (参照2021-07-20).
- [5] 中小企業庁: 中小企業白書 2019年版, p.50, 入手先 <[https://www.chusho.meti.go.jp/pamflet/hakusyo/2019/PDF/chusho/03Hakusyo\\_part1\\_chap4\\_web.pdf](https://www.chusho.meti.go.jp/pamflet/hakusyo/2019/PDF/chusho/03Hakusyo_part1_chap4_web.pdf)> (参照2021-07-20).
- [6] 日本銀行: 業況判断の推移, 短観(概要) - 2020年3月-, p.10, 入手先 <<https://www.boj.or.jp/statistics/tk/gaiyo/2016/tka2003.pdf>> (参照2021-07-20).
- [7] 経済産業省, 厚生労働省, 文部科学省: 令和元年度 ものづくり基盤技術の振興施策, p.9, 入手先 <<https://www.meti.go.jp/press/2020/05/20200529001/20200529001-3.pdf>> (参照2021-07-20).
- [8] 経済産業省, 厚生労働省, 文部科学省: 令和2年度 ものづくり基盤技術の振興施策, p.7, 入手先 <[https://www.meti.go.jp/report/whitepaper/mono/2021/pdf/honbun\\_1\\_1\\_1.pdf](https://www.meti.go.jp/report/whitepaper/mono/2021/pdf/honbun_1_1_1.pdf)> (参照2021-07-20).
- [9] 日本銀行: 業況判断の推移, 短観(概要) - 2021年3月-, p.10, 入手先 <<https://www.boj.or.jp/statistics/tk/gaiyo/2021/tka2103.pdf>> (参照2021-07-20).
- [10] 労働政策研究・研修機構: プレスリリース ものづくり産業におけるDX(デジタルトランスフォーメーション)に対応した人材の確保・育成や働き方に関する調査, p. 16, 入手先 <<https://www.jil.go.jp/press/documents/20210709.pdf>> (参照2021-07-20).
- [11] 中小企業庁: 中小企業白書 小規模企業白書 2021年版, p.49, 入手先 <[https://www.chusho.meti.go.jp/pamflet/hakusyo/2021/PDF/chusho/03Hakusyo\\_part1\\_chap1\\_web.pdf](https://www.chusho.meti.go.jp/pamflet/hakusyo/2021/PDF/chusho/03Hakusyo_part1_chap1_web.pdf)> (参照2021-07-20).
- [12] 東京商工リサーチ: 特別企画 第14回「新型コロナウイルスに関するアンケート」調査, p.1, 入手先 <<https://img03.en25.com/Web/TSR/%7B39ca8863-bf33-44ce-8ec6->

- 1c4817983b53%7D\_20210318\_TSRsurvey\_CoronaVirus.pdf〉 (参照2021-07-20).
- [13] 中小企業庁：中小企業白書 小規模企業白書 2021年版, p.165, 入手先  
〈[https://www.chusho.meti.go.jp/pamflet/hakusyo/2021/PDF/chusho/04Hakusyo\\_part2\\_chap2\\_web.pdf](https://www.chusho.meti.go.jp/pamflet/hakusyo/2021/PDF/chusho/04Hakusyo_part2_chap2_web.pdf)〉 (参照2021-07-20).
- [14] 労働政策研究・研修機構：プレスリリースものづくり産業におけるDX（デジタルトランスフォーメーション）に対応した人材の確保・育成や働き方に関する調査, p.2, 入手先  
〈<https://www.jil.go.jp/press/documents/20210709.pdf>〉 (参照2021-07-20).
- [15] 中小企業庁：中小企業白書 小規模企業白書 2021年版, p.166, 入手先  
〈[https://www.chusho.meti.go.jp/pamflet/hakusyo/2021/PDF/chusho/04Hakusyo\\_part2\\_chap2\\_web.pdf](https://www.chusho.meti.go.jp/pamflet/hakusyo/2021/PDF/chusho/04Hakusyo_part2_chap2_web.pdf)〉 (参照2021-07-20).
- [16] 経済産業省, 厚生労働省, 文部科学省：令和2年度 ものづくり基盤技術の振興施策, p.131, 入手先  
〈<https://www.meti.go.jp/press/2021/05/20210528002/20210528002-3.pdf>〉 (参照2021-07-20).
- [17] 経済産業省, 厚生労働省, 文部科学省：平成30年度 ものづくり基盤技術の振興施策, p.17, 入手先  
〈[https://www.meti.go.jp/press/2019/06/20190611002/20190611002\\_02.pdf](https://www.meti.go.jp/press/2019/06/20190611002/20190611002_02.pdf)〉 (参照2021-07-20).
- [18] 日本インダストリアル・エンジニアリング協会：IE とは, 日本インダストリアル・エンジニアリング協会, 入手先  
〈[https://www.j-ie.com/about/about\\_ie/](https://www.j-ie.com/about/about_ie/)〉 (参照2021-06-15).
- [19] 経済産業省, 厚生労働省, 文部科学省：令和2年度 ものづくり基盤技術の振興施策, p.144, 入手先  
〈<https://www.meti.go.jp/press/2021/05/20210528002/20210528002-3.pdf>〉 (参照2021-07-20).
- [20] 経済産業省, 厚生労働省, 文部科学省：令和2年度 ものづくり基盤技術の振興施策, p.169, 入手先  
〈<https://www.meti.go.jp/press/2021/05/20210528002/20210528002-3.pdf>〉 (参照2021-07-20).
- [21] 荒川雅裕：次世代IE技術の発展と人材育成, 日本インダストリアル・エンジニアリング協会, IEレビュー, Vol.61, No.5, p.7 (2020) .
- [22] 人見勝人：現代IEの基礎概念, 日本機械学会, 第 67 巻, 544号, pp.692-699 (1964) .
- [23] 斎藤衛：IE手法(作業分析含む), 品質, Vol. 32, No. 3, p.62 (2002) .
- [24] 桑田秀夫：生産管理概論, pp.48-50, 日刊工業新聞社 (1990) .

## 第2章 インダストリアル・エンジニアリングにおける課題

### 2.1 インダストリアル・エンジニアリングの定義

日本インダストリアル・エンジニアリング協会（日本IE協会）によると、インダストリアル・エンジニアリング（IE）とは、価値とムダを顕在化させ、資源を最小化することによりその価値を最大限に引き出そうとする見方・考え方であり、それを実現する技術であると定義されている。また、仕事のやり方や時間の使い方を工夫することにより、豊かで実りある社会を築くことを狙いとしており、製造業に限らず、物流・農水産業・サービス・医療・金融・公共団体業務、さらには日常生活の領域など幅広く活用されている手法である[1]。また、同協会の連携団体であるアメリカIE協会（Institute of Industrial & Systems Engineers : IISE）では、IEとは、人・モノ・設備・エネルギーの総合されたシステムの設計・改善・確立に関する活動であり、そのシステムから得られる結果を明確にし、予測し、かつ評価するために工学的な解析・設計の原理や方法とともに、数学・物理学・社会科学の専門知識と技術とを利用することであると定義されている[2]。同じく、同協会の連携団体であるオーストラリアIE協会

（Institute of Industrial Engineers Australia : IIEA）では、IEについて、品質と人の満足が満たされるという条件下において、利用可能なリソースから最高の成果を得るために、数学・物理・生理学・社会・コンピューターサイエンスの専門知識とスキルを利用して分析を行い、そこから得られる結果を特定・予測・評価し・成果を測定することであると定義されている[3]。

このように、IEとは、従来からの勘や経験に基づいた主観的な測定や予測、評価方法ではなく、様々な分野の知識とスキルに基づいた科学的アプローチを用いて客観的に事象をとらえ、そこに含まれている課題を解決（改善）していくための手法として定義されている。また、このようなIE活動を行うエンジニアのことはインダストリアル・エンジニア（IEr）と呼ばれている。

### 2.2 インダストリアル・エンジニアリングの発展過程

IEの発展の歴史について紐解くと、IEへのニーズが高まったのは1875年以降のアメリカにおけるTaylor（F. W. Taylor, 1856~1915）とGilbreth（F. B. Gilbreth, 1868~1924）の活動がその起源とされている[4]。その当時の時代背景としては、それまでの資本拡大による経済成長が厳しくなり、企業間の競争が激しさを増すなかで、経営内部の合理化・充実化を目指す以外に企業にとっての対処方法がなくなり、その結果として「能率増進運動」と呼ばれる合理化運動が進められた。そのような状況のなかで、従来の経験と勘のみに頼っていた経営に対して、「科学的」な視点で最初にアプローチしたのがTaylorであるとされている[5]。この「科学的」とは、労働者や経営者による主観ではなく、客観性がある基準を用いた評価方法を指しており、Taylorは時間を客観性のある道具として用いることにより客観的な標準時間を定め、その結果として、「一日の公平な仕事量」が決められると考えた[6][7]。

一方、Gilbrethはレンガ積み職人の無数にある作業方法のなかで、「最良の方法・一番よい作業方法は世界に1つしかない」と考え、作業時の動作に関する解析と改善について追求した。職人の作業を観察

すると、各自がそれぞれ異なる方法で作業を行っており、また同じ動作を繰り返しているわけではなく、作業順序も異なっていることに着目し、動作研究として一連の動作を分析し、そこから最良の作業方法を追求する研究を行った[8][9].

以上のように、Taylorは時間研究として「一日の公正な仕事量」を提唱し、Gilbrethは動作研究として「唯一最善の方法」を提唱した。IEの発展の歴史については、その考え方と適応範囲によりいくつかの見方ができるが、時間研究と動作研究を起源とするIEの発展過程について、上野は図 2-1のとおり示している[10]。この図を見ると、Taylorの時間研究はその後、標準時間前決め法（Predetermined Time Standard System : PTS）[11]へと進化し、Gilbrethの動作研究は作業簡素化計画へと進化を遂げたが、時間研究と動作研究は二者択一の問題ではなく、密接不可分という認識を持たれている[12]。また、これらを補足し、総合的展開を目指したのがMaynardとその協力者たちで、彼らはこれをMethods EngineeringやMotion and Time Study, Work Studyと呼んだ[13].

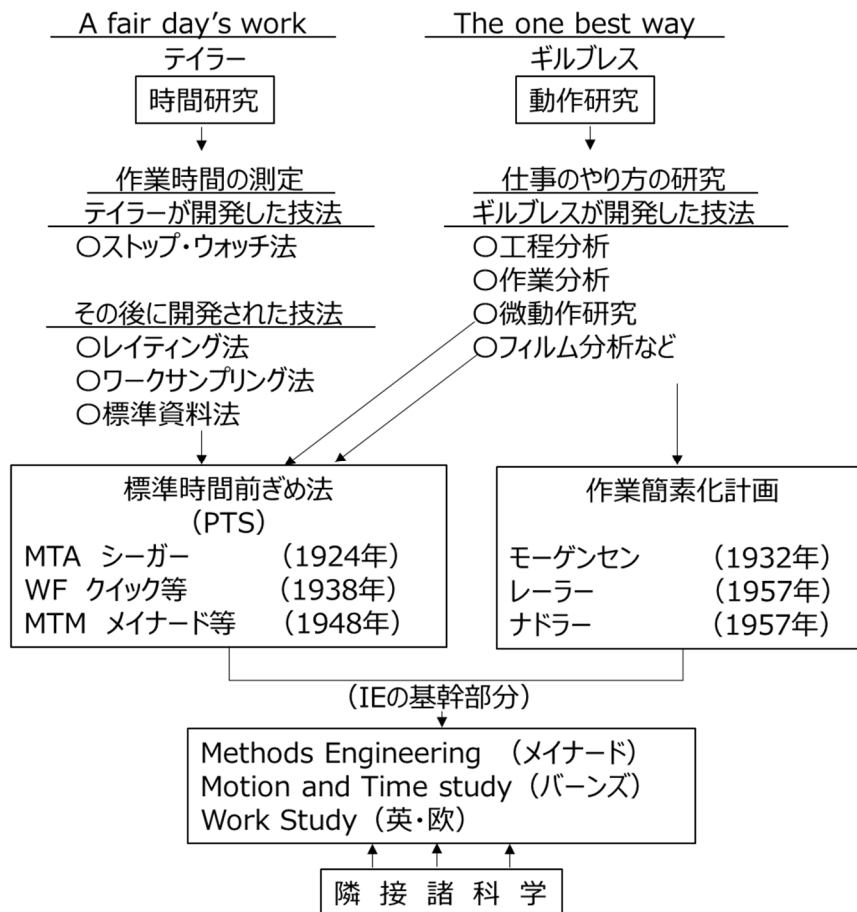


図 2-1 動作時間研究の発展過程

また、1960年代初期のIEにおけるMethods Engineering（作業方法技術）の主な機能と手法の関係性について、松田らにより図 2-2が示されている[14]。一部については近年の技術や用語に置き換わり、

分析には統計的手法も利用されるケースもでてきたが[15]，最適な作業方法の研究・開発を行い，これを標準化し，標準作業時間を定め，作業者の指導・訓練を行いながら改善活動を進めていくという点，また，そこには人間工学としての要素も重要であるという点に，当時も現在も変わりはない。

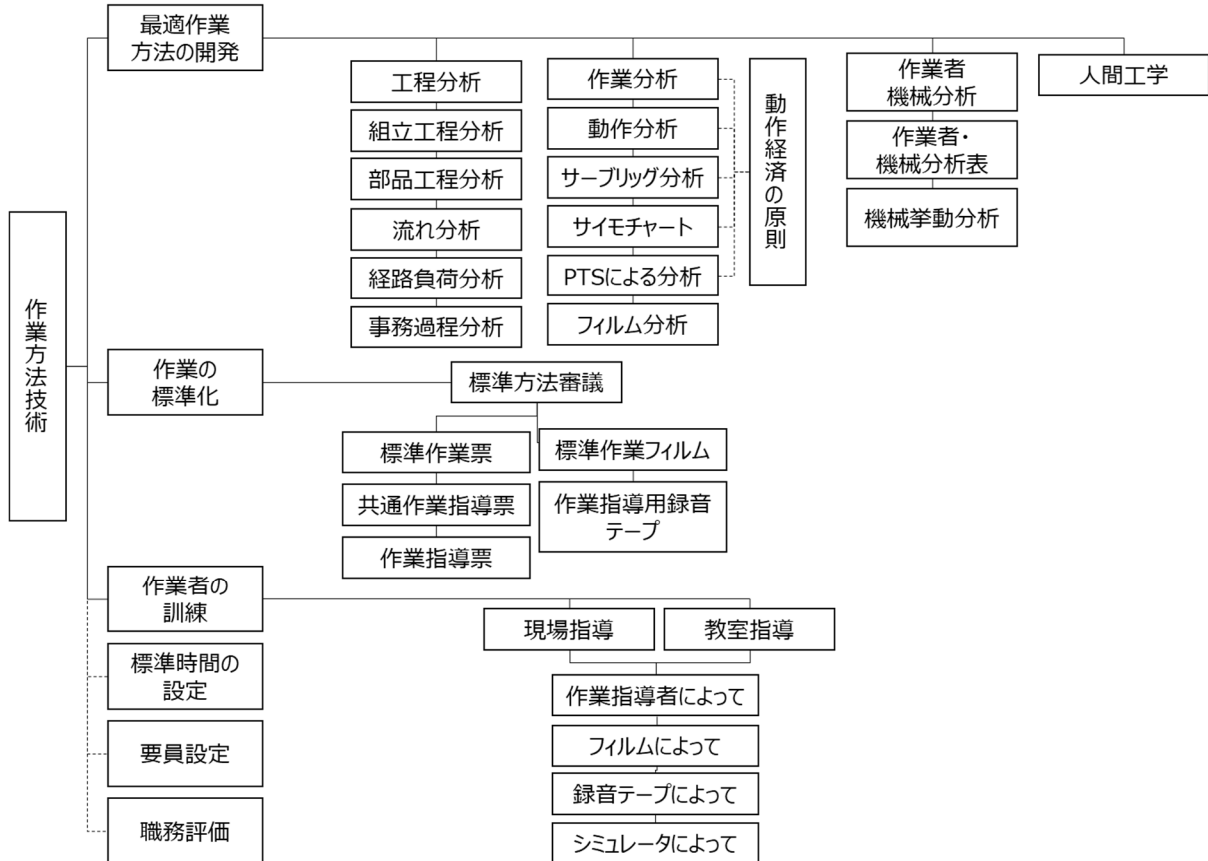


図 2-2 作業方法技術の機能と手法

その後，IEはその範囲を拡大し，経営科学／オペレーションズ・リサーチを中心とする数理的計画法や，情報通信技術，システム・エンジニアリングを包摂しながら経営工学として発展してきた[16][17]。このように発展してきたIEであるが，作業の効率化に対して科学的にアプローチしてきた「時間研究」や「動作研究」は，現在に至っても製造現場における改善のための基本として利用され続けている。

### 2.3 インダストリアル・エンジニアリングにおける分析的アプローチ

IEの概念の見方については，「伝統的IE」，「近代的IE」，「狭義のIE」，「広義のIE」があることは1章で述べた。本研究では，製造現場における労働生産性の向上を目指しているため，それを目的とした「伝統的」且つ「狭義」のIEにおける分析的アプローチについて，主要ないくつかの分析手法を紹介する。

- ①工程分析：工程の概要をつかみ，モノや人を軸に分析し，改善への手がかりを考える[18][19].
- ②稼働分析：人や機械がどのような要素に時間をかけているかを明らかにする[20][21].
- ③動作研究：作業方法を人間の身体部分の動きの面から分析して，最良の方法を設定する[22][23].
- ④時間研究：仕事を要素に分割し，時間を尺度としてそれらを測定・改善・標準化する[24][25].
- ⑤連合作業分析：人と機械または，人同士の組合せ作業を時間的側面から分析し，効果的な方法へ組み替える[26][27].
- ⑥ライン・バランス分析：ライン内各工程の所要時間の差をなくしスムーズな流れを設計する[28].
- ⑦マテリアル・ハンドリング：物の移動，積み下ろし，取り付けなどの作業を改善する[29].
- ⑧プラント・レイアウト：建屋，機械設備，原材料，作業員などの配置を最適に計画する[30].
- ⑨事務工程分析：事務の業務の流れを工場の生産活動に見立て，測定・評価して改善する[31].

なお，上記のような分析手法のなかで，筆者の研究では，時間研究と動作研究を対象範囲としているため，それぞれの概要について以下に示す。

### 2.3.1 時間研究

時間研究とは，作業をある要素に分割し，時間を尺度としてそれらの要素を測定・評価して改善する手法とされている[32][33]. 杉原らのIE手法活用についてのアンケート調査に対する回答結果[34]によると，時間研究の用途としては以下の項目があげられており，作業現場における改善活動に幅広く用いられていることが分かる。

- ① 標準作業，標準作業時間を設定するため
- ② ムダを省き作業能率を向上するため
- ③ 稼働率を向上するため
- ④ 機械化，自動化を推進するため
- ⑤ 段取り作業を改善するため
- ⑥ 省人化を図るため
- ⑦ 工程の設計，作業編成をするため

このように時間研究の活用の用途は多岐にわたるが，人による作業や設備の稼働に含まれるムダな要素を発見し，その排除や改善を行うことと，標準作業や標準作業時間などの標準化を行うことが，主な用途とされている。筆者の経験上，前者は，作業員個人の勘や経験に基づいて作業がなされている製造現場の改善を行う場合に用いられ，後者はこれらの改善が実施されている製造現場における標準化の設定と維持に用いられる場合が比較的多いという印象である。しかし，標準化が行われた製造現場においても，改善の要素が潜んでいる場合や，標準作業から逸脱した作業手順で作業を行う作業員も存在しており，改善と標準化は一体で繰り返される必要があるものと考えられる。

時間研究の進め方については，はじめに改善の目的と対象となる作業を定め，対象の作業を構成する要素として，目的に応じて区分される一連の作業を指す細かい要素作業に分割する。次に予備

調査として観測の計画や分割した要素作業に関する適否、観測の回数などを確認する。そのうえで本観測として対象の作業について要素作業ごとに時間を測定し、その記録を残す。最後にこの結果を用いて分析を行い、改善や標準化へ繋げる。

ここで、測定から分析、標準作業時間設定までの具体的な流れについて示す。時間研究における測定方法はストップ・ウォッチを用いて作業時間を測定するストップ・ウォッチ法と、録画したVTRを用いて測定を行うVTR法がある。また、近年はITツールの普及により、動画撮影機器を使い、録画した動画に対してパソコン上で分析するソフトウェア[35][36]も開発され、活用している企業も少なくない[37]。これらを利用して、対象の作業について要素作業ごとに時間を測定し、記録として残す。

次に、時間研究における分析の方法について示す。作業改善を目的とした分析においては、はじめに作業時間の構成内容を分析し、不要な動作や時間のかかる要素作業、時間にバラツキが見られる要素作業などを把握する。具体的には、以下のような点について検討が行われる[38]。

- 要素作業や動作の必要性から、不要な動作や作業を遅らせる動作がないか、必要な動作だけで再編成すると、どれだけの時間短縮が期待できるか
- 要素作業の時間値から、どの要素作業に時間が多くかかっているか、ネックの作業は何か
- 要素作業の時間のバラツキから、その原因は何か、時間がかかる場合とかからない場合の違いは何か
- 作業間での作業時間の違いから、どの要素作業の動作や方法の差が原因か、熟練者と未熟練者との違いは何か、治具などの利用で簡単にできる方法はないか

分析の際にこれらの検討を行ったうえで、ムダとされる部分を対象に改善が行われる。

また、標準作業や標準作業時間の設定を目的とした分析の場合は、上記の分析のための検討に加え、作業における最小時間値が示す作業条件と作業方法を明らかにし、測定された時間を参考に標準化を行う。標準作業時間については、その仕事に適正を持つ熟練した作業者が、決められた条件のもとで1単位の仕事を完成する際にかかる時間であり、標準作業方法を時間という単位で計数化したものである。決められた条件とは、良好な作業環境、所定の作業条件、必要な余裕などを指す[39]。標準作業時間の構造は図 2-3に示すとおり、実際の作業に要する時間と準備に要する時間に対して、それぞれレーティングと余裕時間を加味して構成される。レーティングとは観測時に観測者が作業ペースを評価し、観測時間に対して係数を用いて補正をかけることにより、作業ペースに基づく時間値（正味時間）に修正を行うためのものである。レーティングによる正味時間は式(1)で求められる。なお、このレーティング係数は、正常な作業ペースを1とした場合の観測対象者の作業ペースの比率を表し、正常ペースより速ければ1より大きい値、遅ければ1より小さい値を用いて補正をかける。また、余裕時間とはサンプリングの結果から求められる作業余裕、職場余裕と理論的な基準によってきめられる人的余裕、疲労余裕で構成される。

$$\text{正味時間} = \text{観測時間} \times \text{レーティング係数} \quad (1)$$



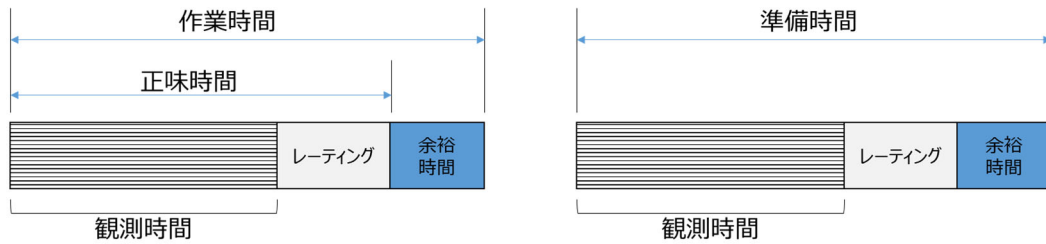


図 2-3 標準作業時間の構成

このようにして求められた標準作業や標準作業時間については、以下のような目的で活用される[40].

- ・ 標準作業時間設定による生産性の向上

製造現場では標準作業時間を決めることが重要とされる場合がある。たとえば、熟練者の作業時間を標準作業時間として設定することにより、作業者が作業時間を意識し、作業ペースが速くなり、その結果として生産性が向上する場合がある。

- ・ 現場の日常管理

標準作業時間を基準として、日常の実際の作業時間と基準との乖離を追及し、時間的な差異が見られた場合には、その原因を調べ、改善をはかる。

- ・ 工数管理への展開

標準作業時間と日報を利用して、稼働率や作業能率を計算し、総合的な生産性を測定するための情報として利用する。

このように、時間研究の結果は、標準作業や標準作業時間として実際の製造現場へ展開され、改善活動に利用される。しかし、製造現場での作業状況を日常的に把握し、日頃の実態に基づいて改善活動を行うためには、製造現場の見回りや作業日報に基づく分析など、人手に頼る方法が多く、IErにとって負担となっているものと考えられる。

### 2.3.2 動作研究

動作研究とは、作業のなかに含まれる人間の身体動作と目の動きを分析し、非効率な動作の排除、動作の組み換えなどにより、最良の方法を設定するための研究とされている[41]。作業とはいくつかの動作の集合体であるため、作業を改善するためには、その構成要素であるそれぞれの動作に対する改善が必要であるという考え方である。動作研究においては、人間の動作そのものの研究という側面とともに、その方法研究の根底に流れている「モーション・マインド」を身に付けることが重要であるとされている。この「モーション・マインド」のことを、動作の違いに気づき、よりよい動作に改善できる心構えであるとして、具体的に以下の4点があげられている[42].

- ①動作の違いに気づく
- ②動作を正しく分析し、対象の動作にひそむ非合理性を見出す
- ③動作の違いを明らかにし、よい動作を判断する

## ④よい動作を設計する

また、動作研究の種類としては、主に以下の2種類があげられる。

## (1)両手作業分析

両手作業分析とは、作業者の両手の動作のプロセスについて時間的に関連性を持たせながら分析、改善する手法とされている[43]。作業者の両手の動作の内容や動作の順序をとらえ、「作業」「移動」「保持」「手持ち」の4つを用いて分類・分析し、動作経済などの原則に照らし合わせて問題点の抽出を行う。「作業」とは手が作業を行っている状態を指し、「移動」とは、手を伸ばしたり物を運んでいる状態を指す。「保持」とは、作業のために物を持っている状態、「手持ち」とは、手が何もしていない状態を指す。そのうえで、 unnecessaryな動作を減らしたり、動作の距離を短くできるようにしたり、動作の向きを円滑で滑らかな方向にしたりなど、動作に対する改善を行うための分析方法とされている[44]。動作経済の原則とは、Gilbrethがレンガ積み作業の動作研究を行い、最も疲労の少ない、経済的な動作を実現するための法則として提起したものであり、その後の研究者により補正、追加され現代に至っている原則である[45]。

## (2)微動作分析

微動作分析とは、作業者のより細かな手足や目の動きを、サーブリック記号と呼ばれる人間の動作を構成する18の基本要素を用いて分類・分析を行い、ムダ・ムラ・ムリを取り除き、動作改善に繋げるための手法とされている[46]。18の要素は動作の基本をなす第1類と、動作を遅れさせる第2類、仕事が進んでいない状態である第3類に分けられる(表 2-1)。

表 2-1 サブリックの構成要素

第1類	第2類	第3類
つかむ	探す	保持する
から手	見つける	避けえぬ遅れ
運ぶ	運ぶ	避けえる遅れ
放す	位置決め	休む
調べる	前置き	
組合せ	考える	
分解する		
使用する		

第1類のなかでも価値を生む要素は「組合せ」「分解する」「使用する」だけであり、それ以外の要素はすべて改善対象とされ、排除を第1に考える[47]。また、価値を生む3要素についても、より楽で、短時間で仕事ができるように考え、改善が行われる[48]。

動作研究の進め方については、はじめにその目的を定める。動作研究の目的としては、動作の中に潜むムダやムリ、ムラを抽出し、効率的な動作へ改善をはかることと、作業仲間や同一作業者の動作のバラツキを分析し、効率的な作業動作の設計を行うこととされており、この目的を明確化したうえで、観測を行うことが大切であるとされている[49]。次に、対象作業を決め、関係者の理解を得た上で現場での作業を観測し、分析を行う。

現場での観測方法については目視観測とVTRを用いた方法が一般的であるが、VTRを活用することにより、速い動作に対する詳細分析や、複数作業の比較分析、熟練者の技能継承などに利用できるというメリットがあげられる[50]。観測の結果については、両手作業分析の場合は「両手作業分析表」などに一連の作業における両手それぞれの動作の内容と、4種類の動作の分類結果が時系列に記録され、微動作分析の場合は、「微動作分析表」などに動作内容とサブリーグ記号が記録される[51][52]。

次に、動作研究に関する分析について、両手作業分析と微動作分析ではそれぞれの分析における視点が異なる。それぞれの分析の視点について示す[53]。

両手作業分析については、主に以下の視点で分析を行う。

- 「移動」「保持」「手待ち」などの付加価値を生まない動作の削減を検討する
- 両手のバランスから、片手が継続して「保持」「手待ち」の状態であれば、両手の左右対象のバランスがとれるよう検討する
- 手作業のウェイトができるだけ少なくなるよう、治具化、機械化、自動化の順で検討する

また、微動作分析については、以下の視点で分析が進められていく。

- サブリーグの構成要素の第1類について、排除できるような要素や動作の順序を変えることにより簡素化できる動作を見つける
- 同じく、第2類について、動作の排除や作業域内の部品や工具の置き方の工夫、治具の考案などを検討する
- 第3類については、排除することを検討する

このように、動作研究においては両手作業分析における「移動」「保持」「手待ち」などの付加価値を生まない動作や、微動作分析における第1類から第3類の中の排除できる動作を抽出し、改善をはかることを狙いとしている。しかし、先にあげた時間研究同様、製造現場での作業時の動作を日常的に把握し、分析を行うことは容易ではない。また、筆者の経験では、両手分析の「作業」や微動作分析の「第1類」の動作においても、熟練技能者と非熟練者の間で動作の相違が多く見られ、人材育成や技能伝承の視点でこれらの把握や改善についても重要であると考えられる。

#### 2.4 インダストリアル・エンジニアリングにおける課題

日本国内へIEが導入されてから長い年月が経つが、藤井らによると大企業においても、IEについて本心に理解し、現場に定着させている企業は少ないとされている。また、センスあるIErを育成するためには、長年の期間が必要であるとし、中小企業に於いては、IEを使いこなせる人材を育てることはもとよ

り、IE的改善手法を活用して改善を行うことは困難なことであると述べている[54]。IErのスキル面の要件としては、基礎的な科学・工学・IE手法の知識以外にも、理解力や洞察力、改善や企画に対する積極的な意欲、他人とのコミュニケーション力や指導力などの総合的な能力があげられるが[55]、特に中小企業にとって、このような能力を持った人材の採用や育成は容易ではないものと考えられる。

また、1990年代のバブル経済崩壊や近年のデフレによる価格破壊にともない、各メーカーは生産拠点を海外に移転し、また価格の安い海外製品との価格競争により、国内の中小製造業は大きな打撃を受けた。その結果、「失われた10年」と言われる期間で、モノづくり現場の熟練技能者とベテラン技術者は大きく減少したとされる。佐藤は、「工場建設・改良計画のできる人」、「加工技術を開発し独自の設備を作ることのできる人」、「治工具設計し最適な作業プロセスを指示できる人」、「作業改善、小集団活動を指導、牽引できる人」、「IT技術を駆使して生産計画／管理のできる人」、「信頼性／品質向上の技術に長けた人」、「モノづくりプロセスを改革できる人」を減少した人材としてあげているが[56]、これらのスキルはIE活動を行うIErにも共通する部分があり、熟練技能者とともIErも大きく減少したものと考えられる。

また、ここ数年のインダストリー4.0やデジタルトランスフォーメーション (DX) , Society5.0におけるモノづくりのデジタル化にともない、IEに対してもデジタル技術の利用が注目を集めている[57]。インダストリー4.0で注目されたセンサーやIoT(Internet of Things)、ビッグデータ、クラウド技術を利用した分析により、これまでブラックボックスだった生産プロセスの状態を把握することができようになり、記述的分析(何が起こったのか)、予測的分析(何が起こるべきか)、規範的分析(どうなっほしいか)を分析することが可能となった[58]。また、世界的に加速しているDXは、人々の生活をあらゆる面でより良い方向に変化させることを目的としており、IEに関してもデジタル技術を応用することにより高度化し、インダストリアル領域に限らずさまざまな領域に拡張することで新たな価値を生み出し、持続可能な社会の実現に貢献することが期待され[59]、Society5.0時代のモノづくりは次世代IEの目指す姿であるとされている[60]。つまり、これからのモノづくりにおいて、デジタル技術を活用して改善活動を効率的に進めていくためには、IEの原理原則に基づく着眼点を持ったIErが、ものの流れ、人による作業、設備の稼働、製品の出来栄えのなかでボトルネックとなるポイントを絞り込み、その仮説に基づいて、デジタル化されたデータの分析を繰り返しながら問題の本質を追及していくことが必要とされる。したがって、デジタル技術を使いこなすことができるようなIErの育成を進めていくことが、これからのものづくり企業の成長に不可欠である[61]。しかし、中小企業を中心とした製造業においては、工場内にIT系の技術者が存在しない場合が多いため、デジタル技術の活用に向けた適切な対応案を出せない企業が存在することが指摘されており、課題となっている[62]。

このようにIEの分野においては、モノづくりの現場を深く理解したIEr自体が不足するという課題と合わせて、これからのIErの育成においては、従来からのIErに必要とされるスキルに加えデジタル技術を活用するためのITスキルが必要となる。しかし、中小企業を中心とした製造業にとっては、そのような人材の採用や育成は容易ではないという課題が存在しており、これらの課題を解決することが、これか

らのIEを利用した労働生産性向上による企業の発展にとっては重要であるものと考えられる。

## 2.5 深層学習を用いたインダストリアル・エンジニアリングにおける課題解決

先に述べたIEにおける課題に対して、大手企業においてはIEとITの両方のスキルを持つ人材の育成や確保が比較的可能であると考えられるが、多くの中小企業においてはこのような人材の育成や確保は容易ではない。

また、労働生産性の向上を目的として、デジタル技術を活用していくうえでは、設備や機器に対するデジタル化だけではなく、作業員（人）に着目したデジタル技術の活用が重要である。このことについて井上は、モノ同士が繋がるIoTが脚光を浴び、様々な企業が積極的に取り組もうとしているなかで、現場で働いている作業員は依然として紙のマニュアルを参照しながら作業を実施し、作業日報や機器の計測結果も紙の報告書に記載されており、作業員がどこで、どのような状態であるかを管理者が常時監視することは難しいと述べている[63]。つまり、工場の設備や機器のデジタル化は進むものの、作業の多くはアナログプロセスであり、人に対するデジタル技術の適用が進んでいないことを指摘している。また、人に対するデジタル技術の適用が容易ではない理由として、日本の現場はこれまで経験・勘・度胸が支配している傾向が強いためであると述べている[64]。このようなアナログプロセスを中心とした製造現場におけるIEに対して、デジタル技術を活用していくためには、単なるセンシングやデータ解析に関する技術や知識だけではなく、製造現場の状況や作業内容を熟知した人材が、その中心的役割を果たすことが重要であると考えられる。

そこで本研究では、製造現場で働く作業員の労働生産性の向上を目的として、ITに関する高度な専門的知識を必要とせず、IErによる改善活動を支援することができる仕組みについて、深層学習技術を用いる方法を研究対象として取りあげた。深層学習技術を利用することにより、従来人間が自然に行うようなタスクをコンピュータに学習させ、実行させることが可能となる。その学習の際に使用する特徴量についても、深層学習ではデータから自動的に特徴量を抽出して学習する仕組みとなっているため、高度なIT技術や統計に関する専門知識を用いた特徴量設計などを必要としない。これにより、ITに関する専門知識を持つ人材が不足している状況においても、深層学習技術がIE改善活動におけるタスクの一部を担当するという形で、効率的に改善活動や人材育成を進めることが可能となるものと考えられる。

## 参考文献

- [1] 日本インダストリアル・エンジニアリング協会：IEとは，日本インダストリアル・エンジニアリング協会，入手先〈[https://www.j-ie.com/about/about\\_ie/](https://www.j-ie.com/about/about_ie/)〉（参照2021-06-15）。
- [2] Institute of Industrial & Systems Engineers : ORIGINS OF IISE, 入手先〈<https://www.iise.org/details.aspx?id=295>〉（参照2021-06-15）。
- [3] Institute of Industrial Engineers Australia : Institute of Industrial Engineers Australia,入手先〈<https://www.engineersaustralia.org.au/Communities-And-Groups/Technical-Societies/Institute-Of-Industrial-EngineersFor>〉（参照2021-06-15）。
- [4] 中村茂弘：現場力を高める簡単IE実践マニュアル，p.25，日刊工業新聞社（2011）。
- [5] 松田武彦，和田栄治，春日井博：現代の経営学4．インダストリアル・エンジニアリング，pp.30-31，中央公論社（1962）。
- [6] 上野一郎：マネジメント思想の発展系譜 テイラーから現代まで，p.30，日本能率協会（1977）。
- [7] 中村茂弘：現場力を高める簡単IE実践マニュアル，p.33，日刊工業新聞社（2011）。
- [8] 上野一郎：マネジメント思想の発展系譜 テイラーから現代まで，pp. 82-83，日本能率協会（1977）。
- [9] 中村茂弘：現場力を高める簡単IE実践マニュアル，pp.35-36，日刊工業新聞社（2011）。
- [10] 上野一郎：マネジメント思想の発展系譜 テイラーから現代まで，p.98，日本能率協会（1977）。
- [11] 梅村守，市川博：動作・時間研究 サブリック分析およびPTS分析を中心として，理学療法のための運動生理，8巻3号，pp.159-164（1993）。
- [12] 上野一郎：マネジメント思想の発展系譜 テイラーから現代まで，p. 97，日本能率協会（1977）。
- [13] 桑田秀夫：生産管理概論，p.50，日刊工業新聞社（1990）。
- [14] 松田武彦，和田栄治，春日井博：現代の経営学4．インダストリアル・エンジニアリング，p.69，中央公論社（1962）。
- [15] Jeffrey E. Fernandez, Michael Halpern, Anand Subramanian, Anil R. Kumar, Brandy F. Ware : Statistical Methods in Time Studies:Industrial Engineering And Health Care, Proceedings 13th on International Conference on Industrial Engineering Theory, Applications and Practice, pp.618-624 (2008).
- [16] 人見勝人：現代IEの基礎概念，日本機械学会誌，67巻，544号，pp.692-699（1964）。
- [17] 辻正重，中邨良樹，井上昌治，田部勉，他：経営工学総論，p.32，ミネルヴァ書房（2010）。
- [18] 杉原寛，嶋本康夫，中村純一，藤山紘一郎，杉浦亘，宮原正典：IE 7つ道具，p.29，日刊工業新聞社（1993）。
- [19] 桑田秀夫：生産管理概論，p.57，日刊工業新聞社（1990）。
- [20] 杉原寛，嶋本康夫，中村純一，藤山紘一郎，杉浦亘，宮原正典：IE 7つ道具，p.57，日刊工業新聞社（1993）。

- 聞社（1993）。
- [21] 桑田秀夫：生産管理概論，p.90，日刊工業新聞社（1990）。
- [22] 杉原寛，嶋本康夫，中村純一，藤山紘一郎，杉浦亘，宮原正典：IE 7 つ道具，p.83，日刊工業新聞社（1993）。
- [23] 桑田秀夫：生産管理概論，p.69，日刊工業新聞社（1990）。
- [24] 杉原寛，嶋本康夫，中村純一，藤山紘一郎，杉浦亘，宮原正典：IE 7 つ道具，p.105，日刊工業新聞社（1993）。
- [25] 桑田秀夫：生産管理概論，p.81，日刊工業新聞社（1990）。
- [26] 杉原寛，嶋本康夫，中村純一，藤山紘一郎，杉浦亘，宮原正典：IE 7 つ道具，p.127，日刊工業新聞社（1993）。
- [27] 桑田秀夫：生産管理概論，p.88，日刊工業新聞社（1990）。
- [28] 杉原寛，嶋本康夫，中村純一，藤山紘一郎，杉浦亘，宮原正典：IE 7 つ道具，p.139，日刊工業新聞社（1993）。
- [29] 杉原寛，嶋本康夫，中村純一，藤山紘一郎，杉浦亘，宮原正典：IE 7 つ道具，p.151，日刊工業新聞社（1993）。
- [30] 杉原寛，嶋本康夫，中村純一，藤山紘一郎，杉浦亘，宮原正典：IE 7 つ道具，p.171，日刊工業新聞社（1993）。
- [31] 杉原寛，嶋本康夫，中村純一，藤山紘一郎，杉浦亘，宮原正典：IE 7 つ道具，pp.187-188，日刊工業新聞社（1993）。
- [32] 川瀬 武志：IEにおける時間，日本機械学会誌，96巻，890号，pp.63-66（1993）。
- [33] 赤木宏匡，阿津沢潔，飯田肇，石井和克，他：ものづくりに役立つ経営工学の事典，p.242，朝倉書店（2013）。
- [34] 杉原寛，嶋本康夫，中村純一，藤山紘一郎，杉浦亘，宮原正典：IE 7 つ道具，pp.105-106，日刊工業新聞社（1993）。
- [35] 日本生工技研：製品情報 TimePrism，入手先〈<https://www.jiet.co.jp/timeprism.php>〉（参照2021-06-15）。
- [36] ブロードリーフ：OTRS，入手先〈<https://www.otrs.jp/>〉（参照2021-06-15）。
- [37] 赤木宏匡，阿津沢潔，飯田肇，石井和克，他：ものづくりに役立つ経営工学の事典，p.242，朝倉書店（2013）。
- [38] 杉原寛，嶋本康夫，中村純一，藤山紘一郎，杉浦亘，宮原正典：IE 7 つ道具，pp.117-118，日刊工業新聞社（1993）。
- [39] 赤木宏匡，阿津沢潔，飯田肇，石井和克，他：ものづくりに役立つ経営工学の事典，p.22，朝倉書店（2013）。
- [40] 杉原寛，嶋本康夫，中村純一，藤山紘一郎，杉浦亘，宮原正典：IE 7 つ道具，pp.121-122，日刊

工業新聞社（1993）。

- [41] 赤木宏匡, 阿津沢潔, 飯田肇, 石井和克, 他: ものづくりに役立つ経営工学の事典, p.244, 朝倉書店 (2013)。
- [42] 赤木宏匡, 阿津沢潔, 飯田肇, 石井和克, 他: ものづくりに役立つ経営工学の事典, p.244, 朝倉書店 (2013)。
- [43] 桑田秀夫: 生産管理概論, p.70, 日刊工業新聞社 (1990)。
- [44] 杉原寛, 嶋本康夫, 中村純一, 藤山紘一郎, 杉浦亘, 宮原正典: IE 7 つ道具, pp.85-88, 日刊工業新聞社 (1993)。
- [45] 松田武彦, 和田栄治, 春日井博: 現代の経営学 4. インダストリアル・エンジニアリング, pp.56-59, 中央公論社 (1962)。
- [46] 梅村守, 市川博: 動作・時間研究 サーブブリック分析およびPTS分析を中心として, 理学療法のための運動生理, 8巻, 3号, pp.159-164 (1993)。
- [47] 杉原寛, 嶋本康夫, 中村純一, 藤山紘一郎, 杉浦亘, 宮原正典: IE 7 つ道具, pp.90-91, 日刊工業新聞社 (1993)。
- [48] 桑田秀夫: 生産管理概論, p.74, 日刊工業新聞社 (1990)。
- [49] 杉原寛, 嶋本康夫, 中村純一, 藤山紘一郎, 杉浦亘, 宮原正典: IE 7 つ道具, pp.87-89, 日刊工業新聞社 (1993)。
- [50] 杉原寛, 嶋本康夫, 中村純一, 藤山紘一郎, 杉浦亘, 宮原正典: IE 7 つ道具, p.86, 日刊工業新聞社 (1993)。
- [51] 杉原寛, 嶋本康夫, 中村純一, 藤山紘一郎, 杉浦亘, 宮原正典: IE 7 つ道具, p.94, 日刊工業新聞社 (1993)。
- [52] 桑田秀夫: 生産管理概論, pp.70-75, 日刊工業新聞社 (1990)。
- [53] 杉原寛, 嶋本康夫, 中村純一, 藤山紘一郎, 杉浦亘, 宮原正典: IE 7 つ道具, p.95, 日刊工業新聞社 (1993)。
- [54] 藤井春雄: IE実践 (やさしくIE普及), 生産管理, 第3巻, pp.126-130 (1996)。
- [55] 松田武彦, 和田栄治, 春日井博: 現代の経営学 4. インダストリアル・エンジニアリング, pp.345-346, 中央公論社 (1962)。
- [56] 佐藤一男: モノづくりは人づくり / モノづくり武士道, 工学教育, 第54巻, pp.11-14 (2006)。
- [57] Roberta Assis Costa, Kelly Alonso Costa, Cecilia Toledo Hernández, Eliane da Silva Christo, Enrico Luigi Moreira Perocco : A Bibliometric Research of Industry 4.0 Opportunities in Industrial Engineering, Br. J. Ed., Tech. Soc., Vol.12, No.4, Apr.-Jun., pp.299-310 (2021).
- [58] Mirjana Misita, Vesna Spasojević Brkić, Dragan D. Milanović, Martina Perišić : Industrial Engineering Methods and Techniques in Industry 4.0, International Conference Industrial Engineering and Environmental Protection 2020 (IIZS 2020), pp. 36-41 (2020).



- [59] 日本インダストリアル・エンジニアリング協会：会長挨拶，日本インダストリアル・エンジニアリング協会，入手先〈<https://www.j-ie.com/about/aisatsu/>〉（参照2021-06-15）.
- [60] 鈴木教洋：デジタル技術で変わるモノづくり，日本インダストリアル・エンジニアリング協会，IEレビュー，Vol.60，No.2，p.2（2019）.
- [61] 吉川浄：IoTを活用したIEで超多品種少量生産の改善を加速，日本インダストリアル・エンジニアリング協会，IEレビュー，Vol.59，No.1，p.2（2018）.
- [62] 荒川雅裕：次世代IE技術の発展と人材育成，日本インダストリアル・エンジニアリング協会，IEレビュー，Vol.61，No.5，p.7（2020）.
- [63] 井上和佳：製造・物流業におけるIoT～AR・ウェアラブル技術の活用～，横断型基幹科学技術研究団体連合，11巻，1号，pp.27-31（2017）.
- [64] 井上和佳：製造・物流業におけるIoT～AR・ウェアラブル技術の活用～，横断型基幹科学技術研究団体連合，11巻，1号，pp.27-31（2017）.

## 第3章 関連研究および基礎理論

本章では、IEをはじめとした製造現場での改善活動に対して、デジタル技術を用いた関連研究に関する手法および考えられる課題と、また、本研究において利用した深層学習技術に関する基礎理論および、研究で用いた手法について述べる。

### 3.1 製造現場でのデジタル技術を利用した改善活動に関する関連研究

近年、製造現場でのIE活動においては、IoTやAIなどのデジタル技術の利用が進められており、作業者の作業監視や動作分析などに関する研究が多く報告されている。これらの研究の目的としては、作業ミスの検知、作業時の動作や時間の分析などがあげられ、製造現場の状況を収集するための機器としては、カメラを利用する方法やウェアラブルセンサを利用する方法が代表的である。また、収集した情報を用いて監視や分析を行う際の手法についてはその目的により多岐にわたっている。以下に関連研究に対してその目的ごとに大きく2つに分け、それぞれの研究における手法と、考えられる課題について述べる。

#### 3.1.1 作業ミスの検出

製造現場での人手による作業に対しては、作業手順の誤りや作業漏れなどのミス、または作業手順に対する恣意的または意図的な逸脱行為が想定される。そのため、日常の作業に対して、あらかじめ設定された標準作業に従って正しく適切に作業を行っていることを監視し、作業ミスや作業手順の逸脱を検出するための手法に関する研究が進められている。

##### (1) 深度カメラを用いたピッキング工程での作業ミス検知

宇田らは、物流事業における配送センターや自動車組立工場において、作業者が棚からの物品の取り出しを行うオフラインピッキング工程を対象として、ピッキングの作業ミスを検知するためのシステムについて研究を行った[1]。この研究では、ピッキングの際に作業者が指定した棚の間口から正しく物品を取り出したか否かを対象としており、手の追跡と手を入れた間口の検出の2つの処理で構成されている。

このシステムは、Microsoft Kinect (Kinect) を使用しているが、対象となるピッキング作業では台車などが障害となって身体の一部が遮られるという理由から、手の追跡方法についてはKinectが提供する人体の骨格を検出するスケルトン検出機能は使用せず、作業者が手に着用した作業グローブの色を検出条件として利用している。具体的な処理については、はじめにKinectが接続されたPC画面上の画像に対し作業グローブの領域を指定し、作業グローブの色のみを抽出するマスクを作成する。そしてKinectで撮影された画像情報を、色相、彩度、明度の成分に分割し、OpenCVの画像処理機能によりヒストグラムのバックプロジェクション（逆投影）を求め色の検出を行う。続いて作業グローブのマ

ク画像を用いて検出領域を絞り込み、作業グローブの色のヒストグラムと一致する場所の座標と領域の情報を得る。次に、手を入れた間口の検出については、Kinectの深度カメラを利用して、Depthセンサ面から作業グローブまでの距離と、Depthセンサ面から間口の開口部の四隅までの距離を求め、その位置関係を用いて認識している。

つまり、本手法は作業グローブの色情報から手を検出し、作業グローブと棚の間口の位置関係からピッキング作業を判断するという手法であるが、実際のピッキングの現場で利用する際には、2点の課題が考えられる。1点目の課題は、本システムに関する有用性である。これは、宇田らの実験においても述べられているが、物品や棚の周囲環境の色とグローブの色の組み合わせによって検出精度が異なる点である。研究では、作業グローブの色は、物品や棚の周囲環境に含まれず、照明による認識率の低下が少ないものを選ぶなど、対象となる現場の状況にあわせて適切なグローブの色を選択する必要性を述べているが、この点は作業現場への制約事項となる場合が考えられる。実験結果においても、一般的に使用されるケースが多いと考えられる白色の軍手については、検出が困難であることが述べられている。したがって、手の検出方法としては、作業グローブの色情報ではなく、作業グローブの特徴となる形状を用いて検出する方法が望ましいものと考えられる。また、宇田らは、Kinectが提供するスケルトン検出機能は、人の全身像が正面からとらえられているときに精度が高く、この研究のように人が斜め方向から撮影されたり、身体の一部が隠れたりする場合には制約があり、適用は困難であるとしているが、このスケルトン検出機能の代わりにKinectで撮影されたRGB画像と深層学習技術を用いることにより、身体の一部などの特定の物体だけを検出することは十分可能である。

2点目の課題は信頼性に関するもので、手と棚の間口の位置関係に基づいて動作を判断する手法における誤検出の問題である。実験において、手が横方向に並んだ隣の間口の前を通過する際の動作に対して、棚から物品を取り出した動作として誤って検出することが記されている。その対策として、作業員に対してピッキングの対象となる間口以外の間口を横切るような方向に手を動かさないように指示することにより回避できるとしている。しかし、システム上の制限や課題の解決を目的として、作業員の動作自体に制約を求めることについては、その制約が作業効率を低下させる可能性も考えられるため、慎重に判断する必要がある。したがって、このような誤検知をなくすためには、手の位置だけではなく、手に持っている物品も含めて検出できるような手法を用いる方が望ましいものと考えられる。

## (2) オプティカルフローを利用した締付作業での作業ミス検知

清水らの研究では、局所的な動きの情報であるオプティカルフローを用いて、組立工程における締め付け作業のポカヨケシステムについて開発を行った[2]。この研究については、製造現場に既に設置されている作業ミス検出用のセンサ類だけでは作業全体の品質を十分に保証することができないとして、既設のセンサ情報と組み合わせるためのシステムとして位置付けられている。

このシステムではカメラで撮影された作業時の映像データをいくつかの動作の単位に分割し、その

動作単位の時系列データである動作シーケンスを評価し、これが作業回数や作業順序など標準作業として定められた条件を満たさない場合に、作業ミスとして検出する。動作シーケンスを求める際には、各フレームに対するオプティカルフローの統計量である動きの大きさと、方向分布を表す方向ヒストグラムを用いて、それらを動作に対する特徴量として利用している。清水らによると、オプティカルフローで抽出したベクトル情報には動きそのものの本質的な情報が含まれており、作業者の肌の色や服装などの影響が少なく、作業者の変化や照明変化に対してロバストであるということが、利点として述べられている。実験においては、複数箇所へのネジの締め作業時の「締める」「戻す」の反復動作を、各フレームのオプティカルフローから抽出した動作特徴量のシーケンスより求め、トルクレンチの締め付け完了信号とあわせて「締め付け忘れ」のポカを検出している。具体的には、作業における動作単位について、始点と終点を結ぶ無駄のない動きであると仮定し、手の動作が必ず静止状態から速度を上げ、ある地点で速度が減少し静止するという過程をもっていることを利用して、動作単位に分割している。つまり、速度の大きさが減少から増加へ切り替わる時点や、零から増加へ変化する時点、減少から零へ変化する時点を作業の分割ポイントとして利用している。また、方向成分としては、オプティカルフローを8つの方向に対するベクトルとして求め、そのヒストグラムを用いている。その実験の結果、従来のポカヨケシステムでは検出することができなかった作業ミスを検出することができたとされている。

清水らが述べているとおり、オプティカルフローから抽出した動きのベクトル情報を用いて動作を把握する方法は服装や照明変化に対して堅牢であるが、想定外の動きに対しては敏感に反応するため、その影響を考慮した場合には、信頼性について課題が考えられる。これについて清水らは、身体や治工具などの微小な揺れについてはノイズとして除去できるとしているが、背景に含まれている検出対象物体以外の物体の動きや作業者のしぐさなど、微小な揺れ以上の動作が生じる場合において、この手法は脆弱である点が懸念される。また、実験に用いた作業動作については、比較的単純な締め付け作業を対象に「締める」→「戻す」の反復動作のパターンを動作シーケンスとして利用しているが、複雑な作業動作が行われるような製造現場で利用する際の有用性や課題については述べられていないため、適用対象となる作業については明らかにされていない。これらの課題を解決するための1つの手法として、各フレーム全体のオプティカルフローではなく、深層学習などにより検出した手や工具などの対象物体を中心とした周辺領域の、部分的なオプティカルフローを用いる方法が考えられる。または、深層学習で検出された物体の位置情報である座標データを、動作に対する時系列特徴量として利用して分類することにより、複雑な動作に対しても適用することが可能であると考えられる。この場合、清水らが指摘したような、作業者の肌の色や服装、作業者の変化や照明変化に対するロバスト性についても、深層学習の学習を重ねることによる汎化性能の向上により、解決することが可能であるものと考えられる。

### (3)姿勢推定ソフトを利用したセル生産ラインでの標準外動作の検知

西田らは、セル生産ラインを対象として、IEにおける改善活動を行うIErによる分析を支援するために、作業者の標準外動作を検知するシステムについて開発を行った[3]。この研究では、改善が必要とされる作業は「やりにくい作業」であり、これは通常の繰返し動作の中にその動作から外れた余分な標準外動作が含まれていることを前提として、これを検知することを目的としている。

このシステムでは、作業者の動作のデータ化、動作時間の計測、動作時間の分析の3つの機能で構成されている。作業者の動作データについては、作業現場に設置したカメラで撮影した動画に対して姿勢推定ソフトを用いて関節座標の時系列データを計測し、その座標データを微分した速度情報を用いている。次に動作時間計測については、あらかじめ定義された基準となる動作の開始および終了点の特徴から、計測した動作データの開始と終了点を推定する方法を用いている。具体的には、基準となるデータと対象の計測データの各点のコストを動的時間伸縮法 (Dynamic Time Warping: DTW) を用いて求めることにより動作の開始と終了点を推定し、その間隔を動作時間としている。次に動作時間分析については、同一動作の動作時間データをヒストグラムで分析することにより、正規分布からの外れ値の有無により標準外動作の発生有無を判断する仕組みである。実験においては、ラベル貼付作業、ネジ締め作業、梱包作業などで構成されるセル生産作業を対象に評価を実施し、ラベル貼り付け作業時とネジ締め作業時に動作時間が正規分布から外れた箇所を検出できたことが述べられている。

以上のように、この研究では標準外の動作を検知することを目的として、作業時の動作時間を自動で計測し動作時間が正規分布から外れた場合を標準外作業として判断する仕組みである。しかし、動作時間の計測方法について、あらかじめ定義された動作の開始および終了点から、計測対象の時系列データに対する開始および終了点を推定する方法を利用しているため、あらかじめ定義されていないような標準外の動作が生じた場合や作業手順自体に誤りがあった場合、どのように検出することができるのかについては述べられていない。つまり、標準外の動作には想定外の動作が含まれている場合が多く、作業に含まれる想定外の動作自体を検出することができなければ、標準外作業を判断することは難しいものと考えられる。したがって、あらかじめ想定されていないような作業や動作に対して、標準外作業として検出するための仕組みについても検討が必要である。

#### (4)3Dカメラを用いた組立作業での作業ミス検知

Faccioらは、3Dカメラを用いた組み立て作業のリアルタイム作業手順監視および、支援システムについて研究を行った[4]。本研究では、Industry4.0の到来を特徴とする現在において、新たに出現する社会的、技術的、経済的条件に迅速に適用する能力が求められ、製造現場での作業者の役割がさらに重要になるなかで、作業者が効率的に作業を行うことができるように支援を行う仕組みを構築することを目的としている。

本手法では、まずDepthカメラと深層学習技術を用いてセンサ視野内の人体のプロファイルを識別し、作業者の関節の三次元位置を検出する。次に、Control Volumeと呼ばれる三次元位置を持つ仮想的な空間オブジェクトを、作業を行う際の監視対象エリア内に準備し、作業者の関節の三次元位置が

Control Volumeのエリアに出入りするかどうかを確認することにより、作業者の動作を把握する。そして、作業者の動作が正しくない場合には、PCの画面上で対象となるControl Volumeの付近にエラーとして情報を示すことにより、作業者へ作業手順をフィードバックする仕組みである。この研究では従来のシステムと比べ、低価格で柔軟性の高い方法として本手法を提案している。

本提案では作業者の関節の三次元位置とControl Volumeの三次元空間との位置関係に基づいて作業や動作の検出を行っている。作業者の関節位置の検出については深層学習技術を利用しているが、Control Volumeの設定については作業に合わせて専用のモデルや設計が必要となり、作業が複雑になるにつれ、Control Volumeの設定も複雑になる。この課題に対しては、作業者の関節位置情報を作業時の動作における特徴として利用し、深層学習で学習を行い作業として分類する方法を用いることにより、Control Volumeのような作業ごとの専用設計を不要とすることが可能であると考えられる。

### 3.1.2 動作分析・時間分析

製造現場での労働生産性向上のためには、作業時の動作や作業時間を把握し、それを分析することにより非効率な作業や動作に対して改善を行う必要がある。作業者の動作や作業時間を分析するための関連研究をあげる。

#### (1)RGBDカメラを用いた作業動作手順の自動記述

渡邊らは、製造現場での作業動作を客観的に且つ、自動記述するための仕組みとして、作業動作を分析するための記述チャートおよび、自動記述を実現するためのセンシングシステムについて研究を行った [5]。記述チャートについては、作業分析に必要な要件をもとに設計されたものであり、動作の種類と時間、手と注視点の軌跡を動作の順番にチャートにしたものである。このチャートを用いることにより動作の順序や両手の使い方、動作時間および動作が行われた位置、注視点の動きを把握することができるとしている。なお、関連研究としては、この研究のなかで、作業中の動作を自動記述する部分を対象として取りあげる。

自動記述の仕組みについては、RGBDカメラと作業帽子に取り付けたレーザーポインターの焦点を利用して、動作手順および注視点の移動軌跡を記述する方法である。具体的な動作手順の自動記述の処理については、はじめにRGBDカメラで撮影した距離画像に対して背景差分と二値化処理、テンプレートマッチングを行い、物体領域に対してクラス分類することにより物体名を付与する。次にクラス分類後の物体領域画像に正方形のセルを格子状に設定して、存在確率マップと呼ばれる領域を生成する。手についても同様に背景差分、二値化処理、ラベリング処理を行い、存在確率マップの生成を行う。そのうえで、手と物体の存在確率マップ間で、それぞれのセル同士の存在確率の重なり具合を積として求め、その値を閾値判定したうえで手と物体の相互干渉を判定し、その結果から、「手の移動」、「物体を持つての手の移動」、「物体同士の組み合わせ」の動作を判断している。

以上のように、本手法の場合は手や物体を検出するためのテンプレートや、手と物体の干渉に対す

る判定の閾値設定など、作業や物体に合わせたルールベースでの専用設計が必要となる。また、把握することができる作業動作については、物体と手の干渉だけでとらえることが可能な範囲に限定され、評価実験においても両手で3種類のブロックを積み重ねるといった単純動作を対象としている。つまり、この手法に対して、部品や工具など様々な物体が存在するような製造現場における有用性については、確認されていない。一方、部品や工具などの物体が存在するような製造現場においても、深層学習技術を用いることにより、画像を用いた学習だけで手や部品などの物体を検出することができる。したがって、対象物体ごとの設計を不要とし汎用的に利用することが可能である。また、検出対象の物体の種類が増えた場合にも、その画像に対する学習を行うことにより、幅広い製造現場で利用することが可能であると考えられる。

## (2)距離画像を用いた両手動作の解析

松原らは、製造現場での改善について支援することを目的として、Kinectを用いた動作解析システムを開発した[6]。Kinectで取得される距離画像を用いて両手の位置を検出し、その位置と作業時に手が通過するエリアの位置関係から作業動作を判定し、作業の動作時間を計測する。両手の位置の検出については、はじめに作業者がいない状況の画像と作業者が作業を行っている状況の距離画像から、その差分画像により人物領域を抽出するマスクを作成する。次に頭部、胴体部、腕部、手部といった身体部位を抽出するために、マスクで抽出された人物領域の距離画像に対して放射リーチ関数(Radial Reach Correlation: RRC)を用いて注目画素から8方向へのリーチ長に関する特徴量の抽出を行い、あらかじめ設計された身体部位の特徴量との閾値判定により判断を行っている。なお、RRCで利用する情報については、距離画像は各画素に距離情報を持っており、この距離情報を濃淡値に対応させた濃淡画像を利用している。

具体的な身体部位の抽出方法については、たとえば丸い頭部を抽出する際には、その形状の特徴について円形をしていることから、8方向のリーチ長の最大と最小の差が閾値より小さい場合を対象とし、手部を抽出する場合は、短いリーチ長が6個以上連続する場合を対象の条件としている。次に、動作の計測については、作業動作ごとに両手が通過すべきエリア(通過判定エリア)をあらかじめ設計しておき、そのエリアに対する手の通過により動作を判断し、動作時間を求めている。複数動作で構成される作業に対しては、作業を構成する動作の始点と終点の位置を、作業の開始から終了までそれぞれ設定し、次に、作業の順番により通過する順序と通過判定エリアを設定する。また、このエリアについては部品箱や組立品の大きさにより異なるため、そのエリアサイズは個々に設定する必要がある。

本研究においては、刃物製品の組み立て工程において実験を行い、標準作業時間を超えた作業を検出することにより、標準作業と異なる動作の発生を検知できることを確認したと述べられている。しかし、本手法については信頼性について課題が考えられる。Kinectで取得した距離画像よりRRCで抽出した特徴量を用いて両手を検出しているが、作業時に使用する部品や治具などの物体を手で掴んでいる状態を想定した場合、背景差分で作成する人物領域のマスクやRRCで抽出した特徴量には、手

以外の特徴が含まれるため、その状況で手を正確に検出するためには、検出用の閾値の設定が難しいものと考えられる。手の検出を行う場合には、深層学習技術による物体検出や人物の骨格検出などを利用する方法が望ましいものと考えられる。

#### (3) ウェアラブルセンサを用いたセル生産ラインの作業動作分析と進捗計測

北澤らは、セル生産ラインにおいて、ウェアラブルセンサから得られたデータを用いて作業者の動作を分析し、生産進捗や作業履歴をリアルタイムで計測するシステムの研究を行った[7]。製品1個単位の生産時間の計測と、作業者の移動・停止の動作推定から作業履歴の1つである異常動作の有無を計測することを実現している。このシステムは、ビーコンと3軸加速度センサが組み合わさったウェアラブルセンサを使用しており、センサから得られたデータに対する効果的なデータ前処理と分析手法の組み合わせを明らかにすることも、この研究の目的としている。製造進捗の計測については、セル生産ラインの作業開始位置付近にビーコンのアンテナを設置し、作業者が持つビーコンの電波受信強度から作業者の位置を推定し、電波受信強度がピークに達する周期を、その作業における製品1個単位での処理時間であるサイクルタイムとして求める方法である。次に、異常動作の計測については、身体に装着された3軸加速度センサの加速度ベクトルを分析し、作業者の移動と停止を推定することにより移動回数を推定し、この移動回数について標準作業における移動回数と乖離がある場合には、作業者の異常動作として検出する仕組みである。なお、この研究における異常動作とは、作業中の部品の落下や作業忘れが発生した際に、部品を拾ったり作業の後戻りをしたりすることで移動回数が増えるような動作を対象としている。

このシステムは、セル生産ラインにおける製品1個単位の生産時間の計測と、ライン内での作業者の移動・停止の動作から異常動作の有無を計測する仕組みであるため、ライン内の個々の作業に対する進捗を把握することはできない。また、着目した異常動作についても、異常が発生した後の対処時の動作である部品を拾ったり、作業の後戻りをしたりする動作を計測しており、作業の異常が発生した時点の状況や内容についてリアルタイムにとらえることはできない。作業の進捗や作業異常をリアルタイムに計測するためには、セル生産ライン内での個々の作業を対象として作業の状況や作業手順の誤り、作業時間の乖離、作業時の動作の特徴などをとらえる必要があるものと考えられる。

#### (4) スマートタグを用いた作業者の動作分類

佐藤らは、製造工程における作業者の動作分析を目的とした行動センシングに関して、ウェアラブル機器やカメラを用いる方法では導入コストや精度が課題であるとし、作業者が簡易に装着可能な小型スマートタグを利用した方法について研究した[8]。小型スマートタグはこの研究グループと企業間で共同開発されたものを利用し、その機能としては、加速度、ジャイロ、温度をセンシングすることが可能である。また、BLE通信機能を持ち、ビーコン電波を受信するとともに、センシングした情報を配信することができる。動作分類モデルについては、SVM、決定木、ニューラルネットの全結合モ



デル (FNN) などを検討した結果として、リカレントニューラルネットワーク (Recurrent Neural Network: RNN) を拡張したLong Short Term Memory (LSTM) を利用した。実験では、作業者のポケットに装着したスマートタグのセンサから収集したデータに対して、「通常歩行」、「台車押し歩行」、「荷物持ち歩行」についての分類評価を行った。

実験の結果からは、小型スマートタグで得られたセンサ値から機械学習モデルを用いて動作分類および作業分析が可能であることを示しているが、実験の対象が単純な歩行動作に限られており、その有用性については十分には示されていない。また、分類精度についてもFNNが約92%だったのに対して、LSTMでは約84%にとどまっており、LSTMを利用することに対する効果については明らかにされていない。

ここまで述べたとおりこれらの関連研究については、対象の作業に合わせた専用の設計が必要となる手法が多く、製造現場の環境に対する制約や前提条件を必要とする研究も見られ、有用性や信頼性についても課題があるものが含まれている。また、これらの関連研究の手法について、実際に製造現場へ導入を進めていくためには、IE担当者に対して高いITスキルが求められる可能性が高いものと考えられる。

一方、筆者の研究では、深層学習技術を用いて製造現場を撮影した画像からEnd-to-Endでの学習を行うことにより、作業の分類や物体の位置情報の取得、動作の特徴抽出を可能とした。具体的には、普段の製造現場の作業時の画像から深層学習の学習を行い、そのネットワークを用いて分類、検出を行うことにより、作業員へのセンサの取り付け、特定の動作に対する制限などの制約事項をなくした。また、この手法については、物体の位置情報の推定や、作業および動作の分類についても、対象となる製造現場や作業に合わせた専用設計を不要とすることができる。これにより、企業内にIT担当者が不在である場合や、IE担当者のITスキルが不足しているような場合でも、製造現場の画像などのデータから学習データを準備し、深層学習のモデルに対して学習を行うことにより、幅広い製造現場で汎用的な仕組みとして利用することが可能となる。

### 3.2 深層学習に関する基礎理論

本研究では、作業の分類や作業動作の特徴抽出に深層学習の技術を利用している。深層学習とは、人間が自然に行うタスクをコンピュータに学習させる機械学習の手法の1つである。従来の機械学習のアルゴリズムでは、データセットの中から、分類や検出などに利用できる特徴量を手作業で抽出し、学習データとして作成する必要があったが、深層学習を利用することにより、データセットのなかから特徴を自動的に抽出することができるようになり、End-to-Endでの学習により利用することが可能となった。

本研究では、深層学習の手法のなかでも画像認識で広く利用されている畳み込みニューラルネットを利用してシステムの開発と評価を行った。4章および5章で述べる研究内容に関連する深層学習の理論および、各研究で利用した技術について概説する。

#### 3.2.1 畳み込みニューラルネット

##### (1) 畳み込みニューラルネットの概要

畳み込みニューラルネット (Convolutional Neural Network : CNN) とは、LeCunらにより研究[9]がすすめられた畳み込み層とプーリング層を含んだ順伝播型ニューラルネットであり、画像認識の分野で優れた性能を発揮した深層学習ネットワークである。CNNの一般的な構成について図 3-1へ示す。

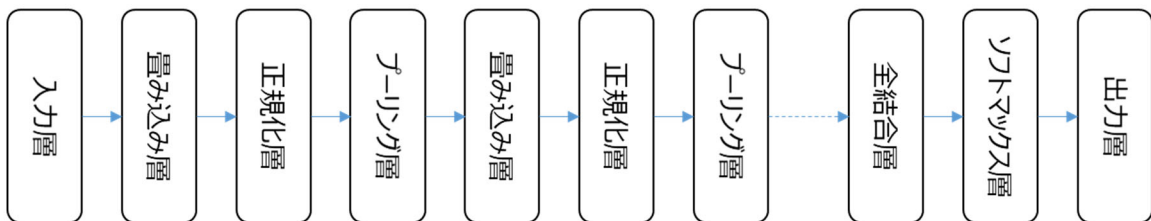


図 3-1 一般的なCNNの構成 (分類モデル)

CNNの特徴は、畳み込み層を通じてデータの中から高次の特徴量を抽出し、プーリング層でその特徴の位置に関する感度を低下させることにより、認識精度にロバスト性を与えるための仕組みの部分である。従来からの順伝播型ニューラルネットは隣接層のユニットのすべてを全結合していたため、各レイヤーのユニットに対する重みの数も非常に多くなり、処理するためには大量のメモリが必要であった。また、入力データに対する局所的な偏りや、データ内での特徴となる位置のバラツキがある場合にも、その影響を受け易いという欠点があった。一方、CNNの場合は隣接層間の特定のユニットのみが結合を持つ構造 (図 3-2) をとることにより、これらの課題を解決したものである。

なお、CNNは画像分類のタスクで優れた性能が認められているが、回帰タスクでも利用されている。分類タスクの場合は、入力情報に対して離散値 (クラス) を予測するが、回帰の場合は、入力情報に対して連続値 (実数) を予測する。筆者の研究においても、画像から作業を分類する際には分類用のネッ

ネットワークを利用し、画像から人の関節位置や工具の位置を検出する際には、回帰用のネットワークを利用した。

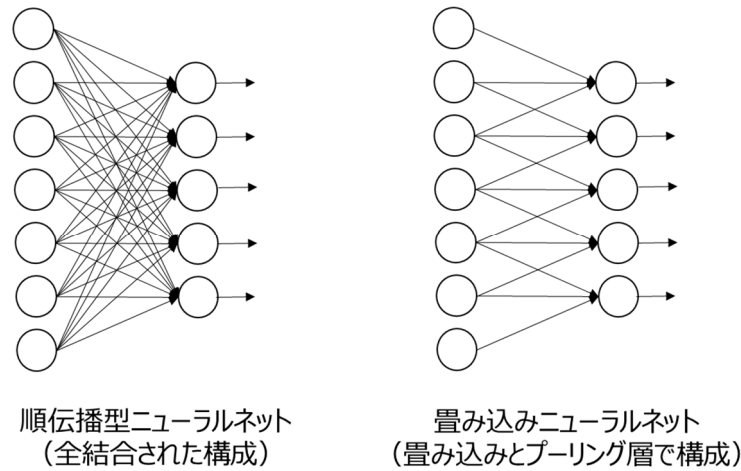


図 3-2 順伝播型ニューラルネットと畳み込みニューラルネットの構成

## (2) CNNにおける受容野のモデル

CNNは、生物の視覚系における神経細胞の受容野の働きを参考にして研究された。外界から眼に取り込んだ光は網膜で電気的な信号に変換され、その後、視覚野に伝達される。視覚野の神経細胞のなかでは、網膜の特定の場所に特定のパターンが入力されると興奮し、それ以外のときには興奮しないという選択的な動きをする。この細胞は単純型細胞と複雑型細胞に分けられ、単純型細胞は、受容野が受け取る光の強さに敏感に反応する特性を持っているが、複雑型細胞は、受け取った感覚信号から特徴を抽出するという特性を持っている [10]。つまり、視野内の特定の方向と位置を持つ物体に反応する単純型細胞と、局所的な位置に対して鈍感ではあるが、視覚的な並行移動、回転、拡大などに対して頑健な特性を持つ複雑型細胞で構成される。岡谷の研究 [11] を参考に、簡単なネットワークモデルを示す (図 3-3)。単純型細胞は図 3-3の(a)のように右の層の1つのユニットが左の層の $4 \times 4$ のユニット群とのみ結合を持ち、そこに(b)のような反応対象のパターンが入力されたときのみ、それに反応して活性化する。複雑型細胞は(c),(d)の単純型細胞を中間層として $3 \times 3$ のユニット群と結合を持ち、このユニット群のうちの1つでも活性化すると、自身も活性化する。この(c)と(d)で中間層のユニットの反応は異なるが、複雑型細胞の反応は同一となる。つまり、中間層である単純型細胞は入力パターンの位置変化に敏感であるが、出力層である複雑型細胞は一定の範囲の位置のズレに対して鈍感である。なお、この中間層と出力層がそれぞれ、畳み込み層およびプーリング層となる。

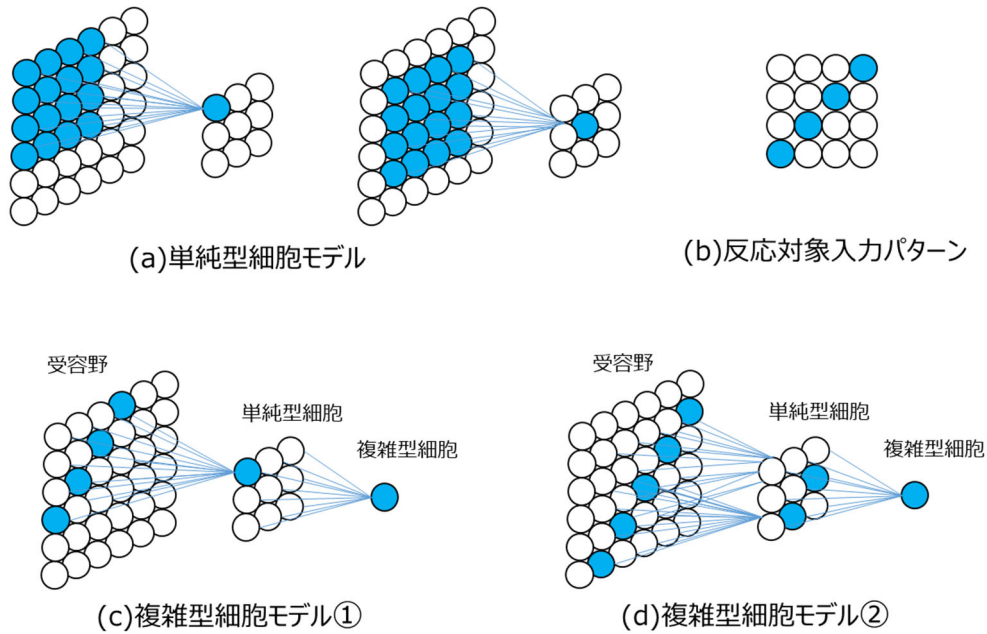


図 3-3 単純型細胞と複雑型細胞のモデル

(3) 畳み込み層とプーリング層

CNNは入力層で受け取ったデータに対して、畳み込み層とプーリング層での変換を繰り返して特徴を抽出し出力する。CNNの畳み込み層では、フィルタのパターンと類似したパターンを入力データから検出し、特徴として抽出する。この畳み込み層の概要と演算方法について、書籍[12]を参考に記述する。

図 3-4のとおり、RGBの3色からなるカラー画像に対するインデックス変換された画素の画素値を入力情報とする場合、第 $l$ 層の畳み込み層に対して1つ前の第 $l-1$ 層から $W \times W \times K$ チャンネル（画像の場合は $K=3$ ）の画素値 $z_{ijk}^{(l-1)}$ （ただし、 $i=0, \dots, W-1, j=0, \dots, W-1, k=0, \dots, K-1$ ）を受け取り、これに画素サイズ $H \times H$ の $M (=3)$ 種類のフィルタ $h_{pqkm}$ （ただし、 $p=0, \dots, H-1, q=0, \dots, H-1, m=0, \dots, M-1$ ）を適用する場合の畳み込み処理の概要である。なお、各フィルタは入力と同一チャンネルでサイズを $H \times H \times K$ とする。ここでは、フィルタごとに並行して計算が行われ、その結果として $u_{ijm}$ が出力される（式(1)）。つまり畳み込みは、画像などの入力情報にフィルタを重ねたときに、画像とフィルタの重なり合う画素どうしの積を求めている。したがって、入力情報にフィルタを重ねる際に、入力情報の範囲からはみ出た位置にはフィルタを配置することができないため、出力された $u_{ijm}$ のサイズは入力である $z_{ijk}^{(l-1)}$ のサイズより小さくなる。出力サイズを入力サイズと同一にするためには、入力情報の周囲を0などでパディングする方法が用いられる。また、図 3-4では省略されているが、式(1)の $b_{ijm}$ はバイアスを示す値である。このようにして求められた $u_{ijm}$ に対して式(2)の活性化関数を適用し、この結果を畳み込み層からの出力とする。

次に、畳み込み層に続くプーリング層で、データの空間的サイズをダウンサンプリングにより集約さ

せる。プーリング処理には領域内の最大値を用いる最大値プーリング (Max Pooling) や、平均値を用いる平均値プーリング (Average Pooling) などがある。プーリング処理は複雑型細胞の働きを持ち、対象となる特徴量に対して多少の位置ズレが生じても、その影響を受けづらくなる。

なお、CNNは画像認識で多く用いられている技術であるが、筆者の本研究では、画像以外に時系列データに対しても、CNNを利用した。

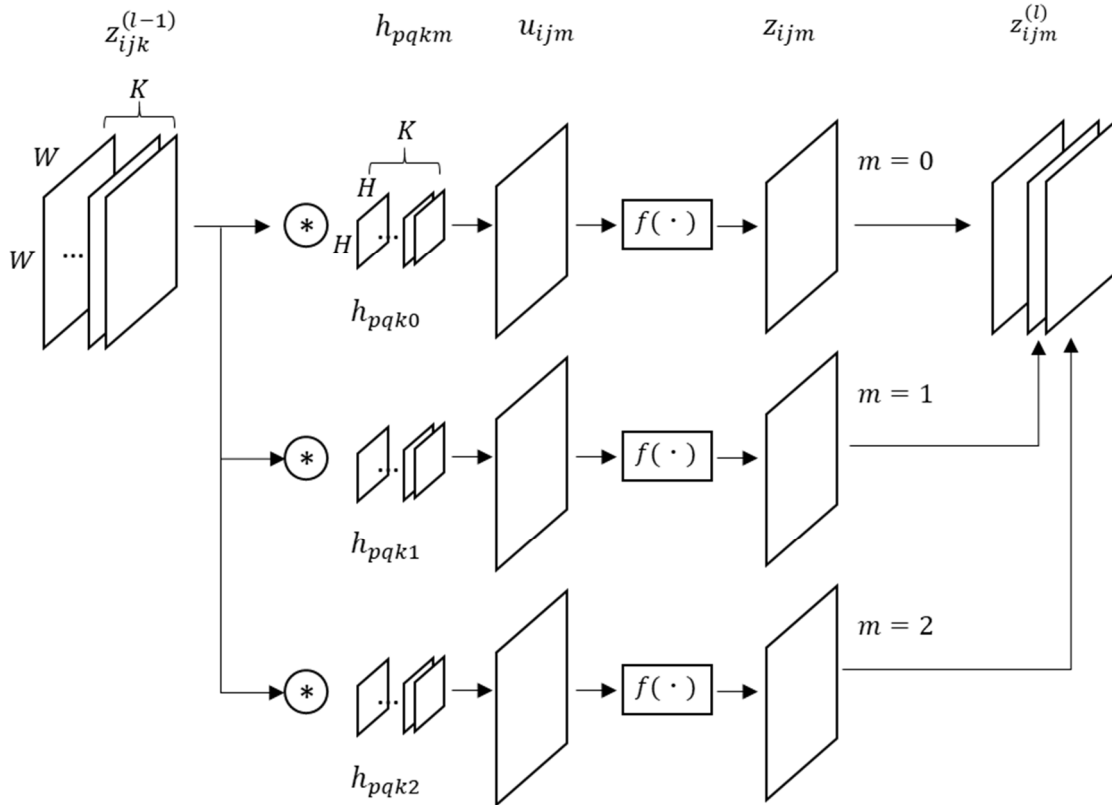


図 3-4 畳み込み層の概要

$$u_{ijm} = \sum_{k=0}^{K-1} \sum_{p=0}^{H-1} \sum_{q=0}^{H-1} z_{i+p,j+q,k}^{(l-1)} h_{pqkm} + b_{ijm} \quad (1)$$

$$z_{ijm} = f(u_{ijm}) \quad (2)$$

### 3.2.2 ネットワークの学習

#### (1) 転移学習

転移学習は、深層学習におけるモデルの構築方法としてよく利用されている学習方法である。事前にあるタスク向けに学習済のネットワークを利用して、類似した新しいタスクに合わせてパラメータの調整を行い新しいタスクの学習の開始点として使用することにより、学習にかかる時間が大幅に短縮できるとともに、必要な学習データも少なくすることができる。事前の学習済モデルは既に重みの学習を行っているため、ゼロからネットワークを学習させるよりも効率的に学習を行うことができる。ただし、

転移学習を効果的に利用するためには、利用するモデルの入力情報が類似した特徴を持つネットワークモデルを利用することに注意する必要がある[13]。なお、今回の筆者の研究においても、画像の分類用のネットワークと物体内の検出対象箇所の予測用のネットワークで転移学習を利用した。

## (2) 転移学習時の学習率

CNNの学習とは、与えられた訓練データをもとに計算される誤差関数を最小化することが目的である。オプティマイザに勾配降下法を用いたネットワークを例にあげると、CNNの学習とは、訓練データの集合(3) (ただし、 $x$ は入力データ、 $d$ は目標出力)をもとに計算される式(4)(5)の誤差関数 $E(w)$ に対して、ネットワークのパラメータ(重みとバイアス)を調整することにより最小化することである。式(4)は回帰用の誤差関数で、 $d_n$ は任意の実数値を成分に持つベクトルを指し、 $y(x_n; w)$ はネットワークからの出力を指す。また、式(5)は分類用の誤差関数で、 $d_{nk}$ は目標出力としてクラス数( $k$ )と同じ成分数を持つベクトルであり、正解クラスに対する成分のみ1をとり、残りはすべて0となる。また、 $y_k(x_n; w)$ はクラス $k$ に対するネットワークからの出力確率である。つまり、その学習のゴールは式(4)および、式(5)に示す誤差関数 $E(w)$ に対し最小解を与える $w = \operatorname{argmin}_w E(w)$ を求めることであるが、 $E(w)$ は一般に凸関数ではなく、大局的な最小解を求めるのは通常不可能とされている。そこで、勾配降下法を用いて $w$ を繰り返し更新する反復計算により最小解を求める。このとき、勾配とは $E$ を $w$ で偏微分したベクトル(6)で表現され( $M$ は $w$ の成分数)、勾配降下法は現在の $w$ を負の勾配方向( $-\nabla E$ )に少しずつ動かすことを何度も繰り返し、現在の重みを $w^{(t)}$ 、動かした後の重みを $w^{(t+1)}$ とすると式(7)のように更新する[14]。ここで $\varepsilon$ は重み $w$ の更新量の大きさを定める定数で、学習率と呼ばれる。

学習率が小さすぎる場合には重みの更新量が小さいため学習が進む速度が低下し、学習に要する時間が大幅に増えることになる(図3-5(a))。一方、学習率が大きすぎる場合には、重みが大きく更新されるため、その結果として発散により学習が不安定になり、振動が生じる場合がある(図3-5(b))。振動とは、オプティマイザでの計算結果による更新処理が大きすぎる場合に、最小値へ向かう途中でオーバーシュートし、反復を繰り返して収束しなくなる状況のことを指す[15]。

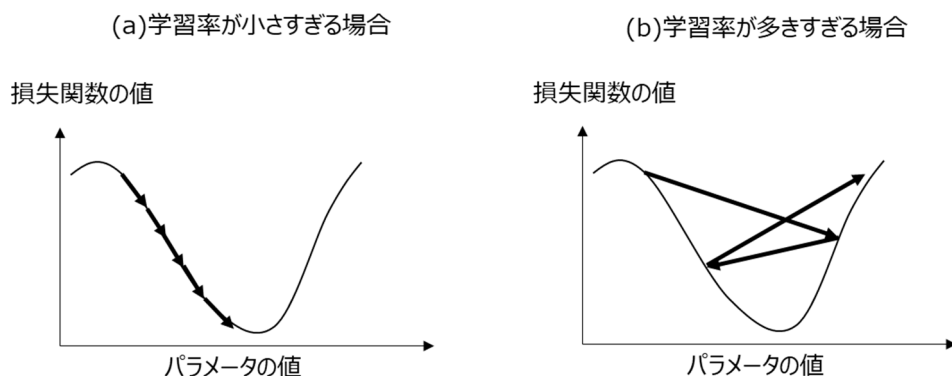


図 3-5 学習率と学習の進み方

なお、転移学習を行う際は、学習済の転移層については既に重みの学習が行われているため重みを固定（凍結）するか、もしくは重みの更新量を小さくし、新たに追加した層に対する重みの更新量を大きくする方法が用いられる（図 3-6）。

本研究においても、転移層の初期学習率を小さくし、新たに追加した層に対する学習率を大きく設定したうえで学習を行った。

$$D = \{(x_1, d_1), \dots, (x_n, d_n)\} \quad (3)$$

$$E(w) = \frac{1}{n} \sum_{n=1}^N \|d_n - y(x_n; w)\|^2 \quad (4)$$

$$E(w) = - \sum_{n=1}^N \sum_{k=1}^K d_{nk} \log y_k(x_n; w) \quad (5)$$

$$\nabla E \equiv \frac{\partial E}{\partial w} = \left[ \frac{\partial E}{\partial \omega_1} \dots \frac{\partial E}{\partial \omega_M} \right]^T \quad (6)$$

$$w^{(t+1)} = w^{(t)} - \epsilon \nabla E \quad (7)$$

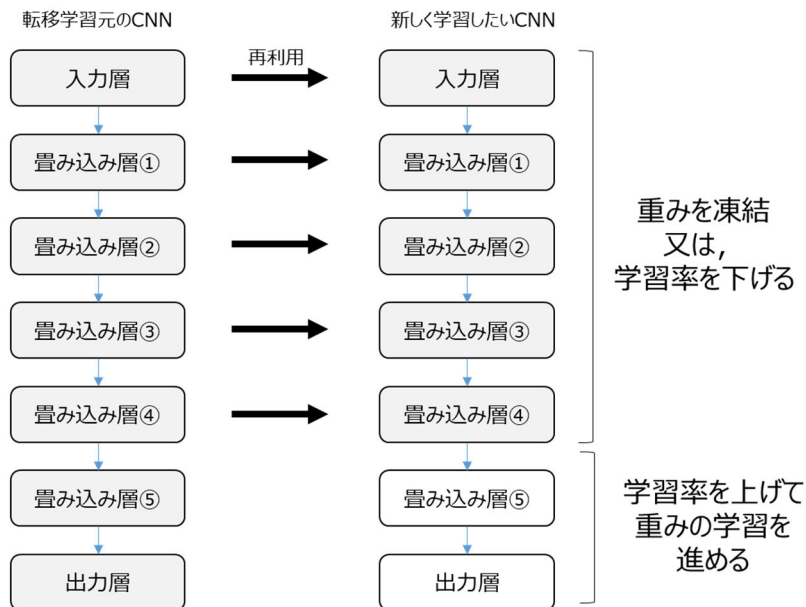


図 3-6 学習済の層の再利用

## (3) 学習用のオプティマイザ

ネットワークの学習の際に用いられる最適化アルゴリズムはオプティマイザと呼ばれ、いくつかの種類が利用されている。オプティマイザの選択は、学習の速度と予測の性能に大きく影響するため、いくつかの研究やベンチマークが行われているが、結果がベンチマークのケースに依存することや、確率論や一般化などの側面から、最適なオプティマイザを適切に説明することは難しいとされている [16][17].

いくつかのオプティマイザについて概説すると、確率的勾配降下法 (Stochastic Gradient Descent : SGD) は、学習用データセットの全てを使って学習を行うバッチ勾配降下法に対して、学習の各ステップで学習用データセットから無作為に一つのインスタンスを選び出し、そのインスタンスだけを使って勾配の計算を繰り返すことにより、学習の効率化をはかるための方法である。一般的には、データを小さなサブセットに分割してサブセット毎に学習を繰り返すミニバッチ学習と合わせて利用される。SGDにモーメンタム項を適用したオプティマイザはモーメンタム項付き確率的勾配降下法 (SGD with Momentum : SGDM) と呼ばれ、慣性最適化のアルゴリズムを用いて最小化へ向けた更新の際に1ステップ前の勾配成分を参考に更新をかけることにより、SGDにおける振動を抑えるとともに、学習の進行速度の向上を狙いとしている。SGDMの重み更新の計算方法は式(8)に示すとおり、前述の式(7)にモーメンタム項を追加したものであり、 $\mu$  は前の勾配成分から現在の反復への寄与率を示す。

$$w^{(t+1)} = w^{(t)} - \varepsilon \nabla E + \mu(w^{(t)} - w^{(t-1)}) \quad (8)$$

また、適応モーメント推定 (Adaptive Moment Estimation : ADAM) [18]は、パラメータの更新の際に、過去の勾配の一次モーメント (平均) と二次モーメント (分散) の指数移動平均を利用するアルゴリズムである。式(9)は一次モーメント、式(10)は二次モーメントを表す式で、 $\beta$  はそれぞれの減衰率を示す。これらのモーメントは0 (ゼロ) のベクトルとして初期化されているため、初期のタイムステップでは値が0 (ゼロ) に偏るといったバイアスが生じる。式(11)と式(12)はこのバイアスを補正するための計算となる。式(13)がADAMにおける重みの更新式となり、 $\gamma$  はゼロ除算を回避するための定数となる。

$$m_t = \beta_1 m_{t-1} + (1 - \beta_1) \nabla E \quad (9)$$

$$v_t = \beta_2 v_{t-1} + (1 - \beta_2) \nabla E^2 \quad (10)$$

$$\hat{m}_t = \frac{m_t}{(1 - \beta_1^t)} \quad (11)$$

$$\hat{v}_t = \frac{v_t}{(1 - \beta_2^t)} \quad (12)$$



$$w^{(t+1)} = w^{(t)} - \frac{\varepsilon \hat{m}_t}{\sqrt{\hat{v}_t + \gamma}} \quad (13)$$

他にも、Momentumと同じようにSGDにおける振動を抑えることを目的として、学習時に勾配の大きさに応じて学習率を調整できるようにした平方根平均二乗伝播（Root Mean Square Propagation : RMSProp）[19]など、オプティマイザのアルゴリズムについては、多くの研究が進められているが、今回の筆者の研究では、論文を参考に、それぞれの転移学習に利用したネットワークモデルの学習に使用されていたオプティマイザである、SGDMおよびADAMをそれぞれ利用した。

#### (4) 学習データの拡張

データ拡張とは学習データセットを人工的に増やす手法のことで、深層学習における汎化性能向上のためのテクニックの一つとされている。具体的には、学習データに含まれる画像に対して、シフト、回転、反転、拡大／縮小、ランダムなアフィン幾何学変換などを行い、データのバリエーションを増やす方法である[20]。これにより、限られたデータに対する過学習が緩和され、学習データが少ない場合でも汎化性能の向上が期待できる。

筆者の研究においても、画像の類似性が高いデータセットに対して、データ拡張処理を利用して学習データとして用いた。

## 参考文献

- [1] 宇田吉広, 吉田和広, 村田嘉利: 深度カメラを用いたピッキング工程における作業ミス検知システム, 情報処理学会論文誌 コンシューマ・デバイス&システム, Vol.6, No.1, pp.63-74 (2016) .
- [2] 清水早苗, 平湯秀和, 浅井博次: カメラを用いたものづくり現場における締め付け作業のポカヨケシステムの開発, 電気学会論文誌, Vol.129, No.5, pp.916-922 (2009) .
- [3] 西田一貴, 音田浩臣: セル生産ラインにおける作業者の標準外動作検知システムの開発, システム制御情報学会論文誌, Vol.33, No.5, pp.149-155 (2020) .
- [4] Maurizio Faccio, Emilio Ferrari, Francesco G. Galizia, Mauro Gamberi, Francesco Pilati : Real-time assistance to manual assembly through depth camera and visual feedback, Proc. 52nd CIRP Conference on Manufacturing Systems, pp. 1254-1259 (2019).
- [5] 渡邊瞭太, 橋本学: 組み立て作業分析のための動作手順と手・注視点の移動軌跡の自動記述システム, 精密工学会誌, Vol.82, No.5, pp.473-480 (2016) .
- [6] 松原早苗, 渡辺博己, 曾賀野健一, 棚橋英樹: 距離画像を用いた両手作業の動作解析システム, 岐阜県情報技術研究報告書, 第18号, pp.24-29 (2018) .
- [7] 北澤正樹: ウェアラブルセンサーを用いた作業者動作分析によるセル生産ラインのリアルタイム生産進捗計測システムの提案, 第12回社会システム部会研究会, pp. 231-238 (2017) .
- [8] 佐藤祐紀, 堀川三好, 岡本東, 中田恵史: 製造工程におけるスマートタグを用いた作業者の動作分類, 情報処理学会第82回全国大会論文集, pp.237-238 (2020) .
- [9] Yann LeCun, Leon Bottou, Yoshua Bengio, Patrick Haffnet : Gradient-based learning applied to document recognition. Proc. of the IEEE, pp. 2278-2324 (1998).
- [10] Pietro Berkes, Laurenz Wiskott : Slow feature analysis yields a rich repertoire of complex cell properties, Speech, and Time-Series, Journal of Vision, Vol.5, pp. 579-602 (2005).
- [11] 岡谷貴之: ディープラーニングと画像認識—基礎と最近の動向—, オペレーションズ・リサーチ, Vol.60, No.4, pp.198-204 (2015) .
- [12] 岡谷貴之: 機械学習プロフェッショナルシリーズ 深層学習, pp.87-89, 講談社 (2015) .
- [13] Aurelien Geron(著), 長尾高弘(訳): scikit-learnとTensorFlowによる実践機械学習, pp.286-287, オライリー・ジャパン (2018) .
- [14] 岡谷貴之: 機械学習プロフェッショナルシリーズ 深層学習, pp.23-24, 講談社 (2015) .
- [15] Aurelien Geron(著), 長尾高弘(訳): scikit-learnとTensorFlowによる実践機械学習, pp.113-114, オライリー・ジャパン (2018) .
- [16] Ashia C. Wilson, Rebecca Roelofs, Mitchell Stern, Nathan Srebro, and Benjamin Recht : The Marginal Value of Adaptive Gradient Methods in Machine Learning, In Advances in Neural Information Processing Systems 30, pp. 4148-4158. Curran Associates, Inc. (2017).

- [17] Frank Schneider, Lukas Balles, Philipp Hennig: DeepOBS : a deep learning optimizer benchmark suite. arXiv preprint arXiv:1903.05499 (2019).
- [18] Diederik P. Kingma, Jimmy Lei Ba.: Adam: a method for stochastic optimization, In ICLR (2015).
- [19] Tijmen Tieleman, Geoffrey Hinton : Lecture 6.5-RMSProp:Dividet the gradient by a running average of its recent magnitude, COURSERA: Neural networks for machine learning, 4(2), pp. 26–31 (2012).
- [20] Aurelien Geron(著), 長尾高弘(訳) : scikit-learnとTensorFlowによる実践機械学習, p.309, オライリー・ジャパン (2018) .

## 第4章 作業分析結果と深層学習を利用した作業手順のリアルタイム監視

### 4.1 はじめに

1章および2章で述べたとおり、製造現場では労働生産性向上を目的として、インダストリアル・エンジニアリング（IE）による改善活動が行われている。筆者はこれまで製造業向けに、IEツールの1つである作業分析ソフトの導入に携わってきた。作業分析ソフトとは、製造現場での作業状況を撮影した映像を使用して時間研究[1]や動作研究[2]を行い、非効率な点や問題点を見つけ出し改善を図ることを主な目的とした分析用のソフトウェアである。分析された結果は標準作業や標準作業時間として製造現場へ展開される。これら標準化された手順が製造現場で正しく守られているかを把握する方法として従来は、日報での確認や現場監督者による見回りが中心であったため、手間や時間がかかるうえに、その結果に対して客観性や正確さに欠ける点が課題とされていた。近年、映像やセンサなどを利用して作業者の作業ミスを自動検知するシステムは考案されているが、この多くは対象となる作業に対する専用設計となっており、また、検知した作業ミスに対して、正しい手順での作業のやり直しまで指示するような汎用的なシステムは報告されていない。

そこで本研究では、「時間研究」として作業分析ソフトで分析された結果から定められた標準作業と標準作業時間を監視用の基準として、製造現場での作業の状況をリアルタイムで監視し、異常が検知された際にその場で作業者へ正しい作業手順をティーチングする手法を開発した。本手法では製造現場での作業状況を画像から認識する方法に、分類用の畳み込みニューラルネット（CNN）を利用する。CNNなどの深層学習を利用する際は、大量の学習データが必要となるために、その準備に多大な労力と時間を要することが課題とされており、学習データ準備の省力化を目的とした研究も行われている[3]。筆者らはこの課題を解決するために、作業分析ソフトで分析した結果とその際に使用した動画から学習データを自動的に生成する手法を開発し、学習データの準備にかかる労力と時間を大幅に削減した。また、製造現場の画像を利用したEnd-to-EndでのCNNの学習によって作業監視の仕組みを実現することにより、多くの製造現場でも利用できる高い汎用性を目指した。

### 4.2 リアルタイム作業手順監視システムの特徴

#### 4.2.1 要求条件

本作業手順監視システムの要件について以下にあげる。

要件1：作業改善を目的とした客観的で合理的な監視基準であること

要件2：CNNの学習データの準備が容易であること

要件3：作業手順誤りや時間乖離を正確に検知できること

要件4：製造現場をリアルタイムで監視可能であること

製造現場の作業の改善を目的とした監視のためには、その監視基準は現場監督者などの主観や勘によるものではなく、IEとして科学的に分析評価され、客観的かつ、合理的に定められたものである必

要がある（要件1）。また、製造現場ではIE活動により改善のサイクルが回されるため、監視基準についても改善に合わせて頻繁に見直しが必要となる。したがって作業の認識に使用するCNNについてもIE活動のサイクルのなかで再学習を行う必要があるため、その学習データは容易に準備できる必要がある（要件2）。また、製造現場で作業手順の誤りや作業時間の乖離が発生した場合には、その結果として、生産性や品質へ与える影響を考慮すると、その検知は正確である必要があり（要件3）、その影響を最小限にとどめるためには問題発生時点でただちに検知し、現場監督者などへ通知するとともに、作業員へ正しい作業手順でティーチングする必要がある（要件4）。

#### 4.2.2 作業手順監視方式

4.2.1 節で示した要件を満足するリアルタイム作業手順監視システムの方式を図 4-1に示す。本方式では、市販されている作業分析ソフトで分析した結果である標準作業と標準作業時間および分析の際に使用した画像により学習したCNNを、監視の際の基準および作業認識の仕組みとして利用することにより、作業改善を目的とした客観的で合理的な監視基準とした（要件1）。また、作業分析結果とその際に使用した動画からCNNの学習データを自動的に生成する仕組みを用いることにより、IEの改善サイクルが繰り返されるなかで、学習用データの準備を容易にした（要件2）。監視のリアルタイム性と検知精度については、比較的スループットが高いとされる学習済のCNNネットワークを選定し、転移学習して利用することにより、限られた学習データで検知精度を高めた（要件3,4）。

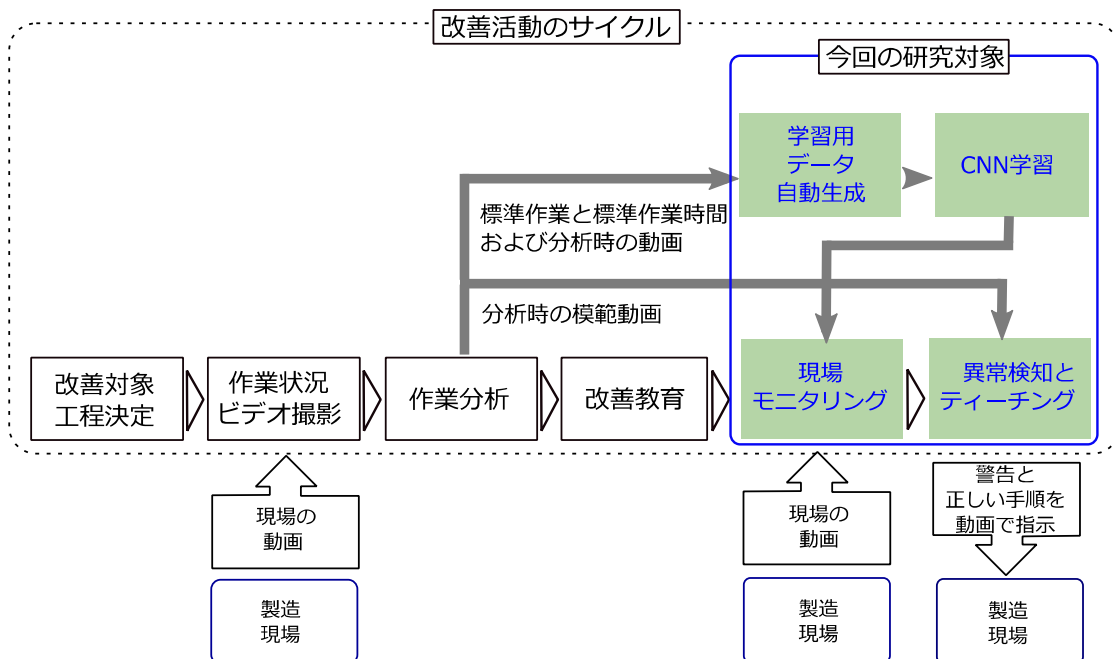


図 4-1 リアルタイム作業手順監視方式

4.2.3 ユースケース

本方式に基づいたユースケースを図 4-2に示す。作業分析の担当者は、製造現場で撮影された動画を用いて作業分析ソフトにより分析を行い、課題や改善ポイントを見つけ、作業の標準化をはかる。その際、作業手順や作業時間の見直しが必要となる場合は、作業手順書を改訂し製造現場の担当者へ教育を行う。次に、分析時の結果である標準作業と標準作業時間および分析に使用した動画は、リアルタイム作業手順監視システムへ取り込まれる。同システムはこれらの情報から学習用のデータを自動的に生成し、その学習データを用いてCNNの学習を行うことにより監視用のモデルを作成または更新し監視を行う。作業者は与えられた作業手順にしたがって作業を行うが、標準作業と異なる順番で作業を行った場合や、作業のムダにより標準作業時間に対して実際にかかった作業時間が一定以上乖離する場合は、同システムが異常として検知し、警告とともに正しい作業手順を動画で作業者へ示す。つまり、本システムが標準作業と標準作業時間および作業時の模範となる画像を覚え、その内容に基づいて製造現場の監視を行い、手順や時間に相違や乖離が検知された際には、作業者に対して正しい作業手順のティーチングを行う仕組みである。

本ユースケースにおいてポイントとなるのは、作業分析結果から作成され製造現場へ展開される作業手順書と、同じく作業分析結果に基づいた監視基準、分析結果から生成された学習データで学習した監視モデルおよびティーチングの際の作業手順について、IEとして科学的に分析された同類の内容を利用できる点であり、作業分析と同期をとりながらIE改善サイクルを回すことが可能となることである。

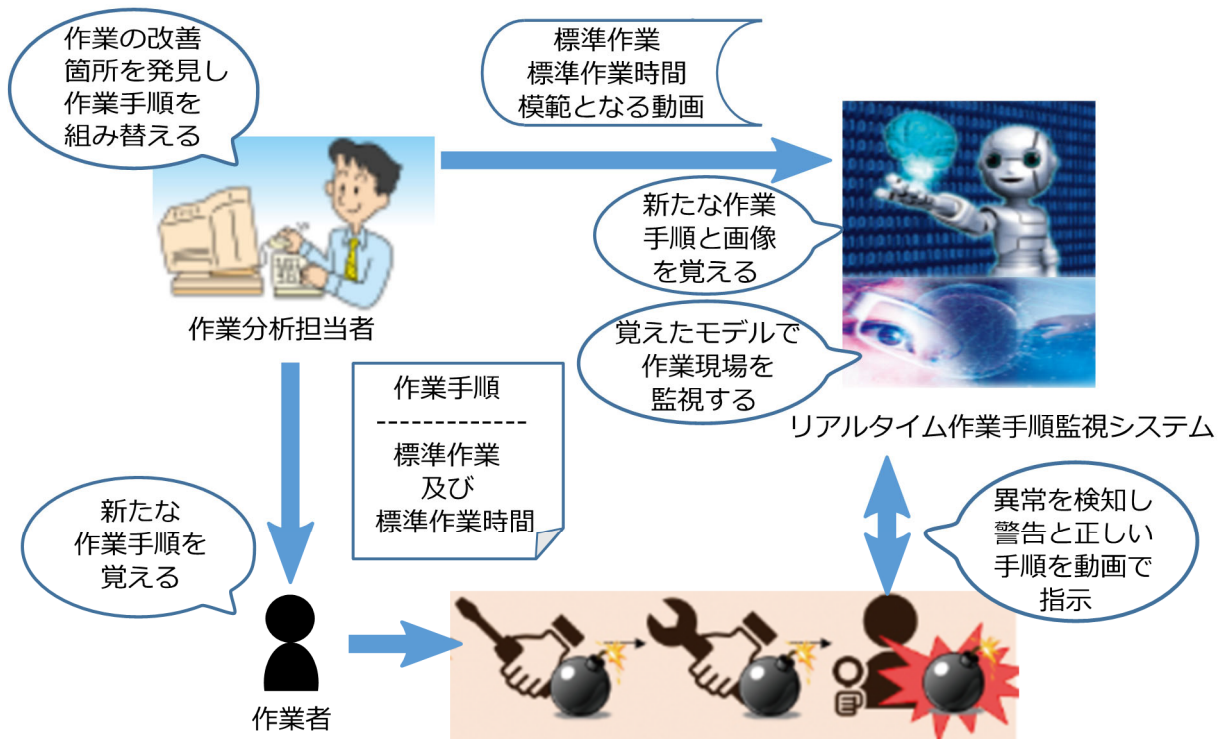


図 4-2 ユースケース

#### 4.2.4 市販の作業分析ソフトでの分析結果の利用

作業分析ソフトについては市販の製品Aを使用した。製品Aでは撮影したビデオ映像を人手による簡単な操作で作業単位に分割し、正確な時間測定や動作の分析が可能である。さらに、ムダな作業の排除や作業手順の組み換えを行い、その結果として標準作業と各作業の標準作業時間が出力される。この分析結果と分析の際に使用された作業動画を本システムで利用する。

#### 4.2.5 リアルタイム性の高いCNNモデルの採用

本システムでは限られた学習データでCNNを効果的に利用できるようにするために、学習済ネットワークを転移学習として利用する。CNNによる分類用の学習済ネットワークモデルとしてはGoogLeNet[4]やResNet[5]など多くの種類が存在するが、今回はリアルタイム性を要件としているため、Canzianiらの評価結果[6]を参考に、比較的スループットが高いとされているAlexNet[7]を採用した。AlexNetは2012年に登場した比較的初期のネットワークモデルであるが、当時のImageNet ILSVRC-2012チャレンジで2位に大差をつけて優勝したモデルである。その構造は、5つの畳込み層（conv1~conv5）と3つの全結合層（fc6, fc7, fc8）からなる深さが8層のニューラルネットワークで、120万枚の高解像度画像を1,000種類のクラスに分類するために学習されたネットワークである（図4-3、図4-4）。

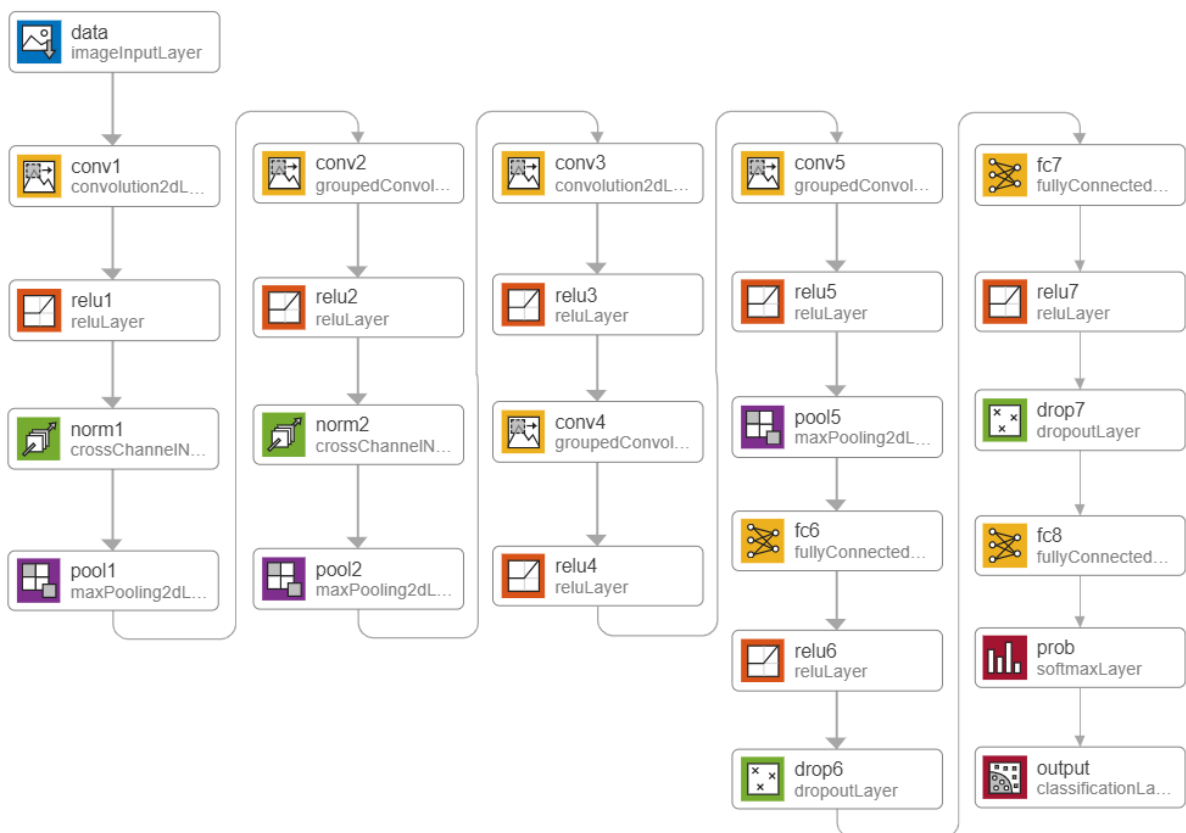


図 4-3 AlexNetのネットワーク構成 (MATLAB Deep Network Designer)



図 4-4 AlexNetの出力部分の構成 (MATLAB Deep Learning Network Analyzer)

AlexNetの5つの畳み込み層の構造としては、第1層で $227 \times 227 \times 3$ の入力画像を $11 \times 11 \times 3$ の96個のカーネルを使用して畳み込む。第2層では第1層の結果を $5 \times 5 \times 48$ の256個のカーネルを使用して畳み込み、第3層では第2層の結果を $3 \times 3 \times 256$ の384個のカーネルを使用して畳み込む。第4層と第5層は、それぞれ前の層の結果を $3 \times 3 \times 192$ の384個、256個のカーネルを使用して畳み込む。また、AlexNetは過学習を緩和するために2つのドロップアウト (Dropout) 層を利用している。Dropoutは2つの全結合層で利用されており、それぞれの重みを50%の割合でランダムにゼロに設定している。さらに、汎化性能をあげるために、局所応答正規化層 (Local Response Normalization : LRN) を第1層と第2層に利用して精度の向上を図っている。LRNは競合正規化ステップとして、最も強く活性化しているニューロンが近隣の特徴量マップの同じ位置のニューロンを活性化させることを禁止するものであり、その結果、異なる特徴量マップが特殊化して互いに離れていき、広い範囲の特徴を探ることにより、究極的には汎化性能を向上させるものである[8]。なお、今回利用したMATLABにおけるLRNレイヤーは、チャンネル単位の局所応答正規化層 (crossChannelNormalizationLayer) と呼ばれる。入力の各要素 $x$ に対する正規化された値 $x'$ について式(1)を使用して計算される。ここで $K$  (バイアス),  $\alpha$  (係数),  $\beta$  (指数) はハイパーパラメータで $ss$ は正規化ウィンドウ (近隣の対象チャンネル) の要素の二乗和となる。また、 $windowChannelSize$ は各要素の正規化に使用されるチャンネルのウィンドウサイズとなり、このサイズが5の場合は、対象チャンネルは自チャンネルと前方と後方にそれぞれ2チャンネルの近傍によって正規化される[9]。なお、今回のネットワークではcrossChannelNormalizationLayerが2カ所 (norm1, norm2) 存在するが、いずれも $K = 1, \alpha = 1e^{-4}, \beta = 0.75, windowChannelSize = 5$ の設定で利用した。

$$x' = \frac{x}{\left(K + \frac{\alpha * ss}{windowChannelSize}\right)^\beta} \quad (1)$$



AlexNetは近年の他のネットワークモデルと比較するとその精度は決して高くはないが、その後登場した各種CNNの原型となった代表的なネットワークモデルである。

### 4.3 リアルタイム作業手順監視システムの構成

#### 4.3.1 ハードウェア構成およびシステム構成

本システムのハードウェア構成は図 4-5のとおりである。CNNの学習およびリアルタイム監視はノートパソコン上で実行され、Webカメラについては小型でフォーカスタイプ固定の一般的な市販品を利用しUSBケーブルでパソコンに接続する。また、アプリケーションについてはMathworks社のMATLAB R2019aとそのオプション製品であるDeep Learning Toolboxを利用して開発した[10]。なお、本システムは製造現場で容易に利用することができるよう、利用するパソコンについてもGPUは使用せず、Webカメラも含め比較的安価に構築することが可能な構成とした。

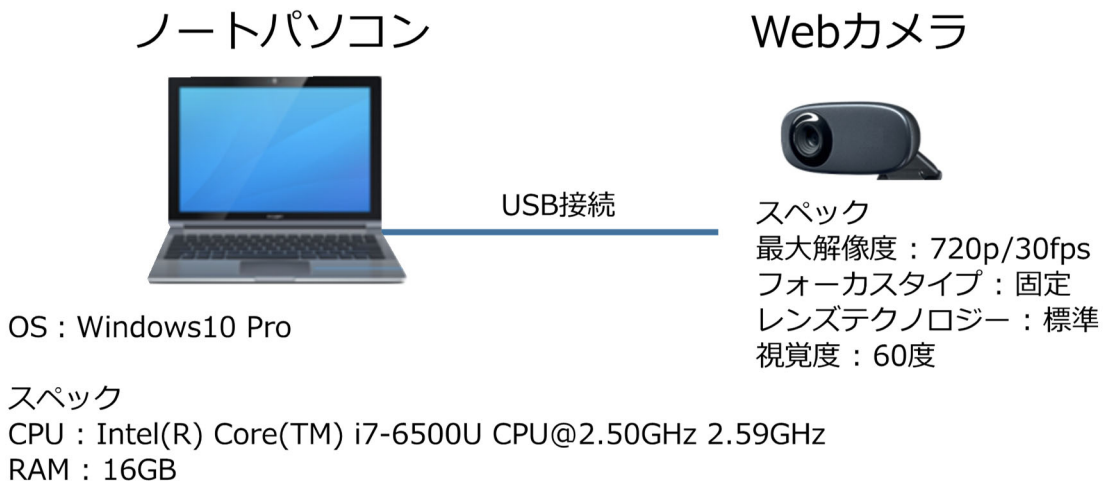


図 4-5 ハードウェア構成

#### 4.3.2 アプリケーション

##### (1)作業分析ソフトからのデータ取り込み

作業分析ソフト製品Aでは、製造現場でビデオ撮影された動画をパソコン上で再生しながら、人手による簡単な操作で作業を分類し、各種分析を行うことができるソフトウェアである。また、分析により問題点の抽出や標準作業・マニュアルの作成も可能である。作業分析を行った結果については、ExcelファイルまたはCSVファイルとして出力することが可能であり、表 4-1にそのサンプルイメージを示す。今回はこの情報をシステムに取り込み、標準作業の手順および、標準作業時間として監視の基準値とする。また、作業分析時に使用した作業動画からCNNの学習用の静止画を抽出する際条件としても本分析結果を利用する。

表 4-1 作業分析ソフトでの出力結果

No.	計測時間 (秒)	作業分類
1	8.9	①金具を運ぶ
2	25.8	②仮止め
3	21.1	③ボルト締め付け
4	22	④マット貼り付け
5	18.2	⑤プレート拭き上げ
6	20.7	⑥金具固定
合計	116.7	

(2)CNN学習用データの生成

CNNの学習用データを生成する際は、表 4-1にある作業順番 (No.) と計測時間、作業分類に基づいて動画から静止画を抽出する (図 4-6)。具体的には、前記(1)で人手により分析された作業分析結果内の各作業に対する計測時間 (各作業の経過秒数) を取得し、その情報と動画の各ビデオフレームに対するタイムスタンプを比較しながら、該当する計測時間内のフレームを静止画として抽出し、あらかじめ準備しておいた各作業の名称を付けたフォルダに各ファイルを保存する処理をシステム化により自動化する。これにより、作業分析結果から学習用データを生成する処理の自動化を実現する。

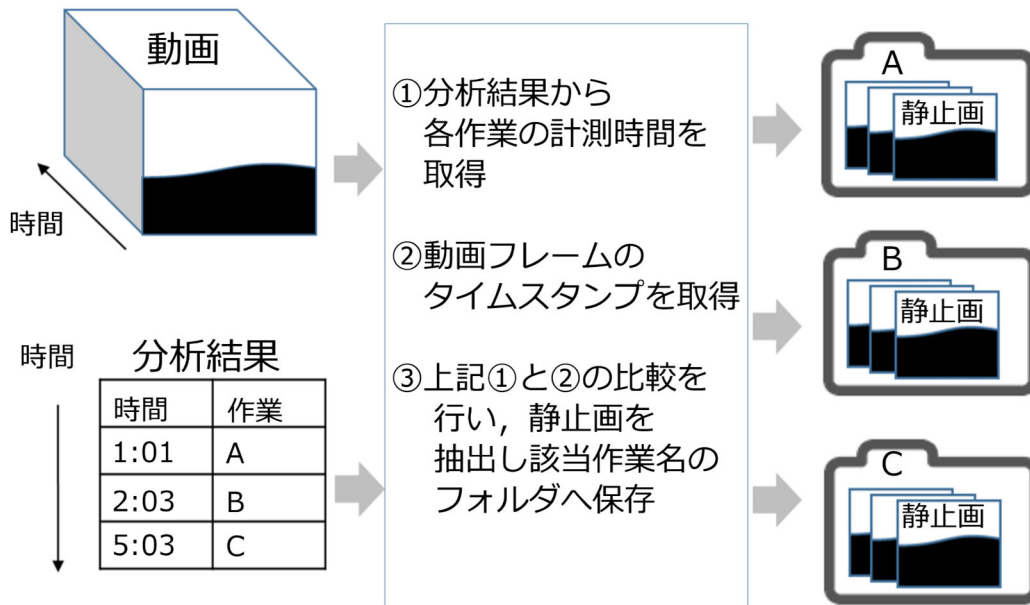


図 4-6 動画からの静止画の抽出処理

(3)CNN転移学習

AlexNetの最後の全結合層およびSoftmax層，出力層について今回の作業分類対象にあわせて分類クラス数の変更を行い，生成した学習用データにより転移学習させることにより，作業認識用のネットワークとする．具体的には，図 4-4に示したAlexNetの最後の全結合(fc8)，ソフトマックス(prob)，分類出力(output)のレイヤーを削除し，そこに図 4-7に示すとおり，今回の分類対象のクラス数と同一となる全結合層 (fc)，ソフトマックス層 (softmax)，分類出力層 (classoutput) を追加し，今回準備した画像を用いて学習を行う．



図 4-7 転移学習用に調整したAlexNetの出力部分 (MATLAB Deep Learning Network Analyzer)

なお，本ネットワークにおける誤差関数は式(2)の交差エントロピー (Cross Entropy) を用いている． $K$ は分類対象のクラスの数， $t_i$ は教師データがクラス*i*に属するインジケータ， $y_i$ はクラス*i*に対するネットワークからの出力確率である．

$$E = - \sum_{i=1}^K t_i \log y_i \quad (2)$$

(4)撮影中の製造現場の作業分類と作業時間集計

製造現場の作業については，カメラで撮影中の画像をフレームごとに静止画として抽出し，学習済のCNNにより各作業クラスに分類を行う．

また，作業時間の計測については，静止画が各作業に分類され，作業が切り替わった際のパソコン内の時刻を用いて経過時間として集計する．

#### (5) 作業手順誤り検知およびティーチング

作業手順誤りの検知については、標準作業に対する作業の順番とCNNで分類された撮影中の作業の順番を比較することで判断を行う。

また、監視対象となる作業において標準作業から大きく逸脱した動作や標準作業とはまったく異なる部品配置など、CNNの学習時に想定していなかった状況が発生した場合、該当作業に対しても手順の誤りとして検知の対象とする。これを判断するために、CNNの分類時の確率スコアを利用する。CNNのsoftmax層から出力された結果は、分類対象となるクラスのなかで最もスコアが高いクラスへ分類されるため、学習データに含まれていない画像に対してもいずれかのクラスへ分類されてしまう。そこで分類時の確率スコアを参照し、一定スコア以下の場合には想定外作業として取り扱う。確率スコアの算出は式(3)に示すが、 $k$ は分類クラスの数、 $s(x)$ はインスタンス $x$ に対する各クラスのスコアからなるベクトル、 $p_k$ はインスタンスに対する各クラスのスコアから推計されたインスタンスがクラス $k$ に属する確率となる。この $p_k$ の値が一定の閾値以下となる場合に、いずれの作業に対してもあてはまる確率が低い想定外作業として扱う。

$$p_k = \sigma(s(x))_k = \frac{\exp(s_k(x))}{\sum_{j=1}^k \exp(s_j(x))} \quad (3)$$

作業時間の乖離に関する検知については、標準作業時間と分類結果から集計された経過時間の比較を行い判断する。

作業手順の誤りもしくは作業時間の乖離が検知された際は、正しい作業の順番と該当する作業手順の動画（作業分析時に使用した模範の動画）をパソコンで再生することにより作業者へのティーチングを行う。

### 4.4 評価実験

#### 4.4.1 実験環境

設備の耐震固定に使われる金具の組立作業について、作業のサイクル全体（金具組立作業）とその中の特定作業（ボルト締め付け作業）を対象に、監視対象の粒度が異なる2種類の模擬環境を準備した。これは、作業分析においては、分析の目的とその性質に応じてそれに適した区分単位の大きさがあり[11]、作業監視対象についても、その区分単位の大きさが異なる場合が考えられるためである。

##### (1) 金具組立作業のサイクル全体に対する実験環境

金具組立作業としては、2種類の金具をボルトで仮止め後にスパナを使ってボルトの締め付けを行い、底面に設備と床を固定するための耐震用のマットを貼り付けた後、床を模擬したプレートに固定するなど、その他の作業あわせて合計6つの作業で構成される実験環境を準備した（図 4-8）。

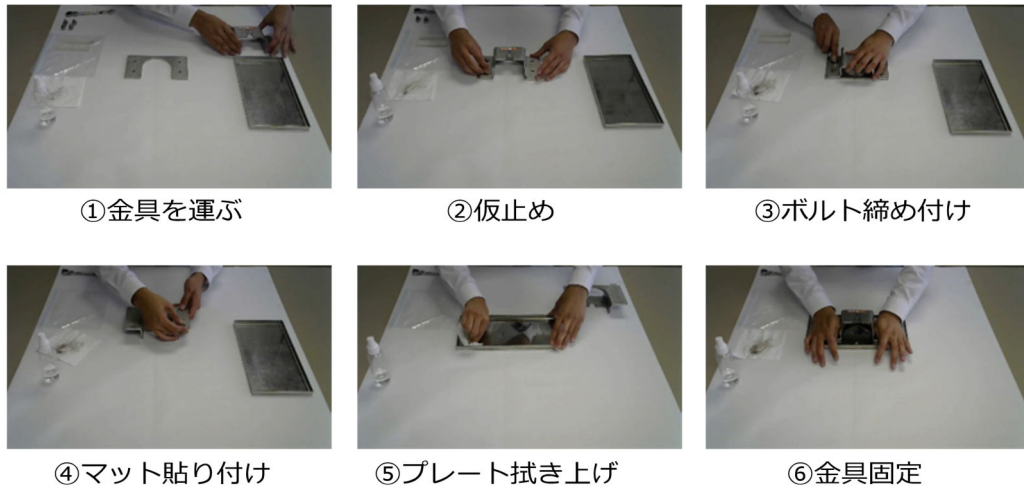


図 4-8 金具組立作業

(2) サイクル中の特定作業に対する実験環境

金具組立作業のサイクル全体の中の特定作業に関する実験環境としては、本組立作業のなかで最も重要な作業とされるボルト締め付けを対象とした。

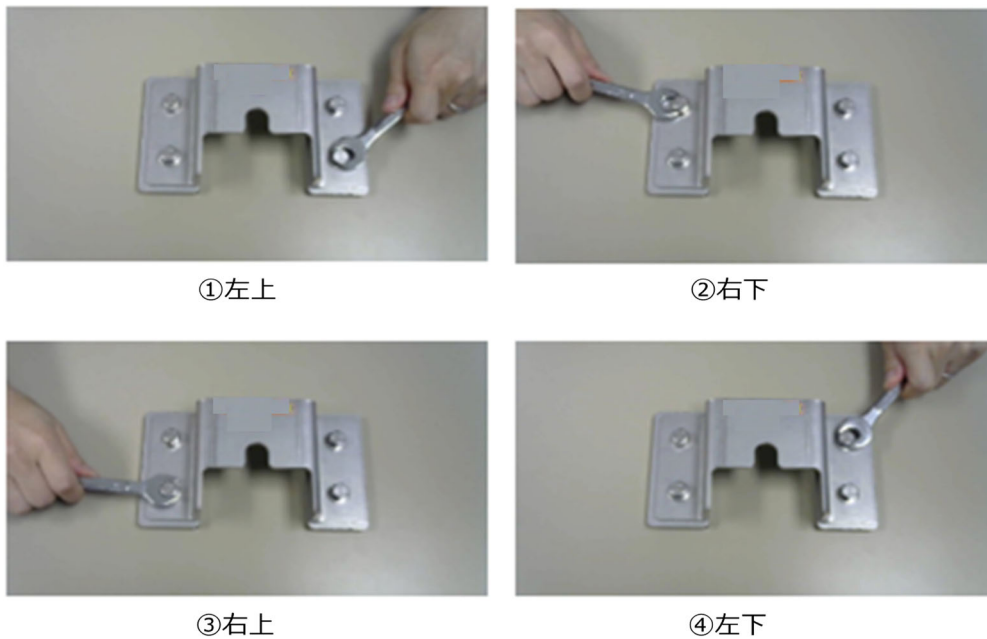


図 4-9 ボルト締め付け作業

図 4-9は作業者を対面方向からカメラで撮影した画像であるが、組み立てる際に金具の間に歪や狂いが生じると、地震の力によりボルトのせん断や引き抜き、金具の破断が生じる恐れがあり、金具を正しく組み立てるためには、ボルトの締め付け順が重要とされている。

#### 4.4.2 実験対象

本システムに対する実験対象については、以下の3点とした。1点目は、CNNにおける静止画の分類精度である。今回の要件の1つである作業手順の誤りおよび作業時間の乖離を正確に検知するためには、対象となる静止画に対する高い分類精度が求められる。本システムで自動生成された静止画で学習したAlexNetを用いて分類精度に関する実験を行った。

2点目は、作業手順や作業時間の乖離に対する検知精度であり、作業手順の誤りや手順の逸脱、標準作業時間との乖離が発生した際の検知の正確さについて実験した。

3点目は、処理のリアルタイム性である。本システムにおいてCNNで1秒間に分類処理できる静止画面数について求めることにより、その処理性能について把握するための実験を行った。あわせて、異常検知時のティーチング用動画再生までのリアルタイム性についても実験を行った。

#### 4.4.3 実験準備

##### (1) 学習データの生成

実験を行うにあたり、金具組立作業とボルト締め付け作業の各動画に対して作業分析ソフトで分析を行い、学習データを準備した。

先に、作業時の動画から学習用の静止画を準備する際の所要時間に関する評価を目的として、ボルト締め付け作業を対象に被験者3名の動画を用いて、本システムを利用して学習データを準備した場合と、従来の手作業で学習データを準備した場合のそれぞれの所要時間について計測を行った結果を表4-2に示す。平均23.91秒の動画に対して、本システムを利用した場合の所要時間は、作業分析ソフトでの分析時間とそれを利用した学習データ生成時間を合計して平均79.33秒となった。これに対して手作業で学習データを準備する際は、動画から静止画を抽出し、手作業でラベル付けを行う必要があるため、要した合計時間は平均363.33秒となった。本システムを利用することにより従来の手作業による学習データの準備と比較すると78%程度の時間削減が可能となった。

表 4-2 学習用データの生成処理性能

	動画時間 (秒)	静止画 (枚)	本システムでの学習データ準備時間 (秒)			手作業での学習データ準備時間 (秒)		
			作業分析ソフト での分析	学習データ 生成	合計	動画から 静止画抽出	手作業での ラベル付け	合計
被験者A	22.39	667	56	19	75	17	334	351
被験者B	20.26	604	54	15	69	15	310	325
被験者C	29.07	866	69	25	94	23	391	414
平均	23.91	712.33	59.67	19.67	79.33	18.33	345.00	363.33

次に、今回の実験に使用した金具組立作業の学習データ (表 4-3) の準備方法について示す。はじめに、熟練者にあたる被験者2名 (被験者A, B) をモデルに3サイクル分の作業を分析し、各サイクルに対してそれぞれ6つの作業に分類した。その分析結果と分析時の動画を本システムに取り込み、図

4-8に示すような静止画を生成した。次に作業ごとに2名の静止画面数の合計と、今回作業ごとに準備する学習用画像数である3,000枚との比率を求め、表 4-3の学習用画像採用率とした。この学習用画像採用率とそれぞれの作業者の作業ごとの静止画面数を掛け合わせて表 4-3の被験者ごとの学習用画像数を算出した。学習画像については、学習用画像採用率が100%未満の場合はそれぞれの作業者の作業ごとの静止画から学習用画像数の画像を無作為標本によりランダムに抽出し、100%を超える場合は同じく無作為標本により静止画からランダムに画像を選び学習用画像数となるようコピーを行った。

また、このような方法で準備した学習データについては、画像の類似性や重複によりデータに対して過適合を起こす要因となる可能性が考えられるため、MATLABの学習データに対するデータ拡張機能 (imageDataAugmenter[12]) を利用して対策を講じた。具体的には、学習用の画像に対して水平方向と垂直方向へそれぞれ20ピクセルの範囲内でランダムに移動させることにより、データのバリエーションを増やすようにした。

なお、被験者Aと被験者Bの間で学習用画像数に差が生じているが、これは各作業者の作業ごとの作業時間が異なっていたことにより静止画面数にも違いが生じたことが原因であり、被験者ごとの静止画面数に対して均一の割合となるように、それぞれの静止画から学習用の画像を抽出したことによるものである。

表 4-3 金具組立作業の学習用データ画像数 (枚)

	データの区分	①金具を運ぶ	②仮止め	③ボルト締め付け	④マット貼り付け	⑤プレート拭き上げ	⑥金具固定	合計
被験者A	静止画面数	819	3,312	2,704	2,068	1,889	3,560	14,352
	学習用画像数	1,490	1,722	1,812	1,654	1,587	2,029	10,294
被験者B	静止画面数	830	2,458	1,773	1,683	1,682	1,704	10,130
	学習用画像数	1,510	1,278	1,188	1,346	1,413	971	7,706
合計	静止画面数	1,649	5,770	4,477	3,751	3,571	5,264	24,482
	学習用画像採用率	182%	52%	67%	80%	84%	57%	-
	学習用画像数	3,000	3,000	3,000	3,000	3,000	3,000	18,000

また、ボルト締め付け作業の学習データについては手元の画像を対象としているため、被験者間で動作や特徴にバラツキが少ないものと仮定し、熟練者である被験者1名(被験者A)をモデルに10作業分を対象とした。分析結果に基づいて表 4-4のとおり、4カ所のボルト締めの動作と、ボルト締め対象箇所間の手や工具の移動を“その他”とした計5つのクラスに分類し、図 4-9のような学習用の静止画をそれぞれ3,000枚準備した。その際、金具組立作業同様に被験者の静止画に対して無作為標本によりランダムに抽出またはコピー、およびMATLABによるデータ拡張処理を適用した。

また、“その他”をクラスに追加した理由は、それぞれの位置に対するボルト締め付けの動作と、それ以外の手の移動などの動作を明確に分けるためである。

表 4-4 ボルト締め付け作業の学習用データ画像数 (枚)

	データの区分	①左上	②右下	③右上	④左下	⑤その他	合計
被験者A	静止画像数	1,430	1,350	1,520	1,480	4,012	9,792
	学習用画像採用率	210%	222%	197%	203%	75%	-
	学習用画像数	3,000	3,000	3,000	3,000	3,000	15,000

(2)CNNの学習

金具組立作業、ボルト締め付け作業に利用したそれぞれのCNNには、オプティマイザにSGDM（モーメント項付き確率的勾配降下法）を用い、モーメント項の寄与率は0.9とした。また、転移層の学習率を減らし、新たに追加した層に対する学習率を増やすために、初期学習レートには $1.00E^{-04}$ を設定し、追加した層の学習レートはその10倍（ $1.00E^{-03}$ ）の学習レートを設定した。

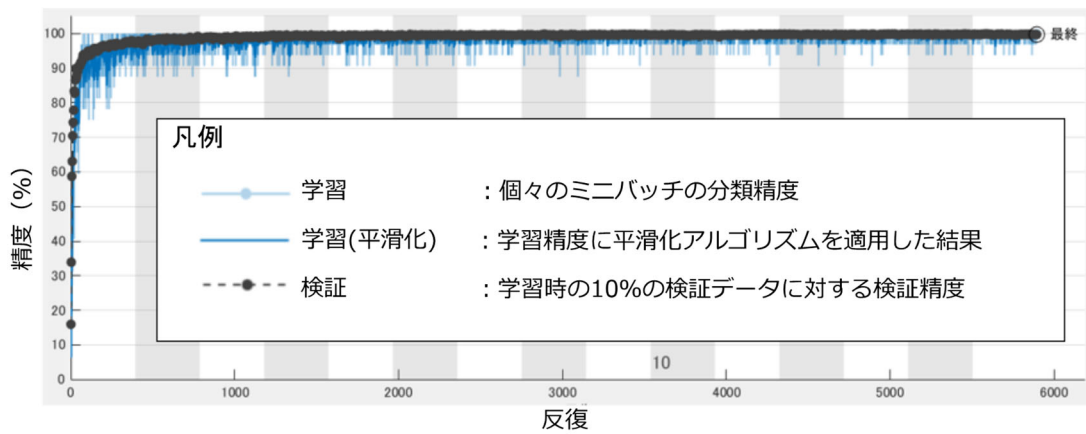


図 4-10 金具組立の学習結果

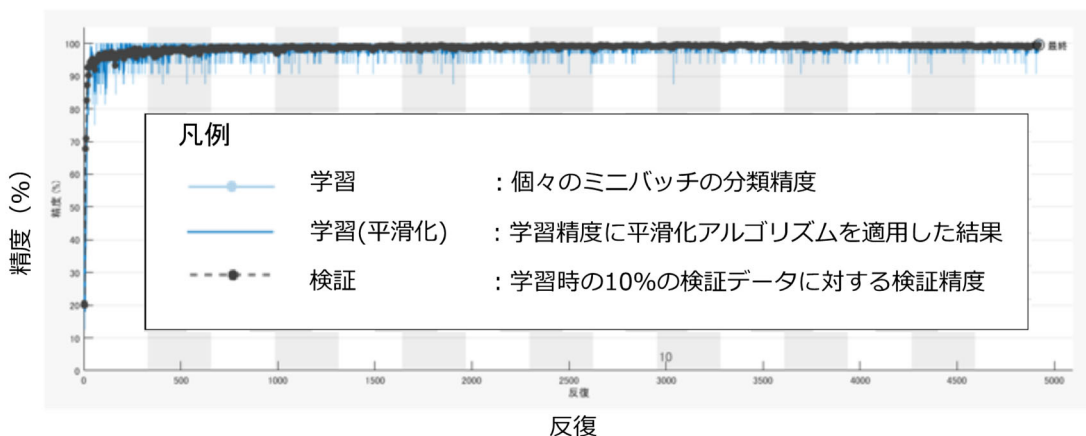


図 4-11 ボルト締め付け作業の学習結果



学習は、ミニバッチサイズ32で15エポック繰り返し、学習データを有効に利用できるよう、エポックごとにデータをシャッフルするようにして学習を行った。図 4-10の金具組立作業に関する学習曲線、図 4-11のボルト締め付け作業に関する学習曲線が示すとおり、学習が進み収束していることが確認できる。学習時のデータの10%を検証データに使用して学習時の検証精度を確認したところ、金具組立作業は99.72%、ボルト締め付け作業は99.53%であった。

#### 4.4.4 実験内容

##### (1) 静止画に対する分類精度

金具組立作業のサイクルに対して、学習データに含まれていない非熟練である被験者2名（被験者C,D）の1サイクル目の作業をテストデータとして静止画に対する分類精度を評価することを実験内容とした。このテストデータについては学習データ同様にあらかじめ作業分析ソフトで生成したものを利用し、AlexNetでの分類結果と正解ラベルを比較することにより分類精度を求める。ボルト締め付け作業についても、学習データに含まれない非熟練の被験者2名（被験者C,D）の作業をテストデータとして準備した。

##### (2) 作業手順誤りおよび作業時間乖離に対する検知精度

作業手順の誤りおよび作業時間の乖離に対する検知精度の実験については、実際に作業手順の誤りが生じる恐れがあるボルト締め付け作業内の手順を対象とした。ボルト締め付けには締め付け箇所（左上->右下->右上->左下）が決まっており、図 4-12のとおりとなる。

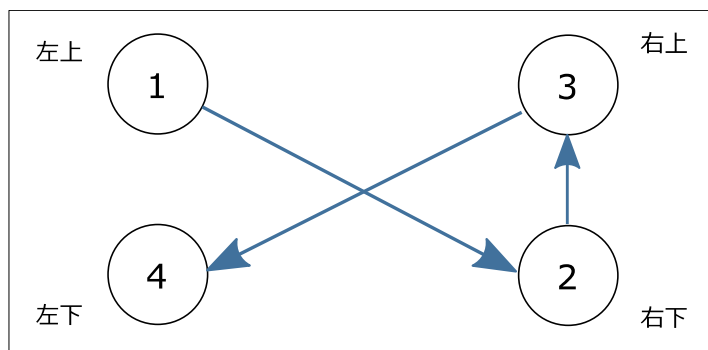


図 4-12 ボルト締め付けの順番

作業手順の誤りに対する検知については、被験者Aおよび学習時に参加していない非熟練の被験者2名（被験者C,D）の作業状況をカメラで撮影しながら、実際に作業手順の誤りが発生した際の検知の正確さについて評価することを実験内容とした。また、標準作業から大きく逸脱した想定外作業に対する検知については、既定の作業と想定外作業の分類時の確率スコアの傾向から異常検知のための閾値を定めることが可能であるかを評価することを実験内容とした。なお、想定外作業については金具組

立作業に対してプレートと工具の位置を左右逆にした状況を模擬的に作り出すために、テスト用の画像を左右反転させた画像を100枚準備して分類スコアの評価を行った。

標準作業時間と実際の作業時間の乖離に関する検知については、熟練者の作業分析結果から定めた標準作業時間に対して、本システムで計測・集計された時間を比較することを実験内容とした。実験の際は作業手順の誤りに対する検知と同様、被験者3名の作業をカメラで撮影しながら実施した。

なお、CNNにおける分類精度が100%には至らないことを想定し、瞬間的な誤分類を対象として同一分類結果が $n$ フレーム連続しなかった場合に、その前の $n \times 2$ フレーム間での最頻値で分類結果の置き換えを行うスライディングウィンドウ方式によるフィルタ処理（図 4-13）を加えた。なお、今回は、 $n = 5$ の条件で評価実験を行った。

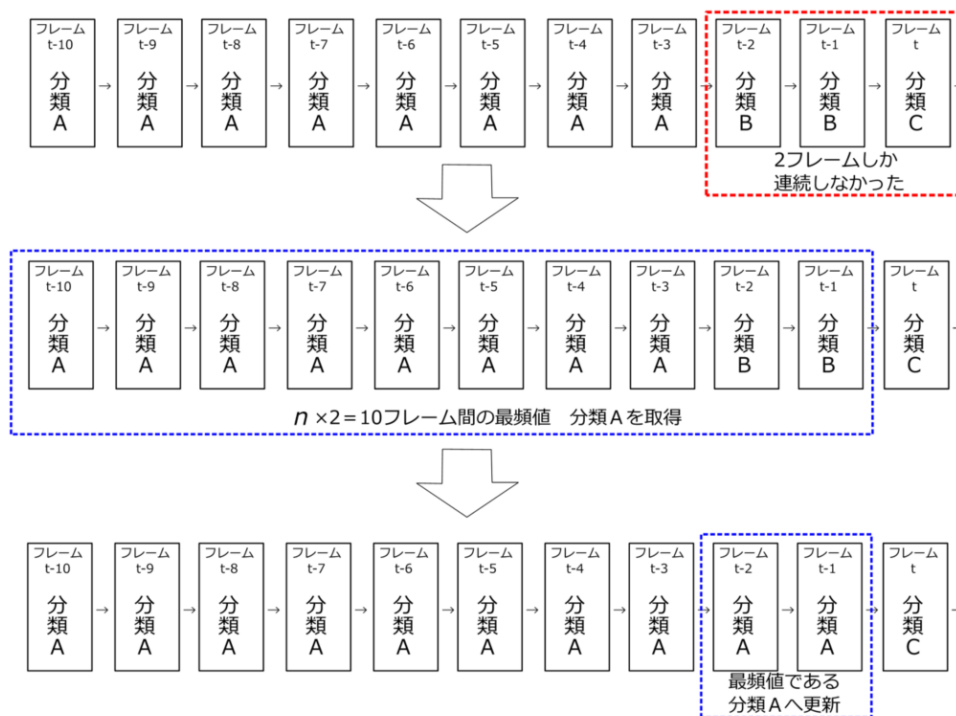


図 4-13 スライディングウィンドウ方式によるフィルタ処理 ( $n = 5$ の場合)

### (3)リアルタイム性

処理のリアルタイム性評価については、CNNでの作業分類処理と、異常検知時のティーチング用動画再生までの処理を実験対象とした。CNNでの分類処理に関しては、前記“(1)静止画に対する分類精度”の実験の際に使用する画像を使用してAlexNetで分類する際に要した処理時間と処理画像数から1秒あたりの平均処理画像数を算出した。異常検知時のティーチング用動画再生までの処理時間については、対象の作業を模範となる作業動画から検索し、該当フレームから再生を開始するまでの処理時間を計測した。

4.4.5 実験結果

(1) 静止画に対する分類精度

金具組立作業の静止画に対する分類精度の実験結果を表 4-5に示す。この表は分類結果を混同行列で表したもので、行（縦）は実際の正解ラベルのクラス、列（横）はCNNで予測されたクラスであり、各クラスの正解率を対角要素で示したものである。全体の結果として、2名平均の分類精度については93.08%であった。

ただし、被験者Cの“①金具を運ぶ”や“⑤プレート拭き上げ”のように、一部の作業において分類精度が低い箇所が見られた。これらは正解の作業に対してその前後の作業に誤分類しているため、作業間で類似した静止画が存在していることが原因である。図 4-14に被験者Cの“④マット貼り付け”に誤分類した“⑤プレート拭き上げ”の画像と“④マット貼り付け”の終了時点の画像を一例として示す。

表 4-5 金具組立作業の分類精度

被験者C	①金具を運ぶ	②仮止め	③ボルト締め付け	④マット貼り付け	⑤プレート拭き上げ	⑥金型固定
①金具を運ぶ	81.17%	17.86%	0.97%	0.00%	0.00%	0.00%
②仮止め	0.00%	93.89%	6.11%	0.00%	0.00%	0.00%
③ボルト締め付け	0.00%	0.00%	97.39%	2.61%	0.00%	0.00%
④マット貼り付け	0.00%	0.24%	7.77%	91.99%	0.00%	0.00%
⑤プレート拭き上げ	0.00%	0.00%	0.00%	8.01%	84.44%	7.55%
⑥固定	0.00%	0.00%	0.00%	0.00%	4.03%	95.97%
平均						90.81%

被験者D	①金具を運ぶ	②仮止め	③ボルト締め付け	④マット貼り付け	⑤プレート拭き上げ	⑥金型固定
①金具を運ぶ	100%	0.00%	0.00%	0.00%	0.00%	0.00%
②仮止め	3%	96.98%	0.00%	0.00%	0.00%	0.00%
③ボルト締め付け	0%	1.27%	96.91%	1.82%	0.00%	0.00%
④マット貼り付け	0%	0.16%	0.78%	96.43%	2.48%	0.00%
⑤プレート拭き上げ	7%	0.00%	0.00%	7.41%	85.19%	0.74%
⑥固定	0%	0.00%	0.00%	0.00%	3.47%	96.53%
平均						95.34%

“④マット貼り付け”に誤分類した  
“⑤プレート拭き上げ”の画像



“④マット貼り付け”の  
作業の終了時点の画像



図 4-14 前後の作業への誤分類ケース

また、ボルト締め付け作業の分類結果を表 4-6に示す。

表 4-6 ボルト締め付け作業に対する分類精度

被験者C	その他	右下	右上	左下	左上
その他	90.13%	3.95%	1.97%	3.95%	0.00%
右下	0.78%	99.22%	0.00%	0.00%	0.00%
右上	3.36%	0.00%	96.64%	0.00%	0.00%
左下	0.98%	0.00%	0.00%	99.02%	0.00%
左上	0.98%	0.00%	0.00%	2.94%	96.08%
平均					96.22%

被験者D	その他	右下	右上	左下	左上
その他	91.79%	4.35%	1.45%	0.97%	1.45%
右下	0.00%	100.00%	0.00%	0.00%	0.00%
右上	1.22%	0.00%	98.78%	0.00%	0.00%
左下	2.99%	0.00%	0.00%	97.01%	0.00%
左上	0.63%	0.00%	0.00%	1.27%	98.10%
平均					97.14%

2名平均の分類精度は96.68%であった。各ボルト締めの対象箇所と“その他”として定義した対象箇所間の移動動作との間で誤分類が見られるが、この原因は前記のとおり類似した静止画が両方に含まれているためである。

(2)作業手順誤りおよび作業時間乖離に対する検知精度

被験者3名をWebカメラで撮影しながら、4カ所のボルトの締め付け順を表 4-7のとおり7パターンで評価した。これは4カ所のボルト締め付け手順全24パターンの中から、正常1パターンと、誤った手順が最初に検知される6パターンである。この6パターンを異常パターンとしてあげた理由は、本システムでは一連の作業手順のなかで最初に誤りが認められた時点を検知対象としているためである。実験の結果、3名の被験者ともに、正常1パターンは正しい手順として判断し、誤った6パターンについては最初に誤りが発生した作業を対象として異常として100%検知することができた。

表 4-7 ボルト締め作業手順確認パターン

パターン	手順1	手順2	手順3	手順4	最初に異常を検知する手順
1	左上	右下	右上	左下	正常のため検知せず
2	左上	左下	-	-	手順2の左下
3	左上	右上	-	-	手順2の右上
4	左上	右下	左下	-	手順3の左下
5	左下	-	-	-	手順1の左下
6	右上	-	-	-	手順1の右上
7	右下	-	-	-	手順1の右下

また、既定の正しい作業と想定外の逸脱した作業に関する確率スコアの比較結果を図 4-15に示す。ヒストグラムa)は既定の作業で平均の確率スコアは98.14%、同じくb)は想定外の作業で平均の確率ス

コアは67.47%となった。このヒストグラムで示されたとおり、両者の確率スコアには大きな開きが見られた。

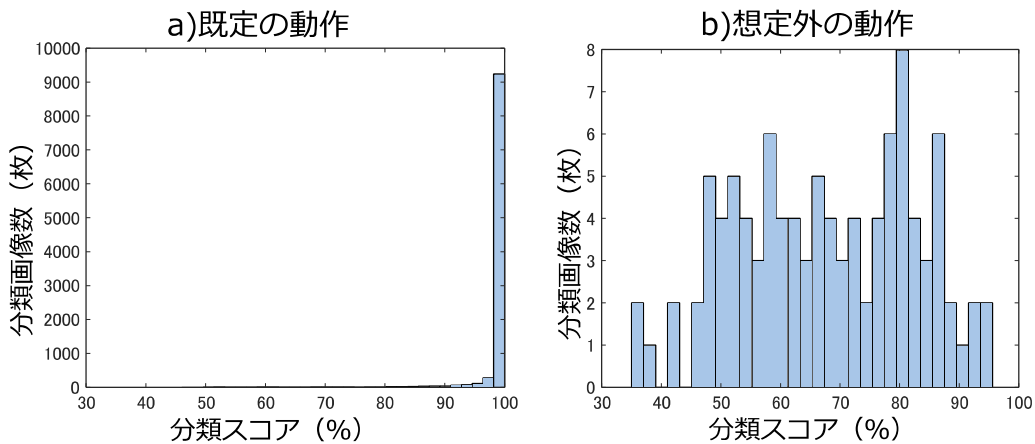


図 4-15 確率スコア比較

標準作業時間に対する実際の作業時間の乖離の比較については表 4-8に示す。今回、標準作業時間をボルト締め1ヵ所あたり4.5秒に定め、実際の作業時間の算出については、Webカメラで撮影中の画像に対してCNNで各作業に分類された際のそれぞれの作業の作業開始時刻から作業終了時刻までの経過時間を用いた。

監視のための閾値については本作業に従事している担当者に対してヒアリングを行い、ボルト締め付け不足や作業に無駄が含まれない範囲として標準作業時間に対して±50%とした。具体的には、標準作業時間の50%以下の時間で作業が完了する場合は、作業が不完全でボルトの締め付けが甘い可能性が考えられるため異常として判断し、標準作業時間の150%以上の時間を要する場合は、作業に無駄があると考えられるため異常として判断するという理由からである。評価の結果、被験者3名ともにそれを超えることはなかった。

表 4-8 標準作業時間と実作業時間比較

	手順 1 (秒)	手順 2 (秒)	手順 3 (秒)	手順 4 (秒)	平均 (秒)
標準作業時間	4.5	4.5	4.5	4.5	4.5
被験者A	3.83	3.62	4.11	5.09	4.16
被験者C	4.86	4.08	4.48	5.17	4.65
被験者D	4.90	3.83	4.56	4.67	4.49
平均	4.53	3.84	4.38	4.98	4.43

### (3)リアルタイム性

AlexNetでの分類に関するリアルタイム性を確認するために、動画をフレームごとに抽出しながらAlexNetで分類した際の1秒あたりの処理数について確認した結果を表 4-9に示す。1秒あたり15回弱

の処理が可能であることが確認できた。

表 4-9 AlexNetでの静止画分類処理数

	所要時間 (秒)	分類処理数 (回)	分類処理数@秒
被験者A	260	3761	14.47
被験者B	214	3322	15.52
被験者C	239	3499	14.64
平均	237.67	3527.33	14.88

また、ティーチング用動画再生までの処理時間については、ボルト締め付け作業の4カ所に対する異常検知から対象作業の作業動画再生までの処理時間を10回計測したところ、いずれも1秒未満であった。

#### 4.5 考察

4.2.1節で示した4つの要求条件に関して、4.3節のシステム構成および4.4節で示した評価実験結果を基に考察する。

##### (1) 作業改善を目的とした客観的で合理的な監視基準であること

作業改善のための分析結果である標準作業と標準作業時間および分析の際の画像を、CNNの学習用データ、監視基準およびティーチングの際の正しい作業手順として利用することにより、本要件をシステムとして実現した。

##### (2) CNNの学習データの準備が容易であること

実験の準備段階で、製品Aの作業分析結果とその際に使用した動画から学習データの準備を行った。本システムを利用して学習データを準備した場合と、従来の手作業で学習データを準備した場合の作業時間について評価した結果、本システムを利用することにより従来に比べ78%程度の時間削減が可能であることが確認できた。従来の方法である手作業でのラベル付けに時間を要した理由は、連続した静止画を目視で見比べながら作業の区切りとなる画像を見つけることが煩雑なためである。一方、筆者らの方法であれば作業分析ソフトとの組み合わせによりこれらの作業を効率的に行うことが可能であり、学習データの準備を容易にした。

製品Aでは分析結果がExcelとCSVで出力されるが、外部にデータを出力する仕組みがないような作業分析ソフトについてもロボティック・プロセス・オートメーション (RPA) を利用することにより分析結果の表示画面から情報の抽出は可能である。本研究の手法を利用することにより、容易に学習データを準備することが可能である。

##### (3) 作業手順誤りや時間乖離を正確に検知できること

CNNによる静止画の分類精度が平均90%以上であったこと、および作業手順ミスを検知が正確であ

ったことにより、本要件を満たすことが確認できた。

標準作業から大きく逸脱した想定外作業を検知する方法については、既定作業と想定外作業に関する確率スコアに大きく開きが見られたため、閾値の設定も可能であると考えられる。今回は閾値を定めたうえでの評価は行わなかったため、今後は閾値の妥当性の評価を行いたい。

また今回、CNNの瞬間的な誤分類に対してはスライディングウィンドウでフィルタをかけ、影響を抑えることができた。連続した一連の動作を正しく分類する方法として、画像を時系列シーケンスデータとして取り扱うリカレントニューラルネットワーク（RNN）の利用が挙げられるため、今後の検討に加えたい。

#### (4)製造現場をリアルタイムで監視可能であること

AlexNetを利用した際の分類については、毎秒15フレーム弱の処理ができることが確認された。本システムを製造現場で利用する場合、Webカメラで撮影しながら約0.07秒周期で画像を分類することができる。誤分類抑制のためのフィルタ処理におけるウィンドウ数 $n = 5$ に指定した場合、判断が5フレーム分遅れることになるが、その場合でも0.35秒遅れで異常の検知を行うことが可能である。また、異常検知時のティーチング用動画の再生についても、1秒未満で処理が可能である。したがって、製造現場で重大な作業ミスが発生した場合においても、即時に誤りを検知し、ティーチングを行うことが可能である。

上記のとおり、本システムの要件に対する有効性は確認できたと考えられるが、リアルタイム性の観点からは、カメラのフレームレートが毎秒30フレームの場合は0.03秒周期で画像を取得することも可能であるため、引き続きCNNモデルの軽量化などによりリアルタイム性の改善をはかりたい。

#### 4.6 まとめと今後の課題

本研究では作業分析結果と深層学習を利用して製造現場の作業手順をリアルタイムに監視し、誤りがあった場合には正しい手順をティーチングすることができる、汎用的な手法について開発を行った。作業改善を目的としたIEにおける作業分析の結果を監視基準として利用する点、その際の模範となった作業動画から監視用CNNの学習データを自動的に生成する点、それらを利用してリアルタイムに作業手順および作業時間の異常検知を行い、異常発生時に正しい作業手順でのティーチングを行う点を特徴としている。本稿では、金具の組立作業の模擬環境で作業分析を行い、分析結果と分析時の動画から生成されたデータで学習したCNNによる静止画に対する分類精度、作業手順および作業時間に対する異常検知精度、処理のリアルタイム性について評価実験を実施し、その有効性について示した。

本研究については、量産品に対する部品組み立て作業や同一製品に対する溶接作業のように、繰り返し行われるルーチン作業に対して応用することは可能であると考えられるが、複雑な製造現場での利用を考慮した場合、画像全体の特徴に基づく検知だけでは十分ではないケースも考えられる。具体的な例をあげると、対象となる製造物の形状が都度異なる場合や、手や工具の移動距離が短い作業、

動作に違いが少ない作業などのように、繰り返される作業のなかで共通した特徴自体が見出しにくいケースが想定される。これらのようなケースに対応するために、対象となる製造物や工具、作業者の手先など身体の一部を物体として検出し、作業分類のための特徴として組み合わせて利用する方法が考えられるため、引き続き検討を行いたい。

また、セル生産方式の製造現場へ適用する場合にも検討が必要である。たとえばU字型にレイアウトされたセル生産現場では、作業者は作業の進捗に合わせてセル内を移動しながら作業を行う。そのため、作業の手順としてもセル内での移動を対象として監視する場合と、セル内の個々の作業場所における作業内容を対象として監視する場合が考えられる。したがって、監視対象に応じて本システムについてセル全体をとらえることができる場所とセル内の個々の作業場所へ設置し、それぞれのシステムから得られた情報を連携することによりセル内での作業手順をとらえる仕組みが必要になるものと考えられる。この情報連携のあり方については引き続き検討を行いたい。



## 参考文献

- [1] 中村茂弘：現場力を高める簡単IE実践マニュアル， p.114， 日刊工業新聞社 (2011).
- [2] 中村茂弘：現場力を高める簡単IE実践マニュアル， p.36， 日刊工業新聞社 (2011).
- [3] 小林尚生， 南野充則：食事画像のラベル付け作業の省力化に関する考察， 第33回人工知能学会全国大会 (2019) .
- [4] Christian Szegedy, Wei Liu, Yangqing Jia, Pierre Sermane, Scott Reed, Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke, Andrew Rabinovich : Going deeper with convolutions, arXiv:1409.4842v1 [cs.CV], pp. 1-12 (2014).
- [5] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun.: Deep residual learning for image recognition, In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 770-778 (2016).
- [6] Alfredo Canziani, Eugenio Culurciello : An Analysis of Deep Neural Network Models for Practical Applications, Proc. 29th Conference on Neural Information Processing Systems, p.6 (2016).
- [7] Alex Krizhevsky, Ilya Sutskever, Geoffrey E. Hinton : ImageNet Classification with Deep Convolutional Neural Networks, Proc. Advances in Neural Information Processing Systems 25 (2012).
- [8] Aurelien Geron(著), 長尾高弘(訳) : scikit-learnとTensorFlowによる実践機械学習, pp.369-370, オライリー・ジャパン (2018) .
- [9] Mathworks : crossChannelNormalizationLayer, Mathworks, 入手先  
〈<https://jp.mathworks.com/help/deeplearning/ref/nnet.cnn.layer.crosschannelnormalizationlayer.html>〉 (参照 2019-07-10).
- [10] Mathworks : MATLAB-計算技術言語, Mathworks, 入手先  
〈<https://jp.mathworks.com/products/matlab.html>〉 (参照2021-05-13).
- [11] 嶋本康夫, 藤山紘一郎, 杉原寛, 他 : IE 7 つ道具, p.20, 日刊工業新聞社 (1993) .
- [12] Mathworks : imageDataAugmenter, Mathworks, 入手先  
〈[https://jp.mathworks.com/help/deeplearning/ref/imagedataaugmenter.html?searchHighlight=imageDataAugmenter&s\\_tid=srchtitle](https://jp.mathworks.com/help/deeplearning/ref/imagedataaugmenter.html?searchHighlight=imageDataAugmenter&s_tid=srchtitle)〉 (参照 参照2021-05-15).

## 第5章 回帰と分類用の深層学習を用いた製造現場での作業動作特徴抽出

### 5.1 はじめに

本章では、IE手法の1つである「動作研究」を対象として、End-to-Endの深層学習技術を用いることにより、作業時の特徴となる動作を抽出することができる汎用的なシステムについて開発を行った。本システムでは、回帰用の畳み込みニューラルネット（CNN）を利用して画像から作業者の姿勢や使用する工具などの画像座標系の座標位置を予測し、それらの位置と移動距離を時系列データとして、分類用のCNNにより作業動作として分類する。つまり、身体の一部や作業時に使用する工具などの物体の位置とその時系列変化から、特徴をあらわす作業動作を抽出する。本システムで抽出された特徴となる作業動作について、熟練技能者と非熟練者との間で比較することにより、その相違点から価値を生み出す動作とそうではない動作を定量的あるいは定性的に示すことが可能となり、効率的に改善活動を行うことができる。

なお、本論文における「汎用的」とは、従来の画像処理やセンサでの計測値に対する解析のように、作業内容や動作内容に合わせた専用の物体検出や動作に関する特徴量についての設計を必要とせず、異なる作業においても必要な情報を製造現場の画像から学習するという共通の手法により、利用することができるという点を指す。

### 5.2 作業動作特徴抽出システムの要件と特徴

#### 5.2.1 要求条件

本システムの目的は、作業時の動画から「動作研究」に役立つような特徴となる作業動作を抽出するための汎用的な仕組みの実現である。熟練技能者と非熟練者から抽出された作業時の特徴をあらわす作業動作を比較することにより、動作の違いや、効率的な動作とムダな動作を相対的に評価することが可能となる。そのために必要な要件について3点あげる。

要件1：身体や工具の検出対象箇所を正確に予測できる

要件2：作業時の動作を正確に分類できる

要件3：異なる作業においても汎用的に利用できる

製造現場の動画から作業者の特徴をあらわす作業動作を抽出するためには、その動作をとらえるための身体の一部や作業時に使用する工具などの物体の位置を、作業時の画像から正確に予測する必要がある（要件1）。

また、「動作研究」の対象となる身体の一部や工具などの物体の動きを、作業に含まれるいくつかの動作として正確に分類する必要がある（要件2）。なお、分類された結果である動作については、その分析手法として、両手作業分析の場合は「作業」「移動」「保持」「手持ち」の4つの動作に分類され[1]、微動作分析の場合は、サブリング記号と呼ばれる18の動作に分類される場合が多いが[2]、熟練技能者と非熟練者間の動作の違いや、効率的な動作、非効率的な動作の分析を行うためには、1つの

動作に対する特徴や、時間あたりの動作の大きさまでとらえることが重要である。

また、異なる作業において汎用的に利用するためには、その実現手段やアルゴリズムについて、従来からの画像処理やセンサを利用した解析手法のような作業内容や動作内容に合わせた専用設計をなくし、作業や動作への依存度を低くする必要がある（要件3）。

### 5.2.2 ユースケース

製造現場における一般的な作業の1つとして、スパナによるボルト締め作業を一例としたユースケースを示す（図 5-1）。

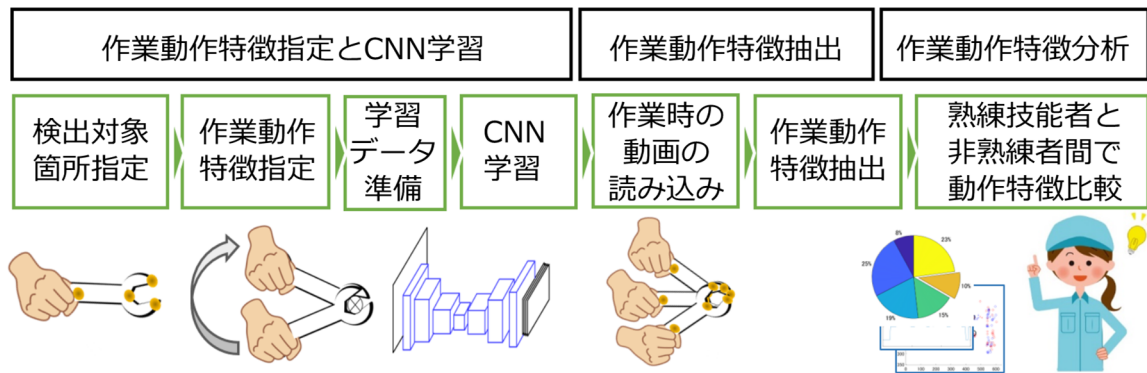


図 5-1 ユースケース

はじめに、ボルト締めの際の一連の動作から、特徴をあらわす動作を把握するために、使用する工具（スパナ）における検出対象箇所と、抽出したい作業動作を定める。次に、作業時の動画を元にしてCNNの学習データを準備しCNNの学習を行う。次に、学習されたCNNを用いて実際の作業時の動画から物体の位置の予測と動作の分類を行い、作業動作の特徴として抽出する。最後に、作業動作の特徴分析として、抽出された結果について、熟練技能者を基準として作業者間で比較することにより、価値を生み出す動作とそうではない動作を明確にし、改善箇所を効率的に把握する。

本ユースケースでポイントとなるのは、把握したい作業動作の特徴にあわせて、対象となる物体の検出対象箇所の指定と、特徴となる動作に関する指定を行うことができる点である。

### 5.2.3 2種類のCNNモデルの採用

本システムでは回帰用と分類用の2種類のCNNを利用している。まず、製造現場の画像から動作を把握するためには、その動作の対象となる身体部位や工具などの物体に関する位置情報が必要となる。たとえば、手作業での溶接作業においては腕や手首の関節位置や溶接治具が対象となり、組立作業においてはスパナや電動ドライバーなどが対象としてあげられるが、これらは工程や作業内容により異なり、その対象は多岐にわたる。これら対象物体の位置情報を画像から得るための手法としては、テ

ンプレートマッチング（パターンマッチング）やSIFT[3]などを利用した画像処理技術が以前から用いられてきたが、これらの手法は対象となる物体ごとに個別にテンプレートの準備やアルゴリズムの構築が必要となるために、導入対象に対して専用に設計されたシステムとなる。これに対して本研究では、物体の位置情報を得るために回帰用のCNNを用いた。作業者の様々な姿勢や多岐にわたる工具に対しても、その製造現場の画像を学習データとしてCNNの学習を行うことにより利用できるため、従来からの画像処理技術を利用した場合と比較すると、作業内容に対する依存度が低く、異なる作業においても汎用的に利用することが可能な手法である。

次に、特徴となる作業動作を抽出するためには、ある一定経過時間の時系列情報をそれぞれの動作に分類する必要がある。本システムでは身体や工具などの物体の検出対象箇所の位置情報および、連続した静止画間の検出対象箇所の移動距離を時系列データとして纏め、分類する仕組みとした。時系列データの分類に関する伝統的なアプローチとしては、時系列データ同士のユークリッド距離や類似度を比較して分類する手法や、動的時間伸縮法（Dynamic Time Warping : DTW）を利用して局所的なデータの歪を補正したうえで時系列データのパターンを比較して分類する方法[4]、幾つかの分類器をアンサンブルで組み合わせた方法[5]などについて研究が進められ、評価も行われてきた[6]。一方で、時系列データに対する分類に対してMultilayer Perceptron (MLP) , Fully Convolutional Networks (FCN) , Residual Network (ResNet) などの深層学習技術を用いる方法についての研究および評価も行われている[7][8]。本研究においても、データから学習することにより、幅広い製造現場でも利用することが可能となるよう、時系列データに対する分類についてもCNNを用いた。

上記のとおり、本システムでは特徴となる作業動作を抽出するために2種類のCNNを用い、End-to-Endで汎用的に利用できるようにした点が特徴である。

#### 5.2.4 本システムの利用条件

本システムの利用については、作業に関する標準化が行われている作業で、見込み生産や繰り返し生産が行われている製造現場を対象とする。本システムは深層学習技術を用いており、学習により一定の汎化性能を満たすことは可能であるが、学習データが限られている場合や、学習時と利用時で撮影環境や撮影条件が大きく異なる場合には、予測精度が低下する可能性が考えられる。したがって、本システムを利用する際の前提条件については、以下のとおりとする。

前提条件1：学習時と利用時でカメラの設置位置は同一

前提条件2：学習時と利用時で作業者の作業位置は同一

前提条件3：学習時と利用時でカメラの製品型式は同一

上記前提条件1について、学習時と利用時が同一製造現場の場合は、同一位置へカメラを設置することにより利用可能であるが、学習時と利用時が同一作業を行う異なる製造現場となる場合には、カメラと対象となる作業場所との位置関係を学習時の作業場所の位置関係と同一に合わせることで利用することが可能となる。

前提条件2については、IEの観点から標準作業が設定され、治工具や作業域内の配置は標準化されており、作業者が作業を行う場所や、治工具および作業域内の配置は標準化されている[9]。したがって、学習時と利用時は同一レイアウトであり作業位置は同一となるため、前提条件を満たすことができる。また、製造現場のレイアウト変更については、工程分析やレイアウト分析により改善の必要性が生じた際に実施するケースが多く、その場合は作業分析により作業手順の見直しなども一緒に行われるため、レイアウト変更後に改めて製造現場の画像を撮影しなおし、各深層学習モデルについても再学習することになる。

前提条件3については、学習時と利用時が同一製造現場であれば、通常は同一カメラを使用することになるが、使用するカメラを交換した場合や、同一作業を行う異なる製造現場で利用する場合には、同一メーカーの同一型式の製品を使用することで利用可能である。通常、画像処理を用いた距離の計測や物体の計測などについては、カメラキャリブレーションを行うことが求められる。カメラキャリブレーションとは、コンピュータ内部に設定したピンホールカメラモデルなどのカメラモデルと実世界のカメラとを関連付け、各種パラメータを推定することである[10]。このパラメータの代表的なもの、被写体の三次元ワールド座標系における座標を三次元のカメラ座標系へ変換する際に利用する外部パラメータと、三次元のカメラ座標系から二次元の正規化画像座標系へ変換する際に利用する内部パラメータがある。また、実際にカメラで撮影された画像はレンズ特性の影響で半径方向や円周方向に歪みが生じ、その歪み係数はカメラ固有の内部パラメータであり個体差が生じる。したがって、カメラで撮影された画像から対象物体のワールド座標系における位置や距離を測定するためには、カメラキャリブレーションによりこれらのパラメータ推定を行い補正する必要がある。一方、今回筆者らが用いた手法については、ワールド座標系の三次元座標ではなく二次元の画像座標系を使用しており、また、計測された物体の座標から直接動作を分類しているわけではなくCNNにより時系列に配置された各物体の座標が構成する特徴により分類を行っているため、同一型式のカメラが持つ製品ごとの個体差はCNNの汎化性能でカバーできる程、微小であると考えられる。したがって、カメラキャリブレーションによるパラメータ推定および補正処理は行わなかった。

上記3つの前提条件については、学習済のモデルについて、同一作業を行う異なる製造現場で利用する場合を含めた前提条件となるが、いずれも深層学習の汎化性能が重要である。したがって、学習済のモデルを、同一作業を行う他の製造現場へ展開する際には、モデルの汎化性能を向上させるために、パラメータの調整やオプティマイザの選択と合わせて、出来るだけ多くの製造現場のデータを学習データとして用いることが望ましい。

### 5.3 作業動作特徴抽出システムの構成

#### 5.3.1 ハードウェア構成およびシステム構成

本システムのハードウェア構成は図 5-2のとおりである。CNNの学習および作業動作の特徴抽出、評価についてはノートパソコンを使用した。Webカメラについては一般的な市販品を利用し、USBケ

ケーブルでパソコンへ接続する。アプリケーションについてはMathworks社のMATLAB R2020aとオープン製品であるDeep Learning Toolboxを利用して開発した[11].

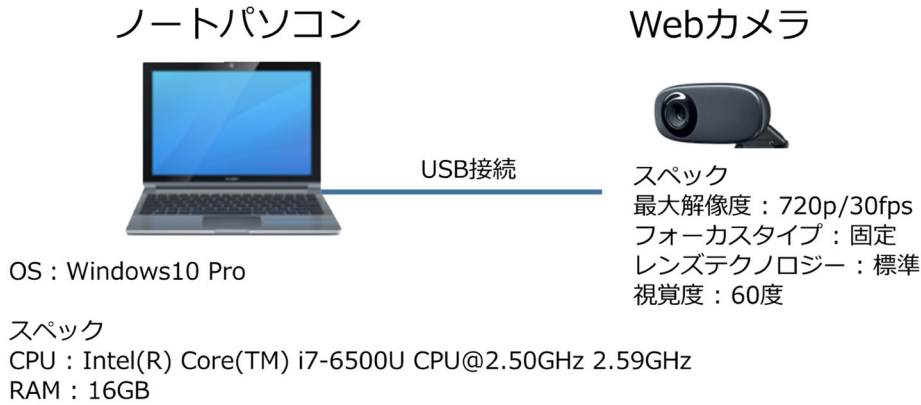


図 5-2 ハードウェア構成

### 5.3.2 身体や工具位置の予測処理

画像から身体や工具などの物体の位置を予測するためのCNNについては, Xiaoらが行った画像からの姿勢検出に関する研究[12]を参考にした. R-CNN (Regions with CNN features) [13]などを用いた物体検出モデルにおいては, たとえば人物を検出する場合には, 検出対象物体である人物が含まれるバウンディングボックスと呼ばれる領域を推定する. 一方, 関節など人物の物体内の位置を予測するためには, その対象が点座標となる. Xiaoらは, CNNを用いて関節位置のヒートマップを推定する方法について取り上げ, この方法をシンプルで効果的なベースラインとして提案している.

筆者の研究においても, 身体や工具などの物体内の位置を予測する必要があるため, Xiaoらの研究を参考に, CNNのネットワークモデルの1つであるResNet-18[14]を用いて, RGBのカラー画像から各検出対象箇所に対して, それぞれ1つのヒートマップを推定する方法を用いた. ResNet-18は深さが18層の畳み込みニューラルネットワークであり, 残差ブロック (図 5-3) を導入することにより, ネットワークのレイヤーが深くなった場合に精度が飽和する劣化問題を解決したモデルとされている.

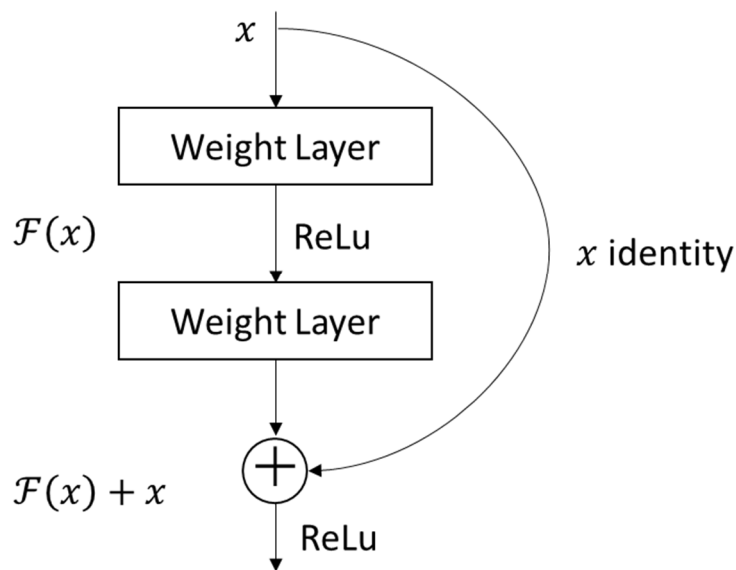


図 5-3 ResNetの残差ブロック

ResNetの手法は、入力画像からそのまま最適な出力を学習するのではなく、ショートカット接続された層の入力情報を参照し、出力と入力との残差が最適になるよう学習することにより最適化し易くしたものである。通常のプレーンなネットワークの場合、学習の目的はターゲット $\mathcal{H}(x)$ をモデリングすることであるが、ResNetの場合は、求める出力写像を $\mathcal{H}(x)$ とした場合、出力 $\mathcal{H}(x)$ と入力 $x$ との残差は残差写像 $\mathcal{F}(x) := \mathcal{H}(x) - x$ で表すことができ、 $\mathcal{H}(x)$ を学習対象の関数だとすると、 $\mathcal{H}(x) = \mathcal{F}(x) + x$ になるように学習を行う。これを残差学習と呼ぶ。今回は、ResNet-18の最後の畳込み層にヒートマップを推定するための3つの転置畳込み層および畳込み層、回帰出力層の追加を行い、転移学習して利用した。

なお、具体的なネットワークについては図 5-4に示すとおり、サイズ $640 \times 352 \times 3$ のRGBのカラー画像1枚に対して、ResNet-18で画像の特徴量を抽出し、その後につづく転置畳込み層でアップサンプリングし、サイズ $160 \times 88$ （縦方向:88, 横方向:160）のヒートマップを検出対象箇所数分出力する回帰出力層を持ったネットワークである。一般的にCNNでは、入力データに対して畳み込み演算とプーリング処理を行い、特徴マップを小さくすることにより性能を向上しているが、特徴マップを小さくすることによりその位置情報の粒度が失われてしまうと言われている。そこで、特徴マップを復元する方法として転置畳込み層を用いてアップサンプリングするというシンプルな手法を用いている。

図 5-4では、4ヵ所を検出対象としてヒートマップが4枚出力され、ヒートマップの1枚目から4枚目までがそれぞれ順番に、4つの検出対象箇所に対応して定義されているモデルである。本ネットワークの出力部分の構造について図 5-5へ示す。

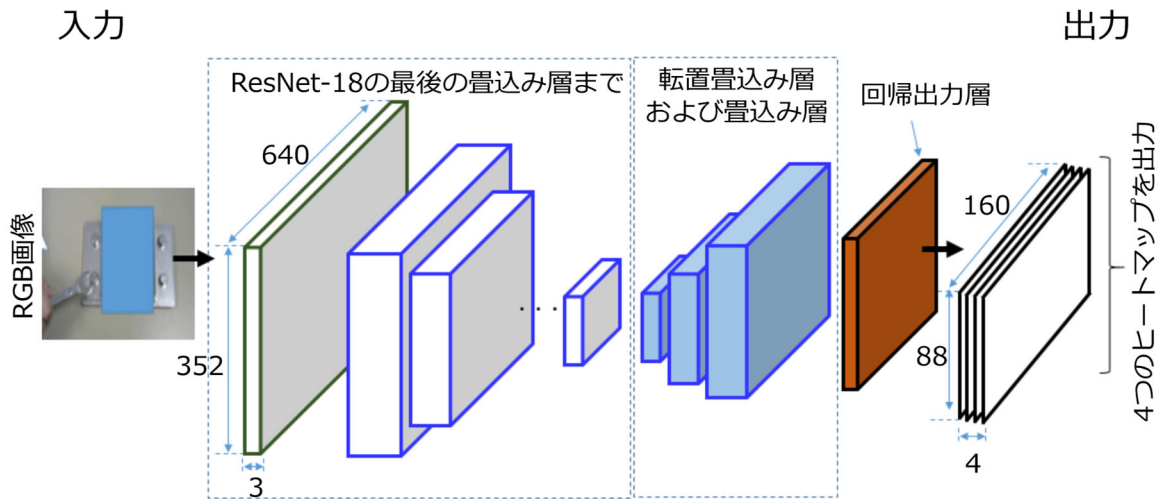


図 5-4 身体や工具位置予測用のネットワーク



図 5-5 転移学習用に調整したResNet-18の出力部分 (MATLAB Deep Learning Network Analyzer)

学習用のヒートマップは、検出対象箇所を中心としてガウシアンフィルタを適用し、最大値を1、最小値を0へ正規化を行ったものを利用する。具体的には、入力用の画像サイズである640×352のサイズに対して、検出対象箇所を中心として標準偏差15の二次元ガウシアンフィルタ処理を行い作成されたヒートマップを、ネットワークの出力サイズと同様の160×88（モデル上は88×160）へリサイズをかける処理を行った。

また、出力されるヒートマップも同様に、予測した対象箇所を最大値に持つヒートマップとなる。一例として、スパナの左先の位置に対する学習用の教師データとネットワークで予測されたデータの



ヒートマップを図 5-6に示す。この図は、 $160 \times 88$ （モデル上は $88 \times 160$ ）のサイズのヒートマップについてメッシュ表面プロットを使ってグラフ化したものである。グラフの $x$ 軸および $y$ 軸はそれぞれ画像座標系における横方向と縦方向を示す。 $z$ 軸については、教師データの場合は検出対象箇所を中心としてガウシアンフィルタを用いて生成した値、予測データの場合はCNNで予測した値となる。図 5-6の予測データについては、正しく予測できた位置に1カ所の大きな山と、その他に誤検出された3カ所の小さな山が見られる。予測の際は、 $88 \times 160$ の二次元配列の中から、もっとも高い山の位置を示す線形インデックスを取得し、それを $88 \times 160$ の二次元配列の添え字へ変換し、各検出対象箇所の座標として扱う。

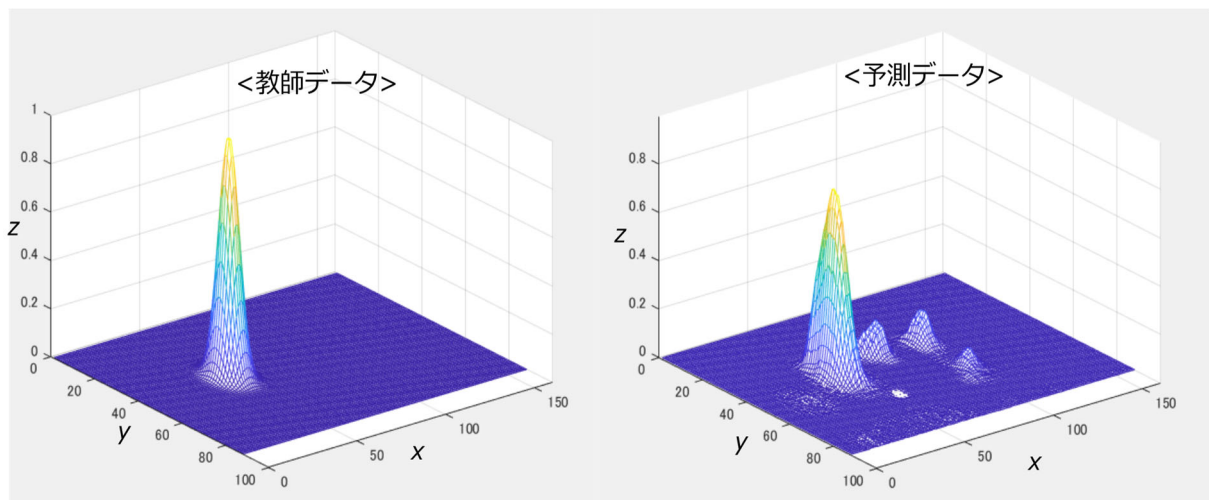


図 5-6 スパナの左先位置に対するヒートマップ

誤差関数は式(1)の平均二乗誤差 (Mean Squared Error) を用いており、回帰出力層の直前の層からの出力値である予測値と、学習に使用した真値である教師データとの間の平均二乗誤差が最小となるように学習されたアルゴリズムである。 $R$ は応答の数、 $t_i$ は応答 $i$ に対する教師データ、 $y_i$ は応答 $i$ に対するネットワークの予測値である。

$$MSE = \frac{1}{R} \sum_{i=1}^R (t_i - y_i)^2 \quad (1)$$

出力された各ヒートマップに対しては、前述のとおり、その中の最大値を示す箇所をヒートマップ内の検出対象箇所の座標として判断する。また、ヒートマップのサイズは、元のRGB画像のサイズに対して1/4のサイズであるため、検出対象箇所の座標を4倍にして元の画像サイズに変換を行い、画像座標系における物体の検出対象箇所の位置として取得する。

### 5.3.3 作業動作の分類処理

作業動作の分類については、時系列データの分類に対してニューラルネットを適用する手法としてWangらの研究[15]および、Fawazらの研究[16]を参考に完全畳み込みニューラルネット（Fully Convolutional Neural Network : FCN）を利用した。具体的には、動画から一定間隔で取得した対象物体の検出箇所に関する座標や移動距離を時系列データとしてCNNに入力し、動作を分類する仕組みとした。本ネットワークはマルチクラス分類を行っており、誤差関数は式(2)の交差エントロピー（Cross Entropy）を用いている。  $K$ は分類対象のクラスの数、  $t_i$ は教師データがクラス  $i$  に属するインジケータ、  $y_i$ はクラス  $i$  に対するネットワークからの出力確率である。

$$E = - \sum_{i=1}^K t_i \log y_i \quad (2)$$

なお、本ネットワークの入力用データについては、縦軸が画像から検出された検出対象箇所ごとの  $x$  座標、  $y$  座標およびフレーム間の時間あたりの移動距離の3変数で構成され、横軸はその時系列変数で構成される。今回は検出対象箇所が4ヵ所であるため、縦軸は検出対象の4ヵ所と3変数を掛け合わせた12変数を使用し、横軸は10フレーム分を時系列に並べて入力情報とした。つまり、10フレームの間のスパナの位置と移動距離を表すデータとなる。この入力サイズは  $10 \times 12 \times 1$  であり、一般的な画像と比べるとサイズが小さいため、CNNについても5層の畳み込み層で構成されたシンプルなネットワーク（図 5-7）とした。

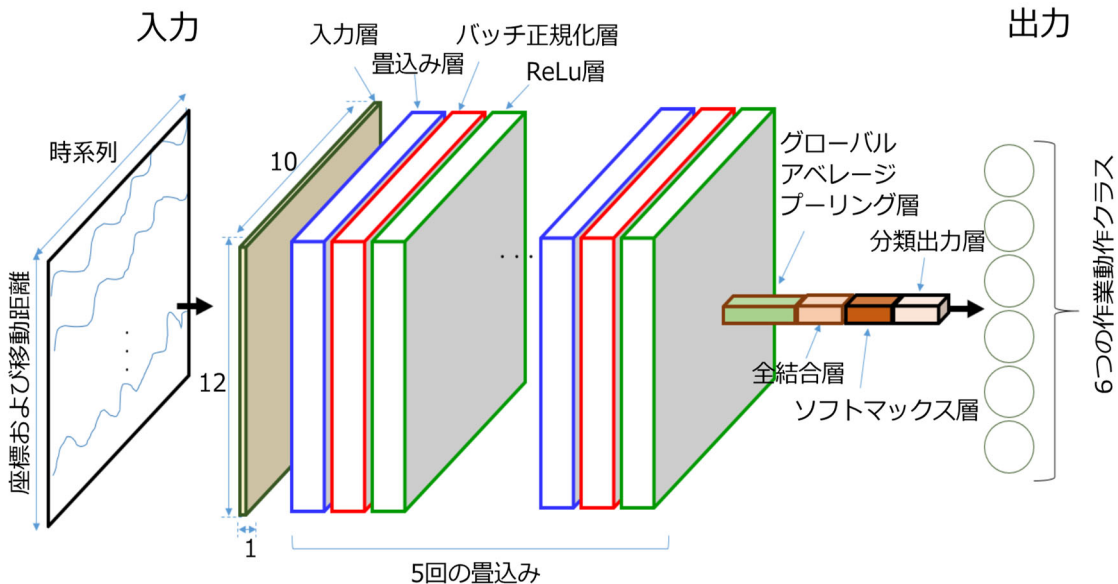


図 5-7 作業動作分類用のネットワーク

今回用いたネットワークの実際の構造については、図 5-8へ示す。



図 5-8 作業動作分類用のネットワークの構造 (MATLAB Deep Learning Network Analyzer)

また、出力情報については、予測されたクラスのラベルと各クラスに対するスコアとなる。また、入力データに関しては、あらかじめ指定したウィンドウサイズ  $n$  を1つの時系列データの単位として、時系列スライディングウィンドウ方式 (図 5-9) を用いて入力情報とした。今回は、ウィンドウサイズ  $n = 10$  として、分類対象のフレームから過去へ10フレーム分遡ったデータを1件の入力データとした。以降、ウィンドウサイズはそのままに、時系列に1フレームずつスライドしながら入力データとして生成した。

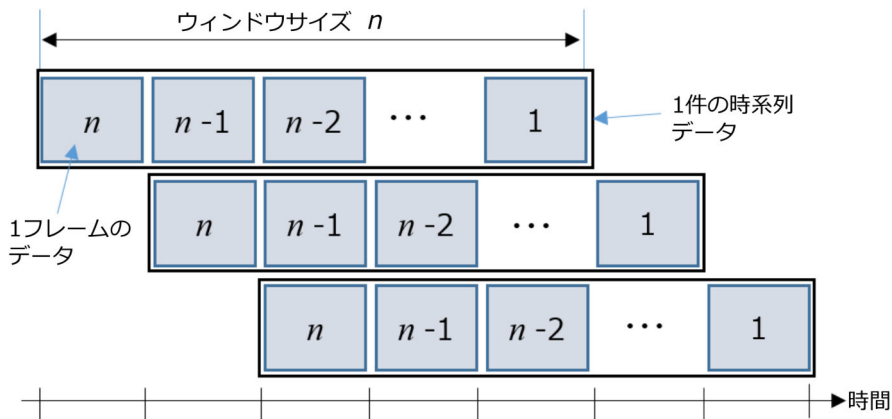


図 5-9 スライディングウィンドウ方式での時系列データ

### 5.3.4 作業動作特徴の分析

CNNから出力された検出対象物体の位置や作業動作の分類結果を作業動作の特徴として、熟練技能者と非熟練者の間で比較評価を行うことにより、非熟練者の動作に対する課題や改善点を把握することが可能となる。

## 5.4 ボルト締め作業での評価実験

### 5.4.1 実験対象

金具組み立て作業における、スパナを用いたボルト締め作業を実験用の環境とし、3点を評価対象として実験を行った。1点目は、作業動作を把握するための対象物体であるスパナの4つの検出対象箇所（図 5-10）の位置に対する予測精度である。2点目は、ボルト締め作業の作業動作に対する分類精度である。この分類については、「動作研究」の手法を参考に、把握したい作業動作にあわせて指定することができるが、今回は熟練技能者と協議した結果、ボルト締め作業の一連の動作に対して、特徴として抽出したい作業動作を図 5-11のとおりに分け、表 5-1のとおり6種類に分類した。なお、図 5-11のインターバルから移動（大・中・小）まではサブブリック記号における「位置決め」に相当し、締付動作は「組合せ」に相当する。なお、今回は熟練技能者と非熟練者との動作特徴の相違を把握するために、より細かい動作に分類した。3点目は、抽出した作業動作の特徴を用いて熟練技能者と非熟練者の間で比較することにより、非熟練者の改善点であるムダな動作を把握することが可能であるかを確認した。



図 5-10 スパナの検出対象箇所

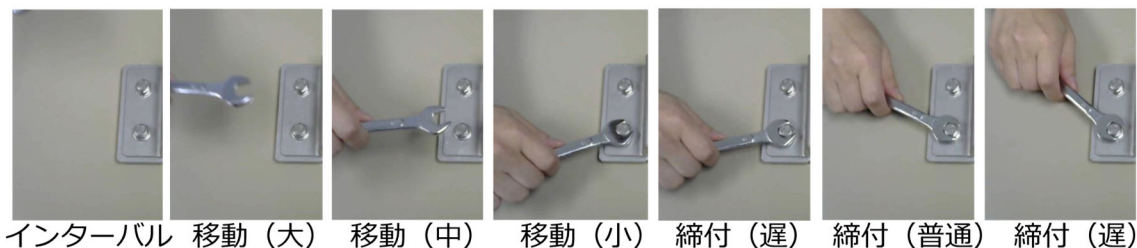


図 5-11 ボルト締め作業動作

表 5-1 ボルト締め作業動作の分類

分類	対象の動作	サブリンク記号
①インターバル	ボルト締め付け動作に入る前や完了後など、スパナが画像内に存在していない状態.	位置決め
②移動 (大)	ボルトの方向へスパナを大きく移動させる動作または、ボルト締め付け後にスパナを大きく移動させる動作.	
③移動 (中)	ボルトの位置へスパナを中程度の速度で移動させる動作または、ボルト締め付け後にスパナを中程度の速度で移動させる動作.	
④移動 (小)	ボルト位置とスパナを正確に合わせる動作または、ボルト締め付け後にスパナをゆっくり移動させはじめる動作.	
⑤締め付 (普)	ボルトをボルト締めに適した速度で締め付ける動作.	組合せ
⑥締め付 (遅)	ボルトをゆっくりした速度で締め付ける動作.	

#### 5.4.2 実験準備

##### (1) スパナ検出対象箇所予測用データの準備

画像からスパナの検出対象箇所を予測するためのCNNの学習データについては、熟練技能者の作業を撮影した画像8,207枚から無作為標本により500枚を選び、それぞれの画像に対して図 5-10に示したスパナの4カ所を対象にヒートマップを作成した。作成したヒートマップと画像を重ね合わせて表示した画像を図 5-12に示す。



図 5-12 検出対象箇所のヒートマップ

なお、このヒートマップは、各検出箇所を中心として、標準偏差15を用いたガウシアンフィルタで生成したものである。また、スパナが画像内に存在しない場合には、すべての値が0（ゼロ）となるヒートマップを準備して学習に用いた。また、テスト用のデータについては、非熟練者Aの作業時の画像667枚、非熟練者Bの作業時の画像866枚から無作為標本によりそれぞれ100枚を選び、同様に準備した。

##### (2) 作業動作分類用データの準備

作業動作分類用の学習データについては、熟練技能者の6回分の作業時の動画から準備した。作業時

の動画フレームに対して、スパナの4ヵ所の位置をスパナ検出対象箇所予測用のCNNを用いて0.1秒周期で取得し、ウィンドウサイズ $n = 10$ の時系列データを1セットとして、無作為標本で4,000件準備した。つまり、1ウィンドウあたり1秒間の動作に相当する時系列データとなる。この学習データに対して、表 5-1にしたがってラベルを割りあてることにより学習用データとした。テストデータについては非熟練者2名の動作を対象に、非熟練者Aは170件、非熟練者Bは234件を同様に時系列データとして準備した。なお、時系列データの一例について、その内訳を図 5-13に示す。

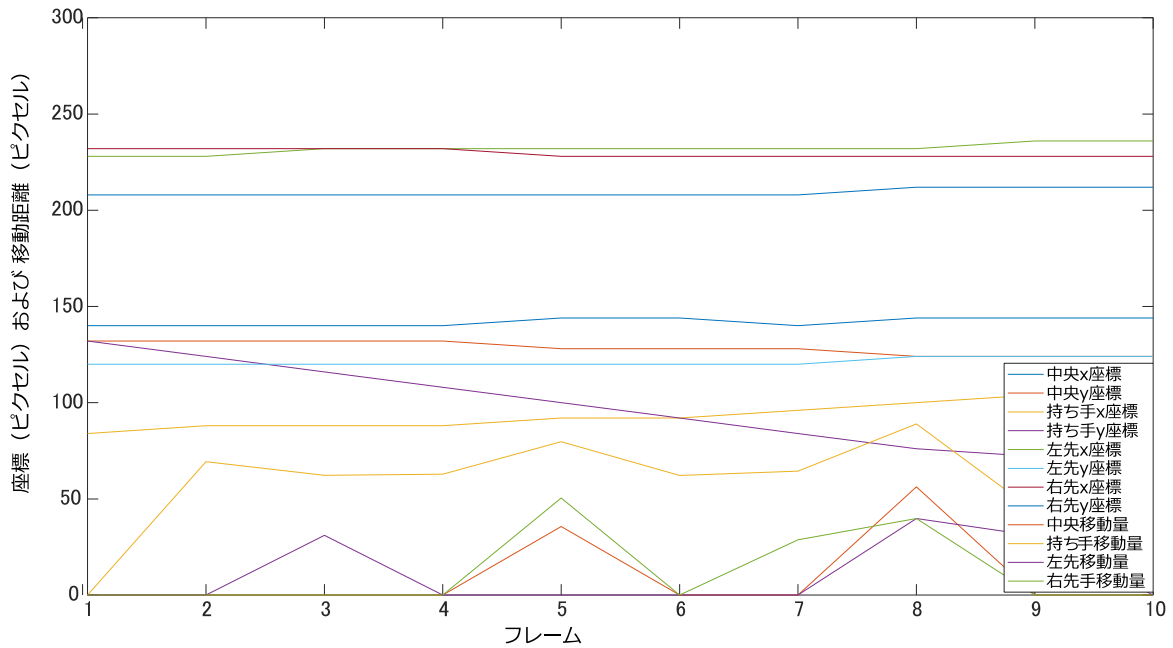


図 5-13 ボルト締付（普）動作時の学習データ

このグラフはボルト締め付け動作中の“締付（普）”の際の時系列データである。x軸は画像のフレーム数で、y軸は画像内の座標および、フレーム間の移動距離をピクセル数で示したものである。内訳としては、検出対象の4ヵ所に対する画像内のx座標およびy座標の8変数のデータと、同じく検出対象箇所のフレーム間の移動距離に関する4変数のデータが10フレーム分で構成されたものである。

### (3) スパナ検出対象箇所予測用CNNの学習

スパナの検出対象箇所に関する予測用のCNNはオプティマイザにはADAMを使用し、ミニバッチサイズ 32で100エポック学習した。図 5-14はその学習曲線となる。横軸が学習の反復回数で、縦軸が予測された位置と正しい位置との平方根平均二乗誤差であり、学習が収束していることが確認できる。

### (4) 作業動作分類用CNNの学習

スパナを使用した際の作業動作分類用のCNNについては、オプティマイザにはADAMを使用し、ミニバッチサイズ 128で500エポック学習した。図 5-15に学習曲線を示す。横軸が学習の反復回数で、

縦軸が分類精度であり，学習が収束していることが確認できる．

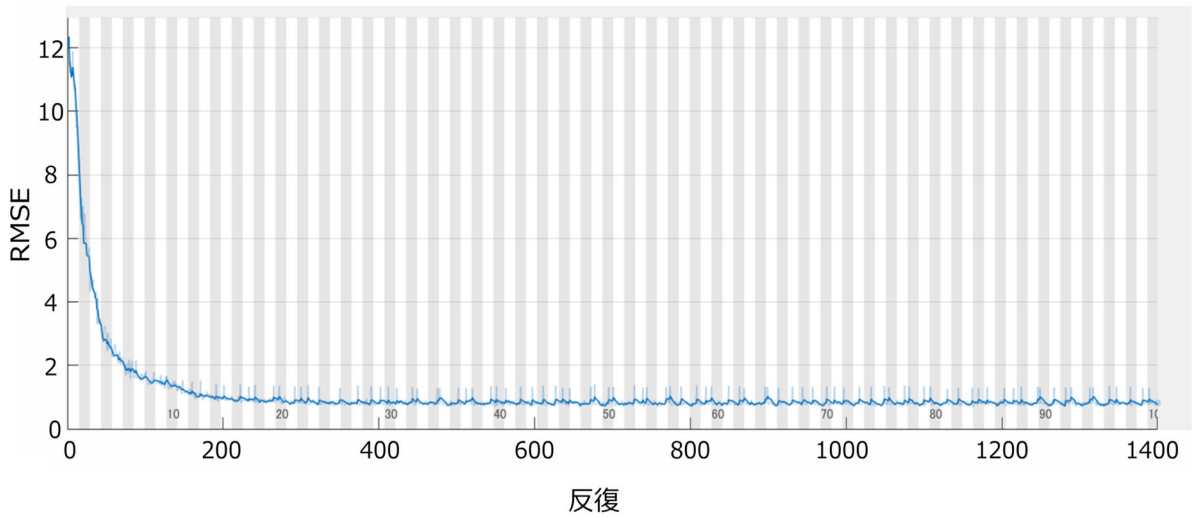


図 5-14 スパナ検出対象箇所予測用CNNの学習結果

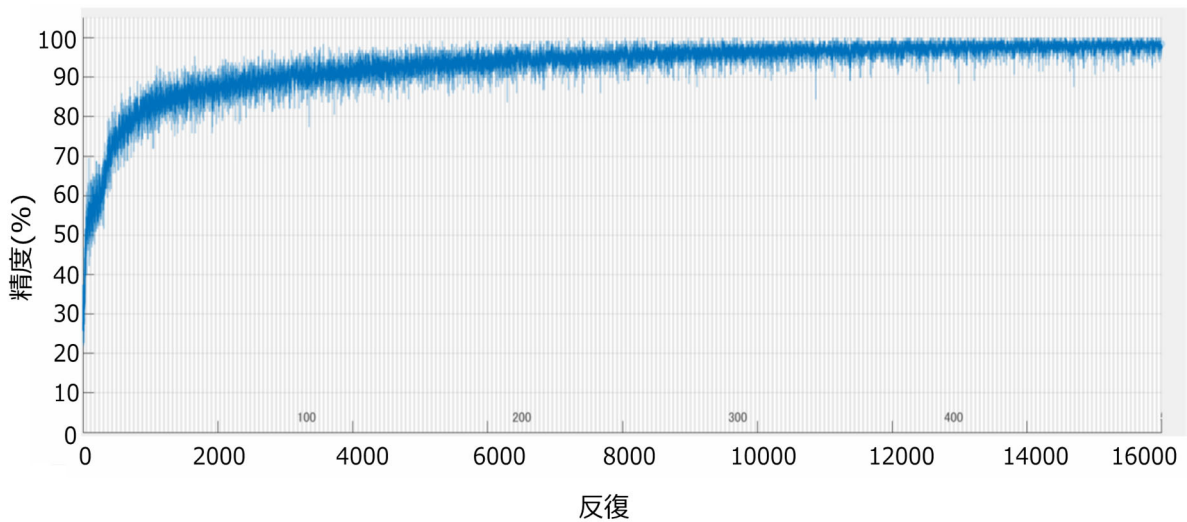


図 5-15 作業動作分類用CNNの学習結果

### 5.4.3 実験内容

#### (1) スパナ検出対象箇所予測精度

スパナの検出対象箇所の子測精度については，学習データに含まれない非熟練の被験者2名の画像に対して，検出対象箇所の子測結果に関する正解率について評価した．具体的には，各画像内のスパナの検出対象4カ所に対して人が目視で指定した座標と，CNNで予測された座標との距離が一定の閾値以下の場合を正解，閾値を上回った場合を不正解として，その正解率を使用して評価した．なお，この評価方法については人物の姿勢推定モデルの評価で利用されている正しいキーポイントの割合

(Percentage of Correct Key-points: PCK) [17]を参考としたが、今回の検出対象物体に適した閾値については、文献などで参考となる数値を確認することができなかった。そのため、本実験用の閾値としてスパナの柄の部分の直径（18ピクセル）を用いた（図 5-16）。これは今回の画像のサイズが横640ピクセル、縦352ピクセルであるため、画像に対して横方向では約3%、縦方向では約5%の誤差に相当する閾値である。また、今後閾値について検討するための参考値として、スパナの直径の半分の長さ（9ピクセル）を閾値とした評価も行った。

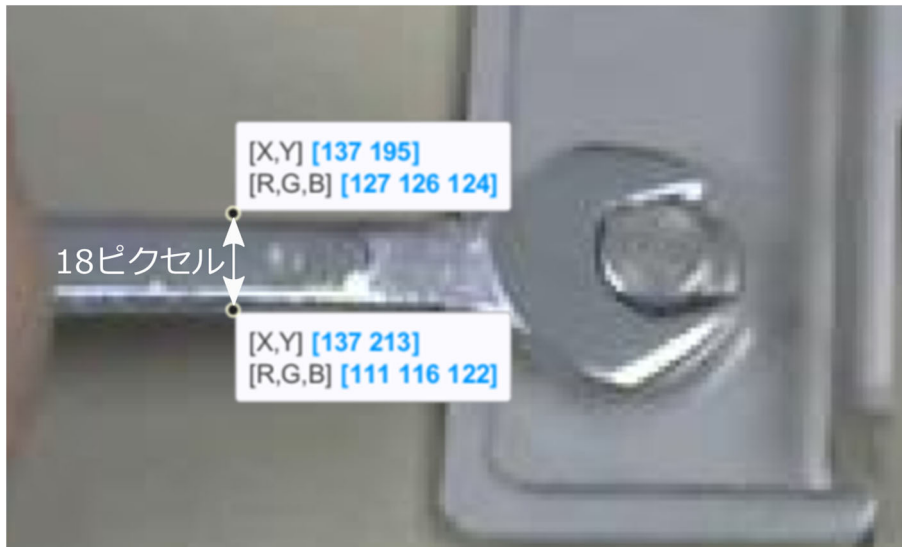


図 5-16 スパナの検出精度の指標

## (2)作業動作分類精度

スパナを使用した際の動作に関する分類精度については、学習データに含まれない被験者2名のそれぞれ6種類の作業動作に対して、あらかじめ分類した作業動作に対するラベルとCNNで分類した作業動作の比較を行い、その分類精度を求めた。

## (3)作業動作特徴の比較評価

ボルト締め作業において、熟練技能者と非熟練者の中で違いが生じやすいとされるスパナを回転させる際の動作を主な評価対象とし、各作業者の動作の相違点について確認することが可能であるかを評価した。なお、この評価についても、学習に使用していないデータを用いて実施した。

### 5.4.4 実験結果

#### (1)スパナ検出対象箇所予測精度

スパナの検出対象箇所に対する予測精度については表 5-2に示す。評価時の閾値にスパナの直径である18ピクセルを用いた場合は非熟練者2名に対する予測精度の平均は98.38%であった。また、参考として用いた閾値にスパナの直径の半分である9ピクセルを用いた場合は、78.50%であった。予測され



た位置と正しい位置との誤差については、2名平均で5.96ピクセルとなった。

表 5-2 スパナの検出対象箇所予測精度

検出箇所	スパナの直径サイズ (18pix) 以内		スパナの直径サイズの50% (9pix) 以内	
	非熟練者A	非熟練者B	非熟練者A	非熟練者B
①左先	94%	99%	57%	72%
②右先	100%	100%	76%	82%
③中央	100%	100%	87%	93%
④持ち手	97%	97%	83%	78%
平均	97.75%	99.00%	75.75%	81.25%

今回実験で用いた全800カ所 (2名\*100枚\*4カ所) の各検出対象箇所に対する予測誤差の分布については図 5-17のヒストグラムで示す。

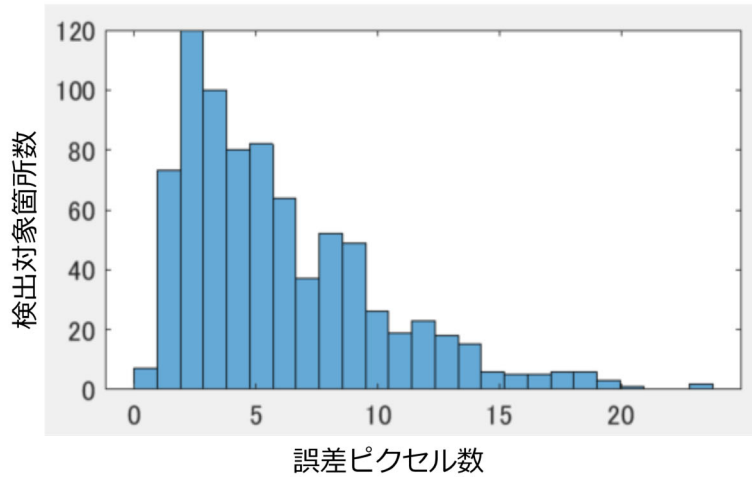


図 5-17 スパナ検出対象箇所予測誤差ヒストグラム

(2)作業動作分類精度

スパナを使用した際の6種類の作業動作の分類精度に関する実験結果を表 5-3に示す。この表は分類結果を混同行列で表したもので、行 (縦) は実際の正解ラベルのクラス、列 (横) はCNNで分類されたクラスであり、各クラスの分類精度を対角要素で示したものである。分類精度については、2名平均で93.50%となった。この混同行列のなかで、非熟練者Aの“③移動 (中)”をはじめ一部誤分類している箇所が見られるが、これは連続した動作の間での動作の切り替わり部分で発生していることが図 5-18で確認できる。この図は正解ラベルとCNNで分類されたラベルを比較したグラフであるが、赤の破線部分が誤分類の主な対象箇所となる。

表 5-3 ボルト締め作業動作分類精度

非熟練者A	①インターバル	②移動(大)	③移動(中)	④移動(小)	⑤締付(普)	⑥締付(遅)
①インターバル	100.00%	0.00%	0.00%	0.00%	0.00%	0.00%
②移動(大)	0.00%	96.55%	3.45%	0.00%	0.00%	0.00%
③移動(中)	0.00%	9.09%	86.36%	4.55%	0.00%	0.00%
④移動(小)	0.00%	0.00%	8.33%	91.67%	0.00%	0.00%
⑤締付(普)	0.00%	0.00%	0.00%	0.00%	92.11%	7.89%
⑥締付(遅)	0.00%	0.00%	0.00%	3.13%	3.13%	93.75%
平均						93.41%

非熟練者B	①インターバル	②移動(大)	③移動(中)	④移動(小)	⑤締付(普)	⑥締付(遅)
①インターバル	100.00%	0.00%	0.00%	0.00%	0.00%	0.00%
②移動(大)	0.00%	93.88%	4.08%	2.04%	0.00%	0.00%
③移動(中)	0.00%	5.26%	89.47%	0.00%	5.26%	0.00%
④移動(小)	0.00%	0.00%	7.32%	92.68%	0.00%	0.00%
⑤締付(普)	0.00%	0.00%	0.00%	0.00%	96.15%	3.85%
⑥締付(遅)	0.00%	0.00%	0.00%	4.76%	5.95%	89.29%
平均						93.58%

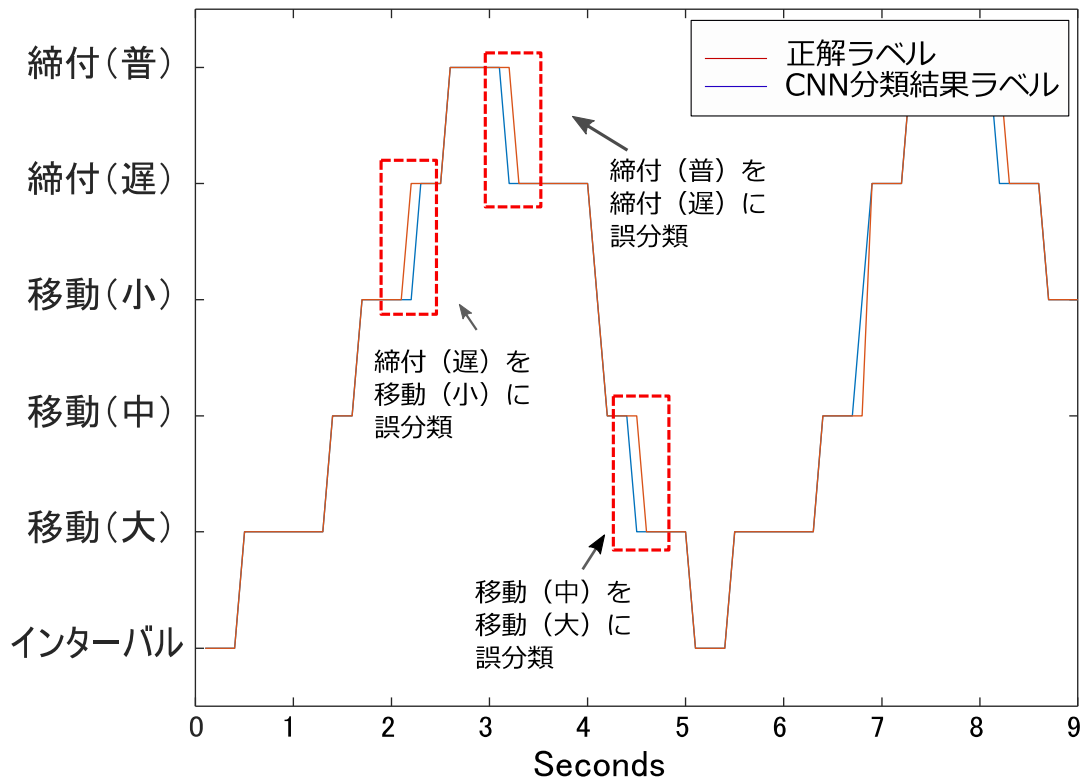


図 5-18 誤分類ケース

(3)作業動作特徴の比較評価

抽出された熟練技能者1名と非熟練者2名の作業動作の特徴を用いて、以下の3点について比較評価を行った。

1点目は各動作が占める時間の割合について、作業2回分を円グラフにして比較した。内容を図5-19に示すが、円グラフの上段は1回目、下段は2回目の作業となる。その結果、熟練技能者は2回分の作業に対する各動作の割合がほぼ同一であることが確認できる。一方、非熟練者2名については2回の作業に対する動作の割合が異なっており、作業ごとに動作時間のバラツキがあることが考えられる。また、非熟練者は締め付けの際に遅い動作の割合が多く、手首の回転だけでボルトを締め付けている可能性を確認することができる。

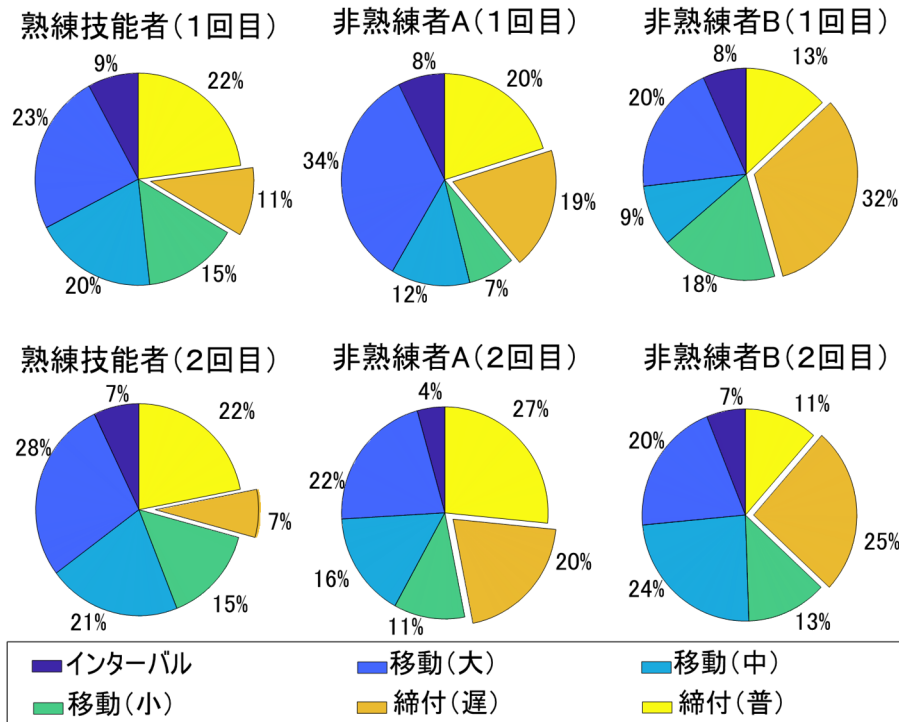


図 5-19 動作内訳の比較

2点目は、ボルト締め動作における各動作の時系列な推移について、同様に熟練技能者と非熟練者2名を対象に15秒間の動作を2回計測した結果をグラフで示す(図5-20)。

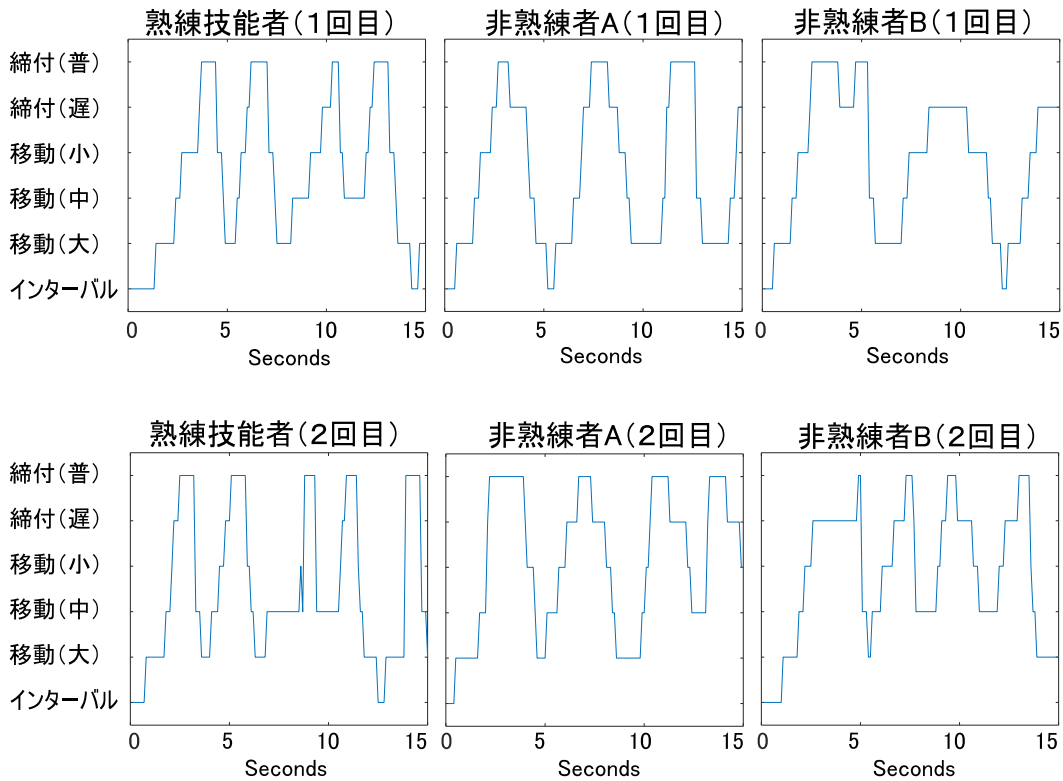


図 5-20 時系列動作推移

ボルト締め動作については、図 5-11に示したとおり、スパナをボルト位置に向けて大きく移動させる動作からはじまり、その移動が中程度、小程度となったあと、ボルト位置でゆっくりとボルトを締め始める。その後、ボルトを締め付けに適した速さで締めたあとに、再度遅い締め付け動作になり、その後スパナを次のボルト位置へ移動させる流れとなる。図 5-20の上段が1回目、下段が2回目となるが、熟練者の動作は1回目、2回目ともに比較的整った周期性が見られる。一方、非熟練者AおよびBについては1回目と2回目で周期に違いが見られる。

3点目として、ボルト締め動作の際のスパナの中央部分(図 5-10の③)と持ち手部分(図 5-10の④)を時系列にプロットしたグラフを図 5-21に示す。青のマーカ―は締付(普)で赤のマーカ―は締付(遅)の動作を示す。

六角形のボルトを締め付ける際はスパナを60度回転させることが効率の良い動作とされている(図 5-22)。熟練技能者については4カ所のボルト位置ともに60度に近い範囲でスパナを回転させているが、非熟練者2名についてはいずれも回転範囲が狭い箇所が見られる。特に非熟練者Bについては、回転範囲の狭さと動作の遅さが著しいことが分かる。

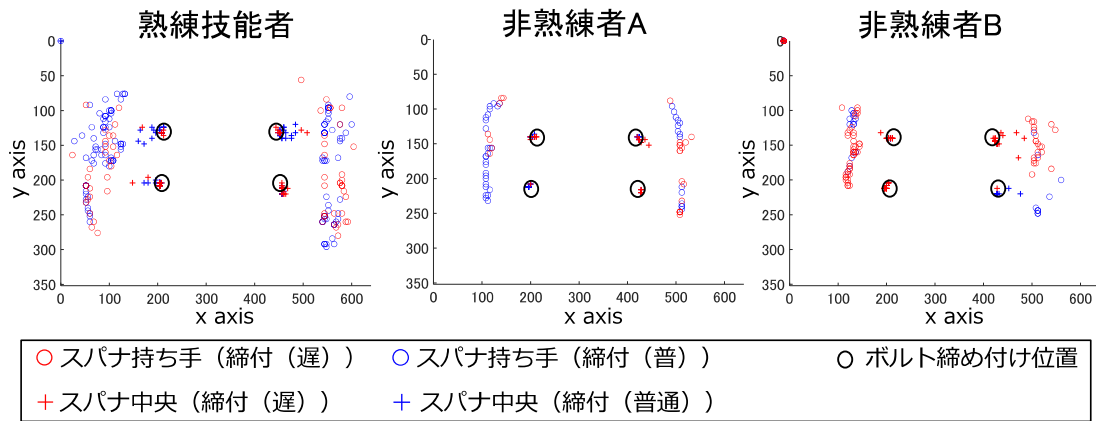


図 5-21 スパナの時系列位置情報

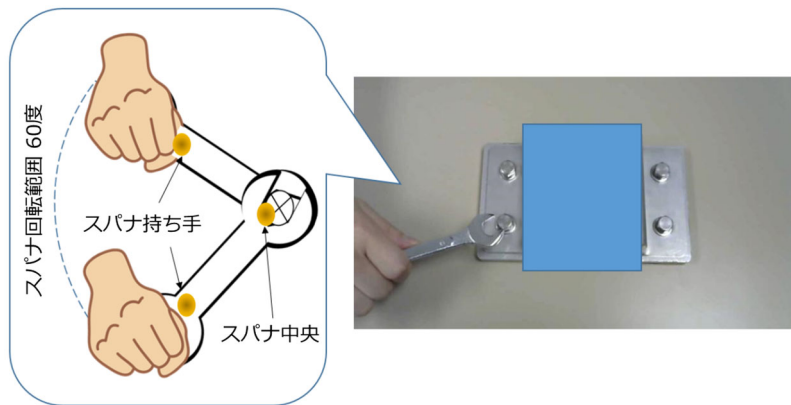


図 5-22 スパナ回転範囲

## 5.5 実用性に関する評価実験

本システムに対する汎用性と課題について確認するために、ボルト締め作業と比べてより動作が複雑である溶接作業を1つの対象として、検出対象箇所の位置に対する予測精度と作業動作に対する分類精度について評価を行った。溶接作業では作業過程での品質判定や製品での破壊検査が困難であるという理由から、技能品質向上を目的として作業者の技能動作の定量評価を行い、技能情報を構造化するための研究[18]が行われている。今回、溶接作業を評価実験の対象として取りあげた理由は、身体動作をともなう技能が重要視されている点から、今後、筆者のシステムの適用先としてより詳細な調査および評価を行いたいためである。したがって、今後の研究へ繋げるために、今回は予測精度と分類精度および、課題を確認することを目的として実験を実施した。

### 5.5.1 実験対象

本実験における検出対象に関しては、画像から作業に関係する可能性がある情報をできるだけ多く取得するために、溶接対象の位置調整に使用するハンマー（4カ所）および溶接治具（4カ所）、作業

者の両腕の各関節（6カ所）を対象とし、合わせて14カ所とした（図 5-23）。

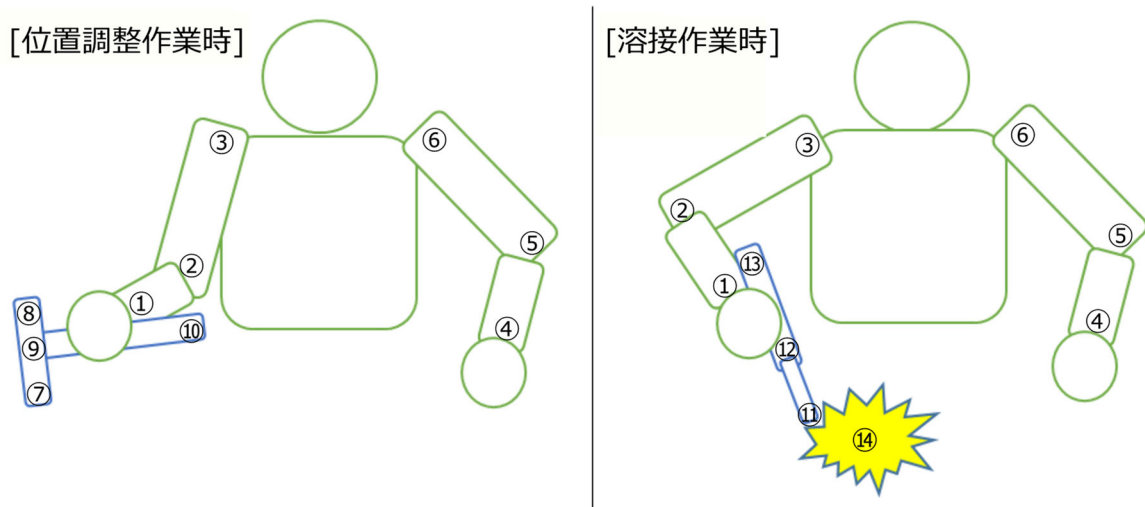


図 5-23 溶接作業時の検出対象箇所

作業動作については、今回の溶接作業を構成する主要な作業動作であるハンマーを用いた溶接対象物体の位置合わせのための“調整”および、溶接治具を用いた“溶接”を抽出の対象とし、他の作業については“その他”として、3種類の作業動作として分類した。また、本実験では、熟練技能者3名（被験者A, B, C）を被験者とした。

### 5.5.2 実験準備

実験準備についてはボルト締め作業と同様に、検出対象箇所予測用のヒートマップと作業動作分類用の時系列データを準備した。具体的な内訳としては、検出対象箇所予測用の学習データについては被験者Aを撮影した画像3,171枚から1,000枚を無作為標本で選び、テスト用データについては被験者Bおよび被験者Cを撮影した画像からそれぞれ100枚を無作為標本で選んだ。なお、14カ所の検出対象箇所については画像に写っている箇所に対してのみラベル付けを行い、画像に写っていない箇所や他の物体で隠れた箇所についてはラベル付け対象から除外した。作業動作分類用の学習データについては、被験者Aの時系列データ3,000件を用い、テストデータについては、被験者Bについては2,127件、被験者Cについては、2,118件を準備した。各、CNNについてはボルト締め作業時と同様にネットワークを準備し、回帰用のネットワークについては予測するヒートマップの出力数を14へ設定し、また、分類用のネットワークについては分類対象となる作業動作のクラス数を3へ設定して、ボルト締め作業時と同様の学習パラメータで学習を行った。

### 5.5.3 実験内容

実験内容についてもボルト締め作業同様に、学習データに含まれない被験者2名の画像に対する検出対象箇所への予測結果に関する正解率と、同じく被験者2名のそれぞれ3種類の作業動作に関する分類精度について確認するための実験を行った。なお、本実験については検出対象箇所への予測精度と作業動作の分類精度の確認を目的とし、作業動作の特徴に関する比較評価については、筆者が溶接作業における作業動作の特徴に関する調査が未実施であるため、本実験から除外した。

### 5.5.4 実験結果

#### (1) 検出対象箇所予測精度

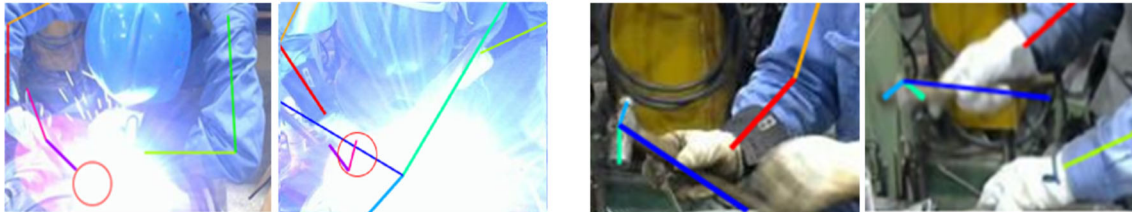
溶接作業における検出対象箇所に対する予測精度について、表 5-4に示す。ボルト締め作業時同様に、正解と判断する際の閾値に18ピクセルを用いた場合の被験者2名に対する予測精度は平均83.92%、閾値に9ピクセルを用いた場合の予測精度は58.44%となった。

表 5-4 溶接時の検出箇所予測精度

検出箇所	18pix以内		9pix以内	
	被験者B	被験者C	被験者B	被験者C
①右手首	98%	96%	92%	91%
②右肘	98%	93%	81%	79%
③右肩	92%	98%	58%	85%
④左手首	85%	95%	77%	78%
⑤左肘	98%	87%	74%	64%
⑥左肩	89%	81%	42%	59%
⑦ハンマー右	83%	91%	43%	79%
⑧ハンマー左	65%	93%	22%	81%
⑨ハンマー中央	96%	95%	52%	88%
⑩ハンマー柄	70%	70%	57%	53%
⑪溶接先端	75%	67%	30%	33%
⑫溶接持手	95%	71%	63%	14%
⑬溶接柄	80%	43%	63%	14%
⑭溶接箇所	87%	60%	23%	40%
平均	<b>86.42%</b>	<b>81.42%</b>	<b>55.57%</b>	<b>61.30%</b>

ボルト締め作業と比較すると予測精度の平均値が低い結果となったが、特に予測精度が低い箇所については、被験者Cの⑪~⑬の溶接治具および、“⑭溶接箇所”があげられる。実際の画像を確認したところ、被験者Cの画像では学習データで使用した被験者Aの画像と比べ溶接の際の溶接アーク光の度合

いが強い画像が含まれており、溶接箇所や治具部分が白飛びしていることが原因であることが確認できた（図 5-24 a-2）。また、ハンマーについても被験者Bの“⑧ハンマー左”や“⑩ハンマー柄”のように予測精度が低い箇所が見られる。これについて確認したところ、ハンマーで叩く際の高速での動作に対して、被写体であるハンマーがブレて撮影されている箇所が対象であった（図 5-24 b-2）。なお、これらの画像はいずれも学習データやテストデータにも存在しており、目視でおおよその位置を対象にラベル付けを行っておいたが、今回の実験では正しく予測することができないケースが見られた。



[a-1 検出できたケース][a-2 白飛びしたケース] [b-1 検出できたケース][b-2 被写体ブレケース]

図 5-24 白飛びするケースと被写体ブレのケース

## (2)作業動作分類精度

溶接作業における3種類の作業動作の分類結果について、表 5-5に示す。作業動作の分類精度についても、ボルト締め作業と比較すると低い結果となった。特に、被験者Cについては、溶接アーク光の影響により検出対象箇所の予測精度が低かったため、“③溶接”に関する分類精度も低い結果となった。

表 5-5 溶接時の作業動作分類精度

被験者B	①その他	②調整	③溶接	被験者C	①その他	②調整	③溶接
①その他	88.89%	10.37%	0.74%	①その他	82.60%	17.15%	0.25%
②調整	20.83%	79.17%	0.00%	②調整	9.29%	90.71%	0.00%
③溶接	11.59%	5.80%	82.61%	③溶接	74.03%	0.00%	25.97%
平均	83.56%			平均	66.43%		

## 5.6 考察

今回のボルト締め作業に関する評価実験では、スパナの検出対象箇所と動作の分類結果を作業動作の特徴として比較評価を行うことにより、熟練技能者と非熟練者の動作の相違を把握することが可能であることを確認した。また、実用性を確認するための溶接作業に関する評価実験では、実用に向けた課題についてのいくつかを把握することができた。5.2.1節で示した3つの要求条件に関して、5.4節および5.5節で示した評価実験に基づき考察する。



### (1)身体や工具の検出対象箇所を正確に予測できること

ボルト締め作業を対象に、回帰用のCNNを用いてスパナの検出対象箇所を予測する実験を行った結果、予測精度については、スパナの柄の部分の直径（18ピクセル）を予測誤差の閾値とした場合で平均98.38%、その半分のサイズ（9ピクセル）を閾値とした場合で平均78.50%であった。誤差の平均は5.96ピクセルで、この数値は画像全体に対して横方向で0.93%、縦方向で1.69%の誤差であり、本ケースにおいて、検出対象物体の位置と形状を予測するためには十分な精度であると考えられる。

### (2)作業動作を正確に分類できること

ボルト締め作業を対象に、回帰用のCNNで予測したスパナの検出対象の位置情報から座標と移動距離を取得し、時系列データとして分類用のCNNを用いて動作の分類を行った結果、分類精度については2名平均で93.50%の精度であった。一部誤分類されている箇所については、動作の切り替わりの部分が曖昧になり誤って分類されているものと考えられる。これは連続した時系列データを取り扱ううえでの課題であるため、対策については引き続き検討を行いたいと考えている。

また、想定外の動作を行った場合について今回評価は行わなかったが、これについては分類時の分類確率スコアについて一定の割合以下の場合を、閾値を用いて判断することにより除外することが可能である。

### (3)異なる作業においても汎用的に利用できること

本システムに対する汎用性と課題に関する評価を目的として、溶接作業を対象に実験を行った。その結果、予測精度については閾値に18ピクセルを用いた場合は平均83.92%、閾値に9ピクセルを用いた場合は平均58.44%となった。精度が低い原因については、溶接アーク光の影響による白飛びと、高速で移動する物体に対する被写体ブレであることが確認された。作業動作の分類精度については、2名平均で75%となり、同様に溶接アーク光や被写体ブレの影響による分類精度の低下が確認された。この2つの課題はいずれも撮影条件としての課題であり対策が必要とされるが、これらの影響を受けていない画像に対しては検出対象箇所の予測精度、作業動作の分類精度ともに比較的高い精度であったため、画像が正しく得られる撮影条件下においては、他の作業においても利用することが可能であると考えられる。

## 5.7 まとめと今後の課題

本研究では、カメラで撮影した動画と2種類の深層学習の技術を用いることにより、画像から「動作研究」に役立つような作業時の特徴となる作業動作を抽出することができる汎用的なシステムを開発した。CNNを利用しEnd-to-Endで学習を行うことにより、一連の処理を実現している点を特徴としている。

本稿では、一般的な作業の1つであるボルト締め作業を対象に実験を行い、作業に使用するスパナ内

の4つの検出対象箇所に基づき、一連の動作を6種類の作業動作に分類し、特徴として抽出した。その結果を用いて熟練技能者と非熟練者を比較することにより、非熟練者の動作の課題を把握することができた。本手法は、製造現場の画像を用いてCNNの学習を行うことにより、多くの作業に対して応用が可能であるが、使用する工具や作業者の身体が画像から取得できることが前提となる。そのため今回評価を行った溶接作業のように、画像の白飛びや被写体ブレが生じる場合、予測や分類精度が低下することになる。今回はこれらに対する対策は行わなかったが、画像の白飛びについてはコントラストの調整や、高速移動物体に対する被写体ブレについては撮影時のシャッタースピードを高くするなどの方法が対策としてあげられるため、調査を行いたい。

今回の評価実験の結果をふまえ、今後以下の3点について引き続き検討を行いたい。1点目は、複雑な作業における技能評価である。今回取り上げた溶接作業を対象として、身体動作に関する技能評価への適用へ向けた研究を行いたい。2点目は、本手法の再利用性に関する研究である。類似した実際の製造現場において今回学習を行ったCNNに対する汎化性能の評価を行いたい。3点目は、本システムの応用に関する研究である。本システムに関しては、技能評価以外の用途として、同一人物における通常時の動作と長時間作業を行った際の動作を比較することにより、疲労度合の検出や作業ミス検出に対しても利用することが可能であると考えられるため、引き続き検討を行いたい。

## 参考文献

- [1] 杉原寛, 嶋本康夫, 中村純一, 藤山紘一郎, 杉浦亘, 宮原正典 : IE 7 つ道具, p.87, 日刊工業新聞社 (1993) .
- [2] 杉原寛, 嶋本康夫, 中村純一, 藤山紘一郎, 杉浦亘, 宮原正典 : IE 7 つ道具, pp. 90-91, 日刊工業新聞社 (1993) .
- [3] David G. Lowe : Object Recognition from Local Scale-Invariant Features, Proceedings of the International Conference on Computer Vision, Volume 2, pp. 1150-1157(1999).
- [4] Donald J. Berndt, James Clifford : Using Dynamic Time Warping to Find Patterns in Time Series, Association for the Advancement of Artificial Intelligence Technical Report, pp. 359-370 (1994).
- [5] Jason Lines, Anthony Bagnal : Time series classification with ensembles of elastic distance measures, Data Mining and Knowledge Discovery 29, pp. 565-592 (2015).
- [6] Anthony Bagnall, Jason Lines, Aaron Bostrom, James Large, Eamonn Keogh : The great time series classification bake off: a review and experimental evaluation of recent algorithmic advances, Data Mining and Knowledge Discovery 31, pp. 606-660 (2017).
- [7] Zhiguang Wang, Weizhong Yan, Tim Oates : Time Series Classification from Scratch with Deep Neural Networks: A Strong Baseline, International Joint Conference on Neural Networks, pp.1578-1585 (2016).
- [8] Hassan Ismail Fawaz, Germain Forestier, Jonathan Weber, Lhassane Idoumghar, Pierre-Alain Muller : Deep learning for time series classification: a review, arXiv:1809.04356v4(2019).
- [9] 宮原正典, 杉原寛, 嶋本康夫, 中村純一, 藤山紘一郎, 杉浦亘 : IE 7 つ道具, 日刊工業新聞, p.85 (1993).
- [10] 米谷竜, 齋藤英雄, 池畑諭, 牛久祥孝 他 : 未来へつなぐデジタルシリーズ37 コンピュータビジョン -広がる要素技術と応用-, p.25, 共立出版 (2018) .
- [11] Mathworks: MATLAB-計算技術言語, Mathworks, 入手先  
〈<https://jp.mathworks.com/products/matlab.html>〉 (参照 2020-05-18).
- [12] Bin Xiao, Haiping Wu, Yichen Wei : Simple Baselines for Human Pose Estimation and Tracking, Proc. European Conference on Computer Vision (2018).
- [13] Ross Girshick, Jeff Donahue, Trevor Darrell, Jitendra Malik : Rich feature hierarchies for accurate object detection and semantic segmentation, IEEE Conference on Computer Vision and Pattern Recognition, ,pp. 580-587 (2014).
- [14] Kaiming He, Xiangyu Zhang, Shaoqing Ren, Jian Sun : Deep residual learning for image recognition, Proc. IEEE conference on computer vision and pattern recognition, pp. 770-778 (2016).
- [15] Zhiguang Wang, Weizhong Yan, Tim Oates : Time Series Classification from Scratch with Deep

Neural Networks: A Strong Baseline, International Joint Conference on Neural Networks, pp.1578-1585 (2016).

- [16] Hassan Ismail Fawaz, Germain Forestier, Jonathan Weber, Lhassane Idoumghar, Pierre-Alain Mulle : Accurate and interpretable evaluation of surgical skills from kinematic data using fully convolutional neural networks, arXiv:1908.07319v1 (2019).
- [17] GitHub : Human Pose Estimation 101, 入手先 <<https://github.com/cbsudux/Human-Pose-Estimation-101#percentage-of-correct-key-points---pck>> (参照 2020-05-18).
- [18] 松浦慶総, 高田一 : 溶接技能教育における情報構造化手法の提案, 第29回人工知能学会全国大会 (2015).

## 第6章 まとめ

日本国内の製造業の状況について、短期的に見ると新型コロナウイルス感染症（COVID-19）の世界的な感染拡大の影響を受け、売上高および営業利益は減少傾向にある。しかし、中期的に見ると、中小企業をはじめとする作業員（人）が中心となって作業を行っている製造現場の多くでは、人材不足や人材育成に関する課題が指摘されており、このような企業が発展を続けていくためには生産性の向上が必要である。1章では、COVID-19による影響を受ける前の時点と現時点をとおして、製造業における人材不足という課題に対する状況について述べた。

労働生産性向上については、以前からインダストリアル・エンジニアリング（IE）による改善活動が行われてきたが、この活動を行うための人材であるIErに対しても同様に、人材不足や人材育成に関する課題が存在する。そこで、IEにおける改善活動に対してデジタル技術を活用して効率化をはかる方法が考えられるが、中小企業を中心とした製造業においては、IEのスキルとITのスキルの両方を持ち合わせた技術者の確保は難しい。2章では、IEに関する主な手法について取り上げるとともに、IEに関する課題として、モノづくりの現場を深く理解したIEの知見を有する人材が不足するという状況と合わせて、デジタル技術を活用するために必要なIT技術を有する人材の育成や確保が困難であるという2つの課題が存在していることを述べた。

3章では、関連研究として製造現場での作業ミスの検知、動作分析、時間分析などにデジタル技術を利用した研究を取り上げた。しかし、これらの関連研究は対象の作業に合わせた専用の設計が必要となる手法が多く、製造現場の環境に対する制約や前提条件を必要とする研究も見られる。つまり、幅広い作業に対して汎用的に利用するという事は難しく、導入する際には、IErに対してITのスキルが求められる可能性が高いものと考えられる。そこで本研究では、IErの人材が限られている場合やIT技術者が不在である場合でも、デジタル技術の活用による効果を得ることができる汎用的な仕組みが必要であると考え、製造現場の画像を用いたEnd-to-Endでの学習により利用することが可能な深層学習技術について着目した。

4章では、IEにおける「時間研究」により定められた標準作業や標準作業時間に基づき、深層学習技術を利用して作業員の作業手順および作業時間をリアルタイムで監視する手法について、評価実験の結果を含めてその有効性について示した。また、5章では、IEにおける「動作研究」を支援する仕組みとして作業員の作業動作の特徴に注目し、動作の特徴を抽出する手法と、その特徴について熟練技能者と非熟練者との間で比較することにより、効率的な動作とムダな動作を相対的に評価することを可能とする手法について、同じく評価実験の結果を含めて有効性について示した。いずれも、製造現場の画像を利用したEnd-to-Endの学習を行うことにより利用することが可能であり、従来からの製造現場に向けて専用に設計された仕組みと比較すると、より多くの製造現場において汎用的に利用することが可能であると考えられる。この手法を用いることにより、中小企業においてもIEによる改善活動や人材育成を効率的に進めることが可能となり、労働生産性の向上に寄与することができるものと考

える。

今後の展望として、以下について引き続き研究を進めていきたいと考えている。

### (1) 他のIE手法に対する深層学習技術の適用

本研究では、IEの分析的アプローチに関する手法の中でも基本とされる「時間研究」と「動作研究」を対象として、深層学習技術を利用したシステムについて開発および評価を行った。いずれも製造現場の作業者の作業および動作を対象とした分析であるが、引き続き他のIE手法に関しても深層学習技術の適用について検討を行いたい。

### (2) 技能レベルの評価

製造現場の人材育成に継続的に取り組んでいくためには、作業者のモチベーションを維持することが重要であるものと考えられる。本研究内容を、非熟練者に対する作業手順の誤りの監視や非効率な動作の把握に留まらず、作業者の技能レベルを評価する仕組みとして応用することにより、作業者は自らの技能の熟練度を把握することが可能となり、技能向上におけるモチベーションにも繋がる可能性が考えられる。技能レベルの評価のためには技能に対する評価基準を明確にする必要があるため、この基準の在り方について調査を行いたい。

### (3) 技能伝承

製造現場の熟練技能者が持つ技能をどのようにして後継者へ伝承していくべきかについては、多くの企業で課題とされている。製造業における技術と技能に関して、2019年版のものづくり白書では、技術については何等かの客観的な表現によって記録された形式知を主体とするものであり、その人を離れても伝えることが可能であるが、技能については人に内在する暗黙知を主体とする能力であり、実際の体験などを通じて人から人へと継承されるものであるとされている[1]。つまり、技能は客観的にとらえることや言語化することが難しいためマニュアルなどで標準化することができず、暗黙知として一部の作業者の内部に存在しているものであり、その伝承においては、OJT (On the Job Training) による方法が中心とされている。このような技能伝承に対して、デジタル技術によりどのような貢献が可能であるか非常に関心が高い内容であるため、今後の研究テーマの1つとして考えたい。

参考文献

- [1] 経済産業省, 厚生労働省, 文部科学省: 2019年版 ものづくり白書, p.199, 入手先  
〈[https://www.meti.go.jp/report/whitepaper/mono/2019/honbun\\_pdf/pdf/honbun\\_01\\_03\\_01.pdf](https://www.meti.go.jp/report/whitepaper/mono/2019/honbun_pdf/pdf/honbun_01_03_01.pdf)  
〉 (参照2021-07-20).