Vol.2020-MBL-96 No.19
Vol.2020-UBI-67 No.19
Vol.2020-CDS-28 No.19
Vol.2020-ASD-19 No.19
2020/9/30

# 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システム

村井浩一[†1]　今井哲郎[†2]　荒井研一[†2]　小林透[†2]

**概要**：製造現場の作業は多岐に渡り，その技能や作業効率を評価するための仕組みについても，それぞれの作業により異なる．モーションセンサや画像処理を利用した従来の方法を用いる場合，各製造現場の作業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．そこで我々は，End-to-End の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動作の特徴を抽出することができる汎用的なシステムを開発した．本システムでは，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作業動作の特徴を抽出する．本システムについて一般的な作業のひとつである，スパナを利用したボルト締め作業を対象として評価実験を行い，その手法の有効性について明らかにした．

**キーワード**：深層学習，作業効率，動作分類，特徴抽出

# Motion Feature Extraction System of the Work in a Manufacturing Site by Regression CNN and Classification CNN

KOICHI MURAI[†1]　TETSUO IMAI[†2]
KENICHI ARAI[†2]　TORU KOBAYASHI[†2]

***Abstract***: There are many different types of a work in a manufacturing site, and methods for evaluating the skill and work efficiency also differs depending on the work. A number of existing systems which are to extract motion features by dedicated methods using motion sensors or processing of images therefore it is difficult to use for general-purpose. In order to use our system for many works in manufacturing sites, we have used end-to-end deep learning models which are prediction of the object location by the regression CNN and classification of motions in a work by the classification CNN to extract motion features. We have evaluated our system in the bolt tightening work using the spanner that is a general operation and have confirmed effectiveness of our method.

***Keywords***: deep learning, work efficiency, motion classification, feature extraction

## 1. はじめに

2020 年度版ものづくり白書によると，日本の経済成長を支える原動力は「人」であり，個々の労働者の労働生産性をより高めていくことが重要であるとされている[1]．製造業ではそのための手段のひとつとして，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．IE とは，「価値とムダを顕在化させ，資源を最小化することで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」とされており，価値を生み出す部分とそうではない部分を明確に区別し，それを定量的あるいは定性的に示すことであるとされている[2]．そのひとつとして，製造現場で働く作業者の労働生産性を高めるためには，作業者の動作を把握する必要がある．現在，映像やセンサを利用して作業時の動作を自動的に抽出するシステムは考案されているが，これらは対象となる作業に合わせた専用の仕組みを用いられている場合が多く，様々な製造現場で汎用的に利用することは難しい．

そこで本研究では，End-to-End の深層学習技術を用いることにより，作業時の動作の特徴を抽出することができる汎用的なシステムを開発した．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検出対象箇所を予測したうえで，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，動作の特徴となる物体の位置に関する時系列変化を用いることにより，把握したい動作の特徴を抽出する．本システムで抽出された作業動作の特徴について，熟練技能者と非熟練者との間で比較することにより，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．

2 章で関連研究について述べ，3 章で本システムの特徴について説明する．4 章で開発したシステム構成について述べた後，スパナによるボルト締め作業での評価実験結果を 5 章で述べ，6 章で考察し，7 章でまとめを行う．

†1 (株)システック井上
Systec Inoue Corporation, Nagasaki 852-8011, Japan
†2 長崎大学大学院工学研究科
Graduate School of Engineering, Nagasaki University, Nagasaki 852-8521, Japan

ⓒ2020 Information Processing Society of Japan

Vol.2020-MBL-96 No.19
Vol.2020-UBI-67 No.19
Vol.2020-CDS-28 No.19
Vol.2020-ASD-19 No.19
2020/9/30

## 2. 関連研究

製造現場での作業者の動作を抽出する研究としては，Microsoft 社の Kinect を利用した方法があげられる．松原らは，製造現場での改善を支援することを目的に，Kinect を用いた動作解析システムを開発した[3]．この研究では Kinect で計測した距離画像から形状の特徴を抽出することにより両手の位置を検出している．具体的には距離画像から身体部位の特徴となる頭部，胴体，腕部，手先等をその形状を特徴として抽出する方法を用いており，対象となる部位ごとに解析作業や設定が必要となる．

また，Kitazawa らはウェアラブルセンサを使用してリアルタイムに作業者の行動を分析するシステムを提案している[4]．この方法を用いる場合，作業者へ加速度センサを装着する必要があり，その取り付け箇所に制約が生じる点が課題である．また，Kitazawa らは画像解析による作業者のセンシングでは，解析パラメータの調整が困難であることを課題として指摘している．そこで我々は，CNN を用いることにより，製造現場の画像から学習したモデルだけで対象物体の位置を予測し，画像に対する解析パラメータの調整やセンサなどのデバイス装着を不要とした．

次に，CNN による動作検出の方法としては，OpenPose を利用した研究があげられる．スポーツを中心とした分野において研究事例が多いが，そのひとつとして中井らは，OpenPose で取得した姿勢データを利用して時系列データを作成し，ロジスティック回帰モデルによりバスケットボールの投入予測を行う仕組みを開発した[5]．CNN を用いることにより対象となる物体の位置を検出する手法は筆者らも同様であるが，筆者らは人物の関節位置だけではなく，製造現場で使用される工具なども検出対象とする必要があるため，OpenPose の利用は行わなかった．また，中井らの研究では時系列データをロジスティック回帰により二値分類を行っているが，筆者らは数種類への分類を必要とするため， CNN によるマルチクラス分類を利用した．

## 3. 作業動作特徴抽出システムの要件と特徴

### 3.1 要求条件

本システムの目的は，作業時の動画から作業動作の特徴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動作の特徴を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が可能となる．そのために必要な要件についてあげる．

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：幅広い製造現場で汎用的に利用できる

製造現場の動画から作業者の動作の特徴を抽出するためには，その動作を捉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，正確に予測する必要がある（要件 1）．さらに，対象となる物体の動きを，作業に含まれる幾つかの動作として正確に分類する必要がある（要件 2）．また，幅広い製造現場で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個々の製造現場への依存度を低くする必要がある（要件 3）．

### 3.2 ユースケース

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図 1 に示す．はじめに，ボルト締めの際の一連の動作から，特徴としたい動作を把握するためのスパナの検出箇所と抽出したい作業動作を定め，作業時の動画を元に CNN の学習データとして準備する．尚，作業動作を定める際には，作業分析用のソフトなどを用いる．次に，CNN の学習を行い，学習された CNN を用いて作業時の動画から，作業動作の特徴を抽出する．最後に，その結果を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その動作に対する特徴としての意味づけを行うことができる点である．

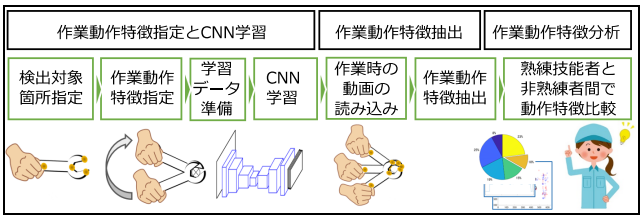



図 1 ユースケース
Fig. 1 Use case.

### 3.3 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．例えば，手作業の溶接作業においては手首の関節位置や溶接棒ホルダが対象となり，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，その対象は多岐に渡る．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[6]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設計やアルゴリズムの構築が必要となるために，導入対象に対しての専用システムとなる．そこで我々は，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渡る工具に対しても，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

ⓒ2020 Information Processing Society of Japan

Vol.2020-MBL-96 No.19
Vol.2020-UBI-67 No.19
Vol.2020-CDS-28 No.19
Vol.2020-ASD-19 No.19
2020/9/30

画像処理技術を利用した場合と比較すると，幅広い製造現場で汎用的に利用することが可能である．

次に，作業動作の特徴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，分類用の CNN により分類する仕組みとした．これを作業動作の特徴として熟練技能者と非熟練者との間で比較を行うことにより，効率的な動作とムダな動作を明らかにすることを可能とした．CNN を用いることにより，対象となる物体の移動距離や角度などを個別に解析する必要が無く，学習により汎用的に利用できる仕組みとした．

上記の通り，本システムでは作業動作の特徴を抽出するために 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．

## 4. 作業動作特徴抽出システムの構成

### 4.1 ハードウェア構成およびシステム環境

本システムのハードウェア構成は図 2 のとおりである．アプリケーションについては MATLAB R2020a とオプション製品である Deep Learning Toolbox を利用して開発した[7].




図 2　ハードウェア構成
Fig. 2 Hardware configuration.

### 4.2 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 らの画像からの人物の姿勢検出に関する研究[8]を参考にした．Xiao らの研究と同様に，CNN のネットワークモデルのひとつである ResNet-18[9]を用いて，検出対象箇所に関するヒートマップを推定する方法を用いた．具体的には 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，ファインチューニングして利用した（図 3）．

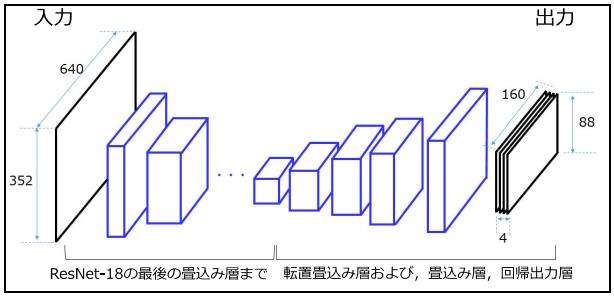



図 3　身体や工具位置予測用の CNN ネットワーク
Fig. 3 CNN for body and tool location prediction.

### 4.3 作業動作の分類処理

作業動作の分類については，Wang らが研究した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方法[10]を参考とした．Wang らは完全畳み込みネットワーク（FCN）を利用して時系列データを分類するための，End-to-End によるシンプルなアプローチとしてこの方法を提案している．今回のシステムでは，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．尚，取り扱うデータが低次元であるため，CNN についても 5 層の畳込み層で構成されたシンプルなネットワークとした（図 4）．

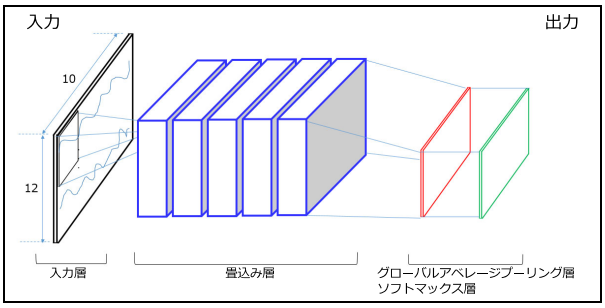



図 4　作業動作分類用のネットワーク
Fig. 4 CNN for motion classification.

具体的なネットワークの入力用となる時系列データの要素としては，それぞれの検出対象箇所に対して，画像内の $x$ 座標，$y$ 座標およびひとつ前の画像からの移動距離を含めた 3 要素を対象とした．このデータを予め指定したウィンドウサイズ $n$ をひと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）を用いて入力情報とした．つまり，ウィンドウサイズ $n$ の間での検出対象箇所に対する画像内での位置情報と移動距離に関する情報を時系列データとして入力し，各動作として分類された結果を出力するネットワークとなる．

ⓒ2020 Information Processing Society of Japan

Vol.2020-MBL-96 No.19
Vol.2020-UBI-67 No.19
Vol.2020-CDS-28 No.19
Vol.2020-ASD-19 No.19
2020/9/30

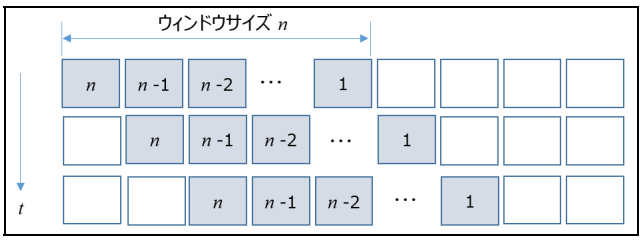



図 5　スライディングウィンドウ方式での時系列データ
Fig. 5 Sliding window method for time series data.

### 4.4 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. 評価実験

### 5.1 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 6）の位置に対する予測精度である．2 点目は，スパナを使用した際の作業動作に関する分類精度である．この分類については，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，今回は作業分析結果に基づき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 7 の通りに分け，表 1 の通り 6 種類に分類した．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．

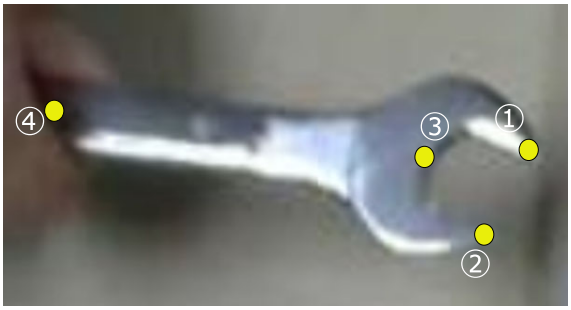



図 6　スパナの検出対象箇所
Fig. 6 Detection points in Spanner.

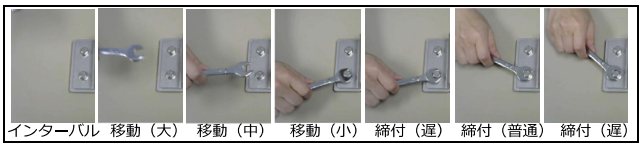



図 7　ボルト締め作業動作
Fig. 7 Bolt tightening motions.

表 1 ボルト締め作業動作の分類
Table 1 Classification of Bolt tightening motions.

| 分類 | 対象の動作 |
|---|---|
| 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． |
| 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． |
| ⑥締付（遅） | ボルトをゆっくりした速度で締め付ける動作． |

### 5.2 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 8207 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 6 に示したスパナの 4 箇所を対象に，図 8 の様なヒートマップを作成した．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，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．

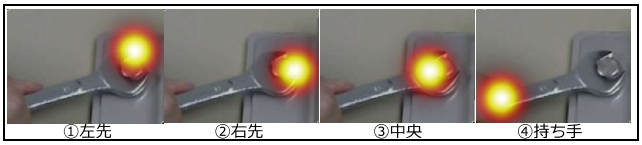



図 8　検出対象箇所のヒートマップ
Fig. 8 Heat map of detection points.

#### (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．動画に対してスパナの 4 箇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，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データとなる．この学習データに対して，表 1 に従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，被験者 B は 234 件を同様に時系列データとして準備した．

尚，時系列データの一例について，その内訳を図 9 に示す．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 箇所に対する画像内の $x$ 座標および $y$ 座標の 8 次元のデータと，同じく検出対象箇所のフレーム間の移動距離に関する

ⓒ2020 Information Processing Society of Japan

Vol.2020-MBL-96 No.19
Vol.2020-UBI-67 No.19
Vol.2020-CDS-28 No.19
Vol.2020-ASD-19 No.19
2020/9/30

4 次元のデータが 10 フレーム分で構成されたものである．

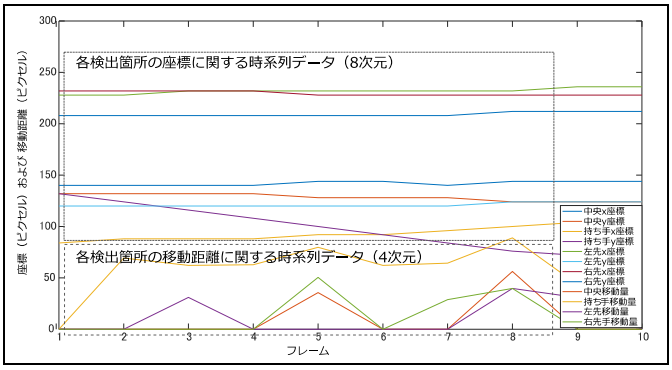



図 9　ボルト締め付け動作時の学習データ
Fig. 9 Training Data for Bolt tightening motion.

### (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 はミニバッチサイズ 32 で 100 エポック学習した．図 10 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．

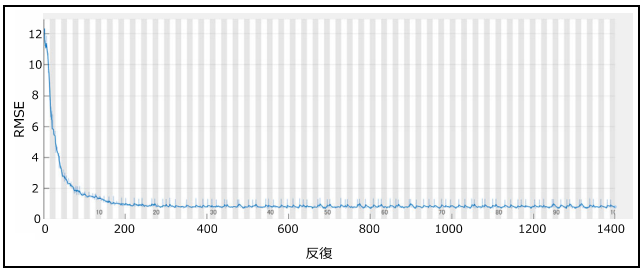



図 10 スパナ検出対象位置予測用 CNN の学習結果
Fig. 10 Learning result of CNN for Spanner location prediction.

### (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 についてはミニバッチサイズ 128 で 500 エポック学習した． 図 11 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．

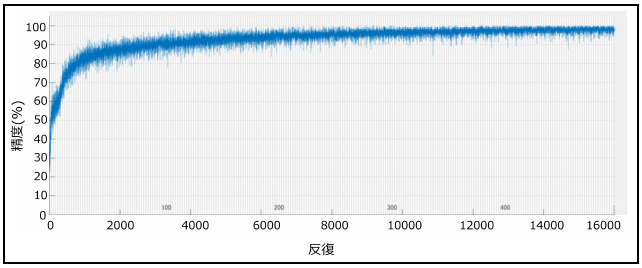



図 11 作業動作分類用 CNN の学習結果
Fig. 11 Learning result of CNN for motion classification.

## 5.3 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予測精度については，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，検出対象箇所の予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 箇所の位置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）[11]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）用いた（図 12）．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 3%，縦方向では 5%の誤差に相当する．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 12　スパナの検出精度の指標
Fig. 12 Index of prediction accuracy for Spanner.

### (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，予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

スパナを利用したボルト締め作業の場合，熟練技能者と非熟練者の間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する．熟練技能者と非熟練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．尚，この確認については，CNN の学習に使用していないデータを用いて実施した．

ⓒ2020 Information Processing Society of Japan

Vol.2020-MBL-96 No.19
Vol.2020-UBI-67 No.19
Vol.2020-CDS-28 No.19
Vol.2020-ASD-19 No.19
2020/9/30

### 5.4 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 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．また，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．今回実験で用いた全 800 箇所（2 名*100 枚*4 箇所）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 13 のヒストグラムで示す．

表 2　スパナの検出箇所予測精度

Table 2 Prediction accuracy of detection points in Spanner.

| 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% |

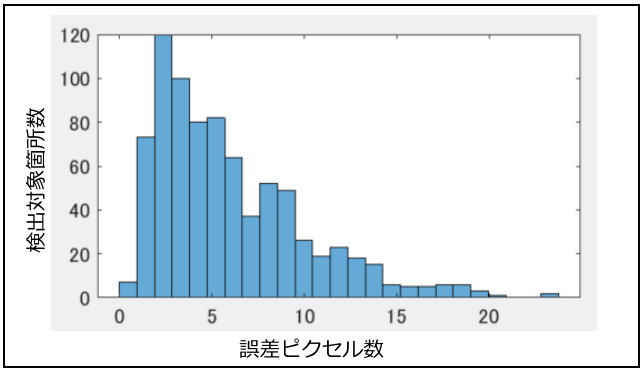



図 13 スパナ検出箇所予測誤差ヒストグラム

Fig. 13 Prediction error histogram of Spanner location.

#### (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 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．この混同行列の中で，非熟練者 A の“③移動（中）”や“④締付（普）”の様に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 14 で確認できる．この図は正解ラベルとＣＮＮ分類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 3　ボルト締め作業動作分類精度

Table 3 Classification accuracy of Bolt tightening motion.

| 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% |

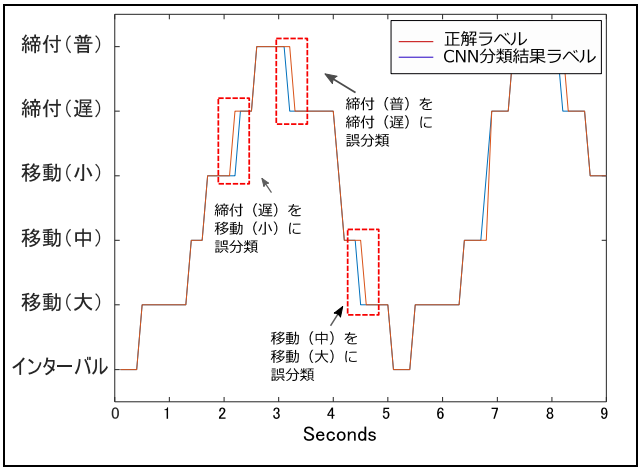



図 14 誤分類ケース

Fig. 14 Misclassification case.

#### (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 15 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．

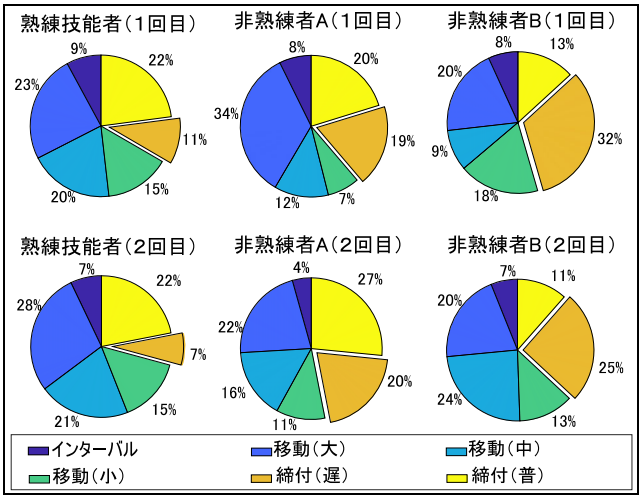



図 15 動作内訳の比較

Fig. 15 Comparison of motion details.

ⓒ2020 Information Processing Society of Japan

Vol.2020-MBL-96 No.19
Vol.2020-UBI-67 No.19
Vol.2020-CDS-28 No.19
Vol.2020-ASD-19 No.19
2020/9/30

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．

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 16）．ボルト締め動作については，図 7 に示した通り，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 16 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 回目共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．

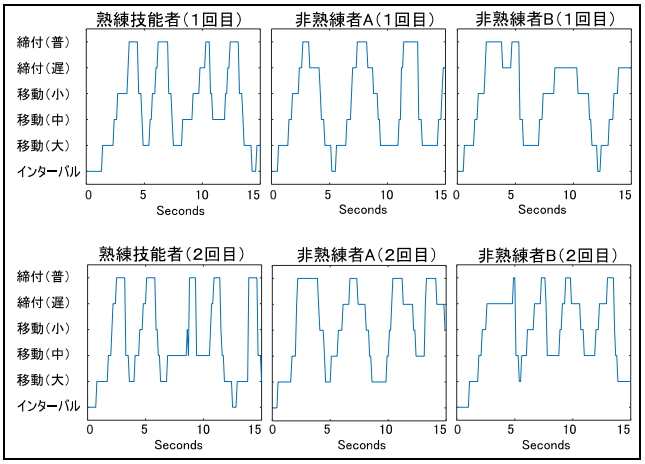



図 16 時系列動作推移

Fig. 16 Time series movement transition.

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 6 の③）と持ち手部分（図 6 の④）を時系列にプロットしたグラフを図 17 に示す．青のマーカーは締付（普）で赤のマーカーは締付（遅）の動作を示す．6 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 18）．熟練技能者については 4 箇所のボルト位置共に 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わかる．

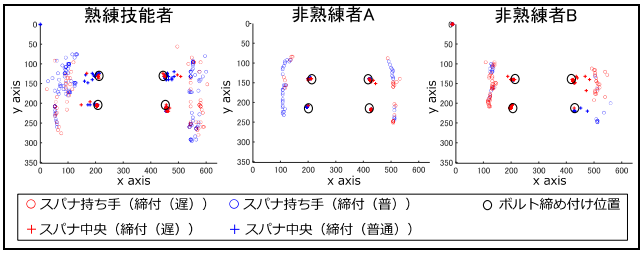



図 17 スパナの時系列位置情報

Fig. 17 Time series position information of Spanner.

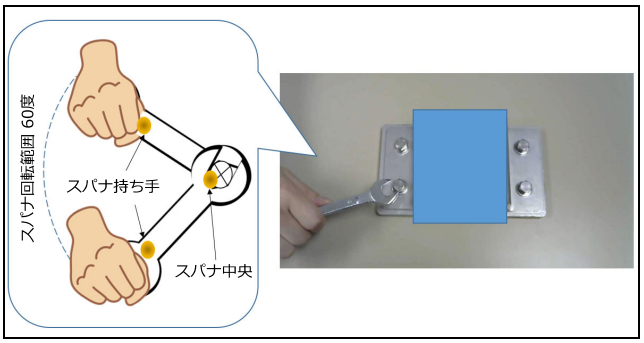



図 18 スパナ回転範囲

Fig. 18 Spanner rotation angle.

## 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．3.1 節で示した 3 つの要求条件に関して，5 章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**

本研究では，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．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 98.38%，その半分のサイズ (9 ピクセル) を閾値とした場合で 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，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**

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．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

ⓒ2020 Information Processing Society of Japan

Vol.2020-MBL-96 No.19
Vol.2020-UBI-67 No.19
Vol.2020-CDS-28 No.19
Vol.2020-ASD-19 No.19
2020/9/30

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) 幅広い製造現場でも汎用的に利用できること**

今回は，ボルト締め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.今回はスパナを用いた作業を対象としたが，他の作業においても,把握したい作業動作の特徴に併せて，製造現場での作業時の画像を用いて CNN の学習を行うことにより汎用的に利用することが可能である．

## 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，画像から作業時の動作の特徴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 で学習を行う事により，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，一般的な作業のひとつである金具組立時のボルト締め作業を対象に評価実験を行い，その有効性について示した．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映像から把握できることが前提となる．本手法が適用可能な具体的な製造現場や作業について，今後調査を行いたい．

また，作業動作の特徴に対する評価については，熟練技能者と非熟練者の間で比較する方法を用いたが，この比較方法については，作業動作に関する時系列データに対して作業効率や熟練度に関するラベルを付加した教師データとすることにより，CNN モデルとすることも可能である．これにより，作業効率の評価に関する殆どの部分を End-to-End の仕組みで実現することも可能となる．分野は異なるがロボット支援による外科手術に対するスキル評価のための研究[12]も進められており，これらを参考に今後さらに研究を進めたい．

## 参考文献

[1] 経済産業省：2020 年版ものづくり白書，p. 127，入手先〈https://www.meti.go.jp/report/whitepaper/mono/2020/honbun_pdf/pdf/all.pdf〉(参照 2020-06-04).
[2] 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〈http://www.j-ie.com/about/about-ie/〉 (参照 2020-06-04).
[3] 松原早苗，渡辺博己，曽賀野健一，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の開発，岐阜県情報技術研究所研究報告，第 18 号，pp. 24-29 (2016).
[4] Kitazawa, M., Takahashi, S., Takahashi, B. T., et al. : Real Time Workers' Behavior Analyzing System for Productivity Measurement Using Wearable Sensor, SICE Journal of Control, Measurement, and System Integration, Vol. 10, No. 6, pp. 536–543 (2017).
[5] 中井眞人，角田善彦，孫財東，ほか：OpenPose によるバスケットボール投入予測，第 32 回人工知能学会全国大会(2018).
[6] Lowe, G. D. : Distinctive Image Features from Scale-Invariant Keypoints, International Journal of Computer Vision (2004).
[7] Mathworks: MATLAB-計算技術言語, Mathworks，入手先〈https://jp.mathworks.com/products/matlab.html〉 (参照 2020-05-18).
[8] Xiao, B., Wu, H., Wei, Y. : Simple Baselines for Human Pose Estimation and Tracking, *Proc. European Conference on Computer Vision* (2018).
[9] He, K., Zhang, X., Ren, S. and Sun, J. : Deep residual learning for image recognition, *Proc. IEEE conference on computer vision and pattern recognition*, pp. 770-778 (2016).
[10] Wang, Z., Yan, W., Oates, T. : Time Series Classification from Scratch with Deep Neural Networks: A Strong Baseline, International Joint Conference on Neural Networks, pp.1578-1585 (2016).
[11] 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).
[12] Fawaz, I. H., Forestier, G., Weber, J., et al. : Accurate and interpretable evaluation of surgical skills from kinematic data using fully convolutional neural networks, International Journal of Computer Assisted Radiology and Surgery, vol.14, pp. 1611-1617 (2019).

ⓒ2020 Information Processing Society of Japan