

製造現場の作業分析結果を利用した カメラと深層学習による作業手順のリアルタイム監視システム

村井浩一^{†1} 今井哲郎^{†2} 荒井研一^{†2} 小林透^{†2}

概要: これまで製造現場の品質向上を目的として、作業ミスをリアルタイムに検知するシステムが提案されているが、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様なシステムは報告されていない。そこで我々は、作業改善のための作業分析結果を利用して、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できる作業手順リアルタイム監視システムを開発した。本システムでは、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監視基準に用いることにより、作業手順の誤りや作業時間の乖離が検知された際に、分析時の模範となった正しい作業手順で作業者にティーチングを行うことを可能とした。さらに、この深層学習に用いる学習データについて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。本システムについて組立作業の模擬環境で実験評価を行い、有効性について明らかにした。

キーワード: 作業分析, リアルタイム監視, 深層学習, Webカメラ, インダストリアル・エンジニアリング

Real-time work procedure monitoring system by deep learning and web camera using results of work analysis at manufacturing

KOICHI MURAI^{†1} TETSUO IMAI^{†2}
KENICHI ARAI^{†2} TORU KOBAYASHI^{†2}

Abstract: The systems to detect work errors in real time have been proposed for the purpose of quality improvement at manufacturing sites, however it has not been reported a system that instructs to rework detected work errors in a correct procedure. Therefore, we have developed a system that monitors work procedures in real time and conducts a teaching in correct procedures, if there are any errors, using results of work analysis for improvement. This system enables to perform a teaching to workers in a correct procedure which is a model in the analysis, when it detects errors in work procedures or deviations of work hours by using a deep learning model learned from standard work and working hours and videos of working in the analysis. Furthermore, by automatically generating learning data used for this deep learning from results of work analysis and videos recording workers' working, the system can remarkably reduce time and labor required for preparation of learning data which is considered as a problem in using deep learning. We conducted an experimental evaluation of this system in a simulated environment of assembly work and clarified its effectiveness.

Keywords: work analysis, real-time monitoring, deep learning, Web camera, industrial engineering

1. はじめに

2019年度版ものづくり白書によると、製造業における企業収益は大企業・中小企業ともに改善傾向にあり製造業が徐々に力を取り戻しつつある中、一方で人材不足が深刻な課題とされている。また、ものづくりに携わる人材が不足している状況では経済成長を支えていくことが難しく、その対策として製造現場での労働生産性をあげることが必要とされている[1]。製造現場での労働生産性向上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(以降IE)が利用されている。IEとは日本インダストリアル・エンジニアリング協会によると、“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すとする見方・考え方であり、それを実現する技術[2]”とされている。筆者はこれまでIEツールのひと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ストップ・ウォッチ法[3]による分析やサブリング分析[4]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。従来、その標準作業や標準作業時間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、その内容および結果に対する客観性や正確さが課題とされていた。既に映像やセンサ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、今後は作業ミスの検知に加え、生産性向上のための作業改善を目的

^{†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

とした IE の改善活動の一環としての作業手順監視システムが必要とされる。そこで本研究では、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および、分析時に使用された作業動画を監視用の基準として、製造現場での作業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員へ正しい作業手順をティーチングするシステムを開発した。本システムでは製造現場での作業状況を画像から判断するモデルに、深層学習のひとつである畳み込みニューラルネットワーク（以降 CNN）を利用する。CNN を利用する際は大量の学習データが必要となるために学習用の画像の準備やラベル付けで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[5]。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。

つまり、筆者らが開発したリアルタイム作業手順監視システムでは、作業分析結果から製造現場を監視するための基準となる情報および CNN の学習データを自動的に生成し、これらを利用することにより、IE 活動の一環としての作業手順および作業時間の異常検知と、その際に作業員へ正しい作業手順をティーチングすることを可能とした。

本研究では製造現場の組立作業を模擬した環境で実験を行い、システムの有効性について明らかにした。

2 章で関連研究について述べ、3 章で本システムの特徴を説明する。4 章で開発したシステム構成について述べた後、5 章で評価実験結果、6 章で評価実験結果を考察し、7 章でまとめを行う。

2. 関連研究

製造現場での作業手順の監視システムとしては、岩手県立大学大学院の宇田らがピッキング工程における作業ミスの検知を目的として Microsoft 社の Kinect を使用したシステムの研究を行っている[6]。この研究ではピッキング時の手の位置について Kinect で撮影した画像と深度センサーで捉える方法を利用しており、手を検知する際に作業時に着用する軍手の色情報を利用して、作業環境に合わせた軍手の色の選定が必要となる。また、棚の間口と手の位置関係を利用してピッキング動作を判断しているため、誤検知を避けるために隣り合った間口に平行に手を動かさない様、作業員の動作にも一定の制約が求められ、制約の有無で誤検知率に差が生じることが評価実験結果として述べられている。一方、筆者らは普段の作業時の画像に基づいて CNN で作業を分類する。上記の様な場合でも“手の移動”と“間口からのピッキング（間口から特定の部品を取り出す動作）”をそれぞれ画像から学習することにより分類が可能となる。従って、作業員への動作の制限については緩和される。

また、組立作業の分析に関する研究として中京大学大学院の渡邊らが、組立作業分析のためのチャートの自動記述およびセンシングシステムを提案している[7]。渡邊らは動作を把握する方法として、手と組立対象物体の映像に対して画像処理を行いそれぞれパーツとして抽出したうえで、画像内のパーツの占有率からそれぞれの存在確率を求め、事前知識として定義したお互いの干渉領域における距離値の変化から判断している。従って、対象となる物体や作業環境毎に個別に条件設定が必要となる。これに対し筆者らは CNN を用いることにより、画像処理による物体の抽出や位置関係の算出、設定にかかる作業を不要とした。

また、改善を目的とした作業分析の研究としては、岐阜県情報技術研究所の渡邊らが作業時間計測を支援するために Kinect と赤外線マーカを使用して作業を捉えるシステムを提案している[8]。渡邊らのシステムでは作業員の両手首に赤外線マーカを取り付け、作業毎に複数設定した動作領域（通過判定エリア）に対するマーカの通過状況を Kinect で捉えることにより、エリア間の通過時間差から動作時間を算出している。そのため作業毎に通過判定エリアの設定が必要となり、作業数の増加に比例して設定も増加する。一方筆者らは、CNN で画像が各作業に分類された際の時刻から作業時間を算出しており、CNN での作業分類結果に応じて自動的に作業時間の計測が可能である。

筆者らの研究では、作業分析ソフトでの分析結果から監視基準となる標準作業と標準作業時間および CNN の学習データを自動的に生成する仕組みにより、作業分析を行う過程で監視に必要なこれらの要素が同時に準備され、監視にかかる設定作業の軽減をはかることを可能とした。また筆者らは作業分析結果を利用して監視を行い、作業手順の誤りやムダによる作業時間の乖離が検知された際に、正しい手順をティーチングするシステムを研究対象としている点で、関連研究とは異なっている。

3. リアルタイム作業手順監視システムの特徴

3.1 要求条件

作業改善を目的とした IE 活動の一環としてのリアルタイム作業手順監視システムの要件は、以下が求められる。

- 要件 1：作業改善を目的とした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：監視において作業員の作業の妨げにならないこと
- 要件 4：作業手順誤りや時間乖離を正確に検知できること
- 要件 5：製造現場をリアルタイムで監視可能であること

作業改善を目的とした IE 活動における監視のためには、その監視基準は IE として分析評価され、定められた内容にする必要がある。（要件 1）。また、製造現場では IE 活動によるサイクルが回され、監視の基準についても頻繁に見直される。従って監視に使用する CNN についても IE 活動のサイクルの中で再学習が必要となるため、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、製造現場へ設置される監視機器が作業者の作業の妨げとなり生産性に悪影響を及ぼす可能性を配慮し、使用する機器は最小限に留める必要がある(要件3)。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、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、影響を最小限に留めるためには問題発生時点で直ちに警告し、作業者へ正しい作業手順でティーチングする必要がある(要件4,5)。

3.2 作業手順監視方式

3.1 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1に示す。本方式では、市販の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動画により学習したCNNを監視の際の基準として利用することにより、作業改善を目的とした監視基準とする(要件1)。また、作業分析結果と動画からCNNの学習データを自動的に生成することにより学習用データの準備を容易にする(要件2)。製造現場のモニタリングについては、取り付け位置の制限が少ない小型のWebカメラを利用することにより作業の妨げにならない様にする(要件3)。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る(要件4,5)。

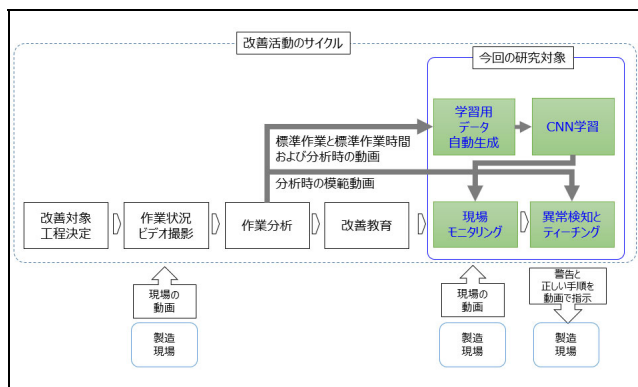


図1 リアルタイム作業手順監視方式

Fig. 1 Method of work procedure monitoring in real time

3.3 ユースケース

本方式に基づいたユースケースを図2に示す。作業分析の担当者は作業分析ソフトで分析を行い、改善のための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。その際の作業分析結果である標準作業と標準作業時間および動画は、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、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従って作業を行うが、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異なる場合は同システムが異常として検知し、警告と共に正しい作業手順を動画と共に作業者へ示す。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へ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、本システムで製造現場を監視する際の基準及びティーチングの際の作業手順を同一の内容にできる点であり、作業分析結果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。

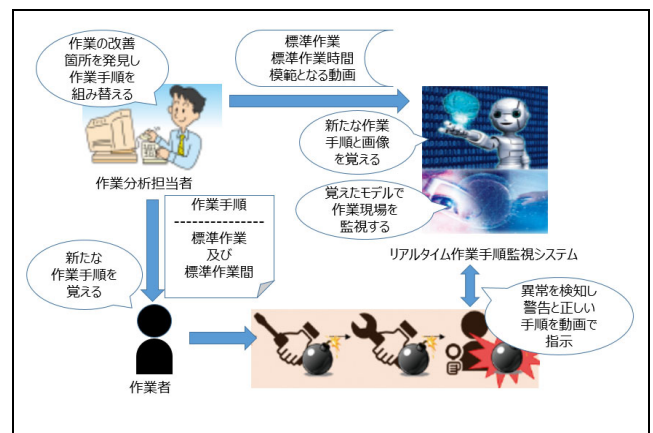


図2 ユースケース

Fig. 2 Use case

3.4 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を使用する。製品Aでは撮影したビデオ映像を、簡単なマウス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。

3.5 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を効率的に学習できる様、学習済ネットワークを転移学習して利用する。CNNの学習済ネットワークモデルとしては多くの種類が挙げられるが、今回はリアルタイムでの異常検知を要件としているため、Alfredo Canzianiらの評価結果[9]を参考に、比較的スループットが高いとされているAlexnet[10]を採用する。Alexnetでの分類精度は他のネットワークモデルと比較すると高いとはされていないが、分類精度とスループットはトレードオフの関係にあるため、今回はスループットを優先した。

4. リアルタイム作業手順監視システムの構成

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 3 の通りである。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 を利用して開発する[11]。



図 3 ハードウェア構成
 Fig. 3 Hardware configuration

4.2 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

製品 A では、作業分析結果を Excel ファイルまたは CSV として出力することができる。表 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の条件としても本分析結果を利用する。

表 1 作業分析ソフトでの出力結果
 Table 1 Result in work analysis software

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 1 にある作業順番 (No.) と作業分類、計測時間に基づいて動画から静止画を抽出する (図 4)。具体的には、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) と動画データのビデオフレームの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、各作業の名称を付けたフォルダに保存する処理をシステム化する。

MATLAB では各フォルダに保存された静止画に対してイメージデータストアと呼ばれるオブジェクトを経由してアクセスするため、学習用のラベル名には各フォルダ名を指定する。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した。

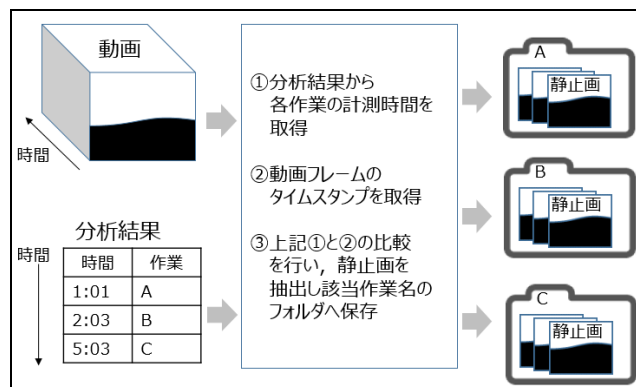


図 4 動画からの静止画の抽出処理
 Fig. 4 Extraction of still images from videos

(3) CNN 転移学習

使用する CNN については、Alexnet の最後の結合層および Softmax 層、出力層について今回のデータにあわせてクラス数の調整を行い、学習させたものを利用する。尚、再利用する転移層の学習率は 0.001 と小さくし、調整した全結合層については、重みとバイアスの学習率をその 10 倍に設定して学習させる。

(4) 撮影中の製造現場の作業分類と作業時間集計

製造現場で行われている作業の分類については、カメラで撮影中の画像をフレーム毎に静止画として抽出し CNN で分類処理する。また、作業時間については、各作業に分類された際の時刻から経過時間を集計する。

(5) 検知およびティーチング

作業手順の誤りに関する検知については、標準作業に対する作業の順番と CNN で分類された作業の順番を比較することで判断を行う。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員へのティーチングを行う。

5. 評価実験

5.1 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とその中の特定作業内の動作を対象に、監視対象の粒度が異なる 2 種類の模擬環境を準備した。

(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、2 種類の金具をボルトで組み立て、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という6つの作業で構成される実験環境を準備した(図5)。

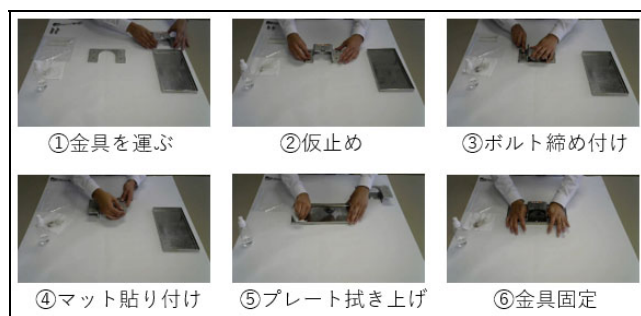


図5 金具組立作業

Fig. 5 Bracket assembly work

(2) サイクル内の1作業内に対する実験環境

1 作業内の動作に関する実験環境として、本組立作業の中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6は作業者を対面方向からカメラで撮影した画像である。組み立てる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。

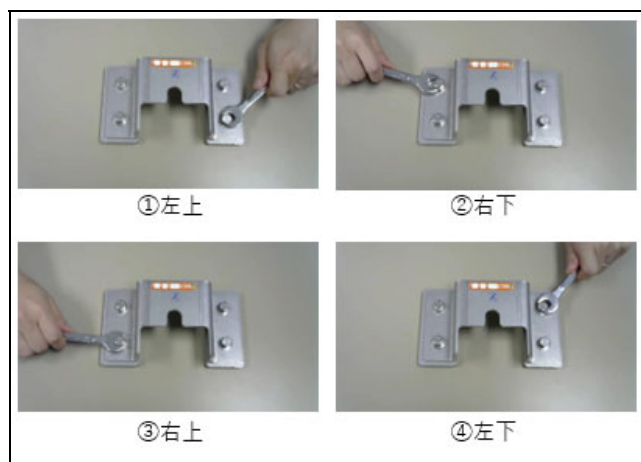


図6 ボルト締め付け作業

Fig. 6 Bolt tightening work

5.2 実験対象

本システムに対する評価項目については、以下の3点を対象とした。1点目は、CNNにおける静止画の分類精度である。今回の要件のひと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、標準作業からの逸脱および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。

5.3 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った。作業時の動画から学習用の静止画を自動生成する処理時間の評価を目的として、ボルト締め付け作業を対象に被験者3名の動画で試したところ、表2に示す通り毎秒平均36枚の静止画を生成できることを確認した。

表2 学習用データの生成処理性能

Table 2 Generation processing performance of data for learning

	動画時間(秒)	静止画(枚)	生成所要時間(秒)	生成枚数@秒
被験者A	22.39	667	19	35.11
被験者B	20.26	604	15	40.27
被験者C	29.07	866	25	34.64
平均	23.91	712.33	19.67	36.22

尚、CNNの学習に使用する金具組立作業の学習データについては、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つのクラスに分類した。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5の様な学習用の静止画を生成し表3の通り作業毎に3,000枚に揃えた。3,000枚に揃える際に被験者Aから先に静止画を生成し、不足分を被験者Bから補ったため被験者間で画像数が異なっている。

表3 金具組立作業の学習用データ画像数(枚)

Table 3 Learning data for bracket assembly work

	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	1,490	1,722	1,812	1,654	1,587	2,029	10,294
被験者B	1,510	1,278	1,188	1,346	1,413	971	7,706
合計	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。分析結果に対して表4の通り、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6の様な学習用の静止画をそれぞれ3,000枚準備した。”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動作を明確に別けるためである。

表4 ボルト締め付け作業の学習用データ画像数(枚)

Table 4 Learning data for bolt tightening work

	①左上	②右下	③右上	④左下	⑤その他	合計
被験者A	3,000	3,000	3,000	3,000	3,000	15,000

(2) CNNの学習

金具組立作業に関する学習結果について図7へ示す。ミ

ニバッチサイズ 32 で 15 エポック学習した。尚、学習時間は 2,688 分で、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、99.72% であった。

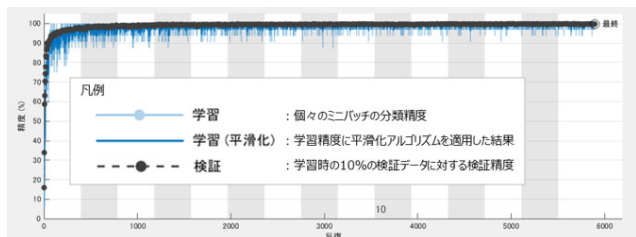


図 7 金具組立の静止画学習結果

Fig. 7 Learning result of bracket assembly

同様に、ボルト締め付け作業の学習時の状況について図 8 に示す。学習時の条件は金具組立作業と同じくミニバッチサイズ 32 で 15 エポック学習した。学習時間は 1,159 分で、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、99.53% であった。

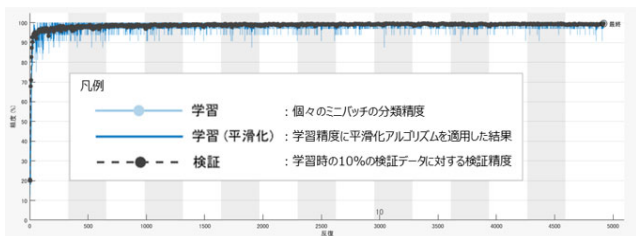


図 8 ボルト締め付けの学習結果

Fig. 8 Learning result of bolt tightening

5.4 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被験者 A および B の 4 サイクル目の作業と、非熟練である被験者 C の 1 サイクル目の作業をテストデータとして静止画の分類精度を算出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、予め作業分析ソフトで分析を行い、作業毎に静止画として抽出した後に正解である作業のラベルを付けたものを利用し、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。

また、ボルト締め付け作業については、被験者毎の特徴にバラツキが少ないという仮定を確認するために、学習データに含まれない非熟練の被験者 3 名の作業をテストデータとして、上記同様に 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算出することを実験内容とした。

(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、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の順番(左上->右下->右上->左下)が決まっており、図 9 の通りとなる。

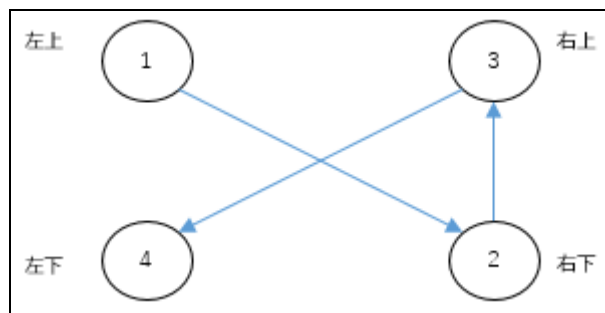


図 9 ボルト締め付けの順番

Fig. 9 Bolt tightening sequence

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、標準作業と実際の作業手順の誤りが発生した際の検知の正確さについて確認することを実験内容とした。尚、CNN における多少の誤分類を想定し、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×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を加えた。尚、今回は N=5 の条件で評価実験を行った。

また、標準作業時間に対する実際の作業時間の乖離に関する実験は、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。

(3) リアルタイム性

処理のリアルタイム性については、本システムにおいて中心となる機能である CNN での作業分類処理を対象として、前記“(1)静止画に対する分類精度”の際に利用した動画を利用して 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することを実験内容とした。

5.5 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 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、3 名平均の分類精度については 93.9% であった。

但し、被験者 B の“①金具を運ぶ”や被験者 C の“⑤プレート拭き上げ”の様に、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 10 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画像の類似性について示す。

表 5 金具組立作業の分類精度

Table 5 Classification accuracy of bracket assembly work

被験者A	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	100.00%	0.00%	0.00%	0.00%	0.00%	0.00%
②仮止め	0.22%	99.78%	0.00%	0.00%	0.00%	0.00%
③ボルト締め付け	0.00%	0.40%	98.95%	0.66%	0.00%	0.00%
④マット貼り付け	0.00%	0.00%	0.32%	93.41%	6.27%	0.00%
⑤プレート拭き上げ	0.00%	0.00%	0.00%	0.00%	100.00%	0.00%
⑥固定	0.00%	0.00%	0.00%	0.00%	1.26%	98.74%
平均						98.48%

被験者B	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	82.99%	16.67%	0.35%	0.00%	0.00%	0.00%
②仮止め	0.00%	98.44%	1.56%	0.00%	0.00%	0.00%
③ボルト締め付け	0.00%	3.83%	94.99%	1.18%	0.00%	0.00%
④マット貼り付け	0.00%	0.00%	8.91%	91.09%	0.00%	0.00%
⑤プレート拭き上げ	0.00%	0.00%	0.00%	5.40%	92.27%	2.34%
⑥固定	0.00%	0.00%	0.00%	0.00%	5.21%	94.79%
平均						92.43%

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%

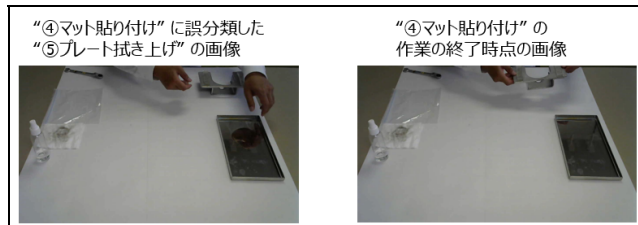


図 10 前後の作業への誤分類ケース
 Fig. 10 Misclassification case between work

また、ボルト締め付け作業に対する静止画の分類精度を表 6 に示す。3 名平均の分類精度は 96.5%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、この原因は前記の通り類似した静止画が両方に含まれているためである。

表 6 ボルト締め付け作業に対する分類精度

Table 6 Classification accuracy of bolt tightening work

被験者A	その他	右下	右上	左下	左上
その他	95.54%	1.79%	0.89%	0.45%	1.34%
右下	0.85%	99.15%	0.00%	0.00%	0.00%
右上	4.35%	0.00%	95.65%	0.00%	0.00%
左下	6.14%	0.00%	0.00%	92.11%	1.75%
左上	1.68%	0.00%	0.00%	0.00%	98.32%
平均					96.15%

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) 作業手順ミス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 7 パターンで評価した結果、正常手順 1 パターンは正しく、誤った手順 6 パターンについては異常として正確に検知することができた。結果を表 7 に示すが、標準作業に対して異なる手順として本システムが検知した際の作業を異常検知対象として纏めたものである。

表 7 ボルト締め作業手順確認結果

Table 7 Bolting procedure check result

パターン	作業順番				被験者A 異常検知対象	被験者C 異常検知対象	被験者D 異常検知対象
	手順 1	手順 2	手順 3	手順 4			
標準作業	左上	右下	右上	左下	-	-	-
1	左上	左下	右上	右下	手順 2	手順 2	手順 2
2	左上	右上	右下	左下	手順 2	手順 2	手順 2
3	左上	右下	右上	左下	なし	なし	なし
4	左上	右下	左下	右上	手順 3	手順 3	手順 3
5	左下	左上	右上	右下	手順 1	手順 1	手順 1
6	右上	左下	左上	右下	手順 1	手順 1	手順 1
7	右下	左上	右上	左下	手順 1	手順 1	手順 1

また、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 8 の通りとなった。今回、標準作業時間をボルト締め一箇所あたり 4.5 秒に定め、各作業時間について CNN での分類結果と、その際の時刻から算出した結果を纏めた表である。今回の実験の際は、標準作業時間に対して±50%を作業時間監視の閾値としていたが、被験者 3 名ともにそれを超えることは無かった。

表 8 標準作業時間と実作業時間比較

Table 8 Comparison of standard time and actual time

	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、動画をフレーム毎に抽出しながら 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 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 9 Alexnet での静止画分類処理数

Table 9 Number of still images classified by Alexnet

	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

6. 考察

3.1 節で示した 5 つの要求条件に関して、4 章のシステム構成および 5 章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動画を CNN の学習用データとして生成し、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、学習用データを毎秒平均 36 枚自動生成することができた。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無い様な作業分析ソフトについても 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り、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) 作業者の作業の妨げにならないこと

今回の評価実験については社内環境で実施したが、作業員へ特別な制約条件を付加すること無く Web カメラ 1 台で有効性が確認できた。製造現場を想定した場合は、カメラ取り付け位置や配線などを検討する必要があるが、取り付け位置の関係でカメラとパソコン間の配線ができない様な場合は、市販のワイヤレス Web カメラも利用できる。

(4) 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、また、作業手順ミスの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。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方式でフィルタ処理を加えて影響を抑えた。また、前後の作業間で類似した静止画を誤分類するケースが認められたが、連続した一連の動作を分類するには、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（以降 RNN）の利用が挙げられる。今回の様に静止画で作業自体が分類できる様な場合においては、CNN とフィルタ処理だけで高い精度が認められたが、さらに精度を高めるためには、RNN についても検討したい。

画像による分類の場合、背景や服装の違いによる分類精度の低下が課題とされるが、本研究の対象となる製造現場では整理・整頓・清掃が習慣的に求められる。従って、同一製造現場において背景が変化することは少なく、作業員も同一の作業服を着用しているケースが多いため、分類精度に与える影響は低くなる。仮に背景の変化や作業服の違いなどの影響が課題となる場合は、画像内で必要となる箇所を画像処理して抽出することで解決できる。今後は作業手順（順番）の誤りだけではなく、標準作業として定義されていない想定外の動作（例えば、両手での作業を片手だけで行った場合）を検知するために、学習データの追加や深層学習のネットワークについて検討したい。

(5) 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。計算上はシステムとして 0.07 秒周期で画像を捉え分類することができる。スライディングウィンドウ方式でのフィルタ処理の連続フレーム数 $N=5$ に指定した場合、5 フレーム分遅れることになるが、その場合でも 0.35 秒周期で異常の検知を行うことが可能である。従って、製造現場で重大な作業ミスが発生した場合においても、即時に誤りを検知することが可能である。

7. おわりに

本研究では、作業分析ソフトでの分析結果を利用して、製造現場の作業手順を監視するためのシステムを開発した。作業改善を目的とした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。

本研究内容を応用することにより、熟練者の作業手順をデジタル化した技能伝承や人材育成に役立てることが期待される。今後はさらに検知精度を向上させるための深層学習の検討や、熟練者の技能を抽出するための手法についての研究を行う予定である。

参考文献

- [1] 経済産業省, 厚生労働省, 文部科学省: 2019 年度版 ものづくり白書, p.17, 2019.
- [2] “IE とは”. <http://www.j-ie.com/about/about-ie/>, (参照 2019-07-10).
- [3] 中村茂弘: 現場力を高める簡単 IE 実践マニュアル, 日刊工業新聞社, p.114, 2011.
- [4] 中村茂弘: 現場力を高める簡単 IE 実践マニュアル, 日刊工業新聞社, p.36, 2011.
- [5] 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, The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019.
- [6] 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, Vol6, No.1, pp.63-74, 2016.
- [7] 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, Vol82, No.5, pp.473-480, 2016.
- [8] 渡辺博己, 生駒晃大, 棚橋英樹: Kinect を用いた作業分析システムの開発, 岐阜県情報技術研究所研究報告第 19 号, pp.41-46, 2018.
- [9] Alfredo Canziani, Eugenio Culurciello, Adam Paszke: AN ANALYSIS OF DEEP NEURAL NETWORK MODELS FOR PRACTICAL APPLICATIONS, p.5, 2017.
- [10] Alex Krizhevsky, Ilya Sutskever, Geoffrey E. Hinton: ImageNet Classification with Deep Convolutional Neural Networks, 2012.
- [11] MATLAB. <https://jp.mathworks.com/products/matlab.html>, (参照 2019-07-10).