

## 機械学習を利用した作業動画からの作業内容推定

長岡技術科学大学 \*井若 玄貴 IWAKA Genki  
 長岡技術科学大学 片岡 翔太郎 KATAOKA Shotaro  
 長岡技術科学大学 稲垣 徹郎 INAGAKI Tetsuro  
 長岡技術科学大学 野中 尋史 NONAKA Hirohumi

### 1. はじめに

作業者の作業効率を向上させることは、作業現場における重要な課題である。従来、そのための方策としてビデオを用いた動画解析が行われており、現場のみならず生産管理における学術的研究においても使用されている[1]。しかし、作業動画の解析は有益ではあるものの多大な労力・時間を費やすものであり、自動化が求められている。

そこで本研究では、動画中の作業者の位置を予測し、そこから作業者の作業内容を判定する機械学習を利用した手法を提案する。

### 2. 関連研究

従来研究において動画以外の情報を用いて作業解析を自動的に行う手法が提案されている。中鶴らのウェアラブルセンサを用いた農作業識別手法の研究[2]は作業の生産効率を向上させるための作業内容を推定する手法の提案をしている。この研究はセンサにより筋電位・加速度・角速度を計測し、それらを用いて作業内容を推定するものである。これに対し、我々が提案する動画データでの作業分析は、以下の2つの点で実用的である。

- ・ センサを道具や複数人の作業者に装着させるよりも、カメラのみで運用可能な動画データの方がコストは抑えられ、導入も容易になる。また作業者への負担にもなりにくい。
- ・ 動画データはセンサデータより情報量が多い。

これらの理由から本研究では動画データから、作業内容の推定を試みた。

### 3. 提案手法

本研究では作業者の位置情報が作業内容の推定において最も重要な情報であると仮説を立てた。その根拠は一般的な工場では作業場所ごとに行う作業がある程度決まっており、その遷移を解析すれば作業内容を推定可能であるためである。

はじめに動画のフレームごとに畳み込みニューラ

ルネットワークを用いて作業者の位置を推定する。次に、得られた位置情報を LSTM(Long Short-Term Memory) [3]によって時系列解析し作業者の作業内容を推定する。

以下の図1に動画データから作業内容を推定する手法をまとめた。

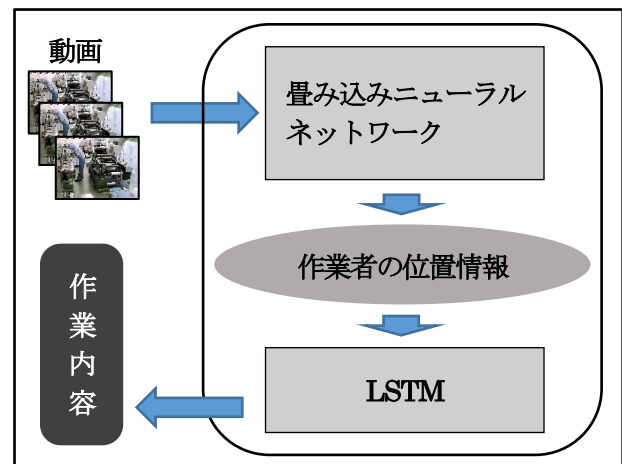


図1. 作業内容推定手法

#### ①作業者の位置情報の推定

作業者の位置情報として、「作業者自身の位置」、「作業者の手の位置」、「作業者の姿勢」の3つを推定する。学習に用いる教師データは以下の方法により作成した。



図2. 作業者の手の位置ラベル

教師ラベルは著者が協力企業における1日分の作業に関する動画データを確認し、作業内容推定のために重要だと思われる位置を選定した。その後、この選定した教師ラベルを作業者に評価してもらい、教師ラベルの有効性を確認した。

上記の図2はアノテーションの際に使用した、作業者の手の位置ラベルの定義を示したものである。

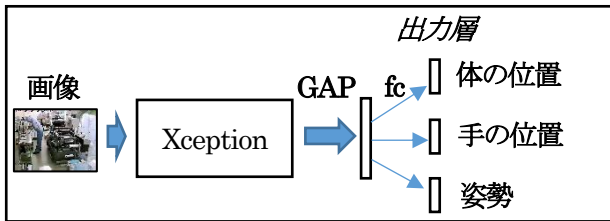


図3. 位置情報推定モデル

図3のモデルはXception[4]をベースに、「体の位置」、「手の位置」、「姿勢」をマルチタスクで学習するように設計した。このモデルでは3つのラベルを同時に推定するため、3つのラベルを別々のモデルで推定するよりも計算量に優れる。

#### ②作業内容の推定

作業内容の推定は、作業が位置の遷移による時系列データであるという仮定から、時系列データの予測に強いLSTMを用いることとした。また、作成した教師データの作業ラベルは69種類となった。

#### 4. 実験

作業者の位置情報の推定のための、畳み込みニューラルネットワークに用いるアノテーションデータは1330枚用意し、そのうちテストデータに400枚、930枚を訓練データとし学習した。テストデータでの検証結果(正解率)は、作業者の体の位置推定は93.5%、手の位置推定は94.25%、姿勢の推定は96.25%と算出された。正解率は以下の式によって定義している。

$$\text{正解率} = \frac{\sum_c^C N_{T_c}}{N} \quad (1)$$

(1)のNはデータの総数、Tは推定と実際のデータが一致したもの、Cはクラス(ラベル)の種類である。

#### 5. 考察と今後の課題

結果から位置情報の推定に関しては高い正解率を示した。しかし、手・体・姿勢の推定結果の間に矛盾が生じることがあった。この問題の原因としては、背景に映り込む作業員以外の人物が推定結果に影響を与えていると考えられる。また、教師データを作成する段階で、無作為に画像データを収集したところクラス間に偏りが生じていた。以下の図4に教師データのクラスの割合を可視化した。

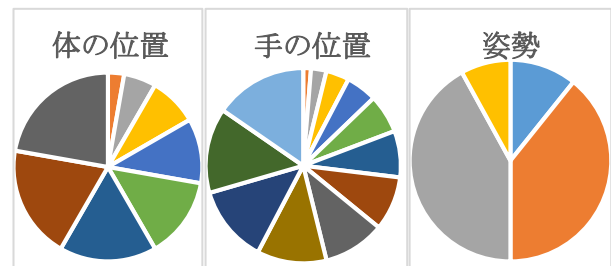


図4. 位置推定のための教師データのクラスの割合

図4から「体の位置」、「手の位置」、「姿勢」全ての教師データに偏りがあることが分かる。このため、各クラスが均等になるように教師データの作成を再検討する必要がある。

「作業内容の推定」を含めて発表当日に詳細な内容を発表する。

#### 参考文献

- [1] Adanna, I. W., and A. Shantharam. "Improvement of setup time and production output with the use of single minute exchange of die principles (SMED)." *International Journal of Engineering Research* 2.4 (2013): 274-277.
- [2] 中鶴 慎二, 福井 類, 割澤 伸一, "ウェアラブルセンサを用いた農作業識別手法の研究", 公益社団法人精密工学会, 2017年度精密工学会春季大会
- [3] Hochreiter, Sepp, and Jürgen Schmidhuber. "Long short-term memory." *Neural computation* 9.8 (1997): 1735-1780
- [4] Chollet, François. "Xception: Deep learning with depthwise separable convolutions." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2017.