論 文 内 容 の 要 旨

氏 名 （ 吉村直也 ）

論文題名	Research on Complex Activity Recognition in Industrial Domains （産業ドメインにおける複雑行動認識に関する研究）

論文内容の要旨

Many manufacturers have introduced robots, sensors, and IoT devices into their industrial sites, and used the collected data from them to improve the productivity. Specifically, sensors that monitor manual works by workers are expected to be used for incorporating the real-world work activity information into the cyberspace, enabling to simulate manufacturing processes with machine and human generated data. Although sensor data are now easily obtained by attaching a smartwatch or other sensors to a human worker, the data represent just a sequence of numerical values, not high-level information that describes the work activities. Therefore, the development of work activity recognition technologies that transform sensor data into high-level work activity information is required.

In this thesis, we aimed to address three challenges for the development of work activity recognition models and their practical applications. The first challenge is the amount of training data. The cost of collecting labeled training data to build a work activity recognition model for each worker in a factory is huge, and it is necessary to build the model with a limited amount of training data to reduce this cost. The second challenge is the usage of data from IoT devices. There are many IoT devices deployed in the site, which can be an important source of information for work activity recognition.
However, there is no dataset of for work activity recognition that includes both sensor data and data from IoT devices. Due to the unique characteristics of IoT device data, it is necessary to explore the ways to combine sensor data with IoT device data that take into account the characteristics of IoT device data.
The third challenge is the interpretability of neural network-based work activity recognition models. Neural networks are considered to be black boxes, but it is important to improve the interpretability of the models in order to introduce the models into the actual industrial sites.

Different workers on a production line are assigned to different tasks. In addition, workers may change the standard work processes depending on their skill level and other factors. Therefore, it is difficult to build a work recognition model for a target worker by using training data collected from other workers. We proposed a lightweight ordered-work segmentation network (LOS-Net), which is a lightweight model that extracts minimal necessary features and can be trained on limited training data from the target worker. We evaluated LOS-Net using data collected from 11 actual factory workers and found that LOS-Net improved recognition performance by about 10% on average (F1-measure macro average) compared to baseline methods.

Many IoT devices have been installed in industrial sites. The usage logs of the IoT devices strongly relates to the user's activities and can be a very important source of information for work activity recognition. However, because the data is generated only at the time when an IoT device is used, it is much sparse in time compared to general sensor data such as acceleration data. Therefore, a method to effectively fuse sensor data with IoT data should be explored. To accelerate activity recognition studies based on the sensor and IoT data fusion, we developed a large-scale multimodal dataset of packing work activity called the "OpenPack dataset." We collected over 53 hours of work activity data from 16 subjects and obtained 20,129 work operation labels and 52,529 action labels. This dataset provides a total of nine data modalities, including sensor data from inertial measurement units (IMUs), depth images, and keypoints, as well as usage logs from IoT devices. In addition, we proposed a method that effectively fuses acceleration data with IoT device data. We evaluated the model on the OpenPack dataset and improved the F1-measure (macro average) by 0.03 pt compared to existing methods.

In addition to the development of work activity recognition methods, we also addressed the issue of the interpretability of activity recognition neural networks. Many recent activity recognition models, including models developed in this study, rely on neural networks. There are few line managers who are familiar with ICT, and it is not easy for them to trust the results of neural networks, which are considered as black boxes. In this study, we applied a method called activation maximization to the sensor-based activity recognition model in order to make the features extracted by the neural network interpretable. Activation maximization is a method of visualizing feature representations extracted by a unit in a neural network by generating an input signal to which the unit of interest strongly responds. In generating such an input signal, the design of the regularization term is important. We proposed two regularization methods that use the unit outputs to determine if anomalous waveforms are generated and penalize them accordingly. Qualitative and quantitative evaluations of the generated waveforms were conducted to confirm their effectiveness.

様式 7

論文審査の結果の要旨及び担当者

氏　名　（　吉村　直也　）		
	（職）	氏　名
論文審査担当者	主 査　准教授	前川　卓也
	副 査　教授	原　隆浩
	副 査　教授	藤原　融
	副 査　教授	松下　康之
	副 査　教授	下條　真司
	副 査　教授	鬼塚　真

論文審査の結果の要旨

提出された論文では、工場やロジスティクスセンタ等の作業員が装着するセンサデバイスから得られたセンサデータを用いた行動認識手法および、行動認識モデルの可視化手法の提案について述べている。作業員による手作業が介在するプロセスは機械のそれに比べて画一的ではなく、効率性向上の余地は大きい。効率性向上のためには、作業ラインの効率化やボトルネックの検出が重要と言える。本論文では、作業ラインの効率化やボトルネック検出に関連の深い作業行動の認識および、作業行動認識モデルを現場に導入するための可視化技術に関して、以下の研究成果について述べている。

１つ目の主な成果として、作業員のウェアラブルデバイスにより得られた加速度センサデータから、作業員の作業行動を教師あり学習により認識する手法を提案している。作業員の作業内容は作業員ごとに異なることが多く、工場などの環境にて教師あり学習のための学習データを大量に用意することは困難である。そのため、少量の学習データで訓練可能な深層学習モデルを提案している。提案モデルは、複雑な作業の長・短期の時間的依存関係を少量のパラメータで捉えるモジュールや作業の境界を検出するモジュールを備えることを特徴とする。実際の工場等で取得されたデータを用いて評価実験を行い、高い精度での認識を達成している。２つ目の主な成果として、ロジスティクス分野における梱包作業の大規模加速度データセットを構築した。産業分野における大規模行動データセットはこれまでに無く、当該分野の研究活動を大きく促進させることが期待される。また、データセットに基づき、加速度データと作業の際に作業員が操作したIoTデバイスのログを効果的に融合する作業行動認識手法を提案している。３つ目の主な成果として、深層学習に基づく行動認識モデルの現場導入を見据えた、行動認識モデルの処理プロセスの可視化技術を提案している。Activation maximization(AM)と呼ばれる手法を用いて、認識モデルのノードが反応する加速度データの波形を可視化した。このとき、ノードのActivation値の正常と考えられる値域を利用した正則化項を提案し、現実の加速度データと近い波形を生成した。

本論文は、作業員を組み込んだ知的な作業プロセスを実現するための基礎技術を実現した。工場作業のような複雑行動を少量の学習データで認識した研究はこれまでになく、既存手法に対して大幅な精度向上を達成している。また、複雑な作業行動に関する大規模データセットを初めて構築し、当該分野の研究コミュニティに大きく貢献した。深層学習に基づく行動認識モデル可視化手法は、ICTに関する知識が限られた管理者に行動認識手法を円滑に展開させる際の重要技術となる可能性がある。本論文により達成された成果は、人間による作業が組み込まれた作業プロセスの効率性を向上させる先端的な情報技術手法の研究として、情報科学に寄与するところが大きい。よって、本論文は博士（情報科学）の学位論文として価値のあるものと認める。