Cyber-Physical Sensing, Modeling, and Control with Augmented Reality for Smart Manufacturing Workforce Training and Operations Management

Students: Md Moniruzzaman[1], Wenjin Tao[2], Ze-Hao Lai[2] , Md. Al-Amin[3], Hao Sun[4]
Investigators : Zhaozheng Yin[1], Ming C. Leu[2], Ruwen Qin[3], Zhihai He[4]
[1]Department of Computer Science, [2]Department of Mechanical & Aerospace Engineering, [3]Department of Engineering Management & Systems Engineering, Missouri University of Science and Technology; [4]Department of Electrical and Computer Engineering, University of Missouri

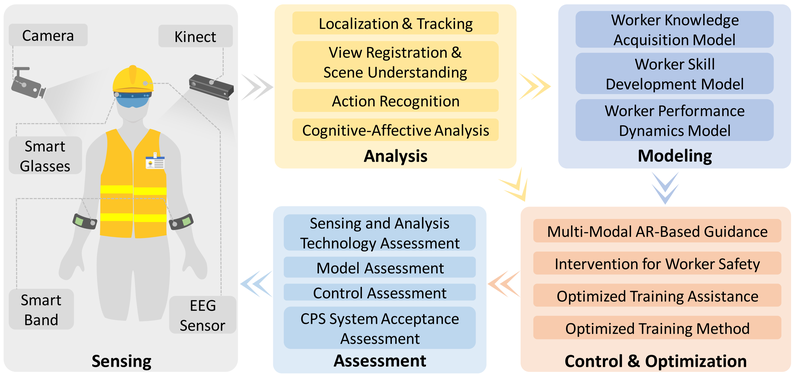
This research is supported by National Science Foundation grant CMMI-1646162.

Objectives

While the U.S. manufacturers are investing tremendous efforts and resources to regain the power and growth of manufacturing, especially in the smart manufacturing, they are confronted by a set of critical and challenging issues:

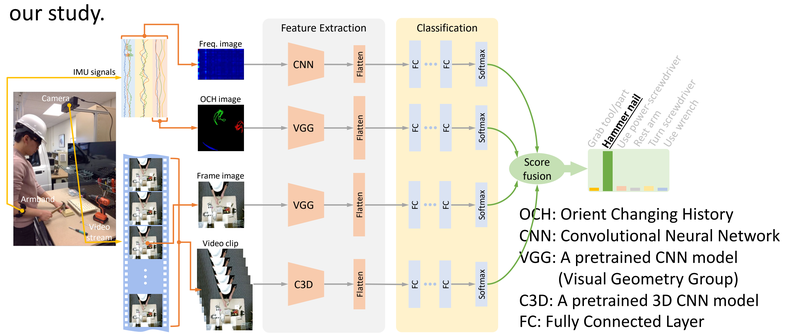
- Lack of workforce with advanced training and skills;
- Need for rapid and individualized training to achieve workforce agility;
- Need for on-the-job personal assistance to improve worker performance and safety.

This project aims to develop an integrated set of cyber-physical methods and tools to sense, understand, characterize, model, and optimize the learning and operations of manufacturing workers, so as to achieve significantly improved efficiency of worker training, effectiveness of behavioral operations management, and safety of front-line workers, for smart manufacturing.



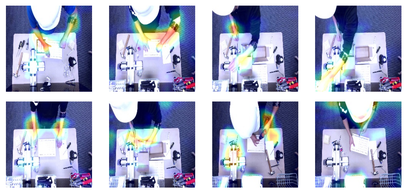
Multi-modal Sensing & Recognition of Worker Activity

Recognition of the worker's activity can be used for quantification and evaluation of the worker's performance, as well as to provide onsite instructions with augmented reality. A multi-modal approach is proposed in our study.



OCH: Orient Changing History
CNN: Convolutional Neural Network
VGG: A pretrained CNN model (Visual Geometry Group)
C3D: A pretrained 3D CNN model
FC: Fully Connected Layer

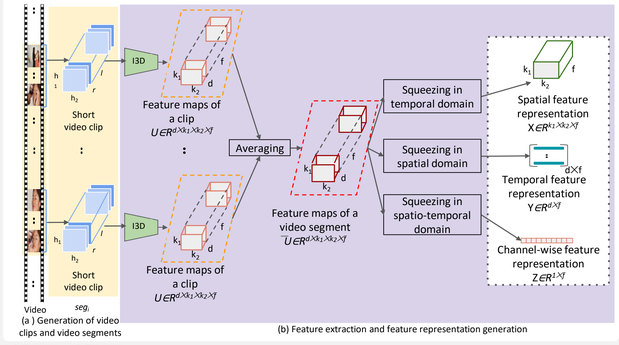
The Class Activation Maps show that the trained model is able to "pay attention" to the areas where hand-tool-part interactions happen.

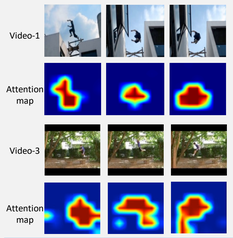


Our multi-modal approach shows superior performance on worker activity recognition over other models.

Action Recognition by Discriminative Feature Pooling (DFP) and Video Segment Attention Model (VSAM)

We introduce a simple yet effective network for human action recognition and address the three main issues: (1) from the convolutional feature maps of a 3D CNN applied on short video clips, which spatio-temporal and channel-wise features should get more attention to highlight the discriminative features related to the action class? (2) which video segment should get more attention to represent an action? and (3) how to train the network from weakly labeled Videos?

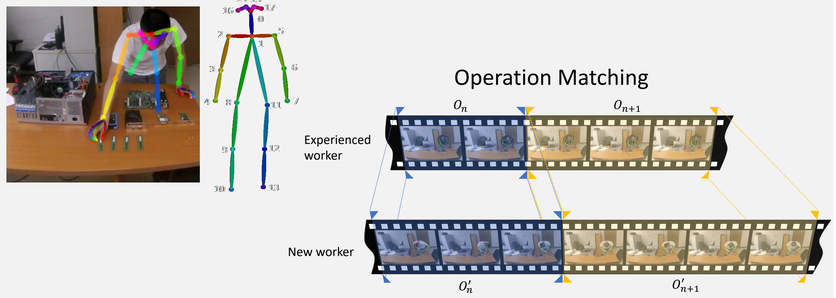




The proposed network addresses the three main issues by using training data obtained from both trimmed and untrimmed videos. Evaluated on three challenging public datasets, our network that integrates the Discriminative Feature Pooling and the Video Segment Attention Model outperforms the current state-of-the-art methods. The proposed network is also efficient and easy to implement.

Evaluation of Workforce Efficiency Based on Human Pose Features

A detection algorithm has been developed to detect the whole operation sequence of a worker who performs a manufacturing task. From the recorded videos, the start point and the end point can be detected and the time spent on the task can be evaluated.



To evaluate the details of each operation, each worker's video (student video) is matched with the fully labelled reference video (teacher video). The start and end points of each operation can be detected. Then the efficiency of each operation can be evaluated.

- Individual workers' learning behavior and operating ability can be analyzed by a cyber-physical system.
- Performance assessment for each operation among the whole manufacturing task can also be achieved.

Action Recognition in Manufacturing Assembly Using Wearable Sensor

In fast changing production systems, manufacturers with the ability to comprehend workers' behavior and assess their operation performance will be far ahead than peers. Action Recognition serves the purpose.

Study Design

- Myo-armband, a wearable sensor, collects acceleration signals
- A CNN architecture works as classifier

Focuses of the study

- Effectiveness of model fusion
- Impact of model calibration and training sample size on model performance in LOO testing

Effectiveness of average fusion

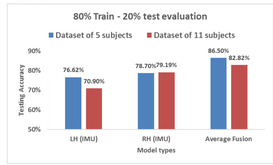


Figure-1(a)

Effectiveness of model calibration

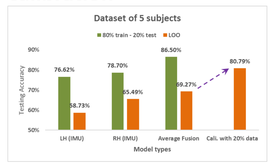


Figure-2(a)

Impact of sample size at the subject level

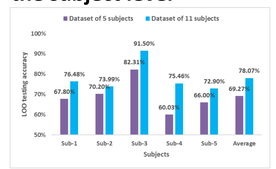


Figure-3

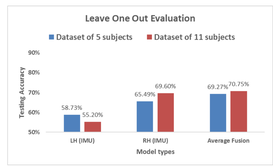


Figure-1(b)

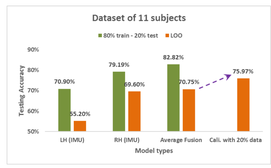
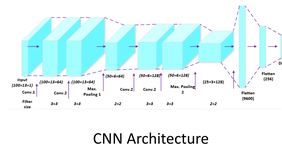


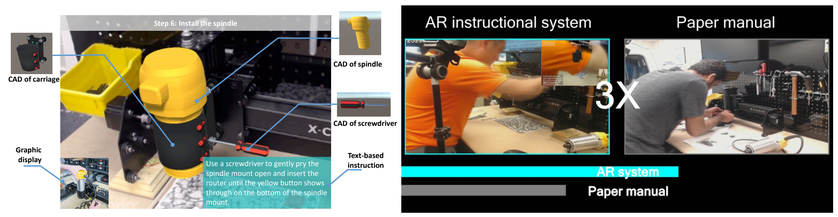
Figure-2(b)



CNN Architecture

Smart Augmented Reality (AR) Instructional System

A smart augmented reality instructional system for mechanical assembly is being developed, which provides in-situ timely instructions with multi-modal augmented reality. The assembly process can be instructed with AR technology in a more immersive and interactive manner.



The Smart AR system can provide assembly instructions through on-site displays with multi-modal AR rendering, which has been demonstrated to reduce 33.2% of task completion time and 32.4% of assembly errors for a spindle assembly task, in comparison to paper manual based instructions.

Summary

A cyber-physical system (CPS) has been developed to:

- Sense and recognize what individual workers are performing in the workplace using deep learning methods;
- Evaluate quantitatively how well individual workers are learning during training or performing tasks in operations;
- Provide multi-modal in-situ task instructions with Augmented Reality (AR).