Detection of Workers' Behaviour in the Manufacturing Plant using Deep Learning

Goh Ching Pang[1*]

[1]Faculty of Computing and Information Technology, Tunku Abdul Rahman University of Management and Technology

***Email:** gohcp@tarc.edu.my

***Received:** 30 November 2023; **Accepted:** 14 December 2023; **Published:** 15 December 2023*

Abstract: In the modern manufacturing landscape, optimizing productivity is a paramount challenge, particularly in dynamic, non-concentrative environments where human activities are diverse and complex. Accurately monitoring and analyzing worker behavior is crucial for enhancing manufacturing processes, but traditional methods fall short in these settings due to their reliance on simplistic global image features and manual classification. Addressing this gap, this paper introduces a groundbreaking vision-based capture technology, integrated into a manufacturing monitoring system. This technology significantly advances productivity by providing a nuanced assessment of worker behavior. It departs from conventional approaches by employing gait recognition techniques, which effectively match input sequences with predefined models. This method adeptly navigates the hurdles of data scarcity, diverse human behaviors, and visual variations typical in manufacturing environments. Utilizing machine learning algorithms, our system learns and detects intricate activities from worker behavior sequences, offering a sophisticated analysis of worker efficiency. The primary aim is to quantify human behavior based on learning rates, thereby facilitating improved production control. Our findings are promising, demonstrating an impressive 99% accuracy in behavior detection. This high level of precision underscores the potential of our technology to transform manufacturing productivity and worker monitoring practices.

Keywords: Deep Learning; TensorFlow; Worker's Behavior; Manufacturing

1. Introduction

An essential approach to attaining success in pressure manufacturing projects is to closely observe and oversee the activities of industry workers. Given that manufacturing projects are characterised by a high demand for manual labour and a reliance on manual tasks, it is crucial to monitor and comprehend the dynamic activities and movements of workers in order to efficiently manage them and enhance productivity. Understanding worker activity and mobility is crucial in environments involving in continuous and repetitive physical activities, like as welding, as it is essential for ensuring and enhancing worker productivity. Labour efficiency has a direct impact on construction productivity in the pressure vessel manufacturing industry, which is inherently reliant on manual labour (Johny & Thenarasu, 2019; Jaysval & Vyas, 2019). Significant labour productivity growth is observed in various sectors, including multi-family housing and industrial production, within the given context.

In 2012, the sections analyzed in the U.S. accounted for less than 10% of the total construction hours (Haugbølle et al., 2019). To improve labor productivity in construction projects, it is essential to evaluate labor productivity at the individual level. However, current methods rely on manual techniques such as direct observation sampling. Before measuring productivity at the individual level, it is necessary to monitor the activities of individual workers. Since manufacturing processes typically involve repetitive physical actions, discerning workers' movements can provide valuable insights into their activities. This study introduces a novel hardware/software framework designed to support real-time operators engaged in manual assembly tasks, thereby enhancing their daily performance in terms of both productivity and quality. A markerless depth camera records human movements within the workstation environment, while an augmented reality application offers visual feedback to assist operators during the training phase and assess their actions.

Worker behaviour recognition can link rapidly to performance. When employees have outperformed in his or her duties, they can highly praise employees for their efforts. This helps them establish an emotional link to the workplace and perform greater in the future performance. The worker wants to know the effort on their hard work and achievements have been appreciated by the management. When a person meets personal goals or job-related objectives, they feel a strong sense of accomplishment. This good feeling will only be spotted when other people also acknowledge and recognize their achievement. In the other side, the management want to surveillance the worker behaviour in order to qualitative of the performance of the worker. With a measured data of real recognition on a regular basis, the management will be more capable of releasing their full potential, such as multitasking skills and increase their productivity.

2. Problem Statement

Currently, there is a reliance on manual approaches, such as visually sampling the process, as indicated by Lee (2023) and Khan (2017). These approaches are both labor-intensive and prone to errors, making it difficult to individually estimate the productivity of several workers using these methods. Therefore, motion capture technology will be employed to replace manual visualisation by digitising the capturing process. Furthermore, the operator will not receive any feedback, resulting in doubt on the accuracy of the procedure (Kim & Cho, 2020; Makihara et al., 2014). This system exhibits suboptimal efficiency and further analysis might be conducted to enhance process efficiency in the future. One significant drawback of this technique was that it made workers unable to achieve a complete feedback loop for assisted assembly, both throughout the learning and processing stages of industrial production processes (Yamashita, 2018; Yan et al., 2018). Therefore, the process of maximising the production can be determined for this specific objective.

3. Methodology

In this project, the framework to create the model of worker behaviour were coded and trained using TensorFlow and Google's Deep Learning System. The data of worker behaviour is captured at Tunas Asal Sdn Bhd and trained the model running on top of TensorFlow as shown in Figure 1. Everything was performed on a local machine in Google Collab. The worker behaviour model is trained in Inception-v3 to recognise the models with stochastic gradient utilizing the TensorFlow distributed machine learning system on a Nvidia Tesla K80 24GB GDDR5 CUDA Cores Graphic Cards. The dataset is divided with ratio of 35 and 65, 35 used to train the model and 65 to test the results. The overview of the process flowchart is shown in Figure 2. The flowchart outlines a machine learning workflow that employs a Convolutional Neural Network (CNN) to classify various human activities. The process begins with the collection of data on activities such as welding, walking, standing still, and crouching. This data is then subjected to preprocessing to ensure it is in the correct format and of the quality required for effective model training. After preprocessing, the data is used to train a CNN, a type of deep learning model renowned for its efficacy in interpreting grid-like data, including images. The performance of the CNN is assessed by its classification accuracy—if the model achieves an accuracy of 80% or higher, it moves on to the next phase; otherwise, the process may need to revisit earlier steps for further refinement. Upon reaching or exceeding the threshold accuracy, the model's results are thoroughly evaluated to ensure they meet the project's standards. The process concludes when the model's performance is satisfactory.

Figure 1: Partial captured dataset for model training

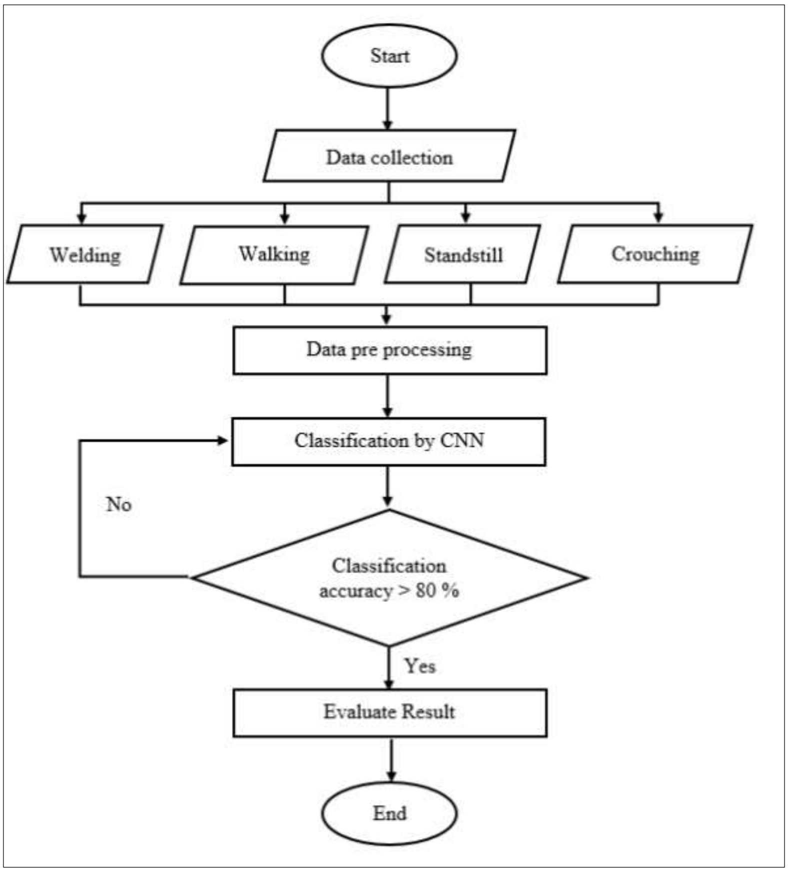


Figure 2: Overview of process flow chart

4. Discussion

The objective of this study is to validate the accuracy of operator motion parameter during accomplished action by analysing and optimizing the worker behaviour in processing stages of industrial manufacturing process. This study reported on applying CNN to a 4-class worker behaviour of data. Table 1 shows the parameters used in the machine learning model. With a total of 22,007,588 parameters, the model's complexity is outlined by the number of elements that it uses to make predictions or decisions. Out of these, only 204,804 parameters are trainable, meaning they can be modified during the training process as the model learns from the data. The vast majority of parameters, amounting to 21,802,784, are non-trainable, indicating that they remain constant during training. This could imply that the model incorporates pre-trained components or uses fixed parameters for certain functions. Additionally, the model includes 4 distinct classifiers or modules for categorizing data, which suggests that it might be used for tasks involving multiple classifications or predictions. The exact nature of these classifiers and their roles within the model would be clearer with additional context regarding the model's specific application and architecture.

Figure 3 shows the accuracy and loss of a machine learning model at two different stages of its training, specifically at 50 epochs and 100 epochs. An epoch in machine learning is a complete pass through the entire training dataset. The upper half of the chart shows the model's loss during validation and training phases. Loss is a measure of how well the model is performing, with lower values indicating better performance. At 50 epochs, the validation loss is approximately 36.43%, and the training loss is about 15.88%. At 100 epochs, both losses have decreased, with the validation loss at around 40.33% and training loss at approximately

18.28%. It's unusual to see the validation loss higher than the training loss; typically, we would expect the training loss to be lower as the model becomes more specialized to the training data. The lower half of the chart details the accuracy of the model, which is the proportion of correct predictions it makes. The validation accuracy at 50 epochs is nearly 99.16%, and the training accuracy is similarly high at 99.34%. At 100 epochs, both accuracies have slightly increased, with validation accuracy at about 99.16% and training accuracy at approximately 99.37%. The results indicate that the model's accuracy is very high and remains stable or slightly improves over time. However, the increase in validation loss with more epochs could be a sign of overfitting, where the model learns the training data too well, including its noise and outliers, which may not generalize well to unseen data. It's important to monitor both accuracy and loss to ensure that the model is improving its performance in a balanced manner without overfitting.

Table 1: The summary of parameter in evaluating the model structure

Description	Value
Total parameters	22,007,588
Trainable parameters	204,804
Non trainable parameters	21,802,784
Number model of classify	4

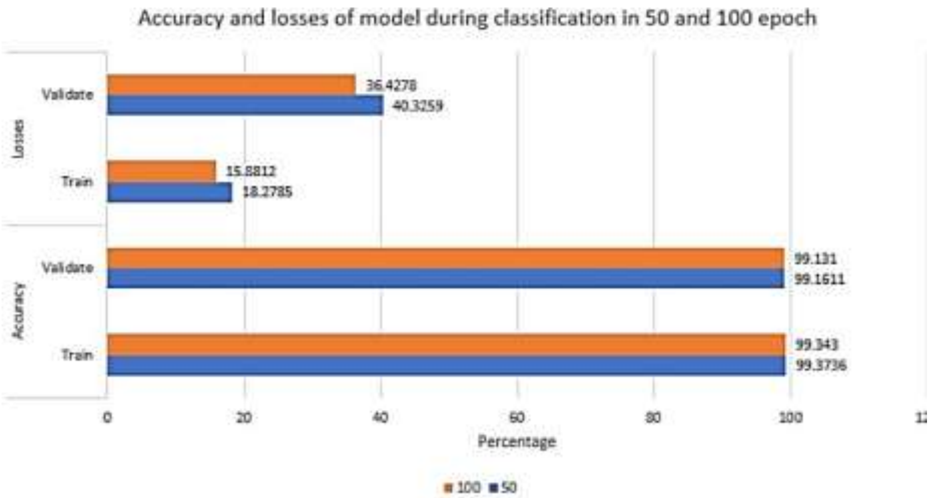
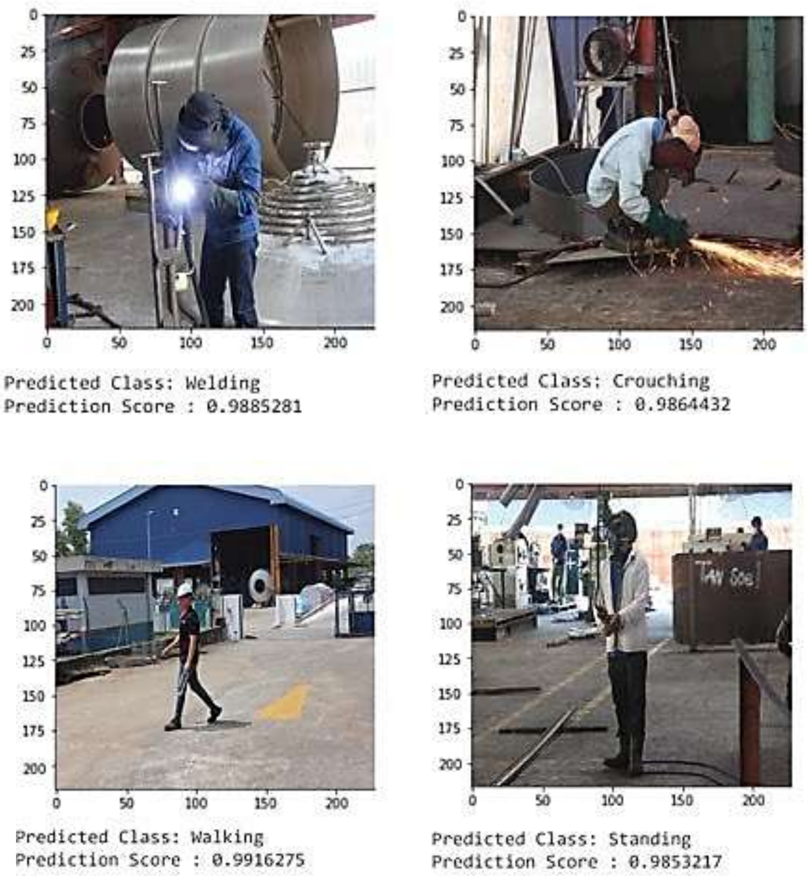


Figure 3: Accuracy and losses of model during classification in 50 and 100 epochs

The worker behavior model achieved the best test accuracy of 99% (Figure 4), surpassing the expected accuracy of 78% achieved by the Inception V3 approach. However, the model suffered from noticeable overfitting due to high losses during training and testing. This issue was addressed by extending the training epochs, which reduced the losses to 4%. Nevertheless, training the model with different epochs revealed limited potential for improving test accuracy. The expected outcome of this study was to achieve an 80% accuracy in the validation of the worker behavior classification model. The current optimized model exceeded this goal, achieving an impressive 99% accuracy. As a result, the objective of this study has been successfully accomplished.

Figure 4: Predictions obtained for images from test data



5. Conclusions

This research presents a framework for modeling and analyzing worker behavior in industrial manufacturing processes using advanced machine learning techniques. TensorFlow and Google's Deep Learning System were employed to code and train the behavior model, which was executed on a local machine in Google Colab. The model utilized Inception-v3 and was trained to recognize various worker actions with the aid of stochastic gradient descent on Nvidia Tesla K80 24GB GDDR5 CUDA Cores Graphic Cards. The dataset was divided into a 35% training and a 65% testing split. The workflow involved data collection, preprocessing, and training using a Convolutional Neural Network (CNN). The CNN's performance was evaluated based on classification accuracy, aiming for 80% or higher. Upon reaching this threshold, the model's results were meticulously assessed to meet project standards. The study's objective was to validate operator motion parameters during industrial manufacturing processes. The CNN was applied to classify four distinct worker behaviors. The model boasted a total of 22,007,588 parameters, with 204,804 being trainable. Notably, the model's accuracy exceeded expectations, achieving a remarkable 99% accuracy rate, although overfitting was initially observed. Extending the training epochs successfully mitigated overfitting, and the model's performance surpassed the study's original goal of 80% accuracy. All in all, this research demonstrates the successful implementation of a CNN-based model to classify and optimize worker behavior in industrial manufacturing processes, achieving a high level of accuracy and addressing initial overfitting concerns.

References

Haugbølle, K., Larsen, J. N., & Nielsen, J. (2019). Construction productivity revisited: towards measuring performance of construction output. Engineering, Construction and Architectural Management, 26(5), 794-813. https://doi.org/10.1108/ECAM-03-2018-0094

Jaysval, J., Goyal, B., & Vyas, T. (2019). Review of productivity improvement in pressure vessel manufacturing industry. Journal of Emerging Technologies and Innovative Research, 6(5), 2349-5162.

Johny, J., & Thenarasu, M. (2019). Productivity enhancement in a pressure vessel manufacturing industry using lean principles. International Journal of Innovative Technology and Exploring Engineering (IJITEE), 8(8), 3272–3279.

Khan, S., Hussain, M., Aboalsamh, H., & Bebis, G. (2017). A comparison of different Gabor feature extraction approaches for mass classification in mammography. Multimedia Tools and Applications, 76, 33-57.

Kim, K., & Cho, Y. K. (2020). Effective inertial sensor quantity and locations on a body for deep learning-based worker's motion recognition. Automation in Construction, 113, 103126. https://doi.org/10.1016/j.autcon.2020.103126

Lee, T. Y., Ahmad, F., & Sarijari, M. A. (2023). Activity sampling in the construction industry: a review and research agenda. International Journal of Productivity and Performance Management. https://doi.org/10.1108/ijppm-10-2022-0507

Makihara, Y., Mansur, A., Muramatsu, D., Uddin, Z., & Yagi, Y. (2014). Multi-view discriminant analysis with tensor representation and its application to cross-view gait recognition. In Proceedings of the 2015 11th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition (FG) (pp. 1-8).

Yamashita, R., Nishio, M., Do, R. K. G., & Togashi, K. (2018). Convolutional neural networks: an overview and application in radiology. Insights into Imaging, 9, 611-629.

Yan, C., Zhang, B., & Coenen, F. (2018). Multi-attributes gait identification by convolutional neural networks. In Proceedings of the 2015 8th International Congress on Image and Signal Processing (CISP) (pp. 642-647).