INTELLIGENT VIDEO ANALYTICS FOR IMPROVING MAN-TO-MACHINE-RATIO (MMR) IN AN ASSEMBLY LINE

Sabari Vishnu Jayanthan Jaikrishnan[1]
Janoah De Los Santos[2]

Operations Excellence, Data Analytics Engineering
Western Digital Corporation, 109 Technology Ave., SEPZ, Laguna Technopark, Binan, Laguna, Philippines.
[1]sabari.vishnu.jayanthan.jaikrishnan@wdc.com,
[2]janoah.de.los.santos@wdc.com

ABSTRACT

Optimizing the Man-to-Machine Ratio (MMR) has emerged as a significant challenge for manufacturing enterprises. MMR serves as a Key Performance Indicator (KPI) directly linked to operational efficiency. A lower MMR signifies a more efficient production process. This paper presents a novel approach to MMR improvement by leveraging cutting-edge technologies such as Machine Learning (ML), Deep Learning (DL), and Video Analytics. These technologies will be employed to automate the real-time quantification of components within testing machine queues and the on-site workforce, encompassing both operators and technicians. By acquiring this critical data, we aim to identify areas with elevated MMR within the manufacturing process. Subsequently, internal interventions can be implemented to achieve a more balanced MMR across the production line.

1.0 INTRODUCTION

The Man-to-Machine Ratio (MMR) is calculated by dividing the total workforce in a factory by the total number of operational machines in the production line. Monitoring and improving MMR is a common industrial practice to optimize the productivity of employees and to make sure the efficiency of all machines is high. For example, an MMR of 1:5 denotes that 1 operator can handle 5 machines at any given point in time. There is no optimum MMR for any industry, it is decided based on a company's requirements and budget.

Every organization experiences limited resources and under-harnessing the available resources leads to poor throughput and efficiency [1]. Many companies are trying to improve their MMR in their ways. Apart from improving efficiency, improving MMR also increases the morale of the workers by keeping them occupied and reduces the time wasted, if any.
In this research, we propose a novel method to handle MMR using Deep learning and live-streaming Video Analytics. By making use of the existing CCTV cameras in our clean rooms, we try to count the number of components that are queued to get tested in our testing machines and count the operators and technicians available to test machines. We use this data obtained from the cameras to assess our MMR.

1.1 Video Analytics

Video Analytics is a technology that uses Artificial Intelligence, Deep Learning, and Computer Vision to process video stream data and extract meaningful information from the processed data. Video Analytics can be implemented in many areas such as enhanced security, monitoring traffic, traffic violation detection, suspicious activity detection, and crowd management. These applications are elaborately reviewed in this paper [2].

A typical video analytic pipeline starts by receiving input video streams in the form of individual frames. Each frame would be received in raw format and would be converted into a machine learning algorithm readable format. Each frame would undergo a sequence of preprocessing steps such as reducing the frame size and converting RGB to BGR. Once preprocessing is done, the frames would be fed to a deep learning model that performs necessary operations such as classification on the frames. Depending on the application and the business requirements, the video analytic pipeline can trigger an alarm and report for any real-time critical events. Furthermore, it can store the data in a database for future analysis and show the output in a real-time live stream.

1.2 Deep Learning

Deep learning, a subfield of Machine Learning (ML) and a branch of Artificial Intelligence (AI), is primarily concerned with the development and implementation of multi-layered Artificial Neural Networks. These networks are composed of multiple layers, each containing numerous interconnected nodes, which are stacked on top of one another. This architecture is designed to mimic the structure and function of the human brain, in which neurons are connected and transmit electrical signals. As data passes through the network, each layer performs specific operations that transform the data, allowing the network to learn and make predictions. Deep learning has found widespread application in a variety of fields, including computer vision, natural language processing, speech recognition, and recommender systems.

This paper is organized as follows: First, the current practice of how organizations are handling MMR and how they are making use of video analytics in their processes has been reviewed. In the next section, the authors introduce a novel pipeline that, using Video Analytics, counts the number of components on the rack before it reaches the testing machines and counts the number of operators and technicians at any given point in time. This is followed by the author's future work recommendations and conclusion.

2.0 REVIEW OF RELATED WORK

Many organizations are focusing on reducing their MMR and much research has been done having this specific problem in mind. In [3], the authors established a rating scale for analyzing the MMR in the Garment industry. They have also created a grading scale to compare all the other companies in the industry and improve their MMR. Though their grading system might be efficient for their company, it is limited to a specific product, and it cannot be easily applied to other companies from other industries.

Simulation can also be used to optimize MMR. In [4], the researchers proved an efficient way of improving MMR by building ratio models using a simulation software called "ProModel". These ratio models use the company's entities that flow through the manufacturing process. While the results of their study facilitate the decision to move headcount from one location to another, running the simulation ratio model in real-time on a large scale is expensive and resource-consuming.

In [5], the researchers proposed an approach that deals with the evaluation problems of the one-person-multi-machine concept in lean methodology. They used this discrete model approach to assign the number of machines to each operator while they were working on the production line, thereby optimizing the MMR. However, the actual analytical method yielded a better result compared to their proposed discrete event simulation model.

There is plenty of literature available in the booming field of video analytics. In [6], the researchers proposed a multi-task classification model that uses a video stream from the main production assembly line to calculate the cycle time and get insights into the variations in the cycle time. However, they used Amazon Web Services (AWS) cloud for data storage and tried to run analytics in EDGE using AWS Greengrass Core. Though this method uses AWS Greengrass Core for EDGE implementations, we need to replace it with real EDGE technology for quicker analysis. In addition to that, the pipeline struggles with human detection which can be fixed by adding more training data and training the model.

The problem specified above was tackled by [7], in which the researchers had taken an edge-centric architecture for constructing real-time smart assistant video analytics services. They achieved this by implementing several pilot services to show the opportunities of industrial video analytics. They prove that edge-centric video analytics has a very high potential to improve smart assistance in Industrial Internet of Things systems.

Though there is much literature on improving MMR and Video analytics, there is no research or experiment that combines both. In this paper, the researchers propose a novel pipeline using Convolutional Neural Networks (CNN) for classifying different types of statuses of the components that are in queue to be tested from the machines and Single Shot Detectors (SSD) to detect the operators and technicians in the clean rooms. Having this information in real-time, the workload can be managed, and MMR can be improved spontaneously in clean rooms. A complete end-to-end edge-level solution is proposed in which an edge hardware called Intel's Neural Compute Stick is employed.

3.0 METHODOLOGY

Our proposed approach has the following steps:

Get the live feed using a camera: The video stream is captured by the CCTV cameras that are available in the clean rooms. The cameras follow RTSP (Real Time Streaming Protocol), and the stream can be H264, which requires low bandwidth but larger memory, or H265 which requires high bandwidth but is more efficient than H264.

Preprocessing the image frames: Intel's Open Vino (2021) [8] framework has been employed. It uses GStreamer, CVAT, and Python programming language to preprocess the incoming video frames into model-readable format. In this case, TensorFlow readable format.

Classify the presence of components: Applying a custom Convolutional Neural Network (CNN) trained with over 80,000 images for 4 different classes. The model is developed using TensorFlow 2.0 [9]. The images are collected at various conditions such as different lighting conditions, and different times of the day, including reflective materials around the machines and different angles of machines. The annotations are done in an open-source application Computer Vision Annotation Tool (CVAT).

Detecting the operators & technicians: The frames are then fed to a Single Shot Detection model (SSD) that detects the operator and technicians in the clean room. This model is trained by transfer learning from the Open Vino Model Zoo [10]. This is a resource-consuming process, so we invested in Intel's Neural Compute Stick chips. These are USB-like chips that act as portable handy GPUs for any device where the pipeline is hosted.

Publish the results: IoT devices usually talk with each other via MQTT (Message Queuing Telemetry Transport). The pipeline publishes the results dynamically that can be subscribed to the database. In the future, we can even trigger an alert or report in real-time by subscribing to this publisher. This allows pub-sub communication between IoT devices.

Live feed output: Bounding boxes are applied to the output frames and using HLS (HTTP Live Stream), a protocol developed by Apple for ultra-low latency live stream to show the output. Nginx is used for load-balancing and rerouting the stream.

Analyze the results in a dashboard: Using TIBCO Spotfire, a data visualization tool, the output captured by the real-time video analytics models is visualized for more historical insights. This opens a door for future developments in terms of maintenance and improving MMR. Finally, this dashboard will be shown in our in-built application for data visualization.

Abnormality triggering: Using the TIBCO Spotfire dashboard, the pipeline will also trigger an alert when an abnormal event happens. The logic for considering an abnormal event hasn't been perfected yet but the researchers use the 3-sigma rule, i.e., anything that crosses the 3-sigma in a stream data is considered as a critical problem.

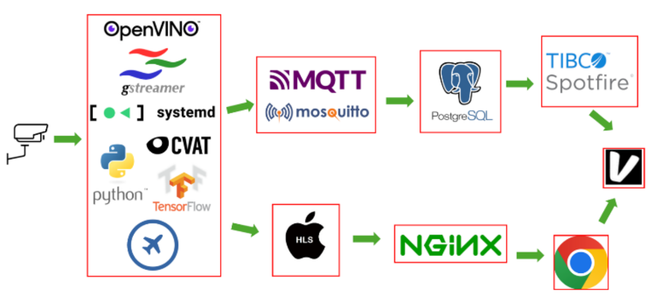


Figure 1. The proposed video analytics pipeline.

4.0 RESULTS AND DISCUSSION

As a result, we were able to track the available components and track the headcount of the operators and technicians at any given point in time in the clean room. With this data, we identified the weak spots and congestion of headcount and improved the MMR significantly.

With this approach, we were able to achieve a 60% improvement in the MMR and machine utilization from 76% to 86%.

Table 1. Accuracy of the models used in the pipeline.

Model	Unit	Accuracy
Convolutional Neural Network (CNN)	%	99.5
Single Shot Detection	%	90

The CNN model is trained with 80,000 images belonging to 4 different classes that detect the active state of the components. The single shot detection model is trained with 2 different classes that detect humans.

5.0 CONCLUSION

In conclusion, a novel method of improving MMR by using CCTV live streams has been proposed. The latest technologies such as Deep learning, machine learning, and video analytics are used to count the number of components that are queued in line to be tested by a tester and count the number of operators and technicians in the clean room area. This measure, coupled with a dashboard that shows the numerical data in the form of easily readable charts, aids in the reduction of MMR and moving the workforce from one place to another in real-time and dynamically.

6.0 RECOMMENDATIONS

The pipeline is scalable and good in terms of the models' accuracies; however, the pipeline relies on Intel's Open VINO. Though it is an open-source application, relying on one software may be difficult to upgrade and maintain. Furthermore, the pipeline requires Intel's Neural Stick to run the object detection model. The operating system must be Ubuntu operating system as the application Open VINO works the best in Ubuntu. The researchers are now working on overcoming the shortcomings by creating a cross-functional, cross-platform, scalable, and affordable video analytics pipeline.

The accuracy of the Single Shot Detection model can be improved by increasing the number of good samples and epochs.

7.0 ACKNOWLEDGMENT

We would like to express our sincere gratitude to Sean Leary, for his invaluable contribution during this framework formation, providing critical invaluable feedback on the methodology, sharing essential data, and offering helpful discussions throughout the research process.

8.0 REFERENCES

[1] C.-H. Cheng and I. E. Olatunji, "Harnessing constrained resources in service industry via video analytics."

[2] Q. Zhang, H. Sun, X. Wu, and H. Zhong, "Edge video analytics for public safety: A review," *Proceedings of the IEEE*, vol. 107, no. 8, pp. 1675–1696, Aug. 2019, doi: 10.1109/JPROC.2019.2925910.

[3] M. Alubel Abtew, T. Berhane, A. Mehtre, and M. Alubel, "Establishing a rating scale for knitted garment industry based on man machine ratio for Ethiopia," 2016. [Online]. Available: https://www.researchgate.net/publication/304942541

[4] Hoay Hoon Ong, "Establishing man-machine ratio using simulation," in *Proceedings of the 2007 Winter Simulation Conference*, Proceedings of the 2007 Winter Simulation Conference, 2007.

[5] Y. Zhou, F. Kong, and T. Chen, "Influence of man-machine ratio on system performance of one-person-multi-machine series production line," 2021.

[6] A. Walker, D. Maeda, and J. Acharya, "Lightweight video analytics for cycle time detection in manufacturing," in *Proceedings - 2021 IEEE International Conference on Big Data, Big Data 2021*, Institute of Electrical and Electronics Engineers Inc., 2021, pp. 3615–3618. doi: 10.1109/BigData52589.2021.9671658.

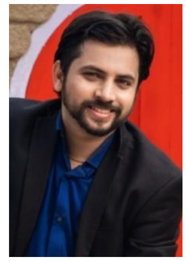
[7] N. A. Bazhenov, A. E. Harkovchuk, and D. G. Korzun, *Edge-Centric Video Data Analytics for Smart Assistance Services in Industrial Systems*.

[8] Intel, "OpenVINO," 2024.

[9] Google, "TensorFlow By Google," 2024.

[10] Intel, "Open Model Zoo," 2023.

9.0 ABOUT THE AUTHOR(S)



Sabari Vishnu Jayanthan Jaikrishnan is a Data Scientist at Western Digital Storage Technologies (Thailand) Limited specializing in Machine Learning, Deep Learning, and Video Analytics projects. He holds a master's degree in Information & Data Science where he researched Automated Machine Learning and a bachelor's degree in Electronics & Communication Engineering where he worked on Digital Signal Processing and Emergency Telemedical Services. His research interests include Generative AI, Machine-to-Machine communication, IIoT, etc.

Janoah De Los Santos is the Director of the Data Analytics Engineering Department at Western Digital Storage Technologies Limited, Philippines. She has more than 10 years of technical experience in the field of Statistics, Business Analytics, and Operations Management. Some of her current interests are Generative AI technology, Data Analytics & Harmonization, and mentoring new Data Scientists.