# Assembly Line Quality Assurance Through Hand Tracking and Object Detection

*Mamuku* Mokuwe[1*], *Yurisha* Goorun[1] and *Gerrie* Crafford[1]

[1]Council for Scientific and Industrial Research, South Africa

**Abstract.** This work presents a vision-based quality assurance system that does assembly line monitoring. The system is developed using machine learning hand tracking and object detection methods to monitor the worker's hand movement while evaluating the correctness of the assembly. Feedback about the order of the steps the worker has taken is continuously shown to the user. This work has the potential to reduce the amount of manual work required for quality assurance in assembly line.

## 1 Introduction

For many years in industrial manufacturing, quality assurance (QA) has relied on human inspectors to spot errors as parts move along the production assembly line. Considering the repetitive nature of assembly operations, human quality assurance inspection can be time consuming and often does not yield the best results, which might result in incomplete products reaching customers. To offer an alternative solution for such problems, this work proposes the use of machine learning (ML) technologies to develop a vision-based QA system that will monitor the hands of the worker and ensure that the correct parts placement and order of assembly is followed.

There are numerous applications of ML in manufacturing such as predictive maintenance, quality assurance and intelligent automation [1, 2]. The key contribution in this paper integrates multiple ML methods to perform hand-tracking and object detection to ensure quality during assembly. The hand tracking vision-based method is used to track the worker's hand movement to ensure a particular order during the assembly is followed. The detection method is used to ensure that the parts placement on the assembly is done correctly. If any inconsistencies are identified, the system alerts the worker about their mistakes to allow them to correct their movement. The QA system using ML methods offers a significant improvement in digital technology.

## 2 Vision based assembly monitoring systems

There exist many studies today that address the challenges and need of quality assurance systems on the production assembly line. The deployment of cameras to capture the workflow

* Corresponding author: windyseipati@hotmail.com

on the assembly line and the use of computer vision, machine learning and deep learning methods to provide insights is one of the common solutions that exits [3]. Some of the reasons for using vision-based methods is the ability to monitor the assembly in real-time through live image capturing thus reducing the burden of physical human inspection.

One example application of the assembly monitoring is through hand tracking to recognise the motion of the worker. The tracking of the worker's hand during assembly process can be used to ensure that the correct order is followed during assembly and that the parts are placed in the correct position. Some work on the hand tracking is based on recognising certain hand gestures of the worker during the assembly process [4]. Recent hand tracking methods use model-based tracking methods that use generative hand models to estimate full joint angle poses that best explain the observed hand position [5].

Other than hand tracking, there exist studies on defect detection of products on the assembly line. Defect detection methods rely on the availability of the data containing those defects. Traditional computer vision methods like the Computer Aided Design (CAD) method are used to perform the assembly monitoring of the product that does not require data collection [6]. More recently, predictive methods are used for quality assurance in manufacturing due to their ability to detect various features from images with near-human accuracy. Predictive methods require sufficient sample data to learn a good feature representation of the defects. Predictive methods such as Naïve bayes and Decision trees are also used for quality inspection of products on the assembly area [7]. Deep learning methods like convolutional neural network (CNN) and generative adversarial network (GAN) can also be used to detect defects [8-9]. The novelty of deep learning methods as opposed to other predictive methods is that these methods incorporate the feature extraction during the training.

# 3 Assembly line quality assurance framework

To facilitate the process of quality assurance on the assembly line, the proposed QA system framework consists of the following functionalities: the experimental setup, data collection, hand tracking and object detection, and user interface and reporting tool. The complete QA system will continuously track the parts supply bins, assembly area, and hand movements to determine if the sequence of hand movement is correct or incorrect, and if a part component is assembled in the wrong location.

## 3.1 Experimental Setup

The setup consists of a single assembly workstation that is intended for the assembly of the continuous positive airway pressure (CPAP) device and is shown in Figure 1. The camera used for image capturing is a low-cost Logitech C930e webcam and it is mounted on a stand facing vertically down to the assembly line workstation.

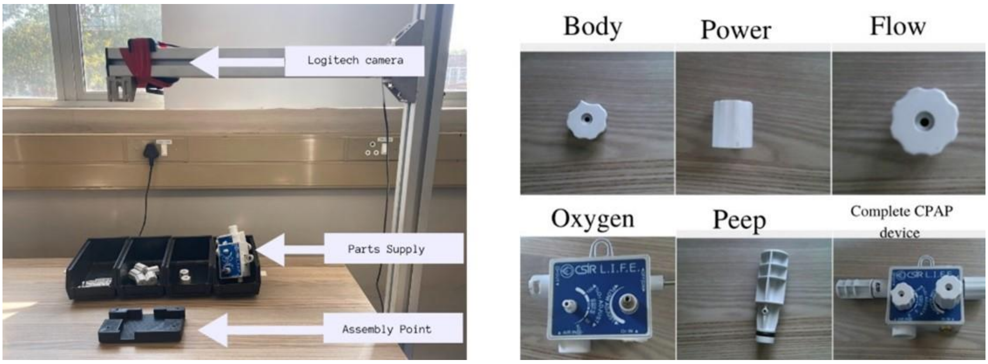


**Figure 1.** Physical setup of the assembly line (left) and the CPAP parts (right).

### 3.2 Data collection and annotation

To develop a state-of-the-art predictive model for quality assurance, quality data about the products on the assembly needs to be collected and processed for modelling. For the dataset, 3,355 images were taken and annotated using a computer vision annotation (CVAT) online tool. In total, 5,723 annotations were created from the captured images. Table 1 shows the number of annotated classes per image and their corresponding annotations per image. 90% of the dataset was used for training and 10% was used for testing.

**Table 1.** Sample data for CPAP components.

| **Class names** | **Total number of annotations per class** | **Images annotation with 1 class** | **Images annotation with 2 objects classes** | **Images annotation with 3 or more object classes** |
|---|---|---|---|---|
| Body | 1204 | 886 | 0 | 318 |
| Power | 1179 | 447 | 219 | 513 |
| Flow | 1112 | 353 | 186 | 573 |
| Oxygen | 1111 | 488 | 186 | 437 |
| Peep | 1117 | 459 | 219 | 439 |
| Year | 5723 | 2633 | 810 | 2280 |

### 3.3 Hand-tracking and object detection components

The full functionality of the proposed QA systems incorporates two components: hand-tracking and object detection. The hand-tracking component is used to monitor the order which the worker follows to assemble parts and the object detection component is used to monitor the correctness of the part being assembled.

The hand-tracking component adopted for this work is a MediaPipe framework developed by Google [10]. MediaPipe Hands is a ML library that can detect and track the hands and fingers. For each hand, 21 joints that describe the hand are detected as shown in Figure 2. To optimize the hand-tracking component for this work, the centroid position of the detected hand fingers is used as identifiers to locate the hand.

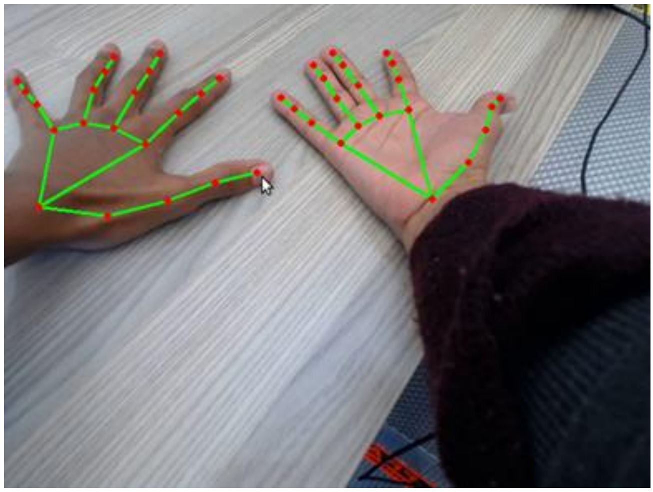

**Figure 2**. Hand tracking using MediaPipe.

To identify whether the hands have been to the correct bin or not, markers around the parts bins and assembly point are drawn with the following colours: white for no hand present, green for the correct execution and red for the incorrect order of execution. The system is always aware of the correct next step the worker must take, therefore, the hands of the worker are observed, analysed, and recorded as correct or incorrect. The output that is displayed to the worker is shown in Figure 3.

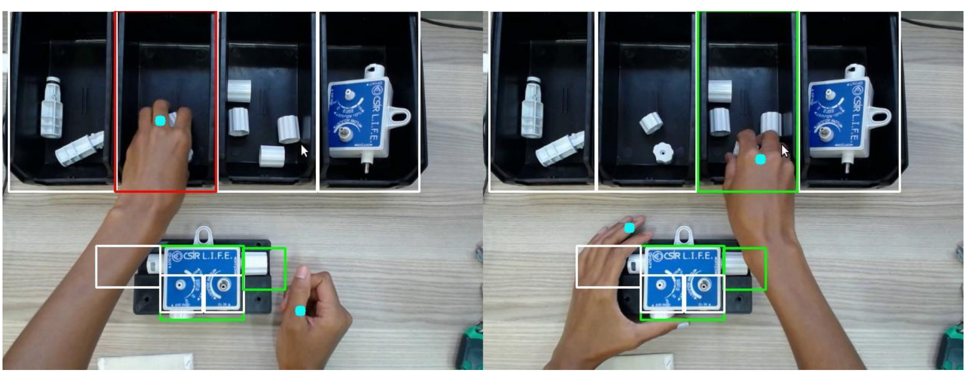


**Figure 3.** The incorrect (left) and correct (right) hand movements demonstration.

For the CPAP object detection, the YOLO (you only look once) object detection model was used for model training [11]. The YOLO model is a type of convolutional neural network that classifies and localises objects on an image. For training the YOLOv3 on the CPAP data, we used the Darknet open-source library [12]. The hyperparameters used for training are shown in Table 2, with parameters that were set to the YOLOv3 default values being omitted.

**Table 2.** Hyper-parameter configuration for the YOLO detection.

| Hyper-Parameter | Value | Description |
|---|---|---|
| Batch | 64 | Number of samples processed every iteration on training set |
| Subdivision | 4 | Fraction of batch size allowed at once per training |
| Steps | 8000, 9000 | Learning rate update policy |
| Max-batches | 10000 | Number of training iteration |
| Classes | 5 | Number of objects for detection |
| Filters | 30 | Number of learned weights for convolutions |

To evaluate the performance of the model over the test dataset, the precision, recall and mean average precision (mAP) metrics were used. The results show that the model's precision, recall and mAP probabilities were 0.96, 0.94 and 0.89 respectively. The high-performance results show that the model detects the correct part objects with high probability. Therefore, the CPAP YOLOv3 model can detect all the parts instances for the assembly of CPAP.

## 3.4 The overall quality assurance system

### *3.4.1 System Framework*

For the overall framework of the QA system, Figure 4 illustrates the systemic design of the QA system. The purpose of the framework is to illustrate how various parts of the system are linked together. The system performs three main functions which are to track the hands of the worker, to detect the components parts and to log the process data to the Manufacturing Execution System (MES).

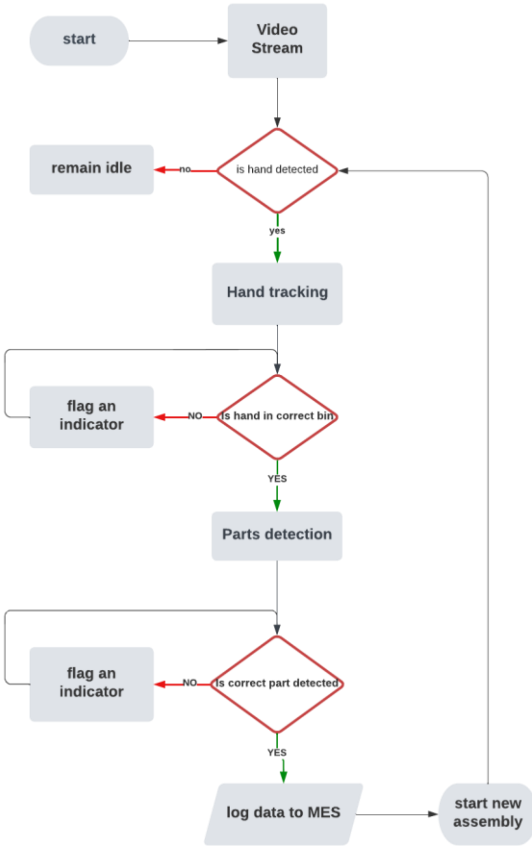


**Figure 4.** QA system framework.

### *3.4.2 User Interface*

The hand tracking together with the CPAP prediction model are integrated to form a QA system used as an assistive tool during the assembly process. The QA system first begins by tracking the hand movement into the correct parts supply bin, then the object detection model is used to determine if the part is assembled correctly. To give interactive feedback to the worker, the user interface for the QA system is designed as shown in Figure 5. As the hand moves into bin parts, the bin markers changes colour to green or red to indicate whether the hand has been to the correct bin or not. At every correct CPAP detection, the assembly marker changes to green and remains green until a CPAP assembly is complete. The user interface also displays information of the completed stages to guide the worker.

https://doi.org/10.1051/matecconf/202237007003

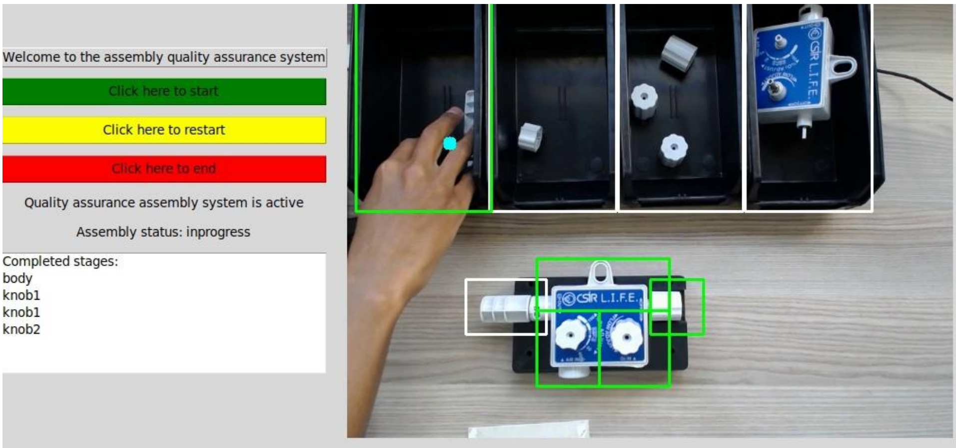


**Figure 5.** QA system user interface.

### *3.4.3 Reporting tool*

For the QA system to be a useful tool in manufacturing, the system must be integrated with the Manufacturing Execution System (MES) [13]. The MES is a software system that enables digitisation of manufacturing processing to deliver information that enables the optimisation of the production line activities from start to finished products and provides functionalities such as product creation, machine monitoring and dashboard. The MES system is put into place to guide, initiate, report, and respond to plant activities as they occur [13].

The QA system was integrated with the MES for progress monitoring on the assembly as shown in Figure 6. The reporting tool captures the status of the QA system: state of the assembly, sensor information about the time taken to complete the assembly of each part component, and historical data analysis about the time taken to assemble each CPAP parts. This information provides great analysis about the assembling of the CPAP and can be used to infer about the system state to tell if it active, idle, or inactive. The historical data can give us information about the time it requires to complete each stage of the CPAP assembly.

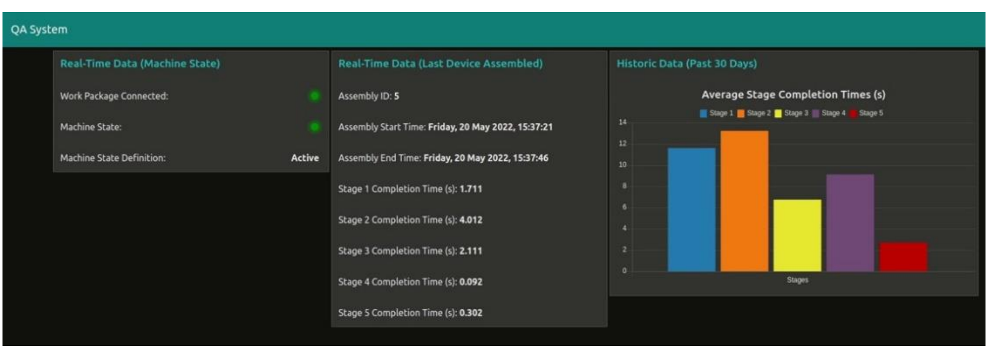


**Figure 6**. The MES reporting tool for the QA system.

## 4 Conclusion

A QA system for monitoring and assisting workers during an assembly process was developed. The system tracks the hands of the worker and displays an indicator for correct or incorrect hand movement. The system also employs a detection model to detect if the parts are assembled correctly or not. Both the hand tracking and object detection models use ML algorithms for their functionality with the hand tracking method being fully end-to-end open-source and the object detector trained locally on the CPAP data. The QA system offers a robust hand tracking and object detection solution capable of monitoring the worker and the quality of assembled component in real time. The QA system has the potential to flag inconsistencies in manual assembly lines and can thereby improve the quality of the assembly lines' output.

For future work, the work will incorporate flexibility to allow the system to monitor the assembly steps in a nonsequential manner and allow the rework of a step at any time. In addition, this work can be extended to include a framework to identify defect free parts components for the delivery of quality products. Therefore, future work will investigate these issues to enable for easy adaption to other projects

## References

[1] R. Cioffi, M. Travaglioni, G. Piscitelli, A. Petrillo, F. De Felice, Artificial Intelligence and Machine Learning Applications in Smart Production: Progress, Trends, and Directions. *Sustainability*, **12**, 492 (2020).
[2] I. Kastanis, M. Hochemer, S. Widmer, P. Schmid, Vision-based Monitoring of Manual
[3] M. Galindo-Salcedo, A. Pertúz-Moreno, S. Guzmán-Castillo, Y. Gómez-Charris, A.R. Romero-Conrado, Smart manufacturing applications for inspection and quality assurance processes, Procedia Computer Science, **198**, 536-541 (2022)
[4] *S. Kaczmarek, S. Hogreve, K. Tracht, Progress Monitoring and Gesture Control in Manual Assembly Systems Using 3D-image Sensors, Procedia CIRP, **37**, 1-6 (2015)*
[5] S. Sridhar, F. Mueller, A. Oulasvirta, C. Theobalt, Fast and Robust Hand Tracking Using Detection-Guided Optimization, (2016)
[6] N. Piero, M. Schmitt, Virtual Commissioning of Camera-based Quality Assurance Systems for Mixed Model Assembly Lines, Procedia Manufacturing, **11,** 914-921 (2017)
[7] *J. Schmitt, J. Bönig, T. Borggräfe, G. Beitinger, & J. Deuse, Predictive model-based quality inspection using Machine Learning and Edge Cloud Computing. Adv. Eng. Informatics, **45**, 101101 (2020)*
[8] A. Deshpande, A. Minai, M. Kumar, One-Shot Recognition of Manufacturing Defects in Steel Surfaces, Procedia Manufacturing, **48**. 1064-1071 (2020)
[9] *M. Mazzetto, M. Teixeira, E.O. Rodrigues, D. Casanova, Deep Learning Models for Visual Inspection on Automotive Assembling Line, (2020)*
[10] C. Lugaresi, J. Tang, MediaPipe: A Framework for Building Perception Pipelines, Computing Research Repository, (2019)
[11] *J. Redmon, S.K. Divvala, R.B. Girshick, A. Farhadi, You Only Look Once: Unified, Real-Time Object Detection. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 779-788, (2016)*
[12] J. Redmon, Darknet: Open-source neural networks in c, (2014)
[13] J. Fraser, S. Cloughley, MES Explained: A high level vision, *MESA International white paper number 6, (1997)*