Empowering Digital Twin for Production Operations with Deep Learning: A Case Study

Eric C.H. Wai
Department of Industrial and Systems Engineering
The Hong Kong Polytechnic University
Hong Kong SAR
chwai2003@hkstar.com

Dr. Carman K.M. Lee
Department of Industrial and Systems Engineering
The Hong Kong Polytechnic University
Hong Kong SAR
ckm.lee@polyu.edu.hk

Abstract

This study aims to establish an intelligent Cyber-Physical Production System (CPPS) with closed-loop control and self-optimization of process parameters. Empowered by a deep learning (DL) agent, the digital twin (DT) simulation can focus on the most significant process routes and eliminate 92.53% of unnecessary exhaustive iterations. On the other hand, DT can provide the most accurate simulation result for fast convergence to the optimal parameters of DL and no need to waiting for the result after the process completed. Though collaborative and autonomous interactions between DT and DL have been proposed recently, the practical integration of DT-DL and CPPS in the Electronic Manufacturing Service (EMS) industry has yet to be validated in official publications. With such novelty approach, the improvement of the production throughput (T/P) and rolled throughput yield (RTY) is 7.513% and 0.86%, respectively, which is on top of the contributions from traditional industrial engineering (IE) productivity tools.

Keywords
Deep Learning, Digital Twin, CPPS, Simulation, Productivity Improvement

1. Introduction

Traditional Industrial Engineering (IE) improvement tools such as Kanban (Reis et al. 2019), Value Stream Mapping (VSM) (Blume et al. 2020), Kaizen (Singh et al. 2009), LEAN system (Bradley and J. R. 2012), DMAIC in 6-sigma program (Burton et al. 2011) can have combined manufacturing productivity increment up to 42%. After adopting those tools and reaching their optimal settings, it is challenging to have dramatic improvement again. Also, some of these tools, like VSM and LEAN systems, are inevitably overlapping each other during implementation. In order to have productivity break-through, the modern Cyber-Physical Production System (CPPS) is adopted for further improvement. Deep Learning (DL) agents can find the missing links between the process parameters and the affected productivity, which humans might not be easily observed (Wang et al. 2018). The Digital Twin (DT) modeling can simulate the DL result (Elisa Negri 2017) and reduce the actual epochs in production iteration for fast convergence to its optimal value. By integrating both essential elements in DL and DT for smart manufacturing, self-configure, self-adapt, and self-learn become possible. Which can significantly increase productivity, speed, flexibility, and efficiency (Jay Lee et al. 2020).

Empowered by Long Short-Term Memory (LSTM) DL, the DT simulation in the CPPS can focus on the most significant process routes and eliminate 92.53% of unnecessary exhaustive iterations. On the other hand, DT can provide the most accurate simulation result for fast convergence to the optimal parameters of DL and no need to waiting for the result after the process completed. Though collaborative and autonomous interactions between DT and DL have been proposed recently (Jay Lee et al. 2020), the practical integration of DT-DL and CPPS in the EMS industry has yet to be validated in official publications. With the DT-DL novelty approach, the improvement of the production throughput (T/P) and rolled throughput yield (RTY) is 7.513% and 0.86%, respectively, which is on top of the contributions from traditional industrial engineering (IE) productivity tools.

The significance and value of this research is providing the practical approaches to bridging the research theories to actual operations of EMS factory. The research outputs might contribute to the industry transformation and towards to Industry 4.0 Smart Factory by CPPS implementation.

2. Literature Review

A digital twin is a virtual representation that serves as the real-time digital counterpart of a physical object or process. Though the concept originated earlier by Grieves and Michael (2016) the first practical definition of digital twin originated from NASA in an attempt to improve the physical model simulation of spacecraft in 2010 (Elisa Negri 2017). Digital twins (DT) are the result of continual improvement in the creation of product design and engineering activities. Product drawings and engineering specifications progressed from handmade drafting to computer-aided drafting/computer-aided design (CAD) to model-based systems engineering. In essence, a computer program that uses real-world data to create simulations that can predict how parameter changes affect the overall performance. These programs can integrate the internet of things, artificial intelligence, and software analytics to enhance the output. With the advancement of machine learning and factors such as big data, these virtual models have become a staple in modern engineering to drive innovation and improve operational efficiency.

Deep learning (DL) is regarded as a breakthrough in artificial intelligence (Wang et al. 2019), which demonstrates outstanding performance in various applications of speech recognition, image reconditions, natural language processing, multimodal image-text, and games (e.g., Alpha-go). Deep learning allows automatic data processing toward highly non-linear and complex feature abstraction via a cascade of multiple layers instead of handcrafting the optimum feature representation of data with domain knowledge. With automatic feature learning and high-volume modeling capabilities, deep learning provides an advanced analytics tool for smart manufacturing in the big data era.

Coronado et al., (2018), combined the DT and Cloud implementation, can provide MES shop management from the mobile device. The manufacturing status, such as production efficiency, machine capacity, and material level, can also be well monitored remotely. The multiple layers model described by Redelinghuys et al. (2019) stated the viable interconnections in different functionality. Leng et al., (2020) have demonstrated the feasibility of reconfiguration and rapid changes in manufacturing systems. Based on the Edge-Fog-Cloud model proposed by Y.H. Pan et al. (2021), the DT can be defined as three layers with different features, such as emulation and simulation, predictive analytics, and remote control.

Integration of the DT and DL has been a novelty approach in the last three years. Jay Lee et al. (2020) proposed architecture of DTDL-CPS implements a service layer that provides intelligence, control, visualization, optimization, PHM, etc. for implementation in the physical system and outlines the development of smart manufacturing with enhanced transparency, cooperation, networking, resilience, and efficiency. Chao Zhang et al. (2021) demonstrated the DL in DT manufacturing cells (DTMC) to optimize the CAD/CAM process. Qiyue Wang et al.(2020) implemented the CNN-based DL and visualized GUI DT for welding process monitoring and control. Pasquale Franciosa et al. (2020) improved the Closed-Loop In-process (CLIP) control with CAE DT simulations and DL enhancement. The first pass yield (FPY) can reach > 96% in laser welding aluminum doors.

Even though the DL-DT collaboration has been proposed or reviewed and well discussed recently, the practical integration of DT-DL and CPPS in actual electronic manufacturing service (EMS) operations has yet to be validated in official publications. In this study, we try to implement the DT-DL at the operational level to fill up the research gap.

3. Methods

In this study, we proposed a Floor-Line-Station-Node model (FLSN) shown in Figure 1 to illustrate the overall DT architecture. From the bottom, the production node (E, Executor) cell is controlled by the Edge computer (S, node selector). They form the foundation of the production stations. On top of the -S-N, we have a Fog Computing layer (MES) facilitator to control the Line level activities (L). Finally, the ERP, WMS, SCM, APS at the cloud level will control the whole production floor level (F). FLSN hierarchy can help the organization manage the production states in detail without sacrificing efficiency. The concept of DL also has been applied in 3 implementation levers (Station, Line, Floor) with three computing approaches (see Table I).

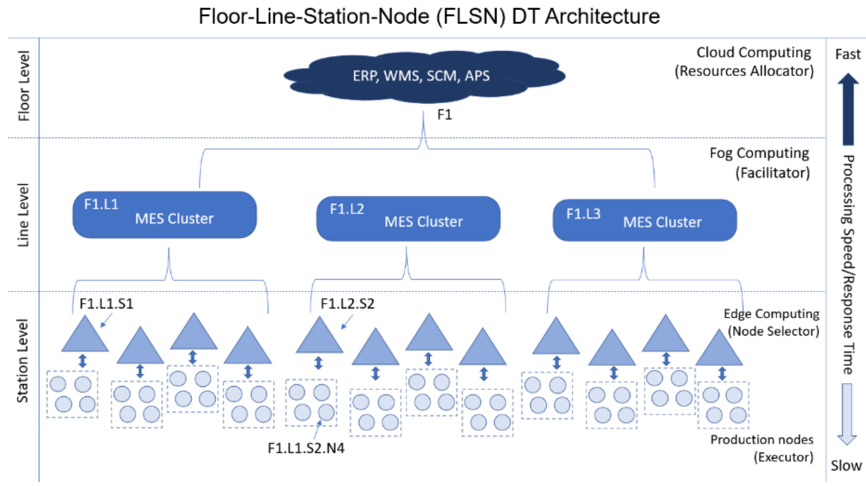


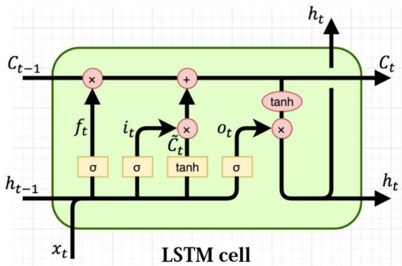
Figure 1. FLSN DT Architecture in 3 computing layers

Table 1. Summary of 3 DT Implementation levels

Implementation Level	**Technique Adopted**	**Deploy mode**
Station Level	Edge Computing	On-premise
Line Level	Fog Computing	On-premise
Floor Level	Cloud Computing	On-cloud

A Recurrent Neural Network (RNN) (Varsamopoulos et al 2018) is part of a broader family of Deep Learning (DL) methods that can produce representations automatically based on raw data in a variety of applications like classification, regression, clustering, and pattern recognition based on this data. RNN is constructed with Long Short-Term Memory (LSTM) cells as nodes. To decide which signals are to be forwarded to another node, an LSTM cell has extra gates, such as the input, forget, and output gates. Recurrent connections between previously hidden layers are indicated by W. The weight matrix U connects the inputs to the hidden layer. Using the current input and the previous hidden state, $\tilde{C}$ is a candidate's hidden state. The input gate multiplies the newly computed hidden state by the gate to become C, the unit's internal memory. In Figure 2, all gates in the LSTM cell are described by equations that describe their behaviours. In the backpropagation of RNNs, using computed gradients to adjust the weight matrix can make the gradients in the network eventually become too small and disappear, or too large and explode, due to many successive multiplications of partial derivatives, which makes it difficult for RNNs to learn information over long distances.

The cell states (i.e., long-term memory) in the LSTM simply require a linear summation operation to pass through the hidden layer, and the gradients can easily move between networks without decay. LSTM also allows the neural network to switch between remembering recent information and information from long ago, allowing the data to decide for itself which information to keep and which to forget.



$$i_t = \sigma\left(x_t U^i + h_{t-1} W^i\right)$$
$$f_t = \sigma\left(x_t U^f + h_{t-1} W^f\right)$$
$$o_t = \sigma\left(x_t U^o + h_{t-1} W^o\right)$$
$$\tilde{C}_t = \tanh\left(x_t U^g + h_{t-1} W^g\right)$$
$$C_t = \sigma\left(f_t * C_{t-1} + i_t * \tilde{C}_t\right)$$
$$h_t = \tanh(C_t) * o_t$$

Figure 2. Structure of the LSTM cell and equations of an LSTM cell. (Varsamopoulos, et al 2018)

In Figure 3, consider that we have ***n*** processing nodes. $f(x)_{\text{pn}}$ is the function to output the yield rate ***Y*** and ***O***. $f(x)_{\text{pn}}$ included environment-related factors $f(x)_{env}$ and machine-dependent factors $f(x)_{M1}$, $f(x)_{M2}$ …… $f(x)_{Mn}$, . Each $f(x)_{Mn}$ included a few hidden coefficients (weighting factors ***Wn***) obtained by DL. Such coefficients will be extracted and used in DT for simulating the results (dotted line path). The process selector will use the DT simulated result to obtain the best productivity route or re-routing in real-time. Once the actual process was completed. A new set of actual data will fit into the DL to update the hidden coefficients again (Solid line path). Fusing virtual and actual data (DL -> DT -> DL loop) will help the DT have a more accurate result and improve the DL training efficiency to converge to optimal value rapidly.

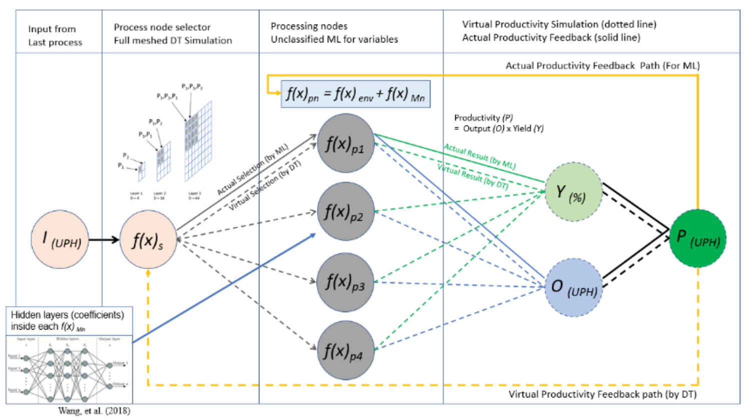


Figure 3. The illustration of Virtual and Actual Productivity feedback to DT Simulation and DL

The DL included three parts; there are input, processing, and output:

Input:

The Overall rolled yield rate for each process node: $\{Y_1,Y_2,Y_3,\ldots,Y_n\}$
The Output for each process node: $\{O_1,O_2,O_3,\ldots,O_n\}$
The parameter set for each process node: $\{\{S_1\},\{S_2\},\{S_3\},\ldots,\{S_n\}\}$

Processing (Connecting hidden layers):

$$f(x)_{\text{Mn}} = Y_n * W_1 + O_n * W_2 + \{ S_{n1} * W_3 + S_{n1} * W_3 + S_{n2} * W_4 + S_{n3} * W_5 \;\ldots.\} \quad (1)$$

Where W_1, W_2, W_3 …. W_n are the weighting factors determined by DL

Output:

The production node-dependent yield rate for each process node:

$$f(x)_{\text{Mn}} = \{ W_1, W_2, W_3, W_4\} \quad (2)$$

The production node independent yield rate for all process nodes:

$$f(x)_{\text{env}} = W_{\text{env}} \tag{3}$$

Many possible factors may affect the performance of the station ($P_{(\text{UPH})}$) which can be machine-dependent ($f(x)_{\text{Mn}}$) or environment-related ($f(x)_{\text{env}}$). It is impossible to list out all of them before defining a precise weighting function $f(x)_{pn}$. Instead, an unclassified DL tool is implemented to obtain the weight of each parameter with a certain amount of accuracy after training. Also, DL can adjust and self-optimize in real-time.

For genialize, the total number of DT simulation Epochs(***D***) from possible processes (***P*** $_{\text{Total}}$) by (***L***) coming production cycles can be calculated below:

$$\boldsymbol{D} = P_{Total\,1} \times P_{Total\,2} \times P_{Total-3} \times \ldots\ldots \times P_{Total\,L} \tag{4}$$

$$\boldsymbol{D} = P_{Total}{}^{L}$$

Where ***D*** = Epochs of DT need to simulate;
P $_{\text{Total}}$ = Possible process can be selected;
L = DT Layers (The production cycle we need to look forward.)

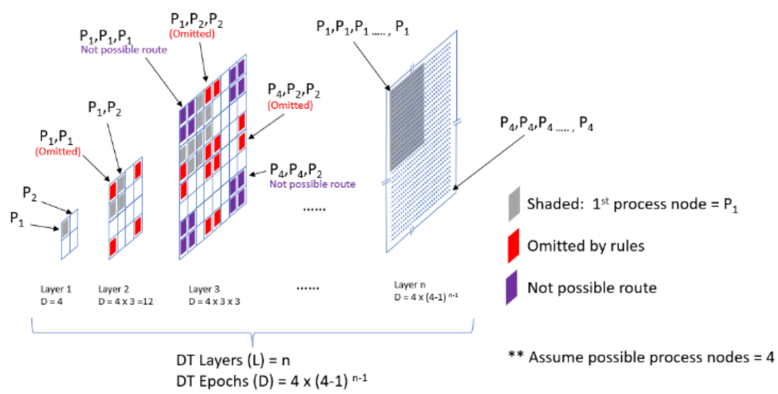


Figure 4. The illustration of Virtual and Actual Productivity feedback to DT Simulation and DL

Considering a production cluster (***P₁, P₂, P₃, P₄***) with a total of four possible process nodes, the objective is to find the maximum productivity by ten coming production cycles with DT simulation; more than 1 million Epochs are required.

Figure 4 describes a reduced DT Epoch, and layers for four process nodes, and the route can be significantly reduced by examining whether the process is the feasible route or not, and the node configuration can be trained by DL.

$$\boldsymbol{D} = 4^{10} = 1{,}048{,}576 \text{ Epochs} \tag{5}$$

However, in a typical application and load balancing consideration, the same process will not be used 2 consecutive times (Figure 7). As a result, the DT Epochs can be reduced as below:

$$\boldsymbol{D} = P_{Total} \times (P_{Total} - 1)^{L-1} \tag{6}$$

For a production cluster with 4 nodes, the epochs can be reduced to 78,372. Which reduced 92.53% of the unnecessary exhaustive iterations after training.

$$\boldsymbol{D} = 4 \times (4-1)^{10-1}$$

$$\boldsymbol{D} = 78{,}372 \; Epochs \tag{7}$$

Depth of the DT Simulation

The processing power of a typical industrial edge computer in our case requires around 0.037 ms per DT epoch simulation. The cycle time is around 16 seconds per process notes, therefore the most appropriate DT Layers is 10 for this application.

Before we can make use of the DL and DT tool, we need to define productivity. Among many KPIs for measuring productivity, Output and Yield loss are the most common and used in this study. The Productivity $\boldsymbol{P_{pn}}$ for each production node is defined as below:

$$\boldsymbol{P_{Pn}} = Output \times (1 - Y_{env}) \times (1 - Y_{Pn}) \tag{8}$$

Where the P_{pn} is the Node Productivity; Y_{env} is the yield loss independent to the Node; Y_{Pn} is the yield loss dependent on the Node.

4. Data Collection

In order to realize the concept of the digital twin (DT), one production line in the research sponsors' company is provided for experimental setup and data collection. The operational information and findings are allowed to share for academic research without customer-related information.

The Station level optimization with DT is illustrated in Figure 5 and the goal of using the DT and DL tools is to maximize productivity in a single station. In this study, a UV Glue Dispensing process is selected. The process of gluing the PCBA from the last process (5.1) to the mechanical part (5.2) together with UV Glue dispenser (6). Then the inline AOI (7) inspects the product and estimates the station's unit per hour (UPH). Then the product will go to casing process (8.1) or scrap (8.2) if failure. Figure 6 shows the physical process of one station. After digitalization, the workstation and the layout can be redesigned with simulation. DL can achieve real-time adjustment process sequence through the formation of the positive feedback loop.

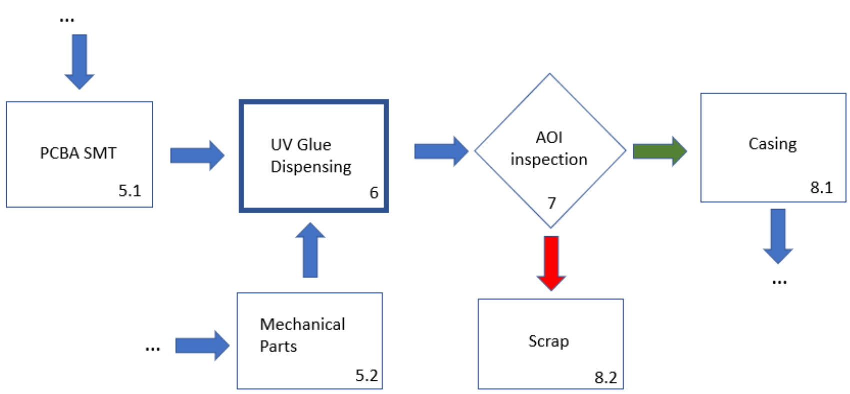


Figure 5. Illustration of process workflow. UV Glue Dispensing process in step 6.

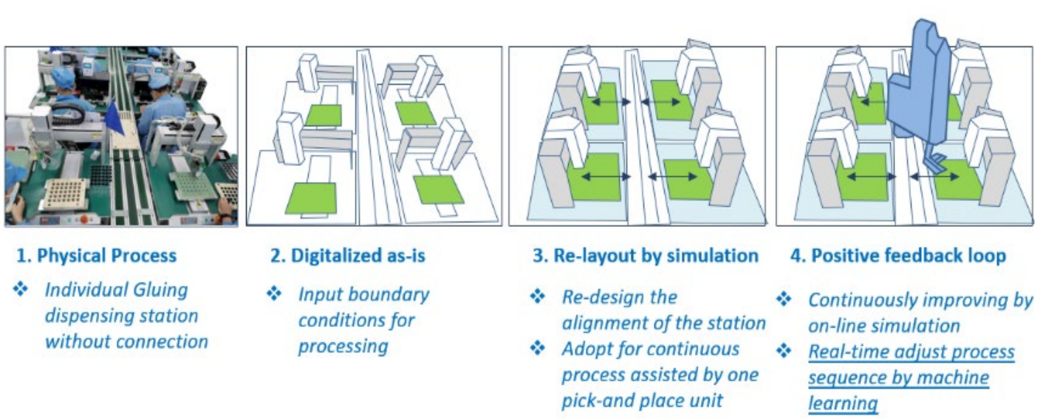


Figure 6. The illustration of re-design the process nodes for DT implementation.

For support DT-DL processing, a GPU-equipped industrial PC is setup for Edge computing. An open-sourced Python package - PyAnsys is used for DT simulation and Keras is used for RNN-LSTM training.

The self-optimization process is carried out based on the DT simulation, DL model, and productivity defined. Table 2 shows the reduced DT epoch and the layer of process nodes in a single cycle. Based on the finding, we can realize each node's effective productivity and normalized cycle time.

Table 2. Illustration of reduced DT Epoch and layers for four process nodes in a single cycle.

	DL Input			**DL Output**		**DT Simulation Factors**	
Node	Node yield loss rate	Parameters Set	Output $_{UPH}$	Nodes independent yield loss = $f(x)_{env}$	Nodes dependent yield loss = $f(x)_{Mn}$	Effective Productivity (UPH)	Normalized Cycle Time (S)
P1	0.0334	Set $_{P1}$	246.3224	0.0121	0.0213	238.0952	15.12
P2	0.0298	Set $_{P2}$	238.3157	0.0121	0.0177	231.2139	15.57
P3	0.0272	Set $_{P3}$	235.1117	0.0121	0.0151	228.7166	15.74
P4	0.0345	Set $_{P4}$	234.0640	0.0121	0.0224	225.9887	15.93

The overall productivity is ranged from **228.5714** to **598.8025** units per hour (UPH) in all possible node sequences (10 cycles) with ten processing layers shown in Table 3. The best ranking with three processing layers is equivalent to the #13 ranking of the 10 processing layers. However, since the yield loss is changing from time to time, the normalized cycle time and productivity will also changing throughout the process.

Table 3. Illustration of all possible Node Sequence (78,732).

Rank	**Node Sequence (10 cycles)**	**Critical Path**	**Cycle time (10 Units)**	**Productivity (UPH)**
#1	P2,P1,P3,P4,P2,P1,P3,P4,P2,**P1**	45 + 0 + $\mathbf{S_{P1}}$	60.12	**598.8024**
#2	P2,P1,P4,P3,P2,P1,P4,P3,P2,**P1**	45 + 0 + $\mathbf{S_{P1}}$	60.12	598.8024
#3	P3,P1,P4,P2,P3,P1,P4,P2,P3,**P1**	45 + 0 + $\mathbf{S_{P1}}$	60.12	598.8024
#4	P3,P1,P2,P4,P3,P1,P2,P4,P3,**P1**	45 + 0 + $\mathbf{S_{P1}}$	60.12	598.8024
….	….			
#13	P1,P2,P3,P4, P1,P2,P3,P4,P1,**P2**	45 + 0 + $\mathbf{S_{P2}}$	60.57	594.3536
….	….			
#78732	P4,P3,P4,P3,P4,P3,P4,P3,P4,**P3**	45 + 96.76 + $\mathbf{S_{P3}}$	157.50	228.5714

The process input from the previous station is 720 UPH, equivalent to 5 seconds of cycle time. The theoretical output of the station is limited by 720 UPH. In the review of the DT optimization, we initially set up the depth of the DT Layer to **3** and run ten cycles; it is found that the reduced possible routes are **36**. The best-selected route is P1,P2,P3,P4,P1,P2,P3,P4,P1,P2, *(Figure 7)* which follow the node rating by DL and fill up sequentially. Overall productivity is **594.3536** UPH for this setup.

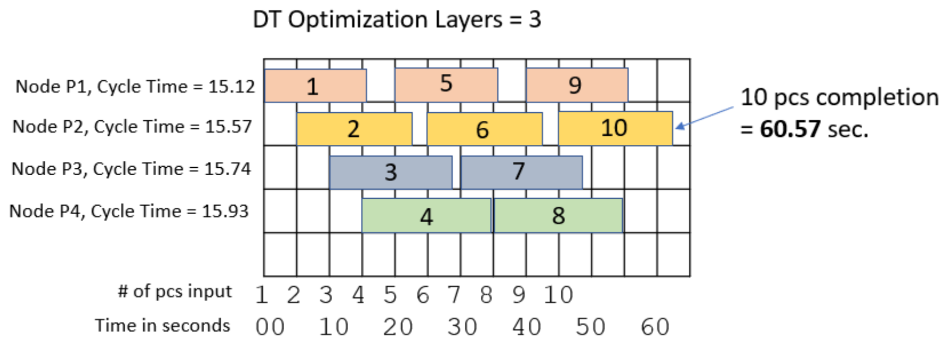


Figure 7. Illustration of productivity by 3 DT optimization layers.

The depth of DT Layer to 10 has been further set up. The reduced possible routes become 78,732. The best-selected route is P2,P1,P3,P4,P2,P1,P3,P4,P2,P1, (Figure 8) which surprisingly does not follow the node rating by DL in starting. However, the overall productivity is improved to **598.8024** UPH. After investigating the details, we found the bottleneck is the last pcs of material inputted at 45th second, and Node P1 can provide the shortest cycle time. In this example, we can demonstrate the DT optimization located some possibility that humans might overlook.

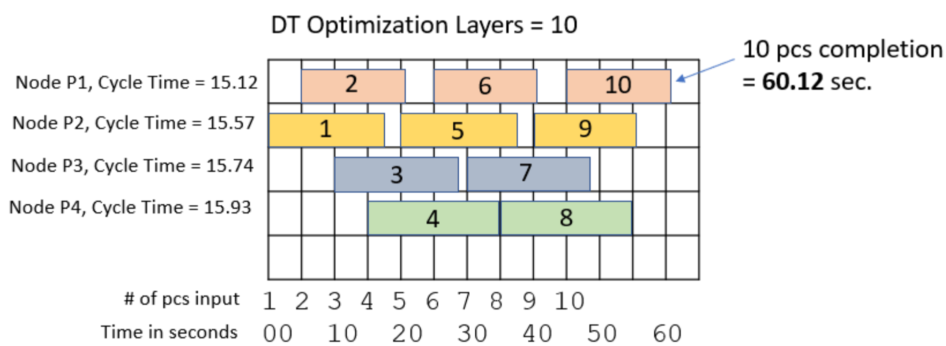


Figure 8. Illustration of productivity by 10 DT optimization layers.

Once station level DT-DL (Edge server) updates the cycle time and production status, the result will be synchronized to the line level MES system (Fog server) for further processing in FLSN architecture.

5. Results and Discussion

5.1 Numerical Results

After implementing the DT-DL in Glue dispensing station, in 20 consecutive production shifts, we found the overall yield has been improved by 0.86% and the station throughput increased by 15.924 UPH (7.513%) (*Table 4*). The standard deviation (StDev) was reduced by 0.15% in yield rate and 0.361 UPH in throughput. The reduction of the standard deviation shows us that the station is more stable than before. We reduced 5.2 hours in 20 production shifts for setup time because the process selector - 4-axis SCARA robot arm can have the self-configuration program and no need for extra manual work. As the routing operation is processing at edge level, the workload and response time is significantly reduced in fog level MES system.

Table 4. Illustration of the improvement after implementing DT and DL

KPI	**Before (July 2021, 20 Production Shift)**	**After (Oct 2021, 20 Production shift)**	**Difference**
Overall Yield Rate	95.836%	96.694%	+ 0.86%
StDev of Yield Rate	0.621%	0.476%	- 0.15%
Station Throughput (T/P)	211.952 UPH	227.876 UPH	+ 15.924 UPH (+7.513%)
StDev of T/P	1.522 UPH	1.161 UPH	- 0.361
Setup Time	22.41 hrs	17.21 hrs	- 5.2 hrs

5.2 Graphical Results

Figure 9 shows the Statical analysis of the Throughput (T/P) improvement. P-value <0.005 in T-Test showing we have a significant difference between "Before" and "After" in the hypnosis test of Throughput (T/P) yield rate in 95% Confidence Interval (CI).

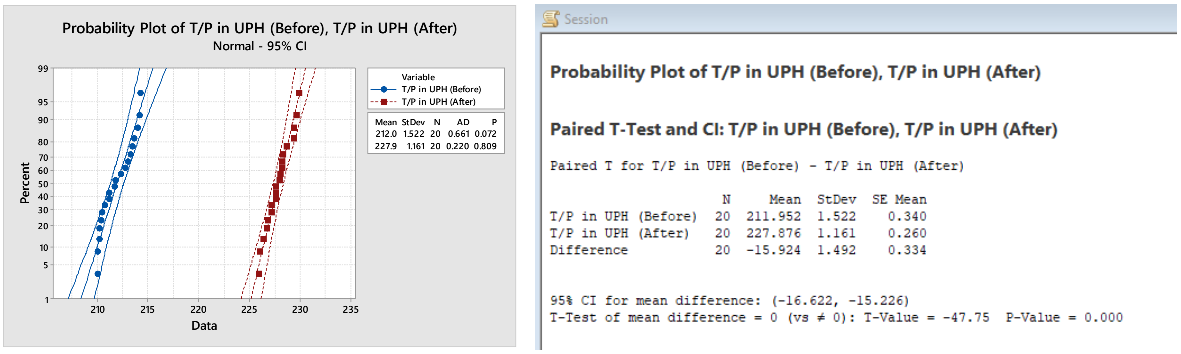


Figure 9. Statistical analysis of the Throughput (T/P) improvement.

5.3 Proposed Improvements

In this study, the DT-DL approach has been proposed and improved productivity at the station level. Although the results are encouraging, we still need more experience and wider implementation to validate the DT-DL integrated model. In future development, more generalized and universal models may be set up for other manufacturing processes such as component insertion, screwing, and ultrasonic welding which may help the manufacturing industry.

5.4 Validation

To validate the DT estimation performance, three consecutive production days are used for benchmarking in *Table 5*. The glue dispensing process is 12 hours per day from 08:00 to 19:59 and is possible to expand to 24 hours. There are 2 break times in the shift. One is a lunch break from 12:00 to 13:30 and a Tea break from 16:15 to 16:30 which is reflected in the duty cycle ratio. DT estimation, actual output, yield loss, and completed products are measured in pcs.

Table 5. Illustration of DT Deviation in 3 consecutive production days

	Production Time Slot	Hour Duty Cycle	DT Estimation	Actual Output	Actual Yield loss	Completed Products	DT Deviation (PCS | %)	
Day 1	08:00:00-08:59:59	1.00	598.8	615	19	596	-2.8	-0.47%
	09:00:00-09:59:59	1.00	598.8	620	18	602	3.2	0.53%
	10:00:00-10:59:59	1.00	598.8	620	18	602	3.2	0.53%
	11:00:00-11:59:59	1.00	598.8	622	19	603	4.2	0.70%
	12:00:00-12:59:59	0.00	0	0	0	0	0	0.00%
	13:00:00-13:59:59	0.50	299.4	309	11	298	-1.4	-0.47%
	14:00:00-14:59:59	1.00	598.8	621	20	601	2.2	0.37%
	15:00:00-15:59:59	1.00	598.8	620	17	603	4.2	0.70%
	16:00:00-16:59:59	0.75	449.1	460	11	449	-0.1	-0.02%
	17:00:00-17:59:59	1.00	598.8	622	18	604	5.2	0.86%
	18:00:00-18:59:59	1.00	598.8	625	20	605	6.2	1.02%
	19:00:00-19:59:59	1.00	598.8	623	21	602	3.2	0.53%
Day 2	08:00:00-08:59:59	1.00	598.8	615	21	594	-4.8	-0.81%
	09:00:00-09:59:59	1.00	598.8	622	18	604	5.2	0.86%
	10:00:00-10:59:59	1.00	598.8	620	18	602	3.2	0.53%
	11:00:00-11:59:59	1.00	598.8	622	17	605	6.2	1.02%
	12:00:00-12:59:59	0.00	0	0	0	0	0	0.00%
	13:00:00-13:59:59	0.50	299.4	308	11	297	-2.4	-0.81%
	14:00:00-14:59:59	1.00	598.8	620	20	600	1.2	0.20%
	15:00:00-15:59:59	1.00	598.8	620	16	604	5.2	0.86%
	16:00:00-16:59:59	0.75	449.1	460	12	448	-1.1	-0.25%
	17:00:00-17:59:59	1.00	598.8	622	18	604	5.2	0.86%
	18:00:00-18:59:59	1.00	598.8	623	21	602	3.2	0.53%
	19:00:00-19:59:59	1.00	598.8	622	21	601	2.2	0.37%
Day 3	08:00:00-08:59:59	1.00	598.8	617	20	597	-1.8	-0.30%
	09:00:00-09:59:59	1.00	598.8	620	16	604	5.2	0.86%
	10:00:00-10:59:59	1.00	598.8	620	18	602	3.2	0.53%
	11:00:00-11:59:59	1.00	598.8	621	18	603	4.2	0.70%
	12:00:00-12:59:59	0.00	0	0	0	0	0	0.00%
	13:00:00-13:59:59	0.50	299.4	310	13	297	-2.4	-0.81%
	14:00:00-14:59:59	1.00	598.8	621	21	600	1.2	0.20%
	15:00:00-15:59:59	1.00	598.8	620	16	604	5.2	0.86%

	16:00:00-16:59:59	0.75	449.1	460	12	448	-1.1	-0.25%
	17:00:00-17:59:59	1.00	598.8	623	20	603	4.2	0.70%
	18:00:00-18:59:59	1.00	598.8	625	21	604	5.2	0.86%
	19:00:00-19:59:59	1.00	598.8	621	22	599	0.2	0.03%

Based on the corporate station level KPI in the research sponsors' company, the target deviation is within 1% of the production planning. Compared with actual production and the DT estimation in 3 consecutive production days, the actual production is 2.04 pcs (or 0.34% of 598.8 pcs) higher than the simulation on average. From the Zhang et al (1990) derived formulas, the Process Capability Index (Cpk) is 1.34 with 1% LSL/USL (Figure. 10), which is equivalent to 29.15 of million hours that may fall out beyond the USL.

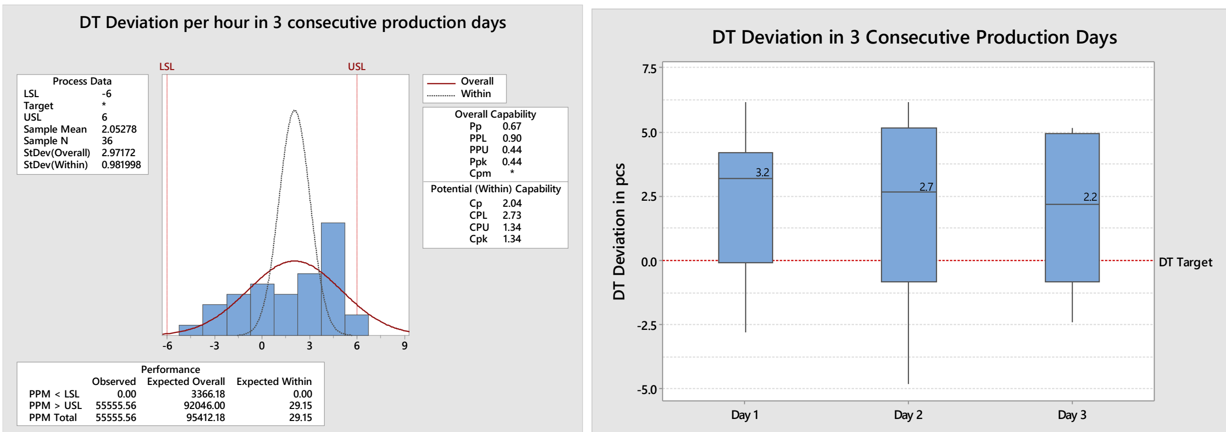


Figure 10. DT Deviation per hour in 3 Consecutive Production Days

The box plot found that the median is 3.2, 2.7, and 2.2 (or 0.53%, 0.45%, and 0.37%) higher than the target. Based on the observations, the DT has slightly underestimated the production output. Having investigated the details with the time-series plot, we found the production output is lower than the DT simulation when production resumed at 08:00, 13:30, and 16:30 (Figure 11). Then the output will ramp up and reach the estimated value. Considering system resource and model complexity, the deviation swing within +/- 1 % is acceptable in practical operations.

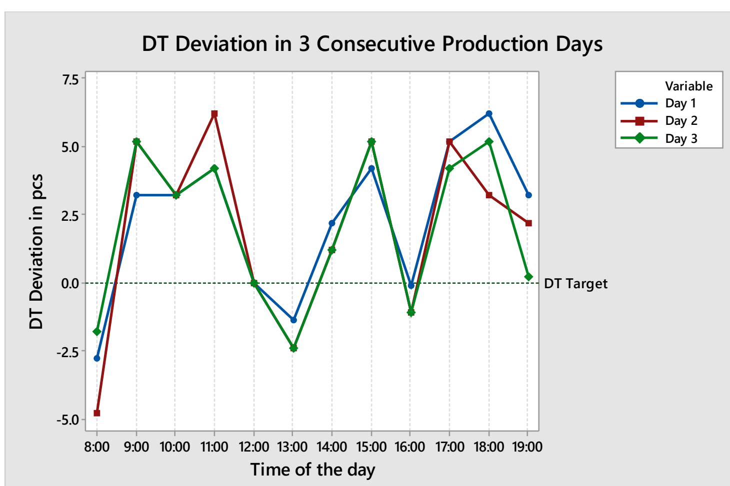


Figure 11. DT Deviation in 3 Consecutive Production Days with time-series plot.

6. Conclusion

We found digital twin for production operations with deep learning can improve the rolled throughput yield (RTY) by 0.86% and the throughput increased by 7.513%. The DT can feedback the simulated data to the DL which can significantly reduce the training time. Although the results are encouraging, we still need more experience and wider implementation to optimize the DT-DL intergraded system.

The DT-DL system can be further extended to the line and floor level with suitable DT and DL strategies. We can implement multiple robotic arms as a process selector for the Line level implementation and use the multiple agents' reinforcement learning strategy. At the floor level, we may implement this approach in cloud APS to schedule the SMT Line. The RNN-based DL can select the best-suited SMT line by product characteristics and process parameters. The DT will further enhance the DL by simulating the production thought-put and yield rate for rapid convergence to the optimized values.

The major contribution of this study is to validate the theoretical DT-DL integration with CPPS in the EMS industry by operational data. Performance and training efficiency can be improved by combining DT and DL. By moving away from traditional methodologies and embracing Industry 4.0, the study case may be able to serve as an example that helps the industry move towards a successful future.

Acknowledgments

We would like to express our gratitude to VTech Communications Limited for their support of this study and their contribution to academic research in this area.

References

Alok B. Patil, Prof. Dr. Kedar H. Inamdar, 2014, Process Improvement using DMAIC Approach: Case Study in Downtime Reduction, INTERNATIONAL JOURNAL OF ENGINEERING RESEARCH & TECHNOLOGY (IJERT) Volume 03, Issue 03 , 2013.

Antony, J., Kumar, M., & Labib, A. . Gearing Six Sigma into UK Manufacturing SMEs: Results from a Pilot Study. The Journal of the Operational Research Society, 59(4), 482–493, 2008.

Blume, S. A. and SpringerLink (Online service) . Resource Efficiency in Manufacturing Value Chains. Sustainable Production, Life Cycle Engineering and Management,: XX, 191 p. 137 , 2020.

Bradley, J. R., Improving business performance with Lean. Supply and operations management collection,: 1 online resource (xii, 255 , 2012.

Burton, T. T. (2011). Accelerating lean six sigma results : how to achieve improvement excellence in the new economy. Ft. Lauderdale, Fla., J. Ross Pub.,: 1 online resource xxii, pp. 409, 2011.

Chao Zhang, Guanghui Zhou, Junsheng Hu, Jing Li, Deep learning-enabled intelligent process planning for digital twin manufacturing cell, Knowledge-Based Systems, Volume 191, 2021, 105247, ISSN 0950-7051

Coronado, Pedro Daniel Urbina, et al. "Part data integration in the Shop Floor Digital Twin: Mobile and cloud technologies to enable a manufacturing execution system." Journal of manufacturing systems 48 , 25-33, 2018.

Elisa Negri ,"A review of the roles of Digital Twin in CPS-based production systems". Procedia Manufacturing. 11: 939–948, 2017.

Grieves, Michael, Origins of the Digital Twin Concept. 2016. 10.13140/RG.2.2.26367.61609.

Lee, J., Azamfar, M., Singh, J. and Siahpour, S. , Integration of digital twin and deep learning in cyber-physical systems: towards smart manufacturing. IET Collab. Intell. Manuf., 2: 34-36, 2020.

Leng, J., Liu, Q., Ye, S., Jing, J., Wang, Y., Zhang, C., ... & Chen, X. , Digital twin-driven rapid reconfiguration of the automated manufacturing system via an open architecture model. Robotics and Computer-Integrated Manufacturing, 63, 101895, 2020.

Pasquale Franciosa, Mikhail Sokolov, Sumit Sinha, Tianzhu Sun, Dariusz Ceglarek, Deep learning enhanced digital twin for Closed-Loop In-Process quality improvement,

CIRP Annals, Volume 69, Issue 1, pp. 369-372, ISSN 0007-8506, 2020,

Qiyue Wang, Wenhua Jiao, YuMing Zhang, Deep learning-empowered digital twin for visualized weld joint growth monitoring and penetration control, Journal of Manufacturing Systems, Volume 57, , Pages 429-439, ISSN 0278-6125, 2020

Redelinghuys, A. J. H., Basson, A. H., & Kruger, K, A six-layer architecture for the digital twin: a manufacturing case study implementation. Journal of Intelligent Manufacturing, 1-20, 2019.

Reis, J., S. Pinelas, N. Melão and SpringerLink (Online service) Industrial Engineering and Operations Management II XXIV IJCIEOM, Lisbon, Portugal, July 18 2013;20. Springer Proceedings

Singh, Jagdeep, and Harwinder Singh. "Kaizen philosophy: a review of literature." IUP journal of operations management 8.2 , 51, 2009.

Tan, S., Hwang, J., & Ab-Samat, H., WITNESS simulation of preventive and corrective maintenance for Surface Mounted Technology (SMT) line. IOP Conference Series. Materials Science and Engineering, 505(1), 12047, 2019.

Varsamopoulos, Savvas & Bertels, Koen & Almudever, Carmen, Designing neural network based decoders for surface codes,2018. https://www.researchgate.net/publication/329362532

Y.H. Pan, T. Qu, N.Q. Wu, M. Khalgui, G.Q. Huang, Digital Twin Based Real-time Production Logistics Synchronization System in a Multi-level Computing Architecture, Journal of Manufacturing Systems, Volume 58, Part B, , Pages 246-260, 2021. ISSN 0278-6125, https://doi.org/10.1016/j.jmsy.2020.10.015.

Wang, J., Ma, Y., Zhang, L., Gao, R. X., & Wu, D., Deep learning for smart manufacturing: Methods and applications. Journal of Manufacturing Systems, 48, 144-156, 2018.

Zhang, Stenback and Wardrop , "Interval Estimation of the process capability index," Communications in Statistics: Theory and Methods, 19(21), 4455-4470, 1990.

Biographies

Eric C.H. Wai is currently the IT Head of VTech Communications Limited. With more than 18 years of EMS industrial experience, he leads more than 30 business transformation and industrial 4.0 projects in the enterprise. He is a certified Six Sigma Black Belt and qualified ISO 13485 auditor. He is pursuing an engineering doctoral degree and obtained his MSc. in Technology Management and Higher Diploma in Manufacturing from The Hong Kong Polytechnic University, BSc. in Applied Physics in City University of Hong Kong. His research interests include Machine Learning, Digital Twin, Cyber-Physical Production Systems (CPPS).

Dr LEE Ka Man is currently an associate professor in the Department of Industrial and Systems Engineering, The Hong Kong Polytechnic University, Hong Kong. She is the program leader of BSc(Hons) Enterprise Engineering with Management. She obtained her PhD and BEng degrees from The Hong Kong Polytechnic University. Her main research areas include Industrial Engineering, Enterprise Resource Planning (ERP), Logistics and Supply Chain Management, Industrial Internet of Things (IIoT), Wireless Sensor and Actuator Network (WSAN), Cloud Computing and Big Data Analytics. As of now, Dr Lee has published over 130 articles in various international journals and seminars. She was awarded Silver Medal in the 47th International Exhibition of Inventions of Geneva in 2019 and Outstanding Paper Award of Emerald Network Awards in 2019. Dr Lee also serves as the Lab-in-Charge of PolyU's Cyber-Physical Systems Laboratory

Abbreviations

AI	Artificial Intelligence
AOI	Automated optical inspection
APS	Advanced Planning System
COT	Changeover Time
DL	Deep Learning
DPPM	Defective Parts Per Million
DT	Digital Twin
IE	Industrial Engineering
LSL	Lower Spec Limit
LSTM	Long Short-Term Memory
M/C	Machine
ML	Machine Learning
OEE	Overall Equipment Effectiveness
OTD	On-time delivery
PCBA	Printed Circuit Board Assembly
RL	Reinforcement Learning
RTY	Rolled throughput yield
RNN	Recurrent Neural Networks
SCADA	Supervisory Control and Data Acquisition
SMT	Surface mounting Technology

T/P	Throughput
UPH	Unit per hour
UPL	Unit per Labor
USL	Upper Spec Limit