

SHL GROUP

APPLICATIONS OF DEEP LEARNING IN MEDICAL DEVICE MANUFACTURING

In this article, Frederick Gertz, PhD, Manager of Data and Process Innovation, and Gilbert Fluetsch, Director, Automation Systems, both of SHL, look at how deep learning can be leveraged in a medical device manufacturing environment.

As buzzwords go, few have had the effect that “deep learning” has had on so many different industries. When deep learning entered the industrial scene, there was much interest and success from companies in various industries. Technology companies such as Google, Microsoft and Apple have always had heavy investment in the area, while traditional pharmaceutical and healthcare companies such as AstraZeneca,¹ Novartis² and Pfizer³ have also drastically increased their spending in areas related to artificial intelligence (AI).

During the early days, the techniques, benefits and limitations of deep learning were less known to most industries and, in particular, presented various challenges to the medical device and manufacturing fields with a somewhat “black box” connotation.

WHAT IS DEEP LEARNING?

The first struggle for industries exploring the deep learning revolution is to understand what deep learning is and how different it is compared with more traditional tools, such as machine learning and computer vision.

Machine learning is a subset of AI research and applications. In the 1960s, many algorithms and techniques were developed specifically for AI research, with the ultimate goal of developing a general intelligence. Unfortunately, many of the techniques fell painfully short of providing that level of capability.⁴ However, several of the mathematical techniques were applied to a smaller subset of problems, such as handwriting recognition, image recognition, etc. – demonstrating great ability at

optimisation and prediction within these domains. Even though some of these tools showed less promise for the development of general AI, they exhibited great potential for many less-aspiring applications. From this, the area of machine learning grew steadily.

In addition, machine learning tools such as principal component analysis (PCA), support vector machines (SVMs) and neural networks were used for a variety of tasks, usually focused on the problem of classification⁵ and optimisation⁶ of existing systems. Engineers in this field focused much of their time on feature engineering, which is simply different representations of data. For example, if a person wanted to track manufacturing performance, they might decide to calculate the average of a process dimension on a weekly basis. This would smooth out fluctuations and make it easier to objectively compare and determine long-term trends in the manufacturing process. Data scientists refer to this simple act of averaging the data as the engineering of a feature.

While data scientists typically develop much more complicated mathematical transformations, this example illustrates an important aspect of the process that machine learning engineers undergo when analysing a data set and building models. In fact, after data wrangling (the cleaning and organising of data to prepare it for processing), feature engineering⁷ is one of the most time-consuming aspects of machine learning – and, in many cases, the most complicated.

So, how does deep learning differ from machine learning (Figure 1)? Deep learning



Dr Frederick Gertz
Manager of Data and
Process Innovation
E: frederick.gertz@shl-group.com



Gilbert Fluetsch
Director, Automation Systems
T: +41 41 368 00 00
E: gilbert.fluetsch@shl-group.com

SHL Medical AG
Gubelstrasse 22
6300 Zug
Switzerland

www.shl.group

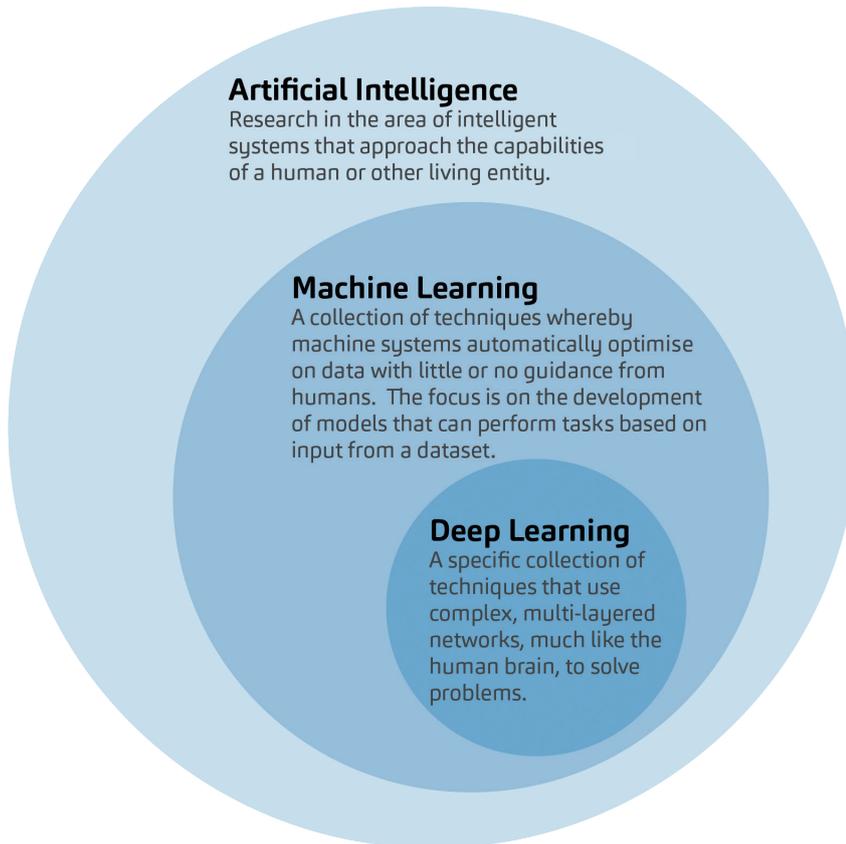


Figure 1: The relationship between AI, machine learning and deep learning.

takes some of the machine learning tools, such as neural networks, and expands their size, allowing them to learn iteratively without the need for “handcrafting” of features. Most importantly, for many applications, these networks require little or no feature engineering, learning on their own from “raw” inputs.

As with many scientific discoveries, the technique quickly expanded and the combination of much larger networks with extensive optimisation, particularly using graphical processing units (GPUs), led to the wider adoption of deep learning. In essence, deep learning can be characterised by two things – larger networks with more nodes,⁸ as well as less emphasis on the need for humans to perform feature engineering.

APPLICATION IN MANUFACTURING ENVIRONMENT

Machine builders in the automation field of assembly equipment have been focusing for years on the mechanics of machines.

Indexing, cam-driven mechanisms, pick and place, and mechanical grippers, etc. have been the primary description of the mechanical functions and processes of automated equipment. Project managers and mechanical engineers have been the driving forces behind these projects, often deciding how much control would be integrated into the equipment. Electromechanical components became available to increase the accuracy and also the precision of an assembly step.

For example, a standard assembly process from the past would be to use a pneumatic cylinder that would drive to a given distance. In the event of interference between two parts, the assembly process continues unless the force of the interference becomes higher than the force of the cylinder. But one could not control or monitor the assembly process. Today, servo motors with load cells are often implemented to achieve precise control of the assembly force, while also enabling engineers to get a read-out of force curves if needed.

This process is known as control engineering, a discipline that has gained much attention compared with just a few years ago. Although the mechanical process of an equipment is still important, more electronics and monitoring of assembly processes result in an increased need for control engineering talents in the development of said equipment. The integration of such different controls generates a wealth of data which can be leveraged by machine builders.

As customers and manufacturing operations push for increased cost-effectiveness and flexibility in manufacturing systems, these same requirements are also pushed onto the control engineers. The additional burden of these new requirements places a strain on all aspects of automation machine design and manufacture – but a particular strain is placed on control engineers, who are now turning to new fields, such as deep learning, to find faster, more flexible means of implementation.

DEEP LEARNING AT SHL

At SHL, we have been working with various machine learning techniques and exploring ways that they can be incorporated into our workflows and used to optimise our processes. In recent collaborations with our Automation Systems department, we repurposed a wealth of data and experience collected over many years of automating the manufacturing process. The Automation Department is a natural partner for the Data and Process Innovation team due to its mature implementations of robotics and computer vision systems, which have allowed for robust deployments and rich datasets.

In recent years, cameras and visual inspection systems have gained popularity as technology evolved. The main challenge has been, and still is, not to define a component or measurement as good or bad – but to identify criteria that lie between pass and fail. Historically, one of the most important components of an inspection system with cameras in an equipment was the lighting – direct, indirect, ring lights, etc. Light sources which “burn” out after a while resulted in an inspection that was no longer accurate.

The implementation of LEDs gave support in overcoming these challenges, while different LED colours were used for different applications. The colour of the LEDs can now be customised based on the

“The main challenge has been, and still is, not to define a component or measurement as good or bad – but to identify a criterion that lies between pass and fail.”

object and the inspection environment – and they therefore represent the ideal light source for the inspection project. Now, those LEDs are able to point out distinctive features on a component, such as angles or specific contours.

Nonetheless, the above does not account for the fact that complicated algorithms, usually hundreds or thousands of lines of code, had to be programmed for an inspection of a component, a measurement of a critical dimension or any other application. The code would often rely on pixel counts of the camera, which in turn was dependent on the resolution of the camera as well as the above-mentioned light source. Hundreds of hours of testing would follow the coding process and, even then, the traditional inspection systems would not provide solutions that could be trusted 100%.

There is an increased push for higher volumes of output with lower investments, necessitating the need to transition from manual assembly to semi- and fully automated assembly. These transitions, especially for products not originally designed for automated assembly, require a rapid collection of data and development of numerous control systems to provide accurate and high-quality assembly of parts. Frequently, this transition can lead to unexpected challenges as well as the need for additional design and inspection requirements to be incorporated into the production process.

One of the focuses of the Automation Department is the development and manufacture of large-scale, fully automated assembly machines. The machines provide SHL with the ability to drastically increase output for high-volume projects while also increasing consistency in the assembly process. At SHL, these machines frequently incorporate multiple vision inspection stations, all of which are enriched with unique data sets from our assembly and testing processes. The Data and Process Innovation (DPI) department investigated the use of deep learning in conjunction with some of these computer vision stations. Below we share an example and some results from one of our studies in the use

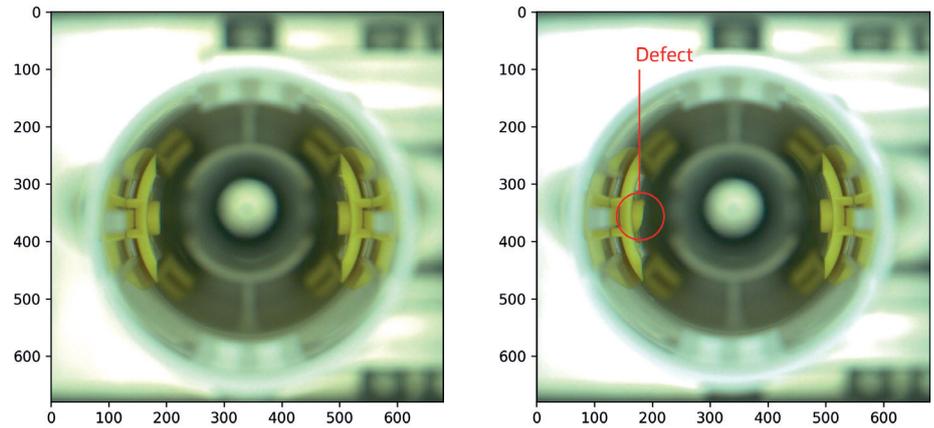


Figure 2: Two images showing a portion of a device assembly. This assembly is made from two parts, an outer white shell and a yellow insert. The left image shows an assembly that has been classified as good and the right image is an assembly that is damaged.

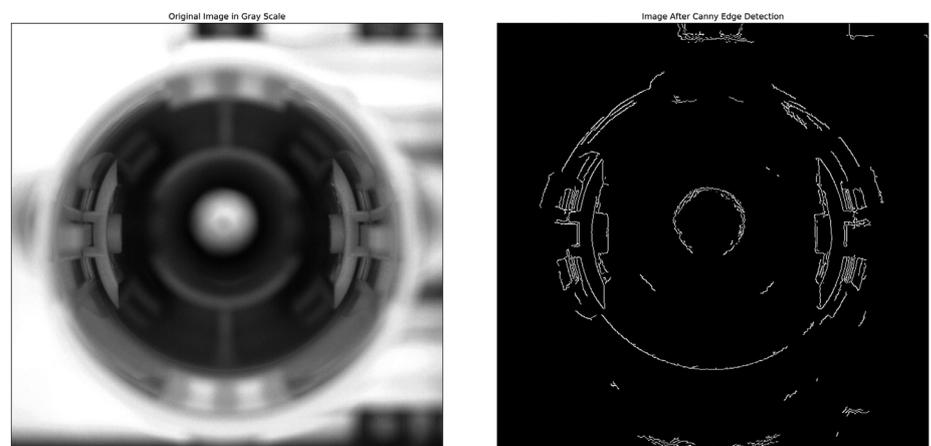


Figure 3: An example of an image after being processed with canny edge detection.

of deep learning for rapid vision inspection development and deployment.

A data set of images was provided to the DPI team for processing. Images were classified by human operators as either being good or bad (with or without defects). This assembly contains a front shell assembly (coloured white) and a yellow internal piece. As the piece is a functional part of the assembly, it is important that the moulded parts both fit together correctly and are free of any moulding or handling defects.

Figure 2 is an example of a good assembly (left) as well as an image of a bad assembly (right), with an annotation in the image to show an example of the defect. It should be noted that this is not the only type of defect; almost an endless possibility

of defects could occur in the production and assembly of components. As such, multiple examples of different defects are required for training.

In a traditional computer vision set-up, the important areas of the images would first be determined through a combination of input from functional specifications and discussions with domain experts. Figure 3 shows the results of a canny edge detection, which would normally be used to reduce the data input into the computer vision algorithm. From this point, an engineer would define the areas of greatest interest and create rules which would be used to determine which parts are within specification and which parts contain defects.

It should be noted that during this process, a certain amount of interpretation is required from the vision engineer. Depending on the part, the process can be quite time consuming, requiring multiple iterations and several validations and revalidations. The technique, like all computer vision techniques, is sensitive to environmental changes and can pose a

“There is an increased push for higher volumes of output with lower investments, necessitating the need to transition from manual assembly to semi- and fully automated assembly.”

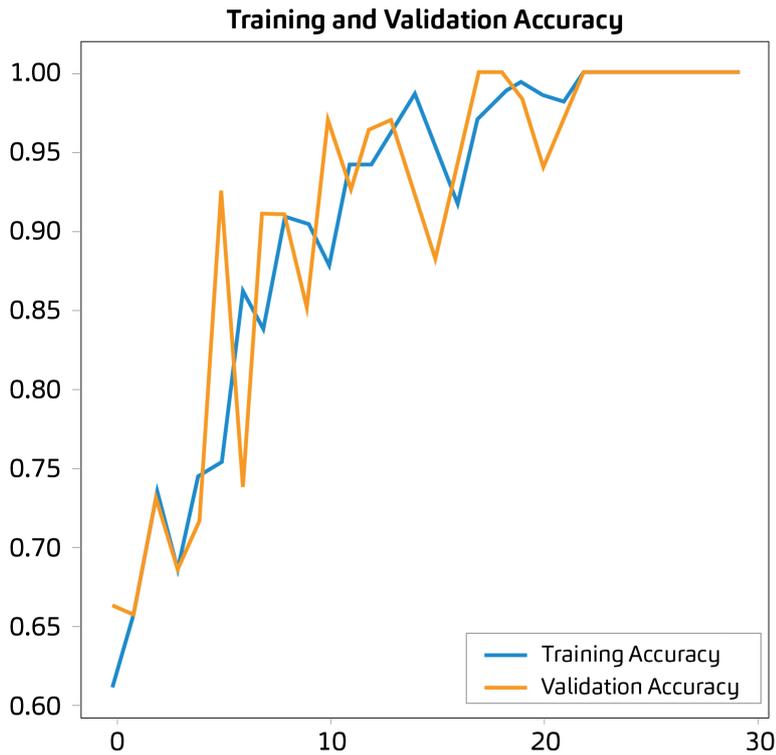


Figure 4: Graphical results showing the evolution of the TensorFlow convolution neural network over 30 iterations. The network uses two sets of files – a training set which the network learns on (blue line) and a validation set which the network tests itself against.

the algorithm was able to provide a very high degree of confidence in its ability to segregate good images from bad.

This technique undoubtedly scales up quite nicely, with new defects easily added to the data set, which can even include difficult-to-define cosmetic defects as well. The ability of the deep learning algorithm is entirely dependent on the input data it receives and, as such, the most time-consuming aspect for some operations would be the data collection and labelling. As a long-time manufacturer and user of automation equipment, SHL has an abundance of validation and production data on hand, which gives us the unique ability to leverage the development of data-intense models such as those used for deep learning algorithms.

LEARNING FROM DEEP LEARNING IMPLEMENTATIONS

This sort of implementation is far from cutting edge, and the example shown does not include the multiple additional steps required to help train for robustness and test the network. However, once the model is developed, production and validation will need relatively strong assurance that it is performing at a high level. At this point, some interpretation of the model becomes very desirable.

Despite the reputation of deep learning models, as mentioned previously, there are avenues that can be used to gain insight into the performance of the model.¹⁰ For example, one common technique used by deep learning engineers who work with images is to have heat-map overlays produced on the images (Figure 5). In these examples, a heat map is created to show regions that the neural network has, through its training, identified as important.

“The ability of the deep learning algorithm is entirely dependent on the input data it receives and, as such, the most time-consuming aspect for some operations would be the data collection and labelling.”

significant risk to the timeline in delivering automated equipment to the end users in manufacturing.

In comparison, we now show the results of the deep learning implementation. The deep learning implementation uses a convolutional neural network through the TensorFlow framework⁹ – an open source, readily available tool for deep learning. The convolutional neural network was defined with the following basic parameters: a convolutional network with nine layers, including three convolution layers and three pooling layers. This network has a total of 236,773,409 parameters, which are all trainable. After defining the network, the images are provided to the network for training, without the need for any transformation like the above-mentioned canny edge detection. Using an unimpressive amount of computational resources, including a readily available

laptop, the training over a data set of about 240 images took approximately 30 minutes.

According to the results (Figure 4), it is apparent that the network was able to learn quite readily

on the images and reach 100% accuracy on the training data set. A test set of images was prepared to check for overtraining and, similarly, 100% accuracy was reached on the test set. Thus, with 30 minutes of work and a relatively small amount of data,

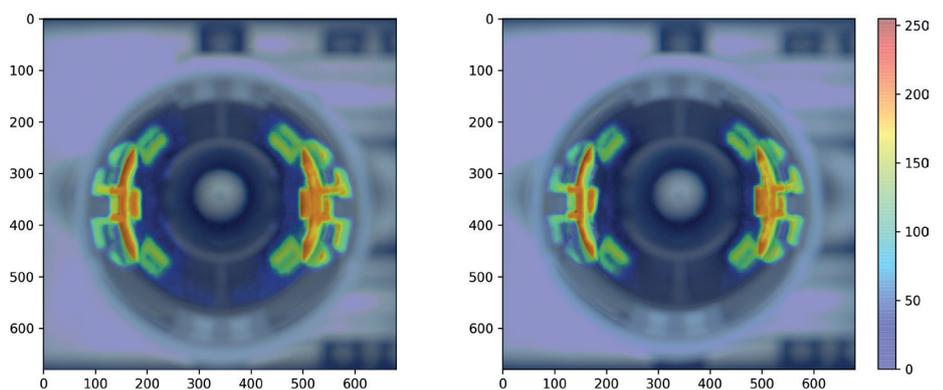


Figure 5: Images from Figure 2 are overlaid with a heat map where the neural network identifies areas of interest based on its training.

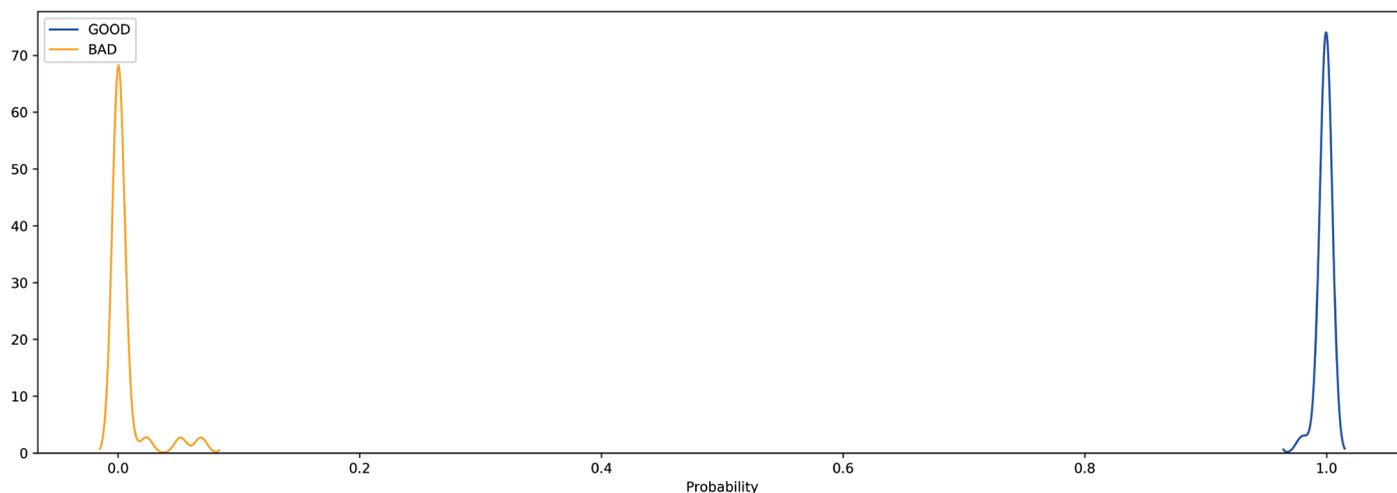


Figure 6: The distribution of good and bad samples, based on results from the neural network. Good samples (coloured blue) show a clear distribution completely opposite to the distribution of bad samples (orange colour).

In the images below, it is clear that the network has correctly defined the areas of importance. It is able to identify that the yellow plastic insert is the region of interest and even shows that the edges of the yellow insert are of even greater importance, performing what can also be viewed as a type of edge detection.

Images such as these can give production engineers a greater deal of confidence that the regions of interest correctly overlap with the areas the production or control engineers would have selected as being the most important. Furthermore, the network does not give a purely binary output. In fact, we can ask the system to provide a continuous probability outcome. In this case, the system assumes that a perfect sample would be scored a one and a bad sample would be scored a zero. We can obtain the distribution results shown in Figure 6, which represent the results for the validation data from the final iteration of the TensorFlow model's training.

We can see from these results that there is a distribution with a strong centre for good samples (blue) at about 0.99 and a strong peak for bad samples (orange) at about 0.02. If we were to evaluate this system as we would many other measurement systems, we could perform a typical statistical analysis and would conclude that even the edge cases of this measurement are quite far removed from the classification boundary of 0.5.

Herein lies another advantage. The 0.5 limit is somewhat arbitrary. It assumes that if the system classifies the part as more likely in one category from another, then that is where it places the classification tag. However, in production, especially risk-averse production such as medical

“We have the ability to not only test and validate the system similar to any other measurement system but also to move criteria based on our own risk assessment and requirements.”

devices, we can tighten the criteria to 0.8 or even 0.9, telling the system that it must be absolutely certain in its classification; otherwise it should reject the part. With a stricter criterion, we have the ability to not only test and validate the system similar to any other measurement system, but also to move criteria based on our own risk assessment and requirements.

CONCLUSION

SHL continues to advance its understanding of the development and validation of these techniques as we explore their deployment into our production processes. Real-world examples derived from internal SHL investigations demonstrate how deep learning can be leveraged in a medical device manufacturing environment. The findings are quite surprising. First, the amount of training data is smaller than might be thought from other deep learning applications.⁸ We assume this is because of the high consistency that is required in the manufacturing environment, resulting in faster learning of the underlying data distributions. Second, the interpretability, the ability of the model to relay parts of its process to engineers, is quite a bit higher than is expected from a true black box.

With these learnings in mind, and with the speed and robustness that these types of techniques offer, SHL expects to see

further adoption of deep learning models in manufacturing, both internally and amongst the medical device manufacturing community as a whole.

ABOUT THE COMPANY

SHL Group is a world-leading solutions provider in the design, development and manufacturing of advanced drug delivery devices such as autoinjectors, pen injectors and advanced inhaler systems. It offers a full range of in-house core competencies and services in the fields of medtech and patient care. With >4,000 employees worldwide, SHL Group consists of several distinct group companies: SHL Medical designs, develops and manufactures advanced drug delivery devices for leading pharma and biotech companies across the globe; SHL Healthcare develops and manufactures equipment solutions for home, hospital and long-term care use; and SHL Technologies provides contract manufacturing and engineering services for the production of complex medtech products.

REFERENCES

1. Chen H, Engkvist O, Wang Y, Olivecrona M, Blaschke T, “The rise of deep learning in drug discovery”. *Drug Discov Today*, 2018, Vol 23, pp1241–1250.

2. "Machine learning poised to accelerate drug discovery". Novartis, May 7, 2018.
3. "IBM and Pfizer to Accelerate Immuno-oncology Research with Watson for Drug Discovery". Press Release, Pfizer, December 1, 2016.
4. Pennachin C, Goertzel B, "Contemporary Approaches to Artificial General Intelligence". *Cogn Technol*, 2007, Vol 8, pp 1–30.
5. Pham D, Afify A, "Machine-learning techniques and their applications in manufacturing". *Proc Inst Mech Eng Part B, J Eng Manufacture*, 2005, Vol 219(5), pp 395–412.
6. Wuest T, Weimer D, Irgens C, Thoben K-D, "Machine learning in manufacturing: advantages, challenges, and applications". *Prod Manuf Res*, 2016, Vol 4, pp 23–45.
7. Khurana U, Samulowitz H, Turaga D, "Feature Engineering for Predictive Modeling using Reinforcement Learning". *32nd AAAI Conference on Artificial Intelligence*, 2017.
8. Goodfellow I, Bengio Y, Courville A, Bengio Y, "Deep Learning". MIT Press, 2016.
9. Abadi M et al, "TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems". *arXiv*, 1603.04467v2, March 2016.
10. Chakraborty S et al, "Interpretability of Deep Learning Models: A Survey of Results". *Proc IEEE SmartWorld/SCALCOM/UIC/ATC/CBDCom/IOP/SCI Conference*, 2017, pp 1–6.

ABOUT THE AUTHORS

Frederick Gertz is the Manager of Data and Process Innovation at SHL Medical. His focus in the company is on facilitating data-driven methods across the organisation and providing unique insights from data using a variety of techniques, including artificial intelligence and deep learning. Prior to SHL, Dr Gertz worked in the medical device start-up space where he focused on bringing novel processes and techniques, including machine learning, into the biotech industry. He holds a PhD in Electrical Engineering from the University of California, Riverside (US) where his research focused on biophysics and spintronics.

Gilbert Fluetsch joined SHL's Automation Systems Department in early 2016. His responsibilities include leading the engineering teams, standardising the existing equipment portfolio and overseeing the development of high-speed assembly and testing machines. Prior to joining SHL, Mr Fluetsch served in various leading roles in engineering, operations and sales management in the medical device and semiconductor industries for almost three decades. He has an MBA in High Technology Management from the University of Phoenix (AZ, US) and a BS in Business Administration from California State University San Marcos (US).

Drug Delivery Partnerships

VIRTUAL

September 21 -24, 2020

Delivered Digitally | US Eastern Daylight Time

Proudly part of **BIOPHARM AMERICA™** **DIGITAL»»**

With our dedication to address the most pressing issues in our current global climate, DDP has undergone a timely transformation, moving to an all-digital format and offering attendees the ability to participate in virtual content, exhibits and networking opportunities. Plus, gain access to on-demand content to view at your own leisure.

Sit in on live keynotes and panels featuring industry experts, including:



USP 800 - How it Will Affect Drug Delivery and Formulations
Lana Shiu, MD
 Executive Director, Medical Device Regulatory Affairs, Global Regulatory Affairs and Safety, Amgen



Pandemic Prevention through Gene Therapy Delivery
Shawn Davis
 Senior Director, Head of Drug Delivery AstraZeneca



Brain Delivery of Biologics – VIA Cerebrospinal Fluid
Shradha Sadekar, Ph.D.
 Scientist, PKPD, Genentech



Patient-Focused Drug Delivery - The View from a New Player
Dr. Vlad Coric, MD,
 CEO and Director, biohaven pharmaceuticals

View the Agenda to see all scheduled live content, virtual networking and on-demand content!

<https://informaconnect.com/biopharm-america/drug-delivery-partnerships/>

The SHL Perspective

Empowering Innovation

SHL's Molly[®] autoinjector is designed to address various fill volumes of complex biologics.

Built upon a proven preconfigured technology, Molly[®] aims to support shortened development timelines for pharmaceutical companies.

