AI-DRIVEN COMPUTER VISION

A DHL perspective on the impact of Computer Vision in logistics

Powered by
DHL Trend Research

CONTENTS

Katja Busch

Chief Commercial Officer DHL &
Head of DHL Customer Solutions and Innovation

With advancements in artificial intelligence (AI), computer vision is at a stage to become an industry-shaping technology and has exceptional promise along the supply chain – for our customers, employees, partners, and certainly the environment.

We at DHL are maintaining our commitment to shaping the Era of Logistics by bringing real-world innovations to the logistics ecosystem, leveraging AI and computer vision technology.

Dr. Klaus Dohrmann

Vice President, Head of Innovation and Trend Research

For good reason, AI is one of today's most hyped topics and it comes in several different shapes and sizes. Arguably the biggest headline-grabbers right now are generative AI tools such as ChatGPT – technology that lets us have human-like conversations with a chatbot. However, another highly significant development in the AI space is AI-driven computer vision, a technology that is already deployed in a range of applications at an increasingly stable and dependable level.

This DHL trend report on AI-driven computer vision in logistics delves into the dynamic intersection of computer vision, artificial intelligence, and logistics, emerging as a compelling arena of transformation. We think there has never been a more exciting time for industries and logisticians to work together to leverage the full potential of computer vision and AI for the benefit of organizations, our colleagues in operations, and for improvements in environmental sustainability.

But, as also outlined in the report, the integration of computer vision into logistics comes with challenges. As with any technological leap, there are considerations of data security, ethical implications, and the need for upskilling the workforce. The convergence of human expertise and AI augmentation requires thoughtful orchestration and collaboration – from all of us! - and we do hope that this report will contribute to this.

As we navigate the terrain of computer vision in logistics through this report, we invite you to explore the depths of this transformative trend. Whether you're a logistics professional, a technology enthusiast, or an advocate for sustainable supply chains, this report offers insights into how computer vision is not only reshaping logistics but also propelling us toward a new era of interconnectedness and efficiency.

By working closely with our customers, jointly developing solutions and copiloting proof-of-concept projects in computer vision, we at DHL are staying ahead of the game.

We believe in innovation beyond potential – there is always a better way to operate, plan, implement, connect, and share. As we seek to improve our own logistics capabilities and those of our customers, we constantly look for fresh approaches and valuable new technologies.

That's why we engage with the brightest tech innovators and disruptors around the world. If a technology development or application can contribute to a better customer experience, higher customer satisfaction levels, improved efficiency, and more sustainable operations, we're interested! You're welcome to explore many opportunity areas in our recently launched virtual 'Warehouse of Innovation.'

We hope this trend report will inspire and guide you and we look forward to collaborating with you in this exciting and potentially transformative field of computer vision in logistics, powered by artificial intelligence.

Computer Vision Looks Ahead

AI-powered computer vision technology appears to have moved through the Gartner Hype Cycle for AI to now enter the Plateau of Productivity. Companies are showing more and more interest in computer vision, and an increasing number of technology providers are ready and able to supply this demand. The global computer vision market is building steadily, with researchers predicting growth from USD $9.40 billion in 2020 to $41.11 billion in 2030, a decade of CAGR at 16%.

As computer vision proves its worth in specific applications around the world, it already looks set to enable many sectors. This trend report highlights the use of computer vision in many areas of logistics, from dimensioning and safety to route optimization and demand prediction. At the same time, it presents application examples from other industries – retail, healthcare, disaster response and recovery, and manufacturing illustrate the enormous potential of this technology in the supply chain.

The DHL Logistics Trend Radar identifies computer vision as a trend that will become part of the standard way of operating in the logistics industry within the next five years, underpinning and driving future logistics successes by enabling more automated and efficient processes as well as sustainable and safe operations.

What is Computer Vision?

AI enables computers to “think” and computer vision allows computers to “see and understand.” Computer vision systems gather information from visual inputs like digital images and videos. By collecting and crunching this visual data using algorithms, these systems can then make recommendations and even take actions.

Since birth, every sighted person has been learning how to tell objects apart, estimate object distance and speed, spot visual anomalies, and interpret what we see. This is the basis of AI-powered computer vision as well.

Computer vision systems, specifically their algorithms, must be trained in the same way, and this is done using visual data. The training process is accelerated by providing vast amounts of digital input. These systems never get tired and can quickly exceed our human capabilities of detecting and reacting to visual inputs. Computer vision accuracy rates for identifying and classifying objects increased from 50% to 99% in less than a decade.

Current Impact on Logistics

AI is already impacting the logistics industry, from chatbots to route optimization and demand prediction. And now, computer vision looks likely to unlock many more opportunities, thanks to technology advances in depth perception, 3D reconstruction, and interpretation of dark and blurred images.

How Computer Vision Systems Learn

Computer vision systems learn by looking at vast quantities of high-quality visual data. They repeatedly analyze this data until they recognize images and learn about any image differences. How is that done? Using two different technologies:

- A type of machine learning called deep learning – this uses algorithms for self-teaching about visual data along with artificial neural networks to find out more and more from the data
- A convolutional neural network (CNN), which breaks images down into tagged labels and performs the math on these labels to repeatedly check prediction accuracy

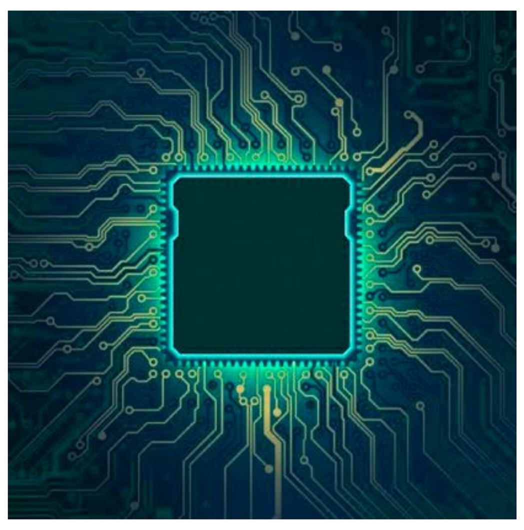
Computer vision is likely to soon unlock many more opportunities, thanks to technology advances in depth perception, 3D reconstruction, and interpretation of dark and blurred images. It's clear that deep learning has moved from the conceptual realm to practical application as many computer vision applications, from facial recognition to self-driving vehicles, make use of it.

Trends Linked to Computer Vision

A wide range of technology trends are linked to computer vision. Here are some key examples from the DHL Logistics Trend Radar.

Interactive AI
This refers to using algorithms that process human user input, like text and speech, to provide a reasonable response.

Read more here



Edge Computing
Featuring decentralized IT architecture, this trend allows the processing of high-quality visual data from the cameras an sensors – at the edge of a network at high speed while keeping the information safe at the source.

Read more here

Digital Twins
When integrated into digital twins, computer vision allows for remote monitoring of physical objects. It can autonomously identify flaws or deviations and promptly initiate corrective actions.

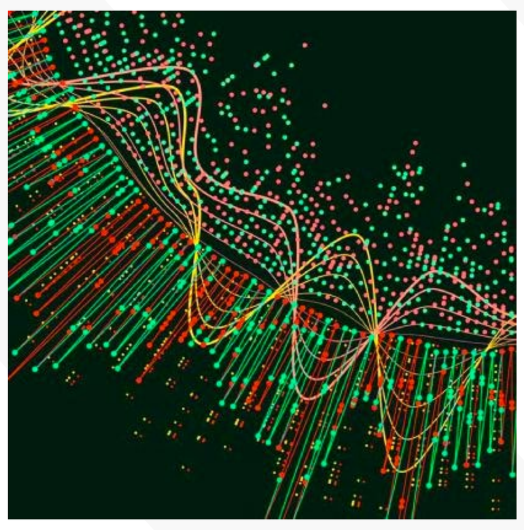
Read more here

Mixed Reality
Computer vision extracts data from images and videos. Mixed reality integrates it into the physical world by creating 3D overlays, providing guidance for many tasks like advanced inspections or complex surgeries.

Read more here

Drones
By implementing deep neural networks, cameras mounted on drones can be trained to detect people and objects. Subsequently, they can analyze the images and communicate the findings in real-time.

Read more here

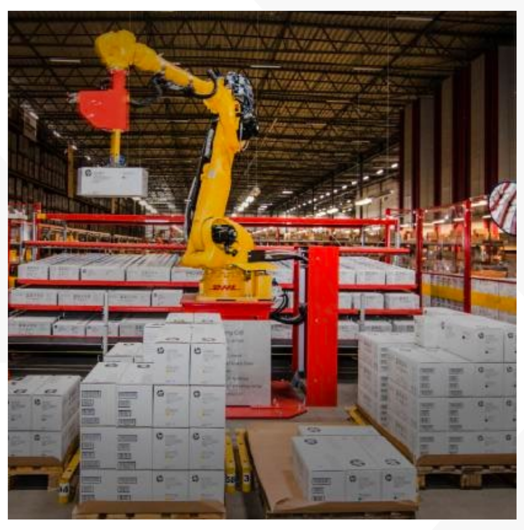


Big Data Analytics
This trend involves analyzing large data to find patterns, track real-time changes and forecast the future. In computer vision, it accelerates processes, enhances productivity, etc.

Read more here

Outdoor Autonomous Vehicles
Computer vision is central to this technology as cameras and sensors combined with object detection algorithms help these vehicles avoid collisions, follow designated routes, and detect obstructions.

Read more here



Robotics
Vision-based simultaneous mapping and localization enable robots to perceive, understand, and react to changes in their surroundings. Applications include plotting routes, mapping unmapped areas and avoiding obstacles.

Read more here

How Computer Vision Creates Value

Today's computer vision systems are deployed in various ways. The most well-known application is image classification. The system sees an image and predicts it belongs to a certain class (e.g., a human, a pair of protective goggles, a forklift).

Another familiar application is object detection, also known as machine vision. The system not only classifies an image but also takes note of (tabulates) its appearance. Once an object has been detected, it can be tracked – object tracking is often done using sequential images and video feeds.

A further application for computer vision systems is content-based image retrieval, to increase the accuracy of search and retrieval of digital images.

Computer vision images are subjected to various processes including:

Image Segmentation
Partitioning into multiple segments to simplify or change the representation into something that is meaningful and easier to analyze.

Blob Checking
Looking for discrete spots of connected pixels as image landmarks; blobs often represent optical targets for observation, robotic capture, or manufacturing checks.

Pattern Recognition
Algorithm-based template matching to find patterns using machine-learning methods.

Image Processing
Stitching, filtering and pixel counting.

Challenges in Applying Computer Vision

Focus

The computer vision model must get highly specific training on a clearly defined problem to solve.

Data Quality

Training models need vast amounts of visual data, and this must be of high quality.

Model Selection

Each computer vision system must use the right model and modeling techniques. Off-the-shelf is not feasible.

User Adoption

To deliver value, the solution must be accepted by all users.

Cybersecurity

Malicious data manipulation can skew analyses and alter AI performance.

Balance

With so much visual data to store, process, analyze, and maintain, companies must balance cost and accuracy.

Best Practice

Machine-learning operations (MLOps) and DataOps best practice is essential, especially controlling data use (versioning).

Privacy

To protect employees and boost acceptance, incorporate privacy measures and comply with GDPR and other laws at outset.

Investment

Costs can include camera upgrades, new tech investment, ongoing platform maintenance, and workforce upskilling and reskilling.

In the next chapter...

Already today, computer vision is proving its worth in a vast range of applications. In the next chapter we explore key solutions that are deployed in retail, healthcare, disaster response and recovery, and manufacturing. With these real-world examples, we are better able to envisage the potential for computer vision applications in logistics and the supply chain.

NON-LOGISTICS USE CASES: RETAIL

Computer Vision in Retail

Retailers are turning to computer vision technology for many reasons: The right solutions can help identify customer movements, purchasing patterns, and busy areas within a store. Also, retailers can leverage computer vision to inform merchandizing, improve store layout, and staff positioning. Computer vision can also help to improve customers' in-store experience, securing their loyalty.

Current Trends

Key trends for online and brick-and-mortar retailers include sustainability and circularity to reduce waste, develop eco-friendly practices, and minimize environmental impact. There are also technologies such as augmented reality and virtual reality which transform how customers interact with products and brands, enhancing virtual and in-store experiences. And other important trends include artificial intelligence (AI) and automation to optimize retail operations, particularly in demand forecasting, inventory management, and delivering personalized customer services.

Long lines in stores can make would-be customers abandon purchases, with an estimated $19 billion lost in sales last year for this reason alone.

Computer Vision Retail Applications

Cashierless Checkout

Checkout technology hasn't changed much in the past two decades but today computer vision technology is capable of fully automating the checkout process to improve the customer experience. A store's camera system captures images of each selected product, and an algorithm is trained to identify the object and its price. The system also associates the checkout transaction with a specific shopper in order to charge the correct person the correct amount. And this system helps prevent theft by observing customer behavior and analyzing patterns, triggering human intervention to assess risk when needed.

Grocery retailer Aldi claims "Smart shopping is the future," piloting its Shop&Go checkout-free experience in selected stores. Many 'just-walk-out' Amazon Go and Amazon Fresh convenience stores have opened over the past seven years. And as more retailers proactively invest in technology to fight the rise in retail crime, Amazon is also selling its solution to other companies. Meanwhile, Australian software startup Tiliter promises an AI-powered solution to protect retailer profits and deliver accurate product recognition with a 5x faster shopping experience.

Inventory Management

On-shelf availability is critical to brick-and-mortar retailers. If shoppers can't find what they're looking for on display, they'll probably select a substitute, but as many as two out of three could leave empty-handed and may buy from a competitor instead. Computer vision technology integrated with the inventory management system can monitor shelves to keep track of inventory levels, detect when items are running out and issue replenishment alerts, and trigger automated reordering and restocking processes.

An AI-powered solution can automatically update brand packaging when it changes by working with synthetic visual data (information that's artificially generated, not produced by real-world events), referencing its own training images and any new product images.

The Californian robotics startup Simbe has developed an innovative autonomous inventory robot featuring inbuilt 3D cameras, enabled by computer vision, that navigates and analyzes the retail store environment for low inventory levels, providing alerts when shelf replenishment is required.

With its ReShelf solution, French retailer Auchan turns everyday shelf images captured by ceiling-mounted cameras into real-time, actionable insights. In-store staff are notified when a product runs low, helping to maintain stock levels. The system can also provide out-of-stock analysis of the entire product supply chain.

Category Management

Retailers create in-store merchandise display schematics or planograms to achieve the required outcomes – more sales, improved category profitability, positive customer experiences, and more. Compliance with these plans means the right product with the right price tag has the right amount of space in the right location and is optimally promoted.

Computer vision helps to understand shopper behavior patterns and this real-time visual data enables the management team to optimize in-store floor layouts and workflow to optimize customer traffic. This technology also contributes sales data for reports that strengthen the negotiating position with suppliers.

Headquartered in Switzerland, Viso offers vision-based footfall analysis, pass-by traffic, and other customer analytics to detect in-store movements and visual data patterns so retailers can clearly see which promotions engage customers and which do not.

Computer Vision in Healthcare

In the life sciences and healthcare sector, computer vision can improve diagnostics, medical treatments, and procedures while also accelerating healthcare research. The technology can give patients an improved overall experience and provide medical professionals with the real-time visual data they need to make better decisions while ensuring safety and cost efficiency.

Current Trends

Examining samples of body tissue in a laboratory is a cornerstone of modern medicine, ensuring correct diagnosis and treatment. But a shortage of pathologists – exacerbated in the post-pandemic era – and rising biopsy volumes are challenging. Meanwhile, advancement in cell and gene therapy has created opportunities for highly personalized patient medication. Also on the increase is remote monitoring of patient condition and compliance with treatment regimes, enabled by digital healthcare and the new healthcare delivery paradigm. Patients as consumers are more proactively involved in their own healthcare decision making and they use more digital and wearable devices to collect information about their health status.

Implementing digital pathology solutions resulted in 21% increase in diagnostic capacity compared to microscopy.

Computer Vision Healthcare Applications

Medical Image Reading and Reporting

Accuracy, speed, and fewer errors. These are the key benefits of computer vision in recognizing patterns to inform diagnoses and improve patient outcomes. Many healthcare facilities now use this AI-based technology to extract and analyze vital visual information from MRI and CT scans, X-rays, ultrasounds, and other medical images.

Computer vision is used mostly for image interpretation and reporting, while enabling health system interoperability, efficient data management, and cloud access to shared visual data.

The overall number of intelligent screening solutions to support diagnosis is growing fast. One example is the AI-powered smart robotic microscope from Indian startup SigTuple. By analyzing visual medical data, this device removes drudgery from routine microscopy, allowing pathologists to spend time where it matters most – with critical patients.

Lithuanian startup Oxipit offers an autonomous AI computer vision tool identifying healthy chest X-rays (those without any abnormality) and producing finalized patient reports without requiring any human intervention at all. This frees up the radiologist's time to focus on more pressing cases.

Early Detection of Diseases

Doctors use their experience and intuition to identify disease, often using tests to verify their diagnoses. But traditional tests don't always reveal the problem, particularly early indicators of disease. Computer vision provides the support that medical professionals require. Systems can go beyond identifying anomalies to also detect disease; they are trained, by learning from images of healthy and diseased tissue types, and algorithms can recognize patterns and notify doctors.

A study from researchers at Google Health in collaboration with Imperial College of London designed an algorithm that shows this technology is more efficient and accurate in diagnosing breast cancer than human radiologists.

Clinicians can send their medical images to the lab of New York startup PreciseDX to receive a detailed report with AI-powered insights within 2-3 days, using this additional information to inform and enhance their usual treatment plans.

The AI-based diagnostic software startup Paige has developed a solution to detect the spread of breast cancer to the lymphatic system – a spread that is most at risk of being missed and a task that's tedious and time-consuming for the pathologist but critical to patient health.

Accurately and swiftly analyzing prostate biopsies is challenging due to increasing case volumes, subjectivity in grades, the small size of certain tumors, the large number of samples per case, and more. The Israel-based medical technology manufacturer Ibex has developed an AI-powered, computer vision solution to help pathologists in improving the detection and grading of prostate cancer.

Enhanced Medical Procedure Efficiency

When performing complex surgeries, medics need to make critical decisions that can profoundly affect the patient. Computer vision can be used to boost surgical success rates. For example, a computer vision system can automate the recording of a surgical procedure and then superimpose this virtual video footage and still images on the surgeon's view in real time as this person undertakes the same type of operation. This can guide and train surgeons as they work and can also be used to support them with a variety of repetitive and typically error-prone processes.

Promising "surgery at your fingertips," health technology startup Touch Surgery, part of Medtronic, has developed an AI-powered surgical video and analytics platform for the operating room. Viewing procedures on a monitor and even on a smartphone, this solution makes it easy to prepare, practice, and teach more than 200 surgical procedures across 17 different specialties and explore new techniques.

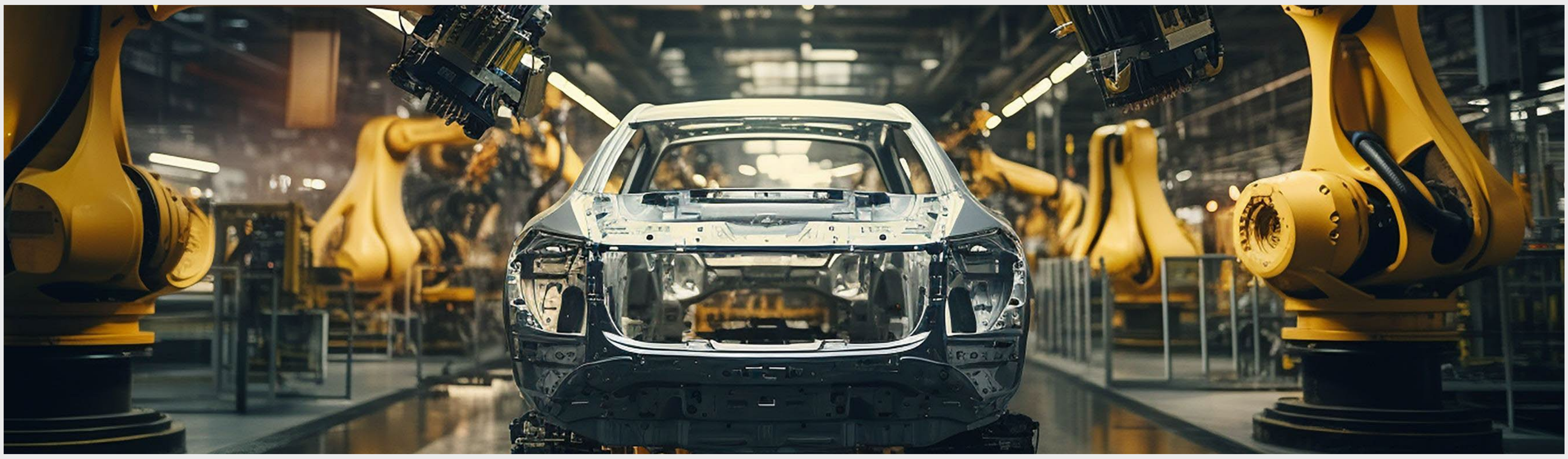
NON-LOGISTICS USE CASES: MANUFACTURING

Computer Vision in Manufacturing

Computer vision can help maintain quality standards in manufacturing, improve equipment monitoring and assessment, boost process efficiency, and support facility surveillance and security. Companies are using this technology to automate tasks that were once done manually so that items come off the production line at the optimal rate and with consistent quality. Computer vision enables manufacturers to inspect even the smallest product details, tracking for damages and faults to reduce the likelihood of error. And it can contribute to workflow optimization, maximizing operational efficiency and minimizing unplanned downtime.

Current Trends

The key priorities of manufacturers have not changed much over time and continue to include sustainability and optimizing the supply chain. Many companies today are assessing the benefits of decentralized manufacturing. They are also exploring digital solutions including AI and computer vision technology.



In 2020, safety concerns regarding door latching mechanisms promoted a global recall of 13 million vehicles from a leading car manufacturer.

Computer Vision Manufacturing Applications

Quality Inspection

One of the most important ways that computer vision is already used in manufacturing today is automating quality checks during production. In the past these types of inspection were done manually by quality control experts, but even the most skilled inspector could make mistakes.

For higher accuracy in quality control and inspection, many manufacturers are choosing to use computer vision. When this is linked with deep learning, the quality inspection system can be trained and retrained to undertake not just one but many different tasks in parallel. These technologies are highly effective and help to make operations more efficient, too.

Novacura, a Swedish startup, puts specialized cameras on the production line where an inspector previously stood. Computer vision is applied for extracting data from the captured images which is further used for automating the quality inspection on the production line. This system can, for example, identify cracks in metal pieces, even faults that can't be seen by the human eye.

For automated quality inspection and control, lovakian-startup Photoneo offers a 3-dimensional machine vision scanner designed for demanding industrial tasks such as precisely inspecting heat exchangers. This solution promises higher accuracy and faster throughput than humanly possible, along with data that is vital to optimizing manufacturing processes.

Equipment Monitoring and Predictive Maintenance

What if equipment faults are invisible to the maintenance team? Often this is the case for the specialized tools used in manufacturing plants – over time, they show signs of wear and can break, risking goods damage and stoppages. Computer vision can be used to find flaws even in tiny machine parts in real time without slowing down production. Machine-learning techniques can be used to identify problems and figure out what's wrong with the equipment – intelligent fault diagnosis – and make predictions, enabling data-driven predictive maintenance for a cost-optimized fix before failure.

SparkCognition, a software startup in Austin, Texas, applies machine-learning algorithms to historical visual data from factory equipment to build a baseline model of what normal operations look like. This is used to analyze video data in real time, identifying and flagging any deviating value – information essential to increasing throughput, preventing quality issues, ensuring operational efficiency, and cutting maintenance costs.

Process Optimization

Computer vision systems can be highly effective at increasing organizational efficiency through process improvement. And often the best outcomes are achieved by combining human skills – sight, intelligence, and brain power – with this technology.

IFM, a German electronics startup, offers a solution that integrates a computing unit, software, and a camera to capture videos and 3D images. Images are captured by the camera installed at the workstations and using computer vision, information regarding the process sequence is extracted from the image and displayed on the monitor placed directly in the field of vision of the worker. Color-coded information like tasks completed and upcoming steps are used to guide the worker to perform tasks without error.

To achieve its objective of ensuring absolute safety in manufacturing processes, Dow Chemicals employed a computer vision system. This system was designed to detect and prevent early leaks and potential contamination. The system's effectiveness relied on a model that underwent training using various annotated images depicting instances of leaks and non-leak scenarios. Furthermore, the model was fed with real-time data from surveillance feeds, enabling it to successfully fulfill its safety objectives.

NON-LOGISTICS USE CASES: DISASTER RESPONSE AND RECOVERY

Computer Vision in Disaster Response and Recovery

From droughts and wildfires to floods and landslides, earthquakes, hurricanes, pandemics, and more, computer vision and machine-learning algorithms can predict certain occurrences before they happen and provide essential warnings. In the immediate aftermath of an uncontrollable event, this technology can assist first responders and recovery experts on the ground, helping them make urgent assessments, scope the damage, and strategize rescue efforts. And over time it can help to coordinate effective reconstruction.

Current Trends

COVID-19 showed everyone the serious implications of being unprepared, from shortages of emergency supplies and other vital resources to the risks of disease transmission. Swift responses are required to the increase in uncontrollable events which can impact communities for months and even years afterwards. Effective event prediction and early warning, supported by computer vision, are saving lives.

If a heatwave is predicted 24 hours in advance, this can help reduce ensuing damage by 30%.

Computer Vision Disaster and Response Recovery Applications

Monitoring Environmental Conditions

Weather models help to improve the accuracy of weather forecasting and prediction. By combining visual data from multiple sources, such as satellite images, weather stations, and ground-level observations, computer vision algorithms can vastly improve and increase how comprehensive and effective these weather models are.

Computer vision can help detect patterns in visual weather data, such as changes in cloud cover, temperature trends, and changes in water levels. These patterns can enable meteorologists to predict future weather conditions and help governments and communities provide early warnings to keep everyone safe. For e.g., IBM and NASA have launched the first open-source geospatial AI foundation model for Earth observation data. Its wide range of potential applications include tracking changes in land use, monitoring natural events, and predicting crop yields.

Based in France, the startup TENEVIA uses computer vision image analysis and numerical modeling to measure, monitor, and forecast environmental conditions. With camera hardware and simulator software, the solution helps forecast high water levels and snow – the cameras record flows, and the simulator analyzes this visual data. Together they create a virtual fence around any flow to analyze water and snow dynamics – and can indicate irregular water accumulation as well as melting glaciers and snow to give advanced flood warnings.

Betterview in California specializes in providing risk information to the insurance industry. To get an immediate, complete picture of wildfire risk, for example, the startup combines aerial imagery, computer vision, and third-party property intelligence. Insurers use this information to streamline efficiency, as well as predict and prevent wildfire-related losses.

Optimizing Relief Resources

Before each response demanding event, it is extremely important to pinpoint where it will occur and predict the likely damage. Computer vision technology can help detect and track the path of hurricanes, volcanic lava flows, and wildfires so that a plan can be made to promptly optimize relief resources. Of course, it also enables rapid and effective evacuation planning.

AiDash, based in India, provides pre-event computer vision solutions to help predict, prepare, and deploy accurately using satellite imagery, real-time weather data, and variations in vegetation data. This detailed intelligence enables organizations to estimate resource requirements in advance.

Measuring Natural Disaster Impact

Computer vision technology can help identify and measure the impact of uncontrollable events. Large satellite image datasets of vegetation and land cover can be sourced from specialist companies. By applying deep-learning methods and high performance-based models, affected areas can be assessed and monitored over time, incorporating recent imagery to recognize any changes. To train the AI model, pre-event images are used, and AI algorithms are applied to identify damage and estimate repair costs.

During Japan's typhoon season, homeowners have been taking and submitting smartphone pictures of any property damage caused by a storm, using an AI solution from Tractable, a New York technology startup. With this simple app-based method, there's no need to send out an appraiser in the wake of a damaging event. Instead, using computer vision technology and machine-learning techniques to enable visual assessment, the insurance company can use these photos for fast, accurate damage appraisal to accelerate each customer's recovery from impact.

Challenges of Implementation

Although these use cases are drawn from various industries and operations, for the most part, they share some common challenges.

A key challenge is the persistent perception that artificial intelligence generally, and computer vision specifically, may be less accurate and less capable than a human. The right balance must be found between extracting meaningful insights, ensuring adequate defenses against hacking and manipulation that can cause AI systems to misclassify objects, and protecting individual privacy.

Wherever cameras are involved, there will always be valid concerns about data privacy and security. In some applications, particularly in the area of healthcare, there is the issue of false positives – a patient may be diagnosed with a disease they don't in fact have. Implementation barriers may arise due to the trustworthiness of AI, if experts cannot explain decisions based on machine learning.

These and other challenges, including the investment needs of technology implementation, are tending to slow the uptake of computer vision technology.

Health and Safety Applications

Accident Prevention

Warehouses, yards, depots, and other busy logistics facilities represent high risk for the health and safety of a workforce. Computer vision and AI algorithms can help monitor and analyze movements of people and vehicles including autonomous robotic devices, forklifts, and trucks. A computer vision system can detect speeding, movement in the wrong direction, parking in the wrong place, and more. It can also identify real-time noncompliance, such as workers not using walkways, and can issue safety alerts for immediate action to minimize risk and reduce unsafe behavior.

With its AI-based unsafe event capture solution, computer vision startup Protex AI collaborated with DHL in a proof-of-concept project to empower EHS teams to take proactive safety decisions. The team succeeded in converting insights into actionable cues to change workflows, improve operational safety, and implement corrective actions.

Ergonomic Pose Improvement

Human pose estimation (HPE) is a computer vision-based technology that classifies joints in the human body, capturing a set of coordinates for each (pairs of key points) that form a skeleton-like representation of the human body and can describe a person's posture and motion. This information is essential for ergonomics, the study of people's efficiency in the working environment.

For example, when a person bends down, they are at the highest risk of back injury. Twisting, turning, lifting things incorrectly, and carrying loads that are too large or heavy are also major hazards.

Working to reduce workplace injuries, tech company TuMeke has developed an AI ergonomic risk assessment platform enabled by computer vision. The user records a smartphone video of an associate doing a warehouse task, such as lifting a box, and this data is analyzed to provide a summary of risk, highlighting risky postures.

Protective Personal Equipment (PPE)

Employers are typically responsible for providing and utilizing personal protective equipment (PPE) at work such as safety helmets, eye protection, and specialist clothing. It's important that employees wear the right PPE correctly at all times. Computer vision technology provides promising solutions to track noncompliance; some systems can even identify the reason for noncompliance, such as the uncomfortable design of a high-visibility vest.

Systems can be trained to identify various PPE types in real time. Analyzing video streams from strategically placed cameras, a system can identify whether associates are wearing appropriate PPE and using it correctly. This helps ensure compliance with safety protocols and avoid incidents. In addition, the system can detect defective and damaged PPE (a major safety hazard) – something that also helps to keep associates safe and secure on the job.

Driver Support

Computer vision can be used to detect signs of human fatigue during long-haul truck driving. For example, facial recognition technology can identify drooping eyelids and changes in facial expression, using a combination of cameras and machine-learning algorithms which analyze images to find patterns that indicate driver fatigue. The system can alert the driver to take a break and even trigger an alarm to warn other road users about this potential danger.

The same technology can be used to detect and issue warnings about correct seat belt use and unauthorized personnel entering the vehicle. In fact, many modern computer vision systems perform multiple tasks simultaneously, such as object detection, facial recognition, and image classification.

A great example of this multitasking is in autonomous vehicles – the vision system can simultaneously detect and classify objects in the environment (such as pedestrians, other vehicles, and obstacles), track their movements, and make decisions based on this information to safely navigate an entire journey.

Israel-based Cipia promises an added layer of intelligence to the in-car and in-cabin automotive environment. Using edge-based computer vision and AI, the company offers a range of sensing solutions including driver monitoring systems and occupancy monitoring systems for better, safer mobility experiences.

Challenges of Implementation

Challenge 1: By definition, a prediction is never 100% correct. If a machine-learning algorithm were to make a mistake about a person's health and safety situation, the cost would be unacceptably high.

Challenge 2: It is difficult for most non-statisticians to fully understand and make practical use of AI confidence percentages and accuracy levels. EHS teams may need help to apply data.

Challenge 3: Computer vision currently struggles to detect some materials such as glass and other transparent materials, as well as things with a shiny surface and materials that change shape or form.

In the warehouse environment, as many as 20,000 US workers are injured in forklift accidents alone each year and 25% of those accidents happen when a forklift overturns.

People and Operations Applications

Heatmaps

Where are the bottlenecks and activity hotspots in the warehouse? Computer vision-based heatmap analysis can help evaluate workflow patterns inside a facility and outside in the yard with contactless, non-invasive methods.

A computer vision-based heatmap uses surveillance camera video feeds to analyze the frequency and duration of movement in different areas by both people and vehicles. Typically, a color-coded overlay is used on original images with, say, red indicating heavy activity in one area and white indicating no activity in another. This helps managers identify bottlenecks and other inefficiencies such as overutilized and underutilized areas. This information can be used to place inventory and equipment in the best locations and increase the efficiency of operations. To measure the effectiveness of any changes, heatmaps can be compared over time.

Headquartered in Switzerland, Viso offers movement heatmap crowd detection to identify human movements trajectories in public spaces and in logistics operations. This computer vision solution can improve operational efficiency by tracking changes over time, finding anomalies and detecting hot spots and bottlenecks.

Head Counting

Warehouse managers need to have enough people to do the required tasks at any given time. If there are too few, tasks may not be completed on time. If there are too many, the company incurs unnecessary staffing costs. And it's essential to comply with safe occupancy levels.

Counting people in a defined space, however, can be a tedious, repetitive task that takes up a lot of time. Instead, computer vision systems can be used to count people and vehicles in any logistics facility 24/7.

This enables logistics organizations to keep track, and analysis of this visual data helps management decide the optimal number of people to perform specific tasks and comply with occupancy requirements.

Indian-startup AIVID has created an AI software using surveillance cameras for automated detection and counting, even distinguishing age and gender, to provide operational insights and patterns for businesses.

Pick Path Optimization

Products in a warehouse must be picked quickly, accurately, and efficiently so people who do this task (pickers) must take the best and most optimal routes during their shift. Their activity can account for a substantial percentage of warehouse operations tasks, so optimizing the pick path provides quantifiable savings in operating costs. It ultimately also results in higher customer satisfaction.

Algorithms alone cannot guarantee success in every picking scenario, but companies are more likely to achieve pick path optimization when computer vision is used.

Camera video feeds inform machine-learning algorithms to identify patterns and trends in the data, making recommendations to change workflows; for example, guiding people to walk shorter distances but still complete their picking tasks by moving the location of some equipment or supplies. These feeds can also show the benefit of changing the sequence in which certain tasks are completed. If a heatmap has revealed high activity in a particular area of the warehouse, this solution may help to understand the workflow in that area so that changes can be made to resolve the bottleneck.

Access Control

Security surveillance using a traditional camera setup requires human interaction and effort to scan through many recordings. Compare this with an intelligent surveillance system which not only records video footage but also leverages advanced algorithms and techniques for real-time motoring, analysis, and detection. Backend AI systems help to rapidly process the video footage and provide rich information which can be used to increase safety and reduce theft.

A computer vision-enabled camera system can act both as a view-only platform and to detect unauthorized entry or intrusion into restricted areas. By analyzing the video feed, the system can identify when someone crosses a virtual boundary or enters a forbidden space. Alerts can be generated immediately, allowing security personnel to respond promptly. Additionally, the system's algorithm can recognize patterns of normal activity and behavior within the monitored environment. When an abnormality is detected, such as a person loitering or running in a restricted area, the system can generate alerts, enabling proactive intervention.

Challenges of Implementation

Challenge 1: For privacy reasons, companies are usually disallowed from monitoring workforce activity solely for that purpose. Data from computer vision monitoring should be used to optimize operations, and this should be explained to the workforce.

Challenge 2: If workplace surveillance methods to track workforce activity are perceived as invasive, it can reduce morale, increase work-related stress, and cause counterproductive work behaviors.

Maintenance Applications

Predictive Maintenance

Computer vision technology can help to consistently and accurately monitor logistics assets and alert maintenance teams to intervene before any issue arises. By analyzing data from various types of equipment, it can also predict when critical assets will require maintenance. This allows managers to schedule repairs and upkeep to prolong asset life and prevent failure.

Providing visibility into asset health, the Delaware-based computer vision startup Clarifai offers solutions that capture AI-based risk predictions. These not only give early warning of potential problems and reduce unscheduled repairs, but also help to delay capital expenditure, minimize human inspections, and ensure business continuity. The company claims maintenance costs can be cut by 25% and downtime by 35%, and users can achieve 10x return on the initial technology investment.

Defect Identification

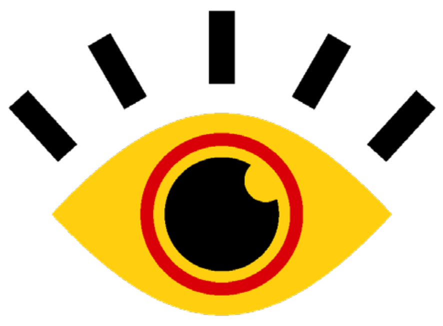
Before widespread deployment of computer vision, defect detection was a labor-intensive, manual operation impacted by human error and requiring round-the-clock employee availability.

Today, probable asset flaws, mistakes, anomalies, and problems can be automatically identified by computer vision. An effective complement to warehouse Gemba Walks, this technology provides valuable additional data when managers walk through the premises gathering observational and interactional information. It can also identify the cost of asset damage and repair, streamlining maintenance processes by providing this data to the asset management system.

Ivisys, a pioneering startup, has created an innovative defect identification solution named Pallet AI. This technology is specifically designed for enhancing the quality inspection process of pallets, effectively identifying defects, and concurrently improving productivity and employee safety.

By employing a sophisticated neural network, the system utilizes a network of cameras to not only identify cracks and holes but also to detect mold and discoloration using advanced pattern recognition techniques. Remarkably, this system is capable of inspecting a range of 250 to 450 pallets per hour.

Challenges of Implementation



Challenge 1: Computer vision systems cannot capture everything. Asset performance could be impacted by qualitative factors – such as how a worker interacts with a device – and this may require human observation.

Challenge 2: Typical warehouse computational power may not be sufficient to cater for the complex requirements of AI algorithm analysis. New IT investments may be needed.

Asset Management Applications

Utilization & Capacity Assessment

When planning capacity to optimize asset utilization, computer vision can deliver quicker insights than the human eye and that of human experience.

This technology can assess the overall space inside trucks and containers to calculate available volume prior to loading – information that helps determine the optimal arrangement of items to maximize loads and minimize wasted space. Measurements can be taken throughout the loading process, enabling real-time data-driven decision-making that saves time, improves efficiency, increases sustainability, and reduces cost.

In the warehouse, computer vision can be used to analyze the dimensions and orientation of pallets and roller cages. This data helps ensure these assets are positioned for optimal load distribution and peak efficiency.

Danish startup Sentispec uses computer vision to manage every point of contact with stock in and out of the warehouse. Instead of allowing partially filled trailers and containers to leave the premises, Sentispec Inspector helps record the densities and fill rates of every load, so the planning office can optimize fill rates.

Asset Counting and Localization

A familiar challenge inside the warehouse is pallets, cages, trolleys, and other assets going missing. It costs time and money to find and return or replace them. Computer vision can be used to count and locate assets, assessing their status in real time to provide visibility and improve efficiency even in warehouse dark zones, where the network signal is weak and tracking sensors may lack connectivity.

For object counting, deep-learning algorithms detect and classify objects in an image or video stream, identifying and analyzing image focus points and repeating this process to count all instances of a specific object. Assets can be identified by type (roller cage, rack, forklift) or by a unique identification code linked either to a single asset or to mul tiple assets within the camera's same field of view.

For localization, a multi-target tracking system using the 'handshake method' is effective. As an asset leaves one camera's field of view, it reappears in the view of another. A backend algorithm analyzes this input to estimate and trace the asset's path throughout the warehouse. The computer vision platform by startup Kibsi uses existing camera networks to track assets in this way and monitor activity within a warehouse. Assets can be georeferenced on a virtual map and warehouse operators can locate assets with accuracy to within an inch.

Fleet Management

Assets outside the warehouse can be monitored 24/7 by an integrated system combining computer vision with surveillance. To restrict yard access to registered vehicles only, cameras can identify each truck and log its entry time, exit time, and number of daily trips. The system can also measure asset usage patterns, including idle time, and use this data to help optimize fleet operations.

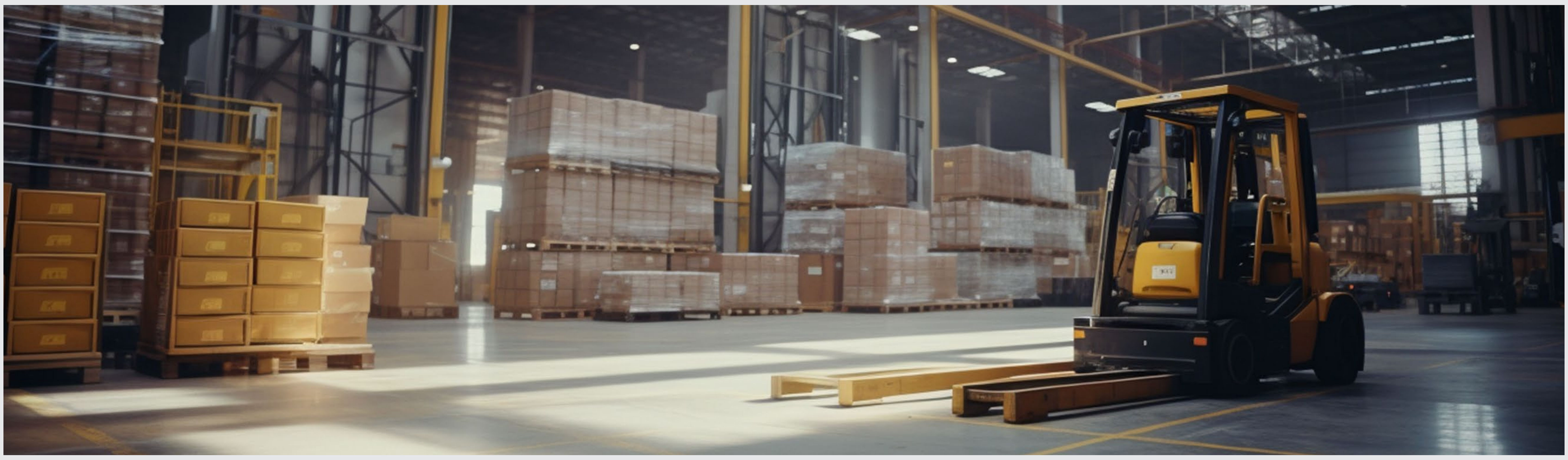
With its real-time monitoring solution, ThinkIQ claims to eliminate the need for a guard to log trucks into and out of a facility. This system is trained to work in all types of lighting and atmospheric conditions, delivering actionable insights to improve fleet management.

Challenges of Implementation

Challenge 1: A camera's field of view can be obstructed; for example, if items are temporarily stacked just inside the truck, this stack may block visibility further in the back.

Challenge 2: Significant data throughput is required for computer vision asset tracking.

Challenge 3: Technical glitches and breakdowns in the computer vision system would likely cause significant asset management difficulties, especially in a busy warehouse.



The projected cost of unscheduled aircraft maintenance is expected to rise globally from $6.57 billion in 2017 to approximately $13.13 billion by 2035.

LOGISTICS USE CASES: SHIPMENT FOCUS

Dimensioning Application

Dimensioning

It's essential to accurately measure either the area or the volume occupied by an object before shipping. Dimensioning enables capacities to be calculated for storage, handling, load utilization planning, and transportation of goods, as well as for shipment billing. And dimensioning helps to optimize the master data on which each warehouse management system (WMS) operates.

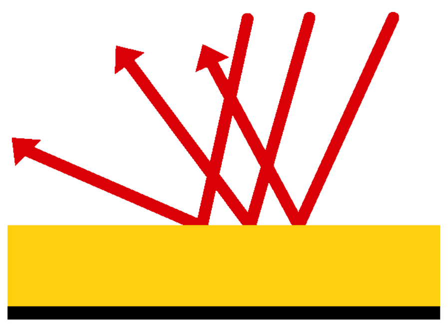
The time dimensioning takes is critical. For example, DHL ships millions of various sized packages a day, most traveling at fast speeds on a conveyor belt, so even a momentary delay in understanding their sizes is likely to significantly stall progress. And not just for speed but also for optimal packaging planning and better use of materials, it is very important to identify odd-shaped parcels mixed in with regularly shaped ones.

The dimensioning process can be automated with computer vision technology, in different types of systems. A fixed system such as the MetriXFreight solution from the German software company Metrilus can continuously monitor a defined measurement area for goods. It starts dimensioning whenever an item arrives at the measurement area from any direction, regardless of orientation and exact positioning.

Based on a handheld terminal, the California-based software company Qboid offers a mobile dimensioning solution that integrates advanced 3D color sensors and computer vision algorithms.

This enables the automation of dimension and shape estimation tasks in more types of items than is possible with a traditional system.

Challenges of Implementation



Challenge 1: Shiny objects cannot be accurately measured, as reflective surfaces and transparent materials that bounce back light pose a challenge for dimensioning technologies.

Challenge 2: Extremely large objects and extremely small objects are difficult to measure with automatic dimensioning systems.

Challenge 3: Accuracy depends on the position and resolution of the camera used for dimensioning.

Compliance Applications

Quality Inspection

When thousands of items of various types, sizes, and shapes are passing through a facility every day, it may not be that easy to spot a damaged item. But if a box is wet, ripped, or bent out of shape, and the products inside are also spoilt in some way, this can jeopardize brand image and reduce consumer satisfaction.

The important thing with shipment quality inspection is to detect damage early on. If a damaged item passes through the facility, it will likely be labeled and assigned a delivery location. This will make it difficult to trace and redirect, as the damaged item will need to be retrieved, repackaged, relabeled, and rerouted.

US manufacturer of machine vision systems, Cognex offers an advanced dimensioning solution that includes damage detection. With a 3D and 2D camera system, it automatically detects damage to boxes such as open flaps, dents, bulges, and crushed parts. It eliminates manual re-routing, enables early intervention, and keeps the shipment line moving efficiently.

Label Detection and Alignment

Consumers are unlikely to buy a food or beverage product or use medication if the label is missing or illegible. Typically, companies are legally obliged to provide a list of ingredients, nutritional information, and a ‘best before’ date with these products and more. Country compliance laws may also require allergen labeling, recyclable materials labeling, and more. Companies can lose money and even face legal action for a bad batch of labels.

Computer vision-based systems can read labels, capturing front and back images of packaging before shipment. This visual data is processed to ensure individual product labels match a master image. Some systems also include an expiry date check system.

If errors or anomalies are detected in the text and graphics of the label itself – maybe the ink is smudged or the label is wrinkled or misaligned – or if a product has been mislabeled, the AI-powered system can trigger a reject mechanism. Detecting label damage early helps to speed up resolution. The item can be extracted from the goods flow and promptly relabeled, resuming its journey in the shortest space of time.

A computer vision system can also identify labels that have been applied to hazardous materials, using this visual information to automatically reroute those items to be repacked in compliance with any hazardous materials shipment requirements. Similarly, it can identify labels on high-priority shipment items that must be delivered to customers within an agreed timeframe.

The Denver-based computer vision startup Visionify offers label checking solutions to ensure a smooth shipping process. These can verify labels on cartons, boxes, and bags of all sizes, shapes, and orientations, whether the item is stationary or in motion.

Barcode Scanning and OCR Capture

Automated product identification takes numerous forms. Barcode lines, dots, and spaces have become a familiar sight by the early 1980s. More recently optical character recognition (OCR) has enabled information from scanned documents and images to be converted into machine-encoded text.

Computer vision systems that adopt barcode scanning and OCR capture can effectively replace time-consuming manual barcode scanner and data input.

PepsiCo uses an advanced AI-powered platform solution from KoiReader Technologies and NVIDIA to enable greater efficiency in reading warehouse labels and barcodes in its fast-moving environments. Labels can be any size, at any angle, and even partially obstructed or damaged.

Also, the technology company Banner offers a range of innovative solutions in this field. These can decode difficult-to-read, low quality, and damaged barcodes, and even codes printed on highly reflective surfaces.

The Belgian technology company Zetes specializes in solutions for supply chain optimization and offers computer vision systems for shipping and loading verification. These can capture a full view of loaded pallets, detecting and decoding multiple barcodes simultaneously and comparing this data to shipping orders.

Operators receive real-time alerts about discrepancies and missing or unreadable barcode labels. Camera gates can then capture images of every pallet, and the system decodes and verifies pallet barcodes, triggering a go/no-go signal for the loading operator. Proof of loading is provided by time-stamped images.

Challenges of Implementation

Challenge 1: Environmental factors can be a challenge for barcode reading. Under poor lighting, shadows, or reflections, algorithm accuracy decreases.

Challenge 2: Not all handwritten labels can be read by OCR technology, and it may be impossible to scan greasy or torn barcode labels. Manual data entry may sometimes be needed.

Shipment Localization: Automatic Sorting

Automatic Sorting

Shipments need to be identified before being sorted according to their type and destination of delivery. This visual classification process can be slow and tedious for workers, and can cause a bottleneck, so it is critically important to automate these tasks. AI and neural networks can be used to segment and pinpoint the location of parcels, envelopes, bags, and more and to accurately determine the placement and alignment of these items.

Camera-based automated sorting systems are used to accelerate the sorting process, and the best results are achieved when utilizing high-quality 3D data. These systems distinguish the different characteristics of each item to enable object classification and quality evaluation. As items arrive on the conveyor belt, they are uniformly separated, spaced, and aligned (a process called singulation). Cameras capture images of each item, which are recognized by AI-powered technology.

The known item is then directed to the correct bin or onto another conveyor belt for further sorting.

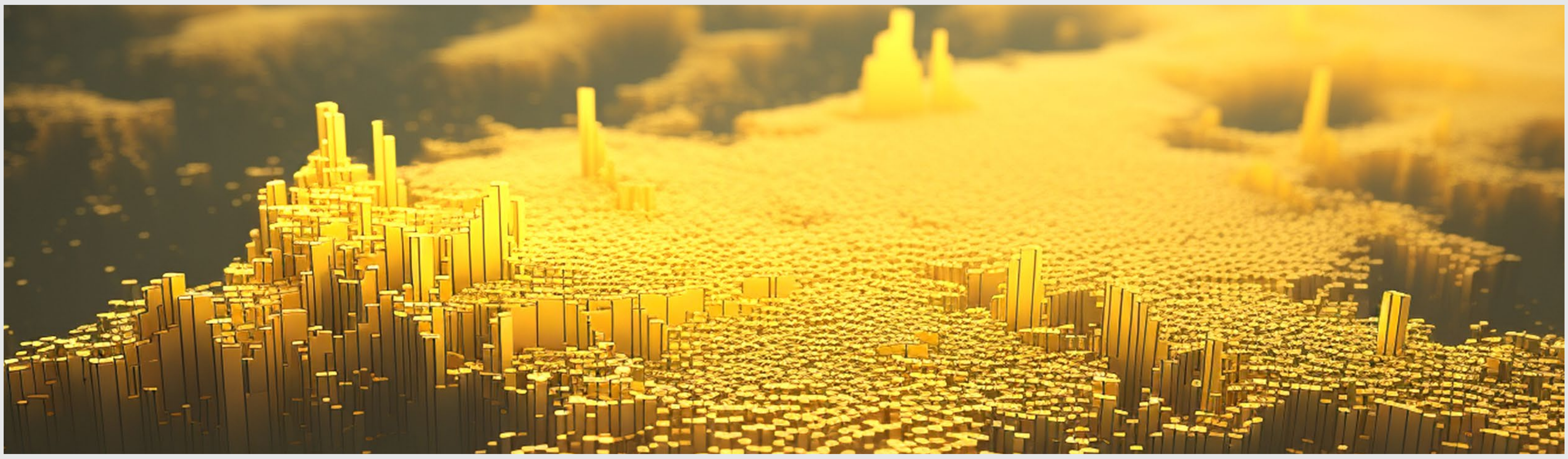
This type of computer vision system matches the image and the sorting decision in real time, and there are three steps to achieving the sorting decision: first, the captured image is converted to a binary image; next, product edge detection is applied; then the result is matched with the base image.

Combining a PhoXi 3D scanner and localization as the eyes of the system and AI algorithms as the brains, Photoneo has achieved an advanced singulation and sorting solution based on a pre-trained neural network that can recognize parcels for pickup with 95% accuracy. It can pick up solid objects with deformed surfaces that change under the slightest pressure, doing so in cycle times of less than 1.3 seconds for throughput of 2,500 non-standardized objects per hour.

Challenges of Implementation

Challenge 1: Not all computer vision singulation systems can handle shiny surfaces.

Challenge 2: It is difficult to achieve fast, accurate automated sorting of disorderly stacked parcels.



To automate the dimensioning task, a computer vision system must express the external surface of an object in terms of data points – a process that may take less than a second in an advanced system (capturing more than 1.5 million data points to form a 3D point cloud). Depending on the chosen method of dimensioning, some systems may only process flat, uniform surfaces while others can, for example, also handle odd shapes.

Conclusion and Outlook

We at DHL recognize that computer vision has reached a level of maturity that now ensures exponential growth in implementation and use cases. And recent advances in computer vision technology – particularly depth perception, 3D reconstruction, and interpretation of dark and blurred images – are broadening the field of applications for computer vision in logistics operations along the supply chain.

In the immediate future, we anticipate increased investment in artificial intelligence (AI), not least because of society's growing interest in generative and interactive AI such as ChatGPT and Alexa. Boosted investment will be essential to further advancement of computer vision solutions, as AI underpins image recognition.

Within the logistics industry today, computer vision has proven and established its value in certain areas such as dimensioning solutions as well as in many health and safety applications as outlined in this report. There remain large numbers of tasks in warehouses, distribution centers, hubs, and the supply chain that can and should be gradually delegated to or supported by computer vision-enabled AI systems. And ultimately these solutions will scale to deliver maximum benefit to logistics organizations.

Computer vision technology is here to stay, and it is fascinating to track the increasing commercial deployment of these AI-powered solutions not only within the logistics industry but also across various other industries.

Several key challenges remain. One is technology acceptance – employees may be suspicious of and resistant to computer vision solutions, believing they are being constantly watched and their personal data is at risk. Compliance with GDPR and other data privacy and protection laws must be ensured, and these regulations may limit the scaling of computer vision solutions across some regions. Cybersecurity must be strengthened to reduce the risk of hacking and malicious manipulation of data which can skew analyses and alter AI performance. Significant improvements in security as well as data privacy are vital to ensuring the trustworthiness of AI.

In terms of the solutions themselves, many are hardware agnostic, but it is often necessary to upgrade older camera technology and there is still a need for more cameras, especially in outdoor operations and for high-definition applications. It is essential to enable effective communication between cameras, sensors, and other elements of the AI-power computer vision system, achieving a seamless view in multi-target, multi-camera object detection and tracking.

Further development is also required in the visual identification of materials such as frosted glass and objects with reflective surfaces. Better cameras will deliver higher quality output resulting in more proficient machine learning, allowing the system that once struggled to recognize a surface to subsequently identify it without any difficulty.

Computer vision represents an intrinsic part of the overall automation and digitalization of logistics, reshaping the future of an industry that unites us all. Its solutions will underpin and drive future logistics to achieve more efficient processes as well as safer and truly sustainable operations.

Today at DHL, we are closely monitoring the maturation and further development of computer vision. And by collaborating with our partners and customers, we intend to drive technological advances that will enable implementation in all logistics operations and across the entire supply chain.

The Team

Dr. Klaus Dohrmann
Project Director & Co-Author

Emily Pitcher
Editor-in-Chief & Co-Author

Maulik Kamdar
Research Lead & Co-Author

Other Contributors
Amy Henshall
Angela Hills
Anna Finkbeiner
Bastiaan Snaterse
Ben Gesing
Christian Lundbak
Dina Falk
Graham Avery
Holger Schneebeck
Julian Selders
Kristin Szekat
Lars Pappe
Maida Ajmal
Noah Tombs
Olande Stols
Paul Schlinkert
Philip Jensen
Santiago Romero
Stefan Fuehner
Susanne Lauer
Torben Pagh
Zineb Darkouch

Advertising and Design Agency
Archetype

Editorial Support
Words Europe

Publisher
DHL Customer Solutions & Innovation

Represented by Dr. Klaus Dohrmann, Vice President,
Head of Innovation & Trend Research DHL

Project Management and Editorial Office
Emily Pitcher

DP DHL Research & Innovation GmbH
Heinrich-Brüning-Straße 5
53113 Bonn, Germany