# Artificial intelligence and machine learning revolutionise railway operations

*Guest article by Ulrike Gollasch*

Artificial intelligence (AI) has experienced enormous development in recent years and has become an important topic in everyday life and in the world of work. Sometimes AI is presented as a threat to human labor or even as a potential danger to society. The concern that AI is smarter than humans is a common theme. AI can help to improve, accelerate and automate processes.

AI is a subfield of computer science. It mimics cognitive abilities by recognizing and sorting information from input data. This intelligence can be based on programmed processes or generated by machine learning. Computers can use AI or machine learning to sift through vast amounts of data to find connections (patterns) within it. A distinction is made between strong and weak AI.

Weak AI aims to learn individual subject areas and to solve specific application problems in them. Learning is an integral part of this. The weak AI must be able to deal with problems where there is some probability of different outcomes. Intelligent behavior is simulated using mathematics and computer science, but it is not about creating consciousness. Already developed and productive solutions belong to the weak AI.

The goal of strong AI is to learn skills in all subject areas and to solve tasks in them. Here, too, learning is an integral part. It is unlikely that it feels emotions. But it can imitate them. A strong AI is as smart as a human and can use its intelligence in the same subject areas. When it comes to strong AI, the developments are not yet as advanced, and there is still no machine that is as smart as a human in all areas.

With machine learning methods, an algorithm learns to perform a task independently through repetition. The machine orientates itself on a given goods criterion and the information content of the data. But unlike conventional algorithms, no solution path is modeled.

A sub-area of machine learning is data mining. This is a process in which large amounts of data are examined for patterns and connections in order to gain valuable insights. The goal is to discover hidden information and trends that can help in decision-making and planning.

Examples of machine learning are image recognition (face and environment recognition), speech recognition (voice assistance systems, voice control), semantic speech recognition (translation software, chatbots), pattern recognition (error detection in vehicles), process optimization (optimized process control).

## AI solutions for the rail industry

Zedas GmbH, a leading provider of software solutions for the digitization of railway processes, is working on various AI solutions to make the railway industry more efficient and safer. The company already offers productive solutions and is intensively researching new technologies. Applications range from optimizing logistics management to predictive maintenance. The advanced technologies should help to improve operations, save costs and increase asset availability.

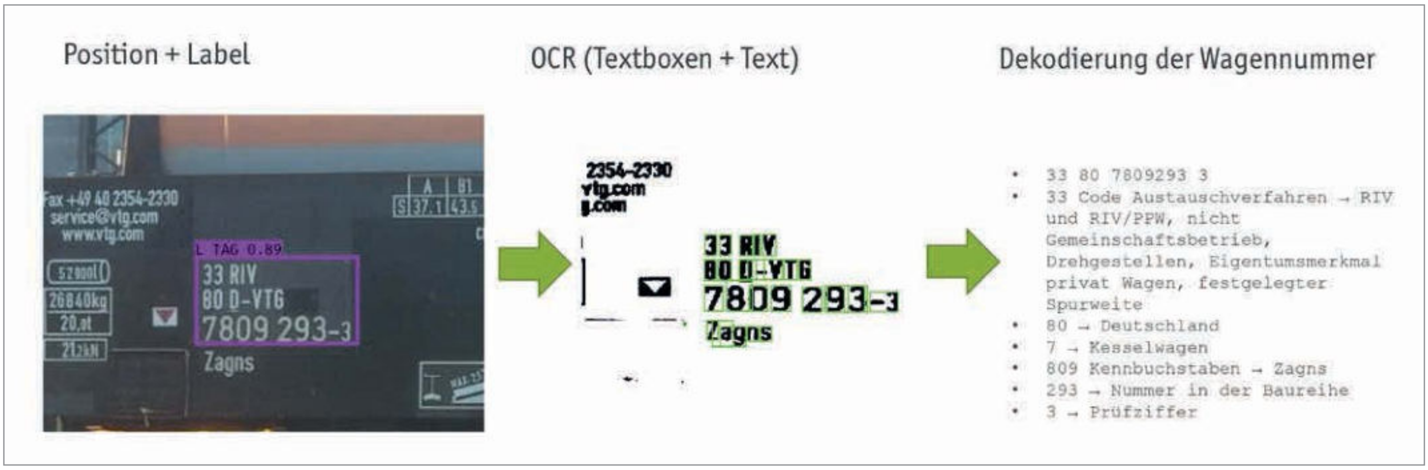
## Paperless train dispatch with AI-based wagon number recognition

With the help of the Train Check app, which was awarded the German Mobility Prize in 2020, wagon inspectors can

**Improving efficiency and safety are already possible today**

The field of application of AI is huge. It can already be used to support a large number of processes in order to improve the efficiency, reliability and safety of rail operations. It can be expected that the areas of application of AI in the railway industry will continue to increase in the coming years.

- **Maintenance planning and condition forecast:** AI-based asset management systems can predict maintenance requirements and intervals by the system learning to recognize certain patterns and relationships and to provide information about an impending defect in a component.
- **Automatic fault detection:** AI-based systems can automatically detect faults in the railway systems. As a result, it can alert personnel at an early stage if deviations from the norm occur, and thus help to avoid potential failures or malfunctions.
- **Workforce planning:** AI-based systems can assist in the planning of human resources by automatically scheduling employees' working hours and conditions, and analyzing historical data on employee performance, workload, etc. to make accurate predictions for future work requirements.
- **Capacity planning:** AI-based systems can support the planning of rail operations and infrastructure capacities by predicting demand based on market analyses, seasonal patterns or political developments, for example.
- **Operational monitoring:** AI-based systems can monitor and analyze data in real time to detect possible deviations from normal operations and to warn employees in good time.
- **Train control:** AI-based systems can automatically control the movement of trains to ensure optimal speed and routing.




Wagon number recognition locates, analyses and decodes the wagon number.

carry out paperless train dispatch directly on the train. The app creates the wagon lists and brake slips and calculates the brakes. In order to reduce manual entries during train dispatch, an AI-based wagon number recognition was developed and integrated into the app. The AI automatically determines the position of the wagon number (UIC number) on the freight wagon, decodes it and uniquely identifies the wagon. This technology is an example of supervised learning with labeled data. A type of machine learning in which an algorithm should learn from input and output data. The algorithm is trained with a data set consisting of input data (features) and the corresponding output data (labels). The goal of the algorithm is to learn a function that maps the input data to the correct output data.

## Training the AI model for wagon number recognition

A collection of photos of freight wagons with UIC wagon numbers formed the basis for teaching the AI model. Care was taken to ensure that there was sufficient photo material of each type of the different spellings, for the most varied positions on the wagon and of all wagon types. The wagon number recognition is a multi-stage procedure. First, an object detector finds the wagon number between the other wagons. Individual characters in the number are then localized and identified using OCR recognition. Finally, the recognized number is broken down into its components, including text recognition, and the recognized components are decoded.

## AI supports rail inspection and maintenance

As part of the mFUND funding program of the Federal Ministry of Transport and Digital Infrastructure, the German Center for Rail Transport Research at the Federal Railway Authority, DB Netz AG Frankfurt a.M., the Federal Institute for Materials Research and Testing Berlin, the Institute for Civil Engineering of the TU Berlin, Vrana GmbH from Rimsting and Zedas GmbH launched a joint research project in December 2021. The aim is to develop an IT tool for maintaining the rail infrastructure, with which the data collected from test runs can be used efficiently through the use of AI.

So far, the maintenance of the rail infrastructure has mostly been based on preventive concepts: Which maintenance measures are to be carried out is decided on the basis of measurement and test data on the condition of the rails from regular inspection trips. So far, however, this data has only been partially automated, but mostly evaluated with a great deal of manual effort and by specially trained personnel.

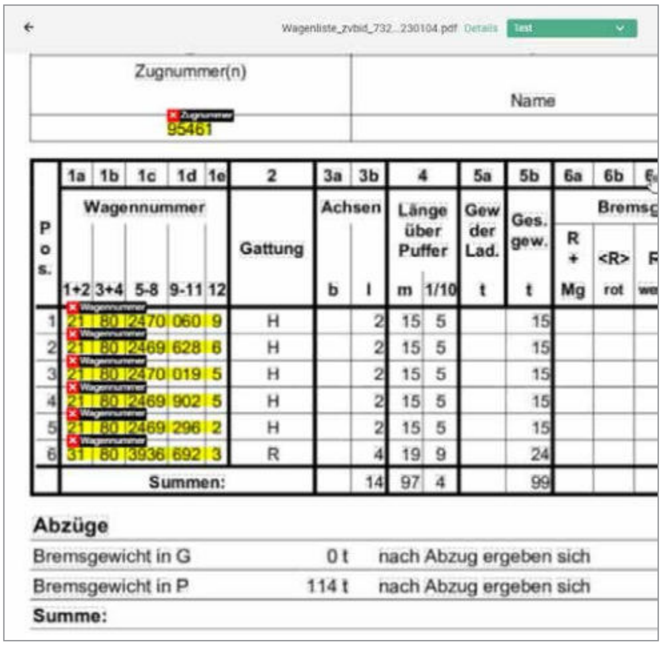
In addition to the pure error detection, a new quality of knowledge gain from the data analyzes should be achieved: a classification of the errors in "risk classes" that express the urgency of the error correction. The effort increases with the size of the route network.

Ultrasonic and eddy current inspection data are important technologies used in the inspection and monitoring of railway infrastructure. Rail head, rail web, rail foot and weld seams are checked for internal defects by means of ultrasound on tracks and points. Surface defects on rails - mostly caused by rolling contact fatigue - can be detected using eddy currents. The criteria for error evaluation are damage depth and error frequency. In the future, AI should support the entire process - from data acquisition from the test systems and analysis to decision-making and visualization.

## Automatic document recognition and processing of railway-specific documents

In the railway industry, a lot of information has to be exchanged across companies that affects logistics management and maintenance management - for example bills of lading or operational release protocols. This information is often available in different formats and has to be typed out manually in order to be able to process it further.

In order to run freight trains, railway companies (RUs) need a large amount of information such as wagon numbers, wagon types, wagon keepers, transport documents and much more. Increasingly, these must be transmitted electronically to various partners - as in the TAF TCM example. With the transmission of the Train Composition Message (TCM) or Passenger Train Composition Message (PTCM), RUs will in future be obliged to send the network operator / railway infrastructure company information on the current composition of trains for dispatching. The transmission of the train content data is based on the train number and must be introduced by all players in the European rail market for both freight and passenger transport. RUs receive information about wagon/loading lists in a wide variety of forms such as PDF, emails, Excel, and photos. The commissioned RU has to type this information in manually or convert it into a uniform format, which is time-consuming and labour-intensive.



Wagenliste_zvbid_732...230104.pdf Details Test

| Zugnummer(n) | Name |
|---|---|
| 95461 | |

| Pos. | 1a Wagennummer 1+2 | 1b 3+4 | 1c 5-8 | 1d 9-11 | 1e 12 | 2 Gattung | 3a Achsen b | 3b l | 4 Länge über Puffer m | 1/10 | 5a Gew der Lad. t | 5b Ges. gew. t | 6a Bremsg R + Mg | 6b <R> rot | 6… R we |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 21 | 80 | 2470 | 060 | 9 | H | | 2 | 15 | 5 | | 15 | | | |
| 2 | 21 | 80 | 2469 | 628 | 6 | H | | 2 | 15 | 5 | | 15 | | | |
| 3 | 21 | 80 | 2470 | 019 | 5 | H | | 2 | 15 | 5 | | 15 | | | |
| 4 | 21 | 80 | 2469 | 902 | 5 | H | | 2 | 15 | 5 | | 15 | | | |
| 5 | 21 | 80 | 2469 | 296 | 2 | H | | 2 | 15 | 5 | | 15 | | | |
| 6 | 31 | 80 | 3936 | 692 | 3 | R | | 4 | 19 | 9 | | 24 | | | |
| | Summen: | | | | | | | 14 | 97 | 4 | | 99 | | | |

| Abzüge | | |
|---|---|---|
| Bremsgewicht in G | 0 t | nach Abzug ergeben sich |
| Bremsgewicht in P | 114 t | nach Abzug ergeben sich |
| Summe: | | |

Document recognition automatically reads out wagon numbers from a wagon list.

Zedas is therefore working with a partner on an AI that reads these documents, evaluates them automatically and extracts the necessary information. A large number of documents must be labeled for AI training. Depending on the format and type, several examples are necessary to train the AI sufficiently. Test documents help the AI to evaluate what has been learned. As a result, the wagon data can be taken over automatically with a simple upload of the document and is thus available almost seamlessly for the following train control and dispatching activities. The next step is to learn how to read out maintenance documents, such as wheel set cards or operational release protocols.

Predictive data mining models are to be integrated into the logistics management software zedas cargo and the asset management software zedas asset in order, for example, to enable cost savings based on the determined data through optimized delay calculations or spare parts logistics analyses.

## Machine learning methods used

- Linear/Nonlinear Regression is a statistical technique used to examine the relationship between a dependent variable and one or more independent variables. In this case, linear / non-linear regression is used to find out whether the weather influences the delay times or whether there is a connection between the damage figures and the vehicle data such as the year of manufacture or the mileage. The goal of linear/nonlinear regression is to find a mathematical relationship that describes the relationship between variables in order to make predictions about future events.

- The k-Nearest Neighbor (KNN) method is one of the simplest machine learning algorithms and is used for object classification. In this case, vehicles are classified as wagons or locomotives based on their specifications such as dead weight and length. For each new object, the ANN algorithm finds the k nearest neighbors in a data set that are most similar and classifies the new object according to the majority decision of the k neighbors.

- Cluster analysis k-means algorithm is a technique used in statistics and machine learning to group similar data points into groups or clusters. The k-means algorithm is a commonly used cluster analysis algorithm. In this case, cluster analysis is used to analyze data on loading and timing relationship to see if there is a relationship between loading and delay. The k-means algorithm groups the data points in such a way that similar data ends up in the same clusters and differs from other clusters.

## Prerequisites, challenges and risks of AI

Data collection is a challenge for many RUs, as data is often insufficiently or incompletely collected. However, precise prediction by AI systems requires large and well-structured data sets of high quality. Incorrect or incomplete data can lead to errors in the predictions. Therefore, data should be checked, cleaned and updated regularly. System decisions must be comprehensible in order to gain the trust of direct employees, but also of society as a whole. Complexity is another challenge when implementing AI technologies in systems and requires careful planning and coordination as well as a deep understanding of the industry in which it is intended to be used.

In summary, the rail industry can benefit from the application of AI and machine learning. AI-based systems can help improve processes, increase the efficiency, reliability and safety of rail operations, and optimize maintenance and operational monitoring. The areas of application for AI in the rail industry are already diverse, and it can be expected that they will continue to increase in the coming years. However, it is important to emphasize that AI does not pose a threat to human labor, but rather can be a support and facilitator in complex tasks. The correct use of AI can therefore help to make the railway industry more future-proof and sustainable. ==

Author: Ulrike Gollasch is Head of Marketing at Zedas GmbH in Senftenberg.

Further information in the presentation by Dr Peter Engel "Machine Learning and AI for optimising maintenance": https://www.zedas. com/de/ki-in-der-bahnbranche