Technology

Applications of Computer Vision in Steel Production Processes

Koji HIRANO*
Sachi ISHITANI
Yuto HIRATA
Tadashi HYODO

Abstract

In order to adapt to labor shortages due to the declining birthrate and aging population, as well as intensified competition world-wide, we are transforming our steel production processes in terms of automation and productivity by using computer vision technologies. This article first explains computer vision based application which supports the analysis of videos captured during work by operators. We next explain automation of various monitoring and confirming works by operators. Development of these applications utilizes fast growing deep learning technology. We then describe the development of a framework to maintain the performance of image recognition tasks, which is important in the deployment of deep learning vision systems.

1. Introduction

Steel products of Nippon Steel Corporation are manufactured through multiple processes such as blast furnace ironmaking, electric arc furnace steelmaking, refining, continuous casting, rolling, and coating. Finished products are inspected for surface and internal defects, mechanically tested, packaged, and shipped by ships or trucks. Over the years, Nippon Steel has pursued labor-saving and automation of manufacturing processes by applying various sensors and automatic control technologies. As a result, even extremely long processes (production lines) that extend up to 1 km can be operated by just a few operators. However, due to the labor force shortage caused by Japan's declining birthrate and aging population and intensified global competition, further labor-saving measures, automation initiatives, and labor productivity improvements are being sought. This paper describes our initiatives to transform our business using image processing technology to address these challenges. Recent advancements such as the affordability of high-definition cameras, the emergence of deep learning technology, and progress in open-source software related to image processing have lowered the development hurdles for image processing technology. This has enabled the increased use of image processing in the steel industry.

This paper first focuses on "operator task analysis" utilizing image processing technology. Video recording of tasks is often used for analyzing current tasks, which serves as a starting point in aiming for further efficiency in various tasks. However, one challenge is the high workload involved in video analysis after recording. In response to this challenge, we have developed technology that uses image processing to support and improve efficiency. Reducing the burden of current task analysis and shortening its duration can lead to rapid business transformations. Additionally, from the standpoint of capturing operator tasks through images, we have also been working on the development and company-wide deployment of a platform named "Anzen Mimamori-kun (Safety Watcher)" designed to ensure worker safety (a tool for remote monitoring and remote task support by managers or others for safe solo work at the steelworks). However, this platform is not covered in this paper.

Next, we describe the automation of monitoring and checking tasks performed by operators. As mentioned above, we have been trying to save labor and automate operations, but operation and quality monitoring and checking are tasks still left to operators. We have been automating these tasks as well. In this paper, we describe "precursor monitoring of abnormal product transportation conditions" as an example of operation monitoring, and "fracture surface identification in impact test" and "check of steel coil products after packaging" as examples of inspection after packaging. We also use "steel ID recognition" in product checking tasks to prevent mix-ups of semi-finished products and products.

Regarding product appearance inspection, please refer to the next technical paper No. 131-07, "Development of Automatic Steel Surface Inspection System toward Digital Transformation," in this issue No. 131 of the Nippon Steel Technical Report (NSTR). The system described in the technical paper No. 131-07 automates the visual checks and judgments that operators make now. This utilizes deep learning, which has progressed significantly in recent years. In

---

* Ph.D., General Manager, Head of Dept., Production Management Research Dept., Intelligent Algorithm Research Center, Process Research Laboratories
20-1 Shintomi, Futtsu City, Chiba Pref. 293-8511

the steel sector, for example, there are disturbances such as differences in brightness between daytime and nighttime and rust. So, it is difficult to create a model that can stably and accurately judge in various situations using conventional image processing technology alone. It was challenging, even for experts in image processing. Using deep learning, it is now possible to consistently obtain highly accurate models by providing many images containing such disturbances and the corresponding correct answer data. Deep learning requires more rigorous and advanced accuracy maintenance than machine learning for structured data. As a corresponding measure, we describe the construction of frameworks to maintain the accuracy of image recognition models at the end. In addition, these image recognition models are developed and operated using the NS-DIG™ and the AIRON-EDGE™. For more information on these common platforms, refer to the technical report No. 131-18, "Platforms to Realize Nippon Steel's DX," in this issue No. 131 of the NSTR.

2. Production and Deployment of "Video Analysis Tool"

2.1 Issues in observing tasks at production sites

"Industrial Engineering (IE)" is widely known as a series of analysis and improvement techniques to optimize production management. It has been used in manufacturing industries, including the steel industry, since the high economic growth period. Among the IE techniques, the "work observation" method is an important means of measuring and recording the actual events at the production site and extracting the issues to be addressed for problem solving. The traditional way of "work observation" is a simple method of measuring and recording work with only a "stopwatch" and "recording paper" brought to the site. To lead to true problem solving with a single observation, the observer needs to have a high level of knowledge as an IE engineer. In recent years, videos recorded with cameras and other devices have come to be used for "work observation" and direct observation by the human eye. The main advantages of using videos are ① the ability to review recorded data multiple times, ② the ability to play back the video slowly as needed, and ③ the ability to observe even in restricted access areas defined for safety reasons. However, full-scale practical use of videos has not been achieved. The challenges of using videos for "work observation" are described below.

Watching videos and preparing record sheets for later work analysis take significant time and effort. Commercial work analysis software using videos is introduced in Reference 1). The software is intended for a limited range of specific tasks and may not be suitable for recording in a wide area, such as in the steel industry production sites.

2.2 "Development of video analysis tool"

To solve the above challenges, we worked on the creation of a new "video analysis tool" as a method suitable for "work observation" at the production site (**Fig. 1**). In creating this tool, we incorporated the following functions to solve the issues mentioned above. We refined the specific functions while verifying them in actual production sites.

Measure 1: Efficient and precise collection of work records

The functions are as follows: "① function to read the combined work of multiple objects accurately (shooting location, people, equipment, instruments)," "② function to automatically read time and other information associated with the information in function ①," and "③ function to view multiple videos shot at the same time simultaneously."

Measure 2: "Automatic video judgment function"

As video judgment functions that can automatically acquire the work status (start to end) of each object from video information, we developed the function of comprehensively judging changes in color, brightness, movement of objects ("optical flow"), and feature quantities using the "computer vision" method and converting these changes into "work observation" records (**Table 1**). Users who view the analysis content on the screen can use it to set the necessary information on the tool screen for setting operations, even if they have no knowledge or experience of programming. This allows for the use of the video analysis tool by operators working in the production sites of the steel industry.

2.3 Application effects of "video analysis tool"

The application results of the above measures are described below.

Fig. 1 Screen of the video analysis tool

Effect 1: Achievement of approximately four times faster work record speed

Using video data of a certain task, we compared ① work records with the “video viewer and spreadsheet software” and ② work records with the “video analysis tool.” As a result, an improvement of approximately four times in work record speed was observed, as shown in **Table 2**.

Effect 2: Almost complete automation of specific work observation

We explain the effects of the “automatic video judgment function” by using the example of “machine scarfing work observation.” This work thermochemically removes defects and impurities on the surface of steel slabs and consists of three steps: “preheating,” “scarfing,” and “extinguishing.” To statistically evaluate the time required for these tasks, two monitors that workers constantly watch were recorded with a video camera from behind for a long time. The “automatic judgment function” was used almost completely to automate the preparation of work records and significantly improved the efficiency of work observation. This automatic judgment method is explained by using figures. First, two judgment regions are defined by coordinates, and the judgment method and conditions are set on the screen for each judgment region (**Fig. 2**).

Table 1 Use cases of automatic video scene recognition in steel manufacturing

Recognition target	Applied recognition function
① Busy/idle durations of a coiler	Optical flow
② Busy/idle durations and moving direction of an overhead crane	Optical flow
③ Arrival/departure of palettes in a pit	Feature matching
④ Human intrusion into a restricted area	Color recognition or combined recognition

Table 2 Comparison with a conventional recording method

Method	Recording speed	Improvement ratio
① A video viewer and a spreadsheet software	20 records/hour	Approx. 4 times
② The developed video analysis tool	81 records/hour	

Next, the judgments of the two judgment regions were combined to judge “preheating,” “scarfing,” and “extinguishing” (**Fig. 3**).

The video analysis tool is currently used at each steelworks for capacity verification and extraction of production issues. Based on the results of the analysis, improvements are implemented to improve productivity and safety. From the perspective of DX personnel development, this tool has also become useful for supporting onsite staff work. In the future, we plan to increase the number of introduction and application sites and accumulate results to improve the video analysis tool further.

3. Precursor Monitoring of Abnormal Product Transportation Conditions[2)]

3.1 Challenges in product transportation

The steel industry has facilities that weld multiple coils and continuously process them in a belt-like form to improve productivity (**Fig. 4**).

Coils are transported using many transfer rolls. Depending on the operating conditions, such as the uneven shape of the coiled strip in the strip width direction or the alignment condition of the transfer rolls, the strip may shift perpendicular to the direction of transport (hereafter referred to as threading). In the worst case, it may collide with the equipment, causing the coil to break (strip breakage) (**Fig. 5**). The frequency of strip breakage is low, but the impact of strip breakage on the operation of the mill is significant. Previously, camera images were monitored by people. When the strip laterally shifted, an action was taken to reduce the threading speed. Humans can detect lateral strip shifts, but constant monitoring is difficult and has been a challenge.

3.2 Construction of precursor monitoring system for abnormal product transportation conditions

We have developed a system that decomposes existing industrial television (ITV) surveillance camera images into 30 still images per

Examined regions	Examined parameters	Recognition conditions
① Light blue frame	Brightness	If [Brightness] ≥ 100, luminescence; otherwise, darkened.
② Green frame	Brightness	If [Brightness] ≥ 200, luminescence; otherwise, darkened.

Fig. 2 Examples of recognition conditions for multiple examined regions

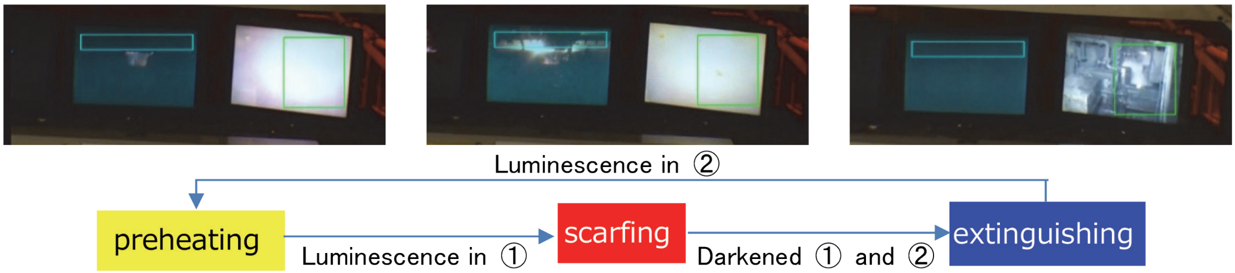


Fig. 3 Examples of decisions by multiple conditions

second, predicts the occurrence probability of lateral strip shift using a deep learning model for each image, processes the prediction results in real-time, graphs them, and displays them on the on-site monitor (**Fig. 6**). It has been evaluated as being capable of detecting the precursor of lateral strip shift 3 to 5 min before the lateral strip shift occurs and as good as on-site operators.

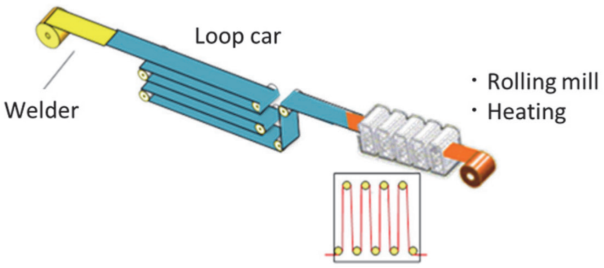


Fig. 4 An example of continuous processing line[2)]

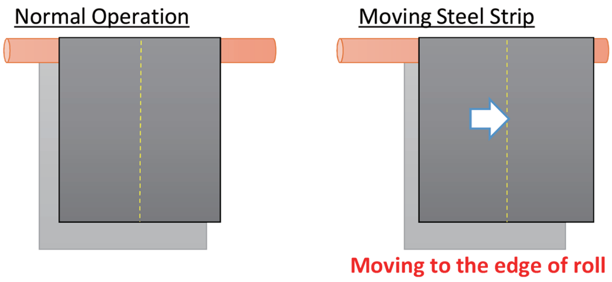


Fig. 5 An example of moving steel strip[2)]

Fig. 6 An example of predictive detection of moving coil[2)]

4. Automatic Identification Technology for Fracture Surfaces in Impact Test

4.1 Challenges in calculating fracture surface ratio in the impact test

Among the various tests for evaluating the quality of steel products, there are tests in which operators derive test results based on visual observation of test specimens. Such tests often involve visual judgment and require high technical skills. It is important to suppress variations between operators. For tests that require a high level of skill in visual observation, we are working to improve the level of quality management by promoting automatic acquisition of test results and automatic import of test results into quality management systems. As an example, the development of automatic fracture surface identification technology for impact tests is described below.

Impact tests such as the Charpy impact test and the Drop Weight Tear Test (DWTT) are conducted to evaluate the toughness of metals such as steel products.[3)] As outlined in **Fig. 7**, a notch is cut in advance in a bar specimen. When hit with a swinging hammer, the specimen is fractured starting from the notch. The fracture surface (as shown in **Fig. 8** (a)) is examined to see whether it is ductile or brittle. The ductile fracture ratio or the ratio of the ductile fracture surface area to the total fracture surface area or the brittle fracture surface ratio or the ratio of the brittle fracture surface area to the total fracture surface area is then determined. (The ductile and brittle fracture ratios are collectively called the fracture surface ratios).

As mentioned above, before we derive the fracture surface ratio, we must determine which regions of the fracture surface are ductile and which are brittle. This judgment requires high technical skill and varies with operators. In addition, since the boundaries between the ductile and brittle surface regions are complicatedly shaped, it requires a long time to calculate the fracture surface ratio accurately.

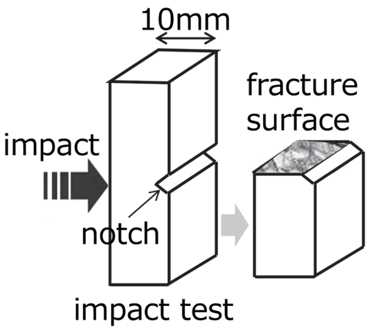


Fig. 7 Schematic of Charpy impact test

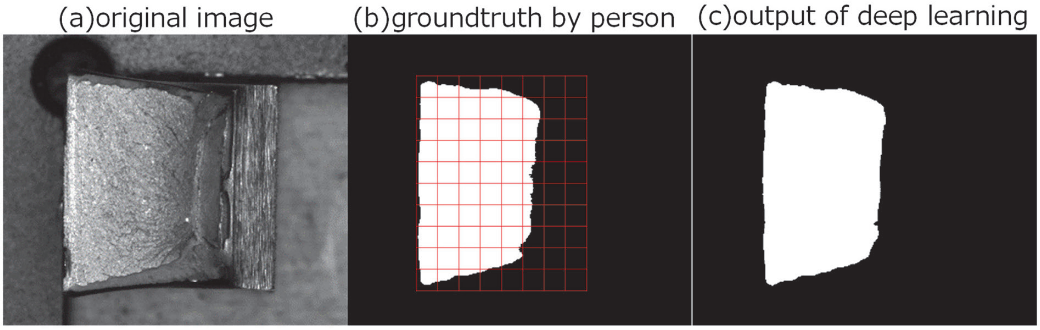


Fig. 8 An example of original Charpy fracture surface image (a), groundtruth of brittle surface given by a person (b), and output of deep learning (c) with respect to the same image (a).

Specifically, as an example of a conventional method, a square grid sheet (red framed in Fig. 8 (b)) is placed on the fracture surface image. Squares that correspond to ductile or brittle fracture surface regions are counted. The ratio of the ductile or brittle fracture square areas to the total fracture surface area is calculated as the fracture surface ratio. The burden of these tasks is a challenge.

4.2 Development of automatic fracture surface identification technology

Therefore, we developed a technology for automatically identifying fracture surfaces using deep learning. As a similar study, Reference 4) proposes a deep learning model that inputs a scanning electron microscope (SEM) image of a fracture surface and outputs the types of fracture structures such as dimples, facets, and striations.[5)] Our development differs because it uses more macroscopic optical microscope images and outputs the fracture surface ratio rather than the types of fractured structure.

As for the design of the deep learning task, we considered constructing a deep learning model to directly output fracture surface ratios from fracture surface images like that reported in Reference 4). To pursue the accuracy of fracture surface ratios, however, we developed a deep learning model[6)] that determines which fracture surface regions are brittle fracture surface regions on a pixel-by-pixel basis and outputs brittle fracture regions in a colored map format.

To train a deep learning model, it is necessary to prepare a large amount of correct groundtruth data. The correct groundtruth data for this task are composed of fracture surface images and region information indicating which fracture surface regions are brittle fracture surface regions. The latter region information must be depicted or marked by a model operator. This marking imposes a large workload.

Therefore, in this development, we did not create a deep learning model after preparing numerous images. Instead, once a certain amount of images and marking data were made available, we created a temporary deep learning model using the images and data. The temporary deep learning model was used to obtain temporary labels for the images collected afterwards, and the temporary labels were manually corrected if required.[6)] The marking load was thus reduced.

When capturing an image with a camera from above the fracture surface, we used an optical system that makes the image axisymmetric along the vertical axis of the fracture surface, which coincides with the camera's optical axis.[6)] This was designed to augment the training data by flipping the image vertically and horizontally.

In addition, during the design of the deep learning model the receptive field[7)] was optimized. The receptive field is the region's size that shows how far from the pixel in question the deep learning model considers when determining whether the given pixel belongs to a brittle fracture region. Even in the current work, people judge whether a certain point belongs to a brittle fracture region by looking at the surrounding region. This optimization corresponds to this judgment procedure by humans.

As a result of improving its accuracy by incorporating the above points, our deep learning model reached a level where it can make judgments equivalent to those of humans on more than 99% of the test specimens on the basis of the number of test specimens. (An example of the inference result by the deep learning model is shown in Fig. 8 (c)). The deep learning model will soon be put into practical use and contribute to improving fracture surface reading work efficiency.

5. Support of Checking Work after Steel Coil Product Packaging

5.1 Challenges in packaging steel coil products

Steel products such as steel coils are packaged and shipped after passing through the manufacturing, inspection, and repair processes. For packaging, the entire product is covered with metal sheets, paper, bands, etc., and multiple labels containing product information and precautions for transportation are affixed to the product (**Fig. 9**).

The contents of these packages are determined according to the characteristics of the products and the requests of customers. Therefore, there are a variety of packaging patterns. There are dozens of labels affixed to products for each packaging pattern. For this reason, we check whether the packaging has been done correctly according to the packaging and labeling patterns that change for each product. Overlooking mistakes in packaging and labeling can lead to the shipment of products that differ from the contracted ones, so careful checking is necessary. If only one worker checks, it is impossible to completely avoid oversights, even if the worker acts with a sense of urgency. Therefore, two or more workers visually checked the packaged products together. However, there was an awareness of the challenge that even with multiple workers involved in the checking process, the workload of conducting checks with concentration continuously maintained can be high for the workers.

5.2 Establishment of a support system for packaging checking work

We developed a technology that automatically checks the contents of a package using images of the packaged product taken with a camera attached to equipment or with a terminal camera. The challenge in image recognition is that in addition to the large number of combinations of packaging and labeling patterns, steel products vary in weight from several to tens of tons. It is important to develop an image recognition model that can handle many conditions, such as handling products of different sizes and accommodating changes in the shooting angle that change with camera installation conditions. Generally, to identify an object from an image, a method can be considered in which, after preprocessing, such as a filter, a feature quantity that characterizes the object to be detected is determined, and a rule for recognition is set based on the distribution of the feature quantity. However, if there are many variations in the images, such as the detection object or other background, designing and setting such features and rules often require a lot of time and effort.[2)] We decided to use deep learning to develop a model for identifying packaging contents. Furthermore, the identification model developed by our company has been improved to a practical

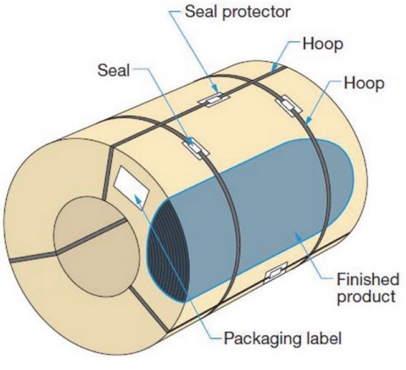


Fig. 9 Coil packaging

level of accuracy through preprocessing and other measures to improve product recognition accuracy. We have ultimately constructed a system that can support post-packaging checking tasks. As a result, by using a deep learning model as a second set of eyes, it is now possible for one worker to check the packaging, which previously required two or more workers.

The peripheral functions we have built this time, such as preprocessing and recognition processing, can be applied to various other tasks, so we have created common components that can be combined to suit specific tasks. Deploying these components to other tasks can shorten the time required to put other image recognition models into practical use.

6. Steel Product ID Recognition Technology[2)]

6.1 Challenges in reading steel product IDs

Our steel products are manufactured at our steelworks, then shipped by ships, etc., and delivered to customers with appropriate transit points in between. In this logistics series, it is necessary to accurately track the products and deliver them to their customers according to the ordered specifications. For this reason, products have their ID printed, stamped, or labeled on their surfaces or sides to identify them. Individual identification based on ID is also important in the multi-step steelmaking processes to process and transport semi-finished products (slabs, billets, coils, etc.) correctly within the steelworks.

Traditionally, ID recognition has been performed by human visual inspection. **Figure 10** shows an example of checking work at a transfer site for steel plate products.[2)] A worker visually reads the ID printed on the side of each product one by one and matches it with a list for each product pile. Even if the worker did not have to take the difficult posture shown in this photo, the workload was still a challenge because the worker was required to continue reading for long periods without losing concentration to read correctly.

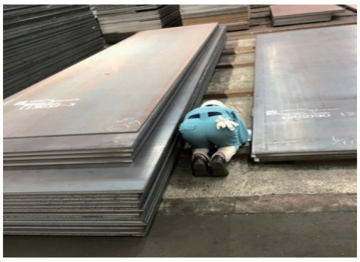


Fig. 10 An example of work to read IDs[2)]
The person in the center is reading ID printed on the side of each plate.

6.2 Development of automatic steel product ID recognition technology

To address this challenge, we have developed technology that uses a smartphone camera to photograph a product's or semi-finished product's ID and automatically recognizes the product ID from the photographed image.

The task of ID recognition is the recognition of letters and numbers. Unlike the environment where an Optical Character Recognition/Reader (OCR) is used to digitize paper documents, there are issues with the shooting environment, such as shadows, the inability to check the presence or absence of text due to insufficient brightness, and halation caused by the sunrise or sunset when shot outdoors. Furthermore, when printed directly on steel, recognition is also affected by the surface properties of the steel, such as rust. Developing a process to make the ID-printed regions visible by devising image processing could be a solution. Still, it is challenging to develop an algorithm that can be applied to all the various conditions, including the environmental disturbances mentioned above.

Therefore, we are moving ahead with the application of deep learning models. The advantage of deep learning is that once training data that corresponds to a variety of conditions is prepared, there is no need to design algorithms that can be applicable to those various conditions one by one manually, and training can be performed automatically. On the other hand, creating highly accurate deep learning models typically requires preparing tens of thousands of image data. To learn images, it is necessary to associate images with information about what characters appear where in each image (annotation). The burden and time required for this manual annotation work becomes a challenge. In addition, the types of IDs for products and semi-finished products are different for each of our products, and tens of thousands of learning data must be prepared for over 100 product types. The total workload and construction period will be enormous.

We have built a common steel product ID recognition foundation model to address the above challenges. As shown in **Fig. 11**, images and annotation information for about ten character types developed in advance are aggregated and the foundation model learned all of them. Data was collected on the company-wide analysis platform NS-DIG™ and training was executed using KAMONOHASHI™ on NS-DIG™. As shown in Fig. 11, although the number of the types of characters is about 10, the model has learned various character types of finished and semi-finished steel products.

When transferring to other new character types, we have confirmed the effect of reducing the amount of training data of the new character type required to reach a certain accuracy by additional

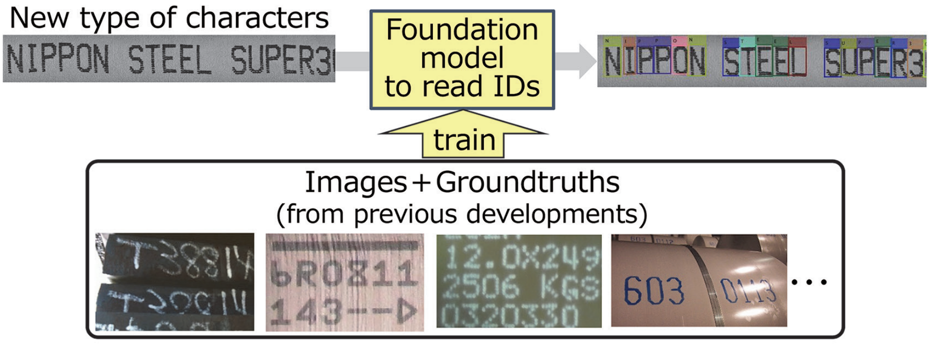


Fig. 11 A foundation model to read IDs on steel

training with the new character types (transfer learning) based on the foundation model. This foundation model can also be used to annotate the new character types. In other words, it is partially annotated by passing an image with the new character type through the foundation model. After that, humans can correct the results and annotate the remaining parts that could not be annotated, significantly reducing the annotation workload and time.

7. Accuracy Maintenance and Management of Image Recognition Model

7.1 Challenges in image recognition model in operation

AI models, such as deep learning models, find features based on vast amounts of data, learn rules, and make inferences accordingly. However, if the prediction target is data that is outside the range of the training data, there is a high possibility that the corresponding rule is not found and a wrong answer is given. As an example, we describe the training of an AI model to recognize the contents of coil packaging labels by installing a camera indoors. Suppose an AI model is trained using images of various packaging patterns and put into practical use. However, if the lighting installed indoors deteriorates during operation and is replaced, the brightness and other aspects of visibility may change. When such environmental changes occur, recognition accuracy generally decreases because data that is different from the data used during training is provided (**Fig. 12**). In this way, AI models tend to become less accurate when the characteristics of the data change during operation. However, if workers use AI models without noticing the decrease in accuracy, they may gradually stop using AI models that provide answers that deviate from actual facts. For AI models to continue to be effective after they are put into practical use, it is necessary to restore the accuracy of the AI model before a fatal drop in accuracy occurs.

7.2 Construction of system for accuracy maintenance and management of AI models

To ensure sustainably effective AI models, it is important to establish a foundation that repeats an improvement cycle that "monitors the accuracy of AI models, detects any decrease in accuracy early, identifies the cause of the decreased accuracy, and performs retraining." However, maintaining and managing the accuracy of image recognition models is more challenging than AI models that use structured data. Therefore, the department that builds the model has to work closely with the business department that deploys the model to build the mechanism. For example, in the case of an AI model that uses structured numerical data, where each numerical value has its own meaning, it is possible to obtain actual result data from sensor data, etc., and automatically manage the discrepancy between predictions and actual results. On the other hand, AI models that use image data may have recognitions based on human visual information as actual result data. Specifically, when an AI model that detects the presence of a person infers that a person is present, only the worker who has confirmed the result can assess whether or not a person is actually in the image. As described above, when it comes to image recognition, it is sometimes difficult to automatically manage the discrepancy between predictions and actual results without human intervention. One possible way to manage the accuracy of AI models put into practical use is to calculate the accuracy daily and manually analyze the latest accuracy trend. However, there are limits to this kind of manual monitoring and management of all of the various AI models.

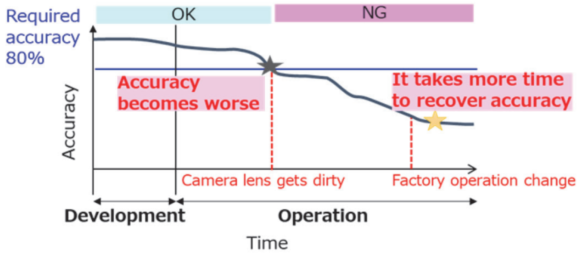


Fig. 12 Changing accuracy of AI model in operation

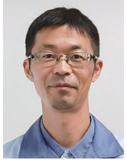
Therefore, we built a system for accumulating the results of an AI model as data and monitoring the accuracy of the AI model according to the accumulated result data and a system for detecting the accuracy decrease of the AI model and supporting the accuracy recovery of the AI model. In this way, we worked to simplify the accuracy management of the image recognition model. These systems are built on the edge computing platform AIRON-EDGE™. We are expanding the use of the image recognition model. Previously, accuracy maintenance and management were mainly performed manually. However, the introduction of automated systems has been increasing the efficiency of model operation and maintenance.

8. Conclusions

In this technical paper, we have described our initiatives to automate work and improve productivity using image processing technology. The technologies discussed here include technologies to support video-based operation analysis by image processing and to automate the visual judgment by operators, such as precursor monitoring of product transportation abnormalities, fracture surface identification, confirmation of steel coil products after packaging, and steel product ID recognition. These technologies contribute to saving labor and improving efficiency in various tasks. These operational tasks are supported by a system that maintains and manages the accuracy of the image recognition model, which was subsequently described. Image-based AI technology continues to advance day by day. We will continue to adopt innovative technologies and promote business transformation.

References

1) Work analysis/business optimization software OTRS (Broadleaf Co., Ltd.) https://www.broadleaf.co.jp/products/other/otrs/
2) Ito et al.: Bull. Iron Steel Inst. Jpn. 26 (2), 69–74 (2021)
3) For example, JIS Z2242, ASTM E23, ISO 148-1, API 5L3, etc.
4) Yamagiwa: J. Soc. Mater. Sci., Jpn. 69 (9), 644–649 (2020)
5) Komai, K. (ed.): Mechanical Materials Science. 9th edition. Kyoto, Society of Material Science, Japan, 1999, p. 173–182
6) Japanese Patent Application Publication 2023-42445. March 27, 2023
7) Okatani: Journal of the Japanese Society for Artificial Intelligence. 31 (2), 169–179 (2016)



Koji HIRANO
Ph.D., General Manager, Head of Dept.
Production Management Research Dept.
Intelligent Algorithm Research Center
Process Research Laboratories
20-1 Shintomi, Futtsu City, Chiba Pref. 293-8511

Yuto HIRATA
Senior Manager, Head of Section
AI Solution Section
ICT System Planning Dept.-II
Information & Communication Technology Div.

Sachi ISHITANI
Manager
AI Solution Section
ICT System Planning Dept.-II
Information & Communication Technology Div.

Tadashi HYODO
Expert, IE Consulting Group
Hirohata Systems Engineering Center
Setouchi Systems Solution Department
Steelmaking System Solutions Division
Steelmaking System Solutions Bureau
NS Solutions Corporation