Technical Report

Development of Automatic Steel Surface Inspection System toward Digital Transformation

Yusuke KONNO*
Hideyuki SHIIBARA
Nobuhiro FURUYA
Takamichi KOBAYASHI
Jun UMEMURA

Abstract

The steel manufacturing process consists of numerous steps from blast furnace to rolling, annealing, plating, and various surface defects may occur due to inclusion in the molten steel or abnormalities in each manufacturing process. If a surface defect that occurs in one process propagates to downstream processes and is discovered in the final process, a large amount of incompatible material is generated between the originating process and the final process. Therefore, visual inspection at the exit of each process has been indispensable, and has traditionally relied on highly skilled inspectors. However, visual inspection lacks quantification and reproducibility of results, and it is difficult to transfer skills. To solve this problem, we developed an automatic surface inspection system to promote automation of visual inspections. By installing an automatic inspection device at the exit of each process and integrating its operation, it has become possible to detect operational and facility abnormalities at an early stage, prevent the generation of large amounts of incompatible materials, and improve yields by reducing defects.

1. Introduction

The steelmaking process can be broadly divided into upstream processes from raw materials such as iron ore and coal through molten iron and molten steel to semi-finished products such as slabs and billets and downstream processes from rolling through heat treating and plating of semi-finished products to finishing to final products. In the case of strip products, the downstream processes consist of hot rolling, pickling, cold rolling, annealing, plating. The time required from slabs to final products is about one week at the shortest. Surface defects caused in the respective processes remain in the final products, except for minor ones. It is desirable that defects caused in the respective processes should be detected by inspection at the exit of the respective processes and that their causes should also be removed.

For this purpose, inspectors have been traditionally stationed at the exit of each process to conduct visual surface inspections (**Fig. 1**). The results of visual inspection are used for quality improvement. To improve the efficiency of data aggregation and management and activate improvement work, we are promoting the introduction of surface inspection systems to accelerate conversion to digital data. In Japan, the use of surface inspection systems began in the 1980s for the final products with little surface contamination and few surface disturbances.[1)]

Fig. 1 Surface inspection by human inspector

Since array-type imaging devices were not yet developed, initial inspection systems mainly used a spot beam as the source of light,

* Leading Researcher, Process Research Laboratories
20-1 Shintomi, Futtsu City, Chiba Pref. 293-8511

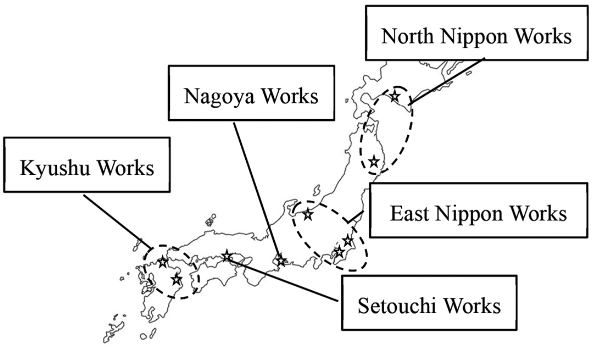


Fig. 2 Status of Nippon Steel's surface inspection system (2023)

scanned the surface with a galvanometer mirror, etc., and received the light through a light guide tube into a phototube. In particular, the method of using a laser as a light source and spatially filtering a diffraction pattern caused by a surface defect using a mask became widely popular in Japan in the 1990s.[2] At that time, analog signal processing was the main signal processing method. Defects were emphasized and detected by bandpass filter processing of one-dimensional phototube output signals and comparator thresholding.

In the 1990s, solid-state imaging devices such as CCDs became practical, and it became possible to image thin surface stains that were difficult to detect with laser methods. This camera imaging method has become mainstream since the 2000s. The standard method for signal processing is to first store a two-dimensional image in memory, then perform spatial filtering, binarization, labeling, feature extraction, and type and grade determination using machine learning methods.[3–5] This signal processing significantly improves the ability to distinguish harmless stains from defects, which extends the range of surface inspection systems to be applied to the processing line before pickling, such as hot rolling and continuous slab casting, where the laser-based systems cannot be applied.[6]

Nippon Steel Corporation began research and development of in-house surface inspection systems in the late 1990s. Since the first system was introduced in 1998, surface inspection systems have been deployed throughout the company. As of 2023, about 100 surface inspection systems are installed at the exit of continuous casters, in hot and cold strip mill processes, and at plate mills and wire rod mills, etc. (**Fig. 2**).

2. Purposes of Surface Inspection Systems

In the steel industry, steel surfaces are inspected for the quality assurance of products to be shipped and for quality control of the production processes.

Quality assurance tasks concerning surface defects are to confirm by inspection that surface defects detrimental to the use of steel products by customers are not included in products to be shipped and to judge whether or not to ship each product. Strip coil products are generally 1 to several kilometers in length, and 1 to 2 m in width. Multiple defects measuring several 100 μm to several meters in size may be present on strip coil surfaces. In the surface treatment process, the task of checking for all defects in steel strip moving at a speed of 10 to 20 km per hour and the task of stopping the strip, grinding the strip surface, and checking for hard-to-see surface irregularities defects exert heavy mental and physical loads on inspectors. Therefore, we have introduced these surface inspection systems to guide inspectors on defect locations and images and to alert them not to overlook defects.

At its steel plants in Japan, Nippon Steel daily manufactures steel products to meet the customer demand. Quality control tasks related to surface defects are divided into routine tasks and improvement tasks. Routine tasks manage the operating conditions of plants from continuous casting through rolling and annealing to plating, reduce the occurrence frequency of surface defects, decrease quality variations, and manage quality trends. Improvement tasks identify the causes of quality abnormalities, verify the effectiveness of countermeasures, and standardize the countermeasures. The starting point of improvement tasks is to identify quality abnormalities. The challenge is to improve the efficiency of information gathering to accurately recognize the distribution of surface defects in coils. The purpose of introducing surface inspection systems is to store data on the locations and causes of surface defects in the coils manufactured on a daily basis, as well as the names, locations, and images of defects. Surface inspection systems are currently used to improve quality abnormalities and manage quality trends to identify conditions for manufacturing new products.

3. Surface Inspection System Configuration

3.1 System configuration

Figure 3 shows the configuration of a surface inspection system developed by Nippon Steel. General-purpose rack-mount x86 computers are used to minimize dependencies on dedicated hardware. In addition, all data such as detecting results and processing parameters are stored on a relational database management system (RDBMS), making it easy to utilize additional data processing as described below.

3.2 Optical system

Various surface defects are caused on steel surfaces. To visualize the defects as images, it is necessary to appropriately design an optical system consisting of light sources and cameras according to the optical characteristics of the defects. Additionally, harmless stains and scale exist on the surface. These can cause false defects (overdetection). Although it is possible to remove false defects in post-processing, it is preferable to remove them at the imaging stage if possible, as this increases the computer load and causes learning data contamination. To this end, we have developed a variety of optical systems, including the multiple camera system shown in **Fig. 4** as a representative example. We use them depending on the process and characteristics of the object to be inspected.

3.2.1 Multiple camera system

The multiple camera system was our first optical system devel-

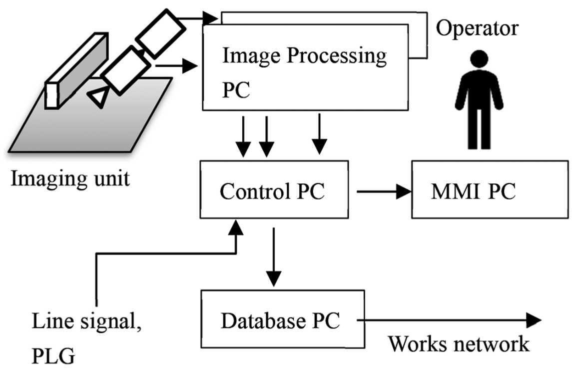


Fig. 3 Diagram of developed surface inspection system

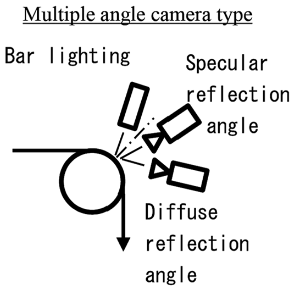
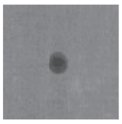
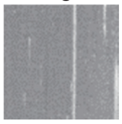
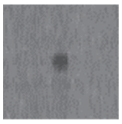


Fig. 4 Major type of inspection optics

Table 1 Relationships between luminance and defect type

	Dirt	Sliver	Scratch
Specular reflection	Dark 	Dark 	Bright 
Diffuse reflection	Dark 	Bright 	Dark 

oped for sheet product inspections. It illuminates the steel strip with a rod-shaped light source (initially a fluorescent lamp and later a metal halide light source-fiber light guide combination). The same location on the steel strip is observed at two or more different angles, which produces specular and some diffuse images. The combination of luminance information is used to improve the defect type determination performance. This is because early image processing had poor features relating to the shape of defects, and it was necessary to compensate by increasing luminance information. **Table 1** shows typical defect and luminance combinations.

3.2.2 Initiatives to improve reliability

Because the optical system is installed on a production line, it is designed to withstand harsh environments such as dust, high temperatures, and splashing water, and to minimize damage from abnormal operations such as strip breakage. Air purging is the basic method for controlling dust. System components in hot rolling lines are housed in water-cooled jackets.

Fluorescent tubes and metal halide lamps have relatively short lifespans, requiring regular illuminance measurements and lamp replacement every few months. With the practical application of power LED lighting, lifespans of several years can now be achieved. This dramatically reduced the maintenance effort. However, care must be taken for device cooling. Nippon Steel is actively moving towards the water cooling of such devices.

3.3 Defect detection using image processing

Image processing is used to detect surface defects from images captured by the optical system described in Section 3.2. The defects are further classified into defect types (**Fig. 5**). Preprocessing involves removing uneven pixel values caused by lenses and lighting (shading correction) and adjusting luminance contrast. These steps are performed to improve the accuracy of subsequent processing, remove factors that have an adverse effect, and emphasize the contours of defect regions.

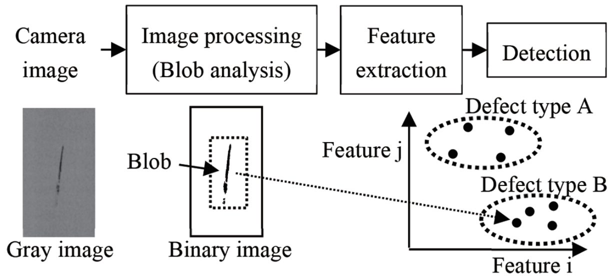


Fig. 5 Diagram of defect detection

In binarization, each pixel is divided into background and foreground (defect candidate) pixels using a preset threshold value. There are two types of threshold values: an upper threshold and a lower threshold. Pixels with greater than or equal to the upper threshold or less than or equal to the lower threshold are considered as the foreground. The other pixels are considered as the background. The threshold values may be determined dynamically based on the variance of pixel values in the image to be processed, or they may be fixed for all images. In either case, the threshold values are uniform for the entire images to be processed. If uneven pixel values such as shading remain in the image, an incorrect result will be produced. Therefore, it is necessary to accurately remove unevenness of pixel values through preprocessing.

The foreground pixels obtained by binarization are grouped by labeling. Grouped pixels are generally called blobs, and these blobs are defect candidates. Therefore, it is necessary to adjust the binarization threshold so that the defective region to be detected is in the foreground. If the threshold is set loosely, a large number of defect candidates can be obtained, therefore the amount of calculation in the subsequent stages increases with the number of defect candidates. This may delay the start of processing subsequent images. If this delay continues, a congestion state occurs, and the image buffer waiting to be processed overflows, resulting in an uninspected region. In other words, the binarization threshold must be determined by making a trade-off between the occurrence of undetected defects and the occurrence of uninspected regions.

The feature values are calculated for the blobs obtained through labeling. By using the subsequent defect discrimination logic based on these feature values, it is determined whether each blob is a defect or not. If the blob is determined as a defect, its type and severity are also determined.

In addition to the general feature values of machine vision, we design dedicated feature values for discriminating specific defects. In addition to the feature values, it is useful to optimize unique filter processing and threshold processing based on prior knowledge of the shape and location of defects. Such dedicated processing facilities are difficult to achieve with commercially available inspection systems. We recognize this as one of the most valuable advantages achieved by in-house inspection systems. In other words, knowledge and expertise about defect characteristics can be implemented into inspection systems freely and at low cost, thereby improving the performance of their inspection systems quickly.

Defect candidates (blobs) are determined using the calculated feature values. Since each defect candidate is a point in a multidimensional space defined by the feature axes, determination is achieved by setting boundaries in this feature space. IF-THEN rules and machine learning methods are commonly used to set the boundaries. The IF-THEN rule is a rule that has feature values and their ranges in the conditional part. Complex boundaries can be achieved by nesting the conditional parts. Machine learning methods can calculate the optimal boundaries by learning training data consisting of feature values and their correct answers. One of the most common machine learning methods is the Support Vector Classifier, which determines the boundaries from the distribution of training data, rather than fitting the boundaries to a parameterized function, in order to maximize the margin between data and boundaries.[7] Furthermore, by introducing a kernel function, it is possible to perform the procedure of converting the feature quantities into a different space more suitable for determining the boundaries without the expensive calculation of converting the feature values. By using the radial basis function (RBF) as a kernel function, it is possible to project feature values onto a nonlinear infinite feature space and to determine the complex boundaries. The concept of constructing defect discrimination logic using these logic tables and support vector machines will be described in Chapter 4.

3.4 Acceleration of software processing

As mentioned above, inspection systems have gained practicality by implementing processing using software that runs on general-purpose computers. This is because software can implement complex processing that is difficult to implement with dedicated image processing hardware, and it also allows the processing to be changed or enhanced to more advanced processing when necessary. Additionally, by developing and deploying our own in-house inspection systems, we were able to make changes and enhancements from the initial implementation by ourselves. On the other hand, as the processing facilities increase and become more complex, there is a problem in that the processing time increases. Although the software can be modified, the computational resources required to execute it remain constant. This inevitably increases the processing time.

Computer performance has improved tremendously in recent years. For example, the number of CPU cores has reached several dozen, and terabytes of memory capacity are common. The Pentium Pro and Pentium II were commonly used when Nippon Steel put its inspection systems into practical use. These processors were single-core devices.[8, 9] Therefore, in addition to using dual CPUs to increase the processing speed of a single computer, we also actively utilized a Single Instruction for Multiple Data (SIMD) instruction set as a means of further increasing speed. Furthermore, by optimizing the alignment of memories that are accessed frequently, such as image buffers, we can use high-speed data transfer instructions that do not go through buffers, and we can tune the CPU and memory at the machine language level using assembly languages. We have built software that maximizes the performance of CPUs and memories.

Until around 2010, improvements in CPU operating frequency automatically improved the processing speed, but it was said that the frequency improvement had reached a plateau, and virtual threading and multi-core technologies emerged in its place.[8, 9] To take advantage of such technology, software that supports multi-threading is required. As mentioned above, since it was assumed that execution would be performed on dual CPUs, the software configuration was designed to be able to be executed in multi-threads from the beginning. Since the design was unified to allocate inputs from multiple cameras to threads on an image-by-image basis, the processing speed increased with the progress of multi-core technology. By simply increasing the number of multiplexed systems, it was possible to obtain the benefits of multi-core technology, such as increasing the sophistication of processing content and processing multiple cameras on a single server PC.

Graphic processing units (GPUs) were mainly used to calculate the RGB values of the pixels to be displayed. But due to requests from the game and CG sectors, GPUs were equipped with such functions as calculating coordinates in three-dimensional space and shading processing based on the distance from the viewpoint.[10] Furthermore, from around 2000, these processes became programmable rather than fixed functions, and this capability is utilized in applications other than graphics computations, which is called general-purpose computing on graphics processing units (GPGPUs).[10] At that time, we considered implementing some of the image processing performed by the CPU on the GPU, but there was still a limit to the processing that could be implemented on the GPU, and it was necessary to cooperate with the CPU. This was not realized because real-time processing was difficult due to the time required to transfer images between the respective memories. However, due to the recent advances in GPUs and the development of implementation and execution environments such as NVIDIA's CUDA that support GPGPUs, real-time inspection processing using GPUs has been utilized since 2018. This has made it possible to perform more complex image processing, such as using higher-resolution images, processing multiple cameras, and applying it to high-speed production lines, contributing to improved defect detection accuracy and an expanded range of defects to be detected. With the application of deep learning to the defect detection and discrimination processing described below, GPUs will be used more than ever before.

3.5 Utilization of deep learning

In order to meet the high quality requirements in recent years, it is necessary to further improve the accuracy of defect detection and discrimination. This can be achieved by continuing the current policy of developing dedicated image processing and feature values individually according to the needs of each process. On the other hand, if technology can be put into practical use that efficiently improves the accuracy of the large number of inspection systems in use, it will become possible to quickly contribute to the advancement of quality control in steel production and the expansion of business fields.

Deep learning is a neural network with a large number of layers. Although the basic technology follows 3-layer networks, backpropagation learning, and neocognitrons from the 1980s, tasks that exhibit human-level or higher accuracy have emerged mainly in image recognition and natural language processing by overcoming difficulties in learning with deep layers, the development of computers and software capable of learning large-scale networks, and the collection of learning data using the internet.[11]

Our research into the application of deep learning to defect detection and discrimination from images began in 2016. Using a deep learning model which discriminates the type of defect for each pixel (**Fig. 6**), we can achieve defect detection and discrimination processing using conventional image processing and feature quantities. We are promoting application as production facilities. This shows that deep learning can achieve sufficient accuracy, but this requires a large amount of training data. To maintain that accuracy, training data must be continuously collected and relearning must be continuously performed. We found that conventional image processing and feature-based methods should be used in combination for defects

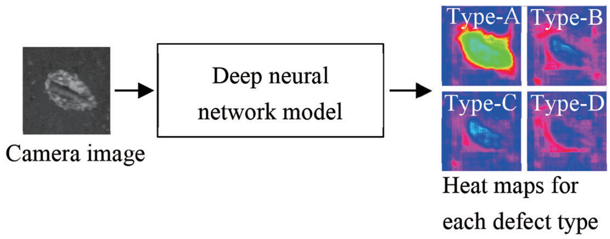


Fig. 6 Structure of deep-learning segmentation model

that occur infrequently or for defects with clear characteristics.

As a means of collecting large amounts of training data, we are collecting training data company-wide and promoting the development of company-wide models. By aggregating the annotated results of images collected using inspection systems installed at our steelworks across the country, a larger amount of training data can be obtained than from a single inspection system. Additionally, by using the developed company-wide models for inspection at the respective steelworks, it is possible to standardize inspection standards and levels.

Under this policy, after standardizing defect types and annotation standards for steel sheets and plates based on images, training data was aggregated and a company-wide model was developed. As a result of studying its application to the pickling process, the detection and discrimination accuracy were significantly improved compared to conventional methods. The decision has already been made to adopt it as production facilities, and we will gradually roll it out to each steelworks.

For tasks that require large amounts of training data, we are also considering applying weakly supervised learning that can use training data with less information. Currently, segmentation tasks are used. This requires a high workload for annotation as it is necessary to colorize pixels in defect regions according to the type of defect. To address this, we have confirmed that accuracy can be improved by also using training data with only the defect type for each defect region in the image.

Additionally, to maintain the accuracy of deep learning models, it is necessary to continually collect training data and retrain the model. It is effective to collect learning data throughout the company. In addition, it is more efficient to invest in computers equipped with high-speed GPUs required for model training if they are placed on a company-wide platform and shared. We are currently proceeding with the development with the aim of constructing such an MLops environment and operating it on Nippon Steel's DX platform NS-DIG™.[12)]

3.6 Data utilization

It was mentioned that surface inspection systems are used to guide inspectors on defect locations and images to prevent defects from being overlooked and also used to improve quality abnormalities and manage quality trends.

The quality of steel is the result of the composition of molten steel and the operation that goes through multiple processes. The process where defects appear on the steel surface may be different from the process where the defects are caused. Therefore, to improve steel quality analysis is performed using operational data for each process and data from the surface inspection systems.

Defects in steel change their appearance every time they pass through each process. Therefore, when a quality abnormality occurs, it is necessary to track specific defect image data for each process and confirm them in a list. This enables the quantification of the appearance change of the defect and identification of the cause of the defect occurrence. We have also been developing functions necessary for on-site improvement activities.

Data from domestic surface inspection systems is stored in a relational database (RDBMS), and some of the data is shared throughout the company. It is used in combination with operational performance data from the steelmaking processes, data of various measuring devices, and steel test result data using Business Intelligence (BI) tools. Defects detected and manufacturing data collected in a process are fed back to the previous process to improve manufacturing conditions and are fed forward to the next process to improve operations and improve work setups.

The image data of the entire length of the steel material captured by the surface inspection systems in each process can be used to identify the process that causes defects by compressing the images in accordance with the rolling ratio and proceeding with the analysis, thereby improving the soundness of the steel surface properties.

4. Adjustment and Performance Maintenance of Surface Inspection Systems

In order to achieve the best performance of surface inspection systems, it is important to adjust them and maintain and manage them after installation at the respective plants. In the sections that follow, we explain the adjustment methods for the surface inspection systems in three steps of imaging, detection, and discrimination, and also discuss how to maintain the performance of the surface inspection systems.

4.1 Imaging: Optical system adjustment

The optical system is adjusted to capture images of the steel strip surface with quality suitable for defect inspection. Specifically, the lighting and camera positions, focus, imaging resolution, amount of overlap between adjacent camera fields of view, lens aperture, camera exposure time, etc., are adjusted to appropriate values for all steel types.

4.2 Detection: Binarization parameter adjustment

Binarization parameter adjustment is performed to detect defect candidate regions from the captured steel strip surface image. Defects are detected by adjusting the binarization luminance threshold described in Section 3.3 to an appropriate value for raw images of the region around defects collected by capturing surface images of product steel strip using the surface inspection system. Since defects that are not detected at this stage remain undetected, it is essential to adjust the non-detection ratio of harmful defects to zero at this stage.

4.3 Discrimination: Logic adjustment

The logic is adjusted that automatically judges the name and harmfulness degree of the defect from the image and feature values of the detected defect candidate region. First, discrimination flow is constructed. A block for discriminating shapes such as dots or vertical lines from the dimensions of defect candidates using IF-THEN rules, a block for discriminating the defect name by shape, and a block for discriminating the severities for each defect name are connected by using machine learning algorithms to construct the overall decision flow. At this point, if the engineer freely creates a discrimination flow, problems such as variations in accuracy and the difficulty of handing over the task to another engineer would occur. So, we established a standard process flow within the company.

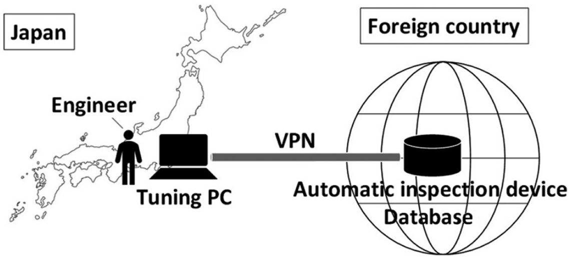


Fig. 7 Tuning and maintenance for overseas works

Next, training data is collected from data on defect candidate regions obtained by the inspection of the steel surfaces with a surface inspection system while the production is in operation. Defect candidate images are set with feature values and are used as training data by adding visually determined defect names and severities. The defect name and severity assigned here are important because they are linked to the accuracy of the discrimination logic and are set based on visual discrimination by an experienced inspector. The discrimination logic can be created by learning the training data collected in this way. This discrimination logic is applied to a surface inspection system to inspect product steel strip, data with incorrect discrimination results is collected, and the correct defect name and severity are assigned as additional training data. The additional training data is then used for additional learning. Accuracy is improved by repeating the cycle of updating the discrimination logic for the surface inspection system. The adjustment is completed when the target accuracy is reached.

When installing a surface inspection system at an overseas plant, a Nippon Steel engineer goes to the site for the initial start-up and performs the adjustment work directly on the surface inspection system. The subsequent maintenance and management are done remotely from within Japan using a virtual private network (VPN) (**Fig. 7**). This not only streamlines the maintenance and management of surface inspection systems at overseas bases, but also meets the demand for high-grade steel required by local production bases of Japanese companies.

5. Future Prospects

5.1 Optical system

Now that the full-length images of the defective regions can be tracked from continuous casting to the product process, product defects that were previously unknown are present, disguised as false defects in surface patterns in semi-finished products on continuous casting and hot rolling lines. To automatically detect such defects at the semi-finished product stage, higher resolution imaging is required, and illumination intensity and camera speed need to be increased.

5.2 Defect detection and discrimination processing

Defect detection and discrimination processing from captured images use a combination of methods that use conventional image processing and feature-based discrimination methods and deep learning models. This is because the conventional methods are advantageous when the characteristics such as defect shapes are clear and stable. On the other hand, deep learning models are increasingly applied to defects difficult to detect and discriminate by conventional image processing. On the other hand, deep learning models require large amounts of training data, and continuous retraining is necessary to maintain accuracy. Construction of a framework that does not depend on individual skills is planned.

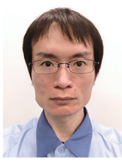
5.3 Data utilization

Utilizing defect data from surface inspection systems and image data of steel surfaces will improve the efficiency of routine tasks related to quality control and final inspection of steel. Using data from surface inspection systems with unified defect discrimination models, we will be able to share surface quality trends within the company and quickly respond to quality issues, thereby enhancing quality control and standardization at domestic and global plants.

References

1) Mashiko, Y.: Kogaku (Japanese journal of optics). 14 (6), 480 (1985)
2) Shimada, M., Morioka, Y.: CAMP-ISIJ. 3, 1330 (1990)
3) Yoshihara, A., Tanaka, H.: CAMP-ISIJ. 8, 370 (1995)
4) Okuno, M., Torao, A., Tomura, Y., Tsujimoto, M., Oi, K.: CAMP-ISIJ. 12, 286 (1999)
5) Hirata, T., Asano, K., Tomura, Y., Tanji, M., Adachi, S.: CAMP-ISIJ. 21, 1126 (2008)
6) Jonker, K.: Proceedings of International Surface Inspection Summit 2011
7) Cristianini, N. et al.: An Introduction to Support Vector Machine and Other Kernel-based Learning Methods. 2000
8) Ando, H.: Technologies Supporting Processors (in Japanese). 2011
9) Ando, H.: Introduction to Computer Architecture (in Japanese). 2018
10) Ando, H.: Technologies Supporting GPU (in Japanese). 2017
11) Okatani, T.: Deep Learning (in Japanese). 2015
12) Nippon Steel Corporation: Press release, https://www.nipponsteel.com/news/20190425_200.html

Yusuke KONNO
Leading Researcher
Process Research Laboratories
20-1 Shintomi, Futtsu City, Chiba Pref. 293-8511



Hideyuki SHIIBARA
Senior Manager
Instrumentation Engineering Section-II
Instrumentation Engineering Dept.
Systems & Control Engineering Div.
Plant Engineering and Facility Management Center

Nobuhiro FURUYA
Chief Manager
Systems & Control Technology Dept.
Systems & Control Engineering Div.
Plant Engineering and Facility Management Center

Takamichi KOBAYASHI
General Manager
Integrated Process Management Dept.
Technical Administration & Planning Div.

Jun UMEMURA
General Manager, Head of Div.
Instrumentation & Control Research Lab.
Process Research Laboratories