files
Jun 20, 2023
9394 words / 51803 characters

HIVMG_4_Plagiarism_Check

Sources Overview

19%

OVERALL SIMILARITY

1	link.springer.com INTERNET	5%
2	Daehyun Baek, Hyeong Soon Moon, Sang-Hu Park. "In-process prediction of weld penetration depth using machine learning-based molte... CROSSREF	3%
3	Rui Yu, Yue Cao, Heping Chen, Qiang Ye, YuMing Zhang. "Deep learning based real-time and in-situ monitoring of weld penetration: Whe... CROSSREF	<1%
4	www.mdpi.com INTERNET	<1%
5	doi.org INTERNET	<1%
6	www.ncbi.nlm.nih.gov INTERNET	<1%
7	Fengjing Xu, Yanling Xu, Huajun Zhang, Shanben Chen. "Application of sensing technology in intelligent robotic arc welding: A review", ... CROSSREF	<1%
8	Guang Yang, Ran Xu, Ye Wang, Yutong Zhu, Feng Ren, Chunyu Li, Jingyi Wu. "Pore-scale numerical simulations of flow and convective ... CROSSREF	<1%
9	Mingchong Han, Aiguo Tan, Jianwei Zhong. "Application of Particle Swarm Optimization Combined with Long and Short-term Memory ... CROSSREF	<1%
10	Qingli Li, Yiting Wang, Hongying Liu, Jianbiao Wang, Fangmin Guo. "A combined spatial-spectral method for automated white blood cel... CROSSREF	<1%
11	Gonen Singer, Maya Golan, Rachel Shiff, Dvir Kleper. "Evaluating the Effectiveness of Accommodations Given to Students with Learnin... CROSSREF	<1%
12	Khaki, Saeed. "Novel Deep Learning-Based Methods for High-Throughput Image-Based Plant Phenotyping and Large Scale Crop Yield ... PUBLICATION	<1%
13	academic.oup.com INTERNET	<1%
14	Polson, Jennifer Sara. "Leveraging Clinical Imaging and Machine Learning Algorithms to Characterize Acute Ischemic Stroke Patients f... PUBLICATION	<1%
15	Weixi Wang, Satoshi Yamane, Qi Wang, Liang Shan et al. "Visual sensing and controlling of the keyhole in robotic plasma arc welding", ... CROSSREF	<1%
16	dokumen.pub INTERNET	<1%
17	hdl.handle.net INTERNET	<1%

18 www.researchgate.net
INTERNET
<1%

19 "Transactions on Intelligent Welding Manufacturing", Springer Science and Business Media LLC, 2019
CROSSREF
<1%

20 uos-app00353-si.soton.ac.uk
INTERNET
<1%

21 Xiaobo Zhang, Xiangguo Xu, Yuxian Zhu. "An improved time delay neural network model for predicting dynamic heat and mass transfe...
CROSSREF
<1%

22 Ziyu Zhao, Weili Zeng, Zhibin Quan, Mengfei Chen, Zhao Yang. "Aircraft Trajectory Prediction Using Deep Long Short-Term Memory Net...
CROSSREF
<1%

23 Zongyao Chen, Jian Chen, Zhili Feng. "Welding penetration prediction with passive vision system", Journal of Manufacturing Processes...
CROSSREF
<1%

24 deepblue.lib.umich.edu
INTERNET
<1%

25 Chaekuk Na, Sun-Pil Kim, Hyo-Gyoung Kwak. "Structural damage evaluation using genetic algorithm", Journal of Sound and Vibration, ...
CROSSREF
<1%

26 Han, Taihao. "Development and Application of Ensemble Machine Learning Algorithms to Enable a Priori, High-Fidelity, and Prompt Pre...
PUBLICATION
<1%

Preprint source
27 Hyo Chang Seo, Hyun Bin Kim, Gi-Byoung Nam, Min Soo Cho, Segyeong Joo. "Real-Time Indirect Continuous Arterial Blood Pressure M...
CROSSREF POSTED CONTENT
<1%

28 Shan-Ben Chen. "Real-Time Control of Weld Pool Dynamics During Robotic GTAW", Lecture Notes in Electrical Engineering, 2009
CROSSREF
<1%

29 www.biomedcentral.com
INTERNET
<1%

30 Jan Adamowski, Christina Karapataki. "Comparison of Multivariate Regression and Artificial Neural Networks for Peak Urban Water-D...
CROSSREF
<1%

31 Safa Riyadh Waheed, Mohd Shafry Mohd Rahim, Norhaida Mohd Suaib, A.A. Salim. "CNN deep learning-based image to vector depictio...
CROSSREF
<1%

32 Yang Yang, Xinyu Chen, Liping Ran, Yunfei Wu, Xurun Yu, Zhaodi Dong, Fei Xiong. "Formation of Protein Bodies and the Response to Ni...
CROSSREF
<1%

33 era.ed.ac.uk
INTERNET
<1%

34 Chunkai Li, Jiaxin Wang, Yue Dai, Yu Shi. "Experimental validation of saliency maps for understanding deep neural networks for weld p...
CROSSREF
<1%

35 Guillermo Bestard, Renato Sampaio, José Vargas, Sadek Alfaro. "Sensor Fusion to Estimate the Depth and Width of the Weld Bead in R...
CROSSREF
<1%

36 Jun Lu, Yang Zhao, Xiaoyu Chen, Jing Han, Zhuang Zhao. "Detection of Reinforcement of Multi-bead and Multi-layer Weld in Additive M...
CROSSREF
<1%

37 Mohammad H. Alshayeji, Sa'ed Abed, Silpa ChandraBhasi Sindhu. "Two-stage framework for diabetic retinopathy diagnosis and disea...
CROSSREF
<1%

38 Zhi-Jun Zhao, Sang-Ho Shin, Dae-Geun Choi, Sang-Hu Park, Jun-Ho Jeong. "Shape-Controlled 3D Periodic Metal Nanostructures Fabri...
CROSSREF
<1%

39 stax.strath.ac.uk
INTERNET
<1%

40 www.open.ac.uk
INTERNET
<1%

41 Do-Yeon Kim, Tae-Kyung Lee, Chul-Jin Park, Sang-Jik Lee, Hae-Do Jeong, Hyoung-Jae Kim. "Evaluation of Cutting Ability of Electroplat...
CROSSREF
<1%

42 Dukun Ding. "Research of neural network for weld penetration control", Applied Mathematics and Nonlinear Sciences, 2022
CROSSREF
<1%

43 G.R.B. Ferreira, H.V.H. Ayala. "Improved Image-Based Welding Status Recognition with Dimensionality Reduction and Shallow Learnin...
CROSSREF
<1%

44 Jung-Dae Lee, Dong-Hwan Park, Rae-Young Kim. "Novel Variable Switching Frequency PWM Strategy for a SiC-MOSFET-Based Electric ...
CROSSREF
<1%

45 Kunxi Xue, Zhijiang Wang, Junqi Shen, Shengsun Hu, Yahui Zhen, Jie Liu, Dingyong Wu, Hui Yang. "Robotic seam tracking system base...
CROSSREF
<1%

46 Kyohei Nozawa, Satoshi Maki, Takeo Furuya, Sho Okimatsu et al. "Magnetic resonance image segmentation of the compressed spinal ...
CROSSREF
<1%

47 Lecture Notes in Computer Science, 2015.
CROSSREF
<1%

48 X.F. Liu, C.S. Wu, C.B. Jia, G.K. Zhang. "Visual sensing of the weld pool geometry from the topside view in keyhole plasma arc welding", ...
CROSSREF
<1%

49 archive.org
INTERNET
<1%

50 scirp.org
INTERNET
<1%

51 www.iieta.org
INTERNET
<1%

52 www.pure.ed.ac.uk
INTERNET
<1%

53 "Computer Vision – ECCV 2018 Workshops", Springer Science and Business Media LLC, 2019
CROSSREF
<1%

54 Alison O'Shea, Gordon Lightbody, Geraldine Boylan, Andriy Temko. "Neonatal seizure detection from raw multi-channel EEG using a full...
CROSSREF
<1%

55 Christian Taylor, Jonathan Guy, Jaume Bacardit. "Prediction of growth in grower-finisher pigs using recurrent neural networks", Biosys...
CROSSREF
<1%

56 Denis, Etienne. "Inductive Knowledge Graph Completion: A Graph Neural Network Approach and the Need for Complimentary Methods...
PUBLICATION
<1%

57 Han, Xu. "Registration of Images with Pathologies.", The University of North Carolina at Chapel Hill, 2021
PUBLICATION
<1%

58 Jaeyong Ju, Daehun Kim, Bonhwa Ku, David K. Han, Hanseok Ko. "Online multi-object tracking with efficient track drift and fragmentat...
CROSSREF
<1%

59 Lecture Notes in Electrical Engineering, 2013.
CROSSREF
<1%

60 Liangrui Zhang, Xi Zhang, Jing Hu, Mingzhou Liu, Lin Ling. "Integrated application model for visual detection of welding quality based ...
CROSSREF
<1%

61 Luo, Hongbo. "Colorectal and Ovarian Cancer Imaging Using Optical Coherence Tomography and Photoacoustic Tomography", Washi...
PUBLICATION
<1%

62 Moon Seong Rang. "APPLICATION OF GREY MODEL AND ARTIFICIAL NEURAL NETWORKS TO FLOOD FORECASTING", Journal of the ...
CROSSREF
<1%

63 N. Chandrasekhar, M. Vasudevan, A. K. Bhaduri, T. Jayakumar. "Intelligent modeling for estimating weld bead width and depth of pene...
CROSSREF
<1%

64 Rui Yu, Joseph Kershaw, Peng Edward Wang, Yuming Zhang. "How to accurately monitor the weld penetration from dynamic weld poo...
CROSSREF
<1%

65 Sheng Su, Juan Wang. "Machine learning prediction of contents of oxygenated components in bio-oil using extreme gradient boosting ...
CROSSREF
<1%

66 Soufiane Hamida, Oussama El Gannour, Bouchaib Cherradi, Hassan Ouajji, Abdelhadi Raihani. "Handwritten computer science words ...
CROSSREF
<1%

67 Wang Cai, Ping Jiang, LeShi Shu, ShaoNing Geng, Qi Zhou. "Real-time monitoring of laser keyhole welding penetration state based on d...
CROSSREF
<1%

68 Wu, Chuan Song, and J.Q. Gao. "Vision-based neuro-fuzzy control of weld penetration in gas tungsten arc welding of thin sheets", Inter...
CROSSREF
<1%

69 app.aws.org
INTERNET
<1%

Preprint source
70 arxiv.org
INTERNET
<1%

71 documents.mx
INTERNET
<1%

72 ebin.pub
INTERNET
<1%

73 epdf.pub
INTERNET
<1%

74 grdspublishing.org
INTERNET
<1%

75 hal-centralesupelec.archives-ouvertes.fr
INTERNET
<1%

76 iopscience.iop.org
INTERNET
<1%

77 krex.k-state.edu
INTERNET
<1%

78 research-information.bris.ac.uk
INTERNET
<1%

79 research.tees.ac.uk
INTERNET
<1%

80 static.frontiersin.org
INTERNET
<1%

81 uknowledge.uky.edu
INTERNET
<1%

82 www.arxiv-vanity.com
INTERNET
<1%

83 www.scribd.com
INTERNET
<1%

84 Sanghoon Kang, Kidong Lee, Minjung Kang, Yong Hoon Jang, Cheolhee Kim. "Weld-penetration-depth estimation using deep learning ...
CROSSREF
<1%

85 Ziquan Jiao, Haijian Qin, Xingyu Gao, Zhiqiang Feng, Yuanzhao Xu, Shanben Chen, Nanhui Shi, Wenjing Liu. "Image Processing and Fe...
CROSSREF
<1%

86 P. Ghanty. "Artificial neural network approach for estimating weld bead width and depth of penetration from infrared thermal image of ...
CROSSREF
<1%

87 S. Amano, Y. Tsujimura, T. Ogawa, T. Shibata. "Development of in-process welding torch position control system using AI technology", ...
CROSSREF
<1%

88 Zongyao Chen, Zhili Feng, Jian Chen. "Key Technologies of Intelligentized Welding Manufacturing", Springer Science and Business Medi...
CROSSREF
<1%

Optimization of weld penetration prediction based on weld pool image and deep learning approach in gas tungsten arc welding

Daehyun Baek[1, 2], Hyeong Soon Moon[1*], Parviz Kahhal[3] and Sang-Hu Park[3*]

[1] Precision Manufacturing and Control R&D Group, Korea Institute of Industrial Technology, Sasang-gu, Busan 46938, Republic of Korea

[2] Graduate School of Mechanical Engineering, Pusan National University, Geumjeong-gu, Busan 46241, Republic of Korea

[3] School of Mechanical Engineering, Pusan National University, Geumjeong-gu, Busan 46241, Republic of Korea

Corresponding Author, Hyeong Soon Moon PhD.,
E-mail: hsmoon@kitech.re.kr, TEL: +82-51-309-7450

Co-corresponding Author, Sang-Hu Park Prof., PhD.,
E-mail: sanghu@pusan.ac.kr, TEL: +82-51-510-1011

Abstract

Gas tungsten arc welding (GTAW) is a popular technology for joining metallic parts with a high stability and quality. However, many GTAW-related processes are conducted manually. Manual welding is time-consuming, and the weld quality strongly depends on the skill of the welder. Although various automatic GTAW methods and systems have been developed, it remains challenging to control the weld quality because of the difficulties in predicting the weld quality, such as the penetration depth and backside bead geometry, during welding. Hence, this paper proposes an accurate and effective method for estimating the penetration depth through weld pool monitoring using a convolutional neural network (CNN) trained on weld pool images. The weld pool images contained several objects, each influencing the prediction accuracy. The CNN architecture and the structure of the fully connected layers (FCLs) also affected the prediction accuracy. To optimize the performance of the estimation model, the effects of each object in the weld pool image and the structure of the CNN architecture were analyzed and evaluated. The structure of the FCLs that outputted a quantitative penetration depth was optimized and evaluated through hyperparameter tuning. With the proposed method, the optimized model could quantitatively predict the penetration depth; the mean absolute error was 0.0516 mm, with an R^2 value of 0.998. An accurate prediction of the penetration depth can be applied to real-time weld quality control to ensure a sound weld back bead.

Keywords

Weld pool monitoring, Deep learning, Weld penetration estimation, Convolutional neural network, Hyperparameter tuning

1. Introduction

Gas tungsten arc welding (GTAW) is a popular welding technology for joining metal parts. Owing to its stability and high quality, GTAW has played an important role in applications such as shipbuilding, pipes, and pressure vessels. It is typically used for thick-plate piping to ensure sound welding quality, and welding is performed manually by highly skilled welders. In recent decades, problems, such as shortage of welders, aging, and time inefficiency, associated with manual welding have necessitated the automation of the GTAW process [1]. Therefore, GTAW processes have been replaced by automatic welding equipment.

For a completely robust automation of the welding process rather than a partial welding automation method involving operator intervention, the welding quality must be monitored and ensured during the welding process. In large-diameter pipes and long-term welding cases, the welding environment varies continuously during the welding process. To obtain a good welding quality in a varying welding environment, the welding parameters must be adaptatively controlled based on the feedback of the current welding quality.

In some studies, methods for monitoring and analyzing the welding parameters, such as the welding current, voltage, and heat input, using various sensors have been proposed [2–5]. Although these methods can respond to constantly varying welding environments, it is difficult to reflect the actual welding state because the measurement made is indirect. The acquired welding information is converted from raw sensor data, which can cause calibration errors. In addition, because data processing is associated with human subjective error, it is difficult to determine whether the actual welding conditions are fully reflected.

The welding process has complex and nonlinear characteristics; thus, it is difficult to optimize the prediction model for welding quality. Machine learning methods, such as support vector machines (SVMs) [6, 7] and artificial neural networks (ANNs) [8-11], can compensate for the nonlinear characteristics of welding quality predictions, resulting in a high accuracy. However, these methods require the acquisition of information from sensors, which inevitably requires complex setups, and there are many restrictions on their use in actual welding processes.

In general manual welding, skilled welders have a knack for predicting the penetration and backside bead states by observing the top-side weld pool. Hence, weld pool image-based methods for predicting the penetration depth or bead geometry have been proposed; these methods only require cameras. Based on weld pool monitoring methods using charge-coupled devices (CCDs) and complementary metal-oxide semiconductors (CMOS), methods for extracting the shape information of the weld pool have been proposed through image processing [12-15]. Image processing-based methods are more direct than prediction methods based on obtaining welding information indirectly. However, these methods are sensitive to changes in the welding environment. Recently, with the development of graphic processing units (GPUs) and deep learning architectures, deep learning approaches, such as CNNs, have been used to predict the weld quality based on weld pool monitoring.

Cheng et al. proposed a method for predicting the penetration state by performing model training on the shape of a laser stripe irradiated on a weld pool surface using a CNN [16]. Keshimiri et al. applied a deep neural network (DNN) to estimate the weld bead geometry. Using a CNN-based deep learning architecture, Li and Cheng et al. predicted the weld penetration state and bead width from dot-patterned laser reflection images of a weld pool [17, 18]. Nomura proposed a prediction model for the burn-through and weld penetration depth in gas metal arc welding (GMAW) using a deep learning model [19]. Image segmentation methods [20, 21] have been applied to accurately extract weld pool geometry from images. Cai [22] extracted the key hole and weld pool geometry information associated with laser welding. Yu [23] recognized the weld pool geometry in GTAW using the U-Net architecture. These approaches can compensate for the restrictive characteristics of image processing-based methods. A deep learning-based prediction model can directly obtain target information without user intervention or preprocessing of the input data, exhibiting a higher prediction accuracy than conventional weld quality estimation models.

However, a quantitative prediction of the welding quality parameters, such as the penetration depth, is essential for an accurate control of the penetration, particularly full and partial penetrations in root-pass welding, using the predicted welding quality as feedback. Therefore, in this study, we developed a deep learning model that can measure the quantitative penetration depth with a high accuracy using a CNN trained on weld pool images. Optimization through training of the deep learning model was performed to improve the CNN model, including the optimization of the weld pool images as input data and of the deep learning architecture through hyperparameter tuning.

Fig. 1 shows a weld pool image used for weld pool monitoring. The captured image contains not only the weld pool but also several objects that can affect the prediction accuracy, such as the arc light, arc light reflected shielding gas, weld bead, and welding torch. In addition, the structure of the deep learning architecture can affect the adequacy of the model training and the accuracy of target detection by the deep learning model. To improve the penetration prediction accuracy, it was necessary to derive an optimized image for penetration prediction by analyzing and evaluating the effect of each object in the weld pool image. The performance of the prediction model was compared and evaluated in terms of the input image and CNN architecture, and the most optimized weld pool image shape and deep learning architecture were determined for a quantitative penetration prediction.

The objective of this study was to construct an efficient and highly accurate prediction model capable of quantitatively measuring the penetration depth using only weld pool images. The proposed method uses only one camera to capture the front-side weld pool images in real time. In this study, VGG-16 [24], ResNet [25], DenseNet [26], and EfficientNet [27] architectures were used to build the penetration prediction model. The CNN architectures were trained on a dataset containing weld pool images and corresponding penetration depths under various welding conditions. The trained CNN models could quantitatively predict the penetration depth using weld pool images acquired by the camera. Finally, the performance of the proposed method was verified through a quantitative comparison between the actual and predicted penetration depths.

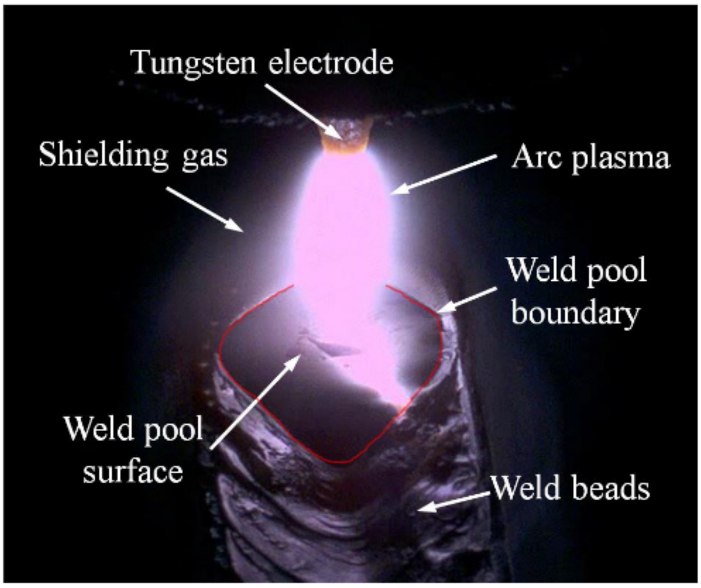

Fig. 1. Weld pool image captured by a weld pool monitoring setup

2. Weld pool monitoring setup

Fig. 2 shows a schematic of the weld pool monitoring setup. The monitoring module was connected to a welding torch, and weld pool images were acquired. The monitoring module contained a CMOS camera, optical lens, and a casing to prevent contaminants, such as fumes and spatters, from affecting the weld. During welding, the welding torch moved along the welding direction (x-axis). Table 1 lists the monitoring parameters.

Fig. 2 shows some of the images captured by the weld pool monitoring setup. Weld pool images were captured under various welding conditions including the voltage, current, and welding speed. The weld pool images were acquired during the welding experiments. Arc plasma and molten pool shapes could be clearly observed in the images. The weld pool images obtained through the monitoring system were correlated with the corresponding penetration depth to configure a dataset that could be used to learn the CNN model for weld pool image-based penetration prediction. The weld pool images obtained through the monitoring system were associated with the corresponding penetration depths to construct a dataset that was used to train a CNN model for penetration prediction.

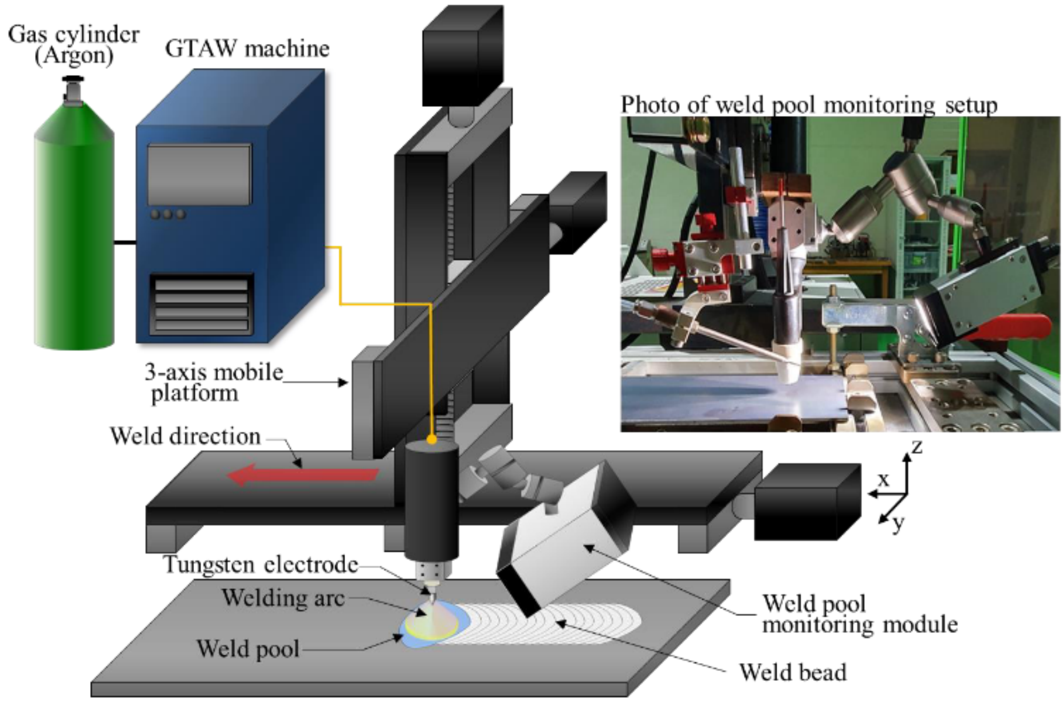

Fig. 2 Weld pool monitoring setup for acquiring weld pool image dataset in gas tungsten arc welding (GTAW)

Table 1 Parameters for weld pool monitoring

Item	Content
Exposure time	10 μs
Aperture	F16
Resolution (H×V)	1280×1024 (1.3 Mpix)
Pixel size (H)	1.315 μm
Pixel size (V)	1.275 μm
Frame rate (fps)	24

3. Estimation model for penetration prediction

Fig. 3 shows the modeling and optimization process for the weld pool image-based penetration prediction developed in this study. The prediction accuracy of the proposed estimation model was affected by the shape of the input image, CNN architecture, and structure of the fully connected layers (FCLs). In this section, the effect of each characteristic on the prediction accuracy was analyzed, and the weld pool image-based penetration prediction model was optimized.

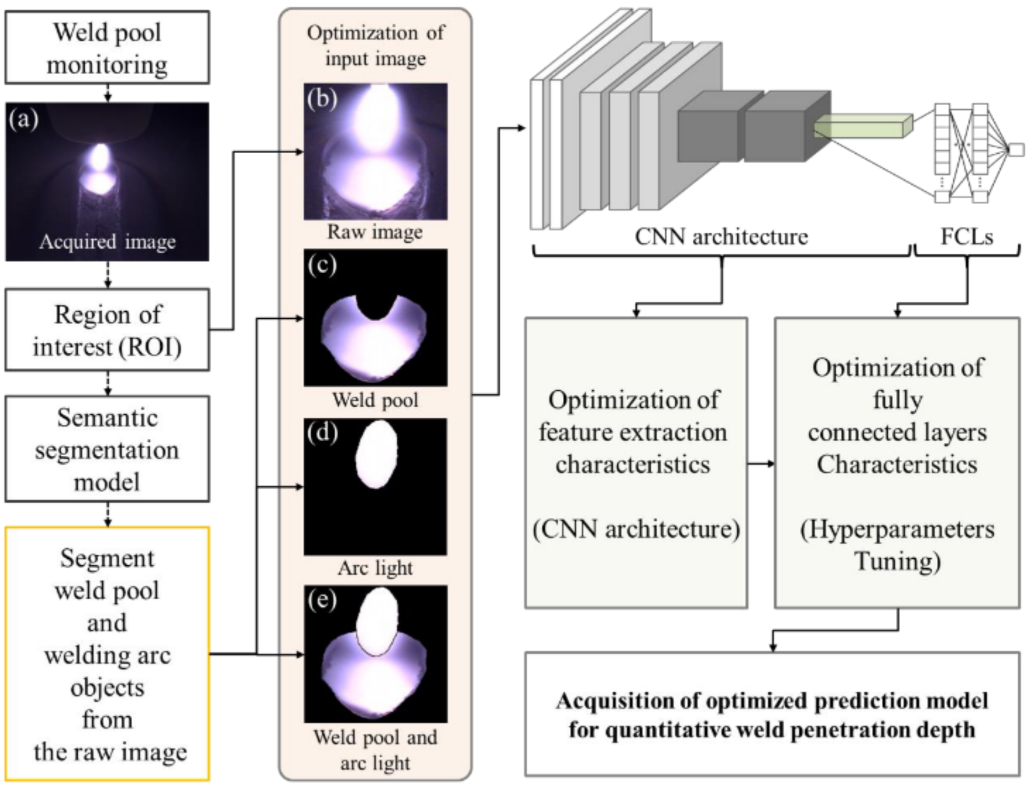

Fig. 3 Overall modeling process for weld pool image-based weld penetration prediction: For the modeling, three optimizations are performed: input image, feature extraction using a deep learning approach, and fully connected structure.

3.1 Dataset configuration of weld pool images

As shown in Fig. 3(a), the raw weld pool image contains not only a weld pool but also various other objects including the welding arc, shielding gas, weld bead, and welding torch. These objects can affect the convolution filter parameters trained during the learning process of the CNN model and affect the prediction of the target. The weld pool shape is closely related to the penetration depth [28]; however, the relationship between the other objects in the weld pool image and the penetration depth remains unclear. Therefore, in this study, various datasets were configured on the basis of the objects, and the most optimized image for penetration prediction was derived by comparing the performances of the CNN models on each dataset (Figs. 3(b)–(d)).

The dataset, according to the object in the weld pool image, contained four types of images: a raw image, a weld pool image, an arc light image, and an image with only the weld pool and arc light. As shown in Fig. 3(b), for the welding torch and welding beads, the objects are filtered by setting a region of interest (ROI) in the raw images. The ROI image contained three main objects: weld pool, arc light, and shielding gas. Multiple datasets were constructed based on these three objects.

Fig. 3(c) shows an image containing only the weld pool. Fig. 3(d) shows an image containing only the arc light. The images captured using the monitoring setup, described in Section 2, were taken tilted at an angle of approximately 45°, and the weld pool was obscured by the arc light object along the length of the arc light. Therefore, the weld pool has a concave upper part. Fig. 3(e) shows an image containing only the arc light and weld pool, without the shielding gas object.

The segmented arc light and weld pool in the images, shown in Figs. 3(c) and (e), were generated as arc light and weld pool objects extracted through the segmentation method proposed in a previous study [29]. Fig. 4 shows an overview of the extraction process of the arc light and weld pool objects using the semantic segmentation model and dataset configuration proposed in a previous study.

The segmentation method was implemented using a CNN-based semantic segmentation architecture with a residual network structure [25], which was trained to segment the arc light and weld pool from the raw image. Conventional object-detection methods that use CNNs detect a target object in the form of a bounding box. In this method, the exact shape of the target object cannot be extracted because the positional information of the target object pixels is lost in the convolutional and flattening layers of the FCLs. The semantic segmentation model classifies labels into pixel units and does not include FCLs. Therefore, the loss of location information is minimized, and the exact shape of the target can be extracted.

The raw weld pool images were inputted to the semantic segmentation model and outputted as three labeled images: arc light, weld pool, and background, as shown in Fig. 4. The binary-labeled map generated through the segmentation model was mapped to the raw image and segmented, as shown in Figs. 3€, (d), a€(e). The segmented images were used for dataset configuration to train the quantitative penetration prediction model.

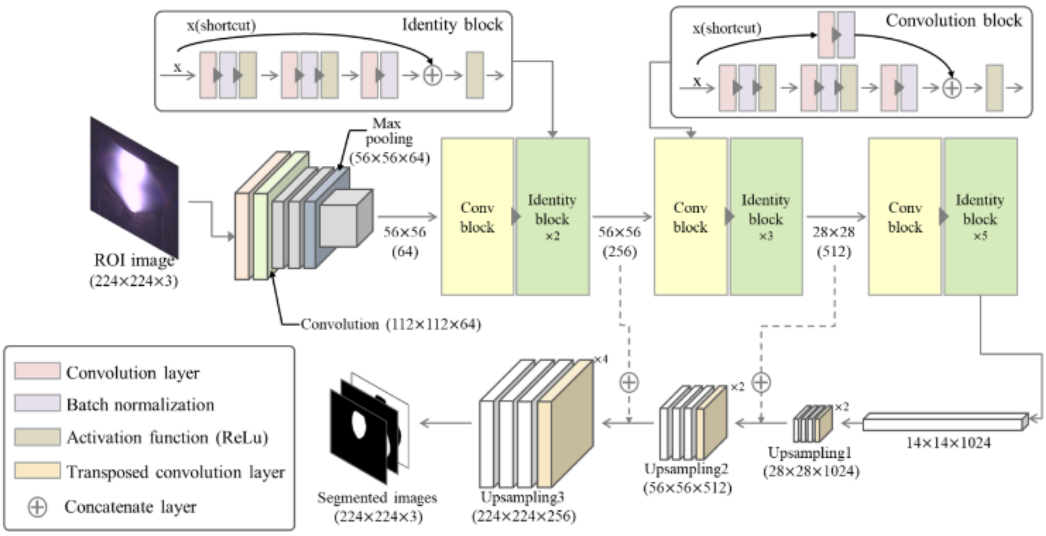

Fig. 4 Semantic segmentation model for extracting weld pool and arc light object from a raw weld pool image. The segmentation model uses a residual network architecture as the feature extractor. Output-labeled images are mapped to a raw image, and a dataset is configured through

the mapped images.

3.2 Construction of a deep learning model for penetration prediction

The configured weld pool image datasets, described in Section 3.1, were inputted to a deep learning model composed of a feature extractor based on the CNN architecture and FCLs. The CNN architecture extracted image features from the input image. The FCLs in the CNN architecture flattened the extracted image features through a flattening layer and recreated them in the form of a 1D array. The 1D array was finally outputted as a quantitative penetration depth.

The extracted image features depend on the characteristics of the CNN architecture, which can affect the prediction output. To verify the universality and practicality of weld pool image datasets configured based on the image objects, it is necessary to compare the performances of various CNN architectures. Therefore, as listed in Table 2, VGG-16 [24], ResNet50 [25], DenseNet [26], and EfficientNet [27], which are widely used CNN architectures with a satisfactory image recognition performance, were selected to compare the performances of the prediction models based on the characteristics of the CNN architecture. The proposed CNN architecture was used to construct image feature extraction layers, and the architecture was evaluated for penetration prediction accuracy.

Table 2 CNN architectures for feature extraction from a weld pool image

Year	CNN	Developed by	Number of training parameters
2014	VGG-16	Simonyan, Zisserman	134.7 million
2015	ResNet-50	Kaiming He	23.9 million
2017	DenseNet-121	Gao Huang et al.	7.2 million
2019	EfficientNet-B4	Mingxing Tan and Quoc V. Le	19 million

The prediction accuracy of the deep-learning model was also significantly affected by the hyperparameters. Unlike weights or biases, the hyperparameters in the deep learning model do not change through model training. The hyperparameters were set before training, and the learning quality was determined according to the set hyperparameters. To optimize the FCLs and target prediction accuracy, the model performance should be verified according to the various hyperparameters.

The prediction accuracy of the deep learning model was significantly affected by the structural characteristics of the FCLs. In particular, the structure of the hidden layers significantly influenced the target accuracy of the object function. Table 3 lists the structural details of the FCLs. FCLs are backpropagation neural networks (BPNNs) having three layers: input, hidden, and output layers. The number of nodes in the input and output layers is fixed and correspond to the number of input and output features. In the FCLs, the depth of the hidden layer and the number of nodes in each hidden layer are the hyperparameters.

The hyperparameter tuning in this study involved the optimization of the depth of the hidden layer and the number of nodes in each hidden layer. Typically, the deeper the hidden layer, the higher the accuracy; however, the number of parameters to be trained will increase significantly. In addition, more hidden layers increase the learning time, processing time, and occurrence of underfitting or overfitting. Therefore, the depth of the hidden layer must be optimized by considering all these factors. The depth range of the hidden layer for the optimization was 2 to 3, and greater depths were not used because of the difficulty in optimizing the model given the large number of the parameters.

The hidden layer nodes increase the nonlinearity of the model due to the increase in the number of parameters, allowing the BPNNs model to achieve a higher accuracy. However, because numerous nodes can cause problems, such as overfitting (e.g., hidden layer depth), the model training needs to be optimized within an appropriate node range.

A total of 288 cases were optimized on the basis of the depth of the hidden layer and the number of nodes; these were too many to be trained and evaluated for performance verification. Therefore, it was necessary to obtain an optimal solution for the estimated function by sampling some of the cases from the total number. We derived an optimal solution for the objective function using a Bayesian optimization algorithm [30]. Bayesian optimization is a global optimization method that determines a global maximum or minimum point for a given function. Compared with grid and random search algorithms, it is easier and more efficient to search for optimized hyperparameters because the Bayesian algorithm derives the optimal solution based on previous trials.

The Bayesian optimization algorithm derives a surrogate model that represents the relationship between the generalization performance and an initial set of hyperparameters. The surrogate model represents the area in which the objective function can be located as a probability distribution, which is a Gaussian process regression. Through the optimization process, new hyperparameters other than the initial points were searched in the distribution until the termination condition was satisfied. An acquisition function was used in this process. The acquisition function recommends a new hyperparameter to be searched through probabilistic calculations. New hyperparameter candidates are typically searched near the searched hyperparameters (exploitation); however, biased optimization results may be obtained depending on the hyperparameter positions of the first few trials. Therefore, biased results are prevented by exploring the hyperparameters in a stochastically uncertain area in the distribution. The Bayesian optimization algorithm determines the optimal hyperparameters by appropriately balancing exploitation and exploration.

Table 3 Structural details of FCLs for hyperparameter tuning

<table>
<tr><td rowspan="2">Two hidden layers</td><td>Input</td><td>Flattening layer</td><td>Hidden layer 1</td><td colspan="2">Hidden layer 2</td><td>Output</td><td>Cases</td><td>Total cases</td></tr>
<tr><td>Output of feature extractor

[1×1×D]</td><td>[1×D]</td><td>1024
512
256
128
64
32
16
8</td><td colspan="2">64
32
16
8</td><td>1</td><td>32</td><td rowspan="3">288</td></tr>
<tr><td rowspan="2">Three hidden layers</td><td>Input</td><td>Flattening layer</td><td>Hidden layer 1</td><td>Hidden layer 2</td><td>Hidden layer 3</td><td>Output</td><td>Cases</td></tr>
<tr><td>Output of feature extractor

[1×1×D]</td><td>[1×D]</td><td>1024
512
256
128
64
32
16
8</td><td>1024
512
256
128
64
32
16
8</td><td>64
32
16
8</td><td>1</td><td>256</td></tr>
</table>

* D: Number of feature maps after the last convolutional filter in CNN

4. Experimental system and scheme

The quantitative penetration depth prediction model using weld pool images was validated and optimized using the optimization process described in Section 3. This section presents the details of the dataset construction for the training and optimization of the deep learning model. The metrics for evaluating the performance of the trained model are also described.

4.1 Weld pool image acquisition and dataset preparation

Table 4 lists the experimental conditions for acquiring the weld pool images. The weld pool experiments were conducted under an automatic welding setup, as shown in Fig. 2. The voltage was set to a fixed value. To obtain various weld pool images, the weld current and speed were set as the variable parameters, and GTAW was performed on the bead-on-plate. A welding machine (Fronius TIG3000) was used for the welding experiments.

Table 4 Experimental conditions for welding

Item	Content
Welding type	GTAW
Workpiece material	SUS304 (100×300)
Workpiece thickness	3 mm
Shielding gas	Argon 100%
Gas flow rate	15 l/min
Electrode tip angle	45°
Weld current	80–180 A
Voltage	12.8 V (fixed)
Welding speed	8–16 cm/min
Travel angle	0°

The input image and penetration depth, which were the target values, were acquired, as shown in Fig. 5. To obtain the input image and penetration depth, n random points were designated in the welded section, excluding the welding start point d_s and endpoint d_e. The image data acquisition (Fig. 5(a)) and designated points (Fig. 5(b)) were matched using the welding speed and distance. The welded specimen was cut at designated points d_n, and the penetration depth p_n at the cross section was measured using an optical microscope, as shown in Fig. 5(c). After welding, a boundary appeared between the welded metal and the base metal. The penetration depth in the case of partial welding was measured as the vertical distance from the base metal surface to the vertical end of the boundary line. During full penetration, a backside bead formed on the rear side of the welded specimen. The full penetration depth was measured as the vertical distance from the base metal surface to the backside bead end. A total of 468 weld pool images and corresponding penetration depths were acquired during the welding experiments. The prepared datasets were separated at a ratio of 75:25 and used as the training and test datasets. Table 5 summarizes the details of the dataset configuration.

As described in Section 3.1, four image datasets were constructed, as shown in Figs. 3(c)–(e), to analyze the effect of objects on the weld pool image. The segmented image datasets were configured using the segmentation method proposed in a previous study [29].

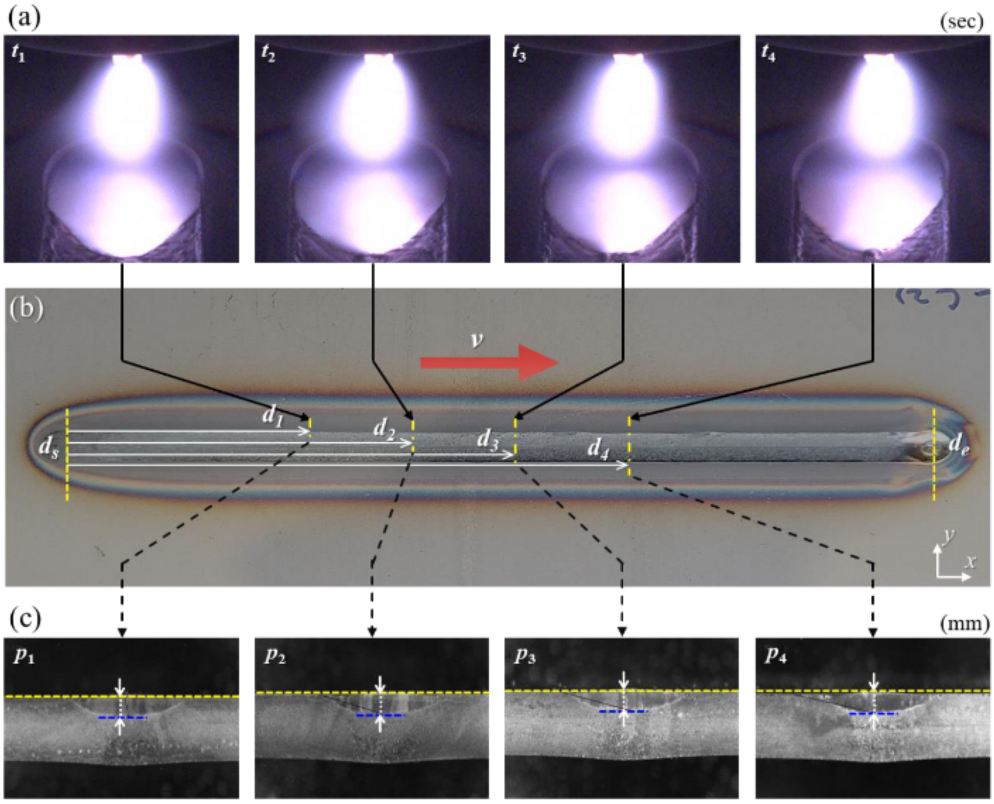

Fig. 5 Data acquisition for training dataset configuration: (a) weld pool monitoring image taken at time t_n, (b) front side of the welded specimen. d_s and d_s are the start and end points of the welding, respectively, and d_n is the distance from d_s to the designated point to measure the penetration depth, (c) weld specimen cross-section at d_n. The penetration depth is measured through an optical microscope, and the penetration depth is the perpendicular distance between the specimen surface and the molten metal boundary ends.

Table 5 Dataset configuration

Item	Content
Input image size	224×224×3
Image format	Color (RGB)
Image dataset type	· Raw image · Weld pool · Arc light · Weld pool + arc light
Number of datasets	468
Number of training datasets	351 (75.0%)
Number of test datasets	117 (25.0%)

4.2 Model Training and Optimization

Three optimization factors were described in Section 3 for optimizing the penetration depth prediction model: the weld pool image dataset, CNN architecture, and hyperparameter tuning. Sixteen major cases were constructed based on four datasets containing weld pool objects and four CNN architectures, and 416 minor cases were constructed based on the hidden layer parameters of the FCLs. The total number of optimization cases was 6,656, which was too high for training and validation. Therefore, a Bayesian optimization algorithm was used for the minor cases, and with this method, the optimized FCL structure for each of the major cases was applied. Table 6 presents the Bayesian optimization parameters.

In this study, except for the hidden layer parameters of the FCLs in the prediction model, the other hyperparameters, such as the batch size, epochs, learning rate, loss function, and optimizer, were fixed, as listed in Table 7. The batch size was set to a mini-batch size of 32, which increased the regularization and minimized the overfitting, although learning tendencies fluctuated more and became more unstable than that in the case of higher batch sizes. Wu et al. [31] studied the performance of deep learning models for various batch sizes. They showed that learning with a batch size of 32 or more generally produced a similar performance; however, the errors tended to increase with a smaller batch size. To minimize the error in the target prediction and conduct appropriate iterations considering the number of training datasets, the deep learning model was trained by selecting a batch size of 32. The epoch was set to be the same for all the cases for quantitative and accurate comparisons. To minimize frequent fluctuations during learning and supplement training stability, the epoch was set to 450.

The loss function applied was the mean squared error (MSE) function, which is most suitable for continuous outputs. The optimizer was the Adam algorithm [32], which was used to minimize fluctuations due to the small batch size; it has a lower dependency on the learning rate than the stochastic gradient descent (SGD) algorithm. The hyperparameters of the Adam algorithm are exponential moving averages β_1 and β_2, which were fixed at 0.9 and 0.999, respectively, for the highest performance [32].

Dataset construction, building, and training of the deep learning models were performed using the Keras library (v.2.4.3) with Python 3.6.5 [33], a library based on Python for modeling various deep learning architectures.

Table 6 Parameters of the Bayesian optimization

Item	Content
Initial points	10
Batch size	32
Maximum trials	25
Epochs per trial	50
Beta	2.6
Objective function	Mean-squared error

Table 7 Training parameters

Parameter	Value
Learning rate	0.001
Epoch	450
Batch size	32
Optimizer	Adam $\beta_1 = 0.9$ $\beta_2 = 0.999$

Loss function	Mean-squared error

4.3 Evaluation of trained model

The trained and optimized models were evaluated using the mean absolute error (MAE) and R^2. In terms of the quantitative and continuous values of the model prediction, the error between the target and predicted values was suitable as a performance metric. R^2, which is the coefficient of determination, describes how well the proposed model fits the target, and the closer the R^2 value is to one, the higher the performance of the model for the target. The two performance metrics can be expressed as in Eqs. (1) and (2), respectively, where n denotes the amount of test data, Y_i is the actual value, $\hat{Y}_i$ is the predicted value, and μ_Y is the average of the actual values. The trained models were quantitatively compared using the MAE and R^2, and the performance was evaluated on the test dataset.

$$MAE = \frac{1}{n} \sum_{i=1}^{n} \left| Y_i - \hat{Y}_i \right|. \quad (1)$$

$$R^2 = \frac{\frac{1}{n}\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2}{\frac{1}{n}\sum_{i=1}^{n}(Y_i - \mu_Y)^2}. \quad (2)$$

To evaluate the performance in more detail, the cumulative frequency of the error between the actual and predicted values was used as the performance metric. The errors varied but were distributed within a certain range. When the error distribution on the test dataset had a high frequency within a narrow error range, the performance of the trained model was considered high. Conversely, when the frequency was high within a wide error range, the performance of the trained model was considered low. In this study, the cumulative frequency $c_{0.05}$ of the errors with a magnitude less than 0.05 mm and a frequency $f_{0.15}$ of errors with a magnitude greater than 0.15 mm were used as the performance indicators.

5. Results and discussion

5.1 Bayesian optimization results

Table 8 presents the Bayesian optimization results. Fig. 6 shows the total weight parameters of each optimized model. Each optimization was performed thrice to compensate for the bias in the optimization results, and the network structure with the highest number of duplicates among the three trials was selected as the optimal structure. If none of the three cases overlapped, optimization was performed twice more to find the most frequent optimization structure.

The bar graphs in Fig. 6 show the total number of weight parameters in the FCLs for each image dataset and CNN architecture. The top of the bar shows the number of neurons in the hidden layer of the FCLs obtained through optimization. Except for the weld pool image dataset, they were optimized to a largely similar structure. Fewer parameters were required to perform training on the weld pool images, because the number of pixels with a valid intensity was reduced due to the absence of shielding gas and arc light objects. The welding arc image dataset had a decrease in the number of valid pixels. However, the number of parameters was relatively high compared with that of the melt-pool image; this was believed to be because arc light objects under different welding conditions do not have visually valid differences, unlike weld pools. This means that a relatively higher number of weighting parameters is required compared with the weld pool dataset to emphasize the subtle visual differences.

The results of the three-layer case were similar to those of the two-layer case. On the welding arc image dataset, the number of neurons in the first hidden layer increased significantly compared with that in the two-layer case. Because a few pixels in the cropped image had valid intensities and the images by welding condition were not visually different, the structure was optimized to compensate for these issues.

Each optimized FCL structure was trained on the basis of the training conditions listed in Table 7, and its performance was evaluated using MAE, R^2, $c_{0.05}$, and $f_{0.15}$.

Table 8 Optimized FCL structure based on the input image and CNN architecture

	CNN architecture	Raw	Weld pool	Arc light	Weld pool and arc light
Two hidden layers	VGG-16	1024-8	32-8	64-8	512-16
	ResNet-50	1024-32	32-8	128-8	512-16
	DenseNet-121	256-8	32-8	256-8	256-16
	EfficienetNet-B4	512-8	64-8	32-8	512-8
Three hidden layers	VGG-16	512-64-16	128-64-8	1024-32-8	1024-256-8
	ResNet-50	512-256-32	64-32-8	1024-128-8	1024-32-8
	DenseNet-121	1024-64-16	1024-64-8	1024-16-8	1024-16-8
	EfficienetNet-B4	64-32-8	64-32-16	1024-32-8	1024-64-8

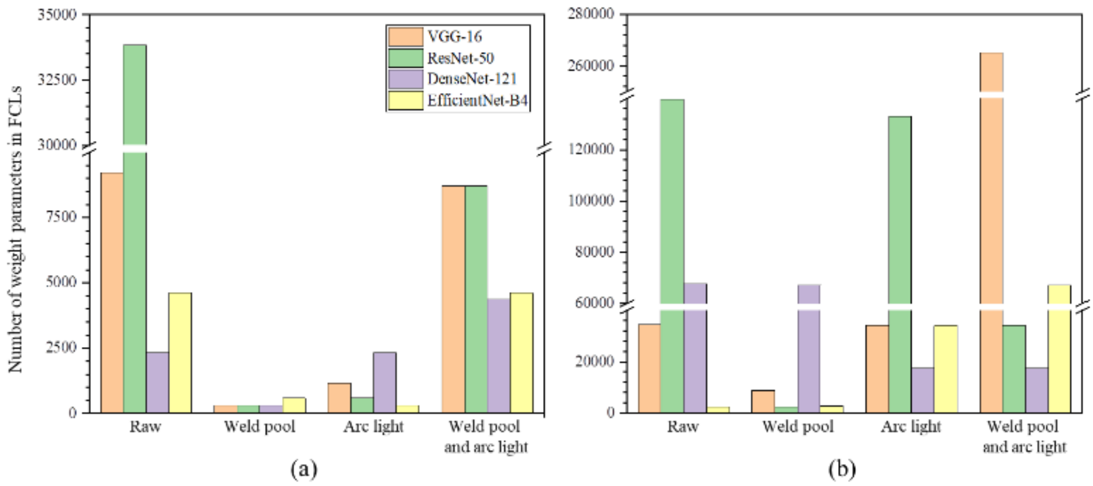

Fig. 6 Number of weight parameters by hyperparameter tuning using Bayesian optimization: (a) models with FCLs including two hidden layers, (b) models with FCLs including three hidden layers

5.2 Training results

Fig. 7 presents the performance metrics MAE and R^2 for the trained model. As shown in Fig. 7, the lower the MAE, the higher the R^2 and performance. Fig. 8 shows the cumulative frequency graph of the error between the penetration depth of the test dataset and the value predicted by the trained model. In Fig. 8, a high cumulative frequency of errors with a magnitude of 0.05 mm or less, $c_{0.05}$, indicates a high performance, while a high frequency of errors with a magnitude above 0.15 mm, $f_{0.15}$, indicates considerable errors in the measurement.

As shown in Figs. 7 and 8, the (a) two-layer and (b) three-layer cases have relatively and extremely low performance when using the welding arc image dataset results, which had extremely high MAE and low R^2 values. $c_{0.05}$ of the two-layer cases on the arc light image dataset was only 0.343, which decreased to 0.229 for the three-layer cases trained with more weighting parameters. The length of the arc light can vary depending on the weld current, particularly the voltage. However, the welding experiments used to obtain the dataset in this study were performed at a fixed voltage, which did not significantly affect the geometric features of the arc light. Even when the current was varied, its effect on the geometric features of the arc light was extremely low.

A visual characteristic of an arc light is that it is a bright, nearly elliptical object with little change in its visual characteristics on varying the welding condition. However, the weld pool shape is highly variable depending on the welding conditions, and the amount of melting per unit of time varies significantly depending on the welding conditions; therefore, the change in its geometric features is prominent. The $c_{0.05}$ values of the DenseNet-121-based model with the highest performance were 0.343 in the two-layer case and 0.229 in the three-layer case; these values were significantly lower than the highest $c_{0.05}$ value obtained on the other image datasets. It was not appropriate to predict the penetration using only arc-light images, and the results of the model trained on the arc-light image dataset were excluded when comparing and evaluating the performance of the trained models.

The VGG-16-based model had a relatively low performance compared with the other architectures, regardless of the image dataset. In Figs. 8(a) and (b), $f_{0.15}$ is high compared with the other architecture-based models. The basic model of the VGG network presented by Simonyan [24] is a large network compared with other architectural models used in this study, and the basic VGG network has three hidden layers with 4,096 neurons each. This implies that the VGG-16-based model requires many parameters for accurate predictions. The other architectures outperformed the VGG-16-based model with fewer FCL parameters. The VGG-16-based model was also excluded from the evaluation of the prediction performance because of its low performance.

The training results of Case 2 obtained on the weld pool image dataset exhibited a relatively higher performance than those obtained on the arc light image dataset. In the graph of Case 2 shown in Fig. 7, both the MAE and R^2 values can be found within the visible range of the graph, unlike the results obtained on the arc light image dataset, and the trend in the cumulative distribution plot, shown in Fig. 8, is similar to the plots obtained on the raw image dataset and weld pool and arc datasets.

The EfficientNet-B4-based model performed relatively well on the weld pool image dataset compared with the other CNN architecture-based models. In particular, in the two-layer case (Case 2 in Fig. 8(a)), $c_{0.05}$ was approximately 2.1 to 2.7 times higher than that in the other CNN architecture cases; this training result demonstrates the characteristics of the efficient net architecture in achieving a high performance even with few training parameters. As shown in Fig. 6(b), the optimization results in the three-layer case indicate that the EfficientNet-B4-based model has relatively fewer parameters than the other image datasets. As in the two-layer case, the model based on EfficientNet-B4 exhibited a relatively higher performance in the three-layer case, evidenced by the relatively lower $f_{0.15}$ value, whereas the $c_{0.05}$ value showed little difference from the results of the other architecture-based models, as shown in Fig. 8(b).

Overall, the model performance on the raw image, weld pool, and arc light image datasets was good. Excluding the VGG-16-based model in the two-layer case, we find that the average MAE of the models on the raw image dataset was 0.0653 mm, which was approximately 36% lower than the average value of the model on the weld pool image dataset and 4.8% lower than that on the weld pool and arc light image datasets. For the three-layer case, the average MAE for the raw image dataset was 0.0576 mm, which was 39% and 29% lower than those for the weld pool image dataset and weld pool and arc light image datasets, respectively, with a uniform distribution for each architecture-based model. The raw image dataset produced a more uniform distribution of the performance metrics than the other two image datasets.

From the results of Case 1 shown in Fig. 8(a), the $c_{0.05}$ value of the EfficientNet-B4-based model was significantly lower, with a difference of approximately 0.13 when compared with the values of the other two architecture models. From the training results of the CNN architecture with FCLs based on two hidden layers on the raw image dataset, it can be concluded that the EfficientNet-B4-based model significantly underperformed the ResNet-50-based and DenseNet-121-based models. The visual difference between the cumulative frequency graphs was relatively negligible in Case 1: the raw image dataset in Fig. 7(b); however, the ResNet-50-based model was found to have a slightly higher performance for cumulative frequency of errors with a magnitude greater than 0.05 mm. Similar to the cumulative frequency graphs, the MAE values of the ResNet-50-based model were 17.3% and 12.3% lower than those of the other two model architectures, respectively.

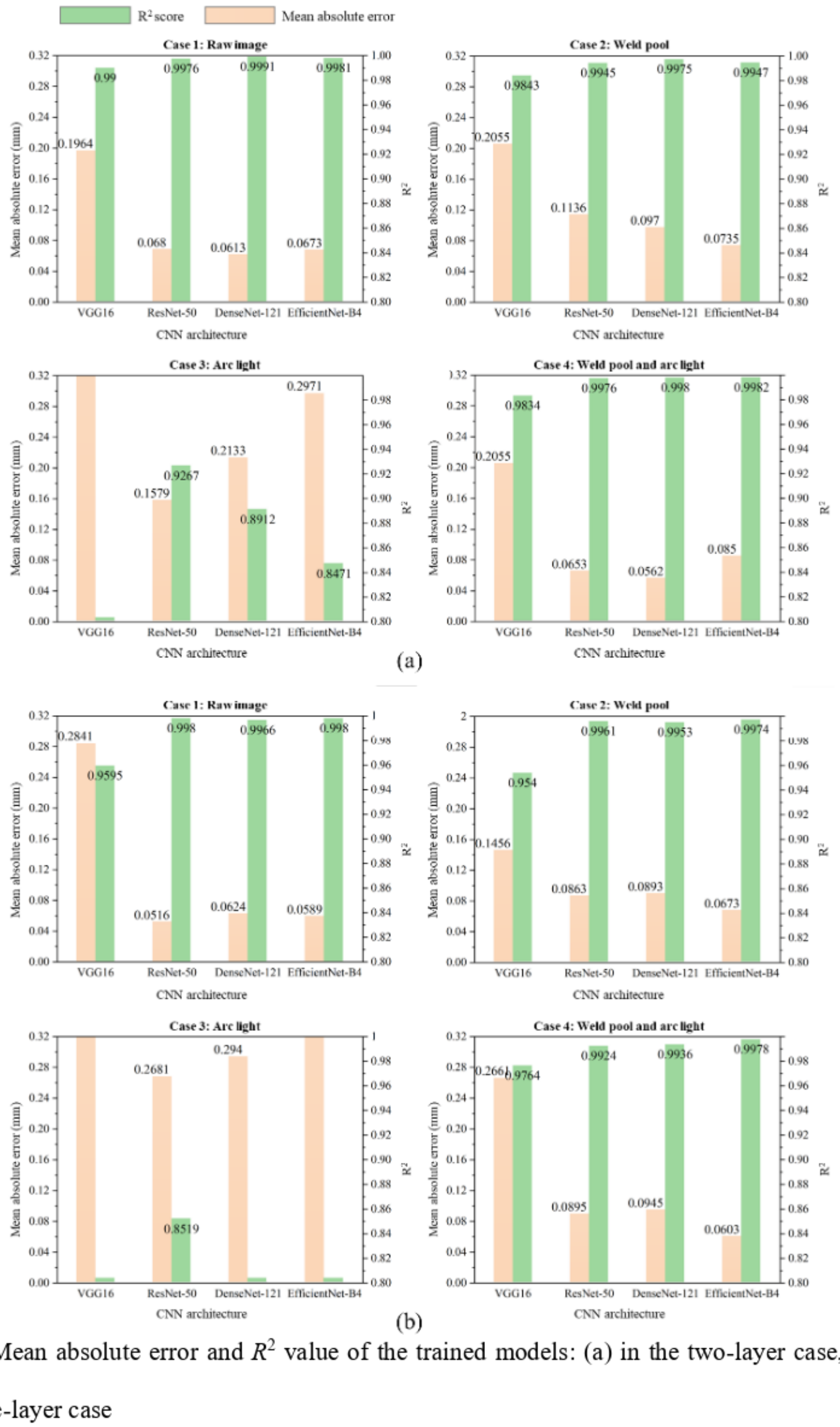

Fig. 7 Mean absolute error and R^2 value of the trained models: (a) in the two-layer case, (b) in the three-layer case

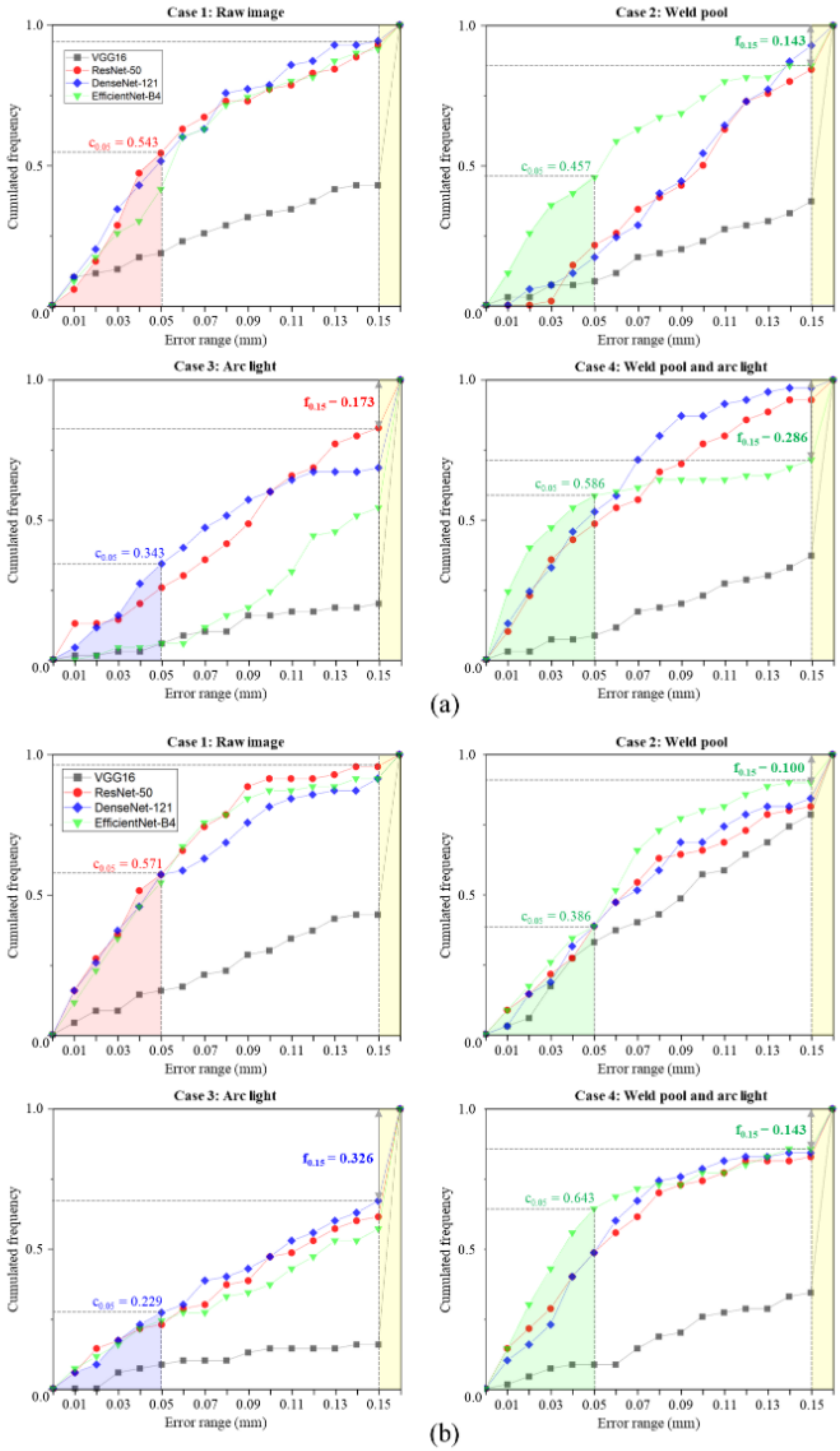

Fig. 8 Cumulative frequency of the error between the actual and predicted penetration depths on the test dataset: (a) in the two-layer case, (b) in the three-layer case

5.3 Performance evaluation

To compare the performances of the prediction models built on various image datasets, models, and FCL structures, several performance metrics and cumulative frequency graphs were used. However, it is difficult to evaluate the performance of the proposed models consistently by comparing each performance metric alone. Therefore, in this study, a method was devised to evaluate the performance by considering all the quantitative performance metrics simultaneously, as shown in Eq. (3). Using Eq. (3), we can derive the performance determinant d for the quantitative penetration prediction. R, c, f, and M in Eq. (3) refer to the four performance metrics used in Figs. 7 and 8: R is the R^2 value, c is $c_{0.05}$, f is $f_{0.15}$, and M is the mean absolute error of the prediction model results. Each variable was normalized between 0 and 1 to make the scale of the variables consistent. Table 8 presents the range of variables for normalization.

In Eq. (3), K_1–K_4 are the priority coefficients. When evaluating the performance of the proposed model, each performance metric was prioritized according to its importance. The priority coefficients are the gain values that determine the weight of each metric variable based on its priority. K_1 and K_2 were set to positive gains, because the higher R and c, the better the model performance. Conversely, K_3 and K_4 were set to a negative gain because the lower M and f, the better the performance of the model.

The R^2 value was the lowest priority of the four performance metrics because the variable was generally high in all the models and had low discrimination for the model performance. However, the other three metrics may be considered important, and the determinant d may vary according to the priority assignment. Therefore, in this study, the performance of the model was evaluated by comparing the deterministic constants for different cases according to the priorities of the performance metrics, as listed in Table 9. The priority gain K_4 for R was fixed at 0.1 as the fourth priority, and six cases were formed depending on the priorities of M, c, and f.

$$d = K_1 M + K_2 c + K_3 f + K_4 R \tag{3}$$

Fig. 9 shows the graphs of the deterministic constants according to the cases set by the priorities in Table 9. In selecting models for calculating d, the training results performed on the arc light image dataset and the VGG-16 architecture were excluded. On the raw image dataset, all the models were selected because the deviation in the MAEs was generally not large, and the visual difference in the cumulative distribution of the errors was not significant. On the weld pool image dataset, the remaining models, except for the EfficientNet-B4-based model, were excluded because of their generally poor performance. The weld pool and arc light image datasets showed some differences in the model selection depending on the number of hidden layers in the FCLs. For the two-layer case, the EfficientNet-B4-based model was excluded because of its high MAE and $f_{0.15}$ when compared to the All models were selected for the raw image dataset case results of other models. In the three-layer case, the EfficientNet-B4-based model was excluded to consider the relatively high MAE of the remaining models. The excluded models had a difference in the MAE of more than 0.02 mm from the highest-performing model with the lowest MAE.

The determinants obtained using Eq. (3) were normalized from 0 to 1 depending on the range of determinants in each priority case, and the determinants were compared within each case because the range of normalization varies by the priority case. In all the cases, the ResNet-50-based model on the raw image dataset had the highest relative rank and included FCLs with three hidden layers, as shown in Fig. 9 1-a. The 1-a model was 1 regardless of the priority case, and it performed best in all cases. In all the cases, the lowest-performing model was the EfficientNet-B4-based model, which included FCLs with two hidden layers for the weld pool and arc-light image datasets, as shown in Fig. 9 3-c. The 3-c model had a normalized determinant of zero in all the cases. The 3-c model had a $c_{0.05}$ of 0.643, which was the highest compared with the performance of all the models. However, owing to the extremely high $f_{0.15}$ and MAE, the determinant was calculated as the lowest value among the model performances.

The model trained on the raw image dataset did not show significant variations between the determinants, even when the performance metrics were prioritized differently. Among the models trained on the weld pool image dataset, only the EfficientNet-B4-based models were selected, as shown in Fig. 2-c. The determinant of Case 2-c was relatively low compared with the determinants of the other models due to the relatively high value of $f_{0.15}$, and errors significantly greater than 0.15 mm also contributed to the low determinant. In particular, the determinant was worse in Cases 2 and 4, where MAE and $f_{0.15}$ were prioritized. On the weld pool and arc light image datasets, the DenseNet-121-based model with two hidden layers exhibited a high performance, which is close to the determinant of the 1-a model with the highest performance.

As a result of the model training and optimization, the raw image, weld pool, and arc length datasets produced a higher performance than the weld pool image dataset using only one object. In other words, the weld pool and welding arc played major roles in predicting the penetration using the welding images. The raw image, weld pool, and arc-light datasets produced some differences in the performance depending on the CNN architecture. In the two-layer case, the ResNet-50- and DenseNet-121-based models outperformed the EfficientNet-B4-based model, whereas in the three-layer case, the other models performed similarly, except for the ResNet-50-based model, which had the highest performance.

Generally, the performance differences were insignificant, except for Models 3-b and 4-a, which had relatively higher performances, and Models 2-c, 3-c, and 5-c, which had significantly lower performances. The number of hidden layers and the resulting increase in the nodes of FCLs did not significantly affect the prediction performance of the model. However, the numbers of hidden layers and nodes affected the performance of the CNN architecture. The performance of the ResNet-50-based model increased when using three hidden layers rather than when using two hidden layers, whereas the performance of the DesNet-121-based model decreased in the three-layer case. As shown in Fig. 7(a), the DenseNet-121 architecture required fewer parameters than the ResNet-50 architecture. In the two-layer case, the DenseNet-121-based model was considered to have a higher training efficiency than the ResNet-50-based model when the learning is done on the weld pool and arc-length image dataset with fewer effective pixels than when it is done on the raw image dataset. Conversely, in the three-layer case, it was estimated that the performance of the ResNet-50-based model was relatively higher owing to the increase in the nonlinearity caused by the large number of parameters. However, because the segmentation process was essential for inputting the weld pool and arc light images, it can be concluded that a ResNet-50-based model, including FCLs with three hidden layers, using raw image as input produced a higher efficiency in terms of the computation time for in-process prediction modeling.

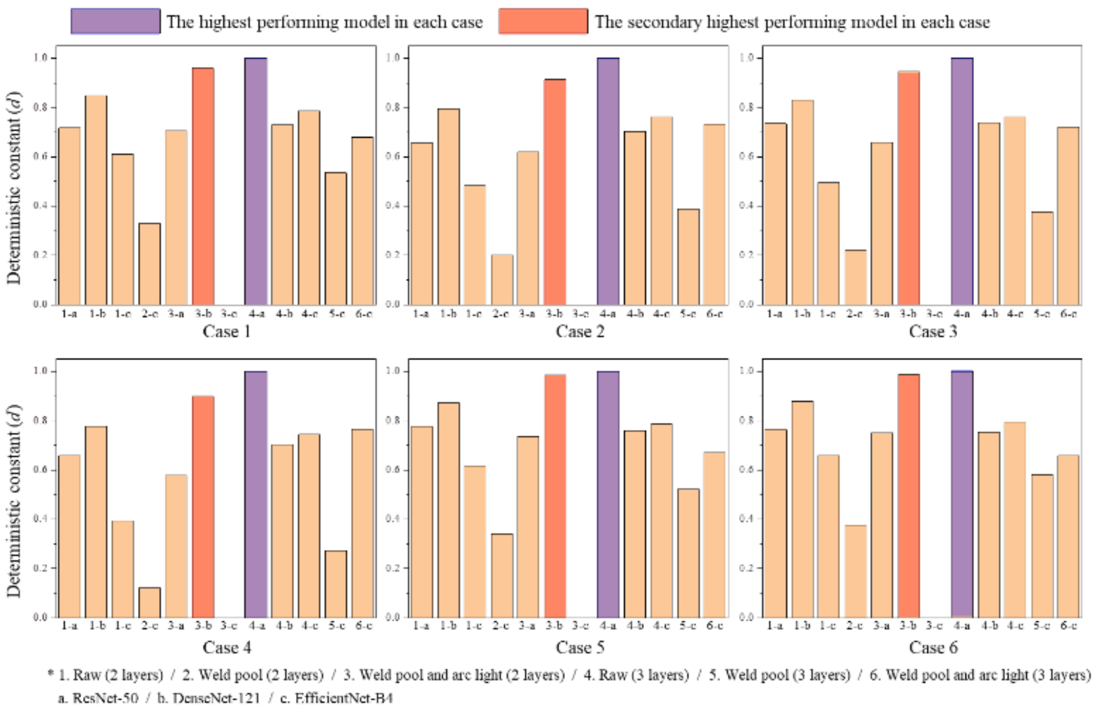

Fig. 9 Cumulative frequency of the error between the actual and predicted penetration depths on the test dataset: (a) Cumulative frequency of the errors in the two-layer case, and (b) Cumulative frequency of the errors in the three-layer case.

Table 9 Cases of determinant calculations according to the priority of performance metrics

	Normalization				Priority factors					
	Min	Max	Variables	Sign	Case 1	Case 2	Case 3	Case 4	Case 5	Case 6
M	0.1	0	K_1	-	0.4	0.4	0.2	0.3	0.2	0.3
c	0.8	0	K_2	+	0.2	0.3	0.4	0.4	0.3	0.2
f	0.3	0	K_3	-	0.3	0.2	0.3	0.2	0.4	0.4
R	1.0	0.99	K_4	+	0.1	0.1	0.1	0.1	0.1	0.1

6. Conclusions

In this study, a deep learning approach and weld pool image-based prediction method that can effectively predict the penetration depth were developed for the in-process quality estimation of automatic welding. Three factors were considered for model optimization: the input image, feature extraction characteristics, and FCL structure. An optimization process for these factors was established and performed, and through this process, the most optimized welding image-based penetration prediction model was established.

The input image of the prediction model, that is, the topside weld pool image, includes various objects in the image; thus, the input image could be optimized by analyzing the effect of each object on the penetration prediction. The topside weld pool images were reconstructed into four image datasets based on the objects. The configured datasets included raw, weld pool, arc light, weld pool, and arc light images. The datasets were constructed using a previously proposed semantic segmentation method to extract the weld pool and arc light objects from the raw image.

Feature extraction characteristics for the penetration prediction using weld pool images were analyzed and compared using CNN architectures such as VGG-16, ResNet-50, DenseNet-121, and EfficientNet-B4architectures, which were mainly used for object detection and classification. The CNN architecture was constructed using FCLs. The FCLs outputted a quantitative and continuous penetration depth from feature vectors, where the input image was transformed through the convolutional layers of the CNN architecture. The quantitative penetration depth output was strongly influenced by the FCL structure.

In this study, a model built by including these three factors was optimized. To optimize the FCLs, the network structure, in terms of the depth of the hidden layer and the number of nodes in each hidden layer, was optimized to a structure suitable for the input image and CNN architecture through hyperparameter tuning using the Bayesian optimization algorithm. The prediction models optimized for the image datasets and CNN architectures were compared and evaluated through a quantitative comparison of the performance metrics for the test dataset.

In terms of the input image, the raw image, weld pool, and arc light image datasets showed relatively low MAEs compared with the weld pool image and arc light image datasets. It is difficult to accurately predict the penetration depth using only one object. In particular, the prediction accuracy of an arc light object, in which the difference in the image characteristics depending on the welding conditions was not large, was extremely low.

Owing to the hyperparameter tuning of the FCL structure, the number of weighting parameters of the optimized structure was relatively low for the weld pool image dataset. As the number of effective pixels in the image was relatively small, it was concluded that the number of required weight parameters was relatively small.

In the optimization of the CNN architecture, the ResNet-50- and DenseNet-121-based models performed well on raw, weld pool, and arc light image datasets, and the EfficientNet-based models showed a relatively high performance on the weld pool image datasets. The parameters of the optimized FCLs for the weld pool image dataset were relatively few compared with those for the other image datasets, indicating the characteristics of the EfficientNet architecture with a relatively high performance when few weight parameters were used.

To evaluate the performance of the optimized model, four performance metrics were used: MAE, R^2, $c_{0.05}$, and $f_{0.15}$. However, it is difficult to evaluate the performance of an optimized model simply by comparing each performance metric individually. Therefore, an equation for a determinant that can evaluate the performance considering all the metrics was devised, and through the determinant, the performance of all the prediction models was compared and verified as a collective and quantitative metric. Six cases were set according to the priority of each metric, and for each case, the determinant of the model was normalized between 0 and 1 for relative comparison.

In all the cases, the ResNet-50-based model with FCLs, including three hidden layers, exhibited the highest performance on the raw image dataset. The ResNet-50-based model with an MAE of 0.0516 mm and $f_{0.15}$ of 0.042 showed the lowest performance among all the optimization cases. $c_{0.05}$ of the model was 0.571, which was relatively lower than that of the EfficientNet-B4 architecture model trained on the weld pool and arc light image datasets. However, the EfficientNet-B4 model had an average determinant of 0.354 for all cases because of the very high $f_{0.15}$ values of 0.143 and 0.286 for the two- and three-layer cases, respectively. This was approximately 43% lower than the average determinant of the ResNet-50-based model. The model with the second-best performance was the DenseNet-121-based model with FCLs, which included two hidden layers.

After the prediction model was optimized, the ResNet-50-based model with three hidden layers trained on the raw image dataset and the DenseNet-121-based model with two hidden layers trained on the weld pool and arc-light image datasets showed the highest performance. The determinants of each model were 1 and 0.95, respectively, and $c_{0.05}$ values on the test dataset were 0.571 and 0.529, respectively, meaning that the ratio of errors with a magnitude less than 0.05 mm exceeded half. In addition, the $f_{0.15}$ values were 0.0429 and 0.0286, respectively, and the ratios of high errors that had the greatest effect on prediction accuracy was very low. However, the DenseNet-121-based model with two hidden layers trained on the weld pool and arc light image dataset also required a segmentation process to extract the weld pool and arc light object, requiring additional computation time to predict the penetration depth. Therefore, to develop an in-process prediction method, an end-to-end model based on the ResNet-50-based model with FCLs, including three hidden layers, can be considered the highest-performing prediction model, which does not require preprocessing steps, such as segmentation.

This study provides an efficient method for in-process weld-quality estimation with a high accuracy by optimizing the input weld pool image, CNN architecture, and FCL structure. Through these proposed methods, weld penetration can be accurately and efficiently predicted using only the weld pool image. In addition, an accurate prediction of the penetration depth can be applied to real-time weld quality control to ensure a sound weld back bead. We also believe that the penetration depth can be automatically controlled based on real-time quantitative feedback.

References

1. Manorathna R. P, Phairatt P, Ogun, P, Widjanarko T, Chamberlain M, Justham L, Marimuthu S, Jackson M. R (2014) Feature extraction and tracking of a weld joint for adaptive robotic welding. 13th International Conference on Control, Automation, Robotics & Vision (ICARCV): 1368-1372. https://doi.org/10.1109/ICARCV.2014.7064515

2. Kovacevic R, Zhang Y. M, Li L (1996) Monitoring of weld joint penetration based on weld pool geometrical appearance. Welding Journal 75(10): 317-329.

3. Bae K, Lee T, Ahn K (2001) An optical sensing system for seam tracking and weld pool control in gas metal arc welding of steel pipe. Journal of Materials Processing Technology 120(1): 458-465. https://doi.org/10.1016/S0924-0136(01)01216-X

4. Wang Z, Zhang C, Pan Z, Wang Z, Liu L, Qi X, Mao S, Pan J (2018) Image segmentation approaches for weld pool monitoring during robotic arc welding. Applied Sciences 8(12): 2445. http://dx.doi.org/10.3390/app8122445.

5. Chen Z, Chen J, Feng Z (2017) Monitoring weld pool surface and penetration using reversed electrode images. Welding Journal 96: 367-375.

6. Kshirsagar R, Jones S, Lawrence J, Tabor J (2019) Prediction of Bead Geometry Using a Two-Stage SVM–ANN Algorithm for Automated Tungsten Inert Gas (TIG) Welds. Journal of Manufacturing and Materials Processing 3(2): 39. https://doi.org/10.3390/jmmp3020039

7. Liu L, Chen H, Chen S (2019) Online Monitoring of Variable Polarity TIG Welding Penetration State Based on Fusion of Welding Characteristic Parameters and SVM. Transactions on Intelligent Welding Manufacturing 2(1): 87-104. https://doi.org/10.1007/978-981-10-8740-0_5

8. Ouafi E, Bélanger R, Méthot J (2011). Artificial Neural Network-based Resistance Spot Welding Quality Assessment System, Metallurgical Research & Technology 108(6): 343-355. https://doi.org/10.1051/metal/2011066

9. Chokkalingham S, Chandrasekhar N, Vasudevan M (2012) Predicting the depth of penetration and weld bead width from the infra red thermal image of the weld pool using artificial neural network modeling. Journal of Intelligent Manufacturing 23: 1995–2001. https://doi.org/10.1007/s10845-011-0526-4

10. Mahadevan R, Jagan A, Pavithran L, Shrivastava A, Selvaraj S (2021) Intelligent Welding by Using Machine Learning Techniques, Materials Today: Proceedings 46(2): 7402-7410. https://doi.org/10.1016/j.matpr.2020.12.1149

11. Korat P, Sama M (2019) Implementation of Artificial Intelligence in TIG. Proceedings of International Conference on Advancements in Computing & Management (ICACM) 2019: 1055-1062. http://dx.doi.org/10.2139/ssrn.3462445

12. Hong Y, Chang B, Peng G, Yuan Z, Hou X, Xue B, Du D (2018) In-process monitoring of lack of fusion in ultra-thin sheets edge welding using machine vision. Sensors 18(8): 2411. https://doi.org/10.3390/s18082411

13. Fang J, Wang K (2019) Weld pool image segmentation of hump formation based on Fuzzy C-Means and Chan-Vese model. Journal of Materials Engineering and Performance 28: 4467-4476. https://doi.org/10.1007/s11665-019-04168-y

14. Jiang C, Zhang F, Wang Z (2017) Image processing of aluminum alloy weld pool for robotic VPPAW based on visual sensing. IEEE Access 5: 21567-21573. https://doi.org/10.1109/ACCESS.2017.2761986

15. Wen H, Zeng J, Bian Z, Hu A, Chu F and Mao Y (2022) Study of weld pool monitoring system based on spatial filtering. Journal of Manufacturing Processes 76: 638-645. https://doi.org/10.1016/j.jmapro.2022.02.044

16. Cheng Y, Wang Q, Jiao W, Yu R, Chen S, Zhang Y, Xiao J (2020) Detecting dynamic development of weld pool using machine learning from innovative composite images for adaptive welding. Journal of Manufacturing Processes 5: 908-915. https://doi.org/10.1016/j.jmapro.2020.04.059

17. Cheng Y, Chen S, Xiao J, Zhang Y (2021) Dynamic estimation of joint penetration by deep learning from weld pool image, Science and Technology of Welding and Joining 26(4): 279-285. https://doi.org/10.1080/13621718.2021.1896141

18. Li C, Wang Q, Jiao W, Johnson M, Zhang Y (2020) Deep learning-based detection of penetration from weld pool reflection images. Welding Journal 99: 239-245. https://doi.org/10.29391/2020.99.022.

19. Nomura K, Fukushima K, Matsumura T, Asai S (2021) Burn-through prediction and weld depth estimation by deep learning model monitoring the molten pool in gas metal arc welding with gap fluctuation, Journal of Manufacturing Processes 61: 590-600. https://doi.org/10.1016/j.jmapro.2020.10.019

20. Shelhamer E, Long J, Darrell T (2017) Fully convolutional networks for semantic

segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence 39: 640-651. https://doi.org/10.1109/TPAMI.2016.2572683

21. Ronneberger O, Fischer P, Brox T (2015) U-Net: Convolutional networks for biomedical image segmentation. Medical Image Computing and Computer-Assisted Intervention (MICCAI 2015): 234–241. https://doi.org/10.1007/978-3-319-24574-4_28

22. Cai W, Jiang P, Shu L, Geng S, Zhou Q (2022) Real-time identification of molten pool and keyhole using a deep learning-based semantic segmentation approach in penetration status monitoring, Journal of Manufacturing Processes 76: 695-707. https://doi.org/10.1016/j.jmapro.2022.02.058

23. Yu R, Kershaw J, Wang P, Zhang Y (2021) Real-time recognition of arc weld pool using image segmentation network, Journal of Manufacturing Processes 72: 159-167. https://doi.org/10.1016/j.jmapro.2022.02.058

24. Simonyan K, Zisserman A (2014) Very Deep Convolutional Networks for Large-Scale Image Recognition. International Conference on Learning Representations (ICLR 2015). https://doi.org/10.48550/arXiv.1409.1556

25. He K, Zhang X, Ren S, Sun J (2016) Deep Residual Learning for Image Recognition. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR): Proceedings 770-778. https://doi.org/10.1109/cvpr.2016.90

26. Huang G, Liu Z, Maaten L, Weinberger K (2017) Densely Connected Convolutional Networks, 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR): Proceedings: 2261-2269. https://doi.org/10.48550/arXiv.1608.06993

27. Tan M and Le Q (2019) EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks. 36th International Conference on Machine Learning (ICML 2019): Proceedings: 6105-6114. https://doi.org/10.48550/arXiv.1905.11946

28. Zhang Y, Cao Z, Kovacevic R (1996) Numerical analysis of fully penetrated weld pools in gas tungsten arc welding. Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science 210(2): 187-195. https://doi.org/10.1243/PIME_PROC_1996_210_185_02

29. Baek D, Moon H, Park S (2022) In-process prediction of weld penetration depth using machine learning-based molten pool extraction technique in tungsten arc welding. Journal of Intelligent Manufacturing. https://doi.org/10.1007/s10845-022-02013-z

30. Brochu E, Cora V, Freitas N (2009) A tutorial on Bayesian optimization of expensive cost functions, with application to active user modeling and hierarchical reinforcement learning. arXiv: 1012.2599. https://doi.org/10.48550/arXiv.1012.2599

31. Wu Y, Johnson J (2021) Rethinking "Batch" in BatchNorm. ArXiv: 2105.07576. https://doi.org/10.48550/arXiv.2105.07576

32. Kingma D, Ba J (2015) Adam: A Method for Stochastic Optimization. 3rd International Conference for Learning Representations (ICLR 2015). https://doi.org/10.48550/arXiv.1412.6980

33. Chollet F (2015) Keras. https://github.com/fchollet/keras