ISSN 2466-2232 (Print)
ISSN 2466-2100 (Online)

Review on the Recent Welding Research with Application of CNN-Based Deep Learning
Part I: Models and Applications

합성곱 신경망기반 딥러닝의 용접연구 적용
Part I: 모델과 활용사례

Kidong Lee*,**, Sung Yi*, Soongkeun Hyun** and Cheolhee Kim*,***,†

*Department of Mechanical and Materials Engineering, Portland State University, OR 97229, USA
**Department of Materials Science and Engineering, Inha University, Incheon, 22212, Korea
***Joining R&D Group, KITECH, Incheon, 21999, Korea

†Corresponding author: chkim@kitech.re.kr, cheol@pdx.edu
(Received September 16, 2020; Accepted November 20, 2020)

Abstract

During machine learning algorithms, deep learning refers to a neural network containing multiple hidden layers. Welding research based upon deep learning has been increasing due to advances in algorithms and computer hardwares. Among the deep learning algorithms, the convolutional neural network (CNN) has recently received the spotlight for performing classification or regression based on image input. CNNs enables end-to-end learning without feature extraction and in-situ estimation of the process outputs. In this paper, 18 recent papers were reviewed to investigate how to apply CNN models to welding. The papers was classified into 5 groups: four for supervised learning models and one for unsupervised learning models. The classification of supervised learning groups was based on the application of transfer learning and data augmentation. For each paper, the structure and performance of its CNN model were described, and also its application in welding was explained.

Key Words: Deep learning, Convolution Neural Network, Image, Welding, Model, Application

1. Introduction

Deep learning is one of the machine learning techniques, a kind of Artificial Intelligence (AI), with increasing applications to a wide range of sectors. Machine learning is not based on explicit rules, it learns rules using datasets, and is categorized into supervised learning, unsupervised learning, and reinforcement learning according to the types of training datasets. Supervised learning uses inputs and correct outputs as training datasets and it is applied to classification and regression problems[1].

Artificial Neural Network (ANN), a type of machine learning, was developed by mimicking human neural networks and consists of an input layer, hidden layers, and an output layer. ANN is further categorized into a shallow neural network (SNN) where there is one hidden layer and a deep neural network (DNN) where there are two or more hidden layers. When machine learning is performed using DNN, it is called deep learning.

As for research that applies ANN to welding, since the error back-propagation algorithm[2] was introduced in the mid-1980s, ANN has been actively applied to a number of welding processes. The examples of application includes prediction of weld quality from control parameters[3,4], suggestion of appropriate control welding parameters from the welding of the desired weld attributes[5], profile extraction from laser vision sensing image[6], and quality and seam tracking using vision sensors[7]. However, there were some problems in expanded application of neural network models at the time. First, the sigmoid function or RBF (Radial Basis Function) function was mainly used as an activation function of neural networks,

and learning of weights far from the output layers was difficult in training of DNN through the back-propagation algorithm. In addition, when images or time series signals were used, learning was performed through feature extraction rather than raw signals without feature extraction, and models that predict outputs from scalar variables such as control parameters were mainly developed, and thus application of the trained model to another production system was intrinsically difficult. When only the control parameters are used as the input parameters of the model, several environmental disturbances (e.g. humidity, temperature, electric power instability) cannot be considered, and it was difficult to set the same system inputs even by setting the same control parameters in other systems.

In recent years, ReLu (Rectified Linear unit) has been proposed as an activation function to enable training of DNN. With the introduction of Convolutional Neural Networks (CNNs)[8], also known as convnets, which are mainly used for image recognition, and Recurrent Neural Network (RNN)[9,10] for use in models using time-series variables, innovative development in deep learning technologies has been achieved including end-to-end learning.

Machine learning has been applied to various welding processes and the examples include the use of deep learning and reinforcement learning in the laser welding control[11], quality prediction through deep learning in arc welding[12], and quality classification through SVM (Supported Vector Machine), a machine learning algorithm [13,14].

In this paper, 18 recently published papers on CNN-based welding research are reviewed. As shown in Fig. 1, these papers were divided into 5 groups based on supervised/unsupervised learning, use of transfer learning, and application of data augmentation in data preprocessing.

2. CNN Structure and Learning Methods

CNN is an artificial neural network that uses convolution operations, and CNN models has shown excellent performance in image recognition because the feature patterns of the images can be extracted by training through convolutional filters. In addition, manual extraction of features is not necessary since features are automatically learned, and models can be constructed based on pre-trained networks. With these advantages, research of applying CNNs for deep learning using images has shown a rapid increase in recent years.

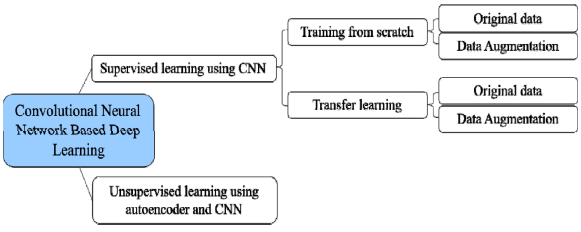

Fig. 1 Classification of papers reviewed

2.1 Basic structure

CNN is composed of convolutional layers, pooling layers, and fully connected layers. In the convolutional layer, inputs within a kernel are connected by specific weights, and a feature map is constructed by information representing each kernel area and passed to the next layer. The convolution operation is performed by sliding the kernel window over the input data at a fixed stride. The output of the convolutional layer can be passed to the next layer as it is, but in most cases, pooling is performed, one of the down sampling operation, to prevent over-fitting and create features that are highly noise-resistant. Typical pooling methods are max pooling and mean pooling, in which the maximum and mean values are output within the specified kernel window, respectively. The fully connected layer is a layer in which nodes of the previous layer and those of the next layer are all connected in the same way as a traditional ANN. In the CNN model, the successive repetition of the convolutional layer and the pooling layer enables inclusion of local features in the image and also represents the global features and these are passed as the input of the fully connected layer, and the fully connected layer produces the final output.

2.2 Key techniques used in the reviewed papers

The techniques adopted for modeling and training are briefly introduced as follows, and more details can be found in References[15].

2.2.1 Activation function

The activation function converts the input signal of a node in an ANN to an output signal, and the identity function, sigmoid, Tanh, ReLu, and softmax function are used as the activation function. In general, ANN performs learning of complex nonlinear phenomena by adopting a nonlinear function as an activation function. Among them, the sigmoid function is useful for converting all values into probabilities, and generates output values between 0 or 1, which can be used for binary classification. The Tanh function has a smooth shape like the sigmoid function, and has the advantage of being able to generate negative values because the output values range between -1 and 1 depending on the input value. The sigmoid function and the tanh function are disadvantageous for use in the learning of DNN be-

cause it is possible that the gradient of the function vanishes by saturation. However, the ReLU function outputs 0 when the input value is negative, and outputs the input value as it is when the input is 0 or positive. The softmax function used in the classification problem converts the input values between 0 and 1 and the sum of the output values is always 1. The softmax function increases the differences in the portions of the input values, making the portion of the largest value even larger and the portions of other values even smaller, and is used as the activation function of the last node for the classification neural network.

2.2.2 Transfer learning

Transfer learning performs learning by using models that has been pre-trained with large-scale data such as VGG, GoogLeNet, and MobileNet, rather than learning a complex CNN structure from scratch in the design of an image classification model using CNN and it allows a accurate model in a relatively time-efficient manner. Transfer learning adopts model optimization through fine-tuning. The fine-tuning strategies can be divided into three methods: retraining the entire model, fine-turning only the parameters of the fully connected (FC) layers, and fine-tuning some convolutional layers and all fully connected layers.

2.2.3 Data augmentation

In deep learning, the more number of training datasets are available, improved accuracy can be achieved. Insufficient number of training datasets leads to over-fitting and ensuring a large number of training datasets is important in the performance. When CNN uses images as input data and the number of training datasets is insufficient, data augmentation can be used to address the problem of insufficient input. Data augmentation is a method of adding artificially processed images to the given training datasets through effects such as flipping, rotation, and changing brightness to the input images. In this paper, for convenience of presentation, the method of using original data without data augmentation is indicated as learning using the original data.

3. Supervised learning with training from scratch

3.1 Original data utilization

Z. Guo et al.[16] proposed a model for classifying normal and defective welds by applying CNN to electric resistance welding in the line pipe manufacturing process, and achieved an accuracy of 99.01%. The cross-section images of the welds were used as input and the sigmoid function was used in the output layer to classify normal and defective welds. The number of hidden layers of the CNN was 5, 7, and 13, and the difference in accuracy according to the network depth was examined. With increasing network depth, the average time per epoch was 261.68 s, 266.24 s, and 267.26 s in case of 5, 7, and 13 layers, respectively, indicating no significant difference but the error rates were 7.89%, 4.93%, and

Table 1 Summary of research using original data and training from scratch

Process	Input	Output	Network	No. of dataset	Ref. No
Electric Resistance Welding	256×256 (px) image from cross-sections	Classification - Weld quality (Good/Defect)	CNN-FCN 13 HL***	11026 (TR*) 304 (TE**)	16
Arc Welding	320×240 (px) image from weld surfaces	Classification - Weld quality (Good, Porosity, Undercut, Spatter)	CNN-FCN 5 HL	120	17
Laser Welding	351 features of in-situ signals from photo diodes, image sensor, and spectrometer	Classification - Weld quality (Good, Blowout, Humping, Undercut)	CNN-FCN 7 HL	7500 (TR) 5000 (TE) from 25 welding runs	18
Gas Tungsten Arc Welding	400×487 (px) in-situ weld pool image	Classification - Weld quality (2, 4, 6 classes)	12 models (6 CNNs & 6 FCNs)	26666 (TR) 6588 (TE)	20
Resistance Spot Welding	128×128 (px) image from weld surfaces	Regression - Strength - Nugget diameter Classification - Fracture mode, Expulsion	CNN-FCN 11 HL	90	21
Laser Welding	121×121 (px) in-situ image of weld pool including keyhole	Classification - Penetration (2, 3, 4 classes)	10 Models (6 CNNs, {HOG, LBP, BOF, or SAE} + SVM)	10,000 (TR) 1000 (TE)	22

* TR: For training; ** TE: For test; *** HL: Hidden Layer

0.99%, showing improved accuracy. It was reported that superior accuracy results were obtained compared to the error rate of 2.3% at human inspection.

A. Kjumaidi et al.[17] proposed a CNN-based defect detection model using photo images of arc welding bead surface as input, and using only 3 hidden CNN layers and 2 hidden FC layer, test accuracy of 95.83% was achieved for 4-class classification of defects (good, porosity, undercut, and spatter).

Y. Zhang et al.[18] performed in-situ measurement of welding signals using multi-sensor technology, and a CNN-based defect detection model was proposed based on the input signals. The multi-sensor system comprised a photodiode, a spectrometer, and two high-speed cameras. The collected photodiode signals were decomposed for each frequency by the Wavelet Packet Decomposition (WPD) method to obtain 320 feature signals. In addition, the spectrometer signals were collected by dividing the wavelength range between 400 nm and 900 nm into 25 bands. The information on the size and location of the keyholes was collected with No. 1 camera with 976 nm laser light and the bandpass filter applied, and the spatter and plume related information was collected from No. 2 camera with a wavelength range of 320 nm-750 nm. After collecting a total of 351 signals from various sensors at a frequency of 500 Hz, the rest were filled with value 1 to form 400 input nodes. For verification on the accuracy of the CNN model, the results were compared with the fully connected network (FCN) model[19] result with same input nodes. In the comparison, the mean error of the FCN model was 9.65% and the mean error of the CNN model was 4.7%, indicating that the CNN model accuracy was superior to that of FCN model. For generalization of the proposed CNN model, the verification of the model was performed by experimenting under four welding conditions. As a result of CNN model classification, error rates of 1.73%, 8.0%, 11.2%, and 8.75% were obtained in each defect class (good, blowout, humping, undercut). When the cases of classification error were analyzed, there were cases where the blowout defect and the undercut defect were confused in detection, or the humping defect and good welds were confused.

D. Bacioiu et al.[20] constructed various DNN models for welding defects classification in GTA welding of aluminum 5083 alloy and compared the model performances. In a total of 60 welding experiments, the in-situ molten pool images were used as input data, and the 6 CNN models and 6 FCN models were developed with defect classifications of 2-class (good, defective), 4-class (good, melt-through, contamination, lack of fusion) and 6-class (good, melt-through, contamination, lack of fusion, misalignment, lack of penetration). Then, the performance was compared according to input image resolution, model depth, and learning rate. In the evaluation of the input image resolution, the comparison of original data images of 800x974 pixels with sampled images at 400x487 and 25x30 pixels confirmed that in the 6-class classification the CNN model showed higher sensitivity to the final classification accuracy than the FCN model. In all models, the accuracy improved with increasing model depth, and learning did not converge at a learning rate of 10^{-3} or below. In all evaluations, CNN models showed higher accuracy than FCN models. S. Choi et al.[21] presented a CNN model that predicts weld quality (tensile shear strength, nugget diameter, fracture mode of welds, expulsion occurrence) through the surface images of resistance spot welds for automobile steel sheets with a tensile strength of 980 MPa. The surface indentation and heat trace correlate with the weld quality, and thus the tensile shear strength, nugget diameter, fracture mode of welds, and expulsion occurrence were predicted with the accuracy of 98.6%, 98.8%, 100%, and 100%, respectively, through the CNN model. The generality of the model was verified through a test using data not included in the training, and the verification result confirmed that the prediction accuracy of the tensile shear strength and nugget diameter was 97.8% and 97.4%, respectively

Z. Zhang et al.[22] used a CNN model to determine the penetration state of tailor-rolled blank (TRB) in-situ laser welding. The applied penetration state classification was 2-class (incomplete penetration, complete penetration), 3-class (incomplete penetration, desirable-complete penetration, overpenetration), and 4-class (incomplete penetration, desirable penetration, complete penetration, overpenetration). Six CNN models with different number of kernels and convolutional layers were trained and evaluated in comparison. It was confirmed that the increased complexity of the model did not necessarily lead to improved accuracy, and the optimal CNN model was determined considering the model complexity in addition to accuracy. Moreover, apart from the model complexity, the image sampling frequency (100 fps, 300 fps, 3000 fps) of the input images that can affect the latency of model was varied to conduct experiments in comparison. Since there was no significant difference in accuracy among models with difference sampling frequencies, 100 fps was selected with the lowest latency. To verify the applicability of the CNN model, it was compared with other methods [(HOG, LBP, Bag of features (BOF), or SAE)+SVM]. As a result of the verification, the CNN model in 2-class classification showed the highest prediction accuracy, and in the case

of 3-class classification, the prediction accuracy of the CNN model was 99%, 93.8%, and 96.5%, respectively, indicating higher performance than other models.

When the penetration state was classified into 4 classes, the accuracy of each model was relatively low, but the stability in the accuracy of the CNN model was high.

3.2 Data augmentation

W. Hou et al.[23] investigated a defect classification model (good, incomplete penetration, pore, slag inclusion, crack) from X-ray images of various welds in an open database GDXray. First, 3503 of 32×32-pixel images were extracted from 88 X-ray images by subsampling. Since the distribution of each type of defect is unbalanced, resampling was performed using ROS (Random Over Sampling), RUS (Random UnderSampling), and SMOTE (Synthetic Minority Oversampling TEchnique) methods. The RUS method is the simplest approach that randomly deletes samples from the majority class to balance the class distribution and the ROS method duplicates samples selected from the minority class and adds to the original dataset to balance the class distribution. In addition, the SMOTE method is a resampling method in which Euclidean distance is calculated to randomly find k-nearest neighbors to create new artificial samples. For classification models, traditional methods of Haralick texture features and HOG (Histograms of Oriented Gradients) were used as well as the deep learning method SSAE (Stacked Sparse Auto-Encoders), and two CNN models with different depths were used. For visual comparison on the results of extracted features through each model, each part was visualized into a 2D map using the t-SNE method to compare the accuracy. From the comparison, it was confirmed that the t-SNE map of the original images was disorderly, and the t-SNE map of the models with Haralick technique and HOG method also showed poor data classification performance. Both SSAE and CNN models showed better performance of classification in the t-SNE map. Between the two CNN models, the one with larger depth showed a better classification performance. Among them, the classification of samples without defects and samples with pores showed the best performance, and crack defects were the most difficult for classification. In the case of the ROS datasets, the mean accuracy of the defect classification for Haralick, HOG, SSAE, CNN1, and CNN2 was 60.9%, 81.6%, 91.4%, 94.7%, and 96.3%, respectively, in the case of the RUS datasets, the accuracy values were 60.7%, 67.5%, and 87.3. %, 76.6%, 79.9%, and for SMOTE datasets, the values were 63.1%, 82.2%, 90.5%, 94.4%, and 97.2%, respectively. In conclusion, they reported that the deep learning method showed better feature extraction performance than the traditional model, and that with SMOTE method, the problem of accuracy reduction from data imbalance can be resolved.

J. Park et al.[24] proposed two-step CNN models for defect detection in engine transmission welds. In the first CNN model, the representation of collected images of the circular welding area of the engine transmission was converted from Polar coordinates to Cartesian coordinates and the center point was predicted. Assuming that the weld width was fixed, the background except the welds was removed, thereby optimizing the input data to the second defect detection CNN model. Compared with Hough circle, a traditional method, the two-step models have showed superior performance

Table 2 Summary of research using data augmentation and training from scratch

Process	Input	Output	Network	Data augmentation	Ref. No
X-ray Inspection	32×32 (px) X-ray images from GDXray-Weld	Classification - Weld quality (2, 5 classes)	5 Models (Haralick, HOG, SSAE, 2 CNNs)	Subsampling ROS**, RUS*** SMOTE*4	23
Transmission Welds	80×256 (px) for center detection 32x352 (px) for defect inspection from weld surfaces	Regression - Center point Classification - Weld quality (Good, NG)	6 Models (HOG+SVM, LBP+SVM, 4 CNNs)	Rotation Translation Scaling Brightness Adj. Contrast Adj.	24
Flux Cored Arc Welding	150×125 (px) in-situ image of weld pool	Regression - 8 molten pool features	CNN-FCN 10 HL*	Horizontal flip Translation Scaling	25
Gas Tungsten Arc Weldng	100×100 (px) in-situ images of weld pool in three-way	Classification - Weld quality (Under-penetration, Full penetration, Burn through)	CNN-FCN 10 HL	Noise addition Rotation	26

* HL: Hidden Layer; ** ROS: Random Over-Sampling; *** RUS: Random Under-Sampling;
*4 SMOTE: Synthetic Minority Oversampling TEchnique

under dark lighting environments and with blurred images. Since there are not many defect data images in the training datasets, image transformation (rotation, translation, scaling) and image distortion (brightness and contrast adjustment) were performed to improve performance. In the second defect detection CNN model, the mini-batch gradient descent method can lead to biased learning due to data imbalance, class-specific batch sampling method with a set ratio of OK and NG images of each batch was used. For the defect detection model, the performance was comparatively evaluated between HOG+SVM, LBP (Local Binary Pattern)+SVM and four CNN models, and all CNN based models showed outstanding performances.

For automation of Flux Cored Arc (FCA) Welding, T. Ashida et al.[25] used a CNN model to extract features of molten pool and calculated the distance between the leading of the molten pool and arc center and the width of the pool to investigate a feedback control system to form back bead. Data augmentation was performed using horizontal flip, translation, and scaling to increase the number of datasets from 2,400 to 12,000. When comparing between the proposed CNN model and the conventional feature-extraction image processing algorithms, the mean prediction error of the leading end position of the molten pool was reduced by 0.3 mm. The seam tracking and back bead formation was verified by implementing the developed feedback control algorithm to the butt weld joint with a joint gap from 3 mm to 10 mm along the weld seam which is also deviated by 10 mm from the pre-taught path.

Z. Zhang et al.[26] developed a CNN model for in-situ weld defects prediction in pulsed gas tungsten arc (GTA) welding of aluminum alloys. In order to increase the training speed and accuracy of the CNN model, a system was devised to collect 3-way information simultaneously from the top front, top back, and back seam. Welding was performed on rectangular, dumbbell-shaped, and ladder-shaped specimens to generate welding defects. Downsampling was performed on the original images of 1392x1040 pixels with the arithmetic mean and thus input images of 100x100 pixels were used. For data augmentation, the number of datasets was increased by adding noise (salt and pepper noise) and randomly applying 15 and 30 degrees rotation to the input images. After the data augmentation, the accuracy improved by 3.88% from 92.3% to 96.18%, verifying the effectiveness of data augmentation. In addition, for optimization of CNN model, the accuracy and the final training speed were compared by varying 4 convolutional layer structures and the number of kernels. The training time and efficiency increased with the number of kernels, and the accuracy increased with the increase in the number of feature images excluding 2D images with only 0 in the pooling layer.

4. Supervised learning with transfer learning

4.1 Research using original data

H. Zhu et al.[27] applied a CNN model for defects (normal, overlap, spatter, porosity) classification on weld surfaces of Gas Metal Arc Welding (GMAW). Transfer learning was performed using LeNet-5, and since the softmax function, which is frequently used in the last layer of the final fully connected layer of the CNN model, has a disadvantage of performance degradation when the number of training datasets is insufficient, Random Forest and SVM classifier were used for comparison of the performance. Preprocessing was performed such as median filtering, image enhancement by gradation processing, and OTSU image thresholding for all collected images

Table 3 Summary of research using original data and transfer learning

Process	Input	Output	Base network	No. of dataset	Ref. No
Gas Metal Arc Welding	80×120 (px) image from weld surfaces	Classification - Weld quality (Normal, Overlap, Spatter, Porosity)	LeNet-5 + Classification (Softmax, Random Forest, SVM)	320 (TR*) 80 (TE**)	27
Laser Welding	150×150 (px) image from weld surfaces	Classification - Weld quality (Normal, Porosity, Level misalignment)	AlexNet, VGG-16, Resnet-50, Densenet-121, MobileNetV3-Large	7217 (TR) 814(TE)	28
Gas Tungsten Arc Spot Weldng	640×480 (px) in-situ image of weld pool	Classification - Weld penetration (Under, Desirable, Excessive penetration)	ResNet 18-layers model	25643 (TR) 2851 (TE)	29
Laser Welding	224×224 (px) in-situ image from a coaxial camera	Classification - Pore or No pore	AlexNet	1971 (TR) 1276 (TE)	30

* TR: For training; ** TE: For Test

of defects on weld surfaces. From the results obtained from softmax, SVM and Random Forest classifiers, the accuracy obtained using the softmax classifier was the lowest and the accuracy of the proposed Random Forest Classifier showed the best performance.

Y. Yang et. al[28] used photographs of laser welding area taken by CMOS digital camera and transfer learning of a CNN model was performed for classification of weld quality into 3 classes (normal, porosity, level misalignment) and 2 classes (qualified, defect). In the basic VGG-16 model, the weight of the convolutional layer was fixed and training was performed only on the weight of the FC layer. The result of this model was compared with the other eight models (pre-optimization AlexNet, Resnet-50, VGG-16, Densenet-121, MobileNetV3-Large and post-optimization pre-AlexNet, pre-Resnet-50, and pre-Densenet-121). The 2-class binary classification achieved accuracy of 90% or higher in all cases. For multiclass classification of 3-class, Resnet-50 and MobileNetV3-Large were the most advanced deep learning algorithm and showed good performance, but 3 GPUs were needed for the training. The optimized VGG-16 model in the study showed excellent performance with training time of 1 hour without having to use GPU.

W. Jiao et al.[29] developed a weld penetration prediction model using CNN in the GTA spot welding process. Images were collected using the camera positioned on the top-side of the specimen, and the penetration state (partial, full and excessive penetration) was predicted through correlation with the area of the bright parts from the camera on the back side. For transfer learning, ResNet model with 18 layers was used and compared with a 9-layer CNN model. The accuracy of the ResNet model with transfer learning was 96.35%, and the test accuracy of the compared CNN model was 92.7%, confirming the superiority in the results obtained with transfer learning.

B. Zhang et al.[30] constructed a model for prediction of the in-situ internal porosity status of aluminum 6061 alloy in laser welding. Consisting of 532 nm light and high-speed camera, in-situ weld-pool images were collected, and the training was performed using porosity occurrence information confirmed through image processing of the longitudinal cross-section. AlexNet-based transfer learning was performed, and most of the pores were successfully detected, but during the verification process, the pores with size of 100 μm or less, and some deep pores were not detected.

4.2 Data augmentation

N. Yang et al.[31] proposed a LeNet-5-based transfer learning CNN model for defect classification using X-ray weld images. From the images acquired through the experiment, images around the weld area were extracted and the contrast in the images was improved by adjusting the gray level. Since there were not enough samples in the dataset, the images were rotated, shifted and blurred and noise was also added to increase the size of the dataset.

For optimization of the CNN model, each size of convolution kernels was varied for evaluation. The larger the kernel size, there was a slight increase in accuracy and the convergence became slower. A function that synthesized LReLU and Softplus with reference to 0 was used for classification to prevent saturation and it was confirmed that this function had superior performance than sigmoid function. When compared with the accuracy values of LeNet-5, ANN and SVM (0.758, 0.897, 0.982), the developed model showed a superior accuracy at 0.993.

C. V. Dung et al.[32] developed a CNN model for detection of fatigue cracks from photographic images of welded joints in a structure. From the images collected from the structure, sub-images of 64x64-pixel were extracted and datasets were constructed by classification according to the crack occurrence. In order to resolve the problem of imbalanced training dataset due to smaller number of images with cracks than images without cracks, data augmentation was performed through rotation, shift, shear, zoom, and flip of the crack images. CNNs of three different structures which are shallow CNN, VGG-16-based BN(Bottle Neck) training, and VGG-16-based fine-turning training were evaluated. In all cases, the results with data augmentation showed

Table 4 Summary of research using data augmentation and transfer learning

Process	Input	Output	Base network	Data augmentation	Ref. No
X-ray Inspection	60×60 (px) X-ray images	Classification - Weld quality (2, 5 classes)	LeNet-5	Rotation, Translation Blurr, Noise	31
Transmission Welds	64×64 (px) image from surfaces	Classification - Crack, No Crack	VGG-16	Rotation, Shift Shear, Zoom, Flip	32

that the accuracy improved by 2% or higher.

5. Unsupervised learning with autoencoder

CNN models are applied to predict or classify specific values in welding process but there are also cases when CNN models are used for feature extraction from acquired images. Autoencoder is unsupervised learning using CNN and reconstructs the original images through encoding and decoding processes. In the process, constraints such as noise are applied so that key features can be applied when the original images are reconstructed. Autoencoder is used when extracting features of the original images in high dimension or when reduced images are needed with the features retained.

J. Guenther et al.[11] applied autoencoder method with CNN for feature extraction in laser welding and extracted 16 features from the input image. The extracted 16 features were used for reinforcement learning, and the weld quality was achieved by controlling the output in laser welding.

A. Muniategui et al.[33] developed a quality determination model of resistance spot welding using a fuzzy algorithm. However, the fuzzy algorithm that uses the original images with high resolution of 1000x800-pixel as input is not efficient in terms of computation time for application in real-time process and the data storage space problem may arise due to the high storage capacity required. To overcome these limitations, an autoencoder was used to reduce the image resolution to 15x15 pixels while retaining the features of the image and the reduced images were applied to the fuzzy algorithm. Finally, quality prediction was achieved at an accuracy of 88% or higher in resistance spot welding.

6. Summary and Outlook

As can be seen from the review in this study, among various types of deep learning techniques, CNN models with images have been actively applied in welding research. The field of Welding is typically dominated by tacit knowledge rather than rule-based explicit knowledge, indicating high applicability of data-based deep learning, and the applications of CNN in welding research are expected to increase even further. The in-situ measurement of waveform-based time series signals have been actively used in determination of weld quality in previous studies, and in the future, there will be increased investigation and adoption of multi-sensor-based deep learning techniques where continuous waveform sensors and image sensors are applied simultaneously. Furthermore, at present, the images at the time of measurement are used for quality classification or regression, but in the future, hybrid models of combining RNN and CNN will be applied, leading to more intelligent models in which the information extracted from images in the past will be transferred to the current state prediction.

Acknowledgement

This research was supported by the MOTIE (Ministry of Trade, Industry, and Energy) in Korea, under the Fostering Global Talents for Innovative Growth Program (P0008750) supervised by the Korea Institute for Advancement of Technology (KIAT)

ORCID: Kidong Lee: http://orcid.org/0000-0001-7565-1965
ORCID: Soong Keun Hyun: http://orcid.org/0000-0002-3434-8465
ORCID: Cheolhee Kim: http://orcid.org/0000-0003-4127-3171

References

1. P. Kim, MATLAB Deep Learning, *Apress*, Seoul, Korea (2017).
https://doi.org/10.1007/978-1-4842-2845-6_5
2. D. E. Rumelhart, G. E. Hinton and R. J. Williams, Learning Representations by Back-Propagating Errors, *Nature*, 323(6088) (1986) 533-536.
https://doi.org/10.1038/323533a0
3. K. andersen, G. E. Cook, G. Karsai and K. Ramaswamy, Artificial Neural Networks Applied to Arc Welding Process Modeling and Control, *IEEE Trans. Ind. Appl.* 26(5) (1990) 824-830.
https://doi.org/10.1109/28.60056
4. G. E. Cook, R. J. Barnett, K. Andersen and A. M. Strauss, Weld Modeling and Control Using Artificial Neural Networks, *IEEE Trans. Ind. Appl.* 31(6) (1995) 1484-1491.
https://doi.org/10.1109/28.475745
5. H. S. Moon and S. J. Na, A Neuro-Fuzzy Approach to Select Welding Conditions for Welding Quality Improvement in Horizontal Fillet Welding, *J. Manuf. Syst.* 15(6) (1996) 392-403.
https://doi.org/10.1016/S0278-6125(97)83053-1
6. K. Bae and S. J. Na, A Study of Vision-Based Measurement of Weld Joint Shape Incorporating the Neural Network, *Proc. Inst. Mech. Eng., Part B: J. Eng. Manuf.* 208(1) (1994) 61-69.
https://doi.org/10.1243/PIME_PROC_1994_208_060_02
7. S. Lee, W. Chang, W. Yoo and S. J. Na, A Study on a Vision Sensor Based Laser Welding System for Bellows, *J. Manuf. Syst.* 19(4) (2000) 249-255.
https://doi.org/10.1016/S0278-6125(01)80004-2
8. A. Krizhevsky, I. Sutskever and G. E. Hinton, Imagenet Classification with Deep Convolutional Neural Networks, *Advances in Neural Information Processing systems,* (2012), Lake Taho, Navada, USA, 1097-1105.

https://doi.org/10.1145/3065386
9. K. Cho, B. Van Merrieboer, C. Gulcehre, D. Bahdanau, F. Bougares, H. Schwenk and Y. Bengio, Learning Phrase Representations Using RNN Encoder-Eecoder for Statistical Machine Translation, *arXiv preprint arXiv: 1406.1078* (2014).
https://doi.org/10.3115/v1/d14-1179
10. J. Weston, S. Chopra and A. Bordes, Memory Networks, *arXiv preprint arXiv:1410.3916* (2014).
11. J. Guenther, P. M. Pilarski, G. Helfrich, H. Shen and K. Diepold, Intelligent Laser Welding Through Representation, Prediction, and Control Learning: an Architecture with Deep Neural Networks and Reinforcement Learning, *Mechatron.* 34 (2016) 1-11.
https://doi.org/10.1016/j.mechatronics.2015.09.004
12. M. S. Kim, S. M. Shin, D. H. Kim and S. Rhee, A Study on the Algorithm for Determining Back Bead Generation in GMA Welding Using Deep Learning, *J. Weld. Join.* 36(2) (2018), 74-81.
https://doi.org/10.5781/JWJ.2018.36.2.11
13. D. Petkovic, Prediction of Laser Welding Quality by Computational Intelligence Approaches, *Optik*, 140 (2017) 597-600.
https://doi.org/10.1016/j.ijleo.2017.04.088
14. J. Chen, T. Wang, X. Gao and L. Wei, Real-Time Monitoring of High-Power Disk Laser Welding Based on Support Vector Machine, *Comput. Ind.* 94 (2018) 75-81.
https://doi.org/10.1016/j.compind.2017.10.003
15. F. Chollet, Deep Learning With Python, *Manning Publications Co.*, New York, NY, USA (2018).
16. Z. Guo, S. Ye, Y. Wang and C. Lin, Resistance Welding Spot Defect Detection with Convolutional Neural Networks, *International Conference on Computer Vision Systems* (2017) Shenzhen, China, 169-174.
https://doi.org/10.1007/978-3-319-68345-4_15
17. A. Khumaidi, E. M. Yuniarno and M. H. Purnomo, Welding Defect Classification Based on Convolution Neural Network (CNN) and Gaussian Kernel, *International Seminar on Intelligent Technology and Its Applications* (2017) Surabaya, Indonesia, 261-265.
https://doi.org/10.1109/ISITIA.2017.8124091
18. Y. Zhang, D. You, X. Gao, N. Zhang and P. P. Gao, Welding Defects Detection Based on Deep Learning with Multiple Optical Sensors During Disk Laser Welding of Thick Plates, *J. Manuf. Syst.* 51 (2019) 87-94.
https://doi.org/10.1016/j.jmsy.2019.02.004
19. Y. Zhang, D. You, X. Gao and S. Katayama, online Monitoring of Welding Status Based on a DBN Model During Laser Welding, *Eng.* 5(4) (2019), 671-678.
https://doi.org/10.1016/j.eng.2019.01.016
20. D. Bacioiu, G. Melton, M. Papaelias and R. Shaw, Automated Defect Classification of Aluminium 5083 TIG Welding Using HDR Camera and Neural Networks, *J. Manuf. Process.* 45 (2019) 603-613.
https://doi.org/10.1016/j.jmapro.2019.07.020
21. S. Choi, I. Hwang, Y. Kim, B. Kang and M. Kang, Prediction of the Weld Qualities Using Surface Appearance Image in Resistance Spot Welding, *Met.* 9(8) (2019) 831.
https://doi.org/10.3390/met9080831
22. Z. Zhang, B. Li, W. Zhang, R. Lu, S. Wada and Y. Zhang, Real-Time Penetration State Monitoring Using Convolutional Neural Network for Laser Welding of Tailor Rolled Blanks, *J. Manuf. Syst.* 54 (2020) 348-360.
https://doi.org/10.1016/j.jmsy.2020.01.006
23. W. Hou, Y. Wei, Y. Jin and C. Zhu, Deep Features Based on a DCNN Model for Classifying Imbalanced Weld Flaw Types, *Measurement*, 131 (2019), 482-489.
https://doi.org/10.1016/j.measurement.2018.09.011
24. J. K. Park, W. H. An and D.J. Kang, Convolutional Neural Network Based Surface Inspection System for Non-Patterned Welding Defects, *Int. J. Precis. Eng. Manuf.* 20(3) (2019) 363-374.
https://doi.org/10.1007/s12541-019-00074-4
25. T. Ashida, A. Okamoto, K. Ozaki, M. Hida and T. Yamashita, Development of Image Sensing Technology for Automatic Welding (Image Recognition by Deep Learning), *Kobelco Technol. Rev.* 37(April) (2019), 77-81.
26. Z. Zhang, G. Wen and S. Chen, Weld Image Deep Learning-Based On-Line Defects Detection Using Convolutional Neural Networks for Al Alloy in Robotic Arc Welding, *J. Manuf. Process.* 45 (2019) 208-216.
https://doi.org/10.1016/j.jmapro.2019.06.023
27. H. Zhu, W. Ge and Z. Liu, Deep Learning-Based Classification of Weld Surface Defects, *Appl. Sci.* 9(16) (2019) 3312.
https://doi.org/10.3390/app9163312
28. Y. Yang, L. Pan, J. Ma, R. Yang, Y. Zhu, Y. Yang and L. Zhang, A High-Performance Deep Learning Algorithm for the Automated Optical Inspection of Laser Welding, *Appl. Sci.* 10(3) (2020) 933.
https://doi.org/10.3390/app10030933
29. W. Jiao, Q. Wang, Y. Cheng and Y. Zhang, End-To-End Prediction of Weld Penetration: A Deep Learning and Transfer Learning Based Method, *J. Manuf. Process.* (2020).
https://doi.org/10.1016/J.Jmapro.2020.01.044
30. B. Zhang, K. M. Hong and Y. C. Shin, Deep-Learning-Based Porosity Monitoring of Laser Welding Process, *Manuf. Lett.* 23 (2020) 62-66.
https://doi.org/10.1016/j.mfglet.2020.01.001
31. N. Yang, H. Niu, L. Chen and G. Mi, X-Ray Weld Image Classification Using Improved Convolutional Neural Network, *AIP Conference Proceedings*, 1995 (2018), P. 020035.
https://doi.org/10.1063/1.5048766

32. C. V. Dung, H. Sekiya, S. Hirano, T. Okatani and C. Miki, A Vision-Based Method for Crack Detection in Gusset Plate Welded Joints of Steel Bridges Using Deep Convolutional Neural Networks, *Autom. Constr.* 102 (2019) 217-229. https://doi.org/10.1016/j.autcon.2019.02.013

33. A. Muniategui, B. Heriz, L. Eciolaza, M. Ayuso, A. Iturrioz, I. Quintana and P. Lvarez, Spot Welding Monitoring System Based on Fuzzy Classification and Deep Learning, *IEEE International Conference on Fuzzy Systems,* Naples, Italy, (2017) 1-6. https://doi.org/10.1109/fuzz-ieee.2017.8015618

합성곱 신경망기반 딥러닝의 용접연구 적용
Part I: 모델과 활용사례

Review on the Recent Welding Research with Application of CNN-Based Deep Learning
Part 1: Models and Applications

이기동*,** · 이 성* · 현승균** · 김철희*,***

*포틀랜드주립대학 기계재료공학과
**인하대 신소재공학과
***한국생산기술연구원 용접접합그룹

1. 서 론

최근 다양한 분야에 적용되고 있는 딥러닝 (Deep learning)은 인공지능(Artificial Intelligence, AI)의 일종인 기계학습(Machine learning)의 하나이다. 기계학습은 명시적 규칙에 기반하지 않고 데이터를 이용하여 규칙을 학습하는 방식으로 학습데이터의 종류에 따라 지도학습, 비지도학습, 강화학습으로 나눠진다. 지도학습은 학습데이터로 입력과 정답을 넣어주는데 분류와 회귀 문제에 적용되고 있다[1].

기계학습의 일종인 인공신경회로망(Artificial Neural Network, ANN)은 인간의 신경망을 모사하여 개발되었으며 입력층(input layer), 은닉층(hidden layer), 출력층(output layer)으로 구성된다. 이때 은닉층이 하나인 경우를 얕은 신경망 (Shallow Neural Network, SNN)으로, 2개 이상인 경우를 심층 신경망 (Deep Neural Network, DNN)으로 구분되며, 심층신경망을 이용하여 기계학습을 수행할 경우를 딥러닝이라 불린다.

ANN을 용접에 적용하려는 연구는 1980년 중반 오차의 역전파 알고리즘[2] 이 소개된 이후 다양한 공정에 활발하게 적용되었다. 제어변수부터 용접품질의 예측[3,4], 원하는 용접품용접부터 적절한 제어변수의 역도출[5], 레이저 비젼센싱 이미지에서 형상 추출[6], 이미지센서를 이용한 품질 및 용접선 추적[7] 등에 활용되었다. 그러나 당시에는 신경망 모델을 확장하여 적용하는데 몇 가지 문제가 있었다. 먼저 신경망의 활성화함수로 시그모이드 (Sigmoid) 함수나 RBF (Radial Basis Function) 함수가 주로 사용되었고, 역전파 알고리즘을 통한 심층 신경망의 학습에서 출력층에서 먼 가중치(weight)의 학습이 어려운 문제가 있었다. 또한 이미지나 시계열신호를 사용하는 경우에서도 특징(feature)을 추출하지 않는 일반화된 신호 보다는 특징 추출을 통해 학습이 이루어졌고, 주로는 제어변수와 같은 스칼라 변수에서 결과를 예측하는 모델이 개발되어 학습된 모델을 다른 생산시스템에서 적용하기는 쉽지 않았다. 모델에서 제어변수만을 이용할 경우에는 환경적인 외란(습도, 온도, 전력의 불안정성 등)을 고려할 수 없으며, 동일한 제어변수를 설정하더라도 서로 다른 시스템에서는 동일한 시스템 입력을 설정하는 것이 쉽지 않았다.

최근에는 활성화함수로 ReLu (Rectified Linear unit)가 제안되어 심층 신경망의 학습이 가능하게 되었으며, 이미지 인식에 주로 사용하는 합성곱 신경망 (Convolutional Neural Network, CNN)[8]과 시계열 변수의 모델에 많이 사용하는 순환신경망(Recurrent Neural Network, RNN)[9,10]이 제안되면서 종단간 (end-to-end) 학습을 포함한 딥러닝기술이 혁신적으로 발전하고 있다.

용접분야에서도 기계학습을 이용한 다양한 작업들이 이루어지고 있는데 레이저 용접제어에서 딥러닝과 강화학습[11], 아크용접에서 딥러닝을 통한 품질 예측[12], 비신경망 기계학습방법인 SVM (Supported Vector Machine)[13,14]을 적용한 용접 품질 분류 등이 적용되고 있다. 본 논문에서는 최근에 발표된 18개의 CNN 기반한 용접연구 논문을 소개하고자 한다. Fig. 1과 같이 지도학습, 전이학습(transfer learning)의 적용여부 및 데이터 전처리에서 데이터증식(data augmen-

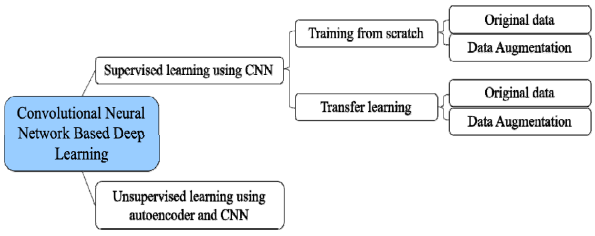

Fig. 1 Classification of papers reviewed

tation) 여부를 기준으로 5개 그룹으로 논문을 구분하였다.

2. CNN 구조 및 학습기법

CNN은 합성곱 연산을 사용하는 인공신경망의 한 종류로 이미지 인식에서 강력한 성능을 발휘하는 딥러닝 모델이다. 합성곱 필터를 통한 학습으로 이미지가 가진 특징 패턴을 추출할 수 있다. 또한 특징을 직접 학습하기때문에 특징을 수동으로 추출할 필요가 없으며, 기존 네트워크를 바탕으로 모델을 구성할 수 있다는 장점이 있어 이미지를 사용하는 딥러닝에 CNN을 사용하는 사례가 급증하고 있다.

2.1 기본 구조

CNN은 합성곱층(convolutional layer)와 풀링층(pooling layer), 완전연결층(fully connected layer)로 구성되어있다. 합성곱층에서는 입력값이 커널(kernel)안의 가중치로 연결되어 이미지의 영역을 강조하는 특성맵을 출력하여 다음 층으로 전달한다. 합성곱의 계산은 일정한 간격(stride)으로 커널윈도를 입력데이터 위로 움직이며 수행된다. 합성곱층의 결과를 그대로 다음 층에 전달할 수 있지만, 대부분의 경우 다운샘플링(down sampling)의 한 종류인 풀링층을 거치면서 과학습(over-fitting)을 막고 노이즈에 강한 피처를 생성한다. 풀링은 최댓값 풀링(max pooling), 평균값 풀링(mean pooling) 등이 있으며 지정된 커널윈도 크기에서 각각 최댓값, 평균값을 출력한다. 완전연결층은 일반적인 신경망과 같은 방식으로 이전 층의 노드와 다음 층의 노드가 모두 연결되는 층이다. CNN 모델에서 합성곱층과 풀링층의 반복은 이미지내의 국부적인 특징을 포함하면서 전역적인 특징도 표현하여 완전연결층의 입력으로 전달하고 완전연결층은 최종 결과를 출력하게 된다.

2.2 리뷰논문들에 사용된 기법들

아래에서 모델링과 학습에 사용된 기법들을 간단히 소개하였으며, 더 자세한 내용은 참고문헌에서 확인할 수 있다[15].

2.2.1 활성화 함수(Activation function)

활성화 함수는 신경망의 노드에서 입력값을 변환하는 함수로 항등함수, 시그모이드, Tanh, ReLu, 소프트맥스(softmax) 함수 등이 사용된다. ANN은 일반적으로 비선형 함수를 활성화함수로 채택하여 복잡한 비선형현상을 학습한다. 그 중 시그모이드함수는 모든 값을 확률로 변환하는 데 유용하며 0 또는 1사이의 출력값이 발생해 이진 분류에 사용할 수 있다. Tanh함수는 시그모이드함수와 같이 매끈한 형태를 가졌으며 입력값에 따라 -1에서 1사이의 출력값이 발생하므로 음의 값이

Table 1 Summary of research using original data and training from scratch

Process	Input	Output	Network	No. of dataset	Ref. No
Electric Resistance Welding	256×256 (px) image from cross-sections	Classification - Weld quality (Good/Defect)	CNN-FCN 13 HL***	11026 (TR*) 304 (TE**)	16
Arc Welding	320×240 (px) image from weld surfaces	Classification - Weld quality (Good, Porosity, Undercut, Spatter)	CNN-FCN 5 HL	120	17
Laser Welding	351 features of in-situ signals from photo diodes, image sensor, and spectrometer	Classification - Weld quality (Good, Blowout, Humping, Undercut)	CNN-FCN 7 HL	7500 (TR) 5000 (TE) from 25 welding runs	18
Gas Tungsten Arc Welding	400×487 (px) in-situ weld pool image	Classification - Weld quality (2, 4, 6 classes)	12 models (6 CNNs & 6 FCNs)	26666 (TR) 6588 (TE)	20
Resistance Spot Welding	128×128 (px) image from weld surfaces	Regression - Strength - Nugget diameter Classification - Fracture mode, Expulsion	CNN-FCN 11 HL	90	21
Laser Welding	121×121 (px) in-situ image of weld pool including keyhole	Classification - Penetration (2, 3, 4 classes)	10 Models (6 CNNs, {HOG, LBP, BOF, or SAE} + SVM)	10,000 (TR) 1000 (TE)	22

* TR: For training; ** TE: For test; *** HL: Hidden Layer

생성되는 장점이 있다. 시그모이드함수와 tanh 함수는 포화되어 기울기가 0이 될 가능성이 있어 심층 신경망의 학습에서 불리하다. 그러나 ReLU 함수는 입력이 0 이하면 0을 출력하고, 0 이상이면 입력값 그대로 출력하는 함수로 계산 비용이 적고 포화되지 않는 강점이 있다. 분류문제에 사용하는 소프트맥스 함수는 입력값을 0과 1 사이의 값으로 정규화하며 출력값들의 총합은 항상 1이 되는 특성을 가진 함수인데, 각각의 입력값의 편차를 확대시켜 큰 값은 상대적으로 더 크게, 작은 값은 상대적으로 더 작게 만들어 분류를 위한 신경망의 마지막 층에 사용된다.

2.2.2 전이 학습

CNN을 사용하는 이미지 분류 모델 설계에서 복잡한 CNN구조를 처음부터 학습하지 않고 기존의 VGG나 GoogLeNet, MobileNet 등 대규모의 데이터로 학습되어 있는 모델을 불러와서 학습하는 방법으로 높은 정확도를 비교적 짧은 시간에 달성할 수 있도록 해준다. 전이학습은 미세 조정(fine-tuning)을 통해 모델을 최적화를 수행한다. 미세 조정 전략은 전체 모델을 재학습하는 방법과 합성곱층 일부와 완전연결 층의 파라미터를 조정한 뒤 학습, 합성곱층은 유지하고 완전연결층의 파라미터만 조정한 뒤 학습하는 세 가지 방법으로 나누어진다. 이 논문에서는 편의상 전이학습을 하지 않고 초기단계부터 훈련(training from scratch)을 하는 경우를 초기훈련을 적용한 학습이라고 표기하였다.

2.2.3 데이터증식(Data augmentation)

딥러닝에서는 학습 데이터가 많을수록 정확도에 유리하고 부족한 학습 데이터는 과학습을 유발하므로 많은 학습 데이터를 필요로 한다. CNN은 입력 데이터로 이미지를 활용하기 때문에 학습 데이터가 부족할 때 데이터 증식을 통해 부족한 입력을 보충해줄 수 있다. 데이터 증식은 입력 이미지에 반전, 회전, 밝기 조정 등의 효과를 주어 다양한 환경의 이미지를 학습 데이터에 추가하는 방법이다. 이 논문에서는 편의상 데이터증식을 하지 않고 원래의 데이터를 이용하는 방식을 원본 데이터 활용 학습이라고 표시하였다.

3. 초기훈련 (training from scratch)을 적용한 지도학습

3.1 원본 데이터 활용

Z. Guo et al.[16]은 라인파이프 제조공정의 전기저항용접에 CNN을 적용하여 정상 용접과 비정상 용접을 분류하는 모델을 제안하였고 정확도 99.01%를 달성하였다. 용접부의 단면이미지를 입력으로 사용했으며 출력층에 시그모이드함수를 사용하여 정상 용접과 비정상 용접을 분류하였다. CNN의 은닉층의 수를 5, 7, 13개로 구성하여 네트워크가 깊어짐에 따른 정확도의 차이를 확인하였다. 네트워크 깊이가 깊어짐에도 불구하고 에포크(epoch)당 평균 시간은 5, 7, 13층 각각 261.68 s, 266.24 s, 267.26 s로 차이가 크지 않았으나 오차는 7.89%, 4.93%, 0.99%로 정확도가 매우 상승하였다. 사람이 분류작업 시 오차인 2.3%와 비교하여 정확도가 높은 결과를 얻을 수 있었다고 보고하였다.

A. Kjumaidi et al.[17]은 아크 용접비드 표면을 촬영한 이미지를 입력으로 사용하여 CNN 결함 분류 모델을 제안하였고, 3층의 CNN 은닉층과 2층의 완전연결층의 은닉층만을 이용해서도 네 가지 분류(양호, 기공, 언더컷, 스패터)에 대해 테스트 정확도 95.83%를 달성하였다.

Y. Zhang et al.[18]은 실시간 레이저 용접 중 다중센서 기술을 사용하여 용접 신호를 계측하고 이를 입력으로 하는 CNN 결함 감지 모델을 제안하였다. 레이저 용접 공정 중 포토다이오드와 스펙트로미터, 카메라를 사용하여 다중센서 시스템을 구축하였다. 수집된 포토다이오드 신호를 WPD 방법(Wavelet Packet Decomposition method)으로 주파수별로 분해하여 320개의 특징신호를 얻었다. 또한 스펙트로미터 신호의 400 nm에서 900 nm 파장 사이를 25개의 밴드로 나누어 수집하였다. 976 nm의 레이저 조명과 밴드패스필터를 적용한 1번 카메라로 키홀의 크기와 위치를 수집하고, 320 nm-750 nm의 파장 대역의 2번 카메라에서는 스패터와 플룸 관련 정보를 수집했다. 다양한 센서로부터 총 351개를 500 Hz 주파수로 수집한 후 나머지는 1로 채워 400개의 입력 노드를 구성했다. 저자는 CNN의 모델의 정확성을 검증하기 위해 같은 입력노드로 구성된 완전연결망(FCN) 모델[19]과 비교 실험하였고 FCN 모델의 평균 오차는 9.65%, CNN 모델의 평균 오차는 4.7%로 CNN 모델이 FCN 모델보다 우수하였다. 제안된 CNN 모델의 일반화를 위해 4가지 용접조건에서 실험하여 검증하였다. CNN 모델 분류 결과 각 분류(양호, 블로우아웃, 험핑, 언더컷)에서 1.73 %, 8.0 %, 11.2 % 및 8.75 % 오차를 나타냈다. 분류 오차가 발생하는 경우를 분석하였을 때, 블로우아웃 결함과 언더컷결함이 혼동되거나 험핑결함과 양호이 혼동되는 경우가 발생하였다.

D. Bacioiu et al.[20]는 알루미늄 5083합금의 GTA

용접에서 용접 결함을 분류하는 다양한 DNN 모델을 구성하고 비교하였다. 총 60번의 용접 실험에서 용접 중 실시간 용융풀 이미지를 입력으로 사용하여 6가지 CNN 모델과 6가지 FCN모델로 2 분류(양호, 불량), 4 분류(양호, 용락, 오염, 융합불량), 6 분류(양호, 용락, 오염, 융합불량, 오정렬, 용입부족) 모델을 구성하고, 입력 이미지 해상도, 모델의 깊이, 학습률(learning rate)에 따른 성능을 비교하였다. 입력의 해상도 평가에서 800×974 픽셀의 원본데이터를 400×487, 25×30 픽셀로 샘플링한 이미지를 비교한 결과 6클래스 분류에서는 400×487 픽셀 샘플링이 더 우수한 성능을 보였으며, CNN 모델이 FCN 모델보다 최종 분류 정확도에 민감한 것을 확인했다. 모든 모델에서 모델이 깊어질수록 정확도가 증가하였고, 10^{-3}이하의 학습률에서는 학습이 수렴하지 않았다. 모든 평가에서 CNN 모델들이 FCN 모델보다 높은 정확도를 보여주었다.

S. Choi et al.[21]은 980 MPa 인장강도를 갖는 자동차용 강판의 저항점부의 용접표면의 이미지를 통해 용접 품질(인장전단강도, 너겟크기, 파단모드, 날림발생 유무)를 예측하는 CNN모델을 연구하였다. 일반적으로 표면의 압흔과 열 자취는 용접품질과 상관관계를 가지고 있으므로 CNN모델을 통해 인장전단강도, 너겟크기, 파단모드, 날림발생 유무를 각각 98.6%, 98.8%, 100%, 100%의 정확도로 예측하였다. 학습에 포함되지 않은 데이터를 활용한 검증평가를 통해 모델의 일반성을 확인하였고, 검증결과 인장 전단강도, 너겟 크기에 대해 각각 97.8%, 97.4%의 정확도를 갖는 것을 확인하였다.

Z. Zhang et al.[22]은 TRB(tailor-rolled blank) 공정 중 레이저 용접의 용입상태를 진단하기 위해 CNN 모델을 사용했다. 용입은 2개(부분용입, 완전용입), 3개(부분용입, 적정-완전용입, 과용입), 4개(부분용입, 적정용입, 완전용입, 과용입)로 분류하고 커널 개수와 합성곱층의 형상이 서로 다른 6개의 CNN 모델을 학습하고 비교평가하였다. 학습결과 복잡성의 증가가 정확도의 향상을 반드시 가져오지 않음을 확인했고, 정확도와 지연시간을 고려할 때 CNN 모델 중 하나가 최적이라고 판단하였다. 또한 CNN 모델의 복잡성을 제외하고 지연시간에 영향을 줄 수 있는 입력 이미지의 촬영 속도(100 fps, 300 fps, 3000 fps)를 비교 실험하였고, 유의미한 차이가 없어 가장 지연시간이 낮은 100 fps를 선택했다. CNN 모델의 유용성을 검증하기 위해 다른 방법들[(HOG, LBP, Bag of features (BOF),or SAE)+SVM]과 비교하였다. 검증 결과 2 분류에서의 CNN 모델이 예측 정확도가 가장 높았고, 3 분류에서도 CNN 모델의 예측 정확도가 각각 99%, 93.8%, 96.5%로 다른 모델에 비해 높았다. 침투 상태가 4 분류일 때는 각 모델의 정확도가 비교적 낮았으나 CNN 모델의 정확도가 가장 안정적이었다.

3.2 데이터 증식

W. Hou et al.[23]은 GDXray의 데이터베이스에 공개된 다양한 용접의 엑스레이 이미지로부터 결함 (양호, 용입부족, 기공, 슬래그 혼입, 균열) 분류모델을 연구하였다. 먼저 88개의 엑스레이 이미지로부터 3503개의 32×32픽셀의 이미지를 서브샘플링(subsampling)하여 추출하였다. 각 결함의 분포가 불균형하므로 ROS (Random OverSampling), RUS(Random UnderSampling), SMOTE(Synthetic Minority Oversampling

Table 2 Summary of research using data augmentation and training from scratch

Process	Input	Output	Network	Data augmentation	Ref. No
X-ray Inspection	32×32 (px) X-ray images from GDXray-Weld	Classification - Weld quality (2, 5 classes)	5 Models (Haralick, HOG, SSAE, 2 CNNs)	Subsampling ROS**, RUS*** SMOTE*4	23
Transmission Welds	80×256 (px) for center detection 32x352 (px) for defect inspection from weld surfaces	Regression - Center point Classification - Weld quality (Good, NG)	6 Models (HOG+SVM, LBP+SVM, 4 CNNs)	Rotation Translation Scaling Brightness Adj. Contrast Adj.	24
Flux Cored Arc Welding	150×125 (px) in-situ image of weld pool	Regression - 8 molten pool features	CNN-FCN 10 HL*	Horizontal flip Translation Scaling	25
Gas Tungsten Arc Weldng	100×100 (px) in-situ images of weld pool in three-way	Classification - Weld quality (Under-penetration, Full penetration, Burn through)	CNN-FCN 10 HL	Noise addition Rotation	26

* HL: Hidden Layer; ** ROS: Random Over-Sampling; *** RUS: Random Under-Sampling;
*4 SMOTE: Synthetic Minority Oversampling TEchnique

TEchnique)방법을 통해 재샘플링(resampling)하였다. RUS 방법은 가장 간단한 방법으로 대다수의 클래스에서 샘플을 무작위로 제거하여 클래스 분포를 유지하는 방법이고, ROS 방법은 소수 샘플에서 선택한 샘플을 복제하여 원래 세트에 추가하여 클래스의 분포를 유지하는 방법이다. 그리고 SMOTE 방법은 유클리드 거리를 계산하여 k-근접 이웃(k-nearest neighbors)를 무작위로 찾아 새로운 인공샘플을 생성하는 재샘플링 방법이다. 분류를 위한 모델로는 전통적 방법인 Haralick 방법, HOG (Histograms of Oriented Gradients) 방법과 딥러닝 방법인 SSAE(Stacked Sparse Auto-Encoders) 방법과 깊이가 다른 CNN모델 두 개를 사용하였다. 각 모델을 통해 추출된 특징 결과를 시각적으로 비교하기위해 t-SNE 방법을 사용해 각 부분을 2D맵으로 시각화 하였으며 정확도를 비교하였다. 원본 이미지의 t-SNE 맵은 무질서하며 Haralick과 HOG 모델의 t-SNE 맵 또한 데이터 분류 가능성이 약한 모습을 확인하였다. SSAE와 CNN 모델 두 개 모두 t-SNE 맵에서 좀 더 명확한 분류 가능성을 보였다. CNN 모델 중에서는 더 깊은 모델에서 더 명확하게 분류되었다. 이 중 결함이 없는 샘플과 기공 샘플의 분류가 가장 명확했고 균열 결함이 가장 식별하기 어려웠다. ROS 데이터셋의 경우 Hararick, HOG, SSAE, CNN1, CNN2 각각 결함에 대한 평균 정확도는 60.9%, 81.6%, 91.4%, 94.7%, 96.3%로 나타났고 RUS 데이터셋의 경우 60.7%, 67.5%, 87.3%, 76.6%, 79.9%, SMOTE 데이터셋의 경우 63.1%, 82.2%, 90.5%, 94.4%, 97.2%로 나타났다. 결론적으로 심층 모델이 전통적인 모델보다 특징 추출 기능이 우수하며, SMOTE 방법을 사용할 경우 데이터 불균형으로 인한 정확도 감소 문제를 해결할 수 있다고 보고하였다.

J. Park et al.[24]은 엔진 트랜스미션 용접부의 결함을 검출하기 위해 2 단계의 CNN 모델을 제안하였다. 첫번째 CNN 모델에서는 수집된 엔진 트랜스미션의 원형 용접부의 이미지를 극좌표에서 직교좌표로 변환하고 중심점을 예측하였으며, 용접부 폭이 고정되었다는 가정하에 용접부를 제외한 배경을 제거하여 두번째 결함 검출 CNN모델로의 입력을 최적화하였다. 전통적인 방법인 휴 원(Hough circle)과 비교하여 어두운 조명 환경과 흐린 이미지에서 강점을 가지는 것을 확인하였다. 학습데이터에서 불량데이터가 많지 않아 이미지 변환(회전, 평행이동, 확대/축소)과 이미지 왜곡(밝기 및 대비 조정)를 하여 성능을 향상하였다. 두번째의 결함 검출 CNN 모델에서는 미니배치 경사하강법이 데이터 불균형에 의해 편향될 수 있어 각 배치의 OK, NG 이미지의 비율을 일정하게 구성하는 샘플링(class-specific batch sampling)을 사용하였다. 결함검출 모델로는 HOG+SVM, LBP(Local Binary Pattern)+SVM과 4가지 CNN을 구성하여 성능을 비교하였다.

T. Ashida et al.[25]은 FCA 맞대기 용접의 자동화를 위해 CNN 모델을 사용하여 용융풀의 특징점 추출하고 용융풀의 선단과 아크중심의 거리, 용융풀의 폭을 계산하여 백비드형상을 제어하는 시스템을 연구했다. 수평대칭, 평행이동, 확대/축소를 이용하여 2,400개 데이터를 12,000개 데이터로 증식하였다. 제안한 CNN 모델과 기존의 이미지 처리기술을 이용한 방법과 비교 평가했을 때 용융풀의 선단위치에 대한 평균 오차를 0.3 mm 만큼 줄일 수 있었다. 3 mm에서 10 mm로 용접선을 따라 갭이 증가하고, 용접선에서 10 mm 벗어난 맞대기 용접부에 해당기술을 적용하여 용융풀을 인식하고 용접속도, 중심 위치를 제어하여 용접선 추적

Table 3 Summary of research using original data and transfer learning

Process	Input	Output	Base network	No. of dataset	Ref. No
Gas Metal Arc Welding	80×120 (px) image from weld surfaces	Classification - Weld quality (Normal, Overlap, Spatter, Porosity)	LeNet-5 + Classification (Softmax, Random Forest, SVM)	320 (TR*) 80 (TE**)	27
Laser Welding	150×150 (px) image from weld surfaces	Classification - Weld quality (Normal, Porosity, Level misalignment)	AlexNet, VGG-16, Resnet-50, Densenet-121, MobileNetV3-Large	7217 (TR) 814(TE)	28
Gas Tungsten Arc Spot Weldng	640×480 (px) in-situ image of weld pool	Classification - Weld penetration (Under, Desirable, Excessive penetration)	ResNet 18-layers model	25643 (TR) 2851 (TE)	29
Laser Welding	224×224 (px) in-situ image from a coaxial camera	Classification - Pore or No pore	AlexNet	1971 (TR) 1276 (TE)	30

* TR: For training; ** TE: For Test

성능과 백비드 형성 능력을 검증하였다.

Z. Zhang et al.[26]은 알루미늄 합금의 펄스 GTA 용접 공정 중 용접 결함 예측을 위해 CNN 모델을 제안했다. CNN 모델의 학습 속도와 정확도를 높이기 위해 용융풀의 상부 중심, 용융풀의 전방 및 용접심의 뒷부분 3방향으로부터 동시에 정보 수집하기 위한 시스템을 개발하여 적용했다. 용접 결함을 만들기 위해 직사각형, 아령형, 사다리형 시편에 대해 용접을 수행하였다. 1392×1040 픽셀의 원본 이미지를 산술 평균으로 다운샘플링하여 100×100 픽셀의 입력 이미지로 사용했다. 데이터 증식을 위해 노이즈(salt and pepper noise)추가와 15도, 30도 회전을 입력 이미지에 무작위로 적용하여 데이터 증식을 하였다. 데이터 증식 후 92.3%에서 96.18%로 정확도가 3.88% 증가하여 데이터 증식의 유효성을 검증하였다. 또한 CNN 모델을 최적화하기 위해 4개의 합성곱층과 커널 개수를 변수로 두고 정확도와 최종 학습 속도를 비교하였다. 학습 시간과 효율성은 커널수에 따라 증가했으며, 풀링층에서 2D 이미지가 0으로만 이루어진 이미지를 제외한 특징 이미지 개수가 많을수록 정확도가 높았다.

4. 전이학습을 이용한 지도학습

4.1 원본 데이터 활용

H. Zhu et al.[27]는 GMAW 용접부 표면 결함(양호, 오버랩, 스패터, 기공)을 분류하기 위해 CNN 모델을 활용하였다. LeNet-5을 이용하여 전이학습을 하였으며, CNN 모델의 마지막 완전연결층의 마지막 층에 많이 사용되는 소프트맥스 함수가 훈련 데이터 수가 부족한 경우 성능이 감소하는 단점을 가지고 있어 랜덤 포레스트(Random Forest)와 SVM 분류기를 사용하여 성능을 비교하였다. 수집된 모든 용접 표면 결함 이미지는 중간값 필터링, 계조 강화, OTSU 임계처리 등의 전처리과정을 거쳤다. 소프트맥스, SVM, 랜덤포레스트 분류기 중 소프트맥스 분류기의 정확도가 가장 낮게 나왔으며, 제안한 랜덤포레스트 분류기의 정확도가 가장 높았다.

Y. Yang et. al[28]은 CMOS 카메라로 촬영한 레이저용접부의 표면을 이용하여 용접품질을 3클래스(정상, 기공, 높이단차)와 2클래스(정상, 불량)으로 분류하기 위해 CNN 모델의 전이학습을 수행하였다. 기본 VGG-16모델에서 합성곱층의 가중치는 고정하고 완전연결층의 가중치만 학습하였다. 그 외 다른 8개의 모델(최적화전 AlexNet, Resnet-50, VGG-16, Densenet-121, MobileNetV3-Large와 최적화 후 pre-AlexNet, pre-Resnet-50, pre-Densenet-121)과 비교되었다. 2클래스 이진분류는 모두 90% 이상의 정확도를 달성했다. 3클래스 다범주 분류에서는 Resnet-50 및 MobileNetV3-Large가 최신 딥러닝 알고리즘으로 성능은 우수하나 3개의 GPU를 이용하여 학습이 필요했다. 연구에서 최적화된 VGG-16모델은 GPU없이 1시간에 학습이 가능하면서도 우수한 성능을 보여주었다.

W. Jiao et al.[29]은 GTA 점용접 공정에서 CNN을 활용한 용입상태 예측모델을 수립하였다. 시편의 상부에 위치한 카메라에서 이미지를 수집하고 용입상태(과소, 적정, 과도 용입)는 이면의 카메라에서의 밝은 부분의 면적과 상관관계를 통해 추정되었다. 전이학습을 위해 18-layer로 구성된 ResNet 모델을 사용하였으며, 9-layer의 CNN 모델과 비교하였다. 환경에서 전이학습된 ResNet 모델의 정확도는 96.35 %이고 비교 CNN 모델의 테스트 정확도는 92.7 %로 전이학습의 장점을 확인하였다.

B. Zhang et al.[30]은 알루미늄 6061합금의 레이저 용접 중 내부기공 여부를 예측하는 모델을 수립하였다. 532 nm 조명과 고속카메라로 구성되어 용접 중 용접풀 이미지를 수집하고, 종단면의 영상처리를 통해 확인한 기공 여부를 이용하여 학습하였다. AlexNet을 기반으로 전이학습을 하였으며, 대부분의 기공을 감지했지만 검증 과정 중 기공의 크기가 100 μm 이하이고, 용입방향으로 깊은 곳에 분포한 일부 기공을 감지하지 못하는 경우가 발생하였다.

4.2 데이터 증식

N. Yang et al.[31]은 용접부 X-ray를 사용한 결함

Table 4 Summary of research using data augmentation and transfer learning

Process	Input	Output	Base network	Data augmentation	Ref. No
X-ray Inspection	60×60 (px) X-ray images	Classification - Weld quality (2, 5 classes)	LeNet-5	Rotation, Translation Blurr, Noise	31
Transmission Welds	64×64 (px) image from surfaces	Classification - Crack, No Crack	VGG-16	Rotation, Shift Shear, Zoom, Flip	32

분류를 위해 LeNet-5를 기반한 전이학습 CNN 모델을 제안하였다. 실험을 통해 획득한 이미지로부터 용접부 근처를 추출하고 계조 조정을 통해 이미지 내 대조를 강화하였다. 데이터 개수가 충분하지 못해 회전, 병진, 블러, 노이즈 추가를 통해 데이터를 증식하였다. CNN 모델의 최적화를 위해 각 합성곱층의 커널크기를 변화하며 평가하였다. 커널 크기가 커질수록 정확도가 매우 약간 증가하고, 수렴속도가 늦어졌다. 분류 함수로 약한 ReLU와 소프트플러스(softplus)를 0을 기준으로 합성한 함수를 사용하여 포화를 방지하였으며, 기존 시그모이드 함수보다 우수한 성능을 확인하였다. 개발 모델은 LeNet-5, ANN, SVM의 정확도(0.758, 0.897, 0.982)와 비교하여 높은 정확도 0.993를 가졌다.

C. V. Dung et al.[32]은 구조물 용접부의 촬영 이미지로부터 피로 균열여부를 판별하기 위한 CNN 모델을 개발하였다. 구조물에서 수집된 이미지에서 64×64 픽셀의 서브이미지로 추출하고 균열여부를 기준으로 분류하여 데이터를 구축하였다. 균열 이미지가 비균열 이미지에 비해 부족하여 데이터 불균형 문제를 해결하기 위해 균열 이미지를 회전, 이동, 뒤틀림, 확대, 뒤집기를 통해 증식하였다. 얕은 CNN, VGG-16 기반 BN (Bottle Neck) 학습, VGG-16 기반 미세조정 학습 세 가지 구조의 CNN을 평가하였다. 모든 경우에서 데이터증식한 경우가 2% 이상 정확도가 향상됨을 확인하였다.

5. 오토엔코더를 이용한 비지도학습

용접 공정에서 특정한 값을 예측하거나 분류하기 위해서 CNN 모델을 적용하는 한편, 얻은 이미지에서 특징을 추출하기 위한 방법으로 CNN 모델을 사용한 경우도 있다. 오토엔코더는 CNN을 활용한 비지도 학습으로 원본 이미지를 엔코딩(encoding)과 디코딩(decoding) 과정을 거쳐 재구성한다. 그 과정에서 노이즈 등의 제약을 주어 원본 이미지를 재구성할 때 핵심 특징들이 추출될 수 있도록 구성된다. 고차원으로 원본이미지의 특징을 추출하거나, 특징을 유지한 압축 이미지가 필요할 때 사용된다.

J. Guenther et al.[11]은 레이저 용접의 특징추출에 CNN을 이용한 오토엔코더기법을 적용하여 입력이미지에서 16개의 특징을 추출할 수 있었다. 추출된 16개의 특징들은 강화학습에 사용되었으며 레이저용접에서 출력을 제어하여 용접품질을 확보하였다.

A. Muniategui et al.[33]는 퍼지 알고리즘을 활용한 저항점용접의 품질판단모델을 개발하였다. 그러나 1000×800 픽셀의 높은 해상도의 원본 이미지를 입력으로 사용하는 퍼지알고리즘은 실시간 공정에 적용하기에 연산시간에 너무 길고 높은 용량으로 인해 저장공간의 부족을 초래할 수 있다. 이를 극복하기 위해 오토엔코더를 이용하여 이미지의 특징을 유지한 채로 이미지의 해상도를 15×15 픽셀로 줄여 퍼지 알고리즘에 적용하였다. 최종적으로 저항점용접에서 88% 이상의 정확도로 품질예측을 달성하였다.

6. 시사점과 전망

본 연구에서 리뷰한 것과 같이 딥러닝 기술 중 이미지를 이용한 CNN의 적용이 용접분야 연구에 활발하게 적용되고 있다. 용접분야는 대표적으로 암묵지(tacit knowledge)의 영역이 넓은 분야로 규칙기반의 형식지보다는 데이터기반의 딥러닝의 적용가능성이 높은 분야이므로 향후 CNN의 적용이 활발할 것으로 예상된다. 용접공정 중 측정한 파형 기반의 연속신호는 기존에도 많이 용접품질 판단에 이용되어 왔으며, 향후에는 연속파형센서와 이미지센서를 동시에 활용하는 멀티센서에 기반한 딥러닝 기술이 보다 활발하게 도입될 것으로 예상된다. 또한 현재는 측정 시점의 이미지를 이용하여 품질판단이나 회귀에 활용하고 있으나 향후에는 RNN-CNN 하이브리드 모델을 적용하여 과거의 이미지로부터 추출한 정보를 현재의 모델에 반영하는 기술의 적용이 예상된다.

Acknowledgement

This research was supported by the MOTIE (Ministry of Trade, Industry, and Energy) in Korea, under the Fostering Global Talents for Innovative Growth Program (P0008750) supervised by the Korea Institute for Advancement of Technology (KIAT)

ORCID: Kidong Lee: http://orcid.org/0000-0001-7565-1965
ORCID: Soong Keun Hyun: http://orcid.org/0000-0002-3434-8465
ORCID: Cheolhee Kim: http://orcid.org/0000-0003-4127-3171

References

1. P. Kim, MATLAB Deep Learning, *Apress*, Seoul, Korea (2017). https://doi.org/10.1007/978-1-4842-2845-6_5
2. D. E. Rumelhart, G. E. Hinton and R. J. Williams, Learning Representations by Back-Propagating Errors, *Nature*, 323(6088) (1986) 533-536. https://doi.org/10.1038/323533a0
3. K. andersen, G. E. Cook, G. Karsai and K. Ramaswamy, Artificial Neural Networks Applied to Arc Welding

Process Modeling and Control, *IEEE Trans. Ind. Appl.* 26(5) (1990) 824-830.
https://doi.org/10.1109/28.60056
4. G. E. Cook, R. J. Barnett, K. Andersen and A. M. Strauss, Weld Modeling and Control Using Artificial Neural Networks, *IEEE Trans. Ind. Appl.* 31(6) (1995) 1484-1491.
https://doi.org/10.1109/28.475745
5. H. S. Moon and S. J. Na, A Neuro-Fuzzy Approach to Select Welding Conditions for Welding Quality Improvement in Horizontal Fillet Welding, *J. Manuf. Syst.* 15(6) (1996) 392-403.
https://doi.org/10.1016/S0278-6125(97)83053-1
6. K. Bae and S. J. Na, A Study of Vision-Based Measurement of Weld Joint Shape Incorporating the Neural Network, *Proc. Inst. Mech. Eng., Part B: J. Eng. Manuf.* 208(1) (1994) 61-69.
https://doi.org/10.1243/PIME_PROC_1994_208_060_02
7. S. Lee, W. Chang, W. Yoo and S. J. Na, A Study on a Vision Sensor Based Laser Welding System for Bellows, *J. Manuf. Syst.* 19(4) (2000) 249-255.
https://doi.org/10.1016/S0278-6125(01)80004-2
8. A. Krizhevsky, I. Sutskever and G. E. Hinton, Imagenet Classification with Deep Convolutional Neural Networks, *Advances in Neural Information Processing systems,* (2012), Lake Taho, Navada, USA, 1097-1105.
https://doi.org/10.1145/3065386
9. K. Cho, B. Van Merrieboer, C. Gulcehre, D. Bahdanau, F. Bougares, H. Schwenk and Y. Bengio, Learning Phrase Representations Using RNN Encoder-Eecoder for Statistical Machine Translation, *arXiv preprint arXiv: 1406.1078* (2014).
https://doi.org/10.3115/v1/d14-1179
10. J. Weston, S. Chopra and A. Bordes, Memory Networks, *arXiv preprint arXiv:1410.3916* (2014).
11. J. Guenther, P. M. Pilarski, G. Helfrich, H. Shen and K. Diepold, Intelligent Laser Welding Through Representation, Prediction, and Control Learning: an Architecture with Deep Neural Networks and Reinforcement Learning, *Mechatron.* 34 (2016) 1-11.
https://doi.org/10.1016/j.mechatronics.2015.09.004
12. M. S. Kim, S. M. Shin, D. H. Kim and S. Rhee, A Study on the Algorithm for Determining Back Bead Generation in GMA Welding Using Deep Learning, *J. Weld. Join.* 36(2) (2018), 74-81.
https://doi.org/10.5781/JWJ.2018.36.2.11
13. D. Petkovic, Prediction of Laser Welding Quality by Computational Intelligence Approaches, *Optik*, 140 (2017) 597-600.
https://doi.org/10.1016/j.ijleo.2017.04.088
14. J. Chen, T. Wang, X. Gao and L. Wei, Real-Time Monitoring of High-Power Disk Laser Welding Based on Support Vector Machine, *Comput. Ind.* 94 (2018) 75-81.
https://doi.org/10.1016/j.compind.2017.10.003
15. F. Chollet, Deep Learning With Python, *Manning Publications Co.*, New York, NY, USA (2018).
16. Z. Guo, S. Ye, Y. Wang and C. Lin, Resistance Welding Spot Defect Detection with Convolutional Neural Networks, *International Conference on Computer Vision Systems* (2017) Shenzhen, China, 169-174.
https://doi.org/10.1007/978-3-319-68345-4_15
17. A. Khumaidi, E. M. Yuniarno and M. H. Purnomo, Welding Defect Classification Based on Convolution Neural Network (CNN) and Gaussian Kernel, *International Seminar on Intelligent Technology and Its Applications* (2017) Surabaya, Indonesia, 261-265.
https://doi.org/10.1109/ISITIA.2017.8124091
18. Y. Zhang, D. You, X. Gao, N. Zhang and P. P. Gao, Welding Defects Detection Based on Deep Learning with Multiple Optical Sensors During Disk Laser Welding of Thick Plates, *J. Manuf. Syst.* 51 (2019) 87-94.
https://doi.org/10.1016/j.jmsy.2019.02.004
19. Y. Zhang, D. You, X. Gao and S. Katayama, online Monitoring of Welding Status Based on a DBN Model During Laser Welding, *Eng.* 5(4) (2019), 671-678.
https://doi.org/10.1016/j.eng.2019.01.016
20. D. Bacioiu, G. Melton, M. Papaelias and R. Shaw, Automated Defect Classification of Aluminium 5083 TIG Welding Using HDR Camera and Neural Networks, *J. Manuf. Process.* 45 (2019) 603-613.
https://doi.org/10.1016/j.jmapro.2019.07.020
21. S. Choi, I. Hwang, Y. Kim, B. Kang and M. Kang, Prediction of the Weld Qualities Using Surface Appearance Image in Resistance Spot Welding, *Met.* 9(8) (2019) 831.
https://doi.org/10.3390/met9080831
22. Z. Zhang, B. Li, W. Zhang, R. Lu, S. Wada and Y. Zhang, Real-Time Penetration State Monitoring Using Convolutional Neural Network for Laser Welding of Tailor Rolled Blanks, *J. Manuf. Syst.* 54 (2020) 348-360.
https://doi.org/10.1016/j.jmsy.2020.01.006
23. W. Hou, Y. Wei, Y. Jin and C. Zhu, Deep Features Based on a DCNN Model for Classifying Imbalanced Weld Flaw Types, *Measurement*, 131 (2019), 482-489.
https://doi.org/10.1016/j.measurement.2018.09.011
24. J. K. Park, W. H. An and D.J. Kang, Convolutional Neural Network Based Surface Inspection System for Non-Patterned Welding Defects, *Int. J. Precis. Eng. Manuf.* 20(3) (2019) 363-374.
https://doi.org/10.1007/s12541-019-00074-4
25. T. Ashida, A. Okamoto, K. Ozaki, M. Hida and T. Yamashita, Development of Image Sensing Technology for Automatic Welding (Image Recognition by Deep Learning), *Kobelco Technol. Rev.* 37(April) (2019), 77-81.

26. Z. Zhang, G. Wen and S. Chen, Weld Image Deep Learning-Based On-Line Defects Detection Using Convolutional Neural Networks for Al Alloy in Robotic Arc Welding, *J. Manuf. Process*. 45 (2019) 208-216. https://doi.org/10.1016/j.jmapro.2019.06.023
27. H. Zhu, W. Ge and Z. Liu, Deep Learning-Based Classification of Weld Surface Defects, *Appl. Sci.* 9(16) (2019) 3312. https://doi.org/10.3390/app9163312
28. Y. Yang, L. Pan, J. Ma, R. Yang, Y. Zhu, Y. Yang and L. Zhang, A High-Performance Deep Learning Algorithm for the Automated Optical Inspection of Laser Welding, *Appl. Sci.* 10(3) (2020) 933. https://doi.org/10.3390/app10030933
29. W. Jiao, Q. Wang, Y. Cheng and Y. Zhang, End-To-End Prediction of Weld Penetration: A Deep Learning and Transfer Learning Based Method, *J. Manuf. Process.* (2020). https://doi.org/10.1016/J.Jmapro.2020.01.044
30. B. Zhang, K. M. Hong and Y. C. Shin, Deep-Learning-Based Porosity Monitoring of Laser Welding Process, *Manuf. Lett.* 23 (2020) 62-66. https://doi.org/10.1016/j.mfglet.2020.01.001
31. N. Yang, H. Niu, L. Chen and G. Mi, X-Ray Weld Image Classification Using Improved Convolutional Neural Network, *AIP Conference Proceedings*, 1995 (2018), P. 020035. https://doi.org/10.1063/1.5048766
32. C. V. Dung, H. Sekiya, S. Hirano, T. Okatani and C. Miki, A Vision-Based Method for Crack Detection in Gusset Plate Welded Joints of Steel Bridges Using Deep Convolutional Neural Networks, *Autom. Constr.* 102 (2019) 217-229. https://doi.org/10.1016/j.autcon.2019.02.013
33. A. Muniategui, B. Heriz, L. Eciolaza, M. Ayuso, A. Iturrioz, I. Quintana and P. Lvarez, Spot Welding Monitoring System Based on Fuzzy Classification and Deep Learning, *IEEE International Conference on Fuzzy Systems,* Naples, Italy, (2017) 1-6. https://doi.org/10.1109/fuzz-ieee.2017.8015618