OXFORD CDE

Journal of Computational Design and Engineering, 2023, **10**, 1423–1442

DOI: 10.1093/jcde/qwad067

Advance access publication date: 6 July 2023

Research Article

Detecting balling defects using multisource transfer learning in wire arc additive manufacturing

Seung-Jun Shin [1], Sung-Ho Hong[2], Sainand Jadhav[3] and Duck Bong Kim[4,*]

[1]School of Interdisciplinary Industrial Studies, Hanyang University, Republic of Korea , 222 Wangsimni-ro, Seongdong-gu, Seoul 04763, Korea
[2]Department of Industrial Data Engineering, Hanyang University, Republic of Korea, 222 Wangsimni-ro, Seongdong-gu, Seoul 04763, Korea
[3]Department of Mechanical Engineering, Tennessee Technological University, 115 W. 10th Street Cookeville, TN 38505, USA
[4]Department of Manufacturing and Engineering Technology, Tennessee Technological University, 920 North Peachtree Ave., Cookeville, TN 38505, USA, USA
*Correspondence: dkim@tntech.edu

Abstract

Wire arc additive manufacturing (WAAM) has gained attention as a feasible process in large-scale metal additive manufacturing due to its high deposition rate, cost efficiency, and material diversity. However, WAAM induces a degree of uncertainty in the process stability and the part quality owing to its non-equilibrium thermal cycles and layer-by-layer stacking mechanism. Anomaly detection is therefore necessary for the quality monitoring of the parts. Most relevant studies have applied machine learning to derive data-driven models that detect defects through feature and pattern learning. However, acquiring sufficient data is time- and/or resource-intensive, which introduces a challenge to applying machine learning-based anomaly detection. This study proposes a multisource transfer learning method that generates anomaly detection models for balling defect detection, thus ensuring quality monitoring in WAAM. The proposed method uses convolutional neural network models to extract sufficient image features from multisource materials, then transfers and fine-tunes the models for anomaly detection in the target material. Stepwise learning is applied to extract image features sequentially from individual source materials, and composite learning is employed to assign the optimal frozen ratio for converging transferred and present features. Experiments were performed using a gas tungsten arc welding-based WAAM process to validate the classification accuracy of the models using low-carbon steel, stainless steel, and Inconel.

Keywords: wire arc additive manufacturing, anomaly detection, transfer learning, convolutional neural networks, domain adaptation

1. Introduction

Additive manufacturing (AM) is the process of joining materials to make parts from 3D model data, usually layer upon layer (ISO/ASTM52900:2021). AM facilitates the fabrication of complex shapes and reduces material waste (Qi *et al.*, 2019). AM processes include binder jetting, material extrusion, material jetting, powder bed fusion, sheet lamination, vat photopolymerization, and directed energy deposition (DED; ISO/ASTM52900:2021). Among them, DED has gained significant attention in the industry, as it enables high-precision fabrication of metal parts. DED uses an energy source to melt a feedstock material on a substrate, and the material is delivered to the energy-focused region (Bikas *et al.*, 2016). Typical DED sub-processes include laser-engineered net shaping, laser solid forming, directed light fabrication, direct metal deposition, electron beam AM, and wire arc additive manufacturing (WAAM; Ngo *et al.*, 2018).

WAAM benefits from arc welding technology, wherein a wired metal is the feedstock, an electric arc is the energy source, and a robot is the deposition operator, as shown in Fig. 1. Based on the energy source, WAAM can be further classified into gas metal arc welding, plasma arc welding, and gas tungsten arc welding (GTAW; Ikram & Chung, 2022). In GTAW, a tungsten inert gas (TIG) torch is attached to the end of the robot arm and supplied with arc energy. The feeding material is supplied by a wire feeder and deposited to form weld beads on the substrate. The major process parameters are travel speed, wire feed speed, and arc current. WAAM has advantages in deposition rate, material utilization, manufacturing cost, and material diversity; however, it has lower precision than laser-based AM (Liu *et al.*, 2020). Furthermore, WAAM has several problems, including lack of process stability and repeatability. These problems must be addressed to enable the broader applicability of WAAM (Jafari *et al.*, 2021).

Process uncertainty results in part quality issues in the WAAM process. This uncertainty primarily originates from the bead formation mechanism and the repeated melting and solidification of the feeding material. Beads are thermally deformed and microstructurally transformed between solid and liquid states due to the non-equilibrium thermal cycles. Factors such as material inhomogeneity, process parameter misassignment, and process setup misalignment can also induce uncertainties. These uncertainties can cause defects, such as pores, cracks, deformation, lack-of-fusion, oxidation, balling, and delamination (Xia *et al.*, 2020).

Quality assurance in WAAM is critical because defects induce issues in geometric dimensioning and tolerancing and poor mechanical properties (Liu *et al.*, 2020). The most feasible approach to mitigate these problems is post-processing treatments, such as heat treatment, interpass cooling, cold rolling, and peening (Vimal *et al.*, 2021). However, these post-processes are time- and

Received: March 19, 2023. **Revised:** June 23, 2023. **Accepted:** June 24, 2023

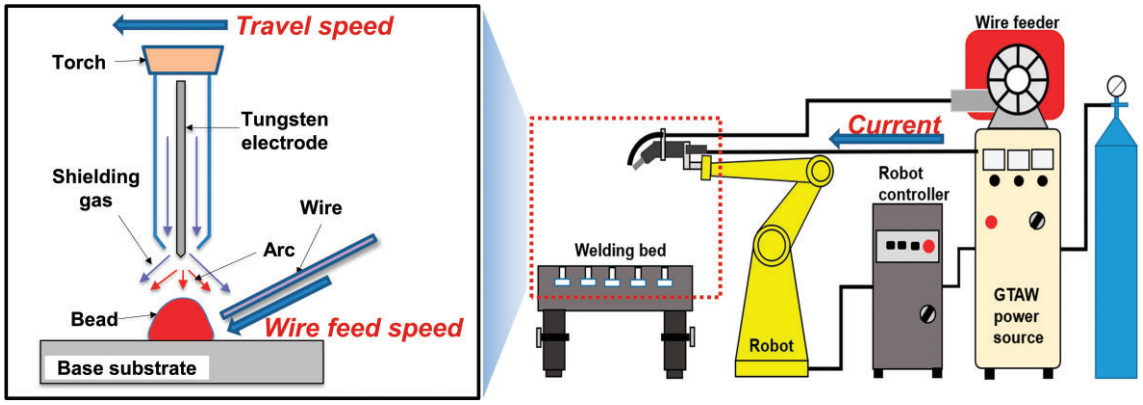


Figure 1: Schematic of gas tungsten arc welding.

resource-intensive, as they require large workloads. Accordingly, anomaly detection has emerged as a feasible and promising technology for quality assurance. Anomaly detection is a solution that detects abnormalities in processes or parts by discovering patterns in data that do not accord with a steady notion of normal behavior (Jin *et al.*, 2019). Anomaly detection ensures reliable quality performance and reduces the workloads of post-processing treatment, as it provides automatic defect detection during the process (Park & Lee, 2022).

The AM community has attempted to develop anomaly detection algorithms that classify the states as normal and abnormal (Wang *et al.*, 2020). Machine learning approaches based on sensor data have recently become a widespread tool for anomaly detection since these data can provide vital information related to heat transfer, thermal transformation, surface formation, structural characterization, and melt pool dynamics during the process (Xia *et al.*, 2020). Machine learning enables a computer to derive mathematical models by learning sensor data to make predictions without being explicitly programmed (Wang *et al.*, 2020).

However, machine learning demands sufficient data to achieve high performance (Liu *et al.*, 2020). This is expensive and time-consuming and may be impeded by technical difficulties because data acquisition systems should be installed, and the data should be obtainable and reliable. These problems can become more critical when the feeding material is new or expensive. Existing machine learning algorithms can be impractical as a result of these issues. Therefore, developing a cost-efficient and robust approach for anomaly detection is necessary.

Transfer learning, a branch of machine learning, is a feasible solution to this issue. It employs the knowledge extracted from source domains or tasks and transfers it to a target domain or task (Pan & Yang, 2009). In this regard, transfer learning enables the implementation of anomaly detection models for certain materials using data and models obtained from different materials. However, transfer learning may cause negative transfer, which must be avoided because it degrades learning performance in the target domain (Pan & Yang, 2009). Negative transfer occasionally occurs in single-source transfer learning due to a data bias dedicated to a specific source or a data dissimilarity between the source and target domains.

This study proposes a multisource transfer learning method to detect balling defects for quality monitoring in WAAM. The proposed method uses convolutional neural network (CNN) models to extract image features from two source materials with sufficient data, then transfers and fine-tunes the models to enable anomaly detection for a target material with insufficient data. In addition, the proposed method employs stepwise learning to extract image features from multiple source materials and composite learning to variably assign frozen ratios to frozen and unfrozen layers, in contrast to typical transfer learning. The method's input consists of image-type voltage data, which have been converted from numerical voltage data to ensure classification accuracy. Its output is a label representing the normal or abnormal state at the present datum. Experiments were performed using an in-house GTAW-based WAAM setup to generate and evaluate models in terms of classification accuracy using the following three materials: low-carbon steel (LCS), stainless steel 316L (STS), and Inconel 625 (INC).

The remainder of this paper is organized as follows. Section 2. summarizes related studies, Section 3 describes the experiments and data, Section 4 presents the method, Section 5 describes the experimental results, and Section 6 concludes the paper.

2. Related Work

2.1 Physical analysis in GTAW

GTAW is a process that deposits weld beads using a non-consumable tungsten electrode. The arc melts the feeding material and base substrate, creating a molten pool that solidifies to form a metallurgical bond (Cho *et al.*, 2022). Current, travel speed, and wire feed speed must be appropriately determined, as they affect the volumetric energy density, which influences bead formation. Equation (1) calculates the volumetric energy density (Q_e) using arc voltage (U) and current (I) and Equation (2) expresses the actual volume of the material concerning travel speed (v_{weld}) and wire feed speed (v_{wire}; Lu *et al.*, 2021).

$$Q_e = \frac{P_e}{V_{ALL}} = \frac{\eta UI}{V_{AM} + V_{pool}} \tag{1}$$

where $V_{ALL}$is the total volume of materials, P_e is the effective input power, V_{AM} is the actual volume of added material, V_{pool} is the surface melting volume under the substrate, and η is the arc efficiency.

$$V_{AM} = A_{cross} v_{weld} = A_{wire} v_{wire} \xi_{wire} \tag{2}$$

where A_{cross} is the cross-sectional area of bead, A_{wire} is the cross-sectional area of welding wire, and ξ_{wire} is the wire material utilization.

From Equations (1) and (2), it can be inferred that an increase in voltage and current results in an increase in energy density, whereas an increase in travel speed and wire feed speed leads to

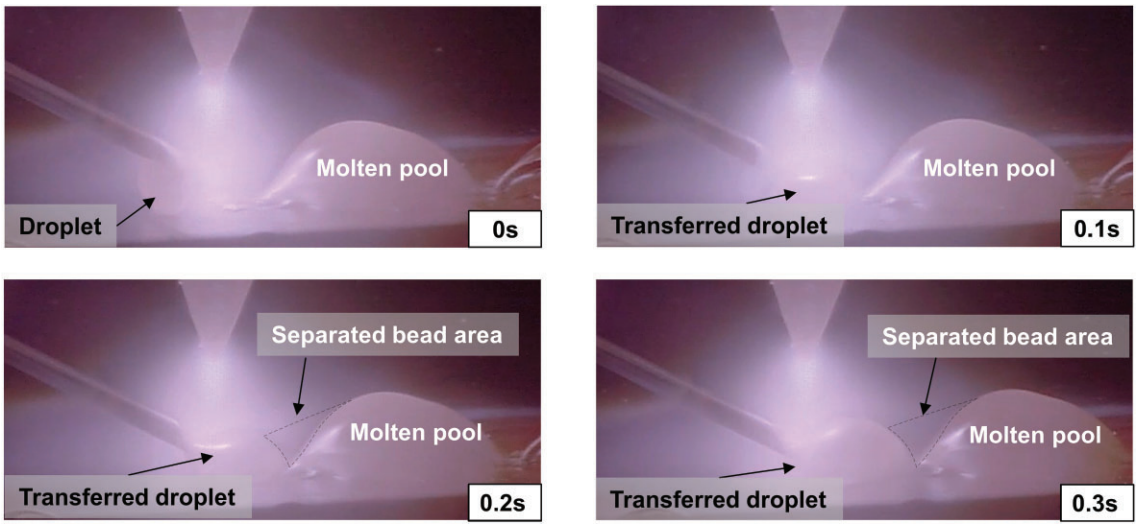


Figure 2: Behaviors of the molten pool.

a decrease in energy density. During the deposition process, current, travel speed, and wire feed speed are constant, while voltage depends on the distance between the tungsten electrode tip and the substrate. Although the average voltage is a typical value, it does not accurately estimate when the voltage values fluctuate. Equation (3) represents a revision of Equation (1), considering the voltage values measured at timestamps (i; Derekar, 2018).

$$Q_e = \frac{\eta \sum_{i=1}^{n} \left(\frac{U_i I_i}{n}\right)}{V_{ALL}} \tag{3}$$

Figure 2 presents snapshots of the molten pool behavior. The feeding material starts to melt in front of the arc center and forms a droplet at the tip of the wire. As the material moves under the arc, the droplet transition occurs continuously when the molten pool is positioned below the arc. The size and area of the molten pools slightly decrease when Q_e is reduced. As expressed in Equations (1) and (2), Q_e decreases when the travel and wire feed speeds increase, whereas Q_e decreases as the current is reduced. An insufficient energy density causes incomplete melting of the material, resulting in the balling effect (Lu *et al.*, 2021). This phenomenon indicates that the molten material does not form a flat layer but manifests spherical beads on the surface due to the formation of separate spherical droplets (Everton *et al.*, 2016). This is regarded as a severe defect in WAAM, as it deteriorates the surface roughness and creates discontinuities during depositing layers (Qiu *et al.*, 2020). Balling should therefore be detected through process monitoring and metrology using sensor data (Everton *et al.*, 2016).

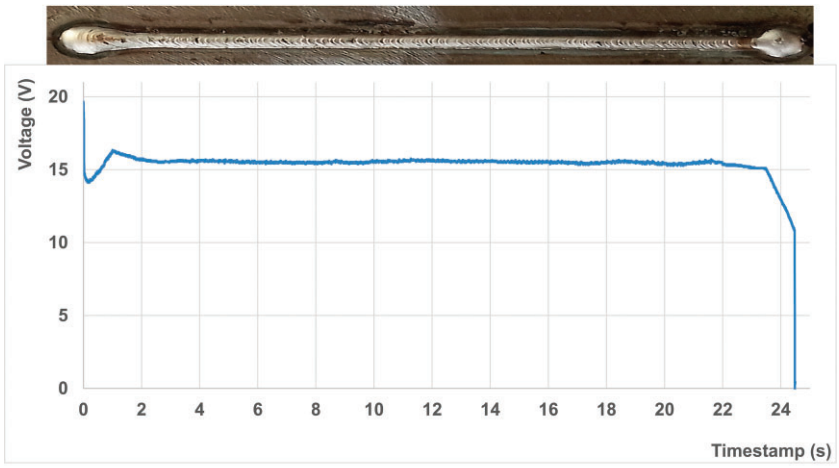
Voltage signals can detect and identify potential balling defects (Li *et al.*, 2022). Figure 3 illustrates the voltage profiles of two beads. The voltage profile tends to exhibit a stable and smooth pattern when the beads are well formed, as seen in Fig. 3a. In contrast, the profile can fluctuate or exhibit instability when the beads are not well-formed, as shown in Fig. 3b. The arc must be emitted uniformly onto the deposition area to sustain a uniform arc length between the electrode and deposited beads (Vimal *et al.*, 2021). The arc voltage at a constant current tends to be inversely proportional to the arc length (Xu *et al.*, 2018). Thus, it increases when the arc length is short during bead formation because of the balling defects. Conversely, the arc voltage decreases when the arc length increases in the separated region between the two balling volumes. These disparate data patterns can be crucial for identifying balling defects in beads. We note that normal and abnormal beads do not always correspond to normal or abnormal voltage patterns. Normal and abnormal beads can frequently exhibit unexpected patterns due to process instability, material impurities, and uncertainties. This voltage unusability makes it difficult to derive simple and intuitive models for anomaly detection.

2.2. Transfer learning in additive manufacturing

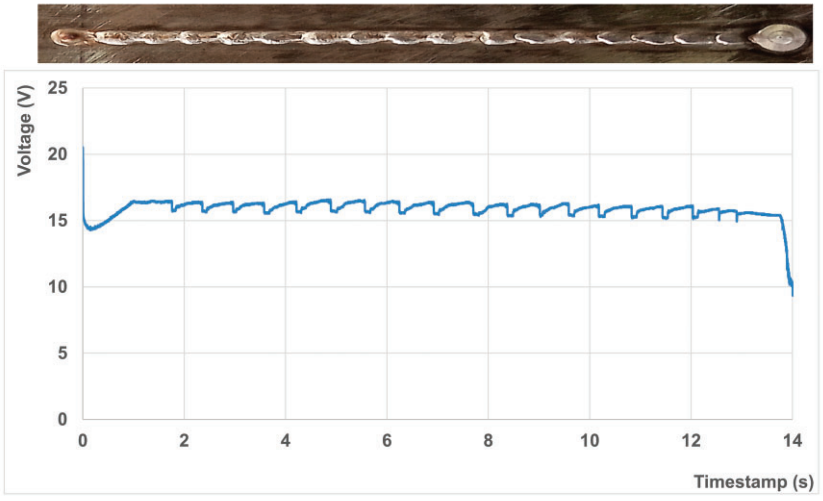
Machine learning, which requires sufficient data to achieve good performance, has been widely used for anomaly detection in AM. However, collecting sufficient data is expensive and challenging in many cases. Furthermore, data collection is unavailable when actual machining is not performed because of the deposition of new materials or part designs (Qin *et al.*, 2022). Transfer learning is a viable solution to resolve data scarcity by transferring knowledge from the source domains or tasks to the target domain or task (Colosimo *et al.*, 2018). The following transfer learning terms are used throughout this study (Pan & Yang, 2009):

- Transfer learning aims to perform the learning task $\mathcal{T}_T$ of the target predictive function $f_T(\cdot)$ in a target domain $\mathcal{D}_T$, using the knowledge in the source domain $\mathcal{D}_s$ and learning task $\mathcal{T}_s$, where $\mathcal{D}_s \neq \mathcal{D}_T$, or $\mathcal{T}_S \neq \mathcal{T}_T$.
- Domain ($\mathcal{D}$) comprises feature space $\mathcal{X}$ and probability function $P(X)$. $\mathcal{D}$ stands for different feature spaces or marginal probability distributions owing to the disparate contexts in which data are created.
- Task ($\mathcal{T}$) consists of label space $\mathcal{Y}$ and predictive function $f(\cdot)$. $\mathcal{T}$ denotes learning tasks, such as regression, clustering, classification, and prediction.
- Source (S) is a knowledge provider $\mathcal{D}_s$ for the target domain or a task $\mathcal{T}_s$ for the target task.
- Target (T) is a consumer $\mathcal{D}_T$ of preceding knowledge in the target domain, or a task of interest $\mathcal{T}_T$.

Figure 4 illustrates the two inductive and transductive transfer learning approaches. Inductive transfer learning induces a predictive model to use in the target domain when $\mathcal{T}_S \neq \mathcal{T}_T$ irrespective of the homogeneity between $\mathcal{D}_s$ and $\mathcal{D}_T$. Transductive transfer learning transduces a predictive model to use in the target domain when $\mathcal{T}_S = \mathcal{T}_T$ but $\mathcal{D}_s \neq \mathcal{D}_T$ (Kim *et al.*, 2022). The latter is



(a) Normal bead (travel speed: 40 cm/min, wire feed speed: 125 cm/min)



(b) Abnormal bead (travel speed: 70 cm/min, wire feed speed: 75 cm/min)

Figure 3: Sample voltage profiles (Inconel 625, current: 200 A).

also known as domain adaptation when $\mathcal{X}_S \neq \mathcal{X}_T$ or $\mathcal{X}_S = \mathcal{X}_T$, but $P(\mathcal{X}_S) \neq P(\mathcal{X}_T)$.

The AM community has applied transfer learning to two broad research streams: geometric shape prediction and anomaly detection. The former aims to improve the shape fidelity of existing and new products. Cheng *et al.* (2017) proposed a shape deviation modeling scheme using transfer learning in fused deposition modeling (FDM). They applied transfer learning to infer the deviation of a new shape from existing deviation models with decomposing the deviation into shape-independent and shape-specific errors. Zhu *et al.* (2018) developed new shape deviation models based on machine learning in FDM and used transfer learning to predict the deviation for a given shape, incorporating knowledge from other shapes. Ferreira *et al.* (2020) adopted the concept of transfer learning, called model transfer, to produce predictive models of geometric deviations for the adaptive model building of new shapes in mask image projection stereolithography. Cheng *et al.* (2021) developed a hybrid transfer learning framework that predicts and compensates in-plane shape deviations of new freeform products using a few fabricated products.

Transfer learning has been combined with machine learning to detect, classify, and predict abnormalities. Gonzalez-Val *et al.* (2020) suggested CNN-based dilution estimation models by learning image data in laser metal deposition and used transfer learning to find defects by adjusting the models to the target laser welding process. Scime *et al.* (2020) designed a machine-to-machine transfer learning method wherein a large amount of labeled data is available in one machine while a small amount of labeled data exists in another. They could transfer the knowledge across diverse AM machines to detect defects such as recoater hopping and streaking, incomplete spreading, swelling, spatter, spot, debris, super-elevation, part damage, and porosity. Ho *et al.* (2021) derived a porosity prediction model using thermal images of the melt pool in selective laser melting and applied transfer learning to improve the prediction accuracy of porosity. Pandiyan *et al.* (2022) suggested deep transfer learning to detect balling, lack-of-fusion, conduction mode, and keyhole pores in laser powder bed fusion. In their demonstration, the models learned from acoustic emission signals in stainless steel could be transferred to generate anomaly detection models for bronze. Their RestNet-18 models achieved average classification accuracies of 86.9% and 84.3% using 5200 and 2600 training samples in bronze, respectively, whereas 5200 samples were trained in stainless steel.

Colosimo *et al.* (2018) stated that statistical transfer learning frameworks need to be further developed to avoid data collection problems in AM. Although the applications of transfer learning

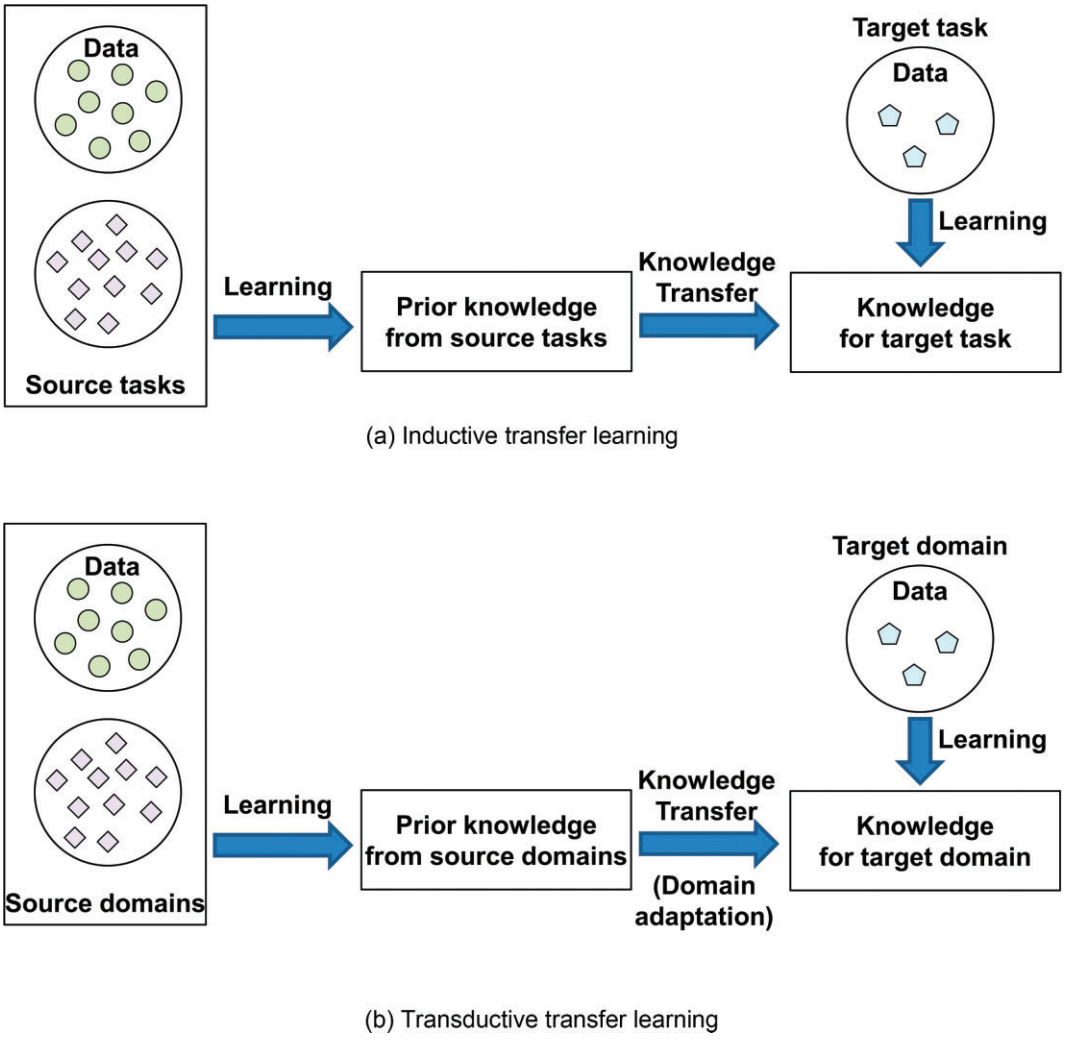


Figure 4: Transfer learning concept.

in AM have increased, they have rarely been reported in WAAM. This study was therefore conducted to develop anomaly detection models for a high-cost material (i.e., target domain) by transferring CNN-based models obtained from low-cost materials (i.e., source domains). Transfer learning was applied for the model transfer in this study. It enables adopting a well-trained CNN network from the source domains using sufficient data and employing this network in the target domain. In particular, features were extracted from the source domains and stored in hidden layers inside the network. They were then adjusted for normal or abnormal classifications in the target domain.

3. Experiments and Data Preparation

3.1. Experiments

Experiments were performed on a GTAW-based WAAM system, as illustrated in Fig. 5. Table 1 presents the system configuration. A TIG torch generates an arc that melts the feeding material and deposits a single-layer bead on a substrate, following the path designated by a robot program created using a robot teaching pendant. While the torch moves along its trajectory to fabricate a bead, a voltage sensor measures and records a set of pairwise timestamps and voltage values. These voltage data are monitored and collected by a data acquisition system. A high-dynamic-range (HDR) camera attached to the torch captures the bead images. This camera is specialized for arc welding and captures high-quality video frames at 50 frames per second (fps) in an environment where sparks and lightning interference commonly occur. These video frame data are also monitored and collected using a data interface.

The three process parameters were set as (1) 75–300 cm per minute (cpm) of wire feed speed with an increment of 25, (2) 10–100 cpm of travel speed with an increment of 10, and (3) 200 amperes (A) of the constant current. The two increments were selected because they could give considerable differences in bead geometry, such as bead height and width, between the current and two adjacent process parameters. A total of 300 beads were fabricated, 100 for each material (LCS, STS, and INC). Figure 6 presents examples of the beads single-layered and deposited on the INC.

3.2. Data preparation

3.2.1. *Data collection*

During the experiments, the voltage and bead image data were collected as single-layer beads were deposited on the substrate. While the torch deposited a feeding material along a bead, voltage data were measured at a time rate of 1 kHz and recorded in a text file. Simultaneously, the bead image data were measured along a bead at 50 frames per second (fps) and recorded in a MPEG-4 video

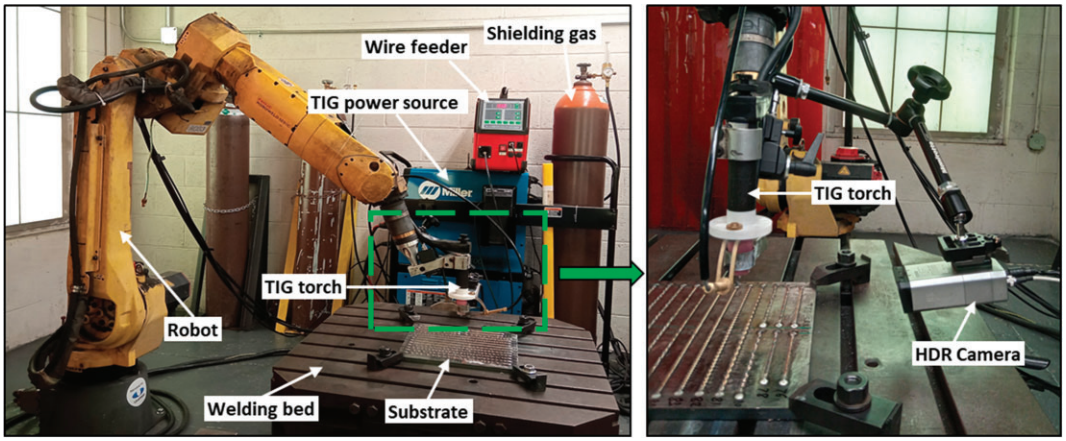


Figure 5: Experimental setup of wire arc additive manufacturing.

Table 1: System configuration for experiments.

System	Sub-system	Model	Description
Gas tungsten arc welding (GTAW)	Robot	Fanuc ArcMate 120iC	- Move the TIG torch with a six-axis movement
	Controller	Fanuc R-30iA	- Control the robot
	Torch	TIG	- Operate arc welding as a robot-hand effector - 70% Argon + 30% Helium shielding gas - 5 mm arc length
	Substrate	LCS	- 30 cm × 30 cm × 1.2 cm size
	Wire feeder	Generic wire feeder	- Feed a wire material
	Power source	Miller Dynasty400	- Supply energy source
Numerical data acquisition	Voltage sensor	Miller Insight ArcAgent Auto	- Measure arc voltage
	Voltage data interface	Miller Insight Centerpoint	- Monitor and acquire arc voltage data
Image data acquisition	Vision sensor	Weldvis WL2-H7machine learning-M35	- HDR camera for monitoring welding arc, molten pool, and weld bead images
	Camera data interface	Elgato Game Capture HD	- Capture and record image data

Figure 6: Examples of single beads.

file, which was re-partitioned into individual JPEG image files at 50 fps. Each image frame includes the region of metal transfer, arc shape, and weld pool, as depicted in Fig. 7. In this study, the voltage data and the bead image data were defined as follows:

- The voltage data indicate time-series voltage data, wherein each data point comprises a pair of timestamps and a numerical voltage value. A voltage profile is represented by a line graph where a set of voltage data is plotted on the time axis along a trajectory, as illustrated in Fig. 3.
- The bead image data represent a set of melt pool image data captured by the HDR camera and formatted as JPEG files. These data were used as a reference for labeling the voltage data.

3.2.2. *Voltage image data conversion*

Classification from time-series data generally depends on the model-, distance-, or feature-driven approaches. However, these approaches have limitations in practice; (1) they require time-series data to satisfy the stationary assumption, (2) the lengths of the time series need to be equal, and (3) the feature selection is difficult without discretization due to information loss (Zhao *et al.*, 2017). Therefore, the image-based approach is a good substitute for the time series classification. Specifically, CNN is an effective image-based learning network for extracting image features automatically from time-series data without prior knowledge (Koprinska *et al.*, 2018).

The time-series voltage data were converted into image-type data on the time domain. Each time-series voltage profile was converted and partitioned into a series of image snapshots along the timestamps. Figure 8 illustrates the process of image conversion. The bandwidth (w) and interval (i) parameters must be identified to build the image snapshots. These parameters correspond to the image size and frequency and are changeable. For example, as shown in Fig. 8, each image was generated consecutively at 1-second intervals ($i = 1$) and comprised numerical voltage samples over a 3-second duration ($w = 3$). The voltage image data were generated from the voltage data as specified below.

- The voltage image data represent image-type data converted from a voltage profile at a given bandwidth and interval. They were entered as the input to the proposed method. For example, Fig. 8 is an example of the voltage image data.

3.2.3. *Data labeling*

Data labeling should be accompanied by supervised learning, wherein parameters of the underlying classification scheme are trained using a dataset with known labels – that is, ground truth – and the trained classification scheme is estimated by generating predicted labels for a testing dataset (Gobert *et al.*, 2018). In supervised learning, each voltage image datum must be labeled with a normal, abnormal, or unclassified label for the ground truth. However, the direct labeling of voltage image data was intuitively difficult; therefore, indirect labeling was performed. This comes from the fact that direct labeling by humans for each voltage image datum is difficult. Meanwhile, the appearance of beads provides explicit decision clues for identifying the ground truth. In this sense, bead labeling is first performed based on appearance, bead image labeling is done using bead labels, and voltage image labeling is then carried out using bead image labels.

The bead labeling stage was performed by two experts, who evaluated all beads based on appearance and labeled them with appropriate classifiers. Here, well-formed beads were assigned as normal, beads that exhibited discontinuities or abnormalities were classified as abnormal, and beads that showed ambiguity between the two cases were assigned the unclassified label and excluded from the training dataset. All labels corresponded to the ground truth for the training data samples. Figure 9 presents an example of bead labeling. As shown in Fig. 9a and b, the beads are labeled normal because they are well-formed. In Fig. 9c and d, the beads are abnormal due to balling. In Fig. 9e and f, the beads are unclassified because of their state transition from normal to abnormal or vice versa on appearance.

The bead image labeling stage was performed using the bead labels above. All bead image labels corresponded to their respective bead labels. However, images were labeled as unclassified if they belonged to the starting and ending points of a bead geometry, as they were associated with a travel speed of zero. Figure 10 presents examples of the two beads shown in Fig. 3. As in the bead labeling stage, all images labeled normal or abnormal were included in the training dataset, whereas images labeled unclassified were excluded from the training dataset.

The voltage image data stage was performed according to the bead image labels via time synchronization. As shown in Fig. 11a and b, the voltage images are classified as normal because their bead images were labeled as normal. In Fig. 11c, the voltage image is abnormal because its bead image was abnormal.

3.2.4. *Dataset preparation*

The training and testing datasets comprising the voltage image data were prepared for learning and validation. A class imbalance problem frequently manifests during dataset preparation in a normal-biased environment, as the dataset may have a skewed ratio of majority to minority samples (Shorten & Khoshgoftaar, 2019). This can negatively affect the classification accuracy owing to overfitting, which induces a small learning error in training, but a high prediction error in testing. Oversampling and undersampling are two established techniques that mitigate class imbalance problems. Whereas the former increases the size of the minority to match the majority, the latter reduces the size of the majority to match the minority (Somasundaram and Reddy 2016).

As expected, more normal than abnormal voltage image data were generated in the experiments. Accordingly, undersampling was applied to eliminate the class imbalance problem by excluding some of the normal-labeled beads. Beads were selected for exclusion under consideration of the balance between the time required to produce normal and abnormal beads.

Table 2 lists the number of data samples in the training and testing datasets. The data samples in each material were obtained individually from 100 beads per material, as explained in Section 3.1. All testing datasets were constructed by a random selection. Each training dataset was further split into training and validation datasets, using a common ratio of 70:30. The validation datasets were used to measure the learning errors, which represent the classification accuracy of the models during learning. The testing datasets were used to measure the prediction errors, which indicate the classification accuracy of the models during use.

4. Proposed Method

This study aims to develop a multisource transfer learning method that derives anomaly detection models learned from multiple source materials and applies them to detect balling defects in a target material. Here, multisource signifies that the source domain contains two or more than two materials (two mate-

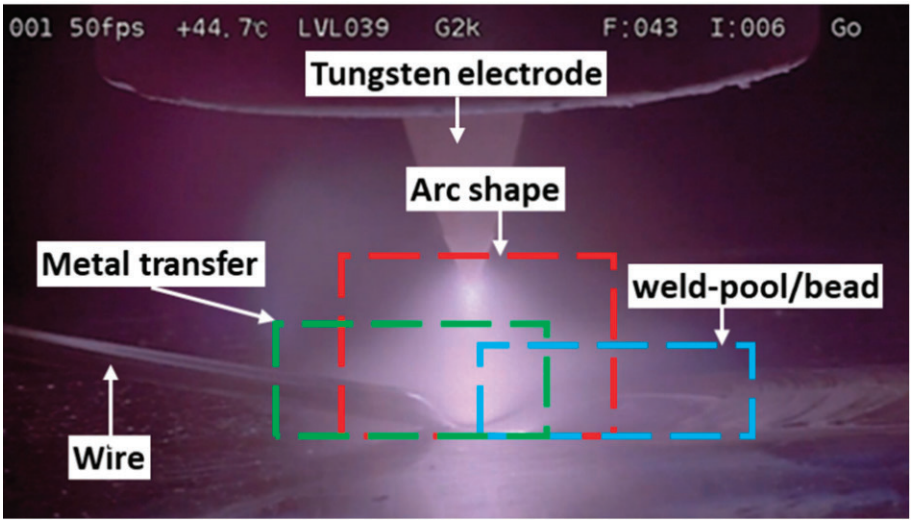


Figure 7: HDR image of the WAAM process.

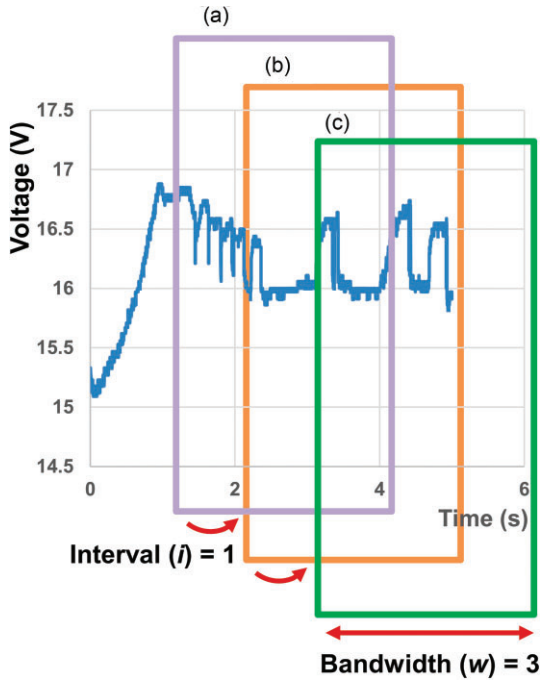


Figure 8: Example of image conversion.

rials in this study). This multisource approach is necessary to avoid the negative transfer. Multisources can reduce the risk of data bias and dissimilarity problems, as mentioned in Section 1, because they can provide the target domain with a positive possibility of data balance and similarity from the source domain. As discussed in Section 3.2.3, the proposed method accounts for a classification problem in supervised learning. It also accounts for domain adaptation, as although the feature spaces are equivalent in the source and target domains, their probability distributions can have a discrepancy owing to the application of different materials.

4.1. Architecture

Figure 12 illustrates the architecture of the proposed method. This method employs the stagewise and composite learning approaches. The stagewise approach involves pre-training a dataset with the source material and extracting image features from each source object in sequence without concatenating the two source datasets. This approach was applied to reduce the training time and increase the flexibility in configuring source materials. A single training session is required if a new material replaces the existing source material. The composite learning approach variably assigns frozen ratios to frozen and unfrozen layers to maximize model performance (Wang & Zhang, 2020; Wang *et al.*, 2020). Composite learning allows the flexible setting of frozen ratios inside the CNN structure to adaptively converge the image features of the prior material with those of the current material.

The proposed method comprises the following processes: (i) in the first stage, voltage image data of the first source material are entered into the CNN structure, which extracts image features from frozen layers; (ii) the extracted image features are transferred to the second stage; (iii) in the second stage, the transferred image features are set to image features at frozen layers inside the CNN; (iv) the voltage image data of the second source material are entered into unfrozen layers, and image features are extracted; (v) the extracted image features are transferred to the third stage; (vi) in the third stage, the image features transferred from the second stage are set to image features in frozen layers; (vii) voltage image data of the target material are adde to the unfrozen layers, and image features are extracted; (viii) the image features are fine-tuned toward the target material in the fully-connected layer; (ix) classification is conducted in the fully-connected layer, which outputs a higher probability of a normal or abnormal state.

4.2. Feature extraction

Image features are extracted using a CNN that learns from voltage image data. An appropriate CNN structure needs to be employed to extract image features efficiently. We chose DenseNet169 as the feature extractor, as it exhibits superior performance in image classification by extracting representative image features automatically through convolutional layers (Wang *et al.*, 2020).

Figure 13 illustrates the feature extraction in DenseNet169. The CNN connects all layers in a feed-forward approach and enables

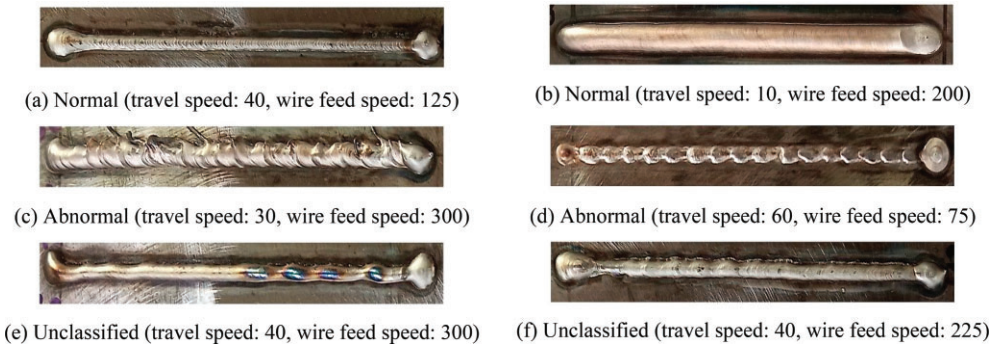


Figure 9: Examples of bead labeling.

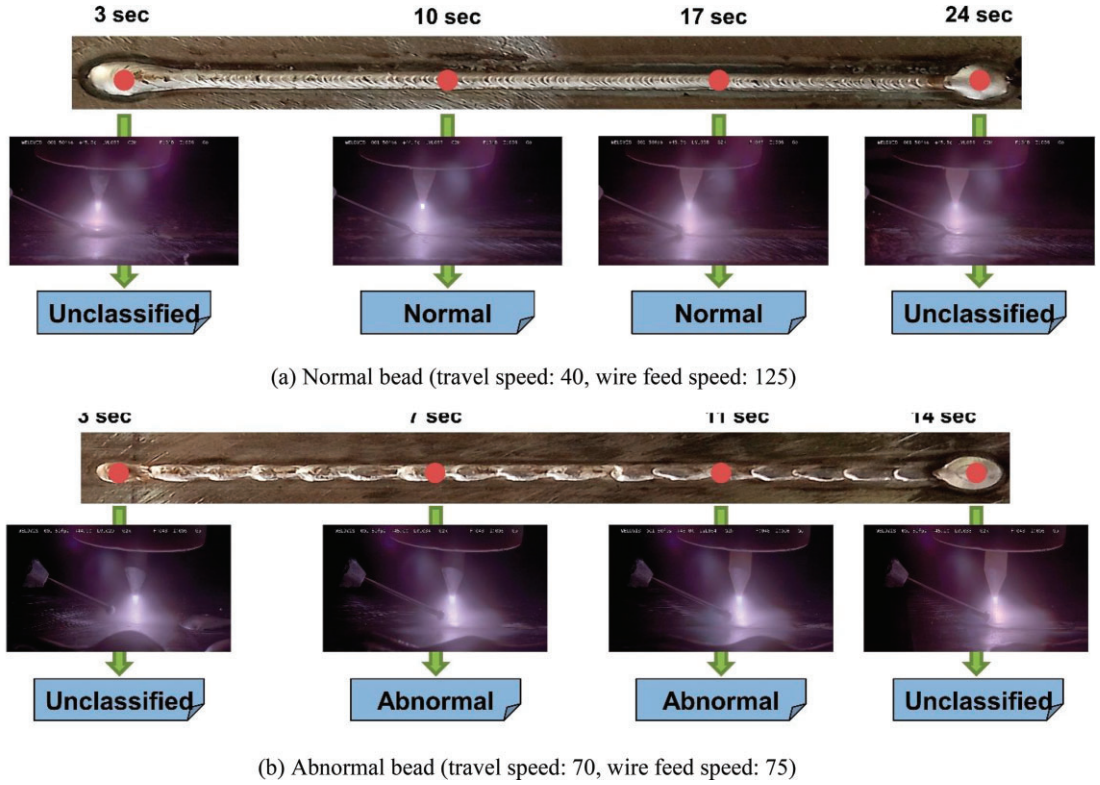


Figure 10: Examples of bead image labeling.

both generic and specific feature learning (Tajbakhsh *et al.*, 2016). The convolution layer comprises convolutional kernels, which split each image into small slices to extract receptive fields. The kernels extract spatial features from image data to provide translation equivariance (Oh *et al.*, 2021). The respective fields tend to determine global cues as they increase along the feature hierarchy (Zheng *et al.*, 2016). The pooling layer reduces input dimensionality and provides spatial invariance to the DenseNet structure. The dense block layer concatenates the features of the preceding layers and differentiates between the original information and any new information added to the network (Khan *et al.*, 2020). The feature maps of all preceding layers become inputs to the subsequent layer. The transition layer comprises a normalization layer, a rectified linear unit (ReLU), a 1 × 1 convolutional layer, and a 2 × 2 average pooling layer to perform downsampling to change the sizes of feature maps. The global average pooling (GAP) layer, located at the end of the structure, decreases the number of parameters to generate a global representation of feature maps (Khan *et al.*, 2020).

4.3. Frozen ratio selection

Composite learning enables the optimization of a frozen ratio to produce the best classification accuracy in transfer learning. No learning is performed in the frozen layers, where pre-trained weights are transferred directly from the preceding stage; meanwhile, learning is performed in the unfrozen layers, where image features are updated by learning new data (Wang & Zheng, 2020).

The setting of frozen ratios needs to be decided in a heuristic manner to determine the optimal ratio among the given scenarios. Figure 14 illustrates the process, and Table 3 lists the frozen ratio settings in the first two stages. These ratios can be assigned differently for each stage. For example, in Case 3, the frozen layers encompass convolution and pooling (CP) to GAP in the first stage,

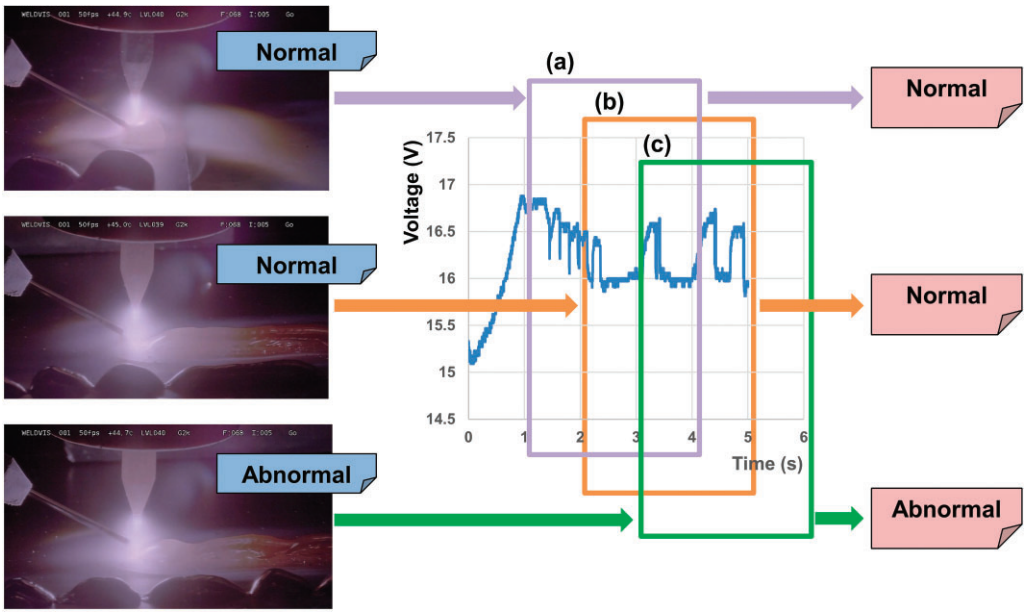


Figure 11: Examples of voltage image labeling.

Table 2: The number of training and testing data samples.

	Training						Testing			
	Training			Validation						
Material	Normal	Abnormal	Total	Normal	Abnormal	Total	Normal	Abnormal	Total	Total
LCS	1224	1385	2609	566	553	1119	2115	1420	3535	7263
STS	1093	844	1937	475	356	831	1328	1332	2660	5428
INC	1128	1379	2507	522	553	1075	1476	1644	3120	6702
Total	3445	3608	7053	1563	1462	3025	4919	4396	9315	19 393

whereas they correspond to CP and the third dense block in the second stage.

Figure 15 shows the learning procedure for frozen and unfrozen layers. In the prior stage, features are created in the frozen layers and transferred to the present stage. In the present stage, the layers that correspond to those frozen in the prior stage do not learn because of feature transfer. The following layers, which associate with the unfrozen layers at the prior stage, combine the transferred features with the initial set of features of the present stage. These layers learn the voltage image data acquired from the present material and derive the output. The loss value of the output is calculated using a loss function through forward propagation. Features are updated through backpropagation as the number of epochs increases until the loss value is minimized or satisfied (Yamashita *et al.*, 2018).

4.4. Classification

Image features are generated by learning voltage image data in the source domain. These features are then shared to derive a transferred model for detecting anomalies in the target domain. The transferred model learns a feature space shared by the source and target domains (Yang *et al.*, 2020). This feature space is transmitted to the fully-connected layer, which acts as a decision-maker for classification in the target domain.

Figure 16 illustrates the classification procedure in the target domain. Feature extraction is conducted on the frozen layers by adopting the image features transferred from the source domain. Meanwhile, learning is performed in the unfrozen layers with new data from the target material. The GAP layer delivers 1664 output nodes derived from feature extraction. The fully-connected layer connects with the output of the preceding layers involved in feature extraction. Then it classifies each voltage image probabilistically by establishing a nonlinear combination of selected features (Khan *et al.*, 2020). This layer uses typical classifiers in machine learning, such as softmax, support vector machines, and artificial neural networks. In this study, softmax was used as the classifier, as expressed in Equation (4),

$$softmax\left(x_i\right) = \frac{\exp\left(x_i\right)}{\sum_{j=1}^{n} \exp\left(x_j\right)}. \quad (4)$$

The fully-connected layer essentially accompanies fine-tuning, which is a process of training the target domain data to update a set of pre-trained weights to adapt to the target domain (Tajbakhsh *et al.*, 2016). Fine-tuning enables the implementation of transfer learning, wherein a small amount of data can be used for feature extraction and classification in the target domain, even though a sufficient amount of data is required for feature extraction in the source domain. The CNN-based model can output a normal or abnormal classifier from the fully-connected layer. As expressed in Equation (4), each classifier is determined probabilistically.

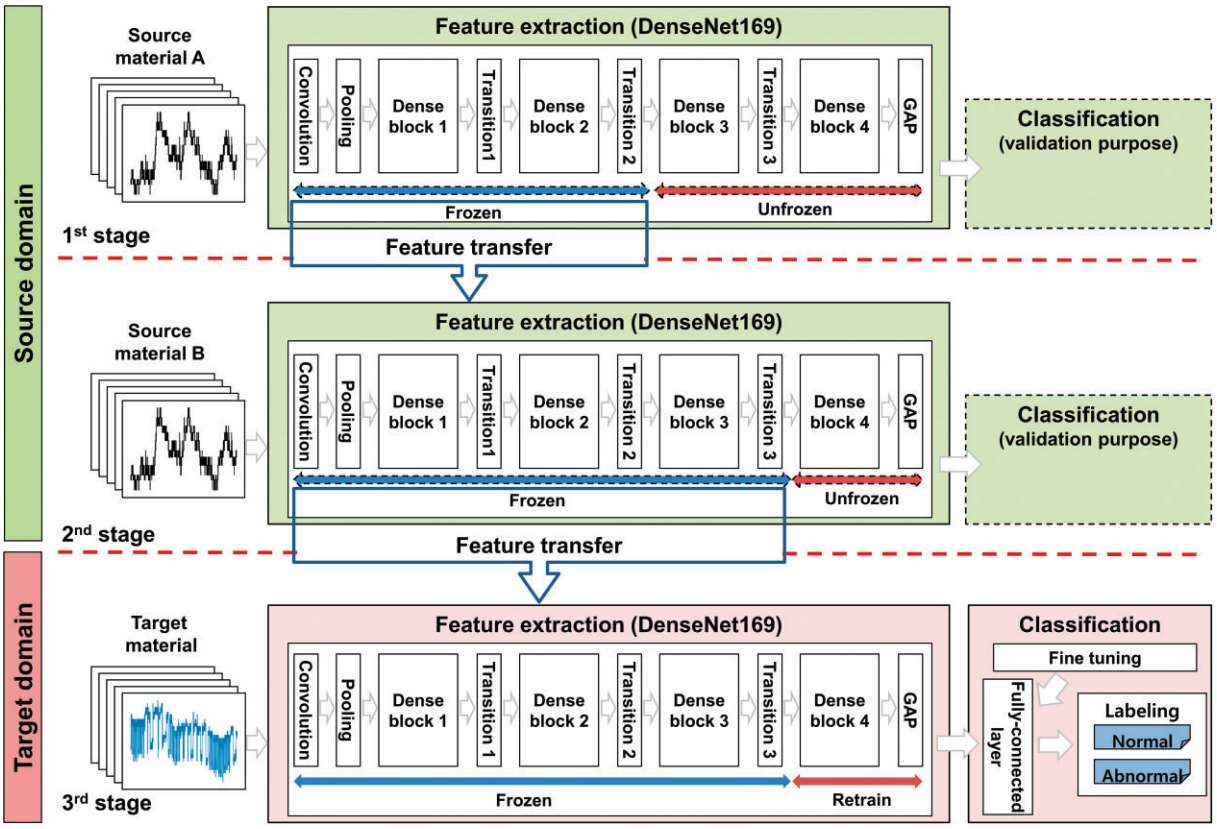


Figure 12: The architecture of the proposed method.

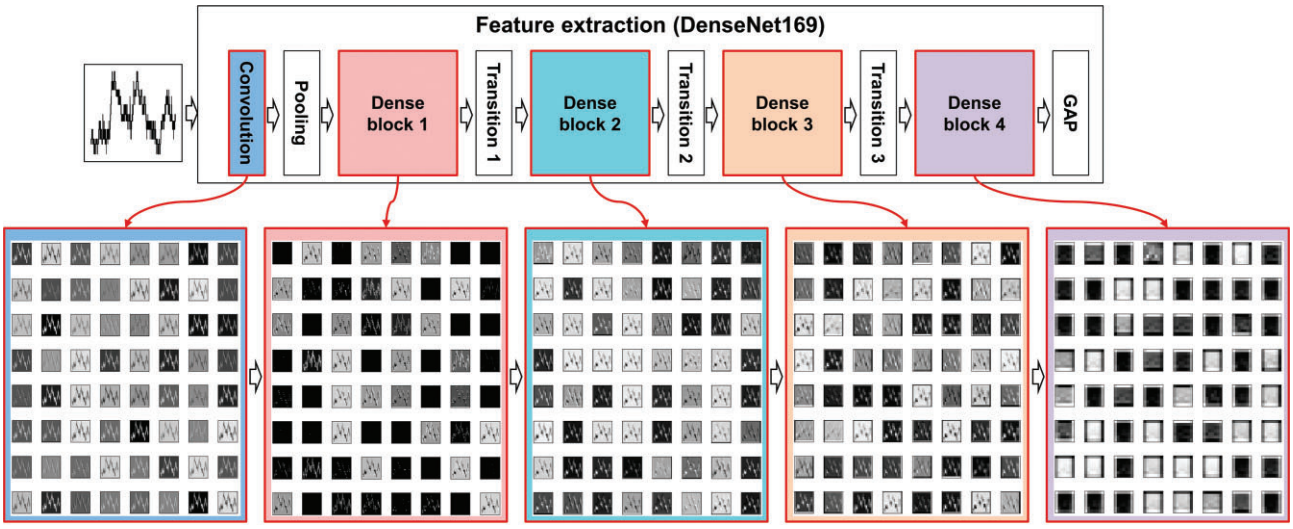


Figure 13: Feature extraction in DenseNet169.

5. Experimental Results

The classification accuracy of the proposed method was validated using the testing dataset. First, the setting of frozen ratios was evaluated to determine the frozen ratio that results in the highest accuracy (see Section 4.3). Next, the classification accuracy was evaluated and compared to the existing methods. The modeling was conducted using Google Colaboratory (Colab) Pro, a cloud-based service for computer vision-based machine learning and deep learning (Gujjar *et al.*, 2021). The Colab Pro features the following specifications: Intel Xeon 2.30GHz CPU, Nvidia P100 GPU, 25.5GB RAM, and Python programming language (Google Colaboratory, 2022; https://colab.research.google.com/). Equation 5 mathematically expresses classification accuracy.

$$classification\ accuracy\ (\%) = 100 \times \frac{the\ number\ of\ data\ samples\ correctly\ classified}{the\ number\ of\ total\ data\ samples} \quad (5)$$

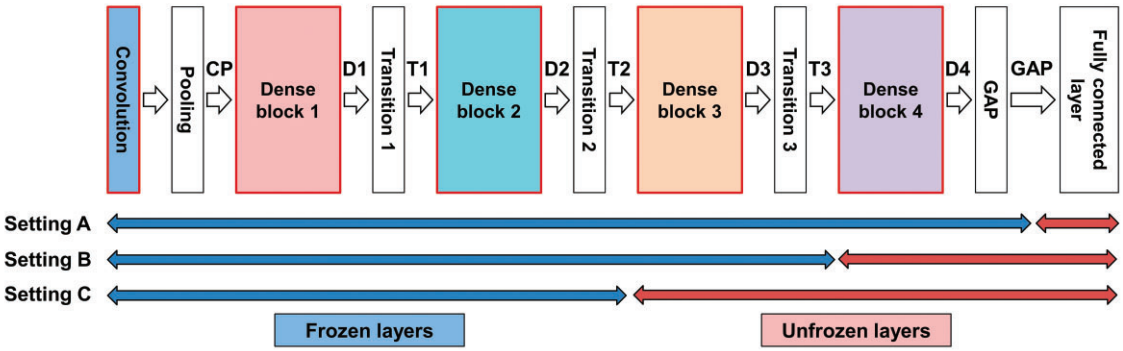


Figure 14: The setting of frozen ratios in transfer learning.

Table 3: Frozen ratio settings.

	First stage		Second stage	
Case No.	Setting	Frozen layer	Setting	Frozen layer
Case 1	A	CP, D1, T1, D2, T2, D3, T3, D4, GAP	A	CP, D1, T1, D2, T2, D3, T3, D4, GAP
Case 2	A	CP, D1, T1, D2, T2, D3, T3, D4, GAP	B	CP, D1, T1, D2, T2, D3, T3
Case 3	A	CP, D1, T1, D2, T2, D3, T3, D4, GAP	C	CP, D1, T1, D2, T2
Case 4	B	CP, D1, T1, D2, T2, D3, T3	A	CP, D1, T1, D2, T2, D3, T3, D4, GAP
Case 5	B	CP, D1, T1, D2, T2, D3, T3	B	CP, D1, T1, D2, T2, D3, T3
Case 6	B	CP, D1, T1, D2, T2, D3, T3	C	CP, D1, T1, D2, T2
Case 7	C	CP, D1, T1, D2, T2	A	CP, D1, T1, D2, T2, D3, T3, D4, GAP
Case 8	C	CP, D1, T1, D2, T2	B	CP, D1, T1, D2, T2, D3, T3
Case 9	C	CP, D1, T1, D2, T2	C	CP, D1, T1, D2, T2

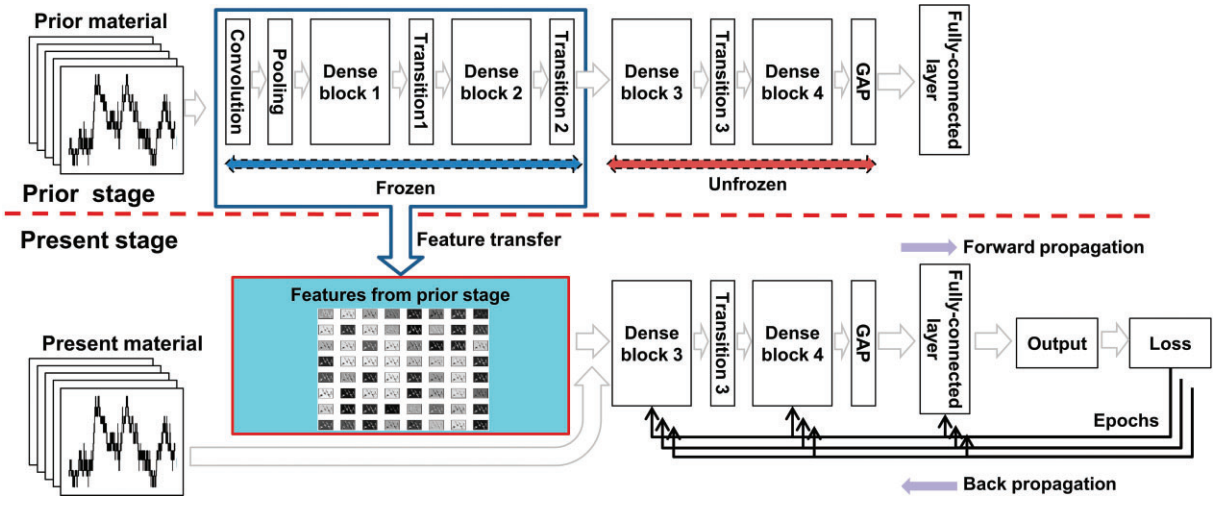


Figure 15: Learning procedure at frozen and unfrozen layers.

5.1. Determination of the frozen ratio

Classification accuracy was measured in the nine frozen ratio setting cases listed in Table 3. The results are presented in Table 4. In the first trial, LCS and STS were used as the first and second source materials, respectively, while INC was used as the target material. The percentage in brackets indicates the portion of data samples in each material dataset used for training. In the second trial, LCS, INC, and STS were used as the first and second source materials and the target material, respectively. The hyperparameters include the Adam optimizer, categorical cross-entropy loss function, 100 epochs, and 32 batch sizes. The epochs and batch sizes were determined heuristically to avoid overfitting and excessive learning time.

Because Case 3 achieved the highest classification accuracy among the nine cases, it was selected as the frozen ratio, which configured Setting A in the first stage and Setting C in the second stage. Hereafter, these settings are used as the default frozen ratios.

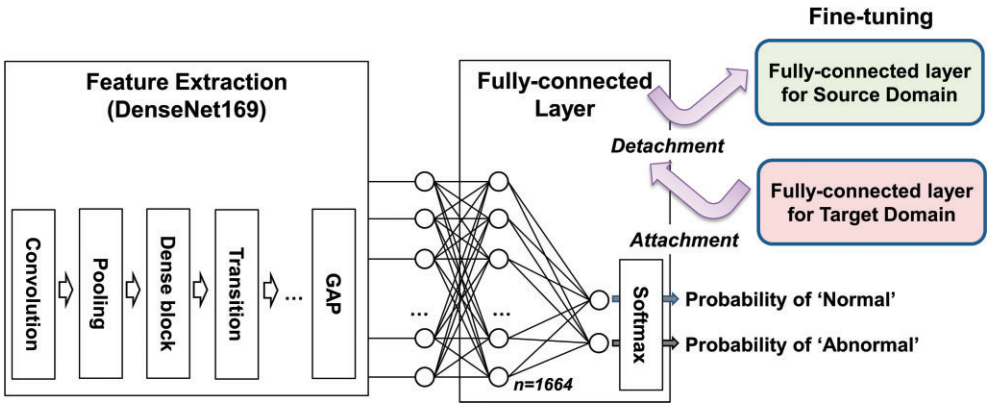


Figure 16: Classification procedure.

Table 4: Classification accuracy (%) in frozen ratio setting.

Trial	1 First source material: LCS (100%) Second source material: STS (30%) Target material: INC (30%)	2 First source material: LCS (100%) Second source material: INC (30%) Target material: STS (30%)	Average
Case 1	87.0	86.6	86.8
Case 2	87.3	88.7	88.0
Case 3	89.7	91.5	90.6
Case 4	85.2	87.6	86.4
Case 5	88.0	88.0	88.0
Case 6	88.8	90.1	89.5
Case 7	87.0	83.8	85.4
Case 8	87.8	91.0	89.4
Case 9	87.3	89.7	88.5

5.2. Evaluation of classification accuracy

5.2.1. *Comparison with existing methods*

Classification accuracy was measured and evaluated for CNN-based models derived using the proposed method and several existing methods. Here, LCS and STS were used as the first and second source materials, and INC was used as the target material. This material setting is reasonable because collecting data from LCS and STS is less expensive than from INC. The hyperparameters include the Adam optimizer, categorical cross-entropy loss function, 30 epochs, and 16 batch sizes. The baseline models are defined as follows:

- The standard machine learning (S-ML) model was trained and tested on the target material and provided classification accuracy of machine learning in the target material. It was therefore useful for a performance comparison between machine learning and transfer learning.
- The single-source transfer learning (S-TL) model was trained solely on LCS and transferred to the target material. This model exemplifies traditional single-source transfer learning.
- The multisource transfer learning (M-TL) model uses the proposed method.

Figure 17 presents the classification accuracy results of the three models with respect to the increase in the number of data samples in the second source material. For example, Fig. 17a presents the classification accuracy when 5% of data samples were trained in the second source material. The x-axis indicates the portion of data samples used for training in the target material. Here, all data samples of the first source material were used for training. Figure 18 summarizes the confusion matrix of the three models when 5% of data samples were used in both the second source and the target materials (the leftmost three points in Fig. 17a). Note that the numbers of total data samples were all 2921 in the three confusion matrix because 199 samples were excluded from the total samples of 3120 during the testing data preprocessing. The number of actual normal and abnormal samples was 1278 and 1643, respectively. The following analytical results were obtained:

- The S-ML models exhibited an S-curve pattern. The classification accuracy was lower when fewer data samples were used. In Fig. 18a, the S-ML model did not show correct classification, as it correctly classified abnormal samples as abnormal but incorrectly classified normal images as abnormal. Meanwhile, it exceeds 90% since 20% of the data samples were used for training. Such S-curve pattern is typical in machine learning, where classification accuracy substantially improves with an increase in the number of data samples. The S-curve pattern implies that the number of data samples was enough to extract features properly from training data as the models achieved the maximum performance and showed no more accuracy increase since 20% of data samples.
- The S-TL models exhibited an average classification accuracy of 84.1% (minimum 83.3%, maximum 85.2%). This implies that single-source transfer learning works desirably, even

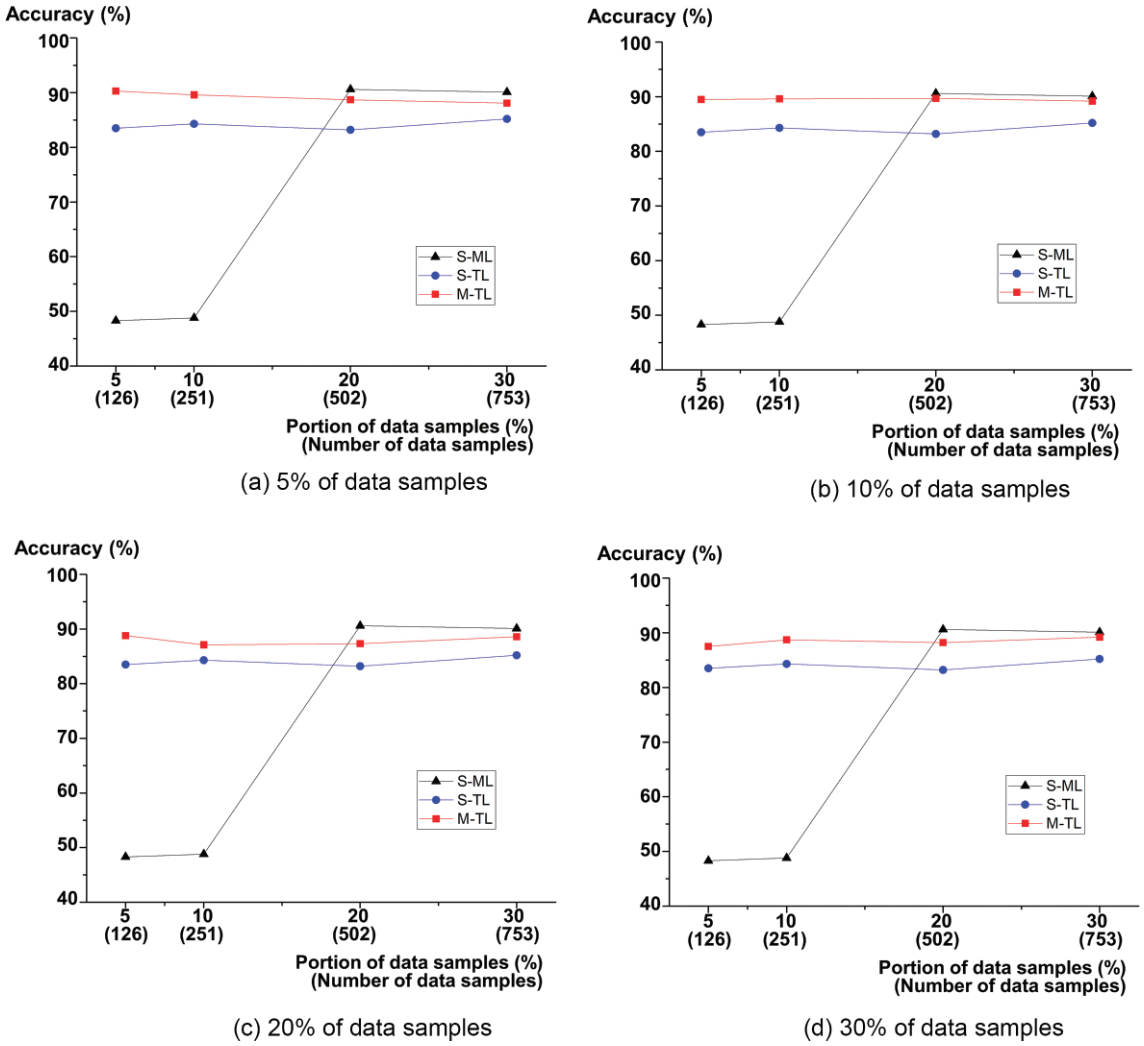


Figure 17: Classification accuracy comparison between methods in INC target material.

S-ML		Predicted value	
		Normal	Abnormal
Actual value	Normal		1278
	Abnormal		1643

(b) S-ML model

S-TL		Predicted value	
		Normal	Abnormal
Actual value	Normal	856	422
	Abnormal	15	1628

(b) S-TL model

M-TL		Predicted value	
		Normal	Abnormal
Actual value	Normal	1128	150
	Abnormal	175	1468

(c) M-TL model

Figure 18: Confusion matrix for 5% of data samples in the source and target materials.

when a small number of data samples is used for training. In Fig. 18b, the S-TL model misclassified actual normal samples as abnormal (422) more than abnormal samples as normal (15).

- The M-TL models scored an average classification accuracy of 88.8% (minimum 87.1%, maximum 90.3%) regardless of the portion of data samples in the target material. In Fig. 18c, the M-TL model exhibited unbiased and small misclassification, while it misclassified normal samples as abnormal (150) and abnormal samples as normal (175). These models show a stable pattern similar to that of the S-TL models, even when a few data samples of the second source material are used for training. Significantly, the M-TL models achieved a higher classification accuracy than the S-TL model. These results demonstrate that the proposed method outperforms traditional single-transfer learning.

5.2.2. *Comparison with material-concatenated transfer learning*

Our research team has developed a single-source and material-concatenated transfer learning method. A similar existing method was developed to detect balling defects in a target material by transferring features extracted from a source material

(Lee *et al.*, 2021). Figure 19 illustrates the model architecture. Although the prior method exemplifies single-source transfer learning, it differs from typical single-source transfer learning since it employs material property features to reflect the characteristics of melting and solidification, including thermal properties, such as thermal conductivity, melting point, and specific heat capacity. These features were concatenated with the image features extracted from single-source transfer learning.

The method above generated the single-source and material-concatenated transfer learning model (Sm-TL). This model extracts image features from the source material, transfers them, and concatenates them with the material property features of the target material. We used LCS as the the source material and INC as the target material.

Our proposed method (M-TL) functions as an extension of Sm-TL that stabilizes and improves model performance. Classification accuracy was measured in all generated models using the same data and experimental setup. The feature extractor is DenseNet169, and the hyperparameters include the Adam optimizer, categorical cross-entropy loss function, 30 epochs, and 16 batch sizes. Figure 20 presents the classification accuracies obtained by the S-TL, Sm-TL, and M-TL models. The blue and red curves represent classification accuracy when 5% and 30% of the STS data samples were used for training, respectively, and the black line denotes the accuracy of Sm-TL models.

- The Sm-TL models achieved an overall classification accuracy of 81.9–89.4%, which exceeds that obtained by traditional S-TL models. This can result from accounting for the melting and solidification mechanisms by concatenating material property features to the image features.
- The Sm-TL models obtained a low classification accuracy of 81.9% when 1% of STS data samples were used. The accuracy increased when the percentage of STS samples exceeded 3%. In contrast, the M-TL models achieved classification accuracies of 88.4% and 88.2% at 30% and 5% of data samples in STS, with 1% of data samples in INC, respectively. These results indicate that models generated by the proposed method can efficiently extract image features even with a few data samples trained in the target material, thereby ensuring stable classification accuracy.

5.3. Additional analysis of the classification accuracy

The classification accuracy was additionally measured and analysed from diverse perspectives to evaluate the proposed method's applicability. The following conditions were analysed: (1) increase of data samples in the second source material, (2) application of various DenseNet structures, and (3) application to a low-cost target material. Note that all data samples in the first source material were used for extracting the image features.

(1) Increase of data samples in the second source material.

The classification accuracy was examined with regard to an increasing amount of data samples trained in the second source material. The following analysis evaluated the effect of the number of data samples in the second source material (STS) on the classification accuracy. Table 5 presents the classification accuracy results for M-TL models trained using 5%, 10%, 30%, and 100% of the STS data samples. The four models did not exhibit significant differences. This implies that the proposed method produces a desirable classification accuracy regardless of the number of training data samples in the second source material.

(2) Application of various DenseNet structures

The classification accuracy was also investigated with respect to the use of DenseNet structures, including DenseNet121, DenseNet169, and DenseNet201. All three are known to be stable structures. Although they are similar, they employ different numbers of parameters in the D3 and D4 layers (Wang & Zheng, 2020). The hyperparameters include the Adam optimizer, categorical cross-entropy loss function, 30 epochs, and 32 batch sizes. The batch sizes were increased from 16 to 32 to avoid overfitting.

Table 6 presents the classification accuracy results for all three DenseNet structures at the frozen ratio settings listed in Table 3. Because the three structures produced similar accuracy values, it is difficult to assert the superiority of DensetNet121 and DenseNet201 to DenseNet169.

(3) Application to a low-cost target material

The classification accuracy was evaluated in Section 5.2 when LCS and STS were used as the first and second source materials, and INC was used as the target material. Here, another classification accuracy was measured to examine whether the proposed method could be similarly applied to a different and low-cost target material. LCS and INC were used as the first and second source materials, respectively, and STS was used as the target material. The setting of hyperparameters was identical to those in Section 5.2.1.

Figure 21 presents the classification accuracy results. Similar to the results in Fig. 17, the M-TL models gained higher classification accuracies than the S-TL models when the portions of data samples were 5% and 10% in the four cases. The similar results at 5% and 10% data portions in the STS show that the proposed method can deliver a good classification accuracy in even small data samples at another target material. However, it was hard to distinguish a superiority of M-TL or S-TL model performance in 20% and 30% of data samples in the STS, different from the results in Fig. 17. The difference in results at 20% and 30% data portions comes from the fact that the settings of frozen ratios and hyperparameters were not optimized to the STS but to the INC target material. This means that frozen ratios and hyperparameters in modeling need to be tuned and customized to a target material to produce a high and stable classification accuracy.

6. Discussion

(1) Applicability of multisource transfer learning

The M-TL models achieved an average classification accuracy of 88.8% across all cases presented in Fig. 17, outperforming the standard S-TL models by 4.7%. Furthermore, the M-TL models exhibited stable classification accuracy even with 10% (251) of the data samples trained in the target material.

These results demonstrate that the features learned from the source domain have high transferability and are transferrable to the target domain for generalizability. Transferability is referred to as the generalizability (Yang *et al.*, 2020). The voltage images that comprise our model's input data are represented by a uniformly-colored line profile plotted on the time and voltage axes. They are simple and homogeneous in visual representa-

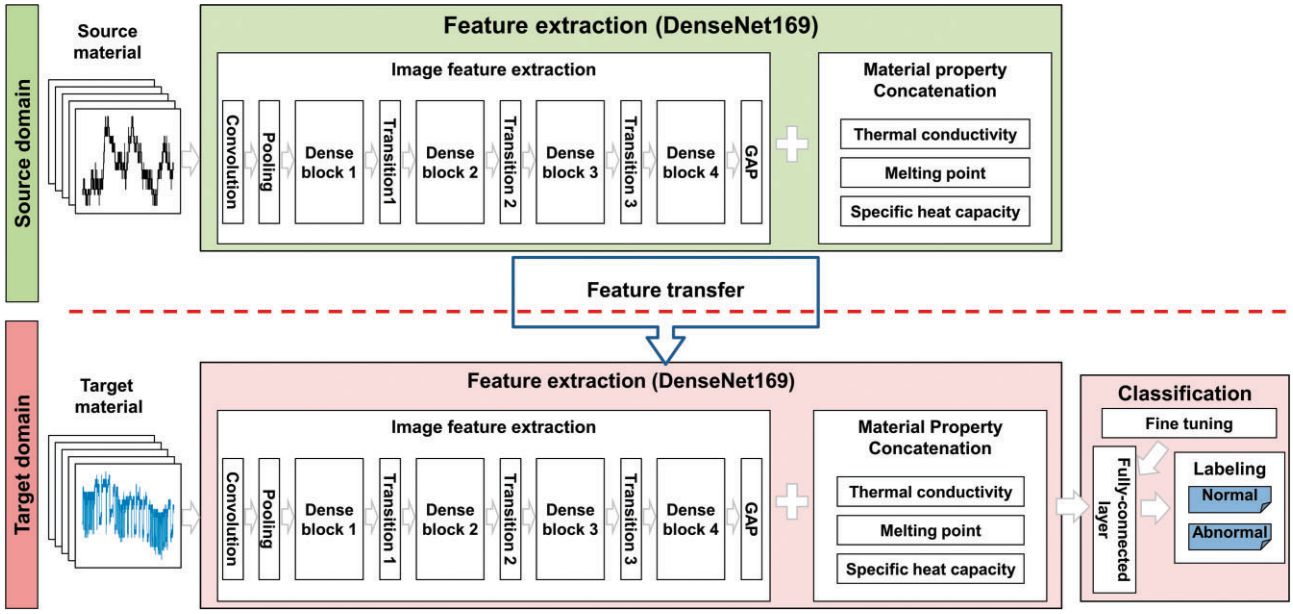


Figure 19: The model architecture of the prior method.

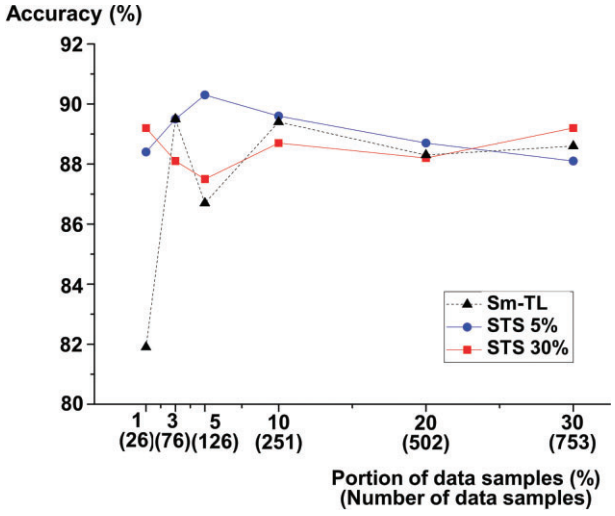


Figure 20: Classification accuracy results between M-TL and Sm-TL.

Table 5: The classification accuracy with regard to STS data samples.

Portion (number) of data samples in second source material	Portion (number) of data samples in the target material				
	5% (126)	10% (201)	20% (326)	30% (502)	Average
5% (97)	90.3	89.6	87.7	88.1	88.9
10% (194)	89.5	89.6	89.7	89.2	89.5
30% (582)	87.5	88.7	88.2	89.2	88.4
100% (1937)	87.3	88.9	88.0	86.6	87.7
Average	88.7	89.2	88.4	88.3	88.6

tion compared with other complex visual data, such as light detection and ranging images in autonomous cars and computed tomography images in medicine. This endows learned features with transferability by reflecting the generic phenomena of melting and solidification. Generalizability facilitates domain adaptation by reducing the data distributional discrepancy between the source and target domains based on a common feature space. In particular, we conjecture that the underlying features from the two source materials provide rigid transferability, as they are not only sequentially extracted using machine learning but also complementarily concatenated with a reasonable set of frozen ratios.

Empirically, we demonstrated the proposed method's practicality by using it to generate CNN-based anomaly detection models for a certain target material with different source materials. Similar results were attained by (Pandiyan *et al.*, 2022), wherein CNN-based models were derived to detect defects in bronze by model transfer from stainless steel in laser powder bed fusion. We therefore believe that transfer learning can be feasible and applicable for implementing material-to-material knowledge transfer for anomaly detection in WAAM.

(2) Resolving data scarcity in machine learning

Machine learning cannot exhibit high classification accuracy in a data-scarce environment that frequently occurs in WAAM, as described in Section 2.2. This should be resolved to deploy machine learning to anomaly detection in WAAM and can be overcome by transfer learning and oversampling. Transfer learning, as the main approach of this study, is a machine learning approach to transfer the features extracted from source materials to a target material. Meanwhile, oversampling can also be considered a data preprocessing technique to resolve data scarcity, although it is used mainly to mitigate class imbalances (Kim *et al.*, 2023). Oversampling increases the number of data samples by augmenting the data samples modified from the original ones. Table 7 presents the pros and cons of transfer learning and oversampling. Transfer learning can deliver fine classification accuracy even using small-sized data; however, it pre-supposes the existence of data obtained from source materials. Oversampling can be applied easily using known techniques; however, it may not be feasible if features are not extracted correctly from augmented data samples

Table 6: The classification accuracy (%) in DenseNet structures.

Structure	Case 1	Case 2	Case 3	Case 4	Case 5	Case 6	Case 7	Case 8	Case 9	Average
DenseNet121	90.1	91.5	92.1	88.8	89.2	90.8	89.5	92.5	92.1	90.7
DenseNet169	87.0	87.3	89.7	85.2	88.0	88.8	87.0	87.8	87.3	87.6
DenseNet201	92.0	92.0	91.0	89.7	88.6	91.4	88.1	91.2	89.3	90.4

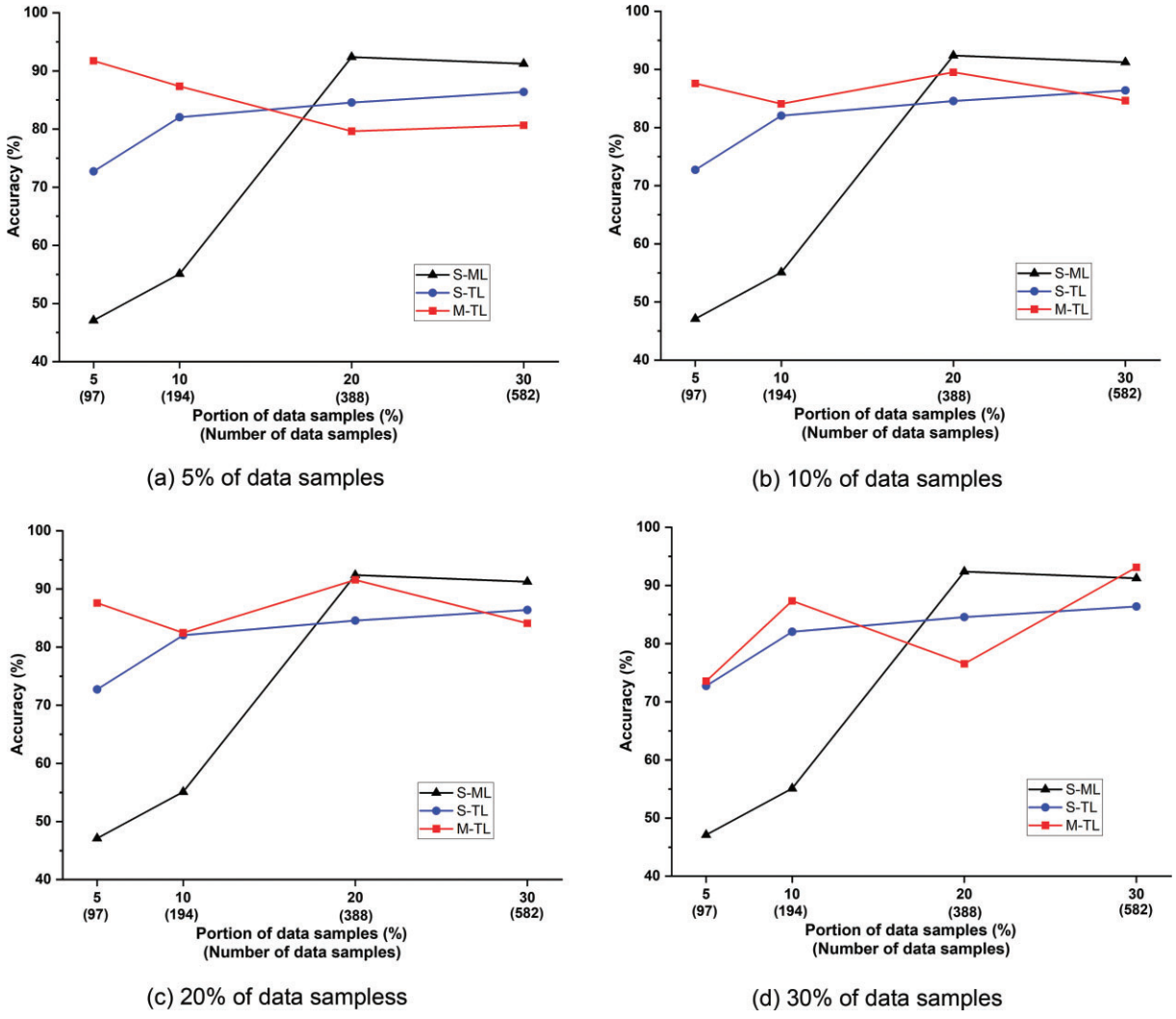


Figure 21: Classification accuracy comparison between methods in STS target material.

due to their irrelevant representation. Hence, choosing a solution to overcome data scarcity is necessary, considering data characteristics, model performances, and application environments.

(3) Usage scenarios of the proposed method

The proposed method can be applied to detect balling defects for quality monitoring in WAAM when data are scarce or unavailable. This application can be practical, particularly when no information about anomaly detection exists for depositing a new material.

Expectedly, the proposed method can be applied to real-time control and process parameter optimization for quality assurance, although they are out-of-scope of this study. The proposed method can be integrated with real-time control for the in situ conversion to normality from abnormality. Real-time control enables a machine to calibrate process parameters adaptively to reduce the difference between intended and unintended quality performance. Furthermore, the proposed method can be considered a modeling technique for optimizing process parameters in terms of quality in WAAM. The process parameter optimization aims to eliminate or minimize the appearance of abnormality during the bead deposition by determining the best process parameters proactively in the process planning phase. The proposed method can contribute to generating anomaly detection models flexibly without the restriction of data scarcity.

7. Conclusions

The following conclusions have been drawn from this study:

Table 7: Pros and cons of transfer learning and oversampling.

Type	Pros	Cons
Transfer learning	- Be useful when data collection is expensive - Make good model performance using small-sized data - Allow to set various materials as a target - Be applicable when oversampling is not feasible	- Should have data obtained from source materials - Cannot outperform machine learnig models in data sufficiency - Should design model architecture
Oversampling	- Make high model performance - Avoid information loss in original data - Be applicable easily using techniques (e.g., data augmentation, Synthetic Minority Over-sampling Technique)	- Should have data obtained from a target material - Should analyse original data to check availability of oversampling - Induce overfitting problems

- A multisource transfer learning method was developed to detect balling defects for quality monitoring in WAAM. This method generates anomaly detection models for a target material by transferring the knowledge learned from voltage image data to two different source materials. Specifically, the proposed method converts numeric-type voltage data to image-type voltage data for the CNN input, extracts image features from source materials using the CNN structure with a variable frozen ratio setting, transfers the image features to the target material, applies the CNN again to extract image features from the target material, and finally fine-tunes and creates the transfer learning model that classifies each voltage image with a normal or abnormal label.
- Experiments were performed in a GTAW-based WAAM system using the LCS, STS, and INC materials. In the experiments, the DenseNet169-based M-TL models achieved a minimum classification accuracy of 87.1%, and a maximum of 90.3%, when 20%–10% and 5%–5% of data samples were used for training in the second source material (STS) and target material (INC). Furthermore, the M-TL models produced good and stable classification accuracies of 88.4%, 89.5%, and 90.3% when 1%, 3%, and 5% of the data samples, respectively, were used for training in the target material under 5% of the data samples in the STS source material. These values are slightly higher than the corresponding 81.9%, 89.5%, and 86.7% accuracies obtained by the Sm-TL models, which were derived from single-source and material-concatenated transfer learning.

We expect the proposed method to provide cost efficiency in developing anomaly detection models in WAAM based on multisource transfer learning, which has rarely been attempted. Because data collection in this field is generally expensive and time-consuming, this approach would be even more practical when high-cost materials (e.g., tungsten and molybdenum) must be deposited. The difficulty of data collection and the unavailability of machine learning are critical problems. The proposed method therefore functions as a cost-efficient alternative for machine learning-based anomaly detection, as it transfers information from sufficient data in low-cost materials to insufficient data in high-cost materials.

However, as training data increased, the proposed method did not achieve higher classification accuracy than machine learning. This was expected due to the nature of transfer learning, as underlying image features cannot match exactly with those of the target domain. In addition, only three materials were tested because of the limited experimental environment. This study demonstrated the feasibility of the proposed method by applying it to the two target materials, including INC and STS. The number of materials was relatively small in the experiments; thus, further investigations using more materials are required to verify the practicality of the proposed method. Moreover, measuring the learning time was out of the scope of this study. Minimizing learning time is critical to ensure in-situ quality monitoring in WAAM.

We plan to implement a hybrid anomaly detection algorithm that applies both machine learning and transfer learning to be adaptively selected, depending on the amount of data. Such a system is expected to support expert decisions. In addition, we aim to establish in-situ quality monitoring in WAAM using this approach. In addition, we plan to develop a learning-based process planning approach that predicts and optimizes performance quality. This objective will necessitate a thorough investigation of the relationships between process parameters and quality performance in WAAM.

Acknowledgments

This research was supported by the MSIT (Ministry of Science, ICT), Korea, under the High Potential Individuals Global Training Program (No. 2021–0-01566) supervised by the IITP (Institute for Information & Communications Technology Planning & Evaluation). The authors of this paper appreciate the continuous support provided by the Center for Manufacturing Research (CMR) and the Department of Manufacturing and Engineering Technology at Tennessee Technological University. This material is based upon work supported by the National Science Foundation under Grant No. 2015693.

Conflict of interest statement

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Bikas H., Stavropoulos P., & Chryssolouris G. (2016). Additive manufacturing methods and modelling approaches: A critical review. *International Journal of Advanced Manufacturing Technology*. **83**, 389–405. https://doi.org/10.1007/s00170-015-7576-2.

Cheng L., Tsung F., & Wang A. (2017). A statistical transfer learning perspective for modeling shape deviations in additive manufacturing. *IEEE Robotics and Automation Letters*, **2**, 1988–1993. https://doi.org/10.1109/LRA.2017.2713238.

Cheng L., Wang K., & Tsung F. (2021). A hybrid transfer learning framework for in-plane freeform shape accuracy control in additive manufacturing. *IISE Transactions*, **53**, 298–312. https://doi.org/10.1080/24725854.2020.1741741.

Cho H. W., Shin S. J., Seo G.-J., Kim D. B., & Lee D. H. (2022). Real-time anomaly detection using convolutional neural network in wire

arc additive manufacturing: Molybdenum material. *Journal of Materials Processing Technology*, **302**, 117495. https://doi.org/10.1016/j.jmatprotec.2022.117495.

Colosimo B. M., Huang Q., Dasgupta T., & Tsung F. (2018). Opportunities and challenges of quality engineering for additive manufacturing. *Journal of Quality Technology*, **50**, 233–252. https://doi.org/10.1080/00224065.2018.1487726.

Derekar K. S. (2018). A review of wire arc additive manufacturing and advances in wire arc additive manufacturing of aluminium. *Materials Science and Technology*, **34**, 895–916. https://doi.org/10.1080/02670836.2018.1455012.

Everton S. K., Hirsch M., Stravroulakis P., Leach R. K., & Clare A. T. (2016). Review of in-situ process monitoring and in-situ metrology for metal additive manufacturing. *Materials & Design*, **95**, 431–445. https://doi.org/10.1016/j.matdes.2016.01.099.

Ferreira R. D. S. B., Sabbaghi A., & Huang Q. (2020). Automated geometric shape deviation modeling for additive manufacturing systems via Bayesian neural networks. *IEEE Transactions on Automation Science and Engineering*, **17**, 584–598. https://doi.org/10.1109/TASE.2019.2936821.

Gobert C., Reutzel E. W., Petrich J., Nassar A. R., & Phoha S. (2018). Application of supervised machine learning for defect detection during metallic powder bed fusion additive manufacturing using high resolution imaging. *Additive Manufacturing*, **21**, 517–528. https://doi.org/10.1016/j.addma.2018.04.005.

Gonzalez-Val C., Pallas A., Panadeiro V., & Rodriguez A. (2020). A convolutional approach to quality monitoring for laser manufacturing. *Journal of Intelligent Manufacturing*, **31**, 789–795. https://doi.org/10.1007/s10845-019-01495-8.

Google Golaboratory. (2022). https://colab.research.google.com/ Accessed 23 February 2022.

Gujjar J. P., Kumar H. P., & Chiplunkar N. N. (2021). Image classification and prediction using transfer learning in colab notebook. *Global Transitions Proceedings*, **2**, 382–385. https://doi.org/10.1016/j.gltp.2021.08.068.

Ho S., Zhang W., Young W., Buchholz M., Al Jufout S., Dajani K., Bian L., & Mozumdar M. (2021). DLAM: Deep learning based real-time porosity prediction for additive manufacturing using thermal images of the melt pool. *IEEE Access*, **9**, 115100–115114. https://doi.org/10.1109/ACCESS.2021.3105362.

Ikram A., & Chung H. (2022). Computational analysis of metal transfer mode, dynamics, and heat transfer under different pulsating frequencies in pulsed wire-arc additive manufacturing. *Journal of Computational Design and Engineering*, **9**, 1045–1063. https://doi.org/10.1093/jcde/qwac043.

ISO/ASTM52900:2021. (2021). *Additive Manufacturing - General principles—Fundamentals and Vocabulary*. ISO/TC 261 and ASTM Committee F42 on Additive Manufacturing Technologies.

Jafari D., Vaneker T. H. J., & Gibson I. (2021). Wire and arc additive manufacturing: Opportunities and challenges to control the quality and accuracy of manufactured parts. *Materials and Design*, **202**, 109471. https://doi.org/10.1016/j.matdes.2021.109471.

Jin B., Tan Y., Nettekoven A., Chen Y., Topcu U., Yue Y., & Sangiovanni-Vincentelli A. (2019). *An encoder-decoder based approach for anomaly detection with application in additive manufacturing. December 16–1, 18th IEEE international conference on machine learning and applications*, 9, Boca Raton, USA, pp. 1008–1015. https://doi.org/10.1109/ICMLA.2019.00171.

Khan A., Sohail A., Zahoora U., & Qureshi A. S. (2020). A survey of the recent architectures of deep convolutional neural networks. *Artificial Intelligence Review*, **53**, 5455–5516. https://doi.org/10.1007/s10462-020-09825-6.

Kim E. S., Lee D. H., Seo G. J., Kim D. B., & Shin S. J. (2023). Development of a CNN-based real-time monitoring algorithm for additively manufactured molybdenum. *Sensors and Actuators A: Physical*, **352**, 114205. https://doi.org/10.1016/j.sna.2023.114205.

Kim Y. M., Shin S. J., & Cho H. W. (2022). Predictive modeling for machining power based on multi-source transfer learning in metal cutting. *International Journal of Precision Engineering and Manufacturing-Green Technology*, **9**, 107–125. https://doi.org/10.1007/s40684-021-00327-6.

Koprinska I., Wu D., & Wang Z. (2018). *Convolutional neural networks for energy time series forecasting. International Joint Conference on Neural Networks (IJCNN)*. Rio de Janeiro, Brazil, July 8–13. https://doi.org/10.1109/IJCNN.2018.8489399.

Lee J. H., Seo G. J., Kim D. B., Lee D. H., & Shin S. J. (2021). *Material-adaptive anomaly detection algorithm using convolutional neural network and transfer learning in wire arc additive manufacturing, KSPE 2021 Autumn Conference*. Busan, Republic of Korea, November 24–26.

Li Y., Polden J., Pan Z., Cui J., Xia C., He F., Mu H., Li H., & Wang L. (2022). A defect detection system for wire arc additive manufacturing using incremental learning. *Journal of Industrial Information Integration*, **27**, 100291. https://doi.org/10.1016/j.jii.2021.100291.

Liu J., Xu Y., Ge Y., Hou Z., & Chen S. (2020). Wire and arc additive manufacturing of metal components: A review of recent research developments. *International Journal of Advanced Manufacturing Technology*, **111**, 149–198. https://doi.org/10.1007/s00170-020-05966-8.

Lu X., Li M. V., & Yang H. (2021). Comparison of wire-arc and powder-laser additive manufacturing for IN718 superalloy: Unified consideration for selecting process parameters based on volumetric energy density. *International Journal of Advanced Manufacturing Technology*, **114**, 1517–1531. https://doi.org/10.1007/s00170-021-06990-y.

Ngo T. D., Kashani A., Imbalzano G., Nguyen K. T. Q., & Hui D. (2018). Additive manufacturing (3D printing): A review of materials, methods, applications and challenges. *Composites Part B: Engineering*, **143**, 172–196, https://doi.org/10.1016/j.compositesb.2018.02.012.

Oh Y., Sharp M., Sprock T., & Kwon S. (2021). Neural network-based build time estimation for additive manufacturing: A performance comparison. *Journal of Computational Design and Engineering*, **8**, 1243–1256. https://doi.org/10.1093/jcde/qwab044.

Pan S. J., & Yang Q. (2009). A survey on transfer learning. *IEEE Transactions on Knowledge and Data Engineering*, **22**, 1345–1359. https://doi.org/10.1109/TKDE.2009.191.

Pandiyan V., Drissi-Daoudi R., Shevchik S., Masinelli G., Le-Quang T., Logé R., & Wasmer K. (2022). Deep transfer learning of additive manufacturing mechanisms across materials in metal-based laser powder bed fusion process. *Journal of Materials Processing Technology*, **303**, 117531. https://doi.org/10.1016/j.jmatprotec.2022.117531.

Park K. B., & Lee J. Y. (2022). Novel industrial surface-defect detection using deep nested convolutional network with attention and guidance modules. *Journal of Computational Design and Engineering*, **9**, 2466–2482. https://doi.org/10.1093/jcde/qwac115.

Qi X., Chen G., Li Y., Cheng X., & Li. C. (2019). Applying neural-network-based machine learning to additive manufacturing: Current applications, challenges, and future perspectives. *Engineering*, **5**, 721–729. https://doi.org/10.1016/j.eng.2019.04.012.

Qin J., Hu F., Liu Y., Witherell P., Wang C. C. L., Rosen D. W., Simpson T. W., Lu Y., & Tang Q. (2022). Research and application of machine learning for additive manufacturing. *Additive Manufacturing*, **52**, 102691. https://doi.org/10.1016/j.addma.2022.102691.

Qiu Y. D., Wu J. M., Chen A. N., Chen P., Yang Y., Liu R. Z., Chen G., Chen S., Shi Y. S., & Li C. H. (2020). Balling phenomenon and cracks in alumina ceramics prepared by direct selective laser melting assisted with pressure treatment. *Ceramics International*, **46**, 13854–13861. https://doi.org/10.1016/j.ceramint.2020.02.178.

Scime L., Siddel D., Baird S., & Paquit V. (2020). Layer-wise anomaly detection and classification for powder bed additive manufacturing processes: A machine-agnostic algorithm for real-time pixel-wise semantic segmentation. *Addittive Manufacturing*, **36**, 101453. https://doi.org/10.1016/j.addma.2020.101453.

Shorten C., & Khoshgoftaar T. M. (2019). A survey on image data augmentation for deep learning. *Journal of Big Data*, **6**, 1–48. https://doi.org/10.1186/s40537-019-0197-0.

Somasundaram A., & Reddy U. S. (2016). *Data imbalance: Effects and solutions for classification of large and highly imbalanced data. Proceedings of the 1st International Conference on Research in Engineering, Computers and Technology (ICRECT)*. Tiruchirappalli, India, September 8–10, pp. 28–34.

Tajbakhsh N., Shin J. Y., Gurudu S. R., Hurst R. T., Kendall C. B., Gotway M. B., & Liang J. (2016). Convolutional neural networks for medical image analysis: Full training or fine tuning?. *IEEE Transactions on Medical Imaging*, **35**, 1299–1312. https://doi.org/10.1109/TMI.2016.2535302.

Vimal K. E. K., Srinivas M. N., & Rajak S. (2021). Wire arc additive manufacturing of aluminum alloys: A review. *Materials Today: Proceedings*, **41**, 1139–1145. https://doi.org/10.1016/j.matpr.2020.09.153.

Wang C., Tan X. P., Tor S. B., & Lim C. S. (2020). Machine learning in additive manufacturing: State-of-the-art and perspectives. *Additive Manufacturing*, **36**, 101538. https://doi.org/10.1016/j.addma.2020.101538.

Wang S. H., & Zhang Y. D. (2020). DenseNet-201-Based deep neural network with composite learning factor and precomputation for multiple sclerosis classification. *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)*, **16**, 1–19. https://doi.org/10.1145/3341095.

Xia C., Pan Z., Polden J., Li H., Xu Y., Chen S., & Zhang Y. (2020). A review on wire arc additive manufacturing: Monitoring, control and a framework of automated system. *Journal of Manufacturing Systems*, **57**, 31–45. https://doi.org/10.1016/j.jmsy.2020.08.008.

Xu F., Dhokia V., Colegrove P., McAndrew A., Williams S., Henstridge A., & Newman S. T. (2018). Realisation of a multi-sensor framework for process monitoring of the wire arc additive manufacturing in producing Ti-6Al-4V parts. *International Journal of Computer Integrated Manufacturing*, **31**, 785–798. https://doi.org/10.1080/0951192X.2018.1466395.

Yamashita R., Nishio M., Do R. K. G., & Togashi K. (2018). Convolutional neural networks: An overview and application in radiology. *Insights Into Imaging*, **9**, 611–629. https://doi.org/10.1007/s13244-018-0639-9.

Yang Q., Zhang Y., Dai W., & Pan S. J. (2020). *Transfer Learning*. Cambridge Univ. Press, Cambridge, UK, p. 228.

Zhao B., Lu H., Chen S., Liu J., & Wu D. (2017). Convolutional neural networks for time series classification. *Journal of Systems Engineering and Electronics*, **28**, 162–169. https://doi.org/10.21629/JSEE.2017.01.18.

Zheng L., Zhao Y., Wang S., Wang J., & Tian Q. (2016). Good Practice in CNN Feature Transfer. *Computer Vision and Pattern Recognition*. https://doi.org/10.48550/arXiv.1604.00133.

Zhu Z., Anwer N., Huang Q., & Mathieu L. (2018). Machine learning in tolerancing for additive manufacturing. *CIRP Annals – Manufacturing Technololgy*, **67**, 157–160. https://doi.org/10.1016/j.cirp.2018.04.119.

Received: March 19, 2023. **Revised:** June 23, 2023. **Accepted:** June 24, 2023