Deep Learning Based Real-Time and In-Situ Monitoring of Weld Penetration: Where we are and what are needed revolutionary solutions?

Rui Yu[1], Yue Cao[1], Heping Chen[2], Qiang Ye[3], YuMing Zhang[1*]

1: Department of Electrical and Computer Engineering and Institute for Sustainable Manufacturing, University of Kentucky, Lexington, KY, USA

2: Ingram School of Engineering, Texas State University, San Marcos, TX 78666, USA

3: Department of Mathematics, University of Kentucky, Lexington, KY, USA

* Corresponding author: yuming.zhang@uky.edu

Abstract: Welding Procedure Designed assures the desired weld penetration be produced under nominal welding conditions. When conditions deviate from the nominal, penetration and other welding outcomes deviate from their desired/targeted ones. To assure the penetration not below minimally necessary, well-designed Procedure should reserve an appropriate margin per estimated condition deviations. An ideal solution is to use relatively small margin but dynamically adjust welding parameters to maintain the penetration at minimally necessary/within the margin. As such, the penetration state should be monitored in real-time during manufacturing. Unfortunately, it occurs underneath workpieces and is not considered directly observable during manufacturing so that its real-time in-situ monitoring is challenging. In the last half century, researchers have focused on finding promising real-time observable phenomena and correlating such phenomena to penetration. This has been difficult as it is unclear what are critical in observed phenomena and trial-and-error iterative processes are practiced proposing features, developing algorithms to calculate them, fitting model from features, and then modifying model, features, or algorithms if fitting accuracy is not acceptable. Such iterative process is not automated, finding/fitting right features/model takes months if not longer, and success is not assured. In particular, algorithms to calculate features vary from features to features and each of them requires extensive tests. Deep learning automates and combines featuring and fitting to maximally use the raw information directly to achieve highest accuracies. Computation is drastically increased but the iterative process is replaced by an automated one so that the time frame is still drastically reduced. This paper reviews various raw information that has been used as the observed phenomena to input into deep learning models and analyzes why they may correlate to penetration; reviews various deep learning models and analyzes why/how they may and are needed to correlate different raw information to penetration; and briefly reviews major techniques that have been used to train deep learning models in penetration prediction. Based on such analyses, this paper identifies major achievements and issues for efforts taken so far to monitor weld penetration using deep learning approaches. Finally, we identify two fundamental issues that require revolutionary solutions in order to move the deep learning technologies from laboratory studies to manufacturing as directions for future efforts. These two issues are analyzed and some preliminary solution directions are proposed.

Content:

1. What is Weld Penetration and Its Specification?

Welding is a method to join two or more pieces of materials together. In the conventional definition, it must involve melting part of the materials being joined [95], in particular melting the solid interfaces that divide them. It is the melting that allows materials from different pieces to move together to bridge their gaps/interfaces, fully or partially. Upon solidification, the liquid bridges become solid and the materials are joined together. After the friction stirring welding (FSW) was invented, this conventional definition

was challenged. The materials at the interfaces are softened and plasticized through friction heat and then forced to move to mix together to bridge the interfaces through pressurization, viscosity, and spatial restriction [96]. The movable plasticized materials are still considered solid and not melted. However, the materials being joined are still mixed together similarly as the materials being joined have been melted so that the FSW is still considered as welding rather brazing [97], soldering [97], riveting [98] etc. where the materials being joined are not melted/moved and the joining is not formed through mixing the materials to be joined. As such, welding is a method to join materials involving mixing of the materials being joined by providing means to mobilize part of them. To ease our discussions, we will stick on the conventional definition for welding based on melting as our focus is fusion welding.

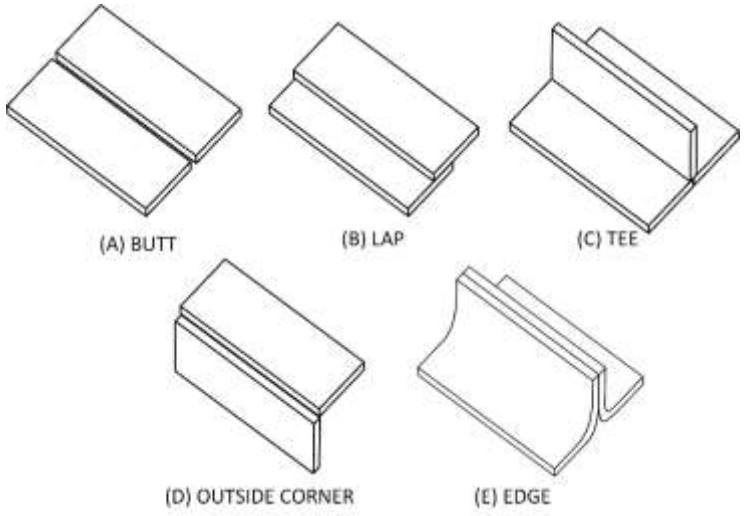


Fig. 1 Five basic joint designs [99].

Fig. 1 shows five basic joint designs. Fig. 2 illustrates their preferred welds. Joint design (A) and (E) shown in Fig. 1 can be welded in ways shown in Fig. 2(a) and (b), while joint design (B)-(D) by that in Fig. 2(c). The solid areas in Fig. 2 are the materials to be joined and dotted lines are the boundaries of welds. The dotted lines are given in three colors: red, green and purple. Red lines are the boundaries of the melting on the materials being joined. Green and purple lines are the boundaries of melted materials with the atmosphere. As such, the red lines define the melting. In particular, in Fig. 2(a) and Fig. 2(b), the materials to be welded, the left and right members, both have a facing edge on their right and left, respectively. Weld is formed through melting the material on the respective facing edges at the same time. In Fig. 2(a), the entire facing edges are melted for the entire thickness t. This is referred to as complete penetration or full penetration. In Fig. 2(b), the facing edges are melt with a depth of d for the upper with $d < t$. This is referred to as incomplete penetration. For a butt joint, a compete penetration is typically required. As melting the facing edge deeply underneath the work-piece faces is in general challenging, the butt joint may be welded from both upper and lower surfaces forming a complete welded joint through making two incomplete penetration welds. In (c), the facing edges are wide and must be filled so that melting filler material must also be a major task while melting the facing edges. In (d) where the weld is completed through multiple passes, the “facing edges” to be melted include the boundary of the solidified material from previous passes. (d) uses a butt joint to illustrate and also applies to fillet joint.

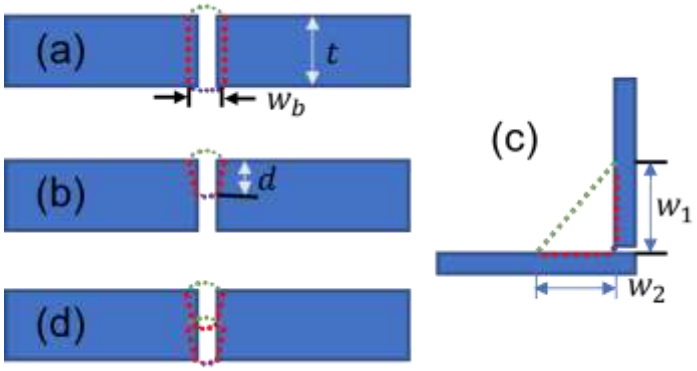


Fig. 2 Welds and requirements. (a) Complete penetration of butt weld; (b) incomplete penetration of butt weld; (c) fillet weld; (d) multiple passes.

In Fig. 2(a), the melting of the facing edges is through the ENTIRE thickness from the upper to the lower surface of the materials as indicated by the red dot-lines and is referred to as complete penetration. In this case, the span from the left edge to the right edge of the melting is referred to as the back-side bead width w_b. As there may be seam tracking error and a gap between the two facing edges, the needed w_b must exceed a minimum. However, a w_b greater than necessary implies excessive heat input, thus a greater distortion and materials property degradation. w_b thus needs to be appropriate per the minimum requirement. The state of the weld penetration for complete penetration can thus be specified by w_b.

In Fig. 2(b), the melting of the facing edges is through from the upper surface for a distance d, rather than the entire thickness t. In this case (incomplete or partial penetration), the weld penetration is specified by d which is termed as the penetration depth. For the joint (E) in Fig. 1, the strength of the welded joint is determined by the penetration depth. To make a complete penetration joint by welding from two sides, the sum of the penetration depths must not be less than t.

For Fig. 2(c), the gap between the facing edge is wide. The weld is formed by bridging the gap using filler. This requires each facing edge be melted in order to fuse with the melted filler material. In this case, the weld sizes w_1 and w_2 are the primary specifications on the resultant weld.

We have assumed that the entire facing edges have been melted, i.e., the entire facing edges in Fig. 2(a) for the complete penetration, the facing edges whose distance to the upper surface is d or less for the incomplete penetration, and the facing edges within the range by w_1 and w_2 in Fig. 2(c) for the fillet weld. If some of the facing edges in such ranges are not melted, it will be a defect referred to as lack of penetration or lack of fusion. Monitoring of weld penetration may thus include (1) detecting if the penetration is complete or incomplete; (2) detecting the penetration depth d or the back-side bead width w_b; and (3) detecting if there is any lack of penetration/fusion. We can refer them together as the **penetration state** to ease the discussion.

We wish to point out that in the cases in Fig. 2(a)-(c), the lack of penetration occurs on the facing edges or the side of the weld. In Fig. 2(d) for multiple pass welding, the surface of the weld from the previous pass must also be melted. The lack of penetration may thus also occur on the bottom of the new weld.

Fig. 2 purposely uses three colors for the dotted lines to depict different kinds of boundaries of the weld pool/welds: red for the liquid-solid interface, due to melting of the solid materials being joined, that has

never been directly visible during and after welding; green for the upper interface of the liquid (melted and mixed materials) with the atmosphere that is directly visible during and after welding; and purple for the lower interface of the liquid with the atmosphere that is difficult to observe during welding but may be seen after welding.

We have verbally defined the penetration state and can introduce four possible state variables: **binary state variable icp** with 1 for incomplete and 0 for complete penetration, **binary state variable lop** with 1 for existence of "lack of penetration" and 0 for none, **quantitative state variable d** for the depth of the penetration that has been defined, and **quantitative state variable w_b** for the back-side bead width that has been defined. The **penetration state** $x = (icp, lop, d, w_b)$, with x stands for the state per the convention, consists of all four variables. For each application, the particular **application penetration state xa** can be defined using an appropriate subset of the four state variable variables.

2. What are the Challenges in Monitoring Weld Penetration?

It is apparent that all four state variables in x can be detected directly from the red (line) interfaces. Variables (icp, d, w_b) can be directly detected from the purple (line) interfaces. However, none of the four penetration state variables can be directly detected from the green (line) interfaces.

The root of the challenges in monitoring the weld penetration is that the red interfaces are not visible as the associated phenomena occur underneath the work-piece and melted materials. Alternative solutions must be sought. Purple interfaces have the direct information to determine all penetration state variables except for lop and may be made directly observable. However, challenges are from meeting other manufacturing goals. For example, one apparent approach to monitor the purple interfaces is to place a sensor such as a camera to view the back-side of the work-piece. This would allow us to determine (icp, w_b)**.** If the gap between the two facing edges is intentionally made relatively large, information for determining d may also become available. However, such solutions require a sensor that accesses the work-piece from the lower surface and cannot be attached to the welding torch/tool that accesses the work-pieces from the upper surface. One can imagine that a second robot be used to carry the sensor and the motion of the robot be synchronized with that which carries the welding torch/tool. This is not implementable for applications where the inner/lower surface is not accessible such as for welding of containers or pipes. For applications that do have the needed access, it may still not be preferred due to the increased complexity and cost. As such, acquiring information from the purple interfaces is not a generally preferred solution to monitor the weld penetration, and a sensor that acquires the raw information from a purple interface has bene referred to as a back-side sensor. As a back-side sensor is not preferred, a possible generally acceptable alternative solution is to acquire the raw information from the green interfaces.

We purposely use green color for the upper gas-liquid interfaces for the relatively good measurability, while purple color for the lower gas-liquid interfaces and red color for the liquid-solid interfaces are for less and poor measurability respectively. As such, we are forced to acquire and use the raw information from the green interfaces as the red interfaces where the raw information is most direct do not have the needed measurability and the measurability for the purple interfaces can only be realized at much increased complexity and cost. We are thus forced to acquire and use the raw information from green interfaces despite the information is the most indirect. A sensor that acquires the raw information from green interfaces, either from the interface self or through the interface to go internally, has been referred to as a top-side sensor or a front-face sensor.

Being the "most indirect" is the basic challenge for top-side sensors that are preferred and generally acceptable for manufacturing. The key is to first find phenomena that are measurable using a top-side sensor while **fundamentally correlating** to the variables in the particular application penetration state xa. Theoretically, the measured phenomena Ξ must be **fully determined** by xa (or the red interfaces), i.e.,

$$\Xi = f(xa) + e \tag{1}$$

with the error $e = 0$ or $|e|$ being acceptably small under the particular application conditions. **However, finding such phenomena are in general uneasy and theoretically proving the existence of the fundamental correlation is in general unrealistic.** We note that most studies in the area only assume such the existence of the correlation and verify using experimental data. The effectiveness depends on the conditions under which the experiments are conducted. Unfortunately, most studies lack a needed reasoning for the assumed existence of such a correlation based on physical analysis and a systematical analysis for what experimental conditions are needed.

The second key is to effectively find the model (1) or actually its inverse f^{-1}:

$$xa = f^{-1}(\Xi) \tag{1A}$$

While xa may have been well defined, the **measured phenomena Ξ are in general very complex**. It is often unknown what in Ξ that actually determine xa. Previous solutions are to propose features $\varphi(\Xi)$ that can be calculated from Ξ and then fit a model from the experimental data. While the model structure may vary, from linear or artificial neural network, the features are hand-crafted manually based on conjectures, analysis of the phenomena, and the ability to extract these features from the measurements. The features are used as inputs of a proposed model $g_\theta(\varphi(\Xi))$ with a given structure g but unknow parameters θ:

$$xa = g_\theta(\varphi(\Xi)) \tag{1B}$$

The model parameters, vector θ, are fitted to minimize a measurement of overall difference for errors

$$g_\theta(\varphi\{\Xi(k)\}) - xa(k) \quad (k = 1, \dots, N) \tag{2}$$

where $\Xi(k)$ is the k^{th} measurement of Ξ. It is apparent the effectiveness of the resultant model in monitoring xa depends on (1) if Ξ contain sufficient raw information to determine xa; (2) if φ (2a) has effectively extracted all the fundamental information that is needed to determine xa from Ξ and (2b) can be accurately computed from Ξ; (3) if the model structure g_θ is appropriate/sufficient for the fundamental correlation between $\varphi(\Xi)$ and xa; (4) if the data $\{\ \Xi(k), xa(k)\ \}$'s $(k = 1, \dots, N)$ are appropriately generated to reflect all the welding conditions for the intended applications; (5) if the fitting is converged. This modeling process with five key steps is shown in Fig. 3. Major challenges are from steps (1)-(3). Fitting convergence is easily checkable and experimental conditions are largely ignored in research papers.

The final fitting error is commonly used as the unconditional measure for the accuracy of the fitted model. It is mistakenly believed that an acceptable fitting error proves the correctness for all previous steps (1)-(5). This is incorrect and it may be correct only if the experimental data $\{\ \Xi(k), xa(k)\}\ (k = 1, \dots, N)$ are generated under all representative conditions for the intended applications. Unfortunately, most studies do not provide justification for the experimental conditions used to generate the data. As such, a small

overall fitting error does not prove (1) the raw information is sufficient; (2) our feature selection is correct; and (3) our model structure is appropriate and sufficient.

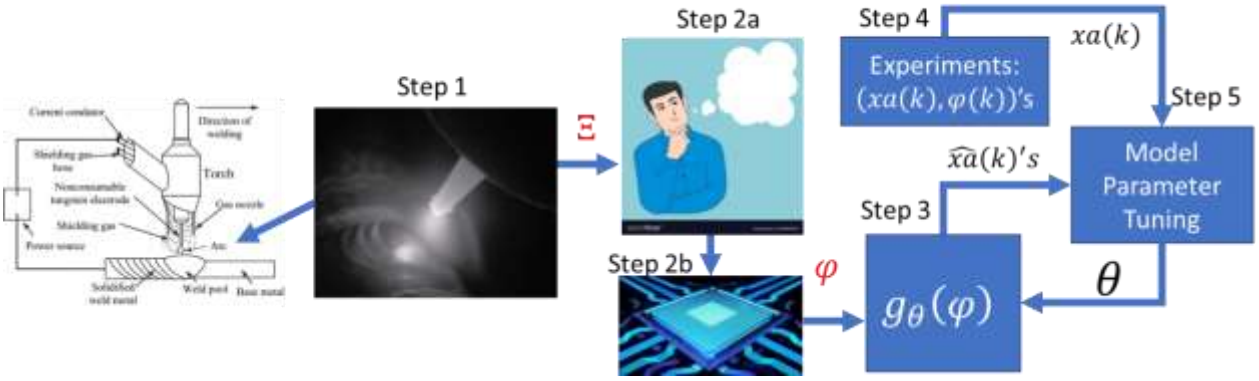


Fig. 3 Process and key elements to develop models to monitor weld penetration.

3. Historical Evolution

The historical evolution in this area can artificially be divided into three stages roughly as shown in Fig. 4. (1) Earlier efforts focused on finding and proposing phenomena that possibly correlate to the weld penetration. The features are hand-crafted and the computation needed from the raw data measured is relatively straightforward. The models are typically linear or simple nonlinear. (2) As the computation power is improved, more complex computation is introduced to draw more complex features from the raw data and the models become more comprehensive. In particular, machine vision and image processing methods are used to extract more complex hand-crafted features and artificial neural networks (ANNs) [100] are used to fit more complex correlation between the features and the penetration. (3) In recent five years, deep learning [101] method has gained application in penetration monitoring modeling by combining and improving Step 2 and 3 in Fig. 3. In particular, the hand-craft approach to manually select the features is replaced by a complex network whose structure allows it to be capable of extracting different features from the raw data by changing the parameters in the structure. In previous models (Stage 1 and 2) including ANNs, the raw data Ξ is converted to/represented by φ in Step 2 and φ is used as model inputs in Step 3. Using a deep learning method, the raw data can be directly used as model input. Step 2 and 3 are combined. Deep learning automatically selects the optimal features results by changing the parameters (rather than manually selecting features and trying again). Combining Step 2 and 3 provides a single step to link the raw data to the weld penetration and an effective method to automate the modeling process.

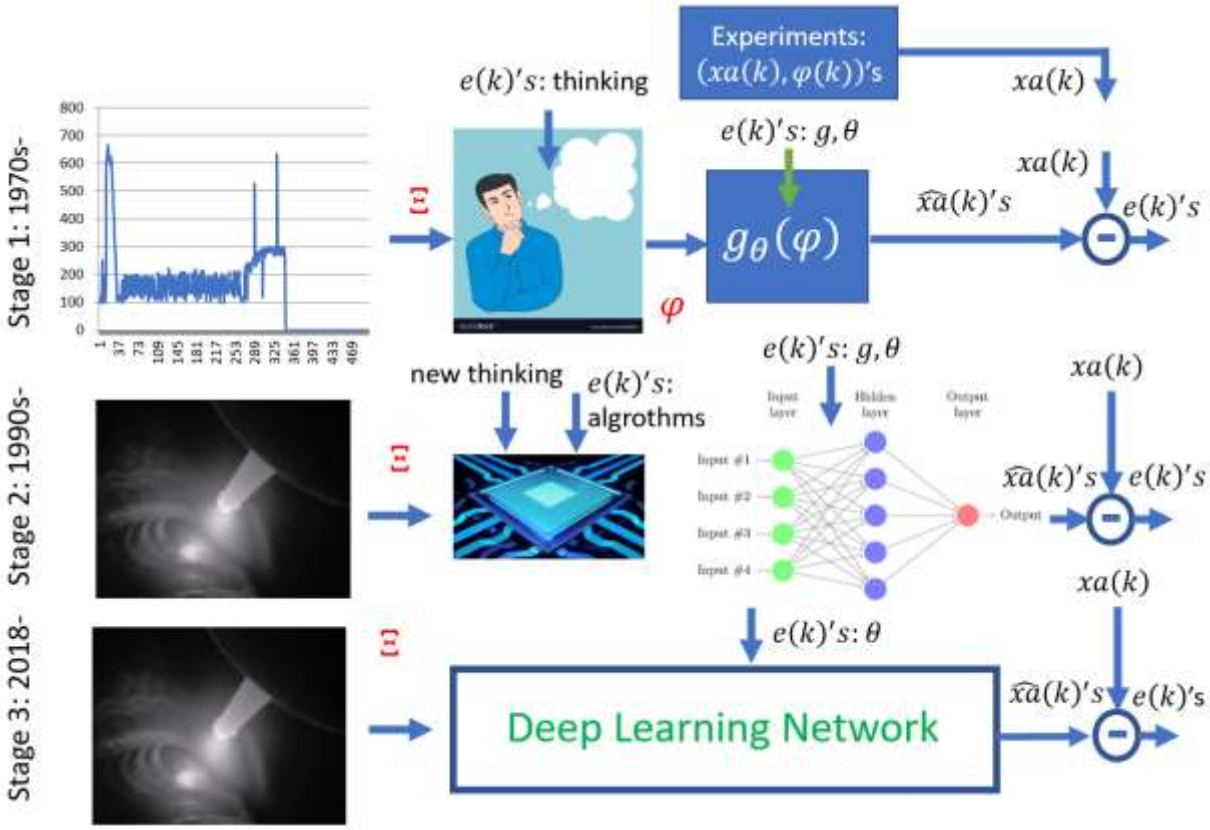


Fig. 4 Historical Evolution of Weld Penetration Monitoring.

The efforts in Stage 1 can date back to 1970s. The major contribution is that the correlation of the physical phenomena proposed/measured with the weld penetration is clearly supported by physical process. Kotecki et al. [102] is considered the first paper that introduced the possibility to corelate the oscillation of weld pool to weld penetration. It simplified the weld pool as a round membrane without thickness and correlated the radius of the round membrane to the natural oscillation frequency. The radius represents the state of the weld penetration. This is the few works where the phenomena Ξ to be measured are clearly correlated to the particular weld penetration state xa. However, this correlation is realized through ignoring the complexity and many other factors. Another representative work is due to Xiao and de Ouden [103] that experimentally demonstrated that the oscillation frequency drastically drops after the penetration transitions from incomplete to complete penetration. This difference is due to the change in the interface of the weld pool bottom from with solid/un-melted material to with atmosphere [104]. Another representative work [105] is from Auburn University that detects the boundary of the weld pool from infrared image based on the physical principle that the emissary of liquid is lower than that of the solid. When the temperature increases from the un-melted work-piece toward the center of the weld pool, the temperature increases. The infrared radiation received by the camera generally increases accordingly. However, when the material changes from solid to liquid, the emissary reduces. A reduced emissivity reduces the infrared radiation when the temperature is the same. There is thus a span where the radiation does not change. The radiation increases when the temperature further increases. There is thus a span where the slope of the radiation is zero that can be used to detect the pool boundary. Although

the pool boundary is not the weld penetration but it is closely related. This is thus another representative that finds relevant phenomena based on a clear physical principle.

The efforts in Stage 1 are featured by foundations for the correlation. They are more fundamental than these in Stage 2 that focus on the tools that made realistic due to the increased computation power. Most of them use the physical phenomena Ξ gained popularity due to the efforts in Stage 1. However, because of the increased computation, more comprehensive features can be calculated from more complex raw and more complex models can be fitted. In particular, images are more complex than welding voltage and current signals. Their features are also more complex and more difficult to be computed. One example is to use the weld pool boundary to predict the weld penetration [106]. In this work, the boundary of the weld pool was obtained from the weld pool image through image processing. The points of the boundary are used to fit a model in a polar coordinate system whose origin is also adaptively identified. The length of the weld pool and the distances at different angles from the origin are calculated to characterize the geometrical appearance of the weld pool (boundary). They are used as the inputs of an ANN to predict w_b. In this case, the length and distances at different angles are the hand-crafted features and their computation is relatively comprehensive. However, there are no fundamental changes in feature extraction and modeling despite the increased complexities. The use of computer vision and ANN both started to become popular in welding in 1990s [107-110].

Hand crafting the features is based on understanding of the physical process why the features proposed may correlate to the penetration state variables to be monitored. The success is checked/verified by the error of fitted models. If the error is larger than expected, another model is fitted or the features are modified. Either change involves manual processes: writing algorithms to fit the parameters in the new model structure, proposing new features that may be more promising, and revising algorithms to extract the new features from the raw data Ξ. Such manual processes are not automated and take time to realize and check the results. More importantly, even if the raw information may have sufficient information to determine the penetration state, if it can be effectively used to actually monitor the penetration depends on human. The complexity of the problem, detecting the red interfaces/activities underneath the surface from indirect observation from the green interfaces limit the ability to succeed. A way to better assure the success in a systematic way would be expected to be a game changer. Deep learning is such a game changer.

In the web of science up to April 11, 2022, within all the records related to “weld penetration”, there are 77 records [1-77] with “machine learning" or "deep learning" or "deep-learning" or "machine-learning" or CNN or RNN where CNN (convolutional neural network) and RNN (recurrent neural network) are the two most widely used networks for deep learning related welding. These references 1-77 are listed per their order from the web of science research. From April 12 to December 31, 2022, there are 15 records added forming our reference 78-92. Among these papers, those that do not automatically extract features from complex phenomena are excluded [18, 38, 45, 50, 53, 58, 63-66, 68-77, 82, 83, 87, 89, 90, 91] from our analysis and discussion.

We wish to specifically mention earlier works prior to 2020 including (1) two conference papers “Seam penetration recognition for GTAW using convolutional neural network based on time-frequency image of arc sound” (2018) [62] and “Convolutional neural network using Bayesian optimization for laser welding tailor rolled blanks penetration detection” (2019) [61]; (2) two journal papers “DeepWelding: A deep learning enhanced approach to GTAW using multisource sensing images” (2029) [14] and “Weld image

deep learning-based on-line defects detection using convolutional neural networks for Al alloy in robotic arc welding" (2019) [13]; and (3) two PhD dissertations (from Google Research for "weld penetration deep learning dissertation" but not in Web of Science) "Chao Li, 2019: WELD PENETRATION IDENTIFICATION BASED ON CONVOLUTIONAL NEURAL NETWORK. University of Kentucky" [93] and "Zarreen Naowal Reza, 2019: Real-time Automated Weld Quality Analysis from Ultrasonic B-scan Using Deep Learning. University of Windsor"[94]. These earlier records suggest (1) studies on deep learning based weld penetration monitoring started no later than 2018; (2) Peer-reviewed journal papers first appeared in 2019. As such, 2018 is considered Year Zero for deep learning based monitoring of weld penetration and 2019 signified its acceptance from peers.

4. Raw Information – Weld Pool Image

Monitoring the weld penetration can be considered as a two-phase process: deciding the phenomena Ξ (Phase 1A) and measuring them (Phase 1B) and developing a model that uses the information from Ξ as the input to predict the penetration state variables (Phase 2). Stage 1 efforts fundamentally contributed to Phase 1A. Phase 1B capabilities have been continuously developed/improved as there are general needs to monitor the welding processes for their current [111], voltage, arc radiation [112], sounds and acoustics [113], temperature fields [114], weld pool and weld pool surface [115-120] etc. In particular, the abilities to image the welding process and weld pool have been greatly enhanced by electronics for high speed, high spatial resolution, and high dynamic range. The welding phenomena can be viewed more clearly and faster despite the poor environment. As human welders are capable of controlling welding process per their visual feedback, more powerful raw information can be provided. The raw information can be from different sources to reflect the welding process from different angles. As we are unable to view the red phenomena (Fig. 2) directly and the physical process that determines the transition from red phenomena to the measured green phenomena Ξ is complex and not fully known, more information in general helps. However, the effective use of the raw information is also more complex. The human crafting based trial-and-error approach as in Stage 1 and Stage 2 for Phase 2 meets increasing difficulties. As deep learning packs all individual, human involved, separate processes into a single automated process, the raw information becomes the major factor determining the success for weld penetration monitoring.

4.1 Justification

Weld pool is undoubtedly the most important weld phenomena and is where all welding phenomena originate [121]. However, this does not directly justify, and none has theoretically proven, that it has sufficient information to determine the penetration state. Many studies used weld pool images to determine the penetration without reasonable justifications for the sufficiency of the raw information. However, it is true that human welders can only concentrate on watching the weld pool during welding and skilled welders are capable of assuring weld penetration in an acceptable range in most cases. They do adjust their operations per their observation on the weld pool. This suggests that the weld pool has reasonable information to determine the penetration in certain operation conditions for example the variations in the manufacturing conditions are in the tolerant range about the nominal ones.

While the weld pool may have the needed raw information to determine the penetration state, it has never been crystally clear what are the critical information in the weld pool that determines the penetration state. Earlier works correlate the penetration state to the oscillation of the weld pool [102, 103], while later works correlate to the weld pool boundary or geometrical appearance (size and shape) [106, 107]. Even for the oscillation and geometrical appearance, it is still not crystally clear what are the

critical information. The frequency of the oscillation has been used as the major critical information to correlate to the penetration [103] so that various methods have been proposed to monitor the oscillation frequency [122, 123] and the frequency has been used a model input (feature of the weld pool) to predict the penetration state. The length and width have been used as the major critical information as the features of the weld pool geometry to correlate to the penetration. More comprehensive approaches are to use a geometrical model or a set of feature parameters (length and relative widths in relation to the width, possibly also the area of the weld pool) [124] to provide more accurate description for the size and shape, i.e., the geometrical appearance the human welder may perceive/feel) [125-128].

As it is not crystally clear what is the critical information, such human crafted features for weld pool vary. Each different set of crafted features may require a different approach to acquire the relevant weld pool raw information. In the pool oscillation based approaches, the raw information has been those that can be used to calculate the frequency of the weld pool including arc voltage [103], arc radiation [122], etc. In the pool boundary based approaches, the raw information is the boundary of the weld pool and different approaches have been used to acquire the needed raw information. A way is to extinguish the arc by short-circuiting the arc on purpose or taking advantage of the natural arc short-circuiting period [129] to image the weld pool. Additional ways include pulsing the current to image during the base current period [130], using pulsed illumination that is much brighter than the radiation of the arc and synchronizing the camera shutter with the illumination pulse [131].

We should realize that the oscillation and boundary are not all the phenomena the human welders may perceive from the weld pool. They were proposed as there are limitations on our imagination and reasoning for more comprehensive features and limitations on our abilities to more completely measure/image the weld pool. For example, most researchers can understand that the 3D surface of the weld pool is more comprehensive and should have more sufficient information to determine the penetration state but more of them used 2D boundary of the weld pool as the 3D surface is more difficult to image and then extract from the image. Two methods have been proposed to measure the 3D surface, stereo vision approach and reflective method. The stereo vision approaches acquire two images from different angles either using two cameras [132] or a camera with a biprism [133]. However, it is still not clear what are the critical information from the 3D surface. A set of characteristic parameters has been proposed to characterize the 3D weld pool surface by introducing/adding a novel 3rd dimensional parameter, termed the convexity that is ratio of the thermally expanded metal above the original work-piece surface normalized by the pool length, to the width and length of the weld pool [134]. This set of characteristic parameters is found to represent the penetration state in full penetration well and has been used to as the input of human welders for their adaptive intelligence [135, 136].

As such, it is not only challenging for researchers to propose promising/selected raw information from the weld pool, but also challenging to develop the associated measurement method to acquire the raw information, to propose the right features, and to process the raw information to extract the features. With all of these efforts, the chance to succeed may not be promising.

Weld pool images may be incorrectly assumed to be the same as human observation. While this assumption cannot be true, they still have abundant information about the weld pool, although not all of them, and many phenomena the human welders may respond to. Calculating any hand-crafted features from them may lose raw information that may be critical in determining the penetration state. If a method is available to directly use the images with the raw information preserved, unintentional loss in

transferring to the hand-crafted features can be avoided. With the development of the deep learning algorithms, models that can be trained convergently become complex. This allows the increase of the dimension in the model inputs. Crafting features from raw data to reduce the dimension of the model inputs and model complexity becomes unnecessary.

As such, direct use of the weld pool images of high dimensions becomes realistic. This provides an opportunity to use "standardized" raw information. While the critical information in the weld pool images that actually determine the specific penetration state variables in specific applications differs from application to application, the method to acquire the weld pool images may be "standardized" rather than "customized". The development of the deep learning algorithms also improves the standardization of the model training. Using deep learning to monitor the penetration becomes (1) determining if the specific penetration state possibly correlates to the weld pool; (2) if yes, conducting experiments to collect (pool image, label) where label is the penetration state variables concerned; (3) choosing a suitable deep learning model to correlate the weld pool images to the labels and train the model using quite standardized training procedure. In many studies, Step (1) is fully ignored and the correlation is assumed and validated by the modeling error.

Table 1: Deep learning-based penetration state monitoring using one wed pool image.

Ref. #	Author	Year	Title	Camera Note	Process	DL Model	Model Output	Other Notes
13	Zhang, ZF	2019.9	Weld image deep learning-based on-line defects detection using convolutional neural networks for Al alloy in robotic arc welding		GTAW	CNN	Penetration state classification	Incomplete, Complete, Burn
49	Zhang, ZH	2020.1	Real-time penetration state monitoring using convolutional neural network for laser welding of tailor rolled blanks		Laser	CNN	Penetration state classification	Conf paper 2019
16	Xia	2020.8	Vision based defects detection for Keyhole TIG welding using deep learning with visual explanation	HDR	K-TIG	CNN	Penetration state classification	Incomplete, Complete, Burn
7	Wang, CY	2020.10	Deep learning-empowered digital twin for visualized weld joint growth monitoring and penetration control		GTAW	CNN	Wb	
41	Jiao, WH	2020.11	Prediction of Weld Penetration Using Dynamic Weld Pool Arc Images		GTAW	CNN	Penetration state classification	Pool and arc in image
8	Nomura	2021.1	Burn-through prediction and weld depth estimation by deep learning model monitoring the molten pool in gas metal arc welding with gap fluctuation		GMAW		d	
12	Jiao, WH	2021.3	End-to-end prediction of weld penetration: A deep learning and transfer learning based method		GTAW	CNN	Penetration state classification	Transfer learning
4	Kershaw, J	2021.9	Hybrid machine learning-enabled adaptive welding speed control		GTAW	CNN	Wb	Control, learning from human
15	Jia, CB	2021.12	Penetration/keyhole status prediction and model visualization based on deep learning algorithm in plasma arc welding		PAW	CNN	Penetration state classification	Transfer learning
40	Wang, ZM	2022.4	Recognition of penetration state in GTAW based on vision transformer using weld pool image		GTAW	ViT-B/16	Penetration state classification	Transfer learning from ImageNet
78	Yu, R	2022.7	How to accurately monitor the weld penetration from dynamic weld pool serial images using CNN-LSTM deep learning model?	HDR	GTAW	CNN-LSTM	w_b	Precision monitoring
92	Zhou FZ	2022.08	Keyhole status prediction based on voting ensemble convolutional neural networks and visualization by Grad-CAM in PAW		PAW	CNNs	d	8 CNNs
80	Baek	2022.10	In-process prediction of weld penetration depth using machine learning-based molten pool extraction technique in tungsten arc welding		GTAW	ResNet	Pool shape for d	
88	Wang ZM	2022.10	Recognition of GTAW weld penetration based on the lightweight model and transfer learning		GTAW	CNN	Penetration State Classification	Transfer learning
29	Oh	2021.6	Deep-learning approach for predicting laser-beam absorptance in full-penetration laser keyhole welding	Two	Laser	ResNet	Object detection	For laser-energy absorptance
23	Kim	2021.8	Deep-learning-based real-time monitoring of full-penetration laser keyhole welding by using the synchronized coaxial observation method	Two	Laser	YOLOv4	Object detection	Image interpretation
26	Li	2022.7	Deep learning-based fusion hole state recognition and width extraction for thin plate TIG welding	Two	GTAW	YOLOV3	Object detection	Image interpretation

The green highlighted studies in Table 1 use weld pool images as inputs of a deep learning model to directly predict penetration state variables and they are listed in the order of their publication time or the time of their first public appearance as indicated in the Web of Science. Listed also include works that detect image features to be used to predict the penetration as highlighted yellow in Table 1. They are listed as references but will not be discussed as they do not directly use weld pool images as the model inputs to predict the penetration.

As all the inputs are the weld pool images, the major difference among these works lies in their outputs. A number of them [12, 13, 15, 16, 40, 49, 88] focused on classification of the penetration state. In addition to incomplete and complete penetration, some of these works also added over penetration or burn-through. Some of these works refer incomplete and over penetration to as defects and this is true for an application that requires complete penetration. A number of later works in Table 1 use weld pool images as inputs to predict the quantitative measures of the penetration, d for incomplete penetration and w_b for the complete penetration. Jiao's work [41] fused the current with two previous weld pool images to consider the dynamic behavior of the welding process, although the number of the images used as the model input is still one. Zhang's work [13] combined two weld pool images taken from two different angles and one back-side image into a single image as the input. Using multiple information source and dynamic information source will be discussed later.

4.2 Instant Weld Pool Image

The work by Zhang [13] is the earliest work published in a peer-reviewed journal in using deep learning to monitor the penetration, regardless of the information source and the actual model structure used. The authors used Fig. 5 to illustrate the difference between conventional and deep learning based approach. The single image on the left used as the input of the deep learning model includes the weld scene taken at a leading angle (right, upper), weld scene behind the pool (right, lower) and back-side of the weld (left). Similar images have been imaged previously in an earlier publication by the corresponding author Chen SB [137]. Fig. 6 shows the visual images associated with different penetration states: incomplete, complete, and over penetration. A CNN model has been successfully established to classify the penetration state from such images.

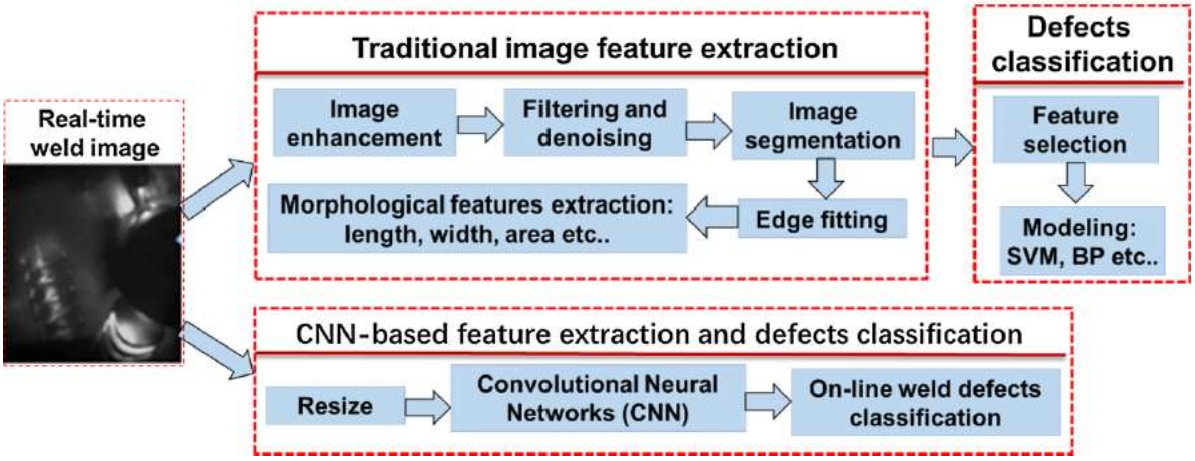


Fig. 5 Illustration of difference between conventional and deep learning based approaches in weld pool image based monitoring [13].

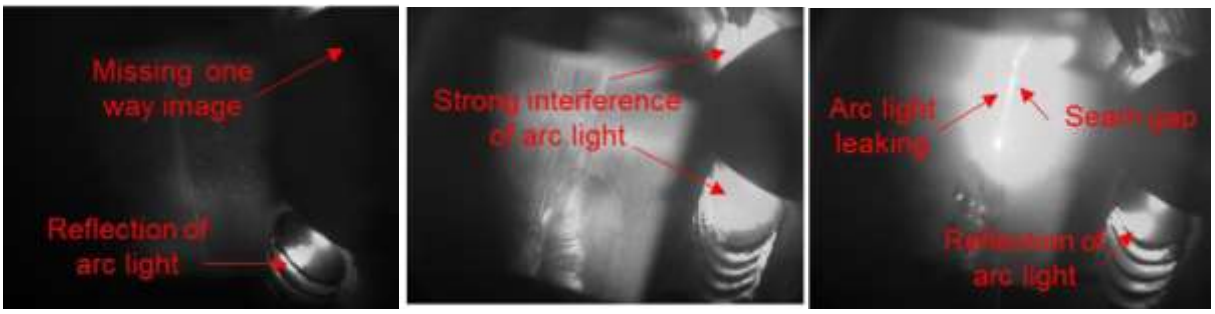


Fig. 6 Images associated with different penetration state [13]. From left to right: incomplete, complete and over penetration.

While the input weld pool image above includes the scene from the back-side (purple in Fig. 2), the work by HZ Zhang [49] only used the weld pool image taken from the front (red interface in Fig. 2). They acquire the weld pool image during laser welding process using a co-axial camera for a co-axial view of the weld pool and keyhole. A co-axial view in laser welding is a standard monitoring technology in laser welding that uses a dichroic mirror whose principle is shown in Fig. 7. The reflection and transmission properties of such a dichroic mirror change with the wavelengths. A proper dichroic mirror allows the laser (wavelength) to be reflected to the work-piece and the lights emitted from the weld pool to pass through to reach the camera. Images in Fig. 8 can thus be obtained co-axially while such a co-axial view in arc welding is less straightforward [138, 139] and is considered a significant innovation. The images on the left clearly demonstrate that they are determined by the penetration state. A CNN model has been trained to classify the penetration state from the co-axial image.

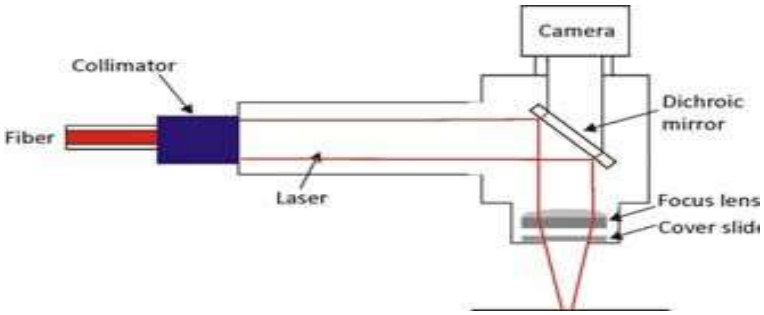


Fig. 7 Co-axial view of laser welding process using a dichroic mirror [49].

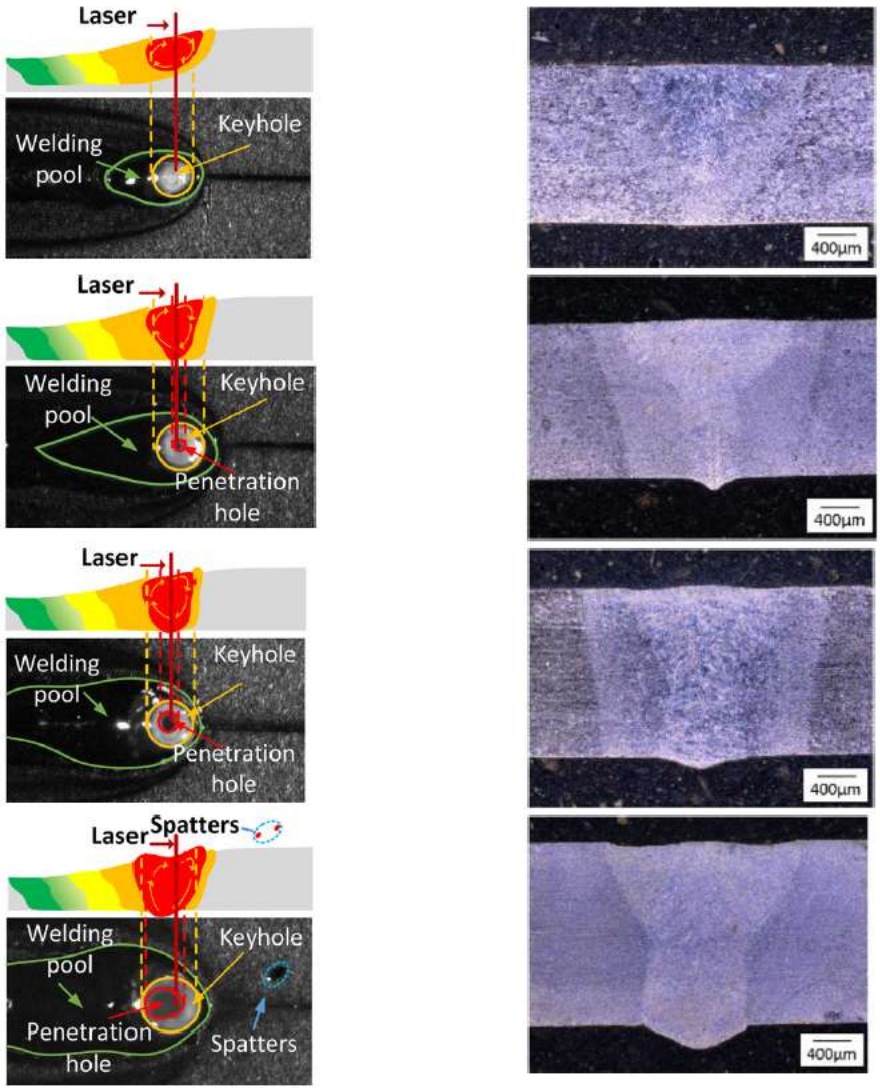


Fig. 8 Co-axial view of weld pool and keyhole during laser welding and corresponding penetration [49].

The above work is for laser welding which easily allows a co-axial view. It is arguable that the correlation between the weld penetration with the weld pool image (from the red interface) is clearer and more definite in a keyhole welding such as laser keyhole welding than in a melt-in process such as GTAW (gas tungsten arc welding). As such, the monitoring of penetration may be more straightforward in laser keyhole process, allowing a clearer image (co-axial view) and having a stronger process feature due to the keyhole, than a GTAW process where the keyhole and co-axial imaging are not typical. Hence, Zhang ZF [13] included the scene from the back-side so that the input image has more sufficient information. In the work by Xia [16], the process to be monitored is a keyhole GTAW, termed K-TIG in literature [140]. This is a special GTAW process where a keyhole is typically needed to assure the penetration using an extremely high current. A coaxial view is more challenging so that the weld pool images were taken non-coaxially (Fig. 9). As can be seen in Fig. 10, the correlation between the weld pool image and penetration state becomes less apparent. However, using a CNN, the weld pool images can still predict the penetration states and detect defects.

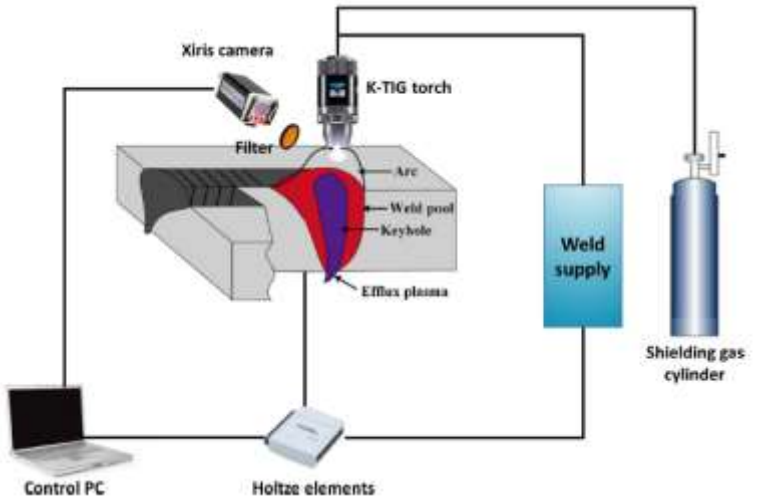


Fig. 9 Monitoring of K-TIG process [16].

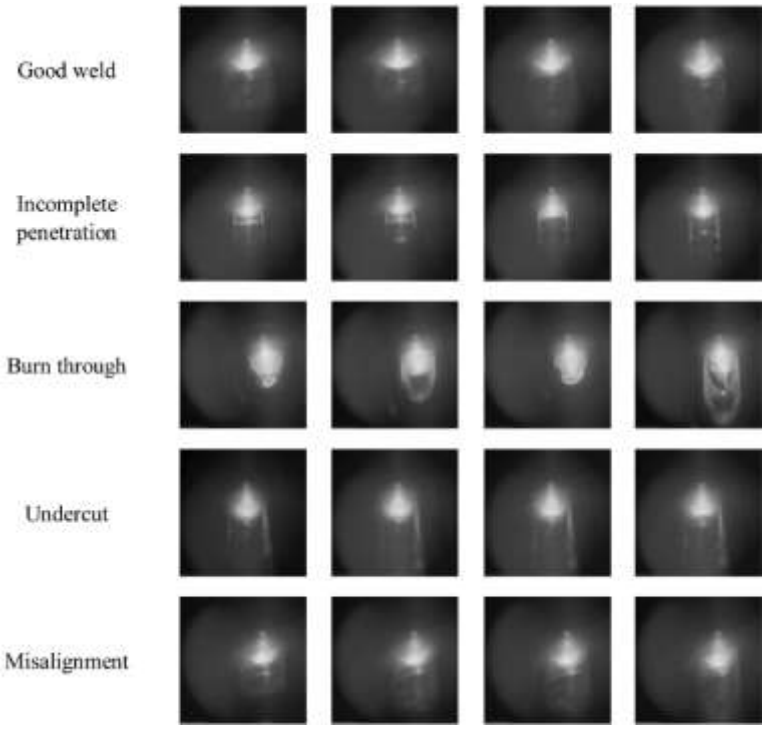


Fig. 10 Weld pool image in K-TIG process acquired using an HDR (high dynamic range) camera [16].

While the earlier works focused on classification, later works in Table 1 that use a single weld pool image as the input of a deep learning model provide quantitative prediction for the penetration state, in particular w_b for complete penetration and d for partial penetration. All these works for the complete penetration were from the University of Kentucky while the work for partial penetration is from Osaka University. For the complete penetration, the label can be easily acquired and paired with the weld pool image so that the acquisition of the data for training, validation and test is relatively straightforward.

"End-to-end prediction of weld penetration: A deep learning and transfer learning based method" by Jiao [12] is a paper that first appeared online in January 2020 but the special issue was published in March

2021. It is for classification of penetration state in GTAW. In the previous work [13] reviewed from GTAW, three images were used including one from the back-side (purple interface) and the information in the image is more apparent for different state. In this work, only one weld pool image is used. The penetration state (label) is measured using a second camera but the label is only used for training and is not part of the input of the prediction model. As the raw information is less and is less apparent than that in Fig. 6 [13], extra effort may be needed. Using a CNN, the accuracy for the classification was 92.7%. To improve the accuracy, Jiao developed a transfer learning approach based on residual neural network (ResNet). This ResNet-based model was first pre-trained on ImageNet dataset with the hope to better extract the feature. The dataset from this work was used to just modify the fully-connected layers. The accuracy was improved to 96.35 %. As shown in Fig. 11, mis-classification only occurred in the transition between two adjacent states.

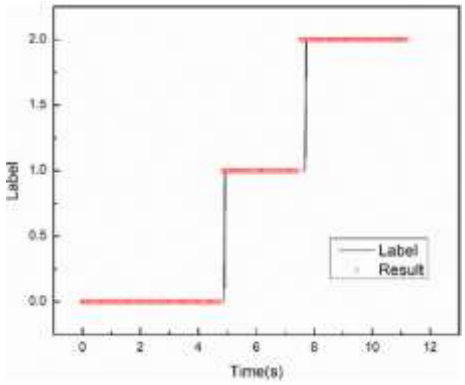


Fig. 11 Illustration of test result for transfer learning [12].

"In-process prediction of weld penetration depth using machine learning-based molten pool extraction technique in tungsten arc welding" [80] uses a ResNet to extract the weld pool boundary from weld pool image. The weld pool shape is then correlated using an ANN to the penetration depth. We note that deep learning is not used to directly correlate the raw image to the penetration but used to segment the weld pool to detect the weld pool shape. This shape is a kind of hand-crafted features and deep learning is used as an image professing method to obtain the proposed features.

4.3 Dynamic Weld Pool Images

As mentioned, the weld pool image (red interface) in typical GTAW (melt-in mode) does not have very apparent features to determine the penetration as these in Fig. 6 (including back-side image from purple interface) or those in Fig. 8 and Fig. 10 where a keyhole is present. [141, 142] analyzed the evolution of the convexity of the weld pool surface due to thermal expansion and penetration. When the penetration is partial, as the weld metal, thus the thermal expansion volume, increases, more metal exceeds the work-piece surface. After the penetration is full, part of the liquid metal flows toward the bottom of the work-piece and this flow increases as w_b increases. The metal exceeds the work-piece surface tends to reduce. As such, a partial penetration weld pool could be relatively small and convex while a full penetration weld pool could be larger and less convex. Such changes should also be reflected in the weld pool image and its dynamic change. This idea has been used to design image sequence recedingly [41] to improve the classification accuracy. The weld pool images are sampled at 30 HZ. When a new image is acquired, this image and the previous two, acquired 5 sampling periods ago and 10 sampling periods ago respectively, form a new image sequence and are used as the 3-chaneel inputs of a CNN to predict the penetration

state. It was found that using image sequence increases the classification accuracy from 92.7% to 94.2%. Dynamic weld pool sequence thus provides another method to increase the classification accuracy in addition to using the transfer learning.

Wang QY [7] is probably the first to quantitatively predict the penetration state from weld pool images using a deep learning model. Per the above principle using dynamic evaluation of weld pool, Wang used two weld pool images, one from the middle of the base current period when the weld pool is small and the penetration reduces to partial and one from the middle of the peak current period when the weld pool has grown, to predict w_b as shown in Fig. 12. The two weld pool images are used as the input of a deep learning model to predict the penetration. The predicted w_b is used to visualize the weld pool to the human operator. The human operator adjusts a virtual torch to control the welding robot [7] whose motion follows that of the virtual torch through a virtual reality system.

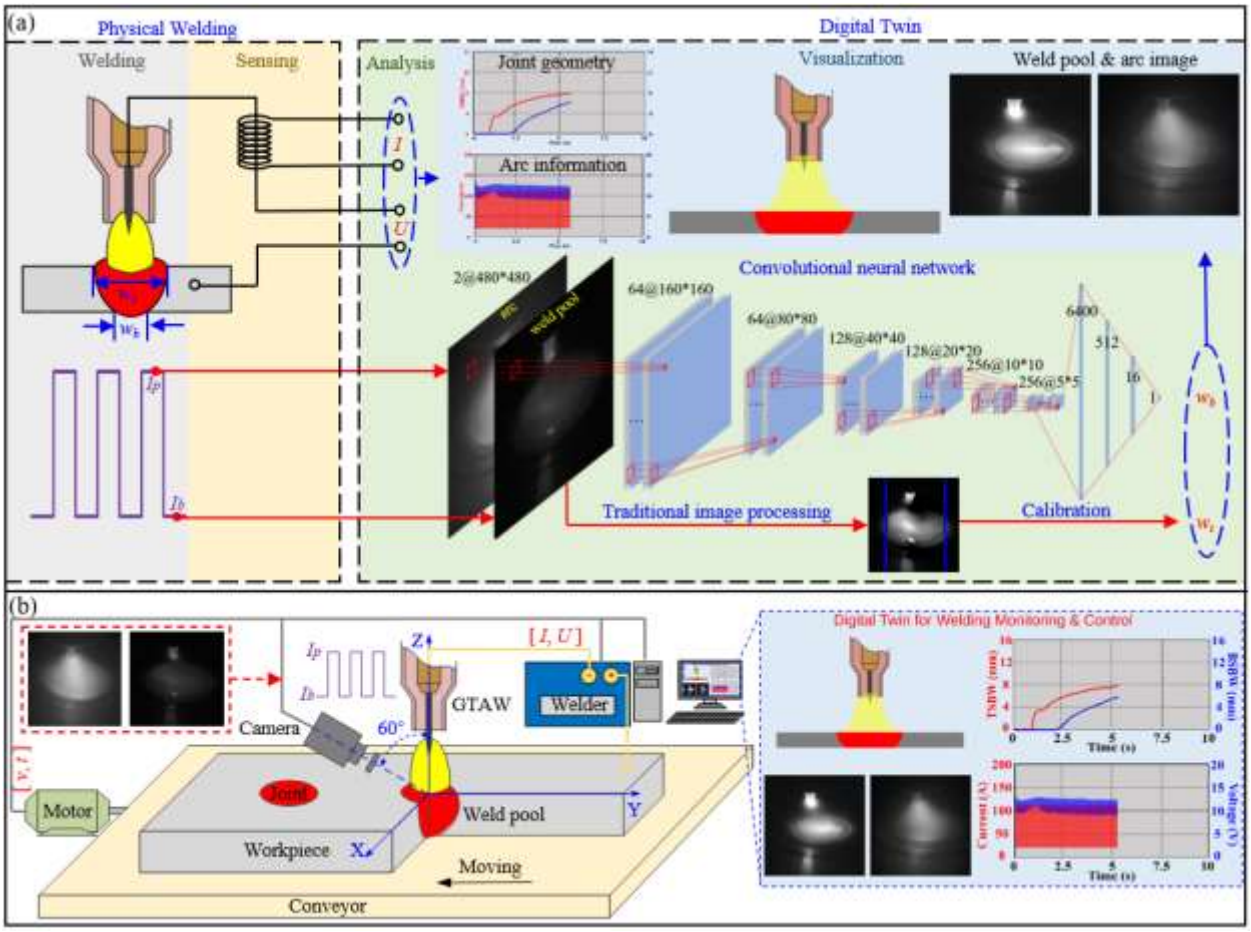


Fig. 12 Digital twin for human-robot collaborative welding based on deep learning predicted penetration [7].

Yu R [78] further explored the contribution from dynamic weld pools. An HDR camera is used to acquire the weld pool image in GTAW. The welding current and travel speed are purposely changed randomly in specified large ranges. Multiple experiments were conducted and the actual current and travel speed waveforms are not exactly the same in different experiments but having the same statistics theoretically. Such large dynamic changes in the major welding parameters (current and travel speed) increase the

challenges in predicting w_b but they simulate the environment where a penetration prediction monitoring and control system operates while other existing works on deep learning based monitoring of penetration state seem having ignored this critical requirement. By looking at Fig. 14, we can observe that the CNN performed poorly in predicting the penetration at the start of welding when using a single weld pool image. The model struggled to capture the drastic dynamic changes of the penetration that occur at the beginning of welding, although it performed reasonably well during welding when the changes were less drastic. By using the dynamic weld pool images recursively using a long-short term memory (LSTM) recurrent neural network (RNN), accurate prediction is observed everything. The average prediction error for w_b is less than 0.3mm while it is 0.54 mm for the CNN using one weld pool image.

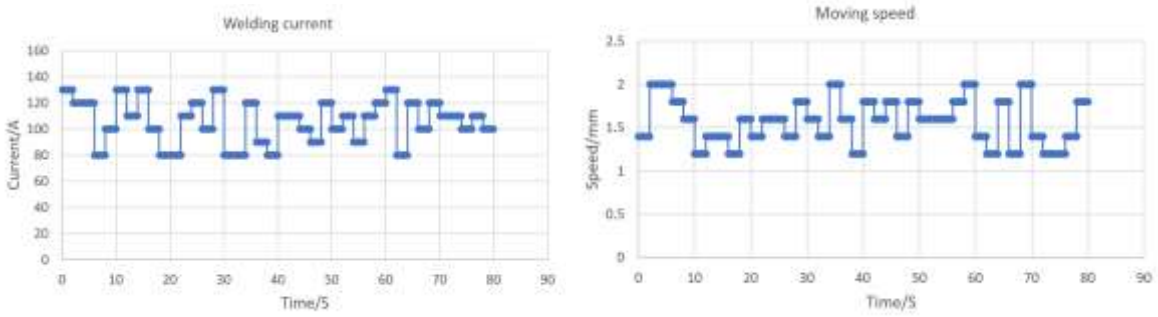


Fig. 13 Welding current and travel speed in [78].

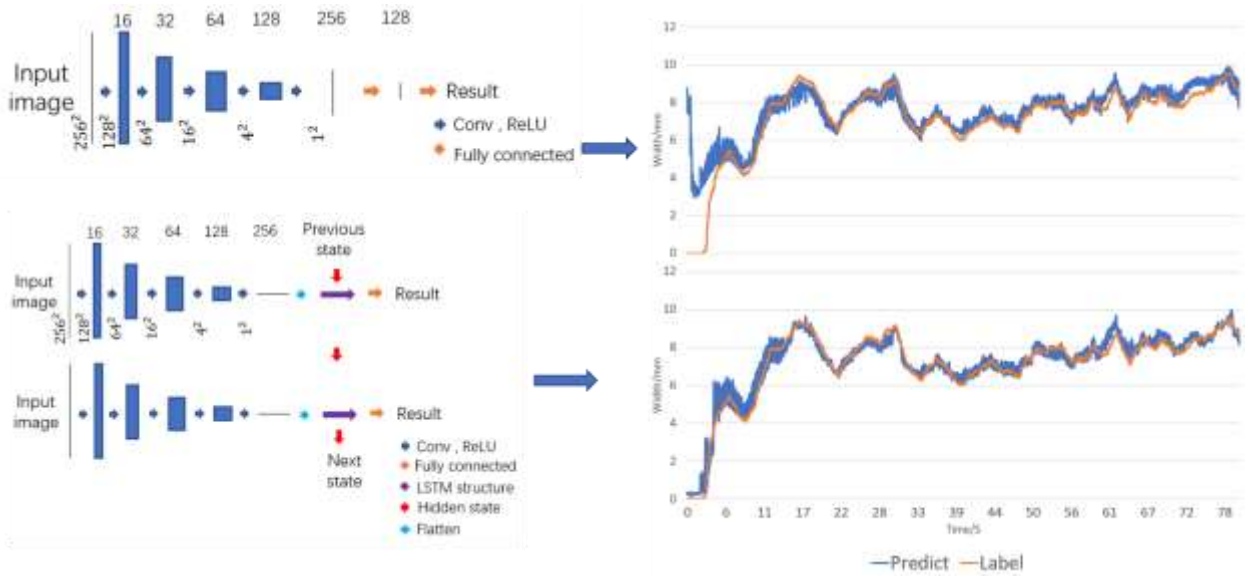


Fig. 14 Deep learning models and their predictions [78].

In summary, it is not crystally clear if the observable weld pool is fully determined by the penetration state and contains the needed, sufficient, raw information to determine the penetration. However, we tend to hypothesize on its sufficiency as we do not have other more promising observable phenomena to use. We believe this is at least partially true in certain applications as human welders can only observe it during welding and skilled welders may assure the penetration in most cases. However, it is unclear what are the critical raw information in the weld pool. Weld pool images appear to provide comprehensive phenomena occurring on the weld pool. With the availability from deep learning, the weld pool image could provide

a standardized input for deep learning models although the specific penetration state is application dependent. However, our observation on human welder capability in assuring weld penetration does not tell how "weld pool" is defined. We found that a single image from the instant when we predict may not have sufficient raw information for non-keyhole process like GTAW. Using the transfer learning can improve the accuracy through more effectively extracting "all possible patterns" in the data. However, to further improve the accuracy, the sufficiency in the raw information needs to be improved. Per the analysis of weld pool and penetration development, it appears that dynamic images are needed. This desires a recursive deep learning model like LSTM-CNN.

5. Raw Information – Images from the Weld Pool

The weld pool images in the above section are general visual observation of the weld pool. The weld pool may also be observed in other ways for its specific properties such as temperature field, 3D surface, pool oscillation, and reflection etc. Such properties also result in images. We refer them to as images from weld pool to distinguish from weld pool images. Table 2 lists papers that use images from the weld pool as inputs for deep learning model to predict the penetration.

Table 2 Images from Weld Pool Used in Deep Learning

Ref. #	Author	Year	Title	Signal Note	Process	DL Model	Model Output	Other Notes
20	Knaak, C	2021.6	A Spatio-Temporal Ensemble Deep Learning Architecture for Real-Time Defect Detection during Laser Welding on Low Power Embedded Computing					
9	Yu, RW	2022.1	Monitoring of back bead penetration based on temperature sensing and deep learning					
79	Buongiorno	2022.7	Inline Defective Laser Weld Identification by Processing Thermal Image Sequences with Machine and Deep Learning Techniques		Laser	Contain CNNs	Defects	Thermal image sqquence
24	Cheng, YC	2020.8	Detecting dynamic development of weld pool using machine learning from innovative composite images for adaptive welding					Two reflective images
28	Cheng, YC	2021.3	Dynamic estimation of joint penetration by deep learning from weld pool image					
47	Ma, GH	2021.4	A vision-based method for lap weld defects monitoring of galvanized steel sheets using convolutional neural network	Laser stripe				On solidified metal

5.1 Temperature Field

Welding and penetration are realized through heating and melting the work-piece. Temperature field reflects the heating and melting and thus provides certain information about the penetration. It is not clear if there is an exact mapping between the temperature field and the penetration. Theoretically, such an exact mapping should not exist and the temperature field from the weld pool and its surrounding area may not have sufficient raw information about the penetration. However, a change in the temperature field can be caused by the weld penetration so that the temperature may be still used to provide an estimate for the penetration in certain circumstances.

Yu RW et al. [9] defined different penetration states in cold metal transfer (CMT) butt welding as shown in Fig. 15. They realized that the measurement from the infrared camera is not emissivity independent. The emissivity on the weld pool surface varies with the slope and is not constant while it is on solid surface. They thus believe that the measurement from the solid surface can better reflect the temperature. As such, they first studied the transverse temperature profile at a fixed position behind the weld pool (along the upper horizontal line in the image in Fig. 16). It was found that this temperature profile is always Gaussian and the distribution of the parameters in the fitted Gaussian distributions is well separable for different states (right in Fig. 16). Two fixed regions of interest (positions and sizes) are selected with the intent for solid regions to be as close to possible to the weld pool. Each of them has been used as the

input to fit a deep learning model to classify the penetration state. They both achieved 100% classification accuracy. They noted that the temperature field is sensitive to surface condition/impurity and recommended to clean the surface before welding.

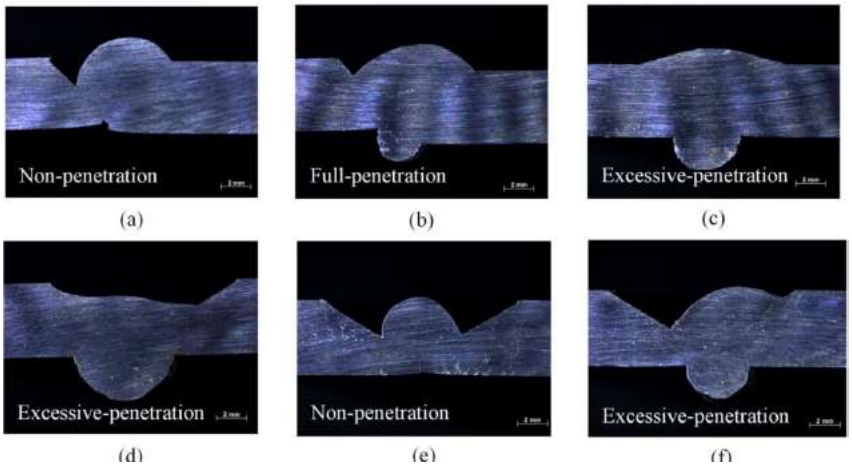


Fig. 15 Penetration states defined and to be detected [19].

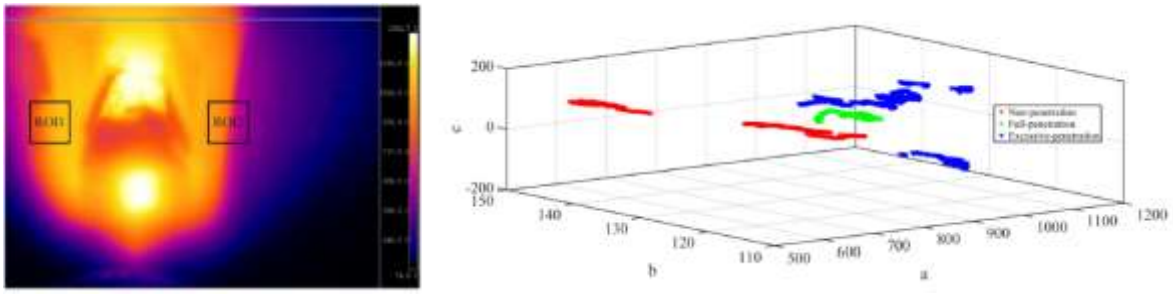


Fig. 16 Positions for temperature profile and regions of interest (left) from the infrared image and distribution of parameters in the fitted Gaussian profiles (right) [19].

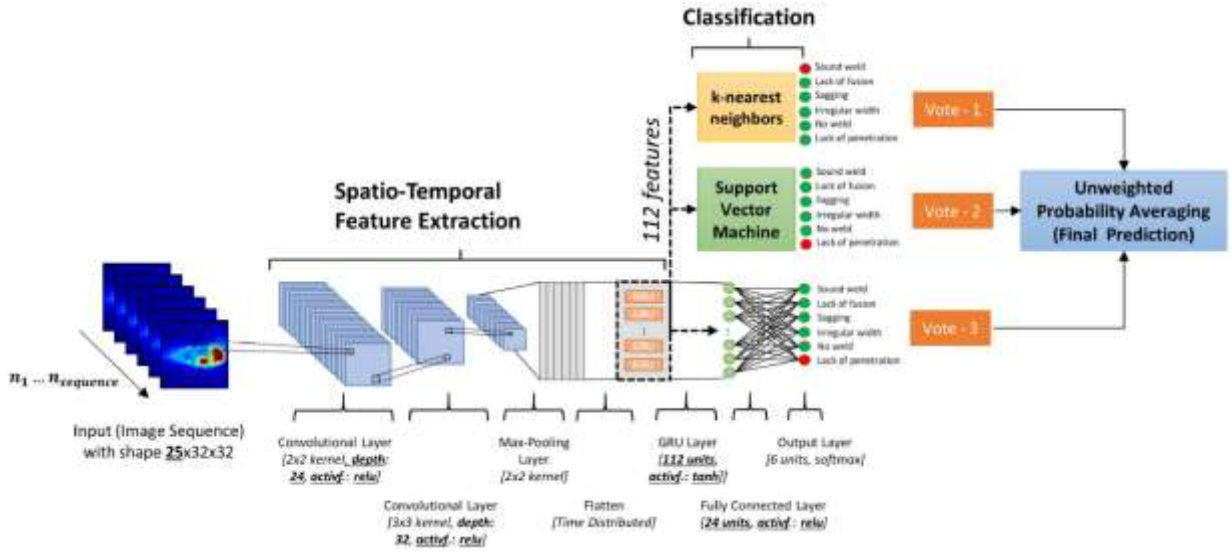


Fig. 17 Detection of laser weld defects using sequential near infrared images [20].

Reference [20] was published before [19] discussed above but it is much more comprehensive. As shown in Fig. 17, a sequence of $n_{sequence} = 25$ near infrared images are used as the input of the deep learning model. The images are processed by a CNN through convolution and pooling to generate 112 features that are inputted into three parallel classifiers through GRU (gated recurrent unit). Each of the three classifiers classifies the weld into one of the six classes to generate a vote. The final prediction is decided from the three votes. The accuracy reaches 97.8%.

Reference [79] acquired a sequence of thermal images; and used them/some of them to conduct cooling analysis to propose 140 features, perform image analysis to propose 78 features; and crop them to do video analysis to propose 5 features. A number of techniques including DT (decision tree), KNN(K-nearest neighbor), SVM, and three CNNs to reduce these features. The reduced set of features are further down selected using DT, KNN, and SVM. Finally, the selected features are used by DT, KNN, and SVM for classification.

5.2 Weld Pool Reflection Images

Weld pool surface is three-dimensional but weld pool images provide only two-dimensional information. There is a possibility of loss of raw information from weld pool to weld pool images. In particular, skilled human welders are able to see it as a 3D subject. As such, weld pool images do not fully preserve the raw information human welders can perceive from the weld pool. This potentially jeopardizes the needed sufficiency in the raw information which is the key in deep learning based penetration monitoring. Studies have been conducted to measure 3D weld pool surface and use it to predict the penetration and study human welders' response to welding process. It was found that the human welders can adjust welding parameters per 3D surface of the weld pool. As such, 3D surface of the weld pool alone may provide the needed information to determine the penetration as shown in [134].

Measuring the 3D surface of the weld pool is challenging. As discussed in Section 4.1, two methods have been proposed to monitor the weld pool surface. However, only its features (characteristic parameters) have been used to predict the penetration [134] and 3D surface has not been used as the input of a deep learning model. As the computation needed for both methods to compute the 3D surface is extensive and the algorithm requires the setting of the measurement system, it is indeed an interest to see if the images can be directly used to predict the penetration without calculating the 3D surface first. This is possible as the 3D information has been included in the images although has not been exclusively computed out. Further, it is also interesting if the change in part of the weld pool surface would be sufficient.

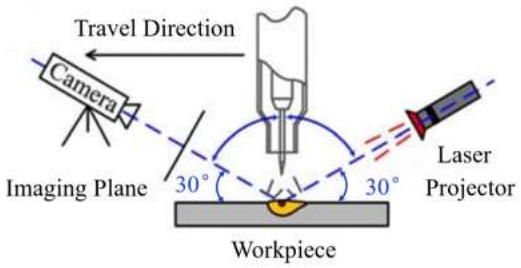


Fig. 18 Imaging of laser reflected from weld pool surface [119].

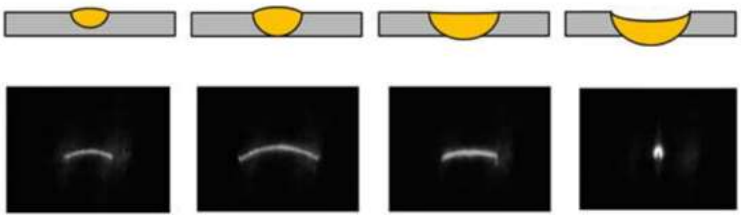


Fig. 19 Reflective images of weld pool under different penetration states.

Fig. 18 shows such a measurement system that projects a laser stripe on the weld pool surface that reflects it like a mirror to the imaging plane [119]. The camera aims at the imaging plane and obtain images with the laser stripe distorted by the shape of the weld pool surface. The shape of the weld pool surface that reflects the projected laser stripe changes with the weld penetration (Fig. 19) [141]. In [24], two reflective images are used as the input of a CNN deep learning model to classify the penetration state. One of the images is the reference one that is taken at the beginning of the welding when the weld pool is just established. Another is the image taken in real-time when the prediction needs to be made. The classification accuracy is 97.5% [24]. In [28], this method is used to predict w_b under varying current and travel speed. As can be seen, despite the simple image without details as in weld pool images, excellent accuracy has been achieved. 3D surface, even if just partial, appears does have the needed raw information for deep learning model to extract useful features.

A laser stripe may also be projected onto the solidified weld to detect the penetration state [143-146]. In [143], this principle was verified. However, in conventional method, the laser stripe needs to be extracted using machine vision algorithm [144]. When there are spatters and arc radiation, this may be challenging. Deep learning approaches have been developed to extract the distorted laser stripes from noisy images [147, 148]. In addition, when the weld is relatively flat, the distortion of the laser stripe is not apparent so that the detection of the weld is challenging [144]. Using a CNN deep learning model, Ma GH et al. have successfully detected the weld defects from the imaged laser stripe projected on the weld behind the weld pool [47].

5.3 Active Pool Oscillation Image

Pool oscillation was among the first proposed to monitor the penetration. As we mentioned, oscillation frequency has been the major feature used to correlate to the penetration. However, the pool is a three-dimensional body whose oscillation is complex and difficult to craft its features. With the availability of the deep learning approaches, it is possible to directly extract its relevant features automatically.

The system in Fig. 20 projects a 19X19 laser dot matrix to the entire weld pool [2]. Image (a) (lower left) is acquired from Camera 1 and Image (b) (lower right) is acquired by Camera 2. Image (a) is used as the input of a CNN and image (b) is used to calculate the label representing the penetration state. As can be seen from Fig. 21, the reflection dots in the image are clear and distribute regularly in the beginning but the number is small. This is because the weld pool is small so that the number of project laser dots being intercepted and reflected by it is small. Also, the weld pool is convex and stable so that the reflection is regular. As the weld pool increases, the oscillation amplitude is larger and shape of the weld pool surface deforms significantly. The change of the mirror shape and oscillation of the mirror alter the shape of the reflected laser dot pattern and the clearness of the laser dots. However, handcrafting features from them

is challenging. Using a CNN, the reflection image can classify the penetration state with 97.6% accuracy through voting from results from consecutive images [2].

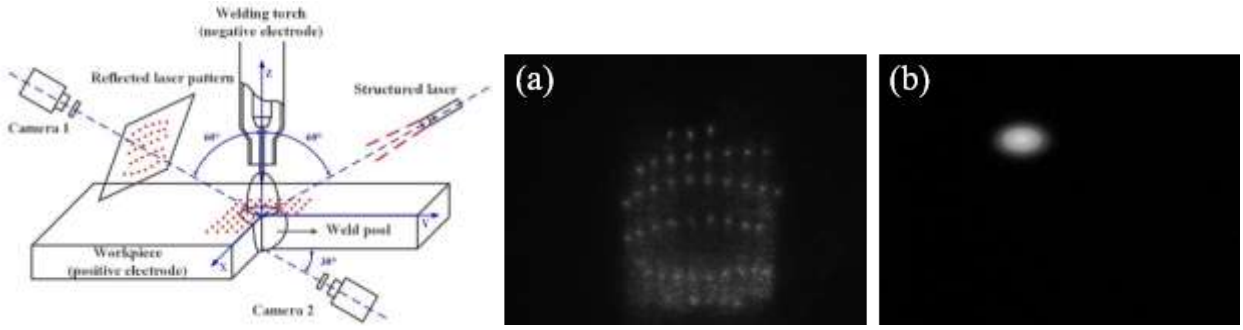
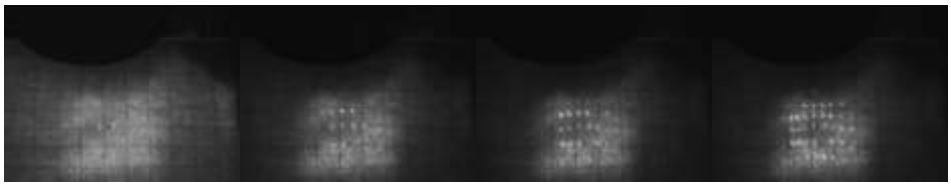
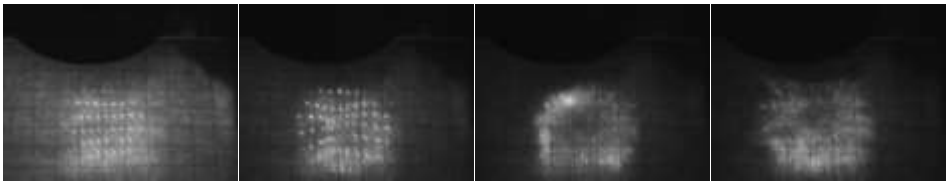


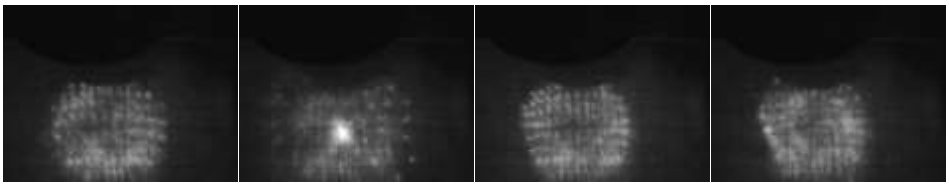
Fig. 20 Active Pool Oscillation Imaging System. Left: system principle; middle reflection image; right: image for penetration label [2].



(250^{th} ms) (500^{th} ms) (750^{th} ms) (1000^{th} ms)



(1250^{th} ms) (1500^{th} ms) (1750^{th} ms) (2000^{th} ms)



(2250^{th} ms) (2500^{th} ms) (2750^{th} ms) (3000^{th} ms)

Fig. 21 Active weld pool reflection images during a welding. The weld pool and penetration increase as the time increases. [104, 149]

In summary, specific images from the weld pool provide specific aspects of the weld pool whose importance/relevance is supported by physical mechanism. As such, although the general visual information is reduced, other features are enhanced. However, as they may still be complex, deep learning provides a meaningful tool to utilize these specific information.

Table 3: Other raw information sources used in deep learning.

Ref. #	Author	Year	Title	Raw Information
5	Oh S	2020	Cross-Section Bead Image Prediction in Laser Keyhole Welding of AISI Steel Using Deep Learning Architectures	Process input images
30	Jin C	2020	Prediction Model for Back-Bead Monitoring During Gas Metal Arc Welding Using Supervised Deep Learning	Welding parameter waveform
62	Ren WJ	2018.9	Seam penetration recognition for GTAW using Convolutional neural network based on time-frequency image of arc sound	Arc sound waveform
44	Ren WJ	2021.2	A Novel Convolutional Neural Network Based on Timex2013:Frequency Spectrogram of Arc Sound and Its Application on GTAW Penetration Classification	Arc sound waveform
19	Rohe, M	2021.12	Detecting Process Anomalies in the GMAW Process by Acoustic Sensing with a Convolutional Neural Network (CNN) for Classification	Acoustic waveform
81	Wu J	2022.9	Penetration Recognition in GTAW Welding Based on Time and Spectrum Images of Arc Sound Using Deep Learning Method	Arc sound waveform
84	Gao YF	2022.12	Weld penetration identification with deep learning method based on auditory spectrum images of arc sounds	Arc sound waveform
86	Zhang Z	2022.7	Deep Learning Empowered Structural Health Monitoring and Damage Diagnostics for Structures with Weldment via Decoding Ultrasonic Guided Wave	Multiple guided Ultrasonic waves

6. Other Raw Information

In addition to weld pool images and images from weld pool, other raw information sources have also been tried to predict the penetration using a deep learning based approach. Table 3 is an up to date list. Similarly as for the review and analysis for weld pool images and images from the weld pool, works that do not directly use raw information as inputs and involve artificial efforts to propose and extract features are not included. In particular, some studies refer networks with multiple hidden layers to as deep learning networks but having extracted features as inputs as deep learning. Such studies are not included. Using support vector, forest tree, decision tree and other machine learning approaches requiring crafted features are also not included.

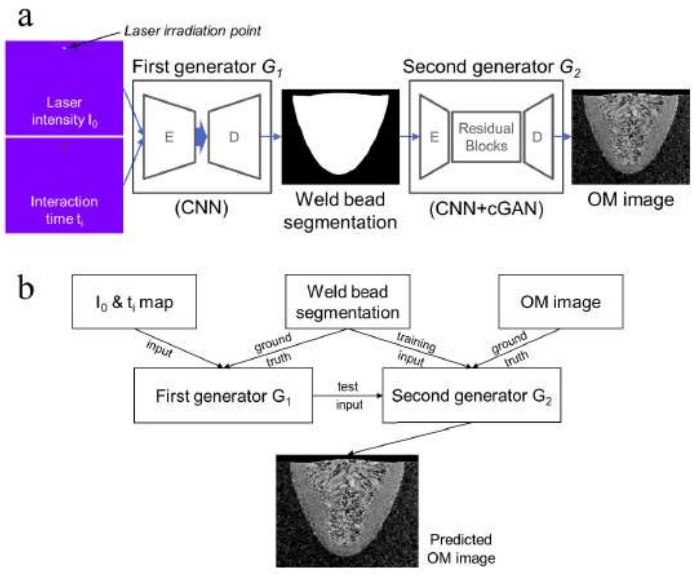


Fig. 22 A deep learning base approach to predict OM image from process parameters. (a) Overview; (b) data flow [5].

6.1 Welding Process Images

Fig. 22 [5] shows an approach to predict the metallography, that is, the optical microscopic (OM) image, of the cross section of laser weld from two input maps/images, including the laser intensity map I_0 and the interaction map of the laser welding process. There are ground trues of the OM images for each set of input parameters including the laser intensity and laser-material interaction time while other welding conditions are kept unchanged. This is not exactly the same as the deep learning models we have introduced so far that use high-dimension complex inputs to predict low-dimension simpler outputs. What to predict is also complex and high-dimensional. Of course, this OM image to be predicted includes the penetration state and additional outcomes from the laser welding process.

Fig. 22 shows two generators. The first generator is to generate the profile of the weld, or the weld bead segmentation in (a), from the two maps (I_0, t_1). The data in box "Weld bead segmentation" in (b) provides the ground true for the weld profile for the used (I_0, t_1). If the generated profiles do not match with the corresponding ground truth ones, the generator (parameters of the network) will be adjusted. For the second generator, it generates the OM image from the profile. If the generated OM images do not match with the ground true ones, the generator will be adjusted.

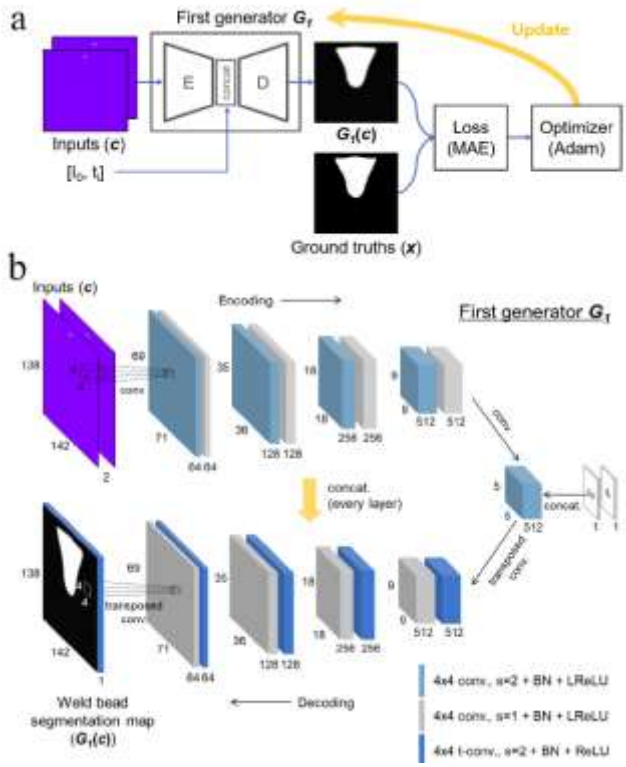


Fig. 23 Encoder-encoder formed generator using generator 1 as example [5].

Fig. 23 shows the principle of the generators, that is similar for both generators, using generator 1 as example. The inputs are two channels of 138X142 images and the output is a binary image of the same dimension – 138X142. The encoder is a CNN without fully connected network so it conducts convolution, reduces dimension and increase channels. The decoder is an inverse of the encoder in the structure but the weights are subject to independent tuning from their counterparts in the encoder. It conducts

transposed convolution, increases dimensions, and reduces the channels. The encoder and decoder are jointly trained. Generator 2 has similar structure except for the input changes from two channels of images to one binary image and the output changes from one binary image to a grey-level image (OM image). It is trained similarly as generator 1.

In the majority of the works analyzed, the inputs are in the form of images and the output is low-dimensional. CNN based deep learning models fit well for this kind of flow of data. In the work demonstrated in Fig. 22 and Fig. 23, an encoder-decoder structure fits with encode to reduce data dimension and decoder to increase the dimension. The dimensions of the inputs and outputs can be flexible. However, in both cases, the raw data are not changed.

6.2 Waveforms

The work by Jin C [30] demonstrates another way to use deep learning. This work is to classify the weld into incomplete and complete penetration while they refer the incomplete penetration to as "Without back bead generation (0)" and the complete penetration to as "With back-bead generation (1)". The raw information they used is the waveform of the welding current an the process metal arc welding (GMAW). However, they did not directly use the waveform as the input but converted the waveform into image of scalogram using the popular Morlet wavelet transform (MWT). If any raw information was lost in the conversion, such conversion would be considered a hand-craft operation and would be out of the scope of this review paper for deep learning. We note that a Fourier transform does not lose any raw information from its time-domain waveform as an inverse Fourier transform will produce exactly the same time-domain waveform. As an improvement of the Fourier Transform for short period of signal that recedes forward to allow separate control on the time and frequency ranges, the MWT also does not lose any raw information. After converting to image, a CNN can be used to process the fully preserved raw information in the waveform although the presentation of the raw information has been changed.

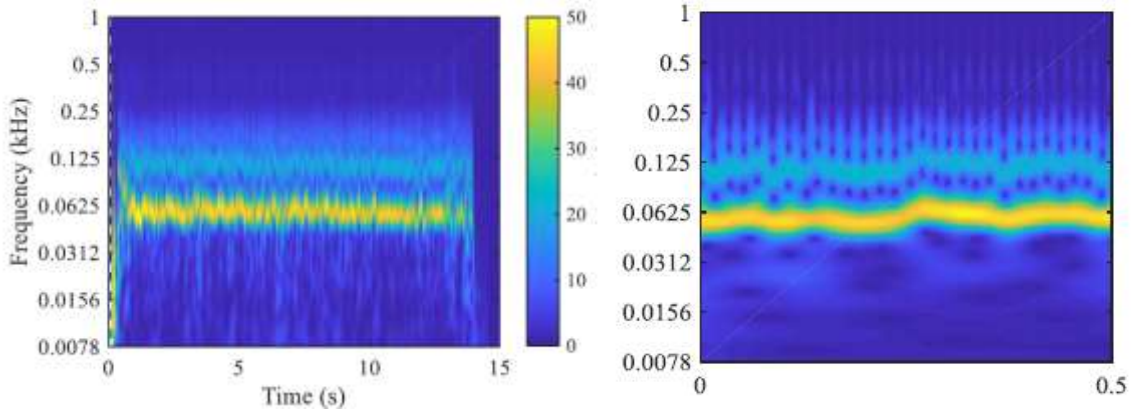


Fig. 24 Scalograms transformed from current waveform [30]. Left: from an experiment; right: from a short period of time.

The left of Fig. 24 shows a scalogram for a welding experiment. For anytime, the color along the corresponding vertical line changes as the frequency changes. The color is the amplitude of the component at the particular frequency. As such, this is a 3D matrix data showing the change of the frequency behavior of the waveform with the time. For real-time prediction, the time range is reduced to 0.5 second in [30] resulting in an image on the right in Fig. 24. The penetration state corresponding to this

period will be the label. The CNN model with the image in the right as input is supposed to predict the label. The CNN model is trained to predict all the labels well using their corresponding images exemplified in Fig. 24 right.

We have to note that the prediction accuracies in [5] using laser process inputs as represented by the two images are 89.0% for penetration depth and 93.6% for weld bead area. The classification accuracies in [30] using the scalogram was 95.7% for cases with back-bead generation and 91.2% for cases without the back-bead generation. Overall, the accuracy appears to be lower than those using previously reviewed raw information sources, weld pool images or images from weld pool. This appears to be well understandable as they use just process inputs (including welding current waveform [30]) rather than process feedback.

Arc sound is another kind of waveform as raw information sources that have been studied as inputs of deep learning models for weld penetration prediction. We have seen that Jin C [30] converted the current waveforms into images. For images based raw information, images at a time already provide complex information that may already be sufficient for certain application; next level of information from their dynamic evolution may not always be needed and are often ignored in studies we have analyzed. However, for waveforms, they provide information only through dynamic evolutions. Otherwise, their instant values do not provide features for “waveforms”. Hence, a time period is needed to provide “waveform” and its possible features. As the features of the waveform are often provided in frequency domain, the time-domain waveform in a time period is needed to conduct frequency features such as the spectrum. This changes the time-domain waveform into frequency-domain “waveform”. When different segments of waveforms are used, multiple frequency-domain waveforms are resulted. Per their time order, these waveforms form an image with its row and column to present frequency and time. This is the time-frequency image. The scalograms in Fig. 24 are a kind of time-frequency images. Time-frequency images can be calculated using different methods. The scalograms and spectrogram are calculated from waveforms using MWT and short-term Fourier transform, respectively. A similar method that uses short time Fourier transform to convert the waveform of the arc sound into image as the input of a deep learning model has also been reported [84].

Ren WJ [62] uses the arc sound waveform in the peak current period to predict the penetration. Such waveform results in the arc sound spectrum. The waveforms from the current and past peak current periods form colored images. Such a colored image is the spectrogram. A 3D data matrix provides the same information as the colored image and is thus another form of the spectrogram. In this way, the image of spectrogram is obtained and represented by a 3D data matrix. For an input in the form of image, a popular deep learning model is the CNN that is also the one used by Ren WJ [44, 62]. They obtained a high classification accuracy of 98.2% for GTAW of aluminum. This appears to be a high accuracy. This implies (1) the arc sound has sufficient raw information for the penetration for this application – pulsed GTAW of aluminum; (2) the deep learning successfully extracted the critical features from the arc sound.

Rohe M [19] studied the possibility to detect the deviation, of the flow rate of the shield gas during GMAW, from the nominal one. The process sound is used as the raw information source and a CNN is trained as the classifier. The study is meaningful as the flow rate of the shield gas affects the process stability. The process stability affects the weld quality through the metal transfer process whose regularity and mode may be detected from the sound. Rohe [19] used a microphone with a fixed distance to the torch/arc to collect the sound from the welding process, resulting in the waveform of the process sound. Again, the waveforms are converted into images of spectrograms as the inputs of the CNN. The CNN classifies the

flow rate of the shield gas and each class of flow rate corresponds to a class of weld attributes: occurrences of the pores and cavities. The overall classification accuracy is 84%.

Wu J [81] uses the arc sound waveform as the raw data to classify the penetration into three classes - full penetration, non-penetration and excessive penetration – for GTAW (Fig. 25). The sound waveform on the left is converted into two types of images, time images and spectrum images, using Adobe Audition software. They claim that human observation can distinguish the penetration class from the waveform (Fig. 26). As such, the needed raw information should be contained in the waveform. "VGG16, AlexNet, and custom convolutional neural network (CNN) were used to extract image features, and softmax was used to classify images for penetration recognition" [81].

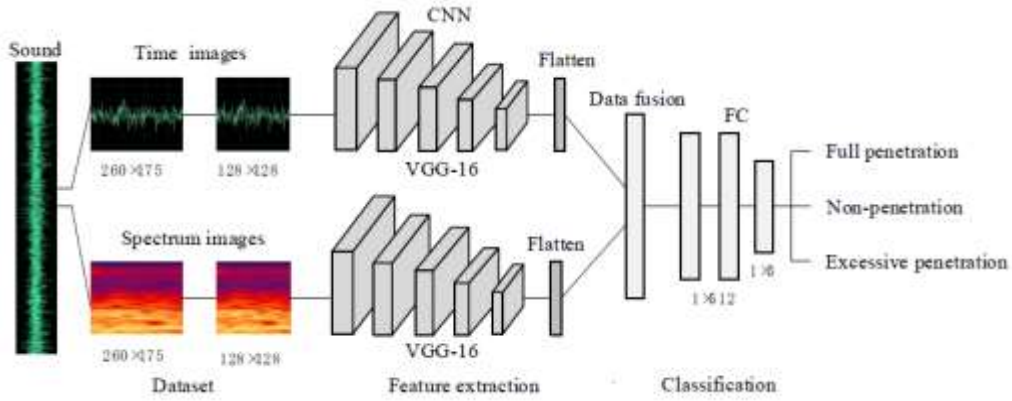


Fig. 25 Classification of weld penetration from arc sound waveform [81].

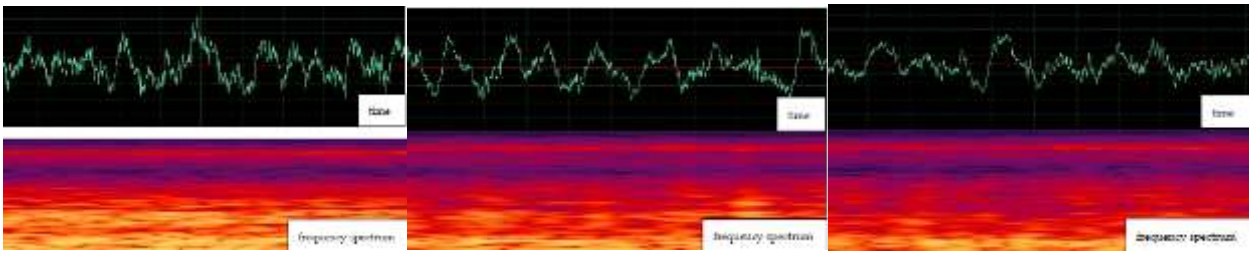


Fig. 26 Correlation between waveform (waveform images) and penetration class [81]. From left to right: waveforms corresponding to full penetration, non-penetration and excessive penetration.

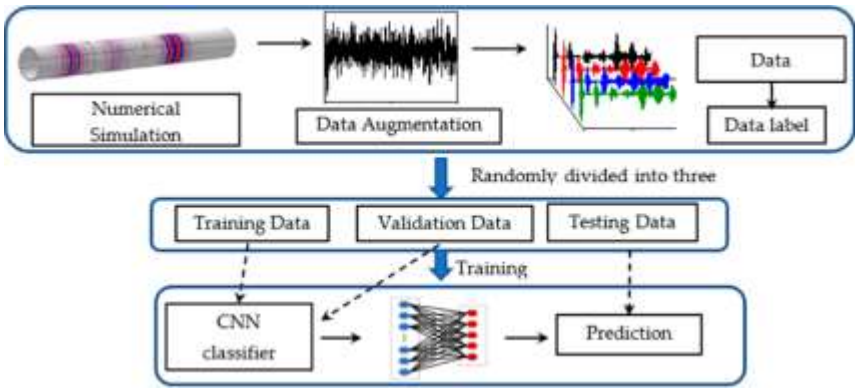


Fig. 27 Damage classification of welded structure from guided ultrasonic waves [86].

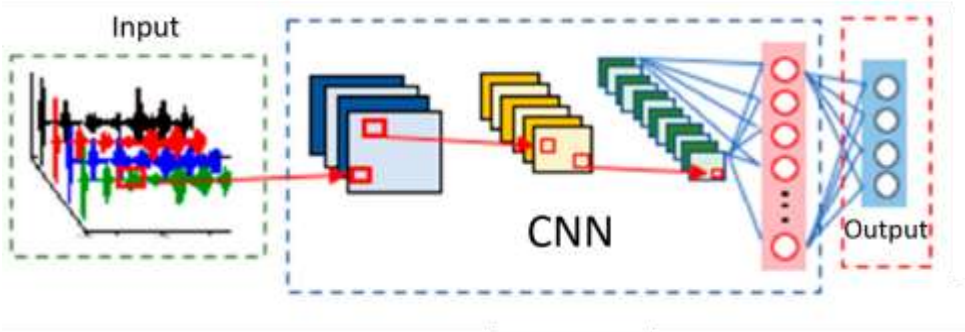


Fig. 28 Formation of the input image by four waves [86].

Zhang Z [86] used four guided ultrasonic waves to form an image as the input of a CNN to monitor the health of welded structures and diagnose possible damages. The health conditions/damages are classified into 16 damage states. As can be seen in Fig. 27, the dataset for training is generated using numerical simulations that each time assume a specific damage state (label) in the specific structure being studied to calculate the guided ultrasonic waves (input of the CNN). Fig. 28 shows the formation of the input image by four waves. After training, the CNN can take the actual guided ultrasonic waves to diagnose if there is a damage and where is the damage (what is the class of the damage in the structure).

In summary, other raw information sources not specific from the weld pool have also be used as the raw information of deep learning models. The accuracies are diverse per the physical correlation between the raw information with the penetration state or the weld quality concerned. If the raw information is not in the form of images, conversion has been conducted to produce its equivalent image form. The most popular deep learning model structure (CNN) that takes images as the inputs can thus be used. If the output is also complex, a decoder has been included into the model [5]. As such, different forms of raw information and concerned outputs can be addressed using a deep learning approach as long as the raw information is not simple. Otherwise, if the raw information is simple and low-dimensional, the deep learning approach would not be needed.

7. Multiple Raw Information Sources

As we have analyzed, one of the most critical requirements for deep learning is the sufficiency of the raw information. This is because deep learning provides possibility to extract the critical information from complex phenomena automatically. That is, complex information may be effectively utilized. However, this is not true for conventional methods that require us to simplify the raw information to utilizable features. As such, complexity and redundancy of the raw information may be less concerned than the sufficiency. As multiple raw information sources tend to help enhance the sufficiency, different raw information sources have been used together to predict the penetration by deep learning models. However, it still needs to point up that using multiple information sources should be justified for the specific application for its benefits against the increased cost due to hardware associated with multiple sensors, accessibility, and increased need for training data.

Table 4: Multiple raw information sources used in deep learning

Ref. #	Author	Year	Title	Raw Information
14	Feng YH	2020.1	DeepWelding: A Deep Learning Enhanced Approach to GTAW Using Multisource Sensing Images	Different Visual Sources
46	Cai W	2021.10	Real-time laser keyhole welding penetration state monitoring based on adaptive fusion images using convolutional neural networks	Combined Visual Images
85	Kang SH	2022.11	Deep learning-based penetration depth prediction in Al/Cu laser welding using spectrometer signal and CCD image	Visual mage, reflection spectrum
57	Jiang R	2021.9	Prediction of penetration based on infrared thermal and visual images during pulsed GTAW process	Visual and infrared
11	Wu D	2020	Visual-Acoustic Penetration Recognition in Variable Polarity Plasma Arc Welding Process Using Hybrid Deep Learning Approach	Visual and Sound

The work done by Feng YH [14] is among the earliest that used deep learning to predict the penetration and probably the first to use raw information from multiple sources in deep learning-based approach for weld penetration monitoring. There are three kinds of images: (1) the active vision is to apply an illustration laser that is brighter than the arc radiation intensity in the wavelength band of the imager; (2) the pass vision is to directly view the weld pool and arc; and (3) the reverse electrode image (REI) [149] is that of the reflection of the tungsten electrode from the weld pool surface. Each of them has been proven to provide information on penetration. In particular, the active vision gives the weld pool boundary clearly that has been used to predict the penetration [153]. The passive vision gives the weld pool image we have discussed first in Section 4 of this paper. The REI is an invention from the authors that has been proven to be capable of predicting penetration. However, such images may not always be ideal. Hence, for all images and all kinds, they are denoised first and judged by a deep learning-based Image Selector to decide on the image quality for acceptability. Accepted denoised images are input into the respective deep learning networks for Image Classification (penetration classification). The same accepted image is used as the input for five parallel deep learning models, AlexNet, DenseNet, ResNet, VGGNet, and CapsNet. For each kind of deep learning model for example AlexNet, it varies for different kind of weld pool image, i.e., active vision, passive vision and REI. For each kind of image, there are five predictions from the five models. Plurality Vote selects the majority that is equal to or greater than 3. As such, for each kind of image, there is a prediction. The final prediction is the majority of three predictions from three kinds of images. In case one kind of image is dropped, there may be two different equal votes. In this case, the vote from the kind of image with greater score may be used as the final prediction. As such, a comprehensive network frame is established to use deep learning to detect the penetration from multiple raw image sources. Without deep learning, handcrafting features from these images may be very difficult.

In the above work [14], the weld pool is imaged in three different ways. In [13], the weld pool is imaged from different views. It may also be considered to use multiple raw information. In [78], the weld pool image from each time is processed by the CNN to generate an intermediate result and the LSTM-CNN outputs the final result based on the immediate results from the present time and previous times. In high energy beam welding such as laser keyhole welding, the metal vapor is rejected from the work-piece. In such processes, the metal vapor is highly dynamic/fluctuating and is not believed to be fundamentally relevant to the weld penetration. The critical information in the image is the keyhole and weld pool. The presence of the metal vapor affects a clear view of the keyhole and weld pool. As such, in the ten images taken continuously by a high speed camera, five images with less vapor can provide better weld pool views. However, vapors still exist in these five images. As the vapors are highly dynamic, they appear differently

in these five images while the weld pool changes much less significantly. They can thus be fused relatively straightforward to form a fusion image. The fusion image can then be used as a single input into a CNN to classify the penetration state. This demonstrates another way to utilize multiple information sources.

While we have discussed using multiple information sources of the same kind, studies have used different kinds of raw information sources to predict the penetration based on deep learning. In [57], visual and infrared images have been used together. As we mentioned, the weld pool is where complex welding phenomena originate. As the weld pool is "red interface" that is not most directly determined by the penetration, missing phenomena from the weld pool reduces the sufficiency of the raw information. As we do not know what phenomena are critical and deep learning is capable of extracting the most relevant features, more complete raw information is helpful. Using the two most important weld pool information, visual and thermal, thus makes good sense. To effectively utilize the information from two different resources, Jiang R [57] developed a dual-input Faster R-CNN (region-convolutional neural network) model. A R-CNN has two stages with the first to identify a subset of regions in the image which may contain the object of interest and the second to classify the object in each region [150].

We have reviewed and analyzed works that use either weld pool images or acoustic/sound waveforms as the input for a deep learning model to predict the penetration. Both of them have achieved reasonable accuracy. We expect that using them together as the raw input would produce even better results. Fig. 29 shows how the weld pool image and acoustic waveform are used together [11]. As can be seen from Fig. 29, a waveform is still converted into image. Two parallel CNNs are used to extract features from the weld pool image and the waveform converted image respectively. The features are fused together as the input of a fully connected neural network to class the penetration. A very high classification accuracy, 98.18%, was achieved.

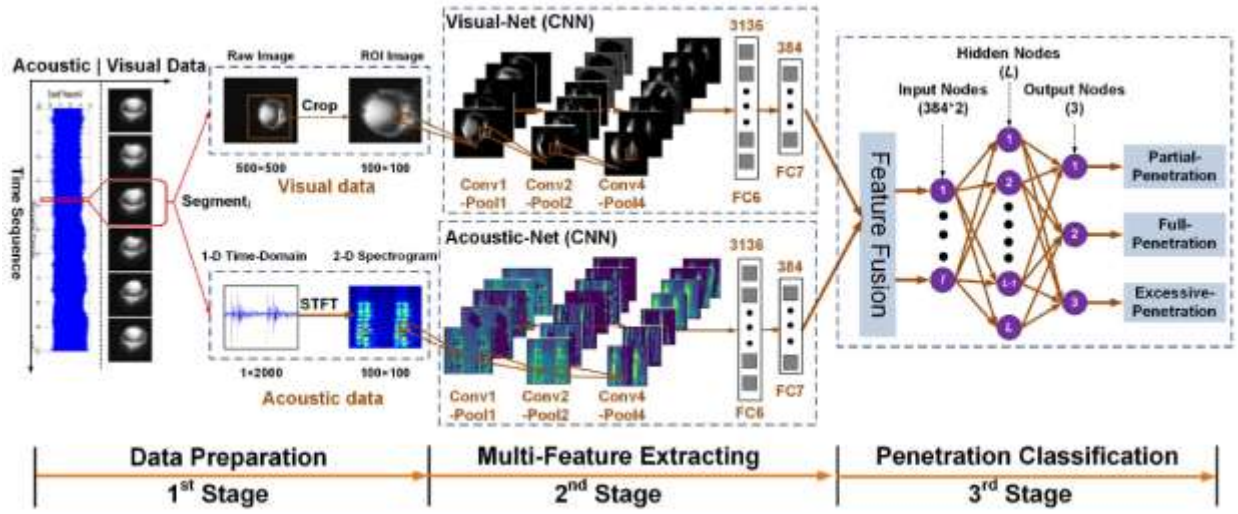


Fig. 29 Using weld pool image and acoustic waveform to classify the penetration based on deep learning [11].

In [85], Kang uses the coaxial image of the keyhole and the waveform of the intensity of laser reflection at a particular wavelength as the inputs to train a deep learning model to predict the penetration depth in overlap welding on an Al/Cu stack using a laser keyhole process. For the training dataset, the label is obtained using an optical coherence tomography (OCT) that can give the keyhole depth.

8. Basic Models and Approaches

We have discussed various kinds/categories of raw information sources proposed in literature for being used to estimate weld penetration using different deep learning models. For each particular application, a selected subset of them, single or multiple, is used. Ideally, this subset is minimally necessary in assuring the sufficiency of the raw information in addition to being realistically implementable and cost effective. The determination and selection of this subset is an art and good understanding of the physics of the subject welding process as well as that of the sensors is critical. However, due to the complexity of the process, the sufficiency is only assumed but not guaranteed.

With the assumed sufficiency of the raw information, an appropriate deep learning model is selected to use the selected raw information (or its pre-processed surrogate) as its input to calculate the weld penetration as the output. The difference with conventional modeling lies in that a deep learning model directly processes the raw information rather than its representative, i.e., the extracted features. Instead, the feature extraction is no longer a separate process from the model fitting but is combined with it. This combination allows the feature selection and corresponding model fitting be automated together. The combined automated process allows us to find the best representative of the raw information and its best corresponding model through automated computation, which is not achievable of the features are hand-crafted as in conventional modeling.

Each deep learning model may be considered a cascade of featuring and fitting. Both of them are parameterized so that the deep learning model is parameterized. Training a deep learning model is to optimize its corresponding set of model parameters. This becomes the same as conventional model fitting that finds optimal values of the model parameters to minimize the difference between the model predictions and the used labels/observations. The key difference between deep learning models and conventional modeling approaches is that deep learning models are designed to extract relevant information from raw data, theoretically maximizing accuracy. In contrast, conventional modeling approaches require human involvement to craft features, which can limit the model's ability to achieve optimal accuracy depending on the effectiveness of the human efforts/experiences.

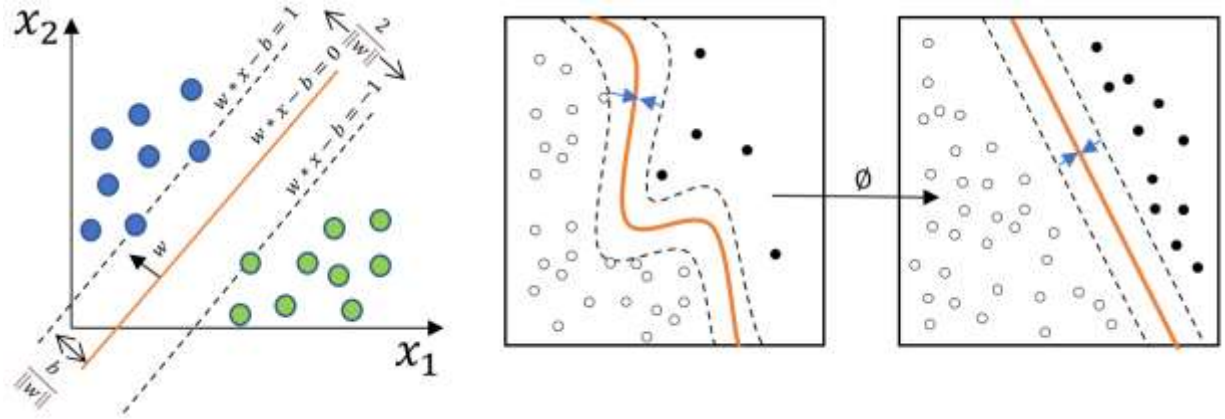
The unique ability of deep learning models is primarily due to both the complexity of the model and its particular structure/mechanism. The complexity assures there are sufficient parameters that can be adjusted so that a model structure may possibly provide versatile, ideally universal, abilities for various natures/dynamics/mappings of the underlying processes being modeled. To provide a good possibility, the model structure must be effective to capture most, or many, of common natures in various processes.

8.1 Classical Machine Learning

Deep learning is part and latest development of machine learning that can be characterized by “finding from data”, rather than analysis of the mechanism of the underlying process/phenomena. The Least Squares, either linear or nonlinear, may be considered the oldest method for “finding from data”. However, it relies on the availability of a model structure that has been obtained from mechanism analysis and is not to make findings primarily from data. For machine learning, the role of mechanism analysis playing in making effective findings reduces. Dependence on the mechanism is reduced by more intensive use of the data to try, exhaust within certain limits, different models (which may also be suggested through mechanism analysis that is labor intensive and the success depends on effectiveness of the efforts) and use the training data set to find the best. To this end, the data must be processed extensively, of

course by computers/machines. As such the learning is made/achieved primarily by a machine and was termed as machine learning [151].

Classical machine learning methods that have been used in penetration monitoring include support vector machine (SVM) [35], decision tree [27], random forest [17, 43], bagging trees [37], and neural networks (NNs) [153] etc. Their slight difference with deep learning lies in that the number of variables that may be adjusted to alter the output of classical machine learning models is relatively small so that their learning/training is relatively easy to converge and that the required number of data is relatively small. Of course the ability to universally model various processes is also relatively limited. SVM and NN are among popular classical machine learning approaches.



(a) Linear classification (b) Non-linear transformation using kernel functions

Fig. 30 SVM ideas

SVM is a supervised learning model that is specifically for classification and regression analysis. Fig. 30(a) shows an example using a two-dimensional feature $\boldsymbol{x} = (x_1, x_2)^T$. The data set used for training, $(\boldsymbol{x}_i, y_i)'s\ (i = 1, \dots, n)$, can be illustrated by the dots in the feature space with blue for $y = 1$ and green for $y = -1$. Classification is to use this data set to find a hyperplane defined by $\boldsymbol{wx} - b = 0$, where $\boldsymbol{w}$ is a 2×1 row vector and b is a scalar, to classify a future given $\boldsymbol{x}$ to $y = 1$ if $\boldsymbol{wx} - b > 0$ or $y = -1$ if $\boldsymbol{wx} - b < 0$. This can be done if the dots in training data set are fully separatable by a hyperplane as shown in Fig. 30(a). If the data set is fully separatable, there must be an infinite number of such hyperplanes. For such a hyperplane $\#j$, each $\boldsymbol{x}_i$ has a distance d_{ij} to it and $d_j = \min_{1 \le i \le n} d_{ij}$ is the separation margin for this hyperplane. The hyperplane with $\max_j d_j$ that achieved the maximal separation margin from the data set is the learned support vector classifier and we are returned with its parameters $(\boldsymbol{w}, b)$. It can be proven that its separation margin is $2/\|\boldsymbol{w}\|$.

If the data set is not fully separable, for any hyperplane $\#j$ there must be dots falling on the wrong side. However, if their distances to hyperplane $\#j$ is small enough, for example not great than a given tolerance $\varepsilon > 0$, we can forgive and not count them in calculating d_j so that $d_j = \min_{1 \le i \le n, d_{ij} > \varepsilon} d_{ij}$. This introduces the soft margin while it is the hard margin when $\varepsilon = 0$. With the concept of the soft margin, this support vector approach can be easily extended to regression where y, to be predicted from $\boldsymbol{x}$ is no longer Boolean but a real number. As such, in the data set $(\boldsymbol{x}_i, y_i)'s\ (i = 1, \dots, n)$, $y_i's$ are the target values for

the hyperplane to approximate. We will still minimize the maximal distance $2/\|\boldsymbol{w}\| = \max_j d_j$ but any hyperplane $\boldsymbol{w}_j\boldsymbol{x}_i - b_j$ with any $d_{ij} = \left|y_i - w_j x_i - b_j\right| > \varepsilon$ will be disqualified. Hence, the maximal fitting error is bounded by $\varepsilon > 0$. In our discussion, $\boldsymbol{wx} - b$ is linear with the original feature space $\boldsymbol{x}$**.** Fig. 30(b) shows a case that is not linearly separable but nonlinearly separatable. In this case, kernels functions may be proposed to make it linearly separable. Kernels can also be used for regression problems. This approach has been used in [35] to predict the backside weld bead width using weld pool surface feature vector consisting of the width, length, and convexity of the weld pool surface as the characteristic parameters of the weld pool surface [134] in GTAW. It has been also used to solve weld penetration monitoring problem in [134].

Skilled welders may maintain the desired weld penetration by adjusting welding parameters based on the observation of the weld pool. However, it is unclear how they make their judgement in estimating the penetration. A common perception is that the width and length of the weld pool may contain major feedback information that welders may have from their observation of the weld pool. However, studies found that the width of the weld pool is not very sensitive to the change in the weld penetration [152]. To comprehensively understand the human welders' responses, the concept of geometrical appearance has been proposed to characterize the perception the welders may have from their observation. It has been proposed to characterize the geometrical appearance using the length of the weld pool and nine rear angles measured at nine positions [153]. It was found that the rear of the weld pool boundary is much more sensitive than the front of the weld pool. The distance from one position to next (for the nine positions from the pool rear) thus increases following a specified schedule. Experiments under different welding conditions and varying welding parameters have been conducted to generate weld pool images and back-side bead width. The total number of the samples for training $N = 7350$. Neural networks (NNs) with one hidden layer have been trained using the above proposed features as well as other comparative features. The number of the inputs, n_1, thus varies from NN model to NN model with different inputs (corresponding to different sets of comparative features). $n_3 = 1$ for all models is the number of elements in the output layer that outputs the back-side bead width. The number of the elements in the hidden layer is determined by $n_2 = \frac{N}{10(n_1+n_3)} \sim \frac{N}{5(n_1+n_3)}$ per [154]. For the NN model trained using the length and 9 rear angles so that $n_1 = 10$, $n_2 = 67 \sim 134$ and $n_2 = 100$ were used. It was concluded "the geometrical appearance of the weld pool, specified by the length and rear angles, contains sufficient information on the weld penetration, judged by the 1 mm critical measurement of the modeling error. However, pure size or pure shape parameters only contain rough information on the weld penetration." The use of NNs assures that possible nonlinear relationships between the proposed features and weld penetration be adequately learned.

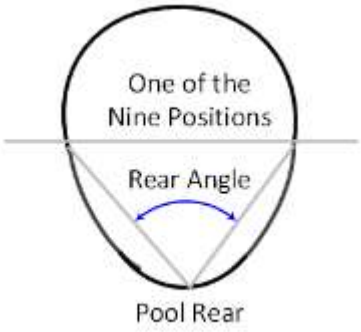


Fig. 31 Pool rear angles/features as the inputs of neural network based weld penetration monitoring using weld pool geometrical appearance [153].

In conventional model fitting, model structures are given and the data are only used to estimate model parameters primarily. One may consider that classical machine learning models do not require particular model structures and have certain abilities to find best relationships hiding in the data. In particular NNs can be considered universal model structures which can provide different relationships by changing the parameters. The SVMs use the worst-case and tolerance as the basis to maximally separate data. As such, the data is used more effectively and classical machine learning may also be considered data-driven approaches. However, they use pre-selected features as their inputs and the features are not automatically selected from the data. There are thus chances the features are not optimally selected and they may not be complex enough to sufficiently capture relevant information in the raw data. As data are not used to optimally select features, the models are relatively less comprehensive requiring less data to train than deep learning networks. As such, when the data size is relatively small and the features can be reasonably appropriately selected, classical machine learning models may outperform deep learning models. However, with the massively increased dataset, the performance of deep learning models will increase significantly with dataset concurrently, while traditional models remain relatively stable as shown in Fig. 32 [155]. As the requirement to train an effective deep learning model is mainly the computation power and sufficient dataset, accompanied by the development of the hardware and algorithms, the only major burden left is that how to acquire an ideal dataset.

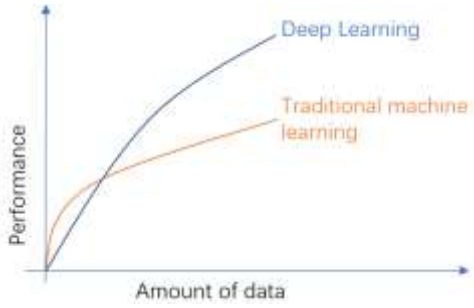
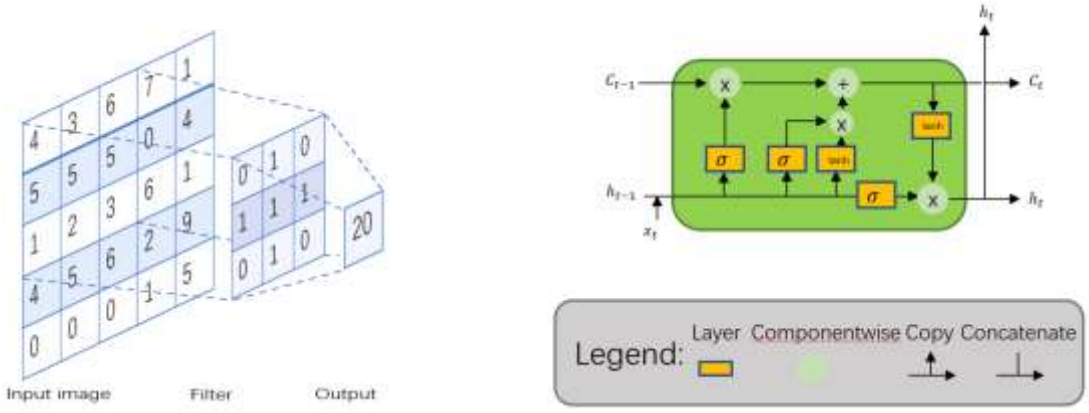


Fig. 32 Deep learning compared with the classical machine learning [155].

8.2 Deep Learning Models

Different machine learning models have abilities to extract relevant information from different categories of raw data sources. If the input data is given by images, CNN is an appropriate model form. Compared

with the classical image processing techniques [156], the convolution (Fig. 33(a)) layer in CNNs could replace the handcraft feature selection algorithm effectively. Multiple convolution process will give model the ability to percept the target features; moreover, it can generalize the target feature's pattern in the model to contest the overfitting problem [157]. If the input data is given by dynamic data, such as a series of images or other sequenced signals, the recurrent neural network (RNN) [158] will then be an appropriate model to process such data. With the unique time delay unit (Fig. 33(b)) and sequential connection from one to next, the long-range structural dependencies of the RNN can simulate time series information [158].



(a) Convolution principle [157] (b) LSTM Basic Cell [158]

Fig. 33 Fundamental elements in Deep Learning Models.

8.2A CNN model

A typical CNN model has multiple convolution and fully connected layers. To illustrate, we can take VGG16 [159] as an example. It contains 13 convolution and 3 fully connected layers. The input image $(224 \times 224 \times 3)$ is first being processed by two continuous convolution layers (both using $3 \times 3 \times 64$ kernels) and a max pooling layer. There thus be 1728+36864 parameters for optimization in the two continuous convolution layers as $(3 \times 3 \times 3) \times 64$=1728 and $(3 \times 3 \times 64) \times 64$=36864. The data outputted from this first section will be processed by next two continuous 3*3 convolution kernels with 128 channels followed by a max pooling layer (73728+147456 parameters here). In the next three sections, each has three continuous convolution layers, all followed by a max pooling layer. Their convolution kernels are $3 \times 3 \times 256$ (294912+589824+589824), $3 \times 3 \times 512$ (1179648+2359296$\times$2), and $3 \times 3 \times 512$ (2359296$\times$3) respectively. After the convolution process finishes, the feature vector as its output is fed into the three fully connected layers as $1 \times 1 \times 4096$, $1 \times 1 \times 4096$, $1 \times 1 \times 1000$, sequentially. During the model training process, all those convolution layers could be considered as for featuring while the fully connected layers perform the fitting. The entire model contains 138M parameters in total. However, despite such a dramatic number of parameters, they will be optimized based on the backpropagation algorithm. Compared with the conventional method, convolution process gives the model some sort of perception ability, which allows the model to transfer the attention on different aspects in the raw information. Accompanied with the iterative training process, the parameters that work best with the fitting procedure can be determined.

Li et al. used a CNN model to detect the penetration state from the weld pool reflection image [2]. The experimental system, images, and labeling are discussed in Fig. 20 and Fig. 21. The weld penetration continuously changes and is classified into six levels. 3550 raw images were collected after designed experiments conducted under various welding conditions. After augmentation (to be discussed further), totally 427,313 images were obtained. 75% of the data were used to train the CNN model whose input is 48X36 image. 90.70% of prediction accuracy has been achieved to classify the weld penetration into six levels [2]. A voting mechanism based on three continuous images increases the classification accuracy to 97.6%. All inaccurate predictions have the actual weld penetration in the borders of two adjacent levels. The accuracy is thus extraordinary as the actual weld penetration changes continuously and does occur sometime in the borders between two adjacent levels.

In such CNN models, the raw information which has been utilized is mainly the top side weld pool images. The raw information is complex representing complex phenomena but the corresponding labels/ground truth such as the back-side bead width can be separately measured in relatively straightforward ways, for example being obtained automatically from the processing of the back-side weld pool image [2]. Large training data may thus be available. However, in other applications such as weld seam/pool tracking, the ground truth/target objects position is not easy to obtain and is contained in the complex raw information/image. The purpose for deep learning is to extract the ground truth so that there are no separate, straightforward alternative ways. In such cases, the ground truth would have to be labeled manually. This limits the data size which could be generated for training. Therefore, to contest the model's overfitting problem, approaches are needed to either increase the training data size or reduce the training models' parameters size.

Limited dataset is only suitable for models with appropriate size. In most cases, welding images share similarities. Compared with natural images, welding images are far less dense in content and information. The image structure is fixed, and semantic information is relatively straightforward. For instance, the weld pool is always located at the center area of the image (provided that the camera is appropriately positioned as required) and is the brightest in the full view. As such, a lite network [160] structure may be sufficient that can be fitted with a small dataset without extensive semantic information.

Yu [161] used a U-net model to recognize the weld pool image in real-time. The U-net structure consists of a contracting path and an expansive path. In the contracting path, the image feeds into a 3×3 convolution process and repeats twice, every time followed by a rectified linear unit (ReLU) [162] and a 2×2 max pooling operation with stride 2 for downsampling and doubling the channels. In the expansive path, every step consists of an upsampling of the feature map that reduces the number of feature channels by half, a concatenation with the correspondingly cropped feature map from the contracting path, and two 3×3 convolutions each of which is followed by a ReLU. The network has 23 convolutional layers in total. The modified U-net structure has only 7.55 M parameters, while the original U-net structure has parameters as 28 M and standard FCN network has parameters around 57 M. With such a small network structure, 213 labeled data pairs were sufficient to finish the training process without overfitting occurrence. The training result reached 99% accuracy and testing result proves that a clear weld pool boundary has be obtained [161].

8.2B RNN model

Different with a CNN, RNN typically takes sequential or time series data as input [158]. It has been commonly used for ordinal problems like natural language processing (NLP), speech recognition, image captioning, etc. [163, 164]. Distinguished from their memory ability, they could utilize the information from prior inputs to influence current input and output. A fully connected RNN typically has three layers: input layer, hidden layer and output layer. Input layer takes the input to the neural network, processes it and passes it onto the hidden layer. There are multiple hidden layers linked together with a standardized activation function and different weights. Each hidden neuron maintains a unique state and uses this state to generate its own output and hidden weights. It passes the output to the output layer and hidden weights to the next hidden neuron. Fig. 34 shows a typical RNN structure, accompanied with the sequential input $x_{(1,2,\ldots,t)}$, the model will maintain a different hidden state as $h_{(1,2,\ldots,t)}$; each hidden neural will generate its own output $y_{(1,2,\ldots,t)}$ and hidden weight $w_{(1,2,\ldots,t)}$ and passes those parameters to their corresponding next position.

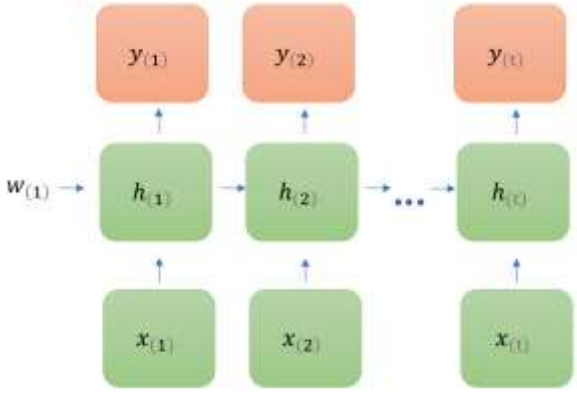


Fig. 34 A typical RNN model structure [158].

Chen has implemented a hybrid network model, as in Fig. 35, CNN-LSTM, to predict the penetration state during the welding process [51]. The penetration state has been divided into six classes as normal penetration, lack of fusion, sag depression, burn through, and misalignments (#1 and #2). In this model, the CNN layer is designed to extract the visual features from the weld pool images. The feature vector from the CNN's output layer concatenates with arc voltage, welding current, arc power, and arc sound to form a 19-dimensional feature vector. The LSTM network will then fuse the extracted 19-dimensional feature and predict the welding states. With the complement information other than the visual feature, the model can predict the penetration states effectively with high accuracy and robust enough [51].

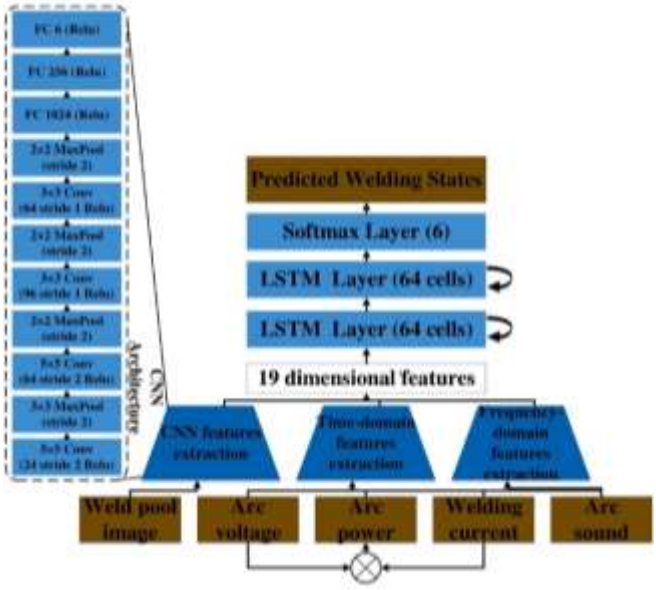


Fig. 35 CNN-LSTM Network model to fuse 19-dimensional feature [51].

Yu also introduced the LSTM structure in the weld penetration prediction task [78]. Different with the Chen's work, it analyzed that the welding process is a dynamic process. The heat input will accumulate over time so that the heat input in the previous time will influence the heat accumulation at the current stage which means the welding current applied a few seconds before will influence the current penetration state. This work more focuses on the visual features and extend the time dimension to ensure the sufficient information, as shown in the Fig. 36. The designed CNN-LSTM model will take 8 images as an input sequence. Each image will have a gap set as 32, i.e., the model will use image series $I(k)$, $I(k-32)$,..., $I(k-224)$ to predict the current backside bead width. With the sampling frequency as 60HZ, 224 images spans in 3.7 seconds which approximately reflect the settling time of the GTAW process [78]

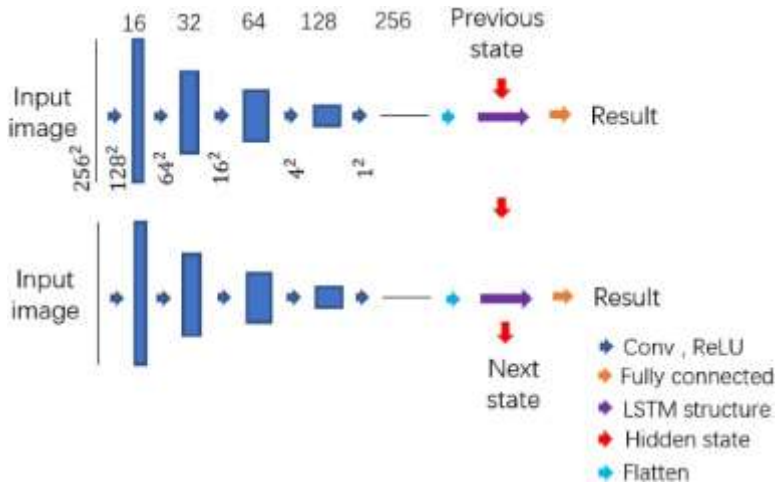


Fig. 36 Hybrid CNN-LSTM Network model to utilize time dimension information [78].

8.3 Other training approaches

All models discussed above only possible assure a theoretical effectiveness. In real applications, the training dataset needs to satisfy certain requirements in order to obtain good training results. The dataset may not be large enough to fit a model with a large parameter size. Also, the dataset may not be various/diverse enough so that the trained model may have poor performance at some specific situations. To contest such problems, this subsection discusses some training techniques which could be utilized to improve the training performance.

In most circumstances, the raw information after being collected will be fed into a machine learning model, directly or after a pre-processing, to obtain the desired output from the trained model. If the quality of the raw data is satisfactory, the data may be used directly. In most circumstances, the structure/dimension/form of raw data may not meet the requirement from the network's input layer hence a pre-processing needs to be performed before the training procedure. For example, most of the existing CNN's input layer requires the input image has a fixed size (e.g., 244 * 244). As such, the input image needs to be cropped or resized to satisfy this requirement from the input layer. Since this kind of requirement is "artificial" and may reduce the training accuracy of the network model, He [165] proposed a "spatial pyramid pooling" (SPP) structure which could generate a fixed length representation regardless of input image's size/scale. Fig. 37 shows the structure of the SPP layer, and by using pooling layers with different kernel sizes this design gives the model an ability to bypass the convolution layer but still can percept the information from the input data. Without size requirement from the convolution layer, extra artificial influence may be better avoided during the data preparation process. Wang QY pre-processed the raw information obtained during the experiment with the principal component analysis before the information is fed into the SVM model.

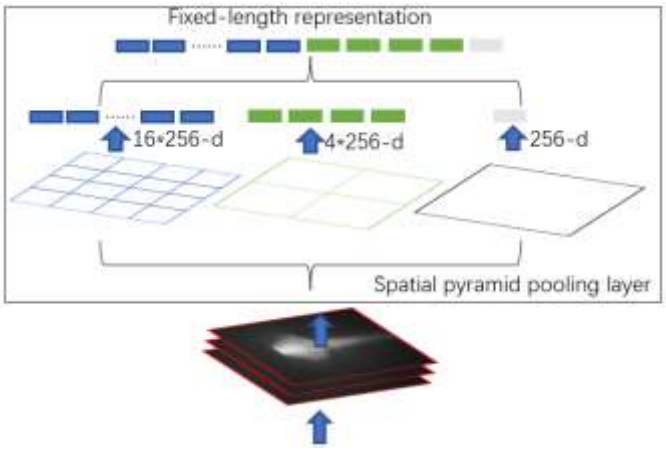


Fig. 37 A network structure with a spatial pyramid pooling layer. Here 256 is the filter number of the conv5 layer, and conv5 is the last convolutional layer [165].

Obtaining data from experiments is a very common practice. However, it only applies to the circumstance if the required data size was not huge. Since the deep learning model needs huge amount of data to contest the overfitting problem, obtaining all the data from experiments is very time consuming and may not be doable in certain applications. Data augmentation is a commonly accepted technique to artificially introduce non-essential information into the data to form new data to add into the dataset. Such information should be ignored by the model but may occur during data acquisition. One example is that

the orientation of objects in the image provided that the orientation is irrelevant to the output to be predicted. This improves the robustness of the trained model against irrelevant information that may appear and also increase the dataset without increasing the cost. By adding modified copies of already existing data or created synthetic data from existing data, data augmentation could increase the amount of data multiple times. As discussed previously, Li enlarged his dataset from 3550 raw images to 427,313 images after the augmentation process [2]. With the augmented dataset, the training model will achieve better performance in most scenarios, however, a massive dataset also means an extended training time. To reduce this time-consuming process, Zheng QH [166] proposed a two-level data augmentation method by replacing the augmented samples at each training iteration and proved that it can still improve the generalization ability of deep learning models and reduce the time cost of the training process. With this method, 1000 augmented images, as an example, 10 images will be added to the training dataset every time, with 100 iterations in total. In each iteration 10 new images will replace the previous 10 images to add into the used training dataset [166].

Another way to contest the insufficient data problem is referred to as Transfer Learning [167]. Per discussion above, deep learning model requires a much larger dataset to train because it has massive parameters to be fit. However, for a class of tasks, in most circumstances, the features in the raw information may be quite similar. This implies that the perception layer of the deep learning model, e.g., the convolution layer in the CNN model, may be quite similar and have similar abilities for similar tasks. Consider the case where a model, trained to detect the dog, swaps its perception layer with that in a model trained to detect the cat. In this case, after the swap, the model will only need a few training iterations to slightly adjust the parameters to achieve its previous accuracy. Since the training process has been reduced, the needed training data will also be reduced. With this ability, under similar application circumstances, transfer learning allows the model to migrate from the large set of existing labeled data (source domain) to small set of specific data (target domain) to train a model suitable for the target domain.

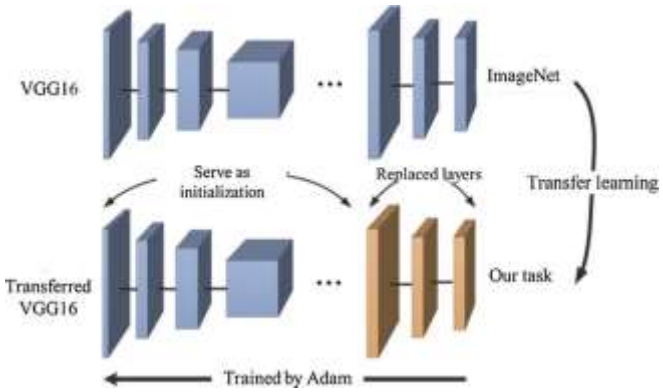


Fig. 38 Transfer learning with VGG16 network [47].

Jiao [12] pre-trained a ResNet model with the open access ImageNet data as the initial model. This model is then further trained with the specific application data. A good prediction accuracy of 96.35% has been achieved for the weld penetration [12]. Ma [47] used a VGG16 structure and trained this model with the ImageNet dataset. When the model was initiated from the scratch, the three fully connected layers of the model in the ImageNet dataset are replaced by three layers with different number of output classes to accommodate the application to classify into 4 classes. The newly constituted model is then trained with

limited laser stripe image database. Since the model's top-level parameters have been frozen, the size of the left parameters in the fully connected layer is relatively small. This could contest the overfitting problem with small datasets. The training process was demonstrated in Fig. 38.

9. Where We Are? Achievements and Issues

Through the analysis for up to data efforts [1-94], we can summarize their major achievements and identify their apparent issues.

9.1 Major Achievements:

Achievement 1 Automating the Modeling Process: Despite the relatively short period of time since 2019, deep learning has become the most studied method to predict penetration. The major reasons include: (1) The penetration occurs underneath the work-piece surface and is not directly measurable during manufacturing. The penetration monitoring problem as defined by the community is to use indirect information to derive the penetration state and applications where the penetration can be directly measured are not the subject of penetration monitoring studies; (2) As the information being used is indirect and typically very complex, the existence of the correlation between the weld penetration and observed phenomena is only assumed rather than theoretically guaranteed. Further, what are critical in the raw information for the weld penetration is difficult to know accurately and conventional methods "guess" different features that may represent the "unknown/unclear" critical information and then tried to fit the target output using the features calculated from the raw information. Guessing features, computing the features, and fitting model are separate processes and overall modeling process is a trail-and-error manual one requiring application specific algorithms to compute features and the algorithms changes with the different features proposed. Deep learning changes the overall modeling process from a manual one to an automated one to find the "best presentative/critical features". Such an automated process greatly improves the probability for better modeling and reduces the threshold for needed training/skills to solve penetration monitoring problems. In particular, conventional methods require solid understanding of the physical mechanism to guess reasonable features and high skills in the signal/image processing/computer vision to calculate the proposed features. Using deep learning, the requirements on these specific skills are reduced.

Achievement 2 Designing Deep Learning Models for the Variety of Raw Information Sources: Various kinds of information and multiple information sources have been used as the raw information. To accommodate the variety in the raw information, various techniques and network structures have been used/designed so the proposed raw information can be fed into a deep learning model. In particular, the model structures used include conventional machine learning models such as ANN and Support Vector [3, 21, 22, 25, 27, 35, 37, 39, 54, 55, 59, 60], CNN (2, 4, 5, 7, 9, 11-16, 19, 20, 24, 28, 33, 34, 39, 41-44, 46, 47, 49, 57, 61, 62, 67], RNN [20, 51, 78], Residual Network [9, 12, 16, 29], YOLOv4 [23, 26], deep belief network (DBN) [36], Generative [5, 14], Encoder [5], and specially designed networks for multiple information sources [11, 14, 46, 57]. The techniques used include Transfer learning [12, 40, 41, 47] and lite networks to reduce the needed labels and model fusion [52].

Achievement 3 Applying to Various Welding Applications: Various applications have been studied. All major fusion welding processes have been studied including gas tungsten arc welding GTAW [2, 4, 7, 12, 13, 24, 26, 28, 33, 37, 40-42, 51, 62], gas metal arc welding GMAW [8, 19, 22, 25, 27, 30, 47, 60], keyhole plasma arc welding PAW [11, 15, 16, 43], laser welding [3, 5, 20, 21, 23, 29, 34, 36, 39, 46, 49, 59, 61, 67].

Deep learning has also been used for process control [4, 67] and the development of digital twin for visualization to human [7].

9.2 Observed Issues from the State-of-the-art works

Issue 1 Lacking Assurance/Justification for the Sufficiency of the Raw Information: A key and fundamental assumption for the effectiveness of a trained deep learning model is that the raw information used is sufficient. Deep learning is just a tool to more effectively extract the critical information that has already been contained in the raw information rather than creating the critical information to generate the most effective features. The prediction accuracy from the trained model reported in all publications is based on the dataset used. The dataset must be representative for the intended application to trust the accuracy tested for the intended application. Otherwise, the sufficiency of the raw information in determining the weld penetration must be analyzed and justified. Unfortunately, out of all the 94 WoS publication records, only four of them include "sufficiency" or "sufficient" in the indexed fields. Further, out of these four publications, "sufficiency" or "sufficient" is related to the raw information [15, 78]. While [15] assumes the weld pool images contain sufficient information, only Yu [78] explicitly addressed the sufficiency issue and justified how serial weld pool images are needed to provide sufficient raw information. As such, the sufficiency is an issue that has been ignored. As it is the foundation for the effectiveness of deep learning based monitoring, all future studies should explicitly justify the sufficiency of the raw information used in determining the weld penetration from the observed phenomena.

Issue 2 Lacking Justification for Experiment Design: Studying the uncovered 94 publications shows that most studies do not specify their intended applications for what are the nominal conditions and what are uncontrollable perturbations that deviate the actual manufacturing conditions from the nominal ones causing the penetration to deviate from the desired. Experiment plan is specified without a sound justification for how perturbations are simulated in the experiments to reflect in the resultant dataset. The applicability, of the trained model using data from such imperfectly planned experiments, to intended application may thus be questionable.

Issue 3 Lacking Verification from Feedback Control of Weld Penetration: Most of the uncovered studies train deep learning models to predict the weld penetration. The accuracy is tested from the used dataset which was generated under artificially designed conditions. However, most studies do not discuss the actual perturbations and if such perturbations would present different phenomena that have not been covered in the training dataset. A method to overcome this is to test the trained deep learning model in control where the control algorithm adjusts the welding parameters. The resultant weld pool and welding process dynamics may be different from the experimental conditions used to generate the training dataset. The deep learning model needs to be capable of predicting under such real dynamic changes.

Issue 4 Lacking Justification for Using Multiple Information Sources: More information in particular more diverse information in general helps improve the sufficiency of the raw information. Theoretically, an effective deep learning model can be trained to effectively extract critical information despite the redundancy without overfitting. However, the model structure will become more complex and training such models not only requires greater but also more diverse datasets. This has been largely ignored in many of the existing studies. More critically, for manufacturing applications, it is desired that sensors are used as little as possible. Most sensors require careful installation and calibration. Increased sensors in general impose more constraints on the resultant welding systems and extra care is needed to use the

systems in manufacturing. As such, careful justification is needed to defend the absolute necessity and extra benefits for using complex and multiple sensors/information sources.

10 What Are the Needed Revolutionary Solutions?

Deep learning promises obtaining the most accurate prediction from given observables by combining and automating featuring and model fitting. However, it does not solve fundamental issues (1) if the labels needed to train the deep learning model are conveniently obtainable; and (2) if the raw information is sufficient. We first analyze the first fundamental issue.

10.1 Training Using Inaccurate Labels:

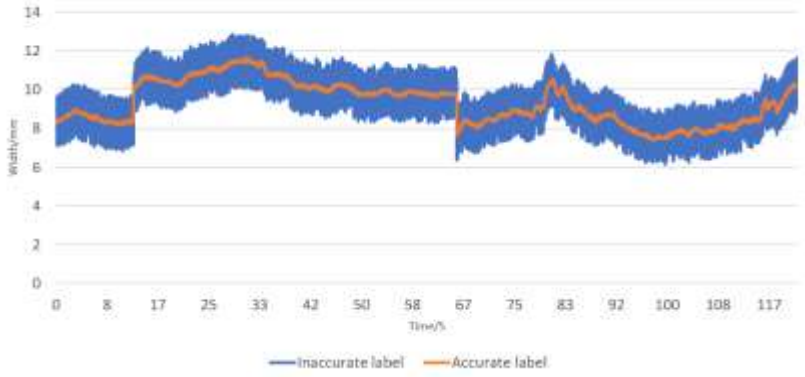
One of most promising observable phenomena for monitoring weld penetration is the images of the weld pool from its observable surface. As such, most of the observable phenomena that are used to monitor weld penetration are based on weld pool images in recent studies. In these studies, weld pool images are directly used as the inputs of deep learning models to predict the penetration. To train such deep learning models, weld penetration labels are needed. Unfortunately, the needed number of labels to train deep learning models is extraordinarily large. Further, the weld penetration occurs underneath the workpiece so that its measurement is difficult. Effectively using deep learning models that potentially can revolutionize intelligent welding manufacturing becomes challenging.

To solve this fundamental issue, in a recent study [168] we proposed to use inaccurate labels that can be conveniently obtained. Denote the set of the weld pool images (observable phenomena) as Ξ and the set of the penetration state labels as X. Deep learning is fit a model, that correlates the weld pool image to the penetration state, from data set (X, Ξ). As X is not easy to obtain, we proposed to use an alternative data set (X^0, Ξ) and the same deep learning training algorithm to train the deep learning model. Here X^0 is the easily obtained "alternative label" that is not accurate and is related to X by $X^0(k) = X(k) + \varepsilon(k)$ $(k = 1, \dots, N)$ where N is the number of samples in the data set. The only requirement for X^0 is that the approximation error ε has zero mean while its variance σ_ε^2 can be large, i.e., the error for X^0 can be large as long as the mean is zero. The size of the dataset N should also be appropriately large per the labeling error's variance in relation with the desired variance of the training error.

To test this proposed inaccurate labeling based training using real penetration data and deep learning model, a preliminary study has been conducted. Three experiments have been performed on mild steel workpiece with dimension 3mmX50mmX250mm using GTAW. The current in each experiment is fixed but the welding speed randomly changes so that the weld pool and penetration change. The weld pool is imaged by an HDR Xiris XVC-1100 to acquire Ξ and the back-side of the work-piece is imaged using a regular camera PointGray to calculate the accurate label X, both at 60HZ. The data from the beginning of the welding in each experiment are excluded from being included into the dataset. The welding time varies from experiment to experiment so that there are 769, 3168 and 3344 pairs of $(X(k), \Xi(k))$'s $(k = 1, \dots, 7281)$ in the dataset from experiment #1, #2, and #3 respectively with $X(k)$ calculated from the back-side image [78]. The accurate labels/back-side bead width of the weld are plotted together using the in Fig. 39. As can be seen, the penetration varies within an experiment and from experiment to experiment. From (X, Ξ), we can generate the corresponding inaccurately labeled dataset (X^0, Ξ) from $X^0(k) = X(k) + \varepsilon(k)$ with $\varepsilon(k)$ be a random number in [-1.4 mm, 1.4 mm] (-20 pixels, 20 pixels) which corresponds to a very large labeling error. The accurate and inaccurate labels are compared in Fig. 39(a). It can be seen that the errors for the inaccurate labels are extra-ordinary large (The variance of the labeling

error $(X - X^0)^T(X - X^0)/N = 0.8mm^2$). Using such very inaccurate labels to train our deep learning model provides a powerful convincing example to test the proposed idea.

The used CNN model repeats convolution and a following pooling twice to covert one image $\Xi(k), 256 \times 256$, nto a 1×64 feature vector $V(k)$. The parameters of the convolution layers are (1, 32, 5, 4, 2) and (32, 64, 5, 4, 2) respectively. Batch normalization and ReLU activation are conducted between each convolution and pooling layer. Finally, the CNN inputs $V(k)$ to a 1×64 fully connected layer to calculate the output $O(k)$ of the CNN to compare with the inaccurate label $X^0(k)$. The first 15% of the data, all of those in experiment #1 and part of those from experiment #2 were used to validate and the rest 85% were used to train the CNN. As can be seen (Fig. 39), with the MSE lose function the model trained using inaccurate labels predicts very accurately despite the large inaccuracy in the used labels. While variance of the labeling error is $0.8mm^2$, the variance of the actual prediction error is $(X - O)^T(X - O)/N = 0.3mm^2$. The effectiveness of the proposed novel idea for using inaccurate labels is powerfully demonstrated through experiments.



(a)

(b)

Fig. 39 Training a deep learning model to predict the penetration using inaccurate labels. (A) Labeling error; (B) Prediction result.

The dataset in the above illustration may not be as large as desired and this can be the case in many applications. We can generate a mirrored dataset from (X^0, Ξ) by using $X^0(k) = X(k) - \varepsilon(k)$ for the given $X(k)$ and the already realized random $\varepsilon(k)$. In this way, we can form dataset (X^0, Ξ) whose size is

doubles to $2N$. This will be the new dataset with inaccurate labels for training that only doubled the data size but also GUARANTEED our requirement #1 for the labeling error (zero mean). It is apparently that we can infinitely increase the size by using different realization for the random labeling error and the mirroring extension so that the data size be $2N$, $4N$, …

While the above demonstrated the effectiveness in using inaccurate labels if the desired zero mean for the error is met, obtaining inaccurate labels meeting this property from experiments may not be straightforward when accurate labels are not available. It is exactly the purpose to find inaccurate labels that can be relatively easily obtained to replace accurate labels that are not easily measurable. While there may be multiple possible ways for a given application, we wish to propose the following way: (1) set the nominal conditions including, but not limited to, work-piece material, thickness, joint design, shielding gas flow rate, etc.; (2) under the nominal conditions, the weld penetration is controlled by the welding current and travel speed after the heat accumulation is realized; we add the travel sped to be used, for example 2 mm/s, as part of the nominal conditions; as such, the weld penetration is only controlled by the welding current and other uncontrolled perturbation after the heat accumulation; if no perturbations exist, the penetration after heat accumulation will be fully determined by the current i so that $X = f(i)$ where f is a deterministic function although it may not be known yet; (3) we try our best to minimize possible perturbations to conduct experiments with different levels of current; as such, $X(k) = f(i) + \varepsilon(k)$ (no difference with to express $X(k) = f(i) - \varepsilon(k)$) where ε is due to the effect from the realized uncontrollable perturbations; in this case $\mathrm{E}(X) = f(i)$ can be used as X^0 as it makes $X(k) = f(i) - \varepsilon(k)$ so that $X(k) = X^0(k) - \varepsilon(k)$ or $X^0(k) = X(k) + \varepsilon(k)$; (4) to assure the heat accumulation condition is met, we should form the dataset $(X^0, \Xi) = (f(i), \Xi)$ by excluding the data from the beginning in each where the inaccurate label is the same in the experiments using the same current i, i.e., $X^0(k) \equiv f(i)$. To assure the zero mean property for the labeling error due to the uncontrolled perturbations, as many as possible experiments under the same nominal conditions/welding current should be considered if possible.

Extra care may be needed to make sure the effect of the realized perturbations is acceptable. A simplest way is probably to simply use the data after the initial heat accumulation period. Another simple way to assure this is to use a simple image processing algorithm to calculate the weld pool width and make sure that the pool width is within acceptable range and such range changes with the level of the current or heat input used. In this case, the weld pool width is considered simple raw information to provide an inaccurate estimation of the penetration. There are many possible other ways to use easily obtained information that may provide inaccurate estimation for the penetration but the prediction error is unbiased (zero mean). The data from which the inaccurate estimation varies about a mean can be used. We may also monitor the back-side surface of the workpiece to acquire back-side images. A deep learning model may be established to use the back-side images that are determined by the weld pool and penetration to classify the current/heat input used. If the model can distinguish the current/input used, the effect from perturbations is acceptable. Data from experiments without a successful classification of the current/heat input should have been caused by large/abnormal perturbations and should be discarded. (Large perturbations theoretically do not affect the effectiveness in using inaccurate labels but such large/abnormal perturbations do not occur often so that realistically the zero mean property will be affected.) Of course, these are just some examples and more approaches may be proposed to determine the acceptability of the experimental data for the zero mean property.

Training a deep learning using inaccurate label $(f(i), \Xi)$ needs to know the function $f(i)$. This can be determined from just a few number of experiments. In case it is not known, the welding current i that determines the penetration under the nominal conditions can be used as the inaccurate label and the

mapping may not be even necessary. For example, we can easily determine the critical current $i^{(c)}$ that is minimally necessary to produce the complete penetration, the maximally allowed current $i^{(m)} > i^{(c)}$ that produce the maximally allowed w_b. A relative penetration measure $\tilde{x} = (\ i - i^{(c)})/(i^{(m)} - i^{(c)})$ may thus be defined with $\tilde{x} > 0$ for complete penetration, $\tilde{x} < 0$ for incomplete penetration, $\tilde{x} > 1$ for over penetration, and $0 < \tilde{x}^* < 1$ as the desired penetration. Of course, $\tilde{x}$ may be calibrated to the absolute measure using experimental data from few number of experiments, much much smaller than that needed to train a deep learning model for penetration prediction. As such, the labeling can be easily and automatically obtained the same as the phenomena to be used as the input of the deep learning model.

10.2 General Approach to Better Assure the Sufficiency:

The mechanism that correlates the proposed input and the penetration changes from process to process, application to application, and raw information to raw inform. Overemphasizing the mechanism reduces the benefits of using deep learning methods as accurate mechanism requires extensive efforts and in-depth knowledge to obtain and success is typically not guaranteed. As such, general approaches/principles that may help better assure the sufficiency of the raw information, thus the sufficiency of the corresponding model structure, are preferred. To this end, we try to find what current practices have ignored against general scientific principles.

For studies that use weld pool images as the raw information source Ξ to predict the penetration xa, $\Xi = f(xa) + e$ is assumed and e is due to uncontrolled perturbations around the nominal conditions and there are no other factors that affect the deterministic correlation between the observed phenomena and the penetration. That is, there are no other systematic variables γ so that $\Xi = f(xa, \gamma) + e$. If there are such γ, a deep learning model $xa = f^{-1}(\Xi)$ as the inverse of $\Xi = f(xa)$ would have systematic error rather than just the uncontrollable random error e. Unfortunately, existing studies have ignored possible γ without a justification/discussion with $\Xi = f(xa) + e$ being naturally assumed.

We can first argue that this natural assumption is in general incorrect, i.e., at least not always correct! This can be easily demonstrated using the pool oscillation as an example. In the classical model [107], the weld pool is approximated by a round membrane without thickness; the mass of the weld pool is determined by the radius of the membrane; the oscillation of the membrane is then approximated by a spring with the mass being determined by the radius. When there is no external force being applied, the spring demonstrates natural oscillation. However, when there is an external force, the oscillation is no longer natural and will be much more complex. The external force is needed to describe the behavior. For arc weld pool, it is subject to continuous application of the arc pressure and other forces. The arc force is so dominant that it dictates the oscillation behavior of the weld pool. The mass, thus the weld pool size and penetration, defines the dynamic system with other (constant) parameters. This system has the penetration determined model, the arc pressure/current determined input and the observed Ξ as the output. As such, the arc pressure which is approximately proportional to the square of the welding current [169] is a significant γ that cannot be ignored for the modeling mechanism based which the raw information and deep learning model are selected/designed. Hence, the deep learning model should in general consider using both Ξ and welding parameters waveforms as the inputs rather than just using Ξ as in most of the existing studies.

We can also argue that the assumed phenomena generation model $\Xi(k) = f(xa(k)) + e$ which serves as the foundation to train a deep learning model as its inverse may not necessarily be sufficient. This model assumes that the phenomena currently observed are only determined by the weld penetration at

the same time. It appears that this ignores the dynamic transition from the bottom of the weld pool where xa occurs to the top of the weld pool where the changed Ξ is observed. In a recent study [170], a Generative Adversarial network (GAN) is trained to generate $\Xi(k)$ from $xa(k)$ as represented by the current backside weld pool image. It is found that the topside image $\Xi(k)$ (middle in Fig. 40) cannot be accurately generated by just using $xa(k)$. The generated images (left) are not alike the real images (middle). When $xa(j)$'s $(j = k-7, \dots, k)$ as represented by the most recent eight backside images (sampled at 60 Hz), the topside image $\Xi(k)$ is generated with a reasonable accuracy. The generated images (right) are highly similar as the real ones (middle). As such, the deep learning model to estimate the weld penetration is better to be based on $\Xi(k) = f(xa(k-7{:}\,k)) + e$ than based on $\Xi(k) = f(xa(k)) + e$. The deep learning model designed based on the former better assumes the sufficiency of the needed raw information and thus needed deep learning model structure.

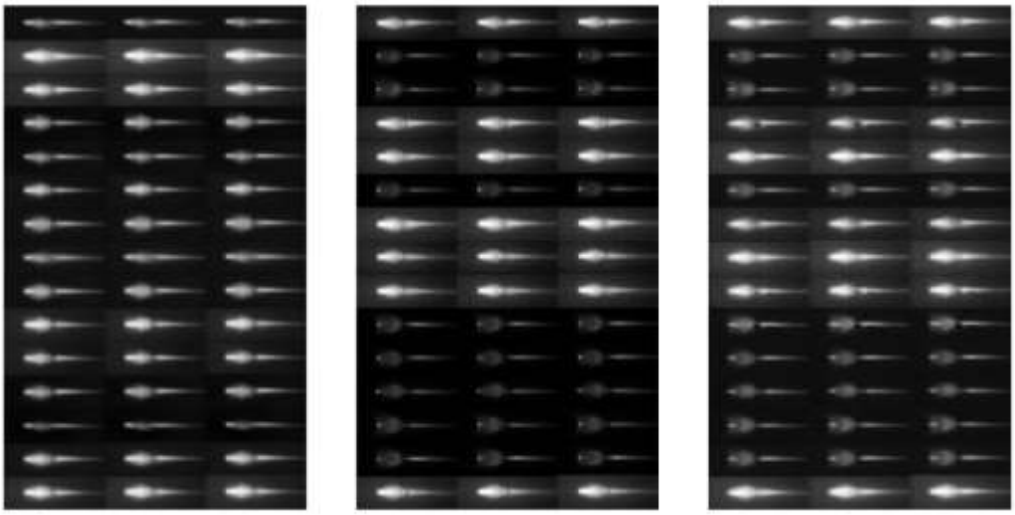


Fig. 40 Comparison of generated results from GAN. Middle: real topside images; Left: generated from one backside image; Right: generated from eight most recent backside images.

11. Summary

- Weld penetration refers to the state of the melting along the facing edges of the workpieces being joined. It occurs underneath the workpieces and is not directly observable from the side of the workpieces where the heat is applied and robot accesses. Monitoring of weld penetration is generally referred to as using indirect observables (welding phenomena) to estimate the state of the penetration.
- Relationships between proposed indirect observables and penetration state are generally very complex. An ideal unique mapping between them in general does not exist and there are other factors affecting their relationships. Fixing such influential other factors results in the nominal conditions. The Designed Welding Procedure is to produce the desired penetration state under the nominal conditions. Monitoring the weld penetration is needed if these factors may deviate from the nominal conditions. Around the nominal conditions, the indirect observables and weld penetration in general are better correlated.

- A key requirement in monitoring the weld penetration is the raw information (indirect observables used) containing sufficient information to determine the penetration. This sufficiency requirement theoretically cannot be guaranteed. Fitting and validation accuracies are just for the used dataset. The ability to apply a fitted model in manufacturing is that the dataset is large and representative for what may occur to affect the relationship between the used observables and penetration state.
- Leep learning accelerates the process to map the observables to the penetration state, i.e., from data to model. Conventional methods hand craft features, develop algorithms to calculate the features from the raw data, propose/change a model structure and fit the model parameters using the features as input and penetration labels as output. If the accuracy is not satisfactory, the features are revised. This is an iterative and trail-and-error process consisting of not automated separate subprocesses. Deep learning changes this into an automated process and thus revolutionizes the process to transmit information from dataset to the final model.
- Various kinds of raw information sources have been tried for being used to predict the penetration state for different processes and applications. The used deep learning model structure largely depends on the information sources. Various training techniques have also been used to improve the training accuracy.
- Despite the success in using deep learning to solve various penetration monitoring problems, we concern with (1) most studies lack justification for the sufficiency of the raw information; (2) most studies lack justification for their experiment plan; (3) most studies have not been tested in feedback control where the conditions may deviate from those used to generate the training dataset; (4) complex and multiple information sources are used without an adequate justification for the absolute needs.
- A major issue that fundamentally affects the ability of deep learning-based penetration monitoring in manufacturing application is the high cost associated with the acquisition of the needed large amount of penetration labels. This paper proposes to use inaccurate labels that can be easily obtained and specified the conditions to assure the effectiveness in using an accurate labeling approach.
- Another fundamental issue in deep learning-based penetration monitoring is the sufficiency of the model structure. While this is in general application dependent, general mechanisms may exist to minimize common errors. This paper proposes initial concepts and calls for further investigations.

Acknowledgement

This work is partially supported by the University of Kentucky Institute for Sustainable Manufacturing, Department of Electrical and Computer Engineering, and Department of Mathematics and the National Science Foundation under Grant CMMI-2024614 and DMS-2208314.

References

1. Wang, Q., Jiao, W., Wang, P. and Zhang, Y., 2021. A tutorial on deep learning-based data analytics in manufacturing through a welding case study. Journal of Manufacturing Processes, 63, pp.2-13.
2. Li, C., Wang, Q., Jiao, W., Johnson, M. and Zhang, Y.M., 2020. Deep learning-based detection of penetration from weld pool reflection images. *Weld J*, *99*(9), pp.239S-245S.
3. Cai, W., Wang, J., Zhou, Q., Yang, Y. and Jiang, P., 2019, February. Equipment and machine learning in welding monitoring: A short review. In *Proceedings of the 5th International Conference on mechatronics and robotics engineering* (pp. 9-15).
4. Kershaw, J., Yu, R., Zhang, Y. and Wang, P., 2021. Hybrid machine learning-enabled adaptive welding speed control. *Journal of Manufacturing Processes*, *71*, pp.374-383.

5. Oh, S. and Ki, H., 2020. Cross-section bead image prediction in laser keyhole welding of AISI 1020 steel using deep learning architectures. *IEEE Access*, *8*, pp.73359-73372.
6. Asif, K., Zhang, L., Derrible, S., Indacochea, J.E., Ozevin, D. and Ziebart, B., 2020. Machine learning model to predict welding quality using air-coupled acoustic emission and weld inputs. *Journal of Intelligent Manufacturing*, pp.1-15.
7. Wang, Q., Jiao, W. and Zhang, Y., 2020. Deep learning-empowered digital twin for visualized weld joint growth monitoring and penetration control. *Journal of Manufacturing Systems*, *57*, pp.429-439.
8. Nomura, K., Fukushima, K., Matsumura, T. and Asai, S., 2021. Burn-through prediction and weld depth estimation by deep learning model monitoring the molten pool in gas metal arc welding with gap fluctuation. *Journal of Manufacturing Processes*, *61*, pp.590-600.
9. Yu, R., He, H., Han, J., Bai, L., Zhao, Z. and Lu, J., 2022. Monitoring of back bead penetration based on temperature sensing and deep learning. *Measurement*, *188*, p.110410.
10. Kim, C., Hwang, S., Chung, J. and Sohn, H., 2020. Development of automatic crack detection technology in welded surface using laser active thermography and CNN deep learning. *Journal of the Korean Society for Nondestructive Testing*, *40*(3), pp.163-173.
11. Wu, D., Huang, Y., Zhang, P., Yu, Z., Chen, H. and Chen, S., 2020. Visual-acoustic penetration recognition in variable polarity plasma arc welding process using hybrid deep learning approach. *IEEE Access*, *8*, pp.120417-120428.
12. Jiao, W., Wang, Q., Cheng, Y. and Zhang, Y., 2021. End-to-end prediction of weld penetration: A deep learning and transfer learning based method. *Journal of Manufacturing Processes*, *63*, pp.191-197.
13. Zhang, Z., Wen, G. and Chen, S., 2019. Weld image deep learning-based on-line defects detection using convolutional neural networks for Al alloy in robotic arc welding. *Journal of Manufacturing Processes*, *45*, pp.208-216.
14. Feng, Y., Chen, Z., Wang, D., Chen, J. and Feng, Z., 2019. DeepWelding: A deep learning enhanced approach to GTAW using multisource sensing images. *IEEE Transactions on Industrial Informatics*, *16*(1), pp.465-474.
15. Jia, C.B., Liu, X.F., Zhang, G.K., Zhang, Y., Yu, C.H. and Wu, C.S., 2021. Penetration/keyhole status prediction and model visualization based on deep learning algorithm in plasma arc welding. *The International Journal of Advanced Manufacturing Technology*, *117*(11), pp.3577-3597.
16. Xia, C., Pan, Z., Fei, Z., Zhang, S. and Li, H., 2020. Vision based defects detection for Keyhole TIG welding using deep learning with visual explanation. *Journal of Manufacturing Processes*, *56*, pp.845-855.
17. Yaseer, A. and Chen, H., 2021. Machine learning based layer roughness modeling in robotic additive manufacturing. *Journal of Manufacturing Processes*, *70*, pp.543-552.
18. Das, D., Das, A.K., Pratihar, D.K. and Roy, G.G., 2021. Prediction of residual stress in electron beam welding of stainless steel from process parameters and natural frequency of vibrations using machine-learning algorithms. *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science*, *235*(11), pp.2008-2021.
19. Rohe, M., Stoll, B.N., Hildebrand, J., Reimann, J. and Bergmann, J.P., 2021. Detecting Process Anomalies in the GMAW Process by Acoustic Sensing with a Convolutional Neural Network (CNN) for Classification. *Journal of Manufacturing and Materials Processing*, *5*(4), p.135.
20. Knaak, C., von Eßen, J., Kröger, M., Schulze, F., Abels, P. and Gillner, A., 2021. A spatio-temporal ensemble deep learning architecture for real-time defect detection during laser welding on low power embedded computing boards. *Sensors*, *21*(12), p.4205.
21. Stadter, C., Schmoeller, M., von Rhein, L. and Zaeh, M.F., 2020. Real-time prediction of quality characteristics in laser beam welding using optical coherence tomography and machine learning. *Journal of Laser Applications*, *32*(2), p.022046.

22. Yin, L., Wang, J., Hu, H., Han, S. and Zhang, Y., 2019. Prediction of weld formation in 5083 aluminum alloy by twin-wire CMT welding based on deep learning. *Welding in the World*, *63*(4), pp.947-955.
23. Kim, H., Nam, K., Oh, S. and Ki, H., 2021. Deep-learning-based real-time monitoring of full-penetration laser keyhole welding by using the synchronized coaxial observation method. *Journal of Manufacturing Processes*, *68*, pp.1018-1030.
24. Cheng, Y., Wang, Q., Jiao, W., Yu, R., Chen, S., Zhang, Y. and Xiao, J., 2020. Detecting dynamic development of weld pool using machine learning from innovative composite images for adaptive welding. *Journal of Manufacturing Processes*, *56*, pp.908-915.
25. Kumar, N.P. and Varadarajan, R., 2021, February. Predictive modelling of the penetration coefficient of cold metal transfer welded joints using machine learning approaches. In *AIP Conference Proceedings* (Vol. 2317, No. 1, p. 020012). AIP Publishing LLC.
26. Li, S., Gao, J., Zhou, E., Pan, Q. and Wang, X., 2022. Deep learning-based fusion hole state recognition and width extraction for thin plate TIG welding. *Welding in the World*, pp.1-19.
27. Moinuddin, S.Q., Hameed, S.S., Dewangan, A.K., Kumar, K.R. and Kumari, A.S., 2021. A study on weld defects classification in gas metal arc welding process using machine learning techniques. *Materials Today: Proceedings*, *43*, pp.623-628.
28. Cheng, Y., Chen, S., Xiao, J. and Zhang, Y., 2021. Dynamic estimation of joint penetration by deep learning from weld pool image. *Science and Technology of Welding and Joining*, *26*(4), pp.279-285.
29. Oh, S., Kim, H., Nam, K. and Ki, H., 2021. Deep-learning approach for predicting laser-beam absorptance in full-penetration laser keyhole welding. *Optics Express*, *29*(13), pp.20010-20021.
30. Jin, C., Shin, S., Yu, J. and Rhee, S., 2020. Prediction model for back-bead monitoring during gas metal arc welding using supervised deep learning. *IEEE Access*, *8*, pp.224044-224058.
31. Cheng, Y., Yu, R., Zhou, Q., Chen, H., Yuan, W. and Zhang, Y., 2021. Real-time sensing of gas metal arc welding process–A literature review and analysis. *Journal of Manufacturing Processes*, *70*, pp.452-469.
32. Zhang, Y.M., Yang, Y.P., Zhang, W. and Na, S.J., 2020. Advanced welding manufacturing: A brief analysis and review of challenges and solutions. *Journal of Manufacturing Science and Engineering*, *142*(11).
33. Cheng, Y.C., Wang, Q.Y., Jiao, W.H., Xiao, J., Chen, S.J. and Zhang, Y.M., 2021. Automated recognition of weld pool characteristics from active vision sensing. *Weld J*, *100*(5), pp.183S-192S.
34. Liu, T.Y., Bao, J.S., Wang, J.L. and Gu, J., 2021. Laser welding penetration state recognition method fused with timing information. *Chinese Journal of Lasers*, *48*(6), p.0602119.
35. Liang, R., Yu, R., Luo, Y. and Zhang, Y., 2019. Machine learning of weld joint penetration from weld pool surface using support vector regression. *Journal of Manufacturing Processes*, *41*, pp.23-28.
36. Cai, W., Jiang, P., Shu, L., Geng, S. and Zhou, Q., 2021. Real-time monitoring of laser keyhole welding penetration state based on deep belief network. *Journal of Manufacturing Processes*, *72*, pp.203-214.
37. Chen, Z., Chen, J. and Feng, Z., 2018. Welding penetration prediction with passive vision system. *Journal of Manufacturing Processes*, *36*, pp.224-230.
38. Zhang, Y., Xiao, J., Zhang, Z. and Dong, H., 2022. Intelligent design of robotic welding process parameters using learning-based methods. *IEEE Access*, *10*, pp.13442-13450.
39. Lee, K., Kang, S., Kang, M., Yi, S. and Kim, C., 2021. Estimation of Al/Cu laser weld penetration in photodiode signals using deep neural network classification. *Journal of Laser Applications*, *33*(4), p.042009.
40. Wang, Z., Chen, H., Zhong, Q., Lin, S., Wu, J., Xu, M. and Zhang, Q., 2022. Recognition of penetration state in GTAW based on vision transformer using weld pool image. *The International Journal of Advanced Manufacturing Technology*, *119*(7), pp.5439-5452.
41. Jiao, W., Wang, Q., Cheng, Y., Yu, R. and Zhang, Y.M., 2020. Prediction of weld penetration using dynamic weld pool arc images. *Welding Journal*, *99*(11), pp.295s-302s.

42. Chen, C., Lv, N. and Chen, S., 2021. Welding penetration monitoring for pulsed GTAW using visual sensor based on AAM and random forests. *Journal of Manufacturing Processes*, *63*, pp.152-162.
43. Wu, D., Hu, M., Huang, Y., Zhang, P. and Yu, Z., 2021. In situ monitoring and penetration prediction of plasma arc welding based on welder intelligence-enhanced deep random forest fusion. *Journal of Manufacturing Processes*, *66*, pp.153-165.
44. Ren, W., Wen, G., Xu, B. and Zhang, Z., 2020. A Novel Convolutional Neural Network Based on Time–Frequency Spectrogram of Arc Sound and Its Application on GTAW Penetration Classification. *IEEE Transactions on Industrial Informatics*, *17*(2), pp.809-819.
45. Devaraj, J., Ziout, A. and Qudeiri, J.E.A., 2021. Grey-Based Taguchi Multiobjective Optimization and Artificial Intelligence-Based Prediction of Dissimilar Gas Metal Arc Welding Process Performance. *Metals*, *11*(11), p.1858.
46. Cai, W., Jiang, P., Shu, L., Geng, S. and Zhou, Q., 2021. Real-time laser keyhole welding penetration state monitoring based on adaptive fusion images using convolutional neural networks. *Journal of Intelligent Manufacturing*, pp.1-15.
47. Ma, G., Yu, L., Yuan, H., Xiao, W. and He, Y., 2021. A vision-based method for lap weld defects monitoring of galvanized steel sheets using convolutional neural network. *Journal of Manufacturing Processes*, *64*, pp.130-139.
48. Chen, Z., Chen, J. and Feng, Z., 2019. 3D Weld Pool Surface Geometry Measurement with Adaptive Passive Vision Images. *Welding Journal*.
49. Zhang, Z., Li, B., Zhang, W., Lu, R., Wada, S. and Zhang, Y., 2020. Real-time penetration state monitoring using convolutional neural network for laser welding of tailor rolled blanks. *Journal of Manufacturing Systems*, *54*, pp.348-360.
50. Chen, S., Teng, X., Sang, X., Zhang, H. and Zhuang, J., 2022. Automatic Recognition of Welding Seam Defects in TOFD Images Based on TensorFlow. *Automatic Control and Computer Sciences*, *56*(1), pp.58-66.
51. Chen, C., Xiao, R., Chen, H., Lv, N. and Chen, S., 2021. Prediction of welding quality characteristics during pulsed GTAW process of aluminum alloy by multisensory fusion and hybrid network model. *Journal of Manufacturing Processes*, *68*, pp.209-224.
52. Liu, T., Zheng, H., Bao, J., Zheng, P., Wang, J., Yang, C. and Gu, J., 2022. An Explainable Laser Welding Defect Recognition Method Based on Multi-Scale Class Activation Mapping. *IEEE Transactions on Instrumentation and Measurement*, *71*, pp.1-12.
53. Leonardo, N., Ronald, T., Andreas, H., Andreas, B. and Heinrich, H., 2012, May. Novel algorithm for the real time multi-feature detection in laser beam welding. In *2012 IEEE International Symposium on Circuits and Systems (ISCAS)* (pp. 181-184). IEEE.
54. Keshmiri, S., Zheng, X., Feng, L.W., Pang, C.K. and Chew, C.M., 2015, September. Application of deep neural network in estimation of the weld bead parameters. In *2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* (pp. 3518-3523). IEEE.
55. Chen, Y., Chen, B., Yao, Y., Tan, C. and Feng, J., 2019. A spectroscopic method based on support vector machine and artificial neural network for fiber laser welding defects detection and classification. *NDT & E International*, *108*, p.102176.
56. Nicolosi, L. and Tetzlaff, R., 2010. Real time control of curved laser welding processes by cellular neural networks (CNN): first results. *Advances in Radio Science*, *8*(C. 2-1/2-2), pp.117-122.
57. Jiang, R., Xiao, R. and Chen, S., 2021. Prediction of penetration based on infrared thermal and visual images during pulsed GTAW process. *Journal of Manufacturing Processes*, *69*, pp.261-272.
58. Kim, J.Y., Lee, D.Y., Lee, J. and Lee, S.H., 2021. Parameter Optimization of Hybrid-Tandem Gas Metal Arc Welding Using Analysis of Variance-Based Gaussian Process Regression. *Metals*, *11*(7), p.1087.
59. Gao, Y., Zhong, P., Tang, X., Hu, H. and Xu, P., 2021. Feature Extraction of Laser Welding Pool Image and Application in Welding Quality Identification. *IEEE Access*, *9*, pp.120193-120202.

60. Na, L., Chen, S.J., Chen, Q.H., Tao, W., Zhao, H. and Chen, S.B., 2021. Dynamic welding process monitoring based on microphone array technology. *Journal of Manufacturing Processes*, *64*, pp.481-492.
61. Zhang, Z., Zhang, Y., Luo, F., Li, J., Lu, C., Zhao, Y., Zhang, H. and Lu, A., 2019, June. Convolutional neural network using Bayesian optimization for laser welding tailor rolled blanks penetration detection. In *International Manufacturing Science and Engineering Conference* (Vol. 58752, p. V002T03A079). American Society of Mechanical Engineers.
62. Ren, W., Wen, G., Liu, S., Yang, Z., Xu, B. and Zhang, Z., 2018, September. Seam penetration recognition for GTAW using convolutional neural network based on time-frequency image of arc sound. In *2018 IEEE 23rd International Conference on Emerging Technologies and Factory Automation (ETFA)* (Vol. 1, pp. 853-860). IEEE.
63. Ozkat, E.C., Franciosa, P. and Ceglarek, D., 2017. Development of decoupled multi-physics simulation for laser lap welding considering part-to-part gap. *Journal of Laser Applications*, *29*(2), p.022423.
64. Zhang, L., Okudan, G., Basantes-Defaz, A.D.C., Gneiting, R.M., Subramaniam, S., Ozevin, D. and Indacochea, E., 2020. Characterization of GMAW (Gas Metal Arc Welding) Penetration Using Ultrasonics. *Materials*, *13*(10), p.2307.
65. Nicolosi, L., Tetzlaff, R., Abt, F., Blug, A. and Höfler, H., 2010, May. A camera based closed loop control system for keyhole welding processes: Algorithm comparison. In *Proceedings of 2010 IEEE International Symposium on Circuits and Systems* (pp. 2043-2046). IEEE.
66. Guo, R., Liu, H., Xie, G. and Zhang, Y., 2021. Weld defect detection from imbalanced radiographic images based on contrast enhancement conditional generative adversarial network and transfer learning. *IEEE Sensors Journal*, *21*(9), pp.10844-10853.
67. Božič, A., Kos, M. and Jezeršek, M., 2020. Power Control during Remote Laser Welding Using a Convolutional Neural Network. *Sensors*, *20*(22), p.6658.
68. Abt, F., Blug, A., Nicolosi, L., Dausinger, F., Höfler, H., Tetzlaff, R. and Weber, R., 2011. Real Time Closed Loop Control of Full Penetration Keyhole Welding with Cellular Neural Network Cameras. *Journal of Laser Micro/Nanoengineering*, *6*(2).
69. Nicolosi, L., Tetzlaff, R., Abt, F., Blug, A. and Höfler, H., 2012, August. Multi-feature detection for quality assessment in laser beam welding: Experimental results. In *2012 13th International Workshop on Cellular Nanoscale Networks and their Applications* (pp. 1-6). IEEE.
70. Blug, A., Abt, F., Nicolosi, L., Heider, A., Weber, R., Carl, D., Höfler, H. and Tetzlaff, R., 2012. The full penetration hole as a stochastic process: controlling penetration depth in keyhole laser-welding processes. *Applied Physics B*, *108*(1), pp.97-107.
71. Du, R., Xu, Y., Hou, Z., Shu, J. and Chen, S., 2019. Strong noise image processing for vision-based seam tracking in robotic gas metal arc welding. *The International Journal of Advanced Manufacturing Technology*, *101*(5), pp.2135-2149.
72. Zhang, D., Zhao, L. and Liu, A.B., 2021. Understanding and controlling the influence of laser energy on penetration, porosity, and microstructure during laser welding. *Chinese Journal of Lasers*, *48*(15), p.1502005.
73. Nicolosi, L., Blug, A., Abt, F., Tetzlaff, R., Höfler, H. and Carl, D., 2011. Real-Time Control of Laser Beam Welding Processes: Reality. In *Focal-Plane Sensor-Processor Chips* (pp. 261-281). Springer, New York, NY.
74. Abt, F., Heider, A., Weber, R., Graf, T., Blug, A., Carl, D., Höfler, H., Nicolosi, L. and Tetzlaff, R., 2011. Camera based closed loop control for partial penetration welding of overlap joints. *physics procedia*, *12*, pp.730-738.
75. Blug, A., Carl, D., Höfler, H., Abt, F., Heider, A., Weber, R., Nicolosi, L. and Tetzlaff, R., 2011. Closed-loop control of laser power using the full penetration hole image feature in aluminum welding processes. *physics procedia*, *12*, pp.720-729.

76. Shang, J., An, W., Liu, Y., Han, B. and Guo, Y., 2020. Oil pipeline weld defect identification system based on convolutional neural network. *KSII Transactions on Internet and Information Systems (TIIS)*, *14*(3), pp.1086-1103.
77. Nicolosi, L., Abt, F., Blug, A., Heider, A., Tetzlaff, R. and Höfler, H., 2011. A novel spatter detection algorithm based on typical cellular neural network operations for laser beam welding processes. *Measurement Science and Technology*, *23*(1), p.015401.
78. Yu, R., Kershaw, J., Wang, P. and Zhang, Y., 2022. How to Accurately Monitor the Weld Penetration From Dynamic Weld Pool Serial Images Using CNN-LSTM Deep Learning Model?. *IEEE Robotics and Automation Letters*, *7*(3), pp.6519-6525.
79. Buongiorno, D., Prunella, M., Grossi, S., Hussain, S.M., Rennola, A., Longo, N., Di Stefano, G., Bevilacqua, V. and Brunetti, A., 2022. Inline defective laser weld identification by processing thermal image sequences with machine and deep learning techniques. *Applied Sciences*, *12*(13), p.6455.
80. Baek, D., Moon, H.S. and Park, S.H., 2022. In-process prediction of weld penetration depth using machine learning-based molten pool extraction technique in tungsten arc welding. *Journal of Intelligent Manufacturing*, pp.1-17.
81. Wu, J., Shi, J., Gao, Y. and Gai, S., 2022. Penetration Recognition in GTAW Welding Based on Time and Spectrum Images of Arc Sound Using Deep Learning Method. *Metals*, *12*(9), p.1549.
82. Bahedh, A.S., Mishra, A., Al-Sabur, R. and Jassim, A.K., 2022. Machine learning algorithms for prediction of penetration depth and geometrical analysis of weld in friction stir spot welding process. *Metallurgical Research & Technology*, *119*(3), p.305.
83. Liu, S., Wu, D., Luo, Z., Zhang, P., Ye, X. and Yu, Z., 2022. Measurement of pulsed laser welding penetration based on keyhole dynamics and deep learning approach. *Measurement*, *199*, p.111579.
84. Gao, Y., Wang, Q., Xiao, J., Xiong, G. and Zhang, H., 2022. Weld penetration identification with deep learning method based on auditory spectrum images of arc sounds. *Welding in the World*, *66*(12), pp.2509-2520.
85. Kang, S., Kang, M., Jang, Y.H. and Kim, C., 2022. Deep learning-based penetration depth prediction in Al/Cu laser welding using spectrometer signal and CCD image. *Journal of Laser Applications*, *34*(4), p.042035.
86. Zhang, Z., Pan, H., Wang, X. and Lin, Z., 2022. Deep Learning Empowered Structural Health Monitoring and Damage Diagnostics for Structures with Weldment via Decoding Ultrasonic Guided Wave. *Sensors*, *22*(14), p.5390.
87. Chianese, G., Franciosa, P., Sun, T., Ceglarek, D. and Patalano, S., 2022. Using photodiodes and supervised machine learning for automatic classification of weld defects in laser welding of thin foils copper-to-steel battery tabs. *Journal of Laser Applications*, *34*(4), p.042040.
88. Wang, Z., Li, L., Chen, H., Lin, S., Wu, J., Ding, T., Tian, J. and Xu, M., 2022. Recognition of GTAW weld penetration based on the lightweight model and transfer learning. *Welding in the World*, pp.1-14.
89. Ferreira, G.R.B. and Ayala, H.V.H., 2022. Improved Image-Based Welding Status Recognition with Dimensionality Reduction and Shallow Learning. *Experimental Mechanics*, pp.1-14.
90. Chianese, G., Franciosa, P., Nolte, J., Ceglarek, D. and Patalano, S., 2022. Characterization of Photodiodes for Detection of Variations in Part-to-Part Gap and Weld Penetration Depth During Remote Laser Welding of Copper-to-Steel Battery Tab Connectors. *Journal of Manufacturing Science and Engineering*, *144*(7).
91. Schmoeller, M., Weiss, T., Goetz, K., Stadter, C., Bernauer, C. and Zaeh, M.F., 2022. Inline Weld Depth Evaluation and Control Based on OCT Keyhole Depth Measurement and Fuzzy Control. *Processes*, *10*(7), p.1422.
92. Zhou, F., Liu, X., Zhang, X., Liu, Y., Jia, C. and Wu, C., 2022. Keyhole status prediction based on voting ensemble convolutional neural networks and visualization by Grad-CAM in PAW. *Journal of Manufacturing Processes*, *80*, pp.805-815.
93. Li, C., 2019. Weld Penetration Identification Based on Convolutional Neural Network. PhD Dissertation, University of Kentucky (USA).

94. Reza, Z.N., 2019. Real-time automated weld quality analysis from ultrasonic B-scan using deep learning. PhD Dissertation, University of Windsor (Canada).
95. Connor, L.P., 1987. Welding Handbook. Vol. I: Welding Technology. *American Welding Society.*
96. Lohwasser, D. and Chen, Z. eds., 2009. *Friction stir welding: From basics to applications*. Elsevier.
97. Schwartz, M.M. and Aircraft, S., 1993. Introduction to brazing and soldering. *ASM International, ASM Handbook.*, *6*, pp.109-113.
98. Li, D., Chrysanthou, A., Patel, I. and Williams, G., 2017. Self-piercing riveting-a review. *The International Journal of Advanced Manufacturing Technology*, *92*(5), pp.1777-1824.
99. https://www.uti.edu/blog/welding/joint-types
100. Jain, A.K., Mao, J. and Mohiuddin, K.M., 1996. Artificial neural networks: A tutorial. *Computer*, *29*(3), pp.31-44.
101. LeCun, Y., Bengio, Y. and Hinton, G., 2015. Deep learning. *nature*, *521*(7553), pp.436-444.
102. Kotecki, D.J., Cheever, D.L. and Howden, D.G., 1972. Mechanism of ripple formation during weld solidification. *WELDING JOURNAL*, *51*(8), p.368.
103. Xiao, Y.H. and Den Ouden, G., 1993. Weld pool oscillation during GTA welding of mild steel. *WELDING JOURNAL*, *72*, pp.428-s.
104. Chen, J., Chen, J., Zhang, K., Feng, Z. and Zhang, Y.M., 2018. Dynamic reflection behaviors of weld pool surface in pulsed GTAW. *Welding Journal*, 97(6): 191s-206s. DOI: 10.29391/2018.97.017
105. Chen, W. and Chin, B.A., 1990. Monitoring joint penetration using infrared sensing techniques. Welding Journal, 69(4), pp.181s-185s.
106. Zhang, Y.M., Kovacevic, R. and Li, L., 1996. Characterization and real-time measurement of geometrical appearance of the weld pool. *International Journal of Machine Tools and Manufacture*, *36*(7), pp.799-816.
107. Kovacevic, R., Zhang, Y.M. and Ruan, S., 1995. Sensing and control of weld pool geometry for automated GTA welding. Journal of Engineering for Industry -Transactions of the ASME, 117(2): 210-222, 1995.
108. Kovacevic, R. and Zhang, Y., 1997. Real-time image processing for monitoring of free weld pool surface. Journal of Manufacturing Science and Engineering-Transactions of the ASME, 119(2): 161-169, 1997.
109. Suga, Y., Shimamura, T., Usui, S. and Aoki, K., 1999. Measurement of molten pool shape and penetration control applying neural network in TIG welding of thin steel plates. *ISIJ international*, *39*(10), pp.1075-1080.
110. Chan, B., Pacey, J. and Bibby, M., 1999. Modelling gas metal arc weld geometry using artificial neural network technology. *Canadian metallurgical quarterly*, *38*(1), pp.43-51.
111. Casais, J.L., Bierzychudek, M.E., Djokic, B. and Parks, H., 2020, August. Calibration of Weld Current Monitors at INTI Argentina. In *2020 Conference on Precision Electromagnetic Measurements (CPEM)* (pp. 1-2). IEEE.
112. Li, P.J. and Zhang, Y.M., 2001. Precision sensing of arc length in GTAW based on arc light spectrum. *J. Manuf. Sci. Eng.*, *123*(1), pp.62-65.
113. Pal, K. and Pal, S.K., 2011. Monitoring of weld penetration using arc acoustics. *Materials and Manufacturing Processes*, *26*(5), pp.684-693.
114. Yang, M., Bai, R., Zheng, H. and Qi, B., 2020. Temperature monitoring and calibration in Ti–6Al–4V molten pool with pulsed arc welding. *Science and Technology of Welding and Joining*, *25*(5), pp.369-376.
115. Ma, X.J. and Zhang, Y., 2011. Gas metal arc weld pool surface imaging: modeling and processing. *Welding Journal*, *90*(5), pp.85s-94s.
116. Wu, C.S., Gao, J.Q. and Wang, D.M., 2011. Observation of weld pool profiles in short-circuiting gas metal arc welding. *Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture*, *225*(10), pp.1873-1887.

117. Yan, Z., Zhang, G., Zhang, X., Gao, H. and Wu, L., 2004. Visual sensing and profile extraction of the weld pool in pulsed gas metal arc welding. *Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture*, *218*(10), pp.1333-1338.
118. Zhang, W., Wang, X. and Zhang, Y., 2013. Analytical real-time measurement of a three-dimensional weld pool surface. *Measurement Science and Technology*, *24*(11), p.115011.
119. Zhang, Y.M., Song, H.S. and Saeed, G., 2006. Observation of a dynamic specular weld pool surface. *Measurement Science and Technology*, *17*(6), p.L9.
120. Shi, Y., Zhang, G., Ma, X.J., Gu, Y.F., Huang, J.K. and Fan, D., 2015. Laser-vision-based measurement and analysis of weld pool oscillation frequency in GTAW-P. *Welding Journal*, *94*(5), pp.176s-187s.
121. Lancaster, J.F., 1986. The Physics of Welding. International Institute of Welding.
122. Andersen, K., Cook, G.E., Barnett, R.J. and Strauss, A.M., 1997. Synchronous weld pool oscillation for monitoring and control. *IEEE Transactions on Industry Applications*, *33*(2), pp.464-471.
123. Shi, Y., Zhang, G., Li, C., Gu, Y. and Fan, D., 2015. Weld pool oscillation frequency in pulsed gas tungsten arc welding with varying weld penetration. In *2015 IEEE international conference on automation science and engineering (CASE)* (pp. 401-406). IEEE.
124. Chokkalingham, S., Chandrasekhar, N. and Vasudevan, M., 2012. Predicting the depth of penetration and weld bead width from the infra red thermal image of the weld pool using artificial neural network modeling. *Journal of Intelligent Manufacturing*, *23*(5), pp.1995-2001.
125. Liu, Y. and Zhang, Y., 2013. Control of 3D weld pool surface. *Control Engineering Practice*, *21*(11), pp.1469-1480.
126. Liu, Y. and Zhang, Y., 2014. Control of human arm movement in machine-human cooperative welding process. Control Engineering Practice, 32, pp.161-171.
127. Liu, Y.K. and Zhang, Y.M., 2015. Supervised learning of human welder behaviors for intelligent robotic welding. *IEEE Transactions on Automation Science and Engineering*, *14*(3), pp.1532-1541.
128. Wu, C.S., Gao, J.Q. and Zhang, M., 2004. Sensing weld pool geometrical appearance in gas—metal arc welding. *Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture*, *218*(7), pp.813-818.
129. Bae, K.Y., Lee, T.H. and Ahn, K.C., 2002. An optical sensing system for seam tracking and weld pool control in gas metal arc welding of steel pipe. *Journal of Materials Processing Technology*, *120*(1-3), pp.458-465.
130. Yan, Z., Zhang, G., Zhang, X., Gao, H. and Wu, L., 2004. Visual sensing and profile extraction of the weld pool in pulsed gas metal arc welding. *Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture*, *218*(10), pp.1333-1338.
131. Zhang, Y.M. and Kovacevic, R., 1998. Neurofuzzy model-based predictive control of weld fusion zone geometry. *IEEE Transactions on Fuzzy Systems*, *6*(3), pp.389-401.
132. Neill, A.M. and Steele, J.P., 2016, July. Modeling and simulation of three dimensional weld pool reconstruction by stereo vision. In *2016 IEEE international conference on advanced intelligent mechatronics (AIM)* (pp. 542-547). IEEE.
133. Liang, Z., Chang, H., Wang, Q., Wang, D. and Zhang, Y., 2019. 3D reconstruction of weld pool surface in pulsed GMAW by passive biprism stereo vision. *IEEE Robotics and Automation Letters*, *4*(3), pp.3091-3097.
134. Zhang, W.J., Liu, Y.K., Wang, X. and Zhang, Y.M., 2012. Characterization of three-dimensional weld pool surface in GTAW. *Welding Journal*, *91*(7), pp.195s-203s.
135. Liu, Y., Zhang, W. and Zhang, Y., 2013. Dynamic neuro-fuzzy-based human intelligence modeling and control in GTAW. *IEEE Transactions on Automation Science and Engineering*, *12*(1), pp.324-335.
136. Zhang, W.J. and Zhang, Y.M., 2012. Modeling of human welder response to 3D weld pool surface: Part I-principles. *Welding Journal*, *91*(11), pp.310s-318s.
137. Chen, S.B., Lou, Y.J., Wu, L. and Zhao, D.B., 2000. Intelligent methodology for sensing, modeling and control of pulsed GTAW: Part 1--Bead-on-plate welding. *Welding Journal(USA)*, *79*(6), p.151.

138. Richardson, R.W., Gutow, D.A., Anderson, R.A. and Farson, D.F., 1984. Coaxial arc weld pool viewing for process monitoring and control. *Welding Journal*, *63*(3), pp.43-50.
139. Brzakovic, D. and Khani, D.T., 1991. Weld pool edge detection for automated control of welding. *IEEE Transactions on Robotics and Automation*, *7*(3), pp.397-343.
140. Feng, Y., Luo, Z., Liu, Z., Li, Y., Luo, Y. and Huang, Y., 2015. Keyhole gas tungsten arc welding of AISI 316L stainless steel. *Materials & Design*, *85*, pp.24-31.
141. Li, X.R., Shao, Z., Zhang, Y.M. and Kvidahl, L., 2013. Monitoring and control of penetration in GTAW and pipe welding. *Welding journal*, *92*(6), pp.190-196.
142. Zhang, Y., Li, X. and Shao, Z., 2017. *Method to monitor and control weld penetration in gas tungsten welding and full-position pipe welding*. U.S. Patent 9,604,301.
143. Zhang, Y.M., Wu, L., Walcott, B.L. and Chen, D.H., 1993. Determining joint penetration in GTAW with vision sensing of weld face geometry. *WELDING JOURNAL*, *72*, pp.463-s.
144. Zhang, Y.M., Kovacevic, R., 1997. "Real-time sensing of sag geometry during GTA welding," Journal of Manufacturing Science and Engineering-Transactions of the ASME, 119(2): 151-160.
145. Zhang, Y.M., Kovacevic, R. and Wu, L., 1996 "Dynamic analysis and identification of gas tungsten arc welding process for full penetration control," Journal of Engineering for Industry - Transactions of the ASME, 118(1): 123-136.
146. Zhang, Y.M., Kovacevic, R. and Li, L., 1996. Adaptive control of full penetration gas tungsten arc welding. *IEEE Transactions on Control Systems Technology*, *4*(4), pp.394-403.
147. Zou, Y., Lan, R., Wei, X. and Chen, J., 2020. Robust seam tracking via a deep learning framework combining tracking and detection. Applied Optics, 59(14), pp.4321-4331.
148. Yang, L., Fan, J., Huo, B., Li, E. and Liu, Y., 2022. Image denoising of seam images with deep learning for laser vision seam tracking. *IEEE Sensors Journal*, *22*(6), pp.6098-6107.
149. Zhang, K., Zhang, Y.M., Chen, J.S. and Wu, S.J., 2017. Observation and analysis of three-dimensional weld pool oscillation dynamic behaviors. *Welding Journal*, *96*(5), pp.143s-153s.
150. Girshick, R., 2015. Fast r-cnn. In *Proceedings of the IEEE international conference on computer vision* (pp. 1440-1448).
151. Samuel, Arthur, 1959. Some Studies in Machine Learning Using the Game of Checkers. IBM Journal of Research and Development. 3 (3): 210–229.
152. Zhang, Y.M., Kovacevic, R. and Wu, L., 1992. Sensitivity of front-face weld geometry in representing the full penetration. *Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture*, *206*(3), pp.191-197.
153. Kovacevic, R., Zhang, Y.M. and Li, L., 1996. Monitoring of weld joint penetration based on weld pool geometrical appearance. *Welding Journal*, *75*(10).
154. Using NeuralWorks: Professional II/Plus and NeuralWorks Explorer, NeuralWare Inc., Pittsburgh, PA, 1993.
155. Mohammedi, R.I., Symagulov, A., Kuchin, Y., Yakunin, K. and Yelis, M., 2021. From classical machine learning to deep neural networks: A simplified scientometric review. Applied Sciences, 11(12), p.5541.
156. D. H. Rao and P. P. Panduranga, "A Survey on Image Enhancement Techniques: Classical Spatial Filter, Neural Network, Cellular Neural Network, and Fuzzy Filter," 2006 IEEE International Conference on Industrial Technology, Mumbai, India, 2006, pp. 2821-2826, doi: 10.1109/ICIT.2006.372671.
157. L. Zhang, L. Zhang and B. Du, "Deep Learning for Remote Sensing Data: A Technical Tutorial on the State of the Art," in *IEEE Geoscience and Remote Sensing Magazine*, vol. 4, no. 2, pp. 22-40, June 2016, doi: 10.1109/MGRS.2016.2540798.
158. Sepp Hochreiter, Jürgen Schmidhuber; Long Short-Term Memory. Neural Comput 1997; 9 (8): 1735–1780. doi: https://doi.org/10.1162/neco.1997.9.8.1735
159. Simonyan, K. and Zisserman, A., 2014. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*.

160. Sharma, S., Ball, J.E., Tang, B., Carruth, D.W., Doude, M. and Islam, M.A., 2019. Semantic segmentation with transfer learning for off-road autonomous driving. *Sensors*, *19*(11), p.2577.
161. Yu, R., Kershaw, J., Wang, P. and Zhang, Y., 2021. Real-time recognition of arc weld pool using image segmentation network. *Journal of Manufacturing Processes*, *72*, pp.159-167.
162. Agarap, A.F., 2018. Deep learning using rectified linear units (relu). *arXiv preprint arXiv:1803.08375*.
163. Miao, Y., Gowayyed, M. and Metze, F., 2015, December. EESEN: End-to-end speech recognition using deep RNN models and WFST-based decoding. In *2015 IEEE Workshop on Automatic Speech Recognition and Understanding (ASRU)* (pp. 167-174). IEEE.
164. Devlin, J., Cheng, H., Fang, H., Gupta, S., Deng, L., He, X., Zweig, G. and Mitchell, M., 2015. Language models for image captioning: The quirks and what works. *arXiv preprint arXiv:1505.01809*.
165. He, K., Zhang, X., Ren, S. and Sun, J., 2015. Spatial pyramid pooling in deep convolutional networks for visual recognition. *IEEE transactions on pattern analysis and machine intelligence*, *37*(9), pp.1904-1916.
166. Zheng, Q., Zhao, P., Li, Y., Wang, H. and Yang, Y., 2021. Spectrum interference-based two-level data augmentation method in deep learning for automatic modulation classification. *Neural Computing and Applications*, *33*(13), pp.7723-7745.
167. Weiss, K., Khoshgoftaar, T.M. and Wang, D., 2016. A survey of transfer learning. *Journal of Big data*, *3*(1), pp.1-40.
168. Yu, R, Chen., Y.Z., Zhang, J., Ye, Q. and Zhang, Y.M., 2023. Monitoring Weld Penetration By Training A Deep Learning Model Using Inaccurate Labels. Automation, Robotics & Communications for Industry 4.0/5.0 (ARCI' 2023): 3rd IFSA Winter Conference, 22-24 FEBRUARY 2023, CHAMONIX-MONT-BLANC, FRANCE
169. de Simas Asquel, G., Bittencourt, A.P.S. and da Cunha, T.V., 2020. Effect of welding variables on GTAW arc stagnation pressure. *Welding in the World*, *64*(7), pp.1149-1160.
170. Edison Mucllari, Rui Yu, Yue Cao, Qiang Ye, YuMing Zhang. “Do We Need a New Foundation to Use Deep Learning to Monitor Weld Penetration?” in review for IEEE Robotics and Automation Letters.