WILEY | Hindawi

Review Article

A Survey of Machine Learning in Friction Stir Welding, including Unresolved Issues and Future Research Directions

Utkarsh Chadha [iD],[1] Senthil Kumaran Selvaraj [iD],[1,2] Neha Gunreddy [iD],[1] S. Sanjay Babu [iD],[3] Swapnil Mishra [iD],[1] Deepesh Padala [iD],[3] M. Shashank [iD],[3] Rhea Mary Mathew,[1] S. Ram Kishore [iD],[1] Shraddhanjali Panigrahi [iD],[4] R. Nagalakshmi,[5] R. Lokesh Kumar,[6] and Addisalem Adefris [iD][7]

[1]School of Mechanical Engineering (SMEC), Vellore Institute of Technology, 632014, Vellore, Tamilnadu, India
[2]Department of Manufacturing Engineering, School of Mechanical Engineering (SMEC), Vellore Institute of Technology, 632014, Vellore, Tamilnadu, India
[3]School of Electronics Engineering (SENSE), Vellore Institute of Technology, 632014, Vellore, Tamilnadu, India
[4]Department of Mechanical Engineering, National Institute of Technology, Raipur, Chattisgarh, India
[5]Department of Computer Science and Engg., Faculty of Engineering and Tech., Kalinga University, Raipur, Chhattisgarh, India
[6]School of Computer Science and Engineering (SCOPE), Vellore Institute of Technology, 632014, Vellore, Tamilnadu, India
[7]School of Mechanical and Automotive Engineering, College of Engineering and Technology, Dilla University, P.O. Box 419, Dilla, Ethiopia

Correspondence should be addressed to Senthil Kumaran Selvaraj; senthilkumaranselvaraj82@gmail.com and Addisalem Adefris; addisalem@du.edu.et

Friction stir welding is a method used to weld together materials considered challenging by fusion welding. FSW is primarily a solid phase method that has been proven efficient due to its ability to manufacture low-cost, low-distortion welds. The quality of weld and stresses can be determined by calculating the amount of heat transferred. Recently, many researchers have developed algorithms to optimize manufacturing techniques. These machine learning techniques have been applied to FSW, which allows it to predict the defect before its occurrence. ML methods such as the adaptive neurofuzzy interference system, regression model, support vector machine, and artificial neural networks were studied to predict the error percentage for the friction stir welding technique. This article examines machine learning applications in FSW by utilizing an artificial neural network (ANN) to control fracture failure and a convolutional neural network (CNN) to detect faults. The ultimate tensile strength is predicted using a regression and classification model, a decision tree model, a support vector machine for defecting classification, and Gaussian process regression (UTS). Machine learning implementation mainly promotes uniformity in the process and precision and maximally averts human error and involvement.

1. Introduction

Friction stir welding is perfect for joining commercially available alloys in aerospace, shipbuilding, electronics, and rails. Most alloys are difficult to weld by traditional methods such as the fusion process. The friction produced by the rub-bing action of the tool on the workpiece produces heat. Low temperatures make the process more energy-efficient while still stopping the workpiece from shrinking [1]. The process reduces the number of workers as it is easy to automate the tool and can be repeated many times. Friction stir welding has been studied extensively over decades, returning to its

infancy. Most research has focused on microstructural properties, material weldability, and mechanical properties after welding. The FSW method has recently been applied to machine learning techniques. Various machine learning models and neural networks effectively detected defects and assessed UTS in friction stir welding. Some of the significant readings and data were studied for further analysis. As a result, many sets of algorithms, including classification, regression, and clustering models, have limited accuracy to a certain extent. The FSW cell environment evaluates K-nearest neighbor, multilayer perceptron, principal component analysis, and random forest methods. Rotational speed, forging speed, travel speed, transverse and longitudinal forces, and specific energy and torque are the process parameters given as input in the machine learning model. The defect index is produced as an output [2]. Process parameters for FSW can be easily determined with less error than 5% using machine learning. Process parameters for FSW can be easily determined with less error percentage of less than 5%. Thus, it proved to be more accurate and robust.

To create a high upright deformity-free weld, RPM of the shoulder-pin arrangement, traverse speed, descending producing power, and apparatus pin configuration should be picked cautiously. To create a high uprightness deformity-free weld, measure factors RPM of the shoulder-pin arrangement, traverse speed, the descending producing power, and apparatus pin configuration should be picked cautiously [3].

2. Understanding the Process

In friction stir welding, measuring a cylinder-shaped device with a profiled probe is pivoted and, bit by bit, inserted between two bits of sheet or plate materials to be welded together to frame a joint. The pieces should be clasped onto a back bar so that the adjoining joint appearances are not pulled separated or pushed strange in some other manner. Wear-resistant welding instruments and workpiece material produce frictional heat. The heat softens the workpiece to reach the necessary temperature, allowing the machine to cross the weld boundary. The subsequent plasticized substance is moved from the top to the following edge of the device. It is joined together by the contact between the shoulder of the device and the pin's top, making a solid phase association between the two sections. Friction stir welding (FSW) is a solid-state joining measure that uses a nonconsumable apparatus to join two confronting plates or workpieces without softening the material. Friction between the pivoting system and the workpiece generates heat, resulting in a soft area in the workpiece near the FSW apparatus. As the apparatus moves, it combines the two separate bits of metal, which produce the hot and mollified metal by utilizing a mechanical pressing factor through the apparatus comparable to joining clay. High-dissolving temperature materials required for FSW apparatus materials require high hardness for scraped area obstruction, alongside substance steadiness and good sturdiness at high temperature. For working with various apparatus materials, advances are progressing quickly, and every material gives explicit benefits to different applications.

2.1. Friction Stir Welding. Figure 1 shows the different regions formed at the time of welding. The heat produced by friction stir welding and the plastic flow causes refined crystallized grains to be formed in the weld zone. In the thermomechanically affected region, recovered grains are seen in the weld. In terms of metalworking zones, there consists 5 stages: the preheat stage followed by forging, extrusion, initial deformation, and the postheat or cool down process [4]. Therefore, the processes involved in FSW can be summarised as given in Figure 2.

2.1.1. FSW of Aluminum Alloys. Friction stir welding can now quickly weld magnesium, copper, aluminum, and stainless steel alloys previously challenging to weld using traditional welding methods. Researchers have found that the FSW process allows materials such as AA 2195 aluminum alloy to be joined quickly, which is usually laborious to undergo a fusion welding process. The quality of weld and stresses can be determined by calculating the amount of heat transferred. The heat lost during the friction stir welding process (FSW) is just 5% (95% efficiency), i.e., the rest of the heat produced is transferred to the workpiece and provides a good quality weld. FSW can now weld even the Al alloy Al7075, which is considered nonweldable. This welding technique's present applications include high-speed train manufacturing, shipbuilding, and even aviation. A cylindrical tool is used in the friction stir butt-welding procedure when two plates with comparatively less thickness are used, and a conical tool is used for thicker plates. This process is also successful in welding plates made of zinc and was considered not easy to weld or even nonweldable by the conventional methods. Aluminum plates that are as much as 100 mm thick can also be joined through this process by double-sided welding. Friction stir welding can also be done underwater since it is a solid-state welding method.

One of the most extensive weld defects is the wormhole defect. One of the leading causes of this defect is increased welding speed while the rotating tool's rpm does not change. Tool geometry also affects this defect significantly. The microstructural changes that occur during friction stir welding are of two types which are the following:

(1) FSW is done on materials with low recrystallization rates (e.g., Al alloys)

(2) FSW looked at materials with faster recrystallization rates (Ag austenitic stainless steels and Ti alloys) and discovered that they have two distinct zones: the HAZ and the stirred zone

The peak temperatures in the stirred zone vary from 0.6 to 0.95 Tm. The tool design influences the peak temperature, material, and operating conditions [6, 7]. The purpose of fabricating components through the FSW of aluminum has opted in recent times. The heat produced from the FSW process is sufficient to produce microstructural variation in hardened alloys. FSW proves to be a highly energy-efficient process. FSW produces high-strength alloys [8].

2.1.2. FSW of Aluminum Alloys with Steel. One of the most significant factors for converting aluminum alloys with steel through conventional methods to friction stir welding was
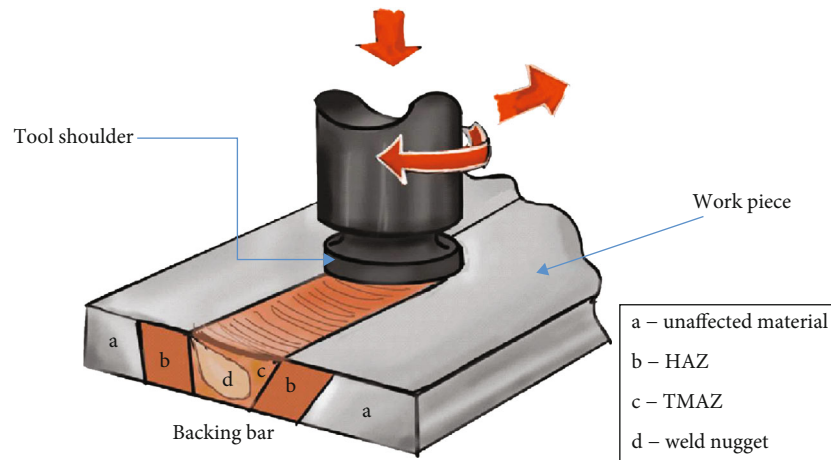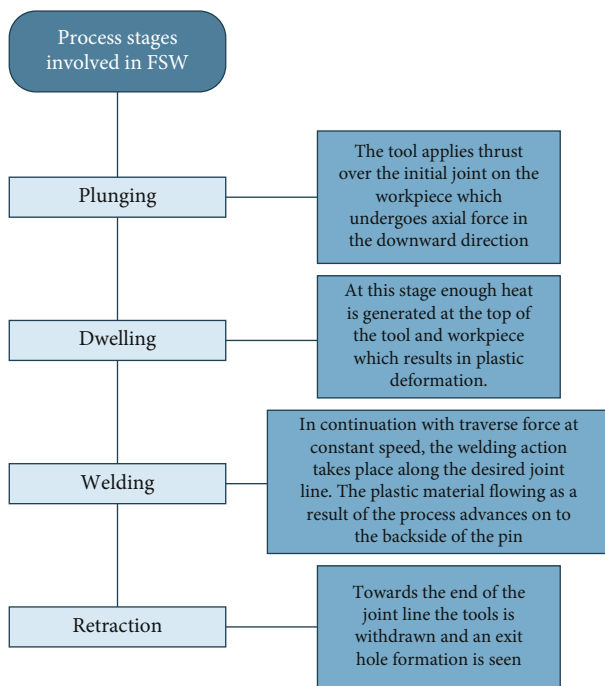
FIGURE 1: Friction stir welding process.

FIGURE 2: Stepwise description of processes involved in FSW [5].

reducing fuel usage and benefitting the solid-state joining process. The stirred zone is majorly formed on aluminum. The attempt to weld Al alloys with steel was made through FSW. The tool is plunged into the aluminum then moved towards the surface of the steel. To prevent overheating, the tool is offset towards the aluminum alloy. At an offset of 0.2 mm, a joint with maximum strength was obtained. It is always accomplished to ensure that the melting temperature of the components to be welded together is still higher than the base material melting point [9]. Aluminum bonds to the steel surface in a semisolid paste-like form because of its heat rotating tool. In this case, the tool pin is plunged into the aluminum's soft surface and not on the two metals' junction. It is done to prevent the insufficient stirring of the

aluminum and steel, which occurs because of the reduced tool life. Also, this method provides minimal wear to the rotating pin as the pin only acts on the soft surface of the aluminum. When the pin rotation speed was set at 250 rpm, the tool life improved drastically, and a good joint was made (at 250 rpm, a joint was produced with 86% tensile strength of the aluminum metal). When the pin rotation speed was set at 100 rpm, the tool was worn out quickly, and a successful weld was not made. When the pin rpm was set to 500, the weld's surface morphology appeared to be the same as when pin rotation of 250 rpm was used, but the joint's strength was much less than that of 250 rpm. The weld could not be completed at an even faster pin rotation of 1250 rpm. Burning was observed at an rpm of 1250. The pin size of the rotating tool should also be just right. Pins less than 2 mm in diameter could not produce a good weld. The aluminum plate must be mounted at the retreating side; otherwise, the weld cannot be produced [10, 11]. The interface region is mainly divided into two sections: the mixed and intermetallic compound layers. A vortex-like structure was observed in the mixed layer. The aluminum part's cooling rate was lower than the steel part's cooling rate. A vortex-like structure was observed on the interface when 304 austenitic stainless steel and 6056 Al alloy were welded together [9]. When X5CrNi18-10 stainless steel and 6013-T4 Al alloy were welded together, the plate thickness for both of these products was 4 mm. The traverse speed was 80 mm/min, and the tool rotation speed was set to 800 rpm. On the stainless steel's surface, coarse austenitic grains were evident. One of the main outcomes of this study was that it was proved that Al 6013 could be welded with different stainless steels by the FSW method. The microstructure of the weld zone was divided into 6 welding zones which are the following:

(1) HAZ in advancing side

(2) Weld nugget

(3) TMAZ in the advancing side

(4) TMAZ in the retreating side

(5) HAZ in the retreating side

(6) Parent aluminum alloy

(7) Parent stainless steel

In [12], it was detected that the amount of flash produced decreased as the speed of the weld was increased. The maximum amount of flash was formed during the starting of slow welding. As the welding speed rises, the solidity level also increases. It was also seen that the cooling rate was much faster at the edges of the weld nugget. It suggests that the weld nugget's hardness was stronger at the edges. A sharp decrease in the hardness value from TMAZ of the steel towards the retreating side's weld nugget was observed. The weld nugget contained coarse stainless steel fragments, resulting in nonuniform weld hardness [13].

2.1.3. FSW Effects of Microstructure and Fatigue.
The microstructures of friction stir welded regions largely vary w.r.t. grain formation among the weld zones, i.e., stir zone, thermomechanically affected zone, heat-affected zone, and the base material. At the interfaces of the shoulder and pin, the heat produced from the friction is the highest because of the close distance to the FSW action, and this temperature decreases as it progresses towards the base material. The temperature ranges in the regions are as follows:

(i) Stir zone peak temperature range: 0.9 to 0.75 melting temperature (Tm)

(ii) Thermomechanically affected zone peak temperature range:0.7-0.6 Tm

(iii) Heat affected zone peak temperature range: 0.55 Tm to surrounding temperatures from the other two zones

The pin height is set according to the width of the plates that have to be welded together, as it was seen that this greatly affects the microstructure of the weld. The stirring action of a pin causes thermomechanical deformation. The flat grains (of parent metal) are attracted to the weld nugget region. When we speak about the disparity in hardness numbers between welds, the top side has lower hardness numbers because it is in contact with the tool shoulder. The backside is connected with the backplate and thus acts as a heat sink. Observations were taken when the rpm was between 300 and 500, and the traverse speed was taken as 120 mm per minute [14]. In the initial deformation stage, flow stress increases as the dislocations multiply. This also causes a rise in the recovery rate as the dislocations start rearranging themselves and producing low-angle boundaries. The flow stress and the recovery rate gradually enter a state of equilibrium. Microstructures can be studied using various methods, including orientation imaging microscopy and transmission electron microscopy [15].

The connection between measure boundaries and the width of remuneration interlayer strips has played a significant role in influencing welded joints' microstructural and mechanical properties [16].

Deformation of the pin is always broader on the retreating side due to the clockwise rotation of the tool. Onion rings or fine-grained bands are seen under polarised light, producing different crystallographic orientations (formed behind the tool). The grain structure in the immediate front of the tool has the highest thermal gradient. The plastic zone was observed to be formed <1 mm in front of the tool instead of the weld's advancing side; the retreating side's deformation zone is more significant. This occurs due to the tool's direction of rotation; as the tool rotates in a clockwise direction, the material is extruded onto the weld's withdrawing side. Because of the tool's clockwise rotation, the parent grains in a thermomechanical affected zone are sheared forward towards the welding direction on the advancing side [17].

When workpiece-2195 aluminum-lithium, rotational speed of 180 revolutions/min was used, the grain structure observed was defined as deformation and mechanical processes. When the material is brushed across the welding tool, it leads to the formation of refined grains that originate from the fcc sheer texture that in turn forms a modified texture [18].

The highest weld temperature is observed to be directly related to the rotation speed of the tool. A sharp rise in maximum temperature is observed till 2000 rpm, and after this, there is a gradual temperature rise. High temperatures and intensive strains cause dynamically recrystallized grains to get nucleated. The grain size of the microstructure was also found to increase with temperature. No significant increase in hardness was observed at different rotation speeds in the welds [19]. Improving the rotational and traverse speeds can lead to a no-defect weld as the welding's final result. Studying the grain TEM microstructure (transmission electron microscopy) is compared with the aluminum workpiece before the welding process [20].

The heat produced by friction stir welding and the plastic flow causes refined crystallized grains to be formed in the weld zone. In the thermomechanically affected region, recovered grains are seen in the weld. Precipitation sequence during FSW processes in the softened regions, where precipitates are dissolved at temperatures above 675 K [21].

The grains formed in the weld center were fine due to the high deformation leading to crystallization. The grains formed from recrystallization are 2-4 micrometers. This solid-state process allows the welding of alloys considered impossible to be welded using traditional welding methods [22].

Table 1 depicts the variations in the formation of microstructures of different metals welded. In the case of carbon steel welding, although martensite formation leads to an increase in strength, it causes a decline in sturdiness and malleability. This can be improved by decelerating the cooling rates to acquire desirable results. The finer the grain size in the stir zone for magnesium alloys, the better the joint hardness improvisation.

2.1.4. Thermal and Mechanical Modeling of FSW.
The friction stir welds' surface structure can correspond well with the same plastic strain disseminations on the top surface [28]. The friction produced by the pin on the material and the friction built by the shoulder on the workpieces' surfaces are all factors that influence the heat increase in FSW. According to some research, pin friction and plastic strain are slightly less than the shoulder's heat [29]. AlSI 1018 steel was welded using a tungsten tool. The thermal conductivity of the tool was found to be around four times the thermal

TABLE 1: Explanation of variation in microstructures for different metals.

S.no.	FSW material composition	The characteristic feature of the microstructure	Ref.
1	Carbon steels	A gradual transformation occurs in the microstructure from austenite to martensite, increasing carbon content above 0.2%. The joint hardness is improved	[23]
2	Stainless steel (304 austenitic stainless steel)	Ferrite form and sigma form were observed. The ferrite form was mainly formed at the grain borders of austenite, which is recrystallized due to the high cooling rates, whereas the sigma phase was seen due to a decrease in cooling rate	[24]
3	AZ61 magnesium alloy	Stir zone: grains are well cultivated as a result of recrystallization. Compared to the base metal, the grain dimensions decrease as it goes onto the other weld zones, i.e., TMAZ and SZ. TMAZ region: distorted grain pattern is observed	[25]
4	Dissimilar aluminum alloy joints	Finer grains are formed near the interfaces of the two dissimilar alloys. The stir zone of dissimilar Al alloys consists of various grain sizes depending on the intermixing of two metals during FSW	[26]
5	Dissimilar magnesium alloys (A5052P-O and AZ31B-O alloys)	The differing microstructure is introduced near the bonded interface. It is visible by the zigzag pattern near the interface. With increasing tool speed, the surface structure turns smoother	[27]

conductivity of steel; therefore, a notable amount of heat was passed on to the tool. More heat is built farther from the rotating tool's axis compared with the heat generated at the axis. It happens because the tool's relative velocity is much more on the outside than on the axis. It can be seen that there occurs a significant amount of heat transfer and plastic flow close to the tool where convective heat transfer was observed to be the main source of heat transfer. It was also seen that there was a slight unevenness of temperature profiles near the tool. The combined impact of the tool rotation direction and the traverse direction was responsible for this [30]. Thermal modeling has been an integral part of friction stir welding since the late 1990s. The most significant cause is that almost all weld features result from a workpiece's thermal history. FSW highly depends on heat generation and flow. Once the solidus temperature was reached, the yield stress was significantly reduced. A self-stabilizing effect occurs when the material's temperature is lower than the solidus temperature. To obtain steady-state conditions, the contact stress should always be balanced with material yield shear stress. It was also seen that the heat generation for material depends on its flow stress, strain, and strain rate for a given temperature [31]. Aluminum alloy 6061-T6 was used to analyze the butt-welding process using a 3-D model based on finite element analysis. Friction between rotating tools, welded plates, and plastic deformation of the material around the tool led to heat transfer. The thermal and mechanical solutions during FEA are integrated to improve the model's accuracy. The X-ray diffraction method was applied to determine residual stresses. The expansion of material during heating of the welded plates led to stress formation in the weld [32]. Experimental procedures are tedious to analyze the welds' complex geometry and properties; therefore, a numerical analysis method is chosen to study the weld joints. The numerical method allows the determination of residual stresses and distortions. Hybrid processes produce a new technique by combining different FSW techniques to produce stronger welds and improve structural behavior [33]. FSW's nature of joints depends mainly on the progression of materials in the mixed

zone. Accordingly, specialists attempted various endeavors to build the progression of material in a mixing region by shifting solder or pin geometry. Preheating the workpiece before welding can be set up is another option. Preheating relaxes materials and expands materials' progression in the mixed region and accordingly diminishes the welding powers and improves productivity. It was seen that in FSW, the strength of joints increments by 8% when contrasted with regular FSW through preheating. Preheating improves the material stream and peripheral change in disfigurement conduct, bringing about the nugget zone's expanding hardness. Inferable from extra warmth, HAZ increases, and TMAZ diminishes [34].

Thermal models support the theory that the connection between temperature proportion and energy is normal for aluminum alloys that ration equal thermal diffusivities. A thermal model can produce trademark temperature bends. A composite's most incredible welding temperature might be assessed beforehand to know its thermal diffusivity, welding boundaries, and apparatus geometry [35].

Base material thickness and welding speed are auxiliary boundaries overseeing torque and temperature. Their impact on force is recognizable when enormous shoulder instruments are utilized and thick welding plates [36].

2.1.5. Defects in FSW. Figure 3 indicates dendrite formation in an ongoing friction stir welding process because of unstable solidification of the alloying materials.

Using conventional methods like fusion welding on aluminum alloys causes many problems like forming blowholes (entrapment of hydrogen gas). It joins materials below their melting point. This way prevents the formation of blowholes as the material never melts. The length of the tool pin always has to be less than that of the plates that are to be welded together. In a study, the tool pin's length was taken as 3.9 mm, and the thickness of the plate was 4 mm. A good joint was obtained when the tool's downforce was set to be 6.9 kN, and the traverse speed was set to be 250 mm/min. A cavity was detected in the joint when traverse speed was
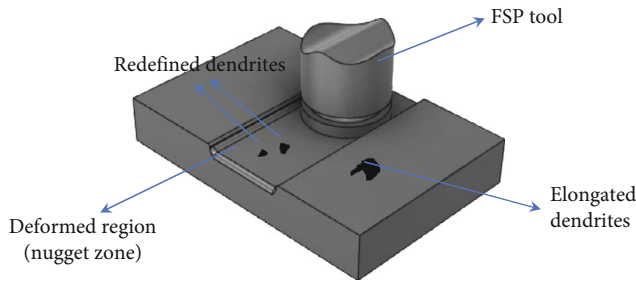
FIGURE 3: Visual representation of dendrite formation in FSW.

increased to 500 mm/min. As the welding speed was increased even further, a groove-like defect was discovered (750 mm/min). When friction stir welding was performed outside of the optimum conditions, mainly three types of defects were observed, which are listed below:

(1) Because of abnormal stirring, a cavity was produced

(2) Because of excessive heat input, flash was produced

(3) Because of insufficient heating, a groove-like defect was seen

The material worked upon was ADC12, and the tool was made up of SKD61. It was observed that when a high down-force of 14.2 kN was used, the range for obtaining the optimum conditions was enlarged [37].

It was seen that weld parameters highly influence defect formation. Defects are also caused due to irregular mixing or disproportionate input of heat. Unnecessary heat generation because of low advancing speed and high rotating speed makes workpiece material soft and can lead to several voids on the advancing side [38]. Considering the correlation between the transverse speed of FSW and the weld defects, if the transverse speed is more than recommended, then it is observed that defects are more prominent. The temperature increases along with it, which contributes to the formation of plastic deformation. The main reason for defect formation with increased transverse speed is that the tool experiences a low, stirring effect.

2.1.6. Tools Used in FSW. Three types of tools are present, namely,

(1) fixed: has a fixed probe length and can weld only a piece with a persistent thickness

(2) adjustable: the tool pin and the shoulder are separate, and any one of the two can be replaced if it breaks

(3) self-reacting comprising three pieces: the probe, shoulder, and bottom shoulder. Backing of the anvil is not required in this one (it is required in the other two types)

There are three types of shoulder end surfaces:

(1) Flat shoulder: the only disadvantage with this type of shoulder is that it does not trap the plasticized material fully, and the material spills out, creating excessive flash

(2) Concave shoulder: this type of shoulder is good at trapping flowing materials and has become quite popular

(3) Convex shoulder

The material friction can also be increased by choosing the suitable material for the shoulder.

Figure 4 demonstrates the flat bottom probe, which is currently most commonly used in the industry. One of the significant disadvantages of a flat bottom probe is that it increases forge force drastically compared to a doom-shaped probe. Tool wear in the flat bottom probe is much faster than the doom bottomed probe. The most optimum condition is that the doom bottom probe's radius is made to be 75% of the tool pin's total diameter. The chances of defect generation increase because of tool wear. The tool material mainly used affects the weld and grain structure quality [39].

The weld quality is dependent on the tool material properties like hardness, thermal conductivity, and fracture toughness. The reactive nature of the tool with air also plays a key role in influencing the joint formation quality [40]. The taper threaded pin tool projected comparably better tensile and flexural strength. On the other hand, the cylindrical threaded pin tool projected more impact strength and hardness [41].

2.1.7. General Properties and Features of Friction Stir Welding. When opposed to other systems, FSW is considered to be environmentally friendly. One of the main factors influencing heat input, material flow, and weld produced is tool geometry. It was noted that the size of the defect decreased as interference increased. The fluid flows into the spinning shoulder as the material's surface is welded on and the tool shoulders rub against each other. The amount of material flow directly depends on the load applied. Onion rings in FSW are formed because of the vertical movement of material and the pin-made material flow's geometric nature [42]. Thermal management plays a crucial role in the welding process. An ideal welding process has strong weld formation. The FSW process is a solid phase process due to which problems such as resolidification could be avoided. The FSW tool design determines the amount of heat transferred, and an ideal tool design must provide stable force for varying plunge depths. The friction stir spot welding method joins lightweight materials that are identical. Also, this process reduces energy consumption to a great extent [43]. To achieve a proper weld, the tool shoulder should always be in proper contact with the material's surface. The plasticized material will spill out if the shoulder even lifts upwards for a small time. Even at low welding speeds, one can achieve a proper weld [44]. Until the melting point, welding pressures are directly proportional to welding temperature. Also, up to a saturation point, the rotational speed directly influences welding temperature, after which the relationship between the two parameters becomes inversely proportional [45]. Friction stir welding processes are suitable for joining
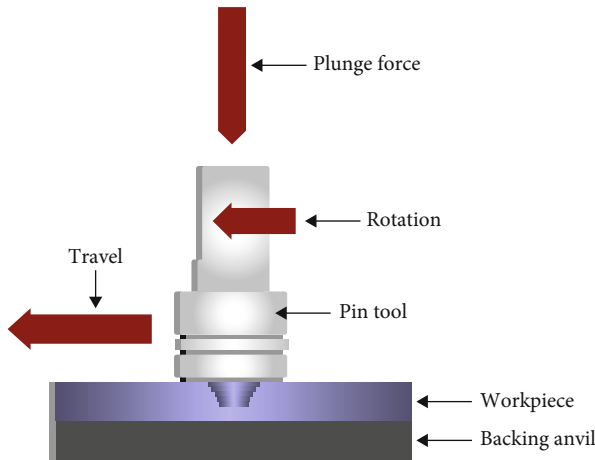
FIGURE 4: The flat bottom tool pin geometry.

alloys used for commercial purposes, such as aerospace. Below are the significant factors that affect weld output:

(1) The rotation rate of the tool

(2) Welding speed

(3) Spindle tilt angle

(4) Target depth

To generate lap joints and butt joints, the FSW process is a feasible and more accessible method. The weld regions of the weld are the nugget region (plastic deformation caused by high temp.), equiaxed recrystallized grains, uncrystallized grains, and the HAZ (heat-affected region). The high melting point and low ductility allow the FSW process to join metals like copper, titanium steel, and magnesium [46]. Hardness depends on the density of dislocation. The tensile strength of FSW copper joints improves the weld's quality compared to electron beam welding [47]. Balram et al. utilized three extraordinary welding boundaries: pivot speed, travel speed, and distance across the shoulder. The last estimation of an outstanding rigidity was 140 MPa. This exploration showed that the calculation of the instrument impacts the miniature underlying properties of the butt joints [48].

Deformation occurs in an equiaxed structure throughout the FSW process while the pin and nib slide along the weld seam. The faying surface tracer is used for tracking the metal flow, which flows in a helical manner due to the gyral motion where the nib and the pin are mounted. The welds' differences caused by the interactivity between the lower and upper regions cause the material to flow in clockwise and counterclockwise directions [49].

Typically, instruments face significant impacts during welding: rough wear, high temperature, and dynamic impacts. Hence, device materials have accompanying properties: high wear opposition, high-temperature strength, temper obstruction, and excellent durability. Two significant FSW instrument configuration fields are geometry and apparatus material [50].

In FSW, properties of the base metal, for example, yield strength, malleability, and hardness, influence the plastic progression of the material under the activity of pivoting nonconsumable apparatus [51].

Apparatus, rpm, and traverse speed offer equivalent impacts on the elastic attributes of the weld. An increment or decline in traverse speed produces more heat, hugely impacting the weld joint's mechanical characteristics [52].

The base material properties control the material stream during welding. The instrument and workpiece interface's contact conditions also autonomously influence base material properties; nonblemished joints might be delivered when the average sticking fraction is more than half [53].

3. Machine Learning Implementations in Friction Stir Casting

Machine learning is a part of computer science that makes computers act without being programmed. Machine learning algorithms can improve the algorithm's ability by increasing the number of samples. ML is now widely used in image recognition, medical, manufacturing, and many other fields. The applications of ML are now coming into use in manufacturing for quality and risk analysis. ML can predict any failures before manufacturers can go into production, preventing financial losses for companies. ML is now popularly being used to optimize the process of friction stir welding. Researchers and manufacturers are now using ML to foresee the weld quality and tensile strength based on process parameters.

3.1. An Overview of Support Vector Machines in Friction Stir Casting. Support vector machine (SVM) is an algorithm used to analyze data either for regression or classification [54]. The capacity of an SVM to achieve data classification patterns that are precise and consistent is its strength. An SVM decision function is a hyperplane used to classify observations into multiple classes based on feature patterns. The features used to infer the hyperplane are derivative data that have been interpolated during the feature selection stage. SVM entails balancing two complementary goals:

(1) Increasing the percentage of correct labels during the classifier's classification of new cases

(2) Ensuring that the classifier can accurately classify fresh data (i.e., improving its reproducibility) [55]

SVM permits a model's generalization capabilities to be maximized. The goal of the structural risk minimization principle is to permit a limit on a model's generalization error to be minimized [56]. The separating and maximum margin hyperplane, soft margin, and kernel function are the four key principles needed to grasp the notion of SVM classification. SVM can be applied to friction stir welding to solve tool misalignment and excessive flash problems. SVM can also be used to find the ultimate tensile strength over a multivariate set of input variables like rotational speed, welding speed, temperature, and various other factors affecting FSW quality. Manufacturers can identify the best possible parameters for FSW in different materials using SVM.

3.1.1. The Separating Hyperplane.
If the data is assumed, one can draw a line on a graph between the two classes. It is easy to categorize separation by a straight line, as shown in Figure 5.

The purpose of support vector machines (SVMs) is to place this hyperplane as far away from both classes' nearest members as feasible [57]. Support vectors are the data points that lie close to the separating hyperplane.

3.1.2. The Maximum Margin Hyperplane.
Many lines separate the two classes, but the nearest expression vector or the support vector is separated from the hyperplane by the maximum distance possible, defined as the margin. The role of the SVM is to maximize the said margin [54]. In FSW, the hyperplane may give us valuable insight into faulty and nonfaulty welds based on different input parameters. Multiple SVM can also determine the effect of all the factors on the final weld, which is also useful information to manufacturers.

3.1.3. The Soft Margin.
If the data chosen is assumed to be linearly separable by a straight hyperplane, it must allow for a certain soft margin that allows certain data to either side of the hyperplane without changing the outcome [54]. FSW may have certain anomalous data points due to defects in metals, although this problem can be overcome if a soft margin is applied to the SVM to account for anomalous data.

3.1.4. The Kernel Function.
Kernel functions are mainly used to increase computing efficiency; they also translate data into higher dimensions to fit different classification situations. As shown in Figure 6, a kernel matrix is used with dimensionality equal to the number of observations when classification is nonlinear. Using the kernel matrix over raw data allows the SVM classifier to train using the matrix, making it easier to achieve the required classification in linear and nonlinear classification scenarios [55]. Support vector machines are very flexible, which is of great use in welding; kernel functions can be picked based on the manufacturer's input parameters or which the researcher is experimenting with. With SVMs, one can opt for a multivariate study or focus on a single parameter as per our application.

3.1.5. Weld Classification Using SVM.
SVM is a tool best used in weld classification and can accurately predict the presence of defects in friction stir welds which proves invaluable to a manufacturer. SVM was able to classify friction stir welds with input of weld surface images [58]. SVM can also smartly be used in combination with ANN to classify as well as locate the defect in the weld and has proven to be one of the most successful in doing the same [59].

3.1.6. Temperature Prediction of Weld Pool Using SVM.
Maintaining the optimal temperature of the weld pool is essential to maintaining a strong weld. Using input parameters such as weld speed, tool rotational speed, and tool angle, an SVM is able to predict the maximum temperature a weld could reach which is invaluable information for a researcher or manufacturer [60]; a modified version of SVM known as LSVM is also used to obtain temperature signals of different frequency bands [61].

3.2. An Overview of Artificial Neural Networks in Friction Stir Casting.
A convolutional neural network takes in a tensor or a multidimensional matrix, which in most cases is an image, as seen in the input layer of Figure 7. In the case of FSW, the CNNs were trained to work with pictures of the weld pool, which were passed through multiple convolution filters, pooling, and activation functions. Each function can be considered a neural network layer, which takes in a tensor and returns another, passing through the subsequent layer. The convolution filters contain weights, which can be trained using backpropagation (gradient descent) [62]. These weights give the network a different perspective of the image, thus enabling it to gain more insights from the input tensor (an image in the case of most CNNs). The tensors, once parsed, are usually flattened or converted to a vector by reshaping them. This makes the data compatible with fully connected layers, containing dense neurons [63] and highlighting more specific details. The final layer, also known as the output layer, is usually a fully connected layer, on top of which sits an activation function, which is decided by the required output of the neural network.

3.2.1. Pooling Layers.
The goal of pooling layers is to compress the data of the input tensor in order to make it easier for the succeeding layer to parse it. This is usually done by max pooling or by average-pooling. The max pool layer takes the largest value in a given region and forwards it to the next layer, whereas the average-pool layer takes the mean value of every cell in the region to forward it to the next layer. The average/maximum values from the regions are arranged similarly to that of the input tensor to maintain logical consistency. This is especially useful in the case of FSW, as the resolution shot by the cameras mounted near the workspace was higher than required, which adds to the computational requirements with less gain in accuracy.

3.2.2. Activation Functions.
Following the universal approximation theorem [64], there is a need to introduce nonlinearity to approximate any relationship between the input and the output by looking at the training data. That is the role of the activation layers. They take inputs from the previous layer, which is usually linear, and transform them into nonlinear equations to better fit the real-world conditions.

3.2.3. Convolution Filters.
Convolution filters can be considered a weighted average over regions of a cuboidal region over the tensor, then placed in its corresponding location in the output tensor. The kernel function in Figure 8 acts on the central point of the image to produce the output. Multiple such images of the same dimension of the kernel are extracted from the previous tensor and arranged in the same order, which is decided by the stride. These weights are initialized to completely arbitrary values but are then trained using gradient descent to maximize the insights it can extract and store from the data. They are also called filters as they filter out parts of the image that do not resemble it as much as compared to the parts of the image that do match with it (which would result in a higher magnitude in the dot product and hence a higher intensity in that corresponding region in the tensor). In the case of FSW, the
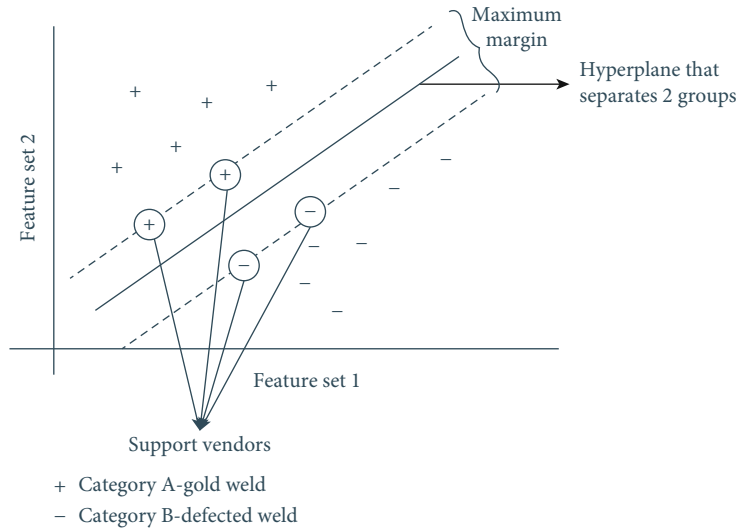
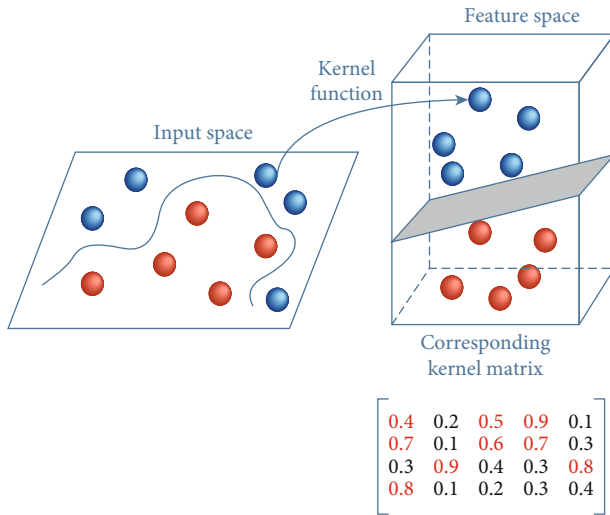FIGURE 5: Representation of the best hyperplane that separates the classes.

FIGURE 6: A representation of a kernel used to translate data to a high-dimensional feature space [55].

filters could resemble the shape of the weld bead or the seam. Shapes of cracks and crevices, common in welds that have gone wrong, could also be stored in the said filters. These filters can also associate a higher weight to a certain channel or color in an image to filter out features that are expressed better in those channels. Such filters could be used along with a thermal or infrared camera, which would give the model much data about the quality of the weld joint. A large number of these filters work in conjunction to filter out a large number of shapes and colors to gather insights before passing it to the next layer to derive more insights. The complexity of the filters and the shapes they look for keep increasing as the data travels deeper through the network, increasing the specificity of the insights gathered from the data while reducing ambiguity. The filters in Figure 9, also known as kernels, are 3×3 filters, as they contain 3 rows and columns. Increasing the filter size reduces the output image size unless

padding is applied. However, it allows CNN to gather more information and store larger patterns to search.

3.2.4. Fully Connected (FC) Layer. FC layers consist of neurons or perceptrons [65]. All neurons in a layer are connected to every other neuron in the previous layer. In other words, each neuron has a vector of weights associated with each of the neurons from the preceding layer. The outputs from the neurons belonging to the previous layers [66] are arranged to form a vector and a dot multiplied by the weights of every neuron. It can be looked at as a linear transformation, the corresponding matrix of which can be formed by aligning the weight vectors of each neuron. Finally, a biasing vector is added to the transformation to ensure more flexibility while training. Fully connected layers usually contain many weights, requiring much computational power to be applied.

Moreover, since all neurons are connected to all the other ones in the previous layer, it is difficult to identify patterns such as shapes and colors that might occur multiple times in the same image, as the weights are not shared. Hence, most CNNs use convolution filters to filter out required data and pass the specific insights over to the FC layers, which process the data further to get the required relationship between the input data and the output. These FC layers are used along with activation functions to perform certain tasks such as regression (predicting the mean tensile strength) or classification (classifying the penetration stage of the friction stir welding pin). When it comes to numerical data, like the rotational and the translational speed of the pin and/or the force applied by the pin, one can use fully connected layers alone to perform regression or classification. Such neural networks are known as fully connected neural networks (FCNNs).

3.2.5. UTS Prediction Using Neural Networks. Applying the UTS prediction neural network considering the input parameters includes rotational speed, translational speed, and the axial force as seen in the input layer from Figure 10. The overall tool fault diagnosis result obtained is 96%. Local discontinuity,
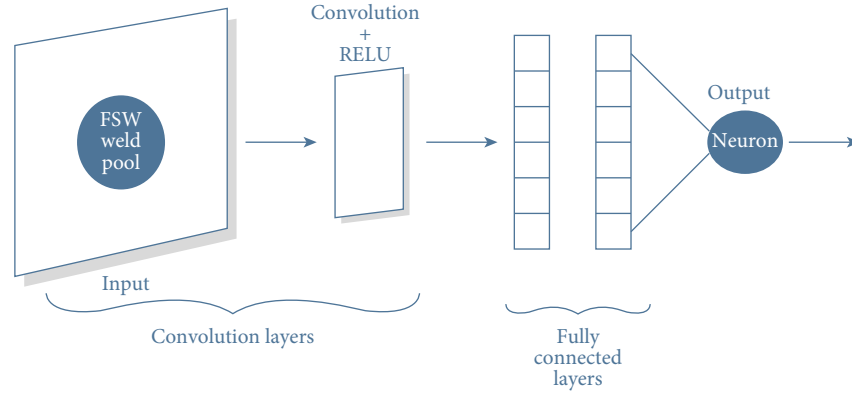
FIGURE 7: Representation of a convolutional neural network (CNN) and its layers.

FIGURE 8: A representation of convolution filters.

which is normal in plasticized materials, is also a root cause of equipment failure, as it provides enough hidden layers in the UTS output. The Bayesian neural network and decision tree classifier with three types of the input dataset with unprocessed welding parameters, numerical model, and computed parameters are inputs. Various other parameters are also taken under the FCNN input layer, including an axial force of the material and welding and rotational speed [67–69]. Every FCNN model was involved in backpropagation on both aspects, with the selected parameter under the input layer keeping tensile strength in the output layer. A fully connected neural network (FCNN) model was produced to investigate and reproduce the relationship between aluminum (Al) plates' FSW boundaries. The input boundaries of the model comprise weld speed and instrument turn speed (TRS). The FCNN model yields include the property boundaries: rigidity, yield strength, prolongation, and HAZ hardness [70]. Then, examining the created support vector machine (SVM) model with the artificial neural network (ANN) model and other basic regression models concludes that the forecast execution of SVR is better than any other general regression models and ANN. The proposed work is adjusted for effective use, progressively demonstrating friction stir welding measures [71].

3.2.6. Fault Detection with Convolution Neural Network. Then, the processing image was primarily performed with convolution neural networks, undergoing various feature extraction with filters. A model is able to detect the defects and faults in the FSW, which usually undergoes three processes, including the mathematical process of convolution, feature extraction to detect the defects at any corner, max pooling to reduce the dimension, the addition of fully connected layers, and then the softmax function to determine the probability of defects and faults [72]. The model determining the correlation

between rotational tool speeds, sample extracted position, and thermal data can be trained to obtain.

3.3. Influence of Regression and Classification Model on FSW Process. Various regression and classification models of sklearn are feasible for application to FSW. To determine and find out process parameters that majorly affect the FSW process, a commonly used supervised learning algorithm for analysis and classification includes logistic, multiple linear regression, SVM, ANNs, and ANFIS; ANFIS performed well to obtain more accuracy and minor error [73]. In the study, all the recent machine learning models were reviewed and compared in the hope of obtaining better results in FSW and FSSW process parameters which include ANNs and SVM.

Each model was compared to another, considering its material, process, input parameters, and DOE. Three distinct parts, like Pearson VII, polykernel (PK), and radial-based kernel function (RBF), are utilized with GPR and SVM regression [74].

The response surface technique is utilized to build up the regression model to foresee the elasticity of joints. The investigation of the different strategies is utilized to get to the ampleness of the created model. The results demonstrated that FSW of aluminum alloys at an apparatus pivot speed of $1050\,min^{-1}$, 40 mm/min welding velocity, and a shoulder width of 17.5 mm would deliver less imperfect joint with higher rigidity [75].

3.4. Fracture Detection with Decision Tree Model. The decision is a widely used supervised learning algorithm for classification problems. It is commonly used to detect fracture and predict UTS for FSW [59, 72]. It is based on the principle that it uses parent nodes to make decisions and leaf nodes for their output. Finally, it classifies or predicts the output with multiple branches.

Decision trees have been widely used next to ANN to determine tool failure and defects. Various parameters prone to effect include flow stress, temperature, torque, and strain rate. Peak temperature and traverse force have less influence on tool failure and can be considered input to the decision tree [67, 69]. The fracture location can also be easily determined with a decision tree with the lowest possible score of 0.5 [59]. Random forest built with multiple decision trees
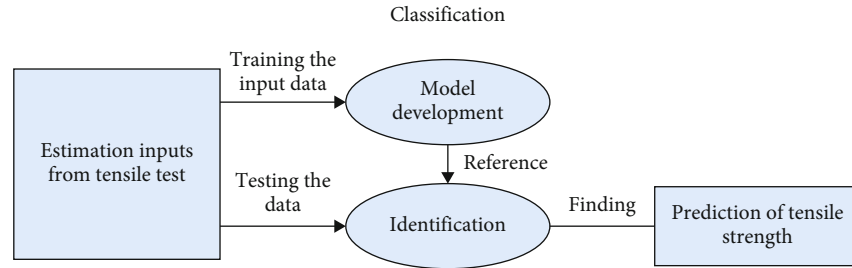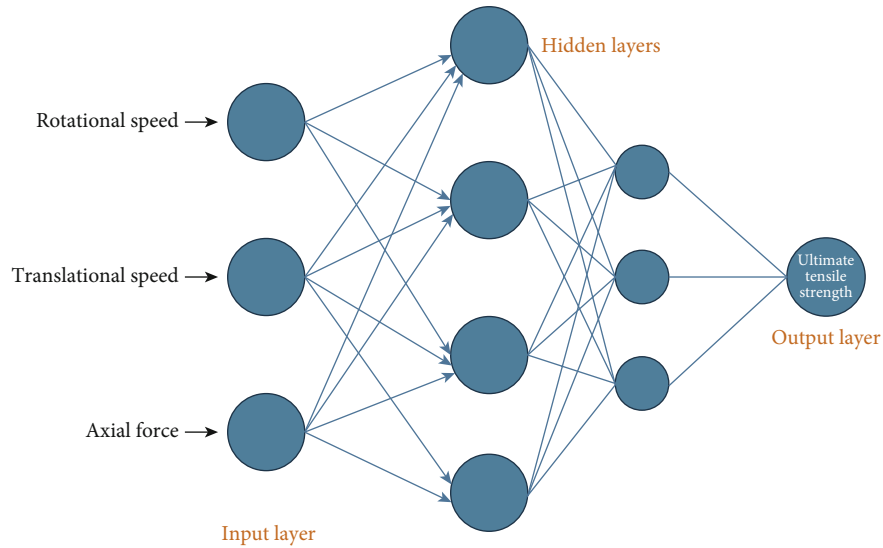
FIGURE 9: Summarisation for SVM prediction.

FIGURE 10: Graphical representation of the neural network used for UTS prediction.

was the best regarding accuracy and errors. Thus, random forest is preferred over decision trees in all cases.

3.5. Reinforcement Learning (RL) for Friction Stir Welding. Reinforcement learning is a method used to teach an agent (usually a neural network) how to perform a certain task by interacting with the environment. The agent is awarded points for going in the right direction, which could be completing subtasks by performing actions in the given set of actions (A) it can perform in its current state. Unwanted behavior results in the agent being given a penalty. The agent aims to reach the goal state while maximizing the reward and minimizing the penalty.

3.5.1. The Environment and the Agent. The RL model [76] consists of a set of discrete and abstract states, S, which is modeled after the dynamic environment that the agent will be interacting with. The elements in S describe the relevant information about the status of the environment and the agent. The agent can use this information to decide its interaction with the environment to achieve the desired state. It does so by performing actions from the set of actions, A. Each element in A is an abstraction of each agent's actions to interact and navigate through the environment. In the case of FSW, the environment contains everything in the

workspace that impacts the final output. It contains elements ranging from perception such as cameras and sensors to tools like the pin, which the agent will use to interact with the environment. As shown in Figure 11, there is a loop where the agent receives the state and a reward from the environment, which it uses to decide the action it has to perform.

3.5.2. Rewards and Penalties. While the agent learns how to reach the goal state(s) or perform a particular task, it requires guidance from the environment, which gives feedback in terms of rewards and penalties. Rewards provide positive reinforcement to encourage the agent to go on the right path. Penalties ensure that the agent does not engage in unwanted actions. The agent starts performing random actions while slowly modeling them by observing the feedback given by the environment. Its goal is to maximize the rewards while minimizing the penalties. Rewards and penalties are the means that the environment uses to communicate the agent's effectiveness. The user can suggest trends, such as penalizing low weld speeds, as it would affect the production time or reward stronger tensile strengths to increase the quality of the products.

3.5.3. The Markov Decision Process (MDP). The MDP [77] is the mathematical framework on top of which the models for the sets which describe the environmental states, actions,
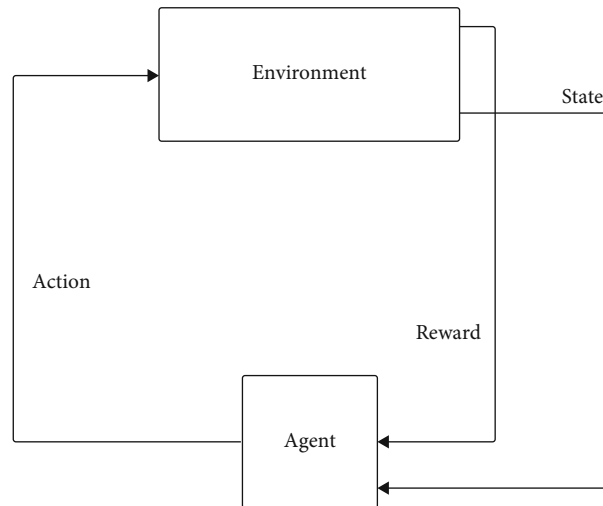
FIGURE 11: RL agent and environment.

rewards, and penalties were built. The MDP is also responsible for the proper training of the agent [76, 77]. It helps find the right trade-off between exploration and exploitation. This ensures that the agent does not exploit one set of actions to get the highest reward with the lowest possible penalty it has encountered in a training session. Rather, it forces the agent to explore the entire set of actions on the off chance that it could find a higher reward in the long run at the expense of a low cost for a short period.

3.6. UTS Prediction with Gaussian Process Regression. It is primarily used to forecast ultimate tensile strength (UTS) by using welding speed as an input parameter for training and constructing the model while eliminating tunnel errors and intermittent irregularities [78, 79]. This technique proves to be more effective in predicting ultimate tensile strength.

4. Results and Discussion

Table 2 depicts the use of various materials and process parameters for the friction stir welding process. The tools are given below: The fixed probe tool consists of a shoulder and probe as a single piece that can only weld at a constant thickness. The adjustable tool comprises a probe and shoulder independently. This allows the tool's free movement about the workpiece's thickness and length, and the probe material can vary and easily be replaced. The bobbin type tool comprises three parts, i.e., shoulder, probe, and bottom shoulder. The flexible probe length between the top and bottom shoulders allows the tool to work on workpieces with various gauge thickness joints [39].

The microstructural region zones are the following:

(i) Nugget region: consists of crystals with axes of almost similar lengths and recrystallized grains at high temperatures

(ii) Thermomechanically affected zone: uncrystallized grains formed at medium temperature

(iii) Heat-affected zone: precipitate coarsening takes place [46]

The formation of equiaxed, fine grains at the welding core and dynamic recrystallization increase fracture durability and mechanical properties. Pressure applied during welding directly influences the welding temperatures. The rotational speed is also directly proportional to welding temperature [45].

The shear friction factor can be considered for process analysis similar to the hot working of metals [82]. To stay away from overinfiltration, the length of the probe should be controlled. Simultaneously, the probe might not have complex calculations; thus, the heat produced from the probe is decreased. This implies that shoulder friction should be reimbursed by giving a significant part of the welding's warmth. Higher rotational and lower cross rates might be utilized for better outcomes, making up for the more significant energy losses. Likewise, the small size of workpieces implies that more critical consideration is expected to guarantee that clamping power does not cause deformation of the welded parts. The tensile strength of Al combinations has been lower than that in parent material for various kinds of welding devices. In contrast, formability has been discovered to be like the base material [83]. It is crucial to consider all the FSW parameters, including fixture clamping condition, traverse speed, shoulder immersing depth, spindle tilt angle, and tool rotation speed [84].

Table 3 depicts the variation in hardness of welds in the case of aluminum alloys w.r.t. tool design, welding speed, and grain size. 304 austenitic stainless steel and A 6056-T4 alloy were friction stir welded. Findings show an aluminum alloy sheet with deformed and stainless steel parts that have diffused. A persistent layer with less than 1-micrometer thickness was built in the middle of the mixed layer and the recrystallized Al alloy [9].

During FSW of A5083 and SS400, a pin rotation speed of 250 rpm, 25 mm/min welding, and a pin offset of 0.2 speed were optimum as they produced a sound weld [10].

TABLE 2: FSW process parameters for various materials used.

S.no.	Material	Rotational speed	Welding speed	Temperature	Tensile strength	Yield stress	Tool material/geometry	References
1	S70C (0.72 wt.% C)	100-800 rpm	25-40 mm/s	723°C	817 MPa	365 MPa	PCBN	[80]
2	1018 steel	450 to 650 revolutions per minute	1 to 4 in./min	Close to 1000°C	463 MPa	310 MPa	Molybdenum-based and tungsten-based alloy tool butt joint	[81]
3	AA2195 plates	240 rpm	2.36 mm/sec	<450°C	600 MPa	570 MPa	M2 tool steel	[7]
4	S12C—carbon steel with a small proportion of pearlite	400 rpm	100 mm/min	790°C	317 MPa	202 MPa	WC-based material	[46]
5	IF steel	400 rpm	100 mm/min	839°C	284	155	WC-based material	[46]
6	S35C	400 rpm	100 mm/min	653°C	533	327	WC-based material	[80]

TABLE 3: FSW grain size variation for different materials w.r.t. design and welding speed.

Type of material to be welded	Tool design	Welding speed (mm/min)	Grain size (μm)	Hardness (HV)	References
Dissimilar AA5083-AA6061 aluminum alloy	Square probe tool	60	6	76.3	[85]
Dissimilar AA5083-AA6061 aluminum alloy	Circular threaded probe tool	40	5	76.87	[85]
Dissimilar AA5083-AA6061 aluminum alloy	Circular threaded probe tool	31.5	6	82.13	[85]
Al5083 and AL6061	A cylindrical tool with tapered pin	45	5.0	43.26	[85]
AA2014-T6 aluminum alloy	Left hand threaded tapered cylindrical pin	40	12	125	[85]
AA2519-T87 aluminum alloy	Straight threaded cylindrical pin	30	17.5	78	[86]
AA2519-T87 aluminum alloy	Taper threaded cylindrical pin	30	15	80	[86]

Ti-6Al-4V friction stir butt-welded plates could be effectively welded using a tool characterized by a bigger tool diameter to plate thickness ratio and lesser shoulder to pin ratio [87].

The aluminum alloy 7050-T7451 was friction stir welded at 100-120 mm per minute traverse speed and 300-500 rotations per minute tool rotation speed. Because fine recrystallized grains formed in the weld nugget region, it was concluded that a sound weld was produced (i.e., dynamic recrystallization occurred). It was also discovered that there was partial recrystallization in the parent alloy's microstructure [14].

Metals joined in a pure solid state have more excellent mechanical properties than those joined in a liquid state [32]. Shoulder-pin-driven material flows are those in which the material is moved layer by layer by the pin. On the other hand, the material shifts as a whole by the shoulder [42].

An additional expansion of the prediction precision is prescribed to improve the preparation information's quality in future examination work. Identifying the pits in the welds utilized for preparing the CNNs through staged exhibit ultrasonics or registered tomography outputs could fundamentally build the precision. However, it will likewise significantly raise the expense for the weld examination. While assessing the exactness accomplished when utilizing artificial neural networks (ANNs), whether the welds were named consistently or fragment astute should be thought of [88].

Various machine learning approaches, including K-nearest neighbor, fuzzy KNN, and along artificial bee colony (ABC), have been used to predict the quality of FSW. Parameters for training the model include spindle rotational speed, plunge force, speed rate, ratio, and empirical force index (EFI). On testing between KNN and FKNN algorithm in which KNN outperformed the latter, by increasing the number of fold accuracy, it also increased accordingly. ABC was employed to improve the classification accuracy further [89].

Machine learning proves very effective in handling the failures and monitoring the FSW process for industrial needs. Various research and discussions were undergoing to improve its potential further and attain maximum out of it. In all sectors, automating processes and working with data help to improve process efficiency. Determining the process factors and causative variables for tool failure of FSW was put to an end by neural network models with high accuracy of over 96% [67].

TABLE 4: Application of various models on friction stir welding process.

Algorithm	Advantages	Disadvantages	Dataset, features, and outputs	Refs.
LeNET; a CNN consists of 5 convolution layers, with filter sizes of 28, 14, 10, and 5, with each layer consisting of 6 filters. Its task was to use the top view of the weld pool to predict the backside bead width as it could not be measured directly due to spatial limitations	(1) The CNN has large filter sizes that let it absorb a lot of generic data for the given features (2) Training will not require a lot of computational resources as it has very few layers	(1) Merely 6 filters per layer may not be enough to capture all the details from the training data (2) Very few filters of sizes 5 or less were used. This prevents the CNN from getting enough breathing space to learn the finer details from the images	Input: an image, which was the top view of the weld pool Output: a floating point number representing the back side bead width Training data: 22687 images Testing data: 2905 images Validation data: 2905 images	[91–93]
A CNN with 5 layers with filter sizes 5 in every layer. The convoluted layers had 8, 16, 32, 64, and 128 layers, respectively. On top of which 3 FC layers were stacked. The goal of the CNN was to classify the images into 3 different penetration stages, as it is difficult for a human to do so	The CNN is given more features to work with, i.e, the top and back side image, which makes the already easy task of classification easier	Batchnorm or dropout layers were not used. That coupled with an overwhelming number of parameters increases the odds of overfitting	Input: 2 images, the top view and the back side view of the welding pool Output: a one hot encoded vector of size 3, with each element representing the penetration stage in the weld pool. Training data: 20573 sets of 2 images Testing data: 2570 sets of 2 images Validation data: 2570 sets of 2 images	[93, 94]
To increase the signal-to-noise ratio (SNR) and fault probing depths of transient thermography inspection for aluminum sheets and FSW, a multilayer perceptron feedforward NN with feature extraction methods is used. Following data preparation, statistical moments, PCT, or TSLM (thermographic signal linear modeling) is used to extract features, and the data is then prepared for supervised learning. The classifying model is a Bernoulli distributed issue. Each network consists of ten input neurons, one hidden layer with a nonlinear activation function, and one neuron with a logistic activation function in the output layer	(1) The ML approach was successful in raising the flaw probing depth to 2.2 mm with a true positive rate of 97.2 percent (2) The CNN is provided with processed data using TSLM, which demonstrates an enhanced signal-to-noise ratio and a considerably more accurate classification at higher depths of the weld	(1) The training data provided for the FS sheet is of lower quality than that provided for the AL sheet; hence, the model appears to be less accurate in the FS sheet than in the AL sheet (2) The algorithm employed for weight optimization does not guarantee that the global minimum will be found	Input: 10 features Output: 0 or 1 where 0 represents no flaw and 1 represents a flaw. Training data: not provided Testing data: not provided Validation data: 30 points	[95]
An RL algorithm implementation using a 4 × 4 Q table. The goal of this algorithm was to identify the parameters, welding speed, and rotational speed to get the best possible outcome or the strongest possible weld. The weld quality was assessed with the help of 2 white LEDs whose light was reflected off of the welds to be captured by a CMOS sensor, the data from which was used to assess the quality and the strength of the weld pool	(1) The RL agent learns completely on its own, without any aid from human observers (2) The algorithm can be used to optimize both the quality of the welds and the process productivity, by penalizing slow welding speeds	The optimization problem was not solved in the most efficient manner. Maximizing the expected value of an infinite sum of random variables is a very time-consuming method	There is not a requirement for a dataset, as RL algorithms learn by exploring the environment. To assess which actions affect the state in order to get a higher reward, while minimizing the penalties. Which essentially means that the algorithm creates its own dataset	[68, 96]
RSM and a fuzzy logic model with four steps. Fuzzy logic is used to forecast weld quality in terms of UTS and percentage elongation	(1) The ANFIS is a hybrid system that combines the benefits of artificial neural networks and fuzzy inference methods. (2) The ANFIS has a number of benefits, including the ability to	(1) ANFIS is limited by the curse of dimensionality which is the increase in error with increase in features; therefore, the model	Input: 4 features; tool pin geometry, TRS, WS, and tool tilt angle Output: weld quality Training data: 31 data points	[58, 97]

TABLE 4: Continued.

Algorithm	Advantages	Disadvantages	Dataset, features, and outputs	Refs.
	capture a process' nonlinear structure, adaptability, and rapid learning capacity	cannot be employed if the programmer was to include a larger number of inputs	Testing data: 5 data points Validation data: 5 data points	
Wavelet analysis and support vector machines are used to classify friction stir welds using weld surface images. Gaussian and polynomial kernels are used in the SVM classification algorithm	(1) The employed Gaussian kernel is more accurate than more commonly used polynomial kernels (2) Multiple SVMs were used for the different outputs	(1) One of the method's drawbacks is that only the most obvious faults are detected during classification	Input: 3 features: TRS (rpm), WS (mm/s), and PD (mm) Output: energy, variance, and entropy Training data: 112 points for each good and defective weld, respectively Testing data: 70 data points from each of the above classes Validation data: 42 data points	[58, 98]
The most important elements influencing the performance of FSSW joints have been discovered. A four-factor, five-level factorial design matrix with a center composite rotatable matrix was used to account for the wide range of features. To determine the influencing nature and ideal condition of the process on TSFL, surface plots and contour plots, which are signals of probable factor independence, were created	(1) RSM is a powerful tool for fine-tuning FSW process parameters (2) The discrepancy in strength between the predicted and experimental values is only 5%. The RSM is more accurate in the modeling and optimization processes	(1) Physical examination of contour plots is required; hence, the method is not fully automated	Input: 4 features: TRS (rpm), PD (mm), plunge rate (mm/min), DT (s) Output: tensile shear fracture load (TSFL) Training data: 16 data points Testing data: 8 data points Validation data: 6 data points	[58, 98, 99]
A Gaussian process reduction (GPR) algorithm was used to predict the maximum temperature that the weld pool would reach	GPR algorithms are very efficient in terms of data required to train the algorithm. One can reach a very good fit without a lot of data. With very less data, the time required to train the model is very less, which makes it perfect for prototyping	The running time of the GPR algorithm scales cubically with n, where n is the number of data points. This makes it hard to work with large datasets	Input: 3 numerical features; weld speed, tool rotational speed, and tool angle Output: peak temperature Training data: 17 data points Testing data: not used Validation data: not used	[60]
An SVM was used to predict the maximum temperature that the weld pool would reach. The output from this would be used to control the parameters that affect the temperature (the input features) so as to reach the optimal temperature and keep the parent metals in the threshold where it stays in the plastic state to ensure the strongest possible weld	SVMs have the same run time irrespective of the size of the training dataset. Moreover, this class of algorithms is very effective when the data is scaled up to higher dimensions	The lack of data increases the chances of overfitting significantly, though it is difficult to overfit an SVM. This makes the reported accuracy unreliable as it may not have the same accuracy when used with a totally new data point	Input: 3 numerical features; weld speed, tool rotational speed, and tool angle Output: peak temperature Training data: 17 data points Testing data: not used Validation data: not used	[60]
A set of 3 FCNNs with 2 hidden layers, each containing 5, 10, and 20 neurons per layer. It was trained to predict the frictional power supplied by the tool by parsing the plunging load, rotational speed, tool diameter, and the interaction time	FCNNs or ANNs are very versatile and can learn to predict any sort of nonlinear relationship between the input features and the outputs. With a large set of hyperparameters to tune to ensure the best fit	One of the major drawbacks of the architectures used is that the increase in the number of neurons increases the accuracy of the prediction but at the expense of generalization. This could have been prevented by using dropouts or batch normalization layers in between the hidden layers	Input: 4 numerical features; plunging load, tool rotation speed, tool diameter, and the interaction time Output: frictional power Training data: 1800 data points, with the train-test-validation split not mentioned	[100]

TABLE 4: Continued.

Algorithm	Advantages	Disadvantages	Dataset, features, and outputs	Refs.
LBP is a descriptor used for classification in computer vision. It is able to distinguish texture and patterns from a particular image and can be used to detect defects in FSW	(1) The algorithm is very light and requires less computing (2) Can be used for real-time applications	(1) May not be able to detect all defects in the weld and still requires physical examination (2) It is unable to extract defects in larger samples	Input: 1 feature; grayscale image Output: region of defect highlighted	[101]
The method uses GLCM (gray-level coocurrence matrix), GLSZM (gray-level size zone matrix), and DWT (discrete wavelet transform) for feature extraction which is then passed through an ANN for automatic weld defect classification	(1) Feature extraction process is effective for higher-order textures (3 or more pixels) (2) The implemented smart system helps NDT experts classify weld defects easily		Input: image of the weld Output: classified defect name Training data: 500 images Testing data: 250 images Validation data: not mentioned	[102]
This method first sends the data through an SVM and then through a naïve Bayes classification algorithm; furthermore, a decision tree algorithm is applied to the dataset, and the dataset is subjected to a random forest algorithm and finally an ANN to detect the fracture location in dissimilar friction welded joint	(1) Study proves that ANN is most successful in classification of defect location when the same dataset is provided to all algorithms	(1) Lack of data to improve the accuracy of the algorithm	Input: 4 features; upset pressure (MPa), tool rotational speed (rpm), burn off length (mm), and friction pressure (MPa). Output: 0 for fracture at upper location and 1 is for fracture at weld Training data: 22 data points Testing data: not mentioned Validation data: validation dataset not used	[59]
In this method wavelet, a packet is used to obtain the temperature signal components of different frequency bands; then, a LSVM (least squares linear system as a loss function) is used as model for classification and parameter optimization using genetic algorithm	(1) Able to replace the inequality restraints of an SVM which helps in learning fast (2) Very efficient for large scale and less complex problems	(1) Inefficient for large scale and less complex problems	Input: 3 features; temperature, rotational speed, and transverse speed Output: one of three classes of coefficient of strength greater than 75%, between 65% and 75%, and less than 65% Training data: 16 data points Testing data: not mentioned Validation data: validation dataset not used	[61]
A logistic model tree (LMT), which took in statistical data corresponding to vibrations from an accelerometer to classify the weld into three classes, namely, good, broken, and air bubble	(1) LMTs combine the best of logistic regression and decision trees to give an accurate model (2) Since there are a lot of features out of which many could be useless, the decision tree part of the algorithm helps in feature selection, while the logistic regression part does the classification	The calculations done to get the weights and trees are very complex, which causes a small change in data to drastically modify the architecture, or the output, which ends up wasting a lot of computing resources	Input: statistical data from the sequential readings from an accelerometer; the mean, median, mode, standard deviation, skewness, variance, maximum, minimum, and count. Output: a probability distribution of the weld being good, broken, or having an air bubble	[103, 104]
AlexNet, a CNN containing 5 layers of filter sizes 5, 3, 3, 3, and 3 with the number of filters being 96, 256, 384, 384, and 256, which took in images from the weld pool and classified them into two classes; ok and not ok	(1) The CNN, which takes image data, gets a very clear idea of the defects. There is a strong correlation between the image of the weld and whether the weld is broken. (2) The CNN outperforms (93.64% accuracy) the RNN (82.72% accuracy), which took in other features to try predicting the same thing	(1) The network, being very complex, requires much time to train. It needed 3 hours to complete 60 epochs of data (2) The lack of batch normalization and dropouts reduces the generalization power of the network	Input: image of the weld pool Output: probability distribution of the weld being ok and not. Training data: 144 data points Testing data: 110 data points	[105, 106]

TABLE 4: Continued.

Algorithm	Advantages	Disadvantages	Dataset, features, and outputs	Refs.
A decision tree was trained to predict the probability of the tool breaking by looking at the raw variables corresponding to the welding process. After which, the tree was dissected to identify the parameters that most affected the tool's chances of breaking and how	The decision tree, albeit a bad model for classification or regression, is very good at identifying and exploiting features which most affect the outcome. While this is one of its major drawbacks in most cases, here, it helps us	The decision tree does not have an understanding of the system behind friction stir welding; rather, it looks at the information gain and the correlation which might not imply causality. Hence, small changes in the training dataset may drastically change the architecture of the tree, thus changing the most important parameter	Input: welding and rotational speeds, tilt angle, axial pressure, shoulder, and pin Radius, plate thickness, and the work plate material properties of thermal diffusivity Output: the probability of the pin breaking	[67]

Table 4 shows the application of various models to the friction stir welding process. The optimum parameter with the required value of UTS avoids any defects. The tensor flow model's prediction of UTS with rotational speed has been researched by considering mean squared error of loss and stochastic gradient descent as an optimizer [67]. Later improvement with the same resulted in better results with ANN (artificial neural network) with sigmoid activation function. In contrast, it depends on significant factors, including tool rotational speed, axial force, and tool traverse speed [68]. This model and the superior result reduce the cost and time of the experiment compared to others. Moreover, Verma et al. analyzed the results by comparing the predicted UTS value vs. the actual UTS value; UTS prediction I FSW are shown in Table 5. When all three techniques were computed, the Gaussian process regression method depicted lower variation between predicted and actual values [90].

To predict parameters such as yield strength and hardness of AA 7075-T6 joints, Maleki et al. implemented the backward propagation method. Under the neural network architecture, the input layers, in this case, are rotational speed, axial force, welding speed, shoulder diameter, pin diameter, and tool hardness. The hidden layers compute to give out output layer data: yield strength, hardness, and tensile strength [90], as shown in Figure 12. The framework has four layers: input, hidden, and output layers. The first layer comprises every one of the information factors. Data from the first layer is prepared throughout two hidden layers, following the output vector figured in the last layer [114].

Deeply discuss the correlation between process parameters and mechanical properties of the welded AA5754 aluminum plates with a simulation model using an artificial neural network. The parameter considered for training the model includes 1. Tool rotational and travel speed is 2. Position of the sample extracted is 3. Thermal data of various analyses, including visual and tensile tests, have been performed to evaluate the effects on output parameters. By implementing the mean absolute percentage error, the outputs for microlevel hardness were obtained to be 0.29%, and that for UTS was found to be 9.57%. This clearly defines the way to analyze, predict, and control all manufacturing technologies [115].

The ANN model has been developed and backpropagation on both aspects with the selected parameter under the input layer, keeping tensile strength in the output layer. The model can also predict the tensile strength of the welded aluminum with the mentioned parameters. The model proved high accuracy and results an R value of 0.99954 for test samples after iterating for 1000 epochs [116].

The tensile test can also be predicted using the SVM model (Figure 9) where the input data is trained, and further, a model is developed which undergoes testing to predict or interpret the final results. Armansyah et al. found that 100% accuracy is obtained w.r.t. to predict tensile strength using the SVM method [117].

Artificial neural networks can also predict the grain size of the joints. To determine the grain size, input parameters considered are temperature, Zener-Hollomon, strain, and strain rate [118]. The end goal is to train the neural network

to have calculated results, and predicted results must have less variation. This, in turn, allows determining joint resistance. The following technique proved to apply on lap, T, and butt joints.

Analyzing the effects of welding speed on UTS with the Gaussian process proved that an increase in welding speed makes prediction more erroneous. It also proved most potent in predicting UTS than the other models [79].

Table 6 indicates that machine learning has impacted almost all welding processes and has made them more efficient in various aspects. Researches were made to monitor defects' early detection and control in friction stir welding. Processing the defective images and extracting the feature lead to proper classification with SVM (support vector machine). On the topic of research, the relationship between surface appearance and tensile strength indicates that weld joints with irregular and imperfect surfaces have substantially lower tensile strength. SVM could predict and classify the good one with greater than 95.8% accuracy [119, 125, 126]. Monitoring helps in reducing defects, especially in aluminum fabrication. SVM strategy is utilized as the example arrangement procedure that measures the similitudes between input information and the information put away in the information base. The entire expectation framework has two significant stages: the training and testing stages. Complete accuracy for every test and training system was observed [112].

Apart from this, image pyramids and image reconstructions were used to analyze the defects on various weld samples. Convolution neural networks are proved to be another best model for detecting defective vs. nondefective welds by processing their images on the production line [72]. This system obtains better results in both offline and online monitoring processes. Extension rate and normal fracture strength of given mechanical segment and image processing calculations can be easily applied for defect identification in the mechanical segments [127].

Various results show that inappropriate setting of rotational speed and other parameters may lead to increased flash formation and surface galling. Bayesian optimization helps obtain better parameters easily with the multitask approach than the single-task one [68].

Flaw detection in FSW with transient thermography has become popular in recent times. Transient thermography was processed using thermographic signal linear modeling (TSLM) and feature extraction without hyperparameters. Researches have been done with halogen lamps of 1600 W with appropriate mounting and orientation for IR thermography experiment in the reflection mode [95]. It clearly shows that NDTE and TSLM techniques could also improve the subsurface flaw-probing depth to further mm.

When FSW considered utilizing computerized reasoning strategies were inspected, it was also noted that more than 81% of the utilized materials were aluminum composites, and 23% were made with divergent materials [117].

Various ML models were investigated to implement optimized techniques to determine the relation between tensile strength and FSW parameters by testing austenitic stainless steel and Ck 45 steel using the ANN method. Studies

TABLE 5: ML models used in UTS prediction.

Algorithm	Advantages	Disadvantages	Dataset, features, and outputs	Refs.
A random forest algorithm to approximate the ultimate tensile strength (UTS) of the weld. It works by building a decision tree, which is a tree in which each node performs a test on the input data to perform the required task (in this case, regression). The model was observed to have a test error (root mean squared error; RMSE) of 6.72, against mPa as a label	The random algorithm is known for being able to gather the most insights from the least possible data without preprocessing or data augmentation. This algorithm performs the best out of the three algorithms (random tree, artificial neural network, and M5P algorithm) that were trained using the same dataset	This performance with lesser data does not scale up. In other words, the performance of the algorithm does not improve with more data. It makes the decision tree more complex and hence takes a longer time to run but fails to increase the model accuracy by a margin as large as a neural network would	Input: 2 floating point numbers; the rotational speed and the feed rate Output: floating point number, the ultimate tensile strength of the weld (in mPa) Training data: 20 data points Testing data: 6 data points Validation data: no validation dataset used	[107–109]
An FCNN consisting of a single hidden fully connected layer. The input features it takes are floating point numbers, which are arranged in a vector. Hence, convolution filters were not required. The model was observed to have a test error (root mean squared error; RMSE) of 20.59, against mPa as a label	The number of neurons, hence weights remain constant throughout the training process. Adding more data will not change that	The meagre dataset provided is not enough to train all the weights in one pass (epoch). A large number of epochs are required, and with such little data, the chances of the model weights getting overfit are very high. Which will significantly reduce its accuracy with new data	Input: 2 floating point numbers; the rotational speed and the feed rate Output: floating point number, the ultimate tensile strength of the weld (in mPa) Training data: 20 data points Testing data: 6 data points Validation data: no validation dataset used	[66, 108, 109]
The M5P algorithm, which works on the M5 tree building algorithm, which is very close to how decision trees work. It differs from a normal decision tree in the sense that it performs linear regression with the leaf nodes of the tree, which in theory takes the best of both worlds. The model was observed to have a test error (root mean squared error; RMSE) of 32.03, against mPa as a label	(1) It takes the best of both decision trees and linear regression models, by attaching weights to the leaves of the decision tree (2) The M5 model tree adds extra steps, namely, pruning and smoothing, to make the model more robust	The data provided is not enough for the M5P model to gather enough insights. The pruning and smoothing steps were not as efficient as they would be if more data were to be provided. The regression layer also suffers from a lack of data to train its weights	Input: 2 floating point numbers; the rotational speed and the feed rate Output: floating point number, the ultimate tensile strength of the weld (in mPa) Training data: 20 data points Testing data: 6 data points Validation data: no validation dataset used	[108–110]
A linear regression algorithm was used, which in essence is a single neuron, i.e., perceptron from a FC layer, which is generally used in neural networks. It has a vector of weights and a bias which it adds to	(1) It is a very light model which will require very less time and computing resources to train. Moreover, since it has very few weights, it is very difficult to overfit	(1) It has very few weights, which may not capture the relation in the best way (2) The model is linear and hence will falter when	Input: 3 floating point numbers, namely, the rotational speed, welding speed, and axial	[65, 111]

TABLE 5: Continued.

| Algorithm	Advantages	Disadvantages	Dataset, features, and outputs	Refs.
return a single floating point value. It is capable of executing regression and binary classification. In this case, it has been used to perform regression	(2) This is the most fitting model for the task as the experiment has only 45 data points, out of which only 32 are being used to train it	made to predict nonlinear relations between the input(s) and the output with the best accuracy possible	force Output: ultimate tensile strength (in mPa) Training data: 32 data points Testing data: 13 data points Validation data: validation dataset not used	[54, 112]
An SVM-based regression algorithm, which accepts the rotational speed and the travel speed (welding speed) of the pin, as well as the category of the material used, which are classified by tensile strength, namely, low tensile strength (LTS) and high tensile strength (HTS). These features were used to predict the mean tensile strength of the weld	SVMs excel in predicting simple relations between a small number of features. Moreover, SVMs do not require as much computational resources as a neural network to learn the relationship between the input and the output data	(1) The testing dataset size was not large enough to assess the accuracy of the model in the best possible way (2) The training data was not ample, which caused the model to overfit	Input: 3 features; rotational, travel speed of the pin, and classification by tensile strength Output: one floating point number depicting the mean tensile strength of the weld Training data: 12 data points (6 from each class) Testing data: 9 data points (6 for LTS and 3 representing HTS) Validation data: validation data not used	
Feedforward neural network combined with a BP algorithm and multiobjective particle swarm optimization Backpropagation is employed in this study with a network that has a two-neuron input layer for each input factor and a one-neuron output layer	(1) Sensitivity analysis was performed to find the parameters that need to be accurately measured. The findings of such an analysis would also provide useful information on the model parameters' "robustness," allowing for better decision-making	(1) In this method, the optimal activation function and the number of neurons were found using a hit and trial method which is time-consuming	Input: 2 features; TRS (rpm) and WS (mm/s) Output: ultimate tensile strength and hardness Training data: 20 data points Testing data: 5 data points Validation data: validation data not used	[58, 113]
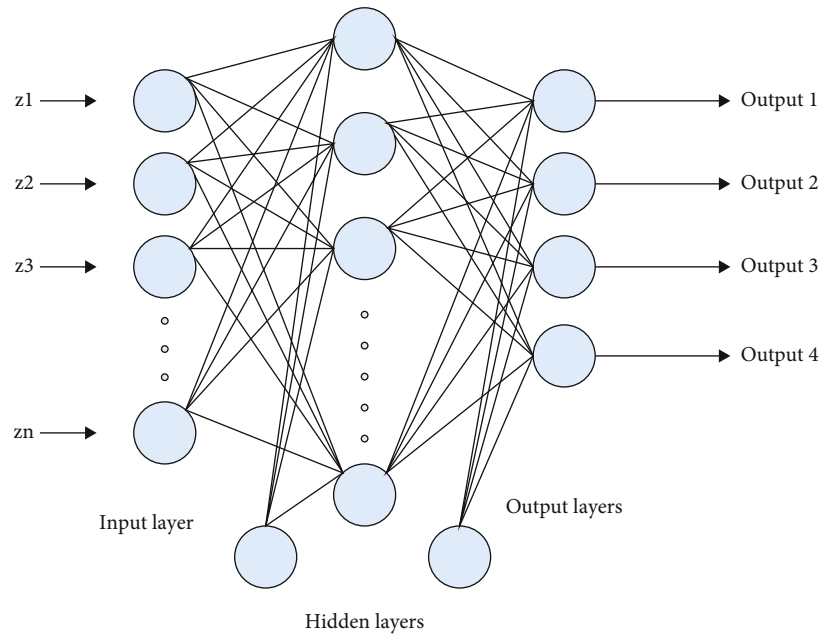
FIGURE 12: Artificial neural network architecture.

TABLE 6: Impact of machine learning on different types of welding process.

Welding type	Target output	Machine learning model used	Parameter considered	Result	Ref.
Resistance spot welding	Prediction of nugget size	Finite element model (FEM)	Temperatures, electrode displacements, and electric voltage drops	Much simulation offers many labeled data due to physics-based simulation data. There is no measurement sensor precision issue in data generated from a simulation model	[119]
Laser beam welding	During deep penetration welding, the capillary depth of the keyhole is calculated	Optical coherence tomography (OCT)	Laser power, welding speed, radiation mode, angle of incidence	Discrete wavelet transformation helps to classify the weld quality based on height parameter, and low pass Chebyshev helps in reducing the amount of process	[120, 121]
Resistance spot welding [52]	Nugget width prediction	Deep neural network	Strength and weight of the material, cost, manufacturability, and environmentalism	In RSW joints, DNN has much higher precision and lower variation for nugget width prediction because RSW data is highly nonlinear and complex	[122]
Ultrasonic metal welding	Prediction of USMW joint strength	Regression model, ANN, and ANFIS	Weld pressure, vibration amplitude, weld time, and T-peel failure loads	The number of neurons, learning rate, and momentum facture of the ANN set to 10, 0.08, and 0.6. MFs are also counted as 4 in the ANFIS system	[123]
Submerged arc welding	Prediction of transient temperature in SAW process	Multilinear regression	Voltage, trolley speed, stick out, current	Proves to be the robust method since computation cost was less and the estimation of transient temperature was good. Also, help to a swift resolution of several decision processes	[124]

A -> best case scenario for KNN
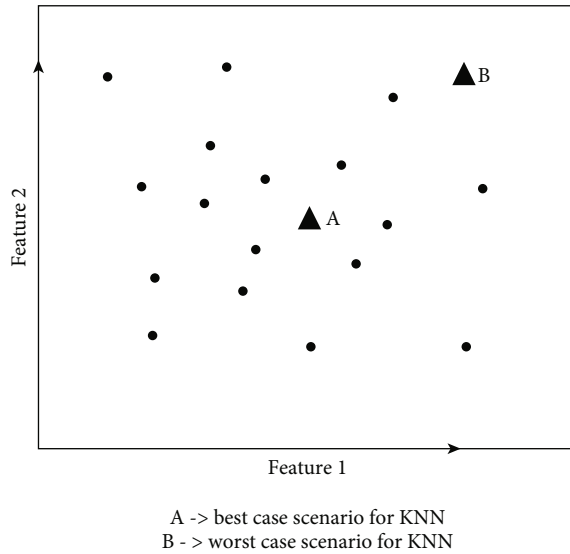B - > worst case scenario for KNN

FIGURE 13: The best- and worst-case scenarios in a KNN.

from Celik et al. investigations showed less variation between the actual and predicted values of UTS [90].

Mishra et al. carried out the convolutional neural network to recognize conventional welded joints and friction stir welded joints' surfaces. Macias et al. set up a connection between the acoustic discharge signals and the difference principle boundaries of friction stir welding measure dependent on artificial neural networks (ANNs) prepared on Levenberg-Marquardt calculation [90].

5. Summary

Machine learning (ML), which is a subset of data science, is a set of algorithms using which a machine learns to predict the outcomes of specific situations with the help of historical data. Unlike expert systems [128], the programmer programs if-else conditions based on their knowledge. It does so without being explicitly programmed to do so. ML applications are used when the relationship between the input and the output is not that clear or too convoluted to spend time designing the system, to derive the formulae and the rules that will govern the expert system.

ML algorithms learn to predict relationships by looking at previous data to build a model of how the relationship might look based on it. Unlike expert systems, they are seldom 100% accurate, but an experienced ML engineer can bring the model's accuracy close to 100% by tuning model hyperparameters, feature engineering, etc.

In friction stir welding (FSW), ML is used to optimize the process by various means. One of the most common use cases of ML in FSW was quality assurance. Various models such as convoluted neural networks (CNNs), support vector machines (SVMs), random forests were given data from cameras (top and side views of the weld pool), or numeric data related to the probe (rotational and translational speeds), in order to predict the bead width, penetration stage, or the tensile strength. These outputs make the process of quality checking and assurance easier and safer.

The insights can be used to improve productivity and quality in factories or parse and slot into a control system to perform minor adjustments on the go to get a better weld.

ML was also used to identify the parameters that most affect the FSW process. This was done using reinforcement learning, in which the agent was trained to set the rotational and translational speeds of the probe to get the best possible weld in terms of time taken and strength. Using RSM (response surface methodology), other implementations were done to find the elements that most impacted the FSW joints. Another use case was dimensionality reduction. The sensor outputs, which were multimodal or multidimensional, had to be reduced before they could be passed through a control system. This was done inherently in many neural networks.

Recent studies in machine learning have proved that it can be used effectively in FSW and is becoming a more popular technique. Researchers are now able to predict responses based on the input parameters. This can be of great help as it gives us an insight into weld quality and strength before the manufacturing process. It can help manufacturers and researchers save materials and time by working on the weld with the insights provided by machine learning algorithms. This review is aimed at gathering all the techniques and algorithms suggested by researchers in one platform.

6. Conclusions

6.1. Open Issues

(i) If we talk of the benefits friction stir welding process gives over the conventional processes, there are many:

   (a) In FSW, the heat input is comparatively lower than in other processes. This reduces the loss of mechanical properties

   (b) Cracking and porosity problems are not experienced in this process as it is a solid-state welding process

   (c) No further machining is required after this process on the weld

   (d) A filler material is not required in this process

Some of the disadvantages include the following:

(a) It is a relatively slow process

(b) This process can only use weld materials that have a low melting point and low strength

(c) Metals joined in a pure solid state have greater mechanical properties than those joined in a liquid state

(d) With a rise in tool velocity, the hardness profile of weld smoothers out near the nugget zone. As a result, inhomogeneity is reduced. Creating a thermal model of FSW also allows the user to predict the weakened zones in the weld

(e) If the right FSW parameters are used, defects such as tunnel, porosity, defective tightness, large or tiny void, surface groove, and crack-like root-flaw can be avoided entirely

(f) Local discontinuity is common in the flow area of plasticized materials and a common cause of equipment failure

(g) The grain size of the microstructure was also found to increase with temperature

(h) Thermal models support the theory that the connection between temperature and energy levels is normal for aluminum alloys that ration comparable thermal diffusivities

(i) Shearing of the grains is in the same direction as pin rotation, and the pin deformation is more on the retreating side because of the clockwise rotation of the tool

(j) The rotational speed significantly differs in the welds' tensile strength formed during FSW of dissimilar alloys. An intermetallic compound is formed between the dissimilar alloys during the welding process. The weld in the middle was exceptionally tough. The compound was formed due to the constitutional liquidation of dismissal alloys during FSW

(k) Prediction of UTS and detection of faults for FSW were cumbersome due to more experimental observations and readings, making it difficult to process on a small scale

(l) The accuracy obtained at the result without any defect was less for FSW of AL and other metals when compared to a few other welding types on the same metal [129]

6.2. ML Algorithms. In FSW, ML is used primarily to achieve 2 outcomes, namely, better-quality assurance and automation. Quality checking was done mainly using CNNs, which worked with photos of the weld pool to ensure the absence of defects. The neural networks also took in nominal and numerical data like the material type or the rotational and translational speed of the FSW pin.

CNNs were used to get the back bead width, the weld's penetration stage using the LeNET architecture, and a generic CNN with 5 layers containing 5×5 shaped filters with the number of filters increasing with depth. Their performances cannot be compared as each model's assigned tasks were different. However, by looking at the performance of the generic CNN, which had a test accuracy of 97.5%, we could surmise that the bead width could have been predicted with much more ease. The dataset was abundant, but the LeNET architecture lacked the required number of filters to learn all the features from the input image that would affect the output.

CNNs were also used for process control. The input from the camera was used to predict the penetration state and hence feed a control system command to control the pin speed and transversal speed to get the best possible UTS. The input from the camera was also used to find defects in the weld, which drastically changes the way quality checking and assurance can be done, thus improving the quality of the end product.

SVM can be effectively implemented to predict the maximum temperature of the weld pool using various input parameters which is an essential quality control measure to maintain the best weld quality.

Another vital implementation of ML in FSW was ultimate tensile strength (UTS) prediction. Various regression models include support vector machines, random forests, and fully connected neural networks. One common issue that all the implementations of UTS prediction had was the lack of data, lacking a validation dataset. Training ML algorithms with fewer data will lead to the weights getting overfitted, which will make the algorithm less accurate when given inputs that are not from the training dataset. The lack of a validation dataset also reduces the potency of the model in the face of previously unseen inputs, as the job of the validation dataset is to eliminate the programmer's bias while setting the model's hyperparameters. The problems above make the reported accuracy metrics slightly unreliable.

However, in the given models, the highest accuracy was observed in the random forest algorithm. Closely following the FCNN and the SVM would easily outperform the random forest giving more data to gather insights and train its weights. Moreover, the least performing algorithm turned out to be the decision tree as the hierarchy of the attributes checked is highly dependent on information gain, which may cause it to neglect some other attributes that seem insignificant when viewed through the scope of the information gain, and the k-nearest neighbor's algorithm, as it does not learn anything but rather memorizes the input data, which makes it perform very badly in edge cases as in Figure 13 and relatively better when the input data is closer to a lot of the data points. While all the models that predict the UTS will benefit from more data, the rate of accuracy will reduce irrespective of the more data being appended for most of them. The only models which will benefit from a large dataset would be the FCNNs, as they have the most versatile and robust training loops.

One of the main problems, i.e., the lack of data, can be remedied by using self-learning or exploratory algorithms, such as RL and Bayesian estimation. As highlighted, these algorithms do not require data to learn the relationship between the input and the output, making it more effective. RL has been used to set parameters such as the rotational and translational speed of the FSW pin to get the best possible weld in the least possible time. These algorithms had very high accuracy rates, which can be trusted as the algorithms explore the environment extensively to gain mastery over controlling the agent in the best possible way. The only issue with RL or other self-learning algorithms is that the computational resources and the time required are very high.

Another limitation in most models is that they are not given information about temperature. Temperature control is critical in FSW, as the metals have to be heated to the right temperature where they are in a plastic state and do not melt

so that they can be cast together with the least possible deformities and reductions to their ultimate tensile strength. The RL model seemed to pick up on that on its own. Nevertheless, input from a thermal camera during the welding process to all the previously mentioned models will drastically increase performances.

6.3. *Future Directions.* One can avoid void formation of FSW with proper tuning of weld parameters such as temperature, torque, and shear stress. The abovementioned decision tree model already proved 90% accuracy [69]. By further extending the dataset size and preprocessing with proper handling of class imbalance, one could obtain more accuracy around 96% by considering the parameters concerning tool pin-like temperature, torque, and maximum shear stress [90]. This will ultimately reduce the chances of defective welds to less than 4%. Recent development in ANN and image processing techniques significantly reduce the cost and time of FSW [67].

Further improvement in the model will provide the scope of making the process much feasible and faster. Rotational, feed rate, and travel speed were the major parameters that affect the ultimate tensile strength. The correlation occurs between the tensile strength and the surface appearance and relates to various input and process parameters [67, 78, 125]. It proves over 95 percent accurate in detecting the good and bad welds, which helps greater in online monitoring and feedback systems [58, 103, 131–136]. A modified LSVM is also being used to obtain temperature signals of different frequency bands another step towards obtaining more data to check what predicts better weld quality [137–144]. SVMs were also used in combination with ANN to classify weld defects as well as locate them when provided with surface weld images which can serve useful to automate quality control. In the application of SVM methods to FSW process characteristics, there is a lot of room for improvement and research gaps. There is a great demand for machine learning algorithms to forecast the behavior of process parameters in FSW, based on the knowledge of prescribed work.

Data Availability

All the data is included in the manuscript.

Conflicts of Interest

The authors declare no conflict of interest.

References

[1] V. P. Singh, S. K. Patel, A. Ranjan, and B. Kuriachen, "Recent research progress in solid state friction-stir welding of aluminium- magnesium alloys: a critical review," *Journal of Materials Research and Technology*, vol. 9, no. 3, pp. 6217–6256, 2020.

[2] F. Nadeau, B. Thériault, and M. O. Gagné, "Machine learning models applied to friction stir welding defect index using multiple joint configurations and alloys," *Proceedings of the Institution of Mechanical Engineers, Part L: Journal of Materials: Design and Applications*, vol. 234, no. 5, pp. 752–765, 2020.

[3] K. Jata and S. Semiatin, "Continuous dynamic recrystallization during friction stir welding of high strength aluminum alloys," *Air force research lab wright-patterson afb oh materials and manufacturing directorate*, vol. 43, no. 8, pp. 743–749, 2000.

[4] R. S. Mishra and Z. Y. Ma, "Friction stir welding and processing," *Materials Science And Engineering: R: Reports*, vol. 50, no. 1-2, pp. 1–78, 2005.

[5] G. K. Padhy, C. S. Wu, and S. Gao, "Friction stir based welding and processing technologies - processes, parameters, microstructures and applications: a review," *Journal of Materials Science & Technology*, vol. 34, no. 1, pp. 1–38, 2018.

[6] G. Çam and S. Mistikoglu, "Recent developments in friction stir welding of Al-alloys," *Journal of Materials Engineering and Performance*, vol. 23, no. 6, pp. 1936–1953, 2014.

[7] Y. J. Chao, X. Qi, and W. Tang, "Heat transfer in friction stir welding—experimental and numerical studies," *Journal of Manufacturing Science and Engineering*, vol. 125, no. 1, pp. 138–145, 2003.

[8] P. L. Threadgill, A. J. Leonard, H. R. Shercliff, and P. J. Withers, "Friction stir welding of aluminium alloys," *International Materials Reviews*, vol. 54, no. 2, pp. 49–93, 2009.

[9] W.-B. Lee, M. Schmuecker, U. A. Mercardo, G. Biallas, and S.-B. Jung, "Interfacial reaction in steel-aluminum joints made by friction stir welding," *Scripta Materialia*, vol. 55, no. 4, pp. 355–358, 2006.

[10] K. Kimapong and T. Watanabe, "Friction stir welding of aluminium alloy to steel," *Welding Journal*, vol. 83, no. 10, article 277, 2004.

[11] T. Watanabe, H. Takayama, and A. Yanagisawa, "Joining of aluminum alloy to steel by friction stir welding," *Journal of Materials Processing Technology*, vol. 178, no. 1-3, pp. 342–349, 2006.

[12] H. Uzun, C. Dalle Donne, A. Argagnotto, T. Ghidini, and C. Gambaro, "Friction stir welding of dissimilar Al 6013-T4 to X5CrNi18-10 stainless steel," *Materials & Design*, vol. 26, no. 1, pp. 41–46, 2005.

[13] A. P. Reynolds, W. Tang, M. Posada, and J. Deloach, "Friction stir welding of DH36 steel," *Science and Technology of Welding and Joining*, vol. 8, no. 6, pp. 455–460, 2003.

[14] K. V. Jata, K. K. Sankaran, and J. J. Ruschau, "Friction-stir welding effects on microstructure and fatigue of aluminium alloy 7050-T7451," *Metallurgical and Materials Transactions A*, vol. 31, no. 9, pp. 2181–2192, 2000.

[15] T. R. McNelley, S. Swaminathan, and J. Q. Su, "Recrystallization mechanisms during friction stir welding/processing of aluminum alloys," *Scripta Materialia*, vol. 58, no. 5, pp. 349–354, 2008.

[16] M. Abu-Okail, A. Abu-Oqail, and M. H. Ata, "Effect of friction stir welding process parameters with interlayer strip on microstructural characterization and mechanical properties," *Journal of Failure Analysis and Prevention*, vol. 20, no. 1, pp. 173–183, 2020.

[17] P. B. Prangnell and C. P. Heason, "Grain structure formation during friction stir welding observed by the 'stop action technique," *Acta Materialia*, vol. 53, no. 11, pp. 3179–3192, 2005.

[18] R. W. Fonda, J. F. Bingert, and K. J. Colligan, "Development of grain structure during friction stir welding," *Scripta Materialia*, vol. 51, no. 3, pp. 243–248, 2004.

[19] Y. S. Sato, M. Urata, and H. Kokawa, "Parameters controlling microstructure and hardness during friction-stir welding of precipitation-hardenable aluminum alloy 6063," *Metallurgical and Materials Transactions A*, vol. 33, no. 3, pp. 625–635, 2002.

[20] G. Liu, L. E. Murr, C. S. Niou, J. C. McClure, and F. R. Vega, "Microstructural aspects of the friction-stir welding of 6061-T6 aluminum," *Scripta Materialia*, vol. 37, no. 3, pp. 355–361, 1997.

[21] Y. S. Sato, H. Kokawa, M. Enomoto, and S. Jogan, "Microstructural evolution of 6063 aluminum during friction-stir welding," *Metallurgical and Materials Transactions A*, vol. 30, no. 9, pp. 2429–2437, 1999.

[22] C. G. Rhodes, M. W. Mahoney, W. H. Bingel, R. A. Spurling, and C. C. Bampton, "Effects of friction stir welding on microstructure of 7075 aluminum," *Scripta Materialia*, vol. 36, no. 1, pp. 69–75, 1997.

[23] F. C. Liu, Y. Hovanski, M. P. Miles, C. D. Sorensen, and T. W. Nelson, "A review of friction stir welding of steels: tool, material flow, microstructure, and properties," *Journal of Materials Science & Technology*, vol. 34, no. 1, pp. 39–57, 2018.

[24] H. Kokawa, S. H. C. Park, Y. S. Sato, K. Okamoto, S. Hirano, and M. Inagaki, "Microstructures in friction stir welded 304 austenitic stainless steel," *Welding in the World*, vol. 49, no. 3-4, pp. 34–40, 2005.

[25] K. Singh, G. Singh, and H. Singh, "Investigation of microstructure and mechanical properties of friction stir welded AZ61 magnesium alloy joint," *Journal of Magnesium and Alloys*, vol. 6, no. 3, pp. 292–298, 2018.

[26] A. Heidarzadeh, S. Mironov, R. Kaibyshev et al., "Friction stir welding/processing of metals and alloys: a comprehensive review on microstructural evolution," *Progress in Materials Science*, vol. 117, article 100752, 2020.

[27] Y. J. Kwon, I. Shigematsu, and N. Saito, "Dissimilar friction stir welding between magnesium and aluminum alloys," *Materials Letters*, vol. 62, no. 23, pp. 3827–3829, 2008.

[28] Z. Zhang and H. W. Zhang, "A fully coupled thermomechanical model of friction stir welding," *The International Journal of Advanced Manufacturing Technology*, vol. 37, no. 3-4, pp. 279–293, 2008.

[29] P. Heurtier, M. J. Jones, C. Desrayaud, J. H. Driver, F. Montheillet, and D. Allehaux, "Mechanical and thermal modelling of friction stir welding," *Journal of Materials Processing Technology*, vol. 171, no. 3, pp. 348–357, 2006.

[30] R. Nandan, G. G. Roy, T. J. Lienert, and T. Debroy, "Three-dimensional heat and material flow during friction stir welding of mild steel," *Acta Materialia*, vol. 55, no. 3, pp. 883–895, 2007.

[31] H. B. Schmidt and J. H. Hattel, "Thermal modelling of friction stir welding," *Scripta Materialia*, vol. 58, no. 5, pp. 332–337, 2008.

[32] C. M. Chen and R. Kovacevic, "Finite element modeling of friction stir welding–thermal and thermomechanical analysis," *International Journal f Machine Tools and Manufacture*, vol. 43, no. 13, pp. 1319–1326, 2003.

[33] X. He, F. Gu, and A. Ball, "A review of numerical analysis of friction stir welding," *Progress in Materials Science*, vol. 65, pp. 1–66, 2014.

[34] S. Verma and J. P. Misra, "A critical review of friction stir welding process," *DAAAM International Scientific Book*, vol. 14, article 266, 2015.

[35] C. Hamilton, S. Dymek, and A. Sommers, "A thermal model of friction stir welding in aluminum alloys," *International Journal of Machine Tools and Manufacture*, vol. 48, no. 10, pp. 1120–1130, 2008.

[36] D. G. Andrade, C. Leitão, N. Dialami, M. Chiumenti, and D. M. Rodrigues, "Modelling torque and temperature in friction stir welding of aluminium alloys," *International Journal of Mechanical Sciences*, vol. 182, article 105725, 2020.

[37] Y. G. Kim, H. Fujii, T. Tsumura, T. Komazaki, and K. Nakata, "Three defect types in friction stir welding of aluminum die casting alloy," *Materials Science and Engineering: A*, vol. 415, no. 1-2, pp. 250–254, 2006.

[38] N. Dialami, M. Cervera, and M. Chiumenti, "Defect formation and material flow in friction stir welding," *European Journal of Mechanics - A/Solids*, vol. 80, p. 103912, 2020.

[39] Y. N. Zhang, X. Cao, S. Larose, and P. Wanjara, "Review of tools for friction stir welding and processing," *Canadian Metallurgical Quarterly*, vol. 51, no. 3, pp. 250–261, 2012.

[40] R. Rai, A. De, H. K. D. H. Bhadeshia, and T. DebRoy, "Review: friction stir welding tools," *Science and Technology of Welding and Joining*, vol. 16, no. 4, pp. 325–342, 2011.

[41] M. K. Gupta, "Effects of tool profile on mechanical properties of aluminium alloy Al 1120 friction stir welds," *Journal of Adhesion Science and Technology*, vol. 34, no. 18, pp. 2000–2010, 2020.

[42] K. Kumar and S. V. Kailas, "The role of friction stir welding tool on material flow and weld formation," *Materials Science and Engineering: A*, vol. 485, no. 1-2, pp. 367–374, 2008.

[43] B. T. Gibson, D. H. Lammlein, T. J. Prater et al., "Friction stir welding: process, automation, and control," *Journal of Manufacturing Processes*, vol. 16, no. 1, pp. 56–73, 2014.

[44] D. Srivastava, U. Chadha, N. Gunreddy et al., "A brief review of tube-to-tube plate welding process," *Materials Today Proceedings*, 2022.

[45] W. Tang, X. Guo, J. C. McClure, L. E. Murr, and A. Nunes, "Heat input and temperature distribution in friction stir welding," *Journal of Materials Processing and Manufacturing Science*, vol. 7, no. 2, pp. 163–172, 1998.

[46] H. Fujii, L. Cui, N. Tsuji, M. Maeda, K. Nakata, and K. Nogi, "Friction stir welding of carbon steels," *Materials Science and Engineering: A*, vol. 429, no. 1-2, pp. 50–57, 2006.

[47] W. B. Lee and S. B. Jung, "The joint properties of copper by friction stir welding," *Materials Letters*, vol. 58, no. 6, pp. 1041–1046, 2004.

[48] R. P. Singh, S. Dubey, A. Singh, and S. Kumar, "A review paper on friction stir welding process," *Materials Today: Proceedings*, vol. 38, pp. 6–11, 2021.

[49] M. Guerra, C. Schmidt, J. C. McClure, L. E. Murr, and A. C. Nunes, "Flow patterns during friction stir welding," *Materials Characterization*, vol. 49, no. 2, pp. 95–101, 2002.

[50] Á. Meilinger and I. Török, "The importance of friction stir welding tool," *Production Processes and Systems*, vol. 6, no. 1, pp. 25–34, 2013.

[51] V. Balasubramanian, "Relationship between base metal properties and friction stir welding process parameters," *Materials Science and Engineering: A*, vol. 480, no. 1-2, pp. 397–403, 2008.

[52] S. M. Senthil, R. Parameshwaran, S. R. Nathan, M. B. Kumar, and K. Deepandurai, "A multi-objective optimization of the friction stir welding process using RSM-based-desirability function approach for joining aluminum alloy 6063-T6

pipes," *Structural and Multidisciplinary Optimization*, vol. 62, no. 3, pp. 1117–1133, 2020.

[53] D. G. Andrade, C. Leitão, N. Dialami, M. Chiumenti, and D. M. Rodrigues, "Analysis of contact conditions and its influence on strain rate and temperature in friction stir welding," *International Journal of Mechanical Sciences*, vol. 191, article 106095, 2021.

[54] S. Suthaharan and S. Suthaharan, "Support vector machine," in *Machine Learning Models and Algorithms for Big Data Classification*, pp. 207–235, Springer, Boston, MA, 2016.

[55] D. A. Pisner and D. M. Schnyer, "Support vector machine," in *Machine Learning*, pp. 101–121, Academic Press, United States, 2020.

[56] J. Cervantes, F. Garcia-Lamont, L. Rodríguez-Mazahua, and A. Lopez, "A comprehensive survey on support vector machine classification: applications, challenges and trends," *Neurocomputing*, vol. 408, pp. 189–215, 2020.

[57] T. Fletcher, "Support vector machines explained," *Tutorial Paper*, pp. 1–19, 2009.

[58] T. Nasir, M. Asmaela, Q. Zeeshana, and D. Solyalib, "Applications of machine learning to friction stir welding process optimization," *Jurnal Kejuruteraan*, vol. 32, no. 1, pp. 171–186, 2020.

[59] A. Mishra, "Machine learning classification models for detection of the fracture location in dissimilar friction stir welded joint," *Applied Engineering Letters*, vol. 5, no. 3, pp. 87–93, 2020.

[60] R. Sandeep and A. Natarajan, "Prediction of peak temperature value in friction lap welding of aluminium alloy 7475 and PPS polymer hybrid joint using machine learning approaches," *Materials Letters*, vol. 308, article 131253, 2022.

[61] H. Wang, D. He, M. Liao, P. Liu, and R. Lai, "Study on quality prediction of 2219 aluminum alloy friction stir welding based on real-time temperature signal," *Materials*, vol. 14, no. 13, p. 3496, 2021.

[62] M. Buscema, "Back propagation neural networks," *Substance Use & Misuse*, vol. 33, no. 2, pp. 233–270, 1998.

[63] C. Decaro, G. Montanari, R. Molinariz et al., "Machine learning approach for prediction of hematic parameters in hemodialysis patients," *IEEE Journal of Translational Engineering in Health and Medicine*, vol. 7, pp. 1–8, 2019.

[64] T. Nishijima, "Universal approximation theorem for neural networks," 2021, http://arxiv.org/abs/2102.10993.

[65] K.-L. Du and M. N. Swamy, "Perceptrons," 2014.

[66] P. Marius, V. Balas, L. Perescu-Popescu, and N. Mastorakis, "Multilayer perceptron and neural networks," *WSEAS Transactions on Circuits and Systems.*, vol. 8, pp. 579–588, 2009.

[67] Y. Du, T. Mukherjee, P. Mitra, and T. DebRoy, "Machine learning based hierarchy of causative variables for tool failure in friction stir welding," *Acta Materialia*, vol. 192, pp. 67–77, 2020.

[68] R. Hartl, J. Hansjakob, and M. F. Zaeh, "Improving the surface quality of friction stir welds using reinforcement learning and Bayesian optimization," *International Journal of Advanced Manufacturing Technology*, vol. 110, no. 11-12, pp. 3145–3167, 2020.

[69] D. Yang, T. Mukherjee, and T. DebRoy, "Conditions for void formation in friction stir welding from machine learning," *npj Computational Materials*, vol. 5, no. 1, pp. 1–8, 2019.

[70] H. Okuyucu, A. Kurt, and E. Arcaklioglu, "Artificial neural network application to the friction stir welding of aluminum plates," *Materials & Design*, vol. 28, no. 1, pp. 78–84, 2007.

[71] B. Das, S. Pal, and S. Bag, "Torque based defect detection and weld quality modelling in friction stir welding process," *Journal of Manufacturing Processes*, vol. 27, pp. 8–17, 2017.

[72] A. Mishra and S. B. Dutta, "Detection of surface defects in friction stir welded joints by using a novel machine learning approach," *Applied Engineering Letters*, vol. 5, no. 1, pp. 16–21, 2020.

[73] M. Asmael, Q. Zeeshan, D. Solyali, and T. Nasir, "Applications of machine learning to friction stir welding process," *Optimization*, vol. 32, pp. 171–186, 2020.

[74] S. Verma, J. P. Misra, J. Singh, U. Batra, and Y. Kumar, "Prediction of tensile behavior of FS welded AA7039 using machine learning," *Materials Today Communications*, vol. 26, article 101933, 2021.

[75] R. Padmanaban, V. Balusamy, and R. Vaira Vignesh, "Effect of friction stir welding process parameters on the tensile strength of dissimilar aluminum alloy AA2024-T3 and AA7075-T6 joints," *Materialwissenschaft und Werkstofftechnik*, vol. 51, no. 1, pp. 17–27, 2020.

[76] L. P. Kaelbling, M. L. Littman, and A. W. Moore, "Reinforcement learning: a survey," *Journal of Artificial Intelligence Research*, vol. 4, pp. 237–285, 1996.

[77] M. Van Otterlo and M. Wiering, *Reinforcement learning and Markov decision processes. In Reinforcement Learning*, Springer, Berlin, Heidelberg, 2012.

[78] S. Verma, M. Gupta, and J. P. Misra, "Performance evaluation of friction stir welding using machine learning approaches," *MethodsX*, vol. 5, pp. 1048–1058, 2018.

[79] R. Hartl, F. Vieltorf, M. Benker, and M. F. Zaeh, "Predicting the ultimate tensile strength of friction stir welds using Gaussian process regression," *Journal of Manufacturing and Materials Processing*, vol. 4, no. 3, p. 75, 2020.

[80] L. Cui, H. Fujii, N. Tsuji, and K. Nogi, "Friction stir welding of a high carbon steel," *Scripta Materialia*, vol. 56, no. 7, pp. 637–640, 2007.

[81] T. J. Lienert, W. L. Stellwag Jr., B. B. Grimmett, and R. W. Warke, "Friction stir welding studies on mild steel," *Welding Journal-New York*, vol. 82, no. 1, p. 1-S, 2003.

[82] S. Kumaraswamy, V. Malik, S. Devaraj, V. K. Jain, and L. Avinash, "A study of shear friction factor in friction stir welding for developing a finite element model and its importance in the context of formation of defect free and defective weld," *Materials Today: Proceedings.*, vol. 45, pp. 299–303, 2021.

[83] K. Sithole and V. V. Rao, "Recent developments in micro friction stir welding: a review," *In IOP conference series: materials science and engineering*, vol. 114, article 012036, 2016.

[84] B. Li, Y. Shen, and W. Hu, "The study on defects in aluminum 2219-T6 thick butt friction stir welds with the application of multiple non-destructive testing methods," *Materials & Design*, vol. 32, no. 4, pp. 2073–2084, 2011.

[85] G. Karrar, A. Galloway, A. Toumpis, H. Li, and F. Al-Badour, "Microstructural characterisation and mechanical properties of dissimilar AA5083-copper joints produced by friction stir welding," *Journal of Materials Research and Technology*, vol. 9, no. 5, pp. 11968–11979, 2020.

[86] S. S. Sabari, S. Malarvizhi, and V. Balasubramanian, "The effect of pin profiles on the microstructure and mechanical properties of underwater friction stir welded AA2519-T87 aluminium alloy," *Engineering*, vol. 11, no. 1, pp. 1–14, 2016.

[87] J. Li, Y. Shen, W. Hou, and Y. Qi, "Friction stir welding of Ti-6Al-4V alloy: friction tool, microstructure, and mechanical

properties," *Journal of Manufacturing Processes*, vol. 58, pp. 344–354, 2020.

[88] R. Hartl, A. Bachmann, J. B. Habedank, T. Semm, and M. F. Zaeh, "Process monitoring in friction stir welding using convolutional neural networks," *Metals*, vol. 11, no. 4, article 535, 2021.

[89] D. J. Huggett, T. W. Liao, M. A. Wahab, and A. Okeil, "Prediction of friction stir weld quality without and with signal features," *International Journal of Advanced Manufacturing Technology*, vol. 95, no. 5–8, pp. 1989–2003, 2018.

[90] A. Mishra, "Understanding machine learning for friction stir welding technology," 2020, https://ssrn.com/abstract=3560731.

[91] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998.

[92] Q. Wang, W. Jiao, P. Wang, and Y. M. Zhang, "A tutorial on deep learning-based data analytics in manufacturing through a welding case study," *Journal of Manufacturing Processes.*, vol. 63, pp. 2–13, 2021.

[93] Y. Chen and Z. Yi, "Adaptive sparse dropout: learning the certainty and uncertainty in deep neural networks," *Neurocomputing*, vol. 450, pp. 354–361, 2021.

[94] Y. Cheng, Q. Wang, W. Jiao et al., "Detecting dynamic development of weld pool using machine learning from innovative composite images for adaptive welding," *Journal of Manufacturing Processes*, vol. 56, pp. 908–915, 2020.

[95] M. Atwya and G. Panoutsos, "Transient thermography for flaw detection in friction stir welding: a machine learning approach," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 7, pp. 4423–4435, 2020.

[96] C. J. Watkins and P. Dayan, "Q-learning," *Machine Learning*, vol. 8, no. 3-4, pp. 279–292, 1992.

[97] S. Shanavas and J. E. R. Dhas, "Quality prediction of friction stir weld joints on aa 5052 H32 aluminium alloy using fuzzy logic technique," *Materials Today: Proceedings*, vol. 5, no. 5, pp. 12124–12132, 2018.

[98] N. N. Bhat, K. Kumari, S. Dutta, S. K. Pal, and S. Pal, "Friction stir weld classification by applying wavelet analysis and support vector machine on weld surface images," *Journal of Manufacturing Processes*, vol. 20, pp. 274–281, 2015.

[99] A. Nooraziah and V. J. Tiagrajah, "A study on regression model using response surface methodology," *Applied Mechanics and Materials*, vol. 666, pp. 235–239, 2014.

[100] F. Lambiase, V. Grossi, S. I. Scipioni, and A. Paoletti, "Prediction of the power supplied in friction-based joining process of metal- polymer hybrids through machine learning," *Journal of Manufacturing Processes*, vol. 68, pp. 750–760, 2021.

[101] A. Mishra, "Local binary pattern defect recognition approach for the friction stir welded AA 1200 and AA 6061-T6 aluminum alloy," *Journal of Mechanical and Energy Engineering*, vol. 4, no. 1, pp. 27–32, 2020.

[102] M. Barstugan, Y. S. Ceran, M. Yilmaz, and N. A. Dundar, "Detection of Defects on Single-Bead Welding by Machine Learning Methods," *IOP Conference Series: Materials Science and Engineering*, vol. 895, no. 1, Article ID 895012012, p. 012012, 2020.

[103] K. Balachandar and R. Jegadeeshwaran, "Friction stir welding tool condition monitoring using vibration signals and random forest algorithm–a machine learning approach," *Materials Today: Proceedings*, vol. 46, pp. 1174–1180, 2021.

[104] N. Landwehr, M. Hall, and E. Frank, "Logistic Model Trees," *Machine learning*, vol. 59, no. 1-2, pp. 161–205, 2005.

[105] P. Rabe, A. Schiebahn, and U. Reisgen, "Deep learning approaches for force feedback based void defect detection in friction stir welding," *Journal of Advanced Joining Processes*, vol. 5, article 100087, 2021.

[106] A. Krizhevsky, I. Sutskever, and G. Hinton, "ImageNet classification with deep convolutional neural networks," *Advances in Neural Information Processing Systems*, vol. 25, 2012.

[107] G. Biau and E. Scornet, "A random forest guided tour," *TEST*, vol. 25, no. 2, pp. 197–227, 2016.

[108] S. Verma, J. P. Misra, and D. Popli, "Modeling of friction stir welding of aviation grade aluminium alloy using machine learning approaches," *International Journal of Modelling and Simulation*, vol. 42, pp. 1–8, 2020.

[109] W. Wang and Y. Lu, "Analysis of the mean absolute error (MAE) and the root mean square error (RMSE) in assessing rounding model," *In IOP conference series: materials science and engineering*, vol. 324, article 012049, 2018.

[110] Y. Wang and I. H. Witten, "Induction of model trees for predicting continuous classes," 1996.

[111] S. A. Vendan, R. Kamal, A. Karan et al., "Supervised machine learning in friction stir welding (FSW)," in *Welding and Cutting Case Studies with Supervised Machine Learning*, pp. 119–185, Springer, Singapore, 2020.

[112] A. Syah, W. Astuti, and J. Saedon, "Development of prediction system model for mechanical property in friction stir welding using support vector machine (SVM)," *Journal of Mechanical Engineering (JMechE)*, vol. 5, pp. 216–225, 2018.

[113] M. H. Shojaeefard, M. Akbari, and P. Asadi, "Multi objective optimization of friction stir welding parameters using FEM and neural network," *International Journal of Precision Engineering and Manufacturing*, vol. 15, no. 11, pp. 2351–2356, 2014.

[114] Y. K. Yousif, K. M. Daws, and B. I. Kazem, "Prediction of friction stir welding characteristic using neural network," *Jordan Journal of Mechanical and Industrial Engineering*, vol. 2, no. 3, 2008.

[115] L. De Filippis, L. M. Serio, F. Facchini, G. Mummolo, and A. D. Ludovico, "Prediction of the Vickers microhardness and ultimate tensile strength of AA5754 H111 friction stir welding butt joints using artificial neural network," *Materials*, vol. 9, no. 11, p. 915, 2016.

[116] A. Mishra, "Artificial intelligence algorithms for the analysis of mechanical property of friction stir welded joints by using python programming," *Przegląd Spawalnictwa*, vol. 92, no. 6, pp. 7–16, 2020.

[117] B. Eren, M. A. Guvenc, and S. Mistikoglu, "Artificial intelligence applications for friction stir welding: a review," *Metals and Materials International*, vol. 27, no. 2, pp. 193–219, 2021.

[118] L. Fratini, G. Buffa, and D. Palmeri, "Using a neural network for predicting the average grain size in friction stir welding processes," *Computers & Structures*, vol. 87, no. 17-18, pp. 1166–1174, 2009.

[119] B. Zhou, T. Pychynski, M. Reischl, and R. Mikut, "Comparison of machine learning approaches for time-series-based quality monitoring of resistance spot welding (RSW)," *Archives of Data Science, Series A (Online First)*, vol. 5, pp. 1–27, 2018.

[120] C. Stadter, M. Schmoeller, L. Rhein, and M. Zaeh, "Real-time prediction of quality characteristics in laser beam welding

using optical coherence tomography and machine learning," *Journal of Laser Applications*, vol. 32, no. 2, article 022046, 2020.

[121] Y. Chen, B. Chen, Y. Yao, C. Tan, and J. Feng, "A spectroscopic method based on support vector machine and artificial neural network for fiber laser welding defects detection and classification," *NDT & E International.*, vol. 108, article 102176, 2019.

[122] S. Zamanzad Gavidel, S. Lu, and J. L. Rickli, "Performance analysis and comparison of machine learning algorithms for predicting nugget width of resistance spot welding joints," *The International Journal of Advanced Manufacturing Technology.*, vol. 105, no. 9, pp. 3779–3796, 2019.

[123] M. Satpathy, S. Mishra, and S. Sahoo, "Ultrasonic spot welding of aluminum-copper dissimilar metals: a study on joint strength by experimentation and machine learning techniques," *Journal of Manufacturing Processes.*, vol. 33, pp. 96–110, 2018.

[124] S. Sarkar, A. Das, S. Paul, K. Mali, R. Sarkar, and A. Kumar, "Machine learning method to predict and analyze transient temperature in submerged arc welding," *Measurement*, vol. 170, article 108713, 2021.

[125] S. Sudhagar, M. Sakthivel, and P. Ganeshkumar, "Monitoring of friction stir welding based on vision system coupled with machine learning algorithm," *Measurement*, vol. 144, pp. 135–143, 2019.

[126] D. Mishra, A. Gupta, P. Raj et al., "Real time monitoring and control of friction stir welding process using multiple sensors," *Science & Technology*, vol. 30, pp. 1–11, 2020.

[127] A. Mishra, "Machine learning approach for defects identification in dissimilar friction stir welded aluminium alloys AA 7075-AA 1100 joints," vol. 4, no. 1, pp. 88–95, 2020.

[128] R. Davis, "Expert systems: where are we? And where do we go from here?," *AI Magazine*, vol. 3, no. 2, pp. 3–3, 1982.

[129] R. Sivasubramani, A. Verma, G. Rithvik, U. Chadha, and S. S. Kumaran, "Influence on nonhomogeneous microstructure formation and its role on tensile and fatigue performance of duplex stainless steel by a solid-state welding process," *Materials Today: Proceedings*, vol. 46, pp. 7284–7296, 2021.

[130] L. Commin, M. Dumont, J. E. Masse, and L. Barrallier, "Friction stir welding of AZ31 magnesium alloy rolled sheets: influence of processing parameters," *Acta Materialia*, vol. 57, no. 2, pp. 326–334, 2009.

[131] Y. S. Sato, S. H. C. Park, M. Michiuchi, and H. Kokawa, "Constitutional liquation during dissimilar friction stir welding of Al and mg alloys," *Scripta Materialia*, vol. 50, no. 9, pp. 1233–1236, 2004.

[132] T. DebRoy and H. K. D. H. Bhadeshia, "Friction stir welding of dissimilar alloys – a perspective," *Science and Technology of Welding and Joining*, vol. 15, no. 4, pp. 266–270, 2010.

[133] A. Mishra, "Local binary pattern for the evaluation of surface quality of dissimilar friction stir welded ultrafine grained 1050 and 6061-T6 aluminium alloys," *ADCAIJ: Advances in Distributed Computing and Artificial Intelligence Journal*, vol. 9, no. 2, pp. 69–77, 2020.

[134] D. Mishra, S. Shree, A. Gupta et al., "Weld defect localization in friction stir welding process," *Welding in the World*, vol. 65, no. 3, pp. 451–461, 2021.

[135] K. Balachandar, R. Jegadeeshwaran, and D. Gandhikumar, "Condition monitoring of FSW tool using vibration analysis - a machine learning approach," *Materials Today: Proceedings*, vol. 27, pp. 2970–2974, 2020.

[136] T. W. Liao, J. Roberts, M. A. Wahab, and A. M. Okeil, "Building a multi-signal based defect prediction system for a friction stir welding process," *Procedia Manufacturing*, vol. 38, pp. 1775–1791, 2019.

[137] N. D. Ghetiya and K. M. Patel, "Prediction of tensile strength in friction stir welded aluminium alloy using artificial neural network," *Procedia Technology*, vol. 14, pp. 274–281, 2014.

[138] S. Saha, "A comprehensive guide to convolutional neural networks — the ELI5 way, towards data science," 2018, https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53.

[139] P. Ganesh, "Types of convolution kernels: simplified," 2019, https://towardsdatascience.com/types-of-convolution-kernels-simplified-f040cb307c37.

[140] R. Amiri, H. Mehrpouyan, L. Fridman, R. K. Mallik, A. Nallanathan, and D. Matolak, "A machine learning approach for power allocation in HetNets considering QoS," in *2018 IEEE international conference on communications (ICC)*, pp. 1–7, Kansas City, MO, USA, May 2018.

[141] R. Karthikeyan and V. Balasubramanian, "Predictions of the optimized friction stir spot welding process parameters for joining AA2024 aluminum alloy using RSM," *International Journal of Advanced Manufacturing Technology*, vol. 51, no. 1–4, pp. 173–183, 2010.

[142] H. R. Baharvandi, N. Ehsani, S. Verma, and N. K. Batra, "Experimental investigation and validation on the effect of nickel addition on properties of the pressureless sintered boron carbide composites using machine learning models," *Ceramics International*, vol. 48, pp. 13205–13215, 2022.

[143] S. Verma, V. Msomi, S. Mabuwa et al., "Machine learning application for evaluating the friction stir processing behavior of dissimilar aluminium alloys joint," *Proceedings of the Institution of Mechanical Engineers, Part L: Journal of Materials: Design and Applications*, vol. 236, no. 3, pp. 633–646, 2022.

[144] S. K. Selvaraj, A. Raj, M. Dharnidharka et al., "A cutting-edge survey of tribological behavior evaluation using artificial and computational intelligence models," *Advances in Materials Science and Engineering*, vol. 2021, 17 pages, 2021.