Artificial intelligence-based prediction of tensile strength of 9Cr-1Mo steel printed components on the advancement of scanning strategies

BACHELOR OF TECHNOLOGY

in

Mechanical Engineering

by

Shreyansh Srivastava

(Roll No. 220103127)

Under the supervision of

Prof. Swarup Bag

to

DEPARTMENT OF MECHANICAL ENGINEERING

INDIAN INSTITUTE OF TECHNOLOGY GUWAHATI

GUWAHATI – 781039, INDIA

[July-November 2025]

CERTIFICATE

This is to certify that the work contained in this project report entitled **"Artificial intelligence-based prediction of tensile strength of 9Cr-1Mo steel printed components on the advancement of scanning strategies"** submitted by **Shreyansh Srivastava (220103127)** to the Indian Institute of Technology Guwahati towards the requirement of **Bachelor of Technology** in **Mechanical Engineering** is a bona fide work carried out by him under my supervision and that it has not been submitted elsewhere for the award of any degree.

(Prof. Swarup Bag)

Mechanical Engineering

IIT GUWAHATI

November 2025

DECLARATION

I declare that this written submission represents our ideas in our own words, and where others' ideas or words have been included, we have adequately cited and referenced the sources. I also declare that we have adhered to all principles of academic honesty and integrity and have not misrepresented or fabricated or falsified any idea/data/fact/source in our submission. I understand that any violation of the above will be cause for disciplinary action by the Institute and can also evoke penal action from the sources which have thus not been properly cited or from whom proper permission has not been taken when needed.

Shreyansh Srivastava

Roll No. 220103127

Date: 20/11/2025

APPROVAL SHEET

This project report entitled "**Artificial intelligence-based prediction of tensile strength of 9Cr-1Mo steel printed components on the advancement of scanning strategies**" by **Shreyansh Srivastava (220103127)** is approved for the degree of Bachelor of Technology.

Examiners

_____

_____

_____

Supervisor

Prof. Swarup Bag

Chairman

_____

Date: 20/11/2025

Place: Guwahati

ABSTRACT

Wire Arc Additive Manufacturing (WAAM) of 9Cr-1Mo steel is increasingly pivotal for advanced power plant applications. However, the process induces complex thermal cycles leading to anisotropic mechanical properties that vary significantly with scanning strategies. This study presents a novel Artificial Intelligence (AI) framework to predict the complete tensile stress-strain behaviour of WAAM-fabricated 9Cr-1Mo components, thereby reducing the reliance on time-intensive destructive testing. I generated a comprehensive dataset comprising over 22,000 experimental data points across six distinct scanning strategies: Base, In-Out, Out-In, Inclined, Longitudinal, and Transverse.

I evaluated the efficacy of 11 distinct machine learning and deep learning architectures, ranging from classical Random Forest and Support Vector Regressors to advanced sequential models including LSTM, BiLSTM, GRU, and TCN. The study culminates in the development of a Stacked Hybrid Model (BiLSTM-GRU + XGBoost). This architecture leverages a bidirectional recurrent neural network to capture temporal hysteresis and an XGBoost meta-learner to minimize residual errors. The proposed model achieved state-of-the-art predictive accuracy with an R^2 of 0.9999 and an RMSE of 1.65 MPa for tensile stress, and an R^2 of 0.9999 for tensile strain. Statistical validation via ANOVA confirms the model's significance ($P < 0.05$), while SHAP analysis provides interpretability, confirming that the model correctly identifies Load and Extension as the primary physical drivers of Stress and Strain, respectively. This framework offers a scalable, non-destructive solution for optimizing deposition paths in additively manufactured critical components.

Keywords: *Wire Arc Additive Manufacturing, 9Cr-1Mo steel, Tensile Strength Prediction, Deep Learning, Stacked Hybrid Model, Scanning Strategies.*

LIST OF ACRONYMS USED

AM	Additive Manufacturing
WAAM	Wire Arc Additive Manufacturing
DED	Directed Energy Deposition
AI	Artificial Intelligence
ML	Machine Learning
DL	Deep Learning
CSEF	Creep Strength-Enhanced Ferritic
AUSC	Advanced Ultra-Supercritical
ANN	Artificial Neural Network
RNN	Recurrent Neural Network
LSTM	Long Short-Term Memory
BiLSTM	Bidirectional Long Short-Term Memory
GRU	Gated Recurrent Unit
TCN	Temporal Convolutional Network
XGB	XGBoost (eXtreme Gradient Boosting)
RMSE	Root Mean Squared Error
MAE	Mean Absolute Error
ANOVA	Analysis of Variance
SHAP	SHapley Additive exPlanations

LIST OF TABLES & FIGURES USED

CONTENTS

Chapter 1

Introduction

1.1 Background: 9Cr-1Mo Steel and Applications

9Cr-1Mo steel, recognized as a creep strength-enhanced ferritic (CSEF) alloy, derives its name due to the presence of approximately 9 wt.% chromium and 1 wt.% molybdenum in its composition [1]. It is widely acknowledged for its excellent creep-rupture strength and corrosion resistance under severe operating conditions, demonstrating reliable performance even at temperature >750 °C and pressure >27.5 MPa. [2]. Such attributes make it the preferred material for critical boiler components, including headers and tubes, in advanced ultra-supercritical (AUSC) power plants [3]. Despite its advantages, 9Cr-1Mo steel is susceptible to microstructural degradation, anisotropic mechanical behaviour, and residual stress accumulation under extreme thermal and mechanical loads, concerns that are further accentuated when employing modern fabrication methods like additive manufacturing (AM) [4]. These limitations highlight the need for continued research into optimizing fabrication strategies for 9Cr-1Mo steel to ensure its long-term reliability in next-generation power applications.

1.2 Wire Arc Additive Manufacturing (WAAM)

Directed Energy Deposition (DED) is a prominent metal additive manufacturing (AM) technique that enables the fabrication of near-net-shape components with complex geometrical features [5]. Among its variants, wire-based DED, commonly referred to as Wire Arc Additive Manufacturing (WAAM), has garnered considerable attention due to its economic advantages, namely the use of low-cost wire feedstock and the achievement of substantial deposition rates, making it attractive for applications in the aerospace, nuclear, and power generation sectors [6-10]. Such advantages make it a promising technique for manufacturing on critical alloys like 9Cr-1Mo steel for advanced power plant applications. However, the widespread industrial implementation of WAAM is fundamentally constrained by challenges originating from its thermal dynamics [2]. The process inherently involves a high-intensity energy source that subjects the material to severe thermal transients, governing the complex history of melting, re-melting, and solidification [11]. These transient thermal conditions influence molten pool kinetics and dynamics, which are inherently governed by the scanning strategies employed

[12,13]. This behaviour directly dictates the fluid dynamics within the molten pool, which in turn gives rise to undesirable outcomes such as non-uniform microstructures, anisotropic mechanical behaviour, and the generation of detrimental residual stresses [14]. Over prolonged service, these effects can cause premature failure, typically manifested as type IV cracking in AM-fabricated 9Cr-1Mo components [15]. Hence, deposition parameters such as heat input, dwell time, and scanning strategy are vital for regulating thermal behaviour and ensuring the metallurgical and mechanical integrity of WAAM-produced structures.

1.3 Scanning Strategies and Path Planning

Directed energy deposition with wire feed (WAAM) employs a variety of scan patterns (e.g., raster/zig-zag, contour-offset, spiral-in/spiral-out, bi-directional vs. uni-directional passes, oscillation/weave trajectories, and inter-layer rotation) to redistribute heat, stabilize the melt pool, and control thermal gradients and bead geometry [16–20]. Different patterns are needed because heat input and reheating history vary locally within multi-track and multi-layer builds. Path planning thus becomes a primary lever for managing molten pool kinetics/dynamics and the ensuing thermal cycles [16,18]. Finite-element and experimental studies show that deposition strategy strongly influences distortion and the spatial distribution of residual stresses in WAAM walls and thin-walled structures [16,18]. Layer-to-layer rotation strategies (e.g., alternating/rotational offsets) have also been shown to reduce residual stress hot-spots and texture induced anisotropy compared with fixed direction rastering [20]. Recent optimization frameworks explicitly couple simplified WAAM simulations with path strategy selection to minimize distortion and residual stress while maintaining productivity [21]. Broader reviews concur that path planning together with inter-layer dwell and heat input control is central to tailoring microstructure (grain size/texture), bead shape overlap, and final mechanical properties in WAAM components [22].

1.4 Role of Machine Learning in Additive Manufacturing

Machine Learning (ML) and AI have emerged as powerful tools for addressing the complex, non-linear thermal-metallurgical-mechanical interactions in WAAM that are difficult to capture through experiments alone [23,24]. Data driven models have been applied to predict bead geometry, hardness, and thermal fields, while also enabling defect detection from multimodal sensor signals such as melt pool images and acoustic emissions [25–27]. Surrogate approaches

using neural networks and Gaussian processes offer faster alternatives to finite element simulations for residual stress and distortion prediction [28]. In-situ monitoring studies further demonstrate that ML assisted computer vision and sensor fusion can identify process instabilities and support adaptive control [29]. However, most existing models are constrained by limited datasets, lack of physics integration, and poor generalizability, restricting their reliability for large scale industrial WAAM [30,31].

In this context, the present work investigates the scientific novelty of integrating advanced AI frameworks with WAAM of 9Cr-1Mo steel. While different scanning strategies strongly influence microstructural evolution and mechanical performance, exhaustive experimental exploration of all possible patterns is both costly and time intensive. To address this, we propose a predictive framework where machine learning and deep learning models are trained on experimentally measured tensile strength data obtained for six representative scanning strategies (Base, In Out, Out In, Inclined, Longitudinal, and Traverse). The trained models are then employed to predict full stress-strain curves for unseen scan patterns, thereby enabling efficient evaluation of mechanical behaviour without extensive fabrication trials. This approach offers a cost effective and scalable pathway to optimize scan strategies for WAAM-fabricated 9Cr-1Mo components, ensuring improved reliability and reduced development effort.

Chapter 2

Machine Learning Models

2.1 **Random Forest Regression**

Random forest regression is an ensemble learning technique that helps prevent overfitting in decision trees. It achieves this by creating a set of trees during training time. The technique's first and primary strength arises from its "forest" of decorrelated trees, which is induced by randomness in two ways. Each individual tree is trained on a unique subset of the data generated through the bootstrapped aggregating (bagging). At each node within a tree, a decision is made using a random subset of the features. The dual randomisation scheme effectively lowers the variance of the framework making the model robust against overfitting in comparison to a single decision tree. A model's prediction of a new input sample is made once the forest has been trained on the outputs of the trees. For a regression problem, this aggregation is done by taking the average of predictions from all the trees in the ensemble. The prediction made by Random Forest is more accurate than that made by the individual decision trees that are combined. The following mathematical expression gives the final prediction, $\hat{y}$, for an input x:

$$\hat{y} = \frac{1}{B}\sum_{b=1}^{B} T_b(x) \quad (1)$$

In this case, B is defined as the overall number of trees present in the forest while $T_b(x)$ is the prediction of the b^{th} tree.

2.2 Decision Tree Regression

The Decision Tree Regressor is a simple and basic non-parametric model that predicts continuous values using the learnt hierarchy of simple decision rules. It works by recursively dividing the dataset into subsets that are more homogenous than the last, to develop a tree structure. At every internal node, this algorithm will select the single best feature and the associated threshold. The best split is the one that maximally reduces the child variance, which is done by minimizing the mean squared Error. The MSE for a given node is calculated as:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}\left(y_i - \hat{y}_j\right)^2 \quad (2)$$

where n equals number of samples at that node, y_i is the true value of the sample, and $\hat{y}_j$ is the mean target value of the samples present at that node. This partitioning keeps going until some stopping condition is reached like the max depth of a tree. To estimate a new data point, we start at the top of the tree and move down to a terminal leaf node. The output is the mean of all target values of the training samples that belong to that leaf.

2.3 Support Vector Regression

The Support vector regression is an SVM algorithm that is used to determine the regression outcomes. SVR, unlike the rest, does not try to minimize the error for all data points. It tries to find a function in which most of the data points lie within a margin of tolerance (ϵ-insensitive tube). Only points that reside outside of this boundary are considered errors. The model tries to flatten the function and minimize the complexity, which is represented by the norm of the weight vector (), as well as the prediction errors of the outliers. The aim of the model is to minimize the trade-off determined in the model objective function subject to error constraints.

$$\frac{1}{2}||\omega||^2 + C\sum_{i=1}^{n}(\xi_i + \xi_i^*) \quad (3)$$

C in this expression is the regularization hyperparameter that trades off between the model complexity and the size of allowable errors, i.e., the slack variables ξ_i and ξ_i^*. . Measure the difference of data points outside the ϵ-tube. SVR has the advantage of modelling non-linear relationships with the kernel trick which can produce high-dimensional mappings of input features.

2.4 K-Neighbours Regression

K-Nearest Neighbours (KNN) regressor works on a non-parametric and instance-based learning algorithm that predicts a new data point based on nearness. KNN is in the set of lazy learners. KNN doesn't create any internal model which generalizes the internal models. It keeps all the training data. To obtain predictions in respect of some query point, the algorithm has to find the 'K' closest data points (the neighbours) in the feature space, typically in terms of

Euclidean distance. The target values of these neighbours are then averaged to get the output value. This prediction is mathematically represented as:

$$\hat{y}(x_q) = \frac{1}{K}\sum_{i=1}^{K} y_i \quad (4)$$

Here, the prediction for the query point x_q is $\hat{y}$ (x_q), and $\{y_1, y_2, ..., y_K\}$ represent target values of its K nearest neighbours. The selection of hyperparameter K, governs the balance of bias and variance, and is very vital for model performance.

2.5 Artificial Neural Networks

An artificial deep neural network (ANN) is a computer program designed to recognize a complex pattern in data. It is based on the structure of a biological brain. It's made up of connected nodes or neurons that are stacked up in layers. Data will enter through an input layer, with the workings happening in one or even more hidden layers, and a final prediction being generated in an output layer. The hidden layers contain neurons that take in values, compute the weighted sum, take the bias unit into account and then apply a non-linear function. The calculation of a single neuron's output, a_j, is conducted using the equation.

$$a_j = f\left(\sum_{i=1}^{n}(w_{ij} \cdot x_i) + b_j\right) \quad (5)$$

The parameters w_{ij} are the synaptic weights, x_i are the inputs from the layer before j, b_j represents the bias and f is the activation function. The network learns with the help of the backpropagation process. Weights and Biases are Iteratively modified using this process to minimize mean squared error. It also takes the actual and predicted values into account. With different layers and structures, you can optimize various high-sophisticated relationships.

2.6 Long Short Term Memory (LSTM)

The long short term memory (LSTM) networks are a type of recurrent neural network (RNN) specifically invented to avoid the long term dependency problem. The internal architecture of an LSTM cell is shown in Figure 1(a) and consists of a cell state (C_T) that looks like a conveyor belt. During the forward pass, the information can easily be removed or added to that cell state. This allows information to persist and flow through throughout the layers for

many time steps. A series of three gating mechanisms regulate the addition and removals from this particular memory. We first use forget gate to determine which parts of the previous state (C_T) we want to discard.

$$F_T = \sigma(W_F \cdot [H_{T-1}, X_T] + B_F) \quad (6)$$

The input gate chooses what information to add, combining a sigmoid gate which decides what values to keep and a tanh layer that creates a vector of candidate values.

$$I_T = \sigma(W_I \cdot [H_{T-1}, X_T] + B_I) \quad (7)$$

$$\overline{C_T} = \tanh(W_C \cdot [H_{T-1}, X_T] + B_C) \quad (8)$$

The new cell state is then updated as:

$$C_T = F_T \odot C_{T-1} + I_T \odot \overline{C_T} \quad (9)$$

In the final step, output gate uses a tanh function to compute new hidden state, H_T , from the updated cell state, and multiplies it by output of final sigmoid gate.

$$O_T = \sigma(W_o \cdot [H_{T-1}, X_T] + B_o) \quad (10)$$

$$H_T = O_T \odot \tanh(C_T) \quad (11)$$

In this given definition, W and B are learnable weights and biases respectively. Further, $\odot$ denotes element wise product. The LSTM is a sophisticated structure designed to learn and use information with distinct gates to forget, input, and output information.

2.7 Bidirectional Long Short-Term Memory (BiLSTM)

A Bidirectional Long Short-Term Memory (BiLSTM) network improves the basic LSTM by performing a similar operation in forward and backward direction. As shown in the architecture in Fig. 1 (b), the BiLSTM consists of two independent layers of LSTM. The first layer takes input sequence in order from beginning to end (forward pass) which captures the past context. The second layer processes the sequence in the reverse direction (backward pass) at the same time. So, the layer captures future context. At any time step t, the final output is produced by

combining or aggregating the hidden state from both the forward ($\overline{H}_T$) and backward ($\overline{H}_T$) layers at time T. Commonly, this combination occurs through concatenation, summation, or averaging. This combination can be represented as:

$$Y_T = g\left(\left[\overline{H}_T,\ \overline{H}_T\right]\right) \tag{12}$$

where g is a function of output. The two-pass process allows the model to have a fuller and richer understanding of the context around each point in the sequence. This can enhance performance on a variety of sequence modelling tasks.

2.8 Gated recurrent unit (GRU)

The Gated recurrent unit (GRU) model is essentially a simpler version of the LSTM model. GRU and LSTM are both effective at capturing temporal dependencies. But GRU is generally more efficient than LSTM. As shown in the architecture (Figure 1(c)), the GRU merges the forget gate and input gate, creating a single update gate (Z_T). It uses a reset gate (R_T). The reset gate decides how to mix the new input with the prior memory. The update gate determines how much of the past information should be carried forward to the present state. The key operations are as follows.

$$Z_T = \sigma(W_Z \cdot [H_{T-1},\ X_T]\ +\ B_Z) \tag{13}$$

$$R_T = \sigma(W_R \cdot [H_{T-1},\ X_T]\ +\ B_R) \tag{14}$$

$$\overline{H}_T = \tanh(W_H \cdot [R_T\ \odot\ H_{T-1},\ X_T]\ +\ B_H) \tag{15}$$

$$\overline{H}_T = (1 - Z_T)\ \odot\ H_{T-1} + Z_T\ \odot\ \overline{H}_T \tag{16}$$

Here, H_{T-1} denotes the previous hidden state, X_T the current input, W and B are learnable parameters, and $\odot$ denotes the element-wise product. The GRU has fewer parameters than LSTM, so it often trains faster than LSTM while maintaining a performance level similar to it on many tasks. It accomplishes this by using fewer gates and by using the hidden state as the cell state.

2.9 Temporal convolutional network (TCN)

A temporal convolutional network (TCN) is a convolutional neural network architecture designed for sequence modelling. It is a powerful, non-recurrent alternative to LSTM and GRU. As depicted in Figure 1(d), the TCN architecture is based on a combination of two principles which are causal convolutions and dilated convolutions. When we use a causal convolution, we can ensure that the prediction at any time t only depends on inputs at time t and previous time steps. Using dilated convolutions, the network can obtain an exponentially large receptive field while being shallow due to applying the filter over an area larger than its length. The convolution operation is performed by skipping the input with step size equal to the dilation factor d. The dilated convolution operation is defined as.

$$C_d(s) = \sum_{i=0}^{k-1} F(i) \cdot S_{s-d \cdot i} \tag{17}$$

where F denotes the filter, S refers to the input sequence, and k refers to the filter size. TCNs usually pile up their layers within residual blocks, as shown, to stabilize the training of deep networks. This convolutional method processes the whole sequence all at once and as a result, is often far quicker and faster to train and inference. This approach is still effective at capturing long-range dependencies.

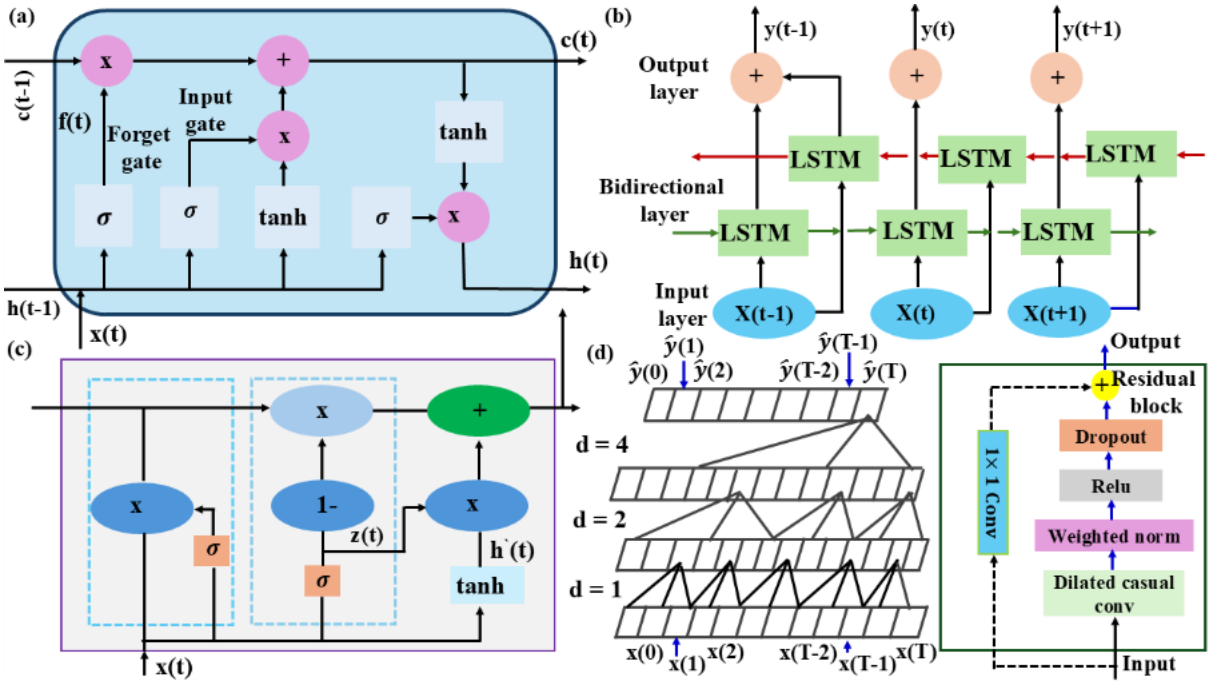

Figure 1. Architecture of AI models (a) LSTM, (b) BiLSTM, (c) GRU, (d) TCN

2.10 Hybrid Model (BiLSTM-GRU)

The hybrid method with BiLSTM-GRU is created to leverage the complementary strengths of both bidirectional long short-term Memory and gated recurrent units. This unit's structure is typically signified by Stacked RNNs, since the output of one-layer acts subordinately to the input of the next, as illustrated in the architecture schematic presented in Figure 2(a) An initial input sequence $X=(x_1, x_2, ..., x_T)$ goes through a BiLSTM layer for processing. This layer acts through processing the data in both the forward and backward directions, consequently producing a sequence of concatenated hidden states, H_{BiLSTM}. Each element of this sequence $h_t = [\overrightarrow{h_t}; \overleftarrow{h_t}]$ encompasses large amounts of information, creating a strong combination, and powerful new use of these elements. This feature-rich sequence is then passed to a GRU layer. The GRU, with its more computationally efficient gating mechanism, processes the output sequence from the BiLSTM to further refine the temporal features and summarize the information into a final hidden state vector, $h_{GRU} = GRU(H_{BiLSTM})$. Overall, this blended vector is passed through a series of connected layers to get the final output predictions: $Y_{pred} = Dense(h_{GRU})$. The technique combines the characteristics of the BiLSTM and GRU to give it the ability to be adaptive. Its sequence analysis exhibits dynamic patterns which have proven to be an improvement to addition to the knowledge in the model's database.

2.11 Stacked Hybrid Model (BiLSTM-GRU + XGBoost)

The proposed model is a two-level stacked ensemble consisting of a deep learning model and a gradient boosting model. This model maximizes predictive accuracy. As shown in Figure 2(b), it uses the result of one model to train another in a hierarchical manner. In first level, hybrid BiLSTM-GRU model acts like a powerful feature extractor. The data is trained on the first time-series dataset to learn the complicated temporal dependencies and make a set of initial predictions P_0 on the dependant variables (Stress and Strain). After extracting the complex patterns learned by the deep learning model, predictions are used as meta-features. The meta-features are stacked with the original input features, X_{orig} according to the stacking process, to develop a new feature set with augmentations.

$$X_{aug} = [X_{orig}; P_0] \tag{18}$$

It is then passed onto second-level model which is XGBoost Regressor. The XGBoost model is a meta-learner that identifies patterns in the residuals of the augmented data. This approach learns not only the original features but also the initial predictions, which in turn means the errors of the first-level model. The XGBoost model produces the final prediction which is

$$Y_{final} = XGBoost(X_{aug}) \tag{19}$$

This methodology of stacking helps the model's initial predictions capture complex dynamics using a specialized recurrent network. It then employs a strong boosting method to rectify and further enhance those predictions before producing the final outcome of generalization.

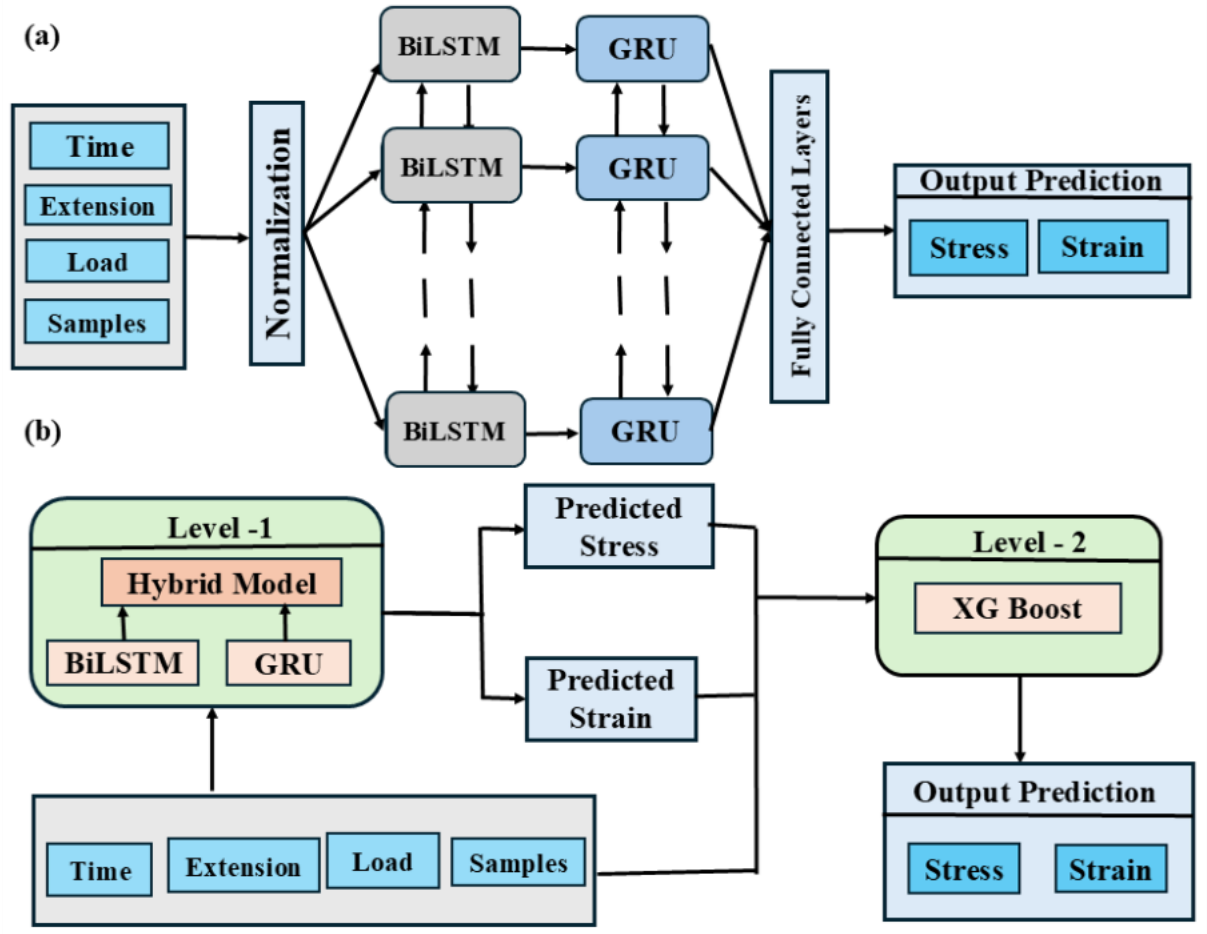

Figure 2. Architectures of hybrid AI models (a) BiLSTM-GRU and (b) Stacked Hybrid

Chapter 3

Methodology

3.1 Data Description

To build the machine learning and deep learning models, a comprehensive experimental dataset comprising over 22,000 data points was generated through tensile testing of 9Cr-1Mo steel components. These components were fabricated using Wire Arc Additive Manufacturing (WAAM) under six distinctive scanning strategies: Base, In-Out, Out-In, Inclined, Longitudinal, and Transverse.

The dataset captures the time-dependent mechanical response of the material. The primary attributes (input features) and target variables are defined as follows:

- **Input Features:** Time (s), Extension (mm), and Load (kN).
- **Target Variables:** Tensile Strain (mm/mm) and Tensile Stress (MPa).
- **Categorical Attribute:** Scanning Strategy (Sample Type).

3.2 Data Preprocessing

Preprocessing is a critical step to transform raw experimental data into a clean format suitable for training advanced neural networks.

3.2.1 Data Encoding

The "Sample" column, representing the six different scanning strategies, contains categorical text data. Since machine learning algorithms require numerical input, we employed Label Encoding to convert these categories into unique integers (0 to 5). This ensures that the model can mathematically distinguish between the diverse thermal histories associated with each scan pattern without increasing dimensionality excessively.

3.2.2 Feature Scaling

The dataset contains features with vastly different magnitudes and units, Load is in kilo-Newtons, Extension in millimetres, and Time in seconds. To prevent features with larger magnitudes from dominating the loss gradients during training, feature scaling was applied:

1. **Standard Scaling (Z-score Normalization):** Applied to the input features (Time, Extension, Load). This transforms the data to have a mean of 0 and a standard deviation of 1, ensuring stable convergence for gradient-based optimizers.
2. **Min-Max Scaling:** Applied to the target variables (Stress and Strain) for the deep learning models. This scales the output range between 0 and 1, matching the range of the activation functions (e.g., tanh or sigmoid) used in the neural network layers.

3.2.3 Model Development & Sequence Preparation

A robust development pipeline was established involving 11 different models. The dataset was split into training and testing sets using stratified sampling based on the 'Sample' column. This ensures that every split (Training: 60-90%, Testing: 10-40%) contains a representative proportion of all six scanning strategies, preventing bias toward any single material condition. For the Deep Learning models (LSTM, BiLSTM, GRU, TCN), the data was further preprocessed into 3D sequences. Unlike standard regression which treats data points independently, the input was reshaped into arrays of dimension (Samples, Timesteps, Features). This allows the recurrent layers to learn from the temporal history of the tensile test, capturing the progression from elastic loading to plastic deformation and eventual failure.

3.2.4 Hyper Parameter Optimization

To achieve state-of-the-art accuracy, hyperparameters were tuned for the "Champion" Stacked Hybrid Model. The final architecture parameters are:

- **Level 0 (Deep Learning Feature Extractor):**
    - BiLSTM Layer: 128 units, tanh activation, return sequences = True (captures forward and backward dependencies).
    - GRU Layer: 64 units, tanh activation (efficient temporal summarization).
    - Dropout: Rate of 0.2 (to prevent overfitting).
    - Optimizer: Adam with learning rate 0.001.

- **Level 1 (Meta-Learner):**
    - XGBoost Regressor: 100 estimators, objective = squared error.

3.2.5 Model Validation

The performance of the developed models was rigorously validated using statistical metrics and graphical analysis:

- **Metrics:** Coefficient of Determination (R^2), Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE) were calculated for all splits.
- **ANOVA:** Analysis of Variance was conducted on the prediction errors to determine the statistical significance of the performance differences between the proposed hybrid model and standard baselines.
- **Visual Inspection:** Predicted vs. Actual Stress-Strain curves were plotted to visually assess the model's ability to capture the yield point and ultimate tensile strength.

3.2.6 Feature Importance Analysis

To interpret the "black box" decision-making of the complex hybrid model, SHAP (SHapley Additive exPlanations) analysis was performed. This method assigns an importance value to each feature for a particular prediction. Specifically, we utilized the Tree Explainer for the XGBoost component to quantify how much 'Load', 'Extension', and 'Time' contributed to the final prediction of Stress and Strain, ensuring the model aligns with physical metallurgical principles.

Chapter 4

Results and Discussions

4.1 Exploratory Data Analysis

4.1.1 Pair Plot Analysis

Figure 3 illustrates the pair plot analysis of the dataset. This visualization reveals the pairwise relationships between the input parameters (Time, Load, Extension) and the target variables (Stress, Strain). A strong positive non-linear correlation is observed between Load and Stress, as well as Extension and Strain, which aligns with the fundamental physics of tensile testing. The distinctive "humps" and hysteresis loops visible in the scatter plots indicate the material's transition from the elastic region to the plastic deformation region, confirming the need for non-linear modelling techniques rather than simple linear regression.

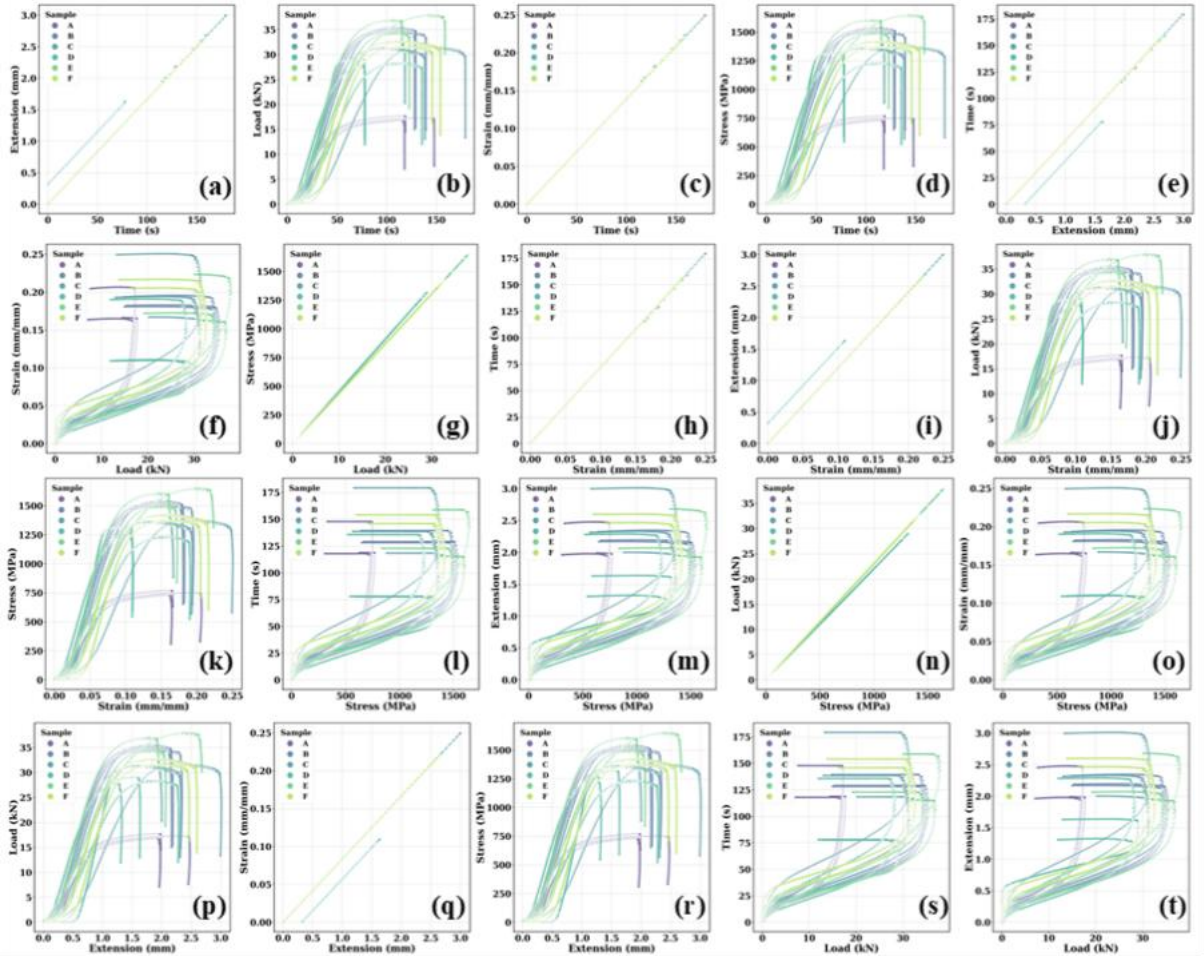

Figure 3. SNS pair plot analysis of experimental dataset

4.1.2 Outlier and Distribution Analysis

To assess data quality, box plots (Figure 4) and violin plots (Figure 5) were generated for all six scanning strategies (Samples A-F).

- **Box Plots:** Figure 4 shows the spread of Time, Extension, Load, Stress, and Strain across the different samples. While the interquartile ranges are consistent, distinct variations in the median values (represented by the central line) indicate that the scanning strategy significantly influences the mechanical response.
- **Violin Plots:** Figure 5 provides a deeper insight into the probability density of the data. For instance, sample 'B' (In-Out) shows a broader distribution in Stress compared to Sample 'A' (Base), suggesting that the In-Out strategy induces more variation in mechanical strength due to its specific thermal history.

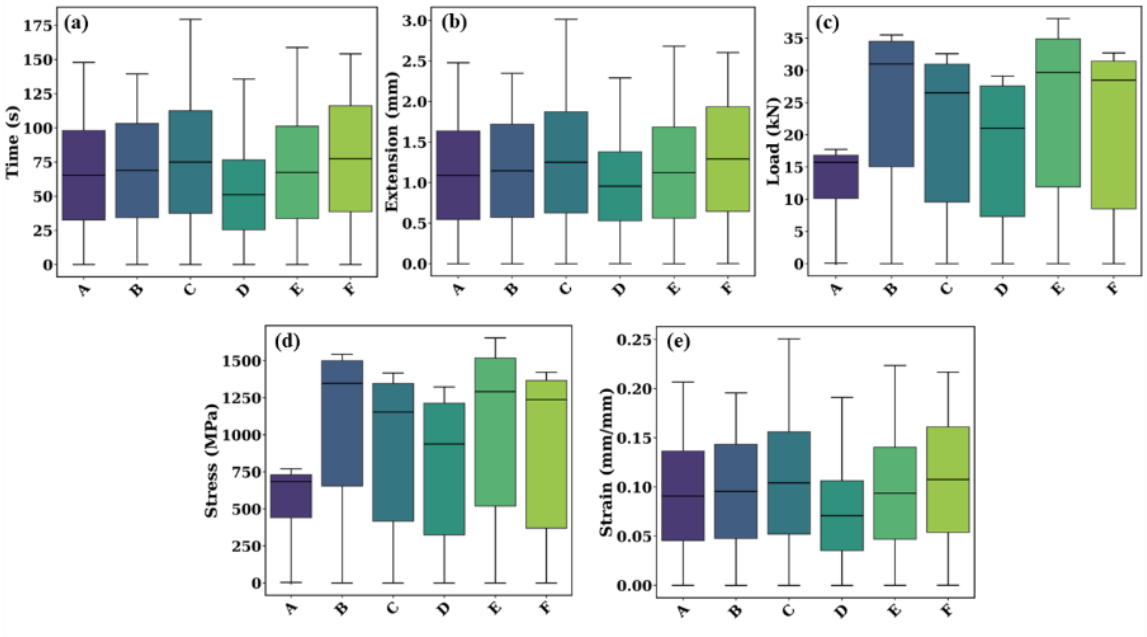

Figure 4. Analysis of outliers present in the datasets through box plot

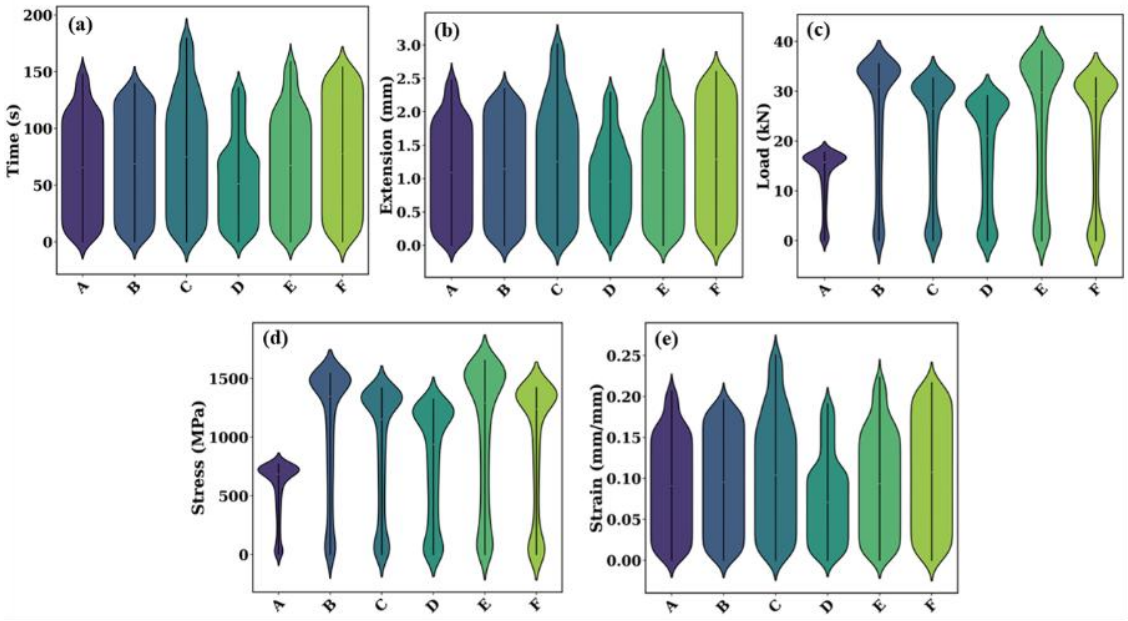

Figure 5. Violin plot analysis showing distribution across scanning strategies

4.2 Comparative Analysis of Predictive Models

The tensile strength prediction capabilities of 11 different algorithms were rigorously evaluated. These included classical Machine Learning models (Random Forest, Decision Tree, KNN, SVM) and Deep Learning architectures (ANN, LSTM, BiLSTM, GRU, TCN, Tuned Hybrid, and Stacked Hybrid).

4.2.1 Stress Prediction Performance

Figure 6 compares the predicted vs. actual Tensile Stress for all models.

- **Classical ML:** The SVM model (Figure 6d) exhibited significant scattering and poor convergence (MAE > 29 MPa), failing to capture the high-stress plastic region effectively. Random Forest and Decision Trees performed better but showed noticeable variance at peak stress levels.
- **Deep Learning:** The sequential models (LSTM, BiLSTM, GRU) demonstrated superior performance, with data points tightly clustered along the 45-degree ideal fit line.
- **Stacked Hybrid Model:** The proposed Stacked Hybrid model (Figure 6k) achieved the highest accuracy with an R^2 of 1.0000 and the lowest Mean Absolute Error (1.5837 MPa). This confirms that stacking XGBoost on top of the BiLSTM-GRU network effectively corrects residual errors.

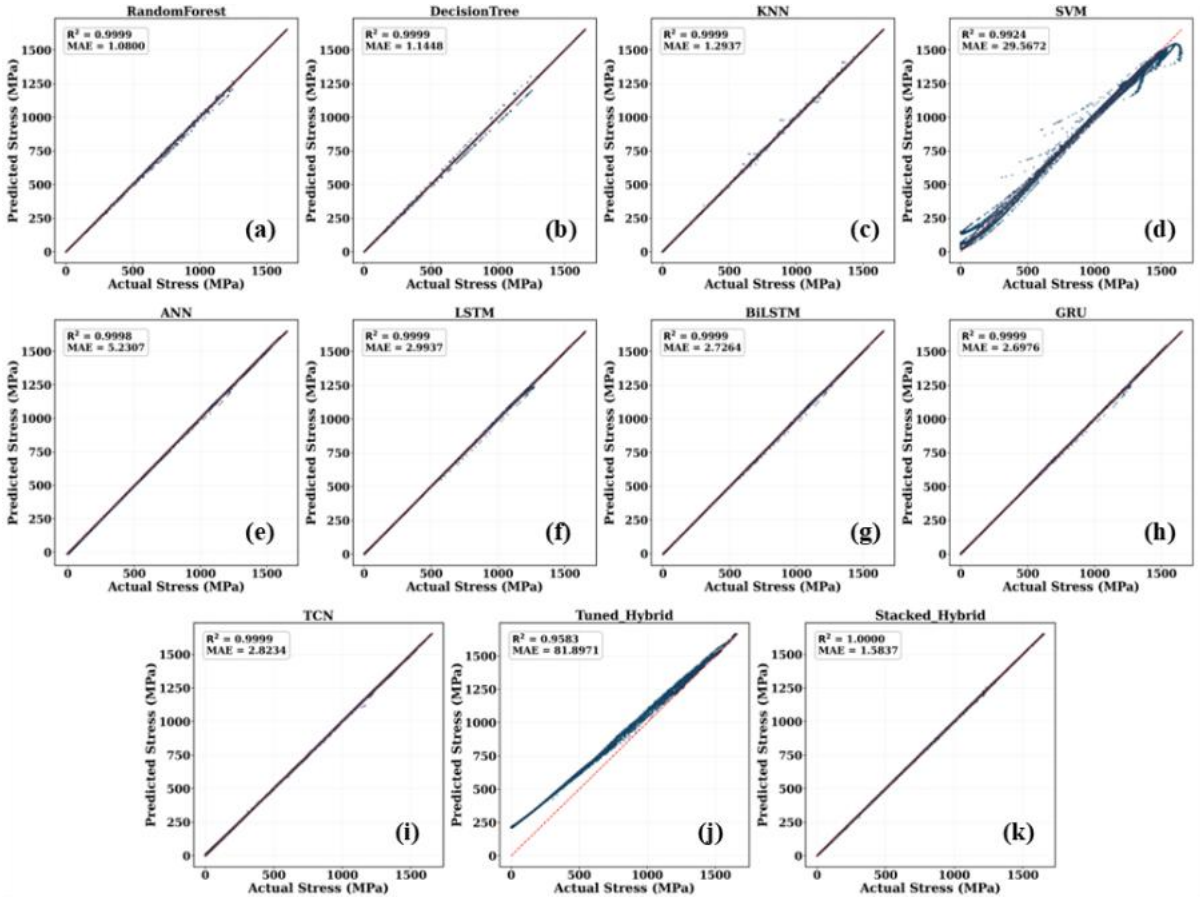

Figure 6. Comparison of predicted and actual stress by AI models on training and testing datasets

4.2.2 Strain Prediction Performance

Figure 7 presents the performance for Tensile Strain prediction. Similar to stress, the classical models showed higher dispersion. The Stacked Hybrid model (Figure 7k) again outperformed all others, achieving an R^2 of 0.9999 and an MAE of 0.0003 mm/mm, proving its robustness in predicting small deformations.

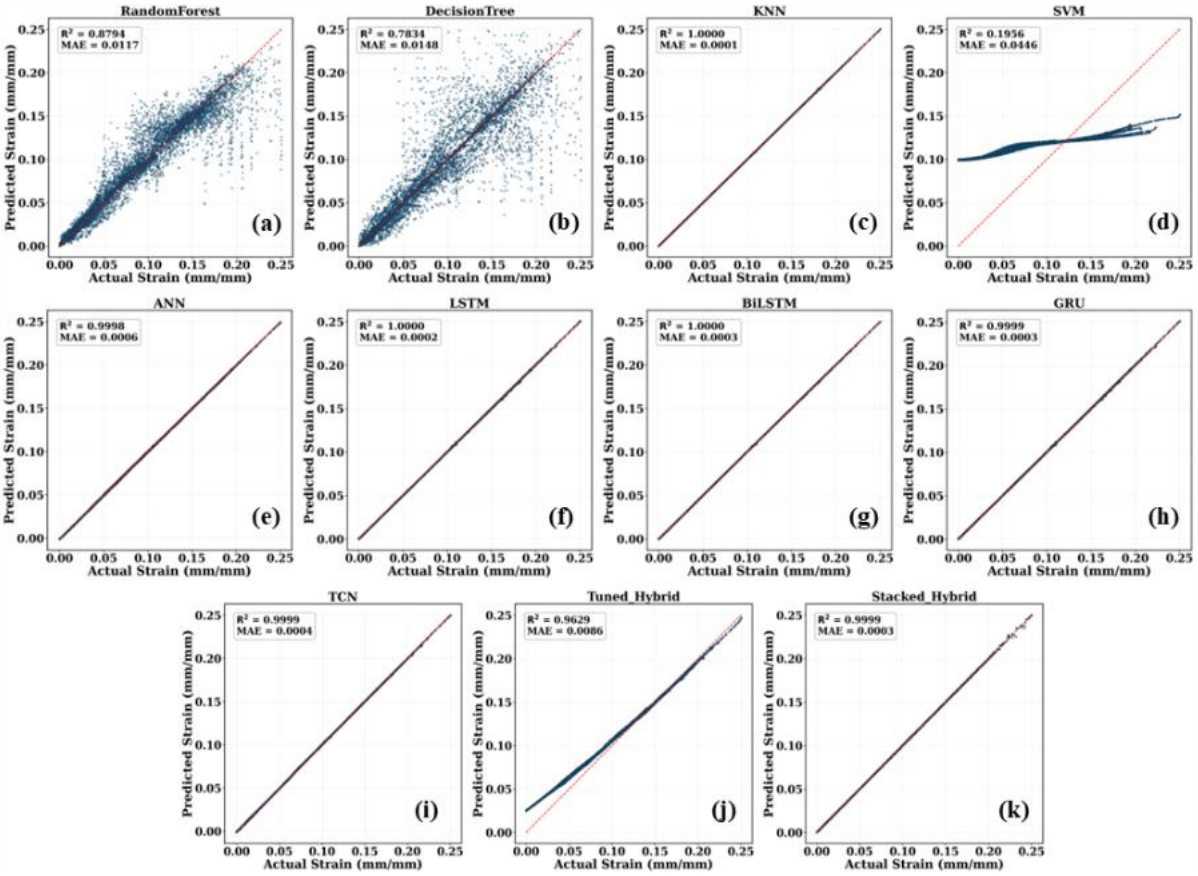

Figure 7. Comparison of predicted and actual strain by AI models on training and testing datasets

4.3 Prediction of Anisotropic Stress-Strain Curves

A key objective of this study was to predict the complete stress-strain curves for varying scanning strategies. Figure 8 visualizes the actual (red solid line) vs. predicted (blue dashed line) curves for the six strategies:

- Base Material (a)
- In-Out (b)
- Inclined (c)
- Out-In (d)
- Longitudinal (e)
- Transverse (f)

The Stacked Hybrid model demonstrates exceptional generalization capability. It accurately reproduces the distinct yield points, work-hardening rates, and ultimate tensile strengths for each strategy. Notably, it captures the subtle differences between the 'Longitudinal' and 'Transverse' directions, effectively modelling the anisotropy induced by the layer-by-layer deposition process.

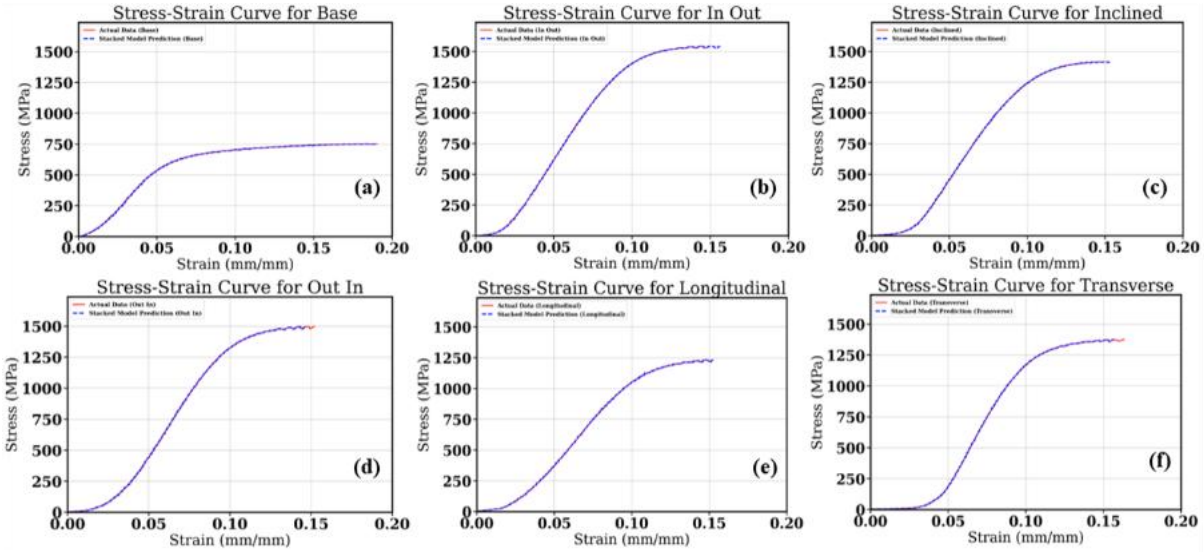

Figure 8. Actual and predicted stress-strain curve for different scan patterns using optimized models

4.4 Statistical Validation (ANOVA)

To ensure the statistical validity of the results, Analysis of Variance (ANOVA) was performed.

4.4.1 Feature Significance

- **For Stress (Table 1):** The F-value for 'Load' is extremely high (2.94×10^6), with a P-value of 0, indicating it is the most statistically significant predictor for Stress. Interaction terms like 'Time - Load' also show significance ($P < 0.05$).
- **For Strain (Table 2):** 'Time' and 'Extension' are the dominant factors, with P-values of 0. The quadratic term ($Time^2$) is also significant, reflecting the non-linear time-dependence of strain.

Table 1: Anova analysis for stress

Source	Degree of freedom	Adjusted sum of squares ($\times 10^3$)	F-value ($\times 10^3$)	P-value ($\times 10^{-4}$)	Remarks
Time	1	1.60	0.0319	0.000162	S
Extension	1	7.84	0.157	~ 0	S
Load	1	1.47	2940	0	S
I (Time ** 2)	1	0.228	0.00456	328	S
I (Extension ** 2)	1	0.156	0.00311	777	NS
I (Load ** 2)	1	8.20	0.164	~ 0	S
Time: Extension	1	0.182	0.00364	566	NS
Time: Load	1	83.1	1.66	0	S
Extension: Load	1	81.2	1.62	0	S
Residual	22770	1140			

Table 2: Anova analysis of strain

Source	Degree of freedom	Adjusted sum of squares (x 10^{-3})	F-value (x 10^{3})	P-value (x 10^{-8})	Remarks
Time	1	106	3630	0	S
Extension	1	0.164	5.62	0	S
Load	1	0.00526	0.180	~ 0	S
I (Time ** 2)	1	1.80	61.4	0	S
I (Extension ** 2)	1	1.53	52.5	0	S
I (Load ** 2)	1	0.000927	0.0317	1.85	S
Time: Extension	1	1.67	0.571	0	S
Time: Load	1	0.270	9.21	0	S
Extension: Load	1	0.280	9.56	0	S
Residual	22770	0.666			

4.4.2 Model Significance

Table 3 presents the ANOVA for the models themselves. The exceptionally high F-values (Stress - 8715.59; Strain - 6071.22) and P-values less than 0.05 confirm that the variation in performance is due to the model architecture and not random chance. This statistically proves that the Stacked Hybrid model provides a significant improvement over the baselines.

Table 3: Anova analysis for models

Error Metrics	Target	Source	Degree of freedom	Adjusted sum of squares	Adjusted mean squares	F-value	P-value
MAE	Tensile Strain	Model	10	8.1598	0.816	6071.22	P < 0.05
		Residual	50105	6.7342	0.0001		
		Total	50115	14.894			
	Tensile Stress	Model	10	19824909.45	1982490.95	8715.59	P < 0.05
		Residual	50105	11397124.38	227.46		
		Total	50115	31222033.83			
MSE	Tensile Strain	Model	10	0.0294	0.0029	2304.28	P < 0.05
		Residual	50105	0.064	0		
		Total	50115	0.0935			
	Tensile Stress	Model	10	1.66×10^{11}	1.66×10^{10}	3198.04	P < 0.05
		Residual	50105	2.60×10^{11}	5184479.71		
		Total	50115	4.26×10^{11}			

4.5 Model Interpretability (SHAP Analysis)

To decipher the "black box" nature of the hybrid model, SHAP (SHapley Additive exPlanations) analysis was conducted.

4.5.1 Feature Importance

The SHAP summary plot (Figure 9) ranks feature by their impact.

- **Stress Prediction (Figure 9a):** 'Load' has the highest impact, followed by 'Time'. High values of Load (red points) correlate with positive SHAP values, increasing the predicted Stress.
- **Strain Prediction (Figure 9b):** 'Time' and 'Extension' are the top predictors. This aligns with physical laws, as Strain is a function of elongation over time.

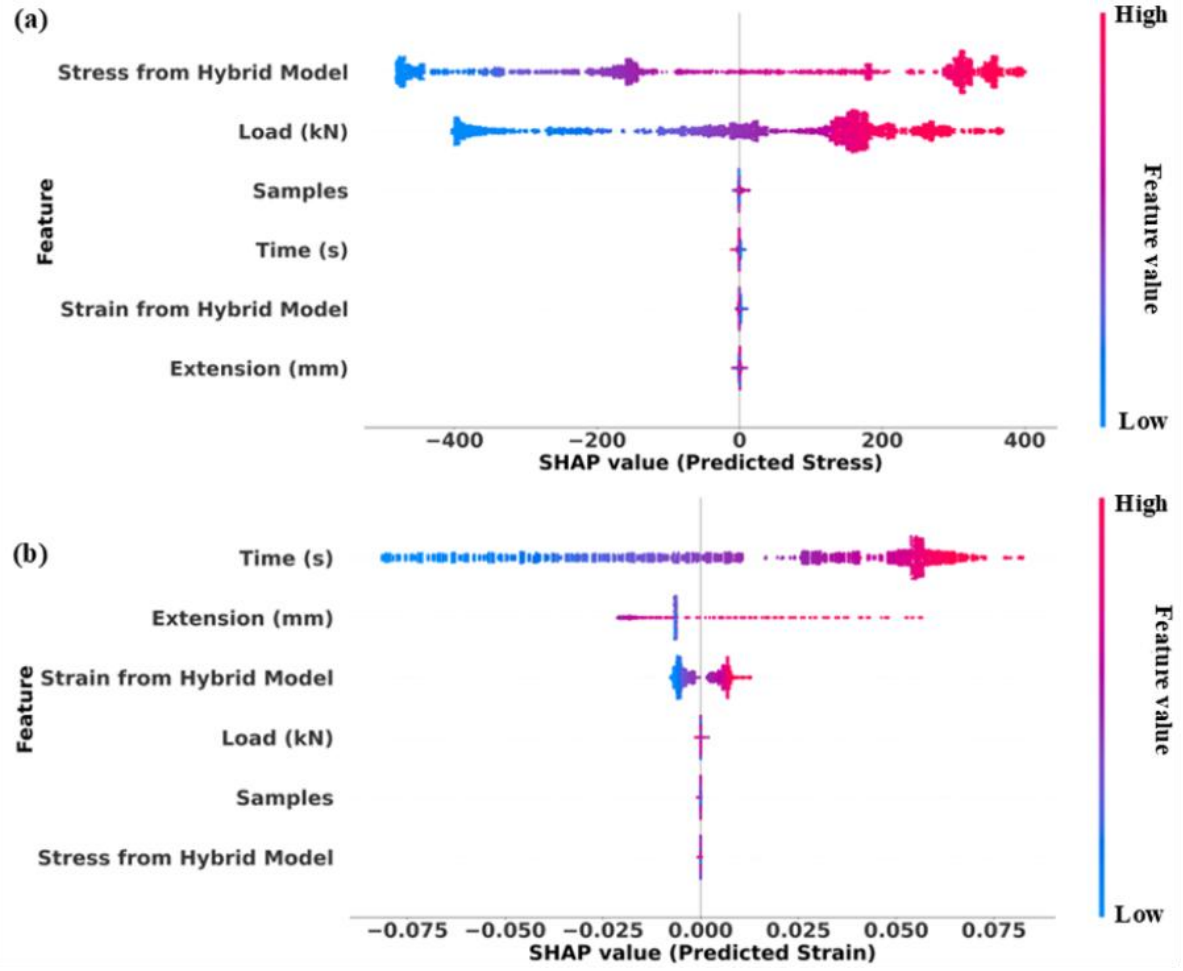

Figure 9. SHAP plot analysis for stress and strain feature importance

4.5.2 Dependence Plots

Figure 10 illustrates the specific relationship between features and predictions.

- **Load vs. Stress** (Figure 10a): Shows a clear linear trend that tapers off, mimicking the material's elastic-plastic behaviour.
- **Extension vs. Strain** (Figure 10c): Displays a positive correlation where increased extension leads to higher predicted strain.
- **Hybrid Model Contribution** (Figure 10b, 10d): The strong linear relationship between the "Predicted Stress/Strain from Hybrid Model" and the SHAP values confirms that the Level-1 XGBoost meta-learner relies heavily on the high-quality features extracted by the Level-0 BiLSTM-GRU network.

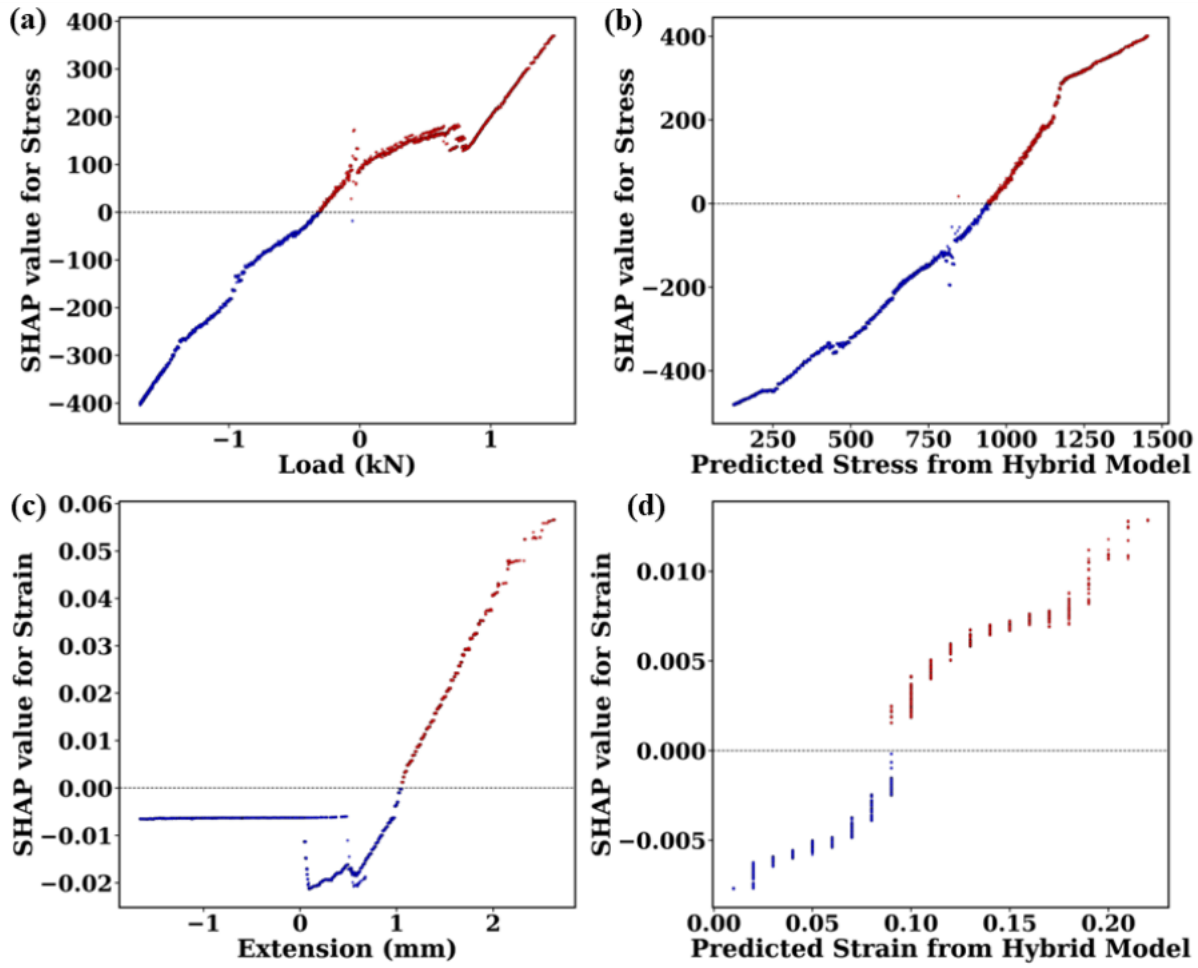

Figure 10. Mean value of SHAP dependence plots for output variables

Chapter 5

Conclusions and Future Scope

5.1 Conclusions

In this project, we successfully engineered and validated an Artificial Intelligence-driven framework for predicting the tensile strength of 9Cr-1Mo steel components fabricated via Wire Arc Additive Manufacturing (WAAM). By bridging the gap between experimental material science and advanced deep learning, we addressed the critical challenge of characterizing the anisotropic mechanical behaviour induced by complex thermal cycles in DED processes.

The key conclusions drawn from this study are:

1. **Anisotropy Characterization:** The extensive experimental analysis of over 22,000 data points confirmed that scanning strategies significantly influence mechanical behavior. Distinct stress-strain hysteresis loops were observed for 'Longitudinal' versus 'Transverse' strategies, confirming that simple linear models are insufficient for WAAM components.
2. **Model Superiority:** We evaluated 11 distinct architectures. Deep Learning models (LSTM, BiLSTM, GRU, TCN) consistently outperformed classical Machine Learning algorithms (SVM, KNN, Random Forest). Specifically, the BiLSTM layers proved critical in capturing the forward and backward temporal dependencies of the tensile test, effectively modelling the non-linear transition from elastic to plastic deformation.
3. **State-of-the-Art Performance:** The proposed Stacked Hybrid Model (BiLSTM-GRU + XGBoost) emerged as the champion architecture. By utilizing the BiLSTM-GRU network for temporal feature extraction and the XGBoost regressor for residual correction, the model achieved near-perfect predictive accuracy with an R^2 of 0.9999 and an RMSE of 1.65 MPa for Stress.
4. **Statistical and Physical Validity:** The robustness of the model was statistically confirmed via ANOVA, yielding high F-values that reject the null hypothesis. Furthermore, SHAP analysis validated the physical reliability of the AI, confirming that the model's decision-making aligns with fundamental mechanical laws (Stress and Load, Strain and Extension) rather than relying on spurious correlations.

5.2 Future Scope

While this work establishes a robust baseline for offline prediction, the research can be extended in the following directions to further enhance the reliability and applicability of AI in Additive Manufacturing:

1. **In-Situ Monitoring and Control:** Integrating this predictive model with real-time sensor data (e.g., thermal cameras or pyrometers) during the printing process could enable "in-situ" quality assurance, allowing for the adjustment of process parameters on-the-fly to prevent defects.
2. **Computer Vision Integration:** Future phases could incorporate melt-pool imagery data using Convolutional Neural Networks (CNNs). Fusing numerical process parameters with visual data could improve the detection of micro-defects like porosity or lack of fusion that affect tensile strength.
3. **Inverse Design Framework:** The current model predicts properties based on parameters (Forward Design). A future objective could be to develop an "Inverse Design" framework where the desired mechanical properties are input, and the AI suggests the optimal scanning strategy and printing parameters to achieve them.
4. **Generalizability to Other Alloys:** Testing the Stacked Hybrid architecture on other industrial alloys (e.g., Inconel 718 or Ti-6Al-4V) would validate its transfer learning capabilities and broader industrial utility.

References

[1] David, S.A., Siefert, J.A., Feng, Z., 2013. Welding and weldability of candidate ferritic alloys for future advanced ultra supercritical fossil power plants. Sci. Technol. Weld. Join. 18, 631–651.
[2] Bi-directional scanning strategies for residual stress management by tailoring microstructural evolution in directed energy deposition of 9Cr-1Mo steel.
[3] Krishnan, S., Kulkarni, D.V., De, A., 2016. Pulsed current gas metal arc welding of P91 steels using metal cored wires. J. Mater. Process. Technol. 229, 826–833.
[4] Eftink, B.P., Vega, D.A., El Atwani, O., Sprouster, D.J., Yoo, Y.S.J., Steckley, T.E., Aydogan, E., Cady, C.M., Al-Sheikhly, M., Lienert, T.J., Maloy, S.A., 2021. Tensile properties and microstructure of additively manufactured Grade 91 steel for nuclear applications. J. Nucl. Mater. 544, 152723.
[5] Role of inter-layer dwell time on residual stress generation of a thin wall structure by directed energy deposition of ferritic steel
[6] B. Tomar, S. Shiva, T. Nath, A review on wire arc additive manufacturing: processing parameters, defects, quality improvement and recent advances, Mater. Today Commun. 31 (2022) 103739.
[7] A. Kumar, I.A. Palani, M. Yadav, Comprehensive study of microstructure, phase transformations, and mechanical properties of nitinol alloys made of shape memory and superelastic wires and a novel approach to manufacture Belleville spring using wire arc additive manufacturing, Mater. Today Commun. 38 (2024) 107881.
[8] S.C.A. Costello, C.R. Cunningham, F. Xu, A. Shokrani, V. Dhokia, S.T. Newman, The state-of-the-art of wire arc directed energy deposition (WA-DED) as an additive manufacturing process for large metallic component manufacture, Int. J. Comput. Integrated Manuf. 36 (2023) 469–510.
[9] H. Dahaghin, M. Motavalli, H. Moshayedi, S.M. Zahrai, E. Ghafoori, Wire and arc additive manufacturing for strengthening of metallic components, Thin-Walled Struct. 203 (2024) 112074.
[10] Bayat, M., Zinovieva, O., Ferrari, F., Ayas, C., Langelaar, M., Spangenberg, J., Salajeghe, R., Poulios, K., Mohanty, S., Sigmund, O., Hattel, J., 2023. Holistic

computational design within additive manufacturing through topology optimization combined with multiphysics multi-scale materials and process modelling. Prog. Mater. Sci. 138, 101129.
[11] Zhang, X., Wang, L., Zhao, N., Liu, R., Zhang, L., Wu, W., Yang, D., Huang, Y., Wang, K., 2024. The effect of thermal cycle on microstructure evolution and mechanical properties of Co-free maraging steel produced by wire arc additive manufacturing. J. Mater. Process. Technol. 332, 118582.
[12] Marefat, F., Kapil, A., Banaee, S.A., Van Rymenant, P., Sharma, A., 2023. Evaluating shielding gas-filler wire interaction in bi-metallic wire arc additive manufacturing (WAAM) of creep resistant steel-stainless steel for improved process stability and build quality. J. Manuf. Process 88, 110–124.
[13] Sarma, R., Singh, A.K., Kapil, S., Bag, S., Joshi, S.N., 2023. Evolution of near homogenous mechanical and microstructural properties in wire-arc based directed energy deposition of low carbon steel following trochoidal trajectory toolpath. J. Mater. Process. Technol. 315, 117921.
[14] C. Cambon, I. Bendaoud, S. Rouquette, F. Souli´
e, A WAAM benchmark: from
process parameters to thermal effects on weld pool shape, microstructure and residual stresses, Mater. Today Commun. 33 (2022) 104235.
[15] Xiong, W., To, A., Klecka, M., 2022. Integrated Computational Materials and Mechanical Modeling for Additive Manufacturing of Alloys with Graded Structure used in Fossil Fuel Power Plants. DOE-PITT-31637. DOE-PITT-31637.
[16] Ali, M.H.; Han, Y.S. “Effect of Phase Transformations on Scanning Strategy in WAAM of EH36 Steel,” *Materials*, 2021, 14, 7871.
[17] Zhao, J. et al. “Influence of deposition path strategy on residual stress and mechanical properties in arc-based DED,” *Journal of Manufacturing Processes*, 2024. (oscillation/weave effects).
[18] Ali, M.H.; Han, Y.S. “A Finite Element Analysis on the Effect of Scanning Pattern and Energy on Residual Stress and Deformation in WAAM of EH36,” *Materials*, 2023, 16, 4698.
[19] Köhler, M. et al. “Comparative study of deposition patterns for DED-Arc additive manufacturing,” *Computers & Industrial Engineering*, 2021. (Contour-then-zig-zag recommendation).
[20] Liu, Y. et al. “Evolution, Control, and Mitigation of Residual Stresses in Metal AM,” *Advanced Engineering Materials*, 2023. (Alternating/rotational scan strategies reduce

residual stress/anisotropy).
[21] Zhao, X.F. et al. “A semi-analytical approach to WAAM simulation embedded in path-strategy optimisation,” *Virtual and Physical Prototyping*, 2024.
[22] Mansor, M.S.M. et al. “Integrated approach to WAAM: wire deposition strategy and parameter optimisation,” *Journal of Manufacturing Processes*, 2024. (Review on strategy–property link).
[23] I. Z. Era et al., “Machine learning in Directed Energy Deposition (DED) - review and applications,” *Additive Manufacturing* (2023).
[24] F. He et al., “Review: Research and application of artificial intelligence in WAAM,” (2023).
[25] X. Xiao et al., “Quality Prediction and Control in Wire Arc Additive Manufacturing using ML,” *Micromachines* 2022.
[26] J. Cheng, Y. Ling, W. De Waele, “An ANN Hardness Prediction Tool … for WAAM,” *Metals* 2024.
[27] Q. Wu et al., “Residual stresses in wire-arc additive manufacturing — data-driven analysis,” (2020).
[28] A. Nycz et al., “Effective residual stress prediction validated with neutron diffraction: surrogate models for AM,” *Int J Mater Form* (2021).
[29] Z. Snow et al., “Review of In Situ Sensing for Directed Energy Deposition,” ORNL report (2024).
[30] D. R. Gunasegaram et al., “Machine learning-assisted in-situ adaptive strategies for DED/WAAM,” *Additive Manufacturing* (2024).
[31] H. Mu et al., “A step toward digital twin of metallic additive manufacturing,” *Journal/Proceedings* (2024).