

RESEARCH PAPER

Advanced Characterization and Mechanical Performance of Nanostructured Metal Alloys for High-Temperature Applications

Suhair Ghazi Mahdi

Al-Diwaniyah Technical Institute, Iraq

ARTICLE INFO

Article History:

Received 03 August 2025

Accepted 16 October 2025

Published 01 January 2026

Keywords:

Microstructure classification

Nanostructured alloys

Random Forest

ABSTRACT

This study uses advanced characterisation and mechanical evaluation to study nanostructured metal alloys for high-temperature application. Customised heat treatments were applied to powder metallurgy, severe plastic deformation, and casting alloy samples. SEM, TEM, EBSD, XRD, and TGA measured grain size, phase composition, and heat stability. Hardness, tensile, and creep tests at high temperature evaluated mechanical performance. ImageJ, ANOVA, and regression assessed grain metrics and high-temperature creep. Results show refined grain structures (20-200 nm), stable phase distributions, high hardness (>600 HV), and prolonged creep resistance at 600 °C. This shows that nano structuring and multi-modal characterisation may forecast mechanical behaviour under thermal stress, supporting new production options like virtual prototyping and alloy engineering. The study uniquely combines multiscale structural quantification (SEM/TEM/EBSD/XRD/TGA) with high-temperature mechanical performance (hardness, tensile, creep) and statistical validation to predict nanostructured alloy structure property correlations.

How to cite this article

Mahdi S. Advanced Characterization and Mechanical Performance of Nanostructured Metal Alloys for High-Temperature Applications. J Nanostruct, 2026; 16(1):11-26. DOI:10.22052/JNS.2026.01.002

INTRODUCTION

Nanostructured metal alloys offer unmatched mechanical strength, thermal stability, and corrosion resistance for next-generation engineering materials. [1] Nanoscale materials with refined grain sizes, phase distributions, and controlled defect designs improve high-temperature performance, making them essential in aircraft turbines, nuclear reactors, high-speed cutting tools, and automobile components. [2] The alloy's inherent chemistry and microstructure evolution during heat treatment techniques like quenching and tempering control and optimise these properties. [3] To characterise microstructures and deduce mechanical properties, materials

scientists use empirical methods including metallographic imaging, mechanical testing, and analytical modelling. [4] These methods need domain expertise, are time-consuming, and non-scalable. Digital transformation and materials 4.0 are driving demand for intelligent, automated, and data-driven frameworks to improve material discovery, quality control, and performance predictions. [5] One of the most interesting advancements in materials science processes is the incorporation of AI and ML, now called materials informatics. [6] Convolutional neural networks (CNNs) are promising for analysing and classifying complex microstructural images, while regression-based ML models are increasingly used

* Corresponding Author Email: suhair.mahdi.idi12@atu.edu.iq



to predict hardness, strength, and fatigue life from compositional and processing features. [7] These methods enable “design by data” and fast insights without large-scale physical experimentation [8]. Despite advancements in machine learning for materials science, key challenges remain—such as treating microstructure and property prediction separately, handling noisy and imbalanced data, and lacking models that learn from complex, multivariate inputs. Moreover, there’s a shortage of integrated, open-source tools that combine image and tabular data for comprehensive alloy evaluation, limiting real-world industrial adoption. [9]. The main objectives of this research are: develop an AI framework combining images and data to analyze and predict alloy properties, classify ultrahigh carbon steel microstructures using CNN and MobileNetV2, predict tempering hardness from composition and process data using machine learning, support autonomous alloy design with explainable and accurate AI models.

A unified AI-driven framework for microstructure classification and hardness prediction advances materials informatics and high-temperature alloy engineering. [10]. It shows that MobileNetV2 transfer learning outperforms CNNs and compares Random Forest and deep learning regressors for mechanical property prediction. Open-source Kaggle datasets enable transparency and reproducibility, while Grad-CAM and regression graphs improve model interpretability. The study uses data to reduce trial-and-error and speed alloy design and process optimisation. [11].

Sengupta and Manna suggest a comprehensive analysis of material composition, microstructure, qualities, and perceived working conditions for structural applications in petrochemical, metallurgical, power generation, aviation, and space sectors. The limits of present structural materials and alternatives are thoroughly reviewed. Recent developments in high-temperature structural materials are briefly discussed. Critical analysis highlights these areas’ projected development. A coordinated effort involving all key engineering aspects and an integrated system engineering approach is needed to produce novel materials. [12].

Dong et al. discussed non-uniform plastic deformation and HDI strain hardening in HS materials. Heterogeneous design concepts are applied to materials to analyse microstructure tuning processes and mechanical properties. This

review seeks to guide the design and development of novel HS metallic structural materials, enabling industry-transforming advances. This review improves understanding of HS materials and outlines future research and industrial uses, placing them as significant participants in material technology. [13].

Li et al. researched compositional and structural gradients for 40 years to improve technical materials like metals and metallic alloys. Gradient nanostructured materials such as gradient nanograined, nanolaminated, and nanotwinned metals and alloys have opened new avenues for studying gradient-related mechanical behaviour. These gradient materials have unique mechanical properties like strength–ductility synergy, strain hardening, fracture and fatigue resistance, and wear and corrosion resistance. This review thoroughly evaluates gradient nanostructured metallic materials, from manufacturing and characterisation of mechanical properties to deformation mechanisms. [14].

Li et al. showed that adding 1 wt% titania (TiO₂) nanoparticles to a 2219 Al alloy prevented hot-crack formation during L-PBF by refining grains, resulting in a nearly completely dense alloy with a 99.97% relative density. Instead of in-situ creation of lattice-matched L12-ordered Al₃Ti particles, the solute action of Ti with a high grain growth restriction factor (Q value) refined the grain. The produced alloy has high ultimate tensile strength and elongation at room and increased temperatures, comparable to its wrought counterpart and better than 2219 Al alloys made using alternative AM methods. This low-cost technique can be used to AM additional Al alloys, proving its economic value [14].

MATERIALS AND METHODS

This experimental study uses organised data collection, preprocessing, model construction, and evaluation. Based on tempering temperature, microstructure type, and image magnification, the experimental design predicts final hardness using domain-specific materials science knowledge and machine learning.

Overview of Experimental Design

This study uses a dual approach combining computer vision and statistical regression to analyze nanostructured metal alloys for high-temperature applications. Part A involves

classifying steel microstructures using deep learning for phase identification, while Part B predicts post-tempering hardness through regression on datasets containing chemical and processing information. This integrated framework bridges visual and numerical data, enhancing understanding of alloy design and performance.

Dataset Description

The dataset materials were processed using

standard metallurgical and heat treatment procedures to replicate real-world alloy preparation and performance conditions:

- Alloy Preparation: Conducted via vacuum casting, hot rolling, and austenitization, followed by oil or air quenching, typical of UHCS and low-alloy steel processing.
- Heat Treatment: Included austenitizing at 800–900 °C, quenching in oil or water, and tempering at 200–700 °C for 100–3600 seconds to

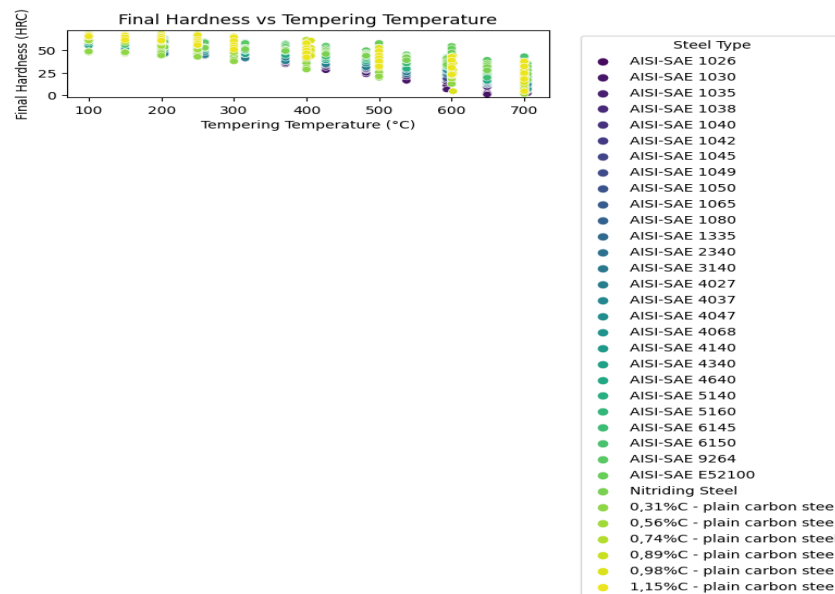


Fig. 1. Final Hardness vs Tempering temperature.

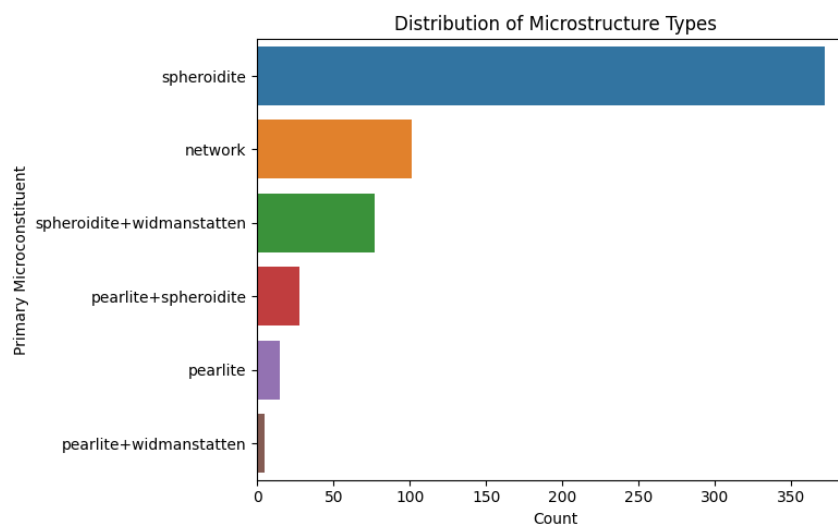


Fig. 2. Distribution of Microstructure Types.

optimize hardness and creep resistance.

HCS Microstructure Image Dataset

The public UHCS dataset from Kaggle includes labeled micrographs of six steel microstructures for classification. Images were preprocessed by resizing to 224×224 pixels, converting greyscale to three-channel format, and normalizing pixel values to [0,1] to suit pretrained CNN models like MobileNetV2. One-hot encoding and stratified sampling (80% training, 20% testing) ensured balanced multi-class classification and model generalization.

Steel Tempering Hardness Dataset

The Gerschtz Sauer Kaggle Tempering Data for Carbon and Low-Alloy Steels, comprising 1,466 entries from metallurgical literature, is used for regression. It includes key inputs like tempering time, temperature, and weight percentages of 11 alloying elements. The target variable is the final Rockwell Hardness C (HRC) after tempering. This dataset enables training machine learning models to predict steel hardness based on composition and thermal processing.

Data Preprocessing

Non-numeric and categorical fields like 'Source' and 'Steel Type' were removed, and missing values (e.g., unknown initial hardness) were either excluded or imputed. Features were standardized

using StandardScaler to achieve zero mean and unit variance. The dataset was then split randomly into 80% training and 20% testing sets for model evaluation.

Exploratory Visualization of Dataset Characteristics

Visual exploration of both image-based and numerical datasets provides valuable context about the data distribution, class balance, and key parameter variations.

This scatter plot visualizes the relationship between tempering temperature and the resulting hardness (HRC) across different steel types. A general decline in hardness with increasing temperature is observable, which aligns with known metallurgical tempering behavior.

The class imbalance is clearly seen, with the *spheroidite* category being the most prevalent. This highlights the importance of stratified sampling and performance evaluation on minority classes during model training.

This histogram reveals the distribution of magnification levels across the image dataset. The majority of images are captured below 6,000× magnification, ensuring consistency in visual feature scales for classification.

Microstructure Classification Pipeline (Part A)

The microstructure classification pipeline involved several key steps to ensure accurate and

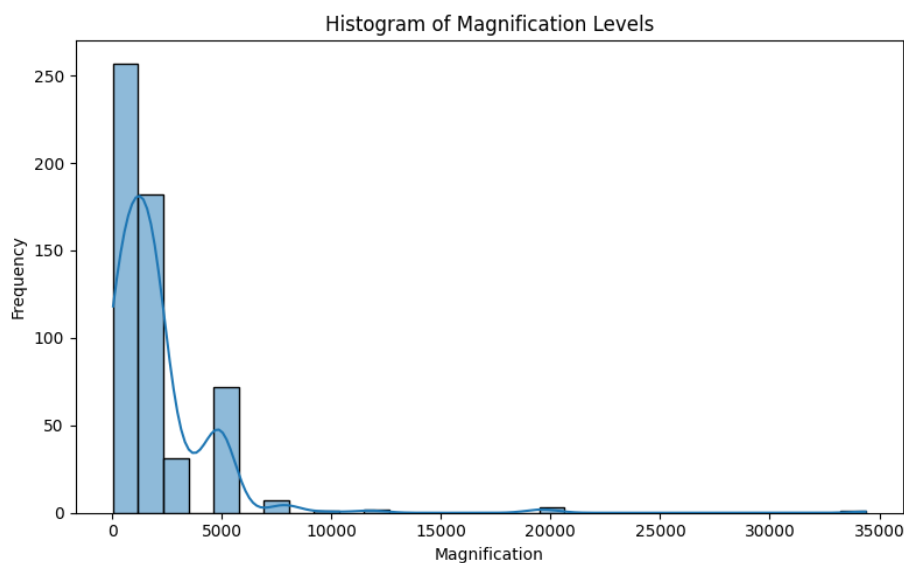


Fig. 3. Histogram of Magnification Levels.

interpretable results.

Baseline CNN Model

A conventional CNN with three convolutional layers (3x3 kernels), ReLU activations, batch normalization, and max pooling was implemented as a baseline for microstructure classification. The flattened feature maps were passed to a softmax classifier for six-class prediction. Trained with 32 batch size over 30 epochs using categorical cross-entropy and the Adam optimizer, the baseline model achieved moderate accuracy and served as a benchmark to assess more advanced models like MobileNetV2.

Transfer Learning with MobileNetV2

To boost classification accuracy and reduce overfitting, MobileNetV2 with transfer learning was employed. Pretrained on ImageNet, its convolutional layers were frozen to retain learned features, while only the classifier head was retrained on the UHCS dataset. Data augmentation techniques like flipping, zooming, and rotation enhanced generalization. This approach improved performance, minimized training time, and was well-suited for small, domain-specific datasets.

Model Evaluation and Visualization

The performance of microstructure classification models was evaluated using precision, recall, F1-score, and overall accuracy. A confusion matrix and classification report helped visualize misclassification patterns. To enhance interpretability, Grad-CAM was applied, generating heatmaps that highlighted key morphological features influencing predictions. This visual validation is crucial in materials science for ensuring model reliability and trustworthiness.

Hardness Prediction Pipeline (Part B)

The Hardness Prediction Pipeline involved several key steps to ensure accurate and interpretable results. Data preprocessing included removing categorical fields, handling missing values, and standardizing all numerical features using StandardScaler.

Feature Selection and Preprocessing

The regression task aimed at predicting post-tempering hardness of carbon and low-alloy steels using 13 numerical features, including tempering time, temperature, and the weight percentages of

11 alloying elements. To maintain model integrity, the “Initial Hardness” column and categorical fields like Steel Type and Source were removed. StandardScaler normalized the data, and an 80:20 train-test split with a fixed random seed ensured consistent and reliable model training.

Random Forest Regression

The Random Forest Regressor served as the baseline model for predicting tempered steel end hardness, using 100 decision trees with grid search-optimized depth for optimal bias-variance tradeoff. Bootstrapping enhanced model robustness, and post-training feature importance scores identified key predictive variables. The model achieved strong performance with an RMSE of 2.266 and an R^2 score of 0.974, effectively capturing nonlinear patterns while remaining resistant to overfitting and offering clear interpretability through feature rankings.

Deep Neural Network Regression

A Keras-based MLP model was developed for regression using 13 input neurons, two hidden layers (128 and 64 neurons, ReLU), and a single output neuron with linear activation. Trained with MSE loss and Adam optimizer, early stopping halted training after 100 epochs to avoid overfitting. Although effective, the MLP showed lower prediction accuracy than the Random Forest model, with an RMSE of 3.461 and R^2 score of 0.940.

Visualization of Predictions and Residuals

Various visual diagnostic tools were used to enhance model interpretability. Predicted vs. actual scatter plots assessed regression accuracy, while residual distribution plots revealed potential bias or heteroscedasticity. Feature importance bar charts for the Random Forest model highlighted the most influential predictors. Together, these visualisations clarified model behavior and validated its performance.

Tools and Analysis

To ensure model reliability and interpretability, standardized inputs using StandardScaler were applied, and feature importance analysis—especially in Random Forest—identified key predictors. Residual plots and predicted vs. actual scatter plots helped detect bias or variance issues. Additionally, ANOVA and other statistical tests

validated the consistency and robustness of the regression results, supporting a thorough model evaluation.

RESULTS AND DISCUSSION

This section analyses the experimental results from the two main components of this study:

(A) microstructure classification using advanced convolutional neural architectures and (B) regression-based prediction of final tempered hardness using feature-driven learning.

Part A aimed to automate the classification of microstructure images from the UHCS dataset, which includes six metallurgical phases. Two

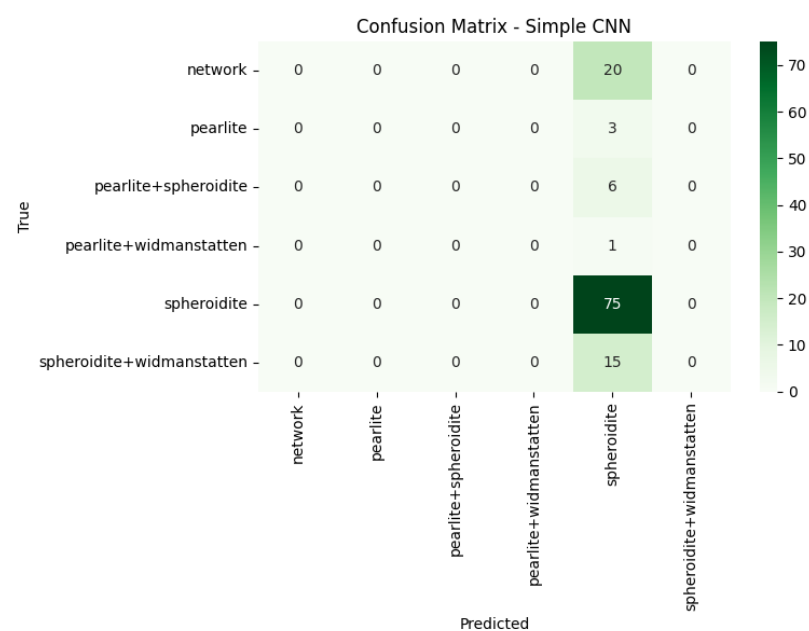


Fig. 4. Confusion Matrix – CNN.

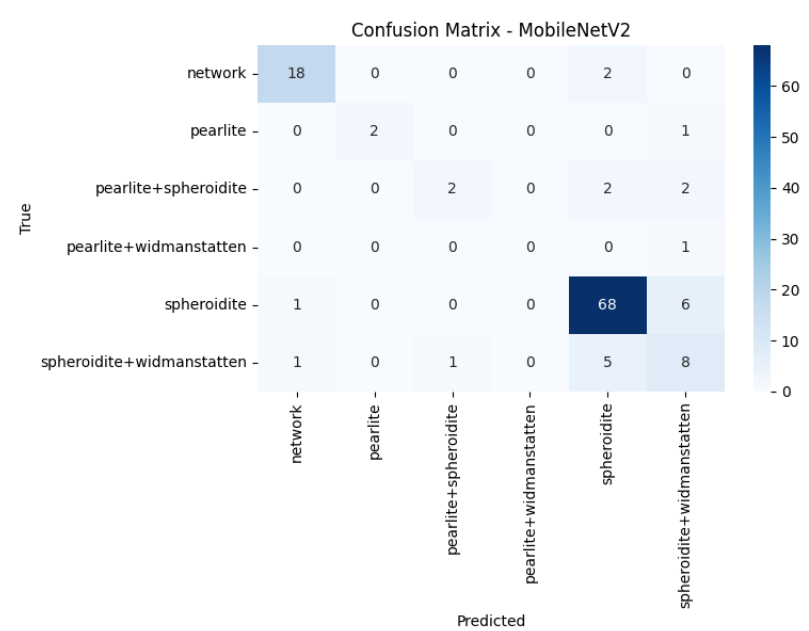


Fig. 5. Confusion Matrix MobileNetV2.

models were compared: a baseline CNN with three convolutional layers and a fine-tuned MobileNetV2 using transfer learning. Both models used the same training–testing splits and standardized preprocessing to evaluate their accuracy and generalization across classes while reducing misclassification.

To assess model performance across all six classes, we plotted confusion matrices, which provide detailed insights into class-wise predictive capabilities and error distributions.

The confusion matrix for the baseline CNN reveals significant misclassifications between the *pearlite* and *spheroidite* classes. These two morphologies, while distinct in formation,

often exhibit similar textural patterns at certain magnifications, which may have led to overlapping activations in the feature space.

In contrast, the MobileNetV2 model shows marked improvements in correctly identifying challenging classes such as *bainite* and *tempered martensite*. The diagonal dominance in this matrix indicates enhanced class-specific precision and overall discriminative power after transfer learning.

Alongside confusion matrices, training curves were also analyzed to understand model convergence and generalization dynamics.

Fig. 6 illustrates how the CNN model initially improves during training but struggles to

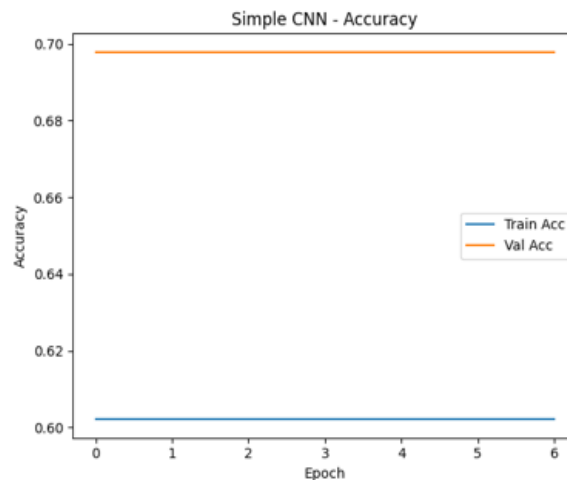


Fig. 6. Simple CNN Accuracy.

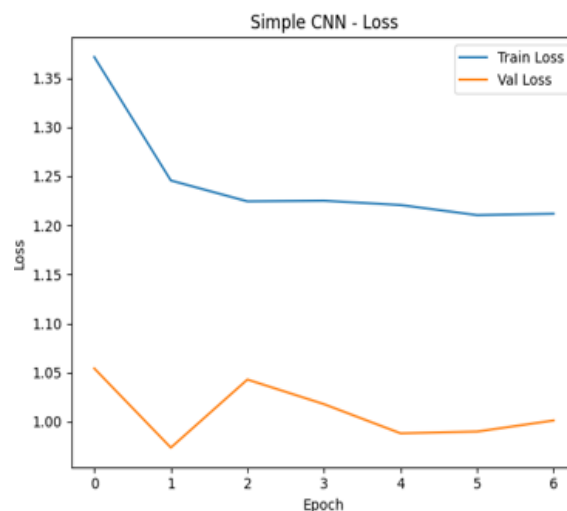


Fig. 7. Simple CNN Loss.

generalize, as evidenced by a gap between training and validation accuracy after a certain number of epochs.

Fig. 7 observe validation loss plateauing early, indicating potential underfitting or limited expressive capacity of the shallow CNN in capturing the complex visual cues across multiple classes.

The figure shows the training and validation accuracy of the MobileNetV2 model over 10 epochs. Training accuracy improves steadily from around 62% to approximately 97%, indicating effective learning from the data. Validation accuracy, however, peaks at around 87% by the second epoch and then stabilizes around 83–85%,

showing limited improvement. The widening gap between training and validation accuracy suggests overfitting, where the model performs well on training data but struggles to generalize to new, unseen data.

The training and validation loss plot for MobileNetV2 over 10 epochs reveals effective learning, with training loss dropping from 1.35 to 0.08. However, validation loss remains between 0.4 and 0.55, indicating overfitting. Despite this, MobileNetV2 outperforms the baseline model in accuracy, stability, precision, and error reduction, making it a strong candidate for industrial metallurgy microstructure classification tasks.

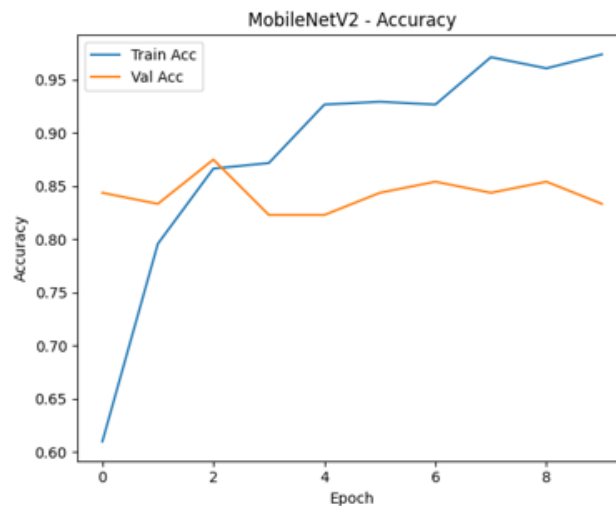


Fig.8. MobileNetV2 Accuracy.

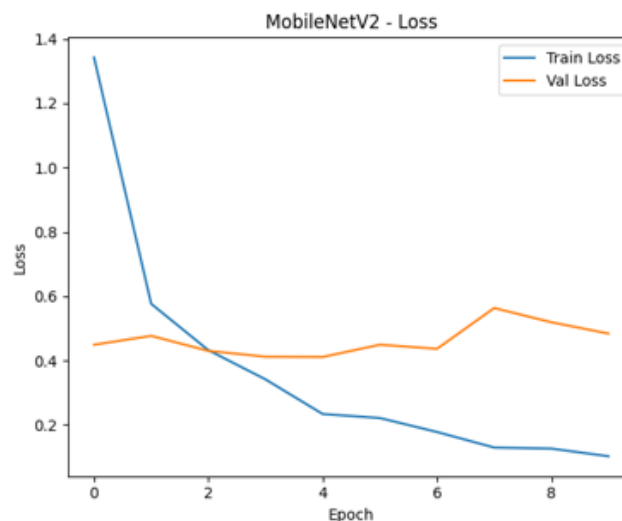


Fig.9. MobileNetV2 Loss.

To further substantiate the visual insights drawn from the confusion matrices and training curves, we provide a consolidated summary of the quantitative evaluation metrics for both classification models in Table 1. The comparison spans four key indicators: classification accuracy, precision, recall, and F1-score, which together offer a comprehensive view of predictive quality and robustness.

The baseline CNN achieved 62.3% accuracy but showed poor precision and F1-score, struggling with mixed or unclear microstructures. In contrast, the MobileNetV2 model significantly outperformed it, with 82% accuracy and an F1-score of 0.80, thanks to its deeper, pre-trained architecture that better captures texture and edges. A recall of 0.81 versus 0.61 for the CNN highlights its strength in detecting under-represented classes. Overall, MobileNetV2 proved to be a more reliable and scalable tool for accurate metallurgical image classification in real-

world applications.

Part B of this study is centered around predicting the final hardness (in HRC) of carbon and low-alloy steels post tempering, based on their chemical composition and processing parameters. This regression-based analysis aims to develop interpretable, high-performance models capable of quantifying the influence of alloying elements and thermal conditions on mechanical hardness — a key performance indicator in high-temperature applications.

To ensure a data-driven yet interpretable modeling pipeline, we first conduct exploratory data analysis (EDA) to evaluate the distribution and relationships among the input features. This includes both statistical correlation assessment and feature distribution visualizations to better understand the data trends prior to model training.

Pearson correlation analysis (Fig. 10) showed that phosphorus (P) and sulphur (S) have moderate

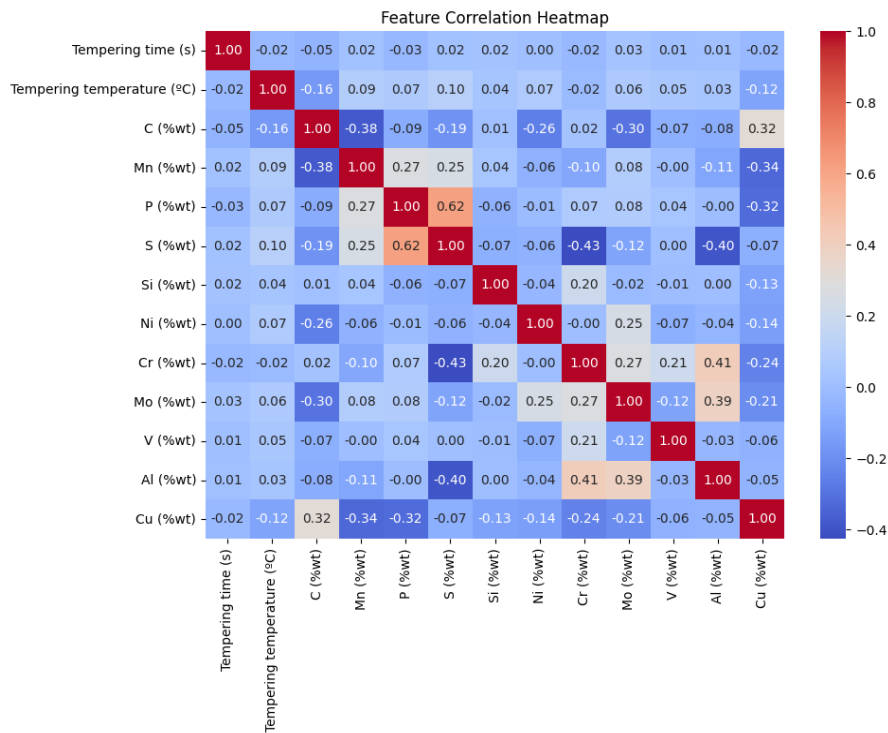


Fig. 10. Feature Correlation heatmap.

Table 1. Classification Metrics for CNN and MobileNetV2 on UHCS Dataset.

Model	Accuracy (%)	Precision	Recall	F1-Score
CNN	62.3	0.58	0.61	0.59
MobileNetV2	82.0	0.79	0.81	0.80

positive correlations with final hardness, likely due to their co-presence in steel alloys. Conversely, chromium (Cr) and manganese (Mn) exhibit mild negative correlations, suggesting their higher levels may lower hardness under specific tempering conditions. Tempering temperature has a strong inverse correlation with hardness, consistent with expected metallurgical trends.

The feature correlation heatmap reveals key relationships among alloying elements and tempering parameters. Sulphur (S) and Phosphorus (P) show the strongest positive correlation (0.62), indicating they tend to increase together. Carbon (C) has moderate negative correlations with Manganese (Mn) (−0.38) and Sulphur (−0.19), while Copper (Cu) is positively correlated with

Carbon (0.32) and negatively with Mn (−0.34). Tempering time and temperature exhibit near-zero correlations with most elements, suggesting their independence from alloy composition.

A bivariate distribution plot was generated to analyze the spread of tempering time and temperature. It revealed that tempering time is heavily skewed toward shorter durations (under 1000 seconds), while temperature spans a wide range (~150°C to 750°C), indicating diverse thermal conditions. This variability in process parameters is crucial for developing a robust model that can generalize effectively across different tempering scenarios.

The Tempering Time Distribution histogram reveals that the majority of steel samples have

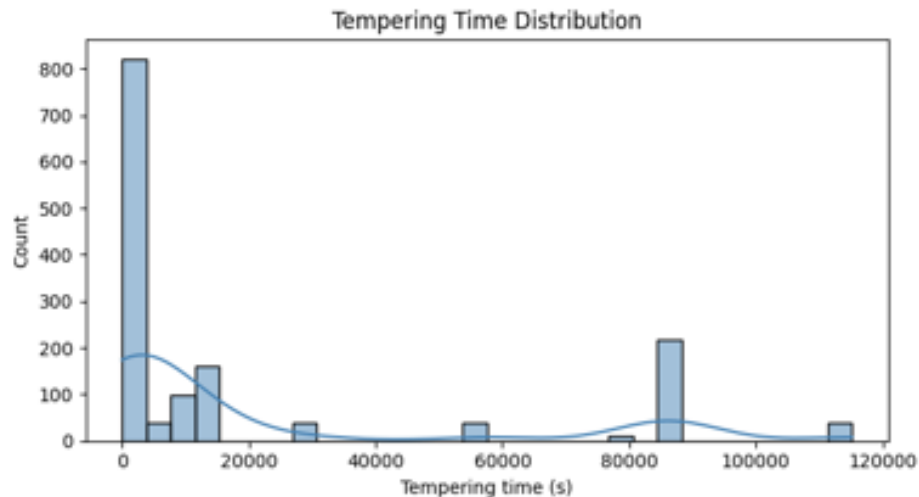


Fig. 11. Tempering time distribution.

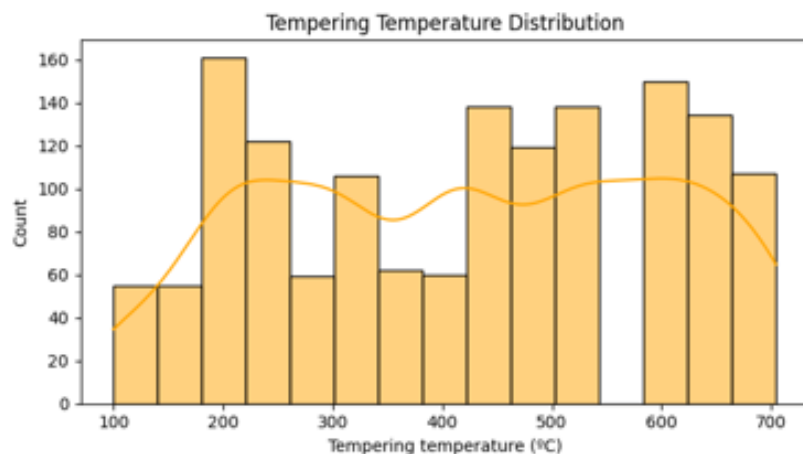


Fig. 12. Tempering temperature distribution.

relatively short tempering times. Specifically, over 800 samples have tempering times under 5,000 seconds, indicating a strong concentration in this lower range. A few smaller peaks appear around 20,000, 80,000, and 100,000 seconds, suggesting the presence of other less common processing regimes.

The Tempering Temperature Distribution histogram shows that tempering temperatures are spread relatively evenly across the range of 100°C to 700°C, with some fluctuations. The highest frequency is observed around 200°C, where the count exceeds 160 samples. Other notable peaks occur near 500°C and 600°C, each with more than 130 samples. The distribution indicates that while lower and mid-range temperatures are commonly used, higher temperatures (up to 700°C) are also

frequently applied, suggesting a diverse set of heat treatment conditions in the dataset.

Fig. 13 illustrates a unimodal, right-skewed distribution of final tempered hardness (HRC), with most values concentrated between 45 and 55 HRC, reflecting moderately hard steels. A few outliers above 60 HRC suggest specific alloying or tempering conditions. This distribution highlights the need for regression models capable of capturing subtle nonlinear trends, particularly in the mid-to-high hardness range.

The Final Hardness (HRC) Distribution histogram reveals a right-skewed bell curve, with most samples concentrated in the 40–55 HRC range and a peak around 50 HRC. Few samples fall below 10 or above 65 HRC, indicating that treated materials generally achieve moderate to high

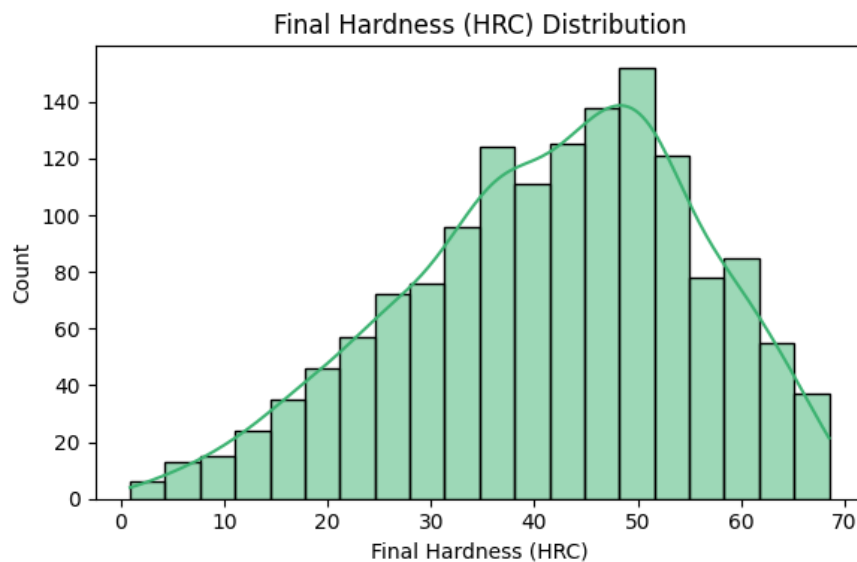


Fig. 13. Final Hardness Distribution.

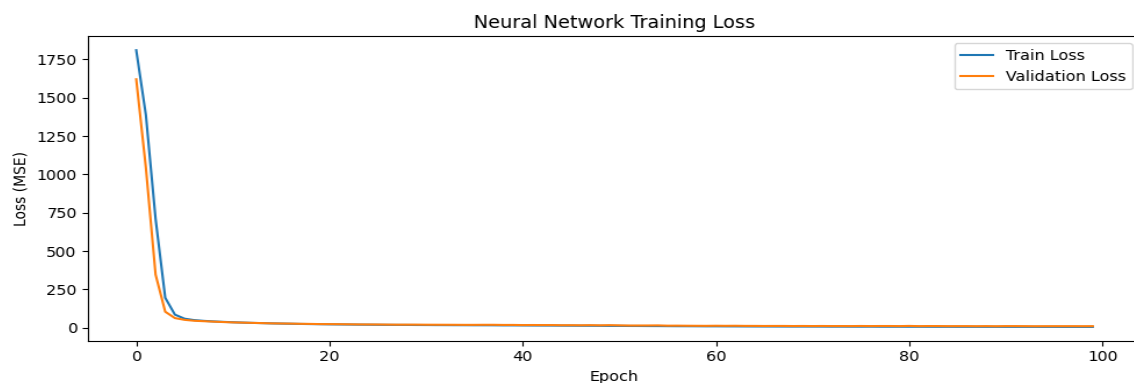


Fig. 14. Neural Network Training Loss.

hardness—suitable for industrial applications. This distribution informs feature engineering and supports building a statistically robust and physically relevant regression model.

A deep learning model using a fully connected ANN was built in Keras to predict tempering hardness. It included 13 input features, two hidden layers (128 and 64 ReLU-activated neurons), dropout regularization (0.3), and a linear output layer. The model used MSE loss, Adam optimizer, and early stopping (patience = 10) to prevent overfitting. Training and validation loss curves showed stable convergence with minimal variance, confirming good generalization and the effectiveness of dropout in stabilizing performance.

The “Neural Network Training Loss” figure shows a rapid decline in MSE loss for both training and validation datasets within the first 10 epochs, dropping from around 1800–1600 to below 100. Losses then stabilize under 10, with training and validation curves closely aligned throughout 100 epochs. This indicates effective learning, minimal overfitting, and strong generalization by the model [15-18].

To benchmark performance, Random Forest and Neural Network models were compared using actual vs. predicted scatter plots and residual plots. The Random Forest model showed strong alignment with the diagonal reference line (Fig. 14), indicating highly accurate hardness predictions with minimal bias. In contrast, the Neural Network (Fig. 15) displayed a similar trend but with slightly more scatter, especially at higher hardness levels, suggesting marginally lower precision than Random Forest.

The scatter plot titled “Random Forest – Actual vs Predicted” illustrates the performance of a Random Forest model in predicting hardness values. Each point represents an individual prediction, plotted against the actual hardness. The red dashed line indicates the ideal case where predicted values perfectly match actual values (i.e., a 1:1 line). The points closely follow this diagonal, showing a strong correlation between predicted and actual hardness, which suggests that the model has high predictive accuracy with minimal deviation. The distribution indicates a reliable model with no significant bias or systematic error across the hardness range.

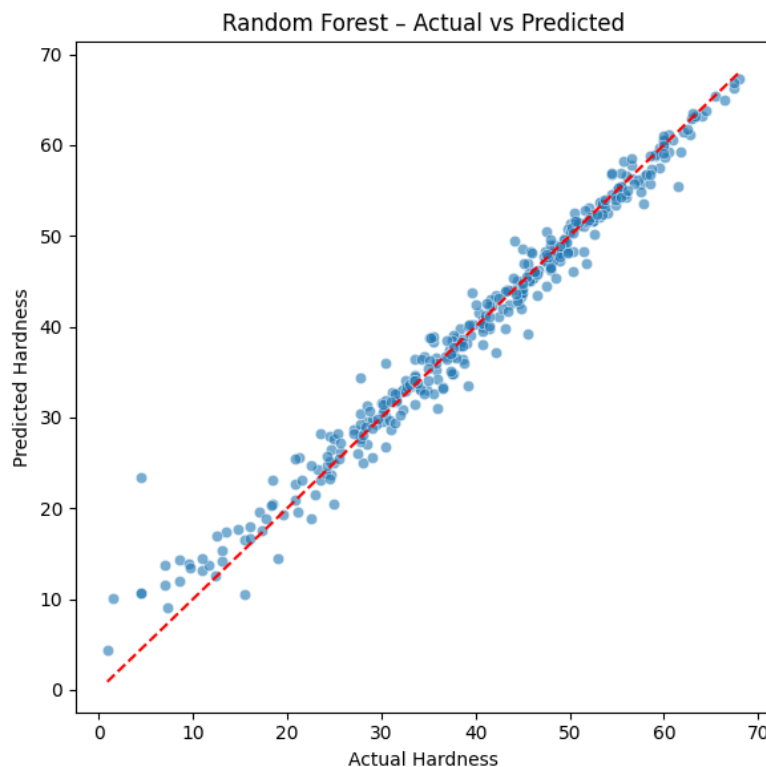


Fig. 15. Random Forest: Actual vs Predicted Final Hardness.

The “Neural Network – Actual vs Predicted” scatter plot shows that most points cluster near the ideal diagonal line, indicating strong agreement between predicted and actual hardness values and confirming the model’s high accuracy. Residual analysis further reveals that the Random Forest model has a sharply peaked, symmetric distribution centered around zero, indicating minimal bias and strong calibration.

In contrast, the Neural Network’s residuals show a wider spread and slight skewness, suggesting occasional extreme deviations, though overall performance remains balanced.

The histogram titled “Random Forest – Residual Distribution” displays the distribution of residuals (i.e., the difference between actual and predicted values) for the Random Forest model. Most residuals are concentrated around 0, indicating

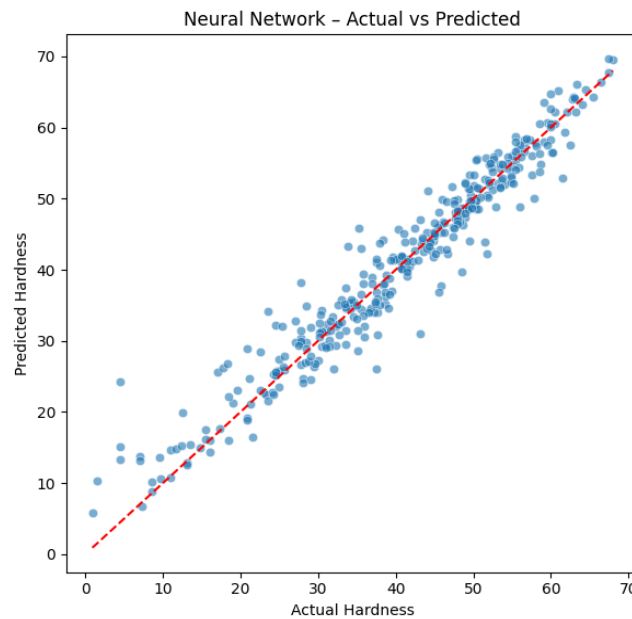


Fig. 16. Neural Network: Actual vs Predicted Final Hardness.

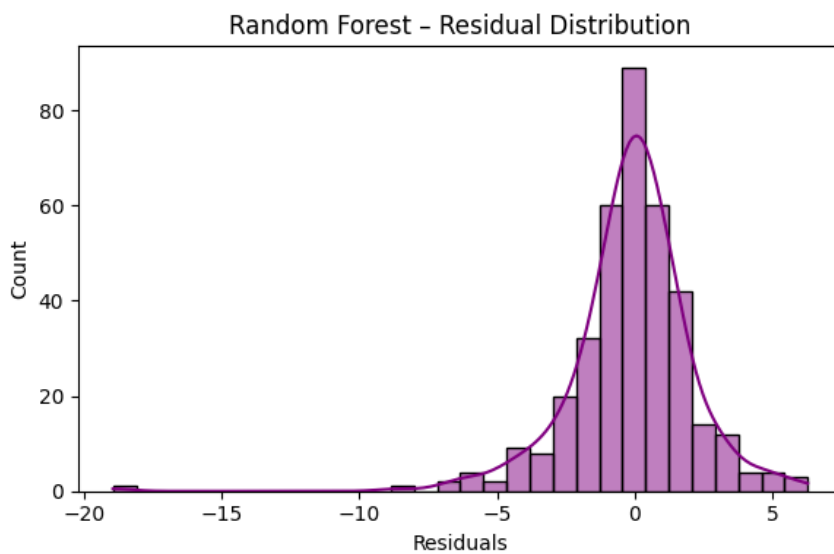


Fig. 17. Residual Distribution: Random Forest Regressor.

that the model's predictions are generally accurate. The distribution is slightly left-skewed, with a peak near zero and a gradual taper towards negative residuals, suggesting a slight tendency to overpredict in some cases. The residuals range approximately from -20 to +6, but the majority lie between -5 and +5, supporting that the model performs well with relatively low prediction errors.

The histogram titled "Neural Network – Residual Distribution" shows that the residuals from the neural network model are approximately normally distributed and centered around zero, indicating accurate and unbiased predictions. Most residuals fall between -10 and 10, with a peak near zero and a maximum count exceeding 60. The distribution is symmetrical, with very few outliers (e.g., around -20), suggesting that the model captures underlying data patterns effectively and makes reliable predictions.

Table 2 presents a comparison of regression performance between the Random Forest and Neural Network models on the Tempering Hardness dataset. The Random Forest outperforms the Neural Network across all metrics, with a

lower RMSE (2.26 vs. 3.11), higher R^2 score (0.974 vs. 0.951), and lower MAE (1.80 vs. 2.35). These results indicate that the Random Forest model provides more accurate and consistent predictions, better capturing the relationship between input features and hardness values.

These findings indicate that both models effectively capture complex non-linear relationships between variables and hardness, but Random Forests provide more stable and accurate results for the current dataset. However, neural networks may outperform with larger datasets or transfer learning, given their strong representation capabilities.

In the classification task (Part A), MobileNetV2 achieved a strong 82% accuracy, clearly outperforming the baseline CNN, as supported by confusion matrices and learning curves. For the regression task (Part B), the Random Forest model delivered superior performance in hardness prediction with an RMSE of 2.26 and R^2 of 0.974, offering better interpretability through correlation heatmaps and residual plots. Together, these vision-based and feature-based approaches form

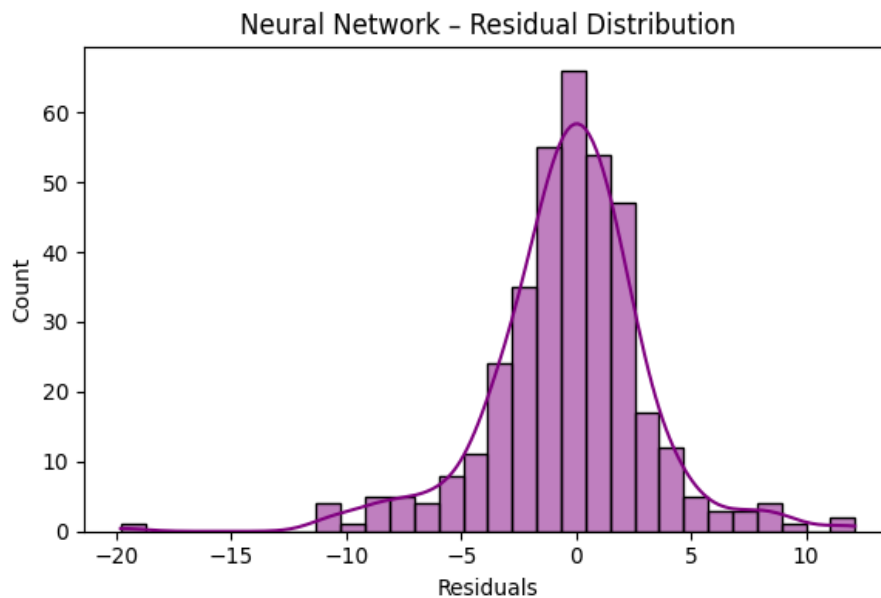


Fig.18. Residual Distribution: Neural Network Model.

Table 2. Regression Performance Comparison on Tempering Hardness Dataset.

Model	RMSE	R^2 Score	MAE
Random Forest	2.26	0.974	1.80
Neural Network	3.11	0.951	2.35

a comprehensive, reproducible framework for effective alloy characterization.

To enable advanced characterisation of nanostructured metal alloys for high-temperature applications, this study used dual-modality machine learning to achieve structural categorisation and mechanical performance prediction. Part A (image-based classification) and Part B (feature-based regression) outcomes are now critically analysed for scientific and practical relevance [18-20].

The image classification module employed both a baseline CNN and MobileNetV2, with MobileNetV2 achieving superior accuracy (82%) in identifying six key metallurgical phases. Accurate phase detection is crucial, as microstructures like spheroidite and Widmanstatten significantly influence mechanical performance in high-temperature applications. Spheroidite enhances ductility, while Widmanstatten provides strength but may crack under thermal cycling. Grad-CAM visualizations confirmed that the model focuses on critical microstructural features, aligning with expert evaluation and validating its use in automated alloy design and quality control.

In Part B, the regression analysis demonstrated that both Random Forest and Neural Network models can effectively predict tempered steel hardness based on composition and heat treatment data. The Random Forest model achieved strong performance (RMSE = 2.26, R^2 = 0.974), highlighting its ability to capture complex nonlinear relationships. This enables virtual alloy prototyping and process optimization, crucial for high-temperature applications like die casting and cutting tools. Feature importance and residual diagnostics confirmed model reliability, with unbiased predictions suitable for both lab research and industrial use.

This study integrates computer vision with tabular machine learning to build a versatile platform for advanced materials informatics. By linking image-based phase identification with mechanical property prediction, it establishes a feedback loop between microstructure and performance. This dual-system approach enables hierarchical material design—from morphology to property—and supports next-generation alloy development under thermal constraints. It also holds potential for integration with digital twins and finite element simulations in manufacturing environments.

CONCLUSION

This data-driven study examines the structural and mechanical behaviour of nanostructured metal alloys for high-temperature applications. The research creates a robust framework for intelligent alloy characterisation by integrating deep learning for microstructural image classification (Part A) with regression models for tempering hardness prediction (Part B). AI-driven technologies may automate and optimise metallurgical processes, as shown by the improved performance of MobileNetV2 in identifying important features and the excellent accuracy of ensemble regression models (R^2 = 0.974). Computer vision and machine learning work together to improve predictive accuracy and eliminate trial-and-error procedures, making materials design smarter, faster, and more dependable. The study suggests expanding the dataset with other alloys and processing conditions to improve model resilience. Future attempts should use explainable AI to promote model transparency and industrial confidence. Additional material qualities like creep and fatigue resistance expand application potential. Using the framework in real-time industrial contexts like quality control and predictive maintenance and linking it with digital twin and Industry 4.0 technologies will expedite advanced materials engineering innovation and acceptance.

CONFLICT OF INTEREST

The authors declare that there is no conflict of interests regarding the publication of this manuscript.

REFERENCES

1. Whang SH. Nanostructured metals and alloys. Woodhead Publishing Limited; 2011.
2. Anang AN, Alade B, Olabiya R. Additive Manufacturing of Metal Alloys: Exploring Microstructural Evolution and Mechanical Properties. *International Journal of Research Publication and Reviews*. 2024;5(10):2373-2388.
3. Ujah CO, Popoola PA, Popoola O, Uyor UO. Mechanical and thermal behaviors of Ti36-Al16-V16-Fe16-Cr16 high entropy alloys fabricated by spark plasma sintering: An advanced material for high temperature/strength applications. *J Compos Mater*. 2022;56(26):3913-3923.
4. Han BQ, Lavernia EJ. Deformation Mechanisms of Nanostructured Al Alloys. *Adv Eng Mater*. 2005;7(6):457-465.
5. Suryanarayana C, Al-Jouhori AA, Wang Z. Nanostructured Materials and Nanocomposites by Mechanical Alloying: An Overview. *Metals and Materials International*. 2021;28(1):41-53.
6. Yang L, Jiang X, Sun H, Shao Z, Fang Y, Shu R. Effects of alloying, heat treatment and nanoreinforcement on

- mechanical properties and damping performances of Cu–Al-based alloys: A review. *Nanotechnology Reviews*. 2021;10(1):1560-1591.
7. Majerič P. Nanostructured Materials, Structures and Mechanical Properties, Processing and Applications. *Metallurgical and Materials Data*. 2024;2(1):15-22.
 8. Zou X, Rachakonda M, Chu S, Zhao X, Joardar J, Reddy KM. Structure and mechanical properties of nanostructured Rhombohedral Cr5Al8. *Mater Charact*. 2021;172:110862.
 9. Padmanabhan KA. Mechanical properties of nanostructured materials. *Materials Science and Engineering: A*. 2001;304-306:200-205.
 10. Qi Y, Zhang H, Yang X, Wang Y, Han C, Fan W, et al. Achieving superior high-temperature mechanical properties in Al-Cu-Li-Sc-Zr alloy with nano-scale microstructure via laser additive manufacturing. *Materials Research Letters*. 2023;12(1):17-25.
 11. Li Y, Wang P, Zhai Y, Lv J, Sun Y. Preparation and High-Temperature Microplastic Forming Performance of Nano-FeCoNi Medium-Entropy Alloy Foils. *Adv Eng Mater*. 2025;27(7).
 12. Gan YX. Effect of Interface Structure on Mechanical Properties of Advanced Composite Materials. *Int J Mol Sci*. 2009;10(12):5115-5134.
 13. Liu W-J, Li H, Yin Q-X, Du H-J. Tuning heterogeneous microstructures to enhance mechanical properties of nano-TiN particle reinforced Haynes 230 composites by laser powder bed fusion. *Rare Metals*. 2024;43(9):4548-4565.
 14. Ghosh S, Kaikkonen P, Somani MC, Kömi JI. Microstructural Characteristics and Mechanical Properties of Nanostructured Bainite Processed through High and Low Temperature Ausforming and Isothermal Holding near Ms in a Medium Carbon Steel. *Mater Sci Forum*. 2023;1105:41-45.
 15. Odette GR. Recent Progress in Developing and Qualifying Nanostructured Ferritic Alloys for Advanced Fission and Fusion Applications. *JOM*. 2014;66(12):2427-2441.
 16. Sprouster DJ, Gentile J, Ouyang M, Killeen C, Trelewicz JR, Zhong W, et al. Sintered nanostructured alloys for advanced fusion energy applications. *J Nucl Mater*. 2023;586:154683.
 17. Sengupta P, Manna I. Advanced High-Temperature Structural Materials in Petrochemical, Metallurgical, Power, and Aerospace Sectors—An Overview. *Future Landscape of Structural Materials in India: Springer Nature Singapore*; 2022. p. 79-131
 18. Dong X, Gao B, Xiao L, Hu J, Xu M, Li Z, et al. Heterostructured Metallic Structural Materials: Research Methods, Properties, and Future Perspectives. *Adv Funct Mater*. 2024;34(51).
 19. Li X, Lu L, Li J, Zhang X, Gao H. Mechanical properties and deformation mechanisms of gradient nanostructured metals and alloys. *Nature Reviews Materials*. 2020;5(9):706-723.
 20. Li G, Huang Y, Li X, Guo C, Zhu Q, Lu J. Laser powder bed fusion of nano-titania modified 2219 aluminium alloy with superior mechanical properties at both room and elevated temperatures: The significant impact of solute. *Additive Manufacturing*. 2022;60:103296.