A Novel Approach to Jominy Profile Prediction Based on 1D Convolutional Neural Networks and Autoencoders that Supports Transfer Learning.

Marco Vannucci[1][0000−0002−6115−1588] and Valentina Colla[1][0000−0002−9574−0575]

TeCIP Institute, Scuola Superiore Sant'Anna, Pisa, Italy
{marco.vannucci,valentina.colla}@santannapisa.it

Abstract. This paper introduces a novel method for the estimation the Jominy profile of steel based on its composition, by combining autoencoders and 1-D Convolutional Neural Networks . The approach has two goals: firstly, to enhance the accuracy of hardenability prediction by exploiting the capability of the 1-D CNN to learn how the chemical composition of steel affects the shape of the Jominy profile; secondly, to use transfer learning to apply the knowledge gained from training on a specific dataset to new types of production with less available data or data with different characteristics as it often occurs in the industrial context. The proposed approach was tested on two industrial datasets aiming to assess the effectiveness of the methods on the two goals achieving satisfactory results.

Keywords: steel hardenability · 1-D convolutional neural networks · transfer learning.

1 Introduction

The Jominy test is important in steelmaking as it allows assessing steel hardenability, which is its ability to be hardened by quenching. This information is crucial in selecting and designing materials for specific applications where hardness, strength, and wear resistance are important factors.

The test involves the heating of one end of a cylindrical steel specimen up to the austenatizing temperature and its subsequent quenching through a water jet. This creates a gradient of cooling rates along the length of the sample, which produces a hardness profile when the sample is tested at fixed distances from the quenched end. The resulting Jominy profile is used to predict the ability of a steel alloy to be hardened by quenching, and is an important factor in determining the suitability of a steel for specific applications. By understanding the steel hardenability, steel producers can optimize the alloying and heat treatment processes to provide the product with the desired properties. This can save time and resources in the production process by avoiding costly trial-and-error methods. Moreover, the Jominy test is widely used as a quality control tool to ensure that the produced steel meets specific requirements. It is also an essential

test for research and development of new steel alloys, as it allows evaluation and comparison of different materials.

The ability of steel to harden is primarily determined by its chemical composition, as the content of certain elements such as C, B, Cr, N, Si, Mo, Nb, V, Ti has a significant impact. While the influence of these elements has been extensively studied, both individually and in combination, there is still some uncertainty regarding the exact relationship between chemical composition and the Jominy profile. This is due to the complex interactions between the chemical elements that affect the cooling behavior and ultimately determine the microstructure of the steel, which in turn affects its hardenability at different distances.

The Jominy end-quench test is very costly and time-consuming, therefore there is a strong economical interest in developing models to estimate hardenability from chemical composition. In this context, the cost of the Jominy test limits the availability of experimental data to set–up and tune such models when new types of steel are designed or produced.

In this paper a novel approach for estimating the Jominy profile from steel composition based on a combination between Autoencoders and Convolutional Neural Networks (CNN) is proposed. This method has a twofold purpose: the first is to improve the accuracy in hardenability prediction by learning how the steel chemical composition affects the *shape* of the Jominy profile through the CNN, and the second is to generalize and embed such knowledge obtained from an arbitrary training dataset to be able to reuse it, via transfer learning, while training the model for new types of productions with lower data availability.

The paper is organized as follows: in Section 2 the current state of art about the Jominy profile prediction is outlined with a special focus on approaches that employ Artificial Intelligence (AI) and, in particular, Artificial Neural Networks (ANNs) in Section 2.2. The proposed approach is described in detail in Section 3. The experimental set–up for the assessment of the method in the light of previously introduced objectives is depicted in Section 4, while the achieved results are reported and discussed in Section 5. Finally, Section 6 is devoted to drawing conclusions and outlining future lines of development of the approach.

2 Related works

The strategical importance of the information gained from the Jominy profile together with its cost led through the years to development of a multitude of models aiming to the prediction of the Jominy profile from the chemical composition. These methods can be grouped according to the exploited approach.

2.1 Numerical approaches

Phase field models simulate the evolution of the microstructure during the heat treatment process [2], and can predict Jominy profiles. These models consider the thermodynamics and kinetics of the phase transformations, and can provide information on the distribution of the different phases in the material which can

be correlated to steel hardenability at the Jominy distances [9] [5]. Similarly, Finite Elements Modelling (FEM) can simulate the heat transfer and phase transformations during the Jominy test, and can provide a detailed picture of the temperature and microstructure evolution [8]. These models require significant computational resources, but can provide a high level of accuracy. Quench Factor Analysis (QFA) involves analyzing the cooling curves and correlating them to the metallurgical response. This technique is used to estimate hardness from simulated cooling curves and has been shown to have a good correlation between the predicted and measured hardness. This approach developed in the 70-ies was adopted in [11] achieving an acceptable accuracy only for high hardness values.

2.2 Data driven approaches

Over time, the availability of information from numerous quenched–end Jominy tests has led to the formation of datasets that include the chemical composition of steels and their respective Jominy profile. These datasets have made the development of data-driven models possible.

Initially, traditional statistical techniques were used to make the first attempts in this area. One example of this is the approach was based on *multiplicators* and was improved and made more general in the following decades by various researchers, such as in [4] where the author proposed multiple statistical and empirical methods to estimate hardenability, while also discussing their advantages and limitations. In addition to these methods, regression analysis has also been employed with favorable results to predict hardenability, focusing not only on the method feasibility and optimization of models accuracy, but also investigating the key variables that impact each hardness value of the Jominy profile [6] paving the road to numerous successive researches. These methods, mainly due to their low complexity and linearity, demonstrated their accuracy only within limited ranges of steel grades. Such models often fail to generalize due to the fact that the relationship between the model parameters and the steel chemical composition is mostly based on empirical analysis and is challenging to learn. Furthermore, these models only offer satisfactory accuracy for a small number of points on the Jominy curve. The individual effect of each alloy element is typically examined, while interactions between them are ignored, resulting in reduced model accuracy. To address these issues, ANNs have been used as a tool to predict the Jominy profile since the 1990s.

Many studies used Multi-Layer Perceptron (MLP) ANNs to estimate the Jominy curve based solely on chemical composition. Some of these studies focused on microalloyed steel grades, while others considered steel for specific applications and included not only chemical composition, but also mechanical properties like yield strength and ultimate tensile strength as input of the ANN. However, all these approaches neglected the correlations between neighboring hardness values in the Jominy profile. To address this issue, in [3] a parametric approach is proposed in which the Jominy profile is represented by a parametric mathematical function of the distance from the quenched end, such as a *quasi–sigmoidal* monotonic decreasing function. Wavelet neural networks are

then applied to correlate the steel chemistry with the function parameters. In [1] a sequential predictor was developed, in which each point of the Jominy profile is predicted by a dedicated ANN that exploits the content of specific chemical elements selected according to theoretical knowledge of the phenomenon and the predicted hardness at other distances. In this latter work a measure of the reliability of the predicted hardness is associated to each point of the profile. Finally, despite the dizzying development of applications based on the use of deep learning, there are still very few works that employ such technology for Jominy profile prediction. Among these, it is worth mentioning [7], in which an optimized Deep Convolutional Neural Network (DCNN) is used and in which the convolutional layer is applied directly to steel chemistry.

3 Proposed approach

In this paper, a novel approach to the design of a Jominy profile predictor using the steel chemical composition is proposed. The model aims at improving the accuracy of existing predictors not through a punctual estimation of hardness at different distances from the specimen quenched-end, but through the determination of the profile shape and by linking it to the steel chemical composition. This kind of approach aims at overcoming some typical deformations of the profile, which are usually fixed during post-processing and result in an accuracy degradation. Furthermore, relating the chemistry to the profile shape instead of single points improves the robustness to outliers within the data.

In addition, the proposed model aims at *storing* the relationship between chemical composition and profile shape to make it available in a transfer-learning context. Such an approach may be useful for training models for new types of steel for which the available datasets are small, or for transferring the model to different steel plants, which are characterized by different ways of collecting experimental data or performing the experimental tests (i.e. different instrumentation) that may alter the aforementioned relationship.

The proposed Jominy profile predictor consists of two main models depicted in Figures 1 and 2:

1. an *autoencoder* named *JAutoencoder* (Fig. 1) that has the task of mapping an arbitrary Jominy profile measured at the 15 standard points (such domain is subsequently referred as D_{J15}) into a latent space L characterized by a significantly smaller dimension;
2. a DNN named *JNetwork* (Fig. 2) that performs the actual prediction of the profile, by mapping the steel chemistry into the 15 points of profile, exploiting the JAutoencoder.

JAutoencoder is a quasi typical autoencoder whose goal is to learn a compressed representation of the input dataset composed by Jominy profiles, such that the original data can be accurately reconstructed from the compressed representation. Its main peculiarity is that the first layer performs a 1-Dimensional

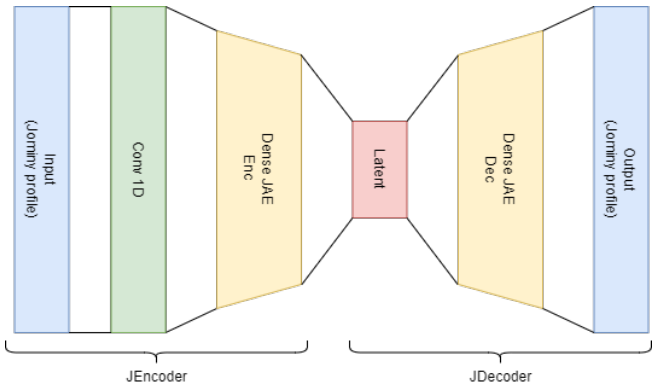

Fig. 1. Schematic architecture of the JAutoencoder.

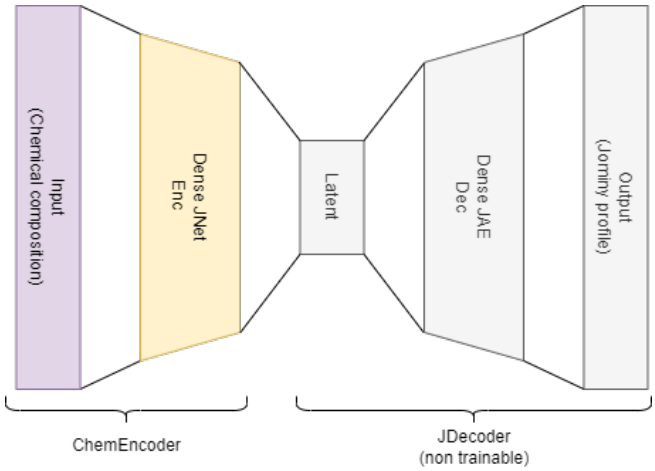

Fig. 2. Schematic architecture of the JNetwork putting into evidence the *ChemEncoder* and the *JDecoder*, non–trainable within this network.

convolution (coupled to a max-padding operator) aiming at extracting knowledge about the shape of the fed Jominy profiles. After the convolutional layer, the autoencoder is composed of different fully connected layers, symmetrical with respect to the latent space L that can represent compactly the original D_{J15} domain, that accounts the shape of the handled Jominy profiles because of the 1D convolutional layer. The fully connected layers of JAutoencoder are activated through the *LeakyRelu* function, which is widely used for autoencoders implementation. In addition, both $L1$ and $L2$ type regularisation are employed to reduce the risk of overfitting.

Once JAutoencoder is trained using experimental data, it is decomposed into an encoder and a decoder, as shown in Figure 1 and denoted as JEncoder and JDecoder, respectively. The latter is exploited for its ability to reconstruct a Jominy profile from a compact representation of it that considers the original shape of the profile. Specifically, JDecoder is embedded within a DNN denoted as *JNetwork* in Figure 2, downstream of a chemical encoder (*ChemEncoder*) that has the task of mapping the chemical composition of a steel to the corresponding L representation. JNetwork, the DNN thus obtained, is connected and takes as input the chemical composition of the steel to return the corresponding predicted Jominy profile. In this context, the JDecoder present within JNetwork is **not trainable** and uses its own internal parameter values as got from JAutoencoder training. JNetwork is then trained with an experimental dataset resulting from Jominy testing that thus includes the chemical compositions of steel and the respective Jominy hardness profiles. This training procedure only affects the *trainable* parameters within the chemical encoder layers, supporting the mapping of chemical compositions into a latent representation that, once fed to the (frozen) JDecoder, leads to the predicted Jominy profile exploiting the knowledge it embeds through transfer learning. The architecture of both the JAutoencoder and the JNetwork were optimized through empirical tests described in Section 4. Optimization involves the number of layers of the two networks, including the dimension of the latent space L, as well as the number of neurons in each layer. The 1D convolutional layer was fine tuned in terms of the number of employed filters, filters dimensions, pooling dimension.

4 Experimental tests

The performance of the proposed approach was evaluated through a test campaign that involves two datasets. The pursued tests aim at evaluating both the model accuracy and its capability of transferring knowledge regarding the Jominy profile shape through the JDecoder sub-system as discussed in Section 3.

4.1 Available datasets

The datasets come from two distinct industrial plants and include a different number of samples: the bigger one, referred as *Dataset A*, is formed by 1500

observations; the other, *Dataset B*, by 250. Observations include steel chemical composition and the result of the Jominy quench-end test.

The two datasets correspond to different products in terms of both chemical composition and mechanical properties. Figure 3 puts into evidence such difference comparing the average Jominy profile throughout the datasets, reported in the figure together with the punctual standard deviation of hardness.

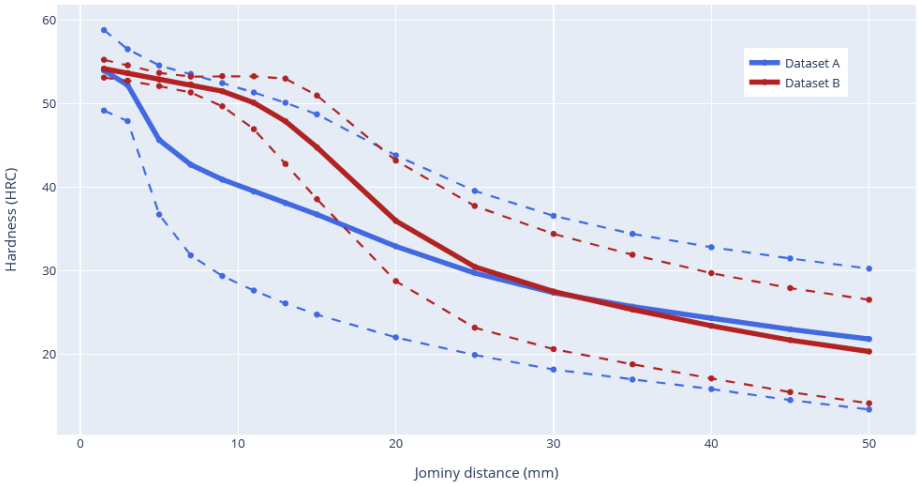

Fig. 3. Comparison between the Jominy profiles present in Dataset A and Dataset B. Average profile is shown together with point–wise standard deviation.

4.2 Tests description

The accuracy in the prediction of Jominy profile from the steel chemical composition of the proposed approach was evaluated mainly based on Dataset A, which contains a number of observations large enough to support the training of the models and a reliable performance assessment on validation and test data. To this aim, Dataset A was preliminary divided into two parts: the first one, composed of 80% of the data, was used for training and validation purposes, the remaining 20% for test. Training and validation data are used for selecting the hyper–parameters values of JAutoencoder and JNetwork. The considered parameters include:

- number of layers and neurons per layer within the JEncoder and JDecoder (symmetric in this work) of the JAutoencoder
- number of layers and neurons per layer within the ChemEncoder of JNetwork
- latent space (L) dimension within JAutoencoder and JNetwork
- number of filters used within the 1D convolutional layer of the JAutoencoder

– dimension of 1D convolutional filters within JAutoencoder
– pooling dimension of the MaxPooling layer following the 1D convolution of the JAutoencoder.

The tested values for the listed hyper–parameters are summarized in Table 1. For each combination of hyper–parameters a 10–fold cross validation is set up and the models are tuned by using training data and evaluated by using the validation data. The combination resulting in a lower average prediction error throughout the points of the Jominy profile is then selected for comparison with other approaches. The ability to use the approach used for the development of JNetwork in a transfer learning context was evaluated through the joint use of Dataset A and Dataset B. Specifically, once the JNetwork was trained as described above using Dataset A, the corresponding JDecoder was extracted and reused in a JNetwork trained using Dataset B that actually exploits the information coming from Dataset A in the context of the original JAutoencoder. The results achieved are compared to those obtained by the sequential predictor presented in [1], which actually is the best performing model for micro–alloyed steel such those present in the available dataset and to a fully–connected ANN optimized in terms of architecture (i.e. layers and number of neurons) similarly to the JNetwork. In the case of knowledge transferability assessment, the proposed approach is compared to models that are trained using only Dataset B to evaluate the benefits of the approach.

Table 1. Values of hyper–parameters tested during the models tuning phase. Layers are described through the number of neurons for each layers in brackets.

Hyper–parameter	**Values**
1D Conv. filters	2,3,5,10
Dimension 1D Conv. filters	2,3,4
Pooling dimension	2,3,4
Latent space dimension	2,3,4,5,6
JEncoder–JDecoder layers	(10,10), (10,10,10), (20,10), (20,20), (20,20,20), (30,30)
ChemEncoder layers	(10,10), (10,10,10), (20,10), (20,20), (20,20,20), (30,30)

5 Results and Discussion

5.1 Base model evaluation

The best performing hyper–parameters combinations resulting from the grid search on the values reported in Table 1 are shown in Table 2 in terms of mean absolute error (MAE) achieved by JAutoencoder and by JNetwork on the validation sets (within the CV framework) and by JNetwork on the test dataset.

In addition to those shown in the table, most of the combinations tested achieve satisfactory results in terms of MAE, in line with the industrial requirements of the application. Among the best ones, it is worth highlighting the low

number of convolutional filters used and their size (2), which is sufficient to extract the necessary features from the Jominy profile likely due to the low number of points the profile itself is formed by. Furthermore, the optimal dimension of the latent space L resulting from the test is limited (4 in best performing cases), which, considering the achieved low JAutoencoder error, highlights that the main characteristics of the Jominy profile shape can be efficiently compacted.

Table 2. Best performing combinations of the hyper–parameters reported oin table 1 achieved while training the models on Dataset A.

1D–Conv filters	**Filters dim.**	**Pool. dim.**	L **dim.**	**JAut layers**	**JNet layers**	**JAut. CV MAE**	**JNet. CV MAE**	**JNet. TS MAE**
3	2	2	4	(20, 20)	(20, 20)	0.22	0.85	0.96
2	2	2	4	(30, 30)	(30, 30)	0,20	0,95	1,04
4	2	3	4	(10, 10)	(20, 20)	0,38	0,98	1,07
2	2	2	5	(20, 20)	(20, 20)	0,22	1,00	1,07
2	2	2	3	(20, 20)	(20, 20)	0,22	1,05	1,13

The best performing model was selected according to the performance of the JNetwork in terms of MAE within the CV (first row of Table 2). The average prediction error (MAE) of this model on the test data within the Dataset A is 0.96 HRC. The above introduced *sequential* model achieves a MAE of 1.10 HRC (+16% with respect to JNetwork) on the same data after being trained with the remaining observations of Dataset A.

A fully connected DNN was tested as well by using the same data. The DNN architecture was optimized testing different number of layers and hidden neurons similarly to what was done for the hyper–parameters optimization of the proposed approach. The best performing set–up in this case, a three layers network with (20,20,20) neurons in the hidden layers respectively and $ReLU$ activation function, achieves a MAE of 1.12 HRC (+17% with respect to the proposed approach). The punctual behaviour of the three compared methods are shown in Figure 4 in terms of prediction error. The barplot shows that the sequential model outperforms the other ones in the initial part of the profile, which are strongly related to a few chemical elements. On the other hand, the JNetwork is much more accurate in the central part of the profile (which is crucial from the industrial point of view) and obtains a more balanced performance throughout the whole profile, demonstrating the correct *learning* of how chemical composition affects the overall shape of the profile. The JNetwork also performs excellently qualitatively as shown in Figure 5, in which two predicted Jominy profiles are compared to the corresponding measured ones are shown as examples.

5.2 Transfer learning evaluation

The effectiveness of the proposed transfer learning approach based on the reuse of the JDecoder component was evaluated according to the procedure described

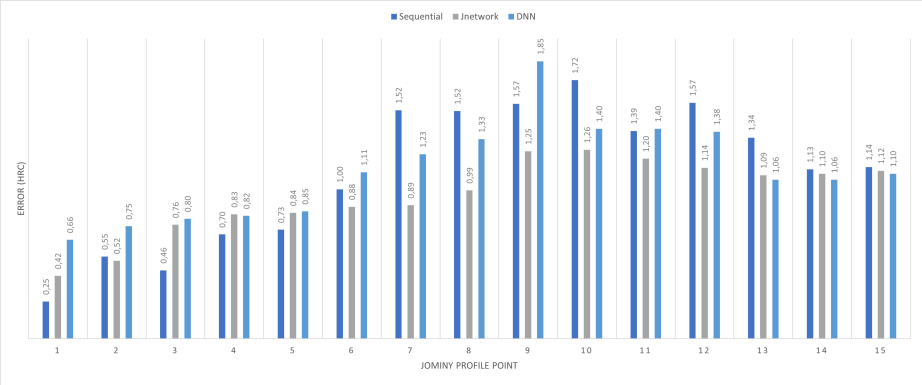

Fig. 4. Point–wise error of JNetwork, sequential model and DNN model on test data after training with Dataset A.

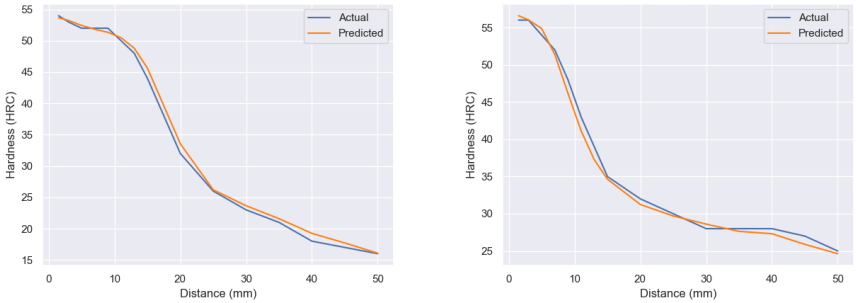

Fig. 5. Examples of two predictions performed by the JNetwork on test data from the Dataset A.

in Section 4. The straightforward approach to feed Dataset B input samples to the JNetwork trained by using only Dataset A led to the achievement of a MAE higher than 4 HRC throughout the profile, which is unacceptable from the industrial point of view. This failure is probably due to the difference, both in terms of chemistry and Jominy profile shape, between the two datasets. This result encourages the application of the proposed approach. The results obtained were compared to those of the sequential model and a fully connected ANN trained using only the data in Dataset B, partitioned into training, validation and testing with proportions 70%, 15%, 15% respectively. The ANN model was optimized in terms of architecture in a similar manner to what was done for Dataset A, obtaining a network of smaller size with respect to the previous one: two layers both holding 10 neurons activated via ReLU. The average error on test data of the JNetwork is 1.09 HRC, while the one of the sequential mode and of the fully connected ANN are 1.34 HRC (23% greater) and 1.40 HRC (29% greater),

respectively. These error values are sensibly higher than the one obtained by the proposed approach. The punctual errors of the three approaches are depicted in Figure 6 that confirms the goodness of the proposed method and puts into evidence that the prediction of the other approaches shows low accuracy in the central region of the profile, where the uncertainty of the shape is higher. The poor performance of these methods is likely due to the low number of samples available for training and validation of the employed ANNs. This latter issue is overcome by the JNetwork approach by transferring useful knowledge from the pre–trained JDecoder.

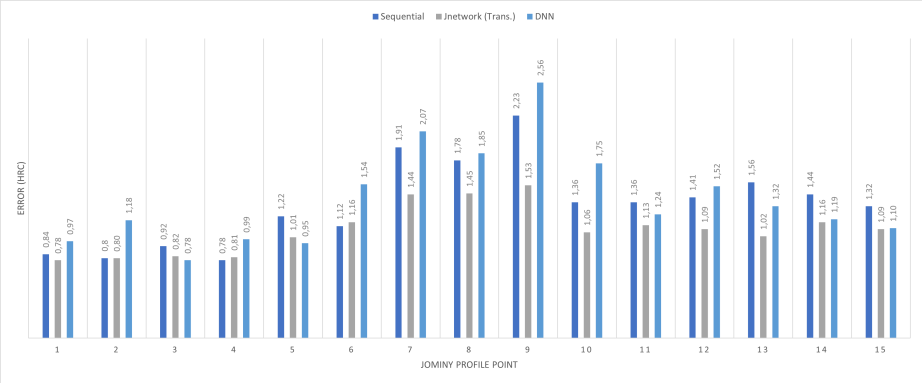

Fig. 6. Point–wise error of JNetwork trained via a transfer learning approach on Dataset B after the training of the JDecoder with Dataset A, sequential model and DNN model on test data after training with Dataset B.

6 Conclusions and future work

The paper presented a new approach to Jominy hardness profile prediction based on the use of an autoencoder using a 1D convolutional layer to learn a compressed encoding of the profile shape. Subsequently, the steel chemistry is mapped into that encoding using a fully connected ANN. The method exploits the idea of *learning* the Jominy profile shape rather than the point value of hardness at standard distances, one of the weaknesses of many existing approaches. In addition, the method is suitable to transfer learning by using the pre-trained JDecoder component and exploiting its ability to reconstruct the complete profile from its compressed encoding. The proposed approach was tested using two different industrial datasets showing very good results both in terms of accuracy of predictions and in the possibility of being used for transfer learning purposes in the rather common cases of experimental data scarcity (i.e. new product types, different plants). These results encourage further developments of this technology that will involve testing in a material design context as in [10], the evaluation of

different and types of autoencoders for profile encoding, and development of a *chemical encoder* that exploits the theoretical knowledge regarding the influence of various chemical elements in different regions of the profile.

References

1. Cateni, S., Colla, V., Vannucci, M., Vannocci, M.: Prediction of steel hardenability and related reliability through neural networks. In: IASTED Multiconferences—Proceedings of the IASTED international conference on artificial intelligence and applications, AIA. pp. 169–174 (2013)
2. Colla, V., Desanctis, M., Dimatteo, A., Lovicu, G., Valentini, R.: Prediction of continuous cooling transformation diagrams for dual-phase steels from the intercritical region. Neural Computing and Applications **33**(23), 16451–16470 (2021)
3. Colla, V., Reyneri, L.M., Sgarbi, M.: Neuro-wavelet parametric characterization of jominy profiles of steels. Integrated Computer-Aided Engineering **7**(3), 217–228 (2000)
4. Doane, D.V.: A critical review of hardenability predictors. Hardenability Concepts with Applications to Steel pp. 351–396 (1977)
5. Kirkaldy, J.: Prediction of microstructure and hardenability in low alloy steels. In: Proceedings of the International Conference on Phase Transformation in Ferrous Alloys, 1983. AIME (1983)
6. Komenda, J., Sandström, R., Tukiainen, M.: Multiple regression analysis of jominy hardenability data for boron treated steels. Steel research **68**(3), 132–137 (1997)
7. Li, C., Yin, C., Xu, X.: Hybrid optimization assisted deep convolutional neural network for hardening prediction in steel. Journal of King Saud University-Science **33**(6), 101453 (2021)
8. Li, M.V., Niebuhr, D.V., Meekisho, L.L., Atteridge, D.G.: A computational model for the prediction of steel hardenability. Metallurgical and Materials transactions B **29**, 661–672 (1998)
9. Saunders, N., Guo, U., Li, X., Miodownik, A., Schillé, J.P.: Using jmatpro to model materials properties and behavior. Jom **55**(12), 60–65 (2003)
10. Vannucci, M., Colla, V.: Automatic steel grades design for jominy profile achievement through neural networks and genetic algorithms. Neural Computing and Applications **33**(23), 16451–16470 (2021)
11. Yazdi, A.Z., Sajjadi, S.A., Zebarjad, S.M., Nezhad, S.M.: Prediction of hardness at different points of jominy specimen using quench factor analysis method. Journal of materials processing technology **199**(1-3), 124–129 (2008)