Pretraining Techniques for Steel Surface Roughness Prediction with Long Thin Spatial Industrial Data

Alex Milne[1,2][0009-0000-9149-4146], Xianghua Xie[1,3][0000-0002-2701-8660], and Gary K.L. Tam[1,3][0000-0001-7387-5180]

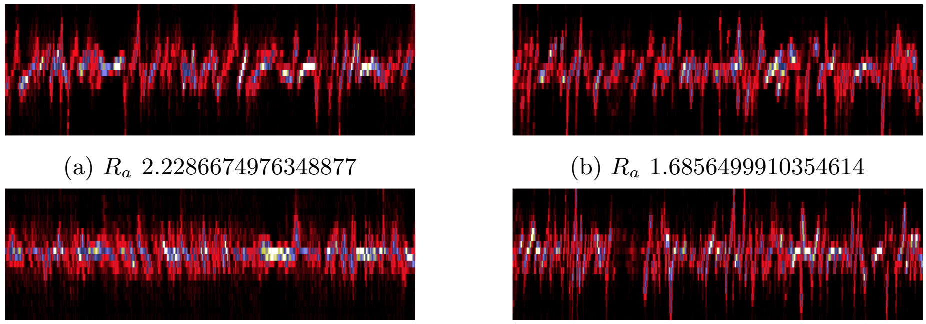
[1] Department of Computer Science, Swansea University, UK
[2] alexjmmilne@gmail.com
[3] {X.Xie,k.l.tam}@swansea.ac.uk

Abstract. Machine learning offers promising advancements in industrial processes, yet collecting labeled samples during production remains challenging. In steel production, the surface roughness R_a parameter of steel coils is crucial, but on-line labeled data collection, with our apparatus, is infeasible, while off-line methods are time-consuming and imperfect. However, unlabeled samples are readily available from on-line production. This paper examines pretraining on a large, unlabeled dataset and its impacts on performance after fine-tuning on a smaller labeled dataset. We use three techniques: (1) contrastive learning, (2) Autoencoder, and (3) Classification of coil ID. We address the challenges posed by the unique structure of the data, comprising 2-dimensional, long and thin arrays. Our results show that our classification pretraining approach improves regression performance and outperforms the baseline.

Keywords: Representation learning · High-aspect-ratio images · Pretext task · Contrastive learning · Autoencoder

Surface roughness, particularly the average roughness parameter R_a, is a critical quality metric in strip steel production, especially for high-value applications such as automotive paneling. Surface texture directly influences formability during pressing and determines the effectiveness of downstream processes like paint adhesion [4]. To achieve target roughness, temper mill operators adjust process variables including roll force, speed, and roll texture. However, conventional methods for measuring R_a rely on stylus-based profilometry, which is performed manually after production. These measurements are time-consuming, spatially limited, and unsuitable for real-time feedback, making mid-coil corrections impossible. This further delays quality verification, disrupts just-in-time workflows, and leads to costly line stoppages and lower customer satisfaction.

To overcome these limitations, this work focuses on a novel laser-based surface scanning system capable of generating structured spatial image data in-line during production. Unlike traditional profilometry, laser scanning is scalable and provides dense coverage of the steel surface. However, the laser system's raw measurements lack the accuracy of stylus measurements and cannot be directly



(a) R_a 2.2286674976348877

(b) R_a 1.6856499910354614

(c) R_a 0.8971018195152283

(d) R_a 1.2550691366195679

Fig. 1: Visualization of grayscale laser measurement scans using colormap `gist_stern`. Each subplot shows a different sample.

aligned with them, due to differences in spatial resolution and sensor characteristics. Accurate labels for roughness parameters like R_a still require destructive off-line measurements on sampled steel sections under controlled laboratory conditions. Consequently, labeled datasets are limited and expensive to obtain, while unlabeled laser scan data is abundant.

This discrepancy creates a compelling opportunity to leverage self-supervised and transfer learning techniques. By pretraining on large volumes of unlabeled laser scan data, models can learn informative representations that improve R_a regression performance when fine-tuned with scarce labeled samples. Representation learning, particularly self-supervised learning (SSL), has proven effective in domains with limited labeled data [14, 22, 18], and has shown broad success in computer vision [5, 15, 8] and time-series modeling [29, 33, 20], including industrial applications [1, 24].

However, our data presents unique challenges. Each laser scan produces a high-aspect-ratio spatial image, a long and narrow 2D array formed from angular reflection intensity measurements across 20 sensors arranged in a semicircle, which scan the steel surface as it moves through the production process. These sensors capture texture variations as the steel moves beneath them. Unlike conventional vision-based square or near-square images of objects or natural scenes, this data consists of structured image representations of laser reflections, capturing surface texture variations rather than traditional visual features. Its non-standard content and dimensions make direct application of standard vision models challenging, requiring specialized processing, augmentations, and strategies to ensure robust surface roughness estimation. Figure 1 shows examples of such reflection profiles.

A key challenge in deploying roughness estimation models in production is that conventional 2D convolutional neural networks (CNNs), though effective for general image tasks, are computationally intensive and fail to meet real-time industrial demands [21]. Prior models struggle to process full-resolution laser scan images at line speed. However, the data's structure, long, narrow, and spatially organized, suggests it is well-suited to 1D convolutional processing, which can efficiently capture localized patterns along the scan direction while reducing computational overhead [21].

Motivated by this and the availability of large-scale unlabeled data, we reformulate the problem using 1D convolutional models for fast, on-line inference. This raises our central question: can 1D convolutional architectures, combined with transfer learning, match or exceed the predictive performance of more complex 2D models while meeting strict industrial real-time speed requirements?

To explore this, we evaluate three pretraining strategies that leverage unlabeled scan data: (1) contrastive self-supervised learning (SSL) using SimCLR [5] and MoCo [15], with domain-specific augmentations tailored to the scan structure; (2) autoencoding via temporal convolutional networks (TCNs) [19] to learn compressed representations; and (3) supervised pretraining on a coil classification task to extract transferable features. These are compared against a supervised baseline trained solely on the limited labeled R_a dataset. Specifically, our contributions are:

- We introduce a novel application of transfer learning, including self-supervised and supervised pretraining, to the task of R_a prediction from laser-based structured image data in steel production.
- We design and evaluate domain-specific augmentations and training protocols for contrastive SSL, autoencoders, and supervised pretraining, and show their effectiveness in improving regression performance with limited labeled data.
- We show state-of-the-art results compared among real-time techniques for steel surface R_a prediction from laser images [21].

1 Related Work

Prediction of R_a Value Prior studies have used machine learning to predict surface roughness in metal processing. Elangovan et al.[12] estimated roughness during lathe turning by analyzing statistical properties like mean and skewness. In steel strip production, Cheri et al.[9] proposed an intelligent control method for cold rolling, using on-line measurements to adjust roll force and control surface roughness. Our previous work [21] extended this by predicting the R_a value directly from laser-based light sensing data, improving accuracy and aligning with stylus-based ground truth standards. This study builds on that work by exploring whether self-supervised pretraining can enhance prediction performance with limited labeled data. Our earlier work [21] is the only deep learning technique that achieves real-time performance. This work shows that we can achieve better performance through pretraining while remaining real-time.

Learning with High Aspect Ratio Data Beyond conventional near-square images commonly used in vision tasks, many domains deal with nonstandard image formats that require specialized processing. For example, panoramic images and rental panoramic radiographs exhibit extreme aspect ratios that challenge standard architectures [13]. Similarly, whole slide histopathology images span exceptionally large spatial resolutions [25], posing challenges for memory

Table 1: Summary of Dataset Statistics

	Labeled Dataset	**Unlabeled Dataset**
Number of samples	775	52,388
Number of steel coils	49	112
Coil Id	Available	Available
Surface roughness labels (R_a)	Available	Not available
Dimension: 20 X T	T = 65,536	T = 50,000
Sampling interval	0.8 μm	1 μm

and computational efficiency even if their aspect ratios are more moderate. Our dataset, consisting of laser reflection profiles in steel production, represents an extreme case, with a resolution of 65536 by 20, making it highly elongated and creating unique demands for representation learning and model scalability. These laser reflection profiles form long, narrow two-dimensional arrays that encode structured surface texture information but lack object-based semantics typical of conventional vision datasets. Prior work [21] explored both 2D and 1D approaches, finding that 2D CNNs are computationally expensive and unsuitable for real-time industrial constraints. Instead, 1D models, widely used in Time Series Extrinsic Regression (TSER) [23] and Time Series Classification (TSC) [17], offer a more suitable approach for this data structure. While high-capacity models such as Vision Transformers (ViTs) have been evaluated, they underperformed compared to Temporal Convolutional Networks (TCNs) [19, 2], which achieved the best results while satisfying industrial speed constraints [21]. These findings highlight the need for architectures that effectively model structured spatial dependencies while remaining computationally feasible for real-time deployment. Building on this, our work focuses on extending transfer learning to structured laser reflection data using large-scale, unlabeled datasets. Through self-supervised pretraining, we improve prediction performance in scenarios with limited labeled data, demonstrating the broader applicability of modern representation learning techniques beyond standard image domains.

Self-Supervised Learning Gui et al. [14] review self-supervised learning (SSL) methods, highlighting Contrastive Learning (CL), which distinguishes unlabeled instances by generating distinct views of a data point. CL originated from instance discrimination [26, 32], with key methods including MoCo v1 [15], MoCo v2 [8], SimCLR v1 [5], and SimCLR v2 [6], all showing near-supervised performance. For time series data, approaches like TimeCLR [29], TS2Vec [30], TF-C [31], BTSF [28], and MF-CLR [11] have been proposed. Additionally, generative models like BEiT [3], Contextualized Learning [7], and Masked Image Modeling [27], often using the Vision Transformer (ViT) [10], have extended SSL to vision tasks. Our work builds on these methods by applying self-supervised pretraining for surface roughness prediction using large-scale unlabeled datasets, aiming to improve predictive performance and speed under limited labeled data conditions, which has not been fully addressed in prior research.

2 Data

This study uses laser reflection data from an industrial scanning system designed for surface roughness measurement on steel coils. The data consists of 20 sensor channels (width) captured by an arc of sensors, varying in length (T) representing measurement increments along the surface. Although the data is not in traditional image format, it exhibits spatial dependencies that can be represented as $20 \times T$ images. Key challenges include the long, narrow aspect ratio (T up to $\sim$65K), surface variability, resolution differences, limited labeled data, and misalignment between stylus and laser measurements.

Two datasets are employed: a small, labeled fine-tuning dataset from a controlled lab environment and a large, unlabeled dataset gathered during production. The labeled dataset contains 775 samples from 49 coils, with R_a ground truth obtained via stylus-based profilometry. Each coil is assigned a mean R_a based on averaged measurements, which are used as labels for corresponding laser scans. Each scan covers a 5.24 cm surface length with a 0.8 μm sampling interval ($T = 65{,}536$), therefore corresponding to a lateral resolution of 0.8 μm. This meets the minimum sampling density recommended by ISO 3274 [16] for roughness evaluation and is sufficient to capture the dominant surface features relevant to R_a estimation. The high number of sampling points ensures spatial continuity and statistical robustness across the scan length.

The unlabeled dataset, used for pretraining, consists of 52,388 samples from 112 coils, capturing laser reflections across a 5 cm surface length with a 1 μm sampling interval ($T = 50,000$). While lacking stylus measurements and with reduced resolution, its broader variability helps enhance pretraining robustness.

Table 1 summarizes dataset characteristics, showcasing the trade-offs between labeled and unlabeled data. High-resolution laser measurements offer sufficient detail for surface roughness regression, and pretraining compensates for the scarcity of labeled data.

3 Methodology

This section outlines the methodology for investigating representation learning approaches to predict surface roughness (R_a). We explore three pretraining strategies: (1) contrastive learning to align views from the same sample while distinguishing different samples, (2) autoencoding for learning compact representations via input reconstruction, and (3) coil classification to capture coil-specific variations linked to R_a.

All approaches use a TCN backbone with 8 layers, 25 channels, and convolutional kernels of size 7. The TCN, which performed well in prior work [21], extracts features for pretraining, followed by fine-tuning for R_a regression. During fine-tuning, the pretrained backbone is transferred and adapted for regression by replacing the pretraining-specific prediction head with a randomly initialized linear layer. This transfer learning approach ensures that all methods leverage the same feature extractor model while isolating the effects of the learned representations rather than the regression layer's capacity.

3.1 Contrastive Learning Approach

We explore contrastive learning to leverage large amounts of unlabeled data and learn robust representations for surface roughness prediction. Contrastive learning helps in extracting useful features by comparing augmented views of the same sample while pushing apart representations of different samples. Our framework uses a TCN backbone to extract features from laser reflection data. Each input sample is split into two non-overlapping views along the length dimension (T), with a 10-pixel gap to avoid edge artifacts, retaining all 20 sensor channels. The TCN processes these views to generate embeddings, minimizing the distance between views of the same sample and maximizing the distance between those from different samples.

To adapt contrastive learning, we use custom augmentations tailored to the dataset. Color jitter (brightness and contrast) simulates sensor variability, Gaussian blur mimics signal diffusion, and high percentile value assignment mitigates bright channel artifacts. Spatial augmentations like flips and Gaussian noise add variability, while the Weakening Threshold transformation emphasizes high-intensity patterns. Row intensity shifts simulate sensor degradation or lens fouling. Augmentations are applied with 50% probability to preserve structure and enhance diversity. Additionally, a queue-based memory mechanism inspired by MoCo [15, 8] helps maintain a large number of negative samples despite memory constraints.

3.2 Autoencoding

We explore the autoencoding approach to learn compressed yet informative representations of laser reflection data by reconstructing inputs from a latent embedding. This method helps capture essential structure while discarding noise, supporting better generalization in downstream tasks. Our custom-designed TCN-based autoencoder, comprising an encoder, bottleneck, and decoder, is employed to learn these representations, tailored to the unique characteristics of the dataset. The encoder uses the TCN backbone to learn a representation, which is then compressed through a custom bottleneck module. The bottleneck includes a convolutional layer with stride 2, followed by a flattening operation and a fully connected layer that reduces the data to a latent embedding of size 512 with 3 bottleneck channels. The bottleneck is reversed to expand the latent embedding back into a higher-dimensional space, and the decoder reconstructs the original input. This design enables meaningful pattern recovery while balancing compression and feature retention.

3.3 Coil Classification

As each sample comes from a specific coil, we explore coil classification as a supervised pretraining task to exploit this naturally occurring label. The model predicts which steel coil a laser sample came from, leveraging within-coil similarity and between-coil variability to learn features useful for downstream R_a

Table 2: Fine-tuning performance after pretraining. Lower RMSE and Max Abs. Error indicate better performance; higher Correlation and Coverage are preferred. SPS (Samples Per Second) is the number of test samples processed per second; SPS $\geq$ 60 is required for real-time steel surface inspection.

Method	**RMSE** ↓	**SPS** ↑	**Correlation** ↑	**Max Abs. Error** ↓	**Coverage (%)** ↑
TCN Baseline	0.0585 ± 0.0061	277.96	0.9336 ± 0.0142	0.1937 ± 0.0402	62.05 ± 4.41
Contrastive SSL	0.0515	277.96	0.9491	0.1529	64.47
Coil Classification	**0.0466**	277.96	**0.9583**	0.1598	**69.74**
TCN Autoencoder	0.0474	277.96	0.9569	**0.1381**	66.45
2D xresnet18*	0.0621	40.38	–	–	–
2D xresnet34*	**0.0453**	22.16	–	–	–
2D ViT Small*	0.1729	39.28	–	–	–

**2D Models from prior experiments on the same data for comparison [21].*

regression. Coil IDs are available for both labeled and unlabeled data. Pretraining is performed on the full unlabeled dataset and the train split of the labeled data. A TCN model is used, with coil ID as the prediction target.

4 Experiment Setup and Evaluation Metrics

The effectiveness of each pretraining strategy is evaluated by fine-tuning on R_a regression and comparing to a baseline trained from scratch using the same architecture and conditions. Fine-tuning uses a randomly initialized regression head and is evaluated on a 20% test split from the labeled dataset. Performance is assessed via RMSE, correlation, max absolute error (worst-case deviation), and prediction coverage—the proportion of predictions within the observed R_a range, reflecting roughness variability. Models are implemented in PyTorch, with pretraining in FastAI and fine-tuning via a custom loop. Experiments run on 3090 GPUs, using batch sizes of 12 (pretraining) and 4 (fine-tuning). For contrastive learning, MoCo [15, 8] enables effective negative sampling with a memory queue.

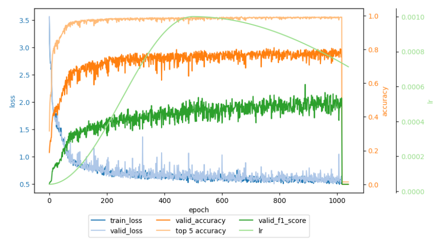
5 Experiment Results

This section compares our pretrained representation learning methods to a fully supervised baseline trained from scratch on the labeled dataset. To assess effectiveness, we track fine-tuning performance across pretrained models saved every 50 epochs, enabling analysis of how pretraining duration affects R_a regression.

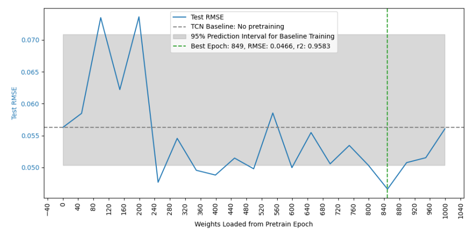
5.1 Transfer learning Performance

Table 2 compares all models, including a supervised TCN baseline averaged over 292 runs to ensure statistical robustness. Pretrained models are evaluated using their best checkpoints (saved every 50 epochs), allowing us to analyze how fine-tuning performance varies across pretraining duration, shown in Figures 2a–2e.

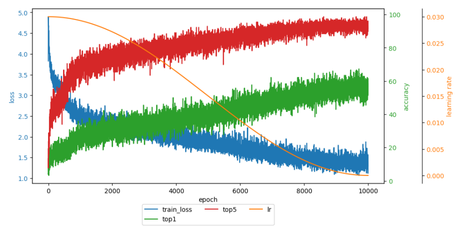
Among pretraining methods, coil classification performs best across most metrics, while the autoencoder achieves the lowest max absolute error, due to



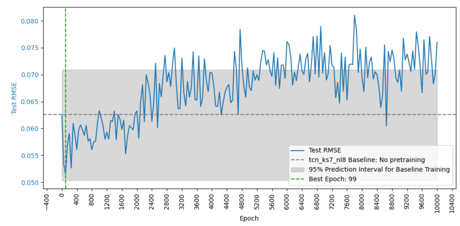
(a) Coil ID Classification Pretraining



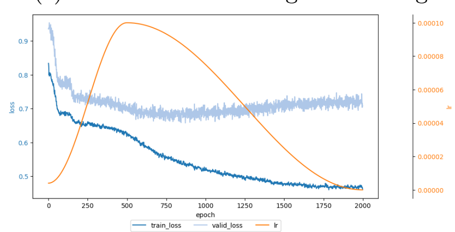
(b) Coil ID Classification Fine-tuning



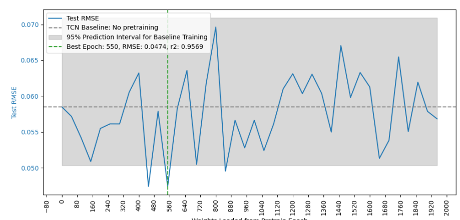
(c) Contrastive Learning Pretraining



(d) Contrastive Learning Fine-tuning



(e) Autoencoder Pretraining



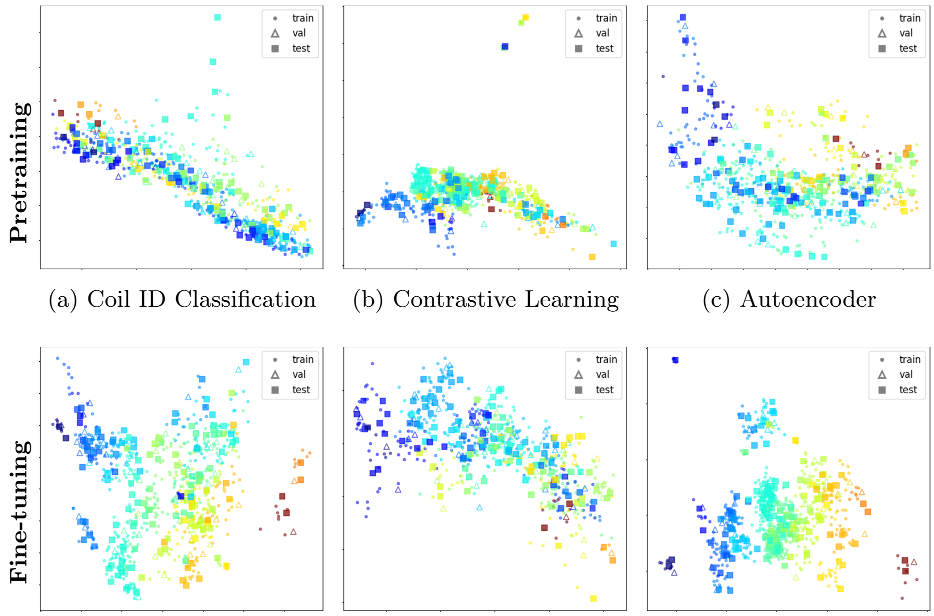
(f) Autoencoder Fine-tuning

Fig. 2: These plots show a comparison of pretraining metrics and the resulting best fine-tune loss when fine-tuning is performed with starting weights from the corresponding pretrained epoch. On the left are the pretraining loss curves, while on the right are the curves for the fine-tune loss. Fine-tuning results show the RMSE on the y-axis; therefore, lower is better. Results are better when these values show a consistent trend lower than the 95% gray box such as in 2b.

a single outlier in coil classification. These results highlight the benefit of pretraining with task-relevant structure while ensuring deployment constraints are met.

An important deployment constraint is inference speed, measured in Samples Per Second (SPS). Real-time steel surface inspection requires SPS ≥ 60, derived from a 300 cm/s line speed and 5 cm sample size. The TCN models exceed this threshold with 277.96 SPS, supporting industrial viability. In contrast, 2D models fall short; xresnet18 achieves 40.38 SPS, while xresnet34 drops to 22.16. Despite similar RMSE, these models are computationally unsuitable for deployment. Our 1D reformulation achieves similar accuracy while running up to 12.5× faster.

Figure 2a shows coil classification performance across epochs. Variability (standard deviation and 95% percentile) from the multiple baseline runs serves



(a) Coil ID Classification (b) Contrastive Learning (c) Autoencoder

(d) Coil ID Classification (e) Contrastive Learning (f) Autoencoder

Fig. 3: These plots show the first two principal components from PCA, reducing the high-dimensional feature space to 2D while preserving key variance. The top row (a,b,c) displays feature representations after pretraining, and the bottom row (d,e,f) after fine-tuning. Each point represents a data sample, with colors indicating ground truth R_a values and shapes representing the train-validation-test splits. Well-structured representations position points with similar R_a values to create a smooth, non-overlapping color gradient.

as a benchmark to assess whether improvements are due to pretrained representations or stochastic factors. Early pretraining (below 250 epochs) underperforms the baseline, but substantial improvements occur after 250 epochs, peaking at 850 epochs and consistently surpassing the baseline's 95% threshold. Other methods show limited or negative trends with extended pretraining. Figure 2c shows minimal improvement from contrastive learning up to 2000 epochs, with performance declining afterward. Similarly, the TCN autoencoder's performance fluctuates randomly, as shown in Figure 2e, with no clear trend based on increased pertaining duration.

5.2 Discussion and Qualitative Analysis

To assess the impact of pretraining on R_a regression, we visualize the learned representations using PCA. The first two components are plotted, with points colored by R_a values to show how well the representation aligns with the target variable. A well-structured representation should form a smooth gradient, with similar R_a values grouped together.

The PCA plots reveal key differences between pretraining methods. Contrastive pretraining shows poor alignment with the R_a gradient, with significant overlap and disorganization, failing to capture regression-relevant features. Autoencoder pretraining results in a well-structured representation with a clear R_a gradient, indicating alignment with the regression task. Coil ID classification pretraining shows moderate structure and a visible gradient, suggesting partial alignment with the regression task. Overall, pretraining methods aligned with the regression task yield more effective representations.

After transfer learning, all methods show improved representation structure. Contrastive learning shows gains, but misalignment persists, with overlapping high and low R_a values, limiting its effectiveness as a pretraining strategy. Autoencoder representations remain highly structured post-fine-tuning, with a near-perfect gradient and minimal overlap, providing a strong initialization for regression. Coil ID classification also improves, showing a strong gradient with occasional discrepancies. These results confirm that autoencoder and Coil ID pretraining significantly enhance fine-tuning, while contrastive pretraining struggles to transfer useful features, highlighting the need for methods aligned with downstream regression goals.

6 Conclusion

This study explores representation learning and transfer learning for improving surface roughness (R_a) predictions from structured laser-based image data in steel production. Unlike conventional near-square vision datasets, our work handles long and thin images, making 2D CNNs computationally infeasible for real-time deployment. Instead, we reframe to a 1D convolutional approach, achieving comparable predictive performance while running 12.5 times faster.

Our findings show that pretraining on large-scale unlabeled laser scan data improves regression performance, demonstrating the value of transfer learning in mitigating limited labeled data while meeting industrial speed constraints. Among the strategies tested, coil ID classification pretraining yields the most significant improvement.

Despite these advancements, variability in pretraining effectiveness highlights the need for aligning pretraining objectives with the regression task. While domain-specific augmentations were applied, the contrastive method led to negative learning effects, reinforcing the importance of methods that preserve structured dependencies in ways relevant to R_a estimation.

This study confirms that transfer learning, when carefully designed, enhances roughness prediction even under real-time constraints. Future work should refine task-aligned pretraining methods, reduce variability, and develop hybrid approaches to optimize feature extraction and inference speed, broadening the applicability of representation learning in industrial settings.

Acknowledgements This work was funded by EPSRC Industrial Case award (EP/V519601/1). For the purpose of open access the authors have applied a Creative Commons Attribution (CC BY) licence to any Author Accepted Manuscript version arising from this submission.

References

1. Abdulrazzaq, M.M., Ramaha, N.T., Hameed, A.A., Salman, M., Yon, D.K., Fitriyani, N.L., Syafrudin, M., Lee, S.W.: Consequential advancements of self-supervised learning (ssl) in deep learning contexts. Mathematics **12**(5), 758 (2024)
2. Bai, S., Kolter, J.Z., Koltun, V.: An empirical evaluation of generic convolutional and recurrent networks for sequence modeling. arXiv:1803.01271 (2018)
3. Bao, H., Dong, L., Piao, S., Wei, F.: Beit: Bert pre-training of image transformers. arXiv:2106.08254 (2021)
4. Bastawros, A.F., Speer, J.G., Zerafa, G., Krupitzer, R.P.: Effects of steel surface texture on appearance after painting. Tech. rep., SAE Technical Paper (1993)
5. Chen, T., Kornblith, S., Norouzi, M., Hinton, G.: A simple framework for contrastive learning of visual representations. In: International conference on machine learning. pp. 1597–1607. PmLR (2020)
6. Chen, T., Kornblith, S., Swersky, K., Norouzi, M., Hinton, G.E.: Big self-supervised models are strong semi-supervised learners. Advances in Neural Information Processing Systems **33**, 22243–22255 (2020)
7. Chen, X., Ding, M., Wang, X., Xin, Y., Mo, S., Wang, Y., Han, S., Luo, P., Zeng, G., Wang, J.: Context autoencoder for self-supervised representation learning. International Journal of Computer Vision **132**(1), 208–223 (2024)
8. Chen, X., Fan, H., Girshick, R., He, K.: Improved baselines with momentum contrastive learning. arXiv:2003.04297 (2020)
9. Cheri, S.H., Wang, J.G., Gu, T.Q.: An online intelligent control method for surface roughness of cold-rolled strip steel. In: 2018 37th Chinese Control Conference (CCC). pp. 8330–8335. IEEE (2018)
10. Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., Dehghani, M., Minderer, M., Heigold, G., Gelly, S., et al.: An image is worth 16x16 words: Transformers for image recognition at scale. arXiv:2010.11929 (2020)
11. Duan, J., Zheng, W., Du, Y., Wu, W., Jiang, H., Qi, H.: MF-CLR: multi-frequency contrastive learning representation for time series. In: Forty-first International Conference on Machine Learning (2024)
12. Elangovan, M., Sakthivel, N., Saravanamurugan, S., Nair, B., Sugumaran, V.: Machine learning approach to the prediction of surface roughness using statistical features of vibration signal acquired in turning. Procedia Computer Science **50**, 282–288 (2015), big Data, Cloud and Computing Challenges
13. Gao, S., Yang, K., Shi, H., Wang, K., Bai, J.: Review on panoramic imaging and its applications in scene understanding. IEEE Transactions on Instrumentation and Measurement **71**, 1–34 (2022)
14. Gui, J., Chen, T., Zhang, J., Cao, Q., Sun, Z., Luo, H., Tao, D.: A survey on self-supervised learning: Algorithms, applications, and future trends. IEEE Trans. Pattern Anal. Mach. Intell. **46**(12), 9052–9071 (Dec 2024)
15. He, K., Fan, H., Wu, Y., Xie, S., Girshick, R.: Momentum contrast for unsupervised visual representation learning. In: 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). pp. 9726–9735 (2020)
16. International Organization for Standardization: Geometrical product specifications (gps)–surface texture: Profile method–nominal characteristics of contact (stylus) instruments. ISO 3274:1996 (1996)
17. Ismail Fawaz, H., Lucas, B., Forestier, G., Pelletier, C., Schmidt, D.F., Weber, J., Webb, G.I., Idoumghar, L., Muller, P.A., Petitjean, F.: Inceptiontime: Finding alexnet for time series classification. Data Mining and Knowledge Discovery **34**(6), 1936–1962 (2020)

18. La Fé-Perdomo, I., Ramos-Grez, J., Mujica, R., Rivas, M.: Surface roughness ra prediction in selective laser melting of 316l stainless steel by means of artificial intelligence inference. Journal of King Saud University-Engineering Sciences **35**(2), 148–156 (2023)
19. Lea, C., Flynn, M.D., Vidal, R., Reiter, A., Hager, G.D.: Temporal convolutional networks for action segmentation and detection. In: proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 156–165 (2017)
20. Malhotra, P., TV, V., Vig, L., Agarwal, P., Shroff, G.: TimeNet: Pre-trained deep recurrent neural network for time series classification. arXiv:1706.08838 (2017)
21. Milne, A., Xie, X.: Steel surface roughness parameter calculations using lasers and machine learning models. The International Journal of Advanced Manufacturing Technology (2024)
22. Saric, T., Simunovic, G., Simunovic, K.: Use of neural networks in prediction and simulation of steel surface roughness. International Journal of Simulation Modelling **12**(4), 225–236 (2013)
23. Tan, C.W., Bergmeir, C., Petitjean, F., Webb, G.I.: Time series extrinsic regression: Predicting numeric values from time series data. Data Mining and Knowledge Discovery **35**, 1032–1060 (2021)
24. Tran, D.H., Nguyen, V.L., Nguyen, H., Jang, Y.M.: Self-supervised learning for time-series anomaly detection in industrial internet of things. Electronics **11**(14), 2146 (2022)
25. Wetstein, S.C., de Jong, V.M., Stathonikos, N., Opdam, M., Dackus, G.M., Pluim, J.P., van Diest, P.J., Veta, M.: Deep learning-based breast cancer grading and survival analysis on whole-slide histopathology images. Scientific reports **12**(1), 15102 (2022)
26. Wu, Z., Xiong, Y., Yu, S.X., Lin, D.: Unsupervised feature learning via non-parametric instance discrimination. In: 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 3733–3742 (2018)
27. Xie, Z., Zhang, Z., Cao, Y., Lin, Y., Bao, J., Yao, Z., Dai, Q., Hu, H.: SimMIM: A simple framework for masked image modeling. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 9653–9663 (2022)
28. Yang, L., Hong, S.: Unsupervised time-series representation learning with iterative bilinear temporal-spectral fusion. In: Chaudhuri, K., Jegelka, S., Song, L., Szepesvari, C., Niu, G., Sabato, S. (eds.) Proceedings of the 39th International Conference on Machine Learning. Proceedings of Machine Learning Research, vol. 162, pp. 25038–25054. PMLR (17–23 Jul 2022)
29. Yang, X., Zhang, Z., Cui, R.: TimeCLR: A self-supervised contrastive learning framework for univariate time series representation. Knowledge-Based Systems **245**, 108606 (2022)
30. Yue, Z., Wang, Y., Duan, J., Yang, T., Huang, C., Tong, Y., Xu, B.: Ts2vec: Towards universal representation of time series. In: Proceedings of the AAAI conference on artificial intelligence. vol. 36, pp. 8980–8987 (2022)
31. Zhang, X., Zhao, Z., Tsiligkaridis, T., Zitnik, M.: Self-supervised contrastive pre-training for time series via time-frequency consistency. Advances in Neural Information Processing Systems **35**, 3988–4003 (2022)
32. Zhao, N., Wu, Z., Lau, R.W.H., Lin, S.: What makes instance discrimination good for transfer learning? (2021)
33. Zheng, X., Wang, T., Cheng, W., Ma, A., Chen, H., Sha, M., Luo, D.: Parametric augmentation for time series contrastive learning. arXiv:2402.10434 (2024)