

HUMBOLDT-UNIVERSITÄT ZU BERLIN  
MATHEMATISCH-NATURWISSENSCHAFTLICHE FAKULTÄT  
INSTITUT FÜR INFORMATIK

# Deep Learning Techniques for Mineral Analysis in X-Ray Fluorescence Spectra

Exposé Master Thesis

eingereicht von: Thomas Meza  
geboren am: 29.05.1997  
geboren in: Berchtesgaden  
Gutachter/innen: Prof. Dr. Ulf Leser  
ausstehend  
eingereicht am: .....

# 1 Introduction

X-ray fluorescence (XRF) spectroscopy is a technique used for example for non-destructive elemental analysis of geological and mineralogical samples. By measuring X-ray fluorescence line positions and intensities of X-ray emissions from a material, XRF provides rapid and cost-effective insights into the elemental composition of geological samples, enabling applications ranging from ore characterization to environmental monitoring [4, 21, 26, 28]. Traditional XRF analysis workflows often encounter challenges such as spectral variability, grain size effects and detector noise. These methods typically require manual preprocessing steps, including baseline correction and peak deconvolution, which are time-consuming and prone to human error [3, 17]. Jones et al. highlight that conventional XRF analysis methods, for example Fundamental Parameters approach, require extensive background knowledge for accurate interpretation [16]. Similarly, Allegretta et al. note that calibrating XRF instruments require reference minerals, making the process both time-consuming and costly - challenges that machine learning methods could potentially overcome [2].

Recent advancements in deep learning have shown significant potential in overcoming these limitations. Techniques such as convolutional neural networks (CNNs) and recurrent neural networks have been applied to enhance classification accuracy, robustness against noise, and segmentation efficiency. Preisler et al. showcased the utility of CNNs for macro-XRF datasets, improving spectral classification even under conditions of high noise, thus reducing errors in material identification [22]. Tang et al. extended this approach by integrating CNNs for 3D mineral identification, highlighting their effectiveness in improving spatial resolution and processing efficiency in noisy environments [24]. Similarly, Andric et al. utilized autoencoder-based models to streamline dimensionality reduction and spectral classification in pigment identification applied on paintings [3].

Despite these advances, adapting these methods to real-world XRF-based mineral segmentation remains challenging. Existing studies highlight limitations in model generalizability and sensitivity to measurement conditions. Xu et al. observed that conventional CNN models for near-infrared spectral imaging suffered from reduced classification accuracy without proper preprocessing, emphasizing the critical role of spectral feature extraction techniques [29]. Preisler et al. highlighted additional challenges specific to XRF analysis. Traditional deconvolution methods often struggle with complex spectra due to scattering and background noise. They further stated that the reliance of synthetic datasets on experimental setup parameters, such as x-ray source configuration, geometry and detector properties, limits their applicability to different instrument setups and sample conditions [22]. Additionally, direct segmentation of mineralogical maps from XRF spectra remains an underexplored area.

This thesis explores the development of a novel framework for mineral identification and segmentation that builds on a baseline CNN model. To improve both accuracy and generalizability, the thesis focuses on comprehensive synthetic data generation and augmentation strategies. Specifically, real-world distortions such as Poisson noise, calibration shifts and Bragg reflections are added to the synthetic spectra to bridge the gap between idealized simulations and field conditions. Additionally, the potential benefit

of integrating attention mechanisms is explored by comparing the performance of the CNN with and without attention.

## 2 Goals and Dataset Description

The dataset used for this thesis will be generated from both real and simulated XRF spectra, providing a robust foundation for spectral classification and image segmentation tasks. Each spectrum represents an energy distribution, with the energy axis (*x-axis*) measured in kiloelectronvolts (keV) and the corresponding X-ray counts (*y-axis*) indicating the intensity of the X-ray signal at each energy level. These spectra provide crucial insights into the elemental composition of the measured mineral.

**Simulated Data** Synthetic spectra are generated using a proprietary algorithm that simulates XRF spectra for user-defined mineral compositions<sup>1</sup>. This algorithm produces synthetic spectra for mineral classes that match the technical format of real data (4096 channels, 0-50 keV range) for direct interoperability. Initial datasets exclude noise and distortions to establish ideal performance benchmarks. Subsequently, real-world distortions — such as Poisson noise, calibration shifts, and Bragg-reflections — are systematically added to the synthetic data. These enhanced synthetic datasets serve as the primary training data, allowing us to evaluate and improve model robustness under conditions that closely mimic real measurements.

**Real Data** Measurements of real-world XRF spectra were conducted under ideal laboratory conditions using the M4 Tornado<sup>2</sup> XRF spectrometer. In the course of the thesis, multiple mineral samples with varying compositions and spectral complexities will be processed, including lithium-bearing minerals, such as granite, spodumene, quartz or pegmatite. The collected data is stored in proprietary .bcf files and structured as spatial maps. Each map represents an image in which each pixel corresponds to an XRF spectrum.

The dataset consists of multiple .bcf measurement files collected across different projects, such as lithium and gold exploration. The maps contained in these files vary not only in image resolution - ranging approximately from  $70 \times 50$  to  $500 \times 500$  pixels - but also in spatial resolution, as the pixel size is determined by the measurement step size (e.g., 30  $\mu\text{m}$ , 100  $\mu\text{m}$ , etc.). As each pixel corresponds to an XRF spectrum, the total number of spectra available for evaluation depends on the specific measurement files used. Initial estimates suggest that across these .bcf files, the dataset contains on the order of millions of spectra, though the exact number depends on the selected samples and their spatial coverage. For example, the initial dataset has a resolution of (70, 50), though subsequent samples vary due to differences in pixel spacing and acquisition settings. An exemplary mineral sample exhibits a resolution of (458, 499), as illustrated in Figure 1 (left).

---

<sup>1</sup>Algorithm provided by Bruker Nano GmbH, developed by Timo Wolff, Fabian Nitsche and Dmitrijs Docenko

<sup>2</sup><https://www.bruker.com/products/m4-tornado-plus>

Each pixel spectrum contains 4096 energy channels, encoding elemental composition information for the corresponding spatial position. The x/y coordinates in these maps represent the physical position ( $\mu\text{m}$ ) of the measurement point on the sample surface where each pixel stores a corresponding spectrum. This enables reconstruction of mineral distribution maps. These spectra serve as benchmarks for evaluating models trained on synthetic datasets. An example of a spatial map with the corresponding spectra of lithium-bearing pegmatite is shown in Figure 1. The images are loaded and displayed using the Hyperspy<sup>3</sup> library.

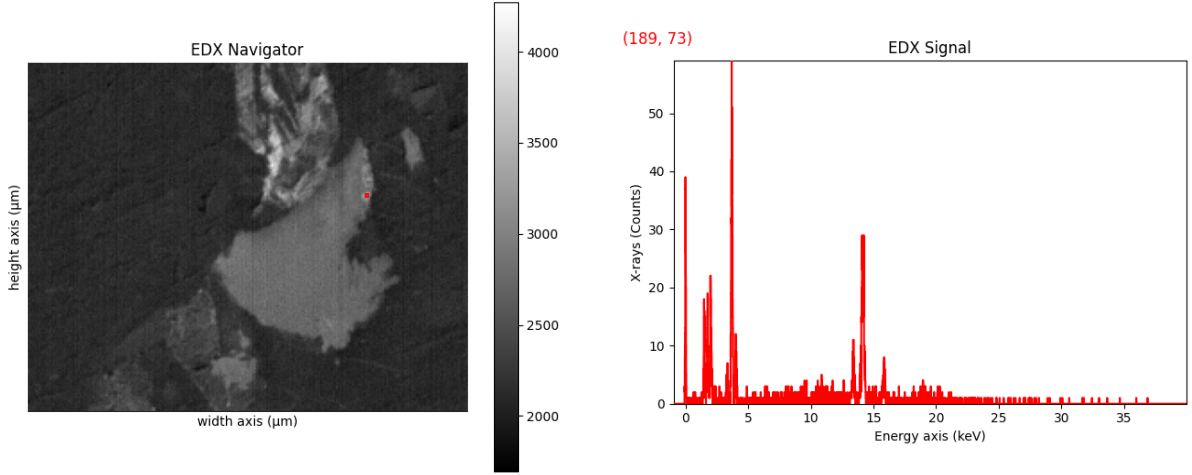


Figure 1: Example dataset of real measured XRF data. On the left, a spatial XRF map of a lithium-bearing pegmatite mineral sample is shown. In this map, each pixel corresponds to an XRF spectrum. The red-marked pixel indicates the location for which the corresponding spectrum is displayed on the right. On the right, the XRF spectrum of that red-marked pixel is presented, revealing characteristic emission peaks that encode the elemental composition at that spatial position.

**Goals** The core objectives of this thesis are:

- Achieving high classification and segmentation accuracy using CNN-based models with attention mechanisms on XRF spectral data.
- Assessing the robustness of the models when faced with noisy, incomplete or out-of-distribution data.
- Introducing real-world conditions into synthetic spectra to evaluate and enhance model robustness and generalizability.
- Comparing model performance on ideal versus noise-rich datasets to assess improvements in stability and practical applicability.

---

<sup>3</sup><https://hyperspy.org>

### 3 Related Work

Recent advancements in deep learning, particularly convolutional neural networks (CNNs), have shown promise in processing spectral data across domains like remote sensing and material science [11, 35]. CNNs excel at extracting hierarchical features from raw inputs, making them effective for high-dimensional datasets [5]. In hyperspectral imaging, CNNs have significantly improved spectral classification accuracy and robustness [30].

Recent advancements in deep learning for spectral data analysis include novel approaches that reduce reliance on traditional preprocessing steps. The DeepSpectra model, introduced by Zhang et al. [35], leverages a convolutional neural network architecture enhanced with the Inception module to perform end-to-end quantitative spectral analysis. Unlike conventional calibration methods, DeepSpectra directly processes raw spectral data, bypassing the need for preprocessing or dimensionality reduction. Yuang et al. [32] proposed a hybrid deep learning model that combines a one-dimensional convolutional neural network (1D-CNN) for feature extraction with an attention-enhanced bidirectional gated recurrent unit (Bi-GRU) for sequential modeling. This architecture effectively captures both local spectral features and long-range dependencies, significantly improving prediction accuracy for moisture content in sand gravel based on near-infrared (NIR) spectroscopy.

In the realm of XRF data, Preisler et al. [22] introduced a deep learning algorithm trained on synthetic datasets for the analysis of macro-XRF data. Applying a CNN to a painting by Raphael, their approach achieved superior accuracy in determining fluorescence line intensities and eliminated artifacts commonly observed in elemental maps generated through traditional deconvolution methods. Similarly, Yang et al. [31] introduced a hybrid framework (CBLA-Net) that combines fractional discrete wavelet transforms (FDWT) for noise reduction with a convolutional neural network with a Squeeze-and-Excitation (SE) attention mechanism for feature extraction, followed by a BiLSTM layer. This demonstrates the potential of CNN-attention models for XRF spectral classification, as attention modules can enhance feature learning by dynamically weighting spectral contributions. Dirks et al. [8] applied autoencoder-based methods for XRF spectral analysis, demonstrating the utility of machine learning in enhancing data interpretation and parameter extraction.

Beyond classification, deep learning has also played a pivotal role in mineral segmentation using XRF and multimodal data fusion. Tang et al. [24] integrated micro-XRF with micro-CT (micro-computed tomography) to enable three-dimensional mineral identification, employing an EfficientU-Net deep learning model to segment ore samples with improved precision. Since classification is often a prerequisite for segmentation, CNN-based spectral classification can provide enhanced feature representations for subsequent mineral mapping. Fang et al. [10] introduced two specialized deep learning architectures for intelligent mineral sorting using X-ray absorption spectroscopy, including a Transformer-LSTM model that effectively combines self-attention with sequential modeling to improve adaptability. Their RNN-overhead-xgboost approach further enhanced fine mineral sorting accuracy, achieving a precision of 99.21%.

Despite these advances, challenges persist. Training CNNs often requires large datasets, prompting reliance on synthetic spectra, which struggle with real-world domain discrepancies. While previous works have independently employed CNNs, attention mechanisms, and sequence models (LSTM, BiGRU, Transformers), this thesis proposes a hybrid framework that first applies CNN models for XRF spectral classification, followed by segmentation using the extracted spectral features. By integrating preprocessing strategies and attention-enhanced architectures, this method aims to enhance mineral segmentation performance under real-world conditions while maintaining computational efficiency.

## 4 Methodology

This thesis proposes the development of a novel machine learning framework for multi-class classification of XRF spectra and mineral segmentation, where each spectrum is assigned to one of several distinct mineral classes. The methodology progresses hierarchically: first classifying individual spectra by mineral type, then generating spatial mineral maps from pixel-wise classifications.

**Preprocessing Stage** Effective preprocessing is a crucial step to improve the quality of XRF spectral data before it is fed into the model. Real XRF spectra often contain noise and redundant information, which can degrade model performance. Synthetic and real data will be processed differently to address their distinct characteristics. The synthetic data starts noise-free and will be augmented with realistic distortions, such as Poisson noise and calibration errors, to mimic real-world conditions. Real data, on the other hand, is collected under ideal lab conditions. This data is preprocessed to reduce inherent noise and standardize inputs. To harmonize synthetic and real spectral data, preprocessing steps include:

- **Baseline Correction and Normalization:** Remove background noise from real data and standardize intensity ranges in both datasets.
- **Targeted Noise Augmentation:** Applied only to synthetic data in order to bridge the gap between idealized simulations and real-world conditions [22, 25].
- **Dimensionality Reduction:** Principal Component Analysis (PCA) will serve as the primary method to retain discriminative spectral features while reducing noise [15]. If PCA underperforms, Variational Autoencoders (VAEs) will be explored to generate denoised spectral reconstructions or latent representations for enhanced feature extraction [3, 9].

The model will be evaluated on both raw and preprocessed spectral inputs to determine the impact of these preprocessing steps on overall performance.

**Classification Stage** This stage addresses a multi-class classification task, where each spectrum is assigned to one mineral class, such as quartz, pegmatite or spodumene. A CNN model will assign each spectrum to a specific mineral class based on its characteristic peaks. CNNs are well-suited for this task due to their ability to learn hierarchical features and identify local spectral patterns critical for distinguishing minerals [6, 13].

The classification methodology will be approached stepwise to balance performance improvements with practical considerations such as dataset size and computational efficiency. The following steps outline the incremental development process:

1. **Baseline CNN model:** Initially, a 1D-CNN architecture will be developed to classify individual spectra into mineral classes. This serves as the control model to establish first benchmarks.
2. **Attention Mechanisms:** To improve the model’s ability to focus on critical spectral regions like elemental peaks, lightweight self-attention modules will be integrated. These mechanisms will dynamically weigh the importance of different ranges based on their relevance to mineral discrimination [14, 34].
3. **Hybrid CNN-BiLSTM Model:** A hybrid architecture combining a CNN and Bidirectional Long Short-Term Memory layers will be tested only if the previous CNN model struggles to capture long-range dependencies in spectral sequences [1, 23, 33, 36].
4. **Validation:** Use stratified splits to evaluate performance, ensuring balanced class representation in training, validation and testing sets.

**Segmentation Stage** Once spectra are accurately classified, the next step is to segment the measured sample based on mineral composition. This segmentation assigns each pixel in the spatial map to a mineral type, providing a visual and quantitative breakdown of the sample’s composition.

1. **Direct Pixel-to-Mineral Mapping:** Assign each pixel to a mineral class directly using the classification results.
2. **Post-Processing:** Apply morphological smoothing, for example erosion or dilation, to the 2D mineral map to reduce isolated misclassifications and refine segmentation [19].
3. **U-Net/DeepLabV3+:** Omitted, unless spatial context, such as mineral zoning patterns, proves critical. Initial segmentation will rely on per-pixel classification [20, 27].

The segmentation results will be used to quantify the percentage of each mineral type present in the sample. This final step enables mineral-specific insights and validates the practical utility of the hierarchical approach [7, 12, 18].

## 5 Evaluation Metrics and Performance Assessment

To ensure a comprehensive evaluation of the proposed deep learning framework, multiple performance metrics will be employed to assess the model's effectiveness across classification and segmentation tasks. The key evaluation metrics include:

- **Classification Accuracy:** Proportion of correctly classified spectra compared to the total number of samples.
- **Segmentation Accuracy:** Pixel-wise accuracy in identifying the correct mineral types.
- **Robustness to Noise:** Performance on datasets with varying levels of synthetic noise.
- **Out-of-Distribution (OOD) Detection:** Ability to identify spectra that differ significantly from the training data.
- **Computational Efficiency:** Evaluation of the model's runtime and resource requirements.

## References

- [1] S. Aburass, O. Dorgham, and J. A. Shaqsi. A hybrid machine learning model for classifying gene mutations in cancer using lstm, bilstm, cnn, gru, and glove. *Systems and Soft Computing*, 6:200110, 2024.
- [2] I. Allegretta, B. Marangoni, P. Manzari, C. Porfido, R. Terzano, O. De Pascale, and G. S. Senesi. Macro-classification of meteorites by portable energy dispersive x-ray fluorescence spectroscopy (ped-xrf), principal component analysis (pca) and machine learning algorithms. *Talanta*, 212:120785, 2020.
- [3] V. Andric, G. Kvascev, M. Cvetanovic, S. Stojanovic, N. Bacanin, and M. Gajic-Kvascev. Deep learning assisted xrf spectra classification. *Scientific Reports*, 14(1), Feb. 2024.
- [4] V. Balaram and S. S. Sawant. Indicator minerals, pathfinder elements, and portable analytical instruments in mineral exploration studies. *Minerals*, 12(4), 2022.
- [5] Bingjie, Xu, Y. Wu, P. Hao, M. Vermeulen, A. McGeachy, K. Smith, K. Eremin, G. Rayner, G. Verri, F. Willomitzer, M. Alfeld, J. Tumblin, A. Katsaggelos, and M. Walton. Can deep learning assist automatic identification of layered pigments from xrf data?, 2022.
- [6] Z. Chang, Q. Zhang, Y. Li, X. Xin, R. Gao, Y. Teng, L. Rao, and M. Sun. Identification method for xrf spectral analysis based on an aga-bp-attention neural network. *Electronics*, 13(3), 2024.
- [7] G. Csurka, R. Volpi, and B. Chidlovskii. Semantic image segmentation: Two decades of research, 2023.
- [8] M. Dirks and D. Poole. Auto-encoder neural network incorporating x-ray fluorescence fundamental parameters with machine learning. *X-Ray Spectrometry*, 52(3):142–150, Mar. 2023.
- [9] C. Doersch. Tutorial on variational autoencoders, 2021.
- [10] Z. Fang, S. Song, H. Wang, H. Yan, M. Lu, S. Chen, S. Li, and W. Liang. Mineral classification with x-ray absorption spectroscopy: A deep learning-based approach. *Minerals Engineering*, 217:108964, 2024.
- [11] M. Grossutti, J. D’Amico, J. Quintal, H. MacFarlane, A. Quirk, and J. R. Dutcher. Deep learning and infrared spectroscopy: Representation learning with a  $\beta$ -variational autoencoder. *The Journal of Physical Chemistry Letters*, 13(25):5787–5793, June 2022.
- [12] C. Guo, M. Szemenyei, Y. Yi, W. Wang, B. Chen, and C. Fan. Sa-unet: Spatial attention u-net for retinal vessel segmentation. In *2020 25th International Conference on Pattern Recognition (ICPR)*, pages 1236–1242, 2021.

- [13] L. He, Y. Zhou, and C. Zhang. Application of target detection based on deep learning in intelligent mineral identification. *Minerals*, 14(9), 2024.
- [14] G. W. Humphreys and J. Sui. Attentional control and the self: The self-attention network (san). *Cognitive Neuroscience*, 7(1–4):5–17, May 2015.
- [15] I. T. Jolliffe. Principal component analysis and factor analysis. *Principal component analysis*, pages 150–166, 2002.
- [16] C. Jones, N. S. Daly, C. Higgitt, and M. R. D. Rodrigues. Neural network-based classification of x-ray fluorescence spectra of artists’ pigments: an approach leveraging a synthetic dataset created using the fundamental parameters method. *Heritage Science*, 10(1), June 2022.
- [17] N. Katsuta, A. Umemura, S. Naito, Y. Masuki, Y. Itayama, M. Niwa, S. iti Sirono, H. Yoshida, and S. ichi Kawakami. Heterogeneity effects in micro-beam xrf scanning spectroscopy of binary powdered mixtures and lake sediments. *Spectrochimica Acta Part B: Atomic Spectroscopy*, 210:106817, 2023.
- [18] F. Krikid, H. Rositi, and A. Vacavant. State-of-the-art deep learning methods for microscopic image segmentation: Applications to cells, nuclei, and tissues. *Journal of Imaging*, 10(12), 2024.
- [19] P. Maragos and L. Pessoa. Morphological filtering for image enhancement and detection. *Analysis*, 13, 01 1999.
- [20] A. Pal, H. M. Rai, M. B. H. Frej, and A. Razaque. Advanced segmentation of gastrointestinal (gi) cancer disease using a novel u-masknet model. *Life*, 14(11), 2024.
- [21] P. J. Potts. X-ray fluorescence analysis. In *Soil and Environmental Analysis*, pages 260–317. CRC Press, 2003.
- [22] Z. Preisler, R. Andolina, A. Busacca, C. Caliri, C. Miliani, and F. P. Romano. Deep learning for enhanced spectral analysis of ma-xrf datasets of paintings. *Science Advances*, 10(39):eadp6234, 2024.
- [23] R. Shelishiyah, D. B. Thiyam, M. J. Margaret, and N. M. M. Banu. A hybrid cnn model for classification of motor tasks obtained from hybrid bci system. *Scientific Reports*, 15(1), Jan. 2025.
- [24] K. Tang, Y. D. Wang, P. Mostaghimi, M. Knackstedt, C. Hargrave, and R. T. Armstrong. Deep convolutional neural network for 3d mineral identification and liberation analysis. *Minerals Engineering*, 183:107592, 2022.
- [25] A. Ulianov, O. Müntener, and U. Schaltegger. The icpms signal as a poisson process: a review of basic concepts. *J. Anal. At. Spectrom.*, 30:1297–1321, 2015.

- [26] C. von Scheffer, F. De Vleeschouwer, G. Le Roux, and I. Unkel. Mineral dust and lead deposition from land use and metallurgy in a 4800-year-old peat record from the central alps (tyrol, austria). *Quaternary International*, 700-701:68–79, 2024. Shifting the Focus: Mountains as Central Places in prehistoric and early historic times.
- [27] C. Wang, P. Du, H. Wu, J. Li, C. Zhao, and H. Zhu. A cucumber leaf disease severity classification method based on the fusion of deeplabv3+ and u-net. *Computers and Electronics in Agriculture*, 189:106373, 2021.
- [28] M. Warlo, G. Bark, C. Wanhainen, I. McElroy, A. Björling, and U. Johansson. Extreme-resolution synchrotron x-ray fluorescence mapping of ore samples. *Ore Geology Reviews*, 140:104620, 2022.
- [29] J.-L. Xu, C. Riccioli, A. Herrero-Langreo, and A. A. Gowen. Deep learning classifiers for near infrared spectral imaging: a tutorial. *Journal of Spectral Imaging*, 9, 2020.
- [30] J. Yang, J. Xu, X. Zhang, C. Wu, T. Lin, and Y. Ying. Deep learning for vibrational spectral analysis: Recent progress and a practical guide. *Analytica Chimica Acta*, 1081:6–17, June 2019.
- [31] W. Yang, F. Li, Q. Zhang, and S. Lyu. An integrated cbia-net with fractional discrete wavelet transform and frequency-based cars to predict heavy metal elements by xrf. *Analytica Chimica Acta*, 1323:343073, 2024.
- [32] Q. Yuan, J. Wang, M. Zheng, and X. Wang. Hybrid 1d-cnn and attention-based bi-gru neural networks for predicting moisture content of sand gravel using nir spectroscopy. *Construction and Building Materials*, 350:128799, 2022.
- [33] C. Zhang, H. Ma, L. Hua, W. Sun, M. S. Nazir, and T. Peng. An evolutionary deep learning model based on tvfemd, improved sine cosine algorithm, cnn and bilstm for wind speed prediction. *Energy*, 254:124250, 2022.
- [34] Q. Zhang, F. Li, and W. Yang. A deep spectral prediction network to quantitatively determine heavy metal elements in soil by x-ray fluorescence. *Journal of Analytical Atomic Spectrometry*, 39(2):478–490, 2024.
- [35] X. Zhang, T. Lin, J. Xu, X. Luo, and Y. Ying. Deepspectra: An end-to-end deep learning approach for quantitative spectral analysis. *Analytica Chimica Acta*, 1058:48–57, Jan. 2019.
- [36] C. Zhao, X. Huang, Y. Li, and M. Yousaf Iqbal. A double-channel hybrid deep neural network based on cnn and bilstm for remaining useful life prediction. *Sensors*, 20(24), 2020.