Enhancing Air Quality Forecasting Through Deep Learning and Continuous Wavelet Transform

Pietro Manganelli Conforti[1][0000−0002−4873−1369], Andrea Fanti[1][0009−0003−0764−3965], Pietro Nardelli[1][0000−0002−9093−1532], and Paolo Russo[1][0000−0002−1886−3491]

DIAG Department, Sapienza University of Rome, Via Ariosto 25, 00185 Roma, Italy
{manganelliconforti,fanti,russo}@diag.uniroma1.it
pietro.nardelli@outlook.com

Abstract. Air quality forecasting plays a crucial role in environmental management and public health. In this paper, we propose a novel approach that combines deep learning techniques with the Continuous Wavelet Transform (CWT) for air quality forecasting based on sensor data. The proposed methodology is agnostic to the target pollutant and can be applied to estimate any available pollutant without loss of generality. The pipeline consists of two main steps: the generation of stacked samples from raw sensor signals using CWT, and the prediction through a custom deep neural network based on the ResNet18 architecture.
We compare our approach with traditional one-dimensional signal processing models. The results show that our 2D pipeline, employing the Morlet mother wavelet, outperforms the baselines significantly. The localized time-frequency representations obtained through CWT highlight hidden dynamics and relationships within the parameter behavior and external factors, leading to more accurate predictions. Overall, our approach demonstrates the potential to advance air quality forecasting and environmental management for healthier living environments worldwide.

Keywords: Air Quality Forecasting · Deep Learning · Continuous Wavelet Transform (CWT).

1 Introduction

The quality of the air we breathe is a critical aspect of our daily lives, directly impacting human health and the overall well-being of our environment. With the continuous rise in urbanization and industrialization, several contaminants have been introduced into the atmosphere that cause adverse changes either directly, by releasing toxic or harmful chemicals into the air, or indirectly, through the disruption of the delicate natural equilibrium, reflected in the composition of the air. Monitoring, forecasting, and preservation of air quality are of uttermost importance, in order to manage the chemical human footprint and adapt to the constant development of technologies. Moreover, accurate analysis through new technologies offers the potential to reduce costs associated with air sensing

and monitoring. By leveraging advanced processing techniques and predictive models, more precise forecasts can be generated, enabling optimized deployment of sensing resources and targeted monitoring efforts.

In recent years, advancements in deep learning techniques combined with signal processing methods have emerged as powerful processing tools for addressing these challenges. In fact, deep neural networks demonstrated exceptional performance in capturing intricate interactions among numerous environmental scenarios, including meteorological conditions predictions [17], pollutant emissions monitoring [5], and environment feature analysis [10]. In addition, transform techniques such as the short-time Fourier transform and the wavelet transform have demonstrated substantial advantages for extracting valuable information about the time-varying frequency properties of one-dimensional signals in several fields of application, such as speech recognition [6], anomaly detection [14], and biomedical signal processing [11]. By transforming these signals into two-dimensional, time-frequency representations, it is possible to conduct a more thorough analysis of the underlying data, exploiting two-dimensional, state-of-the-art deep models with their accumulated knowledge via transfer learning [18].

In this research, we specifically focus on the Continuous Wavelet Transform (CWT) and explore the impact of different mother wavelets on enhancing the time-frequency analysis of air chemical concentration and other environmental measurements. Moreover, we investigate the potential synergies between deep learning techniques and transform-based methods in order to improve air quality forecasting. Therefore, the aim of this research is threefold, with a specific focus on the subsequent key areas:

- Utilizing a deep learning pipeline along with a proposed convolutional compression module to process the transformed data. We stack multiple time-frequency representations and fed them as input to the network, enabling the estimation of an air pollutant, (e.g., CO), starting from the other measured elements (e.g., C6H6, CO, NO2, NOx, NMHC, O3, AH, RH, T).
- Exploring the methodology for generating informative two-dimensional representations from several one-dimensional data segments and the effects of different mother wavelets employed by the CWT to enhance the time-frequency analysis. By comparing the images produced by different wavelets, we aim to identify the one that captures a broader range of temporal features and thus improving the final accuracy of air quality predictions.
- We demonstrate the possibility of accurately predicting one pollutant starting from a subset of the available measurements. This could enable the production of cheaper air sensors that can measure fewer chemical concentrations to obtain the same final results

The rest of the paper is organized as follows: firstly Sec.2 will provide a description of the current state of the art, discussing the related works concerning air quality analysis and prediction and the use of mathematical transforms to analyze one-dimensional signals. Thereafter, in Sec. 3 we present the work methodology employed and the full model pipeline. Following will be described the experimental setup in Sec.4, where are shown the hyperparameters employed

in the 1D and 2D cases, the dataset composition, and the accuracy measures used. Finally, in Sec.5 a comparison table of all the research findings is provided. The last section Sec.6 will conclude the paper with future works and additional considerations.

2 Related Work

In recent years, deep learning has witnessed remarkable success across various domains. However, the performance of these models heavily relies on the quality of data representation, and this is where the power of transformations comes into play. Akansu et al.[1] provide a comprehensive assessment of the diverse and evolving applications of wavelet transforms across various fields. In order to identify complex mixtures in Raman spectroscopy, [12] propose a novel scheme based on a CWT and a deep neural network. The multi-label network model is used for classifying complex mixtures with improved performance thanks to CWT that decomposes the desired molecular information and noise from the Raman spectrum. Another approach [13] utilizes the Short-Time-Fourier transform (STFT) and a new convolutional neural network for lung cancer diagnosis from spectrochemical analysis. Similarly but in a different domain, in [16] authors utilize a deep neural network to detect audio operations using two types of transform techniques: STFT and Modified Discrete Transform. They improve audio authentication for forensics, successfully recognizing spoofed voices with high accuracy.

Recently, the field of air quality assessment has witnessed significant advancements through the application of machine learning techniques. We highlight the key works and contributions in this domain, outlining the methodologies and outcomes of these studies. Liu et al. [7] focused on predicting both the Air Quality Index and air pollutant concentrations through the application of various machine learning algorithms such as random forest regression and support vector regression. The work of Ly et al. [8] investigates the predictions of two important compounds in air pollution: NO_2 and CO. The authors address the regression problem by combining fuzzy logic with neural networks and a metaheuristic optimization chosen between simulated annealing and particle swarm optimization algorithms. Tackling the problem of noisy data, Sabzekar et al. [15] increase the robustness of the support vector regression, by converting constraints into fuzzy inequalities. In the work of Ma et al. [9] recurrent neural networks (RNNs) are exploited for time-series regression to tackle air quality estimation. They introduced a new RNN family, called Particle Filter RNNs, that is able to model the uncertainty by approximating a set of weighted particles, updated through to a particle filtering algorithm. Finally, Chen et al. [2] exploit a Gaussian process for multi-output prediction proposing a framework with a novel multivariate Student-t process regression model.

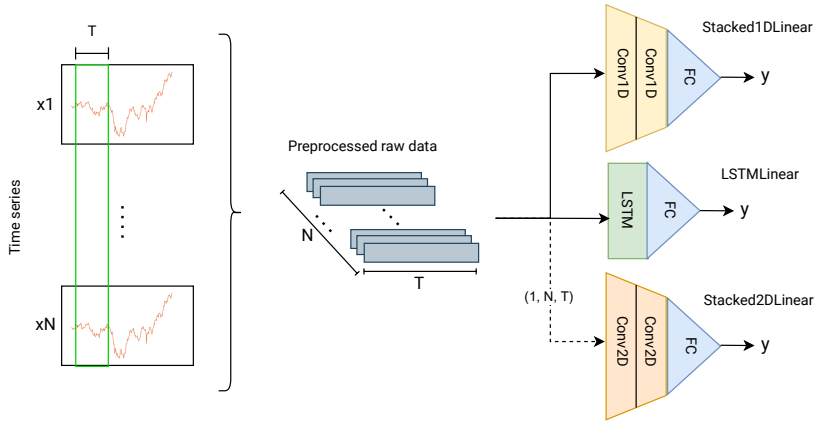

Fig. 1. Schema of the three baseline models, Stacked1DLinear, LSTMLinear, and Stacked2DLinear, employed in the 1D analysis.

3 Methodology and data

The task we address through our pipeline is the accurate estimation of the mean future values of a single pollutant, accordingly to specific time windows. More in detail, given hourly measurements of the input parameters over a certain week, our task is to predict the average concentration of a target pollutant on the day which follows that week. Starting from the list of chemical elements inside the analyzed dataset, we have chosen CO to be predicted from a selection of these elements. However, the proposed methodology is agnostic with respect to the target pollutant and thus can be applied to estimate any of the other available pollutants without loss of generality. We compared our approach (which is described in 2D analysis subsection) with some traditional approaches (1D analysis subsection) which work directly on the one-dimensional signals. A detailed comparison in Section 4 will show the performance differences between the different techniques.

3.1 1D analysis

In the field of air quality estimation, the analysis of raw time-series data has been widely employed by researchers as a standard approach. For this reason, we outline the development of three baseline models for air quality forecasting, emphasizing the usage of raw time-series data. The visual representation of the three baseline models can be seen in figure 1.

As a first step, the raw data undergoes a preprocessing stage, detailed in section 4.1. The resulting output naturally forms a time series due to the sequential nature of the data. To effectively model the time-dependent patterns, we adopt a first baseline approach, called Stacked1DLinear, that utilizes a sequence of one-dimensional convolutional layers. Another model suitable for this

type of data is the Long Short-Term Memory (LSTM) architecture, which is able to capture features from the sequential information contained in the data. This architecture (LSTMLinear) utilizes relevant information from earlier time steps, thereby enhancing its capacity to make accurate predictions for air quality values. In our third baseline approach, called Stacked2DLinear, we explored a different strategy by utilizing 2D convolutions to process the time series data. This approach benefits from the ability of convolutional layers to detect spatial patterns across different parameters and temporal patterns along the time axis. In our pursuit of a robust and comprehensive baseline comparison, we intentionally adopted the 2D convolution approach to closely resemble the Conv2D part of our proposed network architecture. This will highlight the contribution of the 2D transformation which we propose in our 2D pipeline.

3.2 2D data generation

To generate 2D images from the 1D time series, we use the Continuous Wavelet Transform (CWT). It is a mathematical technique used for analyzing and processing signals or time series data in various scientific and engineering applications, providing a multidimensional analysis of both time and frequency components. This can be fundamental to highlight specific spectral features of air chemical concentrations and meteorological parameters hidden in the input data, showing the time-frequency representations of the signals dynamics and their relationship. Furthermore, by employing these two-dimensional representations, the CWT can be synergistically integrated with a 2D deep neural network. This integration allows us to harness both the informative time-frequency representation and the robust generalization capabilities of visual features present in pre-trained deep neural networks.

Given a positive scale parameter a, a translation value b, a mother wavelet function $\psi(t)$, and an input signal $x(t)$, the mathematical formulation [4] used by this technique is defined as:

$$X_\psi(a,b) = \langle x(t), \psi_{\mathrm{ab}}(t) \rangle = \int_{-\infty}^{+\infty} x(t)\psi_{\mathrm{ab}}(t)dt$$

Where ψ_{ab} is the complex conjugate of the mother wavelet ψ and $X_\psi(a,b)$ is the time-frequency representation of the signal. Usually, the wavelets are brief, low-energy oscillatory signals. The mother wavelet function produces daughter wavelets based on the a and b parameters, which can be exploited to examine various frequency components at different levels [19]. Depending on input data, mother wavelets may create quite diverse time-frequency representations, revealing different trends. To analyze this behavior, in Section 5 we compare the different performances obtained by employing three popular mother wavelets: `morlet`, `morlet2`, and `ricker`. We exploit the CWT to produce a so-called *scalogram* of a signal, i.e. a heat map where the axes are time t and frequency f, and the color represents the magnitude or phase of the CWT $X_\psi(f,t)$. In our case, we chose to only use the magnitude, to allow us to generate a single

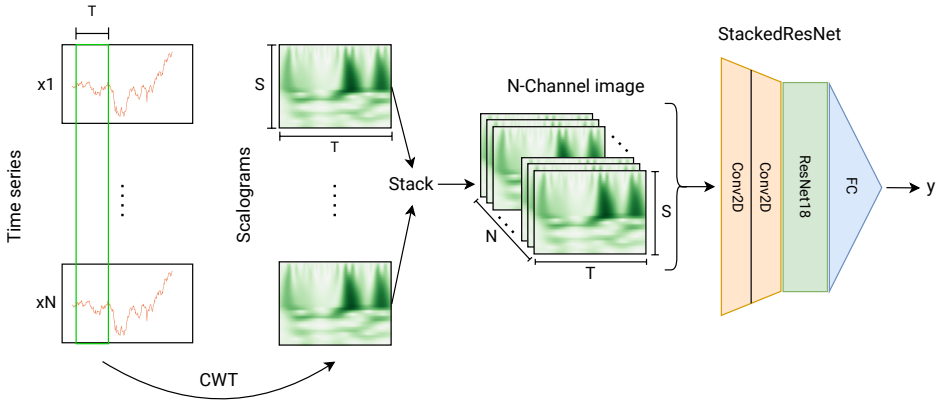

Fig. 2. The full 2D analysis pipeline. T is the window size, S is the number of scales, and N is the number of input signals.

gray-scale image from a single signal. The domain of the f axis is also referred to as *scales*, and its size as the *number of scales*. To increase the number of samples from a single time series, we split the time series into signal windows of fixed size, which are allowed to overlap for a fixed number of time steps in order to further increase the dataset size. Each window is then passed independently through the CWT, and the resulting magnitude is then used to produce the gray-scale image associated with that window by plotting its scalogram. Note that the resulting images are composed of a single channel. This means that, given a window size of N and the number of scales M, the resulting image will have dimensions (N, M). To split the time series into signal windows we use a sliding window, meaning that subsequent windows are partially overlapping. We call the temporal distance between the beginning of a signal window and the beginning of the next one the *stride* of the sliding window. More formally, let us assume the time series has length $\hat{T}$, the window size is T and the stride is S. Then, if $\hat{x}_t$ is the t-th value in the time series, the j-th element of the i-th window x_{ij} is computed as follows: $x_{ij} = \hat{x}_{iS+j}$ and thus the i-th signal window x_i will be $x_i = (\hat{x}_{iS}, \hat{x}_{iS+1}, \ldots, \hat{x}_{iS+N})$. If $\hat{N}$ is not a multiple of N, then the last signal window will be discarded since its size would be smaller than N. Finally, we would like to remark that the previous procedure is applied along the number of pollutant signals employed as input (e.g. if 11 different pollutant measurements are provided as input, for a single signal window 11 different scalograms are produced).

3.3 2D analysis

Convolutional Neural Networks pre-trained on ImageNet provide powerful visual features which can be further specialized on the task of interest; for this reason, their usage is able to increase the accuracy of our regression task. However, those deep models require the input image to be represented as a three-channels image (usually in the RGB format). In our case, for a given signal window we have a number of scalograms equal to the number of pollutants measured and provided as input. For this reason, as a first step we compose each input sample by stacking the images corresponding to the Wavelet transform of the same window of the different input time series. More specifically, given the images I_j corresponding to the scalograms of the magnitudes of the CWTs of the i-th temporal windows of the parameters $j = 1, \ldots, N_j$, each of dimensions (N, M), the corresponding input sample will be a 3D tensor T of dimensions (N_j, N, M), such that $(T_j) = I_j$, or equivalently, a (N, M) image with N_j channels. In order to be able to feed this tensor to our pre-trained ResNet18 model, we provide the tensor as input to a sequence of two convolutional layers, with the last layer producing as output a $(3, N, M)$ tensor which has the correct shape in order to be provided to the backbone model for producing the output. Figure 2 reports the architecture of the proposed 2D pipeline.

4 Experimental setup

4.1 Dataset

The dataset used to assess the validity of our method is the "Air Quality Dataset" provided by De Vito et al. [3]. It contains measurements of air pollutant concentration parameters and environmental parameters on a main road of an Italian city. More specifically, it consists of a collection of time series, one for each air pollutant or meteorological parameter. For some parameters, two types of measurements are available: one from a sensor tower (which we refer to as GT), and the other representing raw readings from cheaper, smaller embedded sensors (which we refer to as PT08); this is because the dataset was originally designed to calibrate the smaller sensor using the measurements from the GT sensors. All time series are aligned, i.e. the t-th samples of all time series correspond to measurements taken at the same real-time. There are 13 parameters in total: 5 from the GT sensors (CO, NMHC, C6H6, NOx, NO2), 5 from the PT08 sensors (CO, NMHC, NOx, NO2, O3), and 3 meteorological measurements (T, AH, RH). The measurements were taken hourly, aside from the first few samples. Some measurements are missing in the original dataset, and are marked with a -200 value. We split the dataset into training and test sets with a ratio of 80/20. We further split the former into training and validations sets with a ratio of 90/10. Thus, after splitting the time series into windows, we obtained 865 samples divided into 623 training samples, 69 validation samples and 173 test samples.

4.2 Implementation Details

Stacked1DLinear (6.2M parameters) and Stacked2DLinear (11M parameters) consist of 2 blocks made of convolutional layers, batch normalizations, and ReLU activations, followed by 2 fully connected layers. The blocks in the Stacked2DLinear are exactly the same as the ones used in the StackedResnet in order to ensure a fair comparison. The LSTMLinear model (13.7M) has 2 LSTM layers as the main architecture with 512 hidden sizes. The LSTM output is fed as input into two fully connected linear layers. StackedResNet (11M trainable parameters) is composed of two 3x3 Conv2D blocks, a ResNet18 pre-trained model used as backbone and a Fully Connected (FC) block. The last layer of the original ResNet18 model is substituted with a new FC layer to allow for the correct output size. (a single floating point number in our case). The first Conv2D layer has 6 output channels, with the second having 3 output channels; both Conv2D layers have a 3x3 kernel with stride 1 and padding 1. We train all the learnable parameters of the resulting architecture, including those of the pre-trained on ImageNet ResNet18 block. To generate the dataset used in all experiments, we used a sliding window size of 168 data points (which correspond to the hours in a week) with a stride of 6. To plot each scalogram, we used $\{1, \ldots, 126\}$ as the scales, resulting in a number of scales of 126. For model optimization, we used the Adam optimizer with a batch size of 8 and an initial learning rate of 1e-5. All models were optimized for 100 epochs, with the weights of the 2 convolutional layers, the ResNet18 and the FC block jointly training (with the ResNet18 weights pre-trained on ImageNet, while the other layers were trained from scratch). Additional implementation details and source code can be found in the GitHub repository [1].

4.3 Performance measures

Two performance measures are employed to evaluate the considered models in the chosen task: mean relative error (MRE), and mean squared error (MSE). More specifically, we computed the mean relative error $\mathrm{RE}(y, \hat{y})$ and mean squared error $\mathrm{MSE}(y, \hat{y})$ over predictions $y = (y_1, \ldots, y_N)$ and targets $\hat{y} = (\hat{y}_1, \ldots, \hat{y}_N)$ as:

$$\mathrm{MRE}(y, \hat{y}) = \frac{1}{N} \sum_{i=1}^{N} \frac{|y_i - \hat{y}_i|}{|\hat{y}_i|} \qquad \mathrm{MSE}(y, \hat{y}) = \frac{1}{N} \sum_{i=1}^{N} (y_i - \hat{y}_i)^2$$

The MSE is also used as the loss function to train all models, which is a common choice in Deep Learning when directly predicting real-valued quantities.

5 Results

We experimented with several combinations of input parameters, both including and excluding CO; results are reported in table 1. For the sake of lighter notation,

[1] https://github.com/PietroManganelliConforti/Deep-Learning-and-Wavelet-Transform-for-Air-Quality-Forecasting

Table 1. 2D pipeline results using different combinations of input parameters. The table is split vertically to separate the configurations that included CO among input parameters from those that didn't include it.

Input variables	Test MRE	Train MRE	Test MSE	Train MSE
GT*	**0.0621**	0.0281	**0.0327**	0.0053
PT*	0.0699	0.0327	0.0367	0.0075
PT*, HT	0.0650	0.0323	0.0372	0.0067
GT*, HT	0.0661	0.0315	0.0351	0.0065
GT*, PT*, HT	0.0806	**0.0275**	0.0658	**0.0049**
HT	0.1023	0.0472	0.0764	0.0145
PT	0.0755	0.0319	0.0416	0.0069
GT	**0.0522**	0.0277	**0.0221**	0.0056
PT*,CO(GT)	0.0577	0.0345	0.0259	0.0075
GT*,CO(PT08)	0.0575	0.0305	0.0281	0.0059
CO(GT)	0.0655	**0.0213**	0.0267	**0.0031**
CO(PT08)	0.0770	0.0268	0.0515	0.0045

Table 2. Performances of our 2D pipeline using different mother wavelet functions. All experiments use the GT* combination of input parameters.

Mother wavelet	Test MRE	Train MRE	Test MSE	Train MSE
Morlet	**0.0621**	**0.0281**	**0.0327**	**0.0053**
Morlet2	0.0762	0.0349	0.0405	0.0073
Ricker	0.0792	0.0337	0.0521	0.0070

Table 3. Performances of our 2D pipeline (from table 1) against the three baseline models. All experiments use the GT* combination of input parameters.

Model	Test MRE	Train MRE	Test MSE	Train MSE
LSTMLinear	0.086	0.0549	0.0466	0.0217
Stacked1DLinear	0.121	0.0340	0.1367	0.0070
Stacked2DLinear	0.137	0.0385	0.1411	0.0103
StackedResNet (Morlet)	**0.0621**	**0.0281**	**0.0327**	**0.0053**

we group some parameters and refer to them as follows. GT stands for all parameters measured with the GT sensor, i.e. CO, C6H6, NOx, NO2. PT stands for all parameters measures with the PT08 sensor, i.e. CO, O3, NOx, NO2, NMHC. HT stands for the 3 atmospheric parameters T, AH, and RH. GT* and PT* indicate the same sets as GT and PT, but excluding their respective CO measurements. As expected, most of the combinations which include CO perform better than those that don't; the exceptions are PT, CO(GT) and CO(PT08). Note, however, that two of these combinations actually only include one parameter as input, which is expected to make the 2D pipeline less effective. This is also supported by the GT* and CO(GT) combinations obtaining similar performances, even though the former doesn't contain the target parameter at all. Moreover, since the measurements from the PT08 sensor are significantly less reliable than those from the GT sensor, using only those measurements instead of CO(GT)

is also expected to result in worse performances. These results also confirm that the GT signals are the most informative, as the best-performing combinations with and without CO are GT and GT*, while as expected the meteorological parameters (HT) seem to be the less informative ones. We also compared different mother wavelet functions, namely Morlet, Morlet2 and Ricker. We used GT* as the set of input parameters, since this was the best-performing combination of inputs that didn't include CO as an input. This still allows us to obtain good performances while not requiring any information on the target parameter itself. Results are reported in table 2. It is clear from these outcomes that the Morlet mother wavelet seems to give the best performances, while the other two give similar performances, which further validates the choice of the Morlet wavelet for all other experiments.

Finally, we evaluated the three 1D baseline approaches described above, to compare them against the results obtained with the 2D pipeline. Again, we used GT* as the set of input parameters. The comparison is reported in table 3. In particular, our 2D pipeline with the Morlet mother wavelet significantly outperforms all 3 baselines, which further validates the usefulness of this methodology for this kind of task. The baseline is mostly outperformed even when using the other wavelets, as seen by comparing the results in tables 3, 2. Our method performs better than 2 out of the 3 baselines even when only using the meteorological parameters as inputs, as seen by comparing results in tables 3, 1.

6 Conclusion

In this study, we explored the application of the Continuous Wavelet Transform in combination with different types of deep neural networks for air quality forecasting using sensor data. Through its employment, we obtained localized time-frequency representations of pollutant and environmental values, thereby highlighting concealed dynamics and relationships within the parameter behavior and the other possible external factors. Exploiting the time-frequency features extracted from three different mother wavelets and leveraging the knowledge acquired by a pre-trained deep neural network, our approach surpassed the performance of conventional one-dimensional signal processing methods.

In our future work, we aim to enhance the functionalities of our system by exploring and testing new pipeline architectures. One potential avenue is to apply an ensemble learning approach using multiple networks, which can lead to improved predictive performance and more robust results. In addition to the CWT, we will also explore other time-frequency signal representation methods, such as the Short-Time Fourier Transform. Moreover, we plan to expand the scope of our predictions beyond the average concentration values for the day following the considered time window. Ultimately, the future work outlined will significantly advance our understanding of the proposed task and improve the accuracy and applicability of our prediction system in order to reduce the cost of air quality management and to promote healthier living environments for communities worldwide.

References

1. Akansu, A.N., Serdijn, W.A., Selesnick, I.W.: Emerging applications of wavelets: A review. Physical Communication **3**(1), 1–18 (Mar 2010). https://doi.org/10.1016/j.phycom.2009.07.001, https://linkinghub.elsevier.com/retrieve/pii/S1874490709000482
2. Chen, Z., Wang, B., Gorban, A.N.: Multivariate Gaussian and Student-t process regression for multi-output prediction. Neural Computing and Applications **32**(8), 3005–3028 (Apr 2020). https://doi.org/10.1007/s00521-019-04687-8, https://doi.org/10.1007/s00521-019-04687-8
3. De Vito, S., Massera, E., Piga, M., Martinotto, L., Di Francia, G.: On field calibration of an electronic nose for benzene estimation in an urban pollution monitoring scenario. Sensors and Actuators B: Chemical **129**(2), 750–757 (2008). https://doi.org/https://doi.org/10.1016/j.snb.2007.09.060, https://www.sciencedirect.com/science/article/pii/S0925400507007691
4. Gargour, C., Gabrea, M., Ramachandran, V., Lina, J.M.: A short introduction to wavelets and their applications. IEEE Circuits and Systems Magazine **9**(2), 57–68 (2009). https://doi.org/10.1109/MCAS.2009.932556
5. Huang, L., Liu, S., Yang, Z., Xing, J., Zhang, J., Bian, J., Li, S., Sahu, S.K., Wang, S., Liu, T.Y.: Exploring deep learning for air pollutant emission estimation. Geoscientific Model Development **14**(7), 4641–4654 (Jul 2021). https://doi.org/10.5194/gmd-14-4641-2021, https://gmd.copernicus.org/articles/14/4641/2021/, publisher: Copernicus GmbH
6. Krawczyk, M., Gerkmann, T.: Stft phase reconstruction in voiced speech for an improved single-channel speech enhancement. IEEE/ACM Transactions on Audio, Speech, and Language Processing **22**(12), 1931–1940 (2014). https://doi.org/10.1109/TASLP.2014.2354236
7. Liu, H., Li, Q., Yu, D., Gu, Y.: Air Quality Index and Air Pollutant Concentration Prediction Based on Machine Learning Algorithms. Applied Sciences **9**(19), 4069 (Sep 2019). https://doi.org/10.3390/app9194069, https://www.mdpi.com/2076-3417/9/19/4069
8. Ly, H.B., Le, L.M., Phi, L.V., Phan, V.H., Tran, V.Q., Pham, B.T., Le, T.T., Derrible, S.: Development of an AI Model to Measure Traffic Air Pollution from Multisensor and Weather Data. Sensors **19**(22), 4941 (Nov 2019). https://doi.org/10.3390/s19224941, https://www.mdpi.com/1424-8220/19/22/4941
9. Ma, X., Karkus, P., Hsu, D., Lee, W.S.: Particle Filter Recurrent Neural Networks. Proceedings of the AAAI Conference on Artificial Intelligence **34**(04), 5101–5108 (Apr 2020). https://doi.org/10.1609/aaai.v34i04.5952, https://ojs.aaai.org/index.php/AAAI/article/view/5952
10. Ma, Z., Mei, G.: Deep learning for geological hazards analysis: Data, models, applications, and opportunities. Earth-Science Reviews **223**, 103858 (Dec 2021). https://doi.org/10.1016/j.earscirev.2021.103858, https://www.sciencedirect.com/science/article/pii/S0012825221003597
11. Manganelli Conforti, P., D'Acunto, M., Russo, P.: Deep learning for chondrogenic tumor classification through wavelet transform of raman spectra. Sensors **22**(19) (2022). https://doi.org/10.3390/s22197492, https://www.mdpi.com/1424-8220/22/19/7492
12. Pan, L., Pipitsunthonsan, P., Daengngam, C., Chongcheawchamnan, M.: Identification of complex mixtures for Raman spectroscopy using a novel scheme based on a new multi-label deep neural network (Oct 2020)

13. Qi, Y., Yang, L., Liu, B., Liu, L., Liu, Y., Zheng, Q., Liu, D., Luo, J.: Accurate diagnosis of lung tissues for 2D Raman spectrogram by deep learning based on short-time Fourier transform. Analytica Chimica Acta **1179**, 338821 (Sep 2021). https://doi.org/10.1016/j.aca.2021.338821, https://linkinghub.elsevier.com/retrieve/pii/S0003267021006474
14. Russo, P., Schaerf, M.: Anomaly detection in railway bridges using imaging techniques. Scientific Reports **13**(1), 3916 (2023)
15. Sabzekar, M., Hasheminejad, S.M.H.: Robust regression using support vector regressions. Chaos, Solitons & Fractals **144**, 110738 (Mar 2021). https://doi.org/10.1016/j.chaos.2021.110738, https://www.sciencedirect.com/science/article/pii/S0960077921000916
16. Saleem, S., Dilawari, A., Khan, U.G.: Spoofed voice detection using dense features of stft and mdct spectrograms. In: 2021 International Conference on Artificial Intelligence (ICAI). pp. 56–61 (2021). https://doi.org/10.1109/ICAI52203.2021.9445259
17. Salman, A.G., Kanigoro, B., Heryadi, Y.: Weather forecasting using deep learning techniques. In: 2015 International Conference on Advanced Computer Science and Information Systems (ICACSIS). pp. 281–285 (Oct 2015). https://doi.org/10.1109/ICACSIS.2015.7415154
18. Tan, C., Sun, F., Kong, T., Zhang, W., Yang, C., Liu, C.: A survey on deep transfer learning. In: Artificial Neural Networks and Machine Learning–ICANN 2018: 27th International Conference on Artificial Neural Networks, Rhodes, Greece, October 4-7, 2018, Proceedings, Part III 27. pp. 270–279. Springer (2018)
19. Tary, J.B., Herrera, R.H., Van Der Baan, M.: Analysis of time-varying signals using continuous wavelet and synchrosqueezed transforms. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences **376**(2126), 20170254 (2018)