ISSN No:-2456-2165

A Survey and Performance Review on Air Quality Data Prediction Techniques in Deep Learning Frameworks

K.Sathya[1] , Dr.T.Ranganayaki[2]
Research Scholar[1]. Associate Professor[2]
Department of Computer Science
Erode Arts and Science College, Erode – 9
Tamil Nadu, India

Abstract:- With increased industry and urbanization, air pollution is becoming an environmental hazard. Air Quality (AQ) is becoming increasingly important for both the environment and humanity. The atmosphere contaminations that cause air pollution like CO_2, NO_2, etc., are produced by the combustion of natural gas, coal and wood, as well as by industry and cars. Air pollution may cause serious diseases such as lung cancer, brain damage, and even death. So, predicting AQ is an important step for the government to take because it is becoming a big problem for human health. To predict AQ, many artificial intelligence frameworks have been developed over earlier days using the different historical data on air pollutants in various regions. This manuscript covers a complete study of various Deep Learning (DeepLearn) frameworks developed to forecast AQ using the different available air pollution databases. First, different AQ prediction frameworks relying on the DeepLearn structures are discussed briefly. After that, a comparative study is conducted to understand the drawbacks of those frameworks and suggest a new solution to predict the AQ accurately.

Keywords:- *Air pollution, Air pollutants, Air quality, Forecasting, Artificial intelligence, DeepLearn.*

I. INTRODUCTION

In recent decades, public are more alarmed regarding air contamination as enterprises expand. The concentration of many types of pollution gas and solid particles, including PM2.5, PM10, NO2, SO2, CO, and O3, has a global influence on individual fitness and economical growth. Air pollution research is particularly significant worldwide and it has always been regarded as a crucial subject in environmental protection [1-2]. As depicted in Figure 1, distinct fundamental components cause air contamination. Such components are categorized into 2 types: primary and auxiliary. The primary components are mostly dependent on air contaminants including hard elements, petroleum combustion, traffic densities and factory emissions. All of such resources have a distinct geographical and sequential distribution. On contrary, auxiliary components are principally fabricated by meteorological information, geography and period [3]. Table 1 presents the major air pollutants and their impacts on both human health and the atmosphere.

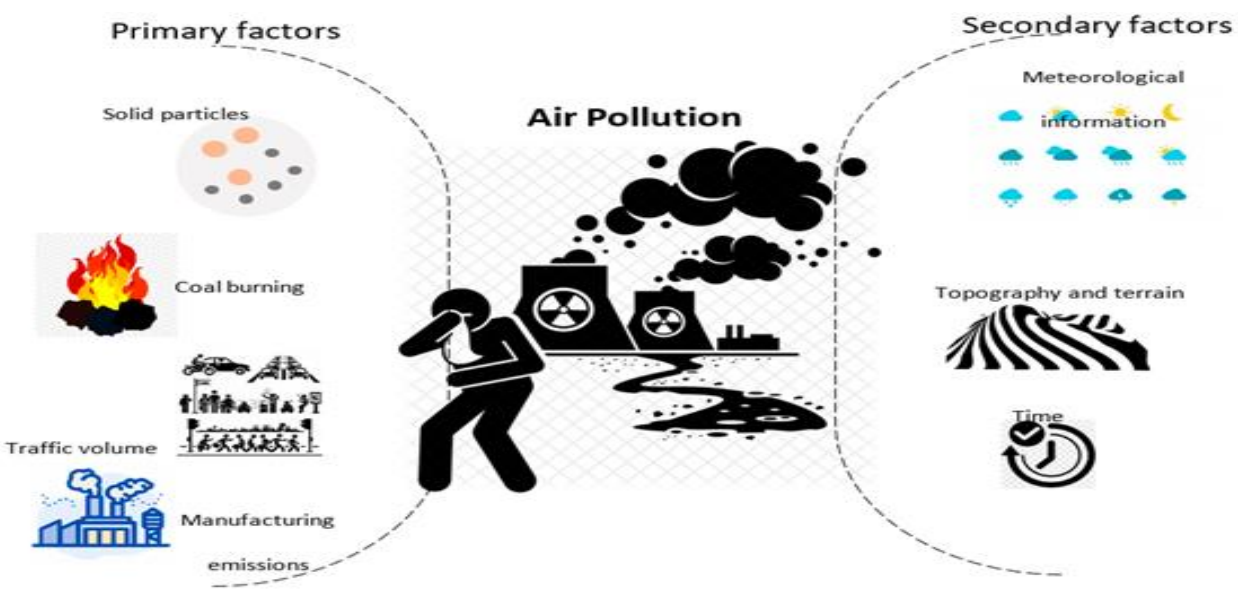

Fig. 1: Factors Responsible for Air Pollution

Pollutants	Impact on Human Health	Impact on Atmosphere
Sulphur Dioxide (SO_2)	1. Breathing issues 2. Irritating to the eyes and skin 3. Pulmonary tumor 4. Coronary disorders and fatality rate	1. Damage vegetation 2. Corrosion 3. Caustic rainfall 4. Plant and water degradation 5. Esthetic deterioration
Incomplete pollutants	1. Rise in tumor and chronic disorder mortality 2. Reduction in respiratory oscillation 3. Causes issues such as pneumonia, influenza, asthma, etc.	1. Chemical accumulation in air and water. 2. Fog in wild regions. 3. Alteration in ecosystem functions. 4. Degradation of ecological facilities.
Lead	1. Destructive behavioral alterations, training problems and chronic brain injury. 2. Impaired taste. 3. Abdominal pain and/or nausea. 4. Sleeplessness.	1. Lead toxicity to soil fauna. 2. Lead toxicity is a general disorder discovered in mammals.
Chlorine	1. Cough, breathing difficulty and hydration in the lobes.	1. Chlorine influences marine ecology. 2. Ozone depletion.
Nitrogen Dioxide (NO_2)	1. NO_2 has the potential to damage the lobes and reduce confrontation to breathing problems such as pneumonia. 2. Lowering the blood O_2 capability. 3. Impaired thyroid gland function. 4. Asthma	1. Negative effects on the biosphere. 2. Mammals subjected to NO_2 can suffer from the cilia damage and alveolar cell rupture. 3. NO_2 is contributor of chemical absorption and O_2, both of which harm vegetation.
Peroxyacetyl Nitrate (PAN)	1. Vision and breathing problems. 2. PAN is more common in those with cardiorespiratory illness.	1. Impacted vegetation. 2. Clouds take it to remote and clean places, resulting in ecosystem deterioration. 3. Deterioration of fabrics, vegetations and foam items. 4. Impacted cultivation.
Formaldehyde	1. Eye irritation, fiery pains in the eyes, and heart problems. 2. Discomfort of the eye, nose and tongue; snoring and cough.	1. It decomposes carbon monoxide and hydrochloric chemicals, which have an impact on the earth. 2. Chemical rainfall.
Carbon monoxide	1. At small quantities, drowsiness in normal individuals and breathlessness in patients with chronic problems. 2. Poor eyesight and dexterity at high doses. 3. Common cold.	1. It influences the amount of methane, O_2 and CO_2. 2. Biofuel combustion.
Hydrogen sulphide	1. It causes neurological, cardiovascular, metabolic and reproductive problems.	1. Impacted marine ecosystem.
Ammonia	1. If mixed with water, it deteriorates the cardiovascular system. 2. Pneumonia, influenza and cough.	1. Land and groundwater alkalinity. 2. Dust emission causes global warming.
Carbon dioxide (CO_2)	1. Migraine, nausea, breathlessness, shivering, fatigue, irregular heartbeat, hypertension, dementia, dyspnea and seizures.	1. Climate changes.
Ozone	1. Lung problems. 2. Chronic bronchitis. 3. Esophagus.	1. Impacted cultivation and leaf shape.

Table 1: Major Air Pollutants and Their Impacts on Human Health and Atmosphere

In smart cities, emphasis is placed on the large-scale usage of smartphones that may be used to track fine-grained air pollution [4]. Currently, agencies are tracking the behavior of pollutants in the air and educating the public about environmental contamination. This information will be useful to persons who are unwell as a result of air pollution. Environmental conservation and ecological civilization are required to overcome some of the most complex environmental problems [5]. By breathing specific gases, both people and animals are impacted by the negative impacts of air pollution, which causes respiratory, cardiological and pulmonary disorders. These changes are also causing global warming and climate change [6-7]. A significant concern for the government is to enhance people's quality of life by taking into account the elements that impact AQ.

Predicting future AQ is becoming increasingly popular since individuals can take greater steps to avoid becoming sick when they know the AQ beforehand. AQ forecasting is also important for any government's emergency management because it allows the administration to configure suitable crisis courses to reduce ecological contamination like restricting the generation and discharges of extremely contaminating endeavours and constraining transportations [8]. Nonetheless, AQ estimation is a hard process and increasing prediction efficacy if decreasing learning period and sophisticated concept in the domain of air contamination prevention. Several investigators have worked on the issue of AQ forecasting over the last few centuries. The authors of these works primarily focused on two kinds of modeling for AQ prediction such as knowledge-based and data-driven [9].

The knowledge-based frameworks primarily rely on physiochemical hypotheses to describe the movement and modification of air contamination substances. Several knowledge-based techniques were developed. But, the effective implementation of those techniques necessitates a strong basis in meteorological and environmental research. Moreover, if the technique has been utilized in various scenarios, the chemical and transportation restrictions may vary, causing the model to produce erroneous findings. To combat these issues, many statistical forecasting techniques have been adopted [10], which execute a mathematical logic and regression analysis. But, those have less efficiency, high computation period and power usage, primarily triggered using the investigation of long-term forecasting information. As well, the variety of factors accountable for air pollution makes it complex to accomplish better prediction accuracy.

Currently, the data-driven techniques for air contamination estimation were introduced because of growing big data and artificial intelligence technologies. AQ forecasting techniques depending on machine learning frameworks have tackled a few limitations of the classical statistical forecasting techniques and have become the mainstream of AQ forecasting investigation [11]. Thus, machine learning-based AQ forecasting systems have attained few better findings. However, those techniques were not suitable for extremely large-scale databases. To solve this challenge, DeepLearn frameworks such as many variants of Convolutional Neural Network (CNN) have been developed and employed in big data analysis [12]. The emergence of DeepLearn technology is significantly improved the efficacy of AQ forecasting. DeepLearn is now the foremost common data-driven approach for automatically extracting and learning the intrinsic properties of varied AQ data. A variety of studies in the past research using DeepLearn for air contamination estimation have yielded promising solutions.

The primary purpose of this paper is to give a comprehensive overview of various DeepLearn frameworks for predicting AQ. A comparison study is additionally offered to highlight the benefits and drawbacks of such frameworks to define research direction. The following sections have been arranged: Section II covers various frameworks designed to predict AQ. Section III compares such frameworks. Section IV concludes the complete study and offers the research direction.

II. REVIEW ON DEEP LEARNING-BASED AQ PREDICTION

Zheng et al. [13] designed a new Multiple Kernel Learning (MKL) framework, which represents the kind of ensemble, kernel and interpretation training strategies to predict the AQ. The mid-align scheme was utilized to train kernels and a boosting scheme was utilized to compute the appropriate kernels.

Wang & Song [14] developed a deep Spatial-Temporal Ensemble (STE) framework, which has 3 major units for AQ prediction. The primary unit was an ensemble scheme with a climate-pattern-based segregation mechanism, which learns many independent frameworks and merges them adaptively. The second unit was used to find spatial association through assessing Granger causalities among locations and creating spatial information as relative stations and relative regions. The final unit was a temporal forecaster depending on deep Long Short-Term Memory (LSTM) to train long and short-term dependencies of AQ.

Soh et al. [15] intended to predict AQ for up to 48 hours based on the mixture of many neural networks such as Artificial Neural Network (ANN), CNN and LSTM to mine spatiotemporal correlations. In this model, the different meteorology data from the past few hours and the data associated with the elevation space were considered to mine terrain effect on AQ. Also, trends from various sites, mined from associations among nearby sites and similar sites in the temporal domain were considered to create the predictive framework.

Zhou et al. [16] developed the Gaussian Process Mixture (GPM) framework that applies unknown factors posterior Hard-Cut (HC) iterative training scheme for the forecasting of fizzy contaminant dose. In this framework, iterative training was adopted and the Maximum-a-Posteriori (MAP) prediction was utilized to establish the best clustering of data that enhances the Expectation-Maximization (EM) training in GPM.

Jiang et al. [17] modeled a new hybridized learning for estimating urban AQ index. First, Wavelet Packet Decomposition (WPD) was conducted to split the actual AQ index information into low-frequency sub-sequences. After that, the Improved Pigeon-Inspired Optimization (IPIO) by means of the Particle Swarm Optimization (PSO) was used to adjust the primary weights and threshold of Extreme Learning Machine (ELM), which was applied to estimate the sub-sequences correspondingly. Also, a Multidimensional Scaling and K-means (MSK) grouping were employed to group the prediction results into high, mid-high, mid-low and low-frequency sub-sequences. Further, Modified ELM (MELM) was adopted to combine the sub-sequences to get the absolute outcomes.

Du et al. [18] exploited 1D-CNN to mine the local trend characteristics and spatial association characteristics. Also, the Bi-directional LSTM (B-LSTM) was used to learn spatiotemporal dependencies. After that, a jointly hybrid deep learner was designed depending on 1D-CNN and B-LSTM for shared interpretation characteristics learning of multivariate AQ associated time-series data.

Aristodemou et al. [19] presented an optimized 3D Variational (3DVar) Data Assimilation (DA) scheme to minimize the incongruity among forecasted contamination doses depending on Computational Fluid Dynamics (CFD). The airstream channel analyses have been executed in the EnFlo metrological airstream channel and the predicted framework was variable, public CFD tool by mesh adaptivity. Then, the optimum super-mesh has been created and utilized in the variational DA task and to interpolate the wind tunnel information.

Liu et al. [20] designed a seq2seq framework called an Attention-based AQ Predictor (AAQP), which uses historical AQ data and climate data to forecast AQ indexes. Initially, the actual Recurrent Neural Network (RNN) in the encoding unit has been substituted by the fully connected encoder to speed up the learning task. As well, location embedding was adopted to support the fully connected encoder to obtain the consecutive correlations amid origin data series. Moreover, the n-step recurrent forecasting was used to resolve the error accumulation triggered by recurrent forecasting.

Tao et al. [21] developed a short-term prediction framework depending on the deep learner for PM2.5 density and the Convolutional-based Bidirectional Gated Recurrent Unit (CBGRU) model that integrates 1D-CNNs and BGRU neural networks. First, the CNN was utilized to perform downsampling of data, which minimizes the dimensionality and enhances the generalization ability of the framework. After that, the reduced-dimensional data was provided to the RNN to extract the data features offered by various information sources in metrological information and find the nonlinear correlation between the time-sequence of PM2.5.

Ma et al. [22] developed a Transfer Learning B-LSTM (TL-BLSTM) framework to predict AQ. First, the raw data was collected and pre-processed to create the time-series instances. After that, the BLSTM was trained to forecast the air qualities. In this model, the BLSTM was trained by the long-term reliance's of PM2.5 and the TL was applied to shift the data trained from low-temporal resolutions to higher ones.

Lv et al. [23] developed Multi-view Transfer Semi-supervised training for AQ Estimation (MTSAE). Primarily, metropolitan and non-metropolitan regions were differentiated by the terrain characteristics. Then, the initial frameworks were trained by the transfer regression scheme via enabling the annotated information from another town. Also, the initial frameworks were improved by the semi-supervised regression scheme using the un-annotated information from the desired town.

An enhanced wind-aware attention strategy with the LSTM structure [24] was designed to estimate PM2.5 absorptions utilizing the effect of wind course and velocity on the alterations of spatiotemporal PM2.5 absorptions in adjacent regions. Initial forecasting for PM2.5 was prepared by the LSTM concerning nearby pollution and an ensemble training depending on eXtreme Gradient Boosting (XGBoost) was applied to merge the initial forecasting with climate prediction to create the secondary stage forecasting of PM2.5.

Zeinalnezhad et al. [25] suggested the Adaptive Neuro-Fuzzy Inference System (ANFIS) model to enhance the efficiency of the day-by-day forecasting of pollutants through time-series data evaluation. A nonlinear multivariate regression framework has been designed and improved to get the minimum loss. First, the database consisting of various pollutants information was gathered from a particular forecasting region in Tehran. Then, the database was partitioned into learning and validation collections. The learning collection was fed to the ANFIS for learning the different pollutant factors and the validation collection was used to predict the AQ.

Janarthanan et al. [26] developed a deep learner to predict the AQ index in Chennai town. First, the database was gathered and pre-processed to substitute missing values as well as eliminate unwanted information. Then, the average, mean square error and standard variance were mined by the Grey Level Co-occurrence Matrix (GLCM). The mixture of Support Vector Regression (SVR) and LSTM framework was applied to categorize the AQ index values.

Mokhtari et al. [27] designed a multi-point deep learner depending on Convolutional LSTM (ConvLSTM) for predicting highly dynamic AQ. In this model, the CNN and LSTM were merged to extract temporal and spatial information characteristics. Also, uncertainty quantification schemes such as Monte-Carlo (MC) dropout and quantile regression were executed on top of this ConvLSTM structure and their efficiencies were analyzed.

III. COMPARATIVE STUDY

This part compares the strengths and drawbacks of the above-discussed DeepLearn frameworks for AQ prediction in Table 1.

Ref. No.	Frameworks	Merits	Demerits	Database	Performance
[13]	MKL	Improved predictive facility for severe conditions of AQ.	It needs more sophisticated models to create a more comprehensive and powerful AQ forecasting system.	Database collected from the Hong Kong and Beijing	Mean Squared Error (MSE) (for future 12 hours PM2.5 in Beijing)=1.536; MSE (for future 12 hours PM2.5 in Hong Kong)=0.609
[14]	Deep STE and LSTM	It can train the long and short-term reliance of AQ.	Its accuracy was very less.	Database composed of AQ data and weather forecast data from 37 stations in Beijing	Root MSE (RMSE)=55.02
[15]	Mixture of ANN, CNN and LSTM	It can be helpful for longer time frame forecasting.	It needs to adjust and eliminate the noise owing to system variances. Also, it needs to collect more chemical factors, which influence PM2.5 elements.	Taiwan and Beijing databases	**Taiwan database:** Mean Absolute Error (MAE) (for 6 hours)=10 **Beijing database:** MAE (for 6 hours)=44.8
[16]	GPM and MAP estimation	It can offer prognostic belief periods to create the forecasting high consistent.	It did not analyze the forecasting of AQ in noisy atmospheres since it was sensitive to noise.	The 101 to 500 instances of the 2-gas sequences are created as the training data and the 501 to 900 instances are created as the test data.	RMSE=0.1241
[17]	WPD, IPIO, MSK and MELM	It exhibits high predictability for short, mid and long-term predictions.	It needs to consider other factors such as seasonality, climate and vacation.	Hourly AQI statistics of Harbin	RMSE (for short-term)=0.1823; RMSE (for mid-term)=0.3308; RMSE (for long-term)=0.547
[18]	1D-CNN and B-LSTM	It has better prediction ability.	It needs to predict the abrupt change of air pollution time series data to enhance the ability of multi-step prediction.	The Beijing AQ database from UCI and urban AQ database gathered in the Urban Air project of Microsoft Re-search	**Beijing PM2.5 database:** RMSE=8.2 **Urban AQ database:** RMSE=9.96
[19]	3DVar DA	Better prediction accuracy and needs less data to minimize the errors.	It needs unbiased frameworks of DA to satisfy hypotheses created in their formulation.	1391 study points, situated downstream of the contaminant source	MSE=3.16×10^{-1}
[20]	AAQP	It has an accurate prediction and a higher training speed.	It focuses only on temporal attention and uses single station data while in real-time, there were correlations among various stations.	The hourly AQ data and hourly weather data of Beijing	**Olympic center station:** MAE=32.848 **Dongsi station:** MAE=41.468
[21]	CBGRU	It can enhance attribute training facilities in time sequence and increase the accuracy considerably.	There were more characteristics needed to increase the prediction efficiency.	The Beijing PM2.5 database in UCI library	RMSE=14.5319
[22]	TL-BLSTM	It can be useful to create and evaluate more proper trends for long-term AQ forecasting.	It needs other influential factors of PM2.5 to increase the efficiency and it has a high training period.	3 years AQ statistics of each forecasting units in the Guangdong province	RMSE=8.5426
[23]	MTSAE	It was able to yield a competitive spatial	The efficiency was affected by different conditions of	Real datasets from Hangzhou, Ningbo, and	RMSE=12 (for urban regions)

		AQ prediction efficiency on non-urban areas.	the local region.	Wuxi	RMSE=20 (for non-urban regions)
[24]	Wind-aware attention strategy with LSTM	It may enhance the efficacy of short and long-term forecasting.	The efficiency was influenced by the missing values in the database and several hyperparameters.	PM2.5 contamination database including climate annotations from the Environmental Protection Administration (EPA) and a climate predict database gathered from the Central Weather Bureau (CWB)	RMSE (for 24 hours)=13.9975
[25]	ANFIS	It achieved the minimum error using the membership function with triangular distribution.	It needs to choose an appropriate fuzzy membership function to get the minimum error.	Information on contaminants involving CO, SO_2, O_3, and NO_2	RMSE=0.1943
[26]	SVR-LSTM	It can be helpful in developing a sustainable community in urban regions.	It needs to consider air pollution control strategies to enhance the town's AQ index values.	Data acquired from Manali, Velachery and Alandur.	RMSE=10.9
[27]	ConvLSTM	It can able to predict AQ at different periods concurrently.	The efficiency of MC dropout greatly relies on dropout values and it has a high cost for high amount of stochastic forward.	Fusion Field Trial (FFT) 2007 real database	MSE (for 28 trials)=0.9; MSE (for 55 trials)=0.33

Table 1: Comparison of Different DeepLearn Frameworks for AQ Prediction

IV. CONCLUSION

This report provides a comprehensive analysis of several AQ forecasting frameworks using DeepLearn structures. According to the findings of this review, numerous academics have expertise in creating DeepLearn frameworks that forecast AQ in specific locations efficiently. Among those different frameworks, ConvLSTM achieves the maximum accuracy in predicting highly dynamic AQ by mining spatial and temporal characteristics. On the other hand, the limitation of this framework is that the efficiency of MC dropout greatly relies on the dropout ratio and it has a high cost while the amount of stochastic forward was high. So, future research will focus on deciding the appropriate dropout rates and the number of stochastic forward to improve the accuracy and lessen the expense of forecasting AQ.

REFERENCES

[1.] India-average AQI by city 2020 | Statista. *Statista*. [Online]. Available: https://www.statista.com/statistics/1115024/india-aqi-by-city/. [Accessed: Apr. 12, 2022]

[2.] WHO. *Air pollution*. [Online]. Available: https://www.who.int/health-topics/air-pollution. [Accessed: Apr. 12, 2022]

[3.] Manisalidis, I., Stavropoulou, E., Stavropoulos, A., & Bezirtzoglou, E. (2020). Environmental and health impacts of air pollution: a review. *Frontiers in Public Health*, 14, 1-13.

[4.] Toma, C., Alexandru, A., Popa, M., & Zamfiroiu, A. (2019). IoT solution for smart cities' pollution monitoring and the security challenges. *Sensors*, *19*(15), 1-24.

[5.] Jiang, X. Q., Mei, X. D., & Feng, D. (2016). Air pollution and chronic airway diseases: what should people know and do?. *Journal of Thoracic Disease*, *8*(1), E31-E40.

[6.] Joshi, M., Goraya, H., Joshi, A., & Bartter, T. (2020). Climate change and respiratory diseases: a 2020 perspective. *Current Opinion in Pulmonary Medicine*, *26*(2), 119-127.

[7.] Sillmann, J., Aunan, K., Emberson, L., Büker, P., Van Oort, B., O'Neill, C., ... & Brisebois, A. (2021). Combined impacts of climate and air pollution on human health and agricultural productivity. *Environmental Research Letters*, *16*, 1-35.

[8.] Bai, L., Wang, J., Ma, X., & Lu, H. (2018). Air pollution forecasts: An overview. *International Journal of Environmental Research and Public Health*, *15*(4), 1-44.

[9.] Mengara Mengara, A. G., Kim, Y., Yoo, Y., & Ahn, J. (2020). Distributed deep features extraction model for AQ forecasting. *Sustainability*, *12*(19), 1-19.

[10.] Taheri Shahraiyni, H., & Sodoudi, S. (2016). Statistical modeling approaches for PM10 prediction in urban areas; a review of 21st-century studies. *Atmosphere*, *7*(2), 1-24.

[11.] Liang, Y. C., Maimury, Y., Chen, A. H. L., & Juarez, J. R. C. (2020). Machine learning-based prediction of AQ. *Applied Sciences*, *10*(24), 1-17.

[12.] Mao, Y., & Lee, S. (2019, August). Deep convolutional neural network for AQ prediction. In *Journal of Physics: Conference Series*, *1302*(3), 1-6.

[13.] Zheng, H., Li, H., Lu, X., & Ruan, T. (2018). A multiple kernel learning approach for AQ prediction. *Advances in Meteorology*, *2018*, 1-15.

[14.] Wang, J., & Song, G. (2018). A deep spatial-temporal ensemble model for AQ prediction. *Neurocomputing*, *314*, 198-206.

[15.] Soh, P. W., Chang, J. W., & Huang, J. W. (2018). Adaptive DeepLearn-based AQ prediction model using the most relevant spatial-temporal relations. *Ieee Access*, *6*, 38186-38199.

[16.] Zhou, Y., Zhao, X., Lin, K. P., Wang, C. H., & Li, L. (2019). A Gaussian process mixture model-based hard-cut iterative learning algorithm for AQ prediction. *Applied Soft Computing*, *85*, 1-12.

[17.] Jiang, F., He, J., & Tian, T. (2019). A clustering-based ensemble approach with improved pigeon-inspired optimization and extreme learning machine for AQ prediction. *Applied Soft Computing*, *85*, 1-29.

[18.] Du, S., Li, T., Yang, Y., & Horng, S. J. (2019). Deep AQ forecasting using hybrid DeepLearn framework. *IEEE Transactions on Knowledge and Data Engineering*, *33*(6), 2412-2424.

[19.] Aristodemou, E., Arcucci, R., Mottet, L., Robins, A., Pain, C., & Guo, Y. K. (2019). Enhancing CFD-LES air pollution prediction accuracy using data assimilation. *Building and Environment*, *165*, 1-15.

[20.] Liu, B., Yan, S., Li, J., Qu, G., Li, Y., Lang, J., & Gu, R. (2019). A sequence-to-sequence AQ predictor based on the n-step recurrent prediction. *IEEE Access*, *7*, 43331-43345.

[21.] Tao, Q., Liu, F., Li, Y., & Sidorov, D. (2019). Air pollution forecasting using a DeepLearn model based on 1D convnets and bidirectional GRU. *IEEE Access*, *7*, 76690-76698.

[22.] Ma, J., Cheng, J. C., Lin, C., Tan, Y., & Zhang, J. (2019). Improving AQ prediction accuracy at larger temporal resolutions using DeepLearn and transfer learning techniques. *Atmospheric Environment*, *214*, 1-10.

[23.] Lv, M., Li, Y., Chen, L., & Chen, T. (2019). AQ estimation by exploiting terrain features and multi-view transfer semi-supervised regression. *Information Sciences*, *483*, 82-95.

[24.] Liu, D. R., Lee, S. J., Huang, Y., & Chiu, C. J. (2020). Air pollution forecasting based on attention-based LSTM neural network and ensemble learning. *Expert Systems*, *37*(3), 1-16.

[25.] Zeinalnezhad, M., Chofreh, A. G., Goni, F. A., & Klemeš, J. J. (2020). Air pollution prediction using semi-experimental regression model and adaptive Neuro-fuzzy inference system. *Journal of Cleaner Production*, *261*, 1-16.

[26.] Janarthanan, R., Partheeban, P., Somasundaram, K., & Elamparithi, P. N. (2021). A DeepLearn approach for prediction of AQ index in a metropolitan city. *Sustainable Cities and Society*, *67*, 1-11.

[27.] Mokhtari, I., Bechkit, W., Rivano, H., & Yaici, M. R. (2021). Uncertainty-aware DeepLearn architectures for highly dynamic AQ prediction. *IEEE Access*, *9*, 14765-14778.