Image-Based Air Quality Estimation Using Convolutional Neural Network Optimized by Genetic Algorithms: A Multi-Dataset Approach

Arshad Ali Khan[1], Mazlina Abdul Majid[2], Abdulhalim Dandoush[3]
Faculty of Computing, Universiti Malaysia Pahang Al-Sultan Abdullah, Pekan, Malaysia[1,2]
Centre for Artificial Intelligence & Data Science, Gambang, Malaysia[2]
University of Doha for Science and Technology (UDST), Doha, Qatar[3]

***Abstract*—Air pollution poses significant threats to human health and the environment, making effective monitoring increasingly essential. Traditional methods using fixed monitoring stations have challenges related to high costs and limited coverage. This paper proposes a new approach using convolutional neural networks with genetic algorithms for estimating air quality directly from images. The convolutional neural network is optimized using genetic algorithms, which dynamically tune hyperparameters such as learning rate, batch size, and momentum to improve performance and generalizability across diverse environmental conditions. Our approach improves performance and reduces the risk of overfitting, thus ensuring balanced and robust results. To mitigate overfitting, we implemented dropout layers, batch normalization, and early stopping, significantly enhancing the model's generalization capability. Specifically, three different open-access datasets were combined into a single training dataset, capturing extensive temporal, spatial, and environmental variability. Extensive testing of the model performance was conducted with a broad set of metrics, including precision, recall, and F1 score. The results demonstrate that our model not only achieves high accuracy but also maintains well-balanced performance across all metrics, ensuring robust classification of different air quality levels. For instance, the model achieved a precision of 0.97, a recall of 0.97, and an overall accuracy of 0.9544 percent, outperforming baseline methods significantly in all metrics. These improvements underscore the effectiveness of Genetic Algorithms in optimizing the model.**

Keywords*—Convolutional neural network; Genetic Algorithm; air quality estimation; image processing

I. Introduction

Air pollution is a major environmental risk that has increasingly become a critical issue, posing significant health threats and adverse effects on the environment [1]. The main group of air pollutants includes particulate matter (PM), specifically classified as PM10 (particles with aerodynamic diameters less than 10 µm) and PM2.5 (particles with aerodynamic diameters less than 2.5 µm), nitrogen dioxide (NO_2), sulfur dioxide (SO_2), oxides of nitrogen (NO_x), and carbon monoxide (CO) [2]. The World Health Organization (WHO) estimates that air pollution contributed to approximately 4.2 million premature deaths worldwide in 2016 [3]. Estimating air pollution emissions is crucial to controlling air pollution [4]. However, many traditional methods developed for this purpose are now outdated and rely on expensive, region-specific fixed stations that often fail to provide comprehensive real-time data.

Recent advances in computer vision and deep learning offer a promising alternative to these conventional methods. The increasing presence of cameras in public spaces, vehicles, and personal devices presents an opportunity to leverage image data for air quality estimation. Convolutional Neural Networks (CNNs), which excel in extracting and analyzing complex visual features, are powerful tools for tasks such as image recognition [5]. They have been increasingly applied in environmental monitoring [6]. In recent years, image-based methods have been proposed to detect air quality, which have demonstrated good accuracy in specific scenarios [7]. However, despite their potential, CNNs often struggle to generalize in diverse environmental conditions due to limited data set diversity and static hyperparameter configurations. Previous studies, such as those conducted by Zhang et al. [8] and Song et al. [9], have successfully demonstrated the feasibility of using CNNs to estimate air pollution levels from images. Recently, numerous articles have been published on estimating air quality from image datasets. However, these approaches often face limitations due to the narrow scope of datasets and challenges in optimizing hyperparameters, which can constrain their broader applicability.

This paper addresses these gaps by presenting a novel approach to air quality estimation that combines the power of CNNs with Genetic Algorithms (GA) to dynamically optimize hyperparameters and enhance model performance. By integrating three diverse open-source datasets, covering a wide range of temporal and spatial variations, our model can perform well under different environmental conditions. This comprehensive dataset includes images captured at different times of the day, in various weather conditions, and in multiple geographic locations, providing a solid foundation for accurate air quality prediction.

A key innovation of our approach shown in Fig. 1 lies in the integration of GA with the CNN framework. The GA component dynamically optimizes CNN's hyperparameter, such as learning rate, batch size, and momentum, to achieve optimal performance. This evolutionary technique allows the model to adapt to various environmental scenarios, significantly improving its accuracy and generalization capabilities. By leveraging this hybrid approach, our model not only achieves high accuracy but also maintains balanced performance across multiple evaluation metrics, making it a powerful tool for real-time air quality monitoring.

The remainder of this paper is structured as follows: Section II reviews related works on air quality estimation, highlighting not only the limitations of traditional methods but also the recent advancements in image-based approaches. Section III details the proposed method, including data collection, preprocessing, and the CNN-GA framework. Section IV presents the experimental setup and the performance evaluation, followed by a discussion of the results obtained in Section V. Finally, Section VI concludes the paper and suggests directions for future research.

II. Related Work

This section reviews both traditional and modern image-based methods used for air pollution estimation.

A. Traditional Methods

Various traditional methods have been developed over the past few decades to estimate air quality. These can be further divided into two major groups, namely, ground-based monitoring and modeling techniques.

In ground-based monitoring, air pollutants are generally monitored using fixed stations installed by environmental or government institutions [10]. Common types of air pollutants monitored include PM2.5, PM10, NO2, SO2, CO, O3, and VOCs [11]. However, this sparse network of regulatory monitoring stations is usually not sufficient for mapping out spatial variations in air pollutants among a considerable population in urban areas. These networks cannot provide high-resolution data for the efficient management of air quality and exposure [11], [12]. Besides, conventional methods of monitoring are costly and cannot capture the temporal-spatial heterogeneity of urban pollution, which restricts their ability to find hotspots of pollution and further management thereof [12].

Common modeling techniques include deterministic and statistical models. Deterministic models use known, based, and expressed mathematical relationships concerning processes underlying CTMs in order to model the emission, transport, transformation, and removal by deposition from the atmosphere [13].Their principal strength is that for sufficiently small scale and homogeneity, they are capable of predicting, with a high spatial resolution, very detailed quantitative data regarding the different complex atmospheric flow phenomena transporting various constituents with pollutants. However, these models presuppose considerable a prior knowledge in the form of reliable and extensive data with respect to atmospheric conditions and sources of pollutants [14]. The use of idealized assumptions and detailed input often makes these models inapplicable and less accurate, especially in regions where small-scale atmospheric data is not available. Besides, deterministic models are computationally intensive, hence unsuitable for real-time applications.

They do not use detailed representations of physical and chemical processes. They rather attempt to find the factorial relationship that exists between a set of factors influencing air pollutant concentrations using statistical techniques. These methods are usually divided into two broad methods: classical methods or traditional machine learning. The important representative classical methods include the ARIMA model [15]–[17]. Among the machine learning methods, ANN is widely used as it simulates the human brain's system for nonlinear sequence modeling. After years of research and application, more advanced versions have evolved for air pollution prediction, such as the Backpropagation Neural Network (BPNN) [18], the Generalized Regression Neural Network (GRNN) [19], and the ensemble ANN approach [20].

Despite their advancements in improving prediction accuracy, statistical methods often struggle to capture complex, nonlinear spatio-temporal correlations and tend to learn only shallow features [14]. Additionally, these models generally perform well only on small-scale datasets, making them less effective for large-scale and dynamically changing air pollutant data that require more sophisticated modeling of spatio-temporal relationships [21], [22]. This limitation also contributes to generalization gaps, where models trained on specific datasets may not perform adequately when applied to new or unseen environments, reducing their overall effectiveness in broad, real-world applications.

B. Image-Based Air Pollution Estimation Methods

Recent advancements in image-based air pollution estimation have leveraged the capabilities of deep learning techniques, particularly Convolutional Neural Networks (CNNs), to significantly enhance the accuracy and efficiency of air quality assessments. These methods provide a scalable, cost-effective alternative to traditional air quality monitoring systems, which are often constrained by high costs and limited spatial coverage.

One prominent approach involves the use of a Double-Channel Weighted Convolutional Neural Network (DCWCN), which processes different parts of an image, such as the sky and buildings, to extract relevant features separately. This technique enhances the accuracy of air quality estimation by focusing on distinct components of the environment, thereby addressing variability in image content due to factors like lighting and weather conditions. The DCWCN architecture includes two separate feature extraction networks for both channels, followed by a feature weights self-learning method that performs weighted feature fusion, combining the extracted features before classification [23].

Zhang et al. [8] developed a convolutional neural network (CNN) and improved both the convolutional layer and classification layer activation functions. They proposed a new activation function, EPAPL, and replaced the traditional SoftMax classifier with a Negative Log-Log Ordinal Classifier in the classification layer. This network was trained using environmental images to predict classifications, and it successfully performed the task of measuring PM2.5 and PM10 levels across six different grades.

One approach integrates Convolutional Neural Networks (CNNs) with regression classifiers to create a hybrid model (CNN-RC) that processes images and HSV (Hue, Saturation, Value) statistics to estimate PM2.5, PM10, and AQI levels. This multi-input multi-output (MIMO) framework has demonstrated significant improvements in estimation accuracy, particularly when handling both daytime and nighttime images. The model's effectiveness is attributed to its ability to deeply learn from high-dimensional datasets and the incorporation of HSV statistics, which play a crucial role in enhancing the estimation reliability by correlating current images with baseline images [24].

The AQC-Net framework, as proposed by Zhang et al. [25], integrates a Convolutional Neural Network (CNN) with a Spatial and Context Attention (SCA) module to create a model that processes images captured by mobile devices to estimate air quality levels such as PM2.5, PM10, and AQI. This deep learning framework leverages ResNet18 for feature extraction, while the SCA module enhances the model's ability to capture global contextual information and inter-channel dependencies. The model has demonstrated significant improvements in classification accuracy, particularly by focusing on the spatial and contextual relationships within images, making it highly effective across various environmental conditions and locations. The model's effectiveness is attributed to its ability to deeply learn from scene images and the integration of the SCA module, which recalibrates feature maps for improved air quality estimation reliability.

This paper presents an innovative method for air quality estimation by integrating Convolutional Neural Networks (CNNs) with Genetic Algorithms (GAs) to dynamically optimize hyperparameters. The CNN is utilized for its robust feature extraction capabilities, enabling it to process images and estimate air quality indicators such as PM2.5, PM10, and the Air Quality Index (AQI). The approach is further strengthened by the amalgamation of three distinct open-source datasets into a single, comprehensive data set, which provides a broad spectrum of temporal and spatial variations for model training.

A significant contribution of this work is the application of GAs to optimize critical CNN hyperparameters, including learning rate and batch size, allowing the model to adapt effectively to diverse environmental conditions. This hybrid CNN-GA approach not only enhances the model's accuracy but also improves its generalization capabilities, making it particularly suitable for real-time air quality monitoring. The model's effectiveness was thoroughly assessed using key performance metrics such as Precision, Recall, F1-Score, and ROC-AUC, where it consistently demonstrated superior accuracy and a well-balanced performance across various environmental scenarios.

III. Proposed Method

This section details the methodology employed in this study, covering data collection, preprocessing, and the CNN-GA proposed model used for air quality estimation.

A. Data Collection and Preprocessing

To develop a robust and generalized CNN model for air quality estimation, we utilized three diverse, publicly available, open-source datasets. The selected datasets represent a diverse array of environmental conditions, including variations in geographical location, weather patterns, and lighting conditions. This diversity is crucial for training a model that can generalize well across different regions and times, making it adaptable for global application. In total, 12,902 images were collected from these datasets. The dataset was split into 80% for training and 20% for validation, ensuring a balanced distribution for model evaluation.

TABLE I. AQI Category Image Count Across Different Datasets

AQI Category	Dataset-A	Dataset-B	Dataset-C
Good	1541	135	58
Moderate	1573	188	52
Unhealthy for Sensitive Groups	2863	29	8
Unhealthy	2622	78	50
Very Unhealthy	2194	26	22
Hazardous	1447	0	16

1) Dataset A [26]: Combined Air Quality Dataset from India and Nepal : includes 12,240 pictures that depict different aspects of air quality in Indian and Nepali cities [26]. All images maintain the same resolution of 224 x 224 each. The images are divided into two categories: the combined dataset and country wise dataset. In this dataset, the folder named" Combined Dataset" focused on categorizing air quality into six categories based on the AQI, namely, Good, Moderate, Unhealthy for Sensitive Groups, Unhealthy, Very Unhealthy, and Hazardous/Severe. This detailed classification offers an extended framework for analyzing air quality in diverse environmental conditions.

2) Dataset B [27]: Smartphone-Based Air Pollution Image Dataset (SAPID) was retrieved from Mendeley Data and is identified as the Smartphone-Based Air Pollution Image Dataset, SAPID [27]. The dataset consists of 456 images displaying various air pollution levels in accordance with the United States Environmental Protection Agency categorization. Images are divided into five AQI classes. This dataset is a very important source for developing and testing computer vision algorithms with the purpose of air quality assessment based on visual data represented by images taken from smartphones, where structured categorization enables detailed analysis and modeling.

3) Dataset C [28]: PM2.5 Image Dataset from Kaggle is provided by Kaggle; the material is entitled "Pictures and Air Quality." It contains images pre-classified into their respective conditions according to the PM2.5 values represented in their PM2.5data.csv file [28]. The 2.5 data have exact concentrations with corresponding images, making it suitable to classify images into normal and polluted classes according to the conventional standard for air. Table I presents the distribution of all images of "Pictures and Air Quality Dataset" that have been prepared according to their corresponding level of PM2.5 concentration.

B. Data Preprocessing

Preparing a dataset for the training of a deep learning model in air quality estimation involves images from different sources and varying dimensions and resolutions. To ensure consistency and quality in the dataset, we implemented a preprocessing pipeline that includes image resizing and quality filtering. The algorithm used for this process is outlined below.

Algorithm 1 is used to preprocess the dataset by standardizing image dimensions and filtering low-quality images. All images were resized to 224 × 224 pixels to ensure uniformity in input data for the deep learning model. The algorithm first iterates through the dataset, verifying file formats and extracting image dimensions before resizing. Next, it applies two quality checks: a uniformity check, which removes nearly blank images using standard deviation analysis, and a sharpness check, which filters out blurry images based on the variance of the Laplacian filter. Only high-quality images that pass both checks are retained and saved in the output directory. This preprocessing step ensures that the dataset contains clear, informative images, improving the accuracy of air quality estimation.

Algorithm 1 Image Resizing and Filtering of Images

Input: A set of images to be resized.
Output: A set of resized and filtered images.
Initialization:
Set $desired\_size \leftarrow (224, 224)$ pixels
Set $uniform\_threshold \leftarrow 5-10$ (filters almost uniform images)
Set $blur\_threshold \leftarrow 50 - 100$ (filters very blurry images)
Prepare $output\_dir$ for saving cropped images
Initialize counters: $resize\_count \leftarrow 0$, $filtered\_count \leftarrow 0$
for each file f_i in f where $i \geq 0$ **do**
  Check File Type: If f_i has extension .png, .jpg, or .jpeg, proceed
  Load Image: Import the image
  Determine Dimensions: Extract image width W and height H
  Set $resize\_width \leftarrow desired\_size[0]$ and $resize\_height \leftarrow desired\_size[1]$
  Resize image to $(resize\_width, resize\_height)$
end for
Quality Filtering:
Uniformity Check:
Convert image to grayscale using `cv2.cvtColor`
Compute standard deviation: $stddev \leftarrow np.std(image)$
if $stddev < uniform\_threshold$ **then**
  Return **True** (filter out the image)
else
  Return **False**
end if
Sharpness Check:
Apply Laplacian filter using `cv2.Laplacian`
Compute variance: $variance \leftarrow laplacian.var()$
if $variance < blur\_threshold$ **then**
  Return **True** (filter out the image)
else
  Return **False**
end if
Save resized and filtered image to $output\_dir$

C. Generalized Convolutional Neural Network (CNN)

CNN is a type of feed-forward Artificial Neural Network (ANN) that is structured using a deep learning algorithm. It has been extensively applied in various domains, including image processing, video recognition, and time series forecasting [29]–[39]. Empirically, CNNs are widely recognized for their robust feature extraction capabilities from images, making them suitable for tasks involving image-based data. This structure is well-suited for the problem of air quality estimation, where extracting complex visual patterns (e.g. particulate matter, pollution indicators in environmental images) is key to classification. The architecture of our CNN is summarized in Table II. The CNN architecture is divided into two primary phases as CNN part: feature extraction and classification, comprising convolutional, pooling, and fully connected layers.

1) Feature Extraction: The feature extraction phase begins with the input image of size 224×224×3 through several convolutional and max-pooling layers. The first convolutional layer applies A set of 96 filters of size 11×11 with a stride of 4 is applied, resulting in an output size of 224×224×96. This operation is mathematically defined as. The operation for each filter is defined as:

$$O_1 = f(W_1 * I + b_1)$$

where W1 represent the weights represent the output, b1 represents the bias of the first convolutional layer, I is the input image, and f is the ReLU activation function. A 3×3 max-pooling operation with a stride of 2 reduces the dimensions to 112×112×96. The output is represented as:

$$P_1 = \max pool(O_1)$$

The second convolutional layer employs a set of 256 filters of size 5×5 with a stride of 1, resulting in an output of 112×112×256 represented by P.

$$O_2 = f(W_2 * P_1 + b_2)$$

This output is further sampled to 56×56×256 via a 3×3 max-pooling operation.

$$P_2 = \max pool(O_2)$$

Three more convolutional layers follow, with varying filter sizes and counts. Each layer applies ReLU activation and batch normalization to stabilize and improve learning. The final feature map is obtained after pooling:

$$P_3 = \max pool(O_5)$$

where O5 is the output from the last convolution layer (56×56×256), and P3 is the result of the third pooling layer, reducing it to 28×28×256.

2) Classification: In the classification phase, the output from the last pooling layer is flattened into a vector of size 50176:

$$F = Flatten(P_3)$$

This vector is passed through a fully connected layer of 4096 units:

$$O_{fc} = f(W_{fc} \bullet F + b_{fc})$$

Subsequently, two dropout layers (with a rate of 0.6) are applied to prevent overfitting. Finally, the output is fed into another fully connected layer with a softmax activation function, providing class probability:

$$P = soft\max(W_{out} \bullet O_{fc} + b_{out})$$

This CNN architecture effectively extracts hierarchical features from the input image and performs classification, making it well-suited for complex image recognition tasks.

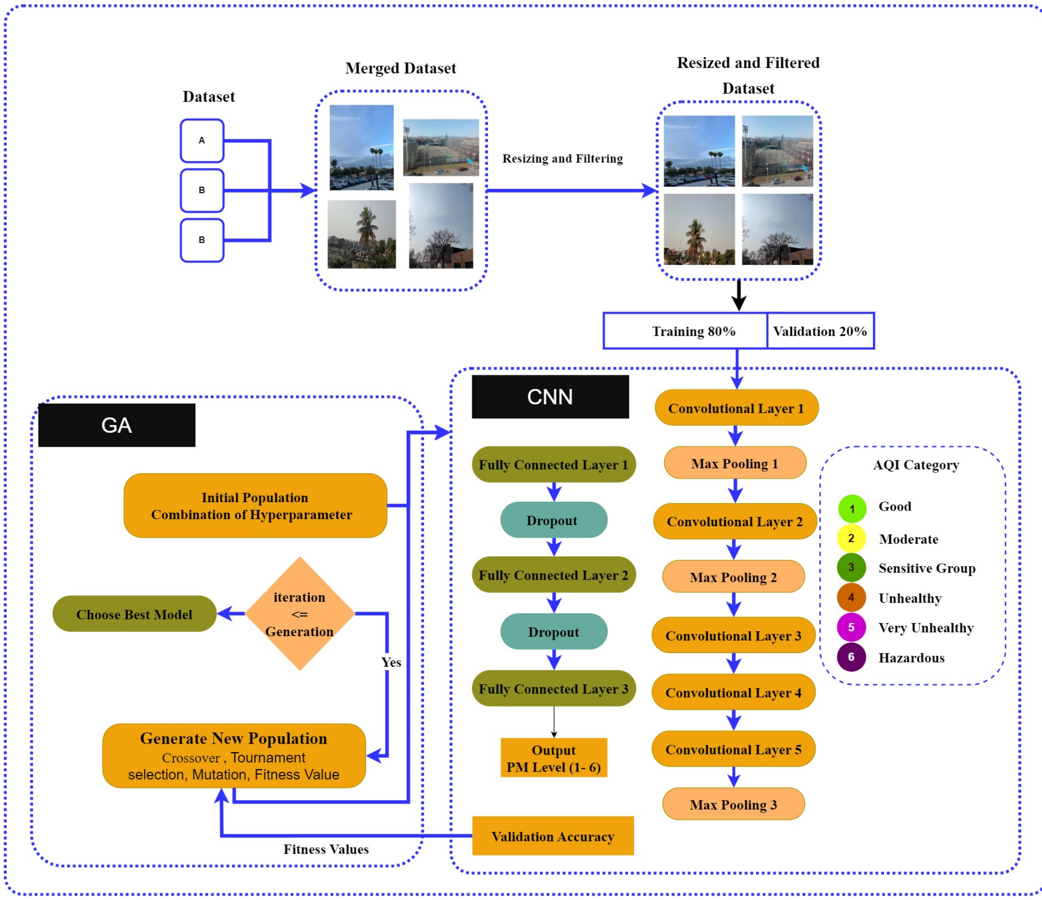

Fig. 1. Proposed model for air quality estimation from images using CNN with GA for hyperparameter optimization.

3) Hyperparameter tuning using GA: Hyperparameters may be defined as the very important parameters set prior to training either a machine learning or deep learning model. Speaking broadly, there exists a division into two types of hyperparameters. One group involves identifying the network structure, where the kernel size and type, stride and padding, number of hidden layers, and activation function determine the hyperparameters. These parameters define the architecture and the complexity of the model. The latter group includes such hyperparameters as how to train the network: a learning rate, momentum, number of epochs, batch size. Hyperparameters responsible for the training process supervise the efficiency and effectiveness of the whole learning process; therefore, this is relevant for convergence and generalization of this model regarding new data. Both model and algorithmic hyperparameters are very important for the optimization of model performance and need to be tuned carefully to derive the best results. Optimization techniques make much difference in the performance of hyperparameter tuning in deep learning by bringing improvements in model accuracy, reduction in computational cost, and enhancing efficiency [40], [41].

The GA is used to carry out the automation of the optimization of hyperparameters related to the training: learning rate, batch size, and momentum. It is an evolutionary technique for hyperparameter tuning, which explores a wide range to find those that allow the maximum CNN performance on the validation dataset. The process begins by initializing an initial population, where each one represents a combination of hyperparameters with the following ranges as shown in Table III:

Fitness evaluation is performed by training the CNN for 30 epochs, using the validation accuracy as the fitness score, calculated as:

$$fitness = \frac{1}{N}\sum_{i=1}^{N} 1(y_i = \hat{y}_i)$$

where N is the number of validation samples, y_i is the true label, and $\hat{y}_i$ is the predicted label. After evaluating fitness, genetic operators are applied. Then, tournament selection is used to choose individuals based on their fitness scores, followed by a two-point crossover to combine parents and generate offspring. The mutation would be done with a probability of 0.2 so as not to lose the diversity in the population. The generated population will evolve over successive generations, ensuring at each step in selection that the best of these formed generations increases the performance of the model at each step.

IV. Results

This section presents the experimental results of our approach. The performance is evaluated on a validation set using evaluation metrics such as precision, recall, and F1-score.

TABLE II. PROPOSED ARCHITECTURE OF CONVOLUTIONAL NEURAL NETWORK

Layer	Output Shape	Filter Size	Number of Filters	Stride	Padding	Activation
Input Layer	224×224×3	-	-	-	-	-
Conv Layer 1	224×224×96	11×11	96	4	Same	ReLU
Max Pooling 1	112×112×96	3×3	-	2	-	-
Conv Layer 2	112×112×256	5×5	256	1	Same	ReLU
Max Pooling 2	56×56×256	3×3	-	2	-	-
Conv Layer 3	56×56×384	3×3	384	1	Same	ReLU
Conv Layer 4	56×56×384	3×3	384	1	Same	ReLU
Conv Layer 5	56×56×256	3×3	256	1	Same	ReLU
Max Pooling 3	28×28×256	3×3	-	2	-	-
Flatten	50176	-	-	-	-	-
Fully Connected	4096	-	-	-	-	ReLU
Dropout	4096	-	-	-	-	-
Fully Connected	4096	-	-	-	-	ReLU
Dropout	4096	-	-	-	-	-
Fully Connected	num_classes	-	-	-	-	Softmax

TABLE III. HYPERPARAMETER RANGES

Hyperparameter	Abbreviation	Range
Learning Rate	learning_rates	[0.001, 0.0005, 0.0001]
Batch Size	batch_sizes	[32, 64, 128, 256]
Momentum	momentum	[0.9, 0.95, 0.99]

A. Experimental Setup

The model was trained in a combination of three open-source datasets, as detailed in the Data Collection and Preprocessing section. The dataset was split into 80% for training and 20% for validation. A set of samples from both training and validation is shown in Fig. 2.

To optimize performance, the GA fine-tuned key hyperparameters such as learning rate, batch size, and momentum based on a range of values selected from prior research. The optimization process ran over 50 generations, with a population size of 20 individuals. The training process was conducted using the TensorFlow and Keras frameworks, and the model was trained on an NVIDIA RTX 3070 GPU for accelerated performance.

B. Model Performance

The performance of the proposed CNN, optimized with GA, was thoroughly evaluated on the test set using a variety of performance metrics, including precision, recall, and F1 score. These results are compared with baseline models, and the learning process is further visualized through training and validation loss and training and validation accuracy graphs.

The model demonstrated strong performance across all pollution categories. The macro-average and weighted-average F1-Scores were both 0.97, indicating balanced performance across different air quality levels. The detailed results are summarized in Table IV.

The overall model accuracy was 95.44%, reflecting a significant improvement compared to the baseline CNN models without GA optimization. The results shown in Table V demonstrate that the proposed CNN-GA model significantly outperformed the baseline CNN model without the GA-based optimizer across all performance metrics, achieving a 17.44% increase in accuracy, a 21.00% increase in precision, and a 21.00% increase in recall.

Fig. 2. A set of samples from training and validation.

V. DISCUSSION

A. Training and Validation Curves

The training and validation loss and accuracy curves further demonstrate the robustness of our model. As shown in Fig. 3, both loss and accuracy stabilized after around five epochs, indicating that the model converged quickly without overfitting.

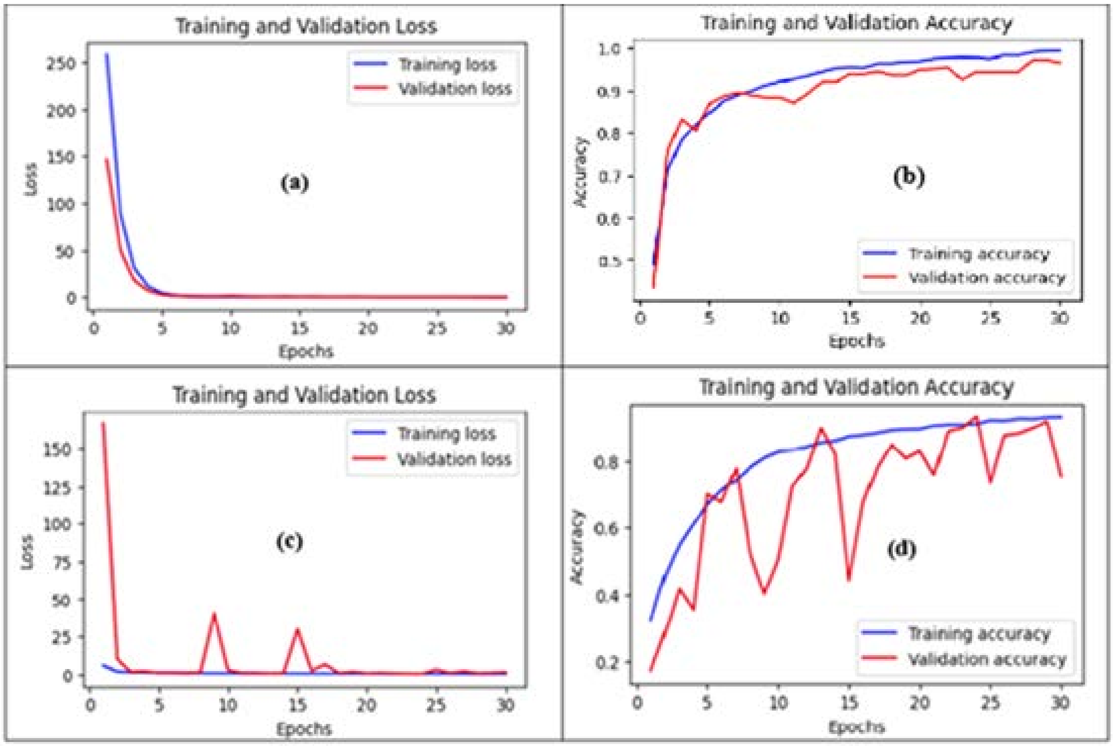

Fig. 3. Proposed model training and validation loss (a), Accuracy (b); Base model training and validation loss (c), Accuracy (d)

TABLE IV. Classification Report for Pollution Levels

Pollution Level	Precision	Recall	F1-Score	Support
Good	0.97	0.96	0.97	348
Moderate	0.97	0.95	0.96	364
Unhealthy	0.95	0.98	0.97	551
Sensitive Groups	0.98	0.95	0.97	581
Very Unhealthy	0.96	0.96	0.96	450
Hazardous	0.97	0.99	0.98	294
Macro Average	0.97	0.97	0.97	2588
Weighted Average	0.97	0.97	0.97	2588
Overall Model Accuracy: 0.9544				

TABLE V. Performance Comparison of Different Models

Model	Accuracy	Precision	Recall	F1-Score
Baseline CNN	0.78	0.76	0.76	0.76
CNN-GA	0.9544	0.97	0.97	0.97

The training accuracy approached near-perfect levels (99%), while the validation accuracy consistently ranged between 95% and 99%, confirming strong generalization capability. In contrast, the base model exhibited noticeable fluctuations in validation accuracy and loss as shown in Fig. 3, with clear signs of overfitting after several epochs, particularly during the later stages of training. Validation loss spiked in certain epochs, while training accuracy continued to improve, indicating that the base model overfit the training data and struggled to generalize to the validation set. This comparison emphasizes the superior generalization ability of the proposed CNN-GA model, as it maintained stable validation performance without significant degradation or divergence from training metrics.

B. *Overfitting Prevention and Generalization*

Several techniques were employed to prevent overfitting and ensure the model generalized well on unseen data. These included a dropout rate of 0.6 to reduce reliance on specific neurons, batch normalization to stabilize training, early stopping to prevent overtraining, and learning rate reduction when validation loss plateaued for finer adjustments. These techniques contributed to the CNN-GA model's ability to maintain a high level of performance across various environmental conditions. The incorporation of GA led to a significant enhancement in the model's performance. The CNN-GA model indeed represented the real improvement to the baseline by several folds along all key performance indicator metrics. Accuracy was increased in CNN-GA by a maximum of 17.44% compared to that proposed by CNN, or, to say precisely, 78% was increased to 95.44%. Precision of the CNN-GA improved by +21 points from that provided by CNN: 0.76 to 0.97; it experienced the very same increase also for recall—by 0.76 to 0.97, also for the F1-Score. The improvement in the results underlines the potential of the proposed GA-based hyperparameter optimization to increase the performance and robustness of air quality estimation from image data, offering a higher generalization ability compared with state-of-the-art methods working with traditional CNN.

VI. Conclusion

This paper presents a new air quality estimation approach using CNN optimized by GA, significantly enhancing predictive accuracy and improving the generalization of the model for a wide range of environmental contexts. The integration of GA within the CNN model allows for dynamic optimization of hyperparameters, which, apart from enhancing performance, may ensure adaptability to diverse spatial, temporal, and environmental conditions. The approach provides a series of limitations with traditional air quality monitoring systems, offering restricted geographic coverage and very high operational costs that cannot provide real-time data.

This work has been done using three different open-source datasets, proving that the model will generalize well for any kind of air quality scenario. These results have been verified using different metrics such as precision, recall, and F1-score, which is considerably better compared to baseline methods; hence, the CNN-GA model is sound and reliable regarding the classification of air quality levels. The scalability of the model at low cost opens a different direction in conducting large-scale monitoring of air quality, which is all-important for protecting public health and the environment.

In future work, we will extend our dataset to more diverse scenes and integrate additional data sources, such as satellite imagery and real-time sensor data, to improve generalization to unseen data.

Acknowledgment

This research is funded by the University Postgraduate Research Grant (PGRS220339), Universiti Malaysia Pahang Al-Sultan Abdullah, Malaysia. The study is also supported by University of Doha for Science and Technology (UDST), Doha, Qatar.

Authors' Contribution

Arshad Ali Khan: Conceptualization, Methodology, Writing-Original draft. Mazlina Abdul Majid: Data curation, reviewing draft, and guidance Abdulhalim Dandoush: reviewing literatures, and proof reading.

References

[1] Satpathy, P., et al. (2024). The Health Menace of Myriad Air Pollutants: An Indian Perspective. In P.K. Padhy, et al. (Eds.), Air Quality and Human Health (pp. 181-202). Springer Nature Singapore https://doi.org/10.1007/978-981-97-1363-9_14

[2] Maji, S., et al. (2023). Health Risks of Major Air Pollutants, their Drivers and Mitigation Strategies: A Review. Air, Soil and Water Research, 16, 11786221231154659. https://doi.org/10.1177/11786221231154659

[3] Organization, W.H. (2021). WHO global air quality guidelines: particulate matter (PM2. 5 and PM10), ozone, nitrogen dioxide, sulfur dioxide and carbon monoxide. World Health Organization. https://apps.who.int/iris/handle/10665/345329.

[4] [4] Hwang, Y., E. Barut, and K. Yeo. (2018). STATISTICAL-PHYSICAL ESTIMATION OF POLLUTION EMISSION. Statistica Sinica, 28, 921-940. http://www.jstor.org/stable/44841931

[5] [5] Krichen, M. (2023). Convolutional Neural Networks: A Survey. Computers, 12, 151. https://doi.org/10.3390/computers12080151

[6] Wu, T.-W., et al. (2023). Applications of convolutional neural networks for intelligent waste identification and recycling: A review. Resources, Conservation and Recycling, 190, 106813. https://doi.org/10.1016/j.resconrec.2022.106813

[7] Wang, Z., F. Wu, and Y. Yang. (2023). Air pollution measurement based on hybrid convolutional neural network with spatial-and-channel attention mechanism. Expert Systems with Applications, 233, 120921. https://doi.org/10.1016/j.eswa.2023.120921

[8] Zhang, C., et al. (2018). End-to-end learning for image-based air quality level estimation. Machine Vision and Applications, 29, 601-615. https://doi.org/10.1007/s00138-018-0919-x

[9] Song, S., et al. (2020). ResNet-LSTM for Real-Time PM2.5 and PM10 Estimation Using Sequential Smartphone Images. IEEE Access, 8, 220069-220082. https://doi.org/10.1109/ACCESS.2020.3042278

[10] Xie, X., et al. (2017). A Review of Urban Air Pollution Monitoring and Exposure Assessment Methods. ISPRS Int. J. Geo-Inf, 6, 389. https://doi.org/10.3390/ijgi6120389

[11] Fattoruso, G., et al. (2020). Site Suitability Analysis for Low Cost Sensor Networks for Urban Spatially Dense Air Pollution Monitoring. Atmosphere, 11, 1215. https://doi.org/10.3390/atmos11111215.

[12] Kumar, P., et al. (2015). The rise of low-cost sensing for managing air pollution in cities. Environment International, 75, 199-205. https://doi.org/10.1016/j.envint.2014.11.019

[13] Li, X., et al. (2017). Long short-term memory neural network for air pollutant concentration predictions: Method development and evaluation. Environmental Pollution, 231, 997-1004. https://doi.org/10.1016/j.envpol.2017.08.114

[14] Zhang, B., et al. (2022). Deep learning for air pollutant concentration prediction: A review. Atmospheric Environment, 290, 119347. https://doi.org/10.1016/j.atmosenv.2022.119347

[15] Zhang, L., et al. (2018). Trend analysis and forecast of PM2.5 in Fuzhou, China using the ARIMA model. Ecological Indicators, 95, 702-710. https://doi.org/10.1016/j.ecolind.2018.08.032

[16] Balachandran, S., et al. (2013). Bayesian-Based Ensemble Source Apportionment of PM2.5. Environmental Science & Technology, 47, 13511-13518. https://doi.org/10.1021/es4020647.

[17] García Nieto, P.J., et al. (2018). PM10 concentration forecasting in the metropolitan area of Oviedo (Northern Spain) using models based on SVM, MLP, VARMA and ARIMA: A case study. Science of The Total Environment, 621, 753-761. https://doi.org/10.1016/j.scitotenv.2017.11.291

[18] Kamal, M.M., R. Jailani, and R.L.A. Shauri.(2006,June). Prediction of Ambient Air Quality Based on Neural Network Technique. 2006 4th Student Conference on Research and Development.Shah Alam, Malaysia.

[19] Antanasijević, D.Z., et al. (2013). PM10 emission forecasting using artificial neural networks and genetic algorithm input variable optimization. Science of The Total Environment, 443, 511-519. https://doi.org/10.1016/j.scitotenv.2012.10.110

[20] Van Roode, S., et al. (2019). An artificial neural network ensemble approach to generate air pollution maps. Environmental Monitoring and Assessment, 191, 727. https://doi.org/10.1007/s10661-019-7901-6

[21] Yan, R., et al. (2021). Multi-hour and multi-site air quality index forecasting in Beijing using CNN, LSTM, CNN-LSTM, and spatiotemporal clustering. Expert Systems with Applications, 169, 114513. https://doi.org/10.1016/j.eswa.2020.114513

[22] Zhang, B., et al. (2020). Constructing a PM2.5 concentration prediction model by combining auto-encoder with Bi-LSTM neural networks. Environmental Modelling & Software, 124, 104600. https://doi.org/10.1016/j.envsoft.2019.104600

[23] [23] Wang, Z., et al. (2019). Air Quality Measurement Based on Double-Channel Convolutional Neural Network Ensemble Learning. IEEE Access, 7, 145067-145081. https://doi.org/10.1109/ACCESS.2019.2945805

[24] Kow, P.-Y., et al. (2022). Real-time image-based air quality estimation by deep learning neural networks. Journal of Environmental Management, 307, 114560. https://doi.org/10.1016/j.jenvman.2022.114560

[25] Zhang, Q., F. Fu, and R. Tian. (2020). A deep learning and image-based model for air quality estimation. Science of The Total Environment, 724, 138178. https://doi.org/10.1016/j.scitotenv.2020.138178

[26] Rouniyar, A., et al. Air Pollution Image Dataset from India and Nepal. https://www.kaggle.com/ds/3152196, 2023 (accessed 18 September,2024).

[27] Wetchayont, P.,Estimated outdoor PM2.5 concentration data by using mobile phone images in Bangkok, Thailand,Mendeley Data,2023 https://data.mendeley.com/datasets/d6g44yftxj.

[28] yunzhenzhang(kingofbabe),Pictures and air quality,Kaggle,2019 https://www.kaggle.com/datasets/yunzhenzhang/pictures-and-air-quality.

[29] Bai, Y., et al. (2019). Hourly PM2.5 concentration forecast using stacked autoencoder model with emphasis on seasonality. Journal of Cleaner Production, 224, 739-750. https://doi.org/10.1016/j.jclepro.2019.03.253

[30] Hamrani, A., A. Akbarzadeh, and C.A. Madramootoo. (2020). Machine learning for predicting greenhouse gas emissions from agricultural soils. Science of The Total Environment, 741, 140338. https://doi.org/10.1016/j.scitotenv.2020.140338

[31] Hatami, N., Y. Gavet, and J. Debayle.(2017). Classification of time-series images using deep convolutional neural networks. Tenth international conference on machine vision.Vienna, Austria.

[32] Kow, P.-Y., et al. (2020). Seamless integration of convolutional and back-propagation neural networks for regional multi-step-ahead PM2.5 forecasting. Journal of Cleaner Production, 261, 121285. https://doi.org/10.1016/j.jclepro.2020.121285

[33] Miao, W., et al.(2020). Efficient and Accurate Classification Enabled by a Lightweight CNN. 2020 5th International Conference on Computer and Communication Systems (ICCCS).Shanghai, China.

[34] Persello, C., et al. (2019). Delineation of agricultural fields in smallholder farms from satellite images using fully convolutional networks and combinatorial grouping. Remote Sensing of Environment, 231, 111253. https://doi.org/10.1016/j.rse.2019.111253

[35] Pyo, J., et al. (2019). A convolutional neural network regression for quantifying cyanobacteria using hyperspectral imagery. Remote Sensing of Environment, 233, 111350. https://doi.org/10.1016/j.rse.2019.111350

[36] Qian, Y., et al. (2020). Coupling cellular automata with area partitioning and spatiotemporal convolution for dynamic land use change simulation. Science of The Total Environment, 722, 137738. https://doi.org/10.1016/j.scitotenv.2020.137738

[37] Wang, Y.-S., L.-C. Chang, and F.-J. Chang. (2021). Explore Regional PM2.5 Features and Compositions Causing Health Effects in Taiwan. Environmental Management, 67, 176-191. https://doi.org/10.1007/s00267-020-01391-5

[38] Yu, S., et al. (2020). Classification of pathogens by Raman spectroscopy combined with generative adversarial networks. Science of The Total Environment, 726, 138477. https://doi.org/10.1016/j.scitotenv.2020.138477

[39] Zhang, C., et al.(2017). Hybrid Measurement of Air Quality as a Mobile Service: An Image Based Approach. 2017 IEEE International Conference on Web Services (ICWS).

[40] A Ilemobayo, J., et al. (2024). Hyperparameter Tuning in Machine Learning: A Comprehensive Review. Journal of Engineering Research, 26, 388-395. https://doi.org/10.9734/jerr/2024/v26i61188

[41] González-Castro, L., et al. (2024). Impact of Hyperparameter Optimization to Enhance Machine Learning Performance: A Case Study on Breast Cancer Recurrence Prediction. Applied Sciences, 14, 5909. https://doi.org/10.3390/app14135909.