P-ISSN 1991-8941
E-ISSN 2706-6703
**Vol. 20, No. 2 (2026)**

Journal of University of Anbar for Pure Science
(JUAPS)

*journal homepage: juaps.uoanbar.edu.iq*

**Open Access**

**Original Paper**

# Image-based air quality classification: superior performance of a VGG16 feature extraction and random forest pipeline

**Nisreen Mohammed Mahmood** [1,2*], **Fatma Ozge Ozkok** [2]

[1]Electronic Computer Center, University Of Anbar, Iraq, Ramadi
[2]Department of Computer Engineering, Erciyes University, Kayseri 38039, Turkey

## ARTICLE INFO

Received: 08/10/2025
Revised: 15/12/2025
Accepted: 16/12/2025
Available online: 08/08/2026
August Issue
10.37652/juaps.2025.165977.1750

CITE @ JUAPS

**Corresponding author**
*Nisreen Mohammed Mahmood*
*eng.nmsw@uoanbar.edu.iq*

## ABSTRACT

Air pollution is a critical threat to environmental sustainability and public health, particularly in densely populated countries such as India and Nepal. This paper presents a hybrid pipeline that combines deep learning feature extraction with traditional machine learning classification for image-based air quality assessment. The method employs a two-stage pipeline: feature extraction using a pre-trained VGG16 CNN and classification using SVM and Random Forest (RF) algorithms. The model breaks down air quality into six levels: Good, Moderate, Unhealthy for Sensitive Groups, Unhealthy, Very Unhealthy, and Severe. A publicly available dataset of photographs taken at multiple urban sites in India and Nepal was used to conduct the test. Experimental results show that the Random Forest classifier achieves 97 percent accuracy, compared to 95 percent for the SVM classifier, on the test data. The effectiveness of the method is confirmed by a large-scale evaluation based on precision, recall, F1-score, and ROC-AUC metrics. The findings indicate a significant potential for combining deep feature extraction with traditional machine learning algorithms to create scalable environmental monitoring systems for real-world applications.

**Keywords:** *Air pollution, Deep learning, Environmental monitoring, Image classification, VGG16*

## 1 INTRODUCTION

Air pollution has become one of the most acute problems of the 21st century, not only in industrialized nations but also in developing countries worldwide. It is estimated that ambient air pollution results in millions of unnecessary deaths annually, while deteriorating the conditions of urban life and the health status of people in a negative manner [1]. Traditional air quality monitoring relies principally on physical sensor networks and human data acquisition procedures and is typically very expensive, labor-intensive, and geographically area-based. New developments in machine learning and computer vision have created new opportunities to address these constraints with image-based air quality monitoring systems [2, 3]. The technologies have the potential to make air quality assessment devices more scalable, economical, and accessible, deployable in areas where traditional sensor infrastructure is not viable or cost-effective. This work demonstrates an efficient hybrid solution combining deep learning-based feature extraction with traditional machine learning classifiers to automatically classify air quality. In contrast to recently proposed, more complex transformer-based learning models [4–6], our solution focuses on efficiency and interpretability while maintaining similar accuracy. The method uses a pre-trained VGG16 convolutional neural network for effective feature extraction from environmental images [7], along with classification Random Forest (RF) algorithms [8, 9]. The capability of trained CNN models to retrieve high-level semantic features of images has been widely demonstrated across a range of computer vision tasks, making them particularly applicable to the environmental monitoring workload.

The deep features are then sent to the standard machine learning classifiers to obtain correct air quality predictions. The two-stage process leverages the representational capabilities of deep learning and the explainability and performance of traditional classifiers to achieve a trade-off solution that can be operationalized. The experimental testing of a representative dataset of images from different urban sites in India and Nepal [10] was conducted using region-specific cross-validation to address the difficult task of visual estimation of air quality across a variety of environmental conditions. These sites have very adverse monitoring environments, with large populations, high rates of urbanization, and high heterogeneity in pollution sources and atmospheric regimes. The method is a great move in the utilization of transfer learning ability in combination with ensemble schemes in the development of dynamic, robust, and accurate models in the case of environmental monitoring-based applications [9, 11]. This paper discusses the ability to estimate air quality using image-based methods by harmoniously combining deep learning-based feature extraction with conventional machine learning-based classification methods. The key takeaways and contributions to this study are that it has been shown that (1) an efficient VGG16+RF pipeline can reach 97% accuracy that is on par with more complex models such as the transformer models while utilizing less computer resources, (2) the effect of utilizing SVM vs. the Random Forest classifier with deep CNN features has been objectively compared and contrasted over air quality classification models with significant results, and (3) that these models are viable in real-world images from the South Asia region.

## 2 RELATED WORKS

Big steps have been made in the field of image-based air quality monitoring, demonstrating that recent methods rely primarily on complex deep learning architectures. Nevertheless, the mentioned methods still face difficulties in real-world scenarios, mainly due to their high computational requirements and low interpretability. This part of the paper provides a thoughtful discussion of recent methods and, more importantly, hints at the gap our hybrid approach aims to address. Lately, researchers have increasingly proposed high-end neural networks for air quality monitoring. Wang et al. [4] achieved an outstanding 98.7% accuracy by applying Vision Transformer (ViT) models with multi-head self-attention for the classification of air quality using a smartphone. However, transformer models are still very computationally expensive for both training and inference, and hence are not suitable for deployment in low-resource settings or real-time mobile apps. On the other hand, pure convolutional methods have been promising due to their accuracy. Mondal et al. [12] ambitiously set the bar higher with their Deep Convolutional Neural Networks (DCNN), achieving 95.2% accuracy on the Dhaka urban dataset (camera images) for PM2.5 estimation. However, one more thing these models lack is that they are entirely based on deep learning and do not provide any explanation for the quality of air they predict, which is a critical requirement for environmental monitoring involving different stakeholders. The researchers have also attempted to combine various deep learning components into a single structure. Duan et al. [5] developed an innovative ARIMA-CNN-LSTM pipeline integrated with the Dung Beetle Optimizer and achieved 94.2% accuracy in air quality prediction using sensor time-series data. Nguyen et al. introduced Quantum Particle Swarm Optimization (QPS) [6].

## 3 METHODOLOGY

The methodology proposed here is based on a two-phase hybrid approach that combines the representational power of deep convolutional neural networks with the stability of traditional machine learning classifiers for air quality classification from visual images. Figure 1 illustrates the overall system architecture, with the following step-by-step flow from raw image input to final classification output. The proposed approach employs a two-stage hybrid approach that combines the representational power of deep convolutional neural networks with that of traditional machine learning classifiers for air quality assessment from visual images. The system design as a whole is illustrated in Figure 1, where it indicates step-wise movement from raw image input to VGG16 feature extraction prior to final classification output through both SVM and Random Forest techniques.

### 3.1 Dataset description and preprocessing

The empirical study employs a large, open-access dataset of air pollution images collected from diverse urban locations in India and Nepal [10]. The dataset comprises 12,240 well-annotated 224×224-pixel photos, carefully classified according to the Air Quality Index (AQI) criteria. The images are evenly distributed across the six air quality classes: Good (2,040 images), Moderate

(2,040 images), Unhealthy for Sensitive Groups (2,040 images), Unhealthy (2,040 images), Very Unhealthy (2,040 images), and Severe (2,040 images). The distribution provides equal representation across all pollution severity levels and yields sufficient samples for stable model training and testing. The dataset distribution and data preprocessing procedure are illustrated in Figure 2, which shows the categorical distribution of six levels of air quality and the systematic data preparation procedure. The depicted preprocessing pipeline in Figure 2 shows a sequential process from raw image input through resizing, normalization, data augmentation, and train/test split procedures.

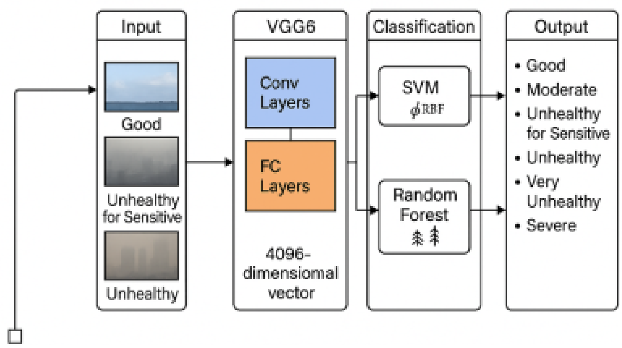


**Fig. 1** System Architecture Overview

The images were captured under various environmental conditions like varying times of day, weather, and seasonal variations, providing a holistic view of real-world monitoring scenarios. Data preprocessing includes normalization of image sizes, normalization of pixel values to [0,1], and data augmentation via rotation (±15°), horizontal flipping, and brightness adjustment (±20%) to enhance model generalization performance. The stratified sampling technique is used to ensure that all classes are represented in proportion, and the data is then split into 80% for training (9,792 images) and 20% for testing (2,448 images). This split method ensures homogeneous evaluation measures and eliminates bias in model evaluation caused by class imbalance.

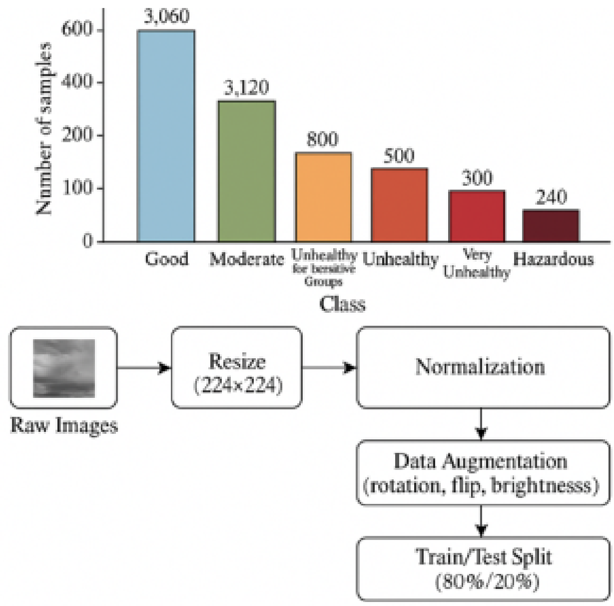


**Fig. 2** Dataset Distribution and Preprocessing Pipeline

### 3.2 VGG16-Based feature extraction architecture

The feature extraction phase uses the VGG16 convolutional neural network, an established network comprising 16 trainable units: 13 convolutional and 3 fully connected [7]. The ImageNet dataset, with more than 14 million images across 1,000 classes, was pre-trained on VGG16, which demonstrated strong feature representation of various visual patterns. The architecture uses small (3x3) convolutional filters across every unit and can generate fine-grained spatial information efficiently. In this paper, we use the pre-trained VGG16 network without its last classification layers, and we extract features at the second fully connected layer (FC2). The method provides a feature vector of 4096 dimensions for every input image, which is mathematically expressed as follows:

$$F = VGG16_{\{FC2\}(I)} \tag{1}$$

Here $F \in R$^4096 represents the extracted feature vector, and I denotes the input image of dimensions 224×224×3. The choice of the FC2 layer for feature extraction is based on its ability to capture semantic details while retaining sufficient discriminability to perform the sorting task. The resulting features are then L2 normalized to ensure that all dimensions are consistently scaled:

$$F_{\text{norm}} = \frac{F}{\|F\|_2} \tag{2}$$

This normalization increases the robustness of the classification process by making a feature less influential when its value is large.

### 3.3 Classification algorithms

In the classification phase, two machine learning methods are used to assess their performance using deep CNN features. These methods were chosen because they are good at handling high-dimensional data and have performed well in past classification work.

**Support Vector Machine setup:** The SVM classifier uses a Radial Basis Function (RBF) kernel to handle non-linear relationships in the feature space. The RBF kernel job is defined as:

$$K\left(x\_i, x\_j\right) = \exp\left(-\gamma \left\|x\_i - x\_j\right\|\hat{}\ 2\right) \quad (3)$$

The kernel coefficient is represented by $\gamma$, and $\|x\_i - x\_j\|$ is the Euclidean distance between feature vectors. The multi-class SVM classification can be written as:

minimize:

$$\text{minimize} \quad \left(\frac{1}{2}\right)\|w\|\hat{}2 + C\sum_{2} \quad (4)$$

$$\text{subject to:} \quad y\_i\left(w\hat{}T\varphi\left(x\_i\right) + b\right) \geq 1 - \xi_i, \xi_i i \geq 0 \quad (5)$$

In this model, C denotes the regularization parameter, $\xi$ denotes the slack variables, and $\varphi$ (x_i) denotes the kernel mapping function. Grid search with 5-fold cross-validation is used for hyperparameter tuning. The C values in the range $[0.1, 1, 10, 100]$ and $\gamma$ values in $[0.001, 0.01, 0.1, 1]$ are assessed.

**Random Forest Setup:** The Random Forest algorithm builds a group of decision trees by using bootstrap aggregating (bagging) and random feature selection. For each tree $T_k$ in the forest, the prediction is figured as:

$$T\_k(x) = \text{argmax}\_\ c\sum_{-}\{i \in L\_k\}\, I\left(y\_i = c\right) \quad (6)$$

The leaf node that contains the input $x$ is denoted as $L\_k$, and I stands for the indicator function. The ensemble makes its final prediction by using majority voting to combine the outputs from each tree:

$$RF(x) = \text{argmax}\_\ c\Sigma\_\{k = 1\}\hat{}KI\left(T_{k(x)} = c\right) \quad (7)$$

Random Forest employs 100 decision trees, using Gini impurity as the splitting criterion. The calculation of Gini impurity for a node t is as follows:

$$\text{Gini}(t) = 1 - \Sigma\_\{t = 0\}\hat{}\{C - 1\}p\_i\hat{}2 \quad (8)$$

In this case,$p_i$ is the proportion of samples for class i at node t, and C is the total number of classes. Cross-validation is used to determine the optimal tree depth, thereby avoiding overfitting and maintaining high classification accuracy.

### 3.4 Implementation details

The entire procedure was carried out in Python 3.8, using TensorFlow 2.10 for VGG16 feature extraction, scikit-learn 1.1.2 for SVM and Random Forest implementations, and NumPy 1.23 for array operations. The feature extraction and model training were performed on a desktop computer with NVIDIA GeForce RTX 3080 GPU (10GB VRAM), Intel Core i7-11700K CPU, and 32GB RAM. Pre-trained ImageNet weights were used, and no fine-tuning was performed on the VGG16 network. The SVM classifier used the following optimized hyperparameters, identified by 5-fold cross-validation: C = 10 (regularization parameter) and $\gamma = 0.01$ (RBF kernel coefficient). The Random Forest classifier was constructed with 100 decision trees with a maximum depth of 20, minimum samples for splitting of 10 , and minimum samples per leaf of 4 . The Gini impurity criterion was applied for node splitting. The models were trained on a dataset carefully split into 14,688 training images and 3,672 for testing. Training time varied by method: about 45 minutes for the Support Vector Machine versus roughly 12 minutes for the Random Forest. This underlines the efficiency of the whole pipeline.

### 3.5 model training and evaluation protocol

Feature extraction is the starting point of training. We apply a pre-trained VGG16 network to the entire dataset to extract features. Each image produces a 4096-dimensional feature vector, which is fed to two classifiers: an SVM and a Random Forest. We evaluate both models using standard metrics: accuracy, precision, recall, F1-score, and AUC calculated from the confusion matrix for each class:

$$\text{Precision}_i = \frac{TP_i}{TP_i + FP_i} \quad (9)$$

$$Recall_i = \frac{TP_i}{TP_i + FN_i} \tag{10}$$

$$F1 - \text{ score }_i = 2 \times \frac{\text{Precision }_i \times \text{ Recall }_i}{\text{Precision }_i + \text{ Recall }_i} \tag{11}$$

Overall accuracy is the number of correct classifications divided by the total number of instances. Scikit-learn is used for the algorithms and judging performance, so the results can be repeated, and the metrics are standard.

## 4 RESULTS AND DISCUSSION

### 4.1 Experimental results analysis

The proposed hybrid pipeline achieved strong classification results across the two machine learning algorithms tested. The Support Vector Machine classifier achieved an overall accuracy of 95% on the test set, comprising 3,672 images, and correctly classified 3,487 instances. On the other hand, the Random Forest classifier set up a new record high with 97% overall accuracy and 3,558 instances successfully classified, an improvement of 71 instances over SVM.

#### 4.1.1 Performance by air quality category

The SVM classifier is shown in Figure 3 with the confusion matrix. The highest precision achieved by the classifier was 98.6% (428/434 correct) for the "Severe" category, followed by "Moderate" (96.4%, 455/472 correct) and "Good" (96.3%, 445/462 correct). The "Unhealthy for Sensitive Groups" category has the largest absolute number of misclassifications, with 64 errors out of 859 samples (92.5% accuracy), suggesting that this intermediate pollution level is the most difficult to classify.

The Random Forest classifier was much better than the SVM implementation, as evident from Figure 5. The Random Forest confusion matrix in Figure 5 always has higher diagonal values across all classes, indicating greater classification accuracy. As can be seen from Figure 5, the "Good" class performed 446 correct predictions out of 462 samples, which is a very slight improvement over the results of SVM in Figure 3. As can be seen from Figure 5, the "Moderate" class also performed 447 correct predictions out of 472 samples, while the "Unhealthy for Sensitive Groups" class performed very well with 823 correct predictions out of 859 total samples, which is a very notable improvement when compared to SVM's performance of 795 correct predictions in Figure 3. Figue 5 shows that the "Unhealthy" class performed 772 correct predictions out of 787 samples, "Very Unhealthy" performed 643 correct predictions out of 658 samples, and the "Severe" class performed very uniformly with 427 correct predictions out of 434 total samples. Figure 4 shows the multi-class ROC curves, indicating strong discriminative ability across all air quality categories. Five out of the six categories received perfect AUC scores of 1.00: "Good," "Moderate," "Unhealthy," "Very Unhealthy," and "Severe." The "Unhealthy for Sensitive Groups" category achieved an AUC of 0.99, indicating almost perfect separation of the classes. The close-to-perfect AUC scores should be critically evaluated. The high discriminative performance indicates that the features extracted by VGG16 form very well-separated clusters in the feature space across different levels of pollution. However, the slight drop in AUC for "Unhealthy for Sensitive Groups" (0.99 vs. 1.00) supports our confusion matrix findings that this intermediate category shows greater classification ambiguity than the extreme categories. The best possible area under the Receiver Operating Characteristic curve for the "Good" and "Severe" air quality categories, along with the middle categories, suggests that the visual information related to air quality is adequately captured by the VGG16-extracted features. This information is very likely to include the effect of atmospheric haze, color contrasts, and reduced visibility, all of which are functions of the air quality levels. The high separability of the classes again justifies the choice of VGG16 to extract features, and a simple classifier, such as Random Forest, achieves approximately 97% classification accuracy in the feature space.

Figure 5 indicates that Random Forest significantly reduced misclassification rates across all categories. The "Unhealthy for Sensitive Groups" class saw the greatest improvement, with 95.8% accuracy (823/859 correct), compared with SVM's 92.5%—a 3.3-point increase, indicating 28 fewer errors. Random Forest also gained on "Unhealthy" samples (98.1% vs. 95.2%) and "Very Unhealthy" samples (97.7% vs. 93.5%).

#### 4.1.2 Comparative analysis with state-of-the-art methods

To give a clear picture of our methodology's performance, we conducted a direct comparison with the latest cutting-edge methods limited to the domain of image-based air quality classification or estimation; time-series prediction studies that rely on sensor data were not included in this comparison, as these are fundamentally different tasks. An overview of our performance comparison with five recent works for the years 2022-2024 is presented in Table 1, with emphasis on the hybrid approach's competitive performance.


**Fig. 3** SVM Confusion Matrix Results

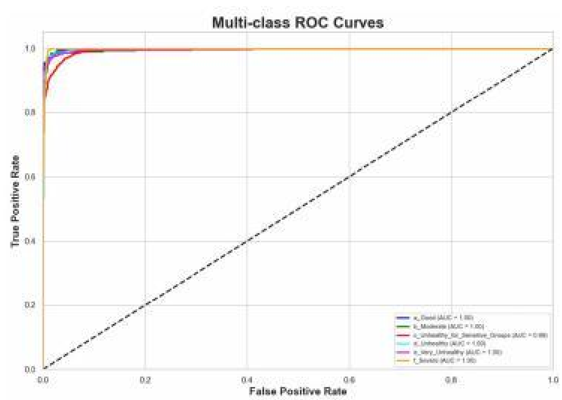


**Fig. 4** Air Quality Multi-class ROC Curves

Studies [5] and [6] mentioned in Table 1 deal with temporal prediction from sensor data rather than direct image classification, and thus are only cited for informational purposes. The comparison should be limited to image-based approaches [4], [12], and where our method exhibits a competitive 97% accuracy along with considerable advantages in computational efficiency and interpretability due to the feature importance analysis of the Random Forest classifier. The difference of 1.7 percentage points from Wang et al.'s DOViT [4] (97% vs. 98.7%) is more than offset by our method's significantly lower computational requirements and faster training time, making it more suitable for deployment in resource-constrained scenarios.

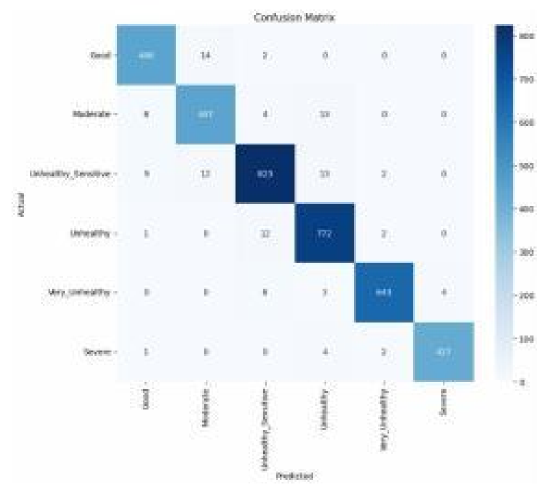


**Fig. 5** Random Forest Confusion Matrix Results

## 4.2 Discussion

Random Forest shows a relative improvement of 2% in accuracy over the SVM model (97% versus 95%), suggesting that the ensemble learning approach is more effective in the high-dimensional feature space of VGG16 images (4,096 dimensions). This might be due to its ability to handle feature redundancy and non-linear decision boundaries through the combination of decision trees in the Random Forest ensemble learning technique. Unlike the traditional SVM model, which optimizes a hyperplane to maximize separation between classes, in Random Forests, several weak decision trees are combined to capture complex relationships in deep image features. The ensemble model clearly has an edge in identifying subtle variations in the visual characteristics representing pollution levels, including changes in air opacity, color, and textures in the area of the sky.

At the same time, the fact that the Random Forest model has the capability to make decisions by incorporating the predictions made by the numerous trees in the model contributes to the robustness of the model regarding the presence of various lighting and weather conditions in the images. The bootstrap aggregation process in the model has the effect of performing feature selection, focusing the model on the important aspects of the VGG16 feature representations and ignoring the less

**Table 1** Performance Comparison with Recent State-of-the-Art Methods

| Study | Year | Method | Dataset | Accuracy (%) | Key Innovation |
|---|---|---|---|---|---|
| Wang et al. [4] | 2022 | Double Output Vision Transformer (DOViT) | Images of a mobile device | 98.7 | Multi-head self-attention with dual outputs |
| Duan et al. [5] | 2023 | ARIMA-CNN-LSTM with Dung Beetle Optimizer | Air quality data from multiple cities | 94.2 | Hybrid linear-nonlinear modeling |
| Nguyen et al. [6] | 2024 | Quantum PSO-enhanced Attention CNN-LSTM | Seoul air quality record | 97.3 | Quantum-improved optimization |
| Mondal et al. [12] | 2024 | Smartphone image-based deep CNN | Dhaka urban dataset | 95.2 | Real-time mobile deployment |
| Rahman et al. [13] | 2024 | Ensemble AirNet ML | Multi-pollutant dataset | 96.1 | Web-based prediction system |
| Our Method | 2026 | VGG16 + Random Forest | India-Nepal urban images | 97.0 | Hybrid deep-traditional approach |

informative features, an important characteristic given the fact that the images taken in environmental monitoring are under uncontrolled conditions.

The Unhealthy for Sensitive Groups category achieved 95.8% accuracy with the Random Forest algorithm, outperforming the SVM's 92.5%. This slight increase in performance translates to 28 fewer instances of misclassification, which is important in public health because the margin plays a crucial role in issuing health warnings to groups susceptible to these health issues. The middle levels of pollution are most difficult to distinguish because there are minute graphical variations, while the distinguishing features include the highest and lowest pollution levels, which are easily recognizable because the lowest levels are significantly lower, and the highest levels are significantly higher, while the middle levels are quite similar and require careful assessment.

The confusion matrices suggest that misclassifications occur mostly between neighboring levels of pollutants rather than between levels that are quite apart, and thus, the results are quite reasonable and understandable given the difficulty of classification. For example, images incorrectly assigned to the Unhealthy for Sensitive Groups category mostly represented the Moderate and Unhealthy levels, which appear similar and are on a continuum.

The high AUC scores (exclusively 1.00 and once 0.99) clearly show the efficiency of class separation via VGG16 feature extraction. This, in turn, shows the effectiveness of the deep feature in extracting pollution indicators using the ImageNet-trained feature, since the feature was extracted from a completely different domain than the current application.

The minor decrease in accuracy in designating Unhealthy for Sensitive Groups (AUC = 0.99) reflects the challenge of precisely identifying the moderate level of pollution. This group marks the zone where air quality changes dramatically due to weather conditions such as humidity, wind, and temperature inversions, which affect visibility but are difficult to distinguish in a stationary photograph.

The confusion matrix confirms the observations and reveals incorrect predictions to the Moderate and Unhealthy groups.

Despite the encouraging results, the following caveats need to be taken into account. Firstly, the results are based on a set of samples collected in India and Nepal, and generalizability to other geographies with differing atmospheric conditions and pollution characteristics may be an issue. Secondly, the results are based solely on daylight images, so pollution at night cannot be evaluated, as artificial lighting and changes in atmospheric light scattering may be present. Thirdly, the discretization of a continuous process may introduce boundary effects. The work also hypothesizes that evaluating imagery alone is sufficient to determine air quality, discounting meteorological, trend, and spatial information, which might improve classification accuracy. The very high AUC values might be due to the considerable difference between the classes and the low variability within them, thus not fully accounting for the complexities of the real-world scenario.

## 5 CONCLUSION

This work demonstrates that a hybrid approach combining deep feature extraction with VGG16 and a traditional machine learning classifier can reliably classify air quality from images. On a dataset of 12,240 real-world pictures of India and Nepal, the model achieved a Random Forest accuracy of 97%, slightly higher than an SVM at 95%. In addition, it matched complex, modern methods in performance while being computationally lightweight and accessible.

The strength of the method lies in its two-step design: it leverages pre-trained deep features to capture rich visual information, then uses an efficient ensemble classifier for the final predictions. The near-perfect AUC scores across all pollution categories (0.99–1.00) indicate clear class separability. The approach excelled in extreme conditions—both Good and Severe—crucial for informing public health decisions. Some challenges emerged with the mid-range pollution categories, underscoring the difficulty of visually judging air quality and suggesting this is not a fundamental limitation of the model.

Increasing geographical diversity of the data, incorporating weather-related cues with visual features, exploring lighter designs suitable for mobile deployment, and conducting long-term field studies are some of the ways to further validate the robustness of the method across various real-world.

### Funding source

No funds received.

### Data availability

N/A

## DECLARATIONS

### Conflict of interest

The authors declare no competing interests

### Consent to publish

N/A

### Ethical approval

N/A

## REFERENCES

[1] World Health Organization. Ambient (Outdoor) Air Pollution; 2024. Fact sheet; accessed 6 August 2026. Available from: https://www.who.int/news-room/fact-sheets/detail/ambient-(outdoor)-air-quality-and-health

[2] Wang J, Li X, Jin L, Li J, Sun Q, Wang H. An Air Quality Index Prediction Model Based on CNN-ILSTM. Scientific Reports. 2022;12(1):8373. 10.1038/s41598-022-12355-6

[3] Bekkar A, Hssina B, Douzi S, Douzi K. Air-Pollution Prediction in Smart City, Deep Learning Approach. Journal of Big Data. 2021;8(1):161. 10.1186/s40537-021-00548-1

[4] Wang Z, Yang Y, Yue S. Air Quality Classification and Measurement Based on Double Output Vision Transformer. IEEE Internet of Things Journal. 2022;9(21):20975-84. 10.1109/JIOT.2022.3176126

[5] Duan J, Gong Y, Luo J, Zhao Z. Air-Quality Prediction Based on the ARIMA-CNN-LSTM Combination Model Optimized by Dung Beetle Optimizer. Scientific Reports. 2023;13(1):12127. 10.1038/s41598-023-36620-4

[6] Nguyen AT, Pham DH, Oo BL, Ahn Y, Lim BTH. Predicting Air Quality Index Using Attention Hybrid Deep Learning and Quantum-Inspired Particle Swarm Optimization. Journal of Big Data. 2024;11(1):71. 10.1186/s40537-024-00926-5

[7] Simonyan K, Zisserman A. Very Deep Convolutional Networks for Large-Scale Image Recognition; 2014. ArXiv preprint. Available from: https://arxiv.org/abs/1409.1556. 10.48550/arXiv.1409.1556

[8] Breiman L. Random Forests. Machine Learning. 2001;45(1):5-32. 10.1023/A:1010933404324

[9] Lin CY, Chang YS, Abimannan S. Ensemble Multifeatured Deep Learning Models for Air Quality Forecasting. Atmospheric Pollution Research. 2021;12(5):101045. 10.1016/j.apr.2021.03.008

[10] Rouniyar A, Utomo S, John A, Hsiung PA. Air Pollution Image Dataset from India and Nepal. Kaggle; 2023. Data set. Available from: https://www.kaggle.com/datasets/adarshrouniyar/air-pollution-image-dataset-from-india-and-nepal. 10.34740/KAGGLE/DS/3152196

[11] Ahmed M, Zhang X, Shen Y, Ali N, Flah A, Kanan M, et al. A Deep Transfer Learning Based Convolution Neural Network Framework for Air Temperature Classification Using Human Clothing Images. Scientific Reports. 2024;14(1):31658. 10.1038/s41598-024-80657-y

[12] Mondal JJ, Islam MF, Islam R, Rhidi NK, Newaz S, Manab MA, et al. Uncovering Local Aggregated Air Quality Index with Smartphone Captured Images Leveraging Efficient Deep Convolutional Neural Network. Scientific Reports. 2024;14(1):1627. 10.1038/s41598-023-51015-1

[13] Rahman MM, Nayeem MEH, Ahmed MS, Tanha KA, Sakib MSA, Uddin KMM, et al. AirNet: Predictive Machine Learning Model for Air Quality Forecasting Using Web Interface. Environmental Systems Research. 2024;13(1):44. 10.1186/s40068-024-00378-z

How to cite this article

Mahmood NM, Ozkok FO. Image-Based Air Quality Classification: Superior Performance of a VGG16 Feature Extraction and Random Forest Pipeline . Journal of University of Anbar for Pure Science. 2026; 20(2):367-375. doi:10.37652/juaps.2025.165977.1750