

## Article

# A Framework for the Estimation of Air Quality by Applying Meteorological Images: Colours-of-the-Wind (COLD)

Kenneth Chan , Paul Matthews and Kamran Munir 

Computer Science Research Centre (CSRC), School of Computing and Creative Technologies, College of Arts, Technology and Environment (CATE), University of the West of England (UWE), Bristol BS16 IQY, UK; paul2.matthews@uwe.ac.uk (P.M.); kamran2.munir@uwe.ac.uk (K.M.)

\* Correspondence: kenneth3.chan@live.uwe.ac.uk

**Abstract:** This paper presents a new framework, “colours-of-the-wind” (COLD), which is designed to estimate air quality based on images from meteorological cameras, data analytics techniques, and the application of deep learning. Existing air quality estimation systems mainly rely on physical monitoring stations, which are limited by the monitoring stations’ physical constraints. Instead of collecting data from environmental monitoring stations, COLD collects air quality data from meteorological cameras. This approach can collect data from any location where a camera could capture a reliable image, which is otherwise not collectable by the physical environmental monitoring station(s). This approach can also avoid bias due to the location of data collection. The system is evaluated by building a prototype based on the Air Quality Health Index from the Hong Kong Observatory. This is one of the unique data-analytic frameworks based on such air quality measurement. The COLD’s air quality estimation is also based on AQHI, the first machine learning-based estimation framework that generates AQHI as the proposed output. Experimental results suggest that the approach adopted by the COLD prototype is feasible and has some promising outcomes. The results also suggest possible parameters for the CNN model used for the training and analyses of the images.

**Keywords:** air pollution; machine learning-based estimation; digital camera; image analysis; air



**Citation:** Chan, K.; Matthews, P.; Munir, K. A Framework for the Estimation of Air Quality by Applying Meteorological Images: Colours-of-the-Wind (COLD). *Environments* **2023**, *10*, 218. <https://doi.org/10.3390/environments10120218>

Academic Editors: William A. Anderson, Ching-Yuan Chang and Sergio Ulgiati

Received: 28 August 2023  
Revised: 16 November 2023  
Accepted: 26 November 2023  
Published: 11 December 2023



**Copyright:** © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Forecasting air quality in an urban area is a well-known research problem. In this regard, a classic approach to this problem is to apply conventional dynamic models to forecast the concentration of urban air pollutants. Recently, various data analytics methodologies have been proposed to predict air pollutants’ density, such that the air quality of a selected area is forecasted using machine learning algorithms (for example: [1–3]); deep learning algorithms (for example: [4–6]); or a combination of multiple algorithms (for example [3] etc.). Such approaches for air quality forecasting rely heavily upon collecting pollutant concentrations at monitoring stations. This method does work, but it has two fundamental issues: (1) physical constraint of a monitoring station and (2) data availability.

Firstly, every monitoring station has some kind of physical limitation (e.g., a concentration of air pollutants could only be collected at a close distance from the monitoring station, and data could not be collected from a long distance). By applying such methods, air quality can only be estimated as far as the visibility goes. The number of air quality monitoring stations is also limited. Nevertheless, monitoring stations cannot be set up at difficult sites, for example, in water-covered areas and locations with no regular power supply.

Secondly, data availability from a monitoring station is mostly limited. Even though a monitoring station is published to the public, there is a limit on refreshing frequency. For example, in Hong Kong, datasets are collected for the COLD prototype. Most air quality data is publicly available, and datasets containing extensive air pollutant measurement data are available. However, the data refreshes hourly. Therefore, the amount of the data is limited to only 24 daily updates.

In order to address these limitations, we propose a novel framework for air quality estimation, named “colours-of-the-wind” (COLD). Instead of relying on extra/other detecting devices [4], COLD has been designed to predict air quality using images from cameras and to provide a short-term estimation of air quality instead of relying on physical monitoring stations, which are very limited due to the physical constraints of monitoring stations. COLD gathers air quality data from meteorological cameras by applying a Convolution Neural Network. This approach can collect data that are otherwise not collectable by the physical monitoring stations and can work for any location wherever images can be captured from a camera.

This paper reports on the image collector and data processing components of the COLD framework. The training data used in COLD are publicly available air quality data sets obtained from the Environmental Protection Department of Hong Kong and images from various locations in Hong Kong collected by meteorological cameras published by the Hong Kong Observatory. Images from the Hong Kong observatory cameras are gathered by web crawling techniques, which are then analysed using a convolution neural network (CNN) algorithm to estimate the air quality of a specific location as a data source for other processes and estimation.

In terms of the methodology applied in this study, first, the impact of various settings on the accuracy is determined that contributes to the data collection, such as the number of possible air quality categories, the effect of daytime and night-time, resolution, and the number of convolutional layers are also tested to ensure proper parameters are implemented. The outcome is then further applied to the image collected from thirteen similar stations, and the accuracy of the analyser has been evaluated.

The rest of the paper is organised as follows: Section 2 summarises and compares existing related work. Section 3 describes the approach taken to accomplish the CNN estimation and introduces the basic architecture of COLD. Section 4 discusses the data collected and the realisation of the COLD architecture using various datasets from Hong Kong. Finally, the last section of this paper presents a discussion and evaluation of the outcomes of the experiments, followed by a conclusion and potential future work.

## 2. Related Literature

One of the notable studies related to data analytics on urban air quality was investigated by Zheng et al. (2013) [5]. Zheng identified a few features affecting air quality in urban areas and performed big data analytics using data sets collected from existing monitoring stations, crowd and participatory sensing, etc.. One of the key features is the methodology based on public and participatory sensing; in Zheng’s case, GPS-equipped taxis can be regarded as mobile sensors probing the travel speed on roads as a data source [5]. However, this approach would not collect air quality for a remote location inaccessible by roads, such as in the middle of a forest or a vast water area. Moreover, data collected by this approach is limited to where cabs are demanded; the data collected focuses primarily on the central business area or locations with significant traffic demand, and the data collected would be biased. Finally, as the cab is in the middle of a street most of the time, there could be a high chance that the traffic heavily influences the collected data and has a possible bias in the dataset. Therefore, an alternative way to collect data that is not biased due to passengers’ demands would improve the quality of the data to be collected.

A case study was performed in 2015 and reported [6] on predicting the concentration of air pollutants and comparing multiple linear regression models, extreme learning machines and feedforward neural networks based on backpropagation. In this study, two air quality monitoring stations on Hong Kong data sets observed by the Hong Kong Observatory (HKO) and Environmental Protection Department (EPD) from 2010 to 2015 were used to evaluate the accuracy of the above-mentioned statistical techniques. Similarly, meteorological parameters were collected daily from 2010 to 2015 [6]. It has been concluded that there are no significant differences between the estimation accuracies of each model. In the same study, extreme learning machines provided a marginally better performance on indicators

related to the goodness of the estimation, such as R2 and RMSE, etc. [6]. This study could only use data collected by two monitoring stations. The data might be biased due to the lack of coverage; it is reasonable to suggest that collecting more information would improve the range of data. The collected data would be more complete, and therefore, an alternative approach to collecting data is recommended.

Wong et al. developed a unique image processing technique for enhancing the capability of Internet video surveillance (IVS) cameras for real-time air quality monitoring [7]. An algorithm based on the fundamental optical theory includes light absorption, scattering, and reflection. Wong et al. further improved their study by combining surveillance cameras with advanced land observing satellite (ALOS) images to map the air quality concentration of the study area [8]. With the advance of machine learning technologies, new approaches have been proposed to study image attributes in recent years. For example, Li et al. [9] proposed estimating haze levels based on colour and transmission information extracted from haze images.

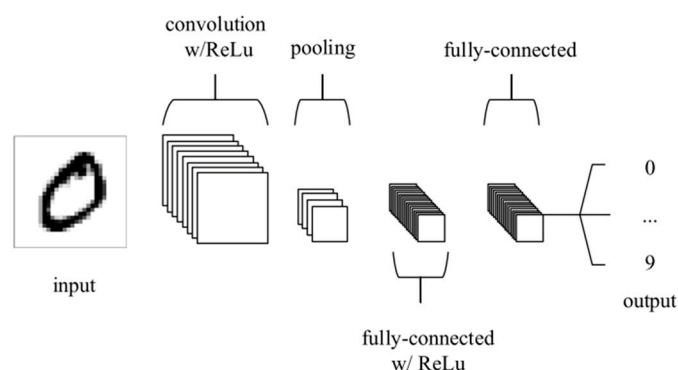
Chakma et al. use a deep convolutional neural network (CNN) to classify raw images into different categories based on their PM<sub>2.5</sub> concentrations [10]. In this study, 8 convolutional layers were adopted, with 2364 images from 3 classes tested, with an average accuracy of 68.74%. The experimental results demonstrate that the CNN approach could estimate image-based PM<sub>2.5</sub> concentration. Liu et al. proposed a method based on support vector regression and comparing similar images with different weather conditions [11] before predicting the value of PM<sub>2.5</sub> by analysing weather conditions such as the colour of the sky and the sun's location during the image taken [12]. Liu estimated the PM<sub>2.5</sub> using clear and cloudy photos from a few Beijing, Shanghai, and Phoenix regions. The proposed method performs well for the images from Beijing and Shanghai but fails for those from Phoenix due to the narrow range of PM<sub>2.5</sub> values. These approaches suggested that analysing images is feasible for collecting air quality-related data.

### 3. Theoretical Background

Theoretically, convolutional neural networks (CNNs) are analogous to traditional ANNs, comprising neurons that self-optimize through learning. From the input raw image vectors to the final output of the class score, the entire network adopts a single perceptive score function (the weight). For colour images, the image would be analysed as a three-dimensional array. The image's red, green, and blue (RGB) components would be treated as three two-dimensional arrays [13]. In practice, this would mean that the input will have a dimensionality of  $n \times n \times 3$  (height, width and colour), leading to a final output layer comprised of a one-dimensional array where  $n$  represents the possible number of classes [14].

CNNs are comprised of three types of layers. These are convolutional layers, pooling layers and fully-connected layers. Figure 1 illustrates the application of the layers to CNN. A convolutional layer applies filters ( $n \times n$  matrix) to each channel to extract characteristics from the input layer. It operates by dividing the input image into smaller pieces of images and convolving them with a specific set of weights [15]. The dot product between the input picture and the filter is computed by multiplying the corresponding values and adding the results to obtain a single scalar value [16].

The pooling layer will then perform down-sampling along the spatial dimensionality of the given input, further reducing the number of parameters. In the image processing domain, it can be considered similar to reducing the resolution. Pooling does not affect the number of filters, and down-sampling does not preserve the position of the information. Therefore, it should be applied only when the presence of information, rather than spatial information, is important [17].



**Figure 1.** A simple CNN architecture [14].

The fully connected layers will then perform the same duties found in standard ANNs and attempt to produce class scores from the activations for classification. This simple transformation method allows CNNs to transform the original input layer-by-layer using convolutional and down-sampling techniques. This will produce class scores for classification and regression purposes.

The filter slides over the input vector to generate the feature map by extracting  $n$  number of features from the input image in a single layer representing distinct features, leading to  $n$  filters and  $n$  feature maps [18].

A convolutional layer applies filters ( $n \times n$  matrix) to the input layer and across the spatial dimensionality of the input to produce a 2D activation map to extract characteristics from the input data [14]. The dot product between the input picture and the filter is computed by multiplying the corresponding values and adding the results to obtain a single scalar value. The convolution for one pixel is calculated according to the formula [19]:

$$G[m, n] = (f * h)[m, n] = \sum_j \sum_k h[j, k] f[m - j, n - k] \quad (1)$$

where  $G[m, n]$  is the output in the next layer,  $f$  is the input image, and  $h$  is the filter matrix and is convolution operation [17].

The dot product of each corresponding value would then be stored as an activation map, which is the output of the convolutional layer. The procedure is repeated until no more sliding is feasible [16]. As the filter could represent a feature, this method could be used to discover the most effective object–description filters.

For the size of the output of the convolution network, it is known that for an image with size  $i \times i$  and a filter with size  $f \times f$ . The size of the output matrix is  $(i - f + 1), (i - f + 1)$  [16].

Every kernel will have a corresponding activation map stacked along the depth dimension to form the full output volume from the convolutional layer. This enables every neuron in a convolutional layer to be only connected to a small region of the input volume. The dimensionality of this region is commonly referred to as the receptive field size of the neuron. The magnitude of the connectivity through the depth is nearly always equal to the depth of the input [14].

The pooling layer aims to gradually reduce the dimensionality of the representation and thus further reduce the number of parameters and the computational complexity of the model.

The pooling layer operates over each activation map in the input and scales its dimensionality [14]. The pooling layer decreases the feature maps' dimensionality while retaining the most important data. This is utilised to reduce the upper layers' complexity by applying the pooling function to the input data. At various pooling levels, various pooling techniques may be applied. The pooling layer in CNN consolidates the feature extracted by the convolution layer and reduces the number of parameters to the next layer [16].

In the CNN architecture, activation functions are used after pooling layers to generate output to be changed or terminated. This layer is used to restrict or oversaturate the

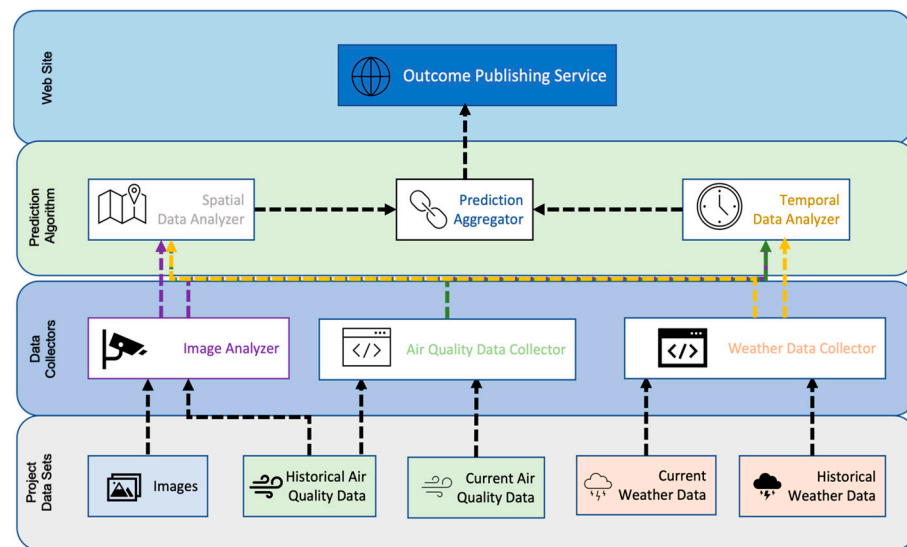
output. The activation function in every neural network serves the essential function of mapping input to output. The input value is calculated by calculating the weighted sum of the neuron input and its bias (if present). This indicates that the activation function determines whether or not to fire a neuron in response to a certain input by generating the matching output. The activation function’s capacity is to differentiate since it enables error backpropagation to train the network [16].

The fully connected layer contains neurons directly connected to the neurons in the two adjacent layers without being connected to any layers within them [14]. Neurons are organised into groups in the fully-connected layer reminiscent of those in traditional neural networks. The node in a connected layer is directly connected to every node in the layer above and below it [16]. The final fully connected layer in the architecture contains the same amount of output neurons as the number of classes to be of the output [13]. The major drawback of a fully-connected layer is that it includes many parameters that need complex computational training examples. Therefore, we try to eliminate the number of nodes and connections. The removed nodes and connection can be satisfied using the dropout technique [17].

The training data in the set would be fitted into the model so that parameters in each layer would be adjusted. One epoch means an entire data set has been run through [18].

**4. Framework Architecture of “Colours-of-the-Wind” (COLD)**

COLD is designed to perform a novel data analytics approach for air quality estimation. Figure 2 illustrates the proposed framework capable of capturing information from various sources, storing data and basic flow, and generating a machine learning-based estimation.



**Figure 2.** End-to-end process of the proposed framework.

COLD crawls images from publicly published meteorological cameras and historical meteorological data to predict air quality. The collected images are further analysed and converted into air quality data, a data source for air quality estimation, by applying a CNN-based image analyser. Groups of components/programs have been developed to achieve this goal, which are detailed as follows:

Data Collectors—These programs collect, process and store data from various sources. This system component is dedicatedly responsible for collecting corresponding information regularly. The data sets are radically available online from multiple sources in varying formats, so multiple programs were required for this information processing. Due to the variation of the data types, COLD has the following groups of data collectors:

1. The air quality data collectors—These collectors collect data related to air quality.

2. The weather data collectors—These collectors collect data related to weather.
3. Historical data collector—These collectors collect historical data for training.
4. Image collectors—These collectors collect images from publicly published cameras.

The image analyser further processes the stored images. The image scraping technique was used to collect this information to become one of the primary data sources for the images. Another essential function for the data collector is to perform necessary data cleansing of the collected data. Herr data cleansing refers to correcting identifiable errors in the data files, including checking data consistency and handling invalid or missing values. The final primary function of the data collector is to store data in the dedicated database and log it.

Analysers—An analyser is designed to estimate air quality according to different variables. In COLD, three types of analysers are designed for this purpose.

1. Temporal analyser—The temporal analyser is responsible for performing the time-related estimation. The primary function of this program is to estimate air quality based on the data collected from the image content in a location and temporal dependency of air quality.
2. Spatial analyser—The primary function of this program is to estimate the spatial transformation of air quality in a place to model the spatial dependency of air quality. As the monitoring stations and cameras have a fixed location, the air quality-related measurements with location information are inserted into an artificial neural model to determine a spatial-based air quality estimation. Interpolation is used for the location between the cameras or air monitoring stations.
3. Estimation aggregator—The estimation aggregator integrates the projections produced by spatial and temporal analysers.

Image analyser—The image analyser analyses the scenes the image collector collects. The primary function of the image analyser is to analyse a composed picture and judge the pollution level. Based on the collected images, data collected from the pictures would pass to the Temporal Predictor and the Spatial Predictor for further estimation.

A three-program set is designed to operate the image analyser, namely:

- (1) the image analyser trainer program;
- (2) the image analyser estimator program; and
- (3) the image analyser image mover program.

Figure 3 illustrates the architecture view of the image analyser, and the logic of each program in the analyser is presented below.

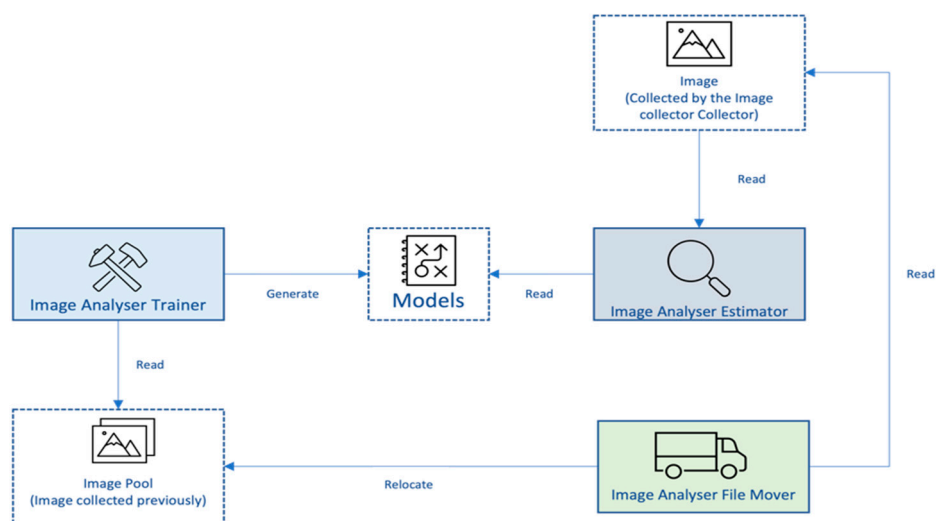
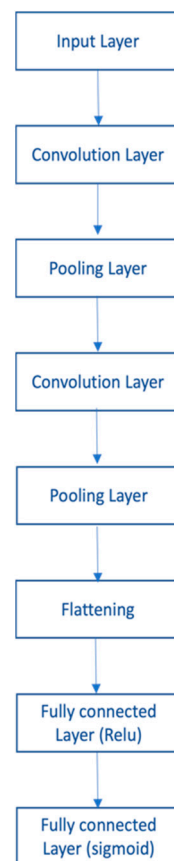


Figure 3. The architectural view of the image analyser programs.

**Image Analyser (Trainer)**—The main aim of the image analyser trainer program is to create trained models based on historical images and air quality records for each camera. To initiate the training, the collected historical images are manually mapped with the air pollution data simultaneously and in location to pair up and become labelled images. These image-air quality pairs are assigned to the trainer and the tester groups. The trainer image-air quality pairs are fitted into a CNN model with two convolution layers with max-pooling layers and a connected layer with two dense layers applied. Figure 4 illustrates the model.



**Figure 4.** A flow chart illustrating the working of the CNN model.

In this model, the convolutional layer recognises patterns from the input images, then further down-sampled by the max-pooling layers. Pooling layers consolidate the features learned by the convolutional layer and reduce the number of parameters by selecting the maximum value from each pool to keep the most prominent features. This would result in computations being further reduced. In general, max pooling is one of the most common types of pooling used. Max pooling is done by applying a max filter to non-overlapping subregions of the initial representation. Max pooling prevents over-fitting by providing an abstracted form of the representation. It also reduces computational cost by decreasing the number of parameters that need to be learned and provides basic translation invariance to the internal representation. The following pseudo-code (Algorithm 1), namely `Image_Analyser_Trainer`, illustrates the idea:

**Algorithm 1:** *Begin-Algorithm Image\_Analyser\_Trainer*

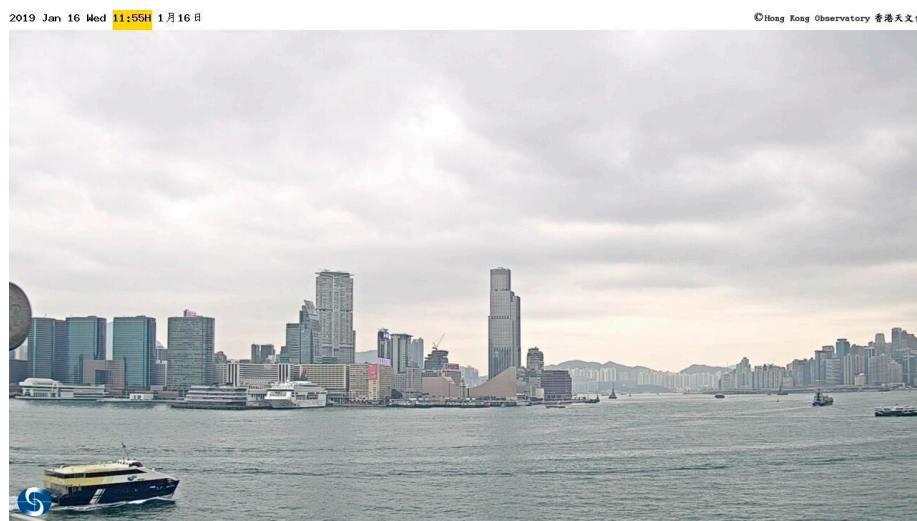

---

```

// Loop at the file structure.
  For all Folders, do the following:
    Load the training dataset ( $P_r$ ) and load the testing dataset ( $P_e$ )
    Obtain the corresponding data classification according to the file structure.
    Create and compile a CNN model ( $M_{cm1}$ ) according to the design in Figure 5
    Fit  $P_r$  into the  $M_{cm1}$ 
    Test  $M_{cm1}$  by  $P_e$  to obtain validation accuracy.
    Store  $M_{cm1}$ 
  Next Folder
End-Algorithm Image_Analyser_Trainer

```

---



**Figure 5.** Good Air Quality([www.hko.gov.hk](http://www.hko.gov.hk) accessed on 1 November 2023).

The tester image-air quality pairs test the corresponding accuracy in the image analyser (Estimator). A well-trained model is used to estimate the air quality of that particular camera. Then, the air quality of the images collected from that camera is calculated. The analyser's output is then stored for future training, and the analysed result(s) are further processed as a part of the data source for air quality estimation.

**Image Analyser (Estimator)**—The image analyser estimator program maps the data collected to the relevant well-trained models created by the trainer program. The estimator collects image files from the temporary folder, where the image collector collects them online.

**Image Analyser (File Mover)**—The image analyser file mover program relocates the data used by the estimator and stores it as a future data source for the trainer program.

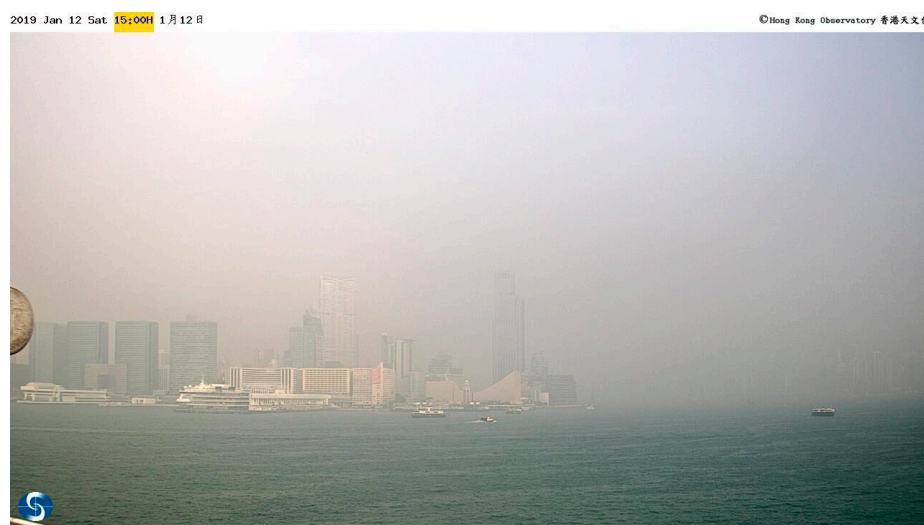
## 5. Dataset

There are three major sets of data that have been used in this research:

- (1) The weather images published by the Hong Kong Observatory;
- (2) The meteorological data collected from the Hong Kong Observatory; and
- (3) The air quality data published by the Environment Protection Department of Hong Kong.

The images used in this research are sets of publicly published weather images from the website of the Hong Kong Observatory. These images are collected from monitoring cameras set up by the Hong Kong Observatory around Hong Kong. For weather monitoring purposes, the Hong Kong observatory published surveillance images and shared the image once every five minutes with the public, which became a key data source for this COLD framework. Figures 5 and 6 illustrate a sample from the camera from one of the many stations used in this study.





**Figure 6.** Poor Air Quality ([www.hko.gov.hk](http://www.hko.gov.hk) accessed on 1 November 2023).

The meteorological data is collected from the Climatological Information Services page of the Hong Kong Observatory's website. Data such as *mean pressure, daily max air temperature, daily mean temperature, daily min air temperature, mean dew point, mean relative humidity, mean amount of cloud* and *total rainfall* are published daily on the observatory's webpage.

The air quality data sets were collected from the Environment Protection Department (EPD) air quality monitoring stations. These stations collect and publish four air pollutants data, namely, ozone (O<sub>3</sub>), nitrogen dioxide (NO<sub>2</sub>), sulphur dioxide (SO<sub>2</sub>) and particulate matter (PM<sub>2.5</sub>/PM<sub>10</sub>) around Hong Kong (<https://www.aqhi.gov.hk/en/what-is-aqhi/about-aqhi.html> accessed on 1 November 2023). In December 2013, the Environment Protection Department launched an index known as the Air Quality Health Index (AQHI) to inform the public of the short-term health risks of air pollution. The AQHIs are reported on a scale of 1 to 10+, with 1 being low health risk and 10+ being severe. The AQHIs are calculated based on the 3-h moving average concentrations of the above four pollutants, and these data could be published from the Air Quality Health Index website on an hourly basis (<https://www.aqhi.gov.hk/en.html> accessed on 1 November 2023).

In this study, the collected images from the Hong Kong Observatory are manually mapped with the air pollution data collected from the Environment Protection Department according to the time taken and the photograph's location to pair up. These image-air quality pairs are then assigned to the image analyser as the trainer and tester groups. The trainer is then submitted to the model (as discussed in Section 3) to train, and the tester is used to test the corresponding accuracy. This study used image scraping to capture and store around 300,000 (three hundred thousand) images from Nov 2018 to Early March 2019. The resolution of the Camera images used in this article is 1280 × 720. As observatory cameras aim to capture weather-related information for the general public, all of the cameras are set in open areas, and the visibility of a specific day determines the distance of the information being captured.

## 6. Evaluation and Discussion

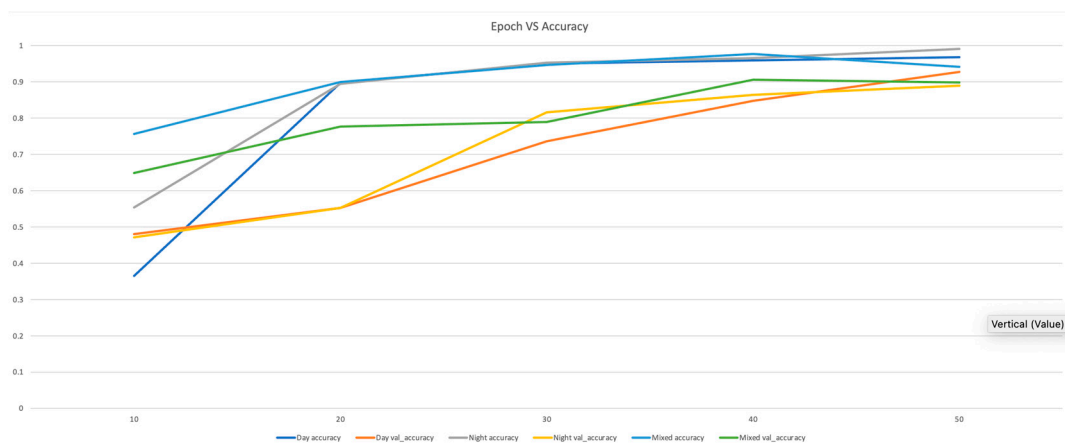
To evaluate the COLD framework, a series of experiments have been conducted, and the *accuracy* and *validation-accuracy* of the Keras class have been used as a measurement metric. Here, the *accuracy* is the accuracy of the model tested by its training data, whereas *validation-accuracy* is the accuracy of a model tested by the testing data. In summary, accuracy is the total amount of correct estimation based on training data (*i.e., the estimated value from the training data equals the actual training data*) divided by the total amount of training data in an epoch. At the same time, *validation-accuracy* is the total amount of correct

estimation based on the testing data (*i.e.*, the estimated value from the testing data equals the actual testing data) divided by the total amount of testing data in an epoch.

One of the central stations was randomly selected to conduct tests in the initial stage. Images from the other thirteen (13) stations were used to verify if the image analysing framework could be used as a general tool for all of these stations. After a series of experiments, the following was learned about the image analysis algorithm.

The first test of the series was to test the feasibility of “using the convolutional neural network (CNN)” (as discussed in the previous section) as a feasible algorithm for analysing images collected by the surveillance camera for data collection purposes. Based on the experimental results, images with *not more than* four air concentration categories (*poor, average, good, excellent*) have a *greater than 75%* accuracy when the program has more than 20 epochs, *i.e.*, 20 times the whole training data sets being run through, in training.

An evaluation experiment is conducted to study the impact of daylight on the images of air pollution property estimation. Three different groups of images were tested in this experiment, and the estimated accuracy was compared. The first set contained mixed day and night images and fit into the program. The second data set only contained day images, whereas the third group contained only night images. A summary of the results was obtained and is depicted in Figure 7:



**Figure 7.** Experimental Results of the Impact of Day Light Test.

For experiment 3, the impact of daylight has been studied. After 30 epochs, all three sets of images got *greater than 95%* accuracy and validation accuracy *above 74%*—the accuracy for pictures taken in the daytime, night time and a mix of both. In our scenario, there is no significant impact on the accuracy of splitting the images taken during the day and night-time. Therefore, we have concluded that breaking the mixed image into day-taken and night-taken groups had no significant difference when conducting image analysis for air quality data collection.

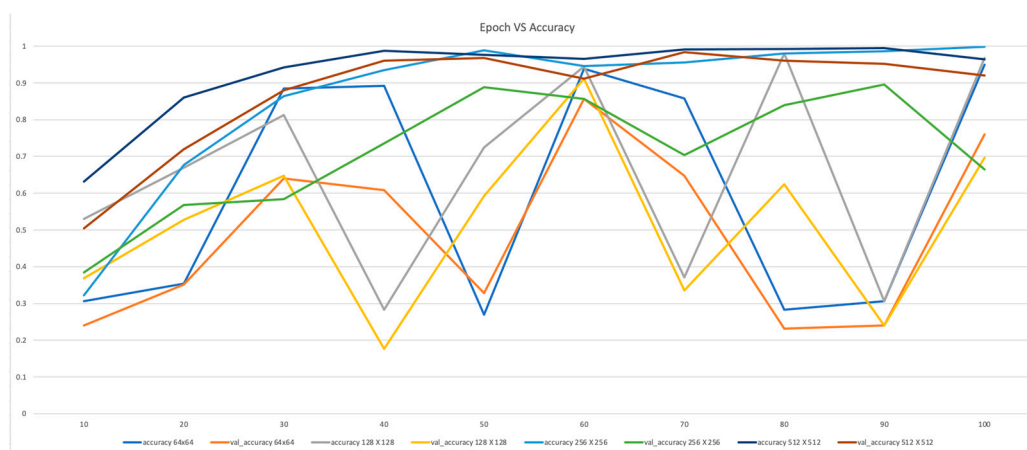
Another experiment is conducted to study seasonal behaviour’s impact on air pollution property estimation images. In this regard, three different groups of images were tested in this experiment, and the estimate’s accuracy was compared. The first set (set 1) contained Autumn Training Data—tested by Winter Testing Data. The second data set (set 2) contained Winter Training Data—tested by Autumn Testing Data. The last data set (set 3) contained both Winter and Autumn data in the training and testing data, which is used as a control set. A summary of the results is presented in Table 1:

**Table 1.** Experimental Results of the Impact of Day Light Test.

	Accuracy	Validation Accuracy
Set 1 (Autumn training, Winter testing)	0.9850	0.4966
Set 2 (Winter training, Autumn testing)	0.9900	0.4832
Set 3 (Control Test)	0.8827	0.8047

This test demonstrates a seasonal dependency on the training and testing data when the training is only on *Autumn* data and testing on only *Winter* data. The validation accuracy drops to 0.4966, and vice versa. However, the control test suggested that the validation accuracy reached 0.8047. This shows a dependency on the season data.

Further experiments were conducted to determine the impact of the image resolution on accuracy. Four input resolutions of  $64 \times 64$ ,  $128 \times 128$ ,  $256 \times 256$ , and  $512 \times 512$  were set up to test the impact of image resolution and accuracy. A summary of the results is presented in Figure 8:



**Figure 8.** Experimental Results of the Impact of Resolution Test.

For the  $64 \times 64$  and  $128 \times 128$  resolutions, the accuracy scatters from 0.2698 to 0.9802, and the validation accuracy ranges from 0.2320 to 0.9120 in these experiments. There is no direct relationship between accuracy/validation accuracy and the epochs. There is no specific pattern for the accuracy, and the output of the  $64 \times 64$  and  $128 \times 128$  input needs reliable precision. It is unsuitable to use as the resolution for air quality estimation in our scenario. The accuracy is relatively stable for the resolution of  $256 \times 256$  and  $512 \times 512$ , and the outcome is more significant than 90% after 40 epochs. The experiment showed that the accuracy of both resolutions of  $256 \times 256$  and  $512 \times 512$  are independent of the number of epochs. The resolution of  $64 \times 64$  and  $128 \times 128$  needs to have reliable accuracy. It is unsuitable to use as the resolution for air quality estimation in our scenario.

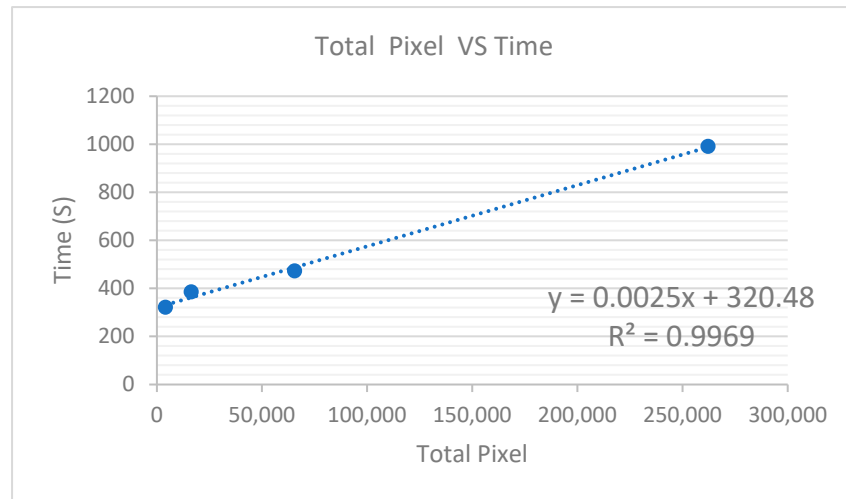
Regarding computational efficiency, Table 2 lists the total time needed to run a 10-epoch training for various resolutions.

**Table 2.** Time (seconds) required to train a ten epoch training for various resolutions.

Resolution	Total Pixel	Time (s)
$64 \times 64$	4096	321
$128 \times 128$	16,384	385
$256 \times 256$	65,536	472
$512 \times 512$	262,144	991

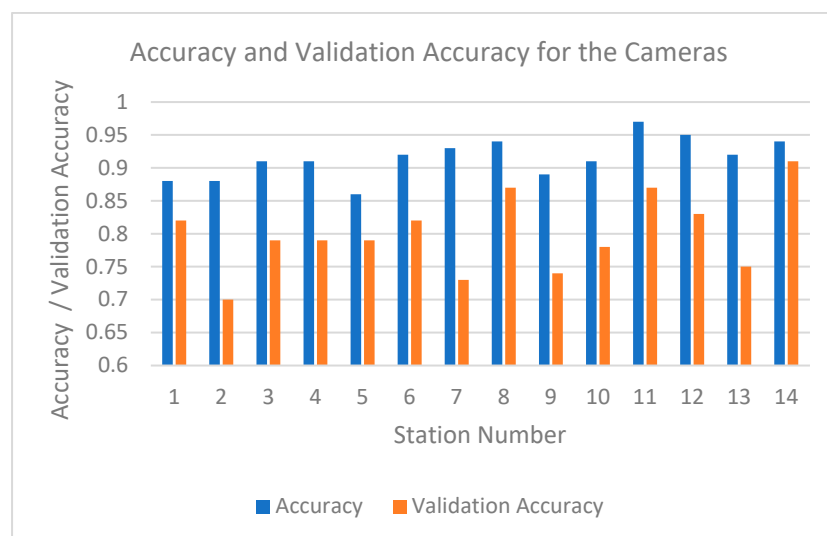
A scatter plot (Figure 9) based on the time and total pixel shows that a linear relationship could be established, the impact of higher resolution images could be estimated

for a 10-epoch training, and the time required could be calculated based on the regression equation (stated above) e.g., for a  $1024 \times 1024$  image, 2942 s (around 50 min) were spent to complete the training.



**Figure 9.** Total Pixel vs. Time (seconds) required to train a 10-epoch training.

This model was applied to experiment with another thirteen (13) sets of images collected from various areas of Hong Kong. The experiment’s results were then used to verify the accuracy of the model of the proposed parameter so that the algorithm’s generalisation could be confirmed. A data set per test collected from a meteorological station (a subset of training and testing data) was set up. A subset of training and testing data was selected randomly for each data set. The accuracy and validation accuracy were measured to see the algorithm’s performance on these datasets. The generalisation of the proposed prototype is verified by data collected from fourteen (14) meteorological cameras. All of these results are summarised in Figure 10.



**Figure 10.** Experimental Results for all 14 Cameras.

The results of the above-explained experiment suggested that the proposed algorithm is a viable solution to analyse the images obtained from the meteorological cameras and generate estimates. The accuracy scatters from 0.85 to 0.97, and the validation accuracy ranges from 0.70 to 0.91 in these tests. This means that all accuracy is higher than 85%, and the validation accuracy is greater than 70%.

In conclusion, based on the experimental outcome, the image analyser of “colours-of-the-wind” could apply a convolutional neural network (CNN) as the learning algorithm. The output of CNN should have four categories. To ensure a reasonable outcome, the input resolution should have at least  $256 \times 256$  with at least 30 epochs of training for the image analyser program.

## 7. Conclusions and Future Work

This paper described a new approach to collecting data from meteorological cameras as a valuable data source for air quality estimation. In this regard, a prototype system, ‘Colours-of-the-wind’ (COLD), is presented along with its sub-systems. The COLD prototype’s breakthrough is the functionality to crawl and convert images into a data source for air quality estimation. In this project, more than three hundred thousand (300,000+) pictures were collected from fourteen (14) meteorological cameras, and based on the outcomes; the presented COLD framework was continuously improved. By applying data analytics and a convolutional neural network, the prototype was built to prove the proposed concept(s). In addition, a series of experiments were performed for the convolutional neural network (CNN) model in order to understand the impacts of various parameters in the algorithm on their accuracy has been studied. The CNN should have an output of four categories; the input resolution should have at least  $256 \times 256$  with at least 30 epochs of training for the image analyser trainer program.

In relation to limitations or possible future directions, we know that certain air pollutants that affect the value of AQHI are invisible to the naked eye, so meteorological cameras cannot capture them. Ozone and  $\text{NO}_2$  are examples of such pollutants that are difficult to detect, and the level of pollutants visible to cameras, such as particulate matter, would dominate the estimation by the camera. This affects the estimation of AQHI in COLD. To further improve this research, a method to detect invisible pollutants by optical means can be investigated to improve COLD. For instance, other optical techniques, such as applying filters or checking light from different wavelengths, can be implemented so that invisible pollutants can also be detected. Second, the research direction that could improve COLD is to have images taken at a shorter distance, such as traffic and security cameras. As the scope of this study was to collect information from existing meteorological cameras that usually take images from a certain distance, a method using short-distance cameras would increase the different variations of data being collected. Another potential research direction is to extend the existing estimation model so that the accuracy may be further improved.

**Author Contributions:** Research and Development—K.C. (R&D lead), P.M. (co-supervisor) and K.M. (director of research); Writing—original draft, K.C.; Writing—review and editing, P.M. and K.M. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work has been partially funded by the University of the West of England (UWE), Bristol, United Kingdom, under their International Collaborations and Partnerships Program.

**Data Availability Statement:** The data supporting this study’s findings are available on Hong Kong Observatory Website, accessible at [https://www.hko.gov.hk/en/wxinfo/ts/index\\_webcam.htm](https://www.hko.gov.hk/en/wxinfo/ts/index_webcam.htm) accessed on 1 November 2023 and the Environmental Protection Department <https://www.aqhi.gov.hk/mobile/gt/aqhi/statistics-of-aqhi/past-aqhi-records.html> accessed on 1 November 2023.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Yang, Z.; Wang, J. A new air quality monitoring and early warning system: Air quality assessment and air pollutant concentration prediction. *Environ. Res.* **2017**, *158*, 105–117. [CrossRef]
2. Madan, T.; Sagar, S.; Virmani, D. Air Quality Prediction using Machine Learning Algorithms-A Review. In Proceedings of the 2020 2nd International Conference on Advances in Computing, Communication Control and Networking (ICACCCN), Greater Noida, India, 18–19 December 2020; pp. 140–145.
3. Julfikar, S.K.; Ahamed, S.; Rehena, Z. Air Quality Prediction Using Regression Models. *Lect. Notes Electr. Eng.* **2021**, *778*, 251–262.

4. Sahoo, L.; Praharaj, B.B.; Sahoo, M.K. Air Quality Prediction Using Artificial Neural Network. *Adv. Intell. Syst. Comput.* **2021**, *1248*, 31–37.
5. Kök, I.; Şimşek, M.U.; Özdemir, S. A deep learning model for air quality prediction in smart cities. In Proceedings of the 2017 IEEE International Conference on Big Data (Big Data), Boston, MA, USA, 11–14 December 2017; pp. 1983–1990.
6. Li, X.; Peng, L.; Yao, X.; Cui, S.; Hu, Y.; You, C.; Chi, T. Long short-term memory neural network for air pollutant concentration predictions: Method development and evaluation. *Environ. Pollut.* **2017**, *231*, 997–1004. [[CrossRef](#)]
7. Umar, I.K.; Nourani, V.; Gökçekuş, H. A novel multi-model data-driven ensemble approach for the prediction of particulate matter concentration. *Environ. Sci. Pollut. Res.* **2021**, *28*, 49663–49677. [[CrossRef](#)] [[PubMed](#)]
8. Padilla, D.A.; Magwili, G.V.; Mercado, L.B.Z.; Reyes, J.T.L. Air Quality Prediction using Recurrent Air Quality Predictor with Ensemble Learning. In Proceedings of the 2020 IEEE 12th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment, and Management (HNICEM), Manila, Philippines, 3–7 December 2020.
9. Zheng, Y.; Liu, F.; Hsieh, H. U-Air: When urban air quality inference meets big data. In Proceedings of the KDD '13: Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Chicago, IL, USA, 11–14 August 2013; pp. 1436–1444.
10. Zhang, J.; Ding, W. Prediction of air pollutants concentration based on an extreme learning machine: The case of Hong Kong. *Int. J. Environ. Res. Public Health* **2017**, *14*, 114. [[CrossRef](#)] [[PubMed](#)]
11. Wong, C.J.; Matjafri, M.Z.; Abdullah, K.; Lim, H.S.; Low, K.L. Temporal air quality monitoring using surveillance camera. *Int. Geosci. Remote Sens. Symp.* **2007**, 2864–2868. [[CrossRef](#)]
12. Wong, C.J.; Matjafri, M.Z.; Abdullah, K.; Lim, H.S. Temporal and spatial air quality monitoring using internet surveillance camera and ALOS satellite image. In Proceedings of the 2009 IEEE Aerospace Conference, Big Sky, MT, USA, 7–14 March 2009; pp. 1–7.
13. Li, Y.; Huang, J.; Luo, J. Using User Generated Online Photos to Estimate and Monitor Air Pollution in Major Cities. *arXiv* **2015**, arXiv:1508.05028.
14. Chakma, A.; Vizena, B.; Cao, T.; Lin, J.; Zhang, J. Image-based air quality analysis using deep convolutional neural network. In Proceedings of the 2017 IEEE International Conference on Image Processing (ICIP), Beijing, China, 17–20 September 2017; pp. 3949–3952.
15. He, K.; Sun, J.; Tang, X. Single image haze removal using dark channel prior. *IEEE Trans. Pattern Anal. Mach. Intell.* **2011**, *33*, 2341–2353. [[PubMed](#)]
16. Liu, C.; Tsow, F.; Zou, Y.; Tao, N. Particle Pollution Estimation Based on Image Analysis. *PLoS ONE* **2016**, *11*, 1–14. [[CrossRef](#)] [[PubMed](#)]
17. Bezdan, T.; Džakula, N.B. Convolutional Neural Network Layers and Architectures. In Proceedings of the SINTEZA 2019 International Scientific Conference on Information Technology and Data Related Research, Novi Sad, Serbia, 20 April, 2019; pp. 445–451.
18. O’Shea, K.; Nash, R. An Introduction to Convolutional Neural Networks. *arXiv* **2015**, arXiv:1511.08458.
19. Bayat, O.; Aljawarneh, S.; Carlak, H.F. International Association of Researchers, Institute of Electrical and Electronics Engineers, and Akdeniz Universitesi. In Proceedings of the 2017 International Conference on Engineering & Technology (ICET’2017): Akdeniz University, Antalya, Turkey, 21–23 August 2017.

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.