Interdisciplinary Description of Complex Systems 21(6), 640-654, 2023

ANALYSIS OF AIR QUALITY PARAMETERS TO ASSESS THE IMPACT ON LAYERS IN POULTRY FARMS USING DEEP LEARNING

Bidri Deepika[1,]*, Nagarathna[1] and Channegowda[2]

[1]P E S College of Engineering
Mandya, India

[2]ZEUS Biotech Pvt. Ltd.
Mysore, India

DOI: 10.7906/indecs.21.6.9
Regular article

Received: 6 April 2023.
Accepted: 16 December 2023.

ABSTRACT

The food security has increased the agriculture production due to satisfying demand of ever-growing population. Due to this growth in population, the demand of protein also increased. A significant amount of population depends upon the chicken and egg to fulfil the demand of protein. The meat and egg production depends on the quality of poultry farming. The presence of air contaminants causes poor air quality within the poultry house which affects health of layers, production of eggs and workers in poultry farm. The proposed work uses data analysis approach and machine learning concept to automatize the process of air quality monitoring in poultry farms. A Convoluted Neural Network Long Short-Term Memory model, along with bidirectional Long Short-Term Memory model is proposed to improve the forecasting performance. This method predicts the Air Quality Index based on air quality parameters. The proposed approach is tested on poultry farm air quality dataset which is collected from different poultry farms. Finally, the obtained performance is compared with existing techniques in terms of RMSE, MAE, MAPE and correlation coefficient.

KEY WORDS

AQI, LSTM, poultry, air quality, agriculture, egg production

CLASSIFICATION

JEL: Q16

*Corresponding author, η: bdeepika@pesce.ac.in; +91 8073310141;
Department of Computer Science & Engineering, P E S College of Engineering,
Mandya-571401, Karnataka, India

INTRODUCTION

A rise in agricultural productivity is required due to concern about food security because of the world's constantly expanding population, which is expected to reach nearly 9,6 billion people by the year 2050 [1]. However, since the early 1980s, the worldwide per-capita grain output has decreased due to environmental production restrictions. Proteins derived from animals are also becoming more and more popular [2]. As a result, it is predicted that by 2050, worldwide meat consumption would have increased by 70 % [3]. According to a research by Henchion et al. [5] on the trends of meat consumption, consumption of poultry meat and poultry meat products has increased, and this trend is expected to continue over the next ten years due to consumers' preferences for white meat, particularly chicken [6]. Animal health, efficiency and sustainable environmental circumstances have become difficult requirements to meet as chicken production has increased and as knowledge of acceptable conditions for animal welfare has grown [7]. As a result, human surveillance is no longer a practical method for monitoring livestock.These issues have been addressed by Precision Livestock Farming (PLF), which offers effective automated solutions while still upholding animal welfare[8]. PLF supports stockmen by keeping track of several bio-processes and bio-reactions pertaining to animal welfare, health, and production [9].

Due to its high protein, low calorie, and low cholesterol content, chicken meat is becoming more and more popular [10]. Nevertheless, environmental factors, disease outbreaks, the breeding process, and active management activities all affect chicken output levels [10]. To stop contagious illnesses, increase productivity, and assure healthy layers, effective chicken health and welfare management is crucial. However, conventional methods for managing the welfare of layer fowl are prone to high labour costs and ineffective resource management, including excessive feed, water and electricity usage. The combination of Internet of Things (IoT) and Machine Learning (ML) has been viewed in this context as one of the most promising technologies for delivering smart poultry farming, continuous data monitoring, and prescriptive analytics in order to address the aforementioned challenges for efficient resource control and optimal decision-making [11]. The IoT enabled systems play a significant role in chicken farming by offering automated monitoring facilities, according to the study given by Raj et al. [12]. However, a number of characteristics, including water, temperature, air and others, are included in the automated monitoring of chicken farms. In this work, air quality parameters such as CO (Carbon Monoxide), NH_3 (Ammonia), PM2.5 (Fine Particulate Matter having 2,5 µm), PM10 (Fine Particulate Matter having 10 µm) and SO_2 (Sulphur dioxide) are studied. These parameters lead to respiratory diseases in poultry layers which affect the weight, feed-conversion ratio and egg production in poultry farms. Generally, the air quality monitoring is the prime aspect of poultry farms in this endeavour to address problems with egg production, boost output and also economy of farmers.Thus, the main aim of this study is to maintain health and welfare of chickens farmed in poultry houses, the air quality need to be monitored which maximizes the egg production and reduces the risk of respiratory diseases [13].

Compute vision-based technologies serve as the foundation for conventional monitoring-based methodologies. These methods rely on frameworks for picture capture, image processing, and detection. However, the production of eggs and meat is also significantly impacted by the air quality [13]. In order to distinguish between appropriate and unsuitable air quality, this work used deep learning modules to monitor the air quality in chicken farms. Air quality forecasting is crucial for taking the necessary actions to increase poultry farming output.The air quality forecasting/prediction is a time series problem where sensors generates the data which may be contaminated due to uncertainty, redundancy, missing values etc. [13]. Due to these issues, the traditional approaches fail to learn the sequential data patterns that generate inaccurate forecasting. Moreover, the traditional machine learning approaches are not suitable for complex real-world scenario. Thus, current research community has motivated by these issues and adopted deep learning techniques due to their significant nature of pattern learning. The deep learning based schemes are widely adopted in various domains such as image classification [23], video summarization [24], Natural Language Processing [25], data mining and many more [26]. Thus, the deep learning schemes are adopted by researchers to obtain the improved accuracy in various domains. To predict air quality the deep learning architecture is suitable for sequential, seasonal, nonlinear and cyclical dependency problems between pollutant data. The long-term dependencies from time-series data can be learned using Long Short-Term Memory (LSTM) technique [27].These techniques achieve the better performance but increasing the finer levels of prediction in air quality monitoring leads to inaccurate forecasting which affects the performance of precision poultry farming. Thus, to overcome these issues, a novel hybrid deep learning approach by using Convolutional Neural Network (CNN) and Bi-LSTM model is introduced. The main contributions of this approach are as follows:

- a brief study about existing techniques of precision poultry farming and air quality monitoring is presented,
- in next phase, a data pre-processing scheme which deals with the missing values and performs the data normalization is presented,
- in next stage, a new hybrid approach by combining the CNN and LSTM model is developed,
- the performance of LSTM model is further improved by incorporating bi-directional LSTM model.

Rest of the article is organized in following sections: literature survey describes the brief literature review regarding recent air quality monitoring techniques in poultry farming, proposed model presents proposed solution, results and discussion presents the experimental analysis where the performance of proposed system is analysed and finally, conclusion presents the concluding remarks about this approach.

LITERATURE SURVEY

This segment describes the brief discussion about recent studies in this field of air quality monitoring and forecasting. Okinda et al. [6] reported the importance of early detection of poultry farming related diseases and presented computer vision based monitoring system to identify the zoonotic infections. This model uses a computer vision based approach for data collection and monitoring. In order to obtain the features, the 2D shape features such as circle, variance, convexity and eccentricity etc. parameters are calculated. Along with these parameters, mobility or walking speed is also considered for analysis. Further, the support vector machine classifier is used to perform the classification.

In [11] Fang et al. also used computer vision based pose estimation strategy for poultry behaviour monitoring and identifying the sickness of the chicken. In this work, authors presented a deep learning based framework to analyze the chicken behaviour. The broiler chicken's feature points are utilised to build the posture skeleton, which is subsequently used to track certain body parts. Additionally, the broiler chicken positions were categorised and identified using the Naïve Bayesian model (NBM). Comparing the postures of categorised broiler chickens, preliminary testing showed that we could distinguish between birds in the standing, walking , running, eating, resting, and preening stages.

Yang et al. [14] studied about the impact of fine particulate matters on chicken health in China. In this study, authors considered the analysis of fine particulate to study about the poultry farm pollution. Lahlouh et al. [15] focused on improving the animal production with the help of broiler house systems. Generally, the animal production is affected due to several parameters such as hygro-thermal parameters such as temperature and relative humidity and contaminant gases such as NH_3, CO_2. In order to perform this operation authors adopted fuzzy logic based approach trained with the help of fuzzy rules. This helps to monitor the chicken health regularly.

Al Assaad et al. [16] reported the issues faced by poultry farmers in the semi-arid climatic region. Therefore, authors presented a comparative analysis of passive cooling systems to meet the indoor air quality. This study presented the impact of NH_3, and CO_2 on poultry farming. Grilli et al. [17] focused on coccidiosis disease detection in poultry farming for precision livestock farming by using combined cheap and specific approaches. This study is based on the Air quality monitoring. Sayour et al. [18] presented wireless sensor network based monitoring system to measure and monitor the air quality. Currently, the WSNs are widely adopted in various real-time systems for data collection, and monitoring such as environment monitoring, weather forecasting etc. In this work, authors mainly focused on temperature, CO_2, NH_3 and humidity monitoring to measure the air quality.

Jayarajan et al. [19] presented IoT based automated approach to monitor the poultry farms because most of the Indian farms use manual and physical methods for this purpose. The IoT based model can help to monitor the temperature, air quality and other parameters which have impact on chickens. This technique monitors these parameters and generates the alert messages to inform to the clients. Similar to this, Lashari et al. [10] also used IoT based system for monitoring the poultry environment. this approach uses humidity, O_2, CO_2, temperature, and NH_3 as the important parameter.

Hofmann et al. [20] reported several conditions of poultry farms which has stressful impact on chickens and their immune system. This review reported the impact of Ammonia, Hydrogen sulphide, Photoperiod, Light Color/Wavelength, and Light Intensity on immune system. High concentration of these gases affects the respiratory and cardiovascular system of the birds. NH_3 is affected due to manure management, temperature and moisture which vary between the houses. Similarly, the H_2S is generated from degradation of liquid manure under anaerobic conditions. Naseem et al. [21] identified Ammonia as an important concern for environmental pollution which is produced from the poultry farming. Production of NH_3 depends on several factors such as bird age, litter type, pH, temperature, ventilation rate and many more. Minimization of NH_3 emission can help to reduce the severe diverse impact of ammonia therefore several types of modern ventilations systems are developed to minimize its emission.

Generally, intensive farming has adopted the antibiotics based solutions for health related complications in farming. This has led to spreading of drugs in environment and contributed to the antibiotic resistance. To mitigate this effect, early detection of disease is highly recommended and several researchers have been working for early disease detection to minimize the use of drugs. Grilli et al. [17] studied about detection of coccidiosis in poultry farms. This method is based on the air quality monitoring system where samples are collected in the presence of coccidiosis. Later, these samples are classified based on the with the help of machine learning approaches. Pereira et al. [22] studied the impact of temperature, ammonia and humidity levels in poultry farming based on the concept of IoT systems. This model mainly concentrates on improving the data transmission method using hardware based system.

PROPOSED MODEL

This section presents the proposed deep learning based solution for air quality prediction and forecasting to improve the poultry farm monitoring. The first phase of this scheme includes data pre-processing where data normalization is employed which helps to minimize the data volatility. In next phase, the CNN and LSTM models are described briefly. Finally, the proposed hybrid CNN-LSTM prediction model is presented.

MISSING VALUE IMPUTATIONS

This stage plays an important role in the field of data analytics. This is used to improve the data representation. During air quality data collection, some sensor may produce faulty values or may lead to missing values in the dataset. Therefore, applying missing value imputation is a primary task. This scheme helps to remove the outlier, imputes the missing values and normalize the input data. In this work, we have collected several data which represent the air quality such as ammonia, particulate matter, CO_2 and SO_2. However, the some of the entries are missing in the dataset which are due to non-working of sensor. First of all, a normalization function is employed to obtain the data in certain range. The normalization function is given as:

$$Y = \frac{(A-M)}{S}, Y = \frac{A-A_{min}}{A_{max}-A_{min}}, \tag{1}$$

where A denotes the actual input data, M denotes the mean of this data, S is the standard deviation, A_{max} and A_{min} represent the maximum and minimum values of attributes. Furthermore, a missing value imputation method is presented by using a combined KNN and correlation computation model. The missing value dataset is represented as given in Table 1.

Table 1. Missing value data representation.

Id	Col1	Col2	Col3	Col4	Col5
0	2,0	5,0	3,0	6,0	4,0
1	9,0	Nan	9,0	0,0	7,0
2	19,0	17,0	Nan	9,0	Nan

Based on this missing data representation, the correlation between attributes is computed as follows:

$$r_{m_2 j} = \frac{\sum_{i=1}^{q} y_{im_2} y_{ij} - \frac{\left(\sum_{i=1}^{q} y_{im_2}\right)\left(\sum_{i=1}^{q} y_{ij}\right)}{q-1}}{\sqrt{\left[\sum_{i=1}^{q} y_{im_2}^2 - \frac{\left(\sum_{i=1}^{q} y_{im_2}\right)^2}{q-1}\right]\left[\sum_{i=1}^{q} y_{ij}^2 - \frac{\left(\sum_{i=1}^{q} y_{ij}\right)^2}{q-1}\right]}}, \tag{2}$$

where $r_{m_2 j}$denotes the correlation coefficient between missing data and variable j, C_m is the case which has the missing value, V_{m_2} denotes the variable which has the missing value, y_{im_2} denotes the variable in any case which has the missing data and y_{ij} is the complete data.

For missing value imputation, KNN imputation method is applied for a specified numbers of k. The Euclidean distance between data, and missing data is computed as follows:

$$dist\left(C_{m_1}, C_i\right) = \sqrt{\sum_{j=1}^{p}\left(y_{m_1 j} - y_{ij}\right)^2}. \tag{3}$$

Here, $dist\left(C_{m_1}, C_i\right)$ denotes the distance between missing value data and complete data, p represents the total number of variables in the dataset and q denotes the total number of cases. The obtained distance values are sorted based on the k values which are closer to the case which has the missing data. By utilizing the aforementioned distance metrics and other parameters, the missing data attribute values can be estimated as:

$$\hat{y}_{m_1,m_2} = \frac{\sum_{a}^{k} y_{a,m_2}}{k}, \tag{4}$$

where $\hat{y}_{m_1,m_2}$ denotes the estimated missing data and y_{a,m_2}represents the data which is the same variable with missing data and closer to the case which is having missing data. With the help of this process, this model generates a pre-processed data which is later used for learning purpose.

PROPOSED HYBRID MODEL

This section presents the description about CNN-LSTM model for prediction and forecasting approaches. The CNN-LSTM air quality prediction model is constructed by combining the characteristics of CNN and LSTM network. When combined with existing models, the CNN-LSTM air quality index prediction model proposed in this paper combines the efficient feature extraction ability of CNN and the ability of LSTM to deal with long time series. The prediction accuracy of model can be improved by learning, analyzing, and processing the historical data through various structures. The CNN-LSTM model was compared with the other models , and the prediction model with the best performance was selected. CNN-LSTM, which is used to extract the complex characteristics received from the different air quality monitoring equipment, may be described simply as a serial connection of CNN and LSTM modules. The air quality forecast also makes use of these features. The CNN layer of the network's initial section takes into account many inputs, including NH_3, CO_2, SO_2, and others. For the CNN layer, this data is regarded as meta data. The whole CNN module has many hidden layers as well as a single output layer that sends the features that were retrieved to the LSTM for further learning. The hidden layer is comprised of convolution layer which helps to generate the activation map of the data, Rectified Linear Unit (ReLU) layer is used to discard the negative values from activation map and pooling layer which is used to reduce the spatial size resulting in controlling the overfitting.

In this process, the convolution layer performs its operations on the incoming multivariate time series data sequence and passes results to next layers. Let us consider that air quality vector is expressed as $x_i^0 = \{x_1, x_2, \ldots., x_n\}$ where n denotes the time window which is normalized as 60 min unit per window. The use of a 60-minute normalization window refers to a specific design choice related to the temporal aspects of the data being processed. It is used to ensure that the input data within each 60-minute segment is on a consistent scale, aiding the learning process. The outcome of first convolution layer can be expressed as:

$$\begin{aligned} y_{ij}^1 &= \sigma\left(b_j^1 + \textstyle\sum_{m=1}^{M} w_m^1, j\mathcal{X}_{i+m-1}^0, j\right), \\ y_{ij}^l &= \sigma\left(b_j^l + \textstyle\sum_{m=1}^{M} w_m^l, j\mathcal{X}_{i+m-1}^0, j\right), \end{aligned} \tag{5}$$

The y_{ij}^1 denotes the output of first convolution layer and y_{ij}^l denotes the output of l^{th} convolution layers, b_j^1 denotes the bias for the j^{th} feature map, w represents the kernel weight, m is index value of filter and σ is the activation function. Moreover, the convolution layer consist of a pooling layer which is used to combine the neuron outputs from one layer to another layer. This layer helps to minimize the space size of feature representation resulting in minimizing the number of network parameters and reducing the computational cost. Here, this model the max-pooling operation which considers the maximum value of each from previous layer. The max-pooling operation can be expressed as:

$$p_{ij}^l = \max_{r \in R} y_{i \times T + r, j}^{l-1}, \tag{6}$$

where R is the pooling size, y is the input data and T denotes the stride value.

Further, the LSTM unit is considered which the second part of this network is. This layer stores the important information obtained from the CNN module. Mainly, the LSTM unit consolidates the memory units which help to update the memory data of previous hidden states. Thus, LSTM has a characteristic to obtain the temporal relationship of long-term multivariate sequence. Further, the output of previous layers is passed to the gated units available in the LSTM network. The LSTM network becomes the prime choice for forecasting problems because of its capacity to handle the gradient vanishing problems when learning using traditional RNN learning approaches. The LSTM unit cab be defined as a group of vector R^d at time step t. Figure 1 illustrates the basic architecture of LSTM unit.

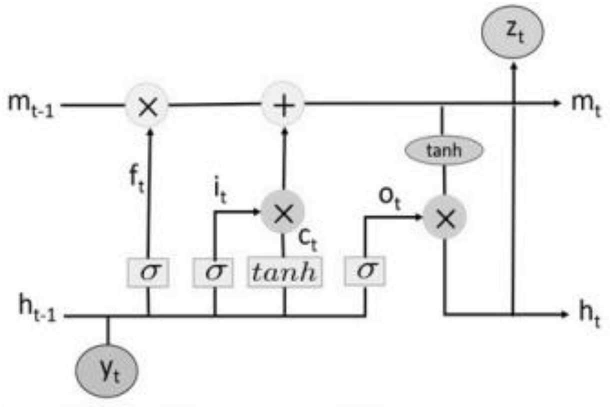

Figure 1. Basic architecture of LSTM Unit.

The other components of LSTM are described further in the text.

Memory cell unit (m_t): it stores the intermediate steps of feature learning process. It can be expressed as:

$$m_t = f_t . m_{t-1} + i_t . c_t . \quad (7)$$

Here, $c_t = \tanh(W_m . [h_{t-1}, y_t] + b_m)$ where t denotes the current time step, $(t-1)$ denotes preceding time step, $h_{(t-1)}$ is the hidden state at time step $(t-1)$, W_m characterise the weight matrix for memory cell, y_t is the input data and b_m is the bias for memory cell unit. The LSTM module contains input gate, forget gate, and output gate. These components are denoted as follows:

- input gate (i_t): it considers the pre-processed air quality data as input from previous layers of CNN. This can be expressed as:
$$i_t = \sigma(W_i . [h_{t-1}, y_t] + b_i), \quad (8)$$
- forget gate (i_t): this is the intermediate gate which resets the old memory data. This can be given as:
$$i_t = \sigma(W_i . [h_{t-1}, y_t] + b_i), \quad (9)$$
- output gate (i_t): this is the final unit of LSTM which generates the final output from learning and prediction steps. This can be given as:
$$o_t = \sigma(W_0 . [h_{t-1} . y_t] + b_0), \quad (10)$$
- finally, the hidden cell state (h_t) used by LSTM as hidden units can be expressed as:
$$h_t = o_t \cdot \tanh(m_t). \quad (11)$$

Generally, these modules process the information from input to output steps in a single direction which suffer from information preserving for accurate forecasting. Thus, to overcome this issue, the proposed model has adopted the bi-directional LSTM architecture which processes the data into two directions as forward and backward. In forward direction, the bi-directional LSTM passes the data from previous index to next point as input whereas in backward direction it passes the data from future inputs to past inputs. This scheme of bi-directional data processing helps to preserve the learned attributes from past inputs and future inputs while processing through different hidden layers. Further, these outputs are processed through the output layer. Figure 2 shows the architecture of bidirectional LSTM.

The forward process is represented as $\vec{h}$ and backward process is denoted as $\overleftarrow{h}$. The outcome of these forward and backward data processing is computed based on the aforementioned conational equations. The final outcome of this bi-directional LSTM is obtained as $Z_T = [Z_{T-k}, Z_{T-k+1}, \dots, Z_{T-1}]$ where each element of this vector is expressed as:

$$z_t = \sigma(\vec{h}, \overleftarrow{h}), \quad (12)$$

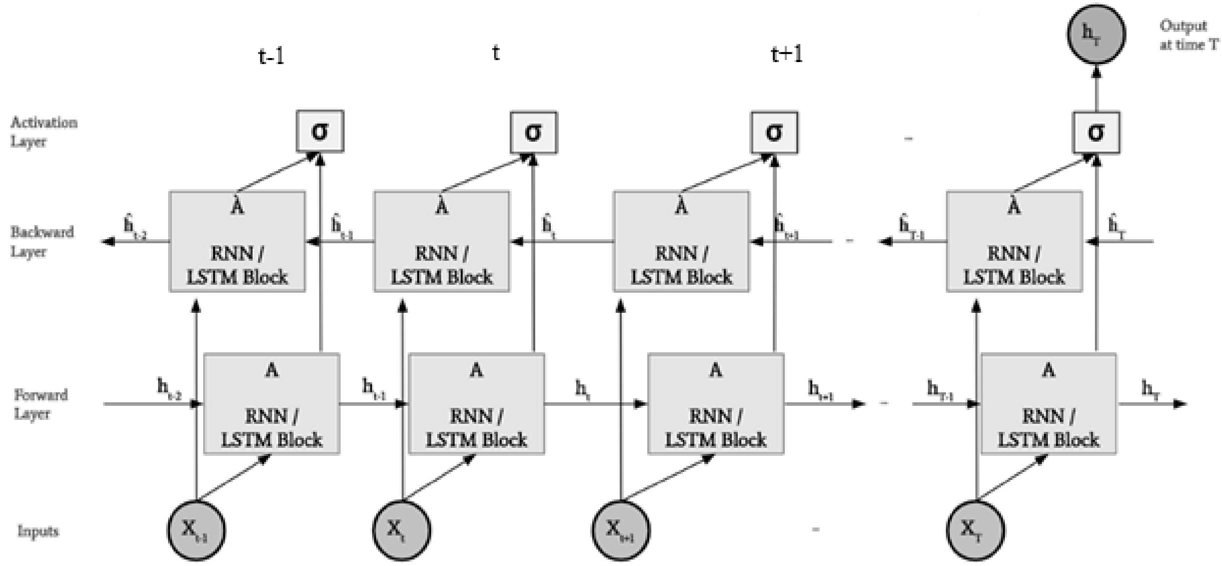

Figure 2. Bi-directional LSTM architecture

where σ is the function which helps to integrate the output of forward and backward passes. Similarly, the output of forward $(\vec{h})$ and backward $(\overleftarrow{h})$ passes can be denoted as:

$$\begin{aligned}\vec{h} &= H\left(W_{y\vec{h}}y_t + W_{\vec{h}\vec{h}}\vec{h}_{t+1} + b_{\vec{h}}\right),\\ \overleftarrow{h} &= H\left(W_{y\overleftarrow{h}}y_t + W_{\overleftarrow{h}\overleftarrow{h}}\overleftarrow{h}_{t+1} + b_{\overleftarrow{h}}\right),\end{aligned} \tag{13}$$

Based on these models of CNN and LSTM, a combined deep learning model is presented to learn the attributes. Below given figure depicts the architecture of proposed model. This model considers the several parameters as input from the data based and processed through the CNN module. These inputs are further transformed into a number of two-dimensional matrices containing time series. The CNN network is then used to extract the features from these matrices. The LSTM receives the output as input. The final prediction result is obtained by decoding the LSTM output using the fully connected layer as shown in Figure 3.

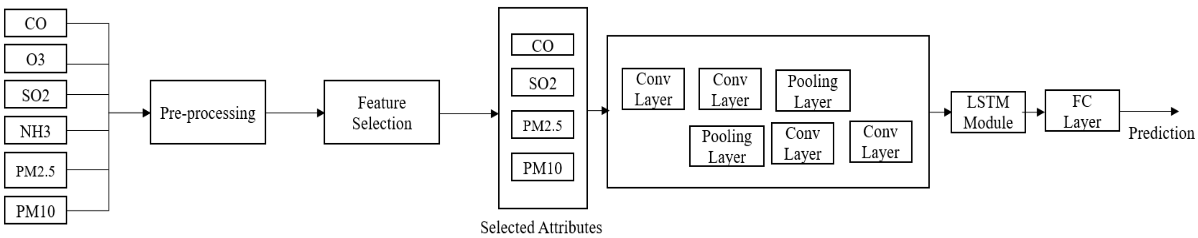

Figure 3. Proposed combined architecture for air quality prediction

It is assumed that the proposed architecture has α layers where β layers are under the training process which is used to train the x_i training data for the y_i labels.

The training process of proposed deep learning model can be expressed as:

$$\begin{aligned}C_i^{\mu} &= g(u^T x_i), \quad \beta = 2,\\ P_i^{\mu} &= g(v^T C_i), \quad w < \beta < \delta - 3,\\ L_i^{\mu} &= g(w^t P_i + d^T L_i), \quad 3 < \beta < \delta - 2,\\ y_i^{\mu} &= f(\zeta^T L_i), \quad \mu = \alpha - 2.\end{aligned} \tag{14}$$

where u denotes the weight matrix of input which corresponds to the convolution layer, v represents the weight matrix from convolution layer to pooling layer, w is the weiht matrix between pooling and LSTM layer, d is the weight of information exchange between LSTM and neurons and finally, ζ characterise the weight of fully connected layer.

RESULTS AND DISCUSSION

This section presents the experimental analysis using proposed approach and compared the performance with various existing techniques. The proposed scheme is applied on air quality data which is collected from poultry farms. This model is trained using NVIDIA RTX 2060 GPU, Intel I core 10th generation processor, 16 GB RAM installed on Linux platform. This implementation includes Keras and Tensorflow libraries for learning along with Adam optimizer.

DATASET ARRANGEMENT

During data collection, several parameters are considered which are collected for a duration of one year from https://kspcb.karnataka.gov.in. Air Quality Index (AQI) is used to measure the quality and how clean the air is and it is a number used by the government agencies to measure the quality of air. Below given Table 2 shows a sample range of these parameters and computed their corresponding threshold to measure the overall AQI.

Table 2. Air Quality Classes and their Corresponding Values.

Air Quality	AQI	PM2.5 (24 hr avg)	PM10 (24 hr avg)	CO	SO2
Good	0-50	0-12	0-54	0-4,4	0-35
Moderate	51-100	12,1-35,4	55-154	4,5-9,4	36-75
Unhealthy	101-150	35,5-55,4	155-254	9,5-12,4	76-185
Very unhealthy	151-200	55,5-150,4	255-354	12,5-15,4	186-304

Let PM2.5 sensor records the 30,1 $\mu g/\text{m}^3$ as average pollution reading for a 24-hour period. The obtained reading is in the range of 12.1-35 which shows moderate air quality index. The AQI for this parameter is computed as follows:

$$I_{PM2.5} = \frac{(I_{High} - I_{Low})}{(C_{High} - C_{low})} \times \left(C_p - C_{low}\right) + C_{low} = \frac{(100-51)}{35.4-12.1} \cdot (30{,}1 - 12.1) + 51 = 89. \quad (15)$$

PERFORMANCE MEASUREMENT PARAMETERS

Based on the previous analysis, the dataset is arranged in 5 different classes which is classified with the help of proposed approach. The classification performance of proposed approach is measured by estimating the confusion matrix. Below given Table 3 shows the confusion matrix for 4 class scenario.

Table 3. Confusion matrix representation

		Predicted Class			
		Good	Moderate	Unhealthy	Very Unhealthy
Actual Class	Good	TN	FP	TN	TN
	Moderate	FN	TP	FN	TN
	Unhealthy	TN	FP	TN	TN
	Very unhealthy	TN	FP	TN	TN

True positive (TP): it shows that the classifier correctly predicts the positive class from the given test set.

True Negative: it shows that the classifier model correctly predicts the negative class from the given test set.

The true negative and true positive values show the accuracy of classifier. However, these categories should match the actual values of TP and TN.

False positive: denotes the classifier model incorrectly predicts the positive class.

False negative: denotes that the classifier mistakenly predicted the negative class.

This confusion matrix facilitates in the computation of the proposed approach’s overall accuracy, precision, specificity, sensitivity, and F-measure. The symbol for accuracy (acc) stands for the rate of proper classification. It is calculated by dividing the total number of predictions by the percentage of true predictions. It may be computed as follows:

$$Acc = \frac{\text{TruePositive+TrueNegative}}{TP+TN+FP+FN}. \tag{16}$$

Similarly, the Recall performance also can be computed with the help of true negative and false positive values. This can be computed as follows:

$$Recall = \frac{TruePositive}{TruePositive+TN}. \tag{17}$$

Then, Precision of the proposed approach is comptuted. It is computed by taking the ratio of True Positive and (True and False) positives.

$$P = \frac{TruePositive}{TruePositive+FP}. \tag{18}$$

Finally, F-measure is computed which is the mean of precision and sensitivity performance. It is expressed as:

$$F = \frac{2*P*Sensitivity}{P+Sensitivity}. \tag{19}$$

Based on these parameters, the performance of proposed approach is measured and compared with other supervised classifiers. Further, the performance of Air Quality forecasting and compared the performance of proposed approach in terms of Root Mean Square Error (RMSE), MAPE, Mean Absolute Error (MAE), and R2 score/coefficient. These parameters can be computed as follows:

$$RMSE = \frac{1}{n}\sqrt{\frac{1}{n} * \sum_{j=1}^{n}\left(A_j - P_j\right)^2}\,, \quad MAE = \frac{\sum_{j=1}^{n}|A_j - P_j|}{n}. \tag{20}$$

Here, A_j represents the actual energy consumption data for a given time sequence t, P_j denotes the predicted value of energy consumption at time step j and n denotes the total number of time steps in given dataset.

COMPARATIVE ANALYSIS

This section presents the comparative analysis in terms of aforementioned parameters and obtained performance is compared with traditional machine learning classifiers. The obtained confusion matrix is presented in Table 4.

Based on the aforementioned performance measurement parameters, the performance of proposed approach is measured. Table 5 and Figure 4 show the outcome for this experiment.

Further, the training and testing ratios are varied and the obtained performance is presented in Table 6 and Figure 5.

This experiment shows the average performance is obtained as 97,5, 0,95, 0,95, and 0,955 in terms of Precision, Recall, and F1-Score, respectively. However, the decreased training and testing ratio leads to reduce the performance but overall comparative analysis shows that proposed approach achieved better performance for 70-30 train-test ratio. Similarly, average performance of different classifiers is also presented in Table 7 and Figure 6.

Table 4. Obtained Confusion Matrix by Using Proposed Approach (for 70 % Training and 30 % Testing)

	Good	Moderate	Unhealthy	Very Unhealthy
Good	97	2	2	1
Moderate	1	93	2	2
Unhealthy	1	3	94	1
Very unhealthy	1	2	2	96

Table 5. Performance of Proposed Approach for Each Class(in percentage)

Class	Truth	Predicted	Accuracy	Precision	Recall	F1- -Score
Good	100	102	98	95	97	96
Moderate	100	98	97	95	93	96
Unhealthy	100	99	97.25	95	94	94
Very unhealthy	100	101	97.75	95	96	96

Table 6. Performance Analysis of Proposed Approach for Varied Train-Test Ratios (in percentage).

Train-Test Ratio	Accuracy	Precision	Recall	F1-Score
50-50	92,25	92,6	92,5	92,0
60-40	95,94	93,3	94,1	94,6
70-30	97,10	97,1	97,2	97,2
80-20	97,25	98,1	98,1	97,8

Table 7. Performance Measurement and Comparison with Different Classifiers (in percentage).

Classifier	Accuracy	Precision	Recall	F1-Score
SVM	90,93	91,0	90,0	90,0
KNN	89,33	89,3	89,3	89,2
Decision tree	96,64	96,6	96,5	96,4
Random Forest	97,10	97,1	97,2	97,2
Proposed Model	97,50	97,5	97,5	97,65

Further, this experiment is extended for air quality forecasting model and measured the performance in terms of RMSE, MAE and MAPE. In this experiment, first of all we measure the performance for air quality forecasting and obtained performance is presented in Table 7.

Table 7. Comparative Performance for Air Quality Forecasting (in percentage).

Technique	RMSE	MAE	MAPE
Linear regression	137,27	85,54	70,30
Extreme Learning	90,109	53,44	65,87
Neural nets	86,23	48,93	57,93
LSTM	17,55	13,57	7,80
LSTM Encoder Decoder Model	6,332	4,36	2,55
CNN-LSTM	19,74	14,97	11,17
Conv-LSTM	7,47	5,55	2,677
bi-directional LSTM	12,71	10,13	3,17
Proposed Model	8,10	6,15	2,11

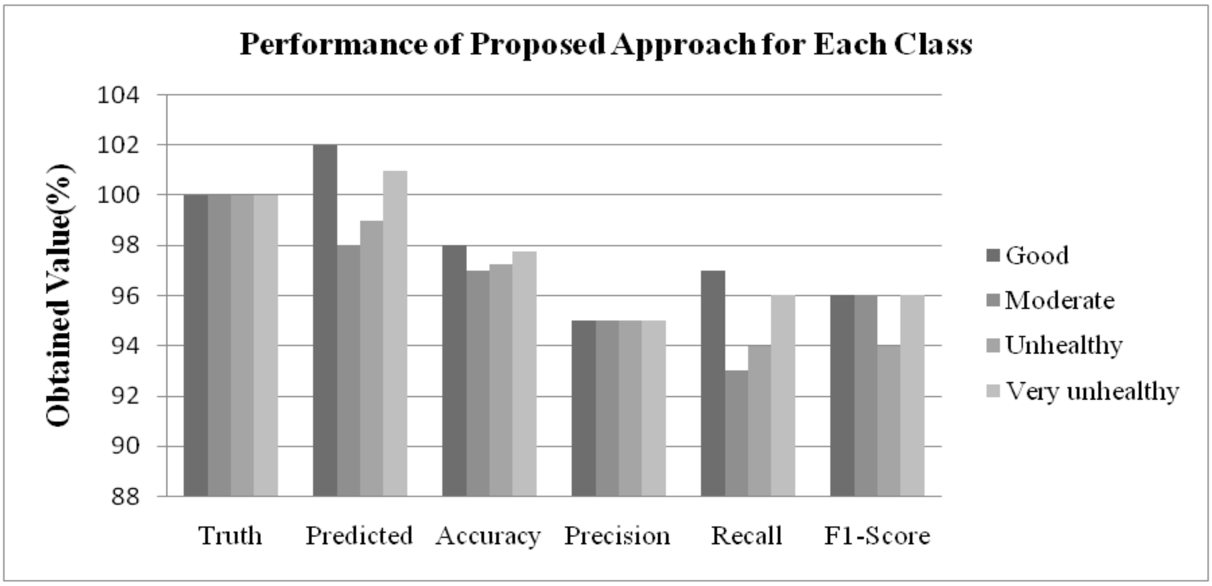

Figure 4. Performance of Proposed Approach for Each Class

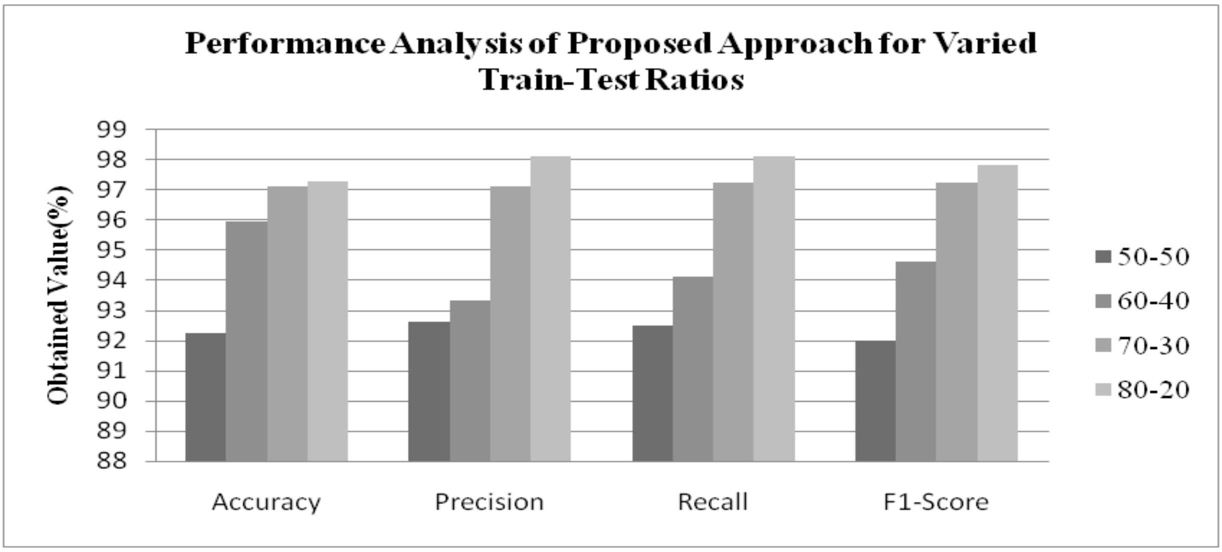

Figure 5. Performance Analysis of Proposed Approach for Varied Train-Test Ratios

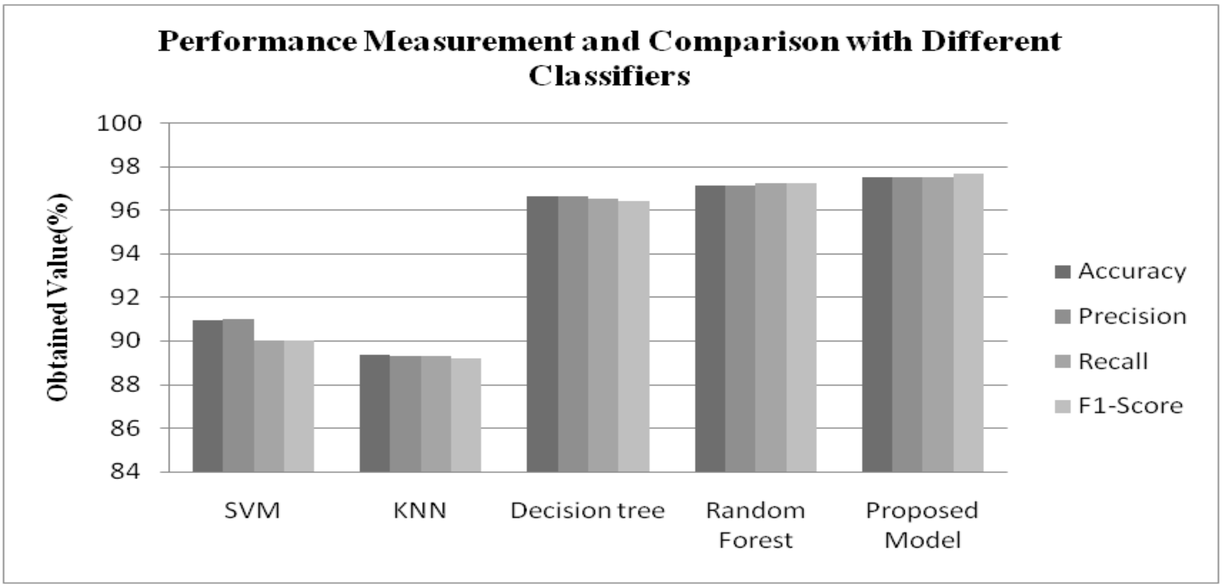

Figure 6. Performance Measurement and Comparison with Different Classifiers.

Finally, a health impact forecasting study on poultry is presented in Table 8 which shows the impact on poultry farms and egg production for different air quality indices. By knowing the air quality index, the poultry farmer can take precautions with respect to health and welfare of chickens farmed in poultry houses. For instance, if air quality is poor or unhealhty, the farmer can provide ventilation and take other actions with respect to welfare of the chickens in poultry farm.

Table 8. Impact of Air Quality on Egg Production.

Air Quality	AQI	Impact on poultry layers in farms
Good	0-50	Quality of air is good for poultry.
Moderate	51-100	Quality of air is moderate. Its an indication for farmers. Moderate effects on poultry layers.
Unhealthy	101-150	Farmers should follow the managemental practices like proper ventilation, litter management etc. This index leads to severe breathing effect on poultry.
Very Unhealthy	151-200	Farmers should take preventive measures to reduce the effect of respiratory diseases which affect the overall health of poultry layers.

CONCLUSION

This article focus on precision poultry farming to cater the health related issues in chickens to increase the meat the egg production. Several methods are present but most of the traditional methods are based on the manual inspection which requires huge amount of time and efforts. Therefore, research community has suggested to incorporate the machine learning based approach for automation of poultry farm monitoring, based on this, computer vision based approaches are introduced however, camera based monitoring systems is a complex task and accuracy also remains a challenging issue. Therefore, a data analysis based approach is presented for poultry farm monitoring by analysing the air quality to improve the health of the chickens. The proposed approach is based on the deep learning based systems where the CNN-LSTM based model is employed to classify the air quality pattern as Good, Moderate, Unhealthy, and Very Unhealthy. The experimental analysis shows that the proposed approach achieves overall accuracy as 97,5 % which can be helpful in appropriate monitoring of poultry chickens.

REFERENCES

[1] Gerland, P., et al.: *World population stabilization unlikely this century*.
Science **346**(6206), 234-237, 2014,
http://dx.doi.org/10.1126/science.1257469,

[2] Russell, E.J.: *World population and world food supplies*.
Routledge, 2019,

[3] Berckmans, D.: *General introduction to precision livestock farming*.
Animal Frontiers **7**(1), 6-11, 2017,
http://dx.doi.org/10.2527/af.2017.0102,

[4] Henchion, M.; McCarthy, M.; Resconi, V.C. and Troy, D.: *Meat consumption: Trends and quality matters*.
Meat Science **98**(3), 561-568, 2014,
http://dx.doi.org/10.1016/j.meatsci.2014.06.007,

[5] Okinda, C., et al.: *A machine vision system for early detection and prediction of sick birds: A broiler chicken model*.
Biosystems Engineering **188**, 229-242, 2019,
http://dx.doi.org/10.1016/j.biosystemseng.2019.09.015,

[6] Berckmans, D.: *Precision livestock farming technologies for welfare management in intensive livestock systems*.
Revue Scientifique et Technique **33**(1), 189-196, 2014,
http://dx.doi.org/10.20506/rst.33.1.2273,

[7] Fournel, S.; Rousseau, A.N. and Laberge, B.: *Rethinking environment control strategy of confined animal housing systems through precision livestock farming*.
Biosystems Engineering **155**, 96-123, 2017,
http://dx.doi.org/10.1016/j.biosystemseng.2016.12.005,

[8] Banhazi, T.M.; Babinszky, L.; Halas, V. and Tscharke, M.: *Precision Livestock Farming: Precision feeding technologies and sustainable livestock production*.
International Journal of Agricultural and Biological Engineering **5**(4), 54-61, 2012,

[9] Lashari, M.H., et al.: *IoT Based poultry environment monitoring system*.
In: 2018 IEEE International Conference on Internet of Things and Intelligence System (IOTAIS). IEEE, pp.1-5, 2018,
http://dx.doi.org/10.1109/IOTAIS.2018.8600837,

[10] Fang, C., et al.: *Pose estimation and behavior classification of broiler chickens based on deep neural networks*.
Computers and Electronics in Agriculture **180**, No. 105863, 2021,
http://dx.doi.org/10.1016/j.compag.2020.105863,

[11] Raj, A.A.G. and Jayanthi, J.G.: *IoT-based real-time poultry monitoring and health status identification*.
In: 2018 11th International Symposium on Mechatronics and its Applications (ISMA). IEEE, pp.1-7, 2018,
http://dx.doi.org/10.1109/ISMA.2018.8330139,

[12] Soliman, E.S. and Hassan, R.A.: *Influence of housing floor on air quality, growth traits, and immunity in broiler chicken farms.*
Advances in Animal and Veterinary Sciences **8**(9), 997-1008, 2020,
http://dx.doi.org/10.17582/journal.aavs/2020/8.9.997.1008,

[13] Schmarje, L.; Santarossa, M.; Schröder, S.M. and Koch, R.: *A survey on semi-, self-and unsupervised learning for image classification*.
IEEE Access **9**, 82146-82168, 2021,
http://dx.doi.org/10.1109/ACCESS.2021.3084358,

[14] Apostolidis, E., et al.: *Video summarization using deep neural networks: A survey*.
Proceedings of the IEEE **109**(11), 1838-1863, 2021,
http://dx.doi.org/10.1109/JPROC.2021.3117472,

[15] Otter, D.W.; Medina, J.R. and Kalita, J.K.: *A survey of the usages of deep learning for natural language processing*.
IEEE Transactions on Neural Networks and Learning Systems **32**(2), 604-624, 2020,
http://dx.doi.org/10.1109/TNNLS.2020.2979670,

[16] Wang, S.; Cao, J. and Yu, P.: *Deep learning for spatio-temporal data mining: A survey*.
IEEE Transactions on Knowledge and Data Engineering **34**(8), 3681-3700, 2022,
http://dx.doi.org/10.1109/TKDE.2020.3025580,

[17] Sharma, e., et al.: *Deep air quality forecasts: suspended particulate matter modeling with convolutional neural and long short-term memory networks*.
IEEE Access **8**, 209503-209516, 2020,
http://dx.doi.org/10.1109/ACCESS.2020.3039002,

[18] Yang, W., et al.: *Detection and analysis of fine particulate matter and microbial aerosol in chicken houses in Shandong Province, China*.
Poultry Science **97**(3), 995-1005, 2018,
http://dx.doi.org/10.3382/ps/pex388,

[19] Lahlouh, I.; Rerhrhaye, F.; Elakkary, A. and Sefiani, N.: *Experimental implementation of a new multi input multi output fuzzy-PID controller in a poultry house system*.
Heliyon **6**(8), No. e04645, 2020,
http://dx.doi.org/10.1016/j.heliyon.2020.e04645,

[20] Al Assaad, D.K., et al.: *A sustainable localised air distribution system for enhancing thermal environment and indoor air quality of poultry house for semiarid region*. Biosystems Engineering **203**, 70-92, 2021, http://dx.doi.org/10.1016/j.biosystemseng.2021.01.002,

[21] Grilli, G., et al. *A pilot study to detect coccidiosis in poultry farms at early stage from air analysis*. Biosystems Engineering **173**, 64-70, 2018, http://dx.doi.org/10.1016/j.biosystemseng.2018.02.004,

[22] Sayour, H.E.M.; Abdel-Megeed, A.H.; Abdel-Khalek, A.M. and Ragai, H.F.: *Evaluation of Real-time Air Quality Monitoring System poultry houses based on wireless sensor network*. In: *1st Conference of Chemistry and Environmental Health*. 2014, http://dx.doi.org/10.13140/2.1.4118.1769,

[23] Jayarajan, P.; Annamalai, M.; Jannifer, V.A. and Prakash, A.A.: *IOT Based Automated Poultry Farm for Layer Chicken*. In: 7th International Conference on Advanced Computing and Communication Systems (ICACCS). IEEE, pp.733-737, 2021, http://dx.doi.org/10.1109/ICACCS51430.2021.9441939,

[24] Hofmann, T., et al.: *Impact of housing environment on the immune system in chickens: a review*. Animals **10**(7), No. 1138, 2020, http://dx.doi.org/10.3390/ani10071138,

[25] Naseem, S. and King, A.J.: *Ammonia production in poultry houses can affect health of humans, birds, and the environment – techniques for its reduction during poultry production.* Environmental Science and Pollution Research **25**(16), 15269-15293, 2018, http://dx.doi.org/10.1007/s11356-018-2018-y,

[26] Pereira, W.F., et al.: *Environmental monitoring in a poultry farm using an instrument developed with the internet of things concept*. Computers and Electronics in Agriculture **170**, No. 105257, 2020, http://dx.doi.org/10.1016/j.compag.2020.105257.