https://doi.org/ 10.33472/AFJBS.6.Si3.2024.436-448

African Journal of Biological Sciences

Journal homepage: http://www.afjbs.com

Research Paper Open Access

Chicken Gender fixation using Machine Learning Techniques

Biswajit Brahma[1]

Senior Data & Visualization Engineer, McKesson Corporation

Narayan Patra[2*]

Department of Computer Science and Information Technology, Siksha 'O'Anusandhan Deemed University, Bhubaneswar, India.

Subrat Kumar Parida[3]

Department of Computer Science and Engineering, Centurion University of Technology and Management, Vizianagaram, Andhra Pradesh, India

Subhendu Kumar Pani[4]

Krupajal Engineering College,Odisha,India

***Corresponding Author-:** Narayan Patra, Department of Computer Science and Information Technology, Siksha 'O'Anusandhan Deemed University, Bhubaneswar, India.

Article History

Volume 6,Issue Si3, 2024

Received:20 Apr 2024

Accepted : 05 May 2024

doi: 10.33472/AFJBS.6.Si3.2024.436-448

Abstract

The demand for poultry meat all over India today encourages to increase the number of farms to meet the consumption capacity of the country. On the other hand, the farm maintains an optimal ratio of female and male chickens to receive good profit and expand the farm's growth. As aresult,modern poultry farmers are now adopting various intelligent technology to make the gender identification of chickenseasy and fast. This paperuses a machine learning (ML) image classification approach to efficiently determine whether the chicken is a hen or a rooster. Further, the dataset wasinitially fitted to the primary image classification ML models and recorded the performance metrics. The convolutional neural network (CNN) model is later considered for investigating better performance than ML models. It is observed that among all the predicted accuracy values, CNN provides the highest accuracy value of 85%

Keywords: confusion matrix, gender identification, image classification, machine learning, poultry farm

1. Introduction

Since the last twenty-year statistical report says that poultry meat production in India is much higher than the production across the globe, due to the population growth in urban and their income capacity, chicken consumption has increased over the year. In India, particularly, the consumption of poultry meat was estimated to be more than four million metric tons in 2022 [1]. On the other hand, as per studies after 2003, though the number of poultry and livestock has been the same, the poultry sector has grown comparatively. In 2019, India claimed the poultry population was more than 800 million, a 16 percent surplus over the last few years.Before five years, the Indian state Tamil Nadu had a population of poultry more than 100 million.

More than 2.5 lacks metric tons were exported by India during 2020-21 to other countries if the poultry product business is concerned [2].Among all the agricultural products, poultry farming is an all-seasonal product for which farmers prefer it. To improve prediction efficiency and commercial profits, gender identification of chickens is more critical. Thus, modern poultry farms are adopting automated sex detection of chicks [3]. The questions arise regarding whether a chicken is a hen or a rooster. There are some methods from which one such way a farmer usually applies for fixing the sex of a chicken, such as waiting until it lays an egg, which means it is a hen; look at the saddle and hackle feathers, which are longer and pointier in roosters; look at the tail shape, which has curved sickle feathers in roosters; look at the breed specific sex characteristics, such as red on the wings for Easter Egger roosters and observe the behaviour, such as crowing, fighting or mating which is more common in rooster [4].

Today it has been observed that most modern poultry farms use artificial intelligence (AI) technology to determine the gender of chickens. Since AI has explored its importance in various sectors such as science, engineering, social media, business, education, entertainment, natural language processing, and healthcare [5], AI could solve many problems in the farming and agricultural sectors economically and within a short period. Due to these reasons, most researchers are now excited to explore the fields where AI is not applied so far.

Nowadays, the advancement of computer technology enlightens ML techniques to make the job easy and cost-effective in diversified research fields [6]. ML is the branch of AI that can train the machine by feeding user data called training data without following any specific algorithm framed by human beings. Intelligently it analyzes data representations and takes action to make predictions and decisions. ML is an essential tool of AI applications because it can build and design models for prediction problems and solve all classification problems based on recognizing the pattern in the input data [7]. The advancement of computers enables machine learning algorithms without following specific instructions, algorithms, and statistical processes to infer from the pattern in the data. The following application-specific learning categories in ML are supervised, semi-supervised, unsupervised, and reinforcement learning [8].

The image classification technique of ML exploited almost all the research filed in computer vision as it can perceive the data content and do its classification. Recently deep learning has been introduced for the problem that could be more effective than others. Thus, CNN addressed this purpose, besides many more classical methods over the years. CNN accepts an input image and predicts a class to which it belongs. In this research, a chicken image is passed to the CNN model; then, this model classifies it and decides whether it is a hen or a rooster. As the physical structure is one of the metrics as shown in Fig.1 below to classify the gender of a chicken, therefore to predict the sex of a chicken, a dataset is generated having two classes of chicken that is hen and rooster. This dataset is fed to the different machine learning models such as Random Forest (RF), K-nearest Neighbor (KNN), Decision Tree (DT), Naïve Bayes (NB), Support Vector Machine (SVM), andCNN [9,10] for gender identification of chicken image dataset. Since deep learning is a more efficient approach in ML for image classification [11,12], CNN is designed and investigated for better results than the other ML models, as discussed above.

According to Y. Uzum et al. [13], the Shape-index (SI) of an egg along with the Random Under-Sampling (RUSBoost) classifier, could predict the gender of a chick.This classifier considers eccentricity, volume, mass, short axis, long axis, ovality, and SI as parameters and classifies the female and male chickensas 80% and 81%, respectively. Vocalizations could be used for determining the gender of one-day-old chicks, as suggested by K. Cuan et al. [14].In this case, extracted audio features could be used as input for the following machine learning models such as CNN, Long Short-TermMemory(LSTM), and Gate Recurrent Unit(GRU). 74.55%, 75.73%, and 76.15%are the accuracies, respectively found from CNN, LSTM, and GRU models. The average accuracy values of gender prediction of chicks are 91.25%, 87.08%, and 88.33%, respectively. Deep neural network with cross-entropy in information theory results in an accuracy of 96.85% by Y. Yao et al. [15]. They have built a chicken gender classification database containing flock images of 800 chickens shot on a farm. Using an object detection network with a gender information label, they included 1000 single images extracted from the flock images.

Zeying Li [3] introduced a method based on chick calls to efficiently to differentiate the sexes of one-day-old chicks. In this case, the calls of chicks are used for fixing their sex using deep learning methods. TheSpectrogram, Cepstrogram, and MFCC+Logfbank are the audio features used as input in CNN, GRN, CRNN,TwoStream, and ResNet-50. The ResNet-50 trained with Spectrogram from three yellow chick's sex provides an accuracy of 95% which is better than all the other results.

Dihua Wu et al. [16] considered the problem of gender detection of chicken by improving the existing ResNet-50 model. They trained and tested the algorithm by collecting publicly available 960 images of hens and roosters from a commercial farm. This algorithm was also compared with AlexNet, GoogleNet, VGG-16, ResNet-18, and DenseNet-201. Among those algorithms, Improved ResNet-50 attains better performance with Accuracy, Precision, Recall, F1 score, and Inference time, respectively of 98.42, 97.92, 98.95, 98.43 percent, and 4.79 ms.

Thavamanin S. et al. [17] used Gray Level Co-occurrence Matrix algorithm (GLCM) for feature extraction and the K-mean algorithm for grouping the cock and hen images. SVM classifier was chosen for training and testing the extracted features. In this case, a dataset of 1000 chicken images is divided into 80:20 percent for training and testing the classifier. They claimed 95.8% of accuracy for classifying both sexes.

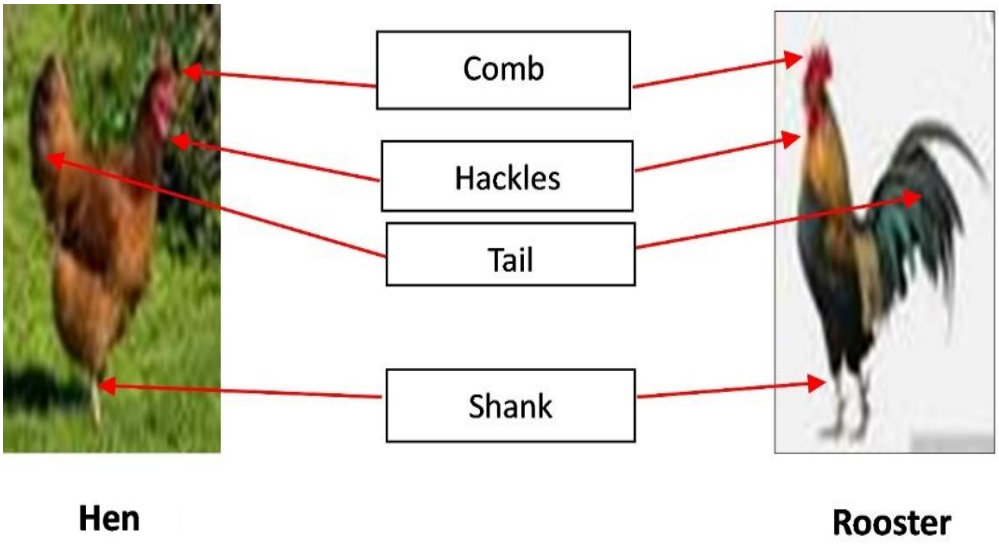

Fig. 1 The visual difference between Hen and Rooster.

In this article, gender identification using the image of a chicken is studied. The overall procedure is described in Fig.2. The forthcoming sections containing literature review, materials and methods, result analysis and discussion, and conclusion with future directions are discussed in section 2, section 3, and section 4, respectively, in this paper.

2. Materials and methods

The gender identification of chickens using the image classification technique is carried out using image datasets of chickens that are publicly available. Since the data size is insufficient to achieve better-predicted value, some downloaded images from the Google site are added. Before using this hybrid nature of the dataset, the data was pre-processed using various existing processes.

- *Data collection*- 499 images of hens and roosters are downloaded from the dataset of the Image CV repository. Since data is not sufficiently available for the application of various classifications of machine learning algorithms, additional 213 images of hens and roosters were downloaded from the Google images website , then created a dataset of chickens by separating hen and rooster in two separate folders with 417 images of hen and 295 rooster images respectively.
- *Data pre-processing*- the dimension of the collected images from the ImageCV repository is 64x64 with three channels Red(R), Green(B), and Blue(B). Since the pictures downloaded from the Google Image site are of various sizes and intensities in color contrast level, resizing all the images to a uniform dimension of 64x64 pixels was required. Fig. 3 displays one of the image processing steps implemented in the proposed work.

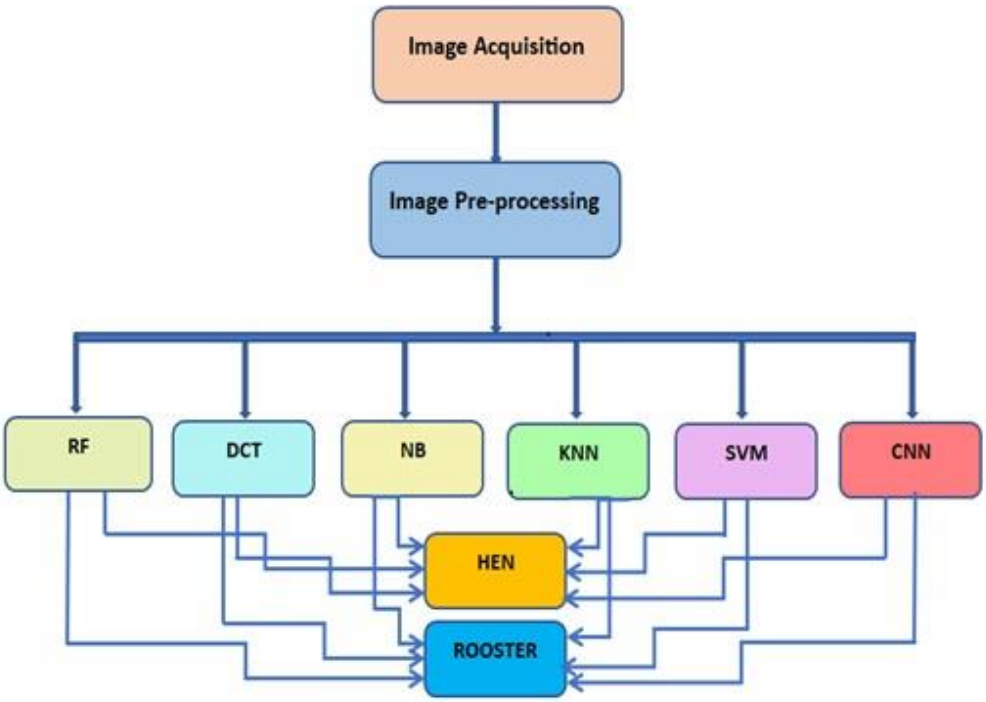

Fig. 2 Overall chicken gender identification pipeline.

Original Hen image

Resized Image

Fig. 3 Image pre-processing.

Flatten the image array to reduce the memory requirement by collapsing multi-dimensional arrays into one-dimensional arrays [18].The dataset is divided into two subsets in the ratio 80:20 corresponding to the training and testing sets.The training dataset and the testing dataset are normalized, dividing each by 255.0.

Five basic ML classifiers are considered for testing the dataset to evaluate the performance. Scikit-Learn library of a Python programming language provides for solving various types of classification problems. Since deep learning recently exploited most of the classification problems, a CNN model is designed, and the model's performance for classification prediction is analyzed.

Random Forest Classifier- RF classifier belongs to the supervised learning category of machine learning and is used for classification, regression, and other problems.

K-nearest neighbor Classifier- KNN classifier is used in pattern recognition, data mining, intrusion detection, and classification problems. It is also a supervised ML algorithm.

Decision Tree Classifier - Most classification and regression problems are handled using a DT classifier, an essential tool of the supervised learning algorithm.

Gaussian Naïve Bayes Classifier-Is based on Baye's Theorem and is preferably used for multi-class classification problems. This learning algorithm belongs to the supervised category. It uses conditional probability theory to perform classification w.r.t input data.

Support Vector Machine-Its accuracy in results and lesser computational time lead to most of the classification and regression problem-solving. This classifier also comes under a supervised learning algorithm.

Convolutional Neural Network- It is an Artificial Neural Network (ANN) using some mathematical operations called CNN is also used for image classification, natural language processing, and image and video recognition. It is a deep learning category of MLtechniques. ANN comprises three layers: the input layer,the hidden layer, and the output layer, as shown in Fig.4.

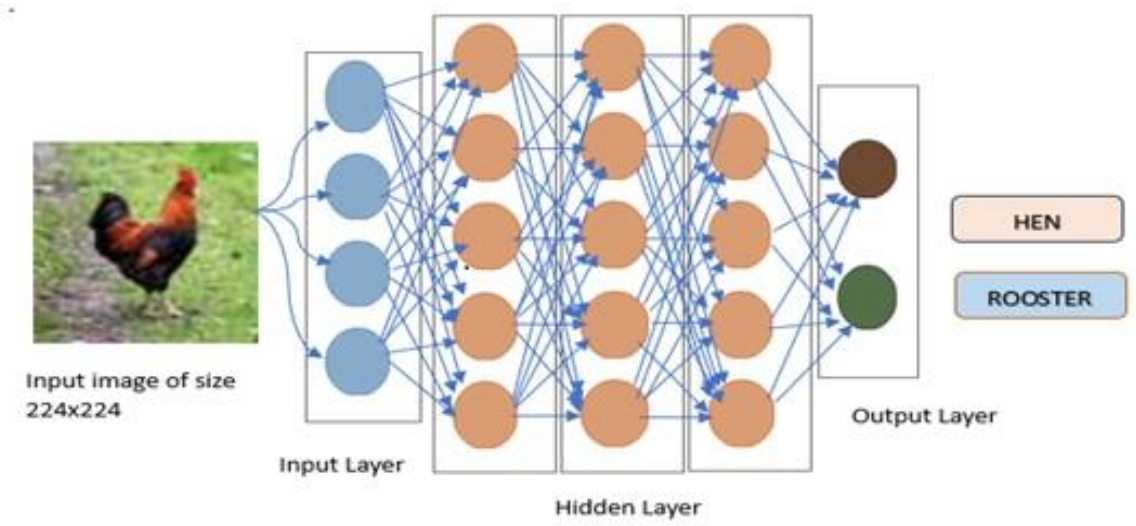

Fig. 4 Artificial neural network.

This paper's proposed CNN model for implementing chicken image classification for gender fixing has nine hidden layers within the input and output layers. There are two convolutional layers, two dropout layers, two max-pooling layers, one flattened layer, and two dense layers, as shown in Fig. 5. In this model, ReLU is used as an activation function in both the convolutional and dense layers. In contrast, the Softmax activation function is used in dense layers at the end of hidden layers. The input shape of the image fed to the first hidden layer is 224x224x3, and the model is passed through trained data using epoch value 100.

The independent and dependent variables in the dataset, called X's and y's, are separated into four subsets known as X_train, X_test, y_train, and y_test using the train_test_splitmethod of the Scikit-Learn library in Python.X's are the independent features used as input to the model, and y's are the dependent variable that are expected outcomes or labels. The training data (X_train,y_train) is passed during the fittingphase so the proposed model can find the parameters that establish a mapping between X's and y's. The correctness of mapping between X's and y's is established using unseen data (test data).The predicted result for X_test is known as the y_pred value. The quality of the prediction value from the model could be quantified using various metrics and scoring. For example, the accuracy score couldjustify the computed accuracy quantity in fractions or the count of correct predictions using the equation (1).

$$\text{Accuracy } (y, \hat{y}) = \frac{1}{n-samples} \sum_{i=0}^{n-samples-1} I(\hat{y}_i = y_i), \text{ where } I(x) \text{ is the indicator function.} \quad (1)$$

The machine learning models are usually evaluated using the model's precision value, recall value, and F1 score through classification reports.

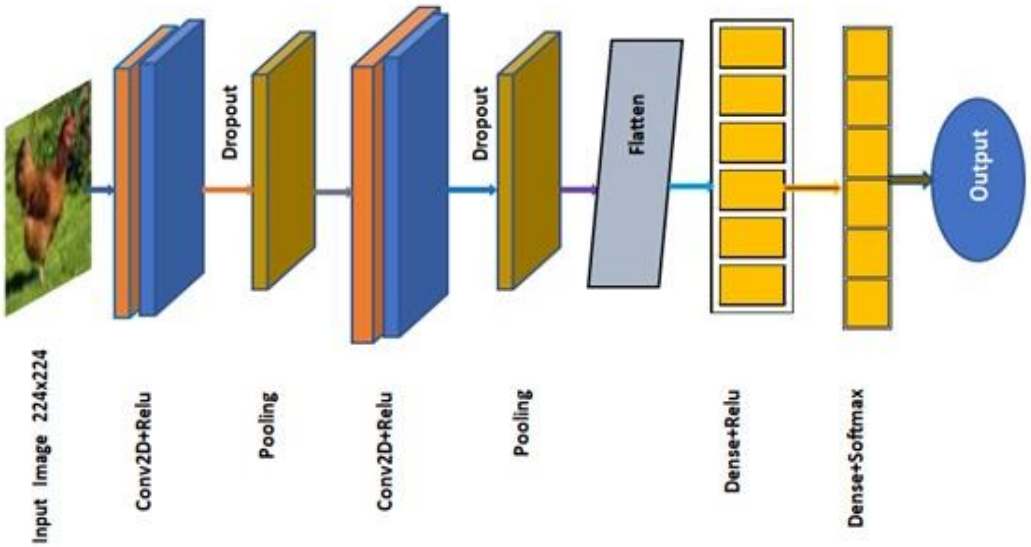

Fig. 5 Convolutional Neural Network with input as chicken image and output decides whether it is Hen or Rooster.

Fig.6shows the layout of the confusionmatrix usedfor illustratingthe classifier's performance depending on True Positive (TP),False Positive (FP), True Negative (TN), and False Negative (FN) values in the case of a binary classification problems [19].

		Actual (True) Values	
		Positive	Negative
Predicted values	Positive	TP	FP
	Negative	FN	TN

Fig. 6 Confusion Matrix.

The mathematical formulae of Precision, Recall, F1-score, and Accuracy are followed in equation (2) .

$$\left.\begin{aligned} \mathrm{Pr}ecision &= \frac{TP}{TP+FP} \\ \mathrm{Re}call &= \frac{TP}{TP+FN} \\ F1-Score &= 2*\frac{\mathrm{Pr}ecison*\mathrm{Re}call}{\mathrm{Pr}ecision+\mathrm{Re}call} \\ Accuracy &= \frac{TP+TN}{TP+TN+FP+FN} \end{aligned}\right\} \quad (2)$$

3. Result analysis and discussions

In this research, the custom dataset is generated by collecting images of chickens from the ImageCV open-source dataset repositoryand some photos downloaded from the Google Image site. Thus, the dataset is formed by data processing, augmentation, and normalization. The standard ML algorithms, such as RF, DCT, SVC, NB, and KNN, are considered to fit

the dataset and predict the performance of image classification-based gender detection for hen or rooster. In addition to the above machine learning algorithms, aCNN is proposed and found better results than the other algorithms. The models' performance evaluation classification reports have been analysed through confusion matrices in Fig.7 to Fig. 12.The performance metrics corresponding to the above models have been shown in Table 1. The graph of comparative studies fortraining vs. validation accuracy concerning Epoch is displayed in Fig.13, and for training vs. validation loss concerning Epoch is depicted in Fig.14. The predicted accuracies from the ML models are compared in Fig.15 for the gender classification of the chicken image dataset. Finally, Fig.16 shows the values of Precision, Recall, and F1-Score obtained from the ML models as discussed above.

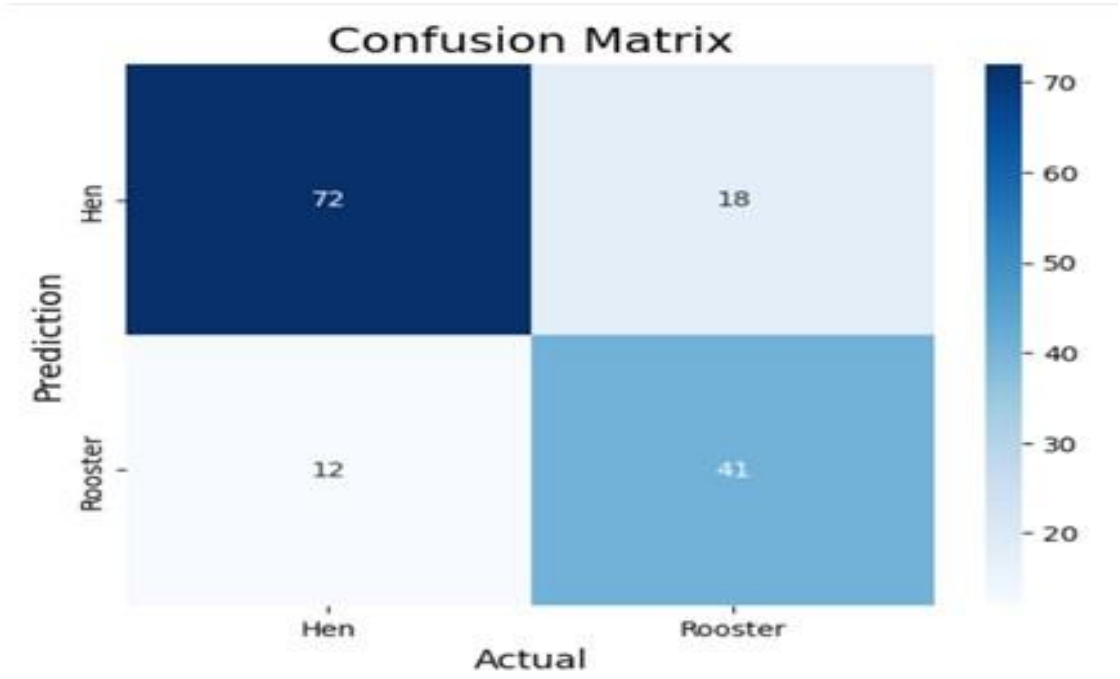

Fig. 7 Confusion matrix for RF classifier.

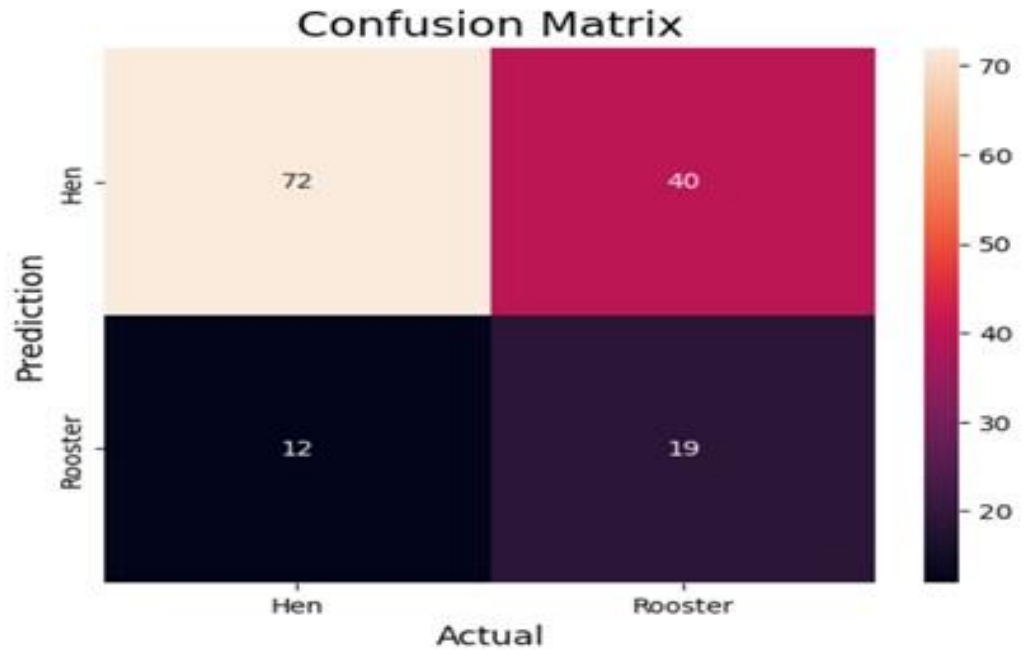

Fig. 8 Confusion matrix for KNN classifier.

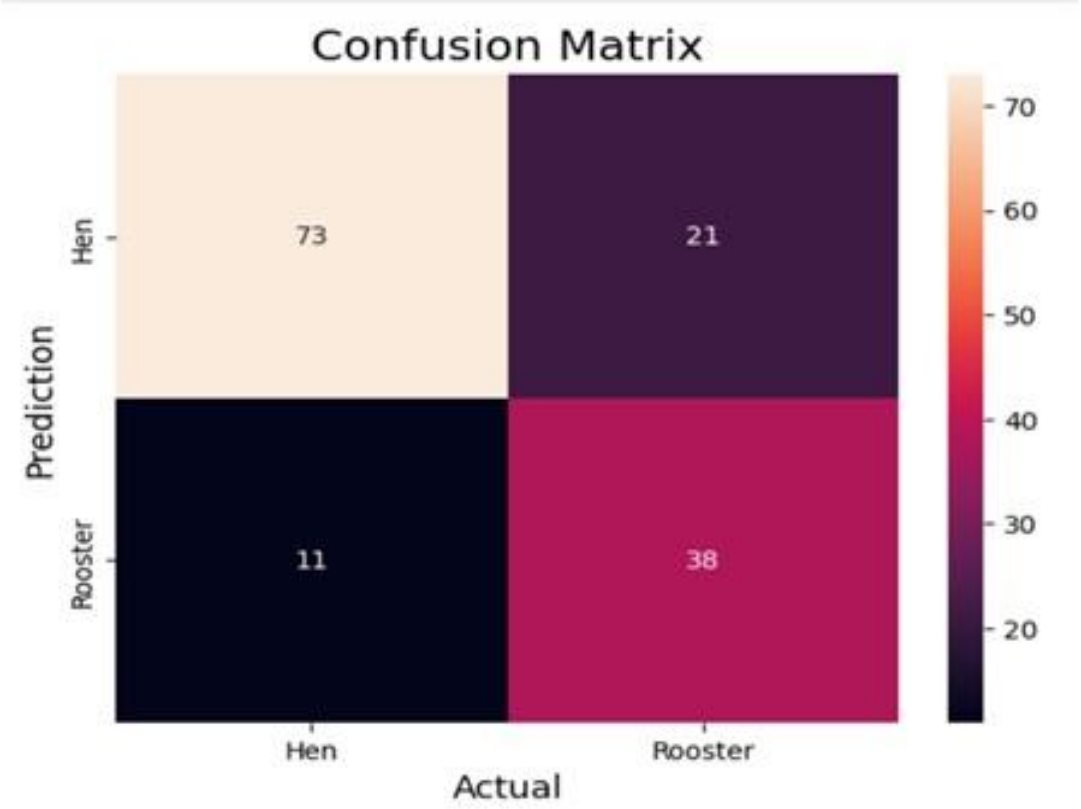

Fig. 9 Confusion matrix for DT classifier.

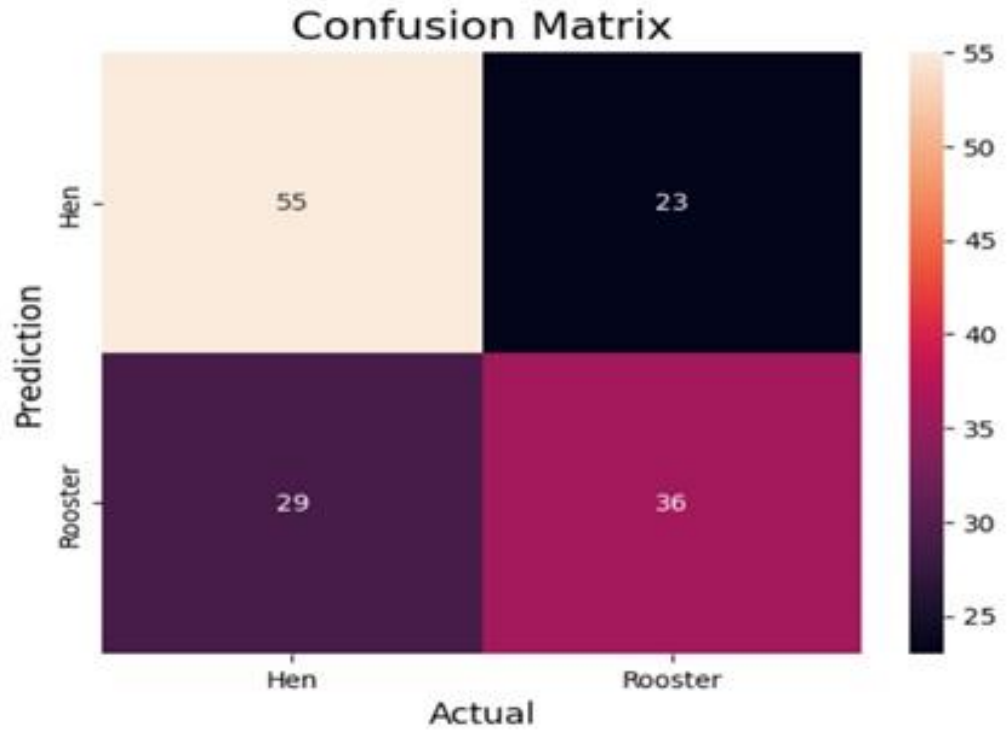

Fig. 10 Confusion matrix for NB classifier.

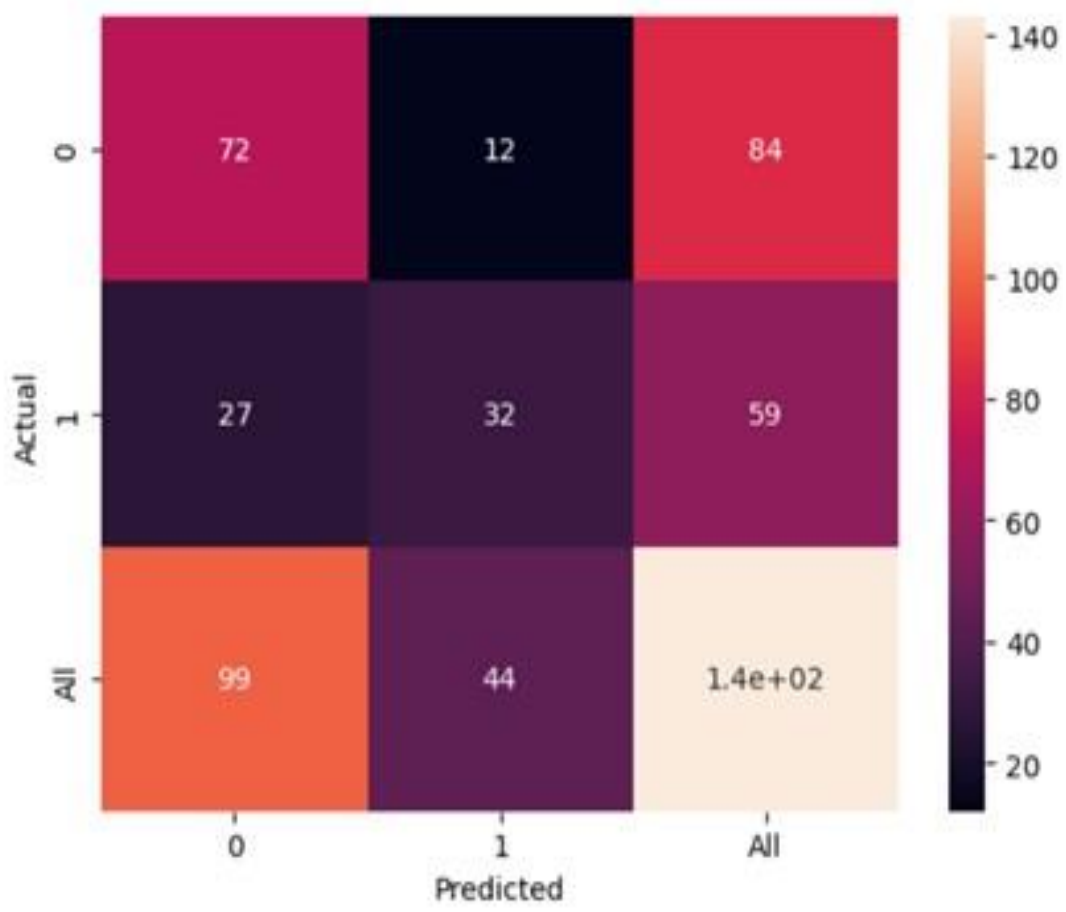

Fig. 11 Confusion matrix for SVC classifier.

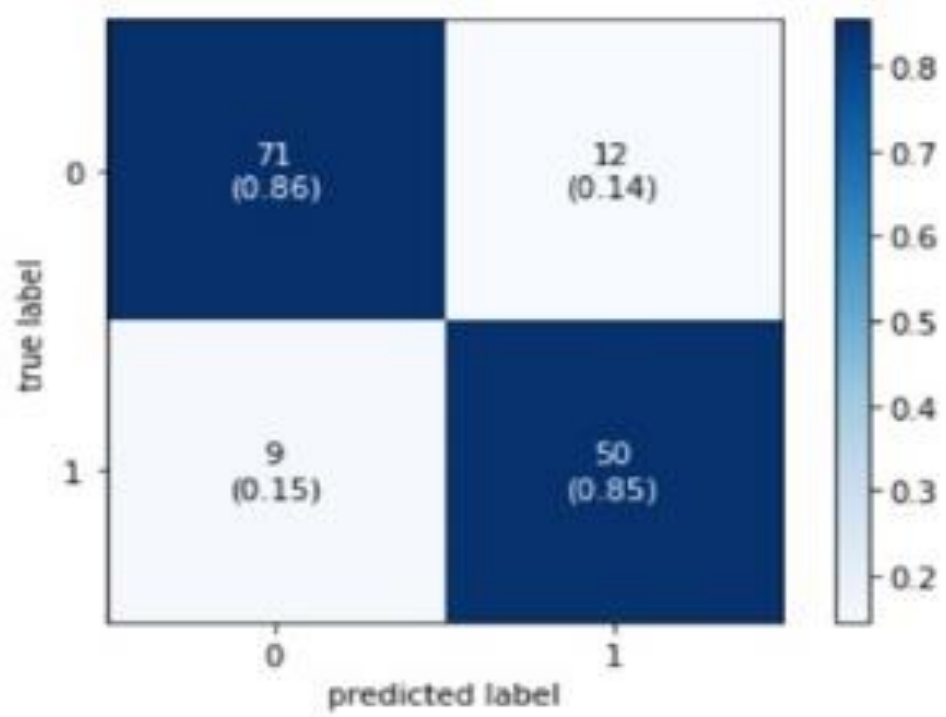

Fig. 12 Confusion matrix for CNN classifier.

TABLE 1. PERFORMANCE OF VARIOUS CLASSIFIERS FOR CHICKEN GENDER IDENTIFICATION.

Machine learning Algorithm	**Class**	**Precision**	**Recall**	**F1-Score**	**Accuracy**
CNN	0	0.89	0.86	0.87	0.85
	1	0.81	0.85	0.83	
RF	0	0.86	0.80	0.83	0.79
	1	0.69	0.77	0.73	
DT	0	0.87	0.78	0.82	0.78
	1	0.64	0.78	0.70	
SVC	0	0.73	0.86	0.79	0.73
	1	0.73	0.54	0.62	
NB	0	0.65	0.71	0.68	0.64
	1	0.61	0.55	0.58	
KNN	0	0.86	0.64	0.73	0.64
	1	0.32	0.61	0.42	

Fig. 13 Analysis of Train vs. Validation accuracy w.r.t Epoch.

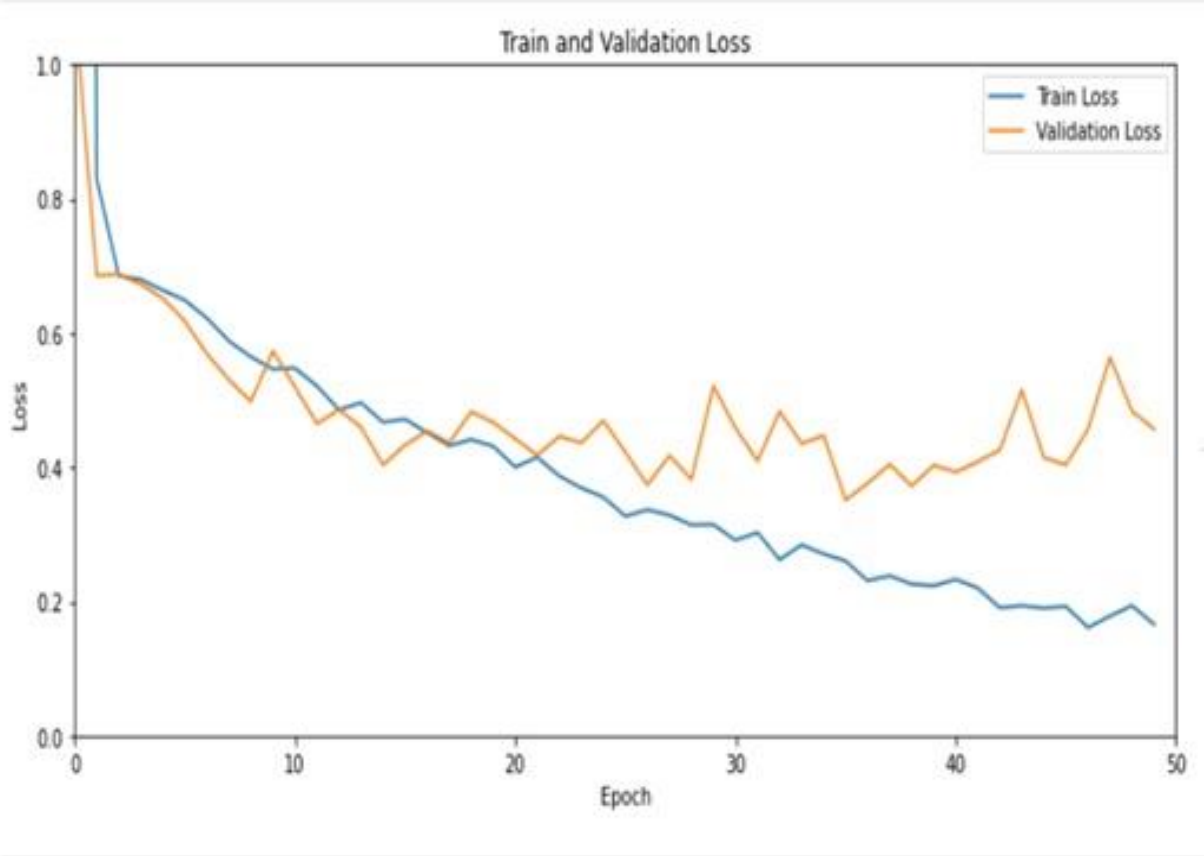

Fig. 14 Analysis of Train vs. Validation Loss w.r.t Epoch.

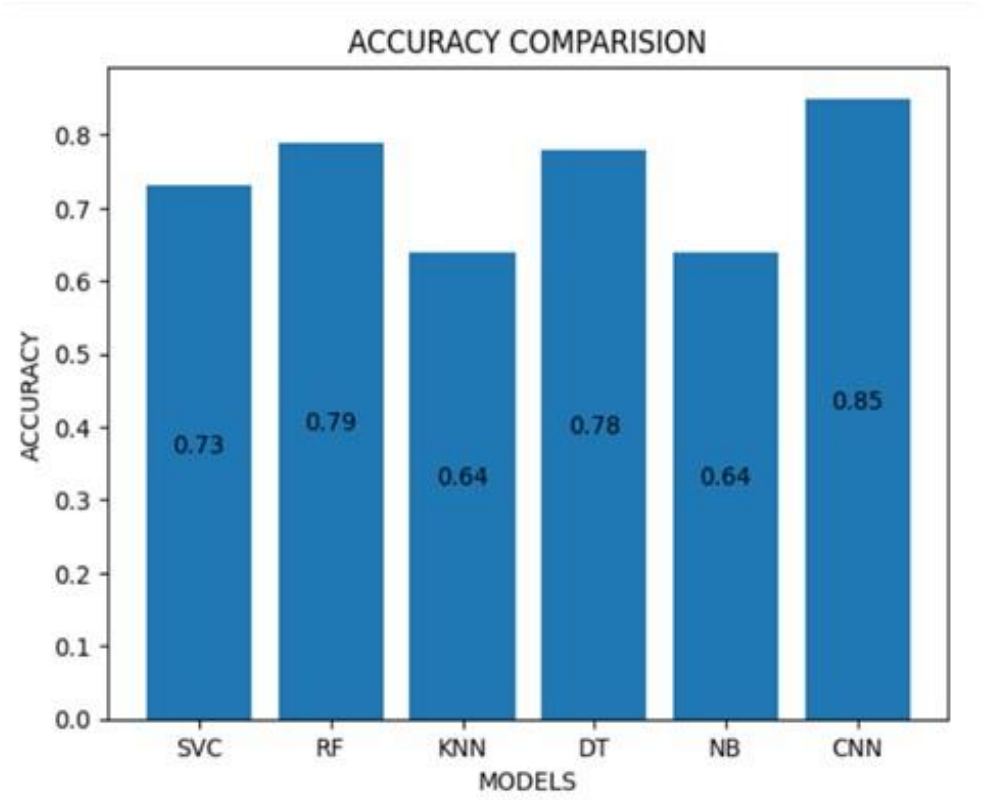

Fig. 15 Accuracy comparison among ML models.

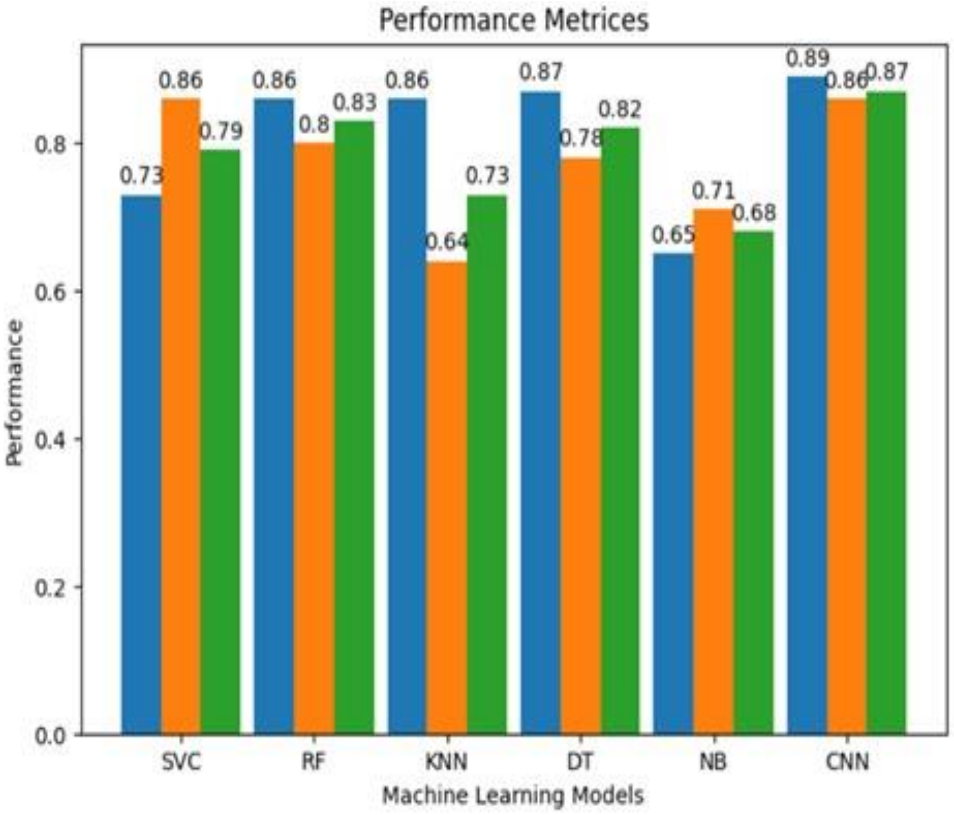

Fig. 16 Performance analysis of ML models.

4. Conclusions

In this paper, gender classification is carried out to identify whether the chicken is a hen or a rooster using ML techniques. The dataset containing two classes of single imagesof a hen and rooster is fitted to the existing five basic algorithms, such as SVC, RF, KNN, DT, NB, and a CNN model. The result evaluated for the prediction probe that CNN is preferable to others. The proposed work finds the accuracy score of CNN is 85%, the highest value among all values tabulated from other ML models. Since CNN needs sufficient data to enhance the performance of prediction performance, more images will be added to the dataset to increase the values of all performance metrics in the forthcoming work.

5. Conflicts of interest

The authors declare no conflict of interest

References

[1] S.Moghana Lavanya, N.Kiruthika, S. Senthilnathan, M.Prahadeeswaran, S Gurunathan,“An overview on Production and Consumption of Chicken Meat in India” , JUST AGRICULTURE multidisciplinary e-newsletter,vol 2, Issue-11,July 2022.

[2] Radha Mani, “Global Agricultural Information Network”, Voluntary Report, September 2021.

[3] Zeying Li, Tiemin Zhang, Kaixuan Cuan, Cheng Fang, Hongzhi Zhao, Chenxi Guan, Qilian Yang, and Hao Qu; “Sex Detection of Chicks Based on Audio Technology and Deep Learning Methods,", Animals, Vol 12, Issuer 22 , Nov 2022.

[4] H. Lukanov, P. Petrov , A. Genchev , E. Halil , N. Ismail ,”Productive Performance of Easter Egger Crosses of Araucana and Schijndelaar Roosters With White Leghorn Hen”, Trakia Journal of Sciences, No 1, pp 72-79, January 2016

[5] Iqbal H. Sarkar, “AI-Based Modelling: Techniques, Applications, and Research Issues Towards Automation, Intelligent and Smart Systems," https://doi.org/10, SN Computer Science,2022.

[6] Dr. O. Obulesu, M. Mahendra, M. Thrilok Reddy, “Machine Learning Techniques and Tools: A Survey”,Proceedings of the International Conference on Inventive Research in Computing Applications (ICIRCA 2018) IEEE Xplore Compliant Part Number: CFP18N67-ART.

[7] Iqbal H. Sarker, “AI-Based Modeling: Techniques, Applications and Research Issues Towards Automation, Intelligent and Smart Systems”, SN Computer Science,Spinger Nature journal, February 2022.

[8] Asongo, A.I, Barma, M, Muazu, H.G; “Machine Learning Techniques, methods, and Algorithms: Conceptual and Practical Insights”, International Journal of Engineering Research and Applications ISSN: 2248-9622, Vol. 11, no.8,pp. 55-64, August 2021.

[9] SehlaLoussaief, Afef Abdelkrim,“Machine Learning framework for image classification," Advances in Science, Technology, and Engineering Systems Journal, Vol. 3, no. 1, pp. 01-10,2018.

[10] Fraunholz, D., Reti, D., Antón, S., Schotten, H.,“Anti-Reconnaissance: Long Short-Term Memory Based Detection, Classification and Mitigation of Hostile Network Exploration. European Conference on Cyber Warfare and Security”, 2018.

[11] Shagun Sharma; Kalpna Guleria; “Deep Learning Models for Image Classification: Comparison and Applications”, 2nd International Conference on Advance Computing and Innovative Technologies in Engineering,2022.

[12] Raza, A., Munir, K., & Almutairi, M., “Ensemble learning-based feature engineering to analyze maternal health during pregnancy and health risk prediction”,PLoS One, 17(11),2022.

[13] Muhammed Kayadan, YunusUzun, “High accuracy gender determination using the egg shape index”, Scientific Reports,Nature article, January 2023.

[14] Kaixuan Cuan, Zeying Li, Tiemin Zhang, Hao Qu, “Gender determination of domestic chicks based on vocalization signals”, Computers and Electronics in Agriculture, July 2022.

[15] Yuanzhou Yao, Haoyang Yu ,Jiong Mu , Jun Li and Haibo Pu, “Estimation of the Gender Ratio of Chickens Based on Computer Vision: Dataset and Exploration”, doi:10.3390/e22070719; Entropy 2020, 22, 719.

[16] Dihua Wu, Yibin Ying, Mingchuan Zhou, Jinming Pan, Di Cui; Improved ResNet-50 deep learning algorithm for identifying chicken; Computers and Electronics in Agriculture. ISSN 0168-1699, Vol. 210, July 2023.

[17] Thavamani S, Vijayakumar J, Sruthi K; “GLCM and K-Means based Chicken Gender Classification”, Smart Technologies, Communication & Robotics, Oct 2021.

[18] [18]Dr.PL. Chithra, P. Bhavani; “A study on various image processing techniques”, International Journal of Emerging Technology and Innovative Engineering, Vol. 5, no.5, May 2019.

[19] [19]Ding, Y., Ma, Q., Liu, F., Zhao, L., Wei, W., “The Potential Use of Salivary miRNAs as Promising Biomarkers for Detection of Cancer: A Meta-Analysis”,PLoS One, 11(11),2016.