Proceedings of Engineering and Technology Innovation, vol. 24, 2023, pp. 73-86

An Effective Supervised Machine Learning Approach for Indian Native Chicken's Gender and Breed Classification

Thavamani Subramani*, Vijayakumar Jeganathan, Sruthi Kunkuma Balasubramanian

Department of Electronics and Instrumentation, Bharathiar University, Coimbatore, India

Received 04 April 2023; received in revised form 07 April 2023; accepted 11 April 2023

DOI: https://doi.org/10.46604/peti.2023.11361

Abstract

This study proposes a computer vision and machine learning (ML)-based approach to classify gender and breed in native chicken production industries with minimal human intervention. The supervised ML and feature extraction algorithms are utilized to classify eleven Indian chicken breeds, with 17,600 training samples and 4,400 testing samples (80:20 ratio). The gray-level co-occurrence matrix (GLCM) algorithm is applied for feature extraction, and the principle component analysis (PCA) algorithm is used for feature selection. Among the tested 27 classifiers, the FG-SVM, F-KNN, and W-KNN classifiers obtain more than 90% accuracy, with individual accuracies of 90.1%, 99.1%, and 99.1%. The BT classifier performs well in gender and breed classification work, achieving accuracy, precision, sensitivity, and F-scores of 99.3%, 90.2%, 99.4%, and 99.5%, respectively, and a mean absolute error of 0.7.

Keywords: native chicken breed classification, gender classification, machine learning algorithms, GLCM, PCA

1. Introduction

Gender and breed predictions for native chickens are important tasks in large-scale poultry production, which require human intervention. Chicken gender identification plays a major role in chicken ratio management, and breed identification is essential in large-scale production to avoid farm-level cross-breeding. With the advancement of machine learning (ML) approaches, ML can be implemented in poultry farms to perform identification activities. This section discusses the types and importance of native chickens, the classification of gender and breed, and the feasible approaches for poultry management.

1.1. Importance of native chicken production

Native chickens have good survival capabilities, and they obtain good yields without any special diet. The native breed can survive with kitchen wastes, insects, and greens [1]. Because of good survivability, harsh environmental adoption, and simple housing facility, rural poultry farmers prefer native breeds' meat and egg. The main motive for native poultry farming is the high monetary value of their meat and eggs. The Tamil Nadu (TN) government also encourages open-source poultry farming with native chicken breeds to maximize income and provide nutritious food to underdeveloped villages. The medium-scale native chicken production is also a key to empowering women and unemployed youths in rural regions of TN [2].

Less production creates demand for native chicken meat and egg in the markets. Therefore, native chicken meat and egg are more expensive than other poultry meat and egg to date. Owing to the benefits for health, consumers are willing to spend more on native breeds of India [3]. Large-scale native-breed poultry farming was introduced to meet the growing demand. In large-scale production houses, the breed selection is based on individual characteristics, including growth period, laying, and

* Corresponding author. E-mail address: thavamani.ei@buc.edu.in

hatching performances. Native Chicken breeds have the advantages of well environmental adaptability, low financial input, suboptimal rearing conditions, abundant protein nutrients in meat and eggs, and genetic diversity. Moreover, native breed hens are good sitters and have good hatching capability and broodiness. [4-5].

1.2. Types of native chicken breeds

The chicken breed is defined according to individual characteristics and features. In India, nineteen varieties of native chicken breeds are registered in the ICAR-National Bureau of Animal Genetic Resources. The following names are the registered chicken breed: Ankaleshwar, Aseel, Busra, Chittagong, Danki, Daothigir, Ghagus, Harringhata black, Kadaknath, Kalasthi, Kashmir Favorolla, Miri, Nicobari, Punjab Brown, Tellichery, Mewari, Kaunayen, Hansli, and Uttara [4].

1.3. Difficulties of identify the chicken gender and breed

In commercially produced large-scale native chicken meat and egg production industries, gender and breed classification is one of the time-consuming tasks for farmers [6]. In manually operated production farms, humans may be easily affected during the poultry pathogens' period [7]. In modernized poultry production industries, these tasks are handled by machines. Since many researchers proposed the artificial intelligence (AI) approach for various management works, this study adopted the ML approach to change the classification of chicken gender and breed from human vision to machine vision.

1.4. Different ML approaches for poultry industry management

The need for human resources in task management will vary based on the poultry production rate, including small, medium, and large scales of native chicken breeds' meat and egg production industries. Backyard farming is unsuitable for large-scale production as it requires more area and human resources for management. In this automated world, different technologies offer reasonable solutions to poultry sector management [8]. ML-based image processing or computer vision approach allows the user to recognize and determine the particular tasks that rely upon the specified tools such as surveillance cameras, sensors, and audio devices.

In general, ML methods have developed from the learning process, and these methodologies need to learn through experiences to perform a particular task [9]. Visual data has trained ML approaches to perform tasks, such as analysis, processing, and recognizing; moreover, various key algorithms are also used for further recognition, object detection, and classification works. In recent years, the poultry industry has started to utilize ML concepts for different purposes, including chicken gender classification [6], chicken floor distribution monitoring [10], and behavior monitoring. These works can be thoroughly distinguished as chicken lameness prediction [11-12], laying hen behavior monitoring [13], vision-based farm management [14], meat safety management [15-17], woody severity identification, and poultry data analysis. Furthermore, earlier chicken disease detection has been implemented to restrict feeding behavior, broiler chicken growth monitoring, and health prediction [18-20]. According to previous studies, the ML approach is suitable for the classification of chicken gender and breed. The proposed ML approach in this study achieves an acceptable accuracy rate through CPU systems.

1.5. Proposed system and contributions

Technological advances have revolutionized the poultry industry's role and structure in India. It became one of the most specialized enterprises in many parts of the country. Many researchers have reported that ML approaches are more suitable and cost-effective for medium and large-scale poultry production industries [6, 10-19]. In both medium and large-scale native chicken meat and egg production farms, the gender and breed identification tasks are done manually. In general, the farmers analyze the gender and breed based on the following features of appearance such as feather color, comb structure, beak shape, skin color, size of the body, and wing. This proposed work is to extract texture and color features from the dataset using the gray-level co-occurrence matrix (GLCM) algorithm to complete this classification work.

2. Literature Review

This section introduces the materials and methods using ML approaches developed by previous researchers. There are different approaches available for poultry task management, such as the classification of chicken gender, monitoring and analysis of chicken activities, and earlier prediction of chicken health, etc. Support vector machine (SVM) and decision tree (DT) classifiers were used to classify the meat sample based on NIR spectral information, pH value, water holding capacity, and meat color [18]. Compared with the SVM classifier, DT obtained effective results of true positive (TP) 0.778, false positive (FP) 0.199, precision 0.747, recall 0.778, and F-measure 0.741 for meat classification.

The feeding and drinking behavior of chicken breeds “Rhode Island Red” and “Barred Plymouth Rock” were analyzed using EthoVision XT 10 software, and the chickens' videos were recorded using infrared-equipped and closed-circuit TV dome cameras with a density of 1.45 pixels/cm^2 or 0.688 cm^2 per pixel [19]. 3D vision and ML approach was developed to detect sick and dead chickens with ZigBee protocol ad-hoc network, and 40,000 mortality sample data was collected through a foot ring sensor [20]. This model measured the 3D displacement of the chicken using a foot ring and calculates the variation through their activity. ML classifiers such as SVM, K-nearest neighbour (KNN), Naive Bayes (NB), DT, and back propagation (BP) were used to analyze the health status of chickens and determine whether they were active, listless, sick, or dead. In this work, BP performed better than the remaining algorithms, for recognition accuracy of active 99.9%, listless 95.2%, sick 95.6%, and dead 100%.

The earlier prediction of chicken lameness using the linear regression (LR) method is necessary to avoid severe chicken injuries [21]. The experiment was carried out on 250 cocks that were 35 days old. The 3D kinetic camera used in this experiment obtained 640 × 480 pixels at a 2 m distance from the object. A double threshold algorithm was used to segment the chicken images with the values of minimum 2,000 and maximum 2,800. The LR classifier obtained an accuracy of 93% in classifying the lame chickens.

A genetic marker with the minimum amount of single nucleotide polymorphism (SNP) was used to identify the Korean native chicken breeds by analyzing the linkage disequilibrium (LD) and genome-wide association study (GWAS) relationship [22]. This work used 283 samples from 20 native Korean chickens. Principle component analysis (PCA) selected the genetic features, and the result of eight ML classifiers was compared. The classifier model AdaBoost (AB) in this research achieved a maximum accuracy of 99.6%.

Mx Control Center v2.4 MOBUTIX AG software and linear real-time model-based chicken health and growth rate monitoring system were developed to identify sick broilers early [23]. The eYeNamic data prediction algorithm was trained and tested using the images obtained from an internet protocol (IP) camera with 1280 by 960 pixels and a 0.5 Hz frame rate for monitoring the chicken. More than five million chicken images were used for training and testing the prediction algorithm. As per the reports, the eYeNamic system could predict sick chickens based on their feeding and drinking behavior and obtained 95.24% accuracy.

Chicken feeding and drinking behavior determine the growth rate [24]. To monitor chickens’ feeding activity, the general linear model was used to assess the relationship between the variables such as occupied area, activity index, and total birds present in an area. The images were analyzed within a rectangular bounding box of 180 × 170 pixels with 16 birds/m^2. The dataset consists of 14,000 broilers aged between 17 to 24 days and obtained 86.4% of prediction accuracy for effective eating.

In feeding behavior analysis, faster region-based convolutional neural network (faster R-CNN) based object detection and SVM-based behavior classification were used [25]. This work used 9,040, 2,260, and 2,826 chicken images for training, validating, and testing. Adjacent frames range from 20 to 150 pixels with five pixels intervals. faster R-CNN-based behavior analyzer obtained 100% precision, recall, and F-score for object detection with 3.1 to 4.2mm root mean square errors (RMSEs). The SVM classifier obtained a value greater than 98% of precision, recall, and F-score with 19 mm RMSEs.

Li et al. [26] followed the faster R-CNN method to examine the spatial and temporal broiler stretching behavior. For this work, 580 × 680 pixels 9,600 images were utilized, including 6,144 for training, 1,536 for validation, and 1,920 for testing the classifier. Two sets of the dataset were used, and the optimal label method with a five-fold cross-validation method was also used to evaluate the detector. The experimental results obtained 99%, 99%, 94%, and 92% accuracy, specificity, recall, and precision in this proposed method.

3. Proposed Methodology

This study extends the investigation using 27 supervised ML classifiers, feature extraction, and feature optimization algorithms for Indian native chicken gender and breed classification work. Fig. 1 depicts six steps for the overall workflow of the proposed work. (1) dataset preparation, (2 and 3) pre-processing, (4) segmentation (5) feature extraction, (6) PCA for feature optimization, training, classifiers training and testing, and performance evaluation.

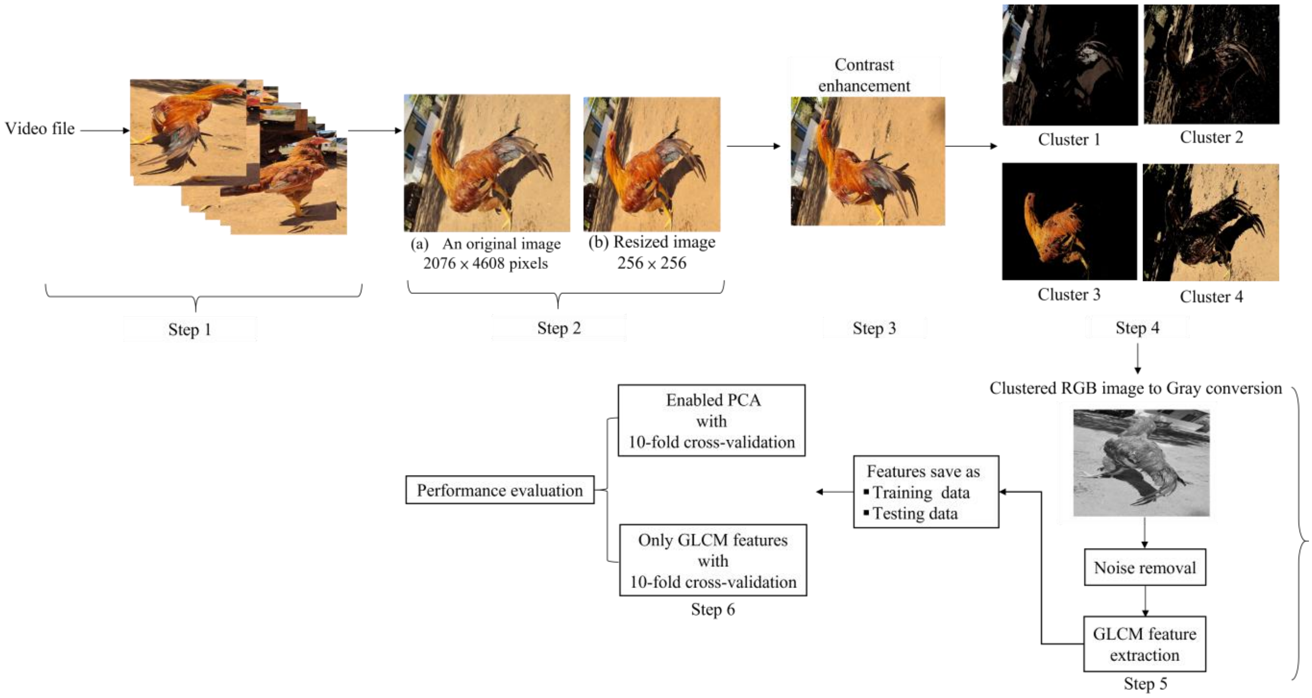

Fig. 1 Classification framework

3.1. Dataset

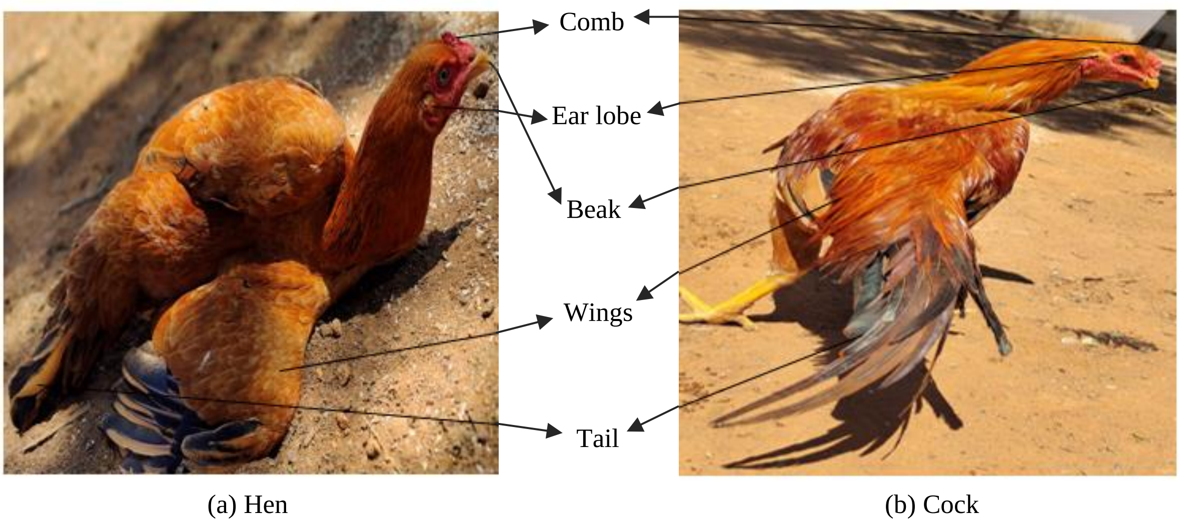

Fig. 2 Common features

Fig. 3 Sample images of the native breed hardware and software requirements

For this investigation, the applied dataset consists of eleven varieties of Indian native chicken images, including Chitta, Chittagong, Hitcari, Kadaknath, Kagar, New Hampshire, Nuri, Parrot Beak Aseel, Peela, Sabja, and Yakub from a rural poultry production houses of Dharmapuri district in TN State. The data collection was done through a smartphone camera, and this proposed work used a 22,000 images dataset that consists of manually labeled chicken gender and breed, including cock and hen in each class. This dataset has 1,000 cock and 1,000 hen images for each breed, 17,600 images for training, and 4,400 for testing (ratio of 80:20).

Figs. 2(a) and (b) indicate common features of hen and cock. Fig. 3 shows the eleven types of native Indian chicken breeds. Each figure mentions the chicken gender: cocks and hens.The classifications of chicken gender and breed initially followed the video-to-image conversion, namely video to frame. The conversion was completed by using free video to JPG converter software (10 frames/sec), and this work was carried out using MATLAB 2021a on a Desktop 3G17UCC with Intel® core ™ i7-7700 processor and 64-bit OS × 64 processors.

3.2. *Pre-processing*

The raw dataset usually contains various noise factors, like irregular intensity distribution resulting from different camera positions during video collection. To avoid diminishing system performance, pre-processing is mandatory after converting the frame from the video files.

In Fig. 1, steps 1 to 4 are important steps in this proposed work. In step 1, the chicken gender and breed dataset is converted from a video file to images (10 frames/sec) indicated. The actual size of the images is different; therefore, optimizing the storage size and minimizing the computational time for image resizing is important. Steps 2 and 3 are contrast enhancement. Step 4 mentions using K-means clustering algorithms to segment the chicken images from the background. In step 5, feature extraction from RGB to grayscale conversion and noise removal techniques are applied.

3.3. *Video-to-image conversion*

Frame extraction is used to segment a long video sequence into the smaller unit for further processing. A smartphone camera was used for data collection, then the dataset was separated manually based on breed features. Here, the Joint Photographic Experts Group converter software application was used for the video file-to-image conversion (10 frames/sec).

3.4. *Image resizing*

In image data, pixel dimension, color depth, and image format will determine the file’s size, which refers to the number of pixels in an image. Image resizing is an essential step in ML. In contrast, image resizing makes the image larger or smaller without removing anything from the original data by changing an image’s height, width, or size. This image-resizing process helps the ML models to train significantly faster. The previous chicken breed images were different in size; in this present dataset, the images are resized in 256 × 256 pixel ranges, and nearest neighbor algorithms were used to resize the images. As the size of the extracted image for texture and color features increased, the values of all features also increased proportionally. Therefore, compared to the size of 126 × 126 pixel range, the optimum size of 256 × 256 used for extraction is better resolution and minimum loss of information.

3.5. *Contrast enhancement*

When resizing the images causes low contrast or blur portions, contrast enhancement techniques are used to overcome them. partial spatial starching algorithms are applied to enhance resized images by stretching and compressing the image without any information loss. This algorithm mapped and stretched the new pixels from a lower threshold value to an upper threshold value, and the remaining pixels were compressed.

3.6. K-means clustering

The K-means clustering algorithm is to cluster the dataset into k-number groups. This clustering function has two stages, the first algorithm calculates the k-centroid, and the next part clusters each point into the nearest centroid points from the dataset. The Euclidean distance method was used to define the nearest centroid (minimum Euclidean distance) point of the given dataset.

Subsequently, the K-means clustering algorithm was used to reduce the sum of the distance from the chicken image to its cluster centroid of all clusters. This process was repeated until getting less error value. Finally, clustered pixels were converted into an image [27]. The clustering images are shown in step 4 in Fig. 1.

3.7. RGB to Grayscale conversion and contrast enhancement

In a pre-processing step, the images were converted from RGB to grayscale, which can be completed by removing the hue and saturation information while remaining the luminance. Improving an image's visibility by adjusting the relative brightness and darkness of the object is called contrast enhancement. The contrast and tone can be modified from the image by mapping the grayscale and adding new values through a grayscale transform.

To enhance the grayscale image contrast, the contrast-limited adaptive histogram equalization (CLAHE) algorithm was utilized. The functionality and efficiency of this algorithm were based on the following parameters, which were the tiles value and clip limits presented in an image. The CLAHE method contrast enhancement was divided into three steps, tile generation, histogram equalization, and interpolation. The image's local contrast was enhanced by the adaptive method. Initially, the algorithm split the images into clear blocks and equalized the histogram. Then, improving the local contrast and edges of all blocks was completed. The main merit of this method is avoiding noise amplification [28].

3.8. Noise removal

Images are distorted due to different types of noise, such as Gaussian noise, salt & pepper noise, Poisson noise, and speckle noise. It may happen during data acquisition, transmission, processing, or coding. The salt & pepper noise will affect the quality of an image. To overcome this issue, the denoising filter can effectively restore appearance in terms of speed, quality, and low complexity [28].

The conversion of images from video files sometimes does simple blurring, embossing, outlining, and sharpening effects; hence, spatial filtering can be applied to identify and clean the object from an image. It replaces every element with a weighted average of its neighborhood pixels. The spatial filtering changes the pixel intensities according to the neighboring pixel intensities for sharpening and smoothing the image.

Based on function, the spatial filter is classified into a smoothing spatial filter for image blurring and noise reduction; and a sharpening filter for removing blurring and highlighting the edges in an image. This filter response is based on neighboring pixels, which is more suitable for removing the impulse or salt & pepper noise from an image.

Image restoration follows the order statistics filter, and this spatial filter function is based on the neighborhood pixels operated by a median filter. The median filter replaces a pixel's value with the gray level median in the neighborhood pixels. This filter has excellent noise-reduction capabilities and effectively reduces salt & pepper and impulse noise without smoothing effects. In this proposed work, the median spatial filter was used for noise removal. The salt & pepper noise reduction method using a median spatial filter can be expressed as:

$$\hat{f}(x, y) = Medium(s,t) \in Sxy\left[g(s,t)\right] \tag{1}$$

where the (x, y) is the spatial domain of image information.

3.9. GLCM algorithm for feature extraction

The features must be derived from the image, created, and stored in a signal format with time and frequency domains. The frequency domain conveys more information about the morphological nature of the image than the spatial domain characteristics [28]. For further work, feature extraction is essential. This work followed the texture and color feature extraction method.

The GLCM algorithms extracted the second-order statistical information from the grayscale level between neighboring pixels representing an image. This feature extraction algorithm is helpful in texture feature analysis by calculating the pairs of pixels with declared values and the spatial relationship between the pictures. The GLCM was constructed by N = 18, 16, and 32 levels of quantization. The image contained objects of a variety of shapes and sizes that were arranged in horizontal and vertical directions. The example specified a set of horizontal offsets that only vary in distance. Second-order metrics analyzed the relationship between pixel pairs. A co-occurrence matrix was used to calculate the texture value presents in images. This matrix is a function of both the angular relationship and distance between two neighboring pixels. It shows the number of occurrences of the relationship between a pixel and its specified neighbor.

GLCM has proved to be a popular statistical method of extracting textural features from images. In the 1973s, Haralick et al. defined the possibilities of the following 14 features that can extract from a single image, including Contrast, Correlation, Energy, Homogeneity, Sum of squares and variance, Entropy, Sum average, Sum entropy, Sum variance, Difference variance, Difference entropy, Information measure of correlation 1, Information measure of correlation 2, and Maximum correlation. The GLCM is the function that can only be performed to extract features on grayscale images. This work extracted the following four features: contrast, correlation, energy, and homogeneity. Table 1 mentions the values of the above-mentioned features for each breed.

Table 1 Four values of GLCM feature samples for each breed

Contrast	Correlation	Energy	Homogeneity	Breed name
0.0813	0.9856	0.2625	0.9712	Kagar cock
0.1441	0.9874	0.1266	0.9357	Kagar hen
0.7335	0.8845	0.0576	0.7652	Nuri cock
0.0597	0.9882	0.2404	0.9702	Nuri hen
0.0591	0.9857	0.1916	0.9705	Peela cock
0.1282	0.9682	0.2147	0.9474	Peela Hen
0.1569	0.9576	0.1537	0.9225	Yakub cock
0.1242	0.9707	0.1687	0.9407	Yakub hen
0.1394	0.9573	0.1678	0.9305	chitta cock
0.2716	0.892	0.2147	0.888	chitta hen
0.0994	0.9645	0.197	0.9526	Chittagong cock
0.2267	0.9622	0.1016	0.8925	chittagong hen
0.1998	0.9752	0.109	0.9071	Hitcari cock
0.1432	0.9531	0.2762	0.931	Hitcari hen
0.1617	0.9796	0.101	0.9217	Sabja cock
0.3202	0.9314	0.1148	0.8561	Sabja hen
0.1279	0.9894	0.1229	0.9378	Kadaknath cock
0.0707	0.993	0.1735	0.9655	Kadakanath hen
0.1806	0.9835	0.2167	0.9208	Parrot Beak Aseel cock
0.3056	0.9498	0.0995	0.8593	Parrot Beak Aseel hen
0.1671	0.979	0.2328	0.9665	New Hampshire cock
0.0398	0.9908	0.5868	0.9811	New Hampshire hen

Contrast: Measures intensity between a pixel and its neighboring pixels presented all over the image. The inertia and variance are called contrast properties. The contrast range of the image is measured in the field between [0 (size)(GLCM,1)-1^2]. The calculation of contrast can be obtained by:

$$Contrast = \sum_{m,n} |m,n|^2 \; p(m,n) \tag{2}$$

where $p(m,n)$ is the spatial relationship between the images.

Correlation: How a pixel related to its neighboring pixels can be obtained using correlation. In the image, the correlation ranges between -1 to 1. The correlation calculation between the pixels presented in an image can be expressed as:

$$Correlation = \sum_{m,n} \frac{(m,\mu m)(n,\mu m)\, p(m,n)}{\sigma m \times \sigma n} \tag{3}$$

where -1 indicates that variables are inversely proportional to each other, and 1 demonstrates that variables are directly proportional to each other.

Energy: In the constant image, if the energy value is 1, the properties and angular second moment are uniform. The energy from the image can be calculated as:

$$Energy = \sum_{m,n} p(m,n)^2 \tag{4}$$

Homogeneity: the returns values that measure the proximity of the distribution of elements in the GLCM to the GLCM diagonal. In images, the homogeneity ranges between 0 and 1, and its value is 1 for a diagonal GLCM. The proximity distribution can be found as:

$$Homogeneity = \sum_{m,n} \frac{p(m,n)}{1+|m,n|} \tag{5}$$

3.10. PCA model for GLCM extracted feature selection

In ML approaches, the PCA model is used for dimensionality reduction [29]. With the help of orthogonal transformation, the PCA converts the observations of correlated features into linearly uncorrelated features, which are called the principal components. This technique reduces the data variance from the given dataset and draws a strong pattern. In this work, the purpose of the PCA model is to select the GLCM extracted features (contrast, correlation, energy, and homogeneity). According to the experimental results of the proposed work, GLCM with four features achieved 99.1% of classification accuracy for BT classifiers. However, the enabled PCA with GLCM features (correlation and homogeneity) achieved a better result of 99.3%. Cross-validation is used to check the ability of the proposed algorithm for new data prediction, which protects against overfitting by portioning the dataset into folders and assessing the accuracy of every folder individually, and this work followed 10-fold cross-validation.

4. Performance Evaluation

In ML techniques, performance evaluation matrices are available for classification works. The balanced dataset used accuracy, recall, and precision matrix for performance evaluation, and the imbalanced dataset used AUC-ROC and Gini coefficient to evaluate the performance. The following matrix evaluated the chicken gender and breed classification work based on accuracy, precision, recall, and F-score. From the classification results, the performance was evaluated based on TP, true negative (TN), FP, and false negative (FN), The evaluation of the classifier's performance can be expressed as:

$$Accuracy = \frac{TN+TP}{FN+TN+TP+FP} \times 100 \tag{6}$$

The accuracy calculation is to identify the appropriate categorization of the classifier by the given samples, and this can be calculated as:

$$Precision = \frac{TP}{TP + FP} \times 100 \tag{7}$$

The ability of optimistic sample prediction by the classifier can be calculated as:

$$Recall = \frac{TP}{TP + FN} \times 100 \tag{8}$$

The recall calculation is to measure the positive sample detection ability of the proposed model. The ratio between the number of optimistic predictions in the total number of positive samples can be calculated as:

$$F\text{-}score = \frac{Recall \times Precision \times 2}{Recall \times Precision} \times 100 \tag{9}$$

5. Experimental Results and Discussion

5.1. Classification results

Supervised ML algorithms were used to recognize the chicken gender and breed with texture and color features extracted by GLCM algorithms. This proposed work performance was tested with two modes, including GLCM with disabled PCA and GLCM with enabled PCA. When PCA is disabled, GLCM defaults to selecting all the GLCM features; When PCA is enabled, the extracted features are selected, optimized, and trained by the classification algorithms. The feature-based homogeneity (used to find similar pixels in a given dataset) and correlation (measures the relationship between the given dataset) obtained high accuracy.

This experiment tested 27 types of classifiers, based on classification results out of 8 class classifier models achieved more than 90% of accuracy. The outcome of the proposed work is that the BT classifier obtained higher performance than the remaining algorithms. The BT classifier obtained 99.1% of accuracy without PCA, and the BT classifier obtained 99.3% accuracy with correlation and homogeneity features enabled PCA.

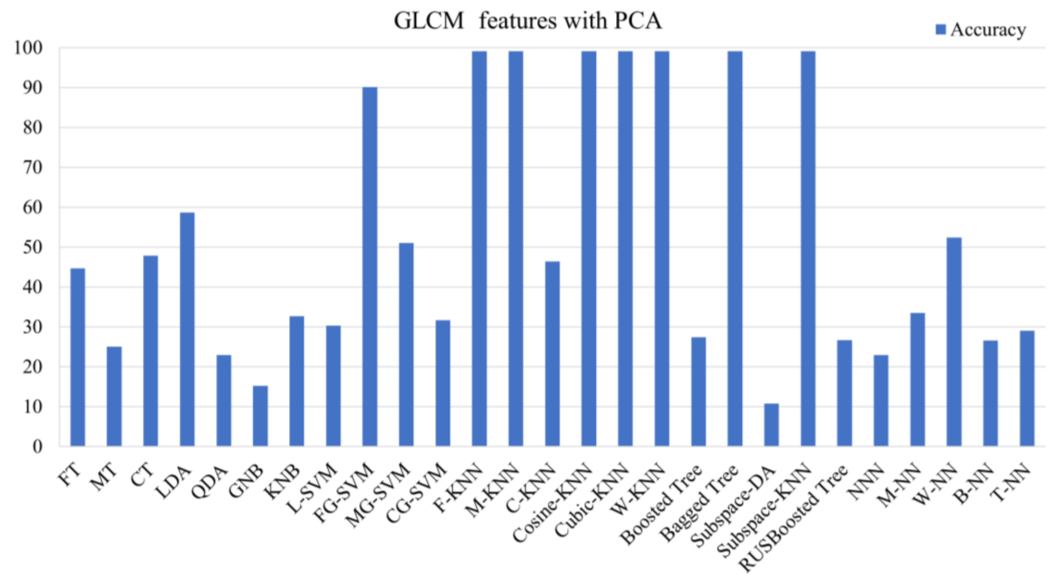

Fig. 4 Classification accuracy of 27 types of classifiers

Based on the classification performance with enabled PCA, the following seven classifiers achieved lower performance metrics than the BT classifiers, including fine Gaussian support vector machine (FG-SVM), fuzzy k-nearest neighbor (F-KNN), multidimensional k-nearest neighbor (M-KNN), Subspace-KNN, weighted k-nearest neighbor (W-KNN), Cosine-KNN, and Cubic-KNN. The FG-SVM obtained 92.4%, 91%, 92%, and 90.1% for accuracy, precision, sensitivity, and F-Score.

The M-KNN model achieved 99.3%, 99.1%, 98.3%, and 98.7% of accuracy, precision, sensitivity, and F-score. F-KNN classifier achieved 99.3% accuracy, 99.4% precision, 99% sensitivity, and 99.2% F-score value. The Subspace KNN model obtained 90.5% accuracy, 99.2 % precision, 99% sensitivity, and 99.1% F-score. W-KNN achieved 99.3%, 99.4%, 99.1%, and 99.2 % of accuracy, precision, sensitivity, and F-score. Cosine-KNN obtained the performance evaluation values of 98.3%, 98.2, 98.2%, and 98.4 for accuracy, precision, sensitivity, and F-score. The Cubic-KNN model achieved 98.4% accuracy, 98.5% precision, 98.1% sensitivity, and 98.2% F-score.

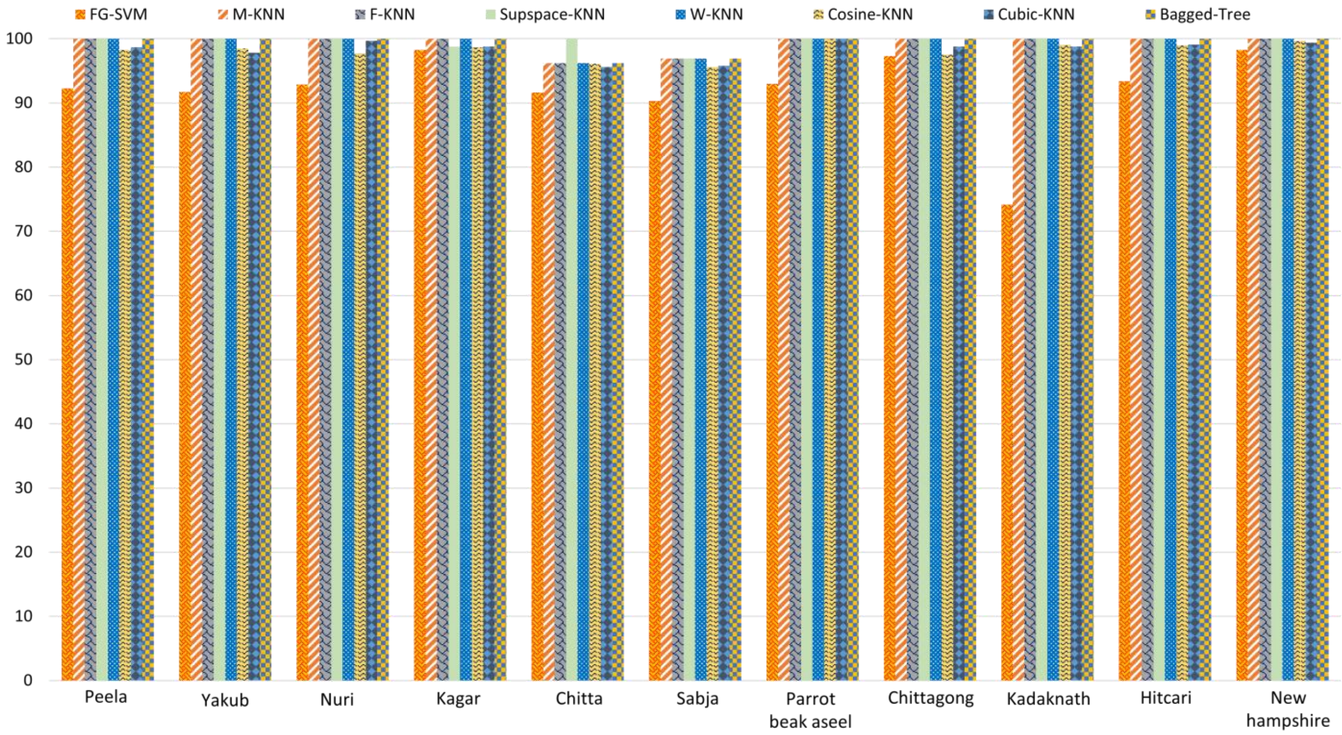

Fig. 5 Native chicken gender and breed classification accuracy of eight types of classifiers

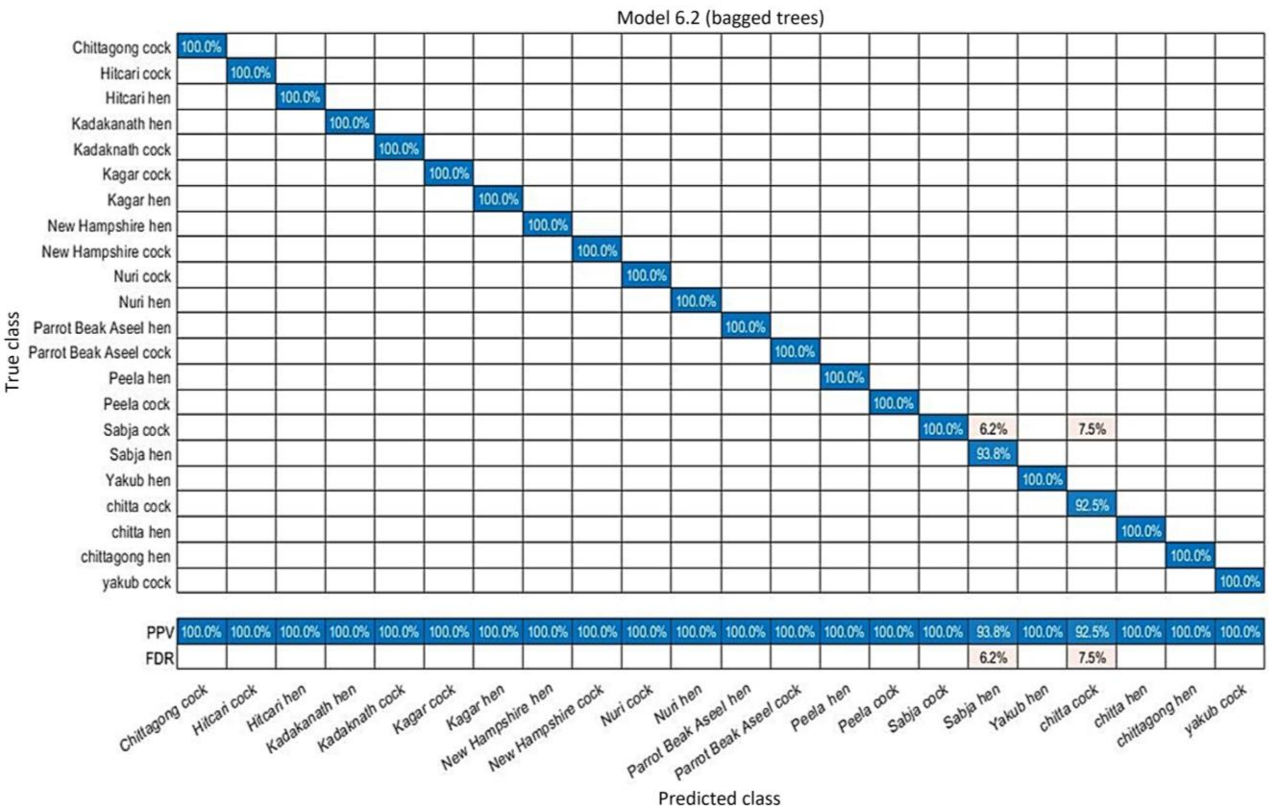

Fig. 6. PPV and FDR percentage of BT classifier

The remaining algorithms performed poorly for the chicken gender and breed classification. Fig. 4 shows the classification accuracy of 27 classifiers GLCM features with an enabled PCA for feature selection. Fig. 5 shows the overall classification accuracy of eleven native breeds. Fig. 6 shows the positive prediction value (PPV) and false detection rate (FDR) of the BT classifier with eleven native chicken breed datasets. Compared to other classifiers, the BT classifier is the best-performed for all gender and breed classification work.

Table 2 shows the type of native chicken gender and breed classification. BT classifier obtained 100% accuracy, precision, sensitivity, and F-score for both cock and hen prediction of Peela, Yakub, Nuri, Kagar, Parrot Beak Aseel, Chittagong, Kadaknath, Hitcari, and New Hampshire breeds. In the Chitta classification, the BT classifier obtained 96.2% of accuracy and precision and 100% of sensitivity and F-score. The Sabja classification obtained 96.9% of accuracy and precision, 94.3% of sensitivity, and 95.5% of F-score.

Table 2 The overall classification performance of the BT classifier

Chicken breed	Accuracy	Precision	Sensitivity	F-Score
Peela	100%	100%	100%	100%
Yakub	100%	100%	100%	100%
Nuri	100%	100%	100%	100%
Kagar	100%	100%	100%	100%
Chitta	96.2%	96.2%	100%	100%
Sabja	96.9%	96.9%	94.3%	95.5%
Parrot Beak Aseel	100%	100%	100%	100%
Chittagong	100%	100%	100%	100%
Kadaknath	100%	100%	100%	100%
Hitcari	100%	100%	100%	100%
New Hampshire	100%	100%	100%	100%
Overall performance	99.3%	90.2%	99.4%	99.5%

5.2. *The accuracy comparison of the model with and without PCA*

This section compares the classification performances of GLCM features and enabled PCA (homogeneity and correlation) models. The chart in Fig. 7 indicates the percentage of accuracy with and without the PCA model for the BT classifier. Accuracy is considered the medium of measuring performance. This proposed method used GLCM for texture and color feature extraction (contrast, correlation, energy, and homogeneity). The GLCM also influenced PCA by selecting extracted features, which enhanced the proposed approach's performance. BT classifier obtained a classification accuracy of 99.1% for GLCM features; however, the enabled PCA with correlation and homogeneity features significantly enhanced the proposed model classification accuracy of 99.3%. Yao et al. [29] proposed a deep learning approach with computer vision, and this work achieved 96.85% accuracy for chicken gender classification. From the results, the proposed model obtained high accuracy for gender and breed classification work.

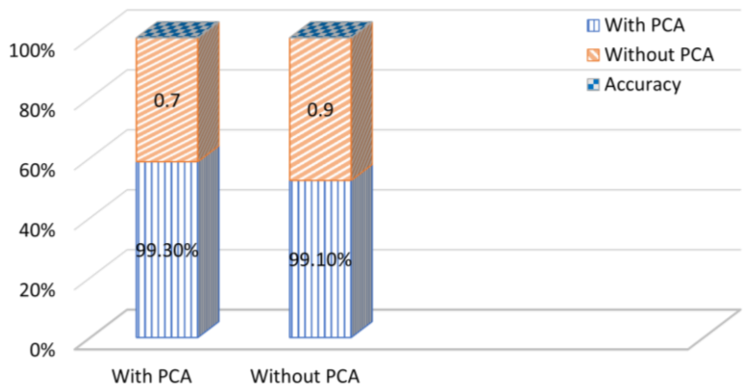

Fig. 7 The obtained accuracy rate of the BT classifier with and without the PCA model

6. Conclusions

A supervised ML approach helps minimize human resources when classifying chicken's gender and breed in the poultry meat and egg production industry. Cock and hen ratio management is an important task in the poultry industry. On the other hand, breed identification can avoid the farm level of cross-breeding.

In this study, a K-means clustering algorithm was used to classify chicken's gender and breed via centroid-based image segmentation, which separated the chicken images from the background. The GLCM approach was employed to extract the second-order statistical feature, including contrast correlation, energy, and homogeneity. After that, the PCA model was utilized to reduce the optimal subset of the above-mentioned features. The combination of GLCM and PCA algorithms can effectively enhance the accuracy of classification.

In addition, the BT classifier achieved accuracy, precision, sensitivity, and F-scores of respective 99.3%, 90.2%, 99.4%, and 99.5% with under 0.7 mean absolute error. The results show that the implementation of ML is useful in the modern large-scale native chicken production industry by utilizing GLCM features with the PCA model and the BT classifier. The classification of chicken's gender and breed with reduced human resources can be accomplished with the proposed work.

This proposed work proves that supervised ML techniques with GLCM futures are suitable for classifying the gender and breed of native chickens. In the future, improving the classification performance without the texture and color features of native chickens can be researched.

Acknowledgment

The authors would like to thank the Department of Science and Technology (Bharathiar University- Grant Number: 12019/IF190264) for funding this project.

Conflicts of Interest

The authors declare no conflict of interest.

Statement of Ethical Approval

All procedures performed in studies involving animals were in accordance with the ethical standards of the institution or practice at which the studies were conducted.

Statement of informed consent

For this type of study, informed consent is not required.

References

[1] M. K. Padhi, "Importance of Indigenous Breeds of Chicken for Rural Economy and Their Improvements for Higher Production Performance," Scientifica, vol. 2016, article no. 2604685, 2016.
[2] M. Kumar, S. P. Dahiya, and P. Ratwan, "Backyard Poultry Farming in India: A Tool for Nutritional Security and Women Empowerment," Biological Rhythm Research, vol. 52, no. 10, pp. 1476-1491, 2021.
[3] R. T. Wilson, "An Overview of Traditional Small-Scale Poultry Production in Low-Income, Food-Deficit Countries," Annals of Agricultural & Crop Sciences, vol. 6, no. 3, article no. 1077, 2021.
[4] S. Haunshi and U. Rajkumar, "Native Chicken Production in India: Present Status and Challenges," Livestock Research for Rural Development, vol. 32, no. 11, article no. 181, November 2020.
[5] M. Kanakachari, H. Rahman, R. N. Chatterjee, and T. K. Bhattacharya, "Signature of Indian Native Chicken Breeds: A Perspective," World's Poultry Science Journal, vol. 78, no. 2, pp. 421-445, 2022.

[6] S. S. Thavamani, J. Vijayakumar, and K. Sruthi, "GLCM and K-Means Based Chicken Gender Classification," Smart Technologies, Communication and Robotics, pp. 1-5, October 2021.
[7] H. M. Hafez and Y. A. Attia, "Challenges to the Poultry Industry: Current Perspectives and Strategic Future after the Covid-19 Outbreak," Frontiers in Veterinary Science, vol. 7, article no. 516, August 2020.
[8] D. K. Dittoe, E. G. Olson, and S. C. Ricke, "Impact of the Gastrointestinal Microbiome and Fermentation Metabolites on Broiler Performance," Poultry Science, vol. 101, no. 5, article no. 101786, May 2022.
[9] C. Okinda, I. Nyalala, T. Korohou, C. Okinda, J. Wang, T. Achieng, et al., "A Review on Computer Vision Systems in Monitoring of Poultry: A Welfare Perspective," Artificial Intelligence in Agriculture, vol. 4, pp. 184-208, 2020.
[10] Y. Guo, L. Chai, S. E. Aggrey, A. Oladeinde, J. Johnson, and G. Zock, "A Machine Vision-Based Method for Monitoring Broiler Chicken Floor Distribution," Sensors, vol. 20, no. 11, article no. 3179, June 2020.
[11] X. Yang, Y. Zhao, G. M. Street, Y. Huang, S. D. Filip To, and J. L. Purswell, "Classification of Broiler Behaviours Using Triaxial Accelerometer and Machine Learning," Animal, vol. 15, no. 7, article no. 100269, July 2021.
[12] P. Pons, J. Jaen, and A. Catala, "Assessing Machine Learning Classifiers for the Detection of Animals' Behavior Using Depth-Based Tracking," Expert Systems with Applications, vol. 86, pp. 235-246, November 2017.
[13] S. M. Derakhshani, M. Overduin, T. G. C. M. van Niekerk, and P. W. G. Groot Koerkamp, "Implementation of Inertia Sensor and Machine Learning Technologies for Analyzing the Behavior of Individual Laying Hens," Animals, vol. 12, no. 5, article no. 536, March 2022.
[14] N. Sharma, R. Sharma, and N. Jindal, "Machine Learning and Deep Learning Applications-A Vision," Global Transitions Proceedings, vol. 2, no. 1, pp. 24-28, June 2021.
[15] A. Siddique, S. Shirzaei, A. E. Smith, J. Valenta, L. J. Garner, and A. Morey, "Acceptability of Artificial Intelligence in Poultry Processing and Classification Efficiencies of Different Classification Models in the Categorisation of Breast Fillet Myopathies," Frontiers in Physiology, vol. 12, article no. 712649, 2021.
[16] S. Barbon, A. P. A. Da Costa Barbon, R. G. Mantovani, and D. F. Barbin, "Machine Learning Applied to Near-Infrared Spectra for Chicken Meat Classification," Journal of Spectroscopy, vol. 2018, article no. 8949741, 2018.
[17] M. Pitesky, J. Gendreau, T. Bond, and R. Carrasco-Medanic, "Data Challenges and Practical Aspects of Machine Learning-Based Statistical Methods for the Analyses of Poultry Data to Improve Food Safety and Production Efficiency," CABI Reviews, vol. 2020, https://doi.org/10.1079/PAVSNNR202015049, January 2020.
[18] X. Zhuang, M. Bi, J. Guo, S. Wu, and T. Zhang, "Development of an Early Warning Algorithm to Detect Sick Broilers," Computers and Electronics in Agriculture, vol. 144, pp. 102-113, January 2018.
[19] B. Milosevic, S. Ciric, N. Lalic, V. Milanovic, Z. Savic, I. Omerovic, et al., "Machine Learning Application in Growth and Health Prediction of Broiler Chickens," World's Poultry Science Journal, vol. 75, no. 3, pp. 401-410, September 2019.
[20] I. Elawady and C. Özcan, "A New Effective Denoising Filter for High Density Impulse Noise Reduction," Turkish Journal of Electrical Engineering and Computer Sciences, vol. 30, no. 4, article no. 15, 2022.
[21] T. Babu, T. Singh, D. Gupta, and S. Hameed, "Optimized Cancer Detection on Various Magnified Histopathological Colon Imagesbased on DWT Features and FCM Clustering," Turkish Journal of Electrical Engineering and Computer Sciences, vol. 30, no. 1, article no. 1, 2022.
[22] P. K. Bhagat, P. Choudhary, and Kh. Manglem Singh, "A Comparative Study for Brain Tumor Detection in MRI Images Using Texture Features," Sensors for Health Monitoring, vol. 5, pp. 259-287, 2019.
[23] B. T. Jijo and A. M. Abdulazeez, "Classification Based on Decision Tree Algorithm for Machine Learning," Journal of Applied Science and Technology Trends, vol. 2, no. 01, pp. 20-28, March 2021.
[24] Y. S. Can, "Stressed or Just Running? Differentiation of Mental Stress and Physical Activityby by Using Machine Learning," Turkish Journal of Electrical Engineering and Computer Sciences, vol. 30, no. 1, article no. 21, 2022.
[25] G. Li, X. Hui, Z. Chen, G. D. Chesser Jr, and Y. Zhao, "Development and Evaluation of a Method to Detect Broilers Continuously Walking around Feeder as an Indication of Restricted Feeding Behaviors," Computers and Electronics in Agriculture, vol. 181, article no. 105982, February 2021.
[26] G. Li, Y. Zhao, Z. Porter, and J. L. Purswell, "Automated Measurement of Broiler Stretching Behaviours under Four Stocking Densities via Faster Region-Based Convolutional Neural Network," Animal, vol. 15, no. 1, article no. 100059, January 2021.
[27] N. Dhanachandra, K. Manglem, and Y. J. Chanu, "Image Segmentation Using K -Means Clustering Algorithm and Subtractive Clustering Algorithm," Procedia Computer Science, vol. 54, pp. 764-771, 2015.
[28] Priyanka and D. Kumar, "Feature Extraction and Selection of Kidney Ultrasound Images Using GLCM and PCA," Procedia Computer Science, vol. 167, pp. 1722-1731, 2020.
[29] Y. Yao, H. Yu, J. Mu, J. Li, and H. Pu, "Estimation of the Gender Ratio of Chickens Based on Computer Vision: Dataset and Exploration," Entropy, vol. 22, no. 7, article no. 719, July 2020.