BIOMETRIC IDENTIFICATION OF WATER BUFFALO USING REAL-TIME FACE RECOGNITION

N. H. Tuvay[1], O. Ermetin[1*] and T. Hayit[2]

[1]Department of Animal Science, Faculty of Agriculture, Yozgat Bozok University, Yozgat 66900, Türkiye
[2]Department of Cardiology, School of Medicine, Wake Forest University, Winston-Salem, NC 27101, United States

*Corresponding author's email: orhan.ermetin@yobu.edu.tr

NHT,ORCID: https://orcid.org/0000-0002-7603-8721; TH,ORCID: https://orcid.org/0000-0001-5367-7988; OE, ORCID:https://orcid.org/0000-0002-3404-0452

ABSTRACT

Face recognition is increasingly used for biometric identification in both humans and animals, offering a non-invasive option for managing hard-to-handle species such as water buffalo. We present a real-time face recognition system based on YOLOv5 to accurately identify individual buffalo in a livestock setting. Our goal is to support precision livestock farming with an efficient and scalable monitoring solution. We also introduce Buffalo-22, a dataset of 4000 augmented face images from eight water buffalo collected at a breeding farm in Yozgat, Turkey. We compare our system with a traditional approach based on Local Binary Patterns in the HSV color space (LBP-HSV) to highlight the advantages of deep learning in challenging agricultural environments. The proposed YOLOv5 model achieved 99.3% average precision at IoU 0.5 (mAP@0.5), while the LBP-HSV model reached 87.8% test accuracy. These results show that convolutional neural networks can deliver accurate, real-time face recognition for water buffalo. The reported performance indicates that the proposed system is a practical tool for automatic and reliable identification on livestock farms.

Keywords: Biometric identification, face recognition, water buffalo, deep learning, YOLOv5

https://doi.org/10.36899/JAPS.2026.6.0130 Published first online August 08, 2026

INTRODUCTION

Animal identification methods are commonly grouped into mechanical, electronic, and biometric categories (Marchant, 2002). In practice, metal or flexible plastic ear tags, subcutaneous passive RFID transponders, and rumen boluses are widely used. However, the application of electronic identification methods requires an expert practitioner and may occasionally lead to severe accidents resulting in animal injury or death (Bugge *et al*., 2011; Lu *et al*., 2014). Hence, biometric identification has begun to gain attention alongside the traditional methods to reduce the risks and requirements and increase manageability. Biometrics leverage unique anatomical and behavioral traits, including visual coat patterns, muzzle prints, iris and retinal patterns, face features, ear vein patterns, bite marks, DNA, and behavior or movement signatures (Nilsson *et al*., 2006; Barry *et al*., 2008; Burghardt, 2008; Erdem and Tuna, 2008; Awad *et al*., 2013; Noviyanto and Arymurthy, 2013; Kumar and Singh, 2014). Animal welfare concerns also require that identification devices or methods should not cause pain or distress. This has increased interest in machine-learning-based frameworks for automatic identification.

Animal biometrics rely on features that can be observed and measured, such as appearance, movement and vocal sounds. These data are usually collected with sensors, compared with stored patterns and then used by other systems or applications (Kühl and Burghardt, 2013; Çelikyürek and Karakuş, 2017).

A growing number of computer vision studies now use deep learning methods. Among the most common are convolutional neural networks (CNNs) and transformers, which have been used for detection, classification, recognition and tracking (Zhang *et al*., 2018; Bezen *et al*., 2020; Erbay and Hayit, 2025). Similar methods have also been tested in livestock research, especially for identifying cattle from face images. Although the reported results are generally strong, one practical problem is the amount of labeled data needed to train these models (Qiao *et al*., 2019; Yang *et al*., 2019; Yao *et al*., 2019; Wang *et al*., 2020; Xu *et al*., 2022). Individual animal recognition can support several farm-management tasks, including disease monitoring, vaccination records, production management, animal tracking and owner identification (Awad, 2016).

In recent years, more studies have applied deep learning to livestock identification. Researchers have used cattle face images, recordings from several cameras and depth data to tell individual animals apart. Many of these studies have reported a high accuracy (Bergman *et al*., 2024; Ermetin and Örnek, 2025; Sharma *et al*., 2025; Yu *et al*., 2025).

Water buffalo (*Bubalus bubalis*) are the second most important dairy species in the world after cattle. Most buffalo studies have focused on genetics, nutrition and breeding. Much less attention has been given to intensive buffalo production systems and to the use of precision livestock farming (PLF) tools on buffalo farms. These technologies could be useful in managing silent estrus, infertility, heat stress and problems related to milking (Kul *et al*., 2018; Ermetin, 2021).

Buffalo are not always easy to approach under farm conditions and in some cases this may involve a safety risk. Reading an ear tag or carrying out a routine procedure can lead to a defensive response, particularly when the animal is handled by someone unfamiliar (Kul, 2020). Housing conditions that allow more natural behavior and reduce unnecessary close contact may help lower fear and aggression and may also reduce problems with milk let-down (Ermetin, 2017). Computer vision offers a way to identify and monitor animals without direct contact. Ear tags can sometimes be read from a distance, but they may also be lost, damaged or difficult to read. Another possibility is to identify each animal from its face and connect this result with records on pedigree, milk yield and breeding history. Developing models directly for water buffalo could therefore make PLF systems easier to use in practice, using either deep learning or traditional machine-learning methods (Ermetin, 2021).

Although several biometric modalities such as muzzle, iris, and retinal patterns have been explored, face recognition for water buffalo has received limited attention. Feature-based methods and simple classifiers remain attractive where resources are constrained and often perform well under controlled conditions, but their accuracy and scalability tend to degrade in dynamic, real-time farm environments.

In this study, we developed a real-time face recognition system for individual water buffalo identification based on YOLOv5, a state-of-the-art deep learning object detection algorithm. We also introduce Buffalo-22, a curated dataset comprising 4000 augmented facial images from eight buffalo collected on a breeding farm in Yozgat, Turkey. Alongside the proposed model, we implement a traditional baseline that combines LBP features from HSV and grayscale images with an SVM classifier, enabling a side-by-side comparison of accuracy, throughput, and practical deployability. We hypothesize that a YOLOv5-based approach trained on Buffalo-22, a curated dataset containing augmented facial images of water buffalo, will outperform a conventional machine learning baseline combining Local Binary Pattern–Hue Saturation Value (LBP-HSV) feature extraction with a Support Vector Machine (SVM) classifier in terms of identification accuracy and real-time efficiency. Furthermore, compared with conventional farm identification methods, the proposed system is expected to provide a fast, low-cost, and robust alternative.

MATERIALS AND METHODS

Dataset

Image Acquisition: In this study, eight healthy water buffalo aged between 2 and 4 years were used on a private water buffalo farm located in Yozgat, Turkey. Face images were captured from different angles in RGB (red, green, blue) format using a Canon EOS 1200D DSLR camera between May and June 2023. Sample images from the data collection phase are presented in Fig. 1.

Table 1. Raw dataset statistics

Buffalo ID	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**Total**
Number of images	422	384	192	301	277	286	193	269	**2324**

Data Annotation and Augmentation: First, the raw images were annotated using the Roboflow labeling tool by manually drawing bounding boxes around the face regions of each water buffalo. Annotations were used to train the face detection model (Dwyer and Gallagher, 2023). Insufficient raw images can lead to overfitting or underfitting during training. To avoid data leakage between the training, validation and test sets, the original non-augmented images were first divided into training, validation and test subsets at a ratio of 70%, 20% and 10%, respectively. This split was performed at the original image level and stratified according to buffalo identity. After this separation, data augmentation was applied only to the training subset using standard techniques, including rotation, horizontal and vertical flipping, and adjustments to brightness and contrast. The validation and test subsets were not augmented and were kept unchanged for unbiased model evaluation. An example of a water buffalo image and its augmented variants is shown in Fig. 2.

Fig. 1. Image acquisition process

As a result of the data collection process, a total of 2324 images were obtained. Individual buffalo were labeled with IDs from 0 to 7. The number of raw (unprocessed) images per animal is summarized in Table 1.

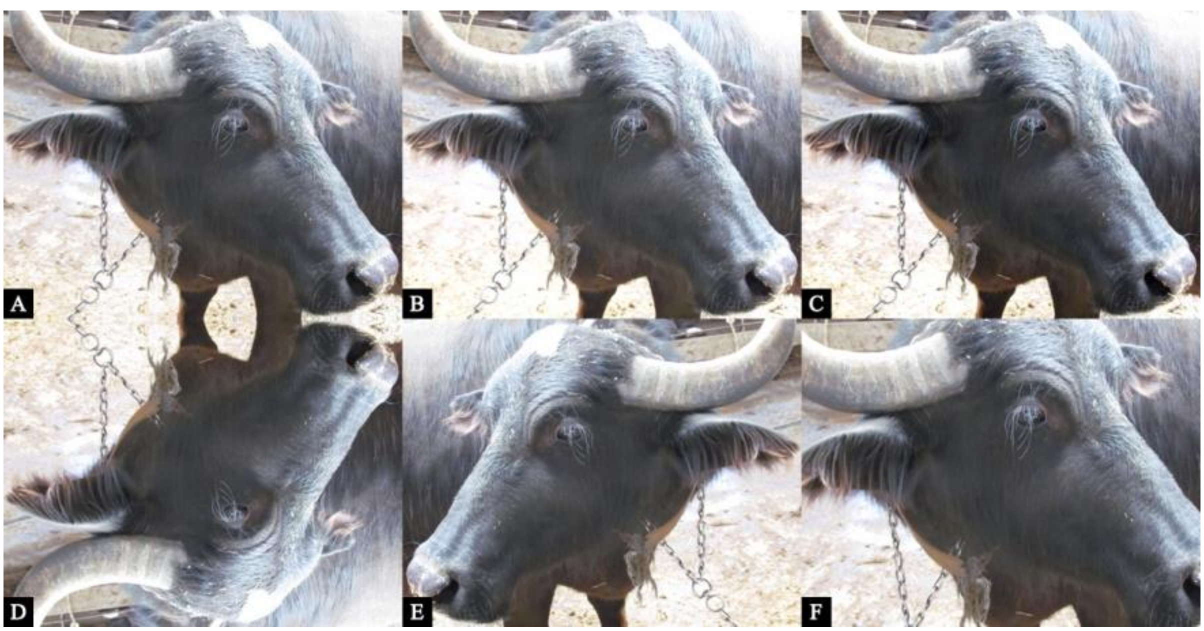

Fig. 2. Examples of augmented water buffalo face images: (A) original image; (B) brightness adjustment; (C) contrast adjustment; (D) vertical flip; (E) horizontal flip; (F) rotation.

After data splitting and training-set augmentation, 500 facial images were obtained for each of the eight water buffalo, resulting in a total of 4000 images. This curated dataset, hereafter referred to as the Buffalo-22 dataset, was used for model training, validation and testing. Summary statistics are presented in Table 2.

Table 2. Buffalo-22 statistics

Buffalo ID	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**Total**
Number of images	500	500	500	500	500	500	500	500	**4000**

Face recognition approaches: The analysis was conducted in two stages to enable a direct comparison between traditional and modern methods: face recognition using a statistical feature extraction approach and face recognition using a deep learning-based approach.

Face recognition using statistical feature extraction: Humeau-Heurtier (2019) reviewed texture feature extraction methods and grouped them into several categories. One of these categories, the statistical approach, has been used in a wide range of computer vision applications, including object detection and image classification, for more than half a century (Haralick *et al*., 1973).

Local Binary Pattern (LBP) feature extraction: Traditional machine learning approaches such as Local Binary Pattern (LBP)-based texture analysis combined with classical classifiers have been widely used in biometric recognition studies due to their low computational cost and effective texture representation capabilities. LBP was proposed by Ojala *et al*. (1996) and has been widely adopted as a texture descriptor due to its computational efficiency and its capability to capture fine-scale image details (Nanni *et al*., 2017). The canonical LBP operator is computed at each pixel by considering the values of a small circular neighborhood with radius R around a central pixel with intensity I_c,, as follows:

$$LBP\ (N, R) = \sum_{n=0}^{N-1} s(I_n - I_c)\, 2^n$$

The resulting descriptor is a histogram of the binary codes, where N denotes the number of pixels in the neighborhood, R is the radius, and the thresholding function is defined as $s(x) = 1$ if $x \geq 0$ and $s(x) = 0$ otherwise.

In this study, uniform LBP operator was used with a radius of R=1 pixel and N=8 neighboring pixels. This configuration creates a 59-bin histogram descriptor for each image channel.

Local Binary Pattern has been widely used as a feature extractor for face recognition tasks (Kumar *et al*., 2015; Niu *et al*., 2021; Liao *et al*., 2023). In this study, LBP was adopted as the traditional feature extraction method.

Color Conversions: Image texture, color and pixel density describe different parts of the visual information contained in an image. These properties can be measured over the full image or only within a selected region (Stockman and Shapiro, 2001). For this reason, earlier studies have used color-based information together with gray-level features (Madiwalar and Wyawahare, 2017; Suttapakti and Bunpeng, 2019; Xian and Ngadiran, 2021). Using texture information from color images has been found to improve classification in several applications. Since digital images can be represented in different color spaces, the choice of color space may also affect the extracted features. Previous studies have obtained good results with the HSV color space, which represents hue, saturation and value separately (Hayit *et al*., 2021; Kurniastuti *et al*., 2021; Hayit *et al*., 2023; Hayit *et al*., 2024).

For this study, the RGB images of water buffalo faces were converted into HSV format. This representation separates color-related information from brightness. A grayscale version of each image was also created so that additional texture information could be included. Features from the HSV channels and the grayscale image were then used together in the analysis. Fig. 3 shows examples of the original RGB image, its HSV form and the corresponding grayscale image.

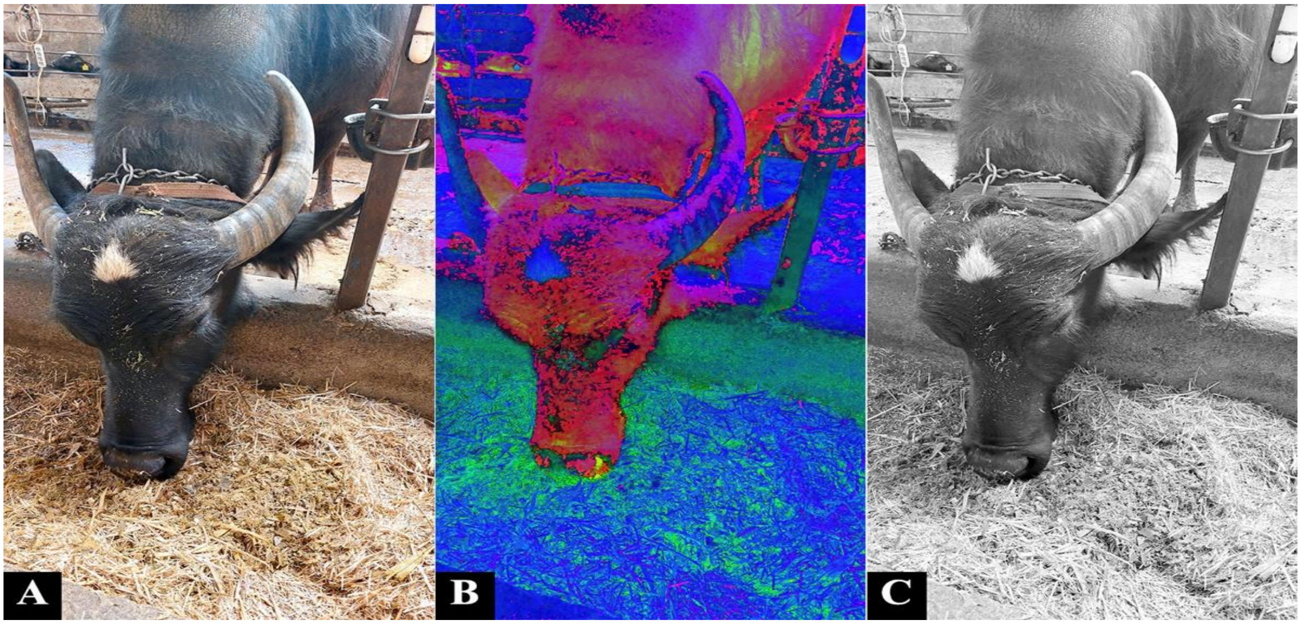

Fig. 3. Comparison of the original image (A), HSV representation (B), and grayscale representation (C).

Support Vector Machine (SVM) classifier: Support Vector Machine (SVM) is one of the classification methods used in supervised machine learning. It has frequently been used in studies on face recognition and texture classification (Vapnik, 2003; Shi *et al*., 2020; Varma *et al*., 2020; Abdullah and Abdulazeez, 2021; Chaabane *et al*., 2022). The method separates classes by finding a decision boundary, usually called a hyperplane, from the features given to the model (Memmedova, 2012). Because face-recognition data are not always linearly separable, a kernel function can be used to represent the data in a higher-dimensional space. In this study, a cubic polynomial kernel was selected because it gave the highest accuracy among the tested kernels.

Face recognition workflow based on statistical feature extraction:

- The traditional face recognition pipeline used in this study can be summarized as follows. Throughout this manuscript, the term "LBP-HSV" refers to the baseline model that combines LBP features extracted from the H (hue), S (saturation), V (value), and grayscale (GS) image channels. The resulting feature vector is subsequently classified using a Support Vector Machine (SVM) classifier.
- Input: Labeled images of water buffalo faces
- Output: Classified water buffalo faces based on the extracted features

Algorithm:

1. Image preprocessing: For each labeled image, crop the region of interest using the bounding box coordinates (x, y, width, height) provided by Roboflow, and resize the cropped image to 640×640×3 pixels, consistent with the YOLOv5 input size.
2. Color space conversion: For each preprocessed image, convert the RGB image into two additional representations: HSV (hue, saturation, value) and grayscale (GS).
3. Feature extraction: For each HSV and GS image, split the HSV image into its three channels, H (hue), S (saturation), and V (value). Apply Local Binary Pattern (LBP) feature extraction to each of the H, S, V, and GS channels, computing 59 features per channel.
4. Feature combination: Concatenate the LBP features extracted from the H, S, V, and GS channels, yielding a 236-dimensional feature vector (59 features per channel × 4 channels) for each image.
5. Classification: Use an SVM classifier to assign a class label to each water buffalo face based on the combined feature vector.

The face recognition workflow based on statistical feature extraction is presented in Fig. 4.

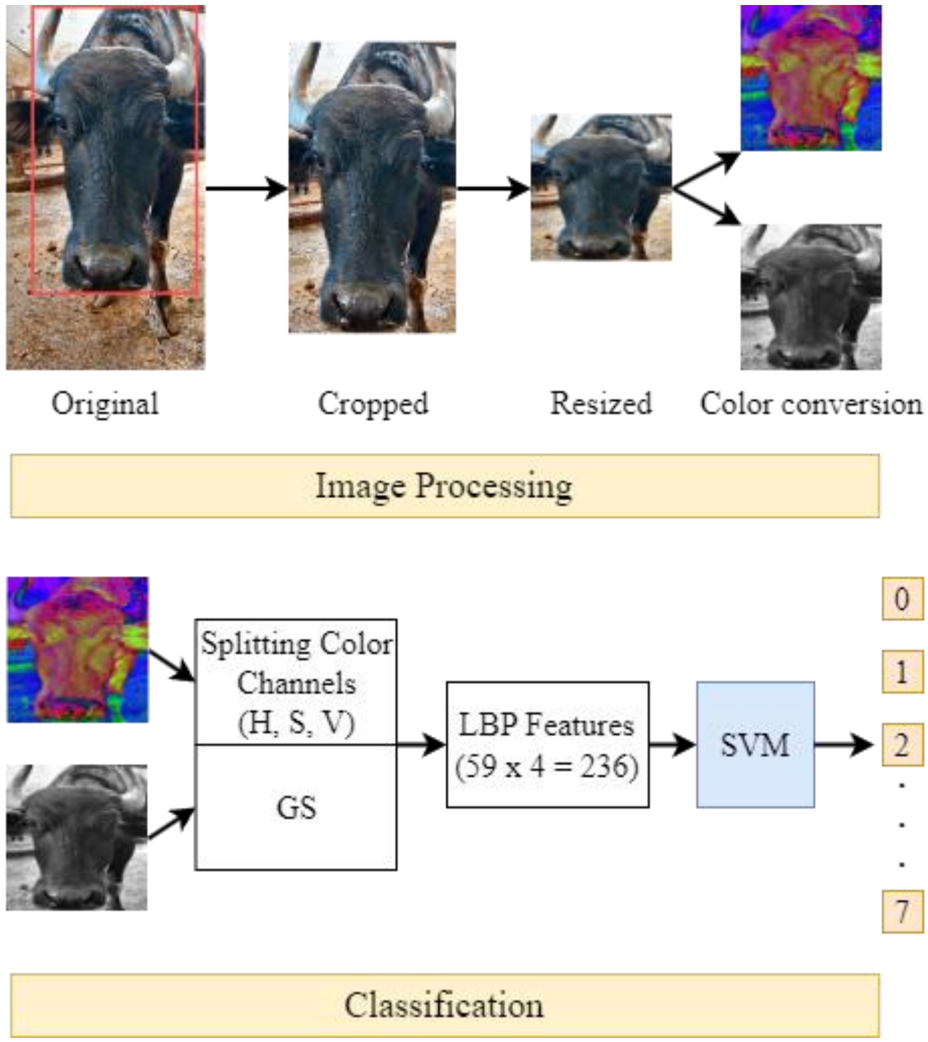

Fig. 4. Graphical representation of face recognition using statistical feature extraction.

Deep learning-based buffalo face detection and identity classification: The YOLO (You Only Look Once) family is primarily a one-stage object detection framework and is widely used in real-time detection tasks due to its high inference speed and practical applicability (Zhao *et al.*, 2020; Kumar *et al.*, 2021; Yadav *et al.*, 2022; Shahid and Yan, 2023; Xu *et al.*, 2023). In the present study, YOLOv5s was used as a multi-class object detection model rather than a classical embedding-based face recognition or re-identification architecture. Each buffalo individual was defined as a separate identity class, and the model was trained to localize the buffalo face and assign it to one of the predefined identity classes in a single step. Therefore, the term "recognition" refers to closed-set, detection-based identity classification of known buffalo individuals. YOLOv5s was preferred because it provides a practical, computationally efficient and real-time framework suitable for livestock farm conditions.

YOLOv5: YOLOv5 is a one-stage object detection model developed by Ultralytics. It is often used in applications that require fast detection because it can locate and classify objects in a single pass through the network (Majeed *et al.*, 2021; Yadav *et al.*, 2022; Zhou, 2022; Xu *et al.*, 2023). In this study, YOLOv5s version 6.0 was implemented in PyTorch and used to detect water buffalo faces and assign an identity class to each detected face. YOLOv5s was selected because its relatively small size makes it suitable for applications where inference speed is important. The model predicts the location of each face together with an objectness score and the probabilities of the identity classes.

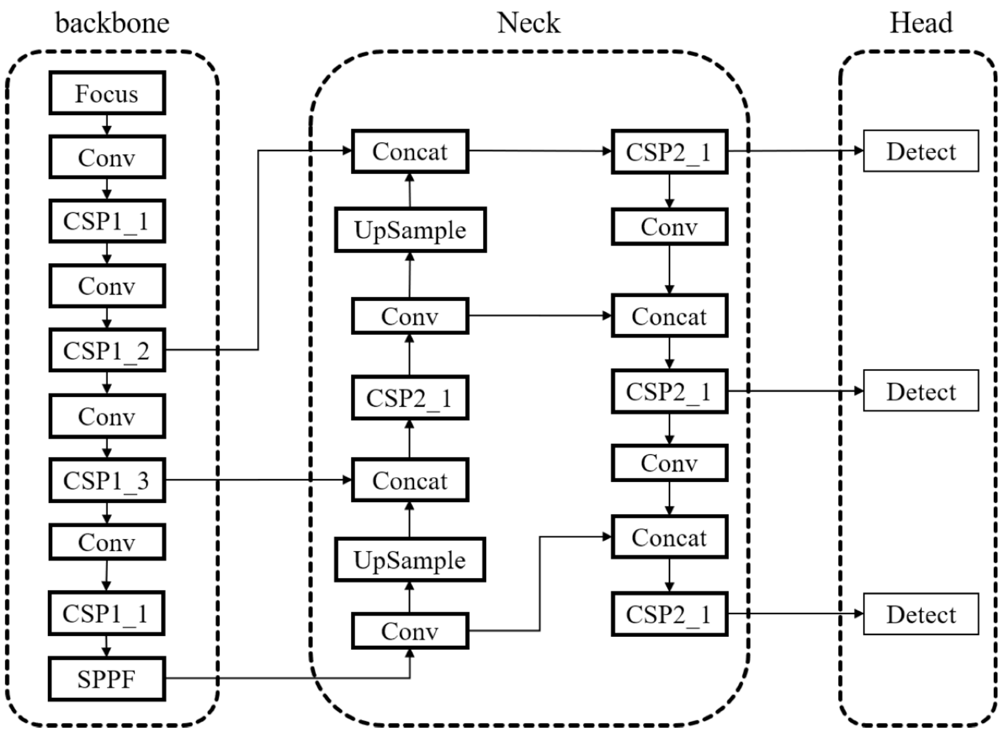

Fig. 5. YOLOv5 Network Architecture (Ma *et al*., 2023)

Activation Function: Activation functions enable a neural network to learn nonlinear relationships from the input data. In version 6.0 of YOLOv5s, the Sigmoid Linear Unit (SiLU) activation function is used primarily in the hidden convolutional layers. SiLU does not zero out small negative values directly, but retains them, which can help in maintaining the gradient flow during training. Corresponding outputs of the detection head such as objectness and class probabilities are sigmoid functions.

Optimization Function: The optimizer updates the model parameters during training based on the calculated gradients. The Adaptive Moment Estimation (Adam optimizer) was used in this study. Adam updates each parameter by estimates of the first and second moments of the gradients. It was selected because it generally provides stable updates and can converge fast in deep learning tasks (Majeed *et al*., 2021).

Loss Function: The model was trained by minimizing several loss terms together. These terms measure the errors in the predicted bounding-box coordinates (x, y, width, height), objectness scores and identity classes. These were summed and used to update the network parameters during backpropagation (Majeed *et al*., 2021; Solawetz, 2021).

Evaluation Metrics: To evaluate the accuracy of the face detection results, we used mean average precision (mAP) and Intersection over Union (IoU). mAP is a standard metric in computer vision for quantifying the performance of object detection algorithms, summarizing precision-recall behavior over all confidence thresholds. IoU measures the overlap between the predicted face region of an object (water buffalo) and the ground-truth annotated face area of the same object (Padilla *et al*., 2021; Bergman *et al*., 2024). mAP is computed as the mean of the class-wise average precision values:

$$\text{mAP} = \frac{1}{C}\sum_{c=1}^{C} AP_c \tag{1}$$

Where AP_c is the average precision of class c, and C is the number of classes.

True positives (TP) and true negatives (TN) are samples that are correctly classified by the model. False positives (FP) are samples that are incorrectly labeled as positive and false negatives (FN) are samples that are incorrectly labeled as negative (Grandini *et al*., 2020; Bergman *et al*., 2024). The F1-score can be interpreted as the harmonic mean of precision and recall, taking its maximum value of 1 and minimum value of 0 (Bergman *et al*., 2024). In this context, precision reflects the model's ability to correctly identify face bounding boxes among all predicted boxes, while recall reflects its ability to detect all ground-truth face bounding boxes (Padilla *et al*., 2021; Bergman *et al*., 2024). Precision, recall, and the F1-score are defined as follows.

$$Precision = \frac{TP}{TP + FP} \tag{2}$$

$$Recall = \frac{TP}{TP + FN} \tag{3}$$

Accuracy and the F1-score were used to evaluate the performance of the face recognition models. They are defined as follows.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{4}$$

$$F1\ Score = \frac{2\ x\ Precision\ x\ Recall}{Precision + Recall} \tag{5}$$

Experimental Setup and Hyperparameters: The Buffalo-22 dataset was used in all training and testing steps. The deep learning experiments were carried out in PyTorch. The software environment consisted of Python 3.9, CUDA 11.3 and cuDNN 8.x. MATLAB was also used for some of the additional analyses. All experiments were run on one NVIDIA RTX A4000 GPU.

Before any augmentation was performed, the original images were separated into training, validation and test groups. The proportions were 70%, 20% and 10%, respectively. The same split procedure was followed for both the LBP-HSV method and YOLOv5s. Augmentation was used only for the training images. The validation and test images were left in their original form so that the same images would not appear in more than one group.

For the LBP-HSV model, multiclass classification was carried out with a one-vs-one SVM and a cubic polynomial kernel. The SVM settings were selected by Bayesian optimization, which was run for a maximum of 30 iterations. YOLOv5s was trained with a batch size of 16 using the Adam optimizer. The starting learning rate was 0.05 and the maximum number of epochs was 150. Training could stop earlier when the validation loss showed no improvement for 10 successive epochs. A weight decay value of 0.0005 was also used as L2 regularization to limit model complexity and reduce variation in the learned parameters.

The hyperparameter values used for YOLOv5s training, including learning rate, batch size and weight decay, were selected based on commonly used YOLOv5 training settings reported in the literature and preliminary experiments aimed at achieving stable model convergence and satisfactory validation performance. No formal grid search or automated hyperparameter optimization procedure was performed.

RESULTS

This section discusses the comparative performance of the traditional and deep learning-based face recognition models. The Buffalo-22 dataset was divided into training, validation, and test sets by using a holdout validation scheme with random sampling and a ratio of 70%, 20%, and 10%, respectively. The number of samples in the training, validation, and test sets for each class is summarized in Table 3.

Table 3. Training, validation, and test groups

Groups	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**
Training	350	350	350	350	350	350	350	350
Validation	100	100	100	100	100	100	100	100
Test	50	50	50	50	50	50	50	50

On the traditional side, 236 LBP-HSV texture descriptor values were extracted from each image in the HSV color space and classified using an SVM classifier. This model achieved a validation accuracy of 88.4 percent and a test accuracy of 87.8 percent.

On the deep learning side, the YOLOv5s model was trained for up to 150 epochs with a batch size of 16. The metric values for the first 5 and last 5 epochs are reported in Table 4. The YOLOv5-based system achieved a mean average precision of 99.3 percent at an IoU threshold of 0.5.

Table 4. First 5 and last 5 epoch data

Epoch	Precision	Recall	mAP@0.5	mAP@0.5:0.95	Total Validation Loss
0	0.12727	0.71596	0.15635	0.099381	0.0045536
1	0.2251	0.68706	0.31847	0.22106	0.0038351
2	0.34601	0.65322	0.44965	0.28446	0.0043228
3	0.59211	0.66742	0.69733	0.47775	0.0042206
4	0.71052	0.73865	0.8094	0.53726	0.0042766
…	…	…	…	…	…
145	0.98917	0.99326	0.99391	0.85402	0.0026172
146	0.9892	0.99326	0.9939	0.85402	0.0026182
147	0.98922	0.99327	0.9939	0.85452	0.0026191
148	0.98927	0.99328	0.99391	0.85465	0.0026197
149	0.98927	0.99328	0.99391	0.85524	0.0026201

Note: Loss (val) represents the total validation loss (box + classification + objectness).

During training, the difference between the predicted boxes and the annotated boxes was followed through the loss values. For YOLOv5, this mainly included box loss and classification loss. Box loss measures errors in the location and size of the predicted boxes, while classification loss measures errors in the assigned buffalo identity. These values were used to monitor the progress of the model while training it. The training and validation loss curves are shown in Fig. 6.

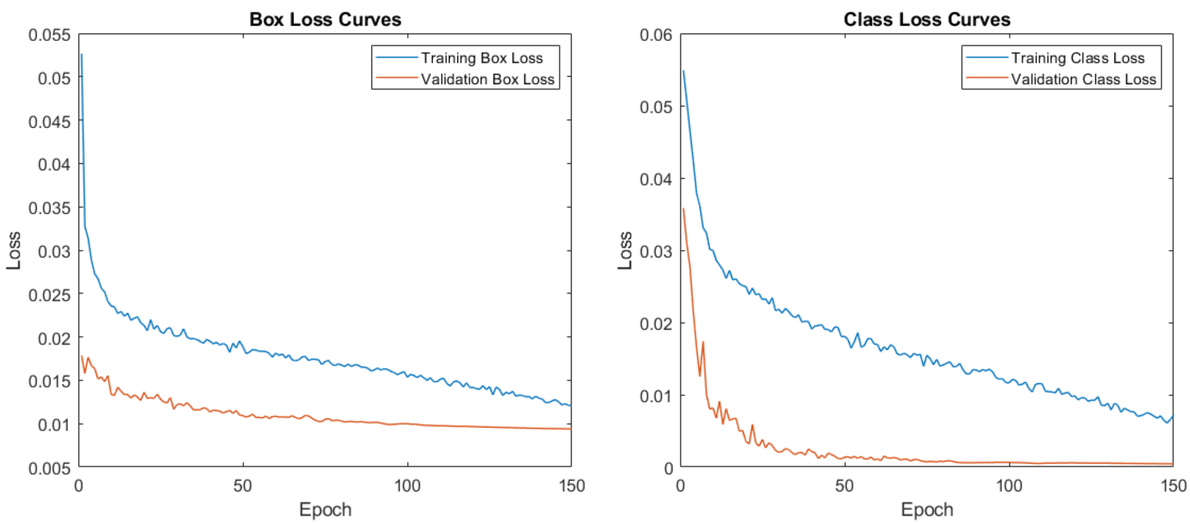

Fig. 6. Training and validation loss values

Performance metrics including precision, recall, and mAP, computed during training on both the training and validation sets, are shown in Fig. 7.

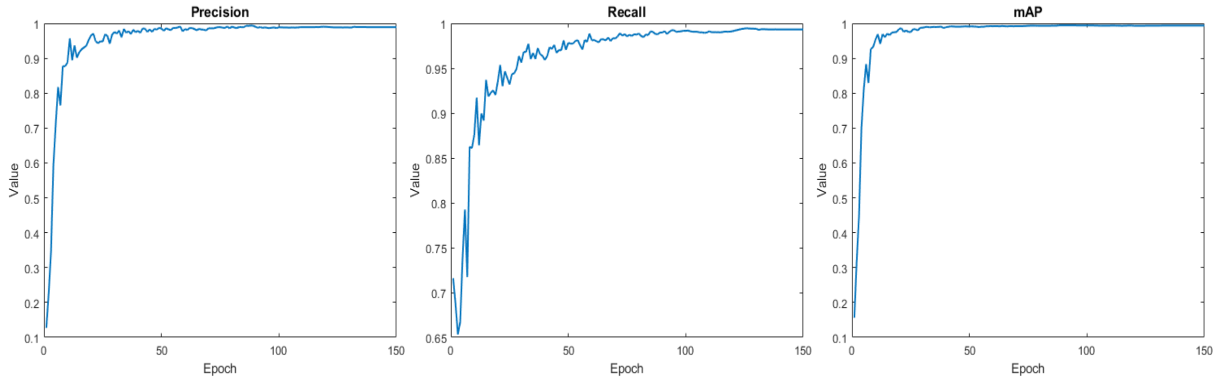

Fig. 7. Performance metrics in the training phase

Fig. 8 and Fig. 9 show example outputs of the system for individual frames, illustrating the predicted bounding boxes and the biometric face recognition results for each identified water buffalo.

Fig. 8. Detected faces of all buffalo using YOLOv5

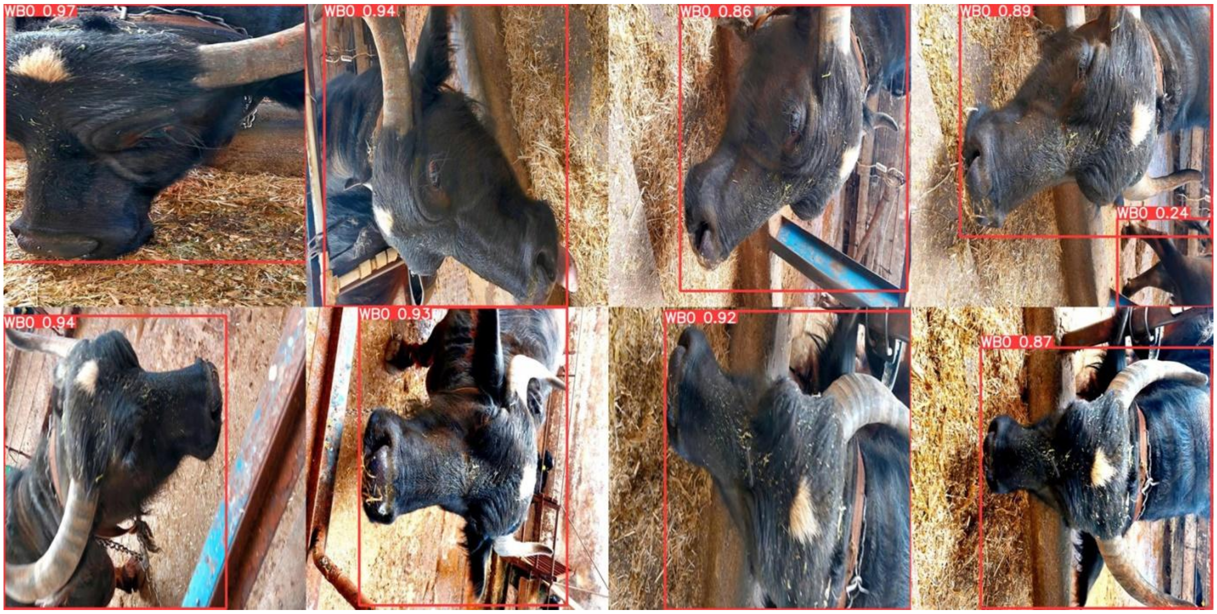

Fig. 9. Detected faces of WB0 (Water Buffalo-0) using YOLOv5

Finally, confusion matrices were used to compare the class-wise performance of the LBP-HSV and YOLOv5 models. The confusion matrices for both models are presented in Fig. 10.

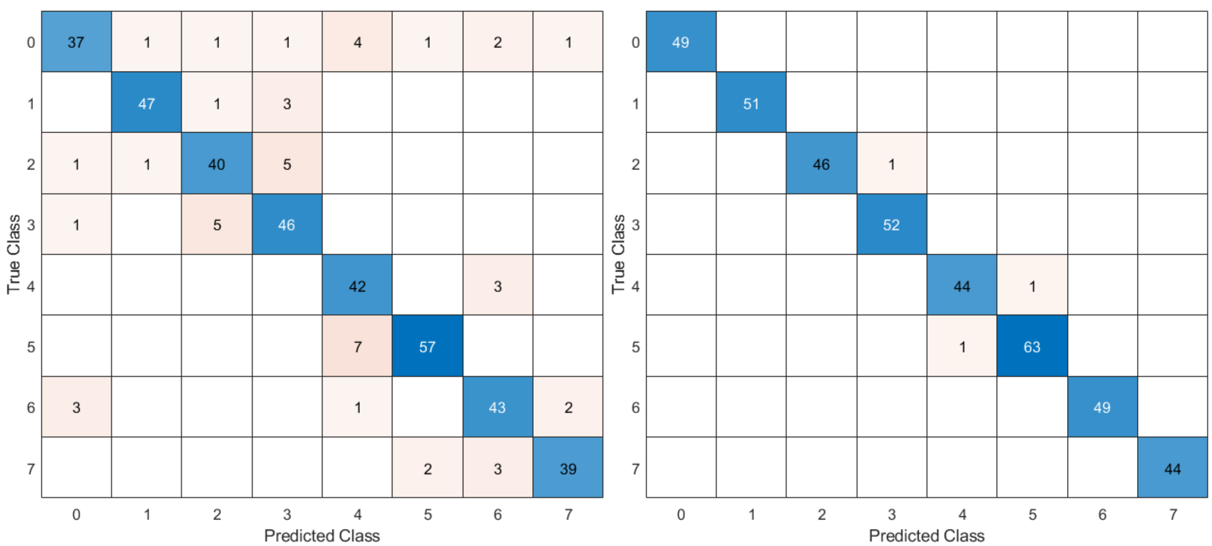

Fig. 10. Confusion matrices of LBP-HSV (left) and YOLOv5 (right) models

DISCUSSION

In this study, we tested a YOLOv5s-based system for identifying individual water buffalo from facial images in the Buffalo-22 dataset. The LBP-HSV method combined with SVM gave an accuracy of 88.4% on the validation set and 87.8% on the test set. These values are close to results reported in earlier biometric studies that used statistical texture features (Shi *et al*., 2020; Varma *et al*., 2020; Abdullah and Abdulazeez, 2021; Chaabane *et al*., 2022). The YOLOv5s model produced an mAP of 99.3% at an IoU threshold of 0.5 and also showed high precision and recall values for the eight buffalo identities (Fig. 7). Based on these results, the model appears suitable for fast, contact-free identification under farm conditions.

YOLOv5s was selected is a compact member of the YOLOv5 family that provides a practical balance between model size, inference speed, and predictive performance. YOLOv5n is generally lighter and faster but has lower model capacity. YOLOv5m, l, and x, on the other hand, have more parameters and require more computation time (Ultralytics, 2022). In our experiments, YOLOv5s gave high mAP, precision and recall values on Buffalo-22 while running on a single RTX A4000 GPU (Fig. 7). Although more recent versions of YOLO are now available, YOLOv5 is still widely used in real-time object detection studies. This was the main reason for using YOLOv5s in the present work.

From the point of view of precision livestock farming, the system could be useful for routine identification and herd management (Kumar *et al*., 2015; Dandıl *et al*., 2019; Yang *et al*., 2019; Yao *et al*., 2019; Bergman *et al*., 2024). Reading ear tags on water buffalo is not always easy and may be unsafe when the handler is not familiar with the animals. Tags may also be lost, damaged or difficult to read under farm conditions. Previous studies on dairy cattle have shown that facial images contain useful biometric information because skin texture and facial shape differ between animals (Kumar *et al*., 2016). The results of the present study show that these facial differences can also be used in water buffalo. The examples in Figs. 8 and 9 show that the animals could be detected and identified without direct contact. If the predicted identity is linked to farm records, information on pedigree, production and breeding could be accessed more quickly. This may reduce the time needed for record checking and support day-to-day management on buffalo farms.

Earlier animal face recognition studies were generally carried out on a small scale and included several species (Burghardt and Campbell, 2007). Subsequent research on cattle achieved high identification accuracy using face images and nose patterns (Kumar *et al*., 2015; Kumar *et al*., 2016; Kumar *et al*., 2018; Dandıl *et al*., 2019). Deep convolutional networks have also been used for classification between different animal species and for feature extraction from large image sets (Khan *et al*., 2020). Faster R-CNN and Mask R-CNN have been applied to animal face recognition, counting and census tasks in cattle and sheep with accuracies reported to be above 90% in controlled or partly controlled conditions (Dandıl *et al*., 2019; McLennan and Mahmoud, 2019; Dutta, 2021). Other studies have dealt with more difficult images, such as animals photographed behind cage bars, using M2Det and transfer learning (Li *et al*., 2020). Data augmentation has also been used to improve classification performance when the available image set was limited (Szegedy *et al*., 2016; Khan *et al*., 2020). The present study differs from much of this earlier work by using YOLOv5s for both face localization and identity classification in water buffalo images collected on a farm. The performance obtained on Buffalo-22 was similar to or higher than many earlier results reported for cattle and sheep, although the small number of animals should be taken into account.

The comparison with the traditional method also showed where hand-crafted features remain useful and where they may be limited. LBP-HSV combined with SVM gave a simple baseline with relatively low computational cost. This agrees with its common use in texture-based face recognition studies (Kumar *et al*., 2015; Niu *et al*., 2021; Liao *et al*., 2023). SVM can still perform well in both linear and nonlinear classification problems involving biometric or texture features (Kaur *et al*., 2022). In this study, however, its accuracy was about ten percentage points lower than the result obtained with YOLOv5s. The confusion matrices in Fig. 10 also show that the LBP-HSV model made more errors between some buffalo identities. This may be because fixed texture descriptors do not represent changes in pose, lighting and facial appearance as well as features learned directly from the images. Previous studies have suggested that combining LBP-type descriptors with CNN features or ensemble classifiers may improve performance further (Kumar *et al*., 2021; Shojaeipour *et al*., 2021; Kaur *et al*., 2022).

Another step is to test ensemble and hybrid approaches in more detail. The combination of the outputs of multiple classifiers, such as bagging and boosting, can be beneficial for unbalanced datasets or datasets with a high degree of variation (Kumar *et al*., 2021). Other studies have also combined hand-crafted texture features such as LBP with CNNs or ensemble classifiers and achieved better classification performance (Shojaeipour *et al*., 2021; Hayıt and Çınarer, 2022; Kaur *et al*., 2022). A similar approach could be applied to water buffalo identification. In this way, the lower computational cost and easier interpretation of hand-crafted features could be used together with the stronger feature-learning ability of deep networks.

Conclusion: Buffalo-22 includes images from only eight animals, which is an important limitation of the study. However, the photographs were taken under different poses, viewing angles and lighting conditions, and the training set was expanded through data augmentation. The results obtained on the validation and test sets indicate that YOLOv5s was able to distinguish the animals in this dataset with high accuracy. Even so, the model should be tested on more animals, on different farms and under a wider range of management conditions before broader conclusions are drawn. The system allows buffalo to be identified without direct contact and without interrupting normal farm activities. It may also reduce the need to approach animals closely to read ear tags. When connected to farm records, the predicted identity could be used to retrieve pedigree, production and breeding information. The results therefore support further testing of YOLOv5s-based identification as a possible tool for water buffalo farms.

Acknowledgements: This study was conducted as part of Niyazi Hayrullah TUVAY's MSc thesis and was supported by TÜBİTAK with Project Number 1002/122E313. The study received ethical approval from the Kırşehir Ahi Evran University Animal Experiments Local Ethics Committee (Decision Date: 24/08/2022, Decision Number: 16).

Conflict of Interest: The authors declare that they have no conflicts of interest.

Author Contribution: Conceptualization: N.H.T. and O.E.; Methodology: T.H.; Software: T.H.; Validation: N.H.T.; Formal analysis: T.H.; Investigation: N.H.T. and O.E.; Resources: O.E.; Data curation: T.H.; Writing-original draft: N.H.T. and O.E.; Writing-review and editing: O.E.; Visualization: T.H.; Supervision: T.H.; Project administration: T.H. and O.E.; Funding acquisition: O.E. All co-authors reviewed the final version and approved the manuscript before submission.

REFERENCES

Abdullah, D.M. and A.M. Abdulazeez (2021). Machine learning applications based on SVM classification: a review. Qubahan Acad. J. 1(2): 81-90. https://doi.org/10.48161/qaj.v1n2a50

Awad, A.I., A.E. Hassanien and H.M. Zawbaa (2013). A cattle identification approach using live captured muzzle print images. In: Advances in Security of Information and Communication Networks. Springer; Berlin, Heidelberg (Germany). pp. 143-152. https://doi.org/10.1007/978-3-642-40597-6_12

Awad, A.I. (2016). From classical methods to animal biometrics: a review on cattle identification and tracking. Comput. Electron. Agric. 123: 423-435. https://doi.org/10.1016/j.compag.2016.03.014

Barry, B., G. Corkery, U. Gonzales-Barron, K. McDonnell, F. Butler and S. Ward (2008). A longitudinal study of the effect of time on the matching performance of a retinal recognition system for lambs. Comput. Electron. Agric. 64(2): 202-211. https://doi.org/10.1016/j.compag.2008.05.011

Bergman, N., Y. Yitzhaky and I. Halachmi (2024). Biometric identification of dairy cows via real-time facial recognition. Animal 18(3): 101079. https://doi.org/10.1016/j.animal.2024.101079

Bezen, R., Y. Edan and I. Halachmi (2020). Computer vision system for measuring individual cow feed intake using RGB-D camera and deep learning algorithms. Comput. Electron. Agric. 172: 105345. https://doi.org/10.1016/j.compag.2020.105345

Bugge, C.E., J. Burkhardt, K.S. Dugstad, T.B. Enger, M. Kasprzycka, A. Kleinauskas, M. Myhre, K. Scheffler, S. Ström and S. Vetlesen (2011). Biometric methods of animal identification. Course notes, Laboratory Animal Science, Norwegian School of Veterinary Science; Oslo (Norway). pp. 1-6.

Burghardt, T. and N. Campbell (2007). Individual animal identification using visual biometrics on deformable coat patterns. In: Proc. 5th Int. Conf. Computer Vision Systems (ICVS), 21-24 March 2007, Bielefeld (Germany). pp. 1-10. https://doi.org/10.2390/biecoll-icvs2007-121

Burghardt, T. (2008). Visual animal biometrics: automatic detection and individual identification by coat pattern. Ph.D. thesis, University of Bristol; Bristol (UK).

Çelikyürek, H. and K. Karakuş (2017). Biometric identification and recording in animal production. Yüzüncü Yıl Üniv. Fen Bilim. Enst. Derg. 22(2): 211-218. (in Turkish).

Chaabane, S.B., M. Hijji, R. Harrabi and H. Seddik (2022). Face recognition based on statistical features and SVM classifier. Multimed. Tools Appl. 81(6): 8767-8784. https://doi.org/10.1007/s11042-021-11816-w

Dandıl, E., M. Turkan, M. Boğa and K.K. Çevik (2019). Recognition of cattle faces using the faster R-CNN. Bilecik Şeyh Edebali Üniv. Fen Bilim. Derg. 6(2): 177-189. https://doi.org/10.35193/bseufbd.592099 (in Turkish).

Dutta, P. (2021). A deep learning approach for animal breed classification: sheep. Int. J. Res. Appl. Sci. Eng. Technol. 9(5): 73-76. https://doi.org/10.22214/ijraset.2021.34050

Dwyer, B. and J. Gallagher (2023). Getting started with Roboflow. Available at: https://blog.roboflow.com/getting-started-with-roboflow/ (accessed on 25 Sept. 2024).

Erbay, H. and T. Hayit (2025). A vision transformer approach for fusarium wilt of chickpea classification. Multimed. Tools Appl. 84(24): 28287-28304. https://doi.org/10.1007/s11042-024-20224-9

Erdem, O.A. and H. Tuna (2008). The constitute of international electronic database for the cattles identifying and application in Türkiye. Eng. Sci. 3(2): 268-287. https://doi.org/10.12739/nwsaes.v3i2.5000067191 (in Turkish).

Ermetin, O. (2017). Husbandry and sustainability of water buffaloes in Turkey. Turkish J. Agric. Food Sci. Technol. 5(12): 1673-1682. https://doi.org/10.24925/turjaf.v5i12.1673-1682.1639

Ermetin, O. (2021). Precision livestock farming: potential use in water buffalo (*Bubalus bubalis*) operations. Anim. Sci. Pap. Rep. 39(1): 19-30.

Ermetin, O. and H.K. Örnek (2025). Deep-learning-based buffalo identification through muzzle pattern images. Arch. Anim. Breed. 68(3): 473-484. https://doi.org/10.5194/aab-68-473-2025

Grandini, M., E. Bagli and G. Visani (2020). Metrics for multi-class classification: an overview. arXiv preprint arXiv:2008.05756. https://doi.org/10.48550/arXiv.2008.05756

Haralick, R.M., K. Shanmugam and I.H. Dinstein (1973). Textural features for image classification. IEEE Trans. Syst. Man Cybern. SMC-3(6): 610-621. https://doi.org/10.1109/TSMC.1973.4309314

Hayit, T., H. Erbay, F. Varçın, F. Hayit and N. Akci (2021). Determination of the severity level of yellow rust disease in wheat by using convolutional neural networks. J. Plant Pathol. 103(3): 923-934. https://doi.org/10.1007/s42161-021-00886-2

Hayıt, T. and G. Çınarer (2022). Classification of Covid-19 based on a combination of GLCM and deep features by using X-ray images. İnönü Üniv. Sağlık Hizm. Meslek Yüksekokulu Derg. 10(1): 313-325. https://doi.org/10.33715/inonusaglik.1015407 (in Turkish).

Hayit, T., H. Erbay, F. Varçın, F. Hayit and N. Akci (2023). The classification of wheat yellow rust disease based on a combination of textural and deep features. Multimed. Tools Appl. 82(30): 47405-47423. https://doi.org/10.1007/s11042-023-15199-y

Hayit, T., A. Endes and F. Hayit (2024). KNN-based approach for the classification of fusarium wilt disease in chickpea based on color and texture features. Eur. J. Plant Pathol. 168(4): 665-681. https://doi.org/10.1007/s10658-023-02791-z

Humeau-Heurtier, A. (2019). Texture feature extraction methods: a survey. IEEE Access 7: 8975-9000. https://doi.org/10.1109/ACCESS.2018.2890743

Kaur, A., M. Kumar and M.K. Jindal (2022). Shi-Tomasi corner detector for cattle identification from muzzle print image pattern. Ecol. Inform. 68: 101549. https://doi.org/10.1016/j.ecoinf.2021.101549

Khan, R.H., K.W. Kang, S.J. Lim, S.D. Youn, O.J. Kwon, S.H. Lee and K.R. Kwon (2020). Animal face classification using dual deep convolutional neural network. J. Korea Multimed. Soc. 23(4): 525-538. https://doi.org/10.9717/kmms.2020.23.4.525

Kühl, H.S. and T. Burghardt (2013). Animal biometrics: quantifying and detecting phenotypic appearance. Trends Ecol. Evol. 28(7): 432-441. https://doi.org/10.1016/j.tree.2013.02.013

Kul, E., G. Filik, A. Şahin, H. Çayıroğlu, E. Uğurlutepe and H. Erdem (2018). Effects of some environmental factors on birth weight of Anatolian buffalo calves. Turkish J. Agric. Food Sci. Technol. 6(4): 444-446. https://doi.org/10.24925/turjaf.v6i4.444-446.1716

Kul, E. (2020). Determination of the best lactation curve model and lactation curve parameters using different nonlinear models for Anatolian buffaloes. Pakistan J. Zool. 52(4): 1291-1297. https://doi.org/10.17582/journal.pjz/20190517090530

Kumar, S. and S.K. Singh (2014). Biometric recognition for pet animal. J. Softw. Eng. Appl. 7(5): 470-482. https://doi.org/10.4236/jsea.2014.75044

Kumar, S., S. Tiwari and S.K. Singh (2015). Face recognition for cattle. In: Proc. 3rd Int. Conf. Image Information Processing (ICIIP), 21-24 December 2015, Waknaghat (India). IEEE. pp. 65-72. https://doi.org/10.1109/ICIIP.2015.7414742

Kumar, S., S. Tiwari and S.K. Singh (2016). Face recognition of cattle: can it be done? Proc. Natl. Acad. Sci. India Sect. A Phys. Sci. 86(2): 137-148. https://doi.org/10.1007/s40010-016-0264-2

Kumar, S., A. Pandey, K.S.R. Satwik, S. Kumar, S.K. Singh, A.K. Singh and A. Mohan (2018). Deep learning framework for recognition of cattle using muzzle point image pattern. Measurement 116: 1-17. https://doi.org/10.1016/j.measurement.2017.10.064

Kumar, S., S.K. Singh and A.K. Singh (2021). Secure and sustainable framework for cattle recognition using wireless multimedia networks and machine learning techniques. IEEE Trans. Sustain. Comput. 7(3): 696-708. https://doi.org/10.1109/TSUSC.2021.3123496

Kurniastuti, I., T.D. Wulan and A. Andini (2021). Color feature extraction of fingernail image based on HSV color space as early detection risk of diabetes mellitus. In: Proc. Int. Conf. Computer Science, Information Technology, and Electrical Engineering (ICOMITEE), 27-28 October 2021, Banyuwangi (Indonesia). IEEE. pp. 51-55. https://doi.org/10.1109/ICOMITEE53461.2021.9650161

Li, N., W. Kusakunniran and S. Hotta (2020). Detection of animal behind cages using convolutional neural network. In: Proc. 17th Int. Conf. Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON), 24-27 June 2020, Phuket (Thailand). IEEE. pp. 242-245. https://doi.org/10.1109/ECTI-CON49241.2020.9158137

Liao, J., Y. Lin, T. Ma, S. He, X. Liu and G. He (2023). Facial expression recognition methods in the wild based on fusion feature of attention mechanism and LBP. Sensors 23(9): 4204. https://doi.org/10.3390/s23094204

Lu, Y., X. He, Y. Wen and P.S. Wang (2014). A new cow identification system based on iris analysis and recognition. Int. J. Biom. 6(1): 18-32. https://doi.org/10.1504/IJBM.2014.059639

Ma, Z., Y. Wan, J. Liu, R. An and L. Wu (2023). A kind of water surface multi-scale object detection method based on improved YOLOv5 network. Mathematics 11(13): 2936. https://doi.org/10.3390/math11132936

Madiwalar, S.C. and M.V. Wyawahare (2017). Plant disease identification: a comparative study. In: Proc. Int. Conf. Data Management, Analytics and Innovation (ICDMAI), 24-26 February 2017, Pune (India). IEEE. https://doi.org/10.1109/ICDMAI.2017.8073478

Majeed, F., F.Z. Khan, M.J. Iqbal and M. Nazir (2021). Real-time surveillance system based on facial recognition using YOLOv5. In: Proc. Mohammad Ali Jinnah University Int. Conf. Computing (MAJICC), 15-17 July 2021, Karachi (Pakistan). IEEE. pp. 1-6. https://doi.org/10.1109/MAJICC53071.2021.9526254

Marchant, J. (2002). Secure animal identification and source verification. JM Communications; Milton Keynes (UK). pp. 1-28.

McLennan, K. and M. Mahmoud (2019). Development of an automated pain facial expression detection system for sheep (*Ovis aries*). Animals 9(4): 196. https://doi.org/10.3390/ani9040196

Memmedova, N. (2012). Early detection of mastitis in dairy cattle using artificial intelligence techniques. Ph.D. dissertation, Graduate School of Natural and Applied Sciences, Selçuk University; Konya (Türkiye). (in Turkish).

Nanni, L., S. Ghidoni and S. Brahnam (2017). Handcrafted vs. non-handcrafted features for computer vision classification. Pattern Recognit. 71: 158-172. https://doi.org/10.1016/j.patcog.2017.05.025

Nilsson, K., T. Rognvaldsson, J. Cameron and C. Jacobson (2006). Biometric identification of mice. In: Proc. 18th Int. Conf. Pattern Recognition (ICPR), 20-24 August 2006, Hong Kong (China). Vol. 4. IEEE. pp. 465-468. https://doi.org/10.1109/ICPR.2006.329

Niu, B., Z. Gao and B. Guo (2021). Facial expression recognition with LBP and ORB features. Comput. Intell. Neurosci. 2021: 8828245. https://doi.org/10.1155/2021/8828245

Noviyanto, A. and A.M. Arymurthy (2013). Beef cattle identification based on muzzle pattern using a matching refinement technique in the SIFT method. Comput. Electron. Agric. 99: 77-84. https://doi.org/10.1016/j.compag.2013.09.002

Ojala, T., M. Pietikäinen and D. Harwood (1996). A comparative study of texture measures with classification based on featured distributions. Pattern Recognit. 29(1): 51-59. https://doi.org/10.1016/0031-3203(95)00067-4

Padilla, R., W.L. Passos, T.L.B. Dias, S.L. Netto and E.A.B. da Silva (2021). A comparative analysis of object detection metrics with a companion open-source toolkit. Electronics 10(3): 279. https://doi.org/10.3390/electronics10030279

Qiao, Y., D. Su, H. Kong, S. Sukkarieh, S. Lomax and C. Clark (2019). Individual cattle identification using a deep learning based framework. IFAC-PapersOnLine 52(30): 318-323. https://doi.org/10.1016/j.ifacol.2019.12.558

Shahid, A.R. and H. Yan (2023). SqueezExpNet: dual-stage convolutional neural network for accurate facial expression recognition with attention mechanism. Knowl.-Based Syst. 269: 110451. https://doi.org/10.1016/j.knosys.2023.110451

Sharma, A., L. Randewich, W. Andrew, S. Hannuna, N. Campbell, S. Mullan, A.W. Dowsey, M. Smith, M. Hansen and T. Burghardt (2025). Universal bovine identification via depth data and deep metric learning. Comput. Electron. Agric. 229: 109657. https://doi.org/10.1016/j.compag.2024.109657

Shi, L., X. Wang and Y. Shen (2020). Research on 3D face recognition method based on LBP and SVM. Optik 220: 165157. https://doi.org/10.1016/j.ijleo.2020.165157

Shojaeipour, A., G. Falzon, P. Kwan, N. Hadavi, F.C. Cowley and D. Paul (2021). Automated muzzle detection and biometric identification via few-shot deep transfer learning of mixed breed cattle. Agronomy 11(11): 2365. https://doi.org/10.3390/agronomy11112365

Solawetz, J. (2021). YOLOv5 v6.0 is here: new nano model at 1666 FPS. Available at: https://blog.roboflow.com/yolov5-v6-0-is-here/ (accessed on 05 Feb. 2024).

Stockman, G. and L.G. Shapiro (2001). Computer vision. Prentice Hall; Upper Saddle River, NJ (USA). pp. 72-89.

Suttapakti, U. and A. Bunpeng (2019). Potato leaf disease classification based on distinct color and texture feature extraction. In: Proc. 19th Int. Symp. Communications and Information Technologies (ISCIT), 25-27 September 2019, Ho Chi Minh City (Vietnam). IEEE. pp. 82-85. https://doi.org/10.1109/ISCIT.2019.8905128

Szegedy, C., V. Vanhoucke, S. Ioffe, J. Shlens and Z. Wojna (2016). Rethinking the inception architecture for computer vision. In: Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR), 27-30 June 2016, Las Vegas, NV (USA). pp. 2818-2826. https://doi.org/10.1109/CVPR.2016.308

Ultralytics (2022). Pretrained checkpoints. Available at: https://github.com/ultralytics/yolov5 (accessed on 12 Feb. 2024).

Vapnik, V. (2003). The nature of statistical learning theory. Springer Science & Business Media; New York (USA). https://doi.org/10.1007/978-1-4757-3264-1

Varma, S., M. Shinde and S.S. Chavan (2020). Analysis of PCA and LDA features for facial expression recognition using SVM and HMM classifiers. In: Techno-Societal 2018: Proc. 2nd Int. Conf. Advanced Technologies for Societal Applications, Vol. 1. Springer; Cham (Switzerland). pp. 109-119. https://doi.org/10.1007/978-3-030-16848-3_11

Wang, H., J. Qin, Q. Hou and S. Gong (2020). Cattle face recognition method based on parameter transfer and deep learning. J. Phys. Conf. Ser. 1453(1): 012054. https://doi.org/10.1088/1742-6596/1453/1/012054

Xian, T.S. and R. Ngadiran (2021). Plant diseases classification using machine learning. J. Phys. Conf. Ser. 1962(1): 012024. https://doi.org/10.1088/1742-6596/1962/1/012024

Xu, B., W. Wang, L. Guo, G. Chen, Y. Li, Z. Cao and S. Wu (2022). CattleFaceNet: a cattle face identification approach based on RetinaFace and ArcFace loss. Comput. Electron. Agric. 193: 106675. https://doi.org/10.1016/j.compag.2021.106675

Xu, W., B. Li, Y. Du and S. Dong (2023). Study on facial recognition method based on YOLOv5. J. Phys. Conf. Ser. 2560(1): 012020. https://doi.org/10.1088/1742-6596/2560/1/012020

Yadav, P.K., J.A. Thomasson, S.W. Searcy, R.G. Hardin, U. Braga-Neto, S.C. Popescu, D.E. Martin, R. Rodriguez, K. Meza, J. Enciso, J. Solórzano Diaz and T. Wang (2022). Assessing the performance of YOLOv5 algorithm for detecting volunteer cotton plants in corn fields at three different growth stages. Artif. Intell. Agric. 6: 292-303. https://doi.org/10.1016/j.aiia.2022.11.005

Yang, Z., H. Xiong, X. Chen, H. Liu, Y. Kuang and Y. Gao (2019). Dairy cow tiny face recognition based on convolutional neural networks. In: Biometric Recognition. Lecture Notes in Computer Science, Vol. 11818. Springer; Cham (Switzerland). pp. 216-222. https://doi.org/10.1007/978-3-030-31456-9_24

Yao, L., Z. Hu, C. Liu, H. Liu, Y. Kuang and Y. Gao (2019). Cow face detection and recognition based on automatic feature extraction algorithm. In: Proc. ACM Turing Celebration Conf. China, 17-19 May 2019, Chengdu (China). pp. 1-5. https://doi.org/10.1145/3321408.3322628

Yu, P., T. Burghardt, A.W. Dowsey and N.W. Campbell (2025). Holstein-Friesian re-identification using multiple cameras and self-supervision on a working farm. Comput. Electron. Agric. 237: 110568. https://doi.org/10.1016/j.compag.2025.110568

Zhang, Q., L.T. Yang, Z. Chen and P. Li (2018). A survey on deep learning for big data. Inf. Fusion 42: 146-157. https://doi.org/10.1016/j.inffus.2017.10.006

Zhao, H., J. Zheng, Y. Wang, X. Yuan and Y. Li (2020). Portrait style transfer using deep convolutional neural networks and facial segmentation. Comput. Electr. Eng. 85: 106655. https://doi.org/10.1016/j.compeleceng.2020.106655

Zhou, Z. (2022). Detection and counting method of pigs based on YOLOV5_Plus: a combination of YOLOv5 and attention mechanism. Math. Probl. Eng. 2022: 7078670. https://doi.org/10.1155/2022/7078670