Journal of Agricultural Sciences (Tarim Bilimleri Dergisi) 2025, 31 (1) : 137 – 150 DOI: 10.15832/ankutbd.1509798

Journal of Agricultural Sciences
(Tarim Bilimleri Dergisi)

J Agr Sci-Tarim Bili
e-ISSN: 2148-9297
jas.ankara.edu.tr

Deep Learning based Individual Cattle Face Recognition using Data Augmentation and Transfer Learning

Havva Eylem Polat[a], Dilara Gerdan Koc[b], Ömer Ertugrul[c], Caner Koc[b]*, Kamil Ekinci[d]

[a]*Ankara University, Faculty of Agriculture, Agricultural Structures and Irrigation Department, 06110, Diskapi, Ankara, TURKEY*
[b]*Ankara University, Faculty of Agriculture, Department of Agricultural Machinery and Technologies Engineering, Ankara, TURKEY*
[c]*Kırşehir Ahi Evran University, Faculty of Agriculture, Department of Biosystems Engineering, Kırşehir, TURKEY*
[d]*Isparta University of Applied Sciences, Faculty of Agriculture, Department of Agricultural Machinery and Technologies Engineering, Isparta, TURKEY*

ARTICLE INFO
Research Article
Corresponding Author: Caner Koc, E-mail: ckoc@ankara.edu.tr
Received: 03 July 2024 / Revised: 13 August 2024 / Accepted: 23 August 2024 / Online: 14 January 2025

Cite this article
Polat H E, Koc D G, Ertugrul Ö, Koc C, Ekinci K (2025). Deep Learning based Individual Cattle Face Recognition using Data Augmentation and Transfer Learning. *Journal of Agricultural Sciences (Tarim Bilimleri Dergisi),* 31(1):137-150. DOI: 10.15832/ankutbd.1509798

ABSTRACT
Accurate identification of cattle is essential for monitoring ownership, controlling production supply, preventing disease, and ensuring animal welfare. Despite the widespread use of ear tag-based techniques in livestock farm management, large-scale farms encounter challenges in identifying individual cattle. The process of identifying individual animals can be hindered by ear tags that fall off, and the ability to identify them over a long period of time becomes impossible when tags are missing. A dataset was generated by capturing images of cattle in their native environment to tackle this issue. The dataset was divided into three segments: training, validation, and testing. The dataset consisted of 15 000 records, each pertaining to a distinct bovine specimen from a total of 30 different cattle. To identify specific cattle faces in this study, deep learning algorithms such as InceptionResNetV2, MobileNetV2, DenseNet201, Xception, and NasNetLarge were utilized. The DenseNet201 algorithm attained a peak test accuracy of 99.53% and a validation accuracy of 99.83%. Additionally, this study introduces a novel approach that integrates advanced image processing techniques with deep learning, providing a robust framework that can potentially be applied to other domains of animal identification, thus enhancing overall farm management and biosecurity.

Keywords: Cattle identification, Deep learning, Face detection, Smart farming

1. Introduction

Developments in technology have led to significant progress in the application of fully automated monitoring and control systems in the field of animal husbandry. In recent times, breeders have demonstrated a preference for intelligent livestock systems that constantly monitor the manner in which animals engage in reproduction, nutrition, health, comfort, and their surroundings (Džermeikaitė et al. 2023). These systems employ a variety of modelling techniques to accomplish this. Further, these systems are able to anticipate significant events such as birth and disease and then take the appropriate precautions in response to those events. Robotic, automated, and artificial intelligence-based tools are utilized in the process of breeding dairy cattle. The ability to monitor the specific requirements of each animal, make appropriate adjustments to the diet of the animals, prevent illnesses, and ultimately improve the overall health of the herd is a greater capacity that breeders possess. For the purpose of improving milk yields and animal welfare, as well as reducing methane emissions from animal waste by thirty percent, digital technological systems are being utilized in livestock enterprises, according to research conducted by scientists (Polat 2022). At the same time that automated systems reduce the amount of work that is performed by humans, they also reduce the amount of time that is spent in a shelter. Consequently, this makes it possible to manage larger herds, which ultimately leads to the development of livestock businesses that are both healthier and more profitable. In order to effectively transmit information regarding product yield and quality, intelligent livestock management makes use of electronic radio frequency identification systems, in addition to herd management software and internet connections. For the purposes of ensuring production, regulating disease, administering vaccinations, monitoring animal well-being, and managing ownership, accurate identification of cattle is essential (Allen et al. 2008). Throughout the course of history, ear tags and tattoos that were used for the purpose of identifying cattle have been prone to experiencing fading, loss, and damage. In comparison to more traditional approaches, RFID systems offer a significant number of advantages and improvements in operational efficiency. However, they also present significant risks to both security and privacy, which makes them susceptible to a variety of vulnerabilities (Awad 2016). According to Ruiz-Garcia & Lunadei (2011) and Kumar et al. (2016), the application of techniques for identifying cattle based on ear tags is a common practice in the management of livestock farms. Managing the transmission of acute diseases and understanding the progression of diseases are both possible outcomes that can be achieved with the assistance of these techniques (Wang et al. 2010). Methods that are based on tags make use of one-of-a-kind identifiers, which may take the form of permanent markings, temporary markings, or electronic devices. According to Awad (2016), ear notching is a technique that involves the removal of a section of an animal's ear or ears, which results in a differentiated shape. Combining the positions of the ear notch on different cattle allows for the identification of specific cattle. Ear notching, on the other hand, can have a negative impact on the well-being of animals, whereas alternative methods of identification are more beneficial to the welfare of animals. According to Noonan et al. (1994), ear notching is a method that is not only limited in its ability to identify specific cattle on a farm, but it is also not feasible for accurately identifying individual cattle while working on large-scale farms. An additional disadvantage of ear tags is that they have the potential to become detached, which makes it impossible to differentiate between different animals (Wang et al. 2010; Awad 2016). There is a specific implementation of object detection known as "face detection" which accurately identifies and localizes target faces in images. At the moment, there is a significant amount of research activity in the field of computer vision that is centered on the detection of objects. According to Xu et al. (2021), this field of research makes it possible to perform more complex undertakings, such as intelligent image recognition and automated person identification. Kusakunniran and Chaiviroonjaroen (2019), Cai & Li (2013), and Xiao et al. (2022) have all reported that in recent times, machine learning and deep learning algorithms have been utilized as alternatives to the conventional methods that have been utilized in the identification of cattle. The convolutional neural network (CNN) is a method of deep learning that is extremely well-liked, as stated by Kaixuan & Dongjian (2015). CNNs have become increasingly common in recent years (Tsai et al. 2018). This surge in popularity can be attributed to the increased capacities of graphics cards. And this research emphasizes the importance of adopting face detection technologies in cattle identification, which can significantly enhance the efficiency of monitoring systems in large-scale farms, thereby reducing labor costs and minimizing human error. To summarise, the following is a list of the primary accomplishments that the current study has achieved:

1. We developed a monitoring system that enables the objective identification of particular animals in order to construct a database specifically for the purpose of recognizing the faces of cattle.

2. The findings of this research demonstrated that transfer learning results in an improvement in the ability to extract features from photographs of cattle faces. In order to improve the accuracy of transfer learning-based cattle face recognition, hyperparameter optimization was performed. Additionally, data augmentation techniques such as random flip, random rotation, random zoom, and others were utilized in order to prevent overfitting. Following the investigation, it was found that the DenseNet201 model had the highest level of performance.

3. The findings indicate that sophisticated computer vision models have been developed and are able to be utilized in the livestock industry. The system that has been proposed, which provides a variety of deployment options and the possibility of future feature enhancements, is designed to eliminate the need for a large number of wearable sensors and physical tags in the future.

2. Material and Methods

2.1. Dataset collection and preparation

The Alaca Livestock and Agriculture Enterprise, which can be found in the village of Seyran in Karacabey, Bursa (Turkey), served as the primary location for collecting data for this study. Numerous breeds, including Holstein Friesian, Montofon, and Simmental, are among those that can be discovered on the farm. Taking pictures of cattle faces both inside and outside was accomplished with the help of an RGB camera that had a resolution of 1920 x 1080 pixels (HD 1080). Creating a face image database for the purpose of cattle recognition was the reason for this action being taken. This study also emphasizes the importance of collecting diverse image data under various environmental conditions to enhance the robustness of the recognition model, thus ensuring accurate identification across different settings. First, the videos of the cattle faces were captured for the purpose of constructing the dataset. After that, the videos were converted to JPEG format by utilizing a free converter from the internet. It was discovered that the dataset initially contained a considerable number of images that were extremely similar to one another; consequently, these images were eliminated manually. In the beginning, there were 43 627 pictures of 30 cattle in the dataset; however, after the selection process was carried out manually, the number of pictures was reduced to 10,326. Following the implementation of image enhancement techniques, the final dataset was comprised of 15 000 photographs of the faces of cattle, captured from thirty different breeds of cattle (Figure 1). The dataset was divided into three sections, namely training, validation, and testing, with a ratio of 64:16:20. By employing a meticulous selection process and subsequent enhancement techniques, the study ensures that the dataset is of high quality, significantly reducing the risk of overfitting and improving model generalization. This arrangement was carried out in the following manner. Since this was the case, 9 600 images were utilized for training purposes, 2 400 images were utilized for validation, and 3 000 images were utilized for testing purposes.

Figure 1- Sample images from the training data set for cattle recognition

2.2. Image augmentation

Image augmentation employs various techniques such as image mixing, generative adversarial networks, random flip, random rotation, random zoom, random height, and random width. Incorporating advanced augmentation techniques such as GANs not only increases the diversity of the training set but also enhances the model's ability to generalize across unseen data, making it more resilient in real-world applications. Although image augmentation is commonly performed with supervision, it has diverse applications (Xu et al. 2022). Rice et al. (2020) and Schmidt et al. (2018) have identified a range of strategies for image augmentation, which include simple techniques like horizontal flipping and random cropping, as well as more advanced methods that leverage unlabeled data for semi-supervised learning. An initial task is created by employing image augmentation techniques, such as predicting the rotation angles and relative positions of image patches (Komodakis & Gidaris 2018; Doersch et al. 2015). Furthermore, augmented images that share similarity with the original can serve as positive examples for contrastive learning (Grill et al. 2020; Caron et al. 2021). Currently, the most widely used approach to improve data involves applying affine image transformations and color corrections, such as rotation, reflection, scaling (zooming in/out), and shearing. There are currently two categories of image augmentation techniques: deep neural network-based black-box methods and conventional white-box methods (Mikołajczyk & Grochowski 2018). The dataset in this study underwent random transformations such as flipping, rotation, zooming, and adjustments to its height and width (Figure 2). The study proposes the integration of semi-supervised learning methods alongside traditional augmentation to further refine the training process, enabling the model to leverage both labeled and unlabeled data efficiently.

Figure 2- Image augmentation process

The dataset also contained images of the left, full frontal, and right faces of cattle, each captured from a distinct viewpoint. Chen et al. (2022) stated that including the dataset scenarios mentioned above ensured a wide range of images, which made the dataset more complex and challenging. As a result, the proposed model became more resilient and capable of generalizing. The shuffle tool is used to input the initial data, including both the original and enhanced data, into the system in a completely random form.

2.3. Deep learning algorithms

A larger neural network with numerous layers, nodes, and activation functions is the basis of the DL based method. DL techniques have recently gained a widespread popularity and are frequently used in image classification tasks. In this study, InceptionResNetV2, MobileNetV2, DenseNet201, Xception, and NasNetLarge DL algorithms were used for the identification of individual cattle faces. The study includes a comparative analysis of multiple state-of-the-art deep learning models, providing insights into their respective strengths and weaknesses in the context of cattle face recognition. Pooling (GlobalAverage-Pooling2D), dropout (Dropout (0.2)) and dense layer were added classifier part of pre-trained TensorFlow models. In many different deep learning frameworks, InceptionResNetV2 has been extensively employed. With 164 layers, it is now widely accepted that larger networks allow for more accurate image comprehension (Bhatia et al. 2019). Using multiple convolution kernels of different sizes can improve the network's adaptability and extract more abundant features on a variety of scales, which is something that the Inception network structure takes into consideration. Furthermore, by applying the model, the Inception network structure can significantly cut down on the model's parameters, which allows it to represent model features accurately with fewer convolution kernels. As a result, the model becomes less complex (Wang et al. 2021).

2.3.1. MobileNetV2

MobileNetV2 is one of the most popular and portable CNNs. An inverted residual and a linear bottleneck are used in this network. It is designed to be used with images and has the ability to generate features and classify them. A total of 154 layers makes up MobileNetV2. Compared to other popular CNN models, MobileNetV2 uses 3.4 million parameters, which is fewer. MobileNetV2's lightweight architecture makes it particularly suitable for deployment in resource-constrained environments, such as farms where computing power may be limited. This enhances the practical applicability of the model in real-world settings. To solve the classification problem, MobileNetV2, a deep neural network, is utilized. (Sandler et al. 2018; Shahi et al. 2022).

2.3.2. DenseNet201

Utilizing a condensed network, the DenseNet201 is able to deliver highly parametric models that are simple to train and that allow for the reuse of features across multiple layers. Because of this feature reuse, performance is enhanced, and the input variety in the subsequent layer is increased. DenseNet201's ability to effectively reuse features across layers not only enhances model accuracy but also significantly reduces the computational cost, making it ideal for large-scale deployments. Within the DenseNet architecture, a straightforward connectivity pattern is utilized in order to establish direct and feed-forward connections between all of the layers. As a consequence of this, each layer transmits its own feature maps to all of the other layers, and these layers also receive additional inputs from all of the layers that came before them (Huang et al. 2017).

2.3.3. Xception

This is an improved version of Inception-v3, which is known as Xception. Instead of the traditional convolution operation, Inception-v3 makes use of a technique known as depth-wise separable convolution. Traditional convolution is divided into two stages: spatial convolution, which handle each input channel independently, and pointwise convolution, which convolves each point using a 1 × 1 kernel. Depth-wise separable convolution is a method that divides traditional convolution into two stages. In order to construct its architecture, the Xception model is comprised of fourteen blocks. The depth-wise separable convolution employed by Xception significantly reduces the model's complexity while maintaining high accuracy, thus providing a balanced approach between computational efficiency and performance. The total number of depth-wise separable convolution layers among these 14 blocks is 33. Additionally, there are three common convolution layers that are distributed throughout these blocks. With the exception of the first and last blocks, each and every block is surrounded by linear residual links (Szegedy et al. 2016).

2.3.4. NasNetlarge

NASNet model employ reinforcement learning-based search strategies. It creates search space by factoring the network into cells and then further breaking it up into multiple blocks. CNN models with varying kernel sizes support a variety of commonly used operations for each block, including convolutions, max pooling, average pooling, dilated convolutions, and depth-wise separable convolutions. By utilizing a reinforcement learning-based approach, NasNetLarge can automatically optimize its architecture for cattle face recognition, potentially discovering novel structures that outperform manually designed models (Zoph et al. 2018; Punn & Agarwal 2021).

2.4. Transfer learning

Increased data sizes have resulted in improved performance of deep learning models. When confronted with a scarcity of data, conventional approaches are frequently superior to deep learning methods. The integration of transfer learning in this study allows for the leveraging of pre-trained models on large-scale datasets, thereby enhancing the model's performance in cattle face recognition even when limited labelled data is available. Traditional learning theory posits that the generalization behaviour of a learning system is contingent upon the nth training case. From this standpoint, deep learning networks exhibit the anticipated behaviour: an increase in training data results in a decrease in test errors (Poggio et al. 2018).

In order to overcome the problem of limited data, one can utilize synthetic data generation or "learning transfer" techniques to augment the dataset by transferring features. Transfer learning is a technique used to address the issue of overfitting by leveraging knowledge acquired from solving one problem and applying it to a similar problem. Adopting this approach is essential for reducing overfitting. The study further refines transfer learning by incorporating hyperparameter optimization, which tailors the model to the specific nuances of cattle face images, thereby enhancing overall accuracy. Transfer learning involves training a model on a large dataset initially, and then using the acquired weights as the starting point for training new models (Khosla & Saini, 2020).

Deep learning models have the ability to achieve zero training error, meaning they can effectively memorize the training set without compromising their ability to generalize (Srivastava et al. 2014). Experimental evidence has shown that augmenting training data by increasing its volume and diversity can effectively mitigate overfitting in modern deep learning tasks that deal with high-dimensional data. Data augmentation is a commonly employed technique that has been demonstrated through empirical evidence to alleviate overfitting. The combination of data augmentation and transfer learning in this study provides a robust framework for improving model generalization, ensuring that the model performs well across different cattle breeds and environmental conditions (DeVries & Taylor 2017; Zhang et al. 2017; Schmidt et al. 2018).

Transfer learning facilitates expedited and more effective resolution of novel problems that bear resemblance to previously addressed ones. Transfer learning differs from traditional machine learning methods by leveraging knowledge from related domains to enhance predictive modelling with diverse data patterns in the current domain. In recent times, computational intelligence has been utilized to improve the effectiveness of transfer learning methods and regulate the process of transferring knowledge in real-world systems. The study highlights the potential of computational intelligence in optimizing transfer learning processes, making the approach more adaptable to real-world scenarios in livestock management (Lu et al. 2015).

Transfer learning involves addressing two fundamental questions: "what specific knowledge should be transferred?" and "what is the most effective method for transferring this knowledge?" Diverse transfer learning algorithms facilitate the transfer of distinct types of knowledge from a source to a target domain, resulting in varied enhancements in the target domain. Identifying the best option to maximize performance improvement requires thorough research or significant expertise. According to educational psychology, it is widely acknowledged that humans develop the ability to decide what to transfer through meta-cognitive reflection on inductive transfer learning practices (Ying et al. 2018).

2.5. Model performance metrics

The most common metrics used to measure the performance of multi-label classifiers currently include the F-score for each class, accuracy, precision, and recall. Each indicator's equation is organized as follows:

As mentioned in the second sentence, the true negative accurately predicts the negative class. The real deal: The positive class utilized in the model is accurately predicted by the true positive. A false positive happens when the model generates an inaccurate prediction regarding the positive class. Negative class prediction (FN) that was computed incorrectly. The incorrectly identified positive class is denoted by the acronym FP, the incorrectly identified negative class by the acronym FN, and the correctly identified positive class by the acronym TP. Precision, recall, and F1 score are the terms that are used to describe the actual values that the system predicts, and they are the components that make up the overall passing assessment.

The term "precision" refers to the proportion of all relevant results that are true positives (TP) or correct predictions. This proportion takes into account both true positives and false positives (FP). In tasks that involve the classification of multiple classes, the average of the classes is denoted by the letter P. In order to achieve precision, the formula is as follows.

$$Precision = \frac{TP}{TP+FP} \tag{1}$$

Recall: False negatives and the percentage of TP from the total amount of TP (FN). Recall is averaged across all classes in problems involving multi-class classification. The recall formula is as follows.

$$Recall = \frac{TP}{TP+FN} \tag{2}$$

The F1 score, which fully reflects the overall index of the model, is the harmonic average of precision and recall:

$$F1 = 2 * \frac{Precision*Recall}{Precision+Recall} \tag{3}$$

The most commonly utilized activation function is SoftMax, which solely considers the accuracy of classification and disregards the inter-class distance. In this study, the researchers have selected the face detection model and the latest loss functions for face identification, as described by Xu et al. (2022), to be used specifically for identifying cattle faces. The selection of SoftMax as the activation function is due to its efficiency in handling multi-class classification problems, which is essential for differentiating among numerous cattle individuals. The hyperparameters that have been chosen are listed in Table 1. These values yielded the most advantageous training results following thorough experimentation. Categorical cross-entropy is a commonly employed loss function in tasks like classification, where the output variable is a categorical variable with multiple classes (Gerdan Koc et al. 2023). The learning rate, which began at 0.01 and decreased correspondingly every 5 epochs to 0.00001, followed a precise applying schedule. The dynamic adjustment of the learning rate helps in fine-tuning the model, ensuring it converges more effectively without overshooting, leading to higher accuracy in cattle face recognition tasks. By lowering the learning rate according to a predefined schedule, learning rate schedules aim to modify the learning rate during training. The models were compared based on their performance on the testing set. The number of epochs without improvement is the patience value, after which training will be stopped. According to the training set and GPU performance, setting the batch size to 32 resulted in better model performance.

Table 1- Hyperparameters of algorithms

Parameters	***Values***
Batch size	32
Epochs	25
Momentum	0.9
Learning rate	0.00001
Metric	Categorical cross entropy
Patience	3
Optimization method	Adam
Activation function	SoftMax

3. Results

A Google Colab notebook, coded in Python, served as the platform for the whole investigation. The system specs for the computer included Windows 10, an Intel® CoreTM i7-10750H CPU, 16 GB of RAM, and a graphics processing unit from NVIDIA called an RTX 2060. Individual cattle face recognition with deep convolutional neural networks trained on augmented datasets was the primary goal of the research. The use of Google Colab with GPU acceleration significantly reduced the training time, enabling rapid experimentation and iteration on various model architectures. Through the utilization of open-source libraries, experiments were carried out with a variety of CNN models, such as InceptionResNetV2, MobileNetV2, DenseNet201, Xception, and NasNetLarge. Each of the three sections of the cattle face dataset consisted of nine thousand six hundred images: training (9,600 images), testing (2,400 images), and validation (three thousand images). The results of training each CNN model with fixed learning rates, epochs, and batch sizes were compared after the training was completed. The accuracy of the training and validation have been summarized in Table 3. By utilizing the Adam optimizer, DenseNet201 was able to achieve the highest validation accuracy possible, which was 99.83%. A validation accuracy of 98.87% was achieved by NasNetLarge, while MobileNetV2 achieved 98.54%, Xception achieved 98.88%, and InceptionResNetV2 achieved 92.54%. The DenseNet201 model's superior performance in validation accuracy highlights its robustness in generalizing from the training data, making it a prime candidate for real-world deployment in cattle face recognition systems (Table 2).

Table 2- Variation of validation accuracy and loss depending on the DL algorithms selected

Algorithms	*Validation Accuracy (%)*	*Validation Loss*
DenseNet201	99.83	0.0836
NasNetLarge	98.87	0.0971
MobilNetV2	98.54	0.1354
Xception	98.88	0.1678
InceptionResNetV2	92.54	0.3665

A value of 0.0836 was obtained for the DenseNet201 architecture, which was the CNN model that experienced the least amount of loss in validation. The low validation loss for DenseNet201 further supports its efficiency and accuracy, indicating minimal overfitting and strong performance across various data splits. Furthermore, according to the outcomes of the training, DenseNet201 was found to be the most efficient architecture with the least amount of loss. InceptionResNetV2 was found to have the highest loss, with a value of 0.3665, according to the findings. The epochs of accuracy for the models that were utilized in the experiments are displayed in Figure 3.

The individual algorithm-based results are presented in Table 3, which includes the average accuracy, macro average, weighted average, precision, recall, and F1 score values that were obtained by each of the models on the test dataset. The total amount of time spent training was determined by starting from the epoch in which the loss values of the models started to increase. This was done because early stopping was utilized during the training process. Implementing early stopping during training ensured that the models did not overfit, preserving their ability to generalize effectively to unseen data. Table 3 presents a comparison of the identification accuracy of the following algorithms: DenseNet201, MobileNetV2, Xception, NasNetLarge, and InceptionResNetV2. The results of the experiments presented in Table 3 make it abundantly clear that the DenseNet201 algorithm performed significantly better than other methods when applied to the cattle face dataset.

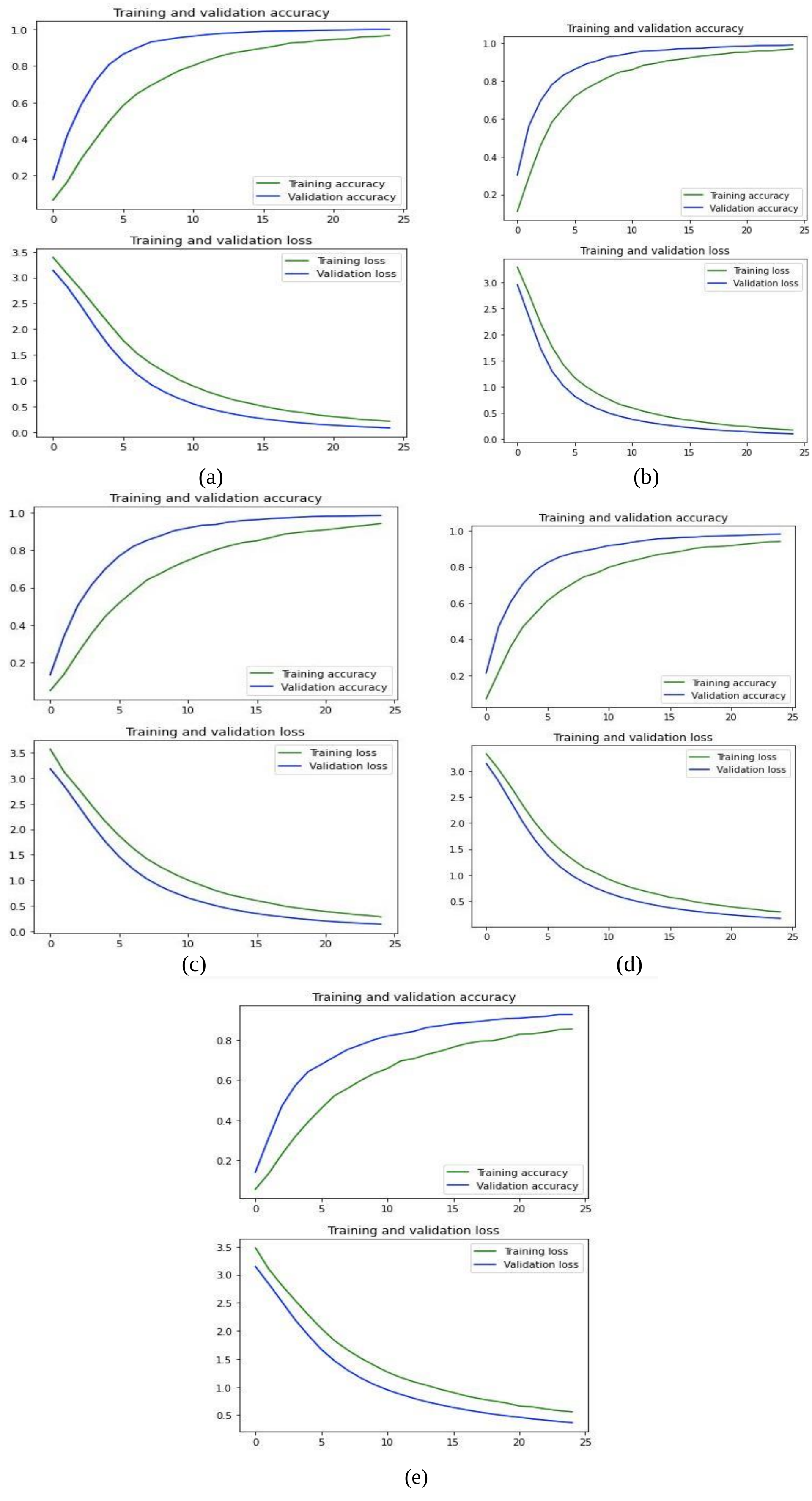

Figure 3- Training and validation accuracy and loss plots of DenseNet201 (a), NasNetLarge (b), MobilNetV2 (c), Xception (d), InceptionResNetV2 (e)

Table 3- Performance metrics and accuracy criteria of the algorithms for test results

Algorithms	*Accuracy (%)*	*Macro avg. Precision (%)*	*Weighted avg. Precision (%)*	*Macro avg. Recall (%)*	*Weighted avg. Recall (%)*	*Macro avg. F1 score (%)*	*Weighted avg. F1 score (%)*	*Total test time (min)*
DenseNet201	99.53	99.52	99.54	99.51	99.53	99.50	99.53	142.5
NasNetLarge	98.96	98.95	98.98	98.96	98.97	98.95	98.97	196.25
MobilNetV2	98.80	98.76	98.81	98.74	98. 80	98.74	98.79	44.6
Xception	98.03	98.01	98.05	97.99	98.03	97.98	98.02	85.42
Inception ResnetV2	93.60	93.42	93.62	93.46	93.60	93.36	93.53	92.1

DenseNet201 not only achieved the highest accuracy but also demonstrated consistent performance across all evaluation metrics, making it the most reliable model for cattle face recognition. According to Table 3, the architectures of the DenseNet201 and NasNetLarge models achieved the highest accuracy (99.53% and 98.96%) and precision (99.52% and 98.95%), respectively, when compared to the other deep learning models that were included in the dataset. With a weighted average precision of 99.53%, the DenseNet201 algorithm achieved the highest level of precision, while the InceptionResNetV2 algorithm achieved the lowest level of precision, which was 93.60%. Values of recall ranged from 99.51% for DenseNet201 to 93.46% for InceptionResNetV2, between the two networks. Similarly, the weighted average recall for DenseNet201 was 99.53%, while the recall for InceptionResNetV2 was only 93.60%. The superior recall of DenseNet201 indicates its ability to accurately identify cattle across a variety of conditions and environments, minimizing the likelihood of missed identifications.

After 196.25 minutes of testing, the NasNetLarge network was found to have the longest test time, while MobileNetV2 was found to have the shortest test time, which was 44.6 minutes. Following the completion of the analysis and reaching a conclusion regarding the training and validation data, the test data were analysed. The results of the test are displayed in Figure 4, which shows five outcomes that were chosen at random.

Figure 4- Results of five random test samples

The use of Convolutional Neural Networks (CNNs) is employed in order to guarantee that the network's "attention" is concentrated on the actual, distinguishing characteristics of the animal, as opposed to other regions of the image that may also contain information that is pertinent (Selvaraju et al. 2016; Hansen et al. 2018). The application of CNNs ensures that the model focuses on key distinguishing features of cattle faces, which is crucial for accurate identification, particularly in complex environments. The Grad-CAM system offers graphical explanations for the decisions made by CNN. Grad-CAM, in contrast to other methods, typically backpropagates the gradient to the final convolutional layer rather than the entire image. This results in the production of a coarse localization map that highlights significant regions of the image. Using Grad-CAM enhances model interpretability, allowing researchers and practitioners to understand which features are being used for decision-making, thus ensuring transparency in the cattle identification process. Figure 5 depicts a method that can be utilized to generate a coarse localization map for a particular class that the network has been trained on. This method is referred to as Gradient-weighted Class Activation Mapping (Grad-CAM). This image illustrates how the algorithm places significant emphasis on the face of the cattle.

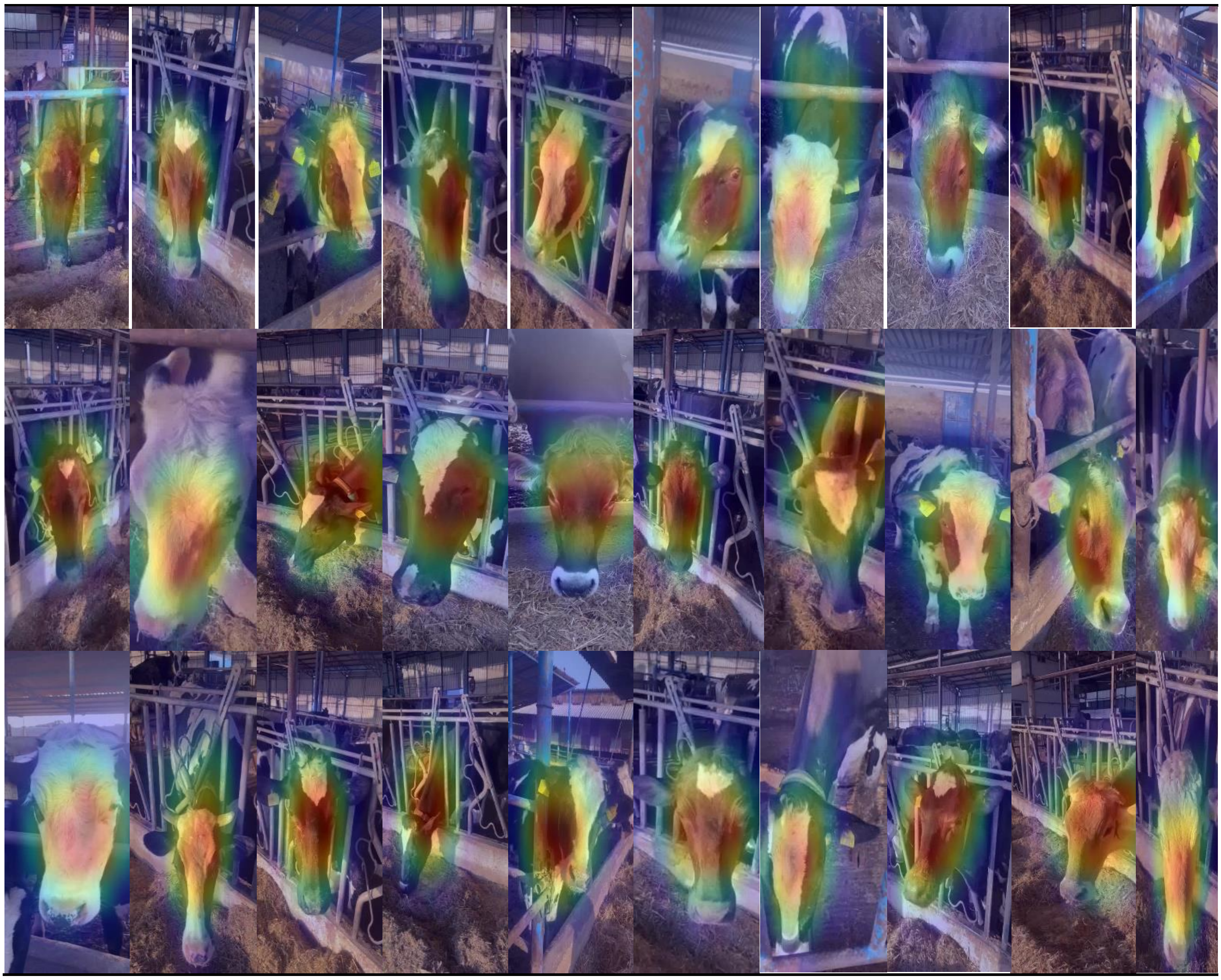

Figure 5- Gradient-weighted Class Activation Maps of cattle's

4. Discussion

A CNN-based model was developed by Qiao et al. (2021) as an alternative to RF-based ear tags for cattle identification. In their study, the researchers used photos from the top and back of the cattle to perform a deep learning analysis. The CNN model Inception-V3 was employed, and they achieved success rates of up to 88% and 99% for categorizing the rear and top photos of 41 calves, respectively. In the current study, five different deep learning methods were used along with 15 000 photos of 30 different cattle faces: InceptionResNetV2, MobileNetV2, DenseNet201, Xception, and NasNetLarge. The DenseNet201 algorithm demonstrated the highest accuracy with a score of 99.53%. Furthermore, the DenseNet201 and other algorithms for identifying livestock by their faces were more accurate at classifying them compared to images taken from the top and rear of the cattle. This finding suggests that facial features provide more reliable and consistent data points for identification than other body parts, which can vary significantly in appearance due to factors like posture, movement, and environmental conditions. Shen et al. (2020) conducted a deep learning study to classify dairy cattle using the YOLO model and the AlexNet model. In the research, 105 cattle were attempted to be classified by side views. The study findings showed that the proposed model had an accuracy of 96.65%. These results outperformed InceptionResNetV2, one of the deep learning algorithms used in our study on cattle facial recognition, in terms of accuracy (93.60%). The accuracy of the algorithms used in the current study varied from 93.60% to 99.53%. The comparative analysis with existing studies highlights the effectiveness of DenseNet201 in achieving higher accuracy, underscoring its potential as a superior model for cattle identification in practical applications.

Li et al. (2022) implemented deep learning algorithms to identify Simmental cows from images of 103 individuals. The models recommended by the researchers, including AlexNet, VGG16, MobileNetV1, SqueezeNet, and CNN, were used in the study. Further-more, AlexNet and the model recommended by the researchers both had the greatest ac-curacy of 98.86% and 98.37%, respectively. The accuracy of the CNN models used in our study on the classification of cattle was 99.53% for DenseNet201. The results showed that the accuracy obtained in our study was higher than the accuracy (98.7%) obtained by Li et al. (2022). The adopted methodologies and the variations in the face patterns of the cattle species used as test subjects might have caused this variance. The methodology and dataset used in this study allowed for more refined and accurate cattle face recognition, which could be attributed to the enhanced image quality and diversity incorporated through data augmentation techniques.

Detecting the muzzle point patterns of cattle and classifying them using deep learning algorithms is another alternative for the classification of dairy cattle (Kumar et al. 2018). In the study, the researchers used 5 000 images collected from 500 samples that were selected using a filtering process. The accuracy levels found in the study using Deep Belief Network (DBN), Stacked Denoising Auto-encoder (SDAE), and CNN deep learning algorithms were 95.99%, 88.46%, and 75.98%, respectively. It was concluded that the DBN-based method can be used for such investigations due to its high accuracy. The classification accuracy achieved in the current study was substantially higher than that of Kumar et al. (2018), which compared the spots on the nose that were utilized to classify cattle. By focusing on facial features rather than muzzle point patterns, this study was able to achieve higher accuracy, suggesting that facial recognition may be a more reliable method for cattle identification. This can be attributed to the fact that the study was built around the cattle face im-ages that we used in our study. The points on the noses of cattle were used as a reference in the study by Kumar et al. (2018). Because they were easier to be distinguished from the photos with a point structure on the nose, images with a larger surface area and cattle faces were chosen.

Andrew et al. (2017) used drones to take upper body pictures of cattle from the air and create a data set in order to identify animals using deep learning techniques. In the study, they attempted to distinguish cattle faces based on the variations in the shape and pattern of their backs. They built their own data sets for the study using the video they captured with the drone. They discovered that using photos of 89 animals after milking resulted in an accuracy of 86.1%, whereas using images of 23 animals while they were grazing resulted in an accuracy of 98.1%. The researchers used faster R-CNN, which is a part of the Caffe deep learning package as a deep learning technique. The accuracy level detected in their study was lower than that of our study. This comparison emphasizes the importance of using high-resolution, close-up images for facial recognition, as it leads to significantly higher accuracy than aerial or distant photography. It might be argued that these variations are caused by continually moving livestock and the shooting angles used for aerial photography. The upper body shots of the cattle were taken from a distance of 5 m. Images that were taken from a distance of around 1 m were used in the samples that made up our data collection. In our study, it can be said that the separation of cattle face images is more decisive than the separation from the back photographs, which is another explanation for these disparities (Andrew et al. 2017).

Similarly, cattle face recognition was studied by Jaddou et al. (2020), who used Sup-port Vector Machine to identify the faces of 702 images of cattle and achieved a 99.00% success rate. The use of CNNs, as opposed to SVMs, allows for the handling of more complex image data and offers scalability to larger datasets, which likely contributed to the higher success rates observed in this study. In their study, Kumar et al. (2018) used 5,000 images of crossbred, Holstein Friesian, hybrid Ongole, and Balinese cattle. For cattle face determination, various methods including batch-CCIPCA, ICA, IND-CCIPCA, ISVM, LDA, LDA-LiBSVM, PCA, and PCA-LiBSVM were used. According to the findings, 95.87% of their study was successful. Lin et al. (2019) achieved 96.76% accuracy using the Fast R-CNN method with 900 images of cattle faces. Guo et al. (2022) used 3,152 Holstein Friesian cattle data points to analyze the YOLO V3-tiny. The accuracy success rate was 90.0%. This study's higher accuracy underscores the effectiveness of DenseNet201 and other advanced CNN models in providing superior performance in cattle face recognition tasks, especially when compared to older or simpler methods. The identification of Simmental and Holstein Friesian cattle faces was studied by Weng et al. in 2022. Deep image processing algorithms like VGG 16, AlexNet, GoogleNet, ResNet34, and two-branch convolutional neural networks (TB-CNN) were used for this purpose. Simmental cattle had an accuracy of 99.85%, while Holstein Friesian cattle had an accuracy of 99.81%. For detecting cattle faces, Yao et al. (2019) used the VGGNet, Inception V2, ResNet50, and Res-Net101 models. Their study had a 98.3% accuracy rate.

5. Conclusions

The fields of object detection and classification have seen significant advancements thanks to the application of deep learning principles. Increasing the gap between classes while simultaneously narrowing the gap between classes is the objective of classification. This field of machine learning for image processing has been dominated by deep learning models. The advancements that have been made in deep learning and image processing have provided an opportunity to broaden the scope of research and applications of plant disease detection and classification through the use of images. It is necessary to have models that are both quick and accurate in order to put into effect effective measures as quickly as possible. The exploration of deep learning models in animal identification, as demonstrated in this study, further expands the horizon for smart farming applications, ensuring more precise and efficient livestock management.

It is dependent on the development of deep learning for object detection and picture processing in order to replace wearable devices such as ear tags with the livestock identification system. This system also reduces the amount of harm of animals. The shift from traditional identification methods, such as ear tags, to facial recognition using deep learning, significantly reduces animal stress and improves welfare, making this approach more humane and sustainable. This study focused on the detection of faces on cattle, which is an essential component of the technology that is expected to emerge in the near future. The goal of this research was to develop a livestock machine vision system that is capable of monitoring individuals. The pretrained DenseNet algorithm, which had a 99.83% accuracy rate, was evaluated using a variety of unstructured scenes. All of these scenes were used to evaluate the algorithm. The use of DenseNet201, with its exceptional accuracy, suggests that this model could serve as a cornerstone in future livestock monitoring systems, enabling precise individual identification even in diverse and challenging environments.

The purpose of this research was to identify the faces of cattle for the purpose of developing intelligent farming systems. A significant increase in the success rates of deep learning can be attributed to the utilization of transfer learning and data augmentation techniques. By integrating transfer learning and data augmentation, this study enhances the robustness and generalizability of the model, making it adaptable to various farming conditions and cattle breeds. The findings indicate that DenseNet performs at the highest level possible, as evidenced by its F1 score of 99.52%, its precision score of 99.50%, and its average total processing time of 142.5 minutes. Transfer learning is an essential component of deep learning because prior CNN models can be improved and retrained to perform new tasks even when there is a lack of labelled data for training. This makes transfer learning an essential part of deep learning. The implementation of transfer learning not only accelerates the training process but also mitigates the issue of overfitting, which is critical in scenarios with limited data availability. One possible factor that may influence the degree to which different deep neural networks generalize across a variety of datasets is the architecture of those networks.

Following the findings of this study, it is possible that the incorporation of robotic and early diagnosis systems on cattle farms will become feasible in the future. It is possible that the development of an infrastructure for artificial intelligence and deep learning will make it possible to quickly address a number of vital parameters pertaining to cattle, such as diseases, live weight, feed consumption, and estrus. The integration of AI and deep learning infrastructure in livestock management promises not only improved health monitoring but also optimized resource usage, contributing to more sustainable farming practices. It is therefore possible that this could make it possible to raise healthy cattle in a more efficient manner.

References

Allen A, Golden B, Taylor M, Patterson D, Henriksen D & Skuce R (2008). Evaluation of retinal imaging technology for the biometric identification of bovine animals in Northern Ireland. Livestock science 116(1-3): 42-52. DOI: https://doi.org/10.1016/j.livsci.2007.08.018

Andrew W, Greatwood C & Burghardt T (2017). Visual localisation and individual identification of holstein friesian cattle via deep learning. In Proceedings of the IEEE International Conference on Computer Vision (ICCV), 2017, pp. 2850-2859. Available from https://openaccess.thecvf.com/content_ICCV_2017_workshops/w41/html/Andrew_Visual_Localisation_and_ICCV_2017_paper.html

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

Bhatia Y, Bajpayee A, Raghuvanshi D & Mittal H (2019). Image captioning using Google's inception-resnet-v2 and recurrent neural network. In 2019 Twelfth International Conference on Contemporary Computing (IC3), IEEE Publish-ing, pp. 1-6. DOI: https://doi.org/10.1109/IC3.2019.8844921

Cai C & Li J (2013). Cattle face recognition using local binary pattern descriptor. In 2013 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference, IEEE Publishing pp. 1-4. DOI: https://doi.org/10.1109/APSIPA.2013.6694369

Caron M, Touvron H, Misra I, Jégou H, Mairal J, Bojanowski P & Joulin A (2021). Emerging properties in self-supervised vision transformers. In Proceedings of the IEEE/CVF International Conference on Computer Vision, pp. 9650-9660. DOI: https://doi.org/10.48550/arXiv.2104.14294

Chen X, Yang T, Mai K, Liu C, Xiong J, Kuan Y & Gao Y (2022). Holstein Cattle Face Re-Identification Unifying Glob-al and Part Feature Deep Network with Attention Mechanism. Animals, 12(8), DOI: https://doi.org/10.3390/ani12081047

DeVries T & Taylor G W (2017). Improved regularization of convolutional neural networks with cutout. arXiv preprint arXiv:1708.04552. DOI: https://doi.org/10.48550/arXiv.1708.04552

Doersch C, Gupta A & Efros A A (2015). Unsupervised visual representation learning by context prediction. In Proceedings of the IEEE International Conference on Computer Vision (ICCV), 2015, pp. 1422-1430. Available from https://www.cv-foundation.org/openaccess/content_iccv_2015/html/Doersch_Unsupervised_Visual_Representation_ICCV_2015_paper.html

Džermeikaitė K, Bačėninaitė D & Antanaitis R (2023). Innovations in Cattle Farming: Application of Innovative Technologies and Sensors in the Diagnosis of Diseases. Animals, 13(5): 780

Fosgate G T, Adesiyun A A & Hird D W (2006). Ear-tag retention and identification methods for extensively managed water buffalo (Bubalus bubalis) in Trinidad. Preventive veterinary medicine, 73(4): 287-296. DOI: https://doi.org/10.1016/j.prevetmed.2005.09.006

Gerdan Koc D, Koc C &Vatandas M (2023). Diagnosis of tomato plant diseases using pre-trained architectures and a proposed convolutional neural network model. *Journal of Agricultural Sciences (Tarim Bilimleri Dergisi)* 29(2): 627-638. doi.org/10.15832/ankutbd.957265

Grill J B, Strub F, Altché F, Tallec C, Richemond P, Buchatskaya E & Valko M (2020). Bootstrap your own latent-a new approach to self-supervised learning. Advances in neural information processing systems, 33, pp: 21271-21284. ISBN: 9781713829546. Available from https://proceedings.neurips.cc/paper/2020/hash/f3ada80d5c4ee70142b17b8192b2958e-Abstract.html

Guo S S, Lee K H, Chang L, Tseng C D, Sie S J, Lin G Z & Lee T F (2022). Development of an Automated Body Temperature Detection Platform for Face Recognition in Cattle with YOLO V3-Tiny Deep Learning and Infrared Thermal Imaging. Applied Sciences, 12(8), DOI: https://doi.org/10.3390/app12084036

Hansen M F, Smith M L, Smith L N, Salter M G, Baxter E M, Farish M & Grieve B (2018). Towards on-farm pig face recognition using convolutional neural networks. Computers in Industry 98: 145-152

Huang G, Liu Z, van der Maaten L &Weinberger K Q (2017). Densely connected convolutional networks. In Proceedings of the IEEE conference on computer vision and pattern recognition. arXiv. pp: 4700-4708. DOI: https://doi.org/10.48550/arXiv.1608.06993

Jiang B, Wu Q, Yin X, Wu D, Song H & He D (2019). FLYOLOv3 deep learning for key parts of dairy cow body detection. Computers and Electronics in Agriculture 166 (2019), DOI: https://doi.org/10.1016/j.compag.2019.104982

Kaixuan Z & Dongjian H (2015). Recognition of individual dairy cattle based on convolutional neural networks. Transactions of the Chinese Society of Agricultural Engineering, 31(5): 181-187. Available from https://www.cabdirect.org/cabdirect/abstract/20153218172

Kang X, Zhang X D & Liu G (2020). Accurate detection of lameness in dairy cattle with computer vision: A new and individualized detection strategy based on the analysis of the supporting phase. Journal of dairy science, 103 (11): 10628-10638. DOI: https://doi.org/10.3168/jds.2020-18288

Khosla C & Saini B S (2020). Enhancing performance of deep learning models with different data augmentation techniques: A survey. In 2020 International Conference on Intelligent Engineering and Management (ICIEM), IEEE, pp. 79-85, DOI: https://doi.org/10.1109/ICIEM48762.2020.9160048

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

Kumar S, Singh S K, Dutta T & Gupta H P (2016). A fast cattle recognition system using smart devices. MM '16: Proceedings of the 24th ACM international conference on Multimedia, October 2016, pp: 742–743, DOI: https://doi.org/10.1145/2964284.2973829

Kumar S, Singh S K, Singh R & Singh A K (2017). Recognition of Cattle Using Face Images. In: Animal Biometrics. Springer, Singapore. DOI: https://doi.org/10.1007/978-981-10-7956-6_3

Kusakunniran W & Chaiviroonjaroen T (2018). Automatic cattle identification based on multi-channel lbp on muzzle images. In 2018 International Conference on Sustainable Information Engineering and Technology (SIET), IEEE Publishing, pp:1-5. DOI: https://doi.org/10.1109/SIET.2018.8693161

Li G, Erickson G E & Xiong Y (2022). Individual Beef Cattle Identification Using Muzzle Images and Deep Learning Techniques. Animals, 12(11), DOI: https://doi.org/10.3390/ani12111453

Li Z, Lei X & Liu S (2022). A lightweight deep learning model for cattle face recognition. Computers and Electronics in Agriculture, 195 (2022), DOI: https://doi.org/10.1016/j.compag.2022.106848

Lu J, Behbood V, Hao P, Zuo H, Xue S & Zhang G (2015). Transfer learning using computational intelligence: A survey. Knowledge-Based Systems, 80 (2015), pp: 14-23, https://doi.org/10.1016/j.knosys.2015.01.010

Mikołajczyk A & Grochowski M (2018). Data augmentation for improving deep learning in image classification problem. In 2018 international interdisciplinary PhD workshop (IIPhDW), IEEE publishing, pp. 117-122, Poland. DOI: https://doi.org/10.1109/IIPHDW.2018.8388338

Noonan G J, Rand J S, Priest J, Ainscattle J & Blackshaw J K (1994). Behavioural observations of piglets undergoing tail docking, teeth clipping and ear notching. Applied Animal Behaviour Science, 39(3-4), 203-213. DOI: https://doi.org/10.1016/0168-1591(94)90156-2

Poggio T, Kawaguchi K, Liao Q, Miranda B, Rosasco L, Boix X & Mhaskar H (2018). Theory of deep learning III: ex-plaining the non-overfitting puzzle. arXiv preprint arXiv:1801.00173. DOI: https://doi.org/10.48550/arXiv.1801.00173

Polat H E (2022) New Technologies in Good Agricultural Practices – Smart Farming (In Turkish). In: Yaldız G, Çamlıca M (Eds.), Innovative Approaches in Medicinal and Aromatic Plants Production. Iksad Publications, Ankara/Turkey pp: 27- 54. ISBN:978-625-8246-33-9

Psota E T, Luc E K, Pighetti G M, Schneider L G, Fryxell R T, Keele J W & Kuehn L A (2021). Development and validation of a neural network for the automated detection of horn flies on cattle. Computers and Electronics in Agriculture, 180 (2021), DOI: https://doi.org/10.1016/j.compag.2020.105927

Punn N S & Agarwal S (2021). Automated diagnosis of COVID-19 with limited posteroanterior chest X-ray images using fine-tuned deep neural networks. Applied Intelligence, 51(5), 2689-2702

Qiao Y, Clark C, Lomax S, Kong H, Su D & Sukkarieh S (2021). Automated Individual Cattle Identification Using Video Data: A Unified Deep Learning Architecture Approach. Front. Anim. Sci., 2 (2021), https://doi.org/10.3389/fanim.2021.759147.

Rice L, Wong E & Kolter Z (2020). Overfitting in adversarially robust deep learning. Proceedings of the 37 th International Conference on Machine Learning, Vienna, Austria, PMLR 119, 2020. pp. 8093-8104. http://proceedings.mlr.press/v119/rice20a

Ruiz-Garcia L & Lunadei L (2011). The role of RFID in agriculture: Applications, limitations and challenges. Computers and Electronics in Agriculture, 79(1), 42-50. DOI: https://doi.org/10.1016/j.compag.2011.08.010

Sandler M, Howard A, Zhu M, Zhmoginov A & Chen L C (2018). MobilenetV2: Inverted residuals and linear bottlenecks. In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 4510–4520. DOI: https://doi.org/10.48550/arXiv.1801.04381

Schmidt L, Santurkar S, Tsipras D, Talwar K & Madry A (2018). Adversarially robust generalization requires more data. In Advances in Neural Information Processing Systems, pp. 5014–5026. https://proceedings.neurips.cc/paper/2018/hash/f708f064faaf32a43e4d3c784e6af9ea-Abstract.html

Selvaraju R R, Das A, Vedantam R, Cogswell M, Parikh D & Batra D (2016). Grad-CAM: Why did you say that? arXiv preprint arXiv:1611.07450

Shahi T B, Sitaula C, Neupane A & Guo W (2022). Fruit classification using attention-based MobileNetV2 for industrial applications. Plos one, 17(2), DOI: https://doi.org/10.1371/journal.pone.0264586

Shen W, Hu H, Dai B Wei X, Sun J, Jiang L & Sun Y (2020). Individual identification of dairy cows based on convolutional neural networks. Multimed Tools Appl (79) 14711–14724, https://doi.org/10.1007/s11042-019-7344-7

Srivastava N, Hinton G, Krizhevsky A, Sutskever I & Salakhutdinov R (2014). Dropout: a simple way to prevent neural networks from overfitting. The journal of machine learning research, 15(1):1929–1958, DOI: https://dl.acm.org/doi/abs/10.5555/2627435.2670313

Szegedy C, Vanhoucke V, Ioffe S, Shlens J & Wojna Z (2016). Rethinking the inception architecture for computer vision. In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 2818-2826. https://www.cv-foundation.org/openaccess/content_cvpr_2016/html/Szegedy_Rethinking_the_Inception_CVPR_2016_paper.html

Tsai H, Ambrogio S, Narayanan P, Shelby R M & Burr G W (2018). Recent progress in analog memory-based accelerators for deep learning. Journal of Physics D: Applied Physics, 51(28), DOI: https://doi.org/10.1088/1361-6463/aac8a5

Wang H, Qin J, Hou Q & Gong S (2020). Cattle face recognition method based on parameter transfer and deep learning. In Journal of Physics: Conference Series, 1453 (2020), IOP Publishing, DOI: https://doi.org/10.1088/1742-6596/1453/1/012054

Wang J, He X, Faming S, Lu G, Cong H & Jiang Q (2021). A Real-Time Bridge Crack Detection Method Based on an Improved Inception-Resnet-v2 Structure. IEEE Access, (9) pp: 93209-93223. DOI: https://doi.org/10.1109/ACCESS.2021.3093210

Wang Z, Fu Z, Chen W & Hu J (2010). A RFID-based traceability system for cattle breeding in China. International Conference on Computer Application and System Modeling (ICCASM 2010), IEEE, Taiyuan. DOI: https://doi.org/10.1109/ICCASM.2010.5620675

Weng Z, Meng F, Liu S, Zhang Y, Zheng Z & Gong C (2022). Cattle face recognition based on a Two-Branch convolutional neural network. Computers and Electronics in Agriculture, 196 (2022), DOI: https://doi.org/10.1016/j.compag.2022.106871

Xiao J, Liu G, Wang K & Si Y (2022). Cow identification in free-stall barns based on an improved Mask R-CNN and an SVM. Computers and Electronics in Agriculture, 194 (2022), DOI: https://doi.org/10.1016/j.compag.2022.106738

Xu B, Wang W, Guo L, Chen G, Li Y, Cao Z & Wu S (2022). CattleFaceNet: A cattle face identification approach based on Retina Face and ArcFace loss. Computers and Electronics in Agriculture, 193 (2022), DOI: https://doi.org/10.1016/j.compag.2021.106675

Xu B, Wang W, Guo L, Chen G, Wang Y, Zhang W & Li Y (2021). Evaluation of deep learning for automatic multi-view face detection in cattle. Agriculture, 11(11), https://doi.org/10.3390/agriculture11111062

Xu M, Yoon S, Fuentes A & Park D S (2023). A comprehensive survey of image augmentation techniques for deep learning. Pattern Recognition, 137 (2023) DOI: https://doi.org/10.1016/j.patcog.2023.109347

Yao L, Liu H, Hu Z, Kuang Y, Liu C & Gao Y (2019). Cow face detection and recognition based on automatic feature extraction algorithm. ACM TURC '19: Proceedings of the ACM Turing Celebration Conference—China, May 2019, Article No.: 95, Pp 1–5, https://doi.org/10.1145/3321408.3322628

Ying W, Zhang Y, Huang J & Yang Q (2018). Transfer learning via learning to transfer. In International Conference on Machine Learning, pp. 5085-5094. Available from https://proceedings.mlr.press/v80/wei18a.html.

Zhang C, Bengio S, Hardt M, Recht B & Vinyals O (2016). Understanding deep learning requires rethinking generalization. arXiv preprint arXiv:1611.03530. DOI: https://doi.org/10.48550/arXiv.1611.03530

Zhang J, Sun G, Zheng K & Mazhar S (2020). Pupil detection based on oblique projection using a binocular camera. IEEE Access, vol: 8, pp: 105754-105765, DOI: https://doi.org/10.1109/ACCESS.2020.3000063

Zoph B, Vasudevan V, Shlens J & Le Q V (2018). Learning transferable architectures for scalable image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 8697-8710. DOI: https://doi.org/10.48550/arXiv.1707.07012